Independent Submission
Request for Comments: 8043
Category: Informational
ISSN: 2070-1721

B. Sarikaya
Huawei USA
M. Boucadair
Orange

January 2017

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

Abstract

This document presents the source-address-dependent routing (SADR) problem space from the host's 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 are needed in view of source-address-dependent routing.

The document is scoped on identifying a set of scenarios for source-address-dependent routing from the host's perspective and analyzing a set of solutions to mitigate encountered issues. The document does not make any solution recommendations.

Status of This Memo

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

This is a contribution to the RFC Series, independently of any other RFC stream. The RFC Editor has chosen to publish this document at its discretion and makes no statement about its value for implementation or deployment. Documents approved for publication by the RFC Editor are not a candidate 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 http://www.rfc-editor.org/info/rfc8043.

Table of Contents

1. Introduction

1.1. Overall Context

BCP 38 recommends ingress traffic filtering to prohibit Denial-of-Service (DoS) attacks. As such, datagrams with source addresses that do not match with the network where the host is attached are discarded [RFC2827]. Preventing packets from being dropped due to 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 address. 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 source-address-dependent-routing (SADR). Source-address-dependent routing can be problematic, 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 the wrong source addresses (see [INGRESS_FIL] for more details).

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

Many use cases have been reported for source/destination routing -- see [SD_RTG]. These use cases clearly indicate that the multihomed host or Customer Premises Equipment (CPE) router needs to be configured with the correct source prefixes/addresses so that it can forward packets upstream correctly to prevent the ingress filtering applied by an upstream Network Provider from dropping the packets.

In multihomed networks, there is a need to enforce source-address-based routing if some providers are performing ingress filtering. This requires that the routers consider the source addresses as well as the destination addresses in determining the packet’s next hop.

1.2. Scope

Based on the use cases defined in [SD_RTG], 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] [SD_RTG_ISIS], OSPF [RFC5340] [SD_RTG_OSPF].

In this document, we describe the scenarios for source-address-dependent routing from the host's 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 "multihomed with multi-prefix" (MHMP). In such case, source address selection is performed by the host while source-dependent routing is enforced by an upstream router.

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). For this case, the host is required to support both source address selection and source-dependent routing to avoid the need for an upstream router to rewrite the IPv6 prefix.

Several limitations arise in multihoming contexts based on NAT and IPv6-to-IPv6 Network Prefix Translation (NPTv6) [RFC6296]; see, for example, [RFC4116]. NPTv6 is out of scope for this document.

This document was initially written to inform the community about the SADR problem space. It was updated to record the various sets of alternate solutions to address that problem space. The 6man WG consensus is documented in [RFC8028].

2. Source-Address-Dependent Routing (SADR) Scenarios

This section describes a set of scenarios to illustrate the SADR problem. Scenarios are listed in order of increasing complexity.

2.1. Multi-Prefix Multihoming

The scenario shown in Figure 1 is a multi-prefix multihoming use case. "rtr" is a CPE router that is connected to two Network Providers, each advertising its own prefixes. In this case, the host may have a single interface, but it receives multiple prefixes from the upstream Network Providers. Assuming that providers apply the 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 scenario, the host does not need to perform source-dependent routing; it only needs to perform source address selection.

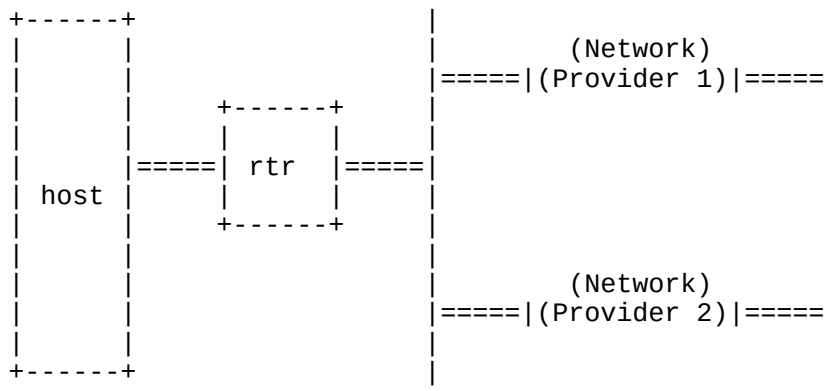

Figure 1: Multihomed Host with Multiple CPE Routers

2.2. Multi-Prefix Multi-Interface

The scenario shown in Figure 2 is multi-prefix multi-interface, where "rtr1" and "rtr2" represent CPE routers 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., they carry the wrong source address, they will get dropped due to ingress filtering.

In order to avoid complications when sending packets and to avoid the need to rewrite the source IPv6 prefix, the host is required to perform both source address selection and source-dependent routing so that the appropriate next hop is selected while taking into account the source address.

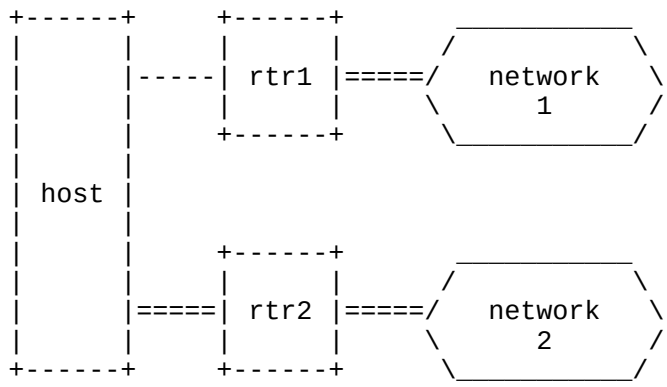

Figure 2: Multiple Interfaced Host with Two CPE Routers

There is a variant of Figure 2 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" provides access directly to the corporate network, and the link from the host to "rtr2" is a secure tunnel (for example, an IPsec tunnel). Therefore, the interface is a virtual interface with its own IP address/prefix assigned by the corporate network.

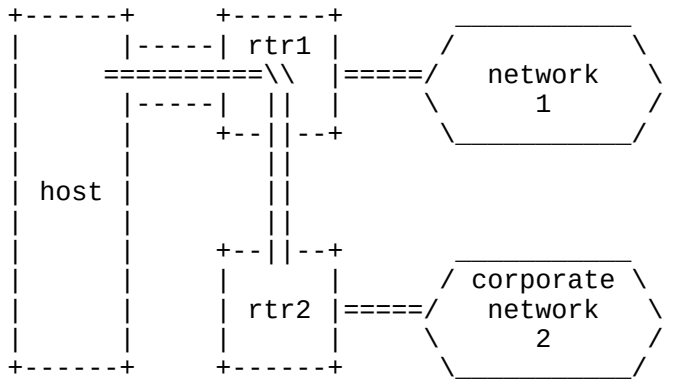

Figure 3: VPN Case

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".

2.3. Home Network (Homenet)

In the homenet scenario depicted in Figure 4, representing a simple home network, there is a host connected to a local network that is serviced with two CPEs that 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 that ingress filtering is used by each provider. This scenario can be considered as a variation of the scenario described in Section 2.2.

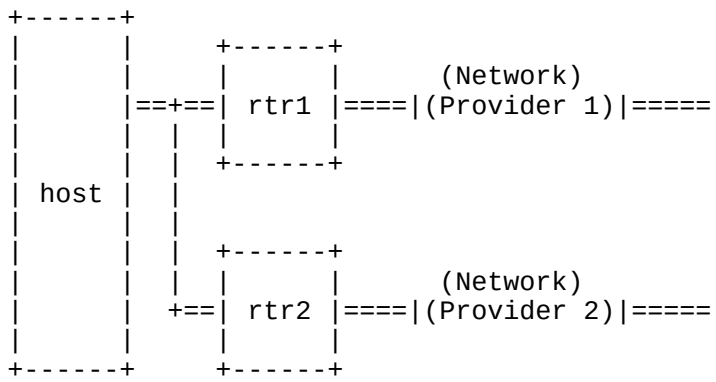

Figure 4: Simple Home Network with Two CPE Routers

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 2.

2.4. Service-Specific Egress Routing

A variation of the scenario in Section 2.1 is specialized egress routing. Upstream networks offer different services with specific requirements, e.g., Voice over IP (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 [SD_RTG].

Both source address selection and source-dependent routing are required to be performed by the host.

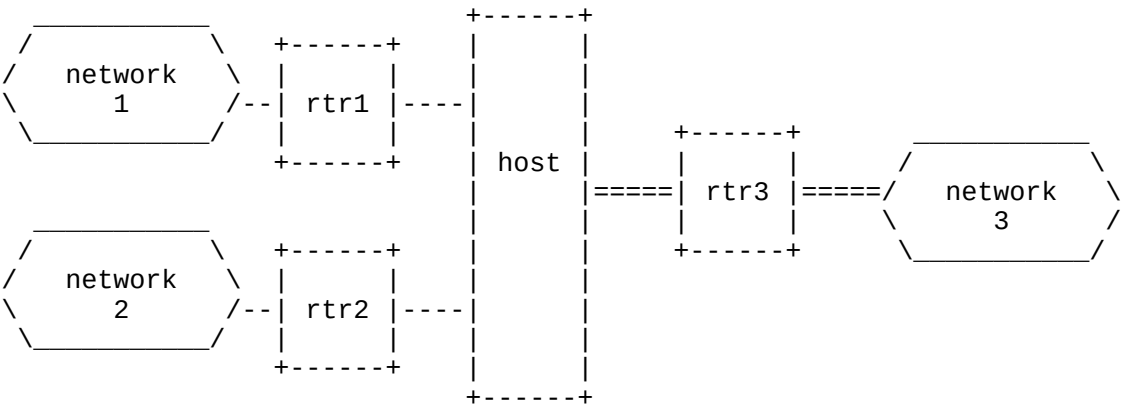

Figure 5: Multi-Interfaced Host with Three CPE Routers

The scenario shown in Figure 5 is a variation of a multi-prefix multi-interface scenario (Section 2.2). "rtr1", "rtr2", and "rtr3" are CPE routers. The networks apply ingress routing. Source-address-dependent routing should be used to avoid dropping any external communications.

3. Analysis of Source-Address-Dependent Routing

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

Section 3.1 presents an analysis of the scenarios in Section 2, and Section 3.2 discusses the relevance of SADR to the provisioning domains.

3.1. Scenarios Analysis

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

Referring to Section 2.1, source address selection is undertaken by the host while source-dependent routing must be followed by "rtr" to avoid packets being dropped. No particular modification is required for next-hop selection at the host.

Referring to the scenario in Figure 2, source-address-dependent routing can present a solution to the problem of when the host wishes to reach a destination in network 2 and the host chooses "rtr1" as the default router. The solution assumes that 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 information for selecting the right source address and next hop when sending packets to remote destinations.

Note that similar considerations apply to the scenario in Figure 5.

In the configuration of the scenario (Figure 1), 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 issues with ingress filtering.

Let us analyze the scenario in Section 2.3. 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., "rtr1" is selected for Provider 2, "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 that the routers will be configured accordingly no matter which routing protocol is used [RFC7788].

This would work, but with some 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 that "rtr1" can send an ICMPv6 Redirect message to the host to direct the traffic to "rtr2". The host would then redirect Provider 2 traffic to "rtr2".

The problem with redirects is that the 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 that 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".

A solution to these issues is to configure the host with the source address prefixes that the next hop supports. In a homenet context, each interface of the host can be configured by its next-hop router, so that all that is needed is to add the information about source address prefixes. This results in the hosts selecting the right router, no matter what.

Source-address-dependent routing in the use case of specialized egress routing (Section 2.4) may work as follows. The specialized service router advertises one or more specific prefixes with appropriate source prefixes, e.g., to the CPE router, "rtr" in Figure 1. The CPE router in turn advertises 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.

3.2. Provisioning Domains and SADR

A consistent set of network configuration information is called a provisioning domain (PvD). In the case of multihomed with multi-prefix (MHMP), more than one provisioning domain is present on a single link. In the 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 [MPvD_NDP] [MPvD_DHCP]. 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 the PvD Container Option. These options are placed in the container options of an explicit PvD.

Explicit PvDs may be received from different interfaces. A 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 does not apply. What matters is that the PvD identity 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 the PvD ID is the same.

This applies to the Router Advertisements (RAs) in which case a multihomed 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 of Alternate Solutions

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 shows that in order to prevent packets from being 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 egress points.

In this section, we provide some alternate solutions vis-a-vis the scenarios presented in Section 2. We start with Rule 5.5 in [RFC6724] for source address selection and the scenarios it solves, and then 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. No recommendation is made in this section.

4.1. Router Advertisement Option

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

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 the Route Prefix with Source Address/Prefix Option. Other options such as the Next-Hop Address with the Route Prefix Option and the Next-Hop Address with the Source Address and Route Prefix Option will be considered in Section 4.2.

As discussed in Section 3.1, the scenario in Figure 4 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 separately for each interface.

The scenario in Figure 4 can be solved by defining a new DHCP option.

4.2. Router Advertisement Option Set

Rule 5.5 for source address selection may 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 the Router Advertisement option set involving third-party next-hop addresses are also out of scope.

4.3. Rule 5.5 for Source Address Selection

One possible solution is Rule 5.5 in [RFC6724], the default rule for source address selection, which recommends selecting the source addresses advertised by the next hop. Considering the above scenarios, we can state that this rule can solve the problem in Figures 1, 2, and 5.

Source address selection rules can be distributed by the DHCP server using the DHCP option OPTION_ADDRSEL_TABLE defined in [RFC7078].

In case of DHCP-based host configuration, the 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 the DHCPv6 address selection policy option defined in [RFC7078].

Rule 5.5, the default rule for source address selection, 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 that 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 that 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. It is required that RAs be sent by the routers and that they contain PIO on all links. It is also required that the hosts remember the source addresses of the routers that sent

PIOs together with the prefixes advertised. This can be achieved by updating redirect rules specified in [RFC4861]. [RFC8028] further elaborates this to specify to which router a host should present its transmission.

The source-address-dependent routing solution is not complete without support from the edge routers. All routers in edge networks need to be required to support routing based on not only the destination address but also the source address. All edge routers need to be required to satisfy filters as defined in BCP 38.

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. References

6.1. 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, DOI 10.17487/RFC2827, May 2000, <http://www.rfc-editor.org/info/rfc2827>.

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

[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, <http://www.rfc-editor.org/info/rfc4861>.

[RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, DOI 10.17487/RFC5340, July 2008, <http://www.rfc-editor.org/info/rfc5340>.

[RFC6296] Wasserman, M. and F. Baker, "IPv6-to-IPv6 Network Prefix Translation", RFC 6296, DOI 10.17487/RFC6296, June 2011, <http://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, <http://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, <http://www.rfc-editor.org/info/rfc7078>.

[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, <http://www.rfc-editor.org/info/rfc8028>.

6.2. Informative References

[INGRESS_FIL] Huitema, C., Draves, R., and M. Bagnulo, "Ingress filtering compatibility for IPv6 multihomed sites", Work in Progress, draft-huitema-multi6-ingress-filtering-00, October 2004.

[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.

[MPvD_DHCP] Krishnan, S., Korhonen, J., and S. Bhandari, "Support for multiple provisioning domains in DHCPv6", Work in Progress, draft-ietf-mif-mpvd-dhcp-support-02, October 2015.

[MPvD_NDP] Korhonen, J., Krishnan, S., and S. Gundavelli, "Support for multiple provisioning domains in IPv6 Neighbor Discovery Protocol", Work in Progress, draft-ietf-mif-mpvd-ndp-support-03, February 2016.

[RFC4116] Abley, J., Lindqvist, K., Davies, E., Black, B., and V. Gill, "IPv4 Multihoming Practices and Limitations", RFC 4116, DOI 10.17487/RFC4116, July 2005, <http://www.rfc-editor.org/info/rfc4116>.

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

[RFC7157] Troan, O., Ed., Miles, D., Matsushima, S., Okimoto, T., and D. Wing, "IPv6 Multihoming without Network Address Translation", RFC 7157, DOI 10.17487/RFC7157, March 2014, <http://www.rfc-editor.org/info/rfc7157>.

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

[RFC7788] Stenberg, M., Barth, S., and P. Pfister, "Home Networking Control Protocol", RFC 7788, DOI 10.17487/RFC7788, April 2016, <http://www.rfc-editor.org/info/rfc7788>.

[SD_RTG] Baker, F., Xu, M., Yang, S., and J. Wu, "Requirements and Use Cases for Source/Destination Routing", Work in Progress, draft-baker-rtgwg-src-dst-routing-use-cases-02, April 2016.

[SD_RTG_ISIS] Baker, F. and D. Lamparter, "IPv6 Source/Destination Routing using IS-IS", Work in Progress, draft-baker-ipv6-isis-dst-src-routing-06, October 2016.

[SD_RTG_OSPF] Baker, F., "IPv6 Source/Destination Routing using OSPFv3", Work in Progress, draft-baker-ipv6-ospf-dst-src-routing-03, August 2013.

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 4 as well as some text for Rule 5.5.

The text on corporate VPN in Section 2 was provided by Brian Carpenter.

Authors' Addresses

Behcet Sarikaya
Huawei USA
5340 Legacy Dr. Building 175
Plano, TX 75024
United States of America

Email: sarikaya@ieee.org

Mohamed Boucadair
Orange
Rennes 35000
France

Email: mohamed.boucadair@orange.com