IPv6 maintenance Working Group (6man)
Internet-Draft
Updates: 4861 (if approved)
Intended status: Standards Track
Expires: September 9, 2015

F. Gont
SI6 Networks / UTN-FRH
R. Bonica
Juniper Networks
W. Liu
Huawei Technologies

March 8, 2015

Validation of IPv6 Neighbor Discovery Options
draft-ietf-6man-nd-opt-validation-00

Abstract

This memo specifies validation rules for IPv6 Neighbor Discovery (ND) Options. In order to avoid pathological outcomes, IPv6 implementations validate incoming ND options using these rules.

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 September 9, 2015.

Copyright Notice

Table of Contents

1. Introduction

IPv6 [RFC2460] nodes use Neighbor Discovery (ND) [RFC4861] to discover their neighbors and to learn their neighbors' link-layer addresses. IPv6 hosts also use ND to find neighboring routers that can forward packets on their behalf. Finally, IPv6 nodes use ND to verify neighbor reachability, and to detect link-layer address changes.

ND defines the following ICMPv6 [RFC4443] messages:

- Router Solicitation (RS)
- Router Advertisement (RA)
- Neighbor Solicitation (NS)
- Neighbor Advertisement (NA)
- Redirect

ND messages can include options that convey additional information. Currently, the following ND options are specified:

- Source link-layer address(SLLA) [RFC4861]
- Target link-layer address (TLLA) [RFC4861]
- Prefix information [RFC4861]
- Redirected header [RFC4861]
- MTU [RFC4861]
- Route Information [RFC4191]
- Recursive DNS Server (RDNSS) [RFC6106]
- DNS Search List (DNSSL) [RFC6106]

This memo specifies validation rules for the ND options mentioned above. In order to avoid pathological outcomes (such as [FreeBSD-rtsold]), IPv6 implementations validate incoming ND options using these rules.

2. Terminology

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].

3. Methodology

Section 4 through Section 11 of this document define validation rules for ND options. These sections also specify actions that are to be taken when an implementation encounters an invalid option. Possible actions are:

- The entire option MUST be ignored. However, the rest of the ND message MAY be processed.
- The entire ND message MUST be ignored

In the spirit of "being liberal in what you receive", the first action is always preferred. However, when an option length attribute is invalid, it is not possible to parse the rest of the ND message, and therefore subsequent ND options should be ignored.

We note that an implementation SHOULD NOT assume a particular length of an option (based on the option type) when it moves to the next option (whether it handles or ignores the current option) and SHOULD always use the length field of the option.

4. The Source Link-Layer Address (SLLA) Option

The SLLA Option is employed with NS, RS, and RA messages. If any other ND message contains an SLLA Option, the SLLA Option MUST be ignored. However, the rest of the ND message MAY be processed. (As per [RFC4861]).

Figure 1 illustrates the SLLA Option:

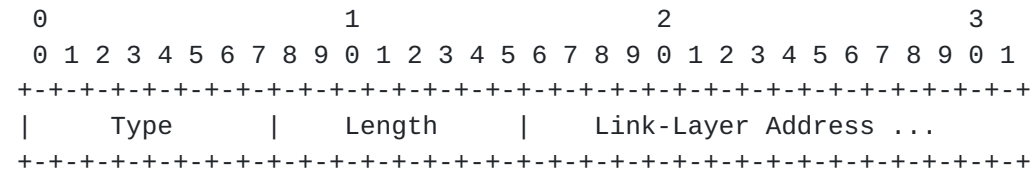

Figure 1: Source Link-Layer Address Option

The Type field is set to 1.

The Length field specifies the length of the option (including the Type and Length fields) in units of 8 octets. The Length field MUST be valid for the underlying link layer. For example, for IEEE 802 addresses the Length field MUST be 1 [RFC2464]. If an incoming ND message does not pass this validation check, the entire ND message MUST be discarded.

The Link-Layer Address field specifies the link-layer address of the packet's originator. It MUST NOT be any of the following:

- a broadcast address (see Appendix B for rationale)
- a multicast address (see Appendix B for rationale)
- an address belonging to the receiving node (see Appendix A for rationale)

If an incoming ND message does not pass this validation check, the SLLA Option MUST be ignored. However, the rest of the ND message MAY be processed.

An ND message that carries the SLLA Option MUST have a source address other than the unspecified address (0:0:0:0:0:0:0:0). If an incoming ND message does not pass this validation check, the SLLA Option MUST

be ignored. However, the rest of the ND message MAY be processed. (As per [RFC4861]).

5. The Target Link-Layer Address (TLLA) Option

NA and Redirect messages MAY contain a TLLA Option. If any other ND message contains an TLLA Option, the TLLA Option MUST be ignored. However, the rest of the ND message MAY be processed. (As per [RFC4861]).

Figure 2 illustrates the Target link-layer address:

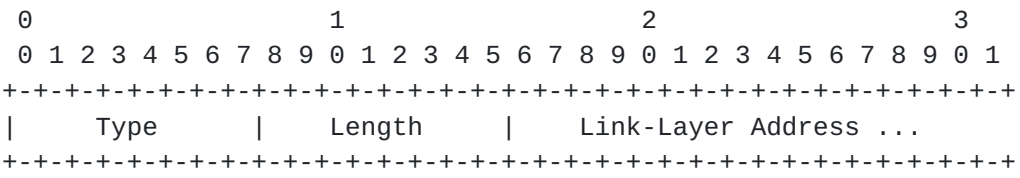

Figure 2: Target link-layer address option format

The Type field is set to 2.

The Length field specifies the length of the option (including the Type and Length fields) in units of 8 octets. The Length field MUST be valid for the underlying link layer. For example, for IEEE 802 addresses the Length field MUST be 1 [RFC2464]. If an incoming ND message does not pass this validation check, the entire ND message MUST be discarded.

An ND message that carries the TLLA option also includes a Target Address. The TLLA Option Link-Layer Address maps to the Target Address. The TLLA Option Link-Layer Address MUST NOT be any of the following:

- a broadcast address (see Appendix B for rationale)
- a multicast address (see Appendix B for rationale)
- an address belonging to the receiving node (see Appendix A for rationale)

If an incoming ND message does not pass this validation check, the TLLA Option MUST be ignored. However, the rest of the ND message MAY be processed.

6. The Prefix Information Option

The RA message MAY contain a Prefix Information Option. If any other ND message contains a Prefix Information Option, the Prefix Information Option MUST be ignored. However, the rest of the ND message MAY be processed. (As per [RFC4861]).

Figure 3 illustrates the Prefix Information Option:

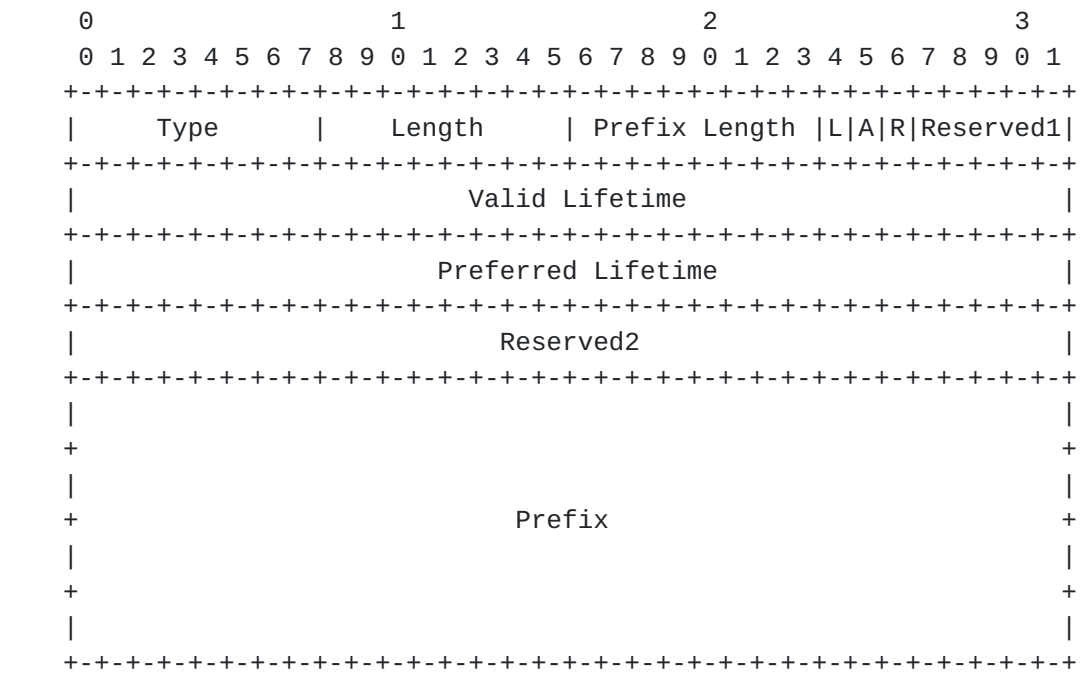

Figure 3: Prefix Information option format

The Type field is set to 3.

The Length field MUST be set to 4. If an incoming ND message does not pass this validation check, the entire ND message MUST be discarded.

As stated in [RFC4861] the Preferred Lifetime MUST be less than or equal to the Valid Lifetime. If an incoming ND message does not pass this validation check, the Prefix Information Option MUST be ignored. However, the rest of the ND message MAY be processed.

The Prefix Length contains the number of leading bits in the prefix that are to be considered valid. It MUST be greater than or equal to 0, and smaller than or equal to 128. If the field does not pass this check, the Prefix Information Option MUST be ignored. However, the rest of the ND message MAY be processed.

The Prefix field MUST NOT contain a link-local or multicast prefix. If an incoming ND message does not pass this validation check, the Prefix Information Option MUST be ignored. However, the rest of the ND message MAY be processed.

7. The Redirected Header Option

The Redirect message MAY contain a Redirect Header Option. If any other ND message contains an Redirect Header Option, the Redirect Header Option MUST be ignored. However, the rest of the ND message MAY be processed. (As per [RFC4861]).

Figure 4 illustrates the Redirected Header option:

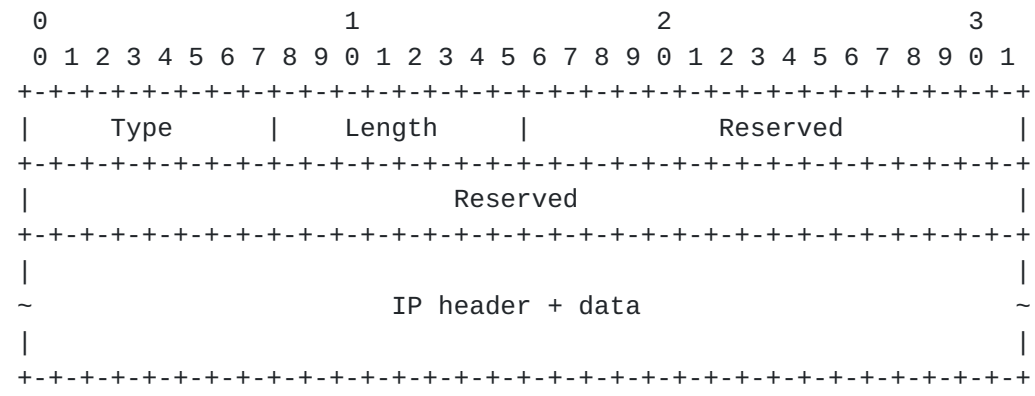

Figure 4: Redirected Header Option format

The Type field is 4.

The Length field specifies the option size (including the Type and Length fields) in units of 8 octets. Its value MUST be greater than or equal to 6. If an incoming ND message does not pass this validation check, the entire ND message MUST be discarded.

The value 6 was chosen to accommodate mandatory fields (8 octets) plus the base IPv6 header (40 octets).

8. The MTU Option

The RA message MAY contain an MTU Option. If any other ND message contains an MTU Option, the MTU Option MUST be ignored. However, the rest of the ND message MAY be processed. (As per [RFC4861]).

Figure 5 illustrates the MTU option:

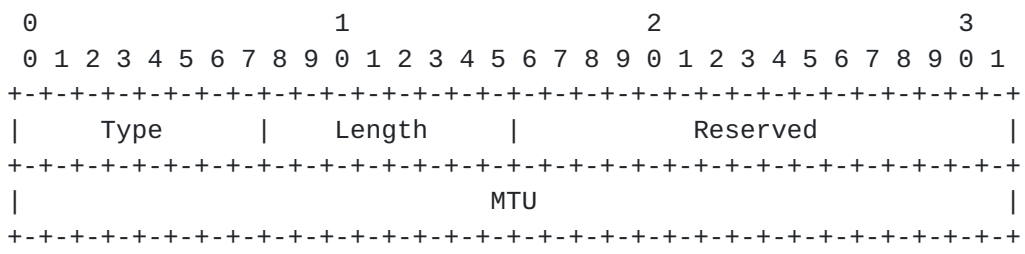

Figure 5: MTU Option Format

The Type field identifies the kind of option and is set to 5.

The Length field MUST BE set to 1 by the sender. If an incoming ND message does not pass this validation check, the entire ND message MUST be discarded.

The MTU field is a 32-bit unsigned integer that specifies the MTU value that should be used for this link. [RFC2460] specifies that the minimum IPv6 MTU is 1280 octets. Therefore, the MTU MUST be greater than or equal to 1280. If an incoming ND message does not pass this validation check, the MTU Option MUST be ignored. However, the rest of the ND message MAY be processed.

Additionally, the advertised MTU MUST NOT exceed the maximum MTU specified for the link-type (e.g., [RFC2464] for Ethernet networks). If an incoming ND message does not pass this validation check, the MTU Option MUST be ignored. However, the rest of the ND message MAY be processed.

9. The Route Information Option

The RA message MAY contain a Route Information Option. If any other ND message contains a Route Information Option, the Route Information Option MUST be ignored. However, the rest of the ND message MAY be processed.

Figure 6 illustrates Route Information option:

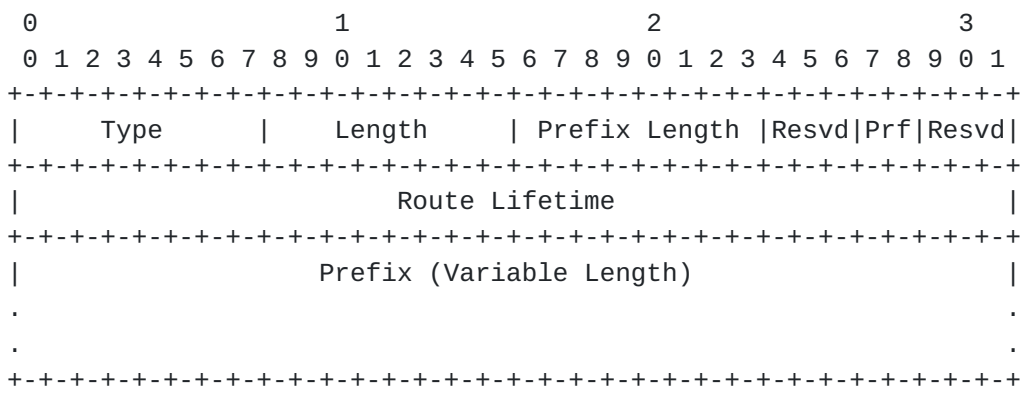

Figure 6: Route Information Option Format

The Type field is 24.

The Length field contains the length of the option (including the Type and Length fields) in units of 8 octets. Its value MUST be at least 1 and at most 3. If an incoming ND message does not pass this validation check, the entire ND message MUST be discarded.

The Prefix Length field indicates the number of significant bits in the Prefix field that are significant. Its value MUST be less than or equal to 128. If the field does not pass this check, the Route Information Option MUST be ignored.

The Length field and the Prefix Length field are closely related, as the Length field constrains the possible values of the Prefix Length field. If the Prefix Length is equal to 0, the Length MUST be equal to 1. If the Prefix Length is greater than 0 and less than 65, the Length MUST be equal to 2. If the Prefix Length is greater than 65 and less than 129, the Length MUST be equal to 3. If an incoming ND message does not pass this validation check, the entire ND message MUST be discarded.

The Prefix field MUST NOT contain a link-local unicast prefix (fe80::/10) or a link-local multicast prefix (e.g., ff02::/64). If an incoming ND message does not pass this validation check, the Route Information Option MUST be ignored. However, the rest of the ND message MAY be processed.

10. The Recursive DNS Server (RDNSS) Option

The RA message MAY contain a Recursive DNS Server (RDNSS) Option. If any other ND message contains an RDNSS Option, the RDNSS Option MUST be ignored. However, the rest of the ND message MAY be processed.

Figure 7 illustrates the syntax of the RDNSS option:

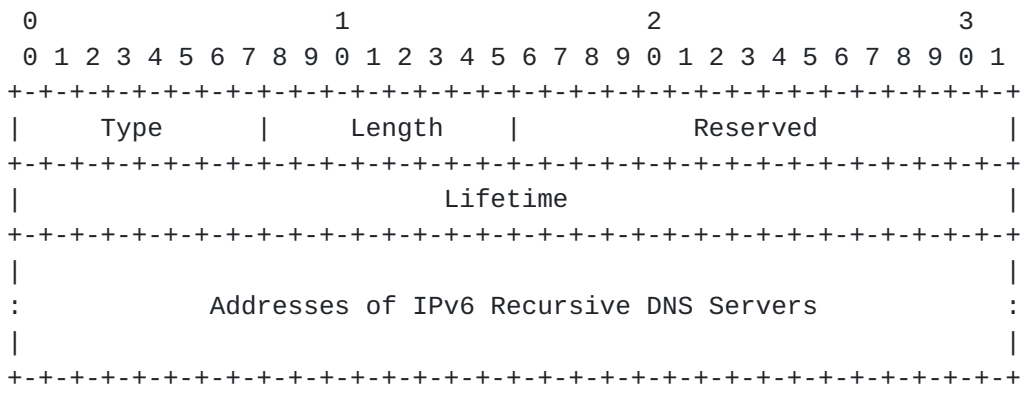

Figure 7: Recursive DNS Server Option Format

The Type field is 25.

The Length field specifies the length of the option (including the Type and Length fields) in units of 8 octets. Its value MUST be greater than or equal to 3. Additionally the Length field MUST pass the following check:

(Length -1) % 2 == 0

Figure 8

If the option does not pass these validation checks, the entire ND message MUST be discarded.

The RDNSS address list MUST NOT contain multicast addresses or the unspecified address. If an incoming ND message does not pass this validation check, the RDNSS Option MUST be ignored. However, the rest of the ND message MAY be processed.

11. The DNS Search List (DNSSL) Option

The RA message MAY contain a DNS Search List (DNSSL) Option. If any other ND message contains a DNSSL Option, the DNSSL Option MUST be ignored. However, the rest of the ND message MAY be processed.

Figure 9 illustrates the syntax of the DNSSL option:

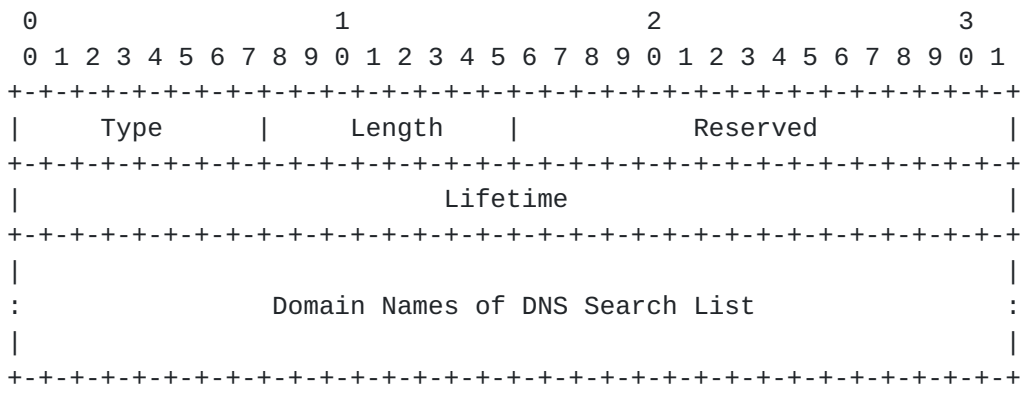

Figure 9: DNS Search List Option Format

The Type field is 31.

The Length field specifies the length of the option (including the Type and Length fields) in units of 8 octets. Its value MUST be greater than or equal to 2. If an incoming ND message does not pass these validation checks, the entire ND message MUST be discarded.

[RFC6106] specifies the valid format of domain suffixes. If a suffix is not validly encoded as specified, the corresponding DNSSL option MUST be ignored.

12. IANA Considerations

There are no IANA registries within this document. The RFC-Editor can remove this section before publication of this document as an RFC.

13. Security Considerations

This document specifies sanity checks to be performed on Neighbor Discovery options. By enforcing the checks specified in this document, a number of pathological behaviors (including some leading to Denial of Service scenarios) are eliminated.

14. Acknowledgements

Thanks to Tomoyuki Sahara and Jinmei Tatuya for their careful review and comments.

15. References

15.1. Normative References

[RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, November 1987.

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

[RFC2464] Crawford, M., "Transmission of IPv6 Packets over Ethernet Networks", RFC 2464, December 1998.

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

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

[RFC4443] Conta, A., Deering, S., and M. Gupta, "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", RFC 4443, March 2006.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[RFC6106] Jeong, J., Park, S., Beloeil, L., and S. Madanapalli, "IPv6 Router Advertisement Options for DNS Configuration", RFC 6106, November 2010.

15.2. Informative References

[FreeBSD-rtsold] FreeBSD, , "rtsold(8) remote buffer overflow vulnerability", 2014, <https://www.freebsd.org/security/advisories/FreeBSD-SA-14:20.rtsold.asc>.

Appendix A. Mapping an IPv6 Address to a Local Router's Own Link-layer Address

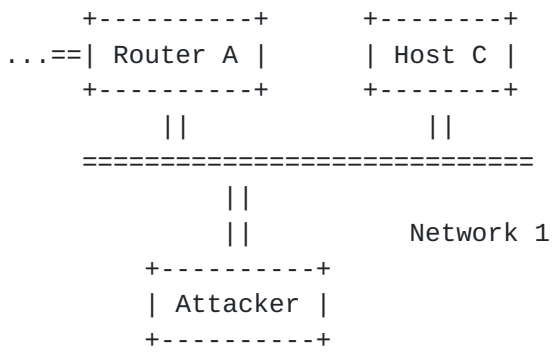

Figure 10: Unicast Forwarding Loop

In Figure 10, an on-link attacker sends Router A a crafted ND message that maps Host C's IPv6 address to the link-layer address of Router A's interface to Network 1. The crafted ND message causes Router A to map Host C's IPv6 address to the link layer address of its own interface to Network 1 and sets up the scenario for a subsequent attack.

A packet is sent to Router A with the IPv6 Destination Address of Host C. Router A forwards the packet on Network 1, specifying its own Network 1 interface as the link-layer destination. Because Router A specified itself as the link layer destination, Router A receives the packet and forwards it again. This process repeats until the IPv6 Hop Limit is decremented to 0 (and hence the packet is discarded). In this scenario, the amplification factor is equal to the Hop Limit minus one.

An attacker can realize this attack by sending either of the following:

- An ND message whose SLLA maps an IPv6 address to the link layer address of the victim router's (Router A's in our case) interface to the local network (Network 1 in our case)

- An ND message whose TLLA maps an IPv6 address to the link layer address of the victim router's (Router A's in our case) interface to the local network (Network 1 in our case)

Appendix B. Mapping a Unicast IPv6 Address to A Broadcast Link-Layer Address

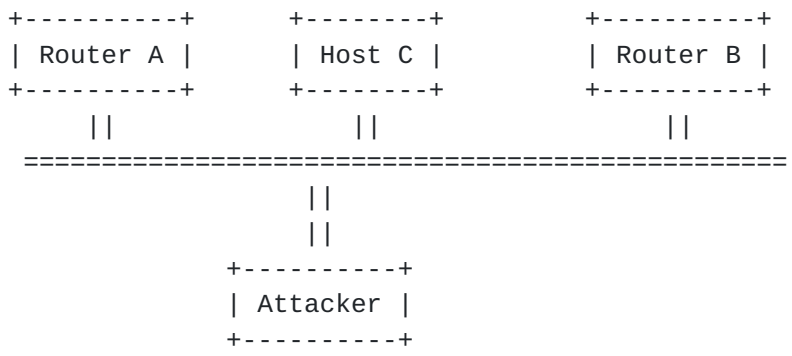

Figure 11: Broadcast Forwarding Loop

In Figure 11, the Attacker sends one crafted ND message to Router A, and one crafted ND message to Router B. Each crafted ND message contains the Target Address set to Host C's IPv6 address, and a TLLA option set to the Ethernet broadcast address (ff:ff:ff:ff:ff:ff). These ND messages causes each router to map Host C's IPv6 address to the Ethernet broadcast address. This sets up the scenario for a subsequent attack.

The Attacker sends a packet to the Ethernet broadcast address (ff:ff:ff:ff:ff:ff), with an IPv6 Destination Address equal to the IPv6 address of Host C. Upon receipt, both Router A and Router C decrement the Hop Limit of the packet, and resend it to the Ethernet broadcast address. As a result, both Router A and Router B receive two copies of the same packet (one sent by Router A, and another sent by Router B). This would result in a "chain reaction" that would only disappear once the Hop Limit of each of the packets is decremented to 0. The equation in Figure 12 describes the amplification factor for this scenario :

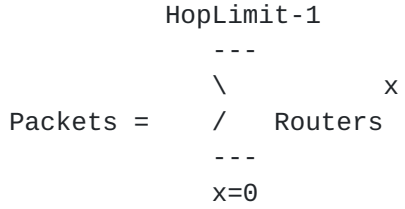

Figure 12: Maximum amplification factor

This equation does not take into account ICMPv6 Redirect messages that each of the Routers could send, nor the possible ICMPv6 "time exceeded in transit" error messages that each of the routers could send to the Source Address of the packet when each of the "copies" of the original packet is discarded as a result of their Hop Limit being decremented to 0.

An attacker can realize this attack by sending either of the following:

- An ND message whose SLLA maps an IPv6 address not belonging to the victim routers to the broadcast link-layer address

- An ND message whose TLLA maps an IPv6 address not belonging to the victim routers to the broadcast link-layer address

An additional mitigation would be for routers to not forward IPv6 packets on the same interface if the link-layer destination address of the received packet was a broadcast or multicast address.

Authors' Addresses

Fernando Gont
SI6 Networks / UTN-FRH
Evaristo Carriego 2644
Haedo, Provincia de Buenos Aires 1706
Argentina

Phone: +54 11 4650 8472
Email: fgont@si6networks.com
URI: http://www.si6networks.com

Ronald P. Bonica
Juniper Networks
2251 Corporate Park Drive
Herndon, VA 20171
US

Phone: 571 250 5819
Email: rbonica@juniper.net

Will (Shucheng) Liu
Huawei Technologies
Bantian, Longgang District
Shenzhen 518129
P.R. China

Email: liushucheng@huawei.com