Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 22, 2014

B. Sarikaya
Huawei USA

June 20, 2014

DHCPv6 Route Options for Source Address Dependent Routing
draft-sarikaya-dhc-dhcpv6-raoptions-sadr-00

Abstract

This document describes DHCPv6 Route Options for provisioning IPv6 routes on DHCPv6 client nodes for source address dependent routing. Using these options, an operator can configure multi-homed nodes where other means of route configuration may be impractical.

Requirements Language

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 RFC 2119 [RFC2119].

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 December 22, 2014.

Copyright Notice

Table of Contents

1. Introduction

The Neighbor Discovery (ND) protocol [RFC4861] provides a mechanism for hosts to discover one or more default routers on a directly connected network segment. Extensions to the Router Advertisement (RA) protocol defined in [RFC4191] allow hosts to discover the preferences for multiple default routers on a given link, as well as any specific routes advertised by these routers. This provides network administrators with a new set of tools to handle multi-homed host topologies and influence the route selection by the host. This ND based mechanism however is sub optimal or impractical in some multi-homing scenarios, e.g. source address dependent routing. Both Router Advertisement options [I-D.sarikaya-6man-next-hop-ra] and DHCPv6 can be used. In networks that deployed DHCPv6, the use of DHCPv6 [RFC3315] is seen to be more viable.

DHCPv6 Route Options defined in this document can be used to configure fixed and mobile nodes in multi-homed scenarios with route information and next hop address. Different scenarios exist such as the node is simultaneously connected to multiple access network of e.g. WiFi and 3G. The node may also be connected to more than one gateway. Such connectivity may be realized by means of dedicated physical or logical links that may also be shared with other users nodes such as in residential access networks.

A document defining topologies and in general providing an overview of the issue of source address dependent routing is TBD.

The solution presented in this document is part of the network configuration information. A consistent set of network configuration is defined as Provisioning Domain (PvD) [I-D.ietf-mif-mpvd-arch]. PvDs or so-called explicit PvDs may include information related to more than one interfaces as is the case in this document. It is important to note that the node has a trust relationship with the PvD, in such a case, it is called trusted PvD. The trust is established using authorization and authentication between the node that is using the PvD configuration and the source that provided that configuration. In this document, we assume that DHCP server can provide trusted PvDs to the hosts.

2. DHCPv6 Based Solution

A DHCPv6 based solution allows an operator an on demand and node specific means of configuring static routing information. Such a solution also fits into network environments where the operator prefers to manage Residential Gateway (RG) configuration information from a centralized DHCP server. [RFC7157] provides additional background to the need for a DHCPv6 solution to the problem.

In terms of the high level operation of the solution defined in this draft, a DHCPv6 client interested in obtaining routing information requests the route options using the DHCPv6 Option Request Option (ORO) sent to a server. A Server, when configured to do so, provides the requested route information as part of a nested options structure covering; the next-hop address; the destination prefix; the route metric; any additional options applicable to the destination or next-hop.

2.1. Default route configuration

A non-trustworthy network may be available at the same time as a trustworthy network, with the risk of bad consequences if the host gets confused between the two. These are basically the two models for hosts with multiple interfaces, both of which are valid, but

which are incompatible with each other. In the first model, an interface is connected to something like a corporate network, over a Virtual Private Network (VPN). This connection is trusted because it has been authenticated. Routes obtained over such a connection can probably be trusted, and indeed it may be important to use those routes. This is because in the VPN case, you may also be connected to a network that's offered you a default route, and you could be attacked over that connection if you attempt to connect to resources on the enterprise network over it.

On the other, non-trustworthy network scenario, none of the networks to which the host is connected are meaningfully more or less trustworthy. In this scenario, the untrustworthy network may hand out routes to other hosts, e.g. those in the VPN going through some malicious nodes. This will have bad consequences because the host's traffic intended for the corporate VPN may be hijacked by the intermediate nodes.

DHCPv6 options described in this document can be used to install the routes. However, the use of such a technique makes sense only in the former case above, i.e. trusted network. So the host MUST have an authenticated connection to the network it connects so that DHCPv6 route options can be trusted before establishing routes.

Server MUST NOT define more than one default route.

2.2. Configuring on-link routes

Server may also configure on-link routes, i.e. routes that are available directly over the link, not via routers. To specify on-link routes, server MAY include RTPREFIX option directly in Advertise and Reply messages.

2.3. Deleting obsolete route

There are two mechanisms that allow removing a route. Each defined route has a route lifetime. If specific route is not refreshed and its timer reaches 0, client MUST remove corresponding entry from routing table.

In cases, where faster route removal is needed, server SHOULD return RT_PREFIX option with route lifetime set to 0. Client that receives RT_PREFIX with route lifetime set to 0 MUST remove specified route immediately, even if its previous lifetime did not expire yet.

2.4. Applicability to routers

Contrary to Router Adverisement mechanism, defined in [RFC4861] that explicitly limits configuration to hosts, routing configuration over DHCPv6 defined in this document may be used by both hosts and routers. (This limitation of RA mechanism was partially lifted by W-1 requirement formulated in [RFC6204].)

One of the envisaged usages for this solution are residential gateways (RG) or Customer Premises Equipment (CPE). Those devices very often perform routing. It may be useful to configure routing on such devices over DHCPv6. One example of such use may be a class of premium users that are allowed to use dedicated router that is not available to regular users.

2.5. Updating Routing Information

Network configuration occassionally changes, due to failure of existing hardware, migration to newer equipment or many other reasons. Therefore there a way to inform clients that routing information have changed is required.

There are several ways to inform clients about new routing information. Every client SHOULD periodically refresh its configuration, according to Information Refresh Time Option, so server may send updated information the next time client refreshes its information. New routes may be configured at that time. As every route has associated lifetime, client is required to remove its routes when this timer expires. This method is particularly useful, when migrating to new router is undergoing, but old router is still available.

Server MAY also announce routes via soon to be removed router with lifetimes set to 0. This will cause the client to remove its routes, despite the fact that previously received lifetime may not yet expire.

Aforementioned methods are useful, when there is no urgent need to update routing information. Bound by timer set by value of Information Refresh Time Option, clients may use outdated routing information until next scheduled renewal. Depending on configured value this delay may be not acceptable in some cases. In such scenarios, administrators are advised to use RECONFIGURE mechanism, defined in [RFC3315]. Server transmits RECONFIRGURE message to each client, thus forcing it to immediately start renewal process.

See also Section 2.6 about limitations regarding dynamic routing.

2.6. Limitations

Defined mechanism is not intended to be used as a dynamic routing protocol. It should be noted that proposed mechanism cannot automatically detect routing changes. In networks that use dynamic routing and also employ this mechanism, clients may attempt using routes configured over DHCPv6 even though routers or specific routes ceased to be available. This may cause black hole routing problem. Therefore it is not recommended to use this mechanism in networks that use dynamic routing protocols. This mechanism SHOULD NOT be used in such networks, unless network operator can provide a way to update DHCP server information in case of router availability changes.

Discussion: It should be noted that DHCPv6 server is not able to monitor health of existing routers. As there are currently more than 60 options defined for DHCPv6, it is infeasible to implement mechanism that would monitor huge set of services and stop announcing its availability in case of service outage. Therefore in case of prolonged unavailability human interverntion is required to change DHCPv6 server configuration. If that is considered a problem, network administrators should consider using other alternatives, like RA and ND mechanisms (see [RFC4861]).

3. DHCPv6 Route Options

A DHCPv6 client interested in obtaining routing information includes the NEXT_HOP and RT_PREFIX options as part of its Option Request Option (ORO) in messages directed to a server (as allowed by [RFC3315], i.e. Solicit, Request, Renew, Rebind or Information-request messages). A Server, when configured to do so, provides the requested route information using zero, one or more NEXT_HOP options in messages sent in response (Advertise, and Reply). So as to allow the route options to be both extensible, as well as conveying detailed info for routes, use is made of a nested options structure. Server sends one or more NEXT_HOP options that specify the IPv6 next hop addresses. Each NEXT_HOP option conveys in turn zero, one or more RT_PREFIX options that represents the IPv6 destination prefixes reachable via the given next hop. Server includes RT_PREFIX directly in message to indicate that given prefix is available directly on-link. Server MAY send a single NEXT_HOP without any RT_PREFIX suboptions or with RT_PREFIX that contains ::/0 to indicate available default route. The Formats of the NEXT_HOP and RT_PREFIX options are defined in the following sub-sections.

The DHCPv6 Route Options format borrows from the principles of the Route Information Option defined in [RFC4191].

3.1. Route Prefix Option Format

The Route Prefix Option is used to convey information about a single prefix that represents the destination network. The Route Prefix Option is used as a sub-option in the previously defined Next Hop Option. It may also be sent directly in message to indicate that route is available directly on-link.

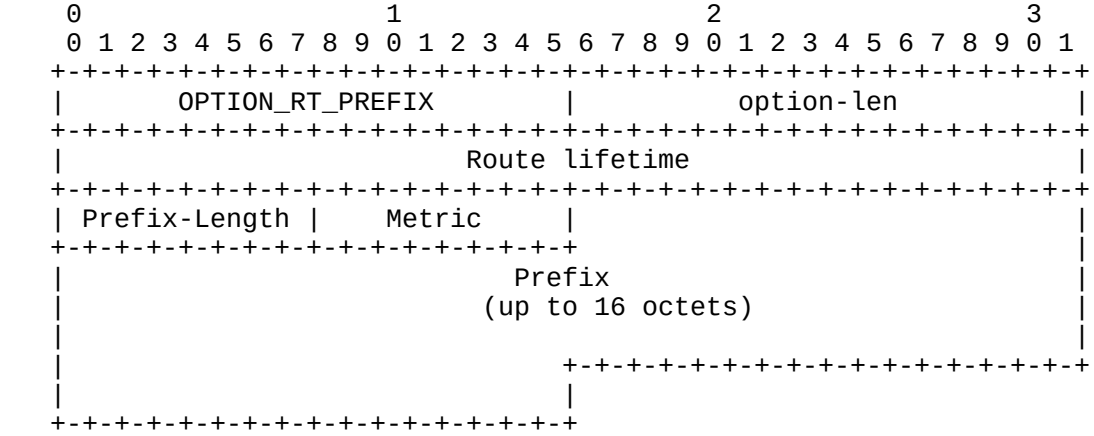

Figure 1: Route Prefix Option Format

option-code: OPTION_RT_PREFIX (TBD2).

option-len: Length of the Route Prefix option including all its sub-options.

Route lifetime 32-bit unsigned integer. Specifies lifetime of the route information, expressed in seconds (relative to the time the packet is sent). There are 2 special values defined. 0 means that route is no longer valid and must be removed by clients. A value of all one bits (0xffffffff) represents infinity. means infinity.

Prefix Length: 8-bit unsigned integer. The length in bits of the IP Prefix. The value ranges from 0 to 128. This field represents the number of valid leading bits in the prefix.

Resvd: Reserved field. Server MUST set this value to zero and client MUST ignore its content.

Metric: Route Metric. 8-bit signed integer. The Route Metric indicates whether to prefer the next hop associated with

this prefix over others, when multiple identical prefixes (for different next hops) have been received.

Prefix: a variable size field that specifies Rule IPv6 prefix. Length of the field is defined by prefix6-len field and is rounded up to the nearest octet boundary (if case when Prefix Length is not divisible by 8). In such case additional padding bits must be zeroed.

Values for metric field have meaning based on the value, i.e. higher value indicates higher preference.

3.2. Next Hop Option Format

Each IPv6 route consists of an IPv6 next hop address, an IPv6 destination prefix (a.k.a. the destination subnet), and a host preference value for the route. Elements of such route (e.g. Next hops and prefixes associated with them) are conveyed in NEXT_HOP option that contains RT_PREFIX suboptions.

The Next Hop Option defines the IPv6 address of the next hop, usually corresponding to a specific next-hop router. For each next hop address there can be zero, one or more prefixes reachable via that next hop.

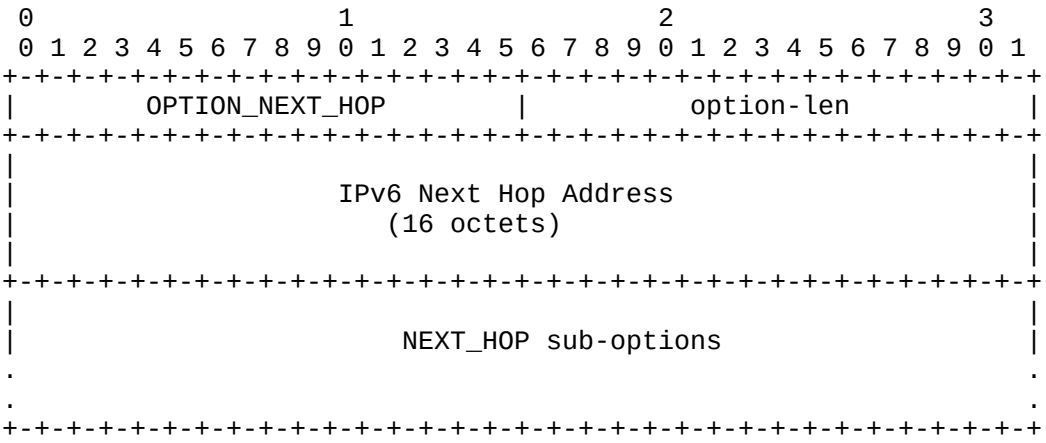

Figure 2: IPv6 Next Hop Option Format

option-code: OPTION_NEXT_HOP (TBD1).

option-len: 16 + Length of NEXT_HOP options field.

IPv6 Next Hop Address: 16 octet long field that specified IPv6 address of the next hop.

NEXT_HOP options: Options associated with this Next Hop. This includes, but is not limited to, zero, one or more RT_PREFIX options that specify prefixes reachable through the given next hop.

NEXT_HOP options: Options associated with this Next Hop. This includes, but is not limited to, zero, one or more SOURCE_AP and RT_PREFIX options that specify prefixes reachable through the given next hop.

3.3. Source Address/Prefix Option Format

Each IPv6 route consists of an IPv6 next hop address, an IPv6 destination prefix (a.k.a. the destination subnet), and a host preference value for the route. Elements of such route (e.g. Next hops and prefixes associated with them) are conveyed in NEXT_HOP option that contains RT_PREFIX suboptions.

The Source Address/Prefix Option defines the source IPv6 prefix/ address that are assigned from the prefixes that belong to this next hop.

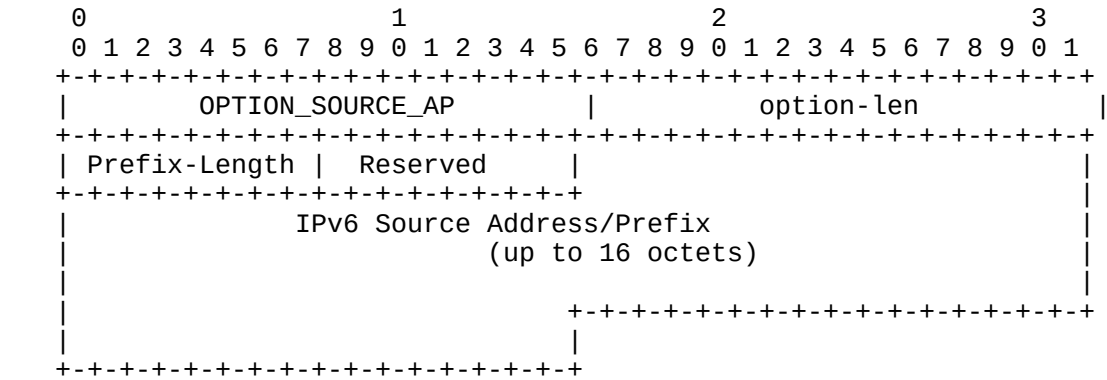

Figure 3: IPv6 Source Address/Prefix Option Format

option-code: OPTION_SOURCE_AP (TBD1).

option-len: 16 + Length of SOURCE_AP options field.

Prefix Length: 8-bit unsigned integer. The length in bits of the IP Prefix. The value ranges from 0 to 128. This field

represents the number of valid leading bits in the prefix. In case of source address this field is set to 132.

Resvd: Reserved field. Server MUST set this value to zero and client MUST ignore its content.

IPv6 Source Address/Prefix: 16 octet long field that specified IPv6 source address or source prefix.

4. DHCPv6 Server Behavior

When configured to do so, a DHCPv6 server shall provide the NEXT_HOP and RT_PREFIX Options in ADVERTISE and REPLY messages sent to a client that requested the route option. Each Next Hop Option sent by the server must convey at least one Route Prefix Option.

Server includes NEXT_HOP option with possible RT_PREFIX suboptions to designate that specific routes are available via routers. Server includes RT_PREFIX options in Next Hop sub-options directly in Advertise and Reply messages to inform that specific routes are available directly on-link.

If there is more than one route available via specific next hop, server MUST send only one NEXT_HOP for that next hop, which contains multiple RT_PREFIX options. Server MUST NOT send more than one identical (i.e. with equal next hop address field) NEXT_HOP option.

When configured to do so, a DHCPv6 server shall send one or more NEXT_HOP options that contain one or more source addresses Figure 3 included in the Next Hop sub-options field. Each Next Hop Address may be associated with zero, one or more Source Prefix that represent the source addresses that are assigned from the prefixes that belong to this next hop. The Next Hop sub-options field MAY contain Route Prefix options that represent the IPv6 destination prefixes reachable via the given next hop as defined in Figure 2. When configured to do so, a DHCPv6 server shall send NEXT_HOP option with Route Prefix option and Source Prefix in the message in the Next Hop sub-options field to indicate that given prefix is available directly on-link and that any source addresses derived from the source prefix will not be subject to ingress filtering on these routes supported by these next hops.

When configured to do so, a DHCPv6 server shall send one or more NEXT_HOP option that specify the IPv6 next hop addresses and source address. Each Next Hop Address option may be associated with zero, one or more Source Address that represent the source addresses that are assigned from the prefixes that belong to this next hop. The Next Hop sub-options field shall contain Source Address Figure 3 and

Route Prefix options Figure 1 that represent the IPv6 destination prefixes reachable via the given next hop. DHCPv6 server shall include Next Hop Address with Source Address and Route Prefix option in Next Hop sub-options field in the message to indicate that given prefix is available directly on-link and that the source address will not be subject to ingress filtering. For the Source Address, Source Address/Prefix option Figure 3 is used with prefix length set to 128.

Each Next Hop Address may be associated with zero, one or more Source Prefix that represent the source addresses that are assigned from the prefixes that belong to this next hop. The option MAY contain Route Prefix options that represent the IPv6 destination prefixes reachable via the given next hop. DHCP server shall include Next Hop Address with Route Prefix option in Next Hop sub-option field defined in Figure 2 in the message to indicate that given prefix is available directly on-link. To indicate that any source addresses derived from the source prefix will not be subject to ingress filtering on these routes supported by these next hops DHCPv6 server shall send two options, Next Hop option with Route Prefix option in Next Hop options field and a Source Prefix option defined in Figure 3.

Servers SHOULD NOT send NEXT_HOP or RT_PREFIX to clients that did not explicitly requested it, using the ORO.

Servers MUST NOT send NEXT_HOP or RT_PREFIX in messages other than ADVERTISE or REPLY.

Servers MAY also include Status Code Option, defined in Section 22.13 of the [RFC3315] to indicate the status of the operation.

Servers MUST include the Status Code Option, if the requested routing configuration was not successful and SHOULD use status codes as defined in [RFC3315] and [RFC3633].

The maximum number of routing information in one DHCPv6 message depend on the maximum DHCPv6 message size defined in [RFC3315]

5. DHCPv6 Client Behavior

A DHCPv6 client compliant with this specification MUST request the NEXT_HOP and RT_PREFIX Options in an Option Request Option (ORO) in the following messages: Solicit, Request, Renew, Rebind, and Information-Request. The messages are to be sent as and when specified by [RFC3315].

When processing a received Route Options a client MUST substitute a received 0::0 value in the Next Hop Option with the source IPv6 address of the received DHCPv6 message. It MUST also associate a

received Link Local next hop addresses with the interface on which the client received the DHCPv6 message containing the route option. Such a substitution and/or association is useful in cases where the DHCPv6 server operator does not directly know the IPv6 next-hop address, other than knowing it is that of a DHCPv6 relay agent on the client LAN segment. DHCPv6 Packets relayed to the client are sourced by the relay using this relay's IPv6 address, which could be a link local address.

The Client SHOULD refresh assigned route information periodically. The generic DHCPv6 Information Refresh Time Option, as specified in [RFC4242], can be used when it is desired for the client to periodically refresh of route information.

The routes conveyed by the Route Option should be considered as complimentary to any other static route learning and maintenance mechanism used by, or on the client with one modification: The client MUST flush DHCPv6 installed routes following a link flap event on the DHCPv6 client interface over which the routes were installed. This requirement is necessary to automate the flushing of routes for clients that may move to a different network.

Client MUST confirm that routers announced over DHCPv6 are reachable, using one of methods suitable for specific network type. The most common mechanism is Neighbor Unreachability Detection (NUD), specified in [RFC4861]. Client SHOULD use NUD to verify that received routers are reachable before adjusting its routing tables. Client MAY use other reachability verification mechanisms specific to used network technology. To avoid potential long-lived routing black holes, client MAY periodically confirm that router is still reachable.

5.1. Conflict resolution

Information received via Route Options over DHCPv6 MUST be treated equally to routing information obtained via other sources. In particular, from the RA perspective, DHCPv6 provisioning should be treated as if yet another RA was received. Preference field should be taken into consideration during route information processing. In particular, administrators are encouraged to read [RFC4191], Section 4.1 for guidance.

To facilitate information merge between DHCPv6 and RA, DHCPv6 options in this document convey the same information specified in [I-D.sarikaya-6man-next-hop-ra].

To facilitate information merge between DHCPv6 and RA, DHCPv6 option RT_PREFIX conveys the same information specified in [RFC4191] albeit on-wire format is slightly different. The differences are:

Metric field is an 8-bit field that conveys the route metric.

RIO uses 128-length prefix field, while DHCPv6 option uses variable prefix length. That difference is used to minimize packet size as it avoid transmitting zeroed octets. Despite slightly different encoding, delivered information is exactly the same.

If prefix is available directly on-link, Route Prefix option is conveyed directly in DHCPv6 message, not within Next Hop option. That feature is considered a superset, compared to RIO.

In short, when DHCPv6 RT_PREFIX option is used alone this specification works in compatibility mode with [RFC4191].

6. IANA Considerations

IANA is kindly requested to allocate DHCPv6 option code TBD1 to the OPTION_NEXT_HOP, TBD2 to OPTION_RT_PREFIX, TBD3 to OPTION_SOURCE_AP. All values should be added to the DHCPv6 option code space defined in Section 24.3 of [RFC3315].

7. Security Considerations

The overall security considerations discussed in [RFC3315] apply also to this document. The Route option could be used by malicious parties to misdirect traffic sent by the client either as part of a denial of service or man-in-the-middle attack. An alternative denial of service attack could also be realized by means of using the route option to overflowing any known memory limitations of the client, or to exceed the client’s ability to handle the number of next hop addresses.

Neither of the above considerations are new and specific to the proposed route option. The mechanisms identified for securing DHCPv6 as well as reasonable checks performed by client implementations are deemed sufficient in addressing these problems.

It is essential that clients verify that announced routers are indeed reachable, as specified in Section 5. Failing to do so may create black hole routing problem.

This mechanism may introduce severe problems if deployed in networks that use dynamic routing protocols. See Section 2.6 for details.

DHCPv6 becomes a complete provisioning protocol with this mechanism, i.e. all necessary configuration parameters may be delivered using DHCPv6 only. It was suggested that in some cases this may lead to decision of disabling RA. While RA-less networks could offer lower operational expenses and protection against rogue RAs, they would not work with nodes that do not support this feature. Therefore such decision is not recommended, unless all effects are carefully analyzed. It is worth noting that disabling RA support in hosts would solve rogue RA problem, it would in fact only change the issue into rogue DHCPv6 problem. That is somewhat beneficial, however, as rogue RA may affect all nodes immediately while rogue DHCPv6 server will affect only new nodes, that boot up after rogue server manifests itself.

Reader is also encouraged to read DHCPv6 security considerations document [I-D.ietf-dhc-sedhcpv6].

8. Acknowledgements

The author acknowledges the work done by his co-authors in MIF WG draft entitled DHCPv6 Route Options.

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3315] Droms, R., Bound, J., Volz, B., Lemon, T., Perkins, C., and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, July 2003.

[RFC3633] Troan, O. and R. Droms, "IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6", RFC 3633, December 2003.

9.2. Informative References

[I-D.ietf-dhc-sedhcpv6]
Jiang, S., Shen, S., Zhang, D., and T. Jinmei, "Secure DHCPv6 with Public Key", draft-ietf-dhc-sedhcpv6-03 (work in progress), June 2014.

[I-D.ietf-mif-mpvd-arch]
Anipko, D., "Multiple Provisioning Domain Architecture", draft-ietf-mif-mpvd-arch-01 (work in progress), May 2014.

[I-D.sarikaya-6man-next-hop-ra]
Sarikaya, B., "IPv6 RA Options for Next Hop Routes", draft-sarikaya-6man-next-hop-ra-02 (work in progress), June 2014.

[RFC3442] Lemon, T., Cheshire, S., and B. Volz, "The Classless Static Route Option for Dynamic Host Configuration Protocol (DHCP) version 4", RFC 3442, December 2002.

[RFC4191] Draves, R. and D. Thaler, "Default Router Preferences and More-Specific Routes", RFC 4191, November 2005.

[RFC4242] Venaas, S., Chown, T., and B. Volz, "Information Refresh Time Option for Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 4242, November 2005.

[RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, September 2007.

[RFC6204] Singh, H., Beebee, W., Donley, C., Stark, B., and O. Troan, "Basic Requirements for IPv6 Customer Edge Routers", RFC 6204, April 2011.

[RFC7157] Troan, O., Miles, D., Matsushima, S., Okimoto, T., and D. Wing, "IPv6 Multihoming without Network Address Translation", RFC 7157, March 2014.

Author’s Address

Behcet Sarikaya
Huawei USA
5340 Legacy Dr.
Plano, TX 75024
United States

Phone: +1 972-509-5599
Email: sarikaya@ieee.org