Internet Engineering Task Force (IETF)
Request for Comments: 6521
Category: Experimental
ISSN: 2070-1721

A. Makela
Aalto University/Comnet
J. Korhonen
Nokia Siemens Networks

February 2012

Home Agent-Assisted Route Optimization between Mobile IPv4 Networks

Abstract

This document describes a home agent-assisted route optimization functionality for the IPv4 Network Mobility Protocol. The function is designed to facilitate optimal routing in cases where all nodes are connected to a single home agent; thus, the use case is route optimization within a single organization or similar entity. The functionality enables the discovery of eligible peer nodes (based on information received from the home agent) and their network prefixes, and the establishment of a direct tunnel between such nodes.

Status of This Memo

This document is not an Internet Standards Track specification; it is published for examination, experimental implementation, and evaluation.

This document defines an Experimental Protocol for the Internet community. This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Not all documents approved by the IESG are a candidate for any level of Internet Standard; see Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc6521.

Table of Contents

1. Introduction and Motivations

Traditionally, there has been no method for route optimization in Mobile IPv4 [RFC5944] apart from an early attempt [MIP-RO]. Unlike Mobile IPv6 [RFC6275], where route optimization has been included from the start, with Mobile IPv4, route optimization hasn’t been addressed in a generalized scope.

Even though general route optimization may not be of interest in the scope of IPv4, there are still specific applications for route optimization in Mobile IPv4. This document proposes a method to optimize routes between networks behind Mobile Routers (MRs), as defined by Network Mobility (NEMO) [RFC5177]. Although NAT and the pending shortage of IPv4 addresses make widespread deployment of end-to-end route optimization infeasible, using route optimization from

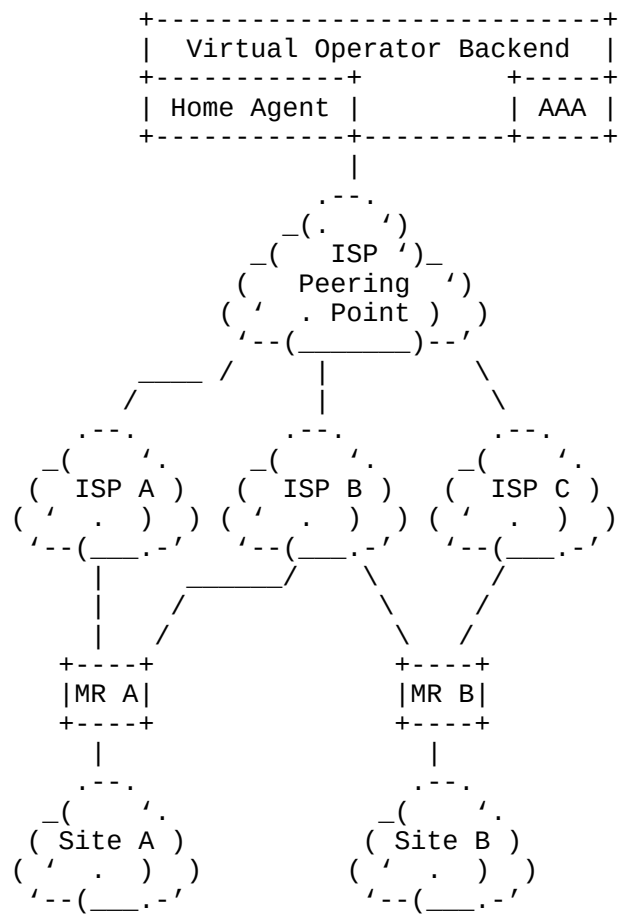
MR to MR is still a practical scenario. Note that the method specified in this document is only for route optimization between MRs; any network prefix not advertised by an MR would still be routed via the home agent, although an MR could advertise very large address spaces, e.g., by acting as an Internet gateway.

A particular use case concerns setting up redundant yet economical enterprise networks. Recently, a trend has emerged where customers prefer to maintain connectivity via multiple service providers. Reasons include redundancy, reliability, and availability issues. These kinds of multihoming scenarios have traditionally been solved by using such technologies as multihoming BGP. However, a more lightweight and economical solution is desirable.

From a service provider perspective, a common topology for an enterprise customer network consists of one to several sites (typically headquarters and various branch offices). These sites are typically connected via various Layer 2 technologies (ATM or Frame Relay Permanent Virtual Circuits (PVCs)), MPLS VPNs, or Layer 3 site-to-site VPNs. With a Service Level Agreement (SLA), a customer can obtain very reliable and well-supported intranet connectivity. However, compared to the cost of "consumer-grade" broadband Internet access, the SLA-guaranteed version can be considered very expensive. These consumer-grade options, however, are not a reliable approach for mission-critical applications.

Mobile IP, especially MRs, can be used to improve reliability of connectivity even when implemented over consumer-grade Internet access. The customer becomes a client for a virtual service provider, which does not take part in the actual access technology. The service provider has a backend system and an IP address pool that it distributes to customers. Access is provided by multiple, independent, possibly consumer-grade ISPs, with Mobile IP providing seamless handovers if service from a specific ISP fails. The drawback of this solution is that it creates a star topology; all Mobile IP tunnels end up at the service provider-hosted home agent, causing a heavy load at the backend. Route optimization between mobile networks addresses this issue, by taking the network load off of the home agent and the backend.

An example network is pictured below:



Virtual Service Provider Architecture Using NEMOv4

In this example case, the organization network consists of two sites that are connected via two ISPs for redundancy reasons. Mobile IP allows fast handovers without the problems of multihoming and BGP peering between each individual ISP and the organization. The traffic, however, takes a non-optimal route through the virtual operator backend.

Route optimization addresses this issue, allowing traffic between Sites A and B to flow directly through ISP B's network, or in case of a link failure, via the ISP peering point (such as the Metropolitan Area Ethernet (MAE), e.g., MAE-West). The backend will not suffer from heavy loads.

The specification in this document is meant to be Experimental, with the primary design goal of keeping the load on the backend to a minimum. Additional design goals include extensibility to a more generalized scope, such as not requiring all MRs to be homed on the same home agent. Experiences are mostly sought regarding applicability to real-world operations, and protocol-specific issues such as signaling scalability, interworking with other Mobile IP extensions not specifically addressed in this document, and behavior of end-user applications over route-optimized paths.

The aforementioned use case is the original application. Moving this specification to Standards Track should be considered after enough deployment experience has been gathered. Besides the aforementioned issues, additional elements that might require refinement based on real-world experiences are delivery of information on networks managed by peer MRs; conducting MR <-> MR authentication; reaction to, and recovery methods for, connectivity breakdowns and other break-before-make topology changes; keepalive timer intervals; formats of signaling extensions; behavior in NAT/firewalled environments; and the prefix and realm compression algorithms.

2. Terms and Definitions

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 [RFC2119].

Care-of Address (CoA)

RFC 5944 [RFC5944] defines a care-of address as the termination point of a tunnel toward a mobile node, for datagrams forwarded to the mobile node while it is away from home. The protocol can use two different types of CoA: a "foreign agent care-of address", which is an address of a foreign agent with which the mobile node is registered, and a "co-located care-of address", which is an externally obtained local address that the mobile node has associated with one of its own network interfaces. However, in the case of Network Mobility, foreign agents are not used, so no foreign CoAs are used either.

Correspondent Router (CR)

RFC 5944 [RFC5944] defines a correspondent node as a peer with which a mobile node is communicating. A CR is a peer MR that MAY also represent one or more entire networks.

Home Address (HoA)

RFC 5944 [RFC5944] defines a home address as an IP address that is assigned for an extended period of time to a mobile node. It remains unchanged regardless of where the node is attached to the Internet.

Home Agent (HA)

RFC 5944 [RFC5944] defines a home agent as a router on a mobile node's home network that tunnels datagrams for delivery to the mobile node when it is away from home and maintains current location information for the mobile node. For this application, the "home network" sees limited usage.

Host Network Prefix

A host network prefix is a network prefix with a mask of /32, e.g., 192.0.2.254/32, consisting of a single host.

Mobility Binding

RFC 5944 [RFC5944] defines Mobility Binding as the association of an HoA with a CoA, along with the lifetime remaining for that association.

Mobile Network Prefix

RFC 5177 [RFC5177] defines a mobile network prefix as the network prefix of the subnet delegated to an MR as the mobile network.

Mobile Router (MR)

RFC 5177 [RFC5177] and RFC 5944 [RFC5944] define a mobile router as a mobile node that can be a router that is responsible for the mobility of one or more entire networks moving together, perhaps on an airplane, a ship, a train, an automobile, a bicycle, or a kayak.

Route Optimization Cache

A Route Optimization Cache is defined as a data structure, maintained by MRs, containing possible destinations for route optimization. The cache contains information (HoAs) on potential CRs and their associated mobile networks.

Return Routability (RR)

Return routability is defined as a procedure to bind an MR's HoA to a CoA on a CR with a degree of trust.

| (Concatenation)

Some formulas in this specification use the symbol "|" to indicate bytewise concatenation, as in A | B. This concatenation requires that all of the octets of the datum A appear first in the result, followed by all of the octets of the datum B.

First (size, input)

Some formulas in this specification use a functional form "First (size, input)" to indicate truncation of the "input" data so that only the first "size" bits remain to be used.

3. Mobile IPv4 Route Optimization between Mobile Networks

This section describes the changed functionality of the HA and the MR compared to the base NEMOv4 operation defined in [RFC5177]. The basic premise is still the same; MRs, when registering with the HA, may inform the HA of the mobile network prefixes they are managing (explicit mode), or the HA already knows the prefix assignments. However, instead of prefix <-> MR mapping information only remaining on the HA and the single MR, this information will now be distributed to the other MRs as well.

Home agent-assisted route optimization is primarily intended for helping to optimize traffic patterns between multiple sites in a single organization or administrative domain; however, extranets can also be reached with optimized routes, as long as all MRs connect to the same HA. The procedure aims to maintain backward compatibility; with legacy nodes or routers, full connectivity is always preserved, even though optimal routing cannot be guaranteed.

The scheme requires an MR to be able to receive messages from other MRs unsolicited -- that is, without first initiating a request. This behavior -- accepting unsolicited messages -- is similar to the registration revocation procedure [RFC3543]. Many of the mechanisms are the same, including the fact that advertising route optimization support upon registration implies the capability to receive Registration Requests and Return Routability messages from other MRs.

Compared to IPv6, where mobile node <-> correspondent node bindings are maintained via Mobility Routing header and home address options, Mobile IPv4 always requires the use of tunnels. Therefore, inter-mobile-router tunnel establishment has to be conducted.

3.1. Maintaining Route Optimization Information

During registration, a registering MR MAY request information on route-optimizable network prefixes. The MR MAY also allow redistribution of information on its managed network prefixes regardless of whether they are explicitly registered or already configured. These are indicated with a Mobile Router Route Optimization Capability Extension; see Section 5.1. If the HA accepts the request for route optimization, this is indicated with a Route Optimization Reply Extension (Section 5.2) in the Registration Reply.

Note that the redistribution of network prefix information from the HA happens only during the registration signaling. There are no "routing updates" from the HA except during re-registrations triggered by handovers, registration timeouts, and specific solicitation. The solicitation re-registration MAY occur if a CR receives a Registration Request from an unknown MR (see Section 3.3.3).

3.1.1. Advertising Route-Optimizable Prefixes

As noted, an HA that supports NEMO already maintains information on which network prefixes are reachable behind specific MRs. The only change to this functionality is that this information can now be distributed to other MRs upon request. This request is implied by including a Route Optimization Capability Extension (Section 5.1) and setting the ’R’ bit.

When an HA receives a Registration Request, standard authentication and authorization procedures are conducted.

If registration is successful and the Route Optimization Capability Extension was present in the Registration Request, the reply message MUST include the Route Optimization Reply Extension (Section 5.2) to indicate that the Route Optimization Capability Extension was understood. Furthermore, the extension also informs the MR whether NAT was detected between the HA and the MR using the procedure in RFC 3519 [RFC3519], which is based on the discrepancy between the requester’s indicated CoA and the packet’s source address.

The reply message MAY also include one Route Optimization Prefix Advertisement Extension, which informs the MR of existing mobile network prefixes and the MRs that manage them, if eligible for redistribution. The networks SHOULD be included in order of priority, with the prefixes determined, by policy, as most desirable targets for route optimization listed first. The extension is constructed as shown in Section 5.5. The extension consists of a list where each MR, identified by its HoA, is listed with corresponding prefix(es) and their respective realm(s).

Each network prefix can be associated with a realm [RFC4282], usually in the form 'organization.example.com'. Besides the routers in the customer's own organization, the prefix list may also include other MRs, e.g., a default prefix (0.0.0.0/0) pointing toward an Internet gateway for Internet connectivity or additional prefixes belonging to possible extranets. The realm information can be used to make policy decisions on the MR, such as preferring optimization within a specific realm only. Furthermore, the unique realm information can be used to differentiate between overlapping address spaces utilized by the same or different organizations concurrently and adjusting forwarding policies accordingly.

In a typical scenario, where network prefixes are allocated to MRs connecting to a single HA, the prefixes are usually either continuous or at least very close to each other. Due to these characteristics, an optional prefix compression mechanism is provided. Another optional compression scheme is in use for realm information, where realms often share the same higher-level domains. These compression mechanisms are further explained in Section 4.

Upon receiving a Registration Reply with a Route Optimization Prefix Advertisement Extension, the MR SHALL insert the MR HoAs included in the extension as host-prefixes to the local Route Optimization Cache if they do not already exist. If present, any additional prefix information SHALL also be inserted into the Route Optimization Cache.

The MR MAY discard entries from a desired starting point onward, due to memory or other policy-related constraints. The intention of listing the prefixes in order of priority is to provide implicit guidance for this decision. If the capacity of the device allows, the MR SHOULD use information on all advertised prefixes.

3.1.2. Route Optimization Cache

MRs supporting route optimization will maintain a Route Optimization Cache.

The Route Optimization Cache contains mappings between potential CR HoAs, network(s) associated with each HoA, information on reachability related to NAT and other divisions, and information related to the RR procedure. The cache is populated based on information received from the HA in Route Optimization Prefix Advertisement Extensions and in registration messages from CRs. Portions of the cache may also be configured statically.

The Route Optimization Cache contains the following information for all known CRs. Note that some fields may contain multiple entries. For example, during handovers, there may be both old and new CoAs listed.

CR-HoA

Correspondent router's home address. Primary key identifying each CR.

CR-CoA(s)

Correspondent router's care-of address(es). May be empty if none known. Potential tunnel's destination address(es).

MR-CoA

Mobile router's care-of address currently used with this CR. Tunnel's source address.

Tunnels

Tunnel interface(s) associated with this CR. The tunnel interface itself handles all the necessary operations to keep the tunnel operational, e.g., sending keepalive messages required by UDP encapsulation.

NAT states

A table of booleans. Contains entries for all pairs of potential MR-CoAs and CR-CoAs that are known to require NAT awareness. The table is populated either statically or based on information received during operation. A setting of true indicates that the MR can establish a UDP tunnel toward the CR, using this pair of CoAs. A received advertisement can indicate that the value should

be set to false for all of the respective CR's CoAs. Settings in this table affect tunnel establishment direction; see Section 3.3.4 and the registration procedure when deciding which CoAs to include in the Care-of Address Extension in the Registration Reply. The existence of an entry mandates the use of UDP encapsulation.

RRSTATEs

Return routability state for each CR-HoA - MR-CoA pair. States are INACTIVE, IN PROGRESS, and ACTIVE. If state is INACTIVE, the RR procedure must be completed before forwarding route-optimized traffic. If state is IN PROGRESS or ACTIVE, the information concerning this CR MUST NOT be removed from the Route Optimization Cache as long as a tunnel to the CR is established.

KRms

Registration management key for each CR-HoA - MR-CoA pair. This field is only used if configured statically -- if the KRm was computed using the RR procedure, it is calculated in situ based on nonces and the router key. If configured statically, RRSTATE is permanently set to ACTIVE.

Care-of nonce indices

If the KRm was established with the RR procedure, contains the care-of nonce index for each MR-CoA - CR-HoA pair.

Care-of keygen token

If the KRm was established with the RR procedure, contains the care-of keygen token for each MR-CoA - CR-HoA pair.

Home nonce indices

If the KRm was established with the RR procedure, contains the Home nonce index for each CR-HoA.

Home keygen token

If the KRm was established with the RR procedure, contains the home keygen token for each CR-HoA.

Network prefixes

A list of destination network prefixes reachable via this CR. Includes network and prefix length, e.g., 192.0.2.0/25. Always contains at least a single entry: the CR-HoA host network prefix in the form of 192.0.2.1/32.

Realms

Each prefix may be associated with a realm. May also be empty, if the realm is not provided by advertisement or configuration.

Prefix_Valid

Boolean field for each prefix - CR-HoA pair, which is set to true if this prefix's owner has been confirmed. The host network prefix consisting of the CR itself does not need validation beyond the RR procedure. For other prefixes, the confirmation is done by soliciting the information from the HA. Traffic for prefixes that have unconfirmed ownership should not be routed through the tunnel.

Information that is no longer valid due to expirations or topology changes MAY be removed from the Route Optimization Cache as desired by the MR.

3.2. Return Routability Procedure

The purpose of the RR procedure is to establish CoA <-> HoA bindings in a trusted manner. The RR procedure for Mobile IPv6 is described in [RFC6275]. The same principles apply to the Mobile IPv4 version: two messages are sent to the CR's HoA -- one via the HA using the MR's HoA, and the other directly from the MR's CoA, with two responses coming through the same routes. The registration management key is derived from token information carried on these messages. This registration management key (KRm) can then be used to authenticate Registration Requests (comparable to Binding Updates in Mobile IPv6).

The RR procedure is a method provided by Mobile IP to establish the KRm in a relatively lightweight fashion. If desired, the KRms can be configured on MRs statically, or by using a desired external secure key provisioning mechanism. If KRms are known to the MRs via some other mechanism, the RR procedure can be skipped. Such provisioning mechanisms are out of scope for this document.

The main assumption on traffic patterns is that the MR that initiates the RR procedure can always send outbound messages, even when behind a NAT or firewall. This basic assumption made for NAT Traversal in [RFC3519] is also applicable here. In the case where the CR is behind such obstacles, it receives these messages via the reverse tunnel to the CR's HoA; thus, any problem regarding the CR's connectivity is addressed during registration with the HA.

The RR procedure consists of four Mobile IP messages: Home Test Init (HoTI), Care-of Test Init (CoTI), Home Test (HoT), and Care-of Test (CoT). They are constructed as shown in Sections 5.6 through 5.9. If the MR has included the Mobile Router Route Optimization Capability Extension in its Registration Request, it MUST be able to accept Return Routability messages. The messages are delivered as Mobile IP signaling packets. The destination address of the HoTI and CoTI messages is set to the CR's HoA, with the sources being the MR's HoA and CoA, respectively.

The RR procedure begins with the MR sending HoTI and CoTI messages, each containing a (different) 64-bit random value -- the cookie. The cookie is used to bind a specific signaling exchange together.

Upon receiving the HoTI or CoTI message, the CR MUST have a secret correspondent router key (Kcr) and nonce. If it does not have this material yet, it MUST produce it before continuing with the RR procedure.

The CR responds to HoTI and CoTI messages by constructing HoT and CoT messages, respectively, as replies. The HoT message contains a home init cookie, current home nonce index, and home keygen token. The CoT message contains a care-of init cookie, current care-of nonce index, and care-of keygen token.

The home keygen token is constructed as follows:

```
Home keygen token = First (64, HMAC_SHA1 (Kcr, (home address |
   nonce | 0)))
```

The care-of keygen token is constructed as follows:

```
Care-of keygen token = First (64, HMAC_SHA1 (Kcr, (care-of address |
   nonce | 1)))
```

Note that the CoA in this case is the source address of the received CoTI message packet. The address may have changed in transit due to network address translation. This does not affect the registration process; subsequent Registration Requests are expected to arrive from the same translated address.

The RR procedure SHOULD be initiated when the Route Optimization Cache's RRSTATE field for the desired CoA with the target CR is INACTIVE. If the state was INACTIVE, the state MUST be set to IN PROGRESS when the RR procedure is initiated. In the case of a handover occurring, the MR SHOULD only send a CoTI message to obtain a new care-of keygen token; the home keygen token may still be valid. If the reply to a registration indicates that one or both of the tokens have expired, the RRSTATE MUST be set to INACTIVE. The RR procedure may then be restarted as needed.

Upon completion of the RR procedure, the Route Optimization Cache's RRSTATE field is set to ACTIVE, allowing for Registration Requests to be sent. The MR will establish a KRm. By default, this will be done using the SHA1 hash algorithm, as follows:

KRm = SHA1 (home keygen token | care-of keygen token)

When de-registering (by setting the Registration Request's lifetime to zero), the care-of keygen token is not used. Instead, the KRm is generated as follows:

KRm = SHA1 (home keygen token)

As in Mobile IPv6, the CR does not maintain any state for the MR until after receiving a Registration Request.

3.2.1. Router Keys

Each MR maintains a Kcr, which MUST NOT be shared with any other entity. The Kcr is used for authenticating peer MRs in the situation where an MR is acting as a CR. This is analogous to the node key (Kcn) in Mobile IPv6. A CR uses its router key to verify that the keygen tokens sent by a peer MR in a Registration Request are the CR's own. The router key MUST be a random number, 16 octets in length, generated with a good random number generator [RFC4086].

The MR MAY generate a new key at any time to avoid persistent key storage. If desired, it is RECOMMENDED that the keys be expired in conjunction with nonces; see Section 3.2.3.

3.2.2. Nonces

Each MR also maintains one or more indexed nonces. Nonces SHOULD be generated periodically with a good random number generator [RFC4086]. The MR may use the same nonces with all MRs. Nonces MAY be of any length, with the RECOMMENDED length being 64 bits.

3.2.3. Updating Router Keys and Nonces

The router keys and nonce updating guidelines are similar to those for Mobile IPv6. MRs keep both the current nonce and the small set of valid previous nonces whose lifetimes have not expired yet. A nonce should remain valid for at least MAX_TOKEN_LIFETIME seconds (see Section 9) after it has first been used in constructing an RR response. However, the CR MUST NOT accept nonces beyond MAX_NONCE_LIFETIME seconds (see Section 9) after the first use. As the difference between these two constants is 30 seconds, a convenient way to enforce the above lifetimes is to generate a new nonce every 30 seconds. The node can then continue to accept keygen tokens that have been based on the last 8 (MAX_NONCE_LIFETIME / 30) nonces. This results in keygen tokens being acceptable MAX_TOKEN_LIFETIME to MAX_NONCE_LIFETIME seconds after they have been sent to the mobile node, depending on whether the token was sent at the beginning or end of the first 30-second period. Note that the correspondent node may also attempt to generate new nonces on demand, or only if the old nonces have been used. This is possible as long as the correspondent node keeps track of how long ago the nonces were used for the first time and does not generate new nonces on every return routability request.

If the Kcr is being updated, the update SHOULD be done at the same time as the nonce is updated. This way, nonce indexes can be used to refer to both Kcrs and nonces.

3.3. Mobile-Correspondent Router Operations

This section deals with the operation of mobile and correspondent routers performing route optimization. Note that in the context of this document, all routers work as both MR and CR. The term "mobile router" applies to the router initiating the route optimization procedure, and "correspondent router" indicates the peer router.

There are two issues regarding IPv4 that are different when compared to Mobile IPv6 route optimization. First of all, since Mobile IPv4 always uses tunnels, there must be a tunnel established between the MR’s and the CR’s CoAs. The CR learns of the MR’s CoA, because it is included in the Registration Request. The MR learns the CR’s CoA via a new extension, "Care-of Address", in the Registration Reply. The second issue is a security consideration: In a Registration Request, the MR claims to represent an arbitrary IPv4 network. If the CR has not yet received this information (HoA <-> network prefix), it SHOULD perform a re-registration with the HA to verify the claim.

An additional aspect is that the MR MAY use a different CoA for different CRs (and the HA). This is useful in situations where the network provides only partial-mesh connectivity and specific interfaces must be used to reach specific destinations. In addition, this allows for load balancing.

3.3.1. Triggering Route Optimization

Since each MR knows the eligible route-optimizable networks, the route optimization between all CRs can be established at any time; however, a better general practice is to conduct route optimization only on demand. It is RECOMMENDED that route optimization be started only when sending a packet that originates from a local managed network (and if the network is registered as route optimizable) and whose destination address falls within the network prefixes of the Route Optimization Cache. With a small number of MRs, such on-demand behavior may not be necessary, and full-mesh route optimization may be in place constantly.

3.3.2. Mobile Router Routing Tables

Each MR maintains a routing table. In a typical situation, the MR has one or more interface(s) to the local networks, one or more interface(s) to wide-area networks (such as those provided by ISPs), and a tunnel interface to the HA. Additional tunnel interfaces become activated as route optimization is being performed.

The routing table SHOULD typically contain network prefixes managed by CRs associated with established route-optimized tunnel interfaces. A default route MAY point to the reverse tunnel to the HA if not overridden by prefix information. The routing table MAY also include additional routes if required by the tunneling implementation.

The routes for the HoAs of any CRs SHOULD also be pointing toward their respective tunnels that are using the optimized path.

If two prefixes overlap each other, e.g., 192.0.2.128/25 and 192.0.2.128/29, the standard longest-match rule for routing is in effect. However, overlapping private addresses SHOULD be considered an error situation. Any aggregation for routes in private address space SHOULD be conducted only at the HA.

3.3.3. Inter-Mobile Router Registration

If route optimization between an MR and a CR is desired, either the RR procedure must have been performed (see Section 3.2), or the KRm must be pre-shared between the MR and the CR. If either condition applies, an MR MAY send a Registration Request to the CR's HoA from the desired interface.

The Registration Request's Source Address and Care-of Address fields are set to the address of the desired outgoing interface on the MR. The address MAY be the same as the CoA used with the HA. The Home Agent field is set to the HA of the MR. The Registration Request MUST be sent to (have a destination address of) the HoA of the CR. The Registration Request MUST include a Mobile-Correspondent Authentication Extension (defined in Section 5.3) and SHOULD include a Mobile Network Request Extension (defined in [RFC5177]). If present, the Mobile Network Request Extension MUST contain the network prefixes, as if registering in explicit mode. If timestamps are used, the CR MUST check the Identification field for validity. The Authenticator field is hashed with the KRm.

The CR replies to the request with a Registration Reply. The Registration Reply MUST include a Mobile-Correspondent Authentication Extension (defined in Section 5.3) and, if a Mobile Network Request Extension was present in the request, a Mobile Network Acknowledgement Extension.

The encapsulation can be set as desired, except in the case where the Route Optimization Cache Entry has NAT entries for the CR, or the MR itself is known to be behind a NAT or firewall. If either condition applies, the Registration Request MUST specify UDP encapsulation. It is RECOMMENDED that UDP encapsulation always be used to facilitate detection of path failures via a keepalive mechanism.

The CR first checks the Registration Request's authentication against Kcr and nonce indexes negotiated during the RR procedure. This ensures that the Registration Request is coming from a valid MR. If the check fails, an appropriate Registration Reply code is sent (see Section 10). If the failure is due to the nonce index expiring, the MR sets RRSTATE for the CR to INACTIVE. The RR procedure MAY then be initiated again.

If the check passes, the CR MUST then check its Route Optimization Cache to determine whether the MR exists and is associated with the prefixes included in the request (i.e., whether prefixes are present

and the 'HA' flag is true for each prefix). Note that the viewpoint is always local; the CR compares CR-HoA entries against the MR's HoA -- from the CR's perspective, the MR is also a "correspondent router".

If the check against the cache fails, the CR SHOULD send a re-Registration Request to the HA with the 'S' (solicitation) bit set, thus obtaining the latest information on network prefixes managed by the incoming MR. If, even after this update, the prefixes still don't match, the reply's Mobile Network Acknowledgement code MUST be set to "MOBNET_UNAUTHORIZED". The registration MAY also be rejected completely. This verification is done to protect against MRs claiming to represent arbitrary networks; however, since the HA is assumed to provide trusted information, it can authorize the MR's claim. If the environment itself is considered trusted, the CR can, as a policy, accept registrations without this check; however, this is NOT RECOMMENDED as a general practice.

If the prefixes match, the CR MAY accept the registration. If the CR chooses to accept, the CR MUST check to determine if a tunnel to the MR already exists. If the tunnel does NOT exist or has wrong endpoints (CoAs), a new tunnel MUST be established and the Route Optimization Cache updated. The reply MUST include a list of eligible CoAs (see Section 5.4) with which the MR may establish a tunnel. The reply MUST also include the Mobile-Correspondent Authentication Extension (see Section 5.3).

Upon receiving the Registration Reply, the MR MUST check to determine if a tunnel to the CR already exists. If the tunnel does NOT exist or has wrong endpoints (CoAs), a new tunnel MUST be established and the Route Optimization Cache updated. This is covered in detail in Section 3.3.4.

The CR's routing table MUST be updated to indicate that the MR's networks are reachable via the direct tunnel to the MR.

After the tunnel is established, the MR MAY update its routing tables to reach all of the CR's Prefixes via the tunnel, although it is RECOMMENDED that time be given for the CR to perform its own, explicit registration. This is primarily a policy decision, depending on the network environment. See Section 6.5.

Due to the fact that the route optimization procedures may occur concurrently at both MRs, each working as each other's CR, there may be a situation where two routers are attempting to establish separate tunnels between them at the same time. If a router with a smaller HoA (meaning a normal 32-bit integer comparison treating IPv4 addresses as 32-bit unsigned integers) receives a Registration

Request (in the CR role) while its own Registration Request (sent in the MR role) is pending, the attempt should be accepted with reply code "concurrent registration" (Value 2). If receiving such an indication, the recipient SHOULD consider the registration a success but only act on it once the peer has completed its own registration.

3.3.4. Inter-Mobile Router Tunnels

Inter-MR tunnel establishment follows establishing standard reverse tunnels to the HA. The Registration Request to the CR includes information on the desired encapsulation. It is RECOMMENDED that UDP encapsulation be used. In the cases of Generic Router Encapsulation (GRE) [RFC2784], IP over IP [RFC2003], or minimal encapsulation [RFC2004], no special considerations regarding reachability are necessary. The tunnel has no stateful information; the packets are simply encapsulated within the GRE, IP, or minimal header.

The tunnel origination point for the CR is its CoA, not the HoA where the Registration Requests were sent. This is different from the creation of the reverse tunnel to the HA, which reuses the channel from registration signaling.

Special considerations rise from using UDP encapsulation, especially in cases where one of the MRs is located behind a NAT or firewall. A deviation from RFC 3519 [RFC3519] is that keepalives should be sent from both ends of the tunnel to detect path failures after the initial keepalive has been sent -- this allows both the MR and CR to detect path failures.

The initial UDP keepalive SHOULD be sent by the MR. Only after the first keepalive is successfully completed SHOULD the tunnel be considered eligible for traffic. If a reply to the initial keepalive is not received, the MR may opt to attempt sending the keepalive to other CoAs provided by the Registration Reply to check whether they provide better connectivity; or, if all of these fail, the MR may perform a re-registration via an alternative interface, or deregister completely. See Section 6.1. Once the initial keepalive packet has reached the CR and a reply has been sent, the CR MAY start sending its own keepalives.

The original specification for UDP encapsulation suggests a keepalive interval default of 110 seconds. However, to provide fast response time and switching to alternate paths, it is RECOMMENDED, if power and other constraints allow, that considerably shorter periods be used, adapting to the perceived latency as needed. However, the maximum amount of keepalives SHOULD at no point exceed

MAX_UPDATE_RATE times per second. The purpose of the keepalive is not to keep NAT or firewall mappings in place but to serve as a mechanism to provide fast response in case of path failures.

If both the MR and the CR are behind separate NATs, route optimization cannot be performed between them. Possible ways to set up mutual tunneling when both routers are behind NATs are outside the scope of this document. However, some of these issues are addressed in Section 6.1.

The designations "MR" and "CR" only apply to the initial tunnel establishment phase. Once a tunnel is established between two routers, either of them can opt to either tear down the tunnel or perform a handover. Signaling messages have to be authenticated with a valid KRm.

3.3.5. Constructing Route-Optimized Packets

All packets received by the MR are forwarded using normal routing rules according to the routing table. There are no special considerations when constructing the packets; the tunnel interface's own processes will encapsulate any packet automatically.

3.3.6. Handovers and Mobile Routers Leaving Network

Handovers and connection breakdowns can be categorized as either ungraceful or graceful, also known as "break-before-make" (bbm) and "make-before-break" (mbb) situations.

As with establishment, the "mobile router" discussed here is the router wishing to change connectivity state, with the "correspondent router" being the peer.

When an MR wishes to join its home link, it SHOULD, in addition to sending the Registration Request to the HA with lifetime set to zero, also send such a request to all known CRs, to their HoAs. The CR(s), upon accepting this request and sending the reply, will check whether the Route Optimization Cache contains any prefixes associated with the requesting MR. These entries should be removed and the routing table updated accordingly (traffic for the prefixes will be forwarded via the HA again). The tunnel MUST then be destroyed. A short grace period SHOULD be used to allow possible in-transit packets to be received correctly.

In the case of a handover, the CR simply needs to update the tunnel's destination to the MR's new CoA. The MR SHOULD keep accepting packets from both old and new CoAs for a short grace period, typically on the order of ten seconds. In the case of UDP

encapsulation, it is RECOMMENDED that the same port numbers be used for both registration signaling and tunneled traffic, if possible. The initial keepalive message sent by the MR will verify that direct connectivity exists between the MR and CR -- if the keepalive fails, the MR SHOULD attempt alternate paths.

If the MR was unable to send the re-Registration Request before handover, it MUST send it immediately after handover has been completed and a tunnel with the HA is established. Since the changing of CoA(s) invalidates the KRm, it is RECOMMENDED that partial return routability be conducted by sending a CoTI message via the new CoA and obtaining a new care-of keygen token. In all cases, necessary tokens also have to be acquired if the existing tokens have expired.

If a reply is not received for a Registration Request to a CR, any routes to the network prefixes managed by the CR MUST be removed from the routing table, thus causing the user traffic to be forwarded via the HA.

3.4. Convergence and Synchronization Issues

The information the HA maintains on mobile network prefixes and the MRs' Route Optimization Caches does not need to be explicitly synchronized. This is based on the assumption that at least some of the traffic between nodes inside mobile networks is always bidirectional. If using on-demand route optimization, this also implies that when a node in a mobile network talks to a node in another mobile network, if the initial packet does not trigger route optimization, the reply packet will.

Consider a situation with three mobile networks, A, B, and C, handled by three mobile routers, MR A, MR B, and MR C, respectively. If they register with an HA in this order, the situation goes as follows:

MR A registers and receives no information on other networks from the HA, as no other MR has registered yet.

MR B registers and receives information on mobile network A being reachable via MR A.

MR C registers and receives information on both of the other mobile networks.

If a node in mobile network C is about to send traffic to mobile network A, the route optimization is straightforward; MR C already has network A in its Route Optimization Cache. Thus, packet transmission triggers route optimization toward MR A. When MR C

registers with MR A (after the RR procedure is completed), MR A does not have information on mobile network C; thus, it will perform a re-registration with the HA on demand. This allows MR A to verify that MR C is indeed managing network C.

If a node in mobile network B sends traffic to mobile network C, MR B has no information on network C. No route optimization is triggered. However, when the node in network C replies and the reply reaches MR C, route optimization happens as above. Further examples of signaling are in Section 8.

Even in the very rare case of completely unidirectional traffic from an entire network, re-registrations with the HA caused by timeouts will eventually cause convergence. However, this should be treated as a special case.

Note that all MRs are connected to the same HA. For possibilities concerning multiple HAs, see Section 6.4.

4. Data Compression Schemes

This section defines the two compression formats used in Route Optimization Prefix Advertisement Extensions.

4.1. Prefix Compression

Prefix compression is based on the idea that prefixes usually share common properties. The scheme is simple delta compression. In the prefix information advertisement (Section 5.5), the 'D' bit indicates whether receiving a "master" or a "delta" prefix. This, combined with the Prefix Length information, allows for compression and decompression of prefix information.

If D = 0, what follows in the "Prefix" field are bits 1..n of the new master prefix, where n is PLen. This is rounded up to the nearest full octet. Thus, prefix lengths of /4 and /8 take 1 octet, /12 and /16 take 2 octets, /20 and /24 take 3 octets, and longer prefix lengths take a full 4 octets.

If D = 1, what follows in the "Prefix" field are bits m..PLen of the prefix, where m is the first changed bit of the previous master prefix, with padding from the master prefix filling the field to a full octet. The maximum value of PLen - m is 8 (that is, the delta MUST fit into one octet). If this is not possible, a new master prefix has to be declared. If the prefixes are equal -- for example, in the case where the same prefix appears in multiple realms -- then one octet is still encoded, consisting completely of padding from the master prefix.

Determining the order of prefix transmission should be based on saving maximum space during transmission.

An example of compression and transmitted data, where network prefixes 192.0.2.0/28, 192.0.2.64/26, and 192.0.2.128/25 are transmitted, is illustrated in Figure 1. Because of the padding to full octets, redundant information is also sent. The bit patterns being transmitted are as follows:

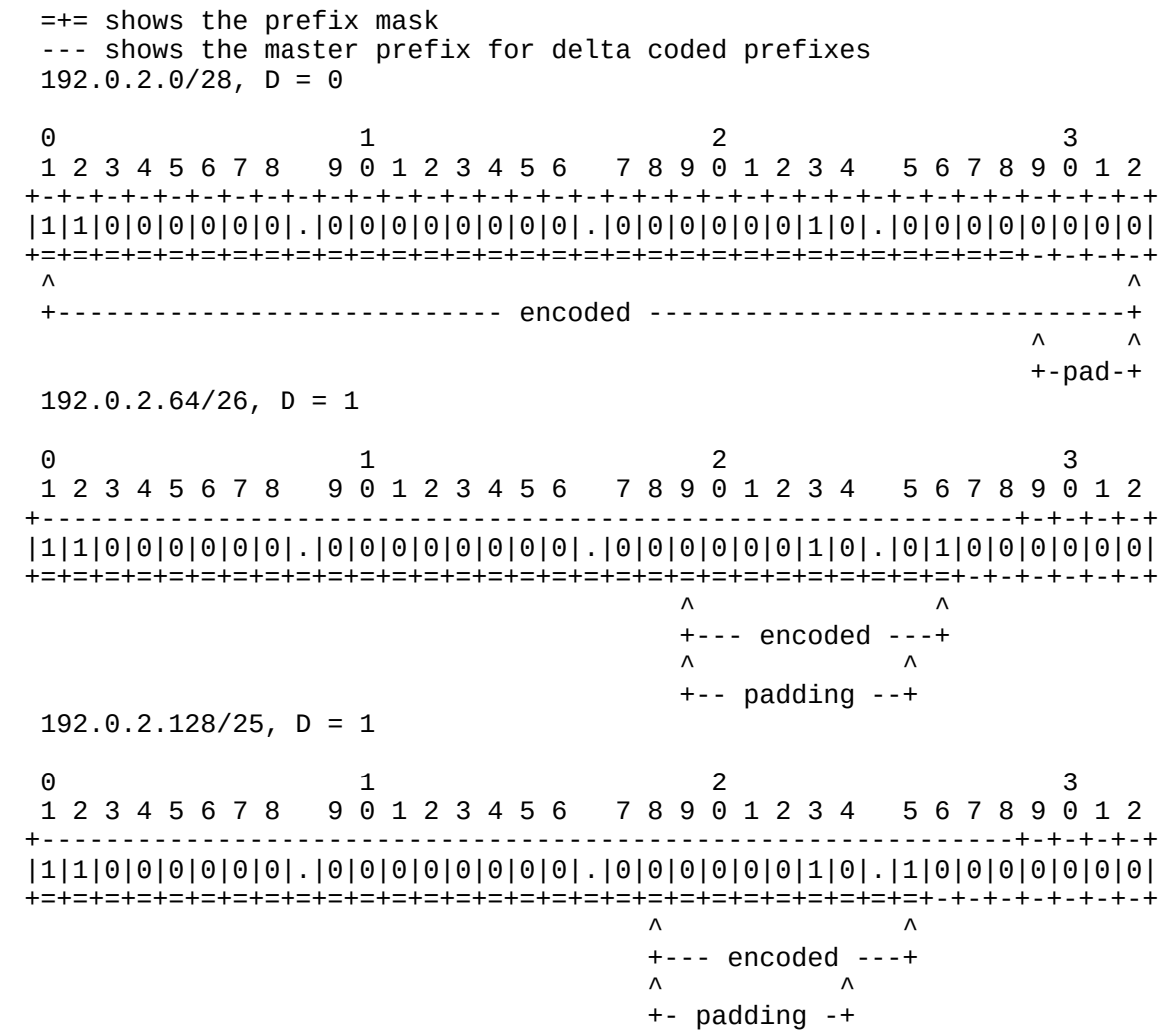


Figure 1: Prefix Compression Example

The first prefix, 192.0.2.0/28, is considered a master prefix and is transmitted in full. The PLen of 28 bits determines that all four octets must be transmitted. If the prefix would have been, e.g., 192.0.2.0/24, three octets would have sufficed, since 24 bits fit into 3 octets.

For the following prefixes, D = 1. Thus, they are deltas of the previous prefix, where D was zero.

192.0.2.64/26 includes bits 19-26 (full octet). Bits 19-25 are copied from the master prefix, but bit 26 is changed to 1. The final notation in binary is "1001", or 0x09.

192.0.2.128/25 includes bits 18-25 (full octet). Bits 18-24 are copied from the master prefix, but bit 25 is changed to 1. The final notation in binary is "101", or 0x05.

The final encoding thus becomes

Prefix	PLen	D	Transmitted Prefix
192.0.2.0/28	28	0	0xc0 0x00 0x02 0x00
192.0.2.64/26	26	1	0x09
192.0.2.128/25	25	1	0x05

It should be noted that in this case the order of prefix transmission would not affect compression efficiency. If prefix 192.0.2.128/25 would have been considered the master prefix and the others as deltas instead, the resulting encoding still fits into one octet for the subsequent prefixes. There would be no need to declare a new master prefix.

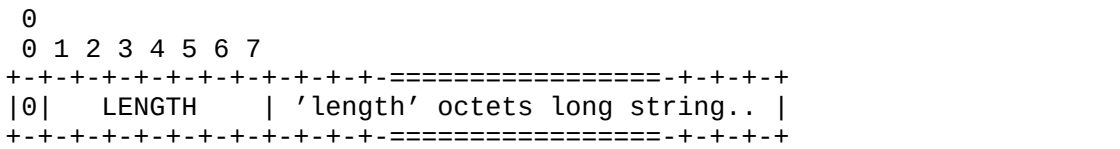
4.2. Realm Compression

4.2.1. Encoding of Compressed Realms

In order to reduce the size of messages, the system introduces a realm compression scheme, which reduces the size of realms in a message. The compression scheme is a simple dynamically updated dictionary-based algorithm, which is designed to compress text strings of arbitrary length. In this scheme, an entire realm, a single label, or a list of labels may be replaced with an index to a previous occurrence of the same string stored in the dictionary. The realm compression defined in this specification was inspired by the RFC 1035 [RFC1035] DNS domain name label compression scheme. Our algorithm is, however, improved to gain more compression.

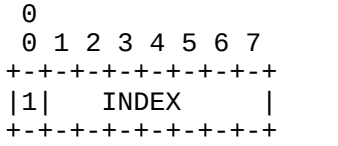
When compressing realms, the dictionary is first reset and does not contain a single string. The realms are processed one by one, so the algorithm does not expect to see them all or the whole message at once. The state of the compressor is the current content of the dictionary. The realms are compressed label by label or as a list of labels. The dictionary can hold a maximum of 128 strings; after that, a rollover MUST occur, and existing contents will be overwritten. Thus, when adding the 129th string into the dictionary, the first entry of the dictionary MUST be overwritten, and the index of the new string will become 0.

The encoding of an index to the dictionary or an uncompressed run of octets representing a single label has purposely been made simple, and the whole encoding works on an octet granularity. The encoding of an uncompressed label takes the form of one octet as follows:



This encoding allows label lengths from 1 to 127 octets. A label length of zero (0) is not allowed. The "label length" tag octet is then followed by up to 127 octets of the actual encoded label string.

The index to the dictionary (the "label index" tag octet) takes the form of one octet as follows:



The above encodings do not allow generating an output octet value of zero (0). The encapsulating Mobile IPv4 extension makes use of this property and uses the value of zero (0) to mark the end of the compressed realm or to indicate an empty realm. It is also possible to encode the complete realm using only "label length" tags. In this case, no compression takes place. This allows the sender to skip compression -- for example, to reduce computation requirements when generating messages. However, the receiver MUST always be prepared to receive compressed realms.

4.2.2. Searching Algorithm

When compressing the input realm, the dictionary is searched for a matching string. If no match could be found, the last label is removed from the right-hand side of the used input realm. The search is repeated until the whole input realm has been processed. If no match was found at all, then the first label of the original input realm is encoded using the "label length" tag, and the label is inserted into the dictionary. The previously described search is repeated with the remaining part of the input realm, if any. If nothing remains, the realm encoding is complete.

When a matching string is found in the dictionary, the matching part of the input realm is encoded using the "label index" tag. The matching part of the input realm is removed, and the search is repeated with the remaining part of the input realm, if any. If nothing remains, the octet value of zero (0) is inserted to mark the end of the encoded realm.

The search algorithm also maintains the "longest non-matching string" for each input realm. Each time the search in the dictionary fails and a new label gets encoded using the "label length" tag and inserted into the dictionary, the "longest non-matching string" is concatenated by this label, including the separating "." (dot, i.e., hexadecimal 0x2e). When a match is found in the dictionary, the "longest non-matching string" is reset (i.e., emptied). Once the whole input realm has been processed and encoded, all possible suffixes longer than one label are taken from the string and inserted into the dictionary.

4.2.3. Encoding Example

This section shows an example of how to encode a set of realms using the specified realm compression algorithm. For example, a message might need to compress the realms "foo.example.com", "bar.foo.example.com", "buz.foo.example.org", "example.com", and "bar.example.com.org". The following example shows the processing of input realms on the left-hand side and the contents of the dictionary on the right-hand side. The example uses hexadecimal representation of numbers.

```
COMPRESSOR:                                  DICTIONARY:

1) Input "foo.example.com"
Search("foo.example.com")
Search("foo.example")
Search("foo")
Encode(0x03,'f','o','o')                     0x00 "foo"
  +-> "longest non-matching string" = "foo"
Search("example.com")
Search("example")
Encode(0x07,'e','x','a','m','p','l','e')     0x01 "example"
  +-> "longest non-matching string" = "foo.example"
Search("com")
Encode(0x03,'c','o','m')                     0x02 "com"
  +-> "longest non-matching string" = "foo.example.com"
                                             0x03 "foo.example.com"
                                             0x04 "example.com"
Encode(0x00)

2) Input "bar.foo.example.com"
Search("bar.foo.example.com")
Search("bar.foo.example")
Search("bar.foo")
Search("bar")
Encode(0x03,'b','a','r')                     0x05 "bar"
  +-> "longest non-matching string" = "bar"
Search("foo.example.com") -> match to 0x03
Encode(0x83)
  +-> "longest non-matching string" = NUL
Encode(0x00)
```

```
3) Input "buz.foo.example.org"
Search("buz.foo.example.org")
Search("buz.foo.example")
Search("buz.foo")
Search("buz")
Encode(0x03,'b','u','z')                    0x06 "buz"
  +-> "longest non-matching string" = "buz"
Search("foo.example.org")
Search("foo.example")
Search("foo") -> match to 0x00
Encode(0x80)
  +-> "longest non-matching string" = NUL
Search("example.org")
Search("example") -> match to 0x01
Encode(0x81)
  +-> "longest non-matching string" = NUL
Search("org")
Encode(0x03,'o','r','g')                    0x07 "org"
  +-> "longest non-matching string" = "org"
Encode(0x00)

4) Input "example.com"
Search("example.com") -> match to 0x04
Encode(0x84)
Encode(0x00)

5) Input "bar.example.com.org"
Search("bar.example.com.org")
Search("bar.example.com")
Search("bar.example")
Search("bar") -> match to 0x05
Encode(0x85)
Search("example.com.org")
Search("example.com") -> match to 0x04
Encode(0x84)
Search("org") -> match to 0x07
Encode(0x87)
Encode(0x00)
```

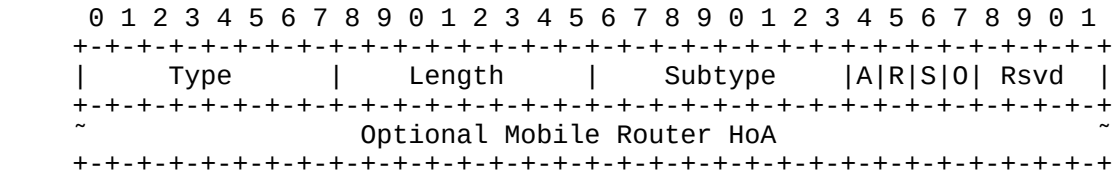
As can be seen from the example, due to the greedy approach of encoding matches, the search algorithm and the dictionary update function are not the most optimal. However, we do not claim that the algorithm would be the most efficient. It functions efficiently enough for most inputs. In this example, the original input realm data was 79 octets, and the compressed output, excluding the end mark, is 35 octets.

5. New Mobile IPv4 Messages and Extensions

This section describes the construction of all new information elements.

5.1. Mobile Router Route Optimization Capability Extension

This skippable extension MAY be sent by an MR to an HA in the Registration Request message.



Type 153 (skippable); if the HA does not support route optimization advertisements, it can ignore this request and simply not include any information in the reply. "short" extension format.

Subtype 1

Reserved Set to zero; MUST be ignored on reception.

A Advertise my networks. If the 'A' bit is set, the HA is allowed to advertise the networks managed by this MR to other MRs. This also indicates that the MR is capable of receiving route optimization Registration Requests. In effect, this allows the MR to work in the CR role.

R Request mobile network information. If the 'R' bit is set, the HA MAY respond with information about mobile networks in the same domain.

S Solicit prefixes managed by a specific MR. The MR is specified in the Optional Mobile Router HoA field.

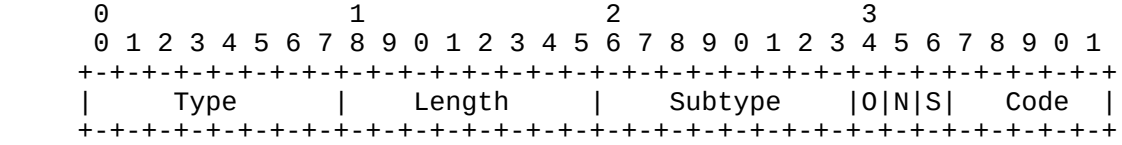
O Explicitly specify that the requesting router is only able to initiate outgoing connections and not accept any incoming connections, due to a NAT device, stateful firewall, or similar issue on any interface. This is reflected by the HA in the reply and distributed in Prefix Advertisements to other MRs.

Optional Mobile Router HoA

Solicited mobile router's home address. This field is only included if the 'S' flag is set.

5.2. Route Optimization Reply

This non-skippable extension MUST be sent by an HA to an MR in the Registration Reply message, if the MR indicated support for route optimization in the registration message and the HA supports route optimization.



Type 49 (non-skippable); "short" extension format.

Subtype 1

O The 'O' flag in the Mobile Router Route Optimization Capability Extension was set during registration.

N NAT was detected by the HA. This informs the MR that it is located behind a NAT. The detection procedure is specified in RFC 3519 [RFC3519] and is based on the discrepancy between the registration packet's source address and indicated CoA. The MR can use this information to make decisions about route optimization strategy.

S Responding to a solicitation. If the 'S' bit was present in the MR's Route Optimization Capability Extension (Section 5.1), this bit is set; otherwise, it is unset.

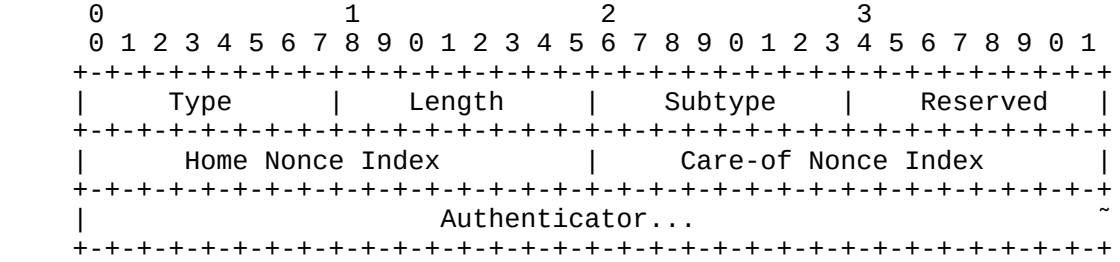
The Reply code indicates whether route optimization has been accepted. Values of 0..15 indicate assent, and values 16..63 indicate that route optimization is not done.

0 Will do route optimization.

16 Route optimization declined; reason unspecified.

5.3. Mobile-Correspondent Authentication Extension

The Mobile-Correspondent Authentication Extension is included in Registration Requests sent from the MR to the CR. The existence of this extension indicates that the message is not destined to an HA, but another MR. The format is similar to the other authentication extensions defined in [RFC5944], with Security Parameter Indexes (SPIs) replaced by nonce indexes.



The Home Nonce Index field tells the CR which nonce value to use when producing the home keygen token. The Care-of Nonce Index field is ignored in requests to remove a binding. Otherwise, it tells the CR which nonce value to use when producing the care-of keygen token. If using a pre-shared key (KRm), the indexes may be set to zero and are ignored on reception.

Type 49 (non-skippable); "short" extension format.

Subtype 2

Reserved Set to zero; MUST be ignored on reception.

Home Nonce Index

Home Nonce Index in use. If using a pre-shared KRm, set to zero and ignored on reception.

Care-of Nonce Index

Care-of Nonce Index in use. If using a pre-shared KRm, set to zero and ignored on reception.

Authenticator

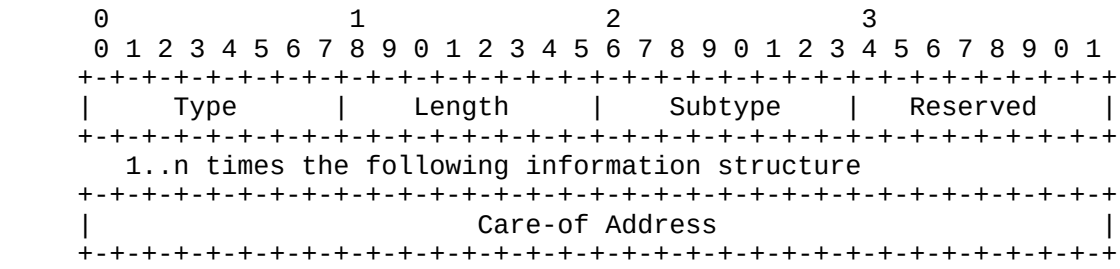
Authenticator field, by default constructed with First (128, HMAC_SHA1 (KRm, Protected Data)).

The protected data, just like in other cases where the Authenticator field is used, consists of

- the UDP payload (i.e., the Registration Request or Registration Reply data),
- all prior extensions in their entirety, and
- the Type, Length, Home Nonce Index, and Care-of Nonce Index of this extension.

5.4. Care-of Address Extension

The Care-of Address Extension is added to a Registration Reply sent by the CR to inform the MR of the upcoming tunnel endpoint.



Type 49 (non-skippable); "short" extension format.

Length Total length of the packet. When processing the information structures, if Length octets have been reached, this is an indication that the final information structure was reached as well.

Subtype 3

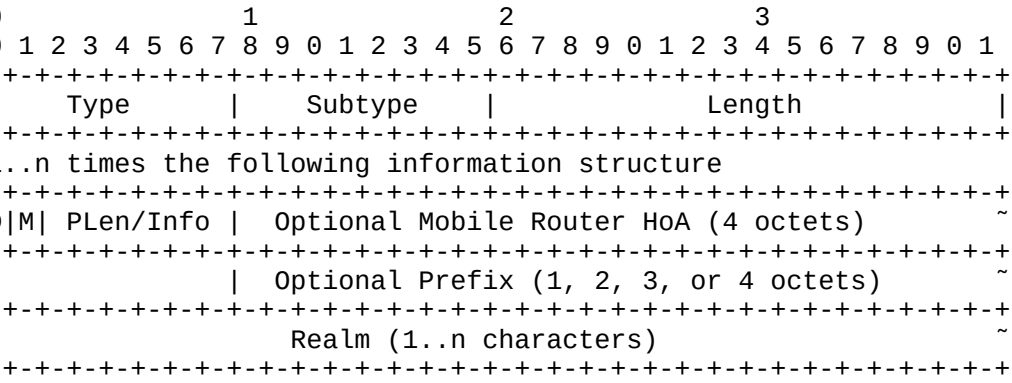
Care-of Address

Care-of address(es) that may be used for a tunnel with the MR, in order of priority. Multiple CoAs MAY be listed to facilitate faster NAT traversal processing.

5.5. Route Optimization Prefix Advertisement Extension

This non-skippable extension MAY be sent by an HA to an MR in the Registration Reply message. This extension is only included when explicitly requested by the MR in the Registration Request message, setting the 'R' flag of the Mobile Router Route Optimization Capability Extension. Implicit prioritization of prefixes is caused by the order of extensions.

The extension contains a sequence of information structures. An information structure may consist of either an MR HoA or a network prefix. Any network prefixes following an MR HoA are owned by that MR. An MR HoA MUST be first in the sequence, since one cannot have prefixes without an MR.



Type 50 (non-skippable); "long" extension format.

Subtype 1

Length Total length of the packet. When processing the information structures, if Length octets have been reached, this is an indication that the final information structure was reached as well.

D Delta. If D = 1, the prefix is a delta from the last Prefix, where D = 0. MUST be zero on the first information structure containing a Prefix; MAY be zero or one on subsequent information structures. If D = 1, the Prefix field is one octet in length. See Section 4.1 for details.

M Mobile Router HoA bit. If M = 1, the next field is Mobile Router HoA, and Prefix and Realm are omitted. If M = 0, the next field is Prefix followed by Realm, and Mobile Router HoA is omitted. For the first information structure, M MUST be set to 1. If M = 1, the 'D' bit is set to zero and ignored upon reception.

PLen/Info

This field is interpreted differently, depending on whether the 'M' bit is set or not. If M = 0, the field is considered to be the PLen field, and the contents indicate the length of the advertised prefix. The 6 bits allow for values from 0 to 63, of which 33-63 are illegal. If M = 1, the field is considered to be the Info field. Permissible values are 0 to indicate no specific information, or 1 to indicate "outbound connections only". This indicates that the target MR can only initiate, not receive, connections on any of its interfaces (apart from the reverse tunnel to the HA). This is set if the MR has explicitly requested it via the 'O' flag in the Mobile Router Route Optimization Capability Extension (Section 5.1).

Mobile Router HoA

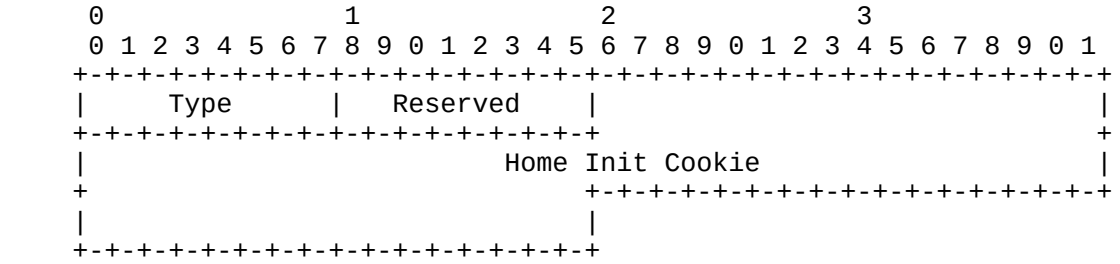
The mobile router's home address. All prefixes in the following information structures where M = 0 are maintained by this MR. This field is present only when M = 1.

Prefix The IPv4 prefix advertised. If D = 0, the field length is PLen bits, rounded up to the nearest full octet. Least-significant bits starting off PLen (and that are zeros) are omitted. If D = 1, the field length is one octet. This field is present only when M = 0.

Realm The Realm that is associated with the advertised Mobile Router HoA and prefix. If empty, MUST be set to '\0'. For realm encoding and an optional compression scheme, refer to Section 4.2. This field is present only when M = 0.

5.6. Home Test Init Message

This message is sent from the MR to the CR when performing the RR procedure. The source and destination IP addresses are set to the MR's HoA and the CR's HoA, respectively. The UDP source port MAY be randomly chosen. The UDP destination port is 434.



Type 24

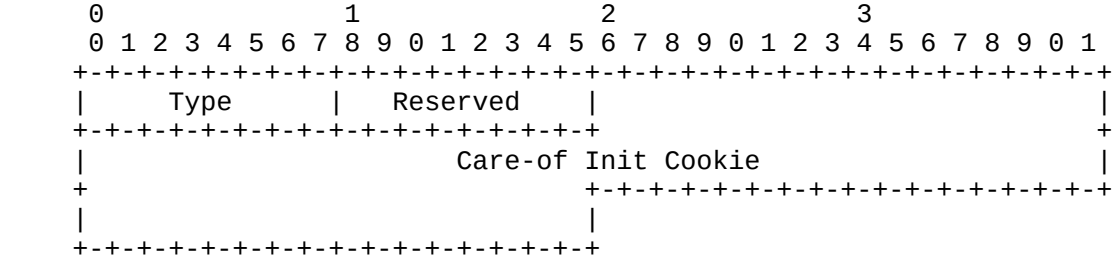
Reserved Set to zero; MUST be ignored on reception.

Home Init Cookie

64-bit field that contains a random value, the Home Init Cookie.

5.7. Care-of Test Init Message

This message is sent from the MR to the CR when performing the RR procedure. The source and destination IP addresses are set to the MR's CoA and the CR's HoA, respectively. The UDP source port MAY be randomly chosen. The UDP destination port is 434.



Type 25

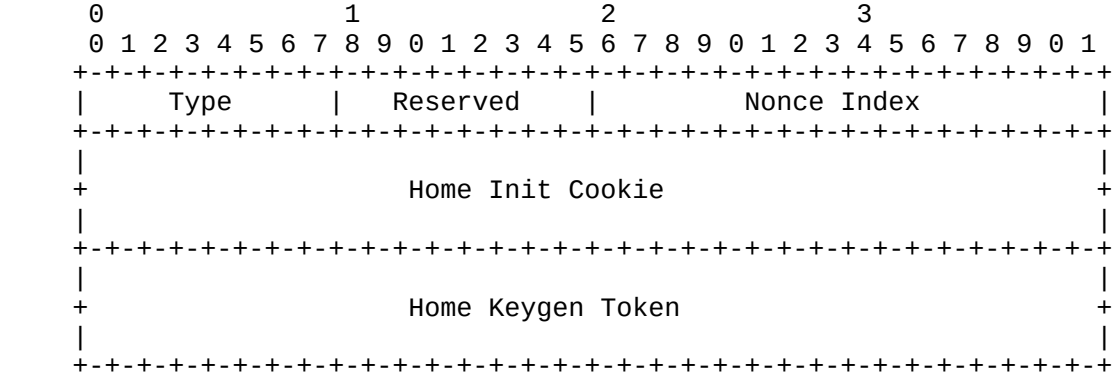
Reserved Set to zero; MUST be ignored on reception.

Care-of Init Cookie

64-bit field that contains a random value, the Care-of Init Cookie.

5.8. Home Test Message

This message is sent from the CR to the MR when performing the RR procedure as a reply to the Home Test Init message. The source and destination IP addresses, as well as UDP ports, are the reverse of those in the Home Test Init message for which this message is constructed. As such, the UDP source port is always 434.



Type 26

Reserved Set to zero; MUST be ignored on reception.

Nonce Index

This field will be echoed back by the MR to the CR in a subsequent Registration Request's authentication extension.

Home Init Cookie

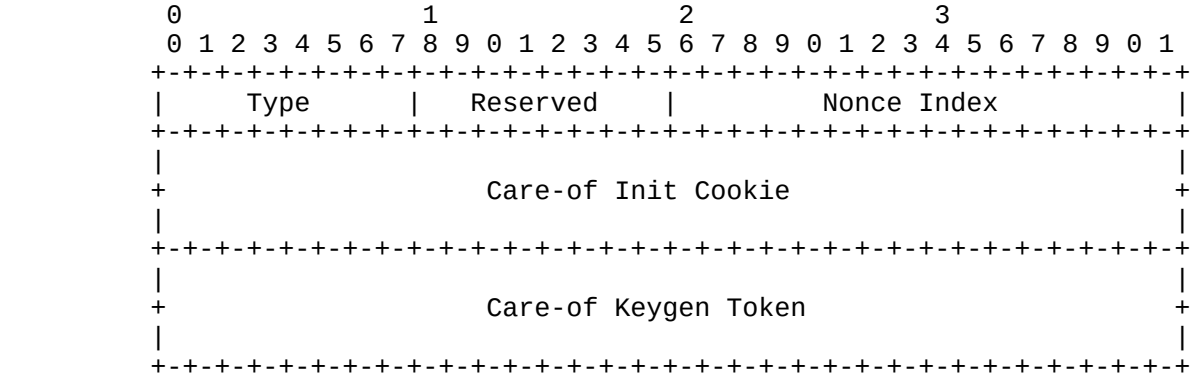
64-bit field that contains a random value, the Home Init Cookie.

Home Keygen Token

This field contains the 64-bit home keygen token used in the RR procedure. Generated from cookie + nonce.

5.9. Care-of Test Message

This message is sent from the CR to the MR when performing the RR procedure as a reply to the Care-of Test Init message. The source and destination IP addresses, as well as UDP ports, are the reverse of those in the Care-of Test Init message for which this message is constructed. As such, the UDP source port is always 434.



Type 27

Reserved Set to zero; MUST be ignored on reception.

Care-of Nonce Index

This field will be echoed back by the MR to the CR in a subsequent Registration Request's authentication extension.

Care-of Init Cookie

64-bit field that contains a random value, the Care-of Init Cookie.

Care-of Keygen Token

This field contains the 64-bit care-of keygen token used in the RR procedure. Generated from cookie + nonce.

6. Special Considerations

6.1. NATs and Stateful Firewalls

Mechanisms described in Mobile IP NAT traversal [RFC3519] allow the HA to work with MRs situated behind a NAT device or a stateful firewall. Furthermore, the HA may also detect whether a NAT device is located between the mobile node and the HA. The MR may also explicitly state that it is behind a NAT or firewall on all interfaces, and this information is passed on to the other MRs with the Info field in the Route Optimization Prefix Advertisement Extension (Section 5.5). The HA may also detect NAT and inform the registering MR via the 'N' flag in the Route Optimization Reply Extension (Section 5.2). In the case where one or both of the routers is known to be behind a NAT or is similarly impaired (not able to accept incoming connections), the tunnel establishment procedure needs to take this into account.

In the case where the MR is behind a NAT (or firewall) and the CR is not, the MR will, when the tunnel has been established, send keepalive messages (ICMP echo requests) through the tunnel. Until a reply has been received, the tunnel SHOULD NOT be considered active. Once a reply has been received, NAT mapping is in place, and traffic can be sent.

The source address may change due to NAT in CoTI and Registration Request messages. This does not affect the process -- the hash values are calculated by the translated address, and the Registration Request will also appear from the same translated address.

Unlike communication with the HA, in the case of route optimization, the path used for signaling is not used for tunneled packets, as signaling always uses HoAs, and the MR <-> CR tunnel is from CoA to CoA. It is assumed that even though port numbers may change, NAT processing rarely allocates more than one external IP address to a single internal address; thus, the IP address seen in the Registration Request and tunnel packets remains the same. However, the UDP source port number may be different in the Registration Request and incoming tunnel packets, due to port translation. This must not cause an error situation -- the CR MUST be able to accept tunneling packets from a different UDP source port than what was used in the Registration Request.

Since MRs may have multiple interfaces connecting to several different networks, it might be possible that specific MRs may only be able to perform route optimization using specific CoA pairs, obtained from specific networks -- for example, in a case where two MRs have an interface behind the same NAT. A similar case may be

applicable to nested NATs. In such cases, the MR MAY attempt to detect eligible CoA pairs by performing a registration and attempting to establish a tunnel (sending keepalives) with each CoA listed in the Registration Reply's Care-of Address Extension. The eligible pairs should be recorded in the Route Optimization Cache. If a tunnel cannot be established with any CoAs, the MR MAY attempt to repeat the procedure with alternative interfaces. The above information on network topology can also be configured on the MRs either statically or via some external feedback mechanism.

If both the MR and the CR are behind two separate NATs, some sort of proxy or hole-punching technique may be applicable. This is out of scope for this document.

6.2. Handling of Concurrent Handovers

If both the MR and the CR move at the same time, this causes no issues from the signaling perspective, as all requests are always sent from a CoA to HoAs. Thus, the recipient will always receive the request and can send the reply. This applies even in break-before-make situations where both the MR and the CR get disconnected at the same time -- once the connectivity is restored, one endpoint of the signaling messages is always the HoA of the respective router, and it is up to the HA to provide reachability.

6.3. Foreign Agents

Since foreign agents have been dropped from work related to Network Mobility for Mobile IPv4, they are not considered here.

6.4. Multiple Home Agents

MRs can negotiate and perform route optimization without the assistance of an HA -- if they can discover each other's existence and thus know where to send registration messages. This document only addresses a logically single HA that distributes network prefix information to the MRs. Problems arise from possible trust relationships; in this document, the HA serves as a way to provide verification that a specific network is managed by a specific router.

If route optimization is desired between nodes attached to separate HAs, there are several possibilities. Note that standard high-availability redundancy protocols, such as the Virtual Router Redundancy Protocol (VRRP), can be utilized; however, in such a case, the HA is still a single logical entity, even if it consists of more than a single node.

Several possibilities exist for achieving route optimization between MRs attached to separate HAs, such as a new discovery/probing protocol or routing protocol between HAs or DNS SRV records, or a common Authentication, Authorization, and Accounting (AAA) architecture. There is already a framework for HA to retrieve information from AAA, so it can be considered the most viable possibility. See Section 6.6 for information on a possible way to generalize the method.

Any discovery/probing protocols are out of scope for this document.

6.5. Mutualness of Route Optimization

The procedure as specified is asymmetric; that is, if bidirectional route optimization is desired while maintaining consistency, the route optimization (RR check and registration) has to be performed in both directions, but this is not strictly necessary. This is primarily a policy decision, depending on how often the mobile prefixes are reconfigured.

Consider the case where two networks, A and B, are handled by MRs A and B, respectively. If the routers are set up in such a fashion that route optimization is triggered when the router is forwarding a packet destined to a network prefix in the Route Optimization Cache, the following occurs if a node in network A starts sending ICMP echo requests (ping packets) to a node in network B.

MR A sees the incoming ICMP echo request packet from the local network destined to network B. Since network B exists in MR A’s Route Optimization Cache, the route optimization process is triggered. The original packet is forwarded via the reverse tunnel toward the HA as normal.

MR A completes the RR procedure and registration with MR B, which thus becomes a CR for MR A. A tunnel is created between the routers. MR B updates its routing tables so that network A is reachable via the MR A <-> MR B tunnel.

The traffic pattern is now such that packets from network B to network A are sent over the direct tunnel, but the packets from A to B are transmitted via the HA and reverse tunnels. The echo reply that the node in network B sends toward network A triggers the route optimization at MR B in similar fashion. As such, MR B now performs its own registration toward MR A. Upon completion, MR B notices that a tunnel to MR A already exists, and updates its routing table so that network A is now reachable via the (existing) MR A <-> MR B tunnel. From this point onward, traffic is bidirectional.

In this scenario, if MR A does NOT wait for a separate route optimization process (RR check and registration) from MR B, but instead simply updates its routing table to reach network B via the tunnel, problems may arise if MR B has started to manage another network, B', before the information has been propagated to MR A. The end result is that MR B starts to receive packets from network A to network B' via the HA and to network B via the direct tunnel. If reverse path checking or a similar mechanism is in use on MR B, some of the packets from network A could be black-holed.

Whether to perform this mutual registration or not thus depends on the situation, and whether MRs are going to start managing additional network prefixes during operation.

6.6. Extensibility

The design considerations include several mechanisms that might not be strictly necessary if route optimization were only desired between individual customer sites in a managed network. The registration procedure (with the optional return routability part), which allows CRs to learn an MR's CoAs, is not strictly necessary; the CoAs could have been provided by the HA directly.

However, this approach allows the method to be extended to a more generic route optimization. The primary driver for having an HA to work as a centralized information distributer is to provide MRs with not only the knowledge of the other routers, but with information on which networks are managed by which routers.

The HA provides the information on all feasible nodes with which it is possible to establish route optimization. If representing a whole mobile network is not necessary -- in effect, the typical mobile node <-> correspondent node situation -- the mechanisms in this document work just as well; the only problem is discovering whether the target correspondent node can provide route optimization capability. This can be performed by not including any prefixes in the information extension -- just the HoA of the MR.

In addition, with route optimization for a single node, checks for whether an MR is allowed to represent specific networks are unnecessary, since there are none.

Correspondent node/router discovery protocols (whether they are based on probing or a centralized directory beyond the single HA) are outside the scope of this document.

6.7. Load Balancing

This design simply provides the possibility of creating optimal paths between MRs; it doesn't dictate what the user traffic using these paths should be. One possible approach in helping facilitate load balancing and utilizing all available paths is presented in [MIPv4FLOW], which effectively allows for multiple CoAs for a single HoA. In addition, per-tunnel load balancing is possible by using separate CoAs for separate tunnels.

7. Scalability

Home agent-assisted route optimization scalability issues stem from the general Mobile IPv4 architecture, which is based on tunnels. Creating, maintaining, and destroying tunnel interfaces can cause load on the MRs. However, the MRs can always fall back to normal, reverse-tunneled routing if resource constraints are apparent.

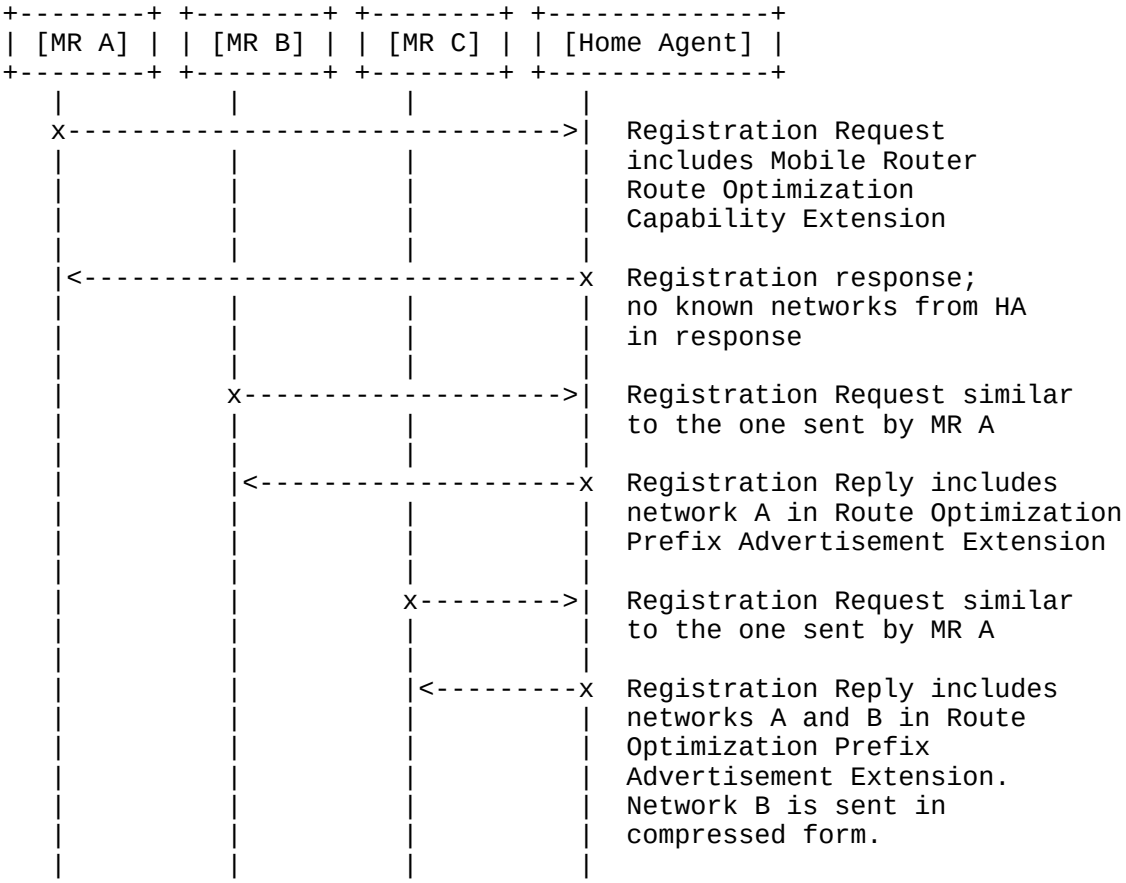
If there are a large number of optimization-capable prefixes, maintaining state for all of these may be an issue also, due to limits on routing table sizes.

Registration responses from the HA to the MR may provide information on a large number of network prefixes. If thousands of networks are involved, the Registration Reply messages are bound to grow very large. The prefix and realm compression mechanisms defined in Section 4 mitigate this problem to an extent. There will, however, be some practical upper limit, after which some other delivery mechanism for the prefix information will be needed.

8. Example Signaling Scenarios

8.1. Registration Request

The following example assumes that there are three mobile routers -- MR A, MR B, and MR C -- each managing network prefixes A, B, and C. At the beginning, no networks are registered with the HA. Any AAA processing at the HA is omitted from the diagram.

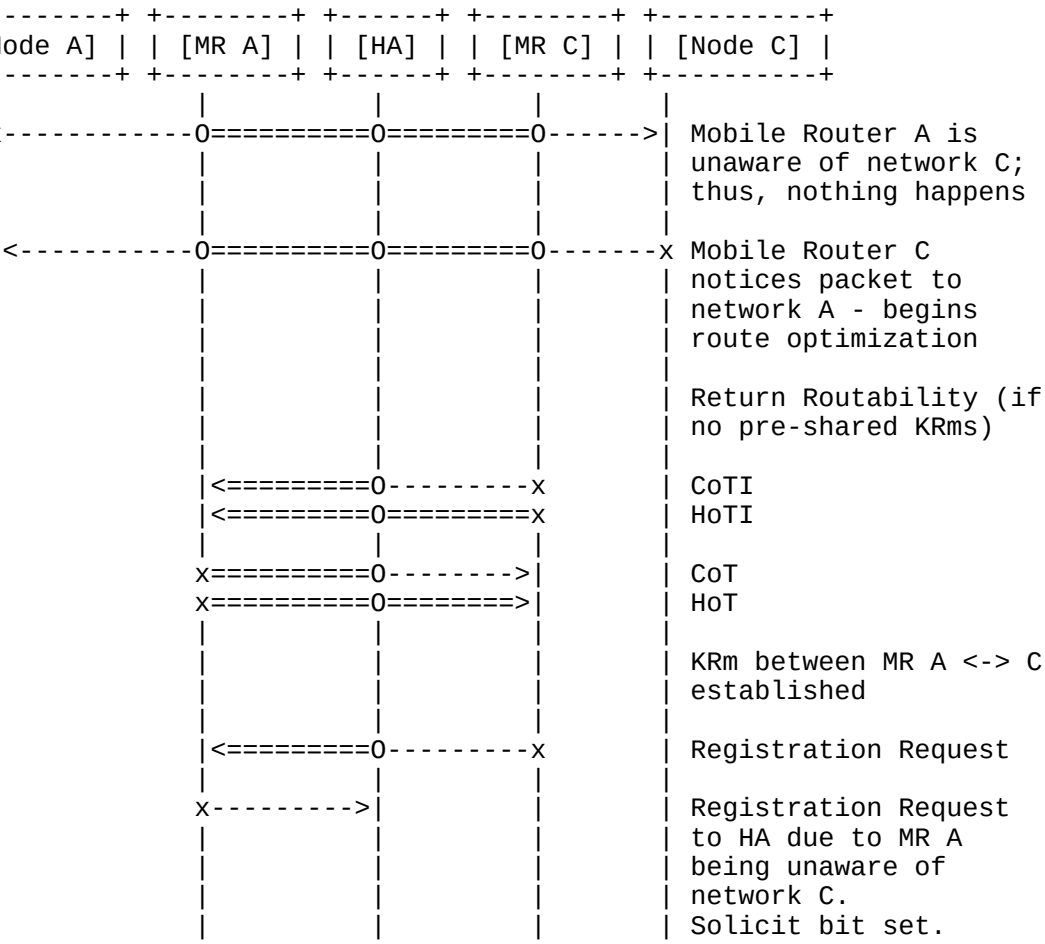


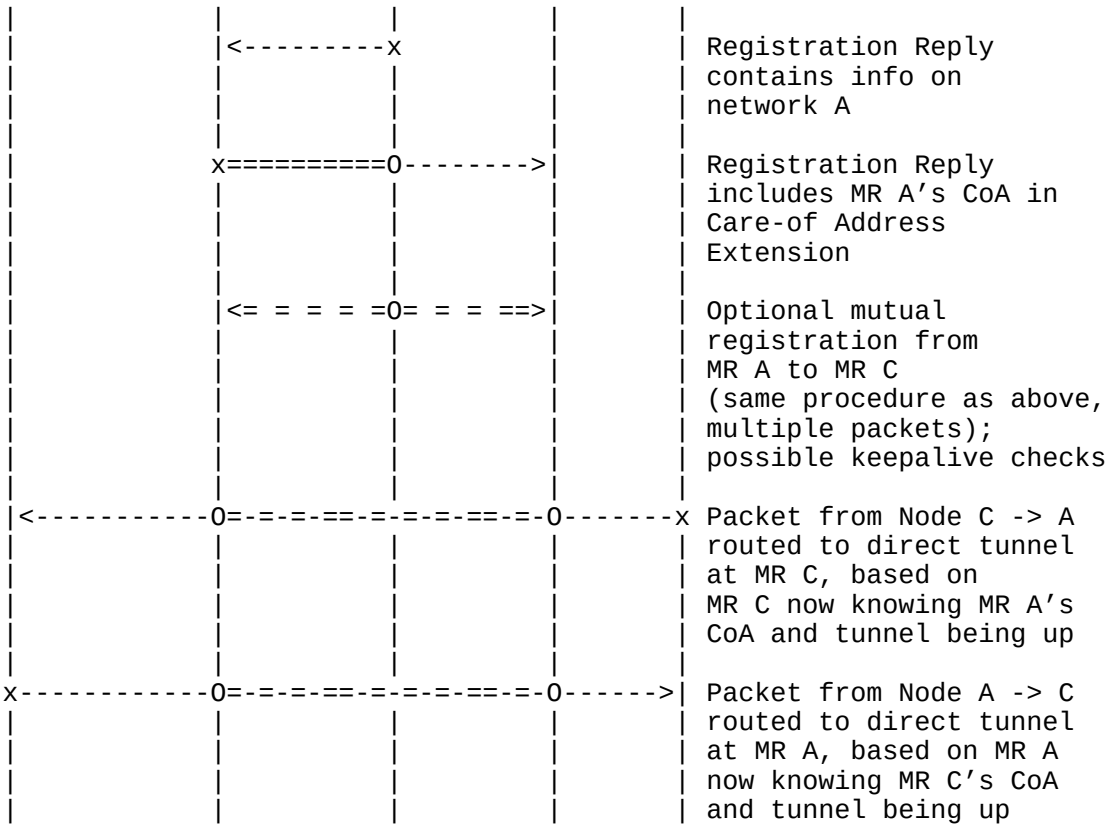
8.2. Route Optimization with Return Routability

The following example has the same network setup as that in Section 8.1 -- three MRs, each corresponding to their respective network. Node A is in network A, and Node C is in network C.

At the beginning, none of the MRs know each other's KRms. If the KRms were pre-shared or provisioned with some other method, the Return Routability messages could be omitted. Signaling as described in Section 8.1 has occurred; thus, MR A is not aware of the other networks, and MR C is aware of networks A and B.

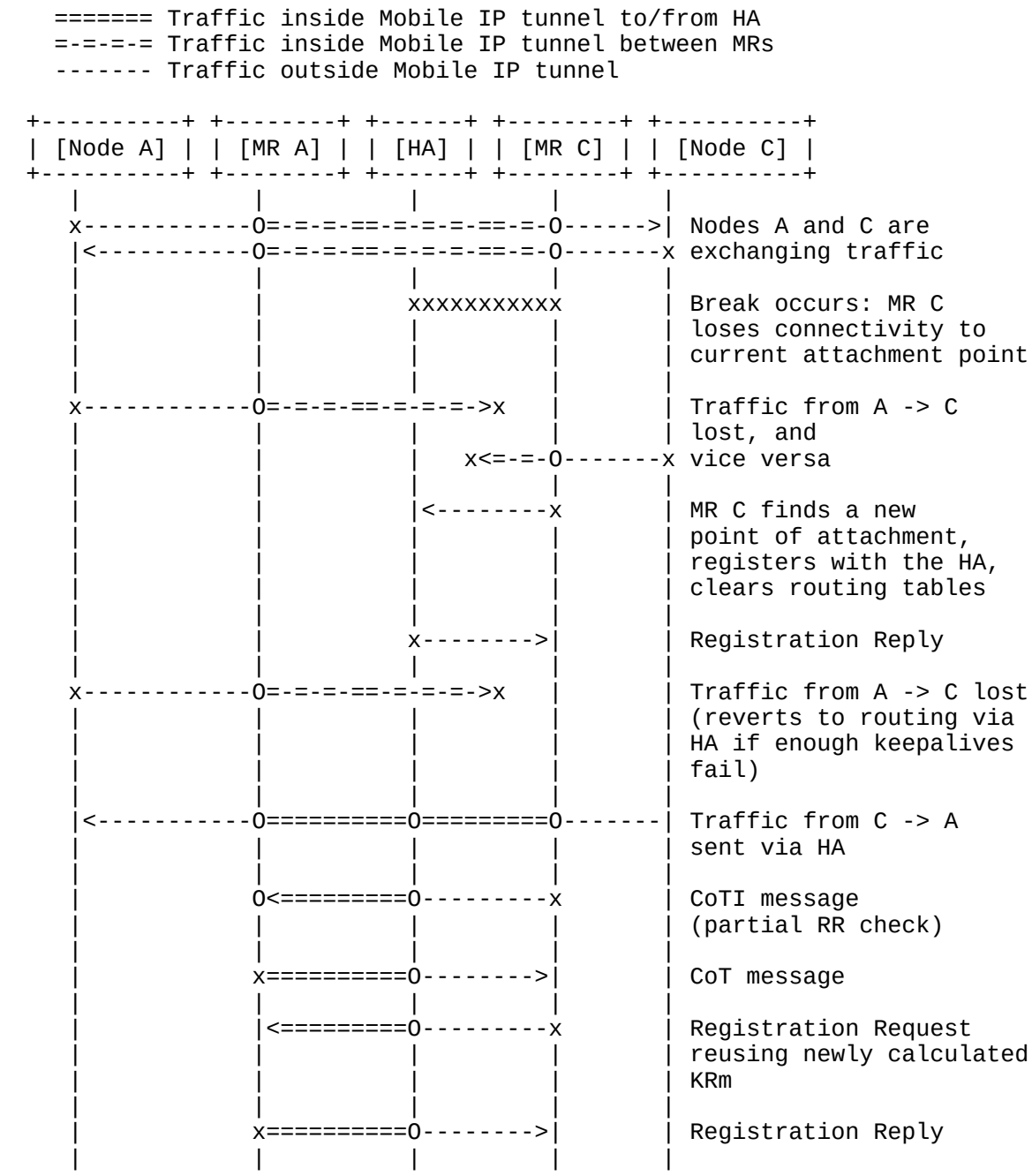
```
  ======= Traffic inside Mobile IP tunnel to/from HA
  =-=-=-= Traffic inside Mobile IP tunnel between MRs
  ------- Traffic outside Mobile IP tunnel
```

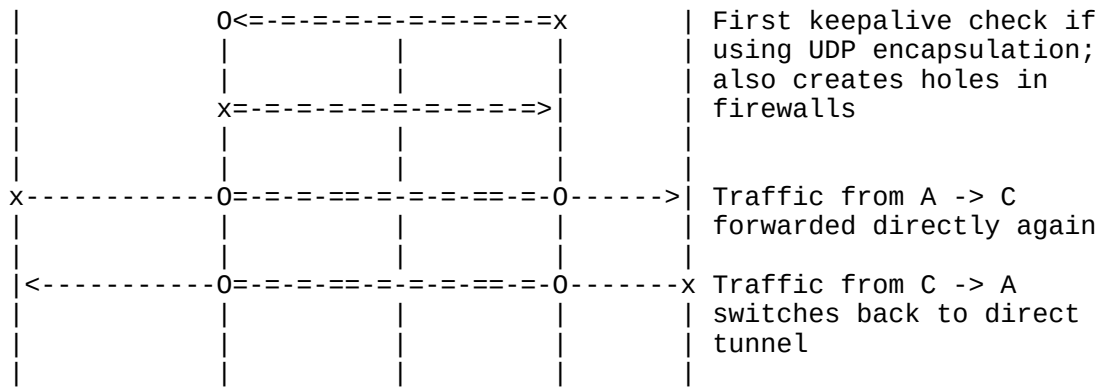




8.3. Handovers

In this signaling example, MR C changes its CoA while route optimization between MR A and MR C is operating and data is being transferred. Cases where the handover is graceful ("make before break") and ungraceful ("break before make") both occur in similar fashion, except that in the graceful version no packets are lost. This diagram considers the case where MR C gets immediate notification of lost connectivity, e.g., due to a link status indication. MR A would eventually notice the breakdown, due to keepalive messages failing.


======= Traffic inside Mobile IP tunnel to/from HA
=-=-=-= Traffic inside Mobile IP tunnel between MRs
------- Traffic outside Mobile IP tunnel
[Node A]
[MR A]
[HA]
[MR C]
[Node C]
Nodes A and C are exchanging traffic
Break occurs: MR C loses connectivity to current attachment point
Traffic from A -> C lost, and
vice versa
MR C finds a new point of attachment, registers with the HA, clears routing tables
Registration Reply
Traffic from A -> C lost (reverts to routing via HA if enough keepalives fail)
Traffic from C -> A sent via HA
CoTI message (partial RR check)
CoT message
Registration Request reusing newly calculated KRm
Registration Reply



9. Protocol Constants

MAX_NONCE_LIFETIME	240 seconds
MAX_TOKEN_LIFETIME	210 seconds
MAX_UPDATE_RATE	5 times

10. IANA Considerations

IANA has assigned rules for the existing registries "Mobile IP Message Types" and "Extensions to Mobile IP Registration Messages", specified in RFC 5944 [RFC5944]. New Mobile IP message types and extension code allocations have been made for the messages and extensions listed in Section 5.

The route optimization authentication processing requires four new message type numbers. The new Mobile IP Message types are listed below, in Table 1.

Value	Name
24	Home Test Init message
25	Care-of Test Init message
26	Home Test message
27	Care-of Test message

Table 1: New Values and Names for Mobile IP Message Types

Three new registration message extension types are required and listed in Table 2. The first type, 153, is skippable and has been allocated from range 128-255. The other two, 49 and 50, are non-skippable and have been allocated from range 0-127, with 49 being of the "short" format and 50 being of the "long" format. None of the messages are permitted for notification messages.

Value	Name
153, 128-255	Mobile Router Route Optimization Indication
49, 0-127	Route Optimization Extensions
50, 0-127	Route Optimization Data

Table 2: New Values and Names for Extensions in Mobile IP Registration Messages

In addition, the registry "Code Values for Mobile IP Registration Reply Messages" has been modified. A new success code, 2, should be allocated as follows:

2 Concurrent registration (pre-accept)

In addition, a new allocation range has been created as "Error Codes from the Correspondent Node", subject to the policy of Expert Review [RFC5226]. The range is 201-210. Three new Registration Reply codes have been allocated from this range. They are specified in Table 3, below:

Value	Name
201	Expired Home nonce Index
202	Expired Care-of nonce Index
203	Expired nonces

Table 3: New Code Values and Names for Mobile IP Registration Reply Messages

Three new number spaces were required for the subtypes of the extensions in Table 2. A new registry, named "Route Optimization Types and Subtypes", has been created with an allocation policy of RFC Required [RFC5226]. The registration entries include Type, Subtype, and Name. Type and Subtype have a range of 0-255. Types are references to registration message extension types. Subtypes are allocated initially as in Table 4, below:

Type	Subtype	Name
153	0	Reserved
153	1	Mobile Router Route Optimization Capability Extension
49	0	Reserved
49	1	Route Optimization Reply
49	2	Mobile-Correspondent Authentication Extension
49	3	Care-of Address Extension
50	0	Reserved
50	1	Route Optimization Prefix Advertisement Extension

Table 4: Initial Values and Names for Registry Route Optimization Types and Subtypes

11. Security Considerations

There are two primary security issues: One issue relates to the RR check, which establishes that a specific CoA is, indeed, managed by a specific HoA. The other issue is trust relationships and an arbitrary router claiming to represent an arbitrary network.

The end-user traffic can be protected using normal IPsec mechanisms.

11.1. Return Routability

The RR check's security has been vetted with Mobile IPv6. There are no major differences, apart from two issues: connectivity check and replay attack protection. The connectivity check is conducted with a separate ICMP message exchange. Replay attack protection is achieved with Mobile IPv4 timestamps in the Registration Request's Identification field, in contrast to the sequence numbers used in Mobile IPv6.

The RR procedure does not establish any kind of state information on the CR; this mitigates denial-of-service attacks. State information is only maintained after a Registration Request has been accepted.

11.2. Trust Relationships

The network of trust relationships in home agent-assisted route optimization solves possible trust issues: An arbitrary CR can trust an arbitrary MR that it is indeed the proper route to reach an arbitrary mobile network.

It is assumed that all MRs have a trust relationship with the HA. Thus, they trust information provided by the HA.

The HA provides information matching HoAs and network prefixes. Each MR trusts this information.

MRs may perform the RR procedure between each other. This creates a trusted association between the MR's HoA and CoA. The MR also claims to represent a specific network. This information is not trustworthy as such.

The claim can be verified by checking the HoA <-> network prefix information received, either earlier, or due to an on-demand request, from the HA. If they match, the MR's claim is authentic. If the network is considered trusted, a policy decision can be made to skip this check. Exact definitions on situations where such decisions can be made are out of scope for this document. The RECOMMENDED general practice is to perform the check.

12. Acknowledgements

Thanks to Alexandru Petrescu for constructive comments and support. Thanks to Jyrki Soini and Kari Laihonen for initial reviews. This work was supported by TEKES as part of the Future Internet program of TIVIT (Finnish Strategic Centre for Science, Technology and Innovation in the field of ICT).

13. References

13.1. Normative References

[RFC2003] Perkins, C., "IP Encapsulation within IP", RFC 2003, October 1996.

[RFC2004] Perkins, C., "Minimal Encapsulation within IP", RFC 2004, October 1996.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2784] Farinacci, D., Li, T., Hanks, S., Meyer, D., and P. Traina, "Generic Routing Encapsulation (GRE)", RFC 2784, March 2000.

[RFC3519] Levkowetz, H. and S. Vaarala, "Mobile IP Traversal of Network Address Translation (NAT) Devices", RFC 3519, April 2003.

[RFC5177] Leung, K., Dommety, G., Narayanan, V., and A. Petrescu, "Network Mobility (NEMO) Extensions for Mobile IPv4", RFC 5177, April 2008.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.

[RFC5944] Perkins, C., Ed., "IP Mobility Support for IPv4, Revised", RFC 5944, November 2010.

13.2. Informative References

[MIP-RO] Perkins, C. and D. Johnson, "Route Optimization in Mobile IP", Work in Progress, September 2001.

[MIPv4FLOW] Gundavelli, S., Ed., Leung, K., Tsirtsis, G., Soliman, H., and A. Petrescu, "Flow Binding Support for Mobile IPv4", Work in Progress, February 2012.

[RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, November 1987.

[RFC3543] Glass, S. and M. Chandra, "Registration Revocation in Mobile IPv4", RFC 3543, August 2003.

[RFC4086] Eastlake 3rd, D., Schiller, J., and S. Crocker, "Randomness Requirements for Security", BCP 106, RFC 4086, June 2005.

[RFC4282] Aboba, B., Beadles, M., Arkko, J., and P. Eronen, "The Network Access Identifier", RFC 4282, December 2005.

[RFC6275] Perkins, C., Ed., Johnson, D., and J. Arkko, "Mobility Support in IPv6", RFC 6275, July 2011.

Authors' Addresses

Antti Makela
Aalto University
Department of Communications and Networking (Comnet)
P.O. Box 13000
FIN-00076 Aalto
FINLAND

EMail: antti.t.makela@iki.fi

Jouni Korhonen
Nokia Siemens Networks
Linnoitustie 6
FI-02600 Espoo
FINLAND

EMail: jouni.nospam@gmail.com