Chapter 18

Handoff Initiation in Mobile IPv6

Torben W. Andersen, Anders Lildballe and Brian Nielsen[!]

Department of Computer Science , Aalborg University, Denmark

18.1 Introduction

The Internet Protocol (IP) is expected to become the main carrier of traffic to mobile and wireless nodes. This includes ordinary data traffic like HTTP, FTP and e-mail, as well as voice, video and other time sensitive data. To support mobile users, the basic Internet protocols have been extended with protocols (Mobile IP) for intercepting and forwarding packets to a mobile and possibly roaming node. Seamless roaming requires that users and applications do not experience loss of connectivity or any noticeable hick-ups in traffic. This is not only important for time sensitive traffic, but also for TCP (Transmission Control Protocol) based traffic, as TCP performance is highly sensitive to packet loss and re-ordering.

It is therefore imperative that a handoff is initiated in such a way that network connectivity is maintained for the longest possible period of time, and that the *handoff latency* and packet loss is minimized. However, little is known about the performance of the Mobile IP protocols in an actual network. In particular, it is not understood how different *handoff initiation algorithms* influence essential performance metrics like the packet loss and the duration of a handoff.

[!] Author for contact, Email: bnielsen@cs.auc.dk

To improve this situation this chapter studies the performance of two basic *handoff initiation algorithms* denoted Eager Cell Switching (ECS) and Lazy Cell Switching (LCS) [1] respectively. The study uses both a theoretical approach that derives a *mathematical model* for handoff latency, and an empirical approach that includes experiments in a *Mobile IPv6 testbed* and an office building. Also the chapter compares ECS and LCS with a novel proactive strategy, Parametric Cell Switching (PCS), that considers link layer information about signal quality.

18.1.1 Mobile IPv6 Operation

A Mobile IP node is associated with a permanent *home network* and is assigned a static IP *home address* of the home network. The home address identifies the node globally. When the node is attached to a *foreign network* it selects one router as its default router, and obtains an additional IP address known as the *care-of address,* which identifies the current location of the mobile node. The network to which the mobile node is currently attached is called its *primary network*. In the basic mode of operation a *correspondent node* sending packets to the mobile node address these to its home address. A router serving as *home agent* must be present at the home network. The home agent is responsible for tunneling IP packets sent to the home address to the mobile nodes current *care-of address*. The result is the triangular routing depicted in Figure 18.1.

Figure 18.1 Basic Mobile IP Operation.

A *handoff* in Mobile IP occurs when the mobile node switches from one (foreign) network to another, and thus obtains a new care-of address. It then registers the new care-of address at the home agent by sending it a *binding update* message. The home agent adds this to its binding cache and replies with a binding acknowledgement. Thus, packets tunneled after the mobile node has changed network but before the binding update has reached the home agent may be lost. The mobile node may also decide to send binding updates to the correspondent node(s) allowing it to address the mobile node directly and thereby bypass the home agent. This avoids triangular routing, and is referred to as *route optimization*.

To discover new networks the mobile node listens for *router advertisements* broadcasted periodically from access routers. A router advertisement contains the network prefix of the advertised network, and a lifetime denoting how long this prefix can be considered valid. The mobile node may then use statefull or stateless auto-configuration to generate its new care-of address. In stateless auto-configuration a host generates its own IP address based on the network prefix learned through router advertisements and the IEEE 802 address of its network interface. Statefull auto-configuration involves additional communication with a DHCP (Dynamic Host Configuration Protocol) Server, but also allows configuration of other network services such as DNS (Domain Name Service).

A *handoff algorithm* has three major responsibilities, 1) detecting and quality assessing available networks, 2) deciding whether to perform a handoff, and 3) executing the handoff. Handoff *initiation* consists of the first two activities. A *seamless handoff* requires that no packets are lost as consequence of the handoff. In general it is also desirable that packets are not reordered, duplicated, or extraordinarily delayed.

18.1.2 Handoff Initiation

The Mobile IPv6 specification [2] contains only a weak specification of handoff initiation algorithms. Two conceptually simple handoff initiation algorithms that have gained considerable interest are Eager Cell Switching (ECS) and Lazy Cell Switching (LCS) [1]. Both operate at the network layer without requiring information from the lower (link) layers.

First consider the scenarios depicted in Figure 18.2. Here the ranges of two wireless networks (1 and 2) are depicted as circles. A mobile user moves from point A to point B. In the situation shown in Figure 18.2a, where the networks do not overlap, no Mobile IP handoff initiation algorithm could avoid losing packets (one might imagine a very elaborate infrastructure where packets were multicasted to all possible handoff targets and that packets could be stored there until the mobile node arrives, but even then a long period without network access would most likely be noted by the user). In contrast, if the networks overlap sufficiently as shown in Figure 18.2b, seamless handoff is possible. Figure 18.2c shows that there are situations where a handoff is possible, but not desirable.

Figure 18.2 A node moves from point A to B. No seamless handoff is possible (a), seamless handoff theoretically possible (b), seamless handoff possible, but should not be performed (c).

ECS *proactively* initiates a handoff every time a new network prefix is learned in a router advertisement. Conversely, LCS acts *reactively* by not initiating a handoff before the primary network is confirmed to be unreachable. When the lifetime of the primary network expires, LCS probes the current default router to see if it is still reachable. If not, a handoff to another network is initiated.

Figure 18.3 Lazy Cell Switching.

Figure 18.4 Eager Cell Switching.

Consider what happens when ECS and LCS are subjected to the movement in Figure 18.2b. Figure 18.3 illustrates a time line for LCS where significant events have been pointed out. The first event is that network 2 gets within range. However, the mobile node cannot in general observe this before it receives a router advertisement from network 2. This results in a *network discovery* delay. LCS does not perform a handoff yet. Next network 1 gets out of range. This cannot in general be detected immediately as this requires active communication with the base station, and gives rise to a *network loss discovery* delay. LCS declares the network unreachable when the lifetime of the last received router advertisement has expired, and the following probing is unsuccessful. LCS then hands off to one of the alternative networks known to it through router advertisements, in this case network 2, and thus establishes a new point of attachment. Thus LCS will loose packets in the handoff latency interval.

The behavior of ECS is illustrated in Figure 18.4. ECS hands off immediately when a new network is discovered. If the mobile node has an interface that is capable of receiving from the old network while attaching to the new, a seamless handoff can be performed provided a sufficient network overlap.

The performance of ECS thus depends on the frequency with which access routers are broadcasting router advertisements. Similarly, LCS also depends on the frequency of broadcasted router advertisements, but additionally depends on the lifetime of network prefixes and probing time. The theory and data needed to decide what handoff initiation algorithms to use in what circumstances, how to tune protocol parameters, and where to put optimization efforts, are missing.

Both ECS and LCS are very simple-minded approaches. *Parametric Cell Switching* (PCS) is our proposal for a more intelligent handoff initiation algorithm that also considers measured signal-to-noise ratio and roundtrip time to access routers. Periodically (currently every 0.5 seconds) the algorithm sends an echo request to the default router at all available networks. The default routers are expected to reply to the echo. These echo requests are sent for three reasons:

- It is only possible to measure the signal-to-noise ratio of a link to a network if there is traffic at the network.
- It can be determined faster that a network has become unreachable than by monitoring the lifetime of network prefixes.
- The roundtrip time is an indication of the capacity of a network.

Parametric Cell Switching only performs handoff when a significantly better network is available [8]. The cost of using the *Parametric Cell Switching* algorithm is a slightly increased network load.

18.1.3 Handoff Performance

There is a number of important metrics that should be considered when evaluating the performance of a handoff initiation strategy as experienced by a mobile node:

> **Handoff latency:** The handoff latency is the period of time where the mobile node is potentially unreachable. In general, it is caused by the time used to discover a new network, obtain and validate a new care-of address, obtain authorization to access the new network, make the decision that a handoff should be initiated, and finally execute the handoff which

involves notifying the home agent of the new care-of address, and awaiting the acknowledgement from the home agent.

Number of performed handoffs: The more handoffs a given strategy will perform in a given scenario the more likely it is that the user will observe them, and the more the network is loaded by signaling messages.

User value: When several networks are candidates as target for a handoff, the one most optimal from the users perspective should be chosen. This may be the network that offers the most bandwidth, cheapest price, the most stable connection etc.

18.1.4 Comparison with Other IP Mobility Schemes

The basic mobility concept of IPv6 is very similar to IPv4. However, the required support for mobility is better embedded in the core IPv6 protocols. Besides expanding the address space, IPv6 provides stateless auto configuration, (proxy) neighbor discovery, and more flexible header extensions and options. For instance, the *home address* option is used in a packet sent by a mobile host to inform the receiver about the mobile node's home address, which is the real source address of the message. The mobile node cannot use its home address as source because a router performing ingress filtering will drop packets with a source address that contains a network prefix different from the prefix of the router's interface at which the packet arrived. The *binding update* option is used to inform the home agent and possibly also corresponding nodes (for route optimization) about its current care-of address. The *routing header* is used to specify a set of intermediate nodes the packet must traverse on its path to its destination. It reduces the tunneling overhead and simplifies treatment of forwarded ICMP (Internet Control Message Protocol) packets.

The main architectural difference is that the *foreign agent* that assigns care-of addresses and forwards tunneled messages received from the home agent to the mobile node in mobile IPv4 is absent in IPv6.

Currently, most work on handoff in Mobile IP is concerned with *handoff execution* using micro mobility schemes such as e.g. HAWAII (Handoff-Aware Wireless Access Internet Infrastructure) [4], and Cellular IP [5]. Micro mobility aims at reducing the network load and handoff latency in environments of limited geographical span but with many frequently migrating nodes by minimizing and localizing the propagation of *binding updates* and *binding acknowledgements*.

For example, in Cellular IP a wireless access network consisting of a collection of interconnected Cellular IP nodes functioning as simple routers and base stations provides Internet access to wireless nodes via a gateway. The required signaling for an intra access network handoff is purely local. Each Cellular IP node maintains a routing cache that maps the home address of the mobile node to the last of its neighbors that forwarded a packet sent by the mobile node, i.e., the cellular IP nodes maintains a soft state path from the mobile node to the gateway, and uses the reverse path as route to the mobile node. Active mobile nodes refresh the path by periodically sending real or dummy packets containing a route update option. After handoff, the new path is established automatically by the transmitted route update packets. The locations of idle nodes are not tracked accurately. Rather idle nodes are paged when needed. This reduces the network load. When handoff to another access network is needed, Mobile IPv6 is used.

In a *hard handoff* the packets transmitted via the old path is lost until the route update reaches the cross over point of the new and old path. Cellular IP also provides a *semi-soft handoff* mechanism that can be used to reduce packet loss during a handoff. The mobile node can request in its route update packet that packets are to be forwarded both along the old and new route, i.e., normal message forwarding is turned into a multicast operation.

Like Mobile IP, handoffs are initiated by the mobile node based on its knowledge of available networks that it has learned from beacon signals similar to Mobile IP router advertisements emitted periodically by base stations. No particular handoff *initiation strategy* seems to be assumed or proposed.

While reducing handoff execution time is important, this is insufficient to obtain seamless handoffs, especially when the network topology is not controllable as is the case in the Internet. It seems beneficial to take a *proactive* approach to handoff initiation that has the potential of reducing the packet loss to zero. This strategy has also proven successful in existing wireless networks such as GSM [6].

18.2 Mathematical Models

Because the focus of the models is the performance of handoff initiation algorithms, they do not include the propagation delay of *binding updates* and *binding acknowledgements* used in the complete handoff procedure. The handoff (initiation) latency is the time from the current primary network gets out of range until an event occurs at the mobile node that triggers it to perform a handoff. To obtain the total handoff latency, the round-trip time to the home agent must be added to the numbers stated in this chapter. A typical number used for the continental US is approximately 200ms.

The goal is to predict the variation in handoff latency and its average as a function of the primary protocol parameters. The variation in latency is important to real-time sensitive traffic because it also indicates the best and worst case amount of time the mobile node risks being unreachable. The idea is therefore to derive a density function from which average and variation of handoff latency can be computed.

18.2.1 Basic Definitions

The mathematical models assume perfect cell boundaries, i.e. getting within range of a new network coincides with leaving the range of the primary network. According to [7] a router must pick a random delay between each broadcast of an unsolicited router advertisement in order to avoid routers synchronizing. The minimum possible time between two consecutive router advertisements is denoted $R_{\min}$ and the maximum time between two consecutive router advertisements $R_{\max}$ meaning that the period between any two router advertisements must be found in the interval $[R_{\min}, R_{\max}]$. An additional simplifying assumption is that all access routers broadcasts with the same frequency range. The time at which a mobile node enters the range of a new network is denoted by C_{time}.

The lifetime of broadcasted network prefixes is denoted by T_1. It is assumed that the time it takes to probe a default router is distributed uniformly within the interval $[Q_{\min}, Q_{\max}]$.

18.2.2 Eager Cell Switching

For ECS the theoretical handoff latency L_{eager} is the period from getting out of range of the primary network until the reception of a router advertisement from a new network. This is illustrated in Figure 18.5a.

Figure 18.5 Model for computing L_{eager} (a), and density function $P_{L_{\text{eager}}}(l)$ for ECS (b).

The goal is a density function for L_{eager} that can be used to derive the average value of L_{eager}. Because the primary variables C_{time} and R are random but dependent, their joint density function $P_{C_{\text{time}},R}(c_{\text{time}},r)$ must first be computed. It can be expressed as

$$P_{C_{\text{time}},R}(c_{\text{time}},r) = P_{C_{\text{time}}|R}(c_{\text{time}} \mid r) \cdot P_R(r) \tag{1}$$

where $P_{C_{\text{time}}|R}(c_{\text{time}} \mid r)$ is the probability distribution for C_{time} given R. As C_{time} is evenly distributed in the interval $[0,R]$ the probability distribution $P_{C_{\text{time}}|R}(c_{\text{time}} \mid r)$ can be calculated as

$$P_{C_{\text{time}}|R}(c_{\text{time}} \mid r) = \frac{1}{r} \cdot 1_{c_{\text{time}} \in [0,r]} \tag{2}$$

where $1_{c_{\text{time}} \in [0,r]}$ is an indicator function with a value of 1 when $c_{\text{time}} \in [0,r]$ and 0 otherwise.

The density function $P_R(r)$ expresses the probability that C_{time} should occur in an interval of size $R = r$. Intuitively, the probability of C_{time} occurring in an interval is proportional to the size of the interval. When the interval size is given as r, the function $f(r) = r \cdot 1_{r \in [R_{\min}, R_{\max}]}$ exhibits this intuitive property. The density function $P_R(r)$ is obtained as $f(r)$ divided with the area of $f(r)$. This yields

$$\begin{aligned} P_R(r) &= \frac{r}{\int_{R_{\min}}^{R_{\max}} r\,dr} \cdot 1_{r \in [R_{\min}, R_{\max}]} \\ &= \frac{2r}{R_{\max}{}^2 - R_{\min}{}^2} \cdot 1_{r \in [R_{\min}, R_{\max}]} \end{aligned} \tag{3}$$

The joint density function $P_{C_{\text{time}},R}(c_{\text{time}},r)$ obtained by combining these results becomes

$$\begin{aligned} P_{C_{\text{time}},R}(c_{\text{time}},r) &= P_{C_{\text{time}}|R}(c_{\text{time}} \mid r)\cdot P_R(r) \\ &= \frac{2}{R_{\max}{}^2 - R_{\min}{}^2}\cdot 1_{c_{\text{time}}\in[0,r]}\cdot 1_{r\in[R_{\min},R_{\max}]} \end{aligned} \tag{4}$$

The density function $P_{L_{\text{eager}}}(l)$ for L_{eager} can be obtained by integrating over the joint density function $P_{C_{\text{time}},R}(c_{\text{time}},r)$ for all possible values of c_{time} and r. As r can be expressed as $r = c_{\text{time}} + l$ (which follows from $l = r - c_{\text{time}}$), we have

$$\begin{aligned} P_{L_{\text{eager}}}(l) &= \int_{-\infty}^{\infty} P_{C_{\text{time}},R}(c_{\text{time}},c_{\text{time}}+l)\,dc_{\text{time}} \\ &= \frac{2}{R_{\max}{}^2 - R_{\min}{}^2}\cdot 1_{l\in[0,R_{\max}]}\left[c_{\text{time}}\right]_{\max(0,R_{\min}-l)}^{R_{\max}-l} \\ &= \begin{cases} \dfrac{2}{R_{\max}+R_{\min}}, & \mathit{if}\ 0 \le l \le R_{\min} \\ \dfrac{2(R_{\max}-l)}{R_{\max}{}^2 - R_{\min}{}^2}, & \mathit{if}\ R_{\min} \le l \le R_{\max} \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{5}$$

In Figure 18.5b the density function $P_{L_{\text{eager}}}(l)$ from Equation 5 is plotted. Observe that when using ECS, the handoff latency is bounded by the value of $R_{\max}$ and that there is the highest probability of obtaining handoff latencies in the range $[0, R_{\min}]$.

Given the density function $P_{L_{\text{eager}}}(l)$ it is easy to obtain the average value of L_{eager}, denoted by $\overline{L}_{\text{eager}}$, by integrating over the product of $L_{\text{eager}} = l$ and $P_{L_{\text{eager}}}(l)$ for all possible values of L_{eager}:

$$\begin{aligned} \overline{L}_{\text{eager}} &= \int_0^{R_{\max}} l \cdot P_{L_{eager}}(l)\, dl \\ &= \frac{R_{\max}{}^3 - R_{\min}{}^3}{3(R_{\max}{}^2 - R_{\min}{}^2)} \end{aligned} \tag{6}$$

18.2.3 Lazy Cell Switching

For LCS the theoretical handoff latency L_{lazy} is the time from leaving the range of the primary network until concluding that the primary network is unreachable. This occurs when the lifetime of the primary network has expired and probing it has failed. Figure 18.6 depicts the used model where the last router advertisement from the primary network was received at time 0. The goal is a density function for L_{lazy} that can be used to calculate the average value of L_{lazy}.

Figure 18.6 Model for computing L_{lazy}.

Intuitively, the handoff latency consists of the remaining lifetime of the primary network plus the probing time Q used to determine that the primary network is unreachable. The remaining lifetime is the lifetime T_1 of the last router advertisement minus the time the network was reachable, C_{time}. The handoff latency L_{lazy} can thus be expressed as

$$\begin{aligned} L_{\text{lazy}} &= \text{Lifetime remaining of primary network} + Q \\ &= T_1 - C_{time} + Q \end{aligned} \tag{7}$$

Because Q is assumed to be uniformly distributed within the interval $[Q_{\min}, Q_{\max}]$ the density function for Q is

$$P_Q(q) = \frac{1}{Q_{\max} - Q_{\min}} \cdot 1_{q \in [Q_{\min}, Q_{\max}]} \tag{8}$$

Using a similar line of reasoning and method of calculation as outlined for ECS, the density function for LCS denoted $P_{L_{\text{lazy}}}(l)$ can be computed. That is, the joint density function of the involved parameters is integrated for all possible values. However, the intermediate calculations are more involved, and only the result is stated in Equation 9. Further details on its derivation can be found in [8].

$$\begin{aligned} P_{L_{\text{lazy}}}(l) &= \int_{-\infty}^{\infty} P_Q(l - l_{np}) \cdot P_{C_{\text{time}}}(T_1 - l_{np})\, dl_{np} \\ &\quad \text{where } P_{C_{\text{time}}}(l) = P_{L_{eager}}(l) \\ &\quad \text{and } l_{np} = T_1 - c_{\text{time}} \end{aligned} \tag{9}$$

The average handoff latency $\overline{L}_{\text{lazy}}$ can be computed by integrating over the density function, or more simply, directly from Equation 7 that states that $L_{\text{lazy}} = T_1 - C_{time} + Q$. The average latency $\overline{L}_{\text{lazy}}$ can be computed as the average remaining lifetime plus the average probing time $\overline{Q}$:

$$\begin{aligned} \overline{L}_{\text{lazy}} &= T_1 - \overline{C}_{time} + \overline{Q} \\ &= T_1 - \frac{R_{\max}^{\;3} - R_{\min}^{\;3}}{3(R_{\max}^{\;2} - R_{\min}^{\;2})} + \frac{1}{2}(Q_{\max} - Q_{\min}) \end{aligned} \quad (10)$$

The average remaining lifetime of the primary network is the lifetime of the last received router advertisement minus the average amount of time that the primary network was reachable. On average the primary network is reachable for an amount of time corresponding to the time used to discover a new network. This corresponds exactly to the average handoff latency for ECS. Thus, $\overline{C}_{\text{time}}$ equals $\overline{L}_{\text{eager}}$ calculated in Equation 6.

18.3 Experimental Results in Testbed

Section 18.2 presented the mathematical models needed to compute the handoff latency as a function of essential protocol parameters. This section presents the design of a Mobile IPv6 testbed and compares the theoretically predicted handoff latency with the handoff latency experienced by a mobile node in the Mobile IPv6 testbed.

18.3.1 The Mobile IPv6 Testbed

The testbed is depicted in Figure 18.7 and consist of four nodes; three routers and one host. The three routers, *iridium*, *platin* and *nikkel*, are assigned an IPv6 prefix for each network device. The mobile node, *lantan*, is manually assigned an IPv6 address at the fec0:0:0:1::/64 network, its home network. When *lantan* is not at its home network, it uses *stateless auto-configuration* to obtain an IPv6 address as its care-of address. The home agent is located at *iridium* that also hosts an application corresponding with an application at the mobile node. The mobile node can roam between the two access routers *platin* and *nikkel*. The link media used in the experiments reported here are standard 802.3 10 Mbit/s Ethernet devices. However, the testbed also runs 802.11b 11 Mbit/s Wireless LAN connections. The connection between the access routers *platin* and *nikkel* allows them to coordinate on whom should offer access to the mobile node. It also allows experiments with route optimizations because it offers an alternative path to the mobile node.

Figure 18.7 Testbed with IPv6 addresses and network prefixes.

All nodes runs FreeBSD version 4.1 [9]. On top of FreeBSD the KAME package [10] is installed. The KAME package includes Mobile IPv6 support and IPsec support. The KAME package installed is the weekly snap-release of 25/9-2000. A snap-release is the newest version of the package and may include functionality that is still under development and is not fully tested. The Mobile IPv6 code supplied with KAME is an example of such functionality. The Mobile IPv6 implementation included in KAME can be configured to use either the ECS or the LCS handoff initiation algorithm.

18.3.2 Experimental Approach

Two different scenarios have been emulated in the testbed. In the *no network overlap scenario* perfect cell boundaries are assumed. A mobile node moving out of the range of one network therefore coincides with the mobile node moving within range of another network. This scenario corresponds to the one for which mathematical models were derived in Section 18.2. In the *network overlap scenario* cell boundaries are overlapping between the two networks and the mobile node is always able to reach at least one network. This scenario was applied to investigate whether ECS was able to avoid packet loss when able to receive and send packets via two networks at the same time. LCS behaves identically in both scenarios.

The mobile node moving in and out of the range of a network has been emulated by operating a firewall at the access routers. When the firewall is enabled, the mobile node is not able to receive or send any traffic through that particular access router. Accordingly, when the firewall is disabled all traffic is allowed to pass to and from the mobile node.

To determine the handoff latency in an experiment we applied the following method:

- UDP packets are sent from the correspondent node to the mobile node. Each packet contains a send timestamp and a sequence number. The UDP packets are send with a random interval between 95 ms and 105 ms. The interval is randomized to make sure the network does not adjust itself to any particular sending frequency.
- The UDP packets are received and time stamped at the mobile node. The sequence number, the send and receive timestamp are stored upon reception of a packet as an entry in a log file.
- A handoff is registered from packets missing in the log file. We compute the measured handoff latency by multiplying the number of lost packets with the average period between sending

packets (0.1 seconds). The precision of the measured latency is thus ±0.1 seconds. If a handoff is performed without losing packets it will therefore not be registered. Both the average and the frequency distribution of handoff latencies can be computed by inspecting the log.

By reducing the interval between UDP packets (increasing frequency) a higher accuracy will be obtained, and the measured latencies will approach the theoretical latencies defined in Section 18.2. The interval of 95 ms to 105 ms was chosen to avoid too many UDP packets being sent. Due to a memory leakage in the KAME Mobile IPv6 software only a limited number of packets can be sent from a correspondent node before it crashes. In the experiments presented in this chapter it was possible to perform 300 to 400 handoffs in sequence before the correspondent node crashed. Further confidence in the mathematical models has been obtained by implementing a simulator in JAVA. This simulator has confirmed the theoretically predicted density functions for a range of configurations for both ECS and LCS.

18.3.3 Overview of Performed Experiments

The Mobile IPv6 testbed has been used to perform the following experiments:

Default configuration: In this experiment the router advertisement interval and network prefix lifetime is set as recommended in [2]. This means a router advertisement interval randomly chosen between 0.5 and 1.5 seconds and a lifetime of 4 seconds. The purpose of this experiment is to reveal handoff latency using the default configuration. In the latest versions of the Mobile Ipv6 specification [3] the minimum time between router advertisements has been reduced from 0.5 seconds to 0.05 seconds. However, this change has little impact on the results of this chapter.

Latency as a function of router advertisement interval: Handoff performance is measured for different router advertisements intervals, but with an identical network prefix lifetime. The purpose of this experiment is to investigate how the interval between sending router advertisements affects the handoff latency.

Latency as a function of network prefix lifetime: In this experiment the handoff latency is measured for different network prefix lifetimes, but with a fixed range for the intervals between sending router advertisements. The purpose is to investigate how the lifetime of router advertisements affects the handoff latency.

All experiments have been performed using both the network overlap and the no network overlap scenario. nly a selection of the empirically obtained results can be presented here. The full set of results is given in [8]. All plots also show the theoretically predicted handoff latency such that the theoretical and empirical results can easily be compared.

18.3.4 Default Settings

First the theoretically predicted probability distributions are compared with the measured frequency distributions using the default configuration of access routers in the no overlap scenario. Next their performance is compared numerically.

The histogram in Figure 18.8a depicts the frequency distribution of the experimentally measured handoff latencies for ECS. The continuous line shows the theoretically predicted density function. The observed handoff latencies lie in the range 0.1 to 1.5 seconds. Also note that no handoff latencies in the interval [0,0.1] are present. This is caused by the experimental setup in which the precision is limited by the frequency of packets from the corresponding node, i.e. latencies below 0.1 seconds cannot be observed. Instead these handoff latencies are recorded in the [0.1,0.2] interval. In conclusion, the experimental results for ECS conforms well to those predicted by the mathematical models.

Figure 18.8 Frequency distribution for handoff latency using ECS and default configuration in the no network overlap scenario (a), and fequency distribution for handoff latency using LCS and default configuration in the no network overlap scenario (b) .

Figure 18.8b depicts the histogram obtained for LCS. Here the observed latencies range from 2.5 seconds to 5 seconds with most values centered around 4 seconds. It can be seen that LCS is generally unable to avoid packet loss as it does not initiate a handoff before after the primary network has become unavailable. In conclusion, the empirically obtained frequency distribution conforms well to the density function predicted by the mathematical model.

Table 18.1 Summary of expected and actual results for the no network overlap setup using default router configuration. The probing time Q for LCS is assumed to be in the interval [0,1].

Handoff strategy	L [s] **(Theory)**			Latency [s]		
	Avg	**Min**	**Max**	**Avg**	**Min**	**Max**
Eager	0.54	0	1.5	0.54	0.10	1.52
Lazy	3.96	2.5	5.0	3.97	2.54	4.97

The comparison of ECS and LCS is summarized in Table 18.1 When no overlap between network ranges exists, ECS yields an average handoff latency of 0.54 seconds. This corresponds to the time it takes to discover the new network. LCS yields a much worse latency with an average of 3.96 seconds. Also best- and worst-case values are higher. With respect to handoff latency ECS outperforms LCS. Furthermore, the experiments with overlapping networks show that ECS is able to avoid packet loss altogether during a handoff, provided that a sufficient overlap between network ranges exists.

However, a disadvantage of ECS is that it always performs a handoff when discovering a new network, whether or not this offers stable connectivity. Consequently, ECS will likely perform unnecessarily many handoffs resulting in an increased network load and loss of connectivity. In conclusion, both ECS and LCS have serious performance lacks, but the performance of ECS indicates that proactive handoff initiation has the potential to avoid packet loss.

18.3.5 Varying Advertisement Frequency

One of the primary protocol parameters is the frequency of router advertisements. Its effect on handoff latency is shown in Figure 18.9 that plots average handoff latency as a function of the interval between broadcasting router advertisements.

Figure 18.9 The average handoff latency as a function of router advertisement interval for ECS and LCS in the no network overlap scenario.

ECS behaves like what intuitively would be expected: A higher frequency implies that networks are discovered sooner, which again implies faster handoffs. Surprisingly however, observe that the LCS latency is actually decreasing when the interval between broadcasting router advertisements is increased. The explanation for this is that the lifetime is fixed at a constant value of 5 seconds in this experiment. This result thus indicates that the handoff latency for LCS can be minimized by configuring access routers with a prefix lifetime very close to the maximum interval between broadcasting router advertisements.

18.4 Optimizing Protocol Configuration

This section demonstrates that the default configuration of access routers proposed in [2] does not result in optimal handoff performance for either ECS, or LCS. In [2] it is suggested that a router should broadcast unsolicited router advertisements distanced by a random period chosen from the interval [0.5,1.5]. This gives an average network load of one router advertisement every second. In later versions of the Mobile IPv6 specification [3] the default configuration has changed the router advertisement period to [0.05,1.5] seconds. This gives a slight increase in average network load to 1.3 router advertisements per second. Using these values our theoretical models predict that the average handoff initiation latency for ECS is 0.5 seconds and LCS is 4 seconds, i.e., a small improvement for ECS and a small drawback for LCS. However, as demonstrated in the following the models can be used to find a new set of parameters that reduce handoff latency *without* increasing the network load. Similarly, the models can be used to find the optimal settings should an increased network load be accepted.

Using the same network load as the suggested rate of an average of one advertisement per second, the average ECS latency $\overline{L}_{\text{eager}}$ can be minimized by adjusting $R_{\min}$ and $R_{\max}$ in Equation 6 subject to the constraint that the sum of $R_{\min}$ and $R_{\max}$ must equal 2. Close inspection reveals that ECS performs best when $R_{\min}$ and $R_{\max}$ are configured with values as close together as possible. Optimal performance for ECS can therefore be obtained when both $R_{\min}$ and $R_{\max}$ are set to a value of 1.

The same method applied to LCS reveals that LCS performs better when $R_{\min}$ and $R_{\max}$ are configured with values far from each other. For LCS optimal performance can therefore be obtained by configuring $R_{\min}$ to a value of 0 and $R_{\max}$ to a value of 2. This is in direct contradiction to the optimal configuration for ECS. However, for LCS the dominating factor for the handoff latency is the lifetime of broadcasted network prefixes. As this lifetime cannot be configured to be lower than the value of $R_{\max}$, a reduction of $R_{\max}$ (which is the case when $R_{\min}$ and $R_{\max}$ is configured to have values close to each other) can also benefit the performance of LCS, if the lifetime is configured close to the value of $R_{\max}$.

In Table 18.2 theoretical values of handoff latency for three different configurations of access routers are shown. The probing time for LCS is assumed to be in the interval [0,1].

Table 18.2 The mathematically predicted handoff latency in the no network overlap scenario for different configurations.

Handoff strategy	Router Configuration			L [s] (Theory)		
	$R_{\min}$	$R_{\max}$	T_1	Avg	Min	Max
Eager	0.5	1.5	4	0.54	0	1.5
Lazy	0.5	1.5	4	3.96	2.5	5.0
Eager	0.9	1.1	1.1	0.5	0	1.1
Lazy	0.9	1.1	1.1	1.1	0	2.1
Lazy	0	2	2	1.83	0	3.0

Observe that both ECS and LCS performs better with the proposed configuration of access routers with $R_{\min} = 0.9$ and $R_{\max} = 1.1$. For ECS the average handoff latency is reduced from 0.54 seconds to 0.50 seconds and the worst-case handoff latency is reduced from 1.5 to 1.1 second. Similarly, for LCS the average handoff latency is reduced from 3.96 to 1.1 seconds and the worst-case handoff latency is reduced from 5.0 to 2.1 seconds. The new proposed settings thus simultaneously improve on average, best, and worst-case for both ECS and LCS.

Observe from the last row in Table 18.2 that for LCS the advantage of configuring $R_{\min}$ and $R_{\max}$ far from each other is out-weighted by the fact that the lifetime T_1 has to be configured at a higher value.

The performance of the new settings has been tried out in the testbed. The experiment confirmed the theoretically predicted values [8].

An alternative to reducing the lifetime of router advertisement messages is to exploit the *advertisement interval option* in router advertisements proposed in [2]. This option contains the maximum time ($R_{\max}$) between router advertisements that mobile nodes should expect. This would allow a mobile node to probe its default router if no router advertisement has been received for a period corresponding to the value of $R_{\max}$. This in effect forces LCS to become more proactive.

18.5 Building Wide Experiment

The following simple experiment compares the handoff performance of ECS, LCS, and Parametric Cell Switching handoff initiation algorithms in a more realistic scenario than the Mobile IPv6 testbed.

18.5.1 Experimental Setup

Three access routers *nikkel*, *blue* and *vismut* are all installed with 802.11b Wireless LAN network devices and are configured with the improved router configuration: $R_{\min} = 0.9$ seconds, $R_{\max} = 1.1$ seconds and $T_1 = 2$ seconds (the value permitted in the testbed nearest the desired 1.1 seconds). These three access routers have been deployed at the Department of Computer Science at 3 different locations. The mobile node *lat11* is carried at normal walking velocity along an approximately 2 minute walk. A single walk was performed for each of the handoff initiation algorithms.

During the experiment a program at *iridium* that acts as home agent and correspondent node, sends UDP packets with intervals randomized between 95 ms and 105 ms. A program at the mobile node *lat11* receives the packets and counts the number of dropped packets and performed handoffs. In addition, it measures the signal-to-noise ratio for each of the three networks once every second during the walk. Figure 18.10 plots the signal-to-noise ratios measured along the path.

Figure 18.10 Measured signal-to-noise ratios from the three access routers when the mobile node is moved along the walking path. The numbers along the x-axis corresponds to different locations along the walking path.

18.5.2 Results

Table 18.3 summarizes the results. Surprisingly, it can be observed that ECS shows the poorest performance. ECS drops many packets presumably because it initiates many handoffs as each network disappears and reappears several times along the path. Often it tries to perform handoffs to networks that are only sporadically available. LCS performs much better. It has reduced the number of handoffs significantly and lost only 10 packets. These were found to be single packets dropped when approaching the maximum range of the primary network.

Table 18.3 Summary of results for building wide experiment.

Handoff strategy	Number of attempted handoffs	Number of lost packets	Average latency [s]	Maximum Latency [s]
Eager	24	48	0.14	0.41
Lazy	3	10	0.10	0.10
Param	2	0	0.00	0.00

PCS shows excellent performance. It keeps the number of handoffs at only two handoffs in this scenario. The number of handoffs is optimal because PCS only initiates a handoff when a network is present which has significantly better signal-to-noise ratio than the primary network and has a stable low round trip time to the access router. In this scenario the signal-to-noise ratio behaves in a way that gives PCS opportunity to handoff to a new network before the connection to the primary network becomes unstable.

More importantly, the results show that it is possible to achieve very good handoff performance using mobile IP, but also that this requires a more intelligent (proactive) initiation algorithms than the proposed eager and lazy cell switching.

18.6 Conclusions and Future Directions

The mathematical models for the Eager Cell Switching (ECS) and the Lazy Cell Switching (LCS) handoff initiation algorithms were found to be able to predict handoff performance. Using a testbed installed with FreeBSD 4.1 and the KAME Mobile IPv6 software, these models were shown to accurately reflect the handoff latency experienced by an actual roaming node. The mathematical models were also used to optimize Mobile IPv6 protocol configuration to reduce the handoff latency without increasing network load due to router advertisements.

The mathematical models have a perspective beyond academic satisfaction and protocol optimization. They can for instance be used to calculate how much cells should overlap. If seamless handoff should be possible they must overlap by at least the duration it takes to initiate and perform a handoff. Translating the amount of overlap measured in time to one measured in geographical distance requires assumptions about the speed with which users move. The models offer a way of relating protocol parameters with assumptions for movement speed and requirements for cell overlap.

Both existing handoff initiation strategies, ECS and LCS, have serious performance lacks, but ECS has the potential to avoid packet loss. But as initial results from the building wide experiment indicate, ECS does not perform well in an actual wireless network because sporadic router advertisements from new but unstable networks barely within reach causes ECS to handoff to the new network. A more intelligent algorithm is needed. Parametric Cell Switching which takes signal quality into account as well as the throughput and price of a link into account, indicates that good performance is possible using Mobile IP. It would be interesting to investigate if the technique can be further improved by additional parameters such as measured bit error rate or media type. One might even consider (base station) topology information that could be broadcasted from access routers or made available upon request.

Further work should include a more complete study than presented here to determine the number of performed handoffs in various cell configurations and movement patterns, possibly performed as a simulation study. Also the behavior in a wide area setting should be investigated more thoroughly.

More work should also be done on how Authentication, Authorization, and Accounting (AAA) can be done such that a limited overhead is added to a handoff. This may involve using techniques like obtaining authorization concurrently with the home agent binding update, using pre-authorization, allowing time limited access while authorization is ongoing, or using smart card technology.

References

[1] Charles E. Perkins, *Mobile IP Design Principles and Practices*, Addison-Wesley, 1998, ISBN 0-13-850505-5.

[2] David B. Johnson and Charles Perkins, *Mobility Support in IPv6*, Internet-Draft, draft-ietf-mobileip-ipv6-13.txt, November 2000. Work in progress.

[3] David B. Johnson and Charles Perkins, *Mobility Support in IPv6*, Internet-Draft, draft-ietf-mobileip-ipv6-16.txt, March 2002.

[4] R. Ramjee, T. La Porta, S. Thuel, K. Varadhan and S.Y. Wang, *HAWAII: A Domain-Based Approach for Supporting Mobility in Wide-area Wireless networks*, 1999.

[5] A.T. Campbell, Gomez, J. Kim, S., Turanyi, Z. Wan, C-Y. and A. Valko, *Design, Implementation and Evaluation of Cellular IP*, IEEE Personal Communications, Special Issue on IP-Based Mobile Telecommunications Networks, June/July 2000.

[6] Michel Mouly, Marie-Bernadette Pautet. *The GSM System for Mobile Communications*.1992, ISBN 2-9507190-0-7.

[7] S. Deering, *ICMP Router Discovery Messages* RFC 1256, September 1991.

[8] Torben Wittrup Andersen and Anders Lildballe, *Seamless Handoff in Mobile IPv6*, Masters Thesis, Department of Computer Science, Aalborg University, Denmark, 2001.

[9] FreeBSD homepage. Available at http://www.freebsd.org.

[10] KAME homepage. Available at http://www.kame.net.

Chapter 18

Handoff Initiation in Mobile IPv6

Torben W. Andersen, Anders Lildballe and Brian Nielsen[!]

Aalborg University
Department of Computer Science
Fredrik Bajersvej 7E
Aalborg Dk-9220, Denmark

Torben.Wittrup@accenture.com

Anders.Lildballe@eu.fkilogistex.com

bnielsen@cs.auc.dk

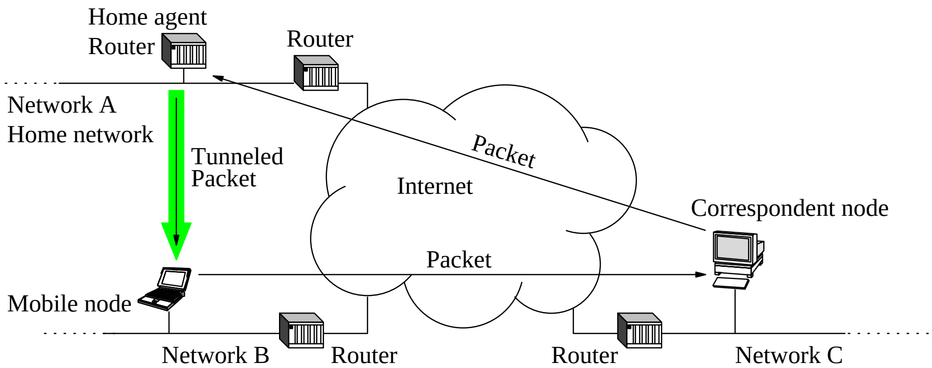

Figure 18.1 Basic Mobile IP Operation

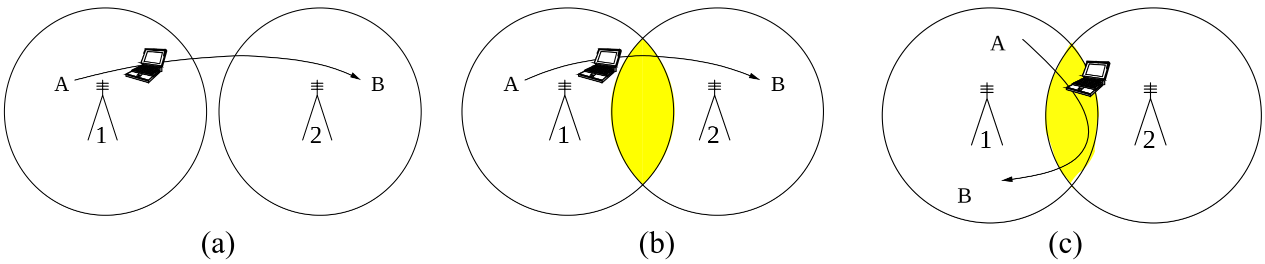

Figure 18.2 A node moves from point A to B. No seamless handoff is possible (a), seamless handoff theoretically possible (b), seamless handoff possible, but should not be performed (c).

[!] Author for contact

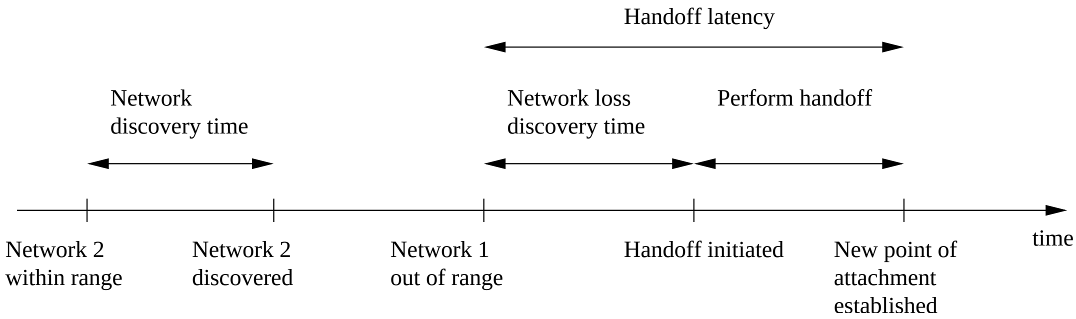

Figure 18.3 Lazy Cell Switching.

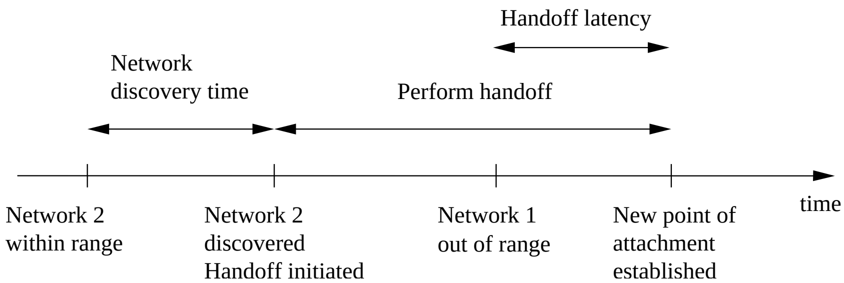

Figure 18.4 Eager Cell Switching.

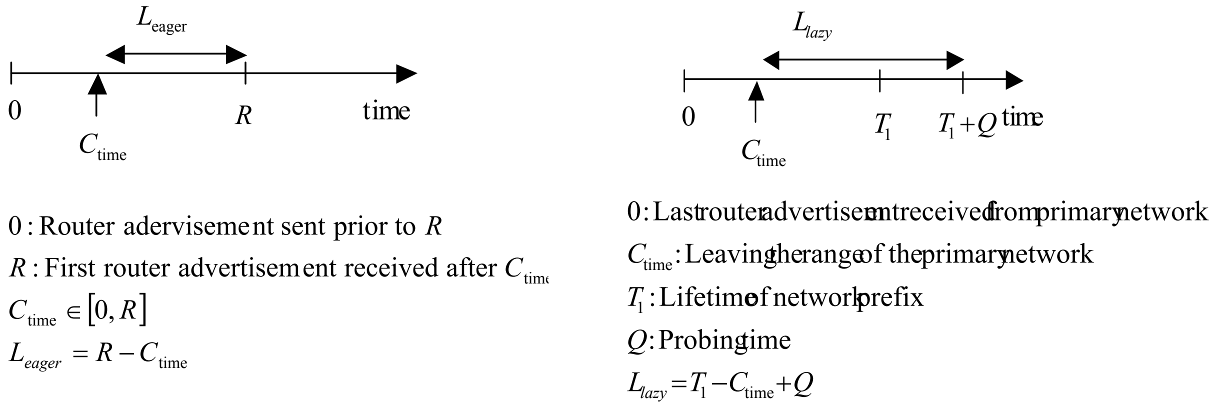

Figure 18.5 Model for computing L_{eager} (a) and model for computing L_{lazy} (b).

.

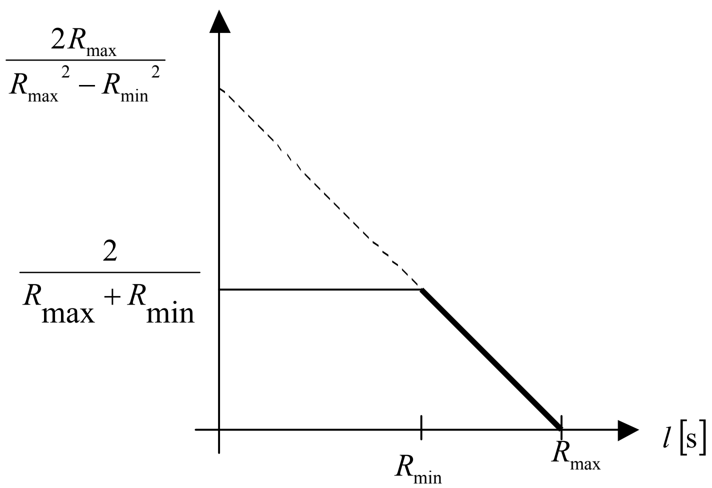

Figure 18.6 Density function $P_{L_{\text{eager}}}(l)$ for ECS

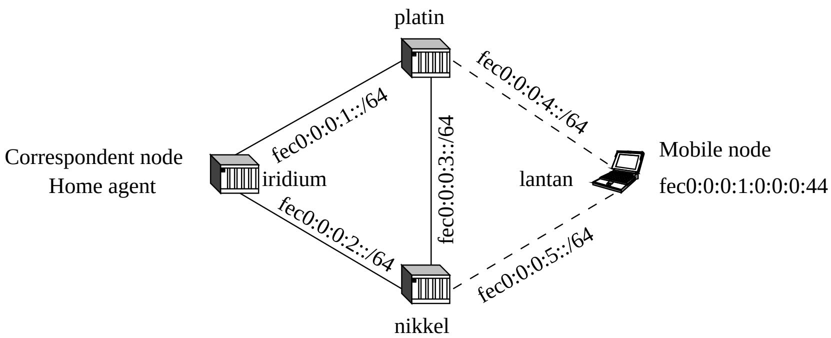

Figure 18.7 Testbed with IPv6 addresses and network prefixes.

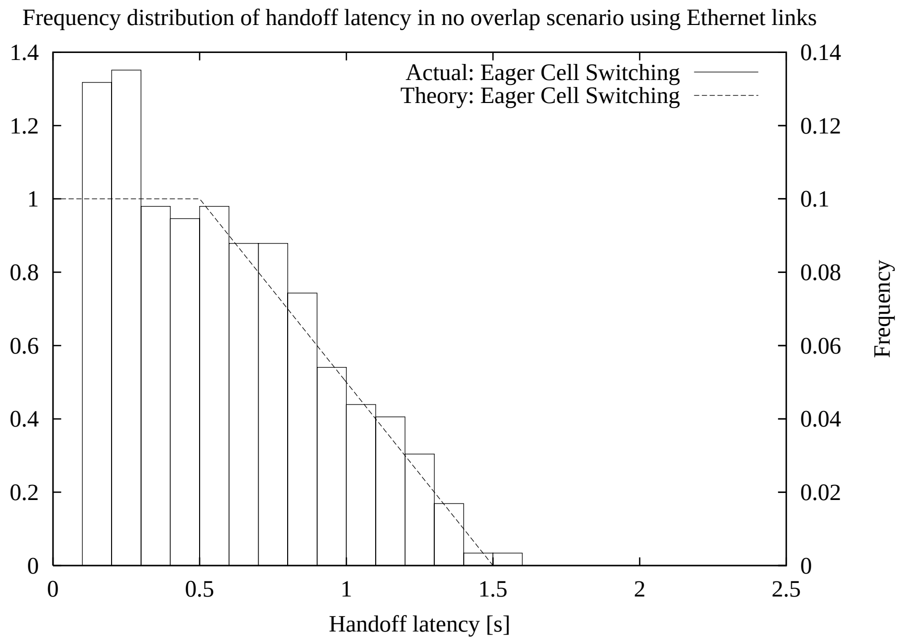

Figure 18.8 Frequency distribution for handoff latency using ECS and default configuration in the no network overlap scenario.

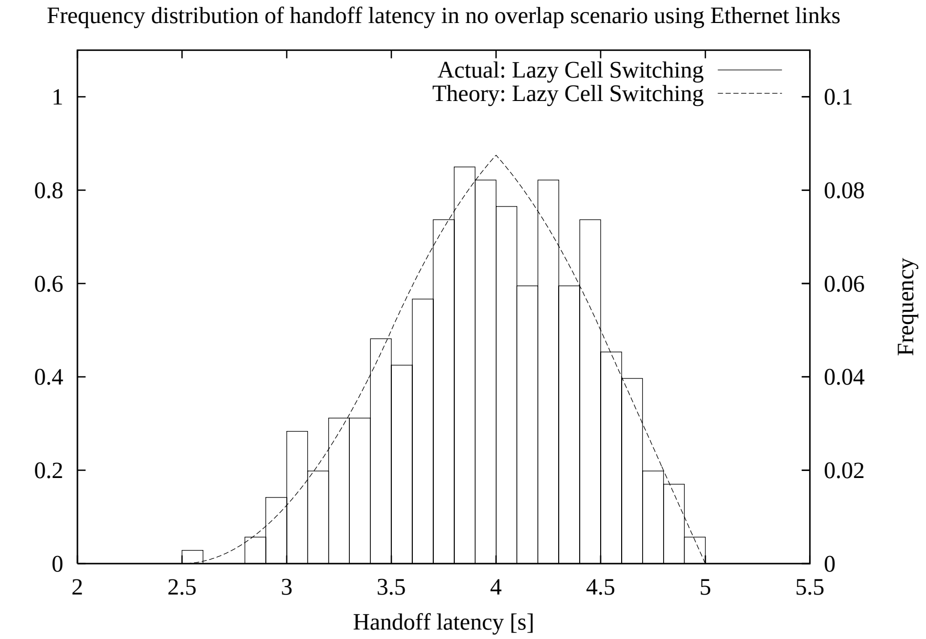

Figure 18.9 Frequency distribution for handoff latency using LCS and default configuration in the no network overlap scenario.

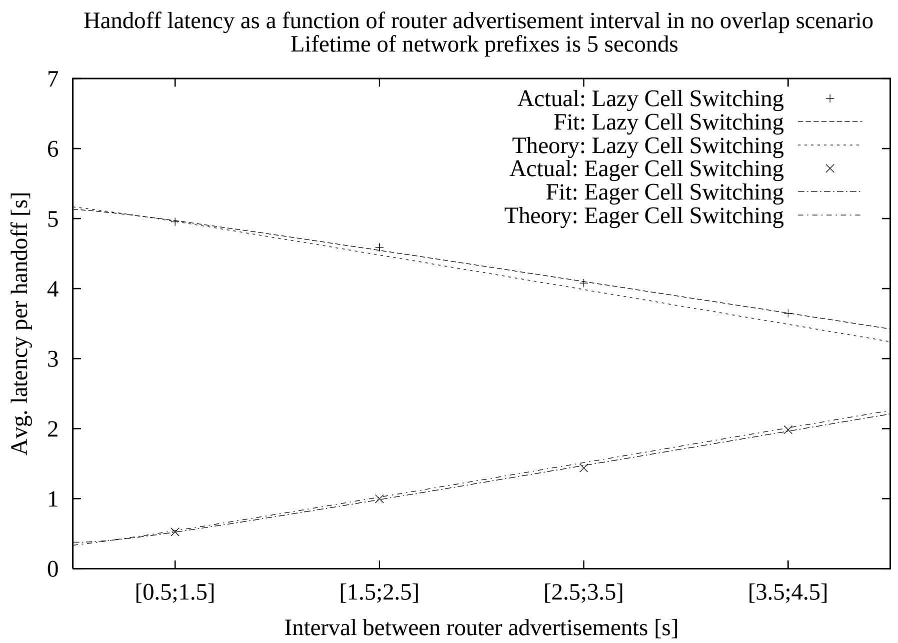

Figure 18.10 The average handoff latency as a function of router advertisement interval for ECS and LCS in the no network overlap scenario.

Figure 18.11 Deployment of access routers within the Department of Computer Science.

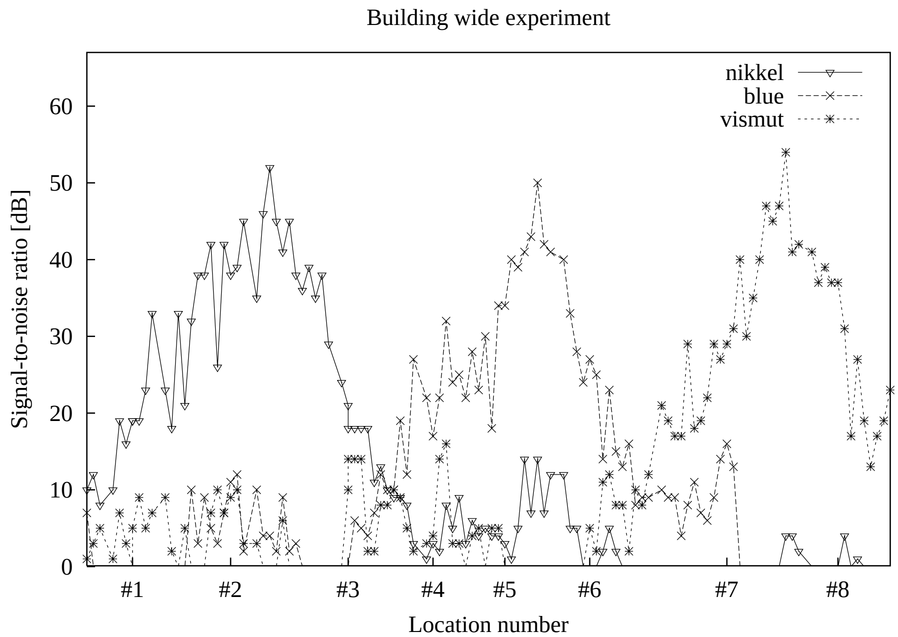

Figure 18.12 Measured signal-to-noise ratios from the three access routers when the mobile node is moved along the path depicted in Figure 18.11

Used Acronyms

IP	Internet Protocol
HTTP	HyperText Transfer Protocol
FTP	File Transfer Protocol
email	Electronic Mail
Mobile IP	Extension of the Internet Protocol to support mobile users
TCP	Transmission Control Protocol
ECS	Eager Cell Switching
LCS	Lazy Cell Switching
PCS	Parametric Cell Switching
IPv6	Internet Protocol version 6
Mobile IPv6	Extension of IPv6 to support mobility
DHCP	Dynamic Host Configuration Protocol
DNS	Domain Name Service
IPv4	Internet Protocol version 4
ICMP	Internet Control Message Protocol
HAWAII	Handoff-Aware Wireless Access Internet Infrastructure
Cellular IP	Internet Protocol extension for cellular networks
GSM	Global System for Mobile Communications (originally Groupe Spéciale Mobile)
LAN	Local Area Network
BSD	Berkeley Software Distribution, the name chosen by the Computer Systems Research Group at University of California, Berkeley, for their Unix distribution
FreeBSD	A free, open source, version of the BSD UNIX Operating System
UDP	User Datagram Protocol
AAA	Authentication, Authorization, and Accounting
UNIX	Derived from UNICS (Uni-plexed information and computing system).
KAME	"Turtle" in Japanese. The KAME Project is a joint effort of seven companies in Japan to provide a free IPv6 and IPsec protocol stack for BSD variants to the world. The project is located in **Ka**rigo**me** village, Fujisawa, Kanagawa JAPAN. (See Also http://orange.kame.net/dev/cvsweb.cgi/kame/FAQ?rev=1.77)

Additional Index Terms

Seamless roaming
seamless handoff
handoff algorithm
Handoff initiation
handoff latency
handoff initiation algorithms
mathematical
Mobile IPv6 testbed
Testbed
Building wide experiment

empirical study
proactive handoff
reactive handoff
home address
foreign network
foreign agent
care-of address
primary network
correspondent node
home agent
triangular routing
handoff performance
binding update
route optimization
router advertisements
stateless auto-configuration
statefull auto-configuration

network discovery delay
network loss discovery
lifetime of network prefixes
probing time
tuning protocol parameters
signal-to-noise ratio
roundtrip time
(proxy) neighbor discovery
home address option
ingress filtering
binding update option
routing header
micro mobility
gateway
network load
hard handoff
semi-soft handoff
cell boundaries
density function
probability distribution
Wireless LAN
802.11b
firewall
sequence number
timestamp
packet loss
Optimizing
Optimization
Performance
access routers
smart card
Authentication

Authorization
Accounting
pre-authorization