IPv6 Operations Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 13, 2008

A. Matsumoto
T. Fujisaki
NTT
R. Hiromi
K. Kanayama
Intec Netcore

October 11, 2007

# Problem Statement of Default Address Selection in Multi-prefix Environment: Operational Issues of RFC3484 Default Rules
# draft-ietf-v6ops-addr-select-ps-02.txt

## Status of this Memo

## Copyright Notice

## Abstract

One physical network can carry multiple logical networks.  Moreover, we can use multiple physical networks at the same time in a host.  In that environment, end hosts might have multiple IP addresses and be required to use them selectively.  Without an appropriate source/

destination address selection mechanism, the host will experience some trouble in communication. RFC 3484 defines both the source and destination address selection algorithms, but the multi-prefix environment considered here needs additional rules beyond those of the default operation. This document describes the possible problems that end hosts could encounter in an environment with multiple logical networks.

Table of Contents

## 1. Introduction

One physical network can carry multiple logical networks. In that case, an end-host has multiple IP addresses. In the IPv4-IPv6 dual stack environment or in a site connected to both a ULA [RFC4193] and Global scope networks, an end-host has multiple IP addresses. These are examples of networks that we focus on in this document. In such an environment, an end-host will encounter some communication trouble.

Inappropriate source address selection at the end-host causes unexpected asymmetric routing, filtering by a router or discarding of packets bacause there is no route to the host.

Considering a multi-prefix environment, destination address selection is also important for correct or better communication establishment.

RFC 3484 [RFC3484] defines both source and destination address selection algorithms. In most cases, the host will be able to communicate with the targeted host using the algorithms. However, there are still problematic cases such as when multiple default routes are supplied. This document describes such possibilities of incorrect address selection, which leads to dropping packets and communication failure.

In addition, the provision of an address policy table is an important matter. RFC 3484 describes all the algorithms for setting the address policy table but address policy provisions are not mentioned. RFC 3484 only defines how to configure the address policy table manually.

### 1.1. Scope of this document

There has been a lot of discussion about "multiple addresses/ prefixes" but the multi-homing techniques for achieving redundancy are out of our scope. Cooperation with a mechanism like shim6 is rather desirable. We focus on an end-site network environment. The scope of this document is to sort out problematic cases of false dropping of the address selection within a multi-prefix environment.

## 2. Problem Statement

### 2.1. Source Address Selection

### 2.1.1.  Multiple Routers on Single Interface

```
                          ==================
                          |    Internet    |
                          ==================
                             |          |
          2001:db8:1000::/36 |          | 2001:db8:8000::/36
                        +----+-+      +-+----+
                        | ISP1 |      | ISP2 |
                        +----+-+      +-+----+
                             |          |
         2001:db8:1000:::/48 |          | 2001:db8:8000::/48
                      +------+---+ +----+-----+
                      | Gateway1 | | Gateway2 |
                      +--------+-+ +-+--------+
                               |     |
          2001:db8:1000:1::/64 |     | 2001:db8:8000:1::/64
                               |     |
                        -----+-+-----+------
                             |
                           +-+----+ 2001:db8:1000:1::EUI64
                           | Host | 2001:db8:8000:1::EUI64
                           +------+
```

[Fig. 1]

Generally speaking, there is no interaction between next-hop determination and address selection.  In this example, when a Host sends a packet via Gateway1, the Host does not necessarily choose address 2001:db8:1000:1::EUI64 given by Gateway1 as the source address.  This causes the same problem as described in the next section 'Ingress Filtering Problem'.

### 2.1.2. Ingress Filtering Problem

```
                        ==================
                        |    Internet    |
                        ==================
                             |       |
          2001:db8:1000::/36 |       | 2001:db8:8000::/36
                        +----+-+   +-+----+
                        | ISP1 |   | ISP2 |
                        +----+-+   +-+----+
                             |       |
         2001:db8:1000:::/48 |       | 2001:db8:8000::/48
                            ++-------++
                            | Gateway |
                            +----+----+
                                 |  2001:db8:1000:1::/64
                                 |  2001:db8:8000:1::/64
                       ------+---+----------
                             |
                           +-+----+ 2001:db8:1000:1::EUI64
                           | Host | 2001:db8:8000:1::EUI64
                           +------+
```

[Fig. 2]

When a relatively small site, which we call a "customer network", is attached to two upstream ISPs, each ISP delegates a network address block, which is usually /48, and a host has multiple IPv6 addresses.

When the source address of an outgoing packet is not the one that is delegated by an upstream ISP, there is a possibility that the packet will be dropped at the ISP by its Ingress Filter. Ingress filtering(uRPF: unicast Reverse Path Forwarding) is becoming more popular among ISPs to mitigate the damage of DoS attacks.

In this example, when the Gateway chooses the default route to ISP2 and the Host chooses 2001:db8:1000:1::EUI64 as the source address for packets sent to a host (2001:db8:2000::1) somewhere on the Internet, the packets may be dropped at ISP2 because of Ingress Filtering.

### 2.1.3. Half-Closed Network Problem

You can see a second typical source address selection problem in a multihome site with global-closed mixed connectivity like in the figure below. In this case, Host-A is in a multihomed network and has two IPv6 addresses, one delegated from each of the upstream ISPs. Note that ISP2 is a closed network and does not have connectivity to the Internet.

```
                           +--------+
                           | Host-C | 2001:db8:a000::1
                           +-----+--+
                                 |
                        ==============  +--------+
                        |  Internet  |  | Host-B | 2001:db8:8000::1
                        ==============  +--------+
                             |            |
           2001:db8:1000:/36 |            | 2001:db8:8000::/36
                        +----+-+   +-+---++
                        | ISP1 |   | ISP2 | (Closed Network/VPN tunnel)
                        +----+-+   +-+----+
                             |       |
           2001:db8:1000:/48 |       | 2001:db8:8000::/48
                            ++-------++
                            | Gateway |
                            +----+----+
                                 |  2001:db8:1000:1::/64
                                 |  2001:db8:8000:1::/64
                       ------+---+----------
                             |
                          +--+-----+ 2001:db8:1000:1::EUI64
                          | Host-A | 2001:db8:8000:1::EUI64
                          +--------+
```

[Fig. 3]

You do not need two physical network connections here. The connection from the Gateway to ISP2 can be a logical link over ISP1 and the Internet.

When Host-A starts the connection to Host-B in ISP2, the source address of a packet that has been sent will be the one delegated from ISP2, that is 2001:db8:8000:1::EUI64, because of rule 8 (longest matching prefix) in RFC 3484.

Host-C is located somewhere on the Internet and has IPv6 address 2001:db8:a000::1. When Host-A sends a packet to Host-C, the longest matching algorithm chooses 2001:db8:8000:1::EUI64 for the source address. In this case, the packet goes through ISP1 and may be filtered by ISP1's ingress filter. Even if the packet is not filtered by ISP1, a return packet from Host-C cannot possibly be delivered to Host-A because the return packet is destined for 2001:db8:8000:1::EUI64, which is closed from the Internet.

The important point is that each host chooses a correct source address for a given destination address as long as NAT does not exist in the IPv6 world.

### 2.1.4. Combined Use of Global and ULA

```
                        ============
                        | Internet |
                        ============
                              |
                              |
                         +----+----+
                         |   ISP   |
                         +----+----+
                              |
              2001:db8:a::/48 |
                         +----+----+
                         | Gateway |
                         +-+-----+-+
                           |     | 2001:db8:a:100::/64
          fd01:2:3:200:/64 |     | fd01:2:3:100:/64
                   -----+--+-   -+--+----
                        |           |
    fd01:2:3:200::EUI64 |           |      2001:db8:a:100::EUI64
                   +----+----+    +-+----+ fd01:2:3:100::EUI64
                   | Printer |    | Host |
                   +---------+    +------+
```

[Fig. 4]

As NAP [I-D.ietf-v6ops-nap] describes, using a ULA may be beneficial in some scenarios. If the ULA is used for internal communication, packets with ULA need to be filtered at the Gateway.

There is no serious problem related to address selection in this case, because of the dissimilarity between the ULA and the Global Unicast Address. The longest matching rule of RFC 3484 chooses the correct address for both intra-site and extra-site communication.

In a few years, however, the longest matching rule will not be able to choose the correct address anymore. That is the moment when the assignment of those Global Unicast Addresses starts, where the first bit is 1. In RFC 4291 [RFC4291], almost all address spaces of IPv6, including those whose first bit is 1, are assigned as Global Unicast Addresses.

### 2.1.5. Site Renumbering

RFC 4192 [RFC4192] describes a recommended procedure for renumbering a network from one prefix to another. An autoconfigured address has a lifetime, so by stopping advertisement of the old prefix, the autoconfigured address is eventually invalidated.

However, invalidating the old prefix takes a long time. You cannot stop routing to the old prefix as long as the old prefix is not removed from the host. This can be a tough issue for ISP network administrators.

```
                              +-----+---+
                              | Gateway |
                              +----+----+
                                   |  2001:db8:b::/64  (new)
                                   |  2001:db8:a::/64 (old)
                         ------+---+----------
                               |
                            +--+-----+ 2001:db8:b::EUI64  (new)
                            | Host-A | 2001:db8:a::EUI64 (old)
                            +--------+

                                [Fig. 5]
```

### 2.1.6. Multicast Source Address Selection

This case is an example of Site-local or Global prioritization. When you send a multicast packet across site-borders, the source address of the multicast packet must be a global scope address. The longest matching algorithm, however, selects a ULA if the sending host has both a ULA and a global address.

### 2.1.7. Temporary Address Selection

RFC 3041 [RFC3041] defines a Temporary Address. The usage of a Temporary Address has both pros and cons. That is good for viewing web pages or communicating with the general public, but that is bad for a service that uses address-based authentication and for logging purposes.

If you could turn the temporary address on and off, that would be better. If you could switch its usage per service(destination address), that would also be better. The same situation can be found when using HA and CoA in a MobileIP network.

## 2.2. Destination Address Selection

### 2.2.1. IPv4 or IPv6 prioritization

The default policy table gives IPv6 addresses higher precedence than IPv4 addresses. There seem to be many cases, however, where network administrators want to control the address selection policy of end-hosts the other way around.

```
                            +---------+
                            | Tunnel  |
                            | Service |
                            +--+---++-+
                               |   ||
                               |   ||
                        ===========||==
                        | Internet || |
                        ===========||==
                             |     ||
                192.0.2.0/24 |     ||
                        +----+-+   ||
                        | ISP  |   ||
                        +----+-+   ||
                             |     ||
               IPv4 (Native) |     || IPv6 (Tunnel)
                192.0.2.0/26 |     ||
                            ++-----++-+
                            | Gateway |
                            +----+----+
                                 |  2001:db8:a:1::/64
                                 |  192.0.2.0/28
                                 |
                       ------+---+----------
                             |
                           +-+----+ 2001:db8:a:1:EUI64
                           | Host | 192.0.2.2
                           +------+
```

[Fig. 6]

In the figure above, a site has native IPv4 and tunneled-IPv6 connectivity. Therefore, the administrator may want to set a higher priority for using IPv4 than using IPv6 because the quality of the tunnel network seems to be worse than that of the native transport.

### 2.2.2. ULA and IPv4 dual-stack environment

This is a special form of IPv4 and IPv6 prioritization. When an enterprise has IPv4 Internet connectivity but does not yet have IPv6 Internet connectivity, and the enterprise wants to provide site-local IPv6 connectivity, a ULA is the best choice for site-local IPv6 connectivity. Each employee host will have both an IPv4 global or private address and a ULA. Here, when this host tries to connect to Host-C that has registered both A and AAAA records in the DNS, the host will choose AAAA as the destination address and the ULA for the source address. This will clearly result in a connection failure.

```
                           +--------+
                           | Host-C | AAAA = 2001:db8::80
                           +-----+--+ A    = 192.0.2.1
                                 |
                        ============
                        | Internet |
                        ============
                             |  no IPv6 connectivity
                        +----+----+
                        | Gateway |
                        +----+----+
                             |
                             | fd01:2:3::/48 (ULA)
                             | 192.0.2.128/25
                            ++--------+
                            | Router  |
                            +----+----+
                                 |  fd01:2:3:4::/64 (ULA)
                                 |  192.0.2.240/28
                       ------+---+----------
                             |
                           +-+----+ fd01:2:3:4::100 (ULA)
                           | Host | 192.0.2.245
                           +------+
```

[Fig. 7]

### 2.2.3. ULA or Global Prioritization

Differentiating services by the client's source address is very common. IP-address-based authentication is an typical example of this. Another typical example is a web service that has pages for the public and internal pages for employees or involved parties. Yet another example is DNS zone splitting.

However, a ULA and IPv6 global address both have global scope, and RFC3484 default rules do not specify which address should be given priority. This point makes IPv6 implementation of address-based service differentiation a bit harder.

```
                            +------+
                            | Host |
                            +-+--|-+
                              |  |
                      ===========|==
                      | Internet | |
                      ===========|==
                            |    |
                            |    |
                       +----+-+  +-->+------+
                       | ISP  +------+  DNS | 2001:db8:a::80
                       +----+-+  +-->+------+ fc12:3456:789a::80
                            |    |
            2001:db8:a::/48 |    |
        fc12:3456:789a::/48 |    |
                       +----+----|+
                       | Gateway ||
                       +---+-----|+
                           |     |    2001:db8:a:100::/64
                           |     |    fc12:3456:789a:100::/64
                         --+-+---|-----
                             |   |
                           +-+---|+ 2001:db8:a:100::EUI64
                           | Host | fc12:3456:789a:100::EUI64
                           +------+
```

[Fig. 7]

## 3. Conclusion

We have covered problems related to destination or source address selection. These problems have their roots in the situation where end-hosts have multiple IP addresses. In this situation, every end-host must choose an appropriate destination and source address, which cannot be achieved only by routers.

It should be noted that end-hosts must be informed about routing policies of their upstream networks for appropriate address selection. A site administrator must consider every possible address false-selection problem and take countermeasures beforehand.

## 4. Security Considerations

When an intermediate router performs policy routing (e.g. source address based routing), inappropriate address selection causes unexpected routing. For example, in the network described in 2.1.3,

when Host-A uses a default address selection policy and chooses an inappropriate address, a packet sent to VPN can be delivered to a location via the Internet. This issue can lead to packet eavesdropping or session hijack.

As documented in the security consideration section in RFC 3484, address selection algorithms expose a potential privacy concern. When a malicious host can make a target host perform address selection, the malicious host can know multiple addresses attached to the target host. In a case like 2.1.4, if an attacker can make Host to send a multicast packet and the Host performs the default address selection algorithm, the attacker may be able to determine the ULAs attached to the Host.

These security risks have roots in inappropriate address selection. Therefore, if a countermeasure is taken, and hosts always select an appropriate address that is suitable to a site's network structure and routing, these risks can be avoided.

## 5. IANA Considerations

This document has no actions for IANA.

## 6. References

### 6.1. Normative References

[RFC3484] Draves, R., "Default Address Selection for Internet Protocol version 6 (IPv6)", RFC 3484, February 2003.

[RFC4193] Hinden, R. and B. Haberman, "Unique Local IPv6 Unicast Addresses", RFC 4193, October 2005.

### 6.2. Informative References

[I-D.ietf-v6ops-nap] Velde, G., "Local Network Protection for IPv6", draft-ietf-v6ops-nap-06 (work in progress), January 2007.

[RFC3041] Narten, T. and R. Draves, "Privacy Extensions for Stateless Address Autoconfiguration in IPv6", RFC 3041, January 2001.

[RFC4192] Baker, F., Lear, E., and R. Droms, "Procedures for Renumbering an IPv6 Network without a Flag Day", RFC 4192, September 2005.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, February 2006.

## Appendix A. Appendix. Revision History

01:

IP addresse notations changed to docmentation address.
Descriptoin of solutions deleted.

02:

Security considerations section rewritten according to comments from SECDIR.

## Authors' Addresses

Arifumi Matsumoto
NTT PF Lab
Midori-Cho 3-9-11
Musashino-shi, Tokyo 180-8585
Japan

Phone: +81 422 59 3334
Email: arifumi@nttv6.net

Tomohiro Fujisaki
NTT PF Lab
Midori-Cho 3-9-11
Musashino-shi, Tokyo 180-8585
Japan

Phone: +81 422 59 7351
Email: fujisaki@nttv6.net

Ruri Hiromi
Intec Netcore, Inc.
Shinsuna 1-3-3
Koto-ku, Tokyo 136-0075
Japan

Phone: +81 3 5665 5069
Email: hiromi@inetcore.com

Ken-ichi Kanayama
Intec Netcore, Inc.
Shinsuna 1-3-3
Koto-ku, Tokyo  136-0075
Japan

Phone: +81 3 5665 5069
Email: kanayama@inetcore.com

## Full Copyright Statement

## Intellectual Property

## Acknowledgment

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).