Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: June 11, 2015

B. Sarikaya
Huawei USA

December 8, 2014

Overview of Source Address Dependent Routing
draft-sarikaya-6man-sadr-overview-03

Abstract

This document presents an overview of source address dependent routing from the host perspective. Multihomed hosts and hosts with multiple interfaces are considered. Different architectures are introduced and with their help, why source address selection and next hop resolution in view of source address dependent routing is needed is explained. The document concludes with a discussion on the standardization work that is needed.

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 June 11, 2015.

Copyright Notice

Copyright (c) 2014 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
2. Terminology 3
3. SADR Scenarios 4
4. Analysis of Source Address Dependent Routing 7
4.1. Scenarios Analysis 7
4.2. Provisioning Domains and SADR 9
5. What Needs to be Done 9
6. Security Considerations 10
7. IANA Considerations 10
8. Acknowledgements 11
9. References 11
9.1. Normative References 11
9.2. Informative References 12
Author's Address 14

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 be 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.

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).

It should be noted that Network Address and Port Translation (NAPT) [RFC3022] in IPv4 and IPv6-to-IPv6 Network Prefix Translation (NPTv6) [RFC6296] in IPv6 implement the functions of source address selection and next-hop resolution and as such they address multihoming (and hosts with multiple interfaces) requirements arising from source address dependent routing [RFC7157]. In this case, the gateway router or CPE router does the source address and next hop selection for all the hosts connected to the router. However, for end-to-end connectivity, NAPT and NPTv6 should be avoided and because of this, NAPT and NPTv6 are left out of scope in this document.

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. SADR Scenarios

Source address dependent routing 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 to distribute default route, next hop as well as source address/prefix information to the host.

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, they will get dropped.

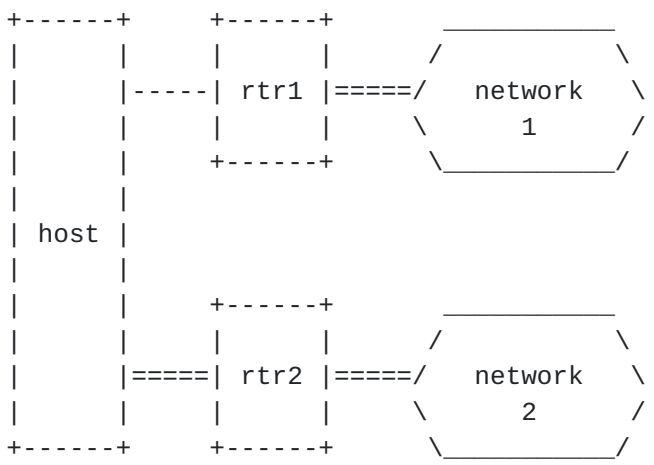


Figure 1: Multihomed Host with Two CPE Routers

Our next use case is shown in Figure 2. This use case is a multi-prefix multihoming use case. rtr is CPE router 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.

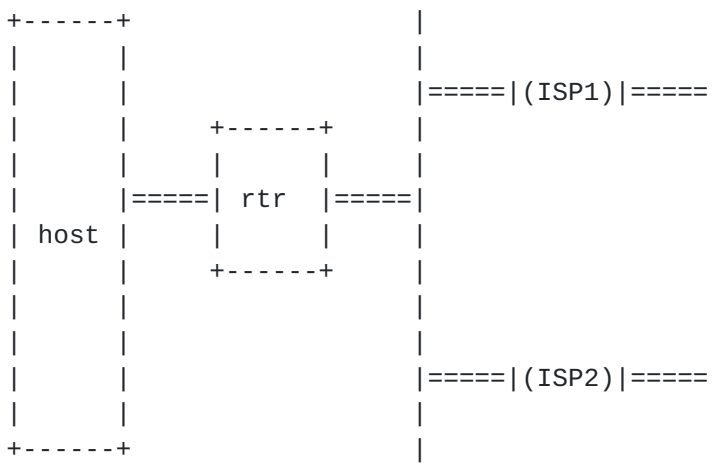


Figure 2: Multihomed Host with Multiple CPE Routers

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].

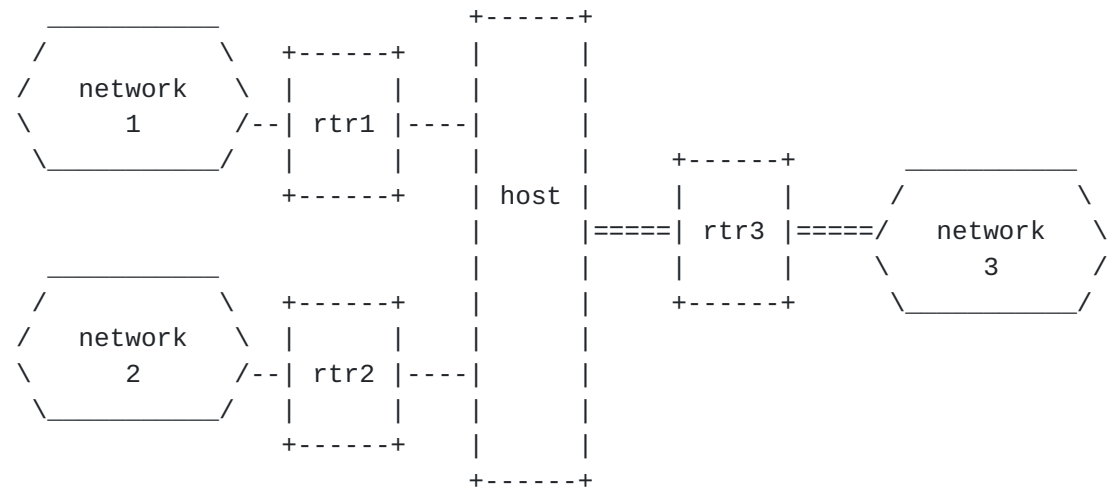


Figure 3: Multihomed Host with Three CPE Routers

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.

In the homenet scenario given in Figure 4, there is a host connected to two CPEs which are connected to ISP1 and ISP2, respectively. Each ISP provides 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 ISP.

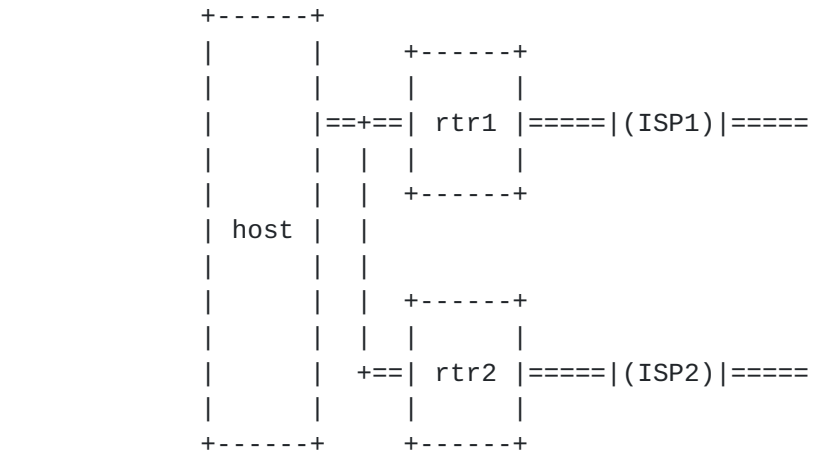


Figure 4: Home Network with Two CPE Routers

The host has to select the source address from the prefixes of ISP1 or ISP2 when communicating with other hosts in ISP1 or ISP2. The next issue is to select the correct next hop router, rtr1 or rtr2 that can reach the right ISP, ISP1 or ISP2.

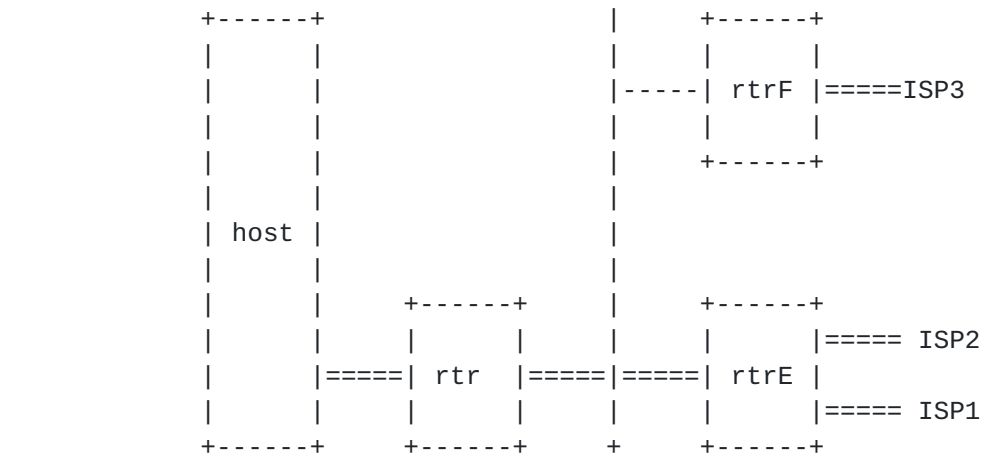


Figure 5: Shim6 Host with Two Routers

The last use case in Figure 5 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.

4. Analysis of Source Address Dependent Routing

In this section we present an analysis of the scenarios of Section 3 and then discuss the relevance of SADR to the provisioning domains.

4.1. Scenarios Analysis

As in [RFC7157] we assume that the routers in Section 3 use Router Advertisements to distribute default route, next hop and source address prefixes supported in each next hop to the hosts or the gateway/CPE router relayes 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 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.

Note that similar considerations apply to the scenario in Figure 3.

In the configuration of the scenario in Figure 2 also it is useful to configure the host with the next hop addresses and 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 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.

Let us analyze the use case in Figure 4. 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 ISP2 rtr1 is selected, rtr1 will forward the traffic to rtr2 to be sent to ISP2 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 ISP2 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 ISP2 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 ISP2. 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 both the next hop and the source address prefixes that the next hop supports. This results in the hosts to select the right router no matter what.

Finally, the use case in Figure 5 shows that even though all the routers may have source address dependent routing support, the packets still may get dropped.

The host in Figure 5 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.

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].

4.2. 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. 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 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 [I-D.ietf-mif-mpvd-arch].

Apart from PvD identity, PvD content may be encapsulated in separate RA or DHCP options called PvD Container Option. 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 an 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.

5. What Needs to be Done

We presented many topologies in which a host with multiple interfaces or a multihomed host is connected to various networks or ISPs which in turn may apply ingress routing. Our scenario analysis showed that in order to avoid packets getting dropped due to ingress routing, source address dependent routing is needed.

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. Source address selection rules can be distributed by DHCP server using DHCP Option OPTION_ADDRSEL_TABLE defined in [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.

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 has to be sent on those interfaces as well using [RFC7078].

There is a need to configure the host not only with the next hops and their prefixes but also with the source prefixes they 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 Next Hop Address with Route Prefix option and Next Hop Address with Source Address and Route Prefix option.

If host configuration is done using DHCP then there is a need to define new DHCP options for source address dependent routing. As mentioned above, DHCP server configuration is interface specific. New DHCP options for source address dependent routing such as orute prefix, next hop address and source prefix need to be configured for each interface separately.

6. Security Considerations

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

7. IANA Considerations

None.

8. Acknowledgements

In writing this document, the author benefited from face to face discussions he had with Brian Carpenter and Ole Troan. Pierre Pfister proposed the scenario in Figure 4 as well as some text for Rule 5.5.

9. References

9.1. Normative References

[I-D.ietf-homenet-hncp]
Stenberg, M. and S. Barth, "Home Networking Control Protocol", draft-ietf-homenet-hncp-02 (work in progress), October 2014.

[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.

[RFC3022] Srisuresh, P. and K. Egevang, "Traditional IP Network Address Translator (Traditional NAT)", RFC 3022, January 2001.

[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.

[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.

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

9.2. 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-mif-mpvd-arch]
Anipko, D., "Multiple Provisioning Domain Architecture", draft-ietf-mif-mpvd-arch-07 (work in progress), October 2014.

[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-00 (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-03 (work in progress), November 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