A Hierarchical Mapping System Approach for ID-to-Locator Resolution

Mukankunga Bisamaza Angel and Choong Seon Hong
Departement of Computer Engineering
Kyung Hee University
1 Seocheon, Giheung, Yongin, Gyeonggi, 449-701 Korea
Email: angel@networking.khu.ac.kr, cshong@khu.ac.kr

***Abstract*—Internet Default Free Zone (DFZ) is facing today a serious scalability problem due to the double use of the current Internet Protocol (IP) name space. IP name space today is used both for location and identification of a host. To solve this problem, Identifier and Locator separation approach was proposed and implemented in many schemes.**

However, the big challenge in the separation approach is how to map a given Identifier to the appropriate Locator. In this paper, we propose a hierarchical mapping system based on Internet hierarchy of IP allocation/assignment authorities. We evaluate our mapping system and show that beside the fact of removing the use of overlay networks scalability, and short lookup time can be achieved.

***Keywords*— Identifier/Locator Separation, Identifiers, Locators, Mapping System**

I. INTRODUCTION

Today's routing and addressing system of the Internet is facing an undeniable problem of scalability [1]. The Default Free Zone (DFZ) of the Internet is made of all the core routers. And these core routers using Border Gateway Protocol (BGP) are aware of how to reach any networks over the Internet. With the growth of the Internet users the content of those core routers has exponentially increased. But this growth is mainly due to some practices like the use of Provider Independent (PI) addresses instead of Provider aggregatable ones (PA), which allow Internet Protocol (IP) addresses aggregation to provide scalability in the DFZ. Also Internet users due to a need of reliability apply multihoming while using PI which also increases the numbers of entries in DFZ elements. And mobility is hardly supported since IP changes during a communication. This mobility creates a need to modify the IP used in the upper level to deliver the packet, and the only way this can be done currently is by interrupting the communication and restarting with the new IP. All these Internet users' changes are reflected in the DFZ because of the dual use of IP that represents the location of a host, as well as its Identity. Therefore, for a network to be globally reachable these modifications need to be reflected in the DFZ hence creating a rapid growth of entries in core routers.

Researchers have agreed that the dual use of IP is one of the main reasons of this scalability problem.

To solve this problem a new Internet architecture, where IP roles are separated into Identifiers and Locators, is designed inside IETF and Routing Research Group (RRG) known as Identifiers/Locators Separation.

This proposed architecture has the advantage of facilitating mobility and multihoming without causing the increase of the Routing table. Many proposals implement the Identifiers/Locator Separation approach, some are host based and others Routers based [2]. Host Identity Protocol is a host based approach of Identifiers/Locators Separation. Here a host identity layer, where the host Identifier is implemented, is added for the upper layer to use, while the IP address is still used at the network layer [16].

And recently, Locator/ID Separation Protocol (LISP) [3] has been proposed in the EITF to support incremental deployment of Identifiers/Locators separation on the current Internet architecture. LISP is a router based approach of Identifiers/Locators aiming to reduce the routing table in the DFZ.

However LISP, while solving the routing table scalability problem comes with a cost that is after separation, how to resolve for a given Identifier the correct Locators. Therefore, there is a need of a mapping system. This one has to be scalable, robust, with a relative short lookup time. For that purpose many proposals of mapping system came up each with pros and cons as we will show in the related works section.

In this paper, we have designed a Hierarchical Mapping System (HMS) based on IP allocation/assignment authorities' hierarchy in the current Internet [17].

HMS hierarchy is made of two components one in the Internet Allocation Number Authority (IANA) level and the second one in the Regional Internet Registry (RIR) level. The main assumption in HMS is that the allocation/assignment of Identifiers (EIDs) is done in the same way as for IP, and they are aggregatable.

The numerical evaluation of HMS in terms of Scalability, and short resolution delay has showed it feasibility. HMS also removes the overlay network burdens compared to previously proposed mapping systems [4], [5], [7].

The rest of the paper is organized as follows;

Section 2 gives an overview of Identifiers/Locator separation. Section 3 presents some related works. In section 4, we describe HMS in details. Section 5 discusses the numerical evaluation of HMS and finally section6 concludes this paper while considering some future works.

II. ID/LOCATOR SEPARATION OVERVIEW

Fig.1 provides a big picture of how the Identifiers/Locators separation works during an inter domain communication between two Hosts source (EIDsrc.) and destination (EIDdest.). The two hosts are attached to their tunnel routers (TR) which are in their turn connected to some Internet Service Provider (ISP) which uses Locators (RLOC) for the inter domain routing purpose. EIDsrc and EIDdest are respectively attached to RLOCsrc and RLOCdest. When EIDsrc want to communicate with EIDdest, it generates the first packet with EIDsrc as the source address and EIDdest as the destination address in the header. This packet is routed inside the domain using the usual routing protocols toward the ITR, if EIDdest belongs to the same domain the packet is forwarded to the destination directly.

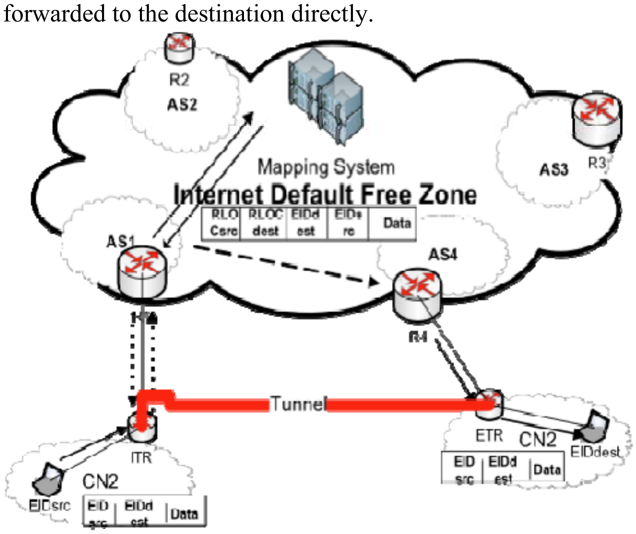

Figure 1. Identifiers/Locators Separation Overview

Otherwise the TR has to find RLOCdest associated to EIDdest using the mapping system, before routing the packet over the Internet towards the EIDdest domain.

In the case of a multihoming on the destination domain, the mapping system will reply with RLOCdest according to the preferences of the destination domain or even sometimes reply with all the available RLOCdest of the destination. Upon the reception of the RLOCdest TR keeps that mapping information into its caching system before it encapsulates the packet with a new header where RLOCsrc (also known as TR's RLOC at the source domain) is the source address and RLOCdest (TR's RLOC at the destination domain) is the destination address and forward it into the Internet. TR at the destination after receiving the packet strips out the outer header and forward the packet inside the destination domain, where it will finally reach EIDdest.

The challenging part of Identifiers/Locator separation is the mapping system design for the resolution of a given EID.

In the next section, we describe previously proposed mapping systems in details.

III. RELATED WORKS

Many proposals provide different ways of designing a mapping system in the Identifiers/Locator Separation approach. Here are some of them based on LISP currently discussed in EITF [3].

In [5], V.Fuller et al propose LISP alternative topology (LISP-ALT), in which the mapping information EID-to-RLOC is stored in a distributed way in different ETR, which are interconnected using an overlay network of ALT-routers. Here, the scalability problem is removed at the cost of the complexity of the overlay network maintenance as well as the inefficient use of network resources. LISP-ALT overlay network uses BGP approach as used in the current Internet; but with a different kind of BGP. Moreover, whenever a host moves from one domain to another the resolver also changes, causing certain instability in LISP-ALT overlay network since the aggregation of EIDs will also be broken.

LISP-CONS [4], distributes the mapping information hierarchically removing the need of BGP routing approach used in LISP-ALT. LISP-CONS proposed the aggregation possibility in high level to achieve scalability but not always the aggregation can be achieved due to mobility therefore, still scalability is not totally assured and again there is a cost of the overlay network created by higher level of the hierarchy.

LISP–NERD [6] main advantage is the absence of caches missing but at the cost of a scalability concern since tunnel routers need to hold the entire mapping information database.

LISP-DHT [7] is based on the Chord DHT overlay network with border routers considered as nodes of the P2P overlay network. Though scalability and robustness is provided due to the use of DHT structure, resolution time remain very critical since it depends on the number of nodes involved in the Chord ring (O (logN).)

All these proposals for LISP mapping system can be applied with some tradeoff since each of them has its pros and cons. One of the cons except for LISP-NERD is the use of an overlay network for the mapping service which comes with the cost of maintenance, physical network most of time different from the topological network which affect the routing performance.

With the objective of removing the burden due to the use of the overlay network in the design of mapping system, we have designed a hierarchical mapping system based on the recent IP allocation/assignment hierarchy (HMS). We also take advantage of LISP-NERD [6] distribution process to reduce the resolution delay. HMS aims to remove the overlay cost as well as providing scalability through a hierarchical mapping system approach which reduces the resolution time.

IV. PROPOSED MAPPING SYSTEM (HMS)

In this section, we describe how HMS has been designed using the hierarchy of IP assignment/allocation authorities [17].

Fig.2 depicts the hierarchy relation between IP assignment authorities and allocation authority which is Internet

Assignment Numbers Authority (IANA). HMS design follows the same hierarchy in how mapping informations are stored and managed.

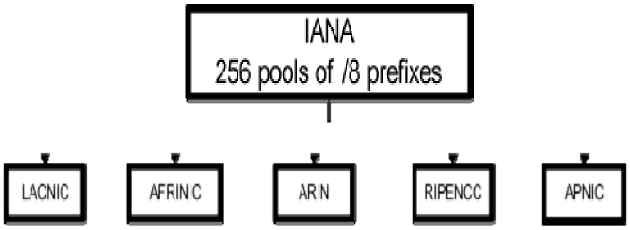

Figure 2. Allocation/assignment Authorities

Today, all IP addresses spaces are managed by IANA, which is in charge of allocating some portion to different region and in those region, these portions are managed by a Regional Internet Registry (RIR). And finally any network in need of IP make a request to the RIR of its region in order to join the Internet. Based on the need a prefix is assigned to that network from the previously allocated portion from IANA.

However, in some case the hierarchy of some region might add the Country Internet Registry (CIR) and/or Local Internet Registry (LIR) which we will not consider for our work.

All assigned prefixes always comes from a previously allocated prefix (/8 for IPV4 or /12 for IPV6). HMS main assumption is that EIDs space will be distributed to different networks in the same way as IP today. IANA level will allocate EIDspace to RIR level and edge networks will get their EIDpref from RIR. Note that EIDspace (a /8 of IPV4 or /12 for IPV6) is larger than EIDpref (prefixes extracted from previously allocated / 8 IPV4 or /12 IPV6)

A. HMS Components

Fig.3 shows the main components of HMS and how they interact.

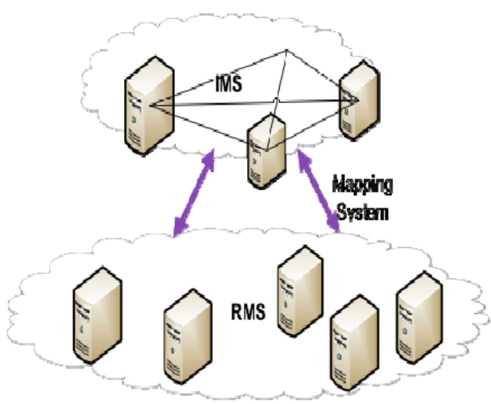

Figure 3. HMS components

The mapping system has two main components; IANA Mapping Server (IMS) and Registry Mapping Server (RMS).

- IMS: stores mapping information of EIDspace which are allocated to registry (EIDspace-to-RMS) and mapped to the RLOC of their authoritative RMS. It is also suppose to share these mappings information with TR of edges networks based on LISP-NERD distribution approach.
- RMS: stores the real mapping information (EIDpref-to-RLOC) of edges network prefixes mapped to the RLOC of the TR they are attached to. RMS is therefore the HMS component responding to mapping requests.

The mapping from IMS in TR points to the registry in charge of the EID space to send the map-request directly to the RMS. IANA and RIR can deploy more than one IMS and RMS respectively to increase robustness and decrease latency while providing load balancing. And for management and efficient communication with TR RIR and IANA can configure an internal multicast group and an internal anycast group local to each region as well as to IANA [11].

HMS supports two main processes respectively the mapping information registration to HMS, and the retrieval process for packet forwarding.

B. Mapping Registration Process

As depicted in the Fig. 4 the mapping information from customer's networks is registered following these steps

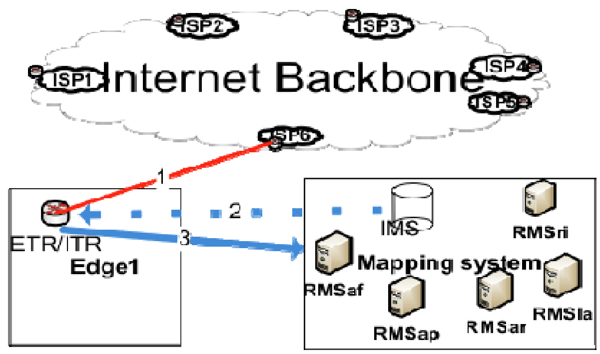

Figure 4. HMS Mapping Information Registration process

Step 1. ITR/ETR receives the connection from the ISP

Step 2. ITR/ETR downloads the IMS mapping information

Step 3. Using the longest prefix match in the downloaded IMS mapping table, ITR/ETR determines to which RMS the mapping information should be sent.

Step 4. After RMS has received the mapping information and it does not exist in the mapping table, the mapping information is stored. In the case it already exists, RMS just updates the mapping table with the new mapping information

C. Mapping Information Retrieval Process

Mapping retrieval takes place for inter domain communication when the TR doesn't have the mapping information in it caching system. Otherwise the packet is directly encapsulated and forwarded to the destination. Fig. 5 depicts the process in four steps; the encapsulation is done after these steps. Differently to the mapping registration process here the network is considered already connected to one ISP. Below is the summary of the mapping retrieval process

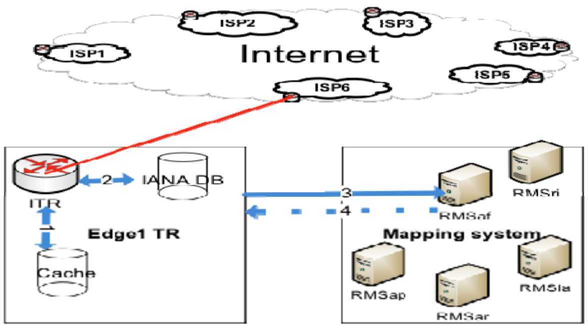

Figure. 5 Mapping Information Retrieval Process

- TR checks the cache in this case there is a cache missed. TR applies the longest prefix on the mappings information downloaded fro IMS. This will point it to theRLOC of RMS in charge of EIDspace to which that EID belongs to.
- Upon the map request reception, RMS uses longest prefix as well to determine to which specific prefix the EID belongs to. RMS sends the map-reply to TR requester when the mapping information is found otherwise a negative reply is sent.

Upon the reception of the mappings information the packet can be forwarded from one edge network to another using a tunnel. Fig. 6 illustrates the packet forwarding process

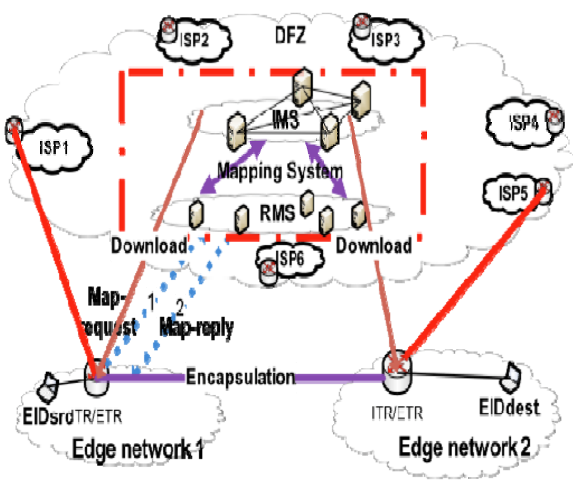

Figure 6. Packet Forwarding using HMS

These processes of HMS have been described with no security consideration. In the next section, we perform evaluation of HMS feasibility in terms of scalability, overhead, and resolution time delay

V. PERFORMANCE ANALYSIS

This section focus on the evaluation of the feasibility of HMS based on parameters like scalability, overhead and, finally resolution delay. With these three parameters, one can determine the efficiency of a mapping system.

A. Scalability

Scalability in the routing table is one of the reason Identifiers/Locators separation was proposed. If the mapping system is not well designed the scalability problem will move from the DFZ of today Internet to the mapping system of Identifiers/Locators separation architecture. To measure the scalability, we adopted LISP-NERD distribution approach [6] metrics.

$$E \quad * (30 + 20 * (R - 1)) \qquad (1)$$

Where E is the number of EIDs and R is the number of Locators attached to the EIDs.

Since in LISP [3] EID address space generation process has not been specified in details we have decided to consider for the evaluation the existing IP namespace in case EID are defines in the same way. However due to the growth of Internet users and the exhaustion of IPV4, we have considered for our evaluation that EIDs could be a subset of the IPV6 due especially to its very large number of addresses range, enough to satisfy the world habitant According to [14] RIR are allocated /12 space of IPV6 address space, which imply that IMS will store a total of 4096 pools of /12 mapped to their respective region. HMS is designed in a way TR downloads from IMS mapping information as in [6]. We evaluate how much data need to be transferred from IMS to TR in case of full allocation using modified equation 1 due to the shorter length of stored prefix in IMS compared to the 64bit used in LISP-NERD [6]. Here we assume that RMS in each directory is identified by one locator.

$$E \quad * \quad 6 \qquad (2)$$

Therefore IMS storage is estimated to 24576Bytes. This space is way much smaller compared to what servers and routers can support today [21], implying that IMS and TR scalability can satisfy. As for TR in case of IPV6 like EIDs space full allocation is still 100 times smaller than what today BGP routers are supporting, leaving enough space for the local EIDs mapping information to be stored

For RMS we consider that EIDs are randomly distributed among registries. Therefore, each RMS will store a 1/5 of the LISP-NERD approximated usable space [6] independently. This result leads us to a conclusion that the space used is way much smaller compared to the approximation of LISP-NERD single database. And moreover the mapping information is stored in terms of prefixes which reduce as well the needed space of RMS compared to when single EID are used. Although [18] shows that in reality some registry can hold more addresses than others depending on the users in that region. In any case RMS will still scale compared to LISP-NERD. From fig. 7 and previous analysis on IMS and TR, we show that scalability in HMS is satisfied in all the participants' components of the mapping service.

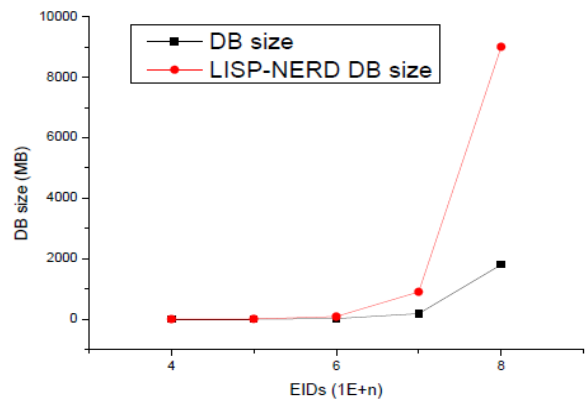

Figure 7. HMS vsLISP-NERD Scalability Performance

B. *Overhead Analysis*

Overhead can be caused by the high frequency of requests and updates. IMS updates its table in case of a new pool assignment and RMS mapping information is updated in case of a provider change which also doesn't happen frequently. Therefore, the overhead at IMS due to TR request for update is bearable, if we consider that TR will have a scheduled time to request updates from IMS. Table1 results shows the time used when a single TR downloads IMS entire mapping information using different bandwidth. However, in realtime traffic IMS will have to reply to simultaneous requests from different TR.

Table 1 TR downloading time from IMS to TR

Bandwidth	1 mbps	10 mbps	100 mbps	1 gbps
Processing time in milliseconds	200	20	2	0.2

We evaluate the processing time in case of 10^2 to 10^6 simultaneous requests from TR to IMS with a 100% utilization of a 1Gbps bandwidth for the transfer. The results in table2 shows an improvement compared to LISP-NERD [6].

Table 2 IMS entries processing time

# Simultaneous Requests	**10 Servers**	**100 Servers**
100	2	0.2
1000	20	2
10000	200	20
100000	2000	200
1000000	20000	2000

Moreover, the processing time can still be improved by the download of only changes in the IMS database instead of the entire database. And the changes happens when a new pool of /12 is allocated to a registry, which happen not before at least 20month for the registry with the highest assignment speed [19].

On the other hand the overhead at RMS is reduced by the caching approach implemented in TR. The fact of resolving a prefix of the all network reduces also the overhead in a sense that if the source domain communicates with more than one node of the destination network the resolution will not take place since the prefix of that domain is already in the caching system. And finally updates in RMS has a very low frequency due to the fact that edge networks don't change frequently their provider, hence their mapping information also remain the same.

The above analysis leads us to the conclusion that overhead due to updates and simultaneous request are manageable compared to the LISP-NERD from which HMS borrows the distribution model.

C. *Resolution Delay*

Through this approach the resolution time is considerably reduced as compared to approaches based on overlay network. There is no specific routing added to mapping retrieval process, therefore we estimates the time using the end to end delay between involved component adapted to today Internet
We have evaluated the resolution delay as follow:

T_1 : Time a packet takes from the PC to the TR; we assume T_1 =1ms

T_2 : Time used by the TR to process the request before generating map request or forwarding the packet, we assume T_2 = 2ms

T_3 : The time map-request takes from the TR to the RMS, we assume that this depends on the RTT between the network and the RMS

T_4 : processing time in the RMS, we assume T_4 = 0,420ms based on [20] estimation with more than 25millions of search/second

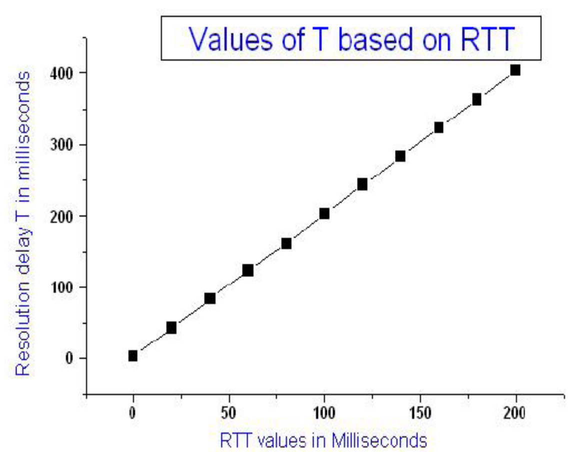

Figure 8. Packet Forwarding using HMS

T_5 : Time the map-reply takes from the RMS to the ITR is same as T_3

Therefore, the resolution delay is obtained as follow:

$$T_{tot} = T_1 + T_2 + T_3 + T_4 + T_5 = 2RTT + 3.420ms \quad (3)$$

We can see that HMS resolution delay depends on the RTT. If the RTT is low then T_{tot} is low but if RTT is high the T_{tot} will be high also. But based on the result on smokeping tool [19] of root DNS RTT, we assume that the results of RTT will always be around 100ms for the highest RTT.

Therefore, the resolution delay T_{tot} = 103,420ms, which is better than the one proposed in [8]. Fig 8 shows the variation of T_{tot} for different values of RTT. But mainly the resolution process burden is released by the caching process, which is always implemented in TR.

VI. CONCLUSION

Identifiers/Locators separation aims to resolve the scalability problem that the Internet routing system is facing currently.

However, this approach requires a high consideration in the design of its mapping system, considered as a vital element of the inter domain packet routing process.

In this paper, we have focused on designing HMS a hierarchical mapping system, based on today IP assignment/allocation hierarchy, for Identifiers resolution to remove overlay network burden associated with previously proposed mapping system.

The evaluation of HMS has shown that scalability, in terms of storage and numbers of entries, is supported. Also overhead and resolution delay imposed to HMS are manageable.

While scalability is achieved in case of stable nodes, mobility will affect HMS scalability especially at the RMS level. This is due to the fact that in case of movement a node will be required to update HMS with its new location in its authoritative RMS. This update will break the prefixof its home network causing the increase of entries in RMS. Our future works will focus in providing a mobility scheme which will not affect the scalability of HMS.

ACKNOWLEDGMENT

This research was supported by the MKE, Korea, under the ITRC support program supervised by the NIPA (NIPA-2010-(C1090-1021-0003)) and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by MEST (No. 2010-0027645). Dr. CS Hong is the corresponding author

REFERENCES

[1] D. Meyer, L. Zhang, and K. Fall ”'RFC 4984: Report from the iab workshop on routing and addressing'”, September 2007.

[2] Ping Dong, Hongchao Wang, Yajuan Qin, Hongke Zhang and Sy-Yen Kuo ”Evaluation of Scalable Routing Architecture Based on Locator/Identifier Separation”', Globecom 2009

[3] D. Farinacci,V. Fuller,D. Meyer and D. Lewis,”'Locator/ID Separation Protocol (LISP)”',draft-ietf-lisp-07, IETF Networking group,october2010

[4] Brim, et al ”'LISP-CONS: A Content distribution Overlay NetworkService for LISP draft-meyer-lisp-cons-04.txt”' April 9, 2008

[5] V. Fuller,D. Farinacci,D. Meyer and D. Lewis, ”'LISP Alternative Topology(LISP+ALT)”',draft-ietf-lisp-alt-04.txt, IETF Networking group, October2010

[6] E. Lear and Cisco Systems GmbH, ”‘NERD: A Not-so-novel EID to RLOC Database”',draft-lear-lisp-nerd-08.txt, IETF Networking group, September 2010.

[7] L. Mathy and L. Iannone,”'LISP-DHT: Towards a DHT to Map Identifiers onto Locators,'” in Re-Architecting the Internet (ReArch), Madrid, Spain,Dec. 2008.

[8] Hongbin Luo, Yajuan Qin and Hongke Zhang,”A DHT-Based Identifierto-Locator Mapping Approach for a Scalable Internet” in IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 20, NO. 12, DECEMBER 2009

[9] Ningning Lu and Hongke Zhang,”BILM: A Novel BGP-based Identifierto-Locator Mapping Approach” in ICIME, 2010.

[10] L. Lannone and O. Bonaventure,”On the Cost of Caching Locator/ID Mappings” Proc. ACM Intl Conf.Emerging Networking Experiments and Technologies (CoNEXT), Dec. 2007

[11] D. Jen, M. Meisel, D. Massey, L. Wang, B. Zhang, and L. Zhang,”APT: A Practical Transit Mapping Service”,IETF Internet draft,draft-jen-apt-01.txt, work in progress, Nov. 2007.

[12] ”RIPE NCC, http://www.potaroo.net/tools/ipv4/index.html” , 2010

[13] ”‘http://wapedia.mobi/en/IPv4addressexhaustion"0; 2010

[14] ”‘http://www.icann.org/en/general/allocation-IPv6-rirs.htm”',2010

[15] ”smokeping.ncren.net”

[16] R. Moskowitz, P. Nikander, P. Jokela, and T. Henderson. Host Identity Protocol. RFC 5201, April 2008

[17] K. Hubbard, M.Kosters, D.Conrad, D. Karrenberg, J. Postel “RFC2050:Internet Registry IP Allocation Guidelines”,November1996

[18] ”'http://www.icann.org/en/general/allocation-IPv4-rirs.html”

[19] ”‘RIPE NCC, http://www.potaroo.net/tools/ipv4/index.html”' , 2010

[20] Masayoshi Kobayash, Tutomu Murase and Atsushi Kuriyama” A Longest Prefix Match Search Engine for Multi-Gigabit IP Processing”

[21] SunMicrosystems, http://www.sun.com/products/sunmd/s20/specifications.jsp