Network Working Group
Internet-Draft
Intended status: Experimental
Expires: August 17, 2014

V. Ermagan
Cisco Systems, Inc.
D. Farinacci
lispers.net
D. Lewis
J. Skriver
F. Maino
Cisco Systems, Inc.
C. White
Logicalelegance, Inc.
February 13, 2014

NAT traversal for LISP
draft-ermagan-lisp-nat-traversal-05.txt

Abstract

This document describes a mechanism for IPv4 NAT traversal for LISP tunnel routers (xTR) and LISP Mobile Nodes (LISP-MN) behind a NAT device. A LISP device both detects the NAT and initializes its state. Forwarding to the LISP device through a NAT is enabled by the LISP Re-encapsulating Tunnel Router (RTR) network element, which acts as an anchor point in the data plane, forwarding traffic from unmodified LISP devices through the NAT.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on August 17, 2014.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction 2
2. Definition of Terms 3
3. Basic Overview 4
4. LISP RTR Message Details 5
4.1. Info-Request Message 5
4.2. LISP Info-Reply 7
4.3. LISP Map-Register Message 9
4.4. LISP Map-Notify 10
4.5. LISP Data-Map-Notify Message 11
5. Protocol Operations 13
5.1. xTR Processing 13
5.1.1. ETR Registration 14
5.1.2. Map-Request and Map-Reply Handling 15
5.1.3. xTR Sending and Receiving Data 16
5.2. Map-Server Processing 16
5.3. RTR Processing 17
5.3.1. RTR Data Forwarding 19
5.4. Example 20
6. Security Considerations 23
6.1. Acknowledgments 24
7. IANA Considerations 24
8. Normative References 24
Authors' Addresses 24

1. Introduction

The Locator/ID Separation Protocol [LISP] defines a set of functions for encapsulating routers to exchange information used to map from Endpoint Identifiers (EIDs) to routable Routing Locators (RLOCs). The assumption that the LISP Tunnel Routers are reachable at their RLOC breaks when a LISP device is behind a NAT. LISP relies on the xTR being able to receive traffic at its RLOC on destination port 4341. However nodes behind a NAT are only reachable through the NAT's public address and in most cases only after the appropriate mapping state is set up in the NAT. A NAT traversal mechanism is needed to make the LISP device behind a NAT reachable.

This document introduces a NAT traversal mechanism for LISP. Two new LISP control messages - LISP Info-Request and LISP Info-Reply - are introduced in order to detect whether a LISP device is behind a NAT, and discover the global IP address and global ephemeral port used by the NAT to forward LISP packets sent by the LISP device. A new LISP component, the LISP Re-encapsulating Tunnel Router (RTR), acts as a re-encapsulating LISP tunnel router [LISP] to pass traffic through the NAT, to and from the LISP device. A modification to how the LISP Map-Register messages are sent allows LISP device to initialize NAT state to use the RTR services. This mechanism addresses the scenario where the LISP device is behind the NAT, but the associated Map-Server [LISP-MS] is on the public side of the NAT.

2. Definition of Terms

LISP Info-Request: A LISP control message sent by a LISP device to its Map-Server.

LISP Info-Reply: A LISP control message sent by a Map Server to a LISP device in response to an Info-Request control message.

LISP Re-encapsulating Tunnel Router (RTR): An RTR is a re-encapsulating LISP Router (see section 8 of the main LISP specification) [LISP]. One function that an RTR provides is enabling a LISP device to traverse NATs.

LISP Data-Map-Notify: A LISP Map-Notify message encapsulated in a LISP data header.

LISP xTR-ID A 128-bit field that, together with a site-ID, can be appended at the end of a Map-Register or Map-Notify message. An xTR-ID is used as a unique identifier of the xTR that is sending the Map-Register and is especially useful for identifying multiple xTRs serving the same site/EID-prefix. A value of all zeros indicate the xTR-ID is unspecified.

LISP site-ID A 64-bit field that, together with a xTR-ID, can be appended at the end of a Map-Register or Map-Notify message. A site-ID is used as a unique identifier of a group of xTRs belonging to the same site. A value of 0 indicate the site-ID is unspecified.

NAT: "Network Address Translation is a method by which IP addresses are mapped from one address realm to another, providing transparent routing to end hosts". "Traditional NAT would allow hosts within a private network to transparently access hosts in the external network, in most cases. In a traditional NAT, sessions are uni-directional, outbound from the private network."

--RFC 2663 [NAT]. Basic NAT and NAPT are two varieties of traditional NAT.

Basic NAT: "With Basic NAT, a block of external addresses are set aside for translating addresses of hosts in a private domain as they originate sessions to the external domain. For packets outbound from the private network, the source IP address and related fields such as IP, TCP, UDP and ICMP header checksums are translated. For inbound packets, the destination IP address and the checksums as listed above are translated." --RFC 2663[NAT].

NAPT: "NAPT extends the notion of translation one step further by also translating transport identifier (e.g., TCP and UDP port numbers, ICMP query identifiers). This allows the transport identifiers of a number of private hosts to be multiplexed into the transport identifiers of a single external address. NAPT allows a set of hosts to share a single external address. Note that NAPT can be combined with Basic NAT so that a pool of external addresses are used in conjunction with port translation." --RFC 2663[NAT]. Transport identifiers of the destination hosts are not modified by the NAPT.

In this document the general term NAT is used to refer to both Basic NAT and NAPT.

While this document specifies LISP NAT Traversal for LISP tunnel routers, a LISP-MN can also use the same procedure for NAT traversal. The modifications attributed to a LISP-Device, xTR, ETR, and ITR must be supported by a LISP-MN where applicable, in order to achieve NAT traversal for such a LISP node. A NAT traversal mechanism for LISP-MN is also proposed in [NAT-MN].

For definitions of other terms, notably Map-Request, Map-Reply, Ingress Tunnel Router (ITR), and Egress Tunnel Router (ETR), please consult the LISP specification [LISP].

3. Basic Overview

There are two attributes of a LISP device behind a typical NAT that requires special consideration in LISP protocol behavior in order to make the device reachable. First, the RLOC assigned to the device is typically not globally unique nor globally routable. Second, the NAT likely has a restrictive translation table and forwarding policy, requiring outbound packets to create state before the NAT accepts inbound packets. This section provides an overview of the LISP NAT traversal mechanism which deals with these conditions. The following sections specify the mechanism in more detail.

When a LISP device receives a new RLOC and wants to register it with the mapping system, it needs to first discover whether it is behind a NAT. To do this, an ETR queries its Map-Server to discover the ETR's translated global RLOC and port via the two new LISP messages: Info-Request and Info-Reply. Once an ETR detects that it is behind a NAT, it uses a LISP Re-encapsulating Tunnel Router (RTR) entity as an anchor point for sending and receiving data plane traffic through the NAT device. The ETR registers the RTR RLOC(s) to its Map-Server using the RTR as a proxy for the Map-Register message. The ETR encapsulates the Map-Register message in a LISP ECM header destined to the RTR's RLOC. The RTR strips the LISP ECM header, re-originates the Map-Register message, and sends it to the Map-Server. This initializes state in the NAT device so the ETR can receive traffic on port 4341 from the RTR. The ETR also registers the RTR RLOC as the RLOC where the ETR EID prefix is reachable. As a result, all packets destined to the ETR's EID will go to its RTR. The RTR will then re-encapsulate and forward the ETR's traffic via the existing NAT state to the ETR.

Outbound LISP data traffic from the xTR is also encapsulated to the RTR, where the RTR de-capsulates the LISP packets, and then re-encapsulates them or forwards them natively depending on their destination.

In the next sections these procedures are discussed in more detail.

4. LISP RTR Message Details

The main modifications in the LISP protocol to enable LISP NAT traversal via an RTR include: (1) two new messages used for NAT discovery (Info-Request and Info-Reply), and (2) encapsulation of two LISP control messages (Map-Register and Map-Notify) between the xTR and the RTR. Map-Register is encapsulated in an ECM header while Map-Notify is encapsulated in a LISP data header (Data-Map-Notify). This section describes the message formats and details of the Info-Request, Info-Reply, and Data-Map-Notify messages, as well as encapsulation details and minor changes to Map-Register and Map-Notify messages.

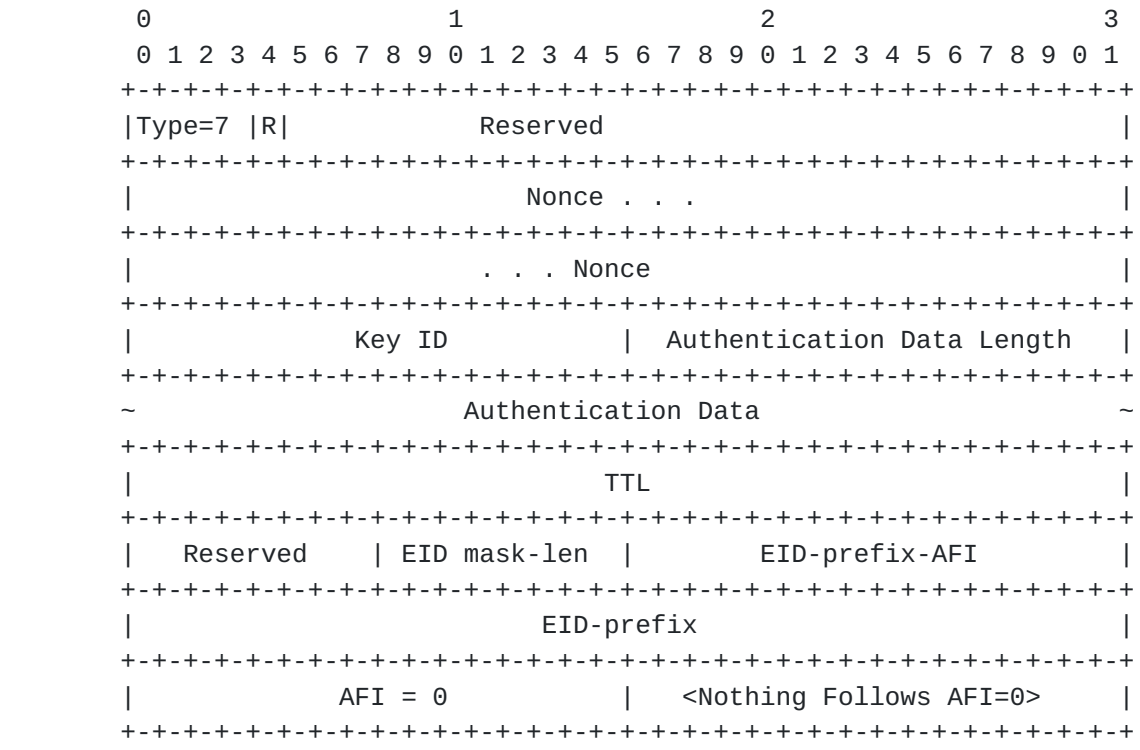
4.1. Info-Request Message

An ETR sends an Info-Request message to its Map-Server in order to

1. detect whether there is a NAT on the path to its Map-Server

2. obtain a list of RTR RLOCs that can be used for LISP data plane NAT traversal.

An Info-Request message is a LISP control message, its source port is chosen by the xTR and its destination port is set to 4342.



LISP Info-Request Message Format

Type: 7 (Info-Request)

R: R bit indicates this is a reply to an Info-Request (Info-Reply). R bit is set to 0 in an Info-Request. When R bit is set to 0, the AFI field (following the EID-prefix field) must be set to 0. When R bit is set to 1, the packet contents follow the format for an Info-Reply, as described below.

Reserved: Must be set to 0 on transmit and must be ignored on receipt.

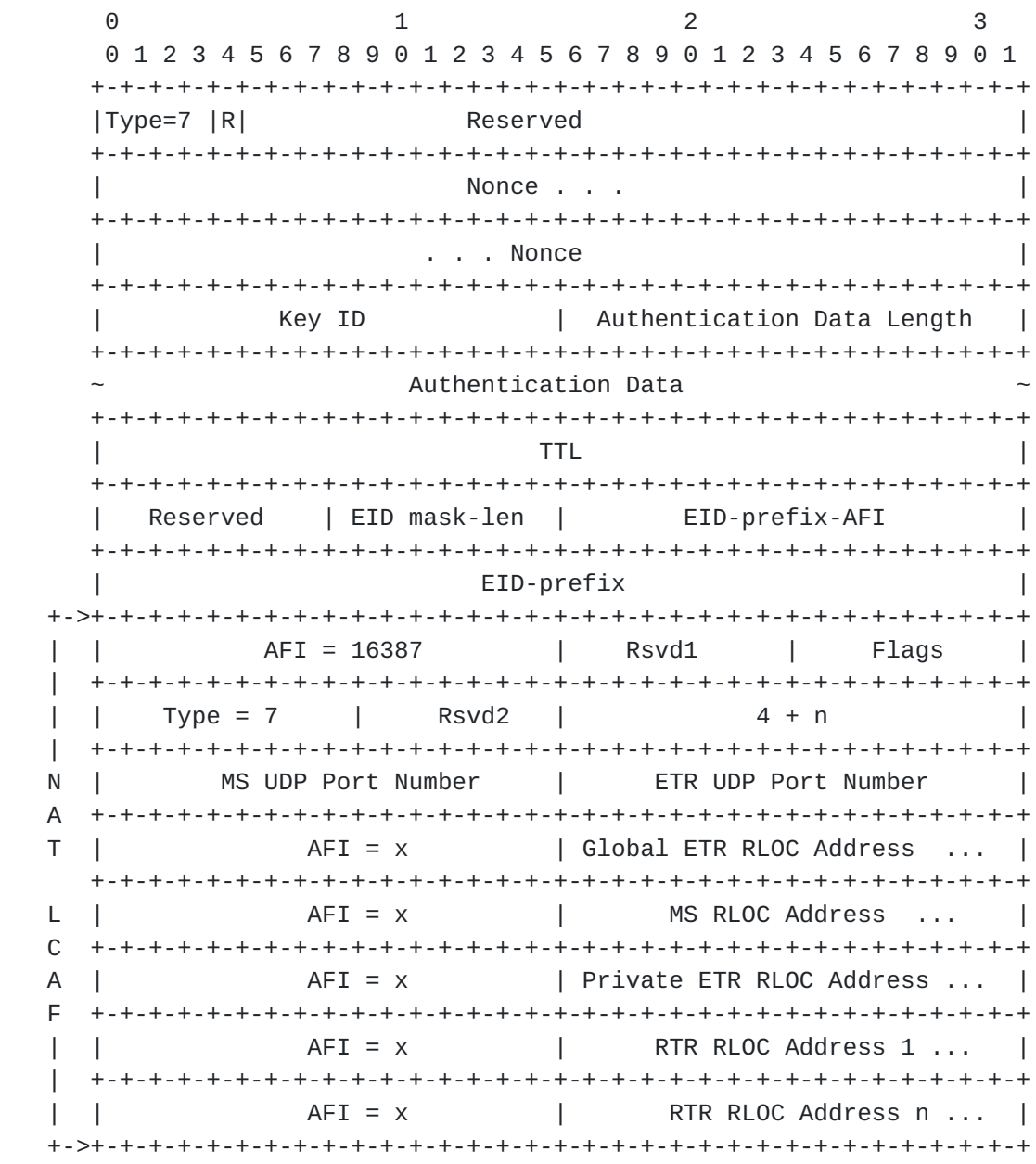
TTL: The time in minutes the recipient of the Info-Reply will store the RTR Information.

Nonce: An 8-byte random value created by the sender of the Info-Request. This nonce will be returned in the Info-Reply. The nonce SHOULD be generated by a properly seeded pseudo-random (or strong random) source.

Descriptions for other fields can be found in the Map-Register section of the main LISP draft [LISP]. Field descriptions for the LCAF AFI = 0 can be found in the LISP LCAF draft [LCAF] .

4.2. LISP Info-Reply

When a Map-Server receives an Info-Request message, it responds with an Info-Reply message. The Info-Reply message source port is 4342, and destination port is taken from the source port of the triggering Info-Request. Map-Server fills the NAT LCAF (LCAF Type = 7) fields according to their description. The Map-Server uses AFI=0 for the Private ETR RLOC Address field in the NAT LCAF.



LISP Info-Reply Message Format

Type: 7 , R = 1, (Info-Reply)

The format is similar to the Info-Request message. See Info-Request section for field descriptions. Field descriptions for the NAT LCAF section can be found in the LISP LCAF draft [LCAF] .

4.3. LISP Map-Register Message

The third bit after the Type field in the Map-Register message is allocated as "I" bit. I bit indicates that a 128 bit xTR-ID and a 64 bit site-ID field are present at the end of the Map-Register message. If an xTR is configured with an xTR-ID or site-ID, it MUST set the I bit to 1 and include its xTR-ID and site-ID in the Map-Register messages it generates. If either the xTR-ID or site-ID is not configured an unspecified value is encoded for whichever ID that is not configured.

xTR-ID is a 128 bit field at the end of the Map-Register message, starting after the final Record in the message. The xTR-ID is used to identify the intended recipient xTR for a Map-Notify message, especially in the case where a site has more than one xTR. A value of all zeros indicate that an xTR-ID is not specified, though encoded in the message. This is useful in the case where a site-ID is specified, but no xTR-ID is configured. When a Map-Server receives a Map-Register with an xTR-ID specified (I bit set and xTR-ID has a non-zero value), it MUST copy the XTR-ID from the Map-Register to the associated Map-Notify message. When a Map-Server is sending an unsolicited Map-Notify to an xTR to notify the xTR of a change in locators, the Map-Server must include the xTR-ID for the intended recipient xTR, if it has one stored locally.

site-ID is a 64 bit field at the end of the Map-Register message, following the xTR-ID. site-ID is used by the Map-Server receiving the Map-Register message to identify which xTRs belong to the same site. A value of 0 indicate that a site-ID is not specified, though encoded in the message. When a Map-Server receives a Map-Register with a site-ID specified (I bit set and site-ID has non-zero value), it must copy the site-ID from the Map-Register to the associated Map-Notify message. When a Map-Server is sending an unsolicited Map-Notify to an xTR to notify the xTR of a change in locators, the Map-Server must include the site-ID for the intended recipient xTR, if it has one stored locally.

A LISP device that sends a Map-Register to an RTR must encapsulate the Map-Register message using an Encapsulated Control Message (ECM) [LISP]. The 6th bit in the ECM LISP header is allocated as the "R" bit. The R bit indicates that the encapsulated Map-Register is to be processed by an RTR. The 7th bit in the ECM header is allocated as the "N" bit. The N bit indicates that this Map-REgister is being relayed by an RTR. When an RTR relays the ECM-ed Map-Register to a Map-Server, the N bit must be set to 1.

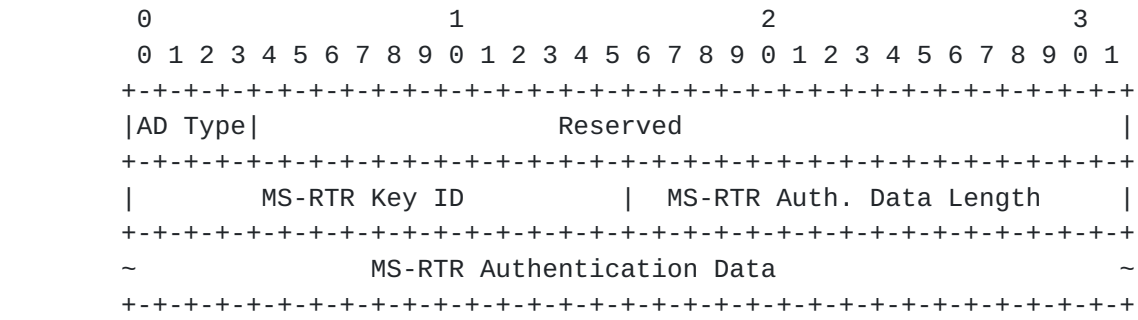
The outer header source RLOC of the ECM is set to the LISP device's local RLOC, and the outer header source port is set to 4341. The

outer header destination RLOC and port are set to RTR RLOC and 4342 respectively. The inner header source RLOC is set to LISP device's local RLOC, and the inner source port is picked at random. The inner header destination RLOC is set to the xTR's Map-Server RLOC, and inner header destination port is set to 4342.

4.4. LISP Map-Notify

The first bit after the Type field in a Map-Notify message is allocated as the "I" bit. I bit indicates that a 128 bit xTR-ID and 64 bit site-ID field is present at the end of the Map-Notify message, following the final Record in the Map-Notify (See Section 4.3 for details on xTR-ID and site-ID). A Map-Server MUST set the I bit in a Map-Notify and include the xTR-ID and/or site-ID of the intended recipient xTR if the associated Map-Register has an xTR-ID and/or site-ID specified, or when the Map-Server has previously cached an xTR-ID and/or site-ID for the destination xTR.

A LISP device that sends a Map-Notify to an RTR must encapsulate the Map-Notify message using an ECM. the 6th bit in the ECM LISP header, allocated as the "R" bit, must be set when the encapsulated Map-Notify is to be processed by an RTR. If the S bit is also set in the Map-Notify ECM header, it indicates that additional MS-RTR authentication data is included after the LISP header in the ECM. If the I bit is also set in the Map-Notify, the xTR-ID and site-ID fields are included in the Map-Notify. If a Map-Server receiving an ECM-ed Map-Register has a shared key associated with the sending RTR, it must generate a Map-Notify message with the S bit in the ECM header set to 1, and with the additional MS-RTR authentication related fields described below.



Changes to LISP Map-Notify Message

AD Type: 2 (RTR Authentication Data)

MS-RTR Key ID: A configured ID to find the configured Message Authentication Code (MAC) algorithm and key value used for the

authentication function. See [LISP] section 14.4 for code point assignments.

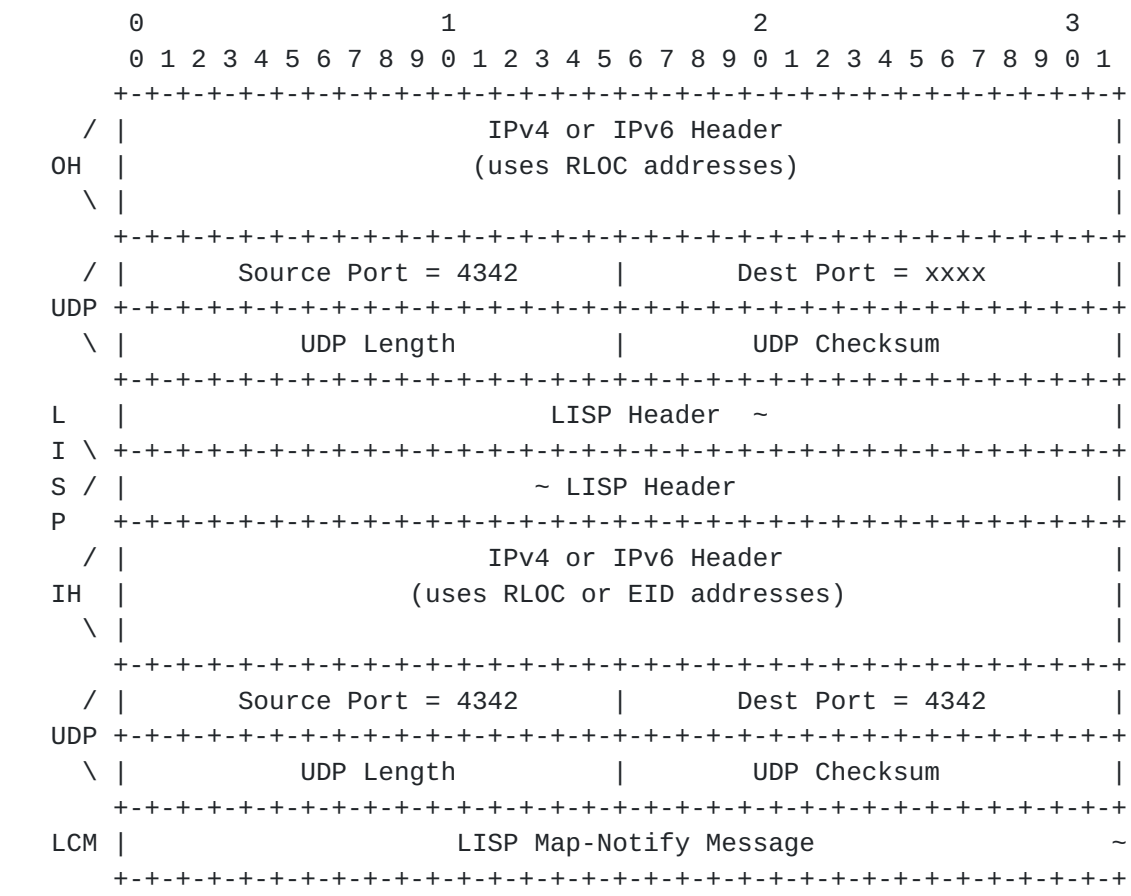
MS-RTR Authentication Data Length: The length in bytes of the MS-RTR Authentication Data field that follows this field. The length of the Authentication Data field is dependent on the Message Authentication Code (MAC) algorithm used. The length field allows a device that doesn't know the MAC algorithm to correctly parse the packet.

MS-RTR Authentication Data: The message digest used from the output of the Message Authentication Code (MAC) algorithm. The entire Map-Notify payload is authenticated. After the MAC is computed, it is placed in this field. Implementations of this specification MUST support HMAC-SHA-1-96 [RFC2404] and SHOULD support HMAC-SHA-256-128 [RFC6234].

For a full description of all fields in the Map-Notify message refer to Map-Notify section in the main LISP draft [LISP].

4.5. LISP Data-Map-Notify Message

When an RTR receives an ECM-ed Map-Notify message with R bit in the ECM header set to 1, it has to relay the Map-Notify payload to the registering LISP device. After removing the ECM header and processing the Map-Notify message as described in Section 5.3, the RTR encapsulates the Map-Notify in a LISP data header and sends it to the associated LISP device. This Map-Notify inside a LISP data header is referred to as a Data-Map-Notify message.



LISP Data-Map-Notify Message

In a Data-Map-Notify, the outer header source RLOC is set to the RTR's RLOC that was used in the associated Map-Register. This is previously cached by the RTR. The outer header source port is set to 4342. The outer header destination RLOC and port are filled based on the translated global RLOC and port of the registering LISP device previously stored locally at the RTR. The inner header source address is Map-Server's RLOC, and inner header source port is 4342. The inner header destination address is set to the LISP device's local RLOC also previously cached by the RTR (See Section 5.3 for details.). The inner header destination port is 4342.

Since a Data-Map-Notify is a control message encapsulated in a LISP data header, a special Instance ID is used as a signal for the xTR to trigger processing of the control packet inside the data header. The Instance ID value 0xFFFFFF is reserved for this purpose. The Instance ID field in a Data-Map-Notify must be set to 0xFFFFFF.

5. Protocol Operations

There are two main steps in the NAT traversal procedure. First, the ETR's translated global RLOC must be discovered. Second, the NAT translation table must be primed to accept incoming connections. At the same time, the Map-Server and the RTR must be informed of the ETR's translated global RLOC including the translated ephemeral port number(s) at which the Map-Server and RTR can reach the LISP device.

5.1. xTR Processing

Upon receiving a new local RLOC, an ETR first has to detect whether the new RLOC is behind a NAT device. For this purpose the ETR sends an Info-Request message to its Map-Server in order to discover the ETR's translated global RLOC as it is visible to the Map-Server. The ETR uses its new local RLOC as the source RLOC of the message. The Map-Server, after authenticating the message, responds with an Info-Reply message. The Map-Server includes the source RLOC and port from the Info-Request message in the Global ETR RLOC Address and ETR UDP Port Number fields of the Info-Reply. The Map Server also includes the destination RLOC and port number of the Info-Request message in the MS RLOC Address and MS UDP Port Number fields of the Info-Reply. In addition, the Map-Server provides a list of RTR RLOCs that the ETR may use in case it needs NAT traversal services. The source port of the Info-Reply is set to 4342 and the destination port is copied from the source port of the triggering Info-Request message.

Upon receiving the Info-Reply message, the ETR compares the source RLOC and source port used for the Info-Request message with the Global ETR RLOC Address and ETR UDP Port Number fields of the Info-Reply message. If the two are not identical, the ETR concludes that its new local RLOC is behind a NAT and that it requires an RTR for NAT traversal services in order to be reachable at that RLOC. An ETR behind other statefull devices (e.g. statefull firewalls) may also use an RTR and the procedure specified here for traversing the statefull device. Detecting existence of such devices are beyond scope of this document.

In the case where an xTR has multiple RLOCs, info-Requests must be sent per each RLOC and the state and processes described below must be followed per each RLOC. The RLOCs included in Map-Register messages in such cases will be the union of the locators resulting from the process below per each RLOC of the xTR, according to the specifics of that interface (whether it is behind the NAT or not).

If there is no NAT on the path identified by an info-Request and an Info-Reply, the ETR registers the associated RLOC with its Map-Server as described in the main LISP draft [LISP].

5.1.1. ETR Registration

Once an ETR has detected that it is behind a NAT, based on local policy the ETR selects one (or more) RTR(s) from the RTR RLOCs provided in the Info-Reply and initializes state in the NAT device in order to receive LISP data traffic on UDP port 4341 from the selected RTR. To do so, the ETR sends a Map-Register encapsulated in an ECM header to the selected RTR(s). The Map-Register message is created as specified in [LISP]. More specifically, the source RLOC of the Map-Register is set to ETR's local RLOC, while the destination RLOC is set to the ETR's Map-Server RLOC, and destination port is set to 4342. The ETR sets the M bit in Map-Register to 1, and it includes the selected RTR RLOC(s) as the locators in the Map-Register message. The ETR can also include its local RLOCs as locators in the Map-Register, including weight and priorities, while setting the R bit to 0 for each local RLOC. This can be used by the RTR for load balancing when forwarding data to a multi-homed xTR behind a NAT. The R bit is set to 1 for all RTR locators included in the Map-Register. The ETR must also set the I bit in the Map-Register message to 1 and include its xTR-ID in t he corresponding field. In the ECM header of this Map-Register the source RLOC is set to ETR's local RLOC and the source port is set to 4341, while the destination RLOC is the RTR's RLOC and the destination port is set to LISP control port 4342. The R bit in the ECM header is also set to 1, to indicate that this EDCM-ed Map-Register is to be processed by an RTR.

This ECM-ed Map-Register is then sent to the RTR. The RTR removes the EMC header, re-originates the Map-Register message, encapsulates the new Map-Register in a new ECM header with R bit set to 0, and sends it to the associated Map-Server. The RTR then encapsulates the corresponding Map-Notify message in a LISP data header (Data-Map-Notify) and sends it back to the xTR.

Upon receiving a Data-Map-Notify from the RTR, the ETR must strip the outer LISP data header, and process the inner Map-Notify message as described in [LISP]. Since outer header destination port in Data-Map-Notify is set to LISP data port 4341, the Instance ID 0xFFFFFF in the LISP header of the Data-Map-Notify is used by the ETR to detect and process the Data-Map-Notify as a control message encapsulated in a LISP data header. While processing the Data-Map-Notify, the xTR also stores the RTR RLOC(s) as its data plane proxy, by storing a default map-cache entry with the RTR RLOC(s) as its locator set. The xTR may map the EID prefix 0/0 to this RTR RLOC(s). This results in the xTR encapsulating all LISP data plane traffic to this RTR. At this point the registration and state initialization is complete and the xTR can use the RTR services. The state created in the NAT device based on the ECM-ed Map-Register and corresponding Data-Map-Notify is used by the xTR behind the NAT to send and receive LISP

control packets to/from the RTR, as well as for receiving LISP data packets form the RTR.

If ETR receives a Data-Map-Notify with a xTR-ID specified, but the xTR-ID is not equal to its local xTR-ID, it must log this as an error. The ETR should discard such Data-Map-Notify message.

The ETR must periodically send ECM-ed Map-Register messages to its RTR in order to both refresh its registration to the RTR and the Map-Server, and as a keep alive in order to preserve the state in the NAT device. RFC 2663 [NAT] points out that the period for sending the keep alives can be set to default value of two minutes, however since shorter timeouts may exist in some NAT deployments, the interval for sending periodic ECM-ed Map-Registers must be configurable.

5.1.2. Map-Request and Map-Reply Handling

The ETR is in control of how to handle the Map-Requests and Map-Replies. If the ETR wants the Map-Server to proxy-reply as described in [LISP], it can register the RTR RLOC(s) as its locator via the ECM-ed Map-Register message. In this case, if the proxy bit is set in the Map-Register, the Map-Server will proxy reply with RTR's RLOC to all Map-Requests for the ETR. As a result all traffic for the ETR is encapsulated to its RTR(s).

If the proxy bit in the ECM-ed Map-Register message is not set, and the ETR chooses to receive Map-Requests, the ETR must also initiate and preserve state in the NAT device to receive LISP control packets from its Map-Server. To do this, the ETR must periodically send Info-Request messages to its Map-Server, and receive Info-Reply messages from the Map-Server. As pointed in RFC 2663 [NAT] the default assumption of two minute period for session lifetime can be used, however since shorter timeouts may exist in some NAT deployments, the interval for sending periodic Info-Requests must be configurable. Furthermore, the ETR must also provide its Map-Server with the ETR's translated global RLOC and port as visible to the Map-Server. To do this, ETR includes a copy of the NAT LCAF section of the Info-Reply message as one of the locators in its Map-Register along with the RTR(s) RLOC(s). The ETR can set the priorities of RTR RLOC(s) in this Map-Register to 255, resulting in the Map Server encapsulating Map-Requests to the ETR's translated global RLOC and port so it can receive them through the NAT device.

If an ETR behind a NAT chooses to receive Map-Requests from the Map-Server, it must send Map-Replies to requesting ITRs. Note that this configuration will result in excessive state in the NAT device and is not recommended. ETR must include its RTR RLOC(s) as its locator set in the Map-Reply in order to receive data through the NAT device.

When an ITR behind a NAT is encapsulating outbound LISP traffic, it must use its RTR RLOC as the locator for all destination EIDs that it wishes to send data to. As such, the ITR does not need to send Map-Requests for the purpose of finding EID-to-RLOC mappings. For RLOC-probing, the periodic ECM-ed Map-Register and Data-Map-Notify messages between xTR and RTR can also serve the purpose of RLOC probes. However, if RLOC-probing is used, no changes are required to the RLOC-probing specification in [LISP], except that the LISP device behind a NAT only needs to probe the RTR's RLOC.

5.1.3. xTR Sending and Receiving Data

When a Map-Request for a LISP device behind a NAT is received by its Map-Server or the LISP device itself, the Map-Server, or the LISP device (ETR), responds with a Map-Reply including RTR's RLOC as the locator for the requested EID. As a result, all LISP data traffic destined for the ETR's EID behind the NAT is encapsulated to its RTR. The RTR re-encapsulates the LISP data packets to the ETR's translated global RLOC and port number so the data can pass through the NAT device and reach the ETR. As a result the ETR receives LISP data traffic with outer header destination port set to 4341 as specified in [LISP].

For sending outbound LISP data, an ITR behind a NAT must use the RTR RLOC as the locator for all EIDs that it wishes to send data to according to the installed default map-cache entry. The ITR then encapsulates the LISP traffic in a LISP data header with outer header destination set to RTR RLOC and outer header destination port set to 4341. This may create a secondary state in the NAT device. ITR must set the outer header source port in all egress LISP data packets to a random but static port number in order to avoid creating excessive state in the NAT device.

If the ITR and ETR of a site are not collocated, the RTR RLOC must be configured in the ITR via an out-of-band mechanism. Other procedures specified here would still apply.

5.2. Map-Server Processing

Upon receiving an Info-Request message a Map-Server first verifies the authenticity of the message. Next the Map-Server creates an Info-Reply message and copies the source RLOC and port number of the Info-Request message to the Global ETR RLOC Address and ETR UDP Port Number fields of the Info-Reply message. The Map-Server also includes a list of RTR RLOCs that the ETR may use for NAT traversal services. The Map-Server sends the Info-Reply message to the ETR, by setting the destination RLOC and port of the Info-Reply to the source

RLOC and port of the triggering Info-Request. The Map-Server sets the source port of the Info-Reply to 4342.

Upon receiving an ECM-ed Map-Register message with the N bit in the ECM header set to 1, the Map-Server removes the ECM header and if the M bit in the Map-Register is set, the Map-Server processes the Map-Register message and generates the resulting Map-Notify as described in [LISP]. The Map-Server encapsulates the Map-Notify in an ECM header and sets the R bit in the ECM header to 1. This indicates that the ECM-ed Map-Notify is to be processed by an RTR. If the Map-Server has a shared secret configured with the RTR sending the Map-Register, the Map-Server also sets the S bit in the ECM header of the Map-Notify and includes the MS-RTR authentication data after the ECM LISP header. See Security Considerations Section for more details. If the I bit is set in the Map-Register message, the Map-Server also locally stores the xTR-ID from the Map-Register, and sets the I bit in the corresponding Map-Notify message and includes the same xTR-ID in the Map-Notify. The ECM-ed Map-Notify is then sent to the RTR sending the corresponding Map-Register.

If a Map-Server is forwarding Map-Requests to an ETR which has registered its RLOC in a NAT LCAF, Map-Server must use the ETR Global RLOC Address and ETR UDP Port as the destination RLOC and port for outer header of the encapsulated Map-Requests. If more than one NAT LCAF is registered for the same EID prefix, the Map-Server must use the NAT LCAF corresponding to the RLOC of this Map-Server.

5.3. RTR Processing

Upon receiving an ECM-encapsulated Map-Register with the R bit set in the ECM header, the RTR creates a map-cache entry for the EID-prefix that was specified in the Map-Register message. The RTR stores the outer header source RLOC and outer header source port, the outer header destination RLOC (RTR's own RLOC), the inner header source RLOC (xTR's local RLOC), the xTR-ID, the weight and priority associated with the xTR's local RLOC that was used to send this Map-Register if present, and the nonce field of the Map-Register in this local map-cache entry. The RTR uses the inner header source address to identify which xTR local RLOC (R bit =0) was used by the xTR to send this Map-Register. The outer header source RLOC and outer header source port is the ETR's translated global RLOC and port number visible to the RTR. Once the registration process is complete, this map-cache entry can be used to send LISP data traffic to the ETR. The inner header source RLOC of the Map-Register is the ETR's local RLOC behind the NAT, and the outer header destination RLOC is the RTR's RLOC used by the ETR. The RTR can later use these fields as the inner header destination RLOC and source RLOC correspondingly, for sending data-encapsulated control messages

(Data-Map-Notify) back to the ETR. The nonce field is used for security purposes and is matched with the nonce field in the corresponding Map-Notify message. This map-cache entry is stored as an "unverified" mapping, until the corresponding Map-Notify message is received.

In the cases where the xTR has multiple RLOCs behind the NAT, and requires the RTR to load balance the traffic across those interfaces, the xTR must include the local RLOCs associated with each interface behind the NAT with the R bit in the locator record set to 0 in the ECM-ed Map-Register sent to the RTR. The RTR uses the weight and priority policies of the RLOCs with R=0 in the Map-Register to load balance the traffic from the RTR to the xTR behind the NAT. The RTR compares the RLOCs with the R bit set to 0 in the Map-Register to the inner header source address of the Map-Register to find the matching RLOC that the xTR used to send the Map-Register from. The RTR associates the weight and priority policies of this local RLOC with the NAT-translated RLOC and xTR-ID for this map-cache entry. For all other local RLOCs included in the Map-Register, that the Map-Register is not originating from, the RTR only updates previously cached weight and priority policies if it already has those local RLOCs previously stored for that EID prefix and xTR-ID. In other words, the RTR only adds new local RLOCs and their weight and priority policies to its cache if the Map-Register is actually originating from that RLOC. The TTL for every map-cache is also only updated when a Map-Register is originating from the same RLOC. However, the weight and priorities of all previously cached local RLOCs will be updated by every Map-Register, whether it is originating from that RLOC or not. The xTR-ID is used to define the Merge domain for these RLOCs. In other words, a Map-Register originating from a unique xTR-ID will always overwrite previously stored policies for that xTR-ID. However it does not modify in any way the policies indicated by any other xTR-ID serving the same EID prefix. As a result, in the case of a renumbering or xTR reboot, the xTR uses its unique xTR-ID to send a new Map-Register, overwriting the previously stored policies for that xTR. Using this method the xTR can immediately remove any RLOCs from the RTR cache that are no longer active. In order to implement this, the RTR must compare the list of local RLOCs in the Map-Register (R=0) with the ones it has previously cached associated with the same xTR-ID. If there is any RLOC previously cached that does not appear in the newly received Map-Register, the RTR must remove that RLOC together with the associated translated RLOC and associated policies, because removal of a local (behind-the-NAT) RLOC also invalidates the NAT-ed address associated with it. .

After filling the local map-cache entry, the RTR strips the outer header and extracts the Map-Register message, re-originates the message by rewriting the source RLOC of the Map-Register to RTR's

RLOC, encapsulated in a new ECM header with the R bit set to 0, and N bit set to 1, and sends the ECM-ed Map-Register to destination Map-Server.

Map-Server responds with a ECM-ed Map-Notify message to the RTR.

Upon receiving an ECM-ed Map-Notify message with R bit set to 1 in the ECM header, if the S bit in ECM header is set to 1, RTR uses the MS-RTR Key ID to verify the MS-RTR Authentication Data included after the ECM header. If the MS-RTR authentication fails, the RTR must drop the packet. Once the authenticity of the message is verified, RTR can confirm that the Map-Register message for the ETR with the matching xTR-ID was accepted by the Map-Server. At this point the RTR can change the state of the associated map-cache entry to verified for the duration of the Map-Register TTL.

The RTR then uses the information in the associated map-cache entry to create a Data-Map-Notify message according to the following procedure: RTR rewrites the inner header destination RLOC of the Map-Notify message to ETR's local RLOC. Inner header destination port is 4342. The RTR encapsulates the Map-Notify in a LISP data header, where the outer header destination RLOC and port number are set to the ETR's translated global RLOC and port number. If more than one ETR translated RLOC and port exists in the map-cache entry for the same EID prefix specified in the Map-Notify, the RTR can use the xTR-ID from the Map-Notify to identify which ETR is the correct destination for the Data-Map-Notify. The RTR sets the outer header source RLOC to RTR's RLOC from the map-cache entry and the outer header source port is set to 4342. The RTR also sets the Instance ID field in the LISP header of the Data-Map-Notify to 0xFFFFFF. The RTR then sends the Data-Map-Notify to the ETR.

If the S bit is set to 0 in the ECM header of the Map-Notify, and the RTR has a shared key configured locally with the sending Map-Server, the RTR must drop the packet. If the S bit is set to 0, and the RTR does not have a shared key configured with the associated Map-Server, according to local policy, the RTR may drop the packet. If the Map-Notify with S bit set to 0 is processed, the RTR must match the nonce field from this Map-Notify with the nonce stored in the local map-cache entry with the matching xTR-ID. If the nonces do not match, the RTR must drop the packet.

5.3.1. RTR Data Forwarding

For all LISP data packets encapsulated to RTR's RLOC and outer header destination port 4341, the RTR first verifies whether the source or destination EID is a previously registered EID. If so, the RTR must process the packet according to the following. If the destination or

source EID is not a registered EID, the RTR can drop or process the packets based on local policy.

In the case where the destination EID is a previously registered EID, the RTR must strip the LISP data header and re-encapsulate the packet in a new LISP data header. The outer header RLOCs and UDP ports are then filled based on the matching map-cache entry for the associated destination EID prefix. The RTR uses the RTR RLOC from the map-cache entry as the outer header source RLOC. The outer header source port is set to 4342. The RTR sets the outer header destination RLOC and outer header destination port based on the ETR translated global RLOC and port stored in the map-cache entry. Then the RTR forwards the LISP data packet.

In the case where the source EID is a previously registered EID, the RTR process the packet as if it is a Proxy ETR (PETR). The RTR must strip the LISP data header, and process the packet based on its inner header destination address. The packet may be forwarded natively, it may be LISP encapsulated to the destination ETR, or it may trigger the RTR to send a LISP Map-Request.

5.4. Example

What follows is an example of an ETR initiating a registration of a new RLOC to its Map-Server, when there is a NAT device on the path between the ETR and the Map-Server.

In this example, the ETR (site1-ETR) is configured with the local RLOC of 192.168.1.2. The NAT's global (external) addresses are from 2.0.0.1/24 prefix. The Map-Server is at 3.0.0.1. And one potential RTR has an IP address of 1.0.0.1. The site1-ETR has an EID Prefix of 128.1.0.0/16.

An example of the registration process follows:

1. The Site1-ETR receives the private IP address, 192.168.1.2 as its RLOC via DHCP.

2. The Site1-ETR sends an Info-Request message with the destination RLOC of the Map-Server, 3.0.0.1, and source RLOC of 192.168.1.2. This packet has the destination port set to 4342 and the source port is set to (for example) 5001.

3. The NAT device translates the source IP from 192.168.1.2 to 2.0.0.1, and source port to (for example) 20001 global ephemeral source port.

4. The Map-Server receives and responds to this Info-Request with an Info-Reply message. This Info-Reply has the destination address set to ETR's translated address of 2.0.0.1 and the source address is the Map-Server's RLOC, namely 3.0.0.1. The destination port is 20001 and the source port is 4342. Map-Server includes a copy of the source address and port of the Info-Request message (2.0.0.1:20001), and a list of RTR RLOCs including RTR RLOC 1.0.0.1 in the Info-Reply contents.

5. The NAT translates the Info-Reply packet's destination IP from 2.0.0.1 to 192.168.1.2, and translates the destination port from 20001 to 5001, and forwards the Info-Reply to site1-ETR at 192.168.1.2.

6. The Site1-ETR detects that it is behind a NAT by comparing its local RLOC (192.168.1.2) with the Global ETR RLOC Address in the Info-Reply (2.0.0.2) . Then site1-ETR picks the RTR 1.0.0.1 from the list of RTR RLOCs in the Info-Reply. ETR stores the RTR RLOC in a default map-cache entry to periodically send ECM-ed Map-Registers to.

7. The ETR sends an ECM encapsulated Map-Register to RTR at 1.0.0.1. The outer header source RLOC of this Map-Register is set to 192.168.1.2 and the outer header source port is set to 4341. The outer header destination RLOC and port are set to RTR RLOC at 1.0.0.1 and 4342 respectively. The R bit in ECM header is set to 1. The inner header destination RLOC is set to ETR's Map-Server 3.0.0.1, and the inner header destination port is set to 4342. The inner header source RLOC is set to ETR's local RLOC 192.168.1.2. In the Map-Register message the RTR RLOC 1.0.0.1 appears as the locator set for the ETR's EID prefix (128.1.0.0/16). In this example ETR also sets the Proxy bit in the Map-Register to 1, and sets I bit to 1, and includes its xTR-ID in the Map-Register.

8. The NAT translates the source RLOC in the ECM header of the Map-Register, by changing it from 192.168.1.2 to 2.0.0.2, and translates the source port in the ECM header from 4341 to (for example) 20002, and forwards the Map-Register to RTR.

9. The RTR receives the Map-Register and creates a map-cache entry with the ETR's xTR-ID, EID prefix, and the source RLOC and port of the ECM header of the Map-Register as the locator (128.1.0.0/16 is mapped to 2.0.0.2:20002). RTR also caches the inner header source RLOC of the Map-Register namely 192.168.1.2, and the outer header destination RLOC of the ECM header in the Map-Register (this would be RTR's RLOC 1.0.0.1) to use for sending back a Data-Map-Notify. RTR then removes the outer header, re-

writes the source RLOC of the Map-Register message to its own RLOC 1.0.0.1, adds a new ECM header with R=0, and N=1, and forwards the Map-Register to the destination Map-Server.

10. The Map-Server receives the ECM-ed Map-Register with N bit set to 1, removes the ECM header, and processes it according to [LISP]. Since Map-Server has a shared secret with the sending RTR, after registering the ETR, Map-Server responds with a ECM-ed Map-Notify with the R bit and S bit both set to 1 in the ECM header and including the MS-RTR authentication data. Since the I bit is set in the Map-Register, the Map-Server also sets the I bit in the Map-Notify and copies the xTR-ID from the Map-Register to the Map-Notify. The source address of this Map-Notify is set to 3.0.0.1. The destination is RTR 1.0.0.1, and both source and destination ports are set to 4342.

11. The RTR receives the ECM-ed Map-Notify and verifies the MS-RTR authentication data. The RTR data-encapsulates the Map-Notify and sends the resulting Data-Map-Notify to site1-ETR with a matching xTR-ID. The outer header source RLOC and port of the Data-Map-Notify are set to 1.0.0.1:4342. The outer header destination RLOC and port are retrieved from previously cached map-cache entry in step 9 namely 2.0.0.2:20002. RTR also sets the inner header destination address to site1-ETR's local address namely 192.168.1.2. RTR sets the Instance ID in the LISP header to 0xFFFFFF. At this point RTR marks ETR's EID prefix as "Registered" status and forwards the Data-Map-Notify to ETR.

12. The NAT device translates the destination RLOC and port of the Data-Map-Notify to 192.168.1.2:4341 and forwards the packet to ETR.

13. The Site1-ETR receives the packet with a destination port 4341, and processes the packet as a control packet after observing the Instance ID value 0xFFFFFF in the LISP header. At this point ETR's registration to the RTR is complete.

Assume a requesting ITR in a second LISP (site2-ITR) site has an RLOC of 74.0.0.1. The following is an example process of an EID behind site2-ITR sending a data packet to an EID behind the site1-ETR:

1. The ITR sends a Map-Request which arrives via the LISP mapping system to the ETR's Map Server.

2. The Map-Server sends a Map-Reply on behalf of the ETR, using the RTR's RLOC (1.0.0.1) in the Map-Reply's Locator Set.

3. The ITR encapsulates a LISP data packet with ITR's local RLOC (74.0.0.1) as the source RLOC and the RTR as the destination RLOC (1.0.0.1) in the outer header.

4. The RTR decapsulates the packet, evaluates the inner header against its map-cache and then re-encapsulates the packet. The new outer header's source RLOC is the RTR's RLOC 1.0.0.1 and the new outer header's destination RLOC is the Global NAT address 2.0.0.2. The destination port of the packet is set to 20002 (discovered above during the registration phase) and the source port is 4342.

5. The NAT translates the LISP data packet's destination IP from to 2.0.0.2 to 192.168.1.2, and translates the destination port from 20002 to 4341, and forwards the LISP data packet to the ETR at 192.168.1.2.

6. For the reverse path the ITR uses its local map-cache entry with the RTR RLOC as the default locator and encapsulates the LISP data packets using RTR RLOC, and 4341 as destination RLOC and port. The ITR must pick a random source port to use for all outbound LISP data traffic in order to avoid creating excessive state in the NAT.

6. Security Considerations

By having the RTR relay the ECM-ed Map-Register message from an ETR to its Map-Server, the RTR can restrict access to the RTR services, only to those ETRs that are registered with a given Map-Server. To do so, the RTR and the Map-Server may be configured with a shared key that is used to authenticate the origin and to protect the integrity of the Map-Notify messages sent by the Map Server to the RTR. This prevents an on-path attacker from impersonating the Map-Server to the RTR, and allows the RTR to cryptographically verify that the ETR is properly registered with the Map-Server.

Having the RTR re-encapsulate traffic only when the source or the destination are registered EIDs, protects against the adverse use of an RTR for EID spoofing.

Upon receiving a Data-Map-Notify, an xTR can authenticate the origin of the Map-Notify message using the key that the ETR shares with the Map-Server. This enables the ETR to verify that the ECM-ed Map-Register was indeed forwarded by the RTR to the Map-Server, and was accepted by the Map-Server.

6.1. Acknowledgments

The authors would like to thank Noel Chiappa, Alberto Rodriguez Natal, Lorand Jakab, Albert Cabellos, Dominik Klein, Matthias Hartmann, and Michael Menth for their previous work, feedback and helpful suggestions.

7. IANA Considerations

This document does not request any IANA actions.

8. Normative References

[LCAF] Farinacci, D., Meyer, D., and J. Snijders, "LISP Canonical Address Format (LCAF)", draft-ietf-lisp-lcaf-03 (work in progress), September 2013.

[LISP-MS] Farinacci, D. and V. Fuller, "Locator/ID Separation Protocol (LISP) Map-Server Interface", RFC 6833, January 2013.

[LISP] Farinacci, D., Fuller, V., Meyer, D., and D. Lewis, "Locator/ID Separation Protocol (LISP)", RFC 6830, January 2013.

[NAT-MN] Klein, D., Hartmann, M., and M. Menth, "NAT traversal for LISP mobile node, In Proceedings of the Re-Architecting the Internet Workshop (ReARCH '10).", 2010.

[NAT] Srisuresh, P. and M. Holdrege, "IP Network Address Translator (NAT) Terminology and Considerations", RFC 2663, August 1999.

[RFC1918] Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, February 1996.

[RFC4632] Fuller, V. and T. Li, "Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan", BCP 122, RFC 4632, August 2006.

Authors' Addresses

Vina Ermagan
Cisco Systems, Inc.

Email: vermagan@cisco.com

Dino Farinacci
lispers.net

Email: farinacci@gmail.com

Darrel Lewis
Cisco Systems, Inc.

Email: darlewis@cisco.com

Jesper Skriver
Cisco Systems, Inc.

Email: jesper@cisco.com

Fabio Maino
Cisco Systems, Inc.

Email: fmaino@cisco.com

Chris White
Logicalelegance, Inc.

Email: chris@logicalelegance.com