Network Working Group
Request for Comments: 2374
Obsoletes: 2073
Category: Standards Track

R. Hinden
Nokia
M. O'Dell
UUNET
S. Deering
Cisco

July 1998

An IPv6 Aggregatable Global Unicast Address Format

Status of this Memo

Copyright Notice

1.0 Introduction

This document defines an IPv6 aggregatable global unicast address format for use in the Internet. The address format defined in this document is consistent with the IPv6 Protocol [IPV6] and the "IPv6 Addressing Architecture" [ARCH]. It is designed to facilitate scalable Internet routing.

This documented replaces RFC 2073, "An IPv6 Provider-Based Unicast Address Format". RFC 2073 will become historic. The Aggregatable Global Unicast Address Format is an improvement over RFC 2073 in a number of areas. The major changes include removal of the registry bits because they are not needed for route aggregation, support of EUI-64 based interface identifiers, support of provider and exchange based aggregation, separation of public and site topology, and new aggregation based terminology.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119].

2.0 Overview of the IPv6 Address

IPv6 addresses are 128-bit identifiers for interfaces and sets of interfaces. There are three types of addresses: Unicast, Anycast, and Multicast. This document defines a specific type of Unicast address.

In this document, fields in addresses are given specific names, for example "subnet". When this name is used with the term "ID" (for "identifier") after the name (e.g., "subnet ID"), it refers to the contents of the named field. When it is used with the term "prefix" (e.g. "subnet prefix") it refers to all of the addressing bits to the left of and including this field.

IPv6 unicast addresses are designed assuming that the Internet routing system makes forwarding decisions based on a "longest prefix match" algorithm on arbitrary bit boundaries and does not have any knowledge of the internal structure of IPv6 addresses. The structure in IPv6 addresses is for assignment and allocation. The only exception to this is the distinction made between unicast and multicast addresses.

The specific type of an IPv6 address is indicated by the leading bits in the address. The variable-length field comprising these leading bits is called the Format Prefix (FP).

This document defines an address format for the 001 (binary) Format Prefix for Aggregatable Global Unicast addresses. The same address format could be used for other Format Prefixes, as long as these Format Prefixes also identify IPv6 unicast addresses. Only the "001" Format Prefix is defined here.

3.0 IPv6 Aggregatable Global Unicast Address Format

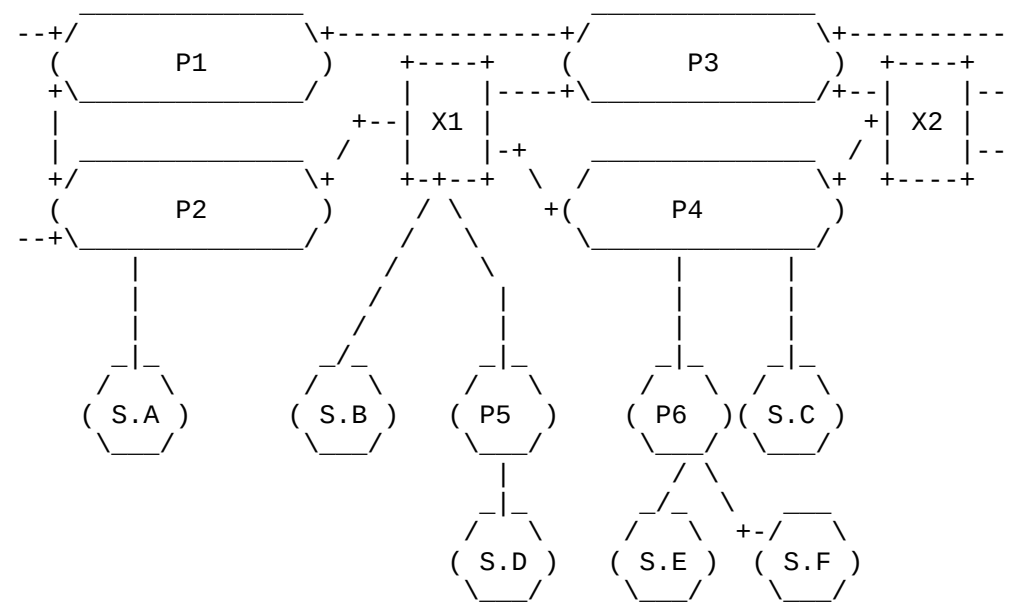
This document defines an address format for the IPv6 aggregatable global unicast address assignment. The authors believe that this address format will be widely used for IPv6 nodes connected to the Internet. This address format is designed to support both the current provider-based aggregation and a new type of exchange-based aggregation. The combination will allow efficient routing aggregation for sites that connect directly to providers and for sites that connect to exchanges. Sites will have the choice to connect to either type of aggregation entity.

While this address format is designed to support exchange-based aggregation (in addition to current provider-based aggregation) it is not dependent on exchanges for it's overall route aggregation properties. It will provide efficient route aggregation with only provider-based aggregation.

Aggregatable addresses are organized into a three level hierarchy:

- Public Topology
- Site Topology
- Interface Identifier

Public topology is the collection of providers and exchanges who provide public Internet transit services. Site topology is local to a specific site or organization which does not provide public transit service to nodes outside of the site. Interface identifiers identify interfaces on links.

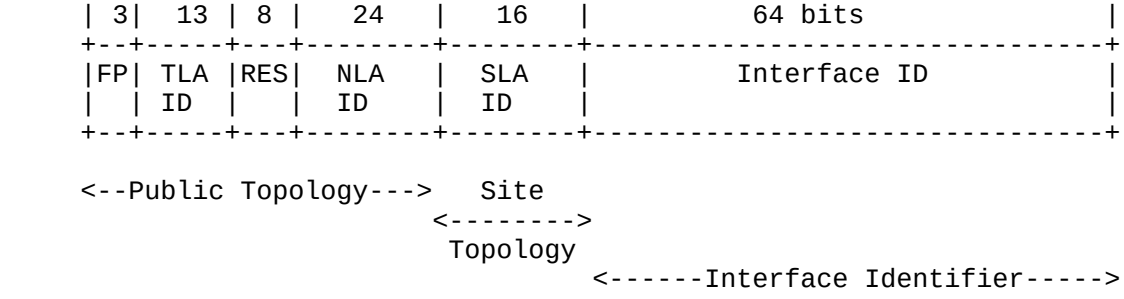


As shown in the figure above, the aggregatable address format is designed to support long-haul providers (shown as P1, P2, P3, and P4), exchanges (shown as X1 and X2), multiple levels of providers (shown at P5 and P6), and subscribers (shown as S.x) Exchanges (unlike current NAPs, FIXes, etc.) will allocate IPv6 addresses. Organizations who connect to these exchanges will also subscribe (directly, indirectly via the exchange, etc.) for long-haul service from one or more long-haul providers. Doing so, they will achieve

addressing independence from long-haul transit providers. They will be able to change long-haul providers without having to renumber their organization. They can also be multihomed via the exchange to more than one long-haul provider without having to have address prefixes from each long-haul provider. Note that the mechanisms used for this type of provider selection and portability are not discussed in the document.

3.1 Aggregatable Global Unicast Address Structure

The aggregatable global unicast address format is as follows:



The following sections specify each part of the IPv6 Aggregatable Global Unicast address format.

3.2 Top-Level Aggregation ID

Top-Level Aggregation Identifiers (TLA ID) are the top level in the routing hierarchy. Default-free routers must have a routing table entry for every active TLA ID and will probably have additional entries providing routing information for the TLA ID in which they are located. They may have additional entries in order to optimize routing for their specific topology, but the routing topology at all levels must be designed to minimize the number of additional entries fed into the default free routing tables.

This addressing format supports 8,192 (2^13) TLA ID's. Additional TLA ID's may be added by either growing the TLA field to the right into the reserved field or by using this format for additional format prefixes.

The issues relating to TLA ID assignment are beyond the scope of this document. They will be described in a document under preparation.

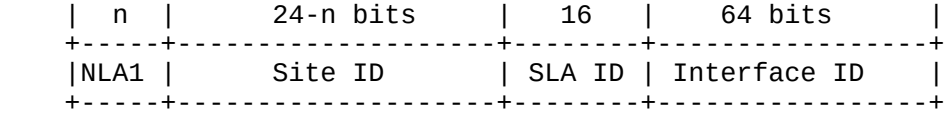
3.3 Reserved

The Reserved field is reserved for future use and must be set to zero.

The Reserved field allows for future growth of the TLA and NLA fields as appropriate. See section 4.0 for a discussion.

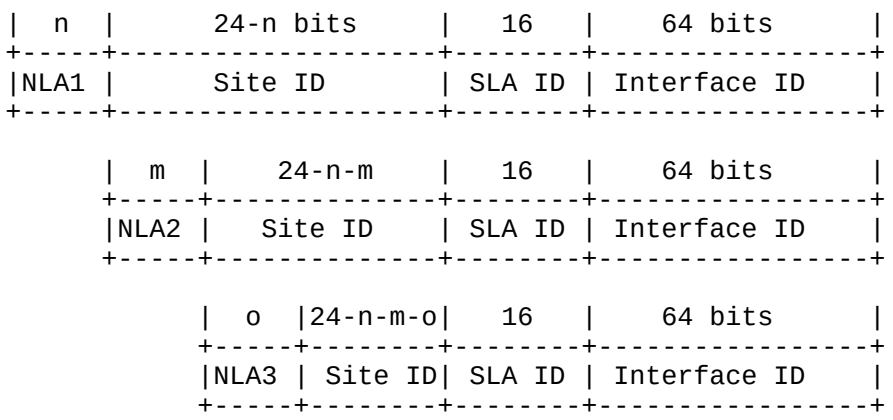
3.4 Next-Level Aggregation Identifier

Next-Level Aggregation Identifier's are used by organizations assigned a TLA ID to create an addressing hierarchy and to identify sites. The organization can assign the top part of the NLA ID in a manner to create an addressing hierarchy appropriate to its network. It can use the remainder of the bits in the field to identify sites it wishes to serve. This is shown as follows:



Each organization assigned a TLA ID receives 24 bits of NLA ID space. This NLA ID space allows each organization to provide service to approximately as many organizations as the current IPv4 Internet can support total networks.

Organizations assigned TLA ID's may also support NLA ID's in their own Site ID space. This allows the organization assigned a TLA ID to provide service to organizations providing public transit service and to organizations who do not provide public transit service. These organizations receiving an NLA ID may also choose to use their Site ID space to support other NLA ID's. This is shown as follows:



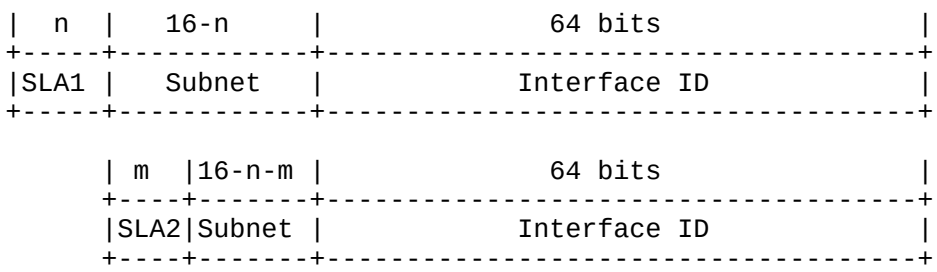
The design of the bit layout of the NLA ID space for a specific TLA ID is left to the organization responsible for that TLA ID. Likewise the design of the bit layout of the next level NLA ID is the responsibility of the previous level NLA ID. It is recommended that organizations assigning NLA address space use "slow start" allocation procedures similar to [RFC2050].

The design of an NLA ID allocation plan is a tradeoff between routing aggregation efficiency and flexibility. Creating hierarchies allows for greater amount of aggregation and results in smaller routing tables. Flat NLA ID assignment provides for easier allocation and attachment flexibility, but results in larger routing tables.

3.5 Site-Level Aggregation Identifier

The SLA ID field is used by an individual organization to create its own local addressing hierarchy and to identify subnets. This is analogous to subnets in IPv4 except that each organization has a much greater number of subnets. The 16 bit SLA ID field support 65,535 individual subnets.

Organizations may choose to either route their SLA ID "flat" (e.g., not create any logical relationship between the SLA identifiers that results in larger routing tables), or to create a two or more level hierarchy (that results in smaller routing tables) in the SLA ID field. The latter is shown as follows:



The approach chosen for structuring an SLA ID field is the responsibility of the individual organization.

The number of subnets supported in this address format should be sufficient for all but the largest of organizations. Organizations which need additional subnets can arrange with the organization they are obtaining Internet service from to obtain additional site identifiers and use this to create additional subnets.

3.6 Interface ID

Interface identifiers are used to identify interfaces on a link. They are required to be unique on that link. They may also be unique over a broader scope. In many cases an interfaces identifier will be the same or be based on the interface's link-layer address. Interface IDs used in the aggregatable global unicast address format are required to be 64 bits long and to be constructed in IEEE EUI-64 format [EUI-64]. These identifiers may have global scope when a global token (e.g., IEEE 48bit MAC) is available or may have local scope where a global token is not available (e.g., serial links, tunnel end-points, etc.). The "u" bit (universal/local bit in IEEE EUI-64 terminology) in the EUI-64 identifier must be set correctly, as defined in [ARCH], to indicate global or local scope.

The procedures for creating EUI-64 based Interface Identifiers is defined in [ARCH]. The details on forming interface identifiers is defined in the appropriate "IPv6 over <link>" specification such as "IPv6 over Ethernet" [ETHER], "IPv6 over FDDI" [FDDI], etc.

4.0 Technical Motivation

The design choices for the size of the fields in the aggregatable address format were based on the need to meet a number of technical requirements. These are described in the following paragraphs.

The size of the Top-Level Aggregation Identifier is 13 bits. This allows for 8,192 TLA ID's. This size was chosen to insure that the default-free routing table in top level routers in the Internet is

kept within the limits, with a reasonable margin, of the current routing technology. The margin is important because default-free routers will also carry a significant number of longer (i.e., more-specific) prefixes for optimizing paths internal to a TLA and between TLAs.

The important issue is not only the size of the default-free routing table, but the complexity of the topology that determines the number of copies of the default-free routes that a router must examine while computing a forwarding table. Current practice with IPv4 it is common to see a prefix announced fifteen times via different paths.

The complexity of Internet topology is very likely to increase in the future. It is important that IPv6 default-free routing support additional complexity as well as a considerably larger internet.

It should be noted for comparison that at the time of this writing (spring, 1998) the IPv4 default-free routing table contains approximately 50,000 prefixes. While this shows that it is possible to support more routes than 8,192 it is matter of debate if the number of prefixes supported today in IPv4 is already too high for current routing technology. There are serious issues of route stability as well as cases of providers not supporting all top level prefixes. The technical requirement was to pick a TLA ID size that was below, with a reasonable margin, what was being done with IPv4.

The choice of 13 bits for the TLA field was an engineering compromise. Fewer bits would have been too small by not supporting enough top level organizations. More bits would have exceeded what can be reasonably accommodated, with a reasonable margin, with current routing technology in order to deal with the issues described in the previous paragraphs.

If in the future, routing technology improves to support a larger number of top level routes in the default-free routing tables there are two choices on how to increase the number TLA identifiers. The first is to expand the TLA ID field into the reserved field. This would increase the number of TLA ID's to approximately 2 million. The second approach is to allocate another format prefix (FP) for use with this address format. Either or a combination of these approaches allows the number of TLA ID's to increase significantly.

The size of the Reserved field is 8 bits. This size was chosen to allow significant growth of either the TLA ID and/or the NLA ID fields.

The size of the Next-Level Aggregation Identifier field is 24 bits.

This allows for approximately sixteen million NLA ID's if used in a flat manner. Used hierarchically it allows for a complexity roughly equivalent to the IPv4 address space (assuming an average network size of 254 interfaces). If in the future additional room for complexity is needed in the NLA ID, this may be accommodated by extending the NLA ID into the Reserved field.

The size of the Site-Level Aggregation Identifier field is 16 bits. This supports 65,535 individual subnets per site. The design goal for the size of this field was to be sufficient for all but the largest of organizations. Organizations which need additional subnets can arrange with the organization they are obtaining Internet service from to obtain additional site identifiers and use this to create additional subnets.

The Site-Level Aggregation Identifier field was given a fixed size in order to force the length of all prefixes identifying a particular site to be the same length (i.e., 48 bits). This facilitates movement of sites in the topology (e.g., changing service providers and multi-homing to multiple service providers).

The Interface ID Interface Identifier field is 64 bits. This size was chosen to meet the requirement specified in [ARCH] to support EUI-64 based Interface Identifiers.

5.0 Acknowledgments

The authors would like to express our thanks to Thomas Narten, Bob Fink, Matt Crawford, Allison Mankin, Jim Bound, Christian Huitema, Scott Bradner, Brian Carpenter, John Stewart, and Daniel Karrenberg for their review and constructive comments.

6.0 References

[ALLOC] IAB and IESG, "IPv6 Address Allocation Management", RFC 1881, December 1995.

[ARCH] Hinden, R., "IP Version 6 Addressing Architecture", RFC 2373, July 1998.

[AUTH] Atkinson, R., "IP Authentication Header", RFC 1826, August 1995.

[AUTO] Thompson, S., and T. Narten., "IPv6 Stateless Address Autoconfiguration", RFC 1971, August 1996.

[ETHER] Crawford, M., "Transmission of IPv6 Packets over Ethernet Networks", Work in Progress.

[EUI64] IEEE, "Guidelines for 64-bit Global Identifier (EUI-64) Registration Authority", http://standards.ieee.org/db/oui/tutorials/EUI64.html, March 1997.

[FDDI] Crawford, M., "Transmission of IPv6 Packets over FDDI Networks", Work in Progress.

[IPV6] Deering, S., and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 1883, December 1995.

[RFC2050] Hubbard, K., Kosters, M., Conrad, D., Karrenberg, D., and J. Postel, "Internet Registry IP Allocation Guidelines", BCP 12, RFC 1466, November 1996.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

7.0 Security Considerations

IPv6 addressing documents do not have any direct impact on Internet infrastructure security. Authentication of IPv6 packets is defined in [AUTH].

8.0 Authors' Addresses

Robert M. Hinden
Nokia
232 Java Drive
Sunnyvale, CA 94089
USA

Phone: 1 408 990-2004
EMail: hinden@iprg.nokia.com

Mike O'Dell
UUNET Technologies, Inc.
3060 Williams Drive
Fairfax, VA 22030
USA

Phone: 1 703 206-5890
EMail: mo@uunet.uu.net

Stephen E. Deering
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA

Phone: 1 408 527-8213
EMail: deering@cisco.com

9.0 Full Copyright Statement