6man Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 22, 2011

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

June 20, 2011

Duplicate Address Detection Proxy
draft-ietf-6man-dad-proxy-01

Abstract

The document describes a mechanism allowing the use of Duplicate Address Detection (DAD) by IPv6 nodes in a point-to-multipoint architecture with "split-horizon" forwarding scheme. Based on the DAD signalling, 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 replies instead of this last one.

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 December 22, 2011.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents

carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction

This document explains why Duplicate Address Detection (DAD) mechanism [RFC4862] cannot be used in a point-to-multipoint architecture with "split-horizon" forwarding scheme. 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.

This document also specifies a function called DAD proxy allowing the use of DAD by the nodes on the same point-to-multipoint domain with "split-horizon" forwarding scheme. 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. However, if it is necessary to cover the scenarios where this assumption is not met, additional solutions could be defined in the future that work in conjunction with the mechanism described here.

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 "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) document [RFC4861] and IPv6 Stateless Address Autoconfiguration document [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 a 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 are the different entities defined above.

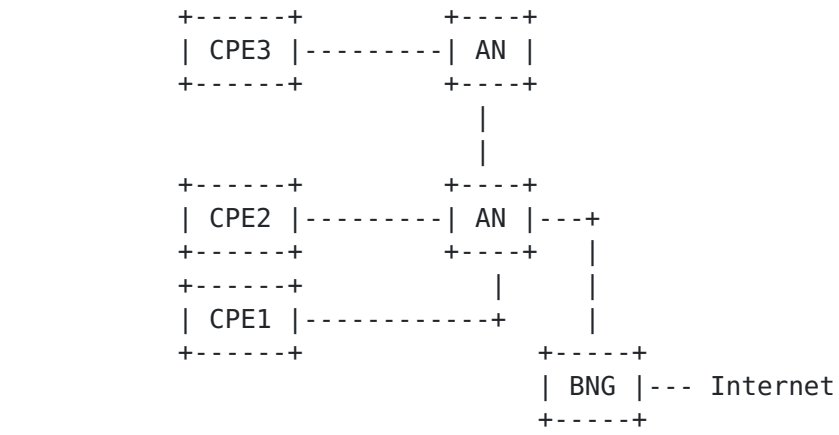


Figure 1: DSL and Fiber access Architecture

3. Why existing IETF solutions are not sufficient?

In a DSL or Fiber access architecture depicted in Figure 1, CPE1,2,3 and the BNG are IPv6 nodes, while AN is a 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 a same link with all CPE, but one CPE is not on a same link with any other CPE.

3.1. Duplicate Address Detection

Duplicate Address Dectection (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 solicited-node multicast address of the tentative address. This NS message is multicasted to other nodes on a 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 are received by other nodes.

However, in a point-to-multipoint network with 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 would be forwarded by AN only to the BNG but not to any other CPE. That said, 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 per [RFC4862] can't work properly without an additional helper.

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. A 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 AN, acting as a ND Proxy, receives a ND message from a CPE, it will forward it to the BNG but none of other CPEs, as only the BNG is on the same link with the CPE. Hence, implementing ND Proxy on 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

[I-D.ietf-6lowpan-nd] defines an optional modification of DAD for a 6LowPAN. When a 6LoWPAN node wants to configure an IPv6 address, it registers that address with one or more of its default router using the Address Registration option (ARO). If this address is already owned by another node, the router informs the 6LoWPAN node this

address cannot be configured.

A problem for this mechanism is that it requires modifications in hosts in order to support the Address Registration option.

3.4. IPv6 Mobility Manager

According to [RFC3775], a home agent acts as a proxy for mobile nodes when these last ones are away from the home network: the home agent defends an 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 a NA message is only received by the sender of the corresponding NS message, for security and scalability reasons. However, the home agent per [RFC3775] multicasts NA messages on the home link and all nodes on this link will receive these NA messages. This shortcoming prevents this mechanism being deployed in DSL or Fiber access networks directly.

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

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 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).

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 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 depending 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 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 performs 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, we will call IPv6-CPE1 the IPv6 address of the existing entry, Link-layer-CPE1 the Link-layer address of that entry and Link-layer-CPE2 the Link-layer address of the CPE which 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 it is associated with another Link-layer address 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. This case could be covered in the future by additional solutions that work in conjunction with the DAD proxy.

- 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 reachibility 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 considering the case as if a packet other than a solicited Neighbor Advertisement was received from IPv6-CPE1. That means that the new entry of the Neighbor Cache MUST contain the following information:

  * IPv6 address: IPv6-CPE1

  * Link-layer address: Link-layer-CPE1

  * State: STALE

  Then 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 NUD (Neighbor Unreachibility Detection) mechanism described in section 7.3.1 of [RFC4861]. This mechanism MUST be triggered as if a packet has 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).

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 which 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 CPE that is performing DAD (CPE2 in Figure 1) with a NA message which 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:

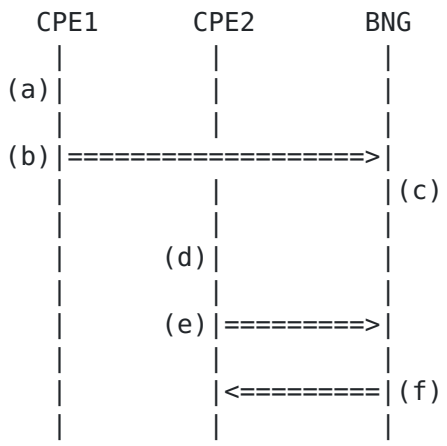
Source Address
: An address assigned to the interface from which the advertisement is sent.

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 generated 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

The BNG and the CPE MUST support the Unicast Transmission on Link-layer of IPv6 Multicast Messages [RFC6085], to be able, respectively, to generate and to process such a packet format.

5. IANA Considerations

No new options or messages are defined in this document.

6. Security Considerations

6.1. Interoperability with SEND

If SEcure Neighbor Discovery (SEND) [RFC3971] is used, the mechanism specified in this document may break the security. Indeed, if an entry already exists and the BNG has to send a reply (cf. Section 4.2.2), the BNG doesn't own the private key(s) associated with to the Cryptographically Generated Addresses (CGA) [RFC3972] to correctly sign the proxied ND messages [RFC5909].

To keep the same level of security, Secure Proxy ND Support for SEND [I-D.ietf-csi-proxy-send] SHOULD be used and implemented on the BNG and the CPEs.

6.2. IP source address spoofing protection

To ensure a protection against IP source address spoofing in data packets, this proposal may be used in combinaison with Source Address Validation Improvement (SAVI) mechanisms [I-D.ietf-savi-fcfs] [I-D.ietf-savi-send].

7. Acknowledgments

TbD

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

[I-D.ietf-6lowpan-nd] Shelby, Z., Chakrabarti, S., and E. Nordmark, "Neighbor Discovery Optimization for Low Power and Lossy Networks (6LoWPAN)", draft-ietf-6lowpan-nd-17 (work in progress), June 2011.

[I-D.ietf-csi-proxy-send] Krishnan, S., Laganier, J., Bonola, M., and A. Garcia-Martinez, "Secure Proxy ND Support for SEND", draft-ietf-csi-proxy-send-05 (work in progress), May 2010.

[I-D.ietf-savi-fcfs] Nordmark, E., Bagnulo, M., and E. Levy-Abegnoli, "FCFS SAVI: First-Come First-Serve Source-Address Validation for Locally Assigned IPv6 Addresses", draft-ietf-savi-fcfs-09

(work in progress), April 2011.

[I-D.ietf-savi-send]
Bagnulo, M. and A. Garcia-Martinez, "SEND-based Source-Address Validation Implementation", draft-ietf-savi-send-05 (work in progress), April 2011.

[RFC3775] Johnson, D., Perkins, C., and J. Arkko, "Mobility Support in IPv6", June 2004.

[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", RFC 4389, April 2006.

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

Appendix A. Open issues

- What happens when the BNG receives a NA message with O-bit set to 1 (e.g. the Link-Layer address of the CPE has changed)?

Authors' Addresses

Fabio Costa
France Telecom Orange
38 rue du General Leclerc
92794 Issy-les-Moulineaux Cedex 9
France

Email: fabio.costa@orange-ftgroup.com

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

Email: jeanmichel.combes@orange-ftgroup.com

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

Email: xavier.pougnard@orange-ftgroup.com

Hongyu Li
Huawei Technologies
Huawei Industrial Base
Shenzhen
China

Email: lihy@huawei.com