

IPv6 Operations
Internet-Draft
Intended status: Informational
Expires: January 3, 2018

J. Linkova
Google
M. Stucchi
July 2, 2017

Using Conditional Router Advertisements for Enterprise Multihoming
draft-linkova-v6ops-conditional-ras-01

Abstract

This document discusses most common scenarios of connecting an enterprise network to multiple ISPs using an address space assigned by an ISP. The problem of enterprise multihoming without address translation of any form has not been solved yet as it requires both the network to select the correct egress ISP based on the packet source address and hosts to select the correct source address based on the desired egress ISP for that traffic.

[I-D.ietf-rtgwg-enterprise-pa-multihoming] proposes a solution to this problem by introducing a new routing functionality (Source Address Dependent Routing) to solve the uplink selection issue and using Router Advertisements to influence the host source address selection. While the above-mentioned document focuses on solving the general problem and on covering various complex use cases, this document describes how the solution proposed in

[I-D.ietf-rtgwg-enterprise-pa-multihoming] can be adopted for limited number of common use cases. In particular, the focus is on scenarios where an enterprise network has two Internet uplinks used either in primary/backup mode or simultaneously and hosts in that network might not yet properly support multihoming as described in [RFC8028].

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 January 3, 2018.

Copyright Notice

Copyright (c) 2017 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. Common Enterprise Multihoming Scenarios	3
2.1. Two ISP Uplinks, Primary and Backup	3
2.2. Two ISP Uplinks, Used for Load Balancing	4
3. Conditional Router Advertisements	4
3.1. Solution Overview	4
3.1.1. Uplink Selection	4
3.1.2. Source Address Selection and Conditional RAs	4
3.2. Example Scenarios	6
3.2.1. Single Router, Primary/Backup Uplinks	6
3.2.2. Two Routers, Primary/Backup Uplinks	7
3.2.3. Single Router, Load Balancing Between Uplinks	9
3.2.4. Two Router, Load Balancing Between Uplinks	10
3.2.5. Topologies with Dedicated Border Routers	10
4. IANA Considerations	12
5. Security Considerations	12
5.1. Privacy Considerations	12
6. Acknowledgements	12
7. References	12
7.1. Normative References	12
7.2. Informative References	14
Appendix A. Change Log	15
Authors' Addresses	15

1. Introduction

Multihoming is an obvious requirement for many enterprise networks to ensure the desired level of network reliability. However, using more than one ISP (and address space assigned by those ISPs) introduces the problem of assigning IP addresses to hosts. In IPv4 there is no choice but using [RFC1918] address space and NAT ([RFC3022]) at the

network edge. Using Provider Independent or PI address space is not always an option as it requires running BGP between the enterprise network and the ISPs). As IPv6 host can, by design, have multiple addresses of the global scope, multihoming using provider address looks even easier for IPv6: each ISP assigns an IPv6 block (usually /48) and hosts in the enterprise network have addresses assigned from each ISP block. However using IPv6 PA blocks in multihoming scenario introduces some challenges, including but not limited to:

- o Selecting the correct uplink based on the packet source address;
- o Signaling to hosts that some source addresses should or should not be used (e.g. an uplink to the ISP went down or became available again).

The document [I-D.ietf-rtgwg-enterprise-pa-multihoming] discusses these and other related challenges in details in relation to the general multihoming scenario for enterprise networks. Unfortunately the proposed solution heavily relies on the rule 5.5 of the default address selection algorithm ([RFC6724]) which has not been widely implemented at the moment this document was written. Therefore network administrators in enterprise networks can't yet assume that all devices in their network support the rule 5.5, especially in the quite common BYOD ("Bring Your Own Device") scenario. However, while it does not seem feasible to solve all the possible multihoming scenarios without relying on rule 5.5, it is possible to provide IPv6 multihoming using provider-assigned (PA) address space for the most common use cases. This document discusses how the general solution described in [I-D.ietf-rtgwg-enterprise-pa-multihoming] can be applied to those two specific cases.

2. Common Enterprise Multihoming Scenarios

2.1. Two ISP Uplinks, Primary and Backup

This scenario has the following key characteristics:

- o The enterprise network is using uplinks to two (or more) ISPs for Internet access;
- o Each ISP assigns IPv6 PA address space for the network;
- o Uplink(s) to one ISP is a primary (preferred) one. All other uplinks are backup and are not expected to be used while the primary one is operational;
- o If the primary uplink is operational, all Internet traffic should flow via that uplink;

- o When the primary uplink fails the Internet traffic needs to flow via the backup uplinks;
- o Recovery of the primary uplink needs to trigger the traffic switchover from the backup uplinks back to primary one.

2.2. Two ISP Uplinks, Used for Load Balancing

This scenario has the following key characteristics:

- o The enterprise network is using uplinks to two (or more) ISPs for Internet access;
- o Each ISP assigns an IPv6 PA address space;
- o All the uplinks may be used simultaneously, with the traffic being randomly balanced between them.

3. Conditional Router Advertisements

3.1. Solution Overview

3.1.1. Uplink Selection

As discussed in [I-D.ietf-rtgwg-enterprise-pa-multihoming], one of the two main problems to be solved in the enterprise multihoming scenario is the problem of the next-hop (uplink) selection based on the packet source address. For example, if the enterprise network has two uplinks, to ISP_A and ISP_B, and hosts have addresses from subnet_A and subnet_B (belonging to ISP_A and ISP_B respectively) then packets sourced from subnet_A must be sent to ISP_A uplink while packets sourced from subnet_B must be sent to ISP_B uplink.

While some work is being done in the Source Address Dependent Routing (SADR) area, the simplest way to implement the desired functionality currently is to apply a policy which selects a next-hop or an egress interface based on the packet source address. Most of the SMB/Enterprise grade routers have such functionality available currently.

3.1.2. Source Address Selection and Conditional RAs

Another problem to be solved in the multihoming scenario is the source address selection on hosts. In the normal situation (all uplinks are up/operational) hosts have multiple global unique addresses and can rely on the default address selection algorithm ([RFC6724]) to pick up a source address, while the network is responsible for choosing the correct uplink based on the source address selected by a host as described in Section 3.1.2. However,

some network topology changes (i.e. changing uplink status) might affect the global reachability for packets sourced from the particular prefixes and therefore such changes have to be signaled back to the hosts. For example:

- o An uplink to an ISP_A went down. Hosts should not use addresses from ISP_A prefix;
- o A primary uplink to ISP_A which was not operational has come back up. Hosts should start using the source addresses from ISP_A prefix.

[I-D.ietf-rtgwg-enterprise-pa-multihoming] provides a detailed explanation on why SLAAC and router advertisements are the most suitable mechanism for signaling network topology changes to hosts and thereby influencing the source address selection. Sending a router advertisement to change the preferred lifetime for a given prefix provides the following functionality:

- o deprecating addresses (by sending an RA with the preferred_lifetime set to 0 in the corresponding POI) to indicate to hosts that that addresses from that prefix should not be used;
- o making a previously unused (deprecated) prefix usable again (by sending an RA containing a POI with non-zero preferred lifetime) to indicate to hosts that addresses from that prefix can be used again.

To provide the desired functionality, first-hop routers are required to

- o send RA triggered by defined event policies in response to uplink status change event; and
- o while sending periodic or solicited RAs, set the value in the given RA field (e.g. PIO preferred lifetime) based on the uplink status.

The exact definition of the 'uplink status' depends on the network topology and may include conditions like:

- o uplink interface status change;
- o presence of a particular route in the routing table;
- o presence of a particular route with a particular attribute (next-hop, tag etc) in the routing table;

- o protocol adjacency change.

etc.

In some scenarios, when two routers are providing first-hop redundancy via VRRP, the master-backup status can be considered as a condition for sending RAs and changing the preferred lifetime value. See Section 3.2.2 for more details.

If hosts are provided with ISP DNS servers IPv6 addresses via RDNSS [RFC8106] it might be desirable for the conditional RAs to update the Lifetime field of the RDNSS option as well.

3.2. Example Scenarios

This section illustrates how the conditional RAs solution can be applied to most common enterprise multihoming scenarios.

3.2.1. Single Router, Primary/Backup Uplinks

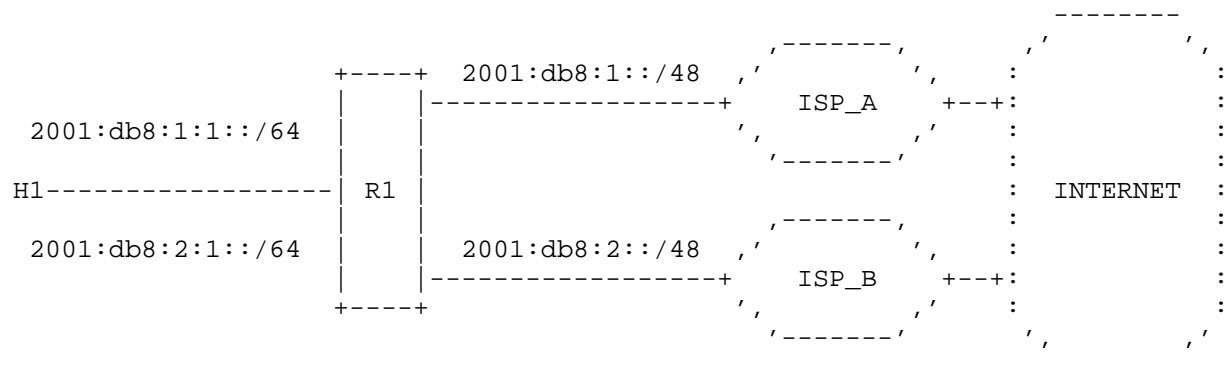


Figure 1: Single Router, Primary/Backup Uplinks

Let's look at a simple network topology where a single router acts as a border router to terminate two ISP uplinks and as a first-hop router for hosts. Each ISP assigns a /48 to the network, and the ISP_A uplink is a primary one, to be used for all Internet traffic, while the ISP_B uplink is a backup, to be used only when the primary uplink is not operational.

To ensure that packets with source addresses from ISP_A and ISP_B are only routed to ISP_A and ISP_B uplinks respectively, the network administrator needs to configure a policy on R1:

```
if {
    packet_destination_address is not in 2001:db8:1::/48 or 2001:db8:2::/48
    packet_source_address is in 2001:db8:1::/48
} then {
    next-hop is ISP_A_uplink
}
if {
    packet_destination_address is not in 2001:db8:1::/48 or 2001:db8:2::/48
    packet_source_address is in 2001:db8:2::/48
}
then {
    next-hop is ISP_B_uplink
}
```

Under normal circumstances it is desirable that all traffic be sent via the ISP_A uplink, therefore hosts (the host H1 in the example topology figure) should be using source addresses from 2001:db8:1:1::/64. When/if ISP_A uplink fails, hosts should stop using the 2001:db8:1:1::/64 prefix and start using 2001:db8:2:1::/64 until the ISP_A uplink comes back up. To achieve the desired behavior the router advertisement configuration on the R1 device for the interface facing H1 needs to have the following policy:

```
prefix 2001:db8:1:1::/64 {
    if ISP_A_uplink is up
        then preferred_lifetime = 604800
    else preferred_lifetime = 0
}

prefix 2001:db8:2:1::/64 {
    if ISP_A_Uplink is up
        then preferred_lifetime = 0
    else preferred_lifetime = 604800
}
```

A similar policy needs to be applied to the RDNSS Lifetime if ISP_A and ISP_B DNS servers are used.

3.2.2. Two Routers, Primary/Backup Uplinks

Let's look at a more complex scenario where two border routers are terminating two ISP uplinks (one each), acting as redundant first-hop routers for hosts. The topology is shown on Fig.2

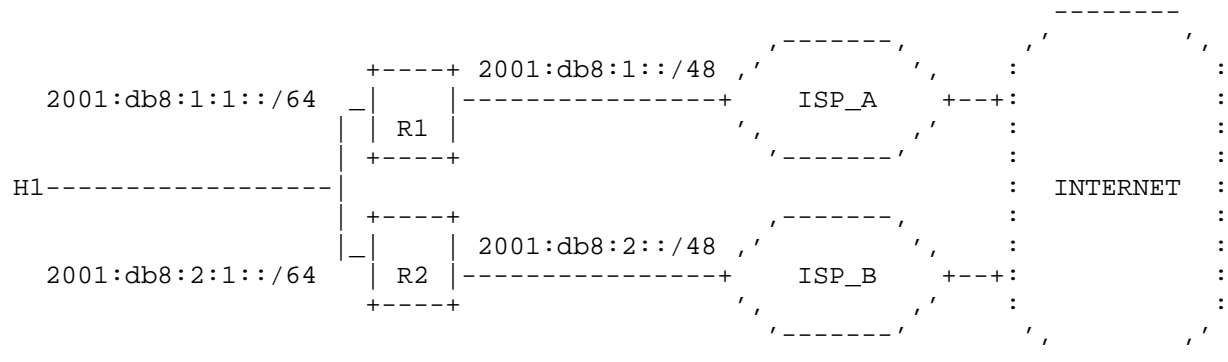


Figure 2: Two Routers, Primary/Backup Uplinks

In this scenario R1 sends RAs with PIO for 2001:db8:1:1::/64 (ISP_A address space) and R2 sends RAs with PIO for 2001:db8:2:1::/64 (ISP_B address space). Each router needs to have a forwarding policy configured for packets received on its hosts-facing interface:

```

if {
    packet_destination_address is not in 2001:db8:1::/48 or 2001:db8:2::/48
    packet_source_address is in 2001:db8:1::/48
} then {
    next-hop is ISP_A_uplink
}
if {
    packet_destination_address is not in 2001:db8:1::/48 or 2001:db8:2::/48
    packet_source_address is in 2001:db8:2::/48
} then {
    next-hop is ISP_B_uplink
}

```

In this case there is more than one way to ensure that hosts are selecting the correct source address based on the uplink status. If VRRP is used to provide first-hop redundancy and the master router is the one with the active uplink, then the simplest way is to use the VRRP mastership as a condition for router advertisement. So, if ISP_A is the primary uplink, the routers R1 and R2 need to be configured in the following way:

R1 is the VRRP master by default (when ISP_A uplink is up). If ISP_A uplink is down, then R1 becomes a backup. Router advertisements on R1's interface facing H1 needs to have the following policy applied:


```
prefix 2001:db8:1:1::/64 {  
  if vrrp_master then preferred_lifetime = 604800  
  else preferred_lifetime = 0  
}
```

R2 is VRRP backup by default. Router advertisement on R2 interface facing H1 needs to have the following policy applied:

```
prefix 2001:db8:2:1::/64 {  
  if vrrp_master then preferred_lifetime = 604800  
  else preferred_lifetime = 0  
}
```

If VRRP is not used or interface status tracking is not used for mastership switchover, then each router needs to be able to detect the uplink failure/recovery on the neighboring router, so that RAs with updated preferred lifetime values are triggered. Depending on the network setup various triggers like a route to the uplink interface subnet or a default route received from the uplink can be used. The obvious drawback of using the routing table to trigger the conditional RAs is that some additional configuration is required. For example, if a route to the prefix assigned to the ISP uplink is used as a trigger, then the conditional RA policy would have the following logic:

R1:

```
prefix 2001:db8:1:1::/64 {  
  if ISP_A_uplink is up then preferred_lifetime = 604800  
  else preferred_lifetime = 0  
}
```

R2:

```
prefix 2001:db8:2:1::/64 {  
  if ISP_A_uplink_route is present then preferred_lifetime = 0  
  else preferred_lifetime = 604800  
}
```

3.2.3. Single Router, Load Balancing Between Uplinks

Let's look at the example topology shown in Figure 1, but with both uplinks used simultaneously. In this case R1 would send RAs containing PIOs for both prefixes, 2001:db8:1:1::/64 and 2001:db8:2:1::/64, changing the preferred lifetime based on particular uplink availability. If the interface status is used as uplink availability indicator, then the policy logic would look like the following:

```
prefix 2001:db8:1:1::/64 {
  if ISP_A_uplink is up then preferred_lifetime = 604800
  else preferred_lifetime = 0
}
prefix 2001:db8:2:1::/64 {
  if ISP_B_uplink is up then preferred_lifetime = 604800
  else preferred_lifetime = 0
}
```

R1 needs a forwarding policy to be applied to forward packets to the correct uplink based on the source address as described in Section 3.2.1.

3.2.4. Two Router, Load Balancing Between Uplinks

In this scenario the example topology is similar to the one shown in Figure 2, but both uplinks can be used at the same time. It means that both R1 and R2 need to have the corresponding forwarding policy to forward packets based on their source addresses.

Each router would send RAs with POI for the corresponding prefix, setting preferred_lifetime to a non-zero value when the ISP uplink is up, and deprecating the prefix by setting the preferred lifetime to 0 in case of uplink failure. The uplink recovery would trigger another RA with non-zero preferred lifetime to make the addresses from the prefix preferred again. The example RA policy on R1 and R2 would look like:

R1:

```
prefix 2001:db8:1:1::/64 {
  if ISP_A_uplink is up then preferred_lifetime = 604800
  else preferred_lifetime = 0
}
```

R2:

```
prefix 2001:db8:2:1::/64 {
  if ISP_B_uplink is up then preferred_lifetime = 604800
  else preferred_lifetime = 0
}
```

3.2.5. Topologies with Dedicated Border Routers

For simplicity reasons all topologies below show the ISP uplinks terminated on the first-hop routers. Obviously, the proposed approach can be used in more complex topologies when dedicated devices are used for terminating ISP uplinks. In that case VRRP

mastership or interface status can not be used as a trigger for conditional RAs and route presence as described above should be used instead.

Let's look at the example topology shown on the Figure 3:

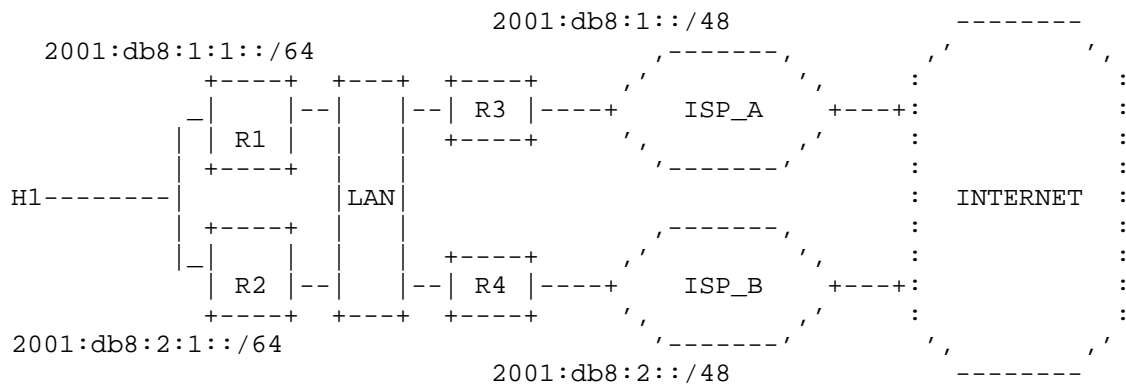


Figure 3: Dedicated Border Routers

For example, if ISP_A is a primary uplink and ISP_B is a backup one then the following policy might be used to achieve the desired behaviour (H1 is using ISP_A address space, 2001:db8:1:1::/64 while ISP_A uplink is up and only using ISP_B 2001:db8:2:1::/64 prefix if the uplink is non-operational):

R1 and R2 policy:

```

prefix 2001:db8:1:1::/64 {
    if ISP_A_uplink_route is present then preferred_lifetime = 604800
    else preferred_lifetime = 0
}
prefix 2001:db8:2:1::/64 {
    if ISP_A_uplink_route is present then preferred_lifetime = 0
    else preferred_lifetime = 604800
}
  
```

For load-balancing case the policy would look slightly different: each prefix has non-zero preferred_lifetime only if the corresponding ISP uplink route is present:

```
prefix 2001:db8:1:1::/64 {  
  if ISP_A_uplink_route is present then preferred_lifetime = 604800  
  else preferred_lifetime = 0  
}  
prefix 2001:db8:2:1::/64 {  
  if ISP_B_uplink_route is present then preferred_lifetime = 0  
  else preferred_lifetime = 604800  
}
```

4. IANA Considerations

This memo asks the IANA for no new parameters.

5. Security Considerations

5.1. Privacy Considerations

6. Acknowledgements

7. References

7.1. Normative References

- [I-D.ietf-rtgwg-enterprise-pa-multihoming]
Baker, F., Bowers, C., and J. Linkova, "Enterprise Multihoming using Provider-Assigned Addresses without Network Prefix Translation: Requirements and Solution", draft-ietf-rtgwg-enterprise-pa-multihoming-00 (work in progress), March 2017.
- [RFC1918] Rekhter, Y., Moskowitz, B., Karrenberg, D., de Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, DOI 10.17487/RFC1918, February 1996, <<http://www.rfc-editor.org/info/rfc1918>>.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<http://www.rfc-editor.org/info/rfc2119>>.
- [RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, DOI 10.17487/RFC2460, December 1998, <<http://www.rfc-editor.org/info/rfc2460>>.
- [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>>.

- [RFC3022] Srisuresh, P. and K. Egevang, "Traditional IP Network Address Translator (Traditional NAT)", RFC 3022, DOI 10.17487/RFC3022, January 2001, <<http://www.rfc-editor.org/info/rfc3022>>.
- [RFC3582] Abley, J., Black, B., and V. Gill, "Goals for IPv6 Site-Multihoming Architectures", RFC 3582, DOI 10.17487/RFC3582, August 2003, <<http://www.rfc-editor.org/info/rfc3582>>.
- [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>>.
- [RFC4193] Hinden, R. and B. Haberman, "Unique Local IPv6 Unicast Addresses", RFC 4193, DOI 10.17487/RFC4193, October 2005, <<http://www.rfc-editor.org/info/rfc4193>>.
- [RFC4218] Nordmark, E. and T. Li, "Threats Relating to IPv6 Multihoming Solutions", RFC 4218, DOI 10.17487/RFC4218, October 2005, <<http://www.rfc-editor.org/info/rfc4218>>.
- [RFC4219] Lear, E., "Things Multihoming in IPv6 (MULTI6) Developers Should Think About", RFC 4219, DOI 10.17487/RFC4219, October 2005, <<http://www.rfc-editor.org/info/rfc4219>>.
- [RFC4862] Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless Address Autoconfiguration", RFC 4862, DOI 10.17487/RFC4862, September 2007, <<http://www.rfc-editor.org/info/rfc4862>>.
- [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>>.
- [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>>.
- [RFC8106] Jeong, J., Park, S., Beloeil, L., and S. Madanapalli, "IPv6 Router Advertisement Options for DNS Configuration", RFC 8106, DOI 10.17487/RFC8106, March 2017, <<http://www.rfc-editor.org/info/rfc8106>>.

7.2. Informative References

- [I-D.ietf-rtgwg-dst-src-routing]
Lamparter, D. and A. Smirnov, "Destination/Source Routing", draft-ietf-rtgwg-dst-src-routing-04 (work in progress), May 2017.
- [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>>.
- [RFC4941] Narten, T., Draves, R., and S. Krishnan, "Privacy Extensions for Stateless Address Autoconfiguration in IPv6", RFC 4941, DOI 10.17487/RFC4941, September 2007, <<http://www.rfc-editor.org/info/rfc4941>>.
- [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>>.
- [RFC5534] Arkko, J. and I. van Beijnum, "Failure Detection and Locator Pair Exploration Protocol for IPv6 Multihoming", RFC 5534, DOI 10.17487/RFC5534, June 2009, <<http://www.rfc-editor.org/info/rfc5534>>.
- [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>>.
- [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>>.
- [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>>.

Appendix A. Change Log

Initial Version: July 2017

Authors' Addresses

Jen Linkova
Google
Mountain View, California 94043
USA

Email: furry@google.com

Massimiliano Stucchi

Email: max@stucchi.ch