Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 27, 2012

W. Dec
Cisco Systems
T. Mrugalski, Ed.
ISC
T. Sun
China Mobile
B. Sarikaya
Huawei USA

February 24, 2012

# DHCPv6 Route Options
# draft-ietf-mif-dhcpv6-route-option-04

## Abstract

This document describes DHCPv6 Route Options for provisioning IPv6 routes on DHCPv6 client nodes. This is expected to improve the ability of an operator to configure and influence a nodes' ability to pick an appropriate route to a destination when this node is multi-homed and 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 August 27, 2012.

## 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 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, where DHCPv6 [RFC3315] is seen to be more viable.

This draft defines the DHCPv6 Route Options for provisioning IPv6 routes on DHCPv6 clients. The proposed option is primarily envisaged for use by DHCPv6 client nodes that are capable of making basic IP routing decisions and maintaining an IPv6 routing table, broadly in line with the capabilities of a generic host as described in [RFC4191].

Throughout the document the words node and client are used as a reference to the device with such routing capabilities, hosting the DHCPv6 client software. The route information is taken to be equivalent to static routing, and limited in the number of required routes to a handful.

## 2. Problem overview

The solution described in this document applies to multi-homed scenarios including ones where the client is simultaneously connected to multiple access network (e.g. WiFi and 3G). The following scenario is used to illustrate the problem as found in typical multi-homed residential access networks. It is duly noted that the problem is not specific to IPv6, occurring also with IPv4, where it is today solved by means of DHCPv4 classless route information option [RFC3442], or alternative configuration mechanisms.

In multi-homed networks, a given user's node may 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. In such multi-homed networks it is quite common for the network operator to offer the delivery of a particular type of IP service via a particular gateway, where the service can be characterised by means of specific destination IP network prefixes. Thus, from an IP routing perspective in order for the user node to select the appropriate gateway for a given destination IP prefix,

recourse needs to be made to classic longest destination match IP routing, with the node acquiring such prefixes into its routing table. This is typically the remit of dynamic Internal Gateway Protocols (IGPs), which however are rarely used by operators in residential access networks. This is primarily due to operational costs and a desire to contain the complexity of user nodes and IP Edge devices to a minimum. While, IP Route configuration may be achieved using the ICMPv6 extensions defined in [RFC4191], this mechanism does not lend itself to other operational constraints such as the desire to control the route information on a per node basis, the ability to determine whether a given node is actually capable of receiveing/processing such route information. A preferred mechanism, and one that additionally also lends itself to centralized management independent of the management of the gateways, is that of using the DHCP protocol for conveying route information to the nodes.

## 3. Motivation

The following section enumerates use cases, both in existing networks and as well as in evisaged future deployments. Usage scenarios are specified here in no particular order. As those use cases are descibed by various network operators, their scenarios may partially overlap.

Discussion: this section is rather long. Nevertheless, there were concerns raised that such option is not needed. Such extensive list can possible solve those concerns.Number of use cases should be limited in future revisions. Alternatively, they can be moved to a separate motivation draft, if needed.

### 3.1. Use cases

Use case 1: In Broadband network environment where the CPE is multi-homed to two upstream edge routers and each router provides connectivity for different types of services for example internet access and Video on Demand (restricted inside a walled garden) and the Service Provide would like to avoid routing on the CPE, there is a need to provision static route entries on RGs/CPEs. Service Provider requires a centralized control/management point for storing the customer's related innformation (IPv6 prefix, IPv6 routes and other provisioned information) and DHCPv6 is a good place for that. Using RA's would require to manually provision the edge router and this operation is not always possible, for example when router is operated by 3rd party. Broadband Forum document WT-124 issue 3 [BBF-WT-124] calls for this draft to solve the problem.

Use case 2: Operators want (approximate) feature parity so that they

can have (approximate) alignment between their operational procedures for v4 and v6, especially in a dual stack network. Having similar mechanisms for both protocols is desired due to lower operational expenses (OPEX).

Use case 3: In cellular networks, it is efficient for the network to configure routing information in central DHCPv6 server to do unified routing policy information. The gateways (GGSN in cellular network) only need to perform DHCPv6 relay. The Option code sent by clients can be used as an indication that host is MIF capable, so that network need not to do such configuration to host without MIF capabilities.

Use case 4: In cellular network, DHCPv6 is used for IPv6 parameter configuration and RA is used for SLACC of handset. This behavior was introduced in 3GPP Release 8 (or earlier). The network gateway in cellular network (e.g., GGSN) can naturally support DHCPv6 extension since the gateway acts as a DHCPv6 relay. However, it is very hard to update those gateways to use RA announcing the route information. The handsets with MIF feature need to visit subscribed/operator provided service. Some traffic is routed to the operator's network through 3G interface instead of to Internet through WiFi. DHCPv6 will be used to configure these specific routes. This use case is described in [THREEGPP-23.853].

Use case 5: PMIPv6 use case in LTE network. In LTE cellular network, both GTP and PMIPv6 are used for mobility management. In GTP, it is a point-to-point link between mobile host and PGW (PDN Gateway). However, in PMIPv6 case, the point-to-point link is between mobile host and SGW(Serving Gateway). The PGW sends /64 prefix to SGW through PBA. The SGW sends RA to mobile host. Route option may be needed when the host is multi-homed if it is simultaneously connected to the cellular network and WiFi or it simultaneously connects to multiple APNs in the cellular network. If RA is used for route configuration, both PGW and SGW(whose number is larger than PGW) need to be updated. Moreover, since a host can only connect to one SGW at a time, the SGW have to keep multiple route information received from different PGWs for one host and send them by RA to the host separately. This makes RA is not favorable in this use case.

Use case 6: WiFi networks. Some WiFi hotspots provide local services ("walled garden"). The route configuration on hosts or RGs is needed to direct some traffic to local network, while other traffic to the Internet. While this can be achieved using Route Information Option (RIO) in RA for all nodes that support [RFC4191], it does not allow doing so on a per-host basis.

Use case 7: VPN network. When a user connects to enterprise VPN

network, the routing of VPN traffic need to be configured. Due to the large number of such VPN networks, we cannot assume all the VPN network only use RA. DHCPv6 provides another choice which may be preferred by the VPN network. This situation is described in [RFC4191], Section 5.2. Hosts that do not support RFC4191 will not operate properly.

Use case 8: Selective walled garden. Figure 1 illustrates the case of two clients connected to a shared link. Both clients are assumed to have global IPv6 addresses and obtain their Internet connectivity via Router2 by means of a configured or a discovered default route. Client 1 however, unlike Client 2, is intended to run a specific application, e.g. VoIP, that is meant to access ServerA by means of Router1 with Server A being otherwise not reachable from the Internet. In addition to the global IP address Client1 may be assigned with another IP address of a more restricted scope for the purpose of communicating with Server A.

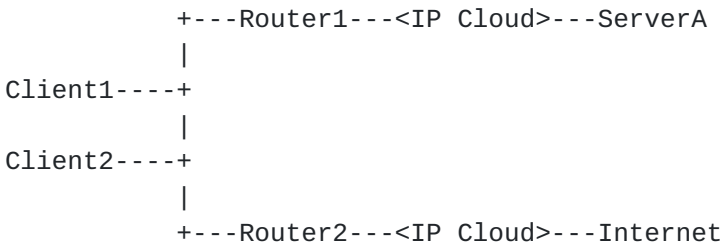


Figure 1: Walled garden scenario

The problem in the above scenario comes down to the fact that in order to reach Server A, Client1 requires to use a more specific route whose next-hop address is Router1. An ICMPv6 based mechanism for disseminating more specific route information, as defined in [RFC4191], disseminates this information via the shared link also to Client2. Often the operator wants to avoid this redundant dissemination to passing to Client2. In addition many operators prefer to be able to manage specific client route information from a centralized repository instead of managing directly such configuration on a router, as is required with the ICMPv6 based scheme. The former requirement is driven by the desire to provide to each client only the information required for their intended role which may be tied to a specific service, as well as to allow the possibility to introduce other routers into the scenario for purposes of load sharing. The requirement for more centralized configuration management is often due to administrative boundaries within an operator's organization as well as an existing operational practice that are in place for IPv4, all of which make router based configuration difficult.

Use case 9: Multihoming problem. A multihomed IPv6 host or gateway needs to solve at least 3 problems to operate properly when more than one link is operational:

1. Source address selection

2. Next-hop selection

3. DNS server selection

Problems one and three are solved by [I-D.ietf-6man-addr-select-opt] and [I-D.ietf-mif-dns-server-selection], respectively. It should be noted that both mechanisms use DHCPv6 as well. This draft attempts to solve problem two. Below is a brief explanation of the problem. See draft [I-D.ietf-v6ops-ipv6-multihoming-without-ipv6nat] for detailed problem analysis, background information and additional discussion regarding the need for a DHCPv6 solution to route information problem and IPv6 multihoming in general (with focus on aforementioned 3 problems).

In multihoming environment, server can restrict assignment of additional prefixes only to hosts that support more advanced next-hop and address selection requirements. (See Section 5.2 of [I-D.ietf-v6ops-ipv6-multihoming-without-ipv6nat]). Obviously this MUST be done on a per-host basis. Information about node capability is obtained via Option Request Option (ORO) in Solicit message, so support for Route Options is also used as means to report node capabilities to a network.

Use case 10: In static networks (i.e. networks that have static routers that are not changing over time, like home network with), such as some enterprise, hosting provider networks or even home network with a single router, it may be possible to stop using RA mechanism and deliver all configuration parameters to hosts using DHCPv6 only. This approach solves the rogue RA problem (i.e. a node that is not an approved router starts announcing RA in a network may hijack traffic from other hosts). This approach may be appealing in some cases, but not in all. For example if there is security association shared between clients and a DHCPv6 server, it may be useful to trust DHCP and disable RA mechanism. Also, environments that need DHCP for extended information, including but not limited to communicating information like DNS servers, hostnames, NTP servers, TFTP boot information and so on are forced to run two protocols increasing complexity and troubleshooting, where we have proof of concept in IPv4 that only one protocol (DHCP) should be needed.

Use case 11: It also has been proposed that route information option may be used as tie breaker in networks that deploy both DHCPv6 route

option and RA. DHCPv6 server could announce routing information along with RA. Legitimate router is also announced over DHCPv6. Host that receives conflicting information over RA may use additional information received from DHCPv6 as a tie breaker. This proposal [nanog-beijnum] was not investigated further.

Use case 12: DHCP-based configuration provides different failure mode than RA. While RA-based configuration works better in networks that offer redundant uplink using separate routers (second router can quickly take over upstream traffic), there are many deployments that cannot use that advantage, because of a single uplink. Current home networks with a single uplink as most obvious example. On the other hand, RA is more severly impacted by rogue entity problem. New rogue RA device may instantly break all other devices on the network. New rogue DHCP server will cause no immediate harm, may cause slow breakage over time, and may in fact never cause any breakage. This is due to the funamental design choices of each protocol and it is hard to make either work the other way.

Use case 13: DHCP-based configuration may use mostly unicast traffic, while RA-based configuration mostly uses multicast. In some environments implementing multicast traffic may be cumbersome, e.g. in WiMAX environment not every subscriber station (SS) supports multicast channels and multicast capability must be emulated by base station (BS) using redundant transmissions. Classic, stateless, multicasted RA is in disadvantage compared to DHCP with standard unicast option enabled. While it is possible to selectively send unicasted RAs to selected subscribers, such architecture is essentially a stateful RA, thus forfeiting major benefit of RA being stateless.

Use case 14: Separated networks. In networks that do not have any routers, two DHCPv6 clients get a global address from DHCPv6 server. They cannot ping each other due to the fact that they do not know prefix that is available on-link. While it is tempting to suggest that separated networks should use link-local addressing, other factors should be taken into consideration. A stateful DHCPv6 may be used as a node monitoring tool, thus having avantage over link-local address usage. The also may be sensor networks that have outside connectivity only sporadically, e.g. uplink is established periodically to gather readings, but most of the time router is powered down for power reasons. Route Option in DHCPv6 could be used to configure on-link routes, while router could announce itself using short-lived RA.

Those requirements and use cases can be summarized as following:

1. In view of the DHCPv6 requirements in several fields, vendor-specific options lead to several segmented definitions. An IETF defined general option is a better choice.

2. Per user/host configuration makes DHCPv6 be used for the on-demand configuration.

3. As there is no well-defined central management system for prefix delegation and routing options va RA, it seems that DHCPv6 is the only available solution. It is better to have a generic option then a bunch of competing vendor options.

4. While this work was initially started with multihoming in mind, it is useful for single interface devices as well.

In a sense this route configuration mechanism makes DHCPv6 complete. Without it, this protocol cannot fully provision all configuration parameters to a host on its own.

### 3.2. Raised concerns

Opponents of this option proposed several alternative approaches. This section attempts to address raised issues.

#### 3.2.1. Vendor-specific option

Claim: During discussion about route configuration, some opponents say that routing information should be defined as vendor specific option.

Response: There are many ISPs, cellular and BBF network operators, CPE vendors, hardware vendors, DHCP implementors that want to implement and deploy this mechanism. Using vendor-specific option would severly limit interoperability and would make adoption and deployment much more complicated.

This solution is not a technology-specific requirement, it is requested by wide variety of companies, so it is not a vendor specific.

#### 3.2.2. Unicast RA

Claim: Some proponents insist that instead of using DHCPv6 solution, RA should be used instead. Some propose to send unicast RA with RIO option on a per-host basis.

Response: While this approach technically does not violate existing specs, it uses RA in a stateful way, thus the benefit of RA being

stateless is lost. Furthermore, it would require deploying additional mechanism, like RADIUS to deliver necessary information about hosts to routers. Authors consider deploying such stateful RA server with RADIUS support more complicated to deploy than the solution it tries to avoid (DHCPv6).

As there is no well-defined central management system for prefix delegation and routing options va RA, it seems that DHCPv6 is the only available solution. It is better to have a generic option than a bunch of competing vendor options.

Another concern raised is that RIO is not menadatory nor optional in 3GPP system and there is currently not support in 29.061 RADIUS or Diameter profile, so use of that alternative is somewhat limited in some cases.

### 3.2.3. DHCPv6 requires client to use one server

Claim: DHCPv6 has less rich semantics as client has to pick one out of all available server.

Response: While that is how currently most clients are implemented, there is nothing in [RFC3315] that mandates that. It is true that DHCPv6 was not designed with several provisioning domains. On the contrary, section 17.1.3 states that "Upon receipt of one or more valid Advertise messages, the client selects one or more Advertise messages based upon the following criteria.". This means that DHCPv6 client can obtain parameters from all available DHCPv6 servers, not just selected one. As such, DHCPv6 may work with overlapping provisioning domains. Authors acknowledge that this possibility is currently rather theoretical, as most known implementations do not take advantage of that possibility.

### 3.2.4. Use VLANs

Claim: There was a proposal to use VLANs as a solution to lack of per-host capability in RA mechanism.

Response: Deploying VLANs complicates network topology much more than adding a single DHCPv6 option. Furthermore in many cases it is not possible to deploy VLANs in any reasonable way, e.g. in multihost environment. Also, low cost devices (e.g. CPE) often do not offer VLAN capabilities, but they are very much capable of supporting DHCPv6. Another objection of estetic nature. Using layer 2 mechanisms to work around limitations in layer 3 is not elegant.

## 4. 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. [I-D.ietf-v6ops-ipv6-multihoming-without-ipv6nat] 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 request 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.

### 4.1. Default route configuration

Defined mechanism may be used to configure default route. Default route is configured using RT_PREFIX option that specifies ::/0 route, included as suboption in NEXT_HOP.

Server MUST NOT define more than one default route.

### 4.2. 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.

### 4.3. 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.

### 4.4. 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.

### 4.5. 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 4.6 about limitations regarding dynamic routing.

### 4.6. 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]).

User is also encouraged to read Section 3.2.

## 5. 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].

## 5.1. 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.

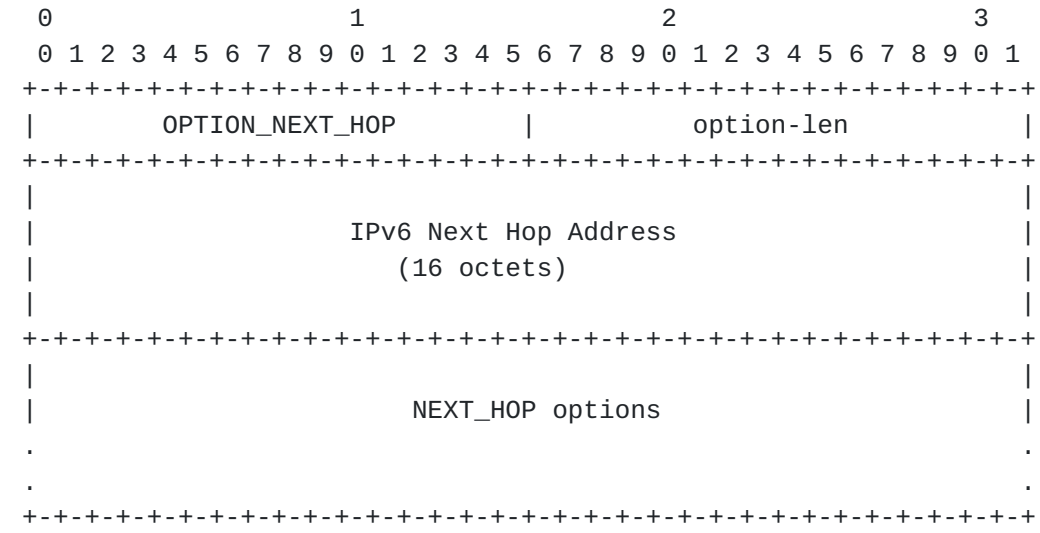


Figure 2: IPv6 Next Hop Option Format

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.

## 5.2. 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.

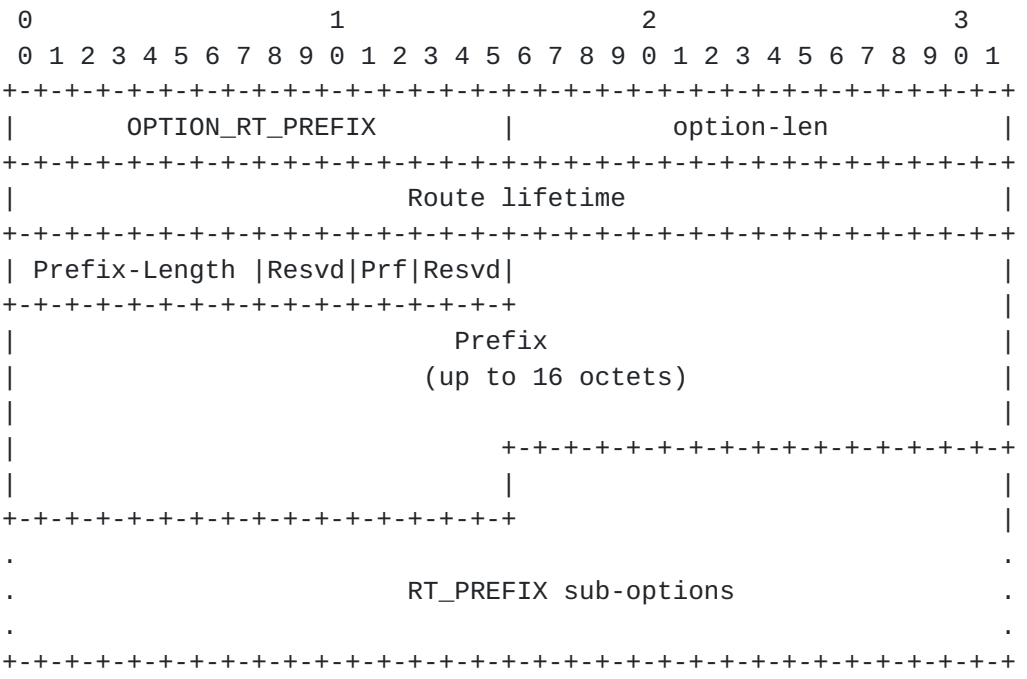


Figure 3: Route Prefix Option Format

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.

Prf(Route Preference): 2-bit signed integer. The Route Preference indicates whether to prefer the router associated with this prefix over others, when multiple identical prefixes (for different routers) have been received. If the Reserved (10) value is received, the Route Information Option MUST be ignored.

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 prefix6-len is not divisible by 8). In such case additional padding bits must be zeroed.

RT_PREFIX options: Options specific to this particular prefix.

Values for preference field have meaning identical to Route Information Option, defined in [RFC4191], Section 2.1:

01 High

00 Medium (default)

11 Low

10 Reserved - MUST NOT be sent

## 6. DHCPv6 Server Behavior

When configured to do so, a DHCPv6 server shall provide the Next Hop and Route 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 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.

Servers SHOULD NOT send Route Option to clients that did not explicitly requested it, using the ORO.

Servers MUST NOT send Route Option 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]

## 7. 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 reachibality verification mechanisms specific to used network technology. To avoid potential long-lived routing black holes, client MAY periodically confirm that router is still reachable.

### 7.1. 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 option conveys the same information as RIO, specified in [RFC4191], albeit on-wire format is slightly different. The differences are:

Metric field (available in previous version of this draft) has been replaced with 2-bit preference field that is in line with RIO information.

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 withing Next Hop option. That feature is considered a superset, compared to RIO.

## 8. IANA Considerations

IANA is kindly requested to allocate DHCPv6 option code TBD1 to the OPTION_NEXT_HOP and TBD2 to OPTION_RT_PREFIX. Both values should be

added to the DHCPv6 option code space defined in Section 24.3 of [RFC3315].

## 9. 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 7. 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 4.6 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-secure-dhcpv6].

## 10. Contributors and Acknowledgements

This document would not have been possible without the significant

contribution provided by: Arifumi Matsumoto, Hui Deng, Richard Johnson, and Zhen Cao.

The authors would also like to thank Alfred Hines, Ralph Droms, Ted Lemon, Ole Troan, Dave Oran, Dave Ward, Joel Halpern, Marcin Siodelski, Alexandru Petrescu, Roberta Maglione, Tim Chown, Brian Carpenter, Dave Thaler, Lorenzo Colitti and Leo Bicknell for their comments and useful suggestions.

This work has been partially supported by Department of Computer Communications (a division of Gdansk University of Technology) and the Polish Ministry of Science and Higher Education under the European Regional Development Fund, Grant No. POIG.01.01.02-00-045/09-00 (Future Internet Engineering Project).

## 11. References

### 11.1. 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.

### 11.2. Informative References

[BBF-WT-124]
Broadband Forum, "BBF WT-124 issue 3", BBF WT-124i3, 2011.

[I-D.ietf-6man-addr-select-opt]
Matsumoto, A., Fujisaki, T., Kato, J., and T. Chown, "Distributing Address Selection Policy using DHCPv6", draft-ietf-6man-addr-select-opt-03 (work in progress), February 2012.

[I-D.ietf-dhc-secure-dhcpv6]
Jiang, S. and S. Shen, "Secure DHCPv6 Using CGAs", draft-ietf-dhc-secure-dhcpv6-04 (work in progress), December 2011.

[I-D.ietf-mif-dns-server-selection]

Savolainen, T., Kato, J., and T. Lemon, "Improved DNS Server Selection for Multi-Interfaced Nodes", draft-ietf-mif-dns-server-selection-07 (work in progress), October 2011.

[I-D.ietf-v6ops-ipv6-multihoming-without-ipv6nat]
Troan, O., Miles, D., Matsushima, S., Okimoto, T., and D. Wing, "IPv6 Multihoming without Network Address Translation", draft-ietf-v6ops-ipv6-multihoming-without-ipv6nat-04 (work in progress), February 2012.

[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.

[THREEGPP-23.853]
Stojanovski, S., "3GPP TR 23.853: Operator Policies for IP Interface Selection (OPIIS)", 3GPP TR 23.853, August 2011, <http://www.3gpp.org/ftp/Specs/html-info/23853.htm>.

[nanog-beijnum]
van Beijnum, I., "", , June 2011, <http://mailman.nanog.org/pipermail/nanog/2011-June/037242.html>.

Authors' Addresses

Wojciech Dec
Cisco Systems
Haarlerbergweg 13-19
1101 CH Amsterdam
The Netherlands

Email: wdec@cisco.com

Tomasz Mrugalski (editor)
Internet Systems Consortium, Inc.
950 Charter Street
Redwood City, CA 94063
USA

Phone: +1 650 423 1345
Email: tomasz.mrugalski@gmail.com

Tao Sun
China Mobile
Unit2, 28 Xuanwumenxi Ave
Beijing, Xuanwu District 100053
China

Phone:
Email: suntao@chinamobile.com

Behcet Sarikaya
Huawei USA
1700 Alma Dr. Suite 500
Plano, TX 75075
United States

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