DHC Work Group
Internet-Draft
Intended status: Standards Track
Expires: June 3, 2021

I. Farrer
Deutsche Telekom AG
Naveen. Kottapalli
Benu Networks
M. Hunek
Technical University of Liberec
R. Patterson
Sky UK Ltd

November 30, 2020

DHCPv6 Prefix Delegating Relay Requirements
draft-ietf-dhc-dhcpv6-pd-relay-requirements-04

Abstract

This memo describes operational problems that are known to occur when using DHCPv6 relays with Prefix Delegation. These problems can prevent successful delegation and result in routing failures. To address these problems, this memo provides necessary functional requirements for operating DHCPv6 relays with Prefix Delegation.

It is recommended that any network operator that is using DHCPv6 prefix delegation with relays should ensure that these requirements are followed on their networks.

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 https://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 3, 2021.

Copyright Notice

Copyright (c) 2020 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 (https://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

For Internet service providers that offer native IPv6 access with prefix delegation to their customers, a common deployment architecture is to have a DHCPv6 relay agent function located in the ISP's Layer-3 customer edge device and separate, centralized DHCPv6 server infrastructure. [RFC8415] describes the functionality of a DHCPv6 relay and Section 19.1.3 mentions this deployment scenario, but does not provide detail for all of the functional requirements

that the relay needs to fulfill to operate deterministically in this deployment scenario.

A DHCPv6 relay agent for prefix delegation is a function commonly implemented in routing devices, but implementations vary in their functionality and client/server inter-working. This can result in operational problems such as messages not being forwarded by the relay or un-reachability of the delegated prefixes. This document provides a set of requirements for devices implementing a relay function for use with prefix delegation.

The mechanisms for a relay to inject routes (including aggregated ones), on its network-facing interface based on prefixes learned from a server via DHCP-PD are out of scope of the document.

Multi-hop DHCPv6 relaying is not affected. The requirements in this document are solely applicable to the DHCP relay agent co-located with the first-hop router that the DHCPv6 client requesting the prefix is connected to, so no changes to any subsequent relays in the path are needed.

2. Terminology

2.1. General

This document uses the terminology defined in [RFC8415], however when defining the functional elements for prefix delegation [RFC8415], Section 4.2 defines the term 'delegating router' as:

> "The router that acts as a DHCP server and responds to requests for delegated prefixes."

This document is concerned with deployment scenarios in which the DHCPv6 relay and DHCPv6 server functions are separated, so the term 'delegating router' is not used. Instead, a new term is introduced to describe the relaying function:

Delegating relay A delegating relay acts as an intermediate device, forwarding DHCPv6 messages containing IA_PD/IAPREFIX options between the client and server. The delegating relay does not implement a DHCPv6 server function. The delegating relay is also responsible for routing traffic for the delegated prefixes.

Where the term 'relay' is used on its own within this document, it should be understood to be a delegating relay, unless specifically stated otherwise.

In CableLabs DOCSIS environments, the Cable Modem Termination System (CMTS) would be considered a delegating relay with respect to Customer Premises Devices (CPEs) [DOCSIS_3.1], Section 5.2.7.2. A Broadband Network Gateway (BNG) in a DSL based access network may be a delegating relay if it does not implement a local DHCPv6 server function [TR-092], Section 4.10.

[RFC8415] defines the 'DHCP server', (or 'server') as:

> "A node that responds to requests from clients. It may or may not be on the same link as the client(s). Depending on its capabilities, if it supports prefix delegation it may also feature the functionality of a delegating router."

This document serves the deployment cases where a DHCPv6 server is not located on the same link as the client (necessitating the delegating relay). The server supports prefix delegation and is capable of leasing prefixes to clients, but is not responsible for other functions required of a delegating router, such as managing routes for the delegated prefixes.

The term 'requesting router' has previously been used to describe the DHCP client requesting prefixes for use. This document adopts the [RFC8415] terminology and uses 'DHCP client' or 'client' interchangeably for this element.

2.2. Topology

The following diagram shows the deployment topology relevant to this document.

```
+
|             ------- uplink ------>
|                                        _    ,--,_
|   +--------+      +------------+   _(   `'      )_    +--------+
+---+   PD   |-------| Delegating |--(    Operator    )---| DHCPv6 |
|   | Client |      |    relay   |   `(_ Network_)'    | server |
|   +--------+      +----------- +      `--'`---'      +--------+
|
|             <----- downlink ------
+                 (client facing)
Client
Network
```

Figure 1: Topology overview

The client requests prefixes via the downlink interface of the delegating relay. The resulting prefixes will be used for addressing

the client network. The delegating relay is responsible for forwarding DHCP messages, including prefix delegation requests and responses between the client and server. Messages are forwarded from the delegating relay to the server using multicast or unicast via the operator uplink interface.

The delegating relay provides the operator's Layer-3 edge towards the client and is responsible for routing traffic to and from clients connected to the client network using addresses from the delegated prefixes.

2.3. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Problems Observed with Existing Delegating Relay Implementations

The following sections of the document describe problems that have been observed with delegating relay implementations in commercially available devices.

3.1. DHCP Messages not being Forwarded by the Delegating Relay

Delegating relay implementations have been observed not to forward messages between the client and server. This generally occurs if a client sends a message which is unexpected by the delegating relay. For example, the delegating relay already has an active PD lease entry for an existing client on a port. A new client is connected to this port and sends a Solicit message. The delegating relay then drops the Solicit messages until it receives either a DHCP Release message from the original client, or the existing lease times out. This causes a particular problem when a client device needs to be replaced due to a failure.

In addition to dropping messages, in some cases the delegating relay will generate error messages and send them to the client, e.g. 'NoBinding' messages being sent in the event that the delegating relay does not have an active delegated prefix lease.

3.2. Delegating Relay Loss of State on Reboot

For proper routing of client traffic, the delegating relay requires a corresponding routing table entry for each active prefix delegated to a connected client. A delegating relay which does not store this

state persistently across reboots will not be able to forward traffic to client's delegated leases until the state is re-established through new DHCP messages.

3.3. Multiple Delegated Prefixes for a Single Client

[RFC8415] allows for a client to include more than one instance of OPTION_IA_PD in messages in order to request multiple prefix delegations by the server. If configured for this, the server supplies one (or more) instance of OPTION_IAPREFIX for each received instance of OPTION_IA_PD, each containing information for a different delegated prefix.

In some delegating relay implementations, only a single delegated prefix per-DUID is supported. In those cases only one IPv6 route for one of the delegated prefixes is installed; meaning that other prefixes delegated to a client are unreachable.

3.4. Dropping Messages from Devices with Duplicate MAC addresses and DUIDs

It is an operational reality that client devices with duplicate MAC addresses and/or DUIDs exist and have been deployed. In this situation, the operational costs of locating and swapping out such devices are prohibitive.

Delegating relays have been observed to restrict forwarding client messages originating from one client DUID to a single interface. In this case if the same client DUID appears from a second client on another interface while there is already an active lease, messages originating from the second client are dropped causing the second client to be unable to obtain a prefix delegation.

It should be noted that in some access networks, the MAC address and/or DUID are used as part of device identification and authentication. In such networks, enforcing MAC address/DUID uniqueness is a necessary function and not considered a problem.

3.5. Forwarding Loops between Client and Relay

If the client loses information about a prefix that it is delegated while the lease entry and associated route is still active in the delegating relay, then the relay will forward traffic to the client which the client will return to the relay (which is the client's default gateway (learned via an RA)). The loop will continue until either the client is successfully re-provisioned via DHCP, or the lease ages out in the relay.

4. Requirements for Delegating Relays

To resolve the problems described in Section 3 and pre-empt other undesirable behavior, the following section of the document describes a set of functional requirements for the delegating relay.

In addition, relay implementers are reminded that [RFC8415] makes it clear that relays MUST forward packets that either contain message codes (Section 19 of [RFC8415]) it may not understand, or contain options that it does not understand (Section 16 of [RFC8415]).

4.1. General Requirements

G-1: The delegating relay MUST forward messages bidirectionally between the client and server without changing the contents of the message.

G-2: The relay MUST allow for multiple prefixes to be delegated for the same client IA_PD. These delegations may have different lifetimes.

G-3: The relay MUST allow for multiple prefixes (with or without separate IA_PDs) to be delegated to a single client connected to a single interface, identified by its DHCPv6 Client Identifier (DUID).

G-4: A delegating relay may have one or more interfaces on which it acts as a relay, as well as one or more interfaces on which it does not (for example, in an ISP, it might act as a relay on all southbound interfaces, but not on the northbound interfaces). The relay SHOULD allow the same client identifier (DUID) to have active delegated prefix leases on more than one interface simultaneously, unless client DUID uniqueness is necessary for the functioning or security of the network. This is to allow client devices with duplicate DUIDs to function on separate broadcast domains.

G-5: The maximum number of simultaneous prefixes delegated to a single client MUST be configurable.

G-6: The relay MUST implement a mechanism to limit the maximum number of active prefix delegations on a single port for all client identifiers and IA_PDs. This value MUST be configurable.

G-7: It is RECOMMENDED that delegating relays support at least 8 active delegated leases per client device and use this as the default limit.

G-8: The delegating relay MUST update the lease lifetimes based on the Client Reply messages it forwards to the client and only expire the delegated prefixes when the valid lifetime has elapsed.

G-9: On receipt of a Release message from the client, the delegating relay MUST expire the active leases for each of the IA_PDs in the message.

4.2. Routing Requirements

R-1: The relay MUST maintain a local routing table that is dynamically updated with leases and the associated next-hops as they are delegated to clients. When a delegated prefix is Released or expires, the associated route MUST be removed from the relay's routing table.

R-2: The delegating relay's routing entry MUST use the same prefix length for the delegated prefix as given in the IA_PD.

R-3: The relay MUST provide a mechanism to dynamically update ingress filters permitting ingress traffic sourced from client delegated leases and blocking packets from invalid source prefixes. This is to implement anti-spoofing as described in [BCP38]. The delegating relay's ingress filter entry MUST use the same prefix length for the delegated prefix as given in the IA_PD.

R-4: The relay MAY provide a mechanism to dynamically advertise delegated leases into a routing protocol as they are learned. When a delegated lease is released or expires, the delegated route MUST be withdrawn from the routing protocol. The mechanism by which the routes are inserted and deleted is out of the scope of this document.

R-5: To prevent routing loops, the relay SHOULD implement a configurable policy to drop potential looping packets received on any DHCP-PD client facing interfaces.

The policy SHOULD be configurable on a per-client or per-destination basis.

Looping packets are those with a destination address in a prefix delegated to a client connected to that interface, as follows:

* For point-to-point links, when the packet's ingress and egress interfaces match.

* For multi-access links, when the packet's ingress and egress interface match, and the source link-layer and next-hop link-layer addresses match.

An ICMPv6 Type 1, Code 6 (Destination Unreachable, reject route to destination) error message MAY be sent as per [RFC4443], section 3.1. The ICMP policy SHOULD be configurable.

4.3. Service Continuity Requirements

S-1: To preserve active client prefix delegations across relay restarts, the relay SHOULD implement at least one of the following:

* Implement DHCPv6 bulk lease query as defined in [RFC5460].

* Store active prefix delegations in persistent storage so they can be re-read after the reboot.

S-2: If a client's next-hop link-local address becomes unreachable (e.g., due to a link-down event on the relevant physical interface), routes for the client's delegated prefixes MUST be retained by the delegating relay unless they are released or removed due to expiring DHCP timers. This is to re-establish routing for the delegated prefix if the client next-hop becomes reachable without the delegated prefixes needing to be re-learned.

S-3: The relay SHOULD implement DHCPv6 active lease query as defined in [RFC7653] to keep the local lease database in sync with the DHCPv6 server.

4.4. Operational Requirements

O-1: The relay SHOULD implement an interface allowing the operator to view the active delegated prefixes. This SHOULD provide information about the delegated lease and client details such as client identifier, next-hop address, connected interface, and remaining lifetimes.

O-2: The relay SHOULD provide a method for the operator to clear active bindings for an individual lease, client or all bindings on a port.

O-3: To facilitate troubleshooting of operational problems between the delegating relay and other elements, it is RECOMMENDED

that a time synchronization protocol is used by the delegating relays and DHCP servers.

5. Acknowledgements

The authors of this document would like to thank Bernie Volz, Ted Lemon, and Michael Richardson for their valuable comments.

6. IANA Considerations

This memo includes no request to IANA.

7. Security Considerations

This document does not add any new security considerations beyond those mentioned in Section 22 of [RFC8213].

If the delegating relay implements [BCP38] filtering, then the filtering rules will need to be dynamically updated as delegated prefixes are leased.

[RFC8213] describes a method for securing traffic between the relay agent and server by sending DHCP messages over an IPsec tunnel. In this case the IPsec tunnel is functionally the server-facing interface and DHCPv6 message snooping can be carried out as described. It is RECOMMENDED that this is implemented by the delegating relay.

8. References

8.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC5460] Stapp, M., "DHCPv6 Bulk Leasequery", RFC 5460, DOI 10.17487/RFC5460, February 2009, <https://www.rfc-editor.org/info/rfc5460>.

[RFC7653] Raghuvanshi, D., Kinnear, K., and D. Kukrety, "DHCPv6 Active Leasequery", RFC 7653, DOI 10.17487/RFC7653, October 2015, <https://www.rfc-editor.org/info/rfc7653>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8415] Mrugalski, T., Siodelski, M., Volz, B., Yourtchenko, A., Richardson, M., Jiang, S., Lemon, T., and T. Winters, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 8415, DOI 10.17487/RFC8415, November 2018, <https://www.rfc-editor.org/info/rfc8415>.

8.2. Informative References

[BCP38] IETF, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing https://tools.ietf.org/html/bcp38", RFC 2827, BCP 38.

[DOCSIS_3.1]
CableLabs, "MAC and Upper Layer Protocols Interface Specification", DOCSIS 3.1, January, 2017", <https://apps.cablelabs.com/specification/CM-SP-MULPIv3.>.

[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>.

[RFC8213] Volz, B. and Y. Pal, "Security of Messages Exchanged between Servers and Relay Agents", RFC 8213, DOI 10.17487/RFC8213, August 2017, <https://www.rfc-editor.org/info/rfc8213>.

[TR-092] Broadband Forum, "Broadband Remote Access Server (BRAS) Requirements Document, August, 2004", <https://www.broadband-forum.org/download/TR-092.pdf>.

Authors' Addresses

Ian Farrer
Deutsche Telekom AG
Landgrabenweg 151
Bonn, NRW 53227
DE

Email: ian.farrer@telekom.de

Naveen Kottapalli
Benu Networks
300 Concord Road
Billerica, MA 01821
US

Email: naveen.sarma@gmail.com

Martin Hunek
Technical University of Liberec
Studentska 1402/2
Liberec, L 46017
CZ

Email: martin.hunek@tul.cz

Richard Patterson
Sky UK Ltd
1 Brick Lane
London E1 6PU
UK

Email: richard.patterson@sky.uk