BESS Working Group
Internet-Draft
Intended status: Standards Track
Expires: July 16, 2017

S. Kumar
D. Kakrania
V. Duna
Juniper Networks
January 12, 2017

EVPN ACCESS SECURITY
draft-surajk-evpn-access-security-00

Abstract

The draft defines a new BGP EVPN route message for syncing DHCP packet contents as well as snoop entry among PEs in an Ethernet Segment (ES). The snoop entry is required to implement Dynamic ARP inspection (DAI), IP Source Guard (IPSG/IPSGv6) and IPv6 Neighbor Discovery Inspection (NDI) access security features.

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 July 16, 2017.

Copyright Notice

Table of Contents

1. Introduction

In EVPN solution where a CE is connected to multiple PEs in All-Active redundancy mode then to support Dynamic ARP inspection (DAI), IP Source Guard (IPSG/IPSGv6) and IPv6 Neighbor Discovery Inspection (NDI) access security features, each PE in the ES needs to build an identical snoop database. This requires exchanging DHCP packet relevant contents as well as complete snoop entry among PEs in the ES. The draft defines a new BGP EVPN route for this.

2. Terminology

CE: Customer Edge device, e.g., a host, router, or switch.

PE: Provider Edge device e.g switch or router.

EVI: An EVPN instance spanning the Provider Edge (PE) devices participating in that EVPN.

Ethernet Segment (ES): When a customer site (device or network) is connected to one or more PEs via a set of Ethernet links, then that set of links is referred to as an 'Ethernet segment'.

Ethernet Segment Identifier (ESI): A unique non-zero identifier that identifies an Ethernet segment is called an 'Ethernet Segment Identifier'

Ethernet Tag ID (ETAG ID): An Ethernet tag identifies a particular broadcast domain, e.g., a VLAN. An EVPN instance consists of one or more broadcast domains.

All-Active Redundancy Mode: When all PEs attached to an Ethernet segment are allowed to forward known unicast traffic to/from that Ethernet segment for a given VLAN, then the Ethernet segment is defined to be operating in All-Active redundancy mode.

3. Access Security Features

3.1. DHCP snooping

DHCP snooping enables the switch (PE), to intercept DHCP messages exchanged between untrusted host (DHCP client) and trusted DHCP server and build an entry for untrusted host in snooping database. A switch builds a DHCP snooping entry by extracting relevant information from snooped DHCP packets. Similarly, a switch builds a DHCPv6 snooping entry by extracting relevant information from snooped DHCPv6 packets. The snoop entry holds the following information

1- Untrusted host MAC address (mac)

2- Untrusted host IP address (ip/ip6)

3- The interface(port) on which untrusted host is connected (intf)

4- Vlan in which untrusted host resides (vlan)

The entry [mac, ip, intf, vlan] is used for DAI, IPSGv4 features Similarly, [mac, ip6, intf, vlan] is used for IPSGv6 and NDI features.

3.2. Dynamic ARP inspection (DAI)

Dynamic ARP inspection (DAI) protects switching devices against ARP spoofing.

DAI inspects Address Resolution Protocol (ARP) packets on the LAN and uses the information in the DHCP snooping database on the switch to validate ARP packets and to protect against ARP spoofing (also known as ARP poisoning or ARP cache poisoning). ARP requests and replies are compared against entries in the DHCP snooping database, and filtering decisions are made based on the results of those

comparisons. When an attacker tries to use a forged ARP packet to spoof an address, the switch checks the sender IP and MAC source addresses in a ARP packet sent from a host attached to an untrusted access interface on the switch. The switch searches an entry [mac, ipv4, intf, vlan] in the snooping database. If the entry is not found in DHCP snooping database, the packet is dropped.

3.3. IP Source Guard (IPSGv4/IPSGv6)

Ethernet LAN switches are vulnerable to attacks that involve spoofing (forging) of source IP addresses or source MAC addresses. These spoofed packets are sent from hosts connected to untrusted access interfaces on the switch.

IP source guard checks the IP source address and MAC source address in a packet sent from a host attached to an untrusted access interface on the switch. The switch searches for the entry [mac, ip/ipv6, intf, vlan] in the snooping database. If the entry is not found in DHCP snooping database, the packet is dropped.

3.4. IPv6 Neighbor Discovery Inspection (NDI)

IPv6 Neighbor Discovery Inspection protects switching devices against ND spoofing.

NDI inspects Neighbor Discovery packets on the LAN and uses the information in the DHCPv6 snooping database on the switch to validate ND packets and to protect against ND spoofing. ND packets are compared against entries in the DHCPv6 snooping database, and

filtering decisions are made based on the results of those comparisons. When an attacker tries to use a forged ND packet to spoof an IPv6 address, the switch checks the IPv6 source address and MAC source address in a ND packet sent from a host attached to an untrusted access interface on the switch. The switch searches the entry [mac, ip6, intf, vlan] in the DHCPv6 snooping database. If the entry is not found in database, the packet is dropped.

4. DHCP Snooping Database

A snoop database is a place holder of snoop entries. A DHCPv4 snoop database contains DHCPv4 snoop entries. Similarly, a DHCPv6 snoop database contains DHCPv6 entries.

A Switch (PE) does not need the complete DHCP packet to build snooping entry. The PE needs some relevant DHCP packet contents as mentioned in section Section 6.1 to build a snoop entry.

4.1. CE Connected to Single PE

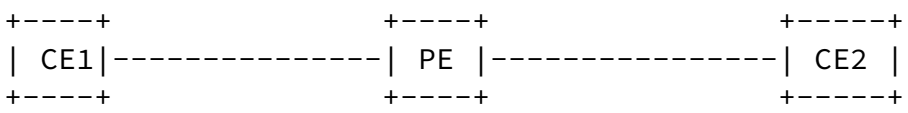


Figure 1

The basic process of DHCP snooping database building consists of the following steps. These steps are mentioned here for better understanding of the document. The scope of document is not to explain complete snooping mechanism.

1. The host (CE1) sends a DHCPDISCOVER packet to request an IP address.

2. The PE forwards the packet to the DHCP server (CE2).

3. The CE2 sends a DHCPOFFER packet to offer an address. If the DHCPOFFER packet is from a trusted interface, the switching device forwards the packet to the DHCP client.

4. The CE1 sends a DHCPREQUEST packet to accept the IP address. The PE adds an [mac, ip, intf, vlan] entry to the database. The entry is considered a placeholder until a DHCPACK packet is received from the server. Until then, the IP address could still be assigned to some other host.

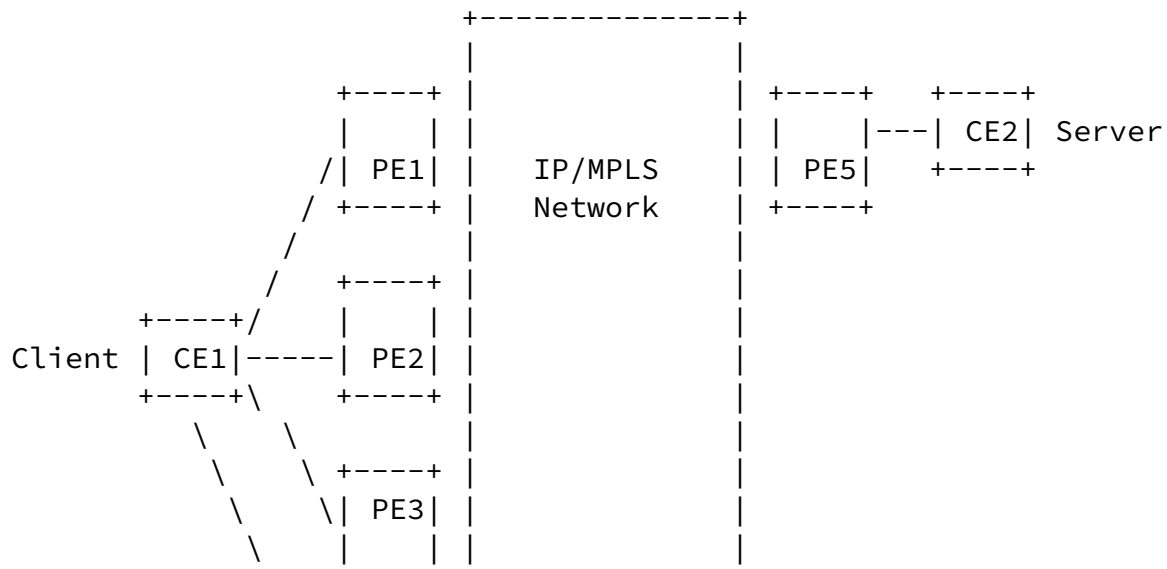
5. The CE2 sends a DHCPACK packet to assign the IP address or a DHCPNAK packet to deny the address request.

6. The PE updates the DHCP snooping database according to the type of packet received.

7. If the switching device receives a DHCPACK packet, it updates lease information for the [mac, ip, intf, vlan] in its database. The entry is deleted upon expiration of lease time.

8. If the PE receives a DHCPNAK packet, it deletes the placeholder.

4.2. CE Connected to Multiple PEs



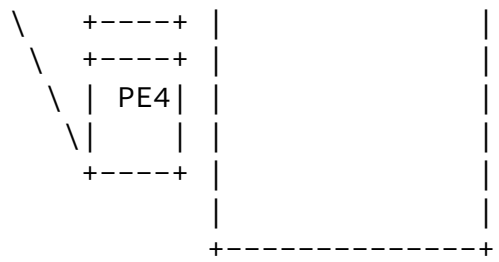


Figure 2

In Figure 2, PE1, PE2, PE3 and PE4 are in All-Active Redundancy Mode in an Ethernet Segment. Each PE advertises BGP Ethernet Segment (ES) route for Redundancy group discovery and also for Designated Forwarder (DF) election. Each PE router connected to an Ethernet Segment, advertises a BGP Ethernet Segment (ES) route that consists of an ESI and ES-Import extended community.

In the above Figure 2, PE1, PE2, PE3 and PE4 have the same ESI value (say ES1). PE1 advertises its ESI value in the Ethernet Segment Route with ES-Import community set to ES1. PE2, PE3, PE4 and PE5 will receive that route but PE2, PE3, PE4 will import this route, since they have a matching ESI value. PE5 will not import this route since it does not have matching ESI. This ensures PE2, PE3, PE4 knows that PE1 is connected to the same CE1 host. The process is repeated for each PE in the ES. Each PEs in the ES comes to know about all other PEs connected to same CE1 in the same ES. The DF election in an ES is done as specified in [RFC7432] section 8.5.

In Figure 2, to build an identical snoop database on each PEs in the ES, each PE needs to extract relevant information from DHCP packets exchanged between Client (CE1) and Server (CE2). But here problem is that all DHCP packets do not go through the same PE. For an example DHCP REQUEST can go through one of the PE say (PE1) and DHCP ACK from

server can go through some other PE say (PE2). Since neither PE1 nor PE2 gets all relevant information of DHCP REQUEST and ACK packets, PE1/PE2 cannot build snooping database.

5. Solution

The draft proposes the two possible solutions for snoop entry [mac, ip, ESI, ETAG ID] creation and synchronization among PEs in an ES in

the All-Active Redundancy Mode. Specific realizations and implementation details (state machines or algorithms, etc.) of below solutions are out of the scope of this document.

5.1. Centralized Mode

The PE acting as DF must be responsible for building the snoop entry and transporting it to all non-DF PE in the ES. The DF PE must also be responsible for withdrawing the entry locally and as well as from all other non-DF remote PEs in the ES. The non-DF PE must neither create nor release the snooping entry by itself. The creation and release of entry is controlled by DF PE in the ES. The PE must use the proposed EVPN DHCP Snoop Advertisement route for exchanging DHCP packet contents as well as complete bindings with other PE in the ES.

When a DF PE receives a DHCP packet from CE, it consumes it locally. When a non-DF PE receives a DHCP packet it extracts relevant information from the packet and transport this information to DF PE using newly proposed EVPN DHCP Snoop Advertisement route. The non-DF PE must not consume the DHCP packet locally.

The DF PE eventually receives all the information that are required to build snooping entry for the untrusted host. The DF PE builds [mac, ip, ESI, ETAG ID] entry and advertise this to all the non-DF PEs in the ES. When the DF PE releases the entry locally then it advertises the withdrawal of the entry to all the non-DF PEs is the ES.

5.2. Distributed Mode

In this mode, PE (DF or non-DF) must be responsible for building and releasing entry independently. The DF PE must be responsible for syncing snoop entry when a new non-DF PE joins the same redundancy group. Unlike Centralized mode, in this mode each PE must release the snoop entry upon expiration of lease time.

When a PE receives a DHCP packet it extracts relevant information from the packet and transport this information to all other PEs in the ES using newly proposed EVPN DHCP Snoop Advertisement route message. Each PE in the ES eventually receives all the information

that are required to build snooping entry for the host. Each PE in

the ES builds [mac, ip, ESI, ETAG ID] snoop entry. Each PE in the ES also receives the relevant DHCP release packet content to release the entry independently.

6. DHCP Snoop Advertisement route

The [RFC7432] defines a new BGP Network Layer Reachability Information (NLRI) called the EVPN NLRI. The format of the EVPN NLRI is as follows:

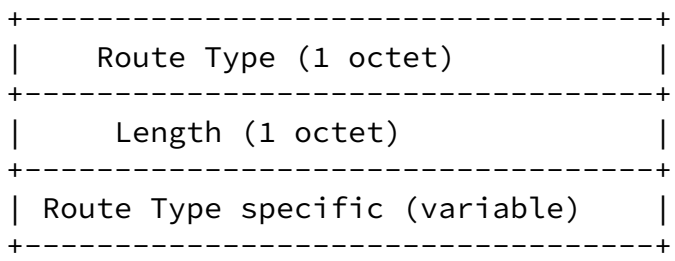


Figure 3

The EVPN NLRI is carried in BGP [RFC4271] using BGP Multiprotocol Extensions [RFC4760] with an Address Family Identifier (AFI) of 25 (L2VPN) and a Subsequent Address Family Identifier (SAFI) of 70 (EVPN). The NLRI field in the MP_REACH_NLRI/MP_UNREACH_NLRI attribute contains the EVPN NLRI (encoded as specified above).

This [RFC7432] defines the following Route Types:

+ 1 - Ethernet Auto-Discovery (A-D) route

+ 2 - MAC/IP Advertisement route

+ 3 - Inclusive Multicast Ethernet Tag rout

+ 4 - Ethernet Segment route

This draft defines a new route (DHCP Snoop Advertisement route) The PE uses this route message for exchanging DHCP packet contents as well as complete bindings with other PE.

+ 5 - DHCP Snoop Advertisement route

An DHCP Snoop Advertisement route type specific EVPN NLRI consists of the following:

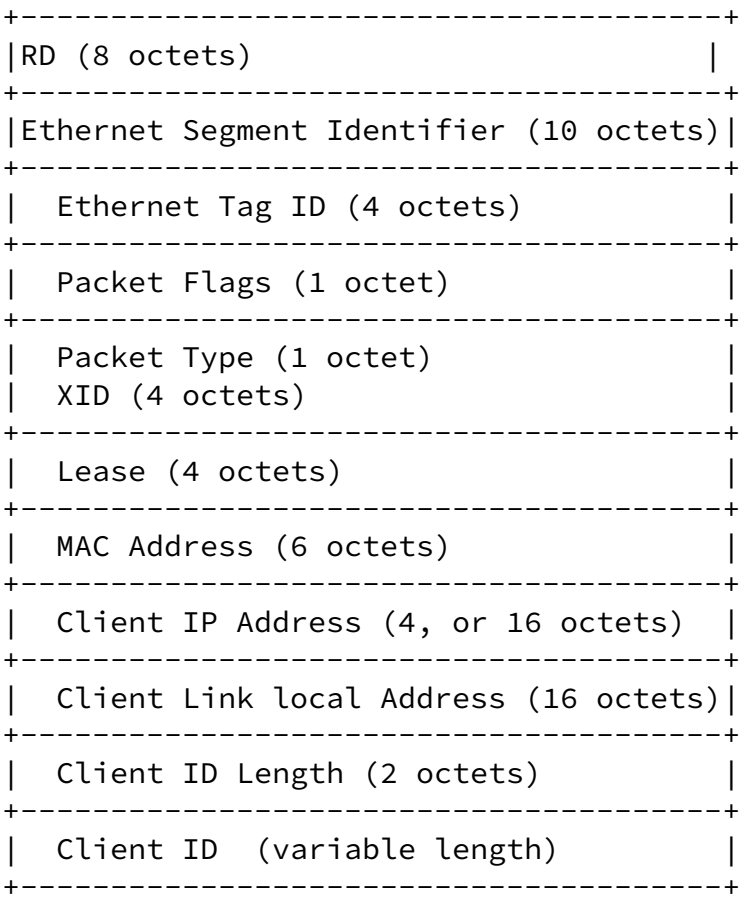


Figure 4

6.1. Constructing DHCP Snoop Advertisement route

Packet Flags:

Packet Flags is one-byte value in the message. The flags bit is as defined below:

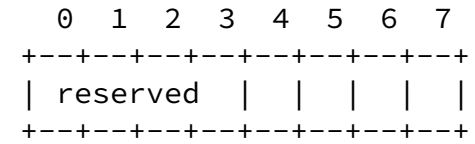


Figure 5

The least significant bit, bit 7 indicates DHCPv6 contents or DHCPv6 snoop entry. This bit is not set for DHCPv4 contents or DHCPv4 snoop entry.

The second least significant bit, bit 6 indicates DHCPv6 Rapid commit

option is enabled.

The third least significant bit, bit 5 indicates DHCPv6 Reply is a NAK. DHCPv6 NAK is extracted from Status Code option of reply packet.

The fourth least significant bit, bit 4 indicates DHCP Snoop Advertisement route contains the snoop entry. If this bit is not set this indicate the DHCP Snoop Advertisement route contains DHCP packet contents.

The lease significant bits 3, 2,1 and 0 are reserved.

Packet Type:

Type of packet, e.g DHCPDISCOVER, DHCPOFFER. This is valid only when bit 4 is not set in Packet Flags.

XID:

Transaction ID, a random number chosen by the client, used by the client and server to associate messages and responses between a client and a server. This is used by the client to match incoming DHCP messages with pending requests. This is valid only when bit 4 is not set in Packet Flags.

Lease:

The period of time IP address is allocated to a client by server.

Mac Address:

Untrusted Client's source mac address

Client IP Address:

Untrusted Client's source IP address. This can be IPv4 or IPv6 based on the Packet Type.

Client Link Local Address:

Untrusted Client's Link Local IPv6 address. This is valid only when bit 7 in Packet Flags is set.

Client ID Length:

Length of client ID in octets. This is valid only when bit 7 in Packet Flags is set.

Client ID:

The Client Identifier option is used to carry a DUID. Each DHCP client and server has a DUID. The DUID is DHCP Unique Identifier. This may be used as key to identify the snoop entry. This filed is valid only when bit 7 in Packet Flags is set.

The Route Distinguisher (RD) SHOULD be a Type 1 RD [RFC4364]. The value field comprises an IP address of the PE (typically, the loopback address) followed by a number unique to the PE

The Ethernet Tag ID:

CE's Ethernet tag value (e.g., CE VLAN ID)

7. Error Handling

The snoop database among PEs in a ES may go out of sync due to some PE going unreachable in the ES. The solution of this problem is out of scope of this draft.

In Centralized mode, If DF PE goes down during the process of building snoop entry, it is possible that the untrusted host gets IP address but no snoop entry gets created on any of the PEs in the ES

8. Security Considerations

Same security considerations as [RFC7432].

9. References

[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, <http://www.rfc-editor.org/info/rfc7432>.

[RFC2131] Droms, R., "Dynamic Host Configuration Protocol", RFC 2131, DOI 10.17487/RFC2131, March 1997, <http://www.rfc-editor.org/info/rfc2131>.

[RFC3315] Droms, R., Ed., Bound, J., Volz, B., Lemon, T., Perkins, C., and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, DOI 10.17487/RFC3315, July 2003, <http://www.rfc-editor.org/info/rfc3315>.

[RFC7209] Sajassi, A., Aggarwal, R., Uttaro, J., Bitar, N., Henderickx, W., and A. Isaac, "Requirements for Ethernet VPN (EVPN)", RFC 7209, DOI 10.17487/RFC7209, May 2014, <http://www.rfc-editor.org/info/rfc7209>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <http://www.rfc-editor.org/info/rfc4271>.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, DOI 10.17487/RFC4760, January 2007, <http://www.rfc-editor.org/info/rfc4760>.

[EVPN-IGMP] Sajassi, A., "https://tools.ietf.org/html/draft-sajassi-bess-evpn-igmp-mld-proxy-01", October 2016.

Authors' Addresses

Suraj Kumar
Juniper Networks
Elnath, Juniper Networks
Bangalore, Karnataka 560036
India

EMail: surajk@juniper.net

Deepak Kakrania
Juniper Networks
Elnath, Juniper Networks
Bangalore, Karnataka 560036
India

EMail: dkakrania@juniper.net

Vijay Kumar Duna
Juniper Networks
Elnath, Juniper Networks
Bangalore, Karnataka 560036
India

EMail: dvijay@juniper.net