Nexus Validation Test Phase 4.2

Table of Contents

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 4.2 network topologies, hardware and software configurations, test procedures and findings.

NVT Phase 4.2 testing is performed on the following networks:

- **VxLAN Solution**: This test profile is developed based on BGP EVPN technology using VxLAN fabric. It is intend to provide host mobility and subnet extension across data centers and simplify network operation using anycast gateway and does not require a first hop redundancy protocol.
- **Service Provider MPLS**: The topology validates the inter DC inter site design with Carriers supporting Carriers CSC backbone for MPLS L3VPN. Two data center networks are simulated. N7710 with F3 card is used in one data center and N7010 with F3 is used in the second network. Testing is focused on the network with N7710. N7710 and N7010 with F3 card are used as aggregation and core switches. ASR9K is used as CSC PE. N5K is used for layer2 access as well as for vPC. ASR1K is used as secondary RRs for redundancy purpose in the network. One POD on DC1 has vPC to MPLS handoff, while the second POD on DC1 does Fabricpath to MPLS handoff.
- **Dynamic Fabric Automation (Vinci) Profile**: DFA profile uses power on auto-provisioning to configure network devices and support distributed gateway function and automated data center interconnect. The topology contains N7000/N7700/N6004 spines, N7000/N6001 leafs, and N7004 border leafs. The topology uses the latest DCNM for cable management and workflow automation. It also uses emulated leafs (Virtual ToR or VToRs) to achieve scale. It supports unicast and multicast for IPv4 and IPv6 hosts and auto configuration based on dynamic frame snooping based on the traffic and VDP running on the vSwitch.
- **ENT**: This test profile primarily is built to validate Enterprise customer profiles. In the first phase, we have validated the TIER I Enterprises customer profile. The test bed is built with a new hardware covering the existing feature set for future deployments. Nexus7000 SUP2E/M2 is used at the core layer and N7700 SUP2E/F2E/F3(40G) at the aggregation layer. The profile also covers interoperability with Nexus6000 and Nexus3000 switches and covers L3Agg with L3 ToR and L3Agg with L2 ToR.
- **MSDC**: This profile focuses on scale requirements of Massively Scalable Data Centers. 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
- **Hosted DC M2 BGP Scale**: In this profile, a pair of N7K ASBRs that peers with provider edge routers and receives Internet feed. N7K also peers with customer access routers in the same AS and redistribute routes. The firewall clusters are connected to the N7Ks through N2K FEXes.
- **M1 VPC**: This test bed focuses on scaling the virtual Port-Channel (vPC] with Nexus 7000. It also covers interoperability with Catalyst 6500. This network uses vPC and PVLAN to deliver high availability to servers connecting to data centers.

Operation: Network management including SNMP poling and inventory collection is performed through 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 by Nexus access switches 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 Validated Scale

2.1 VxLAN Solution

Scale

Feature	Scale
Maximum VLANs	**40**
VLANs per TOR	**14**
VRFs per ToR	**10**
Max VNI per ToR	**24**
Max VNI per network	**10 VRFs + 30 VLANs = 40**
Max underlay Multicast Groups	**10 (to support broadcast + multicast traffic**
Max ToR per 'network'	**116 (112 ToRs per POD and 4 on peer POD)**
Max VTEP peers per 'network'	**116 (112 ToRs per POD and 4 on peer POD)**
Overlay MAC scale	**16K per POD * 2 = 32 K per 'network'**
Overlay IPv4 routes	**16K per POD * 2 = 32 K per 'network'**
LISP dynamic routes	**6600**
Undelay v4 routes	**400 across all PODs**
Undelay v6 routes	**none**

Hardware

	Model No.	NVT 4.2
N77K	N77-SUP2E / N77-F324FQ-25	7.2.1
N9K	N9K-C9396PX / N9K-M12PQ	7.0.3
CSR1000v	Na	15.4(3)S1
ASR1K	ASR1000-RP2/ ASR1000-SIP10/ SPA-10X1GE-V2	155-1.S1
ASR9K	A9K-RSP440-TR/ A9K-40GE-E/ A9K-2x100GE-SE	5.3.0

2.2 Service Provider MPLS F3

Scale

Feature	Scale

Feature	Scale
6VPE VRFs (dual stack)	100
ARP addresses on each switch	21,120
Global Multicast routes	100
HSRP groups	1,000
IGMP groups	100
L3 Physical interfaces	300
L3VPN v4 VRFs	100
MVPN VRFs	10
Number of IPv6 Host	2,000
L3 Port channel	2
SVI	1,000
Total Multicast prefixes in MVPN	100
Total number of IPv4 prefixes in default routing table with 2 IGP Path	200
Total VPNv4 Prefixes from IBGP peers with 2 IGP Path	8,000
Total VPNv6 Prefixes from IBGP peers with 2 IGP Path	2,000
vPC links	8
VPNv4 Prefixes across 6 VRFs with one IGP Path	6,000
VPNv6 Prefixes across 6 VRFs with one IGP Path	2,000
Fabricpath links per spine	5

Hardware

	Model No.	Image Version
N77K	**N77-SUP2E/ N77-F324FQ-25**	7.2.1
N7K	**N7K-SUP2E/ N7K-F312FQ-25**	7.2.1
N5K	**N5K-C5548UP-SUP**	7.0(5)N1(1)
ASR9K	**ASR9001-RP/ A9K-MPA-20X1GE**	5.1.1
ASR1K	**ASR1000-RP2/ ASR1000-**	3.14

	ESP20/ SPA-8X1GE-V2	

2.3 Dynamic Fabric Automation (Vinci) Solution

Scale

Spine/SuperSpine	
VRFs (vni)	Transit-Mode
SVI/BDI	Transit-Mode
VLAN/BD	Transit-Mode
BGP Route-Reflector Peer	180
Fabric Core Ports (Ports to leaves)	192
Fabric BFD Sessions	12
Fabric BGP Sessions	192
FabricPath Layer-2 BFD	12

Leaf	
Fabric Facing	
Fabric Core Ports	16
Fabric BFD Sessions	16
Fabric BGP Sessions	1 RR
VRFs (vni)	410
SVI/BDI (Core)	410
VLAN/BD (Core)	410
Host Facing	
VLAN (vn-segment)	1640
SVI/BDI	1640
VPC+/VPC	96
FEX	NA
Physical/Virtual Server IP Addresses (ARP/ND) – 1:6 IPv4 to IPv6 ratio	
IPv4 Routes (/32 + Subnet Route)	6k
IPv6 Routes (/128 + Subnet Route)	4k
Multicast Routes (S,G)	2k
Total VNI/vn-segment per Leaf	2050

BORDER LEAF	
Fabric Facing	
Fabric Core Ports	16
Fabric BFD Sessions	16
Fabric iBGP Sessions	1
VRFs (vni)	1000
SVI/BDI (Core)	1000
VLAN/BD (Core)	1000
Edge-Router Facing	
VLAN (vn-segment)	1000
Sub-Interface per BorderLeaf (if VRF-lite)	1000

IPv4 Routes (/32 + Subnet Route)	17K*
IPv6 Routes (/128 + Subnet Route)	5K*
Multicast Routes (S,G)	1K*
eBGP/MP-BGP peers	1000
VPC+/VPC	0
Labels/VRF	Option B
Total VNI/vn-segment per Border Leaf	1000

Hardware

	Model No.	Image Version
N77K	**N7F-SUP2E / N77-F324FQ-25 / N77-F248XP-23E / N7F-F248XP-23E**	7.2(0)D1(1)
N7K	**N7K-SUP2E / N7K-F312FQ-25 / N7K-F248XP-25E / N7K-F248XP-25**	7.2(0)D1(1)
N6K	**N6K-C6001-64P-SUP / N6K-C6001-M4Q**	7.1(1)N1(1)
N6K	**N6K-C6004-96Q-SUP / N6K-FIXED-LEM / N6K-C6004-M12Q**	7.1(0)N1(1)
N6K	**N6K-C6004-96Q-SUP / N6K-FIXED-LEM**	7.0(2)N1(1)
Cisco Prime	**Data Center Network Manager**	7.2(0.14)

2.4 Enterprise 1

Scale

Feature	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

Hardware

Platform	Model No.	Software
N7000	SUP2E, M2	7.2.1
N7700	SUP2E, F2E, F3(40G with breakout)	7.2.1
N6000	N6K-C6004-96Q-SUP , N6K-C6001-64P-SUP	7.1.0.N1
N3000	N3K-C3548P-10G-SUP, N3K-C3048TP-1GE-SUP	6.0.2

2.5 MSDC Scale Profiles

Scale

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

2.6 Hosted DC M2 BGP Scale

Scale

Feature	Scale
VLAN / SVI	400
Port Channel Members	8
HSRP	200
VRRP	200

BGP Neighbor	100 eBGP [on Individual links] 50 iBGP
BGP Peer Groups	25
OSPF Neighbor	20
Unicast Route	655K
LACP	9
vPC Config-Sync	120

Hardware

Equipment Model	**Quantity**
C7009 Chassis	2
N7K-M224XP-23L	18
Catalyst 6504	2
ASR9K	2
N3548	7
N3048	4

2.7 M1 vPC

Scale

Feature Max	M1 vPC
Primary VLAN	140
Secondary VLAN	280
Phy. Ports used for PVLAN	468
Port-channel	388
Port-channel in PVLAN	388
VPC PO with PVLAN	256
Host mode	32
Promiscuous Access	64
Promiscuous Trunk	186
Trunk Secondary	186

3 ISSU Matrix

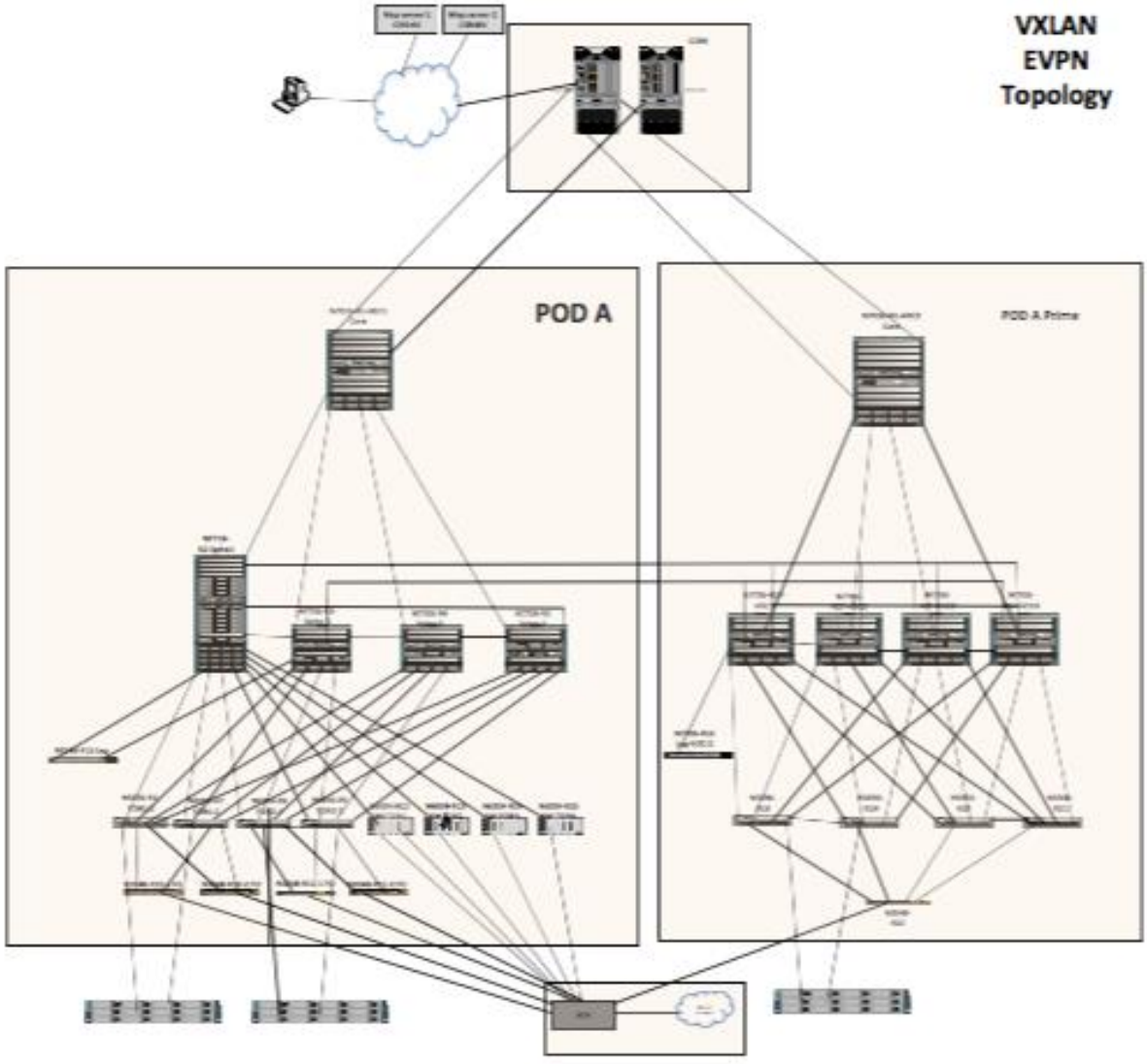
Image	ISSU/COLD BOOT	MPLS Profile 1	ENT	MSDC	M2 BGP Scale	M1v PC
7.2.0 > 7.2.1	ISSU	PASS	PASS	PASS	PASS	
7.2.0 + SMU > 7.2.1	ISSU		PASS	PASS		
7.2.1 > 7.2.1 upg	ISSU	PASS	PASS	PASS		PASS
6.2.10 > 7.2.1	ISSU		PASS		PASS	
6.2.12 > 7.2.1	ISSU		PASS			
7.2.0 > 7.2.1	COLD BOOT	PASS	PASS			
7.2.1 > 7.2.1 upg	COLD BOOT		PASS			

6.2.10 > 7.2.1	COLD BOOT		PASS			
6.2.12 > 7.2.1	COLD BOOT		PASS			
7.2.1 > 6.2.10	COLD BOOT		PASS			
7.2.1 > 6.2.12	COLD BOOT		PASS			

4 Profile Details

4.1 VxLAN Solution

Topology



4.2 Hardware and Software Overview

	Model No.	NVT 4.2
N77K	N77-SUP2E / N77-F324FQ-25	7.2.1
N9K	N9K-C9396PX / N9K-M12PQ	7.0.3
CSR1000v	Na	15.4(3)S1
ASR1K	ASR1000-RP2/	155-1.S1

	ASR1000-SIP10/ SPA-10X1GE-V2	
ASR9K	A9K-RSP440-TR/ A9K-40GE-E/ A9K-2x100GE-SE	5.3.0

4.2.1 VxLAN Solution Design Review

The DC design is structured in pairs of PODs between which mobility and segmentation should be supported. The choice is to use VXLAN with an EVPN control plane within each POD and for east-west connectivity across peer PODs; LISP is required for optimization of North-South traffic. There are 4 aggregation boxes per POD that use SUP2E and F3 line cards and 40GE interfaces. Each aggregation box is connected to all ToRs. The 4 aggregation boxes are also connected to each other in a ring. Each Rack has two ToRs in a vPC array. vPC runs from the physical hosts to the ToRs.

Communication between PODs in the same DC happens over the core. North-South Communication in and out of the PODs also happens over the core. Segmentation is only stretched between peer-PODs; segmentation is not stretched outside of the PODs to the Core. Only one VRF (the production VRF) is routable over the core. Thus, inter-POD and North-South communication is limited to the production VRF only.

East-West communication is optimized between pairs of PODs across DCs.
Each POD will have a peer POD at the remote DC and the aggregation switches of these peer PODs will be directly connected over fiber. The following connections will be established: POD1 to POD1'

East-West communications between Peer PODs will follow the low latency/high-capacity direct connection between Aggregation switches. Communication across PODs that are not peers will happen over the core only for the "Production VRF"

Design Choices

- Spine to Core is eBGP, peering is interface based. On Spine only Production_vrf routes are routable in core
- Core is single AS
- OSPF will be running between all Spines in the ring topology
- Spines are iBGP peering over interface to TORs. Loopback are redistributed with network statement. All four spines are IPv4 RR to all ToRs
- Each N9396 are connected to all four Spines. Connection is 40G links. Some connection uses are port-channel with 2-3 40G member link
- Legacy TORs has 10 VRFs. Correspondingly Spine will have 10 VRF-lite for Intra-POD legacy to fabric traffic
- Legacy TORs eBGP peer to Spine thru "fake iBGP" by manipulating "local AS" in as-path. So BGP peering is iBGP, but local AS are different so route get redistributed to MP-iBGP L2VPN EVPN.
- Spines are route reflector for IPv4 and L2VPN EVPN
- All leafs in the POD are EVPN RR clients to Spins. IBGP peering is with loopback
- Total 10 VRF in system, one is Production VRF which will carry hosts that will require inter-POD and N-S (i.e branch to host) access
- Other 9 VRFs will be connecting within the POD or peer-POD only. No N-S access for these VRFs. There is no route-leaking into Production VRF. These 9 VRFs require connectivity to legacy ToRs also

- VM hosts requiring mobility are separated into different VLANs (we call it mobility VLAN)
- Production VRF will have 10 Mobility Vlan with total of 6600 host and 500 Non-mobility vlans with 4 Vlan each in each TORs (125 TORs X 4 Vlans = 500)
- Each Mobility Vlan will have 3 host (10 Vlan x 3 host X 125 Racks = 3750 hosts) in POD A and 2850 hosts in POD A"
- For Non Mobility Vlans, 4 Vlan X 50 host X 125 TORs = 25000 hosts in POD A
- Inter-peer-POD is EBGP interface peering over high speed direct link. VXLAN tunnels are TOR to TOR
- For N-S for Production VRF, VXLAN tunnel will terminate on Spine and L3VRF handoff to production-vrf handoff
- IPv4 and L2VPN EVPN iBGP peering between inter-spine ring links and Set local pref low for inter spine link
- POD A has 250 emulated TORs using IXIA eVPN emulation
- Each TORs N9K have 40 host VPCs with N3K fan out switches
- 10 Legacy TORs are connected to PODA
- 2 map server (CSR1KV) connected with DDT to Branch router ASR1K.
- LISP traffic will load balance between 4 spines by the map-server
- Spine we have configured route-map to allow only non-mobility prefixes to get advertise to core and branch for south to north traffic

4.3 SP MPLS Solution

4.3.1 Network Logical Topology Design Overview

MPLS Solution: The focus of this test bed is to test the MPLS solution proposed to the data center customers in general. Scale and topology are aligned to the deployment proposed to the to a specific customer. Solution testing is done for inter DC (layer 3 DCI with MPLS) with CSC backbone. Testing is primarily focused on MPLS features configured on N77K with F3 card positioned as aggregation and core switches in the data center network. This network uses vPC towards aggregation on one POD and Fabric Path based vPC+ towards aggregation on another POD to deliver highly available unicast and multicast services and CSC is in the MPLS backbone. ASR9Ks are used as CSC PEs.
The overall goal is to identify any stability issues when configured with multiple features and measure traffic convergence metric with key triggers such as SSO, process crashes/restarts, VDC reload and OIR with multiple features configured at pre-defined scales. The results in this document are closely tied with the setup utilized for this profile. These results may vary depending upon the deployment and optimizations tuned for the convergence.

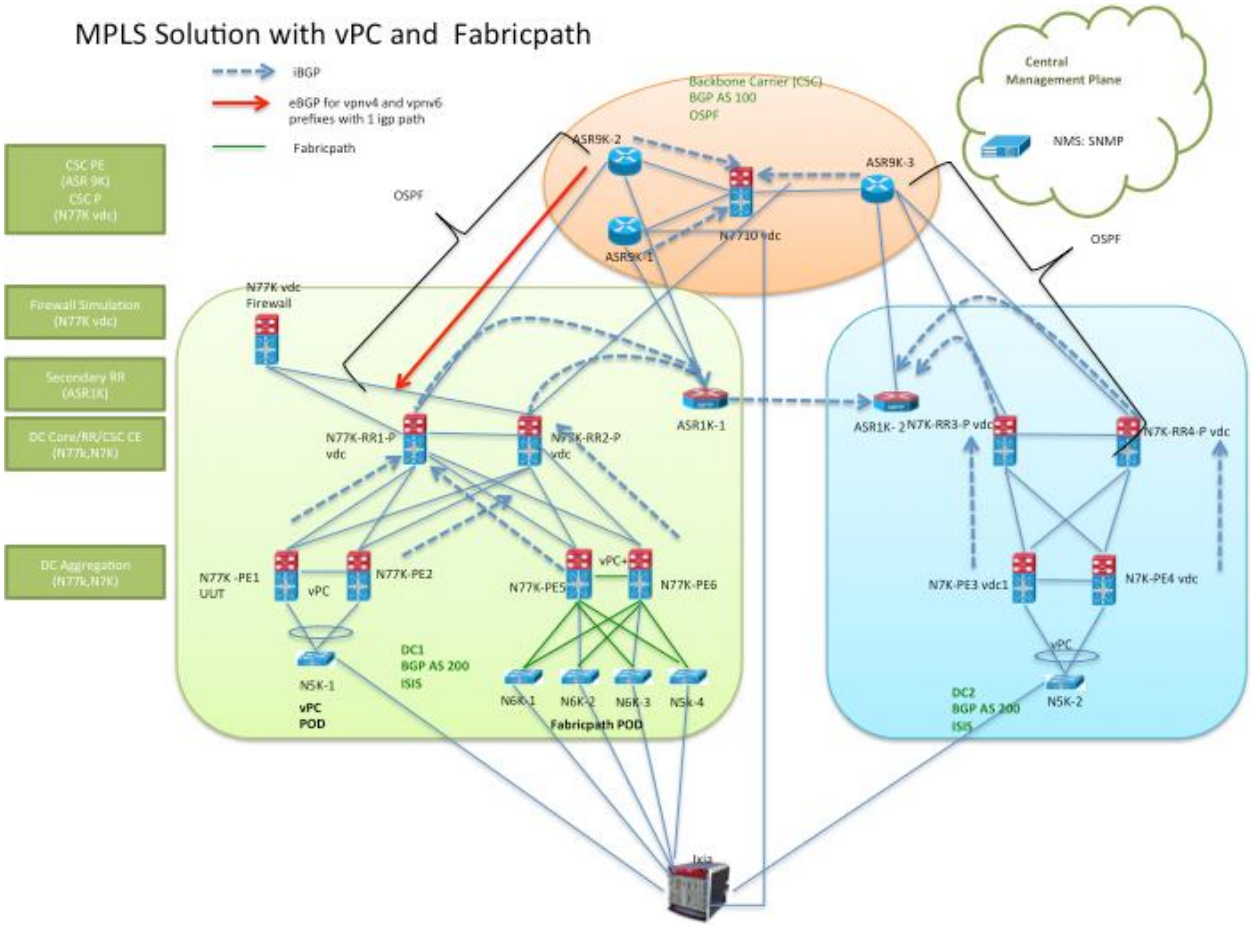
The topology validates the inter DC inter site design with CSC backbone for MPLS L3VPN. Two data center networks are simulated. N7710 with F3 card is used in one data center and N7010 with F3 is used in the second network. Testing is focused on the network with N7710.

N7710 and N7010 with F3 card are used as aggregation and core switches. ASR9K is used as CSC PE. N5K is used for layer2 access as well as for vPC. ASR1K is used as secondary RRs for redundancy purpose in the network as shown in Figure1.

4.3.2 Description of the test network

- N7710 and N7010 with F3 card are used as aggregation and core switches - Features configured are L3VPN with VRFs, ISIS, iBGP, HSRP, SVIs, vPC, LDP, 6vPE, mvpn, multicast.
- ASR9K is used as CSC PE – Features configured are OSPF, iBGP, LDP, VRF.
- N5K is used for layer2 access as well as for vPC.
- ASR1K is used as secondary RRs for redundancy purpose in the network - Features configured are OSPF, iBGP, LDP
- Ixia is used to generate host routes (/32) as well for end to end traffic validation.

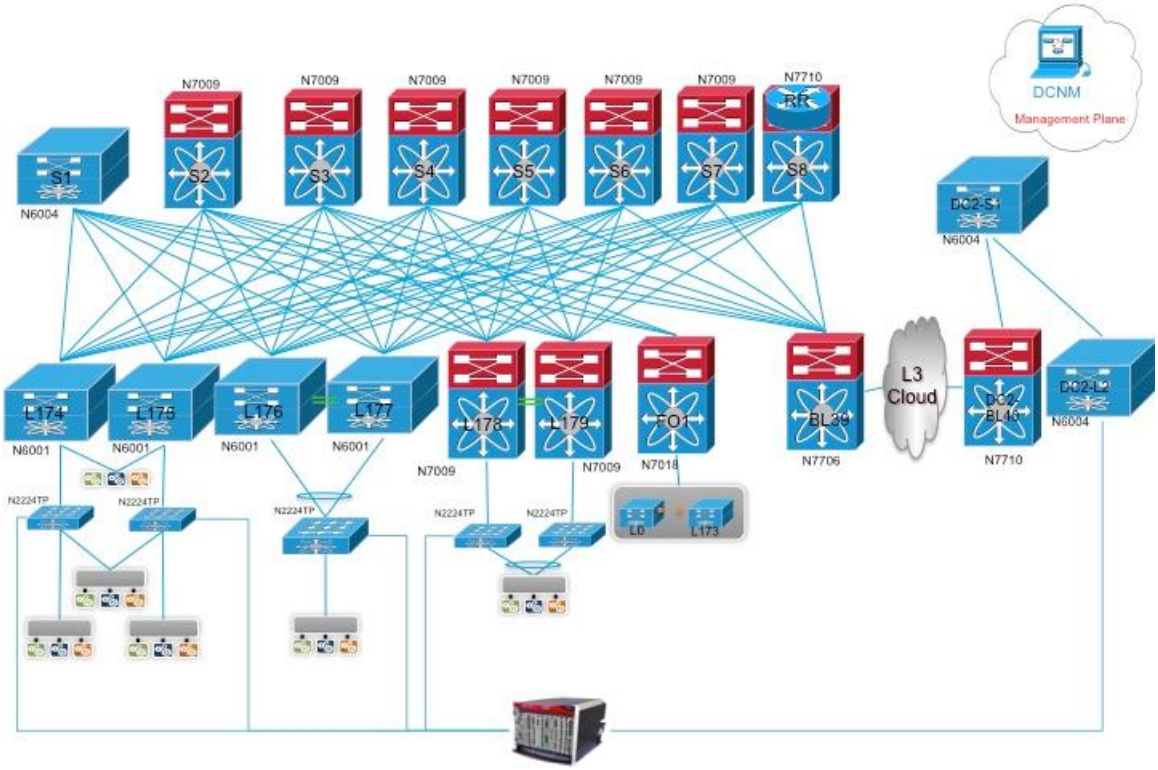
4.3.3 Topology



4.4 Dynamic Fabric Automation (Vinci)

Topology

SSTE Vinci-Fabric Encapsulation Test Topology



4.4.1 Topology Design Review

There are 9 Spines, 180 Leafs and 2 Border Leafs in this topology. Spine S1 and RemoteDC-Spine are Nexus 6004 Switches. Spine S8 is Nexus 7710 Chassis with F3 Card and Spine S2 to S7 are VDCs within Nexus 7009 Chassis with F2e Cards. Leafs L174 to L177 are Nexus 6001 Switches. Leafs L178 and Leaf 179 are Nexus 7009 Chassis with F3 Cards and Leaf 0 to Leaf 173 are emulated using Titanium+ VTORs. Border 39 and Border Leaf 40 are Nexus 7706 and Nexus 7710 Chassis respectively with F3 cards.

Fabric Connectivity

Real Leafs (L174-L179) and Border Leafs (BL38-BL40) are connected to 8 spines. Spine8 is connected to 174 Emulated Leafs (L0-L173) with the help of Fan out switch (Nexus 7018). UCSs used for emulating VTORs are connected to Fan out switch with a Trunk Link as described in Vinci Hybrid Cluster approach.

Server Leaf Connectivity

For configuring VM, we use Emulated hosts. IXIA traffic generator will be used for emulating host with VDP and ARP/ND Support connected to Real Leafs.

DCI and L3VPN Connectivity at Border Leaf

BGP based VRF Lite will be configured as border Leaf node and DCI node. BGP MPLS VPN will be configured between DCI Peers. IXIA emulation will be used to emulate remote branch routes.

4.4.2 Hardware and Software Overview

	Model No.	Image Version
N77K	**N7F-SUP2E / N77-F324FQ-25 / N77-F248XP-23E / N7F-F248XP-23E**	7.2(1)D1(1)
N7K	**N7K-SUP2E / N7K-F312FQ-25 / N7K-F248XP-25E / N7K-F248XP-25**	7.2(1)D1(1)
N6K	**N6K-C6001-64P-SUP / N6K-C6001-M4Q**	7.1(1)N1(1)
N6K	**N6K-C6004-96Q-SUP / N6K-FIXED-LEM / N6K-C6004-M12Q**	7.1(0)N1(1)
N6K	**N6K-C6004-96Q-SUP / N6K-FIXED-LEM**	7.0(2)N1(1)
Cisco Prime	**Data Center Network Manager**	7.2(0.14)

4.5 ENT

4.5.1 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 and 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 provide legacy L2 & L3 and vPC leg port channel connectivity with peer devices.

Access layer switches in this setup are 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

4.6 Configuration Details

The following configurations are applied to the test network:

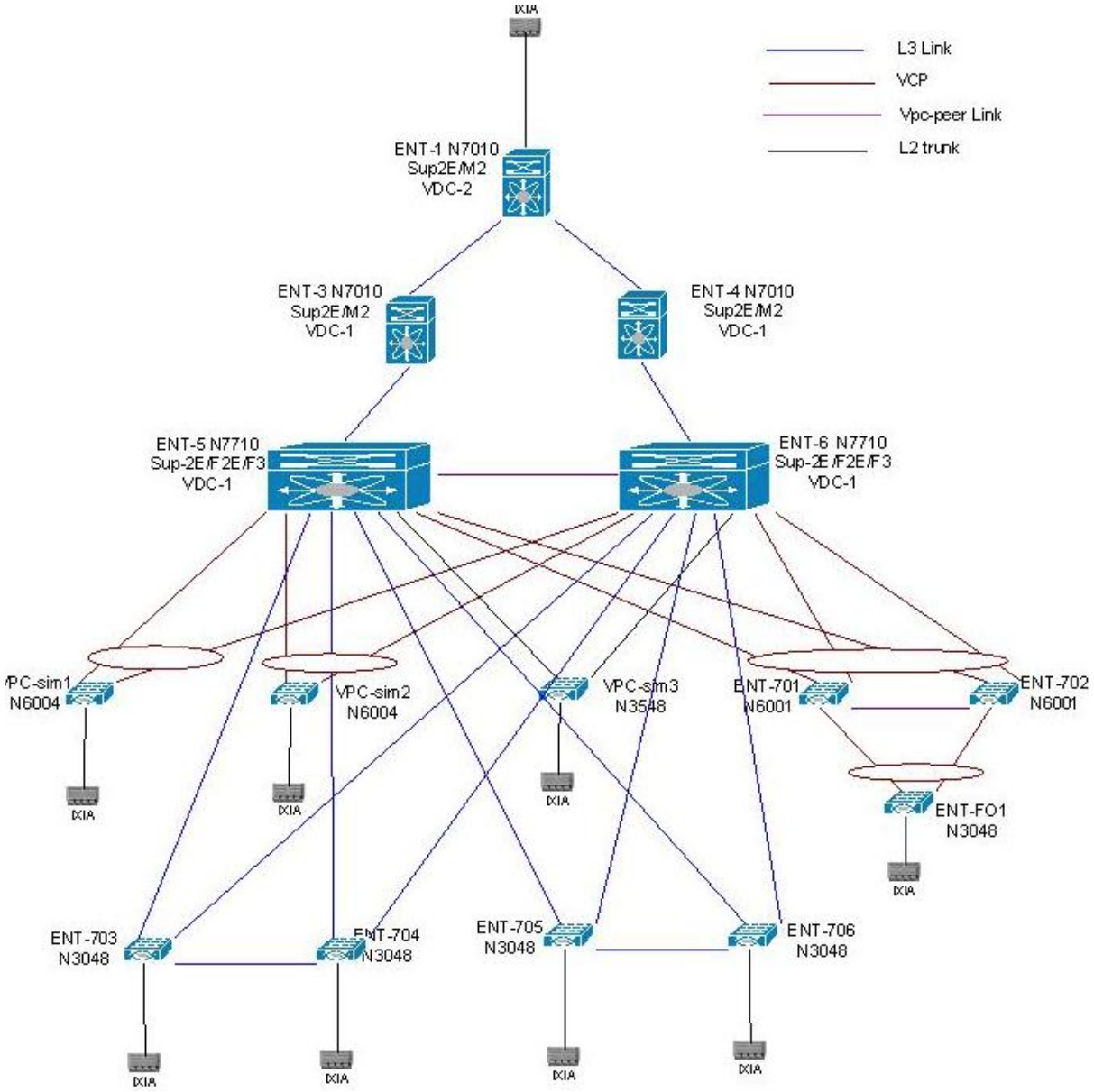
- Common system control, management and accounting: Common system features like SSH, TACACS+, Syslog, SNMP, NTP, SPAN, DNS and Management VRF are configured
- BGP: eBGP is configured between the core switches and the public cloud.
- PIM-SM: PIM Sparse Mode/PIM Any Source Multicast is deployed across 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 bellow. 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 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.

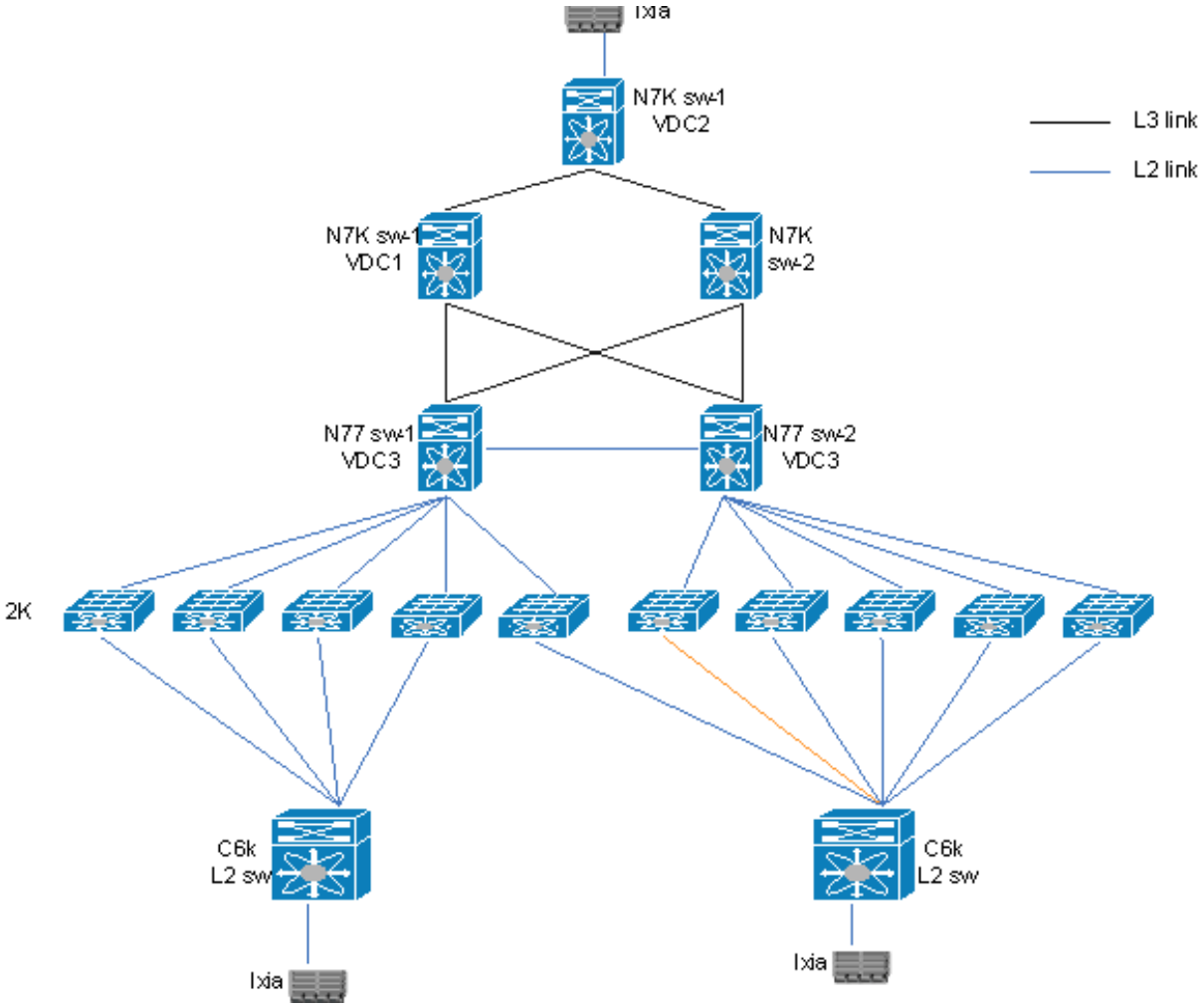
- LACP: LACP is used for link aggregation to form port-channels across the network.

- Separate VDC created in Aggregation N7700 switch to validate additional test scenarios. 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 pv4/ipv6 traffic

- FEX Active-active : Uplink ports as port-channel connected to each vpc peer in vpc mode. All FEX host ports configured as L2 trunk allowing one unique vlan. Unicast ipv4 and multicast ipv4 traffic flow in N-S direction.

4.7 ENT Topology





4.8 MSDC Scale

The topologies and test cases validate the requirement of feature and scale in a Massive Scalable Datacenter (MSDC) deployment. This is achieved by using the Nexus 7010, Nexus 3k, and Nexus 5548 with protocols such as OSPF, BGP, and BFD to create various profiles.

The following features are covered in profile 1:-

- BGP or OSPF
- BFD with default timers (50ms) or 250ms
- SNMP
- NTP
- SYSLOG

The following features are covered in profile 2-1:

- OSPF / OSPFv3
- BFD (v4 and v6) with timers of 250ms
- vPC
- Multicast with MSDP (v4)
- IGMP
- ACL
- DHCP relay

- High availability
- SNMP
- NTP
- SYSLOG

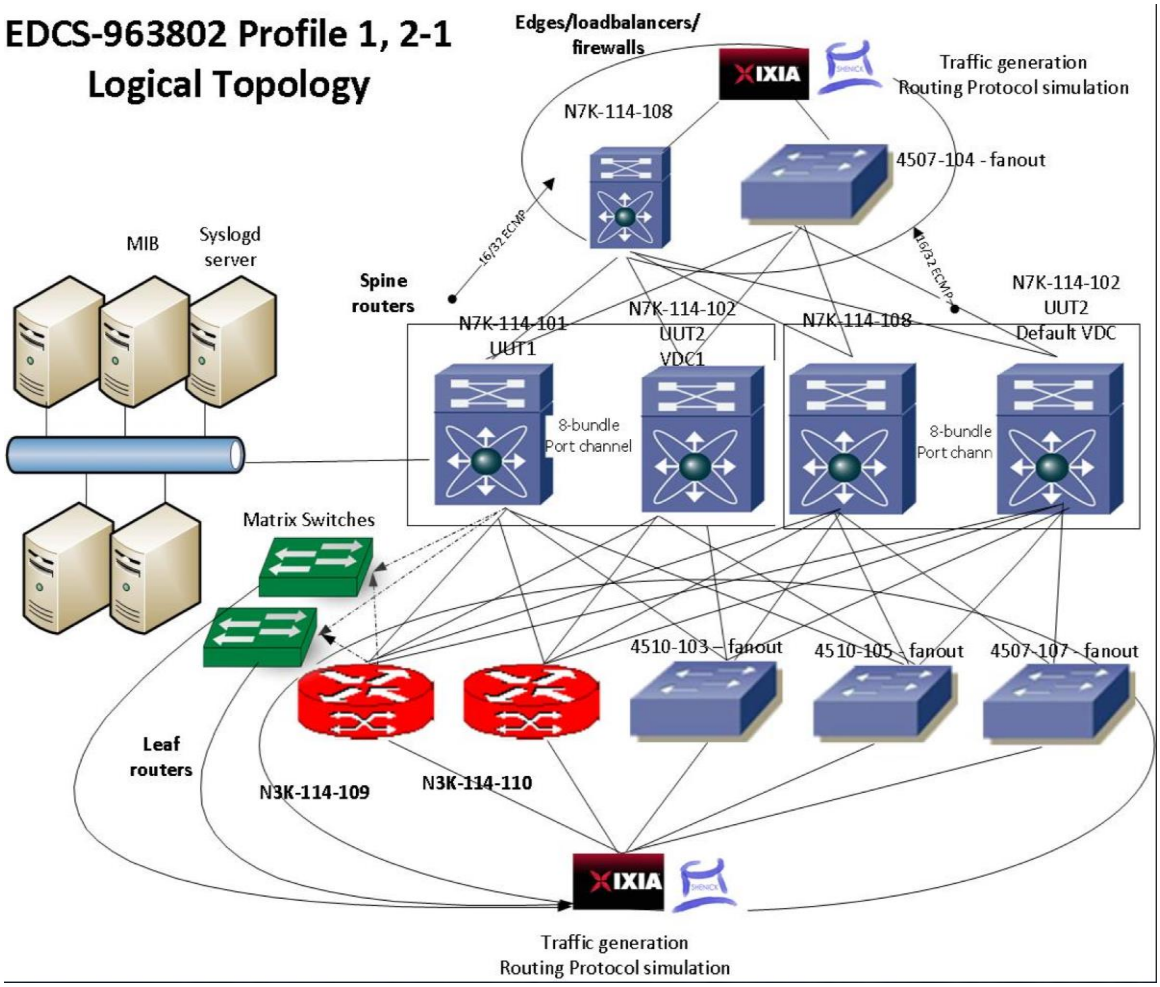
4.8.1 Topology Design Overview

MSDC covered 2 main profiles – profile 1 with BGP/OSPF and profile 2-1

1) Profile 1 – leaf node on a 7018 with 16 F2, running either BGP or OSPF as routing protocol with BFD using different bfd timers, default (50ms) or 250ms.
2) Profile 2-1 – L2/L3 on a 7018 with 2 M2, running ospf/ospfv3 as routing protocol with BFD using 250ms with vpc, multicast.

Roles and Routers/devices

Roles	Routers and Devices
Spine node	N7K-114-101(UUT1), N7K-114-102 default VDC (UUT2) and VDC1 and N7k-114-108 VDC1. 16/32 ECMP is between spine node and gateway/load balancers and firewall.
Leaf node/ToR	N3K and simulated by 4510-1-fanout/4510-2-fanout/4510-3 fanout + Ixia
Gateway, Load balancers, Firewall	Simulated by N7K-114-108 default VDC and 4507-1-fanout + Ixia



4.9 Hosted DC BGP M2 Scale

A pair of N7Ks are used as Edge Router/ASBR. They are connected to provider edge routers with eBGP peering. Within the Autonomous System, these N7K pairs peer with customer access routers. The edge routers also connect to firewall clusters. The firewalls are connected using N2K FEXes.

Base Configuration in Core

- Each core routers (ASR9K) will be connected to Edge Router DCI with 8 members (10 GB) port-channel
- IBGP on each Edge Router links and advertise internet table (655K) to SP POD Customer Access Routers

Base Configuration in SP POD

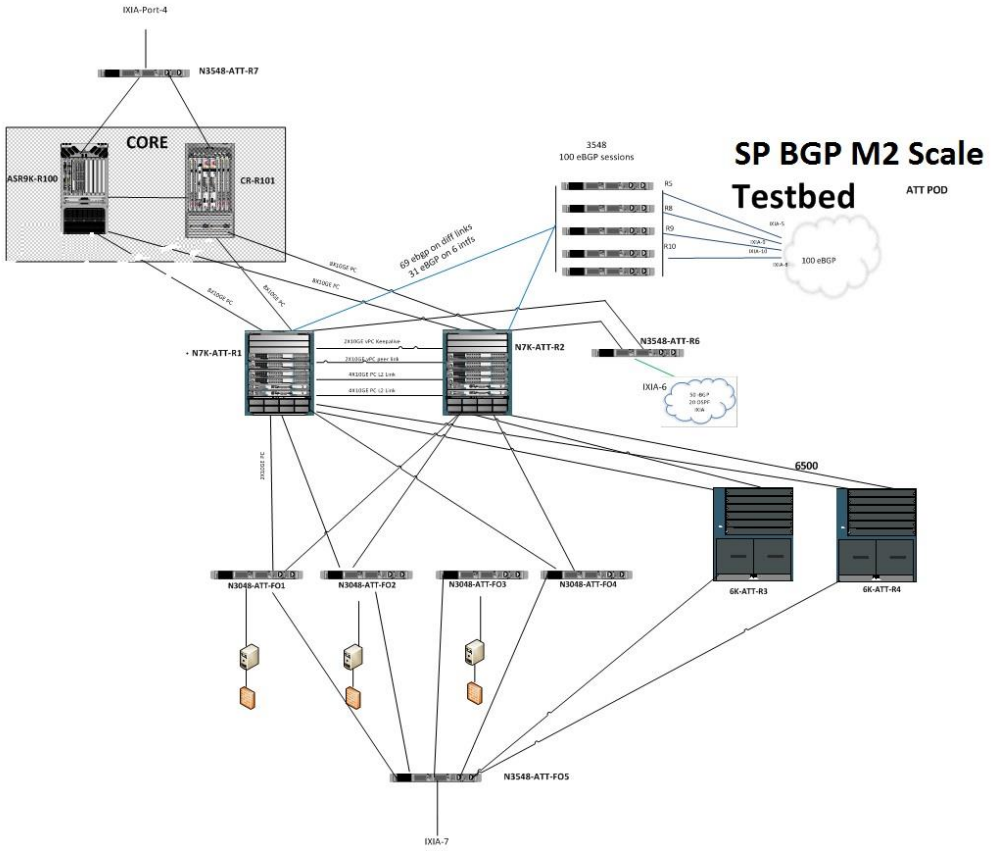
- L3 (IPv4/IPv6) Port-channel interface with 8 members 10GB links (between R1 and R2) to each core routers (ASR9K)
- 1 vPC keep alive link with 2 members links (between R1 and R2)
- 2 vPC peer links with 4 members links (between R1 and R2) one for vPC for L2 and seconds vPC for L3)
- 50 iBGP sessions from R1 and R2 to remote PE emulation to IXIA. 655K internet routing table will be advertise to all 50 iBGP links
- HSRP and VRRP for IPv4 and IPv6

- Each Edge Routers (R1 and R2) are connected to 4 L3 links to 2 6500 switch each
- 100 eBGP sessions from R5 and R6 to remote PE emulation to IXIA. 655K internet routing table will be advertise to all 100 BGP links
- 20 OSPF sessions between R1/R2 and IXIA [thru R6]

- Aggregation Layer: It's a vPC setup with HSRP on all SVIs. vPC enhancements are also covered. UDLD is enabled on all interfaces.
- Core Layer: BGP and OSPF prefixes are injected into this layer. eBGP configured towards north of the traffic generator. Between Agg and Core devices, OSPF and iBGP is configured to carry traffic
- Two sets of traffic profiles will be provisioned for testing. One for regular testing (particularly convergence testing), and the other set for background traffic.

- North-South (N-S) — indicates traffic flows between Ixia ports connected to the core routers and Ixia ports connected to the access switches.
- East-West (E-W) — indicates traffic flows (inter- and intra-VLAN traffic) between Ixia ports connected to different access switches.

Topology:



4.10 M1 vPC Scale

4.10.1 Hardware and Software Overview

Platform	Model No.	Software
N7K	**N7K SUP2E M1**	7.2.0.
C6K	**VS-SUP720-10G**	122-33.SXJ1

4.10.2 Network Logical Topology Design Overview

The topology validates high-available networks that depict the various private VLAN feature implementations in order to provide an idea of the private VLAN scale numbers supported. This is achieved by using the Nexus 7000 and Catalyst 6500 switches.

Figure 3 illustrates the network built around 2 Nexus 7000 switches with Sup2e and M1 modules. The topology contains:

- Nexus 7000 with vPC to Catalyst 6500 VSS switches for access configured.
- Nexus 7000 connected to Catalyst 6500 switch with classical Port-channels.
- Nexus 7000 connected to Catalyst 6500 switches using orphan ports.

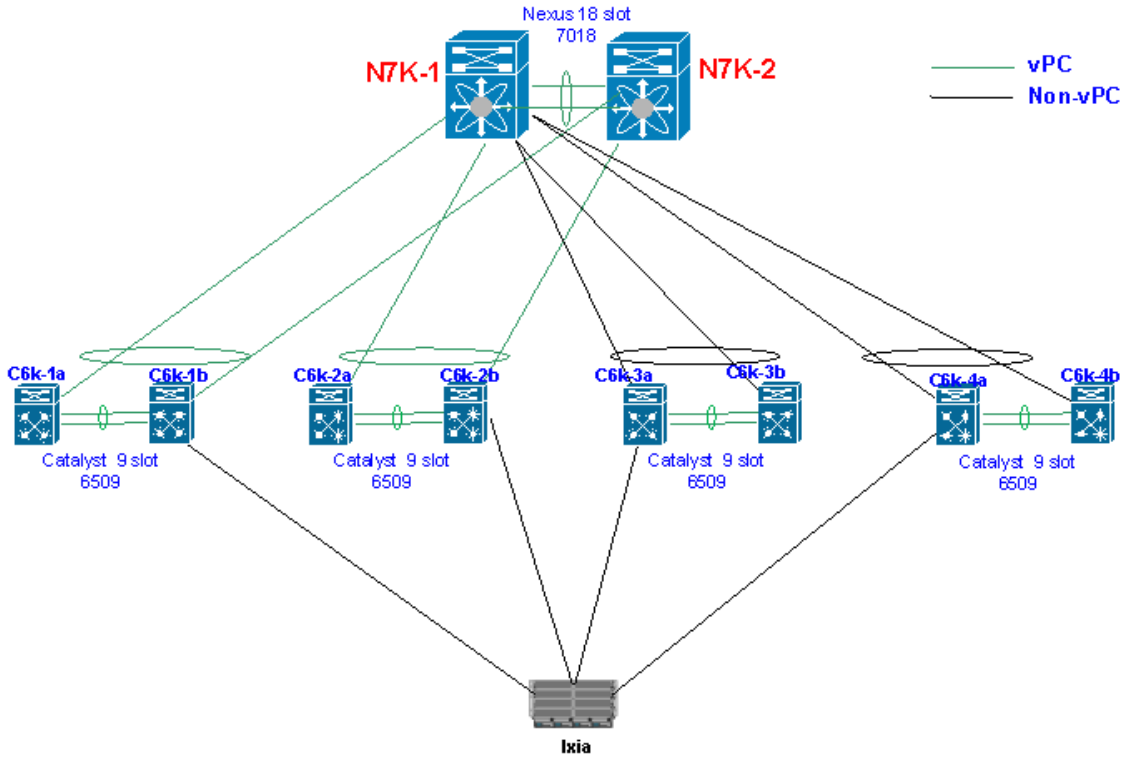
4.10.3 Configuration Details

The following configurations are applied to the test network:

- Common system control, management and accounting: Common system features like SSH, Syslog, SNMP, NTP and Management VRF are configured.
- vPC: vPC technology is deployed in the network between the N7k and the Catalyst VSS switches as shown in the figure 3.

- 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 PortFast Edge are configured on the access ports towards hosts.
- LACP: LACP is used for link aggregation to form port-channels across the network
- PVLAN: PVLAN is configured in the network and is the main focus of testing. The following PVLAN components are covered in the network:
  - PVLAN primary and secondary vlan(Community and Isolated)
  - Promiscuous trunk and secondary trunk on vpc
  - PVLAN host on classic port-channel.
  - PVLAN promiscuous trunk, secondary trunk on classic Port-channel

4.11 M1 vPC Scale Topology



4.12 DC1, ENT1 and M1 vPC Feature / Scale Coverage

Feature Max	DC1	ENT	M1 vPC
F1	2		
M1	5		
VDC	4		
FEX	6		
FEX-HIF	96		
ACL/ACE	500	500	
PO Members	2		
VLAN / SVI	100	2000	
VLAN per FEX	10		
FP VLAN	50		
MAC	7K	1K	
vPC	120		
VLAN per vPC	1,10,16		
HSRP v2	100	2000	
OSPF Neighbor	16		
BGP Neighbor	2	18	
PIM Neighbor	256	2000	
PBR Sequence	20	200	
Unicast Route	10K	22K	

Multicast Route	500	3K	
Multicast Source	20		
Multicast OIF	10		
Multicast Group	100	390	
Fabric Extender	3		
FP IS-IS adjacencies	6		
FP Number of Switch ID's	20		
Primary VLAN	20		140
Secondary VLAN	25		280
Phy. Ports used for PVLAN			468
Port-channel	125		388
Port-channel in PVLAN	18		388
VPC PO with PVLAN	10		256
Host mode	18		32
Promiscuous Access	2		64
Promiscuous Trunk	2		186
Trunk Secondary	7		186
PVLAN	50	60	
vPC Config-Sync		120	
Netflow		84	
WCCP		58	

5 NVT Findings/Conclusion/Recommendations

Assigned/New	➔	*Still working on fixes and may be seen in CCO image*
Unreproducible	➔	*Not seen in CCO image may have fixed by other code fixes.*
Verified/Resolved	➔	*Fixed in CCO image*
Closed	➔	*System limitation and behavior will remain the same*

6 Appendix

6.1 VxLAN Solution Configuration Guide

The following configurations are divided as below

- Spine
- Leaf
- LISP Map server
- LISP Branch router

Spine configuration guide N77K

- Feature set requirement
  - Install feature-set fabric # Installed feature-set fabric

- Feature-set fabric # Enable/Disable fabric
- feature bgp # Enable/Disable Border Gateway Protocol
- feature lisp #Enable/Disable Locator/ID Separation Protocol
- feature pim #Enable/Disable Protocol Independent Multicast
- feature ospf # Enable/Disable Open Shortest Path First Protocol
- feature interface-vlan #Enable/Disable interface vlan
- feature lacp #Enable/Disable LACP
- feature bfd # Enable/Disable BFD
- feature nv overlay # Enable/Disable NV Overlay
- feature vni # Enable/Disable Virtual Network Segment (VNI)

- Global configuration command
  - nv overlay evpn # Enable/Disable Ethernet VPN (EVPN)
  - system bridge-domain #Configure Bridge-domain ID
  - vni <range> # configure VNI range
  - ip pim rp-address <> # Configure static RP for group range for BUM Traffic
  - bridge-domain <> # configure Bridge-domain ID
- VRF Configuration
  - vrf context <> # Create VRF and enter VRF mode
  - Vni <> # configure L3 VNI for VXLAN
  - rd auto # Generate RD automatically
  - address-family ipv4 unicast # Configure IPv4 address family
  - route-target import <> # Import Target-VPN community
  - route-target import <> evpn # Specify Target for EVPN routes
  - route-target export <> # Export Target-VPN community
  - route-target export <> evpn # Specify Target for EVPN routes
  - route-target both auto # Export And Import Target-VPN community
  - route-target both auto evpn # Specify Target for EVPN routes

- Map the bridge domain to L3 VNI
  - bridge-domain <> # Bridge-domain ID
  - member vni <> # configure L3 VNI associated with VRF

- Bridge-domain interface (each VRF must have 1 bridge-domain)
  - interface bdi <> # Bdi interface
  - no shutdown # Enable/disable an interface
  - vrf member <> # Configure VRF parameters
  - ip forward # Enable ip forwarding on interface

- Port-channel interface (for all L3 links)

  - interface port-channel1 # Port Channel interface
  - mtu 9216 # Configure mtu for the port
  - bfd interval <>min_rx <> multiplier <> # BFD interval commands
  - no ip redirects #disable IP directed-broadcast
  - ip address <> # Configure IP address on interface
  - ip pim sparse-mode #onfigures sparse-mode PIM on interface
  - Interface ethernet<> # Ethernet IEEE 802.3z
  - mtu 9216 # Configure mtu for the port

- channel-group 1 mode active #Configure port channel parameters
- no shutdown # Enable/disable an interface

- NVE interface
  - interface nve1 # NVE interface
  - no shutdown # Enable/disable an interface
  - source-interface <> # NVE Source-Interface
  - host-reachability protocol bgp # Configure host reachability advertisement
  - member vni <> associate-vrf # Associate vni with a vrf

- BGP Global configuration
  - router bgp 1 # Border Gateway Protocol (BGP)
  - address-family ipv4 unicast # Configure unicast address-family
  - network <loopback > # Configure an IP prefix to advertise
  - redistribute direct route-map <passall> # Directly connected
  - maximum-paths <> # Number of parallel paths (EBGP)
  - maximum-paths ibgp <> # Number of parallel paths (IBGP)
  - address-family l2vpn evpn # Configure L2VPN EVPN address-family
  - route-map <passall> permit 10 # Permit any
  - vrf <> # Production VRF
  - address-family ipv4 unicast # Configure unicast address-family
  - network 0.0.0.0/0 # Default route
  - advertise l2vpn evpn # advertise l2vpn evpn routes
  - redistribute direct route-map passall # Directly connected

- IBGP (Spine and Leaf configuration) – Suggestion to configure IPv4 over interface and EVPN peering over loopback

  - template peer IBGP-IPv4 # Template configuration for peer parameters
  - bfd # Bidirectional Fast Detection for the neighbor
  - remote-as 1 # Specify AS Number of the neighbor
  - password 3<> # Configure a password for neighbor
  - address-family ipv4 unicast # Configure unicast address-family
  - send-community # Send Community attribute to this neighbor
  - route-reflector-client # Configure a neighbor as Route reflector client
  - route-map SELF out # Apply route-map to neighbor
  - route-map SELF permit 10 # route-map name
  - set ip next-hop peer-address # Use peer address (for BGP only)
  - neighbor <> remote-as 1 # neighbor IBGP peering
  - inherit peer IBGP-IPv4 # Inherit a template
  - update-source <> # Specify source of BGP session and updates

  - template peer IBGP-EVPN # Template configuration for peer parameters
  - bfd # Bidirectional Fast Detection for the neighbor
  - remote-as 1 # Specify AS Number of the neighbor
  - password 3 <> #Configure a password for neighbor
  - update-source loopback0 # Specify source of BGP session and updates
  - address-family l2vpn evpn # Configure L2VPN EVPN address-family

- send-community extended #Send Standard and Extended Community attributes
- route-reflector-client # Configure a neighbor as Route reflector client
- neighbor <> remote-as 1 # neighbor IBGP peering
- inherit peer IBGP-EVPN # Inherit a template

- EBGP (Inter-POD Spine) – Suggestion to configure IPv4 over interface and EVPN peering over loopback
  - template peer EBGP-IPv4 # Template configuration for peer parameters
  - bfd # Bidirectional Fast Detection for the neighbor
  - remote-as 10 # Specify AS Number of the neighbor
  - password 3 <> #Configure a password for neighbor
  - address-family ipv4 unicast # Configure unicast address-family
  - send-community extended # Send Community attribute to this neighbor
  - neighbor <> remote-as 10 #neighbor EBGP peering
  - inherit peer EBGP-IPv4 # Inherit a template
  - template peer EBGP-EVPN # Template configuration for peer parameters
  - bfd # Bidirectional Fast Detection for the neighbor
  - remote-as 10 # Specify AS Number of the neighbor
  - password 3 <> #Configure a password for neighbor
  - update-source loopback0 # Specify source of BGP session and updates
  - ebgp-multihop 2 # Specify multihop TTL for remote peer
  - address-family l2vpn evpn # Configure L2VPN EVPN address-family
  - send-community extended #Send Standard and Extended Community attributes
  - route-map UNCHGD out # Apply route-map to neighbor
  - route-map UNCHGD permit 10 # route-map name
  - set ip next-hop unchanged # Use unchanged address (for eBGP session only)
  - neighbor <> remote-as 10 #neighbor EBGP peering
  - inherit peer EBGP-EVPN # Inherit a template

- IBGP (Spine and Legacy TOR configuration) – Legacy TORs eBGP peer-to-spine through "fake iBGP" by manipulating "local AS" in as-path. So BGP peering is iBGP, but the local AS are different, so the route gets redistributed to MP-iBGP L2VPN EVPN.
  - Router bgp 1 #Border Gateway Protocol (BGP)
  - VRF <> # Production VRF
  - neighbor <> remote-as 201 # neighbor EBGP peering
  - local-as 202 no-prepend replace-as eBGP # Specify the local-as number for the neighbor, Do not prepend the local-as number to updates from the eBGP

neighbor, Prepend only the local-as number to updates to eBGP neighbor

- address-family ipv4 unicast # Configure unicast address-family

- IBGP (Intra-POD Spine over Ring links) – OSPF over ring links an IPv4 and EVPN over loopback
  - router ospf 1 # Open Shortest Path First (OSPF)
  - router-id <loopback> # Set OSPF process router-id

  - interface Ethernet1/22
  - mtu 9216 # Configure mtu for the port
  - bfd interval <> min_rx <> multiplier 3 # BFD interval
  - no bfd echo # disable BFD Echo
  - no ip redirects # disable IP directed-broadcast
  - ip address <> # Configure IP address on interface
  - ip router ospf 1 area 0.0.0.0 # enable ospf on the interface
  - no shutdown # Enable/disable an interface

  - neighbor <> remote-as 1 #neighbor IBGP peering
  - update-source loopback0 # Specify source of BGP session and updates
  - address-family ipv4 unicast # Configure unicast address-family
  - send-community extended #Send Standard and Extended Community attributes
  - route-map LOCAL_PREF in #Apply route-map to neighbor
  - address-family l2vpn evpn # Configure L2VPN EVPN address-family
  - send-community extended #Send Standard and Extended Community attributes
  - route-map LOCAL_PREF in # Apply route-map to neighbor

  - route-map LOCAL_PREF permit 10 #route-map name
  - set local-preference 50 #BGP local preference path attribute (configure less than default 100)

  - EBGP (core link) – setup route-map which will advertise only non-mobility subnets

    - Router bgp 1 #Border Gateway Protocol (BGP)
    - VRF <> # Production VRF
    - neighbor <> remote-as 11 #neighbor IBGP peering
    - address-family ipv4 unicast #Configure unicast address-family
    - route-map WAN-map out # Apply route-map to neighbor

    - route-map WAN-map permit 10 #route-map name
    - match ip address prefix-list wan-list # Match entries of prefix-lists

    - ip prefix-list wan-list seq 5 permit <subnet> # only allow non-mobility prefixes toward core and LISP branch

LISP Configuration on Spine (N7K)

- vrf context <> # Production VRF
- ip lisp etr #Configures LISP Egress Tunnel Router (ETR)
- lisp instance-id 1000 #Configures Instance-ID for global data-mappings
- ip lisp itr map-resolver <ip add> #To interact with Map-Resolver (primary)
- ip lisp itr map-resolver <ip add> #To interact with Map-Resolver (secondary)
- ip lisp etr map-server <ip add>key <> # Authentication key used with Map-Server (Primary)
- ip lisp etr map-server <ip add>key <> # Authentication key used with Map-Server (secondary)
- lisp dynamic-eid 101 # Configure dynamic-EIDs for roaming
- database-mapping <Mobility network> < core link address> priority 1 weight #configure EID-prefix and locator-set for dynamic-EID
- register-route-notifications tag 1 # Register more-specific routes of the database-mapping EID-prefix to Map-Server, Include only routes with this BGP tag

LISP configuration on Map server (CSR1kv)

- router lisp #Locator/ID Separation Protocol
- ddt authoritative instance-id 1000 <Subnet> # ddt authoritative to allow (VXLAN Mobility subnet
- map-server-peer <ip address> # IPv4 Peer map-server locator address (secondary)
- map-server-peer <ip address> # IPv4 Peer map-server locator address (Primary)

- ddt # Enable ddt
- site <name> # LISP site name
- authentication-key <key> # password
- eid-prefix instance-id 1000 <subnet> accept-more-specifics # allow VXLAN Mobility subnet
- eid-prefix instance-id 1000 <subnet> accept-more-specifics # allow Branch subnet

- ipv4 map-server # Configures a LISP Map Server

- ipv4 map-resolver # Configures a LISP Map Resolver (MR)

LISP configuration on Branch (ASR1K)

- router lisp # Locator/ID Separation Protocol
- eid-table default instance-id 1000 #Configures Instance-ID for global data-mappings
- database-mapping <Mobility network> < core link address> priority 1 weight #configure EID-prefix and locator-set for dynamic-EID
- ipv4 itr map-resolver <ip address> #To interact with Map-Resolver (primary)
- ipv4 itr map-resolver <ip address> #To interact with Map-Resolver (secondary)
- ipv4 itr # Configures a LISP Ingress Tunnel Router (ITR)
- ipv4 etr map-server <ip address> key <key>#Authentication key used with Map-Server (Primary)
- ipv4 etr map-server <ip address> key <key>#Authentication key used with Map-Server (secondary)
- ipv4 etr # Configures a LISP Egress Tunnel Router (ETR)
- ipv4 map-cache-limit <> # Address family specific map cache configuration

Leaf configuration guide N9K

- Feature set requirement
  - feature bgp # Enable/Disable Border Gateway Protocol
  - feature pim #Enable/Disable Protocol Independent Multicast
  - feature interface-vlan #Enable/Disable interface vlan
  - feature lacp #Enable/Disable LACP
  - feature bfd # Enable/Disable BFD
  - feature nv overlay # Enable/Disable NV Overlay
  - feature vni # Enable/Disable Virtual Network Segment (VNI)
  - feature vn-segment-vlan-based # Enable/Disable VLAN based VN segment
  - feature vpc # Enable/Disable VPC (Virtual Port Channel)
- Global configuration command
  - nv overlay evpn # Enable/Disable Ethernet VPN (EVPN)
  - ip pim rp-address <> # Configure static RP for group range for BUM Traffic
  - fabric forwarding anycast-gateway-mac <> # Anycast Gateway MAC of the Switch
- Vlan to vn-segment mapping configuration command (Note it is require for each L2 VLAN and L3 VLAN per VRF)
  - vlan <> # Vlan commands
  - vn-segment <> # VN Segment id of the VLAN
- VRF Configuration
  - vrf context <> # Create VRF and enter VRF mode
  - Vni <> # configure L3 VNI for VXLAN

- rd auto # Generate RD automatically
- address-family ipv4 unicast # Configure IPv4 address family
- route-target import <> # Import Target-VPN community
- route-target import <> evpn # Specify Target for EVPN routes
- route-target export <> # Export Target-VPN community
- route-target export <> evpn # Specify Target for EVPN routes
- route-target both auto # Export And Import Target-VPN community
- route-target both auto evpn # Specify Target for EVPN routes

- Map the bridge domain to L3 VNI
  - bridge-domain <> # Bridge-domain ID
  - member vni <> # configure L3 VNI associated with VRF

- VPC domain <>

  - VPN domain # Specify domain
  - peer-switch # Enable peer switch on vPC pair switches
  - role priority <> # Configure priority to be used during vPC role (primary/secondary) election
  - system-priority <> # Configure system priority
  - peer-keepalive destination <>> source <> vrf <> # Keepalive/Hello with peer switch
  - peer-gateway # Enable L3 forwarding for packets destined to peer's gateway mac-address
  - ipv6 nd synchronize # Display Neighbor Discovery interface information
  - ip arp synchronize # CFS synchronize

- VPC Keepalive link

  - interface port-channel <> # Port Channel interface
  - mtu 9216 # Configure mtu for the port
  - vrf member VPC # Set interface's VRF membership
  - ip address <> # Configure IP address on interface

- VPC peer-link
  - interface port-channel<> # Port Channel interface
  - switchport mode trunk # Port mode trunk
  - switchport trunk allowed vlan <> #Set allowed VLAN characteristics when interface in trunking mode
  - spanning-tree port type network # Consider the interface as inter-switch link
  - vpc peer-link # Specify if this link is used for peer communication

- VPV leg port-channel (host connected link)

  - interface port-channel<> # Port Channel interface
  - switchport mode access # Port mode access

- switchport access vlan<> # Set access mode characteristics of the interface
- spanning-tree bpdufilter enable # Enable BPDU filtering for this interface
- mtu 9216 # Configure mtu for the port
- vpc <> #Virtual Port Channel configuration

- Vlan interface configuration

  - interface Vlan101 # Vlan interface
  - no shutdown # Enable/disable an interface
  - mtu 9216 # Configure mtu for the port
  - vrf member<> # Configure VRF parameters
  - no ip redirects #Disable Send ICMP Redirect messages
  - ip address <> # Configure IP address on interface
  - no ipv6 redirects # Disable sending ICMPv6 Redirect messages
  - fabric forwarding mode anycast-gateway # Anycast Gateway Forwarding Mode

- Vlan interface configuration mapped to L3 VNI

  - interface vlan <> # Vlan interface
  - no shutdown # Enable/disable an interface
  - mtu 9216 # Configure mtu for the port
  - vrf member <> # Configure VRF parameters
  - no ip redirects # Configure IP address on interface
  - ip forward #Enable ip forwarding on interface

- Port-channel interface (for all L3 links)

  - interface port-channel1 # Port Channel interface
  - mtu 9216 # Configure mtu for the port
  - bfd interval <>min_rx <> multiplier <> # BFD interval commands
  - no ip redirects #disable IP directed-broadcast
  - ip address <> # Configure IP address on interface
  - ip pim sparse-mode #Configures sparse-mode PIM on interface
  - Interface ethernet<> # Ethernet IEEE 802.3z
  - mtu 9216 # Configure mtu for the port
  - channel-group 1 mode active #Configure port channel parameters
  - no shutdown # Enable/disable an interface

- NVE interface
  - interface nve1 # NVE interface
  - no shutdown # Enable/disable an interface
  - source-interface <> # NVE Source-Interface
  - host-reachability protocol bgp # Configure host reachability advertisement
  - member vni <L3 VNI> associate-vrf # Associate L3 vni with a vrf
  - member vni <L2 VNI> #NVE VN-Segment Membership
  - mcast-group <> #NVE Multicast Group

- BGP Global configuration
  - router bgp 1 # Border Gateway Protocol (BGP)

- address-family ipv4 unicast # Configure unicast address-family
- network <loopback > # Configure an IP prefix to advertise
- maximum-paths ibgp <> # Number of parallel paths (IBGP)
- address-family l2vpn evpn # Configure L2VPN EVPN address-family
- route-map <passall> permit 10 # Permit any
- vrf <> # Production VRF
- address-family ipv4 unicast # Configure unicast address-family
- advertise l2vpn evpn # advertise l2vpn evpn routes
- redistribute direct route-map passall # Directly connected

- IBGP (Spine and Leaf configuration) – Suggestion to configure IPv4 over interface and EVPN peering over loopback

- template peer IBGP-IPv4 # Template configuration for peer parameters
- bfd # Bidirectional Fast Detection for the neighbor
- remote-as 1 # Specify AS Number of the neighbor
- password 3<> # Configure a password for neighbor
- address-family ipv4 unicast # Configure unicast address-family
- send-community # Send Community attribute to this neighbor
-
- neighbor <> remote-as 1 # neighbor IBGP peering
- inherit peer IBGP-IPv4 # Inherit a template
- update-source <> # Specify source of BGP session and updates

- template peer IBGP-EVPN # Template configuration for peer parameters
- bfd # Bidirectional Fast Detection for the neighbor
- remote-as 1 # Specify AS Number of the neighbor
- password 3 <> #Configure a password for neighbor
- update-source loopback0 # Specify source of BGP session and updates
- address-family l2vpn evpn # Configure L2VPN EVPN address-family
- send-community extended #Send Standard and Extended Community attributes
- neighbor <> remote-as 1 # neighbor IBGP peering
- inherit peer IBGP-EVPN # Inherit a template

- IBGP inter vPC SVI link over peer-link (Only for IPv4 peering) and add vlan id to peer link allow list

- interface Vlan3901 # Vlan interface
- no shutdown # Enable/disable an interface
- bfd interval <>min_rx <> multiplier <> # BFD interval commands
- no bfd echo # Disable Echo function for all address families
- no ip redirects #disable IP directed-broadcast
- ip address <> # Configure IP address on interface
- no ipv6 redirects # Disable sending ICMPv6 Redirect messages
- ip pim sparse-mode #Configures sparse-mode PIM on interface

- router bgp <> #Border Gateway Protocol (BGP)
- address-family ipv4 unicast # Configure unicast address-family

- network <> # Add SVI interface address
- neighbor <> remote-as 1 # neighbor IBGP peering
- bfd # Bidirectional Fast Detection for the neighbor
- password 3 <> # Configure a password for neighbor

- update-source <svi id) #Specify source of BGP session and updates
- address-family ipv4 unicast # Configure unicast address-family
- send-community extended #Send Standard and Extended Community attributes
- route-reflector-client # Configure a neighbor as Route reflector client
- route-map NHS out #Apply route-map to neighbor

- route-map NHS permit 10 #route-map name
- set local-preference 50 #BGP local preference path attribute (configure less than default 100)
- set ip next-hop peer-address # Use peer address (for BGP only)

- EVPN Configuration for each L2 VNI
  - Evpn # Enter EVPN configuration mode
  - vni <> l2 # Configure Ethernet VPN ID
  - rd auto # VPN Route Distinguisher Auto
  - route-target import auto # Import Target-VPN community and Generate RT automatically
  - route-target import <> # RT extcommunity in aa:nn format
  - route-target export auto # Export Target-VPN community and Generate RT automatically
  - route-target export <> #RT extcommunity in aa:nn format

6.2 Service Provider MPLS Configuration Guideline

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 >
  - power redundancy-mode ps-redundant *# set the power in redundancy mode*

- no system admin-vdc *# disable system admin-vdc*
- install feature-set fabricpath *# enable feature-set for fabricpath*
- install feature-set mpls *# enable feature-set for mpls*
- vdc <VDC_name> id 1 *# spawns a new named VDC*
  - limit-resource module-type <module-type> *#enable <module-type> modules for the given VDC*
  - allow feature-set fabricpath *# allows fabricpath feature-set to be enabled on the given VDC*
  - allow feature-set mpls *# allows mpls feature-set to be enabled on the given VDC*

  - allocate interface <interface-ranges> *# allocate one of more interface ranges to the given VDC.* ***When assigning any range of interfaces the maximum granularity applicable for the "allocate interface" command is limited to the layout of the port-group for each specific type of module.*** *Attaching to the module and using "show hardware internal dev-port-map" allows to display the mapping between front panel ports and ASIC instances.*

  - limit-resource u4route-mem minimum 200 maximum 200 *# allows to allocate more shared memory for the IPv4 routes. To estimate the amount of memory to be allocated for both IPv4 or IPv6 routes, the following two commands can be used: "show ip route sum"on each VRF and "show routing <ip|ipv4|ipv6> memory estimate routes <total-number-of-routes> next-hop <N>". Likewise, the same method can be applied for multicast allocation.*

- control-plane
  - service-policy input test--copp-policy-strict *# enables the test—copp-policy-strict for the control-plane*
    copp copy profile strict prefix mvpn
    class-map type control-plane match-any mvpn-copp-class-normal
    match exception ip multicast rpf-failure
    match exception ipv6 multicast rpf-failure
    control-plane *# apply the modified mvpn class to the control-plane to reduce the effects of PIM assert issue* **(CSCut68318)**
    service-policy input mvpn-copp-policy-strict

- snmp-server user admin network-admin auth md5 0x213ea412ed9e450340d412f4c9741a25 *# enables user admin with network-admin privileges and md5 authentication*

- ip pim auto-rp forward listen *# enables to forward auto-rp messages"*
- ip msdp originator-id <interface> *# enables the originator-id for MSDP messages*
- ip msdp peer <remote-peer-address> connect-source <interface> *# establish MSDP peering with a given remote peer*

- spanning-tree vlan <vlan-ranges> priority <priority> *# enables spanning tree for the specified list of vlan ranges*

- vrf context <vrf-name> *# create a new vrf*
  - ip pim rp-address <rp-address> group-list <mcast-groups/mask> *# configure static RP for the specified set of multicast groups within the given vrf*
  - rd <N:M> *# configure the route-discriminator for the given vrf. When deployed in a vPC/vPC+ scenarios it is recommended to use different RDs on the two vPC peers.*
  - mdt default <mcast-group-address> *# configure the MDT default tunnel for the given vrf*
  - mdt data <mcast-group-address/mask> threshold <rate> *# configure a set of MDT data tunnels for the given vrf depending on the size of the configured mask*
  - address-family ipv4 unicast *# configure IPv4 unicast AF*
  - route-target import <NN:M>
  - route-target export <NN:M>

- vpc domain <N> *# configure vPC domain number N*
  - peer-switch *# enables peer-switch on the two vPC peers to act as a single device for the STP*
  - peer-keepalive destination <remote-peer> source <local-peer> vrf keepalive *# configure the peer-keepalive on the specified vrf.*
  - delay restore 180 *# delays the leg bringup for 3 minutes to guarantee the convergence of the routing protocols to minimize NS traffic disruption*
  - peer-gateway *# to allow the vPC peer device to act as the active gateway for packets addressed to the other peer device router MAC*
  - fabricpath multicast load-balance *# enables load-balancing of multicast traffic between the two vPC+ peer devices.*
  - fabricpath switch-id 1 *# defines the emulated switch-id for the fabricpath vPC pair*
  - config-sync *# enable config synchronization between the two vPC peers*
  - ip arp synchronize *# enable ARP synchronization between the two vPC peers*

- Switched Virtual Interfaces (SVIs):
  - interface VlanN
  - mtu 9216 *# enable jumbo frames*
  - vrf member <vrf-name> *# assign the SVI to a given vrf*
  - no ip redirects *# prevent the router to send redirects messages to the clients (ICMP)*
  - ip address <ip-address>/<24 bit mask>
  - ipv6 address <ipv4-address>/<64 bit mask>
  - ip ospf passive-interface *# disable the routing updates on the speficied OSPF interface preventing the formation of any OSPF adjacency*
  - ip router ospf 1 area 0.0.0.0 *# enable OSPF on the specified interface and place it in area 0*
  - ip pim sparse-mode *# enable PIM sparse mode*
  - hsrp version 2 *# set the HSRP version two*
  - hsrp G *# defines the HSRP group G for the specified interface*

- preempt delay minimum 120 *# delays the preemption for two minutes upon HSRP switchover to minimize network disruptions*
- priority 101 forwarding-threshold lower 1 upper 101 *#define the priority for this HSRP peer*
- ip <hsrp-ip-address> *# HSRP virtual IP address*
- hsrp Gv6 ipv6 *# defines the HSRPv6 group Gv6 for the specified interface*
- preempt delay minimum 120
- priority 101 forwarding-threshold lower 1 upper 101
- ip <hsrpv6-ip-address>

- LACP: LACP is used for link aggregation to form port-channels across the network:
  - interface port-channelN *# create a port-channel N*
  - mtu 9216 *# enable jumbo frames on the port-channel N*
  - ip address <ip-address>/<24 bit mask>
  - ip ospf network point-to-point *# define the OSPF network type as p2p to minimize the time to form the OSPF adjacency*
  - ip router ospf 1 area 0.0.0.0 *# enable OSPF on the specified interface and place it in area 0*
  - ip pim sparse-mode *# enable PIM sparse mode*

  - interface port-channelM *# create the vPC peer-link as a switched port-channel carrying all the necessary VLANs.*
  - switchport
  - switchport mode trunk
  - switchport trunk allowed vlan 11-510,2001-2500
  - spanning-tree port type network *# defines the network type as network for the vPC peer-link*
  - mtu 9216 *# enable jumbo frames*
  - vpc peer-link *# defines the port-channel M as the vPC peer-link*

  - interface port-channel R
  - switchport
  - switchport mode trunk
  - switchport trunk allowed vlan 11-60
  - mtu 9216 *# enable jumbo frames*
  - vpc R *# defines the port-channel R as a vPC leg*

- MPLS LDP configuration:
  - mpls ldp configuration
  - router-id Lo0 force *# defines the loopback interface 0 as router-id for all the LDP updates*

- OSPF/LDP configurations:

  - router ospf 1 *# set the OSPF process id as one*
  - router-id <router-id> *# set the OSPF router-id*
  - mpls ldp autoconfig area 0.0.0.0 *# turn on LDP on all the OSPF interfaces configured in the backbone area (refer to section 5 for further information)*

- BGP: iBGP is configured between the core switches and the two route-reflectors. iBGP is also configured between each PE on each site and the two RRs:
    - router bgp 1 *# defines the BGP AS as one*
    - graceful-restart-helper *# enables the graceful-restart-helper*
    - address-family ipv4 unicast
    - maximum-paths ibgp 2 *# enable ICMP for the AF IPv4 unicast*
    - template peer RR-CLIENT *# defines the template for the RR iBGP peering*
    - remote-as 1 *# iBGP session*
    - update-source loopback0 *# uses the loopback interface zero as the source of the BGP updates and messages*
    - address-family ipv4 unicast
    - route-reflector-client *# set the template to be a client for the route-reflector cluster in the AF IPv4 unicast*
    - address-family vpnv4 unicast
    - send-community extended
    - address-family vpnv6 unicast
    - send-community extended *# allows the extended attributes for both AF VPNv4 and VPNv6*
    - address-family ipv4 mdt *# enables MVPN services*
    - neighbor <RR1-loopback-interface-0>
    - inherit peer RR-CLIENT
    - neighbor <RR2-loopback-interface-0>
    - inherit peer RR-CLIENT *# enable iBGP peering with the two route-reflectors configured in a cluster*
    - vrf vrf1 *# defines the VPNv4 parameters such as route injections, route redistributions, ECMP, etc. etc.*
    - address-family ipv4 unicast
    - network <loopback interface>/<32 bit mask> *# inject the loopback interface to be advertised to the other iBGP peers. This interface is used as customer RP interface*
    - network <SVI1-ip-address>/<24 bit mask>
    - …
    - network <SVIN-ip-address>/<24 bit mask> *# inject the SVI subnets that belong to the specified vrf into iBGP*
    - maximum-paths ibgp 2 *# enable ECMP*
    - address-family ipv4 unicast

- IGMPv2: 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*

- 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 *# procedure to install and enable fabricpath feature set*
    - fabricpath topology 12
    - member vlan 2001-2500 *# defines the set of VLANs used for the vPC+ access switch (CE12)*

        - fabricpath topology 6
        - member vlan 11-260
        - fabricpath topology 7
        - member vlan 261-510
        - vlan 11-510,2001-2500
        - mode fabricpath *# defines the 2 fabricpath topologies deployed for the 2 CE devices (CE6 and CE7)*
        - fabricpath switch-id 106 *# defines the local FP switch-id*
        - vpc domain 100 *# enters the vPC/vPC+ parameters*
        - fabricpath multicast load-balance
        - fabricpath switch-id 1 *# defines the emulated FP switch-id. This is common between the two vPC peers*
        - interface port-channel7 *# This interface is used as vPC peer-link. It has to carry both FP and vPC+ VLANs and it has to be configured in FP mode*
        - switchport mode fabricpath
        - fabricpath isis metric 800 *# since in steady state it is not desired to forward traffic through the vPC peer-link, it is recommended to set the ISIS metric the highest in the FP domain*
        - fabricpath topology-member 6
        - fabricpath topology-member 7
        - fabricpath topology-member 12 *# FP topology definitions*
        - interface port-channel100
        - switchport mode fabricpath

fabricpath isis metric 10 *# in order to minimize the root changes for multicast traffic, it is recommended to set the ISIS metric in all the FP interfaces. This will minimize the impact on the traffic convergence upon network disruptions*

6.3 DC1 Configuration Guideline

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 comprises 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

6.4 ENT1 Configuration Guide

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

6.5 M1 vPC Scale Configuration Guideline

The following configurations are applied to the test network:

- Common system control, management and accounting: Common system features like SSH, Syslog, SNMP, NTP and Management VRF are configured.
  - 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 >

- vPC: vPC technology is deployed in the network between the N7k and the Catalyst VSS switches as shown in the figure 3.
  - 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*
    - 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 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 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*

- 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*

- PVLAN: PVLAN is configured in the network and is the main focus of testing. The following PVLAN components are covered in the network:
    - PVLAN primary and secondary VLAN(Isolated)
    - Secondary Trunk and promiscuous trunk on vPC
    - Private VLAN promiscuous access and Private vlan host on classis port-channel.
    - Private VLAN promiscuous trunk, secondary trunk on classic Port-channel

    - 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
        - switchport private-vlan trunk allowed vlan 1 *# configure native vlan*
        - switchport private-vlan association trunk <primary> <secondary>
    - interface port-channel <port-channel> *# configure port-channel*
        - switchport
        - switchport mode private-vlan trunk promiscuous
        - switchport private-vlan mapping trunk <primary secondary1, secondary2,…> *# configure primary to secondary mapping*
    - interface port-channel <port-channel> *# configure port-channel*
        - switchport
        - switchport mode private-vlan promiscuous *# configure promiscuous access*
        - switchport private-vlan mapping <primary secondary> *# configure primary to secondary mapping*
    - interface port-channel <port-channel> *# configure port-channel*
        - switchport
        - switchport mode private-vlan host *# configure host*

switchport private-vlan host-association <primary secondary> *# configure primary to secondary host-association*