

Network Working Group
Internet-Draft
Intended status: Informational
Expires: January 27, 2008

T. Narten
IBM
July 26, 2007

Routing and Addressing Problem Statement
draft-narten-radir-problem-statement-00.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/1id-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 27, 2008.

Copyright Notice

Copyright (C) The IETF Trust (2007).

Abstract

Problem statement for the route scaling problem.

Table of Contents

1.	Introduction	3
2.	Terms and Definitions	4
3.	Background	6
3.1.	Technical Aspects	6
3.2.	Business Aspects	7
3.3.	Alignment of Incentives	8
3.4.	Table Growth Targets	8
4.	Pressures on Routing Table Size	10
4.1.	Traffic Engineering	10
4.2.	Multihoming	11
4.3.	End Site Renumbering	12
4.4.	Acquisitions and Mergers	12
4.5.	RIR Address Allocation Policies	12
4.6.	Dual Stack Pressure on the Routing Table	13
4.7.	Internal Customer Routes	14
4.8.	IPv4 Address Exhaustion	14
5.	Pressures on Path Computation Load	15
5.1.	Interconnection Richness	15
5.2.	Multihoming	15
5.3.	Traffic Engineering	15
5.4.	Questionable Operational Practices?	16
5.4.1.	Rapid shuffling of prefixes	16
5.4.2.	Anti-Route Hijacking	16
5.4.3.	Operational Ignorance	16
5.5.	RIR Policy	17
6.	Problem Statement	18
7.	Security Considerations	19
8.	Acknowledgments	20
9.	Informative References	21
	Author's Address	22
	Intellectual Property and Copyright Statements	23

1. Introduction

Prompted in part by the recent IAB workshop on Routing & Addressing [RAWS-REPORT], there has been a renewed focus on the problem of routing scalability within the Internet. The issue itself is not new, with discussions dating back at least 10-15 years [GSE, ROAD,...].

This document attempts to define the "problem", with the aim of describing the essential aspects so that the community has a way of evaluating whether proposed solutions actually address or impact the underlying problem or "pain points" in a significant manner.

2. Terms and Definitions

Default Free Zone (DFZ): That part of the Internet where routers maintain full routing tables. Many routers maintain only partial routes, having explicit routes for "local" destinations plus a "default" for everything else. For such routers, building and maintaining routing tables is relatively simple because the amount of information learned and maintained can be small. In contrast, routers in the DFZ maintain complete information about all reachable destinations, which currently number in the hundreds of the thousands of entries.

Routing Information Base (RIB): The data structures a router maintains that hold the information about destinations (i.e., prefixes) and paths to those destinations. The amount of state information maintained is dependent on a number of factors, including the number of individual prefixes, the number of BGP peers, the number of distinct paths, etc. The RIB may also include information about unused ("backup") paths for a given prefix as well as the active path(s) used for forwarding. The RIB is typically constructed from specialized hardware components, which have different (and higher) cost properties than the hardware typically used to maintain the FIB.

Forwarding Information Base (FIB): The actual table used while making forwarding decisions for individual packets. The FIB is a compact, optimized subset of the RIB, containing only the information needed to actually forward individual packets, i.e., mapping a packet's destination address to an outgoing interface and next-hop. The FIB only stores information about paths actually used for forwarding; it typically does not store information about backup paths. The FIB is typically constructed from specialized hardware components, which have different cost properties compared to the hardware typically used to maintain the FIB.

Traffic Engineering (TE): In this document, "traffic engineering" refers to the current practice of inbound, inter-AS traffic engineering. This is accomplished by placing more specific routes in the routing table and/ or increasing the frequency of routing updates in order to control inbound traffic at the boundary of an Autonomous system (AS).

Provider Aggregatable (PA) address space Address space that an end site obtains from an ISP's address block. The main benefit of PA address space is that reachability to all of a provider's customers can be achieved by advertising a single "provider aggregate" address prefix into the DFZ, rather than needing to

Narten

Expires January 27, 2008

[Page 4]

announce individual prefixes for each customer. An important disadvantage is a requirement that the customer return those addresses (and renumber) when changing providers.

Provider Independent (PI) address space Address space that an end site obtains directly from a Regional Internet Registry (RIR) for numbering its devices. The main advantage (for the end site) is that it does not have to return those addresses (and renumber its site) upon changing providers. However, PI address blocks are not aggregatable and thus each individual PI assignment results in an individual prefix being injected into the DFZ.

3. Background

Within the DFZ, both the size of the RIB and FIB and the overall update rate have historically increased at a greater than linear rate. Specifically:

- o The number of individual prefixes that are propagated into the DFZ is increasing at a super-linear rate. The reasons behind this increase are varied and discussed below. Because each individual prefix requires resources to process, any increase in the number of prefixes requires a corresponding increase in resources. Each individual prefix that appears in routing updates requires state in the RIB (and possibly the FIB) and consumes processing resources when updates related to the prefix are received.
- o The overall rate of routing updates is increasing, requiring routers to process updates at an increased rate or converge more slowly if they cannot. The rate increase is driven by a number of factors (discussed below). It should be noted that the overall routing update rate is dependent on two factors: the number of individual prefixes and the mean per-prefix update rate. While it is clear that the overall number of prefixes is increasing super-linearly, further study is needed to determine whether the mean per-prefix update rate is increasing as well [need reference].

This super linear growth presents a scalability challenge for current and/or future routers. There are two aspects to the challenge. The first one is purely technical: can we build routers (i.e., hardware & software) actually capable handling the load, both today and going forward? The second challenge is one of economics: is the cost of developing, building and deploying such routers economically sustainable, given current and realistic business models that govern how ISPs operate as businesses?

Finally, the scalability challenge is aggravated by the lack of any limiting architectural upper-bound on the growth rate and a weakening of traditional social constraints on the growth rate that have helped restrain growth so far. Going forward, there is considerable uncertainty whether future growth rates will continue to be sufficiently constrained so that the routing load can be handled by routers available at that time.

3.1. Technical Aspects

The technical challenge of building routers relates to the resources needed to process a larger and increasingly dynamic amount of routing information. More specifically, routers must maintain an increasing amount of associated state information in the RIB, they must be

Narten

Expires January 27, 2008

[Page 6]

capable of populating a growing FIB, they must perform forwarding lookups at line rates (while accessing the FIB) and they must be able to initialize the RIB and FIB at boot time. Moreover, this activity must take place within acceptable time frames (i.e., the overall system must converge and stabilize within an acceptable time period). Finally, the hardware needed to achieve this cannot have unreasonable power consumption or cooling demands.

3.2. Business Aspects

Even if it is technically possible to build routers capable of meeting the technical and operational requirements, it is also necessary that the overall cost to build, maintain and deploy such equipment meet reasonable business expectations. ISPs, after all, are run as businesses. As such, they must be able to plan, develop and construct viable business plans that provide an acceptable return on investment (i.e., one acceptable to investors).

While the IETF does not (and cannot) concern itself with business models or the profitability of the ISP community, the cost of running the routing subsystem as a whole is directly influenced by the routing architecture of the Internet, which clearly is the IETF's business. Further, because cost implications are part of each and every engineering decision, controlling or limiting the overall cost of running the routing subsystem (through architectural decisions) is part of the IETF's fundamental charter. Consequently, having the IETF continue with an architectural model that places unbounded cost requirements on critical infrastructure represents an undue risk to the future of the Internet as a whole.

One aspect of planning concerns the assumptions made about the expected usable lifetime of purchased equipment. Businesses typically expect that once deployed, equipment can remain in use for some projected amount of time (e.g., 3-5 years). Upgrading equipment earlier than planned is more easily justified (as an unplanned expense) when a new business opportunity is enabled as a result of an upgrade. For example, an upgrade might be justified by an ability to support increased traffic or an increase in the number of customer connections, etc., where the upgrade can translate into increased revenue. In contrast, it is more difficult to justify unplanned upgrades in the absence of corresponding customer benefit (and revenue) to cover the upgrade cost. It is generally desired that deployed equipment remain usable over its planned lifetime. An increase in the resources required to support larger or more dynamic routing tables is viewed as a sort of "unfunded mandate", in that customers do not expect to have to pay more just to continue retaining the same level of service as before, i.e., having all destinations be reachable as was the case in the past. This

Narten

Expires January 27, 2008

[Page 7]

undermining of planning is particularly problematic when the increase in routing demand originates external to the ISP, and the ISP has no way to control or limit it (e.g., the increased demand comes from being part of the DFZ).

From a business perspective, it is desirable to maintain or increase the useful lifespan of routing equipment, by improving the scaling properties of the routing and addressing system.

3.3. Alignment of Incentives

Today's growth pattern is influenced by the scaling properties of the current system. If the system had better scaling properties, we would be able support and enable more widespread usage of certain applications such as multihoming and traffic engineering. Currently the system does not allow everyone to multihome, as there are some barriers to multihoming due to operational practices that try to strike a balance between the amount of multihoming and preservation of routing slots. It is desirable that the routing and addressing system exert the least possible back pressure on end user applications and deployment scenarios, to enable the broadest possible use of the Internet.

One aspect of the current architecture is a misalignment of cost and benefit. Injecting individual prefixes into the DFZ creates a small amount of "pain" for those routers that are part of the DFZ. Each individual prefix has a small cost, but the aggregate sum of all prefixes is significant, and leads to the core problem at hand. Those that inject prefixes into the DFZ do not generally pay the cost associated with the individual prefix -- it is carried by the routers in the DFZ. But the originator of the prefix receives the benefit. Hence, there is misalignment of incentives between those receiving the benefit and those bearing the cost of providing the benefit. Consequently, incentives are not aligned properly to produce a natural balance between the cost and benefit of maintaining routing tables.

3.4. Table Growth Targets

A precise target for the rate of table size or routing update increase that should reasonably be supported going forward is difficult to state in quantitative terms. One target might simply be to keep the growth at a stable, but manageable growth rate so that the increased router functionality can roughly be covered by improvements in technology (e.g., increased processor speeds, reductions in component costs, etc.).

However, it is highly desirable to significantly bring down (or even

reverse) the growth rate in order to meet user expectations for specific services. As discussed below, there are numerous pressures to deaggregate routes. These pressures come from users seeking specific, tangible service improvements that provide "business-critical" value. Today, some of those services simply cannot be supported to the degree that future demand can be reasonably be expected because of the negative implications on DFZ table growth. Hence, valuable services are available to some, but not all potential customers. As the need for such services becomes increasingly important, it will be difficult to deny such services to large numbers of users, especially when some "lucky" sites are able to use the service and others are not.

4. Pressures on Routing Table Size

There are a number of factors behind the increase in the quantity of prefixes appearing in the DFZ. From a theoretical perspective, the number of prefixes in the DFZ can be minimized through aggressive aggregation [[RFC4632](#)]. In practice, strict adherence to the CIDR principles is difficult.

4.1. Traffic Engineering

Traffic engineering (TE) is the act of arranging for certain Internet traffic to use or avoid certain network paths (that is, TE attempts to place traffic where capacity exists, or where some set of parameters of the path is more favorable to the traffic being placed there).

Outbound TE is typically accomplished by using internal IGP metrics to choose the shortest exit for two equally good BGP paths. Adjustment of IGP metrics controls how much traffic flows over different internal paths to specific exit points for two equally good BGP paths. Additional traffic can be moved by applying some policy to depreference or filter certain routes from specific BGP peers. Because outbound TE is achieved via a site's own IGP, outbound TE does not impact routing outside of a site.

Inbound TE is performed by announcing a more specific route along the preferred path that "catches" the desired traffic and channels it away from the path it would take otherwise (i.e., via a larger aggregate). At the BGP level, if the address range requiring TE is a portion of a larger address aggregate, network operators implementing TE are forced to de-aggregate otherwise aggregatable prefixes in order to steer the traffic of the particular address range to specific paths.

TE is performed by both ISPs and customer networks, for three primary reasons:

- o First, to match traffic with network capacity, or to spread the traffic load across multiple links (frequently referred to as "load balancing").
- o Second, to reduce costs by shifting traffic to lower cost paths or by balancing the incoming and outgoing traffic volume to maintain appropriate peering relations.
- o Finally, TE is sometimes deployed to enforce certain forms of policy (e.g., government traffic may not be permitted to transit through other countries).

TE impacts route scaling in two ways. First, inbound TE can result in additional prefixes being advertised into the DFZ. Second, Network operators usually achieve traffic engineering by "tweaking" the processing of routing protocols to achieve desired results, e.g., by sending updates at an increased rate. In addition, some devices attempt to automatically find better paths and then advertise those preferences through BGP, though the extent to which such tools are in use and contributing to the routing load is unknown.

In today's highly competitive environment, providers require TE to maintain good performance and low cost in their networks.

4.2. Multihoming

Multihoming refers generically to the case in which a site is served by more than one ISP [[RFC4116](#)]. Multihoming is used to provide backup paths (i.e., to remove single points of failure), to achieve load-sharing, and to achieve policy or performance objectives (e.g., to use lower latency or higher bandwidth paths). Multihoming may also be a requirement due to contract or law.

Multihoming can be accomplished using either PI or PA address space. A multihomed site advertises its site prefix into the routing system of each of its providers. For PI space, the site's PI space is used, and the prefix is propagated throughout the DFZ. For PA space, the PA site prefix may (or may not) be propagated throughout the DFZ, with the details depending on what type of multihoming is sought.

If the site uses PA space, the PA site prefix allocated from one of its provider's (whom we'll call the Primary Provider) is used. The PA site prefix will be aggregatable by the Primary Provider but not the others. To achieve the same level of multihoming as described in the case with PI addresses above, the PA site prefix will need to be injected into the routing system of all of its ISPs, and throughout the DFZ. In addition, because of the longest-match forwarding rule, the Primary Provider must also advertise and propagate the individual PA site prefix; otherwise, the path via the primary provider (as advertised via the aggregate) will never be selected due to the longest match rule. For the type of multihoming described here, where the PA site prefix is propagated throughout the DFZ, the use of PI vs. PA space has no impact on the load placed on the routing system. The increased load is due entirely to the need to propagate the site's individual prefix into the DFZ.

The demand for multihoming is increasing [XXX do we have data to cite?]. The increase in multihoming demand is due to the increased reliance on the Internet for mission and business-critical applications (where businesses require 7x24 availability for their

Narten

Expires January 27, 2008

[Page 11]

services) and the general decrease in cost of Internet connectivity.

4.3. End Site Renumbering

It is generally considered painful and costly to renumber a site, with the cost proportional to the size and complexity of the network and most importantly, to the degree that addresses are stored in places that are difficult in practice to update. When using PA space, a site must renumber when changing providers. Larger sites object to this cost and view the requirement to renumber akin to being held "hostage" to the provider from which PA space was obtained. Consequently, many sites desire PI space. Having PI space provides independence from any one provider and makes it easier to switch providers (for whatever reason). However, each individual PI prefix must be propagated throughout the DFZ and adds to the DFZ routing load.

It should be noted that while larger sites may also want to multihome, the cost of renumbering drives some sites to seek PI space, even though they do not multihome.

4.4. Acquisitions and Mergers

Acquisitions and mergers take place for business reasons, which usually have little to do with the network topologies of the impacted organizations. When a business sells off part of itself, the assets may include networks, attached devices, etc. A company that purchases or merges with other organizations may quickly find that its network assets are numbered out of many different and unaggregatable address blocks. Consequently, individual organizations may find themselves unable to announce a single prefix for all of their networks without renumbering a significant portion of its network.

Likewise, selling off part of a business may involve selling part of a network as well, resulting in the fragmentation of one address block into two (or more) smaller blocks. Because the resultant blocks belong to different companies, they can no longer be advertised by a single aggregate and the resultant fragments may need to be advertised individually into the DFZ.

4.5. RIR Address Allocation Policies

ISPs and multihoming end sites obtain address space from RIRs. As an entity grows, it needs additional address space and requests more from its RIR. In order to be able to obtain additional address space that can be aggregated with the previously-allocated address space, the RIR must keep a reserve of space that the requester can grow into

in the future. But any reserved address space cannot be used for any other purpose. Hence, there is an inherent conflict between holding address space in reserve to allow for the future growth of an existing allocation and using address space efficiently. In IPv4, there has been a heavy emphasis on conserving address space and obtaining efficient utilization. Consequently, insufficient space has been held in reserve to allow for the growth of all sites and some allocations have had to be made from discontinuous address blocks. For IPv6, a greater emphasis has been placed on aggregation.

4.6. Dual Stack Pressure on the Routing Table

The recommended IPv6 deployment model is dual-stack, where IPv4 and IPv6 are run in parallel across the same links. This has two implications for routing. First, although alternative scenarios are possible, it seems likely that many routers will be supporting both IPv4 and IPv6 simultaneously and will thus be managing both IPv4 and IPv6 routing tables within a single router. Second, for sites connected via both IPv4 and IPv6, both IPv4 and IPv6 prefixes will need to be propagated into the routing system. Consequently, dual-stack routers will maintain both an IPv4 and IPv6 route to reach the same destination.

It is possible to make some simple estimates on the approximate size of the IPv6 tables that would be needed if all sites reachable via IPv4 today were also reachable via IPv6. In theory, each autonomous system (AS) needs only a single aggregate route. This provides a lower bound on the size of the fully-realized IPv6 routing table. (As of July 2007, [CIDR4] states there are 25,836 active ASes in the routing system.)

A single IPv6 aggregate will not allow for inbound traffic engineering. End sites will need to advertise a number of smaller prefixes into the DFZ if they desire to gain finer grained control over their IPv6 inbound traffic. This will increase the size of the IPv6 routing table beyond the lower bound discussed above. There is reason to expect the IPv6 routing table will be smaller than the current IPv4 table, however, because the larger initial assignments to end sites will minimize the de-aggregation that occurs when a site must go back to its upstream address provider or RIR and receive a second, non-contiguous assignment.

It is possible to extrapolate what the size of the IPv6 Internet routing table would be if widespread IPv6 adoption occurred, from the current IPv4 Internet routing table. Each active AS (25,836) would require at least one aggregate. In addition, the IPv6 Internet table would also carry more specific prefixes for traffic engineering. Assume that the IPv6 Internet table will carry the same number of

Narten

Expires January 27, 2008

[Page 13]

more specifics as the IPv4 Internet table. In this case one can take the number of IPv4 Internet routes and subtract the number of CIDR aggregates that they could easily be aggregated down to. As of July 2007, the 229,789 routes can be easily aggregated down to 150,018 CIDR aggregates [CIDR4]. That difference yields 79,771 extra more specific prefixes. Thus if each active AS (25,836) required one aggregate, and an additional 79,771 more specifics were required, then the IPv6 Internet table would be 105,607 prefixes.

4.7. Internal Customer Routes

In addition to the Internet routing table, networks must also carry their internal routing table. Internal routes are defined as more specific routes that are not advertised to the DFZ. This primarily consists of prefixes that are a more specific of a provider aggregate (PA) and are assigned to a single homed customer. The DFZ need only carry the PA aggregate in order to deliver traffic to the provider. However, the provider's routers require the more specific route to deliver traffic to the end site.

This could also consist of more specific prefixes advertised by multi-homed customers with the no-export community. This is useful when the fine grained control of traffic to be influenced can be contained to the neighboring network.

For a large ISP, the internal IPv4 table can be between 50,000 and 150,000 routes. During the dot com boom some ISPs had more internal prefixes than there were in the Internet table. Thus the size of the internal routing table can have significant impact on the scalability and should not be discounted.

4.8. IPv4 Address Exhaustion

The IANA and RIR free pool of IPv4 addresses will be exhausted within a few years. As the free pool shrinks, the size of the remaining unused blocks will also shrink and unused blocks previously held in reserve for expansion of existing allocations or otherwise not used due to their smaller size will be allocated for use. Consequently, as the community looks to use every piece of available address space (no matter how small) there will be an increasing pressure to advertise additional prefixes in the DFZ.

5. Pressures on Path Computation Load

This section describes a number of trends and pressures that are contributing to the overall load of computing Internet paths. The previous section described pressures that are increasing the size of the routing table. Even if the size could be bounded, the amount of work needed to maintain paths for a given set of prefixes appears to be increasing.

5.1. Interconnection Richness

The degree of interconnectedness between ASes has increased in recent years. That is, the Internet as whole is becoming "flatter" with an increasing number of possible paths interconnecting sites [ref? gih has observed this]. As the number of possible paths increase, the amount of computation needed to find a best path also increases. This computation comes into effect whenever a change in path characteristics occurs, whether from a new path becoming available, an existing path failing, or a change in the attributes associated with a potential path. Thus, even if the total number of prefixes were to stay constant, an increase in the interconnection richness implies an increase in the resources needed to maintain routing tables.

5.2. Multihoming

Multihoming places pressure on the routing system in two ways. First, an individual prefix for a multihomed site (whether PI or PA) must be propagated into the routing system, so that other sites can find a good path to the site. Even if the site's prefix comes out of a PA block, an individual prefix for the site needs to be advertised so that the most desirable path to the site can be chosen when the path through the aggregate is sub-optimal. Second, a multi-homed site will be connected to the Internet in more than one place, increasing the overall level of interconnection richness. If an outage occurs on any of the circuits connecting the site to the Internet, those changes will be propagated into the routing system. In contrast, a singly-homed site numbered out of a Provider Aggregate places no additional routing load in the DFZ as the details of the connectivity status to the site are kept internal to the provider to which it connects.

5.3. Traffic Engineering

The mechanisms used to achieve multihoming and inbound Traffic Engineering are the same. In both cases, a specific prefix is advertised into the routing system to "catch" traffic and route it over a different path than it would otherwise be carried. When

multihoming, the specific prefix is one that differs from that of its ISP or is a more-specific of the ISP's PA. Traffic Engineering is achieved by taking one prefix and dividing into a number of smaller and more-specific ones, and advertising them in order to gain finer-grained control over the paths used to carry traffic covered by those prefixes.

Traffic Engineering increases the number of prefixes carried in the routing system. In addition, when a circuit fails (or the routing attributes associated with the circuit change), additional load is placed on the routing system by having multiple prefixes potentially impacted by the change, as opposed to just one.

5.4. Questionable Operational Practices?

Some operators are believed to engage in operational practices that increase the load on the routing system.

5.4.1. Rapid shuffling of prefixes

Some networks try to assert fine-grained control of inbound traffic by modifying route announcements frequently in order to migrate traffic to less loaded links quickly. The goal of this is to achieve higher utilization of multiple links. In addition, some route selection devices actively measure link or path utilization and attempt to optimize inbound traffic by withholding or depreferencing certain prefixes in their advertisements. In short, any system that actively measures load and modifies route advertisements in real time increases the load on the routing system, as any change in what is advertised must ripple through the entire routing system.

5.4.2. Anti-Route Hijacking

In order to reduce the threat of accidental (or intentional) hijacking of its address space by an unauthorized third party, some sites advertise their space as a set of smaller prefixes rather than as one aggregate. That way, if someone else advertised a path for the larger aggregate (or a small piece of the aggregate), it will be ignored in favor of the more specific announcements. This increases both the number of prefixes advertised, and the number of updates.

5.4.3. Operational Ignorance

It is believed that some undesirable practices result from operator ignorance, where the operator is unaware of what they are doing and the impact that has on the DFZ.

The default behavior of most BGP configurations is to automatically

propagate all learned routes. That is, one must take explicit configuration steps to prevent the automatic propagation of learned routes. In addition, it is often significant work to figure out how to (safely) aggregate routes (and which ones to aggregate) in order to reduce the number of advertisements propagated elsewhere. While vendors could provide additional configuration "knobs" to reduce leakage, the implementation of additional features increases complexity and some operators may fear that the new configuration will break their existing routing setup. Finally, leaking routes unnecessarily does not generally harm those with the misconfiguration, hence, there is less motivation to address the problem.

5.5. RIR Policy

RIR address policy has direct impact on the routing load because address policy determines who is eligible for a PI assignment (which impacts how many are given out in practice) and the size of the assignment (which impacts how much address space can be aggregated within a single assignment). If PI assignments for end sites did not exist, then those end sites would not advertise their own prefix directly into the global routing system; instead their address block would be covered by their provider's aggregate. That said, RIRs have adopted PI policies in response to community demand, for reasons described elsewhere (e.g., to support multihoming and to avoid the need to renumber). In short, RIR policy can be seen as a symptom rather than a root cause.

6. Problem Statement

There is a need for an approach to routing and addressing that:

1. Reduces the growth rate of the DFZ routing load, where the routing load is dependent on:
 - A. The number of individual prefixes in the DFZ
 - B. The update rate associated with those prefixes.
2. Allows any end site wishing to multihome to do so
3. Supports ISP and enterprise TE needs
4. Allows end sites to switch providers while minimizing configuration changes to internal end site devices.
5. Provides meaningful benefits to the parties who bear the costs of deploying and maintaining the technology.

The problem statement in this document has purposefully been scoped to focus on the growth of the routing update function of the DFZ. Other problems that may seem related, but do not directly impact the route scaling problem are not considered to be "in scope" at this time. For example, Mobile IP [[RFC2002](#)], [RFC3775](#)] and NEMO [[RFC3963](#)] place no pressures on the routing system. They are layered on top of existing IP, using tunneling to forward packets via a care-of addresses. Hence, "improving" these technologies (e.g., by having them leverage a solution to the multihoming problem), while a laudable goal, is not considered a part of this problem statement.

7. Security Considerations

None.

8. Acknowledgments

The initial version of this document was produced by the Routing and Addressing Directorate (<http://www.ietf.org/IESG/content/radir.html>). The membership of the directorate at that time included Marla Azinger, Vince Fuller, Vijay Gill, Thomas Narten, Erik Nordmark, Jason Schiller, Peter Schoenmaker, and John Scudder.

9. Informative References

- [RFC2002] Perkins, C., "IP Mobility Support", [RFC 2002](#), October 1996.
- [RFC3963] Devarapalli, V., Wakikawa, R., Petrescu, A., and P. Thubert, "Network Mobility (NEMO) Basic Support Protocol", [RFC 3963](#), January 2005.
- [RFC4632] Fuller, V. and T. Li, "Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan", [BCP 122](#), [RFC 4632](#), August 2006.

Author's Address

Thomas Narten
IBM

Email: narten@us.ibm.com

Full Copyright Statement

Copyright (C) The IETF Trust (2007).

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.

Acknowledgment

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).

