Mobile IP
Internet Draft   P. McCann
Document: draft-mccann-mobileip-80211fh-01.txt   Lucent Technologies
Expires: April 2003   October 2002

Mobile IPv6 Fast Handovers for 802.11 Networks

Status of this Memo

Abstract

This document describes how a Mobile IPv6 Fast Handover [2] could be implemented on a link layers conforming to the 802.11 suite of specifications [3].

Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC-2119 [4].

Table of Contents

1. Introduction

The Mobile IPv6 Fast Handover protocol [2] has been proposed as a way to minimize the interruption in service experienced by a Mobile IPv6 node as it changes its point of attachment to the Internet. Without such a mechanism, a mobile node cannot send or receive packets from the time that it disconnects from one point of attachment in one subnet to the time it registers a new care-of address from the new point of attachment in a new subnet. Such an interruption would be unacceptable for real-time services such as Voice-over-IP.

Note that there may be other sources of service interruption that may be "built-in" to the link-layer handoff. For example, a recent study has concluded that the 802.11 beacon scanning function may take several hundred milliseconds to complete [5] during which time sending and receiving IP packets is not possible. This sort of interruption may present an obstacle to real-time service deployment that needs further optimization; however, such optimization is outside the scope of this document.

The basic idea behind a Mobile IPv6 fast handover is to leverage information from the link-layer technology to either predict or rapidly respond to a handover event. This allows IP connectivity to be restored at the new point of attachment sooner than would otherwise be possible. By tunneling data between the old and new access routers, it is possible to provide IP connectivity in advance of actual Mobile IP registration with the home agent or correspondent node. This removes such Mobile IP registration, which may require time-consuming Internet round-trips, from the critical path before real-time service is re-established.

The particular link-layer information available, as well as the timing of its availability (before, during, or after a handover has occurred), differs according to the particular link-layer technology in use. This document gives a set of deployment examples for Mobile IPv6 Fast Handovers on 802.11 networks. We begin with a brief overview of relevant aspects of basic 802.11 [3]. We examine how and when handover information might become available to the IP layers that implement Fast Handover, both in the network infrastructure and on the mobile node. Finally, we give details on how the proposed

Mobile IPv6 Fast Handover protocol would work in this environment and evaluate the feasibility of the different IP-layer fast handover mechanisms available.

2. Terminology

This document borrows all of the terminology from Mobile IPv6 Fast Handovers [2], with the following additional terms from the 802.11 specification [3] (some definitions slightly modified for clarity):

Access Point (AP): Any entity that has station functionality and provides access to the distribution services, via the wireless medium (WM) for associated stations.

Association: The service used to establish access point/station (AP/STA) mapping and enable STA access to the Distribution System.

Basic Service Set (BSS): A set of stations controlled by a single coordination function, where the coordination function may be centralized (e.g., in a single AP) or distributed (e.g., for an ad-hoc network). The BSS can be thought of as the coverage area of a single AP.

Distribution System (DS): A system used to interconnect a set of basic service sets (BSSs) and integrated local area networks (LANs) to create an extended service set (ESS).

Extended Service Set (ESS): A set of one or more interconnected basic service sets (BSSs) and integrated local area networks (LANs) that appears as a single BSS to the logical link control layer at any station associated with one of those BSSs. The ESS can be thought of as the coverage area provided by a collection of APs all interconnected by the Distribution System. It may consist of one or more IP subnets.

Inter-Access Point Protocol (IAPP): A protocol defined for use between access points [6] at handover time that allows for the old association with the old AP to be deleted, and for context to be transferred to the new AP.

Station (STA): Any device that contains an IEEE 802.11 conformant medium access control (MAC) and physical layer (PHY) interface to the wireless medium (WM).

3. Deployment Architectures for Mobile IPv6 on 802.11

In this section we describe the two most likely relationships between Access Points (APs), Access Routers (ARs), and IP subnets that are possible in an 802.11 network deployment. Here we consider only the infrastructure mode [3] of 802.11. A given STA may be associated with one and only one AP at any given point in time; when a STA moves out of the coverage area of a given AP it must handover (re-associate) with a new AP. It is important to understand that 802.11 offers great flexibility, and that handover from one AP to another does not necessarily mean a change of AR or subnet.

```
                  AR                              AR
            AR     |    AR                   AR    |     AR
              \    |   /                       \   |    /
               Subnet 1                         Subnet 2
             /  /  |  \  \                    /  /  |  \  \
            /  /   |   \  \                  /  /   |   \  \
           /   |   |   |   \                /   |   |   |   \
        AP1  AP2  AP3  AP4  AP5          AP6  AP7  AP8  AP9  AP10
```

Figure 1: An 802.11 deployment with relay APs.

Figure 1 depicts a typical 802.11 deployment with two IP subnets, each with three Access Routers and five Access Points. Note that the APs in this figure are acting as link-layer relays, which means that they transport Ethernet-layer frames between the wireless medium and the subnet. Each subnet is implemented as a single LAN or VLAN. Note that a handover from AP1 to AP2 does not require a change of AR because all three ARs are link-layer reachable from a STA connected to any AP1-5. Therefore, such handoffs are outside the scope of IP-layer handover mechanisms. However, a handoff from AP5 to AP6 would require a change of AR, because the STA would be attaching to a different subnet. An IP-layer handover mechanism would need to be invoked in order to provide low-interruption handover between the two ARs.

```
                                Internet
                               /    |   \
                              /     |    \
                             /      |     \
                           AR      AR      AR
                           AP1     AP2     AP3
```

Figure 2. An 802.11 deployment with integrated APs/ARs.

Figure 2 depicts an alternative 802.11 deployment where each AP is integrated with exactly one AR. In this case, every change of AP would result in a necessary change of AR, which would require some IP-layer handover mechanism to provide for low-interruption handover between the ARs. Also, the AR shares a MAC-layer identifier with its attached AP.

In the next section, we examine the steps involved in any 802.11 handover. Subsequent sections discuss how these steps could be integrated with an IP-layer handover mechanism in each of the above deployment scenarios.

4. 802.11 Handovers in Detail

An 802.11 handover takes place when a STA changes its association from one AP to another ("re-association"). This process consists of the following steps:

1. The STA performs a scan to see what APs are available. The result of the scan is a list of APs together with physical layer information, such as signal strength.
2. The STA chooses one of the APs and performs a join to synchronize its physical and MAC layer timing parameters with the selected AP.
3. The STA requests authentication with the new AP. For an "Open System", such authentication is a single round-trip message exchange with null authentication.
4. The STA requests association or re-association with the new AP. A re-association request contains the MAC-layer address of the old AP, while a plain association request does not.
5. If operating in accordance with the IAPP [6], the new AP performs a lookup based on MAC-layer address to obtain the IP address of the old AP by consulting a local table or RADIUS server. It opens a UDP or TCP connection, protected by IPSec encryption, to the old AP. Via the secure connection, it informs the old AP of the re-association so that information about the STA is deleted from the old AP. Note that IAPP can only be invoked based on a re-association message, as the plain association message does not contain the old AP's MAC-layer address.
6. The new AP sends a Layer 2 Update frame on the local LAN segment to update the learning tables of any connected Ethernet bridges.

Note that in most existing 802.11 implementations, steps 1-4 are performed by firmware that is on-board the 802.11 PCMCIA card. This might make it impossible for the host to take any actions (including sending or receiving IP packets) before the handoff is complete.

During step 5, IAPP is used to communicate with the old AP. The IPSec tunnel between the two APs is originally established with key distribution via RADIUS, but can be subsequently re-used for different MNs that may need to handover between the same pair of APs. Note that the SA is between the pair of APs and has nothing to do with any security association that might be in place between the MN and either of the APs. During IAPP operation, link-layer context may be transferred from the old AP to the new AP. The IAPP defines a container for context information. However, no such context has currently been defined or standardized by IEEE.

Also note that there is no guarantee that an AP found during step 1 will be available during step 2 because radio conditions can change dramatically from moment to moment. The STA may then decide to associate with a completely different AP. Usually, this decision is implemented in firmware and the attached host would have no control over which AP is chosen.

There is no standardized trigger for step 1. It may be performed as a result of decaying radio conditions on the current AP or at other times as determined by local implementation decisions. Usually this step will be performed autonomously by firmware in the NIC using proprietary scanning algorithms.

The coverage area of a single AP is known as a Basic Service Set (BSS). Note that both APs in the above description are considered to belong to the same Extended Service Set (ESS). This is to trigger a re-association (rather than plain association) from the STA, which contains information about both the old and new AP. All APs should therefore broadcast the same ESSID. If two APs belong to different administrative domains, this may require some inter-domain coordination of the ESSID. Otherwise, there may not be sufficient information to construct the link-layer triggers required by some handover mechanisms.

A change of BSS within an ESS may or may not require an IP-layer handover, depending on whether the APs provide STAs access to different or the same IP subnets. The next two sections detail how each mechanism from the Mobile IPv6 Fast Handover specification might accomplish the necessary IP-layer reconfiguration. First we consider Anticipated handover and then move on to Tunnel-based handover.

5. Anticipated Handover

Because all 802.11 handovers are mobile initiated, the network-initiated Anticipated Handover is not applicable to 802.11.

In mobile initiated Anticipated Handover, the MN first sends a Router Solicitation for Proxy (RtSolPr) to the oAR containing the link-layer address of the new Access Point. This would happen between steps 1 and 2 from Section 4. Note that for this to be possible, the MLME-SCAN.request primitive (See Section 10.3.2.1 of the 802.11 specification [3]) must be available to the host, and the card firmware must not make autonomous handover decisions. The oAR maps the new AP's link-layer address into the IP address of the nAR that should be used by the MN on the new link. Note that this requires a mapping table to be maintained at oAR, either by manual configuration or with the use of unspecified discovery protocols. Then, the oAR determines whether stateful or stateless addressing is used by nAR. For stateless addresses, the oAR picks an nCoA on the new subnet (using the MN's interface identifier) and proposes it to the nAR using HI/HACK. For stateful addresses, the oAR must request an address from nAR with the HI/HACK exchange. The oAR returns a Proxy Router Advertisement (PrRtAdv) to the MN. This PrRtAdv may be sent in parallel with HI/HACK, in the case of stateless address configuration, but must be serialized after HI/HACK in the case of stateful address configuration. The MN then sends a Fast Binding Update (F-BU) to the oAR with a binding to the new care-of address (nCoA).

At this point the MN should move to nAR (steps 2-6 from Section 4). Note that here we assume the host can send IP layer messages such as F-BU prior to step 2, which implies that the interface firmware did not autonomously skip to step 2 without permission from the host. Once re-associated with the new AP, the MN will hopefully receive the F-BACK indicating that the oAR received its F-BU and also that the nCoA is valid. This message is sent on both the old and new links because the MN is in transit between them. Packets from the oAR will be forwarded to nAR based on the F-BU. If it doesn't receive the F-BACK right away, the MN retransmits the F-BU and indicates its presence to the nAR with a Fast Neighbor Advertisement (F-NA). The nAR should return a Router Advertisement containing a Neighbor Advertisement Acknowledgement (NAACK) indicating whether the nCoA is valid. If not, the MN can continue to use oCoA as a source address for packets while it obtains a valid nCoA. Either the F-BACK or the Router Advertisement informs the MN which link-layer address to use as its default router on subsequent outbound packets.

Note that Anticipated Handover requires that the MN send a RtSolPr and receive a PrRtAdv prior to executing the layer-two handover.

Otherwise, the MN will not have any information about the new subnet, and will need to begin neighbor discovery and care-of address configuration from scratch once it has completed the layer-two handover. There is no guarantee that the RtSolPr/PrRtAdv exchange will complete especially in a radio environment where the connection to the old AP is deteriorating rapidly. Also, there is no guarantee that the MN will actually attach to the given new AP after it has sent the F-BU to the oAR, because changing radio conditions may cause nAR to be suddenly unreachable. The precise impact of these factors in an Anticipated Handover can only be evaluated after experimentation in a particular deployment.

6. Tunnel-based Handover

In a Tunnel-based Handover, the oAR and nAR collaborate to establish a bi-directional edge tunnel (BET) in reaction to a layer-2 handover event. In an 802.11 network, this event would be step 4 from Section 4 (target trigger) or perhaps step 5 at the old AP (source trigger). If the network looks like Figure 2, where the APs are integrated with the ARs, then the L2-TT (or L2-ST) is available at nAR (or oAR) through some internal interface. However, if the network is deployed like Figure 1, then some network message will need to be sent from the new AP (or old AP) to nAR (or oAR). This message might be the object of future standardization efforts. Note also that there may be several ARs present on the new subnet, and the new AP must choose one to which to deliver the trigger, which becomes nAR. The Layer 2 Update frame sent by the new AP might be of some assistance in constructing L2-TT; however, this message is broadcast to all ARs on the new subnet and does not indicate which one is to be chosen as the endpoint of the tunnel. Also, it does not contain the MAC address of the old AP that would enable discovery of oAR.

The AR that received the trigger sends a HI message to the other AR, who in turn responds with a HACK. Note that this requires a mapping table to be maintained, similar to the one for Anticipated handover, which yields the IP address of an AR given the link-layer address of an AP. This table must be maintained manually or with the aid of some unspecified discovery protocol. The re-association provides L2-LD and L2-LU triggers to oAR and nAR, respectively. At this point the BET is established and traffic is tunneled between the two ARs so that the MN continues to receive service, using oCoA, without the need to exchange any messages immediately before, during, or immediately after the handoff. At some future time, the MN may obtain an nCoA and register from the new network, perhaps using completely standard Mobile IPv6 mechanisms to do movement detection and registration.

Note that the MN must somehow obtain the link-layer address of nAR before service can resume, so that it has a link-layer destination address for outgoing packets (default router information). In the deployment illustrated in Figure 2, this would be exactly the AP's MAC layer address, which can be learned from the link-layer handoff messages. However, in the case of Figure 1, this information must be learned through other means currently unspecified. Also note that even in the case of Figure 2, the MN must somehow be made aware that it is in fact operating in a Figure 2 network and not a Figure 1 network. One option might be the Candidate Access Router (CAR) discovery protocol [7] currently being worked in the Seamoby working group. Interestingly, this information is also available from the PrRtAdv message, although its use is currently prohibited in tunnel-based handover. A MN could conceivably obtain advertisements from all neighboring APs well in advance of the handover, even if it intended to use a Tunnel-Based instead of Anticipated handover.

Note that the BET is established at the behest of layer-2 messages. Because this is a redirection of the MN's traffic, care must be taken to ensure that the layer-2 messages are secure. This issue is discussed in more detail in Section 7.

For now we do not discuss the Handoff to Third (HTT) mechanism of a Tunnel-based handover. Its configuration and security implications are similar to the basic scheme.

7. Security Considerations

As stated in the Mobile IPv6 fast handover specification, the security considerations of Anticipated Handover are very similar to those required of any Mobile IPv6 Binding Update message. The oAR and MN are assumed to have a security association for the Binding Updates, which also provides authentic PrRtAdv messages to the MN. However, creating such a security association for a roaming MN is still an open problem. Also, security must be established between all possible (oAR, nAR) pairs so that PrRtAdv/HI/HACK messages may be authenticated. This might be achieved through manual configuration or automatic discovery, using whatever means was used to set up the mapping table discussed in Section 5.

Similar to Anticipated handover, Tunnel-based handover also requires a secure means to establish neighbor-mapping tables, so that tunnels can be established securely between oAR and nAR based on the L2 triggers. In addition, the security of a Tunnel-based handover depends on the link-layer security in place. This is because a BET is established and MN traffic is redirected purely in reaction to link-layer handoff messages. Note that step 3 from Section 4 could potentially provide some security; however, due to the identified weaknesses in WEP shared key security [8], there is currently no

authentication algorithm for step 3 that is both standardized and secure.

It may be the case that many deployments are configured as "Open Systems", which will rely instead on higher-layer authentication such as 802.1X Port-Based Network Access Control [9], or, ultimately, the future output of the PANA working group [10]. According to published standards, such authentication techniques would happen only after association or re-association takes place, which leaves the re-association messages unprotected. This would allow malicious nodes to redirect traffic to a different subnet in a Tunnel-based handover environment, or to a different AP on the same subnet even in an Anticipated handover environment. Work is currently underway to better integrate 802.1X with 802.11 [11] but it is not yet complete.

The 802.1X standard [9] defines a way to encapsulate EAP on 802 networks (EAPOL, for "EAP over LANs"). With this method, the client and AP participate in an EAP exchange which itself can encapsulate any of the various EAP authentication methods. The EAPOL exchange can output a master key, which can then be used to derive transient keys, which in turn can be used to encipher/authenticate subsequent traffic. It is possible to use 802.1X pre-authentication [11] between a STA and a target AP while the STA is associated with another AP; this would enable authentication to be done in advance of handover, which would both protect the re-association message and allow fast resumption of service after roaming. However, because EAPOL frames carry only MAC-layer instead of IP-layer addresses, this is currently only specified to work within a single subnet, where IP layer handoff mechanisms are not needed anyway. In our case (roaming across subnet boundaries) the 802.1X exchange would need to be performed after roaming to, but prior to re-association with, the new AP. This would introduce additional handover delay while the 802.1X exchange takes place, which may also involve round-trips to RADIUS or Diameter servers.

Perhaps faster cross-subnet authentication could be achieved by leveraging the context transfer features of the IAPP to carry security credentials [12], or with the use of pre-authentication using PANA. To our knowledge this sort of work is not currently underway in the IEEE. The security considerations of these new approaches would need to be carefully studied.

8. Conclusions

The Mobile IPv6 Fast Handoff specification presents two alternative protocols for shortening the period of service interruption during a change in link-layer point of attachment. This document has

attempted to show how each may be applied in the context of 802.11 access networks.

There are currently serious security problems in the published specifications that define the 802.11 handover process that must be fixed before even intra-subnet mobility can be considered secure. Tunnel-based handovers would depend on these mechanisms to secure cross-subnet mobility. In-progress specifications may fix these problems but may also introduce additional delay for handover across different subnets. Usually, only the APs themselves are aware that good link-layer security is in place. This information could be made available to ARs with the use of a new protocol (e.g., [13]), but as such mechanisms are prone to be link-layer specific, we recommend that work on Tunnel-based handovers be progressed in the IEEE rather than the IETF.

Anticipated handover places requirements that messages be exchanged over the wireless link prior to handover, during a period that is normally under the control of low-level firmware. The performance impact of this requirement, and of the failure to meet it in certain radio conditions, must be critically evaluated with experimental data. Also, given a particular firmware implementation of handover, it may be impossible for a host to send the required IP-layer messages at the proper time.

Both schemes rely on unspecified mechanisms for mapping AP L2 addresses into AR IP addresses (Anticipated and Tunnel-based) or AR L2 addresses (Tunnel-based). This problem is arguably more severe with Tunnel-based handovers, especially on networks like Figure 1, because the MN itself does the unspecified mapping and it cannot be handled by manual configuration. In Anticipated handover, the oAR must be configured with this information so that it can send the proper PrRtAdv to the MN.

The relationship between the PrRtAdv and Candidate Access Router discovery protocols needs further study. Some similar functionality seems to be provided by each and it may not be necessary to standardize both mechanisms independently.

For these reasons, we recommend that the draft be refocused to concentrate on the specification and security considerations for the F-BU and F-BACK messages only. This allows for updating the oAR with the current MN location under any circumstance, whether the handover is anticipated or not. The other mechanisms outlined in the draft either need more support from the link layer, and should be moved into the IEEE, or require further study to determine their relationship with other work in the IETF.

References

1 Bradner, S., "The Internet Standards Process -- Revision 3", BCP 9, RFC 2026, October 1996.

2 Dommety, G. (editor), Yegin, A., Perkins, C., Tsirtsis, G., El-Malki, K., and Khalil, K., "Fast Handovers for Mobile IPv6", draft-ietf-mobileip-fast-mipv6-04.txt, March 2002. Work In Progress.

3 "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications", ANSI/IEEE Std 802.11, 1999 Edition.

4 Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

5 Mitra, A., Shin, M., and Arbaugh, W., "An Empirical Analysis of the IEEE 802.11 MAC Layer Handoff Process", CS-TR-4395, University of Maryland Department of Computer Science, September 2002.

6 "Recommended Practice for Multi-Vendor Access Point Interoperability via an Inter-Access Point Protocol Across Distribution Systems Supporting IEEE 802.11 Operation", IEEE Std 802.11f/D4, July 2002. Work In Progress.

7 Krishnamurthi, G. (editor), "Requirements for CAR Discovery Protocols", draft-ietf-seamoby-card-requirements-01.txt, August, 2002. Work In Progress.

8 Borisov, N., Goldberg, I., and Wagner, D., "Intercepting Mobile Communications: The Insecurity of 802.11", Proceedings of the Seventh Annual International Conference on Mobile Computing and Networking, July 2001, pp. 180-188.

9 "Port-Based Network Access Control", IEEE Std 802.1X-2001, October, 2001.

10 Penno, R. (editor), Yegin, A., Ohba, Y., Tsirtsis, G, and Wang, C., "Protocol for Carrying Authentication for Network Access (PANA) Requriements and Terminology", draft-ietf-pana-requirements-02.txt, June 2002. Work In Progress.

11 "Draft Supplement to IEEE 802.11: Specification for Enhanced Security", IEEE Std 802.11i/D2.2, July 2002. Work In Progress.

12 Aboba, B., and Moore, T., "A Model for Context Transfer in IEEE 802", draft-aboba-802-context-02.txt, April, 2002. Work In Progress.

13 Yegin, A., "Link-layer Triggers Protocol", draft-yegin-l2-triggers-00.txt, June 2002. Work In Progress.

Acknowledgments

Thanks to Bob O'Hara for providing explanation and insight on the 802.11 standards. Thanks to James Kempf and Erik Anderlind for providing comments on an earlier draft.

Author's Address

Pete McCann
Lucent Technologies
Rm 9C-226R
1960 Lucent Lane
Naperville, IL 60563
Phone: +1 630 713 9359
Fax: +1 630 713 1921
Email: mccap@lucent.com

Intellectual Property Statement

At the time of submission the author is not aware of any intellectual property rights that pertain to the implementation or use of the technology described in this document. However, this does not preclude the possibility that Lucent Technologies, Inc. or other entities may have such rights. The patent licensing policy of Lucent Technologies, Inc. is on file with the IETF Secretariat.

The IETF takes no position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; neither does it represent that it has made any effort to identify any such rights. Information on the IETF's procedures with respect to rights in standards-track and standards-related documentation can be found in BCP-11. Copies of claims of rights made available for publication and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such