Network Working Group
Internet-Draft
Intended status: Experimental
Expires: April 21, 2011

A. Makela
Aalto University
J. Korhonen
Nokia Siemens Networks
October 18, 2010

Home Agent assisted Route Optimization between Mobile IPv4 Networks
draft-ietf-mip4-nemo-haaro-02

Abstract

This document describes a Home Agent assisted Route Optimization functionality to 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 single organization or similar entity. The functionality adds possibility to discover eligible peer nodes based on information received from Home Agent, Network Prefixes they represent, and how to establish direct tunnel between such nodes.

Table of Contents

1. Introduction and motivations

Traditionally, there has been no method for route optimization in Mobile IPv4 [RFC3344] apart from an early attempt [I-D.ietf-mobileip-optim]. Unlike Mobile IPv6 [RFC3775], where Route Optimization has been included from the start, with Mobile IPv4 route optimization hasn't been addressed in a generalized scope.

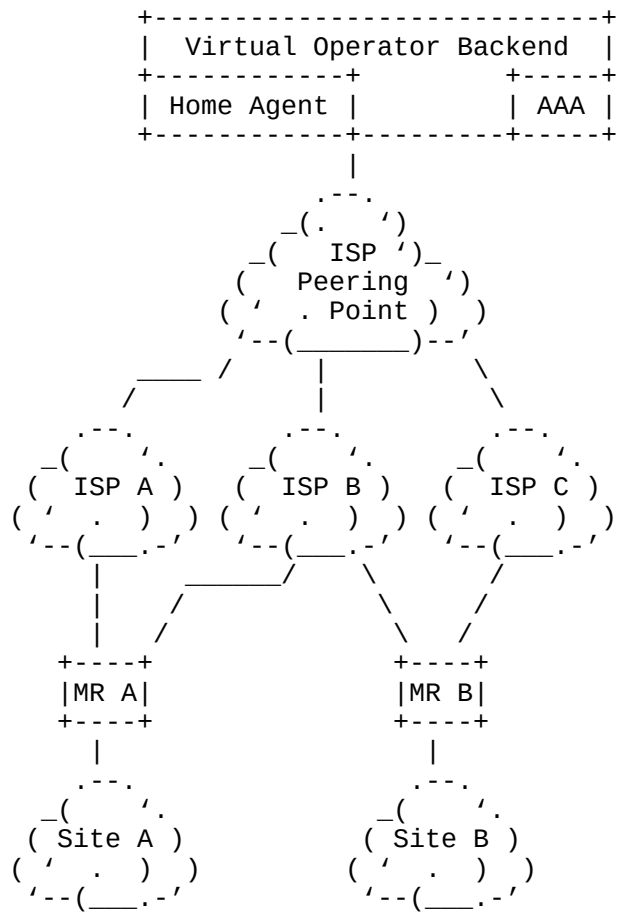
Even though general route optimization may not be of interest in the scope of IPv4, there are still specific applications for Route Optimization in Mobile IPv4. This document proposes method to optimize routes between networks behind mobile routers, as defined by NEMO [RFC5177]. Although NAT and pending shortage of IPv4 addresses makes widespread deployment not feasible, using Route Optimization only in routers is still a practical scenario.

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 multi-homing scenarios have traditionally been solved by using such technologies as multihoming BGP. However, a more lightweight and economical solution is desirable.

From service provider perspective a common topology for 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 PVCs), MPLS VPNs or Layer 3 site-to-site VPNs. With a Service Level Agreement, a customer can obtain a 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 reliable approach for mission-critical applications.

Mobile IP, especially mobile routers, 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 heavy load at the backend. Route Optimization between mobile networks addresses this issue, by taking network load off the home agent and the backend.

An example network is pictured below:



Virtual service provider architecture using NEMOv4

In this example case, organization network consists of two sites, that are connected via 2 ISPs for redundancy reasons. Mobile IP allows fast handovers without problems of multi-homing 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 through ISP B's network, or in case of a link failure, via the ISP peering point (such as MAE-WEST). The backend will not suffer from heavy loads.

The primary design goal is to limit the load to the backend to minimum. Additional design goals include extensibility to a more generalized scope, beyond the need of a single, coordinating Home Agent.

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 3344 [RFC3344] defines 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 care-of address: 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 which 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 care-of addresses are used either.

Correspondent Router (CR)

RFC 3344 [RFC3344] defines a Correspondent node as a peer with which a mobile node is communicating. Correspondent Router is a peer Mobile Router which MAY also represent one or more entire networks.

Home Address (HoA)

RFC 3344 [RFC3344] defines 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 3344 [RFC3344] defines Home Agent as a router on a mobile node’s home network which 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

Network Prefix with the mask of /32. e.g. 192.0.2.254/32, consisting of a single host.

Mobility Binding

RFC 3344 [RFC3344] defines Mobility Binding as the association of Home Address with a Care-of address, along with the lifetime remaining for that association.

Mobile Network Prefix RFC 5177 [RFC5177] defines Mobile Network Prefix as the network prefix of the subnet delegated to a Mobile Router as the Mobile Network.

Mobile Router (MR)

Mobile Router as defined by RFC 5177 [RFC5177] and RFC 3344 [RFC3344]. They 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

Data structure maintained by Mobile Routers on possible destinations for Route Optimization. Contains information (Home Addresses) on potential Correspondent Routers and their associated Mobile Networks.

Return Routability, RR

Procedure to bind a Mobile Router’s Home Address to a Care-of address on a Correspondent Router with a degree of trust.

3. Mobile IPv4 route optimization between mobile networks

This section describes the changed functionality of Home Agent and Mobile Router compared to the base NEMOv4 operation defined in [RFC5177]. The basic premise is still the same; Mobile Routers, when registering to the Home Agent, either inform the Home Agent of the mobile network prefixes they are managing (explicit mode) or get prefixes assigned by Home Agent (implicit mode). However, instead of prefix <-> Mobile Router mapping information only remaining on the Home Agent and the single Mobile Router, this information will now be distributed to the other Mobile Routers as well.

The Home Agent-assisted Route Optimization is primarily intended for helping to optimize traffic patterns between multiple sites in an single organization or administrative domain; However, extranets can also be reached with optimized routes, as long as all Mobile Routers connect to the same Home Agent. The procedure aim to maintain backwards compatibility; With legacy nodes or routers full connectivity is always preserved even though optimal routing cannot be guaranteed.

The schema requires a Mobile Router to be able to receive messages from Home Agent and other Mobile Routers unsolicited - that is, without first initiating a request. This behavior is similar to the registration revocation procedure [RFC3543]. Many of the mechanisms are same - including the fact that advertising route optimization support upon registration implies capability to receive registration requests and return routability messages from other Mobile Routers.

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 joining Mobile Router MAY request information on route-optimizable network prefixes. The Mobile Router MAY also allow redistribution of information on its managed network prefixes regardless whether they are explicit or implicit (statically configured or assigned by Home Agent). These are indicated with Mobile Router Route Optimization capability extension, see Section 5.1. If the Home Agent accepts the request for Route Optimization, this is indicated with Route Optimization Reply extension (Section 5.2) in the registration reply.

Note that the redistribution of network prefix information from the Home Agent happens only during the registration signaling. There are no "routing updates" from Home Agent except during re-registrations triggered by handovers, registration timeouts and specific solicitation. The solicitation re-registration MAY occur if a Correspondent Router receives a registration request from a unknown Mobile Router (see Section 3.3.3).

3.1.1. Advertising route-optimizable prefixes

As noted, NEMO-supporting Home Agent already maintains, and in some cases assigns, information on which network prefixes are reachable behind specific Mobile Routers. Only change to this functionality is that this information can now be distributed to other Mobile Routers

upon request. This request is implied by including Route Optimization capability extension, Section 5.1, and setting the 'R' bit.

When a Home Agent receives a registration request, standard authentication and authorization procedures are conducted.

If registration is successful and the Route Optimization capability information extension was present in the registration request, the reply message MUST include Route Optimization Reply extension (Section 5.2) to indicate whether Route Optimization was accepted. Furthermore, the extension also informs Mobile Router if NAT was detected between Home Agent and the Mobile Router using the procedure in RFC 3519 [RFC3519], which is based on the discrepancy between requester's indicated Care-of address and packet's source address.

The reply message MAY also include one route optimization prefix advertisement extension which informs the Mobile Router of existing mobile network prefixes and the Mobile Routers 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 Mobile Router, identified by Home Address, is listed with according prefix(es) and their respective realm(s).

Each network prefix can be associated with a realm, usually in the form 'organization.example.com'. Besides the routers in customer's own organization, the prefix list may also include other Mobile Routers, e.g. Default prefix (0.0.0.0/0) pointing towards Internet gateway for Internet connectivity, and possible extranets. The realm information can be used to make policy decisions on the Mobile Router, such as preferring optimization within specific realm only.

In a typical scenario where Network Prefixes are allocated to Mobile Routers connecting to a single Home Agent, 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 same higher-level domains. These compression mechanisms are further explained in Section 4.

Upon receiving registration reply with a Route Optimization prefix advertisement extension, the Mobile Router SHALL insert the Mobile Router Home Addresses included in the extension as a host-prefixes to the local Route Optimization Cache if they do not already exist. If present, any additional prefixes information SHALL also be inserted to the Route Optimization Cache.

The Mobile Router MAY discard entries from a desired starting point onwards, 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 Mobile Router SHOULD use information on all advertised prefixes.

3.1.2. Route Optimization cache

Mobile routers supporting route optimization will maintain a Route Optimization Cache.

The Route Optimization Cache contains mappings between potential Correspondent Router HoA's, network(s) associated with each HoA, network topology, and Return Routability procedure-related information. The Cache is populated based on information received from Home Agent in Route optimization prefix advertisements, and in registration messages from Correspondent Routers. Portions of the cache may also be configured statically.

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

CR-HoA

: Correspondent Router's Home Address. Primary key identifying each CR.

CR-CoAs

: Correspondent Router's Care-of Address(es). May be empty if none known. Potential tunnel's destination address(es).

MR-CoAs

: Mobile Router's Care-of Address(es) used with this Correspondent Router. Tunnel's source address.

Tunnels

: Tunnel interface(s) associated with this Correspondent Router. 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, set for all pairs of potential MR-CoA's and CR-CoA's which require NAT awareness and the behavior is known, populated either statically or based on discovery. If set to true, the MR can establish a UDP tunnel towards the CR, using this pair of CoA's. A received advertisement can indicate this to be set initially false for all respective CR's CoA's. Affects tunnel establishment direction, see Section 3.3.4 and the registration procedure in deciding which care-of-address to include in Care-of-address extension in registration reply. If the entry exists, mandates use of UDP encapsulation.

RRSTATEs

Return routability state for each CR-HoA and MR-CoA pair. States are INACTIVE, IN PROGRESS and ACTIVE. If state is INACTIVE, return routability procedure must be completed before forwarding route-optimized traffic. If state is IN PROGRESS or ACTIVE, the information concerning this Correspondent Router MUST NOT be removed from Route Optimization Cache as long as tunnel to the Correspondent Router 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 Return Routability procedure, they are calculated in-situ based on nonces and router key. If configured statically, RRSTATE is permanently set to ACTIVE.

Care-of nonce indexes If the KRm was established with Return Routability procedure, contains the Care-of nonce index for each MR-CoA - CR-HoA pair.

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

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

Home keygen token If the KRm was established with Return Routability procedure, contains the Home keygen token for each CR-HoA.

Network Prefixes

A list of destination network prefixes reachable via this Correspondent Router. 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 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 Correspondent Router itself does need validation beyond Return Routability procedure. For other prefixes, the confirmation is done by soliciting the information from HA. Traffic for prefixes which 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 Mobile Router.

3.2. Return routability procedure

The purpose of return routability procedure is to establish Care-of-Address <-> Home Address bindings in a trusted manner. The return routability procedure for Mobile IPv6 is described in [RFC3775]. Same principles apply to the Mobile IPv4 version: Two messages are sent to Correspondent Router's Home Address, one via Home Agent using Mobile Router's Home Address, and the other directly from the Mobile Router CoA, with two responses coming through same routes. 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 Return Routability procedure is a method provided by Mobile IP protocol to establish the KRm in a relatively lightweight fashion.

If desired, the KRms can be configured to Mobile Routers statically, or using an desired external secure key provisioning mechanism. If KRm's are known to the Mobile Routers via some other mechanism, Return Routability procedure can be skipped. Such provisioning mechanisms are out of scope for this document.

Assumption on traffic patterns is that the Mobile Router that initiates the RR procedure can always send outbound messages, even when behind NAT or firewall. This basic assumption made for NAT Traversal in [RFC3519] is also applicable here. In case the Correspondent Router is behind such obstacles, it receives these messages via the reverse tunnel to CR's Home Address, thus any problem regarding the CR's connectivity is addressed during the registration to the Home Agent.

3.2.1. Router keys

Each Mobile Router maintains a 'correspondent router key', Kcr, which is MUST NOT be shared with any other entity. Kcr is used for authenticating peer Mobile Routers in the situation where mobile router is acting as a CR. This is analogous to node key, Kcn, in Mobile IPv6. Correspondent Router uses router key to verify that the keygen tokens sent by Mobile Router in registration request are its own. The router key MUST be a random number, 16 octets in length.

The Mobile Router MAY generate a new key at any time to avoid persistent key storage. If desired, it is RECOMMENDED to expire the keys in conjunction with nonces, see Section 3.2.3.

3.2.2. Nonces

Each Mobile Router also maintains one or more indexed nonces. Nonces should be generated periodically with a good random number generator. The Mobile Router may use same nonces with all Mobile Routers. 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 ones in Mobile IPv6. Mobile Routers keep both the current nonce and small set of valid previous nonces whose lifetime have not expired yet. Nonce should be kept acceptable for at least MAX_TOKEN_LIFETIME (see Section 9) seconds after it has first been used in constructing a return routability response. However, the correspondent router 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 a time ago the nonces were used for the first time, and does not generate new nonces on every return routability request.

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

The Return Routability procedure consists of four Mobile IP messages: Home Test Init, Care-of Test Init, Home Test and Care-of Test. They are constructed as shown in Section 5.6 through Section 5.9. If the Mobile Router has included the Mobile Router 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 is set to Correspondent Router's HoA.

The return routability procedure begins with the Mobile Router sending HoTI and CoTI messages, each containing a 64-bit random value, the cookie. The cookie is used to bind specific signaling exchange together.

Upon receiving the HoTI or CoTI message the Correspondent Router MUST have a secret Kcr and nonce. If it does not have this material yet, it MUST produce it before continuing with the return routability procedure.

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

Return Routability procedure SHOULD be initiated only when the Route Optimization Cache's RRSTATE field for the desired Care-of Address with target Correspondent Router is INACTIVE. When Return Routability procedure is initiated, the state MUST be set to IN PROGRESS. In case of handover occurring, the Mobile Router 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 has expired, the RRSTATE MUST be set to INACTIVE. The Return Routability procedure may then be restarted as needed.

Upon completion of Return Routability procedure, the Routing Optimization Cache's RRSTATE field is set to ACTIVE, allowing for registration requests to be sent. The Mobile Router will establish a registration management key KRm by default using MD5 hash algorithm:

KRm = MD5 (home keygen token | care-of keygen token)

When de-registering (by setting time to zero), care-of keygen token is not used. Instead the Registration management key is generated as follows:

KRm = MD5 (home keygen token)

Like in Mobile IPv6, the Correspondent Router does not maintain any state for the Mobile Router until after receiving a registration request.

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 Mobile Router and Correspondent Router. The term "Mobile Router" applies to the router initiating the Route Optimization procedure, and "Correspondent Router" indicates the peer router.

Especially compared to Mobile IPv6 route optimization there are two issues that are different regarding IPv4. First of all, since Mobile IPv4 always uses tunnels, there must be a tunnel established between MR and CR's Care-of addresses. The Correspondent Router learns of Mobile Router's Care-of address as it is provided by the Registration Request. The Mobile Router learns Correspondent Router's Care-of address by a new extension, "Care-of Address", in registration reply. Second issue is rising from security standpoint: In a registration request, the Mobile Router 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 to Home Agent to verify the claim.

Additional aspect is that Mobile Router MAY use different Care-of-Address for different Correspondent Routers (and Home Agent). This is useful in situations where 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 Mobile Router knows the eligible route-optimizable networks, the route optimization between all Correspondent Routers can be established at any time; However a better general practice is to conduct Route Optimization on-demand only. It is RECOMMENDED to start Route optimization only be started when receiving a packet where destination address is in a locally managed prefix (and the prefix is registered as route optimizable) and source address exists in the network prefixes of Route Optimization Cache. With small number of Mobile Routers, 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 Mobile Router maintains a routing table. In a typical situation, the Mobile Router has one or more interface(s) to the local networks, one or more interface(s) to wide-area networks (such as provided by ISPs), and a tunnel interface to the Home Agent. Additional tunnel interfaces become activated as Route Optimization is being performed.

The routing table SHOULD typically contain Network Prefixes managed by Correspondent Routers associated with established route-optimized tunnel interfaces. In addition, host-routes to Correspondent Routers' Care-of addresses SHOULD be associated with the assigned to the physical interfaces assigned with corresponding MR-CoA. If the tunneling method does not require such host-routes, these can be omitted. A default route MAY point to the reverse tunnel to the Home Agent if not overridden by prefix information.

The route for the Home Address of Correspondent Router SHOULD also be pointing towards the optimized tunnel.

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 address SHOULD be considered an error situation. Any aggregation for routes in private address space SHOULD be conducted only at HA.

3.3.3. Inter-Mobile Router registration

If route optimization between Mobile Router and Correspondent Router is desired, either Return Routability procedure must have been performed (See Section 3.2), or key KRm must be pre-shared between the Mobile and Correspondent Router. If either condition applies, a Mobile Router MAY send a registration request to the Correspondent

Router's HoA from desired interface.

The registration request's source address and Care-of address field are set to the address of desired outgoing interface on the Mobile Router. The address MAY be same as the Care-of address used with Home Agent. The registration request MUST include Mobile-Correspondent Authentication extension defined in Section 5.3 and SHOULD include 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 Correspondent Router MUST check the identification field for validity. The registration request MUST include Home Address. The Authenticator field is hashed with the key KRm.

The Correspondent Router relies to the request with a Registration Reply. The registration reply MUST include Mobile-Correspondent Authentication extension defined in Section 5.3 and, if 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 Correspondent Router, or the Mobile Router itself is known to be behind NAT or firewall. If either of the conditions apply, registration request MUST specify UDP encapsulation. It is RECOMMENDED to always use UDP encapsulation to facilitate detecting of path failures via keepalive mechanism.

The Correspondent Router first checks the registration request's authentication against Kcr and nonce indexes negotiated during Return Routability procedure. This ensures that the registration request is coming from a correct Mobile Router. If the check fails, an appropriate registration reply code is sent (see Section (Section 10). If the failure is due to nonce index expiring, the Mobile Router sets RRSTATE for the CR to INACTIVE. Return routability procedure MAY then be initiated again.

If the check passes, the Correspondent Router MUST check whether the Mobile Router already exists in it's Route Optimization Cache and is associated with the prefixes included in the request (Prefixes are present and Flag HA is true for each prefix).

If the check against the cache fails, the Correspondent Router SHOULD send a re-registration request to Home Agent with the 'S' (solicitation) bit set, thus obtaining the latest information on Network Prefixes managed by incoming Mobile Router. 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 can also be rejected completely. This verification is done to protect against Mobile Routers claiming to represent arbitrary networks; However, since Home Agent is assumed to provide trusted information, it can authorize Mobile Router's claim. If the environment itself is considered trusted, the Correspondent Router can, as a policy, accept registrations from without this check; however, this is NOT RECOMMENDED as a general practice.

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

Upon receiving the registration reply, the Mobile Router MUST check if a tunnel to the Correspondent Router already exists. If the tunnel does NOT exist, or has wrong endpoints (CoAs), a new tunnel MUST be established and Route Optimization Cache updated. This is covered in detail in Section 3.3.4.

The Correspondent Router's routing table MUST be updated to include the Mobile Router's networks are reachable via the direct tunnel to the Mobile Router.

After the tunnel is established, the Mobile Router MAY update it's routing tables to reach all Correspondent Router's Prefixes via the tunnel, although it is RECOMMENDED to wait for the Correspondent Router to perform it's 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 two Mobile Routers, each working as each other's Correspondent Router, 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 Home Address (meaning a normal 32-bit integer comparison treating IPv4 addresses as 32-bit unsigned integers) receives a registration request (in CR role) while its own registration request (sent in MR role) is still pending, the attempt should be rejected with reply code "concurrent registration". If receiving such an indication, the recipient should not attempt to re-register again until a grace period has passed without route optimization occurring.

3.3.4. Inter-Mobile Router tunnels

Inter-Mobile Router tunnel establishment follows establishing standard reverse tunnels to the Home Agent. The registration request to Correspondent Router includes information on the desired encapsulation. It is RECOMMENDED to use UDP encapsulation. In the cases of GRE [RFC2784], IP over IP [RFC2003] or minimal encapsulation [RFC2004] no special considerations regarding the 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 Correspondent Router is its Care-of Address, not the Home Address where the registration requests were sent. This is different from creation of the Reverse Tunnel to Home Agent, which reuses the channel from registration signaling.

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

The initial UDP keepalive SHOULD be sent by the MR. Only after first keepalive is successfully completed, SHOULD the tunnel be considered eligible for traffic. If reply to the initial keepalive is not received, the MR may opt to attempt sending the keepalive with other Care-of addresses provided by the registration reply to check whether they provide better connectivity, or if all of these fail, perform a re-registration via alternative interface, or deregister completely. See Section 6.1. Once the initial keepalive packet has reached the CR and reply has been sent, the CR MAY start sending it’s 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, to use considerably shorter periods, adapting to the perceived latency as needed. However, the maximum amount of keepalives should at no point exceed MAX_UPDATE_RATE times in second. The purpose of keepalive is not to keep NAT or firewall mappings in place, but serve as a mechanism to provide fast response in case of path failures.

If both the Mobile Router and the Correspondent Router are behind separate NATs, route optimization cannot be performed between them. Possibilities to set up mutual tunneling when both routers are behind NAT, 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 valid KRm.

3.3.5. Constructing route-optimized packets

All packets received by the Mobile Router 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, "Correspondent Router" being the peer.

When a Mobile Router wishes to leave network, it SHOULD, in addition to sending the registration request to the Home Agent with lifetime set to zero, also send such a request to all known Correspondent Routers. The Correspondent Router(s), upon accepting this request and sending the reply, will check if it's Route Optimization Cache contains any prefixes associated with the requesting Mobile Router. These entries should be removed and routing table updated accordingly (traffic for the prefixes will be forwarded via the Home Agent 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 Correspondent Router simply needs to update the tunnel's destination to the Mobile Router's new Care-of Address. Mobile Router SHOULD keep accepting packets from both old and new care-of Addresses for a short grace period, typically in the order of ten seconds. In the case of UDP encapsulation, the port numbers SHOULD be reused if possible.

If the Mobile Router was unable to send the re-registration request before handover, it MUST send it immediately after handover has been completed and tunnel with the Home Agent is established. Since Care-of Address(es) changing invalidates the Krm, at it is

RECOMMENDED to conduct partial Return Routability by sending CoTI message via the new Care-of-Address and obtaining new care-of keygen token. In all cases, necessary tokens have to also be acquired if the existing ones have expired.

If a reply is not received for a registration request to a Correspondent Router, any routes to the network prefixes managed by the Correspondent Router MUST be removed from the routing table, thus causing the user traffic to be forwarded via the Home Agent.

3.4. Convergence and synchronization issues

The information the Home Agent maintains on Mobile Network prefixes and the Mobile Routers' Route Optimization Caches do not need to be explicitly synchronized. This is based on the assumption is 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, C handled by three Mobile Routers, MR A, MR B and MR C respectively. If they register to a Home Agent in this order, the situation goes as follows:

MR A registers; Receives no information on other networks from HA, as no other MR has registered yet.

MR B registers; Receives information on mobile network A being reachable via MR A.

MR C registers; Receives information on both of the other mobile networks.

If a node in mobile network C receives traffic from mobile network A, the route optimization is straightforward; MR C already has network A in its Route Optimization Cache. Thus, packet reception triggers Route Optimization towards MR A. When MR C registers to MR A (after Return Routability procedure is completed), MR A does not have information on mobile network C; Thus it will perform a re-registration to the Home Agent on-demand. This allows MR A to verify that MR C is indeed managing network C.

If a node in mobile network B receives to traffic from mobile network C, MR B has no information on network C. No route optimization is triggered. However, when the node in network B 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, the re-registrations to the Home Agent caused by timeouts will eventually cause convergence. However, this should be treated as a special case.

Note that all Mobile Routers are connected to same Home Agent. For possibilities concerning multiple Home Agents, 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

The 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 a new master prefix, where n is PLen. This is rounded up to nearest full octet. Thus, prefix lengths of /4 and /8 take 1 octet, /12 and /16 take 2 octets, /20 and /24 three, and larger than that 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 previous master prefix, with padding from master prefix filling the field to full octet. Maximum value of Plen-m is 8 (that is, delta MUST fit into one octet). If this is not possible, a new master prefix has to be declared.

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

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 are illustrated in Figure 1. Because of the padding to full octets, redundant information is also sent. The bit-patterns being transmitted are:

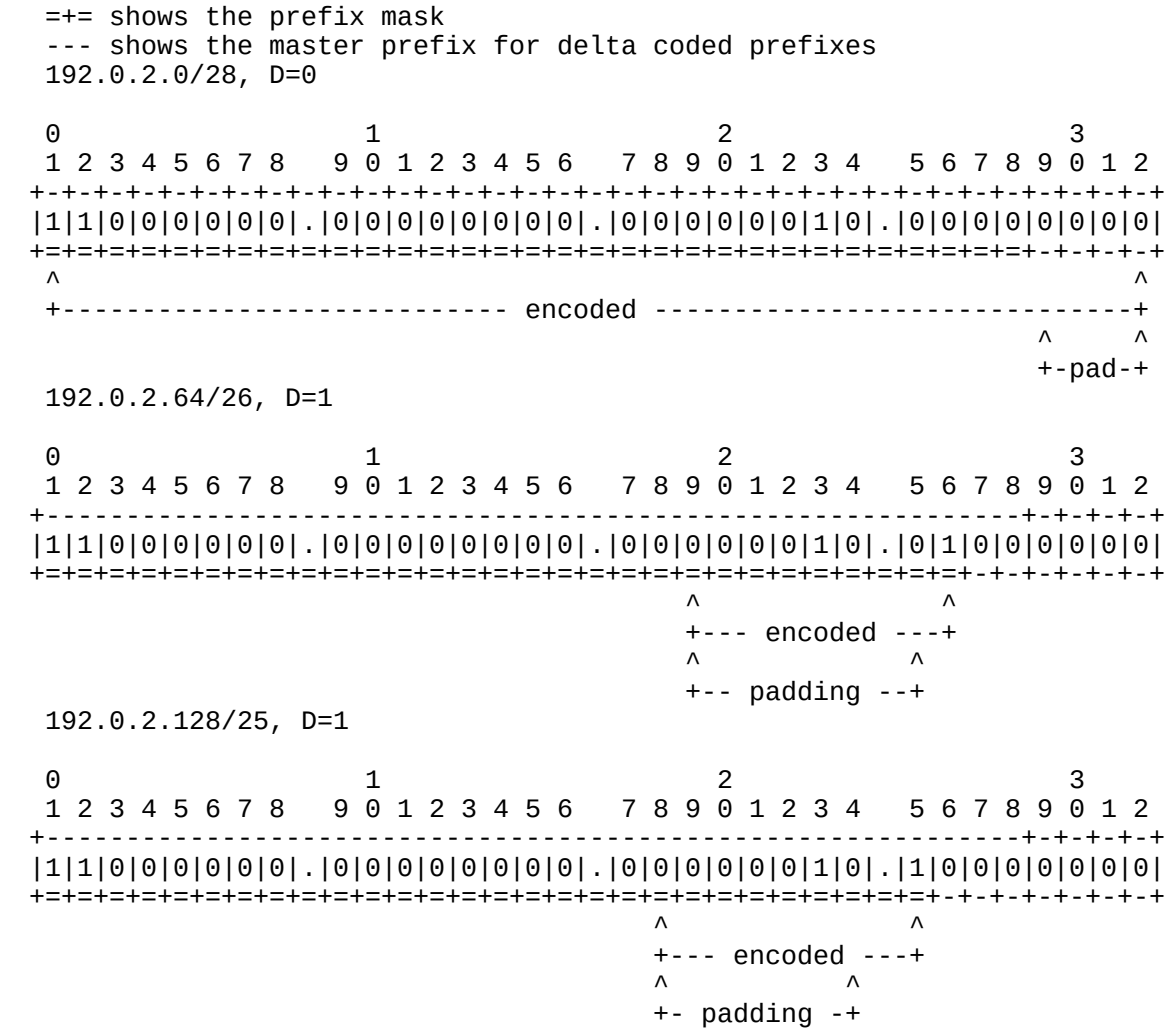


Figure 1: Prefix Compression Example

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, the 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 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 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

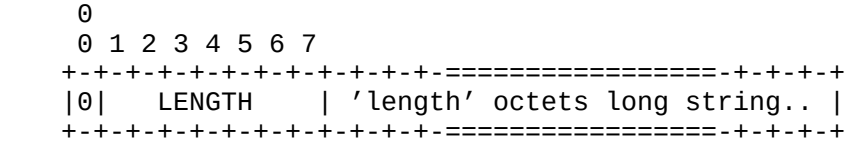
4.2.1. Encoding of compressed realms

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 arbitrary length text strings. 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. 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 maximum 128 strings. Thus, when adding the 129th string into the dictionary, the dictionary MUST first be reset to the initial state (i.e. Emptied) and the index of

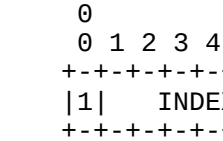
the string will become 0.

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 a one octet:



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 a one octet:



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 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 encoded realm.

The search algorithm also maintains the "longest non-matching string" for each input realm. Each time the search in 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 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 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 the greedy approach of encoding matches, the search algorithm and the dictionary update function is not the most optimal one. However, we do not claim the algorithm

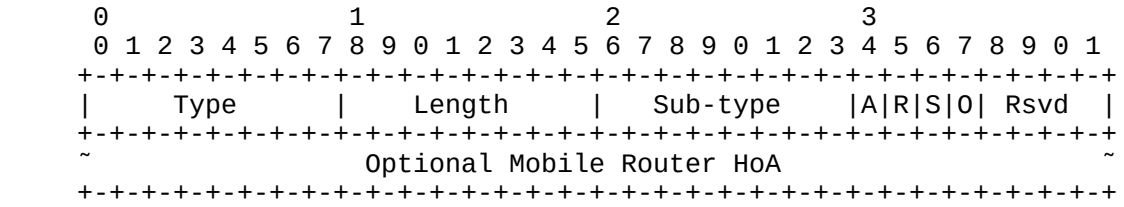
would be the most efficient. It functions efficiently enough for most inputs. In this example, the original input realm data was 79 octets and the compressed output excluding the end mark is 35 octets.

5. New Mobile IPv4 messages and extensions

This section describes the construction of all new information elements.

5.1. Mobile router Route optimization capability

This skippable extension MAY be sent by a Mobile Router to a Home Agent in the registration request message.



Type TBA_T1. Skippable; If Home Agent does not support route optimization advertisements, it can ignore this request and simply not include any information in the reply. "Short" extension format.

Sub_Type TBA_ST1_1

Reserved Set to zero, MUST be ignored on reception.

A Advertise my networks. If 'A' bit is set, the Home Agent is allowed to advertise the networks managed by this Mobile Router to other Mobile Routers. This also indicates that the Mobile Router is capable of receiving route optimization binding updates. In effect, this allows the Mobile Router to work in Correspondent Router role.

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

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

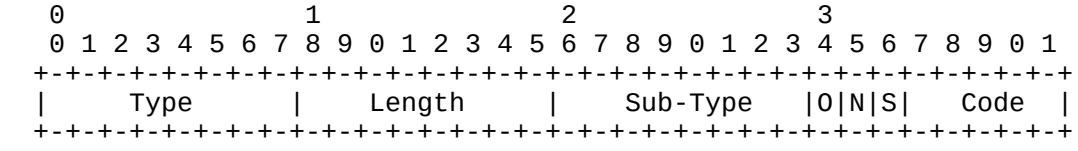
O Explicitly specifying the requesting Router is only able to initiate outgoing connections, not accept any incoming ones, due to NAT device, stateful firewall, or similar issue on any interface. This is reflected by the Home Agent in the reply, and distributed in Prefix Advertisements to outer Mobile Routers.

Optional Mobile Router HoA

Solicited Mobile Router's Home Address.

5.2. Route optimization reply

This non-skippable extension MUST be sent by a Home Agent to a Mobile Router in the registration reply message, if Mobile Router indicated support for Route Optimization in registration message and Home Agent supports Route Optimization.



Type TBA_T2 (Non-skippable), "short" extension format

Sub-Type TBA_ST2_1.

O The 'O' flag in Mobile Router Optimization capability extension was set during registration.

N Presence of NAT was detected by Home Agent. This informs the Mobile Router that it is located behind NAT. The detection procedure is specified in RFC 3519 [RFC3519], and is based on the discrepancy between registration packet's source address and indicated Care-of Address. The Mobile Router can use this information to make decisions about Route Optimization strategy.

S Responding to a solicitation. If 'S' bit was present in Mobile router Route optimization capability extension (Section 5.1), this is set, otherwise unset.

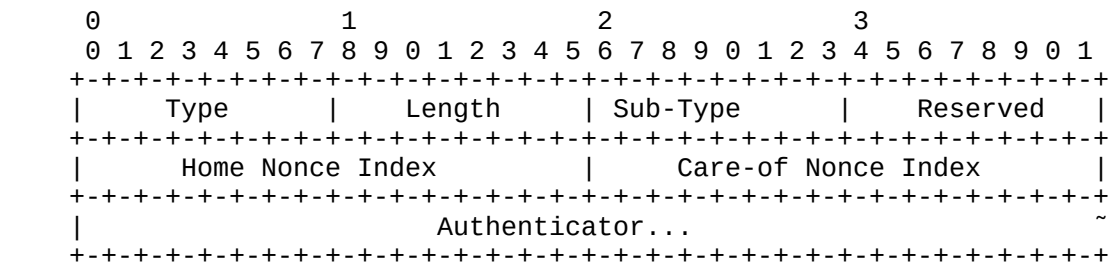
The Reply code indicates whether Route Optimization has been accepted. Values of 0..15 indicate assent and values 16..63 indicate Route Optimization is not done.

0 Will do Route Optimization

16 Route Optimization declined, reason unspecified.

5.3. Mobile-Correspondent authentication extension

Mobile-Correspondent authentication extension is included in registration requests sent from Mobile Router to Correspondent Router. The existence of this extension indicates that the message is not destined to a Home Agent, but another Mobile Router. The format is similar to the other Authentication Extensions defined in [RFC3344], with SPIs replaced by Nonce Indexes.



The Home Nonce Index field tells the Correspondent Router 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 Correspondent Router which nonce value to use when producing the Care-of Keygen Token.

Type TBA_T2 (non-skippable). "Short" extension format.

Sub-Type TBA_ST_2_2

Reserved Set to zero, MUST be ignored on reception.

Home Nonce Index

Home Nonce Index in use.

Care-of Nonce Index

Care-of Index in use.

Authenticator

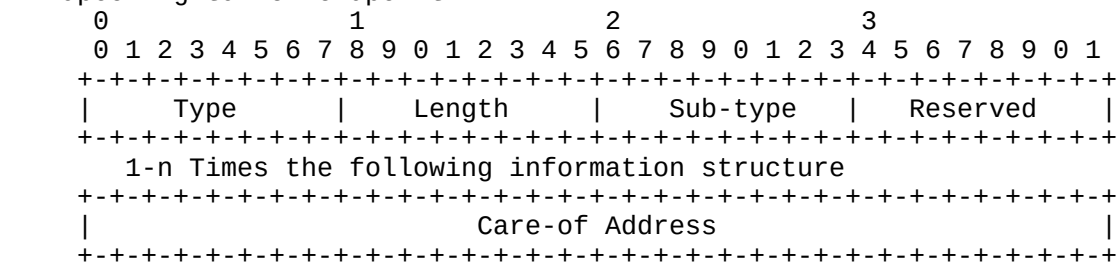
Authenticator field, by default constructed with HMAC_MD5 (KRm2, Protected Data)

The protected data, just like on other cases where Authenticator 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, and Nonce Indexes of this Extension.

5.4. Care-of address Extension

The Care-of Address extension is added to a registration reply sent by the Correspondent Router to inform the Mobile Router of the upcoming tunnel endpoint.



Type TBA_T2 (Non-skippable), "short" extension format

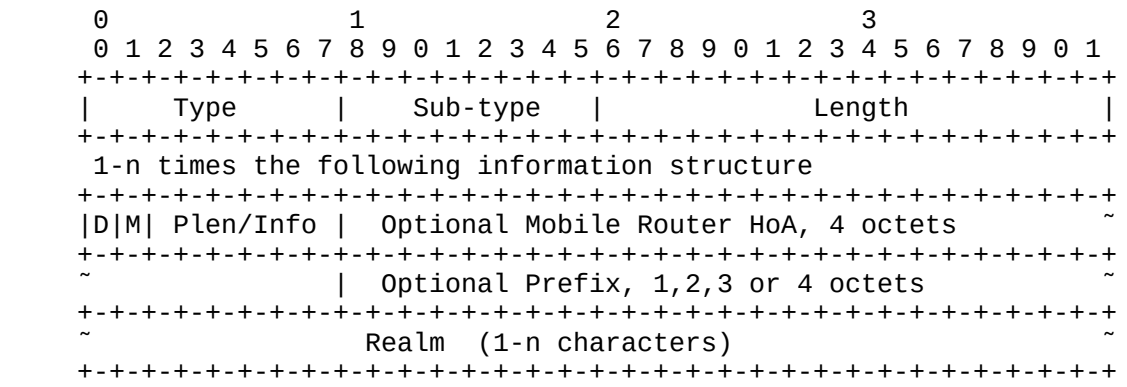
Sub-Type TBA_ST_2_3

Care-of Address

Care-of address(es) which may be used for tunnel with Mobile Router, in order of priority. Multiple CoA's MAY be listed to facilitate faster NAT traversal.

5.5. Route optimization prefix advertisement

This non-skippable extension MAY be sent by a Home Agent to a Mobile Router in the registration reply message. The extension is only included when explicitly requested by the Mobile Router in the registration request message. Implicit prioritization of prefixes is caused by the order of extensions.



Type TBA_T3 (Non-skippable), "long" extension format

Sub-Type TBA_ST3_1

D Delta. If D=1, the prefix is a delta from last Prefix where D=0. MUST be zero on first information structure, 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 M is set or not. If M=0, this indicates the length of the prefix advertised. 6 bits, allows for values from 0 to 63, of which 33-63 are illegal. If M=1, the Information field can be set to zero to indicate no specific information, or to 1 to indicate "outbound connections only". This indicates that the target Mobile Router can only initiate, not receive, connections on any of it’s interfaces (apart from the reverse tunnel to HA). This is set if the Mobile Router has explicitly requested it by the ’O’ flag in Mobile router Route optimization capability extension (Section 5.1).

Mobile router HoA

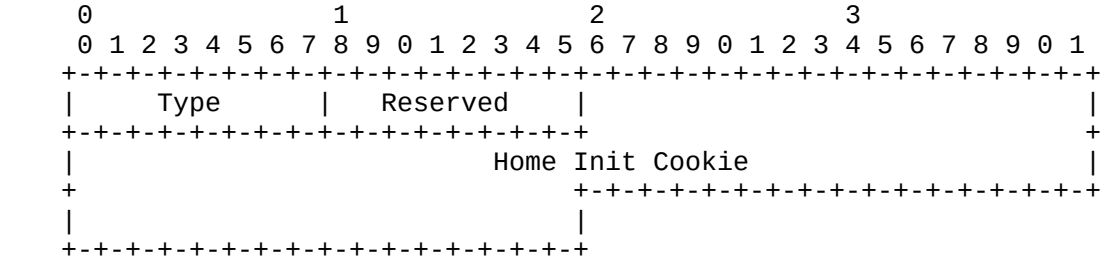
Mobile Router’s Home address. All prefixes in the following information structures where M=0 are maintained

by this Mobile Router. 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 nearest full octet. Least-significant bits starting off Plen (and are zeros) are omitted. If D=1, 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 optional compression scheme, refer to Section 4.2. This field is present only when M = 0.

5.6. Home-Test Init message



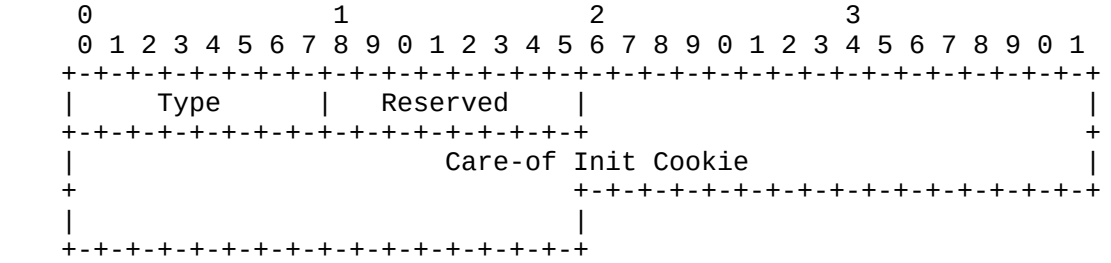
Type TBA_MIP1

Reserved Set to zero, MUST be ignored on reception.

Home Init Cookie

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

5.7. Care-of-Test Init message



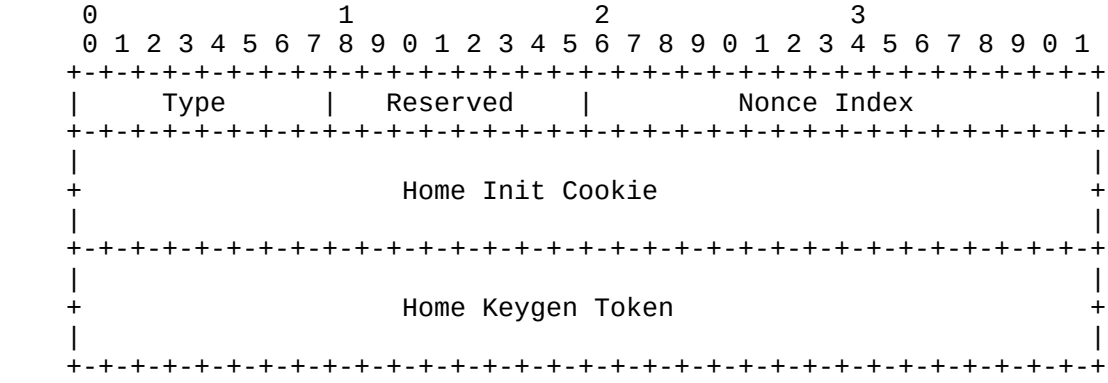
Type TBA_MIP2

Reserved Set to zero, MUST be ignored on reception.

Care-of Init Cookie

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

5.8. Home Test message



Type TBA_MIP3

Reserved Set to zero, MUST be ignored on reception.

Nonce Index

This field will be echoed back by the Mobile Router to the Correspondent Router in a subsequent registration request's authentication extension.

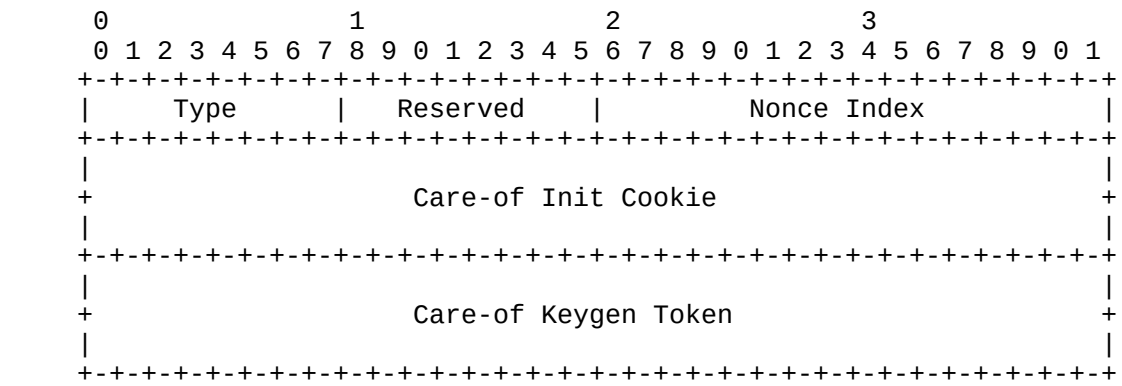
Home Init Cookie

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

Home Keygen Token

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

5.9. Care-of test message



Type TBA_MIP4

Reserved Set to zero, MUST be ignored on reception.

Care-of Nonce Index

This field will be echoed back by the Mobile Router to the Correspondent Router in a subsequent registration requests' authentication extension.

Care-of Init Cookie

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

Care-of Keygen Token

This field contains the 64 bit home keygen token used in the Return Routability procedure.

6. Special Considerations

6.1. NATs and stateful firewalls

Mechanisms described in MIP NAT traversal [RFC3519] allow the Home Agent to work with Mobile Routers situated behind a NAT device or a stateful firewall. Furthermore, the Home Agent may also detect whether NAT device is located between the Mobile Node and the HA. Mobile Router may also explicitly state it is behind a NAT or firewall on all interfaces, and this information is passed on to the other Mobile Routers with the information field in Route optimization

prefix advertisement extension (Section 5.5). Home Agent may also detect presence of NAT and informs the registering Mobile Router with the 'N' flag in Route Optimization Reply extension (Section 5.2). In the case of one or both of the routers is known to be behind NAT or similarly impaired (not being able to accept incoming connections), the tunnel establishment procedure SHOULD take this into account.

In the case where Mobile Router is behind NAT (or firewall) and Correspondent Router is not, the Mobile Router will, when 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 reply has been received, NAT mapping is in place and traffic can be sent.

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 in communication with the Home Agent, in the case of Route Optimization the path used for signaling is not used for tunneled packets, as signaling always uses Home Addresses, and 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 Registration Request and incoming tunnel packets due to port translation. This must not cause an error situation - the Correspondent Router MUST be able to accept tunneling packets from a different UDP source port than what was used in the Registration Request.

Since Mobile Routers may have multiple interfaces connecting to several different networks, it might be possible that specific Mobile Routers may only be able to perform Route Optimization using specific Care-of-address pairs, obtained from specific networks, for example in a case where two Mobile Routers have an interface behind same NAT. Similar case may be applicable to nested NATs. In such cases, Mobile Router MAY attempt to detect eligible Care-of-Address pairs by performing a registration and attempting to establish a tunnel (sending keepalives) with each Care-of-Address listed in the Registration Reply's Care-of-Address extension. The eligible pairs should be recorded in Route Optimization cache. If tunnel cannot be established with any CoA's, the Mobile Router MAY attempt to repeat the procedure with alternative interfaces. The above information on network topology can also be configured to the Mobile Routers either statically or via some external feedback mechanism.

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

6.2. Handling of concurrent handovers

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

6.3. Foreign Agents

Since Foreign Agents have been dropped from Network Mobility for Mobile IPv4 work, they are not considered here.

6.4. Multiple Home Agents

Mobile Routers can negotiate and perform route optimization without the assistance of Home Agent - if they can discover each others existence and thus know where to send registration messages. This document only addresses a logically single Home Agent that distributes network prefix information to the Mobile Routers. Problems arise from possible trust relationships; In this document the Home Agent 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 Home Agents, there are several possibilities. Note that standard high availability redundancy protocols, such as VRRP, can be utilized; However, in such case the Home Agent is still a single logical entity even if consisting of more than a single node.

Several possibilities exist for achieving Route Optimization between Mobile Routers attached to separate Home Agents, such as a new discovery/probing protocol, routing protocol between Home Agents or DNS SRV records, or a common AAA architecture. There already is a framework for HA to retrieve information from AAA so it can be considered as the most viable possibility. See Section 6.6 for information on possibility 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 Mobile Routers A and B respectively. If the routers are set up in such a fashion that Route Optimization is triggered when a packet is received from a Network Prefix in Route Optimization Cache, the following occurs if a node in network A starts sending ICMP echo requests (pinging) a node in network B.

MR B sees the incoming ICMP echo request packet, which is travelling inside the reverse tunnel to the Home Agent. MR B sees that the destination is in network B, and furthermore, source is in network A which exists in the cache. This triggers Route Optimization processing. Until RO is active, the ping packets (echo requests and replies) are routed via the reverse tunnel.

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

The traffic pattern is now 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 Home Agent and reverse tunnels. MR A now performs its own registration towards MR B. Upon completion, MR A notices that a tunnel to MR B already exists, but updates its routing table so that network B is now reachable via the MR A <-> MR B tunnel. From this point onward, traffic is bidirectional.

In this scenario, if MR A does NOT perform a separate route optimization (RR check and registration), 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 propagated to MR A. The end result is that MR B starts to receive packets for network B’ via the Home Agent and for network B via direct tunnel. If Reverse Path checking or similar mechanism is in use on MR B, packets from network A could be black holed.

Whether to perform this mutual registration or not thus depends on the situation, and whether Mobile Routers are going to start managing

additional Network Prefixes during operation.

6.6. Extensibility

The design considerations include several mechanisms which might not be strictly necessary if Route Optimization would only be desired between individual customer sites in a managed network. The registration procedure (with the optional Return Routability part), which allows for Correspondent Routers to learn Mobile Router's Care-of Addresses is not strictly necessary; The CoA's could have been provided by HA directly.

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

The Home Agent 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 - only problem is discovering if 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 address of Mobile Router.

In addition, with Route Optimization for single node, checks on whether a Mobile Router 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 Home Agent) are outside the scope of this document.

6.7. Load Balancing

The design simply provides possibility to create optimal paths between Mobile Routers; It doesn't dictate what should be the user traffic using these paths. One possible approach in helping facilitate load balancing and utilizing all available paths is presented in [I-D.ietf-mip4-multiple-tunnel-support], which effectively allows for multiple Care-of addresses for a single Home Address. In addition, per-tunnel load balancing is possible by using separate Care-of-Addresses 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 Mobile Routers. However, the MRs can always fall back to normal, reverse tunnelled routing if resource constraints are apparent.

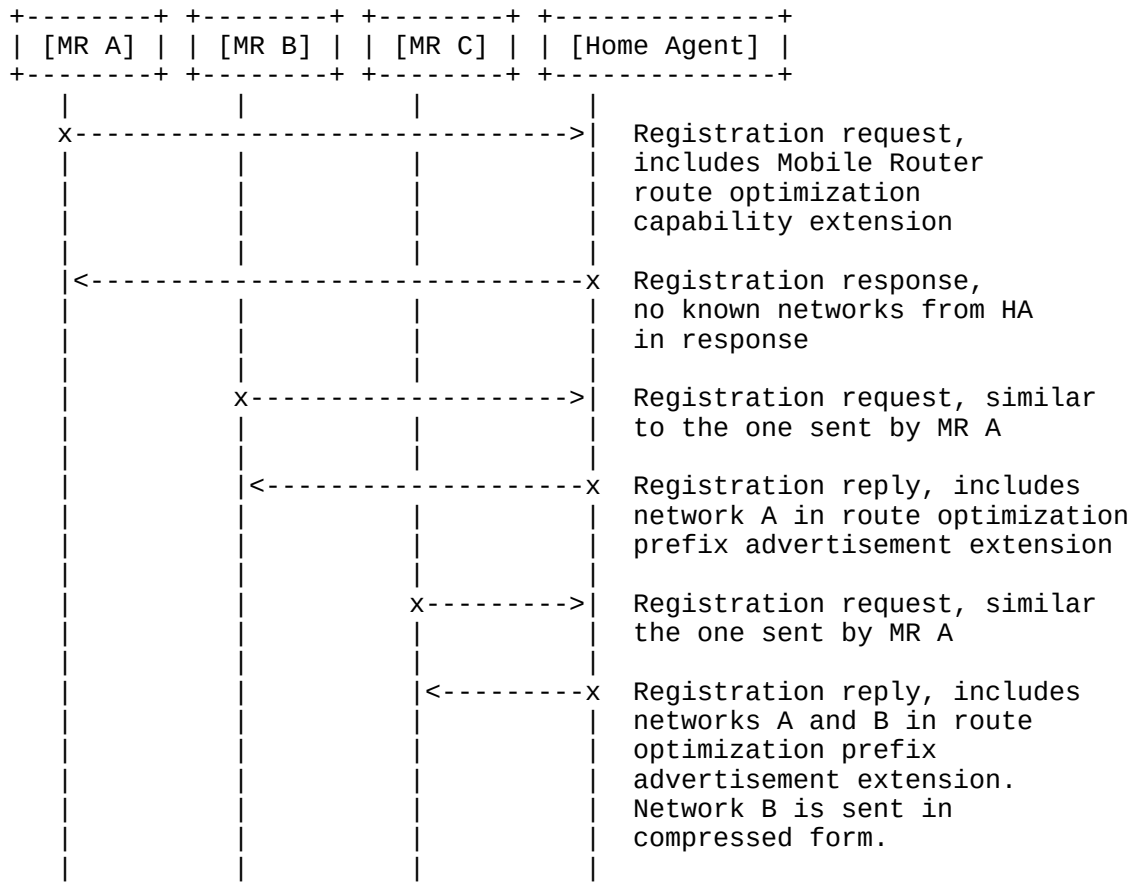
If there is a large number of optimization-capable prefixes, maintaining state for all of these may be an issue also, due to limits on routing table sizes.

Registration responses from Home Agent to Mobile Router may provide information on 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 mitigates this problem to an extent. There will, however, be some practical upper limit after which point some other delivery mechanism for the prefix information will be needed.

8. Example signaling scenarios

8.1. Registration request

The following example signaling assumes that there are three Mobile Routers, MR A, B, C, each managing network prefixes A, B, and C. At the beginning, no networks are registered to the Home Agent. Any AAA processing at the Home Agent is omitted from the diagram.

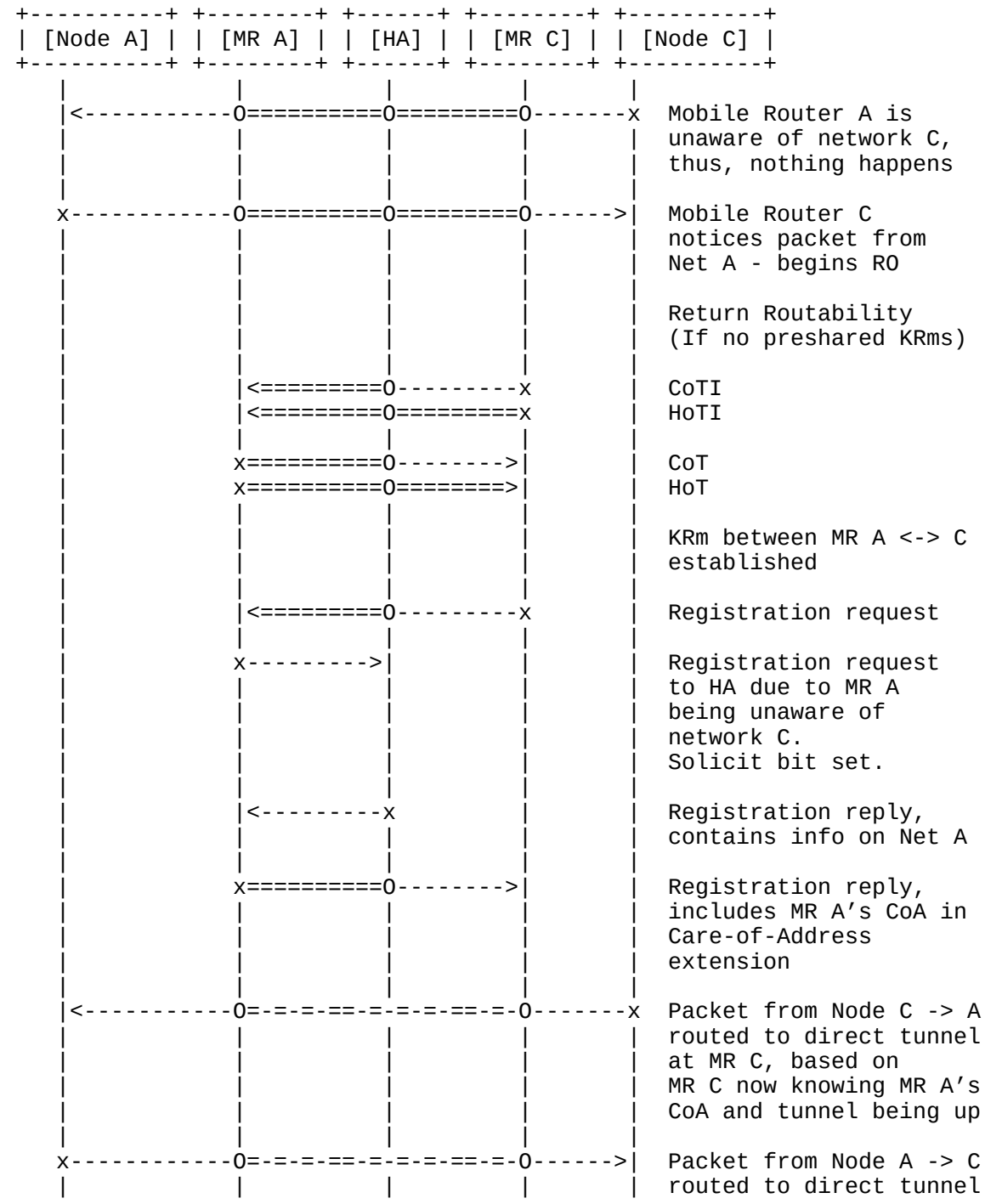


8.2. Route optimization with return routability

The following example signaling has same network setup as in Section 8.1 - Three mobile routers, each corresponding to their respective network. Node A is in network A and Node C is in network C.

At the beginning, no mobile routers know KRm's of each other. If the KRm's would be pre-shared or provisioned with some other method, the Return Routability messages can be omitted. Signaling 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.

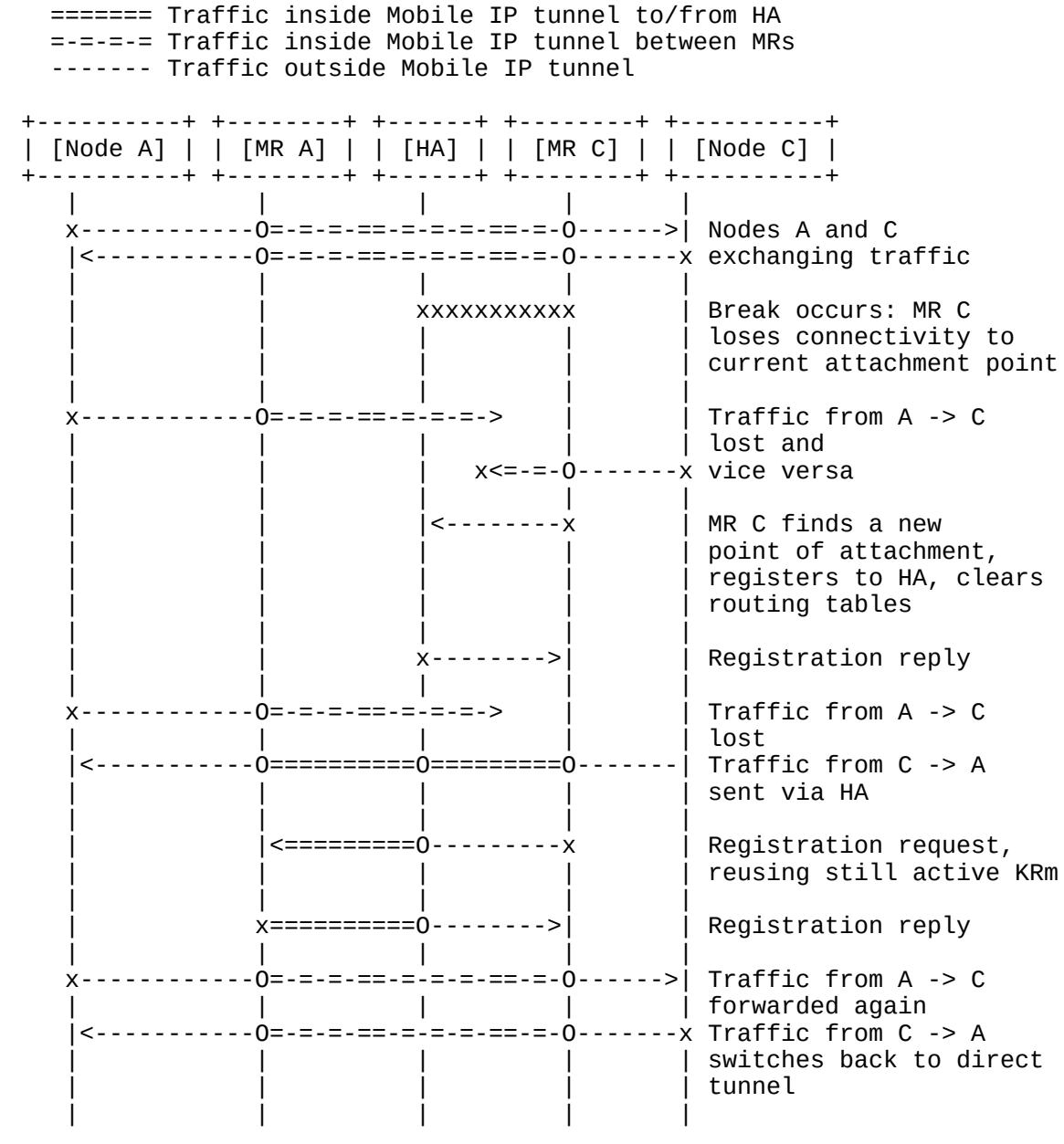
======= Traffic inside Mobile IP tunnel to/from HA
=-=-=-= Traffic inside Mobile IP tunnel between MRs
------- Traffic outside Mobile IP tunnel


[Node A]
[MR A]
[HA]
[MR C]
[Node C]
Mobile Router A is unaware of network C, thus, nothing happens
Mobile Router C notices packet from Net A - begins RO
Return Routability (If no preshared KRms)
CoTI
HoTI
CoT
HoT
KRm between MR A <-> C established
Registration request
Registration request to HA due to MR A being unaware of network C. Solicit bit set.
Registration reply, contains info on Net A
Registration reply, includes MR A's CoA in Care-of-Address extension
Packet from Node C -> A routed to direct tunnel at MR C, based on MR C now knowing MR A's CoA and tunnel being up
Packet from Node A -> C routed to direct tunnel

```
    |            |          |         |       |  at MR A, based on MR A
    |            |          |         |       |  now knowing MR C's CoA
    |            |          |         |       |  and tunnel being up
```

8.3. Handovers

In this example signaling, MR C changes care-of address while Route Optimization between MR A is operating and data is being transferred. Both cases where the handover is graceful ("make before break") and ungraceful ("break before make") occur in similar fashion, except in the graceful version no packets get lost.


======= 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
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 to HA, clears
routing tables
Registration reply
Traffic from A -> C
lost
Traffic from C -> A
sent via HA
Registration request,
reusing still active KRm
Registration reply
Traffic from A -> C
forwarded again
Traffic from C -> A
switches back to direct
tunnel

9. Protocol constants

MAX_NONCE_LIFETIME	240 seconds
MAX_TOKEN_LIFETIME	210 seconds
MAX_RR_BINDING_LIFETIME	420 seconds
MAX_UPDATE_RATE	5 times

10. IANA Considerations

IANA has assigned rules for the existing registries "Mobile IP Messages" and "Mobile IPv4 numbers" in RFC 3344 [RFC3344]. Numbering spaces for Mobile IP messages and for Extensions that may appear in Mobile IP control messages (those sent to and from UDP port number 434) should be modified.

New Mobile IP control message extension and message type values are needed for the messages and extensions listed in Section 5. The Route Optimization authentication processing requires four new message type numbers. In addition, there is a skippable extension which requires it’s own type number. The rest of the new extensions are non-skippable, and grouped under two new types as subtypes. Other type is for extensions in "short" format and other for single extension in "long" extension format.

New Mobile IP registration reply code values are needed for responses from Correspondent Routers. The Route Optimization requires three new reply codes. In addition, a new allocation guideline for "Correspondent Router reply codes" are needed.

The new MIP message types are listed below:

Value	Name
TBA_MIP1	Home-Test Init message
TBA_MIP2	Care-of-Test Init message
TBA_MIP3	Home Test message
TBA_MIP4	Care-of Test message

Table 1: New Values for Mobile IP Message types

The new MIP control message extension types are listed below:

Value	Name
TBA_T1, 128-255	Mobile router Route optimization indication
TBA_T2, 0-127	Route Optimization Extensions
TBA_T3, 0-127	Route Optimization data

Table 2: New Values and Names for Extensions in Mobile IP Control messages

Three new number spaces have been created for the Values and Names for the Sub-Type for Route Optimization-related Extensions. This number spaces are initially defined to hold the following entries, allocated by this document:

Value	Name
TBA_ST1_1	Mobile router Route optimization capability
TBA_ST2_1	Route optimization reply
TBA_ST2_2	Mobile-Correspondent authentication extension
TBA_ST2_3	Care-of address Extension
TBA_ST3_1	Route optimization prefix advertisement

Table 3: New Values and Names for the Sub-type Route Optimization Extension

Note to RFC Editor: this section may be removed on publication as an RFC.

Three new registration reply codes have been created for Code Values for Mobile IP Registration Reply Messages. Following values are added:

Value	Name
TBA_C1	Expired Home nonce Index
TBA_C2	Expired Care-of nonce Index
TBA_C3	Expired nonces
TBA_C4	Concurrent registration

11. Security Considerations

There are two primary security issues: Other relates to return routability check, which establishes that a specific Care-of address is, indeed, managed by a specific Home Address. Other issue is trust relationships and arbitrary router claiming to represent arbitrary network.

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

11.1. Return Routability

The Return Routability check's security has been vetted with Mobile IPv6. There are no large differences apart from requiring a separate ICMP message for connectivity check, and replay attack protection, which in this case uses Mobile IPv4 timestamps in registration request's identification field instead of sequence numbers.

The Return Routability procedure does not establish any kind of state information on the Correspondent router, mitigating 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 solve the issues where arbitrary Correspondent Router can trust an arbitrary Mobile Router that it is indeed the proper route to reach an arbitrary mobile network.

It is assumed that all Mobile Routers have a trust relationship with the Home Agent. Thus, they trust information provided by Home Agent.

The Home Agent provides information matching Home Addresses and network prefixes. Each Mobile Router trusts this information.

Mobile Routers may perform Return Routability procedure between each other. This creates a trusted association between Mobile Router Home Address and Care-of Address. The Mobile Router also claims to represent a specific network. This information is not trustworthy as such.

The claim can be verified by checking the Home Address <-> network prefix information received, either earlier, or due to on-demand request, from the Home Agent. If they match, the Mobile Router'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 decision can be made are out of scope of this

document. The RECOMMENDED general practice is to perform the check.

12. Acknowledgements

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.

[RFC3344] Perkins, C., "IP Mobility Support for IPv4", RFC 3344, August 2002.

[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.

13.2. Informative References

[I-D.ietf-mip4-multiple-tunnel-support] Gundavelli, S., Leung, K., Tsirtsis, G., Soliman, H., and A. Petrescu, "Flow Binding Support for Mobile IPv4", draft-ietf-mip4-multiple-tunnel-support-00 (work in progress), August 2010.

[I-D.ietf-mobileip-optim]

Perkins, C. and D. Johnson, "Route Optimization in Mobile IP", September 2001.

[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.

[RFC3775] Johnson, D., Perkins, C., and J. Arkko, "Mobility Support in IPv6", RFC 3775, June 2004.

Authors' Addresses

Antti Makela
Aalto University
P.O. BOX 13000
FIN-00076 Aalto
FINLAND

Phone: +358 9 451 5590
Email: antti.makela@tkk.fi

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

Email: jouni.nospam@gmail.com