Hewlett Packard Enterprise

HPE aruba networking

AOS-CX 10.13 Update:

Multicast Update: PIM Bidirectional (PIM-Bidir) & Multicast Optimizations

November 2023

Agenda

1	Overview
2	Use Cases
3	Details
4	Configuration
5	Best Practices
6	Troubleshooting

Overview

Overview Multicast Terminology

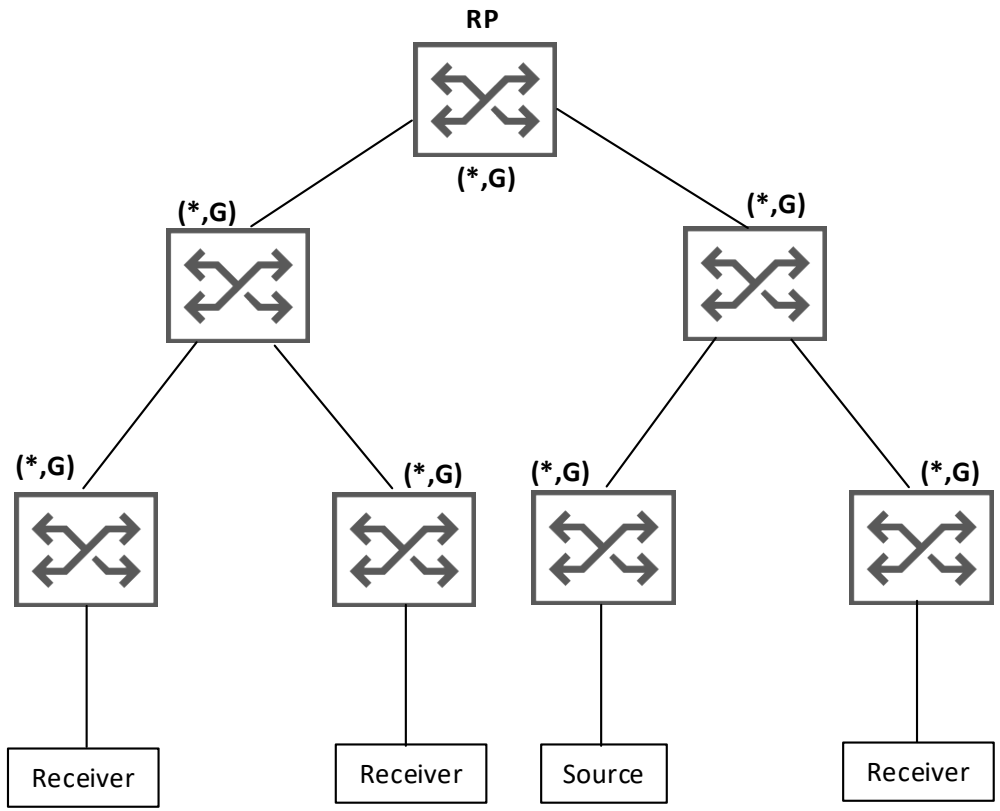
Term	Definition
PIM-SM	Protocol Independent Multicast Sparse Mode
PIM-Bidir	Protocol Independent Multicast Bidirectional
IGMP	Internet Group Management Protocol –three versions, used buy receivers who want to acquire a multicast stream
RP	Rendezvous Point –builds a multicast map of sources and receivers and other RPs, and is root of the shared tree
SPT	Shortest Path Tree AKA Source Tree
RPT	Rendezvous Point Tree AKA Shared Tree
(S) DR	Designated Router - send PIM join/prune messages, register source to RP and forward multicast, sometimes called the Source DR see also FHR. (S) DR **NOT** Applicable to PIM Bidir
FHR	First Hop Router - closest to source interchangeably referred to as DR or Source DR in some cases
LHR	Last Hop Router - connected to receivers, IGMP information is collected on receivers and forwarded upstream [hop by hop] to RP. The LHR is AKA Receiver DR
(R) DR	Receiver DR (NOT applicable to PIN Bidir) , see LHR
Querier	Querier, communicates with receivers using IGMP to determine multicast group membership see also LHR and (R) DR
RPF	Reverse Path Forwarding –Forwarding multicast traffic away from source toward receiver. Check multicast source with unicast routing table if it's a match accept only that traffic flow.
RPF (check)	Reverse Path Forwarding –multicast loop /duplication prevention for PIM SM. Forwarding multicast traffic away from source toward receiver. Check multicast source with unicast routing table if it's a match accept only that traffic flow.
OIL	Outgoing Interface List - Interface facing downstream towards receivers build the OIL after RPF check
DF	Designated Forwarder used in PIM-Bidir to forwards multicast sources downstream on link and off link to RP
DF (E)	Designated Forwarder (Election) used in PIM-Bidir each segment or link becomes DF based on best unicast routing metric to RP
RPTF	RP Tree Forwarding in PIM-Bidir the RP forwards multicast traffic to the receivers along the receiver side RP/Shared Tree

Overview

- **Part 1 of 2**
  - **PIM Bidirectional for IPv4**

- Part 2 of 2
  - Multicast Optimizations
    - Querier wait time
    - Handling of new and old joins
    - Preprograming (*,G) bridge entries

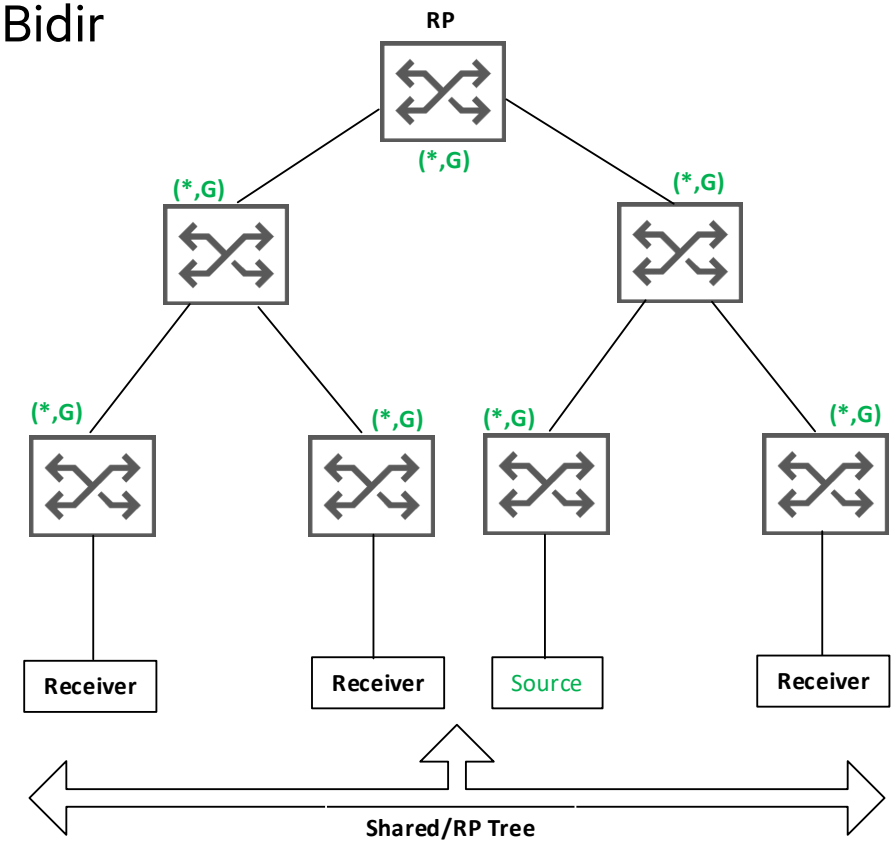
Overview PIM Bidirectional (PIM Bidir)



- New PIM Bidirectional for IPV4
  - Supported platforms 10000, 9300 ,8360, 8325 ,8100 ,6400 ,6300 and 6200
- PIM Bidirectional creates multicast trees between source and receivers without using source specific states
- (S,G) entries are **NOT** used or maintained
- Source traffic is always sent to the RP and then towards receivers
- Benefit
  - removing Shortest/Source Path Tree (SPT) and (S,G) entries less resources and overhead on switch resources
  - Can build multicast networks that scale better than PIM-SM when many to many highly distributed source/receiver

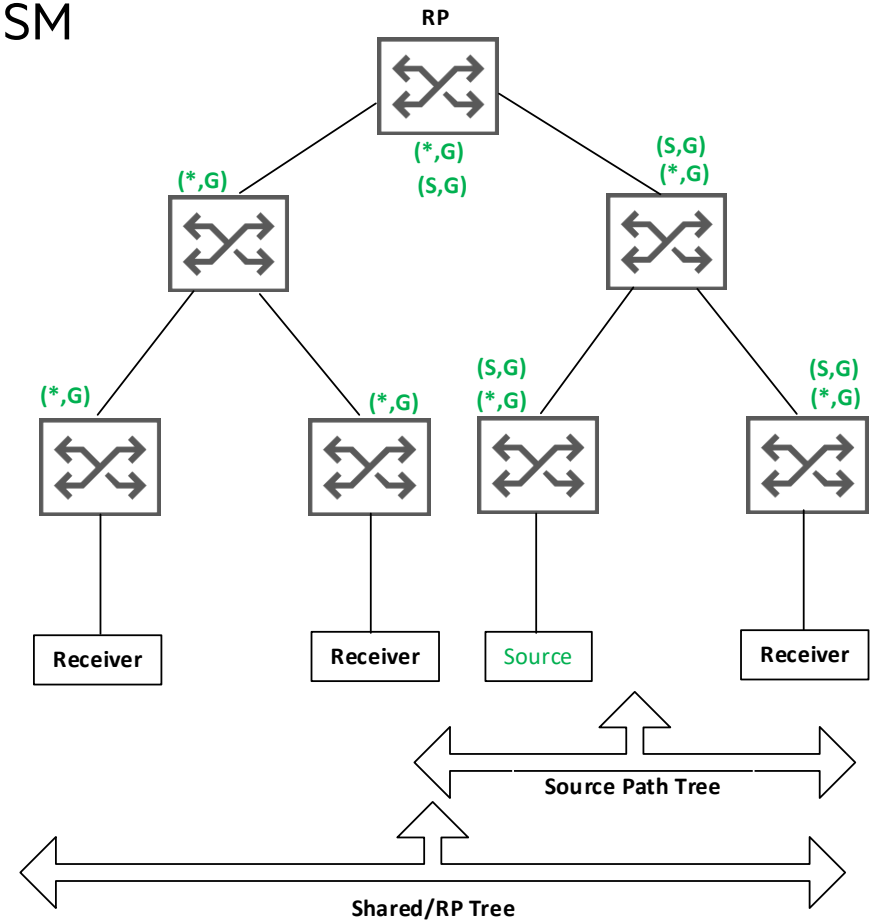
PIM Bidir and PIM SM comparison

PIM Bidir



- PIM Bidir creates multicast Shared /RP Trees Only (*, G) entries
- Shared/RP Tree is bidirectional
- Source traffic is always sent to the RP and then towards receivers
- RP is both control and data plane

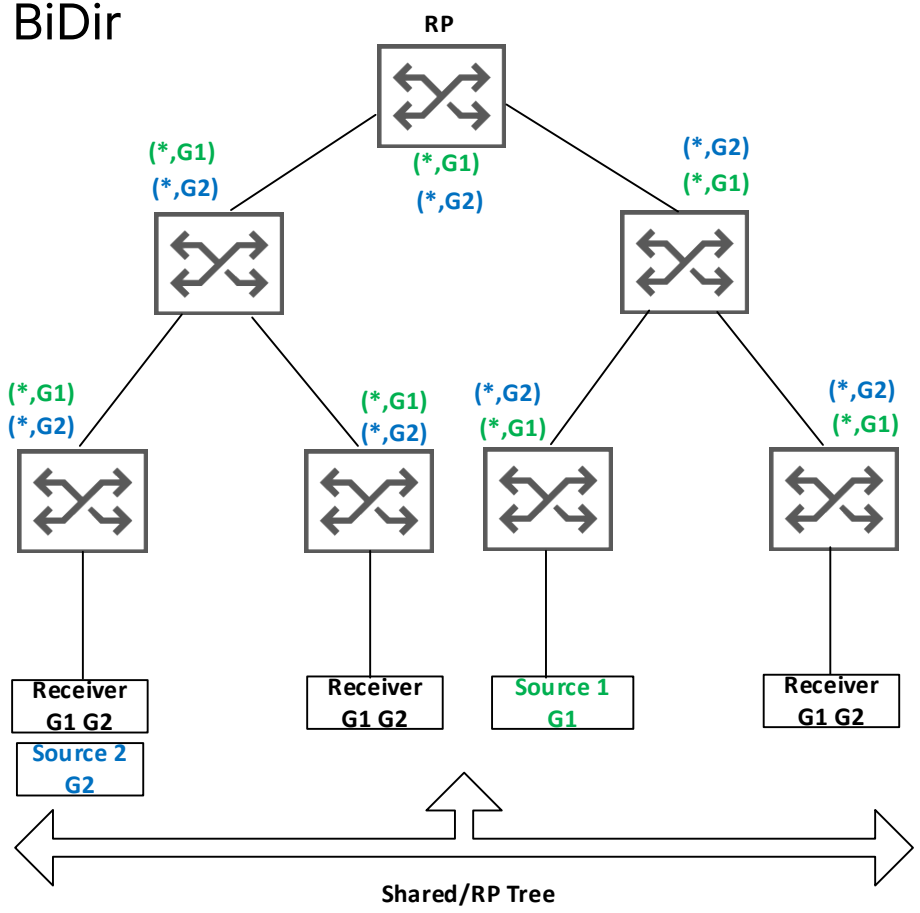
PIM SM



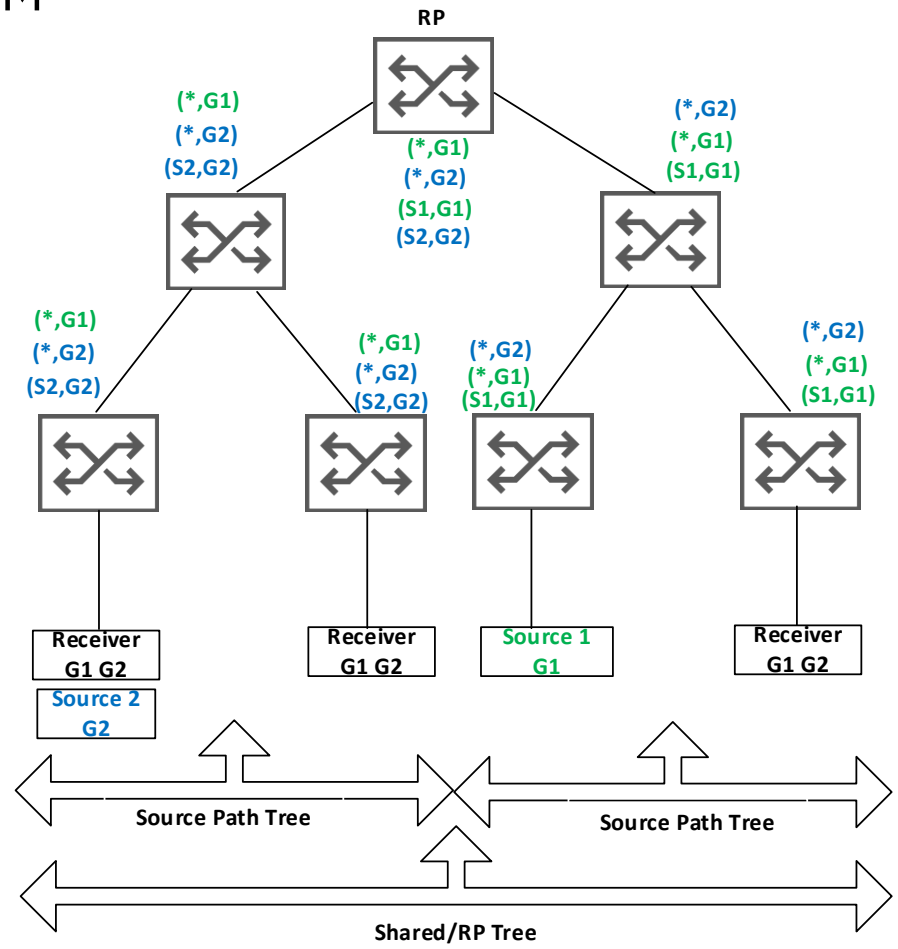
- PIM SM creates multicast Shared /RP Trees (*, G) entries
- Shared/RP Tree is unidirectional
- Source/Shortest Path Trees also maintained (S,G)
- Potential of optimal traffic paths
- RP is control plane

PIM Bidir and PIM SM comparison

PIM BiDir

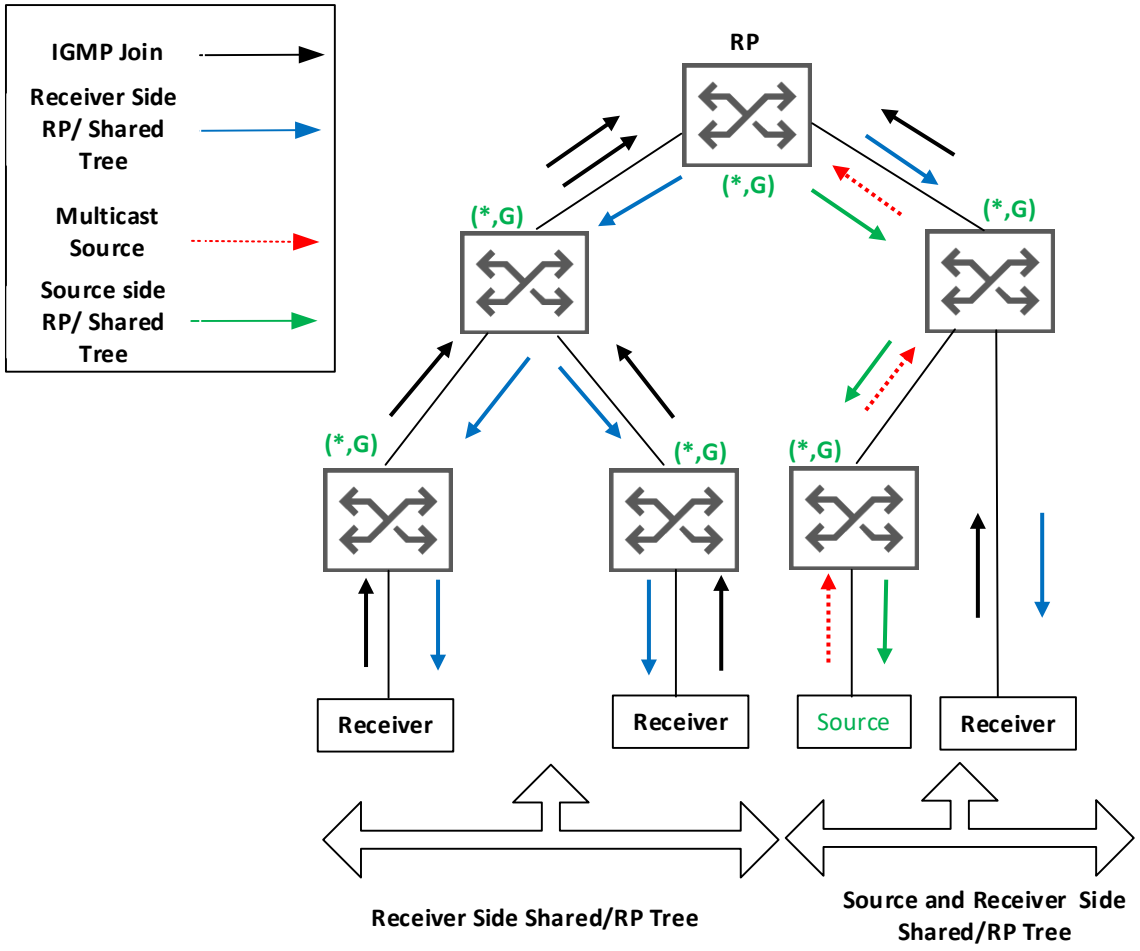


PIM SM



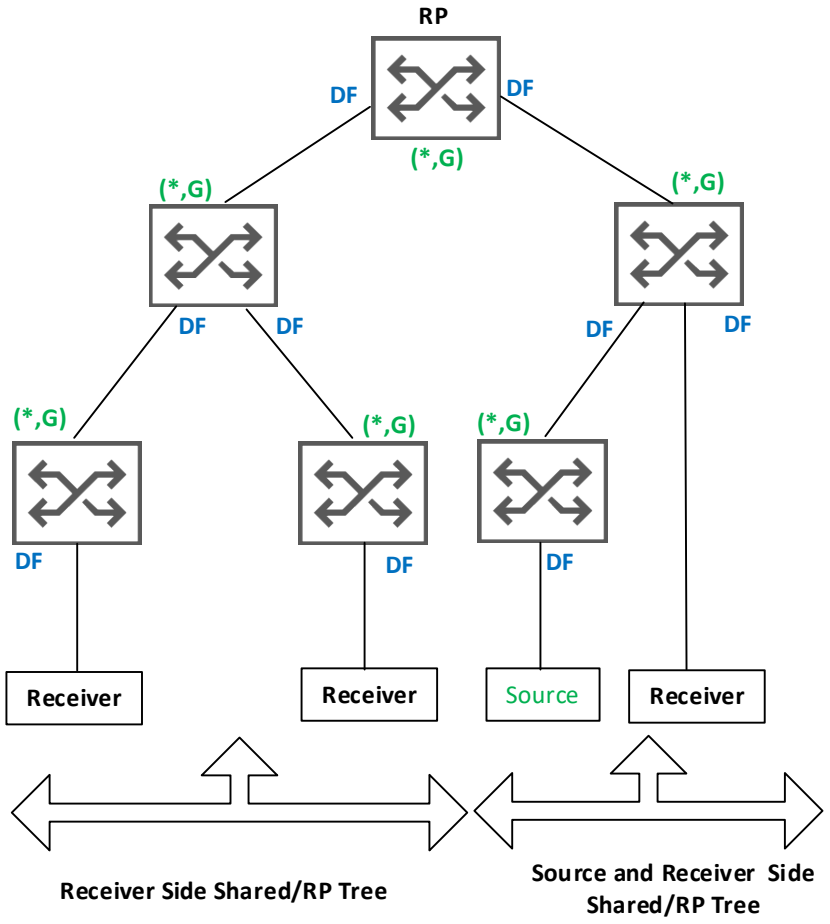
- As source and receiver distribution increases:
  - Increased resources and overhead on switch resources in PIM-SM to maintain (S,G) and (*,G) entries
  - Lesser resources and overhead on switch resources in PIM-Bidir as only (*,G) entries maintained

PIM Bidir building RP/ Shared Tree



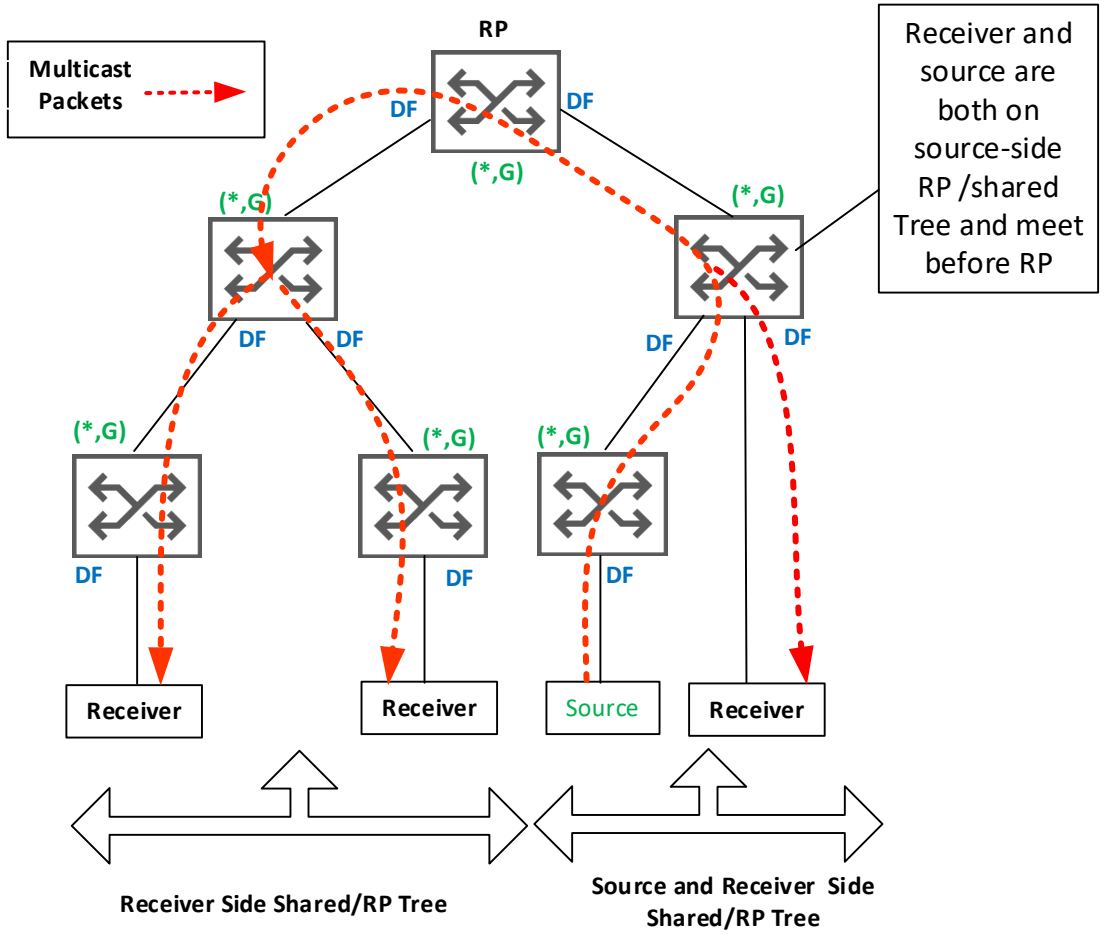
- PIM Bidir builds source side and receiver side RP/Shared Trees
- Receiver side:
  - IGMP joins forwarded to RP by Designated Forwarder (DF see later)
  - IGMP joins forwarded hop by hop to RP
  - routers along the path from the local receiver (DF) to the RP form an RP/Shared Tree branch. Each branch adds a (*, G) entry to its FIB.
  - Receiver prune messages (not shown) similarly sent towards RP remove the branch and (*,G) entries from FIB are aged out
- Source side:
  - Multicast packets are sent unconditionally to RP
  - routers along the path from the local source (DF) to the RP form an multicast route RP/Shared Tree branch.
- Other :
  - In the graphic we also see a source and receiver side branch (right) formation prior to the RP router which is typical
  - If no receiver side branches are formed at the RP multicast traffic will be dropped at the RP
  - Active sources will always send to the RP. PIM-Bidir has no register start/stop

PIM Bidir Operation DF and DFE



- DF –Designated Forwarder
  - DF, forwards multicast sources towards the RP
  - Each multicast router will have a DF **per** RP
  - Each segment or link will have a single DF. Per RP if multiple RPs used
  - DF is the only router on the segment that is allowed to send multicast traffic
    - Downstream onto its link
    - Picks Upstream packets off link to forward to RP
  - DF used to form loop free shared tree routed at RP
  - RPF is **not** used for loop prevention but used check correct interface towards RP
  - DF receiver side also forwards IGMP joins (seen earlier) upstream to RP (combined in OIL)
- DFE-DF Election is per segment or link and per defined RP
  - DF Election is based on the unicast routing table best metric to reach the RP for each link /segment, in a tie break highest IP address will become the DF

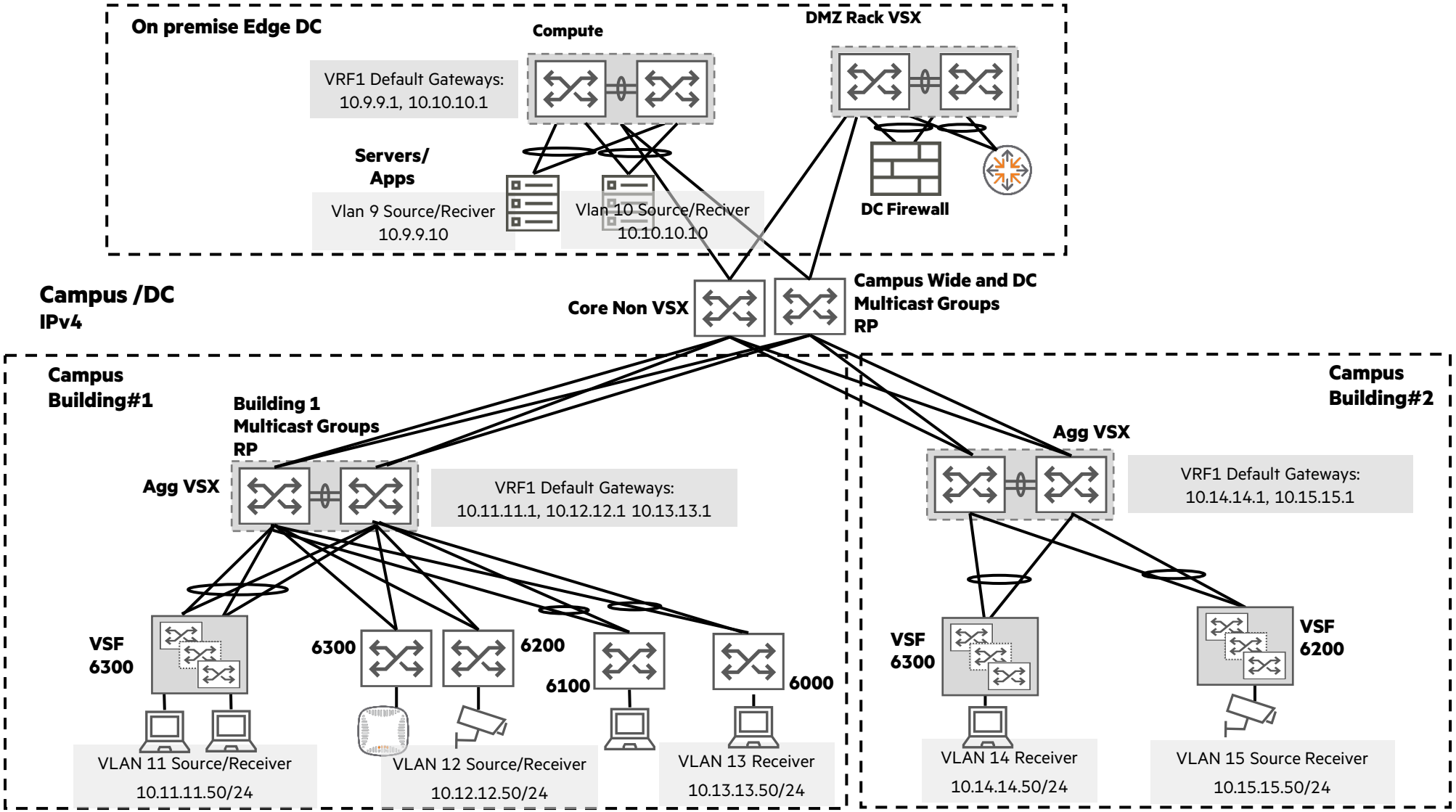
PIM Bidir Operation



- Multicast packets flow to receiver side shared tree via the RP (DF forwarding)
- RP is in the Control and Data plane for PIM Bidir
- If receiver and source are on the source side RP/ Shared Tree and meet prior to the RP as shown
  - multicast to the receiver are directly forwarded by the intermediate router, instead of by the RP.
  - A copy is still always sent to RP to facilitate multicast for receiver side shared tree
- If no active receivers RP will drop/not forward multicast

Use Cases

Campus/Edge DC Use Case



- Sources and receivers are highly distributed across the Enterprise
- Sources can also be receivers, and many to many multicast flows.
- Example applications: UPnP/ Simple Source Discovery Protocol [SSDP], Multi player games with chat, Trading channels & market data systems, CCTV with high dense source and receiver's

Configuration

Configuration between L2 Access to VSX Aggregation

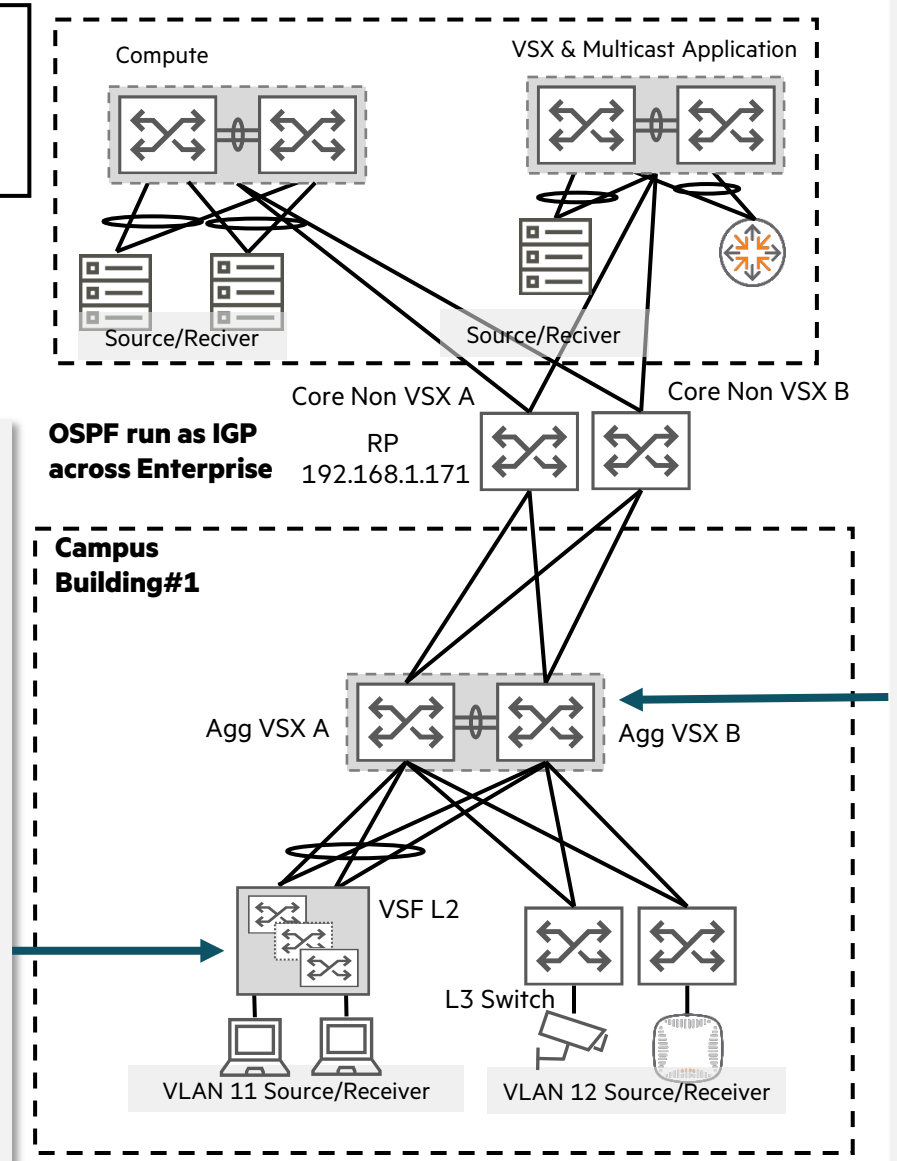
IMPORTANT NOTE:

Only partial and relevant configurations to subject matter feature shown for brevity:

Configuration L2 Access switch to Aggregation VSX

L2 access enable IGMP snooping per VLAN

```
hostname VSF-L2
!
vlan 11
    ip igmp snooping enable
!
interface lag
    no shutdown
    no routing
    vlan trunk native 1
    vlan trunk allowed 8-11
!
interface 1/1/52
      no shutdown
      lag 10
!
interface 1/1/1
    no shutdown
    no routing
    vlan access 11
```




```
Hostname Agg-VSX-B
!
vlan 11
    ip igmp snooping enable
!
interface lag 10 multi-chassis
    description downlink to VSF
    vlan trunk native 1
    vlan trunk allowed 8-11
    lacp mode active
!
interface lag 256
    description ISL
    vlan trunk native 1
    vlan trunk allowed 1,8-11
interface 1/1/52
   description to VSF stack
no shutdown
      lag 10
!
interface vlan 11
    ip address 192.168.11.1/24
    active-gateway ip mac 12:00:00:00:19:69
    active-gateway ip 192.168.11.1
    ip igmp enable
    ip pim-bidir enable
!
router pim
    enable
    rp-address 192.168.1.171
    active-active
```

L2 access enable IGMP snooping per VLAN

Per access VLAN enable IGMP and PIM Bidir

PIM routing and static RP defined on each L3 node

For HA enable active-active on VSX pair (optional)

Configuration between L3 Access to VSX Aggregation

IMPORTANT NOTE:

Only partial and relevant configurations to subject matter feature shown for brevity:

Configuration L3 Access switch to Aggregation VSX

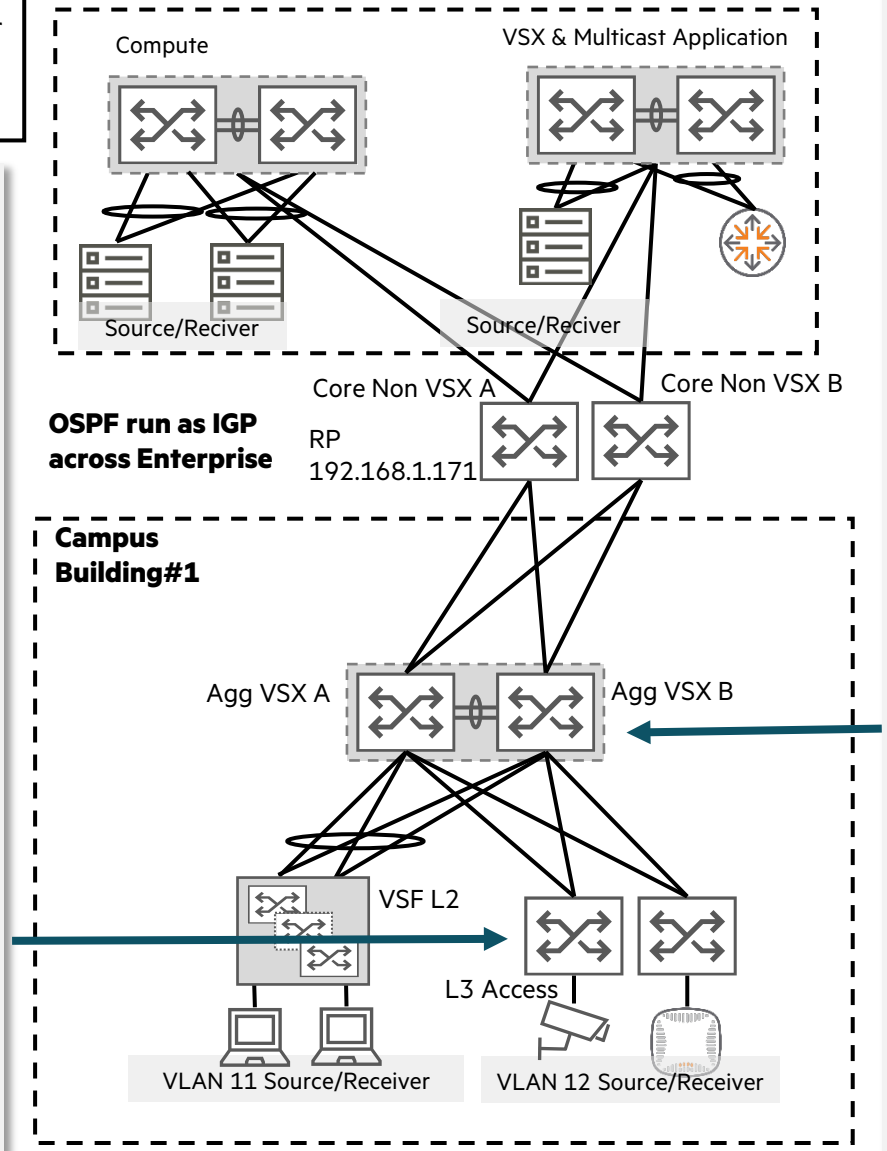
L2 access enable IGMP snooping per VLAN

```
hostname L3-Switch
!
vlan 12
    ip igmp snooping enable
!
interface 1/1/1
    vlan access 12
!
interface 1/1/51
    description uplink to Agg-VSX
    ip address 192.168.2.35/31
    ip ospf 1 area 0.0.0.0
    ip ospf network point-to-point
    ip pim-bidir enable
!
interface vlan 12
ip address 192.168.12.1/24
ip pim-bidir enable
ip igmp enable
!
router pim
    enable
    rp-address 192.168.1.171
```

PIM Bidir per L3 interface

Per access VLAN enable IGMP

PIM routing and static RP defined on each L3 node




```
hostname-Agg-VSX
!
vlan 256
    description transit vlan
!
interface lag 256
    description ISL
    vlan trunk native 1
    vlan trunk allowed 1,8-11,256
interface 1/1/51
    description to L3 Switch
    ip address 192.168.2.34/31
    ip ospf 1 area 0.0.0.0
    ip ospf network point-to-point
    ip pim-bidir enable
interface vlan 256
    description transit vlan
    ip address 192.168.2.64/31
    ip ospf 1 area 0.0.0.0
    ip ospf cost 1
    ip ospf network point-to-point
    ip pim-bidir enable
!
router pim
    enable
    rp-address 192.168.1.171
    active-active
```

ISL transit VLAN for VSX

PIM Bidir per L3 interface

Lower IGP cost on transit VLAN for VSX

PIM routing and static RP defined on each L3 node

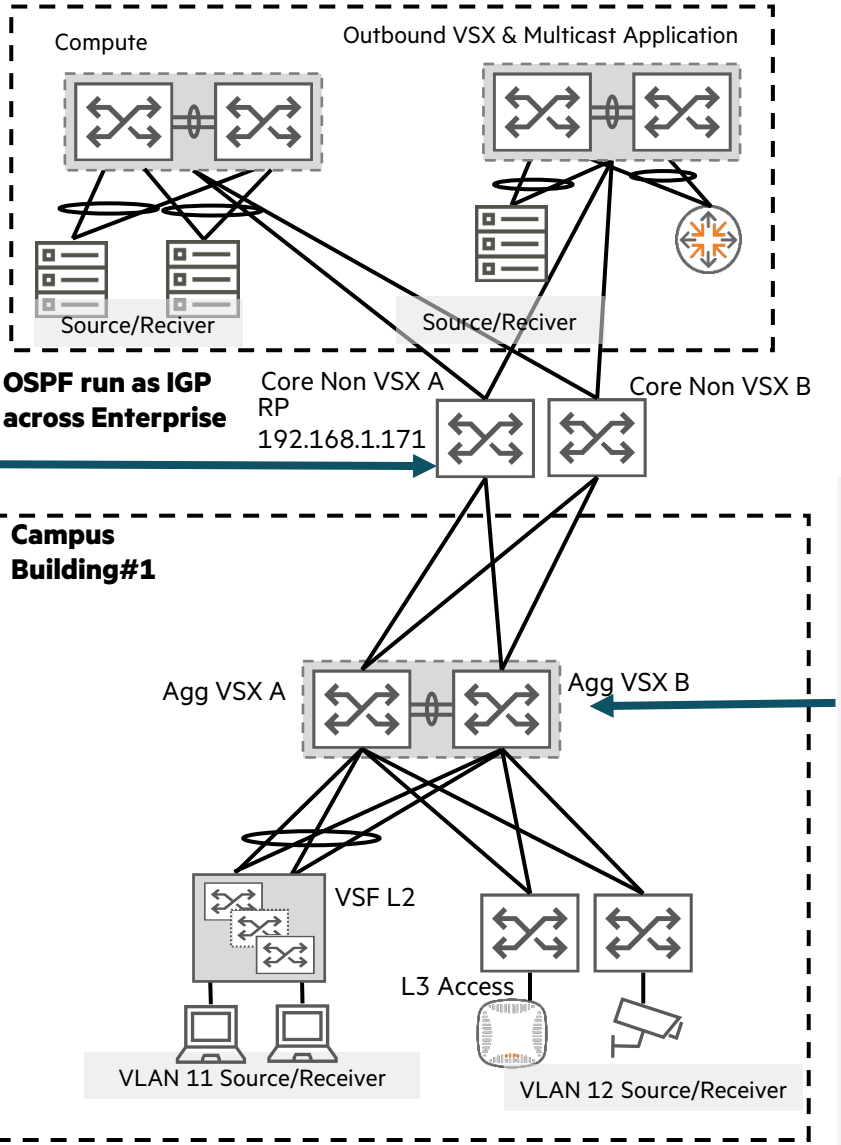
For HA enable active-active on VSX pair (optional)

Configuration between RP (Core Non VSX) to VSX Aggregation

IMPORTANT NOTE:

Only partial and relevant configurations to subject matter feature shown for brevity:

Configuration RP (Core Non VSX) to Aggregation VSX




```
hostname Core-Non-VX-A
!
interface 1/1/51
    description to Agg-VSX
    ip address 192.168.2.40/31
    ip ospf 1 area 0.0.0.0
    ip ospf network point-to-point
    ip pim-bidir enable
!
interface loopback 171
    ip address 192.168.1.172/31
    ip ospf 1 area 0.0.0.0
    ip pim-bidir enable
!
router pim
    enable
    rp-address 192.168.1.171
```

PIM Bidir per L3 interface

Loopback for RP

PIM routing and static RP defined on each L3 node

```
Hostname Agg-VSX-B
!
interface 1/1/49
    description to core non VSX
    ip address 192.168.2.41/31
    ip ospf 1 area 0.0.0.0
    ip ospf network point-to-point
    ip pim-bidir enable
!
router pim
    enable
    rp-address 192.168.1.171
    active-active
```

PIM Bidir per L3 interface

PIM routing and static RP defined on each L3 node

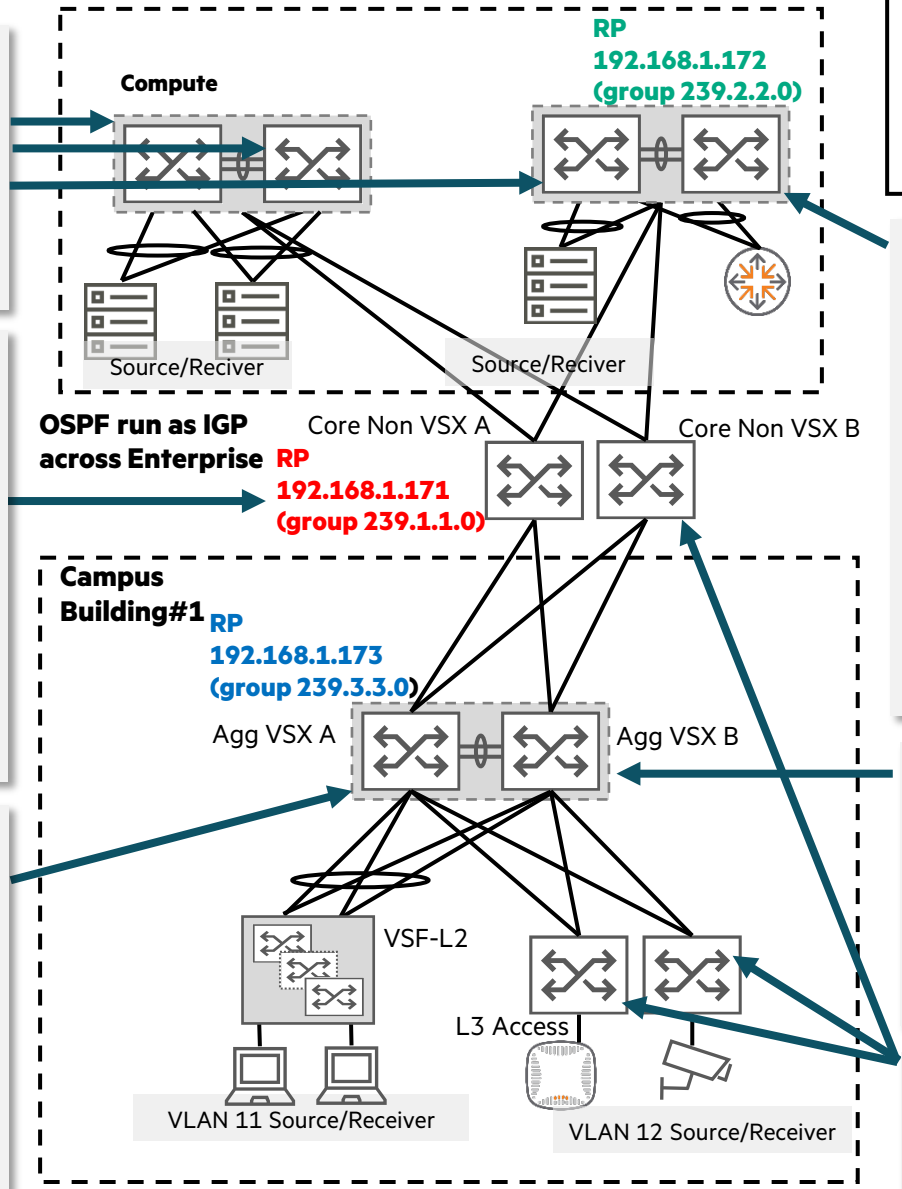
For HA enable active-active on VSX pair (optional)

Configuration distributed RP for traffic load

```
router pim
    enable
    rp-address 192.168.1.171 239.1.1.0/24
    rp-address 192.168.1.172 239.2.2.0/24
    rp-address 192.168.1.173 239.3.3.0/24
    active-active
```

```
interface loopback 171
ip address 192.168.1.171/32
ip ospf 1 area 0.0.0.0
ip pim-bidir enable
!
router pim
    enable
    rp-address 192.168.1.171 239.1.1.0/24
    rp-address 192.168.1.172 239.2.2.0/24
    rp-address 192.168.1.173 239.3.3.0/24
```

```
interface loopback 171
ip address 192.168.1.173/32
ip ospf 1 area 0.0.0.0
ip pim-bidir enable
!
router pim
    enable
    rp-address 192.168.1.171 239.1.1.0/24
    rp-address 192.168.1.172 239.2.2.0/24
    rp-address 192.168.1.173 239.3.3.0/24
    active-active
```



IMPORTANT NOTE:

Only partial and relevant configurations to subject matter feature shown for brevity:

RP traffic distribution for 3 different groups in default/global routing table

```
interface loopback 171
ip address 192.168.1.172/32
ip ospf 1 area 0.0.0.0
ip pim-bidir enable
!
router pim
    enable
    rp-address 192.168.1.171 239.1.1.0/24
    rp-address 192.168.1.172 239.2.2.0/24
    rp-address 192.168.1.173 239.3.3.0/24
    active-active
```

```
router pim
    enable
    rp-address 192.168.1.171 239.1.1.0/24
    rp-address 192.168.1.172 239.2.2.0/24
    rp-address 192.168.1.173 239.3.3.0/24
    active-active
```

```
router pim
    enable
    rp-address 192.168.1.171 239.1.1.0/24
    rp-address 192.168.1.172 239.2.2.0/24
    rp-address 192.168.1.173 239.3.3.0/24
```

- All PIM routers must have all known RPs present per VRF

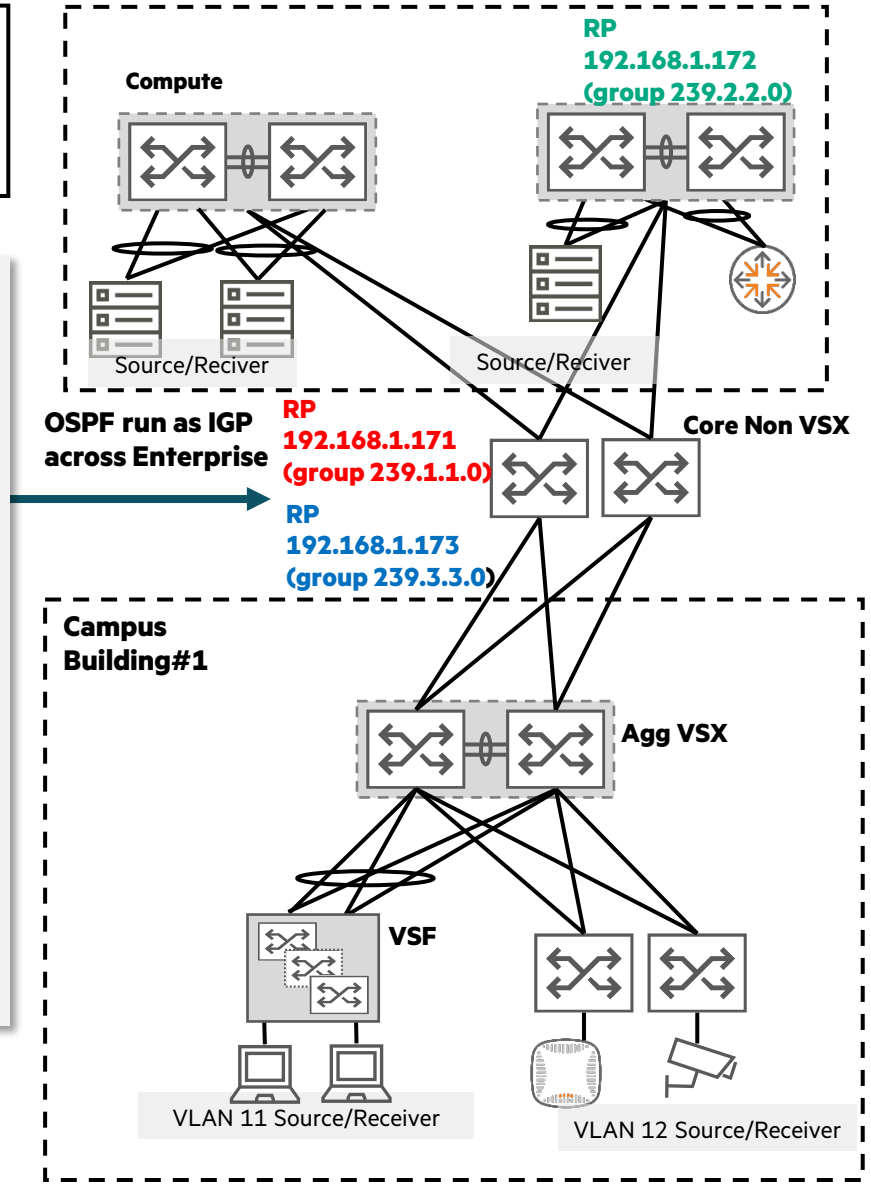
Configuration more than one RP on same device

IMPORTANT NOTE:

Only partial and relevant configurations to subject mater feature shown for brevity:

More than one RP on the same device

```
interface loopback 171
description group RED
ip address 192.168.1.171/32
ip ospf 1 area 0.0.0.0
ip pim-bidir enable
!
interface loopback 173
Description group BLUE
ip address 192.168.1.173/32
ip ospf 1 area 0.0.0.0
ip pim-bidir enable
!
router pim
    enable
    rp-address 192.168.1.171 239.1.1.0/24
    rp-address 192.168.1.172 239.2.2.0/24
    rp-address 192.168.1.172 239.3.3.0/24
```

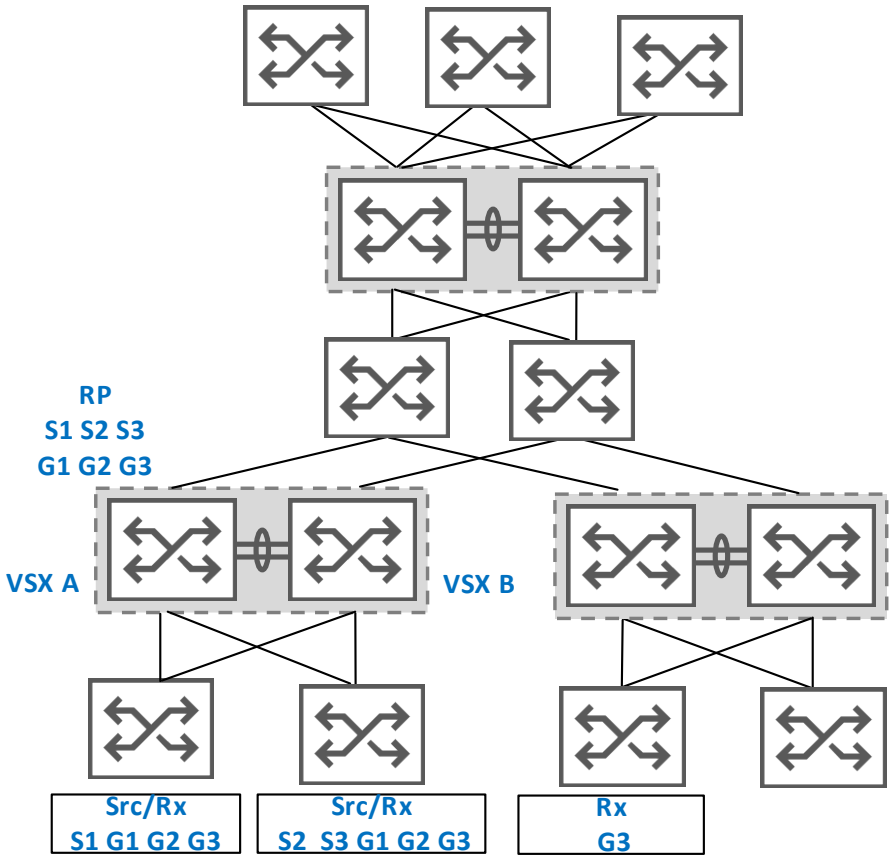


- If multicast groups cannot be summarized, and need to use the same switch for RP a separate loopback must be used per RP group set

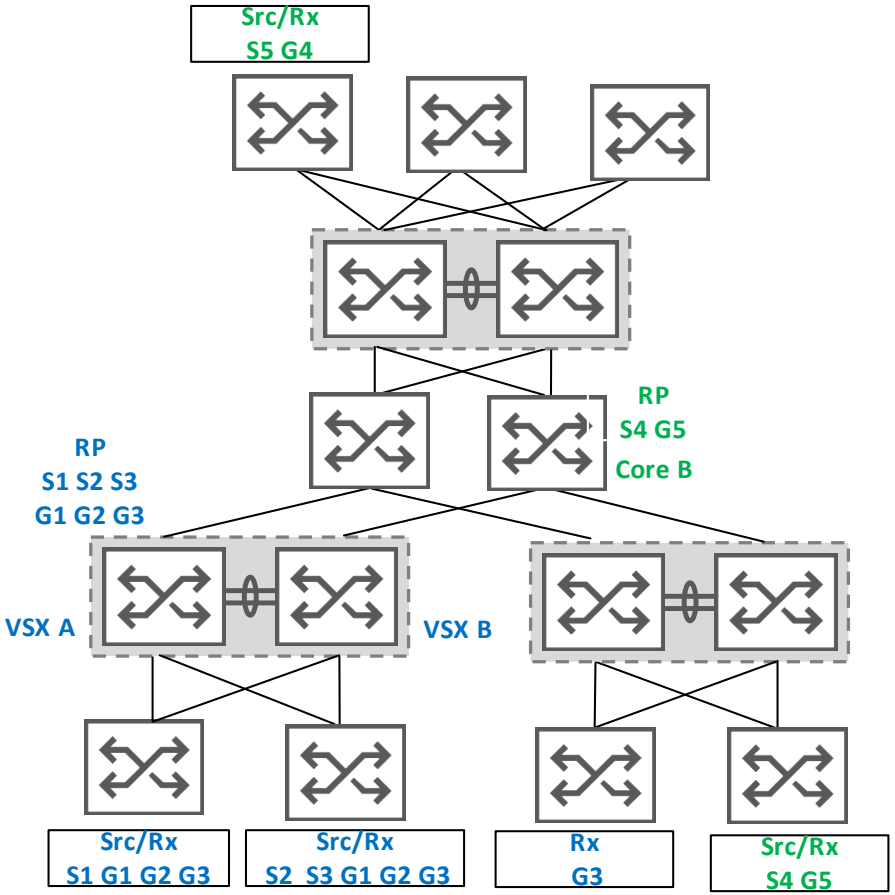
Best Practices

Best Practices

- Place RP based on optimal traffic flows, understand flows.
  - RP placed on VSX A for Sources 1-3 and Groups 1 to 3 in this vicinity



- Use new RP to further optimize and distribute flows
  - RP placed on Core B for Sources 4-5 and Groups 4-5 between north south



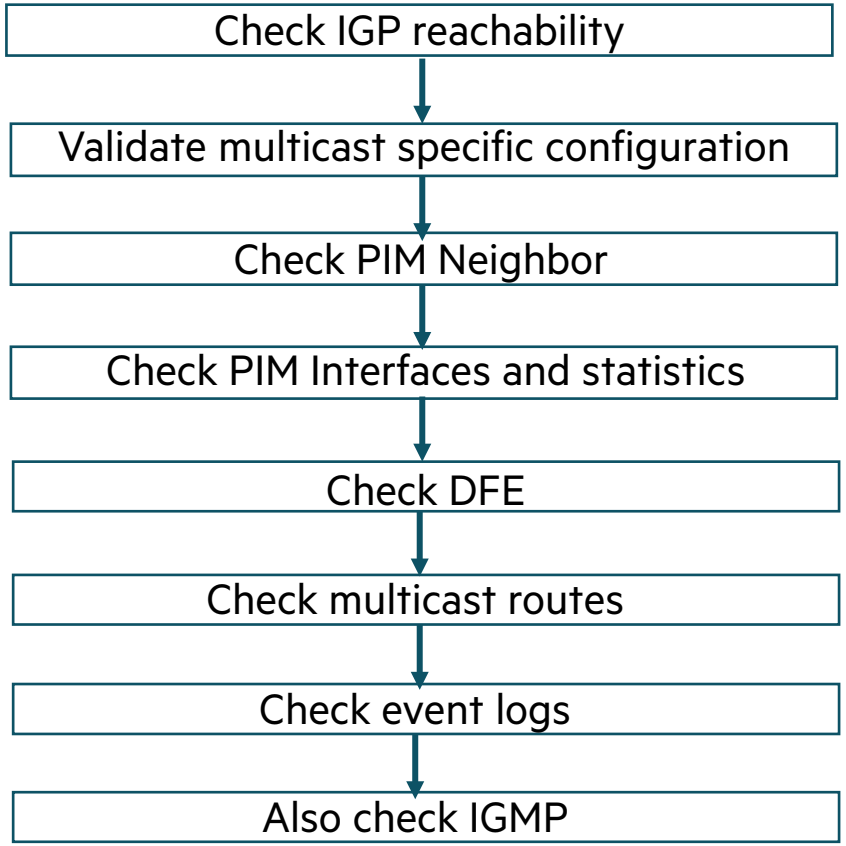
- Use loopback interface for RP
- When using VSX incorporate a transit VLAN with lower cost across ISL

Troubleshooting

Troubleshooting

- Use new PIM Bidir and existing verification commands as shown to troubleshoot

```
switch#show run pim

switch#show ip pim neighbor

switch#show ip pim interface brief

switch#show ip pim interface

show interface 1/1/3 | inc Mul

switch#show ip pim dfe

switch#show ip mroute

switch# show events -d pim-bidid

switch# diagnostics
switch# diag-dump pim-bidid basic

show ip igmp groups
```



Troubleshooting

Verify PIM configuration

```
show run pim
router pim
    enable
    rp-address 192.168.1.171 239.1.1.0/24
    rp-address 192.168.1.172 239.2.2.0/24
    rp-address 192.168.1.173 239.3.3.0/24
interface 1/1/51
    ip pim-bidir enable
interface loopback171
    ip pim-bidir enable
interface loopback172
    ip pim-bidir enable
interface vlan164
    ip pim-bidir enable
interface vlan180
    ip pim-bidir enable
```

Verify PIM Bidir and neighbour

```
show ip pim neighbor

PIM Neighbor

VRF                        : default
Total number of neighbors : 1

IP Address                 : 192.168.2.34
Interface                  : 1/1/51
Up Time (HH:MM:SS)         : 3 days 00:47:34
Expire Time (HH:MM:SS)     : 00:01:20
DR Priority                : NA
Hold Time (HH:MM:SS)       : 00:01:45
Bidir Capable              : True
```

Troubleshooting

Check and evaluate interfaces operating mode

```
show ip pim interface brief

-----------------------------------------------------------------------------------------
           VRF : default                          Total number of interfaces : 5
-----------------------------------------------------------------------------------------
Interface      IP Address            DR Address         Neighbor     Mode      VSX
                                                        count                  Role
-----------    ------------------    ---------------    ---------    ------    --------
1/1/51         192.168.2.35/31       Nil                1            bidir     N/A
vlan180        192.168.10.81/28      Nil                0            bidir     N/A
loopback172                          N/A                N/A          bidir     N/A
loopback171                          N/A                N/A          bidir     N/A
vlan164        192.168.10.65/28      Nil                0            bidir     N/A
-----------------------------------------------------------------------------------------
```

Check and evaluate interfaces in detail

```
show ip pim interface

PIM Interfaces

VRF: default
Total Number of interfaces : 5

Interface  : 1/1/51
Neighbor count : 1
IP Address : 192.168.2.35/31
Mode       : bidir
Designated Router : NA
Proxy DF              : false
Hello Interval (sec)  : 30
Hello Delay (sec)     : 5
Override Interval (msec)  : 2500              Lan Prune Delay     : Yes
Propagation Delay (msec)  : 500               Configured DR Priority    : NA
Operational DR Priority   : NA
Neighbor Timeout          : 83

Interface  : vlan180
Neighbor count : 0
IP Address : 192.168.10.81/28
Mode       : bidir
!output omitted
```

Troubleshooting

Check and evaluate interfaces for multicast traffic levels

```
show interface 1/1/51
interface 1/1/51 is up
 Admin state is up
!output omitted
 Rate collection interval: 300 seconds

 Rate                                RX                  TX        Total (RX+TX)
 ---------------- -------------------- -------------------- --------------------
 Mbits / sec                     107.36               21.40               128.76
 KPkts / sec                     203.34               29.60               232.94
   Unicast                         0.00                0.00                 0.00
   Multicast                     203.34               29.60               232.94
   Broadcast                       0.00                0.00                 0.00
 Utilization %                     1.07                0.21                 1.28

 Statistic                           RX                  TX                Total
 ---------------- -------------------- -------------------- --------------------
 Packets                        3448635              503051              3951686
   Unicast                            2                   2                    4
   Multicast                    3448633              503049              3951682
```

Check DF Election per RP

```
show ip pim dfe

PIM BIDIR DFE Information

VRF                          : default
Total number of DFE entries : 9

Interface   RPA             DF Winner       Metric          Uptime(HH:MM:SS)
---------   ---------       ---------       ------          -------------------
1/1/51      192.168.1.171   192.168.2.34    2               3 days 01:02:18
vlan180     192.168.1.171   192.168.10.81   12              3 days 01:01:58
vlan164     192.168.1.171   192.168.10.65   12              1 days 11:52:50
vlan180     192.168.1.172   192.168.10.81   22              1 days 10:03:17
vlan164     192.168.1.172   192.168.10.65   22              1 days 10:03:16
1/1/51      192.168.1.172   192.168.2.34    12              1 days 10:03:17
vlan180     192.168.1.173   192.168.10.81   10              3 days 01:01:58
1/1/51      192.168.1.173   192.168.2.34    0               3 days 01:02:19
vlan164     192.168.1.173   192.168.10.65   10              1 days 11:52:50
```

Troubleshooting

Check multicast routing and OIL

```
show ip mroute
IP Multicast Route Entries

VRF : default
Total number of entries : 8

Group Address           : 239.1.1.1
Source Address          : any
SSM Mroute              : False
Neighbor                :
Uptime                  : 01:21:38
State                   : route
Incoming interface      : vlan164
Outgoing Interface List :
Interface       State
-----------     ----------
1/1/51          forwarding
vlan180         forwarding
! Output omitted
```

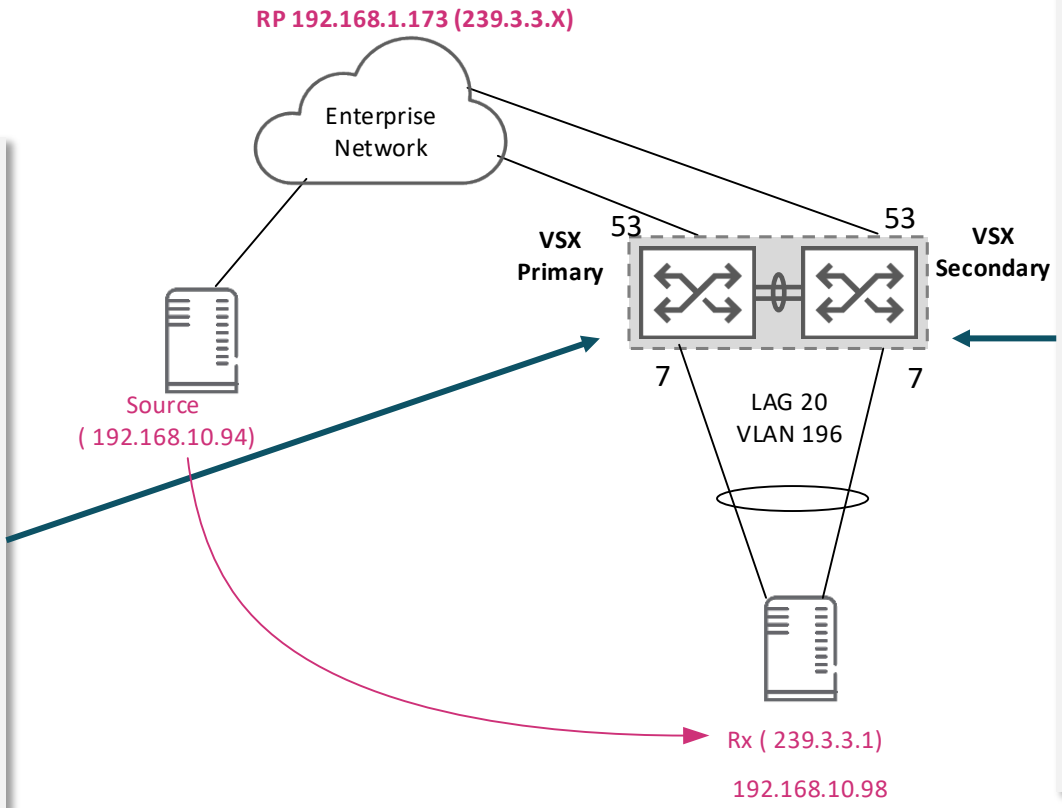
Scan and examine event logs for PIM Bidir

```
show events -d pim-bidid
--------------------------------------------------
Event logs from current boot
--------------------------------------------------
2023-10-09T21:02:18.375398+00:00 8100 pim-bidid[1041]: Event|15304|LOG_INFO|||PIM Bidi Enabled on Interface vlan100
2023-10-09T21:02:18.382854+00:00 8100 pim-bidid[1041]: Event|15304|LOG_INFO|||PIM Bidi Enabled on Interface loopback172
2023-10-09T21:02:18.383625+00:00 8100 pim-bidid[1041]: Event|15304|LOG_INFO|||PIM Bidi Enabled on Interface vlan116
2023-10-09T21:02:18.383786+00:00 8100 pim-bidid[1041]: Event|15304|LOG_INFO|||PIM Bidi Enabled on Interface loopback170
2023-10-09T21:02:18.383933+00:00 8100 pim-bidid[1041]: Event|15304|LOG_INFO|||PIM Bidi Enabled on Interface loopback171
2023-10-09T21:02:18.384489+00:00 8100 pim-bidid[1041]: Event|15304|LOG_INFO|||PIM Bidi Enabled on Interface loopback173
2023-10-09T21:02:18.386702+00:00 8100 pim-bidid[1041]: Event|15304|LOG_INFO|||PIM Bidi Enabled on Interface 1/1/26
2023-10-09T21:02:18.387095+00:00 8100 pim-bidid[1041]: Event|15302|LOG_INFO|||Interface 1/1/26 state Down
2023-10-09T21:02:18.387473+00:00 8100 pim-bidid[1041]: Event|15304|LOG_INFO|||PIM Bidi Enabled on Interface 1/1/3
2023-10-09T21:02:18.388477+00:00 8100 pim-bidid[1041]: Event|15302|LOG_INFO|||Interface 1/1/3 state Down
2023-10-09T21:02:18.388923+00:00 8100 pim-bidid[1041]: Event|15304|LOG_INFO|||PIM Bidi Enabled on Interface 1/1/25
2023-10-09T21:02:18.389352+00:00 8100 pim-bidid[1041]: Event|15302|LOG_INFO|||Interface 1/1/25 state Down
!output omitted
```

Troubleshooting

- VSX Mroute Active-Active
  - Check “Proxy” multicast forwarding states synchronized between VSX peers
  - Check multicast routing OIL
  - VSX Primary will be Designated Forwarder
  - VSX Secondary will be Proxy Designated Forwarder in non failure scenarios

```
VSX-Primary#
show ip mroute 239.3.3.1

IP Multicast Route Entries

VRF : default
Total number of entries : 1

Group Address                    : 239.3.3.1
Source Address                   : any
Neighbor                         :
Incoming interface               : 1/1/53
Multicast Routing Protocol       : PIM-BIDIR
Unicast Routing Protocol         : OSPF
Uptime (HH:MM:SS)                : 1 days
04:56:02
Downstream Interface
Interface        State           VSX Role
-----------      ----------      --------
vlan196          forwarding      DF
```




```
VSX-Secondary#
show ip mroute 239.3.3.1
IP Multicast Route Entries

VRF : default
Total number of entries : 2

Group Address                    : 239.3.3.1
Source Address                   : any
Neighbor                         :
Incoming interface               : vlan196
Multicast Routing Protocol       : PIM-BIDIR
Unicast Routing Protocol         : OSPF
Uptime (HH:MM:SS)                : 00:01:04

Group Address                    : 239.3.3.1
Source Address                   : any
Neighbor                         :
Incoming interface               : 1/1/53
Multicast Routing Protocol       : PIM-BIDIR
Unicast Routing Protocol         : OSPF
Uptime (HH:MM:SS)                : 1 days
03:12:32
Downstream Interface
Interface        State           VSX Role
-----------      ----------      --------
vlan196          forwarding      Proxy DF
```

Troubleshooting

Run Bidir diagnostic , for TAC consumption

```
diagnostics
diag-dump pim-bidid basic
==========================================================================
[Start] Feature pim-bidid Time : Thu Oct 12 15:19:32 2023
==========================================================================
--------------------------------------------------------------------------
[Start] Daemon pim-bidid
--------------------------------------------------------------------------
BIDI APP Info
==================
app data: flow poll timer left: 2

BIDI Vrfs
==================
VRF Id  : 0
VRF Name: default
Table Id: 0
Pim Iface Count: 1
PIM Status: False

BIDI PIM Routers
==================
vrf: 0, pim_status: enabled, active-active: disabled
Static RP List:
Group Range: 239.1.1.0/24, Static RP Uuid: a072c2cd-9e39-4dd7-8be0-1405c9535677
```

Check IGMP to get the full holistic picture from a receiver perspective

```
#show ip igmp group 239.2.2.1
IGMP group information for group 239.2.2.1
Interface Name   : vlan164
VRF Name         : default

Group Address    : 239.2.2.1
Last Reporter    : 192.168.10.66

                                V1        V2        Sources   Sources
Vers Mode Uptime    Expires   Timer     Timer     Forwarded Blocked
---- ---- --------- --------- --------- --------- --------- --------
2         1d 10h 57m 2m 28s              2m 28s
# show ip igmp int vlan 164

IGMP Configured Version  : 3
IGMP Operating Version   : 3
Querier State              : Querier
Querier IP [this switch]   : 192.168.10.65
Querier Uptime             : 1d 12h 12m
Querier Expiration Time    : 1m 58s
IGMP Snoop Enabled on VLAN : True

Active Group Address   Vers Mode Uptime    Expires
---------------------- ---- ---- --------- ---------
239.2.2.1              2         1d 10h 57m 4m 20s
```

Thank you

© 2023 Hewlett Packard Enterprise Development LP