A Scalable Addressing Allocation Architecture for IPv6
draft-azinger-scalable-addressing-01

Abstract

This document presents a scalable architecture for assigning and aggregating IPv6 address space. The current IPv4 assignment and addressing architecture has been successful in helping to scale the IPv4 routing architecture. This same architecture, when carried forward to IPv6, will help to ensure that the IPv6 routing architecture is sustainable.

Status of This Memo

Copyright Notice

Table of Contents

---

1. **Introduction**   TOC

This document presents a scalable architecture for assigning and
aggregating IPv6 address space. The current IPv4 addressing aggregation
strategy was defined in [RFC1519] (Fuller, V., Li, T., Yu, J., and K.
Varadhan, "Classless Inter-Domain Routing (CIDR): an Address Assignment
and Aggregation Strategy," September 1993.) (and updated in [RFC4632]
(Fuller, V. and T. Li, "Classless Inter-domain Routing (CIDR): The
Internet Address Assignment and Aggregation Plan," August 2006.)) and
the IPv4 address allocation architecture was defined in [RFC1518]
(Rekhter, Y. and T. Li, "An Architecture for IP Address Allocation with
CIDR," September 1993.). A similar address allocation architecture was

proposed for IPv6 in [RFC1887] (Rekhter, Y. and T. Li, "An Architecture for IPv6 Unicast Address Allocation," December 1995.). The objective of this document is to update the previous documents and provide the best current guidance on an address allocation architecture to help manage the growth of routing tables in IPv6.

The Internet has continued to evolve and the demands placed on its infrastructure continue to grow at an increasing rate. While there are a number of contributing factors, there are a few key elements that have led to a concerning escalation in routing table growth and have made scalability an area of serious concern for network operators. Effort must be put forward to minimize the impact of IPv6 deployment to the routing subsystem. Two key aspects of this system include routing table churn composed of routing advertisements and withdrawals and the routing table size as measured by the number of entries in the DFZ, the Default-Free zone. While retaining current Internet practices, this document addresses the problem of routing table size by examining steps to minimize the impact of Multi-homing and Traffic Engineering, two widely implemented features that provide enhanced network resiliency and traffic path control.

2. Organizational, technical, and policy issues

2.1. Delegation to IANA

[RFC2860] (Carpenter, B., Baker, F., and M. Roberts, "Memorandum of Understanding Concerning the Technical Work of the Internet Assigned Numbers Authority," June 2000.) is a Memorandum of Understanding (MoU) between IETF and ICANN that delegates the technical work of assignment and allocation of addresses to IANA (a function of ICANN) on behalf of the IETF and IRTF (see sections 1 and 4.3). The MoU directs IANA to comply with "the criteria and procedures specified in RFCs, including Proposed, Draft, and full Internet Standards and Best Current Practice documents" (section 4.1).

Technical disputes in this agreement are first directed to the IESG, and if not resolved, arbitrated by the IAB (sections 4.1, R 4.2).

The document specifically stipulates that policy issues around IP addressing are outside of the scope of the MoU (section 4.4).

2.2. Delegation to NRO

ICANN in turn delegates block allocation to the Number Resource Organization (NRO) [ASOMoU] (, "ICANN Address Supporting Organization (ASO) MoU," .). The NRO is composed of the Regional Internet Registries (RIRs) [NROMoU] (, "NRO Memorandum of Understanding," .).

2.3. Technical issues vs. policy issues

The documents cited above make it clear that policy issues are delegated from the IETF through ICANN and IANA to the RIRs. The same documents make it equally clear that the IETF is still responsible for technical matters and procedures.
The routing subsystem is a key component of the Internet. Without routing, no packets are delivered. The routing subsystem consists of multiple routing protocols as well as the procedures for utilizing these protocols to provide coherent and timely routing services. The procedures for deploying and operating the routing subsystem are characterized in the routing architecture. Specifying the routing architecture is a technical issue.
A key issue in the routing architecture is the scalability of the architecture. If the architecture fails to scale, then the routing subsystem can fail, either in its basic functionality, from a performance perspective, or from a cost perspective. For the routing architecture to scale, the amount of data propagated within the routing subsystem must be limited, as the routing subsystem must ultimately have a hardware instantiation, and unlimited hardware simply does not exist.
As technology progresses over time, the intrinsic capabilities of hardware improves. There are multiple dimensions for improvement, each with their own trend lines and projections. Based on these trends, we can reasonably expect that hardware scalability will improve over time, tracking these trends. As a result, for the routing architecture to scale, it must operate within the growth rates of these trends. Thus, ensuring that the routing subsystem scales at no more than an appropriate rate of growth is a technical issue.
The scalability of the routing subsystem is wholly dependent on the addressing architecture for the network. Routing views the network as a graph composed of nodes and edges between those nodes. Carrying information about each node and edge of the graph does not scale, so routing works by creating abstractions of entire subgraphs. The most productive abstractions happen when the subgraph is closely topologically related, such as when all of the nodes in the subgraph are interconnected by the edges in the subgraph. We can further create hierarchies of abstractions to get further scalability.

Poorly chosen abstractions that do not align with the topology of the network result in abstractions that do not reduce the data burden on routing, as additional data is needed for path computation. Thus, the correct procedures for creating abstractions in the topology are also a technical issue.

To manipulate data about the network in a convenient manner, we give names to nodes in the network, where each name is an address. If names are assigned in a manner that is consistent with the hierarchy of abstractions, then a single name can also be used to represent a subgraph of the network, such as a single site. These types of names are known as prefixes or aggregates and the overall procedure for assigning addresses to nodes and aggregates to subgraphs is the addressing architecture for the network. Clearly, the correct application and operation of the addressing architecture is central to the operation and ongoing scalability of the routing architecture.

The operation of the addressing architecture involves both technical issues and policy issues. Procedures for determining subgraph boundaries and subsequent prefix allocations are clearly driven by the technical requirements of the architecture. From a technical perspective, the graph is composed of many different types of nodes. These include leaf nodes, multiply connected nodes with a minimal number of edges, and densely connected nodes. Determining the current connectivity of a site is clearly a technical issue. Judging how that site might evolve in the future and thus the role that the site should play in the addressing architecture is a matter of policy.

These issues are simply examples. There are likely to be many more, some of which may be contentious. Completely separating the technical issues from the policy issues is decidedly non-trivial and is most likely an inefficient exercise. In reality and for pragmatic purposes, it is necessary that all issues be resolved in a consistent and compatible manner. The end goal is clearly shared: the successful operation of a scalable routing and addressing architecture.

This document deals with the technical aspects of the addressing architecture.

---

3. Contributing factors to the scalability problem   [TOC]

There are several factors that work against routing table scalability. A full description of the contributing factors and views can be read in [I-D.narten-radir-problem-statement] (Narten, T., "On the Scalability of Internet Routing," February 2010.). The exhaustion of the unassigned IPv4 address space is the principal motivator resulting in two of the key growth drivers. The first driver is the presence of increasingly longer prefixes in the DFZ. Over the years the longest prefix generally accepted globally has increased from a relatively small number of classful prefixes to a preponderance of classless /24 CIDR prefixes. As

IPv4 address availability diminishes, more Internet users are and will
continue to push their providers to route even longer prefixes
externally that in the past were filtered. This is something that we
must look to minimize and find ways to deter as much as possible for
IPv6.

The second driver resulting from IPv4 address exhaustion is the rapid
uptake in IPv6 deployment by providers and end users. This adoption,
while clearly in the best interest for the long term viability of the
Internet, contributes a unique set of challenges that must be addressed
to promote efficient routing table growth. Some of these challenges
visible today are the liberal assignment of Provider Independent (PI)
space to end users, micro-allocations, and critical network
infrastructure allocations by the various RIRs.

These drivers have increased the need for more guidance on the
addressing architecture in order to limit the number of unnecessary
entries in the global routing table. The future impact of this
increased pressure on routing table growth is an area of immediate
concern.

---

4. Aggregation

The common method for reducing state on both internal and external
routing tables is through aggregation of information. Borrowing from
experience gained in operating IPv4 networks, in order for aggregation
to succeed in reducing the global routing system growth rate, the IPv6
address assignment process needs to make aggregation of routing
information along topological lines. In general, the topology of the
network has not changed since IPv4 CIDR and even with IPv6 the topology
of the network is still determined by the service providers who have
built it. Topologically significant address assignments are necessarily
service-provider oriented.

Start of Excerpt from [RFC4632] (Fuller, V. and T. Li, "Classless
Inter-domain Routing (CIDR): The Internet Address Assignment and
Aggregation Plan," August 2006.)

> The assignment of prefixes is intended to roughly follow the
> underlying Internet topology so that aggregation can be used to
> facilitate scaling of the global routing system. One implication of
> this strategy is that prefix assignment and aggregation is generally
> done according to provider-subscriber relationships, since that is
> how the Internet topology is determined. [Section 3]
>
> Aggregation is simple for an end site that is connected to one
> service provider: it uses address space assigned by its service
> provider, and that address space is a small piece of a larger block
> allocated to the service provider. No explicit route is needed for

the end site; the service provider advertises a single aggregate route for the larger block. This advertisement provides reachability and routeability for all the customers numbered in the block.

There are two, more complex, situations that reduce the effectiveness of aggregation:

> *An organization that is multi-homed. Because a multi-homed organization must be advertised into the system by each of its service providers, it is often not feasible to aggregate its routing information into the address space of any one of those providers. Note that the organization still may receive its address assignment out of a service provider's address space (which has other advantages), but that a route to the organization's prefix is, in the most general case, explicitly advertised by all of its service providers. For this reason, the global routing cost for a multi-homed organization is generally the same as it was prior to the adoption of CIDR. A more detailed consideration of multi-homing practices can be found in [RFC4116] (Abley, J., Lindqvist, K., Davies, E., Black, B., and V. Gill, "IPv4 Multihoming Practices and Limitations," July 2005.).

> *An organization that changes service provider but does not renumber. This has the effect of "punching a hole" in one of the original service provider's aggregated route advertisements. CIDR handles this situation by requiring that the newer service provider to advertise a specific advertisement for the re-homed organization; this advertisement is preferred over provider aggregates because it is a longer match. To maintain efficiency of aggregation, it is recommended that an organization that changes service providers plan eventually to migrate its network into a prefix assigned from its new provider's address space. To this end, it is recommended that mechanisms to facilitate such migration, such as dynamic host address assignment that uses [RFC2131] (Droms, R., "Dynamic Host Configuration Protocol," March 1997.)), be deployed wherever possible, and that additional protocol work be done to develop improved technology for renumbering. [Section 4.1]

End of Excerpt from [RFC4632] (Fuller, V. and T. Li, "Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan," August 2006.)
It is important to recognize that some efficiency can still be gained with multi-homed sites (and in general, for any site composed of multiple, logical IPv6 networks).

Start of Excerpt from [RFC4632] (Fuller, V. and T. Li, "Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan," August 2006.)

> By allocating a contiguous power-of-two block address space to the site (as opposed to multiple, independent prefixes), the site's routing information may be aggregated into a single prefix. Also, since the routing cost associated with assigning a multi-homed site out of a service provider's address space is no greater than the old method of sequential number assignment by a central authority, it makes sense to assign all end-site address space out of blocks allocated to service providers.
>
> It is also worthwhile to mention that since aggregation may occur at multiple levels in the system, it may still be possible to aggregate these anomalous routes at higher levels of whatever hierarchy may be present. For example, if a site is multi-homed to two relatively small providers that both obtain connectivity and address space from the same large provider, then aggregation by the large provider of routes from the smaller networks will include all routes to the multi-homed site. The feasibility of this sort of second-level aggregation depends on whether topological hierarchy exists among a site, its directly-connected providers, and other providers to which they are connected; it may be practical in some regions of the global Internet but not in others. [Section 4.1]

End of Excerpt from [RFC4632] (Fuller, V. and T. Li, "Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan," August 2006.)

---

5. Allocation plan

Allocations of shorter prefixes are best provided to network service providers from their regional registries. RIR initial and subsequent allocation policy to service providers should allow for a minimum of 2 years worth of usage based on historical or business plan projections. Organizations should be assigned appropriate subnets from their network service providers larger aggregate allocations that are in turn appropriately sized for organizations wishing to multi-home.
Start of Excerpt from [RFC4632] (Fuller, V. and T. Li, "Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan," August 2006.)

> Hierarchical delegation of addresses in this manner implies that sites with addresses assigned out of a given service provider are, for routing purposes, part of that service provider and will be routed via its infrastructure. This implies that routing information

about multi-homed organizations (i.e., organizations connected to more than one network service provider) will still need to be known by higher levels in the hierarchy.

A historical perspective on these issues is described in [RFC1518] (Rekhter, Y. and T. Li, "An Architecture for IP Address Allocation with CIDR," September 1993.). Additional discussion may also be found in [RFC3221] (Huston, G., "Commentary on Inter-Domain Routing in the Internet," December 2001.). [Section 4.2]

End of Excerpt from [RFC4632] (Fuller, V. and T. Li, "Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan," August 2006.)

Similarly to the days of classful routing, IPv6 is following the same historical path of giving PI assignments. It is in the interests of the network infrastructure to document a best practice for obtaining IPv6 addresses, and it is recommended that most, if not all, network numbers be distributed through service providers. Using the process proposed in this document will support this from becoming a growing problem and will also reduce the scalability concerns core engineers face and the workload for Regional Registries.

---

6. Current Statistics and Projections TOC

The good news is that IPv6 has started growing at a significant rate. The bad news is that IPv6 has started growing at a significant rate. Table 1 shows the observed growth for 2009.

	Jan '09	Dec '09	Growth
Prefix count	1,600	2,460	54%
Roots	1,310	1,970	50%
More Specifics	290	490	69%
AS Count	1,220	1,830	50%
Transit	300	390	30%
Stub	920	1,440	56%

Table 1: IPv6 Routing Table Statistics for 2009 [Huston] (Huston, G., "BGP in 2009," .).

There are several salient points that should be extracted from this table. The first, and foremost, is that the routing table is now growing rapidly. At 54% growth, this is faster than Moore's law would accommodate. The roots are prefixes that have no 'less specifics' in the routing table. Even at 50% growth per year, this number exceeds Moore's law. More specifics are typically injected to support traffic engineering or multi-homing.

The AS count growth shows the number of new organizations participating in BGP. Transit ASes are routing domains that have multiple peer ASes. Stub ASes are routing domains that have only a single peer AS.

6.1. Analysis

These numbers show that 610 new organizations have joined IPv6 routing. Of these new organizations, 85% are stub ASes. The new organizations are injecting 860 new prefixes. Of these, 76% are root prefixes. Since any new AS must inject at least one prefix into routing to be counted, there would appear to be a very high correlation between new stub ASes and new root prefixes. From this, it seems reasonable to conclude that the bulk of the new root prefixes are injected by stub ASes. Further, since it seems unlikely that most of these stub ASes will turn into transit ASes in the future, it also seems reasonable to conclude that these organizations are actually end-user organizations who are injecting routes based on their PI address assignments.
Thus, the bulk of the routing table growth appears to be due to PI prefix injection.

6.2. Projections

Given the high state of flux in the deployment of IPv6, it seems difficult to conclude that the statistics from 2009 will be representative of future routing table growth. Thanks to the influx of new users who are being forced onto IPv6 by the impending IPv4 runout, there are plausible arguments that would suggest that growth could accelerate. There are also plausible arguments that suggest that as IPv6 deployment reaches ubiquity, that the growth might curtail in a logistic S-curve. Lacking more data, it is difficult to clearly argue that either of these results is inevitable.
It is possible, however, to look at the implications of the current growth rate, if it is sustained at the 2009 rate of 54%. Table 2 shows this growth rate:

Year	Size
2009	2,460
2010	3,788
2011	5,834
2012	8,985
2013	13,836
2014	21,308

2015	32,814
2020	284,225
2025	2,461,879
2040	1,599,843,323

Table 2: 54% growth rate, extrapolated

6.3. Impact of Scalable Addressing

With the adoption of the plan outlined here, growth of the routing table in a default-free router is greatly reduced since most new address assignments will come from one of the large blocks allocated to the service providers. This plan recognizes the continued need for multi-homing and the requirement to offer multi-homing via IPv6. Due to this requirement multi-homing will be the main reason for the continued growth of the routing table size but not because of independent subnet statements based solely on the desire for independence.

---

7. Protocol

This document requires that all parties implement routing protocols for IPv6 as previously published for IPv4 in [RFC4632] (Fuller, V. and T. Li, "Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan," August 2006.).

---

8. Rules for Route Advertisements

This document requires that all parties follow the rules for route advertisements for IPv6 as previously published in [RFC1887] (Rekhter, Y. and T. Li, "An Architecture for IPv6 Unicast Address Allocation," December 1995.) and as similarly published for IPv4 in [RFC4632] (Fuller, V. and T. Li, "Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan," August 2006.).

---

9. Responsibility of configuration and aggregation

This document requires that all parties take responsibility of configuration or aggregation for IPv6 as previously published [RFC1887]

[(Rekhter, Y. and T. Li, "An Architecture for IPv6 Unicast Address Allocation," December 1995.)](#) and as similarly published for IPv4 in [[RFC4632] (Fuller, V. and T. Li, "Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan," August 2006.)](#).

---

10. Procedural Changes [TOC](#)

It is possible that some organizations will need to alter their filters to follow the guidance of this document. This is minimal and should not be considered an issue.

---

11. Recommendations [TOC](#)

Internet Registries should begin to hand out large IPv6 blocks to network service providers in order to accommodate both their growth and their customers' growth. In addition Internet Registries should severely limit or eliminate the amount of PI assignments in order to help facilitate the decrease in routing table growth. Service providers will allocate address blocks from their aggregates to their customer organizations with multi-homing requirements. Implementation and deployment of these modifications should occur immediately.

---

12. Security Considerations [TOC](#)

The recommendations in this document create no new security concerns.

---

13. IANA Considerations [TOC](#)

This document makes no requests to IANA.

---

14. Acknowledgements [TOC](#)

The authors would like to extend their thanks to the authors of [[RFC1887] (Rekhter, Y. and T. Li, "An Architecture for IPv6 Unicast Address Allocation," December 1995.)](#) and [[RFC4632] (Fuller, V. and T.](#)

Li, "[Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan," August 2006.)](#) (and by extension, to the authors of [[RFC1519] (Fuller, V., Li, T., Yu, J., and K. Varadhan, "Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy," September 1993.)](#)). Much of that work has been incorporated directly into this document as it is conceptually identical and simply translated to IPv6 herein.

---

15. References   [TOC](#)

15.1. Normative References

[TOC](#)

[RFC4632]	Fuller, V. and T. Li, "[Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan](#)," BCP 122, RFC 4632, August 2006 ([TXT](#)).

15.2. Informative References

[TOC](#)

[RFC1518]	Rekhter, Y. and T. Li, "[An Architecture for IP Address Allocation with CIDR](#)," RFC 1518, September 1993 ([TXT](#)).
[RFC1519]	Fuller, V., Li, T., Yu, J., and K. Varadhan, "[Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy](#)," RFC 1519, September 1993 ([TXT](#)).
[RFC1887]	Rekhter, Y. and T. Li, "[An Architecture for IPv6 Unicast Address Allocation](#)," RFC 1887, December 1995 ([TXT](#)).
[RFC2131]	Droms, R., "[Dynamic Host Configuration Protocol](#)," RFC 2131, March 1997 ([TXT](#), [HTML](#), [XML](#)).
[RFC2860]	Carpenter, B., Baker, F., and M. Roberts, "[Memorandum of Understanding Concerning the Technical Work of the Internet Assigned Numbers Authority](#)," RFC 2860, June 2000 ([TXT](#)).
[RFC3221]	Huston, G., "[Commentary on Inter-Domain Routing in the Internet](#)," RFC 3221, December 2001 ([TXT](#)).
[RFC4116]	Abley, J., Lindqvist, K., Davies, E., Black, B., and V. Gill, "[IPv4 Multihoming Practices and Limitations](#)," RFC 4116, July 2005 ([TXT](#)).

[I-D.narten-radir-problem-statement]	Narten, T., "On the Scalability of Internet Routing," draft-narten-radir-problem-statement-05 (work in progress), February 2010 (TXT).
[Huston]	Huston, G., "BGP in 2009."
[ASOMoU]	"ICANN Address Supporting Organization (ASO) MoU."
[NROMoU]	"NRO Memorandum of Understanding."

Authors' Addresses

	Marla Azinger
	Frontier Communications Corporation
	Vancouver, WA
	USA
EMail:	marla.azinger@frontiercorp.com
URI:	http://www.frontiercorp.com/
	Tony Li
	Cisco Systems
	170 West Tasman Dr.
	San Jose, CA 95134
	USA
Phone:	+1 408 853 9317
EMail:	tony.li@tony.li
	Jason Weil
	Cox Communications
	1400 Lake Hearn Dr.
	Atlanta, GA 95134
	USA
Phone:	+1 404 269 6809
EMail:	jason.weil@cox.com