ï»¿

Internet Engineering Task Force (IETF)
Request for Comments: 9047
Category: Standards Track
ISSN: 2070-1721

J. Rabadan, Ed.
S. Sathappan
K. Nagaraj
Nokia
W. Lin
Juniper

June 2021

Propagation of ARP/ND Flags in an Ethernet Virtual Private Network (EVPN)

Abstract

This document defines an Extended Community that is advertised along with an Ethernet Virtual Private Network (EVPN) Media Access Control (MAC) / IP Advertisement route and carries information relevant to the Address Resolution Protocol (ARP) / Neighbor Discovery (ND) resolution so that an EVPN Provider Edge (PE) implementing a proxy-ARP/ND function in broadcast domains (BDs) or an ARP/ND function on Integrated Routing and Bridging (IRB) interfaces can reply to ARP Requests or Neighbor Solicitation (NS) messages with the correct information.

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 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at https://www.rfc-editor.org/info/rfc9047.

Copyright Notice

Copyright (c) 2021 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 (https://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

Authors' Addresses

1. Introduction

An EVPN MAC/IP Advertisement route can optionally carry IPv4 or IPv6 addresses associated with a MAC address. Remote PE routers can use this information to populate their ARP or ND tables on IRB interfaces or their proxy-ARP/ND tables in BDs. 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 or 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 address and MAC address combination ND entry via EVPN (denoted by IPv6->MAC), the PE would not know if that particular IPv6->MAC pair belongs to a router or a host or if that address is an anycast address, as this information is not carried in the EVPN MAC/IP Advertisement routes.

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 (NA) messages with the correct flag information.

The flags are carried in the EVPN Address Resolution Protocol and Neighbor Discovery (ARP/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: Address Resolution Protocol

ND: Neighbor Discovery protocol, specified in [RFC4861]

PE: Provider Edge router

CE: Customer Edge router

IRB: Integrated Routing and Bridging interface

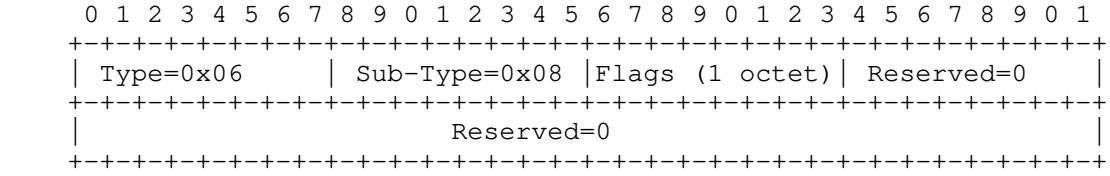
Proxy-ARP/ND: A function on the EVPN PEs by which received ARP Requests or NS messages are replied to locally by the PE, without the need to flood the requests to remote PEs in the BD. In order to reply to ARP Requests or NS messages, the PE does a lookup on an ARP/ND table, which is a collection of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

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

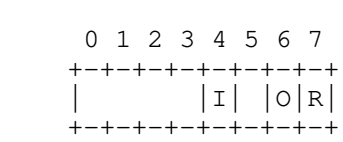
Familiarity with the terminology in [RFC4861] and [RFC7432] is expected.

2. The EVPN ARP/ND Extended Community

This document defines a transitive 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.



Flags field:



The following flags are defined in the Flags field, the 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 an IPv6 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 will the PE 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 always use this O flag value 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 that was sent in an EVPN MAC/IP Advertisement route (along with the Extended Community) is a configured ARP/ND entry. In this case, the IP address in the EVPN MAC/IP Advertisement route can only be bound together with the MAC address specified in the same route, and not with any other MAC addresses received in a different route without the I flag set.

Bits 0-3 and 5 are not assigned by this document. They MUST be set to zero and ignored on receipt.

The reserved fields are set to 0 and ignored by the receiver.

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 that supports the proxy-ARP/ND (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 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:

* Advertised MAC/IP Advertisement routes for IPv6->MAC entries MUST include one (and only one) 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.

* MAC/IP Advertisement routes for IPv4->MAC entries MAY include one 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.

* 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.

* This Extended Community does not change the procedures described in [RFC7432]. Specifically, the procedures for advertising the MAC Mobility Extended Community along with the MAC/IP Advertisement route are not changed.

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:

* Only one EVPN ARP/ND Extended Community is expected to be received along with an EVPN MAC/IP Advertisement route. If more than one ARP/ND Extended Community is received, the PE MUST consider only the first one on the list for processing purposes and MUST NOT propagate the rest of the ARP/ND Extended Communities.

* 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. Otherwise, they are processed as follows.

* R and O flag 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 flag values to the ND entry in the ND or proxy-ND table and propagate the value of the R and O flags from the ARP/ND Extended Community to the 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.

* 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 absence of the EVPN ARP/ND Extended Community in a MAC/IP Advertisement route indicates that the IP->MAC entry is not an "immutable binding".

- Receiving multiple EVPN MAC/IP Advertisement routes with the I flag set to 1 for the same IP but a different MAC address is considered a misconfiguration or a transient error condition. If this happens in the network, a PE receiving multiple routes (with the I flag set to 1 for the same IP and a different MAC address) SHOULD update the IP->MAC entry with the latest received information. Note that if a configured IP1->MAC1 changes to point to a new MAC address, i.e., IP1->MAC2, the EVPN MAC/IP Advertisement route for IP1->MAC1 will be withdrawn before the EVPN MAC/IP Advertisement route for IP1->MAC2 is advertised.

- A PE originating an EVPN MAC/IP Advertisement route for IP1->MAC1 with the I flag set to 1 MAY also originate the route with the "Sticky/static flag" 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 immutable and static IP1->MAC1 is received from a different PE, one of the two routes will be selected. This is analogous to the case described in Section 15.2 of [RFC7432] when two MAC/IP routes with the static flag set are received, and the PE likewise MUST alert the operator of such a situation.

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.

As an example of the use of the I flag, consider PE1, PE2, and PE3 attached to the same BD. PE1 originates an EVPN MAC/IP Advertisement route for IP1->MAC1 with the I flag set to 1 later on, PE2 also originates an EVPN MAC/IP Advertisement route IP1->MAC1 with a higher sequence number and the I flag set to 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. For some reason, if PE3 originates an EVPN MAC/IP Advertisement route for IP1->MAC2 with the I flag set to 0 (even with a higher sequence number), then the EVPN PEs in the BD will 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.

When the I flag is set to 1, a given IP is assumed to be always bound to the same MAC address; therefore, the mobility procedures described in [EXTENDED-MOBILITY] for "Host IP move to a new MAC" will not apply.

4. Security Considerations

The same security considerations described in [RFC7432] apply to this document. In general, it is worth noting that the use of proxy-ARP/ND in EVPN BDs may add some security risks. Attackers can make use of ARP/ND messages to create state in all the PEs attached to the same BD as the attacker and exhaust resources in those PEs. Therefore, additional security mechanisms may be needed. Some examples of such additional security mechanisms are limiting the number of proxy-ARP/ND entries per BD and/or per port or closely monitoring the rate at which hosts create dynamic proxy-ARP/ND entries.

In addition, this document adds pieces of information that impact the

way ARP/ND entries are installed in ARP/ND and/or proxy-ARP/ND tables and, therefore, impacts 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 to Neighbor Solicitations for the IPv6 with NA messages containing the wrong R and O flags. For example, as specified in [RFC4861], the receiver of an NA message with O not set will not update its existing cache entry for the IP->MAC; hence, the communication between the owner of the IP address and the receiver of the NA message with the wrong O flag will fail. Similarly, the receiver of an NA message with the wrong R flag may update its Default Router List by incorrectly adding or removing an entry, which could, for example, lead to sending traffic to a node that is not a router, causing the traffic to be dropped.

The I flag, or Immutable ARP/ND Binding flag, is a useful security tool, allowing an operator to ensure 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, in which a malicious host injects Gratuitous ARPs or unsolicited NAs for that IP address with a different MAC address, will not succeed in programming the ARP/ND and proxy-ARP/ND tables and therefore the spoofer will not receive the traffic.

5. IANA Considerations

IANA has changed the name for Sub-Type Value 0x08 in the "EVPN Extended Community Sub-Types" registry [IANA-BGP-EXT-COMM] to the following:

Sub-Type Value	Name	Reference
0x08	ARP/ND Extended Community	RFC 9047

Table 1: Updated Value in the "EVPN Extended Community Sub-Types" Registry

IANA has created the "ARP/ND Extended Community Flags" registry, where the following initial allocations have been made:

Flag Position	Name	Reference
0-3	Unassigned	
4	Immutable ARP/ND Binding Flag (I)	RFC 9047
5	Unassigned	
6	Override Flag (O)	RFC 9047
7	Router Flag (R)	RFC 9047

Table 2: Initial Values of the "ARP/ND Extended Community Flags" Registry

The registration policy for this registry is Standards Action [RFC8126]. This registry is located in the "Border Gateway Protocol (BGP) Extended Communities" registry [IANA-BGP-EXT-COMM].

Note that the flag position 5 is left unassigned and not used in this specification since it was previously requested by [EVPN-IP-MAC-PROXY].

6. References

6.1. Normative References

[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>.

[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>.

[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>.

6.2. Informative References

[EVPN-IP-MAC-PROXY] Bickhart, R., Lin, W., Drake, J., Rabadan, J., and A. Lo, "Proxy IP->MAC Advertisement in EVPNs", Work in Progress, Internet-Draft, draft-rbickhart-evpn-ip-mac-proxy-adv-01, 24 January 2020, <https://tools.ietf.org/html/draft-rbickhart-evpn-ip-mac-proxy-adv-01>.

[EXTENDED-MOBILITY] Malhotra, N., Ed., Sajassi, A., Pattekar, A., Lingala, A., Rabadan, J., and J. Drake, "Extended Mobility Procedures for EVPN-IRB", Work in Progress, Internet-Draft, draft-ietf-bess-evpn-irb-extended-mobility-05, 15 March 2021, <https://tools.ietf.org/html/draft-ietf-bess-evpn-irb-extended-mobility-05>.

[IANA-BGP-EXT-COMM] IANA, "Border Gateway Protocol (BGP) Extended Communities", <https://www.iana.org/assignments/bgp-extended-communities>.

[RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <https://www.rfc-editor.org/info/rfc8126>.

Acknowledgments

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

Authors' Addresses

Jorge Rabadan (editor)
Nokia
777 Middlefield Road
Mountain View, CA 94043
United States of America

Email: jorge.rabadan@nokia.com

Senthil Sathappan
Nokia
701 E. Middlefield Road
Mountain View, CA 94043
United States of America

Email: senthil.sathappan@nokia.com

Kiran Nagaraj
Nokia
701 E. Middlefield Road
Mountain View, CA 94043
United States of America

Email: kiran.nagaraj@nokia.com

Wen Lin
Juniper Networks

Email: wlin@juniper.net