Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: November 7, 2016

R. Sharma, Ed.
A. Banerjee
R. Sivaramu
Cisco Systems

May 6, 2016

Multi-site EVPN based VXLAN using Border Gateways
draft-sharma-multi-site-evpn-00

Abstract

This document describes the procedures for interconnecting two or more BGP based Ethernet VPN (EVPN) sites in a scalable fashion over an IP-only network. The motivation is to support extension of EVPN sites without having to rely on typical Data Center Interconnect (DCI) technologies like MPLS/VPLS for the interconnection. The requirements for such a deployment are very similar to the ones specified in RFC 7209 -- "Requirements for Ethernet VPN (EVPN)".

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 November 7, 2016.

Table of Contents

1. Introduction

BGP based Ethernet VPNs (EVPNs) are being used to support various VPN topologies with the motivation and requirements being discussed in detail in RFC7209 [RFC7209]. EVPN has been used to provide a Network Virtualization Overly (NVO) solution with a variety of tunnel encapsulation options over IP as described in [DCI-EVPN-OVERLAY]. EVPN used for the Data center interconnect (DCI) at the WAN Edge is discussed in [DCI-EVPN-OVERLAY]. The EVPN DCI procedures are defined for IP and MPLS hand-off at the site boundaries.

In the current EVPN deployments, there is a need to segment the EVPN domains within a Data Center (DC) primarily due to the service architecture and the scaling requirements around it. The number of routes, tunnel end-points, and next-hops needed in the DC are larger than some of the hardware elements that are being deployed. Network operators would like to ensure that they have means to have smaller sites within the data center, if they so desire, without having to have traditional DCI technologies to inter-connect them. In essence, they want smaller multi-site EVPN domains with an IP backbone.

Network operators today are using the Virtual Network Identifier (VNI) to designate a service. However, they would like to have this service available to a smaller set of nodes within the DC for administrative reasons; in essence they want to break up the EVPN domain to multiple smaller sites. An advantage of having a smaller footprint for these EVPN sites, implies that the various fault isolation domains are now more constrained. It is also feasible to have features that can re-use the VNI space across these sites if desired. The above mentioned motivations for having smaller multi-site EVPN domains are over and above the ones that are already detailed in RFC7209 [RFC7209].

In this document we focus primarily on the VXLAN encapsulation for EVPN deployments. We assume that the underlay provides simple IP connectivity. We go into the details of the IP/VXLAN hand-off mechanisms, to interconnect these smaller sites, within the data center itself. We describe this deployment model as a scalable multi-site EVPN (MS-EVPN) deployment. The procedures described here go into substantial detail regarding interconnecting L2 and L3, unicast and multicast domains across multiple EVPN sites.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. Terminology

- Border Gateway (BG): This is the node that interacts with nodes within a site and with nodes that are external to the site. For example, in a leaf-spine data center fabric, it can be a leaf, a Spine, or a separate device acting as gateway to interconnect the sites.

- All-Active Border Gateway: A Virtual set of shared Border Gateways (or Next-hops) acting as Multiple entry-exit points for a site.

- Single-Active Border Gateway: A Virtual set of unique border Gateways (or Next-hops) acting as a Multiple entry-exit points for a site.

- A-D: Auto-discovery.

3. Multi-Site EVPN Overview

In this section we describe the motivation, requirements, and framework of the multi-site EVPN enhancements.

3.1. MS-EVPN Interconnect Requirements

In this section we discuss the requirements and motivation for interconnecting different EVPN sites within a data center. In general any interconnect technology has the following requirements:

a. Scalability: Multi-Site EVPN (MS-EVPN) should be able to interconnect multiple sites in a scalable fashion. In other words, interconnecting such sites should not lead to one giant fabric with full mesh of end-to-end VXLAN tunnels across leafs in different sites. This leads to scale issues with respect to managing large number of tunnel end-points and a large number of tunnel next-hops. Also a huge flat fabric rules out option of ingress replication (IR) trees as number of replications becomes practically unachievable due to the internal bandwidth needed in hardware.

b. Multi-Destination traffic over unicast-only cloud: MS-EVPN mechanisms should be able to provide an efficient forwarding mechanism for multi-destination frames even if the underlay inter-site network is not capable of forwarding multicast frames. This requirement is meant to ensure that for the solution to work there are no additional constraints being requested of the IP network. This allows for use of existing network elements as-is.

c. Maintain Site-specific Administrative control: The MS-EVPN technology should be able to interconnect fabrics from different Administrative domains. It is possible that different sites have different VLAN-VNI mappings, use different underlay routing protocols, and/or have different PIM-SM group ranges etc. It is expected that the technology should not impose any additional constraints on the various administrative domains.

d. Isolate fault domains: MS-EVPN technology hand-off should have capability to isolate traffic cross site boundaries and prevent defects to percolate from one site to another. As an example, a

broadcast storm in a site should not lead to meltdown of all other sites.

e. Loop detection and prevention: In the scenarios where flood domains are stretched across fabrics, interconnecting sites are very vulnerable to loops and flood storms. There is a need to provide comprehensive loop detection and prevention capabilities.

f. Plug-and-play and extensibility: Addition of new sites or increasing capacity of existing sites should be achievable in a completely plug-and-play fashion. This essentially means that all control plane and forwarding states (L2 or L3 interconnect) should be built in downstream allocation mode. MS-EVPN should not pose any maximum requirements on the scale and capacity, it should be easily extendable on those metrics.

3.2. MS-EVPN Interconnect concept and framework

EVPN with an IP-only interconnect is conceptualized as multiple site-local EVPN control planes and IP forwarding domains interconnected via a single common EVPN control and IP forwarding domain. Every EVPN node is identified with a unique site-scope identifier. A site-local EVPN domain consists of EVPN nodes with the same site identifier. Border gateways on one hand are also part of site-specific EVPN domain and on other hand part of a common EVPN domain to interconnect with Border Gateways from other sites. Although a border gateway has only a single explicit site-id (that of the site it is a member of), it can be considered to also have a second implicit site-id, that of the interconnect-domain which has membership of all the BG's from all sites that are being interconnected. This implicit site-id membership is derived by the presence of the Border A-D route announced by that border gateway node.

These border gateways discover each other through EVPN Border A-D routes and act as both control and forwarding plane gateway across sites. This will facilitate site-specific nodes to visualize all other sites to be reachable only via its Border Gateways.

We describe the MS-EVPN deployment model using the topology below. In the topology there are 3 sites, Site A, Site B, and Site C that are inter-connected using IP. This entire topology is deemed to be part of the same Data Center. In most deployments these sites can be thought of as pods, which may span a rack, a row, or multiple rows in the data center, depending on the size of domain desired for scale and fault and/or administrative isolation domains.

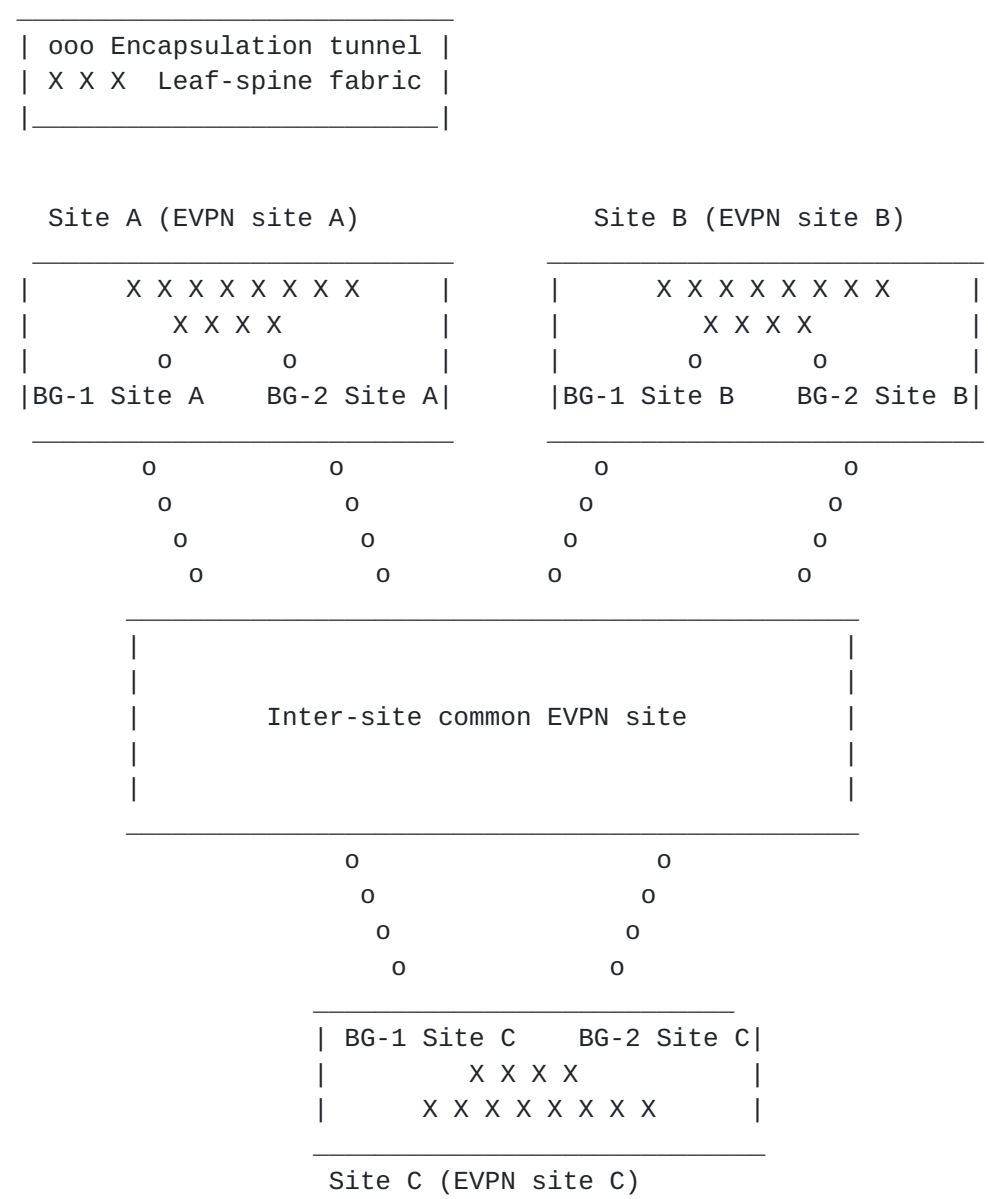

Figure 1

In this topology, site-local nodes are connected to each other by iBGP EVPN peering and Border Gateways are connected by eBGP Muti-hop EVPN peering via inter-site cloud. We explicitly spell this out to ensure that we can re-use BGP semantics of route announcement between and across the sites. There are other BGP mechanisms to instantiate this and they are not discussed in this document. This implies that each domain has its own AS number associated with it. In the topology, only 2 border gateway per site are shown; this is more for ease of illustration and explanation. The technology poses no such limitation. As mentioned earlier, site-specific EVPN domain will consists of only site-local nodes in the sites. A Border Gateway is

logically partitioned into site specific EVPN domain towards the site and into common EVPN domain towards other sites. This facilitates them to acts as control and forwarding plane gateway for forwarding traffic across sites.

EVPN nodes with in a site will discover each other via regular EVPN procedures and build site-local bidirectional VXLAN tunnels and multi-destination trees from leaves to Border Gateways. Border Gateways will discover each other by Border A-D routes (defined in Section 4.1) and build inter-site bi-directional VXLAN tunnels and Multi-destination trees between them. We thus build an end-to-end bidirectional forwarding path across all sites by stitching (and not by stretching end-to-end) site-local VXLAN tunnels with inter-site VXLAN tunnels.

In essence, a MS-EVPN fabric is proposed to be built in complete downstream and modular fashion.

- Site-local Bridging domains are interconnected ONLY via Border Gateways with Bridging domains from other sites. Such interconnect do not assume uniform mappings of mac-vrf VNI-VLAN across sites and stitches such bridging domains in complete downstream fashion using EVPN route advertisements.

- Site-local Routing domains are interconnected ONLY via Border Gateways with Routing domains from other sites. Such interconnect do not assume uniform mappings of IP VRF-VNI across sites and stitches such routing domains in complete downstream fashion using EVPN route advertisements.

- Site-local Flood domains are interconnected ONLY via Border Gateways with flood domains from other sites. Such interconnect do not assume uniform mappings of mac-vrf VNI across sites (or mechanisms to build flood domains with in site) and stitches such flood domains in complete downstream fashion using EVPN route advertisements. It however do not exclude possibility of building an end-to-end flood domain, if desired for other reasons.

The above architecture satisfies the constraints laid out in Section 3.1. For example, the size of a domain may be made dependent on the route and next-hop scale that can be supported by the deployment of the network nodes. There are no constraints on the network that connects the nodes within the domain or across the domains. In the event multicast capability is available and enabled, the nodes can use those resources. In the event the underlay is connected using unicast semantics, creation of ingress replication lists ensure that multi-destination frames reach their destinations. The domains may have their own deployment constraints, and the

overlay does not need any form of stretching. It is within the control of the administrator with respect to containing fault isolation domains. The automated discovery of the border nodes needs no further configurations for existing deployed domains.

4. Multi-site EVPN Interconnect Procedures

In this section we describe the new functionalities in the Border Gateway nodes for interconnecting EVPN sites within the DC.

4.1. Border Auto-Discovery Route

These routes are generated by Border Gateways and imported by leafs and Border Gateways. These routes serve following purpose:

- Discover Border Gateways from same site. This will help in finding designated forwarder for inter-site Multi-destination traffic. Once designated forwarder election is complete, inter-site Multi-destination traffic will be forwarded by DF winner.

- Discover Border Gateways from other sites. This will help in deciding which VXLAN tunnels should be terminated for inter-site traffic. Along with the Type 3 routes, this may help in optimal traffic flow within the common core for multi-destination frames.

A Border A-D route type specific EVPN NLRI is defined as follows. It is proposed to be a new route type in EVPN NLRI defined in RFC7432 [RFC7432].

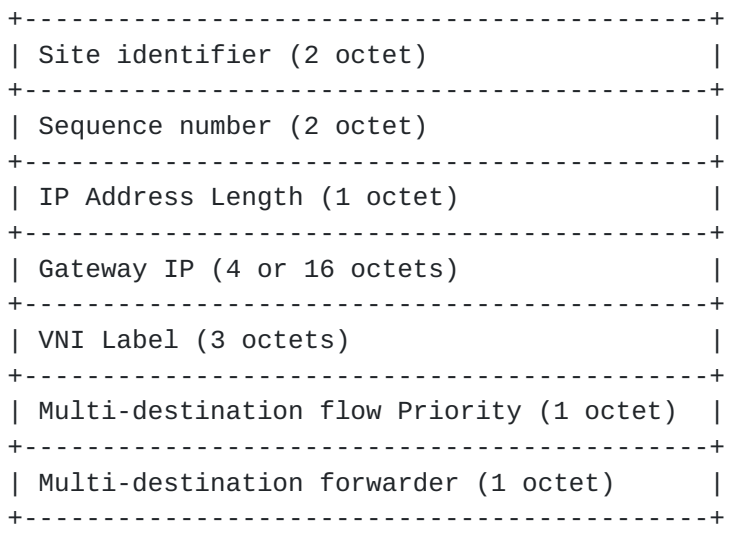

Figure 2

- Site Identifier: This is used to distinguish A-D routes received from border gateways in same site or in different sites. Border gateways discover each other by processing these A-D routes from different sites. These site identifier can be explicitly configured or the BGP Autonomous system (AS) number can be automatically carried as the site-identifier.

- Sequence number: Monotonically increasing sequence number added by Border Gateway while sending A-D route. In case there are multiple Border A-D routes, the one with the highest sequence number is honored while processing.

- IP Address Length: Number of bytes in the Gateway IP field, 4 bytes for IPv4 address or 16 bytes for IPv6 address.

- Gateway IP: This is the unique IP address of the Border gateways. This Gateway IP will be used to build Multi-destination trees.

- VNI Label: This is the MAC-VRF VNI or the IP-VRF VNI.

- Multi-destination flow Priority : This field is optional and is 0 if not used. This field can be used to assist in forwarder election for multi-destination traffic by assigning higher priority among border gateways of same site. This forwarder election is per MAC-VRF or IP-VRF VNI.

- Multi-destination forwarder : This field is set to TRUE once DF election is complete for Multi-destination traffic and announcing Border Gateway is the DF winner.

These A-D routes are advertised with mac-VRF and IP-VRF RTs depending on whether the VNI carried is a mac-VRF VNI or an IP VRF VNI.

After a Border Gateway is provisioned, Border A-D routes will be announced after some delay interval from all border gateways. This will provide sufficient time to learn Border A-D routes from other Border Gateways.

Border Gateways between same site will run a Designated forwarder election per MAC-VRF VNI for multi-destination traffic across the site. Border A-D routes coming from different site will not trigger DF election and will only be cached to terminate VXLAN tunnels from such border gateways.

Multi-destination flow priority will be assigned (based on optional policies) to prefer a border gateway for DF election per MAC or IP VRF VNI for multi-destination traffic and will be used in DF election to prefer higher priority border gateway as forwarder.

As has been defined in the specifications, Type 2, Type 3, and Type 5 routes carry downstream VNI labels. These A-D routes will help to pre-build VXLAN tunnels in the common EVPN domain for L2, L3, and Multi-Destination traffic. Also these A-D routes will help in correlating next-hop of EVPN routes and will facilitate in rewriting next-hop attributes before re-advertising these routes from other sites to a given site. This provides flexibility to keep different VNI-VLAN mapping in different sites and still able to interconnect L3 and L2 domains.

All control plane and data plane states are interconnected in a complete downstream fashion. For example, BGP import rules for a Type 3 route should be able to extend a flood domain for a VNI and flood traffic destined to advertised EVPN node should carry the VNI which is announced in Type 3 route. Similarly Type 2, Type 5 control and forwarding states should be interconnected in a complete downstream fashion.

4.2. Border Gateway Provisioning

Border Gateway nodes manage both the control-plane communications and the data forwarding plane for any inter-site traffic. Border Gateway functionality in an EVPN site SHOULD be enabled on more than one node in the network for redundancy and high-availability purposes. Any external Type-2/Type-5 routes that are received by the BGs of a site

are advertised to all the intra-site nodes by all the BGs. For internal Type-2/Type-5 routes received by the BG's from the intra-site nodes, all the BGs of a site would advertise them to the remote BG's, so any L2/L3 known unicast traffic to internal destinations could be sent to any one of the local BG's by remote sources. For known L2 and L3 unicast traffic, all of the individual border gateway nodes will behave either as single logical forwarding node or a set of active forwarding nodes. This can be perceived by intra-site nodes as multiple entry/exit points for inter-site traffic. For unknown unicast/multi-destination traffic, there must be a designated forwarder election mechanism to determine which node would perform the primary forwarding role at any given point in time, to ensure there is no duplication of traffic for any given flow (See Section 4.2.1).

4.2.1. Border Gateway Designated Forwarder Election

In the presence of more than one Border Gateway nodes in a site, forwarding of multi-destination L2 or L3 traffic both into the site and out of the site needs to be carried out by a single node. This DF election could be done independently by each candidate border gateway, by subjecting an ordered "candidate list" of all the BG's present in the same site (identified by reception of the Border A-D routes per-VNI with the same site-id as itself) to a hash-function on a per-VNI basis. All the candidate border gateways of the same site are required to use a uniform hash-function to yield the same result. Failure events which lead to a BG losing all of its connectivity to the IP interconnect backbone should trigger the BG to withdraw its Border A-D route(s), to indicate to other BG's of the site that it is no longer a candidate BG. Also there is a possibility of configuring policies to prefer a Border gateway over others and pick as DF winner.

There are two modes proposed for Border gateway provisioning.

4.2.2. All-active Border Gateway

In this mode all border gateways share same gateway IP and rewrite EVPN next-hop attributes with a shared logical next-hop entity. However these Gateways will maintain unique gateway IP to facilitate building IR trees from site-local nodes to forward Multi-Destination traffic. EVPN Type 2, Type 5 routes will be advertised to the nodes in the site from all border gateways and Border gateway will run DF election per VNI for Multi destination traffic. Type 3 routes will be advertised by all Border gateways but only DF will forward inter-site traffic.

This mode is useful when there is no preference between different border-gateways to forward traffic from different VNIs. Standard data plane hashing of VXLAN header will load balance traffic among Border Gateways.

Additionally, it is recommended that border gateway be enabled in the All-active mode wherein the BG functionality is available to the rest of the network as a single logical entity (as in Anycast) for inter-site communication. In the absence of capability for All-active, the BG could be enabled as individual gateways (Single-Active BG) wherein a single node will perform the active BG role for a given flow at a given time.

4.2.3. Multi-path Border Gateway

In this mode, Border gateways will rewrite EVPN Next-hop attributes with unique next-hop entities. This provides flexibility to apply usual policies and pick per-VRF, per-VNI or per-flow primary/backup border Gateways. Hence, an intra-site node will see each BG as a next-hop for any external L2 or L3 unicast destination, and would perform an ECMP path selection to load-balance traffic sent to external destinations. In case an intra-site node is not capable of performing ECMP hash based path-selection (possibly some L2 forwarding implementations), the node is expected to choose one of the BG's as its designated forwarder. EVPN Type 2, Type 5 routes will be advertised to the nodes in the site from all border gateways and Border gateway will run DF election per VNI for Multi destination traffic. Type 3 routes will be advertised by all Border gateways but only DF will forward inter-site traffic.

4.3. EVPN route processing at Border Gateway

Border gateways will build EVPN peering on processing A-D routes from other Border gateways. Route targets MAY be auto-generated based on some site-specific identifier. If BGP AS number is used as site-specific identifier, import and export route targets can be auto-generated as explained in RFC7432 [RFC7432]. This will facilitate site-local nodes to import routes from other nodes in same site and from its Border Gateways. Also this will prevent routes exchange between nodes from different sites. However, in this auto-generated scheme, import mechanism on Border Gateway should be relaxed to allow unconditional import of Border A-D routes from other border gateways. Also the routes which are imported at Border Gateway and re-advertised should implement a mechanism to avoid looping of updates should they come back at Border Gateways.

Type 2/Type 5 EVPN routes will be rewritten with Border Gateway IP, Border Gateway system mac as next-hop and re-advertised. Only EVPN

routes received from discovered Border gateways with different site identifiers will be rewritten and re-advertised. This will avoid rewriting every EVPN update if border gateways are also acting as Route reflector (RR) for site-local EVPN peering. Also this will help in interoperating MS-EVPN fabric with sites which do not have Border Gateway functionality.

There are few mechanisms suggested below for re-advertising these inter-site routes to a site and provide connectivity of inter-site hosts and subnets.

- All routes everywhere : In this mode all inter-site EVPN Type2/ Type5 routes are downloaded on site-local leafs from Border Gateways. In other words, every leaf in the MS-EVPN fabric will have routes from every intra-site and inter-site leafs. This mechanism is best-fit for the scenarios where inter-site traffic is as volumonous as intra-site flow traffic. Also this mechanism preserves usual glean processing, silent host discovery and unknown traffic handling at the leafs.

- Default routing to Border Gateways : In this mode, all received inter-site EVPN Type 2/Type 5 routes will be installed only at Border Gateways and will not be advertised in the site. Border Gateways will inject Type 5 default routes to site-local nodes and avoid re-advertising Type 2 from other sites. This mode provides scaling advantage by not downloading all inter-site routes to every leaf in MS-EVPN fabric. This mechanism MAY require glean processing and unknown traffic handling to be tailored to provide efficient traffic forwarding.

- Site-scope flow registry and discovery : This mechanism provides scaling advantage by downloading inter-site routes on-demand. It provides scaling advantages of default routing with out need to tailor glean processing and unknown traffic handling at the leafs. Leafs will create on-demand flow registry on their border Gateways and based on this flow registry border gateways will advertise Type 2 routes in a site. In other words, assuming that we have a trigger to send the EVPN routes that are needed by the site for conversational learning from the Border Gateways, we can optimize on the control plane state that is needed at the various leaf nodes. Hardware programming can be further optimized based on actual conversations needed by the leaf, as opposed to to the ones needed by the site. We will describe a mechanism in the appendix with respect to ARP processing at the Border Gateway.

Type 3 routes will be imported and processed on border gateways from other border gateways but MUST NOT be advertised again. In both modes (All-active and Multipath), Type 3 routes will be generated and

advertised by all Border Gateways with unique gateway IP. This will facilitate building fast converging flood domain connectivity inter-site and intra-site and on same time avoiding duplicate traffic by electing DF winner to forward multi-destination inter-site traffic.

4.4. Multi-Destination tree between Border Gateways

The procedures described here recommends building an Ingress Replication (IR) tree between Border Gateways. This will facilitate every site to independently build site-specific Multi destination trees. Multi-destination end-to-end trees between leafs could be PIM (site 1) + IR (between border Gateways) + PIM(site 2) or IR-IR-IR or PIM-IR-IR. However this does not rule out using IR-PIM-IR or end-to-end PIM to build multi-destination trees end-to-end.

Border Gateways will generate Type 3 routes with unique gateway IP and advertise to Border Gateways of other sites. These Type 3 routes will help in building IR trees between border gateways. However only DF winner per VNI will forward multi-destination traffic across sites.

As Border Gateways are part of both site-specific and inter-site Multi-destination IR trees, split-horizon mechanism will be used to avoid loops. Multi-destination tree with Border gateway as root to other sites (or Border-Gateways) will be in a separate horizon group. Similarity Multi-destination IR tree with Border Gateway as root to site-local nodes will be in another split horizon group.

If PIM is used to build Multi-Destination trees in site-specific domain, all Border gateway will join such PIM trees and draw multi-destination traffic. However only DF Border Gateway will forward traffic towards other sites.

4.5. Inter-site Unicast traffic

As site-local nodes will see all inter-site EVPN routes via Border Gateways, VXLAN tunnels will be built between leafs and site-local Border Gateways and Inter-site VXLAN tunnels will be built between Border gateways in different sites. An end-to-end VXLAN bidirectional forwarding path between inter-site leafs will consist of VXLAN tunnel from leaf (say Site A) to its Border Gateway, another VXLAN tunnel from Border Gateway to Border Gateway in another site (say site B) and Border gateway to leaf (in site B). Such arrangement of tunnels are very scalable as a full mesh of VXLAN tunnels across inter-site leafs is substituted by combination of intra-site and inter-site tunnels.

L2 and L3 unicast frames from site-local leafs will reach border gateway using VXLAN encapsulation. At Border gateway, VXLAN header is stripped out and another VXLAN header is pushed to sent frames to destination site Border Gateway. Destination site Border gateway will strip off VXLAN header and push another VXLAN header to send frame to the destination site leaf.

4.6. Inter-site Multi-destination traffic

Multi-destination traffic will be forwarded from one site to other site only by DF for that VNI. As frames reach Border Gateway from site-local nodes, VXLAN header will be popped and another VXLAN header (derived from downstream Type3 EVPN routes) will be pushed to forward frame to destination site border gateway. Similarly destination site Border Gateway will strip off VXLAN header and forward frame after pushing another VXLAN header towards the destination leaf.

As explained in Section 4.4, split horizon mechanism will be used to avoid looping of inter-site multi-destination frames.

4.7. Host Mobility

Host movement handling will be same as defined in RFC7432 [RFC7432]. When host moves, EVPN Type 2 routes with updated sequence number will be propagated to every EVPN node. When a host moves inter-site, only Border gateways may see EVPN updates with both next-hop attributes and sequence number changes and leafs may see updates only with updated sequence numbers. However in other cases both Border gateway and leafs may see next-hop and sequence number changes.

5. Convergence

5.1. Fabric to Border Gateway Failure

If a Border Gateway is lost, Border gateway next-hop will be withdrawn for Type 2 routes. Also per-VNI DF election will be triggered to chose new DF. DF new winner will become forwarder of Multi-destination inter-site traffic.

5.2. Border Gateway to Border Gateway Failures

In case where inter-site cloud has link failures, direct forwarding path between border gateways can be lost. In this case, traffic from one site can reach other site via border gateway of an intermediate site. However this will be addressed like regular underlay failure and traffic terminations end-points will still stay same for inter-site traffic flows.

6. Interoperability

The procedures defined here are only for Border Gateways. Therefore other EVPN nodes in the network should be RFC7432 [RFC7432] compliant to operate in such topologies.

As the procedures described here are applicable only after receiving Border A-D route, if other domains are connected which are not capable of such multi-site gateway model, they can work in regular EVPN mode. The exact procedures will be detailed in a future version of the draft.

7. Isolation of Fault Domains

Isolation of network defects requires policies like storm control, security ACLs etc to be implemented at site boundaries. Border gateways should be capable of inspecting inner payload of packets received from VXLAN tunnels and enforce configured policies to prevent defects percolating from one part to rest of the network.

8. Loop detection and Prevention

This has already been addressed in the Section 4.2.1. We are in essence using the Designated Forwarder and Split Horizon procedures to break loops in this network.

9. Acknowledgements

This authors would like to thank Max Ardica, Lukas Krattiger, Anuj Mittal, Lilian Quan, Veera Ravinutala, for their review and comments.

10. IANA Considerations

TBD.

11. Security Considerations

TBD.

12. References

12.1. Normative References

[DCI-EVPN-OVERLAY]
A. Sajassi et. al., "A Network Virtualization Overlay Solution using EVPN", 2016, <https://tools.ietf.org/html/draft-ietf-bess-evpn-overlay-02>.

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

[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>.

12.2. Informative References

[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>.

Appendix A. Additional Stuff

TBD.

Authors' Addresses

Rajesh Sharma (editor)
Cisco Systems
170 W Tasman Drive
San Jose, CA
USA

Email: rajshr@cisco.com

Ayan Banerjee
Cisco Systems
170 W Tasman Drive
San Jose, CA
USA

Email: ayabaner@cisco.com

Raghava Sivaramu
Cisco Systems
170 W Tasman Drive
San Jose, CA
USA

Email: raghavas@cisco.com