Network Working Group   B. Sarikaya
Internet-Draft   Huawei USA
Intended status: Standards Track   September 4, 2014
Expires: March 8, 2015

                  Overview of Source Address Dependent Routing
                      draft-sarikaya-6man-sadr-overview-00

Abstract

   This document presents an overview source address dependent routing.
   Different architectures are introduced and with their help, why
   source address dependent routing is needed is explained.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

   BCP 38 recommends ingress traffic routing to prohibit Denial of
   Service (DoS) attacks, i.e. datagrams which have source addresses
   that do not match with the network where the host is attached are
   discarded [RFC2827]. Avoiding packets to be dropped because of
   ingress filtering is difficult especially in multihomed networks
   where the host receives more than one prefix from the connected
   Internet Service Providers (ISP) and may have more than one source
   addresses. Based on BCP 38, BCP 84 introduced recommendations on the
   routing system for multihomed networks [RFC3704].

   Recommendations on the routing system for ingress filtering such as
   in BCP 84 inevitably involve source address checks. This leads us to
   the source address dependent routing. Source address dependent
   routing is an issue especially when the host is connected to a
   multihomed network and is communicating with another host in another
   multihomed network. In such a case, the communication can be broken
   in both directions if ISPs apply ingress filtering and the datagrams
   contain wrong source addresses
   [I-D.huitema-multi6-ingress-filtering].

   Hosts with simultaneously active interfaces receive multiple prefixes
   and have multiple source addresses. Datagrams originating from such
   hosts carry greats risks to be dropped due to ingress filtering.
   Source address selection algorithm needs to careful to try to avoid
   ingress filtering on the next-hop router [RFC6724].

   Many use cases have been reported for source/destination routing in
   [I-D.baker-rtgwg-src-dst-routing-use-cases]. These use cases clearly
   indicate that the multihomed host or Customer Premises Equipment
   (CPE) router needs to be configured with correct source prefixes/
   addresses so that it can route packets upstream correctly to avoid
   ingress filtering applied by an upstream ISP to drop the packets.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
"SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
document are to be interpreted as described in [RFC2119].

3. Source Address Dependent Routing

In multihomed networks there is a need to do source address based
routing if some providers are performing the ingress filtering
defined in BCP38 [RFC2827]. This requires the routers to consider
the source addresses as well as the destination addresses in
determining the next hop to send the packet to.

Based on the use cases defined in
[I-D.baker-rtgwg-src-dst-routing-use-cases], the routers may be
informed about the source addresses to use in routing using
extensions to the routing protocols like IS-IS defined in
[ISO.10589.1992] [I-D.baker-ipv6-isis-dst-src-routing] and OSPF
defined in [RFC5340] [I-D.baker-ipv6-ospf-dst-src-routing]. In this
document we describe the use cases for source address dependent
routing from the host perspective.

There are two cases. A host may have a single interface with
multiple addresses (from different prefixes or /64s). Each address
or prefix is connected to or coming from different exit routers, and
this case can be called multi-prefix multihoming (MPMH). A host may
have simultaneously connected multiple interfaces where each
interface is connected to a different exit router and this case can
be called multi-prefix multiple interface (MPMI).

Consistent set of network configuration information is called
provisioning domain (PvD). In case of multi-prefix multihoming
(MPMH), more than one provisioning domain is present on a single
link. In case of multi-prefix multiple interface (MPMI), elements of
the same domain may be present on multiple links. PvD aware nodes
support association of configuration information into PvDs and use
these PvDs to serve requests for network connections, e.g. chosing
the right source address for the packets. PvDs can be constructed
from one of more DHCP or Router Advertisement (RA) options carrying
such information as PvD identity [I-D.ietf-mif-mpvd-ndp-support].
PvDs constructed based on such information are called explicit PvDs
[I-D.ietf-mif-mpvd-arch].

Apart from PvD identity, PvD content may be defined in separate RA or
DHCP options. Examples of such content are defined in
[I-D.sarikaya-6man-next-hop-ra] and

[I-D.sarikaya-dhc-dhcpv6-raoptions-sadr]. They constitute the
content or parts of the content of explicit PvD.

Explicit PvDs may be received from different interfaces. Single PvD
may be accessible over one interface or simulatenously accessible
over multiple interfaces. Explicit PvDs may be scoped to a
configuration related to a particular interface, however in general
this may not apply. What matters is PvD ID provided that PvD ID is
authenticated by the node even in cases where the node has a single
connected interface. Single PvD information may be received over
multiple interfaces as long as PvD ID is the same. This applies to
the router advertisements (RAs) in which case a multi-homed host
(that is, with multiple interfaces) should trust a message from a
router on one interface to install a route to a different router on
another interface.

In DHCP, DHCP server can configure only the interface of the host to
which it is directly connected. For example, source address
selection rules distributed by [RFC7078] apply only the interface
over which the DHCP Option OPTION_ADDRSEL_TABLE is received. In
order for it to apply on other interfaces the option has to be sent
on those interfaces as well.

Default source address selection Rule 5.5 in [RFC6724] recommends to
select source addresses advertized by the next hop. These address
selection rules can be distributed site-wide using DHCP as in
[RFC7078]. Default routers informing next hop router addesses and
source prefixes supported by these next hops such as in
[I-D.sarikaya-6man-next-hop-ra] goes inline with this rule.

4. SADR Scenarios

The use case shown in Figure 1 is multi-prefix multi interface use
case where rtr1 and rtr2 represent customer premises equipment/
routers (CPE) and there are exit routers in both network 1 and
network 2. The issue in this case is ingress filtering. If the
packets from the host communicating with a remote destination are
routed to the wrong exit router, i.e. carry wrong source address will
get dropped.

Source address dependent routing can present a solution to this
problem. The solution should start with the correct configuration of
the host. The host should be configured with the next hop addresses
and the prefixes supported in these next hops. This way the host
having received many prefixes will have the correct knowledge in
selecting the right source address and next hop when sending packets
to remote destinations.

Host configuration can be made using either router advertisements or
DHCP. The choice depends on how the network is set up.

```
+------+      +------+         _____
|      |      |      |        /           \
|      |-----| rtr1 |=====/   network   \
|      |      |      |     \       1      /
|      |      +------+      _____/
|      |
| host |
|      |
|      |      +------+         _____
|      |      |      |        /           \
|      |=====| rtr2 |=====/   network   \
|      |      |      |     \       2      /
+------+      +------+      _____/
```
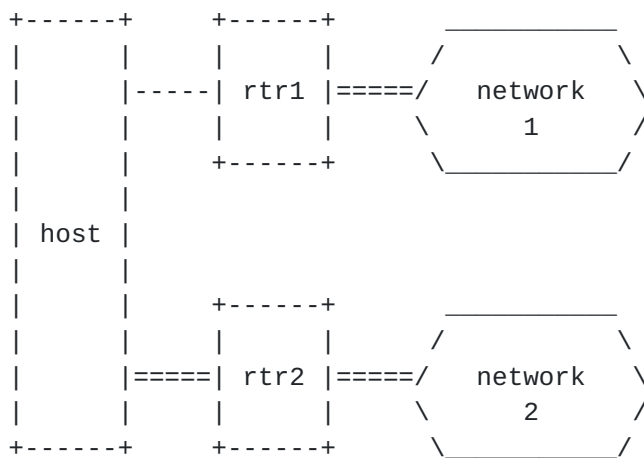
Figure 1: Multihomed Host with Two CPE Routers

Our next use case is shown in Figure 2. This use case is a multi-
prefix multihoming use case. rtr is CPEs which is connected to two
ISPs each advertising their own prefixes. In this case, the host may
have a single interface but it receives multiple prefixes from the
connected ISPs. Assuming that ISPs apply ingress filtering policy
the packets for any external communication from the host should
follow source address dependent routing in order to avoid getting
dropped.

In this configutation also it is useful to configure the host the
next hop addresses and the prefixes they support. This will enable
the host to select the right prefix when sending packets to the right
next hop and avoid any ingress filtering.

```
          +------+              |      +------+
          |      |              |      |      |
          |      |              |------|(ISP1)|=====
          |      |   +------+   |      |      |
          |      |   |      |   |      +------+
          |      |=====|  rtr  |=====|
          | host |   |      |   |
          |      |   +------+   |
          |      |              |      +------+
          |      |              |      |      |
          |      |              |=====|(ISP2)|=====
          |      |              |      |      |
          +------+              |      +------+
```

               Figure 2: Multihomed Host with Multiple CPE Routers

   A variation of this use case is specialized egress routing. Upstream
   networks offer different services with specific requirements, e.g.
   video service. The hosts using this service need to use the
   service's source and destination addresses. No other service will
   accept this source address, i.e. those packets will be dropped
   [I-D.baker-rtgwg-src-dst-routing-use-cases].

   Source address dependent routing in this use case may work as
   follows. The specialized service router advertize one or more
   specific prefixes with appropriate source prefixes, e.g. to the CPE
   Router, rtr in Figure 2. The CPE router in turn advertizes the
   specific service's prefixes and source prefixes to the host. This
   will allow proper configuration at the host so that the host can use
   the service by sending the packets with the correct source and
   destination addresses.

```
          _____           +------+
        /           \   +------+  |      |
       /   network   \  |      |  |      |
       \     1       /--| rtr1 |----|      |
        _____/   |      |  |      |
   _____          |      |  |      +------+
  host |     |    |    /           \
                        |    |=====| rtr3 |=====/    network   \
                        |    |     |      |    \        3      /
   _____          |    +------+     |    +------+   _____/
        /           \   +------+  |      |
       /   network   \  |      |  |      |
       \     1       /--| rtr1 |----|      |
        _____/   |      |  |      |
                        +------+  |      |
                                  +------+
```
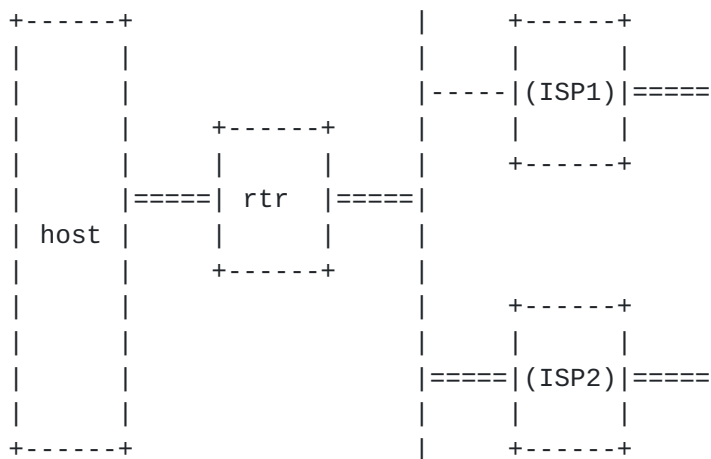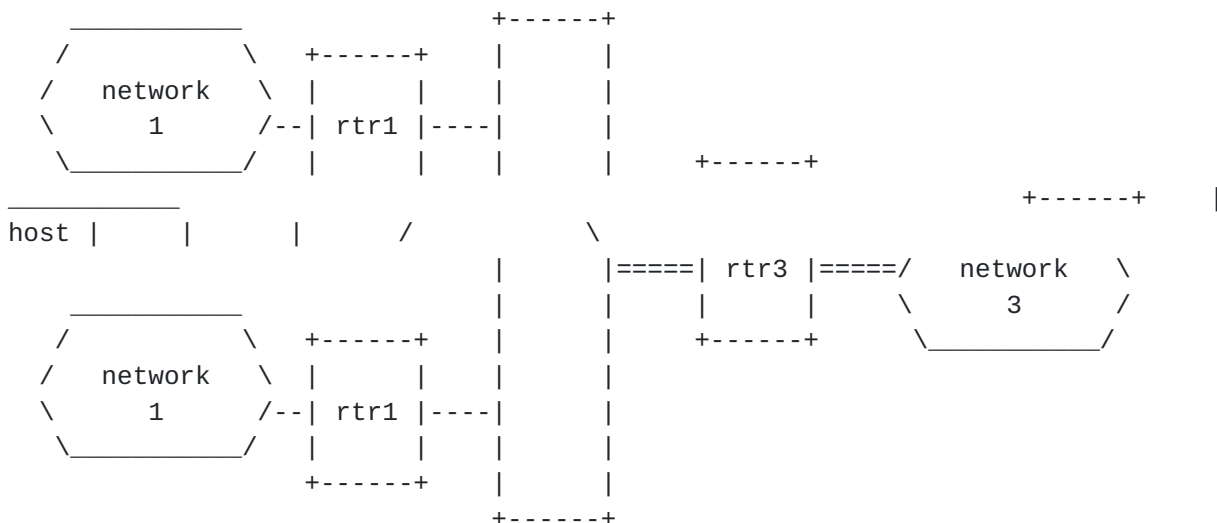
Figure 3: Multihomed Host with Three CPE Routers

Next use case is shown in Figure 3. It is a variation of multi-
prefix multi interface use case above. rtr1, rtr2 and rtr3 are CPE
Routers. The networks apply ingress routing. Source address
dependent routing should be used to avoid any external communications
be dropped.

```
    +------+                  |        +------+
    |      |                  |        |      |
    |      |                  |------| rtrF |=====ISP3
    |      |                  |        |      |
    |      |                  |        +------+
    |      |                  |
    | host |                  |
    |      |                  |
    |      |      +------+     |        +------+
    |      |      |      |     |        |      |===== ISP2
    |      |=====| rtr  |=====|=====| rtrE |
    |      |      |      |     |        |      |===== ISP1
    +------+      +------+     +        +------+
```
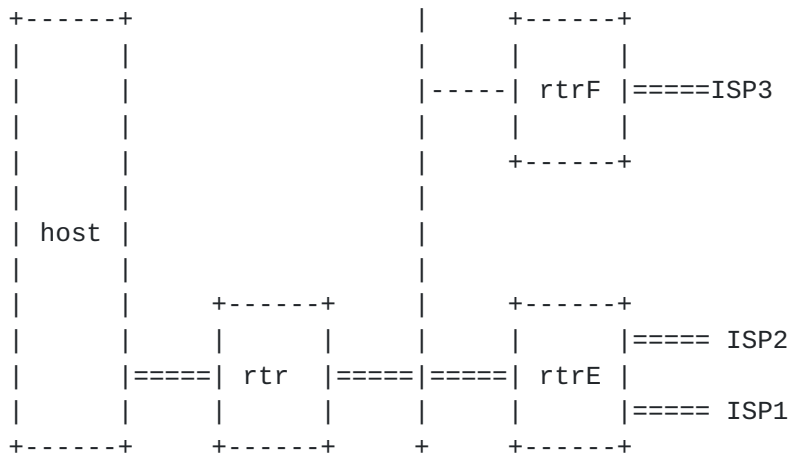
Figure 4: Shim6 Host with Two Routers

The last use case in Figure 4 is also a variation of multi-prefix
multihoming use case above. In this case rtrE is connected to two
ISPs. All ISPs are assumed to apply ingress routing. The host
receives prefixes from each ISP and starts communicating with
external hosts, e.g. H1, H2, etc. H1 and H2 may be accessible both
from ISP1 and ISP3.

The host receives multiple provider-allocated IPv6 address prefixes,
e.g. P1, P2 and P3 for ISP1, ISP2 and ISP3 and supports shim6
protocol [RFC5533]. rtr is a CPE router and the default router for
the host. rtr receives OSPF routes and has a default route for rtrE
and rtrF.

The host starts external communication with H1 and sends the first
packet with source address P3::iid. Since rtr has a default route to
rtrE it will use this default route in sending the host's packet out
towards rtrE. rtrE will route this packet to ISP1 and the packet will
be dropped due to the ingress filtering.

This use case shows that even though all the routers had source
address dependent routing support, the packet still got dropped. A
solution to this issue could be that rtrE having multiple routes to
H1 could use the path through rtrF and could direct the packet to the
other route, i.e. rtrF which would reach H1 in ISP3 without being
subject to ingress routing
[I-D.baker-6man-multiprefix-default-route].

5. Security Considerations

   This document describes some use cases and thus brings no new
   security risks to the Internet.

6. IANA Considerations

   None.

7. Acknowledgements

   In writing this document, the author benefited from face to face
   discussions he had with Brian Carpenter and Ole Troan.

8. References

8.1. Normative References

   [ISO.10589.1992]
              International Organization for Standardization,
              "Intermediate system to intermediate system intra-domain-
              routing routine information exchange protocol for use in
              conjunction with the protocol for providing the
              connectionless-mode Network Service (ISO 8473), ISO
              Standard 10589", ISO ISO.10589.1992, 1992.

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

   [RFC2629] Rose, M., "Writing I-Ds and RFCs using XML", RFC 2629,
              June 1999.

   [RFC2827] Ferguson, P. and D. Senie, "Network Ingress Filtering:
              Defeating Denial of Service Attacks which employ IP Source
              Address Spoofing", BCP 38, RFC 2827, May 2000.

   [RFC3704] Baker, F. and P. Savola, "Ingress Filtering for Multihomed
              Networks", BCP 84, RFC 3704, March 2004.

   [RFC3971] Arkko, J., Kempf, J., Zill, B., and P. Nikander, "SEcure
              Neighbor Discovery (SEND)", RFC 3971, March 2005.

   [RFC4191] Draves, R. and D. Thaler, "Default Router Preferences and
              More-Specific Routes", RFC 4191, November 2005.

   [RFC4605] Fenner, B., He, H., Haberman, B., and H. Sandick,
              "Internet Group Management Protocol (IGMP) / Multicast
              Listener Discovery (MLD)-Based Multicast Forwarding
              ("IGMP/MLD Proxying")", RFC 4605, August 2006.

   [RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman,
              "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861,
              September 2007.

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

   [RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF
              for IPv6", RFC 5340, July 2008.

   [RFC5533] Nordmark, E. and M. Bagnulo, "Shim6: Level 3 Multihoming
              Shim Protocol for IPv6", RFC 5533, June 2009.

   [RFC6724] Thaler, D., Draves, R., Matsumoto, A., and T. Chown,
              "Default Address Selection for Internet Protocol Version 6
              (IPv6)", RFC 6724, September 2012.

   [RFC7078] Matsumoto, A., Fujisaki, T., and T. Chown, "Distributing
              Address Selection Policy Using DHCPv6", RFC 7078, January
              2014.

8.2. Informative References

   [I-D.baker-6man-multiprefix-default-route]
              Baker, F., "Multiprefix IPv6 Routing for Ingress Filters",
              draft-baker-6man-multiprefix-default-route-00 (work in
              progress), November 2007.

   [I-D.baker-ipv6-isis-dst-src-routing]
              Baker, F., "IPv6 Source/Destination Routing using IS-IS",
              draft-baker-ipv6-isis-dst-src-routing-01 (work in
              progress), August 2013.

   [I-D.baker-ipv6-ospf-dst-src-routing]
              Baker, F., "IPv6 Source/Destination Routing using OSPFv3",
              draft-baker-ipv6-ospf-dst-src-routing-03 (work in
              progress), August 2013.

   [I-D.baker-rtgwg-src-dst-routing-use-cases]
              Baker, F., "Requirements and Use Cases for Source/
              Destination Routing", draft-baker-rtgwg-src-dst-routing-
              use-cases-00 (work in progress), August 2013.

   [I-D.huitema-multi6-ingress-filtering]
              Huitema, C., "Ingress filtering compatibility for IPv6
              multihomed sites", draft-huitema-multi6-ingress-
              filtering-00 (work in progress), October 2004.

   [I-D.ietf-mif-mpvd-arch]
              Anipko, D., "Multiple Provisioning Domain Architecture",
              draft-ietf-mif-mpvd-arch-03 (work in progress), August
              2014.

   [I-D.ietf-mif-mpvd-ndp-support]
              Korhonen, J., Krishnan, S., and S. Gundavelli, "Support
              for multiple provisioning domains in IPv6 Neighbor
              Discovery Protocol", draft-ietf-mif-mpvd-ndp-support-00
              (work in progress), August 2014.

   [I-D.sarikaya-6man-next-hop-ra]
              Sarikaya, B., "IPv6 RA Options for Next Hop Routes",
              draft-sarikaya-6man-next-hop-ra-02 (work in progress),
              June 2014.

   [I-D.sarikaya-dhc-dhcpv6-raoptions-sadr]
              Sarikaya, B., "DHCPv6 Route Options for Source Address
              Dependent Routing", draft-sarikaya-dhc-dhcpv6-raoptions-
              sadr-00 (work in progress), June 2014.

   [RFC6106] Jeong, J., Park, S., Beloeil, L., and S. Madanapalli,
              "IPv6 Router Advertisement Options for DNS Configuration",
              RFC 6106, November 2010.

Author's Address

   Behcet Sarikaya
   Huawei USA
   5340 Legacy Dr. Building 175
   Plano, TX 75024

   Email: sarikaya@ieee.org