Internet Engineering Task Force
Internet-Draft
Updates: 4861 (if approved)
Intended status: Standards Track
Expires: August 24, 2013

M. Smith
IMOT

February 20, 2013

Mitigating IPv6 Neighbor Discovery DoS Attack Using Stateless Neighbor Presence Discovery
draft-smith-6man-mitigate-nd-cache-dos-slnd-06

Abstract

One of the functions of IPv6 Neighbor Discovery is to discover whether a specified neighbor is present. During the neighbor presence discovery process state is created. A node's capacity for this state can be intentionally exhausted to perform a denial of service attack, known as the "Neighbor Discovery DoS Attack". This memo proposes a stateless form of neighbor presence discovery to prevent this Neighbor Discovery DoS Attack.

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 24, 2013.

Table of Contents

1. Introduction

One of the functions of IPv6 Neighbor Discovery [RFC4861] is to discover whether a specified neighbor is present. Neighbor presence discovery occurs when a packet needs to be sent to a specified neighbor for which presence hasn't previously been determined.

During neighbor presence discovery, state is created to support the discovery process. The amount of state created is directly proportional to the number of neighbors being discovered at the time. The total possible state that can be created is limited to the lower of the node's state capacity or the size of the IPv6 address space for use by potential neighbors.

To provide operational convenience and simplicity, most IPv6 Interface Identifiers are 64 bits in length [RFC4291]. This results in a common IPv6 subnet prefix length of 64 bits, covering 2^64 addresses. This large IPv6 subnet address space provides an opportunity for an attacker to exhaust a node's capacity for state created during neighbor presence discovery. The consequences of this state exhaustion attack are likely to be a denial of service. New neighbor presence discovery transactions may fail, despite the neighbor existing, and knowledge of existing neighbors' presence may be discarded. This attack is known as the "Neighbor Discovery DoS Attack" [RFC3756].

This memo proposes a stateless form of neighbor presence discovery to prevent this Neighbor Discovery DoS Attack. It takes advantage of hosts' ability to recover from packet loss in the network, necessary because of IPv6's best effort nature. This method would be used when a node's neighbor presence discovery state capacity reaches a medium to high threshold of use, suggesting a Neighbor Discovery DoS Attack is occurring.

This method does not require any changes to neighbors, or changes to Neighbor Solicitation or Neighbor Advertisement messages. An optional enhancement for router implementations is to identify a set of packet sources as trusted not to conduct the DoS attack, and to continue to provide these trusted packet sources with traditional and stateful neighbor presence discovery service.

[RFC4861] calls the neighbor presence discovery function "Address Resolution". This name seems somewhat inaccurate, as it suggests that the discovery of the presence of neighbors is only necessary for links with link-layer addresses. Neighbor presence discovery is necessary on all types of links, as functions such as generating ICMPv6 Destination Unreachable, Address Unreachable messages, or Neighbor Unreachability Detection [RFC4861], cannot be performed if

the presence of a neighbor is assumed by implication of a prefix length [RFC5942], rather than observed or actively tested. Address resolution, for links that require it, occurs as part of the neighbor presence discovery process.

If approved, this memo updates [RFC4861].

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

Neighbor Presence Discovery (NPD): Discovery of the presence of a specified neighbor. For link-layers with addresses, address resolution is performed as part of the presence discovery process.

Stateful Neighbor Presence Discovery (SFNPD): Traditional neighbor presence discovery specified in [RFC4861]. This form of Neighbor Presence Discovery creates state for each potential neighbor for which presence is being discovered.

Stateless Neighbor Presence Discovery (SLNPD): The form of Neighbor Presence Discovery described in this memo. Per-potential neighbor state is not created during Neighbor Presence Discovery.

IGP: Interior Gateway Protocol, such as OSPF [RFC5340].

EGP: Exterior Gateway Protocol, such as BGP [RFC4271].

Node: A device that implements Neighbor Presence Discovery. Both hosts and routers are nodes.

3. The Best Effort Nature of IPv6

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 from packet loss [RFC1958].

If an application requires reliable packet delivery, it will need to utilise locally implemented reliable transport layer protocols such as TCP and SCTP, or implement its own reliability mechanisms. These reliability mechanisms will usually involve packet loss detection and

retransmission. Alternatively, the application needs to accept the possibility and consequences of packet loss.

4. Opportunities for Stateless Neighbor Presence Discovery

During traditional and stateful NPD, 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 presence discovery process, so that they can be sent if the neighbor is present,
- to generate ICMP Destination Unreachable, Address Unreachable messages to the NPD trigger packet(s') origin host(s) should the specified neighbor not be present.

Stateless NPD sacrifices these functions and the related state when a Neighbor Discovery DoS Attack appears to be occurring.

4.1. Neighbor Advertisement Validation

The purpose of Neighbor Advertisement validation, during NPD, is to ensure that the receiver of the Neighbor Advertisement has previously been interested in the presence of the neighbor, expressed by sending a Neighbor Solicitation.

Stateless NPD abandons the state used to enforce a Neighbor Solicitation/Neighbor Advertisement transaction. It accepts Neighbor Advertisements without being able to ensure that they correspond to a previous Neighbor Advertisement. The received Neighbor Advertisement either updates existing neighbor presence information, or creates new neighbor presence information.

Updating nodes' existing neighbor presence information via unsolicited multicast Neighbor Advertisements is already permitted by [RFC4861]. While operating, Stateless NPD in effect allows unsolicited unicast Neighbor Advertisements, as knowledge of sending the previous Neighbor Solicitation is abandoned.

By making the NPD process stateless, hosts and routers would be protected against a Neighbor Discovery DoS Attack launched from a

host against itself, or launched from a host against a remote subnet on a router. However, while stateless NPD is operating, hosts and routers would now be vulnerable to a DoS attack from their own on-link neighbors, as the neighbors could send many unsolicited unicast Neighbor Advertisements for non-existent neighbors. These Neighbor Advertisements would be accepted without question, and false neighbor presence information would be created.

Considering that the set of on-link neighbors will be significantly limited compared to the set of possible off-link attackers (such as those on the wider Internet), may be better known due to geographic proximity or link-layer authorisation, and will have a vested interest in any on-link routers continuing to operate, sacrificing Neighbor Advertisement validation during NPD is a worthwhile compromise when a Neighbor Discovery DoS Attack appears to be occurring.

4.2. Optimisation Functions

The remaining uses of stateful NPD 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 NPD than are stored for transmission when NPD 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 optimisations, as they are not reliable. They can be sacrificed if necessary, as the original packet source will retransmit its packets, reinitiating NPD, or accept that packet loss has occurred. This retransmission or acceptance of packet loss provides the opportunity to perform a stateless form of Neighbor Presence Discovery, if there is evidence that a Neighbor Discovery DoS Attack is occurring.

5. Stateless Neighbor Presence Discovery

5.1. SLNPD Variables

To perform stateless NPD, five variables are maintained:

SLNPD Flag - This flag indicates whether or not the interface will perform SLNPD if necessary. By default, this flag should be set to on.

SLNPD Activate Threshold - This variable specifies the threshold when

stateless NPD is activated. The threshold specifies a neighbor cache utilisation level. It is expressed as a percentage, with a default value of 80%. It may either be a per-interface or node global variable depending on whether the neighbor discovery implementation has per-interface neighbor caches or a global neighbor cache used by all interfaces. In the case of per-interface neighbor caches, for convenience, an implementation may maintain a global SLNPD Activate Threshold variable, used when the per-interface SLNPD Activate Threshold value is set to 0.

SLNPD Active Flag - This flag indicates whether or not the interface is currently performing SLNPD. It is maintained for each interface on the node.

SLNPD Neighbor Advertisement Acceptance Time ("SLNPD NA Accept. Time") - This variable holds the time remaining during which apparent or actual unsolicited unicast Neighbor Advertisements will continue to be accepted, after SLNPD has become inactive. It is measured in milliseconds, and is used to implement a count-down-to-zero timer. It is maintained for each interface on the node.

SLNPD Neighbor Solicitation Rate Limit ("SLNPD NS Rate Limit") - This variable specifies a maximum threshold for multicast Neighbor Solicitations when the interface is performing SLNPD, specified in packets per second. It is a per-interface variable, as different interfaces may have different thresholds. The rate value should be an appropriate portion of the multicast packet per second capabilities of the interface link technology, to ensure multicast capacity remains for other uses. A packet per second rate corresponding to 10% of the link's multicast capability would be typical. For convenience, a node may maintain a global SLNPD NS Rate Limit that is used when an interface specific SLNPD NS Rate Limit is set to 0.

5.2. SLNPD Processing

The stateless NPD process may occur once a node 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 (if necessary for the link-layer type), the packet is forwarded out the link-layer interface to its destination.

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

If the SLNPD Flag is on, and the SLNPD Active flag is off, traditional stateful NPD is performed.

If the SLNPD Flag is on, and the SLNPD Active flag is on, stateless NPD is performed as follows:

1. The node determines if sending a multicast Neighbor Solicitation would exceed the SLNPD NS Rate Limit for the outgoing interface. If the SLNPD NS Rate Limit would be exceeded, discard the packet and do not proceed any further.

2. A multicast Neighbor Solicitation is sent by the node for the destination address in the packet. The packet is then discarded. (An implementation memory optimisation would be to record the packet destination address and then discard the packet before building and sending the corresponding Neighbor Solicitation).

3. As some later point in time, the node is likely to receive a unicast Neighbor Advertisement, for a previously sent Neighbor Solicitation.

4. If the SLNPD Active Flag is on, or the SLNPD Active Flag is off and the SLNPD NA Accept. Time is greater than zero, the node either:

5.

   * Updates an existing but incomplete neighbor cache entry, created as part of a previous stateful NPD transaction.

   * Creates a new entry in its neighbor cache using the information received in the unicast Neighbor Advertisement. Stateless NPD is now complete.

6. If the SLNPD Active Flag is off and the SLNPD NA Accept. Time is zero, the node performs traditional stateful NPD processing of the received Neighbor Advertisement.

The utilisation of the neighbor cache needs to be measured to determine if it crosses the SLNDP Activate Threshold. If the utilisation increases above the SLNDP Activate Threshold, the SLNPD Active Flag is switched on, and if it decreases below the SLNDP Activate Threshold, the SLNPD Active Flag is switched off. Neighbor cache utilisation should be measured and compared to the SLNDP Activate Threshold when:

- o entries are added to the neighbor cache, during either stateful or stateless NPD,

- o entries are removed from the neighbor cache when Neighbor Unreachability Detection discovers the neighbor has become unreachable.

When the SLNPD Active Flag is switched from on to off, the SLNPD NA Accept. Time is reset to the value of the node's RETRANS_TIMER value [RFC4861] multiplied by the node's MAX_MULTICAST_SOLICIT value [RFC4861]. A system timer is then started to decrement SLNPD NA Accept. Time down to zero. This timer provides the opportunity for outstanding SLNPD transactions to complete after SLNPD has become inactive. When the SLNPD Active Flag is switched from off to on, if the timer is operating it can be cancelled.

5.3. Optional Enhancements

5.3.1. Selective SLNPD

When a Neighbor Discovery DoS Attack appears to be occurring, it could be useful to continue to provide traditional stateful NPD service to hosts that are considered unlikely to initiate or participate in the DoS attack. These hosts could be considered trusted hosts, while the remaining set of hosts are untrusted.

The determination of whether a host is trusted or untrusted would take place when NPD is determined to be necessary, during the stateless NPD process. The determination of trust is made based on attributes of the packets that trigger the NPD process. If none of the packet attributes indicate either a trusted or untrusted host, or the value(s) of the packet attribute(s) cannot be trusted, then the source host is considered untrusted.

There are two basic packet attributes that an enhanced implementation should provide mechanisms to use to classify a packet source as trusted or untrusted:

- o source address

- o ingress interface

An implementation may be able to reuse its existing packet classification mechanisms to determine trust, such as those used to implement network QoS. This would mean that other packet attributes, such as Traffic Class, Flow Label [RFC6437], the CALIPSO option [RFC5570] or MPLS label values [RFC3031], could also be used to determine packet source trustworthiness.

5.3.1.1. Source Address

The source address of the packet that has triggered the NPD process can be used to determine the trust level of the origin host. The information used to classify the source address can come from two possible sources:

- an operator configured prefix list
- the network's routing information

5.3.1.1.1. Operator Configured Prefix List

An operator configured prefix list consists of a static 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.

How this list is evaluated would be implementation dependent, however it is likely to be either sequential from first to last entry, or using a longest match algorithm.

In most cases, 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. There may be some cases where even packets with ULA source addresses cannot be trusted; in these cases the prefix list should be empty by default. The likely deployment role for the implementation would be a factor in this decision.

5.3.1.1.2. Routing Information

The network's routing information can be used to distinguish trusted and untrusted packet 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 changes to the operator configured prefix list on individual routers when trusted prefixes 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 packet sources matching internal routes being trusted, and all other packet sources being untrusted.

In more complex routing environments, such as those using one or more

IGPs and an EGP such as BGP [RFC4271], 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, internal routes may be tagged with one or more BGP communities to indicate they are also trusted prefixes.

There may be cases where a subset of the internal routes need to be considered untrusted, despite them being propagated internally via a routing protocol. These routes will likely be for links at the edge of the local network, where untrusted hosts can be attached without local network control or authorisation. These routes need to be labelled as untrusted, and that information propagated to all routers within the local network. Route labelling mechanisms such as OSPF's External Route Tag [RFC5340] or a BGP community could be used for this purpose.

A default route sourced from a routing protocol should never be used as a trusted packet source route. If a router's operator wishes to trust all packet sources, they should specify the prefix that covers all IPv6 addresses, ::/0, as an operator configured trusted prefix. (The ::/0 prefix is only a default route when used as routing information.)

Implementations should provide simple and convenient methods to use the network's routing information to distinguish between trusted and untrusted packet source prefixes.

5.3.1.2. Ingress Interface

A packet's ingress interface on the router could be used to determine whether stateful or stateless NPD takes place. Interfaces on the router would be labelled as trusted or untrusted.

The default trust level for interfaces would be up to the router's implementer. Considerations could be the likely deployment scenario for the router implementation (e.g., residential Internet access, or within an enterprise network), and the type of interface (e.g., an interface type that is usually used to attach the router to the Internet, such as an ADSL interface, would be labelled untrusted). These default interface trust assignments should be easy to change.

6. Acknowledgements

Oliver Ddin provided review and comments, and suggested the use of ingress interface and other more general packet attributes to

determine packet source trust.

Ray Hunter provided review and comments, and initiated the discussion that resulted in this memo using the term and more specifically focusing on Neighbor Presence Discovery.

Igor Lubashev and Deb Banerjee provided review and comments, and identified a hysteresis issue around the SLNPD Activate Threshold.

Review and comments were provided by Matthew Moyle-Croft and Karl Auer.

This memo was prepared using the xml2rfc tool.

7. Security Considerations

This memo proposes a security mitigation method for an off-link sourced Neighbor Discovery DoS Attack.

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

Use of attributes that are carried within a packet to distinguish trusted and untrusted sources in Section 5.3.1 is based on the assumption that the values of these attributes can be trusted, meaning that they have been set by trusted packet sources. If it is possible that these packet attribute values may have been forged, then their possible source should be considered untrusted during the Stateless Neighbor Presence Discovery procedure, if the Selective Stateless NPD enhancement has been implemented.

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 packet sources untrusted
- misc. minor text updates

draft-smith-6man-mitigate-nd-cache-dos-slnd-05, structural changes, 2012-11-08

- moved opportunities for SLNPD section to before SLNPD description
- spit SLNPD into basic functionality and optional enhancements
- use of ingress interface and other more general packet attributes to determine trust

draft-smith-6man-mitigate-nd-cache-dos-slnd-06, better problem definition, 2012-02-20

- rephrase problem as one of neighbor presence discovery
- don't ignore Neighbor Advertisements that may be part of a previous stateful neighbor discovery transaction
- use a count down timer to allow outstanding SLNPD transactions to complete
- mention issues regarding trusting packet attributes

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

[RFC1958] Carpenter, B., "Architectural Principles of the Internet", RFC 1958, June 1996.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

[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.

[RFC4271] Rekhter, Y., Li, T., and S. Hares, "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, January 2006.

[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.

[RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, July 2008.

[RFC5570] StJohns, M., Atkinson, R., and G. Thomas, "Common Architecture Label IPv6 Security Option (CALIPSO)", RFC 5570, July 2009.

[RFC5942] Singh, H., Beebee, W., and E. Nordmark, "IPv6 Subnet Model: The Relationship between Links and Subnet Prefixes", RFC 5942, July 2010.

[RFC6437] Amante, S., Carpenter, B., Jiang, S., and J. Rajahalme, "IPv6 Flow Label Specification", RFC 6437, November 2011.

Author's Address

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

Email: markzzzsmith@yahoo.com.au