Internet Area Working Group
Internet-Draft
Intended status: Standards Track
Expires: 8 January 2026

D. Lamparter
NetDEF, Inc.
T. Fiebig
MPI-INF

7 July 2025

DHCPv4 Option for IPv4 routes with IPv6 nexthops
draft-equinox-intarea-dhcpv4-route4via6-02

Abstract

As a result of the shortage of IPv4 addresses, installations are increasingly recovering IPv4 addresses from uses where they are not strictly necessary. One such situation is in establishing next hops for IPv4 routes, replacing this use with IPv6 addresses. This document describes how to provision DHCP-configured hosts with their routes in such a situation.

// This draft lives at https://github.com/eqvinox/dhc-route4via6

Table of Contents

1. Introduction

IPv4 is currently (and will likely be for some time) in a situation where IPv4 addresses are in short supply, but services still need to be made available to users that do not yet have IPv6 connectivity. In some cases, even the service side may not have IPv6 support yet. In other cases some aspect of the service precludes using proxy-style service delivery with translation technologies on either or both sides. This leads to a need for fine-grained deployment of IPv4 connectivity with minimum wastage of addresses.

A particularly interesting improvement enabled by the extension described here is the complete removal of IPv4 addresses from first-hop routers acting as DHCPv4/v6 relays, while still providing IPv4 connectivity. In this scenario, the relay (assumed colocated with the router) has no IPv4 address to use to communicate with the client. An almost-working solution for this case is presented by [DHCPv6] with the [DHCP4o6] transport method. Since this mechanism

encapsulates IPv4 DHCP messages, all related IPv4 configuration can be carried. However, DHCPv4 does not support a way to encode an IPv6 default gateway or other routes, which is necessary in this case.

If the router and relay are not co-located, the relay may have an IPv4 address while the router does not. In this case, the option described in this document could be carried in a plain IPv4 DHCP message.

Note that the changes described in this document are to DHCPv4, not DHCPv6.

2. 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. Extended static route function

This document defines a list-style DHCPv4 option. Each item describes a pair of IPv4 destination prefix and an associated nexthop. As per usual DHCPv4 processing ([DHCP-LONGOPT]), multiple instances of the option are concatenated before splitting the result up into individual pairs.

1. each pair is processed as one unit, building up a list of destination prefixes and next-hops. It is expected that the list will in most cases present only one destination prefix (the default route) with one or two next-hops. However, DHCPv4 clients processing the option MUST support processing multiple pairs of distinct destination prefix and nexthop.

2. if one and the same destination prefix (equal address and prefix length) occurs multiple times, they MUST be merged into a single route with multiple nexthops (equal-cost multipath routing). The client MAY impose a limit to the number of nexthops, and if that number is exceeded discard an arbitrary choice of excess nexthops. The client MUST NOT reject the route in its entirety if the number of nexthops is exceeded. (Note the limit is permitted to be 1, i.e. no support for multiple nexthops.)

3. the client MUST silently ignore repeats of the same pair of destination prefix and nexthop. The server SHOULD NOT send such an option value.

4. if the nexthop is a link-local address, it MUST be accepted and associated with the link the DHCP packet was received on. This also applies when using the DHCP packet's source address when it is a link-local address. In most cases, "link" will mean "interface", however in some cases the client may have further information on when it is communicating on the same link (e.g. 802.11 SSID, or configured link aggregation) and apply that instead.

3.1. Applicable next-hop behavior

Outlined in [IANA-IPv6], not all IPv6 addresses are valid for use when encoded as next-hop and some have specific functionality [IANA-IPv6-SPECIAL] attached to them as follows:

1. the unspecified-address nexthop indicates that the destination prefix in the pair should use the DHCP packet's source address as nexthop. When [DHCP4o6] is in use, hosts MUST retrieve the IPv6 source address of the DHCPv6 packet carrying the DHCPV4-RESPONSE message.
   // TODO: does it really make sense to support IPv4 here? Maybe only
   // allow this with DHCP4o6?

2. the Discard-Only Address block (0100::/64) [DISCARD] MAY be used to express unreachable destinations, in particular if only limited but not global IPv4 connectivity is available. If this is used, it MUST be the only next-hop paired with the destination prefix in question. Clients SHOULD ignore the destination prefix entirely if this condition does not hold. If a client is unable to mark destinations as unreachable in its routing table, it MAY ignore the destination prefix and SHOULD indicate a client configuration issue in its administrative interfaces.

3. any unicast IPv6 address MAY be used as next-hop. This specifically also covers link-local addresses, which the client MUST support and MUST associate with the link that it has received the DHCP packets on.

4. // TODO: is ::ffff:192.0.2.123 an IPv4 nexthop? Is this worth
   // supporting explicitly, and then saying that the other static route
   // / default gateway options should be ignored?

5. the following types/ranges of addresses are invalid and MUST NOT be used; no client behavior is specified if any are present in a container:

* the loopback address (::1)
  // TODO: express other directly-connected IPv4 hosts with this?

* any multicast address (ff00::/8)

* any address with a reserved allocation

3.2. Applicable destination prefix behavior

Some IPv4 prefixes, due to their function given in [IANA-IPv4], do not make sense to use with this option. DHCPv4 servers MUST NOT encode and DHCPv4 clients MUST ignore the following prefixes as well as any more-specific prefixes within them:

* 0.0.0.0/8 (note that 0.0.0.0/0 is less specific than this, and thus valid)

* 127.0.0.0/8

* 224.0.0.0/4

* 255.255.255.255/32

Behavior for 240.0.0.0/4 is outside the scope of this document.

4. Expected host behavior

The option described in this document is intended to be implemented on hosts supporting IPv4 routes with IPv6 nexthops as described in [v4overv6]. Hosts that do not support the behavior described there MUST NOT request and MUST ignore the option described in this document.

Hosts that support [v4overv6] behavior and acquire their configuration from [DHCP] SHOULD implement the option described here.

4.1. Singular address assignment

While not limited to this case, this option is expected and intended to be used with assigning a singular IPv4 address to a DHCPv4 client. This implies that the Subnet Mask option defined in [DHCP-OPT] will have the value 255.255.255.255.

DHCPv4 clients implementing the option described in this document MUST process such a Subnet Mask option value as assigning a single address. There is no network or broadcast address for this "single-sized" pseudo-subnet. No IPv4 addresses are expressed to be on-link

for the purposes of [ARP] (though they MAY become so due to additional, e.g. local configuration assigning additional addresses to the interface.)

Whether the address is bound to the interface or host (strong vs. weak host model), and whether to perform or skip [DADv4] for the address is beyond the scope of this document.

4.2. Overlapping routes from other sources

[RFC3442] documents a mechanism to communicate a set of routes and their nexthops over DHCP. The original DHCP "router" option (code 3) may communicate a default router. If either of these options is used, the routes communicated may overlap.

To get consistent and unsurprising behavior, this document places the follwing expectations on the host:
// TODO: redundant paragraph/merge with text above, needs some // merging/editing.

* Routes that describe distinct destination prefixes MUST be handled independently. This includes routes that differ only in prefix length. As a result, the routing table MAY contain a mix of IPv4 routes with IPv4 nexthops as well as IPv6 nexthops. Standard longest prefix match behavior MUST be observed.

* If routes with the same destination prefix are described both with previously existing methods as well as the options documented here, the route described by the latter MUST be used and the routes with IPv4 nexthops MUST be discarded. This notably includes "unreachable" routes described here; a route with an IPv4 nexthop for such a destination MUST still be discarded.

* Multiple routes for the same destination prefix with different nexthops of the same address family SHOULD be combined into a single route for equal-cost multipath behavior, if the host supports this. If ECMP routes are not supported, the host MUST deterministically choose one of the routes. This MAY be done by using the first or last option as seen in DHCP packet order, or by choosing the numerically lowest or highest nexthop.

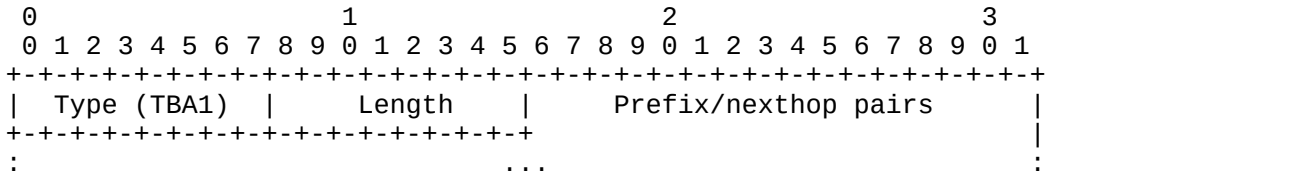
4.3. Default route

The default route is expressed here as a route for 0.0.0.0/0, which is also implied by the absence of any destination prefix suboption. There is no distinct special encoding for a default gateway, any nexthop for 0.0.0.0/0 MUST be treated as if it were a default gateway.

4.4. Routes clashing with the connected subnet

```
// (only applicable if NOT assigning a single IPv4 address as /32)
// TODO: determine what behavior is reasonable here.  (The client is
// likely to be given a /32 subnet mask anyway.)
```

5. DHCP Option encoding

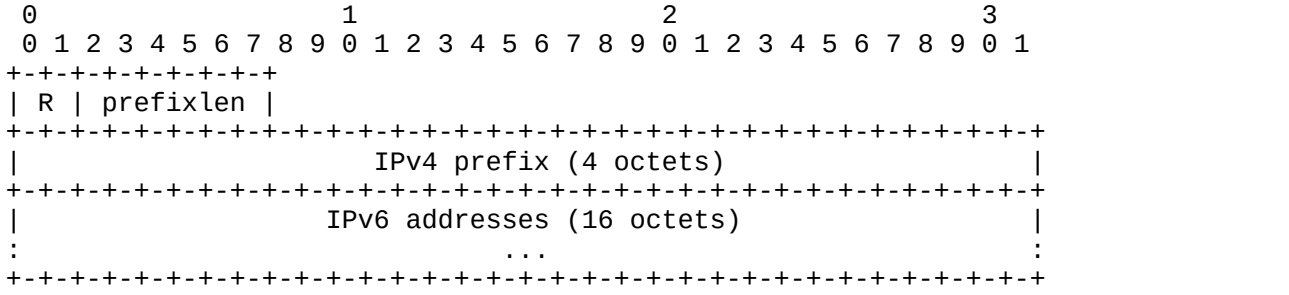


Type TBA1 (field defined in [DHCP-OPT])

Length as defined in [DHCP-OPT]

Prefix/nexthop pairs zero or more items as defined below. Note multiple options are concatenated before processing.

5.1. Destination prefix and nexthop pair



R Reserved bits. MUST be sent as zero, MUST be ignored on receipt.

prefixlen IPv4 prefix length, integer value from 0 to 32 (inclusive)

IPv4 prefix The route’s destination prefix, with unused bits zeroed.

Validity and behavior for specific values is described in Section 3.2.

IPv6 addresses The IPv6 addresses specifying the nexthop for this route. Refer to Section 3.1 for valid values and associated behavior.

6. Security Considerations

TBD

7. Privacy Considerations

TBD

8. IANA Considerations

A codepoint from the "BOOTP Vendor Extensions and DHCP Options" registry is requested for use with the container option described in Section 5.
// Editor note: 2 places of TBA1

A registry is requested to be created for the sub-options in the option above.
// TBD: proper wording for this, and fill in values 1 & 2

9. References

9.1. 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>.

[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>.

[DHCP] Droms, R., "Dynamic Host Configuration Protocol", RFC 2131, DOI 10.17487/RFC2131, March 1997, <https://www.rfc-editor.org/rfc/rfc2131>.

[DHCP-OPT] Alexander, S. and R. Droms, "DHCP Options and BOOTP Vendor Extensions", RFC 2132, DOI 10.17487/RFC2132, March 1997, <https://www.rfc-editor.org/rfc/rfc2132>.

[DHCP-LONGOPT] Lemon, T. and S. Cheshire, "Encoding Long Options in the Dynamic Host Configuration Protocol (DHCPv4)", RFC 3396, DOI 10.17487/RFC3396, November 2002, <https://www.rfc-editor.org/rfc/rfc3396>.

[v4overv6] Chroboczek, J., Kumari, W. A., and T. H淡iland-J淡rgensen, "IPv4 routes with an IPv6 next hop", Work in Progress, Internet-Draft, draft-chroboczek-intarea-v4-via-v6-03, 20 January 2025, <https://datatracker.ietf.org/doc/html/draft-chroboczek-intarea-v4-via-v6-03>.

[DISCARD] Hilliard, N. and D. Freedman, "A Discard Prefix for IPv6", RFC 6666, DOI 10.17487/RFC6666, August 2012, <https://www.rfc-editor.org/rfc/rfc6666>.

[IANA-IPv4] IANA, "IPv4 Address Space Registry", <https://www.iana.org/assignments/ipv4-address-space/>.

[IANA-IPv6] IANA, "Internet Protocol Version 6 Address Space", <https://www.iana.org/assignments/ipv6-address-space/>.

[IANA-IPv6-SPECIAL] IANA, "IPv6 Special-Purpose Address Registry", <https://www.iana.org/assignments/iana-ipv6-special-registry/>.

9.2. Informative References

[RFC3442] Lemon, T., Cheshire, S., and B. Volz, "The Classless Static Route Option for Dynamic Host Configuration Protocol (DHCP) version 4", RFC 3442, DOI 10.17487/RFC3442, December 2002, <https://www.rfc-editor.org/info/rfc3442>.

[DHCPv6] 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>.

[DHCP4o6] Sun, Q., Cui, Y., Siodelski, M., Krishnan, S., and I. Farrer, "DHCPv4-over-DHCPv6 (DHCP 4o6) Transport", RFC 7341, DOI 10.17487/RFC7341, August 2014, <https://www.rfc-editor.org/info/rfc7341>.

[ARP] Plummer, D., "An Ethernet Address Resolution Protocol: Or Converting Network Protocol Addresses to 48.bit Ethernet Address for Transmission on Ethernet Hardware", STD 37, RFC 826, DOI 10.17487/RFC0826, November 1982, <https://www.rfc-editor.org/rfc/rfc826>.

[DADv4] Cheshire, S., "IPv4 Address Conflict Detection", RFC 5227, DOI 10.17487/RFC5227, July 2008, <https://www.rfc-editor.org/rfc/rfc5227>.

Acknowledgements

The authors would like to acknowledge and thank Tomek Mrugalski for very extensive comments, and in particular pointing out the proper way to use DHCP options.

Comments and feedback has been received and appreciated from Ole Troan.

Example encoded options

TBD: outdated examples removed, will be re-added

Revision history (TO BE REMOVED)

* -02: just use a straight up list of fixed size items (prefix+NH) rather than an excessively complicated suboption.

* -01: scrap single-option encoding, use container instead, and reference special-purpose IPv6 addresses (e.g. for discard)

* -00:

Authors' Addresses

David 'equinox' Lamparter
NetDEF, Inc.
San Jose,
United States of America
Email: equinox@diac24.net, equinox@opensourcerouting.org

Tobias Fiebig
Max-Planck-Institut fuer Informatik
Campus E14
66123 Saarbruecken
Germany
Phone: +49 681 9325 3527
Email: tfiebig@mpi-inf.mpg.de