Network Working Group
Internet-Draft
Updates: rfc4191, rfc4861 (if approved)
Intended status: Standards Track
Expires: November 13, 2017

F. Templin, Ed.
Boeing Research & Technology
J. Woodyatt
Google

May 12, 2017

Route Information Options in IPv6 Neighbor Discovery
draft-templin-6man-rio-redirect-03.txt

Abstract

The IPv6 Neighbor Discovery (ND) protocol provides a Router Solicitation (RS) function allowing nodes to solicit a Router Advertisement (RA) response from an on-link router, a Neighbor Solicitation (NS) function allowing nodes to solicit a Neighbor Advertisement (NA) response from an on-link neighbor, and a Redirect function allowing routers to inform nodes of a better next hop neighbor on the link toward the destination. This document specifies backward-compatible extensions to IPv6 ND messages to support the discovery of more-specific routes.

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 November 13, 2017.

Copyright Notice

Table of Contents

1. Introduction 2
2. Terminology 3
3. Motivation 3
4. Route Information Options in IPv6 Neighbor Discovery Messages 4
4.1. Classical Redirection Scenario 4
4.2. RIO Redirection Scenario 6
4.2.1. Router Specification 6
4.2.2. Source Specification 6
4.2.3. Target Specification 7
4.2.4. Operation Without Redirects 8
4.2.5. Multiple RIOs 8
4.2.6. Why NS/NA? 8
4.2.7. RIOs in RS Messages 9
5. Implementation Status 9
6. IANA Considerations 9
7. Security Considerations 9
8. Acknowledgements 10
9. References 11
9.1. Normative References 11
9.2. Informative References 11
Appendix A. Link-layer Address Changes 12
Appendix B. Interfaces with Multiple Link-Layer Addresses . . . 12
Authors' Addresses 12

1. Introduction

"Neighbor Discovery for IP version 6 (IPv6)" [RFC4861] (IPv6 ND) provides a Router Solicitation (RS) function allowing nodes to solicit a Router Advertisement (RA) response from an on-link router, a Neighbor Solicitation (NS) function allowing nodes to solicit a Neighbor Advertisement (NA) response from an on-link neighbor, and a Redirect function allowing routers to inform nodes of a better next hop neighbor on the link toward the destination. Further guidance for processing Redirect messages is given in "First-Hop Router Selection by Hosts in a Multi-Prefix Network" [RFC8028].

"Default Router Preferences and More-Specific Routes" [RFC4191] specifies a Route Information Option (RIO) that routers can include in RA messages to inform recipients of more-specific routes. This document specifies a backward-compatible extension to allow nodes to

include RIOs in other IPv6 ND messages to support the dynamic discovery of more-specific routes.

2. Terminology

The terminology in the normative references applies.

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]. Lower case uses of these words are not to be interpreted as carrying RFC2119 significance.

3. Motivation

An example of a good application for RIO is the local-area subnets served by the routers described in "Basic Requirements for IPv6 Customer Edge Routers" [RFC7084]. While many customer edge routers are capable of operating in a mode with a dynamic routing protocol operating in the local-area network, the default mode of operation is typically designed for unmanaged operation without any dynamic routing protocol. On these networks, the only means for any node to learn about routers on the link is by using the Router Discovery protocol described in [RFC4861].

Nevertheless, hosts on unmanaged home subnets may use "IPv6 Prefix Options for DHCPv6" [RFC3633] (DHCPv6 PD) to receive IPv6 routing prefixes for additional subnets allocated from the space provided by the service provider, and operate as routers for other links where hosts in delegated subnets are attached. Hosts may even learn about more specific routes than the default route by processing RIOs in RA messages according to the rules for Type "C" hosts described in [RFC4191].

However, due to perceptions of the security considerations for hosts in processing RIOs on unmanaged networks, the default configuration

for common host IPv6 implementations is not Type "C" behavior. Accordingly, on typical home networks the forwarding path from hosts on one subnet to destinations on every off-link local subnet always passes through the customer edge router, even when a shorter path would otherwise be available through an on-link router. This adds costs for retransmission on shared LAN media, often adding latency and jitter with queuing delay and delay variability. This is not materially different under the scenarios described in "IPv6 Home Networking Architecture Principles" [RFC7368] except that routers may use an interior dynamic routing protocol to coordinate sending of RIOs in RA messages, which as explained above, are not processed by typical hosts.

In increasingly common practice, nodes that receive prefix delegations may connect an entourage of "Internet of Things" back end devices. The node may therefore appear as a router from the perspective of the back end devices but behave as a host on the link from the perspective of receiving Redirects and without participating in a dynamic routing protocol. Instead, the node sends initial packets with a source address taken from one of the node's delegated prefixes via a default or more-specific route with a router on the link as the next-hop. The router may return a Redirect message with an RIO if there is a target node on the link that would be a better next-hop for the destination. The target may itself be a holder of prefix delegations that behaves in a similar fashion as the source node. Examples of where such relationships apply include civil aviation networks, unmanned aerial vehicle networks and enterprise networks that host mobile end user devices (e.g., cell phones, tablets, laptops, etc.).

By using RIOs in IPv6 ND messages, the forwarding path between subnets can be shortened while accepting a much narrower opening of attack surfaces on general purpose hosts related to the Router Discovery protocol. The basic idea is simple: hosts normally send packets for off-link destinations to their default router unless they receive ND Redirect messages designating another on-link node as the target. This document allows ND Redirects additionally to suggest another on-link node as the target for one or more routing prefixes, including one with the destination. Hosts that receive RIOs in ND Redirect messages then unicast NS messages to the target containing those RIOs, and process the unicast NA messages the target sends in reply. If hosts only process RIOs in NA messages when they have

previously unicast them in NS messages to the targets of received ND Redirect messages, then hosts only process RIO at the initiative of routers they already accept as authoritative.

4. Route Information Options in IPv6 Neighbor Discovery Messages

The RIO is specified for inclusion in RA messages in Section 2.3 of [RFC4191], while the neighbor discovery functions are specified in [RFC4861]. This specification permits routers to include RIOs in other IPv6 ND messages so that recipients can discover a better next hop for a destination *prefix* instead of just a specific destination. This specification therefore updates [RFC4191] and [RFC4861], as discussed in the following sections.

4.1. Classical Redirection Scenario

In the classical redirection scenario there are three actors, namely the Source, Router and Target as shown in Figure 1:

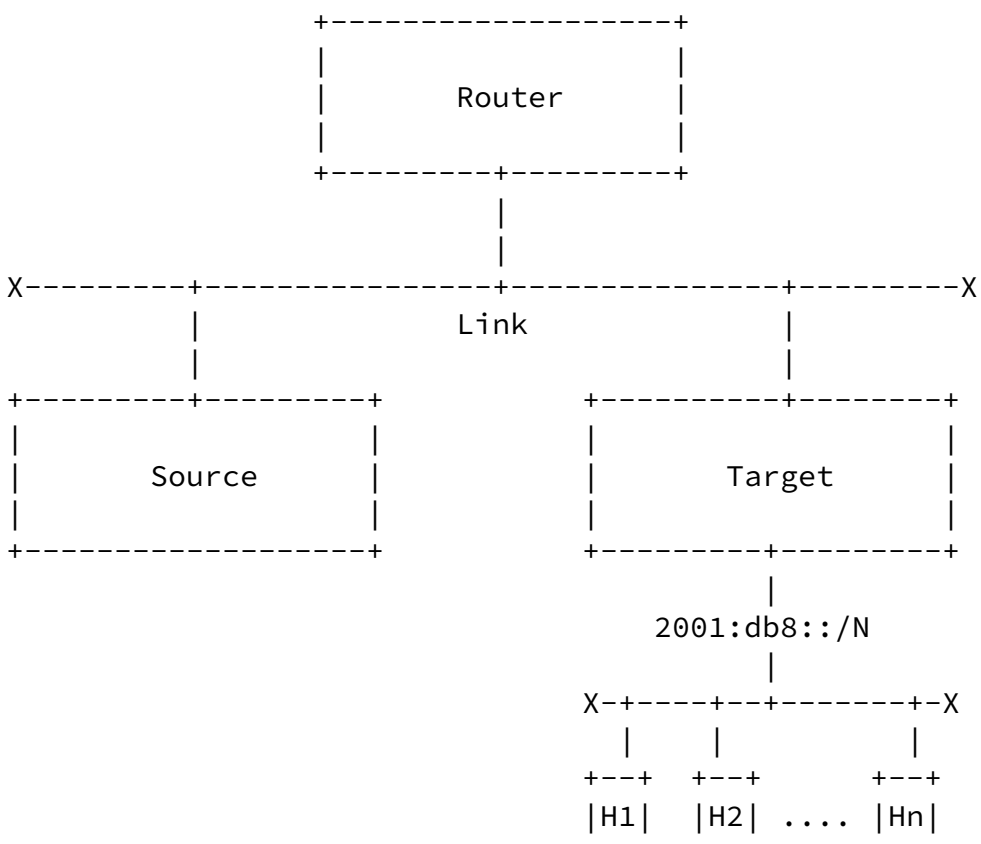

Figure 1: Classical Redirection Scenario

In addition, the Target may be a router that connects an arbitrarily-complex set of IPv6 networks (e.g., as depicted by 2001:db8::/N in the figure) with hosts H(i).

In this scenario, the Source initially has no route for 2001:db8::/N and must send initial packets destined to correspondents H(i) via a first-hop Router. Upon receiving the packets, the Router forwards the packets to the Target and may also send a Redirect message back to the Source with a destination address corresponding to the packet that triggered the Redirect, the target link-local address and the target link-layer address. After receiving the message, the Source may begin sending packets destined to H(i) directly to the Target, which will then forward them to its connected networks.

This specification augments the classical Redirection scenario by allowing the Router to include an entire prefix (e.g., 2001:db8::/N) in an RIO option in the Redirect message, and thereafter allowing the Source to include an RIO in an NS message and the Target to include an RIO in its NA response. The following sections present this "augmented" RIO redirection scenario.

4.2. RIO Redirection Scenario

In the RIO redirection scenario, the Source sends initial packets via the Router the same as in the classical scenario. When the Router receives the packets, it searches its routing tables for a route that is assigned to the Target and that covers the destination address of the packet. The Router then includes this prefix in an RIO in a Redirect message to send back to the Source. When the Source receives the Redirect message, it includes the RIO in an NS message to send to the Target. When the Target receives the NS message, it first verifies that the prefix included in the RIO is indeed one of its own prefixes. If so, the Target fills in the Prefix, Route Lifetime and Prf values in the RIO option, and returns the option in a unicast NA message reply to the Source. The Source can then install the prefix in the RIO option in its routing table. The

following sections present more detailed specifications for the Router, Source and Target.

4.2.1. Router Specification

When the Router receives a packet from the Source that is destined to a host in an IPv6 network aggregated by the Target, the Router searches its routing table for a prefix that covers the destination address(e.g., 2001:db8::/N, as depicted in Figure 1). The Router then prepares a Redirect message with the Destination field set to the packet's IPv6 destination address, with the Target field set to the link-local address of the Target, with a Target Link-Layer Address option (TLLAO) set to the link-layer address of the Target, and with an RIO that includes a Prefix for the destination with Route Lifetime and Prf set to 0. The Router then sends the Redirect message to the Source (subject to rate limiting).

4.2.2. Source Specification

According to [RFC4861], a Source that receives a valid Redirect message updates its destination cache per the Destination Address and its neighbor cache per the Target Address. According to [RFC4191], Sources can be classified as Type "A", "B" or "C" based on how they process RIOs, where a Type "C" Source updates its routing table per any RIO elements included in an RA message. Finally, according to [RFC8028], a Type "C" Source operating on a Multi-Prefix Network with multiple default routes can make source address selection decisions based on information in its routing table decorated with information derived from the source of the RIO element.

In light of these considerations, this document introduces a new Type "D" behavior for Sources with the same behavior as a Type "C" Source, but which also process RIO elements in Redirect and NA messages, and

include RIO elements in NS messages. Type "D" Sources process Redirect messages with RIO elements by first verifying that the Prefix in the first RIO matches the Destination address. If the Destination address does not match the Prefix, the Source discards the Redirect message. Otherwise, the Source updates its neighbor cache per the Target Address and its destination cache per the Destination Address the same as for classical redirection. Next, the Source MAY send an NS message containing an RIO option to the Target

to elicit an NA response.

When the Type 'D' Source receives the solicited NA message from the Target, if the NA includes an RIO with a Prefix matching the one that it received in the Redirect message the Source installs the Prefix in its routing table including the Route Lifetime and Prf values, and with the Target's address as the next hop.

After the Source installs the Prefix in its routing table, it MAY then begin sending packets with destination addresses that match the Prefix directly to the Target Instead of sending them to the Router. The Source SHOULD decrement the Route Lifetime and MAY send new NS messages to receive a fresh Route Lifetime (if the Route Lifetime decrements to 0, the Source instead deletes the route from its routing table). The Source MAY furthermore delete the route at any time and again allow packets to flow through the Router which may send a fresh Redirect. The Source should then again test the route by performing a unicast NS/NA exchange with the Target the same as described above.

After updating its routing table, the Source MAY receive an unsolicited NA message from the Target with an RIO with new Route Lifetime and/or Prf values. If the RIO Prefix is in its routing table, and if the Route Lifetime value is 0, the Source deletes the corresponding route.

After updating its routing table, the Source MAY also receive a Destination Unreachable message from the Target with Code 0 ("no route to destination"). If so, the Source again deletes the corresponding route from its routing table.

4.2.3. Target Specification

When the Target receives an NS message from the Source containing an RIO, it examines the Prefix to see if it matches one of the prefixes in its prefix list. If so, the Target copies the RIO into a unicast NA message to send back to the Source and fills in the Route Lifetime and Prf fields with values that are consistent with its prefix list. The Target then sends the NA message back to the Source.

At some later time, the Target may either alter or deprecate the

corresponding prefix in its prefix list. If the Target has sent solicited NA messages with RIO options to one or more Sources, the Target SHOULD send unsolicited NA messages with RIOs that include the Prefix and with Route Lifetime set to 0. If the Target receives packets with destination addresses that do not match a prefix in its prefix list, the target sends a Destination Unreachable message to the Source with Code 0 ("no route to destination"), subject to rate limiting.

4.2.4. Operation Without Redirects

If the Source has some way to determine the Target's link-local address without receiving a Redirect message from the Router, the Source MAY send an NS message with an RIO directly to the Target with the Prefix field set to the destination address of an IPv6 packet and with Prefix Length set to 128.

When the Target receives the NS message, it prepares an NA response with an RIO that includes a Prefix and Prefix length for one of its prefixes that covers the destination address. The Target then sends the NA message to the Source.

Before accepting the NA message, the Source must have assurance that the Target is authoritative for its claimed Prefix and Prefix length (e.g., through an authoritative intermediate node that examines the message and drops it if the claims are invalid).

4.2.5. Multiple RIOs

If a Redirect includes multiple RIOs, the Source only checks the destination address for a match against the Prefix in the first RIO.

If an NA message includes multiple RIOs, the Source only accepts those Prefixes for which it has some way of knowing that the Target is the correct next hop (e.g., via a Redirect).

If an NS message includes multiple RIOs, the Target only responds to those Prefixes with matching entries in its prefix list.

4.2.6. Why NS/NA?

Since [RFC4191] already specifies the inclusion of RIOs in RA messages, a natural question is why this document advocates the use of NS/NA instead of RS/RA?

First, NS/NA exchanges used by the IPv6 Neighbor Unreachability Detection (NUD) procedure are unicast-only whereas RA responses to RS

messages are typically sent as multicast. Since this mechanism must operate only between the Source and Target without disturbing any other nodes on the link, the use of unicast-only exchanges is required.

Second, the IPv6 ND specification places restrictions on minimum delays between RA messages. Since this mechanism expects an immediate advertisement from the Target in response to the Source's solicitation, only the NS/NA exchange can satisfy this property.

Third, the RA message is the "swiss army knife" of the IPv6 ND protocol. RA messages carry numerous configuration parameters for nodes on the link, including Cur Hop Limit, M/O flags, Router Lifetime, Reachable Time, Retrans Time, Prefix Information Options, MTU options, etc. The Target must not advertise any of this information to the soliciting Source.

Finally, operators are deeply concerned about the security of RA messages - so much so that they deploy link security mechanisms that drop RA messages originating from nodes claiming to be an authoritative router for the link.

4.2.7. RIOs in RS Messages

This document permits a source host to include RIOs in RS messages in order to solicit RIOs in the corresponding RA messages from a trusted router. The RS/RA RIO exchange is conducted in the same fashion as for NS/NA exchanges, with the exception that RA messages may be returned as multicast and/or may also include other configuration information for the link.

5. Implementation Status

The IPv6 ND functions and RIOs are widely deployed in IPv6 implementations.

6. IANA Considerations

This document introduces no IANA considerations.

7. Security Considerations

The Redirect message validation rules in Section 8.1 of [RFC4861] require recipients to verify that the IP source address of the Redirect is the same as the current first-hop router for the specified ICMP Destination Address. Recipients therefore naturally

reject any Redirect message with an incorrect source address.

Other security considerations for IPv6 ND messages that include RIOs are the same as specified in Section 11 of [RFC4861]. Namely, the protocol must take measures to secure IPv6 ND messages on links where spoofing attacks are possible.

A spoofed ND message containing no RIOs could cause corruption in the recipient's destination cache, while a spoofed ND message containing RIOs could corrupt the host's routing tables. While the latter would seem to be a more onerous result, the possibility for corruption is unacceptable in either case.

"IPv6 ND Trust Models and Threats" [RFC3756] discusses spoofing attacks, and states that: "This attack is not a concern if access to the link is restricted to trusted nodes". "SEcure Neighbor Discovery (SEND)" [RFC3971] provides one possible mitigation for other cases.

"IPv6 Router Advertisement Guard" [RFC6105] ("RA Guard") describes a layer-2 filtering technique intended for network operators to use in protecting hosts from receiving RA messages sent by nodes that are not among the set of routers regarded as legitimate by the network operator.

A soliciting node must have some form of trust basis for knowing that the advertising node is authoritative for the prefixes it includes in RIOs. For example, when an NS/NA exchange is triggered by the receipt of a Redirect, the soliciting node can verify that the RIOs in the NA message match the ones it received in the Redirect message.

8. Acknowledgements

Joe Touch suggested a standalone draft to document this approach in discussions on the intarea list. The work was subsequently transferred to the 6man list, where the following individuals provided valuable feedback: Mikael Abrahamsson, Zied Bouziri, Brian Carpenter, Steinar Haug, Christian Huitema, Tatuya Jinmei, Tomoyuki Sahara.

Discussion with colleagues during the "bits-and-bites" session at IETF98 helped shape this document. Those colleagues are gratefully

acknowledged for their contributions.

This work was sponsored through several ongoing initiatives, including 1) the NASA Safe Autonomous Systems Operation (SASO) program under NASA contract number NNA16BD84C, 2) the FAA SE2025 contract number DTFAWA-15-D-00030, 3) the Boeing Information Technology (BIT) MobileNet program, and 4) the Boeing Research & Technology (BR&T) enterprise autonomy program.

9. References

9.1. Normative References

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, DOI 10.17487/RFC0791, September 1981, <http://www.rfc-editor.org/info/rfc791>.

[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>.

[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>.

[RFC4443] Conta, A., Deering, S., and M. Gupta, Ed., "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", RFC 4443, DOI 10.17487/RFC4443, March 2006, <http://www.rfc-editor.org/info/rfc4443>.

[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>.

9.2. Informative References

[RFC3633] Troan, O. and R. Droms, "IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6", RFC 3633, DOI 10.17487/RFC3633, December 2003, <http://www.rfc-editor.org/info/rfc3633>.

[RFC3756] Nikander, P., Ed., Kempf, J., and E. Nordmark, "IPv6 Neighbor Discovery (ND) Trust Models and Threats", RFC 3756, DOI 10.17487/RFC3756, May 2004, <http://www.rfc-editor.org/info/rfc3756>.

[RFC3971] Arkko, J., Ed., Kempf, J., Zill, B., and P. Nikander, "SEcure Neighbor Discovery (SEND)", RFC 3971, DOI 10.17487/RFC3971, March 2005, <http://www.rfc-editor.org/info/rfc3971>.

[RFC6105] Levy-Abegnoli, E., Van de Velde, G., Popoviciu, C., and J. Mohacsi, "IPv6 Router Advertisement Guard", RFC 6105, DOI 10.17487/RFC6105, February 2011, <http://www.rfc-editor.org/info/rfc6105>.

[RFC7084] Singh, H., Beebee, W., Donley, C., and B. Stark, "Basic Requirements for IPv6 Customer Edge Routers", RFC 7084, DOI 10.17487/RFC7084, November 2013, <http://www.rfc-editor.org/info/rfc7084>.

[RFC7368] Chown, T., Ed., Arkko, J., Brandt, A., Troan, O., and J. Weil, "IPv6 Home Networking Architecture Principles", RFC 7368, DOI 10.17487/RFC7368, October 2014, <http://www.rfc-editor.org/info/rfc7368>.

[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. Link-layer Address Changes

Type "D" hosts send unsolicited NAs to announce link-layer address changes per standard neighbor discovery [RFC4861]. Link-layer address changes may be due to localized factors such as hot-swap of

an interface card, but could also occur during movement to a new point of attachment on the same link.

Appendix B. Interfaces with Multiple Link-Layer Addresses

Type "D" host interfaces may have multiple connections to the link; each with its own link-layer address. Type "D" nodes can therefore include multiple link-layer address options in IPv6 ND messages. Neighbors that receive these messages can cache and select link-layer addresses in a manner outside the scope of this specification.

Authors' Addresses

Fred L. Templin (editor)
Boeing Research & Technology
P.O. Box 3707
Seattle, WA 98124
USA

Email: fltemplin@acm.org

James Woodyatt
Google
3400 Hillview Ave
Palo Alto, CA 94304
USA

Email: jhw@google.com