Network Working Group   F. Templin, Ed.
Internet-Draft   Boeing Research & Technology
Intended status: Informational   June 14, 2018
Expires: December 16, 2018

Multi-Addressing Considerations for IPv6 Prefix Delegation
draft-templin-v6ops-pdhost-21.txt

Abstract

IPv6 prefixes are typically delegated to requesting routers which assign them to their downstream-attached links and networks. The requesting node can provision the prefix according to whether it acts as a router on behalf of any downstream networks and/or as a host on behalf of its local applications. In the latter case, the requesting node can use portions of the delegated prefix for its own multi-addressing purposes. This document therefore considers prefix delegation consideations for both the classic routing and various multi-addressing use cases.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

IPv6 Prefix Delegation (PD) entails 1) the communication of a prefix from a server to a requesting router, 2) a representation of the prefix in the network's Routing Information Base (RIB) and the first-hop router's forwarding information base (FIB), and 3) a control messaging service to maintain prefix lifetimes. Following delegation, the prefix is available for the requesting router's exclusive use and is not shared with any other nodes. This document considers multi-addressing considerations for requesting nodes that acts as a router on behalf of any downstream networks and/or as a host on behalf of its local applications.

For nodes that connect downstream-attached networks (e.g., a cellphone that connects a "tethered" Internet of Things (IoT), a laptop computer with a complex internal network of virtual machines, etc.), the classic routing model applies as shown in Figure 1:

```
                               .---------.
                            ,-(           )-.
                          (   +----------+   )
                         (    |Server 'S'|    )
                          (   +----------+   )
                           (   Network 'N'  )
                            `-(__________)-'
                                   |
                        +----------+----------+
                        | first-hop router 'F'|
                        +----------+----------+
                                   |
                     upstream link |
                                   |
                        +----------+----------+
                        |  upstream interface |
                        +---------------------+
                        |                     |
                        |requesting router 'R'|
                        |    (Prefix 'P')     |
                        |                     |
                        +--+-+--+-+--+-----+--+
                        |A1| |A2| |A3| ... |Aj|
                        +--+-+--+-+--+-----+--+
                        |downstream interfaces|
                        +----------+----------+
                                   |
          internal and/or external |
              downstream links     |
       X----+-------------+--------+----+---------------+---X
            |             |             |               |
       +---++-+--+   +---++-+--+   +---++-+--+     +---++-+--+
       |   |Ak|  |   |   |Al|  |   |   |Am|  |     |   |A*|  |
       |   +--+  |   |   +--+  |   |   +--+  |     |   +--+  |
       | host H1 |   | host H2 |   | host H3 | ... | host Hn |
       +---------+   +---------+   +---------+     +---------+

          <-------------- Downstream Network ------------->
```

Figure 1: Classic Routing Model

In the classic routing model, requesting router 'R' has one or more upstream interfaces and connects zero or more internal and/or external downstream networks. When 'R' requests a prefix delegation, the following sequence of events transpires:

o Server 'S' located in network 'N' delegates prefix 'P' via first-hop router 'F' to requesting router 'R'.

- o 'P' is injected into the RIB for 'N', and 'F' configures a FIB entry with 'R' as the next hop.

- o R' receives 'P' and assigns zero or more addresses 'A(*)' taken from 'P' to its downstream interfaces

- o 'R' advertises zero or more sub-prefixes taken from 'P' in RA messages to hosts 'H(i)' on downstream networks.

- o 'R' delegates zero or more sub-prefixes taken from 'P' to requesting routers in downstream networks.

- o 'R' acts as a router for hosts 'H(i)' on downstream networks and as a host on behalf of its local applications.

This document also considers the case when 'R' uses portions of 'P' for its own internal multi-addressing purposes. [RFC7934] provides Best Current Practice (BCP) motivations for the benefits of multi-addressing, while an operational means for providing nodes with multiple addresses is given in [RFC8273]. The following multi-addressing alternatives for delegated prefixes compliment this framework while providing greater efficiency since no duplicate address queries over the upstream link are needed (see:Section 3).

In a first alternative, when requesting node 'R' receives prefix 'P', it can assign addresses taken from 'P' to downstream virtual interfaces (e.g., a loopback) as shown in Figure 2:

```
                            x
                            |
             upstream link  |
                            |
                +-----------+----------+
                |  upstream Interface  |
                +----------------------+
                |                      |
                | requesting node 'R'  |
                |                      |
                +--+-+--+-+--+-----+--+
                |A1| |A2| |A3| ... |An|
                +--+-+--+-+--+-----+--+
                |  virtual interfaces  |
                +----------------------+
```

Figure 2: Address Assignment to Downstream Virtual Interfaces

In a second alternative, 'R' could assign Pv6 addresses taken from 'P' to the upstream interface over which the prefix was received as shown in Figure 3:

```
                                 x
                                 |
                   upstream link |
                                 |
                      +----------+----------+
                      |  upstream interface |
                      +--+-+--+-+--+-----+--+
                      |A1| |A2| |A3| ... |An|
                      +--+-+--+-+--+-----+--+
                      |                     |
                      | requesting node 'R' |
                      |                     |
                      +---------------------+
```

Figure 3: Upstream Interface Address Assignment

In a third alternative, 'R' could assign IPv6 addresses taken from 'P' to its local applications which appear as "psuedo" virtual interfaces as shown in Figure 4:

```
                                 x
                                 |
                   upstream link |
                                 |
                      +----------+----------+
                      |  upstream Interface |
                      +---------------------+
                      |                     |
                      | requesting node 'R' |
                      |                     |
                      +--+-+--+-+--+-----+--+
                      |A1| |A2| |A3| ... |An|
                      +--+-+--+-+--+-----+--+
                      |     Applications    |
                      +---------------------+
```

Figure 4: Application Addresssing Model

With these IPv6 PD-based multi-addressing considerations, the node can configure an unlimited supply of addresses to make them available for local applications without requiring coordination with other nodes on upstream interfaces. The following sections present considerations for nodes that employ IPv6 PD mechanisms.

2. Terminology

The terminology of the normative references apply, and the terms "node", "host" and "router" are the same as defined in [RFC8200].

The following terms are defined for the purposes of this document:

shared prefix
: an IPv6 prefix that may be advertised to more than one node on the link, e.g., in a Router Advertisement (RA) message Prefix Information Option (PIO) [RFC4861]. The router that advertises the prefix must consider the prefix as on-link so that the IPv6 Neighbor Discovery (ND) address resolution function will identify the correct neighbor for each packet.

individual prefix
: an IPv6 prefix that is advertised to exactly one node on the link, where the node may be unaware that the prefix is individual and may not participate in prefix maintenance procedures. The router that advertises the prefix can consider the prefix as on-link or not on-link. In the former case, the router performs address resolution and only forwards those packets that match one of the node's configured addresses so that the node will not receive unwanted packets. In the latter case, the router can simply forward all packets matching the prefix to the node which must then drop any packets that do not match one of its configured addresses. An example individual prefix service is documented in [RFC8273].

delegated prefix
: an IPv6 prefix that is explicitly conveyed to a node for its own exclusive use, where the node is an active participant in prefix delegation and maintenance procedures. The first-hop router simply forwards all packets matching the prefix to the requesting node. The requesting node associates the prefix with downstream and/or internal virtual interfaces (i.e., and not the upstream interface).

3. Multi-Addressing Considerations

IPv6 allows nodes to assign multiple addresses to a single interface. [RFC7934] discusses options for multi-addressing as well as use cases where multi-addressing may be desirable. Address configuration options for multi-addressing include StateLess Address AutoConfiguration (SLAAC) [RFC4862], Dynamic Host Configuration Protocol for IPv6 (DHCPv6) address configuration [RFC3315], manual configuration, etc.

Nodes configure addresses from a shared or individual prefix and assign them to the upstream interface over which the prefix was received. When the node assigns the addresses, it is required to use Multicast Listener Discovery (MLD) [RFC3810] to join the appropriate solicited-node multicast group(s) and to use the Duplicate Address Detection (DAD) algorithm [RFC4862] to ensure that no other node configures a duplicate address.

In contrast, a node that configures addresses from a delegated prefix can assign them without invoking MLD/DAD on an upstream interface, since the prefix has been delegated to the node for its own exclusive use and is not shared with any other nodes.

4. Multi-Addressing Alternatives for Delegated Prefixes

When a node receives a delegated prefix, it has many alternatives for provisioning the prefix to its local interfaces and/or downstream networks. [RFC7278] discusses alternatives for provisioning a prefix obtained by a User Equipment (UE) device under the 3rd Generation Partnership Program (3GPP) service model. This document considers the more general case when the node receives a delegated prefix explicitly provided for its own exclusive use.

When the node receives the prefix, it can distribute the prefix to internal (virtual) or external (physical) downstream networks and configure zero or more addresses for itself on downstream interfaces. The node then acts as a router on behalf of its downstream networks.

The node could instead (or in addition) use portions of the delegated prefix for its own multi-addressing purposes. In a first alternative, the node can assign as many addresses as it wants from the prefix to downstream virtual interfaces.

In a second alternative, the node can assign as many addresses as it wants from the prefix to the upstream interface over which the prefix was received.

In a third alternative, the node can assign addresses taken from the delegated prefix to its local applications. The applications themselves then serve as virtual interfaces, i.e., instead of using a traditional virtual interface such as a loopback. (Note that, in the future, the practice of assigning unique non-link-local IPv6 addresses to applications could obviate the need for transport protocol port numbers.)

In these multi-addressing cases, the node assigns the prefix itself to a virtual interface so that unused portions of the prefix are correctly identified as unreachable. The node then acts as a host on

behalf of its local applications even though neighbors on the upstream link consider it as a router.

5. Address Autoconfiguration Considerations

Nodes autoconfigure addresses according to Section 6 of IPv6 Node Requirements [I-D.ietf-6man-rfc6434-bis].

Nodes configure at least one non-link-local adddress, i.e., for network management and error reporting purposes.

Nodes recognize the Subnet Router Anycast address [RFC4291] for each delegated prefix. Therefore, the node's use of the Subnet Router Anycast address must be indistinguishable from the behavior of an ordinary router when viewed from the outside world.

6. MLD/DAD Implications

When a node configures addresses for itself from a shared or individual prefix, it performs MLD/DAD by sending multicast messages over the upstream interface to test whether there is another node on the link that configures a duplicate address. When there are many such addresses and/or many such nodes, this could result in substantial multicast traffic that affects all nodes on the link.

When a node configures addresses for itself from a delegated prefix, it can configure as many addresses as it wants but need not perform MLD/DAD for any of the addresses over the upstream interface. This means that the node can configure arbitrarily many addresses without causing any multicast messaging over the upstream interface that could disturb other nodes.

7. Dynamic Routing Protocol Implications

Nodes that receive delegated prefixes can be configured to either participate or not participate in a dynamic routing protocol over the upstream interface. When there are many nodes on the upstream link, dynamic routing protocol participation might be impractical due to scaling limitations, and may also be exacerbated by factors such as node mobility.

Unless it participates in a dynamic routing protocol, the node initially has only a default route pointing to a neighbor via an upstream interface. This means that packets sent by the node over an upstream interface will initially go through a default router even if there is a better first-hop node on the link.

8. IPv6 Neighbor Discovery Implications

When a node receives a shared or individual prefix with "L=1" and has a packet to send to an IPv6 destination within the prefix, it is required to use the IPv6 ND address resolution function over the upstream interface to resolve the link-layer address of a neighbor that configures the address. When a node receives a shared or individual prefix with "L=0" and has a packet to send to an IPv6 destination within the prefix, if the address is not one of the node's own addresses it sends the packet to a default router since "L=0" makes no statement about on-link or off-link properties of the prefix [RFC4861].

When a node receives a delegated prefix, it acts as a simple host to send Router Solicitation (RS) messages over the upstream interface (i.e., the same as described in Section 4.2 of [RFC7084]) but also sets the "Router" flag to TRUE in its Neighbor Advertisement messages. The node considers the upstream interface as a non-advertising interface [RFC4861], i.e., it does not send RA messages over the upstream interface. The node further does not perform the IPv6 ND address resolution function over the upstream interface, since the delegated prefix is by definition not to be associated with the interface.

In all cases, the current first-hop router may send a Redirect message that updates the node's neighbor cache so that future packets can use a better first-hop node on the link. The Redirect can apply either to a singleton destination address, or to an entire destination prefix as described in [I-D.templin-6man-rio-redirect].

9. ICMPv6 Implications

The Internet Control Message Protocol for IPv6 (ICMPv6) includes a set of control message types [RFC4443] including Destination Unreachable (DU).

According to [RFC4443], routers should return DU messages (subject to rate limiting) with code 0 ("No route to destination") when a packet arrives for which there is no matching entry in the routing table, and with code 3 ("Address unreachable") when the IPv6 destination address cannot be resolved.

According to [RFC4443], hosts should return DU messages (subject to rate limiting) with code 3 to internal applications when the IPv6 destination address cannot be resolved, and with code 4 ("Port unreachable") if the IPv6 destination address is one of its own addresses but the transport protocol has no listener.

Nodes that obtain and manage delegated prefixes per this document observe the same procedures as described for both routers and hosts above.

10. Prefix Delegation Services

Selection of prefix delegation services must be considered according to specific use cases. An example service is that offered by DHCPv6 [RFC3633]. An alternative service based on IPv6 ND messaging has also been proposed [I-D.pioxfolks-6man-pio-exclusive-bit].

Other, non-router, mechanisms may exist, such as proprietary IPAMs, [I-D.ietf-anima-prefix-management] and [I-D.sun-casm-address-pool-management-yang].

11. IANA Considerations

This document introduces no IANA considerations.

12. Security Considerations

Security considerations for IPv6 Neighbor Discovery [RFC4861] and any applicable PD mechanisms apply to this document. Nodes that receive delegated prefixes need not perform MLD/DAD procedures on their upstream interfaces, meaning that they can avoid introducing multicast messaging congestion on the upstream link. Also, routers that delegate prefixes keep only a single neighbor cache entry for each prefix delegation recipient, meaning that the router's neighbor cache cannot be subject to resource exhaustion attacks.

For shared and individual prefixes, if the router that advertises the prefix considers the prefix as on-link the IPv6 ND address resolution function will prevent unwanted IPv6 packets from reaching the node. For delegated prefixes and individual prefixes that are not considered on-link, the router delivers all packets that match the prefix to the unicast link-layer address of the node (i.e., as determined by resolution of the node's link-local address) even if they do not match one of the node's configured addresses. In that case, the node may receive unwanted IPv6 packets via an upstream interface that do not match either a configured IPv6 address or a transport listener. The node then drops the packets and observes the "Destination Unreachable - Address/Port unreachable" procedures discussed in Section 9.

The node may also receive IPv6 packets via an upstream interface that do not match any of the node's delegated prefixes. In that case, the node drops the packets and observes the "Destination Unreachable - No route to destination" procedures discussed in Section 9. Dropping

the packets is necessary to avoid a reflection attack that would cause the node to forward packets received from an upstream interface via the same or a different upstream interface.

In all cases, the node must decide whether or not to send DUs according to the specific operational scenario. In trusted networks, the node should send DU messages to provide useful information to potential correspondents. In untrusted networks, the node can refrain from sending DU messages to avoid providing sensitive information to potential attackers.

13. Acknowledgements

This work was motivated by discussions on the v6ops list. Mark Smith, Ricardo Pelaez-Negro, Edwin Cordeiro, Fred Baker, Ron Bonica, Naveen Lakshman, Ole Troan, Bob Hinden, Brian Carpenter, Joel Halpern, Albert Manfredi, Dusan Mudric, Paul Marks, Joe Touch and Alex Petrescu provided useful comments that have greatly improved the document.

This work is aligned with the NASA Safe Autonomous Systems Operation (SASO) program under NASA contract number NNA16BD84C.

This work is aligned with the FAA as per the SE2025 contract number DTFAWA-15-D-00030.

This work is aligned with the Boeing Information Technology (BIT) MobileNet program and the Boeing Research & Technology (BR&T) enterprise autonomy program.

14. References

14.1. Normative References

[RFC3315] Droms, R., Ed., Bound, J., Volz, B., Lemon, T., Perkins, C., and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, DOI 10.17487/RFC3315, July 2003, <https://www.rfc-editor.org/info/rfc3315>.

[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, <https://www.rfc-editor.org/info/rfc3633>.

[RFC3810] Vida, R., Ed. and L. Costa, Ed., "Multicast Listener Discovery Version 2 (MLDv2) for IPv6", RFC 3810, DOI 10.17487/RFC3810, June 2004, <https://www.rfc-editor.org/info/rfc3810>.

[RFC4443] Conta, A., Deering, S., and M. Gupta, Ed., "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", STD 89, RFC 4443, DOI 10.17487/RFC4443, March 2006, <https://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, <https://www.rfc-editor.org/info/rfc4861>.

[RFC4862] Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless Address Autoconfiguration", RFC 4862, DOI 10.17487/RFC4862, September 2007, <https://www.rfc-editor.org/info/rfc4862>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

14.2. Informative References

[I-D.ietf-6man-rfc6434-bis]
Chown, T., Loughney, J., and T. Winters, "IPv6 Node Requirements", draft-ietf-6man-rfc6434-bis-08 (work in progress), March 2018.

[I-D.ietf-anima-prefix-management]
Jiang, S., Du, Z., Carpenter, B., and Q. Sun, "Autonomic IPv6 Edge Prefix Management in Large-scale Networks", draft-ietf-anima-prefix-management-07 (work in progress), December 2017.

[I-D.pioxfolks-6man-pio-exclusive-bit]
Kline, E. and M. Abrahamsson, "IPv6 Router Advertisement Prefix Information Option eXclusive Flag", draft-pioxfolks-6man-pio-exclusive-bit-02 (work in progress), March 2017.

[I-D.sun-casm-address-pool-management-yang]
Sun, Q., Xie, C., Boucadair, M., Peng, T., and Y. Lee, "A YANG Data Model for Address Pool Management", draft-sun-casm-address-pool-management-yang-00 (work in progress), March 2017.

[I-D.templin-6man-rio-redirect]
Templin, F. and j. woodyatt, "Route Information Options in IPv6 Neighbor Discovery", draft-templin-6man-rio-redirect-06 (work in progress), May 2018.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, DOI 10.17487/RFC4291, February 2006, <https://www.rfc-editor.org/info/rfc4291>.

[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, <https://www.rfc-editor.org/info/rfc7084>.

[RFC7278] Byrne, C., Drown, D., and A. Vizdal, "Extending an IPv6 /64 Prefix from a Third Generation Partnership Project (3GPP) Mobile Interface to a LAN Link", RFC 7278, DOI 10.17487/RFC7278, June 2014, <https://www.rfc-editor.org/info/rfc7278>.

[RFC7934] Colitti, L., Cerf, V., Cheshire, S., and D. Schinazi, "Host Address Availability Recommendations", BCP 204, RFC 7934, DOI 10.17487/RFC7934, July 2016, <https://www.rfc-editor.org/info/rfc7934>.

[RFC8273] Brzozowski, J. and G. Van de Velde, "Unique IPv6 Prefix per Host", RFC 8273, DOI 10.17487/RFC8273, December 2017, <https://www.rfc-editor.org/info/rfc8273>.

Appendix A. Change Log

<< RFC Editor - remove prior to publication >>

Changes from -20to -21:

- o Re-worked classic routing model section
- o Included multi-addressing case where addresses may be assigned to applications
- o Removed strong/weak end system discussions

Changes from -19 to -20:

- o figure 1 updates to show Server as being somewhere in the network
- o Introductory material to show relation to other RFCs on multi-addressing

Changes from -18 to -19:

- o added new section on Prefix Delegation Services

Changes from -17 to -18:

- o re-worked discussion on the prefix delegation service in Section 1

- o updated figures in Section 1

Changes from -16 to -17:

- o added supporting text in the introduction to discuss the Delegating Router's relationship with the Requesting Router and with supporting intrastructure in the operator's network

- o updated figures in introduction to include representation of operator's network

- o added new section on Address Autoconfiguration Considerations

Author's Address

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

Email: fltemplin@acm.org