EVPN
Next Generation of L2 VPNs

Jack W. Parks, IV
Advanced Technologies
Juniper Networks
jwparks@juniper.net

Legal Disclaimer: This statement of product direction sets forth Juniper Networks’ current intention, and is subject to change at any time without notice. No purchases are contingent upon Juniper Networks delivering any feature or functionality depicted on this statement.
Data Centers are extending beyond traditional boundaries due to:

- Extending Operating System, File System clusters, Database clusters
- Virtual/Physical machine mobility due to load sharing, disaster prevention
- Legacy devices/applications with embedded IP addressing
- Time to deployment and operational reasons
- Extend DC to solve power/heat/space limitations
DCI VPLS DEPLOYMENT OPTIONS

VPLS Multi-Homing

- >1 VPLS devices
- VPLS controlled Active-Standby
- Per VLAN

VPLS with MC-LAG Active-Standby

- >1 VPLS devices
- MC-LAG controlled Active-Standby on LAN
- Per VLAN

VPLS with MX Virtual Chassis

- One VPLS device
- Active forwarding through all links of LAG
EVPN Overview
NEW TERMS

**EVI:** An EVPN instance spanning across the PEs participating in that VPN

**MAC-VRF:** A Virtual Routing and Forwarding table for MAC addresses on a PE for an EVI

**Ethernet Segment Identifier (ESI):** If a CE is multi-homed to two or more PEs, the set of Ethernet links that attaches the CE to the PEs is an 'Ethernet segment'. Ethernet segments MUST have a unique non-zero identifier, the 'Ethernet Segment Identifier'.

**Ethernet A-D:** Ethernet Auto-Discovery Route is a specific EVPN NLRI.
EVPN REQUIREMENTS

EVPN provides VLAN Extension over a shared IP/MPLS network.
*Improves on VPLS*

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>All-Active Multi-Homing</td>
<td>All available paths should be used (CE-PE, PE-PE)</td>
</tr>
<tr>
<td>Better Control Over MAC Learning</td>
<td>MAC learning happens in control plane</td>
</tr>
<tr>
<td>ARP/ND Flooding Minimization</td>
<td>Additional attributes added during MAC advertisement</td>
</tr>
<tr>
<td>L3 Egress Traffic Forwarding</td>
<td>Usage of Default Gateway Extended Community</td>
</tr>
<tr>
<td>Reducing Unknown Unicast Flooding</td>
<td>By using MAC learning in control plane</td>
</tr>
</tbody>
</table>
ADVANTAGES OF EVPN

Traffic Engineering, HA, fast recovery (transport)

MPLS or IP

High scale multi-tenancy across common transport (service tag)

All Active forwarding links on WAN and LAN (LAG)

Control Plane based information exchange / control (policy based control)
EVPN FORWARDING OVERVIEW

MX Series

MPLS or IP

detours

BGP Control Plane based learning on WAN

DP learning over LAN

Hash based LB on Ethernet switch

P2P connections for unicast traffic

MPLS transport label(s) including detour or IP transport label

Service label

Ethernet Frame

P2MP connections for multicast or unknown traffic

MAC1..............LAN Ports
MAC11.............MPLS nexthop

MAC2..............LAN Ports
MAC22.............MPLS nexthop

VLAN 1
MAC1

VLAN 2
MAC2

MX Series

LAG

VLAN 1
MAC11

VLAN 2
MAC22

MX Series

LAG

Copyright © 2009 Juniper Networks, Inc. www.juniper.net
EVPN INFORMATION EXCHANGE OVERVIEW

- EVPN advertises MAC (L2) and IP (ARP) bindings for each segment along with service tags
  - Allowing Control Plane based L2 and ARP learning
  - Minimizes flooding across WAN
  - Allows proxy-ARP to respond queries locally
- IRB MAC address exchange allows same gateway MAC address across sites
  - VM mobility: egress traffic optimization

EVPN reachability advertisement

- Route Distinguisher
- ESI
- Ethernet Tag
- MAC Address
- IPv4 or IPv6 Address
- Service Tag
EVPN AUTO-DISCOVERY CAPABILITY

- EVPN has built-in auto discovery for ease of configuration
- Ethernet Segment Identifier (ESI) allows multi-homing of EVPN routers on the same site
  - Built-in L2 loop prevention mechanism w/out blocking any forwarding interfaces
  - Built-in Split Horizon for L2 BUM
- Auto-discovery of Ethernet Tags (VLANs) on Ethernet Segment
- A Designated Forwarder (DF) is elected (can be per VLAN)
  - Other DCB becomes a backup designated forwarder (BDF)
  - Required for L2 BUM

Auto Discovery message per L2 Segment

- Route Distinguisher
- ESI
- Ethernet Tag
- Service Tag
EVPN Specifics
ETHERNET SEGMENT IDENTIFIER (ESI)

If CE is multi-homed to two or more PEs, the set of Ethernet links constitutes an “Ethernet Segment”.

Only A/P multi-homing is supported in 13.2R1. 
A/A support is roadmap

An Ethernet Segment MUST have a non-reserved ESI that is unique network wide. ESI can be auto-provisioned (roadmap)
EVPN L2 LOOP ELIMINATION CAPABILITIES

EVPN provides Active-Active and Active-Standby multi-homing options

Built-in L2 Loop Prevention capabilities

- Ethernet Segment Identifier (ESID)
  - Per VLAN / BD on the CE facing interface
  - Needed for all multi-homed deployments – to identify a (virtual) LAN instance

- Designated Forwarder Function
  - DF elected for a given ESID – designated L2-BUM authority
  - DF generates a (Split Horizon) MPLS label and distributes to all PEs – for each ESID
  - Non-DFs can send L2 BUM to MPLS network
    - Using this Split Horizon MPLS label – DFs identify own ESID and drop the packet
  - DFs send L2 BUM to MPLS network
    - Non-DFs drop the L2-BUM by default

- LAG
  - Required for Active-Active multi-homing
  - CE based loop prevention, single L2-BUM packet forwarding function
LOAD BALANCE TRAFFIC ACTIVE/ACTIVE PEs (roadmap)

EVPN introducing a concept of Aliasing.

Each PE signals that it has reachability to a given Ethernet segment (using Ethernet A-D Route)

Remote PE should install all PEs as next-hop which are attached to the same Ethernet Segment
FAST CONVERGENCE IN ACTIVE/BACKUP (Roadmap)

EVPN introducing a concept of Backup-Path.

Each PE signals that it has reachability to a given Ethernet segment (using Ethernet A-D Route)

Remote PE should install backup paths to all further PEs which have reachability to particular Ethernet Segment
ARP PROXY *(Roadmap)*

PE can snoop ARP messages for locally attached hosts.

MAC/IP binding can be then redistributed to other PEs by using MAC Advertisement Route.
IRB support within EVPN (roadmap)

IRB allows to forward not only L2 but L3 traffic as well on the same PE

In case of multiple locations (e.g. DC locations) it is desired to use local forwarding for L3 traffic to avoid trombone effect

Each PE that acts as a Default GW for a given EVPN should advertise its Default GW IP and MAC address using MAC Advertisement Route (with Default Gateway Extended Community).

All receiving PE should reply to all ARP requests received to this IP address and should forward traffic destined to this MAC address locally
# VPLS VS EVPN

<table>
<thead>
<tr>
<th>Desirable L2 extension attributes</th>
<th>VPLS</th>
<th>E- VPN</th>
</tr>
</thead>
<tbody>
<tr>
<td>VM Mobility without renumbering L2 and L3 addresses</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Ability to span VLANs across racks in different locations</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Scale to few 100K hosts within and across multiple DCs</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Policy-based flexible L2 topologies similar to L3 VPNs</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Multiple points of attachment with ability to load-balance VLANs</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Active-Active points of attachment with ability to load-balance flows</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Multi-tenant support (secure isolation, overlapping MAC, IP addresses)</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Control-Plane Based Learning</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Minimize or eliminate flooding of unknown unicast</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Fast convergence from edge failures based on local repair</td>
<td></td>
<td>✔</td>
</tr>
</tbody>
</table>
Standards Track – IETF

BGP MPLS Based Ethernet VPN


Requirements for Ethernet VPN (EVPN)

everywhere