

BESS Working Group  
Internet Draft  
Category: Standards Track

K. Patel  
Arrcus  
A. Sajassi  
Cisco  
J. Drake  
Z. Zhang  
Juniper Networks  
W. Henderickx  
Nokia

Expires: May 22, 2019

October 22, 2018

**Virtual Hub-and-Spoke in BGP EVPNs**  
**draft-keyupate-bess-evpn-virtual-hub-01**

Abstract

Ethernet Virtual Private Network (EVPN) solution is becoming pervasive for Network Virtualization Overlay (NVO) services in data center (DC) applications and as the next generation virtual private LAN services in service provider (SP) applications.

The use of host IP default route and host unknown MAC route within a DC is well understood in order to ensure that leaf nodes within a DC only learn and store host MAC and IP addresses for that DC. All other host MAC and IP addresses from remote DCs are learned and stored in DC GW nodes thus alleviating leaf nodes from learning host MAC and IP addresses from the remote DCs.

This draft further optimizes the MAC and IP address learning at the leaf nodes such that a leaf node within a DC only needs to learn and store MAC and IP addresses associated with the sites directly connected to it. A leaf node does not need to learn and store MAC and IP addresses from any other leaf nodes thus reducing the number of learned MACs and IP addresses per EVI substantially.

The modifications provided by this draft updates and extends [RFC7024](#) for BGP EVPN Address Family.

Status of this Memo

This Internet-Draft is submitted to IETF 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 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."

The list of current Internet-Drafts can be accessed at  
<http://www.ietf.org/lid-abstracts.html>

The list of Internet-Draft Shadow Directories can be accessed at  
<http://www.ietf.org/shadow.html>

## Copyright and License Notice

Copyright (c) 2015 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 (<http://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

|                      |                                                              |                    |
|----------------------|--------------------------------------------------------------|--------------------|
| <a href="#">1.</a>   | Introduction . . . . .                                       | <a href="#">4</a>  |
| <a href="#">2.</a>   | Requirements Language . . . . .                              | <a href="#">5</a>  |
| <a href="#">3.</a>   | Terminology . . . . .                                        | <a href="#">5</a>  |
| <a href="#">4.</a>   | Routing Information Exchange for EVPN routes . . . . .       | <a href="#">5</a>  |
| <a href="#">5.</a>   | EVPN unknown MAC route . . . . .                             | <a href="#">6</a>  |
| <a href="#">5.1.</a> | Originating EVPN Unknown MAC Route by a V-Hub . . . . .      | <a href="#">6</a>  |
| <a href="#">5.2.</a> | Processing VPN-MAC EVPN unknown Route by a V-SPOKE . . . . . | <a href="#">6</a>  |
| <a href="#">5.3.</a> | Aliasing . . . . .                                           | <a href="#">7</a>  |
| <a href="#">5.4.</a> | Split-Horizon & Mass Withdraw . . . . .                      | <a href="#">8</a>  |
| <a href="#">6.</a>   | Forwarding Considerations . . . . .                          | <a href="#">8</a>  |
| <a href="#">6.1.</a> | IP-only Forwarding . . . . .                                 | <a href="#">8</a>  |
| <a href="#">6.2.</a> | MAC-only Forwarding - Bridging . . . . .                     | <a href="#">8</a>  |
| <a href="#">6.3.</a> | MAC and IP Forwarding - IRB . . . . .                        | <a href="#">8</a>  |
| <a href="#">7.</a>   | Handling of Broadcast and Multicast traffic . . . . .        | <a href="#">9</a>  |
| <a href="#">7.1.</a> | Split Horizon . . . . .                                      | <a href="#">10</a> |
| <a href="#">7.2.</a> | Route Advertisement . . . . .                                | <a href="#">10</a> |



|                        |                                                                       |                    |
|------------------------|-----------------------------------------------------------------------|--------------------|
| <a href="#">7.3.</a>   | <a href="#">Designated Forwarder in a Cluster . . . . .</a>           | <a href="#">11</a> |
| <a href="#">7.4.</a>   | <a href="#">Traffic Forwarding Rules . . . . .</a>                    | <a href="#">11</a> |
| <a href="#">7.4.1.</a> | <a href="#">Traffic from Local ACs . . . . .</a>                      | <a href="#">12</a> |
| <a href="#">7.4.2.</a> | <a href="#">Traffic Received by a V-hub from Another PE . . . . .</a> | <a href="#">12</a> |
| <a href="#">7.4.3.</a> | <a href="#">Traffic received by a V-spoke from a V-hub . . . . .</a>  | <a href="#">12</a> |
| <a href="#">7.5.</a>   | <a href="#">Multi-homing support . . . . .</a>                        | <a href="#">12</a> |
| <a href="#">7.5.1</a>  | <a href="#">Domain-wide Common Block (DCB) Label . . . . .</a>        | <a href="#">13</a> |
| <a href="#">7.5.2</a>  | <a href="#">Local Bias . . . . .</a>                                  | <a href="#">13</a> |
| <a href="#">7.6.</a>   | <a href="#">Direct V-spoke to V-spoke traffic . . . . .</a>           | <a href="#">13</a> |
| <a href="#">8.</a>     | <a href="#">ARP/ND Suppression . . . . .</a>                          | <a href="#">13</a> |
| <a href="#">9.</a>     | <a href="#">IANA Considerations . . . . .</a>                         | <a href="#">14</a> |
| <a href="#">10.</a>    | <a href="#">Security Considerations . . . . .</a>                     | <a href="#">14</a> |
| <a href="#">11.</a>    | <a href="#">Acknowledgements . . . . .</a>                            | <a href="#">14</a> |
| <a href="#">12.</a>    | <a href="#">Change Log . . . . .</a>                                  | <a href="#">15</a> |
| <a href="#">13.</a>    | <a href="#">References . . . . .</a>                                  | <a href="#">15</a> |
| <a href="#">13.1.</a>  | <a href="#">Normative References . . . . .</a>                        | <a href="#">15</a> |
| <a href="#">13.2.</a>  | <a href="#">Informative References . . . . .</a>                      | <a href="#">15</a> |
| <a href="#">14.</a>    | <a href="#">Authors' Addresses . . . . .</a>                          | <a href="#">15</a> |



## **1. Introduction**

Ethernet Virtual Private Network (EVPN) solution is becoming pervasive for Network Virtualization Overlay (NVO) services in data center (DC) applications and as the next generation virtual private LAN services in service provider (SP) applications.

With EVPN, providing any-to-any connectivity among sites of a given EVPN Instance (EVI) would require each Provider Edge (PE) router connected to one or more of these sites to hold all the host MAC and IP addresses for that EVI. The use of host IP default route and host unknown MAC route within a DC is well understood in order to alleviate the learning of host MAC and IP addresses to only leaf nodes (PEs) within that DC. All other host MAC and IP addresses from remote DCs are learned and stored in DC GW nodes thus alleviating leaf nodes from learning host MAC and IP addresses from the remote DCs.

This draft further optimizes the MAC and IP address learning at the leaf nodes such that a leaf node within a DC only needs to learn and store MAC and IP addresses associated with the sites directly connected to it. A leaf node does not need to learn and store MAC and IP addresses from any other leaf nodes thus reducing the number of learned MACs and IP addresses per EVI substantially.

[RFC7024] provides rules for Hub and Spoke VPNs for BGP L3VPNs. This draft updates and extends [[RFC7024](#)] for BGP EVPN Address Family. This draft provides rules for Originating and Processing of the EVPN host unknown MAC route and host default IP route by EVPN Virtual Hub (V-HUB). This draft also provides rules for the handling of the BUM traffic in Hub and Spoke EVPNs and handling of ARP suppression.

The leaf nodes and DC GW nodes in a data center are referred to as Virtual Spokes (V-spokes) and Virtual Hubs (V-hubs) respectively. A set of V-spoke can be associated with one or more V-hubs. If a V-spokes is associated with more than one V-hubs, then it can load balanced traffic among these V-hubs. Different V-spokes can be associated with different sets of V-hubs such that at one extreme each V-spoke can have a different V-hub set although this may not be desirable and a more typical scenario may be to associate a set of V-spokes to a set of V-hubs - e.g., topology for a DC POD where a set of V-spokes are associated with a set of spine nodes or DC GW nodes.

In order to avoid repeating many of the materials covered in [[RFC7024](#)], this draft is written as a delta document with its sections organized to follow those of that RFC with only delta description pertinent to EVPN operation in each section. Therefore, it is assumed that the readers are very familiar with [[RFC7024](#)] and



EVPN.

## **2. Requirements Language**

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" are to be interpreted as described in [[RFC2119](#)] only when they appear in all upper case. They may also appear in lower or mixed case as English words, without any normative meaning.

## **3. Terminology**

ARP: Address Resolution Protocol  
BEB: Backbone Edge Bridge  
B-MAC: Backbone MAC Address  
CE: Customer Edge  
C-MAC: Customer/Client MAC Address  
ES: Ethernet Segment  
ESI: Ethernet Segment Identifier  
IRB: Integrated Routing and Bridging  
LSP: Label Switched Path  
MP2MP: Multipoint to Multipoint  
MP2P: Multipoint to Point  
ND: Neighbor Discovery  
NA: Neighbor Advertisement  
P2MP: Point to Multipoint  
P2P: Point to Point  
PE: Provider Edge  
EVPN: Ethernet VPN  
EVI: EVPN Instance  
RT: Route Target

Single-Active Redundancy Mode: When only a single PE, among a group of PEs attached to an Ethernet segment, is allowed to forward traffic to/from that Ethernet Segment, then the Ethernet segment is defined to be operating in Single-Active redundancy mode.

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

## **4. Routing Information Exchange for EVPN routes**

[RFC7432] defines multiple Route Types NLRI along with procedures for





advertisements and processing of these routes. Some of these procedures are impacted as the result of hub-and-spoke architecture. The routing information exchange among the hub, spoke, and vanilla PE's are subject to the same rules as described in [section 3 of \[RFC7024\]](#). Furthermore, if there are any changes to the EVPN route advertisements and processing from that of [\[RFC7432\]](#), they are described below.

## **5. EVPN unknown MAC route**

[Section 3 of \[RFC7024\]](#) talks about how a V-hub of a given VPN must export a VPN-IP default route for that VPN and this route must be exported to only the V-spokes of that VPN associated with that V-hub. [DCI-EVPN] defines the notion of the unknown MAC route for an EVI which is analogous to a VPN-IP default route for a VPN. This unknown MAC route is exported by a V-hub to its associated V-spokes. If multiple V-hubs are associated with a set of V-spokes, then each V-hub advertises it with a distinct RD when originating this route. If a V-spoke imports several of these unknown MAC routes and they all have the same preference, then traffic from the V-spoke to other sites of that EVI would be load balanced among the V-hubs.

### **5.1. Originating EVPN Unknown MAC Route by a V-Hub**

[Section 7.3](#) of the [\[RFC7024\]](#) defines procedures for originating a VPN-IP default route for a VPN. The same procedures apply when a V-hub wants to originate EVPN unknown MAC route for a given EVI. The V-hub MUST announce unknown MAC route using the MAC/IP advertisement route along with the Default Gateway extended community as defined in [section 10.1](#) of the [\[RFC7432\]](#).

### **5.2. Processing VPN-MAC EVPN unknown Route by a V-SPOKE**

Within a given EVPN, a V-spoke MUST import all the unknown MAC routes unless the route-target mismatch happens. The processing of the received VPN-MAC EVPN default route follows the rules explained in the [section 3](#) of the [\[RFC7024\]](#). The unknown MAC route MUST be installed according to the rules of MAC/IP Advertisement route installation rules in [section 9.2.2 of \[RFC7024\]](#).

In absence of any more specific VPN-MAC EVPN routes, V-spokes installing the unknown MAC route MUST use the route when performing ARP proxy. This behavior would allow V-Spokes to forward the traffic towards V-Hub.



### 5.3. Aliasing

[RFC7432] describes the concept and procedures for Aliasing where a station is multi-homed to multiple PEs operating in an All-Active redundancy mode, it is possible that only a single PE learns a set of MAC addresses associated with traffic transmitted by the station.

[RFC7432] describes the concepts and procedures for Aliasing, which occurs when a CE is multi-homed to multiple PE nodes, operating in all-active redundancy mode, but not all of the PEs learn the CE's set of MAC addresses. This leads to a situation where remote PEs receive MAC advertisement routes, for these addresses, from a single NVE even though multiple NVEs are connected to the multi-homed station. As a result, the remote NVEs are not able to effectively load-balance traffic among the NVEs connected to the multi-homed Ethernet segment.

To alleviate this issue, EVPN introduces the concept of Aliasing. This refers to the ability of a PE to signal that it has reachability to a given locally attached Ethernet segment, even when it has learnt no MAC addresses from that segment. The Ethernet A-D per-EVI route is used to that end. Remote PEs which receive MAC advertisement routes with non-zero ESI SHOULD consider the MAC address as reachable via all NVEs that advertise reachability to the relevant Segment using Ethernet A-D routes with the same ESI and with the Single-Active flag reset.

This procedure is impacted for virtual hub-and-spoke topology because a given V-spoke does not receive any MAC/IP advertisements from remote V-spokes; therefore, there is no point in propagating Ethernet A-D per-EVI route to the remote V-spokes. In this solution, the V-hubs terminate the Ethernet A-D per-EVI route (used for Aliasing) and follows the procedures described in [RFC7432] for handling this route.

There are scenarios for which it is desirable to establish direct communication path between a pair of V-spokes for a given host MAC address. In such scenario, the advertising V-spoke advertises both the MAC/IP route and Ethernet A-D per-EVI route with the RT of V-hub (RT-VH) per [section 3 of \[RFC7024\]](#). The use of RT-VH, ensures that these routes are received by the V-spokes associated with that V-hub set and thus enables the V-spokes to perform the Aliasing procedure.

In summary, PE devices (V-hubs in general and V-spokes occasionally) that receive EVPN MAC/IP route advertisements (associated with a multi-homed site) need to also receive the associated Ethernet A-D per-EVI route advertisement(s) in order for them to perform Aliasing procedure.



#### **5.4. Split-Horizon & Mass Withdraw**

[RFC7432] uses Ethernet A-D per-ES route to a) signal to remote PEs the multi-homing redundancy type (Single-Active versus All-Active), b) advertise ESI label for split-horizon filtering when MPLS encapsulation is used, and c) advertise mass-withdraw when a failure of an access interface impacts many MAC addresses. This route does not need to be advertised from a V-spoke to any remote V-spoke unless a direct communication path between a pair of spoke is needed for a given flow.

Even if communication between a pair of V-spoke is needed for just a single flow, the Ethernet A-D per ES route needs to be advertised from the originating V-spoke for that ES which may handle tens or hundreds of thousands of flows. This is because in order to perform Aliasing function for a given flow, the Ethernet A-D per-EVI route is needed and this route itself is dependent on the Ethernet A-D per-ES route. In such scenario, the advertising V-spoke advertises the Ethernet A-D per-ES route with the RT of V-hub (RT-VH) per [section 3 of \[RFC7024\]](#).

In summary, PE devices (V-hubs in general and V-spokes occasionally) that receive EVPN MAC/IP route advertisements (associated with a multi-homed site) need to also receive the associated Ethernet A-D per-ES route advertisement(s).

### **6. Forwarding Considerations**

#### **6.1. IP-only Forwarding**

When EVPN operates in IP-only forwarding mode using EVPN Route Type 5, then all forwarding considerations in [section 4 of \[RFC7024\]](#) are directly applicable here.

#### **6.2. MAC-only Forwarding - Bridging**

When EVPN operates in MAC-only forwarding mode (i.e., bridging mode), then for a given EVI, the MPLS label that a V-hub advertises with anUnknown MAC address MUST be the label that identifies the MAC-VRF of the V-hub in absence of a more specific MAC route. When the V-hub receives a packet with such label, the V-hub pops the label and determines further disposition of the packet based on the lookup in the MAC-VRF. Otherwise, the MPLS label of the matching more specific route is used and packet is forwarded towards the associated NEXTHOP of the more specific route.

#### **6.3. MAC and IP Forwarding - IRB**



When a EVPN speaker operates in IRB mode, it implements both the IP and MAC forwarding Modes (aka Integrated Routing and Bridging - IRB).

On a packet by packet basis, the V-spoke decides whether to do forwarding based on a MAC address lookup (bridge) or based on a IP address lookup (route). If the host destination MAC address is that of the IRB interface (i.e., if the traffic is inter-subnet), then the V-spoke performs an additional IP lookup in the IP-VRF. However, if the host destination MAC address is that of an actual host MAC address (i.e., the traffic is intra- subnet) , then the V-spoke only performs a MAC lookup in the MAC-VRF. The procedure specified in [Section 6.1](#) and [Section 6.2](#) are applicable to inter-subnet and intra-subnet forwarding respectively. For intra- subnet traffic, if the MAC address is not found in the MAC-VRF, then the V-spoke forwards the traffic to the V-hub with the MPLS label received from the V-hub for the unknown MAC address. For the Inter- subnet traffic, if the IP prefix is not found in the IP-VRF, then the V-spoke forwards the traffic to the V-hub with the MPLS label received from the V-hub for the default IP address.

## **7. Handling of Broadcast and Multicast traffic**

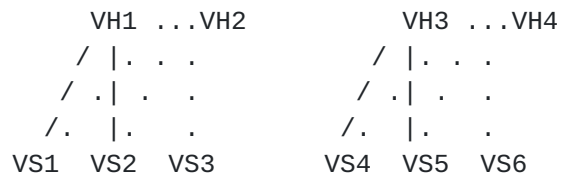
Just like that V-spoke to V-spoke known unicast traffic is relayed by V-hubs, V-spoke to V-spoke BUM traffic can also relayed by V-hubs. This is especially desired if Ingress Replication (IR) would be used otherwise for V-spokes to send traffic to other V-spokes. This way, a V-spoke can unicast BUM traffic to a single V-hub, who will then relay the traffic. This achieves Assisted Replication, and reduces multicast state in the core. Note that a V-hub may relay traffic using MPLS P2MP tunnels or BIER as well as IR. While a V-spoke may use P2MP tunnels or BIER to send traffic to V-hubs, this specification focuses on using IR by V-spokes.

In this particular section, all traffic refers to BUM traffic unless explicitly stated otherwise. The term PE refers to a V-hub or V-spoke when there is no need to distinguish the two.

Consider the following topology, where V-spokes VS1/2/3 are associated with V-hubs VH1/2 in one cluster, and V-spokes VS4/5/6 are associated with V-hubs VH3/4 in another cluster. Note that the lines/dots in the diagram indicate association, not connection.







### 7.1. Split Horizon

When VH1 relays traffic that it receives from VS1, in case of IR it MUST not send traffic back to VS1, and in case of P2MP tunnel it must indicate that traffic is sourced from VS1 so that VS1 will discard the traffic. In case of IR with IP unicast tunnels, the outer source IP address identifies the sending PE. In case of IR with MPLS unicast tunnels, VH1 must advertise different labels to different PEs, so that it can identify the sending PE based on the label in the traffic from a V-spoke.

If MPLS P2MP/multicast tunnels (including VXLAN-GPE and MPLS-over-GRE/UDP) are used by a V-hub to relay traffic, an upstream allocated (by the V-hub) label MUST be imposed in the label stack to identify the source of the V-spoke. The label is advertised as part of the PE Distinguisher (PED) Label Attribute of the Inclusive Multicast Ethernet Tag (IMET) route from the V-hub, as specified in [Section 8 of \[RFC 6514\]](#).

Notice that an "upstream-assigned" label used by a V-hub to send traffic with on a P2MP tunnel to identify the source V-spoke is the same "downstream-assigned" label used by the V-hub to receive traffic on the IR tunnel from the V-spoke. Therefore, the same PED Label attribute serves two purposes. With [\[RFC 6514\]](#), a PED label may only identify a PE but not a particular VPN. Here the PED label identifies both the PE and a particular EVI/BD. A V-spoke programs its context MPLS forwarding table for the V-hub to discard any traffic with the PED label that the V-hub advertised for this V-spoke, or pop other PED labels and direct traffic into a corresponding EVI for L2 forwarding.

Note that a V-hub cannot use VXLAN/NVGRE multicast tunnels to relay traffic because if the V-hub uses the source V-spoke's IP address in the outer IP header (for the purpose of identifying the source V-spoke), multicast RPF would fail and the packets will be discarded.

### 7.2. Route Advertisement



As with other route types, IMET routes from V-hubs are advertised with RT-VH and RT-EVI so they are imported by associated V-spokes and all V-hubs. They carry the PED Label attribute as described above.

IMET routes from V-spokes are advertised with RT-EVI so they are imported by all V-hubs. They also carry PED Label attribute for multi-homing split horizon purpose if and only if V-hubs uses IR to relay traffic.

If a V-hub uses RSVP-TE P2MP tunnel, IR, or BIER to send or relay traffic, all other PEs (V-hubs or V-spokes) will receive traffic directly because the V-hub sees all PEs. If a V-hub uses mLDP P2MP tunnel to send or relay traffic, only its associated V-spokes and all V-hubs will see the V-hub's IMET route and join the tunnel announced in the route. Another V-hub need to relay traffic to its associated V-spokes that are not associated with this V-hub.

For that V-hub to announce the mLDP relay tunnel in its cluster, it needs to advertise a (\*,\*) S-PMSI AD route, as specified in [BUM-PROCEDURE]. The route is advertised with the RT-VH for that cluster, and associated V-spokes will join the tunnel announced in the S-SPMI AD route.

### **7.3. Designated Forwarder in a Cluster**

When there are multiple V-hubs in a cluster, a V-spoke in that cluster decides by itself to which V-hub to send traffic. If the receiving V-hub uses mLDP tunnel to relay traffic, V-hubs in other clusters need to further relay traffic, but only one V-hub in each cluster can do so. As a result, a DF must be elected among the V-hubs for each cluster.

The election is similar to DF election in [RFC 7432](#), with the following differences.

- o Instead of using Ethernet Segment route to discover the PEs on a multi-homing ES, the IMET route are used to determine the V-hubs in the same cluster - they all carry the same pair of RT-EVI and RT-VH, and advertises the unknown mac route.
- o Instead of using VLAN to do per-VLAN DF election, the Local Administration Field of the RT-EVI is used to do per-EVI DF election.

### **7.4. Traffic Forwarding Rules**

When a PE needs to forward received traffic from local Attachment Circuits (ACs) or remote PEs to local ACs, it follows the rules in [RFC 7432](#), except that traffic sourced from this local PE but relayed



back on a p2mp tunnel is discarded. It may also need to forward to other PEs, subject to rules in the following sections.

#### **7.4.1. Traffic from Local ACs**

Traffic from a V-hub's local ACs is forwarded using the tunnel announced in its IMET route, as specified in [RFC 7432](#). In case of an mLDP tunnel, the traffic need to be relayed by V-hubs of other clusters to their associated V-spokes. For other tunnel types, no relay is needed.

Traffic from a V-spoke's local ACs is forwarded to an associated V-hub of its choice. In case of MPLS IR, the label in the V-hub's IMET route's PED attribute corresponding to this V-spoke is used.

#### **7.4.2. Traffic Received by a V-hub from Another PE**

When a V-hub receives traffic from an associated V-spoke, it needs to relay to other PEs, using the tunnel announced in its IMET route. In case of IR or BIER, the source V-spoke, which is determined from the incoming label or source IP address, is excluded from the replication list. In case of a P2MP tunnel, the popped incoming label is imposed again to identify the source PE, before the tunnel label is imposed.

When a V-hub receives traffic from another V-hub on a P2MP tunnel, and the tunnel is announced in an IMET route carrying the same RT-VH as this V-hub is configured with, it does not need to relay the traffic. Otherwise, the traffic is from a V-hub in a different cluster, and this V-hub needs to relay to its associated V-spokes, if and only if it is the DF for this cluster, using the tunnel announced in its (\*,\*) S-PMSI route carrying its RT-VH.

When a V-hub receives traffic from another V-hub via IR or BIER, it does not further relay the traffic as that V-hub can reach all PEs.

#### **7.4.3. Traffic received by a V-spoke from a V-hub**

In case of P2MP tunnel, the V-spoke discards the traffic if the label following the tunnel label identifies the V-spoke itself.

### **7.5. Multi-homing support**

Consider that an ES spans across two V-spokes in the same cluster and the V-hub uses MPLS IR to relay traffic. With ESI Label split horizon method, a source V-spoke uses the ESI label advertised by the V-hub for the ES, and the V-hub must change that to the ESI label advertised by receiving v-spokes when it relays traffic. That means V-hubs must advertise ESI labels for all multi-homing segments, even



when they're not on those segments. They must also do double label swap (EVI/BD label and ESI label) or mac lookup when relaying traffic.

There are two methods detailed below to avoid that complexity. Either one MAY be used.

#### **7.5.1 Domain-wide Common Block (DCB) Label**

[[draft-zzhang-bess-mvpn-evpn-aggregation-label](#)] proposes for all PEs on an MHES to use the same ESI label allocated from a Domain-wide Common Block. Not only does that have the advantages described in that document, but also It avoids the MHES complexity with Virtual Hub and Spoke as mentioned above, because the V-Hubs do not need to care about the ESI label at all any more.

#### **7.5.2 Local Bias**

If DCB labels cannot be used, then Local Bias can be used even For EVPN MPLS. The PED label following the mpls transport tunnel label or BIER header identifies the PE that originated the traffic in addition to identifying the EVI/BD.

If a V-hub uses P2MP or BIER to relay traffic, the PED label is one of the labels in the PE Distinguisher Label attribute in the V-hub's IMET route, allocated by the V-hub for the source V-spoke.

If a V-hub uses IR to relay traffic, for each V-spoke that it relays to, the PED label advertised by that receiving V-spoke for the source V-spoke needs to be imposed by the V-hub. For that purpose, each V-spoke must include the PED Label attribute in its IMET route, to advertise different labels for different PEs. It discovers the PEs that it needs to advertise labels for via the PED label Attribute in the V-hub's IMET route.

#### **7.6. Direct V-spoke to V-spoke traffic**

It may be desired for allow direct V-spoke to V-spoke traffic in a cluster, without the relay by a V-hub. To do that, V-spokes advertise their IMET routes with both RT-VH and RT-EVI. Forwarding rules will be specified in future revisions.

### **8. ARP/ND Suppression**

[RFC7432] defines the procedures for ARP/ND suppression where a PE can terminate gratuitous ARP/ND request message from directly connected site and advertises the associated MAC and IP addresses in an EVPN MAC/IP advertisement route to all other remote PEs. The





remote PEs that receive this EVPN route advertisement, install the MAC/IP pair in their ARP/ND cache table thus enabling them to terminate ARP/ND requests and generate ARP/ND responses locally thus suppressing the flooding of ARP/ND requests over the EVPN network.

In this hub-and-spoke approach, the ARP suppression needs to be performed by both the EVPN V-hubs as well V-spokes as follow. When a V-Spoke receives a gratuitous ARP/ND request, it terminates it and stores the source MAC/IP pair in its ARP/ND cache table. Then, it advertises the source MAC/IP pair to its associated V-Hubs using EVPN MAC/IP advertisement route. The V-Hubs upon receiving this EVPN route advertisement, create an entry in their ARP/ND cache table for this MAC/IP pair.

Now when a V-Spoke receives an ARP/ND request, it first looks up its ARP cache table, if an entry for that MAC/IP pair is found, then an ARP/ND response is generated locally and sent to the CE. However, if an entry is not found, then the ARP/ND request is unicasted to one of the V-hub associated with this V-spoke. Since, the associated V-hub keeps all the MAC/IP ARP entries in its cache table, it can formulate and ARP/ND response and forward it to that CE via the corresponding V-spoke.

## **9. IANA Considerations**

There is no additional IANA considerations for PBB-EVPN beyond what is already described in [[RFC7432](#)].

## **10. Security Considerations**

All the security considerations in [[RFC7432](#)] apply directly to this document because this document leverages [[RFC7432](#)] control plane and their associated procedures - although not the complete set but rather a subset.

This draft does not introduce any new security considerations beyond that of [[RFC7432](#)] and [[RFC4761](#)] because advertisements and processing of B-MAC addresses follow that of [[RFC7432](#)], and processing of C-MAC addresses follow that of [[RFC4761](#)] - i.e, B-MAC addresses are learned in control plane and C-MAC addresses are learned in data plane.

## **11. Acknowledgements**

The authors would like to thank Yakov Rekhter for initial idea discussions.



## **12.    Change Log**

Initial Version:    Sep 21 2014    Original Name: [draft-keyupate-evpn-virtual-hub-00.txt](#)

## **13.    References**

### **13.1.    Normative References**

- [RFC7024]    Jeng, H., Uttaro, J., Jalil, L., Decraene, B., Rekhter, Y., and R. Aggarwal, "Virtual Hub-and-Spoke in BGP/MPLS VPNs", [RFC 7024](#), October 2013.
- [RFC7432]    A. Sajassi, et al., "BGP MPLS Based Ethernet VPN", [RFC 7432](#) , February 2015.

### **13.2.    Informative References**

- [RFC7080]    A. Sajassi, et al., "Virtual Private LAN Service (VPLS) Interoperability with Provider Backbone Bridges", [RFC 7080](#), December 2013.
- [RFC7209]    D. Thaler, et al., "Requirements for Ethernet VPN (EVPN)", [RFC 7209](#), May 2014.
- [RFC4389]    A. Sajassi, et al., "Neighbor Discovery Proxies (ND Proxy)", [RFC 4389](#), April 2006.
- [RFC4761]    K. Kompella, et al., "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling", [RFC 4761](#), January 2007.
- [OVERLAY]    A. Sajassi, et al., "A Network Virtualization Overlay Solution using EVPN", [draft-ietf-bess-evpn-overlay-01](#), work in progress, February 2015.

## **14.    Authors' Addresses**

Keyur Patel  
Cisco Systems  
170 W. Tasman Drive  
San Jose, CA 95134, US  
Email: [keyupate@cisco.com](mailto:keyupate@cisco.com)

Ali Sajassi



Cisco  
170 West Tasman Drive  
San Jose, CA 95134, US  
Email: sajassi@cisco.com

Yakov Rekhter  
Juniper Networks, Inc.  
Email: yakov@juniper.net

John E. Drake  
Juniper Networks, Inc.  
Email: jdrake@juniper.net

Zhaohui Zhang  
Juniper Networks, Inc.  
Email: zzhang@juniper.net

Wim Henderickx  
Nokia  
Email: wim.henderickx@nokia.com

