University of Würzburg
Institute of Computer Science
Research Report Series

Mapping Systems for Loc/ID Split Internet Routing

Michael Menth[1], Matthias Hartmann[1], Michael Höfling[1]

Report No. 472 May 2010

[1] University of Würzburg
Institute of Computer Science
Chair of Communication Networks
Würzburg, Germany
Email: {menth,hartmann,hoefling}@informatik.uni-wuerzburg.de

Mapping Systems for Loc/ID Split Internet Routing

Michael Menth, Matthias Hartmann, Michael Höfling
University of Würzburg, Institute of Computer Science
Am Hubland, D-97074 Würzburg, Germany
Email: {menth,hartmann,hoefling}@informatik.uni-wuerzburg.de

***Abstract*—The locator/identifier split is believed to be an important principle for scalable Internet routing. In many proposals, an intermediate node at the border of an edge network queries a mapping system to find routing locators for endpoint identifiers and adds this information to each outgoing packet. Such a mapping system must be fast, reliable, secure, and should be able to relay initial packets. In this paper, we present FIRMS, a novel two-level mapping system that fulfills these requirements and evaluate its scalability. We propose a classification of mapping systems, provide a comprehensive review, and compare these mapping systems with FIRMS.**

I. Introduction

Today's IP addresses are both endpoint identifiers (EIDs) as they give names to end systems, and routing locators (RLOCs) as they carry the information about the location of the end system in the Internet. The coupling of both functions currently causes multiple problems in the Internet. Users usually receive IP addresses from the IP number space of their Internet service providers (ISPs). If they change ISPs, they get addresses from a different IP number space of their new ISPs. Thus, renumbering of customer equipment is required. If users keep their IP addresses while changing ISPs, their changed location in the Internet must be reflected in inter-domain routing. Hence, BGP needs to update the routing tables worldwide. This leads to increased BGP signalling rate, fragmented IP number space, and increased BGP routing tables, which are just some obvious drawback of coupling EID and RLOC function in IP addresses.

The locator/identifier (Loc/ID) split principle is expected to overcome the presented problem and in particular scaling issues in the Internet [1], [2]. The addressing consists of two separate parts: the RLOC and the EID. The EID suffices to locate the end system within the edge network and the RLOC indicates the location of the edge network in the Internet. A mapping system glues both together. When a user changes its edge network, the mapping system is updated with the new EID-to-RLOC information. Applications know only EIDs and address packets with a destination EID. Depending on the routing architecture, either the source node or an intermediate node queries the mapping system for the EID-to-RLOC information to add RLOCs to EID-addressed packets in order to make them routable over the Internet. When an intermediate node queries RLOC information, the mapping system must be very fast as packets wait or are dropped until the mapping is available. As the DNS is likely too slow for that purpose, another solution is needed. If the mapping system becomes a vital part of the Internet architecture, it must be resilient to outages, secure, and fast.

In this work, we present FIRMS, a distributed "Future InteRnet Mapping System" which supports routing architectures based on Loc/ID split where intermediate nodes query the mapping system. It requires that EIDs are assigned by authorities in prefix blocks to their owners, just like IP addresses are assigned today. FIRMS is resilient in the sense that parts of the system can fail without causing significant service interruptions. It is secure, fast, and offers intermediate nodes a relay service for packets with yet unknown RLOCs. We believe that FIRMS is an interesting base for a mapping system in the future Internet if a LISP-like architecture prevails. We present a simple performance analysis comparing FIRMS with other approaches. We suggest a classification of mapping systems, use it for a comprehensive review of existing proposals, and compare them with FIRMS.

This paper is organized as follows. In the next section, we explain routing proposals implementing the Loc/ID split in more detail. In Section III we propose a classification for mapping systems. Section IV describes FIRMS in detail. The performance analysis is given in Section V. Section VI reviews other mapping system proposals and compares them with FIRMS. Section VII concludes this work.

II. Routing Architectures Based on Loc/ID Split

In this section we present two classes of future Internet routing approaches: one where hosts query the mapping system for EID-to-RLOC information and another where intermediate nodes query the mapping system. The first may use the DNS as mapping system while the second one needs a faster solution.

A. Loc/ID Split with Mapping Lookup in Hosts

The Identifier/Locator Network Protocol (ILNP) [3], [4] implements the Loc/ID split in hosts. IPv6 addresses are split into two separate fields. The high-order bits serve as RLOC and the low-order bits as EID. ILNP assumes that nodes have upgraded networking stacks and that applications use only DNS names to designate other devices. To communicate with them, a node queries the DNS for new I- and L-records that return the EIDs and the RLOCs for DNS names.

This work was funded by the Federal Ministry of Education and Research of the Federal Republic of Germany (support code 01 BK 0800, G-Lab). The authors alone are responsible for the content of the paper.

The clean-slate Hierarchical Architecture for Internet Routing (HAIR) [5] implements Loc/ID split in hosts, too. With hierarchical IP [6], hosts query the mapping information, insert it into packets, and intermediate nodes only rearrange the order of the information in the headers. The host identity protocol (HIP) [7], [8] also implements the Loc/ID split. It introduces a host identifier tag (HIT) as a location-independent designator of a node, which is used instead of its IPv6 address as identifier on the transport layer. A mapping service maps HITs to IP addresses. An extension to it is the Hierarchical Host Identity Tag Architecture (HHIT) [9].

When RLOCs information is added to packets in hosts, packets can wait until the EID-to-RLOC mapping is returned from the mapping system. This is similar to the resolution of domain names to IP addresses before first packets of a flow can be sent. The DNS can even be reused to provide EID-to-RLOC mappings.

B. Loc/ID Split with Mapping Lookup in Intermediate Nodes

With Loc/ID Separation Protocol (LISP) [10], [11], [12], EIDs are routable only within LISP domains, but not in the global Internet. Packets are tunnelled from the ingress tunnel router (ITR) in the source network to the egress tunnel router (ETR) in the destination network. To that end, the xTRs (ITR or ETR) have globally routable RLOCs as addresses. Two nodes in the same LISP domain communicate with each other just like in the current Internet. When a node in a LISP domain communicates with a node in another LISP domain, packets are sent to the ITR, the ITR queries the mapping system for the EID-to-RLOC mapping of the destination EID, and sends them to the ETR with the respective RLOC over the Internet. The ETR decapsulates the packets which are forwarded according to their EID to their destination. Interworking techniques with the non-LISP Internet exist [13].

LISP is currently under standardization in IETF and already deployed in pilot networks. Also the majority of the Loc/ID split proposals we know add RLOCs to EID-addressed packets in intermediate nodes: Six/One router [14], GLI-Split [15], APT [16], a Novel DHT-Based Network Architecture for the Next Generation Internet [17], the Node Identity Architecture [18], and RANGI [19].

When RLOCs are added to packets in intermediate nodes on the communication path, packets need to be stored, relayed to default nodes that know how to forward them, or dropped until the EID-to-RLOC mapping information is available. The mappings can be stored in a local cache, but cache misses occur. Therefore, EID-to-RLOC queries by intermediate nodes add two new challenges for mapping systems. First, the response time of the mapping system must be very short to minimize the impact of cache misses. Second, it is advantageous if the mapping system offers a relay service so that intermediate nodes can forward packets with missing RLOCs over the mapping system to the destination in order to avoid packet loss or extensive delay. The DNS does not meet these requirements as is too slow and cannot relay packets so that a new mapping system is needed.

III. Classification of Mapping Systems

We first clarify assumptions about EIDs and RLOCs. Then, we propose a classification of mapping systems which shows the design space and helps to distinguish FIRMS from other approaches.

A. Assumptions about EIDs and Mappings

EIDs in the Loc/ID split context should be globally unique. Their uniqueness can be achieved through administrative or statistical means. IP or Ethernet addresses are examples for the first category. Numbers authorities assign address prefixes to organizations which may further partition their obtained address space and assign it to nodes. This leads to structured EIDs. As an alternative, EIDs may be randomly created like in HIP [7]. If they are sufficiently long, the probability for the creation of the same EIDs is very small. These EIDs are unstructured and we call their address space flat. Combinations of hierarchically assigned prefixes and random suffixes have been proposed in [9], [19]. They are semi-structured.

Unstructured EIDs of an organization cannot be aggregated by a common prefix and, therefore, need an EID-to-RLOC mapping per EID even if they have all the same RLOC. When structured EIDs with a common prefix have also the same RLOC, the mapping information can be condensed to an EID-prefix-to-RLOC mapping. The current version of LISP even requires that all EIDs of an assigned EID prefix have the same RLOC. However, then mobility cannot be supported with Loc/ID split. Other routing proposals based on Loc/ID-split assume structured EIDs but individual EID-to-RLOC mappings. FIRMS requires structured EIDs to be scalable but supports per-EID mappings.

B. Classification of Mapping Systems

A map-base is a mapping database. In Figure 1 we propose four categories of mapping systems: direct map-bases, map-request forwarding overlays, two-level mapping systems, and DNS-based mapping systems. We briefly explain their basic structure and operation. ITRs may locally cache mappings to reduce the request loads [20] on any type of mapping system.

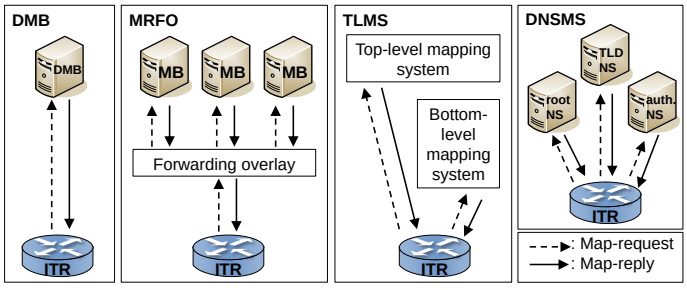


Fig. 1. Classification of mapping systems.

1) Direct Map-Bases (DMBs): DMBs keep the entire mapping information in a single node. There may be only a single central DMB, mirrors thereof in edge networks, or all mappings may be even downloaded to all ITRs. In a

hierarchical system, the global mapping information could be partitioned among multiple DMBs and some mechanism ensures that they receive map-requests only for EIDs they are responsible for.

2) Map-Request Forwarding Overlays (MRFOs): A MRFO consists of distributed map-bases, each only responsible for a partial set of EIDs, and a forwarding overlay. ITRs send map-requests to the forwarding overlay which carries them to the appropriate map-bases. The map-bases return map-replies directly to the ITRs. An ITR just needs to know an entry point to the forwarding overlay, but not the address of an appropriate map-base.

3) Two-Level Mapping Systems (TLMS): In a TLMS, the EID space is partitioned and several so-called bottom-level mapping systems keep EID-to-RLOC mappings for the different EID sets. A top-level mapping system stores EID-to-bottom-level-MS mappings for any EID. An ITR queries the top-level mapping system and uses the result to retrieve the actual EID-to-RLOC mapping from the appropriate bottom-level mapping system. Both the top- and the bottom-level-mapping systems can be implemented as DMBs or MRFOs. An EID-to-bottom-level-MS mapping is likely to be more stable than an EID-to-RLOC mapping. Thus, the update rate at the top-level mapping system is relatively small. Moreover, EID-to-bottom-level-MS mappings may remain longer in a local cache of an ITR which reduces the request frequency at the top-level mapping system. In addition, for structured or semi-structured EIDs, the top-level mapping system may store only EID-prefix-to-bottom-level-MS information which reduces the number of entries. All three aspects improve scalability.

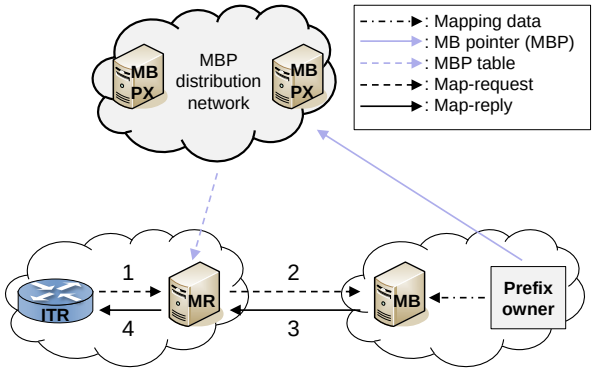
4) DNS-Based Mapping Systems (DNSMS): DNSMSs take advantage of a hierarchical EID structure and essentially work like the DNS. An ITR sends a recursive query to DNS server for an EID-to-RLOC mapping. This server either has the mapping in its cache and returns it or starts an iterative query to find an authoritative DNS server which returns the mapping. In contrast to TLMS, a DNSMS may have multiple hierarchy levels and also DNS servers may use caches to reduce the number of their map-requests.

IV. The FIRMS Architecture

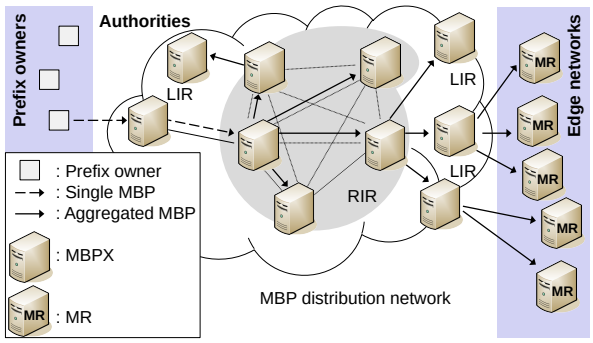
In this section we propose the "Future InteRnet Mapping System" (FIRMS) which is a two-level mapping system. We describe its architecture, specify its operation, and discuss its resilience and security features. We use the LISP's nomenclature (EID, RLOC, ITR, ETR), but FIRMS is not limited to LISP.

A. General Idea

Figure 2(a) illustrates the basic structure and operation of FIRMS. EIDs are assigned to their owners in prefix blocks and each prefix owner provides a map-base (MB) holding the EID-to-RLOC mappings for all its EIDs. The operation of the MB may be delegated to a specialized company. A map-base pointer (MBP) is a data structure containing information about the MB. The prefix owner registers this information in the global MBP distribution network which collects all MBPs and constructs a global MBP table. Each ITR is configured with a map-resolver (MR). The MR registers at the MBP distribution network and receives a copy of the global MBP table. When the ITR requires an EID-to-RLOC mapping for an EID, it sends a map-request to its MR. The MR looks up the address of the responsible MB in its local copy of the MBP table and forwards the map-request to that MB. The MB returns a map-reply containing the desired EID-to-RLOC mapping to the MR which forwards it to the ITR. If a non-existing mapping is queried, a negative map-reply is returned. This design requires that MRs and MBs have globally reachable RLOC addresses. In the following we present ITRs and MRs as two different entities because they have different functionality. However, the MR functionality may be integrated in an ITR which saves communication overhead and simplifies the design.



(a) Basic operation of FIRMS.



(b) Propagation of MBP updates.

Fig. 2. FIRMS concept

B. Map-Base Pointer Distribution Network

We explain how MBPs are distributed from prefix owners to MRs. We assume that EIDs are assigned in a similar way as IP addresses are assigned today; many routing proposals even assume that EIDs are IP addresses. IANA delegates IP address blocks to the five regional Internet registries (RIRs): AfriNIC, APNIC, ARIN, LACNIC, and RIPE NCC. They delegate subsets thereof to local Internet registries (LIRs). Both RIRs and LIRs partition the address space in prefix blocks and assign prefixes to organizations (prefix owners).

Every RIR or LIR runs a map-base pointer exchange node (MBPX). Figure 2(b) shows that the MBPX of a LIR (LIR-MBPX) is connected to the MBPX of its RIR, and the MBPXs of the RIRs (RIR-MBPX) are fully meshed. This constitutes the MBP distribution network. The prefix owner adds, changes, or removes MBPs for its EID prefixes at the MBPX of its LIR or RIR. An LIR-MBPX forwards this data to its superordinate RIR-MBPX. The RIR-MBPX collects the MBPs for all EID prefixes under its control and compiles a regional MBP table. The MBP tables are exchanged among all RIR-MBPXs so that each of them has a copy of the global MBP table. They push this information to their subordinate LIR-MBPXs which forward it to all MRs that have registered for that service. An involvement of RIRs or LIRs for the support of Internet services is not uncommon. For instance, RIRs and LIRs play an active role for reverse DNS lookup.

To facilitate incremental updates to MBP tables, the RIR-MBPX collects individual MBP updates from prefix owners over some time and provides sequentially numbered aggregated updates. It pushes them to the other RIR-MBPXs and its subordinate LIR-MBPX. When an RIR-MBPX, LIR-MBPX, or MR receives such an update, it applies the changes to its local copy of the MBP table and forwards the updates to all its subordinate LIR-MBPXs or MRs. The numbering of the updates contributes to the consistency of all MBP tables. If an update is received with an unexpected number, missing updates are detected and can be requested.

C. Mapping Retrieval

To minimize query overhead, ITRs and MRs have local caches for EID-to-RLOC mappings. To avoid stale information, mappings are automatically purged from the caches after their time-to-live has expired. Figure 3 illustrates how EID-to-RLOC mappings are retrieved in combination with caches. When the ITR requires a mapping, it first checks its cache and can often retrieve it immediately. In case of a cache miss, the ITR sends a map-request to the MR.

When the MR receives a map-request from an ITR, it first searches its cache and, if successful, sends a map-reply back to the ITR. If unsuccesful, the MR selects an appropriate MB for the EID from its local copy of the global MBP table and sends a map-request to that MB. The MR keeps a state for the requested EID so that a map-reply can later be returned to the requesting ITR. The state is removed when the MR returns the requested information to the ITR or when a timer expires.

When the MB receives a map-request, it retrieves the EID-to-RLOC mapping from its database and sends it back to the MR in a map-reply. The MR stores the mapping of the map-reply in its cache and sends a map-reply back to the ITR which also stores the mapping in its cache. The caches at the ITRs and MRs minimize the retrieval time for the mappings and reduce the frequency of map-requests. Performance issues of caches have been discussed in [20].

We propose several enhancements to improve the speed and scalability of the mapping retrieval.

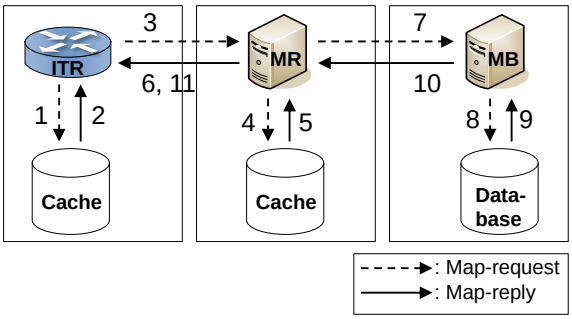


Fig. 3. Cascading mapping retrieval.

- MRs and ITRs should limit the rate of map-requests for the same EID to avoid outgoing map-request storms.
- Every EID has its own RLOC. If EIDs of a common prefix block have the same RLOC, their EID-to-RLOC mappings may be aggregated to a single EID-prefix-to-RLOC mapping. This saves storage in caches and databases and reduces the rate of map-requests.
- If an MR serves only a single ITR, their caches are likely to have the same content so that advantage cannot be taken from the cache at the MR. Hence, several ITRs should be configured with the same MR. Then, the MR may be able to serve an ITR's map-request from its cache with EID-to-RLOC mappings that have been requested earlier by other ITRs.
- Alternatively, the MR functionality may be integrated in ITRs. This saves communication overhead and simplifies the overall structure. Then, the MR is mainly an interface to logically separate ITR and MR functionality within the same physical node.

D. Packet Forwarding

When an ITR receives a packet addressed to an outbound EID, it tries to retrieve the EID-to-RLOC mapping from its local cache and, if successful, tunnels the packet to the ETR whose RLOC was given in the mapping. In case of a cache miss, the ITR retrieves the mapping over the network which is a time-consuming process. This can happen for the first packet of a communication session when a new flow to a previously not contacted EID is established. The arrival rate of such packets is most likely rather low. In contrast, when traffic is shifted from one ITR to another, the rate of packets with missing RLOCs can be very high. This can happen, for example, when the primary ITR of a networks fails, when the internal routing is changed, or when load balancing policies change. There are three options to handle such outbound packets in the meanwhile: they can be dropped, stored, or relayed to another node that knows how to forward them.

When the ITR drops packets, most applications will resend them, and by then the mapping is hopefully available in the ITR's local cache. This might work for the first packet of a communication, but even this packet can be important, e.g., the initial SYN packet of a TCP connection setup. Losing the

first packet can significantly impede the communication setup. When a large number of flows is shifted from another ITR, an immense number of packets is dropped until a mapping can be retrieved from the MS.

As an alternative, the ITR stores the packet until the requested mapping returns from the MR. Then, the ITR can add the RLOC to the packet and send it. This option requires a large buffer to store such packets. Additional logic is needed to continue the processing of the packets as soon as the missing mappings arrive or to drop them when a timer expires. The buffer may overflow and packets may be lost, especially when packets arrive at a high rate. This gives rise to potential attacks where attackers send packets to the ITR with yet unknown destination EIDs. Thus, this option requires complex engineering and still cannot avoid packet loss.

Packet relaying to another node that knows how to forward the packet seems a promising idea because it avoids the drawbacks of dropping and storing. Therefore, it has been proposed also for other mapping systems [21], [22], [23]. Figure 4 illustrates how packet relaying can be realized in FIRMS. Normally, the ITR has the EID-to-RLOC mapping in its cache and tunnels the packet to the ETR. In case of a cache miss, the ITR tunnels the packet to the MR. If the MR finds the required mapping in its cache, it tunnels the packet to the ETR. Otherwise, the MR tunnels the packet to the appropriate MB. The MB has the mapping in its database and tunnels the packet to the ETR. This design has several nice properties.

- Only the MR and the MB are involved in the relay process. They are operated by the sender's network and the prefix owner or on behalf of them so that these elements have incentives in forwarding the data. In particular, no elements of public infrastructure or other private networks are involved. This is different in other proposals where relayed packets are transmitted over an overlay network [21], [22], [23].
- If the MB is collocated with the destination network of the EID and near the ETR, the path of the relayed packets is hardly stretched.
- Relayed packets can be interpreted as implicit map-requests and save explicit map-requests. That means, MRs or MBs not only tunnel the relayed packets to ETRs when they have appropriate mappings, they also respond with map-replies. When an ITR relays multiple packets with the same EID, map-reply storms may occur and appropriate measures should be taken to avoid them (see Section IV-C).

E. Extensions for Resilience

Any system component in FIRMS can fail. We propose simple extensions and show that FIRMS can cope with the failure of any component through replication. FIRMS is also compatible with resilience methods presented in the LISP context.

1) FIRMS-Specific Extensions: RLOCs can become unreachable. If an edge network is multihomed, it is reachable over alternative RLOCs that also appear in the EID-to-RLOC mappings. When an ITR detects problems with an RLOC, it marks the particular RLOC in its cache as unreachable for a while and uses an alternative RLOC instead.

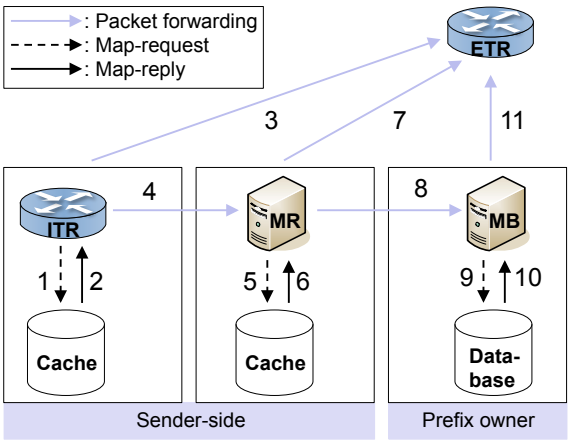


Fig. 4. Cascading packet forwarding.

MRs can fail. ITRs can be configured with several MRs. When an ITR detects the failure of an MR, it marks the MR as unreachable for a while and contacts another configured MR.

MBs can fail. A prefix owner can have multiple MBs with identical mappings and record their addresses in the MBP. When an MR detects the failure of an MB, it marks the MB as unreachable for a while and contacts an alternative MB whose address is given in the MR's local copy of the MBP table.

MBPXs can fail. As a consequence, MRs do not receive updates for the MBP table in time. An MR can register at several MBPXs, and if one of them fails, the MR still receives updates from the other MBPXs.

2) Enhancements with Support of the LISP Encapsulation Header: The LISP encapsulation header reserves four bytes as "locator status bits" [10]. These bits correspond to an ordered list of RLOCs in the EID-to-RLOC mapping and indicate which of them are operational. The prefix owner can change this information at the MBs to give the ITRs a hint which RLOCs are currently reachable.

Another approach proposes that EID-to-RLOC mappings are equipped with version numbers to facilitate detection of outdated information. LISP-versioning [24] proposes that the ITR adds the current version number for the mapping of the source EID in the LISP encapsulation header. The ETR examines the version number in the encapsulation header of incoming packets and compares them with the version number in the corresponding mappings stored in the local cache of the collocated ITR. If the mapping in the local cache is outdated, the ITR sends a map-request for the respective EID to update the mapping in its cache. This mechanism helps to keep track of mapping changes.

F. Extensions for Security

The MR must rely on the authenticity of the MBPs and the EID-to-RLOC mappings. Therefore, security features are needed to ensure that only the owner can change for a prefix the MBP at the MBPX of the LIR/RIR and the mapping data in the MBs, and that the mappings from the MB reach the MR without changes.

Figure 5 visualized the security concept. In [25] an extension to the ITU-T X.509 v3 standard for a public-key infrastructure has been proposed that allows to bind a list of IP prefixes to the subject of a so-called resource certificate. It is already provided for use by APNIC [26]. We propose to use it for FIRMS to transfer the right-to-use for IP prefixes from IANA through the RIRs and LIRs to prefix owners. Then, prefix owners can authenticate themselves as the rightful owners of their EID prefixes. They use this feature when they add, modify, or remove EID-to-RLOC mappings at the MBs or when they add, modify or remove MBPs at the MBPX of the LIRs/RIRs.

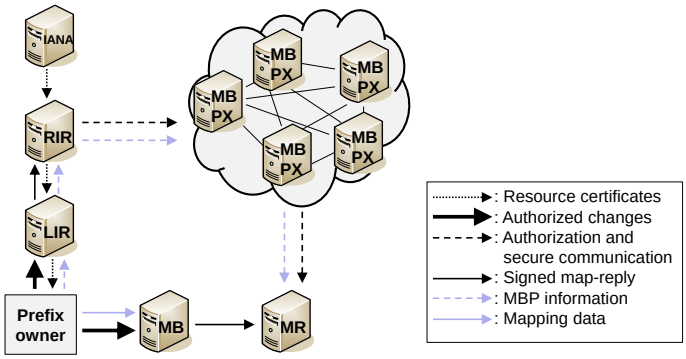


Fig. 5. Security concept for FIRMS.

We assume that neighboring MBPXs trust each other. They authenticate each other and use transport layer security (TLS) [27] to exchange updates about MBPs. MRs should trust the MBPXs from which they want to receive the MBP information via a push service which can also be enhanced by TLS. Therefore, the MRs can trust the information in their global MBP tables.

An MR must be able to verify whether the mapping data obtained from the MB is authentic. To that end, the public key of the MB is also added to the trusted MBP information. The MB signs the mapping data including a time stamp with its private key and sends this information in a map-reply to the MR. The MR can validate it with the public key of the MB which is available in the local copy of its global MBP table. Successful validation tells the MR that the mapping data has been sent by the queried MB, they are recent, and they have not been changed.

An MR receives the MBPs over secured connections from trusted MBPXs that have checked the authenticity of the prefix owners which have registered the MBP information. Therefore, the MR can immediately use the public keys of the MBs in the MBPs to validate map-replies. As a consequence, the time required for the validation of map-replies is short and predictable. This was a major goal. Let us assume a more intuitive design alternative where prefix owners use their hierarchical resource certificates to prove the validity of MBPs instead of transporting them over secured connections through the MBP distribution network to the MRs. Then, an MR needs to recursively validate a certificate chain before it can trust the MBP information. This adds to the lookup delay of EID-to-RLOC mappings which may become significant. Another disadvantage is that MRs burden Internet registries with a tremendous load of validation requests. In particular in this light, our chosen design is fast and efficient.

G. Distinction of FIRMS versus Other Approaches

We classify FIRMS as a two-level mapping system (TLMS). The local MBP table serves as top-level mapping system and the MBs are the bottom-level mapping systems. As the global MBP table is downloaded to the MR, the first lookup can be done locally and does not cause communication overhead.

FIRMS is significantly different from direct map-bases (DMBs). The MBs under the control of the prefix owners constitute a massively distributed system while DMBs store the global mapping information in a single location. This makes FIRMS more scalable than DMBs and facilitates updates.

Map-request forwarding overlays (MRFOs) send map-request over a forwarding overlay. This takes longer than sending map-requests directly to an appropriate MB. Having control over the MRFO infrastructure or at least parts of it means having control over the Internet because map-requests could be selectively blocked. This is not possible with FIRMS where map-requests and map-replies are carried like normal traffic between MRs and MBs. Moreover, it is not clear whether the operators of the MRFO are willing to relay data packets. Apart from that, it is hard to provide a security infrastructure which can validate map-replies without any communication overhead as the ITR does not know the responding MB a priori.

The difference between FIRMS and DNS-based mapping systems (DNSMS) is that an MR in FIRMS definitely knows the responsible MB while a DNS server possibly knows only another DNS server that knows the authoritative DNS server. Thus, the resolution might take longer. Furthermore, DNSMS are hierarchically structured and top-level servers respond only to iterative queries, i.e., they only respond to them with the address of delegation servers but do not forward them. This implies that data packets cannot be relayed by a DNSMS. Like with MRFOs, it is hard for DNSMS to provide a security infrastructure without extra overhead for the validation of map-replies.

V. Performance Analysis

In this section we estimate the expected loads on various system components in FIRMS and show that they are in a manageable order of magnitude. Other mapping systems have a higher load on critical parts of their infrastructure.

A. Record Sizes

We calculate the size of EID-to-RLOC and MBP records in FIRMS. We assume that both EIDs and RLOCs have the same format as IPv6 addresses which are 16 bytes long. Edge networks can take advantage of multihoming more easily with Loc/ID split. Being connected to more than 4 ISPs has only limited benefit [28]. Therefore, we assume that nodes are usually connected to the Internet over three providers which results in an average number of three RLOCs per EID-to-RLOC record. This record contains additional information like a time-to-live (5 bytes), some traffic engineering attributes (10 bytes), and a security signature with a timestamp (16+5 bytes) so that its average size is about 100 bytes.

MBPs consist of an EID prefix (8 bytes), the RLOCs (16 bytes each) and public keys (64 bytes each) of the corresponding MBs, and some additional attributes for traffic engineering (10 bytes). For resilience and load balancing purposes, each EID prefix should have two separate highly available MBs so that we assume two MBs per MBP. This sums up to an average size of 178 bytes per MBP record.

B. Storage Requirements

We estimate the storage requirements of a MB and for the MBP table in FIRMS and for a database in a direct map-base (DMB). The current number of prefixes on the Internet is about $n_{pref} = 10^6$ [29] while the current number of hosts is about 10^9 [30]. This leads to an average number of $n_{pref}^{EIDs} = 10^3$ hosts per prefix. With the Internet of things and other novel applications, we assume that the number of hosts (and EIDs) per EID prefix will dramatically increase in the future. The same holds for the number of EID prefixes.

A MB needs to store on average $n_{pref}^{EIDs} = 10^3$ EID-to-RLOC mappings (100 Kbyte) today and a multiple of them in the future. But that does not seem a critical value. The MBP table keeps $n_{pref} = 10^6$ MBP entries (178 Mbyte) and a multiple in the future. Also that seems feasible. The database of a DMB stores all $n_{pref} \cdot n_{pref}^{EIDs} = 10^9$ global EID-to-RLOC mappings (100 Gbyte) and a multiple in the future. This is already a rather challenging order of magnitude given the fact that these values need to be updates quite often and possible distributed to a large number of mirrors in realtime.

C. Update Loads

We calculate the update load in a MB and for the MBP table in FIRMS and for the database of a DMB. The validity of a MBP can outlast the contract between a customer network and an ISP when the MB is outsourced to special MB providers so that the RLOCs of the MBs do not change. A recent study showed that only 32 percent of small and medium companies changed their provider in 2008 [31]. Thus, we assume that prefix owners change their MBPs every 3 years which also includes key updates for the MBs. With 10^6 prefixes, this leads to an average update rate of 38 prefixes or 6.77 Kbytes per hour. Even a large multiple seems quite feasible in particular as MBP updates are aggregated and not sent individually as this rough calculation assumes. RIRs have between 1000 and 6500 LIRs below them. Hence, they need to push 1.05 Gbytes daily towards the LIRs. This is a large amount of data but breaks down to a continuous upload rate of 12.2 Kbytes/s so that even a large multiple is not problematic.

EID-to-RLOC mappings are less stable when nodes become increasingly mobile. We assume that an EID changes its mapping once a month. This is rather an average over all nodes than a typical value since some devices are significantly more mobile than others. A MB in FIRMS which is responsible for a single EID prefix with 10^3 EID-to-RLOC mappings encounters 33 updates per day. This is more than feasible even if a MB stores the mappings for multiple EID prefixes and if the number of EID is orders of magnitude larger. In contrast, the database of a DMB with 10^9 entries faces 386 updates per second which is significant as they need to be propagated to all mirrors in almost realtime to support mobility. Thus, the update load of the database of a DMB can be problematic.

D. Map-Request Loads

We estimate a lower bound for the expected rate of global map-requests which originate from edge networks due to cache misses. Currently, about 183 million domain names are registered. We use data from VeriSign (.com, .net) and Denic (.de) to estimate the worldwide DNS query load at second-level domains. VeriSign currently experiences an average load of 38 billion DNS queries per day for the 91 million registered .com and .net domains [32]. DENIC has about 7 billion queries for their 12 million .de domains [33]. We sum up these values and extrapolate them for all 183 million domains, leading to a total average load of 80 billion queries per day, 925,000 queries per second, or about 740 Mbit/s when we assume the size of a map-request packet to be 100 bytes including all headers. Peak request rates are about twice as high. This is only a lower bound for EID-to-RLOC requests. Results from DNS queries are usually cached on different levels of the resolution hierarchy so that the queries at the big registrars heavily underestimate the real number of DNS requests issued by hosts. Most mapping systems do not have similar hierarchical caching systems and, therefore, the expected rate for worldwide map-requests that cannot be resolved from local caches is by orders of magnitude larger, in particular in the future.

High loads of map-requests are problematic for mapping systems with strong hierarchies. Obviously, a centralized DMB requires a lot of CPU and transmission capacity to answer the worldwide map-requests; mirrored DMBs can handle the request load much better. However, also map-request forwarding overlays (MRFOs) with hierarchical structures such as LISP+ALT [34] are likely to have a few nodes facing an extremely large load of map-requests. This is not costly for its operator and creates a political issue: the worldwide mapping system should not be controllable by a few ISPs. It is better when critical infrastructure is in public hands and causes only moderate operational costs. With FIRMS, the worldwide request load is not problematic since a MR handles only the

load originating at an ITR and the MB handles only the request load for a single or a few EID prefixes.

E. Resolution Delay

DMBs can resolve map-requests very quickly since all worldwide mappings are available at least within the same domain when mirrors exist. To resolve a map-request, FIRMS takes about one round trip time from the MR in the source network to the MB which is most likely located in or near the destination network. In contrast, general MRFOs experience a significant path stretch when map-requests are carried over several logical hops in the overlay. Additional end-to-end delay possibly accumulates from less efficient forwarding in the overlay compared to simple packet forwarding.

VI. Comparison of Mapping Systems

We review a large number of mapping systems in the context of Loc/ID split based routing, classify them into the categories presented in Section III, and compare them with FIRMS.

A. Direct Map-Bases (DMBs)

We compare FIRMS with general DMBs and give examples for them.

1) Comparison of FIRMS and General DMBs: DMBs hold the global EID-to-RLOC mapping information in a central map-base which may be replicated to mirrors. FIRMS also collects a global MBP table, but that table stores information per EID prefix rather than per EID which significantly reduces the storage requirements. Moreover, MBP information is more stable than EID-to-RLOC mappings. Changes of EID-to-RLOC mappings need updates of all mirrors of an DMB which is quite an effort so that frequent mapping changes should be avoided. This is different with FIRMS. Changes of EID-to-RLOC mappings are performed only in the very few MBs of a prefix owner and can be done quickly without causing scalability concerns. Only changes of MBP data need to be globally distributed, but this information is rather stable. Hence, FIRMS has clear advantages over DMBs regarding memory requirements and mapping dynamics.

2) LISP-NERD: The "Not-so-novel EID to RLOC Database" for LISP (LISP-NERD) [35] assumes that EIDs are assigned to organizations by authorities and these organizations run a map-base (called NERD) with authoritative mappings. One or several of such authorities exist. An ITR is configured with the addresses of possibly several authoritative NERDs and pulls the entire mapping information from them upon system start. To facilitate incremental updates, changes to the NERD are associated with a version number and a change file. ITRs regularly poll the NERDs for their latest version numbers and download and apply the change files to their local database if needed. All information sent from the NERDs to the ITRs is digitally signed using X.509 certificates. As all mappings are locally available at the ITRs, cache misses and querying delay cannot occur. This was a major design goal.

The current assumption in LISP is that all EIDs of an assigned prefix have the same EID-to-RLOC mapping so that NERDs store in fact EID-prefix-to-RLOC mappings. As a result, a global NERD has the same number of entries as the MBP table in FIRMS. Extending NERD to finer mapping granularity leads to scalability problems while FIRMS easily copes with EID-to-RLOC mappings by design.

3) APT: APT is "A Practical Tunnelling architecture" [16] and comes with a tunnelling design from ITRs to ETRs, a mapping distribution system [36], and a failure handling design. Like in LISP, EID prefixes are mapped to RLOCs. APT's mapping system assumes that each ISP has a default mapper (DM), i.e., a mirror with the global mapping information. DMs of neighboring ASes know each other and exchange mapping information via a mapping dissemination protocol using signed messages. The prefix owners inject the mapping information into the DMs of their ISPs. Whenever new information is available, DMs push it to their neighboring DMs. When an ITR encounters a cache miss for a packet destined to an unknown EID, the ITR sends the packet to the DM of its own domain. The DM relays chooses a single RLOC, returns it to the ITR, and relays the packet to an appropriate ETR. This is a significant difference to NERD where all ITRs have the full mapping information and cache misses are avoided.

4) IVIP's Fast Push Mapping System: IVIP is an alternative to LISP [37] and has its own "fast-push" mapping system [38]. So-called root update authorization systems (RUAS) are the source of the mapping information. Each of them is responsible for a different EID prefix. They partition the EID address space and assign it to user organizations. They also store RLOCs for micronets which are arbitrarily long EID prefixes, i.e., possibly also EIDs, on behalf of the prefix owner. The RUAS push these micronet-to-RLOC mappings over a fast mapping distribution network to multiple full database query servers (QSDs) that are queried by ITRs. The QSD can be part of ITRs or standalones. Thus, this concept is similar to NERD or APT depending on which of the both versions is chosen. However, it supports a finer mapping granularity than NERD and APT. To minimize micronet-to-RLOC mapping updates by user organizations, updates should be charged by the RUAS. In contrast to other approaches, only one RLOC is stored per micronet. This is possible since IVIP assumes that edge networks hire third parties to effect realtime updates to the mapping system to take advantage of multihoming for inbound traffic engineering and service restoration in case of ITR/ETR failures.

B. Map-Request Forwarding Overlays (MRFOs)

We compare FIRMS with general MRFOs and give examples for them.

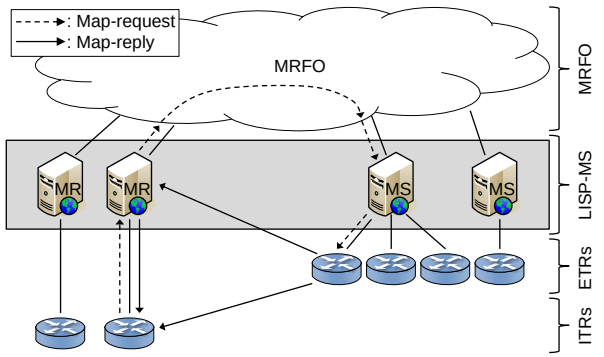
1) Comparison of FIRMS and General MRFOs: MRFOs carry map-requests from ITRs to map-bases. ITRs must trust the nodes in the overlay that they forward their map-requests correctly and do not drop them as the ITR is not necessarily a customer of the ISPs running the MRFO. Hence, appropriate business models are needed which is an unsolved problem. Another option is that the MRFO is run by public authorities. It carries a lot of map-request traffic and possibly also relayed packets. This makes its operation costly which is a problem

for this deployment option. When packets are relayed over the MRFO, they possibly experience a much longer path than packets sent over the normal inter-domain path. Therefore, packet re-ordering is quite likely which can cause problems for some applications. Parts of the MRFO can fail or be attacked which is especially dangerous in case of a hierarchical or semi-hierarchical overlay structure. This is problematic since customers cannot increase the availability of the MRFO by themselves. Hence, MRFOs require backup concepts to avoid service degradation in failure cases.

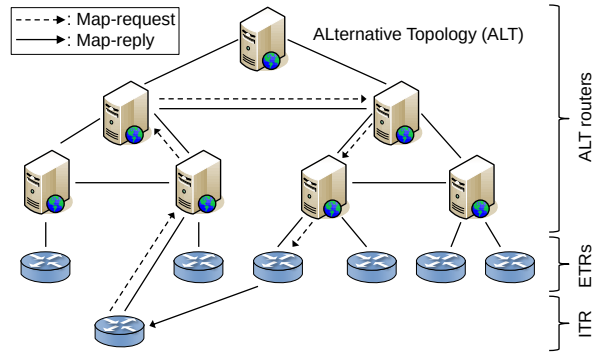
In FIRMS, the MR is under the control of the source network operator and the MB works on behalf of the prefix owner of the destination address. Thus, they have natural incentives to forward map-requests and relayed traffic to enable communication between source and destination. Map-requests are carried from the MR to the MB on the normal inter-domain path so that customer-provider relationships guarantee that traffic is not dropped. If the prefix owner chooses a MB that is located near most of its devices, then the path stretch for relayed packets is likely to be small. In FIRMS, all critical system elements can be replicated by the source network or the prefix owner. So, they can trade between availability of the mapping system and expenses for backup devices. Those are clear advantages of FIRMS compared to MRFOs.

2) LISP-MS: A wide-spread assumption in LISP is that ETRs are authoritative sources of the mappings and perform the MB function. They generally belong to the customer network and are configured by its administrator with all EID-prefix-to-RLOC mappings for the site's EID prefixes. A MRFO may glue ITRs and the distributed ETR-map-bases together, but possibly requires special protocols for communication with them. LISP-MS [39] does not propose an actual mapping system. Figure 6(a) shows that it provides an interface hiding the specifics of different MRFOs from ITRs and ETRs using map-resolvers and map-servers. ITRs send map-requests to map-resolvers which inject them into the MRFO. ETRs register the EID prefixes they are responsible for with map-servers. The map-servers receive from the MRFO appropriate map-requests and forward them to the ETRs. The ETRs respond map-replies either directly to ITRs or to map-resolvers. In the latter case, a map-resolver can cache the mappings and also respond to map-requests with map-replies from the cache.

3) LISP+ALT: The LISP ALternative Topology (LISP+ALT) [34] is an MRFO. In combination with LISP-MS it is the currently preferred mapping system for LISP. So-called ALT routers build an overlay network, the ALT, which is visualized in Figure 6(b). They are associated with EID prefixes and connected in a semi-hierarchical manner with respect to these prefixes. Shortcuts are possible on the same hierarchy level. The bottom-most ALT routers must be connected for each of its associated EID prefixes to at least one authoritative ETR. ALT routers communicate to peering ALT routers via BGP and exchange aggregated EID prefixes that can be reached through them. In contrast to normal inter-domain routing, ALT routers possibly aggregate



(a) LISP-MS consists of map-resolvers (MRs) and map-servers (MSs). They hide specifics of the MRFO from ITRs and ETRs which have the mappings.



(b) LISP+ALT is a semi-hierarchical MRFO.

Fig. 6. LISP currently uses LISP-MS with LISP+ALT.

prefixes received via BGP before forwarding them.

Map-requests are addressed to the queried EID. ITRs tunnel these requests to an ALT router. The ALT router forwards the map-request based on its destination address either to subordinate, peering, or superordinate ALT routers according to the prefix information provided via BGP. Eventually the map-request reaches the appropriate ETR which responds a map-reply directly to the ITR. The operation of LISP+ALT is very efficient since routers do not need to process the packets in a special way, they just forward them. In case of a cache miss at the ITR, packets can also be carried over the ALT, but this is deprecated in the current LISP proposal.

4) LISP-CONS: LISP-CONS stands for "Content distribution Overlay Network Service for LISP" [21] and was a predecessor to LISP+ALT. LISP-CONS does not necessarily use BGP for communication between nodes of the hierarchy. Map-replies are returned from the ETRs back to the ITRs over the overlay network which is also different in LISP+ALT.

5) EMACS-LISP: EMACS-LISP stands for "EID Mappings Multicast Across Cooperating Systems" for LISP [22]. ETRs join multicast groups for all EID prefixes they are responsible for. If that prefix is X.Y.A.B/n, the address of the corresponding multicast group is, e.g., 238.1.X.Y. In case of a cache miss for EID X.Y.A.B, the ITR sends the data packet to the corresponding multicast group so that all ETRs of that group receive it. All ETRs having appropriate mappings for

the requested EID can respond with a map-reply. However, when data packets are relayed over this structure, only one of these ETRs should deliver the packet to avoid duplicates at the destination. This approach has several drawbacks. Up to 2^{16} multicast groups need to be maintained in BGP and a lot of unnecessary extra traffic is generated through multicast delivery.

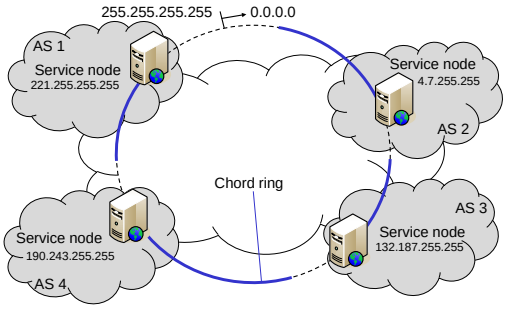


Fig. 7. LISP-DHT uses a modified Chord-ring as a MRFO.

6) LISP-DHT: LISP-DHT stores mappings in a distributed hash table (DHT) [40]. Figure 7 illustrates that map-bases join a Chord ring as so-called service nodes to build a DHT. They have an ID that determines their position within the ring structure. Some modifications are applied to standard Chord. The ID of a service node is the highest number in the EID prefix for which it is responsible. Thus, service node IDs and EIDs are taken from the same number space. The unhashed EIDs of map-requests are used for message forwarding in the DHT. Thus, a map-request is carried within the DHT over several hops to the service node with the smallest ID that is at least as large as the requested EID. These changes ensure that map-requests are forwarded to the service nodes that are responsible for them so that they can answer a map-reply to the requesting ITR. An important feature of LISP-DHT is that prefix owners keep control over the mappings as they are kept local in the service nodes. If a service node is responsible for several EID prefixes, it has several IDs and is connected to the Chord ring at several positions. To prevent malicious nodes from EID prefix hijacking, joining service nodes must be authenticated as the rightful owners of their EID prefixes. For that purpose, the use of X.509 resource certificates [26] is proposed similarly as in FIRMS. To inject map-requests, ITRs join the Chord ring as stealth nodes which do not participate in message forwarding or other critical tasks. To address resilience concerns, LISP-DHT uses backup nodes providing the same mappings as normal service nodes.

7) ER-MO: In [23], a mapping and relaying system is presented which combines techniques similar to LISP+ALT and LISP-DHT. A customer network stores the mappings for its EIDs in a map-base which is part of a mapping overlay (MO) very similar to LISP-DHT. However, Kademlia is used instead of Chord as DHT, and mappings are stored per EIDs instead of per EID prefix. Thus, a map-base joins the DHT as a service node once for each EID under its control. This induces significant management overhead. The relaying system consists of EID routers (ER) which learn EID-prefixes from ETRs via BGP and relay packets if needed.

8) CoDoNS: CoDoNS stands for Cooperative Domain Name System [41]. It is proposed as a substitute for the DNS and is implemented based on a DHT called Pastry. CoDoNS distributes mapping information multiple times in the DHT to achieve an access time of practically O(1). Large organizations should participate in CoDoNS with at least two nodes. These nodes store data from other organizations and the organization's own data are probably stored on nodes of other organizations. This property is hard to accept in practice which is also an argument against the straightforward use of DHTs as a mapping system.

9) HIP-DHT: HIP-DHT [42] is intended for a HIP-context and implements two mapping services for different purposes. The HIT lookup service maps a text name to a HIT, and the address lookup service maps HITs to RLOCs. In both cases, a simple DHT can be used and the authors specify how their concept works with OpenDHT. Normally, the function of the HIT lookup service is performed by DNS and the function of the address lookup service is performed by a rendezvous server. HIP-DHT is an alternative to DNS and particularly useful when legacy DNS servers do not support HIP. In [42] security concerns are listed pointing out potential map-reply spoofing attacks leading to stale information or mapping pollution since authentication is not required to register new or already existing mappings in the system.

C. Two-Level Mapping Systems (TLMS)

We explain why FIRMS can be classified as TLMS and how it distinguishes from other TLMS proposals. Then, we review other TLMSs.

1) FIRMS as a TLMS: FIRMS is a TLMS. The global MBP table serves as top-level mapping system for EID-to-MBP mappings and can be seen as a local copy of a DMB. The MB serves as bottom-level mapping system for EID-to-RLOC mappings and can also be seen as a DMB. FIRMS distinguishes from the following TLMSs because the query of the top-level mapping system does not involve communication overhead. FIRMS is the only TLMS that provides a sound security and resilience concept.

2) RANGI: In the "Routing Architecture for the Next Generation Internet" (RANGI) [19], the name space of host identifiers (HI) is partitioned by prefixes among administrative domains (ADs). HIs consist of two parts: the globally unique AD ID which is possibly assigned by some central numbers authority like IANA and a cryptographical part that is generated as a hash containing the AD ID and a public key value like in HIP. An AD takes care that the HIs under its control are unique. RANGI uses a hierarchical DHT to map HIs to RLOCs. A top-level DHT guides map-requests to bottom-level DHTs using the AD ID in the HI. The bottom-level DHTs uses the unstructured cryptographical part of the HI to resolve the actual mapping and send map-replies to ITRs. The RANGI design implies that all map-requests are carried

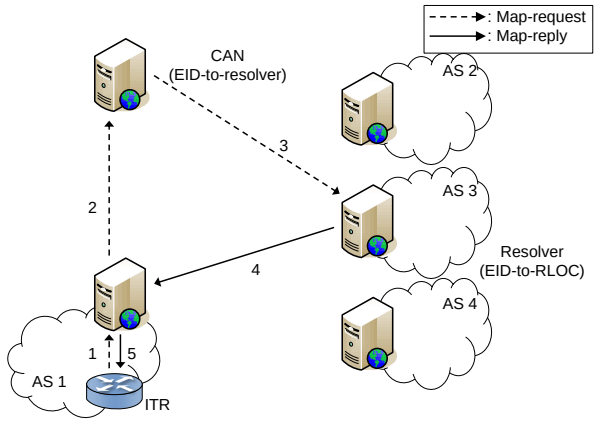
through the top-level DHT. Decoupling the lookup removes load from the top-level mapping system. Nodes can first query the top-level mapping system for the relatively stable AD-ID-to-bottom-level-DHT pointer and cache this information so that future map-requests can directly go to appropriate bottom-level DHTs. FIRMS goes one step further. The map-resolvers have the global MBP table so that the top-level mapping system does not need to be queried remotely which simplifies relaying of packets. However, as hosts resolve the HI-to-RLOC information, delay and relay requirements are relaxed so that a DNS-based mapping system like in ILNP would also suffice.

3) DHT-MAP: DHT-MAP [43] supports a flat identifier space. Figure 8(a) shows the structure of DHT-MAP. Each autonomous system (AS) runs a resolver (MB in FIRMS) which stores the AS-specific EID-to-RLOC mappings for the EIDs supported within the AS. The resolver represents the bottom-level mapping system and can be classified as DMB with partial knowledge. The top-level mapping system stores EID-to-resolver mappings (MBP in FIRMS). It consists of a MRFO that is implemented as a content addressable network (CAN) which is a special type of DHT. ITRs are connected to a resolver. When an ITR encounters a cache miss, it sends a map-request including the packet to the resolver. If the resolver knows the EID-to-RLOC mapping, it tunnels the packet to the ETR and returns a map-reply to the ITR; otherwise, it sends the map-request including the packet into the CAN. The CAN node that is responsible for the requested EID may have different EID-to-resolver mappings, chooses one of them, and forwards the map-request to that resolver. This resolver has an appropriate EID-to-RLOC mapping, tunnels the packet to the ETR, and sends a map-reply to the requesting resolver which forwards it to the requesting ITR.

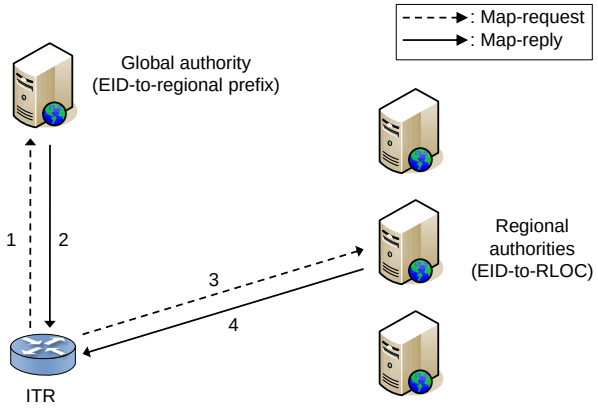
In contrast to FIRMS, DHT-MAP can support flat EID spaces. However, map-replies in DHT-MAP contain only RLOCs of the AS belonging to the responding resolver which is a strong limitation for multihoming. DHT-MAP does not have a comprehensive security concept, it has a more complex resilience concept, it leads to longer lookup delays and, therefore, to more lost, delayed, or relayed traffic.

4) ID/Locator Distributed Mapping Server: The mapping and relaying system presented in [44] is a TLMS similar to DHT-MAP but uses different top- and bottom-level mapping systems. The bottom-level mapping systems in ASes consist of DHTs instead of DMBs. Map-requests that cannot be served by the AS-local DHT are forwarded to a border server. Border servers of different ASes exchange with each other the EID-prefixes under their control via BGP and build a MRFO which serves as top-level mapping system instead of a CAN. Here, EID aggregation is important for scalability. Neighboring EIDs in different ASes due to mobility might be problematic.

5) Hierarchical Internet Mapping Architecture: HiiMap [45] also supports a flat EID space but assumes that EIDs are under the control of a region. Figure 8(b) shows that a single global authority (top-level mapping system) stores a "regional prefix" for each EID which points to a regional authority (bottom-level mapping system) that stores the EID-to-RLOC



(a) DHT-MAP: map-requests are served by AS-local resolvers (bottom-level mapping systems). If the resolver cannot find the requested EID in its own database, it forwards the map-request over the CAN (top-level mapping system) to an appropriate resolver in another AS which answers the query.



(b) HiiMap: first, the ITR queries the global authority if the required EID-to-regional-prefix mapping is not in its cache; then it queries the regional authority for the EID-to-RLOC mapping.

Fig. 8. Two-level mapping systems (TLMSs) supporting flat EID spaces.

mapping for all EIDs controlled by a region. An ITR queries the global authority for the "regional prefix" and after its reception the ITR queries the regional authority for the EID-to-RLOC mapping. HiiMap does not specify the implementation of the top- and bottom-level mapping system.

Compared to FIRMS, HiiMap's advantage is that it supports a flat EID space while FIRMS is applicable only for structured EIDs. However, HiiMap is less scalable than FIRMS since the global authority keeps mappings per EID. This leads to larger storage requirements and possibly cause a performance bottleneck for updates and map-requests. HiiMap requires two queries to obtain the EID-to-RLOC mapping and causes more lookup delay than FIRMS. It cannot relay packets when the ITR encounters a cache miss for the EID-to-regional-prefix mapping, and it has no concept for security and resilience.

6) LISP-TREE: LISP-TREE [46] is also a TLMS. It assumes that the EID address space is partitioned among regional EID registrars (RERs) which allocate parts of their EID space to local EID registrars (LERs). LERs further allocate EID

space to other LERs or customers. To be compliant with LISP, EID-to-RLOC mappings are stored by authoritative ETRs which serve as bottom-level mapping systems for EID-prefixes. Furthermore, LISP-TREE uses the MR/MS interface of LISP-MS to communicate with ITRs and authoritative ETRs in order to hide the top-level mapping system.

LISP-TREE uses a tree-like overlay structure of LISP-TREE servers (LTSs) to assist MRs to find the authoritative ETR for a given EID. The root LTSs are run by the RERs and store the information about their /8 prefixes. Lower level LTSs that control more specific EID-prefixes are run by the corresponding LERs. MRs are configured with the root LTSs and iteratively query LTSs to eventually find the authoritative ETR for a given EID before they query it for the actual EID-to-RLOC mapping. Intermediate results about LTS are cached so that the MR must query the root LTSs only rarely. LISP-TREE is based on existing DNS implementations which makes it very scalable and security is provided by the use of DNSSEC [47]. Nevertheless, we classify LISP-TREE as TLMS because it uses DNS technology only to resolve EID-to-authoritative-ETR mappings while DNSMS would use DNS technology to resolve EID-to-RLOC mappings. FIRMS shares with LISP-TREE the hierarchical assignment of EIDs and the TLMS principle. In contrast to FIRMS, LISP-TREE is not able to forward packets for which ITRs lack appropriate EID-to-RLOC mappings and requires on average more lookups than FIRMS to find authoritative ETRs (map bases).

D. DNS-Based Mapping Systems (DNSMS)

We review examples for DNSMS that have been proposed for mapping lookup by hosts and intermediate nodes and discuss these solutions.

1) Use of DNSMS for EID-to-RLOC Lookup in Hosts: The Identifier/Locator Network Protocol (ILNP) [3], [48] proposes to define a new resource record (RR) type that holds an RLOC for a DNS name (see Section II-B). However, this not applicable when intermediate nodes like ITRs query the mappings since they do not have the DNS name which is needed for the lookup.

In the HIP context, IPs are mapped to HITs so that a HIT-to-IP mapping service is needed. The authors of [49] propose the DNS system for that task. They postulate the "hit-to-ip.arpa" domain in which HITs are denoted like IPv6 addresses within "ipv6.arpa" for reverse DNS. Since HITs are not hierarchically structured, all HITs need to be known by top-level servers that are run by authorities. The authors give evidence that DNS servers are powerful enough for their purpose. Since improved mobility is an objective of HIP, HIT-to-IP mappings are likely to change often. As updates of DNS records take orders of magnitude longer than retrievals, a two-level hierarchy is introduced. The entries in the top-level DNS servers just refer to second-level DNS servers. These entries are likely to stay the same for long time. As a result, top-level servers experience fewer updates which reduces the infrastructure expenses for authorities. This also provides direct control over the actual HIT-to-IP mapping to the HIT owner which is important to support mobility.

2) Use of DNSMS for EID-to-RLOC Lookup in Intermediate Nodes: Reverse DNS (rDNS) performs a lookup of RRs based on given IP addresses. In combination with RLOC RRs, a service returning IP-to-RLOC mappings can be implemented. The prefix owner can set up an authoritative DNS server with the RLOC RRs for his EID prefix and register the address of this delegation server with the authority from which it has received his EID prefix. Thereby, the prefix owner has still control over the mappings. This idea has been sketched for the LISP context in [50] and in [51]. However, it did not prevail since the existing DNS infrastructure should not be burdened with another heavy service. Moreover, for many people this approach did not seem sufficiently robust and powerful to be applied as a mapping system in a Loc/ID split context where intermediate nodes query the mappings.

DNS has been proven to be a powerful and scalable architecture, but it has not been secure. Security has recently been added [47] and clients trust the received data when they are signed by the authoritative DNS server. However, if the client does not trust the public key of the authoritative DNS, it must first validate that key before it can validate the actual data. Thus, the client needs to iteratively validate the trust chain up to a common trust anchor. Since this can become a time-consuming action, we intentionally took a different security approach for FIRMS where tracking a trust chain is not needed to validate map-replies.

The DNS is not suitable for relaying packets without RLOCs. Iterative resolution of EID-to-RLOC mappings forces resolvers to store relayed packets until they have received and validated the EID-to-RLOC mappings from the authoritative DNS server. As an alternative, only recursive queries could be used so that packets without RLOCs are passed from one DNS server to another until they reach the authoritative DNS server. This server finally tunnels the packet to an appropriate ETR. However, this option establishes states in top-level DNS servers and loads them with packet forwarding which both raise performance concerns.

VII. Conclusion

New routing architectures implementing Loc/ID split have been proposed for the Internet. Most of them assume that a mapping system is queried for EID-to-RLOC mappings by an intermediate node at the border of an edge network. Therefore, DNS is not appropriate for that purpose. We have presented FIRMS, a fast two-level mapping system. It includes security and resilience features as well as a relay service for initial packets of a flow when intermediate nodes encounter a cache miss for the EID-to-RLOC mapping. We have implemented a proof-of-concept for FIRMS in the G-Lab experimental facility and it showed its operation. Our performance analysis showed that FIRMS scales significantly better than centralized mapping systems with respect to storage requirements and update rates. We proposed four categories of mapping systems and used them to provide a comprehensive review. FIRMS has

structures in common with many other mapping system, but clearly differs in its overall design and stands out in the sum of the achieved benefits.

ACKNOWLEDGEMENTS

The authors thank Phuoc Tran-Gia for the support by G-Lab, David Hock, Dominik Klein, Steve Uhlig, Lixia Zhang, Luigi Iannone, Christian Vogt, Scott Brim, Tony Li, Benno Overeinder, Roland Bless, Erik Nordmark, Jeffrey Ahrenholz, Wolfgang Mühlbauer, Anja Feldmann, Oliver Hanka, Christopher Spleiss, Tim Neubert, Oleg Ponomarev, Marcelo Bagnulo, Robin Whittle, Eliot Lear, David Oran, Vince Fuller, David Meyer, and Brian Carpenter for valuable input and stimulating discussions.

REFERENCES

[1] D. Meyer, L. Zhang, and K. Fall, "RFC4984: Report from the IAB Workshop on Routing and Addressing," Sep. 2007.

[2] B. Quoitin, L. Iannone, C. de Launois, and O. Bonaventure, "Evaluating the Benefits of the Locator/Identifier Separation," in *ACM MobiArch*, Kyoto, Japan, Aug. 2007.

[3] R. Atkinson, S. Bhatti, and S. Hailes, "ILNP - Identifier/Locator Network Protocol," http://ilnp.cs.st-andrews.ac.uk/, 2009.

[4] ——, "ILNP: Mobility, Multi-Homing, Localised Addressing and Security through Naming," *Telecommunication Systems*, vol. 42, 2009.

[5] A. Feldmann, L. Cittadini, W. Mühlbauer, R. Bush, and O. Maennel, "HAIR: Hierarchical Architecture for Internet Routing," in *ACM ReArch*, Rome, Italy, Dec. 2009.

[6] P. Frejborg, "Hierarchical IPv4 Framework," http://tools.ietf.org/id/draft-frejborg-hipv4-03.txt, Oct. 2009.

[7] R. Moskowitz and P. Nikander, "RFC4423: Host Identity Protocol (HIP) Architecture," May 2006.

[8] R. Moskowitz, P. Nikander, P. Jokela, and T. Henderson, "RFC5201: Host Identity Protocol," 2008.

[9] S. Jiang, "Hierarchical Host Identity Tag Architecture," http://tools.ietf.org/html/draft-jiang-hiprg-hhit-arch-02, May 2009.

[10] D. Farinacci, V. Fuller, D. Meyer, and D. Lewis, "Locator/ID Separation Protocol (LISP)," http://tools.ietf.org/html/draft-ietf-lisp, Jul. 2009.

[11] D. Meyer, "The Locator/Identifier Separation Protocol (LISP)," http://www.1-4-5.net/ dmm/lisp/, 2008.

[12] ——, "The Locator Identifier Separation Protocol (LISP)," *The Internet Protocol Journal*, vol. 11, no. 1, pp. 23–36, Mar. 2008.

[13] D. Lewis, D. Meyer, D. Farinacci, and V. Fuller, "Interworking LISP with IPv4 and IPv6," http://www.ietf.org/internet-drafts/draft-ietf-lisp-interworking-00.txt, May 2009.

[14] C. Vogt, "Six/One Router: A Scalable and Backwards Compatible Solution for Provider-Independent Addressing," in *ACM MobiArch*, Seattle, WA, USA, Aug. 2008.

[15] M. Menth, M. Hartmann, and D. Klein, "Global Locator, Local Locator, and Identifier Split (GLI-Split)," in *currently under submission*, 2009.

[16] D. Jen, M. Meisel, D. Massey, L. Wang, B. Zhang, and L. Zhang, "APT: A Practical Tunneling Architecture for Routing Scalability," UCLA CS Department, Tech. Rep. 080004, Mar. 2008.

[17] O. Hanka, C. Spleiss, G. Kunzmann, and J. Eberspächer, "A Novel DHT-Based Network Architecture for the Next Generation Internet," in *International Conference on Networking (ICN)*, Cancun, Mexico, Mar. 2009.

[18] S. Schuetz, R. Winter, L. Burness, P. Eardley, and B. Ahlgren, "Node Identity Internetworking Architecture," http://tools.ietf.org/id/draft-schuetz-nid-arch-00.txt, Sep. 2007.

[19] X. Xu, "Routing Architecture for the Next Generation Internet (RANGI)," http://tools.ietf.org/id/draft-xu-rangi-02.txt, Jul. 2009.

[20] L. Iannone and O. Bonaventure, "On the Cost of Caching Locator/ID Mappings," in *ACM CoNext*, Dec. 2007.

[21] S. Brim, N. Chiappa, D. Farinacci, V. Fuller, and D. Lewis, "LISP-CONS: A Content distribution Overlay Network Service for LISP," http://tools.ietf.org/html/draft-meyer-lisp-cons, Apr. 2008.

[22] S. Brim, D. Farinacci, D. Meyer, and J. Curran, "EID Mappings Multicast Across Cooperating Systems for LISP," http://tools.ietf.org/html/draft-curran-lisp-emacs-00, Nov. 2007.

[23] G. Chen et al., "An Incremental Deployable Mapping Service for Scalable Routing Architecture," http://tools.ietf.org/html/draft-chen-lisper-mo-01.txt, Jul. 2009.

[24] L. Iannone, D. Saucez, and O. Bonaventure, "LISP Mapping Versioning," http://tools.ietf.org/html/draft-iannone-lisp-mapping-versioning-00, Mar. 2009.

[25] C. Lynn, S. Kent, and K. Seo, "RFC3779: X.509 Extensions for IP Addresses and AS Identifiers," Jun. 2004.

[26] G. Huston and G. Michaelson, "Resource Certification - A Public Key Infrastructure for IP Addresses and AS's," Asia Pacific Network Information Centre (APNIC), Draft, Nov. 2008.

[27] T. Dierks and E. Rescorla, "RFC5246: The Transport Layer Security (TLS) Protocol Version 1.2," Aug. 2008.

[28] A. Akella, B. Maggs, S. Seshan, A. Shaikh, and R. Sitaraman, "A Measurement-Based Analysis of Multihoming," in *ACM SIGCOMM*, Karlsruhe, Germany, Aug. 2003.

[29] RIPE NCC, "RIS Statistics Report," http://www.ris.ripe.net/weekly-report/, 2009.

[30] Internet System Consortium, "The ISC Domain Survey," https://isc.org/solutions/survey, 2009.

[31] P. Martin, "Zen Internet UK Small Medium Enterprise (SME) survey," Shape the Future Limited, Tech. Rep., Nov. 2008.

[32] Verisign, "The Domain Name Industry Brief," Jun. 2009.

[33] DENIC, "Der Nameserverdienst der DENIC," http://www.denic.de/hintergrund/nameservice.html, Jun. 2009.

[34] D. Farinacci, V. Fuller, D. Meyer, and D. Lewis, "LISP Alternative Topology (LISP+ALT)," http://tools.ietf.org/html/draft-ietf-lisp-alt, May 2009.

[35] E. Lear, "NERD: A Not-so-novel EID to RLOC Database," http://tools.ietf.org/html/draft-lear-lisp-nerd, Apr. 2008.

[36] D. Jen, M. Meisel, D. Massey, L. Wang, B. Zhang, and L. Zhang, "APT: A Practical Transit Mapping Service," http://tools.ietf.org/id/draft-jen-apt-01.txt, Nov. 2007.

[37] R. Whittle, "IVIP - A New Routing and Addressing Architecture for the Internet," www.firstpr.com.au/ip/ivip/, 2008.

[38] ——, "Ivip Mapping Database Fast Push," http://tools.ietf.org/id/draft-whittle-ivip-db-fast-push, Aug. 2008.

[39] D. Farinacci and V. Fuller, "LISP Map Server," http://tools.ietf.org/html/draft-ietf-lisp-ms, Oct. 2009.

[40] L. Mathy and L. Iannone, "LISP-DHT: Towards a DHT to Map Identifiers onto Locators," in *ACM ReArch*, Madrid, Spain, Dec. 2008.

[41] V. Ramasubramanian and E. G. Sirer, "The Design and Implementation of a Next Generation Name Service for the Internet," in *ACM SIGCOMM*, Portland, OR, USA, 2004.

[42] J. Ahrenholz, "HIP DHT Interface," http://tools.ietf.org/html/draft-ahrenholz-hiprg-dht-04, Mar. 2009.

[43] H. Luo, Y. Qin, and H. Zhang, "A DHT-Based Identifier-to-Locator Mapping Scheme for a Scalable Internet," *IEEE Transactions on Parallel and Distributed Systems*, vol. 20, no. 10, Oct. 2009.

[44] F. Hu and J. Luo, "ID/Locator Distributed Mapping Server," http://tools.ietf.org/id/draft-hu-lisp-dht-00.txt, Oct. 2009.

[45] O. Hanka, G. Kunzmann, C. Spleiß, J. Eberspächer, and A. Bauer, "HiiMap: Hierarchical Internet Mapping Architecture," in *International Conference on Future Information Networks*, Beijing, China, Oct. 2009.

[46] L. Jakab, A. Cabellos-Aparicio, F. Coras, D. Saucez, and O. Bonaventure, "LISP-TREE: A DNS Hierarchy to Support the LISP Mapping System," *accepted for IEEE Journal on Selected Areas in Communications*, 2010.

[47] R. Arends, R. Austein, M. Larson, D. Massey, and S. Rose, "RFC4080: DNS Security Introduction and Requirements," Mar. 2005.

[48] R. Atkinson and S. Bhatti, "An Introduction to the Identifier-Locator Network Protocol (ILNP)," in *London Communications Symposium (LCS)*, London, UK, Jul. 2006.

[49] O. Ponomarev and A. Gurtov, "Embedding Host Identity Tags Data in DNS," http://tools.ietf.org/id/draft-ponomarev-hip-hit2ip, Mar. 2009.

[50] D. Farinacci, D. Oran, V. Fuller, and J. Schiller, "Locator/ID Separation Protocol (LISP2) [DNS-based Version]," http://www.dinof.net/d̃ino/ietf/lisp2.ppt, Nov. 2006.

[51] C. Vogt, "DNS Map – A DNS-Based Resolution System for IP Address Mappings," Technical Report, Feb. 2008.