Multihoming with Locator/ID Separation Protocol: An Experimental Testbed

Soroush Haeri and Rajvir Gill
Simon Fraser University
Vancouver, British Columbia, Canada
Email: {shaeri, rajvirg}@sfu.ca

Marilyn Hay and Toby Wong
BCNET
Vancouver, British Columbia, Canada
Email: marilyn.hay@ubc.ca, toby.wong@bc.net

Ljiljana Trajković
Simon Fraser University
Vancouver, British Columbia, Canada
Email: ljilja@sfu.ca

Abstract—The exponential growth of the Routing Information Base (RIB) of the Internet's Default-Free Zone (DFZ) routers has raised concerns about non-scalability of the current Internet's routing architecture. The main reason is that Internet addresses currently carry information about both the identity and location (physical connection point) of devices connected to the Internet. The Locator/ID Separation Protocol (LISP) has been introduced to potentially remedy this non-scalability by splitting the location and identity of devices. In this paper, we present the architecture of a deployed testbed that is multihomed using LISP. We investigate LISP performance as a multihoming solution in terms of load balancing and traffic routing in the case of link failures.

I. INTRODUCTION

In the current Internet addressing architecture, the Internet Protocol (IP) namespace is overloaded with information about the routing location and the identity of network devices. Locator/identifier overload of the IP address semantics is a recognized cause of the Internet's scalability issues including the exponential growth [1] of the routing tables in the Internet's Default-Free Zone (DFZ). The DFZ includes Autonomous Systems (ASes) whose routers do not require a default path for sending data to their destinations. The *Locator/Identifier split* has been proposed to achieve a scalable routing infrastructure [2].

Various methods have been proposed to achieve the locator/identifier split. These methods are based on two main techniques: address rewriting and map-and-encap. In the case of address rewriting, the IP addresses of the source and destination are replaced with the source and destination locators, respectively. The Six/One protocol [3] and the Identifier-Locator Network Protocol (ILNP) [4] are examples of protocols based on address rewriting. In the case of map-and-encap technique, a new header is inserted in a packet. This new header includes the source and destination locators. Examples of the map-and-encap based schemes are the Locator/Identifier Separation Protocol (LISP) [5] and the Internet Vastly Improved Plumbing (IVIP) [6]. The advantage of the map-and-encap approach is that the packet that has been generated by a host remains unchanged when traveling through the Internet core.

LISP has received considerable attention from the Internet community. LISP related documents published by the LISP working group of the Internet Engineering Task Force (IETF) [7] have evolved into RFCs. A LISP pilot network has been deployed worldwide [8]. It consists of more than 140 sites located in 25 countries. LISP maintains the existing routing scheme and provides a routing locator design for inter-domain routing. Since LISP decouples location and identifier, it may improve the current routing architectures. It may also decrease the burden of heavy updates of the Routing Information Base (RIB). LISP may be implemented in a network with minimal modifications. Hence, we were interested in exploring deployment of LISP in BCNET [9].

LISP operates in both the data and control planes. LISP introduces two namespaces: Endpoint Identifiers (EIDs) that identify connection endpoints and Routing Locators (RLOCs) that identify the connection point of a node to the Internet. EIDs are only locally routable while RLOCs are globally routable addresses. Packets are encapsulated and decapsulated in the data plane. An Ingress Tunnel Router (ITR) that resides in a LISP site performs EID-to-RLOC mappings and is responsible for encapsulating the outgoing packets. An Egress Tunnel Router (ETR) accepts LISP encapsulated packets from the Internet, decapsulates them, and forwards them to end devices. The LISP control plane provides the EID-to-RLOC mapping infrastructure. The EID-to-RLOC mapping associates a given identifier to a set of locators. These locators provide various reachability paths for a given identifier [5].

A multihomed site has IP (Layer 3) connections to multiple transit providers. Multihoming using the current Internet routing infrastructure is achieved by using either the Border Gateway Protocol (BGP) or the Network Address Translation (NAT). A multihomed site announces its IP address space to the upstream transit providers using BGP. The LISP architecture offers flexible multihoming solutions. In this paper, we deploy a multihomed testbed using LISP and evaluate LISP capabilities in terms of load balancing and recovering the multihomed site from link failures.

This paper is organized as follows: In Section II, we describe multihoming and challenges involved in multihoming sites using the legacy Internet infrastructure. The architecture of the deployed multihomed testbed, description of the test cases, and the measurement results are presented in Section III. In Section IV, we conclude that LISP is a viable multihoming solution that requires no additional configuration or changes in the existing network infrastructure.

II. MULTIHOMING IN IP NETWORKS

A multihomed site may have connections to multiple transit providers and, hence, it may rely on more than one
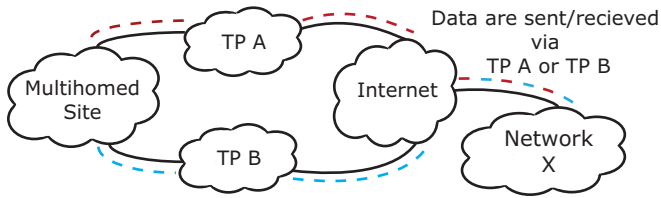

Fig. 1. An illustration of a multihomed site and its connections to upstream transit providers TP A and TP B.

connection to upstream transit providers. Hence, if one of the connections to a transit provider is unavailable, data are routed to the destination via an alternate provider. An example of a multihomed site is shown in Fig. 1. Traffic from the source to the destination may flow through either one or both transit providers labeled TP A and TP B. Multihomed sites deployed within the current Internet routing infrastructure employ either BGP or NAT [11].

In the case of BGP multihoming, a multihomed site announces its IP address space to its upstream transit providers using BGP. For example, in Fig. 1, the site prefixes are announced to both the Transit Provider A (TP A) and the Transit Provider B (TP B). The upstream transit providers then disseminate the routing information to the Internet. In the case of an upstream link failure, BGP advertises the change in the network topology.

Two common Internet address allocation schemes are: Provider Independent (PI) and Provider-Aggregatable (PA) addressing. PI addresses are globally-unique addresses that are usually assigned by a Regional Internet Registry (RIR). PI addresses introduce additional prefixes into the global routing system and, therefore, there are limitations on acquiring PI addresses especially for sites whose address space is smaller than /24 [10]. In contrast, PA addresses are globally-unique addresses that are assigned to a customer by a transit provider. The PA addresses are usually aggregated into a larger routable prefix by the transit provider.

PI addresses are independent of the upstream transit providers and, therefore, they cannot aggregate routes with PI addresses. Consequently, using PI addresses when multihoming with BGP increases the overhead of the BGP routing tables of the upstream transit providers [11]. Using PA addresses for multihoming with BGP requires cooperation and consent from the transit providers. For the example shown in Fig. 1, if the multihomed site possesses an aggregatable address space that belongs to TP A, the site cannot simply announce its prefixes to TP B without TP A's consent. Although there are addressing techniques that may ameliorate this situation, [12], they impose various other limitations [11].

NAT was originally introduced to enable local networks to reuse IP addresses inside their networks [13]. NAT translates a public Internet address to internal local network addresses and vice versa. In the case of multihoming with NAT, the PA IP addresses are often assigned by the upstream transit providers. A NAT multihomed site may possess IP addresses that belong to one or several upstream providers. NAT is employed to map the addresses that are assigned by various upstream providers to the network internal address space. IP addresses of the incoming packets are mapped into the local network

addresses while the addresses of the outgoing packets are mapped into the public Internet addresses. NAT separates the internal addresses of the network from the addresses assigned by transit providers. Hence, a site need not change its local address space if it changes its upstream providers. Multihoming with NAT does not have the addressing requirements of BGP multihoming. However, the routing process outside the site cannot be controlled. This makes NAT multihoming more suitable for small organizations that do not require route control [11].

III. THE MULTIHOMED LISP TESTBED

We evaluate the LISP performance for site multihoming by deploying a new testbed located in the Communication Networks Laboratory at the Simon Fraser University Burnaby campus in British Columbia, Canada. The diagram of the deployed testbed is shown in Fig. 2. It consists of a workstation running Ubuntu operating system. The workstation is connected via a dedicated fiber link to a LISP router located at the BCNET transit exchange center in downtown Vancouver. BCNET provides high-speed optical advanced network to British Columbia's higher education and research institutes and operates the Optical Regional Advanced Network (ORAN) [9]. Primary BCNET backbone is a 10 Gbps Ethernet network with backup 1 Gbps links. BCNET is connected to its transit providers via 1 Gbps and 10 Gbps links [14]. The LISP router (xTR) is connected via two connections to two BCNET subnets. These subnets route the traffic through distinct transit providers. Therefore, the single workstation located in Burnaby is a multihomed site.

The deployed testbed employs the LISP pilot network [8] infrastructure for control plane operations. The EID address space A.B.C.16/28 is assigned by the LISP pilot network. We evaluated performance of the testbed by measuring the distributions of traffic over the site connections. We also analyzed the LISP behavior in case of link failures. For testing purposes, we performed a peer-to-peer BitTorrent file transfer of *The Zeitgeist: Moving Forward* [15] motion picture. The size of the transferred file was 6.7 GB. The peer-to-peer file transfer was completed within 28 min and 15 s. During the file transfer, we repeatedly altered the distributions of traffic on the site connections. We also tested various link failures by shutting down the network interfaces of the LISP router. We captured the inbound and outbound traffic to and from the multihomed site, respectively. The captured traffic traces contain only the headers of the transmitted packets. Average traffic loads that are routed through Subnet A and Subnet B are shown in Figure 3.

During the intervals when one of the links is down, the entire traffic is routed through the available connection. The results show that traffic is distributed according to the weights assigned to the outgoing links and the network delivers the expected throughput. For example, during the first time interval of the file transfer (from 00:00 to 03:50), the distributions of traffic over Subnet A and Subnet B are 70% and 30%, respectively. During the second time interval (from 03:50 to 05:55), the link to Subnet A is disabled. During the first interval, distributions of the transferred data over the links to Subnet A and Subnet B are approximately 70% and 30%, respectively. During the second interval, the entire traffic flows
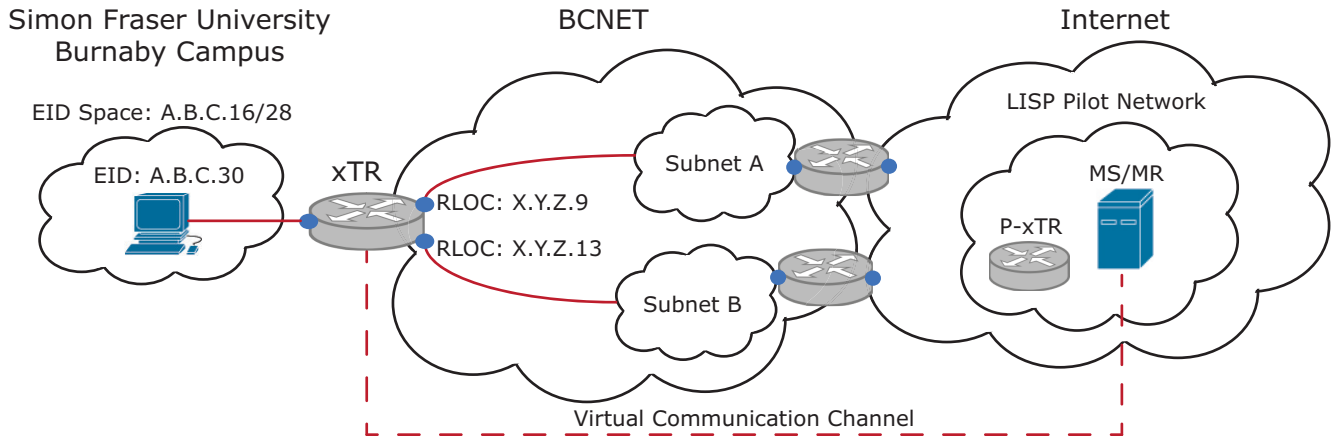
Fig. 2. Network diagram of the deployed testbed. It consists of a single workstation located at the Simon Fraser University Burnaby campus in British Columbia, Canada. The workstation is connected via a dedicated fiber link to a LISP router (xTR) located at the BCNET transit exchange center in Vancouver. The LISP router is connected via two connections to two BCNET subnets. These subnets route the traffic through distinct transit providers. The testbed uses the LISP pilot network's Map-Servers (MS), Map-Resolvers (MR), and Proxy-ETR/ITR (P-xTR). The EID address space A.B.C.16/28 is assigned by the LISP pilot network.
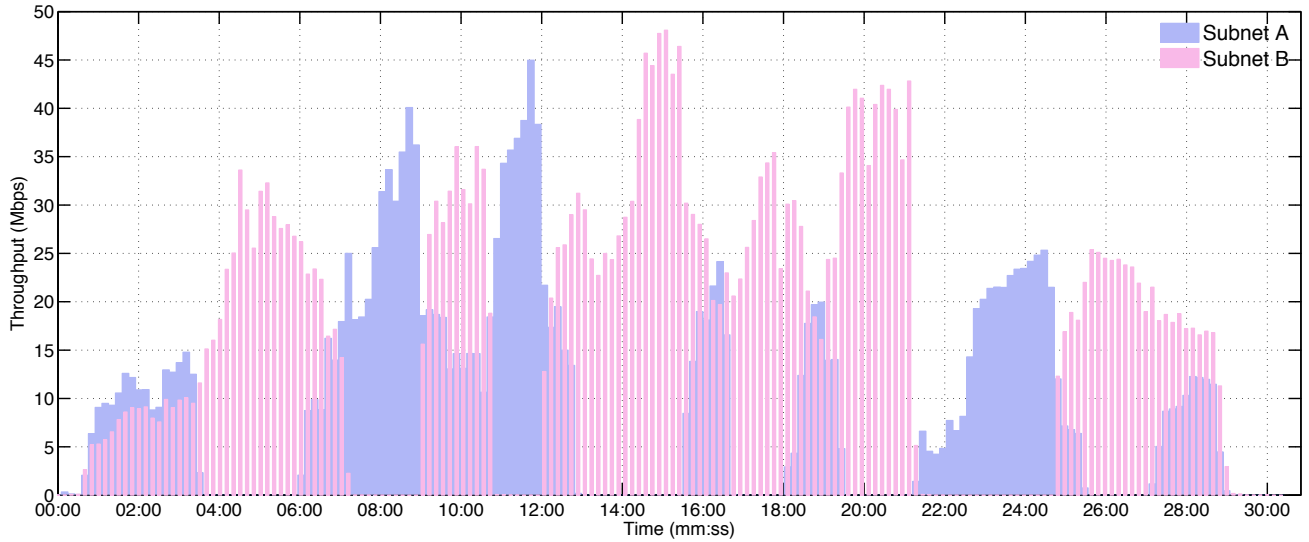

Fig. 3. Average throughput of the multihomed network routed through Subnet A and Subnet B. Each bar represents the average throughput (Mbps) in a 10-second interval. The overlapping areas indicate time intervals when the links to both subnets are operational. During the time intervals when both upstream links are operational, traffic is routed through both Subnet A and Subnet B based on the configured traffic distributions between the two links.

through Subnet B. The throughput of Subnet B increases and it is equal to the sum of the throughputs of Subnet A and Subnet B during the first time interval. The measurements shown in Fig. 3 are consistent with the assigned distributions of traffic. There are no traffic disruptions in the case of a link failure and the entire traffic is routed through the available connection.

The ratios of the traffic routed through Subnet A and Subnet B to the total traffic load are shown in Fig. 4 (top) and Fig. 4 (bottom), respectively. The dashed line is the assigned profile shown while the solid line is the profile generated based on the measured traffic. The dashed line has been adjusted to account for delay in manually recording the time instances when traffic distribution changes. The graphs illustrate that LISP successfully balances the traffic load based on the distributions of traffic configured for each connection. Furthermore, LISP reacts almost instantaneously to changes in the distribution of traffic and balances the traffic on the

available connections according to the configured distributions, as illustrated in Fig. 4. The link to Subnet A fails at 13:05 and recovers at 15:25. As a result of the link failure, the entire traffic is routed via Subnet B. However, when the link to Subnet A recovers at 15:25, 50% of the traffic is routed via Subnet A while the remaining 50% is routed via Subnet B.

IV. CONCLUSION

In this paper, we presented the architecture of a deployed testbed that is multihomed using LISP. We investigated the performance of LISP in terms of load balancing and traffic routing in the case of link failures by capturing the traffic from and to the multihomed site. LISP provides a flexible solution for site multihoming. Multihoming using LISP does not require PI addresses. The analysis of the captured traffic shows that LISP successfully balances the load on connections to the multihomed site according to the predefined distribution
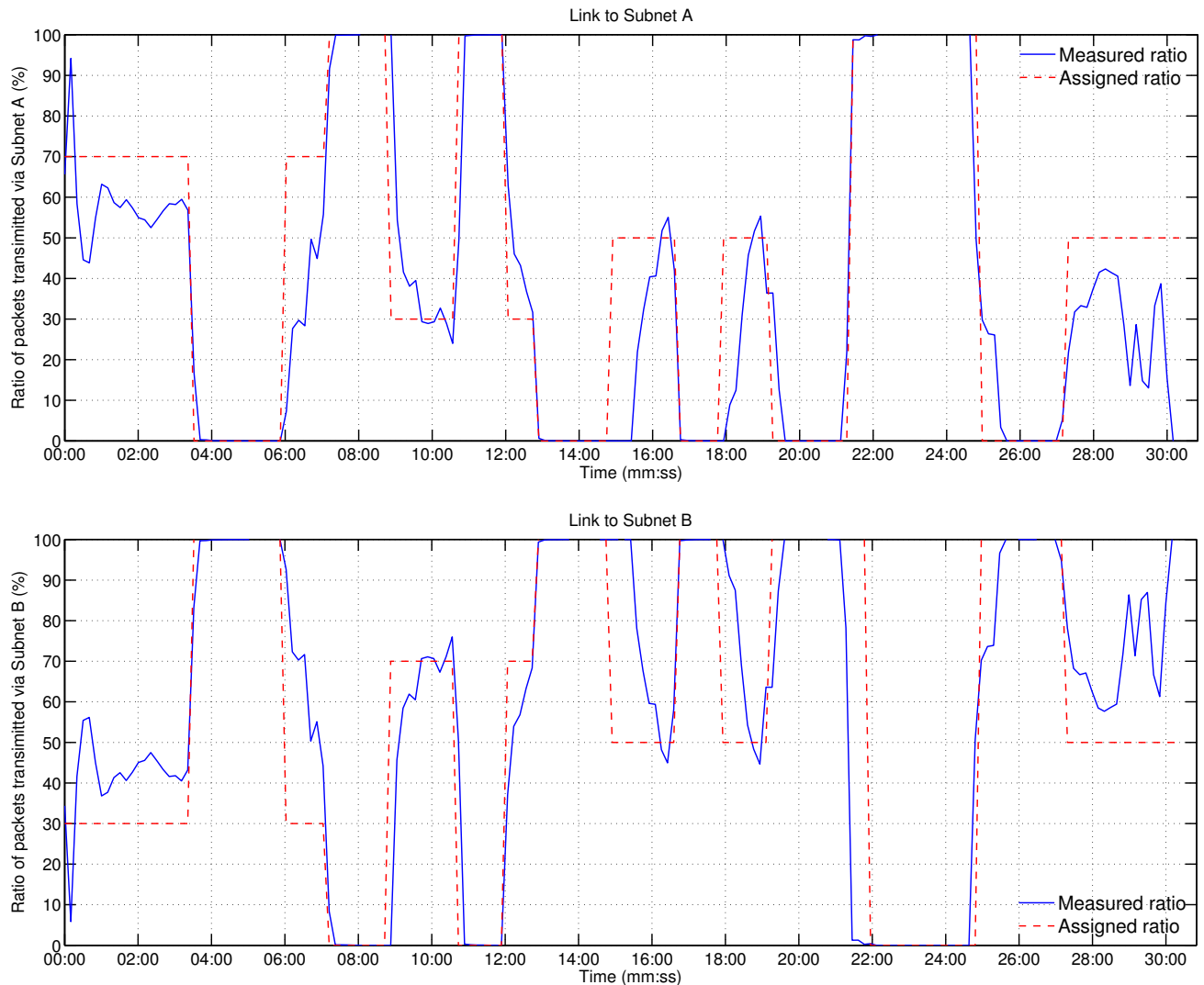
Fig. 4. The ratio of the traffic routed through Subnet A (top) and Subnet B (bottom) to the total traffic load. The dashed line is the assigned traffic profile while the solid line is the profile generated based on the measured traffic. The dashed line has been adjusted to account for delay in manually recording the time instances.

of traffic among the connections. In the case of a link failure, there was no disruptions of the incoming traffic to the testbed and the entire traffic was successfully routed through the alternate connection. Scalability of LISP implementations, LISP's impact on reducing the size of Routing Information Base, and new management challenges remain open issues.

REFERENCES

[1] G. Huston. (Mar. 1, 2014) Analysis of BGP routing table dynamics. [Online]. Available: http://bgp.potaroo.net/.

[2] D. Meyer, L. Zhang, and K. Fall, "Report from the IAB workshop on routing and addressing," *IETF*, RFC 4984, Sept. 2007.

[3] C. Vogt, "Six/One: a solution for routing and addressing in IPv6," *IETF*, Internet Draft, Nov. 2007.

[4] R. Atkinson, "ILNP—identifier/locator network protocol," *IETF*, Internet Draft, Aug. 2010.

[5] D. Farinacci, V. Fuller, D. Meyer, and D. Lewis, "The locator/ID separation protocol (LISP)," *IETF*, RFC 6830, Jan. 2013.

[6] R. Whittle, "IVIP (internet vastly improved plumbing) architecture," *IETF*, Internet Draft, Sept. 2010.

[7] (Mar. 1, 2014) Locator/ID Separation Protocol IETF Working Group. [Online]. Available: http://datatracker.ietf.org/wg/lisp/.

[8] (Mar. 1, 2014) LISP pilot network. [Online]. Available: http://www.lisp4.net/.

[9] (Mar. 1, 2014) BCNET: Shared IT services for higher education and research in british columbia. [Online]. Available: http://www.bc.net/.

[10] J. Abley, K. Lindqvist, E. Davies, B. Black, and V. Gill, "IPv4 multihoming practices and limitations," *IETF*, RFC 4116, July 2005.

[11] X. Liu and L. Xiao, "A survey of multihoming technology in stub networks: current research and open issues," *IEEE Network*, vol. 21, no. 3, pp. 32-40, May 2007.

[12] Y. Rekhter and T. Li, "An architecture for IP address allocation with CIDR" *IETF*, RFC 1518, Sept. 1993.

[13] K. Egevang and P. Francis, "The IP Network Address Translator (NAT)," *IETF*, RFC 1631, May 1994.

[14] T. Farah, S. Lally, R. Gill, N. Al-Rousan, R. Paul, D. Xu, and Lj. Trajković, "Collection of BCNET BGP traffic," in *Proc. 23rd International Teletraffic Congress*, San Francisco, CA, Sept. 2011, pp. 322–323.

[15] (Mar. 1, 2014) The Zeitgeist Film Series. [Online]. Available: http://www.zeitgeistmovie.com.