MIF Working Group
Internet-Draft
Intended status: Informational
Expires: January 14, 2011

J. Laganier
Qualcomm Inc.
G. Montenegro
Microsoft
J. Korhonen
Nokia Siemens Networks
T. Savolainen
Nokia
Z. Cao
China Mobile

July 13, 2010

MIF Current Practice Analysis
draft-cao-mif-analysis-01

Abstract

This document analyzes whether the problems encountered by a multi-homed host are satisfactorily addressed by mechanisms currently implemented in operating systems.

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 January 14, 2011.

Copyright Notice

Table of Contents

1. Introduction

A multihomed host have multiple provisioning domains via virtual and/or physical interfaces. A multihomed host receives node configuration information from each of its access networks, through various mechanisms such as DHCP, PPP's IPCP and IPv6 Router Advertisements. When the multihomed host receives various configuration objects (e.g., DNS server address, default gateway, address selection policies) with values that differ from one administrative domain to another, the node has to decide which one to use or how to reconcile them.

Issues regarding how the multi-homed host uses the configuration objects have been addressed in [I-D.ietf-mif-problem-statement]. Current practices of how the various implementations handle these problems are introduced in [I-D.ietf-mif-current-practices]. This document analyzes whether the problems encountered by a multi-homed host are satisfactorily addressed by mechanisms implemented in operating systems.

2. Terminology

The following terms are used throughout this document:

Multihomed Host: A host that is attached to one or more networks via one or more virtual and/or physical interfaces.

3. Problem Analysis

We group the problems raised in [I-D.ietf-mif-problem-statement] into specific categories as per the subsections below.

3.1. Naming and Addressing

1. The operating systems has node-scoped DNS server addresses but the DNS server addresses provided by a given domain are only reachable through that domain.

2. The answers to DNS queries returned by the DNS server of a given domain are only valid and/or reachable within that domain (e.g., split horizon DNS) but the operating system treats these answers as valid on any domain.

3. Private IPv4 addresses [RFC1918] and Unique Local IPv6 Unicast Addresses [RFC4193] are reachable from within a given domain (i.e., they are site-scoped) but the operating system doesn’t know the domain boundary and treats these as reachable on any domain (i.e., they have global scope.)

4. Private IPv4 addresses [RFC1918] are only unambiguous within a given domain but the operating system doesn’t know the domain boundary and cannot associate a Private IPv4 Address to a given domain and thus treats those as valid on any domain.

3.2. Routing

1. Routing tables entries to ambiguous subnet prefixes in [RFC1918] addressing space are only unambiguous within a given domain but the operating system doesn’t distinguish routes to the same prefixes belonging to different communication domains, thus leading to use of the wrong outbound interface and wrong destincation gateway.

2. Routing tables entries with an ambiguous next hop IP address in [RFC1918] addressing space are only to be used within a given domain but the operating system doesn’t necessarily know which was the communication domain thus leading to use of the wrong outbound interface and wrong destination gateway, and/or communication failure if no destination gateway is reachable at the destination address or if the destination gateway has no upstream route to the final destination of the packet.

3. Host implementations usually do not implement the [RFC1122] model where the Type-of-Service are in the routing table which could be use to choose between routes with same longest prefix match and

same metrics but different Type-of-Service characteristics, e.g., low delay, high throughput.

3.3. Reachability

1. Ingress filtering can prevent communication when a node sends packets from a source address allocated from a given domain to a (default) router in another domain.

2. Strong host model implementaion can cause incoming packets to be discarded when they are sent to a destination address assigned to one of the interface of the node that is not the interface on which the packet is incoming.

3. There is no interface between a router and a host for the router to indicate that there is no default route but only specific routes to some prefixes. As a result, a node that discovers a router assumes that any destination is reachable, which might not always be the case: in some case only connectivity to destination in the domain is available, and other destinations are unreachable, e.g., walled gardens, corporate intranets, etc.

3.4. Domain Selection

1. Application usually does not specify to which domain they want to communicate. When the destination has an unambiguous address the domain can sometimes be derived from that. This is however not the case when the destination is an ambiguous address from [RFC1918].

2. Some applications require domain affinity. There should be some way to set it either by the application itself or by the system on behalf of the application. Therefore the system should be cognizant of domains.

3.5. Configuration and Policy

1. Operating system does not keep separate, per domain copies of same configuration objects (e.g., DNS server addresses, NTP server addresses, ..) and thus these are either overwritten by the operating system when received from multiple provisioning domains, or ignored when not received on a so-called primary interface.

2. There’s no way yet to handle multiple policies coming from different domains. E.g., corporate node usage typically means that the corporation issues some policy on that Wi-Fi interface (and others as well). In this case, the carrier and corporation

domains and their policies will overlap over the Wi-Fi interface. Having a common policy language might help to detect and reason about such conflicts, but conflict resolution is another problem. Ultimately, the issue are the different authorities on these domains (e.g., user at home, admin at corporation and carrier for wireless broadband) and how they resolve their conflicts in the overlap situations. Note: Domains and their policies may span multiple interfaces. There is a fixed hierarchy of domains and their authorities, but the top authority may decide to delegate to others certain parts of the system and to their policies, as long as these don’t conflict with his. A conflict resolution that respects the hierarchy is needed.

4. Current Practice Analysis

4.1. Mobile Handset Operating Systems

4.1.1. Nokia S60 3rd Edition, Feature Pack 2

The following problems occurs:

Naming and Addressing:

Routing:

Reachability:

Domain Selection:

Configuration and Policy:

4.1.2. Microsoft Windows Mobile 2003 Second Edition

The following problems occurs:

Naming and Addressing:

Routing:

Reachability:

Domain Selection:

Configuration and Policy:

4.1.3. RIM BlackBerry

The following problems occurs:

Naming and Addressing:

Routing:

Reachability:

Domain Selection:

Configuration and Policy:

4.1.4. Google Android

The following problems occurs:

Naming and Addressing:

Routing:

Reachability:

Domain Selection:

Configuration and Policy:

4.1.5. Qualcomm AMSS

The following problems occurs:

Naming and Addressing:

Routing:

Reachability:

Domain Selection:

Configuration and Policy:

4.1.6. Arena Connection Manager

The following problems occurs:

Naming and Addressing:

Routing:

Reachability:

Domain Selection:

Configuration and Policy:

4.1.7. Access Selection

The following problems occurs:

Naming and Addressing:

Routing:

Reachability:

Domain Selection:

Configuration and Policy:

4.2. Computer Operating Systems

4.2.1. Microsoft Windows

The following problems occurs:

Naming and Addressing:

Routing:

Reachability:

Domain Selection:

Configuration and Policy:

4.2.2. Linux and BSD-based Operating Systems

The following problems occurs:

Naming and Addressing: 1, 2, 3, 4

Routing: 1, 2, 3

Reachability: 1, 2, 3

Domain Selection: 1, 2

Configuration and Policy: 1, 2

4.2.3. Apple MacOS X

The following problems occurs:

Naming and Addressing:

Routing:

Reachability:

Domain Selection:

Configuration and Policy:

5. Security Considerations

TBD.

6. IANA Considerations

This document does not require any IANA actions.

7. Informative References

[I-D.ietf-mif-current-practices]
Wasserman, M. and P. Seite, "Current Practices for Multiple Interface Hosts", draft-ietf-mif-current-practices-02 (work in progress), June 2010.

[I-D.ietf-mif-problem-statement]
Blanchet, M. and P. Seite, "Multiple Interfaces Problem Statement", draft-ietf-mif-problem-statement-05 (work in progress), July 2010.

[I.D-MIF-DNS]
Savolainen, T., "DNS Server Selection on Multi-Homed Hosts", February 2010, <draft-savolainen-mif-dns-server-selection-02.txt (work in progress)>.

[RFC1918] Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, February 1996.

[RFC4193] Hinden, R. and B. Haberman, "Unique Local IPv6 Unicast Addresses", RFC 4193, October 2005.

Authors' Addresses

Julien Laganier
Qualcomm Incorporated
5775 Morehouse Drive
San Diego, CA 92121
USA

Phone: +1 858 858 3538
Email: julienl@qualcomm.com

Gabriel Montenegro
Microsoft

Email: gmonte@microsoft.com

Jouni Korhonen
Nokia Siemens Networks
Linnoitustie 6
FI-02600 Espoo
FINLAND

Email: jouni.nospam@gmail.com

Teemu Savolainen
Nokia
Hermiankatu 12 D
FI-33720 Tampere
FINLAND

Email: teemu.savolainen@nokia.com

Zhen Cao
China Mobile

Email: zehn.cao@chinamobile.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 1, 2011

W. Dec, Ed.
Cisco Systems
T. Mrugalski
Gdansk University of Technology
T. Sun
China Mobile
B. Sarikaya
Huawei USA

September 28, 2010

DHCPv6 Route Option
draft-dec-dhcpv6-route-option-05

Abstract

This document describes DHCPv6 Route Options for provisioning IPv6 routes on nodes with DHCPv6 clients. This is expected to improve the ability of an operator to configure and influence a node's 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 April 1, 2011.

Copyright Notice

Table of Contents

1. Introduction

The Neighbor Discovery (ICMPv6) protocol [RFC4861] provides a mechanism for hosts to discover one or more default routers on a directly connected network segment. Extensions to the 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 allows network administrators to better 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 is seen to be more viable.

This draft defines the DHCPv6 Route Option 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 following scenario is used to illustrate the problem as found in 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 gateways. 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. 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 RG configuration information from a centralized DHCP server. [I-D.troan-multihoming-without-nat66] 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 option 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. The overall DHCPv6 design follow a similar approach to that used in the design of the IA_NA, IA_TA and IA_PD options in [RFC3633]

4. DHCPv6 Route Option

A DHCPv6 client interested in obtaining routing information includes the OPTION_IA_RT in its DHCPv6 Option Request Option (ORO) sent to a server. A Server, when configured to do so, provides the requested route information using the OPTION_IA_RT option. So as to allow the route option to be both extensible, as well as conveying detailed info for routes, use is made of a nested options structure. An IA_RT conveys one or more OPTION_NEXT_HOP options that specify the IPv6 next hop addresses. Each OPTION_NEXT_HOP conveys in turn one or more OPTION_RT_PREFIX options that represents the IPv6 destination prefixes reachable via the given next hop. The Formats of the OPTION_IA_RT, OPTION_NEXT_HOP and OPTION_RT_PREFIX are defined in the following sub-sections

The DHCPv6 Route Option format borrows from the principles of the Route Information Option defined in [RFC4191]. One notable exception with respect to [RFC4191] is however that a Route Lifetime element is not defined. The information conveyed by the DHCPv6 Route Option is considered valid until changed or refreshed by general events that trigger DHCPv6 or route table state changes on a node, thus not requiring a specific route lifetime. In the event that it is desired for the client to request a refresh of the route information (and other stateless DHCPv6 options), use of the generic DHCPv6 Information Refresh Time Option, as specified in [RFC4242] is envisaged.

4.1. DHCPv6 Route Option Format

To separate routing information from other options conveyed in a DHCPv6 message, the DHCPv6 Route Option is defined and is used to convey to a client one or more IPv6 routes. 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 IA_RT's options, rather than in the IA_RT option itself.

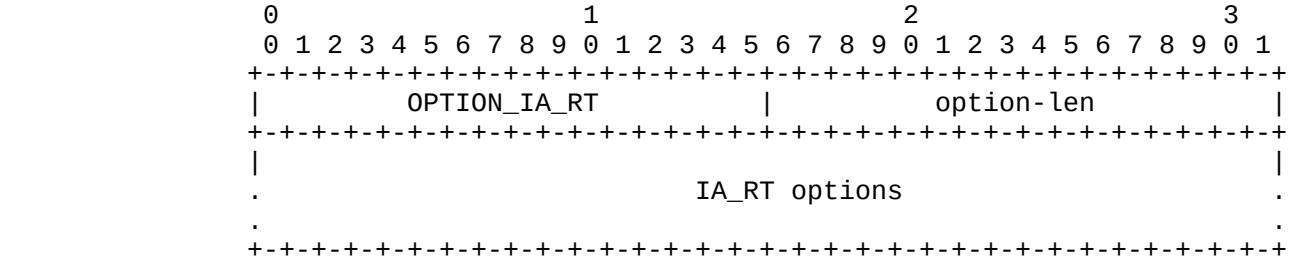

Figure 1: IPv6 Routes Option Format

option-code: OPTION_IA_RT (TBD).

option-len: Length of the IA_RT options field.

IA_RT options: Options associated with this IA_RT. This includes, but is not limited to, OPTION_NEXT_HOP options that specify next hop addresses.

The Route option MUST NOT appear in the following DHCPv6 messages: Solicit, Request, Renew, Rebind, Information-Request. The Route Option MAY appear in ADVERTISE and REPLY messages.

Discussion: Traditionally, grouping options (IA_NA, IA_TA and IA_RD) contain an identifier field (IAID) that must be unique among

identifiers generated by one client. It is used to differentiate between several options of the same type (e.g. several IA_NA options) that may be used simultaneously. However, it is assumed that client will never use more than one IA_RT option therefore such an identifier is not needed.

4.2. Next Hop Option Format

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 are one or more prefixes reachable via that next hop.

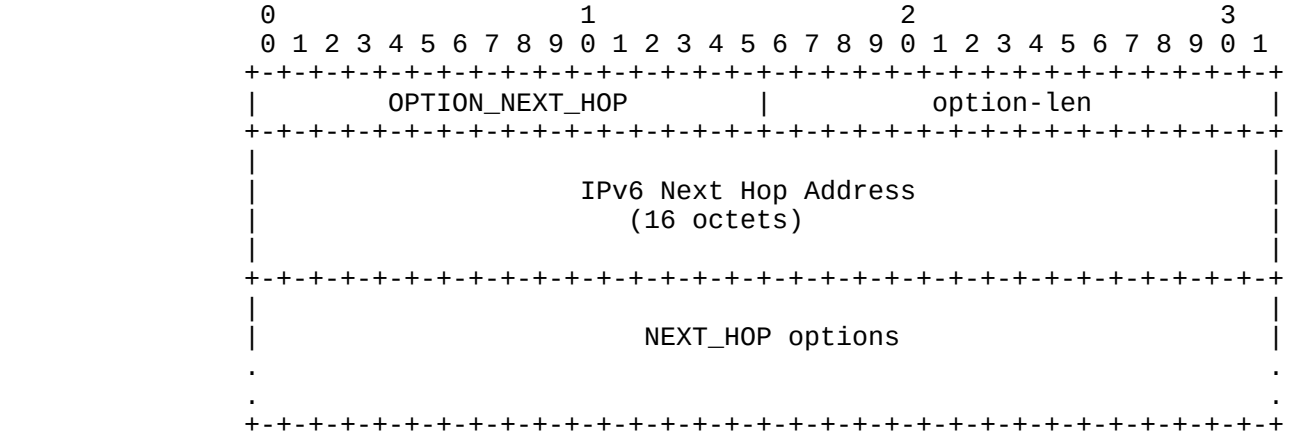

Figure 2: IPv6 Route Option Format

option-code: OPTION_NEXT_HOP (TBD).

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, OPTION_RT_PREFIX options that specify prefixes available via specified next hop.

4.3. 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.

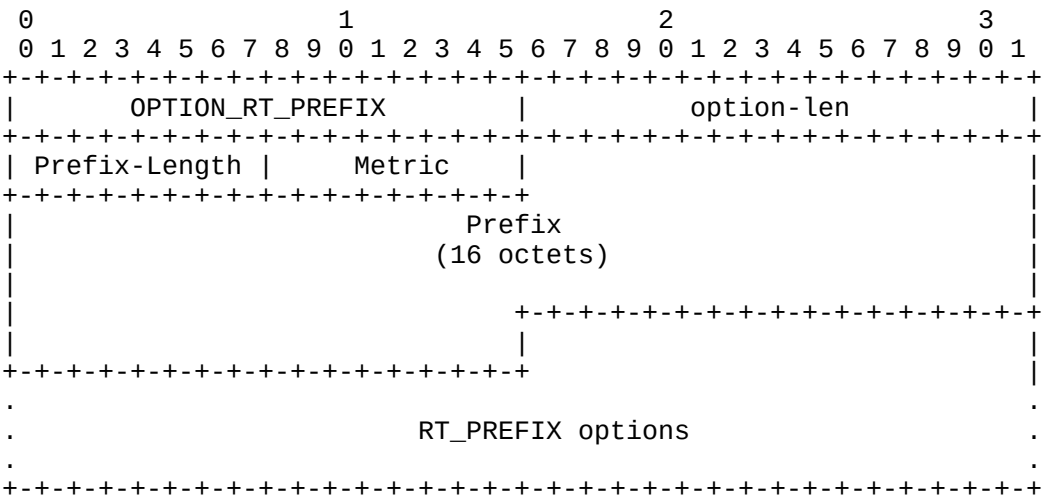

Figure 3: Route Prefix Option Format

option-code: OPTION_RT_PREFIX (TBD).

option-len: 18 + length of RT_PREFIX options.

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.

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: Fixed length 16 octet field containing an IPv6 prefix.

RT_PREFIX options: Options specific to this particular prefix.

5. DHCPv6 Server Behavior

When configured to do so s DHCPv6 server shall provide the Routes Option 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.

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].

Discussion: How should server indicate that there are no specific routes for this particular client? The reasonable behavior is to return empty IA_RT option, possibly with Status Code indicating Success. Another approach could be to simply not return any IA_RT option.

6. DHCPv6 Client Behavior

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

When processing a received Route Option 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 MAY 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.

7. IANA Considerations

A DHCPv6 option number of TBD for the introduced Route Option. IANA is requested to allocate three DHCPv6 option codes referencing this document: OPTION_IA_RT, OPTION_NEXT_HOP and OPTION_RT_PREFIX.

8. 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.

9. Contributors and .Acknowledgements

This document would not have been possible without the significant support and contribution to its development provided by: Arifumi Matsumoto, Hui Deng, Richard Johnson, Zhen Cao.

The authors would like to thank Alfred Hines, Ralph Droms, Ted Lemon, Ole Troan, Dave Oran and Dave Ward for their comments and useful suggestions.

10. References

10.1. 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.

10.2. Informative References

[I-D.troan-multihoming-without-nat66] Troan, O., Miles, D., Matsushima, S., Okimoto, T., and D. Wing, "IPv6 Multihoming without Network Address Translation", draft-troan-multihoming-without-nat66-01 (work in progress), July 2010.

[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.

Authors' Addresses

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

Email: wdec@cisco.com

Tomasz Mrugalski
Gdansk University of Technology
Storczykowa 22B/12
Gdansk 80-177
Poland

Phone: +48 698 088 272
Email: tomasz.mrugalski@eti.pg.gda.pl

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: