INTERNET-DRAFT R. Hinden, Ipsilon Networks
June 12, 1997 M. O'Dell, UUNET
S. Deering, Cisco

An IPv6 Aggregatable Global Unicast Address Format

<draft-ietf-ipngwg-unicast-aggr-01.txt>

Status of this Memo

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 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 both sites that connect directly to providers and sites that connect to exchanges. Sites will have the choice to connect to either type of aggregation entity.

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.

```
        ______________                  ______________
    --+/              \+--------------+/              \+----------
      (       P1       )    +----+    (       P3       )  +----+
      +\______________/     |    |----+\______________/+--|    |--
      |                  +--| X1 |                       +| X2 |
      | ______________  /   |    |-+    ______________  / |    |--
      +/              \+    +-+--+  \  /              \+  +----+
      (       P2       )     / \     +(      P4        )
    --+\______________/     /   \      \______________/
           |               /     \           |      |
           |              /       |          |      |
           |             /        |          |      |
          _|_          _/_       _|_        _|_    _|_
         /   \        /   \     /   \      /   \  /   \
        ( S.A )      ( S.B )   ( P5  )    ( P6  )( S.D )
         \___/        \___/     \___/      \___/  \___/
                                  |          / \
                                 _|_       _/_  \   ___
                                /   \     /   \  +-/   \
                               ( S.E )   ( S.F )  ( S.G )
                                \___/     \___/    \___/
```

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 [EXCH] (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:

```
| 3 |  13 |    32     |   16   |          64 bits               |
+---+-----+-----------+--------+--------------------------------+
|FP | TLA |   NLA*    |  SLA*  |         Interface ID           |
+---+-----+-----------+--------+--------------------------------+


<--Public Topology--->   Site
                      <-------->
                       Topology
                               <------Interface Identifier----->
```

Where

```
FP           Format Prefix (001)
TLA          Top-Level Aggregator
NLA*         Next-Level Aggregator(s)
SLA*         Site-Level Aggregator(s)
INTERFACE ID Interface Identifier
```

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

3.2 Top-Level Aggregator

Top-Level Aggregators (TLA) are the top level in the routing hierarchy. Default-free routers must have a routing table entry for every active TLA. They may have additional entries, 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's. Additional TLA may be added by using this format for additional format prefixes. The addition of another FP will add another 8,192 TLA's.

3.2.1 Assignment of TLAs

TLAs are assigned to organizations providing public transit topology. They are specifically not assigned to organizations only providing leaf or private transit topology. TLA assignment does not imply ownership. It does imply stewardship over valuable Internet property.

The IAB and IESG have authorized the Internet Assigned Numbers Authority (IANA) as the appropriate entity to have the responsibility for the management of the IPv6 address space as defined in [ALLOC].

The IANA will assign small blocks of TLAs to IPv6 registries. The registries will assign the TLAs to organizations meeting the requirements for TLAs. When the registries have assigned all of their TLAs they can request that the IANA give them another block. The blocks do not have to be contiguous. The IANA may also assign TLAs to organizations directly.

Organizations assigned TLAs are required to meet the following requirements:

- Must have a plan to offer public native IPv6 service within 6 months from assignment. Plan must include plan for NLA allocation.

- Plan or track record providing public internet transit service on fair, reasonable, and non-discriminatory terms, to other providers. TLAs must not be assigned to organizations that are only providing leaf service even if multihomed.

- Must provide registry services on fair, reasonable, and non-discriminatory terms, for the NLA address space it is responsible for under its TLA. This must include both sites and next level providers.

- Must provide transit routing and forwarding to all assigned TLAs on fair, reasonable, and non-discriminatory terms. Organizations are not allowed to filter out any specific TLA's (except temporarily for diagnostic purposes or emergency repair purposed).

- Periodically (interval set by registry) provide to registry utilization statistics of the TLA it has custody of. The organization must also show evidence of carrying TLA routing and transit traffic. This can be in the form of traffic statistics, traceroutes, routing table dumps, or similar means.

Organizations which are given custody of a TLA and fail to continue to meet these may have the TLA custody revoked.

3.3 Next-Level Aggregator(s)

Next-Level Aggregator(s) are used by TLA's to create an addressing hierarchy and to identify sites. The TLA can assign the top part of the NLA 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:

```
      |  n  |      32-n bits      |   16   |     64 bits     |
      +-----+---------------------+--------+-----------------+
      |NLA1 |       Site          |  SLA*  | Interface ID    |
      +-----+---------------------+--------+-----------------+
```

Each TLA receives 32 bits of NLA* space. This NLA* space allows each TLA to provide service to about as many organizations as the current IPv4 internet can support total nodes.

The TLAs may also support NLAs in their own Site ID space. This allows the TLAs to provide service to organizations providing public transit service and organizations who do not. The organizations providing public transit service become NLA's themselves. These NLAs may also choose to use their Site ID space to support other NLAs. This is shown as follows:

```
      |  n  |      32-n bits      |   16   |     64 bits     |
      +-----+---------------------+--------+-----------------+
      |NLA1 |       Site          |  SLA*  | Interface ID    |
      +-----+---------------------+--------+-----------------+

            |  m  |    32-n-m     |   16   |     64 bits     |
            +-----+---------------+--------+-----------------+
            |NLA2 |    Site       |  SLA*  | Interface ID    |
            +-----+---------------+--------+-----------------+

                  |  o  |32-n-m-o|   16   |     64 bits     |
                  +-----+---------+--------+-----------------+
                  |NLA3 |  Site   |  SLA*  | Interface ID    |
                  +-----+---------+--------+-----------------+
```

The NLA delegation works in the same manner as CIDR delegation in IPv4 [CIDR]. TLAs are required to assume registry duties for the NLAs. Each level of NLA is required to assume registry duties for the next level NLA.

The design of the bit layout of the NLA space for a specific TLA is left to the organization responsible for that TLA. Likewise the design of the bit layout of the next level NLA is the responsibility of the previous level NLA. It is recommended that organizations assigning NLA address space use "slow start" allocation procedures as is currently done with IPV4 CIDR blocks.

The design of an NLA 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 assignment provides for easier allocation and attachment flexibility but results in larger routing tables.

3.4 Site-Level Aggregator(s)

The SLA* 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* field support 65,535 individual subnets.

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

```
      |  n  |   16-n     |              64 bits                |
      +-----+------------+-------------------------------------+
      |SLA1 |   Subnet   |            Interface ID             |
      +-----+------------+-------------------------------------+

            | m  |16-n-m |              64 bits                |
            +----+-------+-------------------------------------+
            |SLA2|Subnet |            Interface ID             |
            +----+-------+-------------------------------------+
```

The approach chosen for how to the structure of an SLA* field is the responsibility of the individual organization.

The number of subnets supported 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.

3.5 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 interface's identifier will be the same as that 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 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, and John Stewart. for their review and constructive comments.

5.0 References

[ALLOC] IAB and IESG, "IPv6 Address Allocation Management", RFC1881, December 1995.

[ARCH] Hinden, R., "IP Version 6 Addressing Architecture", Internet Draft, <draft-ietf-ipngwg-addr-arch-00.txt>, May 1997.

[AUTO] Thompson, S., Narten T., "IPv6 Stateless Address Autoconfiguration", RFC1971, August 1996.

[CIDR] Fuller, V., T. Li, K. Varadhan, J. Yu, "Supernetting: an Address Assignment and Aggregation Strategy", RFC1338.

[ETHER] Crawford, M., "Transmission of IPv6 Packets over Ethernet Networks", Internet Draft, <draft-ietf-ipngwg-trans-ethernet-00.txt>, March 1997.

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

[EXCH] Hinden, R., Huitema, C. "Internet Exchanges", document under preparation.

[FDDI] Crawford, M., "Transmission of IPv6 Packets over FDDI Networks", Internet Draft, <draft-ietf-ipngwg-trans-fddi-net-00.txt>, March 1997.

[IPV6] Deering, S., Hinden, R., Editors, "Internet Protocol, Version 6 (IPv6) Specification", RFC1883, December 1995.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", RFC2119, BCP14, March 1997.

6.0 Security Considerations

Documents of this type do not directly impact the security of the Internet infrastructure or its applications.

7.0 Authors' Addresses

Robert M. Hinden
Ipsilon Networks, Inc.
232 Java Drive
Sunnyvale, CA 94089
USA
phone: 1 408 990-2004
email: hinden@ipsilon.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