ï»¿

Enterprise Multihoming Using Provider-Assigned IPv6 Addresses without
        Network Prefix Translation: Requirements and Solutions

Abstract

   Connecting an enterprise site to multiple ISPs over IPv6 using
   provider-assigned addresses is difficult without the use of some form
   of Network Address Translation (NAT). Much has been written on this
   topic over the last 10 to 15 years, but it still remains a problem
   without a clearly defined or widely implemented solution. Any
   multihoming solution without NAT requires hosts at the site to have
   addresses from each ISP and to select the egress ISP by selecting a
   source address for outgoing packets. It also requires routers at the
   site to take into account those source addresses when forwarding
   packets out towards the ISPs.

   This document examines currently available mechanisms for providing a
   solution to this problem for a broad range of enterprise topologies.
   It covers the behavior of routers to forward traffic by taking into
   account source address, and it covers the behavior of hosts to select
   appropriate default source addresses. It also covers any possible
   role that routers might play in providing information to hosts to
   help them select appropriate source addresses. In the process of
   exploring potential solutions, this document also makes explicit
   requirements for how the solution would be expected to behave from
   the perspective of an enterprise site network administrator.

Status of This Memo

   This document is not an Internet Standards Track specification; it is
   published for informational purposes.

   This document is a product of the Internet Engineering Task Force
   (IETF). It represents the consensus of the IETF community. It has
   received public review and has been approved for publication by the
   Internet Engineering Steering Group (IESG). Not all documents
   approved by the IESG are candidates for any level of Internet
   Standard; see Section 2 of RFC 7841.

   Information about the current status of this document, any errata,
   and how to provide feedback on it may be obtained at
   https://www.rfc-editor.org/info/rfc8678.

Table of Contents

1. Introduction

     Site multihoming, the connection of a subscriber network to multiple
     upstream networks using redundant uplinks, is a common enterprise
     architecture for improving the reliability of its Internet
     connectivity. If the site uses provider-independent (PI) addresses,
     all traffic originating from the enterprise can use source addresses
     from the PI address space. Site multihoming with PI addresses is
     commonly used with both IPv4 and IPv6, and does not present any new
     technical challenges.

     It may be desirable for an enterprise site to connect to multiple
     ISPs using provider-assigned (PA) addresses instead of PI addresses.
     Multihoming with provider-assigned addresses is typically less
     expensive for the enterprise relative to using provider-independent
     addresses as it does not require obtaining and maintaining PI address
     space nor does it require running BGP between the enterprise and the
     ISPs (for small/medium networks, running BGP might be not only
     undesirable but also impossible, especially if residential-type ISP
     connections are used). PA multihoming is also a practice that should
     be facilitated and encouraged because it does not add to the size of
     the Internet routing table, whereas PI multihoming does. Note that
     PA is also used to mean "provider-aggregatable". In this document,
     we assume that provider-assigned addresses are always provider-
     aggregatable.

     With PA multihoming, for each ISP connection, the site is assigned a
     prefix from within an address block allocated to that ISP by its
     National or Regional Internet Registry. In the simple case of two
     ISPs (ISP-A and ISP-B), the site will have two different prefixes
     assigned to it (prefix-A and prefix-B). This arrangement is
     problematic. First, packets with the "wrong" source address may be
     dropped by one of the ISPs. In order to limit denial-of-service
     attacks using spoofed source addresses, [BCP38] recommends that ISPs
     filter traffic from customer sites to allow only traffic with a
     source address that has been assigned by that ISP. So a packet sent
     from a multihomed site on the uplink to ISP-B with a source address
     in prefix-A may be dropped by ISP-B.

     However, even if ISP-B does not implement BCP 38, or ISP-B adds
     prefix-A to its list of allowed source addresses on the uplink from
     the multihomed site, two-way communication may still fail. If the
     packet with a source address in prefix-A was sent to ISP-B because
     the uplink to ISP-A failed, then if ISP-B does not drop the packet,
     and the packet reaches its destination somewhere on the Internet, the
     return packet will be sent back with a destination address in prefix-
     A. The return packet will be routed over the Internet to ISP-A, but
     it will not be delivered to the multihomed site because the site
     uplink with ISP-A has failed. Two-way communication would require
     some arrangement for ISP-B to advertise prefix-A when the uplink to
     ISP-A fails.

     Note that the same may be true of a provider that does not implement
     BCP 38, even if his upstream provider does, or of a provider that has
     no corresponding route to deliver the ingress traffic to the
     multihomed site. The issue is not that the immediate provider
     implements ingress filtering; it is that someone upstream does (so
     egress traffic is blocked) or lacks a route (causing blackholing of
     the ingress traffic).

     Another issue with asymmetric traffic flow (when the egress traffic
     leaves the site via one ISP, but the return traffic enters the site
     via another uplink) is related to stateful firewalls/middleboxes.
     Keeping state in that case might be problematic, even impossible.

     With IPv4, this problem is commonly solved by using private address

space described in [RFC1918] within the multihomed site and Network
Address Translation (NAT) or Network Address/Port Translation (NAPT)
[RFC2663] on the uplinks to the ISPs. However, one of the goals of
IPv6 is to eliminate the need for and the use of NAT or NAPT.
Therefore, requiring the use of NAT or NAPT for an enterprise site to
multihome with provider-assigned addresses is not an attractive
solution.

[RFC6296] describes a translation solution specifically tailored to
meet the requirements of multihoming with provider-assigned IPv6
addresses. With the IPv6-to-IPv6 Network Prefix Translation (NPTv6)
solution, an enterprise can use either Unique Local Addresses
[RFC4193] or the prefix assigned by one of the ISPs within its site.
As traffic leaves the site on an uplink to an ISP, the source address
is translated in a predictable and reversible manner to an address
within the prefix assigned by the ISP on that uplink. [RFC6296] is
currently classified as Experimental, and it has been implemented by
several vendors. See Section 8.2 for more discussion of NPTv6.

This document defines routing requirements for enterprise
multihoming. This document focuses on the following general class of
solutions.

Each host at the enterprise has multiple addresses, at least one from
each ISP-assigned prefix. As discussed in Section 6.1 and in
[RFC6724], each host is responsible for choosing the source address
that is applied to each packet it sends. A host is expected to be
able to respond dynamically to the failure of an uplink to a given
ISP by no longer sending packets with the source address
corresponding to that ISP. Potential mechanisms for communicating
network changes to the host are Neighbor Discovery Router
Advertisements [RFC4861], DHCPv6 [RFC8415], and ICMPv6 [RFC4443].

The routers in the enterprise network are responsible for ensuring
that packets are delivered to the "correct" ISP uplink based on
source address. This requires that at least some routers in the site
network are able to take into account the source address of a packet
when deciding how to route it. That is, some routers must be capable
of some form of Source Address Dependent Routing (SADR), if only as
described in Section 4.3 of [RFC3704]. At a minimum, the routers
connected to the ISP uplinks (the site exit routers or SERs) must be
capable of Source Address Dependent Routing. Expanding the connected
domain of routers capable of SADR from the site exit routers deeper
into the site network will generally result in more efficient routing
of traffic with external destinations.

This document is organized as follows. Section 4 looks in more
detail at the enterprise networking environments in which this
solution is expected to operate. The discussion of Section 4 uses
the concepts of Source-Prefix-Scoped Routing advertisements and
forwarding tables and describes how Source-Prefix-Scoped Routing
advertisements are used to generate source-prefix-scoped forwarding
tables. A detailed description of generating source-prefix-scoped
forwarding tables is provided in Section 5. Section 6 discusses
existing and proposed mechanisms for hosts to select the default
source address to be used by applications. It also discusses the
requirements for routing that are needed to support these enterprise
network scenarios and the mechanisms by which hosts are expected to
update default source addresses based on network state. Section 7
discusses deployment considerations, while Section 8 discusses other
solutions.

2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
"SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and
"OPTIONAL" in this document are to be interpreted as described in
BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all
capitals, as shown here.

3. Terminology

PA (provider-assigned or provider-aggregatable) address space: a
   block of IP addresses assigned by a Regional Internet Registry
   (RIR) to a Local Internet Registry (LIR), used to create
   allocations to end sites. Can be aggregated and present in the
   routing table as one route.

PI (provider-independent) address space: a block of IP addresses
   assigned by a Regional Internet Registry (RIR) directly to end
   site / end customer.

ISP: Internet Service Provider

LIR (Local Internet Registry): an organization (usually an ISP or an
   enterprise/academic) that receives its allocation of IP addresses
   from its Regional Internet Registry, then assigns parts of that
   allocation to its customers.

RIR (Regional Internet Registry): an organization that manages the
   Internet number resources (such as IP addresses and autonomous
   system (AS) numbers) within a geographical region of the world.

SADR (Source Address Dependent Routing): routing that takes into
   account the source address of a packet in addition to the packet
   destination address.

SADR domain: a routing domain in which some (or all) routers
   exchange source-dependent routing information.

Source-Prefix-Scoped Routing/Forwarding Table: a routing (or
   forwarding) table that contains routing (or forwarding)
   information only applicable to packets with source addresses from
   the specific prefix.

Unscoped Routing/Forwarding Table: a routing (or forwarding) table
   that can be used to route/forward packets with any source address.

SER (Site Edge Router): a router that connects the site to an ISP
   (terminates an ISP uplink).

LLA (Link-Local Address): an IPv6 unicast address from the fe80::/10
   prefix [RFC4291].

ULA (Unique Local IPv6 Unicast Address): an IPv6 unicast address
   from the FC00::/7 prefix. They are globally unique and intended
   for local communications [RFC4193].

GUA (Global Unicast Address): a globally routable IPv6 address of
   the global scope [RFC4291].

SLAAC (IPv6 Stateless Address Autoconfiguration): a stateless
   process of configuring the network stack on IPv6 hosts [RFC4862].

RA (Router Advertisement): a message sent by an IPv6 router to
   advertise its presence to hosts together with various network-
   related parameters required for hosts to perform SLAAC [RFC4861].

PIO (Prefix Information Option): a part of an RA message containing
   information about IPv6 prefixes that could be used by hosts to
   generate global IPv6 addresses [RFC4862].

RIO (Route Information Option): a part of an RA message containing
   information about more specific IPv6 prefixes reachable via the
   advertising router [RFC4191].

4. Enterprise Multihoming Use Cases

4.1. Simple ISP Connectivity with Connected SERs

   We start by looking at a scenario in which a site has connections to
   two ISPs, as shown in Figure 1. The site is assigned the prefix

2001:db8:0:a000::/52 by ISP-A and prefix 2001:db8:0:b000::/52 by ISP-B. We consider three hosts in the site. H31 and H32 are on a LAN that has been assigned subnets 2001:db8:0:a010::/64 and 2001:db8:0:b010::/64. H31 has been assigned the addresses 2001:db8:0:a010::31 and 2001:db8:0:b010::31. H32 has been assigned 2001:db8:0:a010::32 and 2001:db8:0:b010::32. H41 is on a different subnet that has been assigned 2001:db8:0:a020::/64 and 2001:db8:0:b020::/64.

```
                                  2001:db8:0:1234::101   H101
                                                          |
                                                          |
2001:db8:0:a010::31                            --------
2001:db8:0:b010::31                     ,-----.      /        \
        +--+   +--+       +----+  ,'       `.   :           :
     +---|R1|---|R4|---+---|SERa|-+   ISP-A    +--+--        :
  H31--+  +--+   +--+   |   +----+  `.       ,'  :           :
        |                            `-----'    : Internet :
        |                                        :          :
        |                                        :          :
        |                                        :          :
        |                           ,-----.      :          :
  H32--+   +--+       |       +----+  ,'       `.  :          :
     +---|R2|---------+---|SERb|-+   ISP-B     +--+--        :
        +--+          |   +----+  `.       ,'  :           :
                      |            `-----'    :           :
                      |                        :           :
        +--+  +--+  +--+                        \         /
  H41------|R3|--|R5|--|R6|                      --------
        +--+  +--+  +--+
```

2001:db8:0:a020::41
2001:db8:0:b020::41
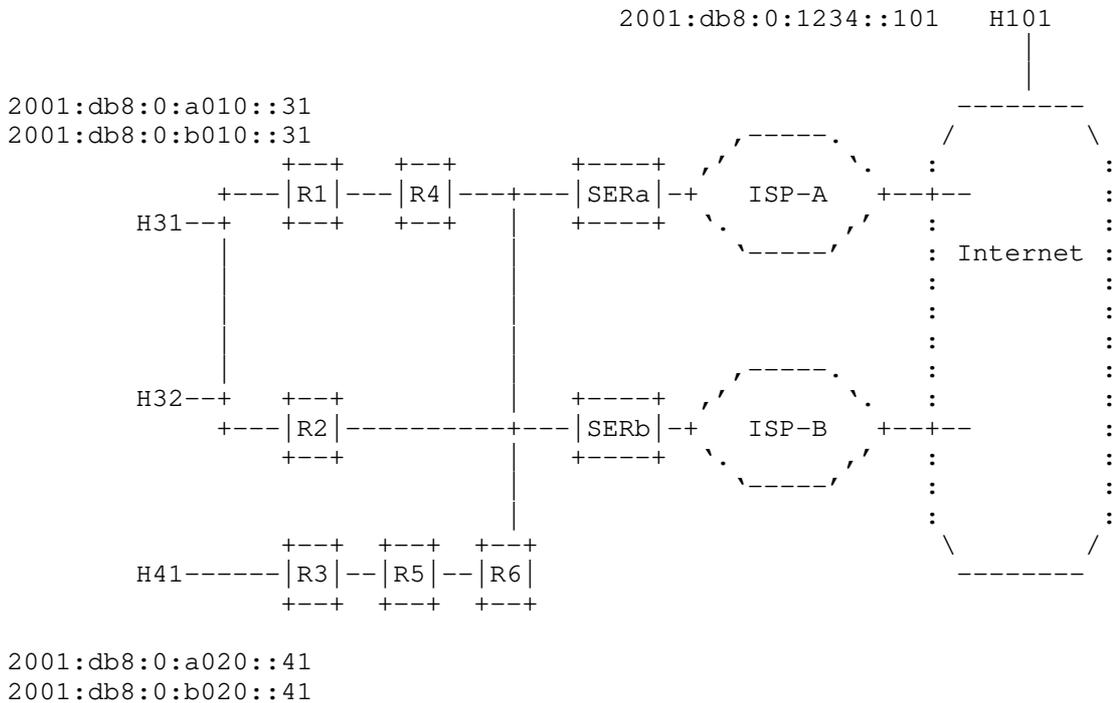
        Figure 1: Simple ISP Connectivity with Connected SERs

We refer to a router that connects the site to an ISP as a site edge router (SER). Several other routers provide connectivity among the internal hosts (H31, H32, and H41), as well as connect the internal hosts to the Internet through SERa and SERb. In this example, SERa and SERb share a direct connection to each other. In Section 4.2, we consider a scenario in which this is not the case.

For the moment, we assume that the hosts are able to select suitable source addresses through some mechanism that doesn't involve the routers in the site network. Here, we focus on the primary task of the routed site network, which is to get packets efficiently to their destinations, while sending a packet to the ISP that assigned the prefix that matches the source address of the packet. In Section 6, we examine what role the routed network may play in helping hosts select suitable source addresses for packets.

With this solution, routers will need some form of Source Address Dependent Routing, which will be new functionality. It would be useful if an enterprise site does not need to upgrade all routers to support the new SADR functionality in order to support PA multihoming. We consider whether this is possible and examine the trade-offs of not having all routers in the site support SADR functionality.

In the topology in Figure 1, it is possible to support PA multihoming with only SERa and SERb being capable of SADR. The other routers can continue to forward based only on destination address, and exchange routes that only consider destination address. In this scenario, SERa and SERb communicate source-scoped routing information across their shared connection. When SERa receives a packet with a source address matching prefix 2001:db8:0:b000::/52, it forwards the packet to SERb, which forwards it on the uplink to ISP-B. The analogous behavior holds for traffic that SERb receives with a source address matching prefix 2001:db8:0:a000::/52.

In Figure 1, when only SERa and SERb are capable of source address dependent routing, PA multihoming will work. However, the paths over which the packets are sent will generally not be the shortest paths. The forwarding paths will generally be more efficient when more routers are capable of SADR. For example, if R4, R2, and R6 are upgraded to support SADR, then they can exchange source-scoped routes with SERa and SERb. They will then know to send traffic with a source address matching prefix 2001:db8:0:b000::/52 directly to SERb, without sending it to SERa first.

4.2. Simple ISP Connectivity Where SERs Are Not Directly Connected

In Figure 2, we modify the topology slightly by inserting R7, so that SERa and SERb are no longer directly connected. With this topology, it is not enough to just enable SADR routing on SERa and SERb to support PA multihoming. There are two solutions to enable PA multihoming in this topology.

```
                                  2001:db8:0:1234::101    H101
                                              |
                                              |
  2001:db8:0:a010::31                                ---------
  2001:db8:0:b010::31        +----+   ,-----.   /        \
           +--+   +--+      |    |  ,'       `. :          :
           +---|R1|---|R4|---+---|SERa|-+  ISP-A   +--+--   :          :
  H31--+   +--+   +--+   |      +----+  `.      ,'  :          :
       |              |                 `-----'    : Internet :
       |             +--+                          :          :
       |             |R7|                          :          :
       |             +--+                          :          :
       |              |             ,-----.        :          :
  H32--+   +--+      |       +----+ ,'       `. :          :
           +---|R2|-----------+---|SERb|-+  ISP-B   +--+--   :          :
           +--+           |      +----+  `.      ,'  :          :
                          |                 `-----'    :          :
                          |                            :          :
           +--+ +--+ +--+                              \        /
  H41------|R3|--|R5|--|R6|                             ---------
           +--+ +--+ +--+                                  |
                                                           |
  2001:db8:0:a020::41           2001:db8:0:5678::501    H501
  2001:db8:0:b020::41
```
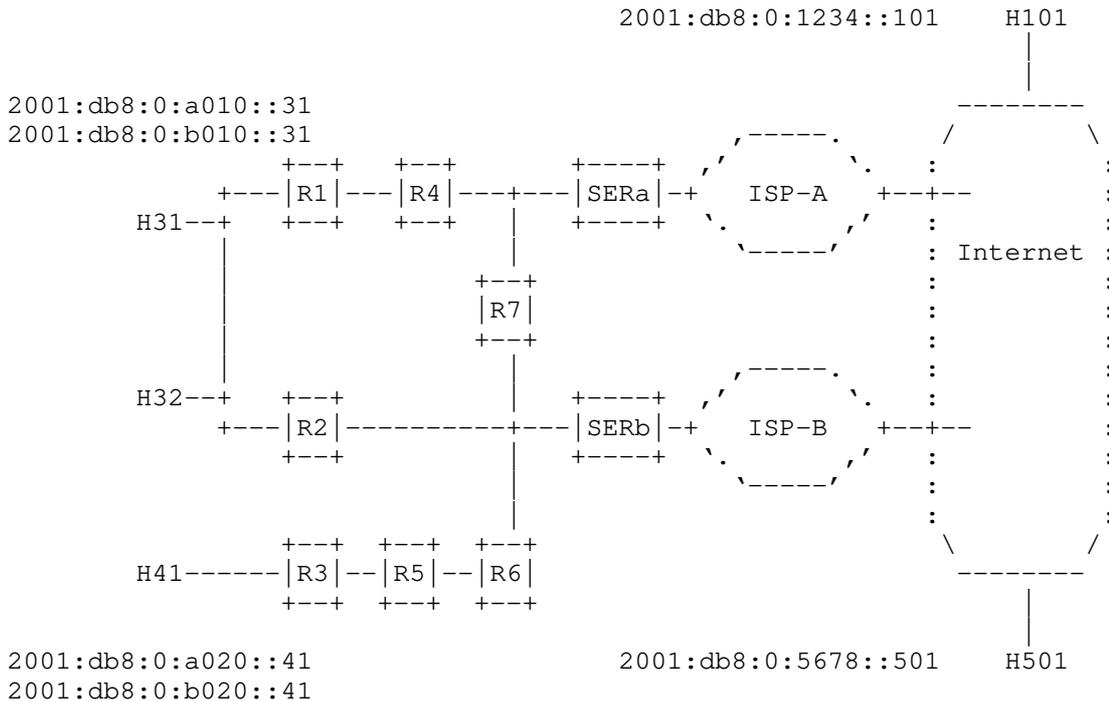
              Figure 2: Simple ISP Connectivity Where SERs Are Not Directly
                                  Connected

One option is to effectively modify the topology by creating a logical tunnel between SERa and SERb by using Generic Routing Encapsulation (GRE) [RFC7676], for example. Although SERa and SERb are not directly connected physically in this topology, they can be directly connected logically by a tunnel.

The other option is to enable SADR functionality on R7. In this way, R7 will exchange source-scoped routes with SERa and SERb, making the three routers act as a single SADR domain. This illustrates the basic principle that the minimum requirement for the routed site network to support PA multihoming is having all of the site exit routers be part of a connected SADR domain. Extending the connected SADR domain beyond that point can produce more efficient forwarding paths.

4.3. Enterprise Network Operator Expectations

Before considering a more complex scenario, let's look in more detail at the reasonably simple multihoming scenario in Figure 2 to understand what can reasonably be expected from this solution. As a general guiding principle, we assume an enterprise network operator will expect a multihomed network to behave as close to a single-homed network as possible. So a solution that meets those expectations where possible is a good thing.

For traffic between internal hosts and for traffic from outside the site to internal hosts, an enterprise network operator would expect there to be no visible change in the path taken by this traffic, since this traffic does not need to be routed in a way that depends on source address. It is also reasonable to expect that internal hosts should be able to communicate with each other using either of their source addresses without restriction. For example, H31 should be able to communicate with H41 using a packet with S=2001:db8:0:a010::31, D=2001:db8:0:b020::41, regardless of the state of uplink to ISP-B.

These goals can be accomplished by having all of the routers in the network continue to originate normal unscoped destination routes for their connected networks. If we can arrange it so that these unscoped destination routes are used for forwarding this traffic, then we will have accomplished multihoming's goal of not affecting the forwarding of traffic destined for internal hosts.

For traffic destined for external hosts, it is reasonable to expect that traffic with a source address from the prefix assigned by ISP-A to follow the path that the traffic would follow if there were no connection to ISP-B. This can be accomplished by having SERa originate a source-scoped route of the form (S=2001:db8:0:a000::/52, D=::/0). If all of the routers in the site support SADR, then the path of traffic exiting via ISP-A can match that expectation. If some routers don't support SADR, then it is reasonable to expect that the path for traffic exiting via ISP-A may be different within the site. This is a trade-off that the enterprise network operator may decide to make.

It is important to understand the behavior of this multihoming solution when an uplink to one of the ISPs fails. To simplify this discussion, we assume that all routers in the site support SADR. We start by looking at the operation of the network when the uplinks to both ISP-A and ISP-B are functioning properly. SERa originates a source-scoped route of the form (S=2001:db8:0:a000::/52, D=::/0), and SERb originates a source-scoped route of the form (S=2001:db8:0:b000::/52, D=::/0). These routes are distributed through the routers in the site, and they establish within the routers two sets of forwarding paths for traffic leaving the site. One set of forwarding paths is for packets with source addresses in 2001:db8:0:a000::/52. The other set of forwarding paths is for packets with source addresses in 2001:db8:0:b000::/52. The normal destination routes, which are not scoped to these two source prefixes, play no role in the forwarding. Whether a packet exits the site via SERa or via SERb is completely determined by the source address applied to the packet by the host. So for example, when host H31 sends a packet to host H101 with (S=2001:db8:0:a010::31, D=2001:db8:0:1234::101), the packet will only be sent out the link from SERa to ISP-A.

Now consider what happens when the uplink from SERa to ISP-A fails. The only way for the packets from H31 to reach H101 is for H31 to start using the source address for ISP-B. H31 needs to send the following packet: (S=2001:db8:0:b010::31, D=2001:db8:0:1234::101).

This behavior is very different from the behavior that occurs with site multihoming using PI addresses or with PA addresses using NAT. In these other multihoming solutions, hosts do not need to react to network failures several hops away in order to regain Internet access. Instead, a host can be largely unaware of the failure of an uplink to an ISP. When multihoming with PA addresses and NAT, existing sessions generally need to be reestablished after a failure since the external host will receive packets from the internal host with a new source address. However, new sessions can be established without any action on the part of the hosts. Multihoming with PA addresses and NAT has created the expectation of a fairly quick and simple recovery from network failures. Alternatives should to be evaluated in terms of the speed and complexity of the recovery mechanism.

Another significant difference between this multihoming solution and
multihoming with either PI addresses or with PA addresses using NAT
is the ability of the enterprise network operator to route traffic
over different ISPs based on destination address. We still consider
the fairly simple network of Figure 2 and assume that uplinks to both
ISPs are functioning. Assume that the site is multihomed using PA
addresses and NAT, and that SERa and SERb each originate a normal
destination route for D=::/0, with the route origination dependent on
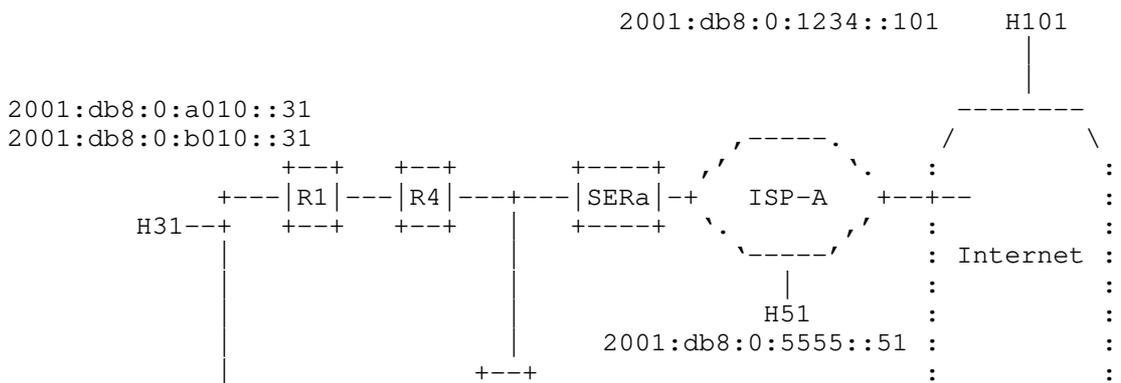the state of the uplink to the respective ISP.

Now suppose it is observed that an important application running
between internal hosts and external host H101 experiences much better
performance when the traffic passes through ISP-A (perhaps because
ISP-A provides lower latency to H101). When multihoming this site
with PI addresses or with PA addresses and NAT, the enterprise
network operator can configure SERa to originate into the site
network a normal destination route for D=2001:db8:0:1234::/64 (the
destination prefix to reach H101) that depends on the state of the
uplink to ISP-A. When the link to ISP-A is functioning, the
destination route D=2001:db8:0:1234::/64 will be originated by SERa,
so traffic from all hosts will use ISP-A to reach H101 based on the
longest destination prefix match in the route lookup.

Implementing the same routing policy is more difficult with the PA
multihoming solution described in this document since it doesn't use
NAT. By design, the only way to control where a packet exits this
network is by setting the source address of the packet. Since the
network cannot modify the source address without NAT, the host must
set it. To implement this routing policy, each host needs to use the
source address from the prefix assigned by ISP-A to send traffic
destined for H101. Mechanisms have been proposed to allow hosts to
choose the source address for packets in a fine-grained manner. We
will discuss these proposals in Section 6. However, an enterprise
network administrator would not expect to interact with host
operating systems to ensure that a particular source address is
chosen for a particular destination prefix in order to implement this
routing policy.

4.4. More Complex ISP Connectivity

The previous sections considered two variations of a simple
multihoming scenario in which the site is connected to two ISPs
offering only Internet connectivity. It is likely that many actual
enterprise multihoming scenarios will be similar to this simple
example. However, there are more complex multihoming scenarios that
we would like this solution to address as well.

It is fairly common for an ISP to offer a service in addition to
Internet access over the same uplink. Two variations of this are
reflected in Figure 3. In addition to Internet access, ISP-A offers
a service that requires the site to access host H51 at
2001:db8:0:5555::51. The site has a single physical and logical
connection with ISP-A, and ISP-A only allows access to H51 over that
connection. So when H32 needs to access the service at H51, it needs
to send packets with (S=2001:db8:0:a010::32, D=2001:db8:0:5555::51),
and those packets need to be forwarded out the link from SERa to ISP-
A.

```
                                 2001:db8:0:1234::101    H101

                                                           |

2001:db8:0:a010::31                                     --------
2001:db8:0:b010::31                       ,-----.      /        \
        +--+    +--+        +----+ ,'       `. :            :
        +---|R1|---|R4|---+---|SERa|-+  ISP-A    +--+--      :
    H31--+  +--+    +--+   |   +----+ `.       ,'  :            :
          |                |           `-----'    : Internet :
          |                |              |        :          :
          |                |             H51       :          :
          |                |      2001:db8:0:5555::51 :        :
          |            +--+                          :          :
```

```
            |                 |R7|                          :            :
            |                 +--+                          :            :
            |                  |                            :            :
            |                  |              ,-----.`      :            :
   H32--+   +--+-----+----+--+-----+       ,’         `.    :            :
       +---|R2|-----+----+--|SERb1|-+     ISP-B     +--+--  :            :
           +--+     |       +-----+   `.        ,’   :            :
                   +--+                  `--|--’      :            :
2001:db8:0:a010::32        |R8|                 |              \        /
                          +--+               ,--|--.          --------
            |         +-----+             ,’        `.           |
            +-------|SERb2|-+          ISP-B    |       |
            |         +-----+         `.       ,’       H501
            |                           `-----’   2001:db8:0:5678
            |                              |               ::501
  +--+   +--+                           H61
H41------|R3|--|R5|           2001:db8:0:6666::61
  +--+   +--+
```
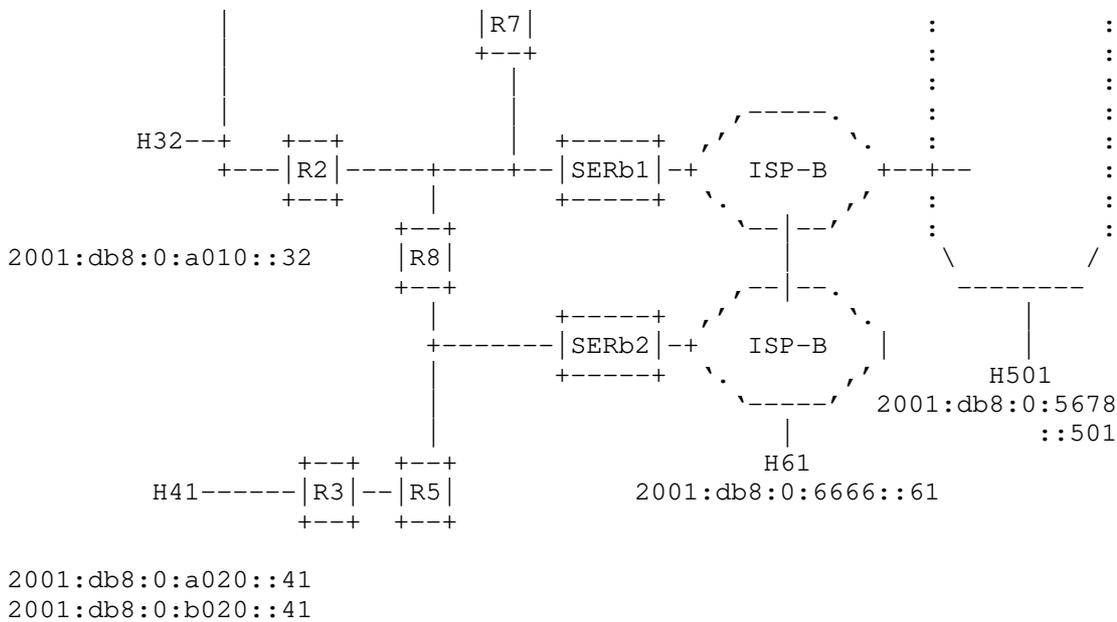
2001:db8:0:a020::41
2001:db8:0:b020::41

Figure 3: Internet Access and Services Offered by ISP-A and ISP-B

ISP-B illustrates a variation on this scenario. In addition to
Internet access, ISP-B also offers a service that requires the site
to access host H61. The site has two connections to two different
parts of ISP-B (shown as SERb1 and SERb2 in Figure 3). ISP-B expects
Internet traffic to use the uplink from SERb1, while it expects
traffic destined for the service at H61 to use the uplink from SERb2.
For either uplink, ISP-B expects the ingress traffic to have a source
address matching the prefix that it assigned to the site,
2001:db8:0:b000::/52.

As discussed before, we rely completely on the internal host to set
the source address of the packet properly. In the case of a packet
sent by H31 to access the service in ISP-B at H61, we expect the
packet to have the following addresses: (S=2001:db8:0:b010::31,
D=2001:db8:0:6666::61). The routed network has two potential ways of
distributing routes so that this packet exits the site on the uplink
at SERb2.

We could just rely on normal destination routes, without using
source-prefix-scoped routes. If we have SERb2 originate a normal
unscoped destination route for D=2001:db8:0:6666::/64, the packets
from H31 to H61 will exit the site at SERb2 as desired. We should
not have to worry about SERa needing to originate the same route,
because ISP-B should choose a globally unique prefix for the service
at H61.

The alternative is to have SERb2 originate a source-prefix-scoped
destination route of the form (S=2001:db8:0:b000::/52,
D=2001:db8:0:6666::/64). From a forwarding point of view, the use of
the source-prefix-scoped destination route would result in traffic
with source addresses corresponding only to ISP-B being sent to
SERb2. Instead, the use of the unscoped destination route would
result in traffic with source addresses corresponding to ISP-A and
ISP-B being sent to SERb2, as long as the destination address matches
the destination prefix. It seems like either forwarding behavior
would be acceptable.

However, from the point of view of the enterprise network
administrator trying to configure, maintain, and troubleshoot this
multihoming solution, it seems much clearer to have SERb2 originate
the source-prefix-scoped destination route corresponding to the
service offered by ISP-B. In this way, all of the traffic leaving
the site is determined by the source-prefix-scoped routes, and all of
the traffic within the site or arriving from external hosts is
determined by the unscoped destination routes. Therefore, for this
multihoming solution we choose to originate source-prefix-scoped
routes for all traffic leaving the site.

4.5. ISPs and Provider-Assigned Prefixes

While we expect that most site multihoming involves connecting to
only two ISPs, this solution allows for connections to an arbitrary
number of ISPs. However, when evaluating scalable implementations of
the solution, it would be reasonable to assume that the maximum
number of ISPs that a site would connect to is five (topologies with
two redundant routers, each having two uplinks to different ISPs,
plus a tunnel to a head office acting as fifth one are not unheard
of).

It is also useful to note that the prefixes assigned to the site by
different ISPs will not overlap. This must be the case, since the
provider-assigned addresses have to be globally unique.

4.6. Simplified Topologies

The topologies of many enterprise sites using this multihoming
solution may in practice be simpler than the examples that we have
used. The topology in Figure 1 could be further simplified by
directly connecting all hosts to the LAN that connects the two site
exit routers, SERa and SERb. The topology could also be simplified
by connecting both uplinks to ISP-A and ISP-B to the same site exit
router. However, it is the aim of this document to provide a
solution that applies to a broad range of enterprise site network
topologies, so this document focuses on providing a solution to the
more general case. The simplified cases will also be supported by
this solution, and there may even be optimizations that can be made
for simplified cases. This solution, however, needs to support more
complex topologies.

We are starting with the basic assumption that enterprise site
networks can be quite complex from a routing perspective. However,
even a complex site network can be multihomed to different ISPs with
PA addresses using IPv4 and NAT. It is not reasonable to expect an
enterprise network operator to change the routing topology of the
site in order to deploy IPv6.

5. Generating Source-Prefix-Scoped Forwarding Tables

So far, we have described in general terms how the SADR-capable
routers in this solution forward traffic by using both normal
unscoped destination routes and source-prefix-scoped destination
routes. Here we give a precise method for generating a source-
prefix-scoped forwarding table on a router that supports SADR.

1. Compute the next-hops for the source-prefix-scoped destination
    prefixes using only routers in the connected SADR domain. These
    are the initial source-prefix-scoped forwarding table entries.

2. Compute the next-hops for the unscoped destination prefixes using
    all routers in the IGP. This is the unscoped forwarding table.

3. For a given source-prefix-scoped forwarding table T (scoped to
    source prefix P), consider a source-prefix-scoped forwarding
    table T', whose source prefix P' contains P. We call T the more
    specific source-prefix-scoped forwarding table, and T' the less
    specific source-prefix-scoped forwarding table. We select
    entries in forwarding table T' to augment forwarding table T
    based on the following rules. If a destination prefix of an
    entry in forwarding table T' exactly matches the destination
    prefix of an existing entry in forwarding table T (including
    destination prefix length), then do not add the entry to
    forwarding table T. If the destination prefix does NOT match an
    existing entry, then add the entry to forwarding table T. As the
    unscoped forwarding table is considered to be scoped to ::/0,
    this process will propagate routes from the unscoped forwarding
    table to forwarding table T. If there exist multiple source-
    prefix-scoped forwarding tables whose source prefixes contain P,
    these source-prefix-scoped forwarding tables should be processed

in order from most specific to least specific.

The forwarding tables produced by this process are used in the
following way to forward packets.

1. Select the most specific (longest prefix match) source-prefix-
    scoped forwarding table that matches the source address of the
    packet (again, the unscoped forwarding table is considered to be
    scoped to ::/0).

2. Look up the destination address of the packet in the selected
    forwarding table to determine the next-hop for the packet.

The following example illustrates how this process is used to create
a forwarding table for each provider-assigned source prefix. We
consider the multihomed site network in Figure 3. Initially we
assume that all of the routers in the site network support SADR.
Figure 4 shows the routes that are originated by the routers in the
site network.

Routes originated by SERa:
(S=2001:db8:0:a000::/52, D=2001:db8:0:5555/64)
(S=2001:db8:0:a000::/52, D=::/0)
(D=2001:db8:0:5555::/64)
(D=::/0)

Routes originated by SERb1:
(S=2001:db8:0:b000::/52, D=::/0)
(D=::/0)

Routes originated by SERb2:
(S=2001:db8:0:b000::/52, D=2001:db8:0:6666::/64)
(D=2001:db8:0:6666::/64)

Routes originated by R1:
(D=2001:db8:0:a010::/64)
(D=2001:db8:0:b010::/64)

Routes originated by R2:
(D=2001:db8:0:a010::/64)
(D=2001:db8:0:b010::/64)

Routes originated by R3:
(D=2001:db8:0:a020::/64)
(D=2001:db8:0:b020::/64)

        Figure 4: Routes Originated by Routers in the Site Network

Each SER originates destination routes that are scoped to the source
prefix assigned by the ISP to which the SER connects. Note that the
SERs also originate the corresponding unscoped destination route.
This is not needed when all of the routers in the site support SADR.
However, it is required when some routers do not support SADR. This
will be discussed in more detail later.

We focus on how R8 constructs its source-prefix-scoped forwarding
tables from these route advertisements. R8 computes the next hops
for destination routes that are scoped to the source prefix
2001:db8:0:a000::/52. The results are shown in the first table in
Figure 5. (In this example, the next hops are computed assuming that
all links have the same metric.) Then, R8 computes the next hops for
destination routes that are scoped to the source prefix
2001:db8:0:b000::/52. The results are shown in the second table in
Figure 5. Finally, R8 computes the next hops for the unscoped
destination prefixes. The results are shown in the third table in
Figure 5.

forwarding entries scoped to
source prefix = 2001:db8:0:a000::/52
==
D=2001:db8:0:5555/64   NH=R7

```
D=::/0                          NH=R7


forwarding entries scoped to
source prefix = 2001:db8:0:b000::/52
==========================================
D=2001:db8:0:6666/64       NH=SERb2
D=::/0                      NH=SERb1


unscoped forwarding entries
==========================================
D=2001:db8:0:a010::/64     NH=R2
D=2001:db8:0:b010::/64     NH=R2
D=2001:db8:0:a020::/64     NH=R5
D=2001:db8:0:b020::/64     NH=R5
D=2001:db8:0:5555::/64     NH=R7
D=2001:db8:0:6666::/64     NH=SERb2
D=::/0                     NH=SERb1
```

Figure 5: Forwarding Entries Computed at R8

The final step is for R8 to augment the more specific source-prefix-
scoped forwarding tables with entries from less specific source-
prefix-scoped forwarding tables. The unscoped forwarding table is
considered as being scoped to ::/0, so both 2001:db8:0:a000::/52 and
2001:db8:0:b000::/52 are more specific prefixes of ::/0. Therefore,
entries in the unscoped forwarding table will be evaluated to be
added to these two more specific source-prefix-scoped forwarding
tables. If a forwarding entry from the less specific source-prefix-
scoped forwarding table has the exact same destination prefix
(including destination prefix length) as the forwarding entry from
the more specific source-prefix-scoped forwarding table, then the
existing forwarding entry in the more specific source-prefix-scoped
forwarding table wins.

As an example of how the source-prefix-scoped forwarding entries are
augmented, we consider how the two entries in the first table in
Figure 5 (the table for source prefix = 2001:db8:0:a000::/52) are
augmented with entries from the third table in Figure 5 (the table of
unscoped or scoped to ::/0 forwarding entries). The first four
unscoped forwarding entries (D=2001:db8:0:a010::/64,
D=2001:db8:0:b010::/64, D=2001:db8:0:a020::/64, and
D=2001:db8:0:b020::/64) are not an exact match for any of the
existing entries in the forwarding table for source prefix
2001:db8:0:a000::/52. Therefore, these four entries are added to the
final forwarding table for source prefix 2001:db8:0:a000::/52. The
result of adding these entries is reflected in the first four entries
the first table in Figure 6.

The next less specific source-prefix-scoped (scope is ::/0)
forwarding table entry is for D=2001:db8:0:5555::/64. This entry is
an exact match for the existing entry in the forwarding table for the
more specific source prefix 2001:db8:0:a000::/52. Therefore, we do
not replace the existing entry with the entry from the unscoped
forwarding table. This is reflected in the fifth entry in the first
table in Figure 6. (Note that since both scoped and unscoped entries
have R7 as the next hop, the result of applying this rule is not
visible.)

The next less specific source-prefix-scoped (scope is ::/0)
forwarding table entry is for D=2001:db8:0:6666::/64. This entry is
not an exact match for any existing entries in the forwarding table
for source prefix 2001:db8:0:a000::/52. Therefore, we add this
entry. This is reflected in the sixth entry in the first table in
Figure 6.

The next less specific source-prefix-scoped (scope is ::/0)
forwarding table entry is for D=::/0. This entry is an exact match
for the existing entry in the forwarding table for the more specific
source prefix 2001:db8:0:a000::/52. Therefore, we do not overwrite
the existing source-prefix-scoped entry, as can be seen in the last
entry in the first table in Figure 6.

```
if source address matches 2001:db8:0:a000::/52
then use this forwarding table
========================================
D=2001:db8:0:a010::/64    NH=R2
D=2001:db8:0:b010::/64    NH=R2
D=2001:db8:0:a020::/64    NH=R5
D=2001:db8:0:b020::/64    NH=R5
D=2001:db8:0:5555::/64    NH=R7
D=2001:db8:0:6666::/64    NH=SERb2
D=::/0                    NH=R7

else if source address matches 2001:db8:0:b000::/52
then use this forwarding table
========================================
D=2001:db8:0:a010::/64    NH=R2
D=2001:db8:0:b010::/64    NH=R2
D=2001:db8:0:a020::/64    NH=R5
D=2001:db8:0:b020::/64    NH=R5
D=2001:db8:0:5555::/64    NH=R7
D=2001:db8:0:6666::/64    NH=SERb2
D=::/0                    NH=SERb1

else if source address matches ::/0 use this forwarding table
========================================
D=2001:db8:0:a010::/64    NH=R2
D=2001:db8:0:b010::/64    NH=R2
D=2001:db8:0:a020::/64    NH=R5
D=2001:db8:0:b020::/64    NH=R5
D=2001:db8:0:5555::/64    NH=R7
D=2001:db8:0:6666::/64    NH=SERb2
D=::/0                    NH=SERb1
```

Figure 6: Complete Forwarding Tables Computed at R8

The forwarding tables produced by this process at R8 have the desired
properties. A packet with a source address in 2001:db8:0:a000::/52
will be forwarded based on the first table in Figure 6. If the
packet is destined for the Internet at large or the service at
D=2001:db8:0:5555/64, it will be sent to R7 in the direction of SERa.
If the packet is destined for an internal host, then the first four
entries will send it to R2 or R5 as expected. Note that if this
packet has a destination address corresponding to the service offered
by ISP-B (D=2001:db8:0:5555::/64), then it will get forwarded to
SERb2. It will be dropped by SERb2 or by ISP-B, since the packet has
a source address that was not assigned by ISP-B. However, this is
expected behavior. In order to use the service offered by ISP-B, the
host needs to originate the packet with a source address assigned by
ISP-B.

In this example, a packet with a source address that doesn't match
2001:db8:0:a000::/52 or 2001:db8:0:b000::/52 must have originated
from an external host. Such a packet will use the unscoped
forwarding table (the last table in Figure 6). These packets will
flow exactly as they would in absence of multihoming.

We can also modify this example to illustrate how it supports
deployments in which not all routers in the site support SADR.
Continuing with the topology shown in Figure 3, suppose that R3 and
R5 do not support SADR. Instead they are only capable of
understanding unscoped route advertisements. The SADR routers in the
network will still originate the routes shown in Figure 4. However,
R3 and R5 will only understand the unscoped routes as shown in
Figure 7.

Routes originated by SERa:
(D=2001:db8:0:5555::/64)
(D=::/0)

Routes originated by SERb1:
(D=::/0)
```

```
Routes originated by SERb2:
(D=2001:db8:0:6666::/64)


Routes originated by R1:
(D=2001:db8:0:a010::/64)
(D=2001:db8:0:b010::/64)


Routes originated by R2:
(D=2001:db8:0:a010::/64)
(D=2001:db8:0:b010::/64)


Routes originated by R3:
(D=2001:db8:0:a020::/64)
(D=2001:db8:0:b020::/64)
```

        Figure 7: Route Advertisements Understood by Routers That Do Not
                                Support SADR

   With these unscoped route advertisements, R5 will produce the
   forwarding table shown in Figure 8.

```
forwarding table
==========================================
D=2001:db8:0:a010::/64     NH=R8
D=2001:db8:0:b010::/64     NH=R8
D=2001:db8:0:a020::/64     NH=R3
D=2001:db8:0:b020::/64     NH=R3
D=2001:db8:0:5555::/64     NH=R8
D=2001:db8:0:6666::/64     NH=SERb2
D=::/0                     NH=R8
```

           Figure 8: Forwarding Table for R5, Which Doesn't Understand
                          Source-Prefix- Scoped Routes

   As all SERs belong to the SADR domain, any traffic that needs to exit
   the site will eventually hit a SADR-capable router. To prevent
   routing loops involving SADR-capable and non-SADR-capable routers,
   traffic that enters the SADR-capable domain does not leave the domain
   until it exits the site. Therefore all SADR-capable routers within
   the domain MUST be logically connected.

   Note that the mechanism described here for converting source-prefix-
   scoped destination prefix routing advertisements into forwarding
   state is somewhat different from that proposed in [DST-SRC-RTG]. The
   method described in this document is functionally equivalent, but it
   is based on application of existing mechanisms for the described
   scenarios.

6. Mechanisms for Hosts To Choose Good Default Source Addresses in a
    Multihomed Site

   Until this point, we have made the assumption that hosts are able to
   choose the correct source address using some unspecified mechanism.
   This has allowed us to focus on what the routers in a multihomed site
   network need to do in order to forward packets to the correct ISP
   based on source address. Now we look at possible mechanisms for
   hosts to choose the correct source address. We also look at what
   role, if any, the routers may play in providing information that
   helps hosts to choose source addresses.

   It should be noted that this section discusses how hosts could select
   the default source address for new connections. Any connection that
   already exists on a host is bound to a specific source address that
   cannot be changed. Section 6.7 discusses the connections
   preservation issue in more detail.

   Any host that needs to be able to send traffic using the uplinks to a
   given ISP is expected to be configured with an address from the
   prefix assigned by that ISP. The host will control which ISP is used
   for its traffic by selecting one of the addresses configured on the

host as the source address for outgoing traffic. It is the
responsibility of the site network to ensure that a packet with the
source address from an ISP is now sent on an uplink to that ISP.

If all of the ISP uplinks are working, then the host's choice of
source address may be driven by the desire to load share across ISP
uplinks, or it may be driven by the desire to take advantage of
certain properties of a particular uplink or ISP (if some information
about various path properties has been made available to the host
somehow. See [PROV-DOMAINS] as an example). If any of the ISP
uplinks is not working, then the choice of source address by the host
can cause packets to get dropped.

How a host selects a suitable source address in a multihomed site is
not a solved problem. We do not attempt to solve this problem in
this document. Instead we discuss the current state of affairs with
respect to standardized solutions and the implementation of those
solutions. We also look at proposed solutions for this problem.

An external host initiating communication with a host internal to a
PA-multihomed site will need to know multiple addresses for that host
in order to communicate with it using different ISPs to the
multihomed site (knowing just one address would undermine all
benefits of redundant connectivity provided by multihoming). These
addresses are typically learned through DNS. (For simplicity, we
assume that the external host is single-homed.) The external host
chooses the ISP that will be used at the remote multihomed site by
setting the destination address on the packets it transmits. For a
session originated from an external host to an internal host, the
choice of source address used by the internal host is simple. The
internal host has no choice but to use the destination address in the
received packet as the source address of the transmitted packet.

For a session originated by a host inside the multihomed site, the
decision of which source address to select is more complicated. We
consider three main methods for hosts to get information about the
network. The two proactive methods are Neighbor Discovery Router
Advertisements and DHCPv6. The one reactive method we consider is
ICMPv6. Note that we are explicitly excluding the possibility of
having hosts participate in, or even listen directly to, routing
protocol advertisements.

First we look at how a host is currently expected to select the
default source and destination addresses to be used for a new
connection.

6.1. Default Source Address Selection Algorithm on Hosts

   [RFC6724] defines the algorithms that hosts are expected to use to
   select source and destination addresses for packets. It defines an
   algorithm for selecting a source address and a separate algorithm for
   selecting a destination address. Both of these algorithms depend on
   a policy table. [RFC6724] defines a default policy that produces
   certain behavior.

   The rules in the two algorithms in [RFC6724] depend on many different
   properties of addresses. While these are needed for understanding
   how a host should choose addresses in an arbitrary environment, most
   of the rules are not relevant for understanding how a host should
   choose among multiple source addresses in a multihomed environment
   when sending a packet to a remote host. Returning to the example in
   Figure 3, we look at what the default algorithms in [RFC6724] say
   about the source address that internal host H31 should use to send
   traffic to external host H101, somewhere on the Internet.

   There is no choice to be made with respect to destination address.
   H31 needs to send a packet with D=2001:db8:0:1234::101 in order to
   reach H101. So H31 has to choose between using S=2001:db8:0:a010::31
   or S=2001:db8:0:b010::31 as the source address for this packet. We
   go through the rules for source address selection in Section 5 of
   [RFC6724].

Rule 1 (Prefer same address) is not useful to break the tie between
source addresses because neither one of the candidate source
addresses equals the destination address.

Rule 2 (Prefer appropriate scope) is also not useful in this scenario
because both source addresses and the destination address have global
scope.

Rule 3 (Avoid deprecated addresses) applies to an address that has
been autoconfigured by a host using stateless address
autoconfiguration as defined in [RFC4862]. An address autoconfigured
by a host has a preferred lifetime and a valid lifetime. The address
is preferred until the preferred lifetime expires, after which it
becomes deprecated. A deprecated address is not used if there is a
preferred address of the appropriate scope available. When the valid
lifetime expires, the address cannot be used at all. The preferred
and valid lifetimes for an autoconfigured address are set based on
the corresponding lifetimes in the Prefix Information Option in
Neighbor Discovery Router Advertisements. In this scenario, a
possible tool to control source address selection in this scenario
would be for a host to deprecate an address by having routers on that
link, R1 and R2 in Figure 3, send Router Advertisement messages
containing PIOs with the Preferred Lifetime value for the deprecated
source prefix set to zero. This is a rather blunt tool, because it
discourages or prohibits the use of that source prefix for all
destinations. However, it may be useful in some scenarios. For
example, if all uplinks to a particular ISP fail, it is desirable to
prevent hosts from using source addresses from that ISP address
space.

Rule 4 (Avoid home addresses) does not apply here because we are not
considering Mobile IP.

Rule 5 (Prefer outgoing interface) is not useful in this scenario
because both source addresses are assigned to the same interface.

Rule 5.5 (Prefer addresses in a prefix advertised by the next-hop) is
not useful in the scenario when both R1 and R2 will advertise both
source prefixes. However, this rule may potentially allow a host to
select the correct source prefix by selecting a next-hop. The most
obvious way would be to make R1 advertise itself as a default router
and send PIO for 2001:db8:0:a010::/64, while R2 advertises itself as
a default router and sends PIO for 2001:db8:0:b010::/64. We'll
discuss later how Rule 5.5 can be used to influence a source address
selection in single-router topologies (e.g., when H41 is sending
traffic using R3 as a default gateway).

Rule 6 (Prefer matching label) refers to the label value determined
for each source and destination prefix as a result of applying the
policy table to the prefix. With the default policy table defined in
Section 2.1 of [RFC6724], Label(2001:db8:0:a010::31) = 5,
Label(2001:db8:0:b010::31) = 5, and Label(2001:db8:0:1234::101) = 5.
So with the default policy, Rule 6 does not break the tie. However,
the algorithms in [RFC6724] are defined in such a way that non-
default address selection policy tables can be used. [RFC7078]
defines a way to distribute a non-default address selection policy
table to hosts using DHCPv6. So even though the application of Rule
6 to this scenario using the default policy table is not useful, Rule
6 may still be a useful tool.

Rule 7 (Prefer temporary addresses) has to do with the technique
described in [RFC4941] to periodically randomize the interface
portion of an IPv6 address that has been generated using stateless
address autoconfiguration. In general, if H31 were using this
technique, it would use it for both source addresses, for example,
creating temporary addresses 2001:db8:0:a010:2839:9938:ab58:830f and
2001:db8:0:b010:4838:f483:8384:3208, in addition to
2001:db8:0:a010::31 and 2001:db8:0:b010::31. So this rule would
prefer the two temporary addresses, but it would not break the tie
between the two source prefixes from ISP-A and ISP-B.

Rule 8 (Use longest matching prefix) dictates that, between two
candidate source addresses, the host selects the one that has longest
common prefix length with the destination address. For example, if
H31 were selecting the source address for sending packets to H101,
this rule would not break the tie between candidate source addresses
2001:db8:0:a101::31 and 2001:db8:0:b101::31 since the common prefix
length with the destination is 48. However, if H31 were selecting
the source address for sending packets to H41 address
2001:db8:0:a020::41, then this rule would result in using
2001:db8:0:a101::31 as a source (2001:db8:0:a101::31 and
2001:db8:0:a020::41 share the common prefix 2001:db8:0:a000::/58,
while for 2001:db8:0:b101::31 and 2001:db8:0:a020::41, the common
prefix is 2001:db8:0:a000::/51). Therefore Rule 8 might be useful
for selecting the correct source address in some, but not all,
scenarios (for example if ISP-B services belong to
2001:db8:0:b000::/59, then H31 would always use 2001:db8:0:b010::31
to access those destinations).

So we can see that of the eight source address selection rules from
[RFC6724], four actually apply to our basic site multihoming
scenario. The rules that are relevant to this scenario are
summarized below.

* Rule 3: Avoid deprecated addresses.

* Rule 5.5: Prefer addresses in a prefix advertised by the next-hop.

* Rule 6: Prefer matching label.

* Rule 8: Prefer longest matching prefix.

The two methods that we discuss for controlling the source address
selection through the four relevant rules above are SLAAC Router
Advertisement messages and DHCPv6.

We also consider a possible role for ICMPv6 for getting traffic-
driven feedback from the network. With the source address selection
algorithm discussed above, the goal is to choose the correct source
address on the first try, before any traffic is sent. However,
another strategy is to choose a source address, send the packet, get
feedback from the network about whether or not the source address is
correct, and try another source address if it is not.

We consider four scenarios in which a host needs to select the
correct source address. In the first scenario, both uplinks are
working. In the second scenario, one uplink has failed. The third
scenario is a situation in which one failed uplink has recovered.
The last scenario is failure of both (all) uplinks.

It should be noted that [RFC6724] only defines the behavior of IPv6
hosts to select default addresses that applications and upper-layer
protocols can use. Applications and upper-layer protocols can make
their own choices on selecting source addresses. The mechanism
proposed in this document attempts to ensure that the subset of
source addresses available for applications and upper-layer protocols
is selected with the up-to-date network state in mind. The rest of
the document discusses various aspects of the default source address
selection defined in [RFC6724], calling it for the sake of brevity
"the source address selection".

6.2. Selecting Default Source Address When Both Uplinks Are Working

   Again we return to the topology in Figure 3. Suppose that the site
   administrator wants to implement a policy by which all hosts need to
   use ISP-A to reach H101 at D=2001:db8:0:1234::101. So for example,
   H31 needs to select S=2001:db8:0:a010::31.

6.2.1. Distributing Default Address Selection Policy Table with DHCPv6

   This policy can be implemented by using DHCPv6 to distribute an

address selection policy table that assigns the same label to
destination addresses that match 2001:db8:0:1234::/64 as it does to
source addresses that match 2001:db8:0:a000::/52. The following two
entries accomplish this.

```
Prefix                  Precedence      Label
2001:db8:0:1234::/64    50              33
2001:db8:0:a000::/52    50              33
```

     Figure 9: Policy Table Entries to Implement a Routing Policy

This requires that the hosts implement [RFC6724], the basic source
and destination address framework, along with [RFC7078], the DHCPv6
extension for distributing a non-default policy table. Note that it
does NOT require that the hosts use DHCPv6 for address assignment.
The hosts could still use stateless address autoconfiguration for
address configuration, while using DHCPv6 only for policy table
distribution (see [RFC8415]). However this method has a number of
disadvantages:

* DHCPv6 support is not a mandatory requirement for IPv6 hosts
   [RFC8504], so this method might not work for all devices.

* Network administrators are required to explicitly configure the
   desired network access policies on DHCPv6 servers. While it might
   be feasible in the scenario of a single multihomed network, such
   approach might have some scalability issues, especially if the
   centralized DHCPv6 solution is deployed to serve a large number of
   multihomed sites.

6.2.2. Controlling Default Source Address Selection with Router
        Advertisements

   Neighbor Discovery currently has two mechanisms to communicate prefix
   information to hosts. The base specification for Neighbor Discovery
   (see [RFC4861]) defines the Prefix Information Option (PIO) in the
   Router Advertisement (RA) message. When a host receives a PIO with
   the A-flag [RFC8425] set, it can use the prefix in the PIO as the
   source prefix from which it assigns itself an IP address using
   stateless address autoconfiguration (SLAAC) procedures described in
   [RFC4862]. In the example of Figure 3, if the site network is using
   SLAAC, we would expect both R1 and R2 to send RA messages with PIOs
   with the A-flag set for both source prefixes 2001:db8:0:a010::/64 and
   2001:db8:0:b010::/64. H31 would then use the SLAAC procedure to
   configure itself with 2001:db8:0:a010::31 and 2001:db8:0:b010::31.

   Whereas a host learns about source prefixes from PIO messages, hosts
   can learn about a destination prefix from a Router Advertisement
   containing a Route Information Option (RIO), as specified in
   [RFC4191]. The destination prefixes in RIOs are intended to allow a
   host to choose the router that it uses as its first hop to reach a
   particular destination prefix.

   As currently standardized, neither PIO nor RIO options contained in
   Neighbor Discovery Router Advertisements can communicate the
   information needed to implement the desired routing policy. PIOs
   communicate source prefixes, and RIOs communicate destination
   prefixes. However, there is currently no standardized way to
   directly associate a particular destination prefix with a particular
   source prefix.

   [SADR-RA] proposes a Source Address Dependent Route Information
   option for Neighbor Discovery Router Advertisements that would
   associate a source prefix with a destination prefix. The details of
   [SADR-RA] might need tweaking to address this use case. However, in
   order to be able to use Neighbor Discovery Router Advertisements to
   implement this routing policy, an extension is needed that would
   allow R1 and R2 to explicitly communicate to H31 an association
   between S=2001:db8:0:a000::/52 and D=2001:db8:0:1234::/64.

   However, Rule 5.5 of the default source address selection algorithm

(discussed in Section 6.1), together with default router preference
(specified in [RFC4191]) and RIO, can be used to influence a source
address selection on a host as described below. Let's look at source
address selection on the host H41. It receives RAs from R3 with PIOs
for 2001:db8:0:a020::/64 and 2001:db8:0:b020::/64. At that point,
all traffic would use the same next-hop (R3 link-local address) so
Rule 5.5 does not apply. Now let's assume that R3 supports SADR and
has two scoped forwarding tables, one scoped to
S=2001:db8:0:a000::/52 and another scoped to S=2001:db8:0:b000::/52.
If R3 generates two different link-local addresses for its interface
facing H41 (one for each scoped forwarding table, LLA_A and LLA_B),
R3 will send two different RAs: one from LLA_A that includes a PIO
for 2001:db8:0:a020::/64, and another from LLA_B that includes a PIO
for 2001:db8:0:b020::/64. Now it is possible to influence H41 source
address selection for destinations that follow the default route by
setting the default router preference in RAs. If it is desired that
H41 reaches H101 (or any destination in the Internet) via ISP-A, then
RAs sent from LLA_A should have the default router preference set to
01 (high priority), while RAs sent from LLA_B should have the
preference set to 11 (low). LLA_A would then be chosen as a next-hop
for H101, and therefore (per Rule 5.5) 2001:db8:0:a020::41 would be
selected as the source address. If, at the same time, it is desired
that H61 is accessible via ISP-B, then R3 should include a RIO for
2001:db8:0:6666::/64 in its RA sent from LLA_B. H41 would choose
LLA_B as a next-hop for all traffic to H61, and then per Rule 5.5,
2001:db8:0:b020::41 would be selected as a source address.

If in the above-mentioned scenario it is desirable that all Internet
traffic leaves the network via ISP-A, and the link to ISP-B is used
for accessing ISP-B services only (not as an ISP-A link backup), then
RAs sent by R3 from LLA_B should have their Router Lifetime values
set to zero and should include RIOs for ISP-B address space. It
would instruct H41 to use LLA_A for all Internet traffic but to use
LLA_B as a next-hop while sending traffic to ISP-B addresses.

The description of the mechanism above assumes SADR support by the
first-hop routers as well as SERs. However, a first-hop router can
still provide a less flexible version of this mechanism even without
implementing SADR. This could be done by providing configuration
knobs on the first-hop router that allow it to generate different
link-local addresses and to send individual RAs for each prefix.

The mechanism described above relies on Rule 5.5 of the default
source address selection algorithm defined in [RFC6724]. [RFC8028]
states that "A host SHOULD select default routers for each prefix it
is assigned an address in." It also recommends that hosts should
implement Rule 5.5. of [RFC6724]. Hosts following the
recommendations specified in [RFC8028] therefore should be able to
benefit from the solution described in this document. No standards
need to be updated in regards to host behavior.

6.2.3. Controlling Default Source Address Selection with ICMPv6

We now discuss how one might use ICMPv6 to implement the routing
policy to send traffic destined for H101 out the uplink to ISP-A,
even when uplinks to both ISPs are working. If H31 started sending
traffic to H101 with S=2001:db8:0:b010::31 and
D=2001:db8:0:1234::101, it would be routed through SER-b1 and out the
uplink to ISP-B. SERb1 could recognize that this traffic is not
following the desired routing policy and react by sending an ICMPv6
message back to H31.

In this example, we could arrange things so that SERb1 drops the
packet with S=2001:db8:0:b010::31 and D=2001:db8:0:1234::101, and
then sends to H31 an ICMPv6 Destination Unreachable message with Code
5 (Source address failed ingress/egress policy). When H31 receives
this packet, it would then be expected to try another source address
to reach the destination. In this example, H31 would then send a
packet with S=2001:db8:0:a010::31 and D=2001:db8:0:1234::101, which
will reach SERa and be forwarded out the uplink to ISP-A.

However, we would also want it to be the case that SERb1 does not
enforce this routing policy when the uplink from SERa to ISP-A has
failed. This could be accomplished by having SERa originate a
source-prefix-scoped route for (S=2001:db8:0:a000::/52,
D=2001:db8:0:1234::/64), and have SERb1 monitor the presence of that
route. If that route is not present (because SERa has stopped
originating it), then SERb1 will not enforce the routing policy, and
it will forward packets with S=2001:db8:0:b010::31 and
D=2001:db8:0:1234::101 out its uplink to ISP-B.

We can also use this source-prefix-scoped route originated by SERa to
communicate the desired routing policy to SERb1. We can define an
EXCLUSIVE flag to be advertised together with the IGP route for
(S=2001:db8:0:a000::/52, D=2001:db8:0:1234::/64). This would allow
SERa to communicate to SERb that SERb should reject traffic for
D=2001:db8:0:1234::/64 and respond with an ICMPv6 Destination
Unreachable Code 5 message, as long as the route for
(S=2001:db8:0:a000::/52, D=2001:db8:0:1234::/64) is present. The
definition of an EXCLUSIVE flag for SADR advertisements in IGPs would
require future standardization work.

Finally, if we are willing to extend ICMPv6 to support this solution,
then we could create a mechanism for SERb1 to tell the host which
source address it should be using to successfully forward packets
that meet the policy. In its current form, when SERb1 sends an
ICMPv6 Destination Unreachable Code 5 message, it is basically
saying, "This source address is wrong. Try another source address."
In the absence of a clear indication which address to try next, the
host will iterate over all addresses assigned to the interface (e.g.,
various privacy addresses), which would lead to significant delays
and degraded user experience. It would be better if the ICMPv6
message could say, "This source address is wrong. Instead use a
source address in S=2001:db8:0:a000::/52".

However, using ICMPv6 for signaling source address information back
to hosts introduces new challenges. Most routers currently have
software or hardware limits on generating ICMP messages. A site
administrator deploying a solution that relies on the SERs generating
ICMP messages could try to improve the performance of SERs for
generating ICMP messages. However, in a large network, it is still
likely that ICMP message generation limits will be reached. As a
result, hosts would not receive ICMPv6 back, which in turn leads to
traffic blackholing and poor user experience. To improve the
scalability of ICMPv6-based signaling, hosts SHOULD cache the
preferred source address (or prefix) for the given destination (which
in turn might cause issues in the case of the corresponding ISP
uplink failure - see Section 6.3). In addition, the same source
prefix SHOULD be used for other destinations in the same /64 as the
original destination address. The source prefix to the destination
mapping SHOULD have a specific lifetime. Expiration of the lifetime
SHOULD trigger the source address selection algorithm again.

Using ICMPv6 Destination Unreachable Messages with Code 5 to
influence source address selection introduces some security
challenges, which are discussed in Section 10.

As currently standardized in [RFC4443], the ICMPv6 Destination
Unreachable Message with Code 5 would allow for the iterative
approach to retransmitting packets using different source addresses.
As currently defined, the ICMPv6 message does not provide a mechanism
to communicate information about which source prefix should be used
for a retransmitted packet. The current document does not define
such a mechanism, but it might be a useful extension to define in a
different document. However, this approach has some security
implications, such as an ability for an attacker to send spoofed
ICMPv6 messages to signal an invalid/unreachable source prefix,
causing a DoS-type attack.

6.2.4. Summary of Methods for Controlling Default Source Address
        Selection to Implement Routing Policy

So to summarize this section, we have looked at three methods for
implementing a simple routing policy where all traffic for a given
destination on the Internet needs to use a particular ISP, even when
the uplinks to both ISPs are working.

The default source address selection policy cannot distinguish
between the source addresses needed to enforce this policy, so a non-
default policy table using associating source and destination
prefixes using label values would need to be installed on each host.
A mechanism exists for DHCPv6 to distribute a non-default policy
table, but such solution would heavily rely on DHCPv6 support by the
host operating system. Moreover, there is no mechanism to translate
desired routing/traffic engineering policies into policy tables on
DHCPv6 servers. Therefore using DHCPv6 for controlling the address
selection policy table is not recommended and SHOULD NOT be used.

At the same time, Router Advertisements provide a reliable mechanism
to influence the source address selection process via PIO, RIO, and
default router preferences. As all those options have been
standardized by the IETF and are supported by various operating
systems, no changes are required on hosts. First-hop routers in the
enterprise network need to be capable of sending different RAs for
different SLAAC prefixes (either based on scoped forwarding tables or
based on preconfigured policies).

SERs can enforce the routing policy by sending ICMPv6 Destination
Unreachable messages with Code 5 (Source address failed ingress/
egress policy) for traffic sent with the wrong source address. The
policy distribution could be automated by defining an EXCLUSIVE flag
for the source-prefix-scoped route, which could then be set on the
SER that originates the route. As ICMPv6 message generation can be
rate limited on routers, it SHOULD NOT be used as the only mechanism
to influence source address selection on hosts. While hosts SHOULD
select the correct source address for a given destination, the
network SHOULD signal any source address issues back to hosts using
ICMPv6 error messages.

6.3. Selecting Default Source Address When One Uplink Has Failed

   Now we discuss whether DHCPv6, Neighbor Discovery Router
   Advertisements, and ICMPv6 can help a host choose the right source
   address when an uplink to one of the ISPs has failed. Again we look
   at the scenario in Figure 3. This time we look at traffic from H31
   destined for external host H501 at D=2001:db8:0:5678::501. We
   initially assume that the uplink from SERa to ISP-A is working and
   that the uplink from SERb1 to ISP-B is working.

   We assume there is no particular routing policy desired, so H31 is
   free to send packets with S=2001:db8:0:a010::31 or
   S=2001:db8:0:b010::31 and have them delivered to H501. For this
   example, we assume that H31 has chosen S=2001:db8:0:b010::31 so that
   the packets exit via SERb to ISP-B. Now we see what happens when the
   link from SERb1 to ISP-B fails. How should H31 learn that it needs
   to start sending packets to H501 with S=2001:db8:0:a010::31 in order
   to start using the uplink to ISP-A? We need to do this in a way that
   doesn't prevent H31 from still sending packets with
   S=2001:db8:0:b010::31 in order to reach H61 at D=2001:db8:0:6666::61.

6.3.1. Controlling Default Source Address Selection with DHCPv6

   For this example, we assume that the site network in Figure 3 has a
   centralized DHCP server and that all routers act as DHCP relay
   agents. We assume that both of the addresses assigned to H31 were
   assigned via DHCP.

   We could try to have the DHCP server monitor the state of the uplink
   from SERb1 to ISP-B in some manner and then tell H31 that it can no
   longer use S=2001:db8:0:b010::31 by setting its valid lifetime to
   zero. The DHCP server could initiate this process by sending a
   Reconfigure message to H31 as described in Section 18.3 of [RFC8415].
   Or the DHCP server can assign addresses with short lifetimes in order

to force clients to renew them often.

This approach would prevent H31 from using S=2001:db8:0:b010::31 to
reach a host on the Internet. However, it would also prevent H31
from using S=2001:db8:0:b010::31 to reach H61 at
D=2001:db8:0:6666::61, which is not desirable.

Another potential approach is to have the DHCP server monitor the
uplink from SERb1 to ISP-B and control the choice of source address
on H31 by updating its address selection policy table via the
mechanism in [RFC7078]. The DHCP server could initiate this process
by sending a Reconfigure message to H31. Note that [RFC8415]
requires that Reconfigure messages use DHCP authentication. DHCP
authentication could be avoided by using short address lifetimes to
force clients to send Renew messages to the server often. If the
host does not obtain its IP addresses from the DHCP server, then it
would need to use the Information Refresh Time option defined in
[RFC8415].

If the following policy table can be installed on H31 after the
failure of the uplink from SERb1, then the desired routing behavior
should be achieved based on source and destination prefix being
matched with label values.

Prefix   Precedence   Label
::/0   50   44
2001:db8:0:a000::/52   50   44
2001:db8:0:6666::/64   50   55
2001:db8:0:b000::/52   50   55

      Figure 10: Policy Table Needed on Failure of Uplink from SERb1

The described solution has a number of significant drawbacks, some of
them already discussed in Section 6.2.1.

* DHCPv6 support is not required for an IPv6 host, and there are
   operating systems that do not support DHCPv6. Besides that, it
   does not appear that [RFC7078] has been widely implemented on host
   operating systems.

* [RFC7078] does not clearly specify this kind of a dynamic use case
   in which the address selection policy needs to be updated quickly
   in response to the failure of a link. In a large network, it
   would present scalability issues as many hosts need to be
   reconfigured in a very short period of time.

* Updating DHCPv6 server configuration each time an ISP's uplink
   changes its state introduces some scalability issues, especially
   for mid/large distributed-scale enterprise networks. In addition
   to that, the policy table needs to be manually configured by
   administrators, which makes that solution prone to human error.

* No mechanism exists for making DHCPv6 servers aware of network
   topology/routing changes in the network. In general, having
   DHCPv6 servers monitor network-related events sounds like a bad
   idea as it requires completely new functionality beyond the scope
   of the DHCPv6 role.

6.3.2. Controlling Default Source Address Selection with Router
        Advertisements

The same mechanism as discussed in Section 6.2.2 can be used to
control the source address selection in the case of an uplink
failure. If a particular prefix should not be used as a source for
any destination, then the router needs to send an RA with the
Preferred Lifetime field for that prefix set to zero.

Let's consider a scenario in which all uplinks are operational, and
H41 receives two different RAs from R3: one from LLA_A with a PIO for
2001:db8:0:a020::/64 and the default router preference set to 11
(low), and another one from LLA_B with a PIO for

2001:db8:0:a020::/64, the default router preference set to 01 (high),
and a RIO for 2001:db8:0:6666::/64. As a result, H41 uses
2001:db8:0:b020::41 as a source address for all Internet traffic, and
those packets are sent by SERs to ISP-B. If SERb1's uplink to ISP-B
fails, the desired behavior is that H41 stops using
2001:db8:0:b020::41 as a source address for all destinations but H61.
To achieve that, R3 should react to SERb1's uplink failure (which
could be detected as the disappearance of scoped route
(S=2001:db8:0:b000::/52, D=::/0)) by withdrawing itself as a default
router. R3 sends a new RA from LLA_B with the Router Lifetime value
set to zero (which means that it should not be used as the default
router). That RA still contains a PIO for 2001:db8:0:b020::/64 (for
SLAAC purposes) and a RIO for 2001:db8:0:6666::/64 so that H41 can
reach H61 using LLA_B as a next-hop and 2001:db8:0:b020::41 as a
source address. For all traffic following the default route, LLA_A
will be used as a next-hop and 2001:db8:0:a020::41 as a source
address.

If all of the uplinks to ISP-B have failed, source addresses from
ISP-B address space should not be used. In such a failure scenario,
the forwarding table scoped S=2001:db8:0:b000::/52 contains no
entries, indicating that R3 can instruct hosts to stop using source
addresses from 2001:db8:0:b000::/52 by sending RAs containing PIOs
with Preferred Lifetime values set to zero.

6.3.3. Controlling Default Source Address Selection with ICMPv6

Now we look at how ICMPv6 messages can provide information back to
H31. We assume again that, at the time of the failure, H31 is
sending packets to H501 using (S=2001:db8:0:b010::31,
D=2001:db8:0:5678::501). When the uplink from SERb1 to ISP-B fails,
SERb1 would stop originating its source-prefix-scoped route for the
default destination (S=2001:db8:0:b000::/52, D=::/0) as well as its
unscoped default destination route. With these routes no longer in
the IGP, traffic with (S=2001:db8:0:b010::31, D=2001:db8:0:5678::501)
would end up at SERa based on the unscoped default destination route
being originated by SERa. Since that traffic has the wrong source
address to be forwarded to ISP-A, SERa would drop it and send a
Destination Unreachable message with Code 5 (Source address failed
ingress/egress policy) back to H31. H31 would then know to use
another source address for that destination and would try with
(S=2001:db8:0:a010::31, D=2001:db8:0:5678::501). This would be
forwarded to SERa based on the source-prefix-scoped default
destination route still being originated by SERa, and SERa would
forward it to ISP-A. As discussed above, if we are willing to extend
ICMPv6, SERa can even tell H31 what source address it should use to
reach that destination. The expected host behavior has been
discussed in Section 6.2.3. Using ICMPv6 would have the same
scalability/rate limiting issues that are discussed in Section 6.2.3.
An ISP-B uplink failure immediately makes source addresses from
2001:db8:0:b000::/52 unsuitable for external communication and might
trigger a large number of ICMPv6 packets being sent to hosts in that
subnet.

6.3.4. Summary of Methods for Controlling Default Source Address
        Selection on the Failure of an Uplink

It appears that DHCPv6 is not particularly well suited to quickly
changing the source address used by a host when an uplink fails,
which eliminates DHCPv6 from the list of potential solutions. On the
other hand, Router Advertisements provide a reliable mechanism to
dynamically provide hosts with a list of valid prefixes to use as
source addresses as well as to prevent the use of particular
prefixes. While no additional new features are required to be
implemented on hosts, routers need to be able to send RAs based on
the state of scoped forwarding table entries and to react to network
topology changes by sending RAs with particular parameters set.

It seems that the use of ICMPv6 Destination Unreachable messages
generated by the SER (or any SADR-capable) routers, together with the
use of RAs to signal source address selection errors back to hosts,

has the potential to provide a support mechanism. However,
scalability issues may arise in large networks when topology suddenly
changes. Therefore, it is highly desirable that hosts are able to
select the correct source address in the case of uplink failure, with
ICMPv6 being an additional mechanism to signal unexpected failures
back to hosts.

The current behaviors of different host operating systems upon
receipt of an ICMPv6 Destination Unreachable message with Code 5
(Source address failed ingress/egress policy) is not clear to the
authors. Information from implementers, users, and testing would be
quite helpful in evaluating this approach.

6.4. Selecting Default Source Address upon Failed Uplink Recovery

The next logical step is to look at the scenario when a failed uplink
on SERb1 to ISP-B comes back up, so the hosts can start using source
addresses belonging to 2001:db8:0:b000::/52 again.

6.4.1. Controlling Default Source Address Selection with DHCPv6

The mechanism to use DHCPv6 to instruct the hosts (H31 in our
example) to start using prefixes from ISP-B space (e.g.,
S=2001:db8:0:b010::31 for H31) to reach hosts on the Internet is
quite similar to one discussed in Section 6.3.1 and shares the same
drawbacks.

6.4.2. Controlling Default Source Address Selection with Router
            Advertisements

Let's look at the scenario discussed in Section 6.3.2. If the
uplink(s) failure caused the complete withdrawal of prefixes from the
2001:db8:0:b000::/52 address space by setting the Preferred Lifetime
value to zero, then the recovery of the link should just trigger the
sending of a new RA with a non-zero Preferred Lifetime. In another
scenario discussed in Section 6.3.2, the failure of the SERb1 uplink
to ISP-B leads to the disappearance of the (S=2001:db8:0:b000::/52,
D=::/0) entry from the forwarding table scoped to
S=2001:db8:0:b000::/52 and, in turn, causes R3 to send RAs with the
Router Lifetime set to zero from LLA_B. The recovery of the SERb1
uplink to ISP-B leads to the reappearance of the scoped forwarding
entry (S=2001:db8:0:b000::/52, D=::/0). That reappearance acts as a
signal for R3 to advertise itself as a default router for ISP-B
address space domain (to send RAs from LLA_B with non-zero Router
Lifetime).

6.4.3. Controlling Default Source Address Selection with ICMP

It looks like ICMPv6 provides a rather limited functionality to
signal back to hosts that particular source addresses have become
valid again. Unless the changes in the uplink specify a particular
(S,D) pair, hosts can keep using the same source address even after
an ISP uplink has come back up. For example, after the uplink from
SERb1 to ISP-B had failed, H31 received ICMPv6 Code 5 message (as
described in Section 6.3.3) and allegedly started using
(S=2001:db8:0:a010::31, D=2001:db8:0:5678::501) to reach H501. Now
when the SERb1 uplink comes back up, the packets with that (S,D) pair
are still routed to SERa1 and sent to the Internet. Therefore, H31
is not informed that it should stop using 2001:db8:0:a010::31 and
start using 2001:db8:0:b010::31 again. Unless SERa has a policy
configured to drop packets (S=2001:db8:0:a010::31,
D=2001:db8:0:5678::501) and send ICMPv6 back if the SERb1 uplink to
ISP-B is up, H31 will be unaware of the network topology change and
keep using S=2001:db8:0:a010::31 for Internet destinations, including
H51.

One of the possible options may be using a scoped route with an
EXCLUSIVE flag as described in Section 6.2.3. SERa1 uplink recovery
would cause the (S=2001:db8:0:a000::/52, D=2001:db8:0:1234::/64)
route to reappear in the routing table. In the absence of that,
route packets to H101 are sent to ISP-B (as ISP-A uplink was down)

with source addresses from 2001:db8:0:b000::/52. When the route
reappears, SERb1 rejects those packets and sends ICMPv6 back as
discussed in Section 6.2.3. Practically, it might lead to
scalability issues, which have been already discussed in 6.2.3 and
6.4.3.

6.4.4. Summary of Methods for Controlling Default Source Address
        Selection upon Failed Uplink Recovery

   Once again, DHCPv6 does not look like a reasonable choice to
   manipulate the source address selection process on a host in the case
   of network topology changes. Using Router Advertisement provides the
   flexible mechanism to dynamically react to network topology changes
   (if routers are able to use routing changes as a trigger for sending
   out RAs with specific parameters). ICMPv6 could be considered as a
   supporting mechanism to signal incorrect source address back to
   hosts, but it should not be considered as the only mechanism to
   control the address selection in multihomed environments.

6.5. Selecting Default Source Address When All Uplinks Have Failed

   One particular tricky case is a scenario when all uplinks have
   failed. In that case, there is no valid source address to be used
   for any external destinations when it might be desirable to have
   intra-site connectivity.

6.5.1. Controlling Default Source Address Selection with DHCPv6

   From the DHCPv6 perspective, uplinks failure should be treated as two
   independent failures and processed as described in Section 6.3.1. At
   this stage, it is quite obvious that it would result in a quite
   complicated policy table that would need to be explicitly configured
   by administrators and therefore seems to be impractical.

6.5.2. Controlling Default Source Address Selection with Router
        Advertisements

   As discussed in Section 6.3.2, an uplink failure causes the scoped
   default entry to disappear from the scoped forwarding table and
   triggers RAs with zero Router Lifetimes. Complete disappearance of
   all scoped entries for a given source prefix would cause the prefix
   to be withdrawn from hosts by setting the Preferred Lifetime value to
   zero in the PIO. If all uplinks (SERa, SERb1 and SERb2) fail, hosts
   either lose their default routers and/or have no global IPv6
   addresses to use as a source. (Note that 'uplink failure' might mean
   'IPv6 connectivity failure with IPv4 still being reachable', in which
   case, hosts might fall back to IPv4 if there is IPv4 connectivity to
   destinations). As a result, intra-site connectivity is broken. One
   of the possible ways to solve it is to use ULAs.

   In addition to GUAs, all hosts have ULA addresses assigned, and these
   addresses are used for intra-site communication even if there is no
   GUA assigned to a host. To avoid accidental leaking of packets with
   ULA sources, SADR-capable routers SHOULD have a scoped forwarding
   table for ULA source for internal routes but MUST NOT have an entry
   for D=::/0 in that table. In the absence of (S=ULA_Prefix; D=::/0),
   first-hop routers will send dedicated RAs from a unique link-local
   source LLA_ULA with a PIO from ULA address space, a RIO for the ULA
   prefix, and Router Lifetime set to zero. The behavior is consistent
   with the situation when SERb1 lost the uplink to ISP-B (so there is
   no Internet connectivity from 2001:db8:0:b000::/52 sources), but
   those sources can be used to reach some specific destinations. In
   the case of ULA, there is no Internet connectivity from ULA sources,
   but they can be used to reach other ULA destinations. Note that ULA
   usage could be particularly useful if all ISPs assign prefixes via
   DHCP prefix delegation. In the absence of ULAs, upon the failure of
   all uplinks, hosts would lose all their GUAs upon prefix-lifetime
   expiration, which again makes intra-site communication impossible.

   It should be noted that Rule 5.5 (prefer a prefix advertised by the
   selected next-hop) takes precedence over the Rule 6 (prefer matching

label, which ensures that GUA source addresses are preferred over
ULAs for GUA destinations). Therefore if ULAs are used, the network
administrator needs to ensure that, while the site has Internet
connectivity, hosts do not select a router that advertises ULA
prefixes as their default router.

6.5.3. Controlling Default Source Address Selection with ICMPv6

   In the case of the failure of all uplinks, all SERs will drop
   outgoing IPv6 traffic and respond with ICMPv6 error messages. In a
   large network in which many hosts attempt to reach Internet
   destinations, the SERs need to generate an ICMPv6 error for every
   packet they receive from hosts, which presents the same scalability
   issues discussed in Section 6.3.3.

6.5.4. Summary of Methods for Controlling Default Source Address
        Selection When All Uplinks Failed

   Again, combining SADR with Router Advertisements seems to be the most
   flexible and scalable way to control the source address selection on
   hosts.

6.6. Summary of Methods for Controlling Default Source Address
      Selection

   This section summarizes the scenarios and options discussed above.

   While DHCPv6 allows administrators to manipulate source address
   selection policy tables, this method has a number of significant
   disadvantages that eliminate DHCPv6 from a list of potential
   solutions:

   1. It requires hosts to support DHCPv6 and its extension [RFC7078].

   2. A DHCPv6 server needs to monitor network state and detect routing
       changes.

   3. The use of policy tables requires manual configuration and might
       be extremely complicated, especially in the case of a distributed
       network in which a large number of remote sites are being served
       by centralized DHCPv6 servers.

   4. Network topology/routing policy changes could trigger
       simultaneous reconfiguration of large number of hosts, which
       presents serious scalability issues.

   The use of Router Advertisements to influence the source address
   selection on hosts seem to be the most reliable, flexible, and
   scalable solution. It has the following benefits:

   1. No new (non-standard) functionality needs to be implemented on
       hosts (except for RIO support [RFC4191], which is not widely
       implemented at the time of this writing).

   2. No changes in RA format.

   3. Routers can react to routing table changes by sending RAs, which
       would minimize the failover time in the case of network topology
       changes.

   4. Information required for source address selection is broadcast to
       all affected hosts in the case of a topology change event, which
       improves the scalability of the solution (compared to DHCPv6
       reconfiguration or ICMPv6 error messages).

   To fully benefit from the RA-based solution, first-hop routers need
   to implement SADR, belong to the SADR domain, and be able to send
   dedicated RAs per scoped forwarding table as discussed above,
   reacting to network changes by sending new RAs. It should be noted
   that the proposed solution would work even if first-hop routers are
   not SADR-capable but still able to send individual RAs for each ISP

prefix and react to topology changes as discussed above (e.g., via configuration knobs).

The RA-based solution relies heavily on hosts correctly implementing the default address selection algorithm as defined in [RFC6724]. While the basic, and the most common, multihoming scenario (two or more Internet uplinks, no 'walled gardens') would work for any host supporting the minimal implementation of [RFC6724], more complex use cases (such as 'walled garden' and other scenarios when some ISP resources can only be reached from that ISP address space) require that hosts support Rule 5.5 of the default address selection algorithm. There is some evidence that not all host OSes have that rule implemented currently. However, it should be noted that [RFC8028] states that Rule 5.5 should be implemented.

The ICMPv6 Code 5 error message SHOULD be used to complement an RA-based solution to signal incorrect source address selection back to hosts, but it SHOULD NOT be considered as the standalone solution. To prevent scenarios when hosts in multihomed environments incorrectly identify on-link/off-link destinations, hosts SHOULD treat ICMPv6 Redirects as discussed in [RFC8028].

6.7. Solution Limitations

6.7.1. Connections Preservation

The proposed solution is not designed to preserve connection state in the case of an uplink failure. When all uplinks to an ISP go down, all transport connections established to/from that ISP address space will be interrupted (unless the transport protocol has specific multihoming support). That behavior is similar to the scenario of IPv4 multihoming with NAT when an uplink failure causes all connections to be NATed to completely different public IPv4 addresses. While it does sound suboptimal, it is determined by the nature of PA address space: if all uplinks to the particular ISP have failed, there is no path for the ingress traffic to reach the site, and the egress traffic is supposed to be dropped by the ingress filters [BCP38]. The only potential way to overcome this limitation would be to run BGP with all ISPs and to advertise all site prefixes to all uplinks - a solution that shares all the drawbacks of using the PI address space without sharing its benefits. Networks willing and capable of running BGP and using PI are out of scope of this document.

It should be noted that in the case of IPv4 NAT-based multihoming, uplink recovery could cause connection interruptions as well (unless packet forwarding is integrated with the tracking of existing NAT sessions so that the egress interface for the existing sessions is not changed). However, the proposed solution has the benefit of preserving the existing sessions during and after the restoration of the failed uplink. Unlike the uplink failure event, which causes all addresses from the affected prefix to be deprecated, the recovery would just add new, preferred addresses to a host without making any addresses unavailable. Therefore, connections established to and from those addresses do not have to be interrupted.

While it's desirable for active connections to survive ISP failover events, such events affect the reachability of IP addresses assigned to hosts in sites using PA address space. Unless the transport (or higher-level protocols) is capable of surviving the host renumbering, the active connections will be broken. The proposed solution focuses on minimizing the impact of failover on new connections and on multipath-aware protocols.

Another way to preserve connection state is to use multipath transport as discussed in Section 8.3.

6.8. Other Configuration Parameters

6.8.1. DNS Configuration

In a multihomed environment, each ISP might provide their own list of
DNS servers. For example, in the topology shown in Figure 3, ISP-A
might provide H51 2001:db8:0:5555::51 as a recursive DNS server,
while ISP-B might provide H61 2001:db8:0:6666::61 as a recursive DNS
server (RDNSS). [RFC8106] defines IPv6 Router Advertisement options
to allow IPv6 routers to advertise a list of RDNSS addresses and a
DNS Search List (DNSSL) to IPv6 hosts. Using RDNSS together with
'scoped' RAs as described above would allow a first-hop router (R3 in
Figure 3) to send DNS server addresses and search lists provided by
each ISP (or the corporate DNS server addresses if the enterprise is
running its own DNS servers. As discussed below, the DNS split-
horizon problem is too hard to solve without running a local DNS
server).

As discussed in Section 6.5.2, failure of all ISP uplinks would cause
deprecation of all addresses assigned to a host from the address
space of all ISPs. If any intra-site IPv6 connectivity is still
desirable (most likely to be the case for any mid/large-scale
network), then ULAs should be used as discussed in Section 6.5.2. In
such a scenario, the enterprise network should run its own recursive
DNS server(s) and provide its ULA addresses to hosts via RDNSS in RAs
sent for ULA-scoped forwarding table as described in Section 6.5.2.

There are some scenarios in which the final outcome of the name
resolution might be different depending on:

* which DNS server is used;

* which source address the client uses to send a DNS query to the
   server (DNS split horizon).

There is no way currently to instruct a host to use a particular DNS
server from the configured servers list for resolving a particular
name. Therefore, it does not seem feasible to solve the problem of
DNS server selection on the host (it should be noted that this
particular issue is protocol-agnostic and happens for IPv4 as well).
In such a scenario, it is recommended that the enterprise run its own
local recursive DNS server.

To influence host source address selection for packets sent to a
particular DNS server, the following requirements must be met:

* The host supports RIO as defined in [RFC4191].

* The routers send RIOs for routes to DNS server addresses.

For example, if it is desirable that host H31 reaches the ISP-A DNS
server H51 2001:db8:0:5555::51 using its source address
2001:db8:0:a010::31, then both R1 and R2 should send RIOs containing
the route to 2001:db8:0:5555::51 (or covering route) in their
'scoped' RAs, containing LLA_A as the default router address and the
PIO for SLAAC prefix 2001:db8:0:a010::/64. In that case, host H31
(if it supports Rule 5.5) would select LLA_A as a next-hop and then
choose 2001:db8:0:a010::31 as the source address for packets to the
DNS server.

It should be noted that [RFC6106] explicitly prohibits using DNS
information if the RA Router Lifetime has expired:

   An RDNSS address or a DNSSL domain name MUST be used only as long
   as both the RA router Lifetime (advertised by a Router
   Advertisement message) and the corresponding option Lifetime have
   not expired.

Therefore, hosts might ignore RDNSS information provided in ULA-
scoped RAs, as those RAs would have Router Lifetime values set to
zero. However, [RFC8106], which obsoletes RFC 6106, has removed that
requirement.

As discussed above, the DNS split-horizon problem and the selection
of the correct DNS server in a multihomed environment are not easy

problems to solve. The proper solution would require hosts to
support the concept of multiple provisioning domains (PvD, a set of
configuration information associated with a network, [RFC7556]).

7. Deployment Considerations

   The solution described in this document requires certain mechanisms
   to be supported by the network infrastructure and hosts. It requires
   some routers in the enterprise site to support some form of SADR. It
   also requires hosts to be able to learn when the uplink to an ISP
   changes its state so that the hosts can use appropriate source
   addresses. Ongoing work to create mechanisms to accomplish this are
   discussed in this document, but they are still works in progress.

7.1. Deploying SADR Domain

   The proposed solution does not prescribe particular details regarding
   deploying an SADR domain within a multihomed enterprise network.
   However the following guidelines could be applied:

   * The SADR domain is usually limited by the multihomed site border.

   * The minimal deployable scenario requires enabling SADR on all SERs
      and including them into a single SADR domain.

   * As discussed in Section 4.2, extending the connected SADR domain
      beyond the SERs to the first-hop routers can produce more
      efficient forwarding paths and allow the network to fully benefit
      from SADR. It would also simplify the operation of the SADR
      domain.

   * During the incremental SADR domain expansion from the SERs down
      towards first-hop routers, it's important to ensure that, at any
      given moment, all SADR-capable routers within the domain are
      logically connected (see Section 5).

7.2. Hosts-Related Considerations

   The solution discussed in this document relies on the default address
   selection algorithm, Rule 5.5 [RFC6724]. While [RFC6724] considers
   this rule as optional, the more recent [RFC8028] states that "A host
   SHOULD select default routers for each prefix it is assigned an
   address in." It also recommends that hosts should implement Rule
   5.5. of [RFC6724]. Therefore, while hosts compliant with RFC 8028
   already have a mechanism to learn about state changes to ISP uplinks
   and to select the source addresses accordingly, many hosts do not
   support such a mechanism yet.

   It should be noted that a multihomed enterprise network utilizing
   multiple ISP prefixes can be considered as a typical multiple
   provisioning domain (mPvD) scenario, as described in [RFC7556]. This
   document defines a way for the network to provide the PvD information
   to hosts indirectly, using the existing mechanisms. At the same
   time, [PROV-DOMAINS] takes one step further and describes a
   comprehensive mechanism for hosts to discover the whole set of
   configuration information associated with different PvDs/ISPs.
   [PROV-DOMAINS] complements this document in terms of enabling hosts
   to learn about ISP uplink states and to select the corresponding
   source addresses.

8. Other Solutions

8.1. Shim6

   The Shim6 protocol [RFC5533], specified by the Shim6 working group,
   allows a host at a multihomed site to communicate with an external
   host and to exchange information about possible source and
   destination address pairs that they can use to communicate. The
   Shim6 working group also specified the REAchability Protocol (REAP)
   [RFC5534] to detect failures in the path between working address
   pairs and to find new working address pairs. A fundamental

requirement for Shim6 is that both internal and external hosts need to support Shim6. That is, both the host internal to the multihomed site and the host external to the multihomed site need to support Shim6 in order for there to be any benefit for the internal host to run Shim6. The Shim6 protocol specification was published in 2009, but it has not been widely implemented. Therefore Shim6 is not considered as a viable solution for enterprise multihoming.

8.2. IPv6-to-IPv6 Network Prefix Translation

IPv6-to-IPv6 Network Prefix Translation (NPTv6) [RFC6296] is not the focus of this document. NPTv6 suffers from the same fundamental issue as any other approaches to address translation: it breaks end-to-end connectivity. Therefore NPTv6 is not considered as a desirable solution, and this document intentionally focuses on solving the enterprise multihoming problem without any form of address translation.

With increasing interest and ongoing work in bringing path awareness to transport- and application-layer protocols, hosts might be able to determine the properties of the various network paths and choose among the paths that are available to them. As selecting the correct source address is one of the possible mechanisms that path-aware hosts may utilize, address translation negatively affects hosts' path-awareness, which makes NTPv6 an even more undesirable solution.

8.3. Multipath Transport

Using multipath transport (such as Multipath TCP (MPTCP) [RFC6824] or multipath capabilities in QUIC) might solve the problems discussed in Section 6 since a multipath transport would allow hosts to use multiple source addresses for a single connection and to switch between those source addresses when a particular address becomes unavailable or a new address is assigned to the host interface. Therefore, if all hosts in the enterprise network use only multipath transport for all connections, the signaling solution described in Section 6 might not be needed (it should be noted that Source Address Dependent Routing would still be required to deliver packets to the correct uplinks). At the time this document was written, multipath transport alone could not be considered a solution for the problem of selecting the source address in a multihomed environment. There are a significant number of hosts that do not use multipath transport currently, and it seems unlikely that the situation will change in the foreseeable future (even if new releases of operating systems support multipath protocols, there will be a long tail of legacy hosts). The solution for enterprise multihoming needs to work for the least common denominator: hosts without multipath transport support. In addition, not all protocols are using multipath transport. While multipath transport would complement the solution described in Section 6, it could not be considered as a sole solution to the problem of source address selection in multihomed environments.

On the other hand, PA-based multihoming could provide additional benefits to multipath protocols, should those protocols be deployed in the network. Multipath protocols could leverage source address selection to achieve maximum path diversity (and potentially improved performance).

Therefore, the deployment of multipath protocols should not be considered as an alternative to the approach proposed in this document. Instead, both solutions complement each other, so deploying multipath protocols in a PA-based multihomed network proves mutually beneficial.

9. IANA Considerations

This document has no IANA actions.

10. Security Considerations

Section 6.2.3 discusses a mechanism for controlling source address
selection on hosts using ICMPv6 messages. Using ICMPv6 to influence
source address selection allows an attacker to exhaust the list of
candidate source addresses on the host by sending spoofed ICMPv6 Code
5 for all prefixes known on the network (therefore preventing a
victim from establishing communication with the destination host).
Another possible attack vector is using ICMPv6 Destination
Unreachable Messages with Code 5 to steer the egress traffic towards
the particular ISP, so the attacker can benefit from their ability
doing traffic sniffing/interception in that ISP network.

To prevent those attacks, hosts SHOULD verify that the original
packet header included in the ICMPv6 error message was actually sent
by the host (to ensure that the ICMPv6 message was triggered by a
packet sent by the host).

As ICMPv6 Destination Unreachable Messages with Code 5 could be
originated by any SADR-capable router within the domain (or even come
from the Internet), the Generalized TTL Security Mechanism (GTSM)
[RFC5082] cannot be applied. Filtering such ICMPv6 messages at the
site border cannot be recommended as it would break the legitimate
end-to-end error signaling mechanism for which ICMPv6 was designed.

The security considerations of using stateless address
autoconfiguration are discussed in [RFC4862].

11. References

11.1. Normative References

   [BCP38]   Ferguson, P. and D. Senie, "Network Ingress Filtering:
               Defeating Denial of Service Attacks which employ IP Source
               Address Spoofing", BCP 38, RFC 2827, DOI 10.17487/RFC2827,
               May 2000, <https://www.rfc-editor.org/info/rfc2827>.

   [RFC1918]   Rekhter, Y., Moskowitz, B., Karrenberg, D., de Groot, G.
               J., and E. Lear, "Address Allocation for Private
               Internets", BCP 5, RFC 1918, DOI 10.17487/RFC1918,
               February 1996, <https://www.rfc-editor.org/info/rfc1918>.

   [RFC2119]   Bradner, S., "Key words for use in RFCs to Indicate
               Requirement Levels", BCP 14, RFC 2119,
               DOI 10.17487/RFC2119, March 1997,
               <https://www.rfc-editor.org/info/rfc2119>.

   [RFC4191]   Draves, R. and D. Thaler, "Default Router Preferences and
               More-Specific Routes", RFC 4191, DOI 10.17487/RFC4191,
               November 2005, <https://www.rfc-editor.org/info/rfc4191>.

   [RFC4193]   Hinden, R. and B. Haberman, "Unique Local IPv6 Unicast
               Addresses", RFC 4193, DOI 10.17487/RFC4193, October 2005,
               <https://www.rfc-editor.org/info/rfc4193>.

   [RFC4291]   Hinden, R. and S. Deering, "IP Version 6 Addressing
               Architecture", RFC 4291, DOI 10.17487/RFC4291, February
               2006, <https://www.rfc-editor.org/info/rfc4291>.

   [RFC4443]   Conta, A., Deering, S., and M. Gupta, Ed., "Internet
               Control Message Protocol (ICMPv6) for the Internet
               Protocol Version 6 (IPv6) Specification", STD 89,
               RFC 4443, DOI 10.17487/RFC4443, March 2006,
               <https://www.rfc-editor.org/info/rfc4443>.

   [RFC4861]   Narten, T., Nordmark, E., Simpson, W., and H. Soliman,
               "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861,
               DOI 10.17487/RFC4861, September 2007,
               <https://www.rfc-editor.org/info/rfc4861>.

   [RFC4862]   Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless
               Address Autoconfiguration", RFC 4862,
               DOI 10.17487/RFC4862, September 2007,

```
                   <https://www.rfc-editor.org/info/rfc4862>.

   [RFC6106]  Jeong, J., Park, S., Beloeil, L., and S. Madanapalli,
              "IPv6 Router Advertisement Options for DNS Configuration",
              RFC 6106, DOI 10.17487/RFC6106, November 2010,
              <https://www.rfc-editor.org/info/rfc6106>.

   [RFC6296]  Wasserman, M. and F. Baker, "IPv6-to-IPv6 Network Prefix
              Translation", RFC 6296, DOI 10.17487/RFC6296, June 2011,
              <https://www.rfc-editor.org/info/rfc6296>.

   [RFC6724]  Thaler, D., Ed., Draves, R., Matsumoto, A., and T. Chown,
              "Default Address Selection for Internet Protocol Version 6
              (IPv6)", RFC 6724, DOI 10.17487/RFC6724, September 2012,
              <https://www.rfc-editor.org/info/rfc6724>.

   [RFC7078]  Matsumoto, A., Fujisaki, T., and T. Chown, "Distributing
              Address Selection Policy Using DHCPv6", RFC 7078,
              DOI 10.17487/RFC7078, January 2014,
              <https://www.rfc-editor.org/info/rfc7078>.

   [RFC7556]  Anipko, D., Ed., "Multiple Provisioning Domain
              Architecture", RFC 7556, DOI 10.17487/RFC7556, June 2015,
              <https://www.rfc-editor.org/info/rfc7556>.

   [RFC8028]  Baker, F. and B. Carpenter, "First-Hop Router Selection by
              Hosts in a Multi-Prefix Network", RFC 8028,
              DOI 10.17487/RFC8028, November 2016,
              <https://www.rfc-editor.org/info/rfc8028>.

   [RFC8106]  Jeong, J., Park, S., Beloeil, L., and S. Madanapalli,
              "IPv6 Router Advertisement Options for DNS Configuration",
              RFC 8106, DOI 10.17487/RFC8106, March 2017,
              <https://www.rfc-editor.org/info/rfc8106>.

   [RFC8174]  Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC
              2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174,
              May 2017, <https://www.rfc-editor.org/info/rfc8174>.

   [RFC8415]  Mrugalski, T., Siodelski, M., Volz, B., Yourtchenko, A.,
              Richardson, M., Jiang, S., Lemon, T., and T. Winters,
              "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)",
              RFC 8415, DOI 10.17487/RFC8415, November 2018,
              <https://www.rfc-editor.org/info/rfc8415>.

11.2.  Informative References

   [DST-SRC-RTG]
              Lamparter, D. and A. Smirnov, "Destination/Source
              Routing", Work in Progress, Internet-Draft, draft-ietf-
              rtgwg-dst-src-routing-07, 10 March 2019,
              <https://tools.ietf.org/html/draft-ietf-rtgwg-dst-src-
              routing-07>.

   [PROV-DOMAINS]
              Pfister, P., Vyncke, E., Pauly, T., Schinazi, D., and W.
              Shao, "Discovering Provisioning Domain Names and Data",
              Work in Progress, Internet-Draft, draft-ietf-intarea-
              provisioning-domains-09, 6 December 2019,
              <https://tools.ietf.org/html/draft-ietf-intarea-
              provisioning-domains-09>.

   [RFC2663]  Srisuresh, P. and M. Holdrege, "IP Network Address
              Translator (NAT) Terminology and Considerations",
              RFC 2663, DOI 10.17487/RFC2663, August 1999,
              <https://www.rfc-editor.org/info/rfc2663>.

   [RFC3704]  Baker, F. and P. Savola, "Ingress Filtering for Multihomed
              Networks", BCP 84, RFC 3704, DOI 10.17487/RFC3704, March
              2004, <https://www.rfc-editor.org/info/rfc3704>.
```

   [RFC4941] Narten, T., Draves, R., and S. Krishnan, "Privacy
              Extensions for Stateless Address Autoconfiguration in
              IPv6", RFC 4941, DOI 10.17487/RFC4941, September 2007,
              <https://www.rfc-editor.org/info/rfc4941>.

   [RFC5082] Gill, V., Heasley, J., Meyer, D., Savola, P., Ed., and C.
              Pignataro, "The Generalized TTL Security Mechanism
              (GTSM)", RFC 5082, DOI 10.17487/RFC5082, October 2007,
              <https://www.rfc-editor.org/info/rfc5082>.

   [RFC5533] Nordmark, E. and M. Bagnulo, "Shim6: Level 3 Multihoming
              Shim Protocol for IPv6", RFC 5533, DOI 10.17487/RFC5533,
              June 2009, <https://www.rfc-editor.org/info/rfc5533>.

   [RFC5534] Arkko, J. and I. van Beijnum, "Failure Detection and
              Locator Pair Exploration Protocol for IPv6 Multihoming",
              RFC 5534, DOI 10.17487/RFC5534, June 2009,
              <https://www.rfc-editor.org/info/rfc5534>.

   [RFC6824] Ford, A., Raiciu, C., Handley, M., and O. Bonaventure,
              "TCP Extensions for Multipath Operation with Multiple
              Addresses", RFC 6824, DOI 10.17487/RFC6824, January 2013,
              <https://www.rfc-editor.org/info/rfc6824>.

   [RFC7676] Pignataro, C., Bonica, R., and S. Krishnan, "IPv6 Support
              for Generic Routing Encapsulation (GRE)", RFC 7676,
              DOI 10.17487/RFC7676, October 2015,
              <https://www.rfc-editor.org/info/rfc7676>.

   [RFC8425] Troan, O., "IANA Considerations for IPv6 Neighbor
              Discovery Prefix Information Option Flags", RFC 8425,
              DOI 10.17487/RFC8425, July 2018,
              <https://www.rfc-editor.org/info/rfc8425>.

   [RFC8504] Chown, T., Loughney, J., and T. Winters, "IPv6 Node
              Requirements", BCP 220, RFC 8504, DOI 10.17487/RFC8504,
              January 2019, <https://www.rfc-editor.org/info/rfc8504>.

   [SADR-RA] Pfister, P., "Source Address Dependent Route Information
              Option for Router Advertisements", Work in Progress,
              Internet-Draft, draft-pfister-6man-sadr-ra-01, 22 June
              2015, <https://tools.ietf.org/html/draft-pfister-6man-
              sadr-ra-01>.

Acknowledgements

Authors' Addresses

   Fred Baker
   Santa Barbara, California 93117
   United States of America

   Email: FredBaker.IETF@gmail.com

   Chris Bowers
   Juniper Networks
   Sunnyvale, California 94089
   United States of America

   Email: cbowers@juniper.net

Jen Linkova
Google
1 Darling Island Rd
Pyrmont NSW 2009
Australia

Email: furry@google.com