Universität Karlsruhe (TH)
Institut für Telematik

TELEMATICS TECHNICAL REPORTS

A Comprehensive Delay Analysis for Reactive and Proactive Handoffs with Mobile IPv6 Route Optimization

Christian Vogt
chvogt@tm.uka.de

January 6, 2006

TM-2006-1

ISSN 1613-849X

`http://doc.tm.uka.de/tr/`

Institute of Telematics, University of Karlsruhe
Zirkel 2, D-76128 Karlsruhe, Germany

Abstract

Optimizations to reduce handoff delays inherent in Mobile IPv6 Route Optimization as well as IPv6 router discovery, address configuration, and movement detection have so far been mostly considered on an individual basis. This document evaluates three integrated solutions for improved handoff experience in surroundings with different preconditions: reactive handoffs with unmodified routers, reactive handoffs with router support, and movement anticipation and proactive handoff management.

1 Introduction

A mode for *Route Optimization* was incorporated into the Mobile IPv6 [1] mobility protocol in an effort to better support applications with real-time requirements on propagation latencies [2]. Route Optimization allows peers, of which either or both may be mobile, to communicate via a direct routing path. This complements the classic approach of routing a mobile node's traffic through a stationary proxy, its *home agent*. However, the problem with Route Optimization is that the reduction in propagation latencies comes at the cost of increased handoff delays. Those are substantial enough to effectively *outrule* meaningful mobility support for real-time applications [3, 4, 5]. Separate from the delays in the mobility protocol are those stemming from the standard IPv6 protocol suite. Delays for router discovery, address configuration, and movement detection are in fact in the order of seconds [6, 7, 8, 9].

A multitude of optimizations have therefore recently been put forth to streamline handoff-related activities and reduce handoff delays [6, 7, 8, 9, 10, 11, 12, 13, 14]. And while mobility support at IP layer has usually been considered a response to link-layer handoff, some of these optimizations facilitate anticipation of movements and proactive handoff preparation [15, 16, 17, 18]. Unfortunately, the optimizations have been studied mostly on an individual basis, and an evaluation of how well they integrate has so far been neglected [14]. This is although a comprehensive reduction of handoff delays can only be achieved with a combination of different optimizations, fine-tuned to seamlessly interoperate with each other.

This document explains the overall handoff procedure in a standard IPv6 deployment from an IP layer's perspective and analyzes to which extent it falls short of expectations. Since the results strongly advise optimization, the document proceeds to present and explore promising enhancements that have recently gained momentum in both the Internet Engineering Task Force and the academic research community. Those are evaluated with respect to their interactions. The document finally proposes three integrated solutions for improved handoff experience in surroundings with different preconditions: reactive handoffs with unmodified routers, reactive handoffs with router support, as well as movement anticipation and proactive handoff management. The document concentrates on mobile nodes with a single interface, although the presented solutions could be conveyed to multi-interfaced mobile nodes as well.

2 Handoffs with Standard IPv6 Neighbor Discovery

A mobile node undergoes an IP-layer handoff, or simply a *handoff*, when it changes IP connectivity. This begins with a change in link-layer attachment, also referred to as a *link-layer handoff*, and includes the discovery of new routers, address configuration, movement detection, and finally Mobile IPv6 registrations. Next is a description of the handoff procedure if routers in the mobile node's visited networks operate the standard IPv6 Neighbor Discovery protocol. This is followed by an analysis of handoff delays caused at IP layer. Additional link-layer latencies and processing delays internal to the nodes are ignored. It is assumed that each link has a single router. Handoff delays may be shorter on links with multiple routers because routers apply rate-limitation and desynchronization delays independently of each other. It is further assumed that the mobile node communicates with a single correspondent node. This simplification will be maintained throughout the document unless the number of correspondent nodes is of particular importance. Figure 1 depicts the entire handoff procedure; table 1 summarizes the performance analysis.

2.1 Router Discovery

A mobile node learns about local routers and on-link prefixes during router discovery. This process is facilitated through Router Advertisement messages, which routers multicast to link-local nodes on a loosely periodic basis. The mobile node may listen for advertisements or, if it is unwilling to wait, actively request one by sending a Router Solicitation message. The IPv6 Neighbor Discovery RFC [19] permits a wide range of frequencies for sending unsolicited Router Advertisement messages. Successive advertisements must be spaced by random times between 3 and 4 seconds at least and between 1350 and 1800 seconds at most.

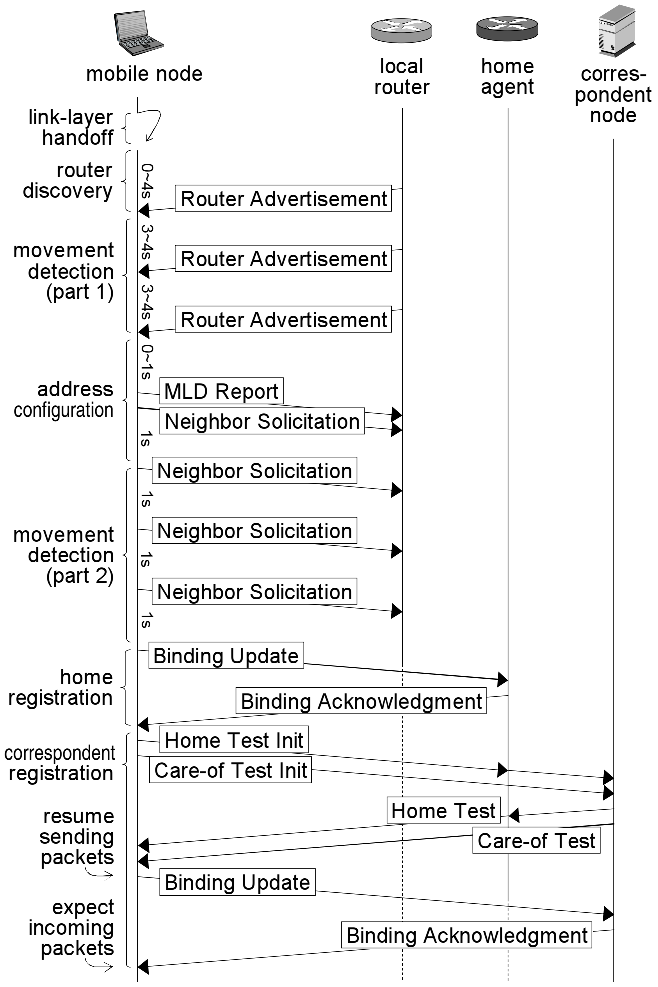

Figure 1: The handoff procedure with standard IPv6 Neighbor Discovery.

The mobile node may slightly increase the advertisement rate by periodically sending a Router Solicitation message. Routers usually transmit solicited advertisements not faster than one every 3 to 3.5 seconds. This is because the solicitations should be sent from the unspecified address since a change in IP connectivity may have invalidated the mobile node's unicast addresses in the meantime. The responding advertisement is consequently sent by multicast transmission, which routers must rate-limit to at most one every 3 seconds. I.e., if the router has already sent another multicast advertisement during the past 3 seconds, it must wait for the remaining time before it transmits the next. In addition, solicited Router Advertisement messages are delayed by a random value between 0 and 500 milliseconds since the solicitation may have synchronized routers. Such desynchronization delays are additive to rate limitations according to implementation guidelines proposed in the IPv6 Neighbor Discovery RFC [19]. It is assumed in figure 1 and table 1 that the router uses advertisement intervals of between 3 and 4 seconds, and that the mobile node does not send Router Solicitation messages.

In theory, the mobile node could send a Router Solicitation message from the link-local address. Routers might then return a solicited Router Advertisement message by unicast and thus bypass any rate limitations for multicast advertisements. However, this approach is impractical for two reasons. First, the IPv6 Neighbor Discovery RFC suggests that routers send a solicited advertisement by multicast even when the solicitation was received from a unicast address. The source address of the solicitation is therefore irrelevant in most cases. Second, the mobile node must verify uniqueness of the link-local address before it sends the solicitation. Address verification involves substantial delays on its own as explained next.

2.2 Address Configuration

A mobile node configures a new global IP address upon receipt of a Router Advertisement message containing an unknown prefix. The mobile node must also re-verify uniqueness of its link-local address if the advertisement suggests a change in IP connectivity. This is necessary even though the link-local address keeps its prefix during handoff, because a new neighbor may already be using the same link-local address. Address configuration and re-verification typically happens as follows in compliance with Stateless Address Autoconfiguration [20]. The mobile node chooses an interface identifier, either randomly or based on the interface's MAC address, and prepends to this the obtained prefix. It then sends a Multicast Listener Discovery (MLD) Report message [21, 22] to subscribe to the solicited-node multicast group corresponding to the new address. The report is deferred by a random desynchronization delay between 0 and 1 second if it was triggered by a multicast advertisement, which is usually the case as mentioned in section 2.1. A method for desynchronization is necessary here because neighboring nodes may respond to the same advertisement. The mobile node then runs the Duplicate Address Detection protocol to verify whether the address is unique: It transmits a Neighbor Solicitation message for the address and, if no responses are received within a period of 1 second, assigns the address to the interface. Stateless Address Autoconfiguration does not automatically recover from an address collision, but the probability for such an event is small enough to make it negligible [23]. Multiple addresses can be configured in parallel, so figure 1 depicts configuration messages only for a single global address and the link-local address.

Routers may specify in transmitted Router Advertisement messages that global addresses be configured in a stateful way. One possible mechanism for this is the Dynamic Host Configuration Protocol version 6 (DHCPv6) [24], through which nodes can request a server to assign them a unique address. DHCPv6 configuration of global addresses requires a valid link-local address, which is always configured through Stateless Address Autoconfiguration. This leads to an initial idle period during which the link-local address is re-verified for uniqueness. Once the link-local address is guaranteed to be unique, the mobile node solicits on-link DHCPv6 servers and listens for responses for a random time between 1.0 and 1.1 seconds. A second, undelayed message exchange actually assigns the mobile node an address. Given a Router Advertisement message that suggests a movement, the DHCPv6 solicitation is likely to be the first that the mobile node sends on the new link. An initial desynchronization delay, randomly chosen to be between 0 and 1 second, must be then applied to the solicitation. Note that figure 1 and table 1 assume the use of stateless address configuration for all addresses.

2.3 Movement Detection

A mobile node implements movement detection to recognize changes in IP connectivity. Such a change implies that the mobile node chooses a new default router, re-verifies uniqueness of its link-local address, invalidates existing global addresses, configures a new care-of address, and initiates Mobile IPv6 registrations. Movement detection is commonly implemented by analyzing the prefixes advertised in Router Advertisement messages and probing reachability of routers considered off-link. When the prefixes in use by the mobile node are no longer seen to be advertised, but new prefixes show up instead, the mobile node would typically decide that it has moved to a different network. On the other hand, received prefixes may also indicate that IP connectivity did not change in spite of a link-layer handoff, e.g., when the mobile node switches access points that connect to the same subnet. Further handoff steps can then be omitted.

Movement detection is complicated by the fact that Router Advertisement messages may include incomplete sets of on-link prefixes. Reception of a single advertisement is therefore usually insufficient to decide whether IP connectivity has changed. There is also no guarantee of a router's advertisement rate. Failure to receive an expected Router Advertisement message does therefore not imply movement either. A typical movement detector would hence draw a possibly premature decision based on a small number of received Router Advertisement messages and, if a change in IP connectivity is assumed, perform Neighbor Unreachability Detection to corroborate this. A reasonable approach would be to use three advertisements.

Neighbor Unreachability Detection [19] cleans up a node's internal state when a former neighbor turns out to be unreachable. In the context of movement detection, a mobile node can use the procedure to actively probe its configured default router when received Router Advertisement messages suggest a change in IP connectivity. Neighbor Unreachability Detection involves transmission of up to three Neighbor Solicitation messages, spaced by 1 second during which responding Neighbor Advertisement messages are awaited. The interval gives the default router a chance to respond and at the same time ensures that the Neighbor Solicitation messages obey the required rate limitations. The potential for packet loss is covered by the retransmissions so that failure to receive a Neighbor Advertisement message can eventually be interpreted as a change in IP connectivity. The second to fourth solicitation in figure 1 is used for Neighbor Unreachability Detection.

The solicitations sent during Neighbor Unreachability Detection cannot originate from the unspecified address. On the other hand, existing global addresses may have been invalidated by a movement at that time. And even though the link-local address keeps its prefix during handoff, the mobile node must re-verify uniqueness of the address before it uses it for Neighbor Unreachability Detection (cf. section 2.2). Neighbor Unreachability Detection is hence preceded by Duplicate Address Detection on the link-local address. Given that addresses for which Duplicate Address Detection is in progress are unavailable for regular communications, the mobile node should not initiate the procedure before a change in IP connectivity has become reasonably likely. An appropriate approach is to wait until at least three consecutive Router Advertisement message with exclusively unknown prefixes have been received. The mobile node then initiates Duplicate Address Detection on the link-local address and subsequently initiates Neighbor Unreachability Detection for the currently configured default router, as illustrated in figure 1. Assuming that the first advertisement causes the mobile node to configure a new global address and send an MLD Report message as part of that, no MLD Report message needs to be transmitted.

2.4 Mobile IPv6 Registrations

If Neighbor Unreachability Detection consolidates a presumed change in IP connectivity, the mobile node chooses a new care-of address and registers this with its home agent and correspondent node. The *home registration* consists of a Binding Update message which notifies the home agent of the new care-of address, and a Binding Acknowledgment message indicating success or failure (cf. figure 1). The mobile node and the home agent are typically administered by the same domain and pre-share credentials to bootstrap an IPsec security association. Both messages can so be authenticated and encrypted.

The *correspondent registration* includes a Binding Update message that conveys the new care-of address to the correspondent node, and an optional Binding Acknowledgment message. (Whether or not the correspondent node sends an acknowledgment is left to the discretion of the mobile node. The mobile node can request one by setting a flag in the Binding Update message.) These cannot generally be protected through IPsec, however, because mobile nodes are neither likely to share authentication credentials with all correspondent nodes they may at some point communicate with, nor is a ”global” public-key infrastructure, available for arbitrary pairs of nodes, expected to come into existence any time soon [25]. The correspondent registration is instead protected through a *return-routability procedure*, based on non-cryptographic verification of a mobile node’s reachability at the home and care-of addresses. This approach is motivated by the following two observations: First, in the context of Mobile IPv6, mobile nodes are identified by home addresses. A reachability test of the home address can therefore authenticate a mobile node. Second, a reachability test of the care-of address prevents redirection-based flooding attacks [25] and so authorizes a mobile node to claim that care-of address.

For the *home-address test*, the mobile node tunnels a Home Test Init message to the home agent, which forwards the message to the correspondent node. The correspondent node returns an unpredictable *home keygen token* to the home address within a Home Test message, and the home agent tunnels this to the mobile node. The *care-of-address test* is a direct exchange between the mobile node and the correspondent node. It consists of a Care-of Test Init message and a Care-of Test message with an unpredictable *care-of keygen token*. Knowledge of the home and care-of keygen tokens proves the mobile node’s ability to receive packets at the home address and care-of address, respectively, and thus enables the correspondent node to bind the two addresses to each other. Specifically, the mobile node authenticates the Binding Update message that it subsequently sends to the correspondent node with a key derived from the received tokens. The correspondent node uses the same key to authenticate the final Binding Acknowledgment message.

2.5 Performance Analysis

The latency of router discovery strongly depends on the configuration of local routers if the mobile node passively listens for multicast Router Advertisement messages. At maximum rates, multicast advertisements are spaced by 3.5 seconds on average, so the mobile node can expect to receive the first one 1.75 seconds after a handoff. This time can be much longer, however, due to the wide range of feasible advertisement intervals. The mobile node avoids the dependency on router configurations if it sends a Router Solicitation message as soon as it arrives on the new link. A new router will then send a Router Advertisement message after a mean desynchronization backoff of 250 milliseconds unless it transmits the advertisement by multicast and has already sent another one throughout the past 3 seconds. In that latter case, the mean delay for the advertisement is 1.75 seconds, comprising the expected time to fill up the minimum interval of 3 seconds between successive multicast advertisements plus the mean desynchronization backoff.

A solicited multicast Router Advertisement message is transmitted with an overall average delay of 3.25 seconds. In scenarios where mobility is high and solicitations are accordingly sent on a frequent basis, the mobile node could expect to receive the first advertisement, be it unsolicited or solicited by a different node, 1625 milliseconds after it arrives on a new link. If the mobile node verifies uniqueness of its link-local address and then sends a Router Solicitation message from this address, the expected time to receive a unicast Router Advertisement message adds up to 1750 milliseconds. This comprises the mean desynchronization delay of 500 milliseconds for the initial MLD Report message, the 1-second latency of Duplicate Address Detection, and an average desynchronization delay of 250 milliseconds for the advertisement itself.

The first Router Advertisement message received after the handoff triggers configuration of new global addresses and movement detection at the same time. As both tasks proceed in parallel, and movement detection takes longer than configuration of global addresses as shown next, the latter has usually no impact on the handoff delay.

The total black-out period for Stateless Address Autoconfiguration consists of a random desynchronization delay for the MLD Report message plus the fixed latency of Duplicate Address Detection. This yields an average of 1.5 seconds if the address is unique. In case the address is already in use by a neighbor, address configuration aborts one link-local round-trip time after the MLD Report message is sent. If DHCPv6 is used to configure global addresses, configuration of global addresses takes 2.05 seconds on average. To this adds the mean delay of 1.5 seconds for re-verifying uniqueness of the link-local address, which must happen in advance. The total delay to configure the global and the link-local addresses is hence 3.55 seconds on average. Obviously, stateless configuration is more efficient than stateful configuration in terms of both latency and signaling overhead.

As discussed in section 2.3, a movement detector should not assume a change in IP connectivity before three subsequent Router Advertisement messages with exclusively unknown prefixes have been received, and Neighbor Unreachability Detection subsequently verifies that the previous default router is no longer available. The delay for the first advertisement is already accounted for by the analysis of router discovery. In addition come the delays for the second and third advertisement. Each of these is 3.5 seconds on average if routers advertise every 3 to 4 seconds and no solicitations are sent. The mobile node hence begins Neighbor Unreachability Detection an expected 7 seconds after router discovery has been accomplished. Neighbor Unreachability Detection itself takes another 3 seconds to complete if IP connectivity has changed. I.e., movement is in this case detected an average of 10 seconds after router discovery is done. When the previous default router is still reachable, Neighbor Unreachability Detection usually concludes after one link-local round-trip time. The average time to receive the second and third advertisement is shorter, namely 3.25 seconds each, if the mobile node solicits multicast Router Advertisement messages. These messages are rate-limited in any case since the first advertisement was solicited just before. Neighbor Unreachability detection hence begins 6.5 seconds and ends 9.5 seconds on average after router discovery is over.

Mobile IPv6 home and correspondent registrations are global message exchanges, whose delays are determined by the round-trip times between the mobile node, home agent, and correspondent node. Let these round-trip times be $RTT(MN, HA)$, $RTT(HA, CN)$, and $RTT(MN, CN)$, where the respective end points are denoted by the straightforward abbreviations in parentheses. With these variables, a home registration concludes after $RTT(MN, HA)$, and the latency of a correspondent registration can be represented as $\max\{RTT(MN, HA) + RTT(HA, CN), RTT(MN, CN)\} + RTT(MN, CN)$, assuming that the home- and care-of-address tests occur in parallel. Given that the path through the home agent is typically longest, the latency of the correspondent registration usually reduces to $RTT(MN, HA) + RTT(HA, CN)$.

The Mobile IPv6 RFC leaves mobile nodes liberties with respect to scheduling signaling and data packets. Figure 1 shows a *conservative* mobile node, which waits for the Binding Acknowledgment message from its home agent before it initiates the return-routability procedure. In contrast, an *optimistic* mobile node could execute the home registration and the return-routability procedure in parallel. An optimistic mobile node would furthermore start sending packets to the correspondent node as soon as the Binding Update message for the correspondent node has been brought on way, whereas a conservative mobile node would use the new care-of address only after reception of an acknowledgment.

It should be noted that figure 1 actually depicts the conservative behavior of the Kame-Shisa [26] Mobile IPv6 implementation for FreeBSD. Mobile IPv6 for Linux (MIPL) [27] differs from this in that a mobile node would send the Care-of Test Init message in parallel with the Binding Update message for the home agent. This does not change delay characteristics unless $RTT(MN, HA) + RTT(HA, CN) < RTT(MN, CN)$, or when a previously acquired, still valid home keygen token redundantizes the home-address test. The performance analyses in this document are based on the Kame-Shisa software.

Conservative mobile nodes avoid a useless return-routability procedure in case the home registration fails. They also do not risk loss of packets sent shortly after a failed Binding Update message. The correspondent node would discard these packets in the face of a mismatching binding due to security measures. This comes at the cost of an additional $RTT(MN, HA) + RTT(MN, CN)$ for outgoing route-optimized packets, and an additional $RTT(MN, HA)$for incoming ones, when both registrations are successful. Optimistic mobile nodes would in the general case perform better. But they may attempt a return-routability procedure in vain or suffer packet loss should the home or correspondent registration fail.

The handoff delay observed by a conservative mobile node for outgoing packet thus amounts to $RTT(MN, HA) + \max\{RTT(MN, HA) + RTT(HA, CN), RTT(MN, CN)\} + RTT(MN, CN)$, and the delay observed by an optimistic mobile node is $\max\{RTT(MN, HA) + RTT(HA, CN), RTT(MN, CN)\}$. The correspondent node is unaware of the new care-of address until it receives the Binding Update message. Its first packet sent to the new care-of address will hence be delivered to the mobile node roughly along with the Binding Acknowledgment message, assuming that one was requested by the mobile node. Thus, a conservative mobile node's observed handoff delay for receiving packets can be calculated as $RTT(MN, HA) + \max\{RTT(MN, HA) + RTT(HA, CN), RTT(MN, CN)\} + RTT(MN, CN)$, and that of an optimistic mobile node amounts to $\max\{RTT(MN, HA) + RTT(HA, CN), RTT(MN, CN)\} + RTT(MN, CN)$.

Kame-Shisa and MIPL are configurable with respect to whether or not the mobile node waits for an acknowledgment from the correspondent node before it sends route-optimized packets. However, neither implementation initiates the home-address tests before the home registration completes, and only MIPL executes the care-of-address test along with the home registration.

Home and care-of keygen tokens are valid for 3.5 minutes after they have been obtained from a correspondent node. The mobile node may therefore be able to omit the home-address test if it has already recently performed one. This reduces the latency of the return-routability procedure, which may so conclude even before the corresponding home registration if both were optimistically initiated at the same time. However, the Mobile IPv6 RFC [1] requires that mobile nodes defer sending a Binding Update message to a correspondent node until an acknowledgment has been received for the home registration. The handoff delays when no home-address test is necessary are thus $RTT(MN, HA) + RTT(MN, CN) + RTT(MN, CN)$ and $RTT(MN, HA) + RTT(MN, CN) + RTT(MN, CN)$ for outgoing and incoming packets, respectively, in case the mobile node is conservative; the delays are $\max\{RTT(MN, HA), RTT(MN, CN)\}$ and $\max\{RTT(MN, HA), RTT(MN, CN)\} + RTT(MN, CN)$ for outgoing and incoming packets, respectively, if the mobile node behaves in an optimistic manner.

These formulas again reflect the behavior of the Kame-Shisa software. But this time, the slightly different conservative behavior in MIPL bears an advantage: Its handoff delay for outgoing packets on a conservative mobile node's side is $\max\{RTT(MN, HA), RTT(MN, CN)\} + RTT(MN, CN)$. Table 1 lists the handoff delays when the home-keygen token acquired during a previous handoff cannot be reused.

3 Handoffs with Increased Advertisement Rates

Router discovery and movement detection are amongst the primary contributors to handoff delay when routers act according to standard IPv6 Neighbor Discovery. This is due to long intervals between transmitted Router Advertisement messages as well as Neighbor Unreachability Detection. In an attempt to improve these conditions, the Mobile IPv6 RFC redefines the minimum transmission interval for Router Advertisement messages to be randomly distributed between 30 and 70 milliseconds. Routers can thus advertise at an average rate of 50 milliseconds. The Mobile IPv6 RFC further defines a new Advertisement Interval option for Router Advertisement messages. A router can use this option to declare an upper bound on its advertisement intervals. E.g., if the router transmits a beacon every 30 to 70 milliseconds, it will advertise an upper bound of 90 milliseconds. (An additional 20 milliseconds are added to the actual maximum interval of 70 milliseconds in order to account for scheduling imprecisions in mobile nodes and routers. The Mobile IPv6 RFC requires this whenever the maximum advertisement interval is smaller than 200 milliseconds.)

These optimizations suggest a different approach to movement detection. Given that the Advertisement Interval option specifies until when the next Router Advertisement message should have been received, a mobile node can consider the absence of one or, more robustly, a small number of expected Router Advertisement messages as an indication of movement without additional Neighbor Unreachability Detection. A reasonable approach would again be to use three advertisements. Not having to perform Neighbor Unreachability Detection is important because the Mobile IPv6 RFC does not change rate limitations for Neighbor Solicitation messages. Router Advertisement messages are hence received on a much faster basis than Neighbor Solicitation messages could be transmitted for the purpose of Neighbor Unreachability Detection. Furthermore, without Neighbor Unreachability Detection, the mobile node does not have to prematurely re-verify uniqueness of its link-local address during the process of movement detection.

The router-discovery and movement-detection optimizations from the Mobile IPv6 RFC reduce the mean time for a mobile node to receive the first Router Advertisement message subsequent to handoff to 25 milliseconds. New global addresses are then configured as needed. The average time to receive the second

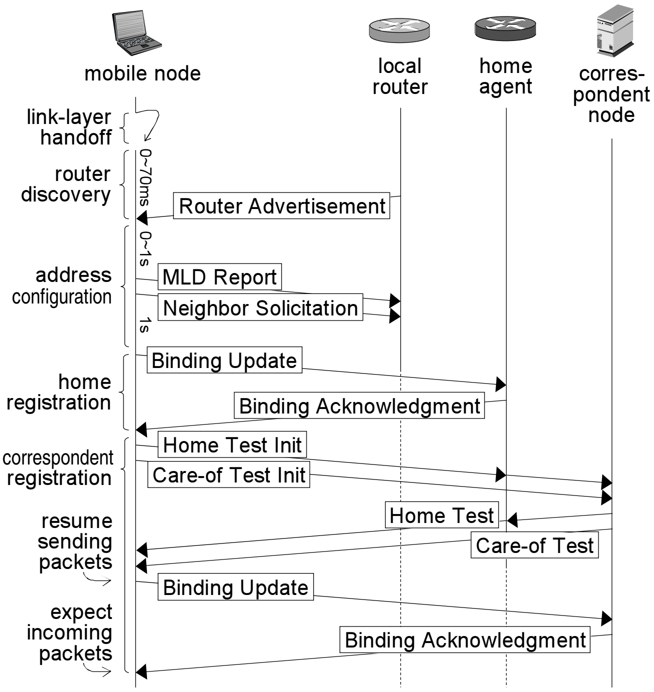

Figure 2: The handoff procedure with increased advertisement rates.

and third advertisement is 50 milliseconds each. Hence, router discovery and movement detection complete after an average 125 milliseconds. This is a substantial advantage over standard IPv6 Neighbor Discovery deployments. Furthermore, the optimizations redundantize transmission of Router Solicitation messages. All a mobile node does is evaluating received Router Advertisement messages. Solicitations would be of no use anyway, because the Mobile IPv6 RFC relaxes rate limitations only for unsolicited Router Advertisement messages. Solicited advertisements must still be limited to one message per 3 seconds.

Configuration of new global addresses starts when the first Router Advertisement message is received. It concurs with the rest of movement detection and, since it dominates with respect to latency, determines the time when Mobile IPv6 registrations begin. Specifically, Mobile IPv6 registrations begin an average of 1.5 seconds after reception of the first advertisement, comprising the desynchronization delay for the initial MLD Report message and the latency of Duplicate Address Detection. Figure 2 illustrates this. The second and third Router Advertisement messages are hidden in the figure as they would overlap with address configuration. Also not shown is the uniqueness re-verification of the link-local address, which begins when the third advertisement is received and overlaps with the configuration of global addresses as well as Mobile IPv6 registrations.

The performance increase of higher advertisement rates comes at the cost of bandwidth. Especially in low-bandwidth, wide-area networks are short advertisement intervals an issue as many users may not frequently leave the geographic area covered by the same IP subnet. A Router Advertisement message with a single Prefix Information option and an Advertisement Interval option is 96 bytes in size, excluding the link-layer frame. At an average transmission rate of 20 beacons per second, this amounts to 15.36 kbps per advertising router. Another disadvantage of multicast advertisements is that they are sent to the all-nodes multicast address, so MLD snooping switches [22] must propagate them to all link segments. By contrast, solicited advertisements consume resources on a single link segment when they are sent to a unicast address.

4 Composing Improved Handoff Procedures

The delays of the standard handoff procedure can significantly impair the quality of real-time applications, even though Route Optimization was built into Mobile IPv6 with an intention to improve support for these applications. The research community has been working to decrease handoff delays for some time now, and a number of proposals have been made. The following proposals can be composed into the improved handoff procedures discussed in this document. Other promising techniques are attended to in section 8.

Handoffs with standard IPv6 Neighbor Discovery			
router discovery			0 to 4 seconds, $\oslash$ = 1.75 seconds
link-local address configuration			1 to 2 seconds, $\oslash$ = 1.5 seconds
global address configuration			0
movement detection			9 to 11 seconds, $\oslash$ = 10 seconds
Mobile IPv6 registrations	outgoing packets	conservative	$RTT(MN, HA) + \max\{RTT(MN, HA) + RTT(HA, CN), RTT(MN, CN)\} + RTT(MN, CN)$
		optimistic	$\max\{RTT(MN, HA) + RTT(HA, CN), RTT(MN, CN)\}$
	incoming packets	conservative	$RTT(MN, HA) + \max\{RTT(MN, HA) + RTT(HA, CN), RTT(MN, CN)\} + RTT(MN, CN)$
		optimistic	$\max\{RTT(MN, HA) + RTT(HA, CN), RTT(MN, CN)\} + RTT(MN, CN)$

Handoffs with increased advertisement rates			
router discovery			0 to 70 milliseconds, $\oslash$ = 25 milliseconds
link-local address configuration			0
global address configuration			1 to 2 seconds, $\oslash$ = 1.5 seconds
movement detection			0
Mobile IPv6 registrations	outgoing packets	conservative	$RTT(MN, HA) + \max\{RTT(MN, HA) + RTT(HA, CN), RTT(MN, CN)\} + RTT(MN, CN)$
		optimistic	$\max\{RTT(MN, HA) + RTT(HA, CN), RTT(MN, CN)\}$
	incoming packets	conservative	$RTT(MN, HA) + \max\{RTT(MN, HA) + RTT(HA, CN), RTT(MN, CN)\} + RTT(MN, CN)$
		optimistic	$\max\{RTT(MN, HA) + RTT(HA, CN), RTT(MN, CN)\} + RTT(MN, CN)$

Table 1: IP-layer delays for handoffs with standard IPv6 Neighbor Discovery as well as for handoffs with increased advertisement rates. The delay of movement detection excludes the delays of router discovery and address configuration in both cases.

4.1 Router Discovery

More sophisticated scheduling intervals in routers can improve router discovery with respect to both bandwidth consumption and efficiency. FastRA [7, 6] permits a mobile node to solicit an immediate Router Advertisement message. This is useful when the mobile node can receive a notification from its local link layer upon a change in link-layer attachment. Based on neighboring routers' link-local addresses and the source address of the solicitation, each router autonomously computes a dynamic ranking indicating which router should respond immediately, and optionally which other routers should send additional advertisements shortly thereafter.

4.2 Address Configuration

The IPv6 working group within the IETF is developing Optimistic Duplicate Address Detection [13] to avoid the handoff delays caused by Stateless Address Autoconfiguration. Optimistic Duplicate Address Detection eliminates the desynchronization delay for the initial MLD Report message and allows for limited use of IP addresses that are yet to be verified for uniqueness. Mobile nodes temporarily change the rules by which they do IPv6 Neighbor Discovery signaling so as to avoid pollution of other nodes' neighbor caches with possibly illegitimate address-resolution information. The technique was about to obtain RFC status at the time of writing this document.

4.3 Movement Detection

The DNA working group within the IETF tackles the problem of slow movement detection with two complementary approaches. The Complete Prefix List protocol [10] works with unmodified routers. A mobile node maintains a list of learned on-link prefixes, obtained by reception of usually multiple Router Advertisement messages. After the list has matured for a while, the mobile node can assume a change in IP connectivity with high confidentiality when a newly received Router Advertisement message exclusively contains prefixes not in the list. However, such predictions are based on potentially incomplete information, so the mobile node might assert movement even when none actually occurred.

The DNA protocol [12] integrates Complete Prefix List and adds to this FastRA for timely transmission of solicited Router Advertisement messages. Furthermore, neighboring routers choose a certain prefix to serve as a *link identifier* and be as such carried in all transmitted Router Advertisement messages. This allows a mobile node to detect changes in IP connectivity based on a single advertisement. Alternatively, the mobile node can explicitly check with routers as part of the solicitation-advertisement exchange whether a network prefix used before a link-layer handoff, as such called a *landmark*, is still valid on the possibly new link.

4.4 Mobile IPv6 Registrations

Early Binding Updates [28, 29] in combination with Credit-Based Authorization [30, 31] changes the timing of the return-routability procedure such that both address tests can be executed outside the performance-critical handoff phase. The techniques thus improve Mobile IPv6 Route Optimization on a purely end-to-end basis. This is realized through five constituent optimizations.

Proactive Home Address Tests. A proactive home-address test is a stand-alone address test by which a mobile node acquires a home keygen token for a future handoff. Proactive home-address tests save a possibly long round trip through the home agent during the critical handoff period. The mobile node invokes proactive home-address tests on a just-in-time basis if its link layer provides a trigger indicating imminent handoff. Alternatively, the mobile node periodically repeats the proactive home-address test whenever the most recently obtained home keygen token is about to expire.

A mobile node may perform proactive home-address tests even when it stays on its home link in order to optimize a future handoff to a foreign link. However, Mobile IPv6 implementations keep received home keygen tokens in a Binding Update List, which they usually remove from memory once a mobile node connects to its home link. To support proactive home-address tests from the home link, an implementation would have to retain existing and create new list entries while the mobile node is at home, just as it does when the mobile node roams away from home.

Concurrent Care-of Address Tests. While home addresses are stable and can as such be tested in a proactive manner, care-of addresses change during handoff. This implies that the care-of-address test must occur after link-layer handoff. Handoff latency can hence be decreased only if the correspondent node allows for limited use of the new care-of address until the mobile node's reachability has been verified. This is the purpose of a concurrent care-of-address test.

Tentative Bindings. The mobile node registers a tentative binding between its home address and an *unverified* care-of address by exchanging Early Binding Update and Early Binding Acknowledgment messages with a correspondent node. The messages are authenticated only with a home keygen token, thus facilitating a subsequent, concurrent care-of-address test. The tentative registration may happen before the link-layer handoff when the movement can be anticipated. Otherwise, it takes place afterwards. The mobile node resumes regular communications as soon as it has dispatched the Early Binding Update message. Once it has executed a concurrent care-of-address test, the mobile node authenticates a standard Binding Update message and registers a *verified* care-of address with the correspondent node. A tentative binding is limited to 10 seconds in lifetime. This should be long enough to bridge the expected duration of the remaining correspondent registration, although the mobile node can refresh the tentative binding just as a regular one. The lifetime limit is meant to decently recover in cases where the tentative registration turns out to be premature. This may at times happen during a proactive handoff (cf. section 7), when the mobile node expects to move to some new link, but eventually understands that it ended up at a different link or simply stayed where it was. A correspondent node would then revert to the home address once the tentative registration expires. Note that the lifetime limit for tentative bindings does not compensate the temporary lack of a reachability check. This is the purpose of Credit-Based Authorization.

Credit-Based Authorization. Well-known security guidelines [25] prohibit a correspondent node to send packets to a care-of address for which reachability has not yet been verified. This is a precaution against malicious nodes which could otherwise trick correspondent nodes into flooding a third party with unrequested packets. The appeal of such *redirection-based flooding attacks* is the potential for significant amplification. E.g., an attacker could accomplish the initial TCP handshake for a voluminous file download through its own address (or home address, for that matter), and then redirect the flow to the address of its victim. The attacker could spoof acknowledgments on behalf of the victim based on the sequence numbers it learned from the initial handshake, but those would be small compared to the full-sized segments that the correspondent node generates. Credit-Based Authorization [30, 31] prevents such misuse of an unverified care-of address as long as the correspondent node does not spend more effort than the mobile node has recently spent. This precludes amplification and so defeats the purpose of redirection-based flooding: An attacker would more effectively flood its victim by sending bogus packets directly.

Keeping the balance between the correspondent node's effort and the mobile node's is technically realized as follows: The correspondent node maintains a byte counter for the mobile node, also called the mobile node's *credit*. This increases with the data volume received from the mobile node and decreases with the data volume sent to the mobile node while the care-of address is unverified. Exponential aging assures that existing credit represents only recent effort of the mobile node. When the correspondent node has a packet for the mobile node, it sends it to the care-of address if the address is either verified or if it is unverified, but the packet size does not exceed the currently available credit. Otherwise, the correspondent node may drop the packet, buffer it until the care-of address becomes verified, or send the packet to the home address. (Packet buffering makes the correspondent node susceptible to memory-overflow attacks and may hence represent a denial-of-service vulnerability on its own. However, where the correspondent node can identify a trustworthy mobile node based on the home address, and the mobile node's reachability at the home address has been verified, packet buffering could be an option.)

Note: The outlined Credit-Based Authorization mode assigns a mobile node new credit based on packets that the correspondent node receives from the mobile node, but reduces the credit based on packets that the correspondent node sends. Applications with strongly asymmetric traffic patterns may work better with an alternative mode [31], in which credit increases with the data volume the mobile node is found to have received. This accommodates any traffic pattern, how asymmetric it may be. The correspondent node can periodically spot check the mobile node's reachability in order to estimate the packet loss on the path to the mobile node. Spot checks are piggybacked to ordinary data packets to minimize overhead.

Parallel Home and Correspondent Registrations. As mentioned in section 2.5, the Mobile IPv6 RFC gives a mobile node the freedom to execute an optimistic return-routability procedure in parallel with the corresponding home registration, but it does not permit the mobile node to continue with a correspondent registration before an acknowledgment has been received from the home agent. This is an issue if the return-routability procedure completes earlier than the home registration, e.g., when a home keygen token from the previous handoff is still valid and no home-address test is necessary. Beyond this, a combination of a proactive home-address test and a concurrent care-of-address test virtualizes the latency of the entire return-routability procedure. The rules of Mobile IPv6 are hence relaxed so as to allow the mobile node to send Early Binding Update messages when the home registration is still pending. Nonetheless, the mobile node waits for an acknowledgment from its home agent before it finally sends standard Binding Update messages to the correspondent node.

4.5 Cross-Layer Interaction

Strict OSI separation of networking protocols precludes synchronization between link- and IP-layer handoffs. Mobile nodes must determine changes in IP connectivity solely based on IP-layer mechanisms such as IPv6 Neighbor Discovery. This was found to substantially delay the handoff procedure. The IEEE hence chartered its 802.21 working group, Media Independent Handover Services [15], in March 2004 to develop a unified interface between different link-layer technologies and IP. This interface defines a set of events and commands that upper-layer mobility protocols can read or issue, respectively, to synchronize their handoff-related activities with the link layer. In its basic form, a mobile node listens for Link Up events so that it can quickly initiate reactive handoff mechanisms at the IP layer when it changes link-layer attachment. Movement anticipation and proactive handoff management (cf. section 7) requires more advanced interaction with the link layer. Here, the mobile node periodically issues a Link Scan command to have its interface search for available links. Feedback about discovered links within reachability is provided by Link Detected events. A Link Handover Imminent event eventually notifies the mobile node about a forthcoming handoff. After appropriate handoff activities have been initiated, the mobile node issues a Handover Initiate command to attach to the new link.

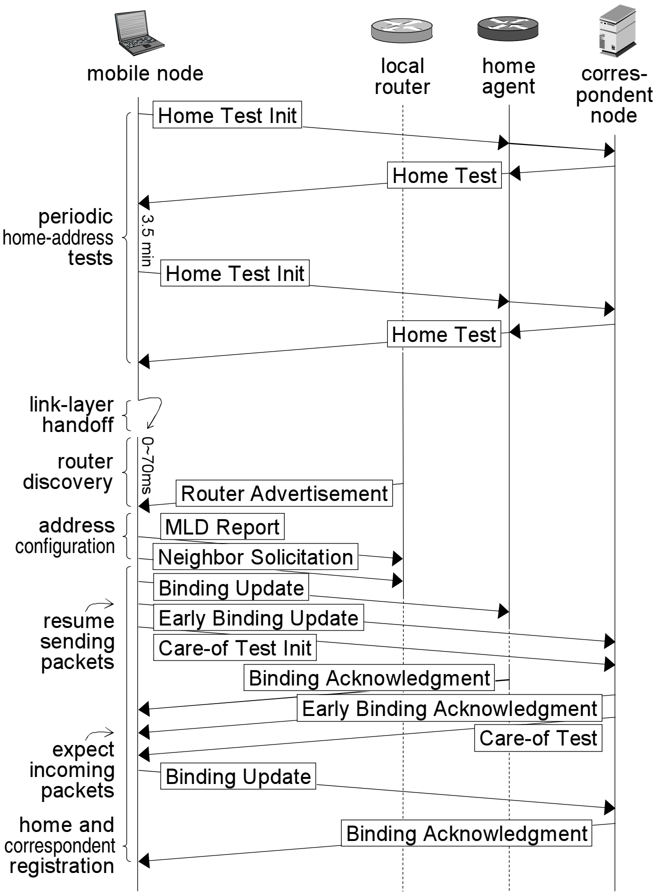

Figure 3: The improved handoff procedure with unmodified routers.

4.6 Off-Link Network Prefix Determination

Proactive handoff management requires a mobile node to somehow obtain prefix information for the target link before it actually attaches to that link. Media Independent Information Services [16], which are part of the afore-mentioned Media Independent Handover Services, provide this information. One parameter of a Link Handover Imminent event is the MAC address of the prospectively new access point. The mobile node leverages Media Independent Information Services to resolve this MAC address to prefix information for the target link. It can then form a new care-of address and register this with its home agent and correspondent nodes. Afterwards, the mobile node issues a Handover Initiate command and attaches to the new link.

Media Independent Information Services as well as other mechanisms for off-link prefix determination (cf. section 8) require special support in the network. This is likely not to be ubiquitously available. In an attempt to provide for more autonomous proactive handoff management, a mobile node can maintain a least-recently-used cache to map the MAC addresses of visited access points onto the on-link prefixes learned at those access points. This technique performs well in scenarios where mobile nodes tend to revisit a rather stable set of access points, e.g., at home or office environments, campuses, conferences, and local shopping centers. Obviously, there is no benefit whenever a mobile node encounters a new access point. Even though this technique is a quite natural approach to autonomous proactive handoff management, the author of this document has so far been unable to locate an equivalent proposal in the literature.

5 Improved Handoffs with Unmodified Routers

The standard handoff procedure described in section 3 can be improved by a combination of end-to-end optimizations in mobile nodes and correspondent nodes. If the mobile node applies Complete Prefix List logic [10], a single Router Advertisement message is in general sufficient for the purpose of movement detection. If the mobile node uses Optimistic Duplicate Address Detection in addition, the delays of Stateless Address Autoconfiguration can also be avoided. Early Binding Updates and Credit-Based Authorization finally reduce the impact that global Mobile IPv6 signaling has on handoff performance. Figure 3 illustrates this handoff procedure, which is described and evaluated in the following. Table 2 juxtaposes performance statistics for this and other improved handoff procedures.

The mobile node periodically executes a proactive home-address test with the correspondent node, refreshing its home keygen token a few seconds ahead of expiry. It should hence know a valid home keygen token when it eventually moves. The mobile node uses Optimistic Duplicate Address Detection to configure a new address upon reception of a Router Advertisement message with an unknown prefix. When the advertisement further suggests a change in IP connectivity, the mobile node initiates uniqueness verification for its link-local address as well, selects a new care-of address, and begins Mobile IPv6 signaling. Optimistic Duplicate Address Detection is usually still in progress at the time home and correspondent registrations begin. The mobile node sends a Binding Update message to the home agent and an Early Binding Update message to the correspondent node. The new care-of address is then available for immediate use.

When the home agent receives the Binding Update message, it sends a Binding Acknowledgment message, and forwards subsequent packets, to the new care-of address. This is standard Mobile IPv6 operation. The correspondent node registers a tentative binding when it receives an Early Binding Update message. The new care-of address is labeled ”unverified”, and the binding is given a lifetime of only 10 seconds. The correspondent node sends an Early Binding Acknowledgment message if one was requested by the mobile node. The acknowledgment is not strictly required during a reactive handoff since loss of the Early Binding Update message would quickly be compensated for by the standard Binding Update message sent shortly afterwards. The correspondent node uses the unverified care-of address to the extent Credit-Based Authorization permits.

The mobile node sends a Care-of Test Init message to the correspondent node along with the respective Early Binding Update message, and the correspondent node returns a Care-of Test message with a new care-of keygen token. This token, in conjunction with the earlier received home keygen token, allows the mobile node to send an authenticated standard Binding Update message to the correspondent node. Upon receipt, the correspondent node changes the status of the care-of address from ”unverified” to ”verified” and extends the binding lifetime to the regular duration of 7 minutes. Use of the care-of address is henceforth no longer governed by Credit-Based Authorization. The correspondent node also sends a Binding Acknowledgment message if one was requested.

Routers transmit unsolicited Router Advertisement messages in random intervals between 30 and 70 milliseconds. They indicate this in Advertisement Interval options. A mobile node can thus expect to receive the first advertisement after an average of 25 milliseconds subsequent to handoff. Complete Prefix List logic can detect changes in IP connectivity based on a single Router Advertisement message without Neighbor Unreachability Detection. This presumes that sufficient prefix information can be collected between successive handoffs, which is very probable, however, given the high advertisement frequency. Movement detection hence does not add to the expected 25 milliseconds required for router discovery.

The delay of a return-routability procedure is masked by the respective tentative correspondent registration. Since the mobile node resumes communications with a correspondent node via the new care-of address immediately after it has sent to it an Early Binding Update message, the registration delay for outgoing packets at the mobile node is eliminated.

A correspondent node tentatively registers the new care-of address when it receives an Early Binding Update message. Assuming that sufficient credit is available, the first packets hence arrive at the new care-of address after a delay of $RTT(MN, CN)$. Lack of credit may cause the correspondent node to send its packets to the mobile node's home address from the very beginning. The first packets will then arrive at the mobile node after a delay of $0.5\,\big(RTT(MN, CN) + RTT(MN, HA) + RTT(HA, CN)\big)$. The statistics in table 2 presume that enough credit is available.

6 Improved Handoffs with Router Support

Unsolicited Router Advertisement messages constitute a permanent trigger for a mobile node to reconsider IP connectivity. This same function can be fulfilled in a more rate-economic and also more efficient way by a combination of cross-layer interaction [32] in the mobile node and IPv6 Neighbor Discovery optimizations in local routers. Both provided, the mobile node can quickly send a Router Solicitation message upon a change in link-layer attachment and receive an immediate Router Advertisement message that allows it to review its current IP configuration. Figure 4 illustrates this handoff procedure.

The optimizations can be realized in two ways. Where routers implement FastRA, mobile nodes must use Complete Prefix List logic to derive prompt movement decisions. Alternatively, both mobile nodes and routers may implement the DNA protocol. Either way, router discovery and movement detection can

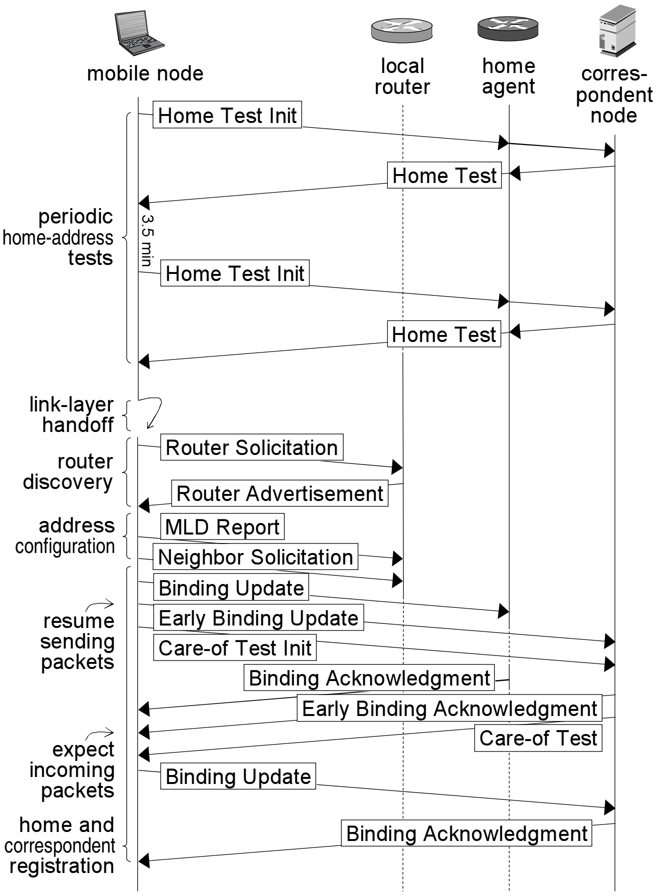

Figure 4: The improved handoff procedure with router support.

be achieved almost instantly. Solicited advertisements are also more rate-economic because they are sent on demand only and do not necessarily consume bandwidth on all segments of a switched link. Address configuration and Mobile IPv6 signaling proceed in the same optimized way as they have been described in section 5.

7 Improved Handoffs with Movement Anticipation

Sophisticated handoff optimizations can eliminate IP-layer delays for packets that a mobile node sends during handoff, yet a minimum black-out period of one round-trip time for incoming packets is inherent in all reactive approaches: It always takes a one-way propagation time to register a new care-of address plus another one-way propagation time for the first payload packet to arrive at that address. Proactive handoff management can eliminate this residual delay.

For a mobile node to anticipate movements and schedule handoff-related activities accordingly, the mobility protocol must be able to issue commands to the link layer and receive events as well as anticipatory information from the link layer. This calls for a bidirectional interface between these layers which goes beyond the unidirectional notification service used in section 6: The mobility protocol must be able to issue commands to the link layer and receive events as well as anticipatory information from it. This can be realized through Media Independent Handover Services. The mobile node must further be able to match link-layer information from a discovered link to network-prefix information for that link. This typically requires a mapping between the MAC address of a discovered access point and the set of on-link prefixes for the network to which this access point connects. Media Independent Information Services may provide this information. Where those, or an equivalent mechanism, do not exist, the mobile node falls back to the reactive handoff described in section 6, but it should cache the retrieved prefix information for later, proactive use (cf. section 4.6).

Figure 5 illustrates the proactive handoff procedure. At some point, the mobile node receives a link-layer notification indicating that the signal strength for its current link attachment has dropped below a

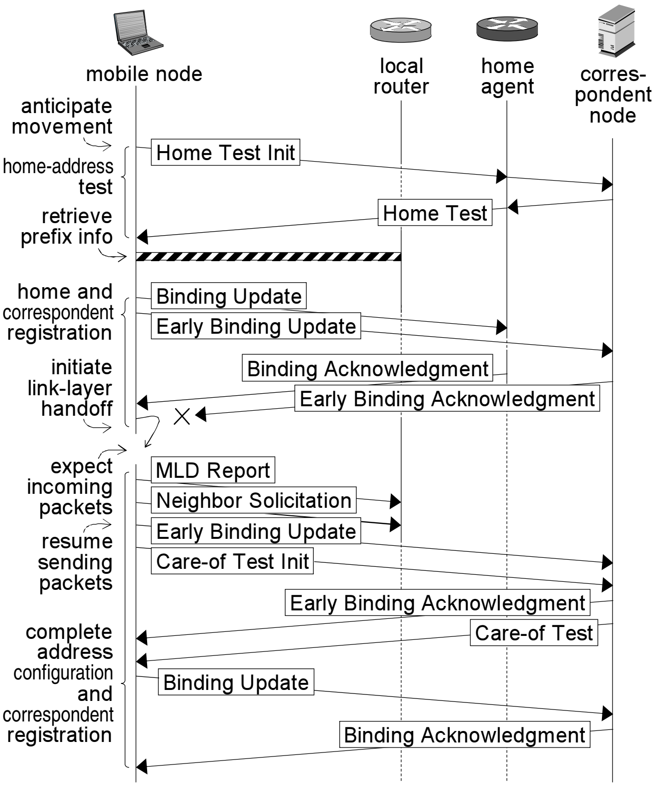

Figure 5: The improved handoff procedure with movement anticipation.

certain threshold. This causes the mobile node to initiate a home-address test and acquire a home keygen token from its correspondent node. Depending on how long the mobile node remains in this pre-handoff stage, it may have to repeat the proactive home-address test to refresh the token. Should a later notification tell that the signal quality has again increased, the periodic tests can be stopped. However, if the signal quality falls further, the mobile node will at some point receive a link-layer notification indicating that a change in link-layer attachment is due. This includes the MAC address of the prospectively new access point. The mobile node then retrieves prefix information for the target link, e.g., through Media Independent Information Services, or checks to see whether it has cached such a mapping from a previous handoff to the same access point.

Based on the prefix of the target link, the mobile node configures a new care-of address and marks this as ”optimistic”. In fact, the mobile node thereby enters Optimistic Duplicate Address Detection, although it defers transmission of related messages until it arrives on the new link. The mobile node then sends a Binding Update message to the home agent and an Early Binding Update message to the correspondent node. Since these messages are sent from the old link, and the IPv6 Source Address field does not contain the new care-of address as it usually does, Alternate Care-of Address options [1] are added to the messages to hold the new care-of address. An interest in an acknowledgment is indicated in both messages. The mobile node uses its old care-of address until it has actually moved to the new link.

When the home agent receives the Binding Update message, it registers the new care-of address, but temporarily continues to also accept packets that the mobile node may send from the old care-of address before moving to the new link [33]. Likewise, correspondent nodes register the new care-of address upon receipt of an Early Binding Update message, but continue to also accept packets from the old care-of address for a while.

The home agent and the correspondent node return their acknowledgments to the old care-of address, but direct subsequent packets to the new care-of address. One of these acknowledgments will cause the mobile node to instruct its link layer to switch to the newly discovered access point. The mobile node implements its own policy as to which acknowledgment it uses for this purpose (cf. section 7.1). When the mobile node eventually arrives on the new link, it begins Optimistic Duplicate Address Detection signaling and sends subsequent data packets via the new care-of address. Some of the acknowledgments are usually lost on the old link, so the mobile node cannot tell whether all registrations were successful. It hence retransmits Binding Update and Early Binding Update messages for any unconfirmed registrations. In figure 5, the registration with the only correspondent node remains unacknowledged on the old link, so the mobile node resends the Early Binding Update message from the new link.

Incoming packets may already be queued at the new access router when the mobile node arrives on the new link, or they arrive shortly, as the home agent and the correspondent node should already be using the new care-of address. The mobile node initiates a concurrent care-of-address test with the correspondent node and finally sends a standard Binding Update message.

In conclusion, router discovery and movement detection become a result of preparatory operations on the old link where the mobile node can anticipate handoffs and discover prefix information for a prospectively new link. An optimistic care-of address is formed prior to handoff on the basis of the obtained prefixes. And although the address is still subject to Duplicate Address Detection on the new link, the mobile node already registers it with its home agent and correspondent node. All this is accomplished on the old link, while regular communications proceed unperturbed. The mobile node can thus resume sending packets immediately after it attaches to the new link, and the delay for packets received from the correspondent node can be eliminated as well. This is shown in table 2.

However, zero handoff delays are possible only when the mobile node makes an appropriate selection with respect to the acknowledgment, received on the old link, upon which it triggers the link-layer handoff. Differences in packet-propagation delays via the old and new routing paths also have an impact. Section 7.1 discusses different approaches to deal with these problems, including scenarios where the mobile node communicates with more than one correspondent node.

7.1 Selecting the Right Time for Link-Layer Handoff

Standard Mobile IPv6 rules cause the home agent and the correspondent nodes to acknowledge Binding Update and Early Binding Update messages to the old care-of address, but to direct subsequent packets to the new care-of address. The mobile node considers one of the arriving acknowledgments as a trigger to initiate the actual link-layer handoff to the new link. The decision which acknowledgment the mobile node should use in this regard has a direct impact on handoff performance, in particular if it communicates with multiple correspondent nodes. If the mobile node switches links too early, it may lose packets on the old link and end up waiting for packets to arrive on the new link. If the mobile node switches too late, it may have to communicate at high power levels on the old link due to fading signal strength, while packets arriving in its absence on the new link are bound to be lost. Address-resolution queues are unlikely to provide a feasible cushion as they are usually short. The IPv6 Neighbor Discovery RFC stipulates a minimum queue length of just a single packet. (The Linux operating system defaults to three packets, while the FreeBSD operating system goes with the minimum.) Though this can in general be increased via application programming interfaces, the hazards of denial-of-service attacks against network infrastructure suggest refrainment from doing so.

A simple heuristic to approach the problem would be to initiate the link-layer handoff when the home registration is acknowledged. The Mobile IPv6 RFC prioritizes home registrations higher than correspondent registration, so this strategy falls in line with the base specification. Earlier reception of an acknowledgment from a correspondent node would be registered as such to avoid an unnecessary retransmission of the respective Early Binding Update message on the new link, but further communications with that correspondent node would be deferred until the mobile node arrives on the new link. This technique is adequate when the mobile node does not frequently use Route Optimization, or when prioritized communications are tunneled via the home agent.

On the other hand, the same heuristic may be a source of decreased handoff performance when Route Optimization is in use. Given that a typical use case for Route Optimization is real-time applications, this heuristic is in fact likely to penalize highly delay-sensitive communications. Circumstances are worst in a so-called *Two-Japaneses-in-America scenario*, where the mobile node and a mobile correspondent node both roam far from their respective homes, but reside close to each other. This may happen, e.g., at a conference venue. Better handoff performance can here be obtained if links are switched upon the arrival of an Early Binding Acknowledgment message. A rudimentary implementation might prompt the link-layer handoff when the first such acknowledgment arrives; a more sophisticated algorithm could choose an acknowledgment based on real-time properties of active applications.

A third strategy launches the link-layer handoff upon the first acknowledgment received, whether it originates with the home agent or with a correspondent node. This approach seeks to combine the advantages of the two previous techniques: When no Route Optimization is used, the home agent's acknowledgment is the only one, and the time for link-layer handoff is chosen well. The result is similar if the home agent is closer to the mobile node compared to the correspondent nodes, possibly in environments where home agents are dynamically assigned based on the logical position of the mobile node. In case an Early Binding

Improved handoffs with unmodified routers		
router discovery		0 to 70 milliseconds, $\oslash$ = 25 milliseconds
address configuration		0
movement detection		0
Mobile IPv6 registrations	outgoing packets	0
	incoming packets	$RTT(MN, CN)$

Improved handoffs with router support		
router discovery		0
address configuration		0
movement detection		0
Mobile IPv6 registrations	outgoing packets	0
	incoming packets	$RTT(MN, CN)$

Improved handoffs with movement anticipation		
router discovery		0
address configuration		0
movement detection		0
Mobile IPv6 registrations	outgoing packets	0 (best case)
	incoming packets	0 (best case)

Table 2: IP-layer handoff delays for the improved handoff procedures under discussion. In all cases, the delay of movement detection excludes the delays of router discovery and address configuration.

Acknowledgment message is received first, a possibly long wait for the home agent's acknowledgment is aborted in favor of route-optimized communications.

Dynamic techniques, where the acknowledgment that determines the time of link-layer handoff is selected based on application characteristics, are certainly superior to algorithms that handle acknowledgments in the order they are received. Humans are limited in the number of concurrent activities they can handle or perceive. This may justify the assumption that at most one, possibly two real-time applications take place at a time. If insight in the nature of active applications is available, the mobile node could identify the peer for which to schedule link-layer handoff would maximize user satisfaction.

Similarly, a handoff decision could be drawn from the ratio between the number of correspondent nodes with which the mobile node uses Route Optimization and the number of correspondent nodes with which it does not. If communications with most correspondent nodes are routed via the home agent, it may be wise to give precedence to the home agent's acknowledgment. Mobile IPv6 implementations must anyway keep state about correspondent nodes that do not support Route Optimization so as to reasonably limit attempts to establish a binding. Depending on the ratio, either the home agent's acknowledgment or an acknowledgment from a correspondent node would trigger the link-layer handoff.

8 Related Research

A number of promising techniques have been proposed to enhance router discovery, address configuration, movement detection, and Mobile IPv6 registrations besides the approaches described in section 4. The Fast Router Discovery [11] proposal suggests that access points replay a cached Router Advertisement message once a node has been associated. This enables immediate router discovery and thereby facilitates expeditious movement detection. Furthermore, the new link-layer support on the network side makes link-layer triggers in mobile nodes superfluous. Fast Router Discovery merges IP-layer and link-layer responsibilities.

In contrast, FastRA, Complete Prefix List, as well as the DNA protocol are plain IP-layer techniques, which may leverage information from the link layer, but leave the link-layer mechanisms themselves unchanged.

With Advanced Duplicate Address Detection [34], routers build up a pool of unique addresses which they then assign to mobile nodes. Duplicate Address Detection is performed for these addresses in advance so that mobile nodes can configure them instantly without further uniqueness verification. The addresses are constructed according to privacy extensions for Stateless Address Autoconfiguration [35].

MLD Duplicate Address Detection [8] makes use of the Multicast Listener Discovery Report messages that mobile nodes send before they begin with standard Duplicate Address Detection. Routers observe from these reports which address ranges include occupied addresses and conclude which addresses are available. They may so be able to give a mobile node clearance to use an address immediately without going through Duplicate Address Detection.

The procedure for proactive handoff management described in section 7 incorporates Media Independent Information Services so that a mobile node can obtain the prefixes of a new link before actually moving to that link. An alternative approach is to adopt the ICMPv6 messages for proxy router discovery defined as part of Fast Handovers for Mobile IPv6 [17]. A mobile node can use proxy router discovery to resolve the MAC address of a discovered access point into prefix and router information for the link to which this access point belongs. The new access router may also be able to check a prospective care-of address for uniqueness. The Candidate Access Router Discovery protocol [36] provides a similar mechanism for proxy router discovery, but does not assist in address configuration.

Depending on the link-layer technology, access points may be able to include prefix information in their beacons, aiding mobile nodes in both movement detection and off-link address configuration. In the special case of IEEE 802.11 networks, this can be realized through specially formatted ESSIDs or new application-specific information elements for management frames [18]. However, augmenting link-layer beacons with IP-layer information may be problematic when link-layer encryption is used. Higher bandwidth consumption may be a separate issue, depending on how much the beacon size increases due to the added contents.

The improved handoff procedures described in this document are based on a combination of Early Binding Updates and Credit-Based Authorization. These optimizations allow communicating nodes to execute the return-routability procedure's home- and care-of-address tests in a way that does not impact handoff performance. Another approach is to replace the address tests with cryptographic authentication. This is possible where mobile and correspondent nodes can be pre-configured to share a secret key or the credentials to bootstrap a security association. Two such proposals are currently under discussion in the IETF. In [37], end nodes are pre-configured with a shared key to compute home and care-of keygen tokens autonomously rather than having to obtain them through the return-routability procedure. [38] uses IPsec and the Internet Key Exchange protocol. Both techniques suffer from a scalability problem, however, given that end nodes must be set up with pair-wise credentials.

Crypto-Based Identifiers [39] address the issue of pre-configuration. A Crypto-Based Identifier has a strong cryptographic binding to the public component of its owner's public/private key pair. Peers can authenticate the owner by testing its knowledge of the corresponding private key. For Mobile IPv6, a Crypto-Based Identifier takes the form of a home address. Such a cryptographically generated address [40, 41, 42] has a standard routing prefix and a hash on the public key as the interface identifier.

All of these optimizations are alternative methods for authentication, but provide no verification of a mobile node's reachability. Technically, they can hence replace a home-address test only. End nodes which trust in the peer's reachability may also omit the care-of-address test, but such trust is unavailable in many important business models. E.g., a mobile-phone operator may be able to configure subscribers with secret authentication keys, but it may not be able to vow that all subscribers use these keys in a trustworthy manner.

Another family of Mobile IPv6 optimizations is based on router support in the mobile nodes' visited networks. Where Fast Handovers for Mobile IPv6 [17] are deployed, a mobile node can request its current default router to establish a bidirectional tunnel to a new care-of address. This allows the mobile node to temporarily communicate through its old care-of address after a handoff, and to register the new care-of address with its home agent and correspondent nodes while doing so. By help of proxy router discovery and assisted address configuration, the mobile node may request the tunnel prior to handoff, provided that it can anticipate movements in advance (cf. section 7). Inbuilt capabilities allow the mobile node to quickly recover in case of an unexpected link break.

Conversely, Media Independent Pre-Authentication [43] uses a bidirectional tunnel between the old care-of address and a prospectively new default router. A mobile node is assigned a new care-of address from remote and initiates home and correspondent registrations before it changes links. The benefits of this approach are similar to those of Fast Handovers for Mobile IPv6 if the overlap between neighboring cells is sufficiently large to permit timely completion of handoff preparations. However, where cell overlaps are small relative to node mobility, postponing global signaling to a stage after handoff is advantageous, since the wireless signal quality is generally higher and more durable then. The strength of Media Independent Pre-Authentication is its ability to pre-authenticate a mobile node to the new network prior to handoff.

Hierarchical Mobile IPv6 [44] enables a mobile node to bind its current *on-link* care-of address to a more stable *regional* care-of address from a Mobility Anchor Point's network located elsewhere in the visited domain. The mobile node sends and receives packets via the regional care-of address through a bidirectional tunnel to the Mobility Anchor Point. It registers the regional care-of address with the home agent and correspondent nodes and updates the Mobility Anchor Point whenever its on-link care-of address changes in the wake of a movement. Movements can so be concealed from the home agent and correspondent nodes as long as the mobile node roves within the same Mobility Anchor Point's service area.

The high performance benefits achievable through Fast Handovers for Mobile IPv6, Media Independent Pre-Authentication, or Hierarchical Mobile IPv6 come at the cost of required upgrades to network infrastructure. Also, these optimizations may not contribute to inter-domain handoffs due to lack of roaming agreements between access providers. End-to-end optimizations do not have such constraints [14]. At the cost of some performance, they provide an independence which can be of advantage in many roaming scenarios.

9 Conclusion

Efficient end-to-end IPv6 mobility support requires optimizations not only for the mobility protocol, but also for IPv6 router discovery, movement detection, and address configuration. This document has taken an in-depth look at the shortcomings of today's protocol standards, explored existing optimizations, examined how those interact, and how they can be combined into complete and efficient mobility solutions. In particular, three integrated procedures have been proposed for improved handoff experience in surroundings with different preconditions: reactive handoffs with unmodified routers, reactive handoffs with router support, as well as movement anticipation and proactive handoff management.

In investigating into handoff procedures, this document has mainly ignored the impacts of link-layer authentication and attachment protocols. This is due to the variety of existing link-layer technologies. However, it can in general be emphasized that link-layer enhancements are equally important as IP-layer enhancements, in particular because IP-layer optimizations as well as cross-layer signaling reduce the overall handoff delay to the delay of a link-layer attachment change. The necessity to improve link-layer handoff management has been recognized and taken on [45, 46, 47, 48, 49].

Acknowledgment

The author thanks Ralf Beck, Roland Bless, Gregory Daley, Mark Doll, Daniel Jungbluth, and Max Laier (in alphabetical order) for valuable discussions on the topic of this document.

References

[1] D. Johnson, C. E. Perkins, and J. Arkko, "Mobility Support in IPv6," IETF RFC 3775, June 2004.

[2] S. Ortiz Jr., "Internet Telephony Jumps off the Wires," *IEEE Computer*, vol. 37, no. 12, pp. 16–19, Dec. 2004.

[3] M. Bernaschi, F. Cacace, G. Iannello, S. Za, and A. Pescape, "Seamless Internetworking of WLANs and Cellular Networks: Architecture and Performance Issues in a Mobile IPv6 Scenario," *IEEE Wireless Communications*, vol. 12, no. 3, pp. 73–80, June 2005.

[4] N. Montavont and T. Noël, "Analysis and evaluation of mobile IPv6 handovers over wireless LAN," *Mobile Networks and Applications*, vol. 8, no. 6, pp. 643–653, Dec. 2003.

[5] ——, "Handover Management for Mobile Nodes in IPv6 Networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 38–43, Aug. 2002.

[6] G. Daley, B. Pentland, and R. Nelson, "Effects of Fast Routers Advertisement on Mobile IPv6 Handovers," in *Proceedings of the International Symposium on Computers and Communication*, vol. 1. IEEE, June 2003, pp. 557–562.

[7] ——, “Movement Detection Optimizations in Mobile IPv6,” in *Proceedings of the IEEE International Conference on Networks*. IEEE, Sept. 2003, pp. 687–692.

[8] N. Moore and G. Daley, “Fast Address Configuration Strategies for the Next-Generation Internet,” in *Proceedings of the Australian Telecommunications, Networks, and Applications Conference*, Dec. 2003.

[9] C. Vogt, R. Bless, M. Doll, and G. Daley, “Analysis of IPv6 Relocation Delays,” Institute of Telematics, Universitaet Karlsruhe (TH), Germany, Technical Report TM-2005-4, Apr. 2005.

[10] J. Choi and E. Nordmark, “DNA with Unmodified Routers: Prefix List Based Approach,” IETF Internet Draft draft-ietf-dna-cpl-01.txt (work in progress), Apr. 2005.

[11] J. Choi, D. Shin, and W. Haddad, “Fast Router Discovery with L2 Support,” IETF Internet Draft draft-ietf-dna-frd-00.txt (work in progress), Oct. 2005.

[12] B. Pentland, Ed., “Detecting Network Attachment in IPv6 Networks (DNAv6),” IETF Internet Draft draft-pentland-dna-protocol3-00.txt (work in progress), Oct. 2005.

[13] N. S. Moore, “Optimistic Duplicate Address Detection for IPv6,” IETF Internet Draft draft-ietf-ipv6-optimistic-dad-07.txt (work in progress), Dec. 2005.

[14] C. Vogt and J. Arkko, “Taxonomy and Analysis of Enhancements to Mobile IPv6 Route Optimization,” IETF Internet Draft draft-irtf-mobopts-ro-enhancements-04.txt (work in progress), Oct. 2005.

[15] “Draft IEEE Standard for Local and Metropolitan Area Networks: Media Independent Handover Services,” Draft IEEE Standard P802.21/D00.01, July 2005.

[16] V. Gupta, Y. Ohba, S. Das, and S. Sreemanthula, “Information Elements,” IEEE Contribution 21-05-0400-00-0000, Nov. 2005.

[17] E. Rajeev Koodli, “Fast Handovers for Mobile IPv6,” IETF RFC 4068, July 2005.

[18] P. Tan, “Recommendations for Achieving Seamless IPv6 Handover in IEEE 802.11 Networks,” IETF Internet Draft draft-paultan-seamless-ipv6-handoff-802-00.txt (work in progress), Feb. 2003.

[19] T. Narten, E. Nordmark, W. A. Simpson, and H. Soliman, “Neighbor Discovery for IP version 6 (IPv6),” IETF Internet Draft draft-ietf-ipv6-2461bis-05.txt (work in progress), Oct. 2005.

[20] S. Thomson, T. Narten, and T. Jinmei, “IPv6 Stateless Address Autoconfiguration,” IETF Internet Draft draft-ietf-ipv6-rfc2462bis-08.txt (work in progress), May 2005.

[21] R. Vida, L. H. M. K. Costa, S. Fdida, S. Deering, B. Fenner, I. Kouvelas, and B. Haberman, “Multicast Listener Discovery Version 2 (MLDv2) for IPv6,” IETF RFC 3810, June 2004.

[22] M. J. Christensen, K. Kimball, and F. Solensky, “Considerations for IGMP and MLD Snooping Switches,” IETF Internet Draft draft-ietf-magma-snoop-12.txt (work in progress), Feb. 2005.

[23] M. Bagnulo, I. Soto, A. Garcia-Martinez, and A. Azcorra, “Random Generation of Interface Identifiers,” IETF Internet Draft draft-soto-mobileip-random-iids-00.txt (work in progress), Jan. 2002.

[24] R. Droms, J. Bound, B. Volz, T. Lemon, C. E. Perkins, and M. Carney, “Dynamic Host Configuration Protocol for IPv6 (DHCPv6),” IETF RFC 3315, July 2003.

[25] P. Nikander, J. Arkko, T. Aura, G. Montenegro, and E. Nordmark, “Mobile IP Version 6 Route Optimization Security Design Background,” IETF RFC 4225, Dec. 2005.

[26] “Kame-Shisa,” Mobile IPv6 for FreeBSD 5.4, Nov. 2005. [Online]. Available: http://www.kame.net/newsletter/20041211/shisa.html

[27] V. Nuorvala, H. Petander, and A. Tuominen, “Mobile IPv6 for Linux (MIPL),” Nov. 2005. [Online]. Available: http://www.mobile-ipv6.org/

[28] C. Vogt, R. Bless, M. Doll, and T. Küfner, “Early Binding Updates for Mobile IPv6,” in *Proceedings of the IEEE Wireless Communications and Networking Conference*, vol. 3. IEEE, Mar. 2005, pp. 1440–1445.

[29] C. Vogt, “Early Binding Updates for Mobile IPv6,” IETF Internet Draft draft-vogt-mobopts-early-binding-updates-00.txt (work in progress), Feb. 2005.

[30] ——, “Credit-based authorization for concurrent ip-address tests,” in *Proceedings of the IST Mobile and Wireless Communications Summit*, Jun 2005.

[31] C. Vogt and J. Arkko, “Credit-Based Authorization for Mobile IPv6 Early Binding Updates,” IETF Internet Draft draft-vogt-mobopts-credit-based-authorization-00.txt (work in progress), Feb. 2005.

[32] A. E. Yegin (Ed.), “Link-layer Event Notifications for Detecting Network Attachments,” IETF Internet Draft draft-ietf-dna-link-information-03.txt (work in progress), Oct. 2005.

[33] H. Petander and E. Perera, “Improved Binding Management for Make Before Break Handoffs in Mobile IPv6,” IETF Internet Draft draft-petander-mip6-mbb-00.txt (work in progress), Oct. 2005.

[34] Y.-H. Han, J. Choi, H.-J. Jang, and S. D. Park, “Advance Duplicate Address Detection,” IETF Internet Draft draft-han-mobileip-adad-01.txt (work in progress), July 2003.

[35] T. Narten and R. Draves, “Privacy Extensions for Stateless Address Autoconfiguration in IPv6,” IETF RFC 3041, Jan. 2001.

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

[37] C. E. Perkins, “Securing Mobile IPv6 Route Optimization Using a Static Shared Key,” IETF Internet Draft draft-ietf-mip6-precfgkbm-04.txt (work in progress), Oct. 2005.

[38] F. Dupont and J.-M. Combes, “Using IPsec between Mobile and Correspondent IPv6 Nodes,” IETF Internet Draft draft-ietf-mip6-cn-ipsec-02.txt (work in progress), Dec. 2005.

[39] C. Castellucia, G. Montenegro, J. Laganier, and C. Neumann, “Hindering Eavesdropping via IPv6 Opportunistic Encryption,” in *Proceedings of the European Symposium on Research in Computer Security, Lecture Notes in Computer Science*. Springer-Verlag, Sept. 2004, pp. 309–321.

[40] G. Montenegro and C. Castelluccia, “Crypto-based Identifiers (CBIDs): Concepts and Applications,” *ACM Transactions on Information and System Security*, vol. 7, no. 1, pp. 97–127, Feb. 2004.

[41] J. Arkko, C. Vogt, and W. Haddad, “Applying Cryptographically Generated Addresses and Credit-Based Authorization to Mobile IPv6,” IETF Internet Draft draft-arkko-mipshop-cga-cba-02.txt (work in progress), Oct. 2005.

[42] G. O’Shea and M. Roe, “Child-Proof Authentication for MIPv6 (CAM),” *ACM SIGCOMM Computer Communication Review*, vol. 31, no. 2, pp. 4–8, Apr. 2001.

[43] A. Dutta, T. Zhang, Y. O. Kenichi Taniuchi, and H. Schulzrinne, “MPA assisted Optimized Proactive Handoff Scheme,” in *Proceedings of the ACM Conference on Mobile and Ubiquitous Systems*. ACM Press, July 2005.

[44] H. Soliman, C. Castelluccia, K. E. Malki, and L. Bellier, “Hierarchical Mobile IPv6 Mobility Management (HMIPv6),” IETF RFC 4140, Aug. 2005.

[45] A. Mishra, M. Shin, and W. Arbaugh, “An empirical analysis of the IEEE 802.11 MAC layer handoff process,” *ACM SIGCOMM Computer Communication Review*, vol. 33, no. 2, pp. 93–102, Apr. 2003.

[46] A. Alimian and B. Aboba, “Analysis of Roaming Techniques,” IEEE Contribution 11-04-0377r1, Mar. 2004.

[47] H. Velayos and G. Karlsson, “Techniques to Reduce IEEE 802.11b MAC Layer Handover Time,” Laboratory for Communication Networks, Department of Microelectronics and Information Technology (IMIT), KTH, Royal Institute of Technology, Stockholm, Sweden, Technical Report TRITA-IMIT-LCN R 03:02, Apr. 2003.

[48] J.-C. Chen, M.-C. Jiang, and Y.-W. Liu, “Wireless LAN Security and IEEE 802.11i,” *IEEE Wireless Communications*, vol. 12, no. 1, pp. 27–36, Feb. 2005.

[49] A. Mishra, M. H. Shin, N. L. Petroni, Jr., T. C. Clancy, and W. A. Arbaugh, “Proactive Key Distribution Using Neighbor Graphs,” *IEEE Wireless Communications*, vol. 11, no. 1, pp. 26–36, Feb. 2004.