UNIVERSITY OF OKLAHOMA
GRADUATE COLLEGE

EVALUATION OF NETWORK MOBILITY SCHEMES FOR TERRESTRIAL AND SATELLITE NETWORKS

A DISSERTATION
SUBMITTED TO THE GRADUATE FACULTY
in partial fulfillment of the requirements for the
Degree of
DOCTOR OF PHILOSOPHY

By

ABU ZAFAR MOHAMMAD SHAHRIAR
Norman, Oklahoma
2012

EVALUATION OF NETWORK MOBILITY SCHEMES FOR TERRESTRIAL AND SATELLITE NETWORKS

A DISSERTATION APPROVED FOR THE
SCHOOL OF COMPUTER SCIENCE

BY

Dr. Mohammed Atiquzzaman, Chair

Dr. John K. Antonio

Dr. Changwook Kim

Dr. Dean Hougen

Dr. Hazem Refai

Acknowledgement

This work has been possible because of a number of individuals. First of all, I would like to express my gratitude to Dr. Mohammed Atiquzzaman for his expert advice and continuous supervision, which made this work possible. Mr. W. D. Ivancic of NASA's Glenn Research Center and Dr. L. Wood of University of Surrey also provided suggestions to improve the dissertation. The technical expertise of Mr. S. M. Huq, Mr. S. Rahman, Mr. R. Qureshi, Mrs. Y. Jahir and Mr. M. S. Hossain led to the improvements of this work. The author would like to acknowledge the financial support of National Aeronautics and Space Administration to carry out this project.

I am thankful to all faculty and staff members in School of Computer Science at the University of Oklahoma for providing a supportive environment for my study. The courses taught by Dr. S. Radhakrishnan, Dr. S. Dhall and Dr. S. Lakshmivarahan gave me the momentum to get back to study. Technical support from Mr. J. Summers and the support from Ms. B. Bledsoe were invaluable. Also, I must acknowledge the logistical support from Ms. C. Yerdon, Ms. E. Pierce and Ms. M. Moore to carrying out this dissertation to completion.

I am truly grateful to my committee members: Dr. John K. Antonio, Dr. Changwook Kim, Dr. Dean Hougen and Dr. Hazem H. Refai. They took their valuable time to review my requirements and competency for the degree and the dissertation. The quality of this dissertation has been improved by virtue of their comments.

Last but not least, I would like to thank my wife Nayeema and my son Muntasir for their understanding and support throughout this work, and my mother who provided moral support.

Contents

List of tables

List of figures

Abstract

NEtwork MObility (NEMO) supports the mobility of multiple Internet-connected devices. However, NEMO Basic Support Protocol suffers from unoptimized route leading to large latency in communication and header overhead. To optimize route, a plethora of schemes have been proposed. These schemes differ in terms of several performance parameters, such as signaling, end-to-end delay and handoff latency. However, no performance or cost evaluation exists in the literature to compare the schemes. In addition, mobility management is required to support the mobility of Internet-connected devices in satellite networks. Existing mobility management solutions for satellite networks are unable to provide connectivity to the Internet when satellites are not directly connected to the ground.

In this dissertation, a comprehensive evaluation of the schemes and a mobility management solution for satellite networks using NEMO are provided. The schemes are classified and compared to choose the optimal class. Using analytical and simulation-based models, the schemes in the chosen class are compared based on the performance parameters. The effect of the parameters on Transmission Control Protocol, the dominant transport protocol in the Internet, is also evaluated. A cost evaluation is performed to determine the network resource consumption of the schemes. Finally, an architecture and extensions of the basic protocol are presented to apply NEMO in satellite networks. This dissertation fosters the application of NEMO to terrestrial and satellite networks by selecting and extending optimal route optimization schemes, and presenting new architecture and protocol.

Chapter 1

Introduction

The advances of wireless technology and the miniaturization of devices have given rise to the demand for Internet connectivity of mobile devices. Examples include devices that connect to the Internet through an onboard Local Area Network, a person carrying several devices which are connected to a Personal Area Network, etc. Internet Protocol (IP) is used to provide Internet connectivity to a device which is identified and located using an IP address. When the device moves between networks in different geographical areas, it has to obtain new addresses. Obtaining a new address requires the use of mobility management protocols to prevent the termination of ongoing sessions as well as unreachability of the deivce.

IP has also been of interest for satellite networks that consist of satellites connecting to each other and ground stations using satellite links. Due to the orbiting motion of satellites, mobility of onboard IP-enabled devices occurs with respect to the Internet. Thus, mobility management of IP-enabled devices onboard satellites are also required.

Internet Engineering Task Force (IETF) standardized host mobility management protocols, such as Mobile IP (MIP) [1] and MIPv6 [2] to maintain session continuity during handover. A summary of the host mobility protocols has been provided by Le et al. [3]. To augment IP with mobility support, these protocols use control messages, called signals, and location management entities to update and track the location of the devices.

Managing the mobility of multiple devices moving together increases the signaling overhead, power consumption and complexity of the manageability. Therefore, host mobility protocols are inefficient to manage the mobility of multiple devices moving together. Moreover, all the devices may not be sophisticated enough to support mobility protocols or have powerful transceivers to communicate with distant access points. Therefore, IETF standardized NEtwork MObility (NEMO) [4] for efficient support of the mobility of multiple devices that are connected as a network called the mobile network. NEMO Basic Support Protocol (BSP) [4] has been proposed to enable communication with the mobile network. However, NEMO BSP suffers from the problem of inefficient route that results in delay in communication and header overhead [5]. Therefore, route optimization is required to solve the problem.

A large number of route optimization schemes have been proposed in the literature [5–11, 11–32]. These schemes raises several performance issues, such as the increase of signaling, degree of optimization, latency of handoff, resource (e.g., memory, processing) requirements, deployability and location transparency. The schemes vary in an effort to trade off the issues that might affect their performance. To find the optimal schemes, a comparative evaluation is required based on the performance difference resulting from the tradeoff. From the evaluations that are found in the literature, it is hard to find which scheme or set of schemes are optimal because either the evaluations do not involve all schemes in one place or the evaluations cannot be assimilated due to the non-homogeneity of the evaluation methods.

Solutions have been proposed to manage the mobility of onboard IP-enabled devices in satellite networks [33–36]. These solutions cannot provide continuous connectivity to the Internet when a satellite loses direct connection with a ground station, although the satellite may have physical connection to a ground station through other satellites.

1.1 Motivation and objectives

Miniaturization of devices and the increase of wireless coverage are leading to mobility of multiple IP-enabled devices, and hence we are to moving towards NEMO. Therefore, there is a need to find optimal route optimization schemes. In an effort to do so, Lim et al. [32] classify the route optimization schemes based on their approaches and provide an analytical model-based comparison among the classes based on handoff latency, resource usage and signaling. However, the classification does not include many recently proposed schemes. Also, the effects of the issues on the performance of the schemes were not investigated. Lim et al.f [37] also presented the issues related to route optimization and the effects of those issues on the infrastructure. However, this does not include the state-of-the-art in the route optimization, and no comparison among the schemes is given. To summarize, existing evaluations do not include state-of-the-art schemes, and they are not comprehensive as they fail to show the effect of the trade off of the issues on the performance of the schemes.

As far as satellite networks are concerned, use of IP in satellite is a reality. Satellites collect earth observing data that are used to monitor flood, wildfire, volcanoes and cryosphere events [38]. At present, IP is being used to transfer imaging data from satellites to the ground to aid in disaster area relief operation [39]. To transfer such data to IP-enabled end users through the Internet, future satellites is expected to contain multiple IP-enabled devices. As the satellites orbit, the IP-enabled devices can be considered as mobile network in motion. Therefore, NEMO can be a natural solution to mange the mobility of these devices to ensure continuous transfer of data to the Internet even when the satellites are not in direct connection with a ground station.

The objectives of this research are as follows:

- The first objective of this research is to perform a comprehensive evaluation of the route optimization schemes. The idea is to review, classify and compare the classes of schemes to select the best performing class of schemes. The selected schemes are then evaluated in terms of various performance parameters to narrow down the selection.

- The second objective of this research is to demonstrate the application of NEMO architecture and protocols to satellite networks for continuous connectivity to the Internet. To achieve the objective, we present an architecture and extension of NEMO Basic Support Protocol to apply NEMO in satellite networks.

1.2 Contribution

To achieve the first objective, this research intends to perform a comprehensive evaluation of the proposed route optimization schemes. We first classify schemes based on the basic idea used for optimizing routes and perform qualitative comparison among the classes. Unlike such previous evaluations, our evaluation is comprehensive, and evaluates the schemes within each class. The performance issues mentioned earlier are used for the comparison that helps to select an optimal class of schemes. Performance difference might exist among the schemes within the selected class. Unlike previous evaluations, we perform further in-depth quantitative evaluations of the selected schemes to determine the performance under various mobility environments. Evaluations reveal inefficiencies of the selected schemes for intra mobile network communications. We extend the schemes to make them efficient. Our work enables one to select optimal schemes for route optimization, show their suitability depending on the mobility environment and increase the efficiency of the optimal performing schemes.

To achieve the second objective, we propose a NEMO architecture for satellite networks and extension of NEMO BSP to take advantage of the multiple connections available in satellite networks. Unlike other architecture and protocols used to connect satellite networks to the Internet, proposed architecture and protocol can provide connectivity to the Internet despite a satellite lose direct connection to a ground station. Thus, results of this research will provide an efficient way for continuous data transfer from IP-enabled devices in satellite networks to terrestrial Internet.

To summarize, contributions of the dissertation are as follows:

- classification and comparison of the route optimization schemes to select the optimal performing class,
- comprehensive evaluation of the schemes within the selected class to find optimal performing schemes,
- development of an analytical framework to show the tradeoff between route optimization and signaling,
- improvement of the schemes within the optimal class for inter and intra mobile network communication, and
- demonstration of NEMO architecture and proposed extension of NEMO BSP for satellite networks.

1.3 Dissertation structure

The rest of the dissertation is organized as follows. Chapter 2 presents a review of NEMO, route optimization schemes, existing evaluations of the schemes, and a classification of the schemes. Chapter 3 presents analytical and simulation model-based performance evaluations of selected schemes, followed by an evaluation of the schemes for TCP-traffic in Chapter 4. In Chapter 5, the effects of increasing the number of communicating hosts in the mobile network are evaluated. Chapter 6 presents the cost evaluation of the schemes. Proposed improvements of the selected schemes are presented in Chapter 7, followed by a presentation of proposed architecture for NEMO in satellite networks in Chapter 8. Finally, concluding remarks are given in Chapter 9.

Chapter 2

Literature Review

NEMO Basic Support Protocol (BSP), which has been proposed to enable communications with the mobile network, suffers from the problem of inefficient route. Therefore, a large number of route optimization schemes have been proposed in recent years. In this chapter, we present an overview of NEMO and problems of NEMO BSP, a survey on the evaluations of the schemes that have been proposed to solve the problems and a classification of the schemes.

2.1 NEMO and its problems

Figure 2.1 shows the NEMO architecture where one or more routers, called Mobile Routers (MRs) (e.g., MR1), act as gateways for the Mobile Network Nodes (MNNs) [4]. When the mobile network moves from one network to another, MRs perform handoff to keep the movement transparent to MNNs. Possible types of MNNs are Local Fixed Nodes (LFNs) that do not move with respect to the mobile network, Local Mobile Nodes (LMNs) that usually reside in the mobile network but can move to other networks, and Visiting Mobile Nodes (VMNs) that belong to other networks but are currently attached to the mobile network. We will refer to LMNs, VMNs and MRs, which use mobility protocols, as mobile nodes. MRs can be MNNs to form a nested mobile network when one mobile network connects to another. A number of MRs connected in series can result in multiple levels of nesting. A Top Level MR (TLMR) attaches directly to the wired network through Access Routers (ARs). In Fig. 2.1, the mobile network under MR1 is nested under the TLMR's mobile network; the MR1's mobile network thus has a nesting level of one.

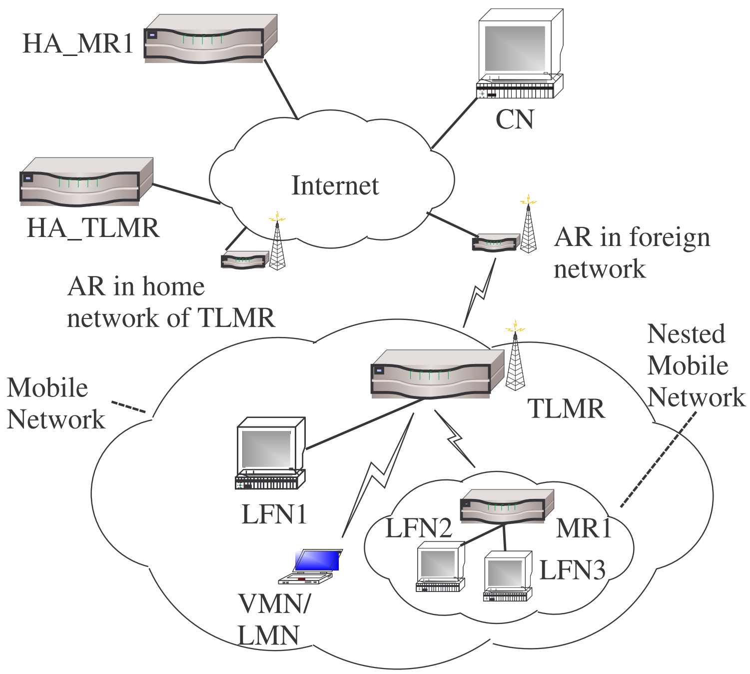

Figure 2.1: Architecture of NEMO showing one level of nesting.

The network to which a mobile network is usually connected is called the home network. An MR is registered with a router, called the Home Agent (HA), in its home network. In Fig. 2.1, HA_TLMR and HA_MR1 are the HAs for TLMR and MR1, respectively. A node that communicates with MNNs is called a Correspondent Node (CN). TLMR has a Home Address (HoA) through which it is reachable in its home network. It is delegated prefixes from its home network to advertise inside its network. When TLMR moves to a foreign network (any network other than the home network), it obtains a new address, called the Care-of-Address (CoA), at the foreign network. TLMR then registers the CoA with HA_TLMR by sending a

Binding Update (BU) that contains a Mobile Router Flag indicating that the TLMR is acting as a router. In response to the BU, HA_TLMR sends a positive Binding Acknowledgement (BA) to indicate that the forwarding to TLMR is set, and creates a binding cache entry that maps TLMR's HoA and prefixes to the CoA of TLMR. Thus, a bidirectional tunnel [40] is created between HA_TLMR and TLMR to tunnel packets exchanged between the CN and MNNs. A nested mobile network is created when MR1 moves under TLMR. MR1 obtains a CoA from the TLMR's prefix, and registers the CoA with HA_MR1 in the same way as described above.

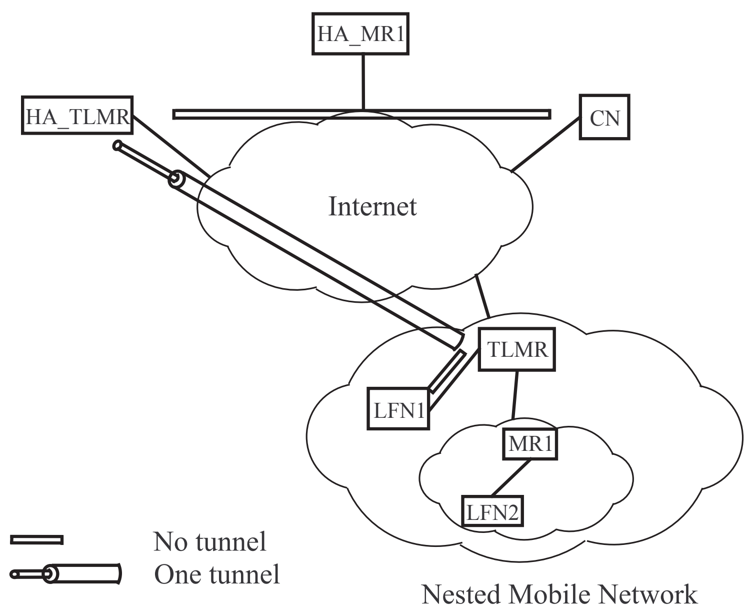

Figure 2.2: Single tunneling when routing packets for LFN1.

Figure 2.2 shows the routing of packets from the CN to LFN1. Since LFN1 obtains its address from TLMR's prefix (delegated by TLMR's home network), the packet is routed towards HA_TLMR. HA_TLMR encapsulates and forwards the packet to TLMR which receives, decapsulates and forwards the packet to LFN1. Packets in the reverse direction take the same path in reverse undergoing encapsulation and decapsulation at TLMR and HA_TLMR, respectively.

Figure 2.3 shows packets going from the CN to LFN2 through multiple tunnels in a nested mobile network. Since LFN2 obtains its address from MR1's prefix (delegated by MR1's home network), the packets are intercepted, encapsulated and tunneled to MR1 by HA_MR1 using MR1's CoA. Since MR1's CoA is obtained from the TLMR's prefix, the packets are intercepted, encapsulated further and tunneled to TLMR by HA_TLMR, resulting in multiple encapsulations. Encapsulated packets, on reaching TLMR, are decapsulated and forwarded to MR1 which again decapsulates the packets and forwards them to LFN2.

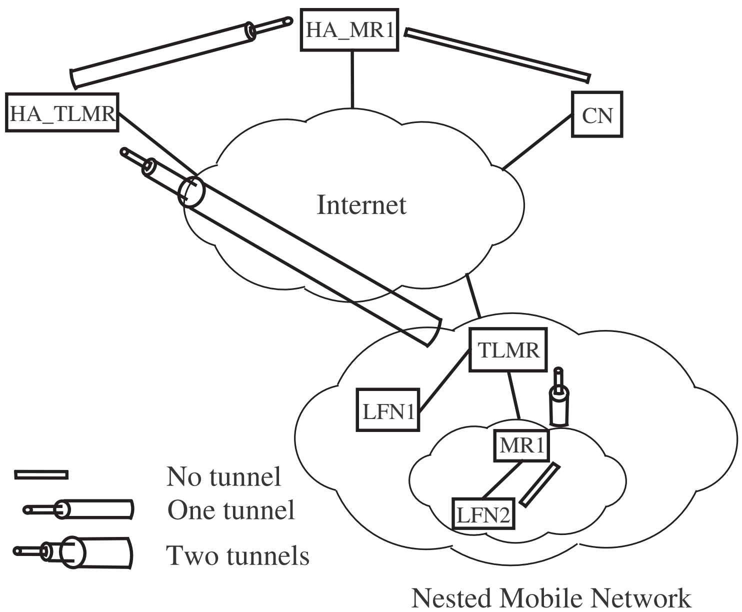

Figure 2.3: Multiple tunneling in a nested mobile network.

It is evident from Figs. 2.2 and 2.3, packets go through one or more bidirectional tunnels between HAs and MRs. Thus, the route traversed by packets may be suboptimal when the mobile network and the CN are in networks that are topologically close but are far away from the home network. The suboptimal route results in inefficiencies, such as large end-to-end delays, additional load on the infrastructure, susceptibility to link failures, etc. that are presented in detail by Ng et. al [41].

The header overhead is another problem that results from the tunneling [41]. As a packet passes through each tunnel, it is encapsulated, resulting in the increase of packet header size. This decreases bandwidth efficiency and increases the chance of fragmentation. Moreover, packets need decapsulated as many times as they are encapsulated requiring additional processing at HAs and MRs. The problems are aggravated when nesting occurs; therefore, route optimization is an active area of research in NEMO.

Route optimization in NEMO requires addressing several challenges which raise issues. The performance and applicability of the schemes, providing the optimization, are affected by those issues [37]. Several schemes, which trade off the gain of the route optimization with their performance and applicability, have been proposed. Route optimization requires bypassing HAs when packets are routed between the CN and MNNs. Bypassing HAs gives rise to the following major challenges which need addressed by the schemes:

- How can a packet destined to an MNN reach TLMR attached to the foreign network?
- How is a packet routed inside the mobile network after reaching TLMR?

The challenges are addressed by the majority of the schemes that focus on optimizing the route between a CN in the wired network and an MNN. Addressing the above-mentioned challenges is insufficient for optimizing the route for communication between two MNNs (intra mobile network communication [41]). In such communications, packets have to traverse through the HAs residing outside the mobile network even though the MNNs are in the same network. The challenge of the route optimization in the intra mobile network case, is how to route packets between two MNNs without sending packets outside the mobile network; some route optimization schemes address this later challenge as well.

Addressing the challenges raises several issues that were reported by Ng et. al [37] as given below.

1. Signaling: When a mobile network moves, only the MR to which the movement is visible needs to perform signaling with its HA. The schemes may require more signaling than NEMO BSP to convey the prefix of the foreign network to the CN, and to find the route from the TLMR to the MNN. Signaling packets compete with data packets for bandwidth, not only inside the mobile network but also in the Internet.

2. Memory requirement: Schemes have to maintain various state information regarding the route and CN-MNN pairs. Maintaining the state information requires memory that can be a limiting factor in memory constrained environments involving small devices like small sensors and PDAs.

3. Degree of optimization: In an effort to trade off issues, such as signaling some schemes allow some non-optimality in the route. To characterize the degree of optimization, we use the terms optimal and near optimal for routes that are optimal and non-optimal (to some extent), respectively.

4. Header overhead: To support NEMO, additional information is put into the header. This results in the degradation of data rate and increase of the chance of fragmentation.

5. Deployability: The schemes propose new functionalities for the hosts, routers in the Internet and mobility entities (e.g., MRs and HAs), and even propose new entities. Changes in the mobility entities are easy to incorporate because they will be introduced in the existing infrastructure. Changes in functionalities of existing hosts and routers in the infrastructure may not be easy to incorporate.

6. Location management: Location management is tracking the location of an MNN to ensure reachability and session continuity. In NEMO BSP and some optimization schemes, location management is performed the by HAs. On the other hand, some schemes propose location management by CNs, TLMRs,

Internet-routers, or by new entities. Location management by HAs, TLMRs or new entities is easily deployable. Location management by CNs or Internet-routers is less prone to failures because of distributed location management. However, it requires changes in existing Internet-routers and hosts raising deployability issue.

7. Location transparency: In NEMO BSP, MNNs (except MRs), and CNs are transparent to location changes. In an effort to optimize route, some of the schemes sacrifice location transparency.

The route optimization schemes that have been proposed to solve the problem of inefficient route vary, as they tradeoff these issues in various ways.

2.2 Evaluation of the schemes

Perera et al. [5] present a comprehensive introduction to NEMO, advantages and limitations of NEMO BSP, a review of the route optimization schemes and future research directions for NEMO. Being an emerging area of research, there have been additional work [6–11,11–31] on NEMO route optimization than those reported in [5]. Although the literature in [5] provides a comprehensive introduction to the problem of route optimization and the solutions, there is no comparison among the schemes. A classification of the route optimization schemes has been shown by Lim et al. [37]. The issues related to route optimization and the effects on the infrastructure are also discussed in this literature. However, this later classification does not include the state-of-the-art in the route optimization, and no comparison among the schemes is given.

Lim et al. [32] classify the route optimization schemes based on their approaches and provide an analytical model-based comparison among the classes. The metrics used for the comparison are the delay to send BU, handoff latency, memory requirements and header overhead. However, the classification does not include many recently proposed schemes. Moreover, the comparisons presented in [32] do not consider the comparisons among the individual schemes in each class to verify that the characteristics of a class represents its constituent schemes. Also, the effects of the metrics the performance of the schemes were not investigated.

In the rest of this chapter, we provide a comprehensive up-to-date summary of the route optimization schemes, and classify and compare the schemes (like in [32]). Unlike that in [32], we also provide a qualitative comparison among the schemes in each class. We also introduce two new classes of schemes.

2.3 A review of route optimization schemes

Based on approaches used, the route optimization schemes that have been proposed can be generally classified as:

- Prefix Delegation
- Hierarchical
- Source routing
- BGP-assisted

In the rest of this section, we present the basic principle of each class, and a description and comparison of the schemes.

2.3.1 Prefix Delegation

In this class, a prefix of the foreign network is delegated inside the mobile network. Mobile nodes obtain CoAs from the prefix and send BUs to their respective HAs and CNs. Therefore, any packet from a CN, addressed to the CoA, reaches the foreign network without going through HAs. For example, as shown in Fig. 2.4, prefix

2001:afce:1ff3:: is relayed by TLMR inside its mobile network. VMN1 and MR1 obtains CoAs 2001:afce:1ff3:110 and 2001:afce:1ff3:11a, respectively. Then MR1 relays the prefix inside its network. The process of obtaining CoAs from foreign network's prefix and provision of routing packets inside the mobile network vary among the schemes in this class resulting in differences in signaling and memory requirement. A comparison based on the differences are presented in Table 2.1.

The concept of prefix delegation is simple, and provides optimal route with low header overhead at the cost of sacrificing location transparency. Moreover, sending BUs to CNs requires additional signaling along with the requirement of protocol support (location management along with HA) from CNs, making the schemes difficult to deploy. Also, the schemes do not focus on intra route optimization.

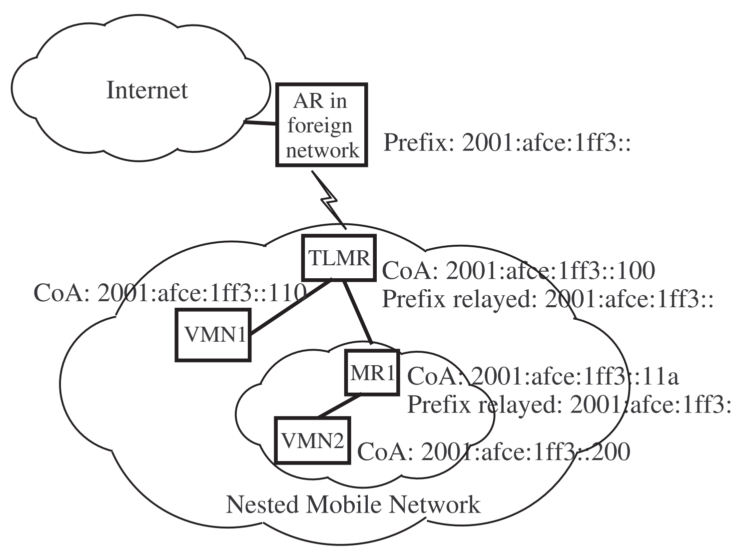

Figure 2.4: Delegation approach for route optimization.

2.3.1.1 Simple Prefix Delegation (SPD)

In Simple Prefix Delegation, proposed by Lee et al. [42], a the foreign network prefix is hierarchically delegated to the MRs. Each MR advertises the delegated prefix

Table 2.1: A comparison of the schemes in prefix delegation class schemes.

Scheme	**Signaling**	**Memory requirement**	**Other overheads**
Simple Prefix Delegation	Medium	Lower	Higher end-to-end delay for LFNs
ND-Proxy	Medium	Low	Additional delay at the start of communication
Optinet	High	High	None
MIRON	High	High	None
Ad hoc-based	Medium	Low	Higher end-to-end delay for LFNs, Flooding of ad hoc protocol messages
OPR	Low	High	Per packet processing
HIP-based	High	High	Per packet processing

inside its own network using Delegated Prefix Option in the header. Since prefixes are hierarchically delegated, packet forwarding inside the mobile network can be done based on the prefix of packets' destination address. This scheme, however, requires a prefix delegator in every mobile network, requiring additional overhead of performing extra functionality related to prefix delegation. Its signaling amount is proportional to the number of mobile nodes and in between low and high (i.e., medium) amount of signaling of other schemes in this class. Memory requirement is low because only attached MRs' prefixes needs to be tracked as the next hop. The advantage of the signaling amount not being high costs incomplete route optimization for LFNs whereas routes for mobile nodes is optimal.

The scheme, proposed by Mimoune et al. [6], is very similar to the Simple Prefix Delegation scheme [42] in terms of delegating prefixes and obtaining care-of-addresses. Unlike the scheme proposed in [42] where CNs are updated by mobile nodes, MRs notify the border routers in home network about the delegated prefix. Border router makes an entry that maps home prefixes of MRs to delegated prefixes, and informs other border routers that eventually inform CNs about the prefixes after the communication starts. CNs configure the MNN's care-of-address by combining its HoA and the prefix informed by the router, and can send packets using the care-of-address. Although the scheme reduces signaing load which is distributed among border routers, functionalities of CNs and routers need to be changed requiring change in infrastructure.

2.3.1.2 Neighbor Discovery Proxy (ND-Proxy)

In this scheme, proposed by Jeong et al. [43], route optimization is achieved by advertising the prefix of the foreign network inside the mobile network. Each MR obtains a care-of-address from the advertised prefix and advertises the prefix inside its mobile network. All mobile nodes use the advertised prefix to obtain care-of-addresses. Routing of packets is different from Simple prefix delegation (where prefixes are hierarchically delegated) because all addresses are obtained from one prefix. When TLMR receives a packet destined to an MNN, and the nexthop for the destination is not present in the routing table, the TLMR makes a neighbor discovery query to find the next hop for the MNN's care-of-address. An MR attached below responds if the MNN's care-of-address that is being sought is attached to the MR. Otherwise, the MR relays the search message to MRs underneath, and replies to the query when an MR underneath responds with the care-of-address being sought. Thus, MRs actually act as proxy for MNNs for neighbor discovery.

Finding the nexthop using neighbor discovery introduces delay at the start of communication. Yet, this scheme has the advantage of not requiring a prefix delegator in every mobile network. Amount of signaling is similar to that of Simple prefix delegation whereas memory requirement is little higher (hence, low instead of lower) than Simple Prefix Delegation because of maintaining routing entries for all communicating MNNs underneath an MR.

In another scheme proposed by Song et al. [44], MRs advertise (like MRs in [43]) foreign network's prefix to attached MRs only. MRs perform route optimization on behalf of attached MNNs by sending BUs to corresponding CNs. Since MRs need send BUs to all CNs and track all CNs, signaling amount and memory requirement for the scheme is high in this scheme.

2.3.1.3 Optimal routing for network mobility (Optinet)

Perera et al. [45] proposed a scheme called Optinet which is similar to Simple Prefix Delegation but with different prefix delegation procedure. Unlike obtaining a prefix directly from the advertised prefix in Simple Prefix Delegation, a DHCP client in an MR obtains a prefix from the network it attaches to (a mobile network or a wired network). Petander et al. [46] extended Optinet (xOptinet) that reduces signaling by restricting the obtaining of the CoAs to only those nodes that are actively communicating with CNs during the handoff. Moreover, xOptinet optimizes the route for LFNs by having MRs perform route optimization signaling on behalf of attached LFNs. Unlike other schemes in Sec. 2.3.1, Optinet requires a DHCP client and a server at every mobile network. Moreover, LFNs' route optimization requires sending BUs to CNs, and tracking LFN-CN communications resulting in high amount of signaling and memory requirement, respectively.

2.3.1.4 Mobile IPv6 Route Optimization for NEMO (MIRON)

In MIRON, proposed by Calderon et al. [47, 48], mobile nodes obtain the care-of-address from the prefix of the foreign network. Upon attachment to an MR, a mobile node obtains a care-of-address from the MR's home prefix, and sends a BU to its HA. The MR intercepts the BU, and notifies the mobile node to obtain a new care-of-address using PANA [49] and DHCP. A mobile node sends a DHCP request to obtain a care-of-address. Instead of relaying the prefix inside the mobile network, this scheme relays the request to the DHCP server at the foreign network. An assigned care-of-address is then relayed back to the mobile node. Relaying is performed by DHCP client and server component in MRs. After obtaining a care-of-address, MR notifies the attached mobile nodes to obtain care-of-addresses. This procedure of obtaining a care-of-address is repeated at each handoff, and takes longer time than it takes in other schemes presented in Sec. 2.3.1. Like xOptinet, MIRON optimizes the route for the LFNs by having the MR perform route optimization signaling on behalf of attached LFNs, and hence signaling and memory requirement is high.

2.3.1.5 Ad hoc-based (Ad hoc)

In Ad hoc-based scheme proposed by Su et al. [7], mobile nodes obtain care-of-addresses from the prefix of the foreign network. Unlike ND-proxy, for routing inside mobile network, the route between the MNN and the AR in foreign network is discovered using an Ad hoc protocol. Route discovery requires flooding of ad hoc protocol messages that consume bandwidth and introduce a delay at the start of a communication or after a communication is interrupted due to handoff. Moreover, Ad hoc network protocols are intended for unstable networks, and does not take advantage of the hierarchical nature of the nested mobile networks. Since it does not optimize route for any node, signaling requirement is similar to Simple Prefix Delegation at the cost of higher end-to-end delay. Due to the maintenance of routing entries for all communicating MNNs, memory requirement is high in this scheme.

2.3.1.6 Optimal Path Registration (OPR)

In OPR proposed by Park et al. [8], like ND-Proxy or Ad hoc-based scheme, the prefix of the foreign network is advertised inside the mobile network. The difference of OPR with ND-Proxy and Ad hoc-based schemes is that the prefix is relayed only to the MRs, resulting in movement transparency for other MNNs. To provide movement transparency, MRs translate prefix of source and destination addresses of outbound and inbound packets of its network. Translated addresses are similar to CoAs. Movement transparency costs additional memory due to the maintenance of a translation table and processing cost per packet for address translation. Signaling is low in OPR than other schemes discussed in Sec. 2.3.1 because of not sending BUs to CNs. To compensate for not sending BUs, CNs are informed of the change in translated address by marking the packet's header. This costs the MR additional processing overhead per packet and high memory due to state management for tracking every CN-MNN communicating pair.

Kim et al. [50] proposed a scheme that improves the performance of OPR by reducing the number of BUs. To reduce the number of BUs, the scheme proposed by Kim requires all HAs to join a multicast group that is managed by either the AR or the HA depending on when the multicast group is formed. For location update, the AR sends a BU to the multicast group that reaches all HAs. However, this scheme requires ARs functionality to be modified for route optimization, and hence not easily deployable.

2.3.1.7 HIP-based (HIP)

A Host Identification Protocol (HIP)-based route optimization is proposed by Novaczki et al. [9,10]. Like MRs in other schemes in prefix delegation class, the mobile Rendezvous Servers (mRSVs) in HIP-NEMO obtain prefixes from the foreign network, and delegate part of the prefixes to attached mRSVs. The prefixes are then advertises inside their mobile network. For route optimization, an mRSV uses the prefix as location identifier of MNNs when sending location updates to CNs and RSVs (acts like HA and DNS), and translates the source/destination address of outgoing/incoming packets. When an mRSV attaches to an AR, it obtains a new prefix, performs location update signaling with CNs and RSVs on behalf of MNNs, and updates the prefix of the attached mRSVs that also do the same. Location updates, attaching to a mobile network, and delegation of signaling to mRSVs are performed according to HIP [51]. Signaling amount for this scheme is the same as MIRON because of sending location updates to all CNs. Memory requirement and per packet processing overhead is like OPR because of the similarity in address translation process. Although the mobility management entities that perform location update and address translation are entities defined in HIP, the basic approach is essentially the same as the other schemes in this class, and hence we include the scheme in this class.

2.3.2 Hierarchical

In the hierarchial class, a packet reaches the foreign network traveling through one HA instead of traveling through all HAs. Unlike prefix delegation-based approach, a nested MR does not send its addresses to CNs. Rather, the MR sends the TLMR's care-of-address or home address to its HA. CNs use the address to send packets to a mobile network node. Packets sent by CNs reach node's HA that tunnels the packets to TLMR's address. Packets tunneled to care-of-address directly reach the foreign network, whereas packets tunneled to home address reach TLMR's HA that tunnels packets to the TLMR. On reaching the TLMR, packets are routed to mobile network nodes by MRs that maintain a routing table containing the mapping of mobile network's prefix to next hop MR.

Figure 2.5 shows an abstract view of the hierarchical approach. TLMR_CoA is passed to HA_MR1 and HA_VMN by MR1 and VMN, respectively. Also, MR1 and VMN send their care-of-addresses to TLMR to enable forwarding inside the mobile network. Therefore, a packet sent to VMN will first reach HA_VMN that tunnels the packet to the TLMR for forwarding towards the VMN. Thus, communication route is divided into two parts: the route between TLMR and HA_VMN, and the route from TLMR to VMN. At least one tunnel always exists between TLMR and HA_VMN.

The schemes in this class mainly differ in the use of TLMR's care-of-address or home address for tunneling, techniques to convey TLMR's address to MRs, and routing of packets inside mobile network resulting in differences in signaling, memory requirement and degree of route optimization. Moreover, depending on the use

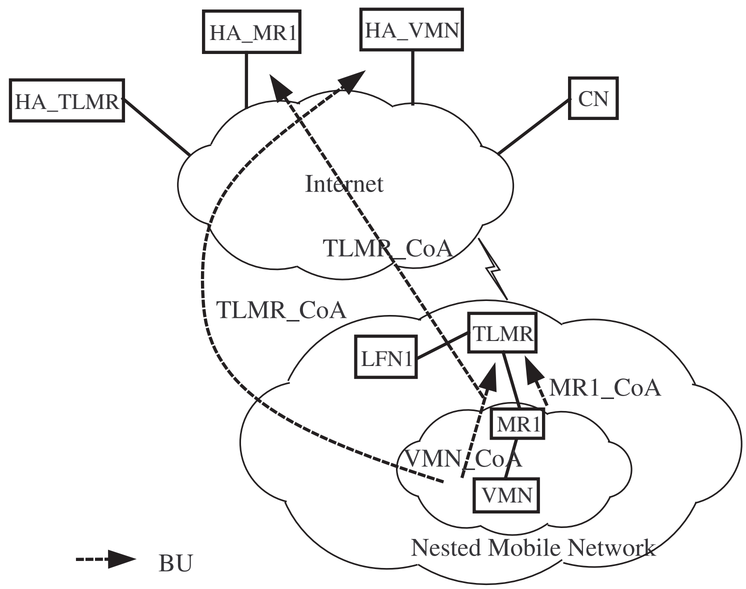

Figure 2.5: Hierarchical approach for route optimization.

of home address or care-of-address of TLMR, the number of tunnels used for communication differs among the schemes; number of tunnels affects degree of route optimization and header overhead. In addition, location management entities also vary among the schemes. A comparative summary based on these differences are presented in Table 2.2.

The schemes in this class require fewer number of signaling than prefix delegation-based schemes because no BU is sent to CNs (except the schemes proposed in [17, 24, 51, 52]). This also makes CNs transparent to the mobility of communicating MNNs, yielding location transparency and easy deployability. Additionally, no BU is sent to the HA for intra mobile network movement because of unchanged TLMR address, resulting in reduced signaling; this resembles Hierarchical MIPv6 [53], and hence the name hierarchical. Moreover, the schemes in this class focus on Intra route optimization. The schemes (except those proposed in [17, 24, 51, 52]) have the

Table 2.2: A comparison of the route optimization schemes in hierarchical class.

Scheme	**Degree of route optimization**	**Tunnels**	**Signaling**	**Memory requirement**	**Location management**
ONEMO	Near optimal	One	Medium	Low	HA and Router in CN's network
ROTIO	Near optimal	Two	Lower	Low	HA, TLMR's HA and TLMR
LRO	Near optimal	One	Low	Low	HA and LMA
χLIN6-NEMO	Near optimal	None	High	High	Mapping agent
HMNR	Near optimal	One	Medium	Low	HA and TLMR
ROAD	Optimal	None	High	Low	HA and CN
HMSRO	Near optimal	One	Low	Low	HA and TLMR or AR
Light-NEMO	Near optimal	One	Low	Low	HA and TLMR
Light-NEMO extended	Optimal	None	High	High	HA and CN
ROPIO	Near optimal	One	Medium	Low	HA and TLMR
HMNB	Near optimal	Two	Lower	Low	HA, TLMR's HA and TLMR
HIP-based	Optimal	None	High	High	TLMR and CN
HMIP-based	Near optimal	Two /three	Low	Low	HA and MAP
MoRaRo	Optimal	None	High	High	TLMR and CN

disadvantage of packets going through one or two tunnels, resulting in near optimal routes and header overhead.

2.3.2.1 Optimized NEMO (ONEMO)

This scheme, proposed by Watari et al. [11], uses extended router advertisement message to convey the care-of-address of the TLMR to nested MRs which send the to the care-of-address to the nearest router of the CN. The router tunnels packets (sent by the CN) to the TLMR. The router that keeps track of the MR's location is discovered by the MR when it receives the first packet through the HA. Discovery is initiated by the MR by sending a message to an anycast address which is configured from the prefix of the CN. Discovery of the router and sending the care-of-address to it require additional signaling; we thus consider it as medium (instead of low) signaling in Table 2.2. The router tunnels the packets, sent by the CN to an MNN, to the TLMR that routes the packets inside the mobile network. Unlike most of the schemes in this class, deployability of the scheme is difficult due to the requirement of support from a router in each CN's network.

2.3.2.2 Route Optimization using Tree Information Option (ROTIO)

In ROTIO, proposed by Cho et al. [12], home address of the TLMR and care-of-addresses of intermediate MRs (MRs between the TLMR and an MNN) are conveyed to the nested MR using router advertisement messages that contain tree information option representing the nesting structure. Each MR appends its care-of-address to the advertisement sent by the TLMR, and relays the advertisement down the nesting level. Thus, an MR knows the care-of-address of intermediate MRs from the advertisement, and sends two BUs: one to its HA to send the HoA of the TLMR, and another to the TLMR to send a list of care-of-addresses of intermediate MRs. Therefore, MR's location is tracked through the HA of the MR, TLMR's HA and the TLMR, and the HA of the MR can tunnel packets to TLMR's home address.

TLMR, knowing the nesting structure of the mobile network from the BUs, can route packets inside the mobile network. The disadvantage of this scheme is packets going through two tunnels- one between the MR's HA and the TLMR's HA, and another between the TLMR's HA and the TLMR. One additional tunnel as compared to ONEMO is compensated by lower amount of signaling.

2.3.2.3 Route optimization for nested mobile network in local mobility domain using local mobility anchors (LRO)

In LRO proposed by Li-hua et al. [13], a prefix used in local mobility domain is advertised to all MRs through extended router advertisements. MRs obtain CoAs from the prefix, and send BUs to their HAs. BUs are also sent to the Local Mobility Anchor (LMA) to send the CoA, HoA, MRs home prefix and address of the HA; therefore, the LMA creates binding entries, and performs location management along with the HA. A packet, sent from a CN to an LFN, reaches the LFN's HA that tunnels the packet to the MR. The packet reaches the LMA that searches the destination (CoA) in the binding entries. On finding the destination, the LMA forwards the packet to the MR through intermediate MRs and routers that already have routing entry (created from BUs sent by the MR to the LMA) for that CoA. On reception of a packet from local domain, the LMA decapsulates the packet to search its binding entry for the prefix of the inner destination. If found then the packet is routed within the local domain; otherwise, the packet is encapsulated again and forwarded to the HA. Therefore, the scheme can handle intra route optimization in a near optimal way but involves one tunneling in all cases.

2.3.2.4 NEMO protocol based on Location Independent Networking in IPv6 (χLIN6-NEMO)

In χLIN6-NEMO proposed by Banno et al. [14,15], MRs obtain the prefix of the foreign network through extended router advertisements, and send the prefix (through

BU) to the Mapping Agent (MA) which acts (e.g., performs location management) like the HA. The MA intercepts the packets that are sent by a CN to an MNN, replaces the prefix of the destination address with the prefix of the foreign network, and forwards the packets to the MNN. Unlike other schemes in this class, packets reach the TLMR through the MA (therefore, near optimal route) without any tunnel because of the prefix replacement procedure. The TLMR forwards the packet inside the mobile network after restoring the prefix of the destination address to MNN's prefix. Location Independent Networking is achieved by always using the prefix of the network at MNN's current location. Translation of prefix is transparent to the transport layer or above where a location independent address, formed by combining a location independent identifier and a prefix, is used. The scheme decreases the chance of single point of failure by employing multiple MAs dispersed in the Internet resulting in increased signaing required to update all MAs. Moreover, memory requirement is high due to TLMR's tracking of MNNs' prefix used for forwarding packets inside the mobile network.

2.3.2.5 Hierarchical Mobile Network Routing (HMNR)

In HMNR proposed by Jeong et al. [54], an extended router advertisement is used to convey TLMR's care-of-address to MRs. Mobile nodes send BUs, containing TLMR's care-of-address, to respective HAs. Therefore, both HA and TLMR in combination keeps track of the location of nodes. Packets sent from a CN to an MNN reach the HA that tunnels the packets to the TLMR. To route packets from the TLMR to the MNN, each MR maintains a routing table that maps the prefix of an MR to the next hop address. The table is constructed by MRs from the BUs sent from MRs below. Memory requirement for the table is low because the number of MRs is small in a mobile network. But, the BUs sent from MRs below in addition to the BUs sent to HAs result in a signaling amount which is higher than LRO. Kim et al. [16] proposed another scheme which is similar to HMNR in terms of conveying TLMR's care-of-address to MRs. Unlike HMNR where the care-of-address is obtained from the prefix of the mobile network's home prefix, MRs obtain care-of-addresses from the home prefix of the TLMR. Moreover, a routing protocol, preferably RIPng [55], is used to route packets inside mobile networks.

2.3.2.6 Route Optimization using Additional Destination-information (ROAD)

ROAD, proposed by Park et al. [17], is very similar to HMNR except that this scheme proposes mobile nodes to send BUs, containing TLMR's care-of-address, to CNs. The care-of-address is used by the CNs as destination address of packets sent to MNNs. BUs also contain node's home address and care-of-address that are put into an additional header of packets sent by CNs. Each MR has a prefix to care-of-address (of lower level MRs) mapping that is used to overwrite the destination and source of incoming and outgoing packets, respectively. Unlike most of the schemes of this class, this scheme avoids tunneling packets through the HA to secure optimal route at the cost of increased signaling due to sending BUs to CNs.

2.3.2.7 Hierarchical Mobility Support for Route Optimization (HMSRO)

Kuo et al. [18] proposed HMSRO which is very similar to the HMNR scheme except the routing table construction process. Unlike HMNR, an MR constructs the routing table using the BUs, sent by MRs to the TLMR, resulting in fewer number of BUs (i.e., signaling) as compared to HMNR. A scheme proposed by Kim et al. [56] is similar to HMSRO with three exceptions - AR's address is used instead of TLMR's care-of-address, source routing is used to route packet inside the mobile network, and mobile nodes can send BUs to CNs resulting in increase of signaling amount.

2.3.2.8 Light-NEMO

Light-NEMO, proposed by Jouaber et al. [52], is similar to HMSRO except the creation of routing entries by intermediate MRs from the BU sent by a mobile node to its HA. Moreover, each MR swaps the source address of the packet by its own care-of-address; eventually the HA gets the TLMR's care-of-address which is used by the HA to tunnel packets to the mobile network. Light-NEMO is extended to remove the HA-TLMR tunnel by having mobile nodes performing MIPv6 route optimization and MRs performing route optimization on behalf on LFNs [57]. Therefore, unlike most of the schemes in this class, signaling and memory requirement is high in this scheme. Unlike HMSRO, in extended Light-NEMO, CNs performs location management along with HAs.

2.3.2.9 Optimization using Prefix Information Option (ROPIO)

In ROPIO, proposed by Lu et al. [19], TLMR's prefix and care-of-address are advertised (using PIO) to nested MRs that obtain care-of-addresses from the TLMR's prefix. A nested MRs send one BU containing its care-of-address to the TLMR, and another containing the TLMR's care-of-address to the HA. Thus, the HA and the TLMR in combination keep track of the MR's location. Packets sent from a CN reaches the HA that tunnels packets to the TLMR. The TLMR decapsulates and tunnels the packets to the nested MR. Packets on the reverse path are tunneled to HA by MR, and are decapsulated by TLMR that checks if the destination prefix is registered with it. If yes then the packet is tunneled to the MR corresponding to the registered prefix (intra route optimization). Otherwise, the packet is tunneled to the HA with with the source address changed to the TLMR's care-of-address. The scheme is similar to HMSRO except the process of conveying the TLMR's care-of-address to nested MRs, and therefore, signaling and memory requirement are similar to HMSRO.

2.3.2.10 Hierarchical Mobile Network Binding (HMNB)

Jeong et al. [20] proposed HMNB which is similar to HMNR and HMSRO with the exception that in HMNB, like ROTIO, MRs send TLMR's home address (instead of care-of-address) to respective HAs. Thus, MRs don't need to send any BU when the TLMR changes network, resulting in less signaling. Disadvantage of this scheme is packets' traversal through two tunnels in contrast to one in HMNR and HMSRO. Depending on the handoff frequency, the scheme proposed by Jeong et al. [58] proposes to switch between HMNR and HMNB to tradeoff signaling with one additional tunneling.

2.3.2.11 HIP-based

Host Identity Protocol (HIP) [51], which supports mobility and multihoming for hosts, is used for NEMO in the scheme proposed by Ylitalo et al. [21]. In HIP, each host uses a unique address at upper layers, and location changes are managed transparently at HIP or lower layers. At the start of communication in HIP, hosts (one may be an MNN) establish a key that is used for location update. The basic principle of HIP-based NEMO is the use of the key to authorize MR to perform location update on behalf of MNNs. Authorization takes place when an MNN joins the mobile network; in nested NEMO, authorization is performed at various levels. When a packet is sent from an MNN to a CN, each MR uses prefix translation of the source address to avoid tunneling. Level by level authorization and prefix translation require maintenance of all HIP sessions going through an MR, and hence high amount of memory. Although the scheme does not directly resemble other schemes in the hierarchical class, the scheme is included here because signaling is performed by the TLMR on behalf of all MNNs. Also, like the most of the schemes in hierarchical class, the TLMR performs location management. Major disadvantages of this scheme are difficulty in deployment due to the requirement of HIP in hosts, high amount of signaling to update CNs, and high memory requirement for TLMR to maintain the states for all HIP sessions going through it.

2.3.2.12 HMIP-based

Route optimization based on HMIPv6 [53] is proposed by Ohnishi et al. [59] where mobile network nodes obtain two care-of-addresses - a Regional care-of-address (RCoA) obtained from the prefix of the Mobility Anchor Point (MAP), and a Local care-of-address (LCoA) obtained from the prefix of the mobile network. The nodes also send two BUs - one to the MAP and another to the HA. The MAP creates a binding entry from the BU that contains the RCoA, LCoA and prefixes of the MR, and also extracts the tree structure (used for routing header) of the mobile network. The HA creates a binding entry from the BU containing the home address and RCoA. Thus, the HA and the MAP keep track of the location of the node.

Packets, sent from a CN to MNNs, reach the HA that tunnels the packets towards the MAP using the RCoA; the MAP uses the LCoA to tunnel the packets towards the mobile network along with the specification of the route inside the mobile network in routing header. Packets sent from the mobile network nodes to the CN are tunneled by each intermediate MR to avoid ingress filtering, and the MAP detunnels the packets before forwarding to the CN. Unlike other schemes in this class, the MAP performs the route optimization functionalities of the TLMR to avoid sending BUs to HAs for the movement under the same domain, incurring a low amount of signaling at the cost of additional tunnels.

The scheme proposed by Kim et al. [60] differs from the scheme proposed by Ohnishi et al. [59] in routing packets inside the mobile network. Packets sent from MNNs to CNs are tunneled by the MR to its HA using the RCoA instead of the LCoA as the source to avoid further tunneling.

Schemes proposed by Park et al. [22] and Hu et al. [23] extend HMIP-based scheme to reduce the number of BUs when the mobile network moves out of the

MAP's domain, and also avoids tunneling that is required to avoid ingress filtering. Extended scheme proposes that VMNs and LMNs send the LCoA (instead of the RCoA) to their HAs using BU resulting in no BU to be sent when mobile network moves under a different MAP.

2.3.2.13 Mobile router-assisted route optimization for NEMO (MoRaRo)

In MoRaRo, proposed by Kafle et al. [24], after receiving the first packet through the HA, a mobile node sends the TLMR's care-of-address to CNs. The mobile network nodes also sends its addresses, and CN's address to the TLMR that creates a binding cache used for routing packets to the nodes, and to send BUs to CNs on behalf of the nodes. Therefore, CNs along with the TLMR performs location management for the nodes. Each MR registers to the MR attached above with its prefix and all prefixes that are reachable through it, and thus MRs are able to route packets inside the mobile network.

Unlike most of the schemes (except HIP-based and ROAD) in this class, CNs can send packets to the nodes without tunneling using the TLMR's care-of-address but at the cost of increased signaling that results from sending BU to CNs. Home addresses of the nodes is put into an additional header in the packet, and used by the TLMR to tunnel the packet to the nodes. Like MIRON, the scheme also proposes route optimization for LFNs by having MRs acting as proxy resulting in high memory requirement for tracking LFN-CN communications.

2.3.3 Source routing

In this class, route optimization is achieved by sending the care-of-addresses of MRs to the CN which, like source routing, inserts the care-of-addresses in the packet header to reflect the nesting structure of the MRs. This however, results in increased header overhead. Packets from the CN reach the TLMR in an optimal route (without going through HAs). Routing within the mobile network is done using the care-of-addresses in the packet header. Memory requirement for routing entries is low because each MR needs to keep track of only the attached MRs as next hop. Schemes in this class notify CNs about the care-of-addresses of MRs in various ways that will be detailed in the descriptions of the schemes. Notification of care-of-addresses to CNs sacrifices location transparency and deployability, and increases signaling. Methods of sending care-of-addresses to CNs result in differences in signaling and overheads. Moreover, the schemes also have different memory requirement for routing packets inside the mobile network, as shown in Table 2.3.

Table 2.3: A comparison of the route optimization schemes in source routing class.

Scheme	**Signaling**	**Memory requirement**	**Other overheads**
S-RO	High	Low	Large delay to converge to optimized route
xMIPv6	Low	Low	One tunnel is required for communication
PCH-based	Low	Low	Requires a router in every network to support the protocol
SIP-based	High	High	None

Figure 2.6 shows the basic principle of the source routing approach where the care-of-addresses of TLMR, MR1 and VMN are inserted in packet's header. Packets, on reaching TLMR, are source routed (using the care-of-addresses) inside the mobile network by TLMR and MR1.

2.3.3.1 Simple Route Optimization (S-RO)

In S-RO, proposed by Kim et al. [61], initially the MRs send their care-of-addresses to their respective HAs. Packets sent from a CN are thus encapsulated by the HAs; MRs decapsulate the packets, and send BUs to the source of decapsulated packets. The CN then gets care-of-addresses of the MRs, and sends packets directly to the TLMR with the list of care-of-addresses in the packet header. This scheme suffers from a large delay for the CN to receive all the care-of-addresses required for

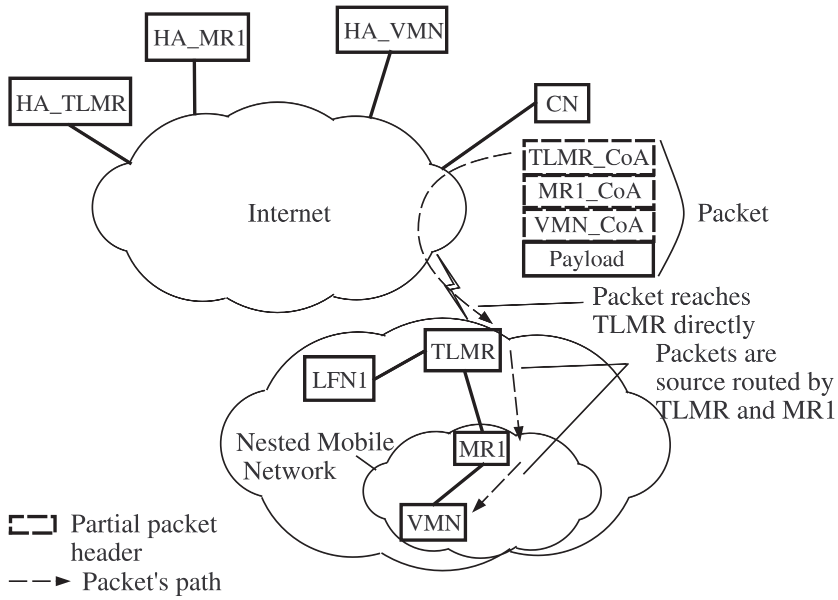

Figure 2.6: Source routing approach.

complete route optimization; this is especially true for higher nesting level. Sending of BUs to HAs and CNs results in a large amount of signaling.

2.3.3.2 xMIPv6

In xMIPv6, proposed by Gu et al. [62], MRs send BUs containing care-of-addresses of MRs above to their corresponding HAs. An MR obtains care-of-addresses of MRs above from the MR to which it is attached. Packets sent from the CN to an MNN reach the HA that inserts the care-of-addresses in the header of packets. Unlike S-RO, xMIPv6 does not need BUs from all MRs, resulting in the advantage of reduced signaling and smaller time for the HA to get care-of-addresses of all MRs above. Unlike other schemes in this class, packets will always go through a tunnel between the HA and the corresponding MR.

2.3.3.3 Path Control Header (PCH)-based

Na et al. [63] proposed a scheme where the care-of-address of the MR is inserted into packets by the corresponding HA, when a packet travels from an MNN to a CN. After passing through all the HAs, the packet's header contains the care-of-addresses of all MRs above. Path control is achieved by a specific router (between the last HA and the CN) that extracts the care-of-addresses to insert in the packet's header sent from the CN to the MNN. Like xMIPv6, this scheme has the advantage of low amount of signaling because of absence of BUs from MRs.

2.3.3.4 Session Initiation Protocol (SIP)-based

Huang et al. [25, 26] proposed a SIP-based [64] route optimization scheme which, unlike other schemes, uses SIP session establishment procedure to discover an optimized route prior to the start of data communication. An MNN (SIP client) sends a SIP invite request to the CN (SIP client) to establish a session. A SIP Home Server (acting like a HA) inserts care-of-addresses of corresponding SIP-network mobility servers (acting like MRs) into the invite request that reaches the CN with the care-of-addresses of the network mobility servers. Route optimization is achieved by the CN inserting the care-of-addresses of the network mobility servers in the packets sent to the MNN. At handoff, the SIP-Network mobility server at the top sends invite requests (through a SIP-foreign server) to all CNs on behalf of MNNs. Sending invite requests results in high volume of signaling as well as high memory requirement due to the tracking of all SIP sessions.

2.3.4 BGP-assisted

Unlike the the schemes described so far, the schemes in this class rely on BGP [65] for mobility management. When the mobile network moves, BGP routers are updated to make necessary changes in the routing tables by making forwarding entries for the prefix of the mobile network. Information regarding the change of route of the mobile network is signaled to few routers that exchange the information with peers using existing routing protocols in the Internet. Therefore, routers contain routing entries to route packets to the mobile network irrespective of its location, and are responsible for location management. Schemes in this class mainly differ (see Table 2.4) in the number of external BGP updates generated, and incurring other overheads for managing intra route optimization.

Table 2.4: A comparison of the route optimization schemes in BGP-assisted class.

Scheme	**Number of external BGP updates**	**Other overheads**
CUIP-NEMO	Medium	Signaling to discover COR
WINMO	Low	Overhead of key management among routers
Multiple P2P connected HA-based	None	Signaling for P2P communication among HAs, and for discovery of the closest HoA

An abstract view of the approach used in this class has been shown in Fig. 2.7. When the TLMR attaches to the AR in the foreign network, the AR injects a BGP update that maps TLMR's prefix (1:3:1::) to AR's address (1::2). BGP router3 in AR's network updates its peers (BGP router1 and BGP router2), accordingly. Therefore, packets sent by the CN will reach a BGP router in its network and will be forwarded to the appropriate BGP router's network where the mobile network resides.

The major advantage of the schemes in this class is the use of no new entity for mobility management. Moreover, CNs are transparent to the change of locations (managed by BGP routers) of MNNs. On the other hand, these schemes will produce a storm of updates (i.e., signaling) in the Internet when the mobile network moves frequently. Moreover, the scalability is also an issue due to the maintenance of routing entries for a large number of mobile networks. Storm of updates and the scalability have been traded off (i.e., reduced signaling and increased scalability) with some additional support from the infrastructure resulting in the difficulty of

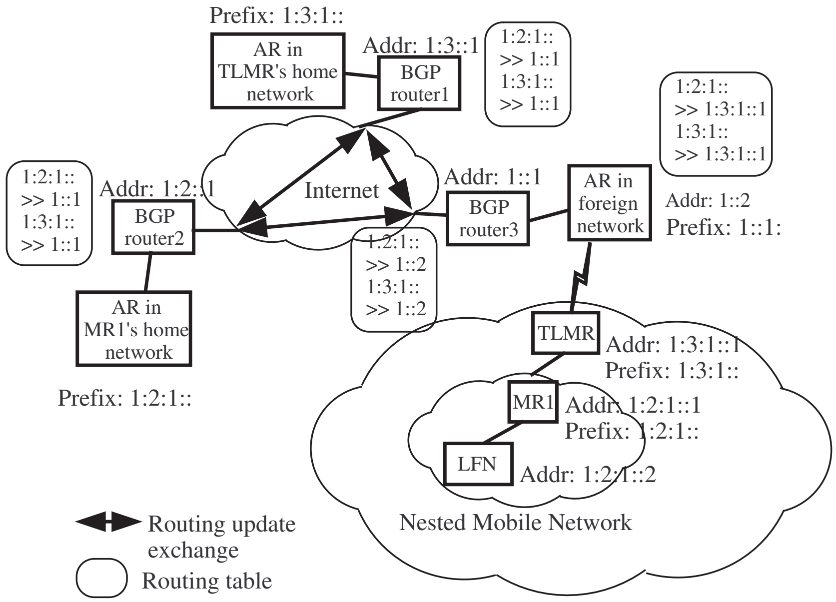

Figure 2.7: BGP-assisted approach.

deployment. Such a trade off also requires packets always traveling through one or more of some designated routers resulting in near optimal routes.

2.3.4.1 Cellular Universal IP for nested network mobility (CUIP-NEMO)

CUIP-NEMO, proposed by Lam et al. [27], is based on Cellular Universal IP (CUIP) [66] where universal addresses are used for a set of mobile nodes that are assumed to be in the same hierarchy of network irrespective of their location. The hierarchy is rooted at a BGP router of the provider network (home network) of mobile nodes, and these routers are directly linked at the network layer. Between the movements of the node there is a Cross-Over Router (COR) that is the first router in the hierarchy common in both previous and the current route. All routers upto the COR is updated with the new route of the mobile node using CUIP signaling. A packet sent to the mobile node is routed towards the CN's nearest BGP router that routes the packet towards the BGP router closest to the COR. Routers before the

COR use prefix-based routing whereas the routers after the COR use inefficient flat routing.

The hierarchical routing structure for host mobility [66] is adapted for mobile networks [27] where ARs are considered as BGP routers, and all MRs are assumed CUIP-enabled routers and hosts. MNNs need not to be aware of CUIP as their packets are handled by MRs. Although this scheme has the advantages of its class, it suffers from the problem of generating frequent updates for routers. The problems will continue to increase with increasing distance of the mobile network from its home network. Moreover, additional signaling is required to discover a COR.

2.3.4.2 Wide-Area IP Network Mobility (WINMO)

The use of BGP for network mobility is proposed by Dul [28] where the AR, upon the attachment of a mobile network, initiates a BGP update announcing the prefix of the mobile network in the Internet. But this may result in large routing tables and large number of update messages because of movement of a large number of mobile networks. To limit the routing table size and number of updates, concepts of mobile prefixes and aggregation routers are introduced in WINMO proposed by Hu et al. [29].

A mobile prefix is used to serve all the mobile networks originated from a particular home network. Mobile prefixes are advertised only by a set of routers called aggregation routers that keeps track of prefixes assigned to the mobile networks. Other routers set the closest aggregation router as next hop for the mobile prefixes. Whenever a mobile network attaches to a new network, a BGP update is injected to announce the prefix of the mobile network. An aggregation router in the new network shares this update with all other aggregation routers. A packet sent to the mobile network reaches a router that forwards the packet to the closest aggregation router. Aggregation router forwards the packet to the appropriate aggregation router in the network to which the mobile network is attached.

For Intra route optimization, a mobile router obtains a care-of-address along with performing authentication with the AR. The care-of-address is used only to route packet efficiently within the network. A packet sent from the mobile network to a CN carries a key which is generated during the authentication by encrypting the prefix of the mobile network and the care-of-address. The key is managed among all the BGP routers and some other additional routers. Route optimization starts when the reply packet containing the key enters the network. A BGP router checks the validity of the key, and forwards the packet after changing the destination address to the care-of-address decrypted from the key.

Although this scheme involves smaller number of routers, and generate smaller number of routing updates, it requires changes in BGP. Moreover, involvement of a small group of routers for the mobility management means routes are not completely optimized. Also, if a CN is unable to recognize the key, the route may not be completely optimized.

2.3.4.3 Multiple P2P connected HA-based route optimization

Cuevas et al. [30] propose deploying multiple HAs that know each other's information (e.g., network, IP, etc.) using P2P [67]. A mobile network has a home HA; but can register with any HA to meet certain performance criteria such as a limit for round trip time. To find a closer HA, an MR sends a special BU to its home HA that responds with a list of HAs closer to current location of the mobile network in terms of the performance criteria. MR selects an HA, obtains a home address, and registers with the selected HA. After registration, the HA initiates a BGP update among routers within the network to install the mapping of the home address to the care-of-address of the MR. These routers tunnel/de-tunnel the packet to/from the mobile networks. The change of the HA takes place only when a mobile network moves out of the current network, and when an MR finds forwarding through current HA's network is not delivering required performance.

Unlike other schemes in this class, this scheme does not require BGP updates outside the network to which the mobile network is attached. But this requires communication, initiated outside, to take place through the home network resulting in unoptimized route. In addition, a large number of mobile networks that are moving frequently can trigger frequent BGP updates along with the problem of a large number of entries in the routing table.

2.3.5 Miscellaneous

This section includes route optimization schemes that do not fall into any of the previous classes described in Secs. 2.3.1 - 2.3.4. The techniques, used for route optimization in the schemes presented in this section, are different than the basic techniques used for route optimization in the classes presented in Secs. 2.3.1 - 2.3.4. A comparison of the schemes is presented in Table 2.5.

Table 2.5: A comparison of the route optimization schemes not belonging to any particular class.

Scheme	**Degree of route optimization**	**Signaling**	**Memory requirement**	**Location management**
ORC-based	Near optimal	Medium	High	Routers in the Internet
RBU-based	Optimal	High	Low	HA and CN
AODV-based	Near optimal	High	High	HA

2.3.5.1 Optimized Route Cache (ORC)-based

Wakikawa et al. [68] proposed an approach where the MR sends BUs to a router in the CN's network, and to the MR attached above (parent MR). The parent MR sends a BU, which maps the mobile network prefix of the MR underneath (child MR) to the parent MR's care-of-address, to the router performing location management. Therefore, packets sent to the mobile network are tunneled to the parent MR's care-of-address by the router that caches the optimized route for tunneling as long as the CN communicates with the MNN. Packets, destined to the child MR's network, are decapsulated and forwarded by the parent MR to the child MR. To route packets when no router in the CNs' network has the mapping, a BU is sent to a router which is in the home network of the mobile network. MRs uses a routing protocol to route packets inside mobile network.

This scheme incurs medium (instead of being high) amount of signaling because all MRs send BUs to routers (instead to all CNs) in CNs network and to parent MRs. Memory requirement is also high because of maintaining routing entries for all MNNs. A major disadvantage of ORC is that it optimizes route for only one level of nesting. Although route from CNs to MNNs is similar to that in hierarchical class for one level of nesting, it is different when the nesting level increases. In addition, unlike the schemes in hierarchical class, TLMR's home address or care-of-address is not conveyed to the nested MRs. Therefore, we have placed this scheme separately in this section.

2.3.5.2 Recursive BU (RBU)-based

Cho et al. [69] proposed a route optimization scheme where BUs, sent by MRs to CNs, are used to recursively process the binding table at the CNs to maintain a route to the TLMR. On reception of a BU having an HoA which is the same as any of the CoAs in the binding table, the CoA in the table is replaced by the received CoA. The CN will eventually have a mapping of the MR's prefix to the TLMR's CoA after receiving BUs from all MRs, thereby enabling sending of packets directly to the TLMR.

Packets are routed inside the mobile network by MRs which maintain routing tables, or by the TLMR broadcasting a route request for route discovery. Memory requirement for routing will be low when routes are discovered dynamically. Amount of signaling in this scheme is high because mobile nodes will send BUs to CNs. Also, it is not specified how the MRs will know about the CN. This scheme resembles schemes in hierarchical class with the difference of conveying the TLMR's HoA or CoA to nested MRs.

2.3.5.3 AODV-based

In AODV-based scheme, proposed by Phang et al. [31], the route between an HA and an MR is established using the AODV protocol. After obtaining a CoA from the attached (above) MR's prefix, the MR uses AODV route request messages to find a route towards its HA. During this route finding process, all MRs between the TLMR and the MR installs the routing entries for routing between the HA and the MR. After a route reply is received from the HA, MR sends a BU to the HA. Packets sent from a CN first reach the HA that tunnels the packet to the MR. Since the route from the HA to the MR is already established by AODV, the packet reaches the MR directly without any further tunneling.

The scheme appears to be very simple; yet, it requires all routers in the Internet, and HAs to support AODV resulting in difficulty of deployment. Moreover, the scheme involves one tunnel for communication along with overhead of burst of messages (i.e., high signaling) in the Internet during handoff due to broadcast of AODV messages. Although AODV is a protocol for Ad hoc networks, we do not include AODV-based scheme in prefix delegation class under Ad hoc-based scheme due to the following reason. The basic principle used in Ad hoc-based scheme is to obtain a CoA from the foreign network prefix contrasting the obtaining of the CoA from MR's prefix in the AODV-based scheme. The scheme also resembles the schemes in hierarchical class in terms of the route except the difference in establishing the route.

2.4 A comparison among the classes

Table 2.6 presents a comparative summary of the classes. Hierarchical schemes are easier to deploy, and also supports efficient intra mobile network communication at the cost of sacrificing complete optimization. The implication is that hierarchical schemes are suitable for mobile networks with large nesting levels and communications are mainly taking place within the mobile network. A mobile network with no nesting or one/two levels of nesting will not be benefitted from the use of hierarchical schemes that trade off allowing one/two tunnels with lower signaling and intra route optimization. In addition, most of the communications in the current Internet are client-server type where MNNs are expected to be clients, and servers are the CNs; this also lowers the significance of Intra route optimization provided by hierarchical schemes.

Table 2.6: A comparison among different classes.

Class	**Degree of route optimization**	**Intra route optimization**	**Signaling**	**Header overhead**	**Deploybility**	**Location transparency**
Prefix delegation	Optimal	No	High	Low	Difficult	No
Hierarchical	Near optimal	Yes	Low	Medium	Easy	Yes
Source routing	Optimal	No	High	High	Difficult	No
BGP-assisted	Near optimal	Yes	Low	Low	Difficult	Yes

Delegation-based and BGP-assisted schemes suite the client-server type communications that prevail in the Internet. Delegation approach is simple, do not introduce any additional overhead on the Internet routing, do not require any support from the infrastructure except the end hosts, and optimize route completely; but exerts additional load on the infrastructure due to high amount of signaling. On the other hand, BGP-assisted approach supports intra route optimization, and requires some support from the infrastructure; but routes may not be completely optimized and excessive mobility will exert load on the infrastructure. Source routing approach is not suitable for mobile networks having large nesting levels due to large header overheads that consume bandwidth which is scarce in wireless environment.

Therefore, we pick the prefix delegation-based schemes as the optimal class of schemes for the communication between the mobile network to the wired network, and consider for further evaluations. Our evaluations will reveal the effects of the differences in the schemes in this class, the effects of signaling on the performance, and limitations of the schemes if any.

2.5 Summary

In this chapter, we have provided an overview of NEMO and its problems followed by a summary of the existing evaluations of the schemes that has been proposed to solve the problems. Then we have provided a classification of the schemes and a comparison of the classes. Based on the comparison, we select a class of schemes for further evaluations.

Chapter 3

Evaluation of prefix delegation-based schemes

Prefix delegation-based schemes differ in the procedure to obtain new addresses from the delegated prefixes, and the degree of optimizing routes depending on the types of MNNs. This results in differences in the performance metrics such as handoff delay, end-to-end delay, signaling volume, and memory consumption. The significance of these differences depends on the number and types of mobile network nodes, number of CNs, nesting level and distance of the mobile network from its home agent. Therefore, it is not obvious which scheme will have the optimal performance, given a mobility scenario and a mobile network characteristics, such as numbers and types of nodes in the mobile network, nesting level and distance from the home network. This necessitates a comparative evaluation of the prefix delegation-based schemes. In this chapter, we present analytical models evaluating the performance, and validate the models using simulation.

To evaluate, we have selected four schemes introduced in Chapter 2 – SPD [42], MIRON [48], OPR [8] and Ad hoc [7]. As far as the differences mentioned in the previous paragraph are concerned, these four schemes are representatives of all prefix delegation-based schemes. The schemes have been evaluated in the literature using either simulation [7], modeling [8] or experimental testbed [48], making it harder to compare the schemes due to differences in evaluation methodology. To facilitate a fair comparison based on a common framework, we developed analytical models for the schemes, and validate the models using simulation.

The models developed here are based on the models developed by Lim et al. [32] (see Appendix C). The model developed to represent signaling (number of BUs) and memory consumption are similar to those developed by Lim et al. [32]. Unlike the models that were developed to represent delays [32], models in this chapter include queuing and contention delays and propagation delays. In addition, we develop models to find the end-to-end delay and time required to obtain CoAs. Moreover, the models of Lim et al. [32] capture only the general characteristics of the prefix delegation-based class and cannot capture the detail characteristics of the individual schemes. In contrast, the models in this chapter are developed for the selected individual schemes of the class and capture the detail characteristics of the schemes to show the differences quantitatively.

The rest of the chapter is organized as follows. Section 3.1 presents the performance critical differences among the selected schemes. Section 3.2 presents the analytical models, followed by numerical results in Sec. 3.3. A comparative analysis of the schemes is presented in Sec. 3.4. Finally, Sec. 3.5 summarizes the findings and types of further analysis to be performed.

3.1 Differences among prefix delegation schemes

Table 3.1 summarizes the differences that affect the performance of the schemes. SPD and Ad hoc do not optimize route for LFNs resulting in higher end-to-end delay due to packets traveling through the HA. End-to-end delay can be significant when mobile network is away from the HA and the nesting level is high. The procedure to obtain CoAs in MIRON may lead to higher handoff delay when the nesting level is high. In MIRON, route optimization for LFNs requires additional signaling whose amount is dependent on number of LFN-CN communicating pairs. The amount of memory required for OPR and its address translation procedure depends

Table 3.1: Differences among the prefix delegation-based schemes.

Schemes	**CoA obtained from**	**LFNs' route optimization**	**Additional BU**	**Memory requirement**	**Security**
SPD	MR	No	No	Depends on number of MRs	Good
MIRON	AR	Yes	Yes, depends on number of CN	Depends on number of MNNs and CNs	Good
OPR	MR	Yes	No	Depends on number of MNNs and CNs	Bad
Ad hoc	MR	No	No	Depends on number of MNNs	Good

on the number of MNN-CN communicating pairs. To quantify the differences for a comparative evaluation of the schemes, we develop analytical frameworks in Sec. 3.2.

3.2 Analytical models

This section presents the models [70] for the selected prefix delegation-based schemes. For convenience, models were developed based on assumptions that do not affect the results as far as the comparison of the schemes is concerned. Assumptions, notations, and the models are presented in this section.

3.2.1 Notations and assumptions

The models for the selected schemes are developed in Sec. 3.2.2. In this section, we introduce the notations that are common for all the schemes, and assumptions under which the models are developed.

3.2.1.1 Assumptions

We make the following assumptions to simplify the development of the models for the selected schemes.

- BUs that are sent periodically to refresh the entries are not considered. This quantity is deterministic and same for all the schemes.
- The number of CNs to which each MNN is communicating is uniform throughout the mobile network.
- Processing capacity of all the nodes are equal. This has little effect on the models because of negligible values of processing delays compared to link delays.
- Time required for packets' processing, such as encapsulation/ de-capsulation, address swapping, table searching, etc. are similar. This assumption does not affect the models because of negligible values of the processing delay when compared to link delays.
- We only consider the movement of the entire mobile network as a whole (assuming no relative movements among the MNNs) to derive the number of BUs and handoff delay. This type of movement of the mobile network is more likely in the real world.
- We assume a hierarchical (with parent-child relationship) and static topology for the nested mobile network, and this assumption was implicitly made in all previous works on NEMO. Link state and prefix can be disseminated efficiently by broadcasting router advertisements down the hierarchy.
- Handoff delay of an MNN can be expressed as the sum of the delay to obtain a CoA after the TLMR hands off to an AR and the location update delays. Since location update delays are equal for the schemes, differences among handoff delays are determined by the differences in the delay to obtain the CoA.

3.2.1.2 Notations common to all schemes

In this section, we describe the notations that have been used to describe the models in Sec. 3.2.2. Since congestion and contention is different during handoff because of the increase in the number of signaling packets, we differentiate them from the congestion and contention delays during times other than handoff. It is also to be noted that the delays due to congestion and contention can be different for the schemes due to differences in signaling. Therefore, we have used different variables to denote the delays due to congestion and contention. The notations used to develop the models are as follows:

T_f^c	=End-to-end delay from an LFN to a CN
T_a	=Delay to obtain the CoA
l	=Nesting Level of an MNN
s_p	=Size of the data packet
s_a	=Size of the router advertisement packet
τ_r	=Average router processing time to process a packet
τ_e	=Average router processing time to encapsulate or decapsulate a packet
p_d, p_w	=Propagation delays for wired and wireless links, respectively
b_w	=Average bandwidth available at a wireless node
$c_d^s(l), c_d^m(l), c_d^o(l), c_d^a(l)$	=Hop delay for data packets in wireless links as a function of l for SPD, MIRON, OPR and Ad hoc, respectively
$c_h^s(l), c_h^m(l), c_h^o(l), c_h^a(l)$	=Hop delay during handoff in wireless links as a function of l for SPD

	MIRON, OPR and Ad hoc, respectively
b_d	=Average bandwidth available at a wired node
m_r	=Memory required by TLMR
n_b, n_c	=Number of BUs and CNs, respectively
n_r, n_f, n_m, n_v	=Number of MRs, LFNs, LMNs and VMNs, respectively in the entire mobile network
$h_a^h, h_a^c, h_h^c, h_h^h$	=Avg. number of hops from an AR to an HA, an AR to a CN, an HA to a CN, and an HA to an HA, respectively
n'_r, n'_f, n'_m, n'_v	=number of MRs, LFNs, LMNs and VMNs, respectively attached to an MR

3.2.2 Models for the selected schemes

In this section, we develop analytical models for the four prefix delegation-based schemes to measure the following metrics:

- *Number of BUs:* Number of BUs is measured by the number of BUs generated from a mobile network during handoff. BUs consume bandwidth in mobile network and wired network, and its amount varies among the schemes depending on the number and types of MNNs.

- *End-to-end delay:* End-to-end delay measures the time taken by a packet sent from an MNN to reach a CN. It is a very crucial performance metric for real time applications, and affects the throughput of acknowledgment-based transport protocols. End-to-end delay is significantly different for the schemes when the mobile network is away from the HA, and when nesting level is high.

- *Memory overhead:* Memory overhead reflects the additional memory required at MRs for route optimization, and it is measured by the number of IPv6 addresses stored in the MR. Memory overhead can be an limiting factor in resource constrained environment, and depends on the number of MNNs.

- *Delay to obtain CoA:* This measures the delay to obtain the CoA during handoff. This delay adds to the handoff delay, varies among the schemes and is a function of the nesting level.

The models are presented in the next four subsections. As we present the models for SPD, we also state the differences of our models from those developed by Lim et al. [32] (Appendix C). However, similar differences apply for the models of MIRON, OPR and Ad hoc schemes.

3.2.2.1 SPD

- **Number of BUs:** Number of BUs is derived in the same way as it was derived by Lim et al. [32]. SPD provides route optimization for all MNNs (except LFNs) that send BUs to CNs and HAs. Thus the number of BUs for each MNN is $(n_c + 1)$, and the number of BUs sent by all MNNs is given by

$$n_b = (n_c + 1)\,(n_r + n_m + n_v)\,. \tag{3.1}$$

  (Number of BUs derived by Lim et al. is $n_b = (n_c + 1)\,(n_r + n_m)$ where they considered only VMNs and MRs. Appendix C contains the details of the models developed by Lim et al. [32].)

- **End-to-end Delay:** End-to-end delay was not modeled by Lim et al. [32], but is important to show the differences of the prefix delegation-based schemes.

End-to-end delay is derived in a similar way the latency of BUs was derived by Lim et al. [32] (please see Appendix C for the latency of BUs model developed by Lim et al.). Unlike the derivation in [32], we use contention and queuing delay in wireless links ($c_d^s(l)$), and propagation delays of links (p_d, p_w) in addition to other parameters. The derivation of the model is discussed next. Since no route optimization is provided for LFNs, packets sent by an LFN are tunneled through its MR's HA. Since other MRs above optimize their route, there is one tunnel only. Therefore, we use the number of hops from a CN to an HA, an HA to an LFN multiplied by sum of propagation delay, transmission delay, per hop delay and processing delay at each router to calculate end-to-end delay which is given by

$$\begin{aligned} T_f^c &= (l+1)\left(\frac{s}{b_w} + p_w + c_d^s(l) + \tau_r\right) \\ &+ \left(h_a^h + h_h^c\right)\left(\frac{s}{b_d} + p_d + \tau_r\right) + 2\,\tau_e. \end{aligned} \tag{3.2}$$

- **Memory overhead for TLMR:** Memory overhead is derived in the same way as it was derived by Lim et al. [32]. Unlike the memory requirement models developed by Lim et al. where the total memory requirement at all entities were derived, we derive the memory requirement only at the TLMR. Although the total memory requirement for all entities may actually be large, the amount of memory required at each entity may be small. However, the memory requirement at the TLMR may be large, as the TLMR is in charge of the entire mobile network.

  In SPD, a prefix is assigned to each attached MR resulting in an entry in the routing table that maps a prefix to the next hop MR. Therefore, memory required is

$$m_r = 2n_r'. \tag{3.3}$$

- **Delay to obtain CoA:** Delay to obtain CoA was not derived by Lim et al. [32]. We derive Delay to obtain CoA in a similar way we derive End-to-end delay. For the derivation, we consider the detail procedure of obtaining the CoA. The TLMR obtains a CoA and prefix from the prefix advertised by the AR. Obtained prefix is then advertised inside the mobile network. All MNNs (except LFNs), on reception of this advertisement, obtain CoAs from the prefix whereas MRs, like TLMR, obtain a prefix to advertise to its MNNs. The delay to propagate the prefix to an MR at a level is the sum of hop delays due to congestion and contention during handoff, processing delay, propagation delay, and transmission delay multiplied by the level:

$$T_a = l\left(\frac{s_a}{b_w} + p_w + c_h^s(l) + \tau_r\right). \tag{3.4}$$

3.2.2.2 MIRON

- **Number of BUs:** MNNs (except LFNs) send BUs to their respective HAs and CNs. In addition, MRs send BUs to the CNs that are communicating with LFNs. Therefore, the number of BUs for MIRON is given by

$$n_b = (n_c + 1)(n_v + n_m + n_r) + n_c n_f. \tag{3.5}$$

- **End-to-end Delay:** End-to-end delay includes propagation delay, transmission delay, per hop delay and processing delay at each router on the optimized route. Additional delay is incurred for LFNs due to MR replacing the source address by its CoA, and placing the LFN's source address in the extension header. End-to-end delay from LFN to CN is given by

$$\begin{aligned} T_f^c = {} & \tau_r(l + h_a^c + 1) + \left(\frac{s_p}{b_w} + p_w + c_d^m(l)\right)(l+1) \\ & + \left(\frac{s_p}{b_d} + p_d\right) h_a^c + \tau_{ad} \end{aligned} \tag{3.6}$$

where τ_{ad} is the average per packet processing time at an MR.

- **Memory overhead for TLMR:** An MR creates a host route entry for each MNN (except LFNs) under it to route packets inside the mobile network. An MR also keeps track of the CN-LFN (attached to MR) pairs. Thus, the memory overhead for MR in MIRON is computed as

$$m_r = 2 \times \left(n_v + n_m + n_r + n_c n'_f\right). \quad (3.7)$$

- **Delay to obtain CoA:** After obtaining the CoA, an MR starts PANA re-authentication phase (requires four messages) [49] to tell the attached MNNs to obtain CoAs. An MNN sends a DHCPv6 request to obtain a CoA from the foreign access network. The request is relayed by the MRs, on the path to the TLMR, towards the foreign network. The DHCPv6 reply, containing the CoA, reach the MNN along the same path. Therefore, the time to obtain a CoA for an MNN at any level is the sum of time required to obtain CoAs by all the MRs on the path to the TLMR, and the time for DHCPv6 request/reply messages exchange. Let, s_n = Size of PANA message, s_h^q = Size of DHCPv6 request message, and s_h^r = Size of DHCPv6 reply message. Then delay to obtain CoA is given by

$$\begin{aligned} T_a &= 4\left(\frac{s_n}{b_w} + p_w + c_h^m(l) + \tau_r\right) l + \left(\frac{s_h^q + s_h^r}{b_w} + 2p_w \right. \\ &\qquad \left. + 2c_h^m(l) + 2\tau_r\right) \sum_{i=1}^{i=l} (i+1). \end{aligned} \quad (3.8)$$

  Note that delay to obtain CoA in MIRON is quadratic in terms of level.

3.2.2.3 OPR

- **Number of BUs:** In OPR, only MRs obtain CoAs, and send BUs to their HAs. No BU is sent to the CN for route optimization. Thus the number of BUs becomes equal to the number of MRs in the mobile network:

$$n_b = n_r. \quad (3.9)$$

- **End-to-end delay:** OPR procedure to register the new translated address with the CN requires table searching at the MR, and binding cache searching at the CN for every packet. Also, the address of the packet is changed by the MR before forwarding it. We combine these three processing costs as OPR processing cost in our model. Therefore, end-to-end delay in OPR is the sum of OPR processing time and the end-to-end delay of MIRON (Eqn. (3.6)) as given below.

$$\begin{aligned} T_f^c &= \tau_r \left(l + h_a^c + 1\right) + \left(\frac{s_p}{b_w} + p_w + c_d^o(l)\right)(l+1) \\ &\quad + \left(\frac{s_p}{b_d} + p_d\right) h_a^c + \tau_{ad} + \tau_{OPR} \end{aligned} \tag{3.10}$$

  where, τ_{OPR} is the OPR processing time at the MR.

- **Memory overhead for TLMR:** In addition to routing entries like SPD, OPR scheme stores a table at each MR for the OPR procedure. For each CN-MNN pair attached to the MR, the table requires an entry containing original address, translated address and the flags. Hence, memory overhead for the MR in OPR scheme is given by

$$m_r = 2n'_r + 3n_c\left(n'_v + n'_m + n'_f + n'_r\right). \tag{3.11}$$

- **Delay to obtain CoA:** This delay is the same as that of SPD, and is given by Eqn. (3.4).

3.2.2.4 Ad hoc

- **Number of BUs:** Like SPD, Ad hoc scheme optimizes route for all MNNs except LFNs. Therefore, number of BUs can be found from Eqn. (3.1).

- **End-to-end Delay:** End-to-end delay for this scheme is equal to the end-to-end delay for SPD ignoring the the additional delay incurred at the start of packet delivery for route discovery using AODV [71], and given by Eqn. (3.2).

- **Memory overhead for TLMR:** When an MNN communicates with one or more CNs, the TLMR has to maintain one routing entry to forward packets for that MNN. Therefore, the memory overhead of this scheme is given by

$$m_r = n_v + n_m + n_r. \tag{3.12}$$

- **Delay to obtain CoA:** An MR obtains a CoA from the advertised prefix followed by the route discovery to the AR using AODV, and advertise the prefix inside its network. Therefore, the delay will be the sum of propagation delay of the prefix, and the route discovery delay. To calculate path discovery delay, we use the number of hops between the AR and an MR which is essentially the level of that MR. Therefore, the delay to obtain CoA is

$$T_a = l\left(\frac{s_a}{b_w} + \frac{s_r^q}{b_w} + \frac{s_r^r}{b_w} + 3\left(p_w + c_h^a(l) + \tau_r\right)\right), \tag{3.13}$$

  where, s_r^q = Size of AODV request message, and s_r^r = Size of AODV reply message.

The models developed in this section are used to compare the schemes using numerical results that are validated by simulation, and analyzed in Sec. 3.3.

3.3 Simulation and numerical results

In this section, we present numerical results obtained from the models developed in Sec. 3.2, and validate the results using ns-2 [72] simulation. Since ns-2 can not be used to validate memory overhead, and the number of BUs is deterministic, we only validate delay to obtain CoA and end-to-end delay. Delay to obtain CoA is the time difference between the time instant the TLMR obtains a CoA and the time instant the lowest level MR obtains a CoA. On the other hand, end-to-end delay was measured by the difference between the time instant of the CN receiving a packet and the time instant of the LFN sending the packet.

3.3.1 Simulation environment

Figure 3.1 shows the simulation topology for a nesting level of two. LFN1 is set as the constant bit rate data source over UDP and CN is the destination. The mobile network moves between ARs resulting in mobile network handoff. Since the results differ among the schemes only for LFNs, we use LFNs in our simulation. IEEE 802.11 is used for all wireless communications. The number of hops between HA_MR3 and the mobile network is varied by varying the number of routers between HA_MR3 and R. Values of parameters used in the simulation and models are summarized in Sec. 3.3.2.

3.3.2 Simulation and model parameters

Table 3.2 shows the values of parameters used in the models and simulation. Values for processing time, bandwidth and propagation delays are those used by Park et al. [8]. Since the delegated prefix translation processing time includes table searching, address changing and copying new address, we set delegated prefix translation processing time as three times of the processing time. Packet sizes for the schemes are taken from the corresponding schemes. Average hop delays (propagation, contention and congestion) in wireless links are obtained from simulation.

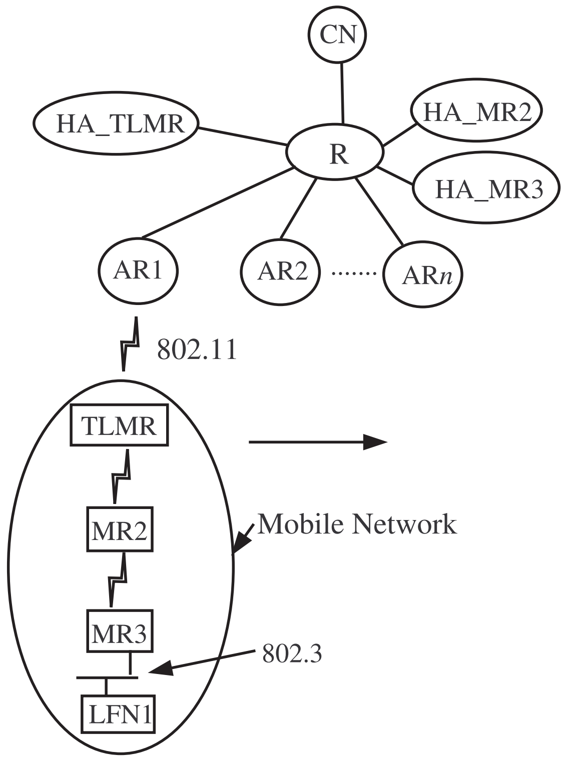

Figure 3.1: Topology used for simulation with two levels of nesting.

Table 3.2: Values of the parameters used in the models.

$s_p = 1000$ bytes	$s_a = 88$ bytes
$\tau_r = 10\mu$s	$s_n = 76$ bytes
$\tau_e = 10\mu$s	$s_h^q = 96$ bytes
$b_w = 10^7$ Mbps	$s_h^r = 184$ bytes
$s_r^q = 88$ bytes	$s_r^r = 84$ bytes
$\tau_{dpt} = 30\mu$s	$\tau_{ad} = 10\mu$s
$p_w = 30/(3 \times 10^8)$ s	$p_d = 1.8$ ms
$c_d^s(l), c_d^m(l), c_d^o(l), c_d^a(l)$ = Obtained from simulation and presented in Fig. 3.6	$c_h^s(l), c_h^m(l), c_h^o(l), c_h^a(l)$ = Obtained from simulation and rresented in Fig. 3.9

3.3.3 Results

3.3.3.1 Number of BUs

Figure 3.2 shows that the number of BUs for SPD, Ad hoc and MIRON increases linearly with the number of CNs. This is because route optimization requires BUs to be sent to each CN. The number of BUs for MIRON is higher than that for Ad hoc and SPD because MIRON optimizes routes for LFNs that requires sending additional BUs to the CN.

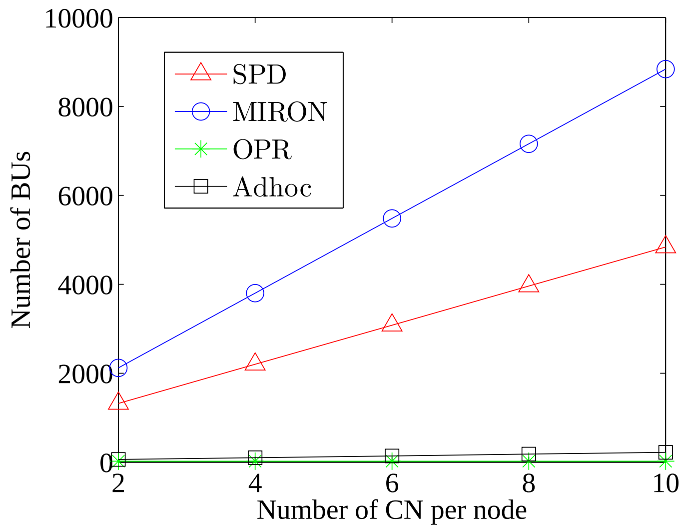

Figure 3.2: Number of BUs generated when TLMR moves, for $n_r = 20$, $n_m = 20$, $n_v = 400$ and $n_f = 400$. Values for SPD and Ad hoc are equal and superimposed. Values for OPR is small and superimposed on the axis.

Figure 3.3 shows that the number of BUs in MIRON increases linearly with the number of LFNs, as BUs are sent for each LFN. The number of BUs in SPD and Ad hoc are constant with respect to number of LFNs which require no BUs, as their routes are unoptimized.

As revealed by Figs. 3.2 and 3.3, the number of BUs in OPR is the lowest and constant because no BUs are sent to CNs.

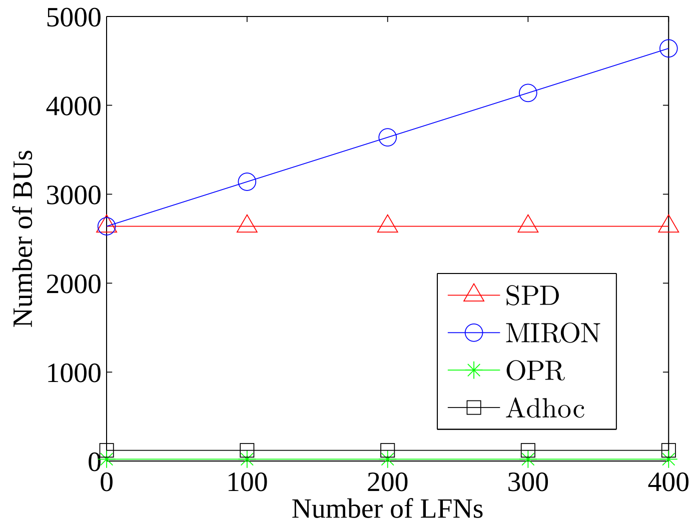

Figure 3.3: Number of BUs generated when the TLMR moves with $n_r = 20$, $n_m = 20$, $n_v = 400$ and $n_c = 5$. Values for SPD and Ad hoc are equal and hence, superimposed. Values for OPR is small, and superimposed with the axis.

3.3.3.2 Memory overhead for TLMR

Figure 3.4 shows the impact of the number of CNs on the memory overhead of the TLMR. The rate of increase for OPR is the highest as the TLMR tracks all CN-MNN (attached) communications. In MIRON, only CN-LFN (attached) communications are tracked. Memory overhead in SPD and Ad hoc is constant with respect to the number of CNs because no tracking of ongoing communications are required. Memory overhead for Ad hoc is a little higher than that of SPD due to memory used to maintain routing entries for all MRs in contrast to SPD's maintaining routing entry for attached MRs only (because of hierarchical prefix delegation).

3.3.3.3 End-to-end delay

Figure 3.5 shows the end-to-end delay as a function of level, where the end-to-end delay increases almost linearly except for the Ad hoc scheme where it tends to be non-linear at higher levels. The end-to-end delay depends on sum of hop delays. Since the hop delay for Ad hoc scheme increases linearly with level (see Fig. 3.6),

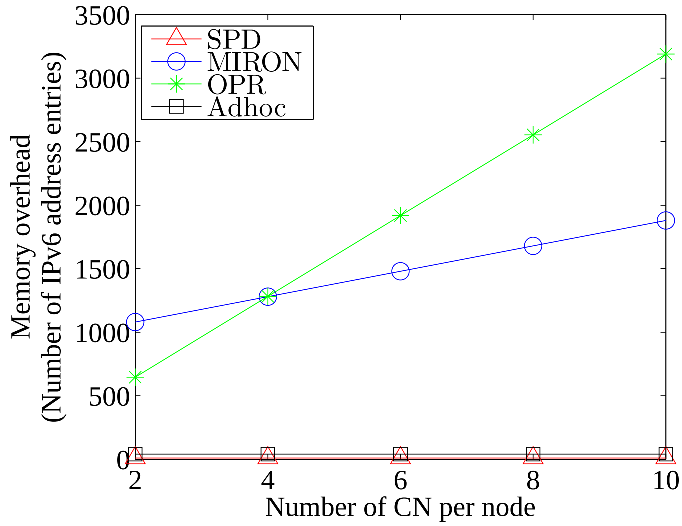

Figure 3.4: Memory overhead with increasing number of CNs.

the increase of the end-to-end delay is quadratic (see Eqn. (3.2)) as a function of the level (see Fig. 3.5). The end-to-end delay for the other schemes is linear as a function of level because increase of the hop delay with the level is insignificant.

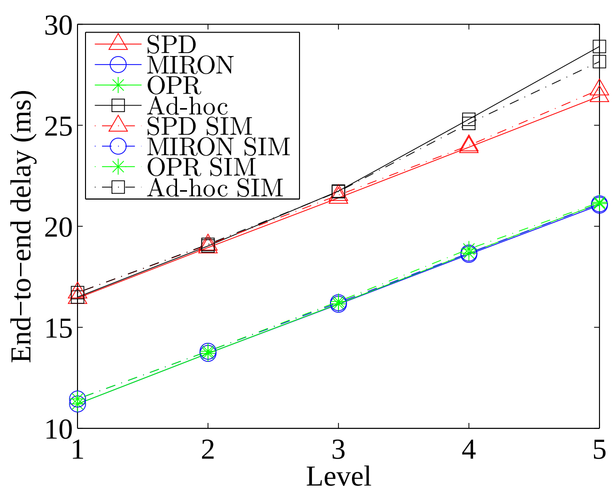

Figure 3.5: End-to-End delay between CN and LFN with increasing nesting level. Data for MIRON and OPR are very close and almost superimposed.

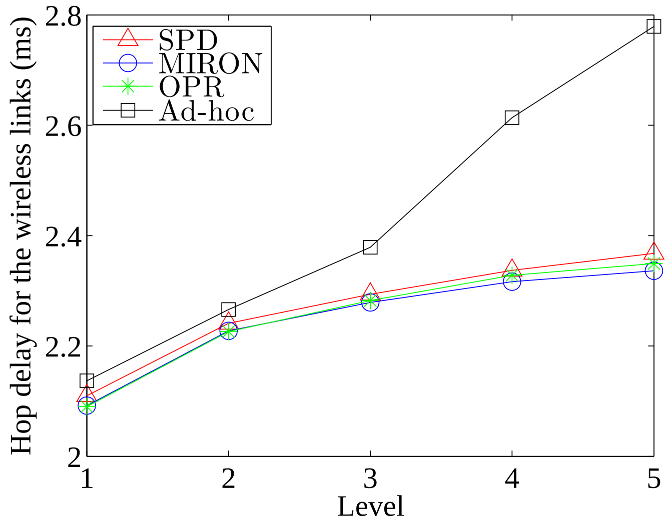

Figure 3.6: Hop delay in the wireless links for different schemes.

Hop delays for the schemes are presented in Fig. 3.6. The reason for higher rate of increase for Ad hoc scheme is the increased contention and congestion due to periodic signaling for updating routes.

End-to-end delay as a function of the number of hops between the mobile network and its HA is shown in Fig. 3.7. End-to-end delays in SPD and Ad hoc schemes increase with increase of the number of hops due to the increase in the route length as packets traverse through the HA. End-to-end delays for MIRON and OPR are independent of the number of hops because packets do not go through the HA.

3.3.3.4 Delay to obtain CoA

Figure 3.8 presents the delay for the lowest level MR to obtain a CoA. For MIRON, the rate of increase is higher than that for others. The higher rate results from the quadratic nature of the delay as a function of level (see Eqn. 3.8) compared to the linear nature (see Eqn. (3.4)) for other schemes. Delay to obtain a CoA in Ad hoc is higher than that of SPD and OPR due to the additional time required to find the route to the AR using AODV request/reply messages after the CoA is obtained. Delay to obtain a CoA in SPD and OPR is similar because of the similar procedure to obtain a CoA. At lower levels, delay in Ad hoc scheme is higher than the delay in MIRON due to two reasons. First, the lack of domination of the quadratic delay of MIRON at lower values of the level. Second, the hop delay of Ad hoc scheme is higher than that of MIRON.

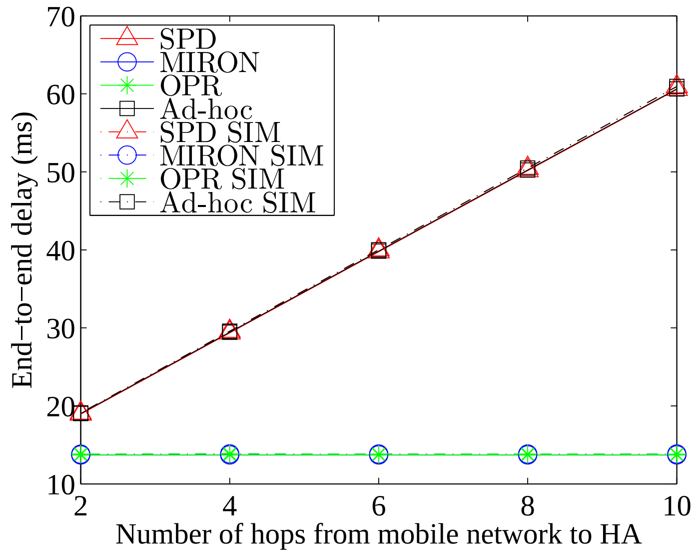

Figure 3.7: End-to-End delay between the CN and the LFN with increasing number of hops between the mobile network and the HA. Values for SPD and Ad hoc, and values for MIRON and OPR are very close and superimposed.

Figure 3.9 shows the hop delays for the schemes during handoff. Hop delays for SPD and OPR are the highest due to the following reason. Hop delay for the first packet (RA, PANA message, etc.), sent from the TLMR to the attached MNNs after the discovery of an AR, is much larger compared to hop delays for the rest of the signaling packets. Since we compute hop delays by taking mean of the hop delays for all signaling packets generated during handoff, it is small when number of signaling packets generated is large, and vice versa. Hop delays are higher in SPD and OPR than that of MIRON and Ad hoc because of smaller number of signaling packets. Hop delay in Ad hoc is higher than MIRON due to more congestion and contention during handoff because of the broadcast of signaling packets.

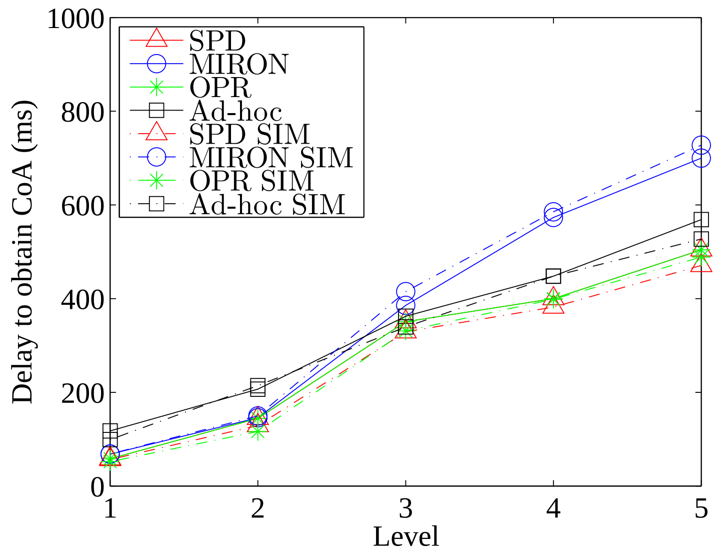

Figure 3.8: Delay to obtain CoA by the lowest level MR for the schemes.

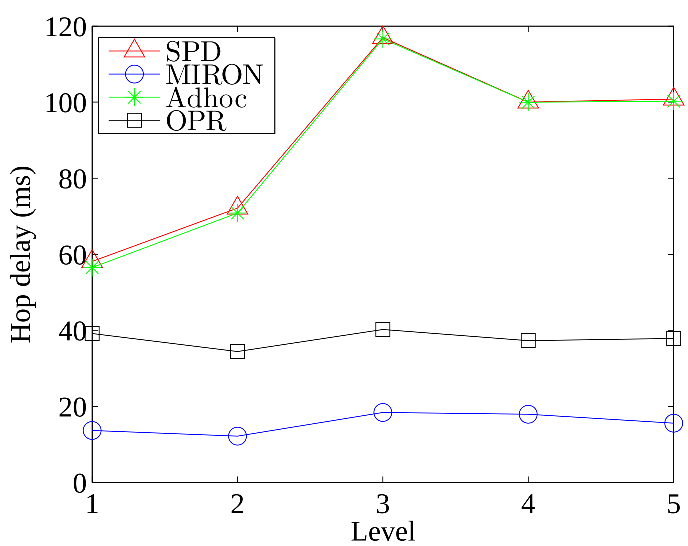

Figure 3.9: Hop delay for different schemes during handoff.

3.4 Comparative analysis of the schemes

Section 3.3.3 presents the metrics (discussed in Sec. 3.2.2) as a function of the level, number of hops between the mobile network and the HA, and number of CNs and LFNs. The metrics can affect the performance, and limit the application of the schemes depending on the handoff frequency (due to the generation of signaling and handoff delay during handoff), number and types of MNNs, distance of the mobile network from the HA, and the level. Among the parameters, handoff frequency is determined by the mobility scenario whereas others are the characteristics of the mobile network. Therefore, in this section, we perform a qualitative comparison of the schemes to discuss the suitability of the schemes under various scenario that affects the parameters and hence, the metrics. We also present a comparative summary of the schemes that shows the major advantages achieved by the schemes, and the cost for achieving the advantages.

3.4.1 Comparison based on mobility characteristics

Results for the delay to obtain the CoA and signaling of the schemes presented in Sec. 3.3.3 can be analyzed for comparison of the schemes under the following mobility scenario.

- Scenario 1 (high mobility): A mobile network, traveling at a high velocity, has low subnet residence time and high handoff frequency. Since OPR generates the minimum number of BUs with low handoff delay, it is the most preferable scheme for this scenario. SPD can be the next choice in high mobility scenario due to the handoff delay similar to OPR along with signaling lower than MIRON.
- Scenario 2 (low mobility): This can be a mobile network in a vehicle moving at a low velocity and thus, having a low handoff frequency. MIRON and OPR will provide better throughput than other schemes, as these schemes optimize

route for all MNNs. Effect of the large delay to obtain CoA on throughput will be low in MIRON because of fewer number of handoffs.

3.4.2 Comparison based on the number of LFNs

Packets to or from LFNs are tunneled in SPD and Ad hoc in lieu of the advantage of having low amount of signalling. Additional bandwidth required due to tunneling headers might be compensated by low bandwidth consumption of small amount of signaling. But tunneling also results in high end-to-end delay that affects performance of real-time traffic and acknowledgement-based transfer protocols. However, the end-to-end delay due to tunneling is small in SPD and Ad hoc when the mobile network is close to home agent. Also, the bandwidth consumption due to BUs will be high in MIRON when number of LFNs is high. Therefore, SPD and Ad hoc will be preferable to MIRON when a mobile network with a large number of LFNs is close to the home network.

3.4.3 Comparison based on memory requirement

In memory-constrained environments (e.g., mobile phones or sensors acting as routers which are also characterized by low processing capability), Ad hoc is the best choice due to low memory overhead. Although SPD has even lower memory overhead than Ad hoc, the requirement for each MR in SPD to be a prefix delegator seems to be infeasible due to additional processing overhead required by MRs to act as prefix delegators.

3.4.4 Comparison: principal advantages and associated cost

Table 3.3 summarizes principal advantages of the schemes and costs to achieve those advantages. Because of tractability reasons, not all the costs mentioned here were modeled in this paper. Comparison shows that there is no single scheme which

Table 3.3: Summary of major performance gains and costs.

Schemes	**Advantages**	**Cost of achieving the advantages**
SPD	Lowest memory overhead.	Prefix delegation involves security, authentication and accounting overheads, and is not possible through extended RA only. It might not be feasible to use all MRs as prefix delegators.
MIRON	1. Optimize route for LFNs 2. Use of standard protocols for CoA configuration.	Increased signaling, memory overhead and delay to obtain CoA.
OPR	Lowest signaling.	1. Memory overhead to track communication of all attached MNNs can increase significantly with the number of actively communicating CNs. 2. Additional processing per packet at MRs and CNs. 3. Inability to optimize route when packets are not sent from the mobile network to CN
Ad hoc	1. Lower expected memory overhead because routes are dynamically discovered at the start of communication (thus no permanent route entry is required.). 2. Resistant to creation of self loop when the MRs move relative to each other like MRs in a MANET	1. Large delay to obtain CoA due to route discovery after handoff. 2. Mobile networks have hierarchical architecture with stable connectivity but ad hoc protocols are designed for networks without hierarchy where connections are intermittent. Therefore, Ad hoc route optimization causes unnecessary signaling for mobile networks.

is the best for all mobility scenarios and mobile network characteristics. Results presented here provide a preliminary idea to select an appropriate prefix delegation-based scheme for a mobile network as a function of parameters, such as the mobility scenario, number and types of MNNs, number of CNs, distance of the mobile router from its HA, and the nesting level. OPR performs better than other schemes, but can optimize the route only when packets flow from the mobile network to the CN. MIRON can perform better at low speeds of the mobile network with a small number of LFNs and small nesting level. SPD and Ad hoc performs better when a mobile network with a large number of LFNs is close to its home network. However, Ad hoc's performance is the worst when the nesting level is more than two.

3.5 Summary

Results of the evaluation show the differences in the end-to-end delay and the delay to obtain the CoA that affects the handoff delay. The end-to-end delay and handoff delay will affect the throughput of the transport protocols, such as Transmission Control Protocol (TCP). Moreover, effects of the handoff delay will be magnified with the increase of the speed of the mobile network. Therefore, we intend to evaluate the throughput of TCP with the prefix delegation-based schemes while using the speed of the mobile network as another parameter.

Chapter 4

TCP-performance evaluation of prefix delegation-based schemes

TCP is the most widely-used data transport protocol in the Internet. The throughput of TCP is inversely proportional to the frequency and duration of interruptions that occur during data transfer. Such interruptions result from the handoffs of devices that move between networks. Also, the throughput of TCP is inversely proportional to the end-to-end delay.

Prefix delegation-based schemes differ in address obtaining procedures resulting in differences in the handoff delay (i.e., duration of the handoff). Effects of the differences in the handoff delay vary with the handoff-frequency that depends on the mobility speed. In addition, schemes differ in optimizing route for LFNs to yield different end-to-end delay when the mobile network is away from the home network (i.e., the end-to-end delay between the home agent and the mobile network increases). Therefore, throughput characteristics of TCP, when used with the schemes, are not obvious as a function of the speed of a mobile network and its distance from the home network.

In this chapter, our goal is to compare the throughput of the prefix delegation-based schemes as a function of the speed of a mobile network and its distance from the home network. However, we exclude the ad hoc-based scheme from the comparison due to its worst performance that was revealed by our evaluations presented in the previous chapter. The rest of the chapter is organized as follows. Section 4.1 presents the differences among the schemes from the perspective of TCP-throughput. Details of simulation and results are discussed in Secs. 4.2 and 4.3. Section 4.4 presents a comparative discussion based on the evaluations followed by concluding remarks in Sec. 4.5.

4.1 Differences among the schemes considering TCP-throughput

The schemes mainly differ in the procedure of obtaining CoAs and in optimizing route for LFNs. These differences, summarized in Table 4.1, affect the throughput of TCP. In MIRON, nodes take longer time to obtain CoAs resulting in a longer handoff delay. Both MIRON and OPR optimize route for LFNs that results in shorter end-to-end delay. Handoff delay causes packet loss and interruptions in communications resulting in degradation of the throughput of TCP. On the other hand, the throughput of TCP is inversely proportional to the end-to-end delay. Therefore, OPR will yield the highest throughput whereas throughput superiority of MIRON and SPD depends on the mobility speed and the end-to-end delay between the HA and the mobile network. The rate of change of throughput with the handoff frequency (determined by speed) is higher for MIRON due to large handoff delay. On the other hand, throughput of SPD changes with the end-to-end delay between the HA and the mobile network whereas that of MIRON is unaffected. Quantitative evaluations of the end-to-end delay, delay to obtain CoAs and throughput under various speed and at different end-to-end delays are presented in subsequent sections.

Table 4.1: Differences among the three schemes that make difference in throughput.

Schemes	**CoA obtaining time**	**LFNs' route optimization**
SPD	Small	No
MIRON	Large	Yes
OPR	Small	Yes

Table 4.2: Values of parameters used in simulations.

Parameter	**Value**
Simulation time	360s
Wired link BW	10Mbps
Wired link delay	1.8ms
Wireless (802.11b) link BW	11Mbps
Wireless range	250m
Ethernet (802.3) BW	10Mbps

4.2 Simulation environment

Figure 4.1 demonstrates a topology (with two levels of nesting) used in simulations. CN is the FTP source over TCP whereas LFN1 (connected to MR3 using Ethernet) is the TCP sink. The mobile network moves between ARs, placed in a horizontal line. Speed of movement is varied between 2 m/s to 30 m/s, and beyond 30 m/s handoffs are too frequent to be practical for real world scenario. Since throughput characteristics (with respect to speed) of all types of MNNs are the same across the schemes, and it is only LFNs' route optimization in which the schemes differ, we use only LFNs in our evaluations. IEEE 802.11 standard is used for wireless communications. To simulate the change of the end-to-end delay between HA_MR3 and the mobile network headed by MR3, we vary the link delay between HA_MR3 and Router. Values of parameters, used in simulations, are presented in Table 4.2.

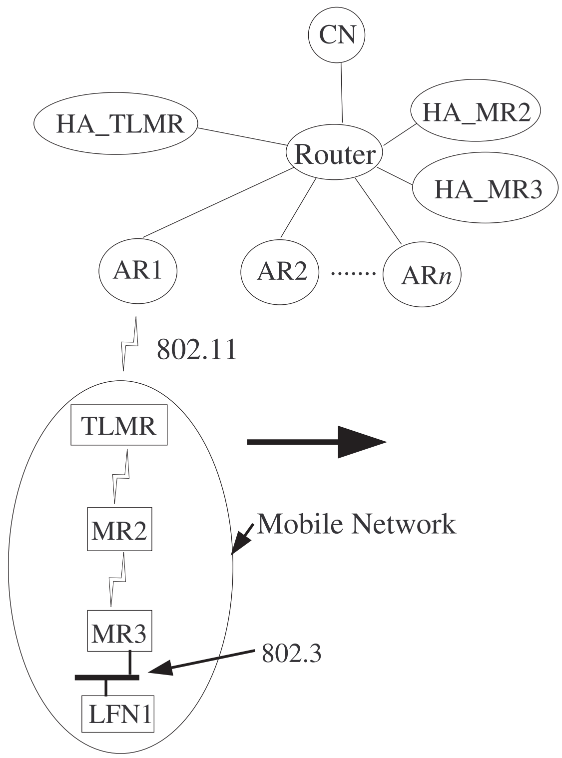

Figure 4.1: Topology used for simulation.

4.3 Results

We measure (at 95% confidence level) end-to-end delays, delays to obtain CoAs and throughput for the mobile network (residing at the lowest level) as a function of the speed and end-to-end delay between the HA and the mobile network. Throughput was measured by the total amount of data received at LFN1. Analysis of results is presented in the following sections.

4.3.1 End-to-end delay for LFNs

End-to-end delay, measured as the difference between the time of sending a packet by CN and the time of receiving the packet by LFN1, is shown in Fig. 4.2. With the increase of the delay between Router and the HA (in Fig. 4.1), the end-to-end delay

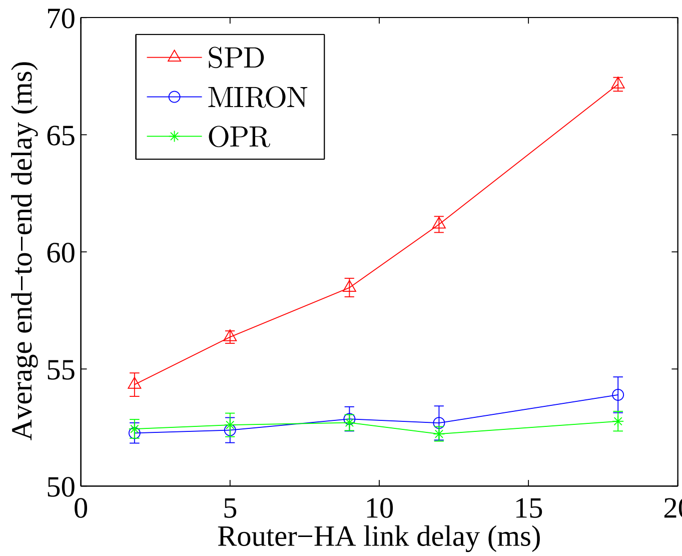

Figure 4.2: Average end-to-end delay between CN and LFN1.

for SPD increases whereas that of the other two schemes is not affected much. The reason for this is described next. Since SPD does not optimize route, packets reach LFN1 through the HA resulting in end-to-end delays of packets dependent on the end-to-end delay between Router and the HA. Other two schemes optimize route for LFNs and therefore, packets reach LFN1 without going through the HA. Thus, end-to-end delays for these two schemes are not affected (except small increase in MIRON) by the end-to-end delay between Router and the HA.

In MIRON, end-to-end delay shows small increase because after handoff, CN may send packets through the HA until a BU is received from the MR. Sending of the BU to CN might be delayed because at the reception of the first packet after handoff, the MR records the source (CN) of the packet for sending the BU (but does not send immediately) which is sent when the next period to send BUs comes. During this delay, packets are sent through the HA resulting in the increase of the end-to-end delay in MIRON. In OPR, there is no such delay, as the MR sends the CoA to CN as soon as the first packet is received after handoff.

4.3.2 Delay to obtain CoA vs. level of nesting

We measure the delay to obtain CoAs for a nested MR (such as MR3 in Fig. 4.1) by measuring the difference between the time instant when TLMR obtains the CoA and the time instant when the nested MR obtains the CoA. This delay which affects the handoff delay is presented in Fig. 4.3.

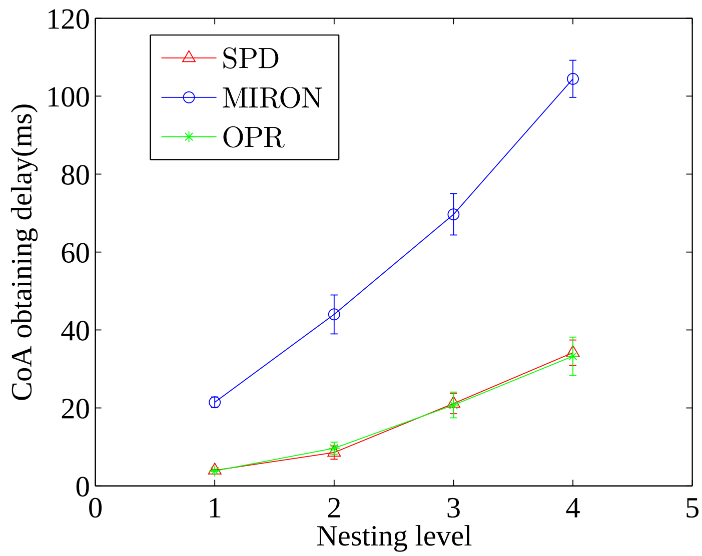

Figure 4.3: CoA obtaining delay of the MR at the lowest level after TLMR hands off.

For SPD and OPR, delays to obtain CoAs are similar where each MR obtains a CoA from the advertised prefix, and immediately advertises a part of the prefix to the attached MRs. All the MRs down the level does the same resulting in a delay which is linearly proportional to the level.

For MIRON, the delay is much larger than that of SPD and OPR. An MR, after obtaining a CoA, requests the attached MRs to obtain CoAs. Attached MRs use DHCPv6 request/reply messages that travels between the MRs and TLMR's foreign network. Therefore, the delay to exchange the messages is proportional to the level of the MRs. Moreover, an MR can start the procedure to obtain a CoA after all the

MRs above has obtained CoAs. Therefore, the delay to obtain a CoA is the sum of delay for all MRs above plus the delay resulting from the exchange of the messages, and hence quadratic in terms of the level.

4.3.3 Change of throughput with speed of mobile network

Figure 4.4 shows the change of throughput with the speed of the mobile network. Throughput falls with increasing speed due to increased handoff frequency. Since handoff delay in SPD and OPR is lower due to a small CoA obtaining time (see Fig. 4.3), and OPR optimizes route for LFNs, throughput for OPR is the highest. Despite having a similar handoff delay, throughput of SPD is lower than OPR's throughput due to higher end-to-end delay (see Fig. 4.2).

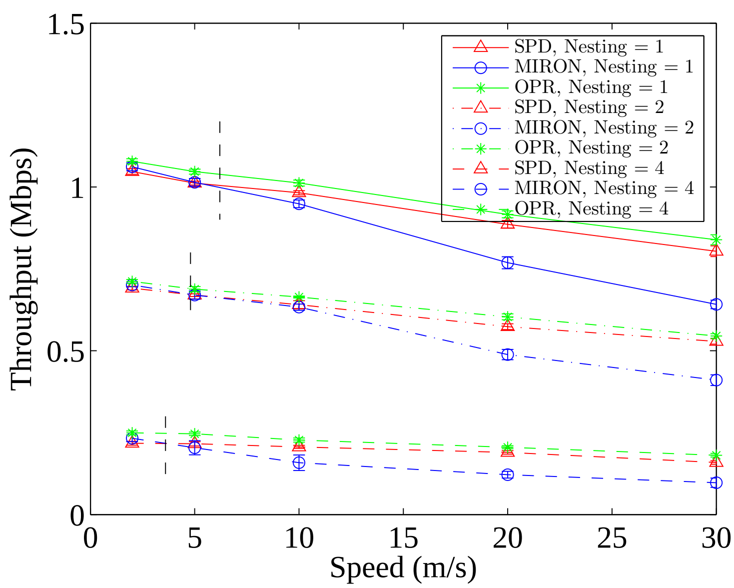

Figure 4.4: Throughput of different schemes for different nesting level. Vertical dashed lines indicates the speed at which throughput of MIRON falls below that of SPD.

Although LFNs' route is optimized, the rate of throughput fall with increasing speed is higher for MIRON because of large handoff delay due to the larger delay to obtain a CoA. Therefore, throughput in MIRON, higher at low speed, falls below the throughput in SPD at higher speed because the rate of throughput fall due to the handoff delay dominates the rate of throughput fall due to the end-to-end delay at higher handoff frequency.

4.3.4 Effect of level on fall of throughput

Another characteristic, observed from Figure 4.4, is that the fall of MIRON's throughput below SPD's throughput occurs at lower speed for higher level. The reason for this is the rate of increase of the delay to obtain a CoA (and hence, the rate of increase of handoff delay) with increase of level is higher for MIRON (see Fig. 4.3). Therefore, at a higher level, the rate of the loss of throughput (due to handoff) of MIRON is higher than other schemes resulting in the fall of MIRON's throughput (below SPD's throughput) at lower speeds.

4.3.5 Throughput vs. end-to-end delay between HA and mobile network

Figure 4.5 shows the effect of end-to-end delay (between the HA and the mobile network) on throughput at two different speeds of the mobile network. Since SPD does not optimize LFNs' route, packets go through the HA. Therefore, increasing the end-to-end delay between the HA and the mobile network results in decreasing throughput because TCP throughput is inversely proportional to the delay. End-to-end delays for the schemes are shown in Fig. 4.2. Throughput of MIRON and OPR is not affected much by virtue of LFNs' route optimization that avoids packets' going through the HA. Although MIRON's throughput is lower than SPD's throughput at higher speed, SPD's throughput falls below MIRON's throughput at higher end-to-end delay between the HA and the mobile network.

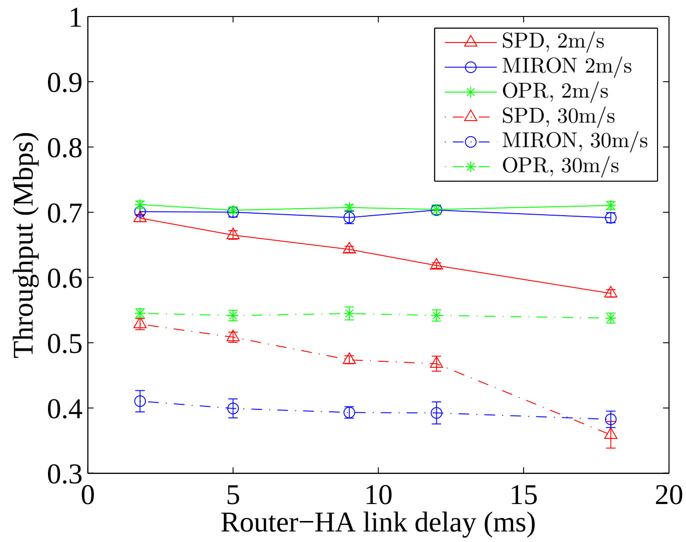

Figure 4.5: Throughput of different schemes at two different speed.

Based on the results presented in Sec. 4.3, we present a comparative discussion of the schemes in Sec. 4.4.

4.4 Comparative discussion

From the results presented in Sec. 4.3, OPR appears to be the best. However, OPR needs packets to flow from the mobile network to the CN for route optimization. If packets are not flowing from the mobile network to the CN, OPR performs like SPD. MIRON always performs better than SPD at lower speed. Even at higher speed, MIRON performs better when mobile network is away from home network in terms of the delay. In addition, MIRON uses a feasible procedure fto obtain CoAs. The CoA obtaining procedure for both OPR and SPD by prefix delegation through router advertisements might not be easily applicable due to accounting and security reasons. Overall, OPR could be the best scheme if route optimization could be enabled when packets are not flowing out of the mobile network. MIRON is a good choice at low speed whereas SPD can be a good choice at high speed when mobile network stays close to the home network.

4.5 Summary

In this chapter, we evaluated the performance of prefix delegation-based schemes for NEMO route optimization. We simulated three schemes, namely MIRON, SPD and OPR, and measured the end-to-end delay, delay to obtain CoA and throughput as a function of the speed of the mobile network and its distance from the home network. Results show that OPR performs the best although OPR's performance is limited due to the inability to optimize routes when packets do not flow from the mobile network to the CN. MIRON, having a feasible solution for obtaining the CoA, performs better at low speeds and nesting levels. On the other hand, SPD can perform better at high speeds when the distance between the home network and the mobile network is small.

Chapter 5

Effect of increasing the number of hosts on TCP-performance

The differences in the route optimization techniques of the prefix delegation-based schemes result in different amount of signaling traffic during handoff depending on the number of LFNs. Since signaling traffic competes for bandwidth with data traffic, throughput of the schemes might vary as a function of the number of LFNs. In this chapter, the throughput of the schemes is evaluated as a function of the number of LFNs. Section 5.1 describes anticipated performance of the schemes. Details of simulations and results are presented in Secs. 5.2 and 5.3. Section 5.4 summarizes the findings from the results.

5.1 Anticipated Performance

MIRON and OPR optimize routes for LFNs resulting in a smaller Round Trip Time (RTT) when compared to SPD that does not optimize routes for LFNs. Differences in the RTT increase as the mobile network moves away from the home network. Since TCP throughput is inversely proportional to the RTT, throughput in SPD will decrease when the mobile network moves away from the home network. Route optimization for LFNs requires more signaling in MIRON than in the other two schemes. The amount of signaling increases with the number of LFNs and the

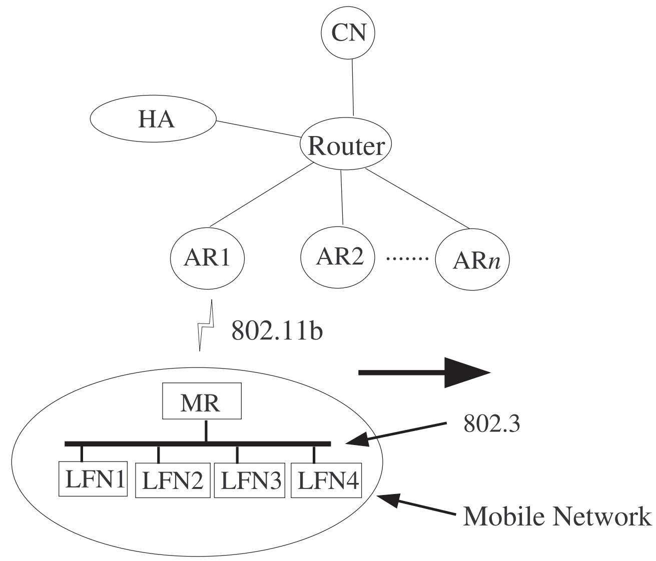

Figure 5.1: Topology used for simulation.

speed of the mobile network. Since signaling packets compete with data packets for bandwidth, throughput in MIRON might decrease when the number of LFNs and the speed of the mobile network increases.

5.2 Simulation Environment

Figure 5.1 demonstrates a topology used in simulations. FTP sources over TCPs are attached to the CN, whereas LFNs (connected to the MR using Ethernet) are the TCP sinks. The mobile network moves between ARs, placed in a horizontal line. Thus, the frequency of handoff is proportional to the speed of the mobile network. As we only intend to observe the effects of the frequency of handoffs, we do not use any particular mobility model to generate the movement of the mobile network. Since the schemes differ in LFNs' route optimization, we use only LFNs in our evaluations. IEEE 802.11b is used for wireless communications. To simulate the change of the distance between the HA and the mobile network, we vary the

Table 5.1: Values of the parameters used in the simulation.

Parameter	**Value**
Simulation time	200s
Wired link bandwidth	100Mbps
Wired link delay	10.0ms
Wireless (802.11b) link bandwidth	11Mbps
Wireless range	250m
Ethernet (802.3) bandwidth	100Mbps
Queue size	50packets
Interval of sending router advertisements	3s
Interval of sending BUs	10s
Lifetime of HA's binding entry, and CN's binding entry in MIRON	12s
Lifetime of CN's binding entry in OPR	1s

Router-HA link delay. Values of parameters used in the simulation are presented in Table 5.1.

5.3 Results

We measure (at 95% confidence level) throughput, RTT, packet drops and handoff latency as a function of the number of LFNs with the speed of the mobile network and its distance from the HA as the parameters. Results are presented in the following subsections.

5.3.1 Aggregated throughput

The aggregated throughput is measured by the total amount of data (TCP packets) received at all LFNs. Figure 5.2 shows the aggregated throughput obtained using the schemes for two different speeds of the mobile network when the Router-HA link delay is 10ms. Although the throughput in SPD is much smaller than that in OPR and MIRON when the number of LFNs is one (1), the throughput obtained using the

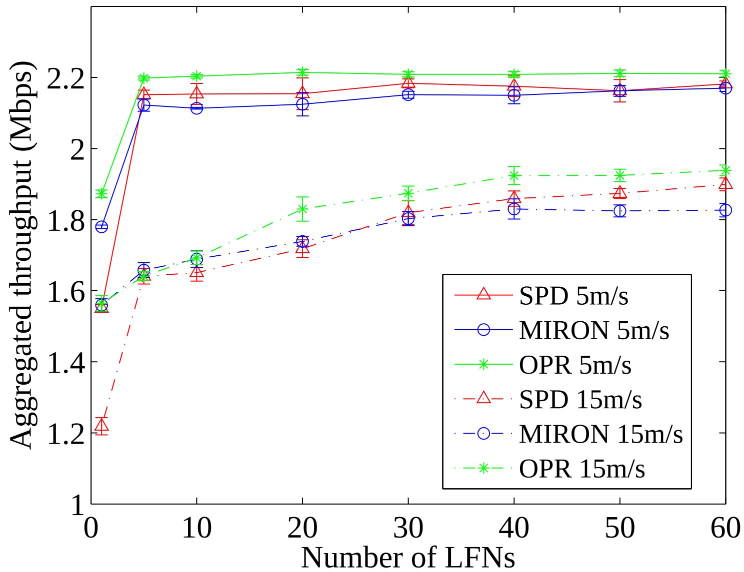

Figure 5.2: Aggregated throughput for two different speeds with Router-HA link delay at 10ms.

schemes is similar when the number of LFNs increases. One of the reasons for the similarity is the increase of the RTT (see Fig. 5.4 and Table 5.2) with the increase of the number of LFNs. Since TCP throughput is inversely proportional to the RTT, the ratio of RTTs in the schemes determines the ratio of throughput. When the number of LFNs is one (1), the ratio of RTTs in SPD, OPR and MIRON is 70:50:50 (approximate). As the number of LFNs increases, the ratio of the RTT in SPD to the RTTs in the other two schemes decreases because the difference in the RTT does not increase much compared to the value of the RTT. For example, when the number of LFNs is 10, the ratio of the RTTs becomes 250:225:230 (approximate). Therefore, the ratio of the throughput follows from the ratio of the RTTs.

Throughput in MIRON is little lower than that in SPD, particularly at high speed and when the number of LFNs is large despite the RTT in MIRON is lower. The reason for the lower throughput in MIRON is the lower sending rate due to the higher drop of TCP acknowledgement (ACK) packets. Also, the throughput in

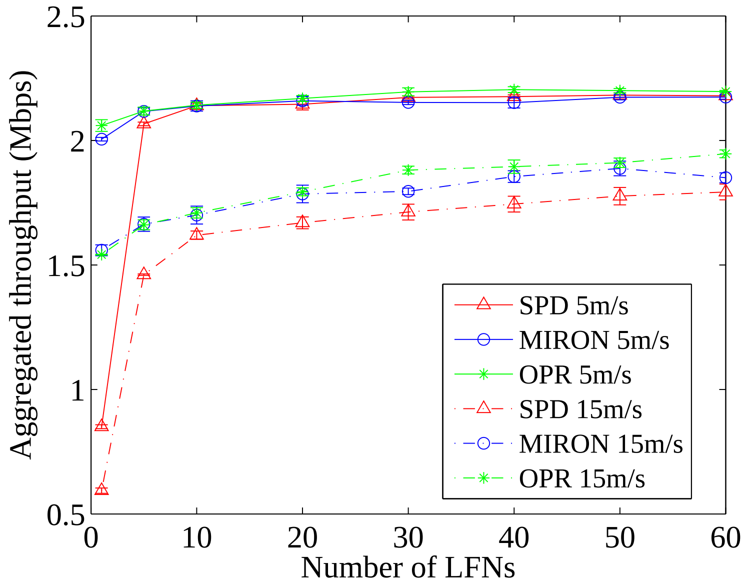

Figure 5.3: Aggregated throughput for two different speeds with Router-HA link delay at 40ms.

MIRON becomes lower than that in SPD when the number of LFNs is large due to the higher data packet drop in MIRON. Drop of packets is discussed in Sec. 5.3.3.

Figure 5.3 shows the aggregated throughput when the Router-HA link delay is 40ms. When the number of LFNs is increased, the difference of the throughput in SPD with that in OPR and MIRON decreases like it decreases for the Router-HA link delay of 10ms case. However, the throughput loss in SPD caused by the unoptimized route in the 40ms case is more than that in the 10ms case due to the increase of the ratio of the RTT in SPD to that in the other two schemes.

5.3.2 RTT

Solid lines in Fig. 5.4 show the RTT for the 15m/s case when the Router-HA link delay is 10ms. As expected, the RTT in SPD is higher than that in OPR and MIRON due to the unoptimized route. However, the difference in the RTTs among

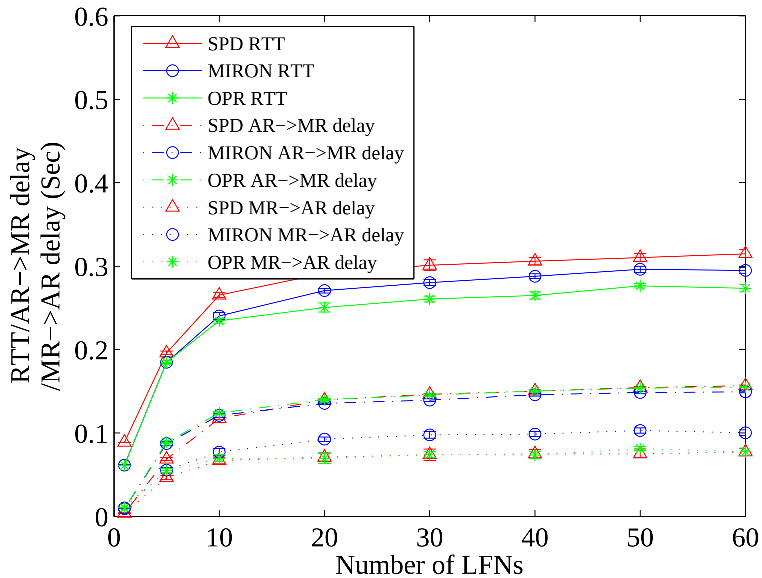

Figure 5.4: RTT, AR to MR and MR to AR delays for 15m/s speed of the mobile network when the Router-HA link delay is 10ms.

the schemes does not increase much when compared to the increase of the RTT. The higher RTT in SPD is due to the Router-MR link delay which is a constant. On the other hand, the increase of RTTs in all schemes is mainly due to the increase of the queuing and contention delays at the wireless link between an AR and the MR. Dash-dot patterned lines in Fig. 5.4 show that the average delay for TCP-packets at an AR is similar, and increases with the increase of the number of LFNs. Dotted lines in Fig. 5.4 show the average delay for ACK-packets at the MR. Table 5.2 summarizes the the delays at the wireless links and the RTTs for three different values of the number of LFNs.

The RTT in MIRON is higher than that in OPR, although both the schemes optimize routes for LFNs. One of the reasons for the higher RTT in MIRON is the higher hop delay from the MR to an AR due to the higher number of packets enqueued at the MR. The total number of packets enqueued and average queue size at the MR are shown in Fig. 5.5 and 5.6, respectively. The number of packets

Table 5.2: Delays at the AR-MR wireless links and the RTT in the three schemes for three values of the number of LFNs when the speed of the mobile network is 15m/s and the Router-HA link delay is 10ms.

Number of LFNs	**Scheme**	**AR to MR delay (ms)**	**MR to AR delay (ms)**	**RTT (ms)**
1	SPD	3.85	4.14	88.86
	OPR	9.67	10.65	61.04
	MIRON	9.34	10.14	61.47
20	SPD	139.59	70.92	291.39
	OPR	139.92	69.59	250.36
	MIRON	135.48	92.55	270.62
60	SPD	156.72	77.15	314.76
	OPR	154.74	77.46	273.37
	MIRON	149.38	100.25	294.78

enqueued and average queue size at the MR are higher in MIRON than that in SPD and OPR due to the higher number of BUs required for LFNs' route optimization. And, the number of BUs in MIRON increases with the increase of the number of LFNs (see Fig. 5.5).

Another reason for the higher RTT in MIRON than that in OPR is the number of packets that travel through the HA i.e., through the unoptimized route (see Fig. 5.7). The RTT for these packets are higher than those traveling through the optimized route. Usually, in MIRON and OPR, packets are expected to bypass the HA as they travel through the optimized route. However, when the binding entry in the CN expires, packets travel through the HA in a similar way they do in SPD. Expiration of binding entries occurs if it is not refreshed within its lifetime. In MIRON, BUs for LFNs' route optimization are sent to the CN in bursts. Therefore, some BUs may get dropped at the queue causing corresponding binding entries to expire. Such drop and expiration events increase with the increase of the number of LFNs resulting in the increase of the number of packets traveling through the HA. Determination of the rate of sending BUs and the lifetime of binding entries for the minimization of

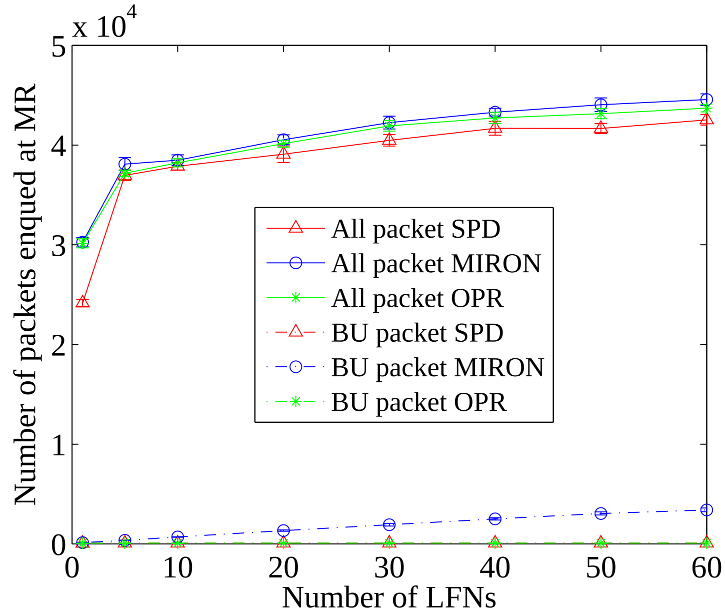

Figure 5.5: Number of packets enqueued at the MR for 15m/s speed case when the Router-HA link delay is 10ms.

the number and duration of expirations requires some kind of tradeoff, and is out of the scope of this paper. In OPR, binding entries are refreshed using the information sent in the OPR header of ACK packets which are usually received at the CN at a higher rate than the rate of sending BUs in MIRON. Therefore, the number and duration of expirations in OPR are smaller than that in MIRON. Hence, in OPR, the number of packets traveling through the HA is small.

Figure 5.8 shows the RTT for the case when the Router-HA link delay is 40ms. The characteristics of the RTT (as a function of the number of LFNs) in this case is similar to that when the Router-HA link delay is 10ms. However, the increase of the RTT from the RTT for the Router-HA link delay of 10ms case differs among the schemes. As expected, the RTT in SPD has increased more than that in the other two schemes due to the use of unoptimized route. Since more packets in MIRON than in OPR traverse through the HA (as explained in the previous paragraph), the

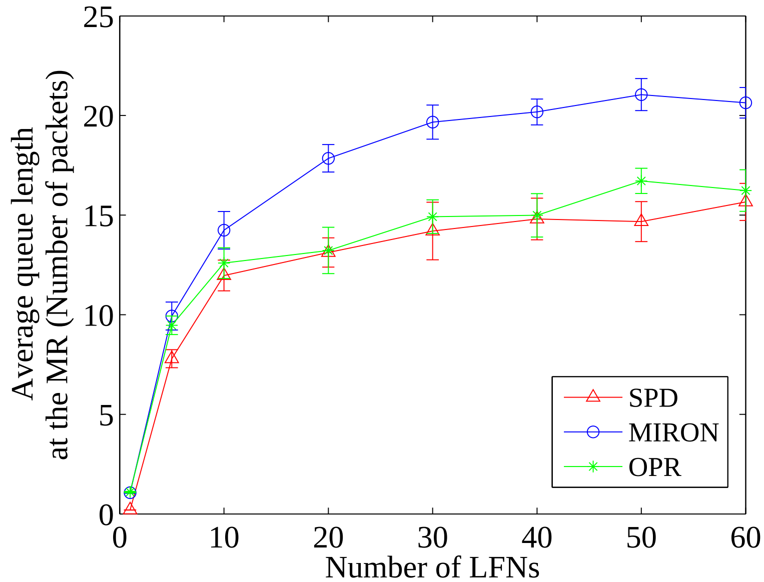

Figure 5.6: Average queue length at the MR for 15m/s case when the Router-HA link delay is 10ms.

RTT in MIRON has increased more than that of OPR. Consequently, the increase of the RTT in OPR is the lowest.

5.3.3 Drop of TCP packets due to expired or dirty binding entries

For 15m/s speed of the mobile network, and 10ms Router-HA link delay case, Fig. 5.9 shows the drop of TCP packets due to expired or dirty binding entries. The drop is measured as the percentage of the packets received at the AR. An expired binding entry results in drops at the HA. Since the destination MNN for packets is not at home, and no binding entry exists for the destination address, the HA is unable to determine the next hop for packets and drops them. Dirty binding entries cause packets to be sent using the old CoA, and therefore, to the old AR that fails to forward the packet because the mobile network has already left this AR's network.

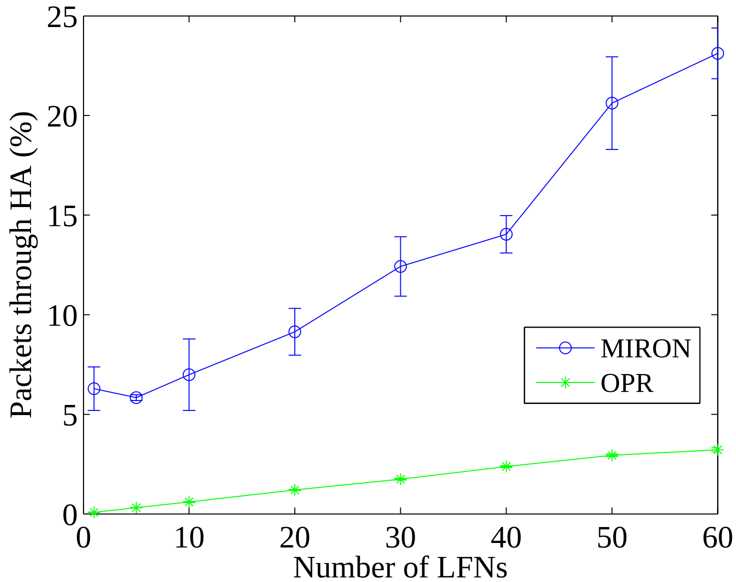

Figure 5.7: Percentage of packets that travel through the HA for 15m/s speed case when the Router-HA link delay is 10ms. In SPD all (100%) packets travel through the HA.

Binding entries become dirty when the binding entries at HAs and CNs are yet to be updated after the handoff. When the number of LFNs is large (more than 30), the MIRON's drop shown in Fig. 5.9 is higher when compared to the drops for the other two schemes. The higher drop occurs because the CN in MIRON sends a higher number of packets to the old AR.

The number of packets sent to the old AR is measured as the percentage of the total number of packets sent from the CN, and is shown in Fig. 5.10. In SPD, packets sent by the CN reach the HA that might tunnel those packets to the old AR if binding entries are dirty. Since the BU to the HA is the one to be sent first after the handoff and the frequency of sending the first few BUs is high [73], the probability for the HA to receive a BU early after the handoff is high. Therefore, binding entries at the HA are updated early after the handoff and hence, the small number of packets sent to the old AR in SPD.

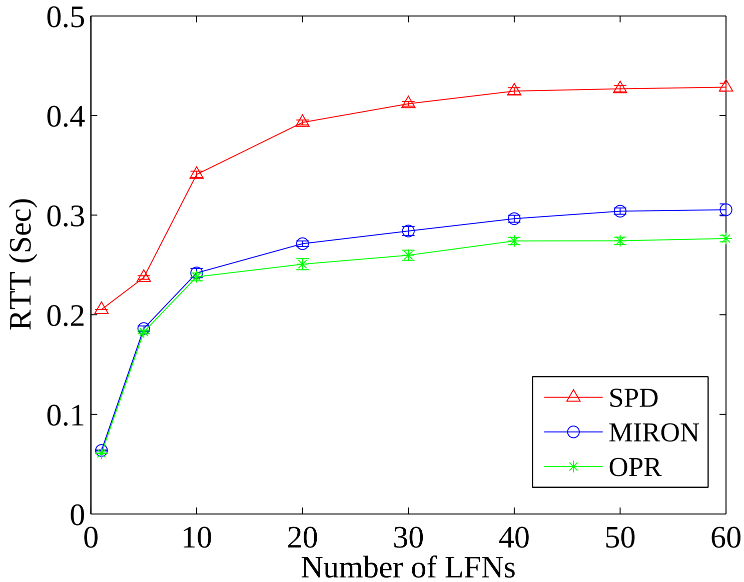

Figure 5.8: RTTs for 15m/s speed of the mobile network when the Router-HA link delay is 40ms.

In MIRON, the CN may send packets to LFNs using old CoAs found in the dirty binding entries. However, when the number of LFNs is large, the chance of receiving a BU very late after the handoff is higher than the chances of that at the HA due to two reasons. First, BUs sent to the CN on behalf of LFNs may get dropped at the queue because BUs are sent at a burst. Particularly, those BUs that are sent later in the sequence have higher chances of getting dropped at the queue. Second, the list of CNs, to which an MR send BUs, may expire due to no packet reception from CNs during the handoff. In this case, the MR cannot not send any BU to the CN after handoff until a packet is received through the HA. Thus, CN's binding entries become dirty, and it continues to send packets to the the old AR as long as binding entries are not expired or BUs are received. Using a small lifetime for binding entries, the chances of sending packets to the old AR, and the duration of the sending can be minimized at the cost of sending BUs at a high rate. In OPR, a small lifetime can be used because binding entries are updated using information

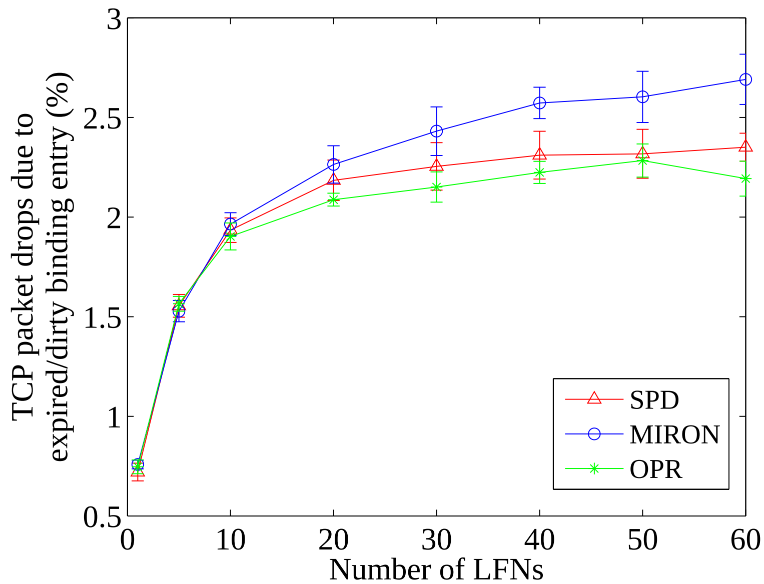

Figure 5.9: For 15m/s speed and Router-HA link delay of 10ms, TCP-packet drops at the AR due to expired or dirty binding entry.

carried in ACK packets that are received at the CN at a rate as high as the sending rate of TCP packets. Therefore, in OPR, the number of packets sent to the old AR is small.

Characteristics of the drop of TCP packets at the AR's queue is similar for all schemes. The characteristics of drops at the queue when the number of TCP flows increases can be found in [74].

5.3.4 Drop of ACK packets at MR's queue

Figure 5.11 shows the drop of ACK packets at MR's queue for 15m/s case when the Router-HA link delay is 10ms. ACK drop was measured as the percentage of incoming packets at the MR. The drop is higher in MIRON than in the other two schemes because of higher signaling in MIRON. Also, the drop in MIRON increases at a higher rate than that in SPD and OPR due to the increase of signaling with

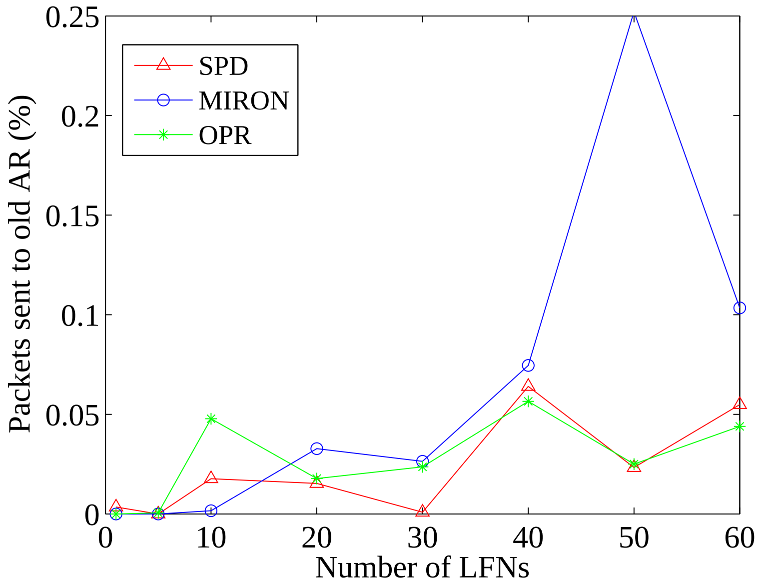

Figure 5.10: Packets sent to the old AR for 15m/s case when the Router-HA link delay is 10ms.

the increase of the number of LFNs. At high speed, drops in MIRON are larger than that in other two schemes due to increase of signaling resulting from increased handoff frequency.

5.3.5 Handoff latency from MR's viewpoint

Figure 5.12 shows the handoff latency from MR's viewpoint for 15m/s and Router-HA link delay of 10ms case. Handoff latency is measured as follows:

Handoff latency = the reception-time of the first router advertisement at the MR from the new AR - the reception-time of the last TCP packet at the MR through the old AR.

Thus, the measured handoff latency can be expressed as the sum of the following components:

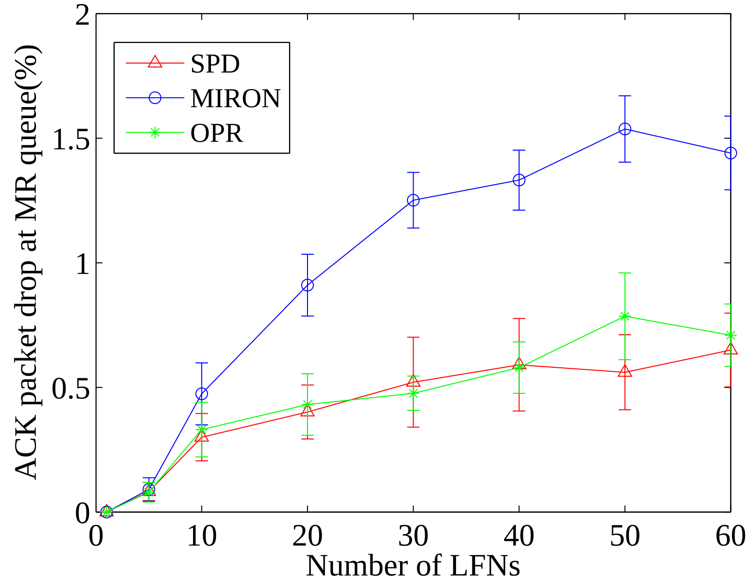

Figure 5.11: Drop of ACK packets at MR's queue for 15m/s case when the Router-HA link delay is 10ms.

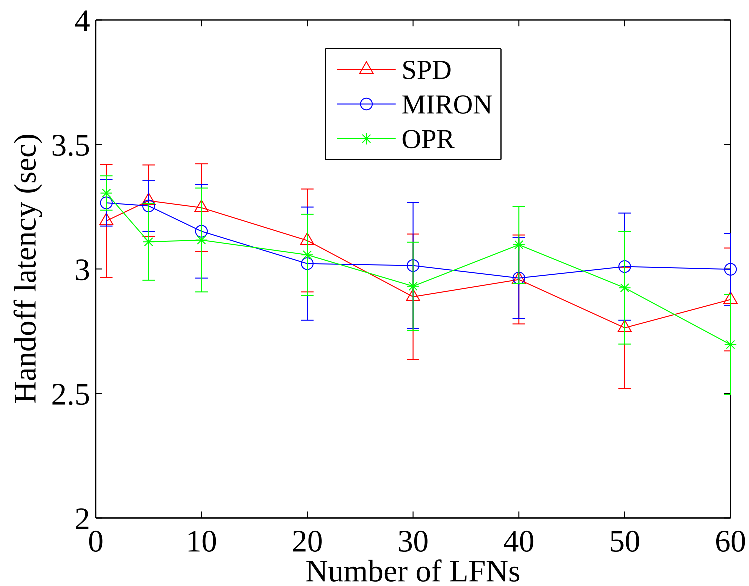

Figure 5.12: For 15m/s speed and Router-HA limk delay of 10ms, handoff latency from the MR's viewpoint.

1. After the reception of the last TCP packet at the MR through the old AR, the time required by the MR to detect that it is out of the coverage of the old AR.

2. After the detection of the coverage outage, the time required for the first router advertisement to be received at the MR.

When the mobile network is out of the old AR's coverage area, it does not receive any router advertisement from the old AR. Therefore, the MR becomes aware of the coverage outage at the expiration of the lifetime (3 seconds in our simulation) of the last router advertisement received by the MR from the old AR. Since the router advertisement is sent with an interval which is uniformly distributed from 0 to 3 seconds and the lifetime is always 3 seconds, the time required for the MR to detect the coverage outage is much larger than the half (1.5 seconds) of the lifetime. After detection of the coverage outage, a router solicitation is sent by the MR. MR's reception of the router advertisement, sent by the new AR in response to the solicitation, completes the router detection process. The time duration between sending the solicitation and receiving the router advertisement is in the order of hundreds of milliseconds. As can be observed from the results, increasing the number of LFNs does not affect the handoff latency of the mobile network.

5.4 Summary

We evaluate the TCP throughput of three prefix delegation-based schemes, namely MIRON, SPD and OPR, proposed for NEMO route optimization. Evaluation was performed as a function of the number of LFNs (fixed hosts in the mobile network) with the speed of the mobile network and its distance from its home network as parameters. Results show that route optimization for LFNs improves the aggregated TCP throughput for the entire mobile network significantly when the number of LFNs is small (not observed in our earlier work [75]) or when the distance from the home network is large. Results also show that additional signaling in MIRON does not reduce throughput significantly. Therefore, the route optimization for LFNs has to be prioritized over the reduction of signaling by not optimizing routes.

Chapter 6

Cost analysis of the NEMO schemes

In NEMO, network parameters (e.g., network size, mobility rate, traffic rate, and distances from mobility agents) influence signaling and routing overheads, resulting from the prefix delegation-based schemes. These overheads include delivery of packets through the partially-optimized route, updating home agents about the change of location, sending updates to hosts with ongoing communication, processing and lookup by mobility agents, and the delegation of prefix. These overheads cost the transmission and processing power at the network (e.g., routers in the network) between end hosts, and at the mobility management entities, such as home agents and mobile routers.

Cost analysis of NEMO protocols have been performed in [76, 77]. They present the signaling cost of NEMO BSP or a similar protocol by constructing analytical models that measure the transmission and processing costs incurred by the signaling packets. Lim et al. [32, 78] performed a cost analysis for the general approaches used for route optimization. However, the analysis presented in [32, 76–78] is unable to show the variations in the costs among the prefix delegation-based schemes adequately. In particular, previous cost analysis do not consider the cost of obtaining prefixes/CoAs and are unable to compare the cost of packet delivery through unoptimized route with other costs.

Our objective is to perform a cost analysis of the prefix delegation-based schemes by developing a framework that captures the tradeoff and the differences among the schemes. We use the term network mobility cost to refer to those costs incurred for sending packets to the hosts inside a mobile network. The notion of costs refers to the use of resources mentioned in [37,41], and is a number-only relative measure for the schemes; the higher the number, the higher the cost.

The rest of the chapter is organized as follows. Analytical cost models are presented in Sec. 6.1. Section 6.2 presents the results and comparison among the schemes followed by a discussion on the results in Sec. 6.3. Finally, Sec. 6.4 summarizes the findings from the work presented in this chapter.

6.1 Cost analysis

This section presents costs to support NEMO for the three representative prefix delegation-based schemes using analytical models. The costs measure the amount of resources being used by the schemes to support NEMO. Our cost analysis resembles the analysis performed in [76,79,80]. Unlike [76,79,80], we introduce costs of prefix delegation or CoA obtention, and effects of nesting on costs that are unique for NEMO.

We use a general NEMO architecture (as shown in Fig. 6.1) that includes LFNs, LMNs, VMNs, multiple visiting mobile networks, and multiple levels of nesting. We consider the cost to send refreshing BUs and the cost of packet delivery. In addition to finding costs incurred at the infrastructure including the mobile network, we show a entity-wise cost evaluation. The HA and the TLMR have been chosen for the entity-wise evaluation because all communications with the mobile network will be through these two entities. Therefore, resource consumptions at these entities are expected to be high, and may become a concern when the resource is limited. For tractability reasons, models were developed based on assumptions. Types of costs

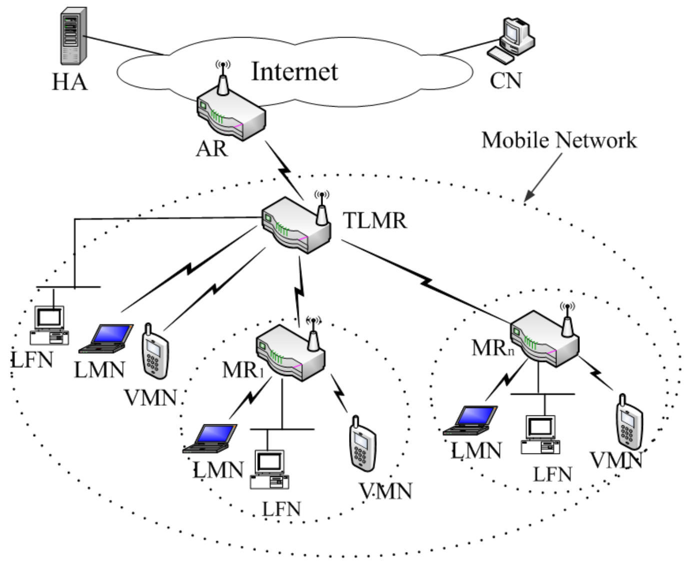

Figure 6.1: Architecture of a mobile network.

analyzed, assumptions, notations, and the models are presented in the following subsections.

6.1.1 Types of costs

We measure the following costs of the schemes:

- *Location update cost:* To maintain reachability, a node sends BUs to the HA to inform its current location whenever it obtains a CoA. Periodic BUs are sent for refreshing the binding entries. Resources (e.g., transmission and processing power, etc.), consumed by these BUs, comprise the location update cost.

- *Session continuity cost:* To continue session through an optimized route, BUs have to be sent to CNs whenever the mobile network changes the point of attachment. Resources consumed by these BUs, comprise this cost. OPR employs a technique other than sending BUs to continue sessions, and the cost incurred by the technique are also included in this type of cost.

- *Packet delivery cost:* To send a packet to the mobile network, the HA has to perform a look up to retrieve the CoA for tunneling towards the mobile network. In addition, HA and MR tunnel/de-tunnel packets. A measure of the processing and transmission power used for look up and tunneling is given by the packet delivery cost. Moreover, transmission power required by original packets are also included in this cost.

- *Prefix/CoA obtention cost:* After handoff, prefixes / CoAs are obtained from the foreign network. Resources consumed by the control messages required to obtain prefixes / CoAs comprise this cost.

6.1.2 Assumptions

For tractability reasons, our models are based on the following assumptions.

- We consider the handoff of the mobile network as a whole. Intra mobile network movements of MRs, and the movements of the mobile nodes inside the network are not considered. This assumption comply with the type of movement of a nested mobile network in a vehicle that actually motivated NEMO.

- Number of VMNs/LMNs and MRs registered with an HA are assumed to be higher than the number of VMNs/LMNs and MRs in the mobile network (by a factor α).

- We assume the worst possible scenario for the analysis, such as, all MNNs are communicating simultaneously, the CN of each session is different. These assumptions were also made in [76, 80].

6.1.3 Notations

To denote the cost terms, we have used the superscript X and the subscript Y to indicate the scheme and the type of cost, respectively. X will be replaced by N, S,

Table 6.1: Some expressions defined to simplify equations.

π_{lk}	$= \psi \log_2(\alpha(N_r + N_m))$
π_{bl}	$= \psi \log_2 N_c$ for VMNs and $\psi \log_2(N_f N_c)$ for MRs
λ_{cs}	$= N_c \lambda_s / S$
λ_{cp}	$= N_c \lambda_s F / P$
f_h	$= \left(1 + \lfloor \frac{T_r}{T_{lf}} \rfloor\right) / T_r$
f_r	$= \lfloor \frac{T_r}{T_{lf}} \rfloor / T_r$

M and O for NEMO BSP, SPD, MIRON and OPR, respectively. Y will be replaced by T, LU, SC, PD, and CO for total, location update, session continuity, packet delivery and prefix / CoA obtention costs, respectively. Some notations are used to denote expressions for simplification of models' representations, and are presented in Table 6.1. The notations that have been used in the models are presented below.

$\Lambda_Y^X =$ Cost of type Y incurred at network for scheme X,

$\Psi_Y^X =$ Cost of type Y incurred at TLMR for scheme X,

$\Phi_Y^X =$ Cost of type Y incurred at HA for scheme X,

$N_r =$ Number of MRs in mobile network,

$N_r^{(i)} =$ Number of MRs at level i,

$N_m =$ Number of mobile nodes in the mobile network,

$N_f =$ Number of LFNs in mobile network,

$N_m^{(i)} =$ Number of LMNs and VMNs at level i,

$N_f^{(i)} =$ Number of LFNs at level i,

$N_c =$ Number of CNs communicating with each node,

$l =$ Nesting Level (hops to TLMR),

$h_{ah} =$ Average number of hops between AR and HA,

$h_{ac} =$ Average number of hops between AR and CN,

$h_{hc} =$ Average number of hops between HA and CN,

$h_{hh} =$ Average number of hops between HA and HA,

$\tau_l =$ Per hop transmission cost for location update,

$\tau_s =$ Per hop transmission cost for session continuity,

$\tau_{dt} =$ Per hop transmission cost for packets without tunnel header,

$\tau_{ip} =$ Per hop transmission cost for tunnel header,

$\tau_{rh} =$ Per hop transmission cost for home address destination option or routing header type 2,

$\tau_d =$ Average transmission cost of DHCPv6 messages,

$\tau_p =$ Average transmission cost of PANA messages,

$\tau_a =$ Average transmission cost of route request-reply messages of AODV protocol,

$\tau_r =$ Transmission cost for the router advertisement,

$\sigma =$ Proportionality constant of transmission cost over wired and wireless network,

$\pi_{lk} =$ Lookup costs,

$\pi_h =$ BU processing cost,

$\pi_t =$ Tunnel processing costs at HA and MR,

$\pi_{rh} =$ Routing header processing cost,

$\lambda_s =$ Average session arrival rate,

$\lambda_{cs} =$ For an MNN, session arrival rate from all CNs,

$S =$ number of sessions,

$\lambda_{cp} =$ For an MNN, average packet arrival rate from all CNs,

$F =$ File size,

$P =$ Maximum transmission unit,

$T_r =$ Subnet residence time,

$T_{lf} =$ Lifetime of binding entry,

$T_{ra} =$ Interval of sending periodic router advertisement,

$f_h =$ The rate of sending BUs per second for both handoff and refreshing,

$f_r =$ The rate of sending BUs per second for refreshing,

$\phi =$ Fraction of MRs acting as TLMR,

$\alpha =$ Ratio of number of mobile nodes registered to the HA to number of mobile nodes in the mobile network.

The models are developed to show the differences in the costs of the schemes from the view point of total cost rather than that of differential cost. Showing only the differences in costs might give an impression of inflated differences. Therefore, we consider all parameters required to compute the costs. However, following are the parameters that are the keys as far as the differences of the schemes are concerned:

- h_{ah} and h_{hc}: These two represent the distance of the mobile network from the home network, and will affect the differences of the costs depending on the degree of optimization.

- The number and types of MNNs, and CNs: Depending on the number of and types of MNNs and CNs, the signaling and the transmission costs among the schemes may vary.

- λ_{cp}: It represents the amount of data exchanged with the mobile network, and will affect the packet delivery cost.

- T_r: The subnet residence time affects the amount of signaling in a scheme and can make difference among the signaling cost of the schemes depending on the number and types of MNNs.

The key parameters are also discussed in Sec. 6.2.

6.1.4 Cost models for the schemes

Analytical models for the costs are presented in the following subsections. We have provided detailed description of the cost terms of NEMO BSP to make readers familiar with the cost terms. Detailed description of the cost terms for other schemes (SPD, MIRON and OPR) can be found at [81].

6.1.4.1 NEMO BSP

- **Location update cost:** After handoff, TLMR sends a BU to the HA to perform the location update, and receives a BA. Handoff occurs every T_r seconds. In addition to the BU sent after handoff, MRs and mobile nodes send refreshing BU $\lfloor \frac{T_r}{T_{lf}} \rfloor$ times during the period of T_r seconds. Therefore, the frequency of sending BUs including BUs sent during handoff is, $f_h = \left(1 + \lfloor \frac{T_r}{T_{lf}} \rfloor\right) / T_r$, and the frequency of sending refreshing BUs is, $f_r = \lfloor \frac{T_r}{T_{lf}} \rfloor / T_r$. BUs sent from MRs and mobile nodes at level i undergoes i number of tunneling resulting in additional transmission cost due to tunnel header. Since all BU/BAs go through the TLMR, the cost at the TLMR is given by

$$\Psi_{LU}^{N} = 2\sigma\tau_l f_h + 2\sigma \sum_{i=1}^{i=l} \left(N_r^{(i)} + N_m^{(i)}\right)(\tau_l + i\tau_{ip} + \pi_t) f_r. \tag{6.1}$$

To find out the cost incurred at the HA due to the location update, we need to consider the updating of the binding cache in addition to the cost mentioned above. Updating the binding cache is required for each MR and mobile node registered to an HA. In addition, tunneled BUs incur a look up cost. Since $(N_r + N_m)$ nodes are managed by the HA, the look up is performed in a table of $(N_r + N_m)$ entries and with a look up key of size equal to the IPv6 address. Assuming a binary search, the look up cost is $\pi_{lk} = \psi \log_2(N_r + N_m)$, where ψ is the cost of the look up per operation. Therefore, cost incurred at HA due to location update becomes

$$\begin{aligned} \Phi_{LU}^{N} &= \phi N_r(2\tau_l + \pi_h) f_h + 2\sum_{i=1}^{l} \left(N_r^{(i)} + N_m^{(i)}\right) \\ &\times \left((\tau_l + i\tau_{ip} + \pi_t) + \pi_{lk} + 0.5\pi_h\right) f_r. \end{aligned} \tag{6.2}$$

The cost of location update for the network includes transmission costs at all hops upto the HA including the costs incurred at MRs and the HA. Transmission costs for MRs and mobile nodes at level i are incurred at $h_{ah} + ih_{hh}$ wired hops and $(i + 1)$ wireless hops. The transmission cost upto the TLMR increases by τ_{ip} at each level due to tunneling, and at each HA it decreases by the same amount. Also, each BU sent from a node at level i undergoes $2i$ number of tunneling and de-tunneling. Therefore, location update cost is

$$
\begin{aligned}
\Lambda_{LU}^{N} = & \left(2(h_{ah}+\sigma)\tau_l+\pi_h\right) f_h + 2\sum_{i=1}^{l}\left(N_r^{(i)}+N_m^{(i)}\right) \\
& \times \Big((i+1)\sigma\tau_l+\sigma\sum_{j=1}^{i} j\tau_{ip}+2i\pi_t+(h_{ah}+ih_{hh})\tau_l \\
& + h_{ah}i\tau_{ip}+\sum_{j=0}^{i-1} jh_{hh}\tau_{ip}+i\pi_{lk}+0.5\pi_h\Big) f_r.
\end{aligned}
\tag{6.3}
$$

where, $2(h_{ah}+\sigma)\tau_l+\pi_h$ includes costs incurred due to BUs/BAs sent by TLMR/HAs, $\sum_{i=1}^{l}\left(N_r^{(i)}+N_m^{(i)}\right)$ includes the number of nodes that send refreshing BUs, $(i+1)\sigma\tau_l+\sigma\sum_{j=1}^{i} j\tau_{ip}+i\pi_t$ includes transmission and tunnel processing costs incurred inside the mobile network and at AR, $i\pi_t+(h_{ah}+ih_{hh})\tau_l+h_{ah}i\tau_{ip}+\sum_{j=0}^{i-1} jh_{hh}\tau_{ip}+i\pi_{lk}+0.5\pi_h$ includes tunnel processing, transmission, and BU processing costs incurred at hops after AR upto HA.

- **Session continuity cost:** Each mobile node sends BUs to (and receive a BAs from) its CNs for session continuity. Since only TLMR's CoA changes during handoff, mobile nodes send only refreshing BUs. Thus, the cost incurred at the TLMR is

$$
\Psi_{SC}^{N} = 2N_c\sum_{i=1}^{l} N_m^{(i)}\left(\sigma\left(\tau_s+i\tau_{ip}\right)+\pi_t\right) f_r. \tag{6.4}
$$

Since BUs are tunneled through the HA, the cost incurred at the HA includes look up, tunneling and transmission costs, and is given as

$$
\Phi_{SC}^{N} = 2N_c\sum_{i=1}^{l} N_m^{(i)}(\tau_s+i\tau_{ip}+\pi_t+\pi_{lk}) f_r. \tag{6.5}
$$

The session continuity cost for the network includes the costs at each hop upto CNs, MNNs and at other MRs, and the cost of updating the binding update list incurred per packet at the VMNs in addition to the above cost, and is given by

$$\begin{aligned}\Lambda_{SC}^{N} =&2N_c \sum_{i=1}^{l} N_m^{(i)} \Bigg((i+1)\sigma\tau_s + \sigma \sum_{j=1}^{i} j\tau_{ip} \\ &+ 2i\pi_t + (h_{ah} + (i-1)h_{hh} + h_{hc})\tau_s + h_{ah}i\tau_{ip} \\ &+ \sum_{j=0}^{i-1} jh_{hh}\tau_{ip} + i\pi_{lk} + 0.5\pi_h \Bigg) f_r + \lambda_{cp}\pi_{bl}N_m,\end{aligned} \tag{6.6}$$

where, $N_c \sum_{i=1}^{l} N_m^{(i)}$ is the number of BUs sent to CNs, $(i+1)\sigma\tau_s+\sigma\sum_{j=1}^{i} j\tau_{ip}+ i\pi_t$ includes transmission and tunnel processing costs incurred inside the mobile network and at AR, $i\pi_t + (h_{ah} + (i-1)h_{hh} + h_{hc})\tau_s + h_{ah}i\tau_{ip} + \sum_{j=0}^{i-1} jh_{hh}\tau_{ip} + i\pi_{lk} + 0.5\pi_h$ includes tunnel processing, transmission, and BU processing costs incurred at hops after AR upto the CN, and $\lambda_{cp}\pi_{bl}N_m$ is the cost for updating the binding update list at VMNs.

- **Packet delivery cost:** Data packets incurs transmission and tunneling cost which is similar to that of BU packets. For each MNN, costs are incurred at a rate proportional to the packet arrival rate, $\lambda_{cp} = N_c\lambda_s F/P$, from all CNs. For the packets sent to mobile nodes, we assume that only the first packet of a session is sent through the HA before a BU is received at the CN, and additional costs are incurred at a rate, $\lambda_{cs} = N_c\lambda_s/S$ for all CNs. TLMR needs to de-tunnel and forward packets to the MNNs at the next level. Additional cost incurred at the TLMR for the first packets sent to mobile nodes is the increased transmission cost for one additional tunnel. Therefore, the cost at the TLMR is

$$\begin{aligned}\Psi_{PD}^{N} =&\lambda_{cp}\Bigg(\sum_{i=1}^{l}\left(N_f^{(i)}+N_m^{(i)}\right)\Big(\sigma(\tau_{dt}+(i-1)\tau_{ip})\\&+\pi_t\Big)+\sigma\tau_{rh}N_m\Bigg)+\sigma N_m(\tau_{ip}+\tau_{dt})\lambda_{cs}.\end{aligned} \tag{6.7}$$

In addition to the transmission cost, costs incurred at the HA are due to look up, tunneling and the transmission cost for one additional tunnel. Therefore, the packet delivery cost at the HA is

$$\begin{aligned}\Phi_{PD}^{N} =&\lambda_{cp}\Bigg(\sum_{i=1}^{l}(N_f^{(i)}+N_m^{(i)})\Big(\tau_{dt}+i\tau_{ip}+\pi_{lk}+\pi_t\Big)\\&+\tau_{rh}N_m\Bigg)+\lambda_{cs}N_m\left(\tau_{dt}+2\tau_{ip}+\pi_{lk}+\pi_t\right).\end{aligned} \tag{6.8}$$

The packet delivery cost for the network can be obtained at each hop similar to the session continuity cost. Additionally, for the first packet sent through the HA of mobile nodes, costs are incurred due to transmission through h_{hh} hops, tunneling, look up and transmission of one additional tunnel header. Therefore, the packet delivery cost for the network is given by

$$
\begin{aligned}
\Lambda_{PD}^{N} &= \lambda_{cp}(costs\, incurred\, per\, packet) \\
&+ \lambda_{cs}(additional\, costs\, incurred\, for\, the\, first\, packet) \\
&= \lambda_{cp}\Bigg(\sum_{i=1}^{l}\left(N_f^{(i)} + N_m^{(i)}\right)\Big(i\pi_{lk} + 2i\pi_t + (h_{ah} + h_{hc} \\
&+ (i-1)h_{hh})\tau_{dt} + ih_{ah}\tau_{ip} + \sum_{j=0}^{i-1} jh_{hh}\tau_{ip} + \sigma\sum_{j=1}^{i} j\tau_{ip} \\
&+ \sigma\tau_{dt}(i+1)\Big) + \sum_{i=1}^{l} N_m^{(i)}\Big(\big(h_{ah} + (i-1)h_{hh} + h_{hc} \\
&+ \sigma(i+1)\big)\tau_{rh} + 2\pi_{rh}\Big)\Bigg) + \lambda_{cs}\sum_{i=1}^{l} N_m^{(i)}\Big(\pi_{lk} + 2\pi_t \\
&+ h_{hh}(\tau_{dt} + \tau_{ip}) + \sigma i\tau_{ip} + h_{ah}\tau_{ip} + ih_{hh}\tau_{ip}\Big).
\end{aligned}
\tag{6.9}
$$

For the subexpression showing the per packet costs, $i\pi_{lk}+2i\pi_t+(h_{ah}+(i-1)h_{hh}+h_{hc})(\tau_{dt}+\tau_{rh})+ih_{ah}\tau_{ip}+\sum_{j=0}^{i-1} jh_{hh}\tau_{ip}$ includes lookup, tunnel processing, transmission costs incurred at hops from the CN until the AR, and $\sigma\sum_{j=1}^{i} j\tau_{ip}+\sigma(\tau_{dt}+\tau_{rh})(i+1)$ includes transmission and tunnel processing costs incurred inside the mobile network for $\sum_{i=1}^{l}\left(N_f^{(i)}+N_m^{(i)}\right)$ VMNs and LFNs that receive packets from CNs. $\big(h_{ah}+(i-1)h_{hh}+h_{hc}+\sigma(i+1)\big)\tau_{rh}+2\pi_{rh}$ includes the additional transmission and processing costs for the home address destination option for VMNs only.

For the subexpression showing the additional costs for the first packet only, $\sum_{i=1}^{l} N_m^{(i)}$ is the number of VMNs that send the first packet through their HA, $h_{hh}(\tau_{dt}+\tau_{ip})$ includes the transmission costs incurred at the hops from the MR's HA upto the VMN's HA, and $\pi_{lk}+2\pi_t$ includes the lookup and tunnel processing costs in the additional HA, and $\sigma i\tau_{ip}+h_{ah}\tau_{ip}+ih_{hh}\tau_{ip}$ includes the transmission costs at the hops from the VMN upto its MR's HA due to one additional tunnel header.

- **Prefix/CoA obtention cost**: After every handoff, only TLMR obtains a CoA from the foreign network. Therefore, costs incurred due to prefix or CoA obtention are zero.

- **Total cost**: Combining the costs presented above, we find the costs incurred at the TLMR, the HA and the network given by Eqns. (6.10), (6.11) and (6.12), respectively.

$$\Psi_T^N = \Psi_{LU}^N + \Psi_{SC}^N + \Psi_{PD}^N \tag{6.10}$$

$$\Phi_T^N = \Phi_{LU}^N + \Phi_{SC}^N + \Phi_{PD}^N \tag{6.11}$$

$$\Lambda_T^N = \Lambda_{LU}^N + \Lambda_{SC}^N + \Lambda_{PD}^N \tag{6.12}$$

6.1.4.2 SPD

- **Location update cost:** In SPD, location update after handoff is performed by each MR and mobile node by sending a BU to the HA, and receiving a BA. In addition to the BU sent after handoff, refreshing BUs are sent periodically. Thus, BUs are sent at a rate given by f_h. Since all BU/BAs go through the TLMR, the cost at the TLMR is given by

$$\Psi_{LU}^S = 2\sigma\tau_l(N_r + N_m)f_h. \tag{6.13}$$

To find the cost incurred at the HA due to the location update, we need to consider the updating of the binding cache in addition to the cost mentioned above. Therefore, cost incurred at HA due to location update becomes

$$\Phi_{LU}^S = (2\tau_l + \pi_h)\,(N_r + N_m)f_h. \tag{6.14}$$

The cost of location update for the network includes transmission costs at all hops upto the HA including the costs incurred at the TLMR and the HA. Transmission costs for all MRs and mobile nodes are incurred at h_{ah} wired hops. For nodes at level i, transmission costs are incurred at $i+1$ wireless hops. Therefore, location update cost is given by

$$\Lambda_{LU}^{S} = \Bigg(2\tau_l \Big((N_r + N_m)\, h_{ah} + \sigma \sum_{i=0}^{l} (i+1) \times \left(N_r^{(i)} + N_m^{(i)}\right) \Big) + (N_r + N_m)\, \pi_h \Bigg) f_h. \tag{6.15}$$

- **Session continuity cost:** In SPD, each mobile node sends BUs to (and receive BAs from) CNs for session continuity. The cost incurred at the TLMR is thus

$$\Psi_{SC}^{S} = 2\sigma\tau_s N_m N_c f_h. \tag{6.16}$$

The session continuity cost for the network also includes costs at each hop upto CNs and at other MRs and VMNs, and is given by

$$\Lambda_{SC}^{S} = 2\tau_s N_c \big(N_m h_{ac} + \sigma \sum_{i=0}^{l} (i+1) N_m^{(i)} + 0.5\pi_h N_c N_m \big) f_h + \lambda_{cp}\pi_{bl} N_m. \tag{6.17}$$

- **Packet delivery cost:** For every packet, sent from a CN to an LFN, the HA of the LFN looks up the binding cache to find the CoA to encapsulate the packet for tunneling. Tunneling and look up costs are incurred at a rate proportional to the packet arrival rate given by λ_{cp}. For the packets sent to mobile nodes, we assume that only the first packet is sent through the HA before a BU is received at the CN while subsequent packets are sent through the optimized route using the home address destination option, and thus, the costs are incurred at a rate given by λ_{cs}. TLMR needs to de-tunnel these packets only for attached LFNs. Therefore, the cost at the TLMR is

$$\begin{aligned}\Psi_{PD}^{S} =& \lambda_{cp}\Big(N_f^{(1)}\pi_t + \sigma\tau_{ip}\big(N_f - N_f^{(1)}\big) \\ &+ \sigma(\tau_{dt}N_f + (\tau_{dt} + \tau_{rh})N_m)\Big) + \sigma\tau_{ip}\lambda_{cs}N_m.\end{aligned} \tag{6.18}$$

The HA needs to perform look up, tunneling and transmit the packet resulting in a cost as

$$\Phi_{PD}^{S} = (\lambda_{cp}N_f + \lambda_{cs}N_m)(\pi_{lk} + \tau_{dt} + \tau_{ip} + \pi_t). \tag{6.19}$$

In addition to the cost incurred at the HA and TLMR, the packet delivery cost for the network have other costs that include the transmission costs at nested MRs and routers upto the CN. For the case of mobile nodes, transmission costs are incurred at each hop between the AR and the CN for all but the first packet. For the case of LFNs and session's first packet of mobile nodes, transmission costs are incurred at each hop from the CN upto the HA, and from the HA upto the AR. For the latter case, additional costs are incurred due to tunnel header at each hop between the HA and the MR for the destination MNN along with the tunneling cost incurred at the MR because it de-tunnels packets. Therefore, the packet delivery cost for the network is given by

$$
\begin{aligned}
\Lambda_{PD}^{S} =& \lambda_{cp}\Big(N_f(\pi_{lk} + 2\pi_t + (h_{ah} + h_{hc})\tau_{dt} + h_{ah}\tau_{ip}) \\
&+ N_m(h_{ac}(\tau_{dt} + \tau_{rh}) + 2\pi_{rh}) + \sigma \sum_{i=1}^{l}(i+1)(\tau_{dt}N_f^{(i)} \\
&+ (\tau_{dt} + \tau_{rh})N_m^{(i)}) + \sigma\tau_{ip}\sum_{i=1}^{l} iN_f^{(i)}\Big) + \lambda_{cs}\Big(N_m(\pi_{lk} \\
&+ 2\pi_t + h_{ah}(\tau_{dt} + \tau_{ip})) + \sigma\tau_{ip}\sum_{i=1}^{l}(i+1)N_m^{(i)}\Big).
\end{aligned} \tag{6.20}
$$

- **Prefix/CoA obtention cost**: In SPD, prefix and CoAs can be obtained from the MR above using DHCPv6 procedures. This requires a request and a reply message, and some processing at the MR for prefix delegation [82]. Since the TLMR delegates prefixes to attached MRs and provide CoAs to attached mobile nodes, the cost incurred at the TLMR is

$$
\Psi_{CO}^{S} = \frac{2\sigma\tau_d\left(N_r^{(1)} + N_m^{(1)}\right)}{T_r}. \tag{6.21}
$$

The cost incurred for the entire mobile network is

$$
\Lambda_{CO}^{S} = \frac{2\sigma\tau_d\left(N_r + N_m\right)}{T_r}. \tag{6.22}
$$

- **Total cost**: Combining the costs presented above, we find the costs of SPD incurred at the TLMR, the HA and the network given by Eqns. (6.23), (6.24) and (6.25), respectively.

$$
\Psi_T^S = \Psi_{LU}^S + \Psi_{SC}^S + \Psi_{PD}^S + \Psi_{CO}^S \tag{6.23}
$$

$$
\Phi_T^S = \Phi_{LU}^S + \Phi_{PD}^S \tag{6.24}
$$

$$
\Lambda_T^S = \Lambda_{LU}^S + \Lambda_{SC}^S + \Lambda_{PD}^S + \Lambda_{CO}^S \tag{6.25}
$$

6.1.4.3 MIRON

- **Location update cost:** Location update for MIRON is similar to that of SPD. Therefore, location update costs for the TLMR, the HA and the network is as follows:

$$\Psi_{LU}^{M} = \Psi_{LU}^{S} \tag{6.26}$$

$$\Phi_{LU}^{M} = \Phi_{LU}^{S} \tag{6.27}$$

$$\Lambda_{LU}^{M} = \Lambda_{LU}^{S} \tag{6.28}$$

- **Session continuity cost:** For session continuity, BUs are sent to CNs by mobile nodes, and by MRs on behalf of the attached LFNs. Thus, the costs for MIRON are similar to the costs of SPD except the additional but identical costs for LFNs. Therefore, the costs incurred at the TLMR and at the network are give by Eqns. (6.29) and (6.30), respectively.

$$\Psi_{SC}^{M} = 2N_c\left(N_f + N_m\right)\sigma\tau_s f_h + \lambda_{cp}\pi_{bl}N_f^{(1)} \tag{6.29}$$

$$\begin{aligned}\Lambda_{SC}^{M} =& 2N_c\Bigg(\left(N_f + N_m\right)\left(h_{ac}\tau_s + 0.5\pi_h\right) + \sigma\tau_s \\ &\times \sum_{i=0}^{l}(i+1)\left(N_f^{(i)} + N_m^{(i)}\right)\Bigg) f_h + \lambda_{cp}\pi_{bl}(N_f + N_m)\end{aligned} \tag{6.30}$$

- **Packet delivery cost:** In MIRON, route optimization is performed for all MNNs. Therefore, packet delivery cost for all MNNs are like that for mobile nodes in SPD. Therefore, the costs for the TLMR, the HA and the network are given by Eqns. (6.31), (6.32) and (6.33), respectively.

$$\Psi_{PD}^{M} = \lambda_{cs}\left(N_f^{(1)}\pi_t + \sigma\tau_{ip}(N_f - N_f^{(1)} + N_m)\right) + \sigma\lambda_{cp}(\tau_{dt} + \tau_{rh})(N_f + N_m) \tag{6.31}$$

$$\Phi_{PD}^{M} = \lambda_{cs}\left(N_f + N_m\right)\left(\pi_{lk} + \tau_{dt} + \tau_{ip} + \pi_t\right) \tag{6.32}$$

$$\begin{aligned}\Lambda_{PD}^{M} = \lambda_{cs}\Bigg(&(N_f + N_m)\Big(\pi_{lk} + 2\pi_t + (h_{ah} + h_{hc}) \times \tau_{dt} \\ &+ h_{ah}\tau_{ip}\Big) + \sigma\tau_{ip}\Big(\sum_{i=1}^{l} iN_f^{(i)} + \sum_{i=1}^{l}(i+1)N_m^{(i)}\Big)\Bigg) \\ &+ \lambda_{cp}\Bigg((h_{ac}(\tau_{dt} + \tau_{rh}) + 2\pi_{rh}) \times (N_f + N_m) + \sigma\Big((\tau_{dt} \\ &+ \tau_{rh})\sum_{i=1}^{l} i(N_f^{(i)} + N_m^{(i)}) + \tau_{dt}N_f + (\tau_{dt} + \tau_{rh})N_m\Big)\Bigg)\end{aligned} \tag{6.33}$$

- **Prefix/CoA obtention cost**:

  Two DHCPv6 messages for each MNN (except LFNs) are forwarded by the TLMR along with the transmission of two PANA messages for attached MRs resulting in the cost incurred at the TLMR as follows:

$$\Psi_{CO}^{M} = \frac{2\sigma}{T_r}\Big((N_r^{(1)} + N_m^{(1)})\tau_p + \left(N_r + N_m\right)\tau_d\Big). \tag{6.34}$$

  For each MNN except the TLMR and LFNs, four PANA messages have to be transmitted, and equal number of replies follow. Moreover, two DHCPv6 messages for each MR and mobile node at level i are transmitted across i number of wireless hops. Therefore, prefix/CoA obtention cost for the network becomes

$$\Lambda_{CO}^{M} = \frac{\sigma}{T_r}\Bigg(8\left(N_r - 1 + N_m\right)\tau_p + 2\sum_{i=0}^{l}(i+1)\left(N_r^{(i)} + N_m^{(i)}\right)\tau_d\Bigg). \tag{6.35}$$

- **Total cost**: Like SPD, the total costs for MIRON are given by Eqns. (6.36), (6.37) and (6.38).

$$\Psi_T^M = \Psi_{LU}^M + \Psi_{SC}^M + \Psi_{PD}^M + \Psi_{CO}^M \tag{6.36}$$

$$\Phi_T^M = \Phi_{LU}^M + \Phi_{PD}^M \tag{6.37}$$

$$\Lambda_T^M = \Lambda_{LU}^M + \Lambda_{SC}^M + \Lambda_{PD}^M + \Lambda_{CO}^M \tag{6.38}$$

6.1.4.4 OPR

- **Location update cost:** In OPR, only MRs obtain CoAs after handoff, and perform location update with the HA. Mobile nodes, being transparent to the mobility, send refreshing BUs only. Therefore, we can find the costs like the previous schemes by considering all BUs sent by MRs, and refreshing BUs sent by mobile nodes.

$$\Psi_{LU}^O = 2N_r\sigma\tau_l f_h + 2N_m\sigma\tau_l f_r \tag{6.39}$$

$$\Phi_{LU}^O = N_r\left(2\tau_l + \pi_h\right) f_h + N_m\left(2\tau_l + \pi_h\right) f_r \tag{6.40}$$

$$\begin{aligned}\Lambda_{LU}^O = &\left(2\tau_l\Big(N_r h_{ah} + \sigma\sum_{i=0}^{l}(i+1)N_r^{(i)}\Big) + N_r\pi_h\right) f_h \\ &+ \left(2\tau_l\Big(N_m h_{ah} + \sigma\sum_{i=0}^{l}(i+1)N_m^{(i)}\Big) + N_m\pi_h\right) f_r\end{aligned} \tag{6.41}$$

- **Session continuity cost:** Since mobile nodes in OPR do not need MIPv6 route optimization, we assume that no BU is sent to CNs. Therefore, the session continuity cost due to the sending of BUs to CNs is zero. But for every packet sent to the CN from each attached MNN at level $(i+1)$, the MR at level i needs to look up the DPT table for the translated address. Size of the

DPT table is proportional to the number of attached LFNs and mobile nodes at level $i+1$. Therefore, the session continuity cost at the TLMR (at level zero) is

$$\Psi_{SC}^{O} = \lambda_{cp}\big(N_f^{(1)} + N_m^{(1)}\big)\big(\psi \log_2(N_f^{(1)} + N_m^{(1)})\big). \tag{6.42}$$

Considering the look up cost for all MRs while assuming equal number of MNNs attached under each MR, the session continuity cost for the network becomes

$$\Lambda_{SC}^{O} = \lambda_{cp}\big(\psi \log_2 \sum_{i=0}^{l} \frac{1}{N_r^{(i)}} \left(N_f^{(i+1)} + N_m^{(i+1)}\right)^2 + (N_f + N_m)(\pi_h + \pi_{rh}) \tag{6.43}$$

where $N_r^{(i)} \neq 0$.

- **Packet delivery cost:** Similar to MIRON, the first packet go through the HA until the CN receives the translated address from the packet sent to the CN in response to the first packet received at an MNN. Therefore, costs for OPR are as follows:

$$\Psi_{PD}^{O} = \Psi_{PD}^{M} \tag{6.44}$$

$$\Phi_{PD}^{O} = \Phi_{PD}^{M} \tag{6.45}$$

$$\Lambda_{PD}^{O} = \Lambda_{PD}^{M} \tag{6.46}$$

- **Prefix/CoA obtention cost**: Prefix obtention procedure is similar to that of SPD except that only MRs obtain the prefix. Therefore, by excluding the cost for mobile nodes from the expressions derived for SPD, we can find the prefix/CoA obtention cost for the TLMR and the network given by Eqns. (6.47) and (6.48), respectively.

$$\Psi^{O}_{CO} = \frac{2\sigma\tau_d N_r^{(1)}}{T_r} \tag{6.47}$$

$$\Lambda^{O}_{CO} = \frac{2\sigma\tau_d N_r}{T_r} \tag{6.48}$$

- **Total cost**: The total costs for OPR are given by Eqns. (6.49), (6.50) and (6.51).

$$\Psi^{O}_{T} = \Psi^{O}_{LU} + \Psi^{O}_{SC} + \Psi^{O}_{PD} + \Psi^{O}_{CO} \tag{6.49}$$

$$\Phi^{O}_{T} = \Phi^{O}_{LU} + \Phi^{O}_{PD} \tag{6.50}$$

$$\Lambda^{O}_{T} = \Lambda^{O}_{LU} + \Lambda^{O}_{SC} + \Lambda^{O}_{PD} + \Lambda^{O}_{CO} \tag{6.51}$$

6.2 Results

In this section, we obtain numerical values for the costs using the expressions derived in the cost analysis section in a simplified format. We present the costs as a function of the number of mobile nodes, the number of MRs, the number of LFNs, the number of CNs, the subnet residence time and the number of hops between entities. The location update and the session continuity costs vary among the schemes depending on the number and types of MNNs and the number of CNs. The number of data packets sent to the mobile network is proportional to the number of CNs to determine the packet delivery cost. In [83], the subnet residence time has been shown to affect the cost. Moreover, the number of hops between various mobility entities determines the packet delivery cost.

Table 6.2: Values of parameters used in the numerical analysis.

Parameter	**Value**	**Parameter**	**Value**
N_m	120	N_f	80
N_r	5	N_c	5
T_r	120 sec	T_{lf}	420 sec
h_{ah}	10	h_{ac}	10
h_{hc}	10	h_{hh}	10
l	2	ϕ	0.1
σ	10	ψ	0.3
τ_l	0.68	τ_s	0.68
τ_{ip}	0.4	π_t	0.4
λ_s	0.01	S	10
F	10240 bytes	P	576 bytes
τ_{dt}	5.76	τ_d	1.4
τ_p	0.56	τ_a	1.56
τ_r	0.72	π_h	0.68
τ_{rh}	0.24	π_{rh}	0.4
α	10		

The default values of the parameters used to obtain the numerical results are shown in Table 6.2. As far as the numbers of MNNs are considered, we consider a large mobile network (e.g., a mobile network onboard a train) with the number of MNNs around 200. We have used α=10. The determination of the actual value of α is not possible since NEMO has not been deployed yet in real operational network. Values of the parameters related to the file-size, packet-size, session arrival rates and the proportionality constant for the wireless network are taken from [76]. The number of hops between various mobility entities is 10 which is reasonable for the networks within USA [84]. Transmission costs are relative and determined based on the packet size assuming unit cost per 100 bytes. Similarly, processing costs, except the lookup cost, are determined assuming unit cost per 100 bytes. The transmission and processing costs are determined following the technique used in [79, 85]. For the lookup cost (Table 6.1), we assume a logarithmic time for the lookup with the proportionality constant as the processing cost per entry.

For the measurement of costs on TLMR, HA, and complete network, we assume a mobile network topology which is simplified from the network shown in Fig. 6.1. Since there exists no standard architecture for NEMO, we are using a generalized topology upon which different prefix delegation-based schemes have been proposed. We are assuming the mobile network to have a two-level hierarchy of Mobile Routers. There is one MR at level 0 or top level (which is the TLMR), hence $N_r^{(0)} = 1$. No LFN, LMN and VMN is connected directly to the TLMR. The TLMR is connected to $N_r^{(1)}$ number of level one routers, so $N_r^{(1)} = N_r - 1$ as there is no other mobile router at level 2. Hence, $N_r^{(2)} = 0$. There is no hosts (mobile or fixed) at level 0, and level 1. So $N_m^{(0)} = N_f^{(0)} = 0$, and $N_m^{(1)} = N_f^{(1)} = 0$. All LFNs and mobile nodes are at level 2, i.e., $N_m^{(2)} = N_m$, and $N_f^{(2)} = N_f$.

6.2.1 TLMR

In this subsection, we present results to show network mobility costs on the TLMR in NEMO BSP, SPD, MIRON and OPR. We vary the number of mobile nodes, the number of mobile routers, the number of LFNs, the subnet residence time, and the number of CNs.

The cost incurred at the TLMR is given by Figs. 6.2 – 6.4 as a function of the number of mobile nodes, subnet residence time and the number of CNs, respectively. The cost associated with delivery of data packets dominates the other costs to determine the characteristics of the total costs. The cost of NEMO BSP is the highest due to the packet delivery cost that results from the transmission cost of multiple tunneled packets. SPD's cost is smaller than OPR because the transmission cost of tunneled packets is incurred only for LFNs.

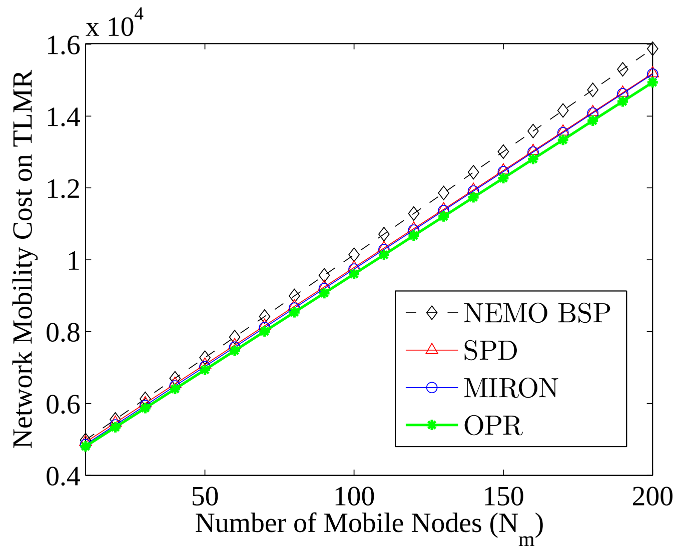

Figure 6.2: Network Mobility Cost on TLMR vs. number of MHs for the four schemes.

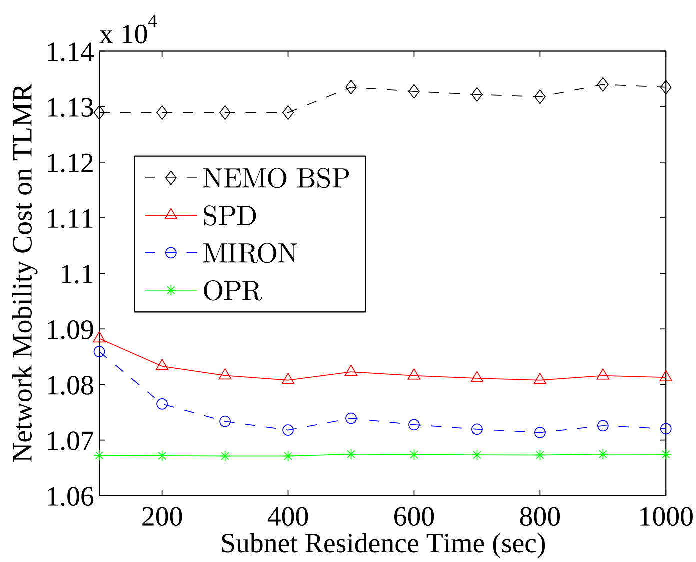

Figure 6.3: Network Mobility Cost on TLMR vs. subnet residence time for the four schemes.

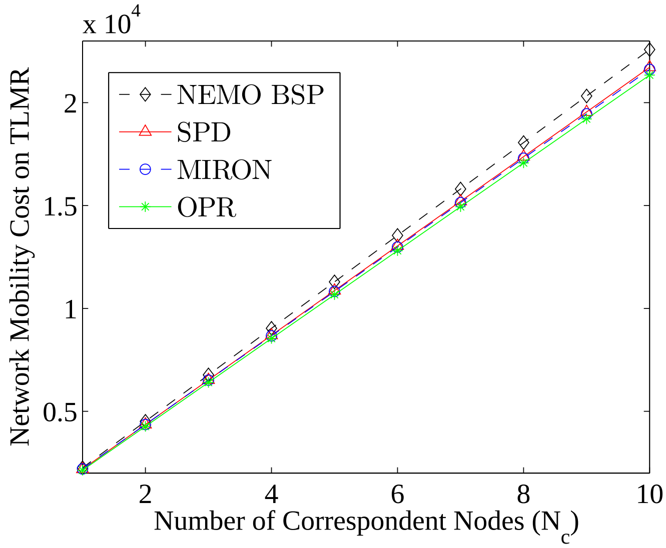

Figure 6.4: Network Mobility Cost on TLMR vs. number of CNs for the four schemes.

The costs of MIRON and OPR are smaller than other schemes. MIRON's cost is little higher than OPR due to the transmission cost incurred for signaling which is required for only MRs in OPR. Also, MIRONs prefix obtention cost is higher than OPR. The costs as a function of subnet residence time (Fig. 6.3) show negligible changes because of the dominance of the packet delivery cost that does not depend on these two parameters.

6.2.2 Home Agent

The effects of the number of mobile nodes, the number of LFNs, and the number of CNs on the cost incurred at the HA are shown in Figs. 6.5 – 6.7, respectively. Like the costs incurred at the TLMR, the cost associated with the packet delivery dominates over other costs. Therefore, the characteristics of the costs at the HA are similar to that at the TLMR except some differences that are explained in the following paragraphs.

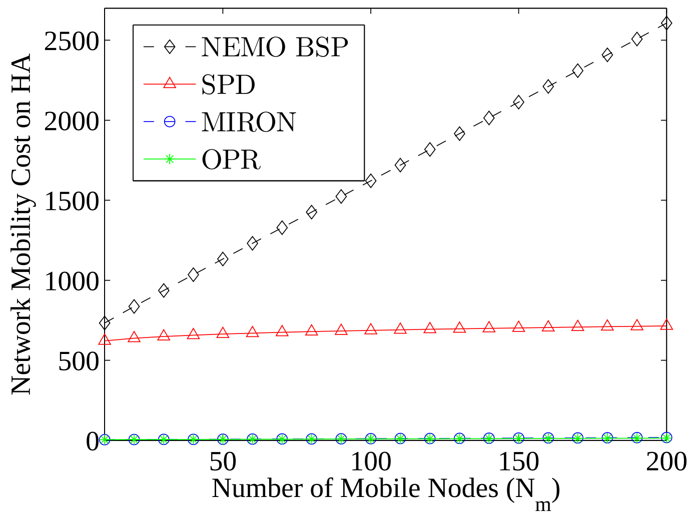

Figure 6.5: Network Mobility Cost of HA vs. number of mobile nodes for the four schemes.

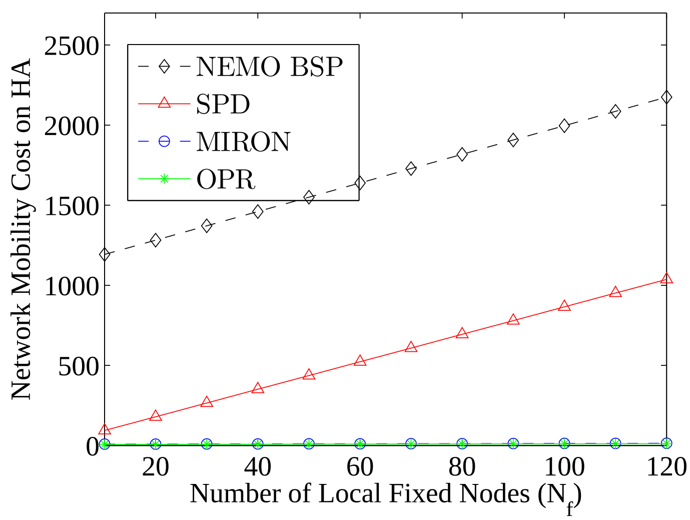

Figure 6.6: Network Mobility Cost on HA vs. number of LFNs for the four schemes.

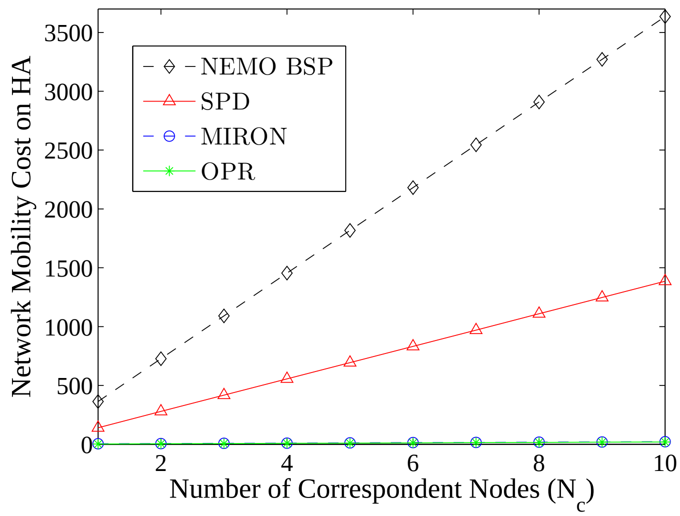

Figure 6.7: Network Mobility Cost on HA vs. number of CNs for the four schemes.

For NEMO BSP, costs increase linearly with the increase of the number of mobile nodes (Fig. 6.5) due to the lookup cost incurred at the HA for tunneling. Lookup cost is proportional to the number of mobile nodes because lookup is required for each mobile node. For SPD, such look up cost is incurred for LFNs only resulting in a negligible (logarithmic) increase rate due to increase of the size of the binding cache.

For MIRON and OPR, the cost is much lower (when compared to the cost incurred at the TLMR) than the costs of other schemes due to the reason described next. Firstly, the dominant look up cost is incurred only for the first packet of a session, thus have negligible effect on the overall increase rate of the cost. Secondly, the location updates sent to the CNs do not incur any cost at the HA.

6.2.3 Complete Network

The cost incurred at the network is given by Figs. 6.8 – 6.11 as a function of the number of mobile nodes, the number of LFNs, subnet residence time, the number of

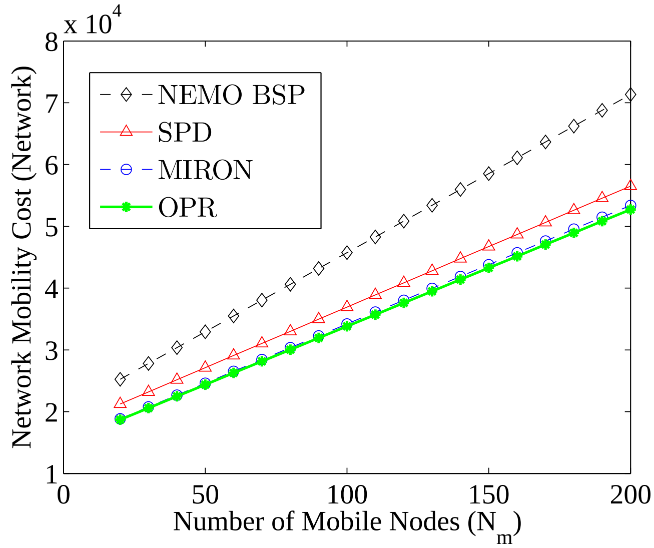

Figure 6.8: Network Mobility Cost on complete network vs. number of mobile nodes for the four schemes.

hops and the number of CNs, respectively. The cost of NEMO BSP is higher than the other schemes due to the higher packet delivery cost that results from multiple tunneling of all packets through the unoptimized route. Since only the first packets of sessions (in contrast to all packets) are tunneled through the unoptimized route, MIRON and OPR incurs the lowest cost.

6.3 Discussions on results

Analysis of the results shows that there is insignificant difference among the schemes as far as the cost incurred at the TLMR is concerned. However, results and the associated discussions also show the domination of the packet delivery cost incurred at the HA and the network due to the processing and the transmission requirements at the HA and the additional route between the AR and the HA. Thus, results suggest not to compromise the route with the signaling if costs incurred at the HA

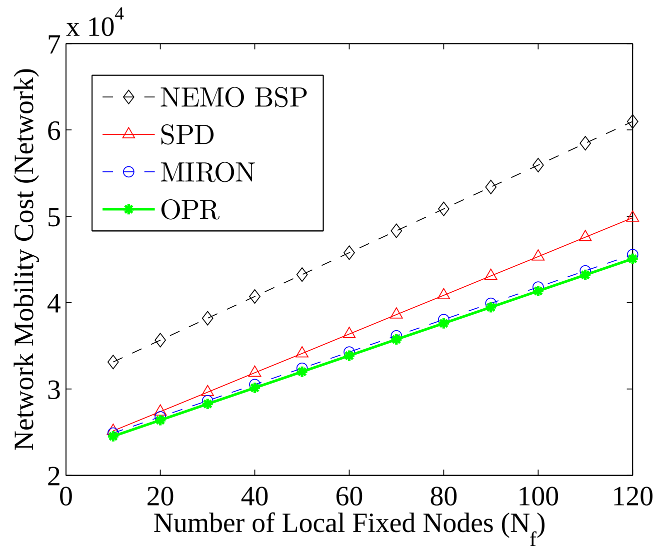

Figure 6.9: Network Mobility Cost on complete network vs. number of LFNs for the four schemes.

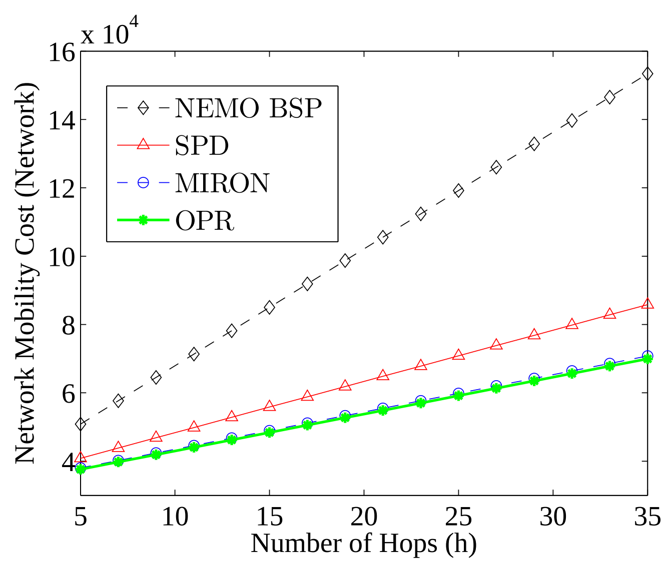

Figure 6.10: Network Mobility Cost on complete network vs. number of hops for the four schemes.

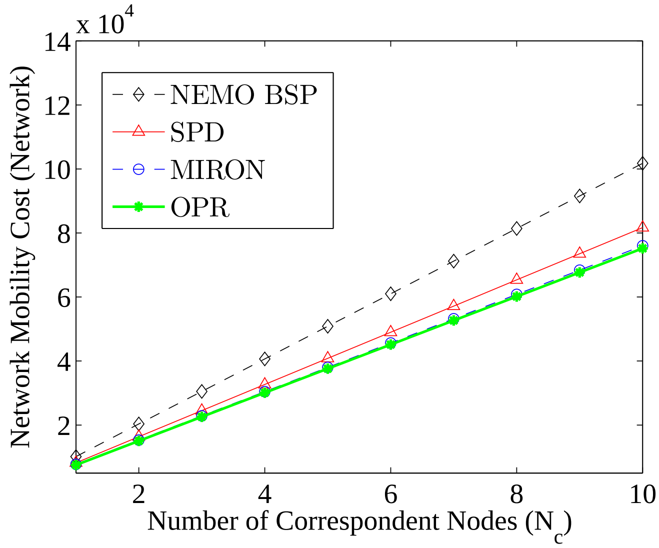

Figure 6.11: Network Mobility Cost on complete network vs. number of CNs for the four schemes.

and the network are to be minimized. However, performance of the schemes need to be considered along with the costs when choosing a scheme.

Signaling is one factor to be considered because it might affect the performance of the schemes when throughput is considered. OPR might be the best scheme because of its low signaling. However, OPR is incapable of optimizing the route when packets do not flow towards the mobile network. In MIRON, amount of signaling is the largest, and the procedure of obtaining CoAs might be a limiting factor when the nesting level is large. The cost computed here have to be traded off with these pros and cons of the schemes.

6.4 Summary

In this chapter, we have developed mathematical models to determine the network mobility costs on various mobility entities of NEMO BSP, and three representative prefix delegation-based NEMO route optimization schemes (SPD, MIRON and

OPR) in terms of network size, mobility rate, distance between mobility agents, and traffic rate. Results show that the effect of packet delivery cost dominates other cost components in the network mobility costs because this cost is incurred per data packet.

Thus, our results lead to an interesting conclusion which is opposite to the general intuition that complete route optimization requires less resources (less cost) than that required for partially-optimized route with the reduction in signaling.

Chapter 7

Improvement of prefix delegation-based schemes for intra mobile network communicaitons

Intra mobile network communications refer to the events of communications that occur between two MNNs in the same mobile network [11]. Evaluations of some schemes (not prefix delegation-based schemes) for such communications can be found in [11, 86]. The comparison of the classes presented in Sec. 2.4 shows the inability of the delegation-based schemes to optimize route for intra mobile network communications. Therefore, our aim is to extend and evaluate the prefix delegation-based schemes for the intra mobile network case. However, this evaluation also applies to the inter mobile network communication case.

The rest of the chapter is organized as follows. Inefficiencies and the extensions of the schemes are presented in Secs. 7.1 and 7.2, respectively. Simulation results are discussed in Sec. 7.3 followed by concluding remarks in Sec. 7.4.

7.1 Inefficiencies of prefix delegation-based schemes

We consider LFN-LFN communication because communication routes for VMNs/LMNs are optimized. In SPD, packets travel through the HAs of both LFNs resulting in a high end-to-end delay. In OPR, LFNs are transparent to mobility and therefore, will not recognize the OPR header used for route optimization. In MIRON, the BU sent to an LFN for optimizing the route will not be recognized resulting in failure of the notification of the CoA. Moreover, in OPR, failure to get the original address of the LFN will result in communication to cease. However, communication can continue in MIRON like in SPD. Thus, none of the schemes can optimize route for LFN-LFN communication.

7.2 Extension of the prefix delegation-based schemes

SPD does not optimize LFNs' route while MIRON and OPR does. We extend the route optimization procedure of OPR and MIRON for LFN-LFN communication, and explain below for LFN1-LFN2 (see Fig. 7.1) communication.

7.2.1 Extension for OPR (xOPR)

In OPR, the first packet, sent by LFN1 to LFN2, will reach MR4 through the HA_MR4. In extended OPR, MR4 will process (like CN) the OPR header in the packet to create binding entry that maps address (address of LFN1) in OPR header to the source address (translated address of LFN1) of the packet. When an outgoing packet (from LFN2 to LFN1) is received, MR4 search the binding entry to find the translated address of LFN1, puts the destination address of the packet into Routing Header Type 2 (RH2) header, and replaces the destination address with the translated address. Since the translated address is obtained from the foreign network's prefix, the packet reach LFN1 without traversing HAs.

7.2.2 Extension for MIRON (xMIRON)

When a packet (from LFN1 to LFN2) is received, MR2 puts its CoA in the source address field whose content is put into Home Address destination Option (HAO), and forwards the packet that will reach MR4 through HA_MR4. At reception of the packet, MR4 (LFN2's MR) adds the source address of the packet in the BU list, replaces the source address with address in HAO which is removed from the packet, and forwards the packet to LFN2. BU, sent from MR4 on behalf of LFN2, will reach MR2 that will create a binding entry that maps LFN2's address to CoA of MR4. MR2 will forward subsequent packets, from LFN1 to LFN2, by replacing destination address with the CoA from the binding entry along with putting the destination address into RH2. Since the destination of packets is the CoA of MR4, packets will be routed by the MRs (without going outside the mobile network) to MR4 which will forward the packets to LFN2 after replacing the destination address with the address in RH2 which is removed.

7.2.3 Performance analysis of the schemes

Since packets in SPD are routed through HAs, end-to-end delay increases with increasing distance between the mobile network and its HA. In xMIRON and xOPR, end-to-end delay is independent of the distance because packets are routed within the mobile network. Thus, TCP throughput, being inversely proportional to end-to-end delay, is different for the schemes when the mobile network is away from its HA.

Route optimization for LFNs affects handoff latency that affects throughput. In SPD, MR tunnels packets using its CoA and the address of HA. Since the address of HA is always available, packets, tunneled after MR obtains a new CoA, can reach the destination. Packets tunneled using old CoA are discarded due to ingress filtering. In xMIRON and xOPR, an MR searches a binding entry for the CoA which is, if found, put as the destination address. Otherwise, the packet is forwarded with its original destination address. Therefore, as long as the binding entry containing the old CoA is not updated or deleted after the handoff, packets are sent using the old CoA as destination address. These packets are dropped because the CoA is no longer in use. Thus, handoff latency of the extended schemes can be different than that of SPD. The difference can result significant variation in the throughput when frequency of handoff (i.e., speed of mobile network) increases.

7.3 Performance evaluation

Simulation environment, analysis of the results and a comparative discussion are presented in the following subsections.

7.3.1 Simulation Environment

Figure 7.1 shows the topology used in simulation. LFN1 is an FTP source over TCP whereas LFN2 is a TCP sink. The mobile network moved between ARs, placed in a horizontal line. Wireless links use IEEE 802.11b (11Mbps) whereas Ethernet (10Mbps) was used for mobile networks. Other values of parameters used in the simulation are presented in Table 7.1.

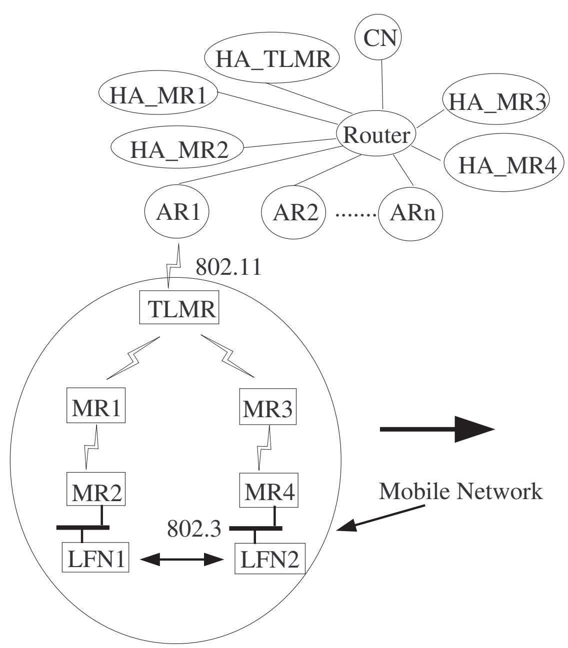

Figure 7.1: Topology used for simulation.

Table 7.1: Values of parameters used in the simulation.

Parameter	**Value**
Simulation time	360s
Wired link BW	10Mbps
Binding entry lifetime for SPD and xMIRON	10s
Binding entry lifetime for xOPR	5s

7.3.2 Results

We measured (at 95% confidence level) end-to-end delay, handoff latency, and throughput at different speeds and delays between the HA and the mobile network. Throughput was measured by the amount of data received at LFN2. Since LFN-LFN communication is not possible in OPR, and MIRON's performance is similar to SPD for LFN-LFN communication, we show results for SPD, xMIRON and xOPR.

7.3.2.1 End-to-end delay and throughput without handoff

End-to-end delay between LFN1 and LFN2 is shown in Fig. 7.2. With the increase of delay between Router and HA, end-to-end delay for SPD increases while that of the other schemes is unaffected. Since SPD does not optimize route, packets reach LFN2 through HAs causing the end-to-end delay to be dependent on the Router-HA link delay. Other schemes optimize route enabling packets to be routed to LFN2 without going through the HAs, and consequently, end-to-end delay is unaffected by the delay. End-to-end delay in SPD is higher because of the same reason. Since the TCP throughput is inversely proportional to end-to-end delay, the characteristics of the throughput, shown in Fig. 7.3, follows from that of end-to-end delay.

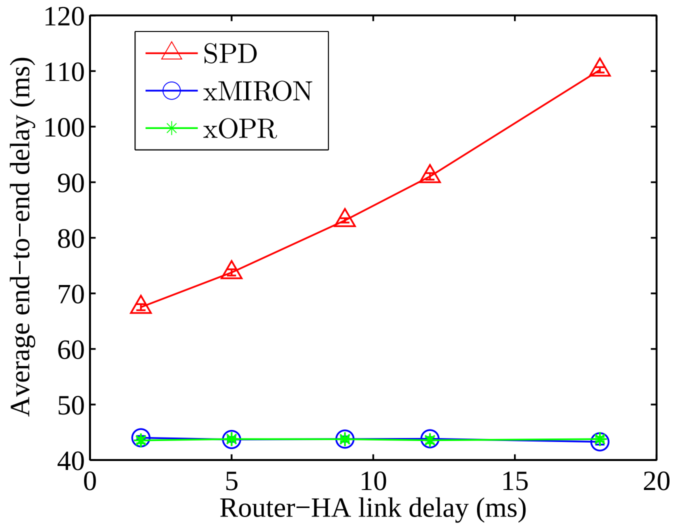

Figure 7.2: End-to-end delay between LFN1 and LFN2.

7.3.2.2 Handoff latency

Events constituting the handoff latency are presented in Fig. 7.4, and explained below.

SPD: Due to suspension of flow of packets during the handoff, the first packet after obtention of the CoA is sent when TCP reaches the next timeout. Therefore, handoff latency, $(t_e - t_s)$ = time to detect AR by TLMR since the last packet received at MR4 + time to propagate AR's prefix to MR4 + delay for TCP to reach the next timeout + delay for the packet to reach MR4 = $(t_d - t_s) + (t_c - t_d) + (t_o - t_c) + (t_e - t_o)$.

On the average, $(t_d - t_s)$ is constant with respect to nesting level with a maximum value of just over 2.3 seconds which is the expiration time for an AR's liveliness. $(t_c - t_d)$ and $(t_e - t_o)$ are proportional to the nesting level, and is in the order of milliseconds. The rest of the handoff latency is due to $(t_o - t_c)$ which is small if

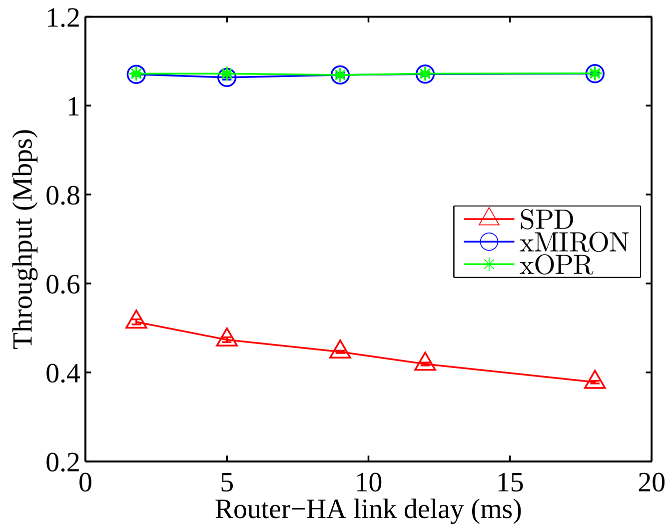

Figure 7.3: Throughput of different schemes without handoff.

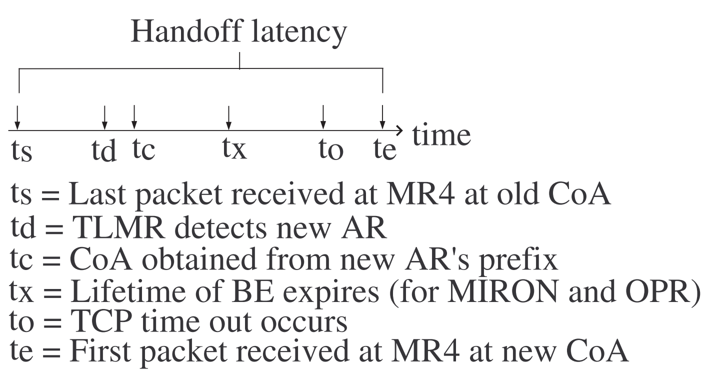

Figure 7.4: Events that occur during handoff.

$(t_c - t_s)$ is small, and can be exponentially large, otherwise. For SPD, small $(t_c - t_s)$ incurs a small handoff latency (Fig. 7.5).

xOPR: Packets sent during handoff cannot reach destination until binding entry is updated or deleted (see Sec. 7.2.3). Binding entry at MR2 is updated when a packet arrives from MR4 which is also awaiting a binding entry update resulting in a deadlock. When the binding entry-lifetime (5 seconds) expires, MR2 forwards packets without modifying the destination address (HoA of LFN2), and packets reach MR4 via HA_MR4 to break the deadlock. Since binding entry-lifetime is updated whenever a packet is received, $(t_x - t_s)$ is always 5 seconds, which is much greater than $(t_c - t_s)$. Therefore, handoff latency, $(t_e - t_s)$ = time to expire binding entry-lifetime + delay of TCP to reach the next timeout + delay for the packet to reach MR4 = $(t_x - t_s) + (t_o - t_x) + (t_e - t_o) = 5 + (t_o - t_x) + (t_e - t_o)$. Hence, handoff latency is more than 5 seconds (Fig. 7.5) which is larger than that of SPD.

xMIRON: Like xOPR, deadlock occurs because BU, sent to old CoA, cannot reach MR2 whereas MR4 cannot send a BU to new CoA of MR2 until it receives a packet from MR2 at the new CoA. Like xOPR, deadlock is broken when MR4 receives a packet via HA_MR4.

binding entry-lifetime is refreshed at reception of BU, received at or before t_s; therefore, $(t_x - t_s)$ can be between 0 to 10 seconds which is the binding entry-lifetime. Thus, when $(t_c - t_s) > (t_x - t_s)$, handoff latency, $(t_e - t_s)$ = Time to detect AR by TLMR since the last packet received at MR4 + time to propagate AR's prefix to MR4 + delay for TCP reaches the next timeout + delay for packets to reach MR4 = $(t_d - t_s) + (t_c - t_d) + (t_o - t_c) + (t_e - t_o)$. Otherwise, handoff latency, $(t_e - t_s)$ = Time to expire binding entry since reception of last packet + delay of TCP reaches the next timeout + delay for the packet to reach MR4 = $(t_x - t_s) + (t_o - t_x) + (t_e - t_o)$.

Since $(t_x - t_s)$ is uniformly distributed between zero and ten, average $(t_e - t_s)$ is expected to be around five. This is not true because large values of $(t_x - t_s)$ make $(t_o - t_x)$ exponentially large, and hence, the average handoff latency (Fig. 7.5) is much larger than five.

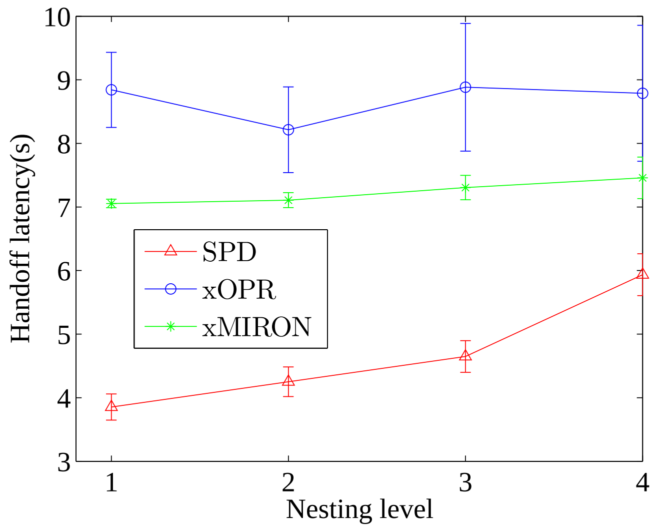

Figure 7.5: Handoff latency of the schemes.

7.3.2.3 Throughput vs speed

Figure 7.6 shows the throughput for different nesting levels. Throughput of SPD is low because of high end-to-end delay. At high nesting levels, throughput of xOPR and xMIRON is close to that of SPD due to increase in the number of wireless hops that start to dominate the effect of unoptimized route.

Throughput decreases with increasing speed due to increasing number of handoffs causing packet loss. For SPD and xMIRON, the rate of decrease is the smallest and the largest, respectively because of the smallest and the largest handoff latency. At high speeds, throughput loss due to handoffs dominates the loss due to high end-to-end delay resulting in throughput of xMIRON to fall below that of SPD.

Throughput of xOPR is close to that of SPD due to high handoff latency which is not high enough to bring the throughput below that of SPD even at high speeds.

7.3.2.4 Throughput vs mobile network's distance from HA

Figure 7.7 presents the throughput as a function of Router-HA link delay. Throughput of SPD decreases with increasing Router-HA link delay while that of xMIRON

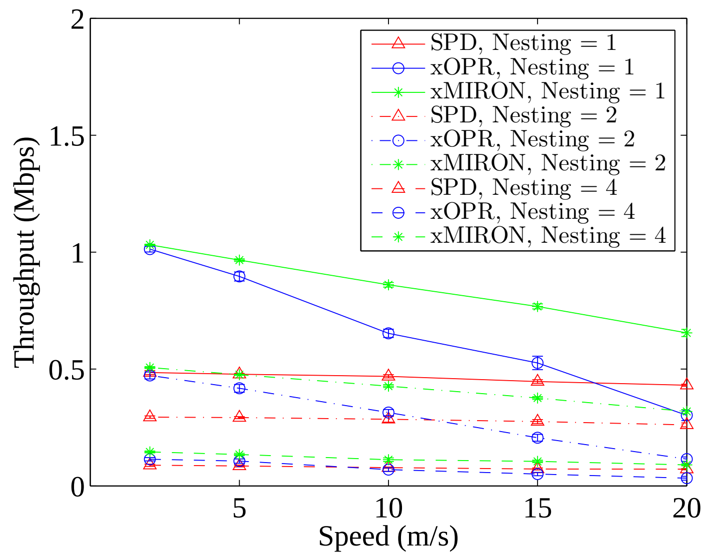

Figure 7.6: Throughput for various nesting level.

and xOPR are unaffected (Fig. 7.3). At low speeds, throughput of SPD is the smallest because of the high end-to-end delay. At high speeds, loss of throughput due to handoff latency dominates even at high Router-HA link delay resulting in the lowest throughput for xMIRON. xOPR has the highest throughput because of small end-to-end delay and handoff latency which is not much larger than that of SPD. At low speeds, xMIRON's throughput is a little less than that of xOPR because of the reason explained below. In xOPR, packets can be sent through the optimized route after the first packet is received after handoff. In xMIRON, after reception of the first packet following a handoff, packets cannot be sent through the optimized route until the BU is sent to the peer when the next event for sending BUs triggers.

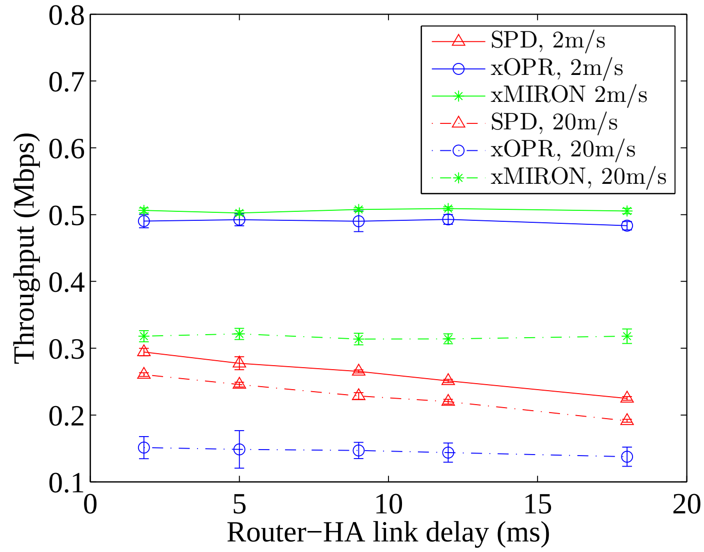

Figure 7.7: Throughput for two extreme speeds.

7.3.3 Comparative discussion

From the results presented in Sec. 7.3.2, xOPR appears to be the best; xOPR, however, needs packets to flow in both direction for route optimization. xMIRON performs better than SPD when nesting level is low except at very high speed. At high nesting levels, performance of the schemes are similar, and xMIRON performs the worst at speeds above 7.5 m/s.

Performance of xOPR and xMIRON degrades faster than that of SPD due to high handoff latency which can be lowered using small binding entry-lifetime. However, small lifetime increases signaling and processing at MRs. In xOPR, lifetime has to be set considering the interval of packet reception to avoid unnecessary expiration. If interval of unidirectional packet flow is high, lifetime has to be high resulting in high handoff latency. Binding entry-lifetimes can be set through BU.

xMIRON uses a feasible CoA obtention process whereas CoA obtention by prefix delegation through router advertisement in xOPR and SPD might not be easily applicable due to accounting, authentication and security requirements [82]. Moreover, xMIRON and xOPR requires additional processing and memory.

7.4 Summary

In this chapter, we evaluated the performance of prefix delegation-based schemes for intra mobile network communication. We measured end-to-end delay, handoff latency and throughput under various speeds at different delays from the HA. Results show that xOPR performs the best, limited due to its inability to optimize route when packets do not flow in both directions. xMIRON, having a feasible solution for CoA obtention, performs better at low speeds. SPD, with the advantage of requiring less resources, is a good choice at high speeds. In addition, the performance loss in xMIRON due to speed dominates over performance loss in SPD due to distance from HA. Overall, xOPR and xMIRON are preferable to SPD at low speed, and vice versa.

Chapter 8

NEMO for satellite networks

A satellite network is a network of satellites and stations on the ground. Satellites connect to each other through Inter Satellite Links (ISLs), and to ground stations through Ground to Satellite Links (GSLs). Depending on the relative movement with respect to the Earth, and on the orbital distance, satellites can be of several types - Geo Stationary Satellites (GEO), Low Earth Orbit Satellite (LEO) and Medium Earth Orbit Satellites (MEO). A network of satellites can involve all types of satellites. Examples of networks involving LEO satellites are constellations of LEO satellites, such as Iridium [87], Globalstar [88], etc.

One of the most important use of satellites is to collect Earth observing data that are used to monitor flood, wildfire, volcanoes and cryosphere events [38]. To monitor the commercial aircrafts' safety, the transfer of real-time data from aircrafts can be another significant job of satellites in future [89]. To transfer such data to the end users through the Internet [90] or to the IP-enabled end users, future satellites will contain multiple IP-enabled devices. At present, IP is being used to transfer imaging data, collected by satellites in the Disaster Monitoring Constellation, to the ground to aid in disaster area relief operations [39].

Low latency, and better coverage of the Earth make constellations of LEO satellites preferable to other types of satellites for the Earth observation. Satellites in a constellation are uniformly placed in several hypothetical orbital planes that are dispersed at equal distance and concentric with the surface of the Earth. A satellite is connected to neighboring satellites through ISLs to form a grid like connectivity. Usually satellites on neighboring planes orbit in the same direction except at the seam position of the constellation where satellites in two neighboring planes (first and last planes) orbits in opposite directions. One of the characteristics of satellite networks is the handoff of satellites with respect to ground stations and neighboring satellites in different planes because of the movement and the limitation of ISLs' connectivity over polar regions. That characteristic and efforts to integrate satellite networks with the Internet demand the use of the IP-based routing and mobility management protocols for satellite networks.

The transfer of data from IP-enabled devices in satellites to the Internet requires the following:

- The routing of packets from the device to a ground station requiring the handling of dynamic topology of satellite networks due to handoffs of space links.
- The routing of packets from the ground station to the Internet host. The Internet routing can handle this.

There have been a number of research efforts to route packets from one satellite to another through multiple satellites. Korcak et al. [91] presented a priority metric based on the amount of packets successfully sent, drop count and queue size, to choose a route from multiple available routes between a source and a destination. Routes are assumed to be found using dynamic virtual topology routing [92]. The later routing protocol relies on satellites predictable movement to statically determine multiple topologies valid for various time duration at various points of the orbiting time. A set of routes are chosen from those topologies. However, a huge amount of memory is required to store all those routes.

Based on fixed logical positions of satellites, Ekici et al. [93] proposed a control overhead free routing in satellite networks. The logical position is determined based on the position of a satellite with respect to its neighboring satellites, and is used to uniquely identify the satellite. Extension of [93] is proposed by Papapetrou et al. [94] for inclinations other than 90 degrees and nonzero phase shift along with the load balancing. Routing decisions are made per packet basis at each satellite. Mapping of IP addresses to logical positions will be required to route packets using the IP. Donner et al. [95] proposed the use of Multi Path Label Switching [96] where ground stations are label edge routers and request label-switched paths when handoffs occur. Chen et al. [97] proposed a routing scheme for IP over SATATM network with the focus on QoS parameters. It takes the advantage of the predictability in the movements of the satellites to achieve the QoS.

For better integration with the Internet, satellite networks could be considered as a mobile network with respect to the Internet, and IP-based mobility management protocols can be used to manage the mobility of IP-enabled devices onboard satellites [98]. There have been a few works on the use of IP and IP-based mobility protocols in satellite networks. National Aeronautics and Space Administration (NASA) has been experimenting with the use of IP in the satellite networks [35]. Application of the MIP to satellite networks has been proposed by Israel et al. [36] where an onboard device is considered as a mobile host with mobility management agents residing in terrestrial networks. However, these host mobility protocols are inefficient for handling the mobility of multiple IP-enabled devices onboard satellites. Moreover, another routing protocol is required along with these mobility management protocols to transfer data to the ground when the satellite is not in direct contact with a ground station.

NEMO can be used to efficiently handle the mobility of multiple devices onboard satellites. Leung et al. [33] presented the application of the IPv4-based mobile network within a single satellite. In the IPv4-based mobile network, a router is used inside a satellite to route packets sent from multiple devices onboard the satellite. Based on concepts similar to NEMO, Shi et al. [34] proposed a satellite constellation network architecture that allows the communication of satellite-hosts with hosts in the Internet through satellite mobile routers while the relative movement of satellites are transparent to the Internet. However, Shi et al. focused on the fast handoff, and their proposed architecture have to rely on another routing scheme for the transfer of data from the satellite to the Internet when a satellite is not in contact with a ground station.

In this chapter, we propose a nested NEMO architecture in the satellite constellation network to handle the mobility of multiple IP-enabled devices onboard satellites and enable continuous transfer of data from satellites to the Internet without relying on any other protocol. Our work differs from earlier works as we use extensions of IPv6-based NEMO BSP for both satellite-satellite and satellite-Internet routing. Architecture of NEMO in satellite networks and the extension of NEMO BSP are presented in Sec. 8.1. Simulation environment used for evaluations of our work is presented in Sec. 8.2. Results from evaluations and a comparisons with an optimal algorithm are presented in Secs. 8.3 and 8.4. Finally, concluding remarks are put forth in Sec. 8.5.

8.1 NEMO for satellite networks

In the following subsections, we present the architecture for NEMO in satellite networks, problems of using NEMO BSP, and its extensions for the proposed architecture.

8.1.1 Architecture

Figure 8.1 shows the architecture for NEMO in satellite networks. Each satellite contains a mobile network connecting the onboard IP-enabled devices, such as LFN1 and LFN2, to an MR onboard. An MR may form a nested mobile network by connecting to another MR through the ISL. TLMRs are directly connected to ARs through GSLs. ARs (e.g., AR1 and AR2) are co-located with ground stations. Thus, multiple isolated nested mobile networks overlay the physical network of satellites. The HAs for the mobile networks reside in the Internet.

The architecture can be extended to support the connectivity for the remote hosts and mobile networks on the ground. In the extended case, a remote host or multiple remote hosts on the ground can become a VMN or a nested mobile network under the mobile network onboard a satellite. When the satellite containing the mobile network moves out of the reach of the remote host or hosts, they can become a VMN or a nested mobile network under another mobile network to maintain a continuous connectivity. Similarly, mobile networks onboard aeroplanes can be considered nested mobile networks under a mobile network onboard a satellite within the range.

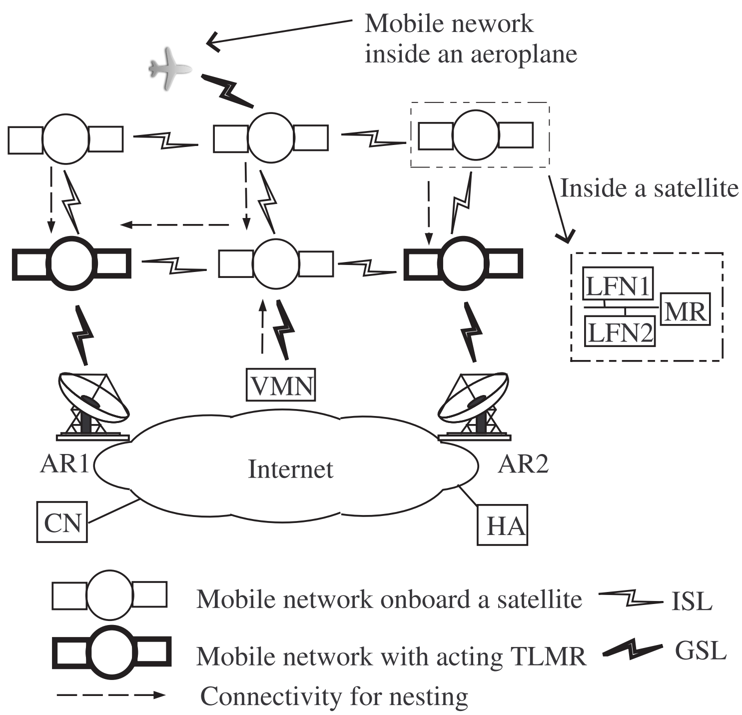

Figure 8.1: Architecture for NEMO in satellite networks.

8.1.2 Limitations of NEMO BSP for satellite networks

In the architecture presented in Sec. 8.1.1, an MR has physical connections to multiple neighboring MRs. Since no method for selecting an MR is specified, NEMO BSP will encounter the following problems:

- **Longer route**: Randomly choosing an MR for nesting might lead to a route having more number of ISLs than other routes that could be used by choosing other MRs. Such a route will be inefficient in terms of the end-to-end delay and the tunneling overhead.

- **Overloaded TLMR**: Routes through different TLMRs will be available for an MR. Randomly choosing an MR for nesting might have some TLMRs overloaded while leaving the others underloaded.

- **Routing loop**: Since an MR might be connected to another MR through multiple routes, a routing loop might be created if care is not taken while choosing an upper level MR for nesting.

Therefore, NEMO BSP can be inefficient for satellite networks, and requires extension.

8.1.3 Extened NEMO BSP for satellite networks

We propose extensions for NEMO BSP to make it efficient for satellite networks. The extensions are presented in the following subsections.

8.1.3.1 Basic principles

We use the following basic principles for the extensions of NEMO BSP for satellite networks:

1. **Achieve minimum nesting level**: Since increasing the number of wireless hops (in this case ISLs), degrades the performance [99], the highest preference

is given to achieve the minimum nesting level. In addition, the minimum level keeps the tunneling overhead and the delay due to the tunneling at the minimum.

2. **Balance load of TLMRs when levels are equal**: This principle is for balancing the load among TLMRs by choosing an MR for the nesting from the neighboring MRs that yield the same level. We use the data rate through a TLMR as its load. This principle reduces the load of a TLMR when an alternative TLMR which yields the same level is available. Note that our load balancing is not based on distribution of traffic through multiple upper level MRs. Rather, it is based on selecting a single upper level MR to send all traffic.

3. **Relaxation of the minimum level constraint when TLMRs get overloaded**: Overloading a TLMR will cause queueing delay and drops. Therefore, an MR is allowed to switch from an overloaded TLMR to an underloaded one if the level of the MR does not exceed the minimum achievable level by a threshold value. We consider overloading of TLMRs because all data from a mobile network exits through them.

4. **Avoid handing off to MRs leading to the current TLMR**: This principle is used to avoid unnecessary handoffs of MRs. If a handoff yields a route that goes through an MR's current TLMR, the handoff does not reduce that TLMR's load. Moreover, due to grid like architecture of satellite networks, such handoff does not reduce the level.

In the extended NEMO BSP, the principles are achieved by having an MR making the decision to handoff to a suitable upper level MR or AR. The changes made to NEMO BSP to execute such an handoff are presented next.

8.1.3.2 Information required for handoff decision making

Information, required to select a suitable upper level MR for handoff, are described below.

Level: The nesting level of the potential upper level MR is required to achieve the minimum level while choosing an MR for the handoff. TLMRs find their level as zero when they attach to an AR that identifies itself to MRs. Other MRs find their level by adding one to the level of their respective upper level MRs.

Load of TLMRs: The load of TLMRs of neighboring MRs are required for load balancing. Following can be used as the metric for the load:

- Data rate through the TLMR: An MR, connected to a TLMR having the minimum data rate through it, can be chosen as the upper level MR. Data rate can be measured at TLMRs by monitoring the incoming packets.

- Number of MRs connected through a TLMR: An MR, connected to a TLMR with the minimum number of MRs, can be chosen as the upper level MR. It is not possible for a TLMR to find the number of connected MRs by just looking at the packets going through it. Explicit messages are required for this requiring additional resources, such as processing power and bandwidth. Moreover, time required to update all MRs about any change of the value of this metric is more than that required for the other metric.

Considering the advantages and the disadvantages of using the two metrics discussed above, we prefer to use the first metric i.e., data rate through TLMRs. Moreover, the data rate will provide better load information when the sending rate of MRs is not uniform.

Downlink bandwidth of TLMRs: Downlink bandwidth of a TLMR is required to know whether it is overloaded or not. In this paper, we assume that each TLMR knows its downlink bandwidth.

Current TLMR's ID: Updated information, used to make handoff decision, may not reach an MR from neighboring MRs, which are under the same TLMR, at the same time instant resulting in inconsistent information. Such information will trigger a handoff in an effort to achieve a level lower than the current one, or to balance the load leading to another upper level MR under the same TLMR and with no change in the level or load of the TLMR. Such a handoff is unnecessary, and can be prevented by checking the current TLMR's ID. TLMR's IP address can be used for this purpose.

8.1.3.3 Handoff decision making

To select an upper level MR/AR for the handoff according to the principles presented in Sec. 8.1.3.1, an MR needs to receive, update and evaluate the information discussed in Sec. 8.1.3.2. Sending, updating and evaluating those information are presented in the following paragraphs.

Sending information to MRs: In NEMO, MRs/ARs express their availability and prefixes to neighboring MRs through periodic router advertisements. We propose the use of router advertisements to disseminate the information mentioned above. An MR includes its level, and its TLMR's load and IP address when the router advertisements are sent. If the MR is acting as a TLMR, it includes its own load. TLMR's load is included otherwise.

Updating information: Each MR keeps track of its level and load, and TLMR's IP address and load. Each MR also maintains a list of neighboring MRs along with their levels, and TLMRs' load, IP addresses and downlink bandwidth. These information are updated whenever a router advertisement is received.

Evaluating information to avoid loop and no connectivity to the Internet: If the MR receiving router advertisements is the upper level MR for MRs sending advertisements, advertisements are ignored to avoid loops. Such loops can occur when the receiving MR looses connectivity with its upper level AR or MR. Also, advertisements announcing no connectivity to the Internet are ignored.

Evaluating information to achieve the minimum level: An MR performs such an evaluation when its level increases or it finds another MR's level decreasing below its current level. While handing off to an MR yielding a level lower than the current level, a check is performed not to overload the TLMR if relaxation of the lowest level is allowed. Also, load balancing at the new level is considered.

Evaluating information to relieve an overloaded TLMR: When a router advertisement is received from a neighboring MR, following conditions are checked before handing off to the MR:

- Current TLMR is overloaded
- Handing off to the advertising MR will not overload its TLMR
- Handoff will yield a level less or equal to the minimum level plus the threshold
- Current TLMR and the TLMR of the advertising MR are different

If all conditions are met, a handoff can be executed. However, handing off immediately after the reception of the router advertisement might result in oscillations (i.e., handing off back and forth between the two upper level MRs).

The oscillation happens when two or more MRs hand off simultaneously to the same advertising MR. Each MR finds the TLMR of the advertising MR underloaded after adding its own load, and therefore, decides to handoff. However, the MRs are unaware of the load added to the load of the TLMR due to the handoff of other MRs. Thus, handoffs of all the MRs might overload the TLMR. Similarly, the MRs might again handoff to get under its previous TLMR in an effort to handoff to an underloaded TLMR. Thus, the check performed not to overload a potential TLMR can not prevent oscillations because of the propagation delay of the updated information to reach the MRs.

To prevent oscillations while selecting a gateway (in NEMO a TLMR) for load balancing in ad hoc networks, several techniques have been proposed in the literature. Jungmin et al. [100] propose nodes to wait a specified amount of time before switching to a gateway. Hoffman et al. [101] suggest switching to a gateway only if it has been used for a specified period of time, and after switching, its load have to be less than the current gateway load by a threshold. However, these techniques fail because similar situation leading to oscillations will occur after the specified period of time.

To prevent oscillations, the synchronous handoffs of MRs need to be prevented. We use the traditional technique of reducing the probability of synchronous actions. When an MR finds that a handoff can relieve its TLMR, it waits for a random time period before executing the handoff. After the waiting period, the handoff is executed if the conditions are met. Thus, some MRs handing off late might have enough time to receive the updated load information of the potential TLMR, and decide against the handoff.

Evaluating information for the load balancing: When a router advertisement is received from a neighboring MR, the following conditions are checked before handing off to the MR:

- whether the level after the handoff will remain the same,

- whether the load of the current TLMR is larger than the TLMR of the advertising MR by a threshold amount,
- after handoff, whether the load of the TLMR of the advertising MR will be smaller than the load of current TLMR by the threshold amount, and
- whether the current TLMR and the TLMR of the advertising MR are different

If all conditions are met, a handoff can be executed following the procedure described in Sec. 8.1.3.3.

8.2 Simulation environment

Figure 8.2 and Table 8.1 presents the topology and the parameter values used for our simulation. We use an iridium constellation where each mobile network onboard a satellite has a different HA on the ground. Nine ground stations and co-located ARs are placed on the ground with 120 degrees separation from each other according to their latitudinal and longitudinal positions. Downlink/uplink capacities are set to 8.134/0.0384Mbps as is currently being used or expected to be used for UK-DMC satellites [39]. ARs, HAs and the CN are connected to a router, R, through wired links. Considering HAs will be located close to the core Internet, we set the R-HA link delays to 1ms. LFNs and the MR are connected by Ethernet (IEEE 802.3 standard) to form a mobile network onboard. LFNs onboard satellites send data to the CN on the ground using a space-friendly transfer protocol called Saratoga [102]. Rationale behind the use of Saratoga was described in our previous work [103]. All Saratoga sources (one in each LFN) send data at the same rate. We refer to the sum of the sending rate of all sources as the aggregate load.

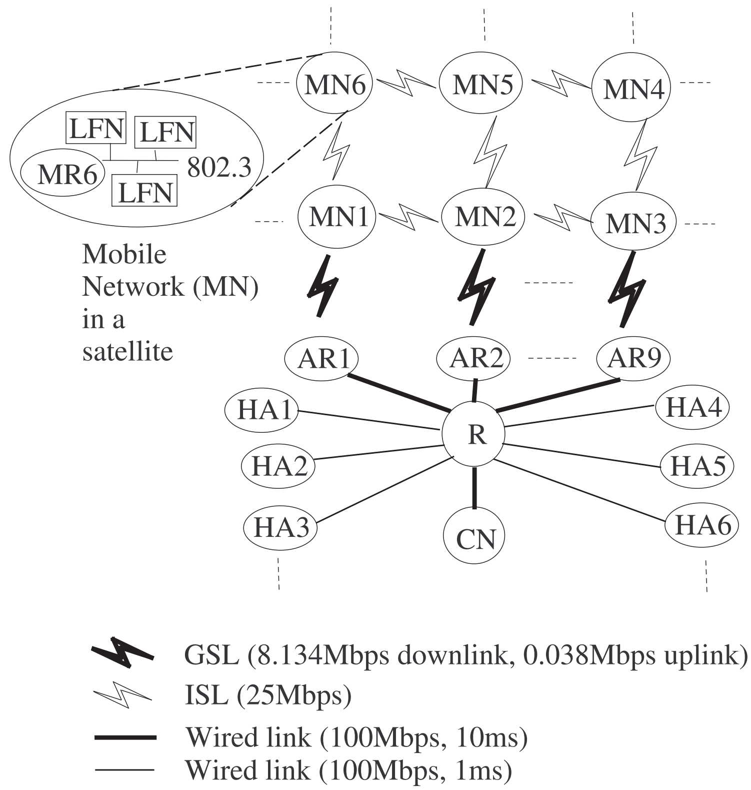

Figure 8.2: Topology used for the simulation.

8.3 Results

To evaluate the performance, we measured the throughput as a function of the time, the incoming data rate at TLMRs, the end-to-end delay, drops at TLMR's queue and overheads. Results are presented in the following subsections.

8.3.1 Throughput as a function of time

Throughput is measured, at one second interval, as the amount of data received at the CN. The objective is to show the continuity of connections at the upper layer despite the movement of the satellites. Figure 8.3 shows the throughput, measured from the data received from all LFNs, as a function of the time. The fall

Table 8.1: Values of parameters used in the simulation.

Number of ground stations	9
Altitude	780km
Orbital inclination	86.4
Elevation mask	8.2
GSL's downlink BW	8.134Mbps
GSL's uplink BW	0.0384Mbps
ISL BW	25Mbps
Ethernet BW	100Mbps
Wired link BW	100Mbps
R-AR and R-CN wired delay	10ms
R-HA wired delay	1ms
Number of cross seam ISLs	0
Queue limit at MRs	200 packets
Number of LFNs per satellite	5
Router advertisement interval	3s
BU interval	10s
Simulation time	1000s

of throughput occurs due to the handoff resulting from the loss of physical links with an AR, or an MR at polar regions. The throughput does not fall to zero because all MRs do not lose connectivity with the CN, simultaneously. Note that the thick line of throughput results from oscillations (variation) of per second throughput because of the variation in the number of packets sent by each LFN in each second to maintain the average data rate. This happens because we specify the rate in mega bits per second which is converted to packets per second.

To better observe the continuity of connections, we present the throughput measured from data received from a single LFN in Fig. 8.4. Vertical lines in Fig. 8.4 show the handoffs due to the loss of physical links. The handoff latency due to the loss of physical links with ARs or MRs is longer than the handoff latency due to other reasons (e.g., load balancing, achieving a lower level). The longer handoff latency is due to the movement detection time which is in the worst case 3 seconds,

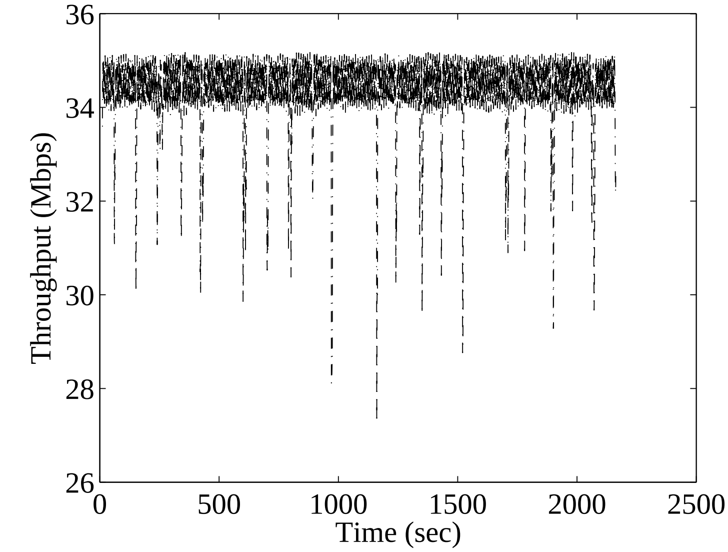

Figure 8.3: Throughput measured from the data received at the CN from all LFNs when aggregate load is 35Mbps.

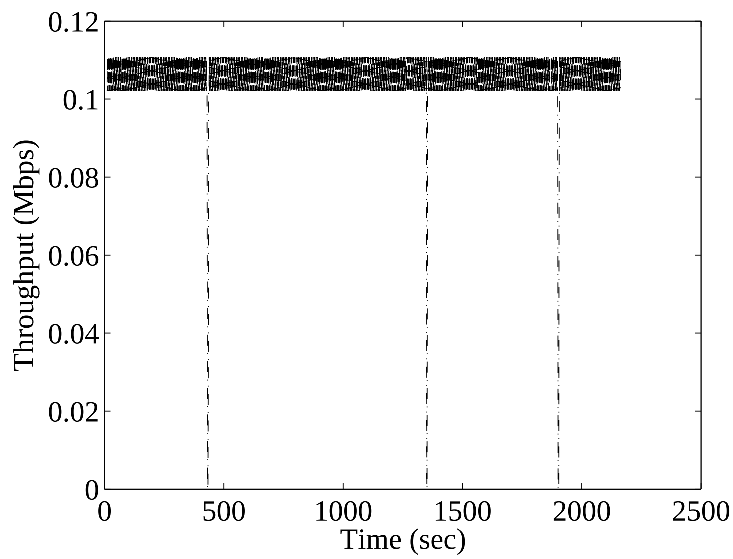

Figure 8.4: Throughput measured from the data received at the CN from one LFN when aggregate load is 35Mbps.

and can be reduced using lower-layer triggers to initiate handoffs. The handoff latency due to the load balancing or achieving a lower level is in the order of milliseconds because of the registration delay only.

8.3.2 Standard deviation of load on TLMRs from the uniform load

The load on a TLMR indicates the amount of incoming data which is to be forwarded to the attached AR. To find the effectiveness of the load balancing, we measure the standard deviation of the load on TLMRs from the uniform load as a function of the aggregate applied load. The average load on TLMRs is measured for each GSL which connects TLMRs to an AR, and expressed as the percentage of the aggregate applied load. The standard deviation of the percentage load from the uniform percentage load (i.e., aggregate load divided equally among the TLMRs) is shown in Fig. 8.5.

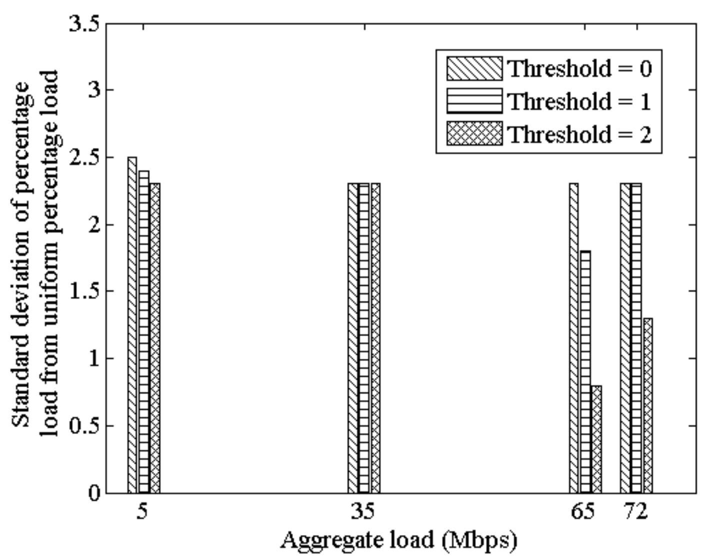

Figure 8.5: Standard deviation of load on TLMRs from the uniform load.

When the aggregate applied load is 5Mbps or 35Mbps which is much smaller than the total capacity (just over 72Mbps) of GSLs connecting TLMRs to ARs, no load balancing is performed. This is because all TLMRs' load are within the threshold limit, and the differences of the load between pairs of potential TLMRs do not exceed the threshold limit. Both threshold limits have to be exceeded to trigger the load balancing event. Therefore, the standard deviation remains the same for various threshold values of the relaxation of the level constraint.

When the aggregate applied load is close to the capacity of GSLs, the increase of the relaxation threshold value reduces the standard deviation of the load. Since at this applied load, some TLMRs get overloaded, relaxing the level constraint allows MRs to switch from overloaded TLMRs, yielding the minimum level, to the underloaded ones yielding higher levels. Thus, the load of TLMRs becomes similar to reduce the standard deviation. An increase of the relaxation threshold allows an increase of the number of TLMRs that switch to underloaded TLMRs, and therefore, decreases the nonuniformity of the load among TLMRs. However, as the aggregate load increases (e.g., 72Mbps in Fig. 8.5), the number of overloaded TLMRs increases. Therefore, the number of underloaded TLMRs yielding levels within the reach of small threshold values (e.g., 1) decreases, and so does the scope to reduce the standard deviation. The average load, as the percentage of the aggregate load, on TLMRs for each AR can be found in Table 8.2 for various aggregate load and level relaxation threshold. It shows that the number of overloaded (when load $> 11.3\%$) TLMRs at load 72Mbps is larger than that of overloaded (when load $> 12.5\%$) TLMRs at load 65Mbps.

8.3.3 End-to-end delay

To find the effects of achieving the minimum level and relaxing the level constraint, we measure the end-to-end delay as the difference between the time of receiving a packet at the CN and the time of sending the packet from the LFN. We compute

Table 8.2: Average load, as percentage of aggregate load, on TLMRs connected to a particular AR.

Aggregate load	**5Mbps**			**35Mbps**			**65Mbps**			**72Mbps**		
	Threshold = 0 (%)	=1 (%)	=2 (%)	=0 (%)	=1 (%)	=2 (%)	=0 (%)	=1 (%)	=2 (%)	=0 (%)	=1 (%)	=2 (%)
AR # = 1	12.8	12.7	12.8	13.1	13.0	13.0	13.9	13.0	12.0	13.0	12.8	13.5
= 2	14.8	14.7	14.6	14.8	14.4	14.4	14.3	13.8	12.7	14.8	14.6	13.6
= 3	15.8	15.5	15.5	15.3	15.6	15.6	15.5	14.5	13.1	15.4	15.5	13.4
= 4	13.3	13.2	13.1	13.0	12.6	12.6	12.2	12.7	12.0	12.8	13.1	11.7
= 5	12.3	12.4	12.4	12.1	12.3	12.3	12.1	12.7	13.0	12.2	12.4	11.9
= 6	9.4	9.2	9.2	9.5	9.6	9.6	9.1	10.0	11.1	9.8	9.2	11.7
= 7	9.1	9.3	9.0	9.4	9.6	9.6	9.4	9.5	11.2	9.4	9.0	10.5
= 8	9.4	9.5	10.0	9.3	9.6	9.6	10.1	10.4	11.5	9.4	10.0	10.8
= 9	10.3	10.6	10.5	10.5	10.4	10.4	10.4	10.6	11.3	10.4	10.5	10.6

the average of the end-to-end delays for all packets for which feedback-packets are received. The results are shown in Fig. 8.6. The end-to-end delay increases with the load due to the queuing delay in overloaded TLMRs. Relaxing the level constraint reduces the end-to-end delay when the load (65Mbps) is below the total capacity of the GSLs because the decrease in the queuing delay is more than the increase in the delay due to the increased tunneling and hops. At a load (72Mbps) almost equal to the capacity, relaxing the level constraint increases the end-to-end delay due to the insufficient or no decrease of the queuing delay compared to the increase of tunneling and hop delay. And, this happens because almost all TLMRs are overloaded leaving a little room for reducing the queuing delay. However, changes in the end-to-end delay resulting from the level relaxation are very insignificant.

8.3.4 Receive-ratio and drops at TLMR's queue

The receive-ratio is measured as the ratio of the number of bytes received at the CN to that of bytes sent from LFNs, and is shown in Fig. 8.7. The objective of the measurement is to observe the effects of relaxing level constraints. The receive-ratio decreases with the increase of the load due to the increase of the drop at the queue of overloaded TLMRs. Drops, as the percentage of the number of incoming packets at TLMR's queue, are shown in Fig. 8.8. Relaxing the level constraints reduces

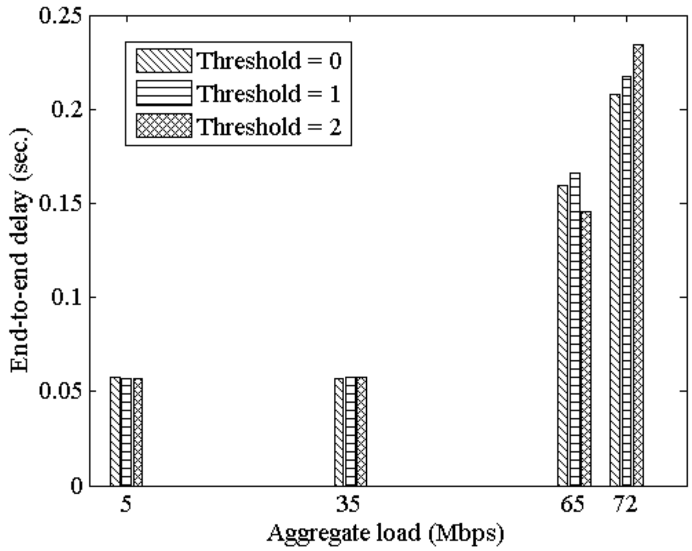

Figure 8.6: Average end-to-end delay from LFNs to the CN.

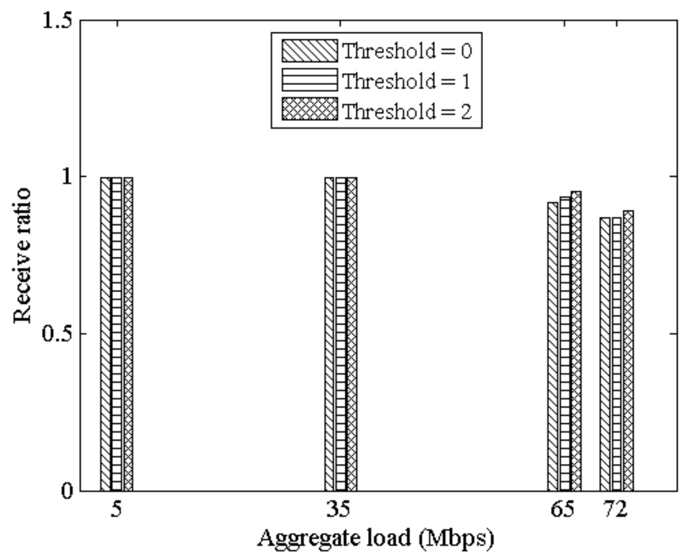

Figure 8.7: Receive ratio for data packets at the CN.

drops by transferring the load from overloaded TLMRs to underloaded ones, and thus, increases the receive ratio. As explained for the previous results, at high load (e.g., 72Mbps), the increase in the number of overloaded TLMRs causes the small level relaxation threshold values (e.g., 1) to fail to increase the receive ratio (or decrease drops).

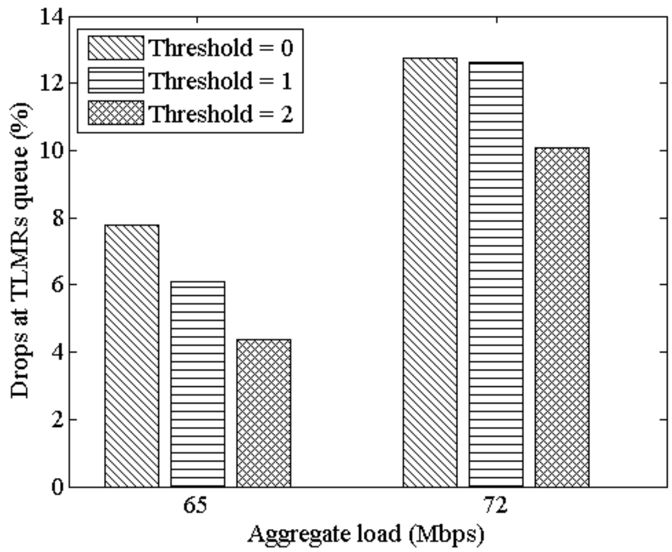

Figure 8.8: Drops at TLMR's queue as a percentage of the number of packets sent. Drops at 5Mbps and 35Mbps aggregate loads are 0 and not shown.

8.3.5 Overhead

For the extended NEMO BSP presented in Sec. 8.1, there could be following two types of overhead:

- Signaling overhead - This overhead results from periodic router advertisements and BUs (ignoring the small overhead due to solicitations sent by MRs).

- Tunneling overhead - This is the overhead due to the transmission of additional bytes of the tunnel header, and varies with the load i.e., data sending rate for a fixed packet size.

We measure signaling overheads as the number of mega bits transmitted per second, and express as the percentage of the aggregate load. Figures. 8.9 and 8.10 show the overhead for the router advertisements and BUs, respectively. The overhead due to router advertisements is higher than that due to BUs because router advertisements are sent every 3 seconds while BUs are sent every 10 seconds (the first five BUs is sent every 1 second). However, the router advertisements consume less bandwidth because it lives only one hop. Finally, both overheads are very insignificant if the aggregate load is high.

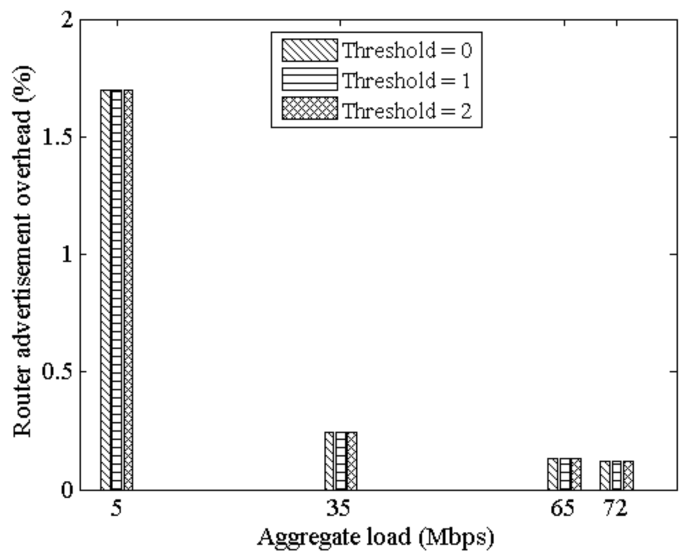

Figure 8.9: Overhead due to router advertisements.

We measure the tunneling overhead as the average additional bytes transmitted for the tunnel header per second per hop, and present in Fig. 8.11. To measure, we add the amount of tunnel headers at each hop traveled by a packet, and divide it by the number of hops to get the average amount of tunnel headers per hop. We

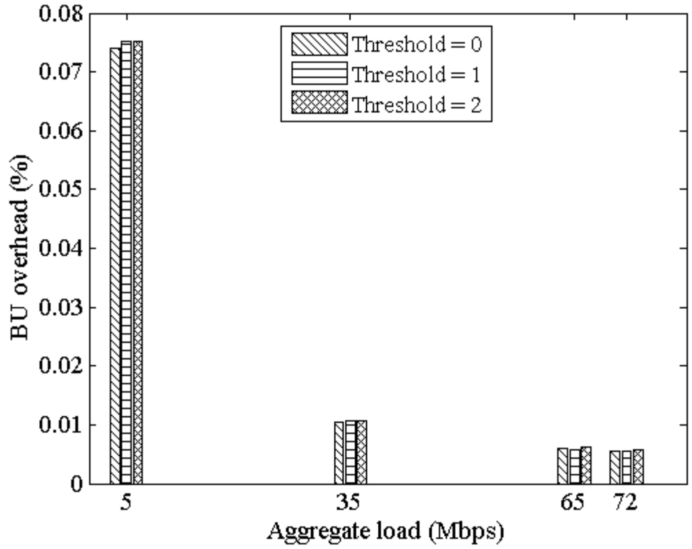

Figure 8.10: Overhead due to BUs.

sum up the average for all packets and then divide by the time to get the per second overhead due to the tunnel header. At load 65 Mbps and 72 Mbps, the overhead increases with the increase of the level relaxation threshold because of the increase of the average level of MRs. The average level increases because some MRs move to levels higher than they would be in at lower threshold values, to relieve overloaded TLMRs. The increase of the level increases the number of tunnels used to deliver packets. At 5 Mbps and 35 Mbps, no TLMR is overloaded, and therefore, no increase of the tunneling overhead as a function of the level relaxation threshold.

8.4 Comparison with an optimal algorithm

In this section, we compare the extended NEMO BSP with an optimal algorithm. For the comparison, we consider only the routing within a satellite constellation including ground stations. Therefore, we compare the extended NEMO BSP with the

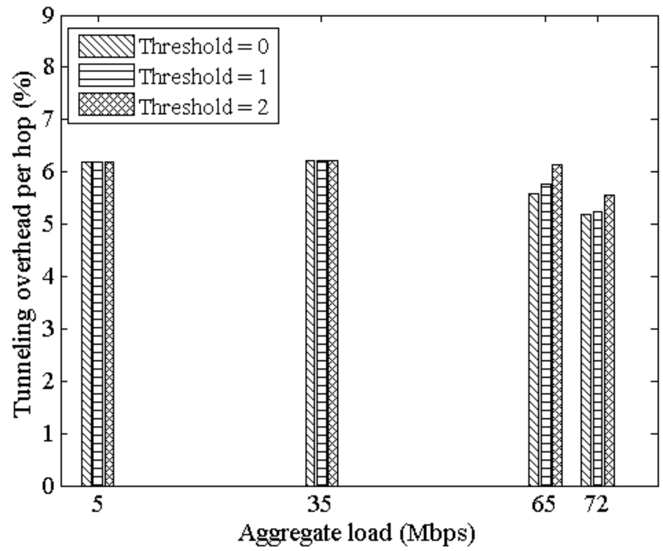

Figure 8.11: Overhead due to the tunneling.

distributed Bellmanford's algorithm which finds the optimal route within a constellation. If the relaxation threshold is zero, the extended NEMO BSP also uses the optimal route within the constellation.

8.4.1 Computational effort when a GSL's connectivity changes

We measure the computational effort to maintain routes despite the changes in the links' connectivity. The distributed Bellmanford's algorithm finds an optimal route between any two satellites (or ground stations) without balancing the load. Therefore, for a fair comparison, we consider the computational effort required by the extended NEMO BSP to find routes only. Also, both algorithms can do computations to re-compute routes when triggered by the lower layer. Therefore, we compute the computational effort required when a link goes on or off. In particular, we consider a GSL's connectivity because the number of MRs/routers (onboard satellites), which are required to do computations due to a change in a GSL's connectivity, will be much more than that are required due to a change in an ISL's connectivity. Also, GSLs' connectivity changes more frequently than ISLs'.

8.4.1.1 Maximum number of times the router advertisements are processed or the algorithm is run when a GSL's connectivity changes

We assume that ground stations are placed in such a way that on the average, an equal number satellites are using the GSL connected to each ground station. Let,

$N_{sat} =$ the number of satellites in the constellation,

$N_g =$ the number of ground stations,

$N_a =$ the maximum number of MRs that are routing through each GSL,

$N_n =$ the number of neighbors of an MR or router onboard a satellite,

$n_{rap} =$ the maximum number of router advertisement processing required when a GSL's connectivity changes.

Then, $N_a = \lceil N_{sat}/N_g \rceil$. In fact in the proposed architecture, N_a is the number of MRs in a nested mobile network under a TLMR connected to an AR through a GSL. Thus, the maximum number of MRs whose route is going to be changed due to the change in the GSL's connectivity is N_a. In the extended NEMO BSP, the MRs that have to process router advertisements are those N_a MRs and their neighbors. The TLMR and its N_n neighbors process the router advertisements. After that $(N_n - 1)$ neighbors of each of the remaining MRs of N_a MRs will process the router advertisements. Thus, n_{rap} is $N_a(N_n - 1) + 2$. On the otherhand, the number of execution of the Bellmanford's algorithm is $N_{sat} * N_n$. All routers onboard all satellites have to update their minimum distance to the ground station, and their neighbors have to run the algorithm when a vector indicting the change in the distance to the ground station is received.

8.4.1.2 Proportionality constant for the number of computations performed

In the extended NEMO BSP, the number of computations, performed by the algorithm that processes the router advertisement, is proportional to N_n. Therefore, the proportionality constant for the number of computations performed in the extended NEMO BSP, $C_{nemo} = n_{rap} * N_n$.

On the otherhand, the number of computations performed by the Bellmanford's algorithm is $(N_{sat} + N_g) * N_n$ which is the dimension of the distance table used in the algorithm. Thus, the proportionality constant for the number of computations performed if the distributed Bellmanford's algorithm is used, $C_{bell} = N_{sat} * N_n * (N_{sat} + N_g) * N_n$.

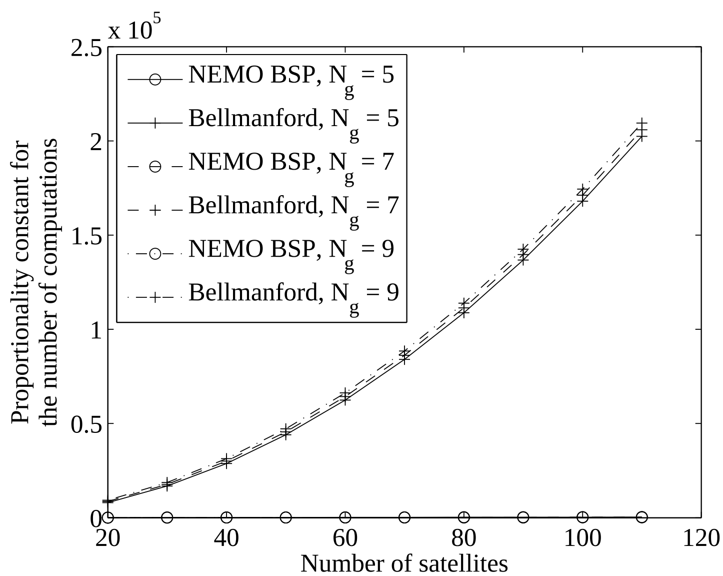

Figure 8.12: The proportionality constant for the number of computations performed when a GSL's connectivity changes.

Figure 8.12 shows the proportionality constant for the number of computations required when a GSL's connectivity changes. The number of computations required in Bellmanford's is much more than that required in the extended NEMO BSP because of the additional computations required to maintain the distance table in Bellmanford's algorithm.

8.4.2 Computational effort when an ISL's connectivity changes

It is difficult to analytically approximate the number of computations required when an ISL's connectivity changes. However, we can comment on the relative measure of the number. In the extended NEMO BSP, MRs that are affected by the change are expected to be at lower levels of the nested mobile networks. Therefore, the number of MRs, whose distance from the TLMR is changed, will be very small, and so is the number of computations. The number of computations in Bellmaford's will be more than that in the extended NEMO BSP. Because every router keeps the record of the distance from every other router, and the distance metric for some of the routers will change due to the change in an ISL's connectivity.

8.4.3 Delay in rerouting packets

The smaller computational effort of extended NEMO BSP than that of the Bellmanford's comes at the price of additional delay in rerouting packets. When a GSL's connectivity changes, an MR in the mobile network under a TLMR can start sending packets along a new route when the registration with the HA is complete. Therefore, the rerouting-delay for an MR is the sum of the time to propagate the router advertisement from the ground station to the MR and the time to perform the registration with the HA. For the distributed Bellmanford's, the rerouting-delay consists of the propagation delay only. Therefore, the rerouting-delay will be more in the extended NEMO BSP than in the Bellmanford's algorithm by an amount equal to the registration-time which can be twice as much as the end-to-end delay shown in Fig. 8.6.

8.4.4 End-to-end delay

End-to-end delay will also be larger in the extended NEMO BSP than in the Bellmanford's algorithm because of the additional delay incurred for packets' traversing through HAs. However, the magnitude of the additional delay will be very small compared to the value of the end-to-end delay which involves large delays in ISLs.

8.5 Summary

In this chapter, we have presented a Network Mobility (NEMO) architecture and an extension of NEMO Basic Support Protocol (BSP) for satellite networks to transfer data from multiple IP-enabled devices onboard satellites to the Internet. In the extended NEMO BSP, a Mobile Router (MR) choose an MR from multiple available MRs while forming a nested mobile network. The choice of an MR tries to achieve the minimum nesting level along with relaxing the level constraint by some threshold when Top Level MRs are overloaded. The proposed NEMO architecture and the extended NEMO BSP can ensure an efficient and continuous transfer of data from satellites to the Internet despite movements of satellites.

The architecture and the protocol enable the satellite network to become an integrated part of the Internet without the use of any other protocol. Results show that when the TLMRs become overloaded, the relaxation of the minimum level constraint can improve the performance in terms of the receive ratio of packets and drops, at the cost of increased tunneling overhead. Therefore, TLMR can deploy the relaxation with a maximum limit on the threshold.

Chapter 9

Conclusion

We have evaluated the Network Mobility (NEMO) schemes for the terrestrial and satellite networks. Our evaluation reveals the best performing schemes from a number of schemes that have been proposed to solve the problem of route optimization in NEMO, improves the performance of the best performing schemes, and provides insight on which aspects of the best schemes have to be prioritized for better performance. We also demonstrate NEMO-architecture and the extension of the basic scheme for NEMO in satellite networks for continuous transfer of data from devices onboard satellites to anywhere in the Internet.

First, we classified a large number of schemes proposed to solve the problem of route optimization in NEMO, and perform comparisons among the classes as well as the individual schemes in each class. Our evaluation reveals the prefix delegation as the best class of schemes in terms of various performance metrics.

Using simulations, we further evaluated four representative schemes of the prefix delegation class – namely SPD, MIRON, OPR and Ad hoc-based. Results indicate that Ad hoc-based performs the worst whereas OPR could perform the best provided the communication is initiated outside the mobile network and packets flow out of the mobile network at a certain interval. We find that the differences between SPD and MIRON can be considered as the trade off of the complete route optimization with the reduction of the amount of signaling. Our evaluation shows that SPD, which does not optimize route for LFNs, can perform as good as MIRON only when the mobile network is topologically close to the home network, and MIRON is the best scheme provided the speed of the mobile network and the nesting level is not too large, and the impact of the signaling in MIRON is insignificant compared to the impact of the unoptimized route in SPD. We evaluate the impact of the signaling on the performance of the schemes by increasing the number of hosts in the mobile network. Results indicate that optimizing route has to be prioritized over the reduction of signaling.

We performed a cost analysis of the schemes to determine the load imposed on the infrastructure by the operation of the schemes. The analysis considers the load due to the consumption of the processing and transmission power at various entities in the network. Results of cost analysis also indicates that optimizing routes has to be prioritized over the reduction of signaling.

Our evaluation also reveals the inability of the prefix delegation-based schemes to optimize route for intra mobile network communications. We proposed extensions for MIRON and OPR to optimize route for intra mobile network case. Results shows that the performance of the extended schemes is improved significantly for the intra mobile network communication.

We also demonstrated the application of the NEMO to satellite networks for continuous transfer of data from multiple devices onboard satellites to anywhere in the Internet. For this purpose, we propose architecture to use NEMO in satellite networks, and necessary extensions for NEMO Basic Support Protocol (BSP) for use in the satellite networks. Unlike the previous architecture of host mobility and NEMO in satellite networks, our architecture can ensure continuous transfer of data from satellite networks to the Internet without using any other intermediate protocols. Extension of the protocol ensures the efficiency of NEMO BSP by taking advantage and coping with the disadvantages of the characteristics of satellite networks.

In summary, NEMO is an efficient technique to manage the mobility of multiple hosts which are moving together provided suitable schemes can be used to solve the problem of the route optimization in NEMO. And, We have shown that which schemes will be suitable to solve the problem, and that which aspects of the suitable schemes have to be prioritized to achieve better performance. In addition, we have shown the way to use NEMO for continuous transfer of data from multiple hosts onboard satellites to anywhere in the Internet. Therefore, NEMO can be considered for managing mobility of multiple hosts in both terrestrial and satellite networks.

Bibliography

[1] C. Perkins, "IP mobility support for IPv4," RFC 3220, Jan. 2002.

[2] D. B. Johnson, C. E. Parkins, and J. Arkko, "Mobility support in IPv6," RFC 3775, Jun. 2004.

[3] D. Le, X. Fu, and D. Hogrefe, "A review of mobility support paradigms for the Internet," *IEEE Communications Surveys & Tutorials*, vol. 8, no. 1, pp. 38–51, First Quarter 2006.

[4] T. Ernst and H-Y. Lach, "Network mobility support terminology," RFC 4885, Nov. 2007.

[5] E. Perera, V. Sivaraman, and A. Seneviratne, "Survey on network mobility support," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 8, no. 2, pp. 7–19, Apr. 2004.

[6] F. Mimoune, F. Nat-Abdesselam, T. Taleb, and K. Hashimoto, "Route optimization for large scale network mobility assisted by BGP," in *IEEE GLOBECOM*, Washington, DC, USA, Nov. 26-30, 2007.

[7] W. Su, H. Zhang, and Y. Ren, "Research on route optimization in mobile networks," in *International Conference on Wireless Communications, Networking and Mobile Computing*, Wuhan, China, Sep. 22-24, 2006.

[8] H. Park, T.J. Lee, and H. Choo, "Optimized path registration with prefix delegation in nested mobile networks," in *International Conference on Mobile Ad-hoc and Sensor Networks*, Wuhan, China, Dec. 13-15, 2005.

[9] S. Novaczki, L. Bokor, G. Jeney, and S. Imre, "Design and evaluation of a novel HIP-based network mobility protocol," *Journal of Networks*, vol. 3, no. 1, pp. 10–24, Jan. 2008.

[10] S. Novaczki, L. Bokor, and S. Imre, "A HIP based network mobility protocol," in *International Symposium on Applications and the Internet*, Hiroshima, Japan, Jan. 15-19, 2007.

[11] M. Watari, T. Ernst, and J. Murai, "Routing optimization for nested mobile networks," *IEICE Transaction on Communications*, vol. E89-B, no. 10, pp. 2786–2793, Oct. 2006.

[12] H. Cho, T. Kwon, and Y. Choi, "Route optimization using tree information option for nested mobile networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 9, pp. 1717–1724, Sep. 2006.

[13] L. Li-hua, X.L. Ma, and L. Yuan-an, "Route optimization solution for nested mobile network in local mobility domain with multiple local mobility anchors," in *International Symposium on Communications and Information Technologies*, Sydney, Australia, Oct. 16-19, 2007.

[14] A. Banno and F. Teraoka, "χLIN6: An efficient network mobility protocol in IPv6," in *International Conference on Information Networking: Advances in Data Communications and Wireless Networks*, Sendai, Japan, Jan. 16-19, 2006.

[15] A. Banno, T. Oiwa, and F. Teraoka, "χLIN6-NEMO: A network mobility protocol based on LIN6," *IEICE Transaction on Communications*, vol. E88-B, no. 7, pp. 2765–2776, Jul. 2005.

[16] S. H. Kim, Y. Y. Ahn, S. H. Kim, and T. I. Kim, "Route optimization using RIPng protocol in nested network mobility," in *8th International Conference Advanced Communication Technology*, Phoenix Park, Korea, Feb. 20-22, 2006.

[17] J. Park, S. Lee, Y. Lee, and H. Choo, "Route optimization with additional destination-information in mobile networks," in *International Conference on Computational Science and Its Applications*, Kuala Lumpur, Malayasia, Aug. 26-29, 2007.

[18] G.S. Kuo and K. Ji, "Novel hierarchical network mobility support protocol with bidirectional end-to-end route optimization solution for nested mobile networks," in *IEEE GLOBECOM*, San Francisco, California, Nov. 27-Dec. 1, 2006.

[19] L.H. Lu and Y.A. Liu, "Route optimization solution based on extended prefix information option for nested mobility network," in *International Conference on Wireless Communications, Networking and Mobile Computing*, Shanghai, P.R. China, Sep. 21-23, 2007.

[20] M. S. Jeong, Y. H. Cho, and J. T. Park, "Hierarchical mobile network binding scheme for route optimization in NEMO," *Wireless Personal Communiations*, vol. 43, no. 3, pp. 975–995, Nov. 2007.

[21] J. Ylitalo, J. Melen, P. Salmela, and H. Petander, "An experimental evaluation of a HIP based network mobility scheme," in *International Conference on Wired/Wireless Internet Communications*, Tampere, Finland, Mar. 12, 2008.

[22] J. Park, T.J. Lee, and H. Choo, "Route optimization with MAP-based enhancement in mobile networks," in *International Conference on Computational Science*, University of Reading, UK, May 28-31, 2007.

[23] J.Y. Hu, C.F. Chou, M.S. Sha, I.C. Chang, and C.Y. Lai, "On the design of micro-mobility for mobile network," in *IFIP International Conference on Embedded and Ubiquitous Computing*, Taipei, Taiwan, Dec. 17-20, 2007.

[24] V. P. Kafle, E. Kamioka, and S. Yamada, "MoRaRo: Mobile router-assisted route optimization for Network Mobility (NEMO) support," *IEICE Transactions on Information and Systems*, vol. E89-D, no. 1, pp. 158–170, Jan. 2006.

[25] C.M. Huang and C.H. Lee, "Signal reduction and local route optimization of SIP-based network mobility," in *IEEE Symposium on Computers and Communications*, Sardinia, Italy, Jun. 26-29, 2006.

[26] C.M. Huang, C.H. Lee, and J.R. Zheng, "A novel SIP based route optimization for network mobility," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 9, pp. 1682–1691, Sep. 2006.

[27] P. Lam, S. Liew, and J. Lee, "Cellular universal IP for nested network mobility," *Computer Networks*, vol. 51, no. 12, pp. 3617–3631, Aug. 2007.

[28] A. Dul, "Global IP network mobility using border gateway protocol," http://www.quark.net/docs/Global_IP_Network_Mobility _using_BGP.pdf, Mar. 2006.

[29] X. Hu, L. Li, Z. M. Mao, and R. Yang, "Wide-area IP network mobility," in *IEEE INFOCOM*, Arizona, USA, Apr. 13-19, 2008.

[30] R. Cuevas, A. Cabellos-Aparicio, A. Cuevas, J. Domingo-Pascual, and A. Azcorra, "fP2P-HN: A P2P-based route optimization architecture for mobile IP-based community networks," *Computer Networks*, vol. 53, no. 4, pp. 528–540, Mar. 2009.

[31] S.Y. Phang, H. Lee, and H. Lim, "Route optimization for NEMO based on AODV," in *International Conference on Advanced Communication Technology*, Phoenix Park, Republic of Korea, Feb. 17-20, 2008.

[32] H.-J. Lim, D. Lee, T. Kim, and T. Chung, "A model and evaluation of route optimization in nested NEMO environment," *IEICE Transaction on Communications*, vol. E88-B, no. 7, pp. 2765–2776, Jul. 2005.

[33] K. Leung, D. Shell, W. D. Ivancic, D. H. Stewart, T. L. Bell, and B. A. Kachmar, "Application of Mobile-IP to space and aeronautical networks," *IEEE Aerospace and Electronic Systems Magazine*, vol. 16, no. 12, pp. 13–18, Dec. 2001.

[34] D. Shi and C. Tang, "The handoff study on satellite networks based on mobility network," in *6th International Conference on ITS Telecommunications*, Chengdu, China, Jun. 21-23, 2006.

[35] W. Ivancic, P. Paulsen, D. Stewart, D. Shell, and L. Wood, "Secure, network-centric operations of a space-based asset - an abridged report," in *Earth-Sun System Technology Conference*, College Park, Maryland, Jun. 28-30, 2005.

[36] D. Israel, R. Parise, K. Hogie, and Ed Criscuolo, "Demonstration of Internet technologies for space communication," http://www.aiaa.org/spaceops2002archive/papers/SpaceOps02-P-T5-18.pdf, 2002.

[37] C. Ng, F. Zhao, M. Watari, and P. Thubert, "Network mobility route optimization solution space analysis," RFC 4889, Jul. 2007.

[38] S. Chien, B. Cichy, A. Davies, D. Tran, G. Rabideau, R. Castano, R. Sherwood, D. Mandl, S. Frye, S. Shulman, J. Jones, and S. Grosvenor, "An autonomous Earth-observing sensorweb," *IEEE Intelligent Systems*, vol. 20, no. 3, pp. 16–24, 2005.

[39] L. Wood, W. Eddy, W. Ivancic, J. McKim, and C. Jackson, "Saratoga: a delay-tolerant networking convergence layer with efficient link utilization," in *International Workshop on Space and Satellite Communications (IWSSC '07)*, Salzburg, Austria, Sep. 13-14, 2007.

[40] A. Conta and S. Deering, "Generic packet tunneling in IPv6 specifications," RFC 2473, Dec. 1998.

[41] C. Ng, P. Thubert, M. Watari, and F. Zhao, "Network mobility route optimization problem statement," RFC 4888, Jul. 2007.

[42] K.J. Lee, J. Park, and H. Kim, "Route optimization for mobile nodes in mobile network based on prefix delegation," in *IEEE 58th Vehicular Technology Conference*, Orlando, Florida, USA, Oct. 6-9, 2003.

[43] J. Jeong, K. Lee, J. Park, and H. Kim, "Route optimization based on ND-Proxy for mobile nodes in IPv6 mobile network," in *IEEE 59th Vehicular Technology Conference*, Milan, Italy, May 17-19, 2004.

[44] J. Song, S. Han, and K. Park, "Route optimization in NEMO environment with limited prefix delegation mechanism," in *International Conference on Computational Science*, University of Reading, UK, May 28-31, 2006.

[45] E. Perera, A. Seneviratne, and V. Sivaraman, "Optinets: An architecture to enable optimal routing for network mobility," in *International Workshop on Wireless Ad-Hoc Networks*, Oulu, Finland, May 31-Jun. 3, 2004.

[46] H. Petander, E. Perera, K.C. Lan, and A. Seneviratne, "Measuring and improving the performance of network mobility management in IPv6 networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 9, pp. 1671–1681, Sep. 2006.

[47] C. J. Bernardos, M. Bagnulo, and M. Calderon, "MIRON: MIPv6 route optimization for NEMO," in *4th Workshop on Applications and Services in Wireless Networks*, Boston, Massachusetts, USA, Aug. 8-11, 2004.

[48] M. Calderon, C. J. Bernardos, M. Bagnulo, I. Soto, and A. de la Oliva, "Design and experimental evaluation of a route optimization solution for NEMO," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 9, pp. 1702–1716, Sep. 2006.

[49] D. Forsberg, Y. Ohba, B. Patil, H. Tschofenig, and A. Yegin, "Protocol for carrying authentication for network access (PANA)," RFC 5191, May 2008.

[50] M. Kim, H. Radha, and H. Choo, "On multicast-based binding update scheme for nemo environments," in *The International Conference on Computational Sciences and its Applications*, Krakw, Poland, Jun. 23-25, 2008.

[51] R. Moskowitz and P. Nikander, "Host identity protocol (HIP) architecture," RFC 4423, May 2006.

[52] B. Jouaber M. Sabeur and D. Zeghlache, "MR-proxy based solution for nested mobile network problems," in *International Symposium on Wireless Personal Multimedia Communications*, Aalborg, Denmark, Sep. 17-22, 2005.

[53] H. Soliman, C. Castelluccia, K. El Malki, and L. Bellier, "Hierarchical mobile IPv6 mobility management (HMIPv6)," RFC 4140, Aug. 2005.

[54] M. S. Jeong and J. T. Park, "Hierarchical mobile network routing: route optimization and micro-mobility support for NEMO," in *Proceedings of International Conference on Embedded and Ubiquitous Computing*, Aizu-wakamatsu, Japan, Aug. 25-27, 2004.

[55] G. Malkin and R. Minnear, "RIPng for IPv6," Internet Draft, Jan. 1997.

[56] Y. Kim, S. Woo, S. Kang, and W. Park, "Local source routing based route optimization in nested mobile networks," in *International Conference on Computational Science and Its Applications*, Glasgow, UK, May 8-11, 2006.

[57] B. Jouaber M. Sabeur and D. Zeghlache, "Light-nemo+: Route optimzation for light-nemo solution," in *IEEE International Conference on Networks*, Singapore, Sep. 13-15, 2006.

[58] M.S. Jeong, Y.H. Cho, and J.T. Park, "Policy-based route optimization for network mobility of next generation wireless networks," in *5th International Conference on Ad-Hoc, Mobile, and Wireless Networks*, Ottawa, Canada, Aug. 17-19, 2006.

[59] H. Ohnishi, K. Sakitani, and Y. Takagi, "HMIP based route optimization method in a mobile network," Internet Draft, Oct. 2003.

[60] K. Kim, D. Lee, J. Y. Ahn, and H. H. Lee, "Hierarchical route optimization in mobile network and performance evaluation," in *Second International Conference on Embedded Software and Systems*, Xi'an, P.R. China, Dec. 16-18, 2005.

[61] H. Kim, G. Kim, and C. Kim, "S-RO: Simple route optimization scheme with NEMO transparency," in *International Conference on Information Networking*, Jeju Island, Korea, Jan. 31-Feb. 2, 2005, pp. 401–411.

[62] Z. Gu, D. Yang, and C. Kim, "Mobile IPv6 extensions to support nested mobile networks," in *18th International Conference on Advanced Information Networking and Application*, Tokyo, Japan, Mar. 29-31, 2004.

[63] J. Na, J. Choi, S. Cho, C. Kim, S. Lee, H. Kang, and C. Koo, "A unified route optimization scheme for network mobility," in *9th International Conference on IFIP TC6*, Delft, Netherlands, Sep. 21-23, 2004.

[64] J. Rosenberg et al., "SIP: Session initiation protocol," RFC 3261, Jun. 2002.

[65] Y. Rekhter, T. Li, and S. Hares, "A Border Gateway Protocol 4 (BGP-4)," RFC 4271, Jan. 2006.

[66] P. Lam, S. Liew, and J. Lee, "Cellular universal IP: A low delay mobility scheme based on universal IP addressing," in *The 8th ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, Montreal, Qc. Canada, Oct. 10-13, 2005.

[67] "Peer-to-peer," http://en.wikipedia.org/wiki/Peer-to-peer.

[68] R. Wakikawa, S. Koshiba, K. Uehara, and J. Murai, "ORC: Optimize route cache management protocol for network mobility," in *International Conference on Telecommunications*, Tahiti, French Polynesia, Feb. 23-Mar. 1, 2003.

[69] H. Cho, E. K. Paik, and Y. Choi, "RBU+: Recursive binding update for end-to-end route optimization in nested mobile networks," in *IEEE International Conference on High Speed Networks and Multimedia Communications*, Toulouse, FRANCE, Jun. 30-Jul. 2, 2004.

[70] R. Qureshi, A. Z. M. Shahriar, and M. Atiquzzaman, "Performance of prefix delegation based route optimization schemes for NEMO," in *IEEE Military Communication*, San Diego, California, Nov. 17-19, 2008.

[71] C. Perkins, E. Belding-Royer, and S. Das, "Ad hoc On-demand Distance Vector (AODV) routing," RFC 3561, Jul. 2003.

[72] K. Fall and K. Varadhan (eds.), "The network simulator *ns*-2: Documentation," http://www.isi.edu/nsnam/ns/ns-documentation.html.

[73] D. B. Johnson, C. E. Parkins, and J. Arkko, "Mobility support in IPv6," RFC 3775, Jun. 2004.

[74] R. Morris, "TCP behavior with many flows," in *IEEE International Conference on Network Protocols*, Atlanta, Georgia, Oct. 29-31, 1997.

[75] A. Z. M. Shahriar and M. Atiquzzaman, "Evaluation of prefix delegation-based route optimization schemes for NEMO," in *IEEE International Conference on Communications*, Dresden, Germany, Jun. 14-18, 2009.

[76] A. S. Reaz, P. K. Chowdhury, and M. Atiquzzaman, "Signaling cost analysis of SINEMO: Seamless end-to-end network mobility," in *Proceedings of First ACM/IEEE International Workshop on Mobility in the Evolving Internet Architecture*, San Francisco, California, Dec. 1, 2006.

[77] K. A. Jalil and J. Dunlop, "Signaling cost analysis of PDE-NEMO," in *World Congress on Science, Engineering and Technology*, Penang, Malaysia, Feb. 25-27, 2009.

[78] H.-J. Lim, M. Kim, J.-H. Lee, and T. M. Chung, "Route optimization in nested NEMO: Classifiction, evaluation, and analysis from NEMO fringe stub perspective," *IEEE Transaction on Mobile Computing*, vol. 8, no. 11, pp. 1554–1572, Nov. 2009.

[79] J. Xie and I.F. Akyildiz, "A novel distributed dynamic location management scheme for minimizing signaling costs in Mobile IP," *IEEE Transactions on Mobile Computing*, vol. 1, no. 3, pp. 163–175, July 2002.

[80] S. Fu and M. Atiquzzaman, "Signaling cost and performance of SIGMA: A seamless handover scheme for data networks," *Wireless Communication and Mobile Computing*, vol. 5, no. 7, pp. 825–845, Nov 2005.

[81] A. Z. M. Shahriar, M. S. Hossain, and M. Atiquzzaman, "A cost analysis framework for NEMO prefix delegation-based schemes," Technical report, TR-OU-TNRL-10-101, University of Oklahoma, January 2010.

[82] S. Miyakawa and R. Droms, "Requirements for IPv6 prefix delegation," RFC 3769, Jun. 2004.

[83] C. Makaya and S. Pierre, "An analytical framework for perfromance evaluation of IPv6-based mobility management protocols," *IEEE Transactions on Wireless Communications*, vol. 7, no. 3, pp. 972–983, Mar. 2008.

[84] Aiguo Fei, Guangyu Pei, Roy Liu, and Lixia Zhang, "Measurements on delay and hop-count of the internet," in *IEEE GLOBECOM - Internet mini-conference*, Sydney, Australia, Nov. 8-12, 1998.

[85] J. Xie and U. Narayanan, "Performance analysis of mobility support in IPv4/IPv6 mixed wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 2, pp. 962–973, Feb. 2010.

[86] S. Baek, J. Yoo, H. Cho, Y Choi, T. Kwon, and Eun K. Paik, "Route optimization for intra communication of a nested mobile network," in *Workshop on Networking in Public Transport*, Waterloo, Ontario, Canada, Aug. 10, 2006.

[87] R. J. Leopold and A. Miller, "The iridium communications system," *IEEE Potentials*, vol. 12, no. 2, pp. 6–9, 1993.

[88] R. A. Wiedeman and A. J. Viterbi, "The globalstar mobile satellite system for worldwide personal communications," in *International Mobile Satellite Conference*, Pasadena, California, Jun. 16-18, 1993.

[89] K. M. Kavi, "Beyond the blackbox," *IEEE Spectrum*, vol. 47, no. 8, pp. 46–51, Aug. 2010.

[90] K. Delin, Y. Chao, and L. Lemmerman, "Earth science system of the future: observing, processing, and delivering data products directly to users," in *IEEE International Geoscience and Remote Sensing Symposium*, Sydney, Australia, Jul. 9-13, 2001.

[91] O. Korcak, F. Alagoz, and A. Jamalipour, "Priority-based adaptive routing in NGEO satellite networks," *International Journal of Communication Systems*, vol. 20, no. 3, pp. 313–333, Jun. 2006.

[92] M. Werner, "A dynamic routing concept for ATM-based satellite personal communication networks," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 8, pp. 1636–1648, Oct. 1997.

[93] E. Ekici, I. F. Akyildiz, and M. D. Bender, "A distributed routing algorithm for datagram traffic in LEO satellite networks," *IEEE Transaction on Networking*, vol. 9, no. 2, pp. 137–147, Apr. 2001.

[94] E. Papapetrou and F-N. Pavlidou, "Distributed load-aware routing in LEO satellite networks," in *IEEE GLOBECOM*, New Orleans, Louisiana, Nov. 30-Dec. 04, 2008.

[95] A. Donner, M. Berioli, and M. Werner, "MPLS-based satellite constellation networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 3, pp. 438–448, Apr. 2004.

[96] E. Rosen, A. Viswanathan, and R. Callon, "Multiprotocol label switching architecture," RFC 3031, Jan. 2001.

[97] J. Chen and A. Jamalipour, "An adaptive path routing scheme for satellite IP networks," *Intl Journal of Communication System*, vol. 16, no. 1, pp. 5–21, Feb. 2003.

[98] A. Z. M. Shahriar, M. Atiquzzaman, and M. S. Rahman, "Mobility managment protocols for next-generation all-IP satellite networks," *IEEE Wireless Communications*, vol. 15, no. 2, pp. 46–54, Apr. 2008.

[99] J. Li, C. Blake, D. S. J. De Couto, H. I. Lee, and R. Morris, "Capacity of ad hoc wireless networks," in *ACM MobiCom*, Rome, Italy, Jul. 16-21, 2001.

[100] S. Jungmin and N. Vaidya, "Load-balancing routing in multichannel hybrid wireless networks with single network interface," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 1, pp. 342–348, Jan. 2007.

[101] F. Hoffman and D. Medina, "Optimum internet gateway selection in ad hoc networks," in *IEEE International Conference on Communications*, Dresden, Germany, Jun. 14-18, 2009.

[102] L. Wood, W. M. Eddy, C. Smith, W. Ivancic, and C. Jackson, "Saratoga: A scalable file transfer protocol," draft-wood-tsvwg-saratoga-11, Mar. 2012.

[103] A. Z. M. Shahriar, M. Atiquzzaman, and W. Ivancic, "Applicability and performance of nemo in satellite," in *Earth Science Technology Conference*, East Adelphi, Maryland, Jun. 24-26, 2008.

Appendix A

Author's Publications

[1] A. Z. M. Shahriar, M. Atiquzzaman, and W. D. Ivancic, "Network mobility in satellite networks: architecture and the protocol," *Accepted for publication in International Journal of Communication Systems*, 2011.

[2] A. Z. M. Shahriar, M. Atiquzzaman, and W. D. Ivancic, "Evaluation of route optimization protocols for nemo in satellite networks (invited)," *Accepted for publication in Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, 2011.

[3] A. Z. M. Shahriar, S. Hossain, and M. Atiquzzaman, "A cost analysis framework for NEMO prefix delegation-based schemes," *IEEE Transactions on Mobile Computing (Online)*, Jun. 2011.

[4] A. Z. M. Shahriar and M. Atiquzzaman, "Performance analysis of prefix delegation-based route optimization schemes: Effects of increasing the number of nodes," *Wireless Personal Communications*, vol. 61, no. 4, Dec. 2011.

[5] A. Z. M. Shahriar, M. Atiquzzaman, W. D. Ivancic, and L. Wood, "A sender-based TFRC for saratoga: a congestion control mechanism for a space-friendly transfer protocol," in *IEEE Aerospace Conference*, Big Sky, Montana, Mar. 5-12, 2011.

[6] A. Z. M. Shahriar, R. Qureshi, and M. Atiquzzaman, "Perfromance of prefix delegation-based route optimization schemes for NEMO," *Wireless Personal Communications*, vol. 56, no. 3, pp. 469–490, Feb. 2011.

[7] S. Hossain, A. Z. M. Shahriar, and M. Atiquzzaman, "Scalabilityanalysis of prefix delegation-based schemes," in *HPSR*, Dallas, Texas, USA, Jun. 13-16, 2010.

[8] A. Z. M. Shahriar, M. Atiquzzaman, and W. Ivancic, "Performance of prefix delegation-based route optimization schemes: intra mobile network case," in *IEEE International Conference on Communications*, Cape Town, South Africa, May 23-27, 2010.

[9] A. Z. M. Shahriar, M. Atiquzzaman, and W. Ivancic, "Route optimization in network mobility: solutions, classification, comparison, and future research directions," *IEEE Communications Surveys & Tutorials*, vol. 12, no. 1, pp. 24–38, First Quarter 2010.

[10] A. Z. M. Shahriar and M. Atiquzzaman, "Evaluation of prefix delegation-based route optimization schemes for NEMO," in *IEEE International Conference on Communications*, Dresden, Germany, Jun. 14-18, 2009.

[11] R. Qureshi, A. Z. M. Shahriar, and M. Atiquzzaman, "Performance of prefix delegation based route optimization schemes for NEMO," in *IEEE Military Communication*, San Diego, California, Nov. 17-19, 2008.

[12] A. Z. M. Shahriar, M. Atiquzzaman, and W. Ivancic, "Performance evaluation of NEMO in satellite networks," in *IEEE Military Communications*, San Diego, California, Nov. 17-19, 2008.

[13] A. Z. M. Shahriar, M. Atiquzzaman, and W. Ivancic, "Applicability and performance of NEMO in satellite," in *Earth Science Technology Conference*, East Adelphi, Maryland, Jun. 24-26, 2008.

[14] A. Z. M. Shahriar, M. Atiquzzaman, and M. S. Rahman, "Mobility managment protocols for next-generation all-IP satellite networks," *IEEE Wireless Communications*, vol. 15, no. 2, pp. 46–54, Apr. 2008.

[15] A. Z. M. Shahriar, M. M. Akbar, M. S. Rahman, and A. H. Newton, "A multi-processor based heuristic for multi-dimensional multiple choice knapsack problem," *Journal of Super Computing*, vol. 43, no. 3, pp. 257–280, Mar. 2008.

[16] A. Z. M. Shahriar and M. Atiquzzaman, "Network mobility in satellite networks," in *NASA Science Technology Conference*, College Park, MD, Jun. 19-21, 2007.

[17] A. Z. M. Shahriar and M. M. Akbar, "A multi-processor based heuristic for multi-dimensional multiple choice knapsack problem," in *International Conference on Computer and Information Technology*, Dhaka, Bangladesh, Dec. 19-21, 2003.

[18] A. Z. M. Shahriar and M. A. Kashem, "Optimal algorithms for star chromatic number of planar graphs," in *International Conference on Computer and Information Technology*, Sylhat, Bangladesh, 1999.

Appendix B

Acronyms

AODV Ad hoc On Demand Vector

AR Access Router

BA Binding Acknowledgement

BGP Border Gateway Protocol

BSP Basic Support Protocol

BU Binding Update

BW Bandwidth

CN Correspondent Node

CoA Care-of-Address

COR Cross-Over Router

DHCP Dynamic Host Configuration Protocol

DNS Domain Name System

FTP File Transfer Protocol

HA Home Agent

HIP Host Identification Protocol

HoA Home Address

IETF Internet Engineering Task Force

IP Internet Protocol

LCoA Local CoA

LFN Local Fixed Node

LMA Local Mobility Anchor

MA Mapping Agent

MCN Mobility Capable Node

MIP Mobile IP

MIPv6 MIP vesrion6

MNN Mobile Network Node

MR Mobile Router

NEMO NEtwork MObility

ORC Optimized Route Cache

P2P Peer to Peer

PANA Protocol for carrying Authentication for Network Access

PCH Path Control Header

PDA Personal Digital Assistant

RA Router Advertisement

RBU Recursive Binding Update

RCoA Regional CoA

RSV Rendezvous Servers

RTT Round Trip Time

SIP Session Initiation Protocol

TCP Transmission Control Protocol

TLMR Top Level MR

UDP User Datagram Protocol

VMN Visiting Mobile Node

Appendix C

Models developed by Lim et al.

The models developed in Chapter 3 are based on models developed by Lim et al. [32]. For the convenience of the comparison, we present Lim et al. [32] models in this appendix. Lim et al. [32] classified route optimization schemes and developed models for those classes. The models measure the number of BUs, memory requirement by mobility entities on the route and the latency of BU. Here we present the models developed by Lim et al. [32] for one class of schemes, called A&S class, that is similar to the prefix delegation-based schemes. For the convenience of the readers, the models are presented using similar notations that have been used to develop the models in this dissertation. The notations used to develop the models are as follows:

T_{bu} =Delay to send a BU to the HA

l =Nesting Level of an MNN

s_{bu} =Size of IPv6 header including control messages of BU packet

s_c =Size of one address entry

τ_r =Average router processing time to process a packet

τ_e =Average Router Processing Time to encapsulate or decapsulate a packet

τ_{bu} =Average processing time to process a BU packet

b_w =Average bandwidth available at a node

m_e	=Memory required at mobile entities on the route
n_b, n_c	=Number of BUs and CNs, respectively
n_r	=Number of visiting MRs in a level
n_m	=Number of visiting mobile nodes (VMNs)in the lowest level
h_a^h, h_t^a	=Avg. number of hops from AR to HA, TLMR to AR

The models are as follows:

- Number of BUs: All MRs and mobile nodes send BUs to their respective HAs and CNs. Therefore, the number of BUs is

  $n_b = (n_c + 1)\,(n_r + n_m)$.

- Latency of BU: Latency of sending BUs measures the time required for a BU to reach from an MNN to the HA. In A&S class, an MNN configures two CoAs and sends the CoAs to the HA. Therefore, the BU packet consists of IPv6 header including control message of BU and two CoAs. Also, the BU packet is encapsulated by the MNN. Considering the transmission and processing delays at each hop, tunnel processing time and the time to process the BU packet, the latency of BU becomes

  $T_{bu} = (\tau_r + \frac{s_{bu}+s_c}{b_w})(l + h_t^a + h_a^h) + 2(\tau_e + \tau_{bu})$.

- Memory requirement at mobility entities: TLMR has to keep track of the next hop MRs and requires n_r routing entries. Other MRs, at each level keep track of next hop using two entries - one for routing table entry and one for the secondary CoA obtained by attached MRs. Therefore, the total amount of entries required is $2lm_r$. Thus, total amount of entries required by all MRs is $n_r(2l + 1)$. The HAs of all MRs also have similar amount of memory requirement to create binding cache entries. Mobile nodes requires n_m entries to store their CoAs. Mobile nodes' HAs require $2n_m$ amount of memory for the binding entries. The CN of all MRs (except TLMR) and mobile nodes creates binding entries for two addresses. Thus the memory requirement for the CNs is $n_c(2n_m + (2l+1)n_r)$. Therefore, the memory requirement becomes

$m_e = s_c(n_r(4l+2) + 3n_m) + n_c(2n_m + (2l+1)n_r)$.