Internet Engineering Task Force (IETF)
Request for Comments: 6957
Category: Standards Track
ISSN: 2070-1721

F. Costa
J-M. Combes, Ed.
X. Pougnard
France Telecom Orange
H. Li
Huawei Technologies

June 2013

Duplicate Address Detection Proxy

Abstract

The document describes a proxy-based mechanism allowing the use of Duplicate Address Detection (DAD) by IPv6 nodes in a point-to-multipoint architecture with a "split-horizon" forwarding scheme, primarily deployed for Digital Subscriber Line (DSL) and Fiber access architectures. Based on the DAD signaling, the first-hop router stores in a Binding Table all known IPv6 addresses used on a point-to-multipoint domain (e.g., VLAN). When a node performs DAD for an address already used by another node, the first-hop router defends the address rather than the device using the address.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc6957.

Table of Contents

1. Introduction

This document specifies a function called Duplicate Address Detection (DAD) proxy allowing the use of DAD by the nodes on the same point-to-multipoint domain with a "split-horizon" forwarding scheme, primarily deployed for Digital Subscriber Line (DSL) and Fiber access architectures [TR-101]. It only impacts the first-hop router and it doesn't need modifications on the other IPv6 nodes. This mechanism is fully effective if all the nodes of a point-to-multipoint domain (except the DAD proxy itself) perform DAD.

This document explains also why the DAD mechanism [RFC4862] without a proxy cannot be used in a point-to-multipoint architecture with a "split-horizon" forwarding scheme (IPv6 over PPP [RFC5072] is not affected). One of the main reasons is that, because of this forwarding scheme, IPv6 nodes on the same point-to-multipoint domain cannot have direct communication: any communication between them must go through the first-hop router of the same domain.

It is assumed in this document that link-layer addresses on a point-to-multipoint domain are unique from the first-hop router's point of view (e.g., in an untrusted Ethernet architecture, this assumption can be guaranteed thanks to mechanisms such as Media Access Control (MAC) address translation performed by an aggregation device between IPv6 nodes and the first-hop router).

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. Background

Terminology in this document follows that in "Neighbor Discovery for IP version 6 (IPv6)" [RFC4861] and "IPv6 Stateless Address Autoconfiguration" [RFC4862]. In addition, this section defines additional terms related to DSL and Fiber access architectures, which are an important case where the solution described in this document can be used:

Customer Premises Equipment (CPE)
: The first IPv6 node in a customer's network.

Access Node (AN)
: The first aggregation point in the public access network. It is considered as an L2 bridge in this document.

Broadband Network Gateway (BNG)
: The first-hop router from the CPE's point of view.

VLAN N:1 architecture
: A point-to-multipoint architecture where many CPEs are connected to the same VLAN. The CPEs may be connected on the same or different Access Nodes.

split-horizon model
: A forwarding scheme where CPEs cannot have direct layer 2 communications between them (i.e., IP flows must be forwarded through the BNG via routing).

The following figure shows where the different entities are, as defined above.

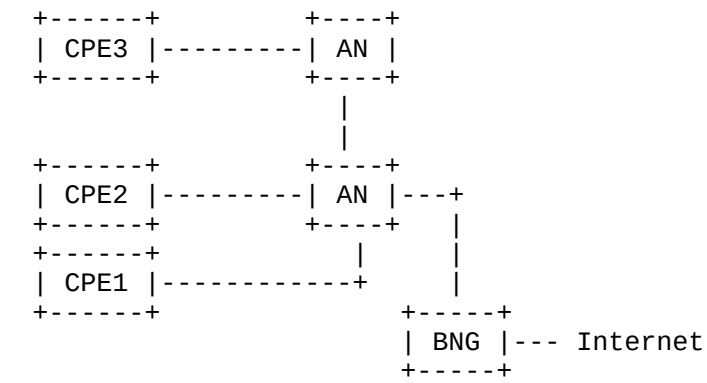


Figure 1: DSL and Fiber Access Architecture

3. Why Existing IETF Solutions Are Not Sufficient

In a DSL or Fiber access architecture depicted in Figure 1, CPE1, CPE2, CPE3, and the BNG are IPv6 nodes, while AN is an L2 bridge providing connectivity between the BNG and each CPE. The AN enforces a split-horizon model so that CPEs can only send and receive frames (e.g., Ethernet frames) to and from the BNG but not to each other. That said, the BNG is on the same link with all CPEs, but a given CPE is not on the same link with any other CPE.

3.1. Duplicate Address Detection

Duplicate Address Detection (DAD) [RFC4862] is performed when an IPv6 node verifies the uniqueness of a tentative IPv6 address. This node sends a Neighbor Solicitation (NS) message with the IP destination set to the solicited-node multicast address of the tentative address.

This NS message is multicasted to other nodes on the same link. When the tentative address is already used on the link by another node, this last one replies with a Neighbor Advertisement (NA) message to inform the first node. So, when performing DAD, a node expects the NS messages to be received by any node currently using the tentative address.

However, in a point-to-multipoint network with a split-horizon forwarding scheme implemented in the AN, the CPEs are prevented from talking to each other directly. All packets sent out from a CPE are forwarded by the AN only to the BNG but not to any other CPE. NS messages sent by a certain CPE will be received only by the BNG and will not reach other CPEs. So, other CPEs have no idea that a certain IPv6 address is used by another CPE. That means, in a network with split-horizon, DAD, as defined in [RFC4862], can't work properly without additional help.

3.2. Neighbor Discovery Proxy

Neighbor Discovery (ND) Proxy [RFC4389] is designed for forwarding ND messages between different IP links where the subnet prefix is the same. An ND Proxy function on a bridge ensures that packets between nodes on different segments can be received by this function and have the correct link-layer address type on each segment. When the ND Proxy receives a multicast ND message, it forwards it to all other interfaces on a same link.

In DSL or Fiber networks, when the AN, acting as an ND Proxy, receives an ND message from a CPE, it will forward it to the BNG but none of the other CPEs, as only the BNG is on the same link with the CPE. Hence, implementing ND Proxy on the AN would not help a CPE acknowledge link-local addresses used by other CPEs.

As the BNG must not forward link-local scoped messages sent from a CPE to other CPEs, ND Proxy cannot be implemented in the BNG.

3.3. 6LoWPAN Neighbor Discovery

[RFC6775] defines an optional modification of DAD for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPAN). When a 6LoWPAN node wants to configure an IPv6 address, it registers that address with one or more of its default routers using the Address Registration Option (ARO). If this address is already owned by another node, the router informs the 6LoWPAN node that this address cannot be configured.

This mechanism requires modifications in all hosts in order to support the ARO.

3.4. IPv6 Mobility Manager

According to [RFC6275], a home agent acts as a proxy for mobile nodes when they are away from the home network: the home agent defends a mobile node's home address by replying to NS messages with NA messages.

There is a problem for this mechanism if it is applied in a DSL or Fiber public access network. Operators of such networks require that an NA message is only received by the sender of the corresponding NS message, for security and scalability reasons. However, the home agent per [RFC6275] multicasts NA messages on the home link and all nodes on this link will receive these NA messages. This shortcoming prevents this mechanism from being deployed in DSL or Fiber access networks directly.

4. Duplicate Address Detection Proxy (DAD-Proxy) Specifications

First, it is important to note that, as this mechanism is strongly based on DAD [RFC4862], it is not completely reliable, and the goal of this document is not to fix DAD.

4.1. DAD-Proxy Data Structure

A BNG needs to store in a Binding Table information related to the IPv6 addresses generated by any CPE. This Binding Table can be distinct from the Neighbor Cache. This must be done per point-to-multipoint domain (e.g., per Ethernet VLAN). Each entry in this Binding Table MUST contain the following fields:

- IPv6 Address
- Link-layer Address

For security or performances reasons, it must be possible to limit the number of IPv6 addresses per link-layer address (possibly, but not necessarily, to 1).

On the reception of an unsolicited NA (e.g., when a CPE wishes to inform its neighbors of a new link-layer address) for an IPv6 address already recorded in the Binding Table, each entry associated to this IPv6 address MUST be updated consequently: the current link-layer address is replaced by the one included in the unsolicited NA message.

For security or performances reasons, the Binding Table MUST be large enough for the deployment in which it is used: if the Binding Table is distinct from the Neighbor Cache, it MUST be at least the same

size as this last one. Implementations MUST either state the fixed size of the Binding Table that they support or make the size configurable. In the latter case, implementations MUST state the largest Binding Table size that they support. Additionally, implementations SHOULD allow an operator to inquire about the current occupancy level of the Binding Table to determine if it is about to become full. Implementations encountering a full Binding Table will likely handle it in a way similar to NS message loss.

It is recommended to apply technical solutions to minimize the risk that the Binding Table becomes full. These solutions are out of the scope of this document.

4.2. DAD-Proxy Mechanism

When a CPE performs DAD, as specified in [RFC4862], it sends a Neighbor Solicitation (NS) message, with the unspecified address as the source address, in order to check if a tentative address is already in use on the link. The BNG receives this message and MUST perform actions specified in the following sections based on the information in the Binding Table.

4.2.1. No Entry Exists for the Tentative Address

When there is no entry for the tentative address, the BNG MUST create one with the following information:

- IPv6 Address field set to the tentative address in the NS message.
- Link-layer Address field set to the link-layer source address in the link-layer header of the NS message.

The BNG MUST NOT reply to the CPE or forward the NS message.

4.2.2. An Entry Already Exists for the Tentative Address

When there is an entry for the tentative address, the BNG MUST check the following conditions:

- The address in the Target Address field in the NS message is equal to the address in the IPv6 Address field in the entry.
- The source address of the IPv6 Header in the NS message is equal to the unspecified address.

When these conditions are met and the source address of the link-layer header in the NS message is equal to the address in the Link-layer Address field in the entry, that means the CPE is still

performing DAD for this address. The BNG MUST NOT reply to the CPE or forward the NS message.

When these conditions are met and the source address of the link-layer header in the NS message is not equal to the address in the Link-layer Address field in the entry, that means possibly another CPE is performing DAD for an already owned address. The BNG then has to verify whether there is a real conflict by checking if the CPE whose IPv6 address is in the entry is still connected. In the following text, we will call IPv6-CPE1 the IPv6 address of the existing entry in the Binding Table, Link-layer-CPE1 the link-layer address of that entry, and Link-layer-CPE2 the link-layer address of the CPE that is performing DAD, which is different from Link-layer-CPE1.

The BNG MUST check if the potential address conflict is real. In particular:

- If IPv6-CPE1 is in the Neighbor Cache and it is associated with Link-layer-CPE1, the reachability of IPv6-CPE1 MUST be confirmed as explained in Section 4.2.3.

- If IPv6-CPE1 is in the Neighbor Cache, but in this cache it is associated with a link-layer address other than Link-layer-CPE1, that means that there is possibly a conflict with another CPE, but that CPE did not perform DAD. This situation is out of the scope of this document, since one assumption made above is that all the nodes of a point-to-multipoint domain (except the DAD proxy itself) perform DAD.

- If IPv6-CPE1 is not in the Neighbor Cache, then the BNG MUST create a new entry based on the information of the entry in the Binding Table. This step is necessary in order to trigger the reachability check as explained in Section 4.2.3. The entry in the Neighbor Cache MUST be created based on the algorithm defined in Section 7.3.3 of [RFC4861], in particular by treating this case as though a packet other than a solicited Neighbor Advertisement were received from IPv6-CPE1. Thus, the new entry of the Neighbor Cache MUST contain the following information:

  * IPv6 address: IPv6-CPE1

  * Link-layer address: Link-layer-CPE1

  * State: STALE

  The reachability of IPv6-CPE1 MUST be confirmed as soon as possible following the procedure explained in Section 4.2.3.

4.2.3. Confirmation of Reachability to Check the Validity of the Conflict

Given that the IPv6-CPE1 is in an entry of the Neighbor Cache, the reachability of IPv6-CPE1 is checked by using the Neighbor Unreachability Detection (NUD) mechanism described in Section 7.3.1 of [RFC4861]. This mechanism MUST be triggered as though a packet had to be sent to IPv6-CPE1. Note that in some cases this mechanism does not do anything. For instance, if the state of the entry is REACHABLE and a positive confirmation was received recently that the forward path to the IPv6-CPE1 was functioning properly (see RFC 4861 for more details), this mechanism does not do anything.

Next, the behavior of the BNG depends on the result of the NUD process, as explained in the following sections.

4.2.3.1. The Result of the NUD Process is Negative

If the result of the NUD process is negative (i.e., if this process removes IPv6-CPE1 from the Neighbor Cache), that means that the potential conflict is not real.

The conflicting entry in the Binding Table (Link-layer-CPE1) is deleted and it is replaced by a new entry with the same IPv6 address, but the link-layer address of the CPE is performing DAD (Link-layer-CPE2), as explained in Section 4.2.1.

4.2.3.2. The Result of the NUD Process is Positive

If the result of the NUD process is positive (i.e., if after this process the state of IPv6-CPE1 is REACHABLE), that means that the potential conflict is real.

As shown in Figure 2, the BNG MUST reply to the CPE that is performing DAD (CPE2 in Figure 1) with an NA message that has the following format:

Layer 2 Header Fields:

Source Address
: The link-layer address of the interface on which the BNG received the NS message.

Destination Address
: The source address in the Layer 2 Header of the NS message received by the BNG (i.e., Link-layer-CPE2).

IPv6 Header Fields:

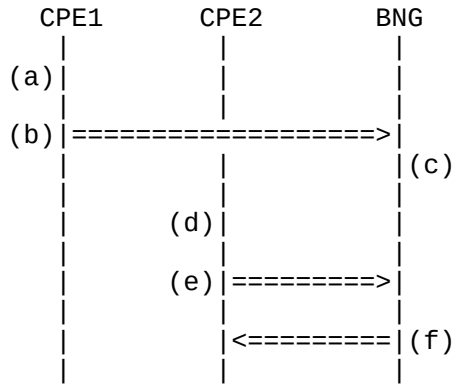
Source Address
An address assigned to the interface from which the advertisement is sent.

Destination Address
The all-nodes multicast address.

ICMPv6 Fields:

Target Address
The tentative address already used (i.e., IPv6-CPE1).

Target Link-layer Address
The link-layer address of the interface on which the BNG received the NS message.



(a) CPE1 generates a tentative address
(b) CPE1 performs DAD for this one
(c) BNG updates its Binding Table
(d) CPE2 generates a same tentative address
(e) CPE2 performs DAD for this one
(f) BNG informs CPE2 that DAD fails

Figure 2: DAD Failure

The BNG and the CPE MUST support the unicast transmission on the link layer of IPv6 multicast messages [RFC6085], to be able, respectively, to generate and to process such a packet format.

5. Manageability Considerations

The BNG SHOULD support a mechanism to log and emit alarms whenever a duplication of IPv6 addresses is detected by the DAD-Proxy function. Moreover, the BNG SHOULD implement a function to allow an operator to access logs and to see the current entries in the Binding Table. The management of access rights to get this information is out of the scope of this document.

6. Security Considerations

6.1. Interoperability with SEND

The mechanism described in this document will not interoperate with SEcure Neighbor Discovery (SEND) [RFC3971]. This is due to the BNG not owning the private key associated with the Cryptographically Generated Address (CGA) [RFC3972] needed to correctly sign the proxied ND messages [RFC5909].

Secure Proxy ND Support for SEND [RFC6496] has been specified to address this limitation, and it SHOULD be implemented and used on the BNG and the CPEs.

6.2. Protection against IP Source Address Spoofing

To ensure protection against IP source address spoofing in data packets, this proposal can be used in combination with Source Address Validation Improvement (SAVI) mechanisms [RFC6620] [SAVI-SEND] [SAVI-MIX].

If SAVI mechanisms are used, the SAVI device is the BNG, and the Binding Anchor for a CPE is its MAC address, which is assumed to be unique in this document (cf. Section 1).

7. Acknowledgments

The authors would like to thank Alan Kavanagh, Wojciech Dec, Suresh Krishnan, and Tassos Chatzithomaoglou for their comments. The authors would like also to thank the IETF 6man WG members and the BBF community for their support.

8. References

8.1. Normative References

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

[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.

[RFC6085] Gundavelli, S., Townsley, M., Troan, O., and W. Dec, "Address Mapping of IPv6 Multicast Packets on Ethernet", RFC 6085, January 2011.

8.2. Informative References

[RFC3971] Arkko, J., Kempf, J., Zill, B., and P. Nikander, "SEcure Neighbor Discovery (SEND)", RFC 3971, March 2005.

[RFC3972] Aura, T., "Cryptographically Generated Addresses (CGA)", RFC 3972, March 2005.

[RFC4389] Thaler, D., Talwar, M., and C. Patel, "Neighbor Discovery Proxies (ND Proxy)", RFC 4389, April 2006.

[RFC5072] Varada, S., Ed., Haskins, D., and E. Allen, "IP Version 6 over PPP", RFC 5072, September 2007.

[RFC5909] Combes, J-M., Krishnan, S., and G. Daley, "Securing Neighbor Discovery Proxy: Problem Statement", RFC 5909, July 2010.

[RFC6275] Perkins, C., Johnson, D., and J. Arkko, "Mobility Support in IPv6", RFC 6275, July 2011.

[RFC6496] Krishnan, S., Laganier, J., Bonola, M., and A. Garcia-Martinez, "Secure Proxy ND Support for SEcure Neighbor Discovery (SEND)", RFC 6496, February 2012.

[RFC6620] Nordmark, E., Bagnulo, M., and E. Levy-Abegnoli, "FCFS SAVI: First-Come, First-Served Source Address Validation Improvement for Locally Assigned IPv6 Addresses", RFC 6620, May 2012.

[RFC6775] Shelby, Z., Chakrabarti, S., Nordmark, E., and C. Bormann, "Neighbor Discovery Optimization for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs)", RFC 6775, November 2012.

[SAVI-MIX] Bi, J., Yao, G., Halpern, J., and E. Levy-Abegnoli, Ed., "SAVI for Mixed Address Assignment Methods Scenario", Work in Progress, May 2013.

[SAVI-SEND] Bagnulo, M. and A. Garcia-Martinez, "SEND-based Source-Address Validation Implementation", Work in Progress, April 2013.

[TR-101] The Broadband Forum, "Migration to Ethernet-Based DSL Aggregation", Issue 2, Technical Report TR-101, July 2011, <http://www.broadband-forum.org/technical/download/TR-101_Issue-2.pdf>.

Appendix A. DAD-Proxy State Machine

This appendix, which is informative, contains a summary (cf. Table 1) of the actions done by the BNG when it receives a DAD-based NS (DAD-NS) message. The tentative address in this message is IPv6-CPE1 and the associated link-layer address is Link-layer-CPE2. The actions are precisely specified in Section 4.2.

Event	Check	Action	New event
DAD-NS message reception.	* No entry for IPv6-CPE1 in the Binding Table.	Create an entry for IPv6-CPE1 bound to Link-layer-CPE2 in the Binding Table.	-
	* Entry for IPv6-CPE1 in the Binding Table.	-	Existing entry.
Existing entry.	* Link-layer-CPE2 bound to IPv6-CPE1 in the Binding Table.	-	-
	* Another link-layer address, Link-layer-CPE1, bound to IPv6-CPE1 in the Binding Table.	-	Conflict?
Conflict?	* IPv6-CPE1 associated to Link-layer-CPE1 in the Neighbor Cache.	-	Reachable?
	* IPv6-CPE1 associated to another link-layer address than Link-layer-CPE1 in the Neighbor Cache.	Out of scope.	-
	* IPv6-CPE1 is not in the Neighbor Cache.	Create an entry for IPv6-CPE1 associated to Link-layer-CPE1 in the Neighbor Cache.	Reachable?

Reachable?	* NUD process is negative.	IPv6-CPE2 is bound to Link-layer-CPE2, instead to Link-layer-CPE1, in the Binding Table.	-
	* NUD process is positive.	A NA message is sent.	-

Table 1: DAD-Proxy State Machine

Authors' Addresses

Fabio Costa
France Telecom Orange
61 rue des Archives
75141 Paris Cedex 03
France

EMail: fabio.costa@orange.com

Jean-Michel Combes (editor)
France Telecom Orange
38 rue du General Leclerc
92794 Issy-les-Moulineaux Cedex 9
France

EMail: jeanmichel.combes@orange.com

Xavier Pougnard
France Telecom Orange
2 avenue Pierre Marzin
22300 Lannion
France

EMail: xavier.pougnard@orange.com

Hongyu Li
Huawei Technologies
Huawei Industrial Base
Shenzhen
China

EMail: lihy@huawei.com