Internet Engineering Task Force (IETF)
Request for Comments: 9469
Category: Informational
ISSN: 2070-1721

J. Rabadan, Ed.
M. Bocci
Nokia
S. Boutros
Ciena
A. Sajassi
Cisco

September 2023

Applicability of Ethernet Virtual Private Network (EVPN) to Network Virtualization over Layer 3 (NVO3) Networks

Abstract

An Ethernet Virtual Private Network (EVPN) provides a unified control plane that solves the issues of Network Virtualization Edge (NVE) auto-discovery, tenant Media Access Control (MAC) / IP dissemination, and advanced features in a scablable way as required by Network Virtualization over Layer 3 (NVO3) networks. EVPN is a scalable solution for NVO3 networks and keeps the independence of the underlay IP Fabric, i.e., there is no need to enable Protocol Independent Multicast (PIM) in the underlay network and maintain multicast states for tenant Broadcast Domains. This document describes the use of EVPN for NVO3 networks and discusses its applicability to basic Layer 2 and Layer 3 connectivity requirements and to advanced features such as MAC Mobility, MAC Protection and Loop Protection, multihoming, Data Center Interconnect (DCI), and much more. No new EVPN procedures are introduced.

Status of This Memo

This document is not an Internet Standards Track specification; it is published for informational purposes.

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). Not all documents approved by the IESG are candidates for any level of Internet Standard; see 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/rfc9469.

Copyright Notice

Table of Contents

1. Introduction

In Network Virtualization over Layer 3 (NVO3) networks, Network Virtualization Edge (NVE) devices sit at the edge of the underlay network and provide Layer 2 and Layer 3 connectivity among Tenant Systems (TSes) of the same tenant. The NVEs need to build and maintain mapping tables so they can deliver encapsulated packets to their intended destination NVE(s). While there are different options to create and disseminate the mapping table entries, NVEs may exchange that information directly among themselves via a control plane protocol, such as Ethernet Virtual Private Network (EVPN). EVPN provides an efficient, flexible, and unified control plane option that can be used for Layer 2 and Layer 3 Virtual Network (VN) service connectivity. This document does not introduce any new procedures in EVPN.

In this document, we assume that the EVPN control plane module resides in the NVEs. The NVEs can be virtual switches in hypervisors, Top-of-Rack (ToR) switches or Leaf switches, or Data Center Gateways. As described in [RFC7365], Network Virtualization Authorities (NVAs) may be used to provide the forwarding information to the NVEs, and in that case, EVPN could be used to disseminate the information across multiple federated NVAs. The applicability of EVPN would then be similar to the one described in this document. However, for simplicity, the description assumes control plane communication among NVE(s).

2. EVPN and NVO3 Terminology

This document uses the terminology of [RFC7365] in addition to the terms that follow.

AC: Attachment Circuit or logical interface associated with a given BT. To determine the AC on which a packet arrived, the NVE will examine the physical/logical port and/or VLAN tags (where the VLAN tags can be individual c-tags, s-tags, or ranges of both).

ARP and NDP: Address Resolution Protocol (IPv4) and Neighbor Discovery Protocol (IPv6), respectively.

BD: Broadcast Domain that corresponds to a tenant IP subnet. If no suppression techniques are used, a BUM frame that is injected in a Broadcast Domain will reach all the NVEs that are attached to that Broadcast Domain. An EVI may contain one or multiple Broadcast Domains depending on the service model [RFC7432]. This document will use the term Broadcast Domain to refer to a tenant subnet.

BT: Bridge Table, as defined in [RFC7432]. A BT is the instantiation of a Broadcast Domain in an NVE. When there is a single Broadcast Domain on a given EVI, the MAC-VRF is equivalent to the BT on that NVE. Although a Broadcast Domain spans multiple NVEs and a BT is really the instantiation of a Broadcast Domain in an NVE, this document uses BT and Broadcast Domain interchangeably.

BUM: Broadcast, Unknown Unicast, and Multicast frames

Clos: A multistage network topology described in [CLOS1953], where all the edge switches (or Leafs) are connected to all the core switches (or Spines). Typically used in Data Centers.

DF and NDF: Designated Forwarder and Non-Designated Forwarder, respectively. These are the roles that a given PE can have in a given ES.

ECMP: Equal-Cost Multipath

ES: Ethernet Segment. When a Tenant System (TS) is connected to one or more NVEs via a set of Ethernet links, that set of links is referred to as an "Ethernet Segment". Each ES is represented by a unique Ethernet Segment Identifier (ESI) in the NVO3 network, and the ESI is used in EVPN routes that are specific to that ES.

Ethernet Tag: Used to represent a Broadcast Domain that is configured on a given ES for the purpose of Designated Forwarder election. Note that any of the following may be used to represent a Broadcast Domain: VIDs (including Q-in-Q tags), configured IDs, VNIs, normalized VIDs, Service Instance Identifiers (I-SIDs), etc., as long as the representation of the Broadcast Domains is configured consistently across the multihomed PEs attached to that ES.

EVI or EVPN Instance: A Layer 2 Virtual Network that uses an EVPN control plane to exchange reachability information among the member NVEs. It corresponds to a set of MAC-VRFs of the same tenant. See MAC-VRF in this section.

EVPN: Ethernet Virtual Private Network, as described in [RFC7432].

EVPN VLAN-Aware Bundle Service Interface: Similar to the VLAN-bundle interface but each individual VLAN value is mapped to a different Broadcast Domain. In this interface, there are multiple Broadcast Domains per EVI for a given tenant. Each Broadcast Domain is identified by an "Ethernet Tag", which is a control plane value that identifies the routes for the Broadcast Domain within the EVI.

EVPN VLAN-Based Service Interface: One of the three service

interfaces defined in [RFC7432]. It is characterized as a Broadcast Domain that uses a single VLAN per physical access port to attach tenant traffic to the Broadcast Domain. In this service interface, there is only one Broadcast Domain per EVI.

EVPN VLAN-Bundle Service Interface: Similar to the VLAN-based interface but uses a bundle of VLANs per physical port to attach tenant traffic to the Broadcast Domain. Like the VLAN-based interface, there is only one Broadcast Domain per EVI.

Geneve: Generic Network Virtualization Encapsulation. An NVO3 encapsulation defined in [RFC8926].

IP-VRF: IP Virtual Routing and Forwarding table, as defined in [RFC4364]. It stores IP Prefixes that are part of the tenant's IP space and are distributed among NVEs of the same tenant by EVPN. A Route Distinguisher (RD) and one or more Route Targets (RTs) are required properties of an IP-VRF. An IP-VRF is instantiated in an NVE for a given tenant if the NVE is attached to multiple subnets of the tenant and local inter-subnet forwarding is required across those subnets.

IRB: Integrated Routing and Bridging. It refers to the logical interface that connects a Broadcast Domain instance (or a BT) to an IP-VRF and forwards packets with a destination in a different subnet.

MAC-VRF: A MAC Virtual Routing and Forwarding table, as defined in [RFC7432]. The instantiation of an EVI (EVPN Instance) in an NVE. A Route Distinguisher (RD) and one or more RTs are required properties of a MAC-VRF, and they are normally different from the ones defined in the associated IP-VRF (if the MAC-VRF has an IRB interface).

NVE: Network Virtualization Edge. A network entity that sits at the edge of an underlay network and implements Layer 2 and/or Layer 3 network virtualization functions. The network-facing side of the NVE uses the underlying Layer 3 network to tunnel tenant frames to and from other NVEs. The tenant-facing side of the NVE sends and receives Ethernet frames to and from individual Tenant Systems. In this document, an NVE could be implemented as a virtual switch within a hypervisor, a switch, or a router, and it runs EVPN in the control plane.

NVO3 tunnels: Network Virtualization over Layer 3 tunnels. In this document, NVO3 tunnels refer to a way to encapsulate tenant frames or packets into IP packets, whose IP Source Addresses (SAs) or Destination Addresses (DAs) belong to the underlay IP address space, and identify NVEs connected to the same underlay network. Examples of NVO3 tunnel encapsulations are VXLAN [RFC7348], Geneve [RFC8926], or MPLSoUDP [RFC7510].

PE: Provider Edge

PMSI: Provider Multicast Service Interface

PTA: PMSI Tunnel Attribute

RT and RD: Route Target and Route Distinguisher, respectively.

RT-1, RT-2, RT-3, etc.: These refer to the Route Types followed by the type numbers as defined in the "EVPN Route Types" IANA registry (see <https://www.iana.org/assignments/evpn/>).

SA and DA: Source Address and Destination Address, respectively. They are used along with MAC or IP, e.g., IP SA or MAC DA.

SBD: Supplementary Broadcast Domain, as defined in [RFC9136]. It is a Broadcast Domain that does not have any Attachment Circuits, only has IRB interfaces, and provides connectivity among all the IP-VRFs of a tenant in the Interface-ful IP-VRF-to-IP-VRF models.

TS: Tenant System. A physical or virtual system that can play the role of a host or a forwarding element, such as a router, switch, firewall, etc. It belongs to a single tenant and connects to one or more Broadcast Domains of that tenant.

VID: Virtual Local Area Network Identifier

VNI: Virtual Network Identifier. Irrespective of the NVO3 encapsulation, the tunnel header always includes a VNI that is added at the ingress NVE (based on the mapping table lookup) and identifies the BT at the egress NVE. This VNI is called VNI in VXLAN or Geneve, Virtual Subnet ID (VSID) in nvGRE, or Label in MPLSoGRE or MPLSoUDP. This document refers to VNI as a generic VNI for any NVO3 encapsulation.

VXLAN: Virtual eXtensible Local Area Network. An NVO3 encapsulation defined in [RFC7348].

3. Why is EVPN Needed in NVO3 Networks?

Data Centers have adopted NVO3 architectures mostly due to the issues discussed in [RFC7364]. The architecture of a Data Center is nowadays based on a Clos design, where every Leaf is connected to a layer of Spines and there is a number of ECMPs between any two Leaf nodes. All the links between Leaf and Spine nodes are routed links, forming what we also know as an underlay IP Fabric. The underlay IP Fabric does not have issues with loops or flooding (like old Spanning Tree Data Center designs did), convergence is fast, and ECMP generally distributes utilization well across all the links.

On this architecture, and as discussed by [RFC7364], multi-tenant intra-subnet and inter-subnet connectivity services are provided by NVO3 tunnels. VXLAN [RFC7348] and Geneve [RFC8926] are two examples of such NVO3 tunnels.

Why is a control plane protocol along with NVO3 tunnels helpful? There are three main reasons:

a. Auto-discovery of the remote NVEs that are attached to the same VPN instance (Layer 2 and/or Layer 3) as the ingress NVE is.

b. Dissemination of the MAC/IP host information so that mapping tables can be populated on the remote NVEs.

c. Advanced features such as MAC Mobility, MAC Protection, BUM and ARP/ND traffic reduction/suppression, multihoming, functionality similar to Prefix Independent Convergence (PIC) [BGP-PIC], fast convergence, etc.

"Flood and learn" is a possible approach to achieve points (a) and (b) above for multipoint Ethernet services. "Flood and learn" refers to "flooding" BUM traffic from the ingress NVE to all the egress NVEs attached to the same Broadcast Domain instead of using a specific control plane on the NVEs. The egress NVEs may then use data path source MAC address "learning" on the frames received over the NVO3 tunnels. When the destination host replies and the frames arrive at the NVE that initially flooded BUM frames, the NVE will also "learn" the source MAC address of the frame encapsulated on the NVO3 tunnel. This approach has the following drawbacks:

* In order to flood a given BUM frame, the ingress NVE must know the IP addresses of the remote NVEs attached to the same Broadcast Domain. This may be done as follows:

  - The remote tunnel IP addresses can be statically provisioned on the ingress NVE. If the ingress NVE receives a BUM frame for the Broadcast Domain on an ingress Attachment Circuit, it will do ingress replication and will send the frame to all the configured egress NVE destination IP addresses in the Broadcast Domain.

  - All the NVEs attached to the same Broadcast Domain can subscribe to an underlay IP multicast group that is dedicated to that Broadcast Domain. When an ingress NVE receives a BUM frame on an ingress Attachment Circuit, it will send a single copy of the frame encapsulated into an NVO3 tunnel using the multicast address as the destination IP address of the tunnel. This solution requires PIM in the underlay network and the association of individual Broadcast Domains to underlay IP multicast groups.

* "Flood and learn" solves the issues of auto-discovery and the learning of the MAC to VNI/tunnel IP mapping on the NVEs for a given Broadcast Domain. However, it does not provide a solution for advanced features, and it does not scale well (mostly due to the need for constant flooding and the underlay PIM states that must be maintained).

EVPN provides a unified control plane that solves the issues of NVE auto-discovery, tenant MAC/IP dissemination, and advanced features in a scalable way and keeps the independence of the underlay IP Fabric; i.e., there is no need to enable PIM in the underlay network and maintain multicast states for tenant Broadcast Domains.

Section 4 describes how EVPN can be used to meet the control plane requirements in an NVO3 network.

4. Applicability of EVPN to NVO3 Networks

This section discusses the applicability of EVPN to NVO3 networks. The intent is not to provide a comprehensive explanation of the protocol itself, but to give an introduction and point at the corresponding reference document so the reader can easily find more details if needed.

4.1. EVPN Route Types Used in NVO3 Networks

EVPN supports multiple Route Types, and each type has a different function. For convenience, Table 1 shows a summary of all the existing EVPN Route Types and their usages. In this document, we may refer to these route types as RT-x routes, where x is the type number included in the first column of Table 1.

Type	Description	Usage
1	Ethernet Auto-Discovery	Multihoming: Used for MAC mass-withdraw when advertised per Ethernet Segment and for aliasing/backup functions when advertised per EVI.
2	MAC/IP Advertisement	Host MAC/IP dissemination; supports MAC Mobility and protection.
3	Inclusive Multicast	NVE discovery and BUM flooding tree setup.

	Ethernet Tag		
4	Ethernet Segment	Multihoming: ES auto-discovery and DF election.	
5	IP Prefix	IP Prefix dissemination.	
6	Selective Multicast Ethernet Tag	Indicate interest for a multicast S,G or *,G.	
7	Multicast Join Synch	Multihoming: S,G or *,G state synch.	
8	Multicast Leave Synch	Multihoming: S,G or *,G leave synch.	
9	Per-Region I-PMSI A-D	BUM tree creation across regions.	
10	S-PMSI A-D	Multicast tree for S,G or *,G states.	
11	Leaf A-D	Used for responses to explicit tracking.	

Table 1: EVPN Route Types

4.2. EVPN Basic Applicability for Layer 2 Services

Although the applicability of EVPN to NVO3 networks spans multiple documents, EVPN's baseline specification is [RFC7432]. [RFC7432] allows multipoint Layer 2 VPNs to be operated as IP VPNs [RFC4364], where MACs and the information to set up flooding trees are distributed by Multiprotocol BGP (MP-BGP) [RFC4760]. Based on [RFC7432], [RFC8365] describes how to use EVPN to deliver Layer 2 services specifically in NVO3 networks.

Figure 1 represents a Layer 2 service deployed with an EVPN Broadcast Domain in an NVO3 network.

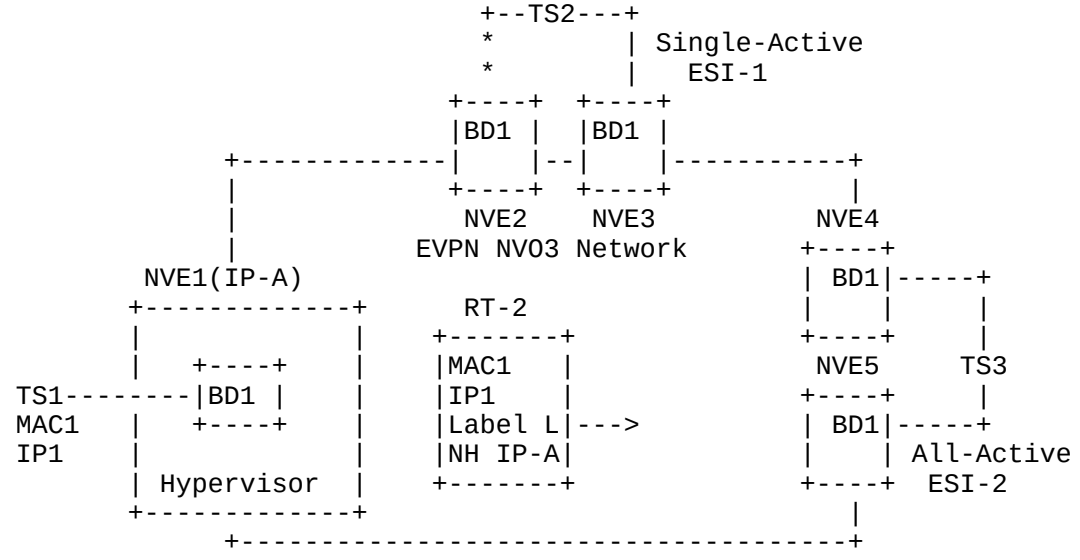

Figure 1: EVPN for L2 in an NVO3 Network - Example

In a simple NVO3 network, such as the example of Figure 1, these are the basic constructs that EVPN uses for Layer 2 services (or Layer 2 Virtual Networks):

* BD1 is an EVPN Broadcast Domain for a given tenant and TS1, TS2, and TS3 are connected to it. The five represented NVEs are attached to BD1 and are connected to the same underlay IP network. That is, each NVE learns the remote NVEs' loopback addresses via underlay routing protocol.

* NVE1 is deployed as a virtual switch in a hypervisor with IP-A as underlay loopback IP address. The rest of the NVEs in Figure 1 are physical switches and TS2/TS3 are multihomed to them. TS1 is a virtual machine, identified by MAC1 and IP1. TS2 and TS3 are physically dual-connected to NVEs; hence, they are normally not considered virtual machines.

* The terms Single-Active and All-Active in Figure 1 refer to the mode in which the TS2 and TS3 are multihomed to the NVEs in BD1. In All-Active mode, all the multihoming links are active and can send or receive traffic. In Single-Active mode, only one link (of the set of links connected to the NVEs) is active.

4.2.1. Auto-Discovery and Auto-Provisioning

Auto-discovery is one of the basic capabilities of EVPN. The provisioning of EVPN components in NVEs is significantly automated, simplifying the deployment of services and minimizing manual operations that are prone to human error.

These are some of the auto-discovery and auto-provisioning capabilities available in EVPN:

* Automation on Ethernet Segments (ESes): An Ethernet Segment is defined as a group of NVEs that are attached to the same Tenant System or network. An Ethernet Segment is identified by an Ethernet Segment Identifier (ESI) in the control plane, but neither the ESI nor the NVEs that share the same Ethernet Segment are required to be manually provisioned in the local NVE.

  - If the multihomed Tenant System or network is running protocols, such as the Link Aggregation Control Protocol (LACP) [IEEE.802.1AX_2014], the Multiple Spanning Tree Protocol (MSTP), G.8032, etc., and all the NVEs in the Ethernet Segment can listen to the protocol PDUs to uniquely identify the multihomed Tenant System/network, then the ESI can be "auto-sensed" or "auto-provisioned" following the guidelines in Section 5 of [RFC7432]. The ESI can also be auto-derived out of other parameters that are common to all NVEs attached to the same Ethernet Segment.

  - As described in [RFC7432], EVPN can also auto-derive the BGP parameters required to advertise the presence of a local Ethernet Segment in the control plane (RT and RD). Local Ethernet Segments are advertised using Ethernet Segment routes, and the ESI-import Route Target used by Ethernet Segment routes can be auto-derived based on the procedures of Section 7.6 of [RFC7432].

  - By listening to other Ethernet Segment routes that match the local ESI and import Route Target, an NVE can also auto-discover the other NVEs participating in the multihoming for the Ethernet Segment.

  - Once the NVE has auto-discovered all the NVEs attached to the same Ethernet Segment, the NVE can automatically perform the Designated Forwarder election algorithm (which determines the NVE that will forward traffic to the multihomed Tenant System/ network). EVPN guarantees that all the NVEs in the Ethernet

Segment have a consistent Designated Forwarder election.

* Auto-provisioning of services: When deploying a Layer 2 service for a tenant in an NVO3 network, all the NVEs attached to the same subnet must be configured with a MAC-VRF and the Broadcast Domain for the subnet, as well as certain parameters for them. Note that if the EVPN service interfaces are VLAN-based or VLAN-bundle, implementations do not normally have a specific provisioning for the Broadcast Domain since, in this case, it is the same construct as the MAC-VRF. EVPN allows auto-deriving as many MAC-VRF parameters as possible. As an example, the MAC-VRF's Route Target and Route Distinguisher for the EVPN routes may be auto-derived. Section 5.1.2.1 of [RFC8365] specifies how to auto-derive a MAC-VRF's Route Target as long as a VLAN-based service interface is implemented. [RFC7432] specifies how to auto-derive the Route Distinguisher.

4.2.2. Remote NVE Auto-Discovery

Auto-discovery via MP-BGP [RFC4760] is used to discover the remote NVEs attached to a given Broadcast Domain, the NVEs participating in a given redundancy group, the tunnel encapsulation types supported by an NVE, etc.

In particular, when a new MAC-VRF and Broadcast Domain are enabled, the NVE will advertise a new Inclusive Multicast Ethernet Tag route. Besides other fields, the Inclusive Multicast Ethernet Tag route will encode the IP address of the advertising NVE, the Ethernet Tag (which is zero in the case of VLAN-based and VLAN-bundle interfaces), and a PMSI Tunnel Attribute (PTA) that indicates the information about the intended way to deliver BUM traffic for the Broadcast Domain.

When BD1 is enabled in the example of Figure 1, NVE1 will send an Inclusive Multicast Ethernet Tag route including its own IP address, an Ethernet-Tag for BD1, and the PMSI Tunnel Attribute to the remote NVEs. Assuming Ingress Replication (IR) is used, the Inclusive Multicast Ethernet Tag route will include an identification for Ingress Replication in the PMSI Tunnel Attribute and the VNI that the other NVEs in the Broadcast Domain must use to send BUM traffic to the advertising NVE. The other NVEs in the Broadcast Domain will import the Inclusive Multicast Ethernet Tag route and will add NVE1's IP address to the flooding list for BD1. Note that the Inclusive Multicast Ethernet Tag route is also sent with a BGP encapsulation attribute [RFC9012] that indicates what NVO3 encapsulation the remote NVEs should use when sending BUM traffic to NVE1.

Refer to [RFC7432] for more information about the Inclusive Multicast Ethernet Tag route and forwarding of BUM traffic. See [RFC8365] for its considerations on NVO3 networks.

4.2.3. Distribution of Tenant MAC and IP Information

Tenant MAC/IP information is advertised to remote NVEs using MAC/IP Advertisement routes. Following the example of Figure 1:

* In a given EVPN Broadcast Domain, the MAC addresses of TSes are first learned at the NVE they are attached to via data path or management plane learning. In Figure 1, we assume NVE1 learns MAC1/IP1 in the management plane (for instance, via Cloud Management System) since the NVE is a virtual switch. NVE2, NVE3, NVE4, and NVE5 are ToR/Leaf switches, and they normally learn MAC addresses via data path.

* Once NVE1's BD1 learns MAC1/IP1, NVE1 advertises that information along with a VNI and Next-Hop IP-A in a MAC/IP Advertisement route. The EVPN routes are advertised using the Route

Distinguisher / Route Targets of the MAC-VRF where the Broadcast Domain belongs. Similarly, all the NVEs in BD1 learn local MAC/IP addresses and advertise them in MAC/IP Advertisement routes.

* The remote NVEs can then add MAC1 to their mapping table for BD1 (BT). For instance, when TS3 sends frames to NVE4 with the destination MAC address = MAC1, NVE4 does a MAC lookup on the Bridge Table that yields IP-A and Label L. NVE4 can then encapsulate the frame into an NVO3 tunnel with IP-A as the tunnel destination IP address and L as the VNI. Note that the MAC/IP Advertisement route may also contain the host's IP address (as shown in the example of Figure 1). While the MAC of the received MAC/IP Advertisement route is installed in the Bridge Table, the IP address may be installed in the Proxy ARP/ND table (if enabled) or in the ARP/IP-VRF tables if the Broadcast Domain has an IRB. See Section 4.7.3 for more information about Proxy ARP/ND and Section 4.3 for more details about IRB and Layer 3 services.

Refer to [RFC7432] and [RFC8365] for more information about the MAC/IP Advertisement route and the forwarding of known unicast traffic.

4.3. EVPN Basic Applicability for Layer 3 Services

[RFC9136] and [RFC9135] are the reference documents that describe how EVPN can be used for Layer 3 services. Inter-subnet forwarding in EVPN networks is implemented via IRB interfaces between Broadcast Domains and IP-VRFs. An EVPN Broadcast Domain corresponds to an IP subnet. When IP packets generated in a Broadcast Domain are destined to a different subnet (different Broadcast Domain) of the same tenant, the packets are sent to the IRB attached to the local Broadcast Domain in the source NVE. As discussed in [RFC9135], depending on how the IP packets are forwarded between the ingress NVE and the egress NVE, there are two forwarding models: Asymmetric and Symmetric.

The Asymmetric model is illustrated in the example of Figure 2, and it requires the configuration of all the Broadcast Domains of the tenant in all the NVEs attached to the same tenant. That way, there is no need to advertise IP Prefixes between NVEs since all the NVEs are attached to all the subnets. It is called "Asymmetric" because the ingress and egress NVEs do not perform the same number of lookups in the data plane. In Figure 2, if TS1 and TS2 are in different subnets and TS1 sends IP packets to TS2, the following lookups are required in the data path: a MAC lookup at BD1's table, an IP lookup at the IP-VRF, a MAC lookup at BD2's table at the ingress NVE1, and only a MAC lookup at the egress NVE. The two IP-VRFs in Figure 2 are not connected by tunnels, and all the connectivity between the NVEs is done based on tunnels between the Broadcast Domains.

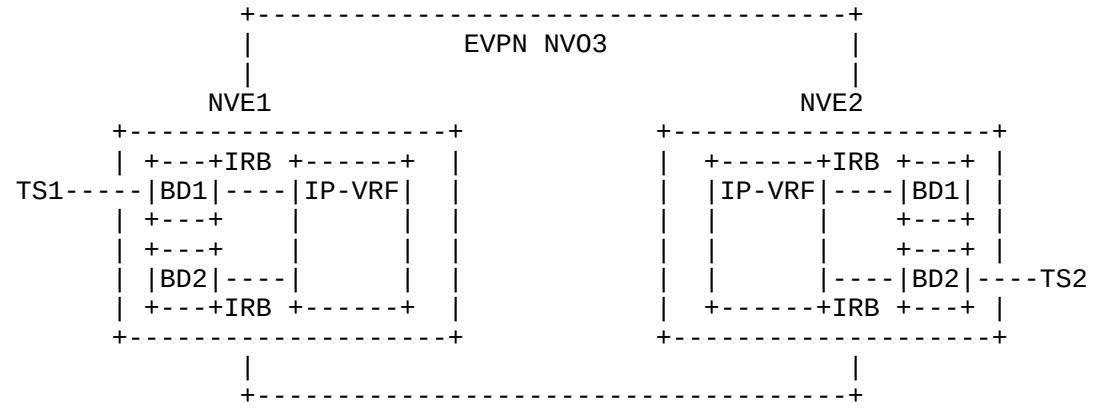

Figure 2: EVPN for L3 in an NVO3 Network - Asymmetric Model

In the Symmetric model, depicted in Figure 3, the same number of data path lookups is needed at the ingress and egress NVEs. For example, if TS1 sends IP packets to TS3, the following data path lookups are required: a MAC lookup at NVE1's BD1 table, an IP lookup at NVE1's IP-VRF, and an IP lookup and MAC lookup at NVE2's IP-VRF and BD3, respectively. In the Symmetric model, the inter-subnet connectivity between NVEs is done based on tunnels between the IP-VRFs.

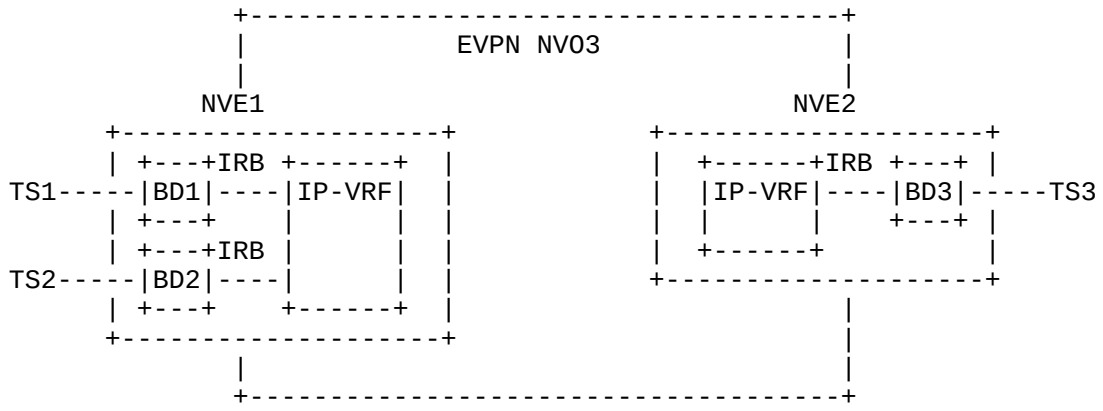

Figure 3: EVPN for L3 in an NVO3 Network - Symmetric Model

The Symmetric model scales better than the Asymmetric model because it does not require the NVEs to be attached to all the tenant's subnets. However, it requires the use of NVO3 tunnels on the IP-VRFs and the exchange of IP Prefixes between the NVEs in the control plane. EVPN uses MAC/IP Advertisement routes for the exchange of host IP routes and IP Prefix routes for the exchange of prefixes of any length, including host routes. As an example, in Figure 3, NVE2 needs to advertise TS3's host route and/or TS3's subnet so that the IP lookup on NVE1's IP-VRF succeeds.

[RFC9135] specifies the use of MAC/IP Advertisement routes for the advertisement of host routes. Section 4.4.1 of [RFC9136] specifies the use of IP Prefix routes for the advertisement of IP Prefixes in an "Interface-less IP-VRF-to-IP-VRF Model". The Symmetric model for host routes can be implemented following either approach:

a. [RFC9135] uses MAC/IP Advertisement routes to convey the information to populate Layer 2, ARP/ND, and Layer 3 Forwarding Information Base tables in the remote NVE. For instance, in Figure 3, NVE2 would advertise a MAC/IP Advertisement route with TS3's IP and MAC addresses and include two labels / VNIs: a label-3/VNI-3 that identifies BD3 for MAC lookup (that would be used for Layer 2 traffic in case NVE1 was attached to BD3 too) and a label-1/VNI-1 that identifies the IP-VRF for IP lookup (that would be used for Layer 3 traffic). NVE1 imports the MAC/IP Advertisement route and installs TS3's IP in the IP-VRF route table with label-1/VNI-1. Traffic, e.g., from TS2 to TS3, would be encapsulated with label-1/VNI-1 and forwarded to NVE2.

b. [RFC9136] uses MAC/IP Advertisement routes to convey the information to populate the Layer 2 Forwarding Information Base, ARP/ND tables, and IP Prefix routes to populate the IP-VRF Layer 3 Forwarding Information Base table. For instance, in Figure 3, NVE2 would advertise a MAC/IP Advertisement route including TS3's MAC and IP addresses with a single label-3/VNI-3. In this example, this MAC/IP Advertisement route wouldn't be imported by NVE1 because NVE1 is not attached to BD3. In addition, NVE2 would advertise an IP Prefix route with TS3's IP address and label-1/VNI-1. This IP Prefix route would be imported by NVE1's IP-VRF and the host route installed in the Layer 3 Forwarding Information Base associated with label-1/VNI-1. Traffic from TS2

to TS3 would be encapsulated with label-1/VNI-1.

4.4. EVPN as Control Plane for NVO3 Encapsulations and Geneve

[RFC8365] describes how to use EVPN for NVO3 encapsulations, such us VXLAN, nvGRE, or MPLSoGRE. The procedures can be easily applicable to any other NVO3 encapsulation, particularly Geneve.

Geneve [RFC8926] is the proposed standard encapsulation specified in the IETF Network Virtualization Overlays Working Group. The EVPN control plane can signal the Geneve encapsulation type in the BGP Tunnel Encapsulation Extended Community (see [RFC9012]).

Geneve requires a control plane [NVO3-ENCAP] to:

* Negotiate a subset of Geneve option TLVs that can be carried on a Geneve tunnel,

* Enforce an order for Geneve option TLVs, and

* Limit the total number of options that could be carried on a Geneve tunnel.

The EVPN control plane can easily extend the BGP Tunnel Encapsulation attribute sub-TLV [RFC9012] to specify the Geneve tunnel options that can be received or transmitted over a Geneve tunnel by a given NVE. [EVPN-GENEVE] describes the EVPN control plane extensions to support Geneve.

4.5. EVPN OAM and Application to NVO3

EVPN Operations, Administration, and Maintenance (OAM), as described in [EVPN-LSP-PING], defines mechanisms to detect data plane failures in an EVPN deployment over an MPLS network. These mechanisms detect failures related to point-to-point (P2P) and Point-to-Multipoint (P2MP) connectivity, for multi-tenant unicast and multicast Layer 2 traffic, between multi-tenant access nodes connected to EVPN PE(s), and in a single-homed, Single-Active, or All-Active redundancy model.

In general, EVPN OAM mechanisms defined for EVPN deployed in MPLS networks are equally applicable for EVPN in NVO3 networks.

4.6. EVPN as the Control Plane for NVO3 Security

EVPN can be used to signal the security protection capabilities of a sender NVE, as well as what portion of an NVO3 packet (taking a Geneve packet as an example) can be protected by the sender NVE, to ensure the privacy and integrity of tenant traffic carried over the NVO3 tunnels [SECURE-EVPN].

4.7. Advanced EVPN Features for NVO3 Networks

This section describes how EVPN can be used to deliver advanced capabilities in NVO3 networks.

4.7.1. Virtual Machine (VM) Mobility

[RFC7432] replaces the classic Ethernet "flood and learn" behavior among NVEs with BGP-based MAC learning. In return, this provides more control over the location of MAC addresses in the Broadcast Domain and consequently advanced features, such as MAC Mobility. If we assume that Virtual Machine (VM) Mobility means the VM’s MAC and IP addresses move with the VM, EVPN’s MAC Mobility is the required procedure that facilitates VM Mobility. According to Section 15 of [RFC7432], when a MAC is advertised for the first time in a Broadcast Domain, all the NVEs attached to the Broadcast Domain will store

Sequence Number zero for that MAC. When the MAC "moves" to a remote NVE within the same Broadcast Domain, the NVE that just learned the MAC locally increases the Sequence Number in the MAC/IP Advertisement route's MAC Mobility extended community to indicate that it owns the MAC now. That makes all the NVEs in the Broadcast Domain change their tables immediately with no need to wait for any aging timer. EVPN guarantees a fast MAC Mobility without flooding or packet drops in the Broadcast Domain.

4.7.2. MAC Protection, Duplication Detection, and Loop Protection

The advertisement of MACs in the control plane allows advanced features such as MAC Protection, Duplication Detection, and Loop Protection.

In a MAC/IP Advertisement route, MAC Protection refers to EVPN's ability to indicate that a MAC must be protected by the NVE receiving the route [RFC7432]. The Protection is indicated in the "Sticky bit" of the MAC Mobility extended community sent along the MAC/IP Advertisement route for a MAC. NVEs' Attachment Circuits that are connected to subject-to-be-protected servers or VMs may set the Sticky bit on the MAC/IP Advertisement routes sent for the MACs associated with the Attachment Circuits. Also, statically configured MAC addresses should be advertised as Protected MAC addresses since they are not subject to MAC Mobility procedures.

MAC Duplication Detection refers to EVPN's ability to detect duplicate MAC addresses [RFC7432]. A "MAC move" is a relearn event that happens at an access Attachment Circuit or through a MAC/IP Advertisement route with a Sequence Number that is higher than the stored one for the MAC. When a MAC moves a number of times (N) within an M-second window between two NVEs, the MAC is declared as a duplicate and the detecting NVE does not re-advertise the MAC anymore.

[RFC7432] provides MAC Duplication Detection, and with an extension, it can protect the Broadcast Domain against loops created by backdoor links between NVEs. The same principle (based on the Sequence Number) may be extended to protect the Broadcast Domain against loops. When a MAC is detected as a duplicate, the NVE may install it as a drop-MAC and discard received frames with source MAC address or the destination MAC address matching that duplicate MAC. The MAC Duplication extension to support Loop Protection is described in Section 15.3 of [RFC7432BIS].

4.7.3. Reduction/Optimization of BUM Traffic in Layer 2 Services

In Broadcast Domains with a significant amount of flooding due to Unknown Unicast and broadcast frames, EVPN may help reduce and sometimes even suppress the flooding.

In Broadcast Domains where most of the broadcast traffic is caused by the Address Resolution Protocol (ARP) and the Neighbor Discovery Protocol (NDP) on the Tenant Systems, EVPN's Proxy ARP and Proxy ND capabilities may reduce the flooding drastically. The use of Proxy ARP/ND is specified in [RFC9161].

Proxy ARP/ND procedures, along with the assumption that Tenant Systems always issue a Gratuitous ARP (GARP) or an unsolicited Neighbor Advertisement message when they come up in the Broadcast Domain, may drastically reduce the Unknown Unicast flooding in the Broadcast Domain.

The flooding caused by Tenant Systems' IGMP / Multicast Listener Discovery (MLD) or PIM messages in the Broadcast Domain may also be suppressed by the use of IGMP/MLD and PIM Proxy functions, as

specified in [RFC9251] and [EVPN-PIM-PROXY]. These two documents also specify how to forward IP multicast traffic efficiently within the same Broadcast Domain, translate soft state IGMP/MLD/PIM messages into hard state BGP routes, and provide fast convergence redundancy for IP multicast on multihomed ESes.

4.7.4. Ingress Replication (IR) Optimization for BUM Traffic

When an NVE attached to a given Broadcast Domain needs to send BUM traffic for the Broadcast Domain to the remote NVEs attached to the same Broadcast Domain, Ingress Replication is a very common option in NVO3 networks since it is completely independent of the multicast capabilities of the underlay network. Also, if the optimization procedures to reduce/suppress the flooding in the Broadcast Domain are enabled (Section 4.7.3) in spite of creating multiple copies of the same frame at the ingress NVE, Ingress Replication may be good enough. However, in Broadcast Domains where Multicast (or Broadcast) traffic is significant, Ingress Replication may be very inefficient and cause performance issues on virtual switch-based NVEs.

[EVPN-OPT-IR] specifies the use of Assisted Replication (AR) NVO3 tunnels in EVPN Broadcast Domains. AR retains the independence of the underlay network while providing a way to forward Broadcast and multicast traffic efficiently. AR uses AR-REPLICATORs that can replicate the broadcast/multicast traffic on behalf of the AR-LEAF NVEs. The AR-LEAF NVEs are typically virtual switches or NVEs with limited replication capabilities. AR can work in a single-stage replication mode (Non-Selective Mode) or in a dual-stage replication mode (Selective Mode). Both modes are detailed in [EVPN-OPT-IR].

In addition, [EVPN-OPT-IR] describes a procedure to avoid sending BUM to certain NVEs that do not need that type of traffic. This is done by enabling Pruned Flood Lists (PFLs) on a given Broadcast Domain. For instance, a virtual switch NVE that learns all its local MAC addresses for a Broadcast Domain via a Cloud Management System does not need to receive the Broadcast Domain’s Unknown Unicast traffic. PFLs help optimize the BUM flooding in the Broadcast Domain.

4.7.5. EVPN Multihoming

Another fundamental concept in EVPN is multihoming. A given Tenant System can be multihomed to two or more NVEs for a given Broadcast Domain, and the set of links connected to the same Tenant System is defined as an ES. EVPN supports Single-Active and All-Active multihoming. In Single-Active multihoming, only one link in the Ethernet Segment is active. In All-Active multihoming, all the links in the Ethernet Segment are active for unicast traffic. Both modes support load-balancing:

* Single-Active multihoming means per-service load-balancing to/from the Tenant System. For example, in Figure 1 for BD1, only one of the NVEs can forward traffic from/to TS2. For a different Broadcast Domain, the other NVE may forward traffic.

* All-active multihoming means per-flow load-balancing for unicast frames to/from the Tenant System. That is, in Figure 1 and for BD1, both NVE4 and NVE5 can forward known unicast traffic to/from TS3. For BUM traffic, only one of the two NVEs can forward traffic to TS3, and both can forward traffic from TS3.

There are two key aspects in the EVPN multihoming procedures:

Designated Forwarder (DF) election:
: The Designated Forwarder is the NVE that forwards the traffic to the Ethernet Segment in Single-Active mode. In the case of All-Active mode, the Designated Forwarder is the NVE that forwards the

BUM traffic to the Ethernet Segment.

Split-horizon function:
Prevents the Tenant System from receiving echoed BUM frames that the Tenant System itself sent to the Ethernet Segment. This is especially relevant in All-Active ESes where the TS may forward BUM frames to a Non-Designated Forwarder NVE that can flood the BUM frames back to the Designated Forwarder NVE and then back to the TS. As an example, assuming NVE4 is the Designated Forwarder for ESI-2 in BD1, Figure 1 shows that BUM frames sent from TS3 to NVE5 will be received at NVE4. NVE4 will forward them back to TS3 since NVE4 is the Designated Forwarder for BD1. Split-horizon allows NVE4 (and any multihomed NVE for that matter) to identify if an EVPN BUM frame is coming from the same Ethernet Segment or a different one. If the frame belongs to the same ESI-2, NVE4 will not forward the BUM frame to TS3 in spite of being the Designated Forwarder.

While [RFC7432] describes the default algorithm for the Designated Forwarder election, [RFC8584] and [EVPN-PREF-DF] specify other algorithms and procedures that optimize the Designated Forwarder election.

The split-horizon function is specified in [RFC7432], and it is carried out by using a special ESI-label that it identifies in the data path with all the BUM frames originating from a given NVE and Ethernet Segment. Since the ESI-label is an MPLS label, it cannot be used in all the non-MPLS NVO3 encapsulations. Therefore, [RFC8365] defines a modified split-horizon procedure that is based on the source IP address of the NVO3 tunnel; it is known as "Local-Bias". It is worth noting that Local-Bias only works for All-Active multihoming, and not for Single-Active multihoming.

4.7.6. EVPN Recursive Resolution for Inter-subnet Unicast Forwarding

Section 4.3 describes how EVPN can be used for inter-subnet forwarding among subnets of the same tenant. MAC/IP Advertisement routes and IP Prefix routes allow the advertisement of host routes and IP Prefixes (IP Prefix route) of any length. The procedures outlined by Section 4.3 are similar to the ones in [RFC4364], but they are only for NVO3 tunnels. However, [RFC9136] also defines advanced inter-subnet forwarding procedures that allow the resolution of IP Prefix routes not only to BGP next hops but also to "overlay indexes" that can be a MAC, a Gateway IP (GW-IP), or an ESI, all of them in the tenant space.

Figure 4 illustrates an example that uses Recursive Resolution to a GW-IP as per Section 4.4.2 of [RFC9136]. In this example, IP-VRFs in NVE1 and NVE2 are connected by a Supplementary Broadcast Domain (SBD). An SBD is a Broadcast Domain that connects all the IP-VRFs of the same tenant via IRB and has no Attachment Circuits. NVE1 advertises the host route TS2-IP/L (IP address and Prefix Length of TS2) in an IP Prefix route with overlay index GW-IP=IP1. Also, IP1 is advertised in a MAC/IP Advertisement route associated with M1, VNI-S, and BGP next-hop NVE1. Upon importing the two routes, NVE2 installs TS2-IP/L in the IP-VRF with a next hop that is the GW-IP IP1. NVE2 also installs M1 in the Supplementary Broadcast Domain, with VNI-S and NVE1 as next hop. If TS3 sends a packet with IP DA=TS2, NVE2 will perform a Recursive Resolution of the IP Prefix route prefix information to the forwarding information of the correlated MAC/IP Advertisement route. The IP Prefix route’s Recursive Resolution has several advantages, such as better convergence in scaled networks (since multiple IP Prefix routes can be invalidated with a single withdrawal of the overlay index route) or the ability to advertise multiple IP Prefix routes from an overlay index that can move or change dynamically. [RFC9136] describes a few

use cases.

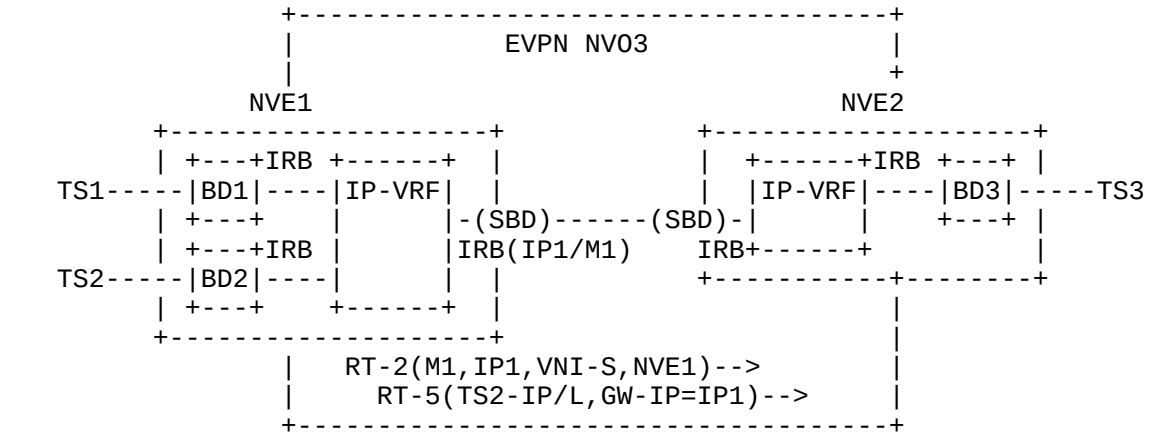

Figure 4: EVPN for L3 - Recursive Resolution Example

4.7.7. EVPN Optimized Inter-subnet Multicast Forwarding

The concept of the Supplementary Broadcast Domain described in Section 4.7.6 is also used in [EVPN-IRB-MCAST] for the procedures related to inter-subnet multicast forwarding across Broadcast Domains of the same tenant. For instance, [EVPN-IRB-MCAST] allows the efficient forwarding of IP multicast traffic from any Broadcast Domain to any other Broadcast Domain (or even to the same Broadcast Domain where the source resides). The [EVPN-IRB-MCAST] procedures are supported along with EVPN multihoming and for any tree allowed on NVO3 networks, including IR or AR. [EVPN-IRB-MCAST] also describes the interoperability between EVPN and other multicast technologies such as Multicast VPN (MVPN) and PIM for inter-subnet multicast.

[EVPN-MVPN-SEAMLESS] describes another potential solution to support EVPN to MVPN interoperability.

4.7.8. Data Center Interconnect (DCI)

Tenant Layer 2 and Layer 3 services deployed on NVO3 networks must often be extended to remote NVO3 networks that are connected via non-NOV3 Wide Area Networks (WANs) (mostly MPLS-based WANs). [RFC9014] defines some architectural models that can be used to interconnect NVO3 networks via MPLS WANs.

When NVO3 networks are connected by MPLS WANs, [RFC9014] specifies how EVPN can be used end to end in spite of using a different encapsulation in the WAN. [RFC9014] also supports the use of NVO3 or Segment Routing (encoding 32-bit or 128-bit Segment Identifiers into labels or IPv6 addresses, respectively) transport tunnels in the WAN.

Even if EVPN can also be used in the WAN for Layer 2 and Layer 3 services, there may be a need to provide a Gateway function between EVPN for NVO3 encapsulations and IP VPN for MPLS tunnels if the operator uses IP VPN in the WAN. [EVPN-IPVPN-INTERWORK] specifies the interworking function between EVPN and IP VPN for unicast inter-subnet forwarding. If inter-subnet multicast forwarding is also needed across an IP VPN WAN, [EVPN-IRB-MCAST] describes the required interworking between EVPN and MVPNs.

5. Security Considerations

This document does not introduce any new procedure or additional signaling in EVPN and relies on the security considerations of the individual specifications used as a reference throughout the document. In particular, and as mentioned in [RFC7432], control

plane and forwarding path protection are aspects to secure in any EVPN domain when applied to NVO3 networks.

[RFC7432] mentions security techniques such as those discussed in [RFC5925] to authenticate BGP messages, and those included in [RFC4271], [RFC4272], and [RFC6952] to secure BGP are relevant for EVPN in NVO3 networks as well.

6. IANA Considerations

This document has no IANA actions.

7. References

7.1. Normative References

[RFC7364] Narten, T., Ed., Gray, E., Ed., Black, D., Fang, L., Kreeger, L., and M. Napierala, "Problem Statement: Overlays for Network Virtualization", RFC 7364, DOI 10.17487/RFC7364, October 2014, <https://www.rfc-editor.org/info/rfc7364>.

[RFC7365] Lasserre, M., Balus, F., Morin, T., Bitar, N., and Y. Rekhter, "Framework for Data Center (DC) Network Virtualization", RFC 7365, DOI 10.17487/RFC7365, October 2014, <https://www.rfc-editor.org/info/rfc7365>.

[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>.

7.2. Informative References

[BGP-PIC] Bashandy, A., Ed., Filsfils, C., and P. Mohapatra, "BGP Prefix Independent Convergence", Work in Progress, Internet-Draft, draft-ietf-rtgwg-bgp-pic-19, 1 April 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-rtgwg-bgp-pic-19>.

[CLOS1953] Clos, C., "A study of non-blocking switching networks", The Bell System Technical Journal, Vol. 32, Issue 2, DOI 10.1002/j.1538-7305.1953.tb01433.x, March 1953, <https://ieeexplore.ieee.org/document/6770468>.

[EVPN-GENEVE] Boutros, S., Ed., Sajassi, A., Drake, J., Rabadan, J., and S. Aldrin, "EVPN control plane for Geneve", Work in Progress, Internet-Draft, draft-ietf-bess-evpn-geneve-06, 26 May 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-bess-evpn-geneve-06>.

[EVPN-IPVPN-INTERWORK] Rabadan, J., Ed., Sajassi, A., Ed., Rosen, E., Drake, J., Lin, W., Uttaro, J., and A. Simpson, "EVPN Interworking with IPVPN", Work in Progress, Internet-Draft, draft-ietf-bess-evpn-ipvpn-interworking-08, 5 July 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-bess-evpn-ipvpn-interworking-08>.

[EVPN-IRB-MCAST] Lin, W., Zhang, Z., Drake, J., Rosen, E., Ed., Rabadan, J., and A. Sajassi, "EVPN Optimized Inter-Subnet Multicast (OISM) Forwarding", Work in Progress, Internet-Draft, draft-ietf-bess-evpn-irb-mcast-09, 21 February 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-bess-

evpn-irb-mcast-09>.

[EVPN-LSP-PING]
Jain, P., Sajassi, A., Salam, S., Boutros, S., and G. Mirsky, "LSP-Ping Mechanisms for EVPN and PBB-EVPN", Work in Progress, Internet-Draft, draft-ietf-bess-evpn-lsp-ping-11, 29 May 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-bess-evpn-lsp-ping-11>.

[EVPN-MVPN-SEAMLESS]
Sajassi, A., Thiruvenkatasamy, K., Thoria, S., Gupta, A., and L. Jalil, "Seamless Multicast Interoperability between EVPN and MVPN PEs", Work in Progress, Internet-Draft, draft-ietf-bess-evpn-mvpn-seamless-interop-05, 13 March 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-bess-evpn-mvpn-seamless-interop-05>.

[EVPN-OPT-IR]
Rabadan, J., Ed., Sathappan, S., Lin, W., Katiyar, M., and A. Sajassi, "Optimized Ingress Replication Solution for Ethernet VPN (EVPN)", Work in Progress, Internet-Draft, draft-ietf-bess-evpn-optimized-ir-12, 25 January 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-bess-evpn-optimized-ir-12>.

[EVPN-PIM-PROXY]
Rabadan, J., Ed., Kotalwar, J., Sathappan, S., Zhang, Z., and A. Sajassi, "PIM Proxy in EVPN Networks", Work in Progress, Internet-Draft, draft-skr-bess-evpn-pim-proxy-01, 30 October 2017, <https://datatracker.ietf.org/doc/html/draft-skr-bess-evpn-pim-proxy-01>.

[EVPN-PREF-DF]
Rabadan, J., Ed., Sathappan, S., Lin, W., Drake, J., and A. Sajassi, "Preference-based EVPN DF Election", Work in Progress, Internet-Draft, draft-ietf-bess-evpn-pref-df-11, 6 July 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-bess-evpn-pref-df-11>.

[IEEE.802.1AX_2014]
IEEE, "IEEE Standard for Local and metropolitan area networks -- Link Aggregation", IEEE Std 802.1AX-2014, DOI 10.1109/IEEESTD.2014.7055197, December 2014, <https://doi.org/10.1109/IEEESTD.2014.7055197>.

[NVO3-ENCAP]
Boutros, S., Ed. and D. Eastlake 3rd, Ed., "Network Virtualization Overlays (NVO3) Encapsulation Considerations", Work in Progress, Internet-Draft, draft-ietf-nvo3-encap-09, 7 October 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-nvo3-encap-09>.

[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, <https://www.rfc-editor.org/info/rfc4271>.

[RFC4272] Murphy, S., "BGP Security Vulnerabilities Analysis", RFC 4272, DOI 10.17487/RFC4272, January 2006, <https://www.rfc-editor.org/info/rfc4272>.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February

2006, <https://www.rfc-editor.org/info/rfc4364>.

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

[RFC5925] Touch, J., Mankin, A., and R. Bonica, "The TCP Authentication Option", RFC 5925, DOI 10.17487/RFC5925, June 2010, <https://www.rfc-editor.org/info/rfc5925>.

[RFC6952] Jethanandani, M., Patel, K., and L. Zheng, "Analysis of BGP, LDP, PCEP, and MSDP Issues According to the Keying and Authentication for Routing Protocols (KARP) Design Guide", RFC 6952, DOI 10.17487/RFC6952, May 2013, <https://www.rfc-editor.org/info/rfc6952>.

[RFC7348] Mahalingam, M., Dutt, D., Duda, K., Agarwal, P., Kreeger, L., Sridhar, T., Bursell, M., and C. Wright, "Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks", RFC 7348, DOI 10.17487/RFC7348, August 2014, <https://www.rfc-editor.org/info/rfc7348>.

[RFC7432BIS] Sajassi, A., Burdet, L., Drake, J., and J. Rabadan, "BGP MPLS-Based Ethernet VPN", Work in Progress, Internet-Draft, draft-ietf-bess-rfc7432bis-07, 13 March 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-bess-rfc7432bis-07>.

[RFC7510] Xu, X., Sheth, N., Yong, L., Callon, R., and D. Black, "Encapsulating MPLS in UDP", RFC 7510, DOI 10.17487/RFC7510, April 2015, <https://www.rfc-editor.org/info/rfc7510>.

[RFC8365] Sajassi, A., Ed., Drake, J., Ed., Bitar, N., Shekhar, R., Uttaro, J., and W. Henderickx, "A Network Virtualization Overlay Solution Using Ethernet VPN (EVPN)", RFC 8365, DOI 10.17487/RFC8365, March 2018, <https://www.rfc-editor.org/info/rfc8365>.

[RFC8584] Rabadan, J., Ed., Mohanty, S., Ed., Sajassi, A., Drake, J., Nagaraj, K., and S. Sathappan, "Framework for Ethernet VPN Designated Forwarder Election Extensibility", RFC 8584, DOI 10.17487/RFC8584, April 2019, <https://www.rfc-editor.org/info/rfc8584>.

[RFC8926] Gross, J., Ed., Ganga, I., Ed., and T. Sridhar, Ed., "Geneve: Generic Network Virtualization Encapsulation", RFC 8926, DOI 10.17487/RFC8926, November 2020, <https://www.rfc-editor.org/info/rfc8926>.

[RFC9012] Patel, K., Van de Velde, G., Sangli, S., and J. Scudder, "The BGP Tunnel Encapsulation Attribute", RFC 9012, DOI 10.17487/RFC9012, April 2021, <https://www.rfc-editor.org/info/rfc9012>.

[RFC9014] Rabadan, J., Ed., Sathappan, S., Henderickx, W., Sajassi, A., and J. Drake, "Interconnect Solution for Ethernet VPN (EVPN) Overlay Networks", RFC 9014, DOI 10.17487/RFC9014, May 2021, <https://www.rfc-editor.org/info/rfc9014>.

[RFC9135] Sajassi, A., Salam, S., Thoria, S., Drake, J., and J. Rabadan, "Integrated Routing and Bridging in Ethernet VPN (EVPN)", RFC 9135, DOI 10.17487/RFC9135, October 2021,

<https://www.rfc-editor.org/info/rfc9135>.

[RFC9136] Rabadan, J., Ed., Henderickx, W., Drake, J., Lin, W., and A. Sajassi, "IP Prefix Advertisement in Ethernet VPN (EVPN)", RFC 9136, DOI 10.17487/RFC9136, October 2021, <https://www.rfc-editor.org/info/rfc9136>.

[RFC9161] Rabadan, J., Ed., Sathappan, S., Nagaraj, K., Hankins, G., and T. King, "Operational Aspects of Proxy ARP/ND in Ethernet Virtual Private Networks", RFC 9161, DOI 10.17487/RFC9161, January 2022, <https://www.rfc-editor.org/info/rfc9161>.

[RFC9251] Sajassi, A., Thoria, S., Mishra, M., Patel, K., Drake, J., and W. Lin, "Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Proxies for Ethernet VPN (EVPN)", RFC 9251, DOI 10.17487/RFC9251, June 2022, <https://www.rfc-editor.org/info/rfc9251>.

[SECURE-EVPN] Sajassi, A., Banerjee, A., Thoria, S., Carrel, D., Weis, B., and J. Drake, "Secure EVPN", Work in Progress, Internet-Draft, draft-ietf-bess-secure-evpn-00, 20 June 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-bess-secure-evpn-00>.

Acknowledgments

The authors thank Aldrin Isaac for his comments.

Authors' Addresses

Jorge Rabadan (editor)
Nokia
520 Almanor Ave
Sunnyvale, CA 94085
United States of America
Email: jorge.rabadan@nokia.com

Matthew Bocci
Nokia
Email: matthew.bocci@nokia.com

Sami Boutros
Ciena
Email: sboutros@ciena.com

Ali Sajassi
Cisco Systems, Inc.
Email: sajassi@cisco.com