

airheads

TECH TALK *LIVE*

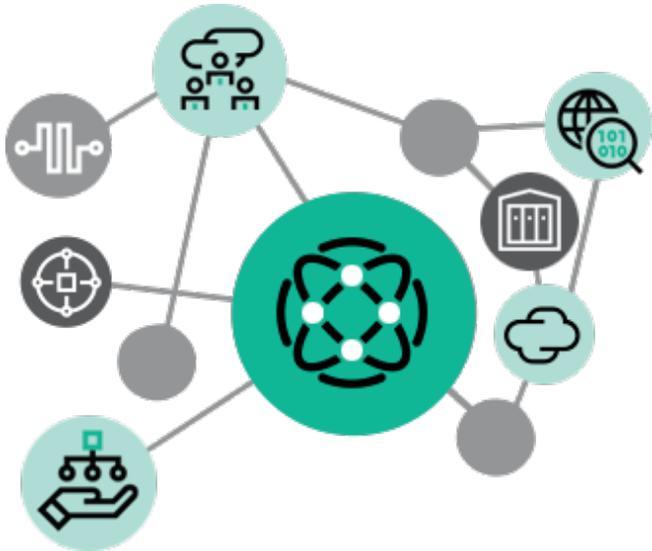
aruba
a Hewlett Packard
Enterprise company

AOS-CX 10.3 VSX Live Upgrade Demo

Vincent Giles - TME

#ArubaAirheads

Agenda



1 Live Upgrade

2 VXLAN/EVPN

3 DCB

4 DHCP Server

5 VSX

6 ERPS

7 1G support

8 VLAN translation

9 MSDP

10 NAE script upgrade

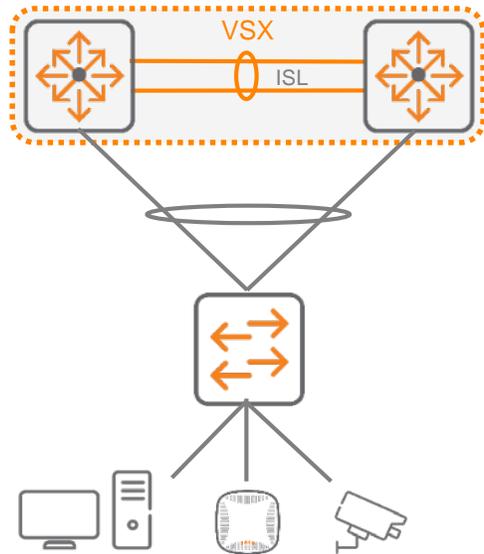
11 Miscellaneous

12 Conclusion

VSX Live Upgrade

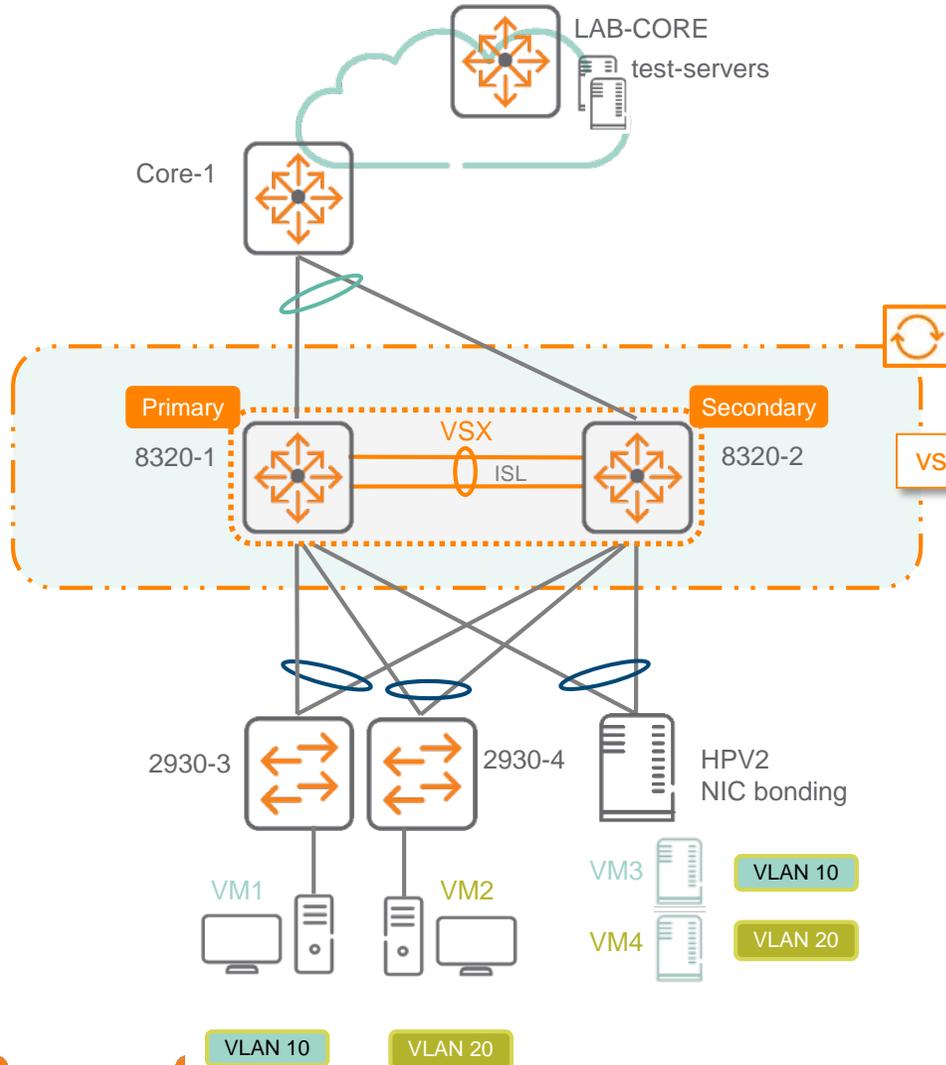
Aruba Virtual Switching Extension

VSX Key Principle: High Availability by design during Live Upgrade



- Bind 2 same AOS-CX switches to operate as one device for L2 but as independent nodes for L3.
- Support for active-active data-path:
 - Active-active L2
 - Active-active L3 unicast
 - Active-active L3 multicast
- Operational simplicity and usability:
 - for configuration
 - for troubleshooting
- Similar VSF benefits with better HA during upgrade

VSX Live Upgrade



```

8320-1# vsx update-software tftp://10.136.40.99/TL_10_03_0020.swi vrf mgmt

Do you want to save the current configuration (y/n)? y
The running configuration was saved to the startup configuration.

This command will download new software to the primary image of both VSX primary and
secondary systems, then reboot them in sequence. The VSX secondary will reboot first,
followed by primary.

Continue (y/n)?y

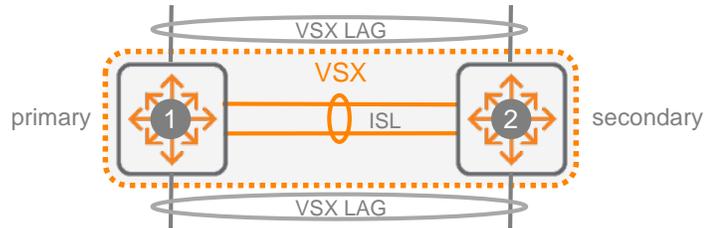
VSX Primary Software Update Status      : Reboot started
VSX Secondary Software Update Status    : Image updated successfully
VSX ISL Status                          : Down
Progress [#####.]

Secondary VSX system updated completely. Rebooting primary.
    
```

Upgrade Process Orchestration

Primary VSX node Progress Status

Step 0



```
8400-1# vsx update-software tftp://15.136.40.99/XL_10_02_0001BN.swi vrf mgmt
Do you want to save the current configuration (y/n)? y
The running configuration was saved to the startup configuration.
```

This command will download new software to the secondary image of both VSX primary and secondary systems, then reboot them in sequence. The VSX secondary will reboot first, followed by primary.

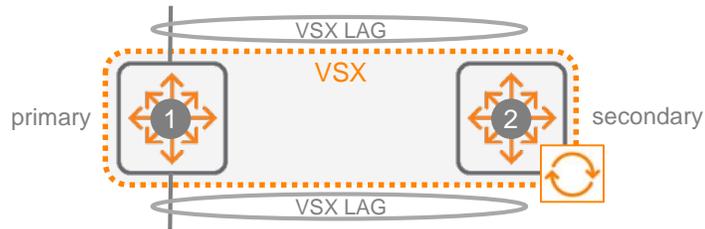
```
Continue (y/n)? y

VSX Primary Software Update Status      : Image download started
VSX Secondary Software Update Status    : Image download started
VSX ISL Status                           : Up
Progress [#####.....]
[]
```

1st impact

Step 1

Failover < 100 ms

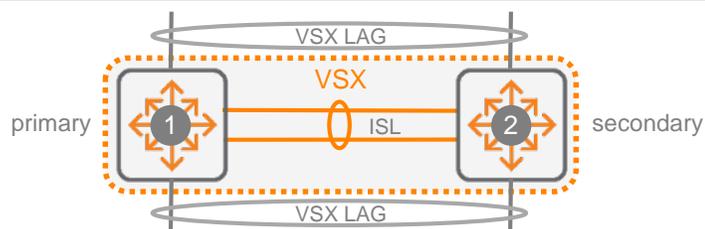


```
VSX Primary Software Update Status      : Waiting for VSX secondary to reboot complete
VSX Secondary Software Update Status    : Rebooting
VSX ISL Status                           : Down
Progress [#####.....]
[]
```

2nd impact

Step 2

Failback < 100 ms

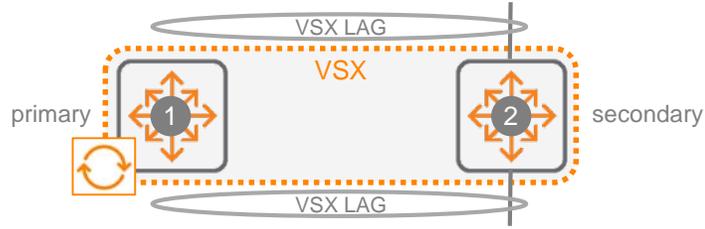


```
VSX Primary Software Update Status      : Waiting for VSX sync to complete
VSX Secondary Software Update Status    : Image updated successfully
VSX ISL Status                           : Up
Progress [#####.....]
[]
secondary# sh vsx status linkup-delay
Initial sync status                      : Completed
Delay timer status                       : Running
Linkup Delay time left                   : 2 minutes 22 seconds
secondary# sh vsx status linkup-delay
Initial sync status                      : Completed
Delay timer status                       : Completed
Linkup Delay time left                   :
```

3rd impact

Step 3

Failover < 100 ms



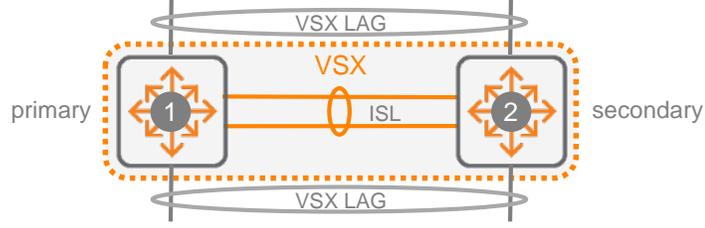
```
VSX Primary Software Update Status      : Reboot started
VSX Secondary Software Update Status    : Image updated successfully
VSX ISL Status                           : Down
Progress [#####.....]

Secondary VSX system updated completely. Rebooting primary.
8400-1# []
```

4th impact

Step 4

Failback < 100 ms

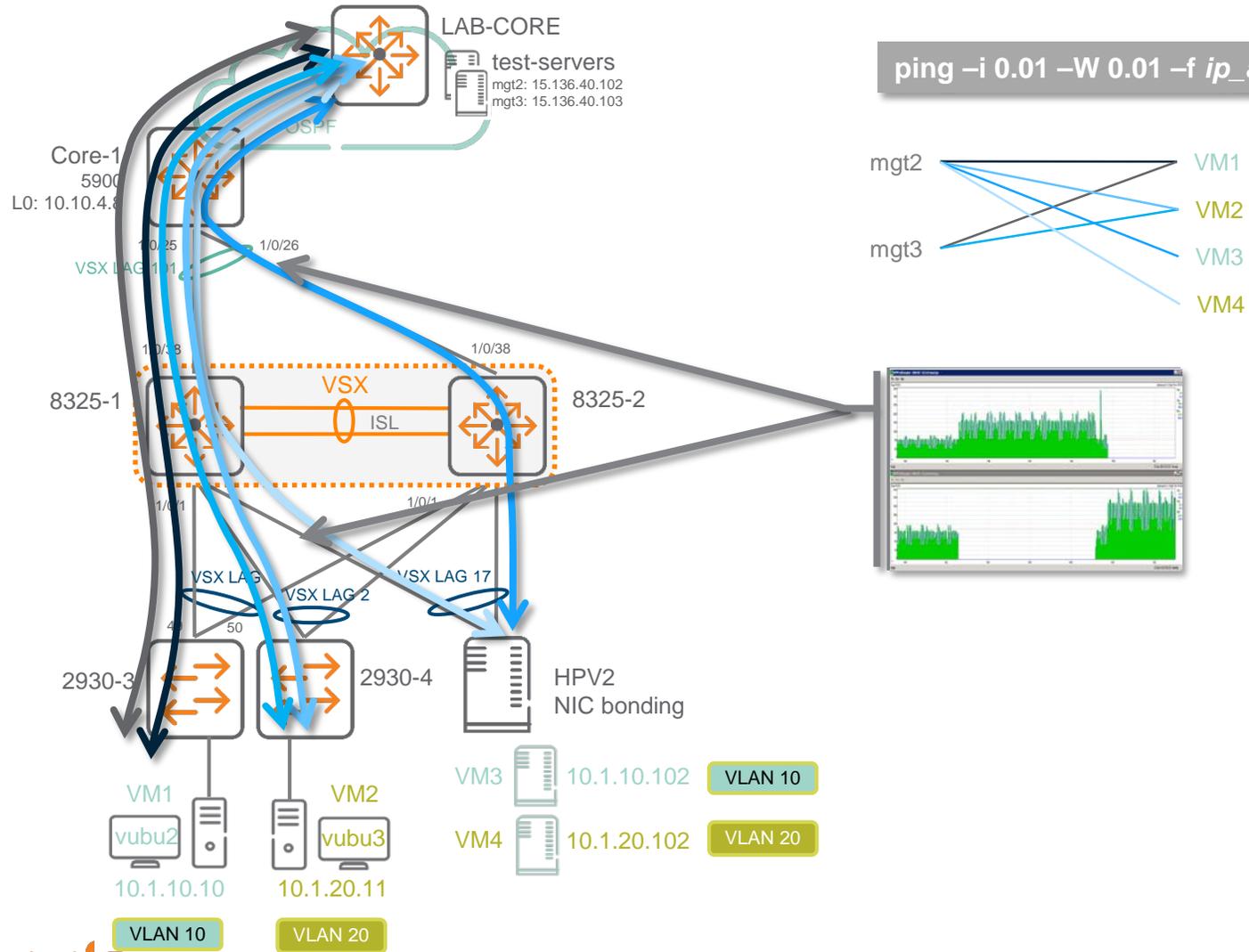


```
8400-1# sh vsx status linkup-delay
Configured linkup delay-timer           : 180 seconds
Initial sync status                     : Completed
Delay timer status                       : Running
Linkup Delay time left                   : 1 minutes 58 seconds

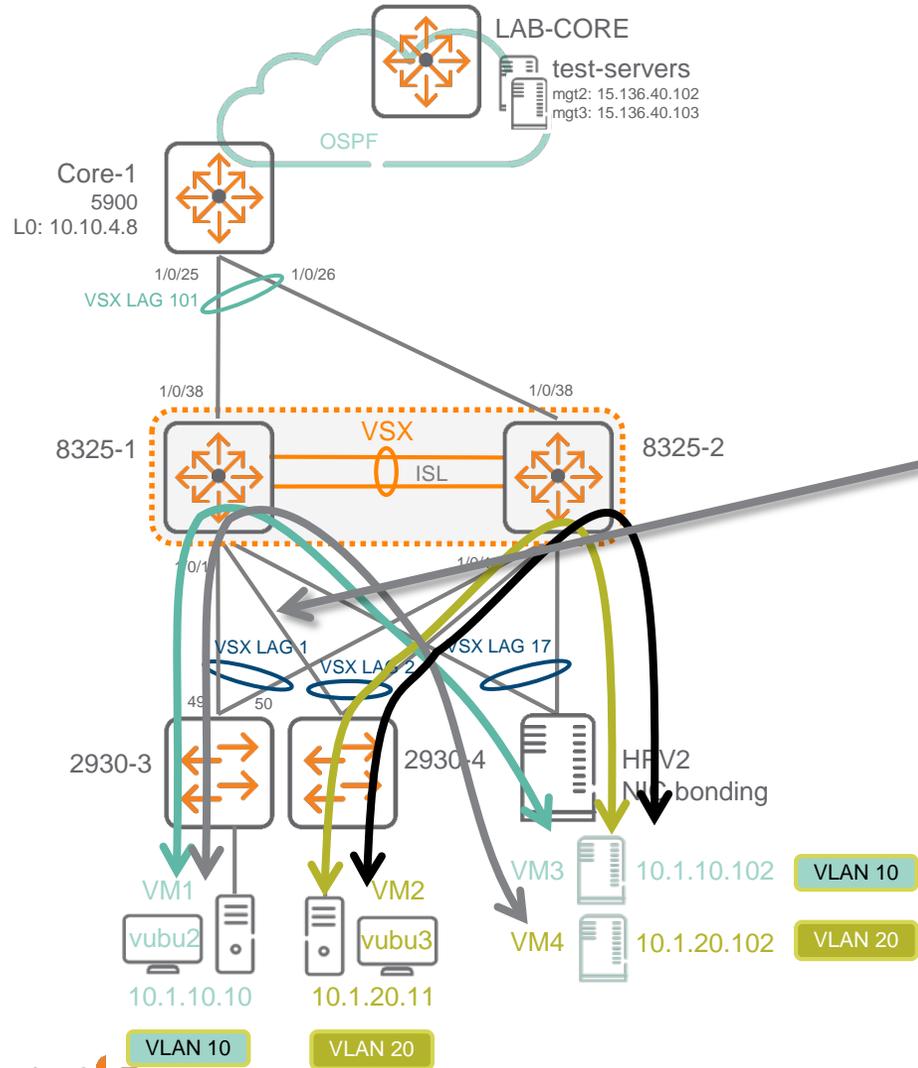
8400-1# sh vsx status linkup-delay
Configured linkup delay-timer           : 180 seconds
Initial sync status                     : Completed
Delay timer status                       : Completed
Linkup Delay time left                   :
```



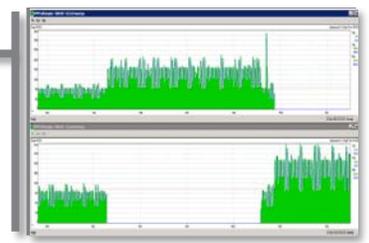
North-South unicast test flows (L3)



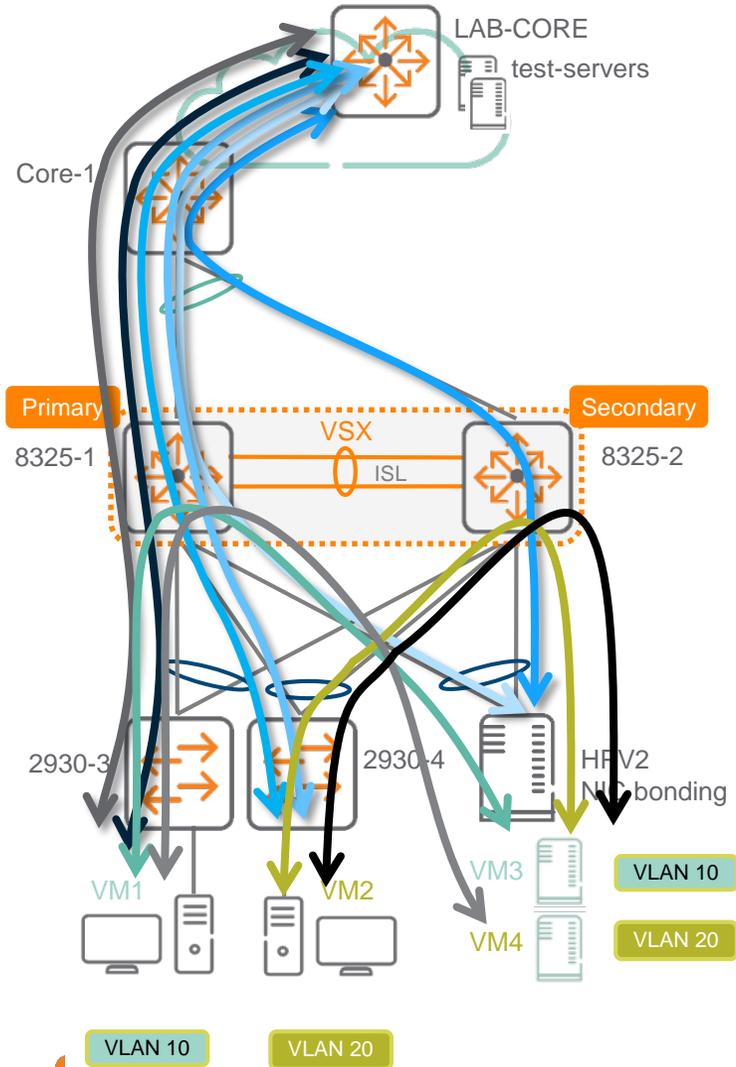
East-West unicast test flows (L2 + L3)



```
ping -i 0.01 -W 0.01 -f ip_address
```



Impact on unicast test flows



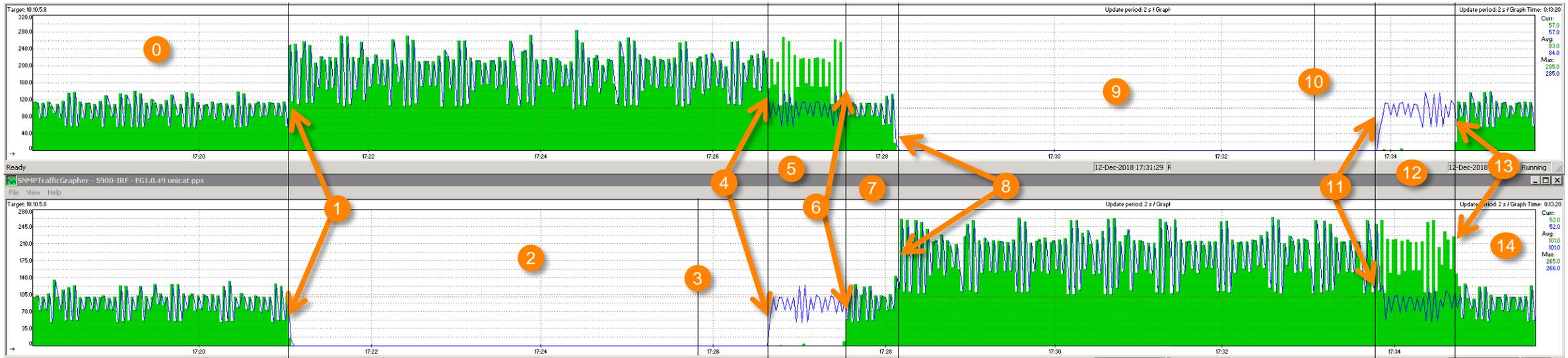
TOTAL IMPACT
=
Secondary failover impact
+
Secondary failback impact
+
Primary failover impact
+
Primary failback impact
=

<200ms

70ms shown @Utrecht

Traffic transition on upstream links

■ outbound
■ inbound

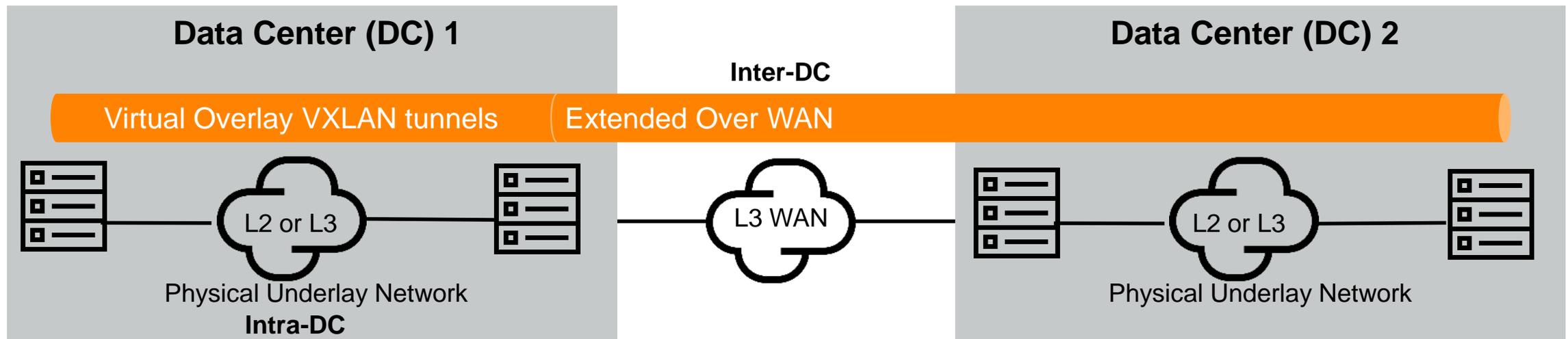


- 0 Nominal situation
Primary and Secondary run version n
- 1 Secondary failover
- 2 Secondary is rebooting
- 3 Secondary starts to join VSX cluster
- 4 Secondary is restarted (end of ISL synchronization)
- 5 Linkup-delay phase
ASIC programming
Upstream LAG is restored
- 6 End of linkup-delay timer
Downstream VSX LAGs are restored
- 7 Secondary runs version n+1
Primary runs version n
Secondary sends "Ready" notification to primary
- 8 Primary failover
- 9 Primary is rebooting
- 10 Primary starts to join VSX cluster
- 11 Primary is restarted (end of ISL synchronization)
- 12 Linkup-delay phase
ASIC programming
Upstream LAG is restored
- 13 End of linkup-delay timer
Downstream VSX LAGs are restored
- 14 Upgrade is completed.
VSX nodes run version n+1

VXLAN / EVPN

VXLAN and Overlay Networking Introduction

- Virtual Extensible Local Area Network (**VXLAN**) is a network encapsulation mechanism that supports up to **16 million virtual network identifiers (VNIs)** over a physical layer 2/3 underlay network for L2 network connectivity and multi-tenancy (vs 4K VLAN limit)
- The endpoints which does VXLAN encapsulation and de-capsulation are called **VTEPs** (VXLAN Tunnel End Point)
- UDP based VXLAN allows traffic to be load shared across multiple equal cost paths (vs TCP based GRE protocol)
- One of the limitations of static VXLAN flood-n-learn is the inherent flooding that is required ensuring that learning happens at the VTEPs.



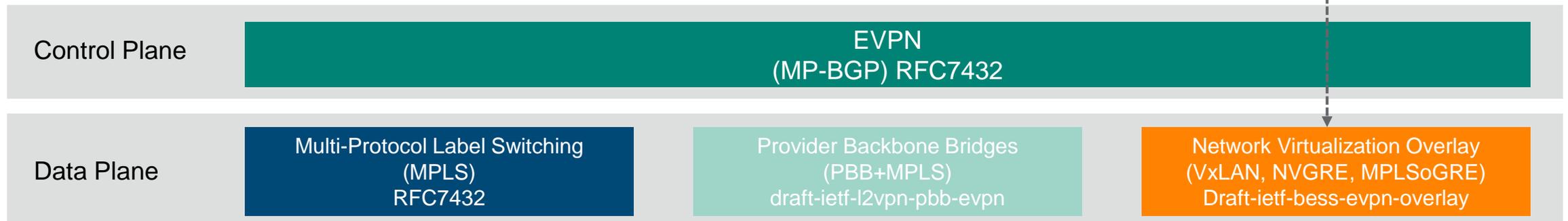
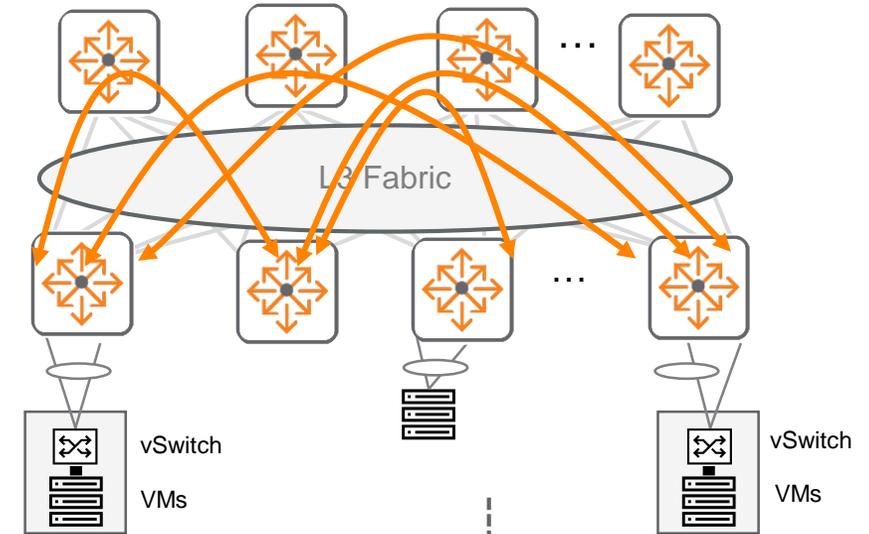
- Provides L2/L3 network connectivity for VMs/Containers/Servers between racks

- Provides L2/L3 network connectivity for VMs/Containers/Servers between DCs

MP-BGP EVPN as a distributed VXLAN control plane

- Resilient & Efficient
- Secure
- Scalable
- Open Standards

Overlay: VXLAN & EVPN
Underlay: Spine/Leaf L3 ECMP (Equal Cost Multi Path Routing)

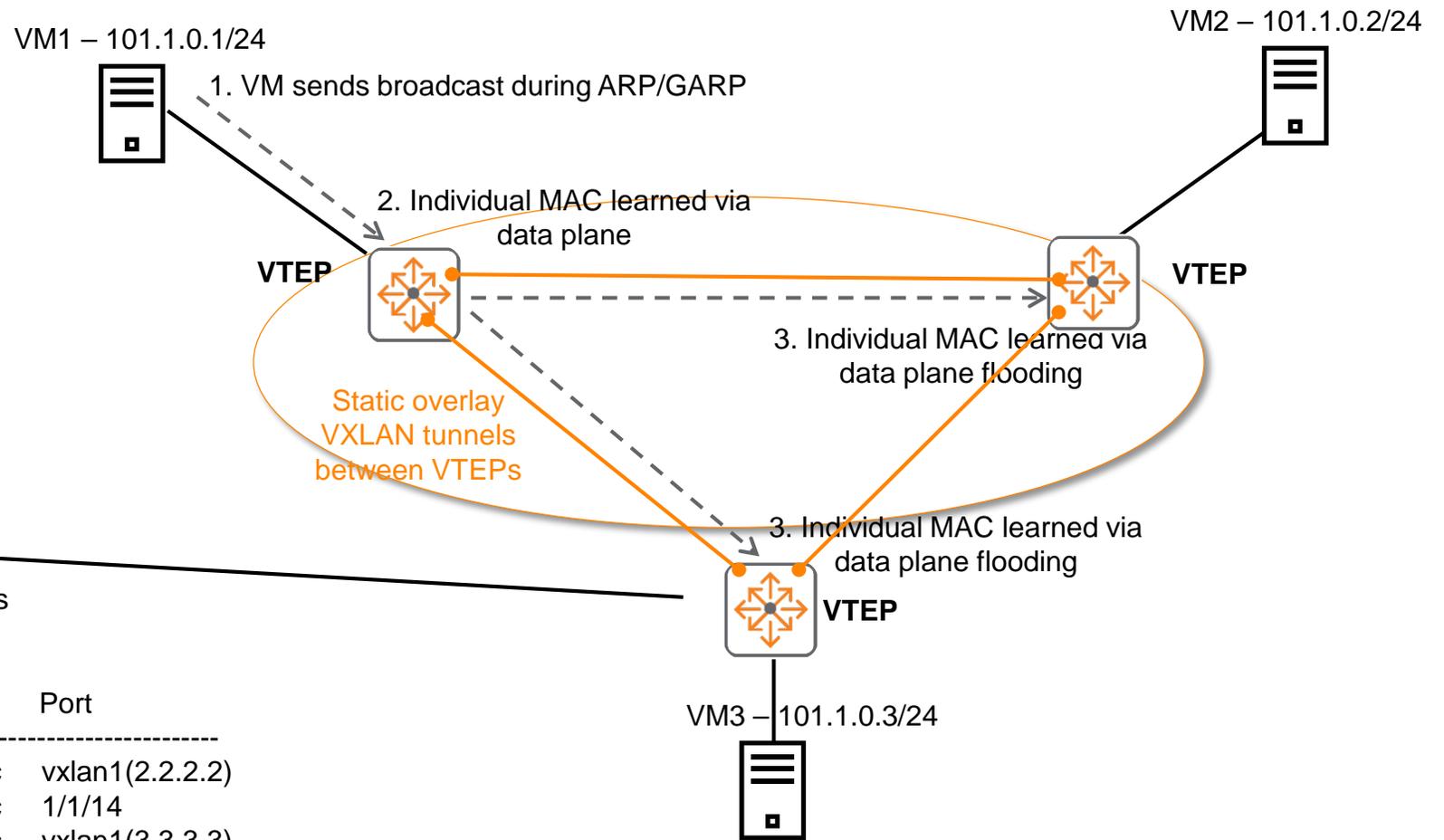


Before EVPN: Static VXLAN Tunnels / MAC Flood & Learn

Sample Configuration

```

vlan 10,20
interface vxlan 1
 source ip 1.1.1.1
 no shutdown
 vni 10
  vlan 10
  vtep-peer 2.2.2.2
  vtep-peer 3.3.3.3
 vni 20
  vlan 20
  vtep-peer 2.2.2.2
  vtep-peer 3.3.3.3
  
```



8325# sh mac-address-table
 MAC age-time : 300 seconds
 Number of MAC addresses : 4

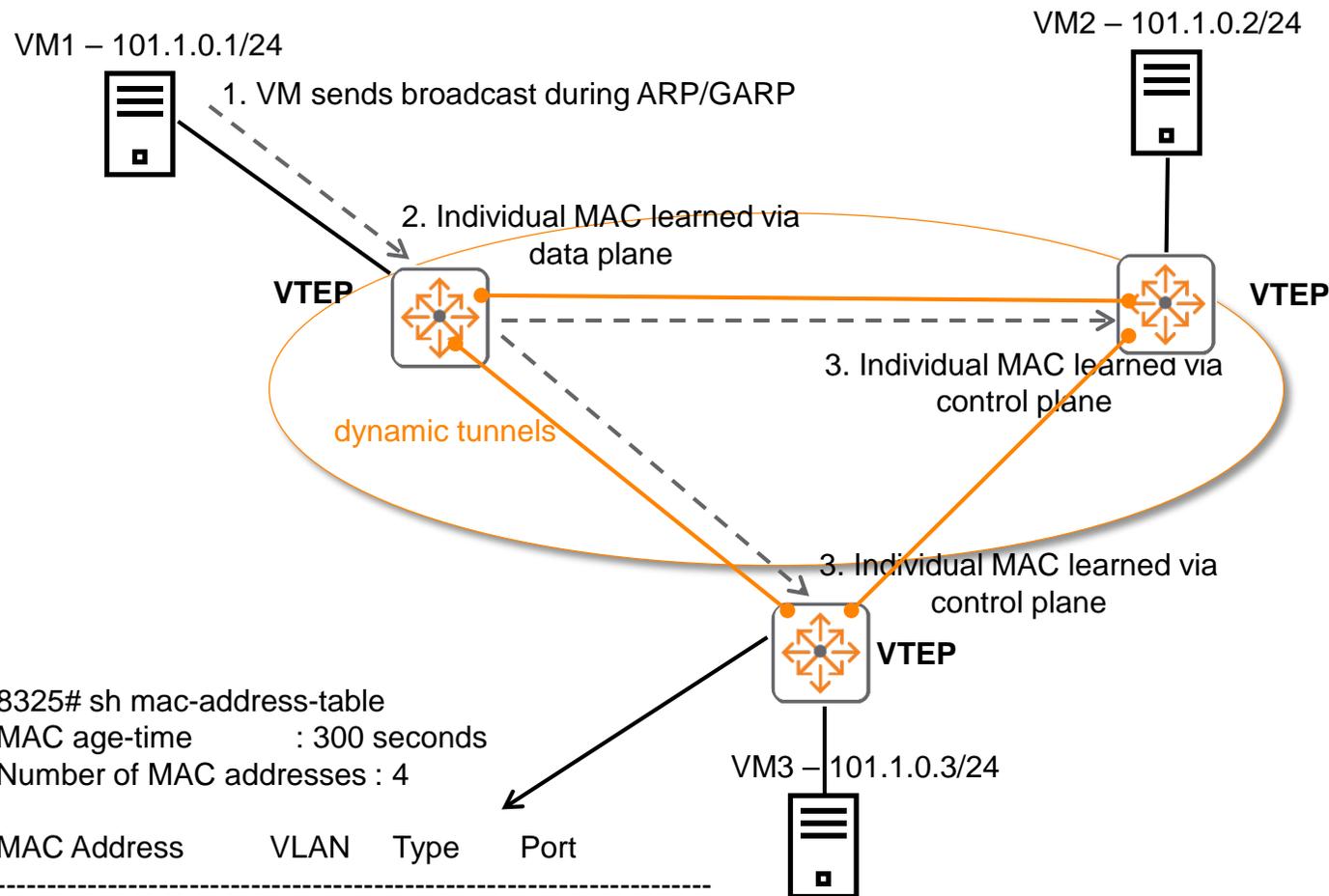
	MAC Address	VLAN	Type	Port
MAC addresses of VMs	d07e-28cf-9900	10	dynamic	vxlan1(2.2.2.2)
	d07e-28cf-9940	10	dynamic	1/1/14
	d07e-28cf-9980	10	dynamic	vxlan1(3.3.3.3)

With EVPN: Dynamic VXLAN Tunnels / Scalable Control Plane Learning

Sample Configuration

```

vlan 10,20
evpn
vlan 10
rd 65001:10
route-target export 65001:10
route-target import 65001:10
vlan 20
rd 65001:20
route-target export 65001:20
route-target import 65001:20
interface vxlan 1
source ip 1.1.1.1
no shutdown
vni 10
vlan 10
vni 20
vlan 20 ← • Manual tunnels not required
           • Scales when there are more remote VTEPs
router bgp 65001
bgp router-id 1.1.1.1
neighbor 2.2.2.2 remote-as 65001
neighbor 2.2.2.2 update-source lo 0
neighbor 3.3.3.3 remote-as 65001
neighbor 3.3.3.3 update-source lo 0
address-family l2vpn evpn
neighbor 2.2.2.2 activate
neighbor 2.2.2.2 send-community
neighbor 3.3.3.3 activate
neighbor 3.3.3.3 send-community
    
```



```

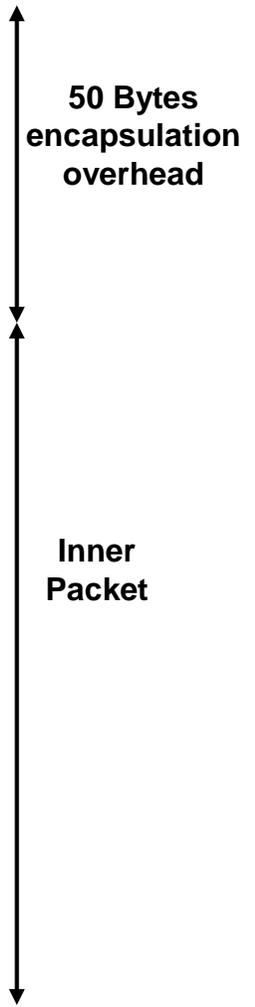
8325# sh mac-address-table
MAC age-time      : 300 seconds
Number of MAC addresses : 4
    
```

MAC Address	VLAN	Type	Port
d07e-28cf-9900	10	evpn	vxlan1(2.2.2.2)
d07e-28cf-9940	10	dynamic	1/1/14
d07e-28cf-9980	10	evpn	vxlan1(3.3.3.3)

VXLAN Data Plane

Frame	Length	Source	Destination	Protocol	Length	Description
46	20.127848	100.1.0.11	100.1.0.1	IPv4	558	any host internal protocol (61)
47	20.127977	100.1.0.11	100.1.0.1	IPv4	558	any host internal protocol (61)
48	20.127978	100.1.0.11	100.1.0.1	IPv4	558	any host internal protocol (61)

- ▶ Frame 46: 558 bytes on wire (4464 bits), 558 bytes captured (4464 bits) on interface 0
- ▶ Ethernet II, Src: 54:80:28:fd:28:00 (54:80:28:fd:28:00), Dst: 54:80:28:fd:54:00 (54:80:28:fd:54:00)
- ▶ Internet Protocol Version 4, Src: 192.168.1.101, Dst: 192.168.1.102 **AOS-CX VTEP IPs**
- ▶ User Datagram Protocol, Src Port: 37706, Dst Port: 4789
- ▼ Virtual eXtensible Local Area Network
 - ▶ Flags: 0x0800, VXLAN Network ID (VNI)
Group Policy ID: 0
 - VXLAN Network Identifier (VNI): 10
 - Reserved: 0
 - ▶ Ethernet II, Src: Cisco_00:00:01 (00:12:01:00:00:01), Dst: CetTechn_00:00:01 (00:11:01:00:00:01)
 - ▼ Internet Protocol Version 4, Src: 100.1.0.11, Dst: 100.1.0.1 **VM IPs (Simulated via IXIA)**
 - 0100 = Version: 4
 - 0101 = Header Length: 20 bytes (5)
 - ▶ Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
Total Length: 494
Identification: 0x0000 (0)
 - ▶ Flags: 0x00
Fragment offset: 0
Time to live: 64
Protocol: any host internal protocol (61)
Header checksum: 0xb0c5 [validation disabled]
[Header checksum status: Unverified]
Source: 100.1.0.11
Destination: 100.1.0.1
[Source GeoIP: Unknown]
[Destination GeoIP: Unknown]
 - ▶ Data (474 bytes)



EVPN Control Plane

```
11 8.683362 192.168.1.101 192.168.1.102 BGP 912 UPDATE Message, UPDATE Message, UPDATE Message, UPDATE Message,
12 8.732378 192.168.1.101 192.168.1.102 TCP 66 179 34003 [ACK] Seq= 854 Len= 545 Win= 30000 Len= 0 TC=1 2510224
▶ Frame 11: 912 bytes on wire (7296 bits), 912 bytes captured (7296 bits) on interface 0
▶ Ethernet II, Src: 54:80:28:fd:28:00 (54:80:28:fd:28:00), Dst: 54:80:28:fd:54:00 (54:80:28:fd:54:00)
▶ Internet Protocol Version 4, Src: 192.168.1.101, Dst: 192.168.1.102
▶ Transmission Control Protocol, Src Port: 179, Dst Port: 34003, Seq: 85, Ack: 85, Len: 846
▶ Border Gateway Protocol - UPDATE Message
▶ Border Gateway Protocol - UPDATE Message
▼ Border Gateway Protocol - UPDATE Message
  Marker: ffffffffffffffffffffffffffffffffff
  Length: 104
  Type: UPDATE Message (2)
  Withdrawn Routes Length: 0
  Total Path Attribute Length: 81
  ▼ Path attributes
    ▶ Path Attribute - ORIGIN: INCOMPLETE
    ▶ Path Attribute - AS_PATH: empty
    ▶ Path Attribute - LOCAL_PREF: 100
    ▶ Path Attribute - EXTENDED_COMMUNITIES
    ▼ Path Attribute - MP_REACH_NLRI
      ▶ Flags: 0x90, Optional, Extended-Length, Non-transitive, Complete
      Type Code: MP_REACH_NLRI (14)
      Length: 44
      Address family identifier (AFI): Layer-2 VPN (25)
      Subsequent address family identifier (SAFI): EVPN (70)
      Next hop network address (4 bytes)
      Number of Subnetwork points of attachment (SNPA): 0
      ▼ Network layer reachability information (35 bytes)
        ▼ EVPN NLRI: MAC Advertisement Route EVPN type 2 route
          Route Type: MAC Advertisement Route (2)
          Length: 33
          Route Distinguisher: 0000fde90000000b (65001:11)
          ▶ ESI: 00 00 00 00 00 00 00 00
          Ethernet Tag ID: 0
          MAC Address Length: 48
          MAC Address: 54:80:28:fd:f4:00 (54:80:28:fd:f4:00) VM MAC
          IP Address Length: 0
          ▶ IP Address: NOT INCLUDED VM IP, only included with L3 VXLAN
```

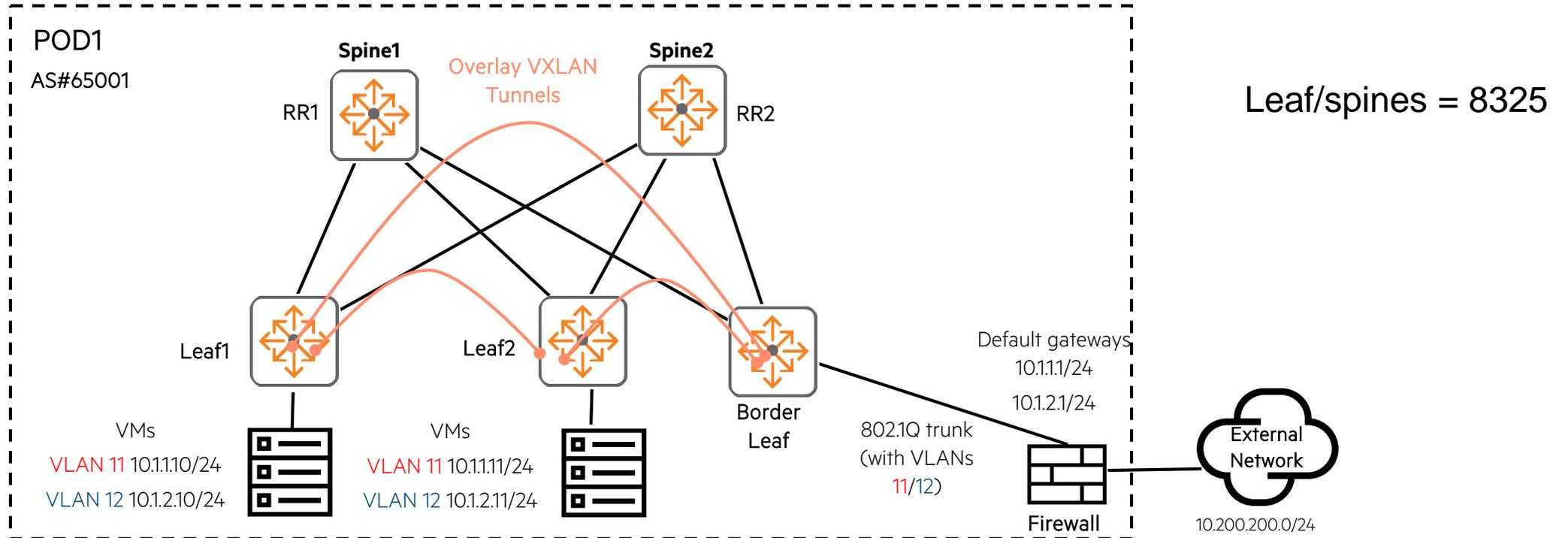
Multiple update messages

Update message info

Benefits of EVPN VXLAN

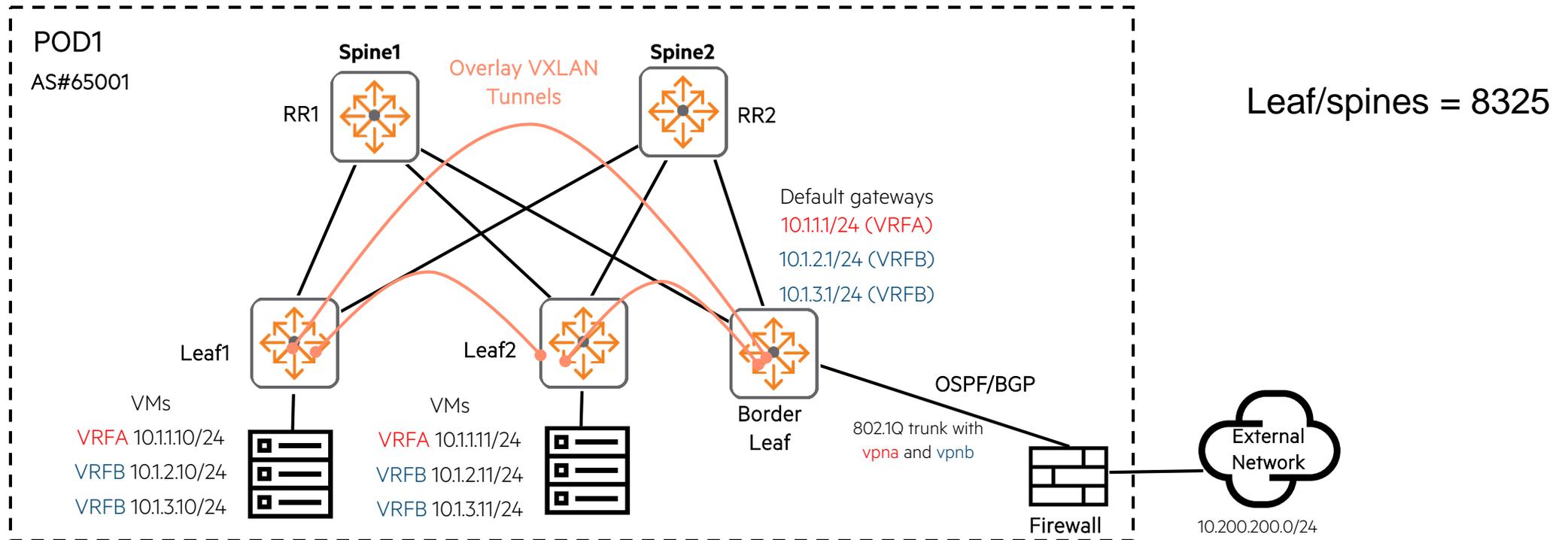
- Overcome 4K VLAN limit
- Stretch L2 Segment across DC
- Reduce flooding traffic
- Mac or VM mobility
- Egress load balancing (ECMP)
- No trunking and no spanning-tree
(VNI's are layer-2 and run across the layer-3 network. To make this possible, VXLAN switches encapsulate layer-2 frames in layer-3 packets).

DC Use Case - Centralized L2 Gateway with VXLAN/EVPN



- Centralized L2 gateway is typically used when centralized firewall functions as default gateway.
- Traffic on the same subnet between VTEPs does not need to traverse border leaf

DC Use Case - Centralized L3 Gateway with VXLAN/EVPN



- Centralized L3 gateway (border leaf) is typically used when Centralized Firewall inspection is required between VRF subnets
- Traffic is placed into different VRFs and sent to firewall for inspection
- Above example shows VRFA (subnet 10.1.1.0/24), VRFB (subnet 10.1.2.0/24, 10.1.3.0/24), traffic will be VXLAN encapsulated at the L2 VTEPs, sent to the default gateways (VRFA 10.1.1.1, VRFB 10.1.2.1/10.1.3.1) and routed out
- The centralized L3 gateway VTEP is able to exchange routes with the centralized firewall via OSPF/BGP
- Traffic on the same subnet between VTEPs does not need to traverse border leaf

AOS-CX 10.3 supports

- The 2 use cases mentioned in previous slides (QA validated)
- Only 8325 supports VXLAN, IBGP EVPN and IBGP Route Reflector functionality
- EVPN type 2 - MAC/IP advertisement route
 - Advertises MAC reachability information and host route information (host ARP or ND information)
- EVPN type 3 - Inclusive multicast Ethernet tag (IMET) route
 - Advertises VTEP and VXLAN mappings for automating VTEP discovery, VXLAN tunnel establishment, and VXLAN tunnel assignment
- IPv4 L3 unicast routing in the overlay network
- IPv4 L2 multicast BUM traffic in the overlay network
- IPv4 VTEPs in the unicast underlay network
- 1:1 VNI/VLAN mapping
- VXLAN/EVPN should always be recommended for production networks
- Static VXLAN only supports L2 VXLAN, cannot scale and is prone to CLI errors (10 VTEPs max as a general rule)

AOS-CX 10.3 does not support

- VSX/VXLAN cannot be used together (VSX & VXLAN interoperability is intended for next major release)
- IPv6 or PIM multicast routing in the overlay network
- IPv6 VTEPs in the underlay network
- Static VXLAN and EVPN VXLAN cannot be used together

8325 VXLAN and EVPN Scale (AOS-CX 10.3)

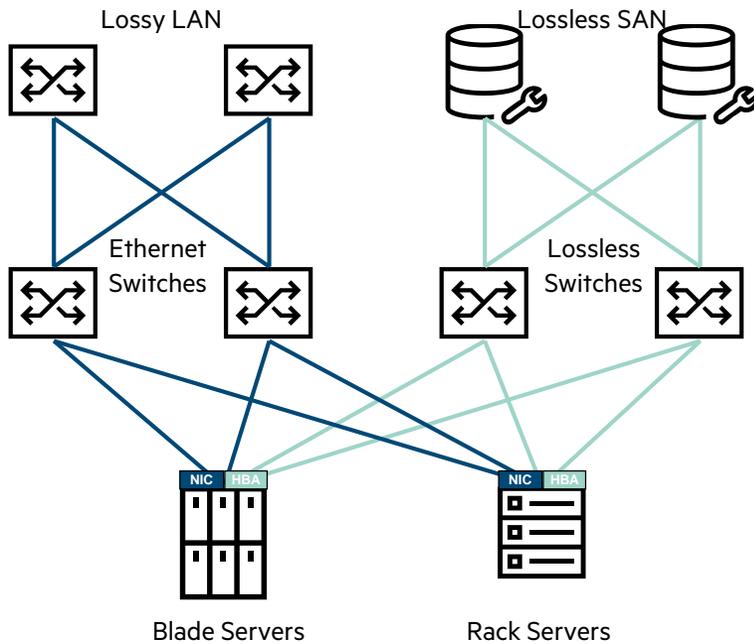
	8325 L3-agg	8325 Leaf	8325 Spine / L3 Core
ARP	120 000	120 000	28 000
MAC	98 304	98 304	32 768
IPv4 Routes	28 672	28 672	131 072
Underlay Hosts*	48 640	12 286	32,768
VXLAN Overlay Hosts*	8192	32,768	12,288
VXLAN VTEP Peers	1,024	1,024	1,024
VXLAN L2 VNI (VLANs)	N/A	4,040	4,040
VXLAN L2 VNI per VTEP	N/A	4,040	4,040
VXLAN VTEP per L2 VNI	N/A	256	256
VXLAN Virtual Ports	4,096	15,872	4,096
EVPN Leaf Nodes	N/A	256	256
EVPN Spine Nodes	N/A	N/A	8
EVPN iBGP RR	N/A	N/A	8

Data Center Bridging DCB

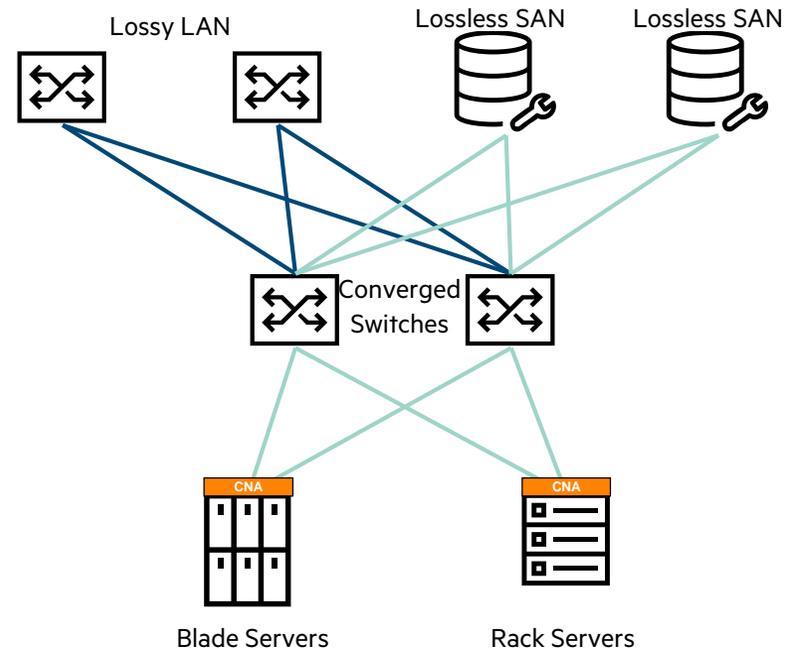
Convergence

Convergence = Using common cabling/switching infrastructure to replace separate server and storage networks.

Traditional Separate SAN



Converged



— Ethernet

— Lossless SAN

Convergence Benefits

- Fewer physical HW switches (4 per ToR vs 2 per ToR)
- Less need for optics/cables
- Fewer NIC/HBAs
- Simplified infrastructure
- \$\$\$ saving!!



Lossless Ethernet uses Data Center Bridging (DCB) protocols:

- **ETS = Enhanced Transmission Selection** for ensure lossless and lossy traffic have a minimum amount of bandwidth
- **PFC = Priority-based Flow Control** informs the devices to not drop any packets in that queue
- **QCN = Quantized Congestion Notification** for monitoring and throttling traffic – L2 only
- **DCBx = Data Center Bridging Exchange** which exchanges information between attached device and simplifies configurations

Not official DCB protocol – but heavily used in lossless IP networks.

- + **IP ECN = Explicit Congestion Notification** informs initiators and targets to throttle traffic in times of congestion – helps to avoid pause frames

Configuring and verifying DCBx

Enable/Disable globally

```
switch(config)# [no] lldp dcbx
```

Enable/Disable on an interface

```
switch(config-if)# [no] lldp dcbx
```

Map iSCSI to priority 5

```
switch(config)# dcbx application iscsi priority 4
```

Verify DCBx status on an interface

```
switch# show dcbx interface 1/1/1
```

```
8325-1(config)# dcbx application
ether          The application 802.3 Ethernet type
iscsi          The iSCSI application (TCP ports 860 and 3260)
tcp-sctp       Traffic for a specified TCP or SCTP port
tcp-sctp-udp   Traffic for a specified TCP or SCTP or UDP port
udp            Traffic for a specified UDP port
```

DCBx example configuration

```
8325-1# show dcbx interface 1/1/1
DCBX admin state: disabled
DCBX operational state : inactive
```

```
8325-1# conf
8325-1(config)# lldp dcbx
8325-1(config)# dcbx application iscsi priority 5
8325-1(config)# end
```

```
8325-1# show dcbx interface 1/1/1
DCBX admin state: enabled
DCBX operational state : active
```

Priority Flow Control (PFC)

Operational state : inactive

Local advertisement:

```
Willing : No
MacSec ByPass Capability : No
Max traffic classes : 1
```

Priority	Enabled
0	False
1	False
2	False
3	False
4	False
5	False
6	False
7	False

Enhanced Transmission Selection (ETS)

Local advertisement:

```
Willing : No
MacSec ByPass Capability : No
Max traffic classes : 8
```

Priority	Traffic Class
0	1
1	0
2	2
3	3
4	4
5	5
6	6
7	7

Traffic Class	Bandwidth Percentage	Algorithm
0	12	ETS
1	12	ETS
2	12	ETS
3	12	ETS
4	12	ETS
5	12	ETS
6	12	ETS
7	16	ETS

Application Priority Map

Local advertisement:

Protocol	Port/Type	Priority
iscsi		5

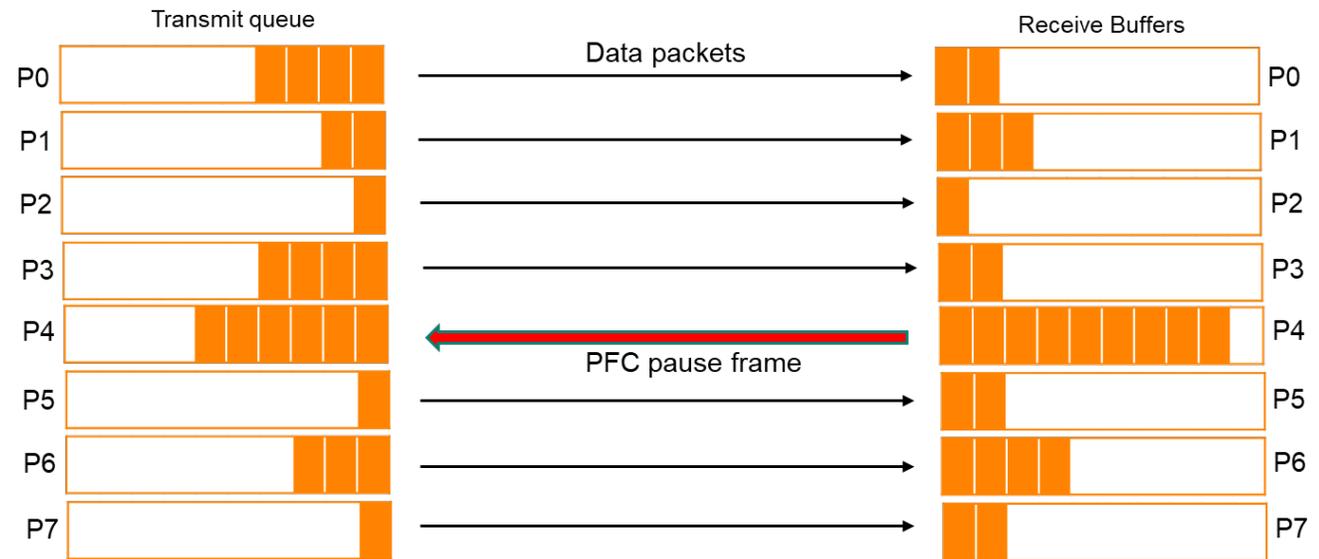
DCBx Summary

- Enable DCBx globally.
- DCBx messages are advertised to peer when LLDP is enabled on the interface and when one of PFC, ETS or Application priority is configured.
- Configure PFC on the interface to support one lossless priority.
- By default all priorities get equal bandwidth. This can be changed using QoS DWRR weights.
- DCB is supported on 8325 platform only.
- Host/target/switch need to support DCBx
- DCB solution is recommended to be used in a lossless iSCSI environment.
- **FCoE / RoCE** are not supported.

Data Center Bridging

Priority Based Flow Control (PFC)

- Primarily used for storage style networking where lossless is required
 - Lossless iSCSI
 - RDMA over Converged Ethernet (RoCE)
 - NVMeOF
 - FCoE (Not supported in AOS-CX)
- PFC works similar to IEEE 802.3x PAUSE for global flow control
 - IEEE 802.3x is not suitable for different flows with different QoS
- PFC enables pause per HW queue on an Ethernet device
- PFC uses the 802.1p CoS values in 802.1Q VLAN tag to differentiate up to eight levels of CoS
- PFC must be enabled on all endpoints and switches in the flow path
- Currently (10.3) supported on 8325



Priority Based Flow Control (PFC)

10.3 Configuration Restrictions

Due to ASIC restriction, PFC configuration changes are not enabled immediately

- The switches must be rebooted to apply PFC configuration when:
 - PFC is configured for the first time
 - An interface's PFC priority is modified
 - PFC is configured on a previously-enabled interface
- Only ONE PFC Priority can be configured per interface
- At most, 3 different PFC Priorities can be configured on the same switch

10.3 – Priority Based Flow Control (PFC)

Configuration

- The existing link-level flow control commands have been extended to add PFC support
- If the configuration cannot take effect immediately the CLI will always display a message

Enabling PFC

```
switch(config-if)# flow control priority 4
```

```
8325-R1-RU27(config-if)# flow-control
  priority  Enable IEEE 802.3Q priority-based flow control
  rx        Enable RX flow control
8325-R1-RU27(config-if)# flow-control priority
  <0-7> The packet priority to flow control
8325-R1-RU27(config-if)# flow-control priority 4
```

The setting will not be applied until configuration is saved to startup-config and the switch is rebooted.

Priority Based Flow Control (PFC)

10.3 - Configuration Best Practices

— Before allowing traffic

- **Configure global QoS COS map using 'qos cos-map' (or qos queue-profile)**
 - Any code point(s) to be used for PFC must be assigned unique local-priorities (i.e., no other code points may be assigned to any local-priority used for PFC)
- **In converged environments configure a 'qos schedule-profile'**
 - The schedule profile determines the order in which queues are selected to transmit a packet
 - Two options:
 - All queues use the same scheduling algorithm (for example, WFQ).
 - The highest queue number uses strict priority, and all remaining (lower) queues use the same algorithm (for example, WFQ).
- **For each interface:**
 - Configure 'qos trust cos'
 - Configure 'flow-control priority'
- **Copy running config to startup-config**
- **Reboot**

- Any interfaces to be used for PFC must already be configured at boot in order for the ASIC memory to be properly configured for lossless PFC.
- Changing the global cos-map and queue-profile is not recommended while the switch is enabled for traffic.

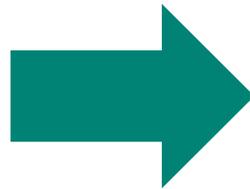
10.3 – Priority Based Flow Control (PFC)

Understanding the “qos cos-map”

Modifying the default CoS-Map

```
switch(config)# qos cos-map 4 local-priority 1 name iSCSI
```

```
8325-R1-RU27# sh qos cos-map default
code_point local_priority color name
-----
0          1             green Best_Effort
1          0             green Background
2          2             green Excellent_Effort
3          3             green Critical_Applications
4          4             green Video
5          5             green Voice
6          6             green Internetwork_Control
7          7             green Network_Control
```



```
8325-R1-RU27(config)# show qos cos-map
code_point local_priority color name
-----
0          0             green
1          0             green
2          0             green
3          0             green
4          1             green iSCSI
5          0             green
6          0             green
7          0             green
```

Changing QoS CoS Map to default

```
switch(config)# no qos cos-map 4
```

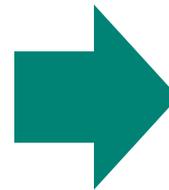
10.3 – Priority Based Flow Control (PFC)

Understanding the “qos schedule-profile”

Creating the schedule profile and assigning weights

```
switch(config)# qos schedule-profile myschedule  
switch(config-schedule)# dwrr queue 1 weight 15
```

```
8325-R1-RU27(config-if)# show qos schedule-profile factory-default  
queue_num algorithm weight max-bandwidth_kbps burst_KB  
-----  
0          dwrr      1  
1          dwrr      1  
2          dwrr      1  
3          dwrr      1  
4          dwrr      1  
5          dwrr      1  
6          dwrr      1  
7          dwrr      1
```



```
8325-R1-RU27(config-schedule)# show qos schedule-profile myschedule  
queue_num algorithm weight max-bandwidth_kbps burst_KB  
-----  
0          dwrr     15  
1          dwrr     15  
2          dwrr      1  
3          dwrr      1  
4          dwrr      1  
5          dwrr      1  
6          dwrr      1  
7          dwrr      1
```

Apply QoS Schedule-Profile to interface

```
switch(config-if)# apply qos schedule-profile myschedule
```

DHCP Server

DHCP Server Details

- Both DHCP Server v4 and v6 support.
- DHCP Server can be configured and controlled per VRF.
- Direct CLI commands for widely used options like DNS Server, Domain name etc.
 - domain-name <name>**
- Raw options support
 - option <2-254> {ascii <string> | hex <string> | ip <ip-address>}**
- Separate CLI context for v4 and v6 configuration for ease of use
- The recommended baseline limit is as follows, individually for IPv4 and IPv6:
 - Pools per VRF – 64**
 - Ranges per pool – 64**
 - Total clients per system – 8192** (i.e. 8192 for IPv4 clients + 8192 for IPv6 clients)
- Mutual exclusiveness with DHCP Relay

DHCP Configuration

```
switch(config)# dhcp-server vrf default
switch(config-dhcp-server)# pool test
switch(config-dhcp-server-pool)# range 10.0.0.2 10.0.0.30
switch(config-dhcp-server-pool)# range 10.0.0.2 10.0.0.254 prefix-len 24
switch(config-dhcp-server-pool)# default-router 10.0.0.1
switch(config-dhcp-server-pool)# lease 00:02:00
switch(config-dhcp-server-pool)# exit
switch(config-dhcp-server)# enable
switch(config-dhcp-server)#
```

- Configure DHCP Server instance on a particular VRF that includes pool and its corresponding range, options if any, lease duration, etc.
- To verify whether the DHCP Server is configured correctly and enabled, check the operational state from the command “show dhcp-server vrf <vrf-name>”. Sample output is below:

```
switch# show dhcp-server vrf default

VRF Name           : default
DHCP Server        : enabled
Operational State  : operational
Authoritative Mode : false

Pool Name          : test
Lease Duration     : 00:02:00
```

VSX Enhancements

VSX Enhancements Summary



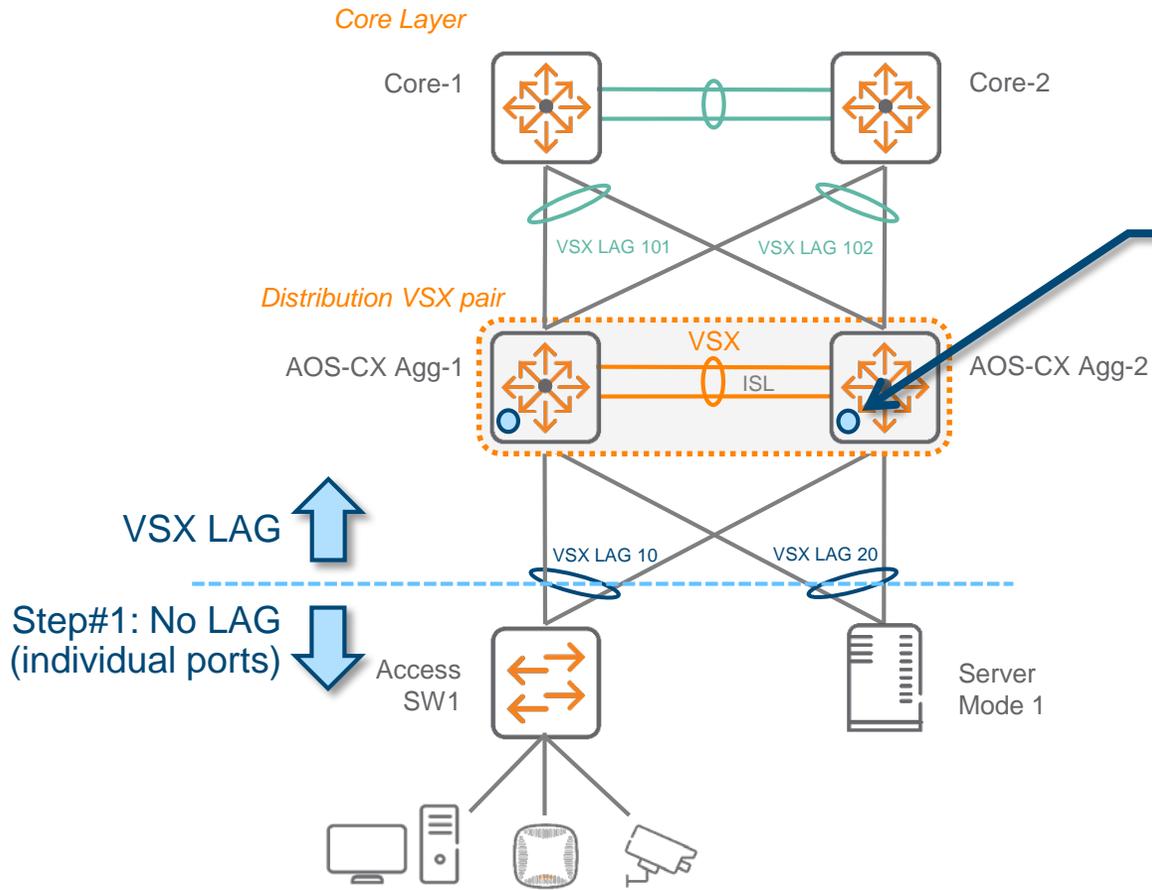
Feature	10.0	10.1	10.2	10.3	Future
VSX + Spanning-tree (MSTP or RPVST+)	No	No	MSTP only	MSTP / RPVST+	
Multicast Active-Active	No	No	Yes: Dual-DR		
VSX static LAG	No	No	Yes		
VSX system-mac	No	No	Yes		
VSX split-recovery	No	No	Yes		
VSX LACP fallback	No	No	Yes: Partial (starting 10.02.0020) ⁽¹⁾		Yes ⁽¹⁾
MVRP (Multiple VLAN Registration Protocol)	No	No	No	No	No
VSX active-gateway and VRRP	No	No	No	Yes: Global co-existence, mutually exclusive per SVI	
VSX active-gateway multinetting	No	No	No	No	Yes
LACP graceful shutdown (during VSX live upgrade)	No	No	No	No	Yes
OSPF and BGP graceful shutdown (during VSX live upgrade)	No	No	No	Yes	
VSX with BGP EVPN VXLAN	No	No	No	No	Yes
Keepalive over OOBM	No	No	No	No	tbd
DHCP relay (active on primary, standby on secondary)	No	Yes			
DHCP server and lease synchronization within VSX	No	No	No	Yes	
Gratuitous ARP on active-gateway (sent by primary)	No	No	Yes: Partial (starting 10.02.0020) ⁽²⁾		
VSX Live Upgrade orchestration from CLI	No	No	Yes		
VSX Live Upgrade orchestration from WebUI	No	No	No	No	tbd
VSX-sync (pseudo single management plane)	No	VLANs, ACLs, Class, Policy	+ feature-group tags	VLAN range sync + new feature-group tags	+ new feature-group tags (routing)
VSX linkup delay optimization	No	No	No	Partial ⁽³⁾	Yes

VSX LAG

LACP fallback

VSX LAG

LACP fallback

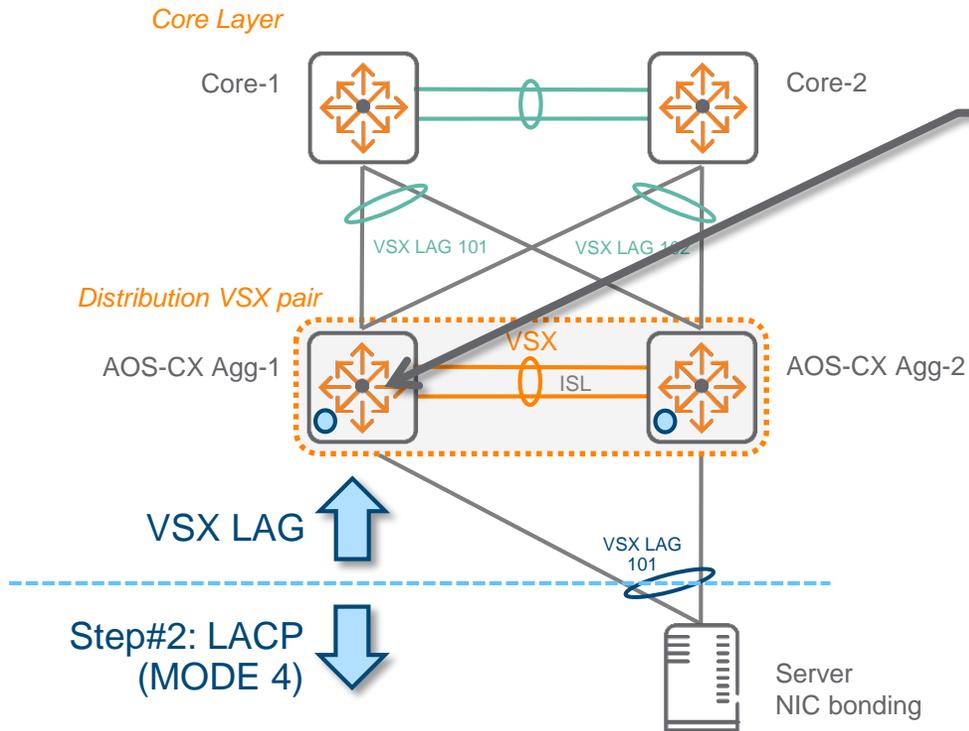


LACP fallback

1. LACP fallback on VSX LAG port will make members of the lag function as non-bonded interfaces when no LACP partner is detected.
2. This configuration is applicable for VSX LAG and ignored otherwise. Even if LACP fallback command is accepted on standard / non-VSX LAG, the fallback feature will work only on VSX LAG (multi-chassis LAG) interface.
3. Once LACP BPDUs are received, LACP status will change from "Individual" to Collecting/Distributing. On server, LACP status will change from Mode 1 (active-backup) to Mode 4 (IEEE 802.3ad).
4. Use-case examples:
 - a) PXE boot: server running PXE software is unaware of NIC teaming and selects any NIC for image download considering it as an individual logical interface. VSX fallback feature ensures that DHCP reply and OS download is performed on the same NIC from which the request was originated.
 - b) vSphere migration from VSS to DVS. During ESXi on-boarding in vcenter, an important step is to migrate the Service-Console to DVS. VSX fallback feature ensures that there is no connectivity loss to vmkernel using same vmnics during VSS to VDS transition.
 - c) ZTP for access switches. This scenario is supported with MSTP.

LACP fallback

Server - once LACP BPDUs are received



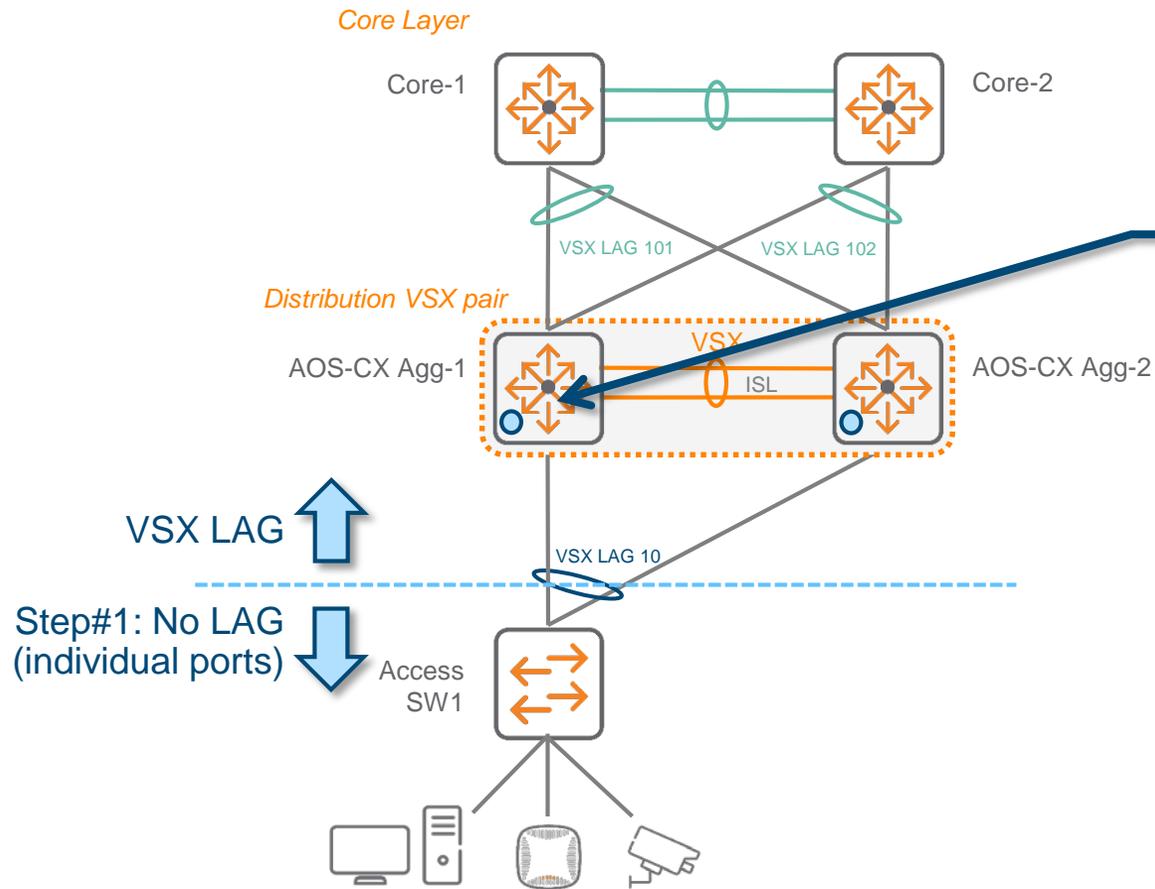
```
8325-1# sh lacp int
State abbreviations :
A - Active          P - Passive          F - Aggregable I - Individual
S - Short-timeout  L - Long-timeout  N - InSync     O - OutofSync
C - Collecting     D - Distributing
X - State m/c expired      E - Default neighbor state

Actor details of all interfaces:
-----
Intf   Aggr   Port  Port  State  System-ID          System Aggr  Forwarding
      Name  Id    Pri   State  ID                 Pri   Key   State
-----
1/1/55 lag1    56    1     ALFNCD 54:80:28:fc:ac:00 65534 1     up
1/1/56 lag1    57    1     ALFNCD 54:80:28:fc:ac:00 65534 1     up
1/1/49 lag49(mc) 49    1     ALFNCD 00:00:00:83:25:01 65534 49    up
1/1/1  lag101(mc) 1     1     ALFNCD 00:00:00:83:25:01 65534 101   up

Partner details of all interfaces:
-----
Intf   Aggr   Port  Port  State  System-ID          System Aggr
      Name  Id    Pri   State  ID                 Pri   Key
-----
1/1/55 lag1    56    1     ALFNCD 54:80:28:fd:42:00 65534 1
1/1/56 lag1    57    1     ALFNCD 54:80:28:fd:42:00 65534 1
1/1/49 lag49(mc) 70    1     ALFNCD 00:00:00:01:01:01 65534 49
1/1/1  lag101(mc) 0     65534 ALFNCD 00:00:00:00:00:00 0     0
```

LACP fallback

Access Switch ZTP

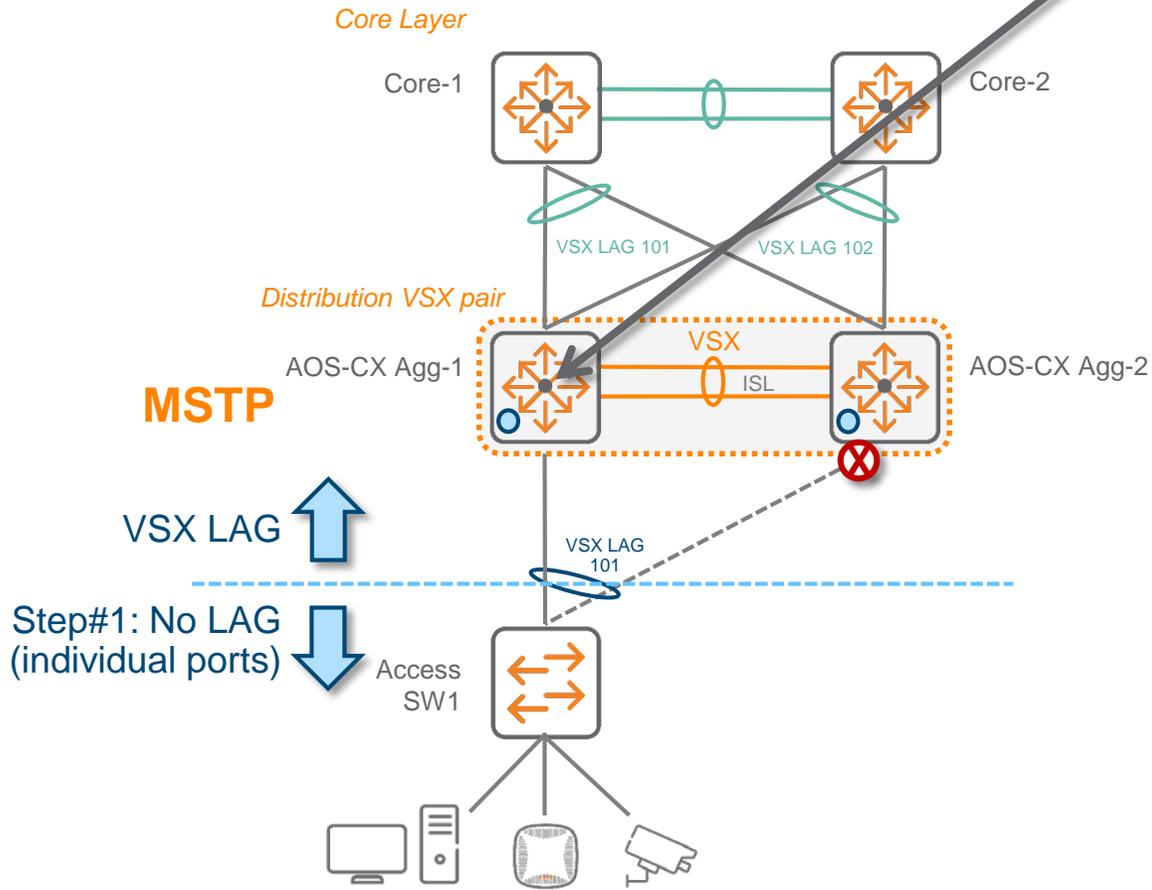


LACP fallback for access switch ZTP

1. It does support any brand of Access Switch.
2. The validated scenario is with factory-default AOS Switch. On AOS:
 - all ports UP
 - all ports member of native VLAN 1 (untagged)
 - no spanning tree running
3. **WARNING:** On AOS-CX, a L2 loop prevention mechanism must be implemented as links individualization of the VSX LAG will create a L2 loop through the access switch that will forward any Ethernet frame.
4. 3 mechanisms:
 - a) MSTP
This is the current supported option. MSTP process learns about the unbinding of the links which are member of the VSX LAG. Consequently, MSTP BPDUs are sent independently on each link. Link to secondary will become blocked.
 - b) RPVST
Not supported yet for this scenario.
 - c) Loop-protect
Not supported yet for this scenario.

LACP fallback

Access Switch ZTP – factory default



8325-1# show lacp interface multi-chassis

State abbreviations :
 A - Active P - Passive F - Aggregable I - Individual
 S - Short-timeout L - Long-timeout N - InSync O - OutofSync
 C - Collecting D - Distributing
 X - State m/c expired E - Default neighbor state

Actor details of all interfaces:

Intf	Aggregate name	Port id	Port Priority	State	System-ID	System Priority	Aggr Key
1/1/49	lag49(mc)	49	1	ALFNCD	00:00:00:83:25:01	65534	49
1/1/1	lag101(mc)	1	1	IE	00:00:00:83:25:01	65534	101

Partner details of all interfaces:

Intf	Aggregate name	Partner Port-id	Port Priority	State	System-ID	System Priority	Aggr Key
1/1/49	lag49(mc)	70	1	ALFNCD	00:00:00:01:01:01	65534	49
1/1/1	lag101(mc)	0	65534	IE	00:00:00:00:00:00	0	0

Remote Actor details of all interfaces:

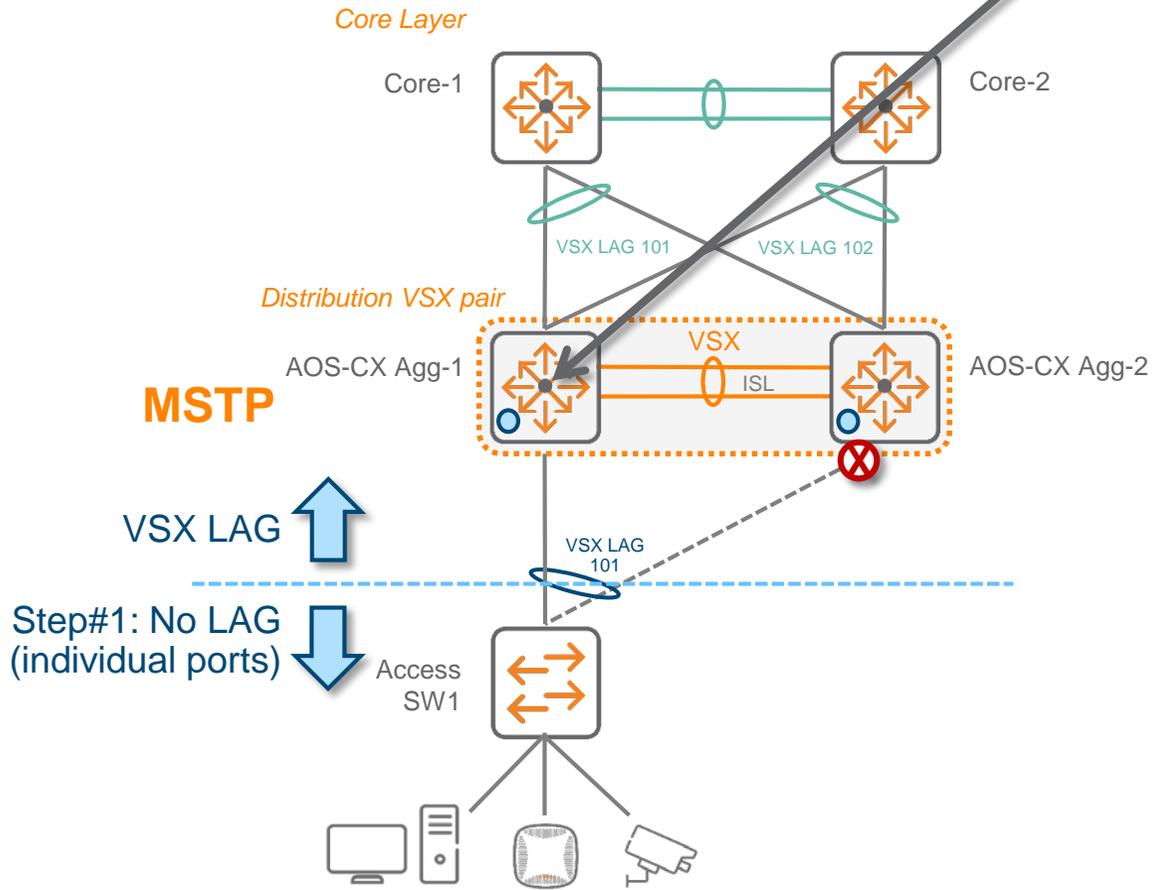
Intf	Aggregate name	Port id	Port Priority	State	System-ID	System Priority	Aggr Key
1/1/49	lag49(mc)	1049	1	ALFNCD	00:00:00:83:25:01	65534	49
1/1/1	lag101(mc)	1001	1	IE	00:00:00:83:25:01	65534	101

Remote Partner details of all interfaces:

Intf	Aggregate name	Partner Port-id	Port Priority	State	System-ID	System Priority	Aggr Key
1/1/49	lag49(mc)	1070	1	ALFNCD	00:00:00:01:01:01	65534	49
1/1/1	lag101(mc)	0	65534	IE	00:00:00:00:00:00	0	0

LACP fallback

Access Switch ZTP – factory default



```
8325-1# sh run spanning-tree
vsx-sync stp
spanning-tree
spanning-tree config-name stp1
spanning-tree config-revision 1
```

```
8325-1# sh spanning-tree
Spanning tree status      : Enabled Protocol: MSTP
```

```
MST0
Root ID   Priority   : 32768
MAC-Address: 00:00:00:83:25:01
This bridge is the root
Hello time(in seconds):2  Max Age(in seconds):20
Forward Delay(in seconds):15
```

```
Bridge ID Priority   : 32768
MAC-Address: 00:00:00:83:25:01
Hello time(in seconds):2  Max Age(in seconds):20
Forward Delay(in seconds):15
```

Port	Role	State	Cost	Priority	Type
lag1	Designated	Forwarding	1	64	point_to_point
lag49	Designated	Forwarding	500	64	point_to_point
lag101	Designated	Forwarding	20000	64	point_to_point

```
8325-1# sh spanning-tree vsx
Spanning tree status      : Enabled Protocol: MSTP
```

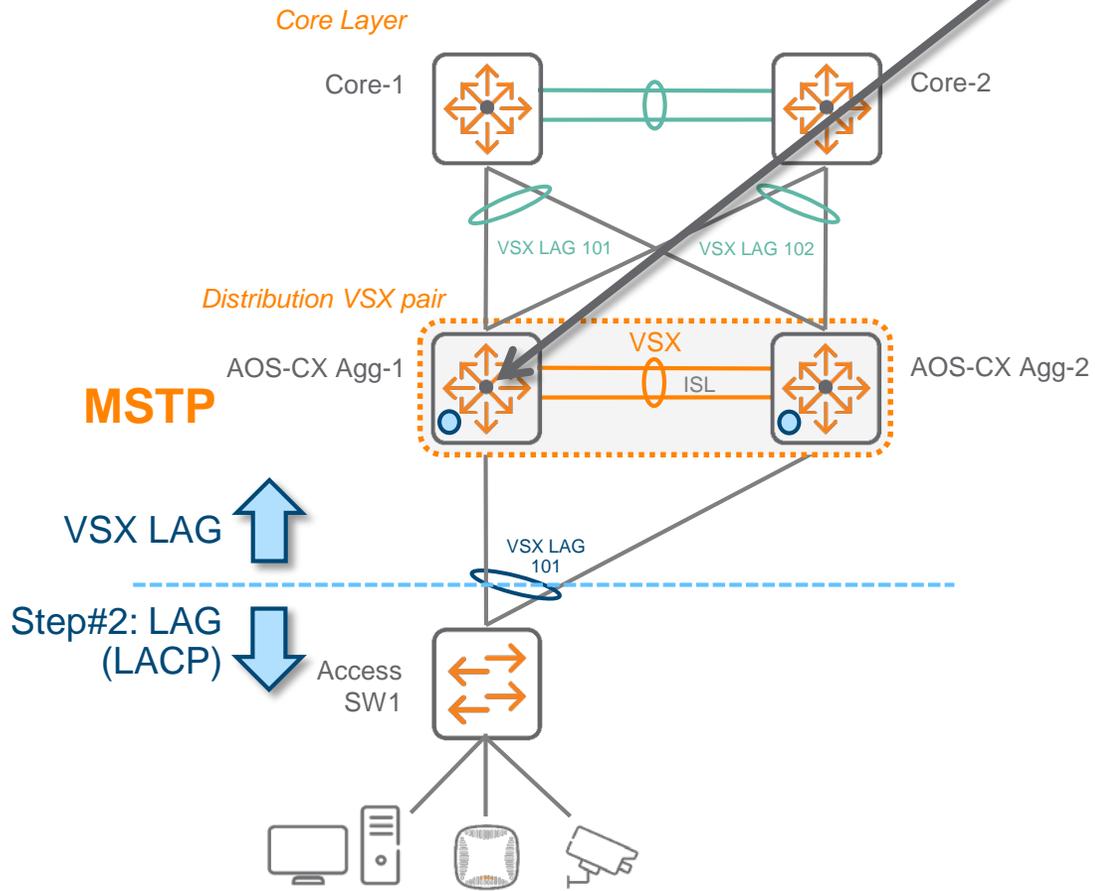
```
MST0
Root ID   Priority   : 32768
MAC-Address: 00:00:00:83:25:01
This bridge is the root
Hello time(in seconds):2  Max Age(in seconds):20
Forward Delay(in seconds):15
```

```
Bridge ID Priority   : 32768
MAC-Address: 00:00:00:83:25:01
Hello time(in seconds):2  Max Age(in seconds):20
Forward Delay(in seconds):15
```

Port	Role	State	Cost	Priority	Type
lag1	Designated	Forwarding	1	64	point_to_point
lag49	Designated	Forwarding	500	64	point_to_point
lag101	Backup	Blocking	20002	64	point_to_point

LACP fallback

Access Switch ZTP - trunk



```
8325-1# sh lacp interfaces multi-chassis
```

State abbreviations :

A - Active P - Passive F - Aggregable I - Individual
 S - Short-timeout L - Long-timeout N - InSync O - OutofSync
 C - Collecting D - Distributing
 X - State m/c expired E - Default neighbor state

Actor details of all interfaces:

Intf	Aggregate name	Port id	Port Priority	State	System-ID	System Priority	Aggr Key
1/1/49	lag49(mc)	49	1	ALFNCD	00:00:00:83:25:01	65534	49
1/1/1	lag101(mc)	1	1	ALFNCD	00:00:00:83:25:01	65534	101

Partner details of all interfaces:

Intf	Aggregate name	Partner Port-id	Port Priority	State	System-ID	System Priority	Aggr Key
1/1/49	lag49(mc)	70	1	ALFNCD	00:00:00:01:01:01	65534	49
1/1/1	lag101(mc)	2	0	ALFNCD	94:f1:28:0c:de:80	56960	532

Remote Actor details of all interfaces:

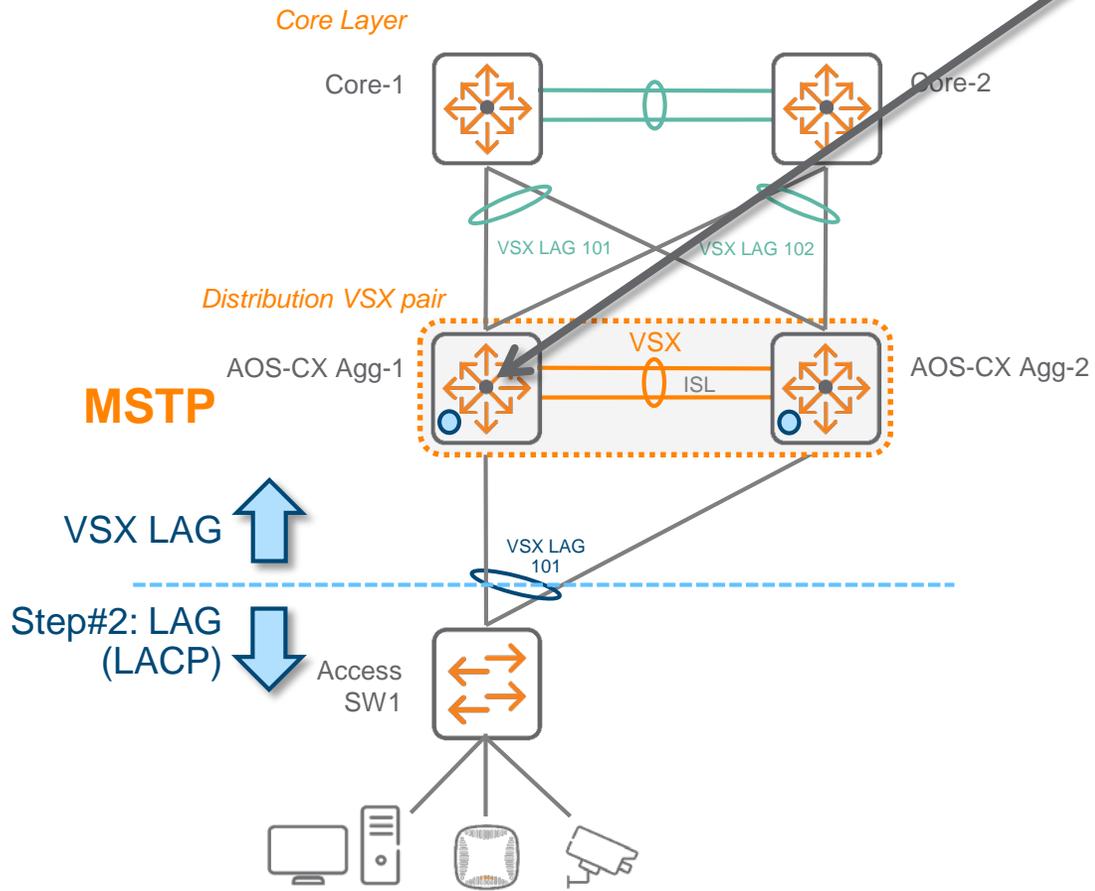
Intf	Aggregate name	Port id	Port Priority	State	System-ID	System Priority	Aggr Key
1/1/49	lag49(mc)	1049	1	ALFNCD	00:00:00:83:25:01	65534	49
1/1/1	lag101(mc)	1001	1	ALFNCD	00:00:00:83:25:01	65534	101

Remote Partner details of all interfaces:

Intf	Aggregate name	Partner Port-id	Port Priority	State	System-ID	System Priority	Aggr Key
1/1/49	lag49(mc)	1070	1	ALFNCD	00:00:00:01:01:01	65534	49
1/1/1	lag101(mc)	1	0	ALFNCD	94:f1:28:0c:de:80	56960	532

LACP fallback

Access Switch ZTP - trunk



```
8325-1# sh spanning-tree
Spanning tree status      : Enabled Protocol: MSTP

MST0
  Root ID   Priority   : 32768
            MAC-Address: 00:00:00:83:25:01
            This bridge is the root
            Hello time(in seconds):2 Max Age(in seconds):20
            Forward Delay(in seconds):15

  Bridge ID Priority   : 32768
            MAC-Address: 00:00:00:83:25:01
            Hello time(in seconds):2 Max Age(in seconds):20
            Forward Delay(in seconds):15

  Port      Role        State      Cost        Priority    Type
  -----
  lag1     Designated Forwarding 1           64         point_to_point
  lag49     Designated Forwarding 500        64         point_to_point
  lag101    Designated Forwarding 20000     64         point_to_point

8325-1# sh spanning-tree vsx
Spanning tree status      : Enabled Protocol: MSTP

MST0
  Root ID   Priority   : 32768
            MAC-Address: 00:00:00:83:25:01
            This bridge is the root
            Hello time(in seconds):2 Max Age(in seconds):20
            Forward Delay(in seconds):15

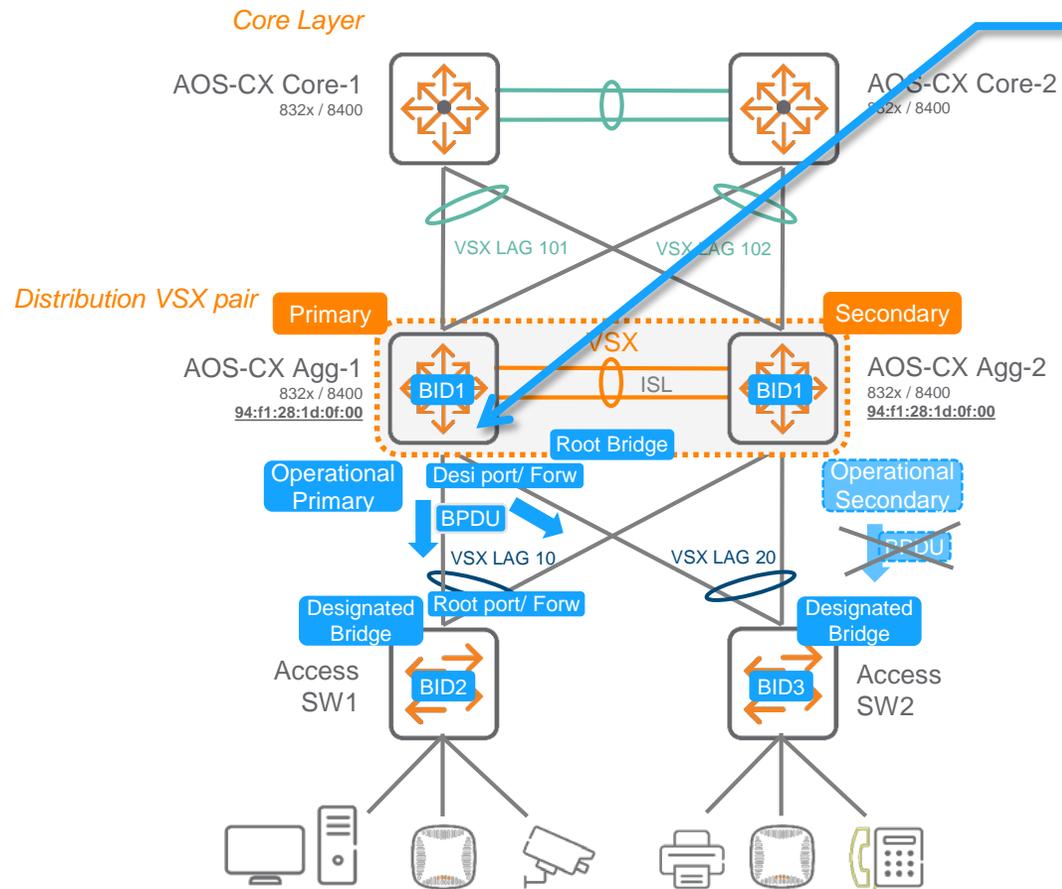
  Bridge ID Priority   : 32768
            MAC-Address: 00:00:00:83:25:01
            Hello time(in seconds):2 Max Age(in seconds):20
            Forward Delay(in seconds):15

  Port      Role        State      Cost        Priority    Type
  -----
  lag1     Designated Forwarding 1           64         point_to_point
  lag49     Designated Forwarding 500        64         point_to_point
  lag101    Designated Forwarding 20000     64         point_to_point
```

VSX and RPVST+

VSX Enhancements

VSX and RPVST+

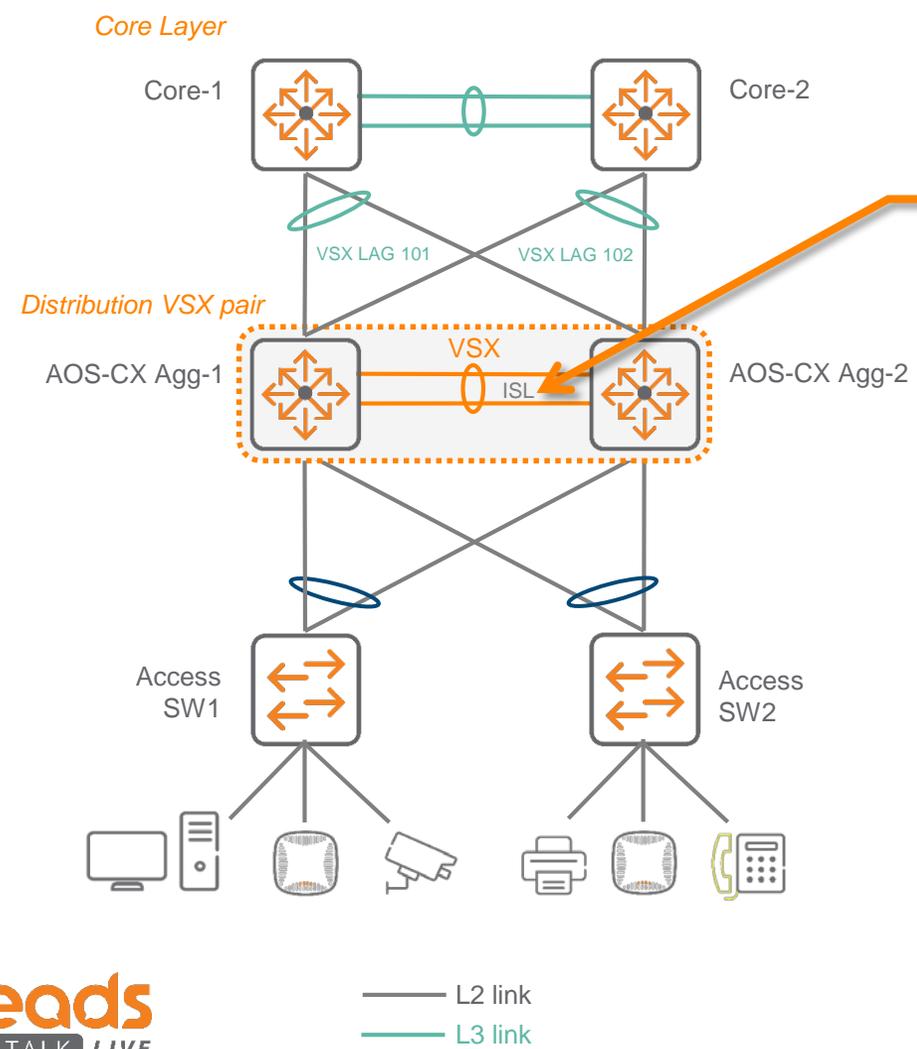


RPVST+ support

1. Thanks to **VSX system-mac**, both VSX switches appear as **a single common spanning-tree Bridge ID** to RPVST+ partner devices upstream and downstream that participate to the same spanning-tree domain.
2. Spanning-tree protocol runs independently on both VSX nodes. (conformance to dual-control plane VSX architecture).
3. The Primary VSX node is responsible to run the protocol for the VSX LAGs. In nominal state, the Primary is "**Operational Primary**" and the secondary is "**Operational Secondary**". In case of Primary VSX node failure, the Secondary VSX node becomes RPVST+ Operational Primary. When Primary VSX node goes back up, it takes back ownership of RPVST+ Operational Primary role.
4. On VSX LAG ports, RPVST+ runs only from the Operational Primary. Operational Secondary never sends STP BPDU on VSX LAG. Only the orphan ports (non VSX LAG) of Operational Secondary run RPVST+.
5. The Operational Secondary hold pre-computed STP information for ready-state switchover thanks to STP states synchronization done by Oper_Pri to Oper_Sec for links member of VSX LAG. That happens as part of the initial sync (LACP, MAC, ARP, STP). During switchover, BPDU to downstream or upstream are sent by new Oper_Prim within the default 6 seconds of spanning-tree BPDU failure detection timer: 3x hello-timer (2s per default).
6. **ISL is always part of STP, non-blocking and sends/receives BPDUs.** (ISL is never blocked, other links will have to block)
7. vsx-sync support STP. STP global configuration on primary VSX node is pushed to secondary node automatically.
8. 64 VLANs are supported.
9. For internal spanning-tree protocol between VSX nodes, Bridge_ID of primary and secondary are derived (-1, +1) from VSX system-mac. Do not use system-mac-1, system-mac and system-mac+1 on any node of the network.

VSX Components

ISL



Inter Switch Link (ISL)

1. Each VSX switch needs to be configured with an ISL link that is **directly connected** to its peer VSX switch.
2. ISL can be a single circuit but aggregated circuits - **LAG is strongly recommended** (up to 8 physical links). Ports must have same speed.
3. Speed could be 10G, 40G or 100G. Prefer 40G or 100G. Example: 2x40G
4. ISL can span **long distances** (transceiver dependent, 40km tested).
5. ISL link is used for data path traffic forwarding, control plane VSX protocol exchange and management plane for peer management.
6. Traffic going over the ISL has **no additional encapsulation**.
7. ISLP is the protocol that runs over ISL and that is used to **synchronize LACP, MAC, ARP, STP, DHCP** and configuration.
8. A **hello packet** is periodically exchanged just to make sure the peer's control plane is alive (range [1..5]s, def 1s). ISL also has a **dead-interval** range of 2..20 and default is 20. If a device does not receive a hello packet from its peer within the dead interval, it treats the peer device as dead and goes for a split detection. ISL port is treated as down when it stays down for the configured **Hold-time** (default=0s) interval.
9. All QoS/ACL policies that can be applied to network ports can be applied to ISL as well.

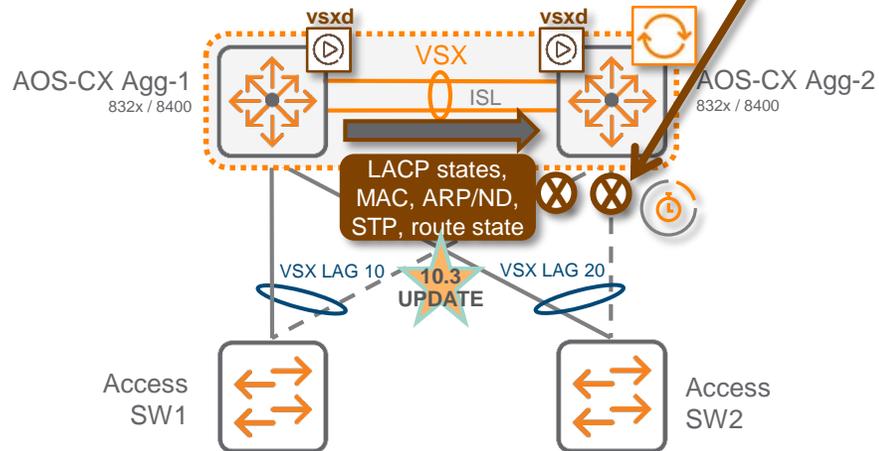
10.3
UPDATE

Linkup Delay

VSX Initial Sync and Linkup Delay

Reminder and new optimization

Linkup Delay



1. When a VSX device is rebooted, it has no entries for MAC, ARP, routes. If downstream VSX LAG ports are activated before all these information are re-learned, traffic is dropped.
2. To avoid traffic drop, VSX LAGs on the rebooted device stay down until restore of LACP, MAC, ARP, STP databases. + routing peering are established and routes are learnt from upstream peer.
3. The learning process has 2 phases:
 - a) Initial Sync Phase:
 - 10.3 UPDATE
 - This is the download phase where the rebooted node learns all the **LACP+MAC+ARP+STP+routing state DB entries** from its VSX peer through ISLP.
 - This Initial Sync timer is dynamic. It is the required time to download DB information from the peer.
 - b) Linkup Delay Phase:
 - This is the duration for:
 - installing the downloaded entries to the ASIC.
 - establishing router adjacencies with core nodes and learning upstream routes.
 - Linkup Delay timer default value is 180s.
4. When both VSX devices reboot, linkup-delay-timer is not used.
5. In order to get upstream router adjacencies established during Linkup Delay, the upstream LAG (ex: LAG 101) has to be excluded from the scope of the Linkup Delay. Until linkup delay timer, all SVIs that VSX LAGs are a member of are kept in a pseudo-shut state.

Linkup Optimization

Initial Sync and Linkup Delay

- **Until 10.2:**

The Initial Sync timer is dynamic, for LACP/MAC/ARP/STP bulk sync completion at DB level.
The linkup-delay timer is static and user configurable (default 180s).

- **In 10.3:**

The Initial Sync timer is fully dynamic, using a bailout-timer value for LACP/MAC/ARP/STP and route_state sync completion at DB level and having auto notification to finish this bailout-timer.
The linkup-delay timer is static and user configurable (default 180s).

Benefits:

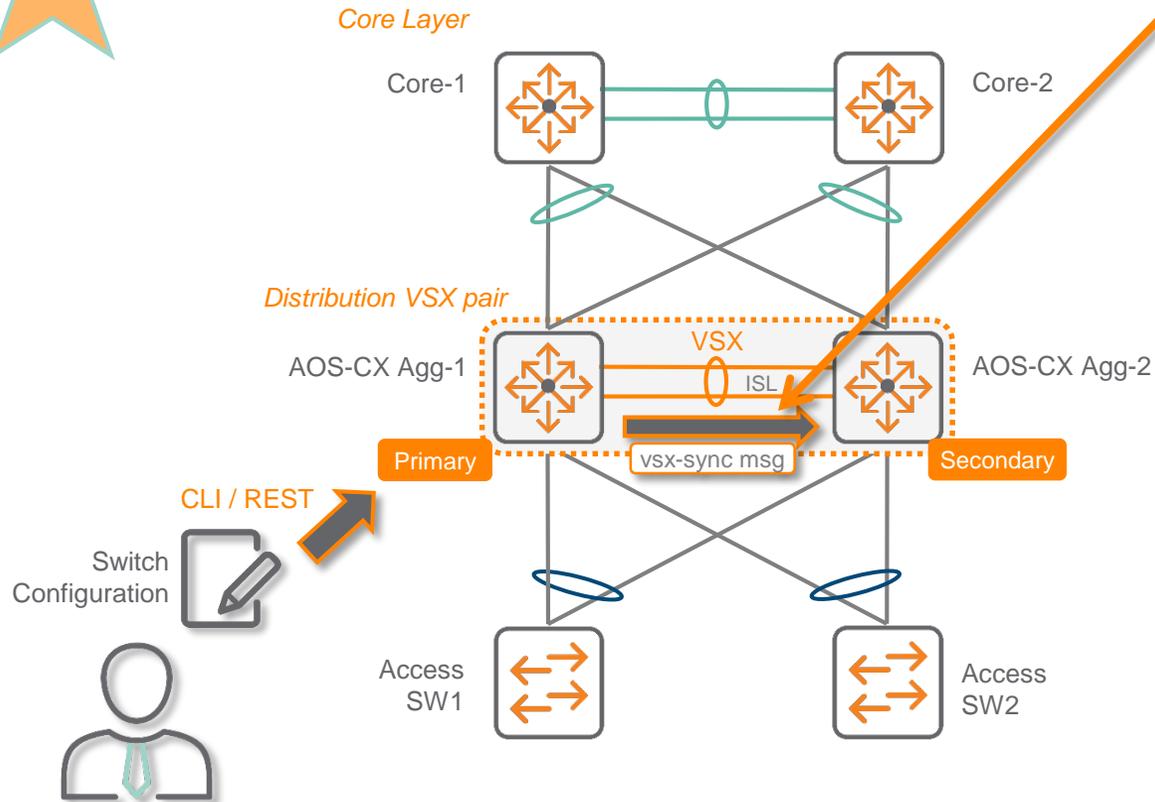
- For complete network readiness, upstream peering and route-learning is completed before VSX LAG is turn on

VSX Sync

VSX Configuration Synchronization

vsx-sync

10.3
UPDATE



vsx-sync

1. Most of the VSX node configuration is the same across the switch pair (exceptions: IP addresses, Router IDs, hostname and a few others).
2. With VSX enabled, VSX configuration synchronization process is enabled by default. (However, nothing is synchronized yet.) Synchronization can be disabled globally.
3. vsx-sync is the CLI attribute that specifies what is to be synced.
4. In nominal condition, vsx-sync can be set only on the Primary switch. vsx-sync is allowed on Secondary if config-sync is disabled or if ISL link is down.
5. With VSX config-sync feature enabled, the configuration item, set with the vsx-sync attribute on the Primary switch, is automatically pushed to the Secondary switch.
6. If the Secondary switch is not available at the time of the Primary device configuration, the configuration will be pushed to the Secondary after it comes up.
7. If a certain configuration synchronization operation failed, it will be indicated in the CLI status commands and the user needs to manually fix the inconsistency.
8. There is no fail-safe mechanism that disables VSX if configuration synchronization fails. (not expected)
9. The following individual features support synchronization (vsx-sync):
 - VLANs, ACLs, object-group
 - Class, QoS, Policies, rate-limits
 - active-gateway, PBR
10. The following group features support synchronization (vsx-sync tag at vsx-sync command level):
aaa, bfd-global, copp_policy, dcb-global, dhcp-relay, dhcp-server, dns, icmp-tcp, lldp, loop-protect-global, mclag_interfaces, qos_global, sflow, snmp, ssh, static_routes, stp-global, time, udp-forwarder, vsx-global

10.3
UPDATE

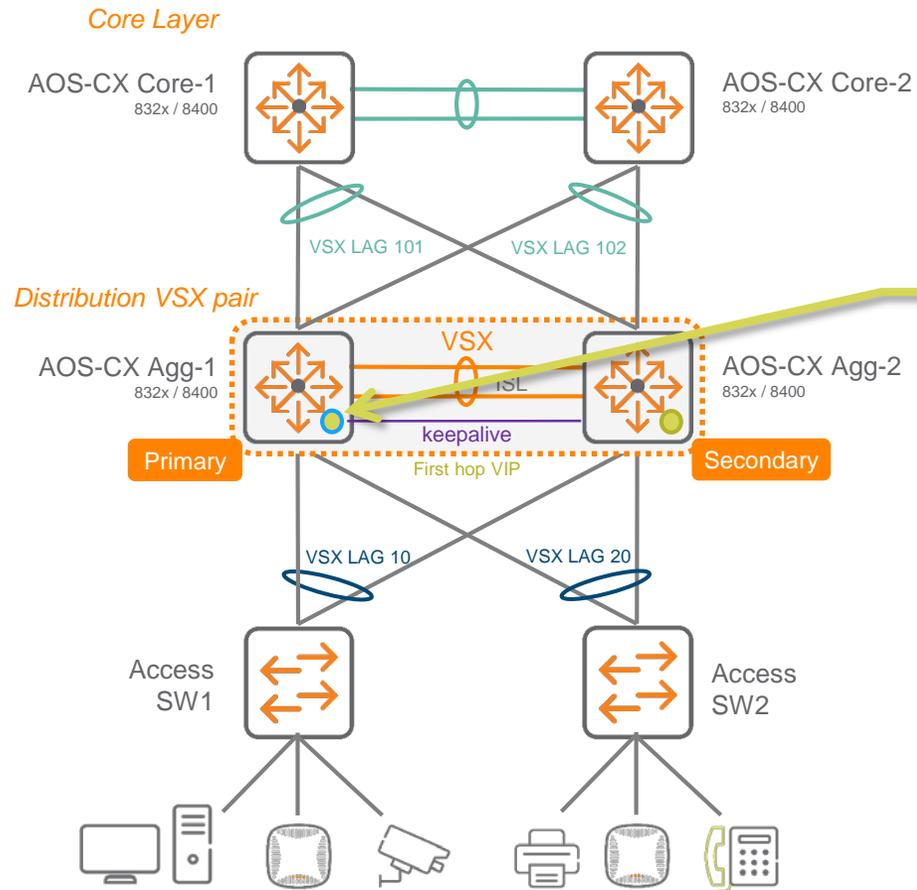
VSX Config Sync

New Feature Groups and associated vsx-sync tags

Feature	Details	Synchronization Scheme:	Application Scheme: vsx-sync tag command
BFD	<ul style="list-style-type: none"> BFD global configuration 	vsx-sync level tag support	<ul style="list-style-type: none"> VSX Context Command: vsx-sync bfd-global
DCB	<ul style="list-style-type: none"> global configurations for DCB features (DCBx, PFC and ETS): lldp dcbx, dcbx application 	vsx-sync level tag support	<ul style="list-style-type: none"> VSX Context Command: vsx-sync dcb-global
DHCP relay	<ul style="list-style-type: none"> All instances 	vsx-sync level tag support	<ul style="list-style-type: none"> VSX Context Command: vsx-sync dhcp-relay
DHCP server	<ul style="list-style-type: none"> All instances 	vsx-sync level tag support	<ul style="list-style-type: none"> VSX Context Command: vsx-sync dhcp-server
ICMP / TCP	<ul style="list-style-type: none"> ICMP and TCP global parameters: ip icmp redirect, throttle, unreachable, ip tcp synack_retries 	vsx-sync level tag support	<ul style="list-style-type: none"> VSX Context Command: vsx-sync icmp-tcp
LLDP	<ul style="list-style-type: none"> All global LLDP parameters (dcbx, timers, TLVs) 	vsx-sync level tag support	<ul style="list-style-type: none"> VSX Context Command: vsx-sync lldp
Loop-protect	<ul style="list-style-type: none"> All global loop-protect parameters (re-enable-timers, transmit-interval) 	vsx-sync level tag support	<ul style="list-style-type: none"> VSX Context Command: vsx-sync loop-protect-global
UDP Forwarder	<ul style="list-style-type: none"> For WOL, enable/disable ip udp-bcast-forward 	vsx-sync level tag support	<ul style="list-style-type: none"> VSX Context Command: vsx-sync udp-forwarder
VSX Setting	<ul style="list-style-type: none"> inter-switch-link hello-interval, dead-interval, hold-time, peer-detect-interval keepalive udp-port, hello-interval, keepalive dead-interval system-mac split-recovery linkup-delay-timer 	vsx-sync level tag support	<ul style="list-style-type: none"> VSX Context Command: vsx-sync vsx-global

VSX and DHCP server

Active-Gateway DHCP Server



Active-standby DHCP Server

1. Only primary VSX node replies to DHCP request with DHCP offer.
2. Secondary VSX node forwards over ISL to Primary VSX switch the DHCP requests received from downstream endpoints.
3. The DHCP lease information are synchronized from Primary to Secondary on to its SystemStateDB.
4. Secondary VSX switch will take over DHCP Server service upon Primary failure detection (ISL down + keepalive down) and continues from where the Primary left it (as it has a copy of the leases allocated by the Primary). When Primary comes back up, it re-learns the additional updates that happened during the time the Secondary was active, takes over from the Secondary and seamlessly continues on from there.
5. Downstream clients receive a single DHCP offer.

VSX Live Upgrade

Enhancements

VSX Live Upgrade

Update-software Enhancements in 10.3

- **Graceful shutdown of protocols during VSX upgrade**

- OSPF, BGP, VRRP protocols on the Secondary and Primary will gracefully shut down and notify their partners to use the peer device for forwarding and then go for a reboot.
- This method will help reducing downtime by avoiding in-flight traffic loss when the device is rebooting and links drop.
- This enhancement mechanism starts between image download and reboot.
- Partial support starting in 10.03.0001
- **Benefit:** drain traffic out of the uplink of the “to-be-rebooted” switch and achieve no packet drop on upstream ROP.

- **LACP Traffic Draining**

- 10.3CPE or future

- **Links UP optimization (specific to update-software)**

- After switch is rebooted, ports which are members of VSX LAGs, are turned ON through an optimized mechanism to maximize uptime during the upgrade.

Protocols Graceful shutdown

Mechanism Details

1. Once the image download is completed on secondary peer, graceful shutdown request of control plane protocols is triggered in the SSDB and a 10 minute timer is started.
2. The protocols which subscribed to this SSDB trigger, start a graceful-shutdown.
3. These protocols update the SSDB once the graceful-shutdown is completed.
4. VSX update-software daemon reboots secondary after gshut completion status is updated in SSDB or after the 10 min timer expires.
5. The same graceful shutdown process will happen on primary before rebooting.
6. If aborted during control plane shutdown, the switch will still go for reboot.

Note: for any reason if switch could not be rebooted, the protocols will be in shutdown state and manual intervention is needed to bring the VSX cluster to normal state.

Ethernet Ring Protection Switching ERPS

ERPS - Ethernet Ring Protection Switching

Standard

ITU-T G.8032/Y1344 – Ethernet Ring Protection Switching

Objectives and principles:

- Use of standard 802 MAC and OAM frames around the ring.
Uses standard 802.1Q (and amended Q bridges), but with xSTP disabled.
- Ring nodes supports standard FDB MAC learning, forwarding, flush behavior and port blocking/unblocking mechanisms.
- Prevents loops within the ring by blocking one of the links (either a pre-determined link or a failed link).
Monitoring of the ETH layer for discovery and identification of Signal Failure (SF) conditions.
Protection and recovery switching within 50 ms for typical rings.
- Total communication for the protection mechanism should consume a very small percentage of total available bandwidth.

Terms and concepts

- **Ring Protection Link (RPL)**: Link designated by mechanism that is blocked during Idle state to prevent loop on Bridged ring
- **RPL Owner**: Node connected to RPL that blocks traffic on RPL during Idle state and unblocks during Protected state
- **Link Monitoring**: Links of ring are monitored using standard ETH CC OAM messages (CFM)
- **Signal Fail (SF)** – Signal Fail is declared when ETH trail signal fail condition is detected
- **No Request (NR)** – No Request is declared when there are no outstanding conditions (e.g., SF, etc.) on the node
- **Ring APS (R-APS) Messages** – Protocol messages defined in Y.1731 and G.8032
- **Automatic Protection Switching (APS) Channel** - Ring-wide VLAN used exclusively for transmission of OAM messages including R-APS messages (*Control-VLAN* in AOS-CX)

Ring Timers

Hold-off timer – (Default – 0 s)

When a new defect or more severe defect occurs (new SF), this event is not to be reported immediately to protection switching if the provisioned hold-off timer value is non-zero. Instead, the hold-off timer is started

WTR timer – (Default – 5 min)

When recovering from an SF, the delay timer must be long enough to allow the recovering network to become stable.

In the revertive mode of operation, the WTR timer is used to prevent frequent operation of the protection switching due to intermittent SF defects.

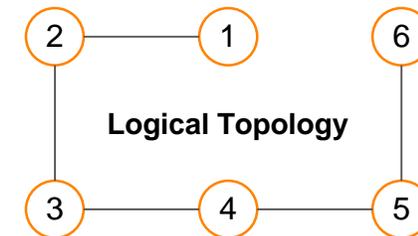
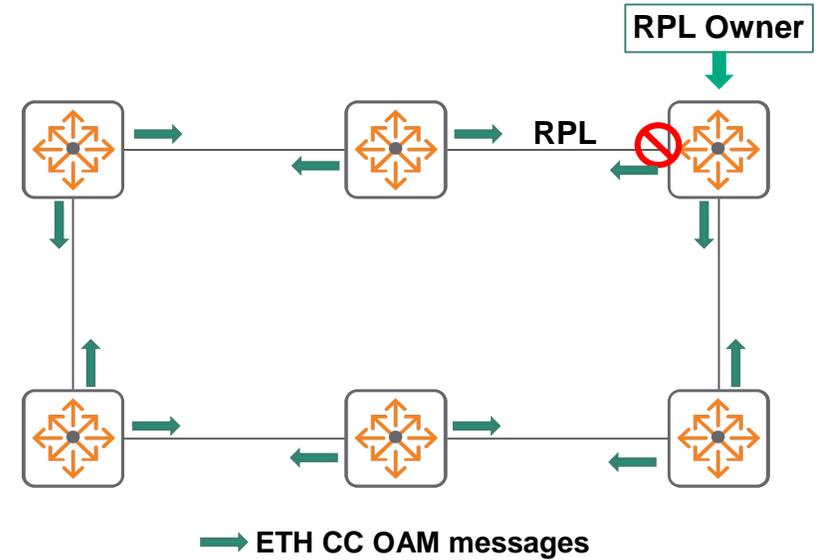
Guard timer – (Default 500 ms)

The guard timer is activated whenever an Ethernet ring node receives an indication that a local switching request has cleared (i.e., local clear SF, Clear).

This timer period should be greater than the maximum expected forwarding delay in which an R-APS message traverses the entire ring.

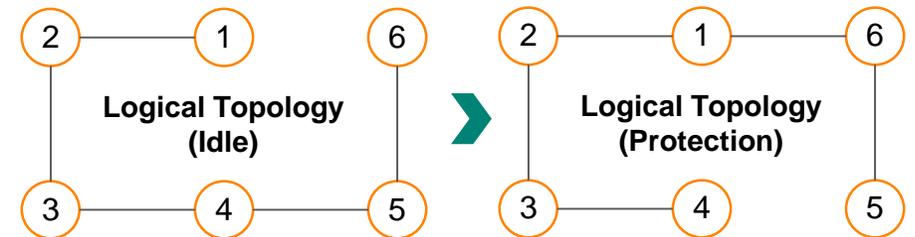
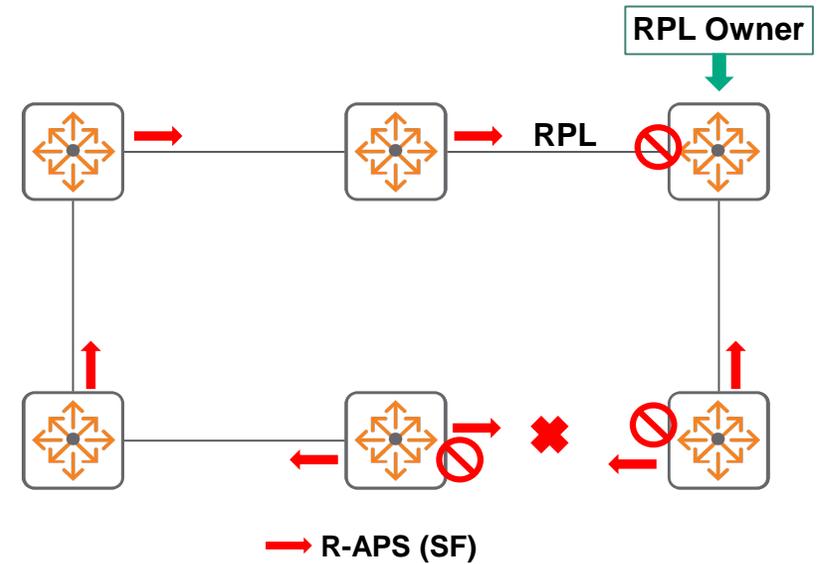
Ring Idle State

1. Physical topology has all nodes connected in a ring
2. ERP guarantees lack of loop by blocking the RPL (link between 6 & 1 in figure)
3. Logical topology has all nodes connected without a loop.
4. Each link is monitored by its two adjacent nodes using ETH CC OAM messages
5. Signal Failure as defined in Y.1731, is trigger to ring protection
 - Loss of Continuity
 - Server layer failure (e.g. Phy Link Down)



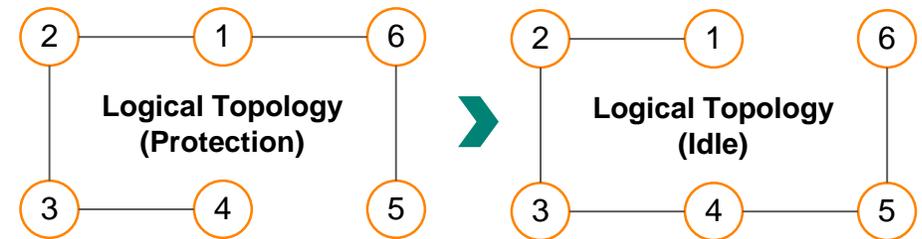
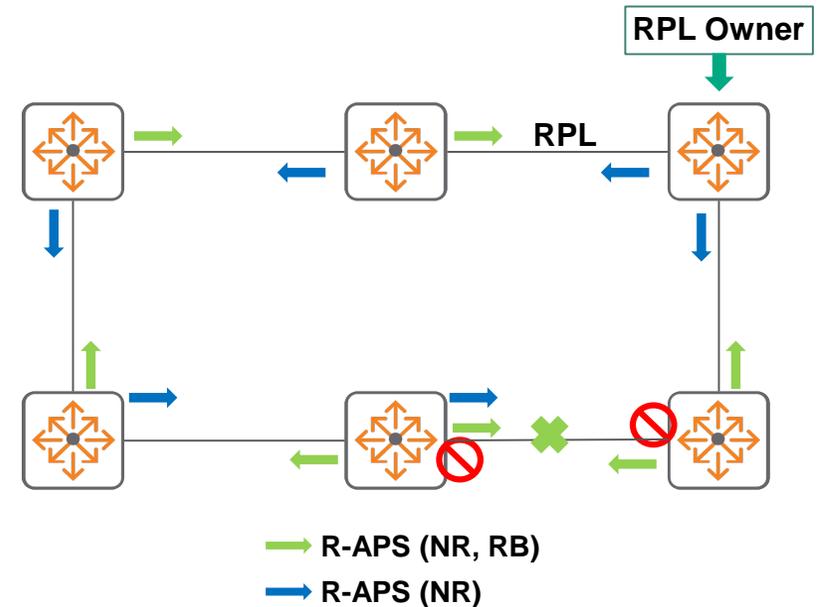
Link or Node Failure

1. Link/node failure is detected by the nodes adjacent to the failure.
2. The nodes adjacent to the failure, block the failed link and report this failure to the ring using R-APS (SF) message
3. R-APS (SF) message triggers
 - RPL Owner unblocks the RPL
 - All nodes perform FDB flushing
4. Ring is in protection state
5. All nodes remain connected in the logical topology



Failure Recovery

1. When the failed link recovers, the traffic is kept blocked on the nodes adjacent to the recovered link
2. The nodes adjacent to the recovered link transmit RAPS(NR) message indicating they have no local request present
3. When the RPL Owner receives RAPS(NR) message it Starts WTR timer
4. D. Once WTR timer expires, RPL Owner blocks RPL and transmits R-APS (NR, RB) message
5. Nodes receiving the message – perform a FDB Flush and unblock their previously blocked ports
6. Ring is now returned to Idle state



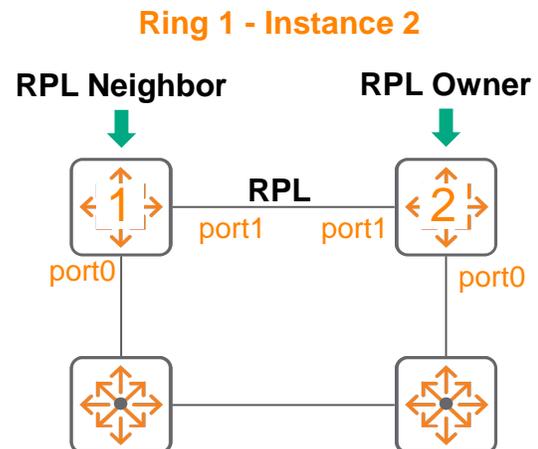
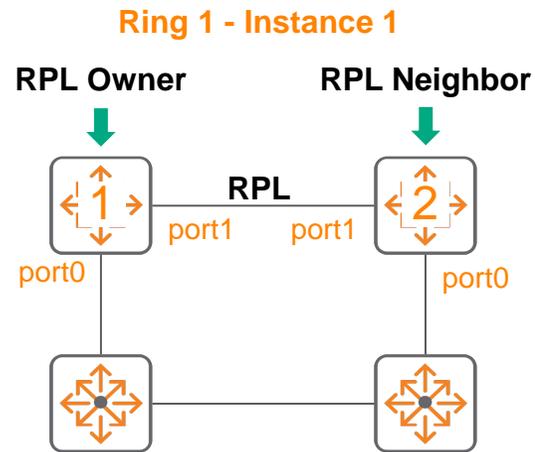
ERPS in AOS-CX - Basic configuration

Switch 1

```
erps ring 1
  port0 interface 1/1/15
  port1 interface 1/1/16
  instance 1
    enable
    control-vlan 16
    protected-vlans 15
    role rpl-owner
    rpl port1
  exit
  instance 2
    enable
    control-vlan 17
    protected-vlans 14
    role rpl-neighbor
    rpl port0
  exit
```

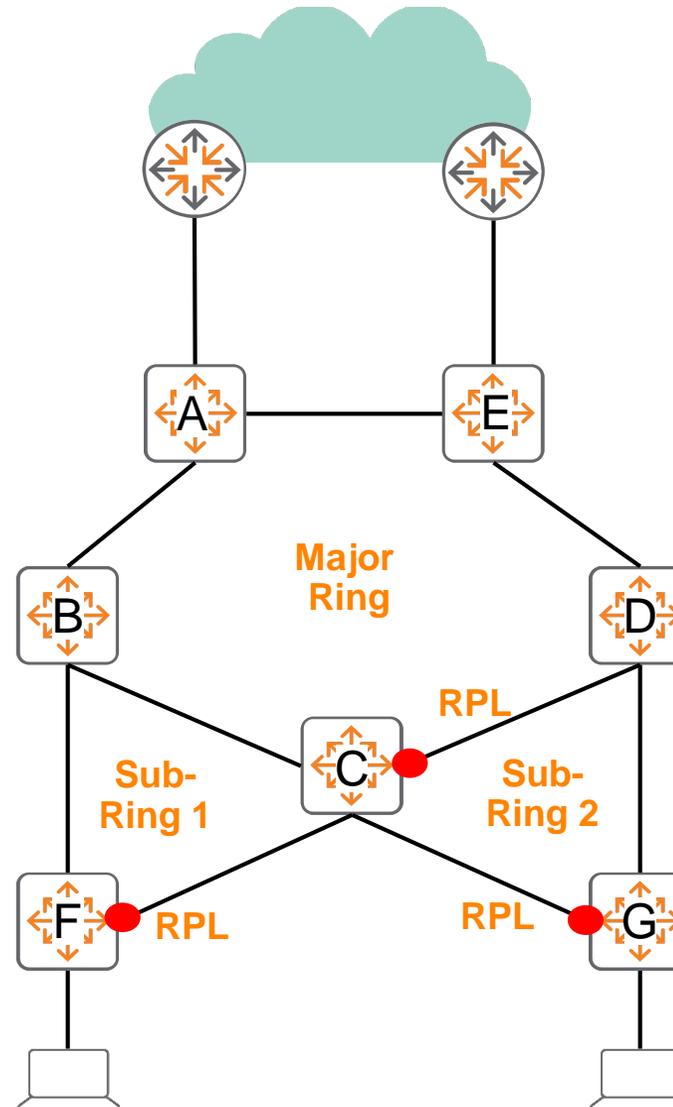
Switch 2

```
erps ring 1
  port0 interface 1/1/16
  port1 interface 1/1/15
  instance 1
    enable
    control-vlan 16
    protected-vlans 15
    role rpl-neighbor
    rpl port1
  exit
  instance 2
    enable
    control-vlan 17
    protected-vlans 14
    role rpl-owner
    rpl port0
  exit
```



Major Rings and Sub-rings

- One or more links of this ring will be part of a Major ring and will be managed by the Major ring itself.
- The nodes where the sub-ring gets connected to the Major ring is referred to as the “interconnection nodes”.
- One of the differences between a sub-ring and a major ring is that the raps-channel control VLAN is not blocked anywhere along the path of the sub-ring, even though the protected data VLANs may be blocked.



1G Interfaces & Port Groups (8325)

Speed

"SPEED" command configures the link speed, duplex, and auto-negotiation settings for an interface

These are the supported combinations of Speed

10-full	10 Mbps, full duplex, no auto-negotiation
10-half	10 Mbps, half duplex, no auto-negotiation
100-full	100 Mbps, full duplex, no auto-negotiation
100-half	100 Mbps, half duplex, no auto-negotiation
1000-full	1000 Mbps, full duplex, no auto-negotiation
auto	Auto-negotiate speed and duplex
10m	Allow interface to link at 10Mbps
100m	Allow interface to link at 100Mbps
1g	Allow interface to link at 1Gbps
2.5g	Allow interface to link at 2.5Gbps
5g	Allow interface to link at 5Gbps
10g	Allow interface to link at 10Gbps
25g	Allow interface to link at 25Gbps
40g	Allow interface to link at 40Gbps
50g	Allow interface to link at 50Gbps
100g	Allow interface to link at 100Gbps

no speed defaults Set interface speed, duplex, and auto-negotiation to defaults

Operate at a fixed speed of 1000 Mbps with full duplex and no auto-negotiation.

```
switch(config)# interface 1/1/1
switch(config-if)# speed 1000-full
```

Configure an interface to advertise only 1Gbps and 10Gbps speeds.

```
switch(config)# interface 1/1/1
switch(config-if)# speed auto 1g 10g
```

8325 - Interface group

Group	Ports	Default speed
1	1 – 12	25Gpbs
2	13 – 24	25Gpbs
3	25 – 36	25Gpbs
4	37 – 48	25Gpbs



```
interface-group <group> speed <10g|25g >  
    10g          Allow 1Gbps and 10Gbps transceivers only  
    25g          Allow 25Gbps transceivers only (default)
```

```
Switch(config)# system interface-group 1 speed 10g  
Changing the group speed will disable all member interfaces that  
do not match the new speed.
```

8325 - 1G Interface restrictions

The 8325-48Y8C switch (JL635A) does not support auto-negotiation with following 1G optical transceivers

model	Type	Support
J4858	SFP-SX	YES
J4859	SFP-LX	YES
J4860	SFP-LH	YES
J9142	SFP-1G-BXD	NO
J9143	SFP-1G-BXU	NO

* These ports must be configured for speed 1000-full to operate correctly

Maximum of 32 1GBase-T Transceiver (J8177D) in 8325 Models JL624A/JL625A; (not applicable to 8325 models JL626A/JL627A)

- The 1GigT will only available in the top 2 rows (i.e. 3rd row not allowed)



Valid ports for J8177D 1GBase-T Transceiver

Front

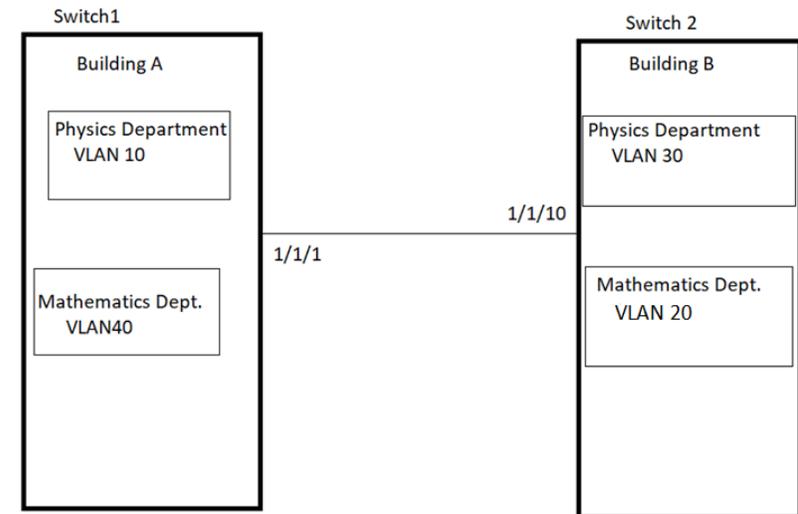
VLAN translation

VLAN Translation

- VLAN translation allows you to configure bidirectional VLAN identifier translation
- This feature helps reduce effort of the network administrator, reconfiguring the legacy VLANs when a merger/Site integration happens
- VLAN translation applies only for tagged packets. The VLAN tag in packet is 'REPLACED'
- Ingress traffic is translated from **vlan1-id** to **vlan2-id**, and the egress traffic will be translated from **vlan2-id** to **vlan1-id**
`SW(config-if)# vlan translate < vlan1-id > < vlan2-id >`

NOTE:

- This command is applicable only for Layer 2 Trunk ports.
- Interface must be Layer 2 physical or LAGs/MCLAGs/VSX interface.
- Routing must be disabled on the interface.
- VLAN1-id or VLAN2-id is not recommended to be native VLAN for the port being configured.



VLAN Translation

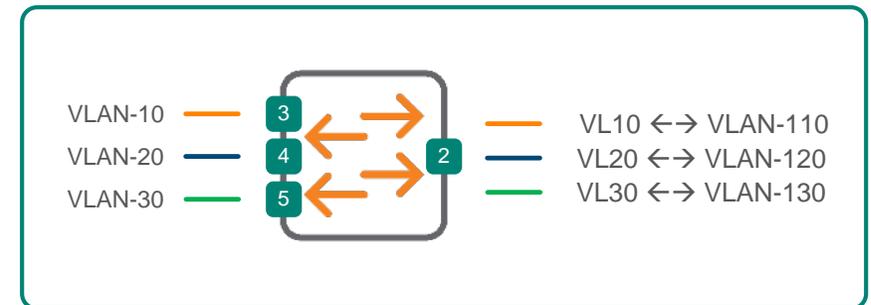
Configuration Example

```
SW(config)# interface 1/1/2
SW(config-if)# no shutdown
SW(config-if)# no routing
SW(config-if)# vlan trunk allowed 10,20,30,110,120,130
SW(config-if)# vlan translate 110 10
SW(config-if)# vlan translate 120 20
SW(config-if)# vlan translate 130 30
```

Verification

```
SW1#show vlan translation
```

```
-----
Interface  VLAN-1      VLAN-2
-----
1/1/2      110             10
1/1/2      120             20
1/1/2      130             30
Total number of translation rules : 3
```



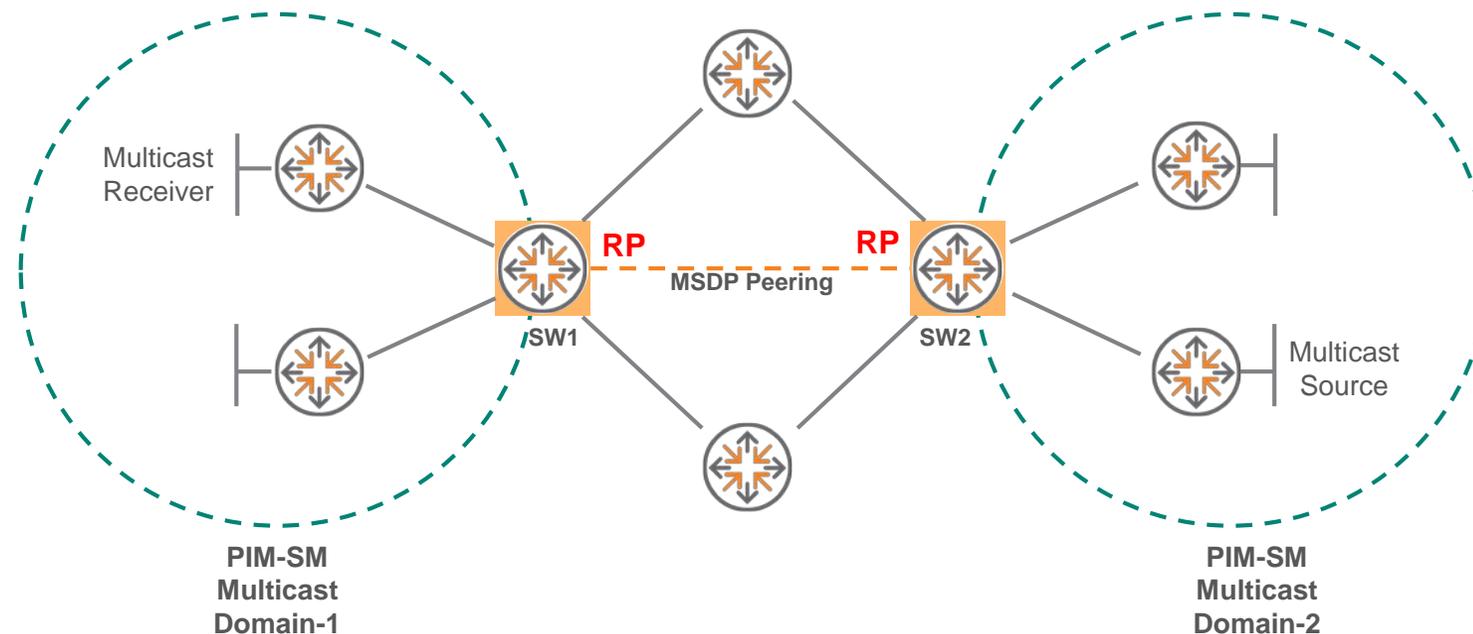
VLAN Translation Details

- 8320 and 8325 can support up to **10k VLAN Translation rules**
- 8400 can support up to **1k unique VLAN Translations rules.**
- This cannot co-exist with RPVST, MVRP.
- This can co-exist with MSTP, VSX, LACP, LLDP, Loop protect, IGMP, MLD.
- This will not effect the VLAN PCP and DEI bits in the packet.
- Native VLAN is not recommended to be part of VLAN translation at the port level.

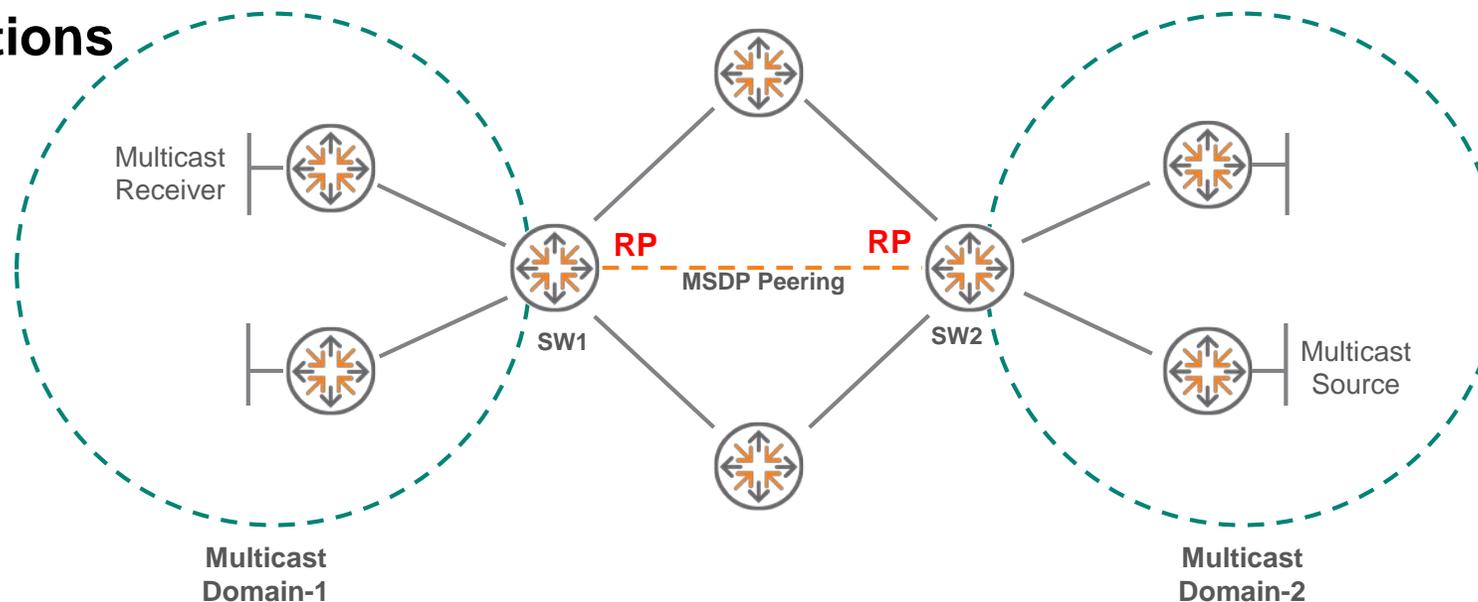
Multicast Source Discovery Protocol (MSDP)

Multicast Source Discovery Protocol (MSDP)

- Connecting multiple multicast (PIM-SM) routing domains to exchange Multicast Source information
- Rendezvous Point (RP) runs MSDP over TCP to dynamically discover multicast sources in other domains
- Each MSDP router peer with other MSDP router to exchange information about active multicast sources within their domains
- MSDP reduces the complexity of interconnecting multiple PIM-SM domains, to share inter-domain source tree information, instead having common shared tree information.



MSDP Configurations



SW1#

```
SW1# configure terminal
SW1(config)# router msdp
SW1(config-msdp)# enable
SW1(config-msdp)# sa-interval 60
SW1(config-msdp)# ip msdp peer 22.22.22.22
SW1(config-msdp-peer)# description "SW2 Peer"
SW1(config-msdp-peer)# enable
SW1(config-msdp-peer)# connect-source loopback1
SW1(config-msdp-peer)# keepalive 60 75
SW1(config-msdp-peer)# connection-retry-interval 30
SW1(config-msdp-peer)# password ARUBA!23
```

SW2#

```
SW2# configure terminal
SW2(config)# router msdp
SW2(config-msdp)# enable
SW2(config-msdp)# sa-interval 60
SW2(config-msdp)# ip msdp peer 11.11.11.11
SW2(config-msdp-peer)# description "SW1 Peer"
SW2(config-msdp-peer)# enable
SW2(config-msdp-peer)# connect-source loopback1
SW2(config-msdp-peer)# keepalive 60 75
SW2(config-msdp-peer)# connection-retry-interval 30
SW2(config-msdp-peer)# password ARUBA!23
```

MSDP Verification

```
SW1# show ip msdp peer 22.22.22.22
```

```
VRF: default
```

```
MSDP Peer: 22.22.22.22
```

```
Connection status
```

```
State: up Resets: 0 Connection Source: loopback 1
```

```
Uptime(Downtime): 0m 25s SA Messages sent: 0
```

```
SA's learned from this peer: 3
```

```
SW1# show ip msdp sa-cache
```

```
VRF: default
```

```
(30.0.0.1, 230.1.1.1) RP: 11.11.11.11 Peer: 22.22.22.22
```

```
(20.0.0.1, 229.1.1.1) RP: 11.11.11.11 Peer: 22.22.22.22
```

```
(10.0.0.1, 229.1.1.1) RP: 11.11.11.11 Peer: 22.22.22.22
```

```
Total entries: 3
```

```
SW1# show ip msdp sa-cache 229.1.1.1
```

```
(20.0.0.1, 229.1.1.1) RP: 11.11.11.11 Peer: 22.22.22.22
```

```
(10.0.0.1, 229.1.1.1) RP: 11.11.11.11 Peer: 22.22.22.22
```

```
Total entries: 2
```

```
SW1#sh ip msdp summary
```

```
VRF: default
```

```
MSDP Peer Status Summary
```

```
Peer address State Uptime(Downtime)Reset Count SA Cou
```

```
22.22.22.22 up 34m 34s 0 50
```

```
1.1.1.1 down 50m 24s 0 0
```

```
SW1# show ip msdp count
```

```
VRF: default
```

```
SA state per Peer counters
```

```
<Peer>:<#SA learned>
```

```
22.22.22.22: 3
```

```
1.1.1.1: 0
```

MSDP Full configuration

```
SW1# show run
router ospf 1
  redistribute connected
  area 0.0.0.0
interface loopback 1
  ip address 11.11.11.11/32
  ip ospf 1 area 0.0.0.0
  ip pim-sparse enable
interface 1/1/1
  no shutdown
  ip address 10.0.0.1/24
  ip ospf 1 area 0.0.0.0
  ip pim-sparse enable
interface 1/1/2
  no shutdown
  ip address 20.0.0.1/24
  ip ospf 1 area 0.0.0.0
  ip pim-sparse enable
router pim
  enable
  rp-address 11.11.11.11

router msdp
  enable
  ip msdp peer 22.22.22.22
    connect-source loopback 1
  enable
  connection-retry-interval 30
```

```
SW2# show run
router ospf 1
  redistribute connected
  area 0.0.0.0
interface loopback 1
  ip address 22.22.22.22/32
  ip ospf 1 area 0.0.0.0
  ip pim-sparse enable
interface 1/1/1
  no shutdown
  ip address 10.2.0.1/24
  ip ospf 1 area 0.0.0.0
  ip pim-sparse enable
interface 1/1/2
  no shutdown
  ip address 20.2.0.1/24
  ip ospf 1 area 0.0.0.0
  ip pim-sparse enable
router pim
  enable
  rp-address 22.22.22.22

router msdp
  enable
  ip msdp peer 11.11.11.11
    connect-source loopback 1
  enable
  connection-retry-interval 30
```

Scale Considerations

- Maximum MSDP Peers – No limits as of now. Don't expect to be a large number
- Maximum Group for a Source Prefix - Configurable (Default No limits). This can be set per VRF
- Maximum SA limit for a Peer – Configurable(Default no Limits). This can be set per peer for a (S,G)

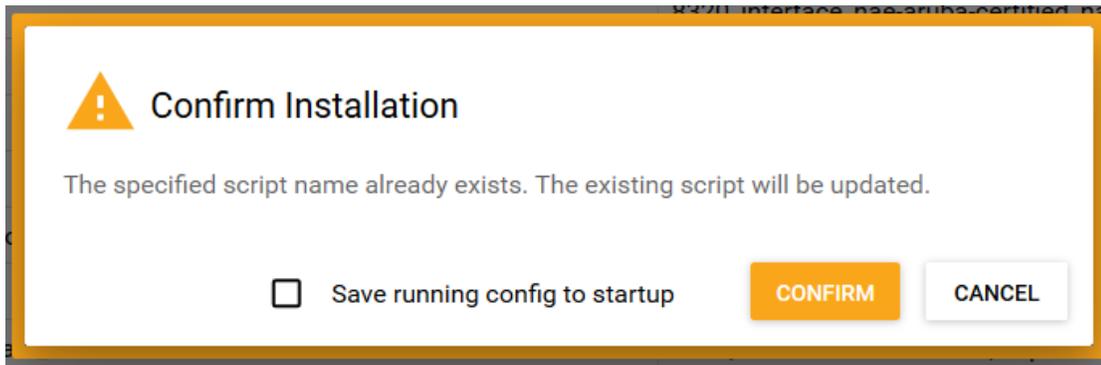
Category	Feature	Scale Numbers
Multicast	IGMP Groups	4 094
Multicast	IPv4 Multicast Routes	4 094

NAE Script Update Feature

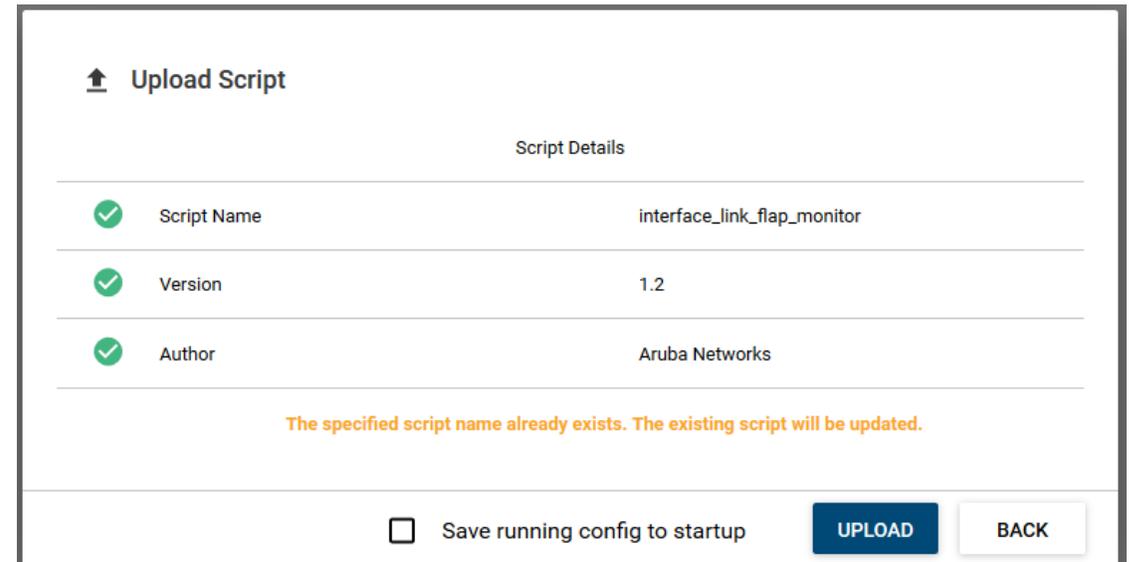
NAE 10.3 Feature – Updating Scripts

Allows for replacing scripts without losing parameter and time series data

Installing Script from ASE



Uploading Script from PC



Updating Scripts with Parameters:

- Customizations on removed parameters will be lost
- New parameters will have default values from script

NAE UI Change

– Versions are stripped from the name of the script and is used to check against

Status	System Created	Name	Version	# Agents	Author
		fault_finder_monitor 1.0	1.0	1	Aruba Networks
		interface_link_flap_monitor 1.2	1.2	0	Aruba Networks



Status	System Created	Name	Version	# Agents	Author
		interface_link_flap_monitor	1.1	1	Aruba Networks
		system_resource_monitor	1.1	1	Aruba Networks

NAE Script REST API Upgrade

```
PUT /system/nae_scripts/{id} ← {id} = script name
{
  "script": "<base64-encoded script>"
}
```

NAE_Script

Show/Hide | List Operations | Expand Operations

GET	/system/nae_scripts	Get a list of resources
POST	/system/nae_scripts	Create NAE Script
DELETE	/system/nae_scripts/{id}	Delete NAE Script
GET	/system/nae_scripts/{id}	Get a set of attributes
PUT	/system/nae_scripts/{id}	Update NAE Script

New Scripts on ASE Since 10.2

- VSX Health Monitor
- IP-SLA Reachability
- Neighbors Decrease Rate Monitor
- Routes Decrease Rate Monitor
- MAC Decrease Rate Monitor

Miscellaneous

- Scale
- IPv6 Link-Local
- CLI, range
- Loop-protect
- SNMP
- OSPF: default originate
- AAA channels
- PKI
- PIM-SM RP ACL
- Global Port Mirror (8400 only)

Aruba 84/83xx Capabilities (10.3)

Competitive Scale – Sell with Confidence

	8400	8325 L3-agg profile	8320 L3-agg profile	8320 L3-core profile
Category	No Shared HW Tables	ARP/ND Table Shared IPv4/IPv6 Route Table Shared	ARP/ND Table Shared IPv4/IPv6 Route Table Shared	ARP/ND Table Shared
IPv4 Routes (system)	1,011,712	28,658	12,288	130,993
IPv6 Routes (system)	524,288	12,289	7,168	32,768
VRFs	64	64	64	64
OSPF Areas (v2/v3)	256	256	256	256
OSPF Interfaces (v2/v3)	256	256	256	256
OSPF Neighbors (v2/v3)	256	256	256	256
BFD Session	256	256	256	256
BGP Neighbors	256	256	256	256
ARP	756,000 ¹	120,000 ³	120,000 ³	14,000
MAC	768,000 ¹	98,304 ³	98,304 ³	32,768
ND	524,000 ²	52,000	52,000	7,000

1. Number of IPv4 Devices/Clients limited to 81,000 for ARP Resolution – 8400 Egress MAC Table Limit
2. Number of IPv6 Devices/Clients limited to 55,000 for ND Resolution – 8400 FEC Table Limit
3. Number of Devices/Clients limited to 43,000 (8320) / 47,000(8325) for ARP/ND Resolution – Egress MAC Table Limit

IPv6 Link Local

Unique address per SVI

- Generate a unique link local address per L3 interface
- SVI, LAG, RoP
- No topology restrictions any longer for multiple SVIs and multiple upstream VSX LAGs due to IPv6 single LL address

CLI

Enhancements

- “Do” not needed for any show command whatever the context.
- Range

Range Contexts

CLI support for Range contexts

- Useful in creating a range of Interfaces / VLANs and applying configurations in one go.
- Some of the most common use cases are:
 - Changing the admin state of any range of interface / VLANs.
 - Configuring an interface as an L2 / L3 interface.
 - Configuring vsx-sync for a range of interfaces / VLANs.
 - For e.g. user can configure 4K SVIs using a single command i.e. by creating a range context.
- Can't be used to apply configurations which need unique value (eg. setting an IP address to an interface / VLAN) need to be done separately or use Netedit
- Configured using CLI
- Range context configurations can be verified by “show running-config”

Range context Config on Physical interfaces/ VLANs/ SVIs (Interface VLANs)

- Range contexts can be created/deleted at global configuration context.
 - 'interface <RANGE>' / 'no interface <RANGE>'
 - 'vlan <RANGE>' / 'no vlan <RANGE>'
 - 'interface vlan <RANGE>' / 'no interface vlan <RANGE>'

- Range can be specified using a hyphen (-) and / or a comma separated list.

```
DUT-01(config)# interface 1/1/1,1/1/7-1/1/8
```

```
DUT-01(config-if-<1/1/1,1/1/7-1/1/8>)#
```

```
DUT-01(config)# vlan 11,12-20
```

```
DUT-01(config-vlan-<11,12-20>)#
```

```
DUT-01(config)# interface vlan 1-3,5-6
```

```
DUT-01(config-if-vlan-<1-3,5-6>)#
```

- On successfully configuring a range of interfaces/ VLANs/ SVIs, the range shall appear on the context prompt (The interface /VLAN/ SVI name(s) does not appear in single interface context).

- Any configuration attempted inside a range context will be applied on each instance (Interface / VLAN) in the range one by one.

- Eg: Configuring "no shutdown" on a physical interface range.

```
DUT-01(config)# interface 1/1/1-1/1/3
```

```
DUT-01(config-if-<1/1/1-1/1/3>)# no shutdown
```

Eg: Error while applying unique value on range

```
DUT-01(config)# interface 1/1/1-1/1/3
```

```
DUT-01(config-if-<1/1/1-1/1/3>)# ip address 10.0.0.1/24
```

```
[1/1/2] Overlapping networks observed for "10.0.0.1/24". Please configure non overlapping networks.
```

```
[1/1/3] Overlapping networks observed for "10.0.0.1/24". Please configure non overlapping networks.
```

Range Context – Best practices

- Interface range must be specified in an increasing order else CLI will throw an error message.

```
DUT-01(config)# interface 1/1/7-1/1/3
```

```
Invalid input: 1/1/7-1/1/3
```

```
DUT-01(config)#
```

- VLAN range can be specified in any order. User is taken to a context with the range displayed on the prompt.

```
DUT-01(config)# vlan 10,9,8
```

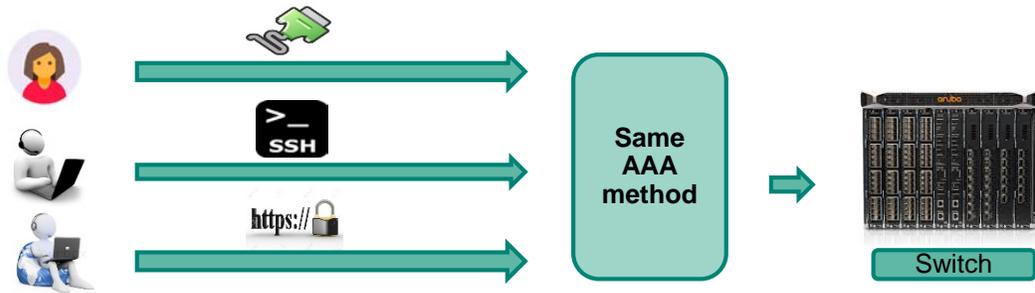
```
DUT-01(config-vlan-<10,9,8>)#
```

- Any Interface VLAN (SVIs) can only be created if corresponding VLAN is already configured. If one or more interface vlans does not have corresponding VLAN(s) configured, they shall not be configured and user will not be taken to the range context.
- While deleting a range of VLANs, ensure corresponding VLAN Interfaces are deleted.
- Avoid giving duplicate numbers in the range as the Interfaces / VLANs are created one after another and Range contexts are not equipped to handle duplicate Interfaces / VLANs. Configurations that are associated with a unique interface / VLAN parameters such as IP address do not apply using range context
- Use Range contexts only when a configuration can be applied to multiple entities without causing any error for e.g. enabling or disabling the interface, configuring mtu value for interfaces.

AAA Channels

Overview of AAA Channels

- This is an extension to the existing AAA feature for the switch management users.
- Prior to 10.3, there was no granularity to configure different AAA methods based on the connection type (channel) of the user.



- Connection-type or channel refers to how the user is accessing the switch.
 - The channels being:
 1. SSH : Users accessing the switch via Secure Shell connection.
 2. Console : Users accessing the switch via console cable.
 3. HTTPS-Server : Users accessing the switch via REST APIs or WebUI.
- AAA channel enables configuration of different AAA methods based on the connection type of the user's session.

Authentication channel separation

CLI comparison between prior to 10.3 and 10.3

Before AAA channel separation

```
aaa authentication login default local | group <list>
```

After AAA channel separation (10.3)

```
aaa authentication login console local | group <list>
```

```
aaa authentication login ssh local | group <list>
```

```
aaa authentication login https-server local | group <list>
```

```
aaa authentication login default local | group <list>
```

Authorization channel separation

CLI comparison between prior to 10.3 and 10.3

Before authorization channel separation

```
aaa authorization commands default none | group <list>
```

After authorization channel separation

```
aaa authorization commands console none | group <list>
```

```
aaa authorization commands ssh none | group <list>
```

```
aaa authorization commands default none | group <list>
```

Accounting channel separation

CLI comparison between prior to 10.3 and 10.3

Before accounting channel separation

```
aaa accounting all default start-stop local | group <list>
```

After accounting channel separation

```
aaa accounting all console start-stop local | group <list>
```

```
aaa accounting all ssh start-stop local | group <list>
```

```
aaa accounting all https-server start-stop local | group <list>
```

```
aaa accounting all default start-stop local | group <list>
```

Authentication channels – Examples

Different authentication method configured for every channel

```
exit
8320# configure terminal
8320(config)# radius-server host 10.0.0.1 key plaintext testing123
8320(config)# tacacs-server host 20.0.0.1 key plaintext testing123
8320(config)# aaa authentication login console local
8320(config)# aaa authentication login ssh group tacacs
8320(config)# aaa authentication login https-server group radius
8320(config)# aaa authentication login default group tacacs radius local
8320(config)#
```

- Local authentication will be used for users logging in via console.
- TACACS authentication will be used for users logging in via SSH.
- RADIUS authentication will be used for users logging in via REST/WebUI.

No explicit authentication configuration for some channels

```
8320(config)# radius-server host 10.0.0.1 key plaintext testing123
8320(config)# tacacs-server host 20.0.0.1 key plaintext testing123
8320(config)# aaa authentication login console local
8320(config)# aaa authentication login default group radius
8320(config)#
```

- Local authentication will be used for users logging in via console.
- There is no explicit authentication list configured for SSH and https-server. Hence the authentication list from default will be used. RADIUS authentication will be used for users logging in via SSH/REST/WebUI.

Best practices/Typical use case

- Typical customer use case of AAA channels would be to set up local AAA via console and set up remote AAA servers (RADIUS/TACACS) for management users logging in to the switch via remote connections such as ssh/console/https-server.
- Console is considered as a trusted channel compared to remote connections as only users with physical access to the switch can use console cable to login to the switch.
- Typical configuration would be:
 - aaa authentication login console local
 - aaa authentication login default radius|tacacs local
 - aaa authorization commands console none
 - aaa authorization commands default tacacs local
 - aaa accounting all console start-stop local
 - aaa accounting all default start-stop group radius|tacacs local
- Above configuration ensures that console users are not impacted by latency to reach AAA servers. This way console users can troubleshoot the switch when there are any issues with the remote AAA server.
- Prior to 10.3, this segregation in AAA configuration was not possible

PKI

PKI

Application

```
8325-1(config)# crypto pki
  application  Configure an application's PKI setting
  certificate  Configure or remove a leaf certificate
  ta-profile   Configure or remove a trust anchor profile
```

```
8325-1(config)# crypto pki application
  hsc          Hardware Switch Controller
  https-server HTTPS Server
  syslog-client Syslog Client
```

Certificate Management

- The Public Key Infrastructure (PKI) feature on an Aruba network switch enables identification and authentication of network entities on the switch.
- It provides configuration and management of digital certificates on the switch, a key component of establishing digital identity within a PKI of a network.
- Each entity in a PKI of a network has its identity validated by a Certificate Authority (CA) that is trusted by all parties.
- The Certificate Management APIs helps the REST clients to generate, install, get and delete certificates from a switch. All certificates are in the PEM format for rel 10.03.

Configuration / Best practices

Combined example of auto-gen custom REST use case.

- Install TA-profile
- POST Generate certificate with pending CSR
- GET generated CSR
- PUT after signing with TA
- Apply certificate with some of the applications (syslog, https-server)

Global Policy (8400 only)

Global Policy: Introduction

- A Global Policy is applied on **all** ports and VLANs i.e. it will be applied, **globally**
 - “Global” is considered another “apply context” similar to port and VLAN
 - Available on 8400 in 10.3
- **Why Global Policy ?**
 - Allows customer to manipulate traffic *regardless* of where it enters the switch
 - No need to apply policy on every port and/or VLAN
 - Elegant and clean CLI
 - Less hardware entries (TCAM) required
- **Use cases :**
 - Need to subject all traffic, regardless of where it enters the switch, to same policy
 - Example use case : Mirror all traffic on a switch to a specific mirror destination

Global Policy: Details

▪ Scalability

- One instance of a global policy
- Max 8000 entries per class
- Max 512 entries per policy

▪ Mutual Exclusions and Interactions with Other Features

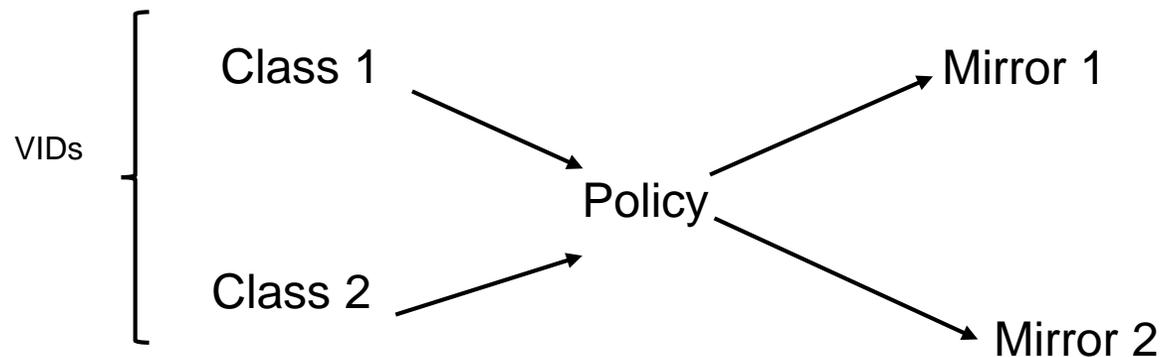
- **Port** policy takes precedence over **Global** policy for the same action
- **VLAN** policy takes precedence over **Global** policy for the same action
- “**drop**” action on global policy and “**mirror**” on a regular policy (or vice versa) can coexist
- When interacting with ACLs, “**drop**” action always wins regardless of whether it is on a global policy or an ACL.

▪ Future planning

- Applying policies to VRF may be another future enhancement
- Global policy being considered for 10.4 on additional platforms

Global Policy: Example

- **2 VLANs** : 2009 and 3500
- **2 Classes** with entries matching each VID
- **1 Policy** with **2 policy entries**, corresponding to each of the 2 classes
- **2 Mirror destinations**
- Global Policy mirrors traffic matching a particular VID onto the corresponding mirror destination
- Traffic will be mirrored **regardless** of which port sends the traffic



```
class ip c1
  10 match any any any vlan 2009 count
class ip c4
  10 match any any any vlan 3500 count
policy p
  10 class ip c1 action mirror 1
  20 class ip c4 action mirror 2
```

Global Policy: Applying and removing

- Global policy is applied/unapplied from the “configure” context
- An extension of port and VLAN policies.
- Class and policy creation/deletion CLI remains the same.
- To apply a global policy, named p :
– apply policy p in `8400X(config)# apply policy p in`
- To un-apply a global policy, named p : `8400X(config)# no apply policy p in`
– no apply policy p in
- Use “**show running-config**” to make sure there are no errors

Global Policy: Hitcounts commands

- To view the hitcounts for a global policy, named p :
 - **show policy hitcounts global**

```
8400X(config)# show policy hitcounts global
Statistics for Policy p:
Global Policy:
      Hit Count  Configuration
10 class ip c1 action mirror 1
      1147685200  10 match any any any vlan 2009 count
20 class ip c4 action mirror 2
      1147685174  10 match any any any vlan 3500 count
* policy statistics are shared among each context type (interface, VLAN).
  For routed ingress, they are only shared within the same VRF.
  Use 'policy NAME copy' to create a new policy for separate statistics.
```

- To clear the hitcounts for a global policy, named p :
 - **clear policy hitcounts global**
 - Clearing should set hitcounts to 0

```
8400X# clear policy hitcounts global
8400X# show policy hitcounts global
Statistics for Policy p:
Global Policy:
      Hit Count  Configuration
10 class ip c1 action mirror 1
      0          10 match any any any vlan 2009 count
20 class ip c4 action mirror 2
      0          10 match any any any vlan 3500 count
* policy statistics are shared among each context type (interface, VLAN).
  For routed ingress, they are only shared within the same VRF.
  Use 'policy NAME copy' to create a new policy for separate statistics.
```

Global Policy: Display command

- Display the configured global policy and apply commands :

8400X (config)# show policy global

```
8400X(config)# show policy global
Direction
      Name
      Additional Policy Parameters
Sequence Comment
      Class Type
              action
-----
Inbound
  10  P
      c1  ipv4
          mirror 1

  20
      c4  ipv4
          mirror 2
```

8400X (config)# show policy global commands

```
8400X(config)# show policy global commands
policy p
  10 class ip c1 action mirror 1
  20 class ip c4 action mirror 2
apply policy p in
```

Global Policy: Best Practices

▪ Best practices for using Global policy

- Only **one** Global Policy can be applied.
- If a second policy is applied globally, it replaces the first one.
- On 8400, Global policy takes precedence over routed-in PBR policy

▪ Best practices for configuring Global Policy

- Global policy can only be applied in ***ingress*** direction
- It does **not** support “routed-in” and “egress” directions
- PBR(Policy based routing) will **not** be supported due to above
- There is no creation/deletion of any context i.e. unlike VLAN and interface contexts, the global context **cannot** be deleted.

Conclusion

- VSX Live Upgrade : <300ms impact
- DC features

Thank You