TECHNICAL WHITEPAPER

aruba
a Hewlett Packard Enterprise company

VSX Configuration Best Practices for Aruba CX 6400, 8320, 8325, 8400

AOS-CX VERSION 10.4

Revision 1.0 - December 2019

CONTENTS

Revision History

The following table lists the revisions of this document:

Revision	Date	Change Description
1.0	December 2019	Initial Release (6400, 8320, 8325, 8400)

3333 Scott Blvd. | Santa Clara, CA 95054
1.844.472.2782 | T: 1.408.227.4500 | FAX: 1.408.227.4550 | info@arubanetworks.com

Overview

VSX or Virtual Switching Extension is a virtualization technology to create a cluster of two Aruba CX switches from the same model 6400, 8320, 8325, or 8400 (not supported on Aruba CX 6300). Different platforms cannot be mixed in the same VSX pair; i.e. 8325 cannot be mixed with 8320. Here is the list of supported combinations:

- Aruba CX 6400: 6405+6405 or 6410+6410. The pair 6405+6410 is not supported.
- Aruba CX 8320: JL479A+JL479A, JL581A+JL581A, JL579A+JL579A, JL479A+JL581A, JL479A+JL579A, JL581A+JL579A
- Aruba CX 8325: JL624A/JL625A+JL624A/JL625A, JL626A/JL627A+JL626A/JL627A, JL624A/JL625A+ JL626A/JL627A
- Aruba CX 8400: 8400+8400

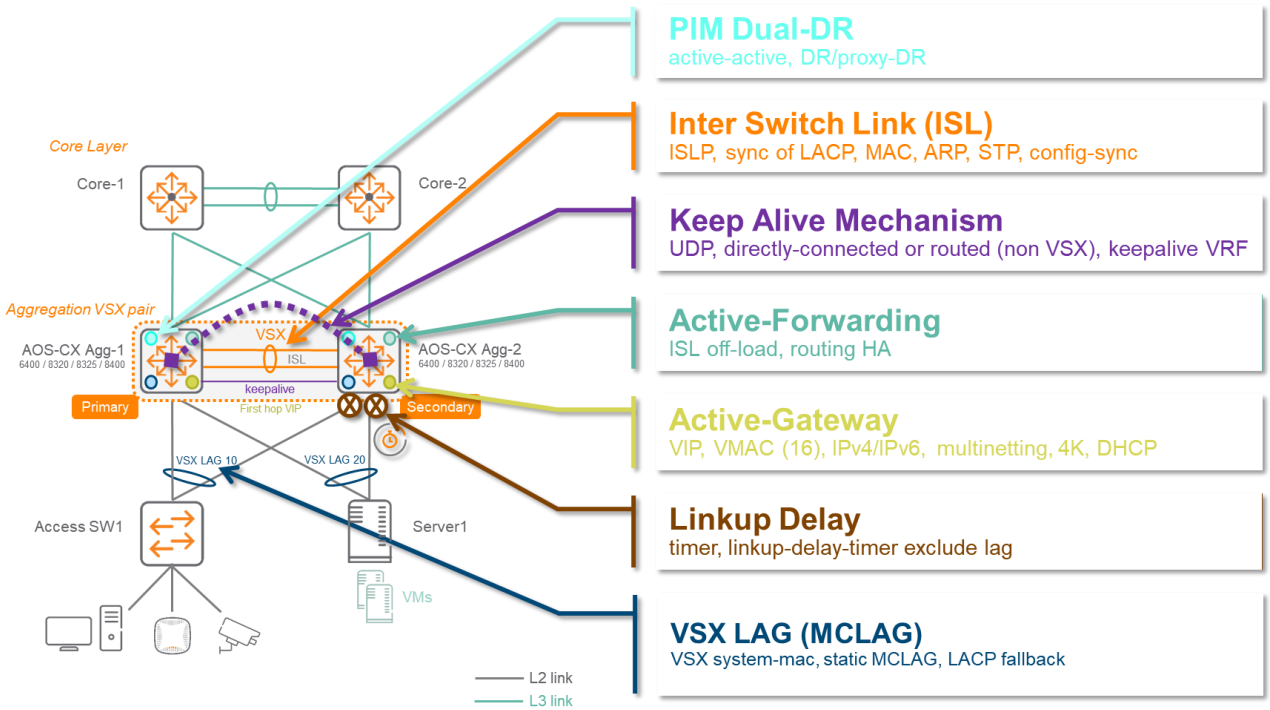
VSX comprises multiple technologies:

- Multi-Chassis Link Aggregation for data-plane virtualization and removal of Spanning-Tree need, with synchronization of the following tables between the two VSX nodes: LACP, MAC, ARP, STP (if used), DHCP.
- Synchronized Management plane: Network admin can choose which particular feature configuration is synchronized from the VSX primary to the VSX secondary. Any show command or REST API call can display information of the VSX peer from the local node with a single SSH session or API call.
- Orchestrated upgrade (VSX Live Upgrade) for sub-second traffic impact during VSX cluster firmware update.
- Independence of the control-planes of the VSX member for protocols High-Availability.

The VSX link aggregation technique is recommended for the Campus aggregation/distribution layer or for Datacenter Top-of-Rack layer. The below topology shows VSX LAG for Campus Access switch uplinks and as well for server NIC attachments in the Datacenter.

VSX Components

Here is a synthetic view of all the VSX components inside a VSX cluster:



Inter Switch Link (ISL)

Both VSX switches are connected through the Inter Switch Link. The ISL is typically a standard LAG (Link Aggregation) of two to eight physical links. The ISL physical ports must be directly connected with no intermediate active L2 nodes, and must have same speed (for example: 2x 100G). The ISLP protocol running over the ISL is used to synchronize multiple software and hardware tables: LACP, MAC, ARP tables, STP states (if configured) and DHCP states (if used). There is no additional encapsulation of data-plane traffic over the ISL, so there is no requirement for specific MTU adjustment due to VSX. However, it is recommended to adjust the MTU of the ISL link to permit transport of jumbo frames if required by end-points. The ISL has its own keepalive mechanism with ISL hello packet, dead-interval and hold-time. Keeping the default values is recommended.

ISLP protocol comprises also a framework for synchronization of the management plane from the VSX primary to the VSX seconday. This "vsx-sync" mechanism is very helpful to avoid configuration human mistakes on the VSX secondary.

VSX LAG

VSX LAG is the Aruba CX name of MCLAG (Multi-Chassis Link Aggregation) technology where two or more links across two switches are aggregated together to form a LAG which will act as a single logical interface. VSX LAG does support all the standard LAG adjustments: timers, L2 or L3 hashing, LACP fallback. It supports both LACP mode active or static mode, and only Layer 2 (i.e. no routed mode).

VSX Keepalive

The VSX keepalive is a UDP probe on port 7678 (configurable) sending hellos between the two VSX nodes to detect a split brain situation. This L3 probe can be established over a direct link or routed path. It is not yet supported over OOBM port. If a split is detected, the VSX secondary will tear down "by feature" all the physical ports which are members of any VSX LAG.

Active-Gateway

This is the default-gateway Virtual IP address of endpoint subnet hosted on both VSX primary and VSX secondary. It comes with a Virtual MAC address. Both VSX nodes reply to ARP request with the same VMAC. The VSX primary periodically sends GARP and Broadcast Hello packets sourced from this VMAC. The VSX primary also relays DHCP request or serves DHCP offer (secondary takes over in case of primary failure).

Active-Forwarding

Active-Forwarding is configured in case of VSX LAGs are used for upstream connectivity and when North-South traffic can reach one VSX node while the destination MAC is actually the other VSX peer (this can happen due to ECMP on upstream routers). In order to avoid such sub-optimum path and unnecessarily traffic load of the ISL, active-forwarding can be optionally configured on the upstream Transit VLAN SVI. When configured, the node receiving packets will process L3 lookup on behalf of the other VSX node and will forward traffic to connected downstream devices without forwarding traffic to ISL.

PIM Dual-DR

PIM DR and PIM routing is handled by the VSX primary. In order to avoid a long convergence time in case of a VSX primary failure, the VSX secondary can also establish PIM peering, send PIM join messages and build shortest path tree for multicast. This is achieved when PIM active-active is configured.

Linkup-Delay

Inside the VSX cluster, the switches synchronize their ARP and MAC ASIC tables. This happens in two phases: a software exchange, and a hardware programming of the ASICs. For this second part, when the VSX node joins the VSX cluster it can take several minutes (depending on network size) to get the complete MAC and ARP tables synchronized and configured. During this synchronization time, the VSX LAG ports are kept down to make sure that there won't be any transient traffic drop. This timer is user-configurable and will depend on the network size.

VSX Features Summary

VSX has been introduced in AOSX-CX 10.1 version and has been improved in each major release.

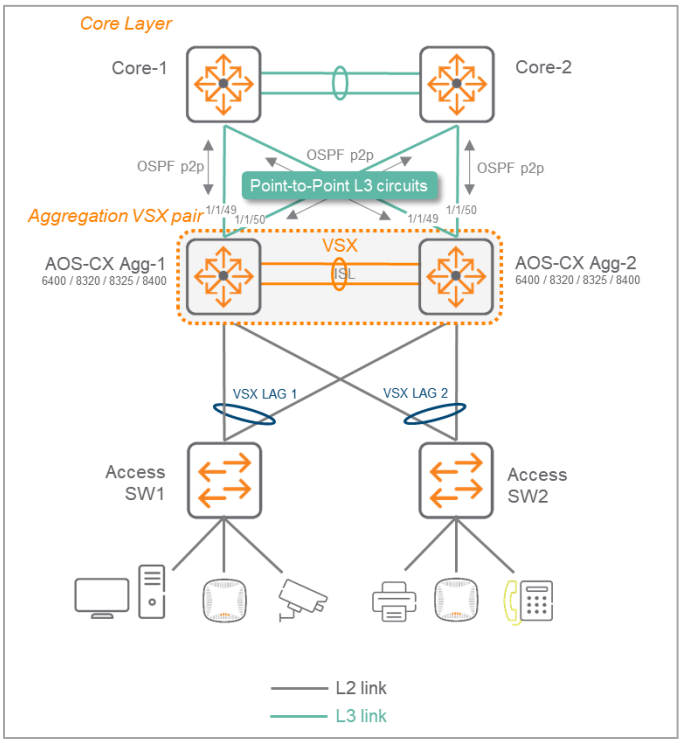
Feature	10.1	10.2	10.3	10.4
VSX + Spanning-tree (MSTP or RPVST+)	No	MSTP only	MSTP / RPVST+	
Multicast Active-Active	No	Yes: Dual-DR		
VSX static LAG	No	Yes		
VSX manual system-mac	No	Yes		
VSX split-recovery	No	Yes		
VSX LACP fallback	No	Yes		
MVRP (Multiple VLAN Registration Protocol)	No			
VSX active-gateway and VRRP	No		Yes: Global co-existence, mutually exclusive per SVI	
VSX active-gateway multinetting	No			Yes
LACP graceful shutdown (during VSX live upgrade)	No			Yes
OSPF and BGP graceful shutdown (during VSX live upgrade)	No		Yes	
VSX with BGP EVPN VXLAN	No			Yes
Keepalive over OOBM	No			
DHCP relay (active on primary, standby on secondary)	Yes			
DHCP server and lease synchronization within VSX	No		Yes	
Gratuitous ARP on active-gateway (sent by primary)	No	Yes		
VSX Live Upgrade orchestration from CLI	No	Yes		
VSX Live Upgrade orchestration from WebUI	No			
VSX-sync (pseudo single management plane)	VLANs, ACLs, Class, Policy	+ feature-group tags	VLAN range sync + new feature-group tags	+ new feature-group tags (ospf, bgp, evpn, vrrp…)
VSX linkup delay optimization	No		Partial	Yes

Topologies and Use-cases

In the following topologies and configuration, the interface ID are for 8325 implementation example. Please adjust these port IDs when using other platforms.

The first consideration of this chapter is the upstream connectivity of the VSX cluster with two options: L3 point-to-point circuits or L2 circuits with VSX LAG. The second consideration is about VSX technology usage on both Aggregation layer and Access/ToR layer.

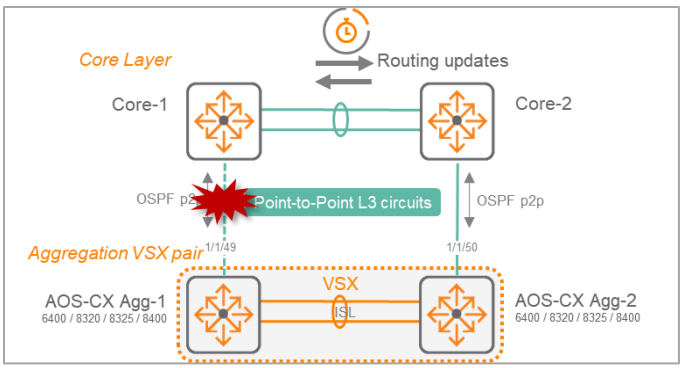
Aggregation VSX with single VRF routing model



Please note that on following topologies, Core-1 and Core-2 can be Aruba CX Switches and, thanks to Aruba adoption of Open Standards, it can also be third party devices like firewalls.

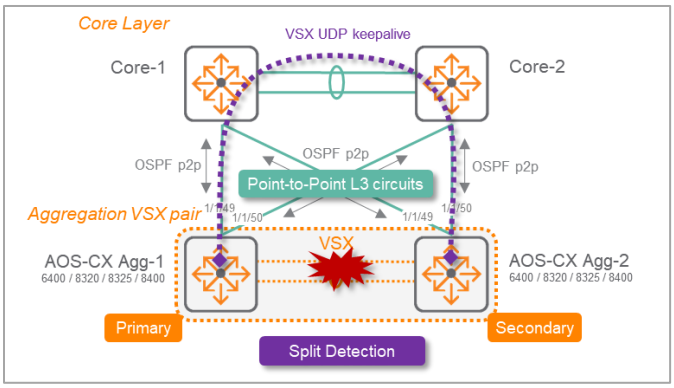
The links between the VSX cluster and Core-1/Core-2 are Layer 3 circuits using Routed ports (here 1/1/49, 1/1/50) on each CX switch. It is recommended to use OSPF as dynamic routing protocol for the IGP (Interior Gateway Protocol) and OSPF peering is configured with point-to-point network-type to skip the DR/BDR election.

It is recommended to implement a full mesh topology with cross connections: Agg-1 connected to both Core-1 and Core-2, and Agg-2 connected to both Core-1 and Core-2.



With a non-meshed topology, if a link fails between Agg-1 and Core-1, there will be some routing convergence to happen and consequently some traffic impact. Even if OSPF spf-trottle timers can be adjusted to minimize such impact to sub-second, this topology should be restricted to environment where full-mesh fibers is not possible.

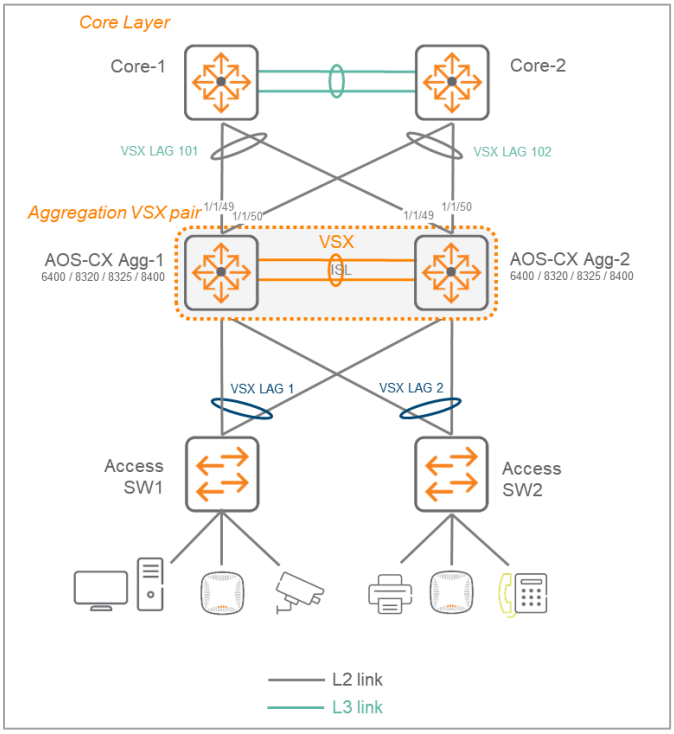
Note: The VSX keepalive UDP traffic is possible over the L3 upstream connectivity.

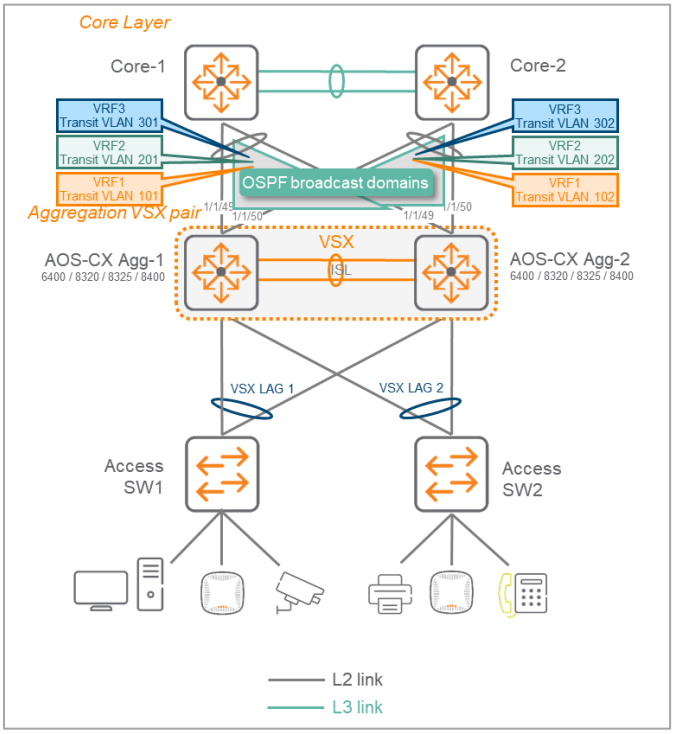


VSX UDP keepalive probes between the VSX primary and the VSX secondary can be routed through upstream Core-1 and Core-2 during a split event.

Aggregation VSX with multiple VRF routing model

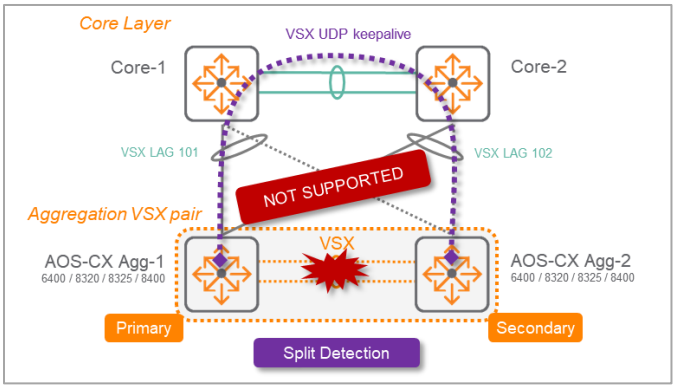
As sub-interface is not yet supported in AOS-CX (10.4 and below), upstream circuits must carry the various VRFs through dedicated Transit VLANs for interconnecting VSX cluster routing domain to upstream Core layer for each VRF. In order to avoid duplication of multiple Transit VLANs per VRF and to minimize SPF routing calculation, the upstream L2 circuits are merged into two upstream VSX LAGs.



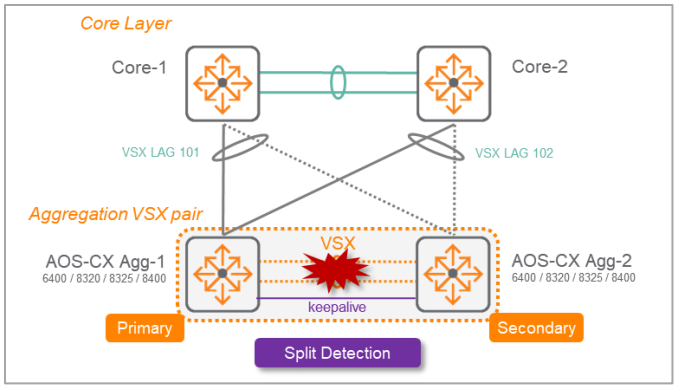


Note: VSX UDP Keepalive is NOT possible over the upstream VSX LAG connectivity. In such scenario, the keepalive would need to be implemented over a dedicated point-to-point L3 circuit that does not need to be directly connected, i.e. some intermediate active equipment can be traversed by the L3 UDP probe as long as it does not use a VSX LAG path.

VSX UDP keepalive would have been possible through L2 point-to-point circuits. However, due the number of associated Transit VLANs to manage (5xnumber of VRFs), this would be the very last implementation option.

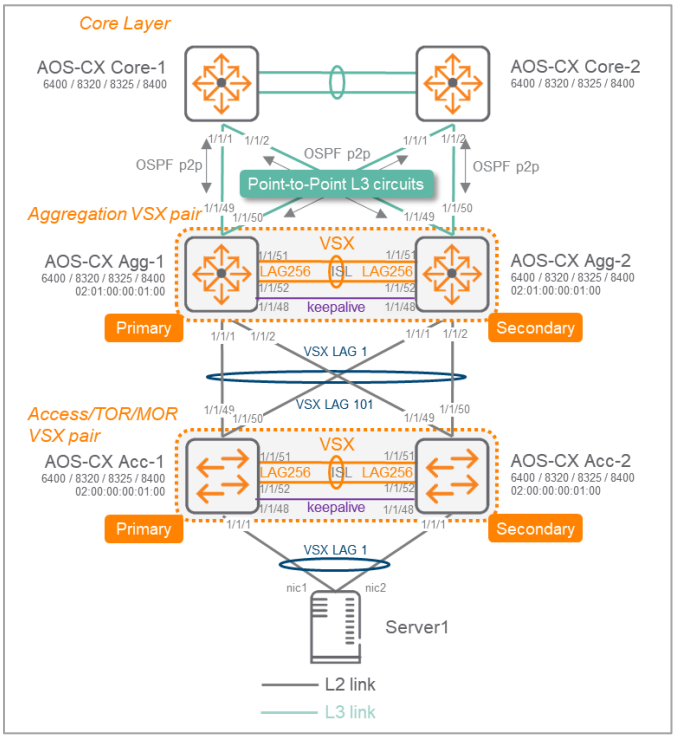


UDP keepalive not supported over VSX LAG



For VSX LAG upstream, UDP keepalive over direct link is recommended

Access VSX to Aggregation VSX



For mission critical campus endpoint or for datacenter servers, VSX technology in the Access/ToR layer might be preferred over VSF. On the left is the corresponding topology with one VSX cluster for the aggregation layer and one VSX cluster for one top-of-rack instance.

The connectivity between the Aggregation cluster and the ToR cluster is built with 4 physical meshed links all part of the same VSX LAG: VSX LAG1 on Aggregation and VSX LAG101 on ToR, providing a single logical link without the need for spanning-tree.

VSX and L2 loop protection mechanisms

With VSX LAG there is a single logical data-plane across multi-chassis which removes the need of spanning-tree as a Layer 2 control-plane to manage forwarding and discarding circuits. When all downlink ports are configured for VSX LAGs, if LACP does not detect partner, then the downlink ports will be blocked, avoiding any possible layer 2 loop. This can be enough protection and considered as the simplest option for some network operations.

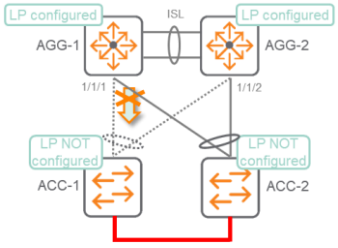
However, some other cases like cabling errors or configuration mistakes during operations may induce the requirement for loop-protection or spanning-tree technologies as an additional protection mechanism. This is the purpose of this section to explain the various options.

Native VLAN trunking exclusion from Access to Aggregation Layer

The best practice for allowed VLANs is to exclude the native VLAN 1 from being propagated. This is a very robust method to avoid Layer2 storm propagation due to potential loop initiated on an access switch. By default, AOS-CX CLI will not include VLAN 1 as allowed VLAN on VLAN trunking. In case of access switch Zero-Touch-Provisioning requirements, VLAN 1 removal has to be performed after ZTP process.

VSX and Loop-protection

The best practice for loop-protect is to configure it only on the access layer without including the uplinks. This practice avoids to isolate the entire access switch if a loop is created between two access switches as exposed on the right side.



Consequently loop-protect is not configured on the aggregation layer in favor of MSTP, especially to protect against such loop scenario that customers would like to protect against.

Loop-protection has a default transmit-interval higher than MSTP, so in case of loop, MSTP will block the port before loop-protection does. Nevertheless, it is a best practice to enable it on all end-point ports as an additional protection mechanism and it may prevent a loop if MSTP is disabled for any reason.

VSX and MSTP

The usage of MSTP in the context of VSX can be compared as an enhanced loop-protect mechanism with more control over loop avoidance parameters. MSTP is the recommended best practice to protect the network infrastructure against mistake or cabling errors. The associated configuration is kept as simple as possible as this is a protection mechanism and not a forwarding control-plane due to VSX LAG benefits. The default instance MST0 is used as the common instance for all VLANs. Although optional, it is recommended to have Root Bridge hosted by VSX for simplicity and easier support during operations. In such a case, consistent approach would lead to configure root-guard on all downstream VSX LAGs to access switches.

VSX and RPVST+

Rapid Per VLAN Spanning-Tree + can be used to protect against layer 2 loops. This use-case is reserved for interoperability with existing devices already running RPVST+. As MSTP is enabled by default on 6300/6400, MSTP is the best practice over RPVST+.

VSX and VXLAN

Since 10.4, VSX supports VXLAN and termination of a VXLAN tunnel with anycast Virtual IP address of the VSX cluster. Configuration Best Practices for VSX and VXLAN are covered in a separate white paper.

VSX Deployment and Configuration – Best Practices

Virtual MAC and System-MAC Guidance

One of the main VSX **best practice is to set VSX system-mac** and not leave it blank with default HW system-mac being used. By doing so, the VSX system-mac is independent from the physical hardware MAC address and in case of hardware replacement of the VSX primary, the new switch can be configured with the same configuration than the previous primary unit with no impact on the VSX secondary as the cluster ID remains unchanged. With such practice, VSX primary HW replacement is hitless for the VSX secondary. (Otherwise the VSX secondary would have to join a new cluster ID, ID from VSX primary, and would turn-off temporary its VSX LAG ports).

Please use locally administered unicast MAC Address when assigning system-mac or active-gateway virtual MAC address. There are 4 ranges reserved for private use for unicast (with second least significant bit of the first octet of the unicast address set to 1). x is any Hexadecimal value.

- x2-xx-xx-xx-xx-xx
- x6-xx-xx-xx-xx-xx
- xA-xx-xx-xx-xx-xx
- xE-xx-xx-xx-xx-xx

In this document, 02:01:00:00:01:00 is used or system-mac and 12:01:00:00:01:00 is used for active-gateway Virtual MAC.

Here is an example proposal to have unique values being used in the administrative domain is the following:

Function	System-mac	Active-gateway Virtual MAC
Access / TOR Layer	02:00:00:00:XX:00	12:00:00:00:XX:0Y
Aggregation	02:01:00:00:XX:00	12:01:00:00:XX:0Y
Core / Spine	02:02:00:00:XX:00	12:02:00:00:XX:0Y

Where XX is reflecting the Unique Cluster ID in the function, and Y is the Virtual MAC ID (0 to 15)

The scope of this VMAC is purely link-local. Consequently, **the same Virtual MAC address value can be used on any L3 VLAN interface (SVI)**.

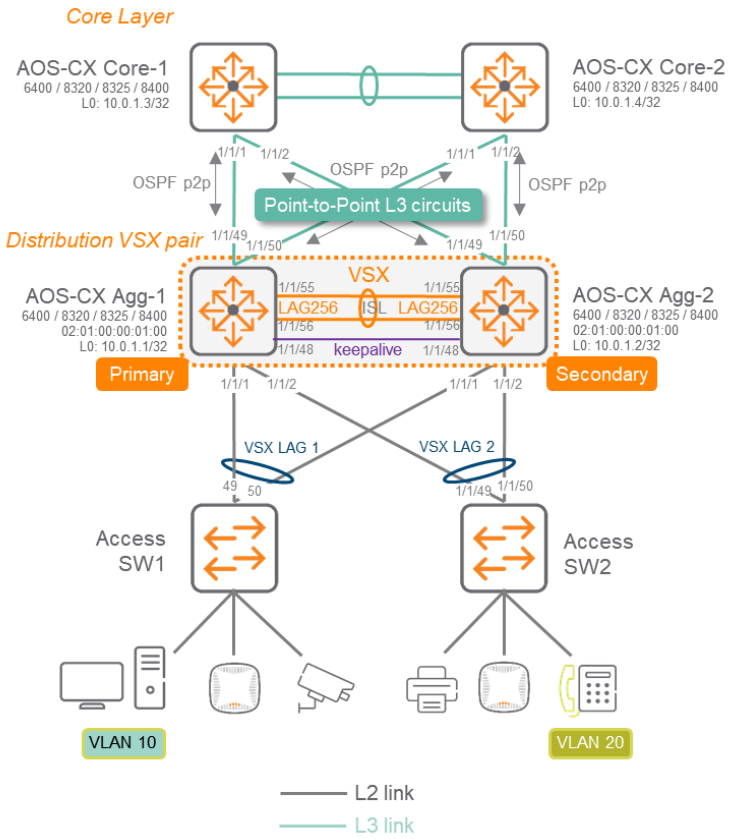
If some servers or systems have dual-attachment to two different SVIs, and the system administrator would like to see distinct MAC addresses for the next-hops over these separate interfaces, then 16 VMACs are available. For dual-stack IPv4 and IPv6, 16 VMACs can be used for IPv4 and the same VMACs can be used for IPv6. It is however a best practice to use only 8 VMACs for IPv4 and 8 different VMACs for IPV6.

Note: any other allocation rules can be chosen according to administrative rules in place by the network operational team. **Multicast or broadcast MAC addresses must not be used for System-mac**.

Aggregation VSX with single VRF routing model

Here is the typical topology:

(default VRF)



Step #0 -Pre-requisite : same firmware release

Please install same version on both CX Switches that will create the VSX cluster. It is better to avoid any version mismatch during the creation of the cluster, as a warning would appear.

AGG-1	AGG-2
AGG-1# show version ---------- ---------- ArubaOS-CX (c) Copyright 2017-2019 Hewlett Packard Enterprise Development LP ---------- ---------- Version : GL.10.04.0001 Build Date : 2019-10-31 12:33:52 PDT Build ID : ArubaOS-CX:GL.10.04.0001:489a60c44c86:201910311907 Build SHA : 489a60c44c86b788edc6808b1e9a4d217f31e3bf Active Image : secondary Service OS Version : GL.01.05.0002 BIOS Version : GL-01-0013	AGG-2# show version ---------- ---------- ArubaOS-CX (c) Copyright 2017-2019 Hewlett Packard Enterprise Development LP ---------- ---------- Version : GL.10.04.0001 Build Date : 2019-10-31 12:33:52 PDT Build ID : ArubaOS-CX:GL.10.04.0001:489a60c44c86:201910311907 Build SHA : 489a60c44c86b788edc6808b1e9a4d217f31e3bf Active Image : secondary Service OS Version : GL.01.05.0002 BIOS Version : GL-01-0013

Step #1 : create LAG for ISL

It is assumed that 2x 40G or 2x50G or 2x100G direct fibers / DACs are already interconnecting AGG-1 and AGG-2.

The best practice for ISL bandwidth is at least 2x40G (QSFP+) or 2x 50G (SFP56) or 2x100G (QSFP28). It is technically possible to use 2x10G or 2x25G; however it is recommended to plan for any uplink failure and associated impact on the bandwidth requirement for the ISL. If the

uplinks from AGG-1 fail, traffic from AGG-1 will be redirected to AGG-2 over the ISL before reaching the upstream layer. This is perfectly fine as long as there is enough bandwidth remaining for the ISL protocol and the control-plane communication. It is recommended to size the ISL bandwidth to be equal to, at least, the sum of uplinks bandwidth of one VSX switch. The best practice rule is to size the ISL bandwidth according to the failure domain target.

The **best practice for ISL physical ports** is to select at least two ports of the same speed (2x40G or 2x50G or 2x100G), and, in case of a chassis, to select these ports from different Line Cards.

AGG-1

```
AGG-1# show lldp neighbor-info

LLDP Neighbor Information
=========================

Total Neighbor Entries          : 3
Total Neighbor Entries Deleted  : 0
Total Neighbor Entries Dropped  : 0
Total Neighbor Entries Aged-Out : 0

LOCAL-PORT  CHASSIS-ID         PORT-ID       PORT-DESC
TTL      SYS-NAME
-------------------------------------------------------
-----------------------
1/1/48      54:80:28:fd:42:00  1/1/48        1/1/48
120      AGG-2
1/1/55      54:80:28:fd:42:00  1/1/55        1/1/55
120      AGG-2
1/1/56      54:80:28:fd:42:00  1/1/56        1/1/56
120      AGG-2
```

AGG-2

```
AGG-2# show lldp neighbor-info

LLDP Neighbor Information
=========================

Total Neighbor Entries          : 3
Total Neighbor Entries Deleted  : 0
Total Neighbor Entries Dropped  : 0
Total Neighbor Entries Aged-Out : 0

LOCAL-PORT  CHASSIS-ID         PORT-ID       PORT-DESC
TTL      SYS-NAME
-------------------------------------------------------
---------------------
1/1/48      54:80:28:fc:ac:00  1/1/48        1/1/48
120      AGG-1
1/1/55      54:80:28:fc:ac:00  1/1/55        1/1/55
120      AGG-1
1/1/56      54:80:28:fc:ac:00  1/1/56        1/1/56
120      AGG-1
```

Here is the associated configuration to create the standard LAG to be used for ISL.

The **best practice for LAG numbering** is to use the last available LAG ID (ie. 256 in AOS-CX10.4) for the ISL, so that LAG ID=1 is used for connecting the Access Switch#1 on port 1/1/1, so that LAG 2 is used to connect the second Access Switch on port 1/1/2, and so on…

The **best practice for VLAN** trunking on the ISL LAG is to permit ALL VLANs, for simpler configuration. Specifying a restrictive list of VLAN IDs is entirely valid if the network admin wants more control.

The **best practice for LACP timers** on the ISL LAG is to keep the default long timer (30s for lacp rate slow).

The **best practice for hashing algorithm** on the LAG is to keep the default l3-src-dst (alternative being l2-src-dst).

The **best practice for MTU** is to configure on all devices the appropriate size to support features such as Dynamic Segmentation as well as other protocols/functions which require MTUs larger than 1500 bytes. Care should be taken to ensure that the IP path from access devices (switches or APs) can provide a MTU of at least 1564 bytes to the mobility controllers. Similarly, for datacenter server connectivity, largest MTU will ensure server jumbo frame traffic over ISL. Recommendation: Ethernet MTU = 9198 bytes.

The **best practice for ACL** is to not set any access-list on the ISL LAG in order to avoid designing complex and unnecessary ACL. The ISL can be seen as a virtual data back-plane and security filtering is processed before or/and after crossing the ISL.

The **best practice for QoS trust mode** on the ISL LAG is to rely on the **qos trust dscp** that is **globally configured** on the Aggregation switches. If not configured globally (which is not the recommendation), qos trust dscp has to be set on the ISL LAG.

AGG-1(config)#

```
qos trust dscp
interface lag 256
    no shutdown
    description ISL link
    no routing
    vlan trunk native 1
    vlan trunk allowed all
    lacp mode active
```

AGG-2(config)#

```
qos trust dscp
interface lag 256
    no shutdown
    description ISL link
    no routing
    vlan trunk native 1
    vlan trunk allowed all
    lacp mode active
```

```
interface 1/1/55
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
interface 1/1/56
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
```

```
interface 1/1/55
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
interface 1/1/56
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
```

Please check that LAG is UP.

AGG-1

```
AGG-1# show interface lag256

 Aggregate lag256 is up
 Admin state is up
 Description : ISL link
 MAC Address                 : 54:80:28:fc:ac:00
 Aggregated-interfaces       : 1/1/55 1/1/56
 Aggregation-key             : 1
 Aggregate mode              : active
 Speed                       : 200000 Mb/s
 L3 Counters: Rx Disabled, Tx Disabled
 qos trust none
 VLAN Mode: native-untagged
 Native VLAN: 1
 Allowed VLAN List: all
 Rx
          239 input packets          29559 bytes
            0 input error                0 dropped
            0 CRC/FCS
 Tx
          260 output packets         32110 bytes
            0 input error                0 dropped
            0 collision
```

AGG-2

```
AGG-2# show interface lag 256

 Aggregate lag256 is up
 Admin state is up
 Description : ISL link
 MAC Address                 : 54:80:28:fd:42:00
 Aggregated-interfaces       : 1/1/55 1/1/56
 Aggregation-key             : 1
 Aggregate mode              : active
 Speed                       : 200000 Mb/s
 L3 Counters: Rx Disabled, Tx Disabled
 qos trust none
 VLAN Mode: native-untagged
 Native VLAN: 1
 Allowed VLAN List: all
 Rx
          276 input packets          34312 bytes
            0 input error                0 dropped
            0 CRC/FCS
 Tx
          255 output packets         31523 bytes
            0 input error                0 dropped
            0 collision
```

Please check that LACP is collecting and distributing (flags should be ALFNCD).

AGG-1 / AGG-2

```
AGG-1# show lacp interfaces

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                             Pri    Key  State
------------------------------------------------------------------------------
1/1/55  lag256     56    1     ALFNCD  54:80:28:fc:ac:00 65534  1    up
1/1/56  lag256     57    1     ALFNCD  54:80:28:fc:ac:00 65534  1    up

Partner details of all interfaces:
------------------------------------------------------------------------------
Intf    Aggr       Port  Port  State   System-ID         System Aggr
        Name       Id    Pri                             Pri    Key
------------------------------------------------------------------------------
1/1/55  lag256     56    1     ALFNCD  54:80:28:fd:42:00 65534  1
1/1/56  lag256     57    1     ALFNCD  54:80:28:fd:42:00 65534  1
```

Please note that at this stage VLAN mode is native-untagged (it will change when ISL function is associated to this LAG).

Step #2 : VSX Keepalive pre-requisite

The best practice for Keepalive connection is to use a direct L3 circuit, which can be a low speed port (1G transceiver is enough, 1GBASE-T works as well) between both VSX nodes. This circuit does not have to be directly connected and the path can include active L2 and L3

equipment. Although this requires an additional dedicated port, it brings simplicity of configuration and operations. In the Appendix D, VSX keepalive over upstream layer 3 routing domain is documented as an alternative for network admins who want to protect from a fiber path cut that would impact ISL and keepalive simultaneously; or when the associated cost of a dedicated port is too high (100G). In case of a chassis (6400 or 8400), if possible, it is recommended to use a port from a different Line Card than the ones used for the ISL ports.

The **best practice for Keepalive routing** is to use a dedicated VRF. This is entirely optional and default VRF can be used as well, typically for the single VRF model with UDP keepalive over the upstream L3 domain. Having a dedicated VRF for Keepalive simplifies the oprations and prevents any impact from routing change on the default VRF.

The **best practice for Keepalive subnet** is to use a /31 subnet as only 2 nodes will communicate together.

Create the dedicated KeepAlive VRF and associated interface.

AGG-1(config)#	AGG-2(config)#

`vrf KA`
 	
`interface 1/1/48`
`  no shutdown`
`  vrf attach KA`
`  description VSX keepalive`
`  ip address 192.168.0.0/31`
	`vrf KA`
	
	`interface 1/1/48`
	`  no shutdown`
	`  vrf attach KA`
	`  description VSX keepalive`
	`  ip address 192.168.0.1/31`

Check IP connectivity between future VSX nodes inside this dedicated "KA" VRF.

AGG-1 / AGG-2

```
AGG-1# ping 192.168.0.1 vrf KA
PING 192.168.0.1 (192.168.0.1) 100(128) bytes of data.
108 bytes from 192.168.0.1: icmp_seq=1 ttl=64 time=0.141 ms
108 bytes from 192.168.0.1: icmp_seq=2 ttl=64 time=0.212 ms
108 bytes from 192.168.0.1: icmp_seq=3 ttl=64 time=0.226 ms
108 bytes from 192.168.0.1: icmp_seq=4 ttl=64 time=0.282 ms
108 bytes from 192.168.0.1: icmp_seq=5 ttl=64 time=0.180 ms

--- 192.168.0.1 ping statistics ---
5 packets transmitted, 5 received, 0% packet loss, time 4094ms
rtt min/avg/max/mdev = 0.141/0.208/0.282/0.047 ms
```

Step #3 : VSX Cluster creation

The **best practice for system-mac** is to set the system-mac manually on the VSX primary switch. Please refer to above section for system-mac values. Here, 02:01:00:00:01:00 is used. The main advantage to set VSX system-mac (and not to leave it blank with default HW system-mac being used) is to be independent from the physical hardware MAC address. In case of hardware replacement of the VSX primary, the new switch can be configured exactly with the same configuration than the previous unit and there will be no impact on the secondary which will remain in the same cluster ID. HW replacement is hitless for the VSX secondary.

The **best practice for inter-switch-link timers** (dead-interval, hello-interval, hold-time, peer-detect-interval) is to keep the default timers (i.e. no specific configuration).

The **best practice for role** (primary or secondary) is to have a meaningful relationship with the switch hostname/identification. Example: AGG-1 is VSX primary and AGG-2 is VSX secondary.

AGG-1(config)#	AGG-2(config)#

`vsx`
`  system-mac 02:01:00:00:01:00`
`  inter-switch-link lag 256`
`  role primary`
`  vsx-sync vsx-global`
	`vsx`
	`  inter-switch-link lag 256`
	`  role secondary`

At this stage few aspects can be highlighted:

- **Best practice for vsx-sync** includes vsx-global. Thanks to this vsx-sync FeatureGroup parameter, the VSX management-plane will synchronize the following VSX settings: 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 automatically **tags the native VLAN** configured on the LAG used for ISL.

AGG-1

```
AGG-1# show vsx status
VSX Operational State
---------------------
  ISL channel             : In-Sync
  ISL mgmt channel        : operational
  Config Sync Status      : in-sync
  NAE                     : peer_reachable
  HTTPS Server            : peer_reachable

Attribute           Local               Peer
------------        --------            --------
ISL link            lag256              lag256
ISL version         2                   2
System MAC          02:01:00:00:01:00   02:01:00:00:01:00
Platform            8325                8325
Software Version    GL.10.04.0001       GL.10.04.0001
Device Role         primary             secondary

AGG-1# show running-config | begin 5 vsx
vsx
    system-mac 02:01:00:00:01:00
    inter-switch-link lag 256
    role primary
    vsx-sync vsx-global

AGG-1# show running-config vsx
vsx
    system-mac 02:01:00:00:01:00
    inter-switch-link lag 256
    role primary
interface lag 256
   description ISL link
   no shutdown
   no routing
    vlan trunk native 1 tag
    vlan trunk allowed all
   lacp mode active
interface 1/1/56
   no shutdown
   mtu 9198
   lag 256
interface 1/1/55
   no shutdown
   mtu 9198
   lag 256

AGG-1# show running-config vsx-sync
Current vsx-sync configuration:
!
!Version ArubaOS-CX GL.10.04.0001
!export-password: default
vsx
    system-mac 02:01:00:00:01:00
    vsx-sync vsx-global
```

AGG-2

```
AGG-2# show vsx status
VSX Operational State
---------------------
  ISL channel             : In-Sync
  ISL mgmt channel        : operational
  Config Sync Status      : in-sync
  NAE                     : peer_reachable
  HTTPS Server            : peer_reachable

Attribute           Local               Peer
------------        --------            --------
ISL link            lag256              lag256
ISL version         2                   2
System MAC          02:01:00:00:01:00   02:01:00:00:01:00
Platform            8325                8325
Software Version    GL.10.04.0001       GL.10.04.0001
Device Role         secondary           primary

AGG-2# show running-config | begin 5 vsx
vsx
    system-mac 02:01:00:00:01:00
    inter-switch-link lag 256
    role secondary
    vsx-sync vsx-global

AGG-2# show running-config vsx
vsx
    system-mac 02:01:00:00:01:00
    inter-switch-link lag 256
    role secondary
interface lag 256
   description ISL link
   no shutdown
   no routing
    vlan trunk native 1 tag
    vlan trunk allowed all
   lacp mode active
interface 1/1/56
   no shutdown
   mtu 9198
   lag 256
interface 1/1/55
   no shutdown
   mtu 9198
   lag 256

AGG-2# show running-config vsx-sync
Current vsx-sync configuration:
!
!Version ArubaOS-CX GL.10.04.0001
!export-password: default
vsx
    system-mac 02:01:00:00:01:00
    vsx-sync vsx-global
```

synchronized

synchronized

At this stage, the VSX cluster is created and show command vsx-peer can be used from any VSX node. However, the cluster is not protected yet against a split. (keepalive is not yes established). Step#3 and Step#4 can be merged into a single step (here they are separated for educational purpose).

AGG-1 / AGG-2

```
AGG-1# show vsx brief
```

```
ISL State                              : In-Sync
Device State                           : Peer-Established
Keepalive State                        : Keepalive-Init
Device Role                            : primary
Number of Multi-chassis LAG interfaces : 0

AGG-1# show vsx brief vsx-peer
ISL State                              : In-Sync
Device State                           : Peer-Established
Keepalive State                        : Keepalive-Init
Device Role                            : secondary
Number of Multi-chassis LAG interfaces : 0
```

Step #4 : VSX keepalive

The **best practice for VSX keepalive timers** (dead-interval, hello-interval) is to keep the default timers (i.e. no specific configuration).

AGG-1(config)#	AGG-2(config)#
vsx keepalive peer 192.168.0.1 source 192.168.0.0 vrf KA	vsx keepalive peer 192.168.0.0 source 192.168.0.1 vrf KA

AGG-1 / AGG-2

```
AGG-1# show vsx brief
ISL State                              : In-Sync
Device State                           : Peer-Established
Keepalive State                        : Keepalive-Established
Device Role                            : primary
Number of Multi-chassis LAG interfaces : 0

AGG-1# show vsx brief vsx-peer
ISL State                              : In-Sync
Device State                           : Peer-Established
Keepalive State                        : Keepalive-Established
Device Role                            : secondary
Number of Multi-chassis LAG interfaces : 0

AGG-2# show vsx status keepalive
Keepalive State           : Keepalive-Established
Last Established          : Thu Nov 28 14:40:35 2019
Last Failed               :
Peer System Id            : 02:01:00:00:01:00
Peer Device Role          : primary

Keepalive Counters
Keepalive Packets Tx      : 5220
Keepalive Packets Rx      : 5220
Keepalive Timeouts        : 0
Keepalive Packets Dropped : 0
```

Step #5 : Configuration-sync and vsx-sync FeatureGroup settings

The **best practice for VSX configuration-sync** is to keep the default enabled configuration-synchronization (no configuration change).

AGG-1	AGG-2
AGG-1# show vsx status config-sync Admin state : Enabled Operational State : Operational Error State : None Recommended remediation : N/A Current time : Thu Nov 28 15:54:17 2019 Last sync time : Thu Nov 28 15:40:27 2019	AGG-2# show vsx status config-sync Admin state : Enabled Operational State : Operational Error State : None Recommended remediation : N/A Current time : Thu Nov 28 16:05:31 2019 Last sync time : Thu Nov 28 15:40:27 2019

The **best practice for vsx-sync global settings** is to use as much as possible the automatic synchronization, to avoid human errors. From the list below, the best practice settings are highlighted (in the context of a traditional Aggregation layer for IPv4). According to specific requirements, more or less parameters can be used from the list.

AGG-1(config)#

```
AGG-1(config-vsx)# vsx-sync ?
  aaa                 Sync all AAA instances
  acl-log-timer       Sync access-list log timer instance
```

```
  arp-security          Sync all ARP security configurations
  bfd-global            Sync all BFD global configuration
  bgp                   Sync all BGP, ip aspath list, community list, prefix
                        list, route map configurations
  copp-policy           Sync all CoPP instances
  dcb-global            Sync global configurations for DCB features (DCBx, PFC
                        and ETS)
  dhcp-relay            Sync all DHCP RELAY instances
  dhcp-server           Sync all DHCPv4-Server and DHCPv6-Server instances
  dhcp-snooping         Sync all DHCPv4-Snooping and DHCPv6-Snooping instances.
  dns                   Sync all DNS instances
  evpn                  Sync all evpn configurations
  icmp-tcp              Sync all icmp and tcp instances
  lldp                  Sync all LLDP instances
  loop-protect-global   Sync all Loop-protect global configuration
  mac-lockout           Sync all mac lockout configurations
  mclag-interfaces      Sync QoS, LACP, Loop-Protect, LAG description, sFlow,
                        STP, Rate-Limits, Vlans, ACLs and Portfilters for MCLAG
                        interface instances
  nd-snooping           Sync all ND-Snooping instances.
  neighbor              Sync all IPv4 and IPv6 static neighbor entries
  ospf                  Sync all OSPF instances
  qos-global            Sync all QoS global instances
  route-map             Sync all ip aspath list, community list, prefix list,
                        route map configurations
  sflow-global          Sync all sFlow global instances
  snmp                  Sync all SNMP instances
  ssh                   Sync all SSH instances
  static-routes         Sync all Static Routes instances
  stp-global            Sync all STP Global Configuration
  time                  Sync all time instances
  udp-forwarder         Sync all UDP FORWARDER instances
  vrrp                  Sync all VRRP instances
  vsx-global            Sync all VSX global configuration
```

In addition of the current vsx-global feature synchronization, the other settings are added to the VSX primary. No configuration for this synchronization features is required on VSX secondary switch which will inherit from VSX primary.

AGG-1(config)#	AGG-2(config)#
vsx vsx-sync aaa acl-log-timer bfd-global bgp copp-policy dhcp-relay dhcp-server dhcp-snooping dns icmp-tcp lldp loop-protect-global mac-lockout mclag-interfaces neighbor ospf qos-global route-map sflow-global snmp ssh stp-global time vsx-global	synchronized

AGG-1 / AGG-2

```
AGG-1# show running-config vsx-sync
Current vsx-sync configuration:
!
!Version ArubaOS-CX GL.10.04.0001
!export-password: default
!
ssh server vrf mgmt
!
vsx
    system-mac 02:01:00:00:01:00
    vsx-sync aaa acl-log-timer bfd-global bgp copp-policy dhcp-relay dhcp-server dhcp-snooping dns lldp loop-protect-
global mclag-interfaces ospf qos-global route-map sflow-global snmp ssh stp-global time vsx-global

AGG-2# show running-config | begin 0 vsx
vsx
    vsx-sync aaa acl-log-timer bfd-global bgp copp-policy dhcp-relay dhcp-server dhcp-snooping dns icmp-tcp lldp loop-
protect-global mac-lockout mclag-interfaces neighbor ospf qos-global route-map sflow-global snmp ssh stp-global time
vsx-global
```

Step #6 : VSX split-recovery

The best practice for VSX split-recovery is to keep the default split-recovery enabled (no configuration change). This best practice might be revisited in case of VSX and VXLAN.

AGG-1	AGG-2
AGG-1# show vsx configuration split-recovery Split Recovery Mode : Enabled	**AGG-2**# show vsx configuration split-recovery Split Recovery Mode : Enabled

Step #7 : VSX linkup-delay-timer

The best practice for VSX linkup-delay-timer for mid-size network (<10k MAC/ARP) is to keep the default timer.

AGG-1 / AGG-2 (mid-size network)

```
AGG-1# show vsx status linkup-delay
Configured linkup delay-timer                                : 180 seconds
Initial sync status                                          : Completed
Delay timer status                                           : Completed
Linkup Delay time left                                       :
Interfaces that will be brought up after delay timer expires :
Interfaces that are excluded from delay timer                :

AGG-2# show vsx status linkup-delay
Configured linkup delay-timer                                : 180 seconds
Initial sync status                                          : Completed
Delay timer status                                           : Completed
Linkup Delay time left                                       :
Interfaces that will be brought up after delay timer expires :
Interfaces that are excluded from delay timer                :
```

The best practice for VSX linkup-delay-timer for large-size network (>10k MAC/ARP) is to set the linkup-delay-timer to the maximum value: 600 (i.e. 600 seconds). VSX will auto-adapt the actual timer based on the completion of tables exchanges and ASICs readiness, so that the time to wait for VSX LAG links being activated is less than or equal to the maximum timer being set.

AGG-1(config)# (large-size network)	AGG-2(config)# (large-size network)
AGG-1(config)# vsx AGG-1(config-vsx)# linkup-delay-timer 600	synchronized

AGG-1 / AGG-2 (large-size network)

```
AGG-1# show running-config vsx
vsx
    system-mac 02:01:00:00:01:00
    inter-switch-link lag 256
    role primary
    keepalive peer 192.168.0.1 source 192.168.0.0 vrf KA
    linkup-delay-timer 600

AGG-2# show run vsx
vsx
    system-mac 02:01:00:00:01:00
    inter-switch-link lag 256
    role secondary
    keepalive peer 192.168.0.0 source 192.168.0.1 vrf KA
    linkup-delay-timer 600

AGG-1# show vsx status linkup-delay
Configured linkup delay-timer                                : 600 seconds
Initial sync status                                          : Completed
Delay timer status                                           : Completed
Linkup Delay time left                                       :
Interfaces that will be brought up after delay timer expires :
Interfaces that are excluded from delay timer                :

AGG-2# show vsx status linkup-delay
Configured linkup delay-timer                                : 600 seconds
Initial sync status                                          : Completed
Delay timer status                                           : Completed
Linkup Delay time left                                       :
Interfaces that will be brought up after delay timer expires :
Interfaces that are excluded from delay timer                :
```

Step #8 : VLANs configuration

The **best practice for VLANs configuration** is to configure the VLANs on the VSX primary with the **vsx-sync** attribute and let the VSX config-sync automatically synchronize the VLANs on the VSX secondary. Here, VLAN 10, 20 to 30 are the endpoints VLANs.

AGG-1(config)#

```
AGG-1(config)# vlan 10,20-30
AGG-1(config-vlan-<10,20-30>)# vsx-sync

AGG-1# show run
Current configuration:
.................................skipped for readability
vlan 10
    vsx-sync
vlan 20
    vsx-sync
.................................skipped for readability
vlan 29
    vsx-sync
vlan 30
    vsx-sync
```

AGG-2(config)#

synchronized

```
AGG-2# show run
Current configuration:
.................................skipped for readability
vlan 10
    vsx-sync
vlan 20
    vsx-sync
.................................skipped for readability
vlan 29
    vsx-sync
vlan 30
    vsx-sync
```

AGG-1 / AGG-2

```
AGG-2# show vlan

-------------------------------------------------------------------------------------------------------------
VLAN  Name                                Status  Reason               Type      Interfaces
-------------------------------------------------------------------------------------------------------------
1     DEFAULT_VLAN_1                      up      ok                   default   lag256
10    VLAN10                              up      ok                   static    lag256
20    VLAN20                              up      ok                   static    lag256
.....................................................................................................skipped for readability
29    VLAN29                              up      ok                   static    lag256
30    VLAN30                              up      ok                   static    lag256
```

Note: if vsx-sync attribute is removed from the configuration element on the VSX primary the configuration item will stay on the VSX secondary without the vsx-sync keyword. Consequently, if an item is then removed from the VSX primary, it will stay on the VSX secondary.

Step #9 : Downstream VSX LAG (MCLAG) configuration

In this section, for simplicity, it is assumed that the connected Access Switch is already configured with uplinks link-aggregation and trunked VLANs. The **best practice for VSX LAG** is to create the multi-chassis lag interface on the VSX primary with all settings and then create the mirrored lag interface on the VSX secondary. LAG interface settings (including description) will be synchronized automatically. Only "no shut" in the lag interface context has to be performed on the VSX secondary. Once the multi-chassis lag interface is created, it can be assigned to the physical port.

The **best practice for allowed VLANs** is to exclude the native VLAN 1 from being propagated. This is a very robust method to avoid Layer2 storm propagation due to potential loop initiated on an access switch. In case of access switch Zero-Touch-Provisioning use-case., this trunking exclusion is performed after the ZTP process.

The **best practice for LAG numbering** is to use LAG ID=1 for connecting the Access Switch#1 on port 1/1/1, LAG 2 used to connect a second Access Switch on port 1/1/2, and so on…

The **best practice for LACP timers** on the VSX LAG is to keep the default long timer (30s = lacp rate slow).

The **best practice for MTU** is to configure on all devices the appropriate size to support features such as Dynamic Segmentation or server jumbo frame. Care should be taken to ensure that the IP path from the access devices (switches or APs) can provide a MTU of at least 1564 bytes to the mobility controllers and that the server jumbo packet of 9000 bytes can be encapsulated. Flexibility should be anticipated to perform VXLAN encapsulation from the access switch (9000+50) or VXLAN encapsulation from the aggregation layer MTU+50. So the

recommended Ethernet MTU is 9100 bytes for the downstream VSX LAG to the access layer and a MTU of 9000 bytes for endpoints or servers. The SVI IP MTU should match the MTU size on the aggregation layer, so the recommended IP MTU is 9100 bytes.

The **best practice for hashing algorithm** on the VSX LAG is to keep the default l3-src-dst (alternative being l2-src-dst). This option has an effect only if at least 2 ports per VSX node are members of the same VSX LAG.

Note: Most of the time the VSX LAG includes only two links: one link from the primary and one link from the secondary. Consequently, hashing algorithm selection has no effect on the traffic path as it is forwarded to the local port of the VSX LAG on the switch receiving the traffic.

AGG-1(config)#

```
interface lag 1 multi-chassis
    description Access-Switch-1 VSX LAG
    no shutdown
    vlan trunk allowed 10,20-30

interface 1/1/1
    no shutdown
    mtu 9100
    description ACC-1
    lag 1
```

AGG-2(config)#

```
interface lag 1 multi-chassis
    no shutdown

interface 1/1/1
    no shutdown
    mtu 9100
    description ACC-1
    lag 1
```

synchronized

AGG-1 / AGG-2

```
AGG-1# show 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  Port    Port     State   System-ID         System   Aggr
        name       id      Priority                           Priority Key
------------------------------------------------------------------------------
1/1/1   lag1(mc)    1       1        ALFNCD  02:01:00:00:01:00 65534    1


Partner details of all interfaces:
------------------------------------------------------------------------------
Intf    Aggregate  Partner Port     State   System-ID         System   Aggr
        name       Port-id Priority                           Priority Key
------------------------------------------------------------------------------
1/1/1   lag1(mc)   49      0        ALFNCD  ec:eb:b8:d0:51:00 20736    532

Remote Actor details of all interfaces:
------------------------------------------------------------------------------
Intf   Aggregate  Port    Port     State   System-ID         System   Aggr
       name       id      Priority                           Priority Key
------------------------------------------------------------------------------
1/1/1  lag1(mc)   1001    1        ALFNCD  02:01:00:00:01:00 65534    1


Remote Partner details of all interfaces:
------------------------------------------------------------------------------
Intf   Aggregate  Partner Port     State   System-ID         System   Aggr
       name       Port-id Priority                           Priority Key
------------------------------------------------------------------------------
1/1/1  lag1(mc)   50      0        ALFNCD  ec:eb:b8:d0:51:00 20736    532
```

The "show lacp interfaces multi-chassis" command is very useful to get a complete status of the local LACP partnership as well as the VSX peer partnership details. Actor = local node, Partner = LACP neighbor (the access switch), Remote Actor = the VSX peer, Remote Partner = LACP neighbor of the VSX peer. Note that the port id of the VSX secondary is equal to 1000+ID_of_the_primary (in the example 1001). ALFNCD LACP state-flags should appear on all entries.

The **best practice for LACP fallback** feature is to enable it on the VSX LAGs for the following use-cases: PXE boot, access switch ZTP, server NIC driver migration from active/standby to LACP. When applied to the VSX primary, LACP fallback is automatically synced on the VSX secondary.

AGG-1(config)#	AGG-2(config)#
`interface lag 1 multi-chassis` `    lacp fallback`	synchronized

`AGG-1# show run`
`Current configuration:`
`.....skipped for readability`
`interface lag 1 multi-chassis`
`    no shutdown`
`    description Access-Switch-1 VSX LAG`
`    no routing`
`    vlan trunk native 1`
`    vlan trunk allowed 10,20-30`
`    lacp mode active`
`    lacp fallback`
	`AGG-2# show run`
	`Current configuration:`
	`.....skipped for readability`
	`interface lag 1 multi-chassis`
	`    no shutdown`
	`    description Access-Switch-1 VSX LAG`
	`    no routing`
	`    vlan trunk native 1`
	`    vlan trunk allowed 10,20-30`
	`    lacp mode active`
	`    lacp fallback`

Further on in this document, lacp fallback is no longer shown as this is reserved for the previous indicated use-cases.

Step #10 : MSTP configuration

The **best practice on Aggregation** layer are:

- No loop-protect (MSTP used instead).
- Use the **default common instance 0**: MST0
- Lower the **spanning-tree priority to 4** to make VSX aggregation the STP root bridge (easier for support)
- Use **root-guard** on all downlinks to prevent any access switches from becoming Root Bridge.
- Keep the default **port-type admin-network**
- Let VSX secondary synchronized by vsx-sync process.

AGG-1(config)#	AGG-2(config)#

`spanning-tree`
`spanning-tree priority 4`
`spanning-tree trap topology-change instance 0`
	synchronized

`interface lag 1 multi-chassis`
`    no shutdown`
`    description Access-Switch-1 VSX LAG`
`    no routing`
`    vlan trunk native 1`
`    vlan trunk allowed 10,20-30`
`    lacp mode active`
`    spanning-tree root-guard`
	synchronized

`AGG-1# show run`
`Current configuration:`
`.....skipped for readability`
`spanning-tree`
`spanning-tree priority 4`
`spanning-tree trap topology-change instance 0`
`.....skipped for readability`
`interface lag 1 multi-chassis`
`    no shutdown`
`    description Access-Switch-1 VSX LAG`
`    no routing`
`    vlan trunk native 1`
`    vlan trunk allowed 10,20-30`
`    lacp mode active`
`    spanning-tree root-guard`
`.....skipped for readability`
	`AGG-2# show run`
	`Current configuration:`
	`.....skipped for readability`
	`spanning-tree`
	`spanning-tree priority 4`
	`spanning-tree trap topology-change instance 0`
	`.....skipped for readability`
	`interface lag 1 multi-chassis`
	`    no shutdown`
	`    description Access-Switch-1 VSX LAG`
	`    no routing`
	`    vlan trunk native 1`
	`    vlan trunk allowed 10,20-30`
	`    lacp mode active`
	`    spanning-tree root-guard`
	`.....skipped for readability`

AGG-1 / AGG-2	

```
AGG-1# show spanning-tree mst 0 int lag1 detail
Port lag1
Port Type : admin-network    Loop Guard  : disable
Link Type : point_to_point   BPDU Filter : disable
Boundary  : internal         BPDU Guard  : disable
Root Guard: enable

Instance       Role           State        Cost       Priority   Vlans mapped
-------------- -------------- ------------ ---------- ---------- ----------
0              Designated     Forwarding   20000      64         1-4094

Port lag1
Designated root address            : 02:01:00:00:01:00
Designated regional root address   : 02:01:00:00:01:00
Designated bridge address          : 02:01:00:00:01:00  Priority : 16384
Multi-Chassis role                 : active
Timers:    Message expires in 0 sec, Forward delay expiry:18, Forward transitions:1
Bpdus sent 19295, received 2
TCN_Tx: 4, TCN_Rx: 2
```

The **best practice on Access** layer are:

- Use loop-protect for all endpoint access ports (not configured on uplinks). Set the re-enable timer to 1hour.
- Keep the **default common instance 0**: MST0
- Keep the **default spanning-tree priority** of 8.
- All endpoint access ports are **admin-edge**, should not receive any BPDU (**BDPU guard**), should not trigger any Topology Change Notification (**tcn-guard**).
- Use loop-protection on all endpoint access ports as an extra-protection mechanism (in case of MSTP BPDUs are filtered by insertion of unmanaged switches which create a loop).
- Use **loop-guard** on all uplinks to prevent any flood due to failure of BPDU reception (fiber strand cut).

ACC-1(config)#

```
spanning-tree
loop-protect re-enable-timer 3600
interface lag 1
    no shutdown
    description UPLINK to AGG
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20
    lacp mode active
    spanning-tree loop-guard
interface 1/1/1
    no shutdown
    description Endpoint1
    no routing
    vlan access 10
    spanning-tree bpdu-guard
    spanning-tree port-type admin-edge
    spanning-tree tcn-guard
    loop-protect

ACC-1# show spanning-tree mst 0 interface lag1 detail
Port lag1
Port Type : admin-network    Loop Guard  : enable
Link Type : point_to_point   BPDU Filter : disable
Boundary  : internal         BPDU Guard  : disable
Root Guard: disable

Instance       Role           State        Cost       Priority   Vlans mapped
-------------- -------------- ------------ ---------- ---------- ----------
0              Root           Forwarding   20000      64         1-4094

Port lag1
```

```
Designated root address             : 02:01:00:00:01:00
Designated regional root address    : 08:00:09:72:61:c6
Designated bridge address           : 02:01:00:00:01:00  Priority : 32768
Timers:    Message expires in 0 sec, Forward delay expiry:1, Forward transitions:9
Bpdus sent 31, received 145466
TCN_Tx: 18, TCN_Rx: 20

ACC-1# show spanning-tree mst 0 interface 1/1/1 detail
Port 1/1/1
Port Type : admin-edge        Loop Guard  : disable
Link Type : point_to_point    BPDU Filter : disable
Boundary  : internal          BPDU Guard  : enable
Root Guard: disable

Instance       Role           State        Cost       Priority   Vlans mapped
-------------- -------------- ------------ ---------- ---------- ----------
0              Designated     Forwarding   20000      128        1-4094

Port 1/1/1
Designated root address             : 02:01:00:00:01:00
Designated regional root address    : 08:00:09:72:61:c6
Designated bridge address           : 02:01:00:00:01:00  Priority : 32768
Timers:    Message expires in 0 sec, Forward delay expiry:1, Forward transitions:1
Bpdus sent 160124, received 0
TCN_Tx: 0, TCN_Rx: 0
```

Step #11 : VSX LAG ACL configuration

If any ACL is used, the best practice is to have ACLs synchronized on secondary through vsx-sync. Any ACL applied on a VSX LAG on the VSX primary will get applied on the VSX secondary as well.

AGG-1(config)#

```
access-list ip IOT-1
    vsx-sync
    !
    5 permit tcp 10.1.0.0/255.255.0.0
10.99.1.0/255.255.255.0 eq 1080
    10 deny any 10.1.0.0/255.255.0.0
10.99.1.0/255.255.255.0
    1000 permit any any any

interface lag 1 multi-chassis
    apply access-list ip IOT-1 in

AGG-1# show run
Current configuration:
.................................skipped for readability
access-list ip IOT-1
    vsx-sync
    !
    5 permit tcp 10.1.0.0/255.255.0.0
10.99.1.0/255.255.255.0 eq 1080
    10 deny any 10.1.0.0/255.255.0.0
10.99.1.0/255.255.255.0
    1000 permit any any any
.................................skipped for readability
interface lag 1 multi-chassis
    no shutdown
    description Access-Switch-1 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20-30
    lacp mode active
    spanning-tree root-guard
    apply access-list ip IOT-1 in
.................................skipped for readability
```

AGG-2(config)#

synchronized

synchronized

```
AGG-2# show run
Current configuration:
.................................skipped for readability
access-list ip IOT-1
    vsx-sync
    !
    5 permit tcp 10.1.0.0/255.255.0.0
10.99.1.0/255.255.255.0 eq 1080
    10 deny any 10.1.0.0/255.255.0.0
10.99.1.0/255.255.255.0
    1000 permit any any any
.................................skipped for readability
interface lag 1 multi-chassis
    no shutdown
    description Access-Switch-1 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20-30
    lacp mode active
    spanning-tree root-guard
    apply access-list ip IOT-1 in
.................................skipped for readability
```

AGG-1 / AGG-2

```
AGG-1# show access-list commands
access-list ip IOT-1
    vsx-sync
```

```
    !
    5 permit tcp 10.1.0.0/255.255.0.0 10.99.1.0/255.255.255.0 eq 1080
    10 deny any 10.1.0.0/255.255.0.0 10.99.1.0/255.255.255.0
    1000 permit any any any
interface lag 1
    apply access-list ip IOT-1 in
```

Step #12 : VSX LAG QoS configuration

QoS Marking being performed on the access layer, the aggregation switch is configure in the global context with **qos trust dscp**. No further configuration is needed as this was already set in step#1.

Step #13 : SVI (VLAN L3 interface) configuration

The **best practice for SVI active-gateway** is to set the active-gateway Virtual IP and Virtual MAC on the VSX primary and get the value synchronized on the VSX secondary with vsx-sync command.

The **best practice for active-gateway VMAC** is to use the **same VMAC for all IPv4 SVIs**. The scope of this VMAC is purely link-local. If some servers or systems have dual-attachment to two different SVIs, and the system administrator would like to see distinct MAC addresses for the next-hops over these separate interfaces, then 16 VMACs are available. For dual-stack IPv4 and IPv6, 16 VMACs can be used for IPv4 and the same VMACs can be used for IPv6. It is however a best practice to use only 8 VMACs for IPv4 and 8 different VMACs for IPV6.

If mutlinetting is used, set one VIP per secondary subnet and **disable ip icmp redirect**.

The **best practice for IP MTU** is to configure on all SVIs the matching size of the L2 MTU: IP MTU recommended value = 9100. This parameter must be identical and manually set on both VSX nodes.

The **best practice for DHCP relay** is to configure the ip helper-address on the VSX primary and let vsx-sync configuring the same on the VSX secondary.

AGG-1(config)#	AGG-2(config)#
`no ip icmp redirect`	synchronized

`interface vlan10`
`  vsx-sync active-gateways`
`  ip mtu 9100`
`  ip address 10.1.10.2/24`
`  ip address 10.2.10.2/24 secondary`
`  active-gateway ip mac 12:01:00:00:01:00`
`  active-gateway ip 10.1.10.1`
`  active-gateway ip 10.2.10.1`
`  ip helper-address 10.99.10.9`
	`interface vlan10`
	`  ip mtu 9100`
	`  ip address 10.1.10.3/24`
	`  ip address 10.2.10.3/24 secondary`

`interface vlan20`
`  vsx-sync active-gateways`
`  ip mtu 9100`
`  ip address 10.1.20.2/24`
`  ip address 10.2.20.2/24 secondary`
`  active-gateway ip mac 12:01:00:00:01:00`
`  active-gateway ip 10.1.20.1`
`  active-gateway ip 10.2.20.1`
`  ip helper-address 10.99.10.9`
	synchronized
	`interface vlan20`
	`  ip mtu 9100`
	`  ip address 10.1.20.3/24`
	`  ip address 10.2.20.3/24 secondary`

`AGG-1# show run`
`Current configuration:`
`no ip icmp redirect`
`.....skipped for readability`
`interface vlan10`
`  vsx-sync active-gateways`
`  ip mtu 9100`
`  ip address 10.1.10.2/24`
`  ip address 10.2.10.2/24 secondary`
`  active-gateway ip mac 12:01:00:00:01:00`
`  active-gateway ip 10.1.10.1`
`  active-gateway ip 10.2.10.1`
`  ip helper-address 10.99.10.9`
`interface vlan20`
	synchronized
	`AGG-2# show run`
	`Current configuration:`
	`no ip icmp redirect`
	`.....skipped for readability`
	`interface vlan10`
	`  vsx-sync active-gateways`
	`  ip mtu 9100`
	`  ip address 10.1.10.3/24`
	`  ip address 10.2.10.3/24 secondary`
	`  active-gateway ip mac 12:01:00:00:01:00`
	`  active-gateway ip 10.1.10.1`
	`  active-gateway ip 10.2.10.1`
	`  ip helper-address 10.99.10.9`
	`interface vlan20`

```
    vsx-sync active-gateways
    ip mtu 9100
    ip address 10.1.20.2/24
    ip address 10.2.20.2/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.20.1
    active-gateway ip 10.2.20.1
    ip helper-address 10.99.10.9
................................skipped for readability
```

```
    vsx-sync active-gateways
    ip mtu 9100
    ip address 10.1.20.3/24
    ip address 10.2.20.3/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.20.1
    active-gateway ip 10.2.20.1
    ip helper-address 10.99.10.9
................................skipped for readability
```

AGG-1 / AGG-2

```
AGG-1# show ip interface vlan10

Interface vlan10 is up
 Admin state is up
 Hardware: Ethernet, MAC Address: 08:00:09:94:00:b8
 IP MTU 9100
 IPv4 address 10.1.10.2/24
 IPv4 address 10.2.10.2/24 secondary
 active-gateway ip mac 12:01:00:00:01:00
 active-gateway ip 10.1.10.1
 active-gateway ip 10.2.10.1
 L3 Counters: Rx Disabled, Tx Disabled
 Rx
          ucast: 0 packets, 0 bytes
          mcast: 0 packets, 0 bytes
 Tx
          ucast: 0 packets, 0 bytes
          mcast: 0 packets, 0 bytes

AGG-1# show ip interface vlan10 vsx-peer

Interface vlan10 is up
 Admin state is up
 Hardware: Ethernet, MAC Address: 08:00:09:b5:e8:2d
 IP MTU 9100
 IPv4 address 10.1.10.3/24
 IPv4 address 10.2.10.3/24 secondary
 active-gateway ip mac 12:01:00:00:01:00
 active-gateway ip 10.1.10.1
 active-gateway ip 10.2.10.1
 L3 Counters: Rx Disabled, Tx Disabled
 Rx
          ucast: 0 packets, 0 bytes
          mcast: 0 packets, 0 bytes
 Tx
          ucast: 0 packets, 0 bytes
          mcast: 0 packets, 0 bytes
```

Step #14 : OSPF configuration

It is a **best practice to create a dedicated Transit VLAN** between the VSX primary and the VSX secondary to exchange routes information for subnets that are not attached to both VSX nodes (ex: loopback addresses of each VSX node). This dedicated Transit VLAN (here VLAN 2) provides better control and will not carry user data traffic in nominal situation or very limited in case of east-west traffic between single-attached endpoints.

There are two strategies to inject endpoint subnets into the routing table: either through OSPF or through BGP

- OSPF: Most of the Campus deployments use OSPF to exchange route information for end-devices. This is simple and can scale very well with appropriate usage of areas. This is the target of this current document.
- BGP: Lot of new DC deployment use BGP as a routing protocol due to the usage of EVPN based VXLAN. Such a design is coming in the Campus as well. Also, for more complex and granular routing engineering, BGP communities and route-map can offer a level of control that OSPF can not provide. This can be exposed in a future white paper.

There are two options to inject end-user subnets into OSPF DataBase: using ospf command on the SVI (VLAN L3 interface), or redistributing the connected into OSPF with route-map control. The **best practice is to use the ospf command on SVI** as offering a simpler configuration like for the area the subnets belongs to. This principle is selected as the OSPF best practice in the following described configuration. More details on OSPF best practices can be found on IP routing configuration guide.

The **best practice for point-to-point interconnectivity subnet** is to use /31 subnet.

The **best practice for OSPF configuration** is to use vsx-sync ospf synchronization option and have OSPF parameters automatically synced on the VSX secondary. As shown on the configuration step, very few elements have to be configured on the secondary. Pay attention that to get such a benefit, the interface ID should be mirrored; i.e. if interface 1/1/49 is used for uplink on the VSX primary, it is strongly recommended to use the same ID 1/1/49 on the VSX secondary, otherwise OSPF synchronization will not synchronized the proper interface.

The **best practice for OSPF cost** is to have VSX primary <-> VSX secondary cost lower than Core-1 <-> Core-2 cost, as it is frequent that the ISL bandwidth is higher than the inter core devices bandwidth. In case of single-attachment subnet on one of the VSX node and non-meshed topology, the traffic from core would be sent to the VSX peer closest to the attached destination, avoiding consuming inter-core bandwidth. Same concept applies for south-to north traffic pattern. In the below example, OSPF cost for Transit VLAN over ISL is set to 50, and 1000 for Core devices. OSPF cost is synchronized from the VSX primary to the VSX secondary, so the importance to use mirrored interface ID.

AGG-1(config)#	AGG-2(config)#
router ospf 1 router-id 10.0.1.1 max-metric router-lsa on-startup passive-interface default graceful-restart restart-interval 300 trap-enable area 0	router ospf 1 router-id 10.0.1.2 synchronized
interface loopback 0 ip address 10.0.1.1/32 ip ospf 1 area 0	interface loopback 0 ip address 10.0.1.2/32
vlan 2 vsx-sync description TRANSIT VLAN	synchronized
interface vlan2 ip address 10.0.2.1/30 ip ospf 1 area 0 no ip ospf passive ip ospf cost 50 ip ospf network point-to-point ip ospf authentication message-digest ip ospf message-digest-key 1 md5 plaintext *yourpass*	interface vlan2 ip address 10.0.2.2/30 synchronized
interface vlan10 ip ospf 1 area 0.0.0.0	
interface vlan20 ip ospf 1 area 0.0.0.0	
interface 1/1/49 no shutdown mtu 9198 description CORE-1 1/1/1 ip mtu 9198 ip address 10.0.0.1/31 ip ospf 1 area 0.0.0.0 no ip ospf passive ip ospf cost 1000 ip ospf network point-to-point ip ospf authentication message-digest ip ospf message-digest-key 1 md5 plaintext *yourpass*	interface 1/1/49 no shutdown mtu 9198 description CORE-1 1/1/2 ip mtu 9198 ip address 10.0.0.5/31 synchronized
interface 1/1/50 no shutdown mtu 9198 description CORE-2 1/1/1 ip mtu 9198 ip address 10.0.0.3/31 ip ospf 1 area 0.0.0.0	interface 1/1/50 no shutdown mtu 9198 description CORE-2 1/1/2 ip mtu 9198 ip address 10.0.0.7/31

```
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 plaintext yourpass

AGG-1# show lldp neighbor-info

LLDP Neighbor Information
=========================

Total Neighbor Entries          : 6
Total Neighbor Entries Deleted  : 0
Total Neighbor Entries Dropped  : 0
Total Neighbor Entries Aged-Out : 0

LOCAL-PORT  CHASSIS-ID         PORT-ID       PORT-DESC
TTL      SYS-NAME
-------------------------------------------------------------
---------------------
1/1/1       08:00:09:72:61:c6  1/1/49        1/1/49
120      ACC-1
1/1/48      08:00:09:9a:9b:10  1/1/48        1/1/48
120      AGG-2
1/1/49      08:00:09:ac:6e:b7  1/1/1         1/1/1
120      CORE-1
1/1/50      08:00:09:4a:f4:ad  1/1/1         1/1/1
120      CORE-2
1/1/55      08:00:09:9a:9b:10  1/1/55        1/1/55
120      AGG-2
1/1/56      08:00:09:9a:9b:10  1/1/56        1/1/56
120      AGG-2

AGG-1# show run
Current configuration:
.................................skipped for readability
router ospf 1
    router-id 10.0.1.1
    max-metric router-lsa on-startup
    passive-interface default
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
.................................skipped for readability
vlan 2
    vsx-sync
    description TRANSIT VLAN
.................................skipped for readability
interface 1/1/49
    no shutdown
    mtu 9198
    description CORE-1 1/1/1
    ip mtu 9198
    ip address 10.0.0.1/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext
AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbL
X
interface 1/1/50
    no shutdown
    mtu 9198
    description CORE-2 1/1/1
    ip mtu 9198
    ip address 10.0.0.3/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
```

synchronized

```
AGG-2# show  lldp neighbor-info

LLDP Neighbor Information
=========================

Total Neighbor Entries          : 6
Total Neighbor Entries Deleted  : 1
Total Neighbor Entries Dropped  : 0
Total Neighbor Entries Aged-Out : 1

LOCAL-PORT  CHASSIS-ID         PORT-ID       PORT-DESC
TTL      SYS-NAME
-------------------------------------------------------------
----------------------
1/1/1       08:00:09:72:61:c6  1/1/50        1/1/50
120      ACC-1
1/1/48      08:00:09:b0:c4:aa  1/1/48        1/1/48
120      AGG-1
1/1/49      08:00:09:ac:6e:b7  1/1/2         1/1/2
120      CORE-1
1/1/50      08:00:09:4a:f4:ad  1/1/2         1/1/2
120      CORE-2
1/1/55      08:00:09:b0:c4:aa  1/1/55        1/1/55
120      AGG-1
1/1/56      08:00:09:b0:c4:aa  1/1/56        1/1/56
120      AGG-1

AGG-2# show run
Current configuration:
.................................skipped for readability
router ospf 1
    router-id 10.0.1.2
    max-metric router-lsa on-startup
    passive-interface default
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
.................................skipped for readability
vlan 2
    vsx-sync
    description TRANSIT VLAN
.................................skipped for readability
interface 1/1/49
    no shutdown
    mtu 9198
    description CORE-1 1/1/2
    ip mtu 9198
    ip address 10.0.0.5/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext
AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1U
bLX
interface 1/1/50
    no shutdown
    mtu 9198
    description CORE-2 1/1/2
    ip mtu 9198
    ip address 10.0.0.7/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
```

```
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext
AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbL
X
.................................skipped for readability
interface vlan2
    ip mtu 9198
    ip address 10.0.2.1/30
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext
AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbL
X
interface vlan10
    vsx-sync active-gateways
    ip mtu 9100
    ip address 10.1.10.2/24
    ip address 10.2.10.2/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.10.1
    active-gateway ip 10.2.10.1
    ip helper-address 10.99.10.9
    ip ospf 1 area 0.0.0.0
interface vlan20
    vsx-sync active-gateways
    ip mtu 9100
    ip address 10.1.20.2/24
    ip address 10.2.20.2/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.20.1
    active-gateway ip 10.2.20.1
    ip ospf 1 area 0.0.0.0
```

```
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext
AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1U
bLX
.................................skipped for readability
interface vlan2
    ip mtu 9198
    ip address 10.0.2.1/30
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext
AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1U
bLX
interface vlan10
    vsx-sync active-gateways
    ip mtu 9100
    ip address 10.1.10.2/24
    ip address 10.2.10.2/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.10.1
    active-gateway ip 10.2.10.1
    ip helper-address 10.99.10.9
    ip ospf 1 area 0.0.0.0
interface vlan20
    vsx-sync active-gateways
    ip mtu 9100
    ip address 10.1.20.2/24
    ip address 10.2.20.2/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.20.1
    active-gateway ip 10.2.20.1
    ip ospf 1 area 0.0.0.0
```

AGG-1 / AGG-2

```
AGG-1# show ip ospf neighbors
OSPF Process ID 1 VRF default
==============================

Total Number of Neighbors: 3

Neighbor ID      Priority  State             Nbr Address       Interface
-------------------------------------------------------------------------
10.0.1.3         n/a       FULL              10.0.0.0           1/1/49
10.0.1.4         n/a       FULL              10.0.0.2           1/1/50
10.0.1.2         n/a       FULL              10.0.2.2           vlan2

AGG-1# show ip route

Displaying ipv4 routes selected for forwarding

'[x/y]' denotes [distance/metric]

0.0.0.0/0, vrf default
        via  10.0.0.2,  [110/1],  ospf
        via  10.0.0.0,  [110/1],  ospf
10.0.0.0/31, vrf default
        via  1/1/49,  [0/0],  connected
10.0.0.2/31, vrf default
        via  1/1/50,  [0/0],  connected
10.0.0.252/31, vrf default
        via  10.0.0.2,  [110/2000],  ospf
        via  10.0.0.0,  [110/2000],  ospf
10.0.0.6/31, vrf default
        via  10.0.2.2,  [110/1050],  ospf
10.0.0.4/31, vrf default
        via  10.0.2.2,  [110/1050],  ospf
10.0.0.1/32, vrf default
        via  1/1/49,  [0/0],  local
```

```
10.0.0.3/32, vrf default
        via  1/1/50,  [0/0],  local
10.0.1.1/32, vrf default
        via  loopback0,  [0/0],  local
10.0.1.2/32, vrf default
        via  10.0.2.2,  [110/50],  ospf
10.0.1.4/32, vrf default
        via  10.0.0.2,  [110/1000],  ospf
10.0.1.3/32, vrf default
        via  10.0.0.0,  [110/1000],  ospf
10.0.2.0/30, vrf default
        via  vlan2,  [0/0],  connected
10.0.2.1/32, vrf default
        via  vlan2,  [0/0],  local
10.1.10.0/24, vrf default
        via  vlan10,  [0/0],  connected
10.1.10.2/32, vrf default
        via  vlan10,  [0/0],  local
10.1.20.0/24, vrf default
        via  vlan20,  [0/0],  connected
10.1.20.2/32, vrf default
        via  vlan20,  [0/0],  local
10.2.10.0/24, vrf default
        via  vlan10,  [0/0],  connected
10.2.10.2/32, vrf default
        via  vlan10,  [0/0],  local
10.2.20.0/24, vrf default
        via  vlan20,  [0/0],  connected
10.2.20.2/32, vrf default
        via  vlan20,  [0/0],  local
```

CORE-1 / CORE-2

```
CORE-1# show ip ospf neighbors
OSPF Process ID 1 VRF default
==============================

Total Number of Neighbors: 3

Neighbor ID      Priority  State             Nbr Address       Interface
-------------------------------------------------------------------------
10.0.1.1         n/a       FULL              10.0.0.1           1/1/1
10.0.1.2         n/a       FULL              10.0.0.5           1/1/2
10.0.1.4         n/a       FULL              10.0.0.253         1/1/3

CORE-1# show ip route

Displaying ipv4 routes selected for forwarding

'[x/y]' denotes [distance/metric]

10.0.0.0/31, vrf default
        via  1/1/1,  [0/0],  connected
10.0.0.252/31, vrf default
        via  1/1/3,  [0/0],  connected
10.0.0.4/31, vrf default
        via  1/1/2,  [0/0],  connected
10.0.0.2/31, vrf default
        via  10.0.0.1,  [110/2000],  ospf
        via  10.0.0.253,  [110/2000],  ospf
10.0.0.6/31, vrf default
        via  10.0.0.253,  [110/2000],  ospf
        via  10.0.0.5,  [110/2000],  ospf
10.0.0.4/32, vrf default
        via  1/1/2,  [0/0],  local
10.0.0.0/32, vrf default
        via  1/1/1,  [0/0],  local
10.0.0.252/32, vrf default
        via  1/1/3,  [0/0],  local
10.0.1.3/32, vrf default
        via  loopback0,  [0/0],  local
10.0.1.1/32, vrf default
        via  10.0.0.1,  [110/1000],  ospf
```

```
10.0.1.4/32, vrf default
        via  10.0.0.253,  [110/1000],  ospf
10.0.1.2/32, vrf default
        via  10.0.0.5,  [110/1000],  ospf
10.0.2.0/30, vrf default
        via  10.0.0.1,  [110/1050],  ospf
        via  10.0.0.5,  [110/1050],  ospf
10.1.10.0/24, vrf default
        via  10.0.0.1,  [110/1100],  ospf
        via  10.0.0.5,  [110/1100],  ospf
10.1.20.0/24, vrf default
        via  10.0.0.1,  [110/1100],  ospf
        via  10.0.0.5,  [110/1100],  ospf
10.2.10.0/24, vrf default
        via  10.0.0.1,  [110/1000],  ospf
        via  10.0.0.5,  [110/1000],  ospf
10.2.20.0/24, vrf default
        via  10.0.0.1,  [110/1000],  ospf
        via  10.0.0.5,  [110/1000],  ospf
```

Step #15 : BGP configuration

Please read the IP routing guide.

Note: VSX provides capability to synchronize the full BGP configuration between the VSX primary and the VSX secondary. Most of the BGP configuration of the VSX secondary is the same than on the VSX primary. Except the configuration for iBGP peering between the VSX nodes inside the cluster, or for remote eBGP upstream peers with neighbor IP address being the physical IP address of theL3 point-to-point circuit (ex: here 1/1/50 with 10.0.0.6/31 remote IP address). In such a case, the specific neighbor parameters are excluded from the VSX configuration synchronization with the following command on the VSX primary only: **neighbor <IP_adress> vsx-sync-exclude**

Step #16 : Multicast configuration

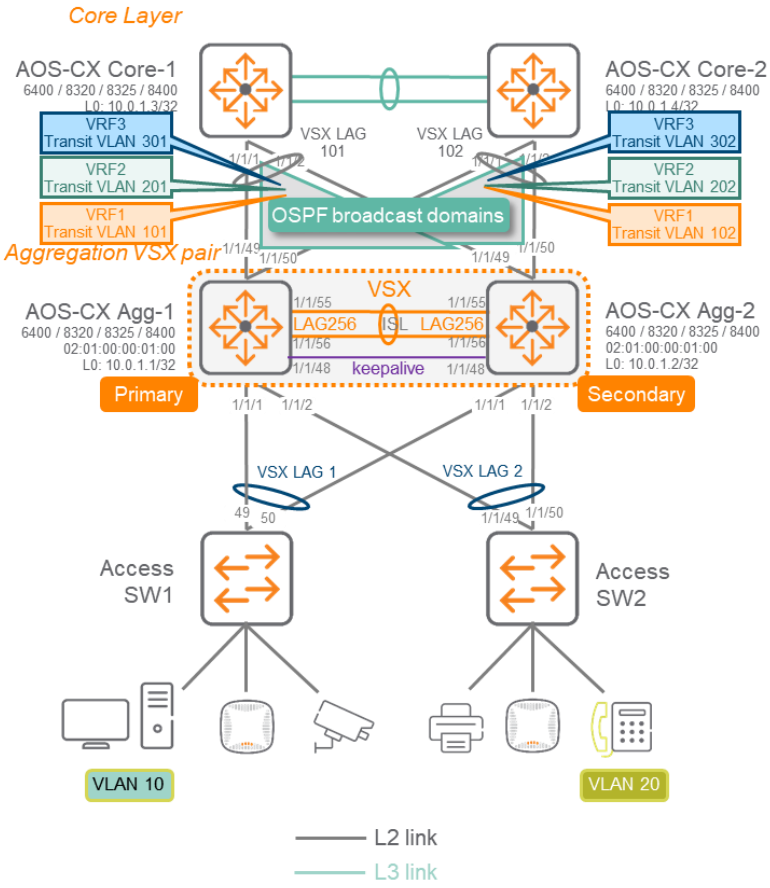
Please read the Multicast guide.

For multicast on VSX cluster, the **best practice is to configure PIM Dual-DR or active/active** under the PIM router command. With active-active command, the proxy-DR will also learn the multicast routes, and will allow fast recovery time if the actual DR fails.

AGG-1(config)#	AGG-2(config)#
router pim enable active-active show ip pim interface vlan10 PIM Interfaces VRF: default Interface : vlan10 IP Address : 10.1.10.2/24 Mode : sparse Designated Router : 10.1.10.2 Proxy DR : false Hello Interval (sec) : 30 Hello Delay (sec) : 5 Override Interval (msec) : 2500 Lan Prune Delay : Yes Propagation Delay (msec) : 500 DR Priority : 200 Neighbor Timeout : 0	router pim enable active-active **show ip pim interface vlan10** PIM Interfaces VRF: default Interface : vlan10 IP Address : 10.1.10.3/24 Mode : sparse Designated Router : 10.1.10.3 Proxy DR : true Hello Interval (sec) : 30 Hello Delay (sec) : 5 Override Interval (msec) : 2500 Lan Prune Delay : Yes Propagation Delay (msec) : 500 DR Priority : 1 Neighbor Timeout : 0

Aggregation VSX with multiple VRF routing model

Here is the typical topology:



It is assumed that VRFs are already created in the configuration. VRF names must be meaningful for the operational support. In the document, VRF1, VRF2, VRF3 are used for obvious generic purpose. VRFs are not synchronized by vsx-sync process, so they have to be created manually on both VSX primary and VSX secondary.

Step #0 to Step#9: follow same steps than for the single VRF scenario

Step #10 : Configure VRF Transit VLANs

The **best practice for upstream routing domain connectivity** is to create one Transit VLAN per VRF per upstream VSX LAG. Consequently, according to the proposed topology, for VRF1, VLAN 101 is created and configured on upstream VSX LAG connecting CORE-1 (VSX LAG 101), VLAN 102 is created on upstream VSX LAG connecting CORE-2 (VSX LAG 102). Similarly, VLAN 201 and 202 are proposed for VRF2 and VLAN 301/302 for VRF3.

Note: for simplicity and readability, only 2 VRFs are documented: VRF1 and VRF2. VRF3 is not included in the configuration to alleviate the document. Configuration for VRF3 or any additional VRFs is similar to VRF1 and VRF2.

These Transit VLANs 101/102 and 201/202 will also serve for VSX intra-cluster routing (unlike the single VRF scenario where a dedicated Transit VLAN (VLAN 2) was created for Transit point-to-point routing between VSX nodes.

AGG-1(config)#	AGG-2(config)#
`AGG-1(config)# vlan 101`	

```
AGG-1(config-vlan-101)# description TRANSIT VLAN VRF1-
CORE1
AGG-1(config-vlan-101)# vsx-sync
AGG-1(config-vlan-101)# vlan 102
AGG-1(config-vlan-102)# description TRANSIT VLAN VRF1-
CORE2
AGG-1(config-vlan-102)# vsx-sync
AGG-1(config-vlan-102)# vlan 201
AGG-1(config-vlan-201)# description TRANSIT VLAN VRF2-
CORE1
AGG-1(config-vlan-201)# vsx-sync
AGG-1(config-vlan-201)# vlan 202
AGG-1(config-vlan-202)# description TRANSIT VLAN VRF2-
CORE2
AGG-1(config-vlan-202)# vsx-sync
```

synchronized

AGG-1

```
AGG-1# show run
Current configuration:
.................................skipped for readability
vlan 10
    vsx-sync
vlan 20
    vsx-sync
.................................skipped for readability
vlan 29
    vsx-sync
vlan 30
    vsx-sync
vlan 101
    vsx-sync
    description TRANSIT VLAN VRF1-CORE1
vlan 102
    vsx-sync
    description TRANSIT VLAN VRF1-CORE2
vlan 201
    vsx-sync
    description TRANSIT VLAN VRF2-CORE1
vlan 202
    vsx-sync
    description TRANSIT VLAN VRF2-CORE2
```

AGG-2

```
AGG-2# show run
Current configuration:
.................................skipped for readability
vlan 10
    vsx-sync
vlan 20
    vsx-sync
.................................skipped for readability
vlan 29
    vsx-sync
vlan 30
    vsx-sync
vlan 101
    vsx-sync
    description TRANSIT VLAN VRF1-CORE1
vlan 102
    vsx-sync
    description TRANSIT VLAN VRF1-CORE2
vlan 201
    vsx-sync
    description TRANSIT VLAN VRF2-CORE1
vlan 202
    vsx-sync
    description TRANSIT VLAN VRF2-CORE2
```

AGG-1 / AGG-2

```
AGG-2# show vlan

-------------------------------------------------------------------------------------------------------------
VLAN  Name                                Status  Reason               Type      Interfaces
-------------------------------------------------------------------------------------------------------------
1     DEFAULT_VLAN_1                      up      ok                   default   lag256
10    VLAN10                              up      ok                   static    lag1,lag256
20    VLAN20                              up      ok                   static    lag1,lag256
.....................................................................................skipped for readability
29    VLAN29                              up      ok                   static    lag1,lag256
30    VLAN30                              up      ok                   static    lag1,lag256
101   VLAN101                             up      ok                   static    lag256
102   VLAN102                             up      ok                   static    lag256
201   VLAN201                             up      ok                   static    lag256
202   VLAN202                             up      ok                   static    lag256
```

Step #11 : Upstream VSX LAG Configuration (VSX LAG 101/102)

In this section, for simplicity, it is assumed that the Core equipment are already configured with link-aggregation and trunked Transit VLANs. The best practice for upstream VSX LAG is to create the multi-chassis lag interface on the VSX primary with all settings and then create the mirrored lag interface on the VSX secondary. LAG interface settings (including description) will be synchronized automatically. Only "no shut" in the lag interface has to be performed on the VSX secondary. Once multi-chassis lag interface is created it is assigned to the physical port.

The best practice for allowed VLANs is to exclude the native VLAN 1 from being propagated and to allow only the Transit VLANs corresponding to the facing Core device: i.e. VSX LAG.101 permitting VLAN 101 and 201, VSX LAG 102 permitting VLAN 102 and 202.

The **best practice example for LAG numbering** is to use LAG ID=101 for connecting the Core-1, LAG ID=102 for connecting Core-2. Any other numbering practice is possible as long as it does not introduce confusion with the downstream VSX LAGs.

The **best practice for LACP timers** on the VSX LAG is to keep the default long timer (30s for lacp rate slow).

The **best practice for MTU** on these upstream VSX LAGs is to configure the maximum value (9198 bytes) like for ISL.

The **best practice for hashing algorithm** on the VSX LAG is to keep the default l3-src-dst (alternative being l2-src-dst), and would have an effect only if at least 2 ports per VSX node are members of the same VSX LAG.

AGG-1(config)#

```
interface lag 101 multi-chassis
    description CORE-1 VSX LAG
    no shutdown
    vlan trunk allowed 101,201
interface 1/1/49
    no shutdown
    mtu 9198
    description CORE-1 1/1/1
    lag 101
interface lag 102 multi-chassis
    description CORE-2 VSX LAG
    no shutdown
    vlan trunk allowed 102,202
interface 1/1/50
    no shutdown
    mtu 9198
    description CORE-2 1/1/1
    lag 102
```

AGG-2(config)#

```
interface lag 101 multi-chassis
    no shutdown
                synchronized
interface 1/1/49
    no shutdown
    mtu 9198
    description CORE-1 1/1/1
    lag 101
interface lag 102 multi-chassis
    no shutdown
                synchronized
interface 1/1/50
    no shutdown
    mtu 9198
    description CORE-2 1/1/2
    lag 102
```

AGG-1 / AGG-2

```
AGG-1# show lldp neighbor-info

LLDP Neighbor Information
=========================

Total Neighbor Entries          : 6
Total Neighbor Entries Deleted  : 0
Total Neighbor Entries Dropped  : 0
Total Neighbor Entries Aged-Out : 0

LOCAL-PORT  CHASSIS-ID         PORT-ID       PORT-DESC     TTL      SYS-NAME
-------------------------------------------------------------------------------
1/1/1       08:00:09:72:61:c6  1/1/49        1/1/49        120      ACC-1
1/1/48      08:00:09:9a:9b:10  1/1/48        1/1/48        120      AGG-2
1/1/49      08:00:09:ac:6e:b7  1/1/1         1/1/1         120      CORE-1
1/1/50      08:00:09:4a:f4:ad  1/1/1         1/1/1         120      CORE-2
1/1/55      08:00:09:9a:9b:10  1/1/55        1/1/55        120      AGG-2
1/1/56      08:00:09:9a:9b:10  1/1/56        1/1/56        120      AGG-2



AGG-1# show 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  Port    Port     State   System-ID         System   Aggr
        name       id      Priority                           Priority Key
-------------------------------------------------------------------------------
1/1/1   lag1(mc)   1       1        ALFNCD  02:01:00:00:01:00 65534    1
1/1/49  lag101(mc) 49      1        ALFNCD  02:01:00:00:01:00 65534    101
1/1/50  lag102(mc) 50      1        ALFNCD  02:01:00:00:01:00 65534    102
```

```
Partner details of all interfaces:
--------------------------------------------------------------------------------
Intf    Aggregate  Partner Port     State   System-ID         System   Aggr
        name       Port-id Priority                           Priority Key
--------------------------------------------------------------------------------
1/1/1   lag1(mc)   49      1        ALFNCD  08:00:09:72:61:c6 65534    1
1/1/49  lag101(mc) 1       1        ALFNCD  08:00:09:ac:6e:b7 65534    1
1/1/50  lag102(mc) 1       1        ALFNCD  08:00:09:4a:f4:ad 65534    1


Remote Actor details of all interfaces:
--------------------------------------------------------------------------------
Intf   Aggregate  Port    Port     State   System-ID         System   Aggr
       name       id      Priority                           Priority Key
--------------------------------------------------------------------------------
1/1/1  lag1(mc)   1001    1        ALFNCD  02:01:00:00:01:00 65534    1
1/1/49 lag101(mc) 1049    1        ALFNCD  02:01:00:00:01:00 65534    101
1/1/50 lag102(mc) 1050    1        ALFNCD  02:01:00:00:01:00 65534    102


Remote Partner details of all interfaces:
--------------------------------------------------------------------------------
Intf   Aggregate  Partner Port     State   System-ID         System   Aggr
       name       Port-id Priority                           Priority Key
--------------------------------------------------------------------------------
1/1/1  lag1(mc)   50      1        ALFNCD  08:00:09:72:61:c6 65534    1
1/1/49 lag101(mc) 2       1        ALFNCD  08:00:09:ac:6e:b7 65534    1
1/1/50 lag102(mc) 2       1        ALFNCD  08:00:09:4a:f4:ad 65534    1
```

Step #12 : VSX linkup-delay-timer exclusion

The **best practice for VSX LAG exclusion for linkup-delay-timer** is to exclude the upstream VSX LAGs that are used for Transit VLANs (in the topology above, VSX LAG 101 and VSX LAG 102). By excluding these upstream VSX LAGs, the routing protocols can establish peering over upstream Transit VLANs as soon as the upstream ports (to L3 Core-1 and Core-2) are UP, without waiting for the linkup-delay timer to complete. The benefit is that routes from L3 core are already learnt when ASIC is ready to forward traffic to downstream.

AGG-1(config)#	AGG-2(config)#
AGG-1(config)# vsx AGG-1(config-vsx)# linkup-delay-timer exclude lag 101,102	**AGG-2(config)# vsx** **AGG-2(config-vsx)# linkup-delay-timer exclude lag 101,102**

AGG-1 / AGG-2 (large-size network)

```
AGG-1# show running-config vsx
vsx
    system-mac 02:01:00:00:01:00
    inter-switch-link lag 256
    role primary
    keepalive peer 192.168.0.1 source 192.168.0.0 vrf KA
    linkup-delay-timer exclude lag 101-102
...
AGG-2# show run vsx
vsx
    system-mac 02:01:00:00:01:00
    inter-switch-link lag 256
    role secondary
    keepalive peer 192.168.0.0 source 192.168.0.1 vrf KA
    linkup-delay-timer exclude lag 101-102

AGG-1# show vsx status linkup-delay
Configured linkup delay-timer                              : 180 seconds
Initial sync status                                        : Completed
Delay timer status                                         : Completed
Linkup Delay time left                                     :
Interfaces that will be brought up after delay timer expires :
Interfaces that are excluded from delay timer              : lag101-lag102

AGG-2# show vsx status linkup-delay
Configured linkup delay-timer                              : 180 seconds
Initial sync status                                        : Completed
Delay timer status                                         : Completed
```

```
Linkup Delay time left                                    :
Interfaces that will be brought up after delay timer expires :
Interfaces that are excluded from delay timer             : lag101-lag102
```

Note: this setting is not synchronized by vsx-sync so it has to be manually set-up on VSX secondary.

Step #13 : MSTP configuration

The **best practice on Aggregation** layer are:

- No loop-protect (MSTP used instead).
- Use the **default common instance 0**: MST0
- Lower the **spanning-tree priority to 4** to make VSX aggregation the STP root bridge (easier for support)
- Use **root-guard** on all downlinks to prevent any access switches from becoming root bridge.
- Keep the default **port-type admin-network**
- Let VSX secondary synchronized by vsx-sync process.

The **best practice for the upstream VSX LAGs to the Core layer is to prevent any BDPU** transmission and reception assuming a continuous forwarding state for these links. The best practice for the Core layer is to avoid any spanning-tree protocol even for such L2 transit connectivity between the core and aggregation layer. The strict provisioning and LACP usage is enough to protect against any cabling mistakes, all the other ports being L3+disabled or LACP+enabled.

AGG-1(config)#	AGG-2(config)#
spanning-tree spanning-tree priority 4 spanning-tree trap topology-change instance 0 interface lag 1 multi-chassis no shutdown description Access-Switch-1 VSX LAG no routing vlan trunk native 1 vlan trunk allowed 10,20-30 lacp mode active spanning-tree root-guard interface lag 101 multi-chassis no shutdown description CORE-1 VSX LAG no routing vlan trunk native 1 vlan trunk allowed 101,201 lacp mode active spanning-tree bpdu-filter interface lag 102 multi-chassis no shutdown description CORE-2 VSX LAG no routing vlan trunk native 1 vlan trunk allowed 102,202 lacp mode active spanning-tree bpdu-filter	synchronized
AGG-1# show run Current configuration: .....skipped for readability spanning-tree spanning-tree priority 4 spanning-tree trap topology-change instance 0 .....skipped for readability interface lag 1 multi-chassis no shutdown description Access-Switch-1 VSX LAG	AGG-2# show run Current configuration: .....skipped for readability spanning-tree spanning-tree priority 4 **spanning-tree trap topology-change instance 0** .....skipped for readability interface lag 1 multi-chassis no shutdown description Access-Switch-1 VSX LAG

```
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20-30
    lacp mode active
    spanning-tree root-guard
.................................skipped for readability
interface lag 101 multi-chassis
    no shutdown
    description CORE-1 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 101,201
    lacp mode active
    spanning-tree bpdu-filter
interface lag 102 multi-chassis
    no shutdown
    description CORE-2 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 102,202
    lacp mode active
    spanning-tree bpdu-filter
```

```
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20-30
    lacp mode active
    spanning-tree root-guard
.................................skipped for readability
interface lag 101 multi-chassis
    no shutdown
    description CORE-1 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 101,201
    lacp mode active
    spanning-tree bpdu-filter
interface lag 102 multi-chassis
    no shutdown
    description CORE-2 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 102,202
    lacp mode active
    spanning-tree bpdu-filter
```

AGG-1 / AGG-2

```
AGG-1# show spanning-tree mst 0 int lag1 detail
Port lag1
Port Type : admin-network    Loop Guard  : disable
Link Type : point_to_point   BPDU Filter : disable
Boundary  : internal         BPDU Guard  : disable
Root Guard: enable

Instance       Role           State        Cost       Priority   Vlans mapped
-------------- -------------- ------------ ---------- ---------- ----------
0              Designated     Forwarding   20000      64         1-4094

Port lag1
Designated root address            : 02:01:00:00:01:00
Designated regional root address   : 02:01:00:00:01:00
Designated bridge address          : 02:01:00:00:01:00  Priority : 16384
Multi-Chassis role                 : active
Timers:    Message expires in 0 sec, Forward delay expiry:18, Forward transitions:1
Bpdus sent 19295, received 2
TCN_Tx: 4, TCN_Rx: 2

AGG-1# show spanning-tree mst 0 int lag101 detail
Port lag101
Port Type : admin-network    Loop Guard  : disable
Link Type : point_to_point   BPDU Filter : enable
Boundary  : internal         BPDU Guard  : disable
Root Guard: disable

Instance       Role           State        Cost       Priority   Vlans mapped
-------------- -------------- ------------ ---------- ---------- ----------
0              Designated     Forwarding   20000      64         1-4094

Port lag101
Designated root address            : 02:01:00:00:01:00
Designated regional root address   : 02:01:00:00:01:00
Designated bridge address          : 02:01:00:00:01:00  Priority : 16384
Multi-Chassis role                 : active
Timers:    Message expires in 1 sec, Forward delay expiry:18, Forward transitions:1
Bpdus sent 1239, received 0
TCN_Tx: 2, TCN_Rx: 0
```

The **best practice on Access** layer are:

- Use loop-protect for all endpoint access ports (not configured on uplinks). Set the re-enable timer to 1hour.
- Keep the **default common instance 0**: MST0
- Keep the **default spanning-tree priority** of 8.

- All endpoint access ports are admin-edge, should not receive any BPDU (BDPU guard), should not trigger Topology Change Notification.
- Use loop-protection on all endpoint access ports as an extra-protection mechanism (in case of MSTP BPDUs are filtered by insertion of unmanaged switches which create a loop).
- Use **loop-guard** on all uplinks to prevent any flood due to failure of BPDU reception (fiber strand cut).

ACC-1(config)#

```
spanning-tree
loop-protect re-enable-timer 3600
interface lag 1
    no shutdown
    description UPLINK to AGG
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20
    lacp mode active
    spanning-tree loop-guard
interface 1/1/1
    no shutdown
    description Endpoint1
    no routing
    vlan access 10
    spanning-tree bpdu-guard
    spanning-tree port-type admin-edge
    spanning-tree tcn-guard
    loop-protect

ACC-1# show spanning-tree mst 0 interface lag1 detail
Port lag1
Port Type : admin-network    Loop Guard  : enable
Link Type : point_to_point   BPDU Filter : disable
Boundary  : internal         BPDU Guard  : disable
Root Guard: disable

Instance       Role           State        Cost       Priority   Vlans mapped
-------------- -------------- ------------ ---------- ---------- ----------
0              Root           Forwarding   20000      64         1-4094

Port lag1
Designated root address            : 02:01:00:00:01:00
Designated regional root address   : 08:00:09:72:61:c6
Designated bridge address          : 02:01:00:00:01:00  Priority : 32768
Timers:    Message expires in 0 sec, Forward delay expiry:1, Forward transitions:9
Bpdus sent 31, received 145466
TCN_Tx: 18, TCN_Rx: 20

ACC-1# show spanning-tree mst 0 interface 1/1/1 detail
Port 1/1/1
Port Type : admin-edge       Loop Guard  : disable
Link Type : point_to_point   BPDU Filter : disable
Boundary  : internal         BPDU Guard  : enable
Root Guard: disable

Instance       Role           State        Cost       Priority   Vlans mapped
-------------- -------------- ------------ ---------- ---------- ----------
0              Designated     Forwarding   20000      128        1-4094

Port 1/1/1
Designated root address            : 02:01:00:00:01:00
Designated regional root address   : 08:00:09:72:61:c6
Designated bridge address          : 02:01:00:00:01:00  Priority : 32768
Timers:    Message expires in 0 sec, Forward delay expiry:1, Forward transitions:1
Bpdus sent 160124, received 0
TCN_Tx: 0, TCN_Rx: 0
```

Step #14 : VSX LAG ACL configuration

See Step #11 for single VRF scenario and configuration example.

Step #15 : VSX LAG QoS configuration

QoS Marking being performed on the access layer, the aggregation switch is configure in the global context with **qos trust dscp**. No further configuration is needed as this was already set in step#1.

Step #16 : SVI (VLAN L3 interface) configuration

Note: While creating the SVI, the associated VRF must be attached to the SVI. The attachment to the VRF is not automatically synchronized and has to be manually set on both VSX nodes.

The **best practice for SVI active-gateway** is to set the active-gateways Virtual IP and Virtual MAC on the VSX primary and get the value synchronized on the VSX secondary with vsx-sync command.

The **best practice for active-gateway VMAC** is to use the **same VMAC for all IPv4 SVIs**. The scope of this VMAC is purely link-local. If some servers or systems have dual-attachment to two different SVIs, and the system administrator would like to see distinct MAC addresses for the next-hops over these separate interfaces, then 16 VMACs are available. For dual-stack IPv4 and IPv6, 16 VMACs can be used for IPv4 and the same VMACs can be used for IPv6. It is however a best practice to use only 8 VMACs for IPv4 and 8 different VMACs for IPV6.

If mutlinetting is used, set one VIP per secondary subnet.

The **best practice for IP MTU** is to configure on all SVIs the matching size of the L2 MTU: IP MTU recommended value = 9100. This parameter must be identical and manually set on both VSX nodes.

The **best practice for DHCP relay** is to configure the ip helper-address on the VSX primary and let vsx-sync configuring the same on the VSX secondary.

AGG-1(config)#	AGG-2(config)#
interface vlan10 vsx-sync active-gateways vrf attach VRF1 ip mtu 9100 ip address 10.1.10.2/24 ip address 10.2.10.2/24 secondary active-gateway ip mac 12:01:00:00:01:00 active-gateway ip 10.1.10.1 active-gateway ip 10.2.10.1 ip helper-address 10.99.10.9 ip ospf 1 area 0.0.0.0	interface vlan10 vrf attach VRF1 ip mtu 9100 ip address 10.1.10.3/24 ip address 10.2.10.3/24 secondary synchronized
interface vlan20 vsx-sync active-gateways vrf attach VRF2 ip mtu 9100 ip address 10.1.20.2/24 ip address 10.2.20.2/24 secondary active-gateway ip mac 12:01:00:00:01:00 active-gateway ip 10.1.20.1 active-gateway ip 10.2.20.1 ip ospf 1 area 0.0.0.0	interface vlan20 vrf attach VRF2 ip mtu 9100 ip address 10.1.20.3/24 ip address 10.2.20.3/24 secondary synchronized
AGG-1# show run Current configuration: .....skipped for readability interface vlan10 vsx-sync active-gateways vrf attach VRF1 ip mtu 9100 ip address 10.1.10.2/24 ip address 10.2.10.2/24 secondary active-gateway ip mac 12:01:00:00:01:00 active-gateway ip 10.1.10.1 active-gateway ip 10.2.10.1	AGG-2# show run Current configuration: .....skipped for readability interface vlan10 vsx-sync active-gateways vrf attach VRF1 ip mtu 9100 ip address 10.1.10.3/24 ip address 10.2.10.3/24 secondary active-gateway ip mac 12:01:00:00:01:00 active-gateway ip 10.1.10.1 active-gateway ip 10.2.10.1

```
    ip helper-address 10.99.10.9
    ip ospf 1 area 0.0.0.0
interface vlan20
    vsx-sync active-gateways
    vrf attach VRF2
    ip mtu 9100
    ip address 10.1.20.2/24
    ip address 10.2.20.2/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.20.1
    active-gateway ip 10.2.20.1
    ip ospf 1 area 0.0.0.0
.................................skipped for readability
```

```
    ip helper-address 10.99.10.9
    ip ospf 1 area 0.0.0.0
interface vlan20
    vsx-sync active-gateways
    vrf attach VRF2
    ip mtu 9100
    ip address 10.1.20.3/24
    ip address 10.2.20.3/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.20.1
    active-gateway ip 10.2.20.1
    ip ospf 1 area 0.0.0.0
.................................skipped for readability
```

AGG-1 / AGG-2

```
AGG-1# show ip interface vlan10

Interface vlan10 is up
 Admin state is up
 Hardware: Ethernet, MAC Address: 08:00:09:94:00:b8
 IP MTU 9100
 IPv4 address 10.1.10.2/24
 IPv4 address 10.2.10.2/24 secondary
 active-gateway ip mac 12:01:00:00:01:00
 active-gateway ip 10.1.10.1
 active-gateway ip 10.2.10.1
 L3 Counters: Rx Disabled, Tx Disabled
 Rx
          ucast: 0 packets, 0 bytes
          mcast: 0 packets, 0 bytes
 Tx
          ucast: 0 packets, 0 bytes
          mcast: 0 packets, 0 bytes

AGG-1# show ip interface vlan10 vsx-peer

Interface vlan10 is up
 Admin state is up
 Hardware: Ethernet, MAC Address: 08:00:09:b5:e8:2d
 IP MTU 9100
 IPv4 address 10.1.10.3/24
 IPv4 address 10.2.10.3/24 secondary
 active-gateway ip mac 12:01:00:00:01:00
 active-gateway ip 10.1.10.1
 active-gateway ip 10.2.10.1
 L3 Counters: Rx Disabled, Tx Disabled
 Rx
          ucast: 0 packets, 0 bytes
          mcast: 0 packets, 0 bytes
 Tx
          ucast: 0 packets, 0 bytes
          mcast: 0 packets, 0 bytes

AGG-1# show vrf
VRF Configuration:
------------------
VRF Name   : default
        Interfaces              Status
        ----------------------------

VRF Name   : KA
        Interfaces              Status
        ----------------------------
        1/1/2                     up

VRF Name   : VRF1
        Interfaces              Status
        ----------------------------
        vlan10                    up

VRF Name   : VRF2
        Interfaces              Status
```

```
        ----------------------------
        vlan20                    up
```

Step #17 : OSPF configuration (including SVIs for Transit VLANs)

The Transit VLANs previously defined (101/102/201/202) are used as OSPF broadcast network domains for routing between both VSX nodes and each core devices. There are 3 OSPF peers per Transit VLANs. Inside this OSPF broadcast network type, The **Best Practice is to set the Core switch as OSPF DR** (Designated Router), leaving both VSX nodes as BDR and DRother. Such DR role hosted on the core will ensure no OSPF impact while performing a VSX upgrade as the secondary and primary will reboot sequentially.

In the current example, Core-1 and Core-2 being AOS-CX, one Transit VLAN per VRF has to be created on Core-1 and Core-2: VLAN 11 for VRF1 and VLAN 12 for VRF2. Both VLAN 11 and 12 are trunked over the L2 point-to-point circuit between Core-1 and Core-2.

There are 2 strategies to inject endpoint subnets into the routing table: either through OSPF or with BGP

- OSPF: Most of the Campus deployments use OSPF to exchange route information for end-devices. This is simple and can scale very well with appropriate usage of areas. This is the target of this current document.
- BGP: Lot of new DC deployment use BGP as a routing protocol due to the usage of EVPN based VXLAN. Such a design is coming in the Campus as well. Also, for more complex and granular routing engineering, BGP communities and route-map can offer a level of control that OSPF can not provide. This can be exposed in a future white paper.

There are two options to inject end-user subnets into OSPF database: using OSPF command on the SVI, or redistributing the connected into OSPF with route-map control. The **best practice is to use the OSPF command on SVI** as offering a simpler configuration like for the area the subnets belongs to. This principle is selected as the OSPF best practice in the following described configuration. More details on OSPF best practices can be found on IP routing configuration guide.

The **best practice for broadcast interconnectivity subnet** is to use /29 subnet.

The **best practice for point-to-point interconnectivity subnet** is to use /31 subnet.

The **best practice for OSPF configuration** is to use vsx-sync OSPF synchronization option and have OSPF parameters automatically synced on VSX secondary. As shown on the configuration step, very few elements have to be configured on secondary.

The **best practice for OSPF cost** is to set values so that the Core-1 to Core-2 link is used for backbone transit whereas AGG-1 to AGG-2 is used for the Aggregation transit. Consequently, it is recommended to set the same cost on all Transit VLANs including the ones between Core-1 and Core-2. In the below example, OSPF cost for all Transit VLANs are set to 50. OSPF cost is synchronized from the VSX primary to the VSX secondary.

OSPF process ID can be the same for all VRFs. it might be useful for some network operators to keep the same process ID on all VRFs as the process ID can be associated to a Private Autonomous System ID.

The **best practice for router-ID** is to keep the same router-ID as being the Loopback 0 IP address. This simplifies a lot the troubleshooting by having the same router-ID for all VRFs when debugging.

Due to upstream VSX LAG and associated ECMP to reach the downstream VSX nodes for L3 lookup, it is a **best practice to configure VSX active-forwarding** on the Transit SVIs between the VSX cluster and the Core layer (SVI 101/102/201/202). This provides two major benefits: this saves ISL bandwidth and avoid sub-optimum data-path, and this provides High-Availability in case of one VSX node failure as the other VSX node will process the L3 lookup on behalf of its peer. VSX active-forwarding is **mutually exclusive with ip icmp redirect** which needs to be disabled.

AGG-1(config)#	AGG-2(config)#

`router ospf 1 vrf VRF1`
`  router-id 10.0.1.1`
`  max-metric router-lsa on-startup`
`  passive-interface default`
	`router ospf 1 vrf VRF1`
	`  router-id 10.0.1.2`
	
	synchronized

```
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
router ospf 1 vrf VRF2
    router-id 10.0.1.1
    max-metric router-lsa on-startup
    passive-interface default
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0

interface loopback 0
    vrf attach VRF1
    ip address 10.0.1.1/32
    ip ospf 1 area 0

no ip icmp redirect

interface vlan101
    vrf attach VRF1
    ip mtu 9198
    ip address 10.0.101.2/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 plaintext yourpass
interface vlan102
    vrf attach VRF1
    ip mtu 9198
    ip address 10.0.102.2/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 plaintext yourpass
interface vlan201
    vrf attach VRF2
    ip mtu 9198
    ip address 10.0.201.2/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 plaintext yourpass
interface vlan202
    vrf attach VRF2
    ip mtu 9198
    ip address 10.0.202.2/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 plaintext yourpass

interface vlan10
    ip ospf 1 area 0.0.0.0
interface vlan20
    ip ospf 1 area 0.0.0.0

AGG-1# show lldp neighbor-info

LLDP Neighbor Information
=========================

Total Neighbor Entries          : 6
Total Neighbor Entries Deleted  : 0
Total Neighbor Entries Dropped  : 0
Total Neighbor Entries Aged-Out : 0
```

```
router ospf 1 vrf VRF2
    router-id 10.0.1.2
```

synchronized

```
interface loopback 0
    vrf attach VRF1
    ip address 10.0.1.2/32
```

synchronized

```
interface vlan101
    vrf attach VRF1
    ip mtu 9198
    ip address 10.0.101.3/29
```

synchronized

```
interface vlan102
    vrf attach VRF1
    ip mtu 9198
    ip address 10.0.102.3/29
```

synchronized

```
interface vlan201
    vrf attach VRF2
    ip mtu 9198
    ip address 10.0.201.3/29
```

synchronized

```
interface vlan202
    vrf attach VRF2
    ip mtu 9198
    ip address 10.0.202.3/29
```

synchronized

```
AGG-2# show  lldp neighbor-info

LLDP Neighbor Information
=========================

Total Neighbor Entries          : 6
Total Neighbor Entries Deleted  : 1
Total Neighbor Entries Dropped  : 0
Total Neighbor Entries Aged-Out : 1
```

```
LOCAL-PORT  CHASSIS-ID         PORT-ID       PORT-DESC
TTL      SYS-NAME
------------------------------------------------------------
--------------------
1/1/1       08:00:09:72:61:c6  1/1/49        1/1/49
120      ACC-1
1/1/48      08:00:09:9a:9b:10  1/1/48        1/1/48
120      AGG-2
1/1/49      08:00:09:ac:6e:b7  1/1/1         1/1/1
120      CORE-1
1/1/50      08:00:09:4a:f4:ad  1/1/1         1/1/1
120      CORE-2
1/1/55      08:00:09:9a:9b:10  1/1/55        1/1/55
120      AGG-2
1/1/56      08:00:09:9a:9b:10  1/1/56        1/1/56
120      AGG-2

AGG-1# show run
Current configuration:
.................................skipped for readability
no ip icmp redirect
vrf KA
vrf VRF1
vrf VRF2
.................................skipped for readability
router ospf 1 vrf VRF1
    router-id 10.0.1.1
    max-metric router-lsa on-startup
    passive-interface default
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
router ospf 1 vrf VRF2
    router-id 10.0.1.1
    max-metric router-lsa on-startup
    passive-interface default
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
.................................skipped for readability
vlan 101
    vsx-sync
    description TRANSIT VLAN VRF1-CORE1
vlan 102
    vsx-sync
    description TRANSIT VLAN VRF1-CORE2
vlan 201
    vsx-sync
    description TRANSIT VLAN VRF2-CORE1
vlan 202
    vsx-sync
    description TRANSIT VLAN VRF2-CORE2
.................................skipped for readability
interface lag 101 multi-chassis
    no shutdown
    description CORE-1 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 101,201
    lacp mode active
    spanning-tree bpdu-filter
interface lag 102 multi-chassis
    no shutdown
    description CORE-2 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 102,202
    lacp mode active
    spanning-tree bpdu-filter
.................................skipped for readability
interface loopback 0
```

```
LOCAL-PORT  CHASSIS-ID         PORT-ID       PORT-DESC
TTL      SYS-NAME
------------------------------------------------------------
--------------------
1/1/1       08:00:09:72:61:c6  1/1/50        1/1/50
120      ACC-1
1/1/48      08:00:09:b0:c4:aa  1/1/48        1/1/48
120      AGG-1
1/1/49      08:00:09:ac:6e:b7  1/1/2         1/1/2
120      CORE-1
1/1/50      08:00:09:4a:f4:ad  1/1/2         1/1/2
120      CORE-2
1/1/55      08:00:09:b0:c4:aa  1/1/55        1/1/55
120      AGG-1
1/1/56      08:00:09:b0:c4:aa  1/1/56        1/1/56
120      AGG-1

AGG-2# show run
Current configuration:
.................................skipped for readability
no ip icmp redirect
vrf KA
vrf VRF1
vrf VRF2
.................................skipped for readability
router ospf 1 vrf VRF1
    router-id 10.0.1.2
    max-metric router-lsa on-startup
    passive-interface default
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
router ospf 1 vrf VRF2
    router-id 10.0.1.2
    max-metric router-lsa on-startup
    passive-interface default
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
.................................skipped for readability
vlan 101
    vsx-sync
    description TRANSIT VLAN VRF1-CORE1
vlan 102
    vsx-sync
    description TRANSIT VLAN VRF1-CORE2
vlan 201
    vsx-sync
    description TRANSIT VLAN VRF2-CORE1
vlan 202
    vsx-sync
    description TRANSIT VLAN VRF2-CORE2
.................................skipped for readability
interface lag 101 multi-chassis
    no shutdown
    description CORE-1 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 101,201
    lacp mode active
    spanning-tree bpdu-filter
interface lag 102 multi-chassis
    no shutdown
    description CORE-2 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 102,202
    lacp mode active
    spanning-tree bpdu-filter
.................................skipped for readability
interface loopback 0
```

```
    vrf attach VRF1
    ip address 10.0.1.1/32
    ip ospf 1 area 0.0.0.0
.................................skipped for readability
interface vlan101
    vrf attach VRF1
    ip mtu 9198
    vsx active-forwarding
    ip address 10.0.101.2/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext
AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan102
    vrf attach VRF1
    ip mtu 9198
    vsx active-forwarding
    ip address 10.0.102.2/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext
AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan201
    vrf attach VRF2
    ip mtu 9198
    vsx active-forwarding
    ip address 10.0.201.2/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext
AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan202
    vrf attach VRF2
    ip mtu 9198
    vsx active-forwarding
    ip address 10.0.202.2/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext
AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX

interface vlan10
    vsx-sync active-gateways
    vrf attach VRF1
    ip mtu 9100
    ip address 10.1.10.2/24
    ip address 10.2.10.2/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.10.1
    active-gateway ip 10.2.10.1
    ip helper-address 10.99.10.9
    ip ospf 1 area 0.0.0.0
interface vlan20
    vsx-sync active-gateways
    vrf attach VRF2
    ip mtu 9100
    ip address 10.1.20.2/24
    ip address 10.2.20.2/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.20.1
    active-gateway ip 10.2.20.1
    ip ospf 1 area 0.0.0.0
```

```
    vrf attach VRF1
    ip address 10.0.1.2/32
    ip ospf 1 area 0.0.0.0
.................................skipped for readability
interface vlan101
    vrf attach VRF1
    ip mtu 9198
    vsx active-forwarding
    ip address 10.0.101.3/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext
AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan102
    vrf attach VRF1
    ip mtu 9198
    vsx active-forwarding
    ip address 10.0.102.3/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext
AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan201
    vrf attach VRF2
    ip mtu 9198
    vsx active-forwarding
    ip address 10.0.201.3/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext
AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan202
    vrf attach VRF2
    ip mtu 9198
    vsx active-forwarding
    ip address 10.0.202.3/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext
AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX

interface vlan10
    vsx-sync active-gateways
    vrf attach VRF1
    ip mtu 9100
    ip address 10.1.10.3/24
    ip address 10.2.10.3/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.10.1
    active-gateway ip 10.2.10.1
    ip helper-address 10.99.10.9
    ip ospf 1 area 0.0.0.0
interface vlan20
    vsx-sync active-gateways
    vrf attach VRF2
    ip mtu 9100
    ip address 10.1.20.3/24
    ip address 10.2.20.3/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.20.1
    active-gateway ip 10.2.20.1
    ip ospf 1 area 0.0.0.0
```

AGG-1 / AGG-2

```
AGG-1# show ip ospf neighbors all-vrfs
OSPF Process ID 1 VRF VRF1
==========================

Total Number of Neighbors: 4

Neighbor ID      Priority  State             Nbr Address       Interface
------------------------------------------------------------------------------
10.0.1.2         1         FULL/BDR          10.0.101.3         vlan101

10.0.1.3         100       FULL/DR           10.0.101.1         vlan101

10.0.1.2         1         FULL/BDR          10.0.102.3         vlan102

10.0.1.4         100       FULL/DR           10.0.102.1         vlan102

OSPF Process ID 1 VRF VRF2
==========================

Total Number of Neighbors: 4

Neighbor ID      Priority  State             Nbr Address       Interface
------------------------------------------------------------------------------
10.0.1.2         1         FULL/BDR          10.0.201.3         vlan201

10.0.1.3         100       FULL/DR           10.0.201.1         vlan201

10.0.1.2         1         FULL/BDR          10.0.202.3         vlan202

10.0.1.4         100       FULL/DR           10.0.202.1         vlan202


AGG-1# show ip route vrf VRF1

Displaying ipv4 routes selected for forwarding

'[x/y]' denotes [distance/metric]

0.0.0.0/0, vrf VRF1
        via  10.0.102.1,  [110/1],  ospf
        via  10.0.101.1,  [110/1],  ospf
10.0.1.1/32, vrf VRF1
        via  loopback0,  [0/0],  local
10.0.1.2/32, vrf VRF1
        via  10.0.101.3,  [110/50],  ospf
        via  10.0.102.3,  [110/50],  ospf
10.0.1.3/32, vrf VRF1
        via  10.0.101.1,  [110/50],  ospf
10.0.1.4/32, vrf VRF1
        via  10.0.102.1,  [110/50],  ospf
10.0.11.0/29, vrf VRF1
        via  10.0.102.1,  [110/100],  ospf
        via  10.0.101.1,  [110/100],  ospf
10.0.101.0/29, vrf VRF1
        via  vlan101,  [0/0],  connected
10.0.101.2/32, vrf VRF1
        via  vlan101,  [0/0],  local
10.0.102.0/29, vrf VRF1
        via  vlan102,  [0/0],  connected
10.0.102.2/32, vrf VRF1
        via  vlan102,  [0/0],  local
10.1.10.0/24, vrf VRF1
        via  vlan10,  [0/0],  connected
10.1.10.2/32, vrf VRF1
        via  vlan10,  [0/0],  local
10.2.10.0/24, vrf VRF1
        via  vlan10,  [0/0],  connected
```

```
10.2.10.2/32, vrf VRF1
        via  vlan10,  [0/0],  local
```

CORE-1 / CORE-2

```
CORE-1# show ip ospf neighbors all-vrfs
OSPF Process ID 1 VRF VRF1
==========================

Total Number of Neighbors: 3

Neighbor ID      Priority  State             Nbr Address       Interface
------------------------------------------------------------------------
10.0.1.1         1         FULL/BDR          10.0.101.2         vlan101

10.0.1.2         1         FULL/DROther      10.0.101.3         vlan101

10.0.1.4         n/a       FULL              10.0.11.2          vlan11

OSPF Process ID 1 VRF VRF2
==========================

Total Number of Neighbors: 3

Neighbor ID      Priority  State             Nbr Address       Interface
------------------------------------------------------------------------
10.0.1.1         1         FULL/BDR          10.0.201.2         vlan201

10.0.1.2         1         FULL/DROther      10.0.201.3         vlan201

10.0.1.4         n/a       FULL              10.0.12.2          vlan12


CORE-1# show ip route vrf VRF1

Displaying ipv4 routes selected for forwarding

'[x/y]' denotes [distance/metric]

10.0.1.2/32, vrf VRF1
        via  10.0.101.3,  [110/50],  ospf
10.0.1.4/32, vrf VRF1
        via  10.0.11.2,  [110/50],  ospf
10.0.1.1/32, vrf VRF1
        via  10.0.101.2,  [110/50],  ospf
10.0.102.0/29, vrf VRF1
        via  10.0.101.3,  [110/100],  ospf
        via  10.0.101.2,  [110/100],  ospf
        via  10.0.11.2,  [110/100],  ospf
10.1.10.0/24, vrf VRF1
        via  10.0.101.3,  [110/150],  ospf
        via  10.0.101.2,  [110/150],  ospf
10.2.10.0/24, vrf VRF1
        via  10.0.101.3,  [110/50],  ospf
        via  10.0.101.2,  [110/50],  ospf
```

Step #18 : BGP configuration

Please read the IP routing guide.

Note: VSX provides capability to synchronize the full BGP configuration between the VSX primary and the VSX secondary. Most of the BGP configuration of the VSX secondary is the same than on the VSX primary. Except the configuration for iBGP peering between the VSX nodes inside the cluster, or for remote eBGP upstream peers with neighbor IP address being the physical IP address of theL3 point-to-point circuit (ex: here 1/1/50 with 10.0.0.6/31 remote IP address). In such a case, the specific neighbor parameters are excluded from the VSX configuration synchronization with the following command on the VSX primary only: neighbor <IP_adress> vsx-sync-exclude

Step #19 : Multicast configuration

Please read the Multicast guide.

For multicast on VSX cluster, the **best practice is to configure PIM Dual-DR or active/active** under the PIM router command: to

AGG-1(config)#

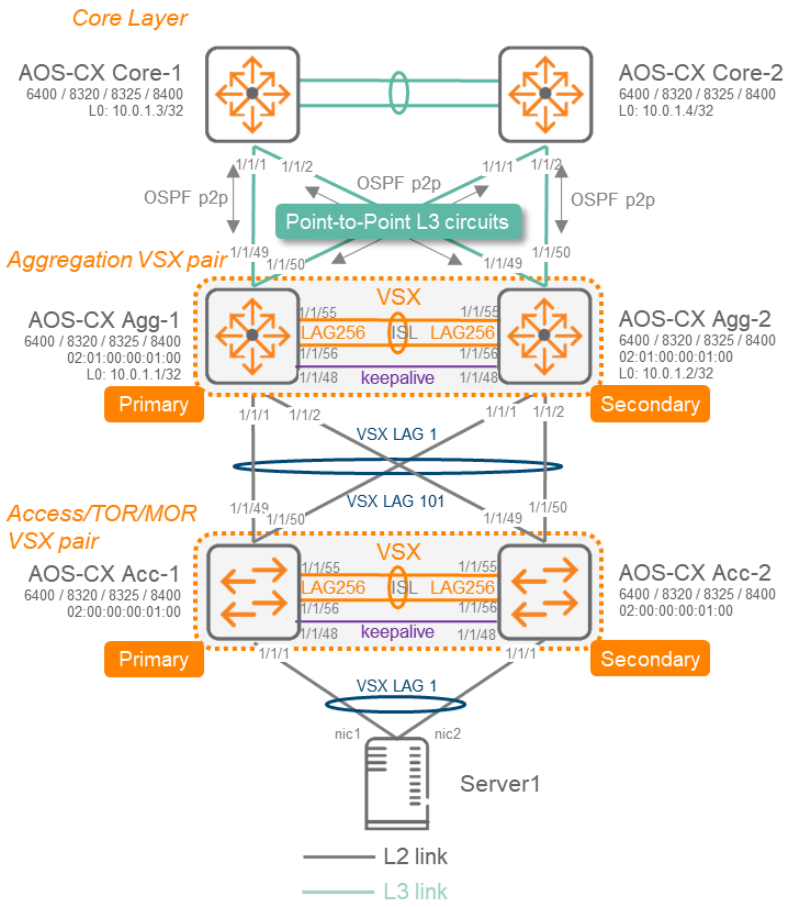
```
router pim vrf VRF1
    enable
    active-active

show ip pim interface vlan10

PIM Interfaces

VRF: VRF1

Interface  : vlan10
IP Address : 10.1.10.2/24
Mode       : sparse

Designated Router : 10.1.10.2
Proxy DR              : false
Hello Interval (sec)  : 30
Hello Delay (sec)     : 5

Override Interval (msec)  : 2500            Lan Prune
Delay     : Yes
Propagation Delay (msec)  : 500             DR Priority
: 200
Neighbor Timeout          : 0
```

AGG-2(config)#

```
router pim vrf VRF1
    enable
    active-active

show ip pim interface vlan10

PIM Interfaces

VRF: VRF1

Interface  : vlan10
IP Address : 10.1.10.3/24
Mode       : sparse

Designated Router : 10.1.10.3
Proxy DR              : true
Hello Interval (sec)  : 30
Hello Delay (sec)     : 5

Override Interval (msec)  : 2500            Lan Prune
Delay     : Yes
Propagation Delay (msec)  : 500             DR
Priority         : 1
Neighbor Timeout          : 0
```

Access VSX to Aggregation VSX

Here is the typical topology:



This scenario is very similar to the Single VRF scenario in term of configuration steps. The below sections will focus on VSX at the Access Layer and repeated steps will be skipped. As this scenario is valid for both a Campus Access layer and for a Datacenter Top-of-Rack layer, we use Access/ToR layer naming convention.

Step #0 to Step#8: follow same steps than for the single VRF scenario

Step #9 : Aggregation - Downstream VSX LAG (MCLAG) configuration

The best practice for VSX LAG is to create the multi-chassis lag interface on the VSX primary with all settings and then create the mirrored lag interface on the VSX secondary. LAG interface settings (including description) will be synchronized automatically. Only "no shut" in the lag interface has to be performed on the VSX secondary. Once multi-chassis lag interface is created it is assigned to the physical port.

The best practice for allowed VLANs is to exclude the native VLAN 1 from being propagated. This is a very robust method to avoid Layer2 storm propagation due to potential loop initiated on an access switch. In case of access switch Zero-Touch-Provisioning use-case., this trunking exclusion is performed after ZTP process.

The best practice for LAG numbering is to use LAG ID=1 for connecting the Access/ToR Switch Cluster#1, LAG ID=2 for connecting the Access/ToR Switch#2, and so on…

The best practice for LACP timers on the VSX LAG is to keep the default long timer (30s for lacp rate slow).

The **best practice for MTU** is to configure on all devices the appropriate size to support features such as Dynamic Segmentation, VXLAN or server jumbo frame. Care should be taken to ensure that the IP path from access devices (switches or APs) can provide a MTU of at least 1564 bytes to the mobility controllers and that 9000 bytes server jumbo packet can be encapsulated. Flexibility should be anticipated to perform VXLAN encapsulation from the access/tor switch (9000+50) or VXLAN encapsulation from aggregation switch with MTU+50. So the recommended Ethernet MTU is 9100 bytes for the downstream VSX LAG to the access layer and MTU of 9000 bytes for endpoints or servers. SVI IP MTU should match the MTU size on the aggregation, so the recommended IP MTU is 9100 bytes.

The **best practice for hashing algorithm** on the VSX LAG is to keep the default l3-src-dst (alternative being l2-src-dst).

AGG-1(config)#

```
interface lag 1 multi-chassis
    description TOR-VSX-1
    no shutdown
    vlan trunk allowed 10,20-30

interface 1/1/1
    no shutdown
    mtu 9100
    description TOR-1
    lag 1
interface 1/1/2
    no shutdown
    mtu 9100
    description TOR-2
    lag 1
```

AGG-2(config)#

```
interface lag 1 multi-chassis
    no shutdown
```

synchronized

```
interface 1/1/1
    no shutdown
    mtu 9100
    description TOR-1
    lag 1
interface 1/1/2
    no shutdown
    mtu 9100
    description TOR-2
    lag 1
```

AGG-1 / AGG-2

```
AGG-1# show 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  Port    Port     State   System-ID         System   Aggr
        name       id      Priority                           Priority Key
------------------------------------------------------------------------------
1/1/1   lag1(mc)   1       1        ALFOE   02:01:00:00:01:00 65534    1
1/1/2   lag1(mc)   2       1        ALFOE   02:01:00:00:01:00 65534    1


Partner details of all interfaces:
------------------------------------------------------------------------------
Intf    Aggregate  Partner Port     State   System-ID         System   Aggr
        name       Port-id Priority                           Priority Key
------------------------------------------------------------------------------
1/1/1   lag1(mc)   0       65534    PLFOEX  00:00:00:00:00:00 65534    0
1/1/2   lag1(mc)   0       65534    PLFOEX  00:00:00:00:00:00 65534    0


Remote Actor details of all interfaces:
------------------------------------------------------------------------------
Intf   Aggregate  Port    Port     State   System-ID         System   Aggr
       name       id      Priority                           Priority Key
------------------------------------------------------------------------------
1/1/2  lag1(mc)   1002    1        ALFOE   02:01:00:00:01:00 65534    1
1/1/1  lag1(mc)   1001    1        ALFOE   02:01:00:00:01:00 65534    1


Remote Partner details of all interfaces:
------------------------------------------------------------------------------
Intf   Aggregate  Partner Port     State   System-ID         System   Aggr
       name       Port-id Priority                           Priority Key
------------------------------------------------------------------------------
1/1/2  lag1(mc)   0       65534    PLFOXE  00:00:00:00:00:00 65534    0
```

```
1/1/1   lag1(mc)    0        65534     PLFOXE   00:00:00:00:00:00 65534     0
```

The "show lacp interfaces multi-chassis" command is very useful to get a complete status of local LACP partnership as well as the VSX peer partnership details. Actor = local node, Partner = LACP neighbor (the access switch), Remote Actor = the VSX peer, Remote Partner = LACP neighbor of the VSX peer. Note that the port id of the VSX secondary is equal to 1000+ID_of_the_primary (in the example 1001).

At this stage, no LACP partners are yet configured, so ALFOE LACP state-flags should appear on all entries.

Step #10 : Access/ToR Layer: ISL, keepalive, VSX, vsx-sync, VLANs

Follow the same steps #0 to #8 for the Access/ToR VSX than the ones completed for the Aggregation VSX.

At this stage you should have the following configuration on the Access/TOR VSX:

TOR1	TOR-2

hostname TOR-1
.....skipped for readability
vrf KA
vlan 1
vlan 10
    vsx-sync
vlan 20
    vsx-sync
interface lag 256
    no shutdown
    description ISL link
    no routing
    vlan trunk native 1 tag
    vlan trunk allowed all
    lacp mode active
interface 1/1/48
    no shutdown
    vrf attach KA
    description VSX keepalive
    ip address 192.168.0.2/31
interface 1/1/55
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
interface 1/1/56
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
vsx
    system-mac 02:00:00:00:01:00
    inter-switch-link lag 256
    role secondary
    keepalive peer 192.168.0.2 source 192.168.0.3 vrf KA
    vsx-sync aaa acl-log-timer bfd-global bgp copp-policy dhcp-relay dhcp-server dhcp-snooping dns icmp-tcp lldp loop-protect-global mac-lockout mclag-interfaces neighbor ospf qos-global route-map sflow-global snmp ssh stp-global time vsx-global
	hostname TOR-2
	.....skipped for readability
	vrf KA
	vlan 1
	vlan 10
	    vsx-sync
	vlan 20
	    vsx-sync
	interface lag 256
	    no shutdown
	    description ISL link
	    no routing
	    vlan trunk native 1 tag
	    vlan trunk allowed all
	    lacp mode active
	interface 1/1/48
	    no shutdown
	    vrf attach KA
	    description VSX keepalive
	    ip address 192.168.0.3/31
	interface 1/1/55
	    no shutdown
	    mtu 9198
	    description ISL physical link
	    lag 256
	interface 1/1/56
	    no shutdown
	    mtu 9198
	    description ISL physical link
	    lag 256
	vsx
	    system-mac 02:00:00:00:01:00
	    inter-switch-link lag 256
	    role secondary
	    keepalive peer 192.168.0.2 source 192.168.0.3 vrf KA
	    vsx-sync aaa acl-log-timer bfd-global bgp copp-policy dhcp-relay dhcp-server dhcp-snooping dns icmp-tcp lldp loop-protect-global mac-lockout mclag-interfaces neighbor ospf qos-global route-map sflow-global snmp ssh stp-global time vsx-global

Step #11 : Access/ToR - Upstream VSX LAG configuration

The steps for configuring upstream VSX LAG of the TOR layer are very similar to the steps for configuring the downlink VSX LAG of the Aggregation layer and the same best practices apply. LAG numbering is assigning high number for upstream.

TOR-1(config)#	TOR-2(config)#

interface lag 101 multi-chassis
    description AGG VSX
    no shutdown
    vlan trunk allowed 10,20
 	
interface 1/1/49
    no shutdown
	interface lag 101 multi-chassis
	    no shutdown
	synchronized
	
	interface 1/1/49
	    no shutdown

```
    mtu 9100
    description AGG-1
    lag 101
interface 1/1/50
    no shutdown
    mtu 9100
    description AGG-2
    lag 101
```

```
    mtu 9100
    description AGG-1
    lag 101
interface 1/1/50
    no shutdown
    mtu 9100
    description AGG-2
    lag 101
```

AGG-1 / AGG-2

```
AGG-1# show 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  Port    Port     State   System-ID         System   Aggr
        name       id      Priority                           Priority Key
------------------------------------------------------------------------------
1/1/1   lag1(mc)   1       1        ALFNCD  02:01:00:00:01:00 65534    1
1/1/2   lag1(mc)   2       1        ALFNCD  02:01:00:00:01:00 65534    1


Partner details of all interfaces:
------------------------------------------------------------------------------
Intf    Aggregate  Partner Port     State   System-ID         System   Aggr
        name       Port-id Priority                           Priority Key
------------------------------------------------------------------------------
1/1/1   lag1(mc)   49      1        ALFNCD  02:00:00:00:01:00 65534    1
1/1/2   lag1(mc)   1049    1        ALFNCD  02:00:00:00:01:00 65534    1


Remote Actor details of all interfaces:
------------------------------------------------------------------------------
Intf   Aggregate  Port    Port     State   System-ID         System   Aggr
       name       id      Priority                           Priority Key
------------------------------------------------------------------------------
1/1/1  lag1(mc)   1001    1        ALFNCD  02:01:00:00:01:00 65534    1
1/1/2  lag1(mc)   1002    1        ALFNCD  02:01:00:00:01:00 65534    1


Remote Partner details of all interfaces:
------------------------------------------------------------------------------
Intf   Aggregate  Partner Port     State   System-ID         System   Aggr
       name       Port-id Priority                           Priority Key
------------------------------------------------------------------------------
1/1/1  lag1(mc)   50      1        ALFNCD  02:00:00:00:01:00 65534    1
1/1/2  lag1(mc)   1050    1        ALFNCD  02:00:00:00:01:00 65534    1
```

Now, all ports are **C**ollecting and **D**istributing traffic: LACP flags are **ALFNCD**. Same command can be used on Access/ToR VSX:

TOR-1 / TOR-2

```
TOR-1# show 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  Port    Port     State   System-ID         System   Aggr
        name       id      Priority                           Priority Key
------------------------------------------------------------------------------
1/1/49  lag101(mc) 49      1        ALFNCD  02:00:00:00:01:00 65534    101
1/1/50  lag101(mc) 50      1        ALFNCD  02:00:00:00:01:00 65534    101
```

```
Partner details of all interfaces:
--------------------------------------------------------------------------------
Intf    Aggregate  Partner Port     State   System-ID         System   Aggr
        name       Port-id Priority                          Priority Key
--------------------------------------------------------------------------------
1/1/49  lag101(mc) 1       1        ALFNCD  02:01:00:00:01:00 65534    1
1/1/50  lag101(mc) 1001    1        ALFNCD  02:01:00:00:01:00 65534    1


Remote Actor details of all interfaces:
--------------------------------------------------------------------------------
Intf   Aggregate  Port    Port     State   System-ID         System   Aggr
       name       id      Priority                          Priority Key
--------------------------------------------------------------------------------
1/1/49 lag101(mc) 1049    1        ALFNCD  02:00:00:00:01:00 65534    101
1/1/50 lag101(mc) 1050    1        ALFNCD  02:00:00:00:01:00 65534    101


Remote Partner details of all interfaces:
--------------------------------------------------------------------------------
Intf   Aggregate  Partner Port     State   System-ID         System   Aggr
       name       Port-id Priority                          Priority Key
--------------------------------------------------------------------------------
1/1/49 lag101(mc) 1       1        ALFNCD  02:01:00:00:01:00 65534    1
1/1/50 lag101(mc) 1002    1        ALFNCD  02:01:00:00:01:00 65534    1



TOR-1# show lacp interface

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

A Actor details of all interfaces:
--------------------------------------------------------------------------------
Intf    Aggr       Port  Port  State   System-ID         System Aggr Forwarding
        Name       Id    Pri                             Pri    Key  State
--------------------------------------------------------------------------------
1/1/49  lag101(mc) 49    1     ALFNCD  02:00:00:00:01:00 65534  101  up
1/1/50  lag101(mc) 50    1     ALFNCD  02:00:00:00:01:00 65534  101  up
1/1/55  lag256     56    1     ALFNCD  08:00:09:72:61:c6 65534  256  up
1/1/56  lag256     57    1     ALFNCD  08:00:09:72:61:c6 65534  256  up


Partner details of all interfaces:
--------------------------------------------------------------------------------
Intf    Aggr       Port  Port  State   System-ID         System Aggr
        Name       Id    Pri                             Pri    Key
--------------------------------------------------------------------------------
1/1/49  lag101(mc) 1     1     ALFNCD  02:01:00:00:01:00 65534  1
1/1/50  lag101(mc) 1001  1     ALFNCD  02:01:00:00:01:00 65534  1
1/1/55  lag256     56    1     ALFNCD  08:00:09:7f:8e:c3 65534  256
1/1/56  lag256     57    1     ALFNCD  08:00:09:7f:8e:c3 65534  256
```

Step #12 : Access/ToR - Server VSX LAG configuration

The steps for configuring downstream/server VSX LAG of the TOR layer are very similar to the steps for configuring the upstream VSX LAGs and the same best practices apply. LAG numbering is assigning low number for downstream.

TOR-1(config)#	TOR-2(config)#
interface lag 1 multi-chassis no shutdown description Server-1 no routing vlan trunk native 1 vlan trunk allowed 10,20	interface lag 1 multi-chassis no shutdown synchronized

```
    lacp mode active

interface 1/1/1
    no shutdown
    mtu 9000
    description Server-1-nic1
    lag 1
```

```
interface 1/1/1
    no shutdown
    mtu 9000
    description Server-1-nic2
    lag 1
```

TOR-1 / TOR-2

```
TOR-1# show 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  Port    Port     State   System-ID         System   Aggr
        name       id      Priority                           Priority Key
------------------------------------------------------------------------------
1/1/1   lag1(mc)   1       1        ALFNCD  02:00:00:00:01:00 65534    1
1/1/49  lag101(mc) 49      1        ALFNCD  02:00:00:00:01:00 65534    101
1/1/50  lag101(mc) 50      1        ALFNCD  02:00:00:00:01:00 65534    101


Partner details of all interfaces:
------------------------------------------------------------------------------
Intf    Aggregate  Partner Port     State   System-ID         System   Aggr
        name       Port-id Priority                           Priority Key
------------------------------------------------------------------------------
1/1/1   lag1(mc)   1       1        ALFNCD  80:c1:6e:80:31:81 65534    1
1/1/49  lag101(mc) 1       1        ALFNCD  02:01:00:00:01:00 65534    1
1/1/50  lag101(mc) 1001    1        ALFNCD  02:01:00:00:01:00 65534    1


Remote Actor details of all interfaces:
------------------------------------------------------------------------------
Intf   Aggregate  Port    Port     State   System-ID         System   Aggr
       name       id      Priority                           Priority Key
------------------------------------------------------------------------------
1/1/1  lag1(mc)   1001    1        ALFNCD  02:00:00:00:01:00 65534    1
1/1/49 lag101(mc) 1049    1        ALFNCD  02:00:00:00:01:00 65534    101
1/1/50 lag101(mc) 1050    1        ALFNCD  02:00:00:00:01:00 65534    101


Remote Partner details of all interfaces:
------------------------------------------------------------------------------
Intf   Aggregate  Partner Port     State   System-ID         System   Aggr
       name       Port-id Priority                           Priority Key
------------------------------------------------------------------------------
1/1/1  lag1(mc)   2       1        ALFNCD  80:c1:6e:80:31:81 65534    1
1/1/49 lag101(mc) 1       1        ALFNCD  02:01:00:00:01:00 65534    1
1/1/50 lag101(mc) 1002    1        ALFNCD  02:01:00:00:01:00 65534    1
```

The **best practice for LACP fallback** feature is to enable it on VSX LAG of VSX primary for the following use-cases: PXE boot, server NIC driver migration from active/standby to LACP. LACP fallback is automatically synced on VSX secondary.

TOR-1(config)#

```
interface lag 1 multi-chassis
    lacp fallback

TOR-1# show run
Current configuration:
.................................skipped for readability
interface lag 1 multi-chassis
    no shutdown
    description TOR-VSX-1
    no routing
    vlan trunk native 1
```

TOR-2(config)#

synchronized

```
TOR-2# show run
Current configuration:
.................................skipped for readability
interface lag 1 multi-chassis
    no shutdown
    description TOR-VSX-1
    no routing
    vlan trunk native 1
```

```
    vlan trunk allowed 10,20-30
    lacp mode active
    lacp fallback
```

```
    vlan trunk allowed 10,20-30
    lacp mode active
    lacp fallback
```

Further on in this section, lacp fallback is no longer shown as this is reserved for the previous indicated use-cases.

Step #13 : MSTP configuration

The **best practice on Aggregation** layer are:

- No loop-protect (MSTP used instead).
- Use the **default common instance 0**: MST0
- Lower the **spanning-tree priority to 4** to make VSX aggregation the STP root bridge (easier for support)
- Use **root-guard** on all downlinks to prevent any access switches from becoming Root Bridge.
- Keep the default **port-type admin-network**
- Let VSX secondary synchronized by vsx-sync process.

AGG-1(config)#

```
spanning-tree
spanning-tree priority 4
spanning-tree trap topology-change instance 0

interface lag 1 multi-chassis
    no shutdown
    description Access-Switch-1 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20-30
    lacp mode active
    spanning-tree root-guard
```

synchronized

synchronized

AGG-2(config)#

```
spanning-tree
spanning-tree priority 4
spanning-tree trap topology-change instance 0
```

AGG-1 / AGG-2

```
AGG-1# show run
Current configuration:
.....................................skipped for readability
spanning-tree
spanning-tree priority 4
spanning-tree trap topology-change instance 0
.....................................skipped for readability
interface lag 1 multi-chassis
    no shutdown
    description Access-Switch-1 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20-30
    lacp mode active
    spanning-tree root-guard
.....................................skipped for readability
```

```
AGG-2# show run
Current configuration:
.....................................skipped for readability
spanning-tree
spanning-tree priority 4
spanning-tree trap topology-change instance 0
.....................................skipped for readability
interface lag 1 multi-chassis
    no shutdown
    description Access-Switch-1 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20-30
    lacp mode active
    spanning-tree root-guard
.....................................skipped for readability
```

AGG-1 / AGG-2

```
AGG-1# show spanning-tree mst 0 int lag1 detail
Port lag1
Port Type : admin-network    Loop Guard  : disable
Link Type : point_to_point   BPDU Filter : disable
Boundary  : internal         BPDU Guard  : disable
Root Guard: enable

Instance       Role           State        Cost       Priority   Vlans mapped
-------------- -------------- ------------ ---------- ---------- ----------
0              Designated     Forwarding   20000      64         1-4094

Port lag1
Designated root address            : 02:01:00:00:01:00
Designated regional root address   : 02:01:00:00:01:00
Designated bridge address          : 02:01:00:00:01:00  Priority : 16384
Multi-Chassis role                 : active
Timers:    Message expires in 0 sec, Forward delay expiry:18, Forward transitions:1
```

```
Bpdus sent 19295, received 2
TCN_Tx: 4, TCN_Rx: 2
```

The **best practice on Access/ToR** layer are:

- Use loop-protect for all endpoint access ports (not configured on uplinks). Set the re-enable timer to 1hour.
- Keep the **default common instance 0**: MST0
- Keep the **default spanning-tree priority** of 8.
- All endpoint access ports are **admin-edge**, should not receive any BPDU (**BDPU guard**), should not trigger any Topology Change Notification (**tcn-guard**).
- Use loop-protection on all endpoint access ports as extra-protection mechanism (in case of MSTP BPDUs are filtered by insertion of unmanaged switches which create a loop).
- Use **loop-guard** on all uplinks to prevent any flood due to failure of BPDU reception (fiber strand cut).

TOR-1(config)#	TOR-2(config)#
spanning-tree spanning-tree trap topology-change instance 0 loop-protect re-enable-timer 3600	synchronized
interface lag 101 multi-chassis no shutdown description AGG VSX no routing vlan trunk native 1 vlan trunk allowed 10,20 lacp mode active spanning-tree loop-guard	synchronized
interface lag 1 multi-chassis no shutdown description Server-1 no routing vlan trunk native 1 vlan trunk allowed 10,20 lacp mode active loop-protect spanning-tree bpdu-guard spanning-tree tcn-guard spanning-tree port-type admin-edge	synchronized
TOR-1# show run Current configuration: .....skipped for readability loop-protect re-enable-timer 3600 spanning-tree spanning-tree trap topology-change instance 0 .....skipped for readability interface lag 1 multi-chassis no shutdown description Server-1 no routing vlan trunk native 1 vlan trunk allowed 10,20 lacp mode active loop-protect spanning-tree bpdu-guard spanning-tree tcn-guard spanning-tree port-type admin-edge interface lag 101 multi-chassis no shutdown description AGG VSX no routing vlan trunk native 1 vlan trunk allowed 10,20	TOR-2# show run Current configuration: .....skipped for readability loop-protect re-enable-timer 3600 **spanning-tree** **spanning-tree trap topology-change instance 0** .....skipped for readability interface lag 1 multi-chassis no shutdown description Server-1 no routing vlan trunk native 1 vlan trunk allowed 10,20 lacp mode active loop-protect spanning-tree bpdu-guard spanning-tree tcn-guard spanning-tree port-type admin-edge interface lag 101 multi-chassis no shutdown description AGG VSX no routing vlan trunk native 1 vlan trunk allowed 10,20

```
    lacp mode active                                    |    lacp mode active
    spanning-tree loop-guard                            |    spanning-tree loop-guard
.................................skipped for readability | .................................skipped for readability
```

TOR-1 / TOR-2

```
TOR-1# show spanning-tree mst 0 interface lag101 detail
Port lag101
Port Type : admin-network     Loop Guard  : enable
Link Type : point_to_point    BPDU Filter : disable
Boundary  : internal          BPDU Guard  : disable
Root Guard: disable

Instance       Role           State        Cost       Priority   Vlans mapped
-------------- -------------- ------------ ---------- ---------- ----------
0              Root           Forwarding   20000      64         1-4094

Port lag101
Designated root address             : 02:01:00:00:01:00
Designated regional root address    : 02:00:00:00:01:00
Designated bridge address           : 02:01:00:00:01:00  Priority : 32768
Multi-Chassis role                  : active
Timers:    Message expires in 0 sec, Forward delay expiry:18, Forward transitions:1
Bpdus sent 5, received 13918
TCN_Tx: 4, TCN_Rx: 2


TOR-1# show spanning-tree mst 0 interface lag1 detail
Port lag1
Port Type : admin-edge        Loop Guard  : disable
Link Type : point_to_point    BPDU Filter : disable
Boundary  : internal          BPDU Guard  : enable
Root Guard: disable

Instance       Role           State        Cost       Priority   Vlans mapped
-------------- -------------- ------------ ---------- ---------- ----------
0              Designated     Forwarding   20000      64         1-4094

Port lag1
Designated root address             : 02:01:00:00:01:00
Designated regional root address    : 02:00:00:00:01:00
Designated bridge address           : 02:01:00:00:01:00  Priority : 32768
Multi-Chassis role                  : active
Timers:    Message expires in 1 sec, Forward delay expiry:18, Forward transitions:2
Bpdus sent 13557, received 515
TCN_Tx: 4, TCN_Rx: 2
```

Step #14 : VSX LAG ACL configuration

Please refer to Step#11 of first scenario (Aggregation VSX with single VRF).

Step #15: VSX LAG QoS configuration

The aggregation switch leverage the global configuration of qos trust dscp. No further configuration is needed as this was already set in step#1. To perform marking at the access layer, please refer to the QoS guide.

Step #16: SVI (VLAN L3 interface) configuration

Please refer to the Step#13 of the first scenario (Aggregation VSX with single VRF).

Step #17: OSPF configuration

Please refer to the Step#14 of the first scenario (Aggregation VSX with single VRF).

Step #18: BGP configuration

Please refer to the Step#15 of the first scenario (Aggregation VSX with single VRF).

Step #19: Multicast configuration

Please refer to the Step#16 of the first scenario (Aggregation VSX with single VRF).

VSX Maintenance and Troubleshooting

VSX show commands

All traditional commands like show interface, show ip route (etc..) are usual commands when troubleshooting. Out of the numerous VSX show commands, here are below the main useful ones.

- "show vsx brief" is the most important command that provides visibility on both ISL state and keepalive.

AGG-1

```
AGG-1# show vsx brief
ISL State                              : In-Sync
Device State                           : Peer-Established
Keepalive State                        : Keepalive-Established
Device Role                            : primary
Number of Multi-chassis LAG interfaces : 1
```

- "show vsx status" is the second most important one as it gives extra synchronization information (config-sync, NAE, API).

AGG-1

```
AGG-1# show vsx status
VSX Operational State
---------------------
  ISL channel             : In-Sync
  ISL mgmt channel        : operational
  Config Sync Status      : in-sync
  NAE                     : peer_reachable
  HTTPS Server            : peer_reachable

Attribute           Local               Peer
-----------         --------            --------
ISL link            lag256              lag256
ISL version         2                   2
System MAC          02:01:00:00:01:00   02:01:00:00:01:00
Platform            8325                8325
Software Version    GL.10.04.0001       GL.10.04.0001
Device Role         primary             secondary
```

- "show vsx status config-sync" can be useful to reveal any configuration mismatch between both nodes.

AGG-1

```
AGG-1# sh vsx status config-sync
Admin state             : Enabled
Operational State       : Operational
Error State             : None
Recommended remediation : N/A
Current time            : Fri Dec 13 12:23:44 2019
Last sync time          : Fri Dec 13 11:08:21 2019


8325-1# show vsx status config-sync
Admin state             : Enabled
Operational State       : Operational
Error State             : Missing reference error
Recommended remediation : A. Execute 'show running-config vsx-sync peer-diff' to determine which lines did not sync
correctly on secondary
                          B. Identify the configuration that is missing from secondary and manually fix via CLI.
Current time            : Fri Dec 13 13:29:32 2019
Last sync time          : Not available
```

- "show running-config vsx-sync peer-diff" is very useful to report the configuration lines creating difference error.

AGG-1

```
AGG-1# show run vsx-sync peer-diff
--- /tmp/running-config-vsx.4de1        2018-05-25 14:55:54.956878984 +0200
+++ /tmp/peer-running-config-vsx.4de1   2018-05-25 14:55:54.951879155 +0200
@@ -1,6 +1,6 @@
 interface lag 1
     vsx-sync vlans
-    description ISL
```

```
+     description isl
      no shutdown
      no routing
      vlan trunk native 1 tag
```

- "show events –d vsx-syncd" command on VSX secondary is an additional source of information to identify mistakes.

AGG-2

```
AGG-2# show events -d vsx-syncd
--------------------------------------------------
Event logs from current boot
--------------------------------------------------
2019-12-11T10:25:23.258398+00:00 AGG-2 vsx-syncd[29481]: Event|7602|LOG_INFO|AMM|-|Configuration sync update :
connected to peer's database
2019-12-11T10:25:23.318891+00:00 AGG-2 vsx-syncd[29481]: Event|7602|LOG_INFO|AMM|-|Configuration sync update :
Initiated configuration sync process
2019-12-11T10:40:00.345125+00:00 AGG-2 vsx-syncd[29481]: Event|7602|LOG_INFO|AMM|-|Configuration sync update : VSX
configuration-sync updated database
2019-12-12T14:13:56.747085+00:00 AGG-2 vsx-syncd[29481]: Event|7601|LOG_ERR|AMM|-|Configuration sync error : Missing
reference in database while syncing configuration. Use "show running-config vsx-sync peer-diff" to help identify the
missing reference.
2019-12-12T14:14:00.380789+00:00 AGG-2 vsx-syncd[29481]: Event|7602|LOG_INFO|AMM|-|Configuration sync update : Missing
Reference Error: Could not find Port 1%2F1%2F3 on secondary VSX database.
```

- "show vsx config-consistency" can be used after the VSX cluster configuration to check all VLANs, STP, and LAG parameters.

AGG-1

```
AGG-1# show vsx config-consistency
Configurations                          Local                                   Peer
------------------                      ------                                  ------
Software Version                        GL.10.04.0001                           GL.10.04.0001
System MAC                              02:01:00:00:01:00                       02:01:00:00:01:00
System Profile                          Basic                                   Basic
ISL hello interval                      1                                       1
ISL dead interval                       20                                      20
ISL hold interval                       0                                       0
Keepalive hello interval                1                                       1
Keepalive dead interval                 3                                       3
Keepalive UDP port                      7678                                    7678

VSX VLAN List
-------------
Local ISL VLANs : 1-2 10 20-30
Peer ISL VLANs : 1-2 10 20-30

VSX Active Forwarding
---------------------
Interface VLANs      : None
Peer Interface VLANs : None

STP Configurations                      Local                                   Peer
------------------                      ------                                  ------
STP Enabled                             Yes                                     Yes
STP Mode                                mstp                                    mstp
MST hello time(in seconds)              2                                       2
MST maximum age(in seconds)             20                                      20
MST maximum hops                        20                                      20
MST Config Name                         02:01:00:00:01:00                       02:01:00:00:01:00
MST Config Revision                     0                                       0
MST Config Digest                       AC36177F50283CD4B83821D8AB26DE62        AC36177F50283CD4B83821D8AB26DE62
MST number of instances                 0                                       0

RPVST VLAN List:
----------------
Local:
Peer :



AGG-1# show vsx config-consistency lacp
Configurations                          Local               Peer
```

```
------------------                          ------              ------

Name                                        lag1                lag1
Loop protect enabled                        false               false
Hash scheme                                 l3-src-dst          l3-src-dst
Qos dscp
Qos cos
Qos trust

VSX VLAN list
10 20-30
Peer VSX VLAN list
10 20-30

STP link-type                               point_to_point      point_to_point
STP port-type                               admin-network       admin-network
STP bpdu-filter                             Disabled            Disabled
STP bpdu-guard                              Disabled            Disabled
STP loop-guard                              Disabled            Disabled
STP root-guard                              Enabled             Enabled
STP tcn-guard                               Disabled            Disabled
```

- "show vsx status linkup-delay" can be useful right after the unit joining the VSX cluster as there is a delay before ports that are members of VSX LAGs become forwarding.

AGG-2

```
AGG-2# show vsx status linkup-delay
Configured linkup delay-timer                                  : 180 seconds
Initial sync status                                            : Completed
Delay timer status                                             : Completed
Linkup Delay time left                                         :
Interfaces that will be brought up after delay timer expires :
Interfaces that are excluded from delay timer                  :
```

- "show lacp interfaces" is the command that provides immediate view of LAGs and VSX LAGs health.

AGG-1

```
AGG-1# show lacp interfaces

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                             Pri    Key  State
------------------------------------------------------------------------------
1/1/1   lag1(mc)   1     1     ALFNCD  02:01:00:00:01:00 65534  1    up
1/1/55  lag256     56    1     ALFNCD  08:00:09:b0:c4:aa 65534  256  up
1/1/56  lag256     57    1     ALFNCD  08:00:09:b0:c4:aa 65534  256  up


Partner details of all interfaces:
------------------------------------------------------------------------------
Intf    Aggr       Port  Port  State   System-ID         System Aggr
        Name       Id    Pri                             Pri    Key
------------------------------------------------------------------------------
1/1/1   lag1(mc)   1     1     ALFNCD  08:00:09:72:61:c6 65534  1
1/1/55  lag256     56    1     ALFNCD  08:00:09:9a:9b:10 65534  256
1/1/56  lag256     57    1     ALFNCD  08:00:09:9a:9b:10 65534  256
```

- "show lacp interfaces multi-chassis" command is very useful to get a complete status of local LACP partnership as well as the VSX peer partnership details. Actor = local node, Partner = LACP neighbor (the access switch), Remote Actor = the VSX peer, Remote Partner = LACP neighbor of the VSX peer. Note that the port id of the VSX secondary is equal to 1000+ID_of_the_primary (in the example 1001). ALFNCD LACP state-flags should appear on all entries.

AGG-1

```
AGG-1# show 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  Port    Port     State   System-ID         System   Aggr
        name       id      Priority                           Priority Key
------------------------------------------------------------------------------
1/1/1   lag1(mc)   1       1        ALFNCD  02:01:00:00:01:00 65534    1


Partner details of all interfaces:
------------------------------------------------------------------------------
Intf    Aggregate  Partner Port     State   System-ID         System   Aggr
        name       Port-id Priority                           Priority Key
------------------------------------------------------------------------------
1/1/1   lag1(mc)   6       1        ALFNCD  08:00:09:72:61:c6 65534    1


Remote Actor details of all interfaces:
------------------------------------------------------------------------------
Intf   Aggregate  Port    Port     State   System-ID         System   Aggr
       name       id      Priority                           Priority Key
------------------------------------------------------------------------------
1/1/1  lag1(mc)   1001    1        ALFNCD  02:01:00:00:01:00 65534    1


Remote Partner details of all interfaces:
------------------------------------------------------------------------------
Intf   Aggregate  Partner Port     State   System-ID         System   Aggr
       name       Port-id Priority                           Priority Key
------------------------------------------------------------------------------
1/1/1  lag1(mc)   7       1        ALFNCD  08:00:09:72:61:c6 65534    1
```

- "show vsx mac-address-table" is used to check that all MAC addresses are properly synchronized between the VSX primary and the VSX secondary.

AGG-1

```
AGG-1# show vsx mac-address-table

MAC Address          VLAN     Type       LocalPort  PeerPort
------------------------------------------------------------
0c:88:77:1b:41:01    10       dynamic    lag1       lag1
00:50:79:66:68:00    10       dynamic    lag1       lag1
0c:88:77:67:57:03    10       dynamic    lag1       lag1
08:00:09:9a:9b:10    21       dynamic    lag256
08:00:09:9a:9b:10    2        dynamic    lag256
08:00:09:9a:9b:10    20       dynamic    lag256
08:00:09:9a:9b:10    23       dynamic    lag256
08:00:09:9a:9b:10    24       dynamic    lag256
08:00:09:9a:9b:10    30       dynamic    lag256
08:00:09:9a:9b:10    29       dynamic    lag256
08:00:09:9a:9b:10    22       dynamic    lag256
08:00:09:9a:9b:10    26       dynamic    lag256
08:00:09:9a:9b:10    1        dynamic    lag256
08:00:09:9a:9b:10    10       dynamic    lag256
08:00:09:9a:9b:10    25       dynamic    lag256
08:00:09:9a:9b:10    27       dynamic    lag256
08:00:09:9a:9b:10    28       dynamic    lag256
08:00:09:b0:c4:aa    24       dynamic               lag256
08:00:09:b0:c4:aa    10       dynamic               lag256
08:00:09:b0:c4:aa    26       dynamic               lag256
08:00:09:b0:c4:aa    28       dynamic               lag256
08:00:09:b0:c4:aa    22       dynamic               lag256
```

```
08:00:09:b0:c4:aa    25       dynamic              lag256
08:00:09:b0:c4:aa    29       dynamic              lag256
08:00:09:b0:c4:aa    21       dynamic              lag256
08:00:09:b0:c4:aa    30       dynamic              lag256
08:00:09:b0:c4:aa    1        dynamic              lag256
08:00:09:b0:c4:aa    20       dynamic              lag256
08:00:09:b0:c4:aa    2        dynamic              lag256
08:00:09:b0:c4:aa    23       dynamic              lag256
08:00:09:b0:c4:aa    27       dynamic              lag256
```

- "show ip interface vlan.." is useful to check VSX active-gateway configuration.

AGG-1

```
AGG-1# show ip interface vlan10

Interface vlan10 is up
 Admin state is up
 Hardware: Ethernet, MAC Address: 08:00:09:b0:c4:aa
 IP MTU 9100
 IPv4 address 10.1.10.2/24
 IPv4 address 10.2.10.2/24 secondary
 active-gateway ip mac 12:01:00:00:01:00
 active-gateway ip 10.1.10.1
 active-gateway ip 10.2.10.1
 L3 Counters: Rx Disabled, Tx Disabled
 Rx
          ucast: 0 packets, 0 bytes
          mcast: 0 packets, 0 bytes
 Tx
          ucast: 0 packets, 0 bytes
          mcast: 0 packets, 0 bytes
```

- "show vsx ip route" is very useful to check the routing table of the cluster, i.e. aggregated view of both VSX nodes.

AGG-1

```
AGG-1# show vsx ip route 0.0.0.0

IPv4 Forwarding Routes

'[x/y]' denotes [distance/metric]

0.0.0.0/0, vrf default
     via  10.0.0.2,  [110/1],  ospf on AGG-1
     via  10.0.0.0,  [110/1],  ospf on AGG-1
     via  10.0.0.6,  [110/1],  ospf on AGG-2
     via  10.0.0.4,  [110/1],  ospf on AGG-2
```

- "**show vsx ip data-path**" provides an aggregated view of the data-path for the given route on the cluster. Very useful.

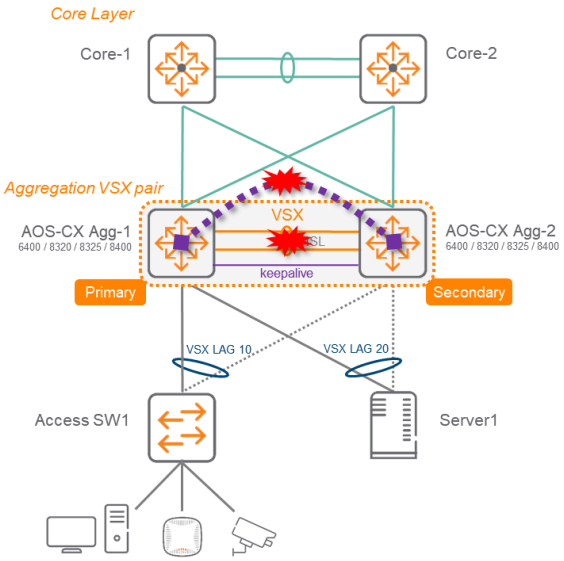
AGG-1

```
AGG-1# show vsx ip data-path 0.0.0.0
IPv4 Data Path Information For 0.0.0.0

Local Device
------------
Route : 0.0.0.0/0
    Egress L3 Interface : 1/1/49

    Egress L3 Interface : 1/1/50

Peer Device
------------
Route : 0.0.0.0/0
    Egress L3 Interface : 1/1/49

    Egress L3 Interface : 1/1/50
```

VSX Split

Split brain detection

Here is a summary of split brain detection scenario:



	Keepalive-Established	Keepalive-Init
ISL "In-Sync"	**Forwarding: OK** **Protection: OK**	**Forwarding: OK** **Protection: NO**
ISL "Out-of-Sync"	**STATE** • ISL is down • Peer is detected and UP **ISSUE** • VSX tables get out of sync **ACTION** • On secondary only: stop forwarding over VSX LAGs that are "Disabled by feature" as well as associated ports	**SPLIT BRAIN** **STATE** • ISL is down • Peer is seen DOWN **ISSUE** • Traffic drop **ACTION** • Both out-of-synced nodes keep forwarding

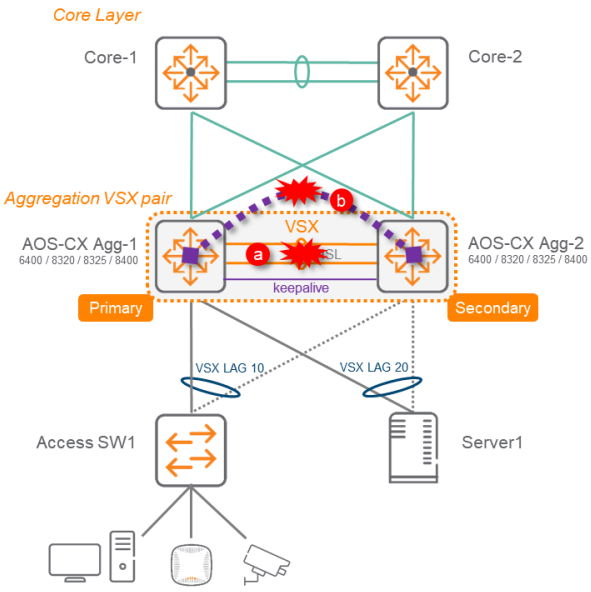
Note: There is no reboot required on VSX secondary when the ISL link is restored and the secondary is joining back the cluster. The VSX secondary LAGs and associated physical ports are brought up after the synchronization time (maximum time being the linkup-delay timer).

The ISL cut has no effect on the state (up/down) of VLAN and associated SVI that are not part of any VSX LAG but part of at least one orphan port.

During ISL cut (before initial sync), if the VSX Secondary node has at least one port that is a member of a VSX LAG then the associated SVI of the VLAN transported by the said VSX LAG is turned OFF/SHUT on the VSX Secondary node, whether or not there is an orphan port carrying that given VLAN.

Split brain

They are two cases that cannot be differentiated: VSX failure (ex: power outage) and ISL+keepalive interruption. In such situation the network administrator would have to choose the strategy: split-recovery ON or OFF. The best practice is to keep the default split-recovery ON as the probability to hit a power outage on VSX primary is higher than a cut of all paths for ISL and keepalive. The table below describes in details the behavior of the VSX secondary.



Failure Scenarios	Split-recovery off	Split-recovery on (default)
b) Keepalive down. ISL up.	No impact. (but loss of split detection)	No impact. (but loss of split detection)
a) ISL down. Keepalive up.	Secondary VSX node tears down VSX LAG member ports	Secondary VSX node tears down VSX LAG member ports
a) ISL down. Keepalive up. b) Then, after sometime[1], keepalive down as well.	Secondary VSX node tears down VSX LAG member ports. Secondary VSX LAGs stays down.	a) Secondary VSX node tears down VSX LAG member ports b) Secondary VSX node restores VSX LAG member ports.
a+b) At the same time[2], ISL down and keepalive down. b) Keepalive restore	a+b) All VSX LAG ports stay up. b) Secondary VSX node tears down VSX LAG member ports	a+b) All VSX LAG ports stay up. b) Secondary VSX node tears down VSX LAG member ports

[1]: enough time for split to be detected between ISL cut and keepalive down events (between 0sec and one hellotime).
[2]: ISL cut and keepalive down events are closed enough so there is no possibility to detect a split (like a power-off).

Switch replacement in the VSX Cluster

To replace the VSX primary or the VSX secondary, follow this steps:

- Make sure all cables are labelled with clear identification
- Unplug power and all fibers, copper cables
- Un-rack the failing unit and rack the replacement unit.
- Power-up the unit.
- Restore switch firmware and configuration and shutdown all ports.
  To perform this task, there are several options. Here is the recommended option using NetEdit:
  - Plug the USB BT dongle on the new switch and plug the OOBM port to a temporary network which can access NetEdit server.
  - Use the CX Mobile application to on-board the replacement unit as a new switch on NetEdit.
  - From NetEdit, upgrade or downgrade the firmware of the new switch to the same version than the running unit of the VSX cluster.
  - In NetEdit, select the last known good configuration of the failing unit, and create a plan based on this configuration for a new deployment on the new unit. Make sure that the OOBM interface will be still accessible with the new configuration (it could be the previous IP of the failing switch as long as it is accessible from NetEdit). Deploy the plan. Commit. (Commit will include saving the running configuration in the start-up configuration).
  - SSH to the replacement switch and shutdown all ports:
    For 8320, 8325: interface 1/1/1-1/1/52 or 56
    shutdown
    For 6400/8400: range can not take all the line cards, so it has to be done per line cards.
  - Once all ports are shutdown, do not save the configuration.

- Move all the transceivers from the old switch to the new switch.
- Connect back all fibers or copper cables to the replacement switch.
- From SSH session to the new swtich, perform the following command:
  "checkpoint rollback startup-config"
- All non VSX LAG ports should turn ON and the unit should immediately join the VSX cluster. After a synchronization time, the VSX LAG ports should forward traffic.
- From NetEdit, if needed, modify the OOBM configuration, and unplug the OOBM port if required. This modification should trigger the switch being red-flagged in NetEdit as the IP address changed. Remove the temporary switch from NetEdit. The initial failing unit on NetEdit should turn green again.
- Save configuration.
- Situation is back to nominal.

VSX Live Upgrade

VSX upgrade can be performed without traffic interruption. The vsx update-software command will orchestrate the full upgrade of the cluster. Here is a quick overview of the steps being performed by this orchestration command:

1. Start two parallel TFTP sessions to download the new firmware to VSX primary and VSX secondary simultaneously
2. Verify firmware and store new firmware on both switches
3. Send graceful shutdown for OSPF, BGP, VRRP and LACP from the VSX secondary and perform traffic redirection to the VSX primary.
4. Reboot secondary
5. After reboot, the secondary joins the VSX cluster with the new version. After HW tables synchronization, the secondary forwards traffic.
6. Once VSX secondary is in steady state, send graceful shutdown for OSPF, BGP, VRRP, LACP from the VSX primary and perform traffic redirection to the VSX secondary.
7. Reboot primary
8. After reboot, the primary joins the VSX cluster with the new version. After HW tables synchronization, the primary forwards traffic.
9. Upgrade is completed.

To proceed with upgrade of the VSX cluster, enter the following command from SSH session:

AGG-1# vsx update-software tftp://<*ip_address*>/*software_name*.swi vrf mgmt

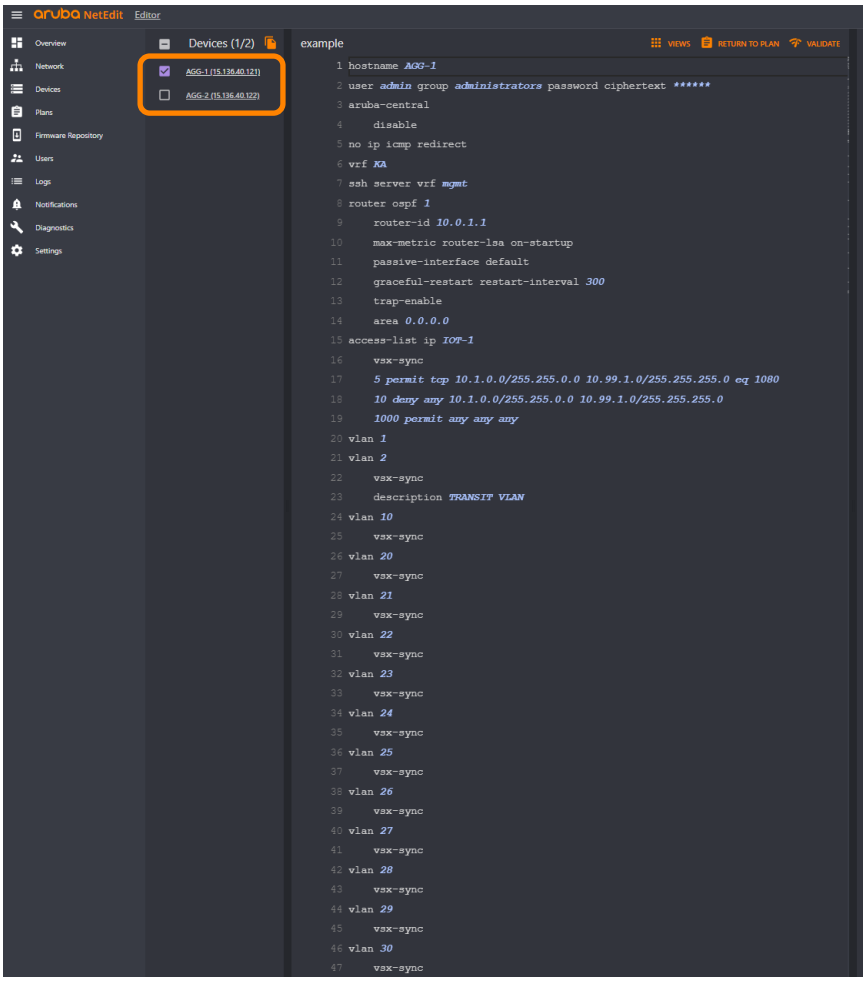
Here we use the OOBM Management VRF. It can be any in-band management VRF including default, as long as the TFTP server is accessible from this VRF.

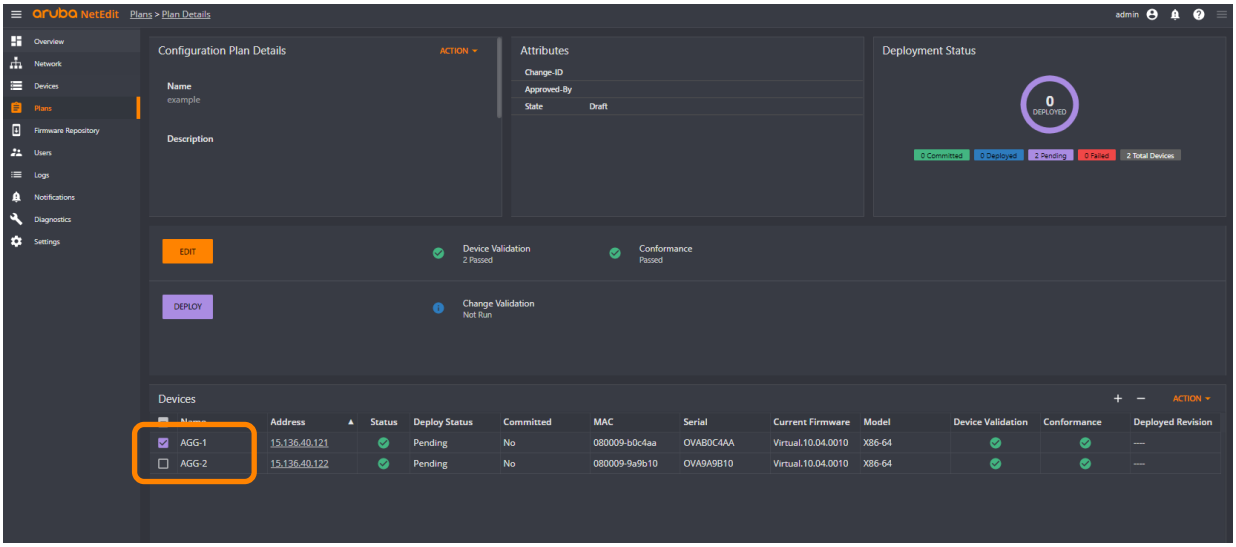
Cumulated impact of the upgrade should be sub-second.

NetEdit

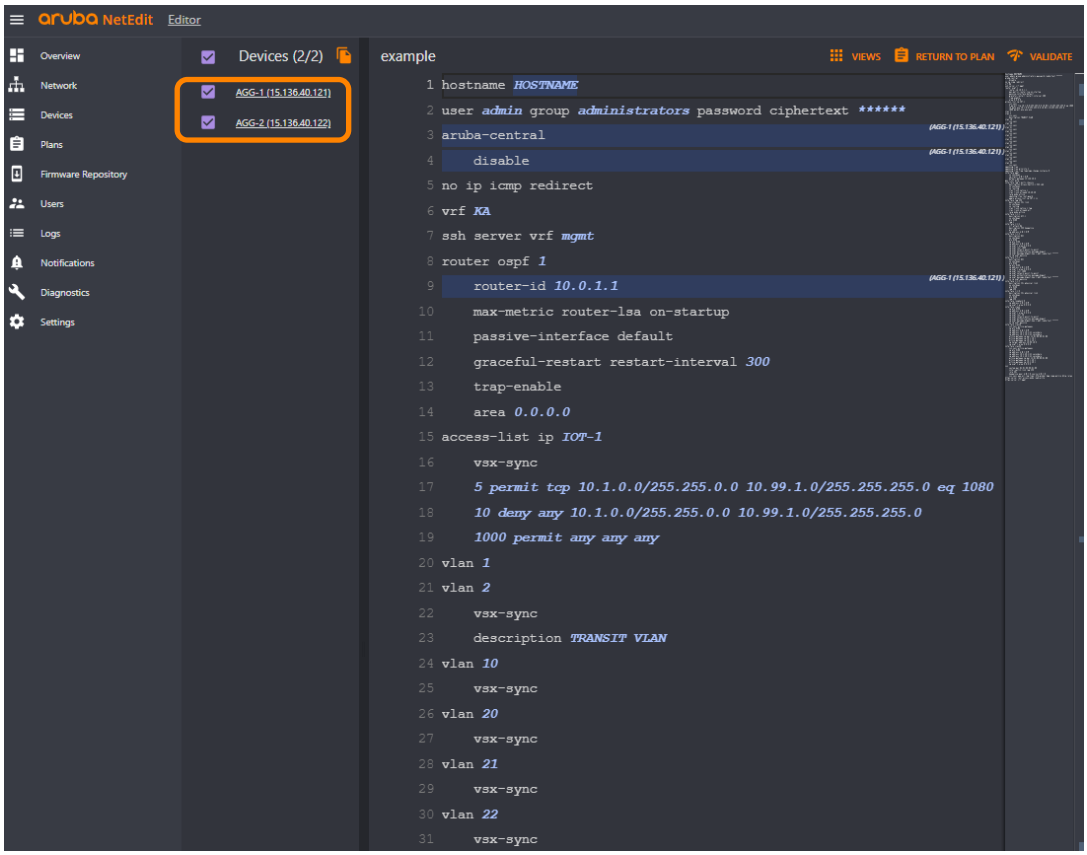
For proper operation with NetEdit and vsx-sync synchronization, the best practice during a NetEdit configuration change is to proceed in 2 phases:

1. Configure with NEtEdit all the parameters that will be synchronized with vsx-sync on the VSX primary only and deploy only on the VSX primary:





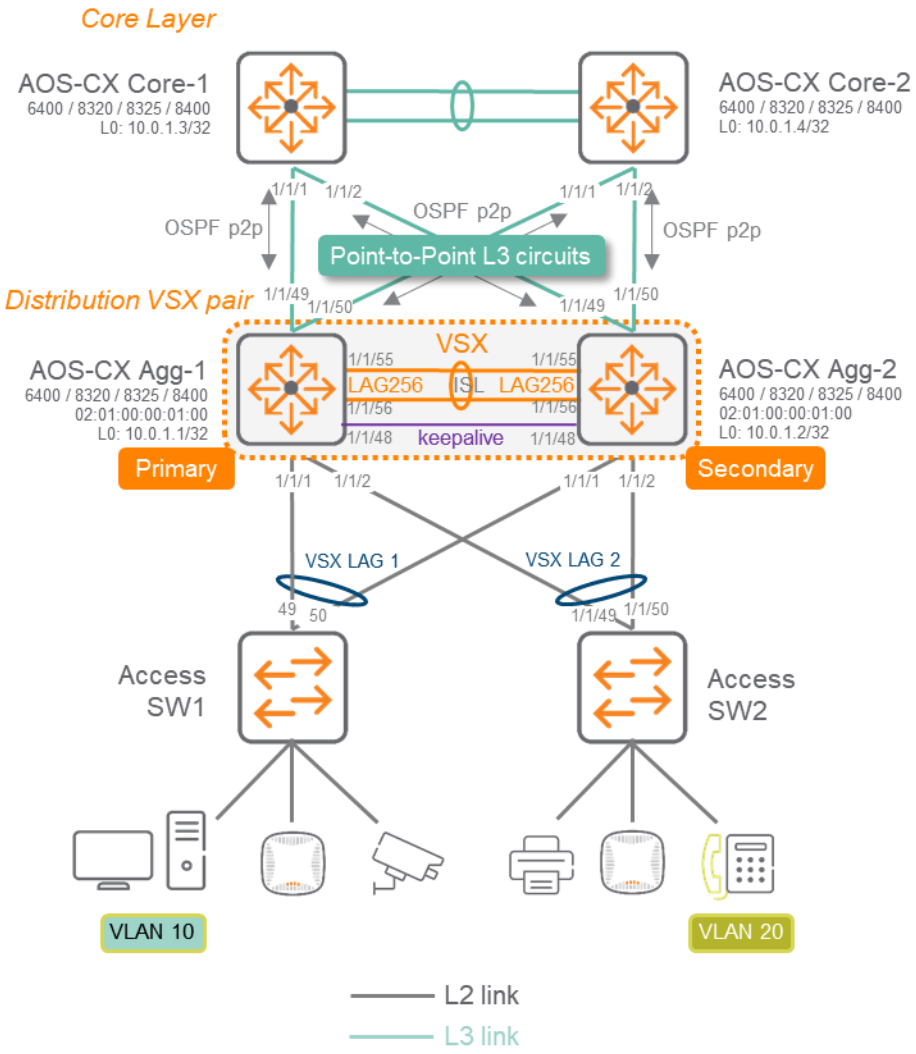
2. After VSX synchronization, return to the plan and select both VSX nodes and continue configuration for items that are not synchronized by vsx-sync.



Then deploy on both and commit.

APPENDIX A – Aggregation VSX with single VRF routing model – Configuration example

Topology



The following configuration examples do not include all the other best practices that are recommended for other aspects like management or authentication as the focus is the VSX configuration and the associated impacts. These other best practices are described in the Campus Validated Reference Design document.. The following examples provide only the extract of the configuration that is required for VSX deployment best practices.

Access Switch Configuration

Access Switch-1 : AOS-S (2930)

```
hostname "ACC-1"
jumbo max-frame-size 9122
trunk 49-50 Trk1 lacp
spanning-tree
spanning-tree bpdu-protection-timeout 3600
spanning-tree Trk1 loop-guard
spanning-tree 1 admin-edge-port
spanning-tree 1 tcn-guard bpdu-protection
```

```
spanning-tree 2 admin-edge-port
spanning-tree 2 tcn-guard bpdu-protection
…
spanning-tree 48 admin-edge-port
spanning-tree 48 tcn-guard bpdu-protection
loop-protect 1-48
loop-protect disable-timer 3600
vlan 1
   name "DEFAULT_VLAN"
   no untagged Trk1
   untagged 1-48,51-52
   no ip address
   exit
vlan 10
   name "VLAN10"
   tagged 1,Trk1
   no ip address
   jumbo
   exit
vlan 20
   name "VLAN20"
   tagged 5,Trk1
   no ip address
   jumbo
   exit
```

Access Switch-2 : AOS-CX (6300)

```
hostname ACC-2
loop-protect re-enable-timer 3600
!
vlan 1,10,20
spanning-tree
interface lag 1
    no shutdown
    description UPLINK to AGG
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20
    lacp mode active
    spanning-tree loop-guard
interface 1/1/1
    no shutdown
    mtu 9000
    description Endpoint1
    no routing
    vlan access 10
    spanning-tree bpdu-guard
    spanning-tree port-type admin-edge
    spanning-tree tcn-guard
    loop-protect
interface 1/1/2
    no shutdown
    mtu 9000
    description Endpoint2
    no routing
    vlan access 10
    spanning-tree bpdu-guard
    spanning-tree port-type admin-edge
    spanning-tree tcn-guard
    loop-protect
interface 1/1/49
    no shutdown
    mtu 9100
    lag 1
interface 1/1/50
    no shutdown
    mtu 9100
    lag 1
```

Aggregation Switch configuration

AGG-1

```
hostname AGG-1
no ip icmp redirect
vrf KA
!
router ospf 1
    router-id 10.0.1.1
    max-metric router-lsa on-startup
    passive-interface default
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
access-list ip IOT-1
    vsx-sync
    !
    5 permit tcp 10.1.0.0/255.255.0.0 10.99.1.0/255.255.255.0 eq 1080
    10 deny any 10.1.0.0/255.255.0.0 10.99.1.0/255.255.255.0
    1000 permit any any any
vlan 1
vlan 2
    vsx-sync
    description TRANSIT VLAN
vlan 10
    vsx-sync
vlan 20
    vsx-sync
vlan 21
    vsx-sync
vlan 22
    vsx-sync
vlan 23
    vsx-sync
vlan 24
    vsx-sync
vlan 25
    vsx-sync
vlan 26
    vsx-sync
vlan 27
    vsx-sync
vlan 28
    vsx-sync
vlan 29
    vsx-sync
vlan 30
    vsx-sync
spanning-tree
spanning-tree priority 4
spanning-tree trap topology-change instance 0
qos trust dscp
interface lag 1 multi-chassis
    no shutdown
    description Access-Switch-1 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20-30
    lacp mode active
    spanning-tree root-guard
    apply access-list ip IOT-1 in
interface lag 256
    no shutdown
    description ISL link
    no routing
    vlan trunk native 1 tag
    vlan trunk allowed all
    lacp mode active
interface 1/1/1
    no shutdown
```

```
    mtu 9100
    description ACC-1
    lag 1
interface 1/1/48
    no shutdown
    vrf attach KA
    description VSX keepalive
    ip address 192.168.0.0/31
interface 1/1/49
    no shutdown
    mtu 9198
    description CORE-1 1/1/1
    ip mtu 9198
    ip address 10.0.0.1/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface 1/1/50
    no shutdown
    mtu 9198
    description CORE-2 1/1/1
    ip mtu 9198
    ip address 10.0.0.3/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface 1/1/55
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
interface 1/1/56
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
interface loopback 0
    ip address 10.0.1.1/32
    ip ospf 1 area 0.0.0.0
interface vlan2
    ip mtu 9198
    ip address 10.0.2.1/30
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan10
    vsx-sync active-gateways
    ip mtu 9100
    ip address 10.1.10.2/24
    ip address 10.2.10.2/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.10.1
    active-gateway ip 10.2.10.1
    ip helper-address 10.99.10.9
    ip ospf 1 area 0.0.0.0
interface vlan20
    vsx-sync active-gateways
    ip mtu 9100
    ip address 10.1.20.2/24
    ip address 10.2.20.2/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.20.1
```

```
    active-gateway ip 10.2.20.1
    ip ospf 1 area 0.0.0.0
vsx
    system-mac 02:01:00:00:01:00
    inter-switch-link lag 256
    role primary
    keepalive peer 192.168.0.1 source 192.168.0.0 vrf KA
    vsx-sync aaa acl-log-timer bfd-global bgp copp-policy dhcp-relay dhcp-server dhcp-snooping dns icmp-tcp lldp loop-
protect-global mac-lockout mclag-interfaces neighbor ospf qos-global route-map sflow-global snmp ssh stp-global time vsx-
global
```

AGG-2

```
hostname AGG-2
no ip icmp redirect
vrf KA
!
router ospf 1
    router-id 10.0.1.2
    max-metric router-lsa on-startup
    passive-interface default
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
access-list ip IOT-1
    vsx-sync
    !
    5 permit tcp 10.1.0.0/255.255.0.0 10.99.1.0/255.255.255.0 eq 1080
    10 deny any 10.1.0.0/255.255.0.0 10.99.1.0/255.255.255.0
    1000 permit any any any
vlan 1
vlan 2
    vsx-sync
    description TRANSIT VLAN
vlan 10
    vsx-sync
vlan 20
    vsx-sync
vlan 21
    vsx-sync
vlan 22
    vsx-sync
vlan 23
    vsx-sync
vlan 24
    vsx-sync
vlan 25
    vsx-sync
vlan 26
    vsx-sync
vlan 27
    vsx-sync
vlan 28
    vsx-sync
vlan 29
    vsx-sync
vlan 30
    vsx-sync
spanning-tree
spanning-tree priority 4
spanning-tree trap topology-change instance 0
qos trust dscp
interface lag 1 multi-chassis
    no shutdown
    description Access-Switch-1 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20-30
    lacp mode active
    spanning-tree root-guard
    apply access-list ip IOT-1 in
interface lag 256
```

```
    no shutdown
    description ISL link
    no routing
    vlan trunk native 1 tag
    vlan trunk allowed all
    lacp mode active
interface 1/1/1
    no shutdown
    mtu 9100
    description ACC-1
    lag 1
interface 1/1/48
    no shutdown
    vrf attach KA
    description VSX keepalive
    ip address 192.168.0.1/31
interface 1/1/49
    no shutdown
    mtu 9198
    description CORE-1 1/1/2
    ip mtu 9198
    ip address 10.0.0.5/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface 1/1/50
    no shutdown
    mtu 9198
    description CORE-2 1/1/2
    ip mtu 9198
    ip address 10.0.0.7/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface 1/1/55
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
interface 1/1/56
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
interface loopback 0
    ip address 10.0.1.2/32
    ip ospf 1 area 0.0.0.0
interface vlan2
    ip mtu 9198
    ip address 10.0.2.2/30
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan10
    vsx-sync active-gateways
    ip mtu 9100
    ip address 10.1.10.3/24
    ip address 10.2.10.3/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.10.1
    active-gateway ip 10.2.10.1
    ip helper-address 10.99.10.9
```

```
    ip ospf 1 area 0.0.0.0
interface vlan20
    vsx-sync active-gateways
    ip mtu 9100
    ip address 10.1.20.3/24
    ip address 10.2.20.3/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.20.1
    active-gateway ip 10.2.20.1
    ip ospf 1 area 0.0.0.0
vsx
    system-mac 02:01:00:00:01:00
    inter-switch-link lag 256
    role secondary
    keepalive peer 192.168.0.0 source 192.168.0.1 vrf KA
    vsx-sync aaa acl-log-timer bfd-global bgp copp-policy dhcp-relay dhcp-server dhcp-snooping dns icmp-tcp lldp loop-
protect-global mac-lockout mclag-interfaces neighbor ospf qos-global route-map sflow-global snmp ssh stp-global time vsx-
global
```

Core Switch Configuration

The configuration of the core layer partially reflects the topology (the routed LAG between Core-1/2 is replaced by a single interface 1/1/3)

Core-1

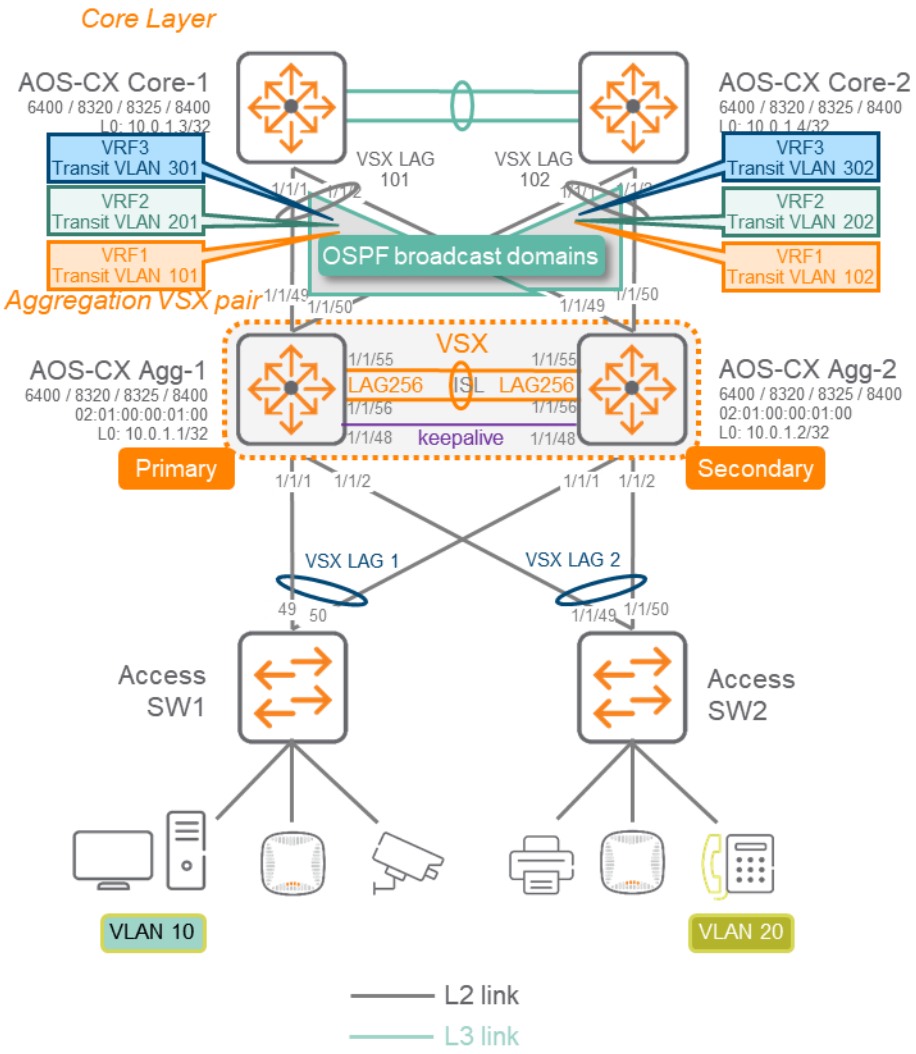
```
hostname CORE-1
!
router ospf 1
    router-id 10.0.1.3
    max-metric router-lsa on-startup
    passive-interface default
    default-information originate always
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
interface 1/1/1
    no shutdown
    mtu 9198
    ip mtu 9198
    ip address 10.0.0.0/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface 1/1/2
    no shutdown
    mtu 9198
    ip mtu 9198
    ip address 10.0.0.4/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface 1/1/3
    no shutdown
    mtu 9198
    ip mtu 9198
    ip address 10.0.0.252/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface loopback 0
    ip address 10.0.1.3/32
    ip ospf 1 area 0.0.0.0
```

Core-2

```
hostname CORE-2
!
router ospf 1
    router-id 10.0.1.4
    max-metric router-lsa on-startup
    passive-interface default
    default-information originate always
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
interface 1/1/1
    no shutdown
    mtu 9198
    ip mtu 9198
    ip address 10.0.0.2/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface 1/1/2
    no shutdown
    mtu 9198
    ip mtu 9198
    ip address 10.0.0.6/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface 1/1/3
    no shutdown
    mtu 9198
    ip mtu 9198
    ip address 10.0.0.253/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface loopback 0
    ip address 10.0.1.4/32
    ip ospf 1 area 0.0.0.0
```

APPENDIX B – Aggregation VSX with multiple VRF routing model – Configuration example

Topology



The following configuration examples do not include all the other best practices that are recommended for other aspects like management or authentication as the focus is the VSX configuration and the associated impacts. These other best practices are decribed in the Campus Validated Reference Design document.. The following examples provide only the extract of the configuration that is required for VSX deployment best practices.

Access Switch Configuration

Access Switch-1 : AOS-S (2930)

```
hostname "ACC-1"
jumbo max-frame-size 9122
trunk 49-50 Trk1 lacp
spanning-tree
spanning-tree bpdu-protection-timeout 3600
spanning-tree Trk1 loop-guard
spanning-tree 1 admin-edge-port
spanning-tree 1 tcn-guard bpdu-protection
```

```
spanning-tree 2 admin-edge-port
spanning-tree 2 tcn-guard bpdu-protection
…
spanning-tree 48 admin-edge-port
spanning-tree 48 tcn-guard bpdu-protection
loop-protect 1-48
loop-protect disable-timer 3600
vlan 1
   name "DEFAULT_VLAN"
   no untagged Trk1
   untagged 1-48,51-52
   no ip address
   exit
vlan 10
   name "VLAN10"
   tagged 1,Trk1
   no ip address
   jumbo
   exit
vlan 20
   name "VLAN20"
   tagged 5,Trk1
   no ip address
   jumbo
   exit
```

Access Switch-2 : AOS-CX (6300)

```
hostname ACC-2
loop-protect re-enable-timer 3600
!
vlan 1,10,20
spanning-tree
interface lag 1
    no shutdown
    description UPLINK to AGG
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20
    lacp mode active
    spanning-tree loop-guard
interface 1/1/1
    no shutdown
    mtu 9000
    description Endpoint1
    no routing
    vlan access 10
    spanning-tree bpdu-guard
    spanning-tree port-type admin-edge
    spanning-tree tcn-guard
    loop-protect
interface 1/1/2
    no shutdown
    mtu 9000
    description Endpoint2
    no routing
    vlan access 10
    spanning-tree bpdu-guard
    spanning-tree port-type admin-edge
    spanning-tree tcn-guard
    loop-protect
interface 1/1/49
    no shutdown
    mtu 9100
    lag 1
interface 1/1/50
    no shutdown
    mtu 9100
    lag 1
```

Aggregation Switch configuration

AGG-1

```
hostname AGG-1
no ip icmp redirect
vrf KA
vrf VRF1
vrf VRF2
!
router ospf 1 vrf VRF1
    router-id 10.0.1.1
    max-metric router-lsa on-startup
    passive-interface default
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
router ospf 1 vrf VRF2
    router-id 10.0.1.1
    max-metric router-lsa on-startup
    passive-interface default
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
access-list ip IOT-1
    vsx-sync
    !
    5 permit tcp 10.1.0.0/255.255.0.0 10.99.1.0/255.255.255.0 eq 1080
    10 deny any 10.1.0.0/255.255.0.0 10.99.1.0/255.255.255.0
    1000 permit any any any
vlan 1
vlan 10
    vsx-sync
vlan 20
    vsx-sync
vlan 21
    vsx-sync
vlan 22
    vsx-sync
vlan 23
    vsx-sync
vlan 24
    vsx-sync
vlan 25
    vsx-sync
vlan 26
    vsx-sync
vlan 27
    vsx-sync
vlan 28
    vsx-sync
vlan 29
    vsx-sync
vlan 30
    vsx-sync
vlan 101
    vsx-sync
    description TRANSIT VLAN VRF1-CORE1
vlan 102
    vsx-sync
    description TRANSIT VLAN VRF1-CORE2
vlan 201
    vsx-sync
    description TRANSIT VLAN VRF2-CORE1
vlan 202
    vsx-sync
    description TRANSIT VLAN VRF2-CORE2
spanning-tree
spanning-tree priority 4
spanning-tree trap topology-change instance 0
qos trust dscp
```

```
interface lag 1 multi-chassis
    no shutdown
    description Access-Switch-1 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20-30
    lacp mode active
    spanning-tree root-guard
    apply access-list ip IOT-1 in
interface lag 101 multi-chassis
    no shutdown
    description CORE-1 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 101,201
    lacp mode active
    spanning-tree bpdu-filter
interface lag 102 multi-chassis
    no shutdown
    description CORE-2 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 102,202
    lacp mode active
    spanning-tree bpdu-filter
interface lag 256
    no shutdown
    description ISL link
    no routing
    vlan trunk native 1 tag
    vlan trunk allowed all
    lacp mode active
interface 1/1/1
    no shutdown
    mtu 9100
    description ACC-1
    lag 1
interface 1/1/48
    no shutdown
    vrf attach KA
    description VSX keepalive
    ip address 192.168.0.0/31
interface 1/1/49
    no shutdown
    mtu 9198
    description CORE-1 1/1/1
    lag 101
interface 1/1/50
    no shutdown
    mtu 9198
    description CORE-2 1/1/1
    lag 102
interface 1/1/55
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
interface 1/1/56
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
interface loopback 0
    vrf attach VRF1
    ip address 10.0.1.1/32
    ip ospf 1 area 0.0.0.0
interface vlan10
    vsx-sync active-gateways
    vrf attach VRF1
    ip mtu 9100
    ip address 10.1.10.2/24
```

```
    ip address 10.2.10.2/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.10.1
    active-gateway ip 10.2.10.1
    ip helper-address 10.99.10.9
    ip ospf 1 area 0.0.0.0
interface vlan20
    vsx-sync active-gateways
    vrf attach VRF2
    ip mtu 9100
    ip address 10.1.20.2/24
    ip address 10.2.20.2/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.20.1
    active-gateway ip 10.2.20.1
    ip ospf 1 area 0.0.0.0
interface vlan101
    vrf attach VRF1
    ip mtu 9198
    vsx active-forwarding
    ip address 10.0.101.2/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan102
    vrf attach VRF1
    ip mtu 9198
    vsx active-forwarding
    ip address 10.0.102.2/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan201
    vrf attach VRF2
    ip mtu 9198
    vsx active-forwarding
    ip address 10.0.201.2/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan202
    vrf attach VRF2
    ip mtu 9198
    vsx active-forwarding
    ip address 10.0.202.2/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
vsx
    system-mac 02:01:00:00:01:00
    inter-switch-link lag 256
    role primary
    keepalive peer 192.168.0.1 source 192.168.0.0 vrf KA
    linkup-delay-timer exclude lag 101-102
    vsx-sync aaa acl-log-timer bfd-global bgp copp-policy dhcp-relay dhcp-server dhcp-snooping dns icmp-tcp lldp loop-
protect-global mac-lockout mclag-interfaces neighbor ospf qos-global route-map sflow-global snmp ssh stp-global time vsx-
global
```

AGG-2

```
hostname AGG-2
no ip icmp redirect
vrf KA
vrf VRF1
```

```
vrf VRF2
!
router ospf 1 vrf VRF1
    router-id 10.0.1.2
    max-metric router-lsa on-startup
    passive-interface default
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
router ospf 1 vrf VRF2
    router-id 10.0.1.2
    max-metric router-lsa on-startup
    passive-interface default
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
access-list ip IOT-1
    vsx-sync
    !
    5 permit tcp 10.1.0.0/255.255.0.0 10.99.1.0/255.255.255.0 eq 1080
    10 deny any 10.1.0.0/255.255.0.0 10.99.1.0/255.255.255.0
    1000 permit any any any
vlan 1
vlan 10
    vsx-sync
vlan 20
    vsx-sync
vlan 21
    vsx-sync
vlan 22
    vsx-sync
vlan 23
    vsx-sync
vlan 24
    vsx-sync
vlan 25
    vsx-sync
vlan 26
    vsx-sync
vlan 27
    vsx-sync
vlan 28
    vsx-sync
vlan 29
    vsx-sync
vlan 30
    vsx-sync
vlan 101
    vsx-sync
    description TRANSIT VLAN VRF1-CORE1
vlan 102
    vsx-sync
    description TRANSIT VLAN VRF1-CORE2
vlan 201
    vsx-sync
    description TRANSIT VLAN VRF2-CORE1
vlan 202
    vsx-sync
    description TRANSIT VLAN VRF2-CORE2
spanning-tree
spanning-tree priority 4
spanning-tree trap topology-change instance 0
qos trust dscp
interface lag 1 multi-chassis
    no shutdown
    description Access-Switch-1 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20-30
    lacp mode active
    spanning-tree root-guard
```

```
    apply access-list ip IOT-1 in
interface lag 101 multi-chassis
    no shutdown
    description CORE-1 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 101,201
    lacp mode active
    spanning-tree bpdu-filter
interface lag 102 multi-chassis
    no shutdown
    description CORE-2 VSX LAG
    no routing
    vlan trunk native 1
    vlan trunk allowed 102,202
    lacp mode active
    spanning-tree bpdu-filter
interface lag 256
    no shutdown
    description ISL link
    no routing
    vlan trunk native 1 tag
    vlan trunk allowed all
    lacp mode active
interface 1/1/1
    no shutdown
    mtu 9100
    description ACC-1
    lag 1
interface 1/1/48
    no shutdown
    vrf attach KA
    description VSX keepalive
    ip address 192.168.0.1/31
interface 1/1/49
    no shutdown
    mtu 9198
    description CORE-1 1/1/2
    lag 101
interface 1/1/50
    no shutdown
    mtu 9198
    description CORE-2 1/1/2
    lag 102
interface 1/1/55
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
interface 1/1/56
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
interface loopback 0
    vrf attach VRF1
    ip address 10.0.1.2/32
    ip ospf 1 area 0.0.0.0
interface vlan10
    vsx-sync active-gateways
    vrf attach VRF1
    ip mtu 9100
    ip address 10.1.10.3/24
    ip address 10.2.10.3/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.10.1
    active-gateway ip 10.2.10.1
    ip helper-address 10.99.10.9
    ip ospf 1 area 0.0.0.0
interface vlan20
    vsx-sync active-gateways
```

```
    vrf attach VRF2
    ip mtu 9100
    ip address 10.1.20.3/24
    ip address 10.2.20.3/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.20.1
    active-gateway ip 10.2.20.1
    ip ospf 1 area 0.0.0.0
interface vlan101
    vrf attach VRF1
    ip mtu 9198
    vsx active-forwarding
    ip address 10.0.101.3/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan102
    vrf attach VRF1
    ip mtu 9198
    vsx active-forwarding
    ip address 10.0.102.3/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan201
    vrf attach VRF2
    ip mtu 9198
    vsx active-forwarding
    ip address 10.0.201.3/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan202
    vrf attach VRF2
    ip mtu 9198
    vsx active-forwarding
    ip address 10.0.202.3/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
vsx
    system-mac 02:01:00:00:01:00
    inter-switch-link lag 256
    role secondary
    keepalive peer 192.168.0.0 source 192.168.0.1 vrf KA
    linkup-delay-timer exclude lag 101-102
    vsx-sync aaa acl-log-timer bfd-global bgp copp-policy dhcp-relay dhcp-server dhcp-snooping dns icmp-tcp lldp loop-
protect-global mac-lockout mclag-interfaces neighbor ospf qos-global route-map sflow-global snmp ssh stp-global time vsx-
global
```

Core Switch Configuration

The configuration of the core layer partially reflects the topology (the routed LAG between Core-1/2 is replaced by a single interface 1/1/3).

Core-1

```
hostname CORE-1
vrf VRF1
vrf VRF2
!
router ospf 1 vrf VRF1
    router-id 10.0.1.3
    max-metric router-lsa on-startup
    passive-interface default
```

```
    default-information originate always
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
router ospf 1 vrf VRF2
    router-id 10.0.1.3
    max-metric router-lsa on-startup
    passive-interface default
    default-information originate always
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
vlan 11
    description Transit CORE VRF1
vlan 12
    description Transit CORE VRF2
vlan 101
    description TRANSIT VLAN VRF1-CORE1
vlan 201
    description TRANSIT VLAN VRF2-CORE1
interface lag 1
    no shutdown
    no routing
    vlan trunk native 1
    vlan trunk allowed 101,201
    lacp mode active
interface 1/1/1
    no shutdown
    mtu 9198
    lag 1
interface 1/1/2
    no shutdown
    mtu 9198
    lag 1
interface 1/1/3
    no shutdown
    mtu 9198
    no routing
    vlan trunk native 1
    vlan trunk allowed 11-12
interface loopback 0
    vrf attach VRF1
    ip address 10.0.1.3/32
    ip ospf 1 area 0.0.0.0
interface vlan11
    vrf attach VRF1
    ip mtu 9198
    ip address 10.0.11.1/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 20
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan12
    vrf attach VRF2
    ip mtu 9198
    ip address 10.0.12.1/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 20
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan101
    vrf attach VRF1
    ip mtu 9198
    ip address 10.0.101.1/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf priority 100
```

```
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan201
    vrf attach VRF2
    ip mtu 9198
    ip address 10.0.201.1/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf priority 100
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
```

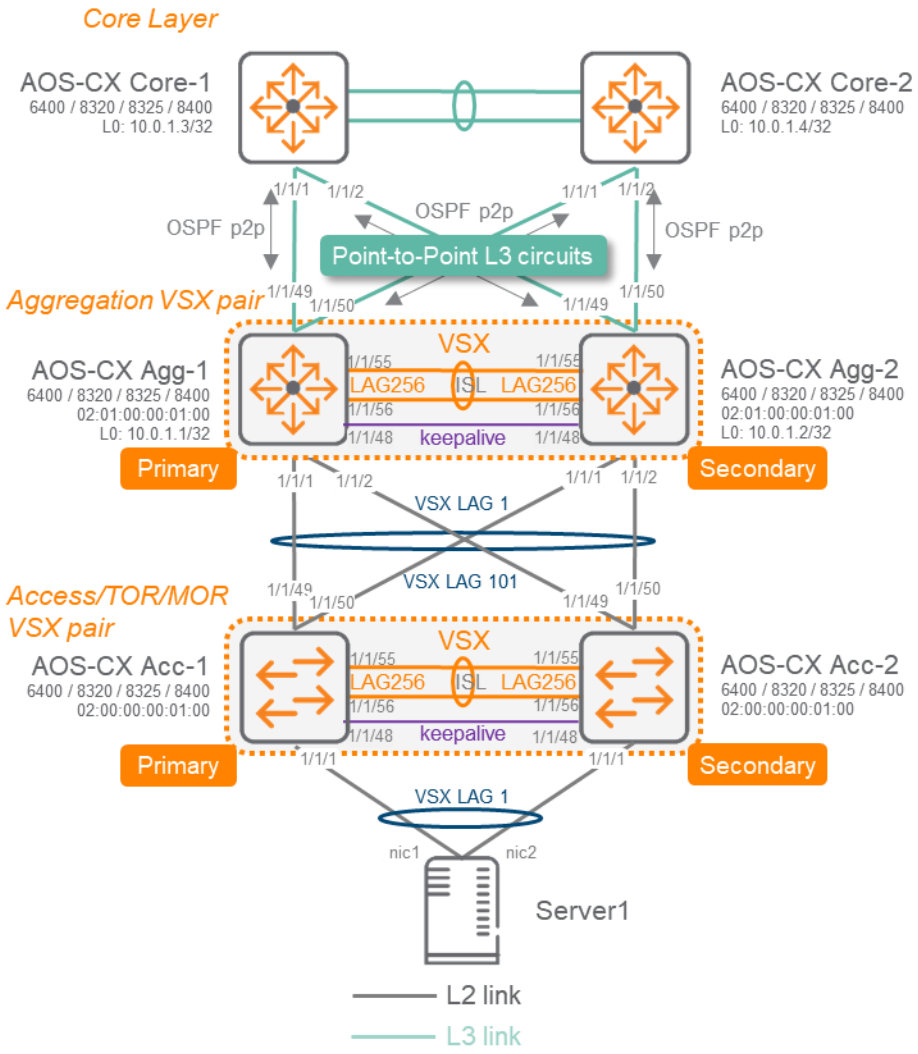
Core-2

```
hostname CORE-2
vrf VRF1
vrf VRF2
!
router ospf 1 vrf VRF1
    router-id 10.0.1.4
    max-metric router-lsa on-startup
    passive-interface default
    default-information originate always
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
router ospf 1 vrf VRF2
    router-id 10.0.1.4
    max-metric router-lsa on-startup
    passive-interface default
    default-information originate always
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
vlan 11
    description Transit CORE VRF1
vlan 12
    description Transit CORE VRF2
vlan 102
    description TRANSIT VLAN VRF1-CORE2
vlan 202
    description TRANSIT VLAN VRF2-CORE2
interface lag 1
    no shutdown
    no routing
    vlan trunk native 1
    vlan trunk allowed 102,202
    lacp mode active
interface 1/1/1
    no shutdown
    mtu 9198
    lag 1
interface 1/1/2
    no shutdown
    mtu 9198
    lag 1
interface 1/1/3
    no shutdown
    mtu 9198
    no routing
    vlan trunk native 1
    vlan trunk allowed 11-12
interface loopback 0
    vrf attach VRF1
    ip address 10.0.1.4/32
    ip ospf 1 area 0.0.0.0
interface vlan11
    vrf attach VRF1
    ip mtu 9198
    ip address 10.0.11.2/29
```

```
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 20
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan12
    vrf attach VRF2
    ip mtu 9198
    ip address 10.0.12.2/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 20
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan102
    vrf attach VRF1
    ip mtu 9198
    ip address 10.0.102.1/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf priority 100
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan202
    vrf attach VRF2
    ip mtu 9198
    ip address 10.0.202.1/29
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf priority 100
    ip ospf cost 50
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
```

APPENDIX C – Access VSX to Aggregation VSX – Configuration example

Topology



The following configuration examples do not include all the other best practices that are recommended for other aspects like management or authentication as the focus is the VSX configuration and the associated impacts. These other best practices are decribed in the Campus Validated Reference Design document.. The following examples provide only the extract of the configuration that is required for VSX deployment best practices.

TOR Switch Configuration

TOR-1

```
hostname TOR-1
loop-protect re-enable-timer 3600
vrf KA
!
vlan 1
vlan 10
    vsx-sync
vlan 20
```

```
    vsx-sync
spanning-tree
spanning-tree trap topology-change instance 0
interface lag 1 multi-chassis
    no shutdown
    description Server-1
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20
    lacp mode active
    loop-protect
    spanning-tree bpdu-guard
    spanning-tree tcn-guard
    spanning-tree port-type admin-edge
interface lag 101 multi-chassis
    no shutdown
    description AGG VSX
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20
    lacp mode active
    spanning-tree loop-guard
interface lag 256
    no shutdown
    description ISL link
    no routing
    vlan trunk native 1 tag
    vlan trunk allowed all
    lacp mode active
interface 1/1/1
    no shutdown
    mtu 9000
    description Server-1-nic1
    lag 1
interface 1/1/48
    no shutdown
    vrf attach KA
    description VSX keepalive
    ip address 192.168.0.2/31
interface 1/1/49
    no shutdown
    mtu 9100
    description AGG-1
    lag 101
interface 1/1/50
    no shutdown
    mtu 9100
    description AGG-2
    lag 101
interface 1/1/55
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
interface 1/1/56
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
vsx
    system-mac 02:00:00:00:01:00
    inter-switch-link lag 256
    role primary
    keepalive peer 192.168.0.3 source 192.168.0.2 vrf KA
    vsx-sync aaa acl-log-timer bfd-global bgp copp-policy dhcp-relay dhcp-server dhcp-snooping dns icmp-tcp lldp loop-
protect-global mac-lockout mclag-interfaces neighbor ospf qos-global route-map sflow-global snmp ssh stp-global time vsx-
global
```

TOR-2

```
hostname TOR-2
loop-protect re-enable-timer 3600
```

```
vrf KA
!
vlan 1
vlan 10
    vsx-sync
vlan 20
    vsx-sync
spanning-tree
spanning-tree trap topology-change instance 0
interface lag 1 multi-chassis
    no shutdown
    description Server-1
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20
    lacp mode active
    loop-protect
    spanning-tree bpdu-guard
    spanning-tree tcn-guard
    spanning-tree port-type admin-edge
interface lag 101 multi-chassis
    no shutdown
    description AGG VSX
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20
    lacp mode active
    spanning-tree loop-guard
interface lag 256
    no shutdown
    description ISL link
    no routing
    vlan trunk native 1 tag
    vlan trunk allowed all
    lacp mode active
interface 1/1/1
    no shutdown
    mtu 9000
    description Server-1-nic2
    lag 1
interface 1/1/48
    no shutdown
    vrf attach KA
    description VSX keepalive
    ip address 192.168.0.3/31
interface 1/1/49
    no shutdown
    mtu 9100
    description AGG-1
    lag 101
interface 1/1/50
    no shutdown
    mtu 9100
    description AGG-2
    lag 101
interface 1/1/55
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
interface 1/1/56
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
vsx
    system-mac 02:00:00:00:01:00
    inter-switch-link lag 256
    role secondary
    keepalive peer 192.168.0.2 source 192.168.0.3 vrf KA
```

```
    vsx-sync aaa acl-log-timer bfd-global bgp copp-policy dhcp-relay dhcp-server dhcp-snooping dns icmp-tcp lldp loop-
protect-global mac-lockout mclag-interfaces neighbor ospf qos-global route-map sflow-global snmp ssh stp-global time vsx-
global
```

Aggregation Switch configuration

AGG-1

```
hostname AGG-1
no ip icmp redirect
vrf KA
!
router ospf 1
    router-id 10.0.1.1
    max-metric router-lsa on-startup
    passive-interface default
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
access-list ip IOT-1
    vsx-sync
    !
    5 permit tcp 10.1.0.0/255.255.0.0 10.99.1.0/255.255.255.0 eq 1080
    10 deny any 10.1.0.0/255.255.0.0 10.99.1.0/255.255.255.0
    1000 permit any any any
vlan 1
vlan 2
    vsx-sync
    description TRANSIT VLAN
vlan 10
    vsx-sync
vlan 20
    vsx-sync
vlan 21
    vsx-sync
vlan 22
    vsx-sync
vlan 23
    vsx-sync
vlan 24
    vsx-sync
vlan 25
    vsx-sync
vlan 26
    vsx-sync
vlan 27
    vsx-sync
vlan 28
    vsx-sync
vlan 29
    vsx-sync
vlan 30
    vsx-sync
spanning-tree
spanning-tree priority 4
spanning-tree trap topology-change instance 0
qos trust dscp
interface lag 1 multi-chassis
    no shutdown
    description TOR-VSX-1
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20-30
    lacp mode active
    spanning-tree root-guard
    apply access-list ip IOT-1 in
interface lag 256
    no shutdown
    description ISL link
    no routing
    vlan trunk native 1 tag
    vlan trunk allowed all
```

```
    lacp mode active
interface 1/1/1
    no shutdown
    mtu 9100
    description TOR-1
    lag 1
interface 1/1/2
    no shutdown
    mtu 9100
    description TOR-2
    lag 1
interface 1/1/48
    no shutdown
    vrf attach KA
    description VSX keepalive
    ip address 192.168.0.0/31
interface 1/1/49
    no shutdown
    mtu 9198
    description CORE-1 1/1/1
    ip mtu 9198
    ip address 10.0.0.1/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface 1/1/50
    no shutdown
    mtu 9198
    description CORE-2 1/1/1
    ip mtu 9198
    ip address 10.0.0.3/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface 1/1/55
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
interface 1/1/56
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
interface loopback 0
    ip address 10.0.1.1/32
    ip ospf 1 area 0.0.0.0
interface vlan2
    ip mtu 9198
    ip address 10.0.2.1/30
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 50
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan10
    vsx-sync active-gateways
    ip mtu 9100
    ip address 10.1.10.2/24
    ip address 10.2.10.2/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.10.1
    active-gateway ip 10.2.10.1
    ip helper-address 10.99.10.9
```

```
    ip ospf 1 area 0.0.0.0
interface vlan20
    vsx-sync active-gateways
    ip mtu 9100
    ip address 10.1.20.2/24
    ip address 10.2.20.2/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.20.1
    active-gateway ip 10.2.20.1
    ip ospf 1 area 0.0.0.0
vsx
    system-mac 02:01:00:00:01:00
    inter-switch-link lag 256
    role primary
    keepalive peer 192.168.0.1 source 192.168.0.0 vrf KA
    vsx-sync aaa acl-log-timer bfd-global bgp copp-policy dhcp-relay dhcp-server dhcp-snooping dns lldp loop-protect-
global mclag-interfaces ospf qos-global route-map sflow-global snmp ssh stp-global time vsx-global
```

AGG-2

```
hostname AGG-2
no ip icmp redirect
vrf KA
!
router ospf 1
    router-id 10.0.1.2
    max-metric router-lsa on-startup
    passive-interface default
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
access-list ip IOT-1
    vsx-sync
    !
    5 permit tcp 10.1.0.0/255.255.0.0 10.99.1.0/255.255.255.0 eq 1080
    10 deny any 10.1.0.0/255.255.0.0 10.99.1.0/255.255.255.0
    1000 permit any any any
vlan 1
vlan 2
    vsx-sync
    description TRANSIT VLAN
vlan 10
    vsx-sync
vlan 20
    vsx-sync
vlan 21
    vsx-sync
vlan 22
    vsx-sync
vlan 23
    vsx-sync
vlan 24
    vsx-sync
vlan 25
    vsx-sync
vlan 26
    vsx-sync
vlan 27
    vsx-sync
vlan 28
    vsx-sync
vlan 29
    vsx-sync
vlan 30
    vsx-sync
spanning-tree
spanning-tree priority 4
spanning-tree trap topology-change instance 0
qos trust dscp
interface lag 1 multi-chassis
```

```
    no shutdown
    description TOR-VSX-1
    no routing
    vlan trunk native 1
    vlan trunk allowed 10,20-30
    lacp mode active
    spanning-tree root-guard
    apply access-list ip IOT-1 in
interface lag 256
    no shutdown
    description ISL link
    no routing
    vlan trunk native 1 tag
    vlan trunk allowed all
    lacp mode active
interface 1/1/1
    no shutdown
    mtu 9100
    description TOR-1
    lag 1
interface 1/1/2
    no shutdown
    mtu 9100
    description TOR-2
    lag 1
interface 1/1/48
    no shutdown
    vrf attach KA
    description VSX keepalive
    ip address 192.168.0.1/31
interface 1/1/49
    no shutdown
    mtu 9198
    description CORE-1 1/1/2
    ip mtu 9198
    ip address 10.0.0.5/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface 1/1/50
    no shutdown
    mtu 9198
    description CORE-2 1/1/2
    ip mtu 9198
    ip address 10.0.0.7/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface 1/1/55
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
interface 1/1/56
    no shutdown
    mtu 9198
    description ISL physical link
    lag 256
interface loopback 0
    ip address 10.0.1.2/32
    ip ospf 1 area 0.0.0.0
interface vlan2
    ip mtu 9198
    ip address 10.0.2.2/30
    ip ospf 1 area 0.0.0.0
```

```
    no ip ospf passive
    ip ospf cost 50
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface vlan10
    vsx-sync active-gateways
    ip mtu 9100
    ip address 10.1.10.3/24
    ip address 10.2.10.3/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.10.1
    active-gateway ip 10.2.10.1
    ip helper-address 10.99.10.9
    ip ospf 1 area 0.0.0.0
interface vlan20
    vsx-sync active-gateways
    ip mtu 9100
    ip address 10.1.20.3/24
    ip address 10.2.20.3/24 secondary
    active-gateway ip mac 12:01:00:00:01:00
    active-gateway ip 10.1.20.1
    active-gateway ip 10.2.20.1
    ip ospf 1 area 0.0.0.0
vsx
    system-mac 02:01:00:00:01:00
    inter-switch-link lag 256
    role secondary
    keepalive peer 192.168.0.0 source 192.168.0.1 vrf KA
    vsx-sync aaa acl-log-timer bfd-global bgp copp-policy dhcp-relay dhcp-server dhcp-snooping dns lldp loop-protect-
global mclag-interfaces ospf qos-global route-map sflow-global snmp ssh stp-global time vsx-global
```

Core Switch Configuration

The configuration of the core layer partially reflects the topology (the routed LAG between Core-1/2 is replaced by a single interface 1/1/3)

Core-1

```
hostname CORE-1
!
router ospf 1
    router-id 10.0.1.3
    max-metric router-lsa on-startup
    passive-interface default
    default-information originate always
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
vlan 1
interface 1/1/1
    no shutdown
    mtu 9198
    ip mtu 9198
    ip address 10.0.0.0/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface 1/1/2
    no shutdown
    mtu 9198
    ip mtu 9198
    ip address 10.0.0.4/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface 1/1/3
```

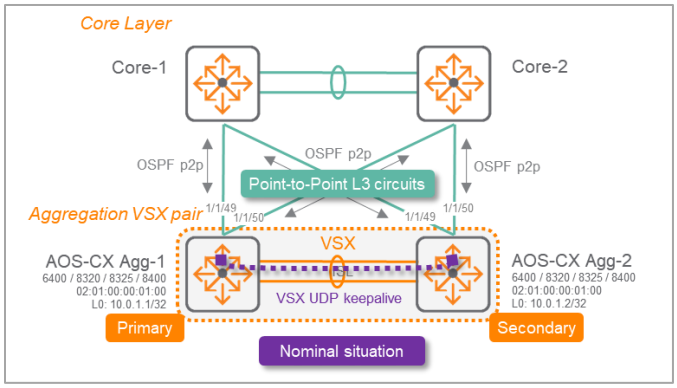
```
    no shutdown
    mtu 9198
    ip mtu 9198
    ip address 10.0.0.252/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface loopback 0
    ip address 10.0.1.3/32
    ip ospf 1 area 0.0.0.0
```

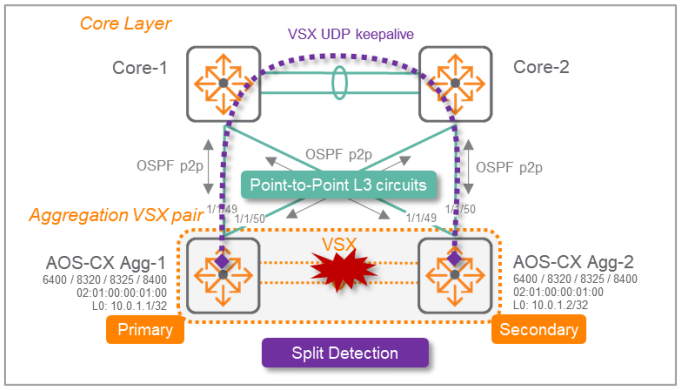
Core-2

```
hostname CORE-2
!
router ospf 1
    router-id 10.0.1.4
    max-metric router-lsa on-startup
    passive-interface default
    default-information originate always
    graceful-restart restart-interval 300
    trap-enable
    area 0.0.0.0
vlan 1
interface 1/1/1
    no shutdown
    mtu 9198
    ip mtu 9198
    ip address 10.0.0.2/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface 1/1/2
    no shutdown
    mtu 9198
    ip mtu 9198
    ip address 10.0.0.6/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface 1/1/3
    no shutdown
    mtu 9198
    ip mtu 9198
    ip address 10.0.0.253/31
    ip ospf 1 area 0.0.0.0
    no ip ospf passive
    ip ospf cost 1000
    ip ospf network point-to-point
    ip ospf authentication message-digest
    ip ospf message-digest-key 1 md5 ciphertext AQBapcc35qrr0SnZBBka0Zgl7scoOzf9+wPnYW36nvk3HA5oBQAAAJb1UbLX
interface loopback 0
    ip address 10.0.1.4/32
    ip ospf 1 area 0.0.0.0
```

APPENDIX D – VSX keepalive over upstream L3 Core nodes

Using a direct circuit for VSX keepalive can be expensive when using 40G or 100G ports on equipment like Aruba 8325 32 ports 40/100G or Aruba 8320 32 ports 40G. To minimize the cost impact of a keepalive dedicated circuit with such switches, it may be preferred to allow the VSX keepalive to get established over the L3 upstream network.





In such scenario, VSX UDP keepalive source and destination IP addresses are the loopback IP addresses. Please note that the keepalive UDP port can be changed to accommodate any security access-list or firewall rules that the UDP probe would cross. Here is the associated configuration:

AGG-1(config)#

```
interface loopback 0
    ip address 10.0.1.2/32
    ip ospf 1 area 0.0.0.0

vsx
    keepalive peer 10.0.1.2 source 10.0.1.1
    keepalive udp-port ?
  <1024-65535>  UDP port (Default:7678)
```

AGG-2(config)#

```
interface loopback 0
    ip address 10.0.1.2/32
    ip ospf 1 area 0.0.0.0

vsx
    keepalive peer 10.0.1.1 source 10.0.1.2
```

AGG-1 / AGG-2

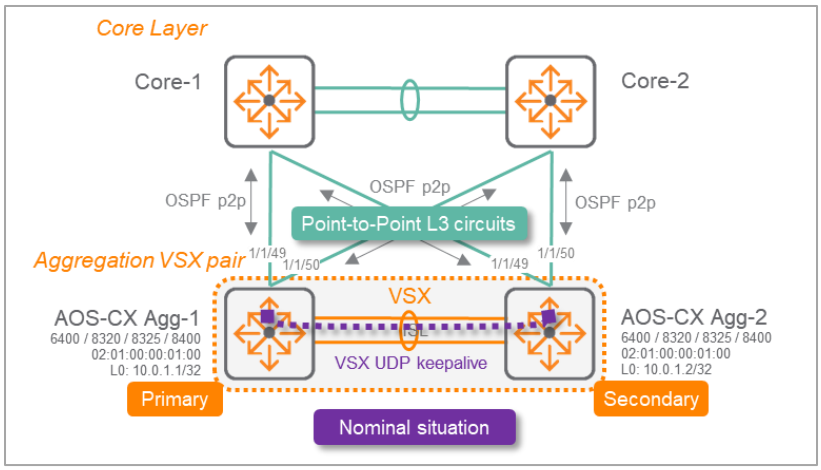
```
AGG-1# show vsx brief
ISL State                              : In-Sync
Device State                           : Peer-Established
Keepalive State                        : Keepalive-Established
Device Role                            : primary
Number of Multi-chassis LAG interfaces : 1

AGG-2# show vsx status keepalive
Keepalive State           : Keepalive-Established
Last Established          : Tue Dec 17 17:38:48 2019
Last Failed               : Tue Dec 17 17:38:26 2019
Peer System Id            : 02:01:00:00:01:00
Peer Device Role          : secondary

Keepalive Counters
Keepalive Packets Tx      : 186
Keepalive Packets Rx      : 161
Keepalive Timeouts        : 0
Keepalive Packets Dropped : 0
```

For such use-case, the following points must be carefully considered:

- No dedicated VRF is used in this scenario. Indeed, introducing a KeepAlive VRF would lead to change the routing model to the scenario where uplinks are VSX LAGs (for obvious Transit VLANs optimization). So the default VRF is used for keepalive traffic like in the configuration example given above.
- In nominal situation, the VSX keepalive UDP traffic goes over the Transit VLAN carried by the ISL. between the VSX nodes.

- In case of ISL failure, it is critical that the traffic between loopbacks can be established over an alternate path in less time than the VSX keepalive dead interval (by default 3 seconds). Otherwise, both ISL and keepalive would be cut which would lead to a cluster split brain. There are several options to establish that alternate path but the simpler configuration and Best Practice is to let OSPF convergence providing that alternate path over the L3 upstream routing domain.

Nominal Situation

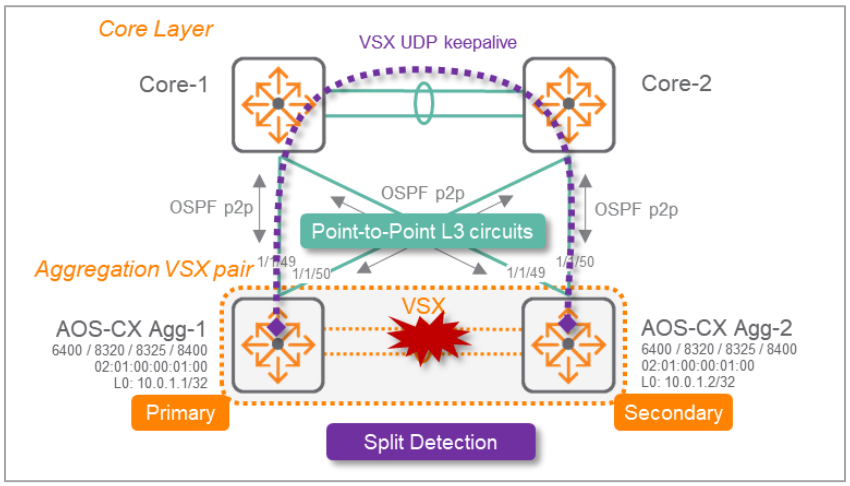


AGG-1	AGG-2
AGG-1# show ip route 10.0.1.2 **Displaying ipv4 routes selected for forwarding** **'[x/y]' denotes [distance/metric]** **10.0.1.2/32, vrf default** **via 10.0.2.2, [110/50], ospf** **AGG-1# show vsx ip data-path 10.0.1.2/32** **IPv4 Data Path Information For 10.0.1.2/32** **Local Device** **------------** **Route : 10.0.1.2/32** **Egress L3 Interface : vlan2** **Peer Device** **------------** **Route : 10.0.1.2/32** **Egress L3 Interface : loopback0** **AGG-1# show vsx brief** **ISL State : In-Sync** **Device State : Peer-Established** **Keepalive State : Keepalive-Established** **Device Role : primary** **Number of Multi-chassis LAG interfaces : 1**	AGG-2# show ip rou 10.0.1.1 Displaying ipv4 routes selected for forwarding '[x/y]' denotes [distance/metric] 10.0.1.1/32, vrf default via 10.0.2.1, [110/50], ospf AGG-2# show vsx ip data-path 10.0.1.1/32 IPv4 Data Path Information For 10.0.1.1/32 Local Device ------------ Route : 10.0.1.1/32 Egress L3 Interface : vlan2 Peer Device ------------ Route : 10.0.1.1/32 Egress L3 Interface : loopback0 AGG-2# show vsx brief ISL State : In-Sync Device State : Peer-Established Keepalive State : Keepalive-Established Device Role : secondary Number of Multi-chassis LAG interfaces : 1
AGG-1 / AGG-2	
AGG-1# traceroute 10.0.1.2	

```
traceroute to 10.0.1.2 (10.0.1.2), 1 hops min, 30 hops max, 3 sec. timeout, 3 probes
  1    10.0.1.2  27.909ms  13.347ms  8.212ms

AGG-2# traceroute 10.0.1.1
traceroute to 10.0.1.1 (10.0.1.1), 1 hops min, 30 hops max, 3 sec. timeout, 3 probes
  1    10.0.1.1  8.276ms  7.542ms  7.304ms
```

As shown above, the traffic between loopback goes directly over the Transit VLAN (here VLAN 2) transported over the ISL.

Split Situation



AGG-1	AGG-2
AGG-1# show ip route 10.0.1.2 **Displaying ipv4 routes selected for forwarding** **'[x/y]' denotes [distance/metric]** **10.0.1.2/32, vrf default** **via 10.0.0.2, [110/2000], ospf** **via 10.0.0.0, [110/2000], ospf** **AGG-1# show vsx ip data-path 10.0.1.2/32** **IPv4 Data Path Information For 10.0.1.2/32** **Local Device** **------------** **Route : 10.0.1.2/32** **Egress L3 Interface : 1/1/49** **Egress L3 Interface : 1/1/50** **VSX peer is not configured** **AGG-1# show vsx brief** **ISL State : Out-Of-Sync** **Device State : Split-System-Primary** **Keepalive State : Keepalive-Established**	AGG-2# show ip route 10.0.1.1 Displaying ipv4 routes selected for forwarding '[x/y]' denotes [distance/metric] 10.0.1.1/32, vrf default via 10.0.0.6, [110/2000], ospf via 10.0.0.4, [110/2000], ospf AGG-2# show vsx ip data-path 10.0.1.1/32 IPv4 Data Path Information For 10.0.1.1/32 Local Device ------------ Route : 10.0.1.1/32 Egress L3 Interface : 1/1/49 Egress L3 Interface : 1/1/50 VSX peer is not configured AGG-2# show vsx brief ISL State : Out-Of-Sync Device State : Split-System-Secondary Keepalive State : Keepalive-Established

```
Device Role                              : primary
Number of Multi-chassis LAG interfaces : 1


AGG-1# show vsx status keepalive
Keepalive State            : Keepalive-Established
Last Established           : Wed Dec 18 15:23:37 2019
Last Failed                : Wed Dec 18 13:06:54 2019
Peer System Id             : 02:01:00:00:01:00
Peer Device Role           : secondary

Keepalive Counters
Keepalive Packets Tx       : 8554
Keepalive Packets Rx       : 8506
Keepalive Timeouts         : 0
Keepalive Packets Dropped : 0
```

```
Device Role                              : secondary
Number of Multi-chassis LAG interfaces : 1


AGG-2# sh vsx status keepalive
Keepalive State            : Keepalive-Established
Last Established           : Wed Dec 18 15:24:11 2019
Last Failed                : Wed Dec 18 13:06:49 2019
Peer System Id             : 02:01:00:00:01:00
Peer Device Role           :

Keepalive Counters
Keepalive Packets Tx       : 8673
Keepalive Packets Rx       : 8552
Keepalive Timeouts         : 0
Keepalive Packets Dropped : 0
```

AGG-1 / AGG-2

```
AGG-1# traceroute 10.0.1.2
traceroute to 10.0.1.2 (10.0.1.2), 1 hops min, 30 hops max, 3 sec. timeout, 3 probes
  1    10.0.0.0  7.035ms  4.419ms  10.271ms
  2    10.0.1.2  15.455ms  9.288ms  14.589ms


AGG-2# traceroute 10.0.1.1
traceroute to 10.0.1.1 (10.0.1.1), 1 hops min, 30 hops max, 3 sec. timeout, 3 probes
  1    10.0.0.4  15.685ms  17.324ms  11.710ms
  2    10.0.1.1  21.630ms  18.922ms  16.793ms
```

As shown in the show ip route and traceroute commands, the traffic between loopbacks is established over the upstream L3 core devices and the VSX keepalive is maintained (not failed). One keepalive packet may be lost due to OSPF convergence time, which takes less than 3 seconds in all cases.

APPENDIX E - Resources and references

VSX introdutction video: https://www.youtube.com/watch?v=8kuhspVwBTk

VSX Technology Brief: https://www.arubanetworks.com/assets/tg/TB_VSX.pdf

VSX configuration guide: https://support.hpe.com/hpsc/doc/public/display?docId=a00091706en_us