Propagation of ARP/ND Flags in EVPN
draft-ietf-bess-evpn-na-flags-05

Abstract

   An EVPN MAC/IP Advertisement route can optionally carry an IPv4 or
   IPv6 addresses associated with a MAC address. Remote PEs can use
   this information to populate their ARP/ND tables on IRB interfaces or
   their proxy-ARP/ND tables in Broadcast Domains (BD). PEs can then
   reply locally (act as an ARP/ND proxy) to IPv4 ARP requests and IPv6
   Neighbor Solicitation messages and reduce/suppress the flooding
   produced by the Address Resolution procedure. However, the
   information conveyed in the MAC/IP route may not be enough for the
   remote PE to reply to local ARP or ND requests. For example, if a PE
   learns an IPv6->MAC ND entry via EVPN, the PE would not know if that
   particular IPv6->MAC pair belongs to a host, a router or a host with
   an anycast address, as this information is not carried in the EVPN
   MAC/IP Advertisement routes. Similarly, other information relevant
   to the IP->MAC ARP/ND entries may be needed. This document defines
   an Extended Community that is advertised along with an EVPN MAC/IP
   Advertisement route and carries information relevant to the ARP/ND
   resolution, so that an EVPN PE implementing a proxy-ARP/ND function
   can reply to ARP Requests or Neighbor Solicitations with the correct
   information.

Status of This Memo

This Internet-Draft will expire on January 28, 2021.

Copyright Notice

Table of Contents

1. Introduction

An EVPN MAC/IP Advertisement route can optionally carry an IPv4 or IPv6 addresses associated with a MAC address. Remote PEs can use this information to populate their ARP/ND tables on IRB interfaces or their proxy-ARP/ND tables in Broadcast Domains (BD). PEs can then reply locally (act as an ARP/ND proxy, as per [RFC7432]) to IPv4 ARP requests and IPv6 Neighbor Solicitation messages and reduce/suppress the flooding produced by the Address Resolution procedure. However, the information conveyed in the EVPN MAC/IP Advertisement route may not be enough for the remote PE to reply to local ARP or ND requests. For example, if a PE learns an IPv6->MAC ND entry via EVPN, the PE would not know if that particular IPv6->MAC pair belongs to a host, a router or a host with an anycast address, as this information is not carried in the EVPN MAC/IP Advertisement routes. Similarly, other

information relevant to the host advertised in the MAC/IP
Advertisement route may be needed.

This document defines an Extended Community that is advertised along
with an EVPN MAC/IP Advertisement route and carries information
relevant to the ARP/ND resolution, so that an EVPN PE implementing a
proxy-ARP/ND function can reply to ARP Requests or Neighbor
Solicitations with the correct information. In particular, the Flags
defined in [RFC4861] can now be conveyed along with a MAC/IP
Advertisement route, so that an egress EVPN PE can issue Neighbor
Advertisement messages with the correct Flag information.

The Flags are carried in the EVPN Address Resolution Protocol (ARP)
and Neighbor Discovery (ND) Extended Community, as described in the
following sections.

1.1. Terminology and Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
"SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and
"OPTIONAL" in this document are to be interpreted as described in BCP
14 [RFC2119] [RFC8174] when, and only when, they appear in all
capitals, as shown here.

EVPN: Ethernet Virtual Private Networks, as in [RFC7432].

BD: Broadcast Domain, also described in [RFC7432].

ARP: refers to the Address Resolution Protocol.

ND: refers to the Neighbor Discovery protocol.

Proxy-ARP/ND: refers to a function on the EVPN PEs by which received
Address Resolution Protocol (ARP) Requests or Neighbor Solicitation
(NS) messages are replied locally by the PE, without the need to
flood the requests to remote PEs in the BD. In order to reply to ARP
Requests or NS messages, the PE does a lookup on an ARP/ND table,
that is a collection of IP->MAC entries learned by the PE.

IP->MAC: refers to an IP address and MAC address combination that
represents a given host and is added to an Address Resolution
Protocol table or Neighbor Discovery table. This document uses
IP->MAC generically for IPv4 and IPv6 addresses. When something is
specific to IPv4, the document will use IPv4->MAC and likewise,
IPv6->MAC will be used when something is specific to IPv6 entries
only.

Familiarity with the terminology in [RFC7432] and [RFC4861] is
expected.

2. The EVPN ARP/ND Extended Community

This document defines a new EVPN Extended Community (Type field value
of 0x06) with a Sub-Type of 0x08, as allocated by IANA. It is
advertised along with EVPN MAC/IP Advertisement routes that carry an
IPv4 or IPv6 address.

```
   0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  | Type=0x06     | Sub-Type=0x08 |Flags (1 octet)| Reserved=0    |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |                          Reserved=0                           |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Flags field:

```
   0 1 2 3 4 5 6 7
  +-+-+-+-+-+-+-+-+
  |       |I| |O|R|
  +-+-+-+-+-+-+-+-+
```

The following Flags are defined in the Flags field, third octet of
the Extended Community:

R - Router Flag (corresponds to Bit 23 of the Extended Community)

Bit 7 of the Flags field is defined as the "Router Flag". When set,
the R Flag indicates that the IPv6->MAC pair advertised in the MAC/IP
Advertisement route along with the Extended Community belongs to a
router. If the R Flag is zero, the IPv6->MAC pair belongs to a host.
The receiving PE implementing the ND function will use this
information in Neighbor Advertisement messages for the associated
IPv6 address. This Flag has no meaning for ARP IPv4->MAC entries and
MUST be ignored when the Extended Community is received with an EVPN
MAC/IP Advertisement route for an IPv4->MAC pair.

O - Override Flag (corresponds to Bit 22 of the Extended Community)
Bit 6 of the Flags field is defined as the "Override Flag". An
egress PE will normally advertise IPv6->MAC pairs with the O Flag
set, and only when IPv6 "anycast" is enabled in the BD or interface,
the PE will send an IPv6->MAC pair with the O Flag = 0. The ingress
PE will install the ND entry with the received O Flag and will use
this information when replying to a Neighbor Solicitation for the
IPv6 address. Similarly to the Router Flag, the Override Flag has no
meaning for ARP IPv4->MAC entries and MUST be ignored when the

Extended Community is received with an EVPN MAC/IP Advertisement
route for an IPv4->MAC pair.

I - Immutable ARP/ND Binding Flag (corresponds to Bit 20 of the
Extended Community)

Bit 4 of the Flags field is defined as the "Immutable ARP/ND Binding
Flag". When set, the egress PE indicates that the IP->MAC pair sent
in an EVPN MAC/IP Advertisement route along with the Extended
Community is a configured ARP/ND entry, and the IP address in the
EVPN MAC/IP Advertisement route can only be bound together with the
MAC address specified in the same route.

Bits 0-3 and 5 are not assigned by this document.

3. Use of the EVPN ARP/ND Extended Community

This section describes the relevant procedures when advertising and
processing the EVPN ARP/ND Extended Community. In all the procedures
below a "PE" must be interpreted as a "PE which supports the ND/ARP
proxy function (introduced by [RFC7432]) and implements the
propagation of the ARP/ND Flags that this document specifies".

3.1. Transmission of the EVPN ARP/ND Extended Community

When an IP->MAC entry is not learned via EVPN, a PE may learn IP->MAC
pairs in the management plane (this will create static entries in the
ARP/ND or proxy-ARP/ND table) or by snooping ARP or Neighbor
Advertisement (NA) messages coming from the CE (this will create
dynamic entries). Those static and dynamic IP->MAC entries will be
advertised in EVPN MAC/IP Advertisement routes that use the EVPN ARP/
ND Extended Community as follows:

o Advertised MAC/IP Advertisement routes for IPv6->MAC entries MUST
   include the ARP/ND Extended Community with the R and O Flag values
   associated with the entry. Those Flag values are either
   dynamically learned (from NA messages) or configured in case of
   static entries.

o MAC/IP Advertisement routes for IPv4->MAC entries MAY include the
   ARP/ND Extended Community. If the EVPN ARP/ND Extended Community
   is advertised along with an EVPN IPv4/MAC Advertisement route, the
   R and O Flags SHOULD be set to zero.

o If an IP->MAC pair is static (it has been configured) the
   corresponding MAC/IP Advertisement route MUST be sent along with
   an ARP/ND Extended Community with the I Flag set.

o This Extended Community does not have any impact on the procedures
   described in [RFC7432], including the advertisement of the MAC
   Mobility Extended Community along with the MAC/IP Advertisement
   route.

3.2. Reception of the EVPN ARP/ND Extended Community

In addition to the procedures specified in [RFC7432] a PE receiving a
MAC/IP Advertisement route will process the EVPN ARP/ND Extended
Community as follows:

o R and O Flags processing:

   * If the EVPN MAC/IP Advertisement route contains an IPv6 address
      and the EVPN ARP/ND Extended Community, the PE MUST add the R
      and O Flags to the ND entry in the ND or proxy-ND table and use
      that information in Neighbor Advertisements when replying to a
      Solicitation for the IPv6 address.

   * If no EVPN ARP/ND Extended Community is received along with the
      route, the PE will add the default R and O Flags to the entry.
      The default R Flag SHOULD be an administrative choice. The
      default O Flag SHOULD be 1.

   * A PE MUST ignore the received R and O Flags for an EVPN MAC/IP
      Advertisement route that contains an IPv4->MAC pair.

o I Flag processing:

   * A PE receiving an EVPN MAC/IP Advertisement route containing an
      IP->MAC and the I Flag set SHOULD install the IP->MAC entry in
      the ARP/ND or proxy-ARP/ND table as an "Immutable binding".
      This Immutable binding entry will override an existing non-
      immutable binding for the same IP->MAC. The absense of the
      EVPN ARP/ND Extended Community in a MAC/IP Advertisment route
      indicates that the IP->MAC entry is not an "Immutable binding".

   * In a situation where a host (with an IP->MAC that is configured
      as Immutable binding in the attached PE) is allowed to move
      between PEs (that is, the associated MAC is non-static), PEs
      can receive multiple MAC/IP advertisement routes for the same
      IP->MAC. In such situations, MAC mobility procedures as in
      [RFC7432] dictate the reachability of the MAC.

   * Receiving multiple EVPN MAC/IP Advertisement routes with I=1
      for the same IP but different MAC is considered a
      misconfiguration.

* As an example, consider PE1, PE2 and PE3 are attached to the same BD. PE1 originates an EVPN MAC/IP Advertisement route for IP1->MAC1 with I=1; later on, PE2 also originates an EVPN MAC/ IP Advertisement route IP1->MAC1 with a higher sequence number and I=1. Then all the EVPN PEs attached to the same BD SHOULD retain their IP1->MAC1 ARP/ND binding but update MAC1's forwarding destination to PE2. If for some reason, PE3 originates an EVPN MAC/IP Advertisement route for IP1->MAC2 (even with a higher sequence number), then the EVPN PEs in the BD SHOULD NOT update their IP1->MAC1 ARP/ND bindings, since IP1 is bound to MAC1 (MAC2 SHOULD still be programmed in the layer-2 BDs). This is considered a misconfiguration in PE3.

* A PE originating an EVPN MAC/IP Advertisement route for IP1->MAC1 with I=1 MAY also originate the route with the Static bit set (in the MAC Mobility Extended Community). In such a case, the IP1->MAC1 binding is not only immutable but it cannot move as well. Even so, if an update for the same IP1->MAC1 immutable and static, is received from a different PE, one of the two routes will be selected, as in the [RFC7432] case where two MAC/IP routes with Static bit are received for the same MAC from different PEs.

* The use of the Flag I=1 assumes that a given IP is always bound to the same MAC address, and therefore the mobility procedures described in [I-D.ietf-bess-evpn-irb-extended-mobility] for "Host IP move to a new MAC" will not apply.

o The R, O and I Flags MUST be ignored if they are advertised along with an EVPN MAC/IP Advertisement route that does not contain an IP (IPv4 or IPv6) address.

4. Security Considerations

The same security considerations described in [RFC7432] apply to this document.

In addition, this document adds pieces of information that impact on the way ARP/ND entries are installed in ARP/ND and/or proxy-ARP/ND tables, and therefore the resolution protocols for IPv4 and IPv6 addresses. For instance, if a given IPv6->MAC binding is configured with the wrong R or O Flags (intentionally or not) on a given PE, the rest of the PEs attached to the same BD will install the wrong information for the IPv6->MAC. This will cause all the PEs in the BD to reply Neighbor Solicitations for the IPv6 with Neighbor Advertisement messages containing the wrong R and O Flags.

The I Flag, or Immutable ARP/ND Binding Flag, introduces a useful
security tool so that an operator makes sure a given IP address is
always bound to the same MAC and that information is distributed to
all the PEs attached to the same BD. ARP/ND spoofing attacks from
hosts injecting Gratuitous ARPs or unsolicited Neighbor Advertisement
messages for that IP address with a different MAC address will not
succeed to be programmed in ARP/ND and proxy-ARP/ND tables and
therefore will avoid attracting traffic to the spoofer.

5. IANA Considerations

This document request that the Name of the currently registered value
for Sub-Type 0x08 in the EVPN Extended Community Sub-Types registry
(https://www.iana.org/assignments/bgp-extended-communities/bgp-
extended-communities.xhtml#evpn) be changed to:

```
+----------+--------------------------+----------------+
| Sub-Type | Name                     | Reference      |
+----------+--------------------------+----------------+
| 0x08     | ARP/ND Extended Community | [this document] |
+----------+--------------------------+----------------+
```

This document also requests the creation of a registry called "ARP/ND
Extended Community Flags" where the following initial allocations are
made:

```
+---------------+------------------------------+----------------+
| Flag position | Name                         | Reference      |
+---------------+------------------------------+----------------+
| 0-3           | Unassigned                   | -              |
| 4             | Immutable ARP/ND Binding Flag | [this document] |
|               | (I)                          |                |
| 5             | Unassigned                   | -              |
| 6             | Override Flag (O)            | [this document] |
| 7             | Router Flag (R)              | [this document] |
+---------------+------------------------------+----------------+
```

The registration procedure for this registry is Standards Action.
This registry should be located in the Border Gateway Protocol (BGP)
Extended Communities general registry
(https://www.iana.org/assignments/bgp-extended-communities/bgp-
extended-communities.xhtml).

6. Acknowledgments

The authors would like to thank Ali Sajassi for his feedback.

7. References

7.1. Normative References

[RFC4861]   Narten, T., Nordmark, E., Simpson, W., and H. Soliman,
            "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861,
            DOI 10.17487/RFC4861, September 2007,
            <https://www.rfc-editor.org/info/rfc4861>.

[RFC7432]   Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A.,
            Uttaro, J., Drake, J., and W. Henderickx, "BGP MPLS-Based
            Ethernet VPN", RFC 7432, DOI 10.17487/RFC7432, February
            2015, <https://www.rfc-editor.org/info/rfc7432>.

[RFC2119]   Bradner, S., "Key words for use in RFCs to Indicate
            Requirement Levels", BCP 14, RFC 2119,
            DOI 10.17487/RFC2119, March 1997,
            <https://www.rfc-editor.org/info/rfc2119>.

[RFC8174]   Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC
            2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174,
            May 2017, <https://www.rfc-editor.org/info/rfc8174>.

7.2. Informative References

[I-D.ietf-bess-evpn-irb-extended-mobility]
            Malhotra, N., Sajassi, A., Pattekar, A., Lingala, A.,
            Rabadan, J., and J. Drake, "Extended Mobility Procedures
            for EVPN-IRB", draft-ietf-bess-evpn-irb-extended-
            mobility-03 (work in progress), May 2020.

Authors' Addresses

   Jorge Rabadan (editor)
   Nokia
   777 Middlefield Road
   Mountain View, CA 94043
   USA

   Email: jorge.rabadan@nokia.com

   Senthil Sathappan
   Nokia
   701 E. Middlefield Road
   Mountain View, CA 94043 USA

   Email: senthil.sathappan@nokia.com

Kiran Nagaraj
Nokia
701 E. Middlefield Road
Mountain View, CA 94043 USA

Email: kiran.nagaraj@nokia.com

Wen Lin
Juniper Networks

Email: wlin@juniper.net