Nexus Validation Test Phase 4.1

1 Introduction

The Cisco Nexus line of data center product hardware and software must pass Cisco's comprehensive quality assurance process, which includes a multistage approach comprising extensive unit test, feature test, and system-level test. Each successive stage in the process adds increasingly higher levels of complexity in a multidimensional mix of features and topologies.

This document describes the NVT Phase 3.7 network topologies, hardware and software configurations, test procedures and findings.

NVT Phase 3.7 testing is performed on the following networks:

- **N9k-GET Enterprise Network**: This network focuses on building and operating a data center with a typical 3-layer design. This network uses vPC and PVLAN to deliver high availability to servers connecting to data centers.

- **N9k-GET Financial Network:** This network focuses on building and operating a data center with a typical spine-leaf design. This network focuses on multicast deliverance for financial and trading floor customers.

- **Data Center 1 (DC1):** This network focuses on building and operating a data center with the Nexus 7000 SUP1 at core and Nexus 7000 SUP2 at aggregation as routing and switching component. It also covers interoperability with the Nexus 5000, Nexus 3000, Nexus 2000,

Catalyst 6500/4500 switches. This network uses virtual Port-Channel (vPC) and Fabric-Path (vPC+) to deliver highly available unicast and multicast services.

- **ENT**: This test profile primarily is built to validate Enterprise customer profiles. In the first phase, we have validated TIER I Enterprises customer profile. The test bed is built with new hardware covering existing feature set for future deployments. Nexus7000 SUP2E/M2 is used at core and N7700 SUP2E/F2E/F3(40G) at aggregation layer. It also covers interoperability with Nexus 6000 and Nexus 3000 switches and covers L3 Agg with L3 ToR and L3 Agg with L2 ToR.
- **MSDC**: This profile focuses on scale requirements of Massively Scalable Data Centers (MSDC). It uses a fully loaded 7018 peering with another 7018 and uses F2 line card and Sup2. Tests were done with BGP and OSPF as IGP protocols.

Operation: Network management including SNMP poling and inventory collection is performed through Data Center Network Management (DCNM) from Cisco and netMRI from Infoblox, TACAS+ authentication and syslog server. NetFlow is configured to export third-party NetFlow Collector Scrutinizer on certain test beds. Real hosts are connected to Nexus access switches by using NIC teaming (in both active mode and On mode). UCS-B series are connected to Nexus access switches thru fabric interconnect. NTP is synced to the server.

2 NVT Scale Validated

2.1 N9500 GET Enterprise and Financial Network

N9500 GET		
Feature/Parameter	**Enterprise**	**Financial**
VLAN	500	500
SVI	250	100
IPv4 hosts x subnet	400	20
vPC	300	100
VLANS x vPC	200	400
Port-channel (excluding peer-link, vPC)	300	100
HSRPv4 version2	250	100
HSRPv6 version2	100	100
VRF	3	3
FEX	7	0
VLAN per FEX	48	0
Primary PVLAN	10	0
Secondary PVLAN	16	0
Physical Port used for PVLAN	121	0
Host Mode	40	0
Promiscuous Access	20	0
Promiscuous Trunk	101	0
Trunk Secondary	3	0
OSPF Peers default VRF	4	4
OSPF Routes/(Path)	20/(2)	20/(2)

eBGP IPv4 Sessions	6	6
BFD neighbors	260	100
PIM neighbors	30	50
PIM ASM groups	200	800
Sources per ASM group	30	10
OIFs per ASM group	20	2
PIM SSM groups	0	10
Sources per SSM group	0	25
OIFs per SSM group	0	20
PIM Bidir groups*	0	1000
Sources per Bidir group*	0	20
OIFs per Bidir group*	0	2
PIM Bidir groups	0	100
Sources per Bidir group	0	20
OIFs per Bidir group	0	20

HW Details: N9K GET	**Model No.**	**SW Version**
N9508	N9K-SUP-A / N9K-X9636PQ / N9K-X9536PQ / N9K-X9464PQ / N9K-X9408PC-CFP2	7.0(3)I2(1)
N9396	N9396 N9K-C9396PX	7.0(3)I2(1)
N7k	N7K SUP2E / F2E	7.2(0)
N5k	N5548 N5K-C5548UP-SUP	7.2(0)N1(1)
C6k	VS-SUP720-10G	12.2(50r)SYS2

2.2 N7000 Data Center 1

HW Details N7000 DC1	Scale
F1	2
M1	5
VDC	4
FEX	6
FEX-HIF	96
ACL/ACE	500
PO Members	2
VLAN / SVI	100
VLAN per FEX	10
FP VLAN	50
MAC	7K
vPC	120
VLAN per vPC	1,10,16
HSRP v2	100
OSPF Neighbor	16
BGP Neighbor	2

PIM Neighbor	256
PBR Sequence	20
Unicast Route	10K
Multicast Route	500
Multicast Source	20
Multicast OIF	10
Multicast Group	100
Fabric Extender	3
FP IS-IS adjacencies	6
FP Number of Switch ID's	20
Primary VLAN	20
Secondary VLAN	25
Port-channel	125
Port-channel in PVLAN	18
VPC PO with PVLAN	10
Host mode	18
Promiscuous Access	2
Promiscuous Trunk	2
Trunk Secondary	7
PVLAN	50

Platform	**Model No.**	**Software**
N7K	N7K SUP2 / F1 / M1	7.2.0
N5K	N5K-C5548UP-SUP	7.0.6.N1.1
N3K	N3K-C3048TP-1GE-SUP	5.0.3.U5.1c
C6K	VS-SUP720-10G	151-1.SY
	WS-SUP720	122-33.SXJ4
C4K	WS-X45-SUP7-E	03.03.02.SG.151-1.SG2
	WS-C4948	150-2.SG6-6.10
N2K	C2224TP; C2248TP; C2232PP	

2.3 N7700 Enterprise 1

N7000/N7700 ENT1	Scale
ACL/ACE	500
VLAN / SVI	2000
MAC	1K
HSRP v2	2000
BGP Neighbor	18
PIM Neighbor	2000
PBR Sequence	200
Unicast Route	22K
Multicast Route	3K

Multicast Group	390
PVLAN	60
vPC Config-Sync	120
Netflow	84
WCCP	58

2.4 N7000 MSDC Scale Profiles

Parameters	**Profile 1 (L3 ToR + L3 Agg)**	**Profile 2-1 (L2 ToR + L3 Agg)**	**Profile 1 (L3 ToR + L3 Agg) w/ RFC5549**
REQ-ID	NXOS-MD-OTT-002-001	NXOS-MD-OTT-002-003	
Description	F2-Series based;	M2-based; Dual-stack;	F2-Series based;
ToR	4948E or N3K	4948E or any other	4948E or any other
Hardware	* Fully loaded F2-series	* 4 M Series	* Fully loaded F2-series
	Single Sup : Test 1 Sup2; Test 2 Sup1	Dual Sup	Dual Sup
VDC	1 VDC	1 VDC	1 VDC
Port channels	8-bundle Port-channel	8-bundle Port-channel	8-bundle Port-channel
L3 ECMP	Test1: 16-way ECMP; Test 2: 32-way ECMP	16-way ECMP	32-way ECMP v4 – 3k host routes and default route
Unicast Routing Protocol	Test1- iBGP 754 Adjacencies; Test 2- OSPF 754 Adj	OSPF, OSPFv3 48 Adj	23 iBGPv4 Adjacencies [leaf nodes that do not support rfc5549 and 16 simulated gateway/loadbalancer/firewall nodes] 755 iBGPv6 Adjacencies that support rfc5549 1 N7700 leaf node that support rfc5549
Multicast Routing Protocols	N/A	PIM-SM, MSDP w/anycast RP	N/A
FHRP	N/A	2000 HSRP groups (1000 HSRPv4 + 1000 HSRPv6)	NA

Fast Detection (for IGP, PIM, FHRP)	Test 1: BFD: 250ms x 3; Test 2: Default Timers	Test 1: BFD: 250ms x 3; Test 2: Default Timers	Test 1: BFD: 250ms x 3; Test 2: Default Timers
Dual Stack	Yes	Yes	Yes, on supporting leaf nodes
Number of /32 & /128 host entries	N/A	80K split (40K IPv4 + 40K IPv6)	N/A
Number of Unicast IGP	10K IPv4 300 IPv6	10K	10K
Number of Multicast Routes (S,G)	N/A	15K	N/A
VLANs	N/A	250	N/A
SVI (same as VLANs)	N/A	250	N/A
ACLs	N/A	250	N/A
ARP/NDP (30seconds ARP refresh rate)	N/A	80K (40K ARP and 40K NDP)	N/A
ARP/NDP Learning Rate	N/A	5000pps	N/A
ARP/NDP Glean rate	N/A	1000pps	N/A
DHCP Relay	N/A	100 pps	N/A
MIB	KPIX	KPIX	KPIX

Hardware

	Model No.	
N7K	**N7K SUP2 / F2**	6.2.14
	N7k SUP2 F2 / M2	6.2.14
N3K	**N3K-C3048TP-1GE-SUP**	6.0(2)U4(0.899)
N3k	**N3K-C3048TP-1GE-SUP**	5.0(3)U5(1)

3 ISSU Matrix

Image	ISSU/COLD BOOT	DC1	ENT	MSDC
6.2.8a > 6.2.14	ISSU			Pass
	COLD BOOT	Pass		
6.2.8b > 6.2.14	ISSU			Pass
	COLD BOOT	Pass	Pass	
6.2.10 > 6.2.14	ISSU	Pass	Pass	Pass
	COLD BOOT			
6.2.12> 6.2.14	ISSU	Pass	Pass	Pass
	COLD BOOT	Pass	Pass	
6.2.14 > 6.2.14.upg	ISSU	Pass	Pass	Pass
	COLD BOOT	Pass	Pass	

4 Profile Details

4.1 N9K GET Enterprise

4.1.1 Network Logical Topology Design Overview

The topologies and test cases validate high-available data center networks in order to provide unified fabric and computing services. This is achieved by using the Nexus 9000 and N7000 Cisco products, with features such as vPC, PVLANs, ACLs and Fabric Extender Modules (FEX).

Description of the Test Network (Core Layer)

The core network is built around two Nexus 9000 chassis (EoR - N9508) and two Nexus 7000 chassis (N7010) with Sup2E.

The four devices are connected as a physical ring. In the N9508 chassis, the following modules had been deployed: N9K-X9636PQ and N9K-X9536PQ

Description of the Test Network (Aggregation Layer)

The aggregation layer is built with two Nexus 9000 chassis (EoR - N9508).

- Two Nexus 9508 (EoR) are fully meshed and connected to the Core network. These chassis contain N9K-X9636PQ, N9K-X9464PX and N9K-X9408PC-CFP2 modules
- In order to enhance port density, the FEX module had been connected to both EoRs.
- The two EoRs are also connected as vPC peers toward the access layer switches.

Description of the Test Network (Access Layer)

The access layer is composed of a pair of Nexus 9396 ToRs and a pair of Nexus 5000, all connected to the aggregation switches (EoRs).

- The Nexus 5000 pair is configured with vPC back-to-back connections to the aggregation switches (EoRs).

- In order to increase the overall port density, FEX modules had been connected to the ToRs.

While the majority of test cases focus on integrated solutions using Nexus switching, modular Catalyst switches are also included for interoperability between NX-OS and IOS.

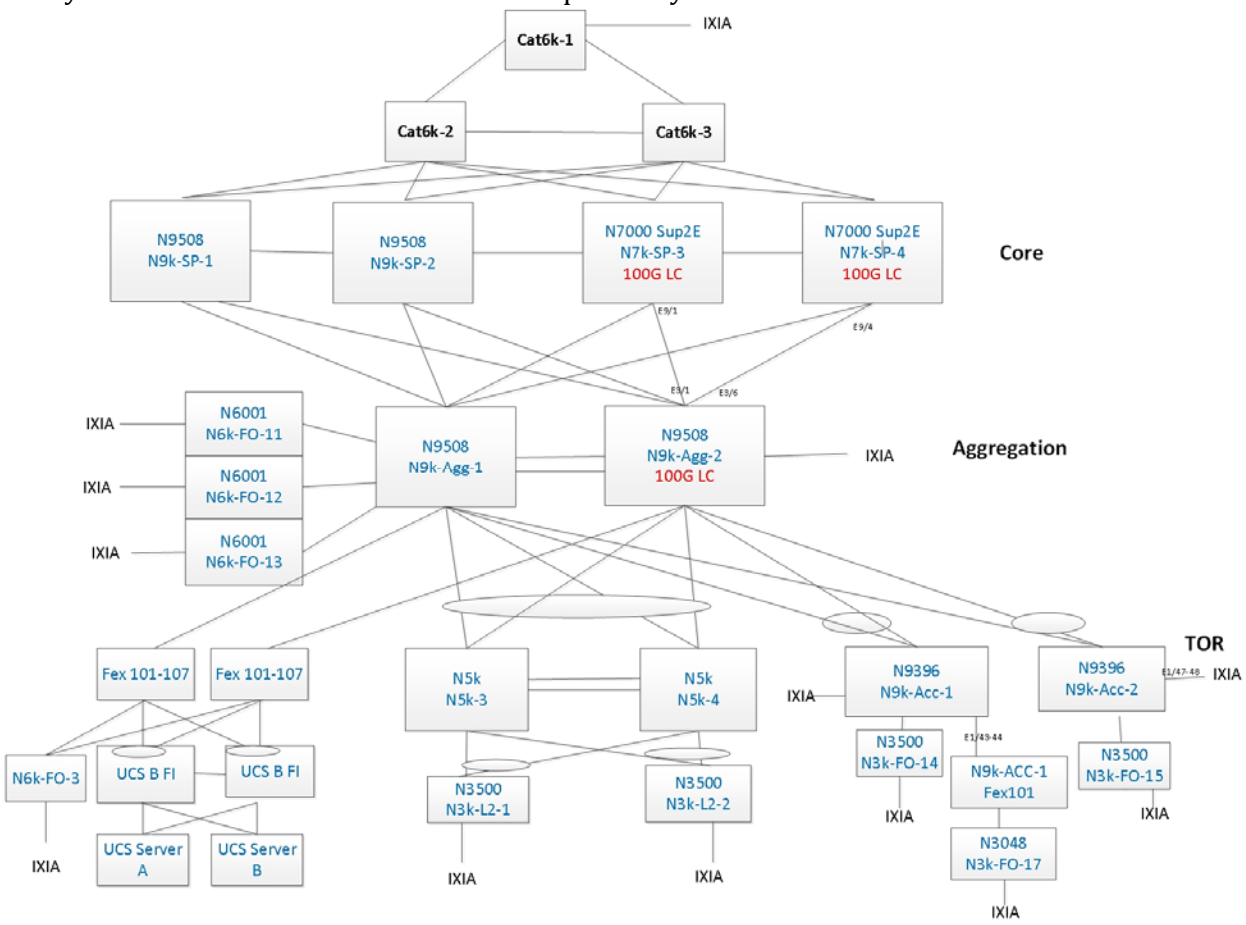


Figure 1: N9K GET Enterprise Topology

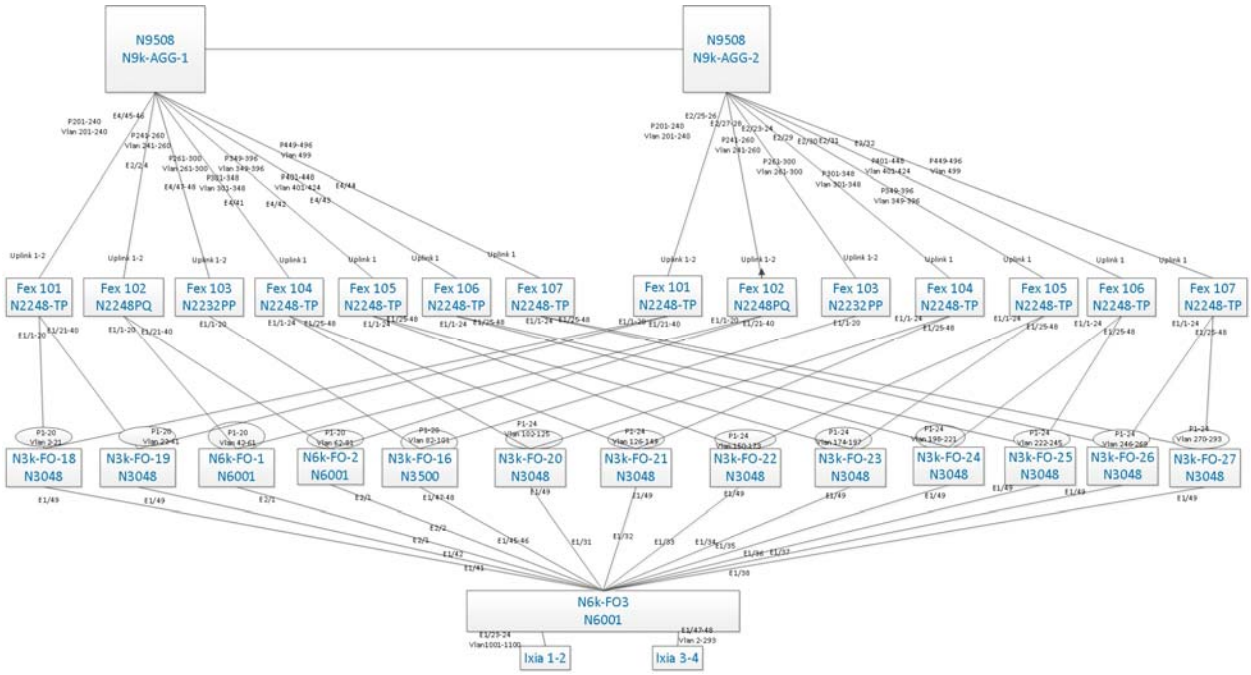


Figure 2: N9K GET Enterprise Topology with FEX

4.1.2 Configuration Details

The following configurations are applied to the test network:

- Common system control, management and accounting: Common system features like SSH, TACACS+, Syslog, SNMP, NTP, SPAN, DNS and Management VRF are configured.
  - feature tacacs+ *# enabling the tacacs+ feature*
  - tacacs server host <ip address> key <0/7> *# configure the tacacs+ server to authenticate users*
  - aaa group server tacacs+ <group name> *# enable server groups for redundancy*
    - server <ip address>
    - use-vrf <vrf_name> *# use-vrf based on server reachability*
  - snmp-server user <user-name> <group-name> auth md5 <pass-phrase> priv <pass-phrase> localizedkey *# SNMP v3 user with authentication enabled*
  - ntp server <ip address> use-vrf management *# enable NTP with server ip address*
  - ntp source-interface mgmt0
  - ntp logging
  - ntp access-group peer ntp-peer *# enable access-list control for NTP peers*
  - ntp access-group query-only ntp-query-on
  - ip domain-name <domain name> *# enable domain-name*
  - interface mgmt0 *# configure mgmt0*
    - ip access-group mgmt-acl in *# enable access-list control for management interface*
    - vrf member management
    - ip address <ip_address >/<mask>

- BGP: eBGP is configured between the core switches and the public cloud and between the distribution switches and the core switches:

- feature bgp *# enable BGP*
- router bgp <autonomous-id> *# BGP autonomous -id*
- router-id <router-id>
- graceful-restart-helper
- log-neighbor-changes
- address-family ipv4 unicast
- maximum-paths <8>
- maximum-paths ibgp <8>
- neighbor <neighbor ip addres> remote-as <remote-AS> *# BGP peer*
- address-family ipv4 unicast
- address-family ipv6 unicast

- OSPF: OSPF is the IGP running across the network. Each distribution-access block is configured as a unique area with the core switches playing the role of the ABR:

  - feature ospf *# enable OSPF for IPv4*
  - router ospf <instance-tag>
  - router-id <router-ID>
  - log-adjacency-changes
  - auto-cost reference-bandwidth 1000000

- PIM-ASM: PIM Sparse Mode/PIM Any Source Multicast is deployed across the network to support multicast features. Static RP with Anycast RP is configured at the core layer:

  - feature pim *# enable PIM*
  - ip pim rp-address <rp-address> group-list <multicast-groups> *# configure static RP for a multicast group range*
  - ip pim anycast-rp <RP-address> <local-source-intf> *# enable anycast RP on all the core routers to synchronize the (S,G) entries between the vPC peers*

- vPC: vPC technology is deployed on the distribution switches N9k-AGG-1 and N9k-AGG-2:

  - feature vpc *# enable vPC*
  - vpc domain <domain-id> *# configure vPC domain-id*
    - peer-switch *# enable peer-switch for faster STP convergence*
    - peer-keepalive destination <ip address> source <ip address> vrf vpc-keepalive *# configure keep-alive link*
    - peer-gateway *# enable peer-gateway to avoid vPC loop*
    - ip arp synchronize *# configure ARP synchronization for faster convergence of ARP tables between the vPC peers*
    - Auto-recovery *# enable auto-recovery in case of network disruption*

- HSRP: HSRP is used as the first hop gateway protocol for hosts:

  - interface <VlanX> *# configure SVI*
  - no ip redirects

- ip address <ip address>/<mask>
- ipv6 address <ipv6 address>/<mask>
- hsrp version 2
- hsrp 1
  - authentication md5 key-string cisco *# enable authentication*
  - preempt delay minimum 200
  - priority 90 forwarding-threshold lower 1 upper 90
  - ip <ip address> *# HSRP IP address*

- IGMP: IGMP is used by hosts to join multicast groups of interest. IGMP snooping is enabled on all switches in the distribution-access blocks to prevent unnecessary flooding of multicast data traffic:

  - ip igmp snooping *# by default enabled on Nexus*

- LACP: LACP is used for link aggregation to form port-channels across the network:

  - feature lacp *# enable LACP, by default LACP is used on all port-channel*

- UDLD: UDLD aggressive mode is configured across the network to detect and prevent unidirectional links:

  - feature udld *# enable feature UDLD*
  - udld aggressive *# UDLD aggressive mode is enabled to fasten the detection a unidirectional link*

- STP: Rapid Spanning Tree Protocol is used to prevent Layer two loops in the distribution-access blocks. The spanning tree root is placed at the aggregation level. Root Guard is configured on the aggregation level to enforce root placement. BPDU Filter, BPDU Guard and PortFast Edge trunk are configured on the access ports towards hosts.

  - interface port-channel <PoX> *# configure port-channel X*
    - switchport
    - switchport access vlan <vlan>
    - switchport mode trunk
    - switchport trunk native vlan <native-vlan>
    - switchport trunk allow vlan <Vlan-range>
    - spanning-tree port type edge trunk *# enable host*
    - spanning-tree bpdufilter enable *# configure bpdufilter*

- DHCP Relay: DHCP reply is configured in distribution switches to forward DHCP packets between host clients and the DHCP server:

  - feature dhcp *# enable feature DHCP*
  - ip dhcp relay *# enable DHCP relay agent*
  - interface VLAN XXX *# configure DHCP server address*
    - ip dhcp relay address x.x.x.x

- PVLAN: PVLAN is configured in the network and is the main focus of testing for N9k GET. The following PVLAN components are covered in the network:

    - PVLAN primary and secondary vlan (isolated and community)
    - PVLAN promiscuous access port and promiscuous trunk port on distribution switches
    - PVLAN host ports on TOR switches
    - PVLAN Secondary trunk port on distribution switches and TOR switches
    - PVLAN primary VLAN SVI on distribution switches

    - feature private-vlan *# enable feature private-vlan*
    - vlan <vlan-id> *# configure primary vlan*
        - private-vlan primary
    - vlan <vlan-id>
        - private-vlan <isolated/community> *# configure secondary vlan*
        - private-vlan association <vlan-id> *# configure association with primary vlan*
    - interface Ethernet <fex/x/y > *# configure secondary trunk port*
        - switchport
        - switchport mode private-vlan trunk secondary
        - switchport private-vlan trunk native vlan xxx *# configure native vlan*
        - switchport private-vlan trunk allowed vlan
        - switchport private-vlan association trunk <primary> <secondary>
    - interface Ethernet <x/y > *# configure promiscuous trunk port*
        - switchport
        - switchport mode private-vlan trunk promiscuous
        - switchport private-vlan trunk native vlan xxx*# configure native vlan*
        - switchport private-vlan trunk allowed vlan
        - switchport private-vlan mapping trunk <primary secondary1, secondary2,…> *# configure primary to secondary mapping*
    - interface Ethernet <x/y > *# configure promiscuous port*
        - switchport
        - switchport mode private-vlan promiscuous *# configure promiscuous access*
        - switchport private-vlan mapping <primary secondary> *# configure primary to secondary mapping*
    - interface Ethernet <fex/x/y > *# configure host port*
        - switchport
        - switchport mode private-vlan host *# configure host*
        - switchport private-vlan host-association <primary secondary> *# configure primary to secondary host-association*

- interface vlan XXX *# configure primary VLAN SVI*
  - private-vlan mapping <secondary vlan>

- Jumbo MTU (9000): Jumbo frames are configured throughout the N9k-GET Enterprise network:

  - policy-map type network-qos jumbo *# enable Jumbo frames on N9k*
    class type network-qos class-default
      mtu 9000
  - system qos *# applies the QoS policy to the control-plane*
    service-policy type network-qos jumbo

- FEX/QoS: is enabled in egress on the EoRs:

  - policy-map type queuing <name-out>
    class type queuing c-out-q3
      bandwidth percent 30
      bandwidth remaining percent 10
    class type queuing c-out-q2
      bandwidth percent 30
      bandwidth remaining percent 40
    class type queuing c-out-q1
      bandwidth percent 20
      bandwidth remaining percent 50
    class type queuing c-out-q-default
      bandwidth percent 10
      bandwidth remaining percent 0
  - interface port-channel101 *# apply the QoS policy to a FEX PO*
    service-policy type queuing output <name-out>
  - hardware access-list tcam region fex-qos 256 *# enable TCAM carving for FEX-QoS entries*

  Whereas the following is configured in ingress on the ToRs:

  - policy-map type queuing <name-in>
    class type queuing c-out-q3
      bandwidth percent 95
    class type queuing c-out-q2
      bandwidth percent 5
    class type queuing c-out-q1
      bandwidth remaining percent 40
    class type queuing c-out-q-default
      bandwidth remaining percent 50

- system qos
  service-policy type qos input <name-in>

- ACL: access-control-lists are configured on some SVIs of the EoRs at the aggregation layer:

  - ip access-list <name>
    statistics per-entry
    10 remark --- UCF Cell1 NOZ4 Vlan 93 ACL ver 3.17 09-12-2013
    20 remark --- allow UCF Cell1 NOZ4 Vlan 93 to secondary VLAN93

  - interface VlanX
    ip access-group <name> in

- Checkpoint/Rollback: configuration(s) can be saved onto the system and they can be retrieved with the rollback option.

- BFD (for BGP/PIM/HSRP): is enabled to fasten the detection of BFD client failures:

  - feature bfd
    bfd interval 250 min_rx 250 multiplier 3
    hsrp bfd all-interfaces
  - interface Ex/y
    ip pim bfd-instance
  - interface <poX>
    ip pim bfd-instance
  - interface vlan <vlanX>
    no bfd echo
    ip pim bfd-instance
  - router bgp <as-number>
    neighbor <neighbor-ip-address>
    bfd

4.2 N9K GET Financial

4.2.1 Network Logical Topology Design Overview

The topologies and test cases validate high-available data center networks in order to provide unified fabric and computing services to financial and trading floor customers. This is achieved by using the Nexus 9000 and N7000, with features such as vPC and the Classical Ethernet access network.

Description of the Test Network (Spine Layer)

The spine network is built around two Nexus 9000 chassis (N9508) and two Nexus 7000 chassis (N7010) with Sup2E.

The four devices are connected as a physical ring. In the N9508 chassis, the following modules had been deployed: N9K-X9636PQ and N9K-X9536PQ.

Description of the Test Network (Leaf Layer)

The leaf layer is built with two pairs of Nexus 9000 chassis (ToR - N9396 and N9372). Both ToR pairs are fully meshed and connected to the spine layer.

- Two ToRs (N9396) are connected as vPC peers toward the access layer switch (N5000).
- Two ToRs (N9372) and one N5000 are connected in a typical triangle switched access network running STP.

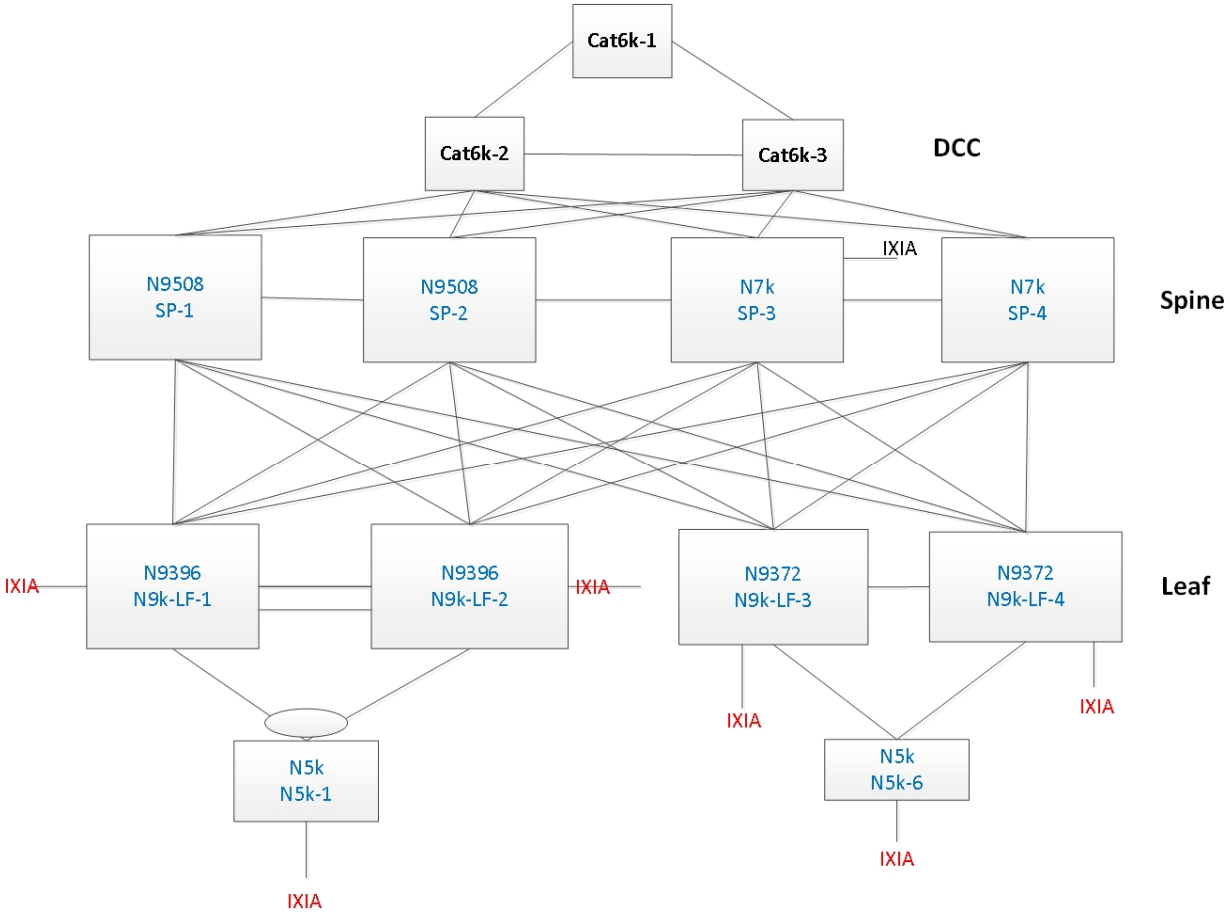


Figure 3: N9K GET Financial Topology

4.2.2 Configuration Details

The following configurations are applied to the test network:

- Common system control, management and accounting: Common system features like SSH, TACACS+, Syslog, SNMP, NTP, SPAN, DNS and Management VRF are configured.
  - feature tacacs+ *# enabling the tacacs+ feature*
  - tacacs server host <ip address> key <0/7> *# configure the tacacs+ server to authenticate users*
  - aaa group server tacacs+ <group name> *# enable server groups for redundancy*
    - server <ip address>
    - use-vrf <vrf_name> *# use-vrf based on server reachability*

- snmp-server user <user-name> <group-name> auth md5 <pass-phrase> priv <pass-phrase> localizedkey *# SNMP v3 user with authentication enabled*
- ntp server <ip address> use-vrf management *# enable NTP with server IP address*
- ntp source-interface mgmt0
- ntp logging
- ntp access-group peer ntp-peer *# enable access-list control for NTP peers*
- ntp access-group query-only ntp-query-on
- ip domain-name <domain name> *# enable domain-name*
- interface mgmt0 *# configure mgmt0*
  - ip access-group mgmt-acl in *# enable access-list control for management interface*
  - vrf member management
  - ip address <ip_address >/<mask>

- BGP: eBGP is configured between the core switches and the public cloud and between the distribution switches and the core switches:

  - feature bgp *# enable BGP*
  - router bgp <autonomous-id> *# BGP autonomous -id*
  - router-id <router-id>
  - graceful-restart-helper
  - log-neighbor-changes
  - address-family ipv4 unicast
  - maximum-paths <8>
  - maximum-paths ibgp <8>
  - neighbor <neighbor ip addres> remote-as 100090 *# BGP peer*
  - address-family ipv4 unicast
  - address-family ipv6 unicast

- OSPF: OSPF is the IGP running across the network. Each distribution-access block is configured as a unique area with the core switches playing the role of the ABR:

  - feature ospf *# enable OSPF for IPv4*
  - router ospf <instance-tag>
  - router-id <ip address>
  - log-adjacency-changes
  - auto-cost reference-bandwidth 1000000

- PIM-ASM: PIM Sparse Mode/PIM Any Source Multicast is deployed across the network to support multicast. Static RP with Anycast RP is configured at the spine layer:

  - feature pim *# enable PIM*
  - ip pim rp-address <rp-address> group-list <multicast-groups> *# configure static RP for a multicast group range*
  - ip pim anycast-rp <RP-address> <local-source-intf> *# enable anycast RP on all the core routers to synchronize the (S,G) entries*

- PIM-SSM: PIM Source Specific Multicast is deployed across the network to support multicast. SSM is supported only on the Classical Ethernet access network:

    - ip pim ssm range 235.3.0.0/24

- PIM-Bidir: PIM Bidir Multicast is deployed across the network to support multicast. BSR RP is located at the spine layer:

    - ip pim bsr bsr-candidate <bsr-candidate-ip-address> priority <priority>
      ip pim bsr rp-candidate <rp-ip-address> prefix-list BIDIR priority <priority> bidir
    - ip pim bsr listen forward
    - hardware access-list tcam region mcast_bidir 512 *# enable TCAM carving for bidir entries*

    Static RP is also being configured and tested with PIM Bidir

- vPC: vPC technology is deployed on the ToR switches (N9396):
  .
    - feature vpc *# enable vPC*
    - vpc domain <domain-id> *# configure vPC domain-id*
        - peer-switch *# enable peer-switch for faster STP convergence*
        - peer-keepalive destination <ip address> source <ip address> vrf vpc-keepalive *# configure keep-alive link*
        - peer-gateway *# enable peer-gateway to avoid vPC loop*
        - ip arp synchronize *# configure ARP synchronization for faster convergence of ARP tables*
        - Auto-recovery *# enable auto-recovery in case of network disruption*

- HSRP: HSRP is used as the first hop gateway protocol for hosts:

    - interface <VlanX> *# configure SVI*
    - no ip redirects
    - ip address <ip address>/<mask>
    - ipv6 address <ipv6 address>/<mask>
    - hsrp version 2
    - hsrp 1
        - authentication md5 key-string cisco *# enable authentication*
        - preempt delay minimum 200
        - priority 90 forwarding-threshold lower 1 upper 90
        - ip <ip address> *# HSRP IP address*

- IGMPv2 and Snooping: IGMPv2 is used by hosts to join multicast groups of interest. IGMP snooping is enabled on all switches in the distribution-access blocks to prevent flooding of multicast data traffic:

    - ip igmp snooping *# by default enabled on Nexus*

- IGMPv3 and Snooping: IGMPv3 is used by hosts to join multicast groups of interest. IGMP snooping is enabled on all switches in the distribution-access blocks to prevent flooding of multicast data traffic:
    - ip igmp version 3
- LACP: LACP is used for link aggregation to form port-channels across the network:
    - feature lacp *# enable LACP, by default LACP is used on all port-channel*
- UDLD: UDLD aggressive mode is configured across the network to detect and prevent unidirectional links:
    - feature udld *# enable feature UDLD*
    - udld aggressive *# UDLD aggressive mode is enabled to fasten the detection a unidirectional link*
- STP: Rapid Spanning Tree Protocol is used to prevent Layer 2 loops in the distribution-access blocks. The spanning tree root is placed on the aggregation level. Root Guard is configured on the aggregation level to enforce root placement. BPDU Filter, BPDU Guard and PortFast Edge trunk are configured on the access ports towards hosts.
    - interface port-channel <PoX> *# configure port-channel*
        - switchport
        - switchport access vlan <vlan>
        - switchport mode trunk
        - switchport trunk native vlan <native-vlan>
        - switchport trunk allow vlan <Vlan-range>
        - spanning-tree port type edge trunk *# enable host*
        - spanning-tree bpdufilter enable *# configure bpdufilter*
- Jumbo MTU (9000): Jumbo frames are configured throughout the N9k-GET Enterprise network:
    - policy-map type network-qos jumbo # enable Jumbo frames
      class type network-qos class-default
      mtu 9000
    - system qos # applies the qos policy to the control-plane
      service-policy type network-qos jumbo
- Checkpoint/Rollback: configuration(s) can be saved onto the system and they can be retrieved with the rollback option.
- BFD (for BGP/PIM/HSRP): is enabled to fasten the detection of BFD client failures:

- feature bfd
  bfd interval 250 min_rx 250 multiplier 3
  hsrp bfd all-interfaces
- interface Ex/y
  ip pim bfd-instance
- interface <poX>
  ip pim bfd-instance
- interface vlan <vlanX>
  no bfd echo
  ip pim bfd-instance
- router bgp <as-number>
  neighbor <neighbor-ip-address>
  bfd

4.3 N7000 Data Center (DC1)

4.3.1 Network Logical Topology Design Overview

The topologies and test cases validate high-available data center networks in order to provide unified fabric and computing services. This is achieved by using the Nexus 7010, Nexus 5548 with features such as vPC and Fabric-Path.

Nexus 7010 installed with F1 and M1 line cards to provide legacy L2 & L3 and vPC leg port channel connectivity with peer devices and F1 ports used for fabric-path port-channel with N5000 peers. Array of Nexus 2000 connected with Nexus 7010 as fabric-extenders (FEX).

Access layer switches in this data center extended to IXIA (Traffic Generator) ports to simulate end hosts/servers to send and receive unicast & multicast traffic

The data center site is built around the Nexus 7010 SUP2 at aggregation. This data center site is split into two halves:

DC101: (vPC)

- Nexus 7000 with vPC to Nexus 5000, C4K/C6K for access
- Nexus 7000 with legacy ether-channel (trunk) with C4K
- Nexus 7000 with Nexus 2000 FEX
- Nexus 7000 with L3 to Nexus 3048

DC102: (vPC+)

- Nexus 7000 (spine) with Fabric-Path to Nexus 5000 (leaf)
- Nexus 7000 with vPC to C6K for access
- Nexus 7000 with L3 to Nexus 3048

DC101

Device	Platform	Position	VPC Status
DC101-5 (VPC Sec.)	N7K	Aggregation	VPC
DC101-6 (VPC Pri.)	N7K	Aggregation	VPC
DC101-7	C6K	Access	VPC
DC101-8	C6K	Access	VPC
DC101-17 (VPC Pri.)	N5K	Access	Dual sided VPC
DC101-18 (VPC Sec.)	N5K	Access	Dual sided VPC
DC101-27 (VPC Pri.)	N5K	Access	Dual sided VPC
DC101-28 (VPC Sec.)	N5K	Access	Dual sided VPC
DC101-37	C4K	Access	Non VPC
DC101-38	C4K	Access	Non VPC
FEX 101, 102, 103	N2K	Access FEX	VPC - HIF
DC101-47	N5K	Access	L3

DC102

Device	Platform	Position	VPC Status
DC102-51 (VPC+ Sec.)	N7K	Aggregation	Fabric-path - Spine
DC102-52 (VPC+ Pri.)	N7K	Aggregation	Fabric-path - Spine
DC102-53 (VPC+ Sec.)	N7K	Aggregation	Fabric-path - Spine
DC102-54 (VPC+ Pri.)	N7K	Aggregation	Fabric-path - Spine
DC102-701 (VPC+ Pri.)	N5K	Access	Fabric-path - Leaf
DC102-702 (VPC+ Sec.)	N5K	Access	Fabric-path - Leaf
DC102-703 (VPC+ Pri.)	N5K	Access	Fabric-path - Leaf
DC102-704 (VPC+ Sec.)	N5K	Access	Fabric-path - Leaf
DC102-705 (VPC+ Pri.)	N5K	Access	Fabric-path - Leaf
DC102-706 (VPC+ Sec.)	N5K	Access	Fabric-path - Leaf
DC102-7011 (VPC+ Pri.)	N5K	Access	Dual sided VPC
DC102-7012 (VPC+ Sec.)	N5K	Access	Dual sided VPC
DC102-17	C6K	Access	VPC
DC102-18	C6K	Access	VPC
DC102-27	C6K	Access	VPC
DC102-47	N3K	Access	L3

4.3.2 Configuration Details

The following configurations are applied to the test network:

- Common system control, management and accounting: Common system features like SSH, TACACS+, Syslog, SNMP, NTP, SPAN, DNS and Management VRF are configured

- BGP: eBGP is configured between the core switches and the public cloud.

- OSPF: OSPF is the IGP running across the network. Each aggregation-access block is configured as a unique area with the core switches playing the role of the ABR.

- PIM-SM: PIM Sparse Mode/PIM Any Source Multicast is deployed across the network to support multicast. Each aggregation-access block is configured with the RP for the locally sourced groups.

- MSDP Anycast RP: MSDP is deployed to exchange source information between Anycast RPs.

- vPC: vPC technology is deployed in the aggregation-access block DC101 as shown in Figure 1. In addition, dual-sided vPC is configured between the Nexus 7000 and Nexus 5000 switches

- FP: FabricPath is deployed in the aggregation block DC102. The spine layer is comprised of Nexus 7000 switches, and the leaf switches are deployed using Nexus 5000 switches.

- VLAN trunking: VLAN trunking is used in the aggregation-access blocks to maintain segregation and security.

- STP: Rapid Spanning Tree Protocol is used to prevent Layer 2 loops in the aggregation-access blocks. The spanning tree root is placed on the aggregation level. Root Guard is configured on the aggregation level to enforce root placement. BPDU Filter, BPDU Guard and Port-Fast Edge are configured on the access ports towards hosts.

- HSRP: HSRP v2 is used as the first hop gateway protocol for hosts; HSRP configured at N7K aggregation VDCs (DC101-5, DC101-6 & DC102-51, DC102-52, DC102-53, DC102-54)

- FEX: Three set of Fabric Extenders (Nexus 2000) FEX101, FEX103 & FEX104 are deployed on Nexus 7000 at DC101

- FEX uplinks are legacy port channels where as FEX host ports are configured as vPC.

- IGMP: IGMP is used by hosts to join multicast groups of interest. IGMP snooping is enabled on all switches in the aggregation-access blocks to prevent flooding of multicast data traffic.

- LACP: LACP is used for link aggregation to form port-channels across the network.

- UDLD: UDLD aggressive mode is configured across the network to detect and prevent unidirectional links

- PVLAN: Promiscuous / Isolated / Community PVLAN configured between Nexus 7000 to Nexus 5000, C6K and C4K over vPC as well as non VPC at DC101. PVLAN traffic configured over Fabric-path PO at DC102

The following configurations are applied to the test network:

- Common system control, management and accounting: Common system features like SSH, TACACS+, Syslog, SNMP, NTP, SPAN, DNS and Management VRF are configured.
  - feature tacacs+ *# enabling the tacacs feature*
  - tacacs server host <ip address> key <0/7> *# configure the tacacs server to authenticate users*
  - aaa group server tacacs+ <group name> *# enable server groups for redundancy*
    - server <ip address>
    - use-vrf <vrf_name> *# use-vrf based on server reachability*
  - snmp-server user <user-name> <group-name> auth md5 <pass-phrase> priv <pass-phrase> localizedkey *# snmp v3 user with authentication enabled*
  - ntp server <ip address> *# enable ntp with server ip address*
  - ip domain-name <domain name> *# enable domain-name*
  - interface mgmt0 *# configure mgmt0*
    - vrf member management
    - ip address <ip_address >

- BGP: eBGP is configured between the core switches and the public cloud.
  - feature bgp *# enable bgp*
  - router bgp <autonomous-id> *# bgp autonomous -id*
  - router-id <router-id>
  - graceful-restart stalepath-time <120>
  - log-neighbor-changes
  - address-family ipv4 unicast
  - redistribute direct route-map <acl-name> *# route-map used for redistribution directly connected subnets*
  - redistribute ospf 1 route-map <acl-name> *# route-map used for redistribution OSPF routes*
  - maximum-paths <8>
  - maximum-paths ibgp <8>
  - neighbor <neighbor ip addres> remote-as 100090 *# BGP peer*
  - address-family ipv4 unicast
  - prefix-list NO_SELF in *# acl configured to restrict prefix import*

- OSPF: OSPF is the IGP running across the network. Each aggregation-access block is configured as a unique area with the core switches playing the role of the ABR.
  - feature ospf *# enable ospf for IPv4*
  - feature ospfv3 *# enable ospf for IPv6*
  - router ospf <instance-tag>
  - router-id <ip address>
  - redistribute bgp <as_no> route-map <acl-name> *# route-map used for redistribution for bgp routes*
  - log-adjacency-changes
  - timers throttle spf 100 200 500
  - timers throttle lsa 50 100 300
  - auto-cost reference-bandwidth 1000000
  - default-metric <1>

- PIM-SM: PIM Sparse Mode/PIM Any Source Multicast is deployed across the network to support multicast. Each aggregation-access block is configured with the RP for the locally sourced groups.
    - feature pim *# enable pim*
    - ip pim rp-address <rp-address> group-list <multicast-groups> *# configure static RP for a multicast group range*
    - ip pim send-rp-announce loopback2 prefix-list <multicast-groups> *# configure candidate auto-rp*
    - ip pim send-rp-discovery loopback2 *# configure auto-rp mapping-agent*
    - ip pim ssm range <> *# configure pim ssm for default range*
    - ip pim auto-rp forward listen *# enable auto-rp messages forwarding*

- MSDP Anycast RP: MSDP is deployed to exchange source information between Anycast RPs.
    - feature msdp *# enable msdp*
    - ip msdp originator-id <interface> *# configure source interface for msdp peering, generally loopback interface*
    - ip msdp peer <ip address> connect-source <interface> *# configure peer address*

- vPC: The vPC technology is deployed in the aggregation-access block DC1-Dist-N7k-101 as shown in Figure 1. In addition, dual-sided vPC is configured between the Nexus 7000 and Nexus 5000 switches
    - feature vpc *# enable vpc*
    - vpc domain <domain-id> *# configure vpc domain-id*
        - peer-switch *# enable peer-switch for faster STP convergence*
        - role priority 200 *# configure priority*
        - peer-keepalive destination <ip address> source <ip address> vrf vpc-keepalive *# configure keep-alive link*
        - peer-gateway *# enable peer-gateway to avoid vPC loop*
        - track <id> *# track the L3 core connectivity to avoid black-hole*
        - ip arp synchronize *# configure arp synchronize for faster convergence of address tables*

- FP: FabricPath is deployed in the aggregation block DC1-Dist-N7k-102. The spine layer is comprised of Nexus 7000 switches and the leaf switches are deployed using Nexus 5000 switches.
    - feature-set fabricpath *# configure feature-set fabricpath*
    - vlan <vlan-range>
        - mode fabricpath *# configure vlan-range in fabric path*
    - fabricpath switch-id <switch-id> *# configure switch-id*
    - vpc domain <domain-id> *# configure vpc domain-id*

        - fabricpath switch-id <vpc+_switch-id> *# configure virtual switch for peers present on the network*
    - interface port-channel <po> *# configure fabricpat interface*
        - switchport mode fabricpath *# configure fabricpath*

- STP: Rapid Spanning Tree Protocol is used to prevent Layer 2 loops in the aggregation-access blocks. The spanning tree root is placed on the aggregation level. BPDU Filter and PortFast Edge are configured on the access ports towards the hosts.
    - interface port-channel <po> *# configure port-channel*
        - switchport
        - switchport access vlan <vlan>
        - spanning-tree port type edge *# enable host*
        - spanning-tree bpdufilter enable *# configure bpdufilter*

- HSRP: HSRP is used as the first hop gateway protocol for hosts.
    - interface Vlan<id> *# configure svi*
    - ip access-group <acl> in *# enable access-list*
    - ip access-group <acl> out
    - no ip redirects
    - ip address <ip address>
    - hsrp version 2
    - hsrp 1
        - authentication md5 key-string cisco *# enable authentication*
        - preempt delay minimum 200
        - priority 200
        - ip <ip address> *# HSRP IP address*

- FEX: Fabric Extenders (Nexus 2000) are deployed on Nexus 7000

- IGMP: IGMP is used by hosts to join multicast groups of interest. IGMP snooping is enabled on all switches in the aggregation-access blocks to prevent flooding of multicast data traffic.
    - ip igmp snooping *# by default enabled on Nexus*

- LACP: LACP is used for link aggregation to form port-channels across the network.
    - feature lacp *# enable LACP, by default LACP is used on all port-channel*

- UDLD: UDLD aggressive mode is configured across the network to detect and prevent unidirectional links
    - feature udld *# enable feature udld*
    - udld aggressive *# udld aggressive mode is enabled to re-establish the connection with the neighbor*

- PVLAN: Private VLAN configured at DC101 between Nexus 7000 VPC peers to:
    - Nexus 5000
    - CAT 6500

  Following Private VLAN modes configured:
    - Promiscuous
    - Isolated (host)

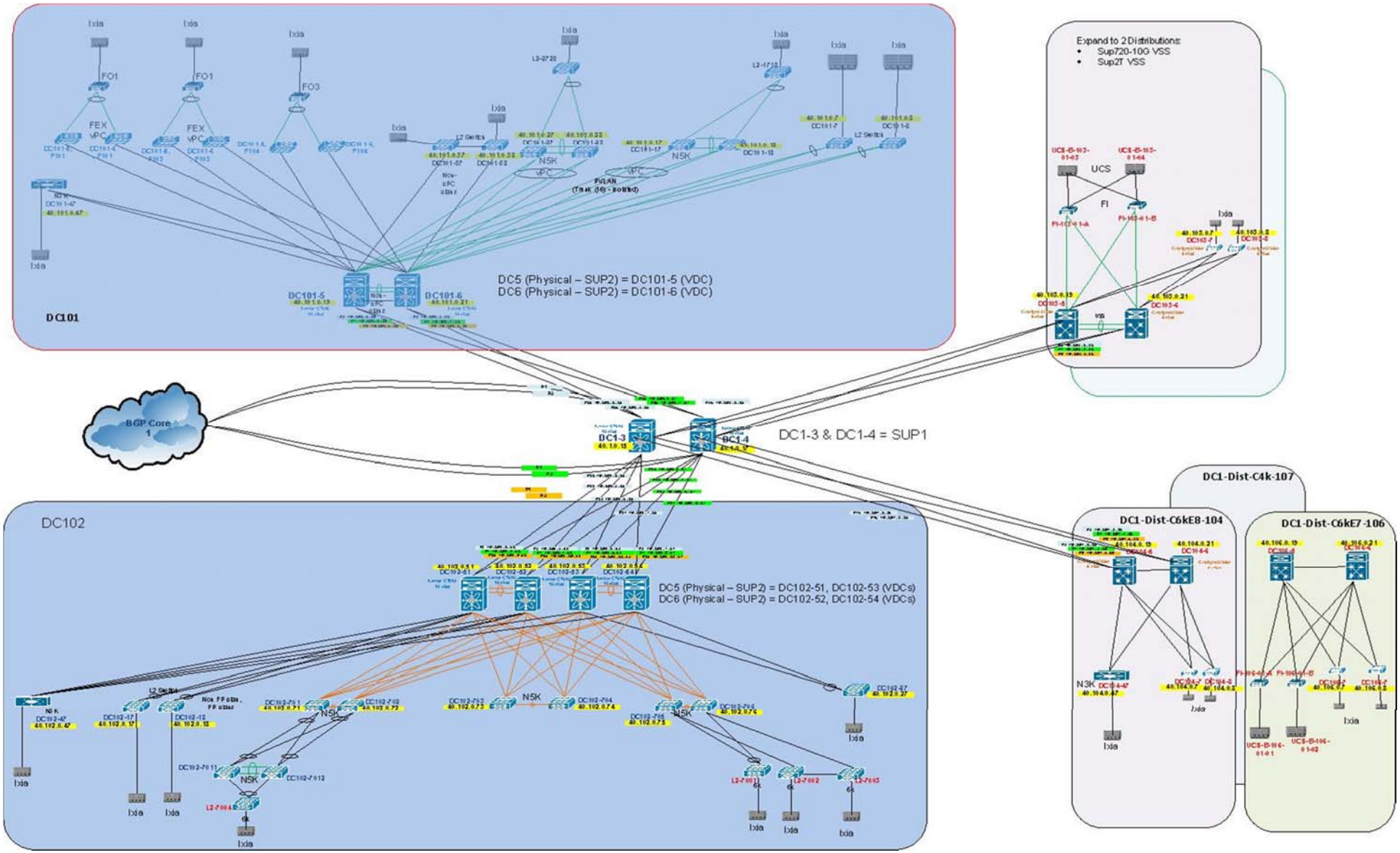
- Isolated (trunk)
- Single PVLAN association in a port-channel (host mode)
- Multiple PVLAN association in a port-channel (trunk mode)
- PVLAN Promiscuous in host mode
- PVLAN Promiscuous in trunk mode

- *feature private-vlan* *# enable feature private-vlan*
- *vlan <vlan-id>* *# configure primary vlan*
  - *private-vlan primary*
- *vlan <vlan-id>*
  - *private-vlan <isolated/community>* *# configure secondary vlan*
  - *private-vlan association <vlan-id>* *# configure association with primary vlan*
- *interface port-channel <port-channel* *# configure port-channel*
  - *switchport*
  - *switchport mode private-vlan trunk secondary #PLAN trunk mode*
  - *switchport private-vlan trunk allowed vlan 1 # configure native vlan*
  - *switchport private-vlan association trunk <primary> <secondary>*

- *interface port-channel <>*
  switchport
  switchport mode private-vlan promiscuous # PVLAN Promiscuous
  switchport private-vlan mapping 1201 1211-1213 # PVLAN mapping
  vpc 71 # assign VPC

- *interface port-channel <>*
  switchport
  switchport mode private-vlan host # PVLAN host mode
  switchport private-vlan host-association 1201 1213 # PVLAN Association
  vpc 81

4.3.2.1 DC1 Topology



4.4 N7700 Enterprise

4.4.1 Network Logical Topology Design Overview

The topologies and test cases validate high-available data center networks in order to provide unified fabric and computing services. This is achieved by using the Nexus 7010, Nexus 7700 with features such as vPC scale, VRF, PVLAN, ACL, NetFlow, WCCP, multicast, dual vPC, etc.

Nexus 7710 installed with F2E and F3 line cards to provide legacy L2 & L3 and vPC leg port channel connectivity with peer devices.

Access layer switches in this setup extended to IXIA (Traffic Generator) ports to simulate end hosts/servers to send and receive unicast & multicast traffic

The data center site is built around the Nexus 7010 with SUP2E at core & Nexus 7710 with SUP2E at aggregation.

ENT1 Aggregation:
Nexus 7710 with 120 VPC PO to N6004
Nexus 7710 with Dual VPC PO to N6001
Nexus 7710 with VRFs on F2 physical ports and 6 ECMP to N3k
Nexus 7710 with VRFs on F3 Port-channels /sub-interface to N3k
Nexus 7710 with PVLAN on F3 and F2 to N3K

ENT1

Device	Platform	Position	Status
Ent-1 (VDC)	N7000	Backbone	L3
Ent-3	N7000	Core	L3
Ent-4	N7000	Core	L3
Ent-5	N7700	Aggregation	VPC
Ent-6	N7700	Aggregation	VPC
VPC-sim1	N6004	VPC simulator	Fanout
VPC-sim2	N6004	VPC simulator	Fanout
VPC-sim3	N3548	Access	No-VPC / Regular L2
Ent-701	N6001	Access	Dual sided VPC
Ent-702	N6001	Access	Dual sided VPC
Ent-703	N3048	Access	L3 with VRF
Ent-704	N3048	Access	L3 with VRF
Ent-705	N3048	Access	L3 with VRF
Ent-706	N3048	Access	L3 with VRF
Ent-FO1	N3048	Fanout	Fanout / Regular L2
Ent-FO2	Cat6k	Fanout	Fanout/ Regular L2
Ent-FO3	Cat6k	Fanout	Fanout/ Regular L2

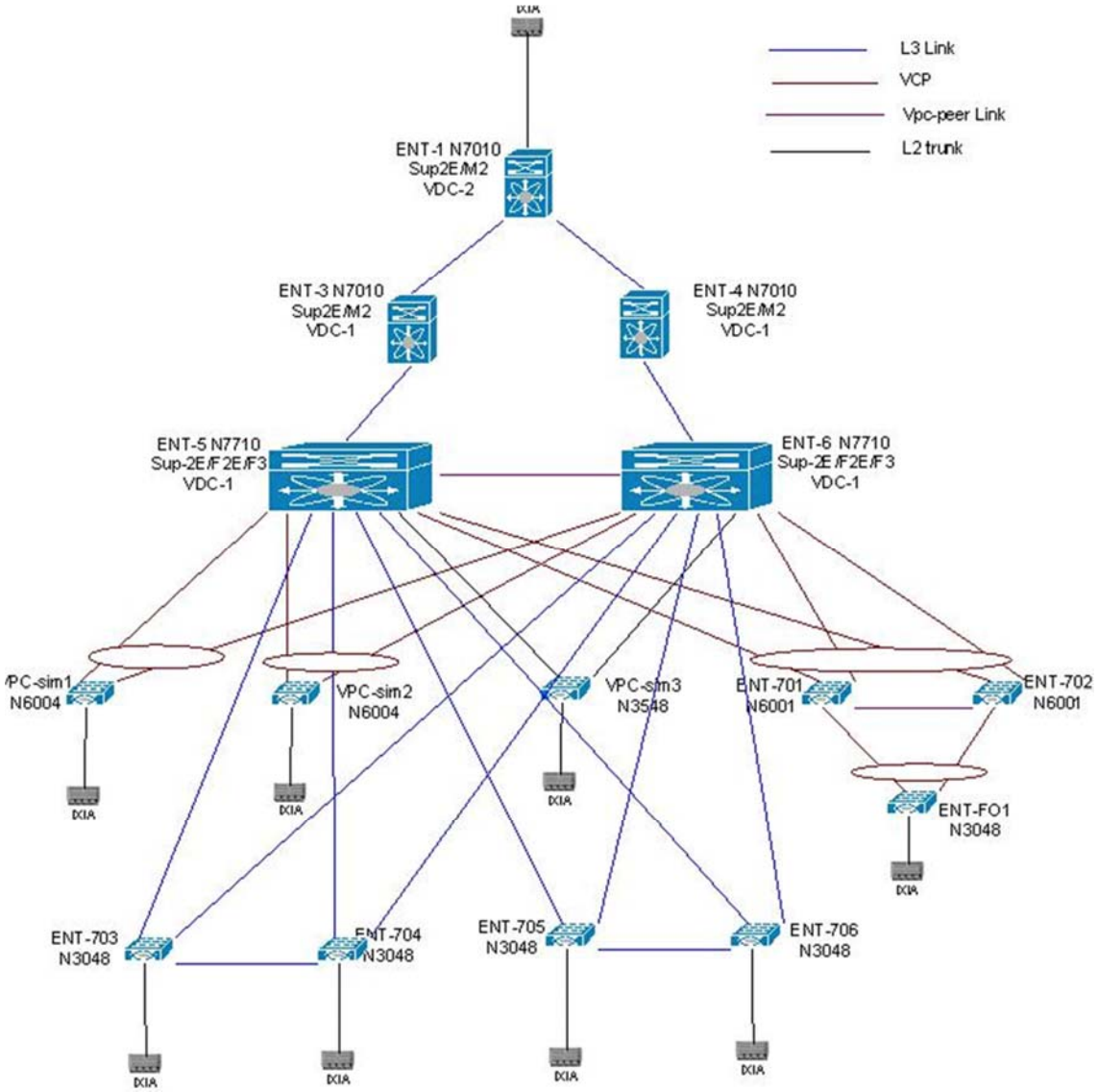


Figure 4: Ent Topology

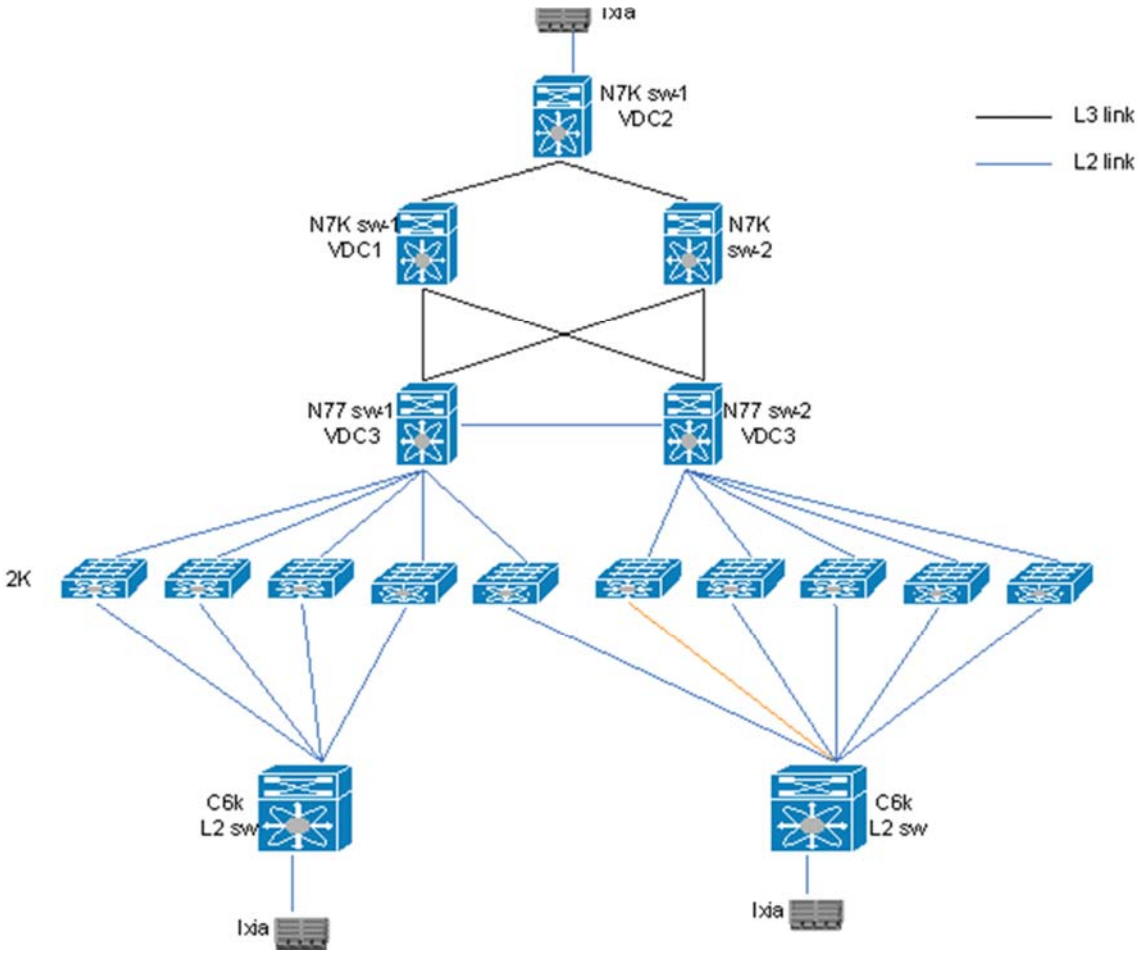


Figure 5: Alternate ENT topology for CFD validation

4.4.2 Configuration Details

The following configurations are applied to the test network:

- Common system control, management and accounting: Common system features like SSH, TACACS+, Syslog, SNMP, NTP, SPAN, DNS and Management VRF are configured
- BGP: eBGP is configured between the core switches and the public cloud.
- PIM-SM: PIM Sparse Mode/PIM Any Source Multicast is deployed across the network to support multicast. Each aggregation-access block is configured with the RP for the locally sourced groups.
- MSDP Anycast RP: MSDP is deployed to exchange source information between Anycast RPs.
- vPC: vPC technology is deployed in the aggregation-access as shown in Figure 1. In addition, dual-sided vPC is configured between the Nexus 7000 and Nexus 6000 switches

- STP: Rapid Spanning Tree Protocol is used to prevent Layer 2 loops in the aggregation-access blocks. The spanning tree root is placed on the aggregation level. Root Guard is configured on the aggregation level to enforce root placement. BPDU Filter, BPDU Guard and Port-Fast Edge are configured on the access ports towards hosts.

- HSRP: HSRP v2 is used as the first hop gateway protocol for hosts; HSRP is configured at N7K aggregation.

- IGMP: IGMP is used by hosts to join multicast groups of interest. IGMP snooping is enabled on all switches in the aggregation-access blocks to prevent flooding of multicast data traffic.

- LACP: LACP is used for link aggregation to form port-channels across the network.

- Separate VDC is created in the Aggregation N7700 switch to validate some CFD bugs for Etisalat. Features enabled in this VDC include legacy L2, HSRP, IPv4/Ipv6 dual stack, FEX HIFs configured as L2 trunk/access port-channel/phy, N-S, E-W ipv4/ipv6 traffic. CFDs verified in 6.2.14 for Etisalat customer are CSCus34965 (STP Crash), CSCum33627(vrrp leaks) & CSCus69185(AAA process crashing)

The following configurations are applied to the test network:

- Common system control, management and accounting: Common system features like SSH, TACACS+, Syslog, SNMP, NTP, SPAN, DNS and Management VRF are configured.
    - feature tacacs+ *# enabling the tacacs feature*
    - tacacs server host <ip address> key <0/7> *# configure the tacacs server to authenticate users*
    - aaa group server tacacs+ <group name> *# enable server groups for redundancy*
        - server <ip address>
        - use-vrf <vrf_name> *# use-vrf based on server reachability*
    - snmp-server user <user-name> <group-name> auth md5 <pass-phrase> priv <pass-phrase> localizedkey *# snmp v3 user with authentication enabled*
    - ntp server <ip address> *# enable ntp with server ip address*
    - ip domain-name <domain name> *# enable domain-name*
    - interface mgmt0 *# configure mgmt0*
        - vrf member management
        - ip address <ip_address >

- BGP: eBGP is configured between the core switches and the public cloud.
    - feature bgp *# enable bgp*
    - router bgp <autonomous-id> *# bgp autonomous -id*
    - router-id <router-id>
    - graceful-restart stalepath-time <120>
    - log-neighbor-changes
    - address-family ipv4 unicast
    - redistribute direct route-map <acl-name> *# route-map used for redistribution directly connected subnets*
    - redistribute ospf 1 route-map <acl-name> *# route-map used for redistribution OSPF routes*

- maximum-paths <8>
- maximum-paths ibgp <8>
- neighbor <neighbor ip addres> remote-as 100090 *# BGP peer*
- address-family ipv4 unicast
- prefix-list NO_SELF in *# acl configured to restrict prefix import*

- OSPF: OSPF is the IGP running across the network. Each aggregation-access block is configured as a unique area with the core switches playing the role of the ABR.
  - feature ospf *# enable ospf for IPv4*
  - feature ospfv3 *# enable ospf for IPv6*
  - router ospf <instance-tag>
  - router-id <ip address>
  - redistribute bgp <as_no> route-map <acl-name> *# route-map used for redistribution for bgp routes*
  - log-adjacency-changes
  - timers throttle spf 100 200 500
  - timers throttle lsa 50 100 300
  - auto-cost reference-bandwidth 1000000
  - default-metric <1>

- PIM-SM: PIM Sparse Mode/PIM Any Source Multicast is deployed across the network to support multicast. Each aggregation-access block is configured with the RP for the locally sourced groups.
  - feature pim *# enable pim*
  - ip pim rp-address <rp-address> group-list <multicast-groups> *# configure static RP for a multicast group range*
  - ip pim send-rp-announce loopback2 prefix-list <multicast-groups> *# configure candidate auto-rp*
  - ip pim send-rp-discovery loopback2 *# configure auto-rp mapping-agent*
  - ip pim ssm range <> *# configure pim ssm for default range*
  - ip pim auto-rp forward listen *# enable auto-rp messages forwarding*

- MSDP Anycast RP: MSDP is deployed to exchange source information between Anycast RPs.
  - feature msdp *# enable msdp*
  - ip msdp originator-id <interface> *# configure source interface for msdp peering, generally loopback interface*
  - ip msdp peer <ip address> connect-source <interface> *# configure peer address*

- vPC: vPC technology is deployed in the aggregation-access block DC2-Dist-N7k-201. In addition, dual-sided vPC is configured between the Nexus 7000 and Nexus 5000 switches.
  - feature vpc *# enable vpc*
  - vpc domain <domain-id> *# configure vpc domain-id*
    - peer-switch *# enable peer-switch for faster STP convergence*
    - role priority 200 *# configure priority*

        - peer-keepalive destination <ip address> source <ip address> vrf vpc-keepalive *# configure keep-alive link*
        - peer-gateway *# enable peer-gateway to avoid vPC loop*
        - track <id> *# track the L3 core connectivity to avoid black-hole*
        - ip arp synchronize *# configure arp synchronize for faster convergence of address tables*

- STP: Rapid Spanning Tree Protocol is used to prevent Layer 2 loops in the aggregation-access block DC-Dist-N7K-201. MSTP is enabled on DC-Dist-N7K-202 for the same purpose wherever applicable. The spanning tree root is placed on the aggregation level. BPDU Filter and PortFast Edge are configured on the access ports towards hosts.
    - interface port-channel <po> *# configure port-channel*
        - switchport
        - switchport access vlan <vlan>
        - spanning-tree port type edge *# enable host*
        - spanning-tree bpdufilter enable *# configure bpdufilter*

- SNMP: SNMP traps are enabled and SNMP scripts are used to collect system information and to monitor potential memory leaks.
- HSRP: HSRP is used as the first hop gateway protocol for hosts.
    - interface Vlan<id> *# configure svi*
    - ip access-group <acl> in *# enable access-list*
    - ip access-group <acl> out
    - no ip redirects
    - ip address <ip address>
    - hsrp version 2
    - hsrp 1
        - authentication md5 key-string cisco *# enable authentication*
        - preempt delay minimum 200
        - priority <priority>
        - ip <ip address> *# HSRP IP address*

- FEX: Multiple types of Fabric Extenders are deployed on Nexus 5000 parent switches.
- IGMP: IGMP is used by hosts to join multicast groups of interest. IGMP snooping is enabled on all switches in the aggregation-access blocks to prevent flooding of multicast data traffic.
    - ip igmp snooping *# by default enabled on Nexus*

- LACP: LACP is used for link aggregation to form port-channels across the network.
    - feature lacp *# enable LACP, by default LACP is used on all port-channel*

- UDLD: UDLD aggressive mode is configured across the network to detect and prevent unidirectional links
    - feature udld *# enable feature udld*
    - udld aggressive *# udld aggressive mode is enabled to re-establish the connection with the neighbor*

- Route MAP for Inter-VRF PBR

- *feature pbr* *# enable feature pbr*
  route-map GLOBAL-to-VRF permit 10 *# define route-map*
  match ip address PBR-GLOBAL-VRF *# match ACL for the route-map*
  set vrf A *# define VRF*

  ip access-list PBR-VRF-GLOBAL *# define ACL for the route-map*
  10 permit ip 151.15.1.0/24 any

  interface Vlan1501 *# Create SVI interface*
  vrf member A
  no ip redirects
  ip address 151.15.1.152/24
  no ipv6 redirects
  ip router ospf 1 area 0.0.0.151
  ip pim sparse-mode
  ip pim dr-priority 100
  ip policy route-map VRF-to-global

hsrp version 2
hsrp 1501
authentication text eCATS
priority 100 forwarding-threshold lower 1 upper 100
timers 1 3
ip 151.15.1.1
ip dhcp relay address 172.28.92.48
ip dhcp relay address 172.28.92.49
no shutdown
mtu 9000

5 NVT Findings/Conclusions/Recommendations

CSCuv45214:
Symptom： In a N9000 system running 7.0(3)I2(1), the failure of the DF might lead to packet duplication up to 25 secs
Conditions: In a typical STP access network with the STP Root located at the N9000 ToR configured also as PIM DF, the failure of the DF itself might lead to packet duplication received by all the receivers attached to any switch in the same STP domain.
Workaround: None, after about 30 secs the traffic properly reconverges to the expected receiving rate.
Severity: Moderate
Status: Assigned
Platform Seen: Nexus 9396
Resolved Releases:
Applicable Releases: 7.0(3)I2(1)

CSCuv72351:
Symptom: Multicast traffic is completely black-holed upon restarting the PIM process.
Conditions: in a N7000 system running 7.2(0) with the RP dynamically advertised (with either Auto-RP or BSR), multicast traffic might be completely dropped up to a minute upon restarting the PIM process. This operation should be hitless regardless the type of RP deployment.
Workaround: Deploy Static RP to minimize the multicast traffic drop (only few secs).
Severity: Severe
Status: Assigned
Platform Seen: Nexus 7000

Resolved Releases:
Applicable Releases: 7.2(0)

CSCuv82339:

Symptom: Sometimes, the FHR fails to properly program the IIF interface in the OIF list causing multicast bidir traffic black-holing.
Conditions: Some network instabilities and/or RP failures might lead the FHR to wrongly program the OIF list.
Workaround: reload the FHR.
Severity: Severe
Status: Assigned
Platform Seen: Nexus 7000
Resolved Releases:
Applicable Releases: 7.2(0)

CSCuv82614:

Symptom: Sometimes, multicast bidir traffic might get duplicated for up to 50 secs.
Conditions: Migrating from BSR to Static RP might cause multicast traffic to loop for 50 secs between the RP and the DF.
Workaround: None
Severity: Moderate
Status: Closed
Platform Seen: Nexus 9000
Resolved Releases:
Applicable Releases: 7.0(3)I2(1)