

Internet Engineering Task Force	F. Hu	
Internet-Draft	J. Luo	
Intended status: Informational	ZTE Corporation	
Expires: April 22, 2010	October 19, 2009	

[TOC](#)

## **ID/Locator Distributed Mapping Server draft-hu-lisp-dht-00.txt**

### **Status of this Memo**

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/1id-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

This Internet-Draft will expire on April 22, 2010.

### **Copyright Notice**

Copyright (c) 2009 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 in effect on the date of publication of this document (<http://trustee.ietf.org/license-info>). Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

### **Abstract**

This draft proposes a DHT based distributed EID-to-RLOC mapping method. The distributed EID-to-RLOC mapping is stored in an overlay, which is different from the data forwarding plane, and made up by DHT rings. The DHT based overlay is deployed in an AS, and the inter-AS communication is through DHT Border Server, which is located in the border of DHT rings. The DHT Border Server is used to flood the cross routing and

forward the cross domain data packet. The DHT based Distributed mapping system is a hybrid push and pull method.

---

## Table of Contents

<a href="#">1.</a>	Introduction
<a href="#">2.</a>	Definition of Terms
<a href="#">3.</a>	DHT based overlay
<a href="#">4.</a>	DHT mapping server Architecture
<a href="#">5.</a>	ID/Locator Mapping Resolution
<a href="#">5.1.</a>	Registration of EID-to-RLOC mapping
<a href="#">5.2.</a>	Update message
<a href="#">5.3.</a>	Leave message
<a href="#">5.4.</a>	Resolving a RLOC for an EID
<a href="#">6.</a>	Routing and forwarding system
<a href="#">6.1.</a>	Intra-AS domain data packet forwarding
<a href="#">6.2.</a>	Inter-AS domain data packet forwarding
<a href="#">7.</a>	Security Considerations
<a href="#">8.</a>	Acknowledgement
<a href="#">9.</a>	References
<a href="#">9.1.</a>	Normative references
<a href="#">9.2.</a>	Informative References
<a href="#">§</a>	Authors' Addresses

---

## 1. Introduction

[TOC](#)

It is common recognized that today's Internet routing and addressing system is facing serious scaling problems, which is much discussed in the Internet Architecture Board (IAB) workshop on Routing and Addressing in Amsterdam. Several proposals have emerged after the workshop. These proposals include LISP [LISP], are based on the idea of "ID/Locator split, and need a mapping mechanism that allows to map identifiers onto locators. Server mapping mechanisms have been proposed [APT][ALT] [LISPDHT].

Differently from [APT], which is a concentration mapping solution and every AS maintains a default server that stores complete EID-to-RLOC mappings, and [ALT], which is in a distributed manner but the EID prefix must be assigned in a hierarchical manner and be aggregated, this draft proposes a DHT based distributed mapping server solution. This document proposes a method of building a logical distributed overlay, which is used to store the EID-to-RLOC in the DHT ring. The overlay is deployed in an AS, and the cross domain communication is through the DHT Border Server which is used to flood the inter-EID prefix and forward the cross domain data packet.

---

## 2. Definition of Terms

[TOC](#)

Endpoint ID (EID): a 32-bit (for IPv4) or 128-bit (for IPv6) value used in the source and destination address fields of the first (most inner) LISP header of a packet. The host obtains a destination EID the same way it obtains a destination address today, for example through a DNS lookup or SIP exchange. Usually, the EID is an IP address. If the host is required to support mobility, the EID should be unique.

Routing Locator (RLOC): the IPv4 or IPv6 address of an egress tunnel router (ETR). It is the output of an EID-to-RLOC mapping lookup. An EID maps to one or more RLOCs.

Ingress Tunnel Router (ITR): a router which accepts an IP packet with a single IP header (more precisely, an IP packet that does not contain a LISP header). The router treats this "inner" IP destination address as an EID and performs an EID-to-RLOC mapping lookup through a mapping service. The router then prepends an "outer" IP header with one of its globally-routable RLOCs in the source address field and the result of the mapping lookup in the destination address field. An ITR maintains a local mapping table that stores some recently used EID-to-RLOC mapping. An ITR also maintains a timer for each EID-to-RLOC mapping. If the timer of an EID-to-RLOC mapping exceeds a predetermined threshold, the ITR will remove the EID-to-RLOC mapping from its local mapping table, and send a mapping leave message to the DHT server. In addition, an ITR may also be an ETR, and vice versa.

Egress Tunnel Router (ETR): a router that accepts an IP packet where the destination address in the "outer" IP header is one of its own RLOCs. The router strips the "outer" header and forwards the packet based on the next IP header found. In general, an ETR receives LISP-encapsulated IP packets from the Internet on one side and sends decapsulated IP packets to site end-systems on the other side

EID-to-RLOC mapping: a binding between an EID and the RLOC-set that can be used to reach the EID. An RLOC-set may contain multiple RLOC, and perhaps the preference to an RLOC.

DHT Server: a node in the DHT system that stores EID-to-RLOC mapping in its main mapping table and backup mapping table.

DHT Border Server: a DHT Server, which is in the border of the LISP DHT Overlay. DHT Border Server not only stores the EID-to-RLOC mapping, but also floods the aggregation EID prefix to other LISP DHT Overlay domains.

---

[TOC](#)

### 3. DHT based overlay

Figure 1 shows the DHT based overlay network, which consists of two parts: one is the internet, and the other is the overlay network. The internet comprises provider networks that provide data forwarding service. The overlay network is on the top of the provider networks, and provides the EID-to-RLOC resolution system, and maintains the EID-to-RLOC mapping system.

DHT based overlay

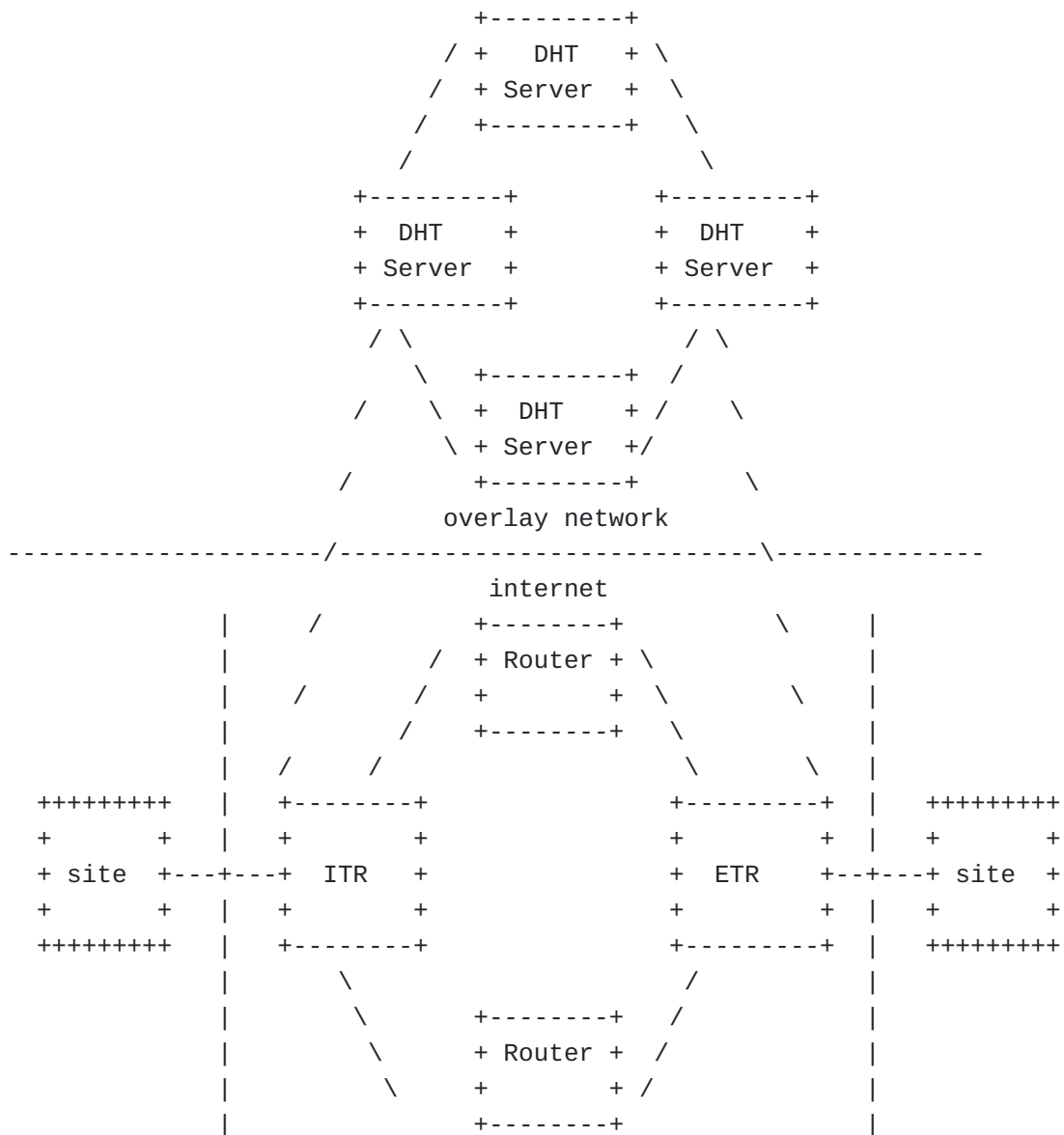


Figure 1

---

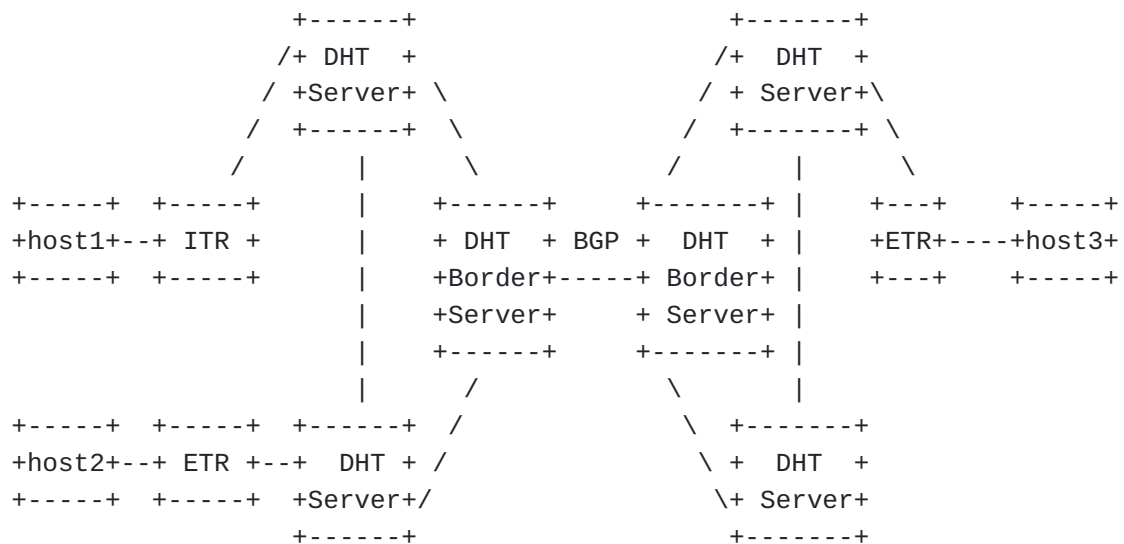
#### 4. DHT mapping server Architecture

[TOC](#)

Figure 2 shows the DHT mapping server Architecture, there are several DHT Servers in the DHT ring, which is used to store the EID-to-RLLOC mapping. The DHT Server maybe a server or router. there is a sepecial DHT Server, which is named as DHT Border Server, and is not only to store the EID-to-RLLOC mapping, but also be used to flood the cross domain routing and forward the cross domain data packet.

---

DHT mapping server Architecture



**Figure 2**

---

#### 5. ID/Locator Mapping Resolution

[TOC](#)[TOC](#)

### 5.1. Registration of EID-to-RLOC mapping

Whenever an end host attaches to an ETR, it should obtain a locator from the ETR and register this EID-to-RLOC mapping to the resolution system. The Registration of EID-to-RLOC mapping is as the following:

- (1) The end host sends a message to its ETR to indicate that it will access the provider networks through the ETR. The message includes the EID of the end host.
- (2) When the ETR receives this message, it first finds in its local mapping cache whether there is a record of the EID-to-RLOC mapping for the EID. If it does, the ETR should check whether the locator in the existing EID-to-RLOC mapping that belongs to the ETR or not. If there is not an EID-to-RLOC mapping for the EID at the ETR<sub>i</sub>'s local cache or the existing locator does not belong to the ETR, the ETR should assign a locator for the EID and stores the new EID-to-RLOC mapping to its local mapping cache. In addition, the ETR sets a timer for the EID-to-RLOC mapping and sends this mapping to the default DHT Server.
- (3) When the default DHT Server receives an EID-to-RLOC mapping for an EID, it hashes the EID to the mapping system. The destination DHT stores the EID-to-RLOC mapping in its mapping table.

---

### 5.2. Update message

[TOC](#)

The end host sends a message periodically to its default ETR. When the ETR receives the message, it finds in its local mapping cache. If it done, ETR updates the timer for the EID-to-RLOC mapping and sends an update message to the default DHT Server of the DHT ring system.

---

### 5.3. Leave message

[TOC](#)

When an end host leaves from an ETR, it may send a message to the ETR to indicate this leave. Or, the ETR may detect the leave of an end node. In any case, the ETR should remove the EID-to-RLOC mapping and send a leave message to the DHT Server system.

---

### 5.4. Resolving a RLOC for an EID

[TOC](#)

When an ITR receives a packet with a destination EID, it should look up in its local cache whether or not there is an EID-to-RLOC mapping for the EID. If it can not finds a RLOC for the EID, which means that it is

the first packet to the destination EID. The process of the resolving is in the following:

- (1) The ITR sends a mapping request to the default DHT Server of the DHT mapping system. The mapping request contains the destination EID information, the data packet, and possibly some signature used for security.
- (2) When the default DHT Server receives the request, it hashes the EID in the mapping system.
- (3) When the destination DHT Server receives the mapping request, it finds the EID-to-RLOC mapping in its main mapping table. If it does, the DHT Server encapsulates the data packet included in the mapping request and sends the new data packet to the ETR according the record of the RLOC. The DHT Server sends a mapping reply to the source ITR.
- (4) When the ITR receives the mapping reply message, it stores this mapping into its local mapping cache. When subsequent packets arrive, the ITR simply encapsulates them and sends the new data packets into the DFZ.

---

## 6. Routing and forwarding system

[TOC](#)

---

### 6.1. Intra-AS domain data packet forwarding

[TOC](#)

The communicated hosts are between one domain, and the EIDs mapping of the hosts are stored in the same overlay, which is called intra-AS domain communication. The principle of intra-AS communication is very simple. The process is described as following:

- (1) Host1 wants to communicate with host2, and sends a IP packet (a IPv4 packet or IPv6 packet, whatever) to its default ITR. The destination IP address and source IP address of the IP packet are the EID address of host2 and host1 respectively;
- (2) When ITR receives the IP packet, it looks up the RLOC of the EID2 in the local cache. If it finds the RLOC, it means that the packet is not the first packet, and skips into step 6 indirectly;
- (3) ITR encapsulates the LISP-Request message and sends to the DHT Server for requesting the RLOC of EID2;
- (4) When DHT Server receives LISP-Request, it looks up the RLOC in the distributed mapping database. Then DHT Server encapsulates the LISP-Reply response message;
- (5) When ITR receives the LISP-Reply message, it stores the EID-to-RLOC mapping into its local cache.
- (6) ITR prepends LISP data packet with a outer IP header. The destination and source IP address of outer IP header are RLOC of ETR and ITR respectively;

- (7)The LISP data packet routes and forwards based on the destination RLOC of the outer IP packet;
  - (8)after the LISP data packet reaches the ETR, it will be unencapsulated in the ETR. The ETR strips the outer IP header and remains the original IP packet;
  - (9)the IP packet forwards to the destination host2.
- 

## 6.2. Inter-AS domain data packet forwarding

[TOC](#)

The communication hosts are between to different domain and the EID mapping of hosts are stored in different DHT overlay, which is inter-AS domain communication. The inter-AS domain communication is based on DHT Border Server. The process is as following:

- (1)Host1wants to communicate with host3, and sends a IP packet(a IPv4 packet or IPv6 packet, whatever) to its default ITR. The destination IP address and source IP address of the IP packet are the EID address of host3 and host1 respectively;
- (2)When ITR receives the IP packet, it looks up the RLOC of the EID3 in the local cache. If it finds the RLOC, it means that the packet is not the first packet, and skips into step 6 indirectly;
- (3)ITR encapsulates the LISP-Request message and sends to the DHT Server for requesting the RLOC of EID3;
- (4)When DHT Server receives the LISP-Request message£~it looks up the RLOC in the distributed mapping database. Because the information of EID is not stored in the local distributed mapping database, the mapping can not find in the local overlay mapping database. The DHT Server encapsulates the LISP-Reply message, and the RLOC information filed is blank;
- (5)When ITR receives the ISP-Reply message with the blank RLOC, it stores the mapping relation of EID and RLOC into its local cache and marks the attribute with external EID information;
- (6)DHT Border Server encapsulates the LISP data packet with prepending a new IP header. The destination and source IP address of the outer IP packet are the RLOC address of DHT Border Server and ITR respectively;
- (7)The LISP data packet forwards to the DHT Border Server based on the destination RLOC of the outer IP header;
- 8)The outer IP packet is striped in the DHT Border Server, while the inner packet is remained. The DHT Border Server looks up in the BGP routing table to match the destination EID;
- (9)The IP packet forwards to the DHT Border Server where is the same domain with host3;
- (10)DHT Border Server looks up the RLOC of the EID3;
- (11)When finds the RLOC, the IP packet will be reencapsulate with the destination and source IP address of outer IP header as RLOC of DHT Border Server and EID3. The inner packet remains unchanged;
- (12)LISP data packet forwards to the ETR;



(13)When the data packet reaches the ETR, the ETR strips the outer IP packet header and remains the inner IP header;  
(14)the inner IP header routes according to the EID, and forwards to the destination host 3.

---

## 7. Security Considerations

[TOC](#)

---

## 8. Acknowledgement

[TOC](#)

---

## 9. References

[TOC](#)

---

### 9.1. Normative references

[TOC](#)

---

### 9.2. Informative References

[TOC](#)

[ALT]	Farinacci, D., Fuller, V., Meyer, D., and D. Lewis, "LISP Alternative Topology (LISP-ALT)," draft-ietf-lisp-alt-01.txt (work in progress), March 2009.
[APT]	Jen, D., Meisel, M., Massey, D., Wang, L., Zhang, B., and L. Zhang, "APT: A Practical Transit Mapping Service," draft-jen-apt-01.txt (work in progress), November 2007.
[LISP]	Farinacci, D., Fuller, V., Meyer, D., and D. Lewis, "Locator/ID Separation Protocol (LISP)," draft-ietf-lisp-00.txt (work in progress), March 2009.
[LISPDHT]	Mathy, L., Iannone, L., and O. Bonaventure, "LISP-DHT: Towards a DHT to map identifiers onto locators," draft-mathy-lisp-dht-00.txt (work in progress), Feb 2008.

---

## Authors' Addresses

[TOC](#)

	Fangwei Hu
	ZTE Corporation

	No.889 Bibo
	Pudong Distric,Shanghai 201203
	China
Phone:	+86-21-68896829
Email:	<a href="mailto:hu.fangwei@zte.com.cn">hu.fangwei@zte.com.cn</a>
	Jian Luo
	ZTE Corporation
	No.68 Zijinghua
	Yuhuatai Distric,Nanjing 200012
	China
Phone:	+86-25-52871234
Email:	<a href="mailto:luo.jian@zte.com.cn">luo.jian@zte.com.cn</a>