Mobile IP Working Group
INTERNET-DRAFT
Expires: April 2001
Karim El-Malki, Ericsson
Hesham Soliman, Ericsson
November, 2000

Fast Handoffs in MIPv6

<draft-elmalki-handoffsv6-01.txt>

Status of this memo

This document is an individual submission to the IETF. Comments should be directed to the authors.

Abstract

This draft describes a method to achieve Fast Handoffs in Mobile IPv6. Fast Handoffs are required in Mobile IPv6 in order to limit the period of service disruption experienced by a wireless Mobile Node when moving between access routers. This requirement becomes even more important when supporting real-time services. Fast Handoffs involve anticipating the movement of MNs and sending multiple copies of the traffic to potential Mobile Node movement locations. Both flat and Hierarchical Mobile IPv6 models are considered. The Hierarchical MIPv6 mobility Management model in [1] already offers improvements to Mobile IP handoffs by providing a local Mobility Anchor Point (MAP) functionality. Some additions are made to the operation of this existing Hierarchical model to achieve Fast Handoffs.

TABLE OF CONTENTS

1. Introduction..2

2. Fast Handoffs.......................................4
2.1 Initiating Fast Handoffs through the "previous" AR...5

3. Fast Handoffs in Hierarchical MIPv6.................7
3.1 Fast Handoffs in a flat MIPv6 architecture...........7
4. Handling Ping Pong in Fast Handoffs.................8
5. Extensions for Fast Handoffs........................9
5.1 Extensions to MIPv6...............................9
5.2 Extensions to Neighbour Discovery...................10
6. Fast Handoffs and DAD..............................10
7. Acknowledgements......................................11

8. References..11

9. Addresses...12

1. Introduction

Fast Handoffs anticipate the movement of wireless Mobile Nodes (MNs) by utilizing simultaneous bindings in order to send multiple copies of the traffic to potential Mobile Node movement locations. In this way, Fast Handoffs coupled to layer 2 mobility can help in achieving seamless handoffs between Access Routers (ARs) by eliminating the delay period required to perform a Registration following a Mobile IP handoff.
An alternative method to perform improved handoffs, namely Smooth Handoffs, is described in [2]. The method for Fast Handoff addresses the need to support services having strict delay bounds (i.e. real-time) which in certain cases may be hard to support if traffic has to be forwarded between ARs using Smooth Handoffs. Also, in the non-realtime case it may be possible that the new AR receives buffered traffic from the previous AR (smooth handoff)and traffic from the CN contemporarily which could cause some out-of-order and delayed packets to be delivered to the MN. In some cases this may affect the performance of higher level protocols (i.e. TCP). This same situation will not arise using Fast Handoffs.

This draft considers both the normal Mobile IPv6 model [2] and the hierarchical Mobile IPv6 model [1]. These are shown in Figure 1 where the Access Points (APs) or Radio Access Networks (RANs) are used to provide a MN with wireless L2 access.

Simultaneous bindings are described in this draft and may be achieved by setting a new, "B" flag in the BU sent by the MN to a

MAP. In this way, the MAP will add a new binding for the MN without removing the existing entry.

Hence packets arriving to the MAP will be tunnelled to both addresses in its Binding Cache.

```
             _________          _________
            |         |        |         |
            |   HA    |--------|  (MAP)  |________
            |_________|        |_________|        \
                                 /  |  \           \
                                                    \
                              ...  ...  ...          \
                                                      \
                          ______/_     _\______        |
                         |        |   |        |       |
                         | AR2/MAP|   | AR1/MAP|       |
                         |________|   |________|       |
                          ____|___     ____|___    ____|___
                         |        |   |        |  |        |
                         |AP/RAN 2|   |AP/RAN 1|  |AP/RAN 3|
                         |________|   |________|  |________|
                              |        ____|___
                                      |        |
                             CN       |   MN   |
                                      |________|
```

Figure 1: Flat (HA only) and Hierarchical (HA and MAP) MIPv6 model

The method to anticipate MN movement by interacting with the wireless L2 is described later in this draft.

The Hierarchical Mobile IPv6 scheme introduced in [1] allows a Mobile Node to perform registrations locally with a MAP in order to reduce the number of signalling messages to the home network and CNs. This achieves a reduction in the signalling delay when a Mobile Node moves between ARs and therefore improves the performance of such handoffs. This draft describes Fast Handoffs in Hierarchical Mobile IPv6 (HMIPv6) as well as a flat network architecture.
When considering a MIPv6 handoff, two different cases can be considered depending on the network architecture:

- The previous and new AR are physically connected
- The previous and new AR are connected via N other nodes or networks

The first case can be considered a subset of the more generic case (second). Hence the solution proposed will be addressing the generic (second) case.

2. Fast Handoffs overview

Fast Handoffs address the need to achieve seamless Mobile IP Handoffs when the MN moves between ARs. This is done by "bicasting" traffic to the "previous" AR and "new" AR while the MN is moving between them. The anticipation of the MN's movement is achieved by tight coupling with Layer 2 functionality which is dependent on the type of access technology used. The coupling between L2 and L3 technologies may occur in the network nodes or the MNs, or both, depending on the access technology. "Bicasting" is achieved through simultaneous bindings, where the MN activates the "B" flag in the MAP registration. When a MAP Registration has the "B" flag set, the receiving MAP, which has an existing binding for the MN, will add the relevant new binding for the MN but will also maintain any existing binding it had for the MN.

Two different handoff scenarios are considered in this draft:

- A MN having to do a handoff between two different ARs with which it can be simultaneously data-connected (eg. Two different access technologies). In this case it may not be essential to request simultaneous bindings. The MN may simply continue using both COAs (on the old and new link) as specified in [2]. A MN can ensure that packets will be arriving at the _new_ interface after receiving a BAck from its HA/MAP/CNs. This can be achieved by requesting a BAck from these nodes. However, since the processing of BUs is not mandated in [1], the MN may need to keep the old interface live for a short period of time to ensure no more packets are addressed to it.

- A MN having to do a handoff between two access routers with which it can not be simultaneously data-connected. This is the more common case where_Fast_Handoffs can be used to achieve seamless mobility.

When the MN has multiple active bindings with a MAP, it may or may not receive multiple copies of the same traffic directed to it. The use of simultaneous bindings does not necessarily mean that the MN is receiving packets contemporarily from multiple sources. This depends on the characteristics of the access (L2) technology. The "bicasting" of packets, combined with the anticipation of the new COA is used to and speed up handoffs by sending a copy of the data to the AR which the MN is moving to. Until the MN actually completes the L2 handoff to the new AR, the data "copy" reaching this AR may be discarded. In this way the total handoff delay is limited to the time needed to perform the L2 handoff. Thus, Fast Handoffs coupled to the L2 access potentially result in loss-less IP-layer mobility. As described in chapter 2.1, depending on the L2 characteristics, it is also possible for an MN to initiate a Fast Handoff through the "previous" AR without having direct access to the "new" AR.

2.1 Initiating Fast Handoffs through the "previous" AR

In the case in which the wireless L2 technology allows the MN to be data-connected to multiple wireless access points simultaneously, the MN may solicit advertisements from ARs before completing a handoff. In this case "bicasting" may not be necessary.

Some existing wireless L2 technologies and their implementations do not allow a MN to be data-connected to multiple wireless access points simultaneously. Thus, in order to perform a Fast Handoff it is necessary for some form of interworking between layers 2 and 3. It should be noted that the method by which an AR determines when a MN has initiated a L2 handoff is outside the scope of this draft and may involve interaction with L2 messaging. Also, the interaction between L2 and L3 should allow the Mobile Node to perform a L2 handoff only after having performed the L3 Fast Handoff described in this draft. That is, the L2 handoff may be performed after the MN's Registration with the "new" AR which produces a simultaneous binding at the MAP. This Registration may be transmitted more than once to reduce the probability that it is lost due to errors on the wireless link. Alternatively, the MN may choose to send a BU to the MAP with the_A flag set.

A Fast Handoff in this case requires the MN to receive "new" router advertisements through the "old" wireless access points, and to perform a registration with the "new" AR through the "old" wireless access point. Two ways of performing this follow.

I. Inter-AR Solicitation

This solution assumes that the AR with which the MN is currently registered is aware of the IP address of the "new" AR which the MN is moving to. The method by which the current AR is informed of this may depend on interaction with L2 and is outside the scope of this draft.

In some wireless networks, an AR may not be closely coupled to the radio link layer protocols. In these scenarios the initiation of the handoff my need to be done by the MN.
Based on L2 indications, the MN may solicit the router for a special advertisement that includes the "new" subnet prefix(es). To indicate the need for such special advertisement, the solicitation message would need to include a new option showing an identifier for the "new" AP. This identifier may then be mapped in the AR to a neighbouring or the current subnet. Hence the appropriate information can be communicated to the MN.

Once the current AR is aware of the address of the AR which the MN will move to, it will solicit the "new" AR for a router advertisemnt. The "new" AR will reply to the current AR by sending it a router advertisement with appropriate extensions. The current AR will then send the anticipated prefixes for the "new" subnet to the MN's unicast address. As a consequence, the MN, being eager to perform new bindings, will send a BU to the MAP to request that packets addressed to it be bicasted to both the "old" and "new" addresses.

It should be noted that problems may arise if the current AR and the "new" AR are not sharing the same link. In this scenario, sending a router solicitation to the "new" AR would mean that the solicitation and advertisement messages need to be routed beyond the subnet scope.

This scenario is not allowed in [5] to avoid ARP-like attacks in IPv4. However, it may be possible for the old AR to send a solicitation message to the new AR if the following conditions are met:

- The source and destination addresses in the message are of site-local or global scope.
- The AR MUST NOT add its link layer address in the solicitation message.
- The solicitation message and the router advertisement both contain sufficient authentication and authorisation information for both routers. This can be achieved by having one common secret shared between a cluster of ARs within a domain. Such secret can be used to process AH which can be added to the solicitation/advertisement messages.

Added security can be established by setting up a permanent tunnell between the two ARs.

If all the above conditions are met it should be possible for router solicitations/advertisemnts to be routed beyond the subnet scope.

II. Piggy-backing Advertisements on L2 messaging

Let us take Figure 1 as an example, where a MN initiates an L2 handoff from AP/RAN1 to AP/RAN2 (Note that it may not be the MN which takes decisions on handoffs). It is assumed that when an L2 handoff is initiated, AP/RAN1 and AP/RAN2 perform L2 messaging procedures to negotiate the L2 handoff. Since the MN is not attached to AP/RAN2 yet, AR2 is unaware of the IP address of the MN and cannot send an advertisement to it. Therefore it is necessary for the L2 procedures to interwork with Mobile IP.

Once a L2 handoff is initiated, such that AP/RAN2 and AP/RAN1 are in communication, it is possible for AP/RAN2 to solicit an advertisement from AR2 and transfer it to AP/RAN1. Once this is received by the MN, the MN can send a BU to the MAP using an AR2 COA even though the MN has no data-connection to AP/RAN2 yet.

The precise definition of such L2 procedures is outside the scope of Mobile IP.

3. Fast Handoffs in HMIPv6

HMIPv6 is described in [1]. Fast IP Handoffs can be achieved in a very simple and efficient manner.
When the MN receives a router Advertisement including a MAP option, as specified in [1], it should perform actions according to the following movement detection mechanisms. In a Hierarchical Mobile IP network such as the one described in this draft, the MN MUST be:

- "Eager" to perform new bindings
- "Lazy" in releasing existing bindings

The above means that the MN will perform bindings with any "new" MAP advertised by the AR (Eager).
The method by which the MN determines whether the MAP is a "new" MAP is described in [1]. However the MN should not release existing bindings until it no longer receives its MAP option or the lifetime of its existing binding expires (Lazy).

If the MN has at least one existing binding with a MAP, additional simultaneous regional registrations will be performed requesting a short lifetime. This is done in order to limit the lifetime of bindings which the MN only needs temporarily and therefore limit bandwidth usage. This is the case when the MN is moving between ARs and uses Fast Handoffs to achieve near loss-less IP mobility. The lifetime of additional "auxiliary" bindings needed for Fast Handoffs is thus limited.

It should be noted that the method described above is applicable to hierarchical and flat architectures. As described in [1], a MAP can exist on any level in the hierarchy, including ARs. Hence, a bicasting request can also be sent to a MAP located in the AR, in the case where no MAPs are located higher in the hierarchy.

3.1 Fast Handoffs in a flat MIPv6 architecture

A flat MIPv6 architecture is one that does not use a static mobility management based hierarchy. In the context of [1], this would mean that no MAP functionality is deployed beyond the on-link AR. In this case the same concepts described above would still apply if the MAP function is used in the AR.

```
             _________          _________
            |         |        |         |
            |   HA    |--------|Internet |________
            |_________|        |_________|        \
                                 /  |  \           \
                                                    \
                              ...  ...  ...          \
                                                      \
                          ______/_     _\______        |
                         |        |   |        |       |
                         |  MAP1  |   |  MAP2  |       |
                         |________|   |________|       |
                          ____|___     ____|___    ____|___
                         |        |   |        |  |        |
                         |AP/RAN 2|   |AP/RAN 1|  |AP/RAN 3|
                         |________|   |________|  |________|
                              |        ____|___
                                      |        |
                             CN       |   MN   |
                                      |________|
```

Figure 1: Flat mobility management architecture

In this scenario a dynamic hierarchy can be established without changing any of the concepts mentioned above. The handoff is anticipated by sending the "new" link's prefixes to the MN before L2 handoff takes place. The MN would then send a new BU to its "old" MAP, located in the "old" AR, and request that all packets addressed to it be bicasted to its current address and the new COA.
In this manner the "old" MAP/AR serves as an anchor point for the MN while it is connected to the new AR/MAP.

This architecture is well suited to an AR serving large coverage areas where IP mobility may not occur very frequently. For other mobility scenarios where handoffs are frequent, a hierarchical mobility management scheme as [1] is more efficient and flexible.

4. Handling ping pong in Fast Handoffs

Ping Pong is a term used mainly in cellular networks to describe the repetetive rapid movement of a mobile terminal between two APs for a short period of time. While this phenomenon may be transparent to the IP layer if both APs belong to the same coverage area of an AR, it is certainly an important issue to address if the APs are associated with two different ARs.

To successfully avoid the negative impacts of ping pong, it is important to avoid sending BUs every time the MN attaches to a new AR. In addition it is important to avoid packet routing defficiences

which may result in packets dropped for the duration of the ping pong.
Fast Handoffs can handle this phenomenon by issuing a bicast request through the "old" AR and before the MN moves to the "new" AR. This will ensure that the MN will continue receiving packets addressed to it irrespective of its current AR. Hence the packet losses due to IP mobility can be reduced to zero.
Since the MN will have an entry in its Binding Update List (BUL) indicating that a BU was sent with a bicasting request, the MN will not need to resend BUs whenever a new router advertisement is received from one of the ARs it is moving in between.

5. Extensions for Fast Handoffs

5.1 Extensions to MIPv6

To allow bicasting from the MAP to take place, a new flag,_B_, is added to the BU message. Upon reception of a BU message with the _B_ flag set, a MAP SHOULD bicast all incoming packets addressed to the MN to its current COA as well as the new COA in the BU requesting the bicast. The new BU message is shown below.

```
  0                   1                   2                   3
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
                                +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                                |  Option Type  | Option Length |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |A|H|R|D|M|B|Res| Prefix Length |        Sequence Number        |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                            Lifetime                           |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |   Sub-Options...
 +-+-+-+-+-+-+-+-+-+-+-+-+-
```

Description of extensions to the BU option:

B If set, it indicates a request for bicasting all traffic received for the MN to its current address as well as the new address in the BU.

Res 2 bit reserved field

5.2 Extensions to Neighbour Discovery

```
       0                   1                   2                   3
       0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |     Type      |    Length     | Prefix Length |L|A|R|P| Res   |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                         Valid Lifetime                        |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                       Preferred Lifetime                      |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                           Reserved                            |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                                                               |
      +                                                               +
      |                                                               |
      +                            Prefix                             +
      |                                                               |
      +                                                               +
      |                                                               |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Modified fields

P   Indicates that the prefix included is the anticipated one on the new subnet. The prefix advertised MUST NOT be treated as an on link prefix.

Res 4-bit reserved field.

6. Fast Handoffs and DAD

Duplicate Address Detection (DAD) was defined in [4] to avoid address duplication on links when stateless address autoconfiguration is used. The use of DAD to verify the uniqueness of a statelessly configured IPv6 address may add delays to a MIPv6 handoff.
The probability of an interface identifier duplication on the same subnet can be considered very low, however it can not be ignored.

In this draft certain precautions are proposed to minimise the effects of a duplicate address occurrence.

In an HMIPv6 network as described in [1], a MN can register with one or more MAPs while moving within one or more MAP domains. Since a MAP domain is likely to include more than one router, when receiving a BU from the MN a MAP can check if other nodes within its domain are using the same interface identifier or address. If address duplication is detected, a MAP MUST reject the BU with the approriate fault code.

If a duplication of the interface identifier is found, a MAP MAY accept the BU and include a new code in the BAck to warn the MN from choosing this interface identifier value.
As a result a MN may change the interface identifier immediately or before moving to a new subnet.
This method guarantees that no address duplication will take place between addresses registered in the MAP.
It should be noted that a MAP higher in the hierarchy may provide better address "coverage" to allow the MN to predict address duplication earlier.

In the case of a flat architecture as decribed in chapter 3.1 a MAP will not be able to know other MN addresses outside its subnet. Since the MN may not register with the "new" MAP until it moves to the new subnet, or it may simply not register for a long time, there is little benefit in using the approach mentioned above.

Hence, to avoid delays due to DAD in a flat mobility management architecture, a MN may choose to continue sending and receiving traffic using its newly formed COA while performing DAD on the new subnet. In the case where a duplication exists, the MN MUST follow the rules in [4].

This issue is not specific to this proposal and may also be addressed in future revisions of [2].

7. References

[1] H. Soliman, C. Castellucia, K. El Malki and L. Bellier "Hierarchical Mobile IPv6 and Fast Handoffs", draft-ietf-mobileip-hmipv6-00.txt (work in progress), September 2000

[2] D. Johnson and C. Perkins, "Mobility Support in IPv6", draft-ietf-mobileip-ipv6-12.txt, February 2000.

[3] K. El Malki and H. Soliman " Fast Handoffs in Mobile IPv4". (work in progress)

[4] S. Thomson and T. Narten "IPv6 Stateless Address Autoconfiguration". RFC 2462.

[5] T. Narten, E. Nordmark and W. Simpson " Neighbour Discovery for IP version 6 ". RFC 2461

8. Acknowledgements

The authors would like to thank Claude Castellucia (INRIA) for his helpful contribution to this draft. In particular, his valuable

contributions for DAD handling. Erik Nordmark provided very useful

comments on handling of inter-AR solicitation as well as DAD handling.

The authors would also like to thank the following members of the working group (in alphabetical order) for their comments and the interesting discussions on this draft: Gopal Dommety (Cisco), Dave Johnson (Rice University), Conny Larsson (Ericsson), Mohan Parthasarathy (Sun), Basavaraj Patil (Nokia), Charles Perkins (Nokia), Mattias Pettersson (Ericsson), Carl Williams (Sun)and Alper Yegin (Sun).

9. Addresses

The working group can be contacted via the current chairs:

Basavaraj Patil	Phil Roberts
Nokia Corporation	Motorola M/S M8-540
6000 Connection Drive	1501 West Shure Drive
Irving, TX 75039	Arlington Heights, IL 60004
USA	USA
Phone: +1 972-894-6709	Phone: +1 847-632-3148
EMail: Raj.Patil@nokia.com	EMail: QA3445@email.mot.com
Fax : +1 972-894-5349	

Questions about this memo can be directed to:

Karim El Malki
Ericsson Radio Systems AB
Access Networks Research
SE-164 80 Stockholm
SWEDEN

Phone: +46 8 7573561
Fax: +46 8 7575720
E-mail: Karim.El-Malki@era.ericsson.se

Hesham Soliman
Ericsson Australia
61 Rigall St., Broadmeadows
Melbourne, Victoria 3047
AUSTRALIA

Phone: +61 3 93012049
Fax: +61 3 93014280
E-mail: Hesham.Soliman@ericsson.com.au