

Internet Research Task Force
Internet-Draft
Intended status: Informational
Expires: January 13, 2009

J. Halpern, Ed.
July 12, 2008

A Taxonomy for New Routing and Addressing Architecture Designs
draft-halpern-rrg-taxonomy-01.txt

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with [Section 6 of BCP 79](#).

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

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

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

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

This Internet-Draft will expire on January 13, 2009.

Abstract

The Routing Research Group is tasked to design a new routing architecture to meet the challenges of scalability in face of pervasive multi-homing and inter-domain traffic engineering. A number of solutions have been proposed. This draft describes a taxonomy for the design space.

Internet-Draft

Design Taxonomy

July 2008

Table of Contents

1.	Introduction	3
2.	Terminology	3
3.	New Math or New Naming	5
4.	Divide and Conquer	6
4.1.	What Mappings?	7
4.2.	How Mapping? or Map Distribution?	7
5.	Tunneling, Rewriting, or Separate Identification	8
5.1.	Tunneling	9
5.2.	Rewriting	10
5.3.	Separate Identification	10
6.	Scoping	12
7.	Version Requirements	12
8.	Mobility	13
9.	Acknowledgments	14
10.	References	14
10.1.	Normative References	14
10.2.	Informative References	14
	Author's Address	14
	Intellectual Property and Copyright Statements	15

1. Introduction

The Routing Research Group is tasked to recommend a new routing architecture to meet the challenges of scalability, multi-homing, and inter-domain traffic engineering. With the approaching exhaustion of IPv4 address space, the discussion and some initial deployment of IPv6 has moved from the back burner to the front stage. However, one of the major issues concerning IPv6 deployment is its potential impact on scalability of the already stressed routing system.

A number of approaches to scaling the Internet's routing system have been submitted. We expect this taxonomy to facilitate discussion of both existing and future proposals, to position each proposal in the design space, and to help evaluation of various design trade-offs in all the proposals. In addition, in order to facilitate both this analysis and research group discussion, this document proposes a set of terminology and meanings for those terms. While this document attempts to avoid redefining existing terms, the discussion space is very crowded, and many terms are already quite overloaded with meanings.

It would be desirable to be able to decompose the solution space being explored into a set of orthogonal dimensions, each of which could be further decomposed. And that is the structure this document attempts to overlay on the space. However, the dimensions are not actually orthogonal, and the choices are not independent. For example, the question of where certain kinds of lookups are done is not completely independent of the question of what kinds of lookups are needed.

Following this introduction, the document has a section on terminology, providing at least the terms as they are used here, and one hopes also providing terms for more general discussion. Following that are a series of sections discussing separate dimensions along which to understand potential solutions to the problems under discussion. These are not evaluation criteria, but

rather ways in which solutions may differ. The main portion of the document ends with a discussion of open issues.

2. Terminology

Core Aggregatable Address -- An address that can be used by routing and which can be included in a collective (i.e. aggregated) advertisement when used in Internet core routing. For this purpose, core is an approximate term which can be thought of as the current default free zone in BGP. Core Aggregatable Addresses are currently globally unique, although some proposals remove that property. (Some

Halpern

Expires January 13, 2009

[Page 3]

Internet-Draft

Design Taxonomy

July 2008

proposals treat the core as just another region. There are also proposals where the core does not need to be aggregatable. Better terminology is still sought.)

Scoped Address -- An address that can be used by routing within some meaningful scope. For most purposes, this scope is distinct from the core of the global Internet, although as mentioned under Core Aggregatable Address, some proposal treat the core as just another scope. It is of course the case that a locally scoped address, even one with a very small routability scope, may be globally unique.

[Editorial Digression: It is understood that even terms such as the above two are actually quite fuzzy. For example, in a hypothetical environment where IPv6 is widely deployed, a globally unique ISP allocated IPv6 address is a Core Aggregatable address. If that same IPv6 addresses are used only for routing within a site, then they are Scoped Addresses. And in a more likely environment where for a long time IPv6 routing makes extensive use of tunnels, but has service providers that route IPv6 addresses and behave like an Internet Core, it is not clear which term would be appropriate to describe an IPv6 address used to direct delivery of a packet.]

Communicating Entity Identifier -- A bit string used to identify an entity participating in the Internet. In the traditional IPv4 and IPv6 Internet, the IP address is used as both the address for the packet forwarding system, and the Communicating Entity Identifier. There is an additional important distinction in the intended use of this term, as compared with current IP addressing. IP addresses name interfaces. When communicating with a host, to the degree that the

address is used to identify the communicating entity, it also constrains the physical interface over which that communication takes places. Thus, it is frequently said that an IP address names an interface. Communicating Entity Identifier names the communicating entity, not just a specific interface of the entity.

Pure Entity Identifier -- A string (typically, but not always binary) used to identify an entity participating in the Internet in a way that is completely insensitive to the connectivity of that entity to the Internet. Note that some routing systems may use such strings for packet forwarding. The determination of whether it is a Pure Entity Identifier is not related to how other systems track or map the identifier. This term is introduced largely to distinguish from a Scoped Identifier. It is also introduced to avoid the free standing term Identifier since, like Address, the term has been used by different folks to mean different things.

Scoped Entity Identifier -- A string (typically but not always binary) used to identify an entity participating in the Internet. A

Scoped Identifier is assigned by a connected region of the Internet, and typically must change if the entity leaves that scope. A scoped identifier may or may not change when the entity moves or the Internet connectivity changes within the scope that has assigned the identifier.

It should be noted that whether an identifier is Scoped or Pure is a property of the assignment / definition process of the identifier as defined by the system. So a Pure Entity Identifier may be used by routing / packet forwarding in some scope as an address. It may well be in fact a scoped address for forwarding purposes. But if the definition of the bit string is such that it does not change when the entity is moved to a different scope on the Internet, then the identifier is a Pure Entity Identifier. Similarly, if the system defines a Communicating Entity Identifier as being assigned by the scope, then it is a Scoped Entity Identifier even if the system permits (and some scopes choose) the use of globally unique arbitrary bits strings as identifiers, since when the Entity moves it has to get an identifier appropriate to the scope it moves into.

Routing Locator -- the bit string used by the routing system in the Internet core to deliver a packet across that core. In most of the

proposals under discussion, this is a Core Aggregatable Address. In some approaches that have been suggested in the past, this may not be aggregatable, or may be a flow identifier, or even something more exotic.

[3.](#) New Math or New Naming

One question that can be asked about a proposal is how much change does it make to the basics of routing as it is practiced today. The choices conceptually range from leaving the system alone through choices such as Nimrod, or alternatively through geographic or AS based ideas, and then to ideas that take a very different mathematical approach to the routing and forwarding system.

Most of the proposals under discussion in the Routing Research Group take as a premise that the ISP based routing system can (may not have to, but is permitted to) remain operating the way it is today. This can be viewed as being based on a conclusion that the current approach is good enough, or on a view that deployability requires leaving some parts of the system fixed, or likely other motivations. These proposals generally rely on splitting the packet destination naming problem (and, of course, source naming) into two parts. One part, often called the identifier space, is used to anchor the communication session. The other part is some form of address that is used to actually deliver the packets. By stretching the notion of

mapping (to allow for a wide range of places that mapping may occur) we can refer to all of these solutions as mapping based solutions.

There are some proposals which attempt to address the dynamics and aggregatability of the routing system by basing it on values, carried in packets, which have somewhat better behaviors than current IP addresses. Such ideas include geographic based addressing and AS based addressing.

There are also theoretical approaches such as compact routing [ed note, ROLF?] which take a very different approach to routing, addressing, and packet forwarding. In theory, these approaches could have highly scalable table sizes while allowing arbitrary bit strings as the names used in packets for destinations.

[4.](#) Divide and Conquer

As was discussed above, most of the solution under discussion attempt to address the systemic routing problems by a divide and conquer approach focused on splitting communication entity identification from the bit strings that are used to forward packets. These forwarding bit strings are called Routing Locators, or RLOCs, in one of the proposals on the table, and that seems a clear term for that function.

Looking at the string used for identifying communicating entities, there seem to be four sorts of approaches to this.

- o Identification by name;
- o Identification by globally unique location insensitive bit string;
- o Identification by globally unique location sensitive bit string;
- o Identification by purely local identifiers;

Of these, the first would be exemplified by using a DNS name as an identifier. Such approaches tend to avoid shipping the identifier in most packets. For example, the DNS name or a string derived from it might be shipped in the first application / transport packet, to create end-to-end state, and only shipped thereafter in packets that modify that state.

The second category of splitting is exemplified by solutions such as GSE, where entities are identified by globally unique bit strings. These IDs are what the terminology section of this document calls Pure Entity Identifiers.

The third category is similar to the second. It differs in that the bit string used for communicating entity identification is assigned by a connected region of the Internet (typically a site.) This simplifies the process of assigning identifiers within the region, since the regional authority can perform the assignment. To some degree, it also simplifies the potential optimization of using the communication identifier for local packet delivery.

The fourth category basically treats the Internet as a collection of regions, with completely different naming in each region. An example of such an approach would be a system which required a description of the path from the source communication domain to the destination communication domain in order to establish communication. The author does not believe any such approaches are under discussion in the research group, but it is included here to try to help complete the taxonomy.

[4.1.](#) What Mappings?

In order to utilize these various forms of splitting of the problem, it is necessary for some devices to be able to determine the correct bit strings to use. There are several mappings which may apply. Not all of these mappings are needed by all of the solutions:

- o Mapping from the application handle (e.g. DNS name) to a communication entity identifier;
- o Mapping from the application handle (e.g. DNS name) to a routing locator;
- o Mapping from a communication entity identifier to an RLOC;
- o Mapping from an RLOC to a communicating entity identifier;

In fact, some of these mappings are not merely unnecessary, but are meaningless for some of the solutions under discussion.

[4.2.](#) How Mapping? or Map Distribution?

In the above description, we focused on the placement of the function to update the information in a packet. A closely related question is how enough information is made available to perform this operation. As discussed below, some solution approaches are designed to avoid needing this operation since it can be a source of complexity. There are two basic approaches to such distribution, and several hybrids.

One approach which has significant simplicity is a pure full push solution. Simply distribute the entire table of mappings to all the

places that could possibly need it. This means that whenever

information is needed, it is available. Conversely, this may need to distribute a lot of information to a lot of places, which introduces scale questions that must be resolved if this kind of approach is to be adopted.

The pure alternative to that approach is have each piece of information stored in a distributed database, and anyone who needs a mapping consults the database to get the needed information. DNS is a classic example of such a database. LISP-CONS proposes such a database, with information distribution for the purpose of steering the information extraction. This family of solutions are often called pull based solutions, as devices only get the information when they need it. Even the purest pull solution actually makes use of caching to reduce the need to request the same information repeatedly. One question with pull based approaches is what latency penalty is incurred to allow the pull to occur. Clearly, when a lookup is needed, traversing such a system takes some noticeable amount of time. Conversely, this lookup is only needed when the first packet for a destination identifier is being handled by a mapper. Subsequent packets, within a reasonable time period, even from distinct sources, will get the benefit of the caching to get effective mapping. Making it more likely that one will get this sort of cache hit is one of the drivers for using scoped identifiers. If someone needs to communicate with a host in a site, it is frequently likely that communication will occur with other hosts in the same site.

It is also practical to use a range of hybrid solutions. Approaches like APT and Ivip use a push based solution to deliver the full information to a subset of all devices, such that every device that needs to perform the mapping has and knows of a nearby device that has the information. This kind of hybrid reduces the scale of the information distribution, while keeping the latency for the mapping function significantly smaller than a pure pull would be likely to need. The question of how much gain this provides depends upon the likelihood of cache hits and misses in the actual edge device, as discussed above.

5. Tunneling, Rewriting, or Separate Identification

In order to ensure that the routing system scales and stabilizes better, without utilizing new mathematical approaches to that system, the solutions under discussion all try to ensure that the address as seen by the routing system in the core of the network is a core aggregatable address. This enables the routing system to utilize the aggregation properties it was designed for in an effective manner.

There are a range of ways that this is achieved in different proposals.

[5.1.](#) Tunneling

One common approach is tunneling. This involves taking the existing packet with the existing information (which appears as addressing information to the packet forwarding system) and modifying the packet by adding new destination (and usually source) addressing information, while preserving the original information.

One aspect of tunnels and tunneling or encapsulation is the question of whether the two ends of the tunnels have shared state. In many VPN solutions that use tunnels, the two ends of explicit shared state to enable cryptographic protections of various kinds. On the other hand, in some tunneling mechanisms used with IPv6 overlays, there is no need for such shared state. Either there is no need for shared information, or anything needed is carried in the packet. It has been suggested that if a tunnel is not using shared state than it is just encapsulation, and should be called encapsulation. The tunneling solutions under consideration for routing scaling all avoid pairwise shared state for the tunnels, as that helps many aspects of the solution. It is for further discussion whether these approaches should be referred to as encapsulation-based approaches rather than tunneling approaches. In this document, the two terms are used interchangeably, according to which term fits the specific context better.

The most common mechanism for this is to add a new IP header, and possibly an intermediate informational header. The intermediate header allows for additional information which can affect the far end behavior. The use of an encapsulating IP header also allows for a different outer header version and inner header version, increasing the versatility of this approach.

Other alternative tunneling mechanisms have been suggested, such as using a new IP option field to carry the old header information. There are three properties that appear relevant for the taxonomy that distinguish tunneling from other approaches.

- o The technique is applied to all data packets. Additional techniques may be used, but this strategy centers on using tunneling for almost all data packets. [Editors note: If tunneling is done such that intra-site packets are not tunneled, we still consider this to apply. Possibly the description should say "all inter-site packets?" But that seems too specific.]

- o The technique makes the data packets larger
- o The technique preserves the original addressing information. This is generally used so that the communicating entities see the same addresses or identifiers. Some tunneling systems introduce an exception to this property for backwards compatibility.

Tunneling generally requires that the device performing the encapsulation be able to perform the mapping from the identifier to the Core Aggregatable Address to be used in the outer, encapsulating, header. Tunneling can be used in conjunction with any of the described placements of the mapping logic, including in the host, as long as the mapping device is performing the encapsulation. Tunneling can be used with any version of IP, or even mixed versions.

[5.2.](#) Rewriting

Tunneling is not the only strategy to ensure that the packet carries Core Aggregatable Addressing information when it is used in the Internet core. Another strategy which still relies on mapping between identifiers and Core Aggregatable Addresses is to rewrite the addressing in place.

At its simplest, this would seem to be a description of NAT. While NAT tends to reverse this (rewriting source address on entry to the Internet core and rewriting destination address on exit from the Internet core) it does match that description. And NAT with its many problems is not going to solve the issues at hand.

However, if the identifying information can be preserved through the NAT, without encapsulation, something effective can be achieved. An example of this is the use of IPv6, where some or all of the upper 8 bytes of the IPv6 address field are rewritten, based on identification information carried in the lower 8 bytes. This has the property that the size of the packet is not affected by the process.

Like Tunneling, rewriting approaches can be used with any placement of the mapping function. When used with mapping in the host, it may

be difficult to distinguish between this sort of approach and a Separate Identification approach, except in terms of the semantic description. To date, this sort of rewriting is only envisioned for use with IPv6.

[5.3.](#) Separate Identification

For systems designed to be deployed in hosts, there is another class of approach. This approach separates out the mapping, and sometimes

Halpern

Expires January 13, 2009

[Page 10]

Internet-Draft

Design Taxonomy

July 2008

even eliminates it entirely. The first distinguishing property of these approaches is that most of the data packets do not carry identifiers on the wire, even in their originating or destination scopes. Instead, the packets carry suitable Core Aggregatable Addresses.

The oversimplified form of this solution family would be to simply declare, as IPv6 initially attempted, that all IP addresses used in packets shall be Core Aggregatable. And to stop there as if that solved the problem. That is clearly insufficient.

Other approaches rely on establishing paired state in the communicating entities so as to enable multiple Core Aggregatable Addresses to be used on individual data packets. Suitable updating packets are used to allow for changes in the set, and provide suitable security. Some of these solutions use an explicit identifier, while others use the first Core Aggregatable Address used as the hook for anchoring a given communication. It has also been suggested to use the original DNS name as the identifier, carrying a reference to it in certain packets where necessary. This does assume that all hosts get DNS names. Hybrid approaches where packets carry extra connection information, but not necessarily full identification in additional headers, driven from the host, make the boundary between these approaches and tunneling more difficult to determine. One complication with these approaches is ensuring that they can work with TCP, SCTP, UDP, and DCCP, as it is not the routing systems job to determine what transport protocol the applications utilize.

One advantage of this sort of Separate Identification is that by leveraging extant host information gathering (DNS lookups, for example) or by using existing information as the communications anchor, the system can avoid the need for custom distribution

techniques to handle the mapping information.

To avoid confusion, there is one other kind of separate identification that should be mentioned, as it is intended that the above candidates NOT be so broad as to include solutions that require state establishment in the network. For example, a solution which established MPLS LSPs from each source to each destination host would clearly be establishing network state. Even solutions which required the communication initiator to establish and maintain state in remote edge devices should probably be considered part of some other space of active state maintenance solutions. Requirements to ensure state in ones own border devices, in order to meet communications control requirements, may well fit within the category of Separate Identification.

Typically, Separate Identification techniques expect (some may not

require) a degree of visibility into the routing connectivity. This may be provided purely be a set of prefix announcements. It may be augmented by observations and probes of traffic behavior. It may also make use of suggested protocols to allow the routing system to provide state and policy hints to the hosts to assist the host processing.

6. Scoping

One of the aspects that makes discussion of these solutions complicated is the uses of scoping. There are two separate but related notions.

Address Scoping is the scoping of address information as used for packet forwarding (i.e. traditional routing.) Routing has always tried to keep local information about reachability local, so as to ensure aggregatability. Success in this endeavor has varied. The approaches discussed here aim to ensure that addresses used in the global or Core Internet scope are highly aggregatable. In conjunction with that, many of the solutions discussed here use different addresses for delivery of packets within a site. These addresses are usually globally unique, but are not usable for packet forwarding outside of a site. Frequently, this involves using the Communicating Entity Identifier as the address for local delivery.

While one could use other addresses, such solutions would likely incur additional mapping for little additional value and have not been explored to date.

Separately, there is the question of the scope of the identifiers used for the communicating entities. The scope of a Communicating Entity Identifier is the scope (or range of places in or attached to the Internet) where the entity can use that identifier. Some solutions, such as GSE, HIP, or Ivip, use identifiers that can be used anywhere in the Internet. The identifier does not change, even if the entity moves. Other solutions use identifiers which are tied to an administrative or routing scope. This allows the routing system to somewhat more easily identify its local entities, and to forward packets using the identifier to those local entities. It also tends to mean that entities which share core Internet reachability will be in an aggregatable block of identifiers, potentially making caching of mapping information more effective.

[7.](#) Version Requirements

Different approaches under consideration make different assumptions about what versions of the underlying communications protocols are

in use by hosts, sites, and the Internet Core. This conceptually includes both the question of IPv4 versus IPv6 and the question of other kinds of modifications to host stacks or various routers. Some of this comes up in terms of the earlier discussion of where functions are placed. It is probably the case that in a pure architectural sense it would be better if these considerations were kept out of a taxonomy of the architectural work. But as the topics come up repeatedly, and seem to affect the view of solutions very strongly, it is mentioned here.

Some of the proposals under consideration are designed to apply to almost any underlying Internet Protocol. In particular, there is a large class of interesting proposals (mostly in the mapping and encapsulation family) which work equally well for IPv4 and IPv6.

Some of the solutions under discussion assume that at least some portions of the infrastructure are using IPv6. Those solutions make use of the larger address size from IPv6 to enable modes of operation

that can not be fit into an existing IPv4 address. Obviously, and such solution must deal with the reality that communication with and over the existing IPv4 network is an essential practical requirement.

As mentioned, while most solutions assume that the current version of host software can be used, some interesting proposals require some degree of host modification. For example, proposals which move the identification of communicating parties to or above the transport protocol obvious will require new versions of host software.

[8.](#) Mobility

One issue that comes up is whether the improvements to the routing system can or should help deal with mobile devices or mobile networks. There is an argument that given that mobility will become more important, and more widespread, it is important to address mobility in the core design of the routing system. Conversely, given that mobility occurs over a range of time a topology scales, with a range of needs, there is an argument that it should be addressed by a range of techniques (potentially including locally mobility management, hierarchical mobility management, and a variety of tunneling and VPN tools.)

It does seem that some of the solutions under discussion offer tools that can help the mobility situation. If globally scoped identifiers are used for communicating entities, those identifiers can serve as a reference to be used by mobility solutions. For mobile networks, potentially including mapping information aggregator in the set of things that move may allow more effective global reachability

information, depending upon the approaches. In general, to date, the routing research group has viewed benefits to mobility as a nice result, but not a driver in the architectural process.

[9.](#) Acknowledgments

This draft owes significant debt to the earlier draft done by Lixia Zhang and Scott Brim [[ZBTaxonomy](#)] that began to address this question.

[10.](#) References

[10.1.](#) Normative References

[10.2.](#) Informative References

[ZBTaxonomy]

Zhang, L. and S. Brim, "A Taxonomy for New Routing and Addressing Architecture Designs", work in progress, [draft-rrg-taxonomy-00.txt](#), March 2008.

Author's Address

Joel M. Halpern (editor)
P. O. Box 6049
Leesburg, VA 20178
US

Email: jmh@joelhalpern.com

This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in [BCP 78](#) and [BCP 79](#).

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.