Universität Karlsruhe (TH)
Institut für Telematik

TELEMATICS TECHNICAL REPORTS

Credit-Based Authorization for Concurrent IP-Address Tests

Christian Vogt
chvogt@tm.uka.de

June 10, 2005

TM-2005-3

ISSN 1613-849X
http://doc.tm.uka.de/tr/

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

Abstract

Route optimization enables mobile nodes to directly communicate with one another. This is an important efficiency benefit of modern mobility protocols like Mobile IPv6 or the Host Identity Protocol. However, route optimization can introduce the possibility for a new type of amplified flooding attacks if designed without care: An attacker may misuse the protocol to trick its peer into redirecting a flow of packets to a false, i.e., a victim's, IP address. A precautionary counter-measure used by various mobility protocols is to first determine whether the right node is present at a new IP address before any data packets are sent to that address. The test can be as simple as a ping carrying some unguessable, to-be-returned piece of data. Yet, an unfortunate side effect of this common approach is that it increases handover latency by one round-trip time, precluding interactive or real-time applications in many scenarios. This paper proposes a credit-based strategy that allows peers to continue communications while *a new IP address is being examined. The optimization is exemplarily applied to Mobile IPv6 and the Host Identity Protocol, for which it reduces handover-signaling delays by 50%.*

1 Introduction

The traditional approach to IP mobility, *bidirectional tunneling*, is to assign to a mobile node a stationary proxy and to relay all traffic through a two-way tunnel between the mobile node and this proxy. A correspondent node communicates with an IP address that routes to the proxy. Mobility can so be handled transparently to the correspondent node, solving the problem with non-mobile legacy hosts. More efficient routing, *route optimization*, can be achieved when the correspondent node itself tracks the mobile node's current location, and peers exchange packets directly. The mobile node updates the state at the correspondent node whenever it moves to another IP subnet. Mobile IPv6 [1] and the Host Identity Protocol (HIP) [2] are two popular IPv6 mobility protocols with support for route optimization.

This paper studies the well-known threat of route-optimization misuse for the purpose of malicious packet redirection and third-party flooding [3]. In such an attack, the perpetrator, who may or may not be really mobile, claims to have moved to a new IP subnet and registers a victim's IP address as its alleged new locality. The correspondent node may so be tricked into spamming the victim with unwanted packets. Flooding attacks have been thoroughly considered during the design of Mobile IPv6 and HIP [4][5]. The common protection mechanism adopted by both protocols is to probe a mobile node's new IP address, subsequent to handover, by sending there some unguessable piece of data which the mobile node must return. The correspondent node refrains from sending any data packets to the new IP address until the probe completes successfully. Naturally, this increases handover latency by one round-trip time. Users of interactive or real-time applications may not accept this additional delay.

On the basis of Mobile IPv6 and HIP, this paper presents *Credit-Based Authorization* (CBA), a generic technique that allows a mobility protocol to securely use a new IP address while this address is being probed. CBA can reduce signaling delays by one round-trip time, which amounts to 50% of the overall latency in the cases of both Mobile IPv6 and HIP. Following this introduction, section 2 further discusses the threat of redirection-based flooding attacks. Sections 3 and 4 introduce the Mobile IPv6 and HIP mobility protocols, respectively. Section 5 explains CBA conceptually, and section 6 described how CBA can be integrated into both Mobile IPv6 and HIP. Related work is discussed in section 7. Section 8 finally concludes this paper.

2 Redirection-based Flooding

This section discusses new possibilities for malicious packet redirection and third-party flooding that poorly designed mobility protocols could introduce. It relates these to similar threats in the non-mobile Internet, discusses the role of packet filtering, and outlines and evaluates the protection mechanism used in Mobile IPv6 and HIP.

It is in general infeasible to assume any sort of pre-existing relationship between two nodes who want to use route optimization. As a consequence, it is often unclear from the correspondent node's perspective whether a mobile node is faithful with respect to its current locality. The mobile node might misuse such unawareness for redirecting packets, the true recipient of which it is, to a victim somewhere else in the Internet. Redirection-based flooding attacks are an attractive means for denial of service because of their enormous potential for amplification at relatively low cost [6]. For instance, a mobility protocol that fails to provide appropriate counter-measures would allow an attacker to setup a TCP connection through the right IP address, learn the initial sequence number this way, and subsequently redirect the connection to a false address. The attacker could influence the rate of misrouted TCP segments through fake acknowledgements that appear to originate at the new connection end point. As the segments are typically much bigger than

their acknowledgements, the correspondent node would spend, unknowingly, much more resources on the flooding attack than the attacker itself.

One may argue that flooding attacks are already a major problem in the non-mobile Internet [7], and that mobility protocols cannot do anything about them. However, it is important to understand that misuse of packet redirection could render such attacks much more accessible: Today's attackers gain amplification by compromising as many Internet nodes as possible through viral software and have them contribute to the attack this way. In contrast, a widely deployed, but improperly designed mobility protocol would allow the attacker to take advantage of the large base of correspondent nodes, making any viral assaults unnecessary.

Most mobility protocols hinder registration of false IP addresses by requiring mobile nodes to put a new IP address in the Source Address field of the registration message's IP header. Filtering techniques on the mobile node's side [8] so get a chance to unveil fraudulent registration attempts. Yet, the problem with verifying IP source addresses at the fringe of the Internet is that it does not fully protect unless applied *universally*. It is questionable whether this will always be the case. As things stand, an attacker can always find a network where no filtering is applied, even though the technique has already been deployed in many places. Mobility protocols should hence provide independent protection against malicious packet redirection.

As previously mentioned, mobility protocols like Mobile IPv6 and HIP require the mobile node to receive and return some random data at a new IP address before any data packets are sent to that address. A successful exchange guarantees that the mobile node either owns the new IP address, or is at least on the path towards it. In either case, any packets sent to the address are routed to, or via, the mobile node. The rationale for considering this evidential that no flooding attack is in the making is that an attacker in such a position would not depend on redirection: It could wage a flooding attack more easily by setting up, say, a TCP connection directly on behalf of its victim.

3 Mobility with Mobile IPv6

This section provides a brief overview on the Mobile IPv6 mobility protocol. The reader may refer to RFC 3775 [1] for the complete specification.

Mobile IPv6 uses two IP addresses per mobile node: a dynamic *care-of address* for the purpose of routing and a static *home address* for end-node identification at stack layers above IP. The addresses are swapped during IP processing in the end nodes so that applications can stick to the home address, whereas routers look at the care-of address. The care-of address routes to the mobile node's actual location. It changes as the mobile node moves. The home address always routes to the mobile node's *home network*. There, a *home agent* serves as a the mobile node's proxy in case the mobile node is away from home and its peer does not support route optimization. The peer then talks to the home address, and packets are bidirectionally tunneled between the home agent and the mobile node. The home agent is also a relay for certain signaling messages, as will be explained below.

When a mobile node finds that it has moved, it configures a new care-of address and registers it with its home agent and correspondent node.[1] Figure 1 illustrates this procedure. The mobile node initiates the *home registration* by sending a Binding Update (BU) message to the home agent. The BU contains the home address as an identifier, the new care-of address to which future packets shall be directed, and some supplementary information. If the home registration succeeds, the home agents returns a Binding Acknowledgement (BA) message. Home registrations are protected through IPsec to prevent misuse by unauthorized nodes.

No IPsec security association generally exists between the mobile node and a correspondent node, so a *correspondent registration* must be protected otherwise. Essentially, a correspondent registration can only be considered safe if the correspondent node has some assurance that the mobile node's home address and care-of address are true. The way this is accomplished in Mobile IPv6 is through probing both addresses in parallel: Triggered by the mobile node's Home Test Init (HoTI) and Care-of Test Init (CoTI) messages, the correspondent node generates a pair of unguessable tokens and sends them back to the mobile node in a Home Test (HoT) message and a Care-of Test (CoT) message, respectively. The HoTI and HoT are routed through the home address, whereas the CoTI and CoT take the direct path. This two-fold message exchange is called the *return-routability procedure*.

It is important to emphasize the different purpose of the two address tests. As the home address identifies the mobile node in transport connections and applications, the home-address test serves to authenticate the

[1] The mobile node may have multiple correspondent nodes at a time, but for simplicity reasons, it is assumed in this paper that there is only a single one.

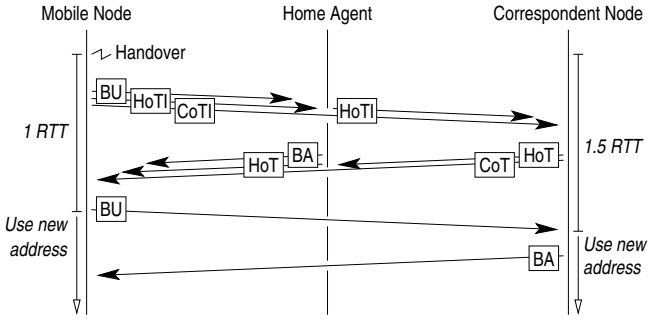

Figure 1: Mobile IPv6 address registration

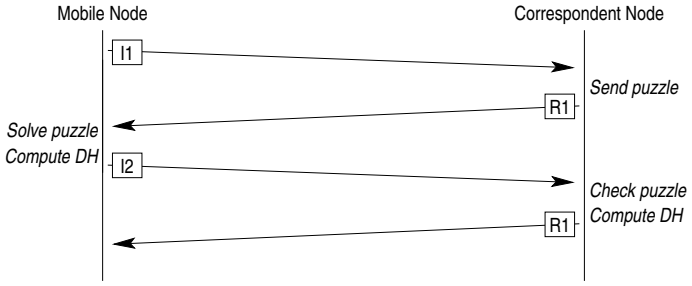

Figure 2: HIP base exchange

mobile node to the correspondent node during a correspondent registration. The care-of-address test, in contrast, provides a reachability test.

The mobile node requires both tokens from the return-routability procedure to form a key with which it can sign the BU to be sent to the correspondent node subsequently. The signature, in turn, allows the correspondent node to decide that the mobile node, first, owns the home address (because one token was sent to that address) and, second, is reachable at the care-of address (to which the other token was sent). The mobile node may optionally request the correspondent node to return a BA for confirmation by setting a flag in the BU.

As the security provided by the return-routability procedure is weak compared to cryptographic mechanisms, a correspondent registration must be refreshed every seven minutes even in the absent of handovers. Tokens are good for only 3.5 minutes. If a new handover occurs within this lifetime, the mobile node may reuse the token from the previous home-address test and omit the HoTI/HoT exchange; otherwise the complete return-routability procedure must be redone.

4 Mobility with HIP

This section outlines the HIP base protocol and its extension for mobility support. The protocol details can be found in specifications [2] and [9].

HIP uses a node's DSA or RSA public key, rather than an IP address, as a *Host Identifier* (HI) at stack layers above IP. A HI is hashed into a 128-bit *Host Identity Tag* (HIT) in order to be syntactically compatible with an IPv6 address. Peers swap their HITs and actual IP addresses within a shim layer between the IP stack and upper layers. The advantage of a HI over an IP address is that the owner can *cryptographically* prove to a peer that its identifier is correct through being able to sign or decrypt messages with the corresponding private key. HIP uses this feature to authenticate a Diffie-Hellman key exchange for IPsec ESP protection of both signaling and data packets.

Contact establishment is shown in figure 2. Either one of the peers, the initiator, starts a *HIP base exchange* by sending the other, the responder, an I1 message. This prompts the responder to send its HI and a public Diffie-Hellman key to the initiator in the R1 message. The initiator, in turn, provides its HI and a public Diffie-Hellman key in the I2 message. Based on their own private and the other's public Diffie-Hellman keys, both peers compute secret shared keying material from which authentication and encryption keys are taken. Finally, the R2 message serves to activate IPsec ESP processing on both sides. The R1, I2, and R2 messages are signed with the HIs. This makes the base exchange robust to man-in-the-middle attacks.

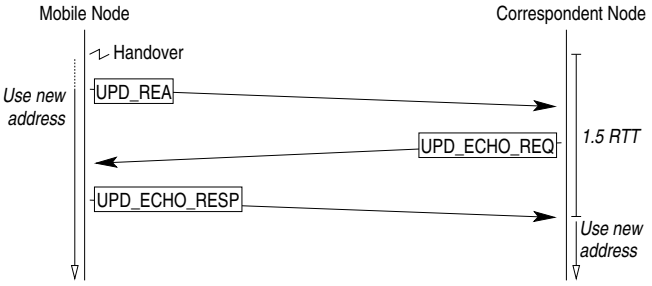

Figure 3: HIP address registration

Since a Diffie-Hellman key exchange includes heavy cryptographic computations, it can potentially be misused for a resource-exhaustion attack against the responder. In such an attack, the perpetrator uses any random number as its public Diffie-Hellman key and makes the responder compute keying material from it. The perpetrator may repeat the attack with different falsified IP source addresses and HIs. Resource-exhaustion attacks can be averted by requiring the initiator to invest some reasonable effort before the responder does so. To this end, the R1 message contains a random number X, a *puzzle*, which the initiator has to concatenate with both peers' HITs and a number Y such that the hash on the concatenation yields a string with a certain minimum of trailing zeros. The R1 message, including its signature, can be pre-computed and reused across multiple base exchanges. The appealing property of a puzzle is that the complexity of solving it grows exponentially with the required count of trailing zeros, whereas a potential solution can be validated very efficiently.

Mobile nodes can change their IP addresses without another run of the base exchange. They may derive new keys at any time, but usually retain their existing ones across multiple handovers. (Mobile IPv6 requires periodic correspondent registrations because the return-routability procedure is weaker, from a security standpoint, than HIP's cryptographic mechanisms.) When the mobile node moves to a different IP subnet, it configures a new IP address and signals it to the correspondent node within an Update message carrying a Readdress parameter (UPD_REA). The UPD_REA holds the corresponding IPsec Security Parameter Index, the new IP address, and some auxiliary data. The IPsec ESP context, within which all signaling takes place, ensures the authenticity of this UPD_REA. Figure 3 illustrates the registration procedure.

A node may register multiple IP addresses with its peer and declare one of them *preferred*. IP-address verification is mandatory only in case the preferred address changes to an address that has not yet been actively used. To verify a mobile node's reachability at a new preferred address, the peer sends an Update message including an Echo Request parameter (UPD_ECHO_REQUEST). The UPD_ECHO_REQUEST includes a random number, which the mobile node must return in an Update message with an Echo Response parameter (UPD_ECHO_RESPONSE).

5 Credit-Based Authorization

Both Mobile IPv6 and HIP require verification of a mobile node's reachability at a new IP address *before* data packets are sent to that address. This inhibits direct use of the new IP address as soon as it becomes available.[2] At the same time, as previously explained, the severity of redirection-based flooding attacks over conventional flooding attacks not so much emanates from packet redirection per se, but rather from the enormous potential for flooding amplification. If no amplification could be gained through redirection, it would just be easier for an attacker to bombard its victim directly. As a consequence, introducing a mechanism that prevents this amplification would afford reasonably secure, *immediate* use of a new IP address subsequent to handover. Such is the approach followed by CBA.

CBA allows a correspondent node to probe a mobile node's reachability at a new IP address while the address is already in active use. This implies two address states: The new IP address is *unverified* until the correspondent node knows the result of the reachability test, and it is *verified* thereafter. The idea of CBA is to restrict the maximum data volume and rate that a correspondent node can send to an unverified IP address to the data volume and rate that the mobile node has sent the other way in the recent past. For this, the mobile node has a credit account at the correspondent node. The account fills up as the mobile node sends packets to the correspondent node. Subsequent to handover, while the new IP address is unverified,

[2]This is obvious in the case of HIP. In Mobile IPv6, one may argue that the care-of-address test runs in parallel with the home-address test, so nodes must wait after a handover anyway. However, as will be explained in section 6, the home-address test may be performed proactively before the handover, leaving only the care-of-address test to be done afterwards.

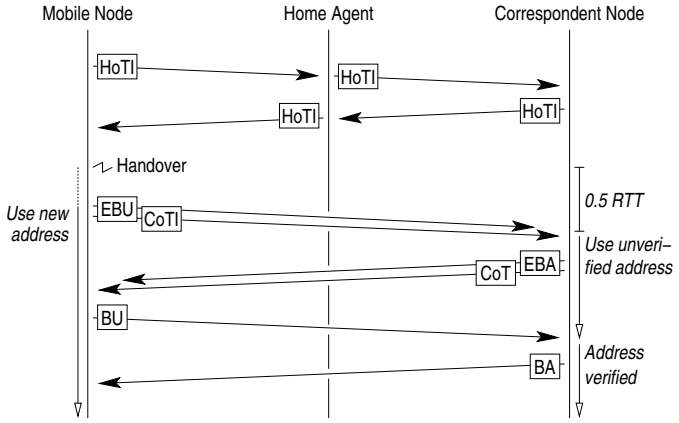

Figure 4: Mobile IPv6 address registration with CBA

packets sent to this address consume the credit. Packets continue to be sent there as long as sufficient credit is left.

The balance of a mobile node's credit account increases and decreases proportionally to the size of the received and sent packets. This guarantees that the data volume sent to an unverified IP address cannot be greater than the previously received data volume. In addition, unused credit withers over time through exponential aging. This helps to bound the rate at which data can be sent to an unverified IP address by the rate at which data was previously received.

Up to this point, the mobile node earns credit for sending packets, while it needs the credit for receiving them. This works well when traffic patterns are for the most part symmetric, as it is in the case of Internet telephony. On the other hand, asymmetric traffic patterns are also very common. File transfers and media streaming, for instance, feature high throughput towards the client, typically the mobile node, and comparably little throughput towards the serving correspondent node.

To prevent a mobile node from running out of credit, CBA can be made a bit more sophisticated. The key observation is that the mobile node invests comparable effort for packet reception as for packet transmission, in terms of bandwidth, memory, and processing capacity. It hence stands to reason that packets received by the mobile node may be taken as the basis for credit allocation, just like the packets that the mobile node sends. The question, though, is how the correspondent node can determine how many of the packets sent to the mobile node are actually received and processed at the other end. The mobile node may position itself behind a low-bandwidth link and deliberately collect more credit than appropriate. It may also fake transport-layer acknowledgements as previously explained. To overcome the issue, a correspondent node needs some feedback on packet-loss conditions along the path towards the mobile node. Such feedback can optionally be provided through *IP-Address Spot Checks*. Here, the correspondent node periodically tags packets that it sends to the mobile node with a random, unguessable token. When the mobile node receives the packet, it stores it in a cache until a local application delivers the next packet for the correspondent node. The mobile node includes all cached tokens in this packet and sends it. The correspondent node keeps an eye on how many of the tokens that it sent the mobile node correctly returns. New credit is allocated proportionally. The preciseness of IP-Address Spot Checks can be traded with overhead through the frequency with which they are exercised.

6 Applying CBA

As Mobile IPv6's care-of-address test is parallelized with the home-address test, it is not necessarily clear how CBA should be applied (and have an advantageous effect on handover latency). The idea is to split the return-routability procedure into its components as shown in figure 4: Since a mobile node keeps its home address across handovers, the home-address test may be completed already before the handover. The test could be initiated periodically or triggered by the local link layer just in time when a handover is imminent. The care-of-address test can then be scheduled independently.

Subsequent to movement, the mobile node immediately informs the correspondent node about its new care-of address with an Early Binding Update (EBU) message. The mobile node signs the EBU with a key that it computes based on the token from the proactive home-address test. Since a token from a care-of-address test is unavailable at this early time, the EBU lacks any reachability information about the mobile

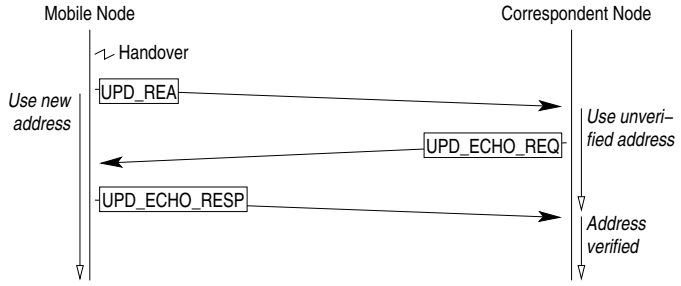

Figure 5: HIP address registration with CBA

node. When the correspondent node receives the EBU, it switches to the new care-of address, labeling it ”unverified”. Packets sent to the mobile node are henceforth subject to credit limitations.

The mobile node initiates the concurrent care-of-address test right after sending the EBU. When the test concludes, the mobile node sends a standard BU to the correspondent node. The correspondent node re-labels the new care-of address ”verified” once it receives the message.

Applying CBA to HIP is more straightforward, because IP-address verification is not parallelized with other protocol tasks. The correspondent node switches to the new, unverified IP address when it receives the UPD_REA, and changes the address status to ”verified” upon reception of the UPD_ECHO_RESPONSE. Figure 5 shows this procedure.

7 Related Work

This section introduces some related approaches to concurrent IP-address tests and reduced handover delays which either have been proposed or are well conceivable.

Correspondent nodes may use *heuristics* for misuse detection in replacement of Credit-Based Authorization. In conjunction with a restrictive lifetime limit for unverified IP addresses, this can prevent, or at least effectually discourage, malicious packet redirection. The challenge here seems to be a feasible heuristic: On one hand, the heuristic must be sufficiently rigid to stop an attacker early on. On the other hand, it should not adversely affect communications with faithful mobile nodes.

In the special case of Mobile IPv6, peers may temporarily resort to *bidirectional tunneling* while a new care-of address is being verified. This technique appeared originally in [10] and has since been used also in [11]. Its performance impact strongly depends on the topological distance between the mobile node and its home agent compared to the distance between the mobile node and the correspondent node. The benefit is high if the mobile node is close to the home agent, but far away from the correspondent node. On the other hand, there may even be a performance penalty in the reverse case.

Local handover support from the mobile node’s access network can bridge end-to-end signaling delays during a handover. For instance, in [12], packets are temporarily routed through a bidirectional tunnel between the mobile node’s old and new access routers while the mobile node updates its home agent and correspondent nodes.

8 Conclusions

Mobility protocols may introduce a new type of flooding attacks if designed without sufficient precautions. Compared to existing flooding techniques, such redirection-based flooding attacks can yield unprecedented amplification at negligible investments from the attacker’s side, jeopardizing not only nodes that participate in the mobility protocol, but the Internet at a whole.

This paper explains the new threat of malicious packet redirection, clarifies why existing security techniques fail to provide full protection, and infers that it is up to the mobility protocols themselves to protect against it. The paper evaluates the standard security mechanism adopted in Mobile IPv6 and HIP, namely, to probe a new IP address before any data packets are sent there. As it becomes evident that this simple approach adversely affects handover latency, the paper proposes a generic, credit-based strategy, Credit-Based Authorization, for secure *concurrent* probing. Finally, the paper explains how Credit-Based Authorization can be integrated into Mobile IPv6 and HIP, where it reduces handover-signaling delays by 50%.

Acknowledgments

For their valuable feedback and advice with respect to the work presented in this paper the author wishes to thank Jari Arkko, Roland Bless, Mark Doll, Tobias Küfner, Pekka Nikander, and Lars Völker.

References

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

[2] R. Moskowitz, P. Nikander, P. Jokela, and T. R. Henderson. Host Identity Protocol. IETF Internet Draft draft-ietf-hip-base, work in progress.

[3] T. Aura, M. Roe, and J. Arkko, "Security of Internet Location Management," pp. 78–87, December 2002.

[4] P. Nikander, J. Arkko, T. Aura, G. Montenegro, and E. Nordmark. Mobile IP version 6 Route Optimization Security Design Background. IETF Internet Draft draft-ietf-mip6-ro-sec, work in progress.

[5] R. Moskowitz and P. Nikander. Host Identity Protocol Architecture. IETF Internet Draft draft-ietf-hip-arch, work in progress.

[6] M. Roe, T. Aura, G. O'Shea, and J. Arkko. Authentication of Mobile IPv6 Binding Updates and Acknowledgments. IETF Internet Draft draft-roe-mobileip-updateauth, work in progress.

[7] V. Paxson, "An Analysis of Using Reflectors for Distributed Denial-of-Service Attacks," *ACM SIGCOMM Computer Communication Review*, vol. 31, no. 3, July 2001.

[8] P. Ferguson and D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing," RFC 2827, May 2000.

[9] P. Nikander, J. Arkko, and T. R. Henderson. End-Host Mobility and Multi-Homing with Host Identity Protocol. IETF Internet Draft draft-ietf-hip-mm, work in progress.

[10] C. Vogt, R. Bless, M. Doll, and T. Kuefner. Early Binding Updates for Mobile IPv6. IETF Internet Draft draft-vogt-mip6-early-binding-updates, work in progress.

[11] W. Haddad, L. Madour, J. Arkko, and F. Dupont. Applying Cryptographically Generated Addresses to Optimize MIPv6 (CGA-OMIPv6). IETF Internet Draft draft-haddad-mip6-cga-omipv6, work in progress.

[12] E. Rajeev Koodli. Fast Handovers for Mobile IPv6. IETF Internet Draft draft-ietf-mipshop-fast-mipv6, work in progress.