AOS-CX 10.10 Update

aruba
a Hewlett Packard Enterprise company

VSX Graceful shutdown for IGMP and MLD

Presenters
- Rahim Raoufi
- Daryl Wan

Agenda

1. Overview
2. Use Cases
3. Details and Caveats
4. Configuration
5. Best Practices
6. Troubleshooting
7. Demo
8. Additional Resources

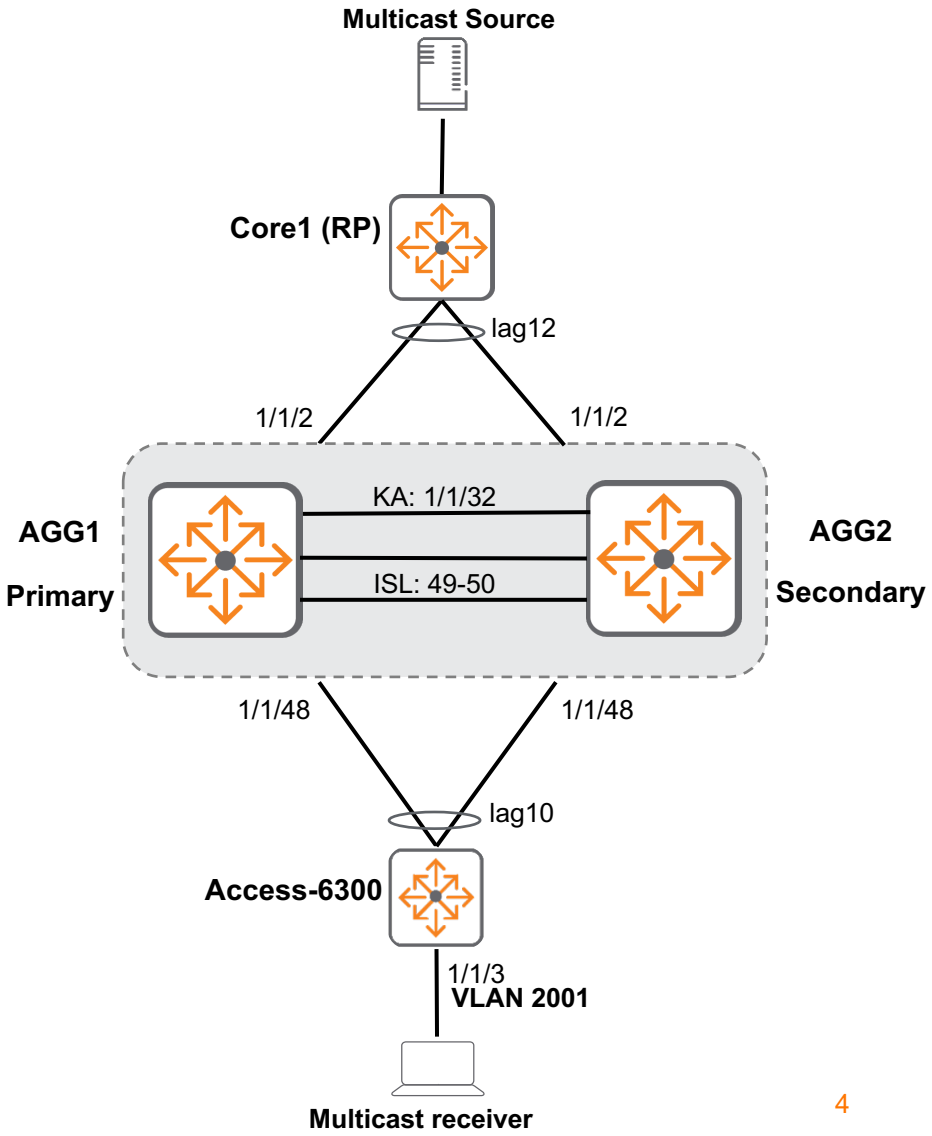
Overview

VSX Graceful shutdown for IGMP and MLD Overview

Querier resiliency in Multicast based VSX Network

IGMP and MLD Querier Offload Functionality

- When Querier offload is enabled, VSX software upgrade/reboot will trigger offloading querier responsibility to the VSX peer that is up and running.
- This is enabled by default.
- This feature is supported for both IPv4 and IPv6.



Supported Platforms

VSX and IGMP/MLD Querier Offload Supported Platforms.

- 6400, 8320, 8325, 8360, 8400, and 10000

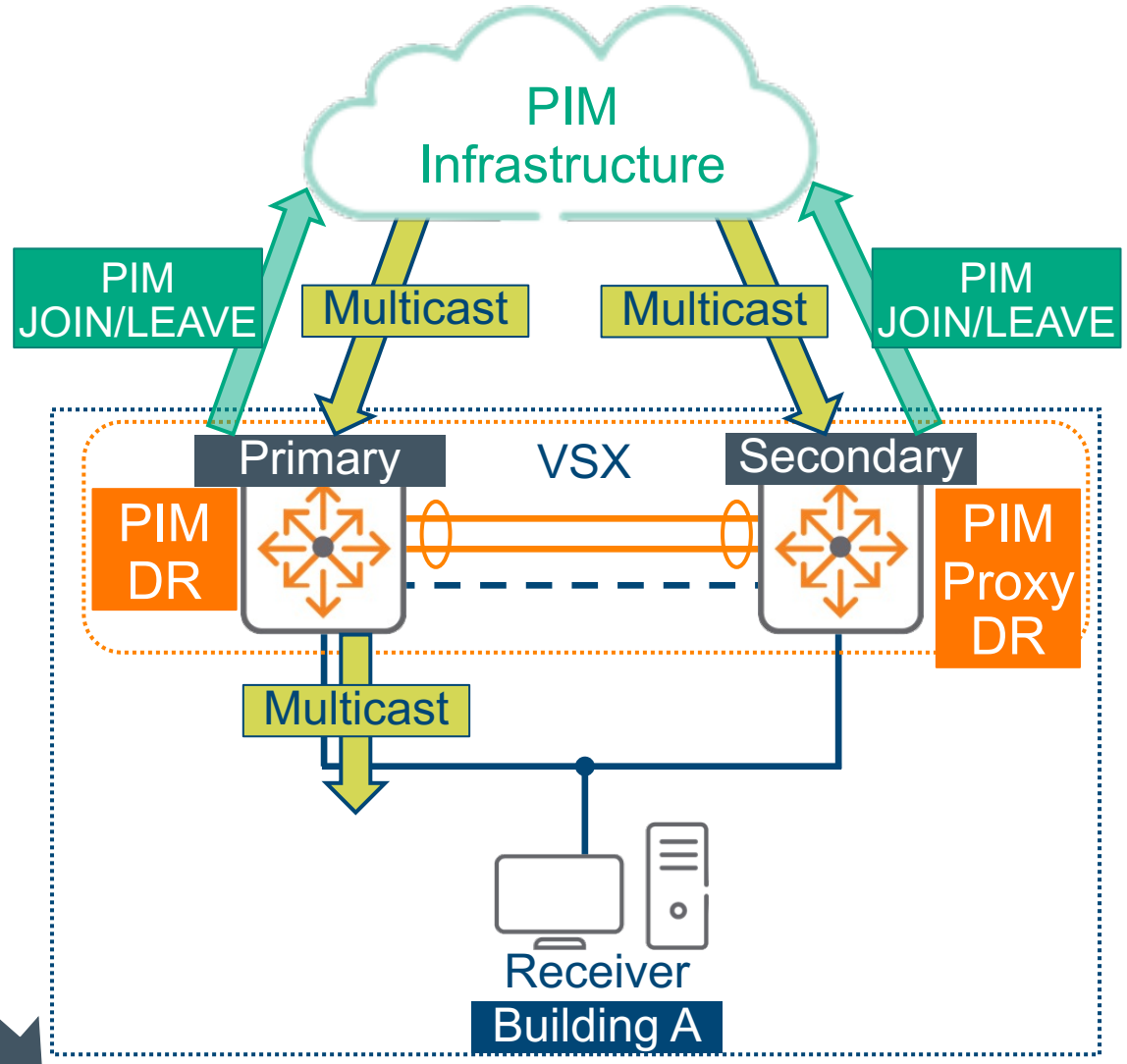
Use Cases

Deployment options and Solutions

VSX and PIM Active-Active review

Behavior

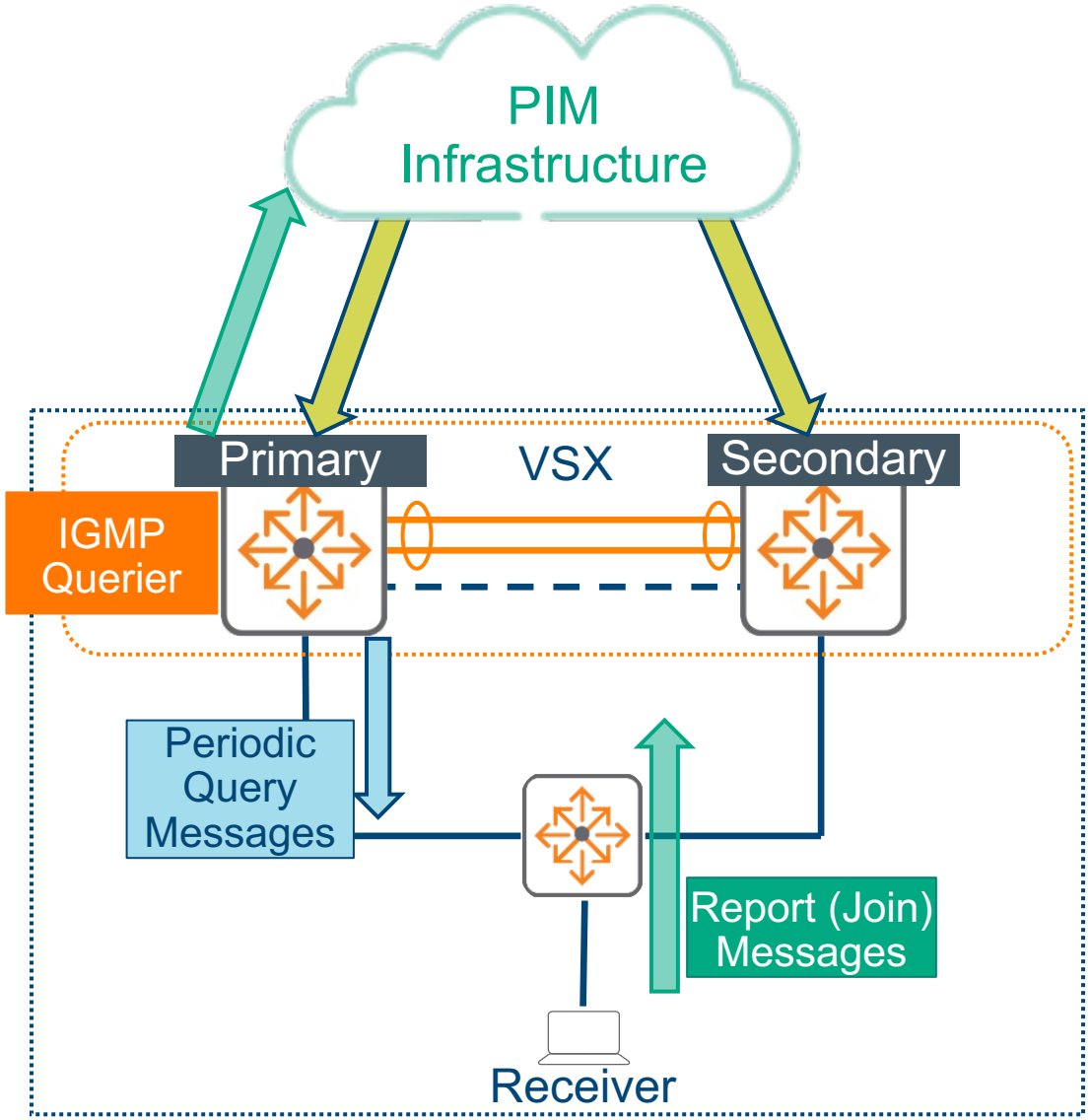
- PIM roles for VSX members:
  - One VSX peer: **PIM DR**
  - One VSX peer: **PIM Proxy DR**
- Both VSX members:
  - Send PIM JOIN/LEAVE messages
  - Establish PIM peering
  - Build the Shortest Path Tree
- **DR** and **Proxy DR** difference
  - Route entries for downstream ports for DR are written with route state and bridge state for Proxy DR
  - The DR is the only device that forwards multicast traffic to downstream ports



Route state means traffic is forwarded
Bridge state means traffic is blocked

Querier Resiliency in a VSX based Multicast network

- Querier's presence in a multicast network is imperative for traffic convergence.
- An IGMP/MLD Querier is chosen among different Routers for a given Subnet/VLAN
- A Querier is the most important node in the L2 Multicast network as it guarantees that periodic Queries are sent to the receiver and the receiver information is not lost in any of the Switches/Routers in the network. The absence of the Querier for a large amount of time can lead to traffic loss.
- During VSX software upgrade, VSX nodes will go for a reboot one after the other and when the node that goes for a reboot is an IGMP querier, we will end up having no queriers in the multicast network. This will cause traffic loss if the joins time out on the non-querier devices. The same applies to a general VSX querier reboot for some other reason.
- While in the process of a software upgrade, there have been many change requests (CRs) where the traffic loss was in the order of minutes before 10.10.



Details and Caveats

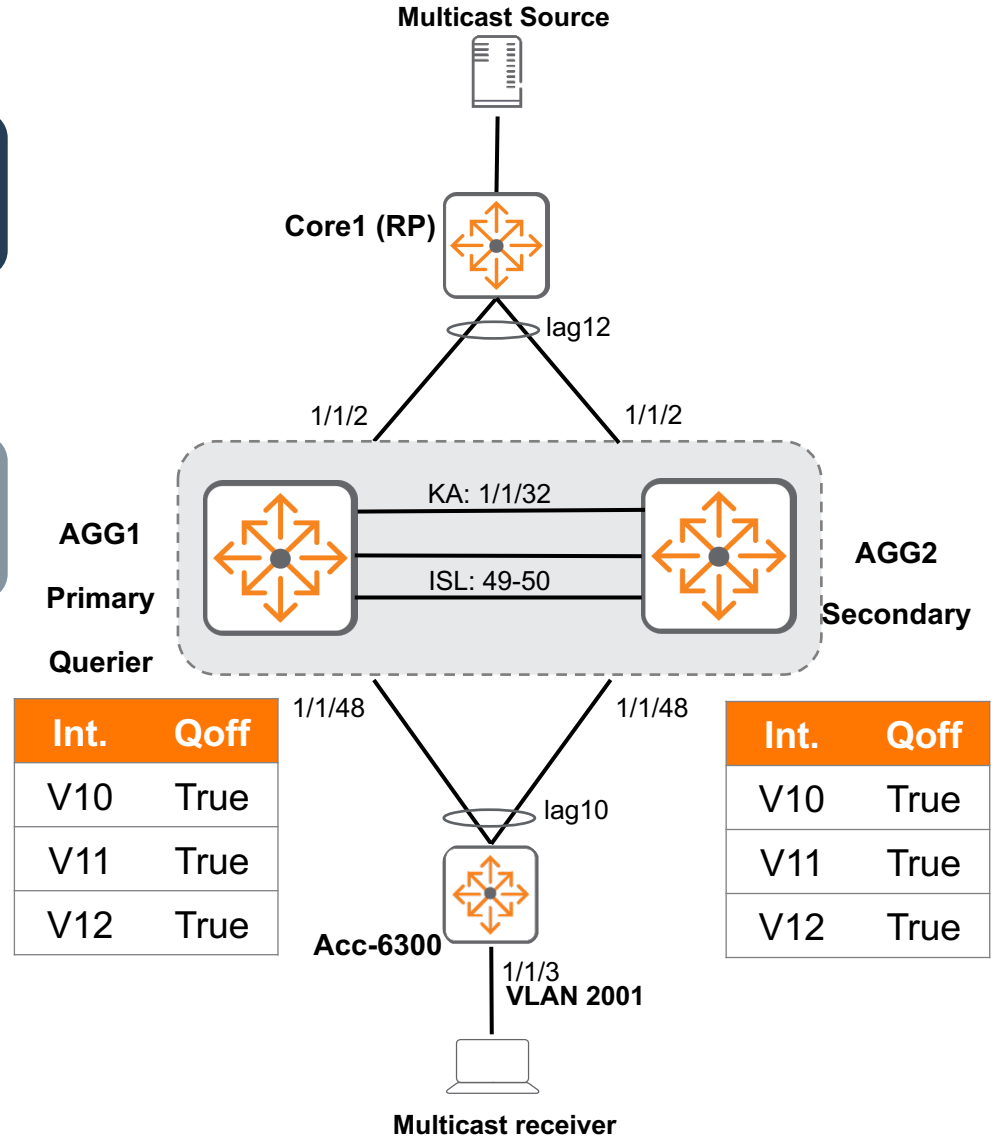
VSX software upgrade

During VSX software upgrade

- Daemon managing IGMP/MLD (MGMD daemon) will subscribe to get notified when the upgrade starts.
- VSX daemon will wait for MGMD daemon before starting the reboot.

When the upgrade notification is received by the MGMD daemon the following actions are taken:

- The VSX node which is going for a reboot, gets the list of interfaces with a querier role set in the MGMD interface state machine.
- The new column in the port table (querier offload) is **set to true** for querier interfaces.
- MGMD daemon will notify VSX daemon to start the reboot once the above step is complete.
- Remote IDL connection is used by the VSX peer which is up and running to read the querier offload column and take over the querier role.
- MGMD daemon will take up the querier role for those interfaces.
- Once the VSX upgrade is completed, the querier offload column will be reset



Int.	Qoff
V10	True
V11	True
V12	True

Int.	Qoff
V10	True
V11	True
V12	True

Multicast Group Membership Discovery protocol (**MGMD**) = IGMP and MLD functionality.

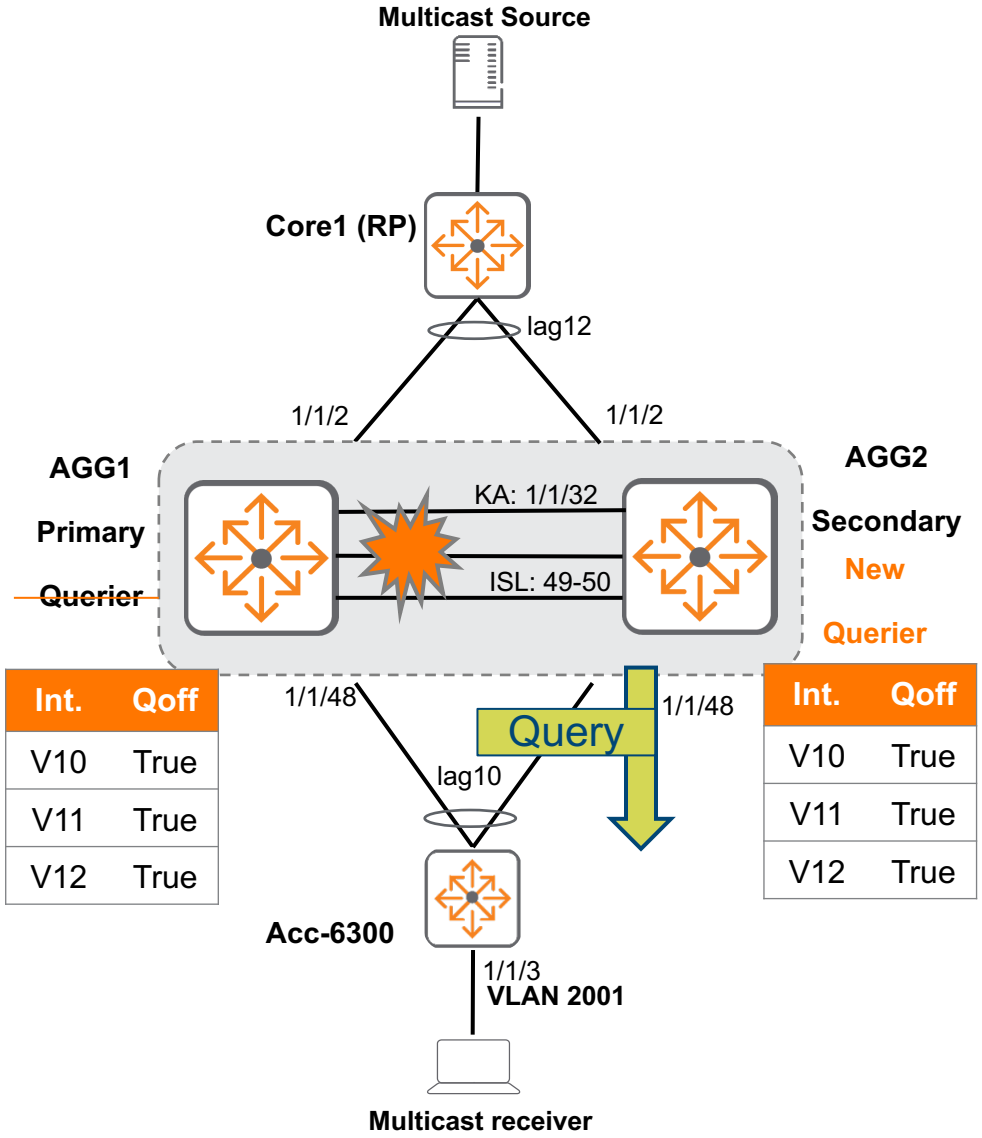
VSX Querier Node Reboot

ISL and keepalive are used to identify peer reboot.

VSX nodes will listen on the ISL link state and keepalive state.

VSX peer reboot is identified when the ISL link state is down, and the keepalive state fails.

The MGMD daemon on up and running VSX node, will take up the querier role for the interfaces which had learned the querier on the ISL port.



New AOS-CX 10.10 Enhancement with Querier Offload

With the new 10.10 enhancement, below are the observations during testing:

- On 10.9, Querier + Proxy DR reboot causes ~30second traffic loss when the querier expires, and querier election is triggered on the node. In this case, traffic loss will not be immediate as the joins will be synced. But only after the querier expires.

- **Sub-second traffic loss in 10.10.**

Wait Times

The wait times for the primary upgrade before multicast graceful shutdown process starts are listed below:

Number of MRoutes	Timer value
0	0
< 1024	120 seconds
< 2048	150 seconds
< 4096	210 seconds
< 8192	300 seconds
< 16284	360 seconds
> 16384	480 seconds

Caveats

VSX Graceful Shutdown Unsupported Interfaces

- **Sub-interface, loopback interfaces, tunnel interfaces such as GRE, VXLAN are not supported**)
- PIM active-active is not currently supported in VXLAN+VSX topology

NOTE: Querier offload functionality will function irrespective of the interface type. However, since active-active is required to get the complete solution, it is also recommended to follow the PIM active-active documentation.

Configuration

Configure MGMD-Querier-offload

mgmd querier-offload

Usage global configuration

```
switch# configure terminal
switch(config)#[no] mgmd querier-offload
```

This is enabled by default

Verification

```
switch# show running-config all | include mgmd


switch# show ip igmp all-vrf
switch# show ipv6 mld all-vrf
```

Best Practices

Best Practices

Recommended configurations for querier offload:

1. Querier offload feature works hand-in-hand with PIM active-active feature.

2. Without PIM active-active, convergence time on querier reboot/software upgrade will be high as there won't be proxy-DR immediately taking over the DR responsibility.

3. Since active-active is required to get to the complete solution, it is also recommended to follow the PIM active-active documentation.

4. This feature is applicable only in the case of VSX and not standalone.

5. In the versions without the querier offload feature, it was recommended to fine-tune “ip igmp robustness” value so that the joins won’t expire during the upgrade. This will not be needed going forward.

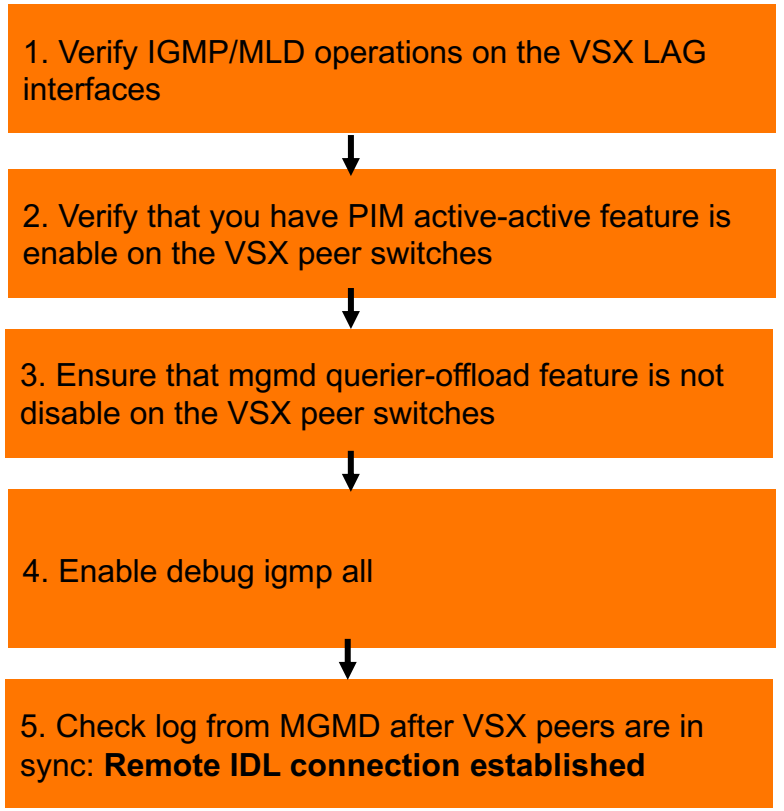
6. The querier offload on VSX upgrade will happen only when both the nodes have querier offload support AOS-CX 10.10 to further releases.

Troubleshooting

VSX Graceful shutdown for IGMP and MLD Troubleshooting

- Have a topology diagram ready
- Ensure IP interface details are included
- Check physical cabling and generate “show tech” when opening a TAC case
- Check network: Using show commands, ensure PIM neighbor between VSX peers, VSX status, and IGMP/MLD on VSX LAG interfaces and fix any issues found.
- Ensure VSX peers are in sync and check logs for MGMD

Recommended troubleshooting flow



1. Verify IGMP/MLD operations on the VSX LAG interfaces

```
BLDG01-AGG01# show ip igmp all-vrfs
VRF Name    : default
Interface   : vlan2001
IGMP Configured Version   : 3
IGMP Operating Version    : 3
Querier State               : Querier
Querier IP [this switch]    : 192.168.201.2
Querier Uptime              : 1d 8h 22m
Querier Expiration Time     : 0m 48s
IGMP Snoop Enabled on VLAN : False

Active Group Address    Vers Mode Uptime    Expires
----------------------- ---- ---- --------- ---------
239.10.10.10            3    EXC  1d 8h 22m 3m 7s
239.255.255.250         3    EXC  1d 8h 22m 3m 5s
```

```
BLDG01-AGG02# show ip igmp all-vrfs
VRF Name    : default
Interface   : vlan2001
IGMP Configured Version   : 3
IGMP Operating Version    : 3
Querier State               : Non-Querier
Querier IP                  : 192.168.201.2
Querier Uptime              : 1d 8h 23m
Querier Expiration Time     : 3m 58s
IGMP Snoop Enabled on VLAN : False

Active Group Address    Vers Mode Uptime    Expires
----------------------- ---- ---- --------- ---------
239.10.10.10            3    EXC  33d 3h 9m 4m 8s
239.255.255.250         3    EXC  5d 5h 9m  4m 8s
```

2. Verify that you have PIM active-active feature is enable on the VSX peer switches

```
BLDG01-AGG01# show ip pim interface vlan2001

PIM Interfaces

VRF: default

Interface  : vlan2001
Neighbor count : 1
IP Address : 192.168.201.2/24
Mode       : sparse
Designated Router : 192.168.201.2
Proxy DR              : true
Hello Interval (sec)  : 30
Hello Delay (sec)     : 5
Override Interval (msec)  : 2500     Lan Prune Delay : Yes
Propagation Delay (msec)  : 500      Configured DR Priority : 0
Operational DR Priority   : 1
Neighbor Timeout          : 90
```

```
BLDG01-AGG01# show ip pim interface vlan2001 vsx-peer

PIM Interfaces

VRF: default

Interface  : vlan2001
Neighbor count : 1
IP Address : 192.168.201.3/24
Mode       : sparse
Designated Router : 192.168.201.3
Proxy DR              : false
Hello Interval (sec)  : 30
Hello Delay (sec)     : 5
Override Interval (msec)  : 2500     Lan Prune Delay  : Yes
Propagation Delay (msec)  : 500      Configured DR Priority : 0
Operational DR Priority   : 4294967295
Neighbor Timeout          : 83
```

3. Ensure that mgmd querier-offload feature is not disable on the VSX peer switches

```
BLDG01-AGG01# show running-config all
Current configuration:
<--OUTPUT OMITTED FOR BREVITY-->
  mgmd querier-offload
```

4. Enable debug igmp all

```
BLDG01-AGG02# sh debug buffer reverse | i offload
2022-04-18:09:30:09.412905|hpe-mgmdd|LOG_DEBUG|AMM|-|IGMP|IGMP_CONFIG|Querier offload event sent for MLD interfaces
2022-04-18:09:30:09.412818|hpe-mgmdd|LOG_DEBUG|AMM|-|IGMP|IGMP_INTERFACE|MGMD Interface Event h=MgmdInterface_ehNoOpImpl i=536870916 s=querier mcast router e=Querier offload
2022-04-18:09:30:09.412780|hpe-mgmdd|LOG_DEBUG|AMM|-|IGMP|IGMP_CONFIG|Event name Querier offload
2022-04-18:09:30:09.412692|hpe-mgmdd|LOG_DEBUG|AMM|-|IGMP|IGMP_CONFIG|Querier offload event sent for IGMP interfaces
2022-04-18:09:30:09.408898|hpe-mgmdd|LOG_DEBUG|AMM|-|IGMP|IGMP_INTERFACE|MGMD Interface Event h=MgmdInterface_ehQuerierOffloadImpl i=536870916 s=non querier mcast router e=Querier offload
2022-04-18:09:30:09.408863|hpe-mgmdd|LOG_DEBUG|AMM|-|IGMP|IGMP_CONFIG|Event name Querier offload
2022-04-18:09:30:09.408837|hpe-mgmdd|LOG_DEBUG|AMM|-|IGMP|IGMP_CONFIG|Querier offload true for interface in route mode
```

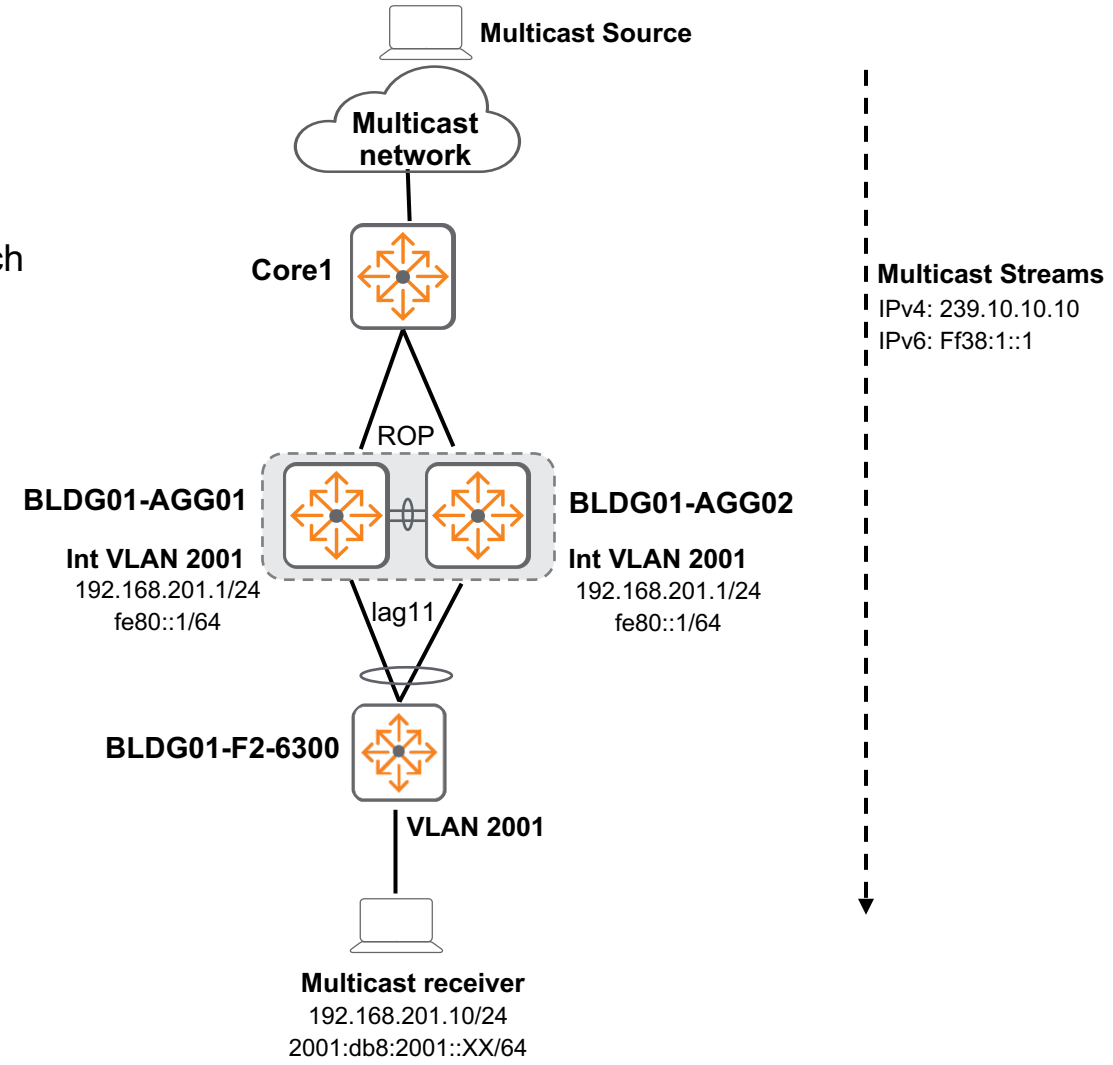
5. Check log from MGMD after VSX peers are in sync: Remote IDL connection established

```
BLDG01-AGG01# show events -r | include mgmd
2022-04-25T08:33:39.124215-07:00 BLDG01-AGG01 hpe-mgmdd[2138]:
Event|2607|LOG_INFO|AMM|1/1|Interface vlan2001: Start IGMP Querier role addr: 192.168.201.2
2022-04-25T08:32:40.119479-07:00 BLDG01-AGG01 hpe-mgmdd[2138]:
Event|2606|LOG_INFO|AMM|1/1|Interface vlan2001: End MLD Querier role
2022-04-25T08:32:40.119334-07:00 BLDG01-AGG01 hpe-mgmdd[2138]:
Event|2604|LOG_INFO|AMM|1/1|Interface vlan2001: Other Querier detected for MLD
2022-04-25T08:32:39.118820-07:00 BLDG01-AGG01 hpe-mgmdd[2138]: Event|2605|LOG_INFO|AMM|1/1|MLD
Querier Election in progress for interface vlan2001 with IP address fe80::d067:2687:d149:cadc
2022-04-25T08:32:35.302088-07:00 BLDG01-AGG01 hpe-mgmdd[2138]: Event|2605|LOG_INFO|AMM|1/1|IGMP
Querier Election in progress for interface vlan2001 with IP address 192.168.201.2
2022-04-25T08:29:31.334414-07:00 BLDG01-AGG01 hpe-mgmdd[2138]:
Event|2623|LOG_INFO|AMM|1/1|Remote IDL connection established.
```

Demo

VSX graceful shutdown for IGMP/MLD demo

– Show minimal multicast traffic impact even when VSX switch which is IGMP querier reboots



Resources

Feature/Solution References

- PIM-SM
  - **https://datatracker.ietf.org/doc/rfc7761/**
  - **PIM Message Type Space Extension and Reserved Bits**
    - **https://datatracker.ietf.org/doc/rfc8736/**

- IGMPv2
  - https://datatracker.ietf.org/doc/rfc2236/
- IGMPv3
  - https://datatracker.ietf.org/doc/rfc3376/
- **IGMPv3/MLDv2 latest version**
  - https://datatracker.ietf.org/doc/rfc4604/

aruba
a Hewlett Packard Enterprise company

Thank you

fardin.rah.raoufi@hpe.com
daryl.wan@hpe.com