

Future internet services based on LIPS technology

Yue Li

▶ **To cite this version:**

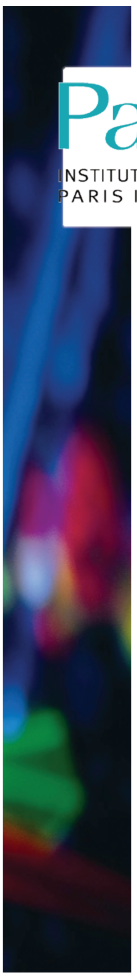
Yue Li. Future internet services based on LIPS technology. Networking and Internet Architecture [cs.NI]. Télécom ParisTech, 2018. English. NNT : 2018ENST0024 . tel-03380758

HAL Id: tel-03380758
https://pastel.hal.science/tel-03380758

Submitted on 15 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

TELECOM ParisTech

2018-ENST-0024

EDITE ED 130

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

Télécom ParisTech

Spécialité “ Informatique et Réseaux”

présentée et soutenue publiquement par

Yue LI

Le 26 Avril 2018

Future Internet Services based on LISP Technology

Directeur de thèse : **Luigi IANNONE**

Jury

M. Albert CABELLOS, Prof., Technical University of Catalonia	Rapporteur
M. Mathieu BOUET, Ingénieur de recherche, Thales communications and Security	Rapporteur
M. Benoit DONNET, Prof., University of Liège	Examinateur
M. Olivier BONAVENTURE, Prof., Université catholique de Louvain	Examinateur
M. Luigi IANNONE, Maître de Conférences, Télécom ParisTech	Examinateur
M. Jean-Louis ROUGIER, Prof., Télécom ParisTech	Examinateur
M. Stefano SECCI, Maître de conférences, Université Pierre et Marie Curie	Examinateur
M. Damien SAUCEZ, Chargé de Recherche, Inria	Invité

T H È S E

Télécom ParisTech
école de l’Institut Mines Télécom – membre de ParisTech
46, rue Barrault – 75634 Paris Cedex 13 – Tél. + 33 (0)1 45 81 77 77 – www.telecom-paristech.fr

Acknowledgements

I would like to express my deep gratitude to my supervisors Luigi Iannone and Jean-Louis Rougier for their guidances during this thesis. They did not save any effort to scarify their skills and expertise for the benefit of my thesis. Their valuable comments and ideas have helped in enriching my thesis work and ameliorating its quality. Also, I'm very grateful to them for hosting me in INFRES department inside Telecom ParisTech in Paris, France. In this occasion, I would also like to thank all the researchers and all the members of INFRES department as well.

Furthermore, I would like to express my sincere gratitudes to the jury members: Prof. Benoit Donnet, Prof. Olivier Bonaventure, Mr. Stefano Secci, and Mr. Damien Saucez for attending my dissertation, and, in particular, the reviewers of my manuscript: Prof. Albert Cabellos and Mr. Mathieu Bouet. Their insightful questions and encouraging remarks on my work made my thesis defense a lifelong memory as well as a learning process. It is my utmost honor to have all these experts reviewing my work.

Moreover, my sincere thanks also go to my colleagues in LISP-Lab project: Chi Dung Phung, Patrick Raad, and in INFRES department: Akram Abouseif, Yimeng Zhao, Wenqin Shao, Hao Qin, XiaoXing Yu, Han Qiu, Zhihan Zhang, Nilesh Vyas. I would never forget the days we worked together. Also, these years in Paris, I made many very good friends like Qin Lu, Sebastien Arthaud, Hao Xu, Weiqi Han. Because of you guys, I had a very colorful life, this would definitely be one of my best memories.

Last but not the least, I would like to thank my boy friend Qipeng Song, my parents Jinxi Li and Jing Wang. No matter what decisions I have made, thank you for your endless love and support. Especially for my mother, I will carry her expect and hope on me to welcome the future life.

Abstract

The *Locator/Identifier Separation Protocol (LISP)* was proposed in 2006 to initially address Internet scalability issues. It is based on a map-and-encap mechanism to split the *who* and the *where* of the current IP addresses. To retrieve the association between them, a new network entity called the *Mapping Distribution System (MDS)* is introduced. Although LISP is currently under standardization in Internet Engineering Task Force (IETF) and is deployed in the wild by two testbeds at the same time, it is still young. It lacks the thorough measurement work to show its real performance in large-scale networks.

In this dissertation, we assess LISP from these different aspects.

- The measurements on the MDS: we continuously measured the MDS for seventeen days. The results show that it is stable and consistent most of the time. Nevertheless, instability and inconsistency are observed although they are rare events. We define a new taxonomy to classify them so to facilitate the deeper analysis.
- Proposing a comprehensive monitor to supervise the MDS: as these instabilities and inconsistencies have never been observed, we propose a more comprehensive LISP monitoring architecture to dynamically supervise the whole MDS. After one full month of testing, our proposed monitor provides more mapping information, and can show the MDS performance by 5 metrics.
- The assessment of LISP interworking performance with legacy Internet: we evaluate interworking by conducting two experiments in the different years. Both results show that LISP is stable although a stretch is introduced. The interworking brings negative effects for short paths, but can be ignored for the intercontinental long-distance transmissions.
- The evaluation of LISP mobility: we analyze the LISP mobility performance in three scenarios. We estimate the handover delay and overhead for the different components implementing LISP. Each scenario has its own advantages and shortcomings. Besides, to facilitate the researchers to verify their proposals or test the new features on LISP mobility, we implement LISP mobility extensions under ns-3.27 for an existing ns-3.24 basic LISP simulator.

Résumé

Le protocole *LISP (Locator/Identifier Separation Protocol)* a été proposé en 2006 dans le but de résoudre les problèmes d'extensibilité de l'Internet. Ce protocole est basé sur le mécanisme de map-and-encap qui consiste à séparer une adresse IP en 2 éléments : le *qui* et le *où*. Une nouvelle entité réseau, le *Mapping Distribution System (MDS)* est alors nécessaire pour faire la réconciliation de ces deux éléments. Malgré le fait que le protocole LISP soit déjà standardisé par l'IETF, et déployé sur deux plateformes de tests, c'est toujours un protocole jeune, et il lui manque encore des métriques précises pour montrer toute son utilité sur de larges réseaux.

Dans cette thèse, nous avons évalué LISP selon trois aspects différents.

- Les mesures sur le MDS: nous avons mesuré de façon continue le MDS pendant 17 jours. Les résultats se sont montrés stables et cohérents au cours du temps. De rares instabilités ont cependant pu être observées au cours d'événements ponctuels. Une nouvelle taxonomie a été définie afin de les classer pour faciliter de futures analyses plus détaillées.
- Implémentation d'un outil de monitoring pour superviser le MDS: les instabilités n'ayant encore jamais été étudiées, nous avons défini une méthode de monitoring dynamique du MDS en temps réel. Après un mois complet de mesures, cette méthode a fourni plus d'informations de mapping et 5 indicateurs de performance du MDS ont été définis.
- L'évaluation des performances d'interconnexion sur l'internet patrimonial: nous avons estimé les performances en réalisant deux expériences au cours de deux années différentes. Les résultats ont montré des performances stables bien qu'un stretch a été observé. L'interconnexion apporte des effets négatifs pour les courtes distances réseau, alors que cela peut être ignoré pour les transmissions longues distances intercontinentales.
- L'évaluation de la mobilité de LISP: l'analyse a été faite dans le cadre de trois scenarii différents. Le délai et la charge des transferts des différents composants implémentant LISP ont été mesurés, chaque scenario ayant ses avantages et ses inconvénients. Afin de faciliter de futures mesures de performances par d'autres chercheurs sur la mobilité de LISP, nous avons implémenté une extension "mobilité LISP" sous ns-3.27 pour un simulateur LISP existant en ns-3.24.

Condensé du manuscrit en français

0.1 Introduction

Depuis la création de l'ARPANET (Réseau d'Agence de Projets de Recherche Avancée) en 1966 [84], Internet a obtenu des résultats significatifs au cours des cinquante dernières années et s'est développé très rapidement. Au début de 2018, les utilisateurs d'Internet dans le monde sont 3,82 milliards, c'est-à-dire qu'environ 40% de la population mondiale dispose aujourd'hui d'une connexion Internet [9]. De plus, nous entrons dans l'ère de l'Internet mobile. Selon les dernières recherches [28], le nombre d'utilisateurs mobiles mondiaux a été supérieur à celui des ordinateurs de bureau depuis 2014. Bien qu'ayant un énorme succès, l'Internet d'aujourd'hui souffre de la surcharge de la sémantique des adresses IP, ce qui entraîne de nombreuses lacunes dans l'application paysage [53]. Dans la pile de protocoles TCP/IP actuelle, l'adresse IP sert à la fois d'identification de l'hôte et d'indication du point d'ancrage dans la topologie du réseau, ce qui entraîne un sérieux problème d'évolutivité du routage. Comme le montre la figure 1.1, la taille de la table BGP, qui est le nombre de préfixes IP dans la table de routage BGP (Border Gateway Protocol) des routeurs DFZ (Default Free Zone), a augmenté à un rythme superlinéaire ces dernières années. Après l'utilisation complète d'IPv6, le problème d'évolutivité du routage sera aggravé, car il a un espace d'adressage beaucoup plus important que IPv4. La maintenance d'une telle table de routage est très lourde, puisqu'une fois qu'un site change ses informations de routage, tous les routeurs ont besoin de se mettre à jour, ce qui provoque un énorme BGP. De plus, des mises à jour fréquentes d'informations de routage entraînent une augmentation du temps de convergence et des taux de perte de paquets.

Pour faire face aux problèmes d'évolutivité d'Internet actuels, le paradigme de séparation Loc/ID est proposé, qui sépare l'espace d'adressage IP unique en deux espaces orthogonaux. De nombreuses solutions basées sur ce mécanisme sont proposées, qui peuvent être classées en deux catégories: basées sur l'hôte et sur le réseau. L'ancien doit modifier la pile de l'hôte. Les architectures représentatives sont: HIP (Host Identity Protocol) [94] [92] [91], Six/One [116], SHIM6 (Multihoming de site par IPv6 Intermédiation) [61] [95], ILNP

(Identificateur-Locator) Protocole de réseau) [37] [34], et LNA (Layered Naming Architecture) [35]. L'approche réseau doit mettre à jour les routeurs, et les propositions représentatives sont: GSE (éléments d'adresse globale, de site et d'extrémité) [96], LISP (Locator/ID Separation Protocol) [47], et TIDR (Routage inter-domaine tunnelé) [30]. Parmi ceux-ci, LISP se démarque par ses systèmes de réseau relativement complets dans les avions de contrôle et de données. LISP a été initialement proposé par Cisco en 2006 et dispose d'un groupe de travail IETF exclusif depuis 2009.

La philosophie clé de LISP est de séparer l'espace d'adressage IP en deux sous-espaces: les localisateurs de routage (RLOC) et les identifiants de point de terminaison (EID). La communication de bout en bout est basée sur les EID, ce qui, dans un contexte intra-domaine, signifie simplement effectuer un routage et un transfert conventionnels. Au contraire, pour la communication de bout en bout entre domaines, les RLOC sont utilisés pour le routage et le transfert. Pour ce faire, LISP effectue une opération de mappage et d'encapsulation en mappant d'abord les EID aux RLOC (via un système de distribution de mappage (MDS) 1 à la demande de type DNS 1 [55]), puis en encapsulant le paquet d'origine pour l'utiliser. RLOC sur l'en-tête externe. L'interréseautage entre les réseaux LISP et les réseaux non-LISP est assuré par deux types de proxies [76]: les routeurs de tunnel Proxy Ingress (PITR) et les routeurs de tunnel d'évacuation Proxy (PETR). Ils sont collectivement désignés par PxTR. Les PITR annoncent dans l'infrastructure de routage BGP l'espace de préfixe EID pour le compte des sites LISP, afin que les sites non-LISP puissent les atteindre. Les PETR sont utilisés comme points de sortie des sites LISP lorsque la destination n'est pas un EID dans un site LISP. Grâce à ce mécanisme, LISP est en mesure d'améliorer l'évolutivité du routage Internet, d'améliorer l'ingénierie du trafic interdomaine, la mobilité VPN multi-locataire légère et la mobilité des VM dans les Data-Centers [110].

LISP étant une technologie prometteuse qui apporte de nombreux avantages à Internet, elle attire l'attention du monde académique et industriel. De nombreuses entreprises, universités et instituts de recherche sont désormais impliqués dans le développement du LISP. Par conséquent, la normalisation et la recherche connexes sont très actives. En termes de déploiement, deux plateformes LISP expérimentales de grande envergure: LISP Beta Network et LISP-Lab sont respectivement déployées sur le monde en 2008 et 2015. A ce jour, tous les efforts portent principalement sur l'analyse de performance théorique ou de simulation, description des architectures LISP, implémentations, extensions, applications et ainsi de suite. Cependant, il manque de travail de mesure complet pour la technologie LISP elle-même. Ces travaux de mesure de la performance donneront à la communauté de recherche une vision claire de la performance du LISP, révéleront les points qui peuvent être améliorés, indiqueront les directions de recherche possibles et amélioreront les performances de deux

plates-formes. Ainsi, les mesures à grande échelle dans des réseaux réalistes sont d'une importance primordiale pour le développement du LISP.

0.2 Motivation

Dans le but d'évaluer le LISP à partir des différents points de vue, nous visons dans cette thèse à répondre aux questions de recherche suivantes, qui n'ont pas été abordées auparavant:

1. Le système de cartographie est essentiel dans LISP, car il détermine où les paquets seront acheminés. Ainsi, fournit-il toujours les mêmes informations de mappage pour la même destination? Est-ce que tous les résolveurs de cartes (MR) du système de cartographie offrent la même réponse en même temps? Les réponses des informations de mappage restent-elles cohérentes si nous les interrogeons à partir des différents points de vue (VP)?
   Nous avons continuellement mesuré le réseau Beta de LISP pendant dix-sept jours. Les mesures montrent que le système de cartographie LISP est stable la plupart du temps. Les réponses des différents MR et VP sont également essentiellement cohérentes. Néanmoins, l'instabilité et l'incohérence sont observées bien que ce soient des événements rares. Nous définissons une nouvelle taxonomie pour classer ces instabilités et incohérences afin de faciliter une analyse plus approfondie. Ce travail est publié dans WNM [81].
2. Comme les résultats de la dernière question montrent qu'il existe une incohérence entre les MR et les VP, est-il possible de mettre en place un moniteur complet pour superviser simultanément l'ensemble du système de cartographie et tous les résolveurs? Nous proposons une architecture de surveillance LISP dynamique, à savoir LISP-Views, afin d'approfondir les connaissances sur LISP et de faciliter les opérations et les dépannages au jour le jour. Après un mois complet de surveillance de l'ensemble du système de cartographie LISP, il démontre que LISP-Views fournit les informations de cartographie complètes. De plus, avec notre plate-forme de surveillance proposée, il est possible d'évaluer plus de métriques de performances du système de cartographie, telles que la fiabilité, la latence et les problèmes de configuration, ce qui permet d'améliorer encore le LISP. Ce travail est publié dans ITC [77].
3. En utilisant l'approche proxy pour communiquer avec l'Internet hérité, un tronçon dans le chemin est introduit. Combien cela représente-t-il dans les performances du monde réel lorsque les plates-formes intègrent l'Internet hérité? Dans quelles con-

ditions un tel étirement a-t-il un impact important sur la performance? Inversement, dans quelle situation ces frais généraux sont-ils si faibles qu'ils peuvent être ignorés? Pour évaluer les performances d'interfonctionnement de LISP avec Internet, nous menons deux campagnes de mesure. Les résultats des deux expériences sont cohérents. Les deux montrent que LISP est stable bien que l'étirement soit introduit en raison de l'utilisation de proxies. Nous trouvons que le proxy introduit en effet des effets négatifs pour les destinations localisées en Europe et en Amérique, mais peut être ignoré pour les transmissions intercontinentales longue distance vers les destinations asiatiques. En outre, nous observons que la position des mandataires est très importante. Si c'est proche des sources ou les destinations peuvent beaucoup diminuer la latence. Ce travail est publié dans l'atelier étudiant INFOCOM [79], l'atelier étudiant ACM SIGCOMM [78] et Elsevier Computer Networks [80].

4. Comme LISP sépare l'identificateur et le point d'attache topologique de l'adresse IP traditionnelle, il est capable de maintenir la communication entre deux hôtes pendant le transfert. LISP peut être implémenté sur les routeurs de frontière ou les terminaux ou même les deux. Ainsi, quel élément de réseau supporte le LISP présentant une meilleure performance, c'est-à-dire, ayant un délai de transfert plus court ou introduisant moins de frais généraux pendant la mobilité? Quels sont les avantages et les inconvénients de chaque cas?

   Pour explorer les performances de mobilité de l'implémentation LISP sur chaque élément de réseau, nous concevons trois scénarios différents: 1) Seul le noeud mobile supporte LISP; 2) Seuls les routeurs de bordure supportent LISP; 3) Les routeurs de nœud mobile et de bordure supportent LISP. L'analyse numérique montre que les scénarios 1 et 3 peuvent réaliser le déplacement du nœud mobile à travers les différents sous-réseaux, alors qu'ils nécessitent un EID permanent supplémentaire pour chaque nœud mobile. Le 1er n'a aucune aide sur la réduction de la taille de la table BGP et le 3ème scénario montre un retard de transfert beaucoup plus long que les deux autres. Le deuxième scénario présente le plus petit délai de transfert et nécessite le moins d'adresses IP, mais il ne peut gérer que la mobilité au sein du même sous-réseau. Pour vérifier l'analyse numérique, nous implémentons un simulateur LISP avec des extensions de mobilité LISP sous ns-3.27 en étendant une implémentation LISP basée sur ns-3 existante.

0.3 Contexte

0.3.1 Architecture LISP

Le *Locator / Identifier Separation Protocol (LISP)* sépare le rôle d'adresse IP traditionnel en deux sous-espaces logiques: (i) *Endpoint IDentifier (EID)* et (ii) *Routing LOCator (RLOC)*. Les deux sont les adresses classiques IPv4 [104] ou IPv6 [45]. L'EID est l'identificateur des hôtes, qui est localement routable dans le site LISP. Normalement, ce sont les adresses des hôtes source et destination. Au sein du même site LISP, les hôtes peuvent communiquer directement entre eux via des EID comme dans un réseau IP hérité. L'hôte reçoit un EID de destination de la même manière qu'il obtient l'adresse IP de destination aujourd'hui, par exemple via une recherche DNS (Domain Name System) [88] ou un protocole SIP (Session Initiation Protocol) [106]. Les RLOC indiquent les points de rattachement des EID dans la topologie Internet, c'est-à-dire, l'adresse source / destination des routeurs de périphérie, derrière lesquels les EID peuvent être trouvés. Les adresses RLOC sont utilisées pour transférer les paquets dans l'Internet central.

Lorsque les paquets sont transférés entre deux sites LISP différents, une encapsulation sur le routeur périphérique est nécessaire: les EID source et distant sont dans l'en-tête de paquet interne, tandis que les RLOC source et destination sont dans l'en-tête externe. Les préfixes EID ne doivent pas être annoncés sur le noyau Internet, donc la taille de la table de routage BGP peut être réduite. De même, dans les réseaux de bout, il n'est plus nécessaire de connaître les informations de routage de l'Internet central. Les opérations d'encapsulation et d'acheminement susmentionnées sont la responsabilité de LISP Data Plane. Plus de détails sont introduits dans Sec. 2.1.1. Pour lier les EID et les RLOC, une nouvelle entité réseau appelée Mapping Distribution System (MDS) est utilisée, dont la partie frontale est constituée de Map Resolver (MR) et de Map Server (MS) [56]. Le MDS est non seulement capable de fournir les informations de mappage entre les EID et les RLOC pour les sites LISP (nous les appelons simplement LISP Map-Replies), mais aussi d'indiquer quelles adresses ne sont pas des sites LISP et n'ont pas de RLOC. de l'Internet hérité (ils s'appellent des réponses de carte négatives). Le plan de contrôle est décrit en Sec. 2.1.2.

0.3.2 Interfonctionnement avec Internet de la légitimité

Bien que les EID soient syntaxiquement identiques aux adresses IPv4 ou IPv6, les routes vers ces adresses ne sont pas annoncées dans le système de routage global. Un mécanisme d'interopérabilité est donc nécessaire pour que les sites non lisp puissent échanger du trafic avec des sites LISP. Ainsi, deux nouveaux éléments de réseau sont introduits: RTPP Proxy

Ingress Tunnel Router (PITR) et LISP Proxy Egress Tunnel Routeur (PETR). Le PITR agit comme un ITL LISP intermédiaire pour les hôtes sur l'Internet hérité. Il est chargé d'annoncer un ou plusieurs préfixes EID hautement agrégés pour le compte des sites LISP dans l'Internet public et reçoit le trafic de l'Internet public pour les encapsuler comme des paquets LISP. Le PETR agit comme un ETR pour le trafic destiné aux sites non LISP. Il reçoit le trafic des ITR, décapsule les paquets LISP en paquets IP conventionnels et les envoie à l'Internet hérité.

0.3.3 LISP mobilité

De nos jours, l'explosion des appareils mobiles qui ont besoin du trafic Internet mobile massif nécessite l'évolution de l'architecture du réseau mobile. Garder la communication pendant l'itinérance sans interruption devient un point clé. Grâce à LISP séparant l'identifiant d'un hôte et le localisateur du point d'attache, la connexion entre deux hôtes peut être maintenue en vie en itinérance. Parce que l'EID, qui est utilisé comme adresse interne, est permanent, alors que seul RLOC (utilisé comme adresse externe) est changé.

En utilisant la mobilité LISP, il est possible 1) permettre aux connexions TCP de rester actives pendant l'itinérance; 2) pour fournir les chemins de données bidirectionnels les plus courts entre un nœud mobile (MN) et un nœud correspondant (CN); 3) pour permettre à MN de communiquer avec un autre MN lorsque les deux sont en itinérance; 4) ne pas exiger de routes à granularité fine dans le réseau central, ni d'éléments de réseau domestique, d'agent étranger ou d'autres éléments de réseau de plan de données à l'appui de la mobilité; 5) il n'y a pas de routage en triangle des paquets de données comme dans Mobile IP [97]; 6) et il n'y a pas de nouveaux en-têtes d'extension IPv6 pour éviter le routage en triangle [71]. LISP Mobile Node (LISP-MN) profite de l'infrastructure LISP pour surmonter les limites imposées par Mobile IP. L'implémentation de LISP-MN est réalisée par Open Overlay Router (OOR) [21], qui sera introduit dans Sec. 3.3.1.

LISP peut être implémenté sur les routeurs de bordure et les terminaux. Selon l'adoption de LISP et de l'hôte / routeur LISP, la mobilité peut être divisée en 4 catégories: LISP-MN dans LISP-Site, LISP-MN dans un site non-LISP, MN dans LISP-Site et MN dans site non-LISP (montré dans l'onglet 2.1). Cependant, le cas "MN dans un site non-LISP" est la mobilité classique actuelle, qui n'a aucun rapport avec LISP. Il y a quelques propositions pour résoudre les problèmes de mobilité pour le handover traditionnel, comme Mobile IP (MIP) [99], Mobility Support dans IPv6 (MIP6) [98] [87], Multipath TCP (MPTCP) [54] etc. Ainsi, dans les sections suivantes, nous présenterons d'abord LISP-MN puis seulement décrirons comment les paquets s'échangent lorsque le transfert se produit pour les trois premiers cas.

0.3.4 LISP Beta Network

Le premier banc d'essai LISP global appelé LISP Beta Network [11] [44], a été déployé depuis 2008. Lancé comme seul banc d'essai LISP, le réseau LISP Beta connaît une croissance régulière et atteint un déploiement expérimental mondial [107] pour acquérir une expérience concrète sur LISP. Initiés par Cisco, les membres du LISP Beta Network ne sont pas seulement des entreprises et des opérateurs majeurs, mais aussi des universitaires, des laboratoires de recherche et des startups offrant des services liés au LISP. Les participants de ce réseau sont situés dans 34 pays différents, avec la plus forte concentration en Europe et en Amérique du Nord.

Le LISP Beta Network est composé du système de cartographie, des routeurs LISP (c'est-à-dire xTRs), des nœuds mobiles, des routeurs de tunnel de réencapsulation (RTR [46]) et des routeurs de tunnel de proxy (PxTR [76]). Le 14 mars 2012, le système de cartographie utilisé par le réseau bêta de LISP est passé de LISP+ALT [57] à LISP-DDT plus flexible [58]. La modification du système de cartographie a considérablement réduit la charge de configuration et de maintenance, tandis qu'une légère dégradation des performances a été observée [107]. Selon la dernière architecture (octobre 2017) [12] illustrée à la Fig. 3.1, LISP Beta Network a 3 racines DDT, 2 TLD DDT, 7 MR/MS disponibles et 6 PxTR. Plus précisément: 2 MR/MS et 2 PxTR sont aux USA, 3 MR / MS et 3 PxTR en Europe, 2 MR/MS et 1 PxTR en Asie. Il utilise principalement deux espaces d'adressage EID, qui appartient à Cisco: 153.16.0.0/16 pour IPv4 et 2610:00D0::/32 pour IPv6 [107], mais il existe également d'autres EID.

0.3.5 LISP-Lab platform

Le projet LISP-Lab visait à construire une plate-forme ouverte, basée sur l'architecture LISP, fournissant un environnement pour effectuer des recherches de haute qualité et soutenir la conception, le développement et l'évaluation approfondie de nouveaux services et cas d'utilisation. La gamme de tâches techniques planifiées dans le projet LISP-Lab, de la mise en réseau cloud à l'accès aux technologies, en passant par la connectivité interdomaines, l'ingénierie du trafic et la gestion de cartographie, stimulant l'innovation au-delà de la technologie LISP. La plate-forme LISP-Lab a été déployée uniquement à l'aide de l'implémentation LISP open source, OpenLISP [23]. Il était coordonné par un consortium français: deux institutions académiques (UPMC, TPT), deux PME du Cloud Networking (Alphalink, NSS), deux opérateurs réseaux (Renater, Orange), deux PME sur Access / Edge Networking (Border 6, Ucopia) et un point d'échange Internet (Rezopole). Dix partenaires externes internationaux ont également rejoint la plateforme. Au moment de la rédaction,

LISP-Lab est constitué de 3 MR / MS et de 2 PxTR situés en France, ainsi que de 13 xTRs mondiaux, et a déjà été interconnecté au réseau LISP Beta avec une racine OpenLISP DDT [59].

0.3.6 LISPmon

LISPmon est la seule plate-forme de surveillance qui supervise le statut LISP public actuel, c'est-à-dire les mappages entre les préfixes EID et les RLOC. Cette plate-forme est développée par le Centre de communication avancé à haut débit de l'UPC (Universitat Politècnica de Catalunya). Il analyse l'ensemble de l'espace d'adressage IPv4 tous les jours, normalement à partir de 7h00 (UTC), envoyant la demande de carte à un MR préféré du réseau LISP bêta (c'est-à-dire que 7 MR sont des candidats à sélectionner). Si ce MR ne répond pas, l'opérateur LISPmon en choisira un autre en remplacement, et relancera les requêtes, pour garantir que la réception des Listes de Réponses LISP se déroule sans heurt entre deux jours consécutifs. Les informations de cartographie EID-RLOC disponibles sont publiées une fois par jour, depuis le début de 2010.

0.4 Évaluation de LISP mapping system

0.4.1 Vue générale des mesures

Pour obtenir une évaluation complète et approfondie de LISP Beta Network Performance, nous avons mené des expériences en utilisant LIG (LISP Internet Groper [52], voir la section 3.2.2 pour les détails). Lorsque LIG est lancé, une demande de carte est envoyée pour une adresse IP de destination, et la réponse de la carte retournée est ensuite stockée pour une analyse plus poussée. Il existe trois types de résultats possibles: (i) *LISP Map-Reply*, où la réponse contient des informations de mappage pour atteindre les sites activés LISP; (ii) *Negative Map-Reply*, une réponse par carte négative, où la réponse indique que l'adresse IP demandée appartient à un préfixe d'un site non LISP; (iii) *No Map-Reply*, Aucune réponse par carte, où aucune réponse n'est reçue dans un certain temps. Une requête est considérée comme réussie si une réponse de carte LISP ou une réponse de carte négative est reçue. Une requête est considérée comme ayant échoué si aucune réponse n'est reçue dans les 5 secondes.

Les expériences sont menées du 2 au 18 juillet 2013. A partir d'un ensemble de différents points de vue (PV), envoyer des *Map-Request* messages à tous les résolveurs de carte du LISP Beta Network, toutes les 30 minutes. L'ensemble des VP était composé d'une partie des réseaux académiques, de l'Internet commercial, de PlanetLab et diffusé en Europe

et aux USA. Les adresses IP de destination comprenaient la première adresse dans chaque préfixe dérivé du projet LISPmon [18], quel que soit le type (c'est-à-dire, les préfixes EID ou les préfixes réguliers). Au cours de cette campagne de mesures, 5 VP et 13 MR/MS étaient disponibles et 613 préfixes ont été interrogés à chaque tour.

Nous avons analysé les réponses aux cartes associées aux EID au lieu de toutes les adresses IP normales. Dans le reste, nous utilisons le terme round expérimental pour identifier un échange Map-Request / Map-Reply complet par LIG pour toutes les destinations sélectionnées pour tous les MR disponibles sur tous les VP existants. Bien que No Map-Reply apparaisse occasionnellement dans certains rounds expérimentaux, nous nous concentrons sur LISP Map-Reply et Negative Map-Reply pour effectuer l'analyse. [1]

0.4.2 Metrics

Le plan de contrôle LISP est chargé d'offrir les informations de mappage qui consistent en un préfixe EID associé à une liste de tuples $<RLOC, Priority, Weight>$. Théoriquement, sauf si la configuration a changé, si le même préfixe EID est interrogé, les réponses à la carte doivent toujours contenir les mêmes informations de mappage, quel que soit le VP à partir duquel la requête est envoyée et quel que soit le MR interrogé. Au cours de nos mesures, cependant, nous avons constaté que les réponses aux cartes ne sont pas toujours identiques à ce qu'elles sont supposées être. Les raisons à l'origine des changements de carte-réponse sont résumées dans le tableau 4.1.

Les différences Map-Reply peuvent être classées en deux grandes catégories: la première est une différence en termes de type Map-Reply, ce qui signifie qu'un mélange de Negative Map-Reply et de LISP Map-Reply se produit; l'autre type de différence est dans l'ensemble RLOC, ce qui signifie qu'au moins l'un des attributs associés à l'ensemble RLOC est différent. Plus précisément, le nombre RLOC indique une différence dans le nombre de tuples RLOC reçus dans la carte-réponse. Les autres sont des différences dans un tuple RLOC, qui se compose d'une adresse RLOC, d'une priorité RLOC et d'une pondération RLOC (présentées à la section 2.1.2).[2] L'état RLOC est dérivé du bit R dans le message de réponse à la carte, qui est défini lorsque l'expéditeur d'une réponse de carte a une route vers le RLOC, ce qui signifie que de son point de vue, le RLOC est actif. Si au moins l'un des six paramètres indiqués dans le Tableau 4.1 transmis dans une réponse à la carte est différent du précédent (dans le temps) ou des autres (entre différents VP ou MR) au même tour d'expérience, la réponse à la carte est marquée comme différent. Sur la base des critères

[1] Nous utilisons "simultanément" ou "en même temps" comme synonyme de "au même tour d'expérience".
[2] Les termes adresse RLOC et RLOC sont utilisés de manière interchangeable.

de la différence Map-Reply, nous proposons deux métriques pour évaluer plus en détail les performances du système de cartographie LISP: (i) *Stability* et (ii) *Consistency*.

Un mappage est considéré comme présentant une stabilité si les réponses de la carte pour un EID spécifique obtenu à partir d'un MR observé à un VP donné (brièvement noté comme un tuple $<EID, MR, VP>$), dans le temps, sont identiques. Une taxonomie de classification pour les cas instables est proposée plus loin dans la Sec. 4.2.

Une mapping présente une cohérence si: *i*) les réponses à la carte pour un EID spécifique de tous les MR observés à l'intention de VP (dénoté comme groupe $<EID, *, VP>$) sont identiques à l'heure actuelle (ce type de cohérence est appelé *Consistency by MR*); *ii*) Les réponses de la carte pour une paire EID-MR spécifique dans tous les VP (désignées comme une paire $<EID, MR, *>$) sont identiques en même temps (appelée *Consistency by VP*). Une taxonomie de classification pour les cas non cohérents est proposée dans la Sec. 4.3.

0.4.3 Résultats

Globalement, lorsque l'on considère l'ensemble des données, 91,49% des observations sont stables, ce qui indique que sur l'ensemble des observations, les changements sont peu fréquents et le MDS est stable. Lorsque la stabilité est considérée sur une base de point de vue, nous observons, comme le montre la Fig. 4.1, que la stabilité est indépendante du point de vue car aucun écart significatif par rapport à la moyenne globale ne peut être observé (ligne pointillée sur la Fig. 4.1) . Néanmoins, alors que ces résultats démontrent que le système de cartographie est généralement stable, dans la suite de Sec. 4.2, nous étudions la raison pour laquelle environ 8,51% des observations présentent une instabilité.

Le pourcentage global de cohérence par MR est de 86,3% (voir la ligne pointillée sur la figure 4.5), ce qui signifie que pour toute l'expérience, l'incohérence entre les réponses par carte envoyées par différents MR n'est pas commune. Cependant, nous pouvons également observer à partir de la figure 4.5 que différents VP montrent des écarts par rapport à la cohérence globale de MR, reflétant le fait que si certains VP reçoivent la même réponse sur carte de différents MR, certains autres reçoivent des réponses incohérentes. Cet écart est très faible (pas plus de 2,77%) entre les différents points de vue, ce qui indique qu'il reste un événement relativement rare. Ces résultats montrent que le système de cartographie est cohérent entre différents MR en général, mais l'incohérence existe en même temps avec un pourcentage autour de 13,7%. Dans la suite de Sec. 4.3.1, nous approfondissons ce problème en explorant les raisons pour lesquelles les quelques observations présentent une incohérence de la part de MR.

Le pourcentage global de cohérence par VP pour l'ensemble de l'ensemble de données de l'expérience est de 90,48%, présenté par la ligne pointillée de la Fig.4.7, qui indique que

les réponses-réponses reçues à différents VP du même MR sont très cohérentes. Cependant, nous pouvons également observer que différents MR montrent des écarts par rapport à la cohérence globale de VP, reflétant le fait que certains MR envoient la même Map-Reply à tous les VP, mais d'autres non. L'écart entre les différents MR est toutefois limité (moins de 12,23%). Par conséquent, nous pouvons certainement affirmer que le système de cartographie montre généralement un degré élevé de cohérence par VP mais présente un nombre non négligeable d'incohérences (9,52%). Dans le reste de la Sec. 4.3.2, en suivant la même approche que pour la cohérence par MR, nous approfondissons cette question en explorant les raisons pour lesquelles certaines observations présentent des incohérences de la part de VP.

Au cours de l'évaluation des performances du système de cartographie LISP, outre les observations sur la stabilité et la cohérence que nous décrivons en détail dans les sections précédentes, nous étudions également des résultats intéressants liés au LISP. Nous les fournissons ici afin qu'ils puissent être utilisés pour améliorer les performances du LISP. Nous avons identifié 8,97% des sites LISP qui utilisent plusieurs RLOC pour être multi-hôtes sur Internet, avec une moyenne de 3 localisateurs distincts. Fait intéressant, lorsque nous comparons les localisateurs avec leur numéro de système autonome, nous avons constaté qu'environ 50% des localisateurs appartiennent à des AS différents, ce qui indique que LISP est utilisé pour interconnecter de manière transparente les sites LISP par l'intermédiaire de plusieurs FAI. Comme le plan de contrôle (c.-à-d. MS/MR et système de cartographie) et le plan de données (c.-à-d. ETR/ITR) sont découplés dans LISP, Map-Replies peut provenir directement des sites LISP ou d'utilitaires partagés. constituant le système de cartographie (c'est-à-dire MS/MR). Cette particularité de LISP est déjà utilisée dans le déploiement actuel, où seulement 88,8% des Map-Replies non négatifs proviennent directement de leurs sites LISP, les 11,2% restants étant issus des MS déléguées répondant au nom de la LISP-sites. Cela permet de réduire le trafic de contrôle observé par les sites LISP et de les aider à mieux évoluer.

En plus des mappages, LISP repose également sur l'encapsulation, ce qui signifie que le trafic LISP peut être transporté sur IPv4 ou IPv6 sans que les hôtes finaux soient à double pile. Fait intéressant, nous avons identifié que pas moins de 8,84% des localisateurs sont IPv6. Cependant, nous n'avons pas observé de cas où un site LISP était accessible en utilisant strictement IPv6 alors que 82,79% des sites ne sont accessibles que via IPv4. Cette divergence tend à confirmer que les opérateurs de réseaux manquent toujours de confiance pour être accessibles uniquement via IPv6. Par conséquent, les 17,21% des sites LISP utilisant explicitement IPv6 incluent toujours au moins une adresse IPv4 pour être joignables.

0.5 LISP-Views: moniteur du LISP mapping system

0.5.1 Motivation

Afin de faire avancer LISP, nous devons comprendre en profondeur le comportement des différentes entités du réseau LISP et puisque le MDS reflète l'état d'un réseau LISP puisqu'il stocke toutes les informations de cartographie, il est essentiel de pouvoir surveiller le MDS. LISPmon a été le premier pas vers un suivi LISP systématique. Cependant, il surveille le MDS juste à partir d'un point de vue (VP) une fois par jour et interroge seulement un MR. Sur les problèmes de MR, LISP-mon doit être reconfiguré manuellement pour surveiller un autre MR. Cependant, les informations de mappage peuvent être instables et incohérentes entre les RM, c'est-à-dire que les informations de mappage changent parfois en un jour et les réponses de carte des différents MR peuvent ne pas coïncider à un instant donné [81]. Il est similaire aux serveurs de fenêtres de recherche BGP, qui ne fournissent pas toujours les mêmes réponses pour une adresse IP car l'ensemble du système de routage n'a peut-être pas convergé ou à cause de politiques de routage. De ce point de vue, LISPmon a quelques limitations car il n'est pas capable de détecter les changements d'informations de cartographie en un jour, et n'est pas capable de montrer les différences entre les MR. Ainsi, nous proposons une nouvelle architecture de surveillance LISP polyvalente, appelée LISP-Views, pour surveiller les déploiements LISP publics, ainsi que pour permettre une évaluation plus poussée des performances de LISP définie par les utilisateurs eux-mêmes. En fait, LISP-Views peut non seulement être utilisé pour surveiller LISP, mais peut également être étendu pour être utilisé dans des réseaux non publics, tels que VxLAN [82].

LISP-Views est une implémentation open source et a été conçu pour atteindre les objectifs suivants:[3]

1. LISP-Viewsquerestallant sur le travailMRsexistingingMDS actuel, inparallel, alors que LISPmon en interroge un seul, visant à construire une vue complète du statut LISP actuel.
2. LISP-Views surveille périodiquement tous les MR avec des intervalles arbitraires, alors que LISP-mon le fait tous les jours, avec pour objectif de fournir des informations sur l'évolution de la cartographie à une granularité plus courte.
3. LISP-Views supervise l'intégralité du MDS sans aucun processus manuel, réagissant automatiquement aux composants défaillants (par exemple, les MR non répondeurs).
4. LISP-Views obtient les informations de mappage de tous les MR du réseau bêta de LISP ainsi que de la plateforme LISP-Lab, tandis que LISPmon préfère tirer parti de la MR de l'ancienne.

[3]Code source disponible sur Github: https://github.com/SeleneLI/LISP-Views

5. LISP-Views est flexible et configurable, avec des utilisateurs capables de définir différentes tâches de surveillance et d'obtenir les différentes mesures, alors que LISP-mon ne fait que publier quotidiennement la liste de mapping.
6. LISP-Views est conçu pour être facilement déployé de manière distribuée.

0.5.2 Déscription de LISP-Views

L'architecture des outils de surveillance LISP-Views proposés est illustrée à la Fig. 5.1. LISP-Views se compose de plusieurs modules avec des fonctions différentes. Les modules Measurement, Report et Raw Data sont déployés sur un serveur centralisé tandis que les modules Crawler, Sonar et Controller peuvent être déployés sur plusieurs VP différents. Pour le moment, LISP-Views vient de se déployer sur un VP à Paris, en France, où le moniteur est développé et le module centralisé déployé. La version distribuée de LISP-Views est en cours de développement. Pour cette raison, nous ne discuterons pas plus en détail de la manière de déployer LISP-Views sur plusieurs VP. Tous les modules sont implémentés en Python et décrits comme suit:

Sonar module: le module principal avec deux fonctions: 1. envoyer les Map-Requests encapsulées LISP et recevoir les Map-Replies de / vers toutes les MR existantes basées sur la norme [47]; 2. stocker les informations reçues dans les rapports et les données brutes à des fins différentes.

L'adresse IP utilisée par Sonar pour interroger MR est sélectionnée soit à partir de la sortie de Crawler, soit à partir des informations enregistrées dans le rapport précédemment produit. Les raisons sont expliquées dans les modules correspondants.

Crawler module: analyse tout l'espace d'adressage IPv4 existant à l'aide de Sonar. Si l'adresse IP interrogée (par exemple 192.0.2.1) n'a pas de réponse de carte, Crawler incrémente l'adresse IP d'un IP comme adresse IP suivante demandée (c'est-à-dire, 192.0.2.2) et laisse Sonar envoyer la demande de carte. Si Sonar obtient Negative ou LISP Map-Reply, étant donné que Reply contient un préfixe (mentionné dans la section 2.1), Crawler définit la requête suivante sur le premier IP au-delà du préfixe renvoyé. Par exemple, si le préfixe renvoyé est 192.0.2.0/25, l'adresse IP interrogée suivante est 192.0.2.128. L'analyse de l'ensemble de l'espace d'adressage IPv4 prend beaucoup de temps et est principalement causée par le temps perdu en attente de non-réponse à la carte. En effet, aucune réponse sur la carte ne signifie attendre 3 secondes, et l'adresse IP suivante est juste augmentée d'une unité, au lieu de sauter un bloc d'adresses IP comme dans le cas de la réponse par carte négative/LISP.

Report module: contient les préfixes EID collectés dans une liste, ainsi qu'une liste des MR qui ont répondu. En rampant, l'espace d'adressage peut prendre beaucoup de temps,

mais nous visons à obtenir le statut de MR le plus souvent possible. Sonar envoie des demandes de carte pour les EID enregistrés dans le rapport uniquement aux MR qui ont déjà répondu. Ainsi, il diminue la possibilité de ne pas recevoir de réponse par carte, ce qui permet d'obtenir le statut LISP dans un délai plus court.

Raw Data module: contient toutes les informations détaillées de Map-Replies pour chaque MR, telles que Map-Reply type, RLOCs, EID-Prefix requis, Round Trip Time (RTT) et la source retournée pour chaque tour spécifier avec TimeStamp (indépendamment de la source de Sonar) . Ainsi, il peut être utilisé pour effectuer une analyse approfondie des performances de MDS. De plus, étant donné que les données brutes sont stockées conformément au MR, il est possible de suivre les performances des MR individuellement.

Measurement module: fournit la composition des mesures demandées (c'est-à-dire, sélectionne différents plug-ins de mesure) sur la base de *Raw Data* et de *Report*.

REST API: est connecté au module de mesure afin que les utilisateurs puissent lancer une expérience personnalisée en définissant le temps d'expérience, les MR surveillés et obtenir les différents aspects du statut LISP. Nous implémentons actuellement l'API REST, donc les résultats des mesures présentées ici ont été obtenus via les lignes de commande.

Controller module: synchronise tous les modules dans LISP-Views et spécifie également l'heure de début et de fin de la production de *Report* et de *Raw Data*. L'intervalle de génération de *Report* et de *Raw Data* peut être modifié en fonction de la capacité de traitement du matériel. Toutefois, étant donné que l'incohérence existe entre les MR, l'intervalle de production des rapports et des données brutes pour chaque MR diffère l'un de l'autre. Ainsi, l'intervalle défini dans le module Contrôle doit couvrir le MR le plus lent.

0.6 Validation de LISP-Views

Afin de valider et d'évaluer l'outil de suivi LISP-Views, nous avons utilisé des données brutes et des rapports collectés pendant un mois (du 0 au 4 septembre 2016 à minuit du 4 octobre 2016), en déployant LISP-Views sur un VP, qui est un xTR de la plateforme LISP-Lab.

L'intervalle pour produire les rapports était de 6 heures, et l'intervalle pour produire les données brutes était de 2 heures. Malheureusement, les deux MR de la plate-forme LISP-Lab ont tout juste répondu aux demandes de carte à la première fois, puis ont arrêté. Ils n'étaient pas en mesure de gérer un grand nombre de requêtes, car les demandes de mappage remplissaient les files d'attente de requêtes, ce qui entraînait une baisse des requêtes de requête. En soi, c'est un succès, car ce bug était inconnu, et les codeurs OpenLISP ont corrigé le problème après nos mesures. Tous les résultats de l'évaluation utilisés dans cette

section ne dépendent que des 6 MR du LISP Beta Network (3 en Europe, 1 aux Etats-Unis et 2 en Asie) et les 6 autres MR étaient en baisse au moment de l'expérimentation. Le but du travail est de valider LISP-Views en évaluant si elle fournit au moins les mêmes résultats que LISPmon.

Dans cette section, nous comparons LISPmon et LISP-Views, afin d'évaluer si les informations fournies sont comparables entre les deux plates-formes de surveillance, validant ainsi LISP-Views. Comme nous l'avons indiqué dans Sec. 5.1.1, à tout moment LISPmon interroge un MR, sur le réseau Beta LISP une fois par jour et commence généralement à 7h00 du matin. Le MR interrogé est parfois modifié lorsqu'il a des problèmes. LISP-Views, cependant, continue d'envoyer des demandes de cartes à tous les 6 MR avec un intervalle de 2 heures tous les jours, afin de récupérer le statut réel réel de chaque MR.

La Fig. 5.2 montre le nombre de Réponses de Carte LISP reçues de LISPmon et de LISP-Views au fil du temps. La sous-figure supérieure fait une comparaison entre les deux moniteurs. La ligne avec point est obtenue à partir des publications quotidiennes de LISPmon, indiquant le nombre de réponses de carte LISP retournées par un MR dans un cycle de surveillance de temps incertain tous les jours. Pour faciliter la comparaison, nous obtenons les résultats de LISP-Views à 8h00, ce qui est le cycle expérimental le plus proche de LISPmon. Comme présenté dans la courbe avec des symboles triangulaires, LISP-Views reflète le fait d'une combinaison du nombre de LISP Map-Reply contenant les différents préfixes EID de tous les MR à un temps d'expérimentation fixe. Dans la plupart des cas, LISP-Views reçoit environ 20 Map-Replies LISP plus que LISPmon, ce qui reflète le fait que les MR ne sont cohérentes à aucun moment en raison de l'existence d'un temps de convergence pour synchroniser les informations de mapping entre elles. Les 5 autres jours, où LISP-Views reçoit très peu de Réponses cartographiques LISP, montrent qu'il y avait des problèmes pour recevoir normalement le statut LISP à ce moment-là, alors que LISPmon présente toujours une tendance plus lisse mais cache la réalité.

Le manque de stabilité peut être mieux apprécié dans la sous-figure du bas de la Fig. 5.2. Cette figure représente le nombre de LISP-Map-Reply (en pointillés) et le mappage global (LISP et Negative Map-Reply, en trait plein) de LISP-Views pour chaque ronde de surveillance de 2 heures pendant toute l'expérience. Bien que le nombre de réponses par carte présenté dans la figure soit toujours la combinaison des résultats de 6 MR lorsque les préfixes EID renvoyés sont différents, cela montre que les MR ne sont pas tout à fait stables. Parfois, les MR fournissent même 0 ou moins de 10 réponses de carte LISP, mais dans le tour d'expérience suivant, le nombre est récupéré. En outre, nous évaluons le nombre total de requêtes réussies toutes les 2 heures. Il présente également le comportement de l'instabilité, où il oscille généralement entre 400 et 500, mais il tombe occasionnellement plus bas que

300. Les lignes pleines et pointillées ont presque la même tendance changeante, indiquant que le changement de Map-Replies est principalement affecté par LISP Map-Replies. Les rondes de surveillance où le nombre de réponses de carte LISP approche de 0 mais où des réponses de carte négatives sont encore reçues indiquent des problèmes au niveau des MR, en particulier pour répondre à la réponse de carte LISP.

Quelques exemples de la manière dont LISP-Views peut être utilisé pour disséquer LISP et obtenir des résultats détaillés sont présentés dans Sec. 5.3.

0.7 Évaluation de la performance d'interfonctionnement LISP grâce à RIPE Atlas

Grâce à deux bancs d'essais: le LISP Beta Network et le projet LISP-Lab, LISP est progressivement déployé dans la nature. Le mécanisme d'interfonctionnement est proposé pour assurer la communication entre les sites parlant le LISP et l'Internet existant. Bien que LISP ait été évalué en termes d'extensibilité [107], de stabilité [81], d'évolution LISP [77] et de délai de résolution des liaisons entre EID et RLOC ([107], [44]), LISP n'a jamais été analysé de l'interfonctionnement avec l'Internet hérité à grande échelle. Cette dissertation comble cette lacune en fournissant une évaluation de latence et une mesure de chemin de routage sur le mécanisme d'interfonctionnement LISP.

0.7.1 Méthodologie de mesure

Le travail est basé sur deux expériences (l'une est réalisée en 2015, l'autre en 2016) en utilisant RIPE Atlas, qui est la plus grande infrastructure de mesure Internet existante pour IPv4 et IPv6. Nous avons déployé une sonde (numéro de sonde RIPE Atlas #22341) avec adresse IPv4 et IPv6 sur la plateforme LISP-Lab d'un institut universitaire à Paris, en France, pour effectuer les mesures actives LISP. Pour IPv4, il utilise à la fois PETR et PITR de LISP-Lab pour communiquer avec l'Internet hérité, et la connexion entre les xTR de sonde et PxTR se fait via un VPN MPLS. Toutefois, le tunnel MPLS ne prenait pas en charge IPv6 à ce moment-là. Ainsi, nous configurons l'ITR de la sonde LISP-Lab en mode natif pour acheminer les paquets vers le cœur Internet pour les cibles IPv6 (c'est-à-dire que PETR n'est pas utilisé pour IPv6). Par conséquent, les paquets IPv6 sortants de cette sonde ne sont pas encapsulés dans des paquets LISP et transmis de manière traditionnelle, mais les paquets renvoyés passent toujours par la plate-forme PITR de LISP-Lab. Nous sélectionnons également d'autres sondes sur RIPE Atlas (LIP6, mPlane, rmd / Gandi) comme éléments de comparaison, une ancre appelée FranceIX est une référence. Dans Dateset 2015, les 4 son-

des sélectionnées ping aux 50 meilleurs sites Alexa toutes les 10 minutes pendant 6 heures. Dans Dateset 2016, nous menons les 4 campagnes d'expérimentation suivantes:

- 5 sondes ping à 500 destinations pendant 15 jours avec un intervalle de 30 minutes (IPv4).
- 5 sondes ping vers 122 destinations pendant 15 jours avec un intervalle de 30 minutes (IPv6).
- 5 sondes traceroute à 500 destinations pendant 15 jours avec un intervalle de 60 minutes (IPv4).
- 5 sondes traceroute à 122 destinations pendant 15 jours avec un intervalle de 60 minutes (IPv6).

0.7.2 Résultats des Ping

Dans nos deux campagnes de mesure à grande échelle, les résultats expérimentaux présentent les résultats cohérents. Nous prenons les résultats du Dataset 2016 en exemple. La figure 6.6 montre le CDF du RTT moyen vers les destinations IPv4 sélectionnées pour Dataset 2016. En tant que point d'ancrage, la sonde FranceIX surpasse les autres sondes dans la plupart des cas, montrant ainsi le plus petit RTT. Avec la plage RTT [0,200] ms, la latence de deux sondes LISP, LISP-Lab et LIP6, est respectivement inférieure de 25 ms et 60 ms à celle de trois autres sondes non LISP. Une telle dégradation des performances RTT est provoquée par le tronçon de longueur de chemin introduit par la technologie proxy (comme mentionné dans la section 2.1), puisque chaque paquet envoyé entre la sonde LISP et l'Internet hérité doit passer par un PxTR. Bien que les deux sondes LISP introduisent l'étirement, la sonde LIP6 est même 50ms plus lente que la sonde LISP-Lab. Etant donné que les sondes LISP-Lab et LIP6 traversent le même PETR situé à Lyon, la dégradation des performances de la sonde LIP6 est due à la sélection PITR pour le trafic d'un réseau Internet vers un réseau LISP. Les PITR utilisés par la sonde LIP6 appartiennent au réseau bêta LISP et il y a au total 6 PITR disponibles. Les PITR annoncent les préfixes Beta LISP à leurs homologues AS en utilisant BGP. Ainsi, différentes destinations sélectionnent différents PITR en fonction de l'emplacement des destinations et conduisent à la réponse par les différents chemins. Par exemple, une destination asiatique utilise normalement le PITR au Japon au lieu d'en choisir un en Europe. Comme indiqué dans Tab. 6.2, LISP Beta Network n'a pas déployé PITR en France, alors que le PITR de LISP-Lab est dans le même pays que sa sonde. Ainsi, le chemin de retour que nous avons mesuré pour la sonde LIP6 est plus long que pour la sonde LISP-Lab. En conséquence, plus la destination est proche des sondes, plus la différence de RTT entre les deux sondes LISP est grande.

Dans la gamme RTT $[200, 500]$ ms, toutes les sondes sauf LIP6 montrent fondamentalement le même RTT. Cela indique que LISP-Lab a une meilleure performance dans le scénario de connexion réseau longue distance. La raison en est que le délai d'étirement de la sonde LISP-Lab à PxTR (à Lyon) peut être ignoré. La Fig. 6.6 confirme cette situation, dans la plage RTT $[200, 500]$ ms, la différence entre deux sondes devient significativement inférieure à celle de l'intervalle $[0, 200]$ ms.

Les résultats et l'analyse susmentionnés coïncident avec la figure 6.3 de l'ensemble de données de 2015, qui couvre uniquement la comparaison liée à la sonde LISP-Lab. Ceci indique que la performance de LISP PxTR est stable: les résultats de mesure ne changent pas avec les différentes destinations à un moment différent.

De l'expérience, les résultats montrent également que la position de PxTR est très importante. Le PxTR soit proche des sources ou des destinations peut beaucoup diminuer la latence. D'une manière générale, LISP est stable par rapport à l'ancre de référence FranceIX, à l'exception de la performance IPv6 de LISP Beta Network. En outre, les performances de LISP-Lab PxTR sont plus fiables que celles de LISP Beta Network, bien que ce dernier dispose de 6 Px-TR mondiaux utilisés pour IPv4 et 2 situés aux États-Unis pour IPv6, alors que LISP-Lab n'a que 1 PxTR pour IPv4 et IPv6. Comparé à l'utilisation de PxTR, le transfert natif sans utiliser LISP diminue la latence, mais pas beaucoup.

0.7.3 Résultats des Traceroute

Lorsque les cibles IPv4 sont en Europe et en Amérique, il y a très peu de cas où le nombre de hops de LISP-Lab ou de LIP6 est le plus petit. Précisément, il y a 7,65% de cas où LISP-Lab a le chemin le plus court par numéro de houblon vers les destinations asiatiques et 6,63% de cas que LIP6 a le chemin le plus court dans l'expérience. Ce résultat coïncide avec le pourcentage du RTT le plus petit représenté sur la figure 6.7. Cette relation cohérente entre le plus petit RTT et le plus court chemin existe également pour l'expérience IPv6, où il n'y a aucune destination pour laquelle le nombre de bonds ou le RTT sont les plus faibles (pour les sondes LISP-Lab et LIP6). Pour IPv6, c'est FranceIX ou Gandi qui ont toujours le chemin le plus court vers toutes les cibles. Pour la sonde LISP-Lab, le nombre de sauts relatif le plus commun est de 12 avec un pourcentage de 21,77%, ce qui reste limité dans la plupart des cas. De l'expérience liée au ping dans Sec. 6.4 et Sec. 6.5 nous savons que l'utilisation de PxTR introduit des frais généraux. Donc, le nombre de sauts relatifs étant généralement positif est raisonnable. Le très grand nombre de sauts relatifs apparaît juste une fois pour 19 et 25, donc nous les considérons comme des valeurs aberrantes. Similaire au cas LISP-Lab, pour la sonde LIP6, le nombre de sauts relatif le plus fréquent est de 13 dans 19,53% des cas.

0.8 Analyse de la mobilité LISP et de la mise en œuvre de ns-3

Le LISP, grâce à son approche de cartographie et d'encapsulation, peut apporter des avantages à la mobilité. Le nœud LISP Mobile (LISP-MN) est basé sur la fonctionnalité LISP de base pour assurer la mobilité à travers les réseaux. Ainsi, LISP peut être implémenté soit sur les routeurs de bordure, soit directement sur les hôtes finaux pour gérer la mobilité. Cependant, il n'y a pas de résultats expérimentaux comparant les avantages et les inconvénients de chaque solution. L'évaluation de la performance de mobilité LISP nécessite les bancs d'essai ou les simulateurs. L'architecture LISP de base est déployée sur LISP Beta Network et sur la plateforme LISP-Lab pour offrir aux chercheurs un environnement expérimental réaliste, mais les deux ne supportent pas LISP-MN. De plus, comme les simulateurs peuvent aider les chercheurs à vérifier rapidement les propositions et à tester de nouvelles fonctionnalités, nous avons implémenté des extensions de mobilité dans un simulateur LISP. Ce chapitre présente la mise en œuvre des extensions de mobilité LISP sous ns-3.27 en s'appuyant sur un simulateur LISP de base open source. En outre, ce chapitre analyse les différents scénarios de mobilité LISP à partir des respects du délai de transfert et de l'overhead du plan de contrôle LISP. Il décrit les caractéristiques de chaque scénario.

0.8.1 Implémentation de LISP sur ns-3

Notre implémentation (voir Fig. 7.1) respecte les normes de mobilité LISP RFC 6830 [47] et LISP [86]. Sans créer un nouveau module ns-3, nous implémentons les fonctionnalités de LISP / MN en modifiant et en étendant deux modules déjà existants de ns-3: *internet* et *internet-apps*. Inspiré par OpenLISP [109], notre implémentation se compose de deux parties principales: le plan de données LISP et le plan de contrôle LISP. La communication entre les données LISP et le plan de contrôle est réalisée via une socket dédiée (c'est-à-dire *LispMappingSocket*) qui hérite de la classe ns-3 *Socket*.

L'implémentation de LISP Data Plane est principalement composée de classes *LispOverIp* et *MapTable* et de leurs sous-classes, ainsi que de certaines classes auxiliaires (e.x. *LispHeader*). En outre, pour prendre en charge les fonctionnalités LISP, les procédures de transmission, de réception, d'acheminement et de distribution des paquets de *Ipv4L3Protcol* sont adaptées en conséquence.

La base de données et le cache LISP sont tous deux modélisés par la même classe *MapTable* qui stocke et gère les informations de mappage EID-RLOC. Cette classe fournit des opérations CRUD (Create, Retrieve, Update, Delete) pour les mappages. Chaque entrée de

mappage dans la base de données et le cache LISP est une instance de *MapEntry*. Par souci de flexibilité, la classe *MapTable* est une classe de base abstraite. Les méthodes CRUD sont implémentées dans sa sous-classe *BasicMapTable*. L'opération de recherche de mappage est une itération directe sur la base de données ou le cache LISP. Il est possible que d'autres utilisateurs fournissent leur propre base de données LISP et implémentent le cache en étendant la classe *MapTable*.

Pour intégrer LISP/MN dans une pile de protocole Internet conventionnelle, une difficulté technique clé est que *Ipv4L3Protcol* devrait être capable de déterminer quand passer un paquet en cours de traitement à une procédure LISP et comment récupérer les informations de mappage associées. A cette fin, une nouvelle classe appelée *LispOverIp* et ses classes étendues sont ajoutées au module ns-3 *internet*. Cette classe est chargée de vérifier s'il est nécessaire de faire des opérations liées à LISP (*NeedEncapsulation ()*, *NeedDecapsulation ()*) et d'encapsuler des paquets IP conventionnels (ie, *LispOutput ()*) ainsi que la décapsulation des paquets LISP (*LispInput ()*). Il contient un pointeur intelligent pointant vers la base de données LISP et le cache LISP (par exemple *MapTable*) sur lequel s'exécute la recherche de mappage.

La mise en œuvre de LISP Control Plane couvre xTR, MR et MS. xTR est inclus dans la classe *LispEtrItrApplication*. Les fonctionnalités de MR et MS sont respectivement implémentées par les classes *MapResolver* et *MapServer*. Les messages du plan de contrôle LISP sont représentés par les classes dérivées de *LispControlMsg*. De plus, pour communiquer avec LISP Data Plane, une classe de socket *LispMappingSocket* est proposée.

Un nœud ns-3 qui exécute *LispEtrItrApplication* est un routeur compatible LISP. Il devrait être capable de communiquer avec *LispOverIp* sur le même nœud (par exemple informer l'événement manquant du cache) et d'autres routeurs compatibles LISP (par exemple Map-Request / Map-Reply). Lorsque la destination RLOC n'est pas trouvée dans le cache de xTR pour un paquet traité, l'événement d'absence de cache se produit et le plan de données LISP (par exemple *LispOverIp*) notifie *LispEtrItrApplication* sur le même noeud de cet événement. Une fois la réception de l'événement de cache manquant dans le plan de données LISP (c'est-à-dire l'objet *LispOverIp*), *LispEtrItrApplication* lance un message de demande de carte vers le système de mappage LISP. Une fois la réception de Map-Reply, le mappage EID à RLOC reçu est inséré dans le cache LISP. En cas de réception de Map-Request, *LispEtrItrApplication* exécute une consultation de base de données sur *MapTable* et génère le mappage de correspondance EID-RLOC correspondant. Lorsque l'application xTR démarre ou que la base de données LISP sur un nœud est mise à jour, l'application xTR envoie un message Map-Register à MS et attend un message Map-Notify. Dans le cas d'une

mise à jour de la base de données LISP, xTR envoie un message SMR (Solicit-Map-Request) à tous les xTR dont RLOC est présent dans son cache.

La classe *MapServer* fournit des fonctionnalités MS. Il gère une base de données LISP gérant les informations de mappage EID vers RLOC. Une fois la réception du message Map-Register, il met à jour la base de données LISP et envoie un message Map-Notify en réponse si nécessaire. Lors de la réception d'un message Map-Request, si le mappage EID-RLOC interrogé est trouvé dans sa base de données, MS transfère cette requête au xTR correspondant sinon envoie un message Negative Map-Reply à l'interrogation xTR. Dans la mise en œuvre actuelle, le rôle de MR est de recevoir le message de demande de carte de xTR et de le transmettre à la MS.

Nous avons adapté l'application client DHCP ns-3 pour prendre en charge LISP / MN. Le client DHCP modifié est capable de communiquer avec *LispEtrItrApplication* exécuté sur le même noeud. Par exemple, après l'attribution d'adresse IPv4, le client DHCP vérifie si le LRLOC est modifié. Si LRLOC est modifié, le client DHCP notifie le *LispEtrItrApplication* en envoyant un message dédié qui contient le mappage EID-LRLOC. *LispEtrItrApplication* est chargé de remplir l'entrée de mappage reçue dans la base de données LISP et d'envoyer un message Map-Register à MS.

En tant que LISP-MN, il possède un EID permanent statique et un RLOC dynamique attribués par le serveur DHCP. Nous utilisons la solution basée sur le dispositif TUN. Dans notre implémentation, au moins deux NIC doivent être installées dans le MN. L'un est un NIC normal tel que *WifiNetDevice*. L'application client DHCP fonctionne sur ce type de carte et donc le LRLOC est affecté à cette carte. L'autre est une carte de type TUN. Le NIC TUN est une carte virtuelle qui devrait réellement appeler *Send ()* d'une autre carte réseau réelle. L'EID permanent est affecté au périphérique TUN.

Rappelez-vous qu'après la procédure DHCP, le nœud sera configuré une passerelle par défaut fournie par le serveur DHCP. La table de routage de LISP-MN est modifiée de sorte que les paquets de la couche application utilisent toujours l'EID comme adresse IP source de l'en-tête IP interne.

0.8.2 Analyse théorique

La mobilité IP sur LISP peut être implémentée soit sur le routeur frontal, soit sur l'hôte final. Pour explorer les caractéristiques de chaque schéma, nous proposons trois scénarios différents avec une topologie spécifique comme exemple et analysons le délai global de transfert et le surcoût du plan de contrôle LISP.

1. LISP-MN dans le site non-LISP (c'est-à-dire, seul l'hôte final prend en charge LISP).
2. MN dans le site LISP (c'est-à-dire, seul le routeur frontal supporte LISP).

3. LISP-MN dans le site LISP (c'est-à-dire que le routeur de frontière et l'hôte final supportent LISP).

Dans le premier scénario, le délai de transfert est de 217 ms et le temps système de transfert est de 6 messages du plan de contrôle LISP. Les avantages de ce scénario, c'est-à-dire, LISP-MN dans un site non-LISP, sont: 1) il est capable de réaliser le transfert à travers différents sous-réseaux; 2) l'analyse numérique indique que le délai global de transfert est faible; 3) donc à l'overhead global. Par rapport aux deux autres scénarios, qui seront décrits en détails dans les sections suivantes, la mobilité de ce scénario ne provoque pas beaucoup de trafic dans le plan de contrôle LISP. Cependant, puisque les routeurs sont toujours les routeurs normaux dans ce scénario, cela ne peut pas aider à réduire la taille de la table de routage BGP, ce qui est le but initial pour motiver LISP. De plus, chaque LISP-MN a besoin d'une adresse IP permanente comme EID, ce qui augmente le fardeau de l'allocation d'adresses IPv4. Chaque EID permanent et son LRLOC doivent être enregistrés dans le système de mappage, ce qui augmente également la taille de la table de mappage LISP.

Dans le second scénario, le délai de transfert est de 212 ms et le temps système de transfert est de 7 messages du plan de contrôle LISP. Etant donné que les routeurs prennent en charge LISP, c'est-à-dire les xTR dans ce scénario, cela peut aider à réduire la taille de la table de routage BGP, ce qui est la motivation initiale de la proposition de LISP. En outre, l'analyse laisse entendre que le délai global de transfert de ce scénario est le plus court. Cependant, il ne peut fournir la mobilité qu'au sein d'un même sous-réseau, ce qui signifie qu'il n'est pas en mesure d'offrir le transfert via différents sous-réseaux. Ainsi, ce scénario est plus approprié pour être appliqué à la mobilité des machines virtuelles dans le centre de données.

Dans le troisième scénario, le délai de transfert est de 1018 ms et le temps de transfert de transfert est de 9 messages du plan de contrôle LISP. Etant donné que les routeurs de MN et de frontière supportent LISP, les avantages de ce scénario, à savoir, LISP-MN dans le site LISP, sont: 1) cela peut aider à réduire la taille de la table de routage BGP; 2) il est capable de réaliser le handover à travers différents sous-réseaux. Cependant, de même que les inconvénients du premier scénario, chaque LISP-MN a besoin d'une adresse IP permanente comme EID, ce qui augmente le fardeau de l'allocation d'adresses IPv4. Chaque EID permanent et son LRLOC doivent être enregistrés dans le système de mappage, ce qui augmente également la taille de la table de mappage LISP. En outre, l'analyse numérique indique que le délai global de transfert est beaucoup plus long que les deux autres scénarios en raison de sa double encapsulation.

0.9 Conclusion et travail futur

0.9.1 Contributions majeures

L'objectif de cette thèse est d'évaluer de manière approfondie le *Locator/Identifier Separation Protocol (LISP)* de différents aspects: mesurer ses performances en réalisant des expériences dans les réseaux réels à grande échelle, analyser numériquement les limites, mettre en œuvre la mobilité LISP en ns-3. Les travaux de mesure sont liés à l'évaluation du système de cartographie et à l'évaluation des performances d'interfonctionnement avec l'Internet existant. En raison des nouvelles découvertes dans le système de cartographie, nous proposons un outil de surveillance LISP plus complet pour répondre aux exigences supplémentaires. L'analyse théorique porte sur la mobilité LISP. Plus précisément, nous analysons les limites basses de transfert pour les différents composants réseau supportant LISP. En outre, nous implémentons les extensions de mobilité LISP sur ns-3.27 en exploitant une implémentation existante. Concrètement, les principales contributions peuvent être résumées comme suit.

Évaluation du LISP mapping system

Nous avons continuellement mesuré le réseau bêta de LISP pendant dix-sept jours pour évaluer la stabilité et la cohérence du système de cartographie. Les expériences sont menées du 2 au 18 juillet 2013, à partir de 5 points de vue (PV), envoyant des messages de *Map-Request* à tous les 13 résonateurs de carte (LIS) du réseau bêta LISP, pour toutes les adresses IP de destination sélectionnées. 30 minutes. Les VP étaient composés d'une partie des réseaux académiques, d'Internet commercial, de PlanetLab et répartis en Europe et aux Etats-Unis. Les adresses IP de destination comprenaient la première adresse dans chaque préfixe dérivé du projet LISPmon [18] quel que soit le type (c'est-à-dire les préfixes EID ou les préfixes réguliers). Les résultats des mesures montrent que le système de cartographie est stable pendant 91,49% du temps, est cohérent respectivement par MR dans 86,3% des cas et par VP dans 90,48% des cas. Comme les instabilités et les incohérences sont observées et pour les étudier dans les détails, nous avons développé une nouvelle taxonomie. Il montre que Map-Reply ne change qu'une seule fois dans les cas les plus instables, et l'incohérence est principalement causée par les différents types de Map-Replies reçues (le mélange de Map-Reply négatif et LISP Map-Reply). Dans l'ensemble, les instabilités et les incohérences sont des événements rares. Enfin, l'utilisation de multi-homing et de dual-stack (c'est-à-dire IPv4 et IPv6) est observée pendant toute l'expérience. Ce travail est publié dans WNM [81].

LISP-Views: moniteur de LISP mapping system

L'évaluation du système de cartographie montre qu'il existe une instabilité et une incohérence, qui sont difficiles à trouver pour le seul moniteur LISPmon actuel LISPmon. Un système de suivi dynamique et complet est nécessaire pour accompagner la croissance du LISP. Nous proposons LISP-Views, une architecture de surveillance LISP à grande échelle polyvalente. LISP-Views permet de réaliser automatiquement des mesures complètes et objectives. Les utilisateurs peuvent le déployer sur n'importe quel VP et définir pour surveiller quels MR. Ils sont également capables de définir le type de mesures. Toutes ces opérations peuvent être effectuées via l'API fournie. Au moment de la rédaction, LISP-Views est déployé sur un VP pour superviser tous les MR existants sur le réseau LISP-Beta et sur la plateforme LISP-Lab. Après avoir exécuté LISP-Views dans la nature pendant plusieurs mois et comparé les résultats de surveillance avec LISPmon, nous confirmons que LISP-Views fournit des informations plus détaillées et plus précises. Nous observons que chaque MR a des comportements différents par rapport aux autres, comme dans le temps de fiabilité et de résolution. Par ailleurs, les mesures ont aidé les codeurs OpenLISP à corriger un bug précédemment inconnu dans la plateforme MR de LISP-Lab. Les résultats graphiques préliminaires, tels que le nombre de LISP Map-Reply, le nombre de RLOC contenus dans chaque LISP Map-Reply, et la réactivité de chaque MR sont affichés sur le site web de LISP-Views [15]. Ce travail est publié dans ITC [77].

Évaluation de la performance d'interfonctionnement LISP grâce à RIPE Atlas

Nous avons mené deux expériences pour évaluer les performances d'interfonctionnement du LISP avec l'Internet hérité. C'est le premier aperçu complet de la performance de LISP PITR / PETR (PxTR). Comme phase d'installation, nous avons sélectionné 4 sondages ping sur les 50 meilleurs sites Alexa toutes les 10 minutes pendant 6 heures le 6 novembre 2015. Dans la seconde expérience, nous avons choisi 5 sondes ping et traceroute pour atteindre respectivement les 500 adresses IPv4 Alexa et 122 IPv6 30 minutes et 60 minutes. Les deux expériences montrent que le PxTR introduit en effet des effets négatifs pour les destinations proches, mais l'impact négatif de PxTR peut être ignoré pour la transmission intercontinentale longue distance. Les résultats montrent également que la position de PxTR est très importante. Le PxTR positionné soit près des sources, soit les destinations peuvent beaucoup diminuer la latence. LISP est généralement stable, à l'exception des performances IPv6 de LISP Beta Network. En outre, les performances de LISP-Lab PxTR sont plus fiables que celles de LISP Beta Network, bien qu'au moment de la rédaction de ce dernier, 6 PxTRs mondiaux soient utilisés pour IPv4 et 2 pour IPv6 aux Etats-Unis, alors que LISP-Lab n'a

que 1 PxTR pour IPv4 et IPv6. Transférer nativement sans utiliser LISP diminue la latence, mais pas beaucoup. L'expérience traceroute de LISP-Lab montre que l'introduction de PxTR apporte bien sûr plus de sauts, mais si le PxTR est bien configuré, pour avoir toujours des pairs vers les destinations, il n'y a plus que 4 sauts par rapport aux paquets encapsulation avec LISP. Ce travail est publié dans l'atelier étudiant INFOCOM [79], l'atelier étudiant ACM SIGCOMM [78] et Elsevier Computer Networks [80].

Analyse de la mobilité LISP et de la mise en œuvre de ns-3

Il existe trois méthodes de prise en charge de la mobilité LISP: l'hôte (c'est-à-dire le LISP-MN), le réseau (c'est-à-dire ITR / ETR (xTR)), à la fois hôte et réseau (c'est-à-dire LISP-MN derrière xTR) mobilité. Dans le premier scénario, il n'y a que MN supportant LISP (LISP-MN dans le site non-LISP); dans le second scénario, il n'y a qu'un routeur de frontière supportant LISP (MN dans le site LISP); et dans le troisième scénario, les routeurs MN et frontière supportant LISP (LISP-MN dans le site LISP). Nous analysons le retard global de transfert et le surcoût du plan de contrôle LISP parmi eux. Le deuxième scénario a le plus petit délai de transfert, mais le MN peut seulement se déplacer dans le même sous-réseau. Les premier et troisième scénarios prennent en charge l'itinérance transparente MN à travers les différents sous-réseaux, mais le premier scénario n'aide pas à réduire la table de routage BGP, et le troisième scénario introduit un délai de transfert beaucoup plus long. Basé sur une implémentation de LISP de base sur ns-3.24, nous l'adaptons d'abord à ns-3.27 (la dernière version au moment de l'écriture). Afin de permettre aux chercheurs de suivre en profondeur l'échange de paquets LISP, nous encodons chaque bit dans les paquets LISP Data Plane afin que Wireshark puisse les résoudre. Enfin, nous implémentons les extensions de mobilité LISP.

0.9.2 Discussion & travaux futurs

Évaluation du LISP mapping system

Il y a deux aspects à faire à l'avenir. L'une est de mener une nouvelle expérience pour voir si quelque chose dans le système de cartographie a changé. L'autre, qui n'a jamais été fait, utilise LISP-DDT Referral Internet Groper (RIG) pour avoir un regard plus profond sur l'arbre LISP-Delegated Database Tree (LISP-DDT).

L'expérience a été menée en 2013, alors que le LISP connaît une croissance dans les années suivantes. L'architecture du système de cartographie a également changé, dans laquelle il y avait 13 MRs pendant la mesure mais seulement 7 MR en 2017. En outre, le réseau LISP Beta est le seul banc d'essai LISP à cette époque. Cependant, la plate-forme

LISP-Lab est ouverte aux expérimentateurs en 2015. Cette dernière se compose de 3 MR et d'interconnexions avec le système de cartographie de LISP Beta Network. Ainsi, il est très nécessaire de mener une nouvelle expérience à l'avenir pour évaluer les dernières performances de stabilité et de cohérence du système de cartographie LISP. Les résultats peuvent être comparés aux résultats du chapitre. 4 pour que le LISP se développe.

Les mapping informations doivent être identiques sur tous les MR en même temps, mais les résultats expérimentaux dans le chapitre. 4 montrent qu'il existe des incohérences entre les MR et les VP. La raison est difficile à explorer par les expériences menées auparavant, qui sont basées sur LIG. Puisque les procédures de transmission des demandes de carte dans le système de cartographie sont transparentes pour les expérimentateurs. Basé sur la théorie, il est probablement causé par le cache de chaque MR stockant les informations de mappage qui ne sont pas synchronisées. Ainsi, tirer parti de RIG pour résoudre les problèmes et d'abord regarder les problèmes incohérents dans le système de cartographie est très utile. Si l'hypothèse est correcte, proposer un mécanisme synchronisé devrait être important pour le système de cartographie.

LISP-Views: moniteur de LISP mapping system

LISP-Views est toujours dans la première phase, il est déployé sur un seul VP et exécute les mesures en utilisant les lignes de commande. Ainsi, il est prévu de déployer les LISP-Views sur plusieurs VP pour reproduire la cohérence des PV comme décrit dans le chapitre. 4. Par ailleurs, les lignes de commande ne sont pas efficaces pour effectuer les expériences, pour implémenter une API REST sur LISP-Views pour paramétrer le temps d'expérience, les MR surveillés, et obtenir les différents aspects du statut LISP à l'avenir. Au moment d'écrire cette dissertation, toutes les adresses IP surveillées via LISP-Views sont IPv4, il devrait être intéressant de tester les comportements IPv6.

Évaluation de la performance d'interfonctionnement LISP grâce à RIPE Atlas

Dans notre expérience, comme la sonde LIP6 ne reçoit pas de réponses ping pour 42 destinations IPv4 et 10 IPv6, pour regrouper ces destinations par géographie et explorer les raisons qui peuvent être des travaux futurs possibles. Comme les paquets IPv6 envoyés par la sonde LISP-Lab sont transmis nativement au lieu d'être encapsulés dans des paquets LISP pendant notre expérience, les configurations doivent être modifiées pour que PETR de LISP-Lab puisse recevoir le trafic IPv6 de la sonde LISP-Lab. Ensuite, les performances de LISP IPv6 peuvent être évaluées. En outre, comme la dégradation des performances observée au chapitre 6 est principalement due à la sélection PITR, il sera utile d'explorer l'interaction en-

tre BGP anycast et LISP interfonctionnement. Si les destinations tirent parti de anycast pour sélectionner le PITR le plus proche, nous supposons que le délai supplémentaire introduit par LISP sera beaucoup diminué.

Analyse de la mobilité LISP et de la mise en œuvre de ns-3

Pour la prochaine étape, nous prévoyons de valider l'analyse numérique sur le simulateur. Comme *Map-Versioning* [67] est un autre mécanisme de mise à jour du cache de mappage, nous pouvons également comparer les performances entre celui-ci et *SMR* que nous présentons dans ce chapitre. 7 par notre simulateur mis en œuvre.

Il y a des directions pour le travail supplémentaire. Le deuxième scénario a le délai de transfert le plus court, mais il ne prend pas en charge l'itinérance MN à travers les différents sous-réseaux. Ainsi, utiliser l'adresse MAC comme EID peut être une option pour que le MN puisse changer de sous-réseau sans interruption de la communication. En outre, puisque le troisième scénario utilise une double encapsulation, ce qui entraîne une forte augmentation du retard, il sera intéressant que le système de mappage puisse traiter différemment la demande de carte SMR et la demande de mappage normale. Lorsque le système de mappage reçoit une demande de carte appelée SMR, s'il trouve que le *RLOC* de l'EID demandé est réellement un *LRLOC*, par exemple, le *LRLOC* de LISP-MN dans notre troisième scénario, il transmet ensuite la demande de carte non seulement à le LISP-MN, mais aussi le xTR2, de sorte que LISP-MN répond au xTR3 avec les informations de mappage de $<EID_{LISP-MN}, LRLOC_{LISP-MN}>$ et xTR_2 répond à xTR_3 avec les informations de mappage de $<LRLOC_{LISP-MN}, RLOC_{xTR_2}>$. De cette façon, xTR_3 n'a besoin qu'une seule fois d'interroger le système de mappage, de sorte que le délai de transfert dû à la procédure liée au LISP peut réduire beaucoup. Le futur travail de mise en œuvre des extensions de mobilité LISP est destiné à supporter IPv6.

Table of contents

0.1 Introduction ix
0.2 Motivation xi
0.3 Contexte xiii
0.3.1 Architecture LISP xiii
0.3.2 Interfonctionnement avec Internet de la légitimité xiii
0.3.3 LISP mobilité xiv
0.3.4 LISP Beta Network xv
0.3.5 LISP-Lab platform xv
0.3.6 LISPmon xvi
0.4 Évaluation de LISP mapping system xvi
0.4.1 Vue générale des mesures xvi
0.4.2 Metrics xvii
0.4.3 Résultats xviii
0.5 LISP-Views: moniteur du LISP mapping system xx
0.5.1 Motivation xx
0.5.2 Déscription de LISP-Views xxi
0.6 Validation de LISP-Views xxii
0.7 Évaluation de la performance d'interfonctionnement LISP grâce à RIPE Atlas xxiv
0.7.1 Méthodologie de mesure xxiv
0.7.2 Résultats des Ping xxv
0.7.3 Résultats des Traceroute xxvi
0.8 Analyse de la mobilité LISP et de la mise en œuvre de ns-3 xxvii
0.8.1 Implémentation de LISP sur ns-3 xxvii
0.8.2 Analyse théorique xxix
0.9 Conclusion et travail futur xxxi
0.9.1 Contributions majeures xxxi
0.9.2 Discussion & travaux futurs xxxiii

List of figures **xliii**

List of tables **xlvii**

Acronyms **xlix**

1 Introduction **1**
1.1 Motivations . . . 1
1.2 Contributions . . . 3
1.3 Manuscript organization . . . 5

2 LISP overview **7**
2.1 LISP architecture . . . 7
2.1.1 LISP Data Plane . . . 8
2.1.2 LISP Control Plane . . . 9
2.1.3 Communication between two LISP-sites . . . 12
2.2 Interworking with legency Internet . . . 14
2.3 Mapping cache update mechanisms . . . 15
2.3.1 Solicit-Map-Request (SMR) . . . 15
2.3.2 Map-Versioning . . . 16
2.4 LISP mobility . . . 17
2.4.1 LISP Mobile Node (LISP-MN) . . . 18
2.4.2 MN mobility in LISP-Site . . . 20
2.4.3 LISP-MN mobility in LISP-Site . . . 20
2.4.4 LISP-MN mobility in non-LISP-Site . . . 21
2.4.5 LISP-MN mobility between LISP-Site & non-LISP-Site . . . 22
2.5 LISP alternative use-cases . . . 23
2.5.1 LISP traffic engineering . . . 23
2.5.2 LISP transition between IPv4 and IPv6 . . . 23
2.5.3 LISP SDN . . . 24

3 Related work **27**
3.1 LISP platforms . . . 27
3.1.1 LISP Beta Network . . . 27
3.1.2 LISP-Lab platform . . . 29
3.2 LISP monitoring tools . . . 29
3.2.1 LISPmon . . . 29
3.2.2 LIG . . . 30

3.2.3 RIG . . . 30
3.2.4 Monitoring tools shortcomings . . . 31
3.3 LISP implementations . . . 31
3.3.1 LISP implementations . . . 32
3.3.2 The LISP simulation gap . . . 34
3.4 LISP mapping system evaluation . . . 34
3.4.1 Current studies . . . 34
3.4.2 The incomplete puzzle of MDS evaluation . . . 34
3.5 LISP network evolution . . . 35
3.5.1 Updating LISP measurements . . . 35
3.6 LISP interworking with legacy Internet . . . 36
3.6.1 PxTRs evaluation: an incomplete picture . . . 37
3.7 LISP mobility . . . 37
3.7.1 Open issue in LISP mobility . . . 38
3.8 Troubleshooting issues . . . 38
3.9 Other LISP related studies . . . 40
3.9.1 LISP cache performance . . . 40
3.9.2 LISP security . . . 40
3.9.3 The Maximum Transmission Unit (MTU) issue . . . 41

4 Evaluation of LISP mapping system 43
4.1 Methodology . . . 44
4.1.1 Measurements overview . . . 44
4.1.2 Metrics . . . 45
4.2 Mapping system stability evaluation . . . 46
4.3 Mapping system consistency evaluation . . . 50
4.3.1 Consistency evaluation by MR . . . 51
4.3.2 Consistency evaluation by VP . . . 54
4.4 Summary . . . 56

5 LISP-Views: LISP mapping system monitor 59
5.1 Proposed monitoring architecture . . . 60
5.1.1 Motivation . . . 60
5.1.2 Description of LISP-Views . . . 61
5.2 LISP-Views validation . . . 63
5.2.1 Methodology . . . 63
5.2.2 LISP-Views vs. LISPmon . . . 63

5.3 Dissecting LISP with LISP-Views 66
5.4 Summary 71

6 Assessing LISP interworking performance through RIPE Atlas 73
6.1 Experiment resources 74
6.1.1 RIPE Atlas 74
6.1.2 Alexa 75
6.2 Measurement methodology 75
6.2.1 Dataset 2015 75
6.2.2 Dataset 2016 76
6.3 IPv4 Ping results from Dataset 2015 79
6.4 IPv4 Ping results from Dataset 2016 82
6.5 IPv6 Ping results 87
6.6 Traceroute-related results 91
6.7 Summary 94

7 Analysis of LISP mobility and ns-3 implementation 97
7.1 NS-3 98
7.2 Basic LISP implementation on ns-3 98
7.3 LISP mobility extensions on ns-3 98
7.3.1 Implementation of LISP Data Plane 99
7.3.2 Implementation of LISP Control Plane 101
7.3.3 Modification of DHCP client to support LISP mobility 103
7.3.4 Integration of TUN net interface card 105
7.4 Theoretical analysis 105
7.4.1 LISP-MN in non-LISP-Site 106
7.4.2 MN in LISP-Site 109
7.4.3 LISP-MN in LISP-Site 111
7.5 Summary 115

8 Conclusion and future work 117
8.1 Major contributions 117
8.1.1 Evaluation of LISP mapping system 117
8.1.2 LISP-Views: LISP mapping system monitor 118
8.1.3 Assessing LISP interworking performance through RIPE Atlas . . . 118
8.1.4 Analysis of LISP mobility and ns-3 Implementation 119
8.2 Discussion & future work 120

8.2.1 Evaluation of LISP mapping system 120
8.2.2 LISP-Views: LISP mapping system monitor 120
8.2.3 Assessing LISP interworking performance through RIPE Atlas . . . 121
8.2.4 Analysis of LISP mobility and ns-3 Implementation 121

References **123**

Appendix A Source Code **133**
A.1 LISP Stability and Consistency analyzer 133
A.2 LISP-Views 133
A.3 LISP measurements on Atlas 133
A.4 LISP implementation on ns-3 134

List of figures

1.1 Growth of the BGP Table - 1994 to Present [5] 2
1.2 Structure of this dissertation 5

2.1 Hierarchical taxonomy of mapping systems (from [65]) 9
2.2 Illustration of the LISP+ALT topology 10
2.3 Illustration of the LISP-DDT topology 11
2.4 LISP architecture overview 13
2.5 LISP architecture with the packets forwarding sequence 14
2.6 Packet sequence of LISP SMR mechanism 16
2.7 Packet sequence of LISP Map-Versioning mechanism 17
2.8 Illustration of LISP-MN in LISP-Site and non-LISP-Site 19
2.9 Packet flow of mobility when LISP-MN in LISP-Site 20
2.10 Packet flow of mobility when LISP-MN in non-LISP-Site 21
2.11 Packet flow of LISP-MN mobility between LISP-Site and non-LISP-Site . . 22
2.12 Illustration of LISP transmission between IPv4 and IPv6 24
2.13 LISP in OpenDaylight (source: [13]) 25

3.1 International LISP Beta Network MDS elements & Proxy xTRs – October 2017 [12] 28
3.2 Illustration of LISP troubleshooting 39

4.1 Overall percentage of stability observed at five different VPs. 46
4.2 CDF of instability frequency. X-axis is instability frequency, indicating the ratio of instability occurrence number for one <EID, MR, VP> tuple over total experiment number. 47
4.3 CDF of stability duration (in number of experimental rounds). X-axis is stability duration, indicating during how many experiment rounds that the mapping for one <EID, MR, VP> tuple is stable. 48
4.4 CDF of the dissimilarity among different types of instabilities. 50

4.5 Overall consistency by MR at 5 different Vantage Points. 52
4.6 Cumulative Distribution Function of inconsistency frequency by MR. X-axis is inconsistency frequency by VP, indicating the ratio of inconsistency occurrence for each <$EID, *, VP$> tuple over total experiment number. . . . 53
4.7 Overall consistency by VP from the different Map-Resolvers. 54
4.8 Cumulative Distribution Function of inconsistency frequency by VP. X-axis is inconsistency frequency by VP, indicating the ratio of inconsistency occurrence for each <$EID, MR, *$> tuple over total experiment number. 55

5.1 Monitoring Architecture 61
5.2 Comparison between LISPmon and LISP-Views. The upper sub-figure shows the number of LISP Map-Reply from LISPmon generally at 7:00 and LISP-Views exactly at 8:00 over days. The bottom sub-figure indicates the LISP Map-Reply and overall Mapping from LISP-Views over time with an interval of 2 hours. 64
5.3 Comparison between the LISP Map-Reply of LISPmon and LISP-Views over days 65
5.4 Reliability of each MR 66
5.5 Median RTT per MR (In the most time, the Negative Map-Reply and overall are overlapped.) 67
5.6 CDF RTT comparison between LISP and Negative Map-Reply 68
5.7 CDF of the number of RLOCs 69
5.8 EIDs by Quantity of Associated RLOCs [15] 70

6.1 Deployment of probes on RIPE Atlas in 2017 74
6.2 Locations of probes and anchor in 2016 77
6.3 CDF of average RTT between different probes from Dataset 2015 79
6.4 Smallest mean RTT (left) and smallest median RTT (right) grouped by continent from Dataset 2015. 80
6.5 Relative mean (left) and median (right) RTT clustered by different continents from Dataset 2015. 81
6.6 CDF of median RTT between different probes (IPv4) from Dataset 2016 . . 82
6.7 Smallest median RTT grouped by continent (IPv4) from Dataset 2016. . . . 84
6.8 IPv4 Relative median RTT clustered by different continents for LISP-Lab (left) and LIP6 (right) from Dataset 2016 85
6.9 CDF of median RTT between different probes (IPv6) from Dataset 2016 . . 87

6.10 IPv6 Relative median RTT clustered by different continents for LISP-Lab (left) and LIP6 (right) from Dataset 2016 88
6.11 Distribution of IPv4 Relative Hops Number for LISP-Lab (left) and LIP6 (right) from Dataset 2016 91
6.12 IPv4 Relative hops number clustered by different continents for LISP-Lab (left) and LIP6 (right) from Dataset 2016 92
6.13 Distribution of IPv6 Relative Hops Number for LISP-Lab (left) and LIP6 (right) from Dataset 2016 93
6.14 IPv6 Relative hops number clustered by different continents for LISP-Lab (left) and LIP6 (right) from Dataset 2016 94

7.1 UML diagram of LISP and LISP mobility implementation. The solid arrow refers to a composition relation, while the blank one refers to a inheritance relation. 99
7.2 State transition diagram of xTR application 102
7.3 LISP-compatible DHCP client state transition diagram 104
7.4 General scenario for LISP mobility architecture 106
7.5 Schema for LISP-MN mobility in non-LISP-Site (SMR-invoked Map-Request is sent to the mapping system) 107
7.6 Schema for MN mobility in LISP-Site (SMR-invoked Map-Request is sent to the mapping system) 110
7.7 Schema for LISP-MN mobility in LISP-Site (SMR-invoked Map-Request is sent to the mapping system) 113
7.8 Schema for LISP-MN mobility in LISP-Site (SMR-invoked Map-Request is sent to the source of SMR) 114
7.9 The handover delay related to LISP (left) and the handover overhead (right) grouped by three LISP mobility scenarios. 115

List of tables

2.1 Catalogs of mobility 18

4.1 Map-Reply Differences 45
4.2 Percentage of Instabilities by Category 49

5.1 Percentage of Mapping Source 68

6.1 Different configurations of probes in 2015 76
6.2 Location of LISP Beta Network PxTRs in 2016 77
6.3 Configuration of two LISP probes in 2016 78
6.4 Different configurations of probes in 2016 78
6.5 Correlation coefficient to FranceIX from Dataset 2015 82
6.6 Correlation coefficient to FranceIX (IPv4) from Dataset 2016 84
6.7 Reliability of each probe (IPv4) from Dataset 2016 86
6.8 Correlation coefficient to FranceIX (IPv6) from Dataset 2016 90
6.9 Reliability of each probe (IPv6) from Dataset 2016 90

7.1 Symbols for numerical analysis 107

Acronyms

ACK ACKnowledgement.

ARPANET Advanced Research Projects Agency Network.

AS Autonomous System.

BGP Border Gateway Protocol.

CDF Cumulative Distribution Function.

CN Correspondent Node.

DFZ Default Free Zone.

DHCP Dynamic Host Configuration Protocol.

DNS Domain Name System.

EID Endpoint IDentifier.

ETR Egress Tunnel Router.

IETF Internet Engineering Task Force.

ITR Ingress Tunnel Router.

LIG LISP Internet Groper.

LISP Locator/Identifier Separation Protocol.

LISP-SN LISP Stationary Node.

LISP-MN LISP Mobile Node.

LISP-DDT LISP Delegated Database Tree.

LISP+ALT LISP Alternative Logical Topology.

LRLOC Local RLOC.

MDS Mapping Distribution System.

MN Mobile Node.

MPTCP Multipath TCP.

MR Map-Resolver.

MS Map-Server.

NFV Network Functions Virtualization.

NIC Network Interface Card.

OOR Open Overlay Router.

PETR Proxy Egress Tunnel Router.

PITR Proxy Ingress Tunnel Router.

PxTR PITR/PETR.

RIG LISP-DDT Referral Internet Groper.

RLOC Routing LOCator.

RTT Round-Trip Time.

SDN Software Defined Network.

SIP Session Initiation Protocol.

SMR Solicit-Map-Request.

VM Virtual Machine.

VP Vantage Point.

xTR ITR/ETR.

Chapter 1

Introduction

1.1 Motivations

Since the ARPANET (Advanced Research Projects Agency Network) emerged in 1966 [85], the Internet has obtained significant achievements over the last fifty years and developed very fast. At the beginning of 2018, the Internet users in the world are 3.82 billions, i.e., around 40% of the world population has an Internet connection today [9]. Moreover, we come into the era of the mobile Internet. According to the latest research [28], the number of global mobile users has been higher than that of desktop since 2014. Although having a huge success, today's Internet is suffering the overloading of IP address semantic, which results in many shortcomings in the application landscape [53]. In the current TCP/IP protocol stack, the IP address serves both as the *who* to identify the host and the *where* to indicate the attachment point in the network topology, leading to a serious routing scalability problem. As shown in Fig. 1.1, the BGP table size, which is the number of IP prefixes in the BGP (Border Gateway Protocol) routing table of DFZ (Default Free Zone) routers, has grown at a super-linear rate in recent years. After full use of IPv6, the routing scalability issue will be aggravated, since it has much larger addressing space than IPv4. Maintaining such routing table is very cumbersome, since once a site changes its routing information, all the routers need to update, which causes a huge BGP churn. In addition, frequent routing information updates result in the increase of convergence time and packet loss rates.

To cope with the current Internet scalability problems, the Loc/ID separation paradigm is proposed, which separates the single IP addressing space into two orthogonal spaces. Many solutions based on this mechanism are proposed, which can be classified into two categories: *Host-based* and *Network-based*. The former one needs to modify the host stack. The representative architectures are: HIP (Host Identity Protocol) [95] [93] [92], Six/One [117], SHIM6 (Site Multihoming by IPv6 Intermediation) [61] [96], ILNP (Identifier-Locator Net-

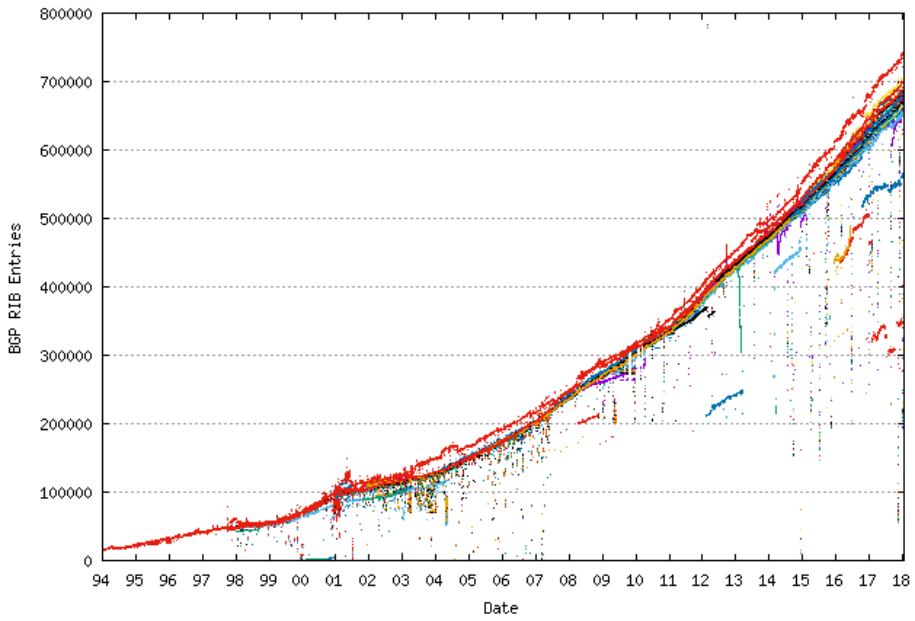


Fig. 1.1 Growth of the BGP Table - 1994 to Present [5]

work Protocol) [37] [34], and LNA (Layered Naming Architecture) [35]. The network-based approach needs to update the routers, and the representative proposals are: GSE (Global, Site, and End-system address elements) [97], LISP (Locator/ID Separation Protocol) [47], and TIDR (Tunneled Inter-domain Routing) [30]. Among them, LISP is outstanding by its relatively comprehensive network systems in both control and data planes. LISP was initially proposed by Cisco in 2006 and has an exclusive IETF WG since 2009.

The key philosophy of LISP is to separate the IP addressing space into two sub-spaces: the *Routing LOCators* (RLOCs) and the *Endpoint IDentifiers* (EIDs). End-to-end communication is based on EIDs, which in an intra-domain context means just performing conventional routing and forwarding. On the contrary, for inter-domain end-to-end communication, RLOCs are used for routing and forwarding. To this end, LISP does a map-and-encap operation, first mapping EIDs to RLOCs (through a DNS-like on-demand Mapping Distribution System (MDS)[1] [55]) and then encapsulating the original packet so to use RLOCs on the outer header. Internetworking between LISP and non-LISP networks is performed through two types of proxies [76]: *Proxy Ingress Tunnel Routers* (PITRs) and *Proxy Egress Tunnel Routers* (PETRs). They are collectively denoted by *PxTR*. PITRs advertise in the BGP routing infrastructure the EID-prefix space on behalf of the LISP sites, so that non-LISP sites can reach them. PETRs are used as exit points from LISP sites when the destination

[1]The terms *Mapping Distribution System (MDS)* and the *mapping system* are used interchangeably in this dissertation.

is not an EID in a LISP site. With help of this mechanism, LISP is able to improve the Internet routing scalability, enable enhanced inter-domain traffic engineering, lightweight multi-tenant carrier-grade VPN, and VMs mobility in Data-Centers [111].

As LISP is a promising technology that brings many benefits to the Internet, it attracts the attention from both academic and industrial world. Many companies, universities and research institutes are now involved in the development of LISP. Hence, the related standardization and research are very active. In terms of deployment, two large-scale experimental LISP platforms: LISP Beta Network and LISP-Lab platform are deployed on the world respectively in 2008 and 2015. Up to date, all the efforts are mainly about theoretic or simulation-based performance analysis, description of LISP architectures, implementations, extensions, applications and so on. However, it lacks of the comprehensive measurement work for LISP technology itself. These performance measurements work will give the research community a clear view about the LISP performance, reveal the points that can be improved, indicate possible research directions, and improve the performance of two platforms. Thus, the large-scale measurements in realistic networks are of paramount importance for the LISP development.

1.2 Contributions

With the goal of evaluating LISP from the different views, in this dissertation we aim at answering the following research questions, which have not been tackled before:

1. Mapping system is essential in LISP, because it decides where the packets will be routed. Thus, does it always provide the identical mapping information for the same destination? Do all of the Map-Resolvers (MRs) of the mapping system offer the same response at the same time? Does the mapping information replies keep consistent if we query them from the different Vantage Points (VPs)?
   We continuously measured LISP Beta Network for seventeen days. Measurements show that the LISP mapping system is stable over most of time. The replies from the different MRs and VPs are also mostly consistent. Nevertheless, instability and inconsistency are observed although they are rare events. We define a new taxonomy to classify these instabilities and inconsistencies so to facilitate a deeper analysis. This work is published in WNM [82].
2. As the results of the last question show that there exists inconsistency between the MRs and VPs, is it possible to implement a comprehensive monitor so to simultaneously supervise the whole mapping system and all the resolvers?

We propose a dynamic LISP monitoring architecture, namely LISP-Views, so to deepen the understandings on LISP and to ease day-to-day operations and troubleshooting. After one full month monitoring the whole LISP mapping system, it demonstrates that LISP-Views provides the complete mapping information. Furthermore, with our proposed monitoring platform, more mapping system performance metrics such as reliability, latency, and configuration issues can be assessed, which helps for further LISP improvements. This work is published in ITC [77].

3. By using proxy approach to communicate with the legacy Internet, a stretch in the path is introduced. How much is it in the real world performance when the platforms integrate with the legacy Internet? Under which conditions such stretch has an important impact on the performance? Conversely, under which situation such overhead is so small that it can be ignored?

   To evaluate LISP interworking performance with legacy Internet, we conduct two measurement campaigns. The results of the two experiments are coherent. Both of them show that LISP is stable although the stretch is introduced due to the use of proxies. We find that the proxy indeed introduces negative effects for the destinations located in Europe and America, but can be ignored for the intercontinental long-distance transmissions to Asia destinations. Besides, we observe that the position of proxies is very important. If it is either near to the sources or the destinations can decrease the latency a lot. This work is published in INFOCOM Student workshop [80], ACM SIGCOMM Student workshop [79], and Elsevier Computer Networks [81].

4. As LISP separates the identifier and the topological attachment point from the traditional IP address, it is able to maintain the communication between two hosts during the handover. LISP can be implemented on the border routers or the terminals or even both of them. Thus, which network element supports LISP showing a better performance, i.e., having a shorter handover delay or introducing less overhead during the mobility? What are the advantages and shortcomings of each case?

   To explore the mobility performance of the LISP implementation on each network element, we design three different scenarios: 1) Only mobile node supports LISP; 2) Only border routers support LISP; 3) Both mobile node and border routers support LISP. The numerical analysis shows that the 1st and 3rd scenarios can achieve the mobile node roam through the different subnets, whereas they require additional permanent EID for each mobile node. The 1st has no help on reducing the BGP table size and the 3rd scenario shows much longer handover delay than the other two's. The 2nd scenario presents the smallest handover delay and needs the least IP addresses, but it can only handle mobility within the same subnet. To verify the numerical anal-

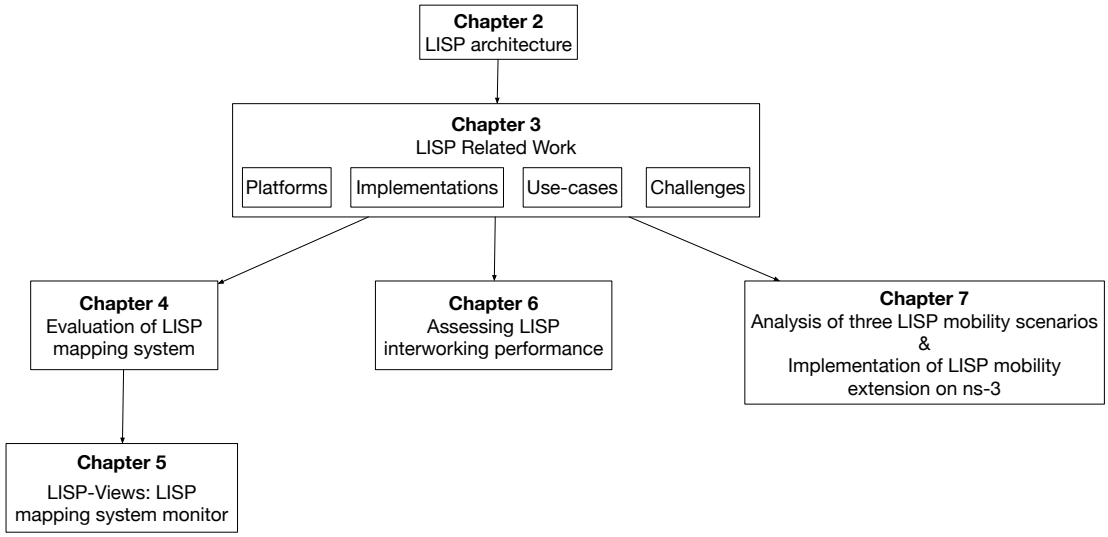


Fig. 1.2 Structure of this dissertation

ysis, we implement a LISP simulator with LISP mobility extensions under ns-3.27 by extending an existing ns-3 based LISP implementation.

1.3 Manuscript organization

The remainder of this dissertation is organized as follows and shown in Fig. 1.2:

- *Chapter 2*: Describes the basic LISP mechanisms, presenting how the packets are exchanged between two LISP-sites and with the legacy Internet. It also shows the different scenario of LISP mobility.
- *Chapter 3*: The challenges brought by LISP, LISP-related studies and missing works are also illustrated. Introduces the LISP current status, such as: the available LISP platforms, the first LISP monitor, the LISP manual management tool, the open source LISP implementations and LISP use-cases.
- *Chapter 4*: This chapter is about the evaluation of LISP Mapping System. In this chapter, we observe that if the mapping system always provides the same mapping information over time, and if we can obtain the identical mapping answers from the different network entities.
- *Chapter 5*: Describes the proposition of LISP-Views: a LISP Mapping System Monitor at large scale. In this chapter, we introduce how this architecture is implemented, how it has been validated and what experimental results it can provide.

- *Chapter 6*: Assesses LISP interworking performance through RIPE Atlas. This chapter presents how the experiments are conducted, how much is the stretch introduced by proxy in the real world, and whether this stretch has a big impact on the performance or it can be ignored.
- *Chapter 7*: Introduces ns-3 Implementation of LISP mobility extension and numerically analyses the performance of different LISP mobility scenarios. In this chapter, we first describe how to implement the LISP mobility extension on the ns-3. Then we define three different LISP mobility scenarios, provide the numerical analysis of each scenario, and compare the advantages and shortcomings between them.
- *Chapter 8*: Concludes the dissertation and discusses perspectives for further work.

Chapter 2

LISP overview

In this Chapter, the LISP principles and its mechanisms to which we refer to in the reminder of the dissertation are introduced. The basic architecture of LISP is introduced by presenting its Data Plane and Control Plane. We detail how the traffic between two LISP-sites is exchanged by using an example. A new network element, which is used to allow communication between LISP-site and the legacy Internet, is also presented, as well as the sequence of packets exchange between them. Then, two mechanisms used to update the LISP mapping cache are illustrated. Finally, this chapter ends with the explanation of LISP mobile node, i.e., a LISP-enabled mobile terminal, and the introduction about how the LISP operations are performed when the LISP-enabled node is behind a LISP-speaking router.

2.1 LISP architecture

The *Locator/Identifier Separation Protocol (LISP)* separates the traditional IP address role into two logical sub-spaces: (i) *Endpoint IDentifier (EID)* and (ii) *Routing LOCator (RLOC)*. Both of them are the conventional IPv4 (32-bit) [105] or IPv6 (128-bit) [45] addresses. The EID is the identifier of hosts, which is locally routable in the LISP-site. Normally they are the addresses of source and destination hosts. Within the same LISP-site, the hosts can directly communicate with each other via EIDs as in a legacy IP network. The host gets a destination EID in the same way it obtains the destination IP address today, for instance through a Domain Name System (DNS) [89] lookup or Session Initiation Protocol (SIP) [107]. RLOCs denote the attachment points of EIDs in the Internet topology, i.e., the source/destination address of the edge routers, behind whom the EIDs can be found. RLOCs addresses are used to forward the packets in the core Internet.

When packets are transferred between two different LISP-sites, an encapsulation on the edge router is needed: the source and remote destination EIDs are in the inner packet

header, while the source and destination RLOCs are in the outer header. The EID-prefixes are not necessary to be announced on the Internet core, so the BGP routing table size can be reduced. Similarly, in stub networks, it is not necessary to know the routing information of core Internet any more. The aforementioned encapsulation and forwarding operations are the responsibilities of LISP Data Plane. More details are introduced in Sec. 2.1.1. To bind the EIDs and RLOCs, a new network entity named Mapping Distribution System (MDS) is used, whose front end is consisting of Map Resolver (MR) and Map Server (MS) [56]. The MDS is not only able to provide the mapping information between the EIDs and RLOCs for the LISP-sites (we simply call them LISP Map-Replies), but also indicate which addresses are not LISP sites and have no RLOCs, meaning that are part of the legacy Internet (these are called Negative Map-Replies). The Control Plane is described in Sec. 2.1.2.

2.1.1 LISP Data Plane

The LISP Data plane mainly takes care of the encapsulation and decapsulation operations done at the source and destination networks. LISP basically provides a level of indirection through a tunneling mechanism over the core Internet (the so called *Default Free Zone (DFZ)*), in order to provide end-to-end communication. More specifically, any host willing to communicate with another host residing at remote LISP-site, uses its own EID as source address and the EID of other host as destination address to generate regular IP packets. These packets are first transferred as usual to the border router, called *Ingress Tunnel Router (ITR)*. ITR performs a mapping lookup in its Cache, or in case of miss, it queries the MDS for the mapping information. Then the ITR encapsulates the regular IP packets into a LISP packet by adding a tunnel header [47], where the RLOC of the ITR is used as source address and the destination RLOC as destination address. After encapsulation, the LISP packets are forwarded over the core Internet. When arriving at the destination border router, called *Egress Tunnel Router (ETR)*, the LISP packets are decapsulated (i.e., the tunnel header is removed) and transferred to the destination host. A device which acts as both an ITR and an ETR is denoted as *xTR*.

Mappings are stored in two data structures on the xTRs: the *LISP Database* and the *LISP Cache*. The LISP Database is populated by configuration and stores all known EID-to-RLOC mappings, for which the EID-Prefixes are behind the xTR. On an ITR, this helps select a source RLOC used in encapsulation. While on an ETR, it allows to verify whether itself is the proper ETR connecting to the destination EID, so that such ETR is able to forward the decapsulated packets to the final destination.

The LISP Cache temporally stores mappings for the EID-prefixes of the remote communicating end-points. On an ITR, it is used to select the destination RLOC of the outer

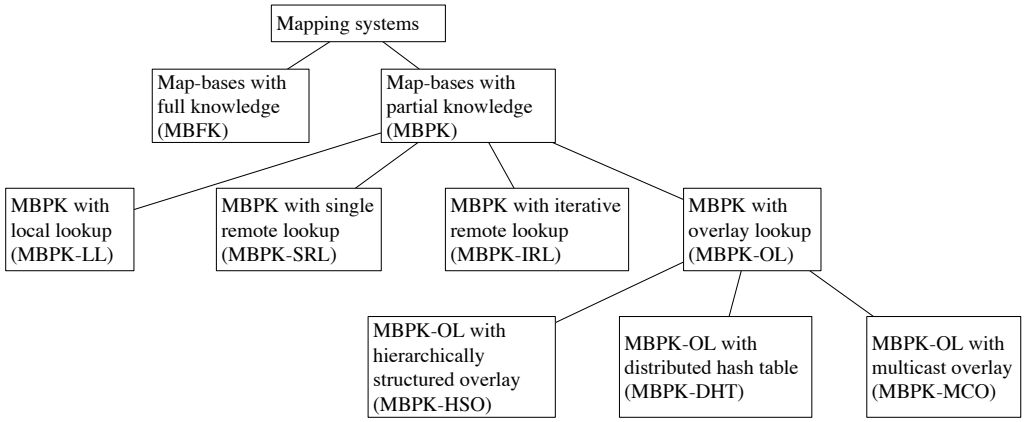


Fig. 2.1 Hierarchical taxonomy of mapping systems (from [65])

header of the LISP-encapsulated packet. While on an ETR, it is used to perform a basic anti-spoof verification. Different from the LISP Database, the LISP Cache is populated on demand. The procedure is triggered by the first packet of a new flow, which can not find a suitable mapping for the destination EID among the mappings stored in the LISP Cache. The mappings in the LISP Cache are purged if not used upon expiration of a time-out ([66, 72, 119, 43]).

2.1.2 LISP Control Plane

The LISP Control plane is mainly tasked to discover the associations of EIDs with RLOCs, i.e., find the *mappings*. A new network entity named *Mapping Distribution System (MDS)*[1] is introduced by LISP to perform such EID-to-RLOC distribution. In particular, a mapping contains an EID prefix and an associated list of RLOC tuples: *<RLOC, Priority, Weight>*. The RLOC refers to the IP address of each xTR interface. Each RLOC is assigned a Priority. The RLOC with the highest priority is preferred. The weight is taken into account only to perform load balancing in the case of several RLOCs having the same priority.

The current BGP-based Internet routing architecture relies on a push model, in which routing information is pushed to the whole Internet. The Locator/Identifier Split-based Internet routing instead relies on a pull model, where the *MDS* provides a mapping upon an explicit query, where routing information is pulled only when actually needed. This paradigm

[1]The terms *Mapping Distribution System (MDS)* and *mapping system* are used interchangeably in this dissertation.

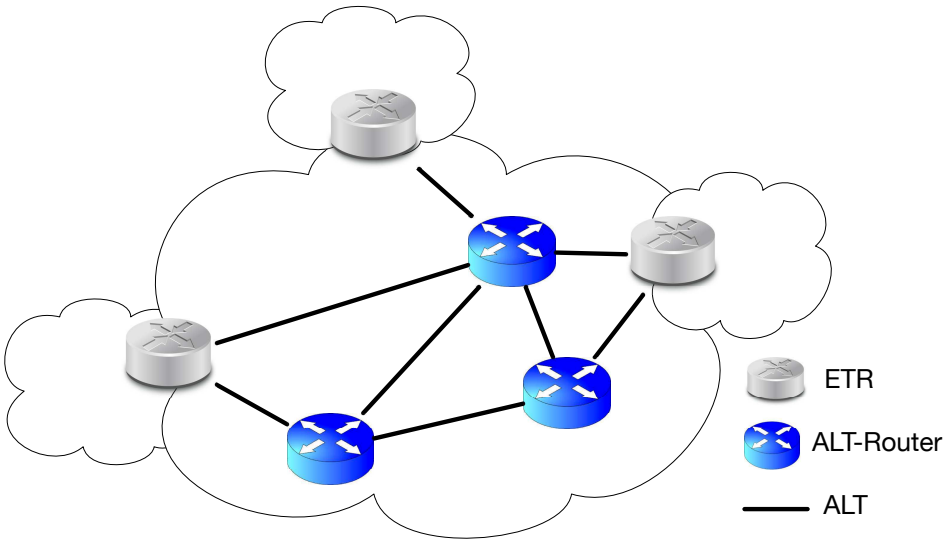


Fig. 2.2 Illustration of the LISP+ALT topology

shift requires making routing information available in an on-demand manner. Fig. 2.1 shows the current possible taxonomy of mapping systems [65]. So far, several mapping systems have been proposed for LISP, such as: LISP-TREE [69], LISP-NERD (Not-so-novel EID to RLOC Database) [75] and LISP-CONS (Content distribution Overlay Network Service for LISP) [39]. However, only two have been deployed: *LISP Alternative Logical Topology (LISP+ALT)* [57] and *LISP Delegated Database Tree (LISP-DDT)* [58].

LISP+ALT

LISP+ALT was the initial mapping system for LISP, where the ETRs store mappings they are authoritative for [108]. The basic idea is to use BGP to construct an overlay, named Alternative Logical Topology (ALT), to establish reachability between ETRs via tunnels (e.g., GRE tunnels [50]). As shown in Fig. 2.2, the routers in the ALT are called ALT-Routers. Each ALT–Router maintains a BGP session with its neighbor and announces the EID–prefixes it is authoritative for. These EIDs are then routable in the ALT. Note that the ALT–Routers only exchange aggregated EID-prefixes that can be reached through them between neighbors, but they do not contain the mapping information. To get a mapping, an ITR constructs a Map-Request for the queried EID by using its own RLOC as source address and the queried EID as destination address, then sends it to an ALT-Router. The Map-Request is forwarded over the ALT and eventually reaches an originator ETR for the EID-

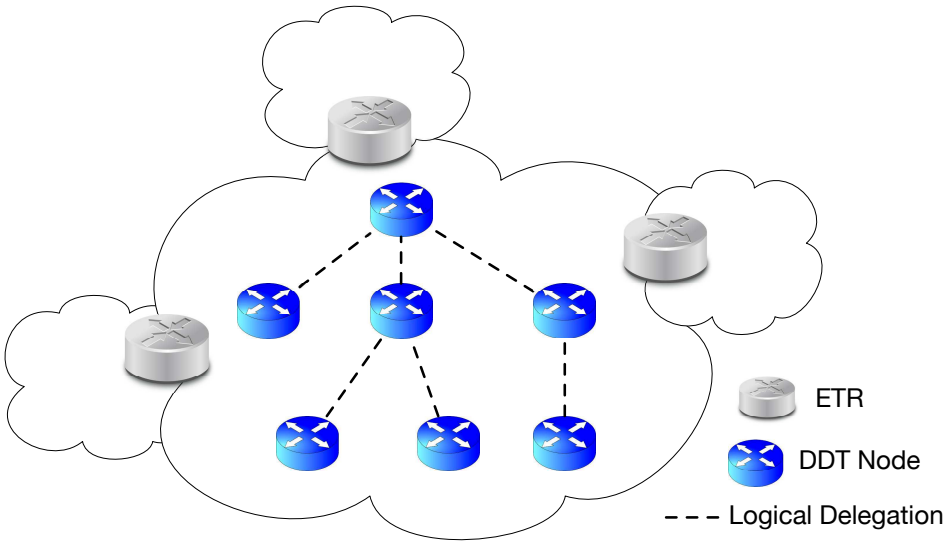


Fig. 2.3 Illustration of the LISP-DDT topology

prefix that matches the destination EID. This ETR resolves the Map-Request and directly sends back a Map-Reply to the inquirer ITR, this time not using the ALT. LISP-ALT belongs to the Map-Bases with Partial Knowledge using Hierarchically Structured Overlay (MBPK-HSO) in Fig. 2.1.

LISP-DDT

LISP-DDT is a hierarchical distributed database that embodies the delegation of authority to provide mappings from LISP EIDs to RLOCs . Its hierarchical organization is shown in Fig. 2.3. It consists of *Map-Resolvers (MRs)* and *Map-Servers (MSes)* to provide a simplified "front end". The MR is a network infrastructure component that accepts LISP Encapsulated Map-Requests from an ITR and resolves the mapping using a DNS-like distributed database. The MS learns of EID-Prefix mapping entries (a list of EID-Prefixes plus a set of RLOC tuples) from ETRs. The hierarchy is maintained as a tree where a root server is responsible for the entire EID space. This space is divided into several portions and each one is managed by one of the root's children. Mapping information is only stored at the tree leaves, which are made of MSes. Intermediate nodes only maintain pointers to their children.

If an ITR cannot find a mapping in its LISP Cache for a new flow, it sends a query, called *Map-Request*, to a MR [56]. The MR determines whether the queried destination IP address

is an EID. If not, a *Negative Map-Reply* is directly returned; otherwise, the root server is queried. The root replies with a pointer to its child node responsible for the queried EID. The process is recursively repeated to find the child root of the sub-tree where a mapping can be retrieved for the queried EID. When the leaf is reached, the mapping is then retrieved by sending a Map-Request to the ETR authoritative for the matching EID-prefix. The ETR in turn directly replies a *Map-Reply* message containing the whole requested mapping information to the ITR initiating the query.

When an ETR first contacts an MS after restarting or changing its database, it sends the Map-Register messages to the MS to publish its EID-prefixes as well as the mapping information with an increased interval, until 20 seconds. For maintaining the association, the registration interval should be expanded to at least 1 minute. If the MS has not received a valid Map-Register within the past 3 minutes, it removes the registration of this ETR. If the "want-map-notify" (M-bit) flag is set in Map-Register, the received MS needs to send back a Map-Notify message to inform the ETR that the Map-Register has been received and processed. If the "proxy Map-Reply" flag (P-bit) is set, the MS answers the Map-Requests to the queried ITR by providing the mapping information on behalf of the ETR. MRs offer an interface to the MDS for the xTRs, so that the whole complex MDS mechanism can be hidden.

2.1.3 Communication between two LISP-sites

We use Fig. 2.4 as an example to describe how the LISP packets are processed in details in the case of the packets exchange between two LISP-sites.

When the host in the AS_s (Autonomous System) communicates with the host in the AS_d, it uses EID_s as source address and EID_d as destination address. After the conventional IP routing to xTR_1 (the function of ITR and ETR are combined together in Fig. 2.4) labeled as *"1"*, ITR_1 first checks in its mapping cache [66] to find out the association between EID_d and its RLOC. If there is no record, it sends a Map-Request to Map-Resolver (MR), labeled as *"2"*. MR searches which Map-Server (MS) stores the mapping information of EID_d and forwards the Map-Request to that MS. If MS is authoritative (i.e., acts as a proxy), it directly returns back a Map-Reply containing the mapping information (this procedure is not labeled in the figure). If not, after forwarding the Map-Request within MDS with the procedure described in Sec. 2.1.2 to MS and then to the destination side (label *"3"*), xTR_1 receives a Map-Reply from one of ETRs of EID_d (label *"4"*). If the Map-Reply shows that the priority of ETR_3 is higher than ETR_4, then the ITR_1 encapsulates the packets by adding $RLOC_{ETR_3}$ as the destination address and $RLOC_{ITR_1}$ as the source address in the outer header and sends the encapsulated packets through the Internet as labeled *"5"* in Fig. 2.4. Once ETR_3 receives

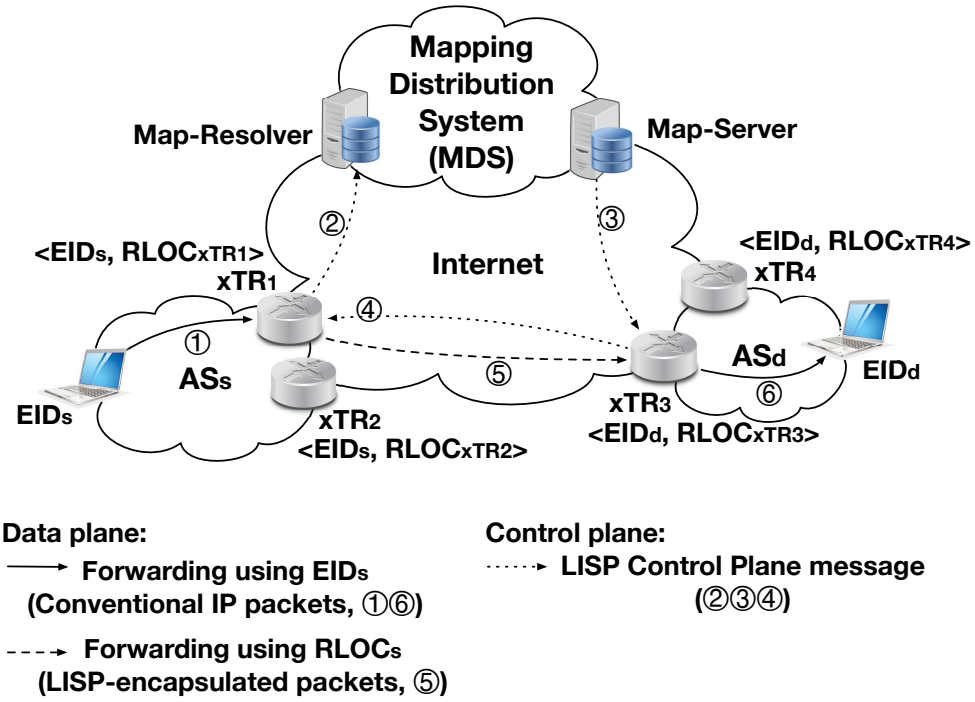


Fig. 2.4 LISP architecture overview

these packets, it decapsulates and forwards them by using the original IP packets (label *"6"*). If the destination site is not a LISP-site, the MR directly returns back the Map-Reply without mapping information to ITR_1.

Based on LISP architecture, there are three types of outcomes for a query:

- *LISP Map-Reply*, means that the queried IP address belongs to a LISP site (i.e., EID) and the Map-Reply contains the mapping information for this site, including: the association between the EID-prefix that queried EID belongs to, and the list of RLOC tuples <RLOC, Priority, Weight> of the destination site.
- *Negative Map-Reply*, means that the prefix covering the queried IP address belongs to a non-LISP site (i.e., conventional site), and the Map-Reply contains no mapping information. We consider these two kinds of replies as successful queries.
- *No Map-Reply*, means that the xTR does not receive any reply during a certain time. According to the standard described in [47], the time-out is set to 3 seconds. But this value can be modified in the implementation to adapt the experiment needs. In this case, we consider the query as failed.

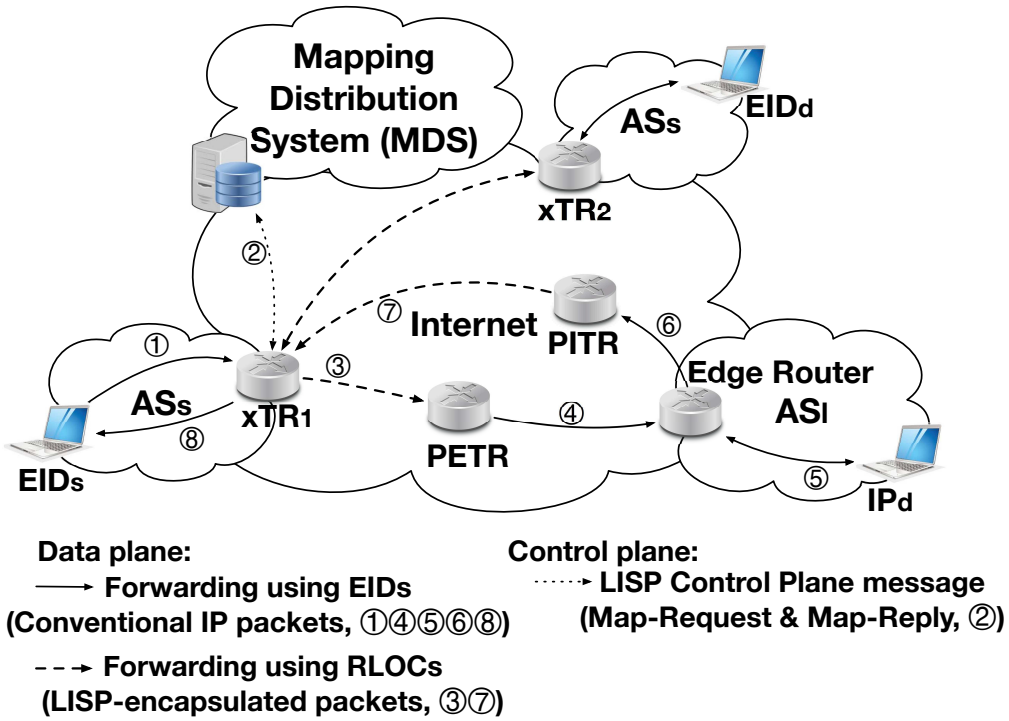


Fig. 2.5 LISP architecture with the packets forwarding sequence

2.2 Interworking with legency Internet

Sec. 2.1.3 describes how the LISP packets are exchanged between two LISP-sites. This section discusses how the packets are processed when a terminal behind LISP-site communicates with a terminal on the legacy Internet.

Although EIDs are syntactically identical to IPv4 or IPv6 addresses, routes to them are not announced in the global routing system, so an interoperability mechanism is needed for non-LISP-speaking sites to exchange traffic with LISP-speaking sites. Thus, two new network elements are introduced: LISP Proxy Ingress Tunnel Router (PITR) and LISP Proxy Egress Tunnel Router (PETR). The PITR acts as an intermediate LISP ITR for the hosts on legacy Internet. It is in charge of announcing one or more highly aggregated EID-Prefixes on behalf of LISP-sites into the public Internet and receives the traffic from the public Internet to encapsulate them as LISP packets. The PETR acts as an ETR for traffic destined to non-LISP sites. It receives the traffic from the ITRs, decapsulates the LISP packets into conventional IP packets and sends them to the legacy Internet.

Fig. 2.5 shows an example for the packets exchanging between LISP-site and non LISP-site (i.e., legacy Internet). The host with EID_s in the AS_s exchanges the packets with the host with the IP_d in the AS_l. With the same procedure introduced in Sec. 2.1.3 labeled as *"1"* and

"2" in Fig. 2.4, the original IP packets are sent from host to xTR_1, and the later one sends the Map-Request to MDS to query the mapping information. Since the destination is in a non-LISP-Site, the MR directly returns back the Map-Reply without mapping information (i.e., Negative Map-Reply) to xTR_1, the packet can either be sent non-encapsulated or xTR_1 encapsulates it using the statically configured RLOC of the PETR as the outer destination address and forwards the LISP packets there. When PETR receives the LISP-encapsulated packets, it performs a decapsulation as an ETR and then forwards the inner IP packets to the legacy destination host by using the normal BGP routing, labeled as procedure *"4"* and *"5"* in Fig. 2.5.

When the packets come back, they are forwarded in a traditional way from the host with address IP_d in AS_l to the PITR, denoted as procedure *"5"* and *"6"* in Fig. 2.5. A PITR shares many characteristics with ITRs, as it also queries the mapping information of destination to the MDS and encapsulates the non-LISP Internet traffic into LISP-encapsulated packets and route them to their destination RLOCs (label *"7"*). Target ETR_1 then dencapsulates the packets and send them to the destination host EID_s. The PETR and PITR can be noted together as PxTR if they are located on a same physical router.

2.3 Mapping cache update mechanisms

Given that LISP mapping distribution is based on a pull model, in case that mapping changes occur in ETRs, the only way that makes remote ITRs be aware of such change and update the corresponding mapping information is to request again the mappings to ETRs. Thus, two mechanisms are proposed to notify ITRs that mapping has changed and they need to get the latest mapping: Solicit-Map-Request and Map-Versionning.

2.3.1 Solicit-Map-Request (SMR)

Soliciting a Map-Request is used by ETRs to tell remote ITRs to update the mappings they have cached. When the mappings in the Database of an ETR changes, the ETR sends the Map-Requests with the SMR bit set (called Solicit-Map-Request (SMR) message) for each RLOC in its Cache. A remote ITR that receives the SMR message will take a check in its Cache whether the source RLOC of the SMR message, i.e., the RLOC of ETR with mappings change is in its Cache. If yes, the remote ITR sends a Map-Request to the MR (as shown in the left hand part of Fig. 2.6) or directly to the sender of SMR (as shown in the right hand part of Fig. 2.6). If the RLOC of sender is not in the Cache of ITR any longer, the later drops the SMR and does not send the Map-Request. The later situation occurs when

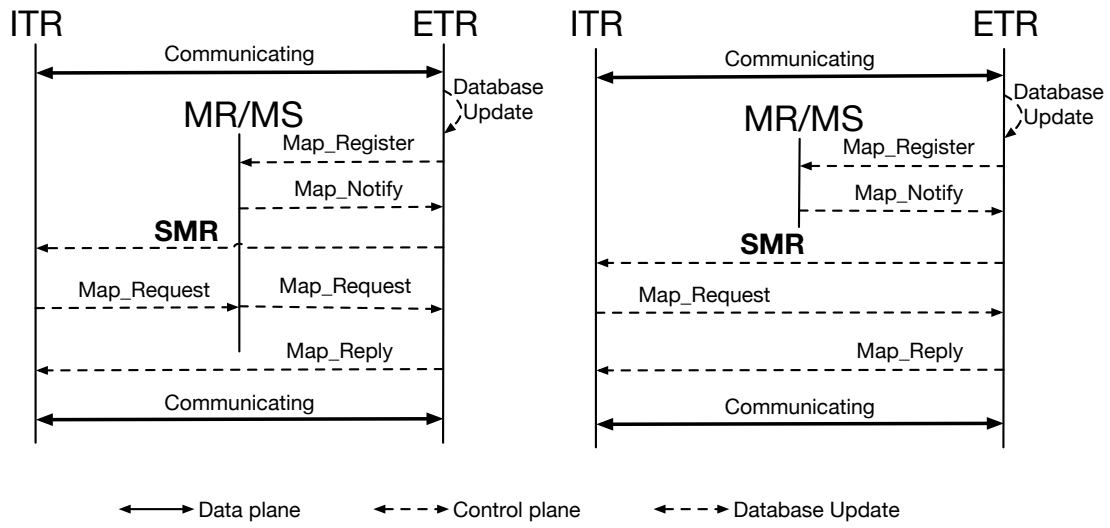


Fig. 2.6 Packet sequence of LISP SMR mechanism

the ETR sends the SMR to all RLOCs stored in its Cache, but the remote ITR actually has not sent the packets to it for a long time. Between sending the SMR message and receiving the new Map-Reply, the ITR continues to use the mapping information previously cached, and that may cause packet loss.

2.3.2 Map-Versioning

Map-Versioning is used to inform through the Data Plane the xTR that the mappings of a remote site changed and update the mapping information with the help of SMR message. Map-Versioning associates a Map-Version number to each LISP mapping information and transports such a version number in the LISP-specific header [67]. When a mapping changes, a new version number, normally incrementally higher than the previous one, is assigned to the updated mapping. When an ITR encapsulates a packet, it selects the version number assigned to the mapping stored in the Database as *source RLOC version number*, and selects the version number assigned to the mapping contained in the Cache as *destination RLOC version number*. When an ETR receives such packets and dencapsulates them, it makes a comparison between the received version number and the local one. If the source map version is higher than the one that ETR stores in its Cache, i.e., the mappings of remote ITR changed, the ETR sends a Map-Request to the MR or directly to the remote ITR to retrieve the latest mapping information (the left figure of Fig. 2.7). If the destination map version is lower than the one contained in the Database of ETR, i.e., the local mappings

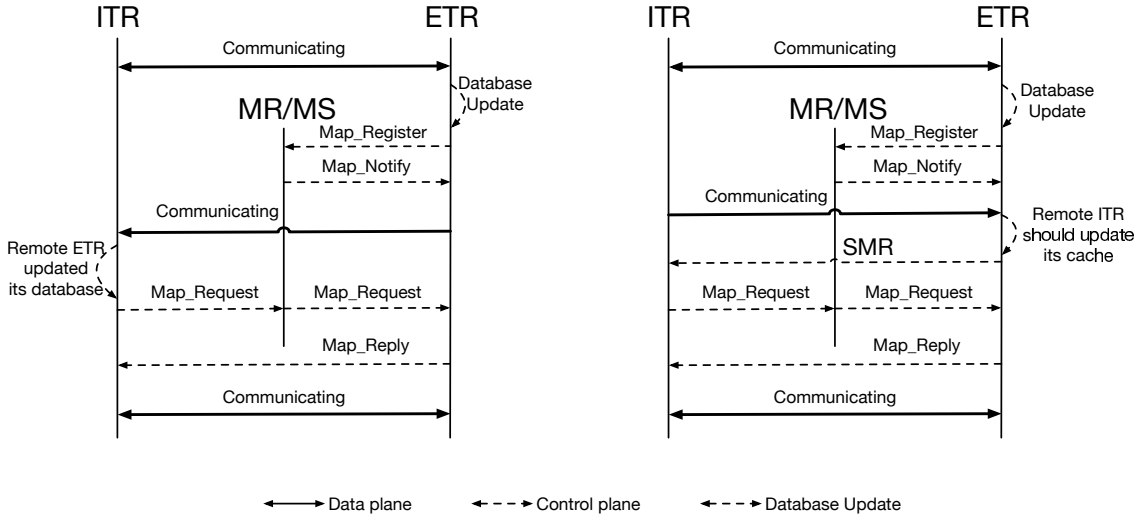


Fig. 2.7 Packet sequence of LISP Map-Versioning mechanism

changed but the remote site does not know, then the ETR sends a SMR to the remote ITR to let it update the cached mapping (the right figure of Fig. 2.7).

2.4 LISP mobility

Nowadays, the explosion of mobile devices that need the massive mobile Internet traffic requires the evolution of mobile network architecture. Keeping the communication during the roaming without any interruption becomes a key point. Thanks to LISP separating the identifier of a host and the locator of the attachment point, the connection between two hosts can be kept alive while roaming. Because the EID, which is used as the inner address, is permanent, whereas only RLOC (used as the outer address) is changed.

By using LISP mobility, it is possible 1) to allow TCP connections to stay alive while roaming; 2) to provide the shortest bidirectional data paths between a Mobile Node (MN) and Correspondent Node (CN); 3) to allow MN to communicate with another MN when both are roaming; 4) not require fine-grained routes in the core network, nor the home-agent, foreign agent or other data plane network elements to support mobility; 5) there is no triangle routing of data packets as is found in Mobile IP [98]; 6) and there is no IPv6 new extension headers to avoid triangle routing [71]. LISP Mobile Node (LISP-MN) takes advantage of the LISP infrastructure so to overcome the limits imposed by Mobile IP. The implementation of LISP-MN is achieved by Open Overlay Router (OOR) [21], which will be introduced in Sec. 3.3.1.

Table 2.1 Catalogs of mobility

Name of catalogs	MN	Site
LISP-MN in LISP-Site	LISP	LISP
LISP-MN in non-LISP-Site	LISP	non-LISP
MN in LISP-Site	non-LISP	LISP
normal mobility	non-LISP	non-LISP

LISP can be implemented on both border routers and terminals. According to the adoption of LISP and host/router supports LISP, mobility can be divided into 4 categories: LISP-MN in LISP-Site, LISP-MN in non-LISP-Site, MN in LISP-Site, and MN in non-LISP-Site (shown in Tab. 2.1). However, the case "MN in non-LISP-Site" is the current conventional mobility, which has no relationship with LISP. There are some proposals to solve the seamless mobility issues for the traditional handover, such as Mobile IP (MIP) [100], Mobility Support in IPv6 (MIP6) [99] [88], Multipath TCP (MPTCP) [54] and so on. Thus, in the following sections, we will present LISP-MN first and then only describe how the packets exchange when the handover occurs for the first three cases.

2.4.1 LISP Mobile Node (LISP-MN)

This section introduces a mobile LISP-speaking node, called LISP Mobile Node (LISP-MN). It can reside behind both LISP-Site and non-LISP-Site. It is able to change its attachment point during the communication. The LISP-MN implements a subset of the standard xTR functionality [49]. It can send Map-Request to MR to get the mapping information of the remote host.

When a LISP-MN resides in a LISP-Site, it is assigned an EID taken from the site's EID-prefix as its RLOC, called *Local RLOC (LRLOC)*. Thus, LISP-MN stores not only its permanent unique EID but also the LRLOC and registers this mapping information to the MS. The conventional IP packets produced by the LISP-MN are encapsulated by itself using the permanent EID as the inner source address and LRLOC as the outer source address. The encapsulated LISP packets are forwarded to xTRs and encapsulated again.

When a LISP-MN resides in a non-LISP-Site, it is assigned an IP address taken from the site's prefix as its *LRLOC*. LISP-MN also need to store both its permanent unique EID and the LRLOC and register this mapping information to the MS. The conventional IP packets produced by the LISP-MN are encapsulated by itself using the permanent EID as the inner source address and LRLOC as the outer source address. The encapsulated LISP packets are forwarded to the border router and natively forwarded to the Internet core.

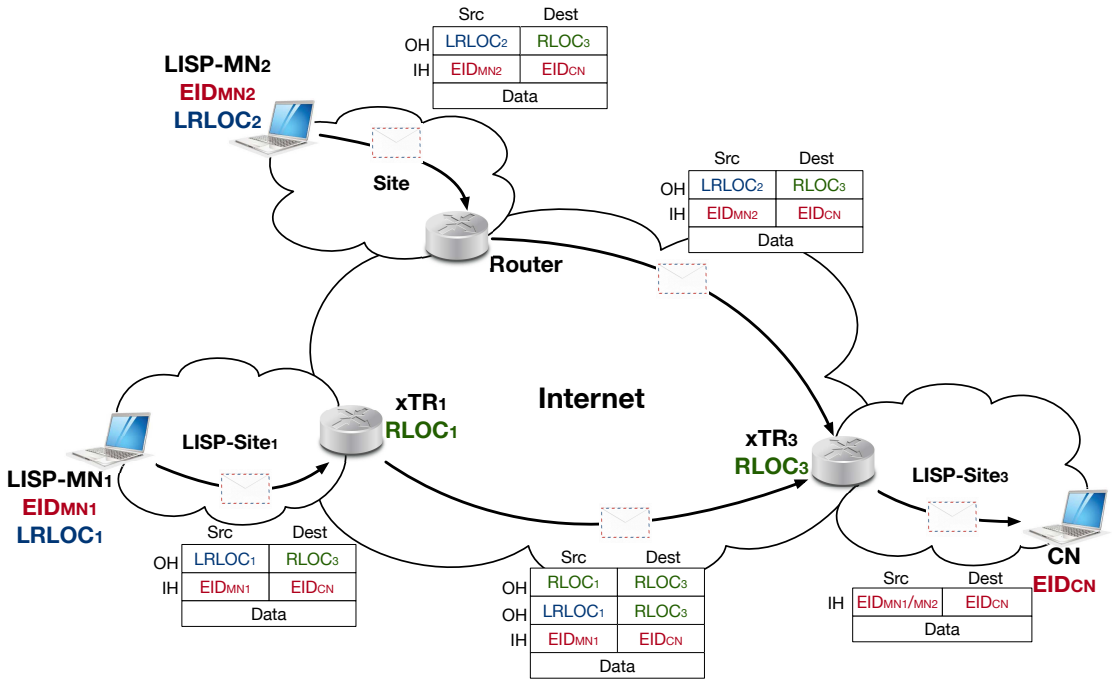


Fig. 2.8 Illustration of LISP-MN in LISP-Site and non-LISP-Site

The packet flow sequence of LISP-MN in LISP-Site and non-LISP-Site are shown in Fig. 2.8. Two LISP-MNs, i.e., LISP-MN_1 and LISP-MN_2 send the packets to a remote Correspondent Node (CN) . The only difference is that LISP-MN_1 resides in a LISP-Site, while LISP-MN_2 resides in a non-LISP-Site. The LISP-MN_1 produces the traditional IP packets with EID_{MN1} as source address and EID_{CN} as destination address. Then it encapsulates by adding $LRLOC_1$ as outer source address and $RLOC_3$ as outer destination address by itself after querying the mapping information. The LISP packets with one time encapsulation are forwarded to xTR_1. The latter gets the mapping information from MR and encapsulates once again the packets by adding the $RLOC_1$ as the outer source address and $RLOC_3$ as the outer destination address, then sends the double LISP encapsulated packets on core Internet. The xTR_3 needs to dencapsulate the packets twice after receiving and verifying the packets, and finally sends to CN.

The situation for LISP-MN_2 is simpler. It also produces the traditional IP packets and encapsulates the packets by itself, then sends to its border router. The *Router* does nothing related to LISP but only forwards the already LISP-encapsulated packets to the Internet core. Thus, the packets are encapsulated only once, whose inner source and destination address is respectively EID_{MN2} and EID_{CN}, outer source and destination address is respectively $LRLOC_2$ and $RLOC_3$.

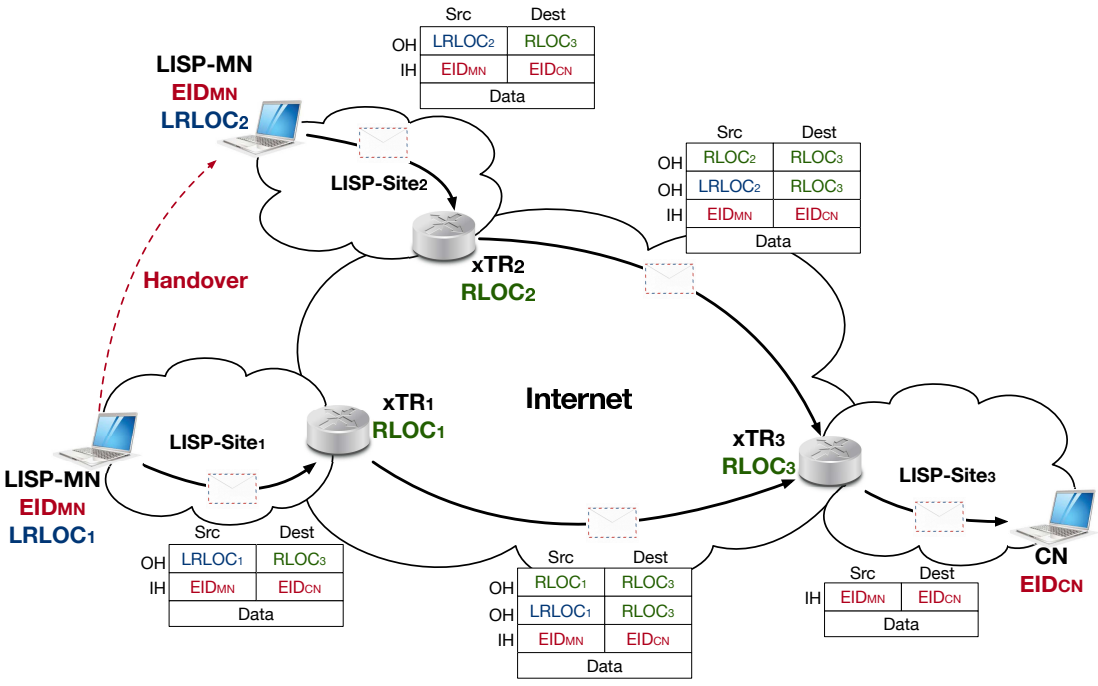


Fig. 2.9 Packet flow of mobility when LISP-MN in LISP-Site

2.4.2 MN mobility in LISP-Site

When the MN is normal host without any LISP function, while the border router supports LISP, i.e., xTR, the packets are exchanged with the remote *CN* as the basic communication between two LISP-Sites introduced in Sec. 2.1.3. The EID of MN is distributed from the EID-prefix of the LISP-Site, if MN moves to a new LISP-Site but keeps the original EID, it can not exchange the packets with its new xTR, since its EID does not belong to the EID-prefix of the new LISP-Site. Thus, in this case, the mobility of MN can only conduct within a subnet instead of roaming through different domains.

2.4.3 LISP-MN mobility in LISP-Site

This section describes how a LISP-MN moves from one LISP-Site to another and the packet flow is shown in Fig. 2.9. At first, packets exchanged between LISP-MN residing in LISP-Site$_1$ and *CN* are double encapsulated LISP packets as mentioned in Sec. 2.4.1. After conducting the handover, LISP-MN still uses its permanent EID_{MN}. However, since it changes the attachment point, xTR_2 distributes a new EID from its EID-prefix to it, and it uses this as its new LRLOC, labeled $LRLOC_2$ in the figure. The LISP-MN should register itself with the new mapping information to the MDS and ask the xTR_3 (xTR of *CN*)

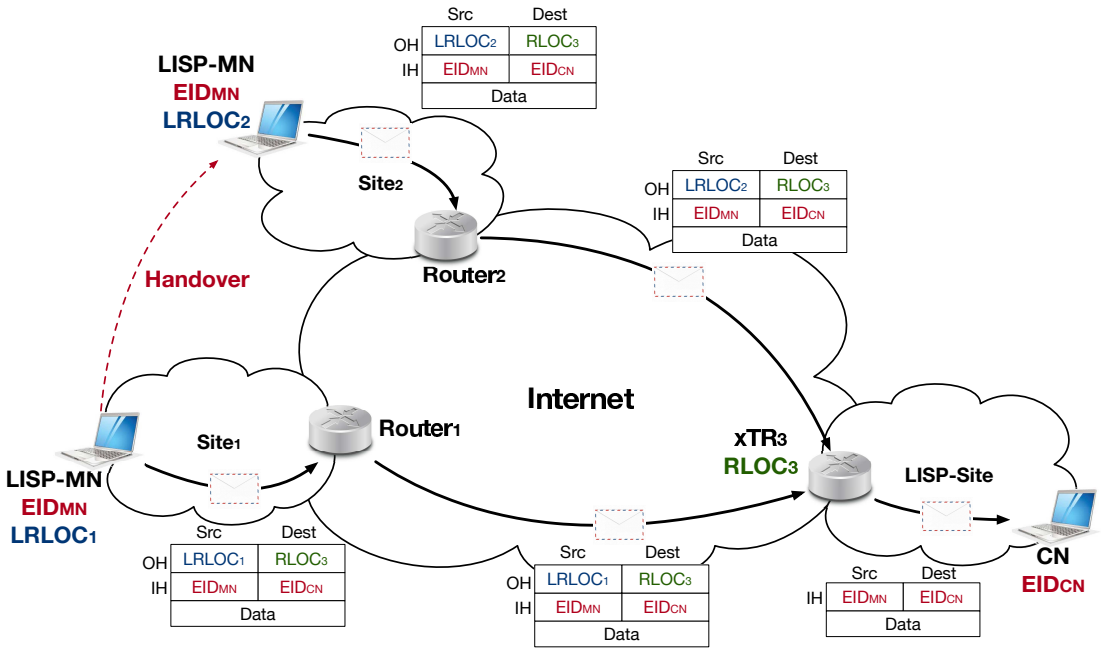


Fig. 2.10 Packet flow of mobility when LISP-MN in non-LISP-Site

to update its LISP-Cache by *SMR* or *Map-Versioning* (introduced in Sec. 2.3). When receiving the packets from LISP-MN, the xTR_2 conducts the second encapsulation and sends the packets to xTR_3. The benefit of such mobility is that the LISP-MN can roam through the different subnets without interrupting the communication with the remote *CN*, because LISP-MN never changes its EID. However, to conduct double lookup may lead to higher delay of handover.

2.4.4 LISP-MN mobility in non-LISP-Site

The Fig. 2.10 illustrates the packet flow of the LISP-MN roaming between two non-LISP-Sites. The $Site_1$ distributes an IP address from its prefix to LISP-MN, and the later uses this as its LRLOC, labeled as $LRLOC_1$ in the figure. The LISP-MN encapsulates the packets by itself and sends to $Router_1$. $Router_1$ just natively forwards the packets to xTR_3. When LISP-MN moves to another site labeled $Site_2$, it receives an IP address from a different prefix and uses it as the new LRLOC. It registers itself to the MDS with the EID_{MN}-to-$LRLOC_2$, and informs the remote xTR_3 (*RLOC* of *CN*) to update the mapping information by sending *SMR* or using *Map-Versioning*. In this scenario, LISP-MN can also permit to roam through the different subnets, but it can also reduce the high latency of double registration during the handover.

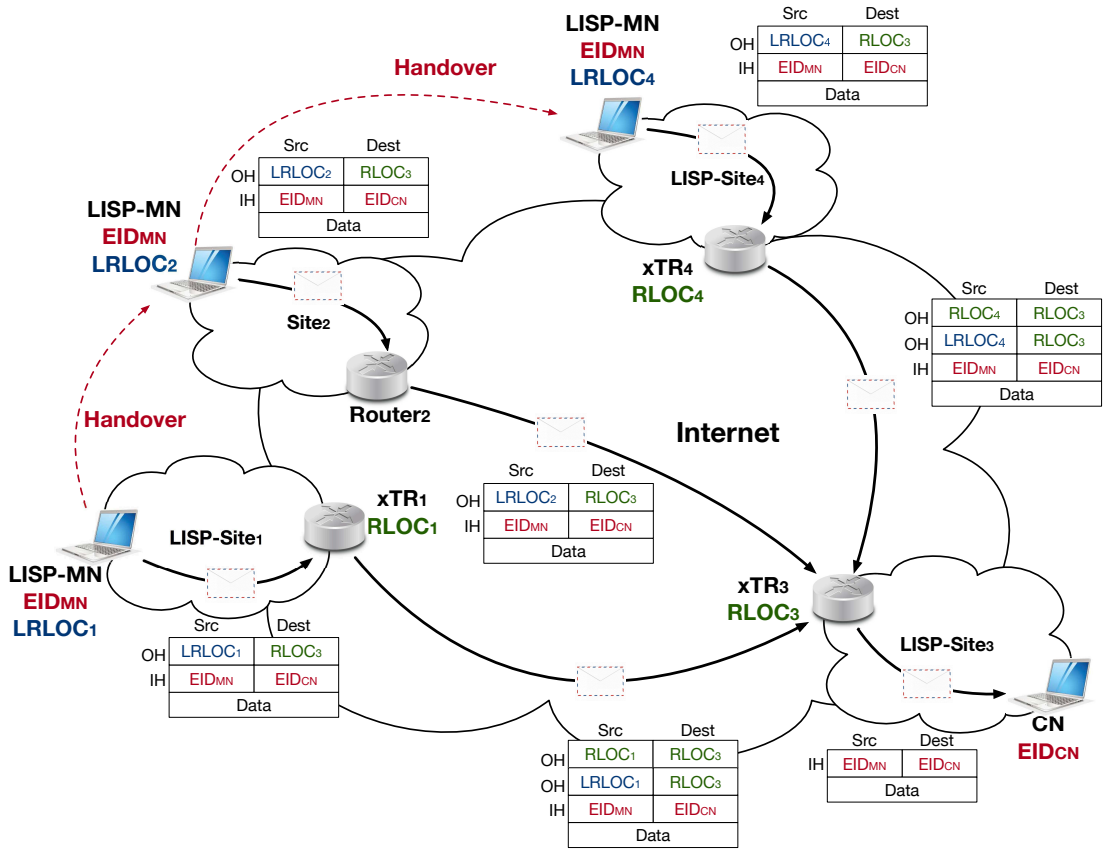


Fig. 2.11 Packet flow of LISP-MN mobility between LISP-Site and non-LISP-Site

2.4.5 LISP-MN mobility between LISP-Site & non-LISP-Site

This section describes how a LISP-MN moves from one LISP-Site (from $LISP\text{-}Site_1$) to non-LISP-Site (to $Site_2$) and vice versa (to $LISP\text{-}Site_4$). Its mechanism combines Sec. 2.4.3 and Sec. 2.4.4, and the packet flow is shown in Fig. 2.11. At first, when the LISP-MN resides in $LISP\text{-}Site_1$, it uses the allocated EID as its $LRLOC$ and exchanges the packets with CN via xTR_1 by double encapsulation. When it moves to the $Site_2$, it receives the assigned IP address from $Router_2$ as its new $LRLOC$. LISP-MN registers its new mapping information $< EID_{MN}, LRLOC_2 >$ to the mapping system and informs the xTR_3 (xTR of CN) to update the mapping in its cache by SMR or $Map\text{-}Versioning$. The $Router_2$ just natively forwards the packets to xTR_3. When the LISP-MN roams behind xTR_4, it uses the allocated EID from xTR_4 as the latest $LRLOC$ and register this mapping information $< EID_{MN}, LRLOC_4 >$ to the mapping system. Also, it tells xTR_3 ($RLOC$ of CN) to update its mapping by using SMR or $Map\text{-}Versioning$. Once the handover procedure is completed, the packets exchanged

between LISP-MN and *CN* will be double encapsulated: once on the LISP-MN itself and once on the xTR_4.

2.5 LISP alternative use-cases

Even though LISP is proposed to solve the Internet scalability issues, it also provides the advantages to the other uses, such as: traffic engineering, transition between IPv4 and IPv6, device's mobility at client-side, virtual machine's mobility in data center, LISP SDN, and so on [40].

2.5.1 LISP traffic engineering

For every mapping information, an EID-prefix associates to a list of RLOC tuples <RLOC, Priority, Weight>, where each RLOC is annotated with a priority and a weight. In the case of multiple RLOCs existing, the ITR selects the one with the highest priority and sends all the LISP encapsulated packet to this RLOC. If several RLOCs all have the highest priority, the traffic is balanced proportionally according to their weight among such RLOCs so to route to multiple ETRs. Traffic engineering in LISP thus helps the mapping owner decide the primary and backup path for its incoming packets, and also allows it to control the traffic load among its links. For example, the received Map-Reply shows that an EID-prefix is associated to <RLOC_1, 10, 50>, <RLOC_2, 10, 20>, <RLOC_3, 10, 20>, <RLOC_4, 10, 10>, then the RLOC_1 gets 50% of the traffic, RLOC_2 and RLOC_3 are respectively assigned to 20% of the traffic, and the last 10% of the traffic are balanced by RLOC_4.

Traffic engineering in LISP can achieve one step further than nowadays BPG. Since every Map-Request contains the source EID of the packet, for which causes the ITR mapping miss and triggers the Map-Request. It is possible that a mapping owner (ETR) gives the different Map-Replies according to the sender of Map-Requests. Thus, the traffic engineering policy may be different to different ITRs although they query the same ETR for the same EID-prefix. This functionality is not available today with BGP, because a domain cannot control exactly the incoming routes from the other domains if they are not direct neighbor.

2.5.2 LISP transition between IPv4 and IPv6

The LISP encapsulation mechanism is designed to support any combination of address families for locators and identifiers, such as: IPv4, IPv6, MAC address, or some other arbitrary elements. It is then possible to bind IPv6 EIDs with IPv4 RLOCs and vice versa. This allows transporting IPv6 packets over an IPv4 network (or IPv4 packets over an IPv6 network)

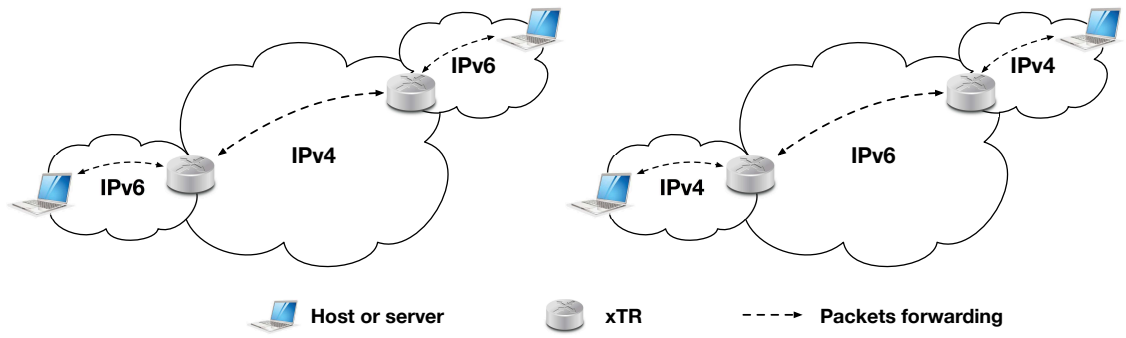


Fig. 2.12 Illustration of LISP transmission between IPv4 and IPv6

by deploying the xTRs at the border of heterogeneous networks, making LISP a valuable mechanism to ease the transition to IPv6.

Examples are as shown in Fig. 2.12, where IPv6 islands are connected via IPv4 infrastructure (on the left) and IPv4 islands are connected through IPv6 infrastructure (on the right). As LISP provides mechanisms for encapsulating IPv6 host packets using IPv4 RLOCs on the xTRs, LISP can then provide IPv6 communication between hosts/servers even though the intermediate network only supports IPv4. Similarly, xTR can encapsulate the IPv4 packets by adding the IPv6 RLOCs to achieve the IPv4 access through the IPv6 infrastructure. In addition, when the LISP infrastructure includes PxTRs, it can provide both IPv4 and IPv6 connectivity to non-LISP IPv4/IPv6 sites (legacy Internet). In summary, it can function in any of four following ways:

- IPv4-to-IPv4 over IPv4
- IPv4-to-IPv4 over IPv6
- IPv6-to-IPv6 over IPv4
- IPv6-to-IPv6 over IPv6

However, LISP does not translate between the IPv4 and IPv6 EIDs.

2.5.3 LISP SDN

As described before, LISP is inherently decoupling Control Plane from Data Plane. LISP moves all control functions onto Mapping System, while keeps data functions at the xTR level.The feature of LISP corresponds to the characteristic of Software Defined Network (SDN), which also decouples Network Control Plane from Data Plane, increases flexibility and development speed of features and functionalists. Thus, this decoupling entitles network operators to build a Software Defined Network (SDN) on top of LISP [106].

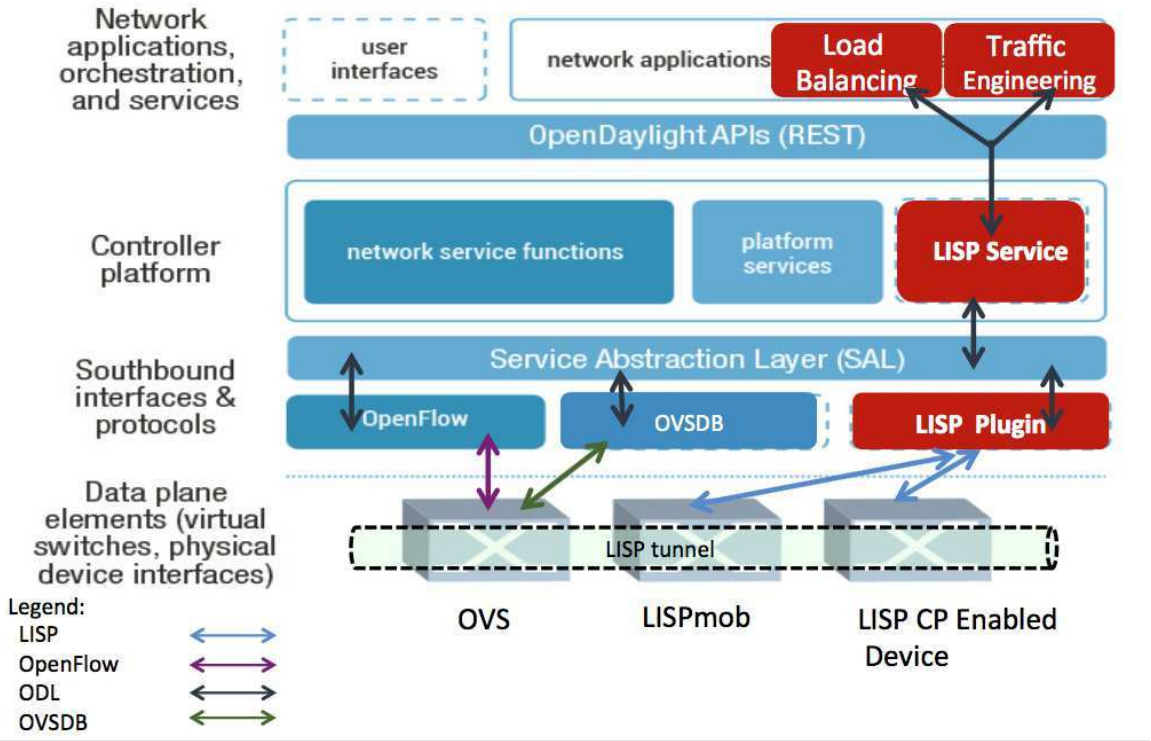


Fig. 2.13 LISP in OpenDaylight (source: [13])

The OpenDaylight Project [22] is a collaborative open source project hosted by The Linux Foundation to promote Software Defined Network (SDN) and Network Functions Virtualization (NFV). It implements the LISP Control Plane functions as shown in Fig 2.13. The data plane elements consist of virtual switches and physical device interfaces, at the aim of constructing the LISP tunnel to transfer the packets. The LISP Plugin is the Southbound interface, which is used to look up LISP policies from LISP Service by sending Map-Request and receiving Map-Reply, and create the LISP tunnel over LISP overlay across the different kinds of data plane devices. The LISP Service is actually implemented as LISP Mapping System to store the various user-defined policies. The only difference is that in OpenDaylight, it is the MS itself that replies to the Map-Requests, instead of forwarding them to any xTR. It is like every mapping has the "proxy Map-Reply" flag (P-bit) set to response on behalf of xTRs. The OpenDaylight APIs (REST) are used to set the policies, such as: the policies for Load Balancing and Traffic Engineering, obtained from the north bound into LISP Service. The policies are defined as the network applications or services by the users.

Network functions such as subscriber-management, content-optimization, security and quality of service, are typically delivered using proprietary hardware appliances embedded

into the network [36]. But the next generation network functions are being implemented as pure software instances running on standard servers. They are unbundled virtualized components of capacity and functionality. LISP-SDN based flow-mapping, dynamically assembles these components to whole solutions by steering the proper traffic in the right sequence to the correct virtual function instance.

Chapter 3

Related work

This chapter first presents the current LISP status, including: two available LISP platforms, the only LISP monitor, two LISP manual management tools, and some private and some open-source implementations. It also introduces the latest research achievements, from the LISP network evolution, the LISP mapping system assessment, the evaluation of interworking between LISP-site and legacy Internet, to the analysis of LISP mobility. Besides presenting the existing works, we also find some limitations in each aspect. Thus, we point out the shortcomings in the LISP monitor and implementations, as well as in the various aspects of LISP performance evaluation. We tackle most of these issues and present them in this dissertation. Further, we list some other LISP challenges and issues at the end of this chapter. Although these fields are not considered in this dissertation, they are also important to help researchers improve LISP technology.

3.1 LISP platforms

This dissertation is focusing on the evaluation of LISP performance in the real environment, thus the large scale experiments on existing platforms are of paramount importance. As such, we introduce first two existing LISP platforms: LISP Beta Network and LISP-Lab platform, before discussing the current research results.

3.1.1 LISP Beta Network

The first global LISP testbed called LISP Beta Network [11] [44], has been deployed since 2008. Started as the only LISP testbed, the LISP Beta network had a steady growth and achieved world-wide experimental deployment [108] with the purpose to gain real-life experience on LISP. Initiated by Cisco, the members of the LISP Beta Network are not only



Fig. 3.1 International LISP Beta Network MDS elements & Proxy xTRs – October 2017 [12]

major companies and operators, but also academics, research laboratories, and startups offering LISP related services. Participants of this network are located in 34 different countries, with the highest concentration in Europe and North America.

The LISP Beta Network is composed of the mapping system, LISP routers (i.e., xTRs), mobile nodes, Re-encapsulating Tunnel Routers (RTR [46]) and Proxy Tunnel Routers (PxTR [76]). On March 14th 2012, the mapping system used by the LISP Beta Network has switched from LISP+ALT [57] to the more flexible LISP-DDT [58]. Changing the mapping system has considerably reduced the configuration and maintenance burden, while a slight performance degradation has been observed [108]. According to the latest architecture (October 2017) [12] shown in Fig. 3.1, LISP Beta Network has 3 DDT Root, 2 DDT TLD, 7 available MR/MSes and 6 PxTRs. More precisely: 2 MR/MSes and 2 PxTRs are in US, 3 MR/MSes and 3 PxTRs are in Europe, 2 MR/MSes and 1 PxTR are in Asia. It mainly uses two EID addressing spaces, which belongs to Cisco: 153.16.0.0/16 for IPv4 and 2610:00D0::/32 for IPv6 [108], but there are also some other EIDs.

3.1.2 LISP-Lab platform

Due to a monolithic control by one single commercial actor, however, LISP Beta network has limitations to explore the innovative services. Thus, some other projects, such as ANR LISP-Lab project [14], focus on developing new enhanced features.

LISP-Lab project aimed at building an open platform, based on the LISP architecture, providing the environment to perform high quality research and support the design, development, and thorough assessment of new services and use-cases. The range of technical tasks planned in the LISP-Lab project, from cloud networking, to access technology, through inter-domain connectivity, traffic engineering, and mapping management, boosting innovation beyond the LISP technology itself. LISP-Lab platform was solely deployed by using open source LISP implementation, OpenLISP [23]. It was coordinated by a French consortium: two academic institutions (UPMC, TPT), two Cloud Networking SME (Alphalink, NSS), two network operators (Renater, Orange), two SMEs on Access/Edge Networking (Border 6, Ucopia) and one Internet eXchange Point (Rezopole). Ten international external partners also joined the platform. At the moment of writing, LISP-Lab consists of 3 MRs/MSes and 2 PxTRs located in France, as well as 13 worldwide xTRs, and has already been interconnected to the LISP Beta Network with an OpenLISP DDT root [59].

3.2 LISP monitoring tools

As the researchers need to leverage on LISP tools to conduct the experiments on the aforementioned LISP platforms, this section introduces the currently existing tools to supervise LISP status or troubleshooting it for some specific purposes.

3.2.1 LISPmon

LISPmon is the only monitoring platform that supervises the current public LISP status, i.e., the mappings between the EID-prefixes and the RLOCs. This platform is developed by the Advanced Broadband Communications Center of UPC (Universitat Politècnica de Catalunya). It scans the whole IPv4 addressing space everyday, normally beginning at 7:00 a.m. (UTC), sending the Map-Request to one preferred MR of the LISP Beta network (i.e., 7 MRs are candidates to be selected). If this MR does not respond, the LISPmon operator will choose another one as a replacement, and start the queries again, to guarantee that the reception of LISP Map-Replies changes smoothly between two consecutive days. The available EID-RLOC mapping information is published once per day, ever since the beginning of 2010.

3.2.2 LIG

LISP Internet Groper (LIG) [52] is a LISP-specific manual management tool, used to obtain the mapping information for a certain EID by querying the LISP mapping system. It can be run by all devices that implement LISP, including xTR, PxTR, MR, MS, as well as by a host system at either a LISP-capable or non-LISP-capable site. A possible syntax for a LIG command is:

$$lig < destination > [source < source >][to < map-resolver >]$$

Where <destination>: is either a Fully Qualified Domain Name (FQDN) or a destination EID for a remote LISP site; source <source>: is an optional source EID to be inserted in the 'Source EID' field of the Map-Request; to <map-resolver>: is an optional FQDN or RLOC address for a MR. When LIG is launched, a Map-Request is sent for a destination IP address, and the returned Map-Reply is displayed as follows:

- The EID-prefix for the site that the queried destination EID matches.
- The locator address of the responder.
- The RLOC tuples <RLOC, Priority, Weight>.
- A round-trip-time (RTT) estimate for the Map-Request/Map-Reply exchange.

LIG is also able to indicate the Negative Map-Replies. Negative Map-Replies could be sent in the following cases: the LIG request was initiated for a non-EID address or there was rate-limiting on the responder.

3.2.3 RIG

LISP-DDT Referral Internet Groper (RIG) [48] is also a LISP-specific manual management tool, but it is not like LIG, which can be used on any MDS to retrieve RLOCs used for packet encapsulation. RIG is used to recursively find all the nodes serving an EID-prefix, specifically within a LISP-DDT mapping database hierarchy. It performs the same operation as that of a MR. More specifically, when RIG command is run, a Map-Request is sent into the DDT hierarchy for a destination EID. Then this Map-Request is recursively forwarded to the DDT nodes one by one based on the EID-prefix and delegation referrals contained in the Map-Referral messages, until the DDT MS containing the mapping information is reached. It can be run on the same devices as LIG. A simple syntax for a RIG command is:

$$rig < eid > to < ddt-node > [follow-all-referrals]$$

Where <eid> is the destination EID being queried; <ddt-node> is the RLOC of any DDT node in the DDT hierarchy; when the keyword [follow-all-referrals] is used, all the referral RLOCs will be queried; if this keyword is not used, one of the referral RLOCs will be selected to descend a branch of the DDT hierarchy. The results of returned Map-Referral messages mainly contain the RLOCs of child nodes storing the destination EID mapping information, as well as the RTT indicating the time spends for getting the referral DDT nodes.

3.2.4 Monitoring tools shortcomings

LIG and RIG are the manual management tools, they are used to troubleshoot specific issues rather than monitor the LISP platforms. LISPmon as the first and only LISP monitor website and was optimal at the initial stage, but it faces some limitations as the LISP Beta Network growths. LISPmon just publishes the mapping information once per day. All the published results are queried from one MR of the LISP Beta Network and captured only by one Vantage Point (VP). Since LISP Beta Network has multiple MRs over the world and was joined by the LISP-Lab platform later, the current monitoring mechanism of LISPmon is not complete. To move LISP forward, it is necessary to fully understand the performance of every LISP network entity. Beyond only publishing the mapping information between EID-prefix and its RLOCs, automatically conducting the comprehensive experiments from the various aspects will help the LISP experimenters improve their research. Thus, we filled this gap by proposing a new LISP monitor tool for large scale deployments, which will be described in Chapter. 5.

3.3 LISP implementations

The most widely used LISP implementation is the one on Cisco routers. However, the source codes of their implementations are not available and thus cannot be extended for LISP research purposes. In this section, we focus on other LISP implementations, which can help the researchers verify their enhancements so to move forward LISP technology. Among these implementations, several are proprietary and some are open-source.

3.3.1 LISP implementations

OpenLISP

OpenLISP [23] [102] is an open source implementation of the LISP protocol. The LISP Data Plane, such as: LISP-Cache and LISP-Database, as well as the encapsulation/decapsulation functions are implemented in the kernel of FreeBSD [7] Operating System. LISP Control Plane is written in C in the user space of FreeBSD and Linux. Everything related to the Control Plane is meant to run in the user space. OpenLISP implements a new type of sockets, called the Mapping Sockets providing an API that can be used by any users space process. OpenLISP was implemented by researchers from Université catholique de Louvain and T-Labs/TU Berlin. At the moment of writing, Pierre et Marie Curie University maintains the Lip6-lisp [10], which is an extended implementation of OpenLISP and it supports fully featured LISP Control-Plane.

Open Overlay Router

Open Overlay Router (OOR) [21], the successor of LISPmob [41], is an open source LISP and LISP mobile node implementation written in C for Linux, Android (only on rooted devices) and OpenWrt that uses LISP or VXLAN [83]/GRE [63] as an overlay for SDN. By interoperating with LISP Beta Network, LISPmob enables mobile hosts to change its network attachment point without losing connectivity, while maintaining the same IP addresses. OOR is maintained by Barcelona Tech University. It focuses on supporting network function virtualization (NFV), e.g., the integration with OpenDaylight [22]. The extended feature set makes OOR attractive for use.

PyLisp

PyLisp [26] is a small but functional implementation of LISP in Python. There is a package of PyLisp implementation on Github, called pylisp [25] providing the means to parse and create LISP data and control messages, as well as command line tools to perform actions like asking a map resolver for a mapping. The functions of PyLisp are not complete, so the intention to the later stage is the implementations for map server, map resolver and DDT node.

jLISP

jLISP [112], proposed in 2016, is an open source implementation of LISP in Java and runs in the user space. It is platform-independent and object-oriented. jLISP is compatible with

the LISP RFCs and is modular, which facilitates configuration for every LISP component, such as: xTRs, RTRs, NTRs, and the mapping-system.

LISP on OMNet++

OMNeT++ [20] is an extensible, modular, C++ simulation library and framework, primarily for building network simulators. A LISP simulator is proposed in [74], which integrates LISP into the INET framework [8] for OMNeT++. This simulator implements basic LISP functions, interworking mechanism and LISP-MN architecture. In addition, this work also support NAT traversal. Since this simulation model is proposed in 2012, the only mapping system supported is LISP-ALT instead of nowadays LISP-DDT. In addition, the simulator OMNeT++ is open source, but this LISP extension is not.

Lispers.net

Lispers.net [16] is a Python implementation aiming to implement the complete feature set of LISP and interoperate with all modern proprietary and open source implementations available today. It can run on Ubuntu, Fedora, CentOS, and Debian Linux distros as well as MacOS, Raspbian, and docker containers. The implementation supports the Data-plane components: ITR, ETR, and RTR, and the Control-plane components: MR, MS, and DDT-Node for 10 different EID address types and 7 different RLOC address types. The first version was released in November 2013. Additional non-LISP-specific functions and behaviors are also added, which make lispers.net a general overlay controller. Although the goal of Lispers.net is the world's richest feature set of LISP, it is not open source.

LISP controller in ONOS

The LISP controller in Open Network Operating System (ONOS), which is an open source SDN OS for service providers [62], is one of South Bound Interfaces (SBI) for SDN controllers. The LISP controller in ONOS is proposed to address the OpenFlow issues, such as: scalability, performance, heterogeneity, interoperability with legacy networks and protocols. An open source SDN controller development project called OpenDaylight [22], includes a sub-project, which is developing the LISP mapping system in SDN environments. The LISP controller centralizes LISP controls into a single entity. It supports LISP as an SDN SBI to work with other protocols such as OpenFlow, and provides APIs for LISP applications. The LISP controller brings many benefits, such as: simplified management, a global view of LISP networks, short information convergence time, and easy development support of LISP based applications.

3.3.2 The LISP simulation gap

At the moment, LISP has some implementations for the different Operation Systems, either proprietary or open source. However, researchers prefer to test new features on simulations rather than real testbeds, because the simulation environment simplifies the studied problem and allows researchers to concentrate on the most critical issues [90]. So, simulation has a dominant capability over the other evaluation tools, when the researchers want to quickly test new features on a larger scale for different network scenarios. Among the recent network simulators, such as ns-2 [38], ns-3 [64], OMNET++ and OPNET [42], basic LISP features have already been added to OMNET++. However, it is not open source, which poses issues for the researchers to easily and flexibly conduct the simulations. In addition, ns-3 is more popular with the researchers than the others [104] and there was an open source project[1] implementing the fundamental LISP on it in 2016. Whereas this project is based on ns-3.24 (the version at that moment), and it does not support features related to LISP mobility. Thus, we decide to leverage on this project, but first adapt to ns-3.27, and realize the extension of LISP mobility. Such contribution will be introduced in Chapter. 7.

3.4 LISP mapping system evaluation

3.4.1 Current studies

The performance of LISP Mapping System is evaluated in [108] and [44] from different aspects. As the LISP Beta Network changed its mapping system from LISP+ALT to LISP-DDT on March 14th 2012, the authors measure the delay of resolving mapping (the latency between sending the Map-Request and receiving the Map-Reply) during this transition period. They find the delay to get the mapping by using LISP+ALT is smaller and more stable than LISP-DDT. However, the manageability of the latter is better than the former, since by using LISP+ALT, maintaining the ALT causes many cumbersome configuration overhead especially as the network growing. Although using LISP-DDT slightly degrades the resolving performance, the delays remain small and the mapping system is able to cope with a significant growth of number of mappings.

3.4.2 The incomplete puzzle of MDS evaluation

The previous measurements of the MDS are mainly depended on the delay, including the different resolving latency that each MR needs to use, and comparing which VPs have longer

[1] Source code is available at [31].

delay to query the mappings. However, the authors have never taken look at the contents of mapping information, i.e., the type of mappings and locators etc.. In addition, there is no work to compare whether the consecutively received Map-Replies change over time. If yes, how they change? If there is any pattern? There are also no comparisons on whether the Map-Replies sent by the different MRs for the same EIDs are identical. Neither for the Map-Replies received by the various VPs. All these evaluations will be discussed in Chapter. 4.

3.5 LISP network evolution

After deploying LISP in LISP Beta Network from 2008, more and more participants joined to the development of LISP. Thus, Saucez et al. [108] look at how the LISP Beta network evolves. The authors observe that the number of LISP mapping was as low as 20 in January 2010, but it has consistently grown to reach 80 mappings in 2012. The number of LISP mappings is four times the initial number, which indicates that LISP network has a regular growth. Within two months of the second part of 2011, the number of negative mappings was doubled, i.e., from about 150 to about 300. The authors suppose that is due to a more fragmented EID-space. The paper also observes some malformed mappings, which belongs to the LISP mappings but have an empty RLOC set. However, this kind of mappings is never more than 0.5% of the total number of mappings within one day. Further, the paper presents that from January 2010 to May 2012, the percentage of mappings using one or two RLOCs has increased. Moreover, there exist mappings with four RLOCs in 2012, whereas they were not present in January 2010.

3.5.1 Updating LISP measurements

Previous works respectively evaluate the LISP network evolution in 2010 and 2012. However later, LISP mapping system is changed from LISP+ALT to LISP-DDT in 2012. Besides, as mentioned in Sec. 3.1.2, LISP-Lab platform is open since 2015. It brings new LISP-sites over the world and interconnects with LISP Beta Network. Further, LISP Beta Network itself also updated their xTRs and changed its network architecture in 2016. Thus, the number of LISP mappings, the fragment of EID-space, the distributions of RLOCs are also changed over time. It is interesting to have a look at them in order to know how the LISP network evolved in recent years. This work will be described in Chapter. 5.

3.6 LISP interworking with legacy Internet

LISP has significant impact on the traffic between LISP-site and non-LISP-site. First, to support the interworking between LISP-site and the legacy Internet, the PxTRs should be set up. The configuration on PxTR does not cause any special technical issue, but using PxTR potentially causes path stretch between source and destination. It makes the packet exchange experience a larger delay and potentially leads to higher packet loss rate. Thus, placing the PxTR near the source/destination of traffic allows the communication between the non-LISP site and the LISP site to have the least path stretch (i.e., the least number of forwarding hops when compared to an optimal path between the sites) [76].

When deploying the PxTRs in the experimental infrastructure, it is better to think about first where to deploy the PxTRs, since the placement in topology has a significant impact on the path stretch. Besides, as the number of PxTRs has a direct impact on the traffic load and the recovery success in case of the failure of a PxTR, we need consider how many PxTRs are needed. Further, we need to think if it is necessary that all the PxTRs advertise for the whole EID space, or it only needs that each of them just announces for the different segmented EID-prefixes.

As more and more deployments of LISP, it becomes necessary that the LISP networks communicate with the legacy Internet. The evaluation of the interworking performance is presented in reference [44]. It chooses 200 VPs on the PlanetLab (a global research network that supports the development of new network services from 2003) [24] located in North America, Asia and Europe by the order of diversity of number of VPs. It also selects 116 EIDs, which have at least 1 corresponding RLOC in the mapping from the daily of LISPmon project [18] published on May 17th 2013. The experiment is conducted from November 4th to 7th 2013. The authors find that the interworking generally has a negative impact compared to not using LISP. The results show that there are 70% of the EID-prefixes having at least 20% of increase in the delay (i.e., Round-Trip Time (RTT)). One PxTR behaves the worst, with 95% of the EID-prefixes using it suffering from a delay increase. Based on BGP, the selection of PITRs is driven by the economical relationship between autonomous systems, instead of their geographical positions. As a result, even if the LISP Beta Network has 9 available PxTRs at the moment, only 4 are used in the experiment. Further, most VPs use PxTRs in Asia although they are located in Europe or America. The most popular PxTR is selected 79% of the time.

3.6.1 PxTRs evaluation: an incomplete picture

The only evaluation of PxTR used for the interworking between LISP and legacy Internet was conducted in 2013. At that time, the unique LISP platform was the LISP Beta Network. Thus, the measurement was only based on its PxTRs and the VPs on the research network PlanetLab. After the addition of the LISP-Lab platform, there is no evaluation about the PxTR anymore. We do not know the recent performance of PxTR on LISP Beta Network and we have no knowledge about if the PxTR of LISP-Lab functions well. Further, the previous assessment was conducted within an experimental/research environment, there is no experience about the interworking performance with real sites, especially for the most popular websites in the world. In the Chapter. 6, we will explore all these issues of LISP interworking performance.

3.7 LISP mobility

To support IP mobility, the conventional LISP could be adopted. The first mobility-compatible LISP extension is LISP-MN [49][94]. LISP-MN is an end-host based IP mobility protocol which requires updating the software of the mobile terminal. In other words, the end-host that uses LISP-MN can be regarded as a small LISP-site.

The main drawbacks of LISP-MN are the double encapsulation that may happen and the need to modify software in end-host. Double encapsulation increases mapping latency, overhead and MTU issues. The modification at end-host hinders the deployment of LISP-MN. To overcome these drawbacks, many research efforts have been made by the LISP community.

The performance of LISP-MN is investigated by M. Menth and et al. [86]. The authors illustrate the performance degradation due to double encapsulation and the triangle route problem under some conditions. They also propose mechanisms such as location-aware LISP mobility node, local mapping systems to improve the performance of LISP-MN. A simple model that help analyze the handoff procedure in IP mobility is proposed by [101]. Besides LISP-MN, the proposed model can be also used for the handoff analysis for other IP mobility protocols such as MIPv6.

To avoid the modification to end-host required in LISP-MN, a network-based and LISP-based IP mobility protocol: LISP-ROAM is proposed in [60]. Their solution is that the network assigns the same IP address regardless of their network attachment point under the cooperation between some additional/modified network components (DHCP servers, authentication services, LISP xTR and servers). The authors implement LISP-ROAM on top of LISPmob and give an experimental evaluation of the performance of LISP-ROAM. As

an effort to avoid double encapsulation in LISP-MN, LISP-NEMO is proposed in [118]. The basic idea of LISP-NEMO originates from NEMO [70]. It introduces a new network element called mobile router (MR) that manages and represents the mobile network. To mitigate the problems in LISP-NEMO and LISP-ROAM, LISP-HNM is proposed in [113]. The authors introduce an access register protocol and an active notification scheme to conventional LISP so that the latter can support both end-host and network mobility.

The packet loss due to IP mobility is also a concern. To reduce the packet loss, Isah et al. [68] propose an improvement to LISP-MN. The basic idea is that those packets generated during handover procedure are buffered to severs close to the MN's new location and forwarding them to the MN on handover completion.

3.7.1 Open issue in LISP mobility

To achieve the seamless handover, LISP can be implemented on the border router or directly on the host to support the mobility. However, these proposals are only limited at the theoretical level. There are no experimental results showing which solution is better based on some metrics, such as: the smaller handover delay or the less packet loss. No evaluations respectively compare the advantages and the disadvantages of each solution. What's more, no matter which solution is adopted to conduct the mobility, SMR and Map-Versioning are the candidates to say to the remote ITR to update its mapping cache. Thus, it is interesting to make a comparison between these two mechanisms, for example modeling the overhead traffic according to the number of MNs and CNs of each mechanism. Chapter. 7 will analyze those scenarios.

3.8 Troubleshooting issues

A big issue in the LISP experimentation is the difficulty of troubleshooting. The most frequent used troubleshooting tool is *traceroute*, but it can not provide the exact path that the packets pass between the xTRs or PxTRs. The experimenters only know how many hops between them. When the host traceroutes a destination, it uses each EID as the source/destination IP address. However, when the packets pass to the ITR, it encapsulates the packets, so the following routers think that the ITR is the source of the packet and will give back the ICMP message to the ITR. In this way, the host, which is the real source receives no response. Until the packets reach to ETR or PETR, they decapsulate the packets to know the real source from the inner header and answer to the EID of host.

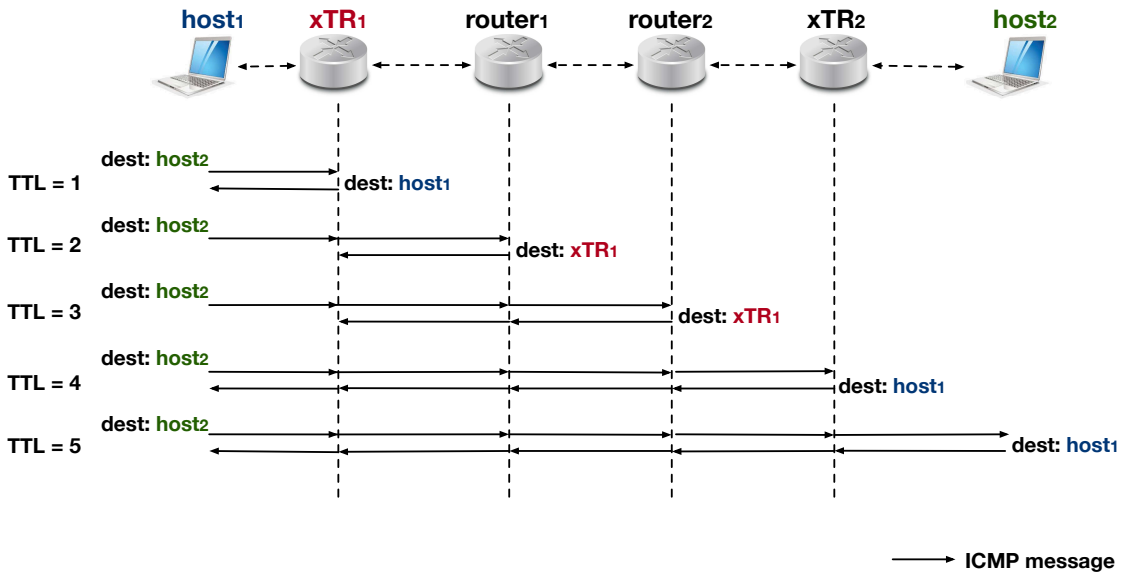


Fig. 3.2 Illustration of LISP troubleshooting

An example of LISP troubleshooting is shown in Fig. 3.2. When $host_1$ traceroutes $host_2$, $host_1$ puts its own EID and the one of $host_2$ respectively as source and destination address. The first UDP message with TTL (time-to-live)=1 expires at xTR_1, then it sends back the ICMP Time Exceeded message to $host_1$. Thus, the $host_1$ knows that its next hop is xTR_1. However, when xTR_1 receives the UDP message with TTL=2, it decreases the TTL value, and encapsulates the UDP message with its own RLOC as source address and forwards to the next hop $router_1$. Since the UDP message that the $router_1$ receives showing xTR_1 is the source instead of $host_1$, the $router_1$ drops the packets and replies back to xTR_1. In this way, the $host_1$ cannot receive any ICMP Time Exceeded messages until the xTR_2 is tracerouted. When the xTR_2 receives the UDP message, it decapsulates the packets and finds that the real traceroute sender is $host_1$. Thus the xTR_2 replies to $host_1$. Same to $host_2$, the received UDP message shows that $host_1$ is the traceroute origin and itself is the destination, so it returns an ICMP Port Unreachable message $host_1$.

To the conventional Internet topology, *traceroute* uses the returned ICMP Time Exceeded messages to build a list of routers that packets traverse until the destination is reached. However, the list of routers between xTRs is hidden to the traceroute sender. It can only know how many hops it needs to arrive the ETR. Thus, when there is a problem in the LISP network, it is hard for the operators to pinpoint the reason as they have only a partial view of the network. The experimenters can leverage on the LISP Cache/Database or querying the MRs to observe the potential problem in the network, but the knowledge remains limited.

This issue brings inconveniences to our experiment, which will be discussed in Chapter 6. In that experiment, we can only find how many hops there are between our VPs and the PxTRs, but are not able to specify the exact routing paths.

3.9 Other LISP related studies

LISP is relatively young and there are many other aspects needed to be taken into consideration, such as: the caching, the LISP security, the MTU issue, and so on. As they are important and necessary to LISP, we list them here. However, since they are not related to this dissertation, they will not be discussed in the following chapters.

3.9.1 LISP cache performance

A thorough analysis of LISP Cache is presented in [66], [72] and [73]. Their work complement and validate each other. The papers show that although the cache timeout can be reduced as small as 60s, the LISP cache maintains a very high efficiency, with more than 99% hit-ratio. As a result, the traffic volume overhead introduced by both the traffic encapsulation and the control traffic to request missing mappings is limited to 6%. Around 70% of cache miss is generated by UDP protocol regardless the port number. The authors argue that is mainly caused by P2P applications or other distributed applications generating the same spread traffic pattern. They also compare the original LISP specification, called vanilla LISP with the symmetric LISP model, proposed by Saucez et al. [72], and show that the later model provides higher security. Besides, they evaluate the relationship between the size of cache and the number of users. It shows that when the number of users is increased to three times, the size of cache only doubles. It means that the average size of the LISP Cache has not a linear relationship with the number of hosts.

As querying the mapping information to the MDS largely increases packet transmission delay, to store the necessary mappings in the LISP Cache can significantly alleviate it. However, as limited capacity of LISP Cache and dynamic changes of mapping information, it still requires smart algorithms to determine which mappings should be reserved or removed for better performance, especially when there are many expired mappings caused by mobility [53].

3.9.2 LISP security

As LISP is born to address the scalability issues of Internet, it is not designed with security in mind. The possible security threats that LISP faces are given in [109]. Although after

the authentication, the MSes only accept the registrations from trusted ETRs, the mapping system can still be attacked. For example, compromised ETRs may de-aggregate their EID-prefixes in order to register more EID-prefixes than necessary to their MSes, and this causes the ITRs to send overwhelmed Map-Requests to the mapping system. The compromised ETRs may also overclaim the prefix they own in order to influence the route followed by Map-Requests for EIDs outside the scope of their legitimate EID-prefix. Besides, the Map-Register stage is vulnerable to masquerading and content poisoning attacks, for example: the ETRs register the fake mapping information to the MSes [33] [91] [32]. By sending the Map-Requests at a high rate can overload the nodes in the MDS and cause the flood of Map-Replies as well. Some bits in the LISP control plane messages also have some security issues, for example: the Locator-Status-Bits can cause a DoS (Denial-of-Service), a replay, a packet manipulation, or a packet interception and suppression attack; the Map-Version number can mount a DoS, an amplification, or a spoofing attack; the routing locator reachability is able to mount a packet manipulation and an amplification attack.

A set of security mechanisms (LISP-SEC) that provides authentication, integrity and anti-replay protection for LISP are introduced by Maino et al. [84]. A number of vulnerabilities should be considered before deploying LISP in large scale [103].

The paper of Aiash et al. [32] analyses the security and functionality of the LISP mapping procedure using a formal methods approach based on Casper/FDR tool. The authors point out several security issues in the protocol such as the lack of data confidentiality and mutual authentication. The reference [33] figures out that the Map-Registration stage is vulnerable to masquerading and content poisoning attacks. Therefore, it presents a new security method for protecting the LISP Registration stage. The proposed method uses the ID-Based Cryptography (IBC), which allows the MDS to authenticate the source of the data.

3.9.3 The Maximum Transmission Unit (MTU) issue

LISP encapsulation adds 36 bytes for IPv4 (i.e., 20 bytes of outer header + 8 bytes of UDP header + 8 bytes of LISP header) and 56 bytes for IPv6 (i.e., 40 bytes of outer header + 8 bytes of UDP header + 8 bytes of LISP header) to the original packet size. As with other protocols that encapsulate or tunnel traffic (e.g., IP Security (IPsec) [114] and generic routing encapsulation (GRE) [51]), it is possible that the resultant encapsulated packet would exceed the MTU somewhere along the transit path. LISP provides both stateful and stateless mechanisms for handling potential MTU issues [6], with the main goals being (i) preventing packets from being dropped, and (ii) preventing the need for ETRs to perform packet reassembly prior to LISP de-encapsulation.

Based on real Internet experiments, [66] and [73] show that although the encapsulation overhead is not negligible, it is very limited (in the order of few percentage points of the total traffic volume). In addition, no major issue due to the MTU is observed on the LISP Beta Network.

Chapter 4

Evaluation of LISP mapping system

As presented in Chapter 2, to retrieve the routing information (RLOCs) for EIDs, LISP relies on a pull model by explicitly querying the Mapping Distribution System (MDS). The performance of the mapping system is of paramount importance for LISP. Thus, we leverage on *LISP Beta Network* [11], to conduct measurement campaigns about the mapping system. Such a worldwide real playground has been used to assess the state of LISP deployment and evolution over the years [108]. Yet, to our best knowledge, nobody at the moment has evaluated the mapping system from the following two standpoints: Stability and Consistency. We thus propose a much deeper study and a more thorough evaluation, focusing on the exploration of all the EID space through active LIG (LISP Internet Groper [52]) measurement. The temporal comparison shows that the mapping system is quite stable over time and a crosswise comparison indicates that the mapping system is consistent between network entities. However, some instabilities and inconsistencies are also identified through the evaluation, which indicates that the mapping system can be improved. Besides proposing a brand new taxonomy in order to classify them, we carry out an in-depth investigation of such cases to provide hints on how to improve the performance of LISP.

The remainder of this chapter is organized as follows. Sec. 4.1 presents an overview of the experiment campaigns and introduces the metrics that we use to evaluate the performance of the mapping system. Sec. 4.2 and Sec. 4.3 respectively describe the mapping system stability and consistency results in details, as well as the instability and inconsistency cases by using the proposed taxonomy, and investigate in depth the causes from different viewpoints. Finally, Sec. 4.4 provides some findings and concludes this chapter.

4.1 Methodology

In order to collect data to carry out our analysis, several measurements campaigns have been performed. Hereafter we provide an overview on the measurements and the metrics used in our evaluation.

4.1.1 Measurements overview

To obtain a comprehensive and in-depth evaluation of LISP Beta Network Performance, we conducted experiments using LIG (LISP Internet Groper [52], see Sec. 3.2.2 for the details). When LIG is launched, a *Map-Request* is sent for a destination IP address, and the returned *Map-Reply* is then stored for further analysis. It exists three possible types of outcome: (i) *LISP Map-Reply*, where the reply contains mapping information to reach the LISP-enabled sites; (ii) *Negative Map-Reply*, where the reply states that the requested IP address belongs to a prefix of a non-LISP site; (iii) *No Map-Reply*, where no answer is received within some amount of time. A query is considered to be successful if either a LISP Map-Reply or a Negative Map-Reply is received. While a query is considered to be failed if no reply is received within 5 seconds.

The experiments are conducted from July 2nd to 18th 2013. From a set of different vantage points (VPs), sending *Map-Request* messages to all the Map-Resolvers of the LISP Beta Network, for all selected destination IP addresses every 30 minutes. The set of VPs was composed by part of academic networks, commercial Internet, PlanetLab and spread across Europe and USA. The destination IP addresses consisted of the first address within each prefix derived from the LISPmon project [18] regardless of the type (i.e., the EID-prefixes or regular prefixes). During this measurement campaign, 5 VPs and 13 MRs/MSes were available and 613 prefixes were queried at each round.

We analyzed the Map-Replies associated with EIDs instead of all regular IP addresses. In the remainder, we use the term *experiment round* to identify a complete *Map-Request*/*Map-Reply* exchange by LIG for all the selected destinations to all the available MRs over all existing VPs. Although No Map-Reply occasionally appears in some experiment rounds, we just focus on LISP Map-Reply and Negative Map-Reply to conduct the analysis. [1]

[1] We use *"simultaneously"* or *"at the same time"* as synonym of *"at a same experiment round"*.

Table 4.1 Map-Reply Differences

Category	Mismatch
Map-Reply Type	Negative Map-Reply and LISP Map-Reply
RLOC set	RLOC number
	RLOC address
	RLOC priority
	RLOC weight
	RLOC state

4.1.2 Metrics

As indicated in Sec. 2.1.2, the LISP Control plane is in charge of offering the mapping information that consists of an EID-prefix associated with a list of $<RLOC, Priority, Weight>$ tuples. Theoretically, unless configuration has changed, if the same EID-prefix is queried, Map-Replies should always contain the same mapping information, no matter from which VP the query is sent, and no matter which MR is queried. During our measurements, however, we found that the Map-Replies are not always identical as they are supposed to be. The reasons causing the Map-Reply changes are summarized in Table 4.1.

The Map-Reply differences can be classified into two broad categories: the first is a difference in terms of *Map-Reply Type*, which means that a mix of Negative Map-Reply and LISP Map-Reply occurs; the other type of difference is in the *RLOC set*, which means that at least one of the attributes associated with RLOC set is different. More specifically, *RLOC number* indicates a difference in the number of RLOC tuples received in the Map-Reply. The others are differences in a RLOC tuple, which consists of RLOC address, RLOC priority, and RLOC weight (presented in Sec. 2.1.2).[2] The RLOC state is derived from the *R-bit* in Map-Reply message, which is set when the sender of a Map-Reply has a route to the RLOC, meaning that from its viewpoint the RLOC is up and running. If at least one of the six parameters shown in Table 4.1 conveyed in a Map-Reply is different from the previous one (over time) or the others (between different VPs or MRs) at the same experiment round, the Map-Reply is marked as different. Based on the criteria of the Map-Reply difference, we propose two metrics to evaluate in more details the performance of the LISP mapping system: (i) *Stability* and (ii) *Consistency*.

A mapping is considered to exhibit *stability* if the Map-Replies for one specific EID obtained from one MR observed at a given VP (briefly denoted as a tuple $<EID, MR, VP>$), over time, are identical. A classification taxonomy for unstable cases is proposed later in Sec. 4.2.

[2]The terms *RLOC address* and *RLOC* are used interchangeably.

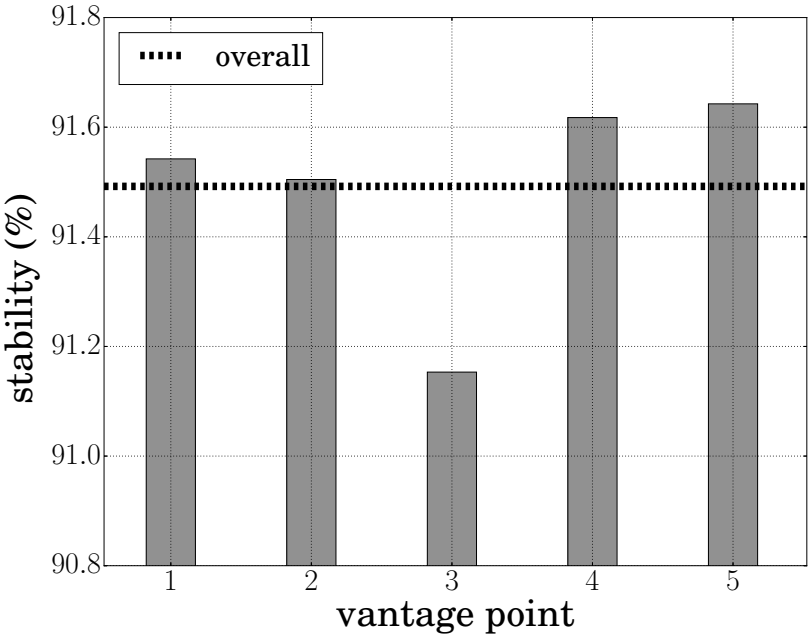


Fig. 4.1 Overall percentage of stability observed at five different VPs.

A mapping exhibits *consistency* if: *i*) Map-Replies for a specific EID from all MRs observed at a given VP (denoted as a group $<EID, *, VP>$) are identical at the same time (this kind of consistency is referred as *Consistency by MR*); *ii*) Map-Replies for a specific EID-MR pair in all VPs (denoted as a pair $<EID, MR, *>$) are identical at the same time (referred as *Consistency by VP*). A classification taxonomy for non-consistent cases is proposed later in Sec. 4.3.

4.2 Mapping system stability evaluation

In this section, we study the stability of the mapping system over time. More precisely, we consider that an EID is *stable* if all mappings retrieved for this EID are identical. If it is not the case, we say there is *instability*.

Overall, when we consider the whole dataset, 91.49% of observations are stable, which indicates that over the full set of observations, changes are infrequent and the MDS is stable. When the stability is considered on a per vantage point basis, we observe, as shown in Fig. 4.1, that the stability is independent of the vantage point since no significant deviation from the overall average can be observed (dashed line in Fig. 4.1). Nevertheless, while these results demonstrate that the mapping system is generally stable, in the following of this section, we investigate the reason why about 8.51% of the observations exhibit instability.

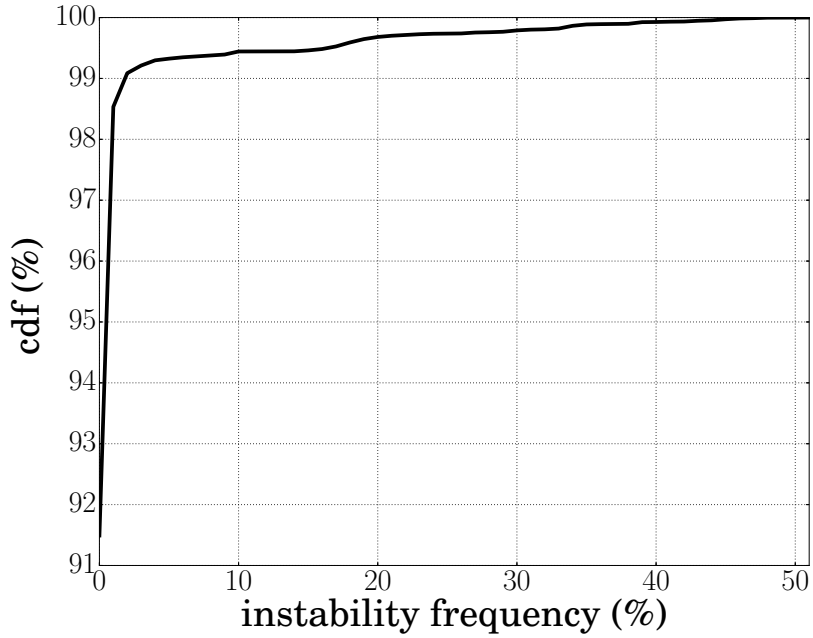


Fig. 4.2 CDF of instability frequency. X-axis is instability frequency, indicating the ratio of instability occurrence number for one <EID, MR, VP> tuple over total experiment number.

Fig. 4.2 complements Fig. 4.1 by showing, for each tuple <EID, MR, VP>, the frequency $f_{<EID,MR,VP>}$ at which mappings are different, by using the following definition:

$$f_{<EID,MR,VP>} = \frac{\text{\# Instabilities}}{\text{\# Experiments}} \times 100\%. \tag{4.1}$$

Thus, if the instability frequency $f_{<EID,MR,VP>}$ is 0%, it means that such a tuple is always stable during the whole experiment, i.e., its Map-Replies are always the same. On the contrary, a tuple <EID, MR, VP> with instability frequency of 100% indicates that its current Map-Reply is never identical to the previous ones, i.e., its Map-Replies change at every experiment round.

Fig. 4.2 shows the Cumulative Distribution Function (CDF) of the mapping change frequency (i.e., instability frequency). It shows the probability where the instability frequency equals to zero is 91.49%, which is coherent with the result in Fig. 4.1 (dashed line) and means that the mapping is globally stable. Apart from the stability case, the others ($100.0\% - 91.49\%$) belong to the mapping changes case, i.e., instability. This CDF also shows when the instability frequency equals to 1%, the CDF reaches to 98.53%, which indicates that among the instability cases, 82.77% ($\frac{98.53-91.49}{100.0-91.49}\%$) of all mapping changes events have an instability frequency less than 1%. We can deduce that mapping changes, if they occur, are rare events. The existence of instability frequency around 50% means that, in

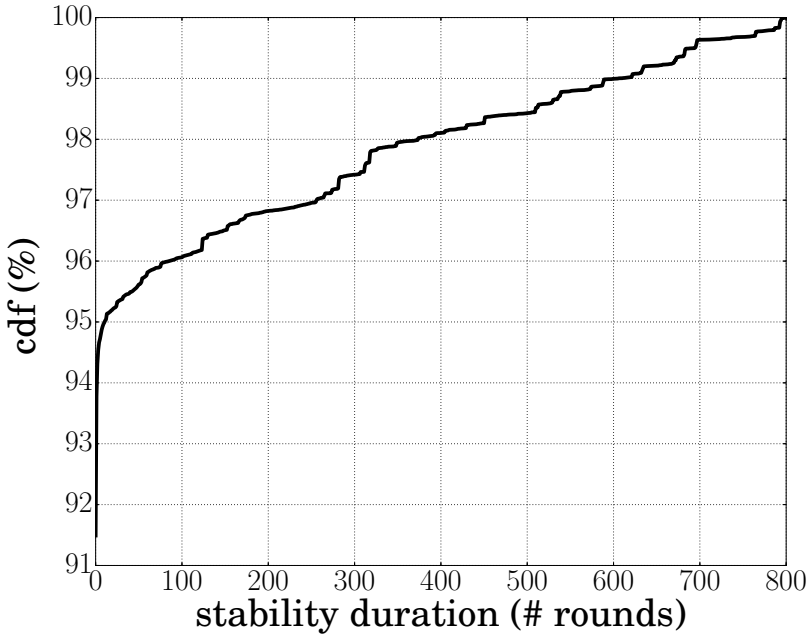


Fig. 4.3 CDF of stability duration (in number of experimental rounds). X-axis is stability duration, indicating during how many experiment rounds that the mapping for one <EID, MR, VP> tuple is stable.

some cases, even though it is rather rare, the mapping changes occurs almost one time every two experiment rounds.

While Fig. 4.2 indicates that mappings can potentially change frequently, Fig. 4.3 shows the CDF of the stability duration, i.e., for how many experiment rounds a mapping remains stable. We observe that the majority of instabilities has a stability duration of 1 experiment round. This fact shows that the mapping change occurs during a measurement period (i.e., 30 minutes, the interval between two experiment rounds). At the other side of the curve, stability lasts up to 800 experiment rounds are observed, indicating that the mapping changed at the very beginning of or the end of the measurements we performed. In the dataset we identified both cases. To further understand the instability, we classify it into four categories:

- *New Deployment*: the Map-Replies of a tuple <EID, MR, VP> mix two different types: Negative Map-Reply and LISP Map-Reply. Since the changes between the different types of Map-Reply are mainly caused by the deployments of new LISP-Sites, we make this case named New Deployment.

- *Statistical outliers*: the Map-Replies of a tuple <EID, MR, VP> change more frequently than usually seen. In our measurements, outliers correspond to mappings that change at least 3 times during a day.

Table 4.2 Percentage of Instabilities by Category

New Deployment	Statistical outliers	Mobility	Reconfiguration
72.36%	7.23%	0%	20.41%

- *Mobility*: the Map-Replies of a tuple <*EID*,*MR*,*VP*> are explicitly tagged with the mobility bit as specified in RFC6830 [47].
- *Reconfiguration*: all cases that fit neither into the New Deployment case nor the Statistical outliers case.

Tab. 4.2 presents the percentage of observations for the four categories. New Deployments dominate the total unstable cases with a percentage of 72.36%, followed by Reconfiguration with 20.41%, and Statistical outliers with 7.23%. Throughout the dataset, we did not observe any Mobility event, probably because the mobility specifications were not yet implemented at the time of measurements. Therefore, we omit the Mobility category in the rest of this chapter.

New Deployments are the most common instabilities and we used spectral analysis with Fourier transforms and auto-correlation to understand if such events are correlated between themselves. This analysis shows that New Deployments are equally spread over the whole measurement period, i.e., their occurrence per day is relatively constant. We performed the same study on the Reconfiguration category and Statistical outliers and reached the same conclusion, i.e., the absence of correlations between events.

Since New Deployments are incidental events that can happen at any time, no particular duration is observed for this category of instabilities. Moreover, as New Deployments account for 72.36% of instabilities, it significantly impacts the duration distributions shown in Fig. 4.3, explaining the relatively smooth evolution of the CDF. On the contrary, Statistical outliers are characterized by very frequent changes and they bias the distribution around the lowest duration values, explaining the relatively high number of observations for small duration.

To characterize the mapping changes happening during instabilities, we derived the following *dissimilarity* metric $d(m_i,m_j) = 1 - \frac{|m_i \cap m_j|}{|m_i \cup m_j|}$, based on the Jaccard similarity coefficient, where m_x is the RLOC set of mapping x. A dissimilarity of 1 indicates that no RLOCs are shared between the two compared mappings, while a dissimilarity of 0 indicates that the RLOC sets are the same. Fig. 4.4 shows the CDF of the mappings' dissimilarity for each unstable <*EID*,*MR*,*VP*> tuple. In our experiment, the dissimilarity results are discrete (as the discrete points shown in the Fig. 4.4) in a range from 0.33 to 1.0. As one can expect, the most likely dissimilarity is 1 in 90.09% of the total number of instabilities. These ob-

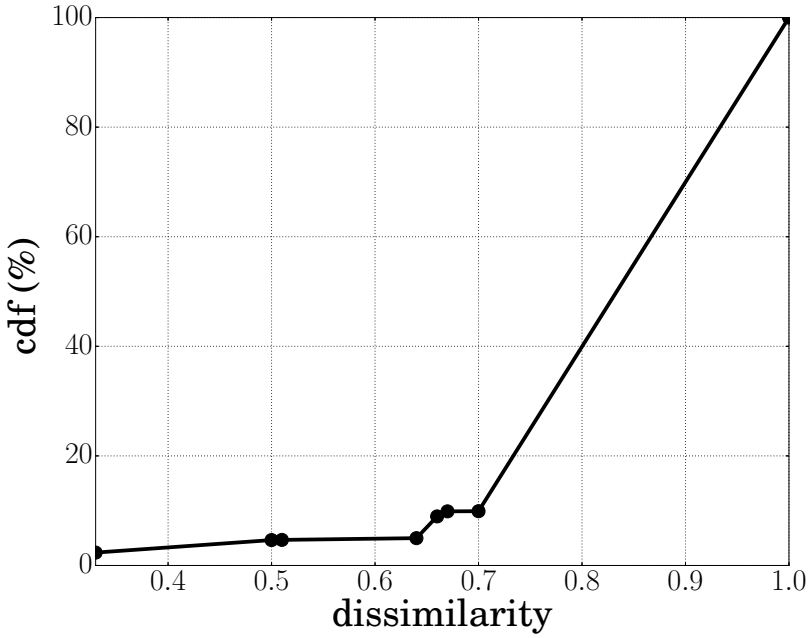


Fig. 4.4 CDF of the dissimilarity among different types of instabilities.

servations mostly come from New deployment case where before turning on the xTR, no RLOC is provided with the Negative Map-Reply (RLOC set is empty), hence, switch to a LISP Map-Reply implies a full change in the RLOC set. As 90.09% is larger than the 72.36% of New deployments share, it indicates as well that complete change of the RLOC set have also been observed during Reconfiguration and Statistical outliers. For the rest, no particular trend can be observed showing that changes of RLOC set can actually take any form.

4.3 Mapping system consistency evaluation

In this section, we discuss the consistency of the mapping system respectively by MR and by VP. More precisely, we consider that a tuple of $<EID, *, VP>$ exhibits consistency, if all the Map-Replies from different MRs, for a certain EID, observed at a given VP are identical at the same time (same experiment round). This kind of consistency is denoted as *consistency by MR*. Similarly, if all the Map-Replies sent by a specific MR, for a certain EID, and observed at different VPs are identical at every experiment round, then we consider that a tuple of $<EID, MR, *>$ is consistent, and denote it as *consistency by VP*. Otherwise, for cases not falling in the above two, we simply say that there exists *inconsistency*.

Overall, throughout the whole dataset, we found that 86.3% of observations of all $<EID, *, VP>$ tuples are consistent, which indicates that non-identical Map-Replies from different MRs are uncommon. Similarly, consistency by VP is observed in 90.48% of observations of all $<EID, MR, *>$ tuples, showing that the Map-Replies received at different VPs are usually identical. In the following, Sec. 4.3.1 will go deeper in the analysis of consistency by MR, proposing as well several types of inconsistency; then Sec. 4.3.2 will provide thorough analysis of consistency by VP.

4.3.1 Consistency evaluation by MR

The overall percentage of consistency by MR is 86.3% (see the dashed line on Fig. 4.5), meaning that across the whole experiment the inconsistency among the Map-Replies sent by different MRs is not common. Yet, we can also observe from Fig. 4.5 that different VPs show some deviations from the overall consistency by MR, reflecting the fact that while certain VPs receive the same Map-Reply from different MRs, some other VPs receive inconsistent responses. Such deviation is very small (no more than 2.77%) among different vantage points, indicating that it remains a relatively rare event. These results show that the mapping system is consistent among different MRs in general, but the inconsistency exists at the same time with a percentage around 13.7%. In the following of this section, we go deeper in this issue, exploring the reasons why the few observations exhibit inconsistency by MR.

We compute the CDF of the inconsistency frequency by MR: $f_{<EID,*,VP>}$, i.e., the probability of inconsistency frequency over all $<EID, *, VP>$ tuples, by using the following definition:

$$f_{<EID,*,VP>} = \frac{\text{\# Inconsistencies by MR}}{\text{\# Experiments}} \times 100\%. \tag{4.2}$$

For a $<EID, *, VP>$ tuple, an inconsistency frequency of 0% means that all Map-Replies from all of the different MRs for one EID are the same at each experiment round throughout the whole measurement campaign. For a $<EID, *, VP>$ tuple, an inconsistency frequency of 100% means that at each experiment round during the whole measurement campaign, there is at least one MR replying a Map-Reply different from the replied by other MRs. The resulting CDF of inconsistency frequency is shown on Fig. 4.6. The value 0% of inconsistency frequency has a probability of 82.38%, indicating that through all the dataset, there are 82.38% of $<EID, *, VP>$ tuples that are always consistent by MR at whichever VP. Yet, it is worth noticing that this value is lower than the overall consistency of 86.3% shown in Fig. 4.5. The reason of such difference lays in the fact that, when we analyze the 0% inconsistency frequency, we just take into account those common EIDs exhibiting a consistency

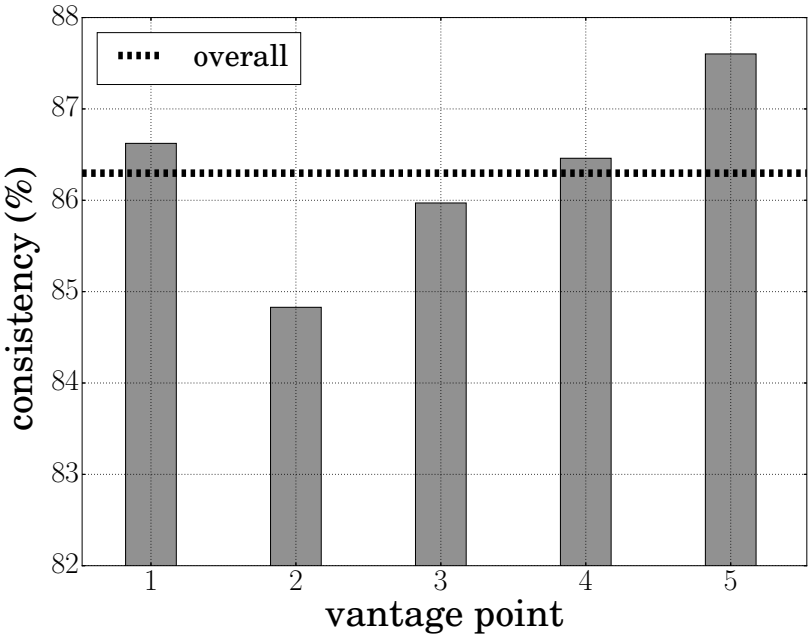


Fig. 4.5 Overall consistency by MR at 5 different Vantage Points.

by MR in all VPs. Hence, 82.38% represents the subset of the EIDs considered in the former percentage and is the lower bound, guaranteeing that 82.38% of <$EID, *, VP$> tuples always have consistent responses, regardless of the VP. In addition, the inconsistency frequency concentrates at a very low range, since a sharp increase occurs between 0% and 2% with a growth of 16.15%. It indicates that the <$EID, *, VP$> tuples classed as inconsistent do actually receive inconsistent responses rarely, hence, demonstrating that the mapping system is generally consistent. Yet, in very few cases (only 0.5%), the inconsistency between different MRs occurs all the time.

So far, the analysis indicates that the mapping system is consistent by MR in general, but the inconsistency indeed exists. To have a closer look to the inconsistency, we further cluster it into two cases:

- **Map-Reply Type**:, The type of Map-Reply provided by the 13 MRs for a specific <$EID, *, VP$> tuple is different, meaning that some MRs provide LISP Map-Reply while others return a Negative Map-Reply. The difference of received Map-Reply type leads the xTRs to use a different way to send the packets, i.e., xTRs receiving a LISP Map-Reply sends LISP-encapsulated packets by using the obtained RLOCs, while xTRs receiving a Negative Map-Reply will forward the conventional packets (i.e., without LISP-encapsulation).
- **RLOC set**: All the Map-Replies returned by the 13 MRs for a <$EID, *, VP$> tuple belong to LISP Map-Reply type, but the associated attributes (cf. Tab. 4.1 from Sec. 4.1)

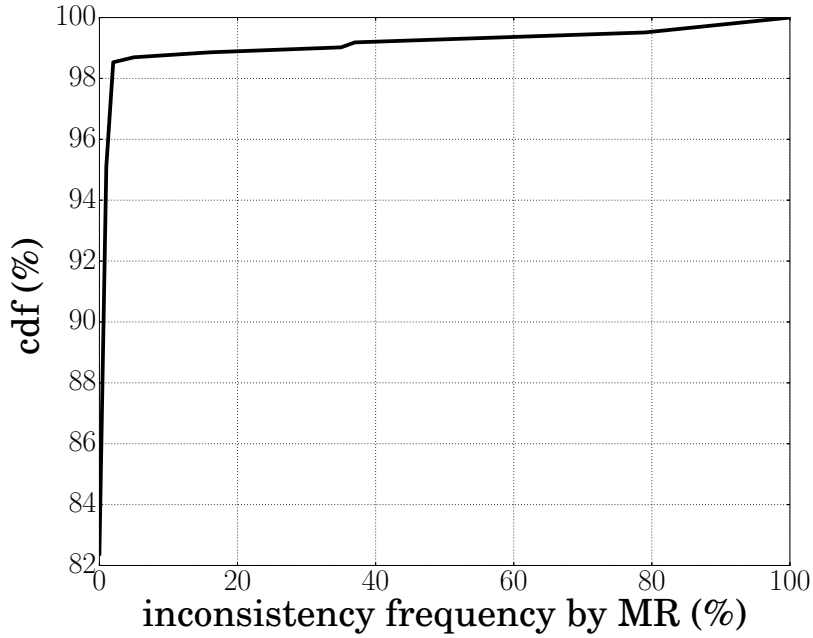


Fig. 4.6 Cumulative Distribution Function of inconsistency frequency by MR. X-axis is inconsistency frequency by VP, indicating the ratio of inconsistency occurrence for each $<EID, *, VP>$ tuple over total experiment number.

are different. For instance, some Map-Replies may contain $RLOC_1$, while some others convey $RLOC_2$ (i.e., the RLOC address is not identical); or some Map-Replies have only one RLOC, whereas others have multiple RLOCs (i.e., the RLOC number is different), and so on. The difference in such Map-Replies could lead the requesting xTRs to select different destination RLOCs.

The large majority of inconsistencies, around 98.34%, are Map-Reply Type, probably because there is an update of mapping information in the mapping system for the New Deployment case. Not all of the 13 MRs simultaneously update the mapping information, meaning that a convergence time is needed so that the global mapping information maintained in all parts of the mapping system is synchronized. If VPs query the mapping system during this convergence period, some MRs will still provide Negative Map-Replies, while other (already updated) MRs will forward the Map-Requests towards the corresponding ETRs, which in turn will send the updated LISP Map-Replies, and vice versa. This situation contributes to the inconsistency frequency around 1%, which means that there is only once or twice inconsistency between Map-Replies from different MRs for one specific EID, since changing the status of xTR is not frequent.

Just few inconsistencies, only 1.66%, are caused by the RLOC set. This scenario usually occurs for those EIDs whose RLOCs change frequently over time (the case of Statistical

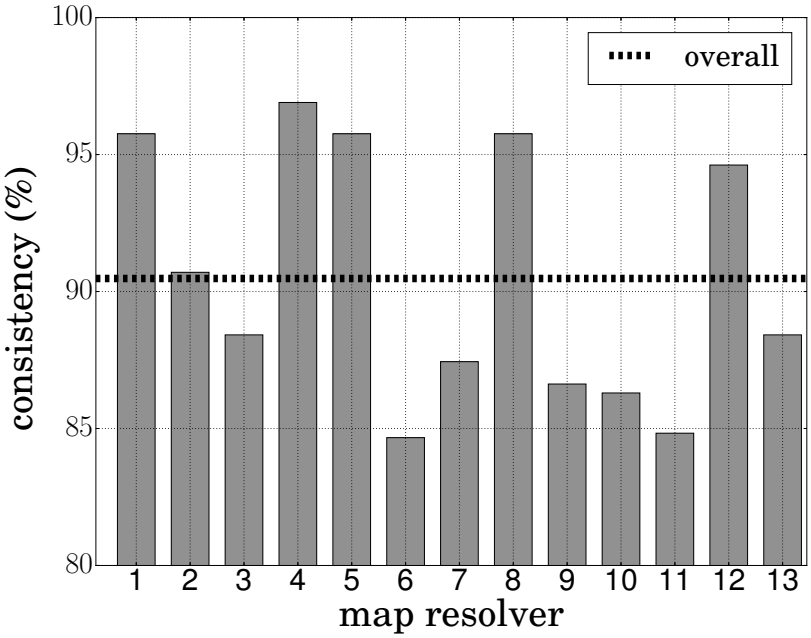


Fig. 4.7 Overall consistency by VP from the different Map-Resolvers.

outliers instability). So, because of the convergence time of the mapping system, it is very difficult to guarantee that every MR can provide the same LISP Map-Reply at every experiment round. This reason for changes in the RLOC set is most probably due to traffic engineering policies. Similarly as the previous case, when different MRs provide a different RLOC set for a specific EID, the requesting xTRs may end up using different RLOCs to reach the same destination EID.

4.3.2 Consistency evaluation by VP

The overall percentage of consistency by VP for the whole experiment dataset is 90.48%, presented by the dotted line in Fig.4.7, which indicates that the Map-Replies received at different VPs from the same MR are highly consistent. Yet, we can also observe that different MRs show some deviations from the overall consistency by VP, reflecting the fact that certain MRs send the same Map-Reply to all VPs, but some do not. The deviation among different MRs is however limited (less than 12.23%). Therefore, we can certainly state that the mapping system shows generally a high degree of consistency by VP but presents a non-negligible amount of inconsistencies (9.52%). In the rest of this section, following the same approach as for the consistency by MR, we go deeper in this issue, exploring the reasons why some observations exhibit inconsistency by VP.

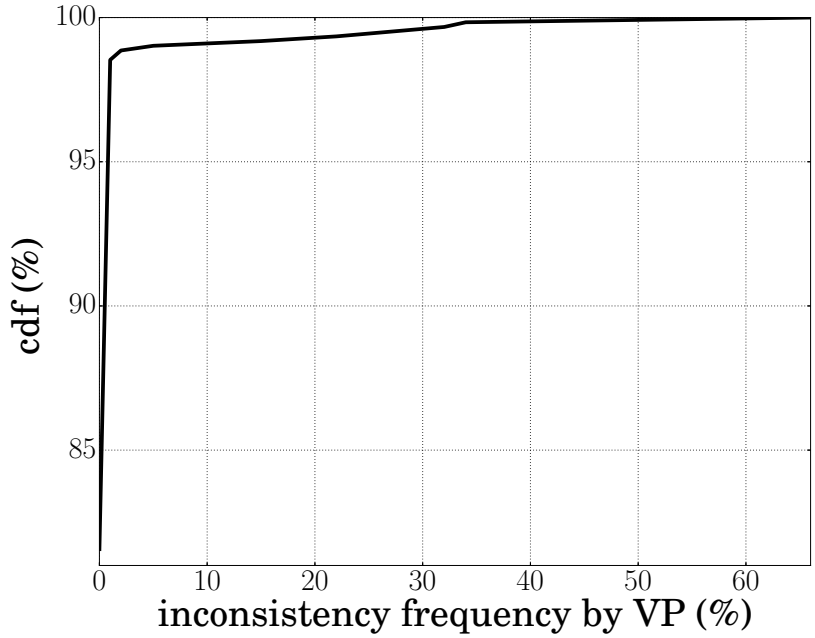


Fig. 4.8 Cumulative Distribution Function of inconsistency frequency by VP. X-axis is inconsistency frequency by VP, indicating the ratio of inconsistency occurrence for each $<EID, MR, *>$ tuple over total experiment number.

Fig.4.8 shows the CDF of the inconsistency frequency by VP: $f_{<EID,MR,*>}$, i.e., the probability of inconsistency frequency for all $<EID, MR, *>$ tuples, by using the following definition:

$$f_{<EID,*,VP>} = \frac{\text{\# Inconsistencies by VP}}{\text{\# Experiments}} \times 100\%. \tag{4.3}$$

Like for the explanation for Fig. 4.6, here an inconsistency frequency by VP of 0% means that all Map-Replies received by all of the VPs from a same MR are totally consistent at every experiment round during the whole experiment campaign, while an inconsistency frequency of 100% means that at least one of VPs received an inconsistent Map-Reply compared to the other VPs at every experiment round. The probability of consistency by VP (when X-axis is 0% in Fig. 4.8) is similar to that by MR, with a value of 81.57%. It can be noticed that this percentage is slightly lower than the overall percentage of consistency by VP in Fig. 4.7. The reason is that here the inconsistent occurrence with 0% only includes those EIDs for which all VPs always show consistency, no matter which MR is actually queried. A sharp increase with a growth of 16.97% is found when X-axis is 1%, which means that although VPs sometimes receive inconsistent Map-Replies at several experiment rounds, the occurrence remains pretty low, hence showing that even by VP the mapping system exhibits a very high degree of consistency. Differently from the largest inconsistency

frequency by MR, which goes up to reaches 100%, the largest inconsistency frequency by VP is only 66%. This phenomenon indicates a fact that the Map-Replies received by different VPs are not inconsistent all the time, at least there are one third of observations show that their Map-Replies are totally same. Thus, the consistency by VP is higher than consistency by MR.

To further explore the inconsistency by VP, we then use same metrics as in Sec. 4.3.1, with the exception that the queried target changes from $<EID, *, VP>$ to $<EID, MR, *>$ tuples.

The most inconsistent case, 89.06%, is due to a mix of Negative Map-Reply and LISP Map-Reply, i.e., demand for a same $<EID, MR, *>$ tuple, at some VPs receive Negative Map-Reply but at others receive LISP Map-Reply. This is caused by the New Deployment case, but differently from the cause for inconsistency by MR, the reason for inconsistency by VP is probably because of Map-Requests sent by different VPs can not be processed exactly at the same time, due to queuing in the MR. Thus, if the queries are sent during the period of mapping information update, some VPs, whose Map-Requests are processed first will have different Map-Reply Type from the latter ones. Inconsistencies caused by different RLOC sets, i.e., different VPs receive LISP Map-Replies with different RLOCs set from the same MR, also exist, contributing for around 10.94%. It is probably caused by a combination of the Statistical outliers instability case and the queuing mechanism in MR. As the RLOC sets provided by a MR for some certain EIDs frequently change, and the Map-Requests sent by different VPs need to be processed one by one, it is very possible that the first processed Map-Requests are replied by the different RLOC sets compared to the latter ones. The difference between Map-Replies is probably due to traffic engineering policies, but the further experiments are needed to explore the deeper reasons for such inconsistency.

4.4 Summary

During the evaluation of LISP mapping system performance, besides the observations about the stability and consistency that we describe in details in the previous sections, we also investigate some interesting LISP-related results. We provide them here so that they can be used to improve LISP performance. We identified that 8.97% of the LISP-sites leverage on multiple RLOCs in order to be multihomed to the Internet, with an average of 3 distinct locators. Interestingly, when we compare the locators with their Autonomous System Number we found that around 50% of the locators in multihoming situations belong to different ASes, which indicates that LISP is used to transparently interconnect LISP-sites by the in-

termediate of multiple ISPs.[3] As the control-plane (i.e., MS/MR and mapping system) and the data-plane (i.e., ETR/ITR) are decoupled in LISP, Map-Replies can originate from the LISP-sites directly or from shared utilities constituting the mapping system (i.e., MS/MR). This particularity of LISP is already used in the current deployment, where only 88.8% of the non-negative Map-Replies are originated directly from their LISP-sites, with the remaining 11.2% being originated by delegated MSes replying on behalf of the LISP-sites. This allows to reduce the control traffic seen by LISP-sites and help them to scale better.

In addition to mappings, LISP also relies on encapsulation, meaning that with LISP traffic can be transported over IPv4 or IPv6 without requiring end-hosts to be dual-stack. Interestingly, we identified that no less than 8.84% of the locators are IPv6. However, we have not observed cases where a LISP site was reachable strictly using IPv6 while 82.79% of sites are reachable only through IPv4. This discrepancy tends to confirm that network operators still lack of confidence being reachable only through IPv6. As a result, the 17.21% of LISP-sites explicitly using IPv6, always include at least one IPv4 address in order to be reachable.

As a solution to cope with the increase of BGP routing table, traffic engineering, and scalability issues, LISP has been widely deployed on the LISP Beta Network experimental testbed for ten years. To ensure accurate and efficient communications, LISP mapping system should be stable and consistent. However, there is no study evaluating the stability and consistency of LISP mapping system. For such purpose, we continuously measured the LISP Beta Network for seventeen days to assess the mapping system from these two aspects. Measurements show that the LISP mapping system is stable and consistent over time. Nevertheless, instability and inconsistency are observed and for studying them in details we developed a new taxonomy. All in all, instability and inconsistency are rare events. At last, we observe utilization of multi-homing and dual-stack (i.e., IPv4 and IPv6) during the whole experiment and plan to use these results to further improve the LISP performance.

[3] We used the Team Cymru BGP database to associate IP addresses to ASNs (see http://www.team-cymru.org/IP-ASN-mapping.html).

Chapter 5

LISP-Views: LISP mapping system monitor

Currently only two LISP experimental platforms have already been deployed in the Internet: LISP Beta Network and LISP-Lab platform. Yet, a single LISP monitoring system called LISPmon [18] supervises the global MDS and publishes the mapping information daily. However, it is known that the mapping information sometimes changes frequently within a day and that the elements constituting the MDS are not always consistent (more details can be found in Chapter 4 in this dissertation) [82]. We hence propose a dynamic versatile LISP monitoring architecture, namely LISP-Views, to overcome these limitations. LISP-Views automatically explores the whole MDS every 2 hours and stores the detailed mapping information, so to facilitate the experimenters to evaluate the LISP comprehensive performance.

We used a one-month long set of traces produced by LISP-Views to evaluate its performance and accuracy, including comparing it with LISPmon. We show that LISP-Views is more accurate than LISPmon since it monitors all the MDS elements in parallel. In addition, thanks to its detailed reporting, LISP-Views allows to assess high level metrics of LISP deployments such as reliability, latency, or configuration issues.

In the remainder, Sec. 5.1.1 introduces LISPmon and explains why we propose a new LISP monitor. Sec. 5.1.2 describes our proposed LISP monitoring architecture in details. Sec. 5.2 validates LISP-Views by comparing with LISPmon. Sec. 5.3 provides the snapshot of what kind of further analysis can be done with our proposal. Finally, Sec. 5.4 concludes the chapter.

5.1 Proposed monitoring architecture

5.1.1 Motivation

In order to move LISP forward, we need to deeply understand the behaviour of the different LISP network entities and since the MDS reflects the status of a LISP network as it stores all the mapping information, it is essential to be able to monitor the MDS. LISPmon was the first step towards a systematic LISP monitoring. However, it monitors the MDS just from one vantage point (VP) once per day and only queries one MR. Upon MR issues, LISPmon must be manually re-configured to monitor another MR. Yet the mapping information may be unstable and inconsistent between MRs, i.e., the mapping information sometimes changes within a day, and the Map-Replies from the different MRs may not coincide at a given time [82]. It is similar to the BGP looking glass servers, which do not always provide the same responses for an IP address as the whole routing system may not have converged or because of routing policies. From such point of view, LISPmon has some limitations since it is not able to detect the changes of mapping information within one day, and is not able to show the differences among MRs. Thus, we propose a new versatile LISP monitoring architecture, called LISP-Views, to monitor public LISP deployments, as well as to enable further performance evaluation of LISP defined by the users themselves. In fact, LISP-Views not only can be used to monitor LISP, but also can be extended to be used in non-public networks, such as VxLAN [83].

LISP-Views is an open source implementation and has been designed to fulfill the following objectives:[1]

1. LISP-Views queries to all the working MRs existing in current MDS, in parallel, while LISPmon just queries one, aiming at building a complete view of the current LISP status.
2. LISP-Views periodically monitors all the MRs with arbitrary intervals,[2] while LISPmon just does it daily, aiming at providing information about the mapping evolution at smaller time granularity.
3. LISP-Views supervises the whole MDS without any manual process, automatically reacting to failing components (e.g., unresponsive MRs).
4. LISP-Views obtains the mapping information from all the MRs of the LISP Beta Network as well as LISP-Lab platform, whereas LISPmon prefers to leverage the MR of the former one.

[1] Source code available on Github: https://github.com/SeleneLI/LISP-Views

[2] This interval can be changed by the experimenters to accord their necessity.

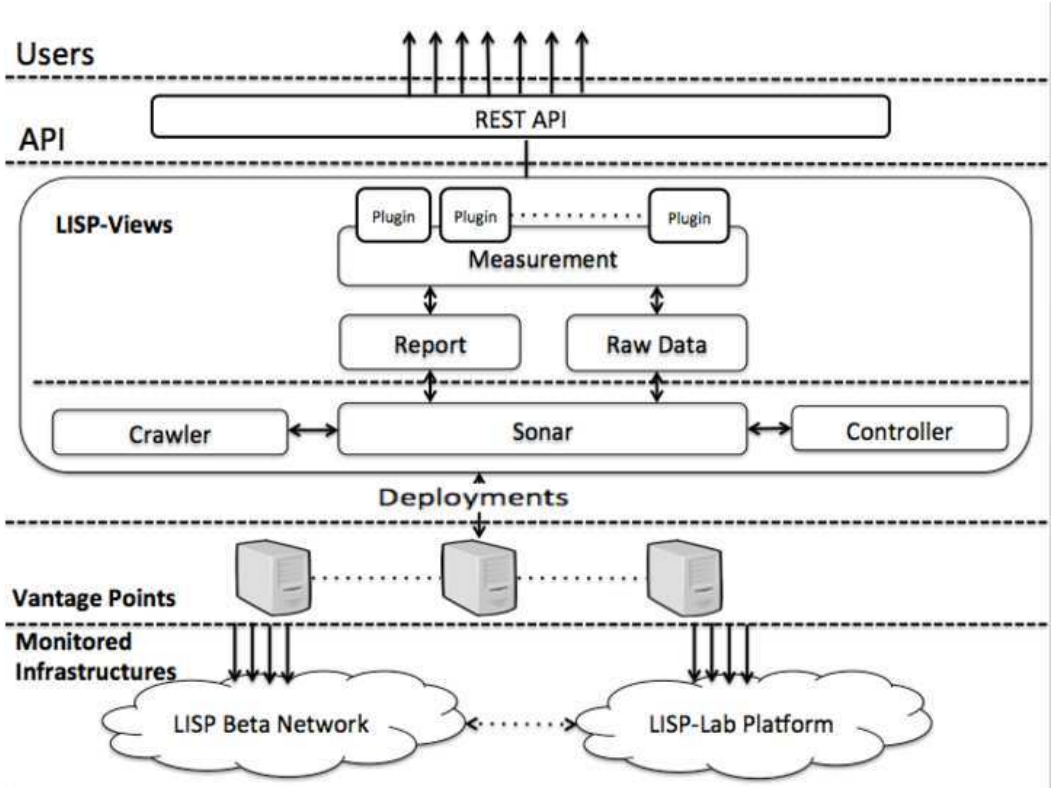


Fig. 5.1 Monitoring Architecture

5. LISP-Views is flexible and configurable, with users able to define different monitoring jobs and get the various measurements, whereas LISPmon just publishes the mapping list daily.
6. LISP-Views is designed to be easily deployed in a distributed manner.

5.1.2 Description of LISP-Views

The architecture of the proposed LISP-Views monitoring tools is depicted in Fig. 5.1. LISP-Views consists of several modules with different functions. The *Measurement*, *Report*, and *Raw Data* modules are deployed on a centralized server while, the *Crawler*, *Sonar*, and *Controller* modules can be deployed on several different *VPs*. As for now, LISP-Views is just deployed on one VP in Paris, France, where the monitor is developed and the centralized module is deployed. The distributed version of LISP-Views is under development, for this reason, we will not discuss further how to deploy LISP-Views on multiple VPs. All the modules are implemented in *Python* and described as follows:

Sonar module: the main module with two functions: 1. sending the LISP encapsulated Map-Requests and receiving the Map-Replies to/from all the existing MRs based on the

standard [47]; 2. storing the received information in *Report* and *Raw Data* with different purposes.

The IP address that *Sonar* uses to query to MRs is selected either from the output of *Crawler*, or from the recorded information in the previously produced *Report*. The reasons are explained in the corresponding modules.

Crawler module: scans all the existing IPv4 addressing space using *Sonar*. If the queried IP (e.g., 192.0.2.1) has no Map-Reply, *Crawler* increments the IP by one as the next queried IP (i.e., 192.0.2.2) and lets *Sonar* send the Map-Request. If *Sonar* obtains Negative or LISP Map-Reply, since the Reply contains a prefix (mentioned in Sec. 2.1), *Crawler* sets the next query to the first IP beyond the returned prefix. For instance if the returned prefix is 192.0.2.0/25, the next queried IP is 192.0.2.128. Scanning the whole IPv4 address space takes very long time, and is mainly caused by time wasted waiting for no Map-Reply. Indeed, no Map-Reply means to wait 3 seconds, and the next queried IP is just increased by one, instead of skipping a block of IP addresses like in the case of Negative/LISP Map-Reply.

Report module: contains the collected EID-prefixes in a list, as well as a list of MRs that answered. As crawling the address space may take long time but we aim to obtain the status of MRs as frequently as possible, *Sonar* sends Map-Requests for the EIDs recorded in *Report* only to MRs that have previously responded. Thus, it decreases the possibility to receive no Map-Reply, so that to get LISP status within a shorter time.

Raw Data module: contains all the detailed information of Map-Replies for each MR, such as Map-Reply type, RLOCs, required EID-Prefix, Round Trip Time (RTT) and the returned source for each round specify with *TimeStamp* (regardless the source of *Sonar*). So, it can be used to perform thorough performance analysis of MDS. Moreover, since the *Raw Data* is stored according to the MR, it is possible to track the performance of the MRs individually.

Measurement module: provides the composition of requested measurements (i.e., select different measurement plug-ins) based on the analysis of *Raw Data* and *Report*.

REST API: is connected to *Measurement module* so that the users can launch a custom experiment by setting the experiment time, the monitored MRs, and obtain the different aspects of LISP status. We are currently implementing the REST API, so the results of measurements presented here were obtained via the command lines.

Controller module: synchronizes all the modules in LISP-Views and also specifies the start and stop time of producing both *Report* and *Raw Data*. The interval of generating *Report* and *Raw Data* can be changed according to the hardware processing capability. However, since the inconsistency between MRs exits, the interval of producing Report and

Raw Data for every MR differs from each other. Thus, the interval set in Control module should cover the slowest MR.

5.2 LISP-Views validation

5.2.1 Methodology

In order to validate and evaluate the LISP-Views monitoring tool, we used raw data and report collected during one month (from 0:00 September 4th to midnight of October 4th 2016), by deploying LISP-Views on one VP, which is an xTR of LISP-Lab platform.

The interval to produce reports was 6 hours, and the interval to produce raw data was 2 hours. Unfortunately, both MRs of LISP-Lab platform just responded to the Map-Requests at the first time and then stopped. They were not able to handle large number of queries, because the Map-Requests did fill the request queues, resulting a drop in the MRs. Per-se this is a success, because this bug was unknown, and the OpenLISP coders fixed the issue after our measurements. All the results of evaluation used in this section only depend on the 6 working MRs of LISP Beta Network (3 in Europe, 1 in US and 2 in Asia) and the other 6 MRs were down at the moment of conducting the experiment. The aim of the work is to validate LISP-Views by assessing if it provides at least the same results as LISPmon.

5.2.2 LISP-Views vs. LISPmon

In this section, we compare LISPmon and LISP-Views, so to assess if the information provided is comparable between the two monitoring platforms, hence validating LISP-Views. As we indicated in Sec. 5.1.1, at any time LISPmon just queries one MR, on LISP Beta Network once per day and generally begins at 7:00 a.m.. The queried MR is changed sometimes when it has issues. LISP-Views, however, keeps sending Map-Requests to all the 6 MRs with an interval of 2 hours everyday, so to retrieve the actual real status of each MR.

Fig. 5.2 shows the number of LISP Map-Replies received from LISPmon and LISP-Views over time. The upper sub-figure makes a comparison between the 2 monitors. The line with point is obtained from the daily LISPmon publications, indicating the number of LISP Map-Replies returned by one MR in an uncertain time monitoring round every day. To facilitate the comparison, we pick up the results from LISP-Views at 8:00 a.m., which is the nearest experimental round to LISPmon. As presented in the curve with triangle symbols, LISP-Views reflects the fact of a combination of the number of LISP Map-Reply containing the different EID-prefixes from all the MRs at a fixed experiment time. In most cases, LISP-Views receives around 20 LISP Map-Replies more than LISPmon, reflects the fact that the

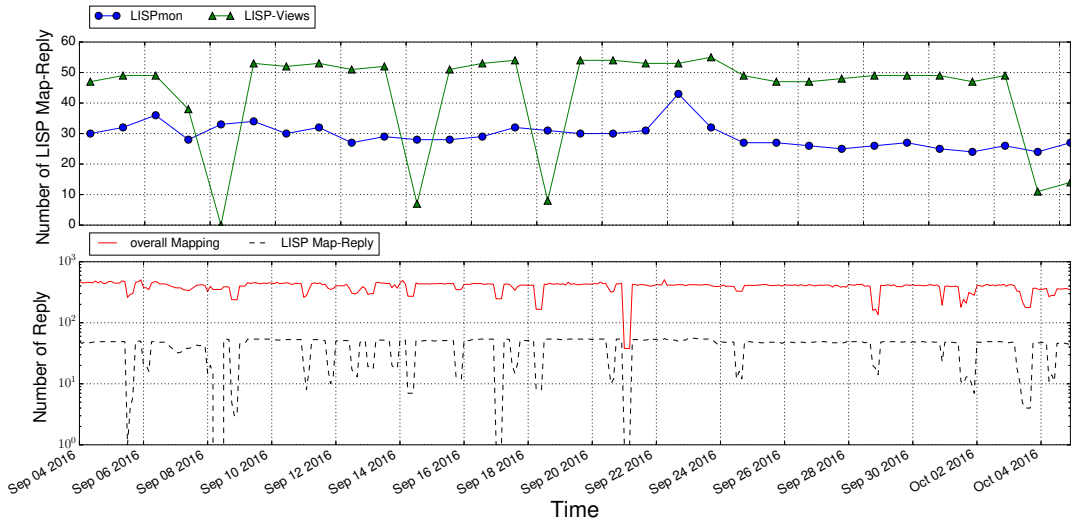


Fig. 5.2 Comparison between LISPmon and LISP-Views. The upper sub-figure shows the number of LISP Map-Reply from LISPmon generally at 7:00 and LISP-Views exactly at 8:00 over days. The bottom sub-figure indicates the LISP Map-Reply and overall Mapping from LISP-Views over time with an interval of 2 hours.

MRs are not coherent at any time due to the existence of convergence time for synchronizing the mapping information among them. The other 5 days, where LISP-Views receives very few LISP Map-Replies, show that there were issues to normally receive LISP status at that time, whereas LISPmon always presents a smoother trend but hides the reality.

The lack of stability can be better appreciated in the bottom sub-figure of Fig. 5.2. Such figure depicts the number of LISP Map-Reply (in dashed line) and the overall Mapping (LISP and Negative Map-Reply, in solid line) from LISP-Views for every 2 hours monitoring round during the whole experiment. Although the number of Map-Replies presented in the figure is still the combination of results from 6 MRs when the returned EID-prefixes are different, it shows that the MRs are not quite stable. Sometimes the MRs even provide 0 or less than 10 LISP Map-Replies, but in the next experiment round, the number is recovered. Besides, we evaluate the total number of successful queries every 2 hours. It also presents the behavior of instability, where it generally oscillates between 400 and 500, but it occasionally drops lower than 300. Both solid and dashed lines have almost same changing trend, indicating that the change of overall Map-Replies is mainly affected by LISP Map-Replies. Monitoring rounds where the number of LISP Map-Replies approaches 0 but where Negative Map-Replies are still received indicate issues at the MRs, especially for answering the LISP Map-Reply.

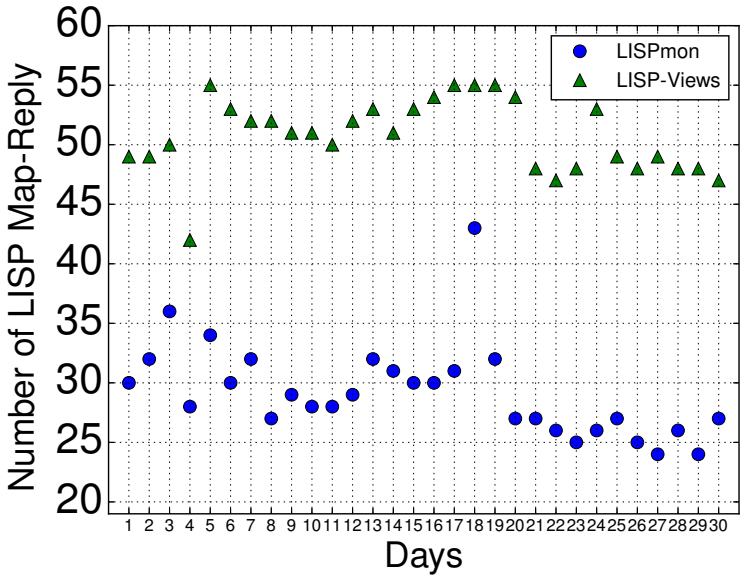


Fig. 5.3 Comparison between the LISP Map-Reply of LISPmon and LISP-Views over days

We compared the number of received LISP Map-Replies within a whole day between LISPmon and LISP-Views over 30 days. Shown in Fig. 5.3, as LISPmon only publishes one record each day, the number (represented by points) is exactly identical to those in Fig. 5.2. LISP-Views, however, provides a combination of LISP Map-Replies not only from all MRs but also from 12 monitoring rounds within one day. We observe that our monitor architecture receives more LISP Map-Replies than LISPmon in all days with a large difference. 83.3% of the time LISP-Views receives more than 20 LISP Map-Replies. The maximum number of LISP Map-Replies that our proposed monitor obtains is 55, while the maximum value for LISPmon is 43. Differently from the upper sub-figure of Fig. 5.2, where LISP-Views occasionally receives nearly 0 LISP Map-Replies, the total number within a whole day is always more than 40, which illustrates that the MRs may not work normally in some rounds but they are able to recover fast.

Since LISP-Views repeats querying simultaneously to every MRs, it is able to report on the status of each MR at any time. It provides more complete mapping information compared to LISPmon and is also able to highlight sporadic issues with MRs.

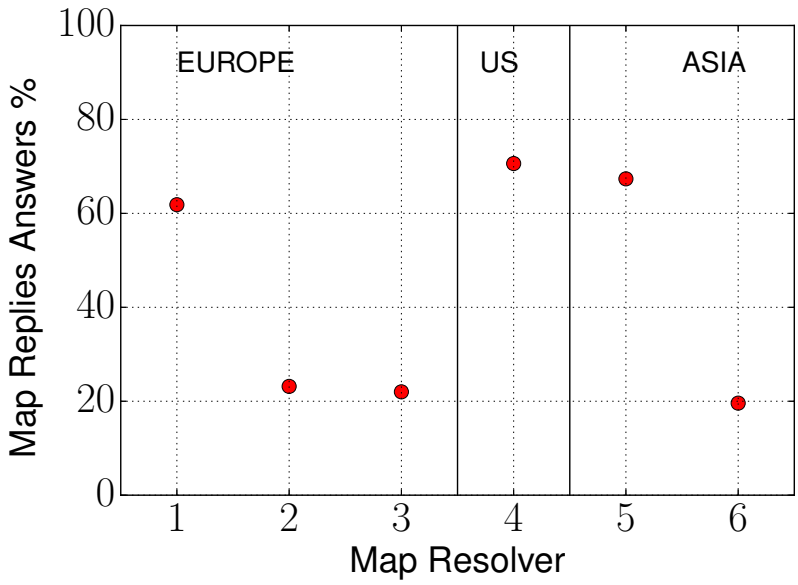


Fig. 5.4 Reliability of each MR

5.3 Dissecting LISP with LISP-Views

After validating LISP-Views in the previous section, this section presents several examples of how LISP-Views can be used to dissect LISP and obtain in-depth results. The experimental dataset used is the one in Sec. 5.2.1.

Fig. 5.4 shows the *Reliability* of each MR during our data collection, by calculating the percentage of successful queries over the total number of Map-Requests. MR1, MR4, and MR5 give the highest reliability values, which are more than 60%. The lowest reliability values, about 20% are observed for MR2, MR3, and MR6. In general, the reliability of each MR is different and low compared to the years of 2012 and 2013 presented in [44], which coincides with the fact that there was a change of the MRs architecture on LISP Beta Network that year, and an updated-software was tested on MRs as well. The low value of reliability is caused by MRs having an unstable behaviour, as shown in Fig. 5.2, the number of Map-Reply changes heavily over time, and especially the number of LISP Map-Reply sometimes drops to 0.

In order to understand the behavior of each MR in terms of latency, we analyze the median RTT obtained from our dataset. The RTT here refers to the Round Trip Time from sending out the Map-Request until receiving the Map-Reply. Fig. 5.5 shows that the best performance come from MR1, MR3, MR4, and MR5; since the overall RTTs are much lower than the others. For MR3, the number of LISP Map-Replies is much higher than the number of Negative Map-Replies, probably because its embedded Map-Server registers

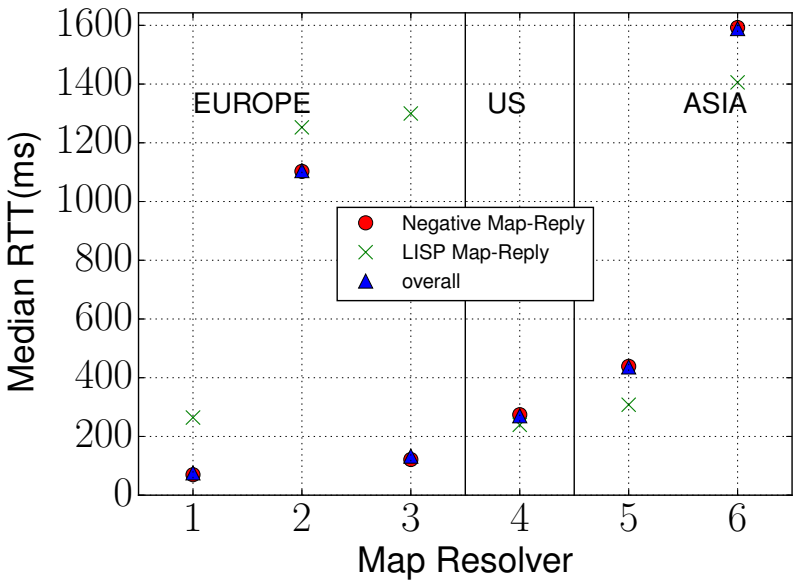


Fig. 5.5 Median RTT per MR (In the most time, the Negative Map-Reply and overall are overlapped.)

less EID-prefixes hence requiring Map-Requests to be forwarded to a remote MS. Then, the Map-Request is forwarded to the xTR, where the remote MS registers and the xTR gives back the Map-Reply. Compared to the Negative Map-Replies, which are normally returned by MR, LISP Map-Replies take longer time, especially in this case. However, MR2 and MR6 present very high latency, particularly for MR2 (located in Europe), but its RTT is even higher than MR5, which is located in Asia. If we focus on the RTT of LISP Map-Replies, only MR1, MR4, and MR5 provide the best behavior, which coincides with the result of Fig. 5.4 on reliability. The high RTTs of LISP Map-Replies are from the other 3 MRs, where the median RTT is around 1300 ms, is caused by the failed queries. Furthermore, it also explains why we can always receive the Negative Map-Replies, but the number of LISP Map-Replies is sometimes quite low in Fig. 5.2. Half of LISP Map-Replies are more than $1300ms$ and partly even more than $3s$. Since our measurement timeout is set to $3s$, it implies that some Map-Replies are dropped by our measurement unit. As the RTT of Negative Map-Replies is generally lower, we can receive more of them. This phenomenon may be very biased by the VP location. In addition, Fig. 5.5 presents that the number of Map-Replies returning from the MRs located in Europe and US is indeed higher than the number of Negative Map-Replies, which is expected. However, both Asian MRs return LISP Map-Replies faster than Negative Map-Replies. The reason of this observation is unclear and requires further exploration.

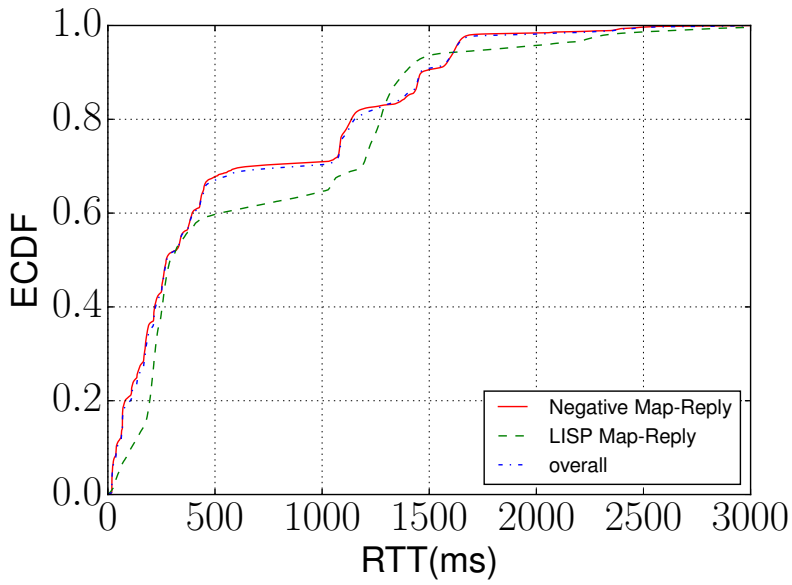


Fig. 5.6 CDF RTT comparison between LISP and Negative Map-Reply

Table 5.1 Percentage of Mapping Source

Map-Reply Type	Map Resolver	xTR	Other
Negative Map-Reply	98.79%	-	1.21%
LISP Map-Reply	0.14%	88.37%	11.49%

We also calculate the Cumulative Distribution Function (CDF) of the RTT for the 6 MRs. Fig. 5.6 provides the CDF of the RTT for Negative, LISP, and overall Map-Replies. As indicated in Sec. 2.1, the LISP MDS needs more time to solve a complete mapping (LISP Map-Reply) than the Negative Map-Reply. We observe that the Negative Map-Reply is indeed faster than the LISP Map-Reply until the RTT reaches 1000 ms, then the behavior changes, i.e., the LISP Map-Reply sometimes becomes faster. What's more, we find that 67.03% of RTTs within 500 ms, 70.33% less than 1000 ms and 90.88% do not exceed 1500 ms for overall Map-Replies. In details, for the Negative Map-Reply, we found that 67.8% of RTT values within 500 ms, for the LISP Map-Reply that 59.7% of RTT values less than 500 ms. This figure almost presents a bi-modal distribution, where 500 ms is a peak and 1300 ms is another one. These two high occurrences of RTTs are exactly the most two frequent latency in Fig. 5.5.

The dataset obtained with our monitoring architecture also shows the information about *Mapping Source*. We explore the source answering the Map-Replies according to different types (LISP or Negative). In the case of LISP Map-Replies, we expect the source of replies to come either from one of the ETRs or the queried MR. On the contrary, for Negative Map-

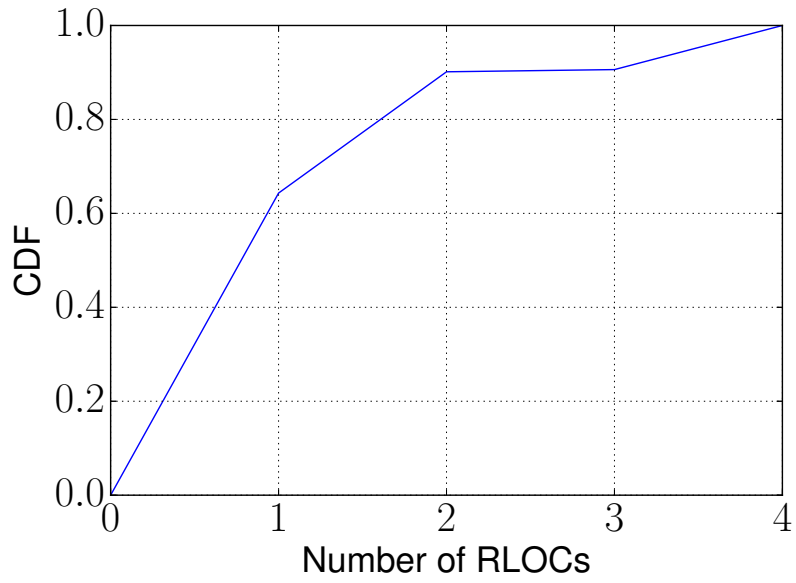


Fig. 5.7 CDF of the number of RLOCs

Replies, replies should just come from the queried MR. Tab. 5.1 presents the percentage of observations for the two types of Map-Replies. For the Negative Map-Reply, 98.79% come from the queried MR and 1.21% come from the other sources. Further, the other sources are actually the other MRs without query and it happens just for a fixed EID-Prefixes, i.e., if we send a Map-Request for one of these EID-Prefixes to a dedicated MR, the Map-Reply always comes from a fixed specific MR. As a conclusion, all the Negative Map-Replies are answered by MRs. For the LISP Map-Reply, we observe that 0.14% come from the queried MR, 88.37% come from one of their ETRs, but 11.49% come from the other sources in different locations. This unexpected behaviour needs further investigation.

The following type of measurement is the distribution of the size of RLOC set, i.e., how many RLOCs are associated to one EID prefix on average. As previously observed, the percentage of mappings using two or less RLOCs has increased between 2010 and 2012 [108]. Fig. 5.7 shows that this trend keeps going on, i.e., more mappings use fewer RLOCs and the maximum number of RLOC is 4. An interesting point is that although LISP is a good candidate to support the increasing multi-homing in Internet, more than 60% LISP users are not multi-homed and among them the majority only has 2 RLOCs. Not only the number of RLOCs of each site does not significantly change, but also the RLOCs themselves remains stable. We measured that the stability reaches 99.8% for all the dataset, i.e., once an EID-prefix – to – RLOCs mapping is decided, it rarely changes. We have not found any mobile LISP sites.

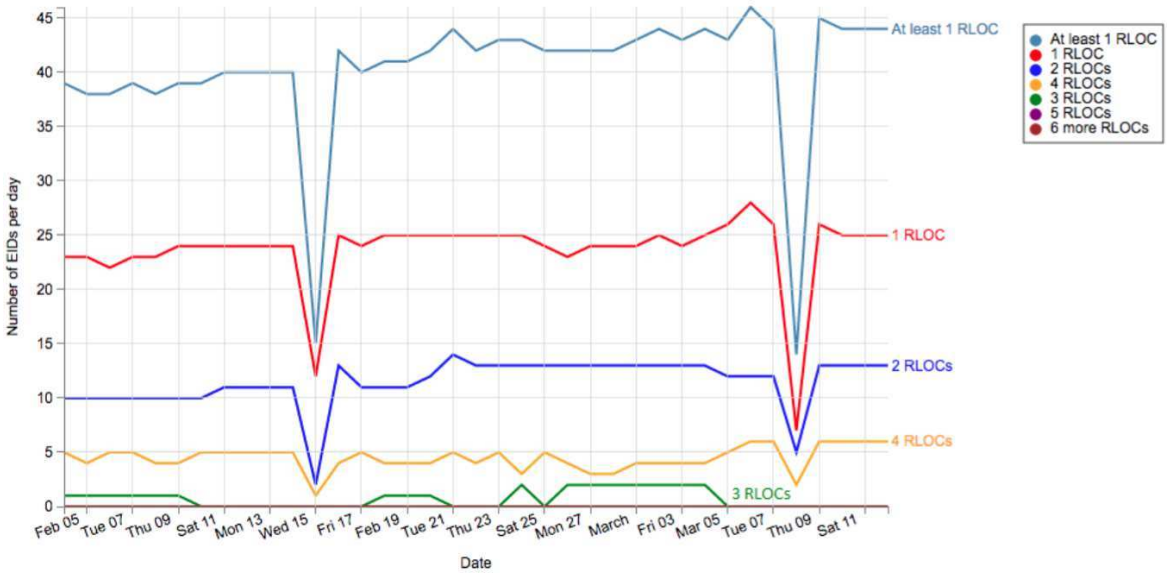


Fig. 5.8 EIDs by Quantity of Associated RLOCs [15]

All the aforementioned observations are based on the experiment made between September and October 2016 as described in Sec. 5.2.1. Later LISP-Views detects that MR2 and MR6 in Fig. 5.5 with very high overall RTTs are down, and three new other MRs are up. After confirming with the operators of the LISP Beta Network, these two MRs are indeed definitely down. The change of the architecture of MRs also proves the accuracy of LISP-Views, while the LISPmon presents a smooth change in its daily reports, hence hiding these facts. Since February 2017 LISP-Views publicly publishes preliminary daily reporting online [15]. Fig. 5.8 is a capture from the website, it is one type of measurement about the number of EID-prefix quantified by the different size of RLOC set. The shown result is an union of all MRs during a whole day, to present the most complete mapping information of the actual LISP deployment. The line on the top indicates the number of LISP Map-Replies, i.e., the Map-Replies with at least 1 RLOC, which coincides to the results shown in Fig. 5.2 and is mainly affected by the EID-prefix with 1 RLOC. Moreover, the composition of the different sizes of RLOC set is also rather identical to the one in Fig. 5.7. The only difference is that in the previous dataset there is no EID-prefix with 3 RLOCs, while in the latest dataset it appears, but not very stable. In addition, the line with 3 RLOCs is almost complementary to the line with 4 RLOCs. It is probably caused by one LISP-site that has 4 xTRs but one being always down, or the LISP-site having 4 interfaces on a same xTR among which one is down. Besides, the number of observed EIDs per day heavily drops two times in the figure. Since the two valley don't drop to zero, we know that the problem doesn't come directly

from the MDS. Instead, the problem comes from issues that occurred within the network of the LISP-Views server itself. This observation highlights the need of distributing VPs.

5.4 Summary

Very little is known about the behavior of LISP in operational environments and it still lacks of troubleshooting tools. Motivated by the only LISP monitor deployed so far, named LISPmon, which records the current LISP status only from a specific MR just once per day, and also by the measurement work presented in Chapter 4, we propose a more dynamic LISP monitoring architecture, namely LISP-Views, so to deepen the understandings on LISP and to ease day-to-day operations and troubleshooting. As LISP-Views aims at being deployed in large scale and dynamic networks, we make a comparison with LISPmon to validate the former one by comparing their behavior during one full month.

It demonstrates that LISP-Views provides more information by discovering more mapping information from all MRs and more complete mapping information of each MR. Furthermore, with our proposed monitoring platform, more mapping system performance metrics such as reliability, latency, and configuration issues can be assessed, which helps for further LISP improvements.

Chapter 6

Assessing LISP interworking performance through RIPE Atlas

Thanks to two testbeds: the LISP Beta Network and the LISP-Lab project, LISP is gradually deployed in the wild. The interworking mechanism is proposed to ensure the communication between LISP-speaking sites and legacy Internet. Although LISP has been evaluated in terms of scalability [108], stability [82], LISP evolution [77], and delay resolving the bindings between EIDs and RLOCs ([108], [44]), LISP was never analyzed from the aspect of the interworking with the legacy Internet at large scale. This dissertation fills this gap by providing a latency evaluation and routing path measurement about LISP interworking mechanism. The work is based on two experiments (one is conducted in 2015, the other one is in 2016) by using RIPE Atlas, which is the largest existing Internet measurement infrastructure for both IPv4 and IPv6.

The experimental results confirm that the use of proxies to connect LISP and non-LISP sites introduce negative effects, which are important for the nearby destinations but can be ignored for the intercontinental long-distance destinations. It also shows that the selection of the proxy location is very important, since having them close either to the sources or to the destinations can decrease a lot the negative stretch. Although LISP introduces some overhead, the performance is quite stable for IPv4. However, the same conclusion does not hold for IPv6. We also observed that the interworking performance of the LISP-Lab platform is more reliable than the LISP Beta Network.

In the remainder, Sec. 6.1.1 and Sec. 6.1.2 introduce the necessary resources on which our experiment leveraged. Sec. 6.2 describes the methodology we used to conduct the experiment. Sec. 6.3 and Sec. 6.4 respectively present the IPv4 ping results obtained from 2015 and 2016. Sec. 6.5 shows the IPv6 ping results and Sec. 6.6 indicates the traceroute results for both IPv4 and IPv6.

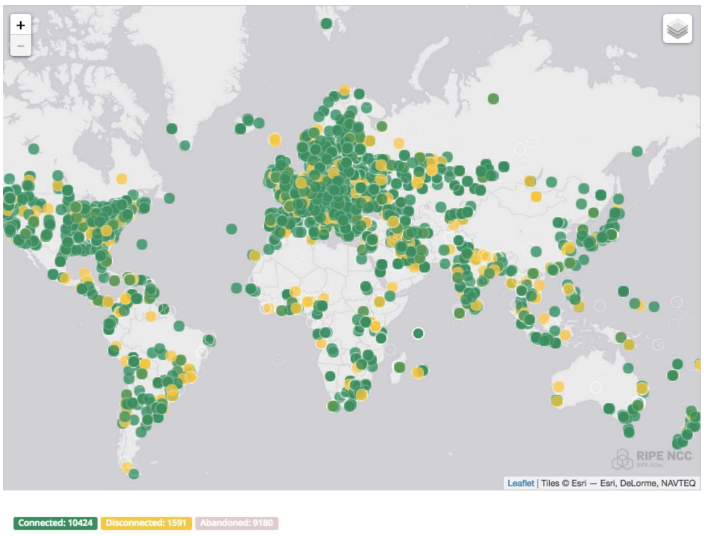


Fig. 6.1 Deployment of probes on RIPE Atlas in 2017

6.1 Experiment resources

6.1.1 RIPE Atlas

RIPE Atlas [27] is the largest Internet measurement infrastructure, consisting of a global network of more than 9000 probes all over the world that measure Internet connectivity, reachability, and provide an understanding of the state of the Internet in real time. The deployment of worldwide probes is shown in Fig. 6.1. From early 2013, the hardware of the probe is a modified TP-Link wireless router (model TL-MR 3020) with a small USB thumb drive in it, but this probe does not support WiFi. Atlas also has more than 200 worldwide anchors, which are enhanced probes, offering more process capacity and sufficient bandwidth to support the larger number of measurements. Thus, the anchors are normally more stable, and can be used as reference. All the probes and anchors whenever are set up, they automatically start executing a set of pre-defined measurements, called built-in measurements. The set of built-in measurements contain *ping*, *traceroute*, *DNS*, *SSL* and some *HTTP* requests, mostly towards well-known targets such as DNS root servers, but also towards some of the RIPE Atlas infrastructure components. Besides, the probes and anchors also provide the user-defined active measurements with the same measurement types. Volunteers all over the world host these small hardware devices to actively measure Internet performance. To avoid malicious attacks, credits are required to launch experiments and the amount of required credits varies according to experiment types. In addition, the maximum allowed number of measurements towards the same target is 10. RIPE Atlas provides a set of RESTful API with which experiment campaign parameters (e.g.,type of measurement, query source and desti-

nation, the duration and interval of experiment etc.) are passed to probes and measurement traces can be retrieved. Based on the available API, one can schedule experiment campaign in an automatic manner.

6.1.2 Alexa

Alexa [1] [2] is a website created by Amazon which provides commercial web traffic data, global rankings, and other information on 30 million websites. The analytic such as website rankings is based on the data collected by a toolbar developed by Alexa and installed within users web browser. The toolbar provides functions such as popup blocker, a search engine,etc. In early 2015, Alexa stated that there had been 10 million downloads of the toolbar. Alexa ranks sites based primarily on tracking a sample set of Internet traffic — users of its toolbar for the Internet Explorer, Firefox and Google Chrome web browsers. Due to its huge sampling space, the website ranking that it published is widely used to evaluate the popularity of websites.

6.2 Measurement methodology

The mechanism of LISP interworking is presented in Sec.2.2 and Coras et Al. [44] describe how the use of PxTRs introduces a stretch in the path between LISP and non-LISP networks. Thus, it is necessary to conduct a large scale experiment in the real world to quantify this overhead. The objective is to answer three questions through such an experiment: how much is the negative impact? Under which conditions such stretch has an important impact on the performance? Conversely, under which situation such overhead is so small that it can be ignored? The follows describe how we conduct our experiments.

6.2.1 Dataset 2015

As the purpose of our experiment campaign is to fully obtain the knowledge about the LISP interworking performance with legacy Internet, we deployed a probe (RIPE Atlas probe number is #22341) with IPv4 and IPv6 address on the LISP-Lab platform inside an academic institute in Paris, France to conduct the LISP-enabled active measurements. For IPv4, it uses both PETR and PITR of LISP-Lab to communicate with the legacy Internet, and the connection between the xTR of probe and PxTR is via a MPLS VPN. However, the MPLS tunnel did not support IPv6 at that time. Thus, we configure the ITR of the LISP-Lab probe natively forward the packets towards the Internet core for IPv6 targets (i.e., PETR is not used for IPv6). As a result, the IPv6 packets outgoing from this probe are

Table 6.1 Different configurations of probes in 2015

Name	using LISP	network type	probe/anchor
LISP-Lab	yes	academic	probe
mPlane	no	academic	probe
rmd	no	industrial	probe
FranceIX	no	industrial	anchor

not encapsulated into LISP packets and natively forwarded in the traditional way, but the returned packets still pass through the PITR of LISP-Lab platform.

The mPlane probe (#13842) resides in an academic network and uses the conventional routing. It allows to compare with non-LISP academic networks. While the rmd probe (#16958) resides in industrial network and also uses the conventional routing. It is chosen in order to compare with non-LISP and non-academic networks. Further, a stable probe with much more measurement capacity is necessary as a reference with all other probes. To this end, the only anchor in Paris named FranceIX (#6118) is selected. It resides in an IXP (Internet Exchange Point) network and does not use LISP. Thus, in this experiment, there are in total 4 probes used as sources to conduct the active measurements. Tab. 6.1 summarizes the different probes.

We are now in a setup phase, where we start with a reduced number of destinations so to first setup the automated experiments. In this first experiment, the selected 4 probes ping to the top 50 Alexa sites every 10 minutes during 6 hours. However, from the experiment we find that there are 14 websites resolved to the same IPv4 addresses. We filtered them, so the assessment presented in Sec. 6.3 are analyzed by the results of 36 top popular websites. With the collected dataset, we evaluate the various performances leveraged on RTT.

6.2.2 Dataset 2016

The new experiment still leverages on the LISP-Lab probe and RIPE Atlas infrastructure, but provides the following new contributions:

- Add another LISP probe connected to the LISP Beta Network as experimental source, to compare LISP-Beta Network and LISP-Lab.
- Enlarge the number of destinations to the first 500 most popular websites on the Alexa ranking [1].
- Consider the performance of IPv6 besides IPv4.
- Use *traceroute*, in complement to latency measurements, in order to have a deeper understanding of the observed behavior.
- Extend the experimental span to 15 days so to study potential periodicity of traffic.



Fig. 6.2 Locations of probes and anchor in 2016

Table 6.2 Location of LISP Beta Network PxTRs in 2016

Number	Continent	Country
1	Europe	Netherlands
2	Europe	Denmark
3	Europe	Norway
4	America	US
5	America	US
6	Asia	Japan

The newly added LIP6 probe (#2403) also connects to academic network behind both LISP Beta Network and LISP-Lab platform. When this probe communicates with the hosts in legacy Internet, it uses PETR of LISP-Lab by configuration, while return traffic goes through one of 6 PITRs of LISP Beta Network according to the BGP behavior. Depending on the BGP announcement to the legacy host, a different PITR is used. The more precise location of PITRs on LISP Beta Network is shown in Tab. 6.2: 3 are in Europe (Netherlands, Denmark and Norway), 2 are in US, and 1 is in Asia (Japan). Both LISP probes have a MPLS tunnel from their ITRs to the PETR at Lyon for IPv4, while for IPv6, the ITR of LIP6 probe is configured to use normal BGP routing to reach the PETR in Lyon (i.e., PETR is used but without MPLS VPN). The use of the MPLS tunnel IPv4 packets experience a shorter path compared to the IPv6 BGP-based one. Thus, theoretically the IPv4 packets sending from the LIP6 probe should arrive to PETR faster than those of IPv6. The configuration of two LISP probes are listed in Tab. 6.3.

Since the rmd probe is not configured with an IPv6 address, we replace it by the Gandi probe (#3141), which also resides in an industrial network and also uses the conventional

Table 6.3 Configuration of two LISP probes in 2016

Probe	PETR in Lyon	PITR
LISP-Lab (IPv4)	Via MPLS Tunnel	Lyon (Via MPLS Tunnel)
LIP6 (IPv4)	Via MPLS Tunnel	LISP Beta Network
LISP-Lab (IPv6)	not used	Lyon
LIP6 for (IPv6)	Via BGP Routing	LISP Beta Network

Table 6.4 Different configurations of probes in 2016

Name	using LISP	network type	probe/anchor
LISP-Lab	yes	academic	probe
LIP6	yes	academic	probe
mPlane	no	academic	probe
Gandi	no	industrial	probe
FranceIX	no	industrial	anchor

routing for both IPv4 and IPv6. As the mPlane probe and the FranceIX anchor satisfy all the requirements of this experiment, we still keep them as the reference. The locations of all the probes and anchor are shown in Fig. 6.2. Tab. 6.4 shows the different configurations of probes.

What we care about is the delay performance and the path of LISP interworking with legacy Internet. Thus, we rely on the ping and traceroute tools provided by RIPE Atlas. As destinations of our measurements, we selected the 500 most popular websites (according to worldwide website ranking provided by Alexa [24]), which reply to IPv4 ping and traceroute. Among them, 122 websites are configured with IPv6 address. Thus, in our experiment, 500 IPv4 and 122 IPv6 addresses are the destinations of ping and traceroute. We develop a Python script to schedule the experiment campaign by using the API provided by RIPE Atlas. The parameters of experiment campaign are as follows: the experiment campaign lasts 15 days (from December 15^{th} to 29^{th} 2016). The 5 chosen probes ping and traceroute to 500 IPv4 and 122 IPv6 addresses. The intervals of ping and traceroute measurements are respectively 30 minutes and 60 minutes. We want to evaluate LISP interworking performance as frequently as possible, but each probe sequentially launches 622 traceroute measurements, which last more than 30 minutes. To avoid the heavy traffic burden and guarantee that all the measurements in a same experimental round can be finished before next round, we set the sampling interval of traceroute at 60 minutes. As a summary, the results presented in this dissertation come from the following 4 experiment campaigns:

- 5 probes ping to 500 destinations during 15 days with an interval of 30 minutes (IPv4).

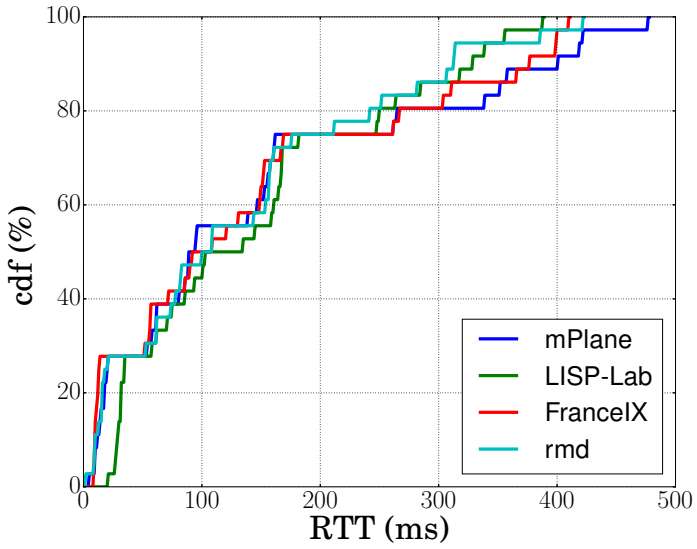


Fig. 6.3 CDF of average RTT between different probes from Dataset 2015

- 5 probes ping to 122 destinations during 15 days with an interval of 30 minutes (IPv6).
- 5 probes traceroute to 500 destinations during 15 days with an interval of 60 minutes (IPv4).
- 5 probes traceroute to 122 destinations during 15 days with an interval of 60 minutes (IPv6).

From the experiment, we find that some websites actually resolve to the same IPv4 (34 over 500) or IPv6 (47 over 122) addresses. Further, for some destinations, at least one probe (mainly the LIP6 probe) does not get any responses of *ping* during the whole experiment (42 for IPv4 and 10 for IPv6). In this chapter, we filtered the anycast destinations and only consider responding destinations. Thus, after cleaning the traces for the above reasons, all the results of *ping* measurements used in the following subsections consists on the 5 chosen probes as the experimental sources and 424 IPv4 and 65 IPv6 responding addresses as the destinations. Differently from the ping measurements, for the *traceroute* dataset, all the successful *traceroute* responses from each probe are kept for further analysis.

6.3 IPv4 Ping results from Dataset 2015

Fig. 6.3 shows the cumulative distribution function (CDF) of average RTTs towards the selected 50 targets. Since the experimental destinations are located all over the world, the range of observed RTTs varies from few milliseconds (ms) to nearly five hundreds ms. When the RTT values are in the small range, especially when RTT is less than 50 ms, the

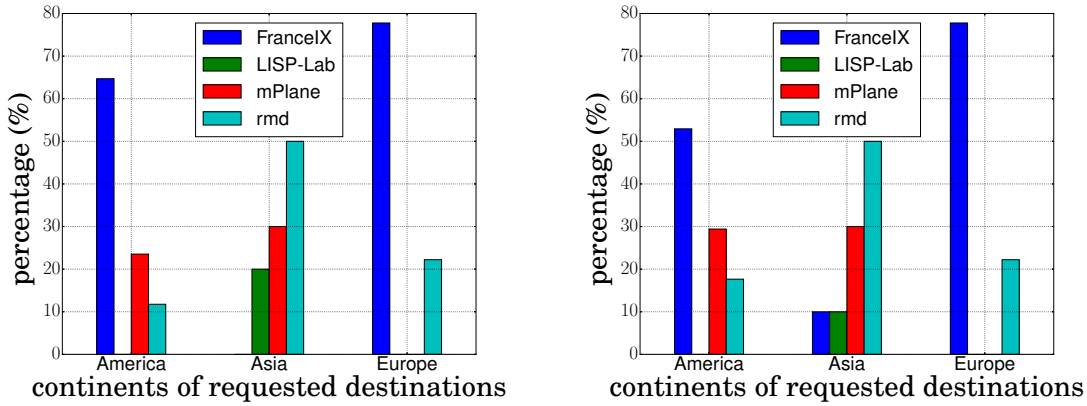


Fig. 6.4 Smallest mean RTT (left) and smallest median RTT (right) grouped by continent from Dataset 2015.

latency from LISP-Lab probe is always higher than the three others and the difference is mainly around 10 ms. This latency difference is caused by the stretch introduced by the proxy technology (cf. Sec.2.2), since every packet sent between LISP-Lab probe and the legacy Internet has to pass through the LISP-Lab PxTR, which is located in Lyon (approximately 400*km* away from the probe). As the RTT increases, actually the difference (surprisingly) decreases. When the RTT around 200 ms is reached, all probes show basically the same RTT. Around such RTT values, the destinations concentrate in North America. Going further in the high range RTT values, i.e., more than 350 ms, when destinations are mainly located in Asia, the probe in the LISP-Lab domain actually shows the lowest RTT. Thus, it indicates that the network connection between the LISP-Lab PxTR to the (asian) intercontinental destinations has a better performance and the stretch delay from LISP-Lab probe to PxTR can be ignored. We wanted to use traceroute to further investigate the reasons, but it cannot be natively used at that moment, since the LISP-encapsulation prevents its correct function. We explored a new way to find out what happens for Dataset 2016, which is described in Sec. 6.6.

We tried to quantify the percentage of times that one probe's RTT is the smallest compared to three others, since the destinations that the different probes contact to are exactly the same. The result shows that in 52.8% of the time, FranceIX shows the smallest RTT to reach the destinations. It is normal that its RTTs are the smallest, since FranceIX is an Internet Exchange Point (IXP), hence well connected, and also acts as one of the anchors of RIPE Atlas, thus with a more powerful hardware. While, only in the 5.6% of the cases, the LISP-Lab probe has the smallest RTT. In this small percentage the main contribution comes from Intercontinental destinations, i.e., when the RTT values belong to high range.

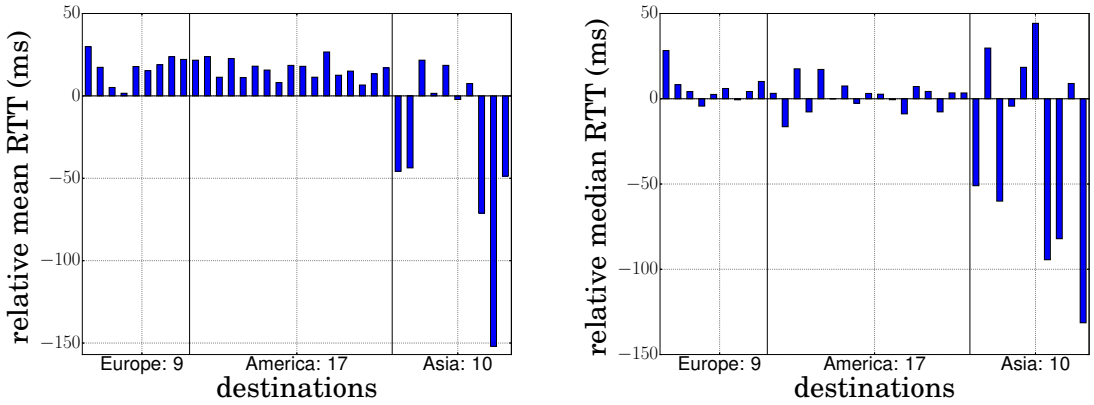


Fig. 6.5 Relative mean (left) and median (right) RTT clustered by different continents from Dataset 2015.

Fig. 6.4 depicts the percentage of times that one probe's RTT is the smallest compared to three other probes grouped by continents where the selected targets are located. When the destinations are in Europe and America, FranceIX is the fastest most of the time. Whereas, the percentage of LISP-Lab is always 0. Its higher RTT is caused by the proxy stretch. When the targets are in Asia, LISP-Lab becomes the fastest with a percentage of 20% (mean RTT) and 10% (median RTT). It indicates that such connection from the LISP-Lab PxTR is faster, so that the stretch can be ignored. The performance of FranceIX is not very stable to the Asian destinations, being the fastest 0% in average, but it is 10% looking at the median RTT. It shows that FranceIX sometimes has extremely high RTT values to Asian destinations.

We also evaluate whether the performance of LISP-Lab is as stable as FranceIX. We define a metric called *Relative RTT (rRTT)* for each destination as:

$$rRTT_{LL}(d) = RTT_{LL}(d) - RTT_F(d) \tag{6.1}$$

where d is the destination, subscriptions LL and F respectively refers to LISP-Lab and FranceIX. The results clustered by continents are shown in Fig. 6.5. The left hand one shows the mean RTT, while the right hand one shows the median RTT. For the European and American targets, LISP-Lab is a little slower than FranceIX but with a stable behavior. On the contrary, for half of the Asian destinations, LISP-Lab is significantly faster than FranceIX. It shows that the network connection between LISP-Lab PxTR and Asian destinations has better performance. Comparing the two subfigures, there is no negative values at all in Europe and America area in left hand figure, but there are some in the right hand one. It indicates that LISP-Lab RTTs to these destinations are very unstable and the variance is quite high.

Table 6.5 Correlation coefficient to FranceIX from Dataset 2015

	LISP-Lab	mPlane	rmd	FranceIX
Coefficient	0.9733	0.9784	0.9646	1.0

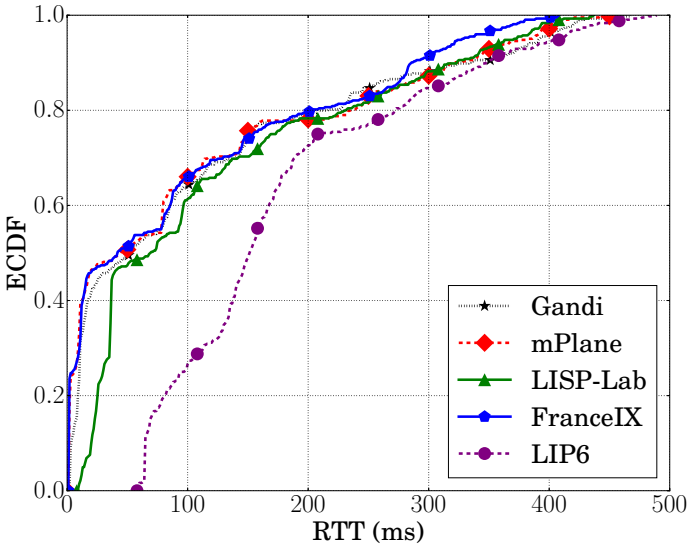


Fig. 6.6 CDF of median RTT between different probes (IPv4) from Dataset 2016

Since FranceIX shows the best RTT most of the time, we tried to evaluate the correlation between the RTT of the other 3 probes and FranceIX. The purpose is to see if RTT measurements are correlated or totally independent. We compute the correlation coefficient for every RTT series of different probes to all destinations with the ones of FranceIX. Tab.6.5 shows the results. The absolute value of coefficient is 1 in the case of a perfect direct linear relationship, whereas 0 in the case of no correlation at all. The correlation coefficient of LISP-Lab is 0.9733 (> 0.8), showing a very high correlation with FranceIX, like the other 2 probes. This means that LISP while certainly introducing some overhead, has quite stable performance that does not deviate from normal network operation.

6.4 IPv4 Ping results from Dataset 2016

Fig. 6.6 shows the CDF of the average RTT toward the selected IPv4 destinations for Dataset 2016. As an anchor, the FranceIX probe outperforms other probes in most cases, hence, showing the smallest RTT. With RTT range $[0, 200]$ ms, the latency from two LISP probes, LISP-Lab and LIP6, are respectively 25 ms and 60 ms slower than three other non-LISP probes. Such a RTT performance degradation is caused by the path length stretch introduced

by the the proxy technology (as mentioned in Sec. 2.1), since every packet sent between LISP probe and the legacy Internet has to pass through a PxTR. Although both LISP probes introduce the stretch, LIP6 probe is even 50ms slower than the LISP-Lab probe. Given that both LISP-Lab and LIP6 probes pass through the same PETR located in Lyon, the performance degradation of LIP6 probe is due to the PITR selection for traffic from legacy Internet to LISP network. The PITRs used by LIP6 probe belongs to the LISP Beta Network and there are in total 6 available PITRs. The PITRs announce the LISP Beta prefixes to their AS peers using BGP. So different destinations select different PITRs depending on where the destinations are and lead to the reply goes through the different paths. For example, an Asian destination normally uses the PITR in Japan instead of selecting one in Europe. As listed in Tab. 6.2, LISP Beta Network has not deployed PITR in France, while the PITR of LISP-Lab is in the same country to its probe. Thus, the return path that we measured for the LIP6 probe is longer than for the LISP-Lab probe. As a result, the closer the destination is located to the probes, the bigger the difference of RTT between the two LISP probes.

Within RTT range $[200, 500]$ ms, all probes except LIP6 show basically the same RTT. It indicates that LISP-Lab has a better performance in scenario of long distance network connection. The reason is that stretch delay from LISP-Lab probe to PxTR (in Lyon) can be ignored. Fig. 6.6 confirms this situation, within RTT range $[200, 500]$ ms, the difference between two probes becomes significantly smaller than that in range $[0, 200]$ ms.

The aforementioned results and analysis coincide with the Fig. 6.3 from the dataset of 2015, which just covers comparison related to LISP-Lab probe. This indicates that the LISP PxTR performance is stable: the measurement results do not change with the different destinations at the different time.

Even with stretch brought by the PxTR, LISP probes still outperform non-LISP probes in some cases according to the measurement results. To identify these cases, worldwide destinations are classified by continent. Fig. 6.7 depicts the percentage of times that one probe has the smallest RTT compared to the other probes grouped by continents where the selected targets are located. The experimental destinations are located in 4 continents, i.e., Europe, North America, South America, and Asia. The *ping* measurement is repeated 720 times as the experiment lasts 15 days and its interval is 30 minutes. The RTTs of the FranceIX anchor and mPlane probe are the smallest most of the time, especially for the European and South American destinations. Only for the Asian destinations the RTT of LISP-Lab probe experiences sometimes the smallest latency for 8.82% of the targets, while the LIP6 probe experiences the smallest latency in 0.98% of the cases. Among the North American destinations, there is just one destination for which the RTT of the LISP-Lab probe is the smallest, while there is no destination for which LIP6 experiences the smallest

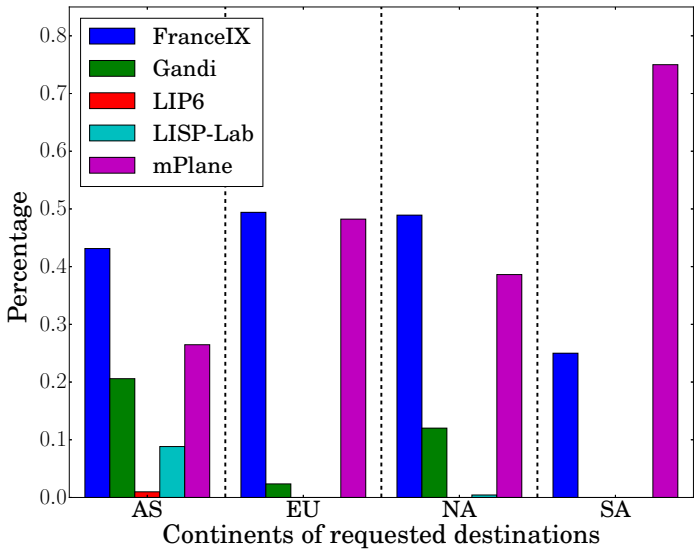


Fig. 6.7 Smallest median RTT grouped by continent (IPv4) from Dataset 2016.

Table 6.6 Correlation coefficient to FranceIX (IPv4) from Dataset 2016

	Gandi	mPlane	LISP-Lab	LIP6	FranceIX
Coefficient	0.9647	0.9707	0.9766	0.8547	1.0

latency. Fig. 6.6 shows that for RTT above 200 ms LISP-Lab performs like the other probes, so that we can question whether it actually has the best performance in some cases, and for which destinations. Fig. 6.7 shows the geographic location of destinations for which LISP-Lab actually is the best performing probe. In particular, the LISP-Lab probe has no additional delay compared to other non-LISP probes for the intercontinental transmission to the Asian destinations. It means that the negative effect of LISP stretch delay is evident for communications with European and American destinations, but can be neglected for Asian intercontinental destinations. This phenomenon is similar to the result shown in [79], which shows only in 5.6% of the cases that the LISP-Lab probe experiences the smallest latency and all these destinations are in Asia. It proves again that the negative effect of LISP PxTR can be ignored for the intercontinental transmission to the Asian destinations (when considering the European vantage point).

Also for the Dataset 2016, we try to evaluate the correlation in term of RTT between FranceIX and other probes. The correlation coefficient of LISP-Lab is 0.9766 (> 0.8), showing a very high correlation level with FranceIX, and even higher than the correlation coefficient of all the other probes. The correlation coefficient of LIP6 is 0.8547, shows a high correlation with FranceIX as well. Both LISP probes having a high correlation with FranceIX means that while certainly introducing some overheads, LISP has quite stable per-

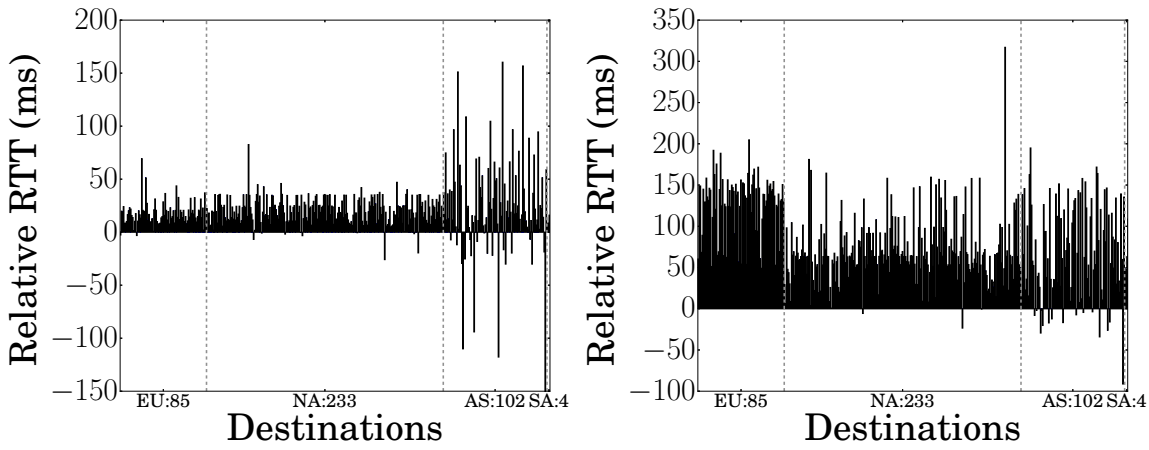


Fig. 6.8 IPv4 Relative median RTT clustered by different continents for LISP-Lab (left) and LIP6 (right) from Dataset 2016

formance that does not deviate from normal network operation. Such result has two sides for the performance of LISP. Good because LISP shows stability. Bad because it means that LISP experiences the same performance and failures like the normal Internet.

Another interesting point is to assess whether the two LISP probes have a stable performance as FranceIX according to the destinations. To this end, we leverage on the same metric of *Relative RTT (rRTT)* (formula 6.1), defined in Sec. 6.3 for each destination as:

$$rRTT_{LL}(d) = RTT_{LL}(d) - RTT_F(d) \tag{6.2}$$

$$rRTT_{L6}(d) = RTT_{L6}(d) - RTT_F(d) \tag{6.3}$$

where d is the destination, subscriptions LL, $L6$ and F respectively refers to LISP-Lab, LIP6 and FranceIX. The results clustered by continents are shown in Fig. 6.8, where the relative RTT for LISP-Lab probe is on the left and the relative RTT for LIP6 probe is on the right. The positive relative RTT indicates that FranceIX is faster, while the negative values indicating that LISP-Lab or LIP6 are faster. In the left side of Fig. 6.8, for the European and American targets, the values are between 10 ms and 20 ms in most cases (i.e., 31.37%), showing that LISP-Lab is a little slower than FranceIX but with a stable behavior. On the contrary, for some Asian destinations, LISP-Lab is significantly either faster (in 20.6% of the cases) or slower than FranceIX and the largest difference reaches 150 ms. It shows as well that for the European and American targets, LISP-Lab probe is stable compared to FranceIX, since the difference is mainly caused by the transmission delay between Paris (location of probe) and Lyon (location of PxTR). But when considering the Asian destinations, LISP-Lab does not show a very stable behavior, where the difference might be very large,

Table 6.7 Reliability of each probe (IPv4) from Dataset 2016

	Gandi	mPlane	LISP-Lab	LIP6	FranceIX
Reliability (%)	99.66	99.65	99.43	77.78	99.62

since the path diversity for long-distance intercontinental transmission is higher, with every path having very different performance. The right figure of Fig. 6.8 shows that for LIP6 probe the difference is almost always positive. For the European destinations, the average relative RTT is around 130 ms and much higher than those in the other 3 continents. All European targets show that the LIP6 probe is significantly slower than the FranceIX anchor. For the American destinations, most relative RTT decreases to around 50 ms but some of them are around 150 ms. For the Asian targets, the relative RTT presents big differences. Some still stay at around 150 ms, some drop to below 50 ms, and there are 5 destinations even show negative relative RTTs. The reasons causing the relative RTTs for LIP6 probe varying a lot are similar to that for LISP-Lab but not totally the same. The same point is that the longest transmission of traffic in our experiment are the destinations located in Asia. Since the stretch can be ignored in the case of intercontinental transmission, both of two LISP probes are not always slower than FranceIX to Asian targets. The different performance between LISP-Lab probe and LIP6 probe is caused by the location of each PxTR. Although the PxTR of LISP-Lab is not close to its probe, but at least they are in a same country. While the LIP6 probe uses the same PETR to LISP-Lab probe but its PITR is not in France, and is even quite far away. As a result, for the shorter transmission, i.e., to the European destinations, the LIP6 probe introduces extremely higher RTTs than the non-LISP probes and even higher than LISP-Lab probe. Thus, the location of PxTR is very important. The PxTR near to either the sources or the destinations can obviously decrease the stretch. At least the probe having a PxTR in the same country shows a better performance compared to the PxTR just in the same continent.

We also want to evaluate the *Reliability* when having the response to each ping measurement. We measure reliability, in our experiment, by simply calculating the percentage of replies over the total number of requests. If every ping measurement is successful, i.e., there is a response having the RTT value for a probe at every experiment round, its reliability is 100%. As shown in Tab. 6.7, except for the LIP6 probe with 77.78%, all the other probes are more than 99%. The lower reliability of LIP6 probe indicates that sometimes its ping measurement is not successful. Since it is much lower than the others, to make sure that the LIP6 probe works well and there is no congestion or misconfiguration on the probe or any of its connected routers, we conduct another experiment letting the LIP6 probe use a normal public IPv4 address pinging the same 500 destinations. In this case, the LIP6 probe

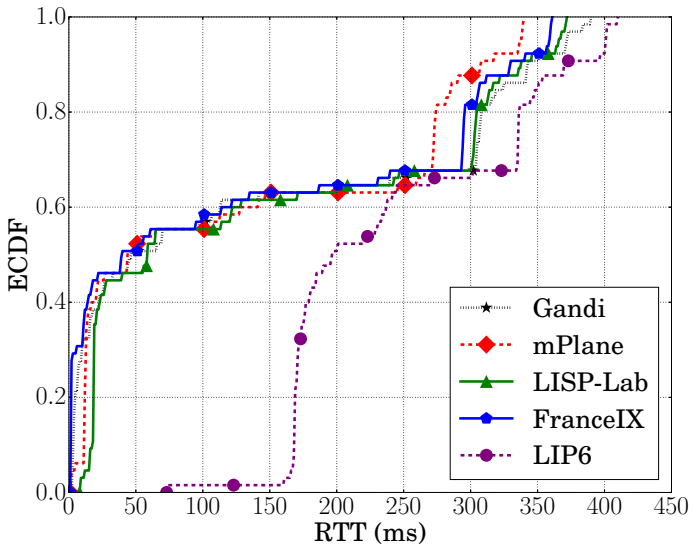


Fig. 6.9 CDF of median RTT between different probes (IPv6) from Dataset 2016

is non-LISP-speaking. The results show that the reliability of LIP6 probe is 98.91%, which is very close to the other probes. Thus, we eliminate the possibility of the problem on the LIP6 probe itself. As the LIP6 probe uses the same PETR to the LISP-Lab probe, while the reliability of LISP-Lab is very high, the difference is caused by the PITRs. The LISP-Lab probe uses the PITR of LISP-Lab platform but the LIP6 probe leverages one of LISP Beta Network. On one hand, LIP6 probe has longer transmission distance causing more risk of losing the packets and thus leads to more losses. On the other hand, we can conclude that the performance of LISP-Lab PxTR is more reliable than that of LISP Beta Network.

Since the experiment in 2015 just lasts 6 hours and shows no periodicity. Thus, this experiment campaign is extended to conduct with a span of 2 weeks to assess whether the RTT measurements have periodicity. Two methods are used to check: Fast Fourier transform (FFT) and Auto-correlation. However, the analysis results also show that none of RTT series from probes to any destinations exhibits periodicity. It means that the traffic does not periodically fluctuate with a certain interval.

6.5 IPv6 Ping results

In this section, LISP interworking performance of IPv6 is evaluated with the same metrics as those used in Sec. 6.4.

The CDF of average RTTs to the selected IPv6 destinations is shown on Fig. 6.9, which is generally similar to the Fig. 6.6 for IPv4. The FranceIX anchor still shows the best

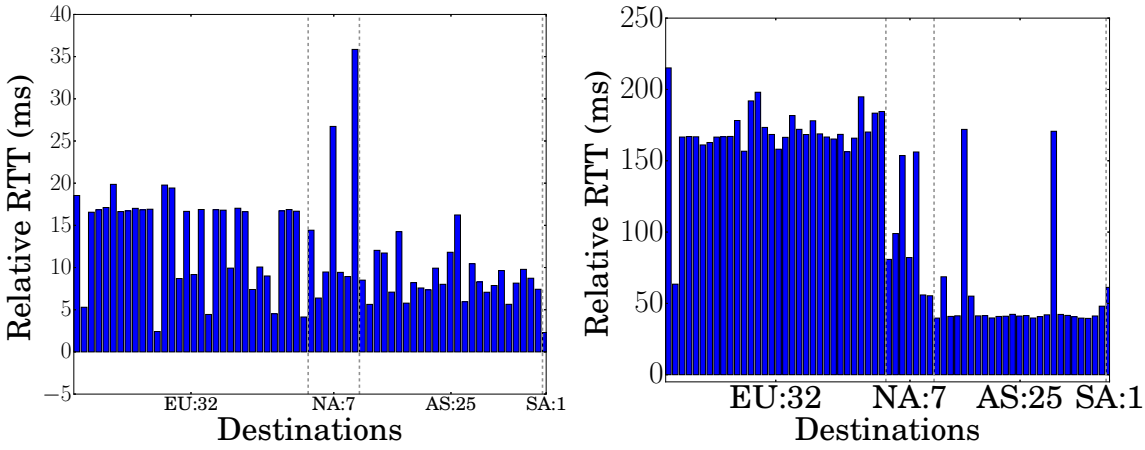


Fig. 6.10 IPv6 Relative median RTT clustered by different continents for LISP-Lab (left) and LIP6 (right) from Dataset 2016

performance with the smallest delay for almost all the destinations. The RTT of LISP-Lab probe is always higher than the other non-LISP probes with RTT range $[0, 30]$ ms and the difference is around 7 ms. The RTTs are almost the same when the values are more than 30 ms. Hence, the performance of IPv6 is similar to IPv4 that the stretch introduced by PxTR is significant when the range of RTT value is small, but can be ignored when the range of RTT value becomes large.

However, in the experiment for IPv6 targets, the traffic produced by LISP-Lab probe are natively forwarded to the destinations instead of going to the PETR in Lyon first, thus the RTT difference compared to the other non-LISP probes becomes smaller compared to those for IPv4. But the latency still exists, caused by the returning traffic that still need to pass through a PITR. Since the latency decreases, it is easier to be ignored for IPv6. That is the reason why the stretch can be ignored when RTT values are higher than just 30 ms for IPv6 while it needs to be higher than 200 ms for IPv4. While the LIP6 probe is always the slowest compared to the other probes, the difference is even 160 ms to the non-LISP probes and 150 ms to the LISP-Lab probe. The reason having such a big difference is not only caused by the bidirectional traffic should pass through the PxTR, but also caused by the route between xTR of LIP6 probe to the PETR at Lyon, which has no tunnel indicating the used routing path longer than the case of having an MPLS tunnel. Similar to IPv4, when RTT values are higher than 250 ms, the difference decreases to 40 ms in most cases.

To further explore why two LISP probes are slower, especially to understand why the LIP6 probe has an extremely large latency compared to the others, we leverage on the Relative RTT between both LISP probes and the FranceIX anchor grouped in four continents as mentioned in Sec. 6.4. In the left figure of Fig. 6.10, which is the relative RTT between

LISP-Lab and FranceIX, all the values are positive, indicating that the LISP-Lab probe is slower than the FranceIX anchor no matter to which destination. For the European and North American targets, the relative RTTs are mostly between 5 ms and 17 ms. But the average relative RTTs decrease to 7 ms for the 25 Asian destinations. It proves the fact that for the nearby transmission, the delay introduced by PxTR is significant, while it effects less for the intercontinental transmission to the Asian destinations. What's more, the absolute relative RTT values for IPv6 are smaller than those for IPv4, because there is only returning traffic passing PITR for IPv6 so to reduce the latency. Thus, natively forwarding is faster than using PETR, but not so much. From the Fig. 6.10 we also know that there are 32 destinations located in Europe and just 7 in North America. It explains the reason why the curve of CDF for LISP-Lab sharply increases when the RTT is in the small range and mixes with the other non-LISP lines so quickly. It is because a half of destinations are not far away from the probes. Thus, a high percentage (around 50%) consists of small RTT values and these RTTs of LISP-Lab have a consistent delay compared to the other non-LISP probes. The mix part is mainly produced by the Asian destinations, to which the RTT values are large and the delay introduced by only the PITR of LISP-Lab is less significant. The relative RTT for the LIP6 probe is shown on the right hand side of Fig. 6.10. The relative RTTs are all positive and much higher, indicating that the LIP6 probe is always quite slower than the FranceIX anchor to whichever target. Especially for the European destinations, the values are around 170 ms. But to the North American and Asian targets, the relative RTTs are obviously smaller. As mentioned in Sec. 6.2.2, the outgoing packets from LIP6 probe still need to go to PETR first and especially the path between its xTR and PETR has no MPLS VPN tunnel. Besides, the main reason that the higher relative RTT values for IPv6 than IPv4 is caused by the LIP6 probe using the PITRs of LISP Beta Network. Different from IPv4 [3], there are only two ASes of IPv6 PITRs and both of them are located in US [4]. As a result, no matter from which destination, all the returning packets need to pass the PITRs in US first and then forward back to the LIP6 probe located in Paris. By consequence, the relative RTTs to the North American destinations are extremely smaller than those to European destinations. Although the traffic returning back from the Asian destinations also pass through the PITRs in US, the distance between the source and destination is quite far, thus it does not effect a lot to the relative RTTs for Asian targets.

Further, in the 84.61% of cases the FranceIX anchor has the smallest RTT in the IPv6 experiment. The LISP-Lab probe and LIP6 are not the fastest to any destinations, even to the Asian targets. It shows that the LISP performance of IPv6 is a little bit worse than IPv4. Since FranceIX has the smallest RTTs in most situations and as an anchor, its higher measurement capacities lead to its higher stability, we also use the correlation between the

Table 6.8 Correlation coefficient to FranceIX (IPv6) from Dataset 2016

	Gandi	mPlane	LISP-Lab	LIP6	FranceIX
Coefficient	0.9565	0.968	0.967	0.3734	1.0

Table 6.9 Reliability of each probe (IPv6) from Dataset 2016

	Gandi	mPlane	LISP-Lab	LIP6	FranceIX
Reliability (%)	98.43	99.97	99.98	82.24	99.99

RTT of the other 4 probes and FranceIX to see if RTT measurements of each probe are as stable as FranceIX. As shown in Tab. 6.8, the correlation coefficient of LISP-Lab is 0.967, which is almost one, indicating although LISP over IPv6 has higher latency than FranceIX caused by the introducing of PxTR, the performance is still stable. However, the correlation coefficient of LIP6 probe is just 0.3734. It is higher than 0.2, but much lower than 0.8, showing that the LIP6 probe is not totally independent from FranceIX, but it has quite low correlation. In fact, the IPv6 RTT series of LIP6 fluctuate a lot over time during the experiment and are not stable.

The Reliability of every probe for IPv6 is almost the same of IPv4 as shown in Tab. 6.9 except for the LIP6 probe. The higher reliability of LISP-Lab probe than the LIP6 probe confirms again that even for IPv6, the PxTR of LISP-Lab is still more reliable than that of LISP Beta Network, although the latter one has 6 world-wide PxTRs in total, from which only 2 can be used for IPv6. The reliability of LISP-Lab probe for IPv6 is a little bit higher than itself for IPv4, showing that using PxTR may decrease the reliability but the effect is very small. The reliability of the LIP6 probe for IPv6 is also higher than itself for IPv4, indicating that the IPv6 performance of PxTR on LISP Beta Network is better than IPv4, mainly thanks to the use of PxTRs in US. Thus from this comparison, we get a conclusion that the two PxTRs in US are more reliable than the other 4 PxTRs of LISP Beta Network.

The periodicity check for IPv6 is also conducted, but no probes show that their RTT measurements to any destinations have periodicity, either. It indicates that the latency of traffic has no relationship with the specified experiment time, i.e., the RTT measurements generally have the same results regardless the experiment time and duration. Thus, the results of these experiments are not occasional phenomenon with the special results, they are reproducible instead.

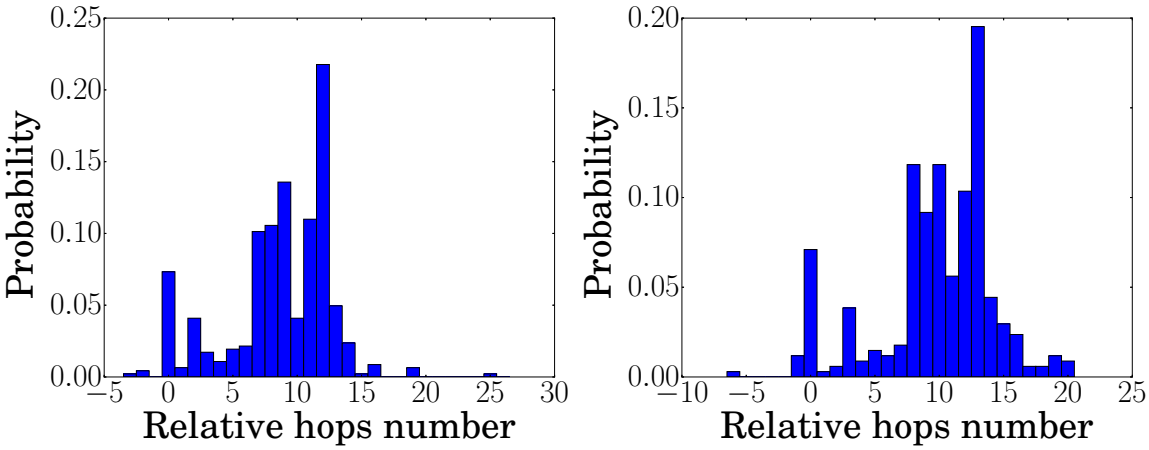


Fig. 6.11 Distribution of IPv4 Relative Hops Number for LISP-Lab (left) and LIP6 (right) from Dataset 2016

6.6 Traceroute-related results

In this section, we use hops number obtained via traceroute from sources (i.e., probes) to destinations as the metric to further evaluate the LISP performance and also further investigate the reasons for the results presented in the ping-related experiments. Since the traceroute can only present the outgoing path, i.e., the routing path from probes to destinations, we focus on the use of the PETR and existence of a VPN between the xTR and the PETR.

When the IPv4 targets are in Europe and America, there are very few cases that the hops number of LISP-Lab or LIP6 is the smallest. Precisely, there is 7.65% of cases that LISP-Lab has the shortest path by hops number to the Asian destinations and 6.63% of cases that LIP6 has the shortest path in the experiment. This result coincides to the percentage of the smallest RTT shown in Fig. 6.7. This consistent relationship between smallest RTT and shortest path exists as well for the IPv6 experiment, where there is no destination at all for which the hop number or the RTT are the smallest (for both LISP-Lab and LIP6 probes). For IPv6, it is FranceIX or Gandi that always have the shortest path to all the targets.

As FranceIX is always the most stable probe and has the shortest path in most cases, we define *Relative Hops Number (rHN)* clustered by different continents to look at the difference between the hops number of LISP probes and FranceIX. The definition is as:

$$rHN_{LL}(d) = HN_{LL}(d) - HN_F(d) \tag{6.4}$$

$$rHN_{L6}(d) = HN_{L6}(d) - HN_F(d) \tag{6.5}$$

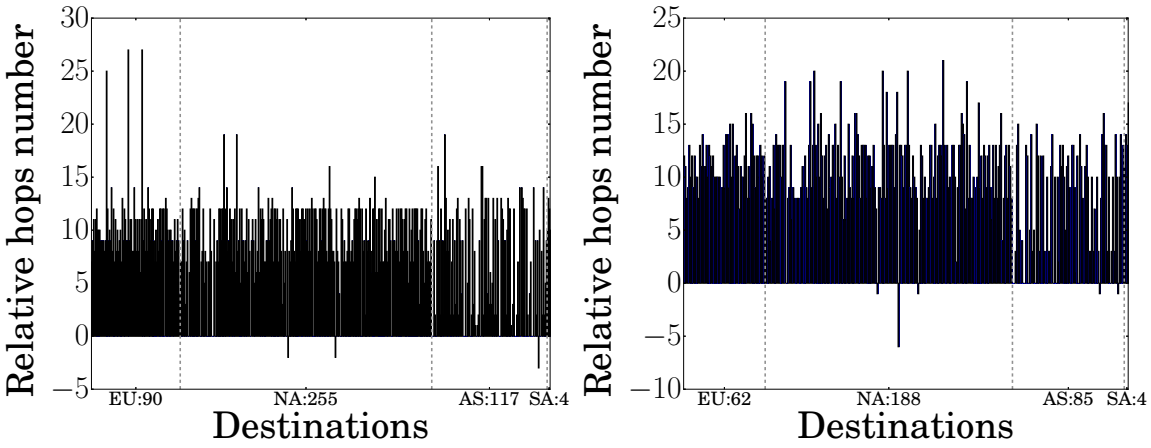


Fig. 6.12 IPv4 Relative hops number clustered by different continents for LISP-Lab (left) and LIP6 (right) from Dataset 2016

where d is the destination, subscriptions LL, $L6$ and F respectively refers to LISP-Lab, LIP6 and FranceIX. The positive value indicates that the hops number of FranceIX is smaller than LISP probes. Fig. 6.11 provides an IPv4 distribution of relative hops number for LISP-Lab on the left and for LIP6 on the right. This figure shows the probability of each relative hops number. For the LISP-Lab probe, the most common relative hops number is 12 with a percentage of 21.77%, hence remaining limited in most of the cases. From the ping-related experiment in Sec. 6.4 and Sec. 6.5 we know that using PxTR introduces some overhead. So the Relative Hops Number being generally positive is reasonable. The very large relative hops number just appears once for 19 and 25, so we regard them as outliers. Similar to the LISP-Lab case, for the LIP6 probe, the most frequent relative hops number is 13 in 19.53% of cases.

In order to understand where the most common relative hops number appears, we produce the relative hops number for each destination clustered by different continents in Fig. 6.12 to complete Fig. 6.11. The left figure is for the LISP-Lab probe, indicating that the relative hops number is almost the same for the European and North American destinations and contributes to the relative hops number with high probabilities in Fig. 6.11. The small relative hops number is mostly produced from the Asian targets. Whereas the outliers are all from the European destinations. The right figure of Fig. 6.12 is for the LIP6 probe, mainly presenting the same result as the LISP-Lab probe, except that the latter rarely has the relative hops number higher than 16, but LIP6 has more.

With the same evaluation metrics of IPv4, the distribution of Relative Hops Number and the Relative Hops Number clustered by different continents for IPv6 are respectively shown in Fig. 6.13 and Fig. 6.14. The left figure of Fig. 6.13 presents that the most common rel-

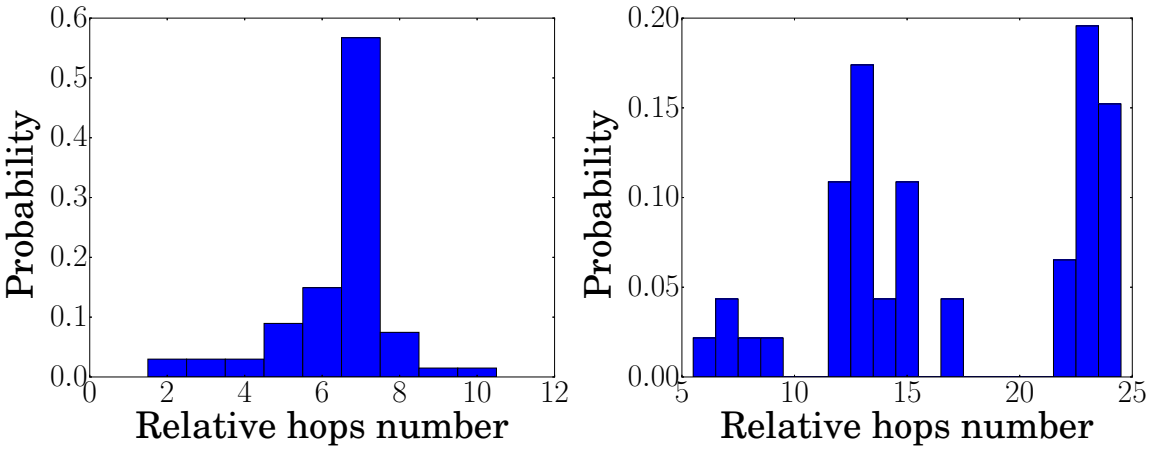


Fig. 6.13 Distribution of IPv6 Relative Hops Number for LISP-Lab (left) and LIP6 (right) from Dataset 2016

ative hops number between the FranceIX anchor and the LISP-Lab probe is 7 in 56.72% of cases, more than a half, much higher than the probability of other relative hops numbers. The range of relative hops number is smaller compared to IPv4 and the most common relative hops number is also smaller, since the traffic does not need to pass the PETR, but is just natively forwarded so the overhead is reduced. The left figure of Fig. 6.14 reveals the truth that 7 relative hops numbers are produced by the majority of targets regardless where the destinations are, instead of like the one of IPv4, where the relative hops number for European and American destinations are more than the Asian's. As the relative hops number from PETR is only 3, but 7 for natively forwarding, it is likely that the PxTR has shorter paths to most destinations, leads to the decreasing of hops number for the LISP-Lab probe. The relative RTTs higher for European and American destinations than those for Asia shown in left figure of Fig. 6.10 are mainly caused by the returning path, since it is natively forward for outgoing and there is no difference by continents. For IPv6, the LIP6 probe has extremely high relative hops number in a range from 12 to 24. The majority of relative hops number is 23 in 19.57% of the cases and followed by 13 with a percentage of 17.39%. The higher relative hops number is caused by not using the VPN between xTR and PETR compared to the IPv4 case. The right figure of Fig. 6.14 visualizes the 23 relative hops number coming from the European destinations and 13 coming from the Asian targets. The form coincides to the relative RTT for IPv6 shown in the right hand part of Fig. 6.10, indicating that the high relative RTTs are caused by the high relative hops number.

Concerning the RTT performance of the LISP-Lab probe to Asian destinations, which sometimes shows even better than FranceIX, we analyze the AS-path (Autonomous System path) from FranceIX and LISP-Lab to Asian destinations that the packets traverse. Since

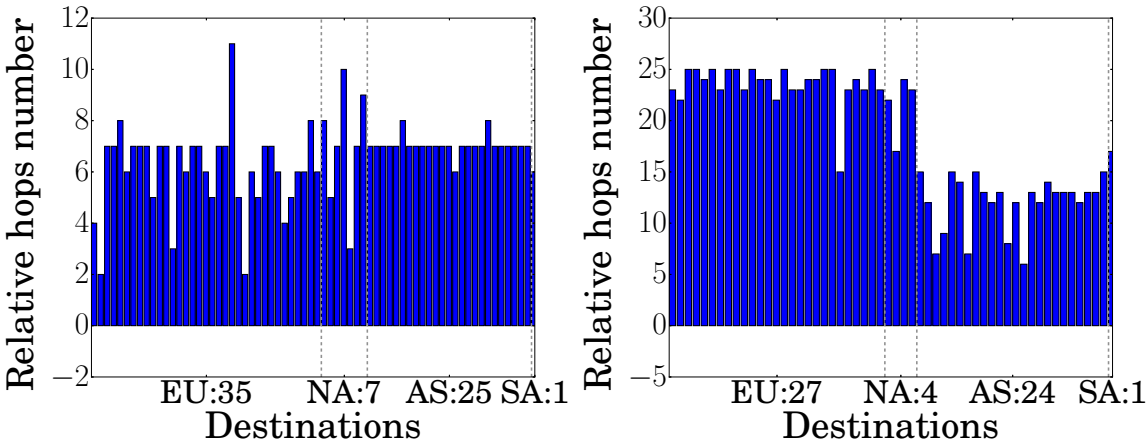


Fig. 6.14 IPv6 Relative hops number clustered by different continents for LISP-Lab (left) and LIP6 (right) from Dataset 2016

all the IPv4 packets from LISP-Lab probe are sent to PETR at Lyon first, so the AS-path discussed in the followings is actually from PETR to the Asian destinations. By comparing the AS-path, we find that FranceIX and LISP-Lab very often take different paths, with only the last 1 or 2 hops in common. Even further, by looking at the geographic location of each hop of *traceroute*, it can be observed that FranceIX and LISP-Lab take the path even in the different directions in most of the cases. In particular, there are 11 destinations out of 117 in total (i.e., 9.4%) for which the packets sent by LISP-Lab pass through US, while the packets sent by FranceIX pass through Eastern Europe. There are 68 destinations (i.e., 58.1%), which show exatly the opposite situation, i.e., packets sent by FranceIX pass through US, while packets sent by LISP-Lab pass through Eastern Europe. This last case is when LISP-Lab has smaller RTT (compared to FranceIX), almost all of the times (with only few exceptions).

6.7 Summary

In this chapter, we conduct a six-hours as well as a two-weeks real network measurements with LISP Beta Network, LISP-Lab platform and RIPE Atlas to provide a comprehensive performance evaluation of LISP interworking. Concretely, we provide a first thorough sight on the performance of LISP PxTR. Since the results of experiment in 2016 coincides to the one in 2015 and are more comprehensive. Thus, we conclude this chapter by presenting the observations of the Dataset 2016.

In our large scale measurement campaign, we take into account 5 probes as sources, 500 IPv4 and 122 IPv6 addresses as destinations, conducting ping and traceroute experiments.

We find that the PxTR indeed introduces the negative effects for the destinations located in Europe and America, but the negative impact of PxTR can be ignored for the intercontinental long-distance transmission to Asia destinations. From the experiment, the results show that the position of PxTR is very important. The PxTR either near to the sources or the destinations can decrease the latency a lot. Generally speaking, LISP is stable compared with the reference anchor FranceIX, except for the IPv6 performance of LISP Beta Network. Further, the performance of LISP-Lab PxTR is more reliable than the one of LISP Beta Network, although the latter has 6 worldwide PxTRs used for IPv4 and 2 located in US used for IPv6, whereas LISP-Lab has only 1 PxTR for both IPv4 and IPv6. Compared to leveraging on PxTR, natively forwarding without using LISP decreases the latency, but not much. The traceroute experiment shows that introducing PxTR of course brings more hops, but if the PETR is well configured, so to always have peers to the destinations, there are only 4 more hops compared to the packets being natively forward by xTR without encapsulating with LISP.

Chapter 7

Analysis of LISP mobility and ns-3 implementation

The *Locator/Identifier Separation Protocol* (LISP), due to its map-and-encap approach, can bring benefits to mobility. LISP Mobile Node (LISP-MN) is based on the basic LISP functionality to provide mobility across networks. Thus, LISP can be implemented either on the border routers or directly on the end hosts to manage mobility. However, there are no experimental results comparing the advantages and disadvantages of each solution. Assessing the LISP mobility performance needs the testbeds or the simulators. The basic LISP architecture is deployed on LISP Beta Network and LISP-Lab platform to offer the researchers a realistic experimental environment, but both do not support LISP-MN. Moreover, since the simulators can help researchers quickly verify the proposals and test new features, we implemented mobility extensions in a LISP simulator. This chapter introduces the implementation of LISP mobility extensions under ns-3.27 leveraging on an open source basic LISP simulator. In addition, this chapter analyzes the different LISP mobility scenarios from the respects of handover delay and overhead of LISP Control Plane. It describes the characteristics of each scenario.

The rest of chapter is organized as follows: Sec. 7.1 and Sec. 7.2 respectively introduces ns-3 and the existing LISP simulator on it. Sec. 7.3 analyzes the design and implementation of our prototype, and afterwards, Sec. 7.4 illustrates three different LISP scenarios supporting mobility, presents their traffic schema, modelizes the handover delay and overhead of LISP Control Plane, and compares their advantages and disadvantages. Sec. 7.5 provides some ideas of evaluation based on the proposed simulation for future work.

7.1 NS-3

ns-3 [19] is a popular and free discrete-event network simulator for networking research. It is developed completely in the C++ programming language. The ns-3 architecture is similar to Linux computers with application, TCP/IP protocol stack, network interface, sockets, etc. ns-3 is very well documented and has an active community which facilitate the researches to adapt ns-3 source code for their researches. Besides, ns-3 offers the possibility to visualize the simulation instance so to allow the users to visually confirm the packets flow as they expect.

7.2 Basic LISP implementation on ns-3

Simulation is becoming more important for deploying new technologies or as a proof of concept of new protocols. In the study of LISP, there exist few simulators based on OMNet++ [116, 115, 74] or based on Java [112]. However, these existing simulators are not open-source, which hinders other researchers to modify or adapt the simulator with respect to their own research purposes.

To our best knowledge, the unique open-source LISP simulator that we found in the literature is the one proposed by Agbodjan [31]. The authors implemented a basic LISP simulator under ns-3.24, but this work can be further polished. For example, the encoding of LISP Control Plane messages does not respect RFC 6830 [47] so that the Wireshark [29] can not resolve the captured results in the correct format. More importantly, its implementation has no support for LISP mobility. By leveraging the source code of Agbodjan, we implement an open source LISP simulator with LISP mobility extension. Besides, we also cover the shortage of the original source code. For example, the encoding of LISP control messages is according to RFC 6830 [47] so that Wireshark can correctly decode these messages for analysis. The case of Negative Map-Reply has been covered. Recall that the work of Agbodjan is still under ns-3.24, but ns-3.24 evolved lots from ns-3 to the latest version ns-3.27. This implementation is under ns-3.27, which allows the other researchers to profit the newest functionalities of ns-3 simulator.

7.3 LISP mobility extensions on ns-3

Our implementation respects LISP RFC 6830 [47] and LISP mobility standards [87]. As a design choice, we implement LISP and LISP mobility functionalities by modifying and extending two already existing modules of ns-3: *internet* and *internet-apps*, instead of creating

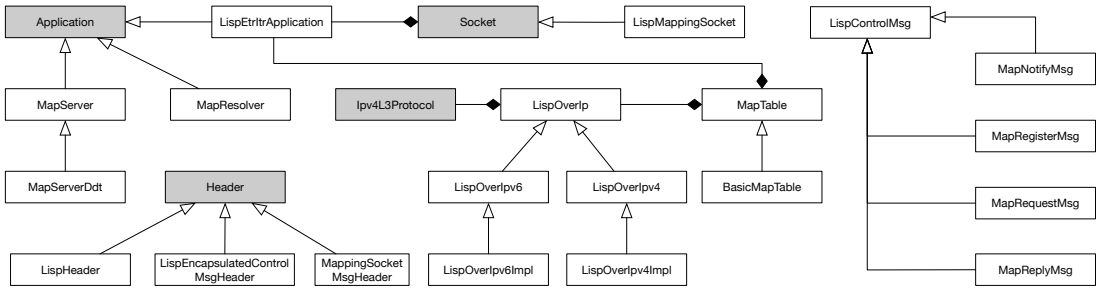


Fig. 7.1 UML diagram of LISP and LISP mobility implementation. The solid arrow refers to a composition relation, while the blank one refers to a inheritance relation.

a new independent module. The justification behind this design is that LISP protocol and legacy Internet module have an interdependent relationship: an IP layer packet is processed by LISP and then passed to IP protocol again. However, this kind mutually dependent relationship between modules is not supported by ns-3. Our implementation consists of two parts: LISP Data Plane and LISP Control Plane. The communication between LISP Data and Control Plane is achieved via a dedicated socket (i.e. *LispMappingSocket*) that inherits from ns-3 *Socket* class. The Data Plane implementation is in "kernel space" (i.e. ns-3's *TCP/IP stack*) and Control Plane is implemented in "user space" (i.e. ns-3 *Application*). Such a design is inspired by that of OpenLISP [110].

The UML diagram of proposed LISP and LISP mobility implementation is illustrated in Fig. 7.1. The darker blocks are classes already in ns-3, while the white blocks refer to the classes that Agbodjan [31] added into ns-3. It should be noted that our implementation keeps the same class names used in the Agbodjan's implementation. Except for the class *Header*, we rewrote the contents of all his classes to support LISP mobility, especially for the class *LispEtrItrApplication*, *LispOverIpv4Impl*, and *BasicMapTables*. This work currently only supports IPv4 at time of this writing. The IPv6 support (i.e. the implementation related to IPv6 such as *LispOverIpv6Impl*) is still in process. In addition, the authentication procedure involved in LISP is not considered in our implementation.

7.3.1 Implementation of LISP Data Plane

The implementation of LISP Data Plane mainly consists of *LispOverIp* and *MapTable* classes and their subclasses, along side with some auxiliary classes (e.g. *LispHeader*). In addition, to support LISP functionalites, *Ipv4L3Protcol*'s packet transmission, reception, forward and delivery procedures are accordingly adapted.

LISP database and cache

Each LISP-speaking node should maintain one LISP database and LISP cache for LISP encapsulation and decapsulation operations. In our implementation, both LISP database and cache are modelized by the same class *MapTable* that stores and manages EID-to-RLOC mapping information. This class is in charge of CRUD (Create, Retrieve, Update, Delete) operations for mappings. Each mapping entry in LISP database and cache is an instance of *MapEntry*. For the sake of flexibility, the class *MapTable* is an abstract base class. The CRUD methods are implemented in its subclass *BasicMapTable*. The mapping search operation is a straightforward iteration over LISP database or cache. It is possible that for other users to provide their own LISP database and cache implementation, for example, which uses more sophisticated mapping entry look up algorithms, by extending *MapTable* class.

In addition, *MapTable* has a callback which allows to send the buffered packet (either LISP Data Plane or Control Plane) due to LISP cache miss event, upon insertion of the required EID-to-RLOC mapping information into LISP cache.

Implementation of LISP encapsulation and decapsulation

To integrate LISP and LISP mobility into conventional Internet protocol stack, one key technical difficulty is that *Ipv4L3Protcol* should be able to determine when passing a packet being processed to LISP-related procedure and how to retrieve the associated mapping information. To this end, a new class called *LispOverIp* and its extended classes are added to ns-3 *internet* module. This class is in charge of checking whether it is necessary to do LISP-related operations (*NeedEncapsulation()*, *NeedDecapsulation()*), and encapsulating conventional IP packets (i.e., *LispOutput()*) as well as decapsulating LISP packets(*LispInput()*). It contains a smart pointer[1] pointing to the LISP database and LISP cache (e.g. *MapTable*) on which executes mapping search.

We take the double encapsulation example shown in Fig. 2.8 to illustrate how LISP encapsulation and decapsulation is implemented. We assume that the required mapping entries during LISP Data Plane operations are already in LISP-MN cache. A LISP-speaking node behind the xTR_1 needs to communicate with CN behind the xTR_3. Thus, a packet should be encapsulated within the considered node and forwarded to the xTR_1. Subsequently, the packet is further encapsulated and forwarded. At the xTR_3, the received packet is decapsulated twice and forwarded. Thus, the example involves packet transmission, forwarding and reception.

[1] A smart pointer is an abstract data type introduced in C++ that simulates a pointer while providing additional features, such as automatic memory management or bounds checking.

Within the LISP-MN node, when the upper layer of IP protocol calls the *Send()* method of *Ipv4L3Protcol*, a packet comes down to IP layer. The *Send()* method is adapted so that it first verifies whether the *LispOverIp* object is present. If yes, some checks are then conducted to determine that this packet should be processed by *LispOutput()* (to encapsulate the packets) or by conventional packet transmission routine. For example, if both source and destination IP address of this packet belong to the same network, the LISP-related process (e.g., encapsulation) is skipped and this packet is processed as in a non-LISP network. Otherwise, EID-to-RLOC mapping information is searched from LISP cache and LISP database on LISP-MN node.

After encapsulation and forwarding, the considered packet is forwarded to xTR_1. Low layer invokes the *Receive()* method of *Ipv4L3Protocol* to pass this packet to IP layer. Since this packet is destinated to a remote CN instead of itself, this packet is processed by the patched *IpForward()* method. With this method, LISP encapsulation is verified. *LispOverIp* looks for the source RLOC (RLOC of xTR_1) and destination RLOC (RLOC of xTR_3) for the outer IP header building. Once this step is done, *Ipv4L3Protocol*'s *Send()* method transmits these encapsulated packet to MAC layer for transmission.

When this packet arrives at xTR_3 which servers CN, xTR_3 finds that the destination of this packet is the node itself, the packet is processed by *LocalDelivery()* method in *Ipv4L3Protocol*. Before passing to transport layer, *LocalDelivery* checks if the packet should be decapsulated. If yes, it is passed to *LispInput()* method, in which the packet is decapsulated and re-injected into the IP stack. That is to say, *Receive()* method is called again after decapsulation operation. If the packet still has LISP header, the aforementioned procedure will repeated until it has no need to be decapsulated. Finally, the packet with source address of LISP-MN and destination address of CN is forwarded to CN.

7.3.2 Implementation of LISP Control Plane

The implementation of LISP Control Plane at least should provide ITR/ETR, MR and MS. In practice, ITR and ETR functionalities are usually placed on a same router called xTR. In our implementation, they are included into class *LispEtrItrApplication*. The functionalities of MR and MS are respectively implemented by class *MapResolver* and *MapServer*. The LISP Control Plane messages (Map-Register, Map-Request, etc.) are represented by the derived classes of *LispControlMsg*. In addition, to communicate with LISP Data Plane, a socket class *LispMappingSocket* is proposed.

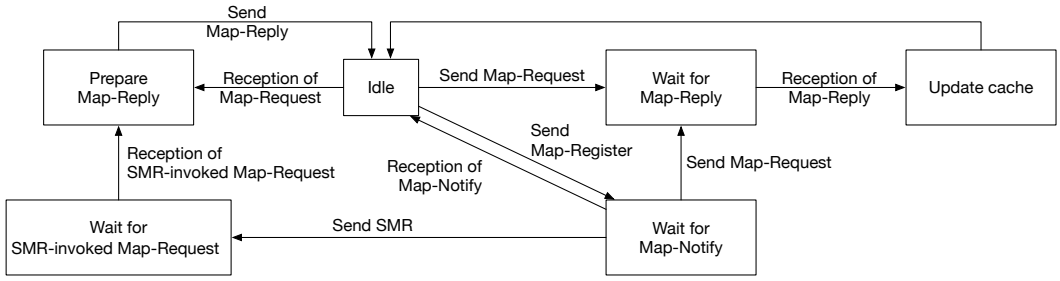


Fig. 7.2 State transition diagram of xTR application

Implementation of xTR functionalities

A ns-3 node that runs *LispEtrItrApplication* is a LISP-compatible router. It should be able to communicate with *LispOverIp* on the same node (e.g. inform cache missing event) and other LISP-compatible routers (e.g. Map-Request/Map-Reply). The state transition diagram is illustrated in Fig. 7.2. When destination RLOC is not found in the cache of xTR (here indicates the xTR functionalities on LISP-MN) for a processed packet, the cache miss event occurs and LISP Data Plane (e.g. *LispOverIp*) notifies *LispEtrItrApplication* on the same LISP-MN node of this event via a *LispMappingSocket* socket. Once reception of cache missing event from LISP Data Plane (i.e. *LispOverIp* object), *LispEtrItrApplication* initiates a Map-Request message to LISP mapping system. Once reception of Map-Reply, the received EID-to-RLOC mapping is inserted into LISP cache. It should be noted that in our implementation, before the reception of Map-Reply message, all transmitted packets with the required RLOC as desination are dropped. There exists one reception: the processing of SMR-invoked Map-Request message. If the RLOC of xTR initiating the SMR (actually a local RLOC) is not found, SMR-invoked Map-Request message is buffered and sent again once the insertion of required mapping information into LISP cache. This is achievable thanks to a callback associated with LISP cache insert operation.

In case of reception of Map-Request, *LispEtrItrApplication* executes a database look up on *MapTable* and generates the corresponding Map-Reply containing EID-to-RLOC mapping.

When xTR application starts or LISP database on a node has information update (e.g. during a handover scenario), xTR application sends a Map-Register message and waits for a Map-Notify message. In case of LISP database update, xTR sends a SMR message to all xTR whose RLOC is present in its cache. According to RFC 6830 [47], the xTR receiving the SMR has two possibilities of reactions about sending SMR-invoked Map-Request: towards a mapping system or the xTR initiating this SMR. Both cases are implemented in

our work and this can be configured by an attribute of *LispEtrItrApplication*. It should be noted that in double-encapsulation, if SMR-invoked Map-Request is directly sent to LISP-MN node whose LRLOC is unknown for xTR, the first SMR-invoked Map-Request cannot be sent due to cache miss. We implement a callback function within *BasicMapTable* which allows to send immediately the buffered SMR-invoked Map-Request upon insertion of required EID-to-RLOC mapping into cache. Note that this mechanism is designed only for LISP Control Plane messages. LISP Data Plane is dropped in case of cache miss.

To support LISP-MN feature, *LispEtrItrApplication* also communicates with DHCP client application. For example, once a LISP-MN obtains an IP address from DHCP server, *LispEtrItrApplication* receives the corresponding EID-to-RLOC mapping and sends a Map-Register message to the Map Server.

Implementation of MR and MS

A node that runs a *MapServer* application is the MS in a LISP-supported network. This class has a smart pointer pointing to a LISP database (i.e. *MapTables*) to store the all EID-to-RLOC mapping information. This application is always listening to UDP port 4342. Once reception of Map-Register message, it retrieves the EID-to-RLOC mapping inside, inserts the latter into LISP database and sends a Map-Notify message as response if necessary. Each time MS receives a Map-Request message, it looks up the required EID within its database. If the EID-to-RLOC mapping is found, map server forwards this request to the corresponding xTR otherwise sends a Negative Map-Reply message to the querying xTR. It is worth to indicating that class *MapServer* is actually an abstract base class. The real functionalities are implemented by its subclasses, which is *MapServerDdt* in our implementation. This design allows the researchers to easily integrate and test their own implementations of MS. In current implementation, the role of MR is to receive the Map-Request message from xTR and forward it to the MS.

7.3.3 Modification of DHCP client to support LISP mobility

To support LISP mobility for IPv4, the intervention of Dynamic Host Configuration Protocol (DHCP) is indispensable. From ns-3.27, DHCP client and server have been implemented in module *internet-apps*. However, the DHCP client of ns-3 is not compatible with LISP. Thus, conventional DHCP client is adapted to support LISP. The state transition diagram of DHCP client is illustrated in Fig. 7.3. When a LISP-MN node roams into the area covered by another xTR, the boot procedure of DHCP client is triggered once that the link state is changed. Afterwards, the DHCP client sequentially pass 'INIT', 'Selecting', 'Requesting' states and

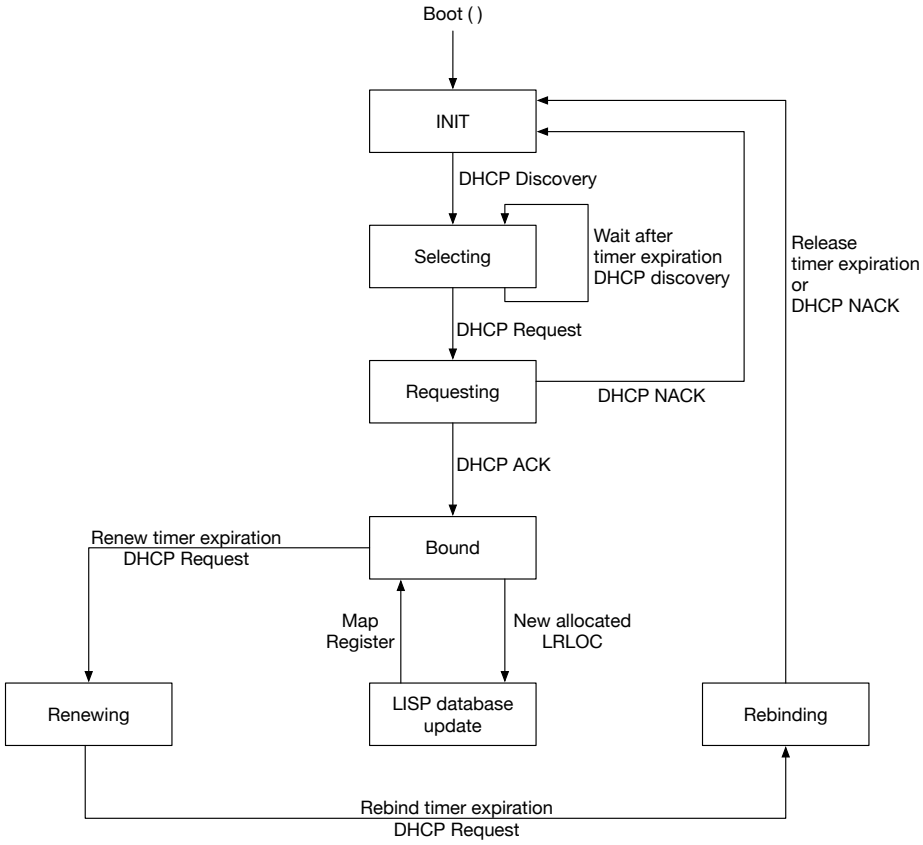


Fig. 7.3 LISP-compatible DHCP client state transition diagram

enter into 'Bound' state after the reception of DHCP ACK message from DHCP server. In 'Bound' state, apart from saving the newly obtained IPv4 address (namely LRLOC) and default gateway, DHCP client should be able to check if the LRLOC is different from the one associated with its EID in its LISP database. If LRLOC is changed, DHCP client enters into state of 'LISP database update'. DHCP client is equipped with a dedicated socket of type *LispMappingSocket*. By this socket, DHCP client notifies the *LispEtrItrApplication* by sending a dedicated message that contains the EID-LRLOC mapping. *LispEtrItrApplication* is in charge of populating the received mapping entry into LISP database and sending a Map-Register message to the Map Server.

To be compatible with DHCP, conventional LISP-related process is also modified. For example, to transmit a DHCP Discovery message (application layer message), its source IP address is set as 0.0.0.0. The *Send()* method of *Ipv4L3Protocol* should be modified so that this message is not processed *LispOverIp*.

7.3.4 Integration of TUN net interface card

To support mobility, a LISP-speaking node actually can be regarded as a small LISP-Site. The xTR functionalities and the DHCP service should be implemented on the LISP-speaking node. The address of MR and MS should be configured. As a LISP-MN, it has a static permanent EID and dynamic RLOC assigned by the DHCP server. To differentiate with conventional RLOC of xTR interface, such kind of RLOC is referred to as the local RLOC (LRLOC). As a design choice, we use the solution based on TUN device, which is also applied by Lispmob [17]. In our implementation, different from a conventional LISP node, at least two NICs should be installed into the node. One is normal NIC such as *WifiNetDevice*. The DHCP client application runs on this kind of card and thus the LRLOC is allocated to this card. The other is a TUN type card. The TUN NIC is a virtual card which should actually invoke *Send()* of another real NIC. The permanent EID is assigned to TUN device.

Recall that after the DHCP procedure, the node will be configured a default gateway provided by the DHCP server. Routing table of LISP-MN are modified so that the packets from application layer always use EID as the source IP address of inner IP header.

7.4 Theoretical analysis

IP mobility leveraging on LISP can be implemented either on the border router or the end host. To explore the characteristics of each scheme, we propose the following three different scenarios and analyze the overall handover delay and LISP Control Plane overhead.

1. LISP-MN in the non-LISP-Site (i.e., only the end host supports LISP).
2. MN in the LISP-Site (i.e., only the border router supports LISP).
3. LISP-MN in the LISP-Site (i.e., both the border router and the end host support LISP).

The network topology for all the scenarios is shown in Fig. 7.4. An MN is initially in the subnet of $Router_1$. It exchanges packets with a remote stationary node CN situated in the LISP-Site of xTR_3. The roles of $Router_1$ and $Router_2$ are slightly different in different scenarios. This will be respectively specified in the following parts (Sec. 7.4.1 to Sec. 7.4.3). To obtain the estimation of LISP mobility handover, we do not consider the delay due to wireless link switch and use an intermediate router to connect all three routers and the mapping system. The connection between MN and $Router_1$ can be either Wi-Fi or wired link. If Wi-Fi link is used, MN will move into the subnet of $Router_2$ at a constant speed. At a certain moment, the Wi-Fi link between MN and $Router_1$ is down, which triggers the handover procedure. Afterwards, MN connects to $Router_2$ and re-establishes the communication with CN node. If they use wired link, the handover procedure can be simulated as follows. The

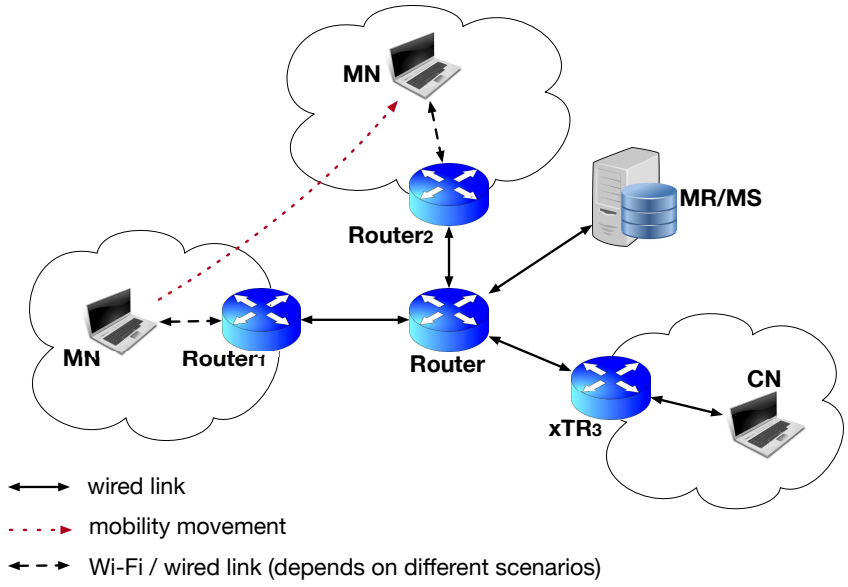


Fig. 7.4 General scenario for LISP mobility architecture

MN have two interfaces respectively connected to $Router_1$ and $Router_2$. At a certain time, the connection to $Router_1$ set to be down and the one with $Router_2$ is set to be up. At the same time, the DHCP client on the interface connected to $Router_2$ runs and this triggers the handover procedure. By using wired link between MN and its routers is just to test the different mechanisms in an ideal situation, and it is not a real scenario.

In this chapter, the overall handover delay related to LISP $D_{overall}$ is defined as the time interval between the first and the last LISP packets during the handover procedure. The overall handover overhead $C_{overall}$ is defined as the number of LISP Control Plane messages exchanged during handover procedure. According to the three following scenarios, the handover delay and overhead consist of different parts. All the necessary delay and LISP overhead of LISP Control Plane during the mobility are listed in Tab. 7.1.

7.4.1 LISP-MN in non-LISP-Site

The first scenario is the LISP-MN in non-LISP-Site, where the border routers are the conventional routers and LISP is only implemented on the mobile end host MN. In our simulation, the LISP-MN with permanent EID is initially placed in the subnet of $Router_1$, with the IP address distributed by $Router_1$ as its RLOC. The remote CN is a conventional stationary end host, residing in a LISP-Site of xTR_3. LISP-MN first has a DHCP procedure when MN moves to $Router_2$, so that the later allocates a new IP address as RLOC. Then LISP-MN

Table 7.1 Symbols for numerical analysis

Symbols	Explanations
$D_{overall}$	Overall handover delay related to LISP
D_{DHCP}	DHCP address configuration delay
$D_{Register}$	Delay of sending Map-Register
D_{Notify}	Delay of receiving Map-Notify
$D_{Request}$	Delay of sending Map-Request to MDS
D_{Reply}	Delay of receiving Map-Reply
$D_{Resolve}$	Delay of resolving mapping information in MDS
D_{SMR}	Delay of sending SMR
$D_{Request_{SMR}}$	Delay of sending a SMR-invoked Map-Request
D_{Link}	Link delay between two network entities
T_{A-B}	Delay of packet transmission between A and B
$T_{timeout_{SMR}}$	Timeout of SMR

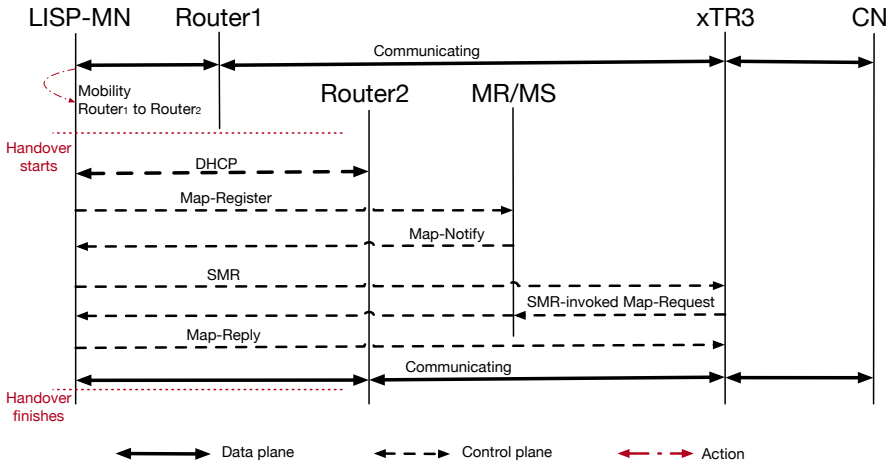


Fig. 7.5 Schema for LISP-MN mobility in non-LISP-Site (SMR-invoked Map-Request is sent to the mapping system)

needs register its new mapping information to the mapping system, and also send a *SMR* to its communicating nodes in its cache (there is only CN in our scenario). xTR_3 sends an SMR-invoked Map-Request to the mapping system, so to obtain the new mapping information of LISP-MN. Afterwards, LISP-MN re-establishes the communication with CN node via $Router_2$. The detailed traffic schema related to the handover procedure is illustrated in Fig. 7.5.

The overall handover delay in this scenario is composed by two parts: the DHCP related delay and LISP related delay. The delays of DHCP procedures consist of LISP-MN sending DHCP Discovery message to $Router_2$, receiving DHCP Offer message, sending DHCP Request message, and receiving DHCP ACK message. The delays of all the LISP procedures

are presented in Fig. 7.5. This chapter only focus on the delay caused by LISP. Thus, the overall handover delay $D_{overall}$ related to LISP of LISP-MN in non-LISP-Site is:

$$D_{overall} = D_{Register} + D_{Notify} + D_{SMR} + D_{Request_{SMR}} + D_{Reply} \tag{7.1}$$

To facilitate the comparison between each scenario, we provide the numerical results for the designed topology as shown in Fig. 7.4. Every link between two network entities in this architecture is set to $1ms$. According to the experimental results of Coras et al.'s paper [44], we set the resolving time in the mapping system as $200ms$. Thus, the $D_{overall}$ of the first scenario in our designed architecture is:

$$\begin{aligned} D_{overall} &= 3T_{MN-MDS} + 2T_{MN-xTR_3} + T_{xTR_3-MDS} + D_{Resolve} \\ &= 3*(3*D_{Link}) + 2*(3*D_{Link}) + 2*D_{Link} + D_{Resolve} \\ &= 17D_{Link} + D_{Resolve} \\ &= 217ms \end{aligned}$$

The handover overhead is 6 messages of LISP Control Plane. It includes 2 signalings of registration when LISP-MN connects to $Router_2$ (Map-Register and Map-Notify), and 4 signalings related to SMR procedure: 1 SMR from LISP-MN to xTR_3, 1 SMR-invoked Map-Request from xTR_3 to the mapping system, 1 Map-Request forwarded by mapping system to LISP-MN, and 1 Map-Reply from LISP-MN to xTR_3.

There are two options when xTR_3 receives *SMR*. It can send the SMR-invoked Map-Request to the mapping system as we described before, or it can directly send the SMR-invoked Map-Request to the source locator address of *SMR* [47]. In this scenario, the source of *SMR* is LISP-MN. Thus, the overall handover delay $D_{overall}$ in this scenario is as follows:

$$D_{overall} = D_{Register} + D_{Notify} + D_{SMR} + D_{Request_{SMR}} + D_{Reply} \tag{7.2}$$

The numerical result is:

$$\begin{aligned} D_{overall} &= 2T_{MN-MDS} + 3T_{MN-xTR_3} \\ &= 2*(3*D_{Link}) + 3*(3*D_{Link}) \\ &= 15D_{Link} \\ &= 15ms \end{aligned}$$

Compared with solution of sending SMR-invoked Map-Request to the mapping system, the overall handover delay of sending the SMR-invoked Map-Request back to the source of *SMR* is smaller, since there is no resolving delay in the mapping system.

In this scenario, the handover overhead associated to LISP Control Plane is 5 messages. Since the SMR-invoked Map-Request is directly sent from xTR_3 to LISP-MN instead of passing the mapping system, it has 1 signaling less than the one sent to the mapping system.

The advantages of this scenario, i.e., LISP-MN in non-LISP-Site, are: 1) it is able to achieve handover through different subnets; 2) the numerical analysis indicates that the overall handover delay is small; 3) so to the overall overhead. Compared to the other two scenarios, which will be described in details in the following sections, the mobility of this scenario does not cause lots of traffic in LISP Control Plane. However, since the routers are still the normal routers in this scenario, it cannot help to reduce the BGP routing table size, which is the initial purpose to motivate LISP. Moreover, each LISP-MN needs a permanent IP address as its EID, which increases the burden of IPv4 address allocation. Each permanent EID and its LRLOC needs register to the mapping system, which also increases the size of LISP mapping table.

7.4.2 MN in LISP-Site

The second scenario is the MN in LISP-Site, where the mobile node MN is conventional and LISP are only implemented on the border routers. In our simulation, the MN is initially placed in the subnet of xTR_1, with the assigned IP address as its EID. The remote CN is same to the first scenario. The MN communicates with CN by encapsulating the packets on xTR_1 and decapsulating the packets on xTR_3, and the MN moves into the coverage of xTR_2 after the simulation begins. Since the communication should not be interrupted during the mobility, this scenario limits the movement of MN being only within the same subnet, i.e., one of the EID-prefixes of xTR_1 is same to one of xTR_2's. Similar to the first scenario, the DHCP procedure is necessary so to trigger the registration of new mapping information, but MN still keeps the former IP address as its EID, instead of xTR_2 distributing a new one to it. Then xTR_2 registers the new mapping information to the mapping system. As the mapping system finds out that the EID of MN has been registered and associated by xTR_1, it sends a Map-Notify to both xTRs. The reason for sending to xTR_2 is the acknowledgment of the reception of its Map-Register. Whereas the purpose to xTR_1 is to tell it that the MN is now mapping with xTR_2 and inform the remote CN to update its mapping information. Since MN used xTR_1 to communicate with CN in the past time, only xTR_1 stored in its cache that to whom MN was exchanging the packets instead of xTR_2. Thus, xTR_1 sends a *SMR* to xTR_3, and xTR_3 sends an SMR-invoked Map-Request to the mapping system, so to obtain the new mapping information of MN. Afterwards, MN re-establishes the communication with CN node via xTR_2. The detailed traffic schema related to the handover procedure is shown in Fig. 7.6.

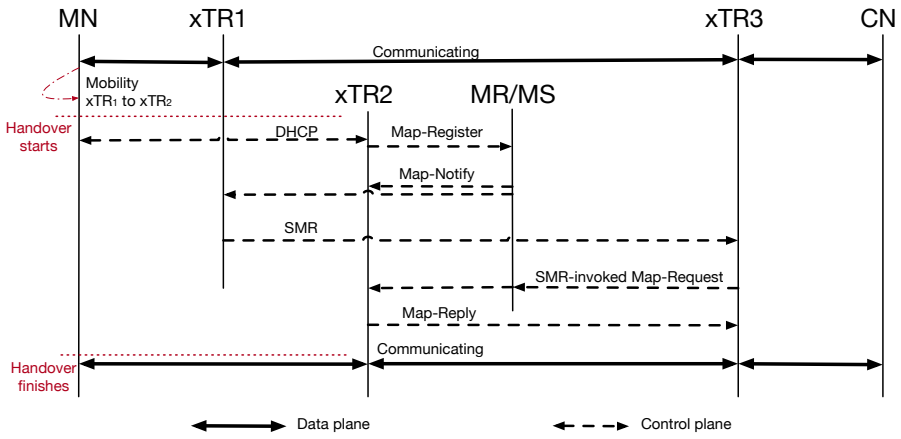


Fig. 7.6 Schema for MN mobility in LISP-Site (SMR-invoked Map-Request is sent to the mapping system)

The overall handover delay $D_{overall}$ in this scenario is almost same to the first one:

$$D_{overall} = D_{Register} + D_{Notify} + D_{SMR} + D_{Request_{SMR}} + D_{Reply} \tag{7.3}$$

The numerical result is as follows:

$$\begin{aligned} D_{overall} &= 2T_{xTR_2-MDS} + T_{MDS-xTR_1} + T_{xTR_1-xTR_3} + T_{xTR_3-MDS} + \\ &\quad D_{Resolve} + T_{xTR_2-xTR_3} \\ &= 2*(2*D_{Link}) + 2*D_{Link} + 2*D_{Link} + 2*D_{Link} + D_{Resolve} + 2*D_{Link} \\ &= 12D_{Link} + D_{Resolve} \\ &= 212ms \end{aligned}$$

The handover overhead in this scenario is 7. Besides 4 messages used for SMR procedure, it needs 3 signalings to complete the registration. There are 1 Map-Register from xTR_2 to the mapping system and 2 messages of Map-Notify: 1 to the xTR_2 and 1 to the xTR_1.

Same to the first scenario having two options, when xTR_3 receives the *SMR* from xTR_1, it can also directly send SMR-invoked Map-Request back to xTR_1, which implies that the xTR_1 puts the new mappings into its database. The overall handover delay $D_{overall}$ is as follows:

$$D_{overall} = D_{Register} + D_{Notify} + D_{SMR} + D_{Request_{SMR}} + D_{Reply} \tag{7.4}$$

The numerical result is:

$$\begin{aligned} D_{overall} &= T_{xTR_2-MDS} + T_{MDS-xTR_1} + T_{xTR_1-xTR_3} + 2T_{xTR_2-xTR_3} \\ &= 2*D_{Link} + 2*D_{Link} + 2*D_{Link} + 2*(2*D_{Link}) \\ &= 10D_{Link} \\ &= 10ms \end{aligned}$$

Since the SMR-invoked Map-Request is not sent to the mapping system, there is no resolving delay. However, in this scenario, as xTR_1 is no longer in charge of MN, how long it stores the CNs for MN in its cache is an important point to discuss. If the expired time is set too long, it wastes the source of xTR_1 and is not necessary. Thus, it still remains as an open issue to seek for an optimal value of timeout that offers an appropriate tradeoff between resources saving and effectively providing the new mapping information.

Similarly to the first scenario, directly responding to the source locator of SMR has 1 signaling message less than the one requesting to the mapping system first. Thus, the overall handover overhead in this case is 6.

Since the routers support LISP, i.e., are the xTRs in this scenario, it can help to reduce the BGP routing table size, which is the initial motivation of proposing LISP. Besides, the analysis hints that the overall handover delay of this scenario is the shortest. However, it can only provide the mobility within a same subnet, which means that it is not able to offer the handover through different subnets. Thus, this scenario is more suitable to be applied for the mobility of virtual machines in the Data Center.

7.4.3 LISP-MN in LISP-Site

The third scenario is the LISP-MN in LISP-Site, where both the border routers and the mobile node MN are implemented LISP. In our simulation, the LISP-MN with permanent EID is initially placed in the subnet of xTR_1, with the IP address allocated by xTR_1 as its LRLOC. The remote CN is still same to the first two scenarios, which is a conventional stationary end host, residing in a LISP-Site of xTR_3. The LISP-MN communicates with CN by double encapsulation. The first encapsulation is on itself and the second time is on the xTR_1. When the LISP packets arrive at xTR_3, it needs decapsulate them twice. LISP-MN first has a DHCP procedure with xTR_2, so that the later assigns it a new IP address as its LRLOC. Then LISP-MN needs register its new mapping information to the mapping system, and also send a *SMR* to all the xTRs of CNs in its cache (actually is to the xTR of CN, i.e., xTR_3 in our scenario). Once xTR_3 receives a *SMR*, it sends an SMR-invoked Map-Request to the mapping system, so to obtain the new mapping information of LISP-MN. Actually,

this mapping information that xTR_3 obtains is the $< EID_{MN}, LRLOC >$ for LISP-MN. It is not able to send the packets to LISP-MN at the moment, since it does not know how to route the packets to the *LRLOC* of LISP-MN, i.e., it lacks the mapping information for *LRLOC*. Only until xTR_3 receives the packets from CN to LISP-MN, which triggers the Map-Request procedure to the mapping system, can xTR_3 know the mapping information of LRLOC $< LRLOC, RLOC_{xTR_2} >$. Now xTR_3 gets the double mapping information for LISP-MN. Afterwards, LISP-MN re-establishes the communication with CN node by passing xTR_2. The detailed traffic schema related to the handover procedure is illustrated in Fig. 7.7.

Since this scenario is double encapsulation that xTR_3 needs to know both inner and outer mapping information of LISP-MN for sending the packets. The overall handover delay $D_{overall}$ in this scenario is larger than the first two scenarios. The $D_{overall}$ is:

$$\begin{aligned} D_{overall} &= D_{Register} + D_{Notify} + D_{SMR} + D_{Request_{SMR}} + D_{Reply} + \\ &\quad D_{Request} + D_{Reply} \end{aligned} \tag{7.5}$$

The numerical result is as follows:

$$\begin{aligned} D_{overall} &= 3T_{MN-MDS} + 2T_{MN-xTR_3} + 2T_{xTR_3-MDS} + 2D_{Resolve} + \\ &\quad T_{MDS-xTR_2} + T_{xTR_2-xTR_3} \\ &= 3*(3*D_{Link}) + 2*(3*D_{Link}) + 2*(2*D_{Link}) + 2D_{Resolve} + \\ &\quad 2*D_{Link} + 2*D_{Link} \\ &= 23D_{Link} + 2D_{Resolve} \\ &= 423ms \end{aligned}$$

The double encapsulation causes not only longer handover delay but also more handover overhead related to LISP Control Plane. Two messages are needed for the registration, 3 signalings are used to obtain the *LRLOC* of *EID*. Besides, two more Map-Requests (one from xTR_3 to the mapping system, one from the mapping system to xTR_2) and one more Map-Reply are required to get the *RLOC* of *LRLOC*. Thus, the handover overhead is 9, which has 3 messages more than the first scenario.

Differently from the first two scenarios, when xTR_3 directly sends the SMR-invoked Map-Request back to the source of *SMR* has smaller overall handover delay, this solution for the third scenario has bigger delay instead. It is caused by the double encapsulation in this scenario while the first two scenarios have only single encapsulation. When xTR_3 receives the *SMR* from LISP-MN, it wants to send SMR-Invoked Map-Request to the LISP-

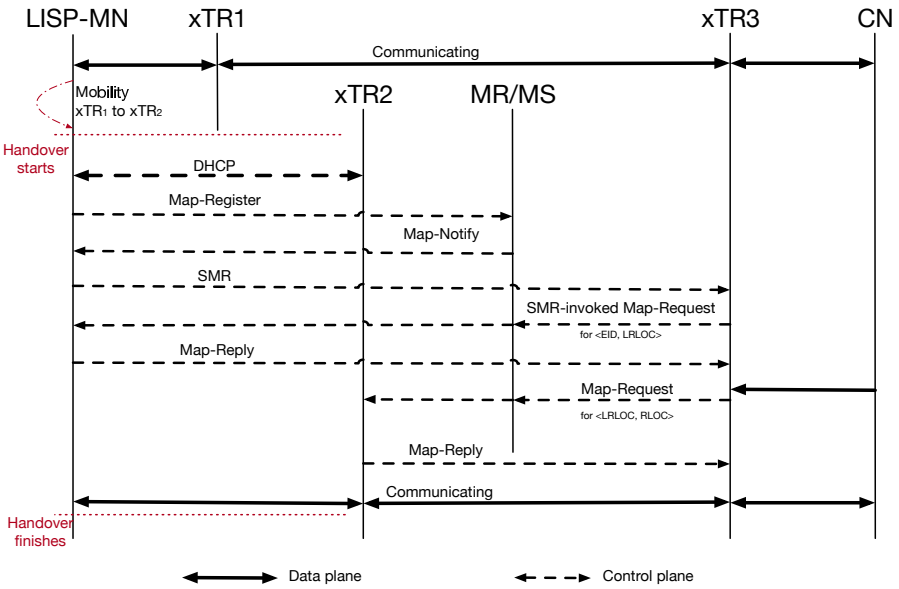


Fig. 7.7 Schema for LISP-MN mobility in LISP-Site (SMR-invoked Map-Request is sent to the mapping system)

MN for the mapping information of $< EID_{MN}, LRLOC >$, but it does not know how to reach to LISP-MN, i.e., it lacks the mapping information of $< LRLOC, RLOC_{xTR_2} >$. Thus, it discards the *SMR* and sends the Map-Request to the mapping system first. Then, the mapping information of $< LRLOC, RLOC_{xTR_2} >$ is stored in its cache. It waits for the next *SMR* so to send the SMR-invoked Map-Request to the LISP-MN. The traffic schema is shown in Fig. 7.8.

The overall handover delay $D_{overall}$ by sending the SMR-invoked Map-Request to the source of *SMR* in this scenario is as follows, where $T_{timeout_{SMR}}$ is the interval to re-send the *SMR* in case of nothing received. We set it to $1s$ in the simulation:

$$\begin{aligned} D_{overall} &= D_{Register} + D_{Notify} + D_{SMR} + T_{timeout_{SMR}} + D_{SMR} + \\ &\quad D_{Request_{SMR}} + D_{Reply} \end{aligned} \tag{7.6}$$

The numerical result is as follows:

$$\begin{aligned} D_{overall} &= 2T_{MN-MDS} + 4T_{MN-xTR_3} + T_{timeout_{SMR}} \\ &= 2 * (3 * D_{Link}) + 4 * (3 * D_{Link}) + T_{timeout_{SMR}} \\ &= 18D_{Link} + T_{timeout_{SMR}} \\ &= 1018ms \end{aligned}$$

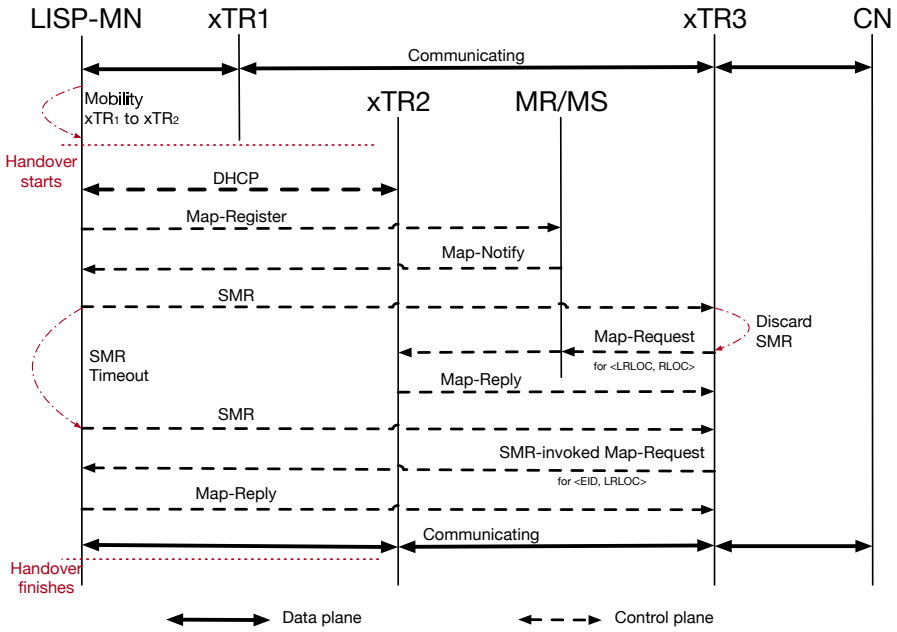


Fig. 7.8 Schema for LISP-MN mobility in LISP-Site (SMR-invoked Map-Request is sent to the source of SMR)

Although when the SMR-invoked Map-Request directly sent to the source of SMR has the different traffic schema from the one sent to the mapping system, the overall handover overhead is still 9. The difference between them is only the order of getting mapping information on xTR_3, where this case is to obtain the outer mapping first but the last case is to get the inner mapping first.

Since both the MN and the border routers support LISP, the advantages of this scenario, i.e., LISP-MN in the LISP-Site, are: 1) it can help to reduce the BGP routing table size; 2) it is able to achieve handover through different subnets. However, same to the shortcomings for the first scenario, each LISP-MN needs a permanent IP address as its EID, which increases the burden of IPv4 address allocation. Each permanent EID and its LRLOC needs register to the mapping system, which also increases the size of LISP mapping table. Besides, the numerical analysis indicates that the overall handover delay is much longer than the other two scenarios due to its double encapsulation.

Depend on our designed topology, all the aforementioned analyzed results are presented in Fig. 7.9. The left hand figure is the numerical overall handover delay related to LISP and right hand figure is about the handover overhead. To be noted, time values are specific to the topology used as an example, while the number of messages do not depend on the specific topology but only on the scenario.

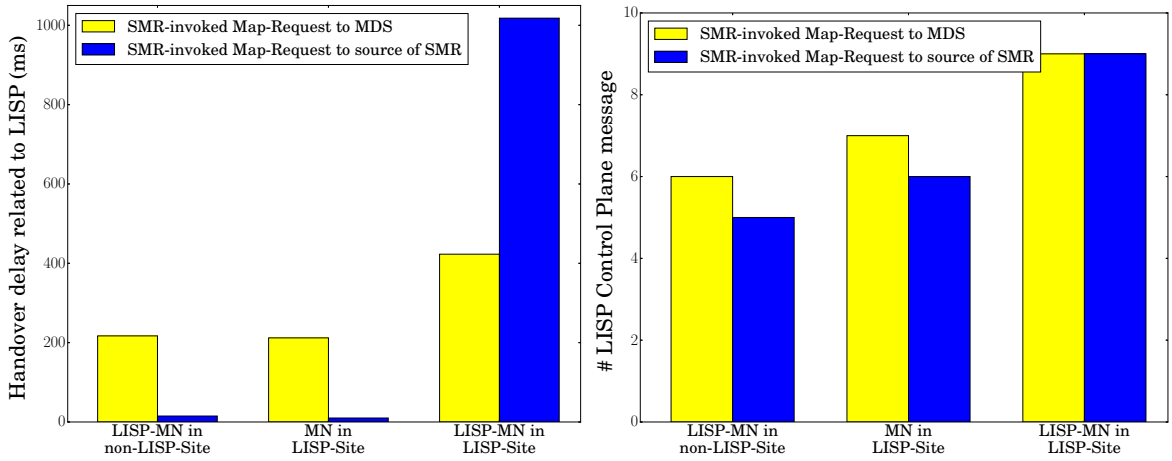


Fig. 7.9 The handover delay related to LISP (left) and the handover overhead (right) grouped by three LISP mobility scenarios.

7.5 Summary

There exist some LISP simulation implementations, but they are proprietary or they do not support the extension of LISP mobility. Further, although measurements on LISP-testbeds can provide real time performance, due to the complicated topological structure, it is somewhat like a black box test, which hinders us to find the exact explanation for some results. This highlights the importance to have an open source simulator for LISP in particular to support LISP mobility functionality. In this chapter, based on an implementation of basic LISP on ns-3.24, we adapt it to ns-3.27 first (the latest version at the moment of writing). To facilitate the researchers to deeply track the exchange of LISP packets, we encode in LISP Data Plane packets so that the Wireshark can resolve them. Finally we implement the LISP mobility extensions on it. As there are three methods to support mobility in LISP: host-based (i.e. LISP-MN), network-based (i.e., xTR), both host-based and network-based (i.e., LISP-MN behinds xTR) mobility. We analyze the overall handover delay and the overhead of LISP Control Plane among them, compare the performance among them by listing their advantages and shortcomings. The source code of both the implementation and the simulation evaluation of all mobility scenarios is available on Github[2].

[2]Source code available on Github: https://github.com/SeleneLI/LISP_ns-3

Chapter 8

Conclusion and future work

In this chapter, we summarize our major contributions and discuss future research directions.

8.1 Major contributions

The objective of this dissertation is to thoroughly evaluate *Locator/Identifier Separation Protocol (LISP)* from different aspects: from measuring its performance by conducting experiments in the large-scale real networks, numerically analyzing the bounds, up to implementing LISP mobility in ns-3. The measurement works are related to the evaluation of mapping system and the assessment of interworking performance with the legacy Internet. Due to the new findings in the mapping system, we propose a more comprehensive LISP monitor tool to address the additional requirements. The theoretical analysis is about the LISP mobility. More precisely, we analyze the low bounds of handover for the different network components supporting LISP. Besides, we implement LISP mobility extensions on ns-3.27 by leveraging an existing implementation. Concretely, the main contributions can be summarized as follows.

8.1.1 Evaluation of LISP mapping system

We continuously measured the LISP Beta Network for seventeen days to assess the stability and consistency of the mapping system. The experiments are conducted from July 2nd to 18th 2013, from 5 Vantage Points (VPs), sending *Map-Request* messages to all the 13 Map-Resolvers (MRs) of the LISP Beta Network, for all selected 613 destination IP addresses every 30 minutes. The VPs were composed by part of academic networks, commercial Internet, PlanetLab and spread across Europe and USA. The destination IP addresses consisted of the first address within each prefix derived from the LISPmon project [18] re-

gardless of the type (i.e., the EID-prefixes or regular prefixes). The results of measurements show that the mapping system is stable during 91.49% of time, is consistent respectively by MR in 86.3% and by VP in 90.48% of the cases. As the instabilities and inconsistencies are observed and for studying them in details we developed a new taxonomy. It shows that Map-Reply only changes once in the most instability cases, and the inconsistency is mainly caused by the different types of received Map-Replies (the mix of Negative Map-Reply and LISP Map-Reply). All in all, instabilities and inconsistencies are rare events. At last, the utilization of multi-homing and dual-stack (i.e., IPv4 and IPv6) are observed during the whole experiment. This work is published in WNM [82].

8.1.2 LISP-Views: LISP mapping system monitor

The evaluation of the mapping system shows that there exist instability and inconsistency, which are hard to be found by the only current LISP monitor LISPmon. A dynamic and complete monitoring system is required to accompany the growth of LISP. We propose LISP-Views, a versatile large scale LISP monitoring architecture. LISP-Views allows to automatically conduct comprehensive and objective measurements. Users can deploy it on whichever VP and define to monitor which MRs. They are also able to define the type of measurements. All these operations can be conducted via the provided API. At the moment of writing, LISP-Views is deployed on one VP to supervise all the existing MRs on both LISP-Beta Network and LISP-Lab Platform. After running LISP-Views in the wild for several months and comparing the monitoring results with LISPmon, we confirm that LISP-Views provides more detailed and accurate information. We observe that every MR has different behaviours compared to others, such as in the reliability and resolving time. Besides, the measurements helped the OpenLISP coders fix a previously unknown bug in the MR of LISP-Lab platform. The preliminary graphic results, such as the number of LISP Map-Reply, the number of RLOCs contained in each LISP Map-Reply, and the responsiveness of each MR are shown on the website of LISP-Views [15]. This work is published in ITC [77].

8.1.3 Assessing LISP interworking performance through RIPE Atlas

We conducted two experiments to evaluate LISP interworking performance with legacy Internet. It is the first thorough insight on the performance of LISP PITR/PETR (PxTR). As a setup phase, we selected 4 probes ping to the top 50 Alexa sites every 10 minutes during 6 hours on November 6^{th} 2015. In the second experiment, we chose 5 probes ping and traceroute to top 500 Alexa IPv4 and 122 IPv6 addresses respectively every 30 minutes and

60 minutes. Both experiments show that the PxTR indeed introduces negative effects for nearby destinations, but the negative impact of PxTR can be ignored for the intercontinental long-distance transmission. The results also show that the position of PxTR is very important. The PxTR positioned either near to the sources or the destinations can decrease the latency a lot. LISP is generally stable, except for the IPv6 performance of LISP Beta Network. Further, the performance of LISP-Lab PxTR is more reliable than the one of LISP Beta Network, although at the time of this writing the latter has 6 worldwide PxTRs used for IPv4 and 2 located in US for IPv6, whereas LISP-Lab has only 1 PxTR for both IPv4 and IPv6. Natively forwarding without using LISP decreases the latency, but not much. The traceroute experiment of LISP-Lab shows that introducing PxTR of course brings more hops, but if the PxTR is well configured, so to always have peers to the destinations, there are only 4 more hops compared to the packets being natively forward by xTR without encapsulating with LISP. This work is published in INFOCOM Student workshop [80], ACM SIGCOMM Student workshop [79], and Elsevier Computer Networks [81].

8.1.4 Analysis of LISP mobility and ns-3 Implementation

There are three methods to support LISP mobility: host-based (i.e. LISP Mobile Node (LISP-MN)), network-based (i.e., ITR/ETR (xTR)), both host-based and network-based (i.e., LISP-MN behinds xTR) mobility. In the first scenario, there is only MN supporting LISP (LISP-MN in the non-LISP-Site); in the second scenario, there is only border router supporting LISP (MN in the LISP-Site); and in the third scenario, both MN and border routers supporting LISP (LISP-MN in the LISP-Site). We analyze the overall handover delay and the overhead of the LISP control plane among them. The second scenario has the smallest handover delay, but the MN can only roam within the same subnet. The first and third scenarios support the MN seamless roaming through the different subnets, but the first scenario does not help in reducing the BGP routing table, and the third scenario introduces a much longer handover delay. Based on an implementation of basic LISP on ns-3.24, we adapt it to ns-3.27 first (the latest version at the moment of writing). To facilitate the researchers to deeply track the exchange of LISP packets, we encode every bit in LISP Data Plane packets so that the Wireshark can resolve them. Finally, we implement the LISP mobility extensions on it.

8.2 Discussion & future work

8.2.1 Evaluation of LISP mapping system

There are two aspects to be done in the future. One is to conduct a new experiment to see if something in the mapping system has changed. The other one, which has never been done is using LISP-DDT Referral Internet Groper (RIG) to have a deeper look at the LISP Delegated Database Tree (LISP-DDT).

The experiment was conducted in 2013, while there is a growth of LISP in the following years. The architecture of the mapping system has also changed, in which there were 13 MRs during the measurement but only 7 MRs in 2017. Besides, the LISP Beta Network is the only LISP testbed at that time. However, the LISP-Lab platform is open to the experimenters in 2015. The later consists of 3 MRs and interconnects with the mapping system of LISP Beta Network. Thus, it is very necessary to conduct a new experiment in the future to evaluate the latest *Stability* and *Consistency* performance of LISP mapping system. The results can be compared to the results in Chapter. 4 so to get the LISP developing trend.

The mapping information should be identical on all the MRs at the same time, but the experimental results in Chapter. 4 show that there exist the inconsistencies between the MRs and VPs. The reason is difficult to explore by the experiments conducted before, which are based on LIG. Since the procedures of forwarding Map-Requests within the mapping system is transparent to the experimenters. Based on the theory, it is probably caused by the cache of each MR storing the mapping information is not synchronized. Thus, leveraging on RIG to troubleshoot and first look at the inconsistent issues within the mapping system is very useful. If the assumption is correct, proposing a synchronized mechanism should be important to the mapping system.

8.2.2 LISP-Views: LISP mapping system monitor

LISP-Views is still in the first phase, it is deployed on only one VP and executed the measurements by using the command lines. Thus, it is expected to deploy the LISP-Views on multiple VPs to reproduce the consistency of the VPs as described in Chapter. 4. Besides, the command lines are not efficient to conduct the experiments, to implement a REST API on LISP-Views for setting the experiment time, the monitored MRs, and obtain the different aspects of LISP status should be done in the future. At the moment of writing this dissertation, all the monitored IP addresses via LISP-Views are IPv4, it should be interesting to test IPv6 behaviors.

8.2.3 Assessing LISP interworking performance through RIPE Atlas

In our experiment, as LIP6 probe does not receive any *ping* responses for 42 IPv4 and 10 IPv6 destinations, to cluster these destinations by geography and explore the reasons can be possible future works. As the IPv6 packets sent from the LISP-Lab probe are natively forwarded instead of encapsulating into LISP packets during our experiment, the configurations should be modified so that PETR of LISP-Lab allows to receive the IPv6 traffic from LISP-Lab probe. Then the LISP IPv6 performance can be evaluated. Furthermore, as the degradation of performance observed in Chapter 6 is mainly due to the PITR selection, it will be helpful to explore the interaction between BGP *anycast* and LISP interworking. If the destinations leverage on the *anycast* to select the nearest PITR, we suppose that the additional delay introduced by LISP will be decreased a lot.

8.2.4 Analysis of LISP mobility and ns-3 Implementation

For the next step, we plan to validate the numerical analysis on the simulator. As *Map-Versioning* [67] is another Mapping Cache update mechanism, we can also compare the performance between it and *SMR* that we present in this Chapter. 7 by our implemented simulator.

There are some directions for the further work. The second scenario has the shortest handover delay, but it does not support the MN roaming through the different subnets. Thus, to use MAC address as the EID can be an option so that the MN is able to change the subnet without the interruption of the communication. Further, since the third scenario uses double encapsulation, which causes a big increase in the delay, it will be interesting if the mapping system can differently process SMR-invoked Map-Request and the normal Map-Request. When mapping system receives an SMR-invoked Map-Request, if it finds that the *RLOC* of the queried EID is actually a *LRLOC*, e.g., the *LRLOC* of LISP-MN in our third scenario, it then forwards the Map-Request not only to the LISP-MN, but also to the xTR_2, so that LISP-MN replies to the xTR_3 with the mapping information of $<EID_{LISP-MN}, LRLOC_{LISP-MN}>$, and xTR_2 replies to the xTR_3 with the mapping information of $<LRLOC_{LISP-MN}, RLOC_{xTR_2}>$. In this way, xTR_3 only needs query to the mapping system once, so the handover delay due to the LISP related procedure can reduce a lot. The future work for the implementation of LISP mobility extensions is intended to support IPv6.

References

[1] Alexa. http://www.alexa.com/.

[2] Alexa top sites ranking. https://www.alexa.com/topsites.

[3] As3943 ipv4 route propagation. http://bgp.he.net/AS3943#_graph4.

[4] As3943 ipv6 route propagation. http://bgp.he.net/AS3943#_graph6.

[5] BGP Routing Table Analysis Reports. See https://bgp.potaroo.net/.

[6] Cisco Locator/ID Separation Protocol (LISP) Q&A. See https://www.cisco.com/c/en/us/products/collateral/ios-nx-os-software/locator-id-separation-protocol-lisp/qa_c67-582925.pdf.

[7] FreeBSD (website). https://www.freebsd.org/.

[8] INET Framework (Homepage). See https://inet.omnetpp.org/.

[9] Internet live stats. Accessed on Jan. 7, 2018. [Online]. Available: http://www.internetlivestats.com/internet-users/.

[10] LIP6-LISP (github). https://github.com/lip6-lisp.

[11] LISP Beta Network. See http://www.lisp4.net.

[12] LISP Beta Network Latest Architecture. See https://www.lisp4.net//images/lisp-ddt.pdf.

[13] LISP for SDN and NFV (OpenDaylight Submmit). See https://events.linuxfoundation.org/sites/events/files/slides/LISP_ODLSummit_2014.pdf.

[14] LISP-Lab. See http://www.lisp-lab.org/.

[15] LISP-Views (website). See http://hunter.lisp-views.org/.

[16] Lispers.net. See https://www.lispers.net/.

[17] LISPmob (website). https://www.openoverlayrouter.org/lispmob/.

[18] LISPmon. See http://www.lispmon.net.

[19] ns-3. See http://https://www.nsnam.org/.

[20] OMNet++ (Homepage). See https://www.omnetpp.org/.

[21] Open Overlay Router (website), the successor of LISPmob. See https://www.openoverlayrouter.org/.

[22] OpenDaylight (website). See https://www.opendaylight.org/.

[23] OpenLISP (website). http://www.openlisp.org/.

[24] Planetlab (website). https://www.planet-lab.org/.

[25] pylisp (Github). See https://github.com/steffann/pylisp.

[26] pylisp (Reference). See https://www.biostat.wisc.edu/~annis/creations/PyLisp/pylisp.html.

[27] RIPE Atlas. https://atlas.ripe.net/.

[28] Smartinsights. Accessed on Jan. 7, 2018. [Online]. Available: http://www.smartinsights.com/mobile-marketing/mobile-marketing-analytics/mobile-marketing-statistics/.

[29] The official website of Wireshark. See https://www.wireshark.org/.

[30] J. Adan. Tunneled inter-domain routing (tidr). *Work in Progress*, 2006.

[31] L. Agbodjan. Towards a lisp simulator. Master's thesis, Université de Liège, 2016. Unpublished master's thesis. The source code of Basic LISP implementation under ns-3.24 is on the bitbucket: https://bitbucket.org/Lionel_Agbodjan/tfe__towards_a_lisp_simulator.

[32] M. Aiash. A novel security protocol for resolving addresses in the location/id split architecture. In *International Conference on Network and System Security*, pages 68–79. Springer, 2013.

[33] M. Aiash, A. Al-Nemrat, and D. Preston. Securing address registration in location/id split protocol using id-based cryptography. In *International Conference on Wired/Wireless Internet Communication*, pages 129–139. Springer, 2013.

[34] R. Atkinson, S. Bhatti, and S. Hailes. Evolving the internet architecture through naming. *IEEE Journal on Selected Areas in Communications*, 28(8):1319–1325, 2010.

[35] H. Balakrishnan, K. Lakshminarayanan, S. Ratnasamy, S. Shenker, I. Stoica, and M. Walfish. A layered naming architecture for the internet. In *ACM SIGCOMM Computer Communication Review*, volume 34, pages 343–352. ACM, 2004.

[36] S. Barkai, D. Farinacci, D. Meyer, F. Maino, V. Ermagan, A. Rodriguez-Natal, and A. Cabellos-Aparicio. Lisp based flowmapping for scaling nfv. Technical report, Internet-Draft draft-barkai-lisp-nfv-10, Internet Engineering Task Force. Work in Progress, 2017.

[37] S. Bhatti and R. Atkinson. Identifier-locator network protocol (ilnp) architectural description. 2012.

[38] L. Breslau, D. Estrin, K. Fall, S. Floyd, J. Heidemann, A. Helmy, P. Huang, S. McCanne, K. Varadhan, Y. Xu, et al. Advances in network simulation. *Computer*, 33(5):59–67, 2000.

[39] S. Brim, N. Chiappa, D. Farinacci, V. Fuller, and D. Lewis. Lisp-cons: A content distribution overlay network service for lisp. draft-meyer-lisp-cons-04, 2008.

[40] A. Cabellos and F. Coras. Internet engineering task force (ietf) d. saucez request for comments: 7834 inria category: Informational l. iannone. 2016.

[41] A. Cabellos, A. R. Natal, L. Jakab, V. Ermagan, P. Natarajan, and F. Maino. Lispmob: mobile networking through lisp. *LISPmob white paper*, 2011.

[42] X. Chang. Network simulations with opnet. In *Simulation Conference Proceedings, 1999 Winter*, volume 1, pages 307–314. IEEE, 1999.

[43] F. Coras, A. Cabellos-Aparicio, and J. Domingo-Pascual. An analytical model for the LISP cache size. In *Proc. IFIP Networking*, May 2012.

[44] F. Coras, D. Saucez, L. Iannone, and B. Donnet. On the performance of the LISP beta network. In *Proc. IFIP Networking*, June 2014.

[45] S. Deering and R. Hinden. Rfc 2460: Internet protocol, 1998.

[46] V. Ermagan, D. Farinacci, C. White, J. Skriver, D. Lewis, and F. Maino. Nat traversal for lisp. 2016.

[47] D. Farinacci and et al. The locator/ID separation protocol (LISP). RFC 6830, Internet Engineering Task Force, January 2013.

[48] D. Farinacci, A. Jain, I. Kouvelas, and D. Lewis. Locator/id separation protocol delegated database tree (lisp-ddt) referral internet groper (rig). Technical report, 2017.

[49] D. Farinacci, D. Lewis, D. Meyer, and C. White. Lisp mobile node, draft-ietf-lisp-mn-01. Technical report, institution=IETF Internet draft, Apr. 2017.[Online]. Available: https://tools.ietf.org/html/draft-ietf-lisp-mn-01, 2017.

[50] D. Farinacci, T. Li, S. Hanks, D. Meyer, and P. Traina. Rfc 2784-generic routing encapsulation (gre). *IETF, March 2000*, 2000.

[51] D. Farinacci, T. Li, S. Hanks, D. Meyer, P. Traina, and R. Glenn. Generic routing encapsulation. *IETF (The Internet Engineering Task Force) Request for Comments RFC*, 2784, 2000.

[52] D. Farinacci and D. Meyer. The locator/ID separation protocol Internet groper (LIG). RFC 6835, Internet Engineering Task Force, 2013.

[53] B. Feng, H. Zhang, H. Zhou, and S. Yu. Locator/identifier split networking: A promising future internet architecture. *IEEE Communications Surveys & Tutorials*, 2017.

[54] A. Ford, C. Raiciu, M. Handley, and O. Bonaventure. Tcp extensions for multipath operation with multiple addresses. Technical report, 2013.

[55] V. Fuller. LISP-DDT: A new mapping database for LISP, June 2012. Nanog 55 Talk.

[56] V. Fuller and et al. Locator/ID separation protocol (LISP) map-server interface. RFC 6833, Internet Engineering Task Force, January 2013.

[57] V. Fuller, D. Farinacci, D. Meyer, and D. Lewis. Locator/ID separation protocol alternative logical topology (LISP+ALT). RFC 6836, Internet Engineering Task Force, January 2013.

[58] V. Fuller, D. Lewis, and A. Ermagan, V. Jain. LISP Delegated Database Tree. Internet Draft, Work in Progress draft-ietf-lisp-ddt-01.txt, Internet Engineering Task Force, March 2013.

[59] V. Fuller, D. Lewis, V. Ermagan, and A. Jain. Lisp delegated database tree. *draft-ietf-lisp-ddt-09, Internet Engineering Task Force*, 2017.

[60] A. Galvani, A. Rodriguez-Natal, A. Cabellos-Aparicio, and F. Risso. Lisp-roam: network-based host mobility with lisp. In *Proceedings of the 9th ACM workshop on Mobility in the evolving internet architecture*, pages 19–24. ACM, 2014.

[61] A. García-Martínez, M. Bagnulo, and I. Van Beijnum. The shim6 architecture for ipv6 multihoming. *IEEE Communications Magazine*, 48(9), 2010.

[62] Y. Han, S. Ryu, Y.-J. Suh, and J. W.-K. Hong. Design and implementation of lisp controller in onos. In *NetSoft Conference and Workshops (NetSoft), 2016 IEEE*, pages 417–422. IEEE, 2016.

[63] S. Hanks, D. Meyer, D. Farinacci, and P. Traina. Generic routing encapsulation (gre). 2000.

[64] T. R. Henderson, M. Lacage, G. F. Riley, C. Dowell, and J. Kopena. Network simulations with the ns-3 simulator. *SIGCOMM demonstration*, 14(14):527, 2008.

[65] M. Hoefling, M. Menth, and M. Hartmann. A survey on mapping systems for locator/identifier split internet routing. *IEEE Communications Surveys and Tutorials*, 15(4):1842–1858, November 2013.

[66] L. Iannone and O. Bonaventure. On the cost of caching Locator/ID mappings. In *Proc. ACM SIGCOMM CoNEXT*, December 2007.

[67] L. Iannone, D. Saucez, and O. Bonaventure. Locator/id separation protocol (lisp) map-versioning. Technical report, 2013.

[68] M. Isah, S. Simpson, Y. Sani, and C. Edwards. Towards zero packet loss with lisp mobile node. In *Computing, Networking and Communications (ICNC), 2017 International Conference on*, pages 265–271. IEEE, 2017.

[69] L. Jakab, A. Cabellos-Aparicio, F. Coras, D. Saucez, and O. Bonaventure. LISP-TREE: a DNS hierarchy to support the LISP mapping system. *IEEE Journal on Selected Areas in Communications*, 28(8):1332 – 1343, October 2010.

[70] S. Jeon, J. Jang, and Y.-H. Kim. Network mobility support in the proxy mobile ipv6 domain. *Network*, 2010.

[71] D. Johnson, C. Perkins, and J. Arkko. RFC 3775. *Mobility support in IPv6*, 2004.

[72] J. Kim, L. Iannone, and A. Feldmann. A deep dive into the LISP cache. In *Proc. IFIP Networking*, May 2011.

[73] J. Kim, L. Iannone, and A. Feldmann. Caching locator/id mappings: An experimental scalability analysis and its implications. *Computer Networks*, 57(4):897–909, 2013.

[74] D. Klein, M. Hartmann, M. Hoefling, and M. Menth. Integration of lisp and lispmn into inet. In *Proceedings of the 5th International ICST Conference on Simulation Tools and Techniques*, pages 299–306. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2012.

[75] E. Lear. Nerd: A not-so-novel endpoint id (eid) to routing locator (rloc) database. 2013.

[76] D. Lewis, D. Meyer, D. Farinacci, and V. Fuller. Interworking between locator/id separation protocol (LISP) and non-LISP sites. RFC 6832, Internet Engineering Task Force, January 2013.

[77] Y. Li, A. Abouseif, L. Iannone, and D. Saucez. Lisp-views: Monitoring lisp at large scale. In *Teletraffic Congress (ITC 29), 2017 29th International*, volume 1, pages 178–186. IEEE, 2017.

[78] Y. Li, A. Abouseif, L. Iannone, and D. Saucez. Lisp-views: Monitoring lisp at large scale. In *Algotel/Cores 2018*, 2018.

[79] Y. Li and L. Iannone. Performance evaluation of locator/identifier separation protocol through ripe atlas. In *Proceedings of the 2016 conference on ACM SIGCOMM 2016 Conference*, pages 561–562. ACM, 2016.

[80] Y. Li and L. Iannone. Using ripe atlas to evaluate the locator/id separation protocol. In *Computer Communications Workshops (INFOCOM WKSHPS), 2016 IEEE Conference on INFOCOM 2016 Conference*, pages 696–697. IEEE, 2016.

[81] Y. Li and L. Iannone. Assessing locator/identifier separation protocol interworking performance through RIPE Atlas. *Computer Networks*, 132:118–128, 2017.

[82] Y. Li, D. Saucez, L. Iannone, and B. Donnet. Stability and consistency of the lisp pull routing architecture. In *Proc. Workshop on Network Measurement (WNM)*, November 2016.

[83] M. Mahalingam, D. Dutt, K. Duda, P. Agarwal, L. Kreeger, T. Sridhar, M. Bursell, and C. Wright. Virtual extensible local area network (vxlan): A framework for overlaying virtualized layer 2 networks over layer 3 networks. Technical report, 2014.

[84] F. Maino, V. Ermagan, A. Cabellos, and D. Saucez. Lisp-security (lisp-sec). *draft-ietf-lisp-sec-14, Internet Engineering Task Force*, 2017.

[85] T. Marill and L. G. Roberts. Toward a cooperative network of time-shared computers. In *Proceedings of the November 7-10, 1966, fall joint computer conference*, pages 425–431. ACM, 1966.

[86] M. Menth, D. Klein, and M. Hartmann. Improvements to lisp mobile node. In *Teletraffic Congress (ITC), 2010 22nd International*, pages 1–8. IEEE, 2010.

[87] D. Meyer, D. Lewis, D. Farinacci, and C. White. Lisp mobile node. *draft-meyer-lisp-mn-16. txt, Internet Engineering Task Force*, 2016.

[88] D. Minoli. Security considerations for mipv6. *Mobile Video with Mobile IPv6*, pages 173–189.

[89] P. Mockapetris. Rfc 1034: Domain names-concepts and facilities, 1987. *URL: ftp://ftp. isi. edu/in-notes/rfc1034. txt*.

[90] M. Möller, R. Bye, K. Bsufka, S. A. Camtepe, and S. Albayrak. From simulation to emulation: an integrated approach for network security evaluation. *Lecture Notes in Informatics*, 192, 2011.

[91] D. Montero, M. Siddiqui, R. Serral-Gracia, X. Masip-Bruin, and M. Yannuzzi. Securing the lisp map registration process. In *Global Communications Conference (GLOBECOM), 2013 IEEE*, pages 2145–2151. IEEE, 2013.

[92] R. Moskowitz, T. Heer, P. Jokela, and T. Henderson. Host identity protocol version 2 (hipv2). Technical report, 2015.

[93] R. Moskowitz and M. Komu. Host identity protocol architecture. *draft-ietf-hip-rfc4423-bis-18*, 2017.

[94] A. R. Natal, L. Jakab, M. Portolés, V. Ermagan, P. Natarajan, F. Maino, D. Meyer, and A. C. Aparicio. Lisp-mn: mobile networking through lisp. *Wireless personal communications*, 70(1):253–266, 2013.

[95] P. Nikander, A. Gurtov, and T. R. Henderson. Host identity protocol (hip): Connectivity, mobility, multi-homing, security, and privacy over ipv4 and ipv6 networks. *IEEE Communications Surveys & Tutorials*, 12(2):186–204, 2010.

[96] E. Nordmark and M. Bagnulo. Shim6: Level 3 multihoming shim protocol for ipv6. Technical report, 2009.

[97] M. O'DELL. Gse-an alternate addressing architecture for ipv6. *http://www. watersprings. org/pub/id/draft-ietf-ipngwggseaddr-00. txt*, 2006.

[98] C. Perkins. RFC 3344. *IP Mobility Support for IPv4*, 8, 2002.

[99] C. Perkins, D. Johnson, and J. Arkko. Mobility support in ipv6. Technical report, 2011.

[100] C. E. Perkins. Mobile ip. *IEEE communications Magazine*, 35(5):84–99, 1997.

[101] D. Phoomikiattisak and S. N. Bhatti. Control plane handoff analysis for ip mobility. In *Wireless and Mobile Networking Conference (WMNC), 2016 9th IFIP*, pages 65–72. IEEE, 2016.

[102] D. Phung, S. Secci, D. Saucez, and L. Iannone. The openlisp control plane architecture. *IEEE Network*, 28(2):34–40, 2014.

[103] A. Raheem, A. Lasebae, M. Aiash, and J. Loo. Supporting communications in the iots using the location/id split protocol: a security analysis. In *Future Generation Communication Technology (FGCT), 2013 Second International Conference on*, pages 143–147. IEEE, 2013.

[104] M. K. Rana, B. Sardar, S. Mandal, and D. Saha. Implementation and performance evaluation of a mobile ipv6 (mipv6) simulation model for ns-3. *Simulation Modelling Practice and Theory*, 72:1–22, 2017.

[105] I. RFC0791. Internet protocol. *J. Postel. September*, 1981.

[106] A. Rodriguez-Natal, S. Barkai, V. Ermagan, D. Lewis, F. Maino, and D. Farinacci. Software defined networking extensions for the locator/id separation protocol. *IETF LISP Working Group Internet-Draft*, 2014.

[107] J. Rosenberg et al. Sip: Session initiation protocol (jun. 2008) retrieved at http://tools. ietf. org/html/rfc3261. *Relevant pages provided*.

[108] D. Saucez and et al. A first measurement look at the deployment and evolution of the locator/id separation protocol. *ACM SIGCOMM Computer Communication Review*, 43(1):37–43, April 2013.

[109] D. Saucez and L. Iannone. O. bonaventure," locator/id separation protocol (lisp) threat analysis. Technical report, RFC 7835, DOI 10.17487/RFC7835, April 2016,< http://www. rfc-editor. org/info/rfc7835.

[110] D. Saucez, L. Iannone, O. Bonaventure, et al. Openlisp: An open source implementation of the locator/id separation protocol. *ACM SIGCOMM Demos Session*, pages 1–2, 2009.

[111] D. Saucez, L. Iannone, O. Bonaventure, and D. Farinacci. Designing a deployable internet: The locator/identifier separation protocol. *IEEE Internet Computing*, 16(6):14–21, 2012.

[112] A. Stockmayer, M. Schmidt, and M. Menth. jlisp: An open, modular, and extensible java-based lisp implementation. In *Teletraffic Congress (ITC 28), 2016 28th International*, volume 1, pages 205–208. IEEE, 2016.

[113] Z. Tang, Y. Zhou, W. Deng, and B. Wang. Lisp-hnm: Integrated fast host and network mobility control in lisp networks. In *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2017 15th International Symposium on*, pages 1–6. IEEE, 2017.

[114] R. Thayer, N. Doraswamy, and R. Glenn. Rfc 2411: Ip security document roadmap. *Internet standard, November*, 1998.

[115] V. Veselỳ, M. Marek, O. Ryšavỳ, and M. Švéda. Multicast, trill and lisp extensions for inet. *International Journal on Advances in Networks and Services Volume 7, Number 3 & 4, 2014*, 2014.

[116] V. Veselỳ and O. Ryšavỳ. Locator/id split protocol improvement for high-availability environment. *ICNS 2015*, page 75, 2015.

[117] C. Vogt. Six/one: A solution for routing and addressing in ipv6. *INTERNET DRAFT, draft-vogt-rrg-six-one-01. txt*, 2007.

[118] Y. Wu, K. Chen, K. Xue, and D. Ni. Nemo-based mobility management in lisp network. In *Wireless Communications and Signal Processing (WCSP), 2014 Sixth International Conference on*, pages 1–6. IEEE, 2014.

[119] H. Zhang, M. Chen, and Y. Zhu. Evaluating the performance on IDLoc mapping. In *Proc. IEEE Global Communications Conference (GLOBECOM)*, November 2008.

List of Publications

- Y. Li, D. Saucez, L. Iannone, and B. Donnet. Stability and consistency of the lisp pull routing architecture. In *Proc. Workshop on Network Measurement (WNM)*, November 2016

- Y. Li and L. Iannone. Using ripe atlas to evaluate the locator/id separation protocol. In *Computer Communications Workshops (INFOCOM WKSHPS), 2016 IEEE Conference on INFOCOM 2016 Conference*, pages 696–697. IEEE, 2016

- Y. Li and L. Iannone. Performance evaluation of locator/identifier separation protocol through ripe atlas. In *Proceedings of the 2016 conference on ACM SIGCOMM 2016 Conference*, pages 561–562. ACM, 2016

- Y. Li, A. Abouseif, L. Iannone, and D. Saucez. Lisp-views: Monitoring lisp at large scale. In *Teletraffic Congress (ITC 29), 2017 29th International*, volume 1, pages 178–186. IEEE, 2017

- Y. Li and L. Iannone. Assessing locator/identifier separation protocol interworking performance through RIPE Atlas. *Computer Networks*, 132:118–128, 2017

- Y. Li, A. Abouseif, L. Iannone, and D. Saucez. Lisp-views: Monitoring lisp at large scale. In *Algotel/Cores 2018*, 2018

Appendix A

Source Code

Some of our implementations are originally written and some are modified from previous open source projects. So we upload related codes to GitHub in my personal repositories as https://github.com/SeleneLI. It is convenient to download and explore the codes. A brief introduction of each repository is provided as follows.

A.1 LISP Stability and Consistency analyzer

https://github.com/SeleneLI/LispMdsAnalyzer

This project is totally originally written to analyze the stability and consistency of LISP mapping system.

A.2 LISP-Views

https://github.com/SeleneLI/LISP-Views

This project is the original codes for LISP-Views architecture.

A.3 LISP measurements on Atlas

https://github.com/SeleneLI/Atlas

This project is used to conduct the experiment on RIPE Atlas and analyze the ping / traceroute results for LISP interworking performance.

A.4 LISP implementation on ns-3

https://github.com/SeleneLI/LISP_ns-3

This project is an implementation of LISP mobility extensions on an existing LISP simulator under ns-3. It also provides the scripts evaluating LISP mobility performance.