

Network Working Group
Internet-Draft
Intended status: Informational
Expires: April 1, 2017

B. Sarikaya
Huawei USA
M. Boucadair
Orange
September 28, 2016

Source Address Dependent Routing and Source Address Selection for IPv6
Hosts: Problem Space Overview
draft-sarikaya-6man-sadr-overview-12

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.

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

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 1, 2017.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	2
1.1.	Overall Context	2
1.2.	Scope	3
2.	Source Address Dependent Routing (SADR) Scenarios	4
2.1.	Multi-Prefix Multihoming	4
2.2.	Multi-Prefix Multi-Interface	5
2.3.	Home Network (Homenet)	6
2.4.	Service-specific Egress Routing	7
3.	Analysis of Source Address Dependent Routing	8
3.1.	Scenarios Analysis	8
3.2.	Provisioning Domains and SADR	10
4.	Discussion on Alternate Solutions	11
4.1.	Router Advertisement Option	11
4.2.	Router Advertisement Option Set	12
4.3.	Source Address Selection Rule 5.5	12
5.	Security Considerations	13
6.	Acknowledgements	13
7.	References	13
7.1.	Normative References	13
7.2.	Informative References	14
	Authors' Addresses	16

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 enforce source address based routing if some providers are performing the ingress filtering. 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). In such case, source address selection is

performed by the host while source-dependent routing is to be 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 requires to support both source address selection and source-dependent routing to avoid the need to rewrite the IPv6 prefix by an upstream router.

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

This document was initially written to inform the community about the SADR problem space. It was updated to record the various set of alternate solutions to address that problem space. The 6man consensus is documented in [I-D.ietf-6man-multi-homed-host].

2. Source Address Dependent Routing (SADR) Scenarios

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

2.1. Multi-Prefix Multihoming

The scenario shown in Figure 1 is a multi-prefix multihoming use case. "rtr" is a 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 upstream Network Providers. 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.

In this scenario, the host does not need to perform source-dependent routing; it does only need to perform source address selection.

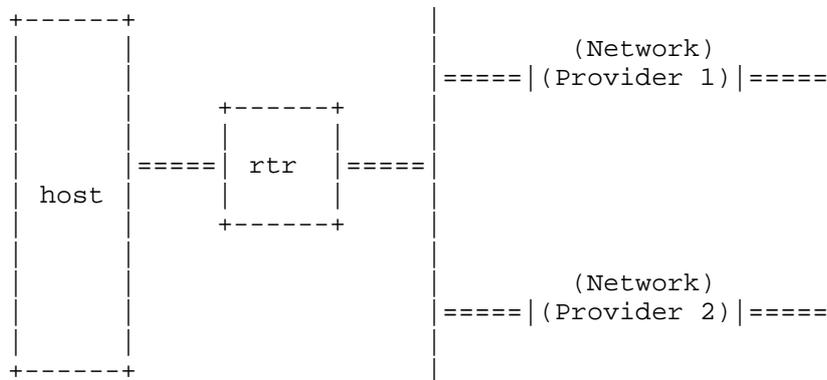


Figure 1: Multihomed Host with Multiple CPE Routers

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., carry wrong source address, they will get dropped due to ingress filtering.

In order to avoid complications to send packets and avoid a need to rewrite the source IPv6 prefix, the host requires to perform both source address selection and source-depending routing so that appropriate next-hop is selected taking into account the source address.

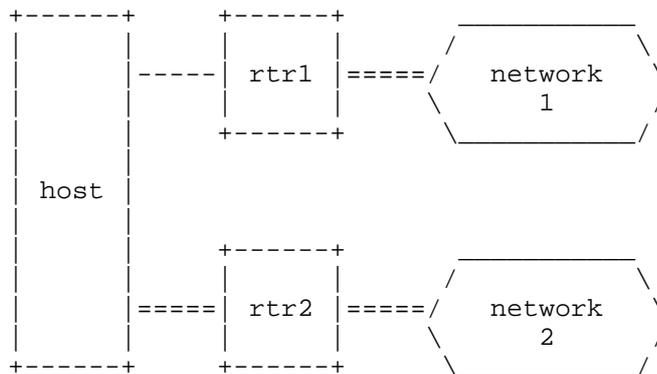


Figure 2: Multiple Interfaced Host with Two CPE Routers

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" 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.

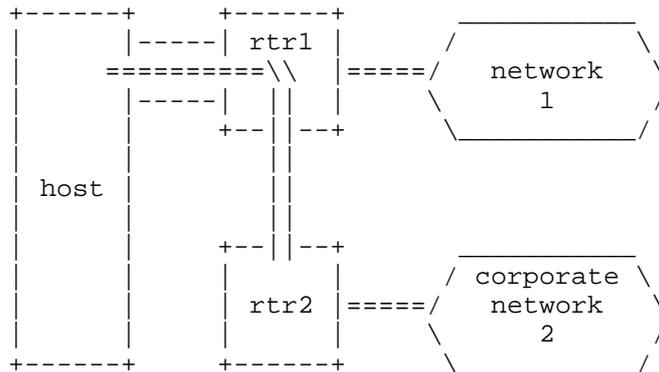


Figure 3: VPN case

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 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. This scenario can be considered as a variation of the scenario described in Section 2.2.

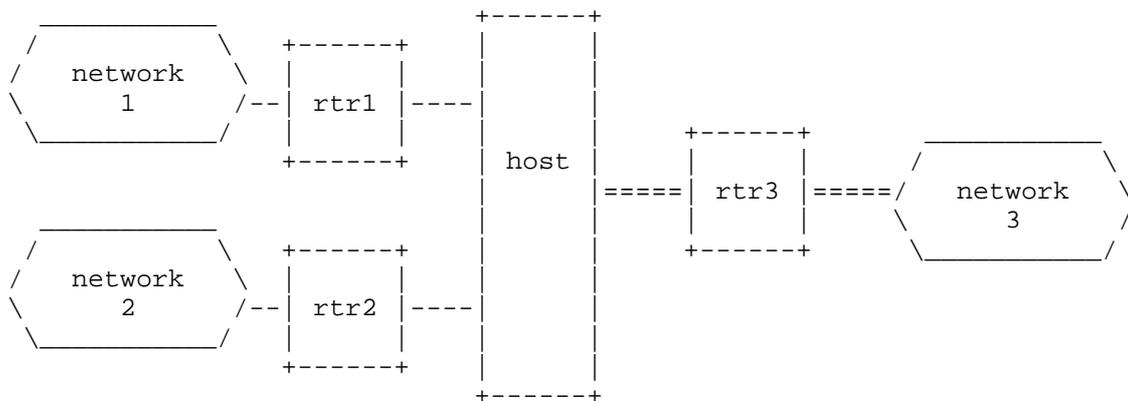


Figure 5: Multiple Interfaced Host with Three CPE Routers

The scenario shown in Figure 5 is a variation of 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 any external communications be dropped.

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, next hop as well as source address/prefix information to the host. As a reminder, while Prefix Information Option (PIO) is defined in [RFC4861], Route Information Option (RIO) is defined in [RFC4191].

Section 3.1 presents an analysis of the scenarios of Section 2 and then 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 Router Advertisements (RAs) 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 Section 2.1, source address selection is undertaken by the host while source-dependent routing must be followed by "rtr" to avoid packets drop. 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 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 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 ingress filtering issues.

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., 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 is used [RFC7788].

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.

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 on source address prefixes. This results in the hosts to select 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

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 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 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 egress points.

In this section, we provide some alternate 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. 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 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 4.2.

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 for each interface separately.

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

4.2. 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.

4.3. 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 advertised by the next hop. Considering the above scenarios, we can state that this rule can solve the problem in Figure 2, Figure 1 and Figure 5.

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. 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].

[I-D.ietf-6man-multi-homed-host] further elaborates this to specify to which router a host should present its transmission.

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 BCP 38 filters.

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. 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 3 was provided by Brian Carpenter.

7. References

7.1. Normative References

- [I-D.ietf-6man-multi-homed-host]
Marcon, J. and B. Carpenter, "First-hop router selection by hosts in a multi-prefix network", draft-ietf-6man-multi-homed-host-09 (work in progress), August 2016.
- [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>>.
- [RFC5533] Nordmark, E. and M. Bagnulo, "Shim6: Level 3 Multihoming Shim Protocol for IPv6", RFC 5533, DOI 10.17487/RFC5533, June 2009, <<http://www.rfc-editor.org/info/rfc5533>>.
- [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>>.

7.2. 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-05 (work in progress), April 2016.
- [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., Xu, M., Yang, S., and J. Wu, "Requirements and Use Cases for Source/Destination Routing", draft-baker-rtgwg-src-dst-routing-use-cases-02 (work in progress), April 2016.

- [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-dhcp-support]
Krishnan, S., Korhonen, J., and S. Bhandari, "Support for multiple provisioning domains in DHCPv6", draft-ietf-mif-mpvd-dhcp-support-02 (work in progress), October 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-03 (work in progress), February 2016.
- [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, 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>>.

Authors' Addresses

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

Email: sarikaya@ieee.org

Mohamed Boucadair
Orange
Rennes 35000
France

Email: mohamed.boucadair@orange.com