# Nexus Validation Test Phase 3.6

# 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 3.6 network topologies, hardware and software configurations, test procedures and findings.

NVT Phase 3.6 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.
- **MPLS Profile 1**: This is divided into access technologies.
  - o Classic Ethernet Access: This network focuses on building and operating an MPLS core network that provides end-to-end L3VPN and MVPN services with the Nexus 7000 Sup2E and Nexus 7700 Sup2E between two different sites: Site 8 is mainly focused on legacy M1 modules whereas Site 10 is built to test and verify the MPLS feature set over the new F3 modules. Both sites use virtual Port-Channel (vPC) to deliver highly available unicast and multicast services.
  - o Fabric Path Access: In addition to the Phase 1 topology, the new Site 6 has been integrated to the MPLS topology. This network uses the new F3 modules with both virtual Port-Channel (vPC+) and FabricPath solutions to deliver highly available unicast and multicast services. Furthermore, the deployment of FabricPath will increase the overall bandwidth (since vPC solution is limited to only two vPC peers) allowing to deploy multi-spine network layer, while eliminating the need for STP and allowing a "plug-and-play" type of network as the addition of new devices at either the spine or the leaf layer does not present any significant traffic loss for the pre-established flows.
- **Service Provider MPLS**: 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
- **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.
- **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

components. 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 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. It 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.
- **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 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 by Nexus access switches using NIC teaming (in both active mode and On mode). UCS-B series are connected to Nexus access switched thru fabric interconnect. NTP is synced to the server.

# 2 NVT Topology and Scale Achieved

## 2.1 VxLAN Solution

*Topology*
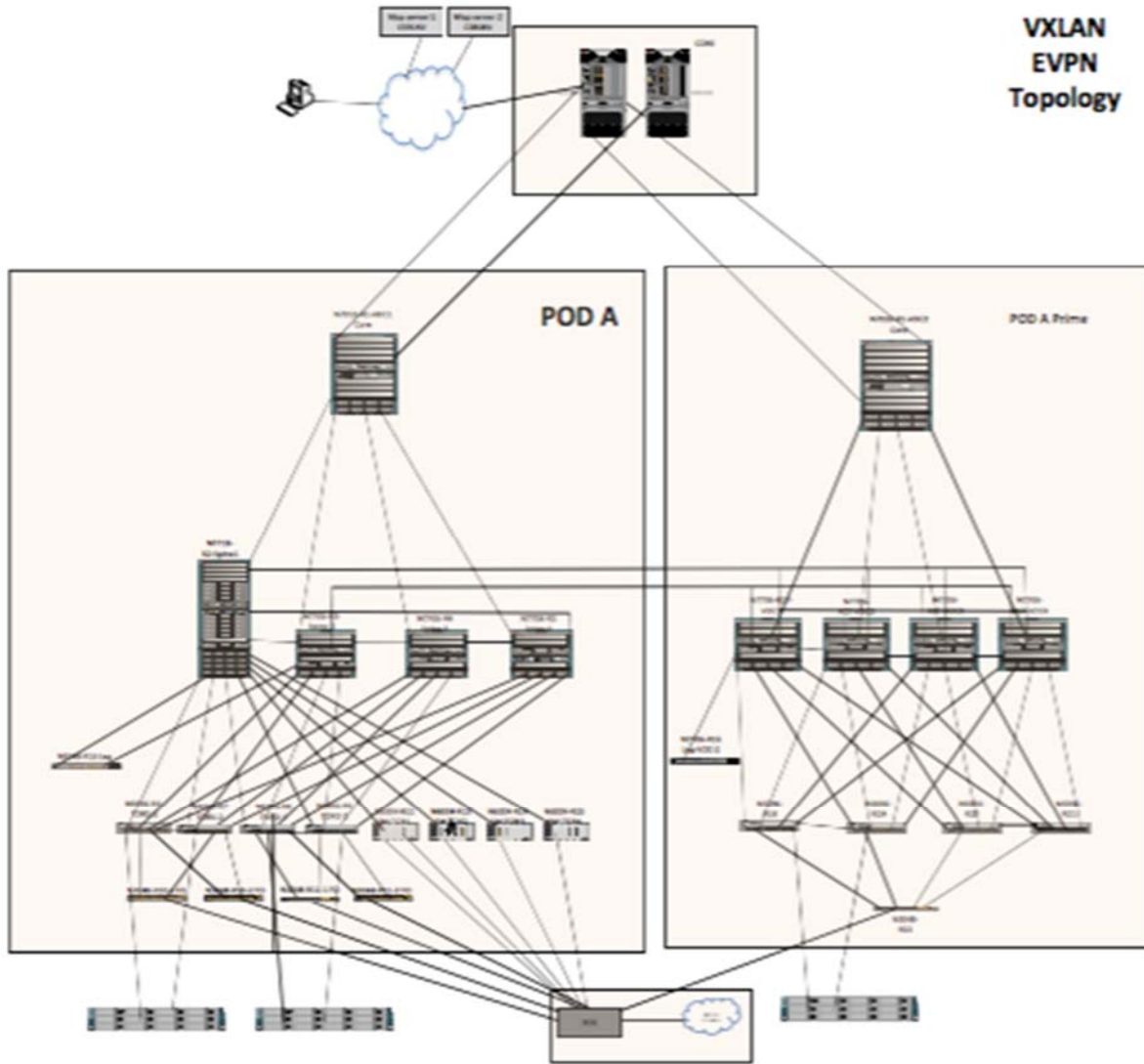


*Scale*

| 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** |

## 2.2   MPLS Profile 1 with F3, M1 and M2 and Fabric Path and CE Access

*Topology*
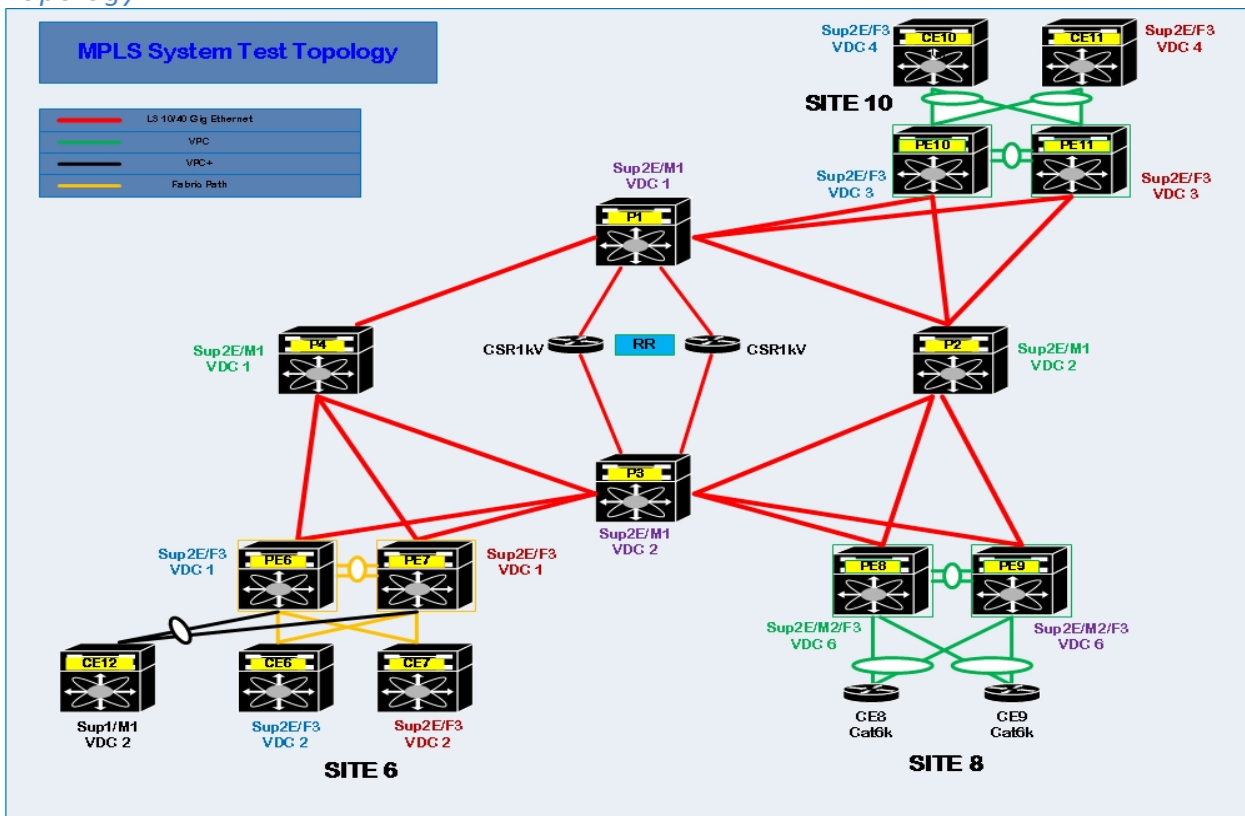


*Scale*

| Feature/Parameter | Classical Ethernet | | Fabric Path |
|---|---|---|---|
| | Site 10 (F3 vPC) | Site 8 (M1 vPC) | Site 6 (F3 FP/vPC+) |
| VLAN | 1000 | 1000 | 1000 |
| SVI | 1000 | 1000 | 1000 |
| IPv4 hosts x subnet | 32 | 32 | 32 |
| Hosts  per VLAN | 47+ | 47+ | 47+ |
| vPC/(vPC+) | 20/na | 2/na | 2/(1) |
| VLANS x vPC/(vPC+) | 50/na | 500/na | 250/(500) |

| | | | |
|---|---|---|---|
| Fabric Path IS-IS adjacencies | na | na | 6 |
| Port-channel (excluding peer-link, vPC&vPC+) | 4 | 4 | 4 |
| HSRPv4 version2 | 1000 | 1000 | 1000 |
| HSRPv6 version2 | 1000 | 1000 | 1000 |
| VRF | 100 | 100 | 100 |
| VLANs per VRF | 10 | 10 | 10 |
| ARP x VRF | 350+ | 350+ | 350+ |
| mVRF | 50 | 50 | 50 |
| PIM neighbor default VRF | 4 | 4 | 4 |
| PIM neighbor x VRF | 36 | 36 | 36 |
| MDT default tunnel | 50 | 50 | 50 |
| MDT data tunnel | 200 | 200 | 200 |
| Number of local customer groups | 10 | 10 | 10 |
| Number of sources x customer group | 8 | 8 | 10 |
| Number of OIF x group | 10 | 10 | 10 |
| OSPF Peers default VRF | 4 | 4 | 4 |
| OSPF Routes/(Path) | 48/(54) | 48/(54) | 48/(54) |
| LDP Sessions | 4 | 4 | 4 |
| iBGP IPv4 Sessions | 2 | 2 | 2 |
| iBGP VPNv4 Sessions | 2 | 2 | 2 |
| iBGP VPNv6 Sessions | 2 | 2 | 2 |

*Hardware*

| Classical Ethernet - vPC | Model No. | Software |
|---|---|---|
| N7K | N7K SUP2E / F3 /M2 / M1XL | 7.2.0 |
| C6K | VS-SUP720-10G | 151-2.SY2 |
| CSR1000v | Na | 15.4.3.S1 |
| Fabric Path – vPC+ | Model No. | |
| N7700 | N7700 SUP2E / F3 40G | 7.2(0) |
| N7000 | N7K SUP1 / M1 | 7.2(0) |

## 2.3 Service Provider MPLS F3

*Topology*



Solution Testing Topology
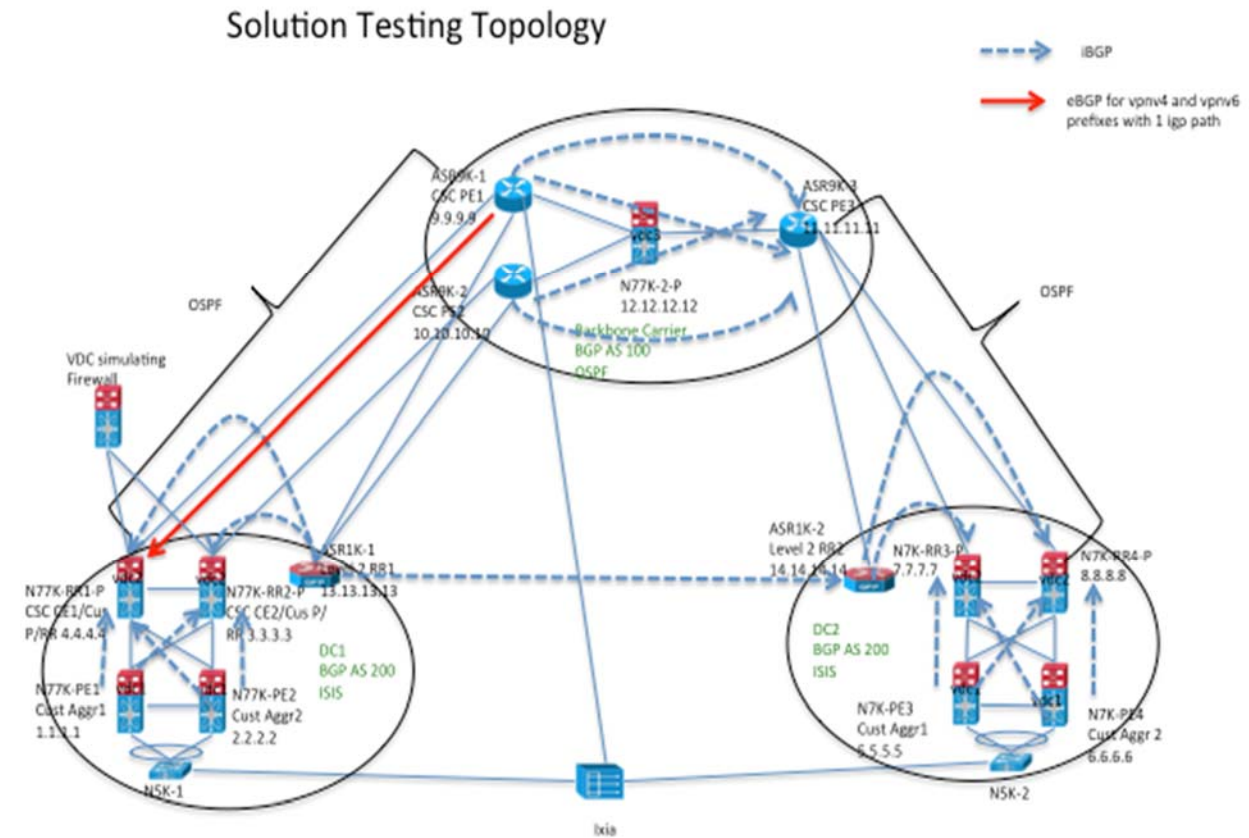
*Scale*

| 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 |

| Feature | Scale |
|---|---|
| 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 |

*Hardware*

| | Model No. | Image Version |
|---|---|---|
| N77K | N77-SUP2E/ N77-F324FQ-25 | 7.2 |
| N7K | N7K-SUP2E/ N7K-F312FQ-25 | 7.2 |
| N5K | N5K-C5548UP-SUP | 7.0(5)N1(1) |
| ASR9K | ASR9001-RP/ A9K-MPA-20X1GE | 5.1.1 |
| ASR1K | ASR1000-RP2/ ASR1000-ESP20/ SPA-8X1GE-V2 | 3.14 |

## 2.4   Dynamic Fabric Automation (Vinci) Solution
*Topology*

**SSTE Vinci-Fabric Encapsulation Test Topology**
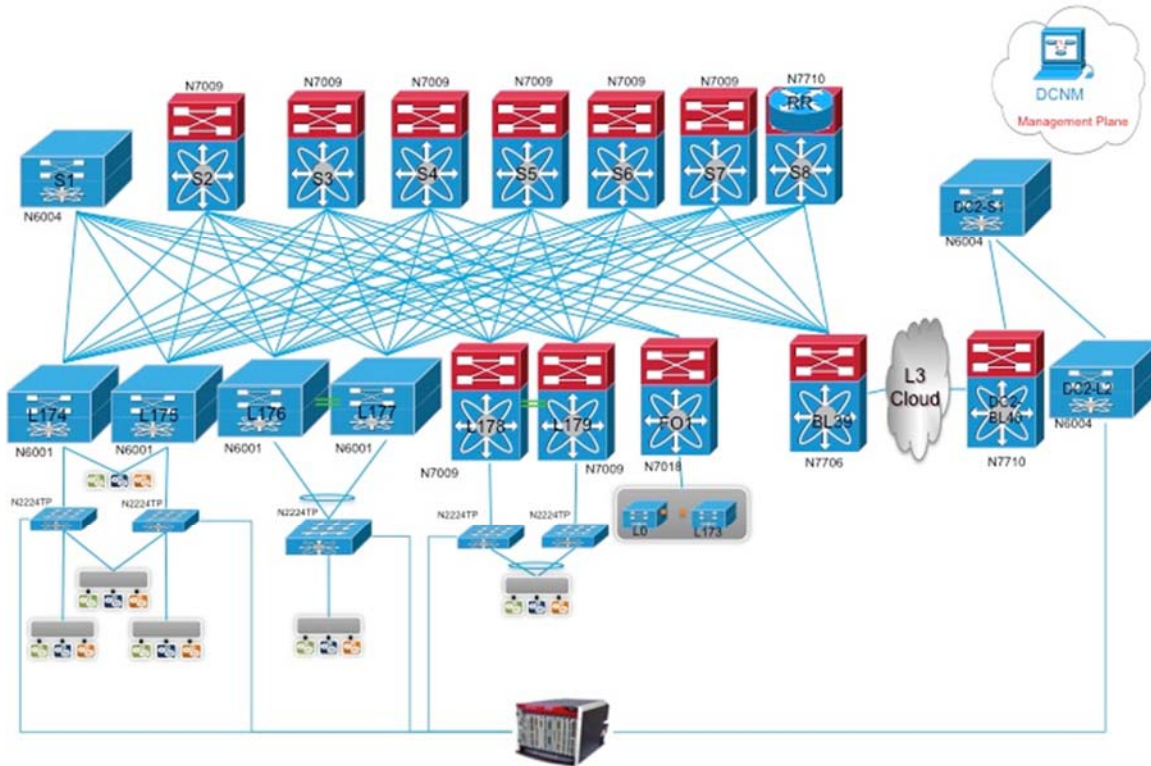


*Scale*

| 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 |

## 2.5   Data Center 1

*Topology*
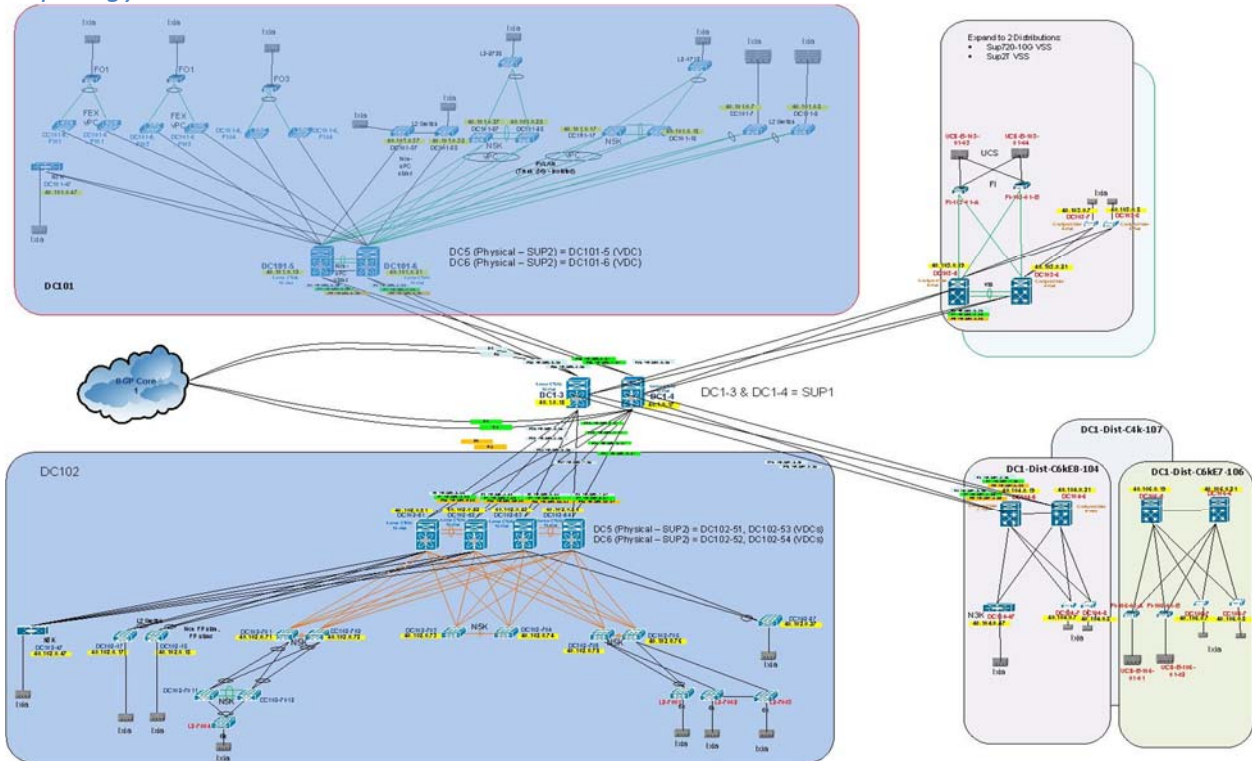


*Scale*

| Feature | 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 |

## 2.6   Enterprise 1

*Topology*
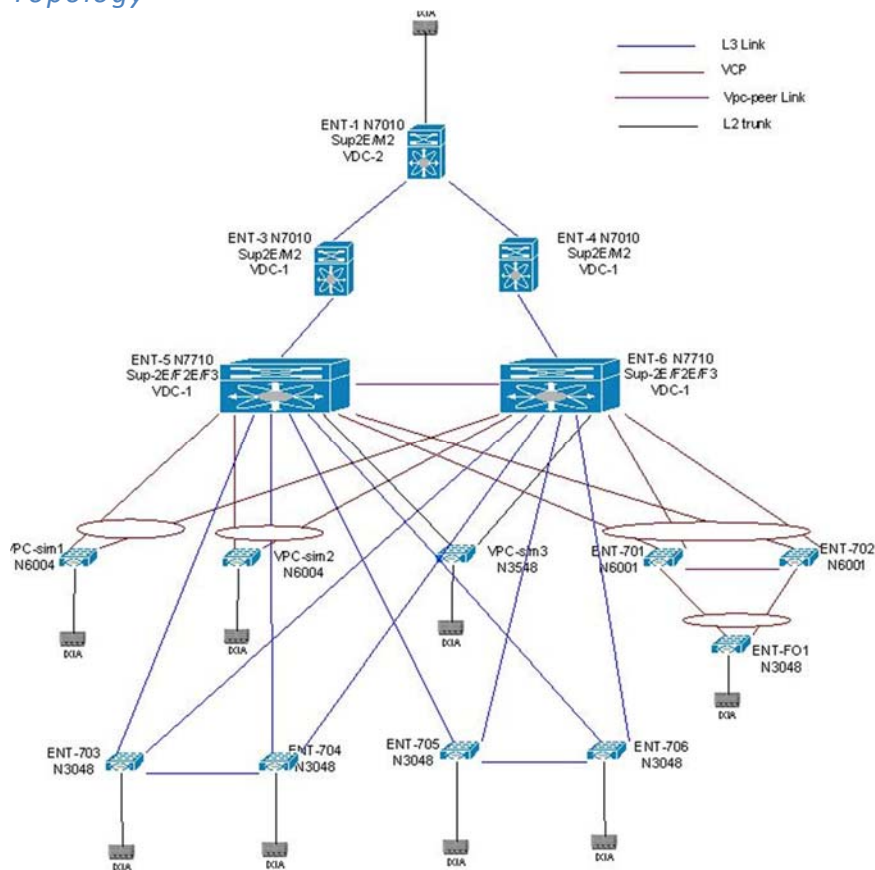


*Scale*

| 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 |

## 2.7 MSDC Scale Profiles

*Topology*



EDCS-963802 Profile 1, 2-1 Logical Topology
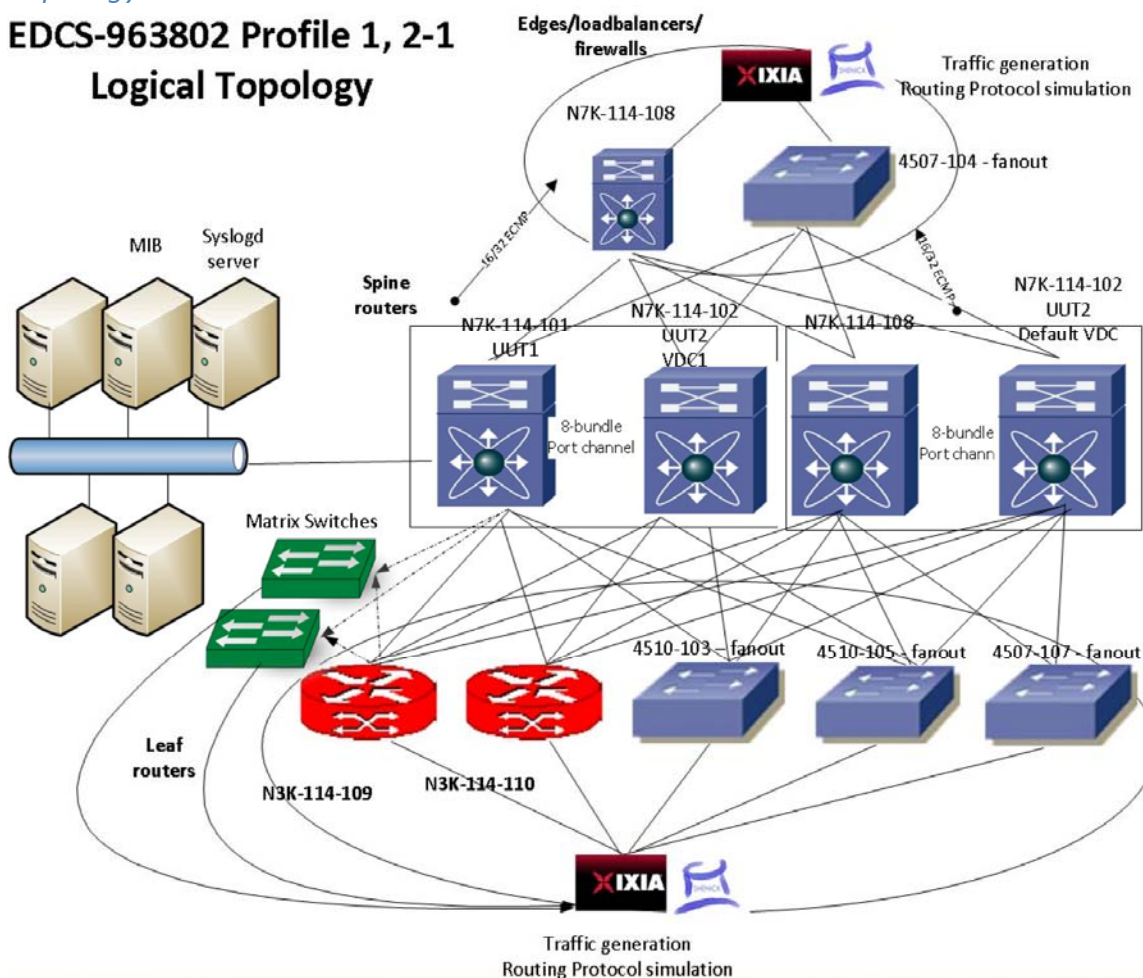
*Scale*

| 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 |

## 2.8  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 |

*Topology*
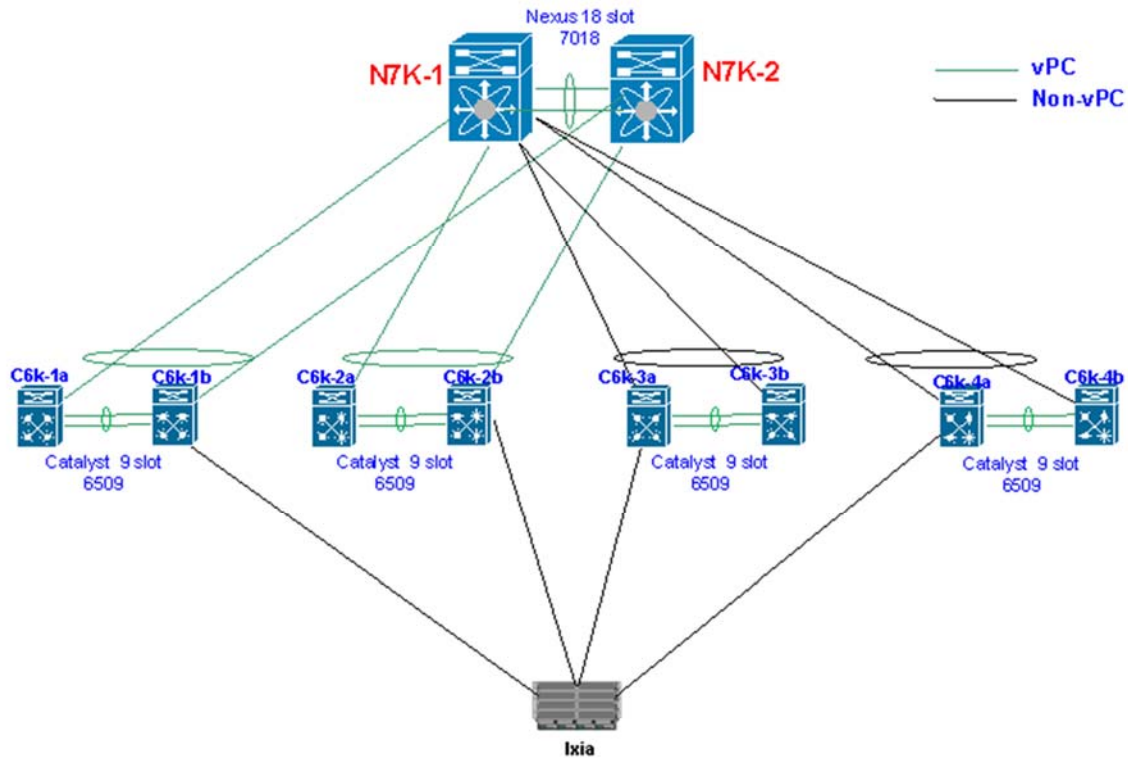


## 3 ISSU Matrix

| Image | ISSU/COLD BOOT | MPLS Profile 1 | DC1 | ENT | MSDC | M1vPC |
|---|---|---|---|---|---|---|
| 7.2.0 >7.2.0.upg | ISSU | PASS | PASS | PASS | PASS | PASS |
| 6.2.12 > 7.2.0 | ISSU | PASS | PASS | PASS | PASS | |
| 6.2.10 > 7.2.0 | ISSU | PASS | | PASS | PASS | |
| 6.2.8a > 7.2.0 | ISSU | | | PASS | | |
| 6.2.8b > 7.2.0 | ISSU | PASS | | PASS | | |
| 7.2.0 > 7.2.0.upg | COLD BOOT | | | PASS | | |
| 6.2.12 > 7.2.0 | COLD BOOT | PASS | | PASS | | |
| 6.2.10 > 7.2.0 | COLD BOOT | PASS | | PASS | | |
| 6.2.8a > 7.2.0 | COLD BOOT | | | PASS | | |
| 6.2.8b > 7.2.0 | COLD BOOT | PASS | PASS | PASS | | PASS |
| 7.2.0 > 6.2.12 | COLD BOOT | | PASS | PASS | | |
| 7.2.0 > 6.2.10 | COLD BOOT | | | PASS | | |
| 7.2.0 > 6.2.8a | COLD BOOT | | | PASS | | |
| 7.2.0 > 6.2.8b | COLD BOOT | | PASS | PASS | | |

# 4   Profile Details

## 4.1   VxLAN Solution

### 4.1.1   Network Topology and 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



BGP Design

eBGP
pfx list allow non-mobility pfxs and legacy pfxs

iBGP lo0
Ipv4, evpn
local_pref 50

eBGP intf ipv4
lo0 evpn

AS 100
local as 200
remote-as 300 prepend

N7700

IGP

"fake"
eBGP ipv4

AS 100
local as 300
remote-as 200 prepend

N300

AS 100

iBGP ipv4 intf
iBGP evpn lo0
Spine: next-hop-peer-addr

N300

iBGP ipv4 intf
RR clients each other
local_pref 50
next-hop-peer-addr

# LISP

Branch

1. GARP
2. BGP EVPN update (MAC, IP)
3. a. eBGP EVPN update
3. b. LISP update MS
4. Branch XTR send traffic to XTR old POD
5. LISP SMR to Branch XTR to check MS
6. Branch XTR query MS
7. Traffic send to new POD XTR

# Data Packet Flow

Native IPv4
VxLAN Encaped
LISP Encaped

MS

#### 4.1.1.1 VxLAN Solution Topology



### 4.1.2 Hardware and Software Overview

|          | Model No.                                     | NVT 3.6    |
|----------|-----------------------------------------------|------------|
| N77K     | N77-SUP2E / N77-F324FQ-25                     | 7.2.0      |
| 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      |

### 4.1.3 VxLAN Solution Configuration Guide

Please see the appendix for configuration details

### 4.1.4 VxLAN Solution Scale

| Feature           | Scale                            |
|-------------------|----------------------------------|
| Maximum VLANs     | 510                              |
| VLANs per TOR     | 14                               |
| VRFs per ToR      | 10                               |
| Max VNI per ToR   | 24                               |
| Max VNI per network | 10 VRFs + 500 VLANs = 510      |

| Max underlay Multicast Groups | 510 (to support broadcast + multicast traffic |
| Max ToR per 'network' | 250  (125 ToRs per POD * 2) = We build in pairs. |
| Max VTEP peers per 'network' | 250  (125 VTEPs per POD * 2) = We build in pairs. |
| 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 |

### 4.1.5   Convergence Results

| Shutdown 1 Port channel from Spine to leaf | Low | High |
| --- | --- | --- |
| East to West L2 VNI Unicast traffic | 44ms | 290ms |
| East to West L3 VNI Unicast traffic ( POD A to POD A) | 42ms | 85ms |
| East to West L3 VNI Unicast traffic ( POD A to POD A') | 24ms | 259ms |
| East to West L2 VNI Multicast traffic (POD A and POD A') | 75ms | 75ms |
| North to Sourth Branch LISP Unicast traffic  - bi-directional | 43ms | 975ms |
| Legacy to North Unicast traffic | 0ms | 0ms |
| Legacy to PODA Unicast | 43ms | 968ms |
| Legacy to PODA' Unicast | 0ms | 0ms |
| Legacy to Legacy Unicast traffic | 0ms | 0ms |
|  |  |  |
| No Shutdown 1 Port channel from Spine to leaf | Low | High |
| East to West L2 VNI Unicast traffic | 0ms | 0ms |
| East to West L3 VNI Unicast traffic ( POD A to POD A) | 0ms | 0ms |
| East to West L3 VNI Unicast traffic ( POD A to POD A') | 27ms | 27ms |
| East to West L2 VNI Multicast traffic (POD A and POD A') | 0ms | 0ms |
| North to Sourth Branch LISP Unicast traffic  - bi-directional | 0ms | 0ms |
| Legacy to North Unicast traffic | 0ms | 0ms |
| Legacy to PODA Unicast | 0ms | 0ms |
| Legacy to PODA' Unicast | 0ms | 0ms |
| Legacy to Legacy Unicast traffic | 0ms | 0ms |
|  |  |  |

| Shutdown all 4 Port channels from Spine to all 4 leafs | Low | High |
|---|---|---|
| East to West L2 VNI Unicast traffic | 99ms | 175ms |
| East to West L3 VNI Unicast traffic ( POD A to POD A) | 155ms | 224ms |
| East to West L3 VNI Unicast traffic ( POD A to POD A') | 56ms | 101ms |
| East to West L2 VNI Multicast traffic (POD A and POD A') | 127ms | 527ms |
| North to Sourth Branch LISP Unicast traffic  - bi-directional | 644ms | 696ms |
| Legacy to North Unicast traffic | 0ms | 0ms |
| Legacy to PODA Unicast | 716ms | 764ms |
| Legacy to PODA' Unicast | 0ms | 0ms |
| Legacy to Legacy Unicast traffic | 0ms | 0ms |
|  |  |  |
| No Shutdown all 4 Port channels from Spine to all 4 leafs | Low | High |
| East to West L2 VNI Unicast traffic | 0ms | 0ms |
| East to West L3 VNI Unicast traffic ( POD A to POD A) | 0ms | 0ms |
| East to West L3 VNI Unicast traffic ( POD A to POD A') | 18ms | 307ms |
| East to West L2 VNI Multicast traffic (POD A and POD A') | 0ms | 0ms |
| North to Sourth Branch LISP Unicast traffic  - bi-directional | 0ms | 0ms |
| Legacy to North Unicast traffic | 0ms | 0ms |
| Legacy to PODA Unicast | 0ms | 0ms |
| Legacy to PODA' Unicast | 0ms | 0ms |
| Legacy to Legacy Unicast traffic | 0ms | 0ms |
| Shutdown all 4 Port channels from leaf to all 4 spines | Low | High |
| East to West L2 VNI Unicast traffic | 150ms | 3.8sec |
| East to West L3 VNI Unicast traffic ( POD A to POD A) | 166ms | 502ms |
| East to West L3 VNI Unicast traffic ( POD A to POD A') | 91ms | 155ms |
| East to West L2 VNI Multicast traffic (POD A and POD A') | 151ms | 151ms |
| North to Sourth Branch LISP Unicast traffic  - bi-directional | 55ms | 1.4sec |
| Legacy to North Unicast traffic | 0ms | 0ms |
| Legacy to PODA Unicast | 55ms | 1.4sec |
| Legacy to PODA' Unicast | 0ms | 0ms |
| Legacy to Legacy Unicast traffic | 0ms | 0ms |

| No Shutdown all 4 Port channels from leaf to all 4 spines | Low | High |
| --- | --- | --- |
| East to West L2 VNI Unicast traffic | 0ms | 0ms |
| East to West L3 VNI Unicast traffic ( POD A to POD A) | 0ms | 0ms |
| East to West L3 VNI Unicast traffic ( POD A to POD A') | 0ms | 0ms |
| East to West L2 VNI Multicast traffic (POD A and POD A') | 101ms | 101ms |
| North to Sourth Branch LISP Unicast traffic  - bi-directional | 0ms | 0ms |
| Legacy to North Unicast traffic | 0ms | 0ms |
| Legacy to PODA Unicast | 0ms | 0ms |
| Legacy to PODA' Unicast | 0ms | 0ms |
| Legacy to Legacy Unicast traffic | 0ms | 0ms |
| | | |
| | | |
| Retstart LISP in N7K Spine | Low | High |
| East to West L2 VNI Unicast traffic | 0ms | 0ms |
| East to West L3 VNI Unicast traffic ( POD A to POD A) | 0ms | 0ms |
| East to West L3 VNI Unicast traffic ( POD A to POD A') | 0ms | 0ms |
| East to West L2 VNI Multicast traffic (POD A and POD A') | 0ms | 0ms |
| North to Sourth Branch LISP Unicast traffic  - bi-directional | 0ms | 0ms |
| Legacy to North Unicast traffic | 0ms | 0ms |
| Legacy to PODA Unicast | 0ms | 0ms |
| Legacy to PODA' Unicast | 0ms | 0ms |
| Legacy to Legacy Unicast traffic | 0ms | 0ms |
| | | |
| Retstart NVE in N7K Spine | Low | High |
| East to West L2 VNI Unicast traffic | 0ms | 0ms |
| East to West L3 VNI Unicast traffic ( POD A to POD A) | 0ms | 0ms |
| East to West L3 VNI Unicast traffic ( POD A to POD A') | 0ms | 0ms |
| East to West L2 VNI Multicast traffic (POD A and POD A') | 0ms | 0ms |
| North to Sourth Branch LISP Unicast traffic  - bi-directional | 0ms | 0ms |
| Legacy to North Unicast traffic | 0ms | 0ms |
| Legacy to PODA Unicast | 0ms | 0ms |
| Legacy to PODA' Unicast | 0ms | 0ms |
| Legacy to Legacy Unicast traffic | 0ms | 0ms |
| | | |
| Retstart L2FM in N7K Spine | Low | High |
| East to West L2 VNI Unicast traffic | 0ms | 0ms |

| | | |
|---|---|---|
| East to West L3 VNI Unicast traffic ( POD A to POD A) | 0ms | 0ms |
| East to West L3 VNI Unicast traffic ( POD A to POD A') | 0ms | 0ms |
| East to West L2 VNI Multicast traffic (POD A and POD A') | 0ms | 0ms |
| North to Sourth Branch LISP Unicast traffic  - bi-directional | 0ms | 0ms |
| Legacy to North Unicast traffic | 0ms | 0ms |
| Legacy to PODA Unicast | 0ms | 0ms |
| Legacy to PODA' Unicast | 0ms | 0ms |
| Legacy to Legacy Unicast traffic | 0ms | 0ms |
| | | |
| | | |
| **Shutdown Peer link between N9K leafs** | **Low** | **High** |
| East to West L2 VNI Unicast traffic | 864ms | 4.7ec |
| East to West L3 VNI Unicast traffic ( POD A to POD A) | 849ms | 5.4sec |
| East to West L3 VNI Unicast traffic ( POD A to POD A') | 136ms | 2.7sec |
| East to West L2 VNI Multicast traffic (POD A and POD A') | 4.6sec | 4.6sec |
| North to Sourth Branch LISP Unicast traffic  - bi-directional | 432ms | 4.9sec |
| Legacy to North Unicast traffic | 0ms | 0ms |
| Legacy to PODA Unicast | 238ms | 3.3sec |
| Legacy to PODA' Unicast | 0ms | 0ms |
| Legacy to Legacy Unicast traffic | 0ms | 0ms |
| | | |
| **No Shutdown Peer link between N9K leafs** | **Low** | **High** |
| East to West L2 VNI Unicast traffic | 57ms | 8.5sec |
| East to West L3 VNI Unicast traffic ( POD A to POD A) | 756ms | 18sec |
| East to West L3 VNI Unicast traffic ( POD A to POD A') | 126ms | 10sec |
| East to West L2 VNI Multicast traffic (POD A and POD A') | 1sec | 1sec |
| North to Sourth Branch LISP Unicast traffic  - bi-directional | 38ms | 19sec |
| Legacy to North Unicast traffic | 0ms | 0ms |
| Legacy to PODA Unicast | 38ms | 12sec |
| Legacy to PODA' Unicast | 0ms | 0ms |
| Legacy to Legacy Unicast traffic | 0ms | 0ms |
| | | |
| **Shutdown all uplinks and Peer link on N9K leafs** | **Low** | **High** |
| East to West L2 VNI Unicast traffic | 712ms | 3.6sec |
| East to West L3 VNI Unicast traffic ( POD A to POD A) | 93ms | 3.8sec |
| East to West L3 VNI Unicast traffic ( POD A to POD A') | 56ms | 2sec |
| East to West L2 VNI Multicast traffic (POD A and POD A') | 3.7sec | 3.7sec |

| | | |
|---|---|---|
| North to Sourth Branch LISP Unicast traffic - bi-directional | 77ms | 2.4sec |
| Legacy to North Unicast traffic | 0ms | 0ms |
| Legacy to PODA Unicast | 77ms | 2.5sec |
| Legacy to PODA' Unicast | 0ms | 0ms |
| Legacy to Legacy Unicast traffic | 0ms | 0ms |
| | | |
| **No Shutdown all uplinks and Peer link on N9K leafs** | **Low** | **High** |
| East to West L2 VNI Unicast traffic | 219ms | 1.6sec |
| East to West L3 VNI Unicast traffic ( POD A to POD A) | 98ms | 4.2sec |
| East to West L3 VNI Unicast traffic ( POD A to POD A') | 165ms | 4.4sec |
| East to West L2 VNI Multicast traffic (POD A and POD A') | 848ms | 848ms |
| North to Sourth Branch LISP Unicast traffic - bi-directional | 73ms | 12sec |
| Legacy to North Unicast traffic | 0ms | 0ms |
| Legacy to PODA Unicast | 40ms | 8.2sec |
| Legacy to PODA' Unicast | 0ms | 0ms |
| Legacy to Legacy Unicast traffic | 0ms | 0ms |
| | | |
| **Reload N9K leafs (just after reload)** | **Low** | **High** |
| East to West L2 VNI Unicast traffic | 714ms | 5.7sec |
| East to West L3 VNI Unicast traffic ( POD A to POD A) | 104ms | 696ms |
| East to West L3 VNI Unicast traffic ( POD A to POD A') | 73ms | 602ms |
| East to West L2 VNI Multicast traffic (POD A and POD A') | 995ms | 995ms |
| North to Sourth Branch LISP Unicast traffic - bi-directional | 78ms | 574ms |
| Legacy to North Unicast traffic | 0ms | 0ms |
| Legacy to PODA Unicast | 93ms | 828ms |
| Legacy to PODA' Unicast | 0ms | 0ms |
| Legacy to Legacy Unicast traffic | 0ms | 0ms |
| | | |
| **When N9K comes up after reload** | **Low** | **High** |
| East to West L2 VNI Unicast traffic | 101ms | 12sec |
| East to West L3 VNI Unicast traffic ( POD A to POD A) | 4sec | 12sec |
| East to West L3 VNI Unicast traffic ( POD A to POD A') | 5.8sec | 7.7sec |
| East to West L2 VNI Multicast traffic (POD A and POD A') | 90ms | 90sec |
| North to Sourth Branch LISP Unicast traffic - bi-directional | 998ms | 13sec |
| Legacy to North Unicast traffic | 0ms | 0ms |
| Legacy to PODA Unicast | 1.5sec | 8.6sec |
| Legacy to PODA' Unicast | 0ms | 0ms |
| Legacy to Legacy Unicast traffic | 0ms | 0ms |

| Reload N7K Spine (just after reload) | Low | High |
|---|---|---|
| East to West L2 VNI Unicast traffic | 110ms | 1.9sec |
| East to West L3 VNI Unicast traffic ( POD A to POD A) | 123ms | 246ms |
| East to West L3 VNI Unicast traffic ( POD A to POD A') | 51ms | 259ms |
| East to West L2 VNI Multicast traffic (POD A and POD A') | 182ms | 182ms |
| North to Sourth Branch LISP Unicast traffic  - bi-directional | 16ms | 272ms |
| Legacy to North Unicast traffic | 207ms | 208ms |
| Legacy to PODA Unicast | 0ms | 2ms |
| Legacy to PODA' Unicast | 39ms | 18sec |
| Legacy to Legacy Unicast traffic | 0ms | 0ms |

## 4.2    MPLS Profile 1 with F3, M1 and M2 and Fabric Path and Classical Ethernet Access

The MPLS profile is performed in two access method technologies.
- Classical Ethernet: Core network, Site8 and Site 10. This network focuses on building and operating an MPLS core network that provides end-to-end L3VPN and MVPN services with the Nexus 7000 Sup2E and Nexus 7700 Sup2E between two different sites: Site 8 is mainly focused on legacy M1 modules, whereas Site 10 is built to test and verify MPLS feature set over the new F3 modules. Both sites use virtual Port-Channel (vPC) to deliver highly available unicast and multicast services.

- Fabric Path: In addition to the above topology, the new Site 6 has been integrated to the MPLS topology. This network uses the new F3 modules with both virtual Port-Channel (vPC+) and FabricPath solutions to deliver highly available unicast and multicast services. Furthermore, the deployment of FabricPath will increase the overall bandwidth (since vPC solution is limited to only two vPC peers) allowing to deploy multi-spine network layer, while eliminating the need for STP and allowing a "plug-and-play" type of network as the addition of new devices at either the spine of the leaf layer does not present any significant traffic loss for the pre-established flows.

### 4.2.1   MPLS Topology Design Review
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 7000 and N7700, with features such as vPC with different Nexus 7000 modules.

#### Core Network
The core network is built around two Nexus 7000 chassis with Sup 2E. Each chassis is configured to have two P routers in two different VDCs (for a total of four P routers).

- Nexus 7000 with distributed port-channels connected in a square topology
- CSR1kv are configured as a route-reflector cluster hence connected to two of the opposite P routers (P1 and P3)

### *Site 10 Network – Classical Ethernet F3*
This site is built around the Nexus 7700 with Sup 2E with F3 modules:

- Two Nexus 7700 with vPC acting as redundant multi-homing PEs, which are connected to two other Nexus 7700 for aggregation/access acting as CEs devices configured on different VDCs.

### *Site 8 Network – Classical Ethernet M1*
This site is built around the Nexus 7000 with Sup 2E with M1 and M1XL modules:

- Two Nexus 7700 with vPC acting as redundant multi-homing PEs which are connected to two Catalyst 6500 for aggregation/access acting as CEs devices

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.

### *Site 6 Network – Fabric Path*
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 N7700, with features such as Fabric Path and vPC+ with F3/40G Nexus modules

In Figure 1, Site 6 is the site that had been deployed to verify and test the FabricPath feature set on the Nexus 7700 with Sup 2E and F3/40G modules.

- Nexus 7700/F3-40G modules with vPC+ to Nexus 7000/M1 modules.
- Nexus 7700/F3-40G modules with vPC to Nexus 7700/F3-40G modules for FabricPath with different topologies.

### 4.2.2   Hardware and Software Overview

| Classical Ethernet - vPC | Model No. | Software |
|---|---|---|
| N7K | N7K SUP2E / F3 /M2 / M1XL | 7.2.0 |
| C6K | VS-SUP720-10G | 151-2.SY2 |
| CSR1000v | Na | 15.4.3.S1 |
| Fabric Path – vPC+ | Model No. | |
| N7700 | N7700 SUP2E / F3 40G | 7.2(0) |
| N7000 | N7K SUP1 / M1 | 7.2(0) |

### 4.2.3  Topology
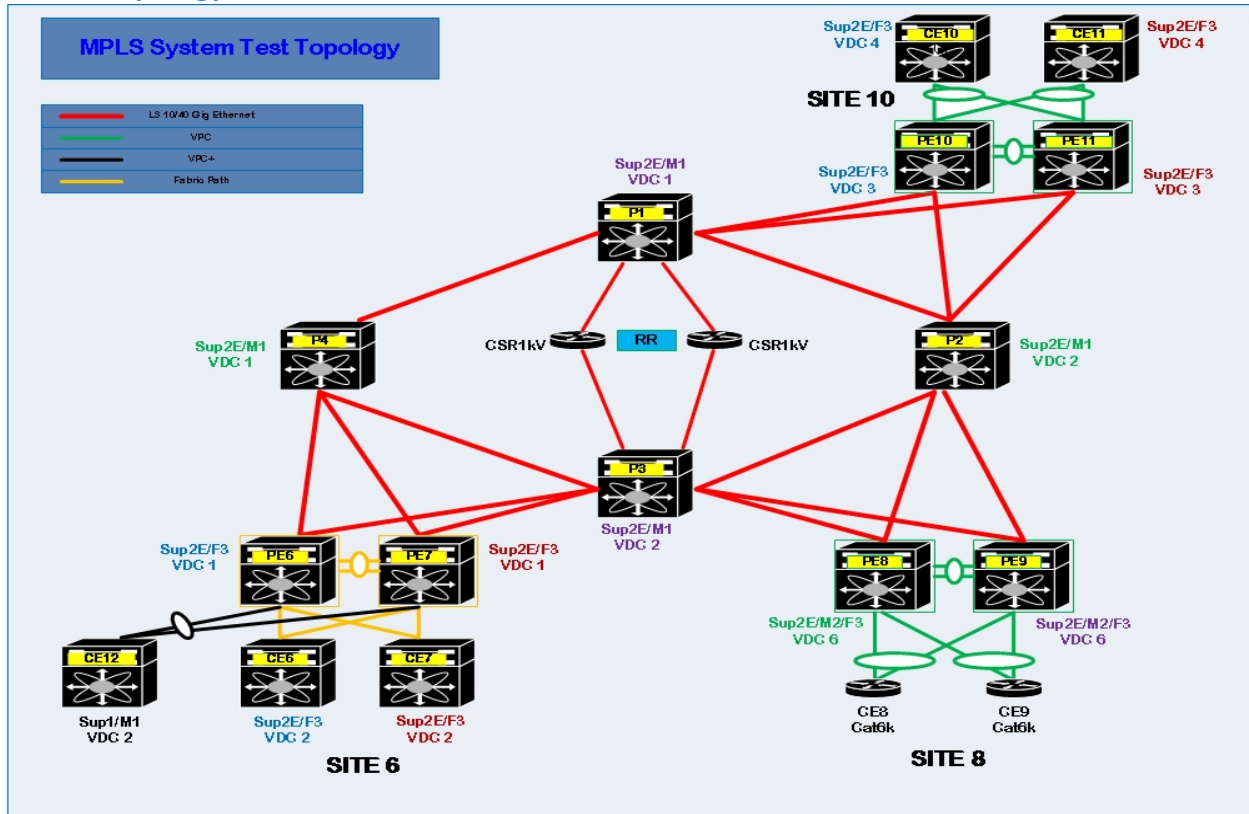


### 4.2.4  SP MPLS Scale Coverage

| | Phase 1 | | Phase 2 |
|---|---|---|---|
| Feature/Parameter | Site 10 (F3 vPC) | Site 8 (M1 vPC) | Site 6 (F3 FP/vPC+) |
| VLAN | 1000 | 1000 | 1000 |
| SVI | 1000 | 1000 | 1000 |
| IPv4 hosts x subnet | 32 | 32 | 32 |
| Hosts per VLAN | 47+ | 47+ | 47+ |
| vPC/(vPC+) | 20/na | 2/na | 2/(1) |
| VLANS x vPC/(vPC+) | 50/na | 500/na | 250/(500) |
| Fabric Path IS-IS adjacencies | na | na | 6 |
| Port-channel (excluding peer-link, vPC&vPC+) | 4 | 4 | 4 |
| HSRPv4 version2 | 1000 | 1000 | 1000 |
| HSRPv6 version2 | 1000 | 1000 | 1000 |
| VRF | 100 | 100 | 100 |
| VLANs per VRF | 10 | 10 | 10 |
| ARP x VRF | 350+ | 350+ | 350+ |
| mVRF | 50 | 50 | 50 |

| | | | |
|---|---|---|---|
| PIM neighbor default VRF | 4 | 4 | 4 |
| PIM neighbor x VRF | 36 | 36 | 36 |
| MDT default tunnel | 50 | 50 | 50 |
| MDT data tunnel | 200 | 200 | 200 |
| Number of local customer groups | 10 | 10 | 10 |
| Number of sources x customer group | 8 | 8 | 10 |
| Number of OIF x group | 10 | 10 | 10 |
| OSPF Peers default VRF | 4 | 4 | 4 |
| OSPF Routes/(Path) | 48/(54) | 48/(54) | 48/(54) |
| LDP Sessions | 4 | 4 | 4 |
| iBGP IPv4 Sessions | 2 | 2 | 2 |
| iBGP VPNv4 Sessions | 2 | 2 | 2 |
| iBGP VPNv6 Sessions | 2 | 2 | 2 |

### 4.2.5 SP MPLS ISSU Matrix

| ISSU/Cold boot | (SUP2E) |
|---|---|
| ISSU 6.2.10 -> 7.2.1 | Pass |
| Cold boot 6.2.10 -> 7.2.1 | Pass |
| ISSU 6.2.8b -> 7.2.1 | Pass |
| Cold boot 6.2.8b -> 7.2.1 | Pass |
| ISSU 6.2.12 -> 7.2.1 | Pass |
| Cold boot 6.2.12 -> 7.2.1 | Pass |

## 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 to deliver highly available unicast and multicast services and CSC 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   Hardware and Software Overview

|  | Model No. | Image Version |
|---|---|---|
| **N77K** | **N77-SUP2E/ N77-F324FQ-25** | 7.2 |
| **N7K** | **N7K-SUP2E/ N7K-F312FQ-25** | 7.2 |
| **N5K** | **N5K-C5548UP-SUP** | 7.0(5)N1(1) |
| **ASR9K** | **ASR9001-RP/ A9K-MPA-20X1GE** | 5.1.1 |
| **ASR1K** | **ASR1000-RP2/ ASR1000-ESP20/ SPA-8X1GE-V2** | 3.14 |

### 4.3.4 Topology



Solution Testing Topology
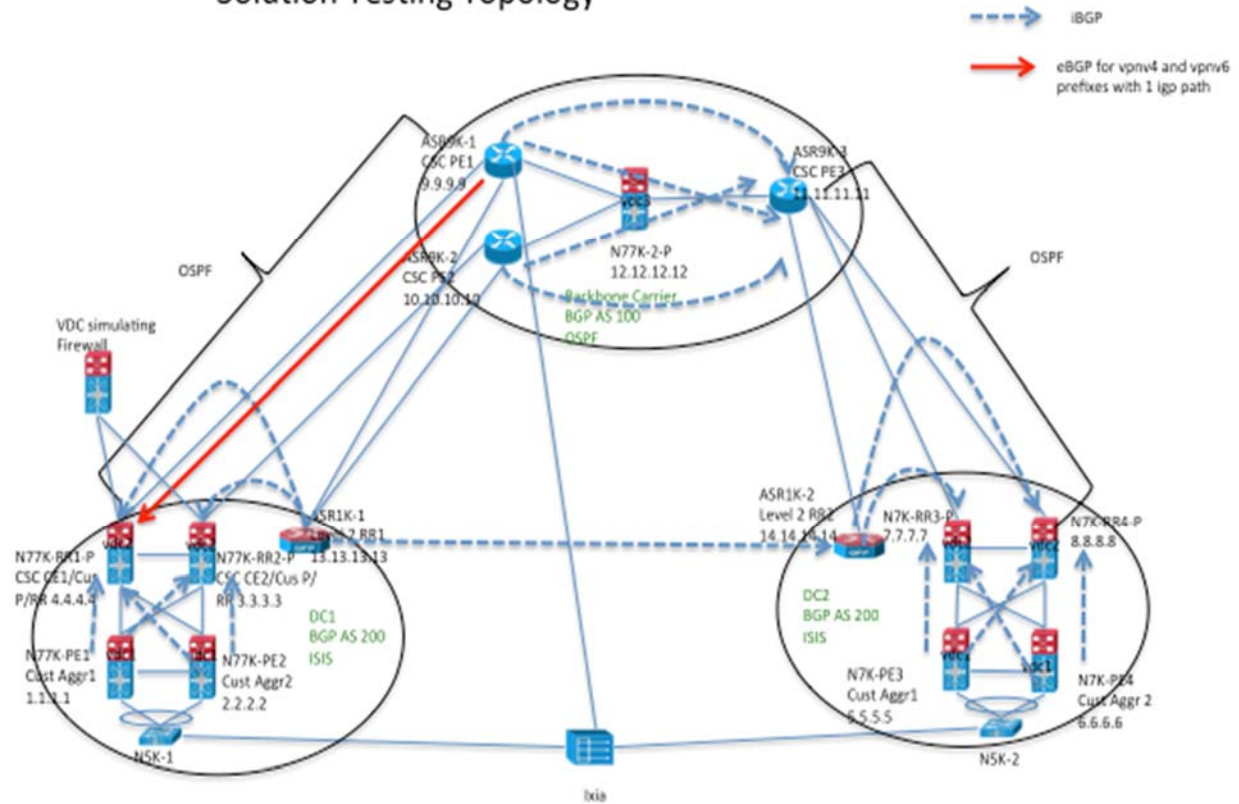
### 4.3.5 Scale

The scale numbers listed below are configured in the N77K-PE1 and N77K-PE2 (as shown in the Topology above) Aggregation switches.

| 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 |

| Feature | Scale |
|---|---|
| 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 |

### 4.3.6 ISSU Matrix

Since this is a new deployment, ISSU from 7.2 to 7.2 upgrade images was tested

### 4.3.7 Traffic and Convergence

As mentioned previously, the overall goal of this testing is to check the stability of the N77K as aggregation and RR cum inline core and measure the traffic convergence during various operational triggers. The traffic convergence measurements are in the attached spreadsheet. All the traffic convergence measurements are focused on N77K used as either aggregation switch or N77K as RR cum inline core. The traffic convergence metrics for each trigger are not consistent every time. So each trigger is repeated 3 to 4 times and the maximum value observed during any of the iteration is recorded as shown in the table below.

The following traffic flow metrics are configured and tested using Ixia. Fixed frame size of 100 is used for all streams.

| Traffic Streams | Description | Traffic Rate |
|---|---|---|
| V4 traffic bet hosts on diff DCs | Bidirectional Flow between 900 v4 hosts | 100K Frames/Sec |
| V6 traffic bet hosts on diff DCs | Bidirectional Flow between 200 v6 hosts | 100K Frames/Sec |
| V4 Scale traffic | Bidirectional flow for rest of the v4 prefixes from remote DC for scale as specified in Section 2 | 100K Frames/Sec |
| V6 Scale Traffic | Bidirectional flow for rest of the v6 prefixes | 100K Frames/Sec |

| | | |
|---|---|---|
| | from remote DC for scale as specified in Section 2 | |
| Multicast (mvpn)with N77K connected to Receivers on Ixia | 100 flows for 100 groups | 10K Frames/Sec |
| Multicast (mvpn) with N77K connected to Source on Ixia | 100 flows for 100 groups | 10K Frames/Sec |
| Global Multicast with N77K connected to Receivers on Ixia | 100 flows for 100 groups | 10K Frames/Sec |
| Global Multicast with N77K connected to Source on Ixia | 100 flows for 100 groups | 10K Frames/Sec |
| V4 Inter VRF traffic | Inter vrf traffic going through firewall within the DC | 200K Frames/Sec |
| V6 Inter VRF Traffic | Inter vrf traffic going through firewall within the DC | 200K Frames/Sec |

| Test Cases | VPNv4 Unicast | | VPNv6 Unicast | | VPNv4 Multicast | | Global v4 Muticast | | VPNv4 Thro Firewall | | VPNv6 Thro Firewall | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Convergence (msec) | Reconvergence(msec) | Convergence(msec) | Reconvergence(msec) | Convergence(msec) | Reconvergence(msec) | Convergence(msec) | Reconvergence(msec) | Converge nce(msec) | Reconvergence(msec) | Converge nce(msec) | Reconvergence(msec) |
| SSO on UUT N77K-PE1 | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A |
| Shut/no shut vpc keepalive port on N77K-PE1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Shut/no shut vpc peer link (port channel) on N77K-PE1 | 19.5 | 0 | 10.7 | 0 | 0 | 0 | 2536 | 816 | 16.6 | 0 | 16.6 | 0 |
| Shut/no shut vpc member link on N77K-PE1 | 90.8 | 477.7 | 38.5 | 451.8 | 210 | 2353.9 | 55 | 1452.9 | 45.6 | 903.5 | 45.6 | 903.5 |
| Shut/no shut the core link on N77K-PE1 connected to N77K-RR2-P | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 232.5 | 0 | 114.9 | 0 |
| Shut/no shut the member port on the L3 Port channel on N77K-PE1 | 22 | 5.9 | 20 | 6 | 33.7 | 17 | 32.4 | 8.9 | 0 | 0 | 0 | 0 |
| Shut/no shut the L3 port channel towards the core on N77K-PE1 | 2403 | 0 | 2405 | 0 | 4351 | 6498 | 2503 | 5231 | 0 | 0 | 0 | 0 |
| Shut/no shut the vpc member ports on the vPC peer N77K-PE2 | 31 | 0 | 12.6 | 0 | 0 | 0 | 0 | 0 | 21.5 | 0 | 21.5 | 0 |
| Shut/no shut the interface on N77K-RR1-P connected to CSC backbone | 1992.9 | 10012.9 | 1969 | 10042 | 1324.7 | 1872 | 1445 | 1172 | 10 | 0 | 11.5 | 0 |
| Shut/no shut the interface on N77K-RR2-P connected to CSC backbone | 220.7 | 2058 | 235.6 | 2105 | 757.5 | 884 | 1083 | 476 | 0 | 0 | 0 | 0 |
| Shut/no shut the interface on N77K-RR1-P connected to vPC peer PE N77K-PE2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Clear ip mroute * vrf all on PE, N77K-PE1 | N/A | N/A | N/A | N/A | 10691 | N/A | 8694.7 | N/A | N/A | N/A | N/A | N/A |
| Clear ip mroute * on N77K RR, N77K-RR1-P | N/A | N/A | N/A | N/A | 497 | N/A | 940 | N/A | N/A | N/A | N/A | N/A |
| Clear ip mroute * on N77K RR, N77K-RR2-P | N/A | N/A | N/A | N/A | 512 | N/A | 864.9 | N/A | N/A | N/A | N/A | N/A |
| Clear ip pim route *  vrf all on PE, N77K-PE1 | N/A | N/A | N/A | N/A | 20582 | N/A | 13209 | N/A | N/A | N/A | N/A | N/A |
| Clear ip pim route * on RR, N77K-RR1-P | N/A | N/A | N/A | N/A | 19041 | N/A | 22302 | N/A | N/A | N/A | N/A | N/A |
| Clear ip pim route * on RR, N77K-RR2-P | N/A | N/A | N/A | N/A | 15881 | N/A | 16060 | N/A | N/A | N/A | N/A | N/A |
| clear ip bgp * on UUT PE, N77K-PE1 | 5107 | N/A | 4572 | N/A | 3239 | N/A | 0 | N/A | 0 | N/A | 0 | N/A |
| clear ip bgp * on secondary RR, ASR1K-1 | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A |
| Clear ip bgp * on primary RR, N77K-RR1-P | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A |
| Clear ip bgp * on primary RR, N77K-RR2-P | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A |
| clear ip ospf neighbor * on primary RR, N77K-RR1-P | 5750 | N/A | 5684 | N/A | 6803 | N/A | 6192 | N/A | 9 | N/A | 10 | N/A |
| clear ip ospf neighbor * on primary RR, N77K-RR2-P | 3357 | N/A | 3346 | N/A | 5864 | N/A | 5041 | N/A | 0 | N/A | 0 | N/A |
| Restart BGP on N77K-PE1 | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A |
| Restart BGP on N77K-PE2 | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A |
| Restart BGP on N77K-RR1-P | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A |
| Restart BGP on N77K-RR2-P | | | | | | | | | | | | |
| Soft OIR of F3 card on PE, N77K-PE1 | 33705 | 33971 | 34065 | 32424.7 | 65367 | 46158 | 65310 | 47537 | 65310 | 8450 | 65310 | 8458 |
| Reload UUT PE, N77K-PE1 | 4266.9 | 14952.8 | 3960.6 | 20548 | 7301 | 32245 | 7143 | 9169.5 | 83 | 69136 | 82 | 58725 |
| Reload UUT PE in vpc sec role, N77K-PE2 | 50 | 3526 | 50 | 4886 | 0 | 12 | 47.5 | 5263 | 43.7 | 14054 | 43.7 | 13721 |
| Reload primary RR, N77K-RR1-P | 4835.9 | 9545 | 4855 | 9571.9 | 8634 | 868 | 5949 | 998.5 | 253 | 7760.5 | 215 | 6757.9 |
| Reload primary RR, N77K-RR2-P | 542 | 253.5 | 614 | 238.5 | 246 | 715 | 128 | 777 | 4919 | 66500 | 2479 | 67122.6 |
| Reload secondary RR, ASR1K-1 | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A | 0 | N/A |

CPU and memory utilization on the N77K-PE1 and N77K-PE2 measured under steady state are shown below. The steady state measurements are done when the system is stable with all the features configured and 100% end to end traffic running. SNMP polling of MPLS MIBs are done continuously in the background.

| Device | Average CPU Utilization (5 min with 16 core CPU) | Memory Utilization |
|---|---|---|
| N77K-PE1 | 1.49 | 6738104K |
| N77K-PE2 | 0.51 | 7957180K |

Due to the redistribution of the ISIS routes into OSPF and vice-versa towards the CSC backbone, routing loops are bound to happen during certain triggers like link failure, vdc reload etc. To avoid this, it is recommended to have route-maps attached such that, for locally learnt routes ISIS has higher priority and for remote routes, OSPF has higher priority.

VDC reload of the PEs and P/RR causes traffic black hole in the order of seconds. This is due to the following:
1> BGP timers default value is 60 sec with holdtime of 180 sec.
2> When the core in the best path is reloaded, the old route on the PE sticks around, with "inaccessible" when the interface is down.   As soon as the restarting router's interface is up, the PE's RIB sends a notification to BGP about the nexthop being "up".  The assumption is that this is coming from the BFD on the interface (not the BFD under BGP).  As a result, that route on PE is re-installed as the best path. This causes a traffic blackhole.

## 4.4   Dynamic Fabric Automation (Vinci)

### 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 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(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) |

### 4.4.3   Topology



SSTE Vinci-Fabric Encapsulation Test Topology
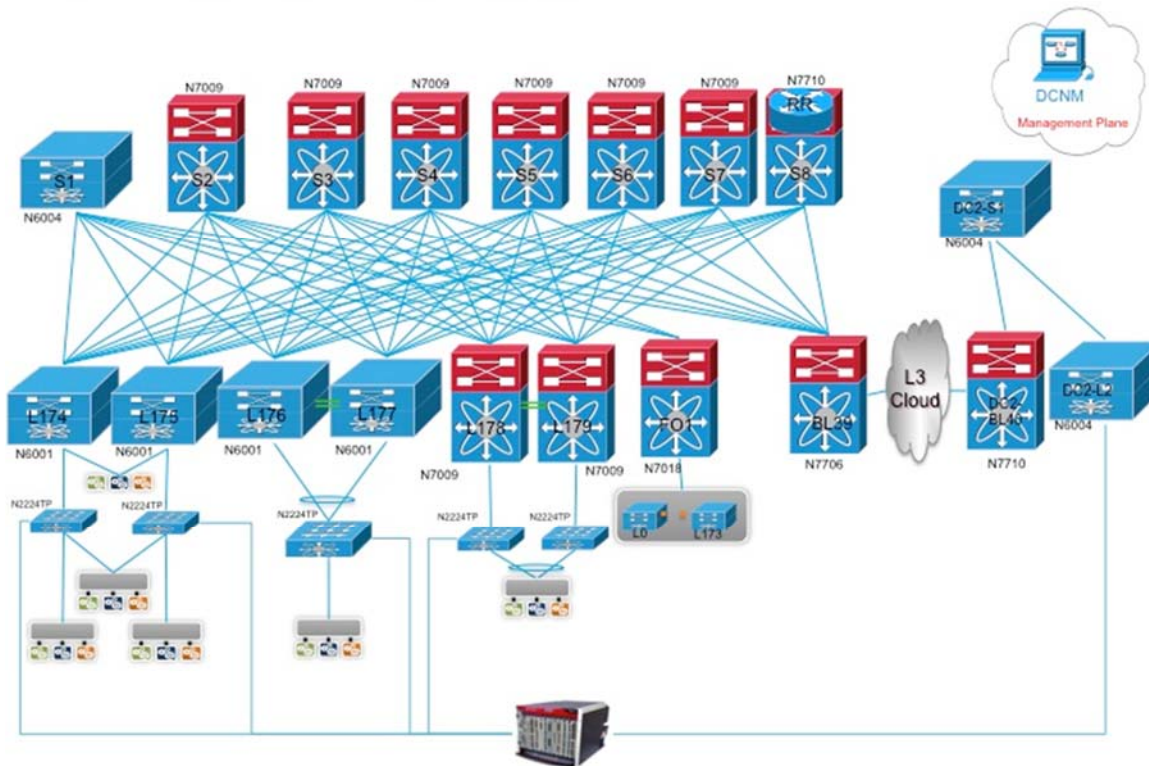
### 4.4.4 Traffic and Metrics

| IPv4 Unicast Traffic | Across Different Leaf/Across Border Leaf |
|---|---|
| IPv6 Unicast Traffic | Across Different Leaf/Across Border Leaf |
| IPv4 Multicast Traffic | Across Different Leaf/Across Border Leaf |

| Device | Average CPU Utilization/Load Average 15mins | Memory Utilization |
|---|---|---|
| Leaf – Primary vPC | 0.34 | Memory usage: 5965956K used |
| Leaf – Secondary vPC | 0.25 | Memory usage: 5967084K used |
| Spine | 0.36 | Memory usage: 5757104K used |
| BL | 0.43 | Memory usage : 6308256K used |

## 4.5 DC1

### 4.5.1 Hardware and Software Overview

| 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 |  |

### 4.5.2 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 5548 with features such as vPC and Fabric-Path.

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

### 4.5.3   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 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 comprises 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

## 4.5.3.1 DC1 Topology
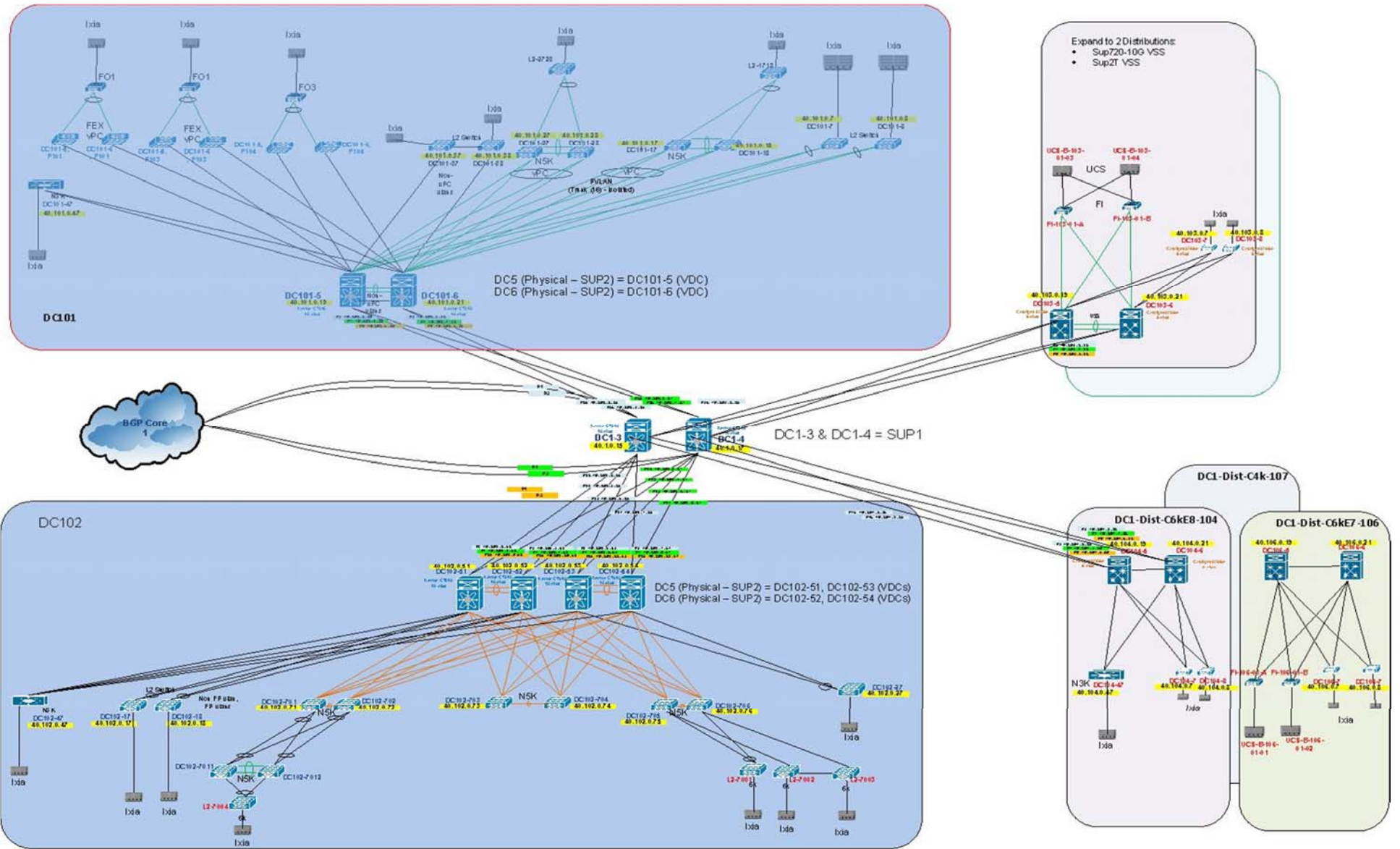
## 4.6 ENT

### 4.6.1 Hardware and Software Overview

| Platform | Model No. | Software |
|---|---|---|
| N7000 | SUP2E, M2 | 7.2.0 |
| N7700 | SUP2E, F2E, F3(40G with breakout) | 7.2.0 |
| 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 |

### 4.6.2 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 |

## 4.7 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 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.

## 4.8 ENT Topology
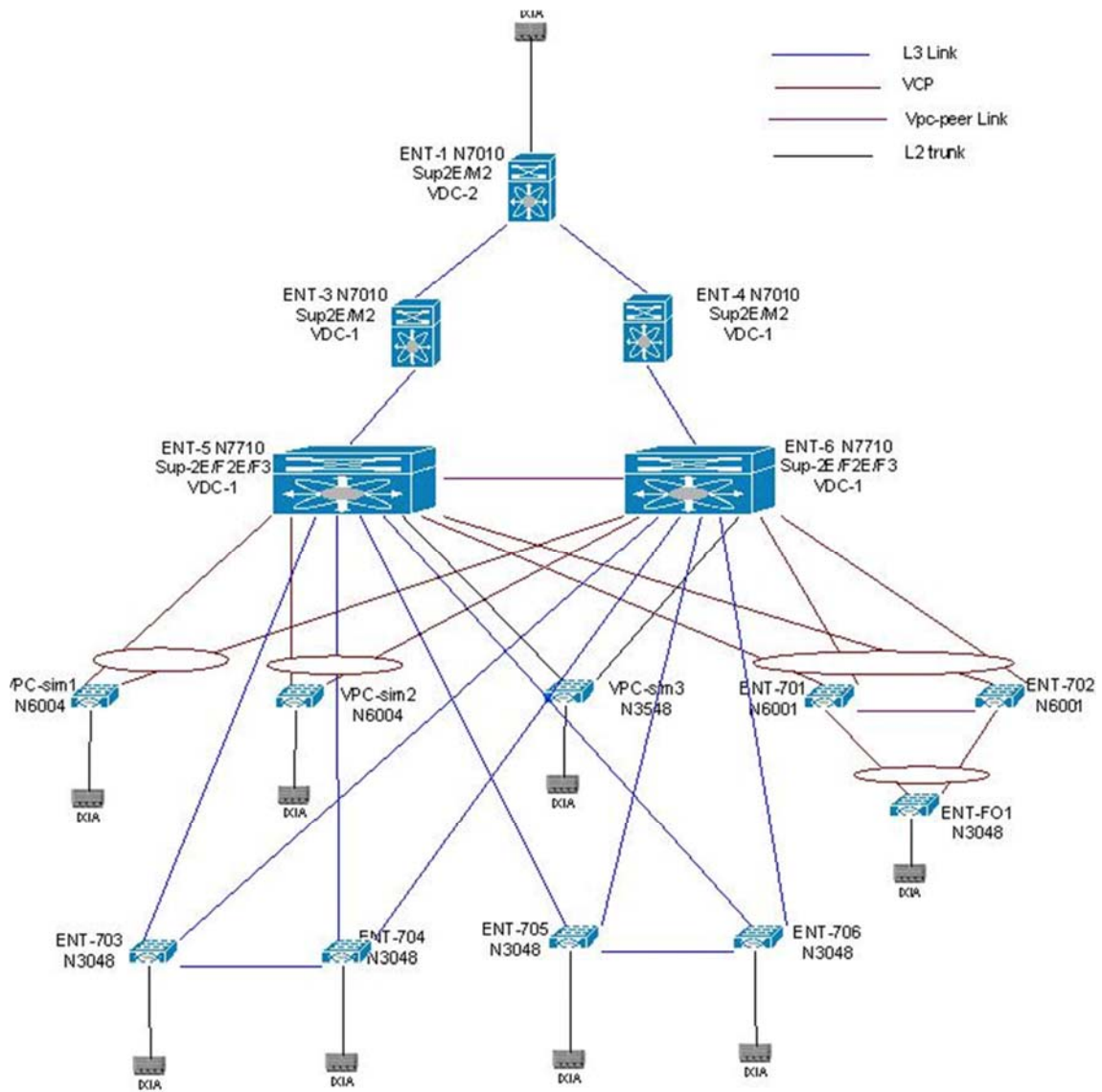
## 4.9  MSDC Scale

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

### 4.9.1  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 |

EDCS-963802 Profile 1, 2-1
Logical Topology
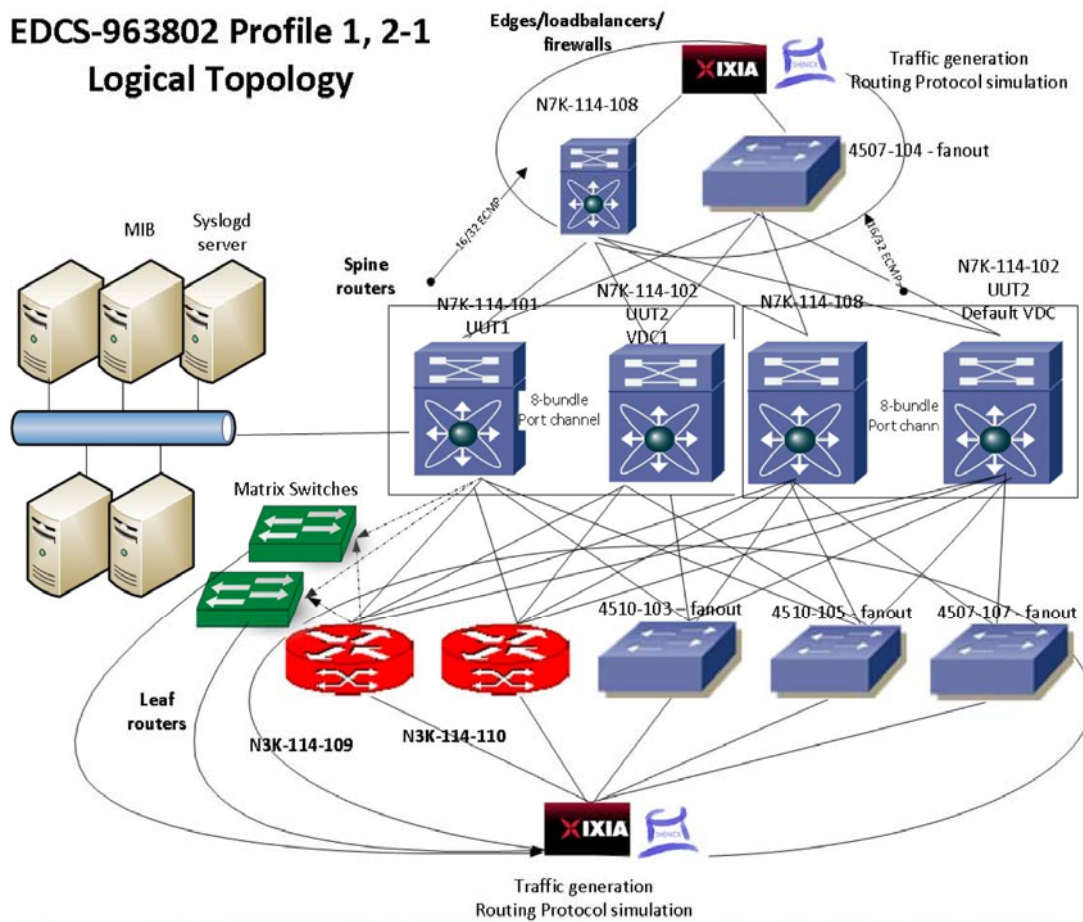
### 4.9.2   ISSU Matrix

| ISSU/Cold boot | |
|---|---|
| ISSU 6.2.8a -> 7.2 | |
| ISSU 6.2.10 -> 7.2 | Pass |
| ISSU 6.2.12 -> 7.2 | Pass |
| ISSU 7.2 -> upg | Pass |

## 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
  - ➤ Private VLAN host on classic port-channel.
  - ➤ Private VLAN promiscuous trunk, secondary trunk on classic Port-channel

## 4.11 M1 vPC Scale Topology
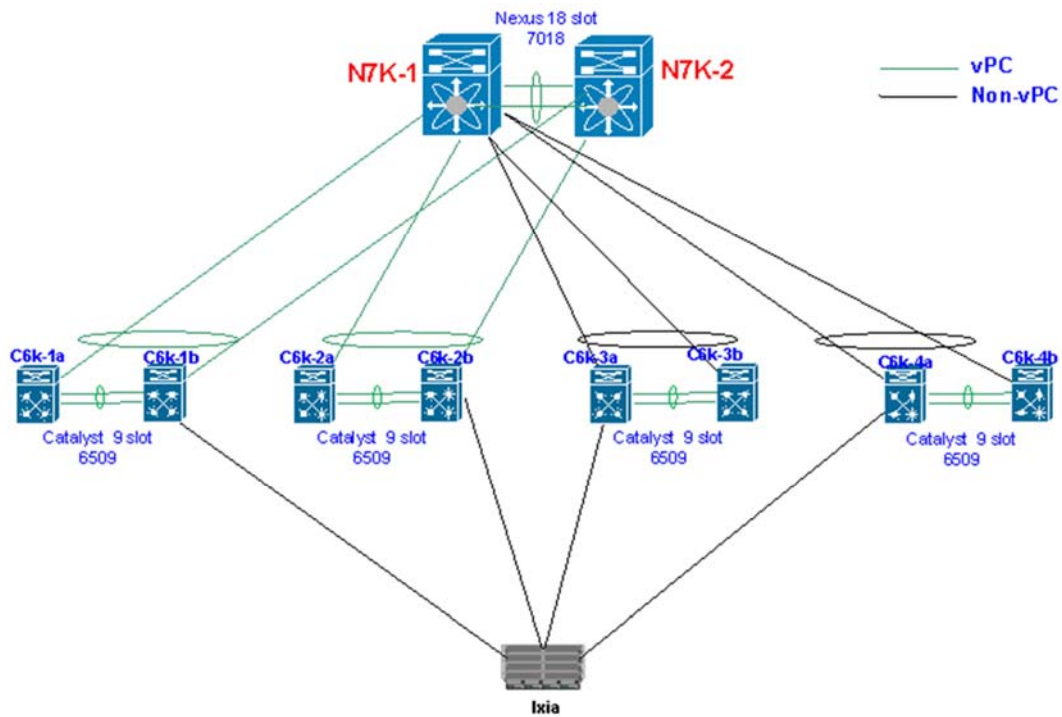
## 4.12 DC1, ENT1 and M1 vPC Feature / Scale Coverage

| 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 | |

## 4.13  ISSU Matrix for DC1, ENT and M1 vPC

| Image | ISSU/COLD BOOT | DC1 | ENT | M1vPC |
|---|---|---|---|---|
| 7.2.0 >7.2.0.upg | ISSU | PASS | PASS | PASS |
| 6.2.12 > 7.2.0 | ISSU | PASS | PASS | |
| 6.2.10 > 7.2.0 | ISSU | | PASS | |
| 6.2.8a > 7.2.0 | ISSU | | PASS | |
| 6.2.8b > 7.2.0 | ISSU | | PASS | |
| 7.2.0 > 7.2.0.upg | COLD BOOT | | PASS | |
| 6.2.12 > 7.2.0 | COLD BOOT | | PASS | |
| 6.2.10 > 7.2.0 | COLD BOOT | | PASS | |
| 6.2.8a > 7.2.0 | COLD BOOT | | PASS | |
| 6.2.8b > 7.2.0 | COLD BOOT | PASS | PASS | PASS |
| 7.2.0 > 6.2.12 | COLD BOOT | PASS | PASS | |
| 7.2.0 > 6.2.10 | COLD BOOT | | PASS | |
| 7.2.0 > 6.2.8a | COLD BOOT | | PASS | |
| 7.2.0 > 6.2.8b | COLD BOOT | PASS | PASS | |

# 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* |

CSCus00094
**Symptom:** IPFIB process might crash upon enabling LDP autoconfig feature.
**Conditions:**  while running the system image 6.2(10) and in presence of F3 modules with MPLS feature set enabled (configuration not supported currently), the IPFIB process might crash upon enabling LDP autoconfig.
**Workaround: None**, deploy 7.2(0) instead.
**Severity:** Severe
**Status:** Closed
**Platform Seen:** Nexus 7700
**Resolved Releases:**
**Applicable Releases:** 6.2(10)


CSCus00165
**Symptom：**  In a N7700 system running 6.2(10), F3 modules might fail to come online due to "*SAC_INTR: Xbar ASIC interrupt occurred: SAC Chico Interrupt Detected*"
**Conditions:**  Under the same conditions of CSCus00094,  upon the IPFIB process crash, all the F3 modules might fail to come online.
**Workaround: None**, deploy 7.2(0) instead.
**Severity:** Severe
**Status:** Closed
**Platform Seen:** Nexus 7700
**Resolved Releases:**
**Applicable Releases:** 6.2(10)


CSCus02115
**Symptom:** M1 failure upon upgrading the system image from 6.2(10) to a 7.2(0) daily image.
**Conditions:**  while upgrading the system image from 6.2(10) to an interim 7.2(0) build, one of the M1 modules failed the bootup sequence.
**Workaround: reload failed module**.
**Severity:** Severe
**Status:** Unreproducible
**Platform Seen:** Nexus 7000
**Resolved Releases:**
**Applicable Releases:** 7.2(0)


CSCus22635
**Symptom:** Some time, the configuration of the IPv6 VIP might not succeed.
**Conditions:** The problem has occurred only ones and it could not be recreated anymore.
**Workaround:** reload the system
**Severity:** Severe
**Status:** Unreproducible
**Platform Seen:** Nexus 7000
**Resolved Releases:**
**Applicable Releases:** 7.2(0)

CSCus31113
**Symptom:** Some times, in a scaled environment, MP-BGP routes might fail to be properly installed in the RIB for one or more VRFs.
**Conditions:** Sometimes, upon system reload in a scaled environment, some routes or some mroutes might not get properly installed in the shared memory.
**Workaround:** **change the setting for the** *"limit-resources u4route-mem minimum NNN maximum NNN"* **and then perform a system reload or a system switchover in order to ensure the new value is properly parsed. To estimate the proper value for the memory allocation, the following command can be used** *"show routing <ip|ipv4|ipv6|multicast> estimate memory estimate route <# of routes|mroutes>"*
**Severity:** Minor
**Status:** Assigned
**Platform Seen:** Nexus 7000
**Resolved Releases:**
**Applicable Releases:**


CSCus50951
**Symptom:** Some times, in a scaled environment, MP-BGP routes might fail to be properly installed in the RIB for one or more VRFs.
**Conditions:** Sometimes, upon system reload in a scaled environment with many host routes, some routes (or some mroutes) might not get properly installed in the shared memory.
**Workaround:** **Restart BGP process or reload the system**
**Severity:** Severe
**Status:** Unreproducible
**Platform Seen:** Nexus 7700
**Resolved Releases:**
**Applicable Releases:**


CSCut60771
**Symptom:** In a MVPN site, multicast traffic sourced locally might get continuously duplicated for the receivers attached to the primary vPC orphan port.
**Conditions:** In a given MVPN site with both local sources and receivers, multicast traffic might get continuously duplicated for all the local groups for the receivers attached to the vPC primary orphan port.
**Workaround: None**
**Severity:** Severe
**Status:** Unreproducible
**Platform Seen:** Nexus 7000
**Resolved Releases:**
**Applicable Releases: 7.2(0)**


CSCuu02810:
Symptom: vlan_mgr timedout during system bringup with copy r s and reload
Conditions:  After copy r & s, reload of system, vdc, results in vlan_mgr timeout errors
Workaround: none
Severity: Severe
Status: Resolved
Platform: Nexus 7000
Resolved Releases: 7.2.0
Applicable Releases: 7.2.0


 CSCuu27998:

Symptom: ("Invalid VLAN ID") communicating with MTS_SAP_VLAN_MGR
Conditions:  Invalid native vlan allocation during a vdc reload.
Workaround: none
Severity: Moderate
Status: Assigned
Platform: Nexus 7000
Resolved Releases:
Applicable Releases: 7.2.0

CSCuu38691:
Symptom: ISSU failed due to /nxos/tmp space exhaustion with logging
Conditions: ISSU fails before pre-check and aborts without doing the pre-check
Workaround: none
Severity: Moderate
Status: Assigned
Platform: Nexus 7000
Resolved Releases:
Applicable Releases: 7.2.0


CSCus21952
Symptom: PVLAN traffic block-hole during COLD BOOT from 7.2.0.512.bin to 6.2.10.bin
Condition:  L3 egress traffic from Trunk secondary port was not working due to dummy bd
configuration for primary vlan.
Workaround :  Use 6.2.12 instead of 6.2.10
Severity:
Status : Resolved and verified in 6.2.12 (fix not available in 6.2.10)
Platform: Nexus 7000
Resolved Release : 6.2.12
Applicable Release :


CSCuq47581
Symptom: VPC Peer-link flap upon ISSU to 7.2.0
Condition:  Object tracking feature in place
Workaround:  Remove object tracking while perform ISSU and add it back
Severity:
Status: CLOSED
Platform: Nexus 7000
Resolved Release:
Applicable Release: 7.2.0

CSCus71454
Symptom: VPC primary leg flap upon on VPC peer link flap
Condition: VPC leg configured as Private VLAN host mode
Workaround: No
Severity: 2
Status: A
Platform: Nexus 7000
Resolved Release:
Applicable Release: 7.2.0

Symptom: FEX HIF configuration miss while COLD BOOT from 7.2.0 to 6.2.x
Condition: COLD BOOT from 7.2.0 to 6.2.x
Workaround:  Re-apply HIF configuration after FEX come online
Severity:
Status:
Platform: Nexus 7000
Resolved Release:
Applicable Release:


CSCuu39696
Symptom: Removing pvlan host-association from one VPC PO flaps all VPC PO
Condition: VPC Po configured as community host
Workaround:  No
Severity: 2
Status: A
Platform: Nexus 7000
Resolved Release:
Applicable Release: 6.2.12 & 7.2.0


# 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
- o   Feature set requirement
  - o   Install feature-set fabric      # Installed feature-set fabric
  - o   Feature-set fabric           # Enable/Disable fabric
  - o   feature bgp                 # Enable/Disable Border Gateway Protocol
  - o   feature lisp                #Enable/Disable Locator/ID Separation Protocol
  - o   feature pim                 #Enable/Disable Protocol Independent Multicast
  - o   feature ospf                # Enable/Disable Open Shortest Path First Protocol
  - o   feature interface-vlan       #Enable/Disable interface vlan
  - o   feature lacp                #Enable/Disable LACP
  - o   feature bfd                  # Enable/Disable BFD
  - o   feature nv overlay           # Enable/Disable NV Overlay
  - o   feature vni                      # Enable/Disable Virtual Network Segment (VNI)
- o   Global configuration command
  - o   nv overlay evpn            # Enable/Disable Ethernet VPN (EVPN)
  - o   system bridge-domain       #Configure Bridge-domain ID
  - o   vni <range>                # configure VNI range
  - o   ip pim rp-address <>         # Configure static RP for group range for BUM
                                       Traffic
  - o   bridge-domain <>           # configure Bridge-domain ID
- o   VRF Configuration

- o vrf context <>      # Create VRF and enter VRF mode
- o Vni <>       # configure L3 VNI for VXLAN
- o rd auto      # Generate RD automatically
- o address-family ipv4 unicast # Configure IPv4 address family
- o route-target import <>  # Import Target-VPN community
- o route-target import <> evpn # Specify Target for EVPN routes
- o route-target export <>  # Export Target-VPN community
- o route-target export <> evpn # Specify Target for EVPN routes
- o route-target both auto  # Export And Import Target-VPN community
- o route-target both auto evpn # Specify Target for EVPN routes

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



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

- o Port-channel interface ( for all L3 links )

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

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

- o BGP Global configuration
  - o router bgp 1      # Border Gateway Protocol (BGP)
  - o address-family ipv4 unicast  # Configure unicast address-family
  - o network <loopback >   # Configure an IP prefix to advertise
  - o redistribute direct route-map <passall> # Directly connected
  - o maximum-paths <>    # Number of parallel paths (EBGP)
  - o maximum-paths ibgp <>   # Number of parallel paths (IBGP)

- o address-family l2vpn evpn        # Configure L2VPN EVPN address-family
- o route-map <passall> permit 10      # Permit any
- o vrf <>                    # Production VRF
- o  address-family ipv4 unicast     # Configure unicast address-family
- o  network 0.0.0.0/0            # Default route
- o  advertise l2vpn evpn          # advertise l2vpn evpn routes
- o  redistribute direct route-map passall  # Directly connected


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


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


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


- o EBGP (Inter-POD Spine) – Suggestion to configure IPv4 over interface and EVPN peering over loopback
  - o template peer EBGP-IPv4       # Template configuration for peer parameters
  - o bfd                  # Bidirectional Fast Detection for the neighbor

  - o remote-as 10             # Specify AS Number of the neighbor
  - o password 3 <>            #Configure a password for neighbor
  - o address-family ipv4 unicast    # Configure unicast address-family
  - o send-community extended      # Send Community attribute to this neighbor
  - o 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

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

- o route-map LOCAL_PREF permit 10 #route-map name
- o set local-preference 50     #BGP local preference path attribute ( configure
  less

      than default 100)

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

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

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

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


## *LISP Configuration on Spine (N7K)*

- o vrf context <>      # Production VRF
- o ip lisp etr       #Configures LISP Egress Tunnel Router
  (ETR)
- o lisp instance-id 1000    #Configures Instance-ID for global data-
  mappings
- o ip lisp itr map-resolver <ip add>  #To interact with Map-Resolver
  (primary)
- o ip lisp itr map-resolver <ip add>  #To interact with Map-Resolver
  (secondary)
- o ip lisp etr map-server <ip add>key <> # Authentication key used with Map-
  Server (Primary)
- o ip lisp etr map-server <ip add>key <> # Authentication key used with Map-

  Server (secondary)
- o lisp dynamic-eid 101    # Configure dynamic-EIDs for roaming
- o database-mapping <Mobility network> < core link address>  priority 1
  weight

#configure EID-prefix and locator-set for dynamic-EID

- o 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)*

- o router lisp                              #Locator/ID Separation Protocol
- o ddt authoritative instance-id 1000 <Subnet> # ddt authoritative to allow (VXLAN

                                        Mobility subnet
- o map-server-peer <ip address> # IPv4 Peer map-server locator address (secondary)
- o map-server-peer <ip address> # IPv4 Peer map-server locator address (Primary)


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

                                                        subnet
- o eid-prefix instance-id 1000 <subnet> accept-more-specifics # allow Branch subnet



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


## *LISP configuration on Branch ( ASR1K)*
- o router lisp                     # Locator/ID Separation Protocol
- o eid-table default instance-id 1000 #Configures Instance-ID for global data-
                                mappings

- o database-mapping <Mobility network> < core link address>  priority 1 weight
                          #configure EID-prefix and locator-set for dynamic-EID

- o ipv4 itr map-resolver <ip address> #To interact with Map-Resolver (primary)
- o ipv4 itr map-resolver <ip address> #To interact with Map-Resolver (secondary)
- o ipv4 itr  #  Configures a LISP Ingress Tunnel Router (ITR)
- o ipv4 etr map-server <ip address> key <key>#Authentication key used with Map-
                                        Server (Primary)
- o ipv4 etr map-server <ip address> key <key>#Authentication key used with Map-
                                        Server (secondary )

- o ipv4 etr                  # Configures a LISP Egress Tunnel Router (ETR)

- o ipv4 map-cache-limit <>      # Address family specific map cache
  configuration

- o Feature set requirement
    - o feature bgp            # Enable/Disable Border Gateway Protocol
    - o feature pim            #Enable/Disable Protocol Independent Multicast
    - o feature interface-vlan    #Enable/Disable interface vlan
    - o feature lacp            #Enable/Disable LACP
    - o feature bfd            # Enable/Disable BFD
    - o feature nv overlay       # Enable/Disable NV Overlay
    - o feature vni            # Enable/Disable Virtual Network Segment (VNI)
    - o feature vn-segment-vlan-based # Enable/Disable VLAN based VN segment
    - o feature vpc            # Enable/Disable VPC (Virtual Port Channel)
- o Global configuration command
    - o nv overlay evpn       # Enable/Disable Ethernet VPN (EVPN)
    - o ip pim rp-address <>     # Configure static RP for group range for BUM
                         Traffic
    - o fabric forwarding anycast-gateway-mac <>   # Anycast Gateway MAC of the Switch

- o Vlan to vn-segment mapping configuration command (Note it is require for each L2
  VLAN and L3 VLAN per VRF)
    - o vlan <>            # Vlan commands
    - o vn-segment <>       # VN Segment id of the VLAN

- o VRF Configuration
    - o vrf context <>         # Create VRF and enter VRF mode
    - o Vni <>            # configure L3 VNI for VXLAN
    - o rd auto            # Generate RD automatically
    - o address-family ipv4 unicast   # Configure IPv4 address family
    - o route-target import <>     # Import Target-VPN community
    - o route-target import <> evpn # Specify Target for EVPN routes
    - o route-target export <>     # Export Target-VPN community
    - o route-target export <> evpn   # Specify Target for EVPN routes
    - o route-target both auto     # Export And Import Target-VPN community
    - o route-target both auto evpn   # Specify Target for EVPN routes

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

- o VPC domain <>

    - o VPN domain         # Specify domain
    - o peer-switch          # Enable peer switch on vPC pair switches

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

<br>

- o VPC Keepalive link

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

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

- o VPV leg port-channel (host connected link)

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

- o Vlan interface configuration

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

- o Vlan interface configuration mapped to L3 VNI

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

- o Port-channel interface ( for all L3 links )

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


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

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


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

    - o template peer IBGP-IPv4              # Template configuration for peer parameters

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


- o   template peer IBGP-EVPN    # Template configuration for peer parameters
- o     bfd                     # Bidirectional Fast Detection for the neighbor
- o     remote-as 1           # Specify AS Number of the neighbor
- o     password 3 <>       #Configure a password for neighbor
- o     update-source loopback0   # Specify source of BGP session and updates
- o     address-family l2vpn evpn   # Configure L2VPN EVPN address-family
- o      send-community extended   #Send Standard and Extended Community attributes
- o   neighbor <> remote-as 1     # neighbor IBGP peering
- o   inherit peer IBGP-EVPN     # Inherit a template
- o IBGP inter vPC SVI link over peer-link (Only for IPv4 peering) and add vlan id to peer link allow list


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


- o   router bgp <>          #Border Gateway Protocol (BGP)
- o   address-family ipv4 unicast   # Configure unicast address-family
- o   network <>           # Add SVI interface address
- o   neighbor <> remote-as 1     # neighbor IBGP peering
- o   bfd                   # Bidirectional Fast Detection for the neighbor
- o   password 3 <>        # Configure a password for neighbor
- o   update-source <svi id)     #Specify source of BGP session and updates
- o   address-family ipv4 unicast   # Configure unicast address-family
- o   send-community extended    #Send Standard and Extended Community attributes
- o   route-reflector-client      # Configure a neighbor as Route reflector client
- o   route-map NHS  out      #Apply route-map to neighbor


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

- o  set ip next-hop peer-address        # Use peer address (for BGP only)


- o  EVPN Configuration for each L2 VNI
  - o  Evpn                              # Enter EVPN configuration mode
  - o  vni <> l2                         # Configure Ethernet VPN ID
  - o   rd auto                          # VPN Route Distinguisher Auto
  - o   route-target import auto         # Import Target-VPN community and Generate
                                          RT automatically
  - o   route-target import <>           # RT extcommunity in aa:nn format
  - o   route-target export auto         # Export Target-VPN community and Generate
                                          RT automatically
  - o   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.
  - o  feature tacacs+                              *# enabling the tacacs feature*
  - o  tacacs server host <ip address> key <0/7>    *# configure the tacacs server to authenticate  users*
  - o  aaa group server tacacs+ <group name>        *# enable server groups for redundancy*
    - ▪ server <ip address>
    - ▪ use-vrf <vrf_name>                           *# use-vrf based on server reachability*
  - o  snmp-server user <user-name> <group-name> auth md5 <pass-phrase> priv <pass-phrase> localizedkey                          *# snmp v3 user with authentication enabled*
  - o  ntp server <ip address>                       *# enable ntp with server ip address*
  - o  ip domain-name <domain name>                  *# enable domain-name*
  - o  interface mgmt0                               *# configure mgmt0*
    - ▪ vrf member management
    - ▪ ip address <ip_address >
  - o  power redundancy-mode ps-redundant mode       *# set the power in redundancy*
  - o  no system admin-vdc                           *# disable system admin-vdc*
  - o  install feature-set fabricpath                *# enable feature-set for fabricpath*
  - o  install feature-set mpls                      *# enable feature-set for mpls*
  - o  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.*

o 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

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

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

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

o 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>

- o 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*

- o 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>

- o 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*

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

- o 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)*

- o 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.
  - o feature-set fabricpath                                *# configure feature-set fabricpath*
  - o vlan <vlan-range>
    - ▪ mode fabricpath                                *# configure vlan-range in fabric path*
  - o fabricpath switch-id <switch-id>                       *# configure switch-id*
  - o vpc domain <domain-id>                                 *# configure vpc domain-id*
    - ▪ fabricpath switch-id <vpc+_switch-id>          *# configure virtual switch for peers present on the network*
  - o 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.
  - o 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.
  - o interface Vlan<id>                                     *# configure svi*
  - o ip access-group <acl> in                              *# enable access-list*
  - o ip access-group <acl> out
  - o no ip redirects
  - o ip address <ip address>
  - o hsrp version 2
  - o 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.
  - o ip igmp snooping                                      *# by default enabled on Nexus*

- LACP: LACP is used for link aggregation to form port-channels across the network.
  - o 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
  - o feature udld                                          *# enable feature udld*

- o 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

  - o *feature private-vlan*      *# enable feature private-vlan*
  - o *vlan <vlan-id>*      *# configure primary vlan*
    - *private-vlan primary*
  - o *vlan <vlan-id>*
    - *private-vlan <isolated/community>*    *# configure secondary vlan*
    - *private-vlan association <vlan-id>*    *# configure association with primary vlan*
  - o *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>*

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

  - o *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.
  - o feature tacacs+      *# enabling the tacacs feature*
  - o tacacs server host <ip address> key <0/7>    *# configure the tacacs server to authenticate  users*
  - o 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.

- o Private VLAN promiscuous trunk, secondary trunk on classic Port-channel

- o feature private-vlan                                              *# enable feature private-vlan*
- o vlan <vlan-id>                                                       *# configure primary vlan*
    - ▪ private-vlan primary
- o vlan <vlan-id>
    - ▪ private-vlan <isolated/community>     *# configure secondary vlan*
    - ▪ private-vlan association <vlan-id>      *# configure association with primary vlan*
- o 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>
- o 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*
- o 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*
- o 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*