TKN Telecommunication Networks Group

Technical University Berlin

Telecommunication Networks Group

Handover Blackout Duration of Layer 3 Mobility Management Schemes

Svetoslav Yankov, Sven Wiethoelter

{yankov,wiethoel}@tkn.tu-berlin.de

Berlin, May 2006

TKN Technical Report TKN-06-002

This work was conducted within the project "POL4G" sponsored by
the German Research Foundation (DFG).

TKN Technical Reports Series

Editor: Prof. Dr.-Ing. Adam Wolisz

Abstract

In this work, we present a study of the existing Layer 3 mobility approaches towards suitability for low blackout times during handovers. Our starting point is the well-known Mobile IP protocol specified by the Internet Engineering Task Force (IETF). Although the protocol supports mobility of mobile computers and hand-held devices in the Internet, it performs poorly when handovers happen while a communication session is active. Enhancements, such as Fast Mobile IP (FMIP) and Hierarchical Mobile IP (HMIP) have been proposed to handle the drawbacks of Mobile IP. We describe these protocols and discuss other extensions to Mobile IP devised by IETF, as well as non-standardized schemes proposed by other research groups. The selection of extensions is focussed on approaches aiming small Layer 3 blackout times applicable for inter-domain and vertical handovers during ongoing communication sessions. The discussed Layer 3 approaches have been considered in previous work extensively before by means of simulations, analysis, and measurements. Since the scenarios of previous work vary greatly, a comparison of the Layer 3 protocols is hardly possible. Therefore, we compare the blackout times of different schemes analytically for one definition of the blackout time and by four specific scenarios.

Contents

Acronyms

AN Access Network

AP Access Point

AR Access Router

ARP Address Resolution Protocol

BS Base Station

CoA Care-of-Address

CBR Constant Bit Rate

CN Correspondent Node

CIMS Columbia IP Micro-Mobility Suite

DHCP Dynamic Host Configuration Protocol

DSR Dynamic Source Routing

DSDV Destination Sequence Distance Vector

FA Foreign Agent

FBACK Fast Binding Acknowledge

FBU Fast Binding Update

FMIP Fast Mobile IP

FNA Fast Neighbor Advertisement

FNAACK Fast Neighbor Advertisement Acknowledgment

GT-ITM Georgia Tech Internetwork Topology Models

HA Home Agent

HACK Handover Acknowledge

HI Handover Initiate

HMIP Hierarchical Mobile IP

IETF Internet Engineering Task Force

LCOA On-Link Care-of-Address

L2 Layer 2

L2-MT Layer 2 Mobile Trigger

L2-ST Layer 2 Source Trigger

L2-TT Layer 2 Target Trigger

L2-LU Layer 2 Link Up Trigger

L2-LD Layer 2 Link Down Trigger

L3 Layer 3

MAP Mobility Anchor Point

MBWA Mobile Broadband Wireless Access

MN Mobile Node

NAR New Access Router

NAT Network Address Translation

NCoA Next Care-of-Address

NOAH NO Ad-Hoc Routing Agent

nFA New Foreign Agent

NLCoA Next Link Care-of-Address

oFA Old Foreign Agent

PAR Previous Access Router

PCoA Previous Care-of-Address

PLCoA Previous Link Care-of-Address

PrRtAdv Proxy Router Advertisement

RCOA Regional Care-of-Address

RtSolPr Router Solicitation for Proxy Advertisement

SCTP Stream Control Transmission Protocol

SINR Signal-to-Noise-and-Interference-Ratio

SIP Session Initiation Protocol

TORA Temporally Ordered Routing Protocol

WWAN Wireless Wide-area Access Network

WLAN Wireless Local-area Access Network

Glossary — IETF Mobile IP Terms

Mobile IP Terms

This is an extract of a part of the Mobile IP terminology, as specified in RFC 3775 [1].

- Mobile Node (MN)
  A node that can change its point of attachment from one link to another, while still being reachable via its home address.

- Home address
  A unicast routable address assigned to a MN, used as the permanent address of the MN. This address is within the MN's home network.

- Correspondent Node (CN)
  A peer node with which a MN is communicating. The CN may be either mobile or stationary.

- Care-of-Address (CoA)
  A unicast routable address associated with a MN while visiting a foreign network.

- Home Agent (HA)
  A router on a MN's home link with which the MN has registered its current CoA.

- Binding
  The association of the home address of a MN with a CoA for that MN, along with the remaining lifetime of that association.

- Registration
  The process during which a MN sends a Binding Update to its HA or CN, causing a binding for the mobile node to be registered.

Handover-Specific Terms

Terms used to Describe Mobile IPv4 Handovers

This is an extract of some of the IPv4 Handover Terminology, as specified in [2].

- Old Foreign Agent (oFA)
  The Foreign Agent (FA) involved in handling the CoA of a MN prior to a Layer 3 (L3) handoff.

- New Foreign Agent (nFA)
  The Foreign Agent (FA) anticipated to be handling a MN's CoA after completion of an L3 handoff.

- Layer 2 (L2) Handover
  A process by which the MN changes from one link-layer (L2) connection to another. For example, a change of wireless Access Point (AP) is an L2 handover.

- Layer 3 (L3) Handover
  Movement of a MN between FAs which involves changing the CoA at Layer 3 (L3).

- Router Solicitation for Proxy Advertisement (RtSolPr)
  A message from MN to oFA requesting information for a potential handover.

- Proxy Router Advertisement (PrRtAdv)
  A messsage from oFA to the MN that provides information about neighboring links. The message also acts as a trigger for network initiated handover.

- Layer 2 trigger
  Information from L2 that informs L3 of particular events before and after L2 handover.

- Layer 2 Mobile Trigger (L2-MT)
  An L2 Trigger that occurs at the MN informing of movement to a certain nFA.

- Layer 2 Source Trigger (L2-ST)
  An L2 trigger that occurs at oFA, informing the oFA that L2 handover is about to occur.

- Layer 2 Target Trigger (L2-TT)
  An L2 trigger that occurs at nFA, informing the nFA that a MN is about to be handed off to nFA.

- Layer 2 Link Up Trigger (L2-LU)
  An L2 trigger that occurs at the MN or nFA, informing that the L2 link between MN and nFA is established.

- Layer 2 Link Down Trigger (L2-LD)
  An L2 trigger that occurs at the oFA, informing the oFA that the L2 link between MN and oFA is lost.

- Low latency handover
  L3 handover in which the period of time during which the MN is unable to receive packets is minimized.

- Low loss handover
  L3 handover in which the number of packets dropped or delayed is minimized.

- Seamless handover
  L3 handover that is both low latency and low loss.

- Network-initiated handover
  L3 handover in which oFA or nFA initiates the handover.
- Mobile-initiated handover
  L3 handover in which the MN initiates the handover.

Terms used to Describe Mobile IPv6 Handovers

This is an extract of some of the IPv6 Handover Terminology, as specified in RFC 4068 [3].

- Access Point (AP)
  A L2 device connected to an IP subnet that offers wireless connectivity to an MN.
- Access Router (AR)
  The MN's default router.
- Previous Access Router (PAR)
  The MN's default router prior to its handover.
- New Access Router (NAR)
  The MN's default router subsequent to its handover.
- Previous Care-of-Address (PCoA)
  The MN's CoA valid on PAR's subnet.
- Next Care-of-Address (NCoA)
  The MN's CoA valid on NAR's subnet.
- Router Solicitation for Proxy Advertisement (RtSolPr)
  A message from the MN to the PAR requesting information for a potential handover.
- Proxy Router Advertisement (PrRtAdv)
  A message from the PAR to the MN that provides information about neighboring links facilitating expedited movement detection. The message also acts as a trigger for network-initiated handover.
- Fast Binding Update (FBU)
  A message from the MN instructing its PAR to redirect its traffic (toward NAR).
- Fast Binding Acknowledge (FBACK)
  A message from the PAR in response to an FBU.
- Fast Neighbor Advertisement (FNA)
  A message from the MN to the NAR to announce attachment, and to confirm the use of NCoA when the MN has not received an FBACK.
- Handover Initiate (HI)
  A message from the PAR to the NAR regarding MN's handover.
- Handover Acknowledge (HACK)
  A message from the NAR to the PAR as a response to HI.

Chapter 1

Introduction

1.1 Motivation

The primary purpose of IP is delivering packets between hosts in the Internet. One of its most important functions is addressing. Every host in the Internet has a unique IP address, which specifies its location. Such an address consists of a network identifier and a host identifier. Thus, IP datagrams are first routed to the network in which the host is located and then to the host itself. This system has one flaw: the IP address is tied directly to the network where the host is located. This is not a problem for wired hosts whose location is hardly ever changed, but it is an issue for mobile hosts. The fact that IP addresses a node as well as specifies its location is known as the semantic overloading problem of IP.

An example for this problem could be a person travelling on a business trip. When she moves with her laptop to a new location (say, from Berlin to San Francisco) the point of attachment to the Internet, and thus the IP address of the laptop changes. When another host decides to start communication with the laptop, it would try to reach it under its home address in the home network. The request would be routed to the router which is responsible for the home network, but since the mobile host is not available there, delivery of the request fails. Mobile IP [1, 4, 5, 6] solves this problem by giving mobile hosts and routers the possibility to forward packets from one location to another.

Another problem arises for mobile nodes: dealing with *"handovers"*. When a mobile host is moving while a communication session is ongoing (for example during a VoIP call) it might move in the range of a new access point or even in a totally different network. The latter case is referred to as an "inter-domain handover". The movement might also be combined with a change of the access technology, such as a movement from a LAN to an UMTS cell. In this case we talk about a movement in a heterogeneous network. In any case, the change of IP address has to be done on-the-fly so that the ongoing session is not discontinued.

1.2 Problem Statement

The main problem which arises with handovers is the "blackout" time during which a Mobile Node (MN) is not able to receive packets. During this time the MN obtains a new IP address and updates its communication partners about the change. The blackout time varies in different handover scenarios. Especially during inter-domain handovers it might reach several hundreds of milliseconds. Such long handover delays can significantly disrupt an ongoing communication session. There are different schemes which try to reduce the blackout time during handovers. They can be categorized in two major groups: Layer 3 (L3) schemes and higher-layer solutions. As the names suggest, these schemes rely on mobility awareness either on Layer 3 or on the higher layers.

1.3 Scope of this Technical Report

This document focusses on the existing variations of the Mobile IP protocol. Our special interest is the handover delay—how fast can a mobile host obtain a new IP address and update its communication partners when changing its point of attachment to the Internet (Layer 3 Mobility).

We focus on inter-domain handover protocols and the handover latencies they achieve. Since the scenarios of previous work vary greatly, a comparison of the Layer 3 protocols is hardly possible. Therefore, we compare the blackout times of different schemes analytically for one definition of the blackout time and by four specific scenarios. After presenting all Layer 3 schemes, we compare them by using a single metric in the second part of Chapter 4.

Higher layer protocols, such as the Session Initiation Protocol (SIP) and the Stream Control Transmission Protocol (SCTP) are focused primarily on mobility management on an end-to-end basis and do not have the potential to achieve short inter-domain handover delays. The communication sessions in these protocols are initiated and maintained through servers. The behavior of the protocols during inter-domain handovers is similar to the standard Mobile IP scheme and perform unsatisfactory when a communication session is active during the handover. Contrary, some enhanced Mobile IP schemes seem to be able to reduce the inter-domain handover delay significantly. Therefore, this work focusses on Mobile IP extensions targeting low delays in inter-domain handovers.

Chapter 2 gives an introduction to Mobile IP. Chapter 3 presents extensions to Mobile IP devised by the IETF, which propose several enhancements to the standard protocol. Then, in Chapter 4 we present handover performance measurements presented in publications related to inter-domain handovers and make a comparative analysis based on a single metric. Finally, Appendix A presents a study of the existing Layer 3 Mobility simulation models implemented as extensions to the ns simulation environment.

Chapter 2

IP Mobility Support for the Internet

2.1 Overview

IP Mobility protocols, such as Mobile IPv4 [4] and Mobile IPv6 [1], allow Mobile Nodes (MNs) to remain reachable while moving around in the Internet. Each MN is always identified by its home address, regardless of its current point of attachment to the Internet. While situated away from its home, a MN is also associated with a Care-of-Address (CoA), which provides information about the MN's current location. Packets destined to the MN's home address are transparently routed to its CoA. For this purpose, the protocol provides for registering the CoA with a "Home Agent (HA)" in the home network. The HA intercepts packets destined for the MN and forwards them to the CoA through a tunnel. After arriving at the end of the tunnel, each packet is then delivered to the MN. Usually, a "Foreign Agent (FA)" in the foreign network would be the endpoint of the tunnel. In this case, the CoA of the MN is the IP address of the FA. Upon receiving tunneled datagrams, the FA decapsulates them and delivers the inner datagram to the MN.

There are two possible modes of communication between the MN and a "Correspondent Node (CN)". The first mode, "bidirectional tunneling", does not require that the CN supports Mobile IP. Packets from the CN are routed to the HA and then tunneled to the mobile node. Packets to the CN are "reverse tunneled" from the MN to the HA and then routed normally from the home network to the CN. Such routing is called "triangular routing".

The second mode, "route optimization", requires the MN to register its current binding at the CN. In this case, packets from the CN can be routed directly to the CoA of the MN and the triangular routing via the HA is avoided.

Since L3 protocols are decoupled from lower, technology-specific layers, Mobile IP is just as suitable for mobility across homogeneous media as for mobility across heterogeneous media. For example, Mobile IP should facilitate node movement from one wireless LAN cell to another, as well as node movement from a wireless LAN cell to an UMTS cell, provided the MN is equipped with the appropriate link layer devices.

2.2 Differences between Mobile IPv4 and Mobile IPv6

Mobile IPv6 shares many features with Mobile IPv4, but is integrated in IPv6 and offers some other improvements. This section summarizes the major differences between Mobile IPv4 and Mobile IPv6. A precise list of all differences can be found in [1].

In Mobile IPv6, there is no need to deploy special routers as foreign agents since the protocol operates in any location, without need of special support from the local router. Further, route optimization is supported as a part of the protocol, rather than as an extension to it. Mobile IPv6 also allows route optimization to coexist efficiently with routers that perform "ingress filtering"—a process in which the routers filter out packets originating from outside the network, but which have a source address indicating origination from inside the network.

Mobile IPv4 uses IP encapsulation to forward packets to the MN while it is away from its home network. Mobile IPv6 uses an IPv6 routing header instead, and thus reduces the overhead.

Another important difference is that Mobile IPv6 uses the IPv6 Neighbor discovery Protocol, as opposed to the Address Resolution Protocol (ARP) used by Mobile IPv4. This decouples the protocol from any particular link layer and improves its robustness.

A general difference between IPv6 and IPv4 is the handling of address configuration. An IPv6 host has the ability to automatically configure its address without the use of a stateful configuration protocol, such as the Dynamic Host Configuration Protocol (DHCP). This mode of operation is called "stateless auto-configuration", as opposed to the "stateful configuration" used in IPv4.

2.3 Mobile IP Protocol Extensions for Handover Latency Minimization

When using Mobile IP, the movement of a MN away from its home link is transparent to transport and higher-layer protocols and applications, since the IP address of the communicating nodes remains the same at all times. Therefore, the MN may easily continue communication with other nodes after moving to a new link. At least in theory. There are cases in which it is not possible for the MN to keep its IP address after a handover. For example when the MN operates in private address networks which are separated from the public Internet by Network Address Translation (NAT) devices, which is not an uncommon case. A solution to this problem is proposed by Levkovetz et al. in [7]. The basic idea is that the MN would use a specific source port for the communication with HA from which the HA would "guess" the real IP address of the MN. Further, the MN would use a dedicated destination port to tell the HA that it is communicating from behind a NAT.

Handovers between subnets served by different FAs (L3 handovers) require a change of the CoA and a succeeding registration of the new CoA with the HA. This process takes some non-zero time to complete as the Registration Request propagates through the network. During this period of time the MN is not able to send or receive IP packets. The latency involved in Layer 3 handovers can be above the threshold required for the support of delay-sensitive or real-time services. IETF is working on several drafts, which propose methods to achieve

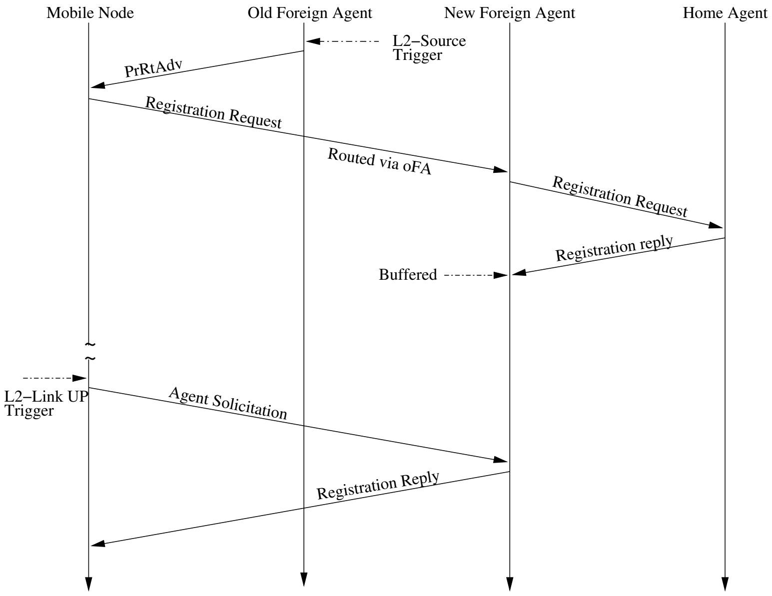

Figure 2.1: Pre-Registration Handover (Network Initiated, Source Trigger)

low-latency Mobile IP handovers [2, 3, 8, 9]. The methods are explained in the following.

2.4 Low Latency Handovers in Mobile IPv4

In [2] Mali et al. describe two techniques, which allow greater support for real-time services on a Mobile IPv4 network by minimizing the period of time when a MN is unable to send or receive IPv4 packets due to the delay in the Mobile IPv4 registration process.

2.4.1 Pre-Registration handover method

This handover method allows the MN to communicate with the new New Foreign Agent (nFA) while still connected to the Old Foreign Agent (oFA). This way, the MN is able to "pre-build" its registration state on the nFA prior to an underlying L2 handover.

The L3 handover can be either network-initiated or mobile-initiated. Accordingly, L2 triggers can be used both in the MN and in the FAs to trigger particular L3 handover events.

- Network-Initiated Handover
  A network initiated handover can be source triggered (Figure 2.1) or target triggered (Figure 2.2), depending on whether oFA (source trigger case) or nFA (target trigger case) receives an L2 trigger informing it about a certain MN's upcoming movement from oFA to nFA. In both cases the mobile node receives a *Proxy Router Advertisement message* (PrRtAdv), which contains information about the nFA. Upon reception of an PrRtAdv message the MN starts registration with nFA by sending it a *Registration Request* message. This message has to be routed through oFA since the MN is not

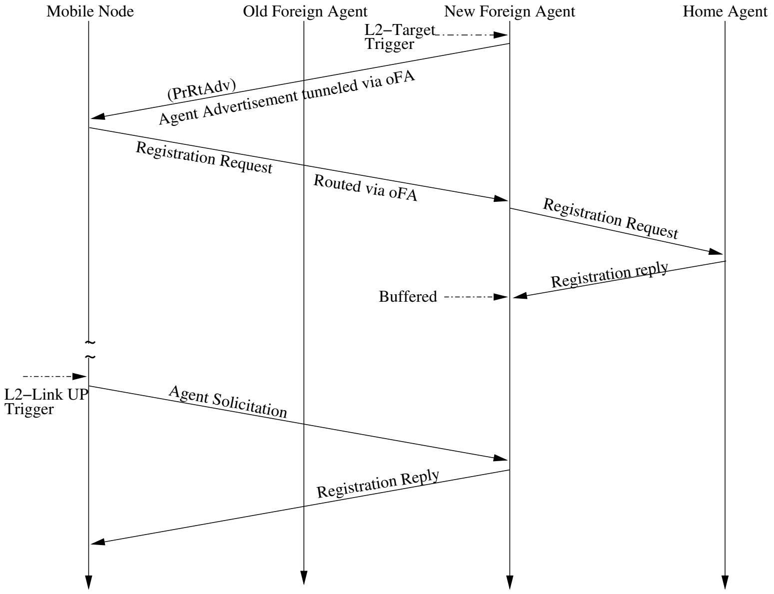

Figure 2.2: Pre-Registration Handover (Network Initiated, Target Trigger)

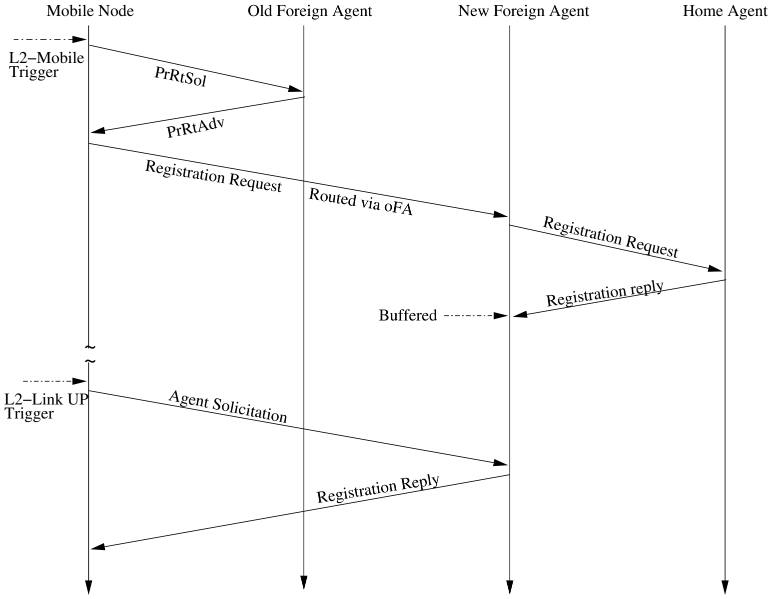

Figure 2.3: Pre-Registration Handover (Mobile Initiated)

directly connected to nFA prior to the L2 handover. The nFA performs the registration of the MN with the HA and buffers the *Registration Reply* until the MN completes the L2 handover and connects to nFA.

- Mobile-Initiated handover
  A mobile-initiated handover (Figure 2.3) occurs when a trigger is received at the MN to inform that it will shortly move to nFA. The L2 trigger contains information such as the nFA's identifier (i.e. its IPv4 address). As a consequence of the L2 trigger, the MN begins registration with nFA by sending the *"Proxy Router Solicitation"* (PrRtSol) message to oFA. The solicitation message must contain an identifier of nFA (i.e. nFA's IPv4 address). oFA replies to the MN with a PrRtAdv message containing the agent advertisement for the requested nFA. In order to expedite the handover, the actual nFA advertisement can be cached by oFA, following a previous communication between the two. Such caching can be done in a pre-soliciting process of known FAs to avoid performing the solicitation during an actual handover procedure. In case that oFA does not have cached information about nFA it has to make an PrRtSol–PrRtAdv exchange with nFA in order to obtain the information. The rest of the registration process is similar to the network-initiated cases.

2.4.2 Post-Registration handover method

This extension proposes the setup of a tunnel between nFA and oFA, thus it allows the MN to continue using its oFA while on nFA's subnet. This enables a rapid establishment of service at the new point of attachment which minimizes the impact on real-time applications. The MN must eventually perform a registration, but it can do this after communication with the nFA is established.

The handover process starts with either oFA or nFA receiving an L2 trigger informing it that a certain MN is about to move from oFA to nFA. In the former case the trigger is called Layer 2 Source Trigger (L2-ST) and in the later case Layer 2 Target Trigger (L2-TT)—to indicate whether the trigger is made in the previous network (source) or the destination network (target) of the MN. The trigger contains the MN's L2 address and an identifier for the other FA (i.e., the other FA's IPv4 address). The two FAs make a Handover Request (HRqst)–Handover Reply (HRply) exchange. The exchange triggers the initialization of a bi-directional tunnel between the two.

The point during the L2 handover in which the MN is no longer connected on a given link is signalled by an Layer 2 Link Down Trigger (L2-LD) trigger at oFA and MN. The completion of the L2 handover is signaled by an Layer 2 Link Up Trigger (L2-LU) trigger at nFA and MN. The trigger is handled as follows:

a.) When oFA receives the L2-LD trigger, it begins forwarding packets to MN through the forward tunnel to nFA.

b.) When the nFA receives the L2-LU trigger, it begins delivering packets tunneled from oFA to MN and forwards outstanding packets from MN using normal routing mechanisms or through a reverse tunnel to oFA or the HA.

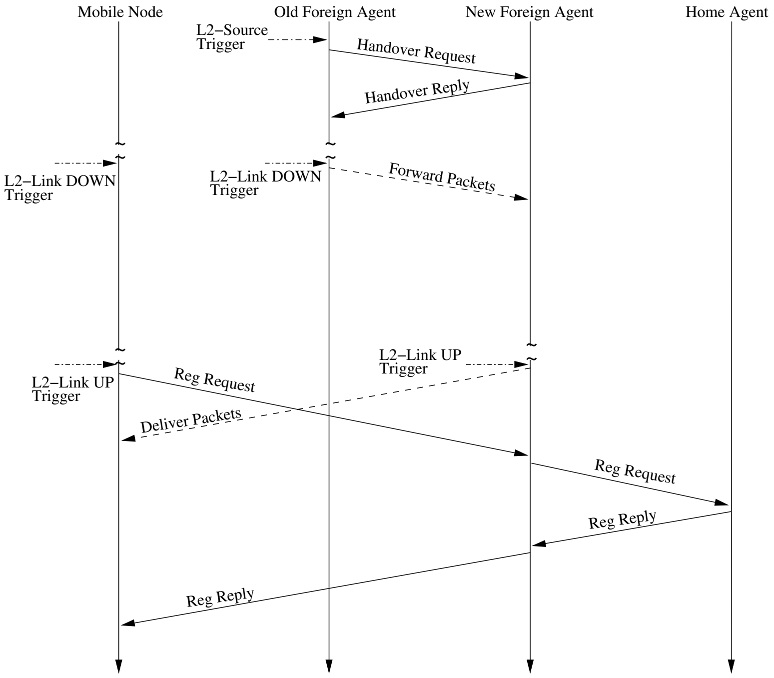

Figure 2.4: Post-Registration Handover (Source Trigger)

c.) When the MN receives the L2-LU, it initiates the registration process with nFA by soliciting an agent advertisement. After registration, the nFA takes over the role of default foreign agent for the MN.

Figure 2.4 shows the Post-Registration process after a source trigger. The only difference in the target trigger case is that nFA initializes the handover.

Chapter 3

Mobile IP Extensions

3.1 Fast Handovers for Mobile IPv6

3.1.1 FMIPv6 Overview

Standard Mobile IPv6 procedures have to deal with the same handover latency problem as Mobile IPv4. In [3], Koodli specifies a protocol to improve handover latency in Mobile IPv6 as [2] does for Mobile IPv4. There are some differences in the terminology used in both documents. The glossary in the beginning of this document lists the important terms used in both protocol specifications. The main difference comes from the lack of necessity to deploy special routers as "foreign agents" in Mobile IPv6, as in Mobile IPv4. Mobile IPv6 operates in any location without any special support required from the local router. Therefore, the routers supporting the MN while it moves in the Internet are just called Access Routers (ARs).

The approaches to reducing the handover latency in Mobile IPv4 and Mobile IPv6 are quite similar. The Fast Handover method is an extension proposed for Mobile IPv6 and resembles a combination of Pre-Registration and Post-Registration. On one hand the MN has the possibility to prepare its registration with New Access Router (NAR) and obtain its Next Care-of-Address (NCoA) while still connected to /acfPAR. This is like the Pre-Registration process described in Section 2.4.1. On the other hand the MN can instruct the PAR to forward packets addressed to its Previous Care-of-Address (PCoA) to its NCoA. This is similar to the bi-directional tunnel used in Post-Registration (described in Section 2.4.2), with the difference that in the fast handover case the MN triggers the forwarding whereas in Post-Registration no action from the MN is required.

3.1.2 Protocol Operation

Fast Handover can be either Network-Initiated or Mobile-Initiated, depending on whether one of the ARs or the MN initiates the handover. The triggers for the handover decision are beyond the scope of the RFC /citekoodli:fmip and are left as options. The two main possibilities are router discovery performed by MN on Layer 3 and a link-specific event (L2 trigger) occurring in the MN or in the network. Both cases result in a decision to perform a handover. In the description of the protocol operation in the following we use generic L2

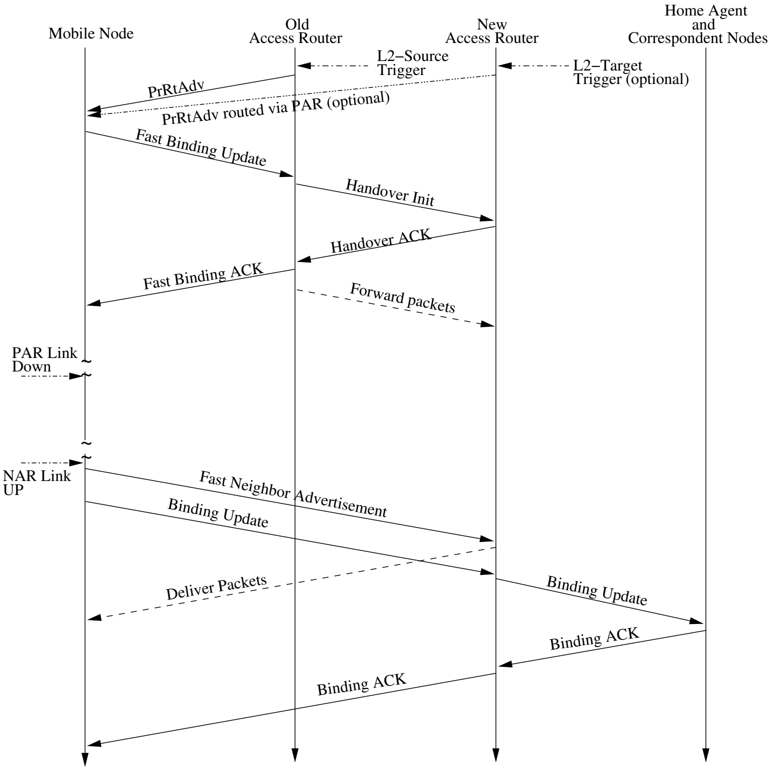

Figure 3.1: Network-Initiated Predictive Fast Handover (source or target trigger)

triggers to trigger the handover decision.

Fast handovers can further be classified in *"predictive"* and *"reactive"*, which resembles the Pre- and Post-Registration separation in Section 2.4:

- Network-Initiated Predictive Fast Handover (Figure 3.1)
  After receiving an L2 source trigger, the PAR sends a PrRtAdv message to the MN. The message contains the link layer address, the IP address, and subnet prefix of the NAR. The trigger could also be received at the NAR (L2-TT) and forwarded to PAR. The PrRtAdv message is a trigger for the MN to start registration with NAR. It sends a *Fast Binding Update (FBU)* message to PAR. The specification in [3] suggests that the MN may formulate a "prospective" NCoA by using the information about the NAR in the PrRtAdv message. The NCoA would be sent in the FBU message to inform PAR about the new CoA of the MN. An optional choice would be to leave the address assignment to the routers. This might be preferable from the perspective of inter-domain handovers (in heterogeneous networks), in order to avoid conflicts with address assignment schemes in the new network the MN might not be aware of.

  In the case of assigned addressing (i.e., addresses are assigned by the router) PAR and NAR have to make a Handover Initiate (HI)–Handover Acknowledge (HACK) exchange. In this exchange the NAR is informed about the imminent handover, upon which it

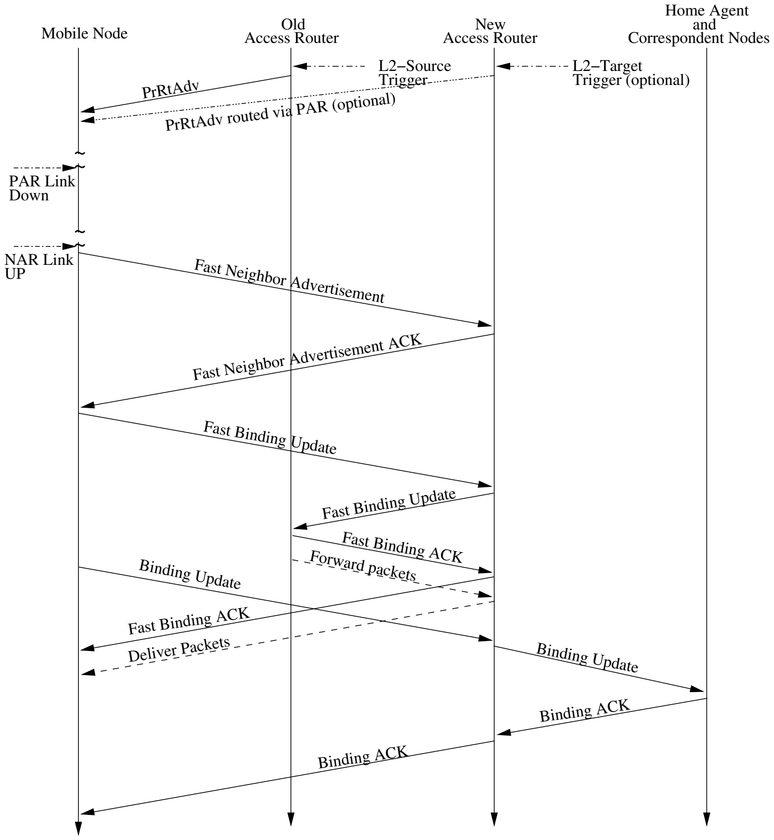

Figure 3.2: Network-Initiated Reactive Fast Handover (source or target trigger)

assigns a NCoA for the MN and sends it back to PAR. Then PAR sends a *Fast Binding Acknowledge (FBACK)* to the MN, informing it about the NCoA and starts forwarding incoming packets addressed to PCoA to NCoA in NAR's subnet. At this time the NAR is still unable to deliver the tunneled packets, since the MN hasn't yet completed the L2 handover. Depending on the implementation, the NAR may start buffering the tunneled packets. When the MN completes the L2 handover, it sends a *Fast Neighbor Advertisement (FNA)* message to NAR, informing it about its arrival. This is a trigger for NAR to start forwarding buffered and new incoming packets to the MN. Meanwhile, the MN starts to send Binding Updates to its HA and CNs to inform them about its new binding.

- Network-Initiated Reactive Handover (Figure 3.2)
  In the reactive handover case, the registration of the MN with NAR is delayed until the L2 handover is completed. This can be either intentional, or can serve as a fall-back case when a preceding predictive handover couldn't complete successfully (for example when the L2 handover was completed before the FBACK message reached the MN).

  Upon arrival in NAR's subnet the MN sends an FNA message to NAR. NAR assigns

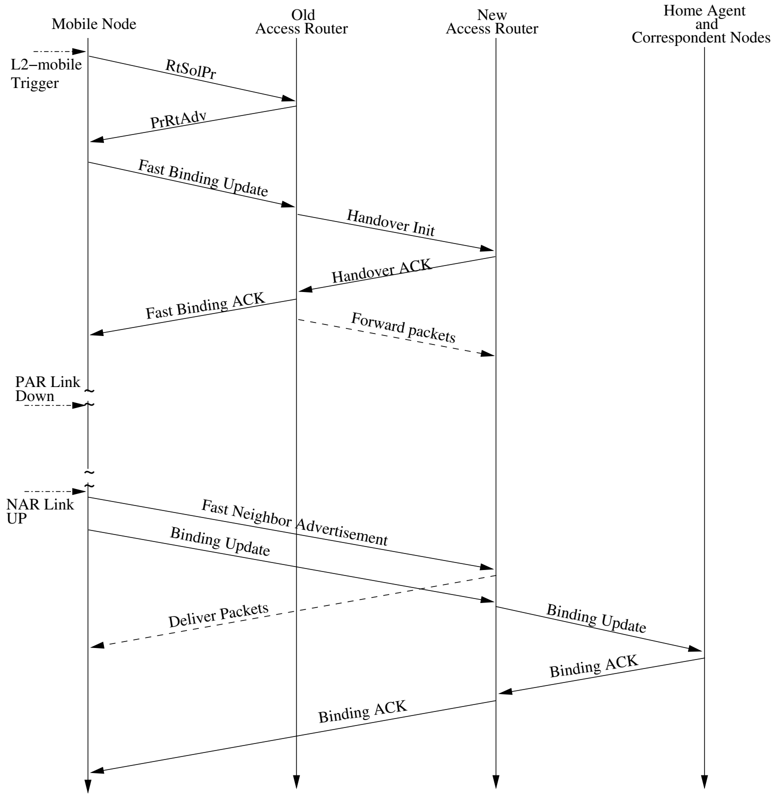

Figure 3.3: Mobile-Initiated Predictive Fast Handover

a NCoA for the MN and answers with a *Fast Neighbor Advertisement Acknowledgment (FNAACK)* message. Then the MN sends the FBU message (or repeats it in case of a preceding failed handover) to PAR, upon which PAR starts forwarding packets addressed to PCoA to NCoA in NAR's network.

- Mobile-Initiated Predictive Handover (Figure 3.3)
  In the mobile-initiated handover case, it is the MN, which receives an L2 trigger about an upcoming movement to a new AR. This would usually be a trigger about a new Access Point (AP) in the area. The MN reacts to the trigger by sending a RtSolPr message to PAR, which contains the AP identifier. PAR has to resolve the identifier to subnet-specific information and answers with a PrRtAdv message, which contains link layer, IP, and prefix information about the NAR. In case that the AP is unknown to PAR, it has to respond indicating that the new AP is unknown. In this case, the MN must stop fast handover protocol operations on the current link, but may conduct a reactive handover from the new link.
  If the new AP is known, the PAR supplies the information about the NAR to the MN through the PrRtAdv message. The rest of the operations is similar to the network-initiated predictive handover.

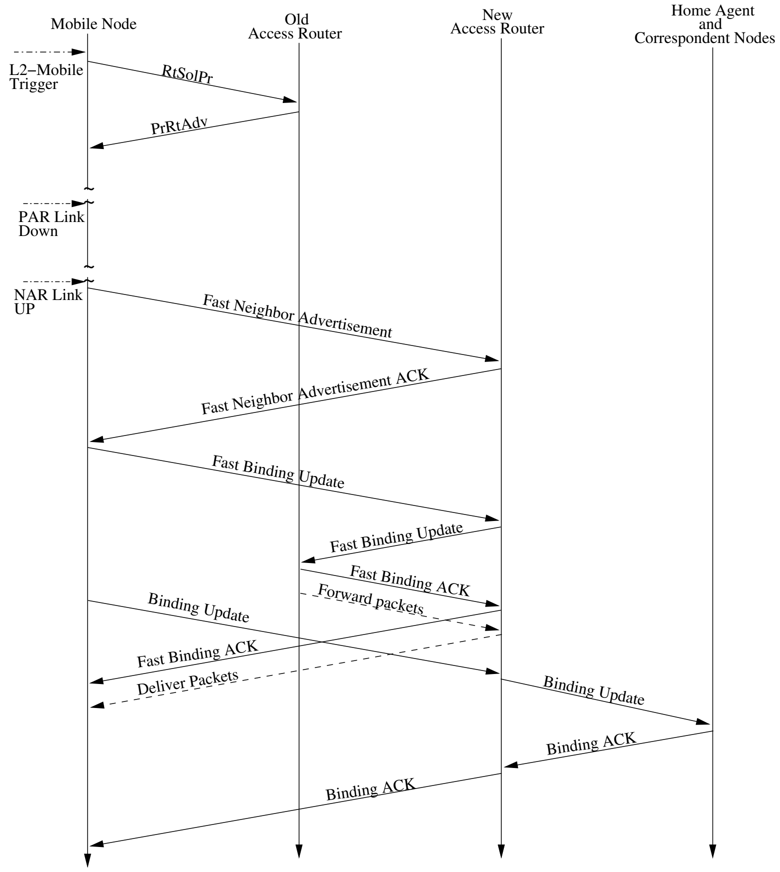

Figure 3.4: Mobile-Initiated Reactive Fast Handover

- Mobile-Initiated Reactive Handover (Figure 3.4)
  The reactive case of mobile-initiated fast handover starts like the predictive case. After the MN receives information about NAR through the PrRtAdv message it may choose to defer registration until the L2 handover is completed (intentionally reactive). As in the network-initiated reactive handover case, the handover may become reactive unintentionally when the predictive handover failed for some reason. Apart from the starting solicitation message, the mobile-initiated and the network-initiated fast handover cases are identical.

3.2 Hierarchical Mobile IPv6 Mobility Management (HMIPv6)

In [8], Soliman et al. suggest an extension to Mobile IPv6 which aims to reduce the amount of signaling between the MN and its CNs during a handover, and to improve the performance in terms of handover speed.

The authors state that the sending and processing of binding updates requires approximately one and a half round-trip times between the MN and each CN, in a best case scenario.

In addition, one round trip time is needed to update the HA. These round trip delays will disrupt active connections every time a handover to a new AR is performed. The RFC proposes the usage of a new Mobile IPv6 node, called Mobility Anchor Point (MAP). A MAP is essentially a local HA situated in the foreign network. It can be located at any level in a hierarchical network of routers. The introduction of the MAP provides an improvement in the following ways [8]:

- "The MN sends Binding Updates to the local MAP rather than the HA and CNs, which are typically further away."
- "Only one Binding Update message needs to be transmitted by the MN before traffic from the HA and all CNs is re-routed to its new location. This is independent of the number of CNs that the MN is communicating with."

For proper operation of HMIPv6, the MN has to be provided with two CoAs. A *Regional Care-of-Address (RCOA)* obtained by the MN from the visited network. This is the address which needs to be communicated to the CNs and HA of the MN. The second address is the *On-Link Care-of-Address (LCOA)*. This is the current address of the MN within the local MAP domain. Acting as a local HA, the MAP will receive all packets addressed to the RCOA of the MN and will encapsulate and forward them to the MN's current LCOA. If the MN changes its LCOA within MAPs domain, it only needs to register the new address with the MAP. The RCOA does not change as long the MN moves within the MAPs domain. This makes the MN's mobility transparent to the CNs it is communicating with.

The MAP domain's boundaries are defined by the ARs advertising the MAP information to the attached MNs. The MAP domain needs not be bounded within a physical domain with the same prefix. Routers from different locations, domains and technologies may participate in a hierarchical MAP domain. Every time the MN detects a movement, it will also detect whether it is still in the same MAP domain. If a change in the advertised MAP's address is received, the MN must obtain a new RCOA and perform binding updates with its CNs and HA by using the standard Mobile IPv6 protocol.

3.3 Fast Handover for Hierarchical Mobile IPv6 (F-HMIPv6)

In an IETF draft, which expired in April 2006 [9], Jung et al. propose a combination of the Fast Handovers and Hierarchical Mobile IP extensions to Mobile IPv6. The scheme is called "Fast Handover for Hierarchical Mobile IPv6" (F-HMIPv6). On one hand, such a combination has the potential to provide a low signaling overhead and delay related to the Binding Update procedure after a L3 handover (addressed by HMIPv6). On the other hand, it can also reduce the latency related to the movement detection and new CoA configuration during the L3 handover (addressed by FMIPv6).

3.3.1 F-HMIPv6 Overview

The authors suggest the following operation procedure of F-HMIPv6:
When a MN enters a new MAP domain, it first performs the HMIPv6 registration procedures with HA and MAP. Later, when the MN moves from a PAR to a NAR within the

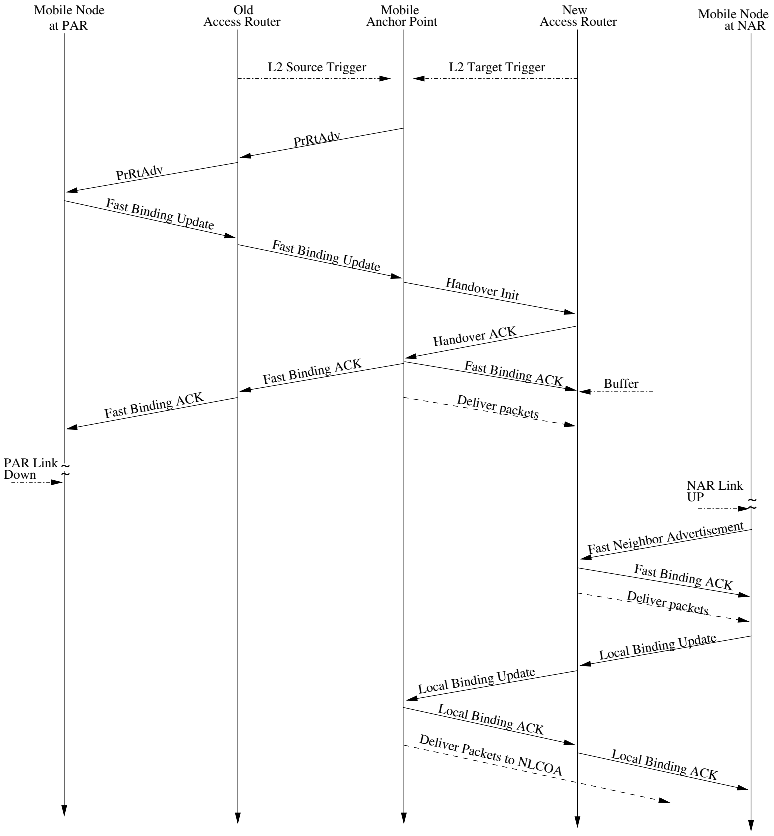

Figure 3.5: Network-Initiated Predictive Fast Hierarchical Handover 1

MAP domain, it will follow the local Binding Update Procedure of F-HMIPv6. During the handover, data packets sent by CNs will be tunneled by the MAP toward the NAR via a bi-directional tunnel, similarly to the FMIPv6 procedure. Optionally, the MAP may start bi-casting packets to PAR and NAR simultaneously. It should be noted that no bi-directional tunnel is established between PAR and NAR.

3.3.2 F-HMIPv6 Operation

Like FMIPv6, F-HMIPv6 can also support network- and mobile-initiated handovers. Further classification in predictive and reactive handovers is also possible, but a reactive handover is only partly reasonable as we'll see by an example.

Network-Initiated Predictive F-HMIPv6 handover

In the network-initiated predictive handover case either PAR (source trigger) or NAR (target trigger) receives an L2-trigger informing them that a MN's moving from PAR to NAR is imminent. The procedure consists of the following steps (see also Figure 3.5):

a.) The AR which received the trigger sends a handover indication to the MAP. The

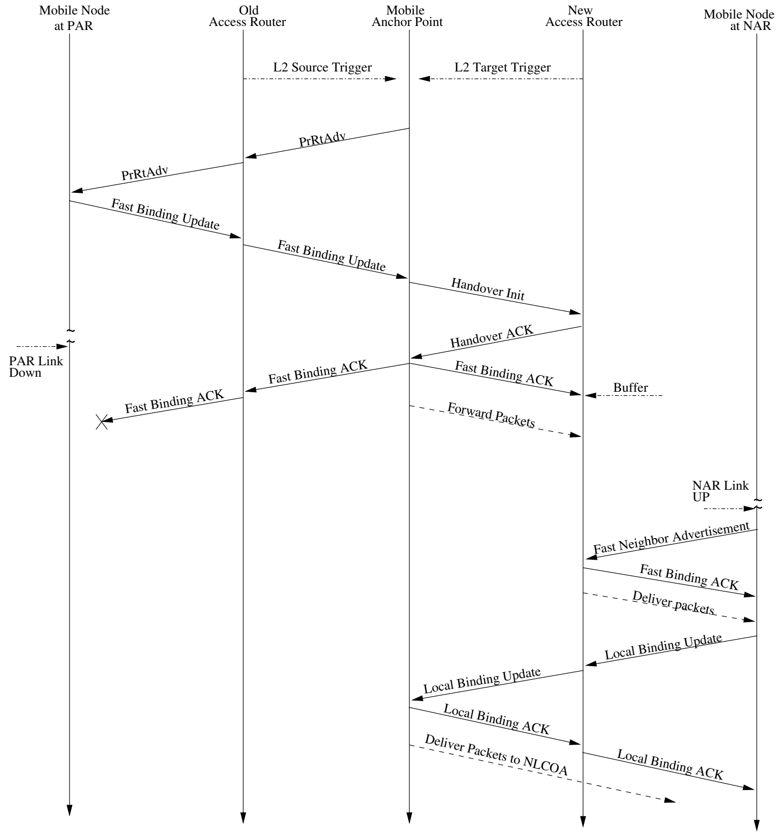

Figure 3.6: Network-Initiated Predictive Fast Hierarchical Handover 2

indication message includes the Previous Link Care-of-Address (PLCoA) of the MN as well as an identifier of the NAR (i.e., link layer or IP address).

b.) The MAP sends a PrRtAdv message to the MN, containing an advertisement of the NAR. The authors state that the message should contain information about the Next Link Care-of-Address (NLCoA) for the MN to use in the NAR region (i.e., NAR's network prefix for stateless auto-configuration or NLCoA for stateful configuration). Similar to the predictive network-initiated handover in the HMIPv6 extension, this approach is somewhat restrictive from the perspective of inter-domain handovers—it assumes that the MAP has sufficient knowledge about the associated NAR in order to assign the NLCoA address for the MN or even leaves the choice to the MN. As in the HMIPv6 case, it seems that leaving this decision to the NAR is best suitable.

c.) The MN sends an FBU message to MAP. The message contains the PLCoA and the IP address of the NAR.

d.) MAP and NAR make a HI–HACK exchange in order to establish a bi-directional tunnel.

e.) MAP sends the FBACK message toward the MN over PLCoA and NLCoA, and begins

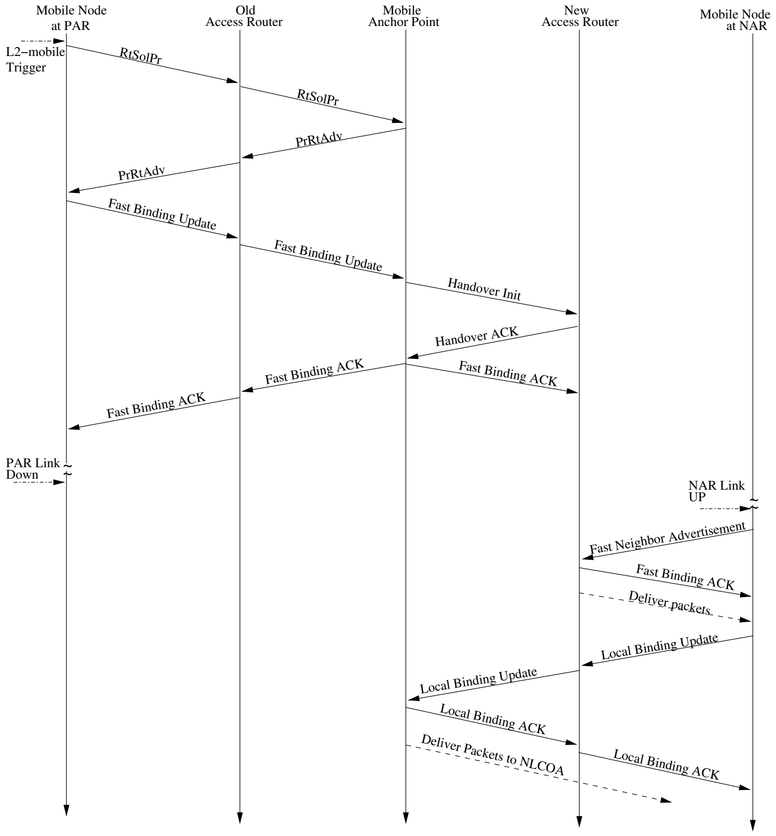

Figure 3.7: Mobile-Initiated Predictive Fast Hierarchical Handover 1

the forwarding of packets addressed to PLCoA to the NAR, via the established tunnel. NAR buffers the FBACK message and also the forwarded messages, in order to deliver them when the MN completes the L2-handover. The reason for sending the FBACK message back to MN via both NLCoA and PLCoA is that the MN might have completed the L2 handover to NAR before the FBACK message is received via PAR (Figure 3.6).

f.) When the MN detects by a Link-UP trigger that it has moved to NAR, it sends an FNA message to NAR.

g.) NAR sends the buffered FBACK message to MN for the case that the MN still hasn't received it. NAR also starts the forwarding of potentially buffered packets and the delivery of new incoming packets to the MN.

h.) MN makes a Local Binding Update (LBU)–Local Binding ACK (LBACK) exchange with the MAP to inform it about its arrival at NAR.

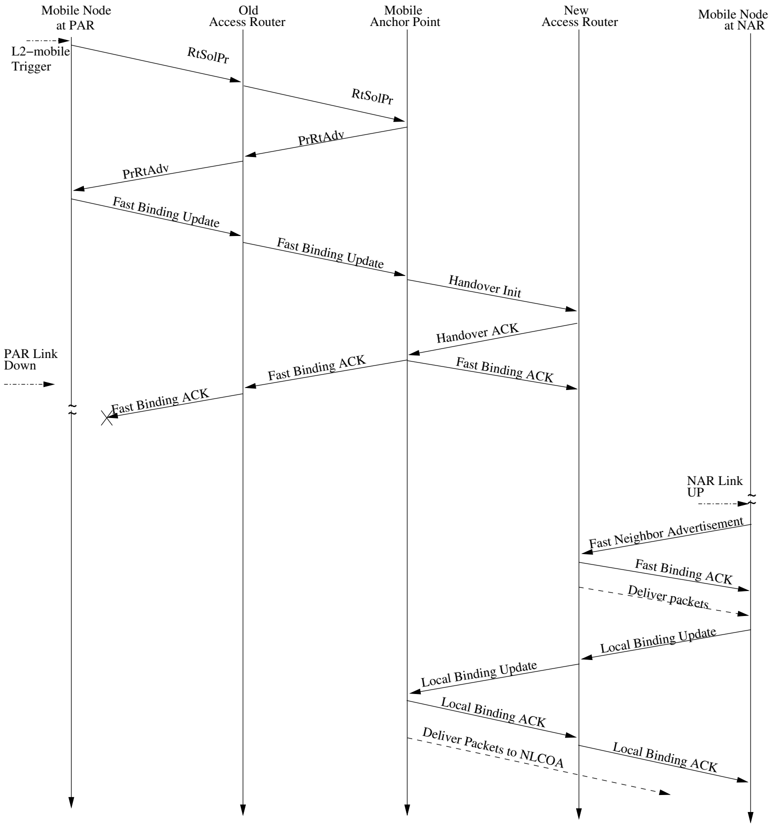

Figure 3.8: Mobile-Initiated Predictive Fast Hierarchical Handover 2

Mobile-Initiated Predictive F-HMIPv6 handover

The operation of a predictive mobile-initiated handover is the same as in the network-initiated case with the difference that it is the MN that receives a L2 trigger informing it about an imminent movement to a NAR. Upon the trigger the MN sends an RtSolPr message to MAP to request information about the NAR. The rest of the message exchange is as listed in 3.3.2. Figures 3.7 and 3.8 depict the operations.

Reactive F-HMIPv6 handover

Figure 3.9 shows why an intentional reactive F-HMIPv6 handover barely makes sense. Since the MN skips the sending of an FBU message while it is still connected to PAR, the establishment of a bi-directional tunnel between MAP and NAR is not fulfilled when MN arrives at NAR. Thus, it is the Local Binding Update message which informs MAP about the new binding. A tunnel establishment at this point is no more necessary, since the MAP can forward packets to NCoA directly.

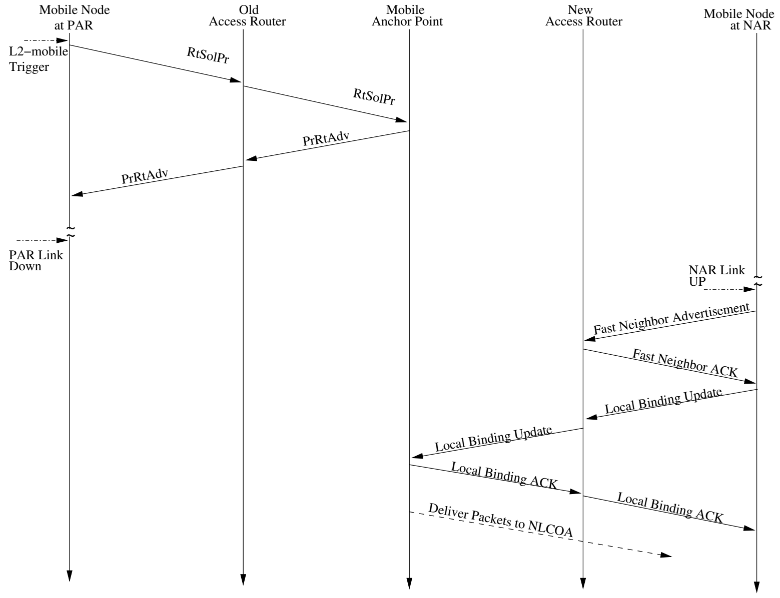

Figure 3.9: Mobile-Initiated Reactive Fast Hierarchical Handover

3.4 Seamless Mobile IP (S-MIP)

This section presents an interesting option that has not been standardized by IETF. Since this approach fits nicely in the F-HMIP discussion, it is presented in the following. In [10], Hsieh et al. propose a seamless handover protocol for hierarchical Mobile IP architectures (S-MIP), which is able to minimize the time during which the MN is unable to send and receive IP packets. S-MIP provides the possibility to further reduce the disruption during a L3 handover—the delivery of packets to the MN while it still has connectivity to PAR is combined with a simultaneous pre-registration procedure with NAR. The paper also suggests the usage of a new node—the "Decision Engine"—whose purpose is to monitor the movement of the MNs and make decisions about their handovers. However, in order to be consistent with the protocol descriptions in the previous sections, we keep the style of using L2 triggers to initiate handovers, without being specific on the movement and L2 events causing the trigger. This doesn't change anything on the handover mechanism described in the following.

3.4.1 S-MIP Overview

The S-MIP protocol can basically be regarded as a further extension to the F-HMIPv6 extension. After the handover decision has been made, every subsequent packet from a CN arriving at the MAP will be duplicated and sent to both the PAR and the NAR simultaneously. These packets will be marked with a Simulcast bit (S bit), as an option parameter in the IP header. Meanwhile, a bi-directional tunnel between PAR and NAR will be created (as opposed to F-HMIPv6, where the tunnel is between MAP and NAR). The NAR maintains a forward-buffer (f-buffer), containing packets forwarded from PAR (f-packets) and a simulcast-buffer

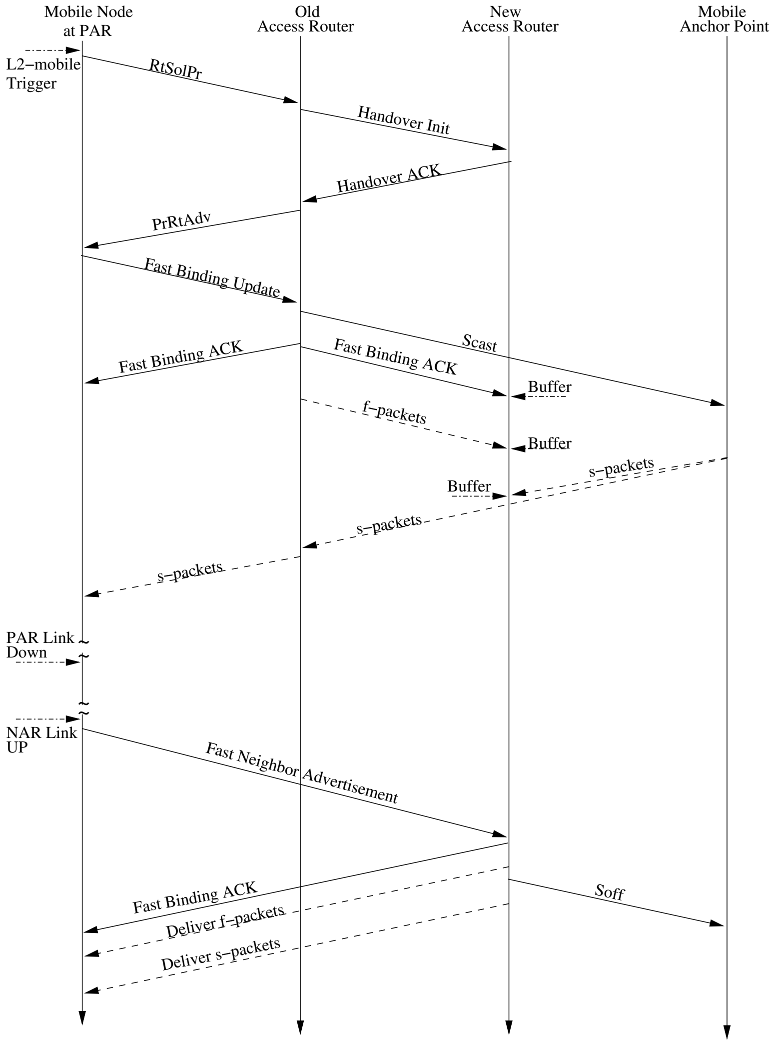

Figure 3.10: Mobile-Initiated Predictive Seamless Mobile IP handover

(s-buffer), containing packets that are marked with the S bit (s-packets). The PAR would forward only such packets to NAR, which don't have the S bit set. In addition, all packets (s/f packets) will be sent on the wireless channel by the PAR. This provides the possibility for the MN to continue receiving packets until the moment in which it loses connectivity to PAR.

3.4.2 S-MIP Operation

As in the previous extensions, the L2 trigger to initiate the L3 handover can be received in the MN (mobile trigger) or in the network (source/target trigger). As usual, the difference is only in the initial messages and we show just one of the cases this time (Figure 3.10).

Mobile-Initiated Predictive S-MIP handover

The procedure begins with the MN receiving an L2 trigger about an imminent handover to a NAR. It sends the RtSolPr message to PAR, which contains identifying information about the NAR. PAR and NAR exchange HI–HACK messages to initiate a bi-directional tunnel and agree on the NLCoA of the MN. Then, PAR answers to the MN with the PrRtAdv message informing it about its NLCoA. The MN initiates the L3 handover by sending the FBU message to PAR. PAR sends a *"Simulcast"* (Scast) message to MAP, which tells MAP to start setting the S bit in packets destined to MN and bi-cast them to PAR and NAR. PAR sends the FBACK message to NLCoA and PLCoA and continues to broadcast all messages coming from MAP. It also forwards the remaining messages which do not have the S bit set to NAR. Those might be some buffered messages, or the last messages coming from MAP before MAP started setting the S bit.

Upon arrival at NAR the MN sends the FNA message to NAR. NAR starts forwarding the f-packets from the f-buffer. When the f-buffer is empty NAR sends a *"Simulcast Off"* (Soff) message to MAP to tell it to stop the bi-casting. Then the NAR delivers the s-packets from the s-buffer to the MN. Meanwhile, MAP stops to set the S bit in packets destined to the MN and makes the usual packet delivery to the MN's current AR.

Chapter 4

Handover Latencies of Selected Non-Standardized Mobile IP Extensions

It is interesting to find out how the different extensions to Mobile IP presented in Chapter 2 perform in comparison to the basic Mobile IP protocol. Several publications have investigated the handover latencies, which apply when using the basic Mobile IP protocol and the Mobile IP extensions [10, 11, 12, 13, 14, 15, 16]. The selection of the publications was made from the perspective of our special interest in inter-domain handover approaches. We selected publications which present inter-domain handover latency results for any of the handover protocols standardized by IETF, or made a proposal for a new inter-domain handover protocol or extension.

In [10, 13], Hsieh et al. compare the S-MIP protocol, which was presented in Section 3.4, with the standardized IETF Extensions. In [11], Chow et al. propose a protocol for macro mobility support in Mobile Broadband Wireless Access (MBWA) networks, which is similar to FHMIPv6. In [12], Gwon et al. propose a protocol called FF-HMIP, which supports inter-MAP handovers. In [14], Kwon et al. suggest an extension to the basic Mobile IP protocol and present analytical results of its performance. In [15], Vivaldi et al. propose a macro-mobility handover scheme based on bi-casting. In [16], Zhang et al. suggest an improvement of the FHMIPv6 protocol which aims to decrease the home registration latency.

This Chapter gives an overview of the performed experiments and summarizes the experienced latencies.

4.1 Assumptions, Handover Latency Definitions, and Results of the Selected Approaches

In general, every handover latency investigation is based on different assumptions about the environment, the topology, the link delays, the definition of handover latency, etc. Although at the end a numerical result is available, it is not possible to compare the results presented in the different papers directly. Therefore, we give a summary of the basic experiment, the latency definition and the topology used in every paper in order to give a context to the presented handover latency result. The latency results presented in each paper are shown in a separate table after each summary. The next section compares the different protocols by means of a common topology scenario.

1.) In [14], Kwon, Gerla et al. present analytical results for inter-domain handover latency under usage of the basic Mobile IP protocol with a *Smooth Handover* extension, which is basically a tunneling functionality between oFA and nFA. The authors state that the handover delay consists of two components: *link layer establishment delay* and *signaling/disruption delay.* They further assume that the link layer establishment delay is negligible compared to the signaling delay and, therefore, concentrate on the signaling delay.

    The following link delays are used in the calculation of the handover latency:

    - t_s: MN ⇔ AR
    - t_{n0}: oFA ⇔ nFA
    - t_h: MN ⇔ HA
    - t_{hc}: HA ⇔ CN
    - t_{mc}: MN ⇔ CN

    The inter-domain handover is initiated when the MN detects the existence of a new domain (for example through a new agent advertisement on layer 3, with an IP of the advertising FA which belongs to a different subnet). The MN is eager to perform a handover and makes an RtSolPr–PrRtAdv exchange with nFA. It is assumed that the MN has an active connection to the CN at this time. The disruption time in which the MN is not able to communicate with CN starts from the moment of handover initialization. The RtSolPr–PrRtAdv exchange takes $2t_s$ time. Then the MN makes a registration of its new CoA with HA ($2t_h$). The nFA catches the registration reply from HA ($2t_h - t_s$) and sends a BU message to oFA (t_{no}). After this moment oFA starts forwarding packets destined to MN to nFA (therefore "smooth handover"). Until the first forwarded packet reaches MN, another ($t_s + t_{n0}$) pass by. This is also the end of the disruption time. The total sum yields

    $$2t_s + (2t_h - t_s) + t_{n0} + (t_s + t_{n0}) = 2t_s + 2_{th} + 2t_{n0}$$

    After oFA receives the BU message from nFA it performs a route optimization by sending a *"Binding Warning"* message to the HA ($t_h - t_s$). HA sends a BU message to CN (t_{hc}). After this moment CN can route packets directly to nFA and the route optimization is completed. The authors consider this instant as the end of disruption time. The total sum of disruption time yields:

    $$2t_s + (2t_h - t_s) + t_{n0} + (t_h - t_s) + t_{hc} + t_{mc} = t_{n0} + 3_{th} + t_{hc} + t_{mc}$$

    In the "Numerical Results" section of [14], the authors assign values to the link delays and present the results for 3 different scenarios:

(a) MN is located in its home network, while the distance between MN and CN is varied. The resulting disruption time is between 80 and 180 ms. The link delays are set as follows:

t_s	$10ms$	MN ⇔ nFA
t_{n0}	$5ms$	nFA ⇔ oFA (in this case oFA = HA)
t_h	$12ms$	MN ⇔ HA
t_{hc}	$t_{mc} - t_h$	HA ⇔ CN
t_{mc}	$25 - 75ms$	MN ⇔ CN (varied parameter)

(b) MN and CN are close to each other while the distance between the MN and its home network varies. The link delays in this case are set as follows:

t_s	$10ms$	MN ⇔ FA
t_{n0}	$5ms$	nFA ⇔ oFA
t_h	$15 - 40ms$	MN ⇔ HA (varied parameter)
t_{hc}	const.	HA ⇔ CN
t_{mc}	$25ms$	MN ⇔ CN

A figure of the delay between MN and its home network presented in the "Numerical Results" section of the paper shows a disruption time between 90 and 190 ms. The link delay between HA and CN is not given and obviously assumed to remain constant, since the disruption time grows linearly with the MN-HA link delay. However, CN and MN are assumed to be close to each other. This should imply that their link delays to HA should grow together, which is not visible in the analysis.

(c) In the third scenario the wireless link delay is being varied. The MN is located in its home network:

t_s	$10 - 60ms$	MN ⇔ FA (varied parameter)
t_{n0}	$5ms$	nFA ⇔ oFA
t_h	$12ms$	MN ⇔ HA
t_{mc}	$2t_s + 10ms$	MN ⇔ CN
t_{hc}	$t_{mc} - t_h$	HA ⇔ CN

The resulting disruption time is between 100 and 400 ms.

The results from the three scenarios are summarized in Table 4.1. Generally, the paper presents a simple means for analysis of handover delays. The presented numerical results have to be considered with caution, since the chosen link delays appear quite optimistic.

Also important is that the analysis considers solely the signaling delay after a handover has been detected on Layer 3. Losing of connectivity and movement detection on layer 2 are not considered.

Table 4.1: MIPv6 Disruption Time Results presented in [14]

Nr.	**Scenario**	**Disruption Time** [ms]
(a)	MN in home network. MN ⇔ CN distance varied	80-180
(b)	MN close to CN. MN ⇔ home network distance varied	90-190
(c)	MN in home network. wireless link delay varied.	100-400

2.) In [12], Gwon et al. present a handover performance analysis of MIP, HMIP, FMIP and FHMIP. Further the authors propose a modified FHMIP protocol (FF-HMIP), which supports inter-MAP handovers. When the MN moves to a new MAP domain it sends an FBU message to its previous MAP from NAR's link. Then, a bi-directional tunnel between previous MAP and NAR is created and the MN is able to receive packets, while its registration in the new domain is completed. Similar to [14], the authors assign link delay times for the different hops in the network. Here is a list of the used link delays:

- t_{up} : Wireless uplink transport latency.
- t_{down} : Wireless downlink transport latency.
- t_{cross} : Transport latency between two ARs in the same Access Network (AN).
- t_{gate} : Transport latency from AR to MAP (gateway router).
- t_{core} : Transport latency from one gateway router MAP (gateway router) to another.
- $t_{internet}$: Transport latency to a node in the Internet.
- t_{L2BO} : link layer blackout.
- t_{DAD} : duplicate address detection time.

The values of the individual link delays are not stated.

The **"IP Layer blackout duration"** is defined as the total time the MN is not reachable due to an IP handover. The blackout durations for the different Mobility Protocols are calculated as follows:

IP layer blackout duration for inter-AN handovers when using MIP or FHMIP:

$$t_{L2BO} + 2t_{up} + 2t_{down} + 2t_{gate} + 2t_{core} + t_{DAD}(+2t_{internet})$$

IP layer blackout duration for inter-AN handovers when using FMIP:

$$t_{L2BO} + 2t_{up} + 2t_{down} + 4t_{core} + 8t_{gate}$$

IP layer blackout duration for inter-AN handovers when using FF-HMIP:

$$t_{L2BO} + 2t_{up} + 2t_{down} + 2t_{gate} + 4t_{core} + 4t_{gate}$$

The authors performed a simulation of a Wireless Wide-area Access Network (WWAN) in which a large number of Base Transceiver Stations (BTSs) and Wireless Local-area Access Network (WLAN) Access Points provide for wireless connectivity. The WWAN and WLAN had optimal hexagonal cellular configuration with circular cells. Adjacent cells had overlapped coverage. A WWAN cell radius was $1000m$ (macro cell), whereas the WLAN consisted cells of $30m$ radius (micro cell). A single BST/AP per each micro/macro cell was assumed.

The experienced "mean IP layer blackout durations" of the different Mobility protocols are plotted out in a single graphic in the paper. The authors make an intra-domain and an inter-domain handover analysis and it is not clear whether the presented figure is for inter-domain or intra-domain handovers or a mix of both.

Table 4.2: IP Layer Blackout Duration Results presented in [12]

Nr.	**Approach Base**	**IP Layer Blackout Duration** [ms]
1	MIPv6	1300
2	HMIPv6	300–500
3	F-HMIPv6	200–400
4	FMIPv6	200
5	FF-HMIP	200

3.) In [13], Hsieh et al. make a comparison of the Handover protocols MIPv6, HMIPv6, FMIPv6, FHMIPv6, and S-MIP. The authors use the topology and link delays shown in Figure 4.1.

Both CN and HA are connected to an intermediate node (N1) with $2ms$ link delay and 100 Mbps links. The link between N1 and the MAP is a 100 Mbps link with 50 ms link delay. The MAP is further connected to the intermediate nodes N2 and N3 with

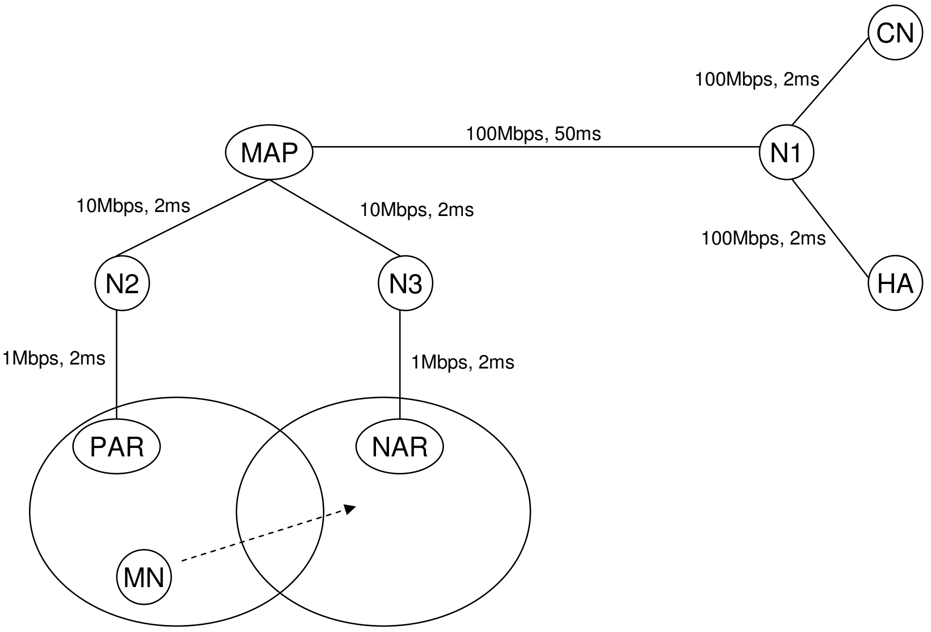

Figure 4.1: Network Topology used in [13]

$2ms$ link delay over 10 Mbps links. N1 and N2 are connected to PAR and NAR with $2ms$ link delay over 1 Mbps links. In order to simulate the handover, the MN starts to move towards NAR at the speed of one m/s. The L2 handover time is modeled by $20ms$. The address resolution time is fixed to $100ms$. This is the time taken for the MN to obtain a new CoA from NAR.

The handover is initialized by the MN as soon as it receives a router advertisement from an unknown router on Layer 3. Thus, no layer 2 triggering is involved in the handover decision.

CN and MN are involved in a TCP session in which a bulk data transfer application transfers packets from CN to MN. The packet size is 512 bytes and the TCP window size is 32. Further, no route optimization is used, which means that all packets are first routed to HA and then tunneled to MN. The handover latency is measured through the disruption of the TCP stream between the communicating stations. As soon as MN starts the registration process with NAR it is no longer able to receive TCP segments arriving from CN. The TCP session remains active and the CN continues to inject TCP segments in the network. When registration is completed, the MN begins to receive out-of-sequence segments. TCP conform, the MN sends (negative) acknowledgment messages, containing the expected sequence number. Upon receiving three such messages, the CN starts retransmission from the requested sequence number upwards. The reception of the first retransmitted segment by the MN is regarded as the end of disruption time. Basically, the sooner the MN starts receiving out-of-sequence segments, the shorter the disruption time will be.

Depending on the TCP implementation the MN might store the out-of-sequence segments while sending negative acknowledgments and waiting for the missing segments. This leads to another interesting effect. After three negative acknowledgments, the CN begins to retransmit the requested segments. When the MN receives all segments which went lost during the handover, it suddenly sends a cumulative acknowledgment acknowledging all segments which were previously received out-of-sequence. Depending on the propagation delay between MN and CN the CN will retransmit a certain number of segments which were already received by the MN, before receiving the first cumulative acknowledgment and continuing with segments having a sequence number above the requested. This degrades somewhat the profit achieved through a tunneling approach, such as FHMIPv6 (only in the case when no buffering is provided at NAR).

In order to avoid the unnecessary retransmission of received segments, the authors suggest an optimization to the S-MIP protocol for the usage together with TCP. The duplicate reception of f-packets is avoided in that the PAR does not broadcast them, but only forwards them to NAR. Thus, the MN will receive all f-packets and all s-packets when it connects to NAR. A figure of the optimized S-MIP handover behavior shows that not a single packet needed to be retransmitted, since the handover was lossless. The disruption consists solely of the L2 handover delay and the address resolution delay, which was about $100ms$ in this case.

The handover delays of the different protocols for this evaluation procedure are tabled in Table 4.3. The simulation environment was NS-2.

Table 4.3: TCP Disruption Time Results presented in [13]

Nr.	**Approach Base**	**TCP Disruption Time** [ms]
1	MIPv6	814
2	HMIPv6	326
3	FMIPv6	358
4	FHMIPv6	270
5	HMIPv6 + bi-casting	268
6	S-MIP	100**

** Although the HO latency is 100 ms, the authors argue that the delay perceived by the sender and the receiver is zero - totally seamless handover.

4.) In [15], Vivaldi et al. propose a macro-mobility handover scheme for HMIP. It is based on bi-casting. During a handover procedure the previous MAP of MN would bi-cast incoming packets to PAR in its own MAP domain and to NAR in the new MAP domain. NAR buffers the forwarded packets and delivers them to MN when it receives connectivity to it.

The handover delay in this paper is measured from the time the MN sends the FBU message to NAR until the time the first packet from CN, routed directly through the new MAP, reaches MN. Additionally, the authors consider the rendezvous time—the time needed for the MN to hear the beacon from NAR after roaming out of PAR's network. The sum of handover delay and rendezvous time gives the complete handover time. The rendezvous time can be up to a beacon interval duration. The authors assume a typical beacon interval duration of $100ms$ in a wide area cellular network. Apparently, the building of the bi-casting group is done before the MN leaves the domain of its previous MAP. How the MN is able to inform the MAP about the address information of NAR before hearing its beacon is not further explained in the paper.

The handover results are derived from NS-2 simulation experiments. CN and MN communicate over UDP. CN starts a Constant Bit Rate (CBR) traffic source, producing fixed length packets of 200 bytes each $20ms$. CN and ARs are connected to the two MAPs via wired links. The wired links have a bandwidth of 100 Mbps and a link delay of $2ms$, which is quite optimistic for far distance communication over the Internet (for the CN-MAP link). The wireless link propagation delay between MN and the ARs is varied from 10ms to 50 ms over a 2Mbps links. The results are shown in Table 4.4.

Table 4.4: Handover Delay+Rendezvous Time Results presented in [15]

Nr.	**Approach Base**	**Handover Delay + Rendezvous Time** [ms]
1	HMIPv6	~330–650
2	Proposed scheme with bi-casting	~140–315

5.) In [11], Chow et al. propose a protocol for macro (and micro) mobility support in Mobile Broadband Wireless Access (MBWA) networks. Handovers are mobile-initiated. The handover decision is based on a comparison between the MN's received Signal-to-Noise-and-Interference-Ratio (SINR) from the serving AP to the neighboring APs. The proposed protocol is similar to FHMIP, although different terminology is used (e.g. domain AR instead of MAP). Handover latency is defined as the elapsed time from the point in time when a MN initiates a handover by sending a request to a NAR, until the MN is able to receive packets from the NAR. This occurs when the MN receives a handover response from the new AP accepting its original request. Further, during inter-domain handovers, a tunnel between previous and new MAP for forwarding of packets is created.

An inter-domain handover starts when the MN decides to switch ARs due to a SINR-based decision (L2 mobile trigger). The MN sends a handover request to NAR from its CoA. The following procedure is equivalent to the reactive mobile-initiated FHMIPv6, as described in Chapter 3.3.2.

The handover latency is measured in simulation experiments conducted with the OPNET simulation environment. The used topology is the same as in Figure 4.1. The wireless link is modeled according to the 802.16e standard. In the numerical results section the authors redefine the handover latency: *"the overall handover latency is defined as the sum of two delays, the delay incurred for obtaining a new CoA and the delay for registering the new CoA with the HA"*. Apparently this definition is used for the evaluation. The overall handover latency of the proposed scheme is $128ms$, according to the results. There is no description of the communication model between MN and CN. The result is compared to the results presented by Hsieh et al. in [13], but it is not clear whether the handover delay is measured in the same way (disruption of the TCP stream). The results are shown in Table 4.5.

Table 4.5: Delay to Obtain and Register new CoA Results presented in [11]

Nr.	**Approach Base**	**Delay to Obtain and Register new CoA** [ms]
1	MIPv6	814
1	MIPv6 with Fast Handover	358
1	HMIPv6	326
1	HMIPv6 with Fast Handover	270
2	Proposed scheme	128

6.) In [16], Zhang et al. propose a modified FHMIP protocol. The resulting mobility management protocol is called "two-way registration". The scheme aims to decrease the home registration latency and hence minimize the disruption caused by macro-mobility handovers.

The inter-domain handover procedure starts when the MN detects a NAR by receiving its agent advertisement on Layer 3. The MN sends a home registration request to the new MAP via NAR. The registration request contains addressing information about the old MAP. The new MAP bi-casts the registration request to HA and the old MAP. Both of them generate a registration reply and send it to the new MAP. The reply, which arrives first, is forwarded to the MN. The second one is discarded. It is assumed that old and new MAP are generally closer to each other than new MAP and HA. Therefore, the registration reply from the old MAP should arrive faster and thus accelerate the HMIP handover procedure.

Upon reception of the handover request, the old MAP also starts forwarding of packets destined to PCoA to the new MAP. The handover latency is measured from the MN's sending of the registration request until the reception of the registration reply (it regards any registration reply as if it comes from HA).

The performance of two-way registration is compared to HMIP performance by using a C++ simulation program (the environment is not further specified). CN and MN are sending 200 bytes PCM audio packets to each other in intervals of $20ms$. The delay time of each link is determined as the sum of the transmission time (packet size /

bandwidth) and the system time (packet waiting and processing time in each router). Bandwidths and system times are set as follows:

- Wireless link bandwidth: 2 Mbps
- Intra-Domain bandwidth: 10 Mbps
- Inter-Domain bandwidth: 100 Mbps
- Wireless link system time: $20ms$
- Intra-Domain link system time: $10ms$
- Inter-Domain link system time: $8ms$

The old and the new MAP are eight hops away from the HA. The distance between the two MAPs is three hops.

A comparison plot of the home registration latency with two-way registration and with basic HMIP shows that the bi-casting of the home registration request is able to reduce the registration time by up to $90ms$ for the given scenario. The results are shown in Table 4.6.

Table 4.6: Handover Latency from Reg. Request to Reg. Reply Results presented in [16]

Nr.	**Approach Base**	**Latency** [ms]
1	HMIPv6	190
2	Two-way registration scheme	85–190

4.2 Analytical Comparison of the Handover Latencies

4.2.1 Choice of the Scenarios

This section compares the different handover approaches by using four scenarios and presenting analytical results of the latencies. In these scenarios, the communication delays (the time needed to generate, propagate and process a message between two hosts) between the participating hosts are kept constant for all protocols, thus making them comparable.

The scenarios involve an inter-domain handover of a mobile node. The topology is shown in Figure 4.2. Two domains, represented by a MAP and an AR are connected to each other and to HA and CN via Internet. In scenarios which do not involve hierarchical mobility management the MAPs act as a normal intermediate node. The assumption is that MN and CN are leading an ongoing CBR VoIP communication. The handover latency is measured from the moment in which the handover is initiated by a registration request (either by MN, PAR or NAR—see different protocols in Chapter 2) until the time the MN receives the first packet from CN via NAR. A further assumption is that the MN makes the L2 handover

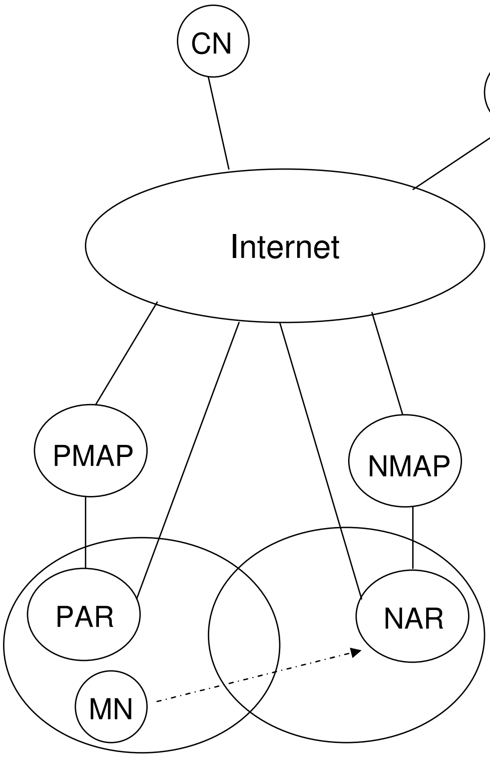

Figure 4.2: Reference Topology

immediately after no further communication of the MN with PAR is required by the handover protocol.

In reality, the communication delays between the involved communicating hosts depend on their geographic locations. We investigate four different constellations:

1.) MN is located close to HA, but far from CN. This implies a short communication delay between MN and HA and a longer delay between MN and CN. An example for this constellation would be when HA is located in the same city as MN and CN is in another city or even in a different country or on another continent.

2.) MN is located close to CN, but far from HA. There is a short communication delay between MN and CN and a long delay between MN and HA. In this case CN is in the same city and HA is at a larger distance.

3.) MN is located far from both CN and HA—a case with long communication delays to both HA and CN.

4.) MN is located close to both CN and HA—a case with short communication delays to both HA and CN.

4.2.2 Choice of Communication Delays

t_{mn_cn}	10–100ms	MN ⇔ CN
t_{mn_ha}	10–100ms	MN ⇔HA
t_{par_ha}	5–95ms	PAR ⇔ HA
t_{nar_ha}	5–95ms	NAR ⇔ HA
t_{mn_par}	2ms	MN ⇔PAR
t_{mn_nar}	2ms	MN ⇔NAR
t_{par_pmap}	2ms	PAR ⇔PMAP
t_{nar_nmap}	2ms	NAR ⇔NMAP
t_{pmap_ha}	5–95ms	PMAP ⇔HA
t_{nmap_ha}	5–95ms	NMAP ⇔HA
t_{pmap_nmap}	5–95ms	PMAP ⇔NMAP
t_{nar_pmap}	5–95ms	NAR ⇔PMAP
t_{par_nar}	10–100ms	PAR ⇔NAR
t_{L2}	20ms	L2-handover

Table 4.7: Communication Delays used in the Analysis

Table 4.7 shows the chosen ranges for the delays. We parameterized the communication delays between hosts which do not belong to the same domain. The communication delays between hosts within the same domain and the L2 handover delay are kept constant and are set in accordance to the assumptions made by Hsieh et al. in [13].

We used ping request data to choose suitable values for the communication delays between hosts from different domains. The usual round-trip-time between hosts within the city of Berlin was about 16ms. This suggests that the communication delay in one direction within the city takes about 8ms. We chose 5ms as a minimal value to include variation cases. The maximal value of 100ms was chosen by performing a ping request between a wired host within the Technical University of Berlin and the web server of the University of Berkeley (berkeley.edu). The round-trip-time was about 190ms, suggesting 95ms in one direction. We chose 100ms to include variation cases.

4.2.3 Analytical Formulas for the Handover Delay Calculations

In the following, we present the formulas used to calculate the handover delays for the different handover schemes. These include the standardized schemes presented in Chapter 2 and the non-standardized ones presented in Section 4.1.

- Standard MIPv6
  In standard MIP, the MN performs registration after switching to NAR's network, without taking any precautions. Depending on whether route optimization is used or not, two handover latencies may be the result:
  - with route optimization
    The MN has to send a binding update directly to the CN. In this case the handover

latency is:

$$t_1 = t_{L2} + 2t_{mn\_nar} + 2t_{mn\_cn} = 44 \ to \ 224ms \tag{4.1}$$

- without route optimization
  When no route optimization is used the MN updates the HA:

$$t_2 = t_{L2} + 2t_{mn\_nar} + 2t_{mn\_ha} = 44 \ to \ 224ms \tag{4.2}$$

- MIPv4 with Pre-Registration.
  In Pre-Registration the registration with NAR is initiated before MN loses connectivity to PAR (Fig. 2.1, 2.2, 2.3). We have again the cases of route optimization and no route optimization.

  - with route optimization
    MN informs CN of its new binding just before starting L2 handover. The handover delay depends only on the distance to CN.

$$t_3 = 2t_{mn\_cn} = 20 \ to \ 200ms \tag{4.3}$$

  - without route optimization
    When no route optimization is used, the MN has to update only HA.

$$t_4 = t_{mn\_par} + t_{par\_nar} + 2t_{nar\_ha} + 2t_{mn\_nar} = 26 \ to \ 296ms \tag{4.4}$$

  In both cases, the L2 handover delay doesn't play a role since the actual registration takes a longer time than the MN needs to complete handover on Layer 2.

- MIPv4 with Post-Registration.
  In Post-Registration the network takes care to organize the registration of MN with NAR before the handover. Further, PAR forwards packets to NAR where they are buffered until MN completes handover (Fig. 2.4.2). Because of this buffering the first packet from CN via NAR can be received immediately after the MN arrives at NAR. Until the MN registers with HA it already may start receiving a stream of buffered packets. The assumption is that the MN starts the L2 handover together with the L2-Source Trigger at PAR.

$$t_5 = 2t_{par\_nar} + 2t_{mn\_nar} = 24 \ to \ 204ms \tag{4.5}$$

- Predictive Fast Handover.
  In predictive fast handover the MN has to wait for a registration reply from PAR and can not start L2 handover before the registration is fully completed (Fig. 3.1, 3.3). Therefore the L2 handover delay has to be taken into account. Due to the tunneling of packets to NAR the MN may start receiving packets from CN immediately after switching to NAR, although the registration with HA is still imminent.

$$t_6 = t_{L2} + 2t_{mn\_nar} = 24ms \tag{4.6}$$

- Reactive Fast Handover
  In the reactive fast handover case the registration starts after the MN moves to the new network (Fig. 3.2).

$$t_7 = t_{L2} + 3t_{mn\_nar} + 2t_{par\_nar} + t_{mn\_nar} = 48 \ to \ 228ms \tag{4.7}$$

- HMIPv6
  HMIPv6 performs like reactive MIP in inter-domain handover scenarios plus additional $2t_{nar\_nmap}$ for the registration with NMAP.
  - with route optimization

$$t_8 = t_{L2} + 2t_{mn\_nar} + 2t_{nar\_nmap} + 2t_{mn\_cn} = 48 \ to \ 228ms \tag{4.8}$$

  - without route optimization

$$t_9 = t_{L2} + 2t_{mn\_nar} + 2t_{nar\_nmap} + 2t_{mn\_ha} = 48 \ to \ 228ms \tag{4.9}$$

- Predictive FHMIPv6
  MN may start L2 handover immediately after sending the FBU message, since the FBACK message can also be delivered by NAR (Fig. 3.5, 3.6, 3.7, 3.8).

$$t_{10} = t_{mn\_par} + t_{par\_pmap} + t_{nar\_pmap} + 2t_{nar\_nmap} + t_{nar\_pmap} + 2t_{mn\_nar} = 22 \ to \ 202ms \tag{4.10}$$

- Reactive FHMIPv6
  The handover is initialized by MN after the movement in NAR's network. NAR has to contact PMAP in order to initialize the tunneling of packets. NAR also has to contact NMAP in order to register a global address for MN, but this can happen in parallel to the communication with PMAP.

$$t_{11} = t_{L2} + 3t_{mn\_nar} + 2t_{nar\_pmap} + t_{mn\_nar} = 38 \ to \ 218ms \tag{4.11}$$

- Predictive S-MIP
  Reactive and Predictive S-MIP are not presented as such in the paper. We model the cases in which handover is initiated before and after the L2 handover. In Predictive S-MIP, the forwarding of packets begins before the MN is able to switch networks on Layer 2.

$$t_{12} = t_{L2} + 2t_{mn\_nar} = 24ms \tag{4.12}$$

- Reactive S-MIP
  In reactive S-MIP NAR would have to contact PAR to initialize the tunneling. NAR also has to contact NMAP in order to register a global address for MN, but this can happen in parallel to the communication with PAR. The procedure is the same as in Reactive Fast Handover.

$$t_{13} = t_{L2} + 3t_{mn\_nar} + 2t_{par\_nar} + t_{mn\_nar} = 48 \ to \ 228ms \tag{4.13}$$

- MIP with smooth handover (presented in [14])

$$t_{14} = 2t_{mn\_nar} + 2t_{mn\_ha} - t_{mn\_nar} + 2t_{par\_nar} + t_{mn\_nar} = 44 \; to \; 404ms \tag{4.14}$$

- FF-HMIP ([12])

$$t_{15} = t_{L2} + t_{mn\_nar} + 2t_{nar\_nmap} + 2t_{nar\_mn} + 2t_{nar\_pmap} + t_{mn\_nar} = 42 \; to \; 222ms \tag{4.15}$$

- HMIP with bi-casting (presented in [15])
  - Predictive case

$$t_{16} = t_{L2} + 2t_{mn\_nar} = 24ms \tag{4.16}$$

  - Reactive case

$$t_{17} = t_{L2} + t_{mn\_nar} + 2t_{nar\_nmap} + 2t_{nar\_mn} + 2t_{nar\_pmap} + t_{mn\_nar} = 42 \; to \; 222ms \tag{4.17}$$

- MBWA micro-macro mobility protocol protocol presented in [11].
  Like FHMIP

- Two-Way Registration (presented in [16])

$$t_{18} = t_{L2} + t_{mn\_nar} + t_{nar\_nmap} + 2t_{nmap\_pmap} + t_{nmap\_nar} + t_{mn\_nar} = 38 \; to \; 218ms \tag{4.18}$$

4.2.4 Results

The handover delays of the handover schemes, dependent on their inter-domain communication delay parameters, are depicted in Figures 4.3 to 4.9. The handover results are also shown numerically in Table 4.8 at the end of this section.

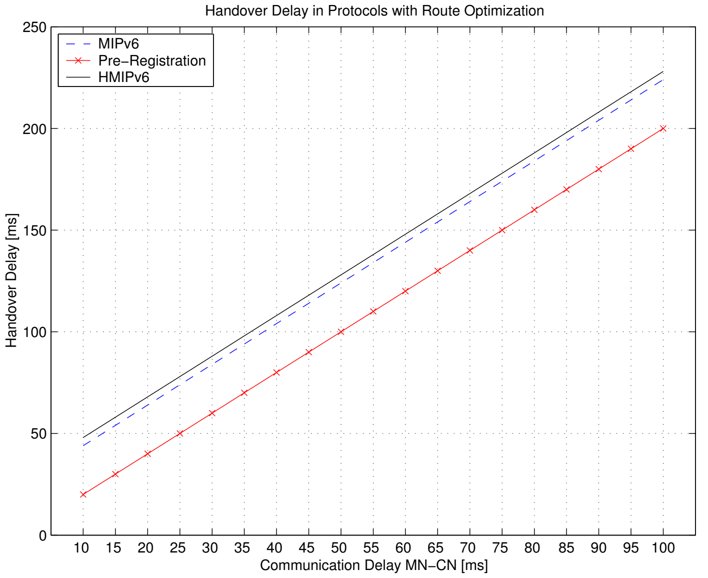

Figure 4.3: Handover Delays in Protocols with Route Optimization

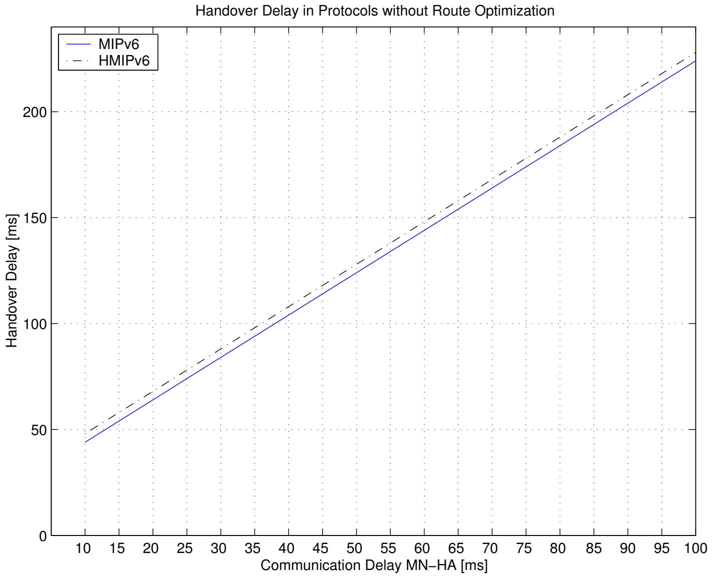

Figure 4.4: Handover Delays in Protocols without Route Optimization

Figure 4.3 shows the handover delay for the MIPv6, Pre-Registration and HMIPv6 schemes, in their versions with route optimization. The handover delay depends only on the communication delay between MN and CN. We noticed that route optimization is not advisable in all cases. This can be seen when we take a look at Figure 4.4, which shows the delay in the schemes MIPv6 and HMIPv6 in their versions without route optimization. The handover delay in this case depends on the communication delay between MN and HA. When HA is close to MN while CN is far away, a handover will finish faster when no route optimization is applied. The reverse is true when CN is close to MN while HA is far from it.

The HMIPv6 protocol reduces the signaling costs during handovers in general. However, it does not bring a major improvement with respect to seamless service continuation during inter-domain handovers, since a change of the MAP is required. In this case the scheme falls back to standard Mobile IP. It even introduces a further delay in comparison to the standard protocol, since NAR has to contact NMAP to obtain a regional CoA for MN. Pre-Registration performs best from this group, since it initiates the handover signaling before performing the L2 handover.

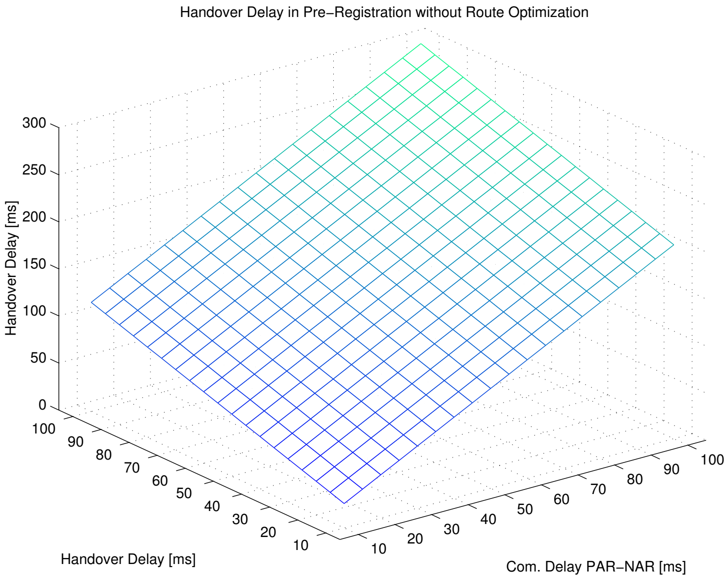

Figure 4.5: Handover Delay in Pre-Registration without Route Optimization

Figures 4.5 and 4.6 show the handover delay performance of schemes which depend on two inter-domain communication delay parameters. The schemes are Pre-Registration (Figure 4.5) and MIP with smooth handover (Figure 4.6) in their versions without route optimization. The dependency is linear on both axes. It is interesting to observe that the PAR–NAR communication delay can significantly increase the handover delay in inter-domain handovers. Whereas in an intra-domain handover the communication delay between PAR and NAR is assumed to be small (in the range of one to two milliseconds), this assumption can not be made when PAR and NAR belong to different domains. This fact is fortified by the possibility of PAR and NAR to use different access technologies, which can further influence

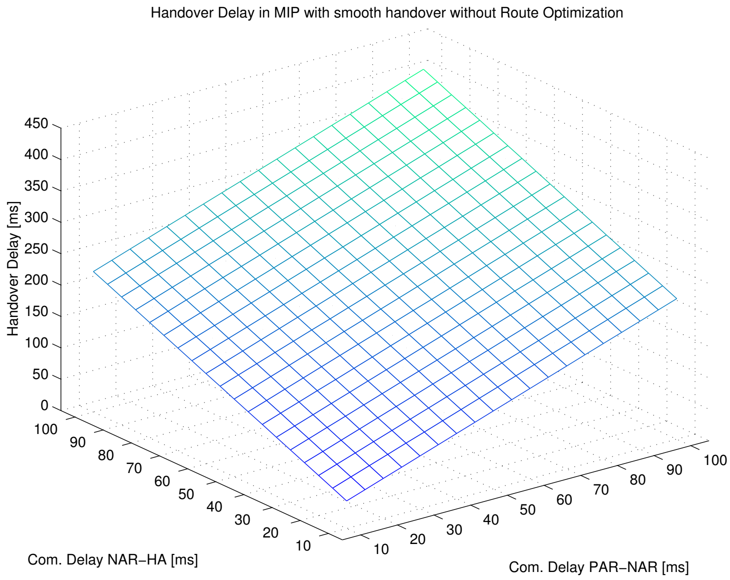

Figure 4.6: Handover Delay in MIP with smooth handover without Route Optimization

the handover. The dependency on communication delays between hosts from both domains is also visible in Figures 4.7, 4.8 and 4.9. Figure 4.7 shows the handover delay achieved by the Post-Registration and Reactive FMIPv6 schemes. The handover delay in these schemes depends on the communication delay between PAR and NAR. The Reactive FMIPv6 curve is also representative for the Reactive S-MIP scheme, since they share the same formula. Figure 4.8 shows the handover delay in the Two-way registration scheme. It depends on the PMAP–NMAP communication delay. Figure 4.9 depicts the handover delay in schemes which depend on the NAR–PMAP communication delay—Predictive FHMIPv6, Reactive FHMIPv6, FF-HMIP, Predictive and Reactive MBWA micro-mobility protocol and reactive HMIP with bi-casting. Predictive FHMIPv6 performs best in this group, since the MN may perform the L2 handover immediately after initiating the handover signaling. The binding acknowledgment is buffered and delivered by NAR when MN connects to it.

The dependency on communication delays between two domains shows that it is important to consider all communication delays between hosts belonging to these domains when dealing with inter-domain handovers.

The best performance in terms of handover delay is shown by the pure FMIPv6 protocol in its predictive version, predictive S-MIP, and predictive HMIP with bi-casting. All of them achieve a constant delay of $24ms$ (Table 4.8).The blackout time in these schemes depends only on intra-domain communication delays, since the inter-domain signaling is completed before the L2 handover. Therefore, they achieve a constant delay in all inter-domain handover scenarios. The important features that put these three schemes on top of the list are the tunneling functionality between old and new domain and the fact that the time-consuming inter-domain signaling is completed before the L2 handover.

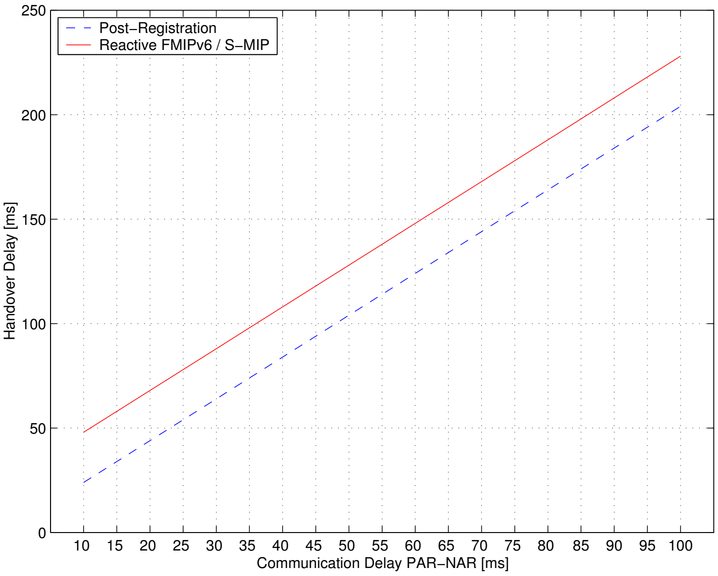

Figure 4.7: Handover Delays in Protocols depending on the PAR-NAR Communication Delays

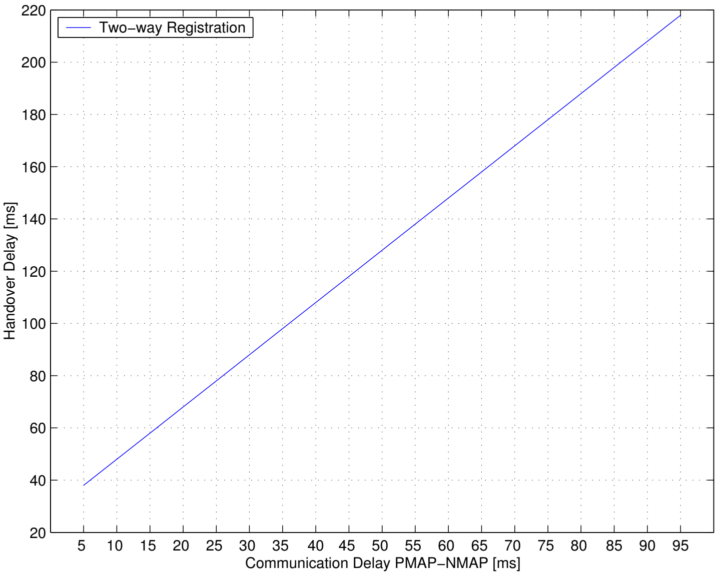

Figure 4.8: Handover Delays in Protocols depending on the PMAP-NMAP Communication Delays

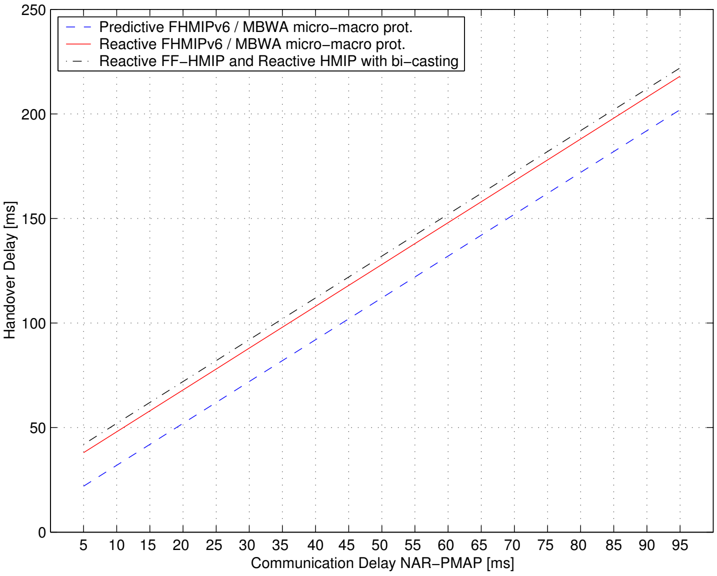

Figure 4.9: Handover Delays in Protocols depending on the NAR-PMAP Communication Delay

Table 4.8: Comparison of Inter-Domain Handover Approaches

Nr.	**Approach**	**Features**	**Latency [ms]**	**Dependency on inter-domain com. delays**
1	MIP	route opt.	44–224	MN–CN
2	MIP	NO route opt.	44–224	MN–HA
3	Pre-Registration	route opt.	20–200	MN–CN
4	Pre-Registration	NO route opt.	26–296	PAR–NAR, NAR–HA
5	Post-Registration		24–204	PAR–NAR
6	FMIPv6	predictive	24	only intra-domain dependency
7	FMIPv6	reactive	48–228	PAR–NAR
8	HMIPv6	route opt.	48–228	MN–CN
9	HMIPv6	NO route opt.	48–228	MN–HA
10	FHMIPv6	predictive	22–202	NAR–PMAP
11	FHMIPv6	reactive	38–218	NAR–PMAP
12	S-MIP	predictive	24	only intra-domain dependency
13	S-MIP	reactive	48–228	PAR–NAR
14	MIP with smooth handover [14]		44–404	PAR–NAR, NAR–HA
15	FF-HMIP [12]		42–222	NAR–PMAP
16	HMIP with bi-casting [15]	predictive	24	only intra-domain dependency
17	HMIP with bi-casting [15]	reactive	42–222	NAR–PMAP
18	MBWA micro-macro mobility protocol [11]		like FHMIP	like FHMIP
19	Two-Way Registration [16]		38–218	PMAP–NMAP

Chapter 5

Summary and Conclusions

This report discussed the existing Layer 3 handover approaches with respect to their blackout times during inter-domain handovers. We first presented extensions to the standard Mobile IP protocol devised by IETF supporting low delays in handover situations. These included Pre-Registration, Post-registration as extensions to HMIPv4 and HMIPv6, FMIPv6, and their non-standardized combination FHMIPv6, as extensions to HMIPv6. We further investigated the published, non-standardized extensions and enhancements suggested by other research groups. The selection of publications was made from the perspective of our main area of interest—low Layer 3 blackout time during inter-domain handovers.

Firstly, each protocol was presented with a description of its operation and assumptions. Secondly, we contributed a simple analytical comparison of all presented handover protocols. The comparison was done by means of four inter-domain handover scenarios with different delays between involved entities. The results have shown that the handover performance of Mobile IP can be improved significantly when the basic protocol is extended by additional functionality. HMIPv6 reduces the signaling overhead during a handover, since MN communicates only with MAP. FMIPv6 reduces the handover delay and the packet loss during a handover, since packets destined to MN are forwarded through a tunnel from PAR to NAR, while MN is updating its communication partners. A combination of HMIPv6 and FMIPv6, called FHMIPv6 integrates the advantages of both approaches. However, HMIPv6, FMIPv6, and their combination, are specified for intra-domain handovers. For inter-domain handovers, a solution to this problem is found in providing tunneling functionality between appropriate nodes in the involved domains—usually PAR–NAR, PMAP–NMAP or PMAP–NAR. Another interesting option is the bi-casting of packets to two ARs. In this approach, the current MAP of the MN would forward packets to both PAR and NAR so that MN can always receive packets—independently from the current status of the handover.

Whereas the HMIPv6 approach is reasonable and offers an enhancement to Mobile IP, it is the FMIPv6 part of the FHMIPv6 combination which brings the major improvement to the handover performance, with respect to the seamless service continuation. The tunneling of packets during handover significantly reduces the blackout time during which the MN is not able to receive packets. In our analysis, the best performance in terms of handover delay was shown by the pure FMIPv6 protocol in its predictive version and S-MIP—a combination of HMIPv6, FMIPv6 and bi-casting suggested by Hsieh et al. in [10]. As expected, the handover delay in inter-domain scenarios further depends on the constellation of the participating entities. Thereby, not only the communication delays between MN, HA and CN play a major role, but also the delays between PAR, NAR, PMAP and NMAP.

The last part of the technical report provides a short overview of existing Mobile IP simulation models for the ns2 simulation environment (see Appendix A). A more detailed description is provided for the FHMIP extension (Hsieh et al. [17]). It provides a lot functionality and gives the possibility to combine standard MIP, FMIP, HMIP and HMIP+FMIP.

Bibliography

[1] D. Johnson, C. Perkins, and J. Arkko. "IP Mobility Support in IPv6". http://www.ietf.org/rfc/rfc3775.txt, June 2004. IETF.

[2] K. Malki. "Low Latency Handoffs in Mobile IPv4" (work in progress). http://www.ietf.org/internet-drafts/draft-ietf-mobileip-lowlatency-handoffs-v4-11.txt, October 2005. IETF.

[3] R. Koodli. "Fast Handovers for Mobile IPv6". http://www.ietf.org/rfc/rfc4068.txt, July 2005. IETF.

[4] C. Perkins. "IP Mobility Support for IPv4". http://ietf.org/rfc/rfc3344.txt, August 2002. IETF.

[5] C. Perkins. "Mobile Networking Through Mobile IP". *IEEE Internet Computing*, Volume 2, Issue 1:58 – 69, Jan.-Feb. 1998.

[6] C. Perkins. "Mobile IP". *IEEE Communications Magazine*, Volume 35, Issue 5:84 – 99, May 1997.

[7] H. Levkowetz, S. Vaarala, and Netsael. "Mobile IP Traversal of Network Address Translation (NAT) Devices". http://rfc3519.x42.com/34402.html, April 2003. IETF.

[8] H. Soliman. "Hierarchical Mobile IPv6 Mobility Management (HMIPv6)". http://www.ietf.org/rfc/rfc4140.txt, August 2005. IETF.

[9] H. Jung. "Fast Handover for Hierarchical MIPv6 (F-HMIPv6)", (work in progress). http://tools.ietf.org/wg/mipshop/draft-jung-mipshop-fhmipv6-00.txt, October 2005. IETF.

[10] R. Hsieh, Z.G. Zhou, and A. Seneviratne. "S-MIP: a seamless Handoff Architecture for Mobile IP". In *INFOCOM 2003. IEEE Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies.*, volume 3, pages 1774 – 1784, 30 March-3 April 2003.

[11] J. Chow and G. Garcia. "Macro- and Micro-Mobility Handoffs in Mobile IP based MBWA Networks". In *IEEE Global Telecommunications Conference, 2004. GLOBECOM '04.*, volume 6, pages 3921 – 3925, 29 Nov.-3 Dec. 2004.

[12] Y. Gwon, J. Kempf, and A Yegin. ”Scalability and Robustness Analysis of Mobile IPv6, Fast Mobile IPv6, Hierarchical Mobile IPv6, and Hybrid IPv6 Mobility Protocols using a Large-Scale Simulation”. In *2004 IEEE International Conference on Communications*, volume 7, pages 4087 – 4091, 20 March-24 June 2004.

[13] R. Hsieh and A. Seneviratne. ”A Comparison of Mechanisms for Improving Mobile IP Handoff Latency for End-to-end TCP”. In *MobiCom 2003*, 14-19 September 2003.

[14] T.T. Kwon, M. Gerla, and S. Das. ”Mobility management for VoIP service: Mobile IP vs. SIP”. *IEEE Wireless Communications*, Volume 9, Issue 5:66 – 75, Oct. 2002.

[15] I. Vivaldi, M.H. Habaebi, B.M. Ali, and V. Prakesh. ”Fast Handover Algorithm for Hierarchical Mobile IPv6 Macro-Mobility Management”. In *The 9th Asia-Pacific Conference on Communications. APCC 2003.*, volume 2, pages 630 – 634, 21-24 Sept 2003.

[16] J. Zhang, D.A.J. Pearce, and T.C. Tozer. ”Two-way Registration: a Fast Handoff Scheme for IPv4 Macro-Mobility Management”. In *15th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, volume 2, pages 1252 – 1256, 5-8 Sept 2004.

[17] R. Hsieh. Fhmip ns extension. http://mobqos.ee.unsw.edu.au/ robert/opcomm/.

[18] The network simulator ns2. http://www.isi.edu/nsnam/ns.

[19] The ns Manual. http://www.isi.edu/nsnam/ns/ns-documentation.html.

[20] A. Festag. ”Optimization of Handover Performance by Link Layer Triggers in IP-Based Networks; Parameters, Protocol Extensions, and API for Implementation, Technical Report”. August 2002.

[21] X. Fu, T. Chen, A. Festag, G. Schaefer, and Karl H. ”SeQoMo Architecture: Interactions of Security, QoS and Mobility Components”. April 2002.

[22] T. Chen, S. Hermann, and Schaefer G. ”Secure, QoS-enabled Mobility Support in All-IP Networks, technical Report”. Juni 2003.

[23] C. Perkins. Mobins2 extension to ns2. http://people.nokia.net/ charliep/mobins2/.

[24] Columbia University. Columbia IP Micro-Mobility Software. http://www.comet.columbia.edu/micromobility/.index.html.

[25] J. Widmer. No ad hoc routing agent. http://icapeople.epfl.ch/widmer/uwb/ns-2/noah/.

[26] T. Ernst. Mobiwan: A ns simulation platform for mobile ipv6 in wide area networks. http://www.ti-wmc.nl/mobiwan2/.

[27] CMU Monarch Group. Wireless and mobility extensions to ns. http://www.monarch.cs.cmu.edu/cmu-ns.html.

[28] I.F. Akyildiz, Jiang Xie, and S. Mohanty. "A Survey of Mobility Management in Next-Generation All-IP-based Wireless Systems". *IEEE Wireless Communications*, Volume 11, Issue 4:16 – 28, Aug. 2004.

[29] W. Mohr and W. Konhäuser. "Access Network Evolution Beyond Third Generation Mobile Communications". *IEEE Communications Magazine*, Volume 38, Issue 12:122 – 133, Dec. 2000.

[30] Helsinki University of Technology. The Dynamics Mobile IP system. http://dynamics.sourceforge.net/.

Appendix A

Existing Mobile IP Simulation Models in NS

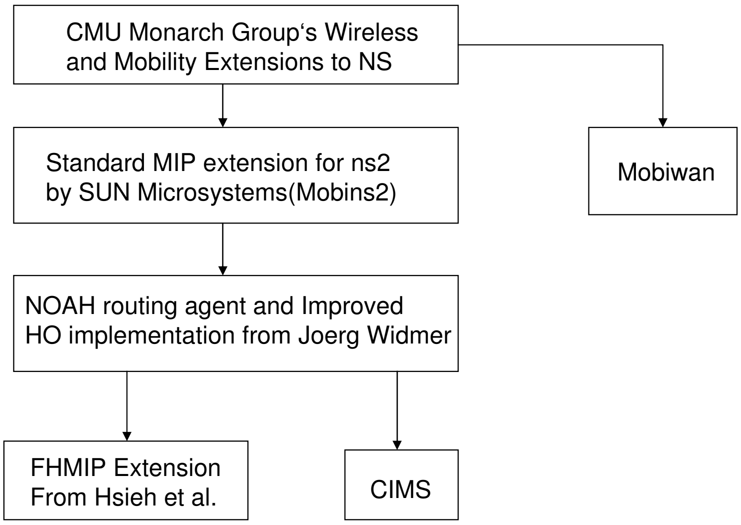

Figure A.1: Dependency of the Extensions

A.1 Wireless and Mobility Extensions to ns2

Wireless functionality in ns was first provided by the CMU Monarch's "Wireless and Mobility Extensions to ns2" [27]. The extensions were developed in 1998 and since then integrated in the mainline ns code.

Mobile nodes in ns are independent entities which are able to compute their own position and velocity as a function of time. Each mobile node can have one or more network interfaces, each of which is attached to a *channel*. Channels are the conduits that carry packets between mobile nodes. When a mobile node transmits a packet on the channel, the channel distributes a copy of the packet to all the other network interfaces on the channel. These interfaces then use a *radio propagation model* to determine if they are actually able to receive the packet. Special *Base Station (BS)* nodes act as gateways between wired and wireless nodes, so that packets can be routed between wired and wireless topologies.

Here is a short list of the features provided by the Wireless and Mobility extensions, organized by ISO network stack layer:

Physical Layer

- Radio Propagation Models - free space, two ray ground and shadowing models.
- Antennas - unity gain omni-directional antenna.
- Network interfaces - shared media interface. This network interface implements a shared media model where, subject to collisions and the propagation model, each node can overhear packets transmitted by others.

Link Layer

On the Link Layer the extensions implement models of the IEEE 802.11 Distributed Coordination Function (DCF) MAC protocol.

Network Layer

On the network layer the extensions implement four routing protocols - Dynamic Source Routing (DSR), Destination Sequence Distance Vector (DSDV), Temporally Ordered Routing Protocol (TORA), ad hoc on demand Distance Vector (AODV).

A.2 Mobile IP Extension for ns

There is a basic Mobile IP extension to ns, which is an integral part of current versions of the simulator. It was developed by SUN Microsystems (Perkins et al. [23]) as the "Mobins2 extension".

The MIP scenario consists of MHs moving between their HAs and FAs. HAs and FAs are BS nodes which routinely send beacons out to the MHs, set up encapsulators and decapsulators as necessary and reply to solicitation messages from MHs. MHs receive and respond to beacons and send out solicitation messages to search for BSs. A solicitation message from a MH provokes the generation of an advertisement message (ad) that is sent directly to the requesting MH. The address of the BS sending out the beacon is used as the CoA of the MH. Thus, as the MH moves from its native to foreign domains, its CoA changes. MHs reply to ads from BSs by means of a registration request. Upon receiving a registration request from an MH the BS checks to see if it is the HA for the MH. If not, it sets up its decapsulator and forwards the registration request towards the HA of the MH.

In case the BS is the HA for the requesting MH but the CoA does not match its own, it sets up an encapsulator and sends a registration request reply to the CoA who has forwarded the registration request. This way, the HA now has a binding of the MH's CoA to which it tunnels any incoming packets for the MH. For this purpose the packets are encapsulated in a packet destined to the FA. The FA's decapsulator receives the packet, removes the encapsulation and sends the inner packet to the MH.

In case the CoA in the registration request matches the address of the HA, it just removes the encapsulator it might have set up—the registration request means that the MH has just come back from some roaming. The HA sends a reply directly to the MH. If the MH does not hear any ads from BSs it starts to send out solicitations. Upon receiving ads , it changes its CoA to the address of the HA/FA it has heard the ad from, and replies back to the CoA with a registration request.

These and other details about the Mobile IP implementation can be found in the ns manual [19].

The MIP extension does not consider any of the other handover protocols.

A.3 NOAH

NOAH [25] is a wireless routing agent that only supports direct communication between BSs and mobile nodes. This allows for the simulation of scenarios where multi-hop wireless routing is undesired.

NOAH also brings some improvements to the basic MIP implementation—support for overlapping service areas of BS and improved handover mechanism through intelligent selection of foreign agents. The NOAH routing agent is used as a basis for two further extensions described in the following sections.

A.4 FHMIP ns Extension

The FHMIP ns extension developed by Hsieh et al. [17] offers models for most of the handover protocols presented in Section 2. It supports MIP, HMIP and the combinations MIP+FMIP and HMIP+FMIP. The implementation is based on the standard MIP and NOAH extensions. The code is extended by adding a special MAPAgent and fast handover functionality. The MAPAgent can be attached to a usual wired node, thus making it a MAP. The MAP acts as a hop between HA and the current FA of the MN. The HA encapsulates packets destined to MN and tunnels them to the MAP. MAP decapsulates the packets and encapsulates them again, by using the address of the FA. Finally the FA decapsulates the packet and delivers it to the MN. The type of the simulation can be adjusted with #define statements in the source code and requires recompilation of ns2 (file mip-reg.cc). There are three types of nodes, whose functionality is implemented in mip-reg.cc—Base-station nodes, MAP nodes and Mobile Nodes.

A.4.1 FHMIP Extension Example Script

The FHMIP extension provides an example script with which one is able to simulate the different scenarios (MIP, HMIP, MIP+FMIP and HMIP+FMIP) depending on the settings in mip-reg.cc . The example script is called simula.tcl. The topology used is the same shown

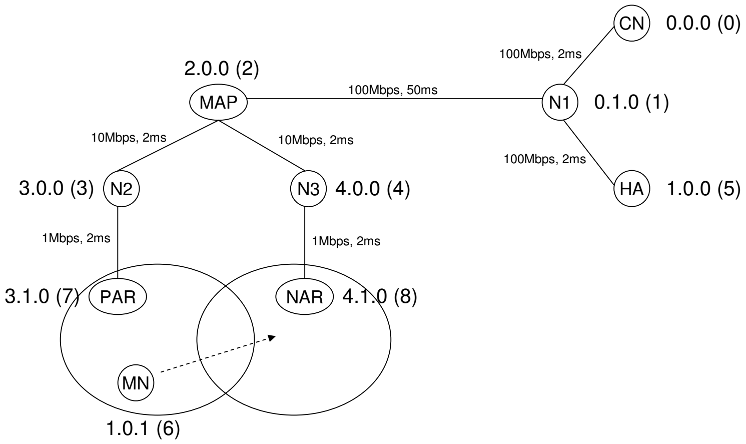

Figure A.2: Example Network Topology with Hierarchical Addresses of the Nodes

in Fig. A.2. All nodes possess a hierarchical address. There are 5 domains - the distribution of the nodes in the domains is shown in Fig. A.3. In the beginning of the simulation the

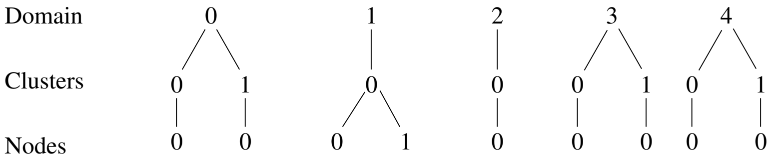

Figure A.3: Node Distribution in Domains

MH is close to HA and starts an ftp session with CN. Six seconds into the simulation the MN is moved abruptly in the vicinity of PAR. We had to change the script to do this with the *setdest* command instead of just changing the position of the node since ns-2.28 was not updating the topological lists properly if the node was moved this way. Ten seconds into the simulation MN starts to move with 1 m/s towards NAR. Until this point in time MN has already registered its CoA with the HA.

HA, PAR and NAR send advertisement beacons (ads) every second. This functionality is provided by the standard Mobile IP implementation. In case no BS is in sight the MN would send solicitation messages, but this case does not appear in this example. Every second, when the MN receives an ad it initiates a registration with the sending BS. It does not matter whether it is already registered with the same BS. The registration is then propagated to HA, also every second. This is actually not standard MIP behavior (according to RFCs), but the standard Mobile IP extension works this way.

The different handover scenarios behave differently when the MN reaches the range of NAR (PAR's and NAR's ranges overlap partly):

1.) MIP.
    As long as MN receives ads from PAR it doesn't react to ads from NAR. When the

connection to PAR is lost (3 seconds timeout), MN sends a registration request to NAR and changes its CoA. In the example scenario the connectivity to PAR went lost one second before the timeout came.

2.) MIP with priority handover.
In this scenario, priorities are assigned to the BSs. In case that a NAR has a higher priority than PAR, the handover is initiated immediately. In the example scenario, NAR had a priority of 4 and PAR a priority of 3. When the first ad from NAR reached MN, it initiated the handover immediately. However, MN had also initiated the update with PAR (like every second) and the next message it received was a confirmation of its old CoA. For some reason it wouldn't accept the new CoA from NAR which came slightly later. These double registrations continued until the MN lost connectivity to PAR and registered successfully with NAR. It turned out that a boolean variable "force_handoff" in mip-reg.cc was not set correctly when handover was initiated. After changing the code the handover ran as expected - MN abandoned its old CoA and registered with NAR. The code change is documented in the installation instructions in Section B.1.

   The registration problem pointed to another interesting behavior. When the MN starts registration with a NAR it ceases receiving packets from PAR. It actually receives them, but drops them intentionally. The reasoning of the developers is that this behavior simulates channel switching. PAR and NAR are supposed to communicate on different channels and, therefore, MN should not be able to receive from PAR during a handover procedure. As soon as MN initiates handover it sets a recvVerifier variable and accepts data packets only from NAR.

3.) HMIP.
The HMIP case behaves as expected. Initially there was the same problem with the registration as in the priority MIP case, but it was resolved after the code change. In this scenario also, the registration of MN is updated every second and is propagated to MAP and HA.

4.) FMIP.
In this scenario the MN starts registration immediately after it receives an ad from NAR (like in priority MIP and HMIP). The registration is initiated via PAR, by the MN sending a RTSOLPR message to it. PAR and NAR then exchange HI–HACK messages and build a tunnel. PAR responds to MN with a PRRTADV message. Then MN sends a registration request to NAR which is forwarded to HA.

   After the tunnel has been built PAR starts forwarding all incoming packets for MN to NAR. Additionally it also broadcasts them on the medium. However, MN drops all packets as long as they come from PAR—to simulate channel switching as described before. Further, NAR has no buffering capabilities, so all packets received at PAR before MN's registration with NAR is completed, are lost.

5.) FHMIP.
The FHMIP functionality is a mix of the FMIP functionality of the extension and the F-HMIPv6 draft [9] (see also Section 3.3). After hearing the ad from NAR, MN sends a RTSOLPR message to PAR. Instead of forwarding the message to MAP (F-HMIPv6

conform) PAR and NAR make a HI–HACK exchange (like FMIP). This is not necessary since they are not going to build a tunnel. Then PAR sends the PRRTADV to MN and MN sends a registration request to NAR. NAR forwards the request to MAP upon which MAP starts forwarding packets destined to MN to NAR. This is not really a tunnel which reduces packet loss since the forwarding starts when the registration is completed (when it would start anyway).

A.4.2 Summary of the FHMIP Extension

This extension offers a good basis for development of Layer 3 Handover models. It already combines the Handover protocols devised by IETF, although some corrections seem to be necessary. Unfortunately, the code doesn't support S-MIP although its designers and the developers of the extension are the same (Hsieh et al. [17], [10], [13]). S-MIP can probably be integrated with little effort.

The extension extends the mainline ns code directly, but keeps the possibility to simulate with the original code (SUN's MIP) through options (the #define statements in mip-reg.cc).

A.5 MobiWan

The Mobiwan project [26] (not active anymore), was concerned about studying mobility of nodes in Wide-area IPv6 networks. The goal was to simulate local mobility (within a single administrative domain) and global mobility (across domain boundaries or sites). The intention of the developers was to simulate MIPv6 and HMIPv6. However, HMIPv6 was not implemented before the end of the project. The project branches off the mainline ns wireless extension from CMU. (Figure A.1).

A main concern of the project was the simulation of large Internet topologies consisting of hundreds of nodes. The first problem was the configuration and management of large topologies in ns. The nodes in ns move in a bounded geographical grid (geographical movement), whereas WAN mobility is more concerned by topological movements. The authors wanted to use the "Georgia Tech Internetwork Topology Models (GT-ITM)" to model large WAN topologies. This is a collection of routines to generate and analyze graphs using a wide variety of models for internetwork topology. The existing translator from GT-ITM to ns was not suitable to manipulate large topologies. The first set of enhancements made by the Mobiwan extensions addressed these issues. It provides a set of tools for topology creation, manipulation and scenario configuration called TOPOMAN, TOPOGEN and SCEN TOOLS. It further extends the NS addressing hierarchy from 3 levels to 4 levels.

The second set of enhancements was meant to deal with WAN Mobility—provide global mobility functionalities and implement the mobility management protocols MIPv6 and HMIPv6. The code does not extend the standard MIP code of ns directly, but seems to use some of it (for example the encapsulation/decapsulation modules). According to the documentation provided in the distribution the MIPv6 implementation offers (at least partially implemented) several features which are not included in standard MIP:

- Binding Updates.
  The MN sends BUs when it obtains a new CoA and when the periodic timer has expired.

The MN sends a BU to all nodes registered in its *Binding Update List* for which the entry is activated by an activation flag.

- Binding Cache and Route Optimization.
  When a node receives a BU, an entry is added in its Binding cache. As a result, the routing table (a set of classifiers) is updated to redirect packets for this node to the CoA specified in the BU (using a routing extension header or encapsulation). The usage of BUs and Binding Caches ends up in route optimization—an option which is missing in standard MIP.

The implementation of HMIPv6 was planned, but not conducted at the end.

The third enhancement of the extension is an implementation of global mobility, i.e. a MN moving from one site to another. All sites are associated with the same geographical grid, but with a distinct channel. As a result, all BSs and MNs in the same site communicate on the same channel, and have the same address prefix. Thus, a MN might change its topological location by switching to a new channel. This is a nice feature which might help to resolve the packet dropping problem in the FHMIP extension. Per documentation global mobility works with only one MN and needs changes to make it work with more MNs.

The most effort in this extension seems to have been spent in the development of the tools for generation and management of large wireless topologies. This is not closely related to the handover latency problem at hand. Most of the offered MIP functionality is already included in the mainline ns code through SUN's extension. Parts of the code seem to be unfinished and not properly tested (the author warns of bugs in several places). Nevertheless, the Binding Update, Binding Cache and triangular routing features could be interesting as an add-on to another model. The idea to implement global mobility through channel switching is very interesting, especially from the perspective of handovers in heterogeneous networks.

A.6 CIMS

CIMS is an extension which offers micro-mobility support [24]. It implements HMIP and two micro-mobility protocols: Cellular IP and Hawaii. Both are not interesting for our research since they address strictly intra-domain handover solutions and use basic Mobile IP for inter-domain handovers.

The extension builds upon the standard MIP implementation of ns and the NOAH routing agent extension. No details about the implemented HMIP protocol were to be found in the documentation.

Appendix B

B.1 Installation of the FHMIP Extension under ns-2.28

The FHMIP extension was developed on ns-2.1b7a. We were able to port the extension to ns-2.28. The extension also needed an installation of the NOAH routing agent.

B.1.1 Installation of NOAH

The NOAH routing extension [25], was implemented on ns-2.26 but the installation went without problems on ns-2.28. The following description of the installation steps is copied from the project's web site:

```
Makefile.in:    add noah/noah.o to OBJ_CC and
tcl/mobility/noah.tcl:  to NS_TCL_LIB
 noah/noah.{h,cc}:  add noah.h and noah.cc to a new subdirectory noah/
tcl/mobility/noah.tcl:  add noah.tcl to tcl/mobility/
tcl/lib/ns-lib.tcl.h:
line 191: add source ../mobility/noah.tcl
line 603ff: add

            NOAH {
                    set ragent [$self create-noah-agent $node]
            }

line 768ff: ad

Simulator instproc create-noah-agent { node } {
    # Create a noah routing agent for this node
    set ragent [new Agent/NOAH]

    ## setup address (supports hier-addr) for noah agent
    ## and mobilenode
    set addr [$node node-addr]\\
\newline
    $ragent addr $addr\\
```

```
    $ragent node $node\\

    if [Simulator set mobile_ip_] {
        $ragent port-dmux [$node demux]
    }
    $node addr $addr
    $node set ragent_ $ragent
    return $ragent
}
```

B.1.2 Installation Steps for the FHMIP Extension

The extension package consists of the following files:

- Makefile.in
- mip.h
- mip.cc
- mip-reg.cc
- fasthandover.h
- fasthandover.cc
- packet.h
- a tcl/lib directory containing the files
  - ns-agent.tcl
  - ns-default.tcl
  - ns-lib.tcl
  - ns-node.tcl
  - ns-packet.tcl

The Makefile.in in the package can be ignored since it is for ns-2.1b7a only and can not be used on ns-2.28.

We made a backup of the ns-2.28 versions of the above files. We also downloaded ns-2.1b7a to make differential comparisons to ns-2.28 and proceeded as follows:

1.) mip.{h,cc} and fasthandover.h can be copied directly to ns-2.28/mobile as is.

3.) For mip-reg.cc some small changes are required. Here is the output of 'diff -wB' of the current version of ns-2.28/mobile/mip-reg.cc and the mip-reg.cc from the extension package.

```
148c146
<   const char *objname = NULL;
---
>    char *objname = NULL;
834c832
<       h->lifetime_ = MIN(reglftm_, agt->lifetime_);
---
>             h->lifetime_ = min(reglftm_, agt->lifetime_);
1002d999
<               force_handoff_ = 1; // <SY>
1021d1017
<               force_handoff_ = 1; // <SY>
1096d1091
<               force_handoff_ = 1; // <SY>
1104d1098
<               force_handoff_ = 1; // <SY>
```

The lines with <SY> comment are added to resolve the registration problem mentioned in A.4.1.

4.) ns-2.28/Makefile.in: include mobile/fasthandover.o to OBJ_CC

5.) ns-2.28/common/packet.h has to be extended by the new packet types. Simple cut and paste doesn't suffice since packet.h has been altered in the mainline ns code. Here is a 'diff -wB' comparison of packet.h after the changes and the unaltered backup version in ns-2.28:

```
37,38d36
< /* modified by Robert Hsieh (UNSW) Jan. Feb. 2003 */
<
169,174d166
<
<       PT_PRRTADV,                         /* RCH */
<       PT_RTSOLPR,                         /* RCH */
<       PT_HI,                              /* RCH */
<       PT_HACK,                            /* RCH */
<
266,271d257
<               // Fast Handover
<               name_[PT_PRRTADV] = "ProxyRouterAdvertisement";
<               name_[PT_RTSOLPR] = "RouterSolisitationProxy";
<               name_[PT_HI] = "HandoverInitiation";
<               name_[PT_HACK] = "HandoverAcknowledgement";
```

1 The files
ns-2.28/tcl/lib/ns-agent.tcl

ns-2.28/tcl/lib/ns-default.tcl
ns-2.28/tcl/lib/ns-lib.tcl
ns-2.28/tcl/lib/ns-mip.tcl
ns-2.28/tcl/lib/ns-node.tcl
ns-2.28/tcl/lib/ns-packet.tcl
are changed in the same way as packet.h: functionality from the extension package versions, which is not available in the ns-2.28 versions has to be included, but the files should not be overwritten directly.