TOMORROW starts here.
Deploying MPLS-based Layer 2 Virtual Private Networks

BRKMPL-2101

Jose Liste – Technical Marketing Engineer
jliste@cisco.com
Agenda

• Motivation and Overview
• Ethernet Point-to-Point L2VPNs
• Ethernet Multipoint L2VPNs
  – VPLS
  – EVPN and PBB-EVPN
• Advanced Topics
  – Resiliency Solutions
  – Load-Balancing
• Deployment Use Cases
• Summary
L2VPN Motivation and Overview
# What is a Layer 2 VPN?

## L2VPN Transport Services

<table>
<thead>
<tr>
<th>TDM</th>
<th>ATM</th>
<th>Frame Relay</th>
<th>Ethernet</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Point-to-Point L2VPN</strong></td>
<td></td>
<td><strong>Multipoint L2VPN</strong></td>
<td></td>
</tr>
<tr>
<td>Circuit Emulation Service over PSN (CESoPSN)</td>
<td>AAL5 over Pseudowire</td>
<td>FR over Pseudowire</td>
<td>Ethernet Virtual Private Line (EVPL)</td>
</tr>
<tr>
<td>Muxed UNI</td>
<td>Muxed UNI</td>
<td>Muxed UNI</td>
<td>Muxed UNI</td>
</tr>
<tr>
<td>Structure Agnostic TDM over Packet (SAToP)</td>
<td>Cell Relay with Packing over Pseudowire</td>
<td>PPP/HDLC over Pseudowire</td>
<td>Ethernet Private LAN (EPLAN)</td>
</tr>
<tr>
<td>Muxed UNI</td>
<td>Muxed UNI</td>
<td>Unmuxed UNI</td>
<td>Unmuxed UNI</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Motivation for L2VPNs

Old and New Drivers

• Network Consolidation (circa 2000)
  – Multiple access services (FR, ATM, TDM) required multiple core technologies

• Enterprise Ethernet WAN Connectivity Services (circa 2005+)
  – Ethernet well understood by Enterprise / SPs
  – CAPEX (lower cost per bit) / Growth (100GE)
  – Layer 2 VPN replacement to ATM/Frame Relay
  – Internet / Layer 3 VPN access (CE to PE)

• Data Center Interconnection (DCI)

• Mobile Backhaul Evolution
  – TDM /PDH to Dual/Hybrid to All-packet (IP/Ethernet)
  – Single (voice + data) IP/Ethernet mobile backhaul universally accepted solution
MPLS Layer-2 Virtual Private Networks

Technology Options

• VPWS services
  – Point-to-point
  – Referred to as Pseudowires (PWs)

• VPLS services
  – Multipoint

• EVPN
  – Multipoint with BGP-based MAC learning

• PBB-EVPN
  – Combines scale tools from PBB (aka MAC-in-MAC) with BGP-based MAC learning from EVPN
Ethernet Point-to-Point L2VPNs
Virtual Private Wire Service (VPWS)
Layer 2 VPN Enabler

The Pseudowire

- L2VPNs are built with Pseudowire (PW) technology
- PWs provide a common intermediate format to transport multiple types of network services over a Packet Switched Network (PSN)
- PW technology provides Like-to-Like transport and also Interworking (IW)
Pseudowire Reference Model

- **Any Transport Over MPLS** (AToM) is Cisco’s implementation of VPWS for IP/MPLS networks
- An **Attachment Circuit** (AC) is the physical or virtual circuit attaching a CE to a PE
- Customer Edge (CE) equipment perceives a PW as an unshared link or circuit

Ref: RFC 3985 Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture, March 2005
Layer 2 Transport over MPLS

Control Connection
- Targeted LDP session / BGP session / Static
  - Used for VC-label negotiation, withdrawal, error notification

The “emulated circuit” has three (3) layers of encapsulation

Tunnelling Component
- Tunnel header (Tunnel Label)
  - To get PDU from ingress to egress PE
  - MPLS LSP derived through static configuration (MPLS-TP) or dynamic (LDP or RSVP-TE)

Demultiplexing Component
- Demultiplexer field (VC Label)
  - To identify individual circuits within a tunnel
  - Could be an MPLS label, L2TPv3 header, GRE key, etc.

Layer 2 Encapsulation
- Emulated VC encapsulation (Control Word)
  - Information on enclosed Layer 2 PDU
  - Implemented as a 32-bit control word
How Are Ethernet Frames Transported?

- Ethernet frames transported without Preamble, Start Frame Delimiter (SFD) and FCS
- Two (2) modes of operation supported:
  - Ethernet VLAN mode (VC type 0x0004) – created for VLAN over MPLS application
  - Ethernet Port / Raw mode (VC type 0x0005) – created for Ethernet port tunneling application
VPWS Traffic Encapsulation

- Three-level encapsulation
- Packets switched between PEs using **Tunnel label**
- **VC label** identifies PW
- **VC label** signaled between PEs
- Optional **Control Word** (CW) carries Layer 2 control bits and enables sequencing

<table>
<thead>
<tr>
<th>Control Word</th>
<th>Encap.</th>
<th>Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATM N:1 Cell Relay</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>ATM AAL5</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Ethernet</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Frame Relay</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>HDLC</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>PPP</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>SAToP</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>CESoPSN</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>
VPWS

Discovery and Signaling Alternatives

• VPWS Signaling
  – LDP-based (RFC 4447)
  – BGP-based (RFC 6624)
• VPWS with LDP-signaling and No auto-discovery
  – Most widely deployed solution
• Auto-discovery for point-to-point services not as relevant as for multipoint
LDP Signaling

Overview

- RFC 4447 defines use of LDP protocol for setting up and maintaining pseudowires
  - Targeted LDP (t-LDP) session between PE routers
- PW label bindings exchanged using LDP Label Mapping messages
- Two Forward Equivalency Classes (FEC) element types defined
  - LDP PWid FEC Element (FEC 128) - Used in manual provisioning scenarios
  - LDP Generalized PWid FEC Element (FEC 129) – Used in auto-discovery scenarios
PW Control Plane Operation

LDP Signaling

1. PW manually provisioned – Remote PE info included.
   - CE-1
     - Interface A
   - CE-2
     - Interface B

2. New targeted LDP session between PE routers established, in case one does not already exist.
   - PE-1
   - PE-2
   - MPLS
   - Local_int = A
   - Remote PE = PE2_ip
   - VC-id <123>
   - Local_int = B
   - Remote PE = PE1_ip
   - VC-id <123>

3. PEs assigns local VC label to PW.
   - Local Label: X
   - Remote Label: Y

4. PEs advertize local VC label using LDP label-mapping message: Label TLV + PW FEC TLV

5. PEs bind remote label for PW with matching VC-id.
   - Local Label: Y
   - Remote Label: X

PW manually provisioned – Remote PE info included.
VPWS Forwarding Plane Processing

Traffic direction

Tunnel label swapping through MPLS cloud

Penultimate Hop Popping (PHP)

VC label disposition

Push Push

VPWS Forwarding Plane Processing

Payload

Label = 34

Label = 28

Payload

Label = 45

Label = 28

Payload

Label = 28

Payload

Cisco Public
Ethernet Multipoint L2VPNs
Virtual Private LAN Service (VPLS)
Virtual Private LAN Service

Overview

• Defines Architecture to provide Ethernet Multipoint connectivity sites, as if they were connected using a LAN

• VPLS operation emulates an IEEE Ethernet switch

• Two (2) signaling methods
  – RFC 4762 (LDP-Based VPLS)
  – RFC 4761 (BGP-Based VPLS)
Virtual Private LAN Service

Reference Model

- **VFI (Virtual Forwarding Instance)**
  - Also called VSI (Virtual Switching Instance)
  - Emulates L2 broadcast domain among ACs and VCs
  - Unique per service. Multiple VFIs can exist same PE

- **AC (Attachment Circuit)**
  - Connect to CE device, it could be Ethernet physical or logical port
  - One or multiple ACs can belong to same VFI

- **VC (Virtual Circuit)**
  - EoMPLS data encapsulation, tunnel label used to reach remote PE, VC label used to identify VFI
  - One or multiple VCs can belong to same VFI
  - PEs must have a full-mesh of PWs in the VPLS core
Virtual Private LAN Service

Operation

- **Flooding / Forwarding**
  - Forwarding based on destination MAC addresses
  - Flooding (Broadcast, Multicast, Unknown Unicast)

- **Split-Horizon and Full-Mesh of PWs for loop-avoidance in core**
  - SP does not run STP in the core

- **MAC Learning/Aging/Withdrawal**
  - Dynamic learning based on Source MAC and VLAN
  - Refresh aging timers with incoming packet
  - MAC withdrawal upon topology changes
Why H-VPLS? Improved Scaling

• Flat VPLS
  – Potential signaling overhead
  – Packet replication at the edge
  – Full PW mesh end-to-end

• Hierarchical-VPLS
  – Minimizes signaling overhead
  – Packet replication at the core only
  – Full PW mesh in the core
VPLS Operation

Loop Prevention

• Core PW – Split Horizon ON
• Spoke PW – Split Horizon OFF (default)
• Split-Horizon Rules
  – Forwarding between Spoke PWs
  – Forwarding between Spoke and Core PWs
  – Forwarding between ACs and Core / Spoke PWs
  – Forwarding between ACs
  – Blocking between Core PWs
VPLS Operation

MAC Address Withdrawal

• Remove (flush) dynamic MAC addresses upon Topology Changes
  – Faster convergence – avoids blackholing
  – Uses LDP Address Withdraw Message (RFC 4762)

• H-VPLS dual-home example
  – U-PE detects failure of Primary PW
  – U-PE activates Backup PW
  – U-PE sends LDP MAC address withdrawal request to new N-PE
  – N-PE forwards the message to all PWs in the VPLS core and flush its MAC address table
Ethernet Multipoint L2VPNs
VPLS Signaling and Auto-Discovery
VPLS

Discovery and Signaling Alternatives

• VPLS Signaling
  – LDP-based (RFC 4762)
  – BGP-based (RFC 4761)

• VPLS with LDP-signaling and No auto-discovery
  – Most widely deployed solution
  – Operational complexity for larger deployments

• BGP-based Auto-Discovery (BGP-AD) (RFC 6074)
  – Enables discovery of PE devices in a VPLS instance
Ethernet Multipoint L2VPNs

VPLS with LDP Signaling and BGP-based AutoDiscovery (BGP-AD)
BGP Auto-Discovery (BGP-AD)

- Eliminates need to manually provision VPLS neighbors

- Automatically detects when new PEs are added / removed from the VPLS domain

- Uses BGP Update messages to advertise PE/VFI mapping (VPLS NLRI)

- Typically used in conjunction with BGP Route Reflectors to minimize iBGP full-mesh peering requirements

- Two (2) RFCs define use of BGP for VPLS AD$^1$
  - RFC 6074 – when LDP used for PW signaling
  - RFC 4761 – when BGP used for PW signaling

(1) VPLS BGP NLRIs from RFC 6074 and 4761 are different in format and thus not compatible, even though they share same AFI / SAFI values
BGP Auto-Discovery in Action

1. L2 VFI with VPN ID 100 is configured PEs on BGP AS 20

2. Packet Format:
   - RD <AS>:<VPN-ID>
   - Prefix: <L2 Router ID>
   - Extended Community
     - RT <AS>:<VPN-ID>
     - VPLS-id <AS>:<VPN-ID>

3. BGP Update message
   - RD 20:100:10.0.0.1/96
   - RT 20:100 VPLS-id 20:100
   - Next Hop: 10.0.0.1

4. I learned about 20:100:10.0.0.1/96 !!!

5. Found matching RT !!!

6. Import NLRI...
   - VPN ID: 100
   - Neighbor LDP ID: 10.0.0.1

7. LDP signaling phase starts
   - Using auto-discovered L2VPN router IDs for a common VPLS-id

PE1
- BGP subsystem
- L2VPN subsystem

PE2
- BGP subsystem
- L2VPN subsystem

Cisco Public
BRKMPL-2101 © 2014 Cisco and/or its affiliates. All rights reserved.
What is Discovered? NLRI + Extended Communities

<table>
<thead>
<tr>
<th>NLRI</th>
<th>Extended Communities</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Source Address</strong> = 1.1.1.10</td>
<td><strong>Source Address</strong> = 2.2.2.20</td>
</tr>
<tr>
<td><strong>Destination Address</strong> = 2.2.2.20</td>
<td><strong>Destination Address</strong> = 1.1.1.10</td>
</tr>
<tr>
<td><strong>Length</strong> = 14</td>
<td><strong>Length</strong> = 14</td>
</tr>
<tr>
<td><strong>Route Distinguisher</strong> = 100:111</td>
<td><strong>Route Distinguisher</strong> = 100:111</td>
</tr>
<tr>
<td><strong>L2VPN Router ID</strong> = 10.10.10.10</td>
<td><strong>L2VPN Router ID</strong> = 20.20.20.20</td>
</tr>
<tr>
<td><strong>VPLS-ID</strong> = 100:111</td>
<td><strong>VPLS-ID</strong> = 100:111</td>
</tr>
<tr>
<td><strong>Route Target</strong> = 100:111</td>
<td><strong>Route Target</strong> = 100:111</td>
</tr>
</tbody>
</table>

**BGP Update Messages**

- **BGP ASN** = 100
- **BGP Rtr ID** = 1.1.1.10
- **BGP neighbor** = 2.2.2.20

- **L2VPN Rtr ID** = 10.10.10.10
- **VPN ID** = 111
- **RT** = auto (100:111)
- **RD** = auto (100:111)
- **VPLS-ID** = auto (100:111)

- **BGP ASN** = 100
- **BGP Rtr ID** = 2.2.2.20
- **BGP neighbor** = 1.1.1.10

- **L2VPN Rtr ID** = 20.20.20.20
- **VPN ID** = 111
- **RT** = auto (100:111)
- **RD** = auto (100:111)
- **VPLS-ID** = auto (100:111)
What is Signaled?

FEC 129

Local and Remote (discovered) L2VPN router ID and VPLS-ID used for PW signaling

BGP ASN = 100
BGP Rtr ID = 1.1.1.10
BGP neighbor = 2.2.2.20

L2VPN Rtr ID = 10.10.10.10
VPN ID = 111
RT = auto (100:111)
RD = auto (100:111)
VPLS-ID = auto (100:111)

BGP ASN = 100
BGP Rtr ID = 2.2.2.20
BGP neighbor = 1.1.1.10

L2VPN Rtr ID = 20.20.20.20
VPN ID = 111
RT = auto (100:111)
RD = auto (100:111)
VPLS-ID = auto (100:111)

AGI = VPLS-ID = 100:111
SAII = Local L2VPN ID = 10.10.10.10
TAII = Remote L2VPN ID = 20.20.20.20

AGI = VPLS-ID = 100:111
SAII = Local L2VPN ID = 20.20.20.20
TAII = Remote L2VPN ID = 10.10.10.10

LDP Generalized PwId
FEC Element (FEC 129)
Ethernet Multipoint L2VPNs
VPLS with BGP-based Signaling and AutoDiscovery
BGP Signaling and Auto-Discovery

Overview

- RFC 4761\(^1\) defines use of BGP for VPLS PE Auto-Discovery and Signaling

- All PEs within a given VPLS are assigned a unique VPLS Edge device ID (VE ID)

- A PE X wishing to send a VPLS update sends the same label block information to all other PEs using BGP VPLS NLRI

- Each receiving PE infers the label intended for PE X by adding its (unique) VE ID to the label base
  - Each receiving PE gets a unique label for PE X for that VPLS

---

(1) VPLS BGP NLRI from RFC 6074 and 4761 are different in format and thus not compatible, even though they share same AFI / SAFI values
BGP Signaling and Auto-Discovery

Label Blocks

- RFC 4761 is primarily based on the concept of Label Blocks
  - Contiguous set of local labels
  - Label Block boundary advertised using BGP VPLS NLRI
- Label Base (LB) – start of label block
- VE Block Size (VBS) – size of label block
- VE Block Offset (VBO) – start of remote VE set

![Label Blocks Diagram]

VE ID (VBO + n) corresponds to Label (LB + n)
Ethernet Multipoint L2VPNs

Ethernet VPN Family Overview

Topic covered in detail in BRKMPL-2333 (THU)
What is xEVPN?

- xEVPN family introduces next generation solutions for Ethernet services
  - BGP control-plane for Ethernet Segment and MAC distribution and learning over MPLS core
  - Same principles and operational experience of IP VPNs

- No use of Pseudowires
  - Uses MP2P tunnels for unicast
  - Multi-destination frame delivery via ingress replication (via MP2P tunnels) or LSM

- Multi-vendor solutions under IETF standardization
Ethernet VPN

Highlights

• Next generation solution for Ethernet multipoint (E-LAN) services

• PEs run Multi-Protocol BGP to advertise & learn Customer MAC addresses (C-MACs) over Core
  – Same operational principles of L3VPN

• Learning on PE Access Circuits via data-plane transparent learning

• No pseudowire full-mesh required
  – Unicast: use MP2P tunnels
  – Multicast: use ingress replication over MP2P tunnels or use LSM

• Under standardization at IETF – draft-ietf-l2vpn-evpn
### Concepts

#### EVPN Instance (EVI)
- EVI identifies a VPN in the network
- Encompass one or more bridge-domains, depending on service interface type
  - Port-based
  - VLAN-based (shown above)
  - VLAN-bundling
  - VLAN aware bundling (NEW)

#### Ethernet Segment
- Represents a ‘site’ connected to one or more PEs
- Uniquely identified by a 10-byte global Ethernet Segment Identifier (ESI)
- Could be a single device or an entire network
  - Single-Homed Device (SHD)
  - Multi-Homed Device (MHD)
- Single-Homed Network (SHN)
- Multi-Homed Network (MHN)

#### BGP Routes
- EVPN and PBB-EVPN define a single new BGP NLRI used to carry all EVPN routes
- NLRI has a new SAFI (70)
- Routes serve control plane purposes, including:
  - MAC address reachability
  - MAC mass withdrawal
  - Split-Horizon label adv.
  - Aliasing
  - Multicast endpoint discovery
  - Redundancy group discovery
  - Designated forwarder election

#### BGP Route Attributes
- New BGP extended communities defined
- Expand information carried in BGP routes, including:
  - MAC address moves
  - C-MAC flush notification
  - Redundancy mode
  - MAC / IP bindings of a GW
  - Split-horizon label encoding

---

**Route Types**
- [1] Ethernet Auto-Discovery (AD) Route
- [2] MAC Advertisement Route
- [3] Inclusive Multicast Route
- [4] Ethernet Segment Route
Ethernet Multipoint L2VPNs
Provider Backbone Bridging (PBB) Ethernet VPN (PBB-EVPN)

Topic covered in detail in BRKMPL-2333 (THU)
PBB Ethernet VPN

Highlights

• Next generation solution for Ethernet multipoint (E-LAN) services by combining Provider Backbone Bridging (PBB - IEEE 802.1ah) and Ethernet VPN

• Data-plane learning of local C-MACs and remote C-MAC to B-MAC binding

• PEs run Multi-Protocol BGP to advertise local Backbone MAC addresses (B-MACs) & learn remote B-MACs
  – Takes advantage of PBB encapsulation to simplify BGP control plane operation – faster convergence
  – Lowers BGP resource usage (CPU, memory) on deployed infrastructure (PEs and RRs)

• Under standardization at IETF – WG draft: draft-ietf-l2vpn-pbb-evpn
Provider Backbone Bridging Overview

• **PBB** (IEEE 802.1ah-2008) defines an architecture that includes
  – \(2^{24}\) service instances (I-SID) per B-VLAN
  – MAC-in-MAC

• **I-Component**
  – Learns & forwards using C-MACs
  – Maintains a mapping table of C-MACs to B-MACs
  – Performs PBB encap/decap on PIP

• **B-Component**
  – Learns & forwards using B-MACs
  – Push / pop B-VLAN on CBP

IB-BEB = I/B-comp Backbone Edge Bridge
I-SID = Backbone Service Instance Identifier
PIP = Provider Instance Port
CBP = Customer Backbone Port

IB-DA = I/B-comp Backbone Edge Bridge
I-TAG = I-TAG
I-SA = I-SA
C-DA = C-DA
C-SA = C-SA
Customer Frame

\[ IB-BEB \]
\[ B-DA / B-SA \]
\[ I-TAG \]
\[ C-DA / C-SA \]
\[ Customer Frame \]

\[ C-DA / C-SA \]
\[ C-DA / C-SA \]
\[ Customer Frame \]
PBB-EVPN Encapsulation

Traffic Direction

Ethernet Access

I-Component

BD
I-SID X

B-Component

BD
EVI aaa

EVPN Forwarder

MPLS

DA (NH router)
SA
E-type (MPLS 0x8847)
PSN MPLS label
EVPN MPLS label
Control Word
PBB Header
Customer Frame

24-bit I-SID inside I-TAG

6B
6B
2B
4B

B-DA
B-SA
E-type (I-TAG 0x88E7)
I-TAG

EVPN MPLS label
Control Word
PBB Header
Customer Frame

Payload E-Type
C-VID
E-type (802.1q 0x8100)
DA
SA

Payload

18B

18B

4B

4B

4B
PBB-EVPN Operation
Multicast Tunnel ID / Endpoint Discovery

1. At start-up, PEs send EVPN Inclusive Multicast routes to signal I-SID membership.

2. I-SID X Flood List
   - Entry 1: PE 2 – label A
   - Entry 2: PE 3 – label B
   - Entry 3: PE 4 – label C

2. I-SID Y Flood List
   - Entry 1: PE 2 – label D
   - Entry 2: PE 4 – label E

Inclusive Multicast route signals MPLS label to be used in the downstream direction.
PBB-EVPN Operation

B-MAC Reachability Advertisement

At start-up, PEs send EVPN MAC Advertisement route for local B-MAC/EVI

1. L2 Routing Information Base (RIB)
   - B-DA2; Next Hop PE2; label F
   - B-DA3; Next Hop PE3; label G
   - B-DA4; Next Hop PE4; label H

2. MAC Advertisement route signals MPLS label to be used in the downstream direction
PBB-EVPN Operation

Multi-Destination Traffic Forwarding (Per-ISID Ingress Replication)

1. Multi-destination Traffic
   • Unknown unicast
   • Broadcast
   • Multicast

   SA: C-MAC1a
   DA: FFFF.FFFF.FFFF

   SA: C-MAC1b
   DA: FFFF.FFFF.FFFF

2. Ingress replication with Per-ISID flooding
   3 copies for I-SID X
   2 copies for I-SID Y

   I-SID X Flood List
   Entry 1: PE 2 – label A
   Entry 2: PE 3 – label B
   Entry 3: PE 4 – label C

   I-SID Y Flood List
   Entry 1: PE 2 – label D
   Entry 2: PE 4 – label E

3. CAM Table I-SID X
   Entry1: C-MAC1a; B-DA1

   CAM Table I-SID Y
   Entry1: C-MAC1b; B-DA1

CE1

SA: C-MAC1a
DA: FFFF.FFFF.FFFF

SA: C-MAC1b
DA: FFFF.FFFF.FFFF

CE2

CE3

CE4

SA: C-MAC1a
DA: FFFF.FFFF.FFFF

SA: C-MAC1b
DA: FFFF.FFFF.FFFF

SA: C-MAC1b
DA: FFFF.FFFF.FFFF

CAM Table I-SID X
Entry1: C-MAC1a; B-DA1

CAM Table I-SID Y
Entry1: C-MAC1b; B-DA1

SA: C-MAC1a
DA: FFFF.FFFF.FFFF

SA: C-MAC1b
DA: FFFF.FFFF.FFFF

SA: C-MAC1b
DA: FFFF.FFFF.FFFF
PBB-EVPN Operation

Known Unicast Traffic Forwarding

1. Known Unicast Traffic

   CAM Table I-SID X
   Entry1: C-MAC1a; local
   Entry2: C-MAC2; B-DA2
   Entry3: C-MAC4; B-DA4

2. Known Unicast delivered to specific remote PEs

   L2 Routing Information Base (RIB)
   B-DA2; Next Hop PE2; label F
   B-DA3; Next Hop PE3; label G
   B-DA4; Next Hop PE4; label H

   L2 Routing Information Base (RIB)
   B-DA2; Next Hop PE2; label F
   B-DA3; Next Hop PE3; label G
   B-DA4; Next Hop PE4; label H

   Known Unicast delivered to specific remote PEs
Introducing PBB-EVPN in Cisco ASR 9000

• Introducing the next-generation of L2VPNs – Provider Backbone Bridging Ethernet VPN (PBB-EVPN)

• Support across Cisco ASR 9000 series router family
  – From ASR9001-S to ASR9922

• Support starting with Cisco IOS-XR release 4.3.2¹ (FCS 09/2013)

• Enhanced Ethernet Line Cards (Typhoon) required as Ingress and Egress linecards

¹ PBB-EVPN support started in IOS-XR 4.3.2 and 5.1.1 releases
Advanced Topics
Resiliency
Two-Way Pseudowire Redundancy and mLACP
Two-Way Pseudowire Redundancy

Overview

• Allows dual-homing of two local PEs to two remote PEs

• Four (4) pseudowires: 1 primary & 3 backup provide redundancy for dual-homed devices

• Two-Way PW redundancy coupled with Multi-Chassis LAG (MC-LAG) solution on the access side
  – LACP state used to determine PW AC state
  – InterChassis Communication Protocol (ICCP) used to synchronize LACP states
E-LINE Availability Model
Active / Backup Access Node Redundancy (mLACP)

- Port / Link Failures

- For VPWS Coupled Mode, attachment circuit (AC) state (Active/Standby) drives PW state advertised to remote peers

Events

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Initial state</td>
</tr>
<tr>
<td>F&lt;sub&gt;A-C&lt;/sub&gt;</td>
<td>Port / Link Failures</td>
</tr>
<tr>
<td>1&lt;sub&gt;A&lt;/sub&gt;</td>
<td>Active PoA detects failure and signals failover over ICCP</td>
</tr>
<tr>
<td>1&lt;sub&gt;B&lt;/sub&gt;</td>
<td>Failover triggered on DHD</td>
</tr>
<tr>
<td>2</td>
<td>Standby link brought up per LACP proc.</td>
</tr>
<tr>
<td>3</td>
<td>Active PoA advertises “Standby” state on its PWs</td>
</tr>
<tr>
<td>4</td>
<td>Standby PoA advertises “Active” state on its PWs</td>
</tr>
</tbody>
</table>

For VPWS Coupled Mode, attachment circuit (AC) state (Active/Standby) drives PW state advertised to remote peers.
E-LINE Availability Model
Active / Backup Access Node Redundancy (mLACP)

- Port / Link Failures (cont.)

- Local site access failure does not trigger LACP failover at remote site (i.e. control-plane separation between sites)

Events

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Initial state</td>
</tr>
<tr>
<td>F_{A-C}</td>
<td>Port / Link Failures</td>
</tr>
<tr>
<td>1_A</td>
<td>Active PoA detects failure and signals failover over ICCP</td>
</tr>
<tr>
<td>1_B</td>
<td>Failover triggered on DHD</td>
</tr>
<tr>
<td>2</td>
<td>Standby link brought up per LACP proc.</td>
</tr>
<tr>
<td>3</td>
<td>Active PoA advertises “Standby” state on its PWs</td>
</tr>
<tr>
<td>4</td>
<td>Standby PoA advertises “Active” state on its PWs</td>
</tr>
<tr>
<td>E</td>
<td>End State</td>
</tr>
</tbody>
</table>

Local site access failure does not trigger LACP failover at remote site (i.e. control-plane separation between sites)
Advanced Topics
Resiliency
ITU-T G.8032 Access Redundancy
ITU-T G.8032 Overview

- **Standards-based protection** switching for Ethernet ring topologies
  - Defined by ITU-T Study Group 15 [G.8032/Y.1344] (v1 – 06/08; v2 – 03/10)
- Ring traffic forwarding based on Ethernet bridging rules – *Layer 2 Rings*
- **Loop avoidance** by blocking of designated ring link under normal conditions
- Uses a **dedicated Control Channel (VLAN)** carrying control messages - *Ring APS*
- Leverages Ethernet CFM / ITU-T Y.1731 for Fault Detection (*CCM*)
- **Single Ring or Multi-Ring network** topologies
- Supports **MAC flushing, load-balancing, revertive / non-revertive switching and administrative switching commands**
E-LINE Availability Model
Ring Access Node Redundancy (G.8032)

- G.8032 Ring Span Failure

- G.8032 Open Ring without R-APS Virtual Channel, terminating on Aggregation Nodes
- VLAN load balancing using two ERP instances with RPL Owners on Aggregation Nodes.

Events

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Initial state</td>
</tr>
<tr>
<td>( F_B )</td>
<td>Ring Span failure</td>
</tr>
<tr>
<td>( 1_{A-B} )</td>
<td>Access switches “A” and “B” detect link failure. Send R-APS Signal Fail (SF) on the ring</td>
</tr>
<tr>
<td>2</td>
<td>Access nodes in the ring flush MAC tables and propagate R-APS SF</td>
</tr>
<tr>
<td>3</td>
<td>RPL owner AGG node receives R-APS and unblocks RPL owner port</td>
</tr>
</tbody>
</table>
E-LINE Availability Model
Ring Access Node Redundancy (G.8032)

- G.8032 Ring Span Failure (cont.)

<table>
<thead>
<tr>
<th>Events</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>RPL owner AGG node receives R-APS SF and unblocks RPL owner port</td>
</tr>
<tr>
<td>4</td>
<td>AGG nodes flush MAC tables. Trigger LDP MAC add withdrawal to VPLS peers</td>
</tr>
<tr>
<td>5</td>
<td>Remote peers flush MAC tables</td>
</tr>
</tbody>
</table>
E-LINE Availability Model
Ring Access Node Redundancy (G.8032)

- G.8032 Ring Span Failure (cont.)

Events

5
Remote peers flush MAC tables

E
End State
Advanced Topics
L2VPN Load Balancing
Load-balancing Questions

• How do we make LERs distribute flows within the same PW across ECMPs?

• How do we make LERs distribute flows within the same PW across members of core-facing bundle interface?

• How do we make LSRs distribute flows within the same PW across ECMPs?

• How do we make LSRs distribute flows within the same PW across members of core-facing bundle interface?
Flow Aware Transport PWs (RFC6391)

• Problem: How can LSRs load-balance traffic from flows in a PW across core ECMPs and Bundle interfaces?

• LSRs load-balance traffic based on IP header information (IP payloads) or based on bottom of stack MPLS label (Non-IP payloads)
  – PW traffic handled as Non-IP payload

• RFC6391 defines a mechanism that introduces a Flow label that allows P routers to distribute flows within a PW
  – PEs push / pop Flow label
  – P routers not involve in any signaling / handling / manipulation of Flow label
L2VPN Load-balancing (E2E Scenario) (1/2)
L2VPN Load-balancing (E2E Scenario) (2/2)

PE router with ECMP and Bundle interfaces

P router with ECMP and Bundle interfaces

P router with ECMP and Non-bundle interfaces

P router without ECMP and Non-bundle interfaces

PE router with Bundle interface as PW attachment circuit (AC)

P router with ECMP and Bundle interfaces
L2VPN Load-balancing (Per-VC LB)

Default - ASR9000 PE with Core-facing Bundle
P rtr load-balances traffic across Bundle members based on VC label; i.e. all traffic from a PW assigned to one member

Default - ASR9000 PE with ECMP
PE load-balances PW traffic across ECMPs based on VC label; i.e. all traffic from a PW assign to one ECMP

Default - ASR9000 P with Core-facing Bundle
PE load-balances traffic across Bundle members based on VC label; i.e. all traffic from a PW assigned to one member

Default - ASR9000 P with ECMP
P rtr load-balances traffic across ECMPs based on VC label; i.e. all traffic from a PW assigned to one ECMP

Svc X – Flow 1
Svc X – Flow 2
Svc X – Flow 3
Svc X – Flow 4
Svc Y – Flow 1
Svc Y – Flow 2
Svc Y – Flow 3
Svc Y – Flow 4
L2VPN Load-balancing (L2/L3 LB)

**PE L2VPN load-balancing knob:**

```
12vpn
load-balancing flow {src-dst-mac | src-dst-ip}
```

**Default - ASR9000 P**

- PW loadbalancing based on VC label; only one ECMP and one bundle member used for all PW traffic

**Two-stage Hash process**

**ASR9000 PE with ECMP**

- PE now load-balances PW traffic across ECMPs based on L2 or L3 payload info; i.e. flows from a PW distributed over ECMPs

**ASR9000 PE with Core-facing Bundle**

- PE now load-balances traffic across Bundle members based on L2 or L3 payload info; i.e. flows from a PW distributed over members

**ASR9000 PE with AC Bundle**

- PE load-balances now traffic across Bundle members based on L2 or L3 payload info
L2VPN Load-balancing (L2/L3 LB + FAT)

PE L2VPN load-balancing knob

PE FAT PW
l2vpn
pw-class sample-class
encapsulation mpls
load-balancing flow-label both

ASR9000 PE
PE now adds Flow labels based on L2 or L3 payload info

ASR9000 PE with ECMP
PE now load-balances PW traffic across ECMPs based on L2 or L3 payload info; i.e. flows from a PW distributed over ECMPs

ASR9000 PE with Core-facing Bundle
PE now load-balancing based on Flow label; i.e. flows from a PW distributed over bundle members

ASR9000 P with Core-facing Bundle
PW loadbalancing based on Flow label; i.e. flows from a PW distributed over bundle members

ASR9000 P with ECMP
P rtr now load-balances traffic across ECMPs based on Flow label; i.e. flows from a PW distributed over ECMPs

ASR9000 P with AC Bundle
PE load-balances now traffic across Bundle members based on L2 or L3 payload info

No new configuration required on P routers

Svc X – Flow 1
Svc X – Flow 2
Svc X – Flow 3
Svc X – Flow 4
Significance of PW Control-Word

Problem:
DANGER for LSR
LSR will confuse payload as IPv4 (or IPv6) and attempt to load-balance based off incorrect fields

Solution:
Add PW Control Word in front of PW payload. This guarantees that a zero will always be present and thus no risk of confusion for LSR

Table:
- RTR DA
- RTR SA
- MPLS E-Type (0x8847)
- PSN MPLS Label
- PW MPLS Label
- PW CW
- DA
- SA
- 802.1q Tag (0x8100)
- C-VID
- Payload E-Type
- Non-IP Payload
Deployment Use Cases

Data Center Interconnect – VPLS on Nexus 7000
Data Center Interconnect with VPLS
Nexus 7000

• Nexus 7000 as DC WAN Edge provides VPLS Multi-Homing with Virtual Port Channel (vPC)

• User configuration sets VFI as primary / secondary on vPC members
  – vPC members can alternate in Active / Standby responsibilities for different VLANs

• PW status signaled as Active / Standby on primary / secondary VFIs respectively
  – Single PW activated to forward traffic between pair of data center sites
  – vPC Peer Link used to forward traffic to / from vPC member with VFI in primary designation
Data Center Interconnect with VPLS

Sample Configuration – Nexus 7000

PE 1

- Primary VFI owner for EVEN vlans
- Secondary owner for ODD vlans

PE 2

- Primary VFI owner for ODD vlans
- Secondary owner for EVEN vlans

Note: Virtual Port Channel (vPC) configuration not shown
Deployment Use Cases

E-LAN with per-flow load-balancing – ASR 9000 (PBB-EVPN)
PBB-EVPN

Multi-Homing Scenarios – All-Active Load-Balancing

• Dual Home Device / Multi Home Device\(^1\) scenarios and All-Active LB
  – A.k.a. Active / Active per-flow (AApF) LB
  – Both PEs forward traffic associated with a given PBB I-SID

• PEs attached to Ethernet Segment using bundle interfaces
  – Single bundle (manual or LACP) configured on CE

• PEs on same segment must share the same source B-MAC and ESI
  – ESI and B-MAC auto-sensed from CE LACP information

• DF election (manual or automatic)

(1) Standard does not limit solution to only dual homing
PBB-EVPN Dual Home Device (DHD)

All-Active (per-FLOW) Load-Balancing

**PE1**
- redundancy iccp group 66
  - mlacp node 1
  - mlacp system priority 1
  - mlacp system mac 0111.0222.0111
  - **mode singleton**
    - backbone interface GigabitEthernet 0/0/0/1

- interface Bundle-Ether25
  - mlacp iccp-group 66

- interface Bundle-Ether25.1 l2transport
  - encapsulation dot1q 777

- l2vpn
  - bridge group gr1
  - bridge-domain bd1
    - interface Bundle-Ether25.1
      - pbb edge i-sid 256 core-bridge core_bd1
  - bridge group gr2
  - bridge-domain core_bd1
    - pbb core
      - evpn evi 1000

- router bgp 64
  - bgp router-id 1.100.100.100
  - **address-family l2vpn evpn**
  - neighbor 2.100.100.100
  - remote-as 64
    - **address-family l2vpn evpn**

**Note:** MPLS / LDP configuration required on core-facing interfaces (not shown)

**PE2**
- should use same RG #
- should use different mlacp node id
- should use same mlacp system mac and system priority

- ICCP in singleton mode (i.e. No peer neighbor configuration)

- PBB I-component and B-component configuration. ISIDs must match on both PEs
- No need to define B-VLAN
  - **Mandatory** EVI ID configuration

- BGP configuration with new EVPN AF

**Auto-sensed B-MAC SA**
**Auto-sensed ESI**
**Auto RD for Segment Route**
**Auto RT for EVI**
**Auto RD for EVI**
**A/A Per-flow LB (default)**
**Auto DF / service carving**
Summary

• MPLS is a mature technology with widespread L2VPN deployments by Service Providers and Enterprises around the globe
  – Ethernet-based WAN services and Data Center Interconnect are key applications driving deployments of L2VPN today

• L2VPNs can be deployed addressing key requirements including: Resiliency, Auto-Discovery, Load-Balancing and OAM

• EVPN / PBB-EVPN are next-generation L2VPN solutions based on BGP control-plane for MAC distribution/learning over the core
MPLS Sessions at Cisco Live

- BRKMPL-1100  Introduction to MPLS
- BRKMPL-2100  Deploying MPLS Traffic Engineering
- BRKMPL-2101  Deploying MPLS-based Layer 2 Virtual Private Networks
- BRKMPL-2102  Deploying MPLS-based IP VPNs
- BRKMPL-2108  Designing MPLS in Next Generation Data Center: A Case Study
- BRKMPL-2333  E-VPN & PBB-EVPN: the Next Generation of MPLS-based L2VPN
- BRKMPL-3101  Advanced Topics and Future Directions in MPLS
- LTRMPL-2102  Enterprise Network Virtualization using IP and MPLS Technologies: Introduction
- LTRMPL-3102  Enterprise Network Virtualization using IP and MPLS Technologies: Advanced
- TECMPL-3100  Unified MPLS - An architecture for Advanced IP NGN Scale
- TECMPL-3200  SDN WAN Orchestration in MPLS and Segment Routing Networks
Participate in the “My Favorite Speaker” Contest

Promote Your Favorite Speaker and You Could be a Winner

• Promote your favorite speaker through Twitter and you could win $200 of Cisco Press products (@CiscoPress)

• Send a tweet and include
  – Your favorite speaker’s Twitter handle @jliste2002
  – Two hashtags: #CLUS #MyFavoriteSpeaker

• You can submit an entry for more than one of your “favorite” speakers

• Don’t forget to follow @CiscoLive and @CiscoPress

• View the official rules at http://bit.ly/CLUSwin
Complete Your Online Session Evaluation

• Give us your feedback and you could win fabulous prizes. Winners announced daily.

• Complete your session evaluation through the Cisco Live mobile app or visit one of the interactive kiosks located throughout the convention center.

Don’t forget: Cisco Live sessions will be available for viewing on-demand after the event at CiscoLive.com/Online
Continue Your Education

- Demos in the Cisco Campus
- Walk-in Self-Paced Labs
- Table Topics
- Meet the Engineer 1:1 meetings
Thank you.
Ethernet Point-to-Point L2VPNs
Virtual Private Wire Service (VPWS)
hostname PE1
!
interface Loopback0
ipv4 address 106.106.106.106 255.255.255.255

l2vpn
xconnect group Cisco-Live
  p2p xc-sample-1
    interface GigabitEthernet0/0/0/2.100
    neighbor 102.102.102.102 pw-id 111

  p2p xc-sample-2
    interface GigabitEthernet0/0/0/2.200
    neighbor 102.102.102.102 pw-id 222

  p2p xc-sample-3
    interface GigabitEthernet0/0/0/6
    neighbor 102.102.102.102 pw-id 333

interface GigabitEthernet0/0/0/2.100 l2transport
encapsulation dot1q 100
rewrite ingress tag pop 1 symmetric

interface GigabitEthernet0/0/0/2.200 l2transport
encapsulation dot1q 999-1010
rewrite ingress tag push dot1q 888 symmetric
hostname PE1
!
interface Loopback0
  ip address 106.106.106.106 255.255.255.255

interface GigabitEthernet2/4.300
  encapsulation dot1q 300
  xconnect 102.102.102.102 111 encapsulation mpls

interface GigabitEthernet2/4
  service instance 10 ethernet
  encapsulation dot1q 300
  rewrite ingress tag pop 1 symmetric
  xconnect 102.102.102.102 111 encapsulation mpls

interface Vlan 300
  xconnect 102.102.102.102 111 encapsulation mpls
  !
interface GigabitEthernet2/4
  switchport mode trunk
  switchport trunk allowed vlan 300

interface Vlan 300
  xconnect 102.102.102.102 111 encapsulation mpls
  !
interface GigabitEthernet2/4
  service instance 10 ethernet
  encapsulation dot1q 300
  rewrite ingress tag pop 1 symmetric
  bridge-domain 300
VPWS (EoMPLS) LDP Signaling
Cisco IOS (Port-based services)

```
hostname PE1
!
interface Loopback0
 ip address 106.106.106.106 255.255.255.255

interface GigabitEthernet2/5
 xconnect 102.102.102.102 222 encapsulation mpls

interface GigabitEthernet2/5
 service instance 1 ethernet
 encapsulation default
 xconnect 102.102.102.102 111 encapsulation mpls

interface Vlan 300
 xconnect 102.102.102.102 111 encapsulation mpls
!
interface GigabitEthernet2/5
 switchport mode dot1q-tunnel
 switchport access vlan 300

interface Vlan 300
 xconnect 102.102.102.102 111 encapsulation mpls
!
interface GigabitEthernet2/5
 service instance 1 ethernet
 encapsulation default
 bridge-domain 300
```

Main interface based xconnect

Service-Instance (EFP) based xconnect (encap default)

Interface VLAN (SVI) based xconnect + Switchport dot1q-tunnel

Interface VLAN (SVI) based xconnect + Service instance BD
VPWS (EoMPLS) LDP Signaling

Cisco IOS / NX-OS (NEW Service-based CLI)

hostname PE1
! interface Loopback0
  ip address 106.106.106.106 255.255.255.255

l2vpn xconnect context sample-xconnect
  member Pseudowire1 102.102.102.102 111 encap mpls
  member GigabitEthernet2/4 service instance 333
!
interface GigabitEthernet2/4
  service instance 333 ethernet
  encapsulation dot1q 300
  rewrite ingress tag pop 1 symmetric

bridge-domain 300
  member Pseudowire2 192.0.0.5 222 encap mpls
  member GigabitEthernet2/4 service instance 333
!
interface GigabitEthernet2/4
  service instance 333 ethernet
  encapsulation dot1q 300
  rewrite ingress tag pop 1 symmetric

vlan 400
  vlan configuration 400
    member Pseudowire2 102.102.102.102 222 encapsulation mpls
!
interface GigabitEthernet2/5
  switchport mode trunk
  switchport trunk allowed vlan 400

NEW
PWs modeled as virtual interfaces. PW and EFPs now members of BD/Xconn context

NEW
Service-based CLI Xconn context / Bridge-Domain or VLAN configurations

For NX-OS
Ethernet Multi-Point L2VPNs
VPLS with LDP Signaling
hostname PE1
!
interface Loopback0
ip address 192.0.0.1 255.255.255.255
!
l2 vfi sample-vfi manual
vpn id 300
neighbor 192.0.0.2 encapsulation mpls
neighbor 192.0.0.3 2222 encapsulation mpls
neighbor 192.0.0.4 3333 encapsulation mpls
!
interface Vlan300
xconnect vfi sample-vfi

interface GigabitEthernet2/4
service instance 333 ethernet
encapsulation dot1q 333
rewrite ingress tag pop 1 symmetric
bridge-domain 300

interface GigabitEthernet2/4
switchport mode trunk
switchport trunk allowed vlan 300
hostname PE1
!
interface Loopback0
  ip address 192.0.0.1 255.255.255.255
!
l2 vfi sample-vfi manual
  vpn id 300
  neighbor 192.0.0.2 encapsulation mpls
  neighbor 192.0.0.3 2222 encapsulation mpls
  neighbor 192.0.0.4 3333 encapsulation mpls
  neighbor 192.0.0.5 5555 encapsulation mpls no-split-horizon
  neighbor 192.0.0.6 5555 encapsulation mpls no-split-horizon
!
interface Vlan300
  xconnect vfi sample-vfi

interface GigabitEthernet2/4
  service instance 333 ethernet encapsulation dot1q 333
  rewrite ingress tag pop 1 symmetric bridge-domain 300

interface GigabitEthernet2/4
  switchport mode trunk
  switchport trunk allowed vlan 300

Bridge-Domain or VLAN/switchport configurations

Spoke PWs
VPLS LDP Signaling / Manual provisioning

Cisco IOS XR

hostname PE1
!
interface Loopback0
  ipv4 address 192.0.0.1 255.255.255.255
!
interface GigabitEthernet0/0/0/14.101 l2transport
  encapsulation dot1q 101
  rewrite ingress tag pop 1 symmetric

l2vpn
  bridge group Cisco-Live
  bridge-domain bd101
  interface GigabitEthernet0/0/0/14.101
  vfi vfi101
    vpn-id 1111
    neighbor 192.0.0.2 pw-id 1111
    neighbor 192.0.0.3 pw-id 2222
    neighbor 192.0.0.4 pw-id 3333

Service-based CLI:
EFPs, PWs and VFI as members of Bridge Domain

VPN ID defined per VFI or on a per-neighbor basis
l2vpn
bridge group Cisco-Live
bridge-domain bd101
interface GigabitEthernet0/0/0/14.101
neighbor 192.0.0.5 pw-id 5555
neighbor 192.0.0.6 pw-id 5555
vfi vfi101
vpn-id 1111
neighbor 192.0.0.2 pw-id 1111
neighbor 192.0.0.3 pw-id 2222
neighbor 192.0.0.4 pw-id 3333

hostname PE1
interface Loopback0
ipv4 address 192.0.0.1 255.255.255.255
interface GigabitEthernet0/0/0/14.101 12transport
encapsulation dot1q 101
rewrite ingress tag pop 1 symmetric

PE1 192.0.0.1
PE2 192.0.0.2
PE3 192.0.0.3
PE4 192.0.0.4

u-PE1 192.0.0.5
u-PE2 192.0.0.6

Core PWs
Full-mesh

Spoke PWs
VPLS LDP Signaling / Manual provisioning
Cisco IOS / NX-OS (NEW Service-based CLI)

hostname PE1
!
interface Loopback0
ip address 192.0.0.1 255.255.255.255

l2vpn vfi context sample-vfi
vpn id 1111
member Pseudowire1 192.0.0.2 encapsulation mpls
member Pseudowire2 192.0.0.3 2222 encapsulation mpls
member Pseudowire3 192.0.0.4 3333 encapsulation mpls
!

bridge-domain 300
member vfi sample-vfi
member GigabitEthernet2/4 service instance 333
!
interface GigabitEthernet2/4
service instance 333 ethernet
encapsulation dot1q 300
rewrite ingress tag pop 1 symmetric

NEW PWs modeled as virtual interfaces. VFI and EFPs now members of BD

NEW Service-based CLI Bridge-Domain or VLAN/switchport configurations

OR

VLAN

vlan 300
vlan configuration 300
member vfi sample-vfi
!
interface GigabitEthernet2/4
switchport mode trunk
switchport trunk allowed vlan 300

Core PWs Full-mesh

CE1

MPLS Core

PW VC id

GigabitEthernet2/4

1111

PE1 192.0.0.1

PE2 192.0.0.2

PE3 192.0.0.3

PE4 192.0.0.4

PW VC id

2222

3333
**H-VPLS LDP Signaling / Manual provisioning**

**Cisco IOS (NEW Service-based CLI)**

```plaintext
hostname PE1
!
interface Loopback0
  ip address 192.0.0.1 255.255.255.255
!
l2vpn vfi context sample-vfi
  vpn id 1111
  member Pseudowire1 192.0.0.2 encapsulation mpls
  member Pseudowire2 192.0.0.3 2222 encapsulation mpls
  member Pseudowire3 192.0.0.4 3333 encapsulation mpls
!
bridge-domain 300
  member vfi sample-vfi
    member Pseudowire4 192.0.0.5 5555 encapsulation mpls
    member Pseudowire5 192.0.0.6 5555 encapsulation mpls
    member GigabitEthernet2/4 service instance 333
!
interface GigabitEthernet2/4
  service instance 333 ethernet
  encapsulation dot1q 300
  rewrite ingress tag pop 1 symmetric

NEW
PWs modeled as virtual interfaces.
VFI, spoke PW, EFP members of BD

NEW
Service-based CLI
Bridge-Domain configurations

Spoke PWs
```

The diagram illustrates the network architecture with highlights on the **NEW Service-based CLI Bridge-Domain configurations** and **NEW PWs modeled as virtual interfaces**.
Ethernet Multi-Point L2VPNs
VPLS with LDP Signaling and BGP-based AutoDiscovery (BGP-AD)
hostname PE1
!
interface Loopback0
ip address 102.102.102.102 255.255.255.255
!
router bgp 100
bgp router-id 102.102.102.102
neighbor 104.104.104.104 remote-as 100
neighbor 104.104.104.104 update-source Loopback0
!
address-family l2vpn vpls
neighbor 104.104.104.104 activate
neighbor 104.104.104.104 send-community extended
exit-address-family

l2 vfi sample-vfi autodiscovery
vpn id 300
vpls-id 100:300
!
interface Vlan300
xconnect vfi sample-vfi

interface GigabitEthernet2/4
service instance 333 ethernet
encapsulation dot1q 333
rewrite ingress tag pop 1 symmetric
bridge-domain 300

interface GigabitEthernet2/4
switchport mode trunk
switchport trunk allowed vlan 300
hostname PE1
!
interface Loopback0
  ip address 102.102.102.102 255.255.255.255
!
l2 vfi sample-vfi autodiscovery
  vpn id 300
  vpls-id 100:300
  neighbor 192.0.0.5 5555 encapsulation mpls no-split-horizon
  neighbor 192.0.0.6 5555 encapsulation mpls no-split-horizon

Manually provisioned Spoke PWs
VPLS LDP Signaling and BGP-AD
Cisco IOS XR

hostname PE1
!
interface Loopback0
  ipv4 address 106.106.106.106 255.255.255.255
!
interface GigabitEthernet0/0/0/2.101 l2transport
capsulation dot1q 101
  rewrite ingress tag pop 1 symmetric

router bgp 100
  bgp router-id 106.106.106.106
  address-family l2vpn vpls-vpws
  neighbor 110.110.110.110
    remote-as 100
    update-source Loopback0
    address-family l2vpn vpls-vpws

l2vpn
  bridge group Cisco-Live
    bridge-domain bd101
      interface GigabitEthernet0/0/0/2.101
        vfi vfi101
          vpn-id 11101
          autodiscovery bgp
          rd auto
          route-target 100:101
          signaling-protocol ldp
          vpls-id 100:101

BGP AS 100
  BGP Auto-Discovery

PE1 106.106.106.106
  PW VC id 100:101

PE2 110.110.110.110
  PW VC id 100:101

PE3 192.0.0.3

PE4 192.0.0.4

CE1

Full-mesh Core PWs
  auto-discovered with BGP-AD
  and signaled by LDP

PW ID = VPLS-id (100:101)
H-VPLS LDP Signaling and BGP-AD / Manual provisioning

Cisco IOS XR
hostname PE1
!
interface Loopback0
 ip address 102.102.102.102 255.255.255.255

12vpn vfi context sample-vfi
 vpn id 3300
   autodiscovery bgp signaling ldp
   vpls-id 100:3300
!
router bgp 100
 neighbor 104.104.104.104 remote-as 100
  update-source loopback 0
  address-family l2vpn vpls
     send-community extended

system bridge-domain 300
!
bridge-domain 300
   member vfi sample-vfi
   member Ethernet2/4 service instance 333
!
interface Ethernet2/4
   service instance 333 ethernet
   encapsulation dot1q 300

BGP AS 100
BGP Auto-Discovery

vlan 300
vlan configuration 300
   member vfi sample-vfi
!
interface Ethernet2/4
   switchport
   switchport mode trunk
   switchport trunk allowed vlan 300
VPLS LDP Signaling and BGP-AD

Cisco IOS (NEW Service-based CLI)

```plaintext
hostname PE1
!
interface Loopback0
  ip address 102.102.102.102 255.255.255.255
!
routing bgp 100
  bgp router-id 102.102.102.102
  neighbor 104.104.104.104 remote-as 100
  neighbor 104.104.104.104 update-source Loopback0
!
address-family l2vpn vpls
  neighbor 104.104.104.104 activate
  neighbor 104.104.104.104 send-community extended
  exit-address-family

l2vpn vfi context sample-vfi
  vpn id 300
  autodiscovery bgp signaling ldp
    vpls-id 100:300
!
bridge-domain 300
  member vfi sample-vfi
  member GigabitEthernet2/4 service instance 333

interface GigabitEthernet2/4
  service instance 333 ethernet
  encapsulation dot1q 333
  rewrite ingress tag pop 1 symmetric
```

Bridge Domain-based Configuration

BGP AS 100
BGP Auto-Discovery
H-VPLS LDP Signaling and BGP-AD / Manual provisioning

Cisco IOS (NEW Service-based CLI)

```
hostname PE1
!
l2vpn vfi context sample-vfi
  vpn id 3300
  autodiscovery bgp signaling ldp
  vpls-id 100:3300

bridge-domain 300
member vfi sample-vfi
member Pseudowire4 192.0.0.5 5555 encapsulation mpls
member Pseudowire5 192.0.0.6 5555 encapsulation mpls
member GigabitEthernet2/4 service instance 333
```

Bridge Domain-based Configuration

Manually provisioned Spoke PWs
Ethernet Multi-Point L2VPNs
VPLS with BGP-based Signaling and AutoDiscovery
hostname PE1
!
interface Loopback0
  ipv4 address 106.106.106.106 255.255.255.255
!
router bgp 100
  bgp router-id 106.106.106.106
  address-family l2vpn vpls-vpws
  neighbor 110.110.110.110
    remote-as 100
    update-source Loopback0
  address-family l2vpn vpls-vpws

l2vpn
  bridge group Cisco-Live
  bridge-domain bd102
    interface GigabitEthernet0/0/0/2.102
    vfi vfi102
      vpn-id 11102
      autodiscovery bgp
        rd auto
        route-target 100:102
        signaling-protocol bgp
        ve-id 5

VE-id must be unique in a VPLS instance
VPLS BGP Signaling and BGP-AD

Cisco IOS (NEW Service-based CLI)

hostname PE1
!
interface Loopback0
  ip address 102.102.102.102 255.255.255.255
!
routing bgp 100
  bgp router-id 102.102.102.102
  neighbor 104.104.104.104 remote-as 100
  neighbor 104.104.104.104 update-source Loopback0
!
  address-family l2vpn vpls
    neighbor 104.104.104.104 activate
    neighbor 104.104.104.104 send-community extended
    neighbor 104.104.104.104 suppress-signaling-protocol ldp
  exit-address-family

l2vpn vfi context sample-vfi
  vpn id 3300
  autodiscovery bgp signaling bgp
    ve id 5
    ve range 10

bridge-domain 300
  member vfi sample-vfi
    member GigabitEthernet2/4 service instance 333
!
interface GigabitEthernet2/4
  service instance 333 ethernet
  encapsulation dot1q 300
  rewrite ingress tag pop 1 symmetric
hostname PE1
!
interface Loopback0
  ip address 106.106.106.106 255.255.255.255
!
router bgp 100
  neighbor 110.110.110.110 remote-as 100
  update-source Loopback 0
  address-family l2vpn vpls
  suppress-signaling-protocol ldp
  send-community extended

l2vpn vfi context sample-vfi
  vpn id 3300
  autodiscovery bgp signaling bgp
  ve id 5
  ve range 10

system bridge-domain 300
!
bridge-domain 300
  member vfi sample-vfi
  member Ethernet2/4 service instance 333
!
interface Ethernet2/4
  service instance 333 ethernet
  encapsulation dot1q 300

VPLS BGP Signaling and BGP-AD
Cisco NX-OS

VE-id must be unique in a VPLS instance

Bridge Domain-based Configuration

OR

VLAN/switchport-based Configuration

BGP AS 100
BGP Signaling and Auto-Discovery

system bridge-domain 300
!
bridge-domain 300
  member vfi sample-vfi
  member Ethernet2/4 service instance 333
!
interface Ethernet2/4
  service instance 333 ethernet
  encapsulation dot1q 300

VLAN/switchport-based Configuration
PBB-EVPN IOS-XR Implementation
Configuration and Examples
PBB-EVPN Single Home Device (SHD)

**PE1**

interface Bundle-Ether1.777 l2transport
   encapsulation dot1q 777

l2vpn
   bridge group gr1
      bridge-domain bd1
         interface Bundle-Ether1.777
            pbb edge i-sid 256 core-bridge core_bd1

   bridge group gr2
      bridge-domain core_bd1
         pbb core
t         evpn evi 1000

router bgp 64
   bgp router-id 1.100.100.100
   address-family l2vpn evpn
   !
   neighbor 2.100.100.100
      remote-as 64
      update-source Loopback0
      address-family l2vpn evpn

Chassis B-MAC SA
Null ESI
Auto RD for Segment Route
Auto RT for EVI
Auto RD for EVI

PBB I-component
Includes I-SID assignment

PBB B-component
No need to define B-VLAN
**Mandatory** - Globally unique identifier for all PEs in a given EVI

BGP configuration with new EVPN AF

**Note:** MPLS / LDP configuration required on core-facing interfaces (not shown)
PBB-EVPN Single Home Device (SHD) with PW access

PE1

l2vpn
  bridge_group gr1
  bridge-domain bd1
  !
  pbb edge i-sid 256 core-bridge core_bd1

bridge group gr2
  bridge-domain core_bd1
    pbb core
      evpn evi 1000

router bgp 64
  bgp router-id 1.100.100.100
  address-family l2vpn evpn
    !
  neighbor 2.100.100.100
    remote-as 64
    update-source Loopback0
    address-family l2vpn evpn

PBB I-component includes:
- Access PW
- I-SID assignment

PBB B-component
No need to define B-VLAN
Mandatory - Globally unique identifier for all PEs in a given EVI

BGP configuration with new EVPN AF

Note: MPLS / LDP configuration required on core-facing interfaces (not shown)
PBB-EVPN Dual Home Device (DHD)

All-Active (per-FLOW) Load-Balancing

PE1
- redundancy iccp group 66
- mlacp node 1
- mlacp system priority 1
- mlacp system mac 0111.0222.0111
  - mode singleton
  - backbone interface GigabitEthernet 0/0/0/1

interface Bundle-Ether25
- mlacp iccp-group 66

interface Bundle-Ether25.1 l2transport
- encapsulation dot1q 777

l2vpn
- bridge group gr1
- bridge-domain bd1
  - interface Bundle-Ether25.1
  - pbb edge i-sid 256 core-bridge core_bd1

bridge group gr2
- bridge-domain core_bd1
- pbb core
  - evpn evi 1000

router bgp 64
- bgp router-id 1.100.100.100
- address-family l2vpn evpn

neighbor 2.100.100.100
- remote-as 64
- address-family l2vpn evpn

PE2 should use same RG #
PE2 should use different mlacp node id
PE2 should use same mlacp system mac and system priority

ICCP in singleton mode (i.e. No peer neighbor configuration)

PBB I-component and B-component configuration. ISIDs must match on both PEs
No need to define B-VLAN

Mandatory EVI ID configuration

BGP configuration with new EVPN AF

Auto-sensed B-MAC SA
Auto-sensed ESI
Auto RD for Segment Route
Auto RT for EVI
Auto RD for EVI
A/A Per-flow LB (default)
Auto DF / service carving

Note: MPLS / LDP configuration required on core-facing interfaces (not shown)
PBB-EVPN Dual Home Device (DHD)

Single-Active (per-Service) Load-Balancing and Dynamic Service Carving

**PE1**

interface Bundle-Ether25.1 12transport
encapsulation dot1q 777

**evpn**

interface Bundle-Ether25

  ethernet-segment
  identifier system-priority 1 system-id 0300.0b25.00ce
  load-balancing-mode per-service

l2vpn

  bridge group gr1
  bridge-domain bd1
  interface Bundle-Ether25.1
  pbb edge i-sid 256 core-bridge core_bd1

  bridge group gr2
  bridge-domain core_bd1
  pbb core
  evpn evi 1000

router bgp 64

  bgp router-id 1.100.100.100
  address-family l2vpn evpn
  neighbor 2.100.100.100
  remote-as 64
  address-family l2vpn evpn

---

**Chassis B-MAC SA** (def.)
- Manual ESI
- Auto RD for Segment Route
- Auto RT for EVI
- Auto RD for EVI
- A/A Per-Service LB
- Auto Service Carving (def.)

**A/A per-service (per-ISID)** load balancing with dynamic Service Carving
- ESI must match on both PEs

**PBB I-component and B-component configuration.**
- ISIDs must match on both PEs
- No need to define B-VLAN
  - Mandatory EVI ID configuration

**BGP configuration with new EVPN AF**

**Note:** MPLS / LDP configuration required on core-facing interfaces (not shown). ICCP (singleton) config (not shown)
Data-Plane considerations for Ethernet transport
How Are Ethernet Frames Transported?

- Ethernet frames transported without Preamble, Start Frame Delimiter (SFD) and FCS

- Two (2) modes of operation supported:
  - Ethernet VLAN mode (VC type 0x0004) – created for VLAN over MPLS application
  - Ethernet Port / Raw mode (VC type 0x0005) – created for Ethernet port tunneling application
Ethernet PW VC Type

- VC type used must match on PEs
- Cisco IOS devices by default will generally attempt to bring up an Ethernet PW using VC type 5
  - If rejected by remote PE, then VC type 4 will be used – VC Type auto-sensing
- Alternatively, Cisco IOS and IOS-XR devices can be explicitly configured to use either VC type 4 or 5

```
7604-2#show running-config
pseudowire-class test-pw-class-VC4
  encapsulation mpls
  interworking vlan
!
pseudowire-class test-pw-class-VC5
  encapsulation mpls
  interworking ethernet
```
Introducing Cisco EVC Framework

Functional Highlights

Flexible service delimiters
- Single-tagged, Double-tagged
- VLAN Lists, VLAN Ranges
- Header fields (COS, Ethertype)

ANY service – ANY port
- Layer 2 Point-to-Point
- Layer 2 Multipoint
- Layer 3

Ethernet Service Layer
- Ethernet Flow Point (EFP)
- Ethernet Virtual Circuit (EVC)
- Bridge Domain (BD)
- Local VLAN significance

VLAN Header operations - VLAN Rewrites
- POP
- PUSH
- SWAP
Encapsulation Adjustment Considerations

EoMPLS PW VC Type and EVC VLAN Rewrites

• VLAN tags can be added, removed or translated prior to VC label imposition or after disposition
  – Any VLAN tag(s), if retained, will appear as payload to the VC

• VC label imposition and service delimiting tag are independent from EVC VLAN tag operations
  – Dummy VLAN tag – RFC 4448 (sec 4.4.1)

• VC service-delimiting VLAN-ID is removed before passing packet to Attachment Circuit processing
Encapsulation Adjustment Considerations
VC 5 and EVC Rewrites

Single-tagged frame
Double-tagged frame

10
10 tag

• POP VLAN 10
• No Push of Dummy tag (VC 5)

PE1
104.104.104.104

MPLS
Pseudowire
VC Type 5

PE2
102.102.102.102

CE-1
104.104.104.104

CE-2
102.102.102.102

10
10 tag

No service-delimiting vlan expected (VC 5)
PUSH VLAN 10

ios-xr

12vpn
pw-class class-VC5
capsulation mpls
transport-mode ethernet

xconnect group Cisco-Live
p2p xc-sample-1
interface GigabitEthernet0/0/0/2.100
neighbor 102.102.102.102 pw-id 111
pw-class class-VC5

interface GigabitEthernet0/0/0/2.100 12transport
capsulation dot1q 10
rewrite ingress tag pop 1 symmetric

 IOS

pseudowire-class class-VC5
capsulation mpls
interworking ethernet

interface GigabitEthernet2/2
service instance 3 ethernet
capsulation dot1q 10
rewrite ingress tag pop 1 symmetric
xconnect 104.104.104.104 111 encap mpls pw-class class-VC5
Encapsulation Adjustment Considerations

VC 4 and EVC Rewrites

Single-tagged frame
Double-tagged frame

IOS-XR

12vpn
pw-class class-VC4
encapsulation mpls
transport-mode vlan

xconnect group Cisco-Live
p2p xc-sample-1
interface GigabitEthernet0/0/0/2.100
neighbor 102.102.102.102 pw-id 111
pw-class class-VC4

interface GigabitEthernet0/0/0/2.100 l2transport
encapsulation dot1q 10
rewrite ingress tag pop 1 symmetric

IOS

pseudowire-class class-VC4
encapsulation mpls
interworking vlan

interface GigabitEthernet2/2
service instance 3 ethernet
encapsulation dot1q 10
rewrite ingress tag pop 1 symmetric
xconnect 104.104.104.104 111 encap mpls pw-class class-VC4
MTU Considerations

• No payload fragmentation supported

• Incoming PDU dropped if MTU exceeds AC MTU

• PEs exchange PW payload MTU as part of PW signaling procedures
  – Both ends must agree to use same value for PW to come UP
  – PW MTU derived from AC MTU

• No mechanism to check Backbone MTU
  – MTU in the backbone must be large enough to carry PW payload and MPLS stack
Ethernet MTU Considerations

Cisco IOS

- Interface MTU configured as largest ethernet payload size
  - 1500B default
  - Sub-interfaces / Service Instances (EFPs) MTU always inherited from main interface

- PW MTU used during PW signaling
  - By default, inherited from attachment circuit MTU
  - Submode configuration CLI allows MTU values to be set per subinterface/EFP in xconnect configuration mode (only for signaling purposes)
  - No MTU adjustments made for EFP rewrite (POP/PUSH) operations
Ethernet MTU Considerations

Cisco IOS XR

• Interface / sub-interface MTU configured as largest frame size – FCS (4B)
  – 1514B default for main interfaces
  – 1518B default for single-tagged subinterfaces
  – 1522B default for double-tagged subinterfaces

• PW MTU used during PW signaling
  – AC MTU – 14B + Rewrite offset
  – E.g. POP 1 (- 4B), PUSH 1 (+ 4B)

By default, sub-interface MTU inherited from Main interface
Sub-interface MTU can be overwritten to match remote AC

XC MTU = 1518 – 14 – 4 = 1500B

interface GigabitEthernet0/0/0/2
description Main interface
mtu 9000

interface GigabitEthernet0/0/0/2.100 12transport encapsulation dot1q 100 rewrite ingress tag pop 1 symmetric mtu 1518

RP/0/RSP0/CPU0:PE1#show l2vpn xconnect neighbor 102.102.102.102 pw- id 11
Group Cisco-Live, XC xc-sample-1, state is down; Interworking none
AC: GigabitEthernet0/0/0/2.100, state is up
  Type VLAN; Num Ranges: 1
  VLAN ranges: [100, 100]
  MTU 1500; XC ID 0x840014; interworking none
  Statistics:
  (snip)
Advanced Topics
Multi-Segment Pseudowire
Multi-Segment Pseudowire

Overview

• Separate IGP processes (or areas) for separate MPLS Access networks

• T-PE – Terminating Provider Edge
  – Customer facing PE, hosting the first or last segment of a MS-PW

• S-PE – Switching Provider Edge
  – Switches control / data planes of preceding and succeeding segments
  – Control Word, sequencing, or original packet header not examined
  – VC labels swapped
  – VC Type, MTU should match end-to-end
  – One or more S-PEs can be used depending on number of segments

• MS-PW uses same signaling procedures and TLVs described in RFC 4447
Multi-Segment Pseudowires

- Tunnel Label
  - VC Label
    - Push
    - Pop
    - Swap
    - Traffic direction

- PW switching point
  - VC labels swapped, new Tunnel label pushed

- Penultimate Hop Popping (PHP)

- VC label disposition

- VC and Tunnel label imposition

- Tunnel Label
  - VC Label
    - Label = 34
      - Label = 28
        - Payload
    - Label = 28
      - Payload
    - Label = 19
      - Label = 45
        - Payload
    - Label = 45
      - Payload
Configuring MS-PWs

**Cisco IOS**

```bash
hostname S-PE
interface Loopback0
 ip address 104.104.104.104 255.255.255.255

l2 vfi sample-ms-pw-1 point-to-point
neighbor 106.106.106.106 222190 encapsulation mpls
neighbor 102.102.102.102 111190 encapsulation mpls
```

```
7604-3#show xconnect peer 102.102.102.102 vcid 111190
Legend:  XC ST=Xconnect State  S1=Segment1 State  S2=Segment2 State
         UP=Up  DN=Down  AD=Admin Down  IA=Inactive
         SB=Standby  HS=Hot Standby  RV=Recovering  NH=No Hardware

XC ST Segment 1                       S1 Segment 2                       S2
-------------------------------------+---------------------------------+------
UP  mpls 106.106.106.106:222190     UP  mpls 102.102.102.102:111190     UP
 -------------------------------------+---------------------------------+------
```

```
7604-3#show xconnect peer 102.102.102.102 vcid 111190 detail
Legend:  XC ST=Xconnect State  S1=Segment1 State  S2=Segment2 State
         UP=Up  DN=Down  AD=Admin Down  IA=Inactive
         SB=Standby  HS=Hot Standby  RV=Recovering  NH=No Hardware

XC ST Segment 1                       S1 Segment 2                       S2
-------------------------------------+---------------------------------+------
UP  mpls 106.106.106.106:222190     UP  mpls 102.102.102.102:111190     UP
 -------------------------------------+---------------------------------+------
```

```
Local VC label 65536
Remote VC label 16029
pw-class:
```

```
Local VC label 65549
Remote VC label 47
pw-class:
```

S-PE labels for each PW segment

BRKMLP-2101 © 2014 Cisco and/or its affiliates. All rights reserved. Cisco Public
Configuring MS-PWs
Cisco IOS XR

hostname S-PE
interface Loopback0
  ipv4 address 106.106.106.106 255.255.255.255

l2vpn
  xconnect group Cisco-Live
    p2p xc-sample-8
      neighbor 102.102.102.102 pw-id 111200
      neighbor 104.104.104.104 pw-id 222200

RP/0/RSP0/CPU0:ASR9000-2# show l2vpn xconnect group Cisco-Live xc-name xc-sample-8
Legend: ST = State, UP = Up, DN = Down, AD = Admin Down, UR = Unresolved,
SB = Standby, SR = Standby Ready, (PP) = Partially Programmed

<table>
<thead>
<tr>
<th>XConnect Group</th>
<th>Name</th>
<th>ST</th>
<th>Description</th>
<th>ST</th>
<th>Description</th>
<th>ST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cisco-Live</td>
<td>xc-sample-8</td>
<td>UP</td>
<td>102.102.102.102 111200 UP</td>
<td></td>
<td>104.104.104.104 222200 UP</td>
<td></td>
</tr>
</tbody>
</table>
Advanced Topics
L2VPN Inter – Autonomous Systems (I-AS)
L2VPN Inter-AS

• Three (3) deployment models
• Option A
  – No reachability information shared between AS
• Option B
  – Minimal reachability information shared between AS
  – ASBR configured as S-PEs (multi-segment PWs)
  – eBGP (IPv4 prefix + label) used to build PSN tunnel between AS
• Option C
  – Significant reachability information shared between AS
  – Single-segment PW signaled across AS boundary
L2VPN Inter-AS Option B

Traffic direction:
- PE1 to ASBR1: Label 34, Label 28
- ASBR1 to PE2: Label 28, Label 19
- PE1 to PE2: Dummy tunnel label, Payload
- PE2 to PE1: Dummy tunnel label, Payload

VC and Tunnel label imposition:
- Push Push: VC labels swapped
- Pop Swap: BGP on NNI used to exchange tunnel label
- Swap Push: No LDP running on NNI
- Pop Pop: Penultimate Hop Popping (PHP)

Label values:
- PE1 to ASBR1: Label = 34, Label = 28
- ASBR1 to PE2: Label = 28, Label = 19
- PE1 to PE2: Label = 88
- PE2 to PE1: Label = 88

Payloads:
- PE1: Label 34, PE2: Label 88
- ASBR1: Label 28, ASBR2: Label 88
- PE1: Payload, PE2: Payload

Penultimate Hop Popping (PHP):
- VC label disposition

PW switching point:
- ASBR1 and ASBR2 exchange tunnel label
L2VPN Inter-AS Option C

**Traffic direction**

<table>
<thead>
<tr>
<th>Tunnel Label</th>
<th>VC Label</th>
<th>Label</th>
<th>Payload</th>
</tr>
</thead>
<tbody>
<tr>
<td>VC and Tunnel label imposition</td>
<td>Push</td>
<td>34</td>
<td>Push</td>
</tr>
<tr>
<td>Push</td>
<td>28</td>
<td>Push</td>
<td></td>
</tr>
<tr>
<td>Push</td>
<td>28</td>
<td>Push</td>
<td></td>
</tr>
<tr>
<td>Push</td>
<td>28</td>
<td>Push</td>
<td></td>
</tr>
<tr>
<td>Push</td>
<td>28</td>
<td>Push</td>
<td></td>
</tr>
<tr>
<td>Push</td>
<td>28</td>
<td>Push</td>
<td></td>
</tr>
</tbody>
</table>

**Penultimate Hop Popping (PHP)**

**BGP on NNI used to exchange tunnel label**

**VC label disposition**

**ASBR** Push of PSN label (if any) VC labels untouched

**ASBR1**

**ASBR2**

**RSVP-TE / LDP/iBGP**

**eBGP IPv4+Label**

**Targeted-LDP**

**CE-1**

**PE1**

**MPLS**

**Pseudowire 1**

**CE-2**

**PE2**

**MPLS**

**ASBR1**

**ASBR2**

**Pseudowire 1**
Advanced Topics
Resiliency
Pseudowire Redundancy
High Availability in L2VPN Networks

Solutions

- IP Fast Re-Route (FRR) / MPLS FRR
  - PSN core failure
- Pseudowire Redundancy:
  - PSN end-to-end routing failure – Redundant PEs
  - PE failure – Redundant PEs
  - Attachment circuit failure – AC Diversity
  - CE failure – Redundant CEs
One-Way Pseudowire Redundancy

Overview

- Allows dual-homing of one local PE to one or two remote PEs
- Two pseudowires - primary & backup provide redundancy for a single AC
- Faults on the primary PW cause failover to backup PW
- Multiple backup PWs (different priorities) can be defined
- Alternate LSPs (TE Tunnels) can be used for additional redundancy
One-Way Pseudowire Redundancy

Failure Protection Points

- **Failure 1** - Core failures handled by IGP re-routing / IP/MPLS FRR do not trigger pseudowire switchover

- **Failure 2** - Loss of route to remote PE as notified by IGP (PE isolation)

- **Failure 3** - Loss of Remote PE

- How to detect PE failures?
  - LDP Fast Failure Detection (FFD) (a.k.a. Route-Watch)
    - Monitors IGP route availability for LDP peer (2-3 sec or sub-sec with Fast IGP)
  - LDP session timeout (default = 3 x 30 sec)
  - BFD timeout (multi-hop PE-to-PE BFD session) (a.k.a. “xconnect client” IOS feature)
Pseudowire Redundancy

Preferential Forwarding Status Bit

• Extensions to PW status codes (RFC 6870)

• Allows PEs to signal local forwarding status of the PW (Active or Standby)

• A PW is selected for forwarding when declared as Active by both PEs

• Minimize service downtime during PW failover
  – Backup PWs always signaled before failures and held in Standby mode

• Allows VCCV capability over a backup PW
  – OAM over backup PWs
  – SP monitors backup PWs prior to its usage
Two-Way Pseudowire Redundancy

Overview

- Allows dual-homing of two local PEs to two remote PEs
- Four (4) pseudowires: 1 primary & 3 backup provide redundancy for dual-homed devices
- Two-Way PW redundancy coupled with Multi-Chassis LAG (MC-LAG) solution on the access side
  - LACP state used to determine PW AC state
  - InterChassis Communication Protocol (ICCP) used to synchronize LACP states
Configuring Pseudowire Redundancy

hostname PE1
interface Loopback0
  ip address 102.102.102.102 255.255.255.255

interface GigabitEthernet2/4
  service instance 170 ethernet
  encapsulation dot1q 170
  rewrite ingress tag pop 1 symmetric
  xconnect 104.104.104.104 170 encapsulation mpls

  backup peer 106.106.106.106 170 170
  mtu 1500

PE1
  102.102.102.102

PE2
  104.104.104.104

PE3
  106.106.106.106

CE1
  Site X

CE2a

CE2b
  Site Y

IP/MPLS

PW VC id

7604-2# show xconnect peer 104.104.104.104 vcid 170
Legend:  XC ST=Xconnect State  S1=Segment1 State  S2=Segment2 State
         UP=Up  DN=Down  AD=Admin Down  IA=Inactive
         SB=Standby  HS=Hot Standby  RV=Recovering
         NH=No Hardware

XC ST  Segment 1  S1 Segment 2  S2
+---------------------------------+          +---------------------------------+
UP pri ac  Gi2/4:170(Eth VLAN)  UP mpls 104.104.104.104:170  UP

Redundant PW configuration

Primary PW in UP state

Redundant PW in Standby state

hostname PE2
interface Loopback0
  ip address 104.104.104.104 255.255.255.255

interface GigabitEthernet2/4
  service instance 170 ethernet
  encapsulation dot1q 170
  rewrite ingress tag pop 1 symmetric
  xconnect 106.106.106.106 170 encapsulation mpls

  backup peer 106.106.106.106 170 170
  mtu 1500

PE2
  104.104.104.104

PE3
  106.106.106.106

CE1
  Site X

CE2a

CE2b
  Site Y

IP/MPLS

PW VC id

7604-2# show xconnect peer 106.106.106.106 vcid 170170
Legend:  XC ST=Xconnect State  S1=Segment1 State  S2=Segment2 State
         UP=Up  DN=Down  AD=Admin Down  IA=Inactive
         SB=Standby  HS=Hot Standby  RV=Recovering  NH=No Hardware

XC ST  Segment 1  S1 Segment 2  S2
+---------------------------------+          +---------------------------------+
IA sec ac  Gi2/4:170(Eth VLAN)  UP mpls 106.106.106.106:170170  SB
Configuring Pseudowire Redundancy

Cisco IOS XR

hostname PE1

interface Loopback0
  ipv4 address 106.106.106.106 255.255.255.255

interface GigabitEthernet0/0/0/2.180 l2transport
  encapsulation dot1q 180
  rewrite ingress tag pop 1 symmetric

l2vpn
  xconnect group Cisco-Live
    p2p xc-sample-6
      interface GigabitEthernet0/0/0/2.180
        neighbor 104.104.104.104 pw-id 180
        pw-class sample-CW-ON
        backup neighbor 102.102.102.102 pw-id 180180
        pw-class sample-CW-ON

RP/0/RSP0/CP0:ASR9000-2# show l2vpn xconnect group Cisco-Live xc xc-sample-6
Sun Apr 15 20:18:50.180 UTC
Legend: ST = State, UP = Up, DN = Down, AD = Admin Down, UR = Unresolved,
        SB = Standby, SR = Standby Ready, (PP) = Partially Programmed

<table>
<thead>
<tr>
<th>XConnect Group</th>
<th>Name</th>
<th>Segment 1</th>
<th>Segment 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Description</td>
<td>ST</td>
<td>Description</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------</td>
<td>-----------</td>
<td>-------------</td>
</tr>
<tr>
<td>Cisco-Live xc-sample-6</td>
<td>G10/0/0/2.180</td>
<td>UP</td>
<td>104.104.104.104</td>
</tr>
<tr>
<td></td>
<td>Backup</td>
<td></td>
<td>102.102.102.102</td>
</tr>
</tbody>
</table>

Primary PW in UP state
Redundant PW in Standby state
Deployment Use Cases

Data Center Interconnect – ASR 9000
Data Center Interconnect with VPLS
ASR 9000 Use Case 1 – nV Edge

- ASR 9000 as DC WAN Edge provides VPLS with Network Virtualization (nV) for DCI applications

- nV and VPLS provides:
  - Single-Chassis (Virtual) Redundancy solution – Network Virtualization Cluster
  - Access Multi-Homing solution with Multichassis EtherChannel
  - Single control and management plane, distributed data plane – single VFI / single PW between DC pairs
  - Flow-based load balancing over Pseudowire using Flow Aware Transport (FAT) PW
  - Scalability (MAC address table, number of VFIs / PWs)

ASR9000 sessions:
BRKARC-2003
BRKSPG-2904
Data Center Interconnect with VPLS

ASR 9000 Use Case 1 – nV Edge Sample Configuration

PE 1
hostname PE1
!
interface Loopback0
 ipv4 address 10.0.0.1 255.255.255.255
interface bundle-ethernet1.1 l2transport
 encapsulation dot1q 80
interface bundle-ethernet1.2 l2transport
 encapsulation dot1q 81
l2vpn
 pw-class sample-flow-lb
 encapsulation mpls
 load-balancing
 load-balancing flow-label
!
bridge group DCI
bridge-domain bd-80
interface bundle-ethernet1.1
vfi vfi1111
neighbor 10.0.0.2 pw-id 1111
 pw-class sample-flow-lb
!
bridge-domain bd-81
interface bundle-ethernet1.2
vfi vfi2222
neighbor 10.0.0.2 pw-id 2222
 pw-class sample-flow-lb

PE 2
hostname PE2
!
interface Loopback0
 ipv4 address 10.0.0.2 255.255.255.255
interface bundle-ethernet1.1 l2transport
 encapsulation dot1q 80
interface bundle-ethernet1.2 l2transport
 encapsulation dot1q 81
l2vpn
 pw-class sample-flow-lb
 encapsulation mpls
 load-balancing
 load-balancing flow-label
!
bridge group DCI
bridge-domain bd-80
interface bundle-ethernet1.1
vfi vfi1111
neighbor 10.0.0.1 pw-id 1111
 pw-class sample-flow-lb
!
bridge-domain bd-81
interface bundle-ethernet1.2
vfi vfi2222
neighbor 10.0.0.1 pw-id 2222
 pw-class sample-flow-lb

Note: nV cluster configuration not shown
Etherchannel configuration incomplete
Data Center Interconnect with VPLS
ASR 9000 Use Case 2 – ICCP-based Service Multi-Homing

- ASR 9000 as DC WAN Edge device provides VPLS with service multi-homing for DCI applications

- Service Multi-homing and VPLS provides:
  - Geo-Redundant dual-home DCI layer solution
  - Active / Active per VLAN load balancing
  - Distributed Control / Management / Data Plane
  - Forwarding state coordination via Inter-Chassis Communication Protocol (ICCP)
Deployment Use Cases
Data Center Interconnect – Catalyst 6500
Data Center Interconnect with VPLS
Catalyst 6500

- DC WAN Edge device (Catalyst 6500) implements VPLS with Advanced-VPLS (A-VPLS) for DCI applications

- A-VPLS provides:
  - Single-Chassis (Virtual) Redundancy solution – Virtual Switching System (VSS)
  - Multichassis EtherChannel (MEC)
  - Flow-based load balancing over WAN using Flow Aware Transport (FAT) PW
  - Simplified configuration
Data Center Interconnect with VPLS
Sample Configuration – Catalyst 6500

Virtual Ethernet interface modeled as Switchport trunk towards VFIs

hostname PE1

! interface Loopback0
 ip address 10.0.0.1 255.255.255.255
!
pseudowire-class sample-class
 encapsulation mpls
 load-balance flow
 flow-label enable

interface virtual-ethernet 1
 transport vpls mesh
 neighbor 10.0.0.2 pw-class sample-class
 switchport
 switchport mode trunk
 switchport trunk allowed vlan 80,81

hostname PE2

! interface Loopback0
 ip address 10.0.0.2 255.255.255.255
!
pseudowire-class sample-class
 encapsulation mpls
 load-balance flow
 flow-label enable

interface virtual-ethernet 1
 transport vpls mesh
 neighbor 10.0.0.1 pw-class sample-class
 switchport
 switchport mode trunk
 switchport trunk allowed vlan 80,81

Note: Complete Virtual Switching System (VSS) / Multichassis EtherChannel (MEC) configuration not shown