BESS Workgroup
Internet Draft

Intended status: Standards Track

J. Rabadan, Ed.
S. Sathappan
K. Nagaraj
Nokia

W. Lin
Juniper

Expires: January 4, 2020 July 3, 2019

Propagation of ARP/ND Flags in EVPN
draft-ietf-bess-evpn-na-flags-04

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 reply locally (act as proxy) to IPv4 ARP requests and IPv6 Neighbor Solicitation messages and reduce/suppress the flooding produced by the Address Resolution procedure. 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 MAC/IP route advertisements. 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 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), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months

This Internet-Draft will expire on January 4, 2020.

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 reply locally (act as proxy) to IPv4 ARP requests and IPv6 Neighbor Solicitation messages and reduce/suppress the flooding produced by the Address Resolution procedure. 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 MAC/IP route advertisements. 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].

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.

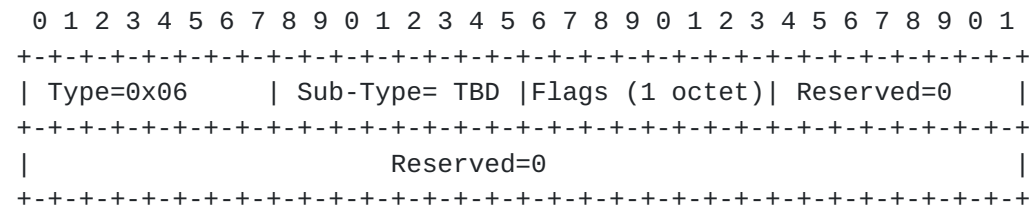
Proxy-ARP/ND: refers to a function on the EVPN PEs by which received

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.

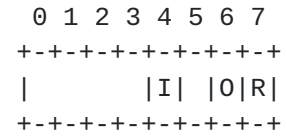
Familiarity with the terminology in [RFC7432] and [RFC4861] is expected.

2. The EVPN ARP/ND Extended Community

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



Flags field:



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 octet is defined as the "Router flag". When set, the R-bit 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-bit 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 is ignored when the extended community is advertised with a MAC/IP route for an IPv4->MAC pair.

O - Override Flag (corresponds to Bit 22 of the extended community)

Bit 6 of the Flags octet is defined as the "Override flag". An egress PE will normally advertise IPv6->MAC pairs with the O-bit set, and only when IPv6 "anycast" is enabled in the BD, the PE will send an IPv6->MAC pair with the O-bit = 0. The ingress PE will install the ND entry with the received O-bit and will use this information when replying to a Neighbor Solicitation for the IPv6 address. This flag is ignored when the extended community is advertised with a MAC/IP 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 octet is defined as the "Immutable ARP/ND binding flag". When set, the egress PE indicates that the IP->MAC pair sent in a MAC/IP route along with the extended community is a configured ARP/ND entry, and the IP address in the MAC/IP 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

An EVPN PE supporting a ND/ARP function and implementing the propagation of the ARP/ND Flags MUST follow this procedure:

a) Transmission of the EVPN ARP/ND Extended Community

A PE may learn the IPv6->MAC pair and its associated ND Flags in the management plane or by snooping Neighbor Advertisement messages coming from the CE. Either way, the PE sends a MAC/IP Advertisement route including the learned IPv6->MAC pair and MUST send the ARP/ND Extended Community carrying its associated "R" and "O" Flags.

If an IPv4->MAC or IPv6->MAC pair has been learned in the management plane (it has been configured) the corresponding MAC/IP Advertisement route SHOULD be sent along with an ARP/ND extended community with the flag I set.

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

b) Reception of the EVPN ARP/ND Extended Community

In addition to the procedures specified in [RFC7432] a PE receiving a MAC/IP Advertisement route containing an IPv6 address and the ND

Extended Community MUST add the R and O Flags to the ND entry for the IPv6->MAC entry and use that information in Neighbor Advertisements when replying to a Solicitation for the IPv6 address.

A PE that implements a proxy-ND function SHOULD have an administrative option to define the default Flag to be used in case no EVPN ND Extended Community is received for a given IPv6->MAC entry. A PE MUST ignore the received R and O Flags for a MAC/IP route that contains an IPv4 address.

A PE receiving a MAC/IP Advertisement route containing an IPv4 or IPv6 address and the I flag set, SHOULD install the IP->MAC entry in the ARP/ND table as "Immutable binding" entry.

In a situation where a host (with a IP->MAC configured as Immutable binding) 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 dictate the reachability of the MAC. Receiving multiple MAC/IP routes with I=1 for the same IP but different MAC is considered a misconfiguration.

For example, consider PE1, PE2 and PE3 are attached to the same BD. PE1 originates a MAC/IP route for IP1->MAC1 with I=1; later on, PE2 also originates a MAC/IP 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 a MAC/IP 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 a MAC/IP 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. Also, note that the use of the flag I=1 assumes that a given IP is always bound to the same MAC address, and therefore some of the mobility procedures described in [EXT-MOBILITY] will not apply.

The flags SHOULD be ignored if they are advertised along with a MAC/IP Advertisement route that does not contain an IP address.

4. Security Considerations

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

document.

5. IANA Considerations

This document requests the registration of a new EVPN Extended Community sub-type:

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 octet" where the following allocations are made:

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

The registration procedure for this registry is Standards Action.

6. References

6.1. 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 RFC2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

6.2. Informative References

[EXT-MOBILITY] Malhotra, N. et al., "Extended Mobility Procedures for EVPN-IRB", Work in Progress, draft-ietf-bess-evpn-irb-extended-mobility-01, June 2019.

7. Acknowledgments

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

Authors' Addresses

Jorge Rabadan (Editor)
Nokia
777 E. 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