6man Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 3, 2011

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

July 2, 2010

Duplicate Address Detection Proxy
draft-costa-6man-dad-proxy-00

Abstract

The document describes a mechanism allowing the use of Duplicate Address Detection (DAD) by IPv6 nodes in a VLAN N:1 with "split-horizon" model DSL architecture. Based on the DAD signalling, the first hop router stores all used addresses on the VLAN in a Binding Table. When a node performs DAD for an address already used by another node, the first hop router replies instead of this last one.

Table of Contents

1. Introduction

This document explains why Duplicate Address Detection (DAD) mechanism [RFC4862] cannot be used in a VLAN N:1 with "split-horizon" model DSL architecture. One of the main reasons is IPv6 nodes on the same VLAN cannot have direct communication: all the messages between them must go through the first hop router.

This document specifies a mechanism allowing the use of DAD by the hosts on the same VLAN. It only impacts the first hop router and it doesn't need modifications on the other IPv6 nodes.

It is assumed in this document that Link-layer addresses on a VLAN 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 the use of "MAC Address Translation" mechanism performed upstream by a 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 the DSL architecture:

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

Access Node (AN)
: The first aggregation point in the DSL 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 many-to-one forwarding scheme 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 model where CPEs cannot have direct layer 2 nor layer 3 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.

```
   +------+         +----+
   | CPE3 |---------| AN |
   +------+         +----+
                      |
                      |
   +------+         +----+
   | CPE2 |---------| AN |---+
   +------+         +----+   |
   +------+            |     |
   | CPE1 |------------+     |
   +------+               +-----+
                          | BNG |--- Internet
                          +-----+
```

Figure 1: DSL Architecture

3. Why IETF solutions don't work?

In a DSL architecture depicted in Figure 1, CPE1,2,3 and BNG are IPv6 nodes, while AN is a special bridge providing links between CPEs and the BNG. AN enforces in a split-horizon mode so that all CPEs can only talk to BNG but not to each other. That said, each CPE is on a same link with BNG, 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 address. This node sends a Neighbour 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 DSL network depicted in Figure 1, split-horizon is implemented on AN to prevent CPEs from talking to each other

directly. All packets sent out from a CPE would be forward by AN only to BNG but none of other CPE nodes. That said, NS messages sent by a certain CPE will be received only by BNG, which will never forward these NS messages to other CPEs. So, other CPEs have no idea that a certain address is used by another CPE. That means, in a network with split-horizon, DAD per RFC4862 can't work properly.

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 forwards received ND messages to other segments with correct-link layer type address. When the ND proxy receives a multicast ND message, it forwards it to all other interfaces on a same link.

In the DSL network depicted in Figure 1, when AN, acting as a ND Proxy, receives a ND message from a CPE, it will forward it to BNG but none of other CPEs, as only BNG is on the same link with the CPE. Hence, implementing ND Proxy on AN doesn't 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 this one to 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 the DSL network depicted in Figure 1. Operators of DSL networks require a NA message is only received by the sender of the corresponding NS message for security reason. 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 a DSL network 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 on a VLAN. Each entry in this Binding Table MUST contain the following fields:

- IPv6 Address

- Link-layer Address

- Creation Time

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.

- Creation Time set to the value of the BNG clock when the entry is created.

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.

o 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.

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 another CPE performs DAD for an already owned address. As shown in Figure 2, the BNG MUST reply to the CPE that has sent the NS message 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.

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.

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

```
     CPE1      CPE2       BNG
       |         |          |
    (a)|         |          |
       |         |          |
    (b)|===================>|
       |         |          |(c)
       |         |          |
       |      (d)|          |
       |         |          |
       |      (e)|=========>|
       |         |          |
       |         |<=========|(f)
       |         |          |
```

(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 of IPv6 Multicast Messages on Link-layer [I-D.gundavelli-v6ops-l2-unicast], 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 [I-D.ietf-csi-sndp-prob].

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

[I-D.gundavelli-v6ops-l2-unicast]
Gundavelli, S., Townsley, M., Troan, O., and W. Dec, "Unicast Transmission of IPv6 Multicast Messages on Link-layer", draft-gundavelli-v6ops-l2-unicast-00 (work in progress), February 2010.

[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.

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", draft-ietf-6lowpan-nd-10 (work in progress), June 2010.

[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-04 (work in progress), May 2010.

[I-D.ietf-csi-sndp-prob]
Combes, J., Krishnan, S., and G. Daley, "Securing Neighbor Discovery Proxy: Problem Statement",

draft-ietf-csi-sndp-prob-04 (work in progress), January 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 Addresses", draft-ietf-savi-fcfs-03 (work in progress), May 2010.

[I-D.ietf-savi-send] Bagnulo, M. and A. Garcia-Martinez, "SEND-based Source-Address Validation Implementation", draft-ietf-savi-send-03 (work in progress), May 2010.

[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.

Appendix A. Open issues

- o A same VLAN on n different interfaces (n > 1) of a BNG?

- o 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)?

- o When to remove a entry from the Binding Table?

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