Internet Engineering Task Force
Internet-Draft
Updates: 4861 (if approved)
Intended status: Standards Track
Expires: April 19, 2013

M. Smith
IMOT
O. Ddin
ZeroLag Communications, Inc.

October 16, 2012

Mitigating IPv6 Router Neighbor Cache DoS Using Stateless Neighbor Discovery
draft-smith-6man-mitigate-nd-cache-dos-slnd-02

Abstract

The IPv6 neighbor discovery cache is vulnerable to a Denial of Service attack that purposely exhausts the state used during the neighbor discovery address resolution process. This attack can be very disruptive when the target is a router. This memo proposes a stateless form of neighbor discovery to be used by routers to mitigate this attack. It does not require any changes to hosts.

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 16, 2013.

Table of Contents

1. Introduction

The IPv6 neighbor discovery cache [RFC4861] is vulnerable to a Denial of Service attack that purposely exhausts the state used during the neighbor discovery address resolution process [RFC3756].

When a router is the target of this attack, an off-link attacker sends traffic towards many non-existent addresses within a prefix attached to the router. This causes the router to create neighbor cache state for neighbor solicitations for these non-existent addresses. The denial of service occurs when the router's neighbor cache state capacity is exhausted due to too many outstanding neighbor solicitations.

Sizing a prefix proportional to the number of attached hosts, rather than using the standard /64 prefix size [RFC4291], would mitigate this attack. However, operational conveniences and benefits such as universal fixed length prefixes and interface identifiers, Stateless Address Auto configuration (SLAAC) [RFC4862] and privacy addresses [RFC4941], and never having to resize the prefix or add secondary prefixes to attach more hosts to the link would be lost.

This memo proposes a stateless form of neighbor discovery to prevent this type of DoS attack on a router. It does not require any changes to the operation of neighbor discovery on hosts. It optionally takes advantage of hosts' ability to recover from packet loss in the network, necessary due to IPv6's best effort nature. This method can be used for unknown or untrusted packet sources, when the router's neighbor cache's state capacity reaches a medium to high threshold of use, suggesting a neighbor cache DoS attack is occurring. Trusted packet sources would continue to be provided with traditional stateful neighbor discovery.

1.1. 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].

2. Terminology

Stateful Neighbor Discovery (SFND): Traditional neighbor discovery, as specified in [RFC4861]. This form of neighbor discovery maintains per packet destination state for all unresolved destinations during the neighbor discovery process. The neighbor cache's state capacity is intentionally exhausted to cause the neighbor cache Denial of Service attack.

Stateless Neighbor Discovery (SLND): The form of neighbor discovery described in this memo. This form of neighbor discovery does not maintain state for unresolved destinations during the neighbor discovery process. SLND has two modes of operation: hashed and unhashed. Hashing refers to any computational function which digests input into a shorter, unique value.

3. Stateless Neighbor Discovery

3.1. SLND Variables

To perform stateless neighbor discovery, four variables are maintained:

SLND Flag - This flag indicates whether or not the interface will perform SLND if necessary.

SLDN Activate Threshold - This variable specifies the threshold when stateless neighbor discovery is activated. The threshold specifies a neighbor cache utilization level. It is expressed as a percentage, with a RECOMMENDED default value of 80%. It may be either a per-interface or router global variable depending on whether the router implementation has per-interface neighbor caches or a global neighbor cache used by all interfaces.

SLND Active Flag - This flag indicates whether or not the interface is performing SLND for untrusted packet sources. It is maintained for each interface on the router.

Trusted/Untrusted Discriminator (TUD) - This variable specifies an arbitrary discriminator that determines whether a datagram is considered trusted or untrusted for the purposes of neighbor discovery. a TUD implementation MUST provide for at least per-source and per-interface discrimination although this MAY be controlled externally such as a packet mark/tag being the basis of the TUD and separate processes of the router providing the infrastructure to set the value by way of a bitmask.

SLND Neighbor Solicitation Rate Limit ("SLND NS Rate Limit") - This variable specifies a threshold for multicast Neighbor Solicitations when the interface is performing SLND, specified in packets per second. This limit MUST be configurable on a per-interface basis but a router MAY provide for configuring a default or global rate for convenience. OPTIONAL granular rate limiting may be achieved by maintaining a table of sources, grouped on a large boundary (/48 or bigger) in order to achieve weighted rate limiting that has a bigger impact on an attacking subnet. If such granular rate limiting is implemented, a router MUST NOT exceed the total interface rate

threshold.

SLND Hash Salt (NDHS) - This is a value that SHOULD be at least 8 octets in length, known only to the router and MUST change periodically to mitigate replay attacks. This value SHOULD NOT change more often than the typical time taken for successful resolution and it is RECOMMENDED that this value change approximately as often as the configured ND timeout.

3.2. SLND Process

The stateless neighbor discovery process may occur once a router has determined the outgoing interface for a packet, and that the packet's destination is on-link.

If the packet's destination address is present in the neighbor cache, and the link-layer address has been resolved, the packet is forwarded to it's destination.

If the packet's destination address is not present in the neighbor cache, and the SLND Flag is off, traditional stateful neighbor discovery is performed for the packet's destination.

If the packet's destination address is not present in the neighbor cache, and the SLND Flag is on, the TUD is referred to in order to ascertain whether the packet is trusted or not.

If the packet is determined to be trusted, traditional stateful neighbor discovery is performed.

If the packet is determined to be untrusted, stateless neighbor discovery is performed.

A router MUST implement unhashed SLND and MAY also implement hashed SLND. Where both SLND modes are implemented, a router MUST allow mode selection on a per-interface basis. Defaulting to hashed SLND is RECOMMENDED.

3.3. Unhashed SLND

The unhashed stateless neighbor discovery process is as follows:

1. The router determines if sending a multicast neighbor solicitation would exceed the SLND NS Rate Limit for the outgoing interface. If the SLND NS Rate Limit would be exceeded, drop the packet and do not proceed any further.

2. A multicast neighbor solicitation is sent by the router for the destination address in the packet. The packet MAY then be dropped. If the packet is instead requeued any subsequent dequeue of the same packet MUST NOT result in the transmission of another solicitation whilst in SLND mode. the packet SHOULD NOT be queued indefinitely. if the packet expires, the router MAY send an ICMPv6 error to the source but given the likely scenario for the activation of SLND, this is not recommended.

3. As some later point in time, the router may receive a unicast neighbor advertisement, for a previously sent neighbor solicitation.

4. If the SLND Active Flag is off and there is no stateful entry regarding the advertisement, the router ignores the neighbor advertisement.

5. If the SLND Active Flag is on, the router creates an entry in it's neighbor cache using the information received in the unicast neighbor advertisement. Stateless neighbor discovery is now complete.

3.4. Hashed SLND

The hashed stateless neighbor discovery process is as follows:

1. The router determines if sending a multicast neighbor solicitation would exceed the SLND NS Rate Limit for the outgoing interface. If the SLND NS Rate Limit would be exceeded, drop the packet and do not proceed any further.

2. The router performs hash function H(x) where x is the concatenation of the IPv6 address to solicit and the NDHS, known only to the router. the hash function SHOULD result in a value of at least 8 octets in length and subsequently be used to replace the lower 8 octets (host portion) of the IPv6 address that would have been the source of the solicitation. When applicable to the link medium, a router MUST include the ICMPv6 option indicating the data-link layer source as part of the solicitation.

3. As some later point in time, the router may receive a unicast neighbor advertisement, for a previously sent neighbor solicitation.

4. If the SLND Active Flag is off or the neighbor already exists in the cache, the router ignores the neighbor advertisement.

5. If the SLND Active Flag is on, the router performs hash function H(y) where y is the concatenation of the advertised IPv6 address and the NDHS. The result of H(x) is then compared to the lower 8 octets (host portion) of the IPv6 destination address in the advertisement. If the result does not match, the advertisement is discarded; otherwise the information in the advertisement is used to update the neighbor cache.

The utilization of the neighbor cache has to be measured to determine if it crosses the SLND Activate Threshold. If the utilization increases above the SLND Activate Threshold, the SLND Active Flag is set, and if it decreases below the SLND Activate Threshold, the SLND Active Flag is unset. Neighbor cache utilization should be measured and compared to the SLND Activate Threshold when:

- entries are added to the neighbor cache, during either stateful or stateless neighbor discovery.

- entries are removed from the neighbor cache when NUD discovers the neighbor has become unreachable or timed out.

4. Consequences of Stateless Neighbor Discovery

During traditional stateful neighbor discovery, state is used to perform the following:

- ensure a received neighbor advertisement corresponds to a previously sent neighbor solicitation

- to retransmit a limited number of neighbor solicitations if previous solicitations remain unanswered

- to store a small number of packets that triggered the neighbor discovery process, so that they can be transmitted if neighbor discovery completes successfully

- to generate an ICMPv6 destination unreachable, address unreachable messages back to the packet source, should the neighbor discovery process fail

Unhashed stateless neighbor discovery sacrifices these functions and the related state to mitigate the neighbor cache DoS attack.

Hashed stateless neighbor discovery retains much of the robustness of stateful ND at a cost of computation time for hash calculation and comparison and the minor risk of replay attacks although this is largely mitigated by appropriate tuning of the frequency by which the NDHS changes. However, the calculated source address may trigger a solicited host to attempt neighbor discovery of that address, thus creating the cost of some spurious ND traffic for the benefit of preventing on-link cache attacks.

4.1. Neighbor Advertisement Validation

Ensuring received neighbor advertisements correspond to previously sent neighbor solicitations prevents on-link nodes from sending unsolicited neighbor advertisements to the router, and then having them added to the router's neighbor cache without validation. This would allow on-link hosts to perform a neighbor cache DoS attack, as they could send many neighbor advertisements for non-existent addresses within the link assigned prefixes, exhausting the neighbor cache capacity.

If neighbor advertisement validation occurs, then the router is vulnerable to an off-link sourced neighbor cache DoS attack, but is not vulnerable to an on-link sourced neighbor cache DoS attack. If neighbor advertisement validation does not occur, then the router is vulnerable to an on-link sourced neighbor cache DoS attack, but is now not vulnerable to an off-link sourced neighbor cache DoS attack.

For best-case performance, Hashed SLND should be performed to provide cache protection against both on-link and off-link attacks.

4.2. Optimization Functions

The nature of IPv6 is best effort, meaning that there is a possibility that packets may be lost as they transit the network, and that IPv6 will not make any attempt to recover lost packets. If an application residing on an IPv6 node requires reliable packet delivery, it will need to utilize locally implemented reliable upper layer protocols such as TCP and SCTP, or implement it's own reliability mechanisms. These reliability mechanisms involve retransmitting packets. Alternatively, the application needs to accept the possibility of packet loss.

The remaining uses of stateful neighbor discovery state are not

assured of success. The limited number of neighbor solicitation retransmissions may not be enough, causing neighbor discovery to fail even though the target node exists. There may be more packets sent that trigger neighbor discovery than are stored for transmission when neighbor discovery completes successfully, causing them to be dropped. The ICMPv6 destination unreachable message may be dropped on the way back to the traffic originating node, perhaps intentionally by a network located firewall.

This means that these functions are useful but not essential optimizations. If necessary, they do not need to be performed, as the packet source will retransmit its packets, reinitiating the neighbor discovery process, or accept that packet loss has occurred. This provides the opportunity to perform a stateless form of neighbor discovery if there is evidence that a neighbor cache DoS attack is occurring, mitigating the off-link sourced neighbor cache DoS attack.

5. Trusted/Untrusted Discriminator

As previously described, the Trusted/Untrusted Discriminator (TUD) is used to determine whether a packet is trusted or untrusted, with trusted datagrams continuing to trigger traditional stateful neighbor discovery services, and untrusted ones receiving stateless neighbor discovery services.

For routers where it may not be operationally convenient or possible to implement comprehensive trusted and untrusted datagram selection, such as on low-end or embedded devices, it would be acceptable to consider all datagram sources untrusted when stateless neighbor discovery is active.

For routers that can support more comprehensive trusted and untrusted datagram discrimination, the following information sources are suggested for the purposes of configuring a default set of discriminators.

5.1. Configured Trusted and Untrusted Prefixes

The first suggested TUD source is an operator configured list of prefixes and their lengths, each with a flag indicating whether traffic with source addresses that falls within the specified prefix is from a trusted or untrusted source, using the longest-match logic that is applied to IPv6 routing decisions.

This list should have a default entry of the ULA prefix (fc00::/7) [RFC4193], flagged as a trusted source. An implementation must allow this entry to be removed. An implementation SHOULD NOT have this default entry if reverse path filtering is not possible or where spoofing is considered trivial.

5.2. Routing Information

The second suggested TUD source is the network's routing information.

The network's routing information can be used to distinguish trusted and untrusted datagram sources. An advantage of using routing information for this purpose is that it will typically be dynamically and automatically distributed to all routers within the network, when dynamic routing protocols are used. This avoids individual routers in the network having to be manually reconfigured with trusted prefixes when subnets are added or removed from the network.

The contents of a stub network's route table is typically all the internal routes for the network, and then a default route used to reach the Internet. The list of internal routes can be used to distinguish between trusted and untrusted sources, with datagram sources matching internal routes being trusted, and all other datagram sources being untrusted.

In more complex routing environments, such as those using one or more IGPs and an EGP such as BGP, there may be other methods available to distinguish between trusted and untrusted sources. For example, routes carried in an IGP could be considered trusted, while routes carried in BGP are untrusted. For a network using BGP to carry all reachability information, except network transit and loopback interface routes, routes may be tagged with one or more BGP communities which indicate internal and therefore trusted prefixes.

A default route SHOULD NOT be used to define a trusted datagram source prefix.

5.3. Default to Untrusted

Finally, should none of the previous trusted or untrusted source prefix information sources match the source address of traffic that would trigger neighbor discovery, the datagram source should be considered untrusted.

6. Acknowledgements

Review and comments were provided by Ray Hunter and Matthew Moyle-Croft.

This memo was prepared using the xml2rfc tool and nroff I-D editor.

7. Security Considerations

This memo proposes a security mitigation for an off-link sourced neighbor cache Denial of Service attack, aimed at a router.

As discussed in Section 4.1, the unhashed method proposed creates an opportunity for an on-link sourced neighbor cache DoS attack, when mitigating the off-link sourced neighbor cache DoS. This is considered to be an acceptable security trade-off.

The hashed method mitigates the above issue at the cost of creating some additional unnecessary ND traffic and potentially having a window of opportunity for a replay attack. The benefit of Hashed SLND is considered to outweigh these concerns.

8. Change Log [RFC Editor please remove]

draft-smith-6man-mitigate-nd-cache-dos-slnd-00, initial version, 2012-09-04

draft-smith-6man-mitigate-nd-cache-dos-slnd-01, more clarity, 2012-10-13

- more comprehensive introduction (problem definition) text
- make it more obvious that hosts don't need to be changed
- low-end/embedded hosts can consider all datagram sources untrusted
- misc. minor text updates

draft-smith-6man-mitigate-nd-cache-dos-slnd-02, major changes, 2012-

10-16

- Add definition of Hashed SLND.
- Expand optional scope for flexibility of SLND
- Address potential security risks

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.

9.2. Informative References

[RFC3756] Nikander, P., Kempf, J., and E. Nordmark, "IPv6 Neighbor Discovery (ND) Trust Models and Threats", RFC 3756, May 2004.

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

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, February 2006.

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

[RFC4862] Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless Address Autoconfiguration", RFC 4862, September 2007.

[RFC4941] Narten, T., Draves, R., and S. Krishnan, "Privacy Extensions for Stateless Address Autoconfiguration in IPv6", RFC 4941, September 2007.

Author's Addresses

Mark Smith
In My Own Time
PO BOX 521
HEIDELBERG, VIC 3084
AU

Email: markzzzsmith@yahoo.com.au

Oliver Ddin
ZeroLag Communications
289 S. Robertson Blvd #441
Beverly Hills, CA 90211
USA

Email: oliver@zerolag.com