Network Working Group
Internet-Draft
Intended status: Informational
Expires: December 11, 2015

B. Sarikaya
Huawei USA
M. Boucadair
France Telecom

June 9, 2015

Source Address Dependent Routing and Source Address Selection for IPv6 Hosts: Problem Space Overview

draft-sarikaya-6man-sadr-overview-07

Abstract

This document presents the source address dependent routing (SADR) problem space from the host perspective. Both multihomed hosts and hosts with multiple interfaces are considered. Several network architectures are presented to illustrate why source address selection and next hop resolution in view of source address dependent routing is needed.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

1.1. Overall Context

BCP 38 recommends ingress traffic routing to prohibit Denial-of-Service (DoS) attacks. As such, 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 networks it is connected to, and consequently 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 to the source address dependent routing (SADR). 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 Network Providers apply ingress filtering and the datagrams contain wrong source addresses (see for more details [I-D.huitema-multi6-ingress-filtering]).

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

Many use cases have been reported for source/destination routing, for example [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 forward packets upstream correctly to avoid ingress filtering applied by an upstream Network Provider to drop the packets.

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.

1.2. Scope

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 for forwarding using extensions to the routing protocols like IS-IS [ISO.10589.1992] [I-D.baker-ipv6-isis-dst-src-routing] or OSPF [RFC5340] [I-D.baker-ipv6-ospf-dst-src-routing].

In this document, we describe the scenarios for source address dependent routing from the host perspective. Two flavors can be considered:

1. A host may have a single interface with multiple addresses (from different prefixes or /64s). Each prefix is delegated from different exit routers, and this case can be called multi-prefix multihoming (MPMH).

2. 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).

Several limitations arise in such NAT- and NPTv6-based ([RFC6296]) multihoming contexts (see for example [RFC4116]). NPTv6 is left out of scope in this document.

2. Source Address Dependent Routing (SADR) Scenarios

SADR can be facilitated at the host with proper next-hop and source address selection. For this, each router connected to different interfaces of the host uses Router Advertisements (RAs, [RFC4861]) to distribute default route, next hop as well as source address/prefix information to the host. As a reminder, Route Information Option is defined in [RFC4191].

2.1. Scenario 1: Multi-Prefix Multi-Interface

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

```
      +------+     +------+       ___________
      |      |     |      |      /           \
      |      |-----| rtr1 |=====/   network   \
      |      |     |      |     \     1       /
      |      |     +------+      \___________/
      |      |
      | host |
      |      |
      |      |     +------+       ___________
      |      |     |      |      /           \
      |      |=====| rtr2 |=====/   network   \
      |      |     |      |     \     2       /
      +------+     +------+      \___________/
```

Figure 1: Multiple Interfaced Host with Two CPE Routers

There is a variant of Figure 1 that is often referred to as a corporate VPN, i.e., a secure tunnel from the host to a router attached to a corporate network. In this case "rtr2" gives access directly to the corporate network, and the link from the host to "rtr2" is a secure tunnel (for example an IPsec tunnel). The interface is therefore a virtual interface, with its own IP address/prefix assigned by the corporate network.

```
      +------+     +------+       ___________
      |      |-----| rtr1 |      /           \
      |     ==========\\  |=====/   network   \
      |      |-----|  ||  |     \      1      /
      |      |     +--||--+      \___________/
      |      |        ||
      | host |        ||
      |      |        ||
      |      |     +--||--+       ___________
      |      |     |      |      / corporate \
      |      |     | rtr2 |=====/   network   \
      |      |     |      |     \      2      /
      +------+     +------+      \___________/
```

Figure 2: VPN case

There are at least two sub-cases:

a. Dedicated forwarding entries are created in the host such that only traffic directed to the corporate network is sent to "rtr2"; everything else is sent to "rtr1".

b. All traffic is sent to "rtr2" and then routed to the Internet if necessary. This case doesn't need host routes but leads to unnecessary traffic and latency because of the path stretch via rtr2.

2.2. Scenario 2: Multi-Prefix Multihoming

Another scenario is shown in Figure 3. This one is a multi-prefix multihoming use case. "rtr" is CPE router which is connected to two Network Providers, each advertising their own prefixes. In this case, the host may have a single interface but it receives multiple prefixes from the connected Network Provider. Assuming that providers 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.

```
      +------+                  |
      |      |                  |        (Network)
      |      |                  |=====|(Provider 1)|=====
      |      |     +------+     |
      |      |     |      |     |
      |      |=====| rtr  |=====|
      | host |     |      |     |
      |      |     +------+     |
      |      |                  |
      |      |                  |        (Network)
      |      |                  |=====|(Provider 2)|=====
      |      |                  |
      +------+                  |
```

Figure 3: Multihomed Host with Multiple CPE Routers

2.3. Scenario 3: Service-specific Egress Routing

A variation of the scenario in Section 2.2 is: specialized egress routing. Upstream networks offer different services with specific requirements, e.g., VoIP or IPTV. 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].

```
    ___________                +------+
   /           \   +------+    |      |
  /   network   \  |      |    |      |
  \      1      /--| rtr1 |----|      |
   \___________/   |      |    |      |     +------+       ___________
                   +------+    | host |     |      |      /           \
                               |      |=====| rtr3 |=====/   network   \
                               |      |     |      |     \      3      /
    ___________                |      |     +------+      \___________/
   /           \   +------+    |      |
  /   network   \  |      |    |      |
  \      2      /--| rtr2 |----|      |
   \___________/   |      |    |      |
                   +------+    |      |
                               +------+
```

Figure 4: Multiple Interfaced Host with Three CPE Routers

The scenario shown in Figure 4 s a variation of multi-prefix multi interface scenario (Section 2.1). "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.

2.4. Scenario 4: Home Network (Homenet)

In the homenet scenario depicted in Figure 5, representing a simple home network, there is a host connected to two CPEs which are connected to providers 1 and 2, respectively. Each network delegates a different prefix. Also each router provides a different prefix to the host. The issue in this scenario is also ingress filtering used by each provider.

```
+------+
|      |     +------+
|      |     |      |      (Network)
|      |==+==| rtr1 |====|(Provider 1)|=====
|      |  |  |      |
|      |  |  +------+
| host |  |
|      |  |
|      |  |  +------+
|      |  |  |      |      (Network)
|      |  +==| rtr2 |====|(Provider 2)|=====
|      |     |      |
+------+     +------+
```

Figure 5: Simple Home Network with Two CPE Routers

The host has to select the source address from the prefixes of Providers 1 or 2 when communicating with other hosts in Provider 1 or 2. The next issue is to select the correct next hop router, rtr1 or rtr2 that can reach the correct provider, "Network Provider 1" or "Network Provider 2".

2.5. Scenario 5: Shim6 Host with Two Router

The last scenario shown in Figure 6 is also a variation of multi-prefix multihoming scenario (Section 2.2). In this case rtrE is connected to two providers. All network providers are assumed to apply ingress routing. The host receives prefixes from each provider and starts communicating with external hosts, e.g., H1, H2, etc. H1 and H2 may be accessible both from provider 1 and provider 3.

```
      +------+                  |     +------+
      |      |                  |     |      |      (Network)
      |      |                  |-----| rtrF |=====(Provider 3)
      |      |                  |     |      |
      |      |                  |     +------+
      |      |                  |
      | host |                  |
      |      |                  |
      |      |     +------+     |     +------+      (Network)
      |      |     |      |     |     |      |=====(Provider 2)
      |      |=====| rtr  |=====|=====| rtrE |      (Network)
      |      |     |      |     |     |      |=====(Provider 1)
      +------+     +------+     +     +------+
```

Figure 6: Shim6 Host with Two Routers

The host receives multiple provider-allocated IPv6 address prefixes, e.g., P1, P2 and P3 for providers 1, 2 and 3 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.

3. Analysis of Source Address Dependent Routing

In this section we present an analysis of the scenarios of Section 2 and then discuss the relevance of SADR to the provisioning domains.

3.1. Scenarios Analysis

As in [RFC7157] we assume that the routers in Section 2 use Router Advertisements to distribute default route and source address prefixes supported in each next hop to the hosts or the gateway/CPE router relays this information to the hosts.

Referring to the scenario in Figure 1, source address dependent routing can present a solution to the problem of the host wishes to reach a destination in network 2 and the host may choose rtr1 as the default router. The solution assumes the host is correctly configured. The host should be configured with 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.

Note that similar considerations apply to the scenario in Figure 4.

In the configuration of the scenario (Figure 3) it is also useful to configure the host with the prefixes and source address 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.

Source address dependent routing in the use case of specialized egress routing may work as follows. The specialized service router advertizes one or more specific prefixes with appropriate source prefixes, e.g., to the CPE Router, rtr in Figure 3. 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.

Let us analyze the scenario in Figure 5. If a source address dependent routing protocol is used, the two routers (rtr1 and rtr2) are both able to route traffic correctly, no matter which next-hop router and source address the host selects. In case the host chooses the wrong next hop router, e.g., for provider 2 rtr1 is selected, rtr1 will forward the traffic to rtr2 to be sent to network provider 2 and no ingress filtering will happen.

Note that home networks are expected to comply with requirements for source address dependent routing and the routers will be configured accordingly, no matter which routing protocol, e.g., OSPF is used [I-D.ietf-homenet-hncp].

This would work but with issues. The host traffic to provider 2 will have to go over two links instead of one, i.e., the link bandwidth will be halved. Another possibility is rtr1 can send an ICMPv6 Redirect message to the host to direct the traffic to rtr2. Host would redirect provider 2 traffic to rtr2.

The problem with redirects is that ICMPv6 Redirect message can only convey two addresses, i.e., in this case the router address, or rtr2 address and the destination address, or the destination host in provider 2. That means the source address will not be communicated. As a result, the host would send packets to the same destination using both source addresses which causes rtr2 to send a redirect message to rtr1, resulting in ping-pong redirects sent by rtr1 and rtr2.

The best solution to these issues is to configure the host with the source address prefixes that the next hop supports. In homenets, each interface of the host can be configured by its next hop router, so that all that is needed is to add the information on source address prefixes. This results in the hosts to select the right router no matter what.

Finally, the scenario in Figure 6 shows that even though all the routers may have source address dependent routing support, the packets still may get dropped.

The host in Figure 6 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 provider 1 and the packet will be dropped due to the ingress filtering.

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 Network Provider 3 without being subject to ingress routing [I-D.baker-6man-multiprefix-default-route].

3.2. Provisioning Domains and SADR

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) environments, 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., choosing 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 and PvD container [I-D.ietf-mif-mpvd-ndp-support], [I-D.ietf-mif-mpvd-dhcp-support]. PvDs constructed based on such information are called explicit PvDs [RFC7556].

Apart from PvD identity, PvD content may be encapsulated in separate RA or DHCP options called PvD Container Option. These options are placed in the container options of an explicit PvD.

Explicit PvDs may be received from different interfaces. Single PvD may be accessible over one interface or simultaneously 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. The authentication of the PvD ID should meet the level required by the node policy. 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.

4. Discussion on Candidate Solution Tracks

We presented many topologies in which a host with multiple interfaces or a multihomed host is connected to various networks or Network Providers which in turn may apply ingress routing. The scenario analysis in Section 3.1 showed that in order to avoid packets getting dropped due to ingress routing, source address dependent routing is needed. Also, source address dependent routing should be supported by routers throughout a site that has multiple exits.

In this section, we provide informative guidelines on different existing and future solutions vis a vis the scenarios presented in Section 2. We start with source address selection rule 5.5 ([RFC6724]) and the scenarios it solves and continue with solutions that state exactly what information hosts need in terms of new router advertisement options for correct source address selection in those scenarios.

4.1. Source Address Selection Rule 5.5

One possible solution is the default source address selection Rule 5.5 in [RFC6724] which recommends to select source addresses advertized by the next hop. Considering the above scenarios, we can state that this rule can solve the problem in Figure 1, Figure 3 and Figure 4.

In using Rule 5.5 the following guidelines should be kept in mind. Source address selection rules can be distributed by DHCP server using DHCP Option OPTION_ADDRSEL_TABLE defined in [RFC7078].

In case of DHCP based host configuration, DHCP server can configure only the interface of the host to which it is directly connected. In order for Rule 5.5 to apply on other interfaces the option should be sent on those interfaces as well using [RFC7078].

The default source address selection Rule 5.5 solves that problem when an application sends a packet with an unspecified source address. In the presence of two default routes, one route will be chosen, and Rule 5.5 will make sure the right source address is used.

When the application selects a source address, i.e., the source address is chosen before next-hop selection, even though the source address is a way for the application to select the exit point, in this case that purpose will not be served. In the presence of multiple default routes, one will be picked, ignoring the source

address which was selected by the application because it is known that IPv6 implementations are not required to remember which next-hops advertised which prefixes. Therefore, the next-hop router may not be the correct one, and the packets may be filtered.

This implies that the hosts should register which next-hop router announced each prefix.

4.2. Router Advertisement Option

There is a need to configure the host not only with the prefixes but also with the source prefixes the next hop routers support. Such a configuration may avoid the host getting ingress/egress policy error messages such as ICMP source address failure message.

If host configuration is done using router advertisement messages then there is a need to define new router advertisement options for source address dependent routing. These options include Route Prefix with Source Address/Prefix Option. Other options such as Next Hop Address with Route Prefix option and Next Hop Address with Source Address and Route Prefix option will be considered in Section 4.3.

As discussed in Section 3.1, the scenario in Figure 5 can be solved by defining a new router advertisement option.

If host configuration is done using DHCP then there is a need to define new DHCP options for Route Prefix with Source Address/Prefix. As mentioned above, DHCP server configuration is interface specific. New DHCP options for source address dependent routing such as route prefix and source prefix need to be configured for each interface separately.

The scenario in Figure 5 can be solved by defining a new DHCP option.

4.3. Router Advertisement Option Set

The source address selection rule 5.5 may possibly be a solution for selecting the right source addresses for each next hop but there are cases where the next hop routers on each interface of the host are not known by the host initially. Such use cases are out of scope. Guidelines for use cases that require router advertisement option set involving third party next hop addresses are also out of scope.

4.4. Other Solutions

So far we have singled out the scenario in Figure 6. All the above solutions do not work in this case. This brings us the issue of IP path probing [I-D.naderi-ipv6-probing].

For a given destination, the host selects a source address and a next hop and sends its packet. When the selected path fails, in case of IP probing, the host can probe all available paths until finding one that works.

The guideline in probing is SADR should be used, i.e., it is a necessary tool. Basically, SADR saves time in eliminating wrong paths, i.e. sending the packets to the wrong exit router. If SADR is not taken into account correctly the host will end up wasting resources trying to explore paths that are certain to fail.

5. Security Considerations

This document describes some use cases and thus brings no additional security risks. Solution documents should further elaborate on specific security considerations.

6. IANA Considerations

None.

7. Acknowledgements

In writing this document, we benefited from the ideas expressed by the electronic mail discussion participants on 6man Working Group: Brian Carpenter, Ole Troan, Pierre Pfister, Alex Petrescu, Ray Hunter, Lorenzo Colitti and others.

Pierre Pfister proposed the scenario in Figure 5 as well as some text for Rule 5.5.

The text on corporate VPN in Section 3 was provided by Brian Carpenter.

8. References

8.1. Normative References

[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.

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

[RFC6296] Wasserman, M. and F. Baker, "IPv6-to-IPv6 Network Prefix Translation", RFC 6296, June 2011.

[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. and D. Lamparter, "IPv6 Source/Destination Routing using IS-IS", draft-baker-ipv6-isis-dst-src-routing-02 (work in progress), October 2014.

[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-01 (work in progress), October 2014.

[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-homenet-hncp]
Stenberg, M., Barth, S., and P. Pfister, "Home Networking Control Protocol", draft-ietf-homenet-hncp-05 (work in progress), June 2015.

[I-D.ietf-mif-mpvd-dhcp-support]
Krishnan, S., Korhonen, J., and S. Bhandari, "Support for multiple provisioning domains in DHCPv6", draft-ietf-mif-mpvd-dhcp-support-01 (work in progress), March 2015.

[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-01 (work in progress), February 2015.

[I-D.naderi-ipv6-probing]
Naderi, H. and B. Carpenter, "Experience with IPv6 path probing", draft-naderi-ipv6-probing-01 (work in progress), April 2015.

[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.

[RFC4116] Abley, J., Lindqvist, K., Davies, E., Black, B., and V. Gill, "IPv4 Multihoming Practices and Limitations", RFC 4116, July 2005.

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

[RFC7157] Troan, O., Miles, D., Matsushima, S., Okimoto, T., and D. Wing, "IPv6 Multihoming without Network Address Translation", RFC 7157, March 2014.

[RFC7556] Anipko, D., "Multiple Provisioning Domain Architecture", RFC 7556, June 2015.

Authors' Addresses

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

Email: sarikaya@ieee.org

Mohamed Boucadair
France Telecom
Rennes 35000
France

Email: mohamed.boucadair@orange.com