IPNGWG Working Group
Internet Draft
draft-ietf-ipngwg-scoping-arch-00.txt
March 2000
Expires September 2000

S. Deerin
Cisco System
B. Haber
Nortel Network
B. Zil
Microsof

IP Version 6 Scoped Address Architecture

Status of this Memo

Abstract

This document specifies the architectural characteristics, expected behavior, and usage of IPv6 addresses of different scopes

1. Introduction

The Internet Protocol version 6 (IPv6) introduces the concept of limited scope addresses to the IP lexicon. While operational practice with IPv4 has included the concept of a private address space (net 10, etc.), the design of IPv6 incorporates such addresses into its base architecture. This document defines terms associated with such addresses and describes mechanisms for their behavior.

2. Definitions

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].

3. Basic Terminology

The terms link, interface, node, host, and router are defined in [RFC 2460]. The definitions of unicast address scopes (link-local, site-local, and global) and multicast address scopes (node-local, link-local, etc.) are contained in [RFC 2373].

4. Address Scope

Every IPv6 address has a specific scope, that is, a topological "distance" within which the address may be used as a unique identifier for an interface. The scope of an address is encoded as part of the address, as specified in [RFC 2373].

For unicast addresses, there are three defined scopes:

- Link-local scope, for uniquely identifying interfaces within a single link only.
- Site-local scope, for uniquely identifying interfaces within a single site only. A "site" is, by intent, not rigorously defined, but is typically expected to cover a region of topology that belongs to a single organization and is located within a single geographic location, such as an office, an office complex, or a campus. A personal

residence may be treated as a site (for example, when the residence obtains Internet access via a public Internet service provider), or as a part of a site (for example, when the residence obtains Internet access via an employer's or school's site).

- Global scope, for uniquely identifying interfaces anywhere in the Internet.

For multicast addresses, there are fourteen possible scopes, ranging from node-local to global (including both link-local and site-local). A node-local multicast address serves as a unique identifier for an interface within a single node only; such an address is used only for "loopback" delivery of multicasts within a single node, for example, as a form of inter-process communication within a computer.

There is an ordering relationship among scopes:

- for unicast scopes, link-local is a smaller scope than site-local, and site-local is smaller scope than global.
- for multicast scopes, scopes with lesser values in the "scop" subfield of the multicast address [RFC 2373, section 2.7] are smaller than scopes with greater values, with node-local being the smallest and global being the largest.

However, two scopes of different size may cover the exact same region of topology, for example, a site may consist of a single link, in which

both link-local and site-local scope effectively cover the same topological "distance".

5. Scope Zones

A scope zone, or a simply a zone, is a connected region of topology of a given scope. For example, the set of links connected by routers within a particular site, and the interfaces attached to those links, comprise a single zone of site-local scope. To understand the distinction between scopes and zones, observe that the topological regions within two different sites are considered to be two DIFFERENT zones, but of the SAME scope.

Addresses of a given (non-global) scope may be re-used in different zones of that scope. The zone to which a particular non-global address pertains is not encoded in the address itself, but rather is determined by context, such as the interface from which it is sent or received.

Zones of the different scopes are defined as follows:

- A node-local zone (for multicast only) consists of a single interface on a node. [Note: node-local scope would have been more accurately named interface-local.]
- A link-local zone (for unicast and multicast) consists of a single link and all the interfaces attached to that link.
- There is a single zone of global scope (for both unicast and multicast), comprising all the links and interfaces in the Internet.
- The boundaries of zones of scope other than node-local, link-local, and global must be defined and configured by network administrators. The only required such boundaries are site boundaries. A site boundary serves for both unicast and multicast.

Zone boundaries are relatively static features, not changing in response to short-term changes in topology. Thus, the requirement that the topology within a zone be "connected" is intended to include links and interfaces that may be only occasionally connected. For example, a residential node or network that obtains Internet access by dial-up to an employer's site may be treated as part of the employer's site-local zone even when the dial-up link is disconnected. Similarly, a failure of a router, interface, or link that causes a zone to become partitioned does not split that zone into multiple zones; rather, the different partitions are still considered to belong to the same zone.

Zones have the following additional properties:

- Zone boundaries cut through nodes, not links. (There are two exceptions: the global zone has no boundary, and the boundary of a node-local zone conceptually cuts through an interface between a node and a link.)
- Zones of the same scope cannot overlap, i.e., they can have no links or interfaces in common.
- A zone of a given scope (less than global) falls completely within zones of larger scope, i.e., a smaller scope zone cannot include more topology than any larger scope zone with which it shares any links or interfaces.

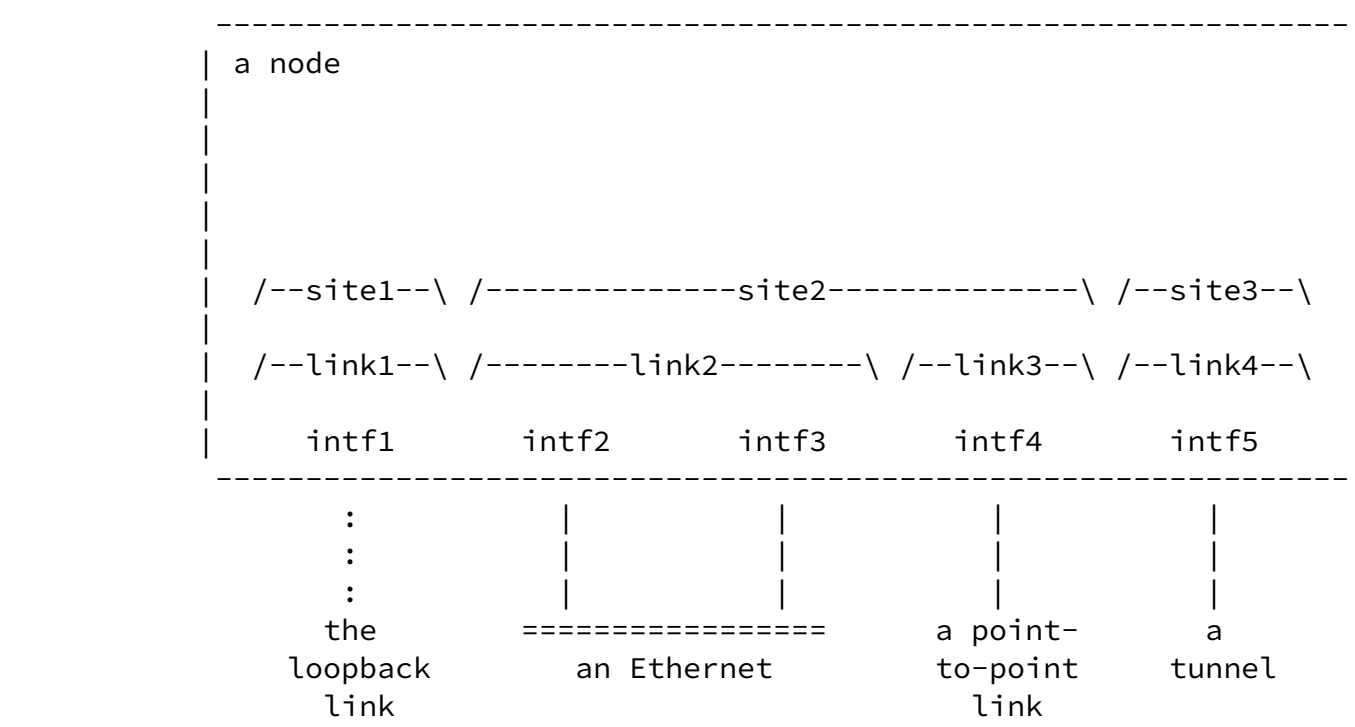
Each interface belongs to one node-local zone, one link-local zone, one site-local zone, and the global zone. Each link belongs to one link-

local zone, one site-local zone, and the global zone. An interface or link only belongs to additional (i.e., multicast) zones if it falls within the configured boundaries of such additional zones.

6. Zone Indexes

Because the same address of a given (non-global) scope can be re-used in different zones of that scope, a node must have a means _- other than examining the address itself _- of associating non-global addresses with particular zones when sending, receiving, or forwarding packets containing such addresses. This is accomplished by assigning a local "zone index" to each zone to which a node is attached. Each attached zone of the same scope must be assigned a different index value; attached zones of different scopes can re-use the same index values.

The assignment of zone indexes is illustrated in the example in the figure below:



This example node has five interfaces:

- A loopback interface, which can be thought of as an interface to a phantom link _- the "loopback link" _- that goes nowhere,
- Two interfaces to the same Ethernet,
- An interface to a point-to-point link, and
- A tunnel interface (e.g., the abstract endpoint of an IPv6-overIPv6 tunnel [TUNNEL], presumably established over either the Ethernet or the point-to-point link.)

It is thus attached to five node-local zones, identified by the interface indexes 1 through 5.

Because the two Ethernet interfaces are attached to the same link, the node is attached to only four link-local zones, identified by link indexes 1 through 4.

It is attached to three site-local zones: one imaginary one to which the loopback interface belongs, one to which the Ethernet and the point-to-point link belong, and one to which the tunnel belongs (perhaps because it is a tunnel to another organization). These site-local zones are identified by the site indexes 1 through 3.

The zone indexes are strictly local to the node. For example, the node on the other end of the point-to-point link may well be using entirely different interface, link, and site index values for that link.

The zone index values are arbitrary. An implementation may use any value it chooses to label a zone so long as it maintains the requirement that the index value of each attached zone of the same scope must be unique within the node. Implementations choosing to follow the recommended basic API [BASICAPI] will also want to restrict their index values to those that can be represented by the sin6_scope_id field of a sockaddr_in6.

An implementation may also support the concept of a "default" zone for each scope. It is convenient to reserve the index value zero, at each scope, to mean "use the default zone". This default index can also be used to identify the zone for any scopes for which the node has not assigned any indexes, such as the various multicast-only scopes.

There is at present no way for a node to automatically determine which of its interfaces belong to the same zones, e.g., the same link or the same site. In the future, protocols may be developed to determine that information. In the absence of such protocols, an implementation must provide a means for manual assignment and/or reassignment of zone indexes. Furthermore, to avoid the need to perform manual configuration in most cases, an implementation should, by default, initially assign zone indexes as follows:

- o A unique interface index for each interface
- o A unique link index for each interface
- o A single site index for all interfaces

Then, manual configuration would be necessary only for the less common cases of nodes with multiple interfaces to a single link, interfaces to different sites, or interfaces to zones of different (multicast-only) scopes.

7. Sending Packets

When an upper-layer protocol sends a packet to a non-global destination address, the node must also identify the intended zone to be used for transmission.

Note that there is one exception to the above statement: when sending to the IPv6 unicast loopback address, ::1, there is no need to identify the intended zone, even though that address is non-global. Conceptually, the unicast loopback address is a link-local address for a node's loopback interface, and is never assigned to any other

interface. Therefore, it unambiguously identifies a single zone of link-scope, that being the phantom loopback link.

Although identification of an outgoing interface is sufficient to identify an intended zone (because each interface is attached to no more than one zone of each scope), that is more specific than desired in many cases. For example, when sending to a site-local unicast address, from host that has more than one interface to the intended site, the upper layer protocol may not care which of those interfaces is used for the transmission, but rather would prefer to leave that choice to the routing function in the IP layer. Thus, the upper-layer requires the ability to specify a zone index, rather than an interface index, when sending to a non-global, non-loopback destination address.

There may also be cases where the upper-layer wishes to restrict the choice of outgoing interface to those belonging to a zone of smaller scope than the destination address. For example, when sending to a site-local destination, the upper-layer may wish to specify a specific link on which the packet should be transmitted, but leave the choice of which specific interface to use on that link to the IP layer. One possible reason for such behavior is that the source address chosen by

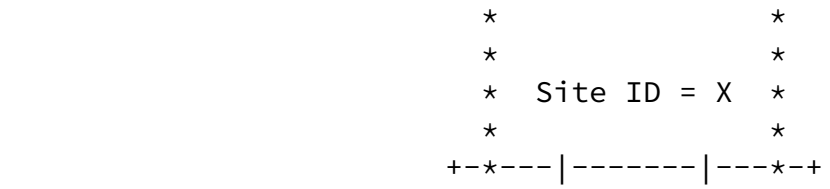
the upper-layer is of smaller scope than the destination, e.g., when using a link-local source address and a site-local destination address. Thus, the upper layer requires the ability, when sending a packet, to specify any zone of scope less than or equal to the scope of the destination address, including the case in which the destination address is of global scope. For this reason, an implementation might find it useful to assign a distinct value for each zone index, so that they are unique across all zones, regardless of scope.

8. Receiving Packets

When an upper-layer protocol receives a packet containing a non-global source or destination address, the zone to which that address pertains can be determined from the arrival interface, because the arrival interface can attached to only one zone of the same scope as the address under consideration.

9. Forwarding Rules and Routing

A single zone router is defined as a router configured with the same zone index on all interfaces. A zone boundary router is defined as a router that has at least 2 distinct zone indices of the same scope.



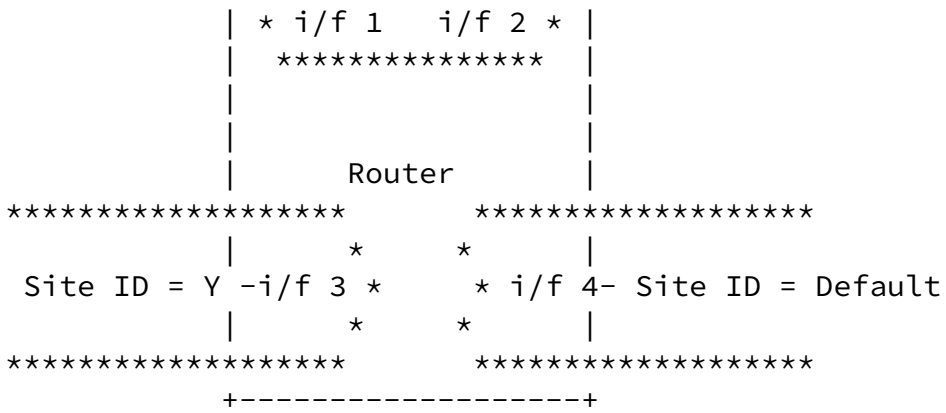


Figure 1: Multi-Sited Router

9.1 Single Zone Routing Protocols

In a single zone router, a routing protocol can advertise all addresses and prefixes, except the link-local prefixes, on all interfaces. This configuration does not require any special handling for scoped addresses. The reception and transmission of scoped addresses is handled in the same manner as global addresses. This applies to both unicast and multicast routing protocols.

9.2 Zone Boundary Unicast Routing

With respect to zone boundaries, routers must consider which interfaces a packet can be transmitted on as well as control the propagation of routing information specific to the zone. This includes controlling which prefixes can be advertised on an interface.

9.2.1 Routing Protocols

When a routing protocol determines that it is a zone boundary router, it must perform additional work in order to protect inter-zone integrity and still maintain intra zone connectivity.

In order to maintain connectivity, the routing protocol must be able to create forwarding information for the global prefixes as well as for all of the zone prefixes for each of its attached sites. The most straightforward way of doing this is to create up to (n+1) forwarding tables; one for the global prefixes, if any, and one for each of the (n) zones.

To protect inter zone integrity; routers must be selective in the forwarding information that is shared with neighboring routers. Routing protocols routinely transmit their routing information to its

neighboring routers. When a router is transmitting this routing information, it must not include any information about zones other than the zones defined on the interface used to reach a neighbor.

As an example, the router in Figure 1 must advertise routing information on four interfaces. The information advertised is as follows:

- Interface 1
  - All global prefixes
  - All site prefixes learned from Interfaces 1 and 2
- Interface 2
  - All global prefixes
  - All site prefixes learned from Interfaces 1 and 2
- Interface 3
  - All global prefixes
  - All site prefixes learned from Interface 3
- Interface 4
  - All global prefixes
  - No site prefixes

By imposing advertisement rules, zone integrity is maintained by keeping all zone routing information contained within the zone.

9.2.2 Packet Forwarding

In addition to the extra cost of generating additional forwarding information for each zone, boundary routers must also do some additional checking when forwarding packets that contain non-global scoped addresses.

If a packet being forwarded contains a non-global destination address, regardless of the scope of the source address, the router must perform the following:

- Lookup incoming interface's zone index
- Perform route lookup for destination address in arrival interface's zone scoped routing table

If a packet being forwarded contains a non-global source address and a global scoped destination address, the following must be performed:

- Lookup outgoing interface's zone index
- Compare inbound and outbound interfaces' zone indices

If the zone indices match, the packet can be forwarded. If they do not match, an ICMPv6 destination unreachable message must be sent to the sender with a code value, code = 2 (beyond scope of source address).

Note that the above procedure applies for addresses of all scopes, including link-local. Thus, if a router receives a packet with a link-local destination address that is not one of the router's own link-local addresses on the arrival link, the router is expected to try and forward the packet to the destination on that link (subject to successful determination of the destination's link-layer address via the Neighbor Discovery protocol [ND]). The forwarded packet may be transmitted back out the arrival interface or out any other interface attached to the same link.

9.3 Scoped Multicast Routing

With IPv6 multicast, there are multiple scopes supported. Multicast routers must be able to control the propagation of scoped packets based on administratively configured boundaries.

9.3.1 Routing Protocols

Multicast routing protocols must follow the same rules as the unicast protocols. They will be required to maintain information about global prefixes as well as information about all scope boundaries that exist on the router.

Multicast protocols that rely on underlying unicast protocols for route exchange (i.e. PIM, MOSPF) will not suffer as much of a performance impact since the unicast protocol will handle the forwarding table generation. They must be able to handle the additional scope boundaries used in multicast addresses.

Multicast protocols that generate and maintain their own routing tables will have to perform the additional route calculations for scope boundaries. All multicast protocols will be forced to handle fourteen additional scooping identifiers above the site identifiers supported in

IPv6 unicast addresses.

9.3.2 Packet Forwarding

The following combinations describe the forwarding rules for multicast:

- Global multicast destination / Global unicast source
- Global multicast destination / Non-global unicast source
- Non-global multicast destination / Global unicast source
- Non-global multicast destination / Non-global unicast source

The first combination requires no special processing over what is currently in place for global IPv6 multicast. The remaining combinations should result in the router performing the same zone index check as outlined for the non-global unicast addresses

9.4 Routing Headers

A node that receives a packet addressed to itself and containing a Routing Header with more than zero Segments Left [RFC 2460, section 4.4] swaps the original destination address with the next address in the Routing Header. Then the above forwarding rules are applied, using the new destination address. An implementation need not, indeed MUST NOT, examine additional addresses in the Routing header to determine whether they are crossing boundaries for their scopes. Thus, it is possible, though generally inadvisable, to use a Routing Header to convey a non-global address across its associated zone boundary.

10. Related Documents

The following list is a set of documents that are related to scoped IPv6 addresses:

- Site Prefixes in Neighbor Discovery, draft-ietf-ipngwg-site-prefixes-03.txt
- An Extension of Format for IPv6 Scoped Addresses, draft-ietf-ipngwg-scopedaddr-format-00.txt
- Default Address Selection for IPv6, draft-ietf-ipngwg-default-addr-select-00.txt

11. Mobility

TBD

12. Security Considerations

The routing section of this document specifies a set of guidelines that allow routers to prevent zone-specific information from leaking out of each site. If site boundary routers allow site routing information to be forwarded outside of the site, the integrity of the site could be compromise

13. References

[RFC 2119] S. Bradner, "Key words for use in RFCs to Indicate Requirement Levels", RFC 2119, BCP14, March 1999.

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

[RFC 2460] Deering, S., and Hinden, R., "Internet Protocol Version 6 (IPv6) Specification", RFC 2460, December 1998.

[TUNNEL] Conta, A., and Deering, S., "Generic Packet Tunneling in IPv6 Specification", RFC 2473, December 1998.

[ICMPv6] Conta, A., and Deering, S., "Internet Control Message Protocol (ICMPv6) for Internet Protocol Version 6 (IPv6)", RFC 2463, December 1998.

[ND] Narten, T., Nordmark, E., and Simpson, W., "Neighbor Discovery for IP Version 6 (IPv6)", RFC 2461, December 1998.

[BASICAPI]

Acknowledgements

Authors' Addresses

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

Phone: +1-408-527-8213
Fax: +1-408-527-8213
Email: deering@cisco.com

Brian Haberman
Nortel Networks
4309 Emperor Blvd.
Suite 200
Durham, NC 27703
USA

Phone: +1-919-992-4439
Email: haberman@nortelnetworks.com

Brian D. Zill
Microsoft Research
One Microsoft Way
Redmond, WA 98052-6399
USA

Phone: +1-425-703-3568
Fax: +1-425-936-7329
Email: bzill@microsoft.com