Design and Automate VXLAN BGP EVPN Fabric with NDFC

Parth Patel, Technical Leader,
Technical Marketing Engineer
Data Center and Provider Connectivity BU
BRKDCN-2918

Cisco Webex App

Questions?
Use Cisco Webex App to chat
with the speaker after the session

How

1. Find this session in the Cisco Live Mobile App

2. Click "Join the Discussion"

3. Install the Webex App or go directly to the Webex space

4. Enter messages/questions in the Webex space

Webex spaces will be moderated
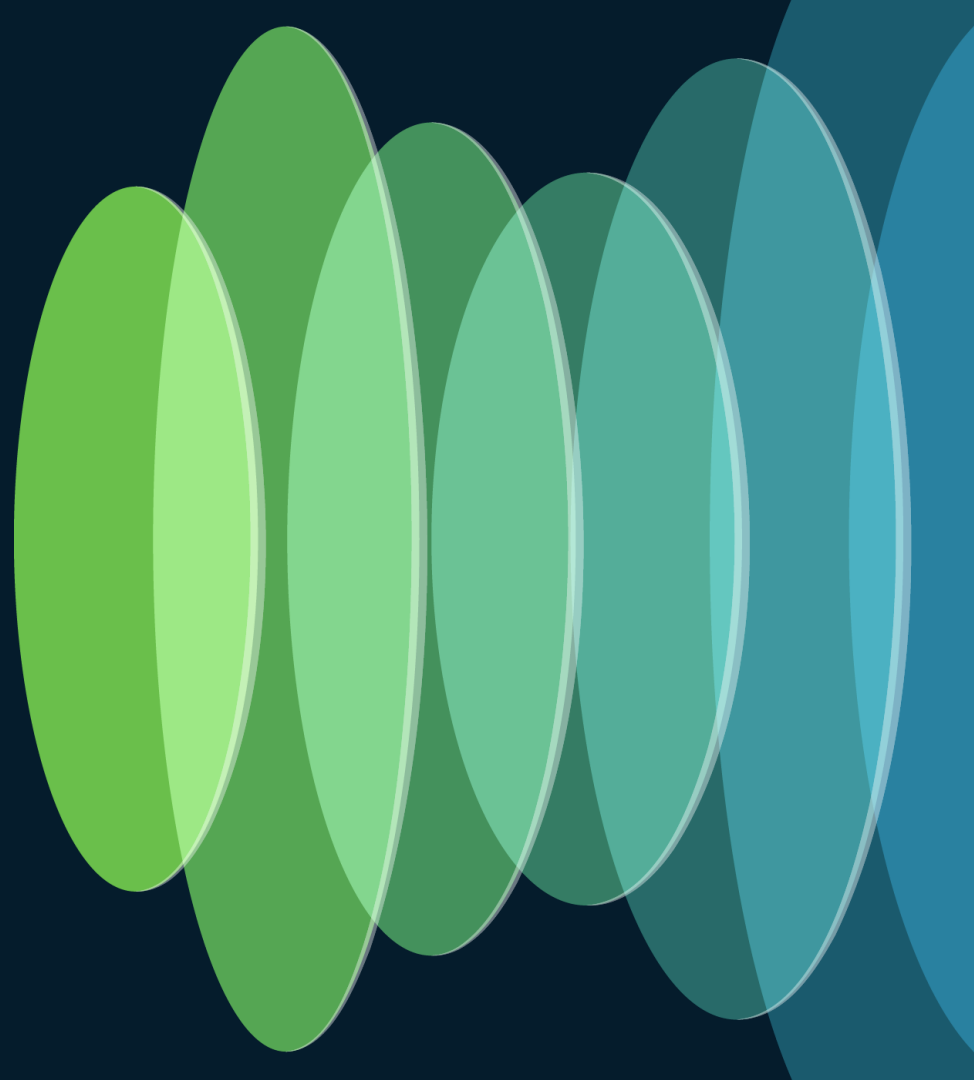by the speaker until June 7, 2024.

CISCO Live!

Agenda

- Flexible Design options for VXLAN EVPN

- External Handoff options for VXLAN EVPN

- Introduction to NDFC

- Automate VXLAN EVPN Single-Site with NDFC
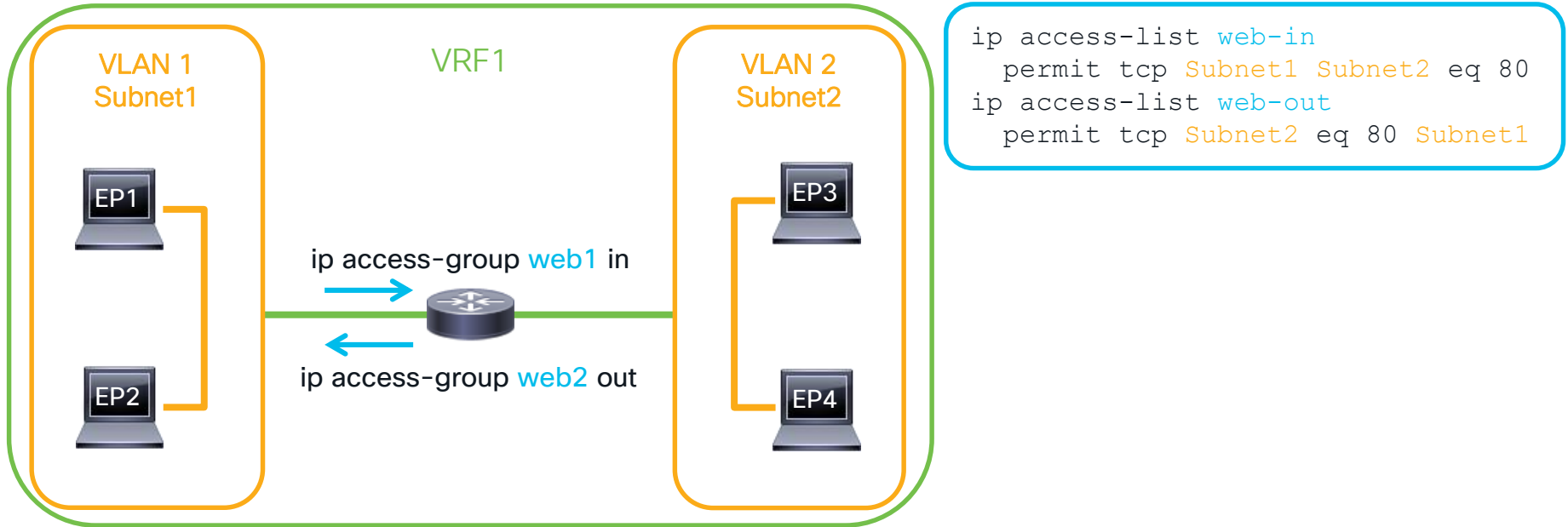
- Conclusion

Data Center
Network
Requirements

What are our basic network requirements?

1) Provide paths for endpoints to communicate at Layer2(MAC) and Layer3(IP)

2) Provide separation of endpoint into Layer2 forwarding domains (VLAN or BD)

3) Routing between IPv4/IPv6 subnets and allow separation of these into multiple VRFs

4) Communication to external L2 networks (DCI)

5) Communication to external L3 networks (WAN)

EP1   VLAN 1   EP2
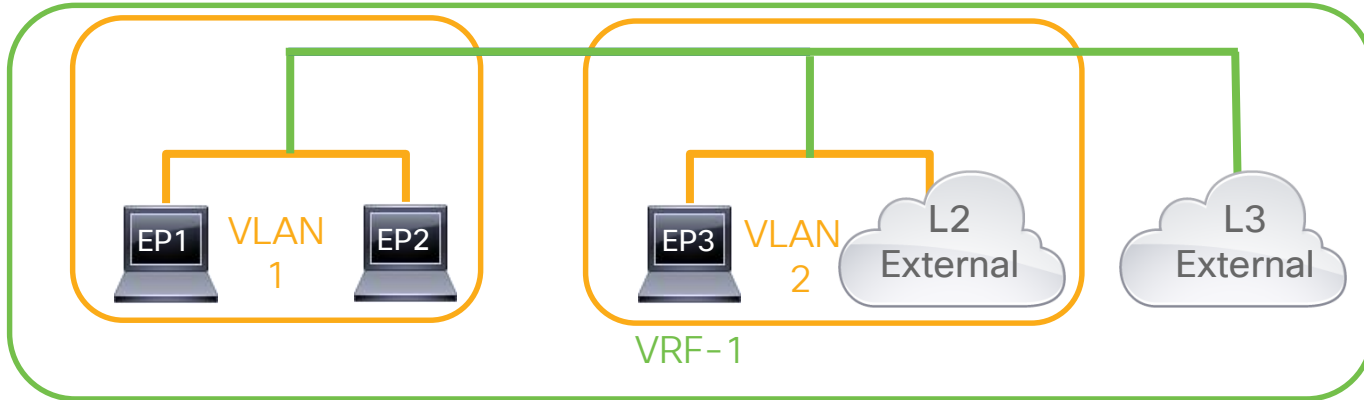
EP3   VLAN 2   L2 External   L3 External

VRF-1
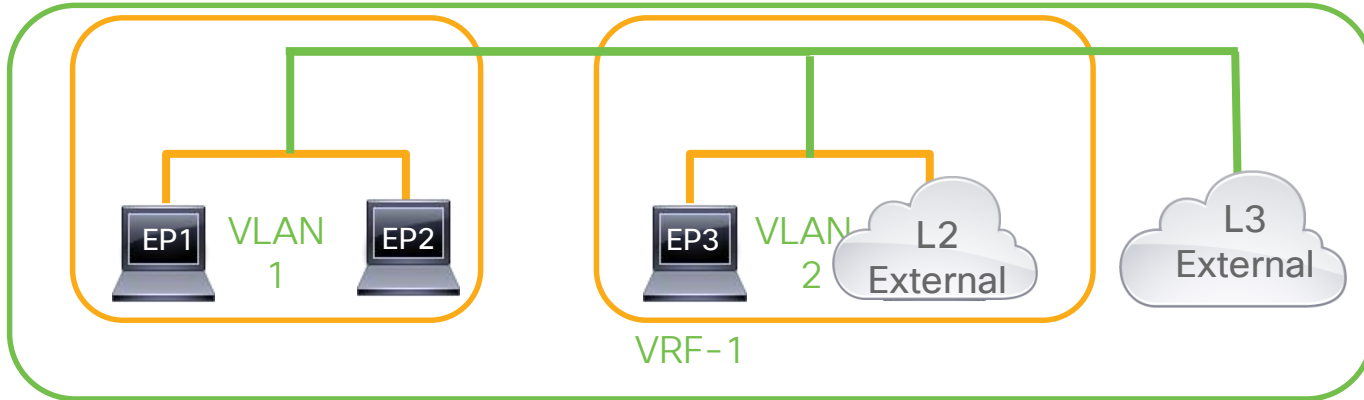
What are our basic network requirements?

6) Allow security policies in order to limit communication to between endpoints to allowed protocols.


```
ip access-list web-in
  permit tcp Subnet1 Subnet2 eq 80
ip access-list web-out
  permit tcp Subnet2 eq 80 Subnet1
```

VLAN 1
Subnet1

VRF1

VLAN 2
Subnet2

EP1

EP3

ip access-group web1 in
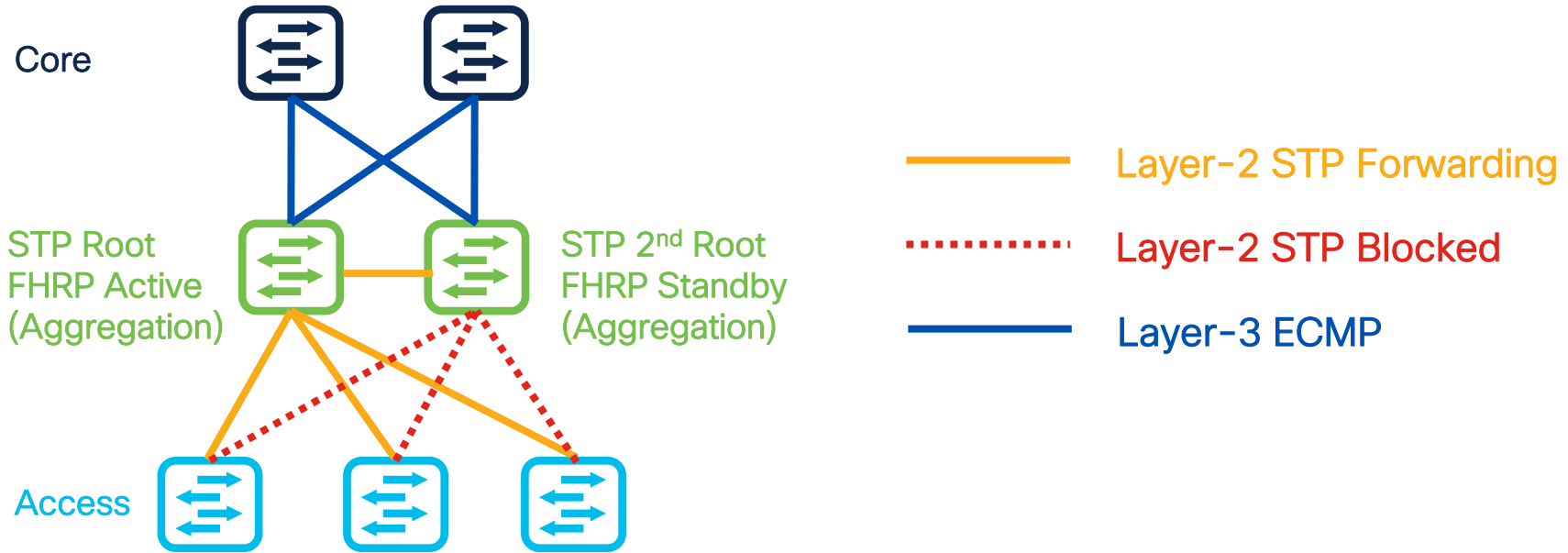
ip access-group web2 out

EP2

EP4

What Physical Topology is required?

- Physical topology must support our endpoint communication (layer-2 / layer-3), and the location of endpoints within the physical network will affect the supporting design/configuration.

VLAN 1 — EP1, EP2

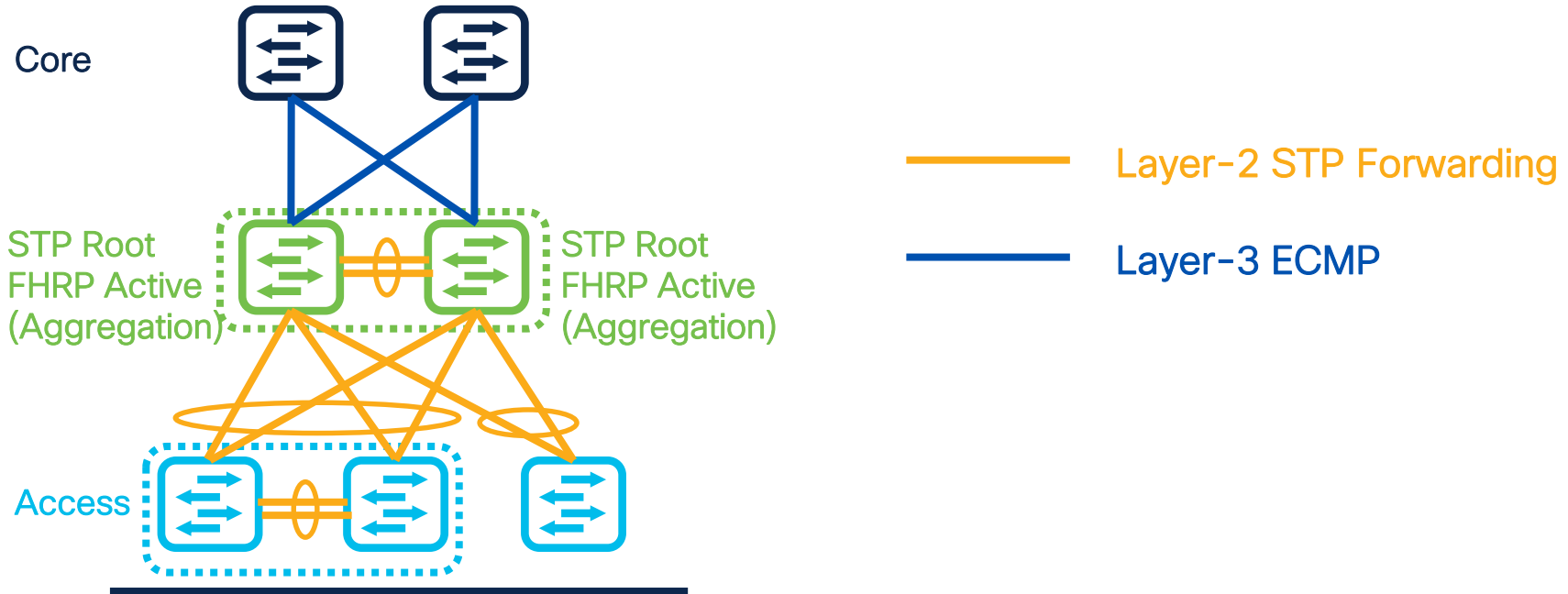VLAN 2 — EP3, L2 External

L3 External

VRF-1

Data Center Network Evolution

Well-Known but Legacy Methods

Core

STP Root
FHRP Active
(Aggregation)

STP 2nd Root
FHRP Standby
(Aggregation)

Access

Classic Spanning-Tree

—— Layer-2 STP Forwarding

········· Layer-2 STP Blocked

—— Layer-3 ECMP

Data Center Network Evolution

Well-Known but Legacy Methods

Core

STP Root
FHRP Active
(Aggregation)

STP Root
FHRP Active
(Aggregation)

Access

VPC and Spanning-Tree

———— Layer-2 STP Forwarding

———— Layer-3 ECMP

Data Center Network Evolution

Well-Known but Legacy Methods

Core

Anycast HSRP (Spine)

Anycast HSRP (Spine)

Leaf

FabricPath (Mac-in-Mac)

Layer-2 FabricPath

Layer-3 ECMP
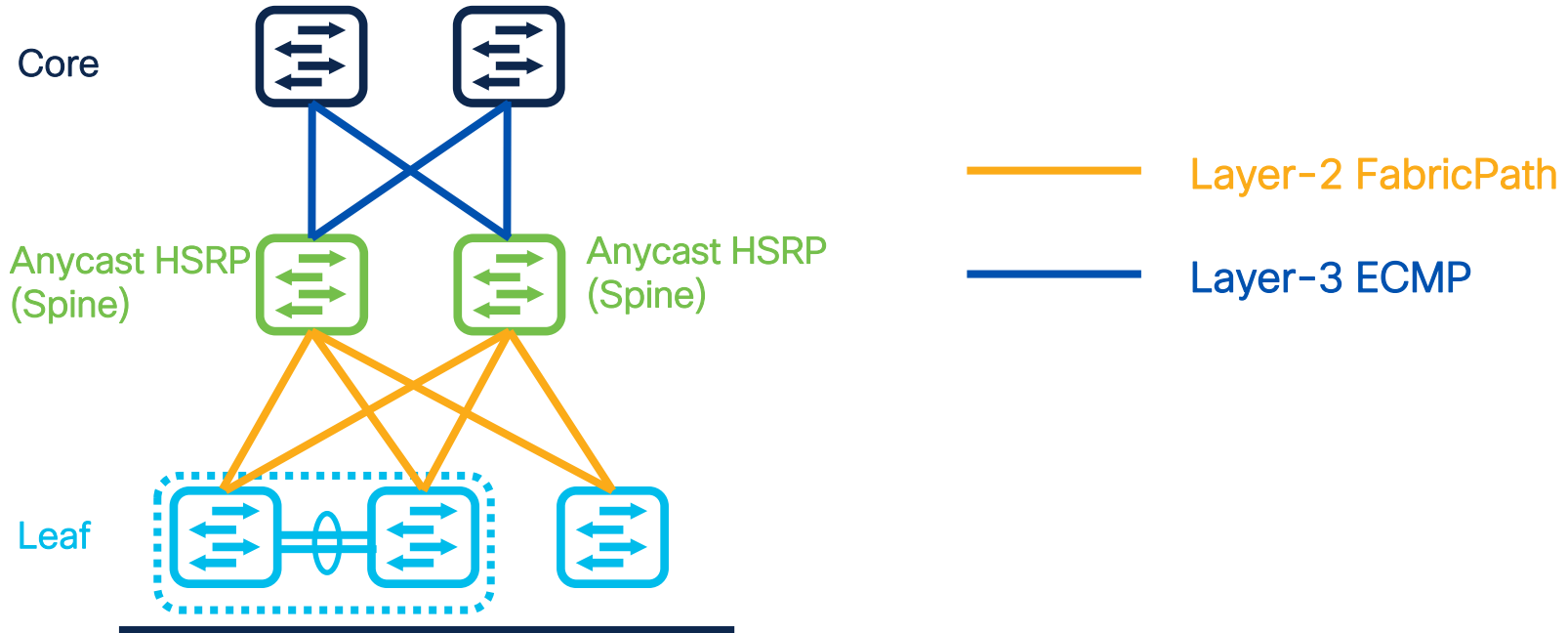
Data Center Network Challenges
Of Legacy Methods

Hierarchical Topology

- Scale-Up with Big Centralized Chassis (Aggregation)

- STP limits full bandwidth utilization

Hair-Pining

- Suboptimal performance, traffic forwarding constrained by spanning-tree rules

- Rigid Network Service Placement (L4-L7)

- Limited Endpoint Mobility

Flood & Learn

- Convergence dependent on Single Tree and MAC Flush (TCN)

- Exposed to Large Broadcast Domains (All Access and Aggregation)

Next Generation– VXLAN BGP EVPN Fabrics

Becoming Industry De-Facto Standard

- An **IGP** is **recommended** for the underlay (OSPF or IS-IS)
  - BGP can also be used if needed

- **BGP EVPN** must be used in the overlay to exchange endpoints information
  - **Spines** act as **route-reflectors**

- **VXLAN** is used to transport endpoint traffic in the fabric
  - **Leafs** are considered **VTEP** as they **encapsulate** and **decapsulate** VXLAN **traffic**

	Underlay
••••••••	BGP EVPN Overlay
R	BGP Route-Reflector
V	VXLAN Tunnel Endpoint (VTEP)
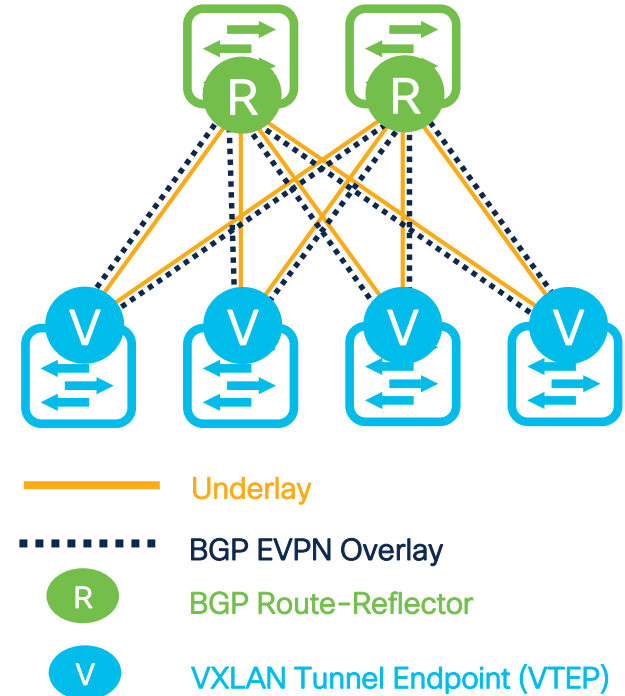
Data Center Network Challenges
Solving it with VXLAN EVPN

Hierarchical Topology

- Scale-Out
  - Add more Spine for bandwidth and redundancy
  - Add more Leaf for port capacity
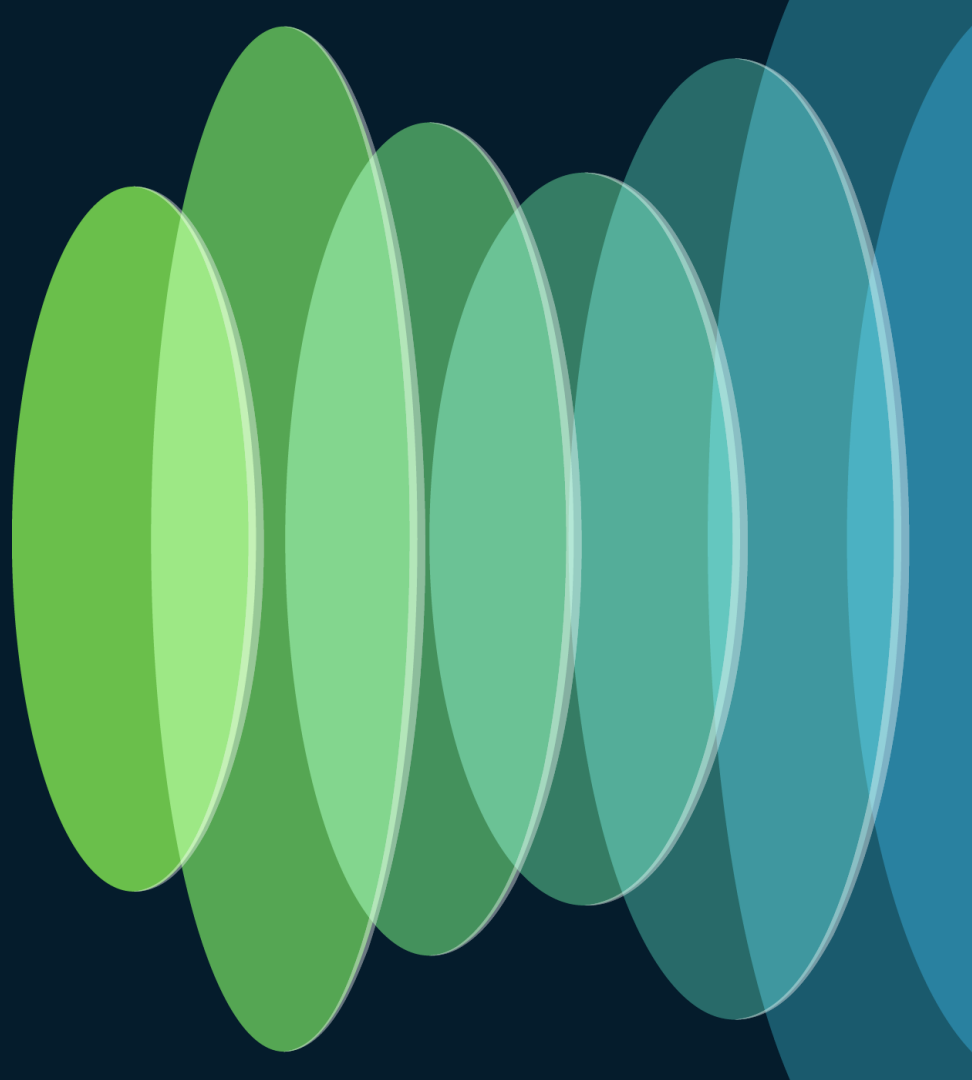- All Links are used (IP ECMP)

No More Hair-Pining

- Default Gateway at every Leaf
- Distributed Anycast Gateway
- Flexible Network Service Placement (L4-L7)
- Pervasive Subnet and Endpoint Mobility
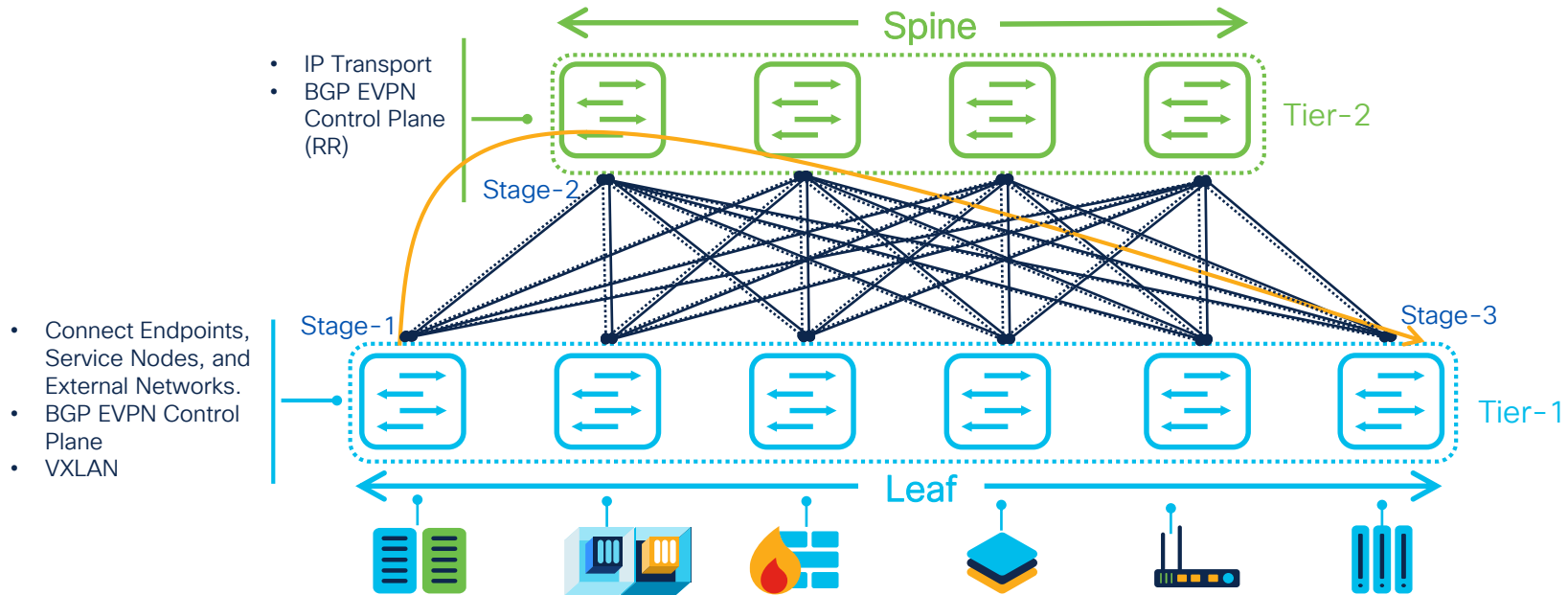
Control-Plane Learned

- Active Learning and Distribution with BGP EVPN
- Reduces the Broadcast Domain by configuring VLANs where needed
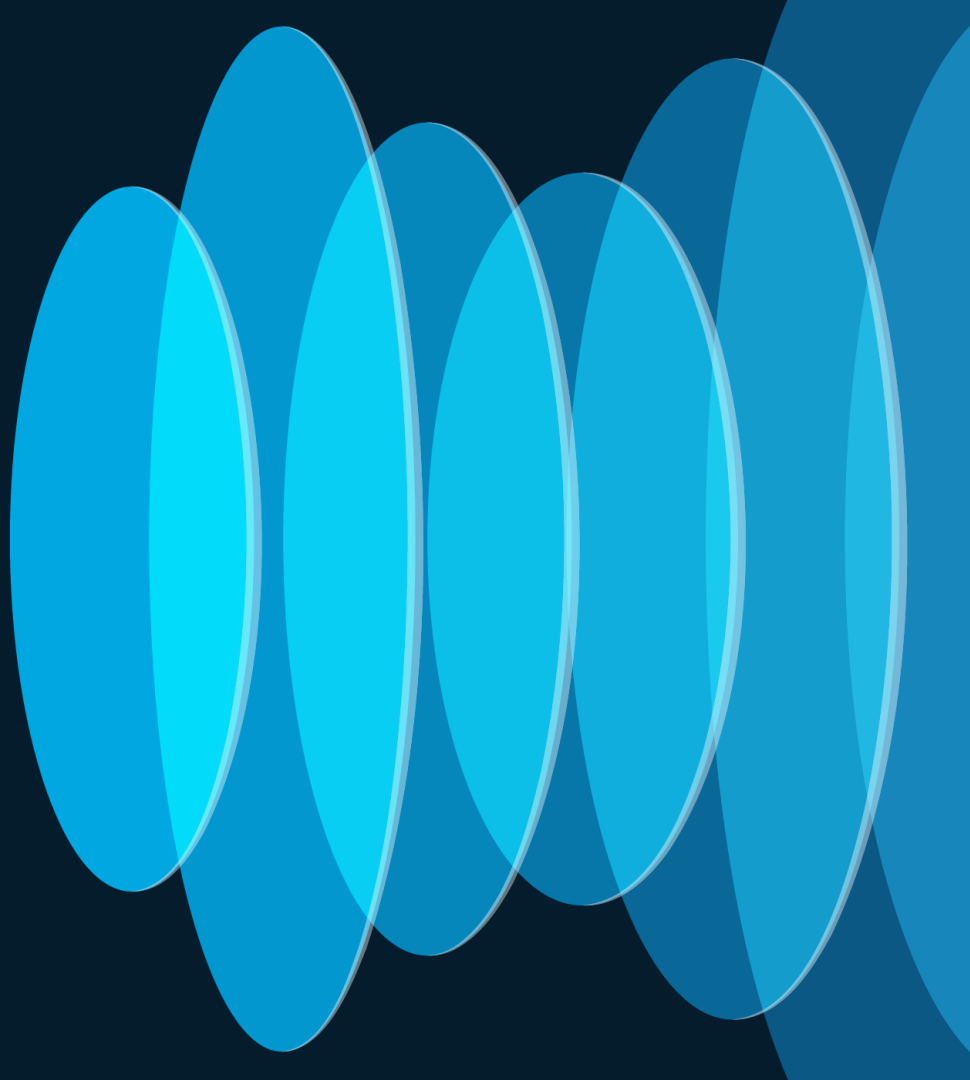
A Leaf and Spine
Paradigm

Kick Start your VXLAN Fabric!
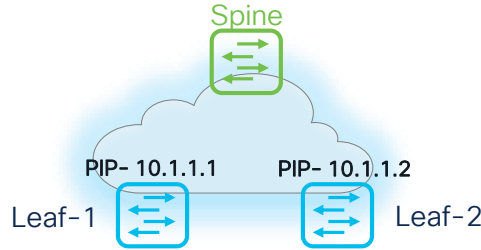
A Leaf and Spine Paradigm

- IP Transport
- BGP EVPN Control Plane (RR)

Spine

Tier-2

Stage-2

Stage-1

Stage-3

- Connect Endpoints, Service Nodes, and External Networks.
- BGP EVPN Control Plane
- VXLAN

Tier-1

Leaf

VXLAN Single-Site
Flexible Design

CISCO Live!

Leaf Node Placement Option

Leaf as Standalone VTEP
- Seen as an individual VTEP (PIP)
- No HA, Orphan hosts only

Spine

PIP- 10.1.1.1   PIP- 10.1.1.2

Leaf-1   Leaf-2

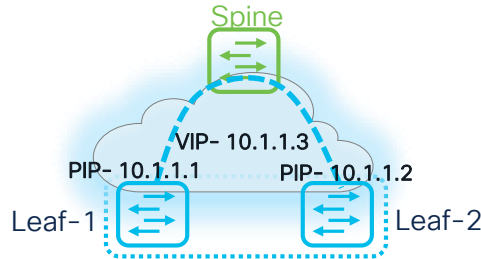| EVPN Route Type	Attachment	Next-hop
Type 2 (Host Routes)	vPC	N/A
	Orphan Port	Advertised by PIP
Type 5 (IP Prefix Routes)	vPC	N/A
	Orphan	Advertised by PIP

Leaf as vPC VTEP
- Seen as a single logical VTEP (VIP)
- Legacy vPC Peer-Link required
- HA with Dual-Attached and Orphan hosts
- Possibility for L4-7 peering

Spine

VIP- 10.1.1.3
PIP- 10.1.1.1   PIP- 10.1.1.2

Leaf-1   Leaf-2

EVPN Route Type	Attachment	Next-hop
Type 2 (Host Routes)	vPC	Advertised by VIP
	Orphan Port	Advertised by VIP
Type 5 (IP Prefix Routes)	vPC	Advertised by VIP
	Orphan	Advertised by VIP
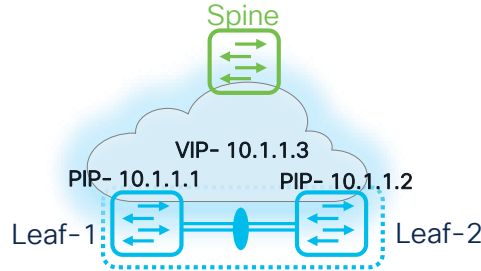
Leaf as vPC Fabric-Peering VTEP
- Seen as a 3 "three" VTEP (PIP + VIP)
- No need for Physical Peer-Link
- More ports available for EPs
- HA with Dual-Attached and Orphan hosts
- Possibility for L4-7 peering

Spine

VIP- 10.1.1.3
PIP- 10.1.1.1   PIP- 10.1.1.2

Leaf-1   Leaf-2

EVPN Route Type	Attachment	Next-hop
Type 2 (Host Routes)	vPC	Advertised by VIP
	Orphan Port	Advertised by PIP
Type 5 (IP Prefix Routes)	vPC	Advertised by PIP
	Orphan	Advertised by PIP

Super-Spine Node Placement Option

✔ Scale-out Multi-Clos Fabric to Interconnect the PODs using Super-Spine

✔ Architecture beyond a single server room. Simpler capacity planning

✔ Clean role separation and uniform reachability from the entire fabric are the major advantages
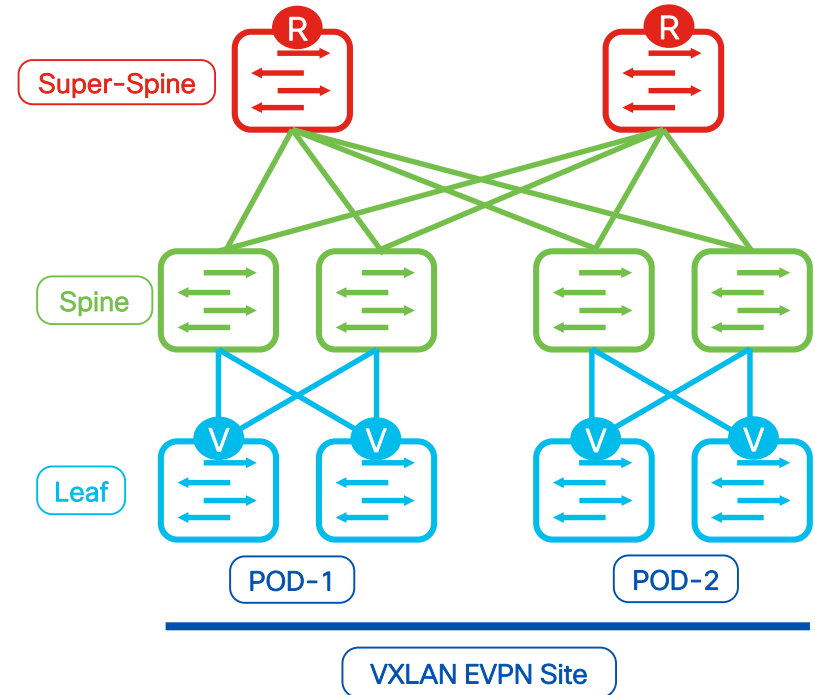
Leaf hosts:
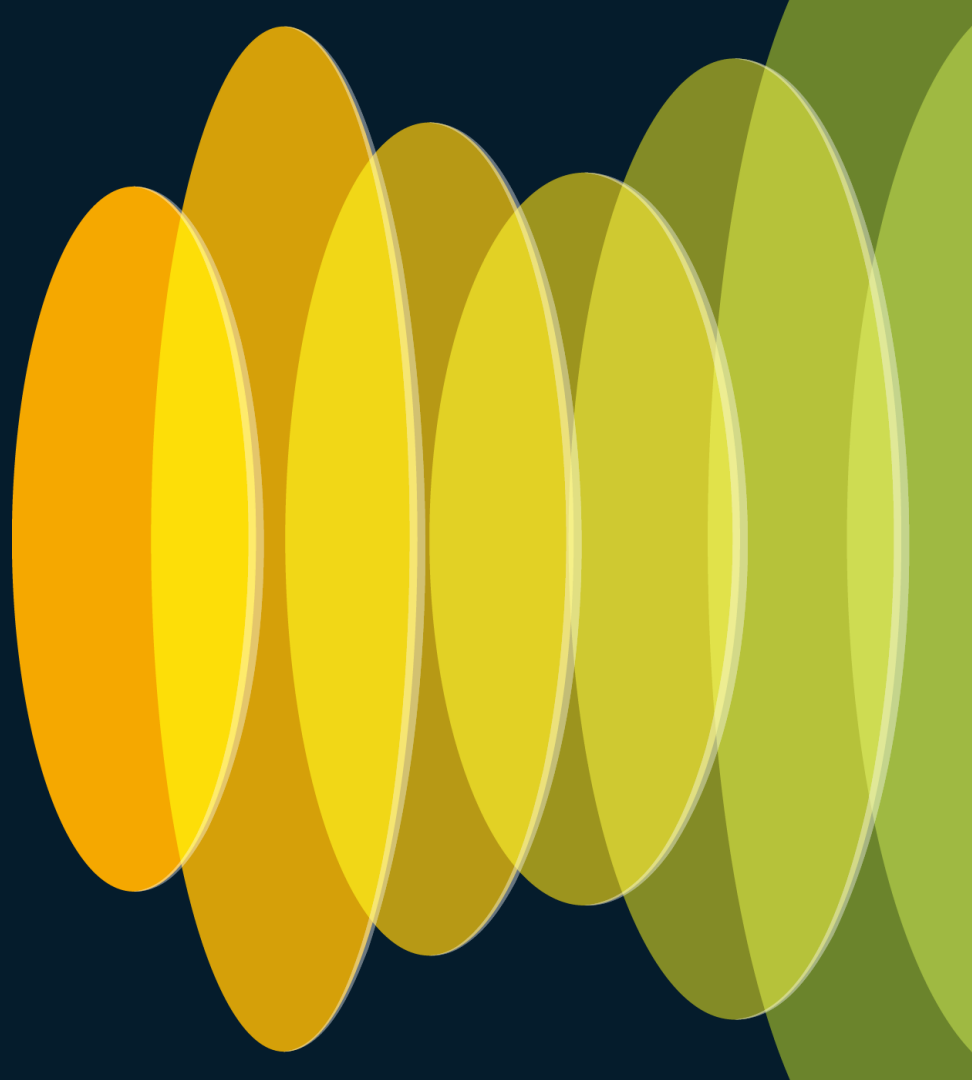- East-West VXLAN (VTEP)

Spine hosts:
- Layer-3 IP Transit

Super-Spine hosts:
- Route Reflector (RR) (iBGP EVPN)
- Rendezvous Point (RP) (Multicast Underlay BUM)

Super-Spine

Spine

Leaf

POD-1   POD-2

VXLAN EVPN Site

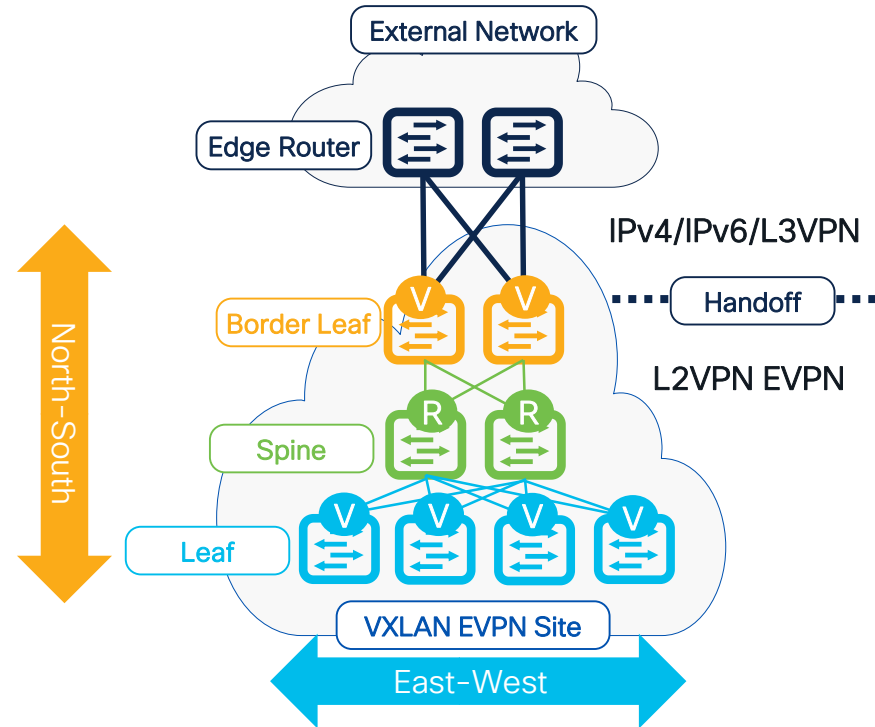# External Handoff
Border Placement

Border Node Placement Option
Border as Leaf (Flexible Design option-1)

- **Leaf** VTEP (**East-West** traffic)
  - Connectivity to endpoints, the first-hop routing
  - Server-to-server traffic

- **Border** Leaf VTEP (**North-South** traffic)

  ✔ Capacity planning only for North-South traffic flows

  ✔ Clean role separation and uniform reachability from the entire fabric are the major advantages

  ✔ Support for Inter-AS option A (VRF-LITE) and seamless VXLAN-MPLS gateway (Border-PE)

  ✔ Optionally it can have directly attached endpoints

Border Leaf hosts:
- North <> South VXLAN (VTEP) <> IP handoff

External Network

Edge Router

IPv4/IPv6/L3VPN

Handoff

Border Leaf

L2VPN EVPN

North-South

Spine

Leaf

VXLAN EVPN Site

East-West

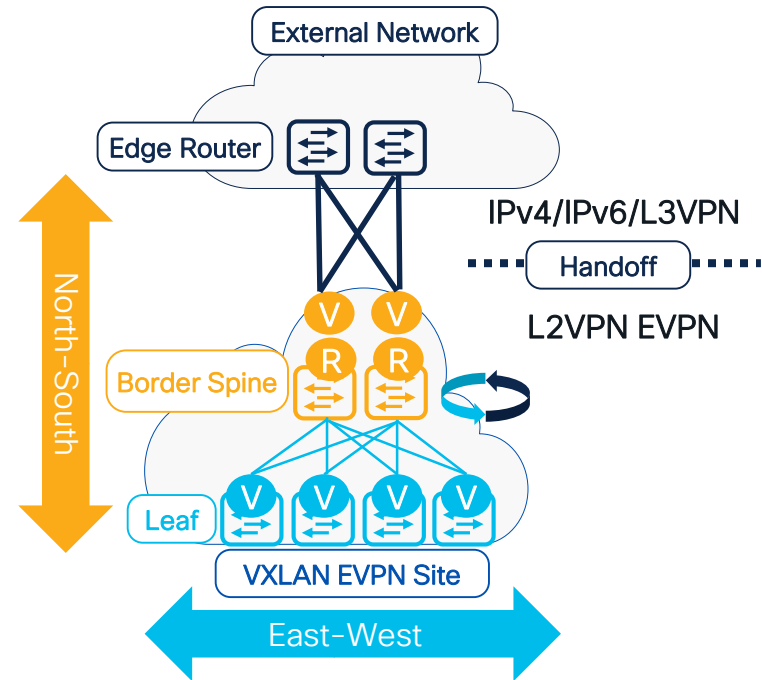# Border Node Placement Option
Border Spine (Flexible Design option-2)

✔️ Flexible option for Small deployments

✔️ Support for Inter-AS option A (VRF-LITE) and seamless VXLAN-MPLS gateway (Border-PE)

⚠️ Extra functional dependency (Border + Spine)

⚠️ Capacity planning needs to accommodate all flows

Border Spine hosts:
- North <> South VXLAN (VTEP) <> IP handoff
- Route Reflector (RR) (iBGP EVPN)
- Optionally Rendezvous Point (RP) (Multicast Underlay BUM)

Ⓥ Ⓥ = VTEP

Ⓡ = RR/RP

External Network

Edge Router

IPv4/IPv6/L3VPN

Handoff

L2VPN EVPN

North-South

Border Spine

Leaf

VXLAN EVPN Site

East-West

Border Node Placement Option
Border on top of Super-Spine (Flexible Design option-3)

✓ Capacity planning only for North-South traffic flows

✓ Clean role separation and uniform reachability from the entire fabric are the major advantages

✓ Support for Inter-AS option A (VRF-LITE) and seamless VXLAN-MPLS gateway (Border-PE)
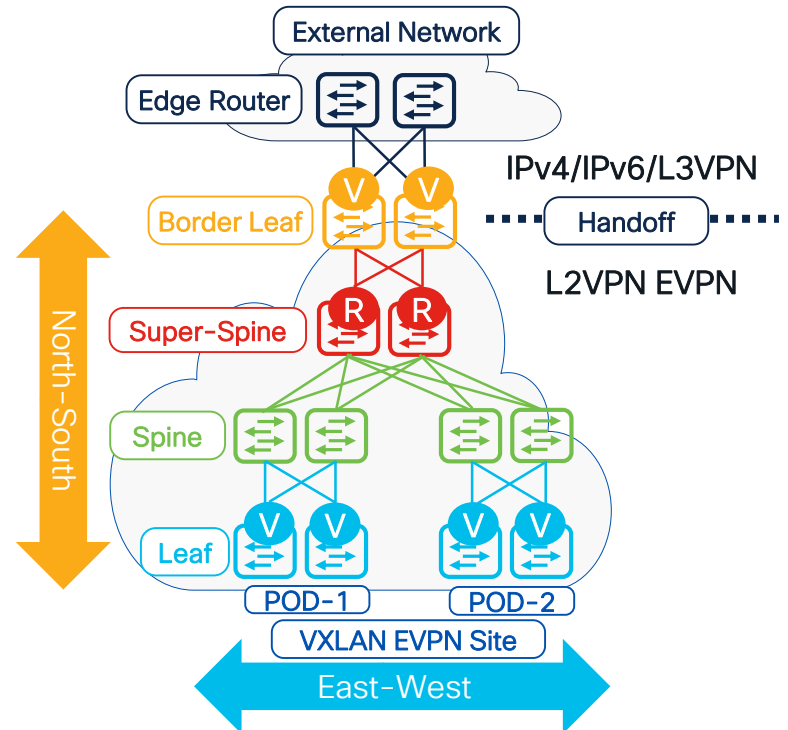
✓ Optionally it can have directly attached endpoints

Border Leaf hosts:
- North <> South VXLAN (VTEP) <> IP handoff

V V = VTEP

R = RR/RP

Border Node Placement Option

Border Super-Spine (Flexible Design option-4)

✔️ Support for Inter-AS option A (VRF-LITE) and seamless VXLAN-MPLS gateway (Border-PE)

⚠️ Extra functional dependency (Border + Spine). **Not recommended** due to Multi-POD failure dependency

⚠️ Capacity planning needs to accommodate all flows

Border Super-Spine hosts:
- North <> South VXLAN (VTEP) <> IP handoff
- Route Reflector (RR) (iBGP EVPN)
- Optionally Rendezvous Point (RP) (Multicast Underlay BUM)

Ⓥ Ⓥ = VTEP

Ⓡ = RR/RP
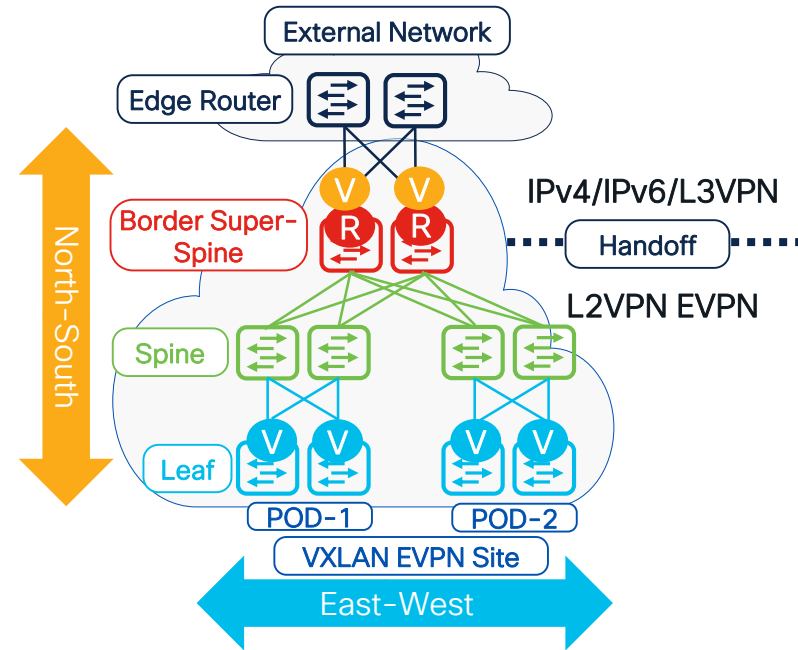
External Handoff
Connectivity
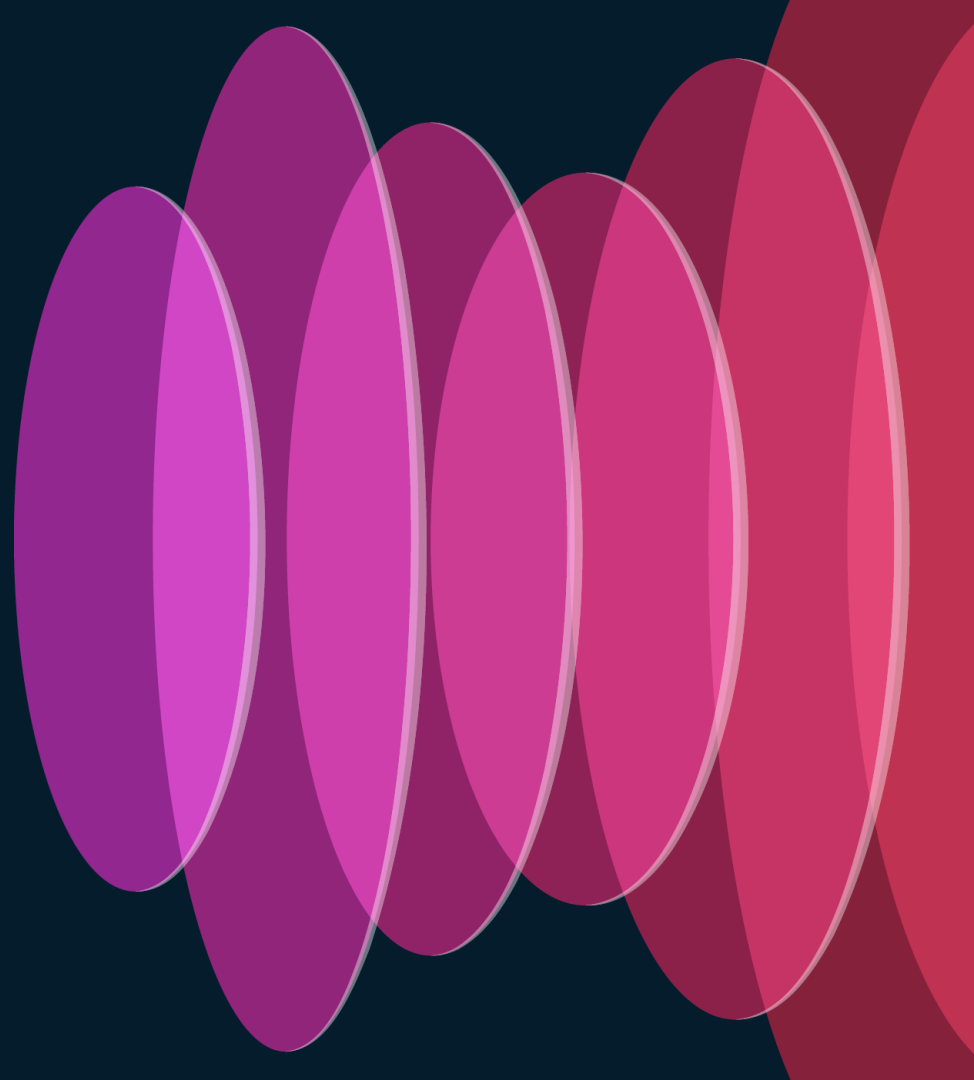
External Layer-3 connectivity options
Inter-AS Option A

✔️ Clear separation of Autonomous Systems

✔️ Simple, Straight forward, and Commonly used

✔️ No need for redistribution

✔️ Easy and Flexible BGP route-filtering mechanisms

✔️ BGP natural loop avoidance

✔️ Structured handoff between the VXLAN BGP EVPN fabric and the external routing domain (Backbone, WAN, Campus, etc.)

⚠️ Not ideal for High scale VRF handoff deployment

Peering Type = Sub-interfaces on physical routed (or L3 Port-channel) interfaces. L3 SVI also supported.
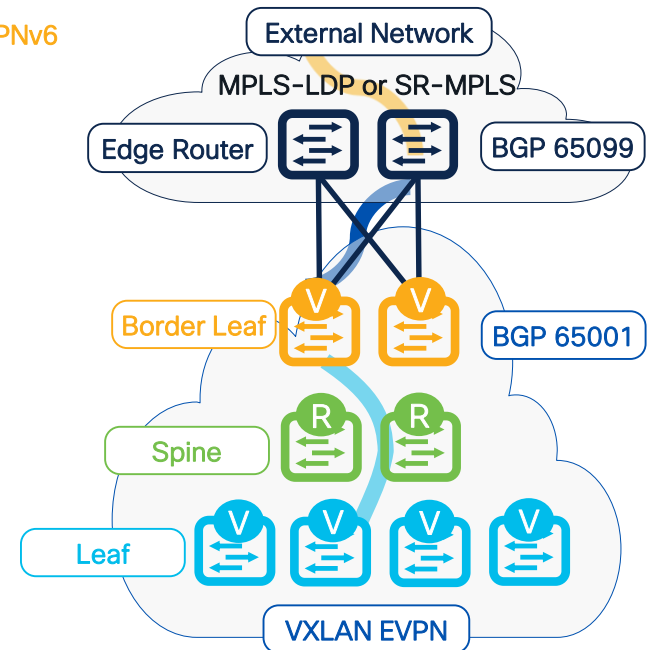
- Sub-interface with dot1q tag to mark the traffic to a specific VRF

- Sub-interface used for eBGP peering and as next-hop

- Per VRF, Per Sub-interface eBGP session
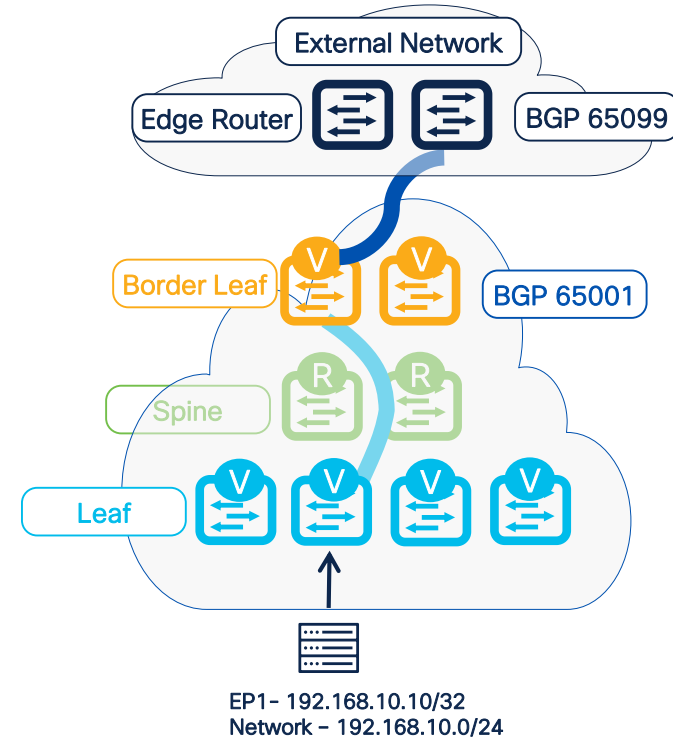
VPNv4/VPNv6

IPv4/IPv6

EVPN

Separated Border + PE (Inter-AS Option A)

External Network

MPLS-LDP or SR-MPLS

Edge Router

BGP 65099

Border Leaf

BGP 65001

Spine

Leaf

VXLAN EVPN

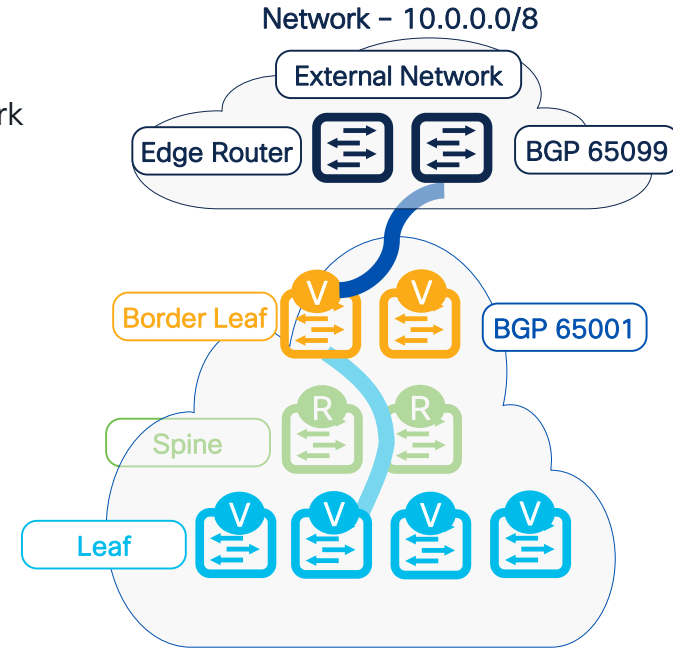# External Layer-3 connectivity options
Inter-AS Option A

1 An EP is connected to Leaf and records an ARP

2 EVPN Type-2 (/32) and Type-5 (/24) are created with Leaf VTEP IP as Next-Hop

3 Leaf advertises EVPN route towards Spine with Route-Target attachment with MAC-VRF (ASN:L2VNI) and IP-VRF (ASN:L3VNI)

4 Border imports EVPN route based on matching Route-Target and downloads the route in BGP-VRF.

5 Border advertises route as IPv4/IPv6 towards External Edge with itself as Next-Hop

External Network

Edge Router   BGP 65099

Border Leaf   BGP 65001

Spine

Leaf

EP1- 192.168.10.10/32
Network – 192.168.10.0/24

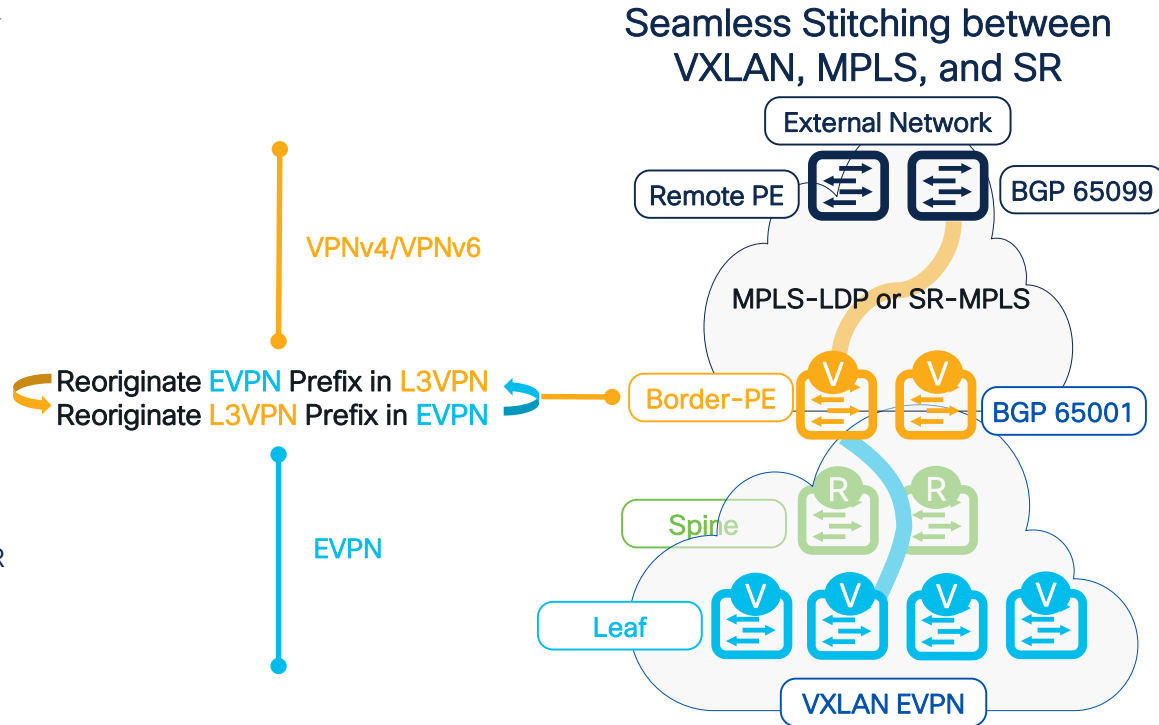# External Layer-3 connectivity options
Inter-AS Option A

1 Edge Router uses per-VRF eBGP session to advertise External network with itself as Next-Hop

2 Border receives External route in BGP-VRF and export it as EVPN Type-5 alongside IP-VRF Route-Target (ASN:L3VNI)

3 All VTEPs (Leaf) learns EVPN Type-5 and install them in BGP-VRF based on matching IP-VRF Route-Target (ASN:L3VNI). From BGP-VRF it downloads in RIB-VRF and Forwarding table

Network – 10.0.0.0/8

External Network

Edge Router   BGP 65099

Border Leaf   BGP 65001

Spine

Leaf

External Layer-3 connectivity options
Seamless Protocol Gateway Model (Border-PE)

✅ Combines two different encapsulations and Address Family, using a "single-box (Border-PE)" instead of a "two-box (CE-PE)" model

✅ VXLAN VTEP Border nodes also becomes a MPLS L3VPN Provider Edge (PE), resulting in a role called Border-PE

✅ Best suited for high scale VRF deployment

✅ Saves CAPEX and OPEX

✅ Seamless stitching between L2VPN EVPN and VPNv4/v6 Address Family

✅ BGP route-filtering mechanisms available

⚠️ Specific Hardware support

- • MPLS LDP: Nexus 3600-R, Nexus 9500-R

- • SR MPLS: Nexus 9300 FX2/FX3/GX/GX2, Nexus 9500-R

Seamless Stitching between VXLAN, MPLS, and SR

External Network

Remote PE

BGP 65099

VPNv4/VPNv6

MPLS-LDP or SR-MPLS

Reoriginate EVPN Prefix in L3VPN
Reoriginate L3VPN Prefix in EVPN
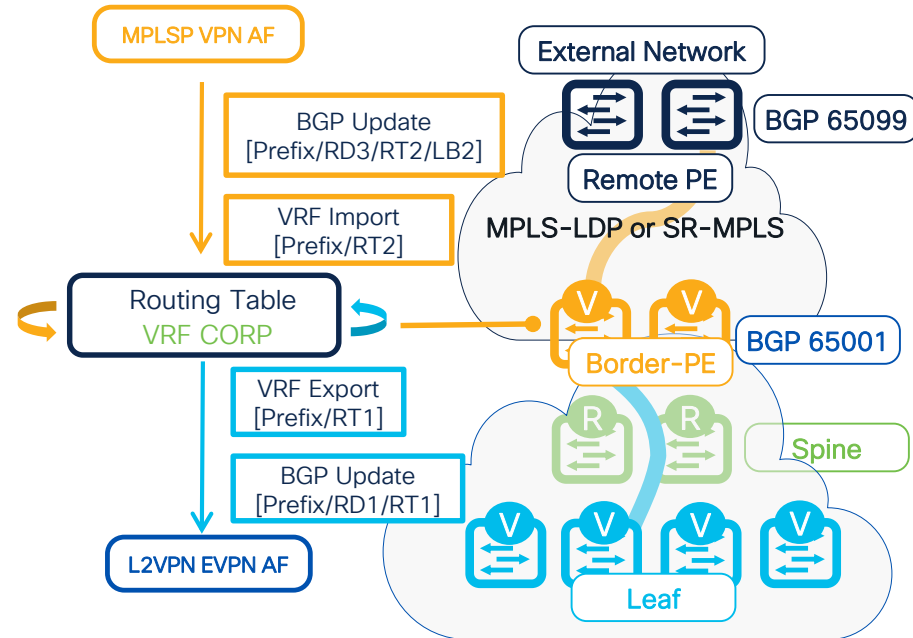
Border-PE

BGP 65001

Spine

EVPN

Leaf

VXLAN EVPN

External Layer-3 connectivity options
Seamless Protocol Gateway Model (Border-PE)

1 EVPN routes are imported into the local VRF instance according to the BGP Route-Target filtering

2 Imported routes are re-exported to the VPN address family, and a local VPNv4/v6 MPLS label is allocated

3 Re-exported routes are Re-originated and advertised to eBGP peers with the allocated VPNv4/v6 MPLS label alongside local RD and RT of Border-PE. The per-VRF VXLAN fabric relevant BGP RT is stripped off

Key	Description
Prefix	IPv4/IPv6 route in VXLAN going towards MPLS
RD1	VPN Route-Distinguisher for VRF CORP on Source Leaf
RT1	BGP Route-Target for VRF CORP
RD2	VPN Route-Distinguisher for VRF CORP on Border-PE
RT2	BGP Route-Target for VRF CORP facing MPLS
LB1	MPLS Label

MPLSP VPN AF

BGP Update
[Prefix/RD2/RT2/LB1]

VRF Export
[Prefix/RD2/RT2]

Routing Table
VRF CORP

VRF Import
[Prefix/RT1]

BGP Update
[Prefix/RD1/RT1]

L2VPN EVPN AF

External Network

BGP 65099

Remote PE

MPLS-LDP or SR-MPLS

BGP 65001

Border-PE

Spine

Leaf
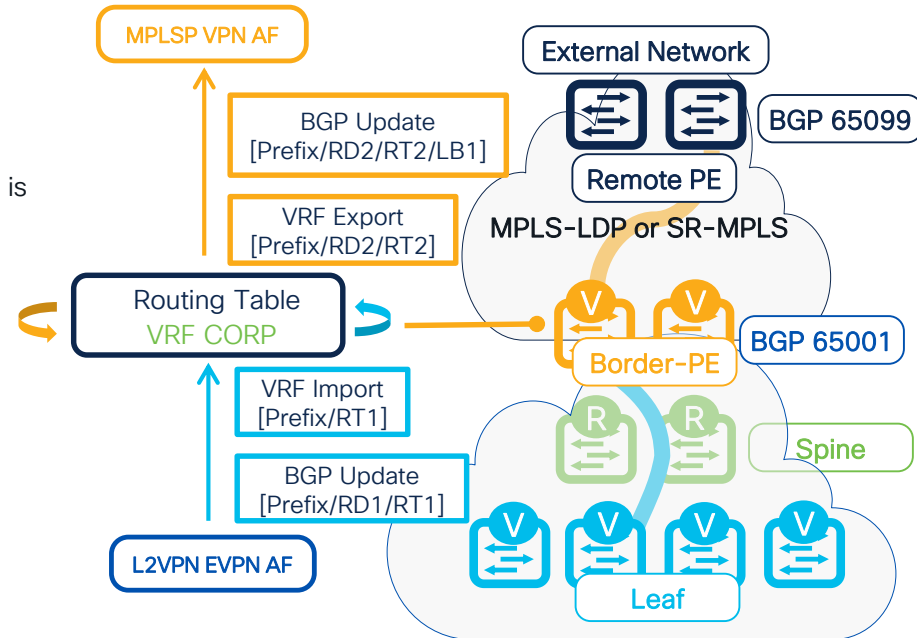
External Layer-3 connectivity options
Seamless Protocol Gateway Model (Border-PE)

1 MPLS routes are imported into local VRF according to Route-Target filtering. Routes are imported into local RIB and FIB with MPLS label.

2 Imported routes are re-exported to the EVPN address family

3 Re-exported routes are Re-originated and advertised to EVPN Spine peers with fabric specific encapsulation information such as VXLAN VNI and local RD/RT. The per-VRF MPLS network relevant BGP RT is stripped off

Key	Description
Prefix	IPv4/IPv6 route in MPLS going towards VXLAN
RD3	VPN Route-Distinguisher for VRF CORP on Remote MPLS PE
RT2	BGP Route-Target for VRF CORP facing MPLS
LB2	MPLS Label
RD1	VPN Route-Distinguisher for VRF CORP on Border-PE
RT1	BGP Route-Target for VRF CORP

MPLSP VPN AF

BGP Update [Prefix/RD3/RT2/LB2]

VRF Import [Prefix/RT2]

Routing Table
VRF CORP

VRF Export [Prefix/RT1]

BGP Update [Prefix/RD1/RT1]

L2VPN EVPN AF

External Network

BGP 65099

Remote PE

MPLS-LDP or SR-MPLS

Border-PE

BGP 65001

Spine

Leaf

Advertise VTEP Primary IP Address
What exactly is Advertise-PIP and Why you need it?

① External routes are injected into VXLAN fabric

② Border advertises External routes as EVPN Type-5 with the BGP Next-Hop of vPC VIP (Anycast)

③ From Leaf perspective the Next-Hop to reach Backbone is Border Anycast vPC VIP

Problem Statement: While ARP/MAC/IPv6 ND entries are synced between the peers of a vPC pair, prefix routes belonging to an individual peer as well as external routes received by a peer are not synced between vPC peer switches. Using the VIP as the BGP next-hop for these routes can cause traffic to be forwarded to the wrong vPC peer and hence be black-holed.

Backbone

vPC VIP
10.13.0.3

PIP-1
10.13.0.1

PIP-2
10.13.0.2

MP-BGP EVPN Table
Backbone →
10.13.0.3

VXLAN EVPN Site
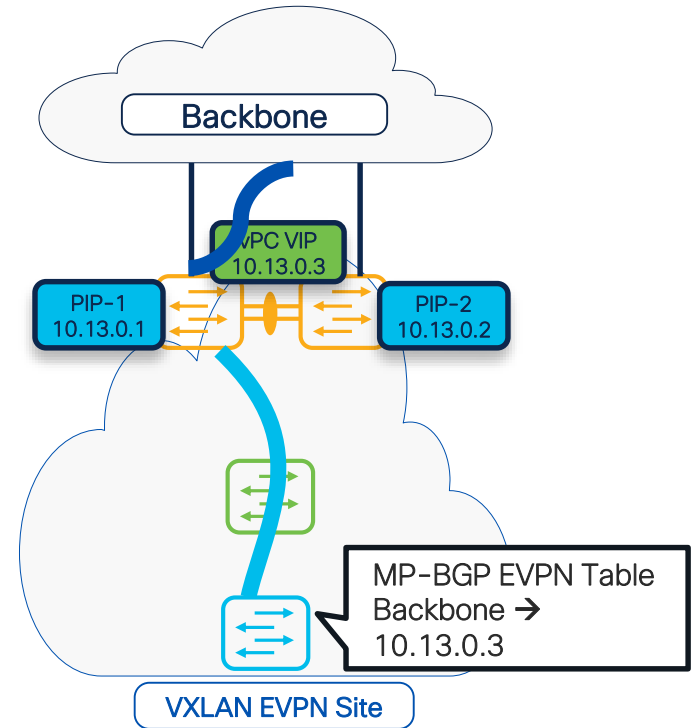
Advertise VTEP Primary IP Address
What exactly is Advertise-PIP and Why you need it?

1 Border-2 losses the link towards Backbone and Border-1 is the only available path towards the Fabric

2 Border-1 continues to advertise External routes as EVPN Type-5 with the BGP Next-Hop of vPC VIP (Anycast)

3 From Leaf perspective the Next-Hop to reach Backbone is Border Anycast vPC VIP. Hence, traffic can hash to either Border-1 or Border-2. If packet hits Border-2, it will drop the traffic!

Note: Border-1 still has an active link towards the Backbone and advertises the routes towards the Spine (RR). Later, the Spine will reflect the route to Border-2, but it will reject it due to Next-Hop being its own IP (VIP 10.13.0.3)

Backbone

vPC VIP
10.13.0.3

PIP-1
10.13.0.1

PIP-2
10.13.0.2

MP-BGP EVPN Table
Backbone →
10.13.0.3

VXLAN EVPN Site
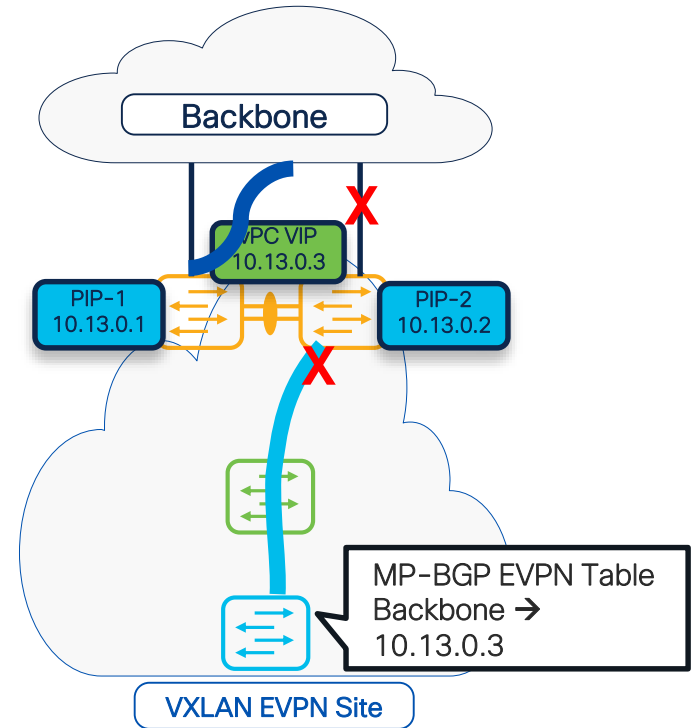
Advertise VTEP Primary IP Address

What exactly is Advertise-PIP and Why you need it?

```
router bgp 65001
  address-family l2vpn evpn
    advertise-pip

Interface nve1
  advertise virtual-rmac
```

On both VPC peer Border Devices

The advertise-pip command lets BGP use the PIP as next-hop when advertising prefix routes or leaf-generated routes if vPC is enabled. With the advertise-pip and advertise virtual-rmac commands, EVPN Type-5 routes are advertised with PIP, and EVPN Type-2 routes are still advertised with VIP. In addition, a virtual MAC will be used with the VIP that is shared by both vPC peers, and individual peer specific system Router MAC will be used with PIP when the advertise-pip feature is enabled. In this way, the traffic will always be destined to the right vPC peer.

Backbone

vPC VIP
10.13.0.3

PIP-1
10.13.0.1

PIP-2
10.13.0.2

MP-BGP EVPN Table
Backbone →
10.13.0.1, 10.13.0.2

VXLAN EVPN Site
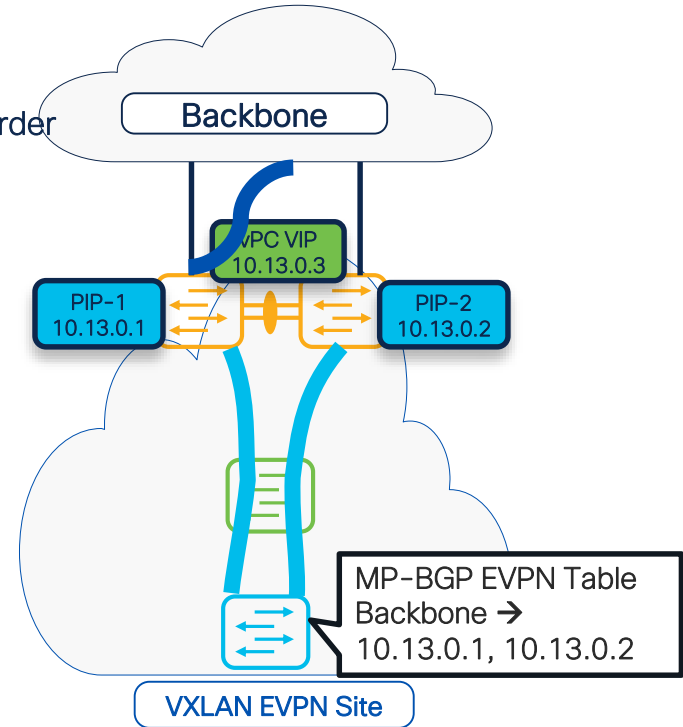
Who can be the
External Edge Router?

Edge Router Placement Option
Nexus and Non-Nexus support

External Network

Edge Router

Border Leaf

Data Center VXLAN EVPN

External Network

Edge Router

Border Leaf

Data Center VXLAN EVPN

External Network

Edge Router

Border Leaf

Data Center VXLAN EVPN

Managed NX-OS as Edge Router
- ✔ Nexus 9000
- ✔ Nexus 7000

Managed IOS-XR as Edge Router
- ✔ Cisco 8000
- ✔ ASR 9000
- ✔ NCS 5500

Managed IOS-XE as Edge Router
- ✔ Catalyst 9000
- ✔ Catalyst 8000
- ✔ ASR 1000
- ✔ CSR 1000

Cool! But What's the Catch?

There is always a catch ☺

(1) This seems really cool! Is it easy to configure?
There are lots of moving parts: OSPF/BGP/VXLAN. Manual configuration can be challenging.

(2) How easy is it to make changes?
You still rely on traditional SSH based management to each device, which can be cumbersome and error prone.

(3) How much Visibility do I have into the network?
Visibility and Troubleshooting is still performed on a "switch-by-switch" basis.

What is Nexus Dashboard Fabric Controller?
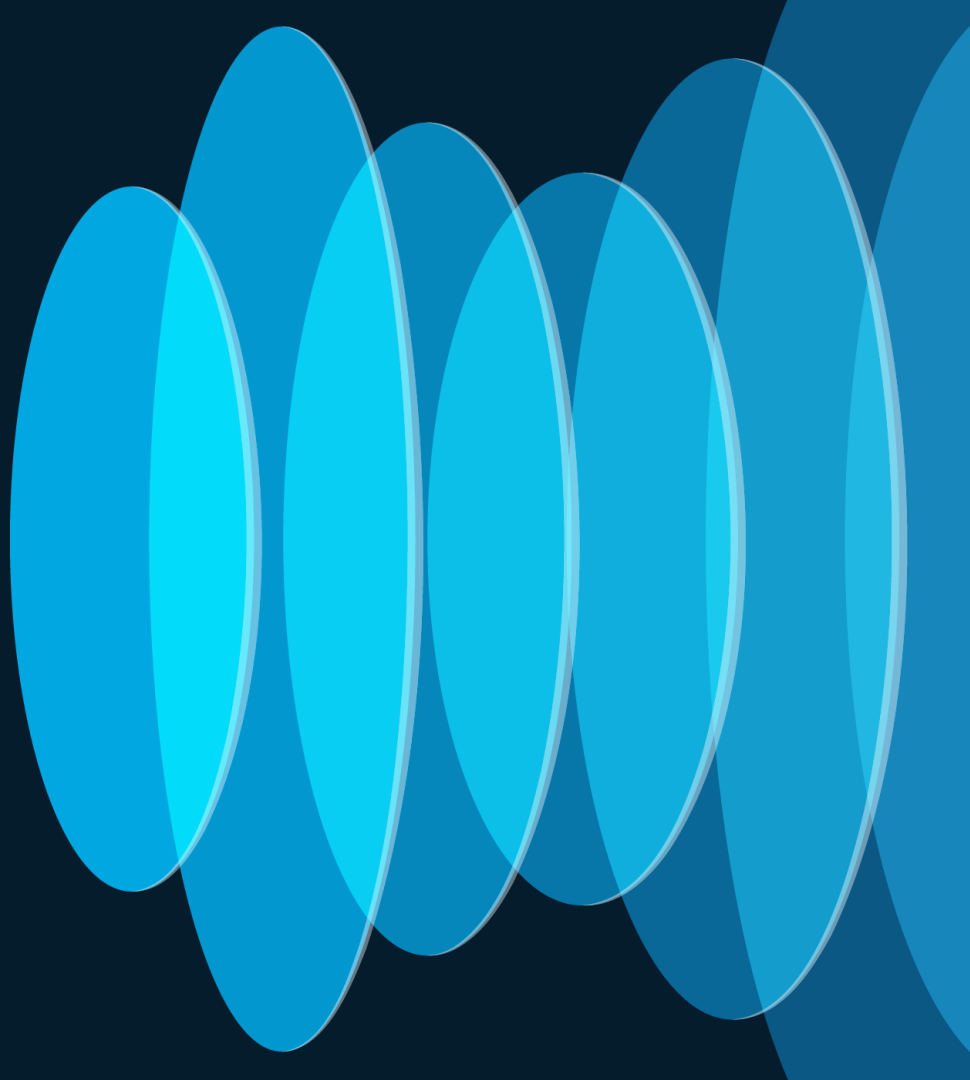
Cisco Nexus Dashboard

Simple to automate, simple to consume

Powering automation
Unified agile platform

Insights

Orchestrator

kubernetes

Fabric Controller

Data Broker

SAN Controller

Consume all services in one place

Private cloud | Third-party Apps | Public cloud
aws
Google Cloud
Azure

Cisco Nexus Dashboard Fabric Controller
App accessed through Cisco Nexus Dashboard
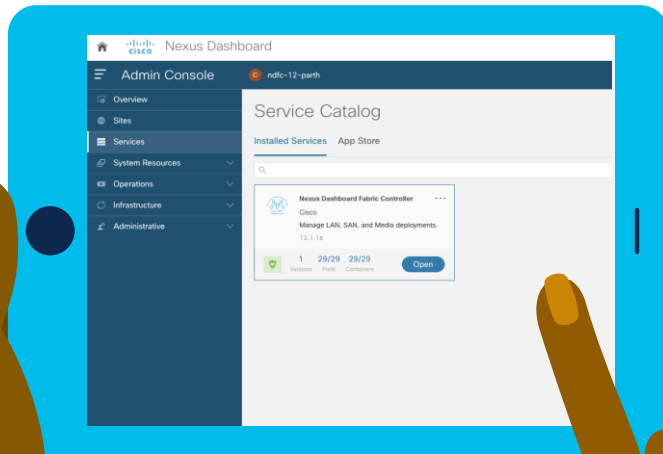
Cisco Nexus Dashboard → Access NDFC
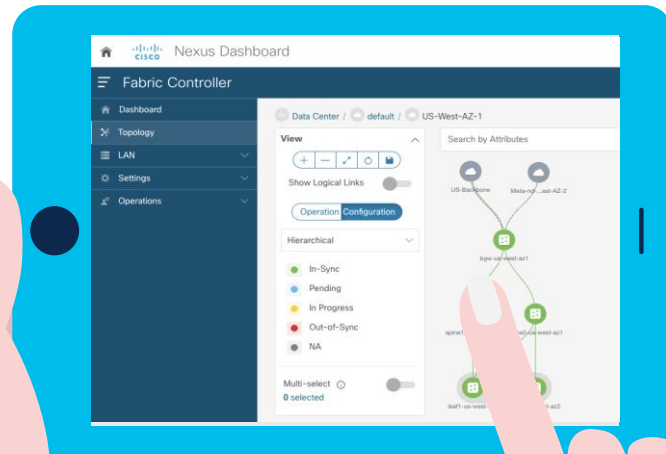
Benefits   Automation   Management and Compliance   Visibility and Monitoring

Cisco Nexus Dashboard Fabric Controller
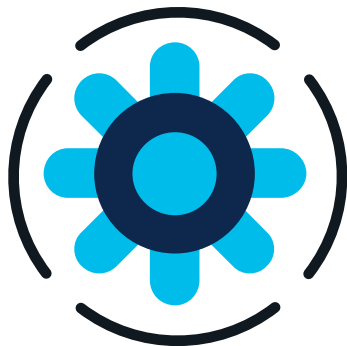
Automation

Accelerate provisioning and simplify deployments

Management

In depth Management and control for all network deployments

Visibility

Get Centralized Visibility and Monitoring views

Automation

Accelerate provisioning from days to minutes

Easy to understand approach to auto-bootstrapping of entire fabric

Rapid Deployment with Fabric Builder best practice templates for VXLAN-EVPN

Optimized for both large deployments and traditional deployment models

Service Insertion and Layer-3 handoff

DevOps friendly

Benefits

Simplify fabric deployments | Developer agility | VXLAN EVPN Multi-Site

Management

Single point for management for data center operations

Optimized for both large deployments and traditional deployment models

Granular RBAC

Image management

RMA

Change Control

Management for non-Nexus platforms

Benefits

Reliability	Compliance	Secure

Visibility & Monitoring

Get comprehensive monitoring

Enhanced topology views

Compute and endpoint visibility

VXLAN OAM support with NDFC

Obtain detailed inventory, health, resource consumption information on devices

End-to-end visibility, monitoring and troubleshooting

Integrate with NDI for Day 2 operations

Benefits

| Intuitive | Deep visibility | Enhanced monitoring |

Cisco NDFC Modes

Make decision at run-time!

Runtime Feature Installer

Easy switch between modes

Fabric discovery for LAN Deployments

Fabric controller for LAN and IPFM Deployments

SAN controller for MDS Fibre Channel Deployments

Fabric Discovery

Run fabric discovery for LAN deployments:
Enable inventory, discovery, monitoring only

Enable Cisco Nexus Dashboard's Day 2 operations capabilities without deploying fabric controller

TIP

You can switch anytime from Fabric Discovery to Fabric Controller Mode

Benefit

Deep visibility into deployments

Fabric Controller

Provides fabric management for multiple types of LAN solutions, including VXLAN-EVPN, and traditional 3-tier LAN deployments

Compliance management ensures that network is in sync with intended deployments and allows users to deploy any corrections

Benefit

Most configurations are automatically done following Cisco Best Practices

SAN Controller

Completely redesigned web-based zoning interface to drastically reduce the cycle time for common administration tasks. Provides IVR zoning function as well, all on the same page.

SAN Insights provides useful data to the administrators so they can be fully aware about the fabric status

Benefit

Transition to a web-based configuration method is made easy

Cisco Nexus Dashboard

kubernetes

Orchestrator

Insights

Fabric Controller

Data Broker

SAN Controller
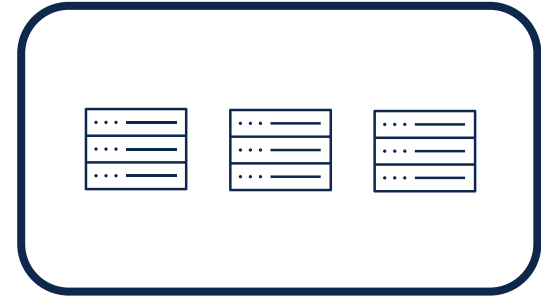
Physical ND Cluster

VM   VM   VM

VM   VM
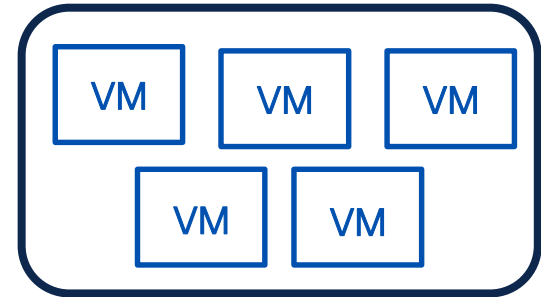
Virtual ND Cluster

Cisco Nexus Dashboard Formats – NDFC

Physical ND Cluster

Each node is a UCS Server with:

2.8GHz AMD CPU
256G RAM
4x2.4TB HDD
960 GB SSD
1.6 TB NVMe drive

TIP For the latest information check the specific scalability guide. 12.1.3b Verified Scalability

Virtual ND Cluster

For NDFC each vND VM must satisfy the following requirements:

Specs	APP
vCPU	16
RAM (GB)	64
SSD (GB)	550

Cisco Nexus Dashboard Scaling – NDFC

Cisco NDFC 12.1(3)

Physical ND Cluster

Full scale for NDFC can be achieved with 3 nodes

Managed mode (VXLAN and BGP fabrics): 500 switches

Managed/Monitor mode (External fabrics): 1000 switches

Overall fabric count: 50

Recommended

Virtual ND Cluster

Full scale for NDFC can be achieved with 5 nodes

Managed mode (VXLAN and BGP fabrics): 400 switches

Managed/Monitor mode (External fabrics): 1000 switches

Overall fabric count: 50

3x vND can support 100 switches in managed mode

VXLAN EVPN (Greenfield)

Switches per Fabric: 200

Overlays: 500 VRF and 2000 Layer-3 Networks OR 2500 Layer-2 Networks

Multi-Site Domain: 30 fabrics

ToR/Leaf: 40 Leaf (VTEP) and 320 ToRs in DC VXLAN EVPN fabric

VXLAN EVPN (Brownfield)

Switches per Fabric: 200

Overlays: 400 VRF and 1050 Layer-3/Layer-2 Networks

Multi-Site Domain: 30 fabrics

TIP In any case at least 3 nodes must be deployed for proper redundancy. 1x vND also supported for Production

Nexus Dashboard
Interface Types

ND Mgmt. Interface

ND Fabric Interface
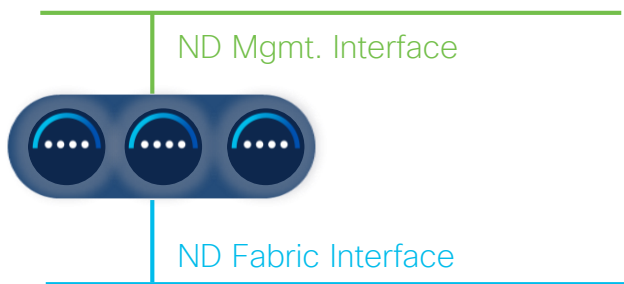
- Each ND node has two interface types:

  - Management Interface: should be dedicated to the management of the ND cluster → connectivity to NTP and DC Proxy servers, Intersight, DNS, ND (and ND Apps) UI access and to perform firmware upgrade (for ND or Apps)

  - Fabric Interface: used for the bring up of the ND cluster (node to node communication) and application to application (NDO, NDI, NDFC, etc.) communication

Cisco Nexus Dashboard Connectivity – NDFC

L2

Data/Fabric Network
bond0br

MAX RTT 50msec

10.1.1.101/24   10.1.1.102/24   10.1.1.103/24

10.2.2.101/24   10.2.2.102/24   10.2.2.103/24
10.2.2.150/24   10.2.2.151/24

MAX RTT 50msec

Management Network
bond1br
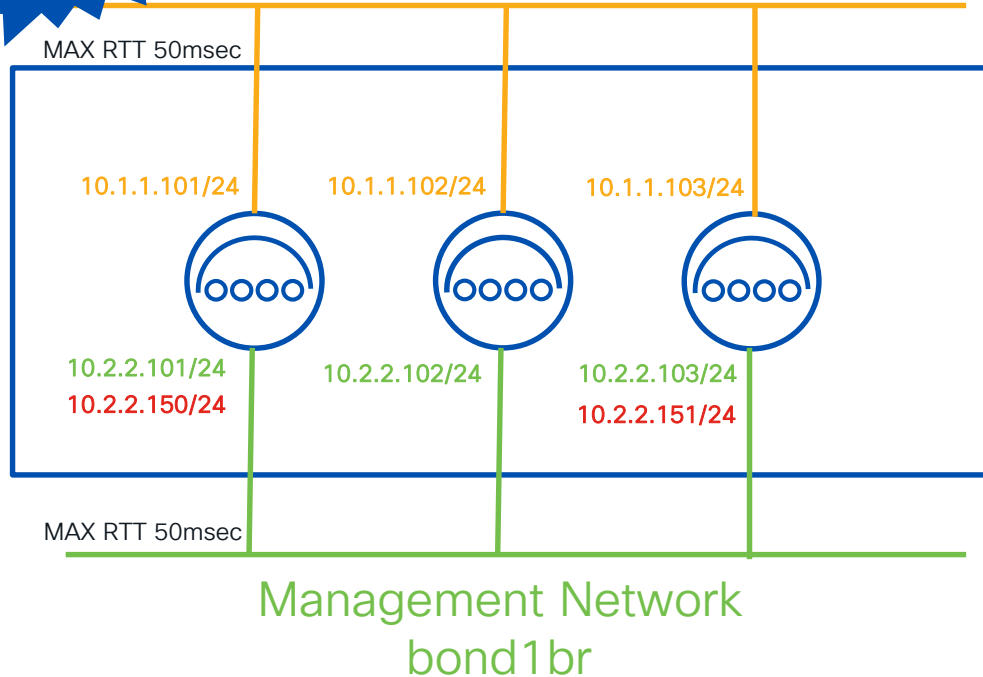
TIP The two interfaces cannot share the same subnet

Intra/Inter APP   PTP

ND Clustering

Switch Access*

DNS   SNMP TRAPS

POAP DHCP
                            NTP
GUI Access

                   DC App Center
CLI via SSH

Intersight

* by default

NDFC Persistent IPs
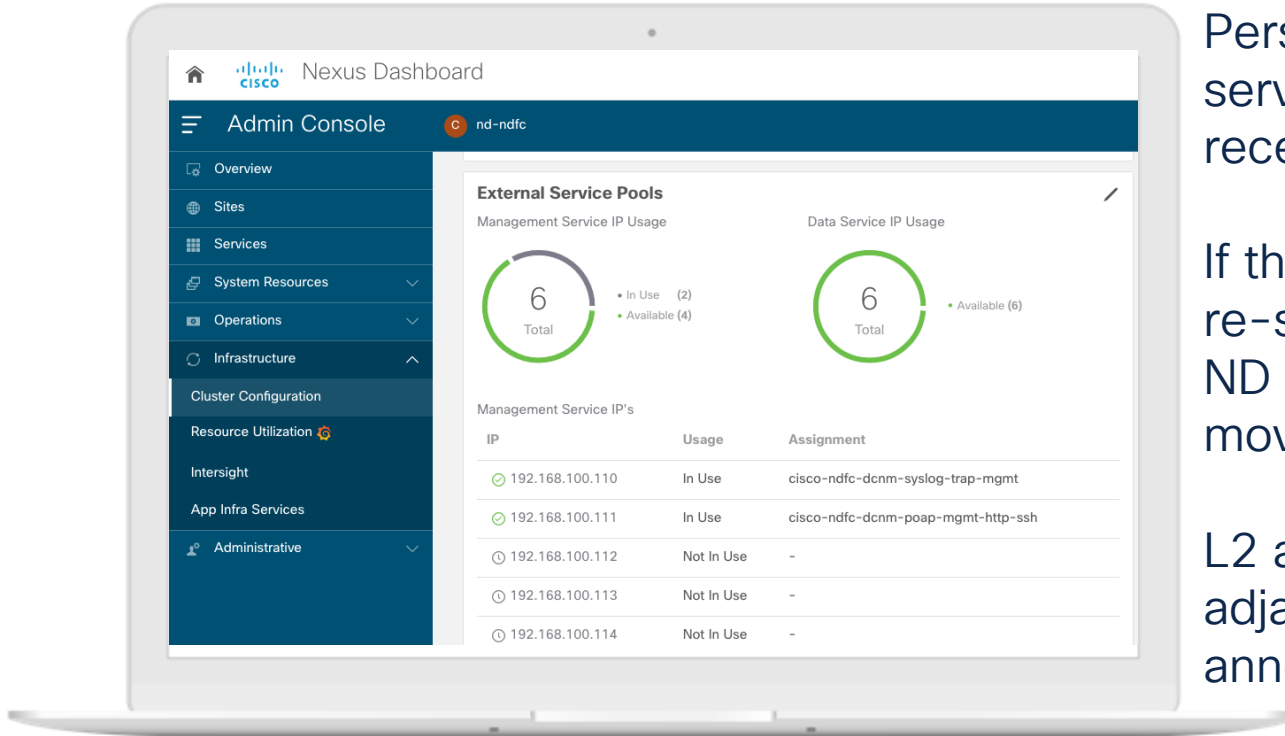
Persistent IPs are tied to a service, like the SNMP trap receiver

If the SNMP trap POD gets re-spawned into a different ND host the sticky IP will be moved there

L2 adjacency uses ARP, L3 adjacency BGP announcements

NDFC Persistent IPs – Normal conditions

Data/Fabric Network
bond0br

10.1.1.101/24   10.1.1.102/24   10.1.1.103/24

10.2.2.101/24   10.2.2.102/24   10.2.2.103/24
10.2.2.150/24   10.2.2.151/24

Management Network
bond1br
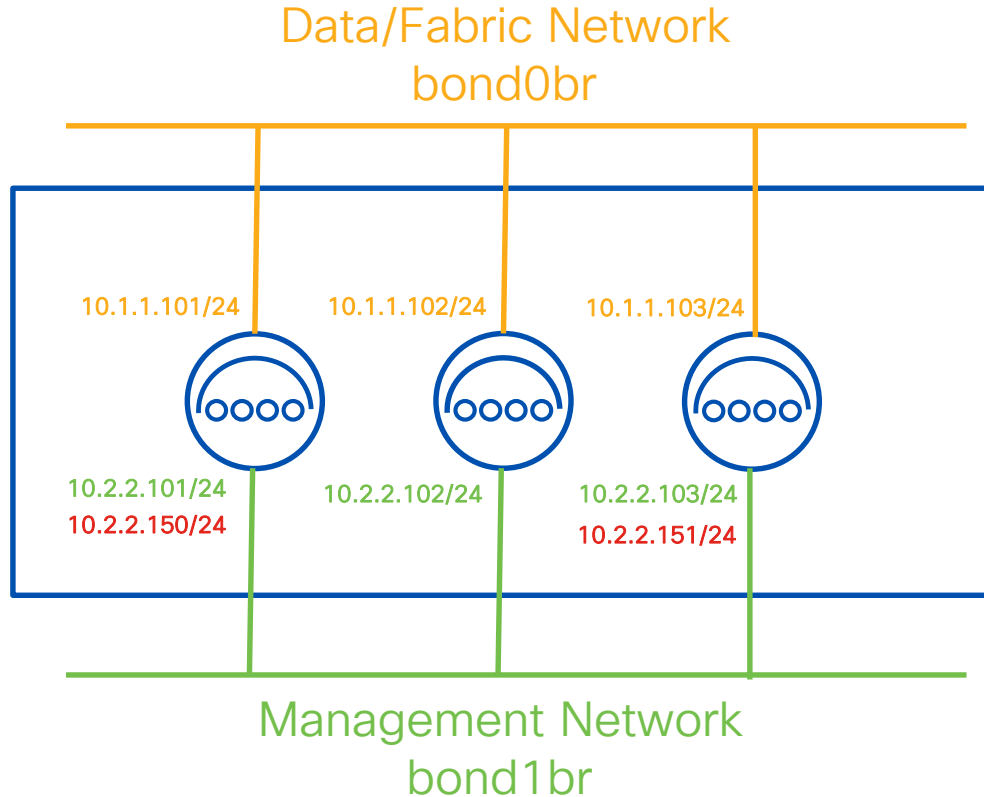
NDFC Persistent IPs – Failover

Data/Fabric Network
bond0br

10.1.1.101/24   10.1.1.102/24   10.1.1.103/24

10.2.2.101/24   10.2.2.102/24   10.2.2.103/24
                  10.2.2.150/24   10.2.2.151/24

Management Network
bond1br
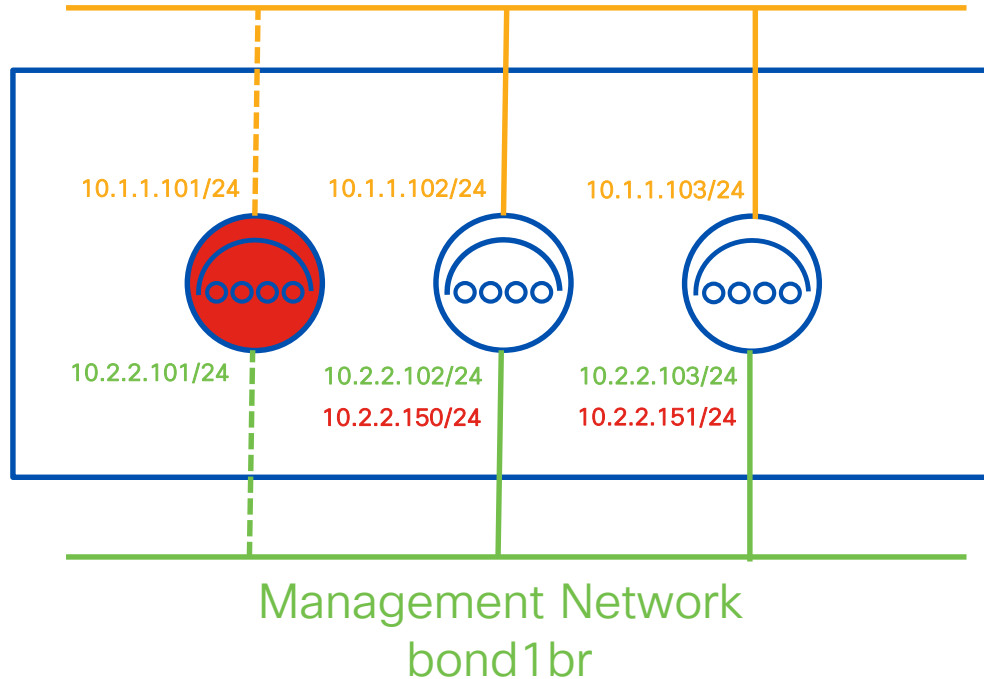
Cisco Nexus Dashboard Connectivity – NDFC

L3

Data/Fabric Network bond0br

MAX RTT 50msec

172.17.20.127/24
10.1.1.101/24

10.1.2.101/24

172.17.20.128/24
10.1.3.101/24

10.2.1.101/24

10.2.2.102/24

10.2.3.103/24

MAX RTT 50msec

Management Network bond1br

This L3 options is valid since 12.1.1e

Each ND node on a different Subnet

BGP Sessions are established for Persistent IP advertisement (no multi-hop)

Persistent IPs must not overlap with ND subnets

Cisco Nexus Dashboard Connectivity – NDFC

Data/Fabric Network bond0br

L3

MAX RTT 50msec

172.17.20.127/24
10.1.1.101/24

10.1.2.101/24

172.17.20.128/24
10.1.3.101/24

10.2.1.101/24

10.2.2.102/24

10.2.3.103/24

MAX RTT 50msec

Management Network bond1br

Cisco Nexus Dashboard Connectivity – NDFC

Data/Fabric Network bond0br

L3

MAX RTT 50msec

172.17.20.127/24

172.17.20.128/24

10.1.1.101/24

10.1.2.101/24

10.1.3.101/24

10.2.1.101/24

10.2.2.102/24

10.2.3.103/24

MAX RTT 50msec

Management Network bond1br

Cisco NDFC Connectivity to the Switches

Use case #1
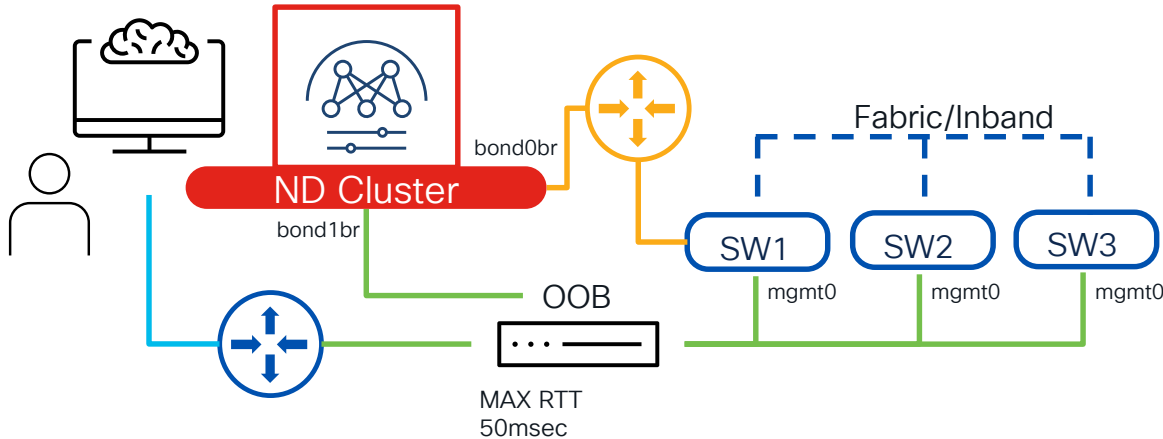
Discovery and Deployment happen via ND Management Interface as that subnet is directly connected

ND Data Interface eventually used for Endpoint Locator Feature (BGP towards Spine RR)

Persistent IPs are allocated on the Management Subnet

Works by default!

ND Mgmt Subnet	ND Data Subnet	Fabric Mgmt 0 Subnet	Fabric Inband Subnet
10.2.2.0/24	10.1.1.0/24	10.2.2.0/24	10.3.3.0/24

Cisco NDFC Connectivity to the Switches

Use case #2
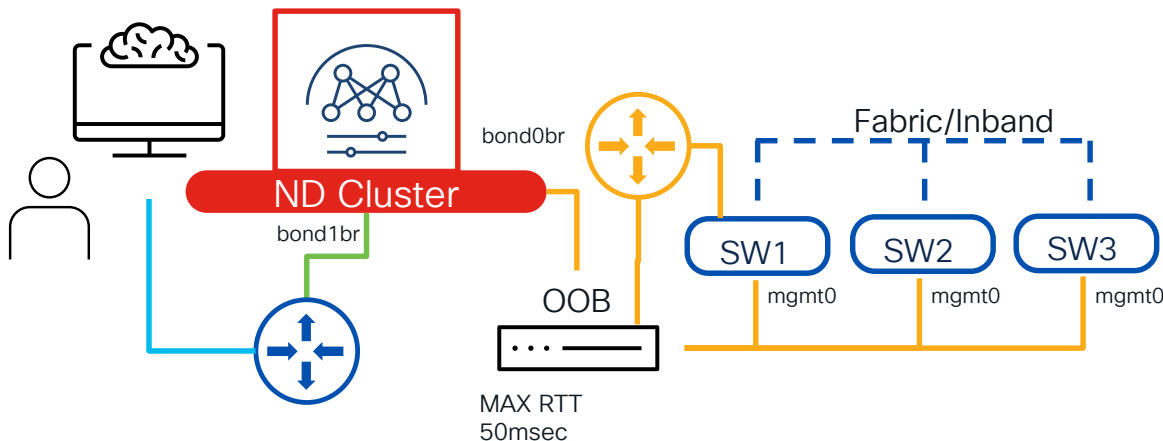
Everything is done over the ND Data Interface as that subnet is directly connected

Persistent IPs are allocated on the Data Subnet

Labels in diagram: bond0br, bond1br, ND Cluster, OOB, MAX RTT 50msec, Fabric/Inband, SW1, SW2, SW3, mgmt0, mgmt0, mgmt0

LAN Device Management Connectivity must be set to Data
(see next slide)

ND Mgmt Subnet	ND Data Subnet	Fabric Mgmt 0 Subnet	Fabric Inband Subnet
10.2.2.0/24	10.1.1.0/24	10.1.1.0/24	10.3.3.0/24

Cisco NDFC Connectivity to the Switches

The change is global for the NDFC Instance

Persistent IPs will be provisioned over ND Data Interface

Settings --> Server Settings --> LAN Device Management Connectivity
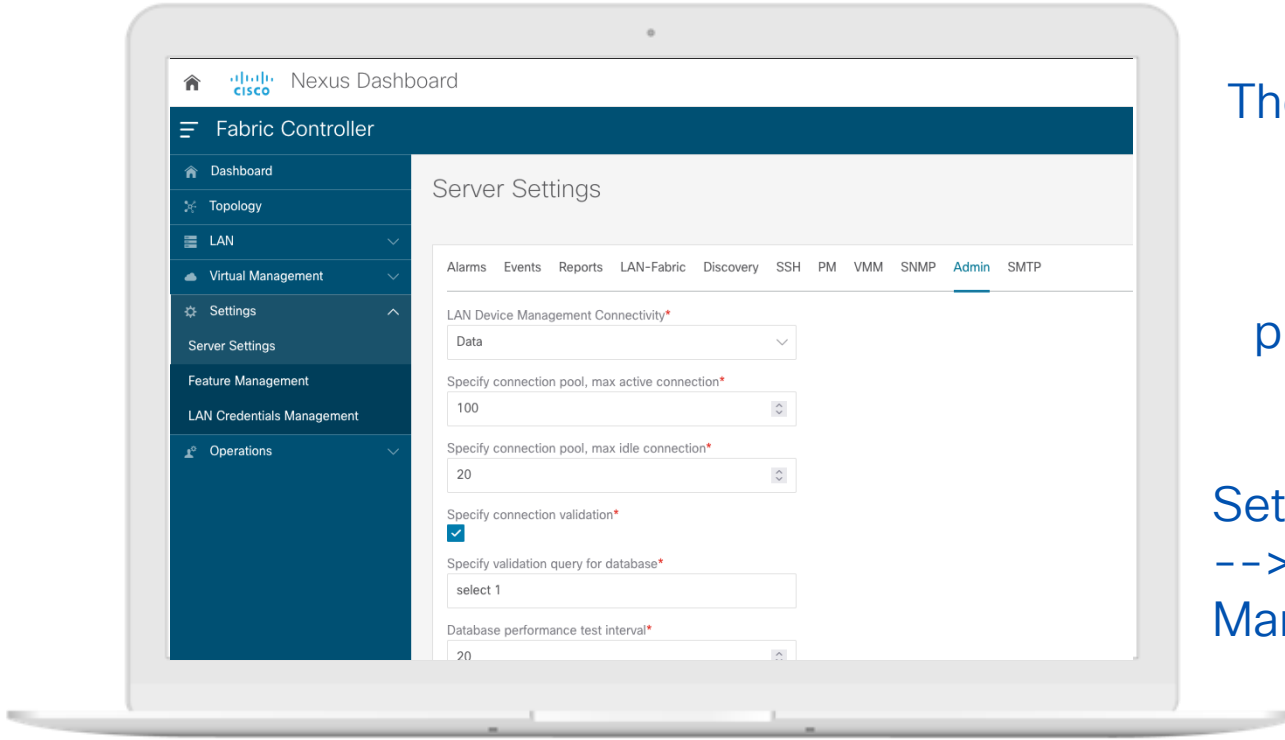
Cisco NDFC Connectivity to the Switches

Use case #3
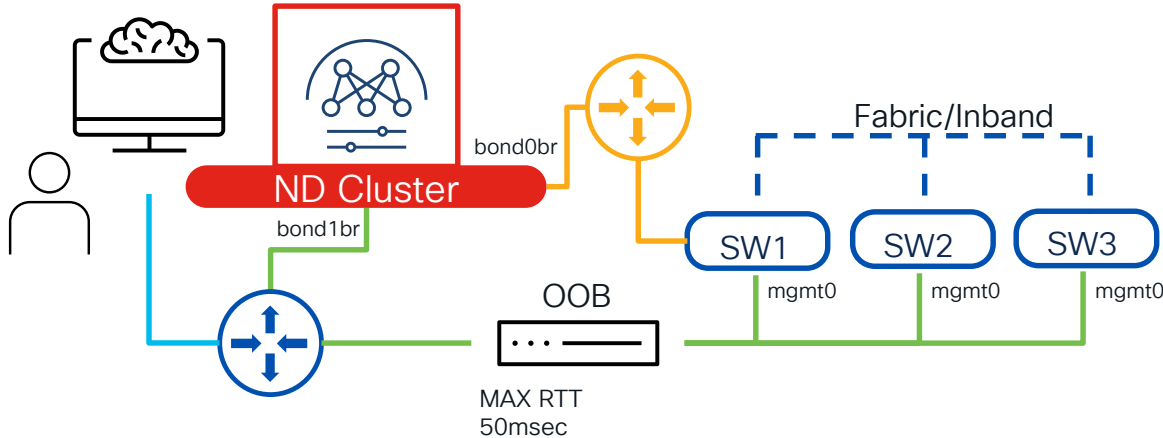
Discovery and Deployment happen via ND Management Interface

ND Data Interface eventually used for Endpoint Locator Feature (BGP towards Spine RR)

Persistent IPs are allocated on the Management Subnet

A static route to 10.4.4.0/24 must be added in ND Management Interface (see next slide)

ND Mgmt Subnet	ND Data Subnet	Fabric Mgmt 0 Subnet	Fabric Inband Subnet
10.2.2.0/24	10.1.1.0/24	10.4.4.0/24	10.3.3.0/24

Cisco NDFC Connectivity to the Switches

Use case #3 continues
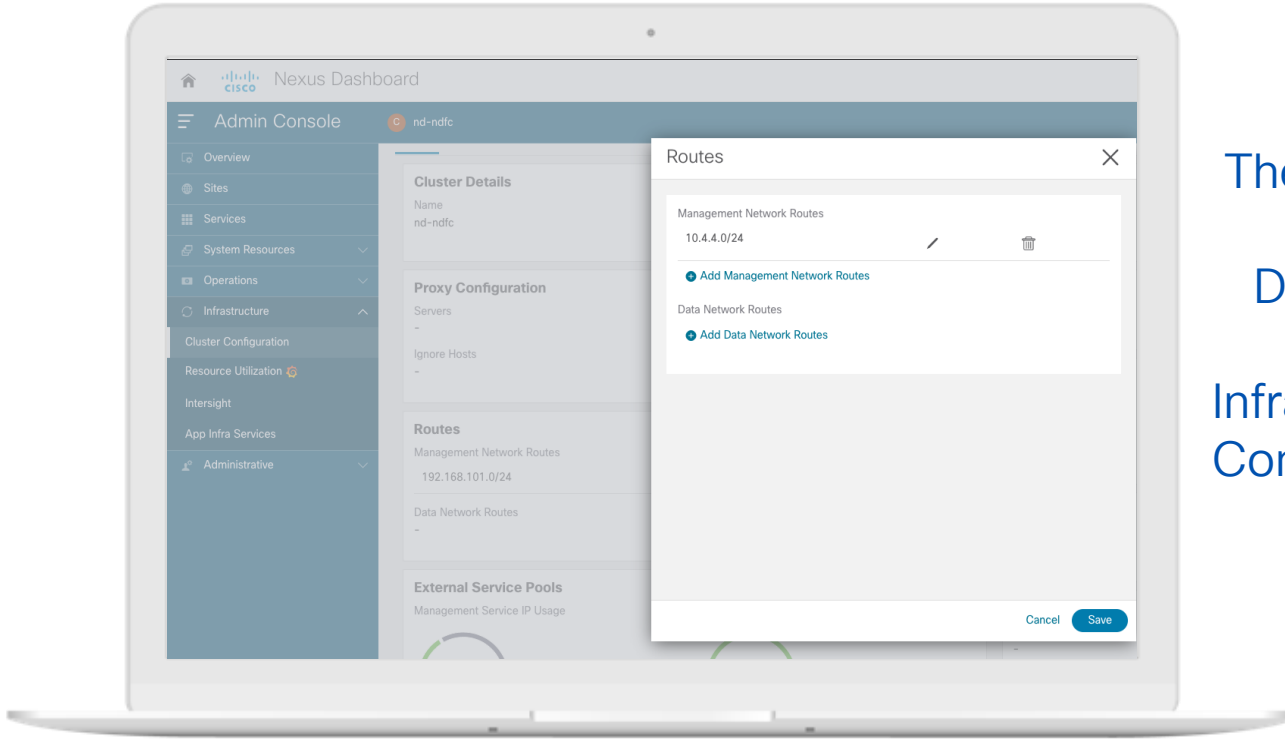
The static route needs to be added in the Nexus Dashboard Control Panel.

Infrastructure --> Cluster Configuration --> Routes

Cisco NDFC Connectivity to the Switches

Use case #4
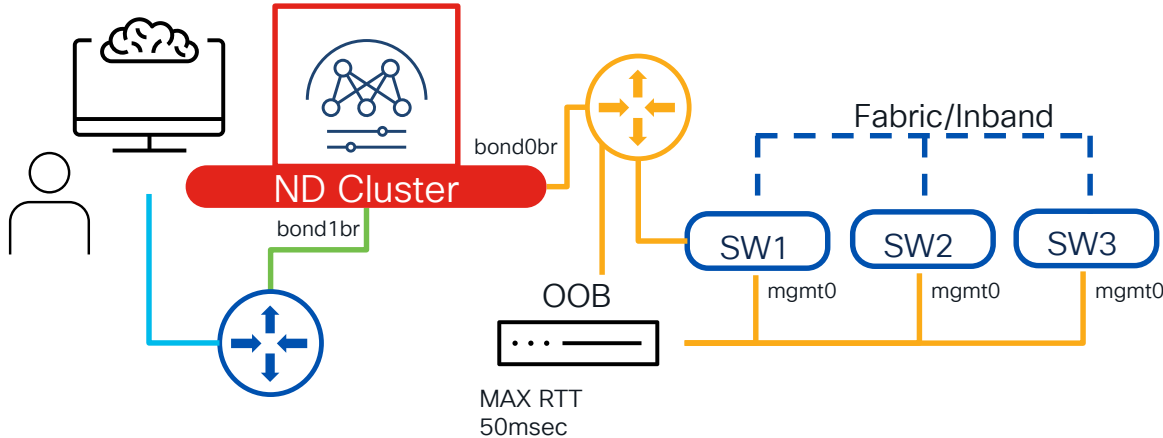
Everything is done over the ND Data Interface as that subnet is directly connected

Persistent IPs are allocated on the Data Subnet

LAN Device Management Connectivity must be set to Data

A static route to 10.4.4.0/24 must be added in ND Data Interface, not for routing but for POAP

Labels in diagram: bond0br, bond1br, ND Cluster, OOB, MAX RTT 50msec, Fabric/Inband, SW1, SW2, SW3, mgmt0, mgmt0, mgmt0

ND Mgmt Subnet	ND Inband Subnet	Fabric Mgmt 0 Subnet	Fabric Inband Subnet
10.2.2.0/24	10.1.1.0/24	10.4.4.0/24	10.3.3.0/24

Why do YOU need NDFC?
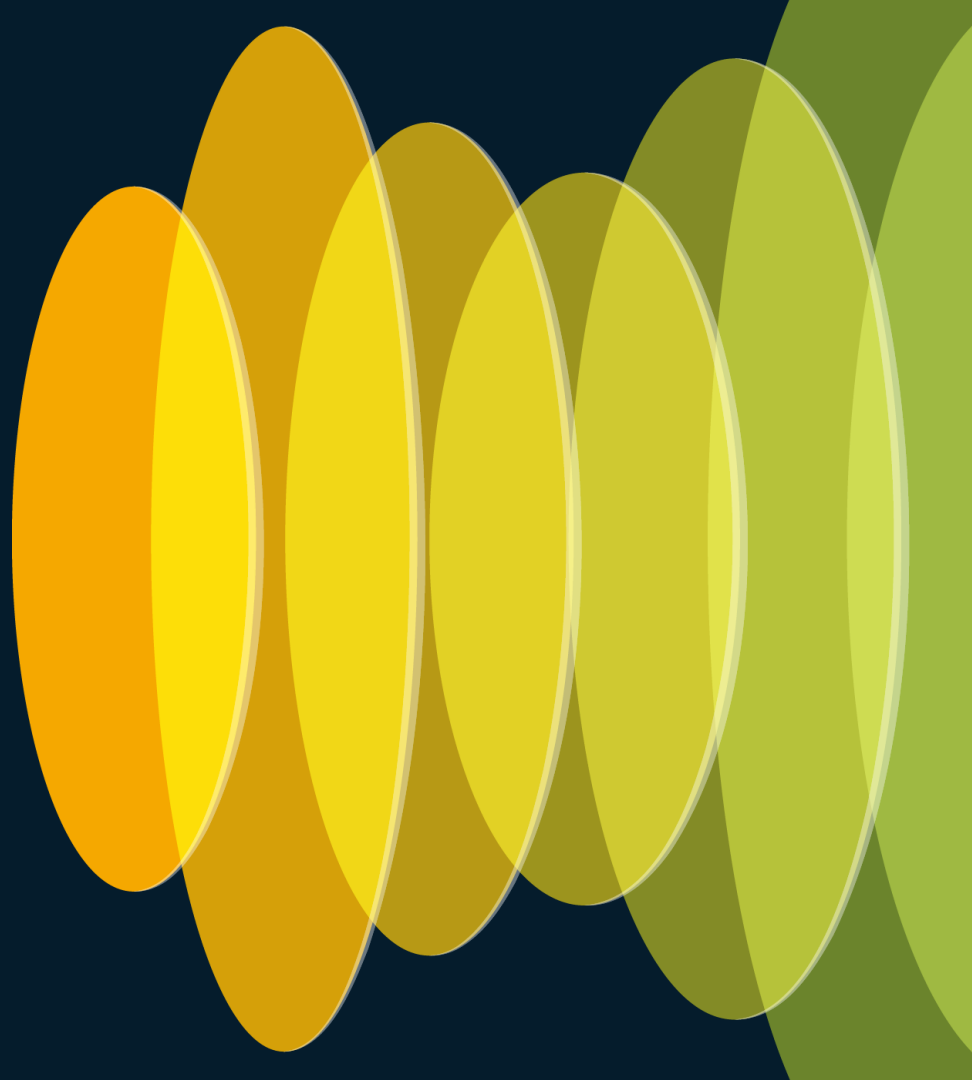
Why NDFC?
Multi-Architecture

Layer-3 MSDC Fabric

Campus VXLAN EVPN

Data Center VXLAN EVPN

VXLAN EVPN Multisite

Traditional 3-Tier LAN
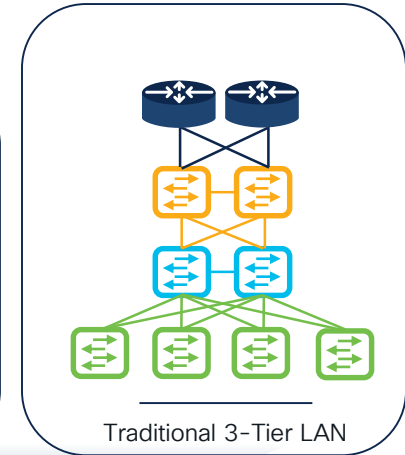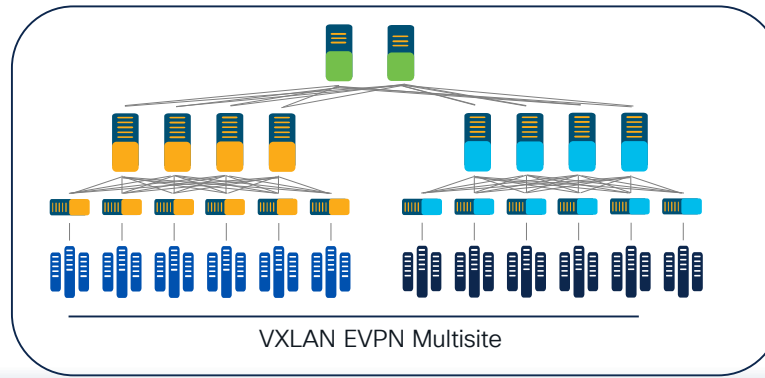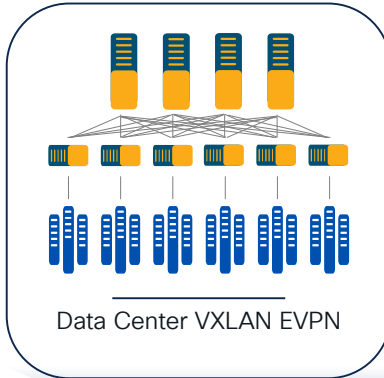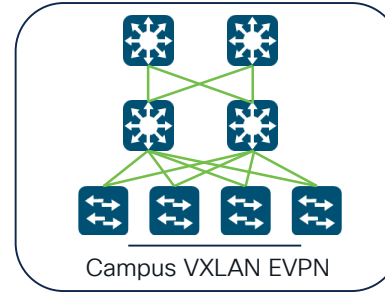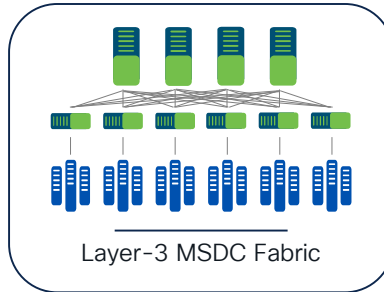
Why NDFC?
Multi-Topology, Multi-Protocol

Rich set of control plane and data plane possibilities available

VXLAN BGP EVPN

vPC

BGP Routed

MPLS

FHRP

STP

VLAN

IS-IS

OSPF

FabricPath

IP

IPFM
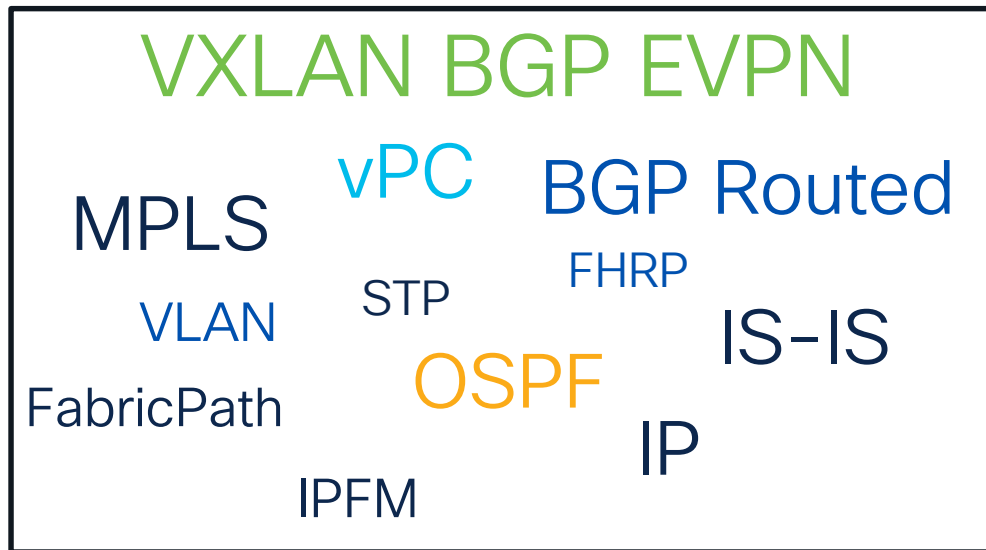
Why NDFC?
Multi-Domain, Multi-Platform

NX-OS Nexus 9000 and 3000

IOS-XE Catalyst 9000

IOS-XR ASR 9000

NX-OS Nexus 7000

IOS-XE ASR 1000

TIP
Supported Hardware and Software might vary depending on NDFC version
Check compatibility matrix 12.1.3b

Why NDFC?

In a nutshell...

Step into SDN via VXLAN BGP EVPN

Config and Compliance across Cisco Products

Single Source of Truth

End to End Automation

Single Pane of Glass for Day-0/Day-1 Provisioning

Multi-OS management and support

Simplify Complex Network Operations

Automate, Manage, and Interconnect Multi-Fabric topologies
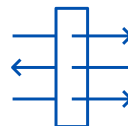
Layer-3 Boundary across Zones, L2/L3 across IOS-XE, NXOS, and Multicast Overlay

Programmability and Orchestration

Automate VXLAN EVPN deployments

Provision a new fabric in minutes

Un-provisioned switches
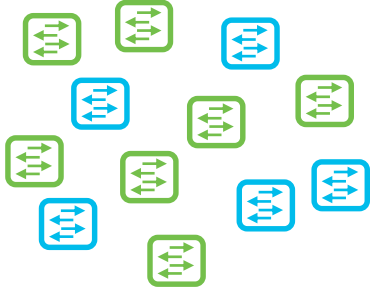
Cisco best practice implemented
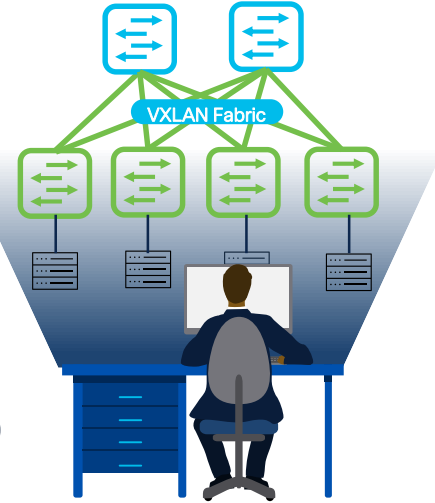
VXLAN Fabric

Fast, automated process

Benefit

Accelerate fabric deployments | Automated consistency | Minimize risk | Support for both Greenfield and Brownfield deployment

Manage and Deploy
VXLAN BGP EVPN
with NDFC

CISCO *Live!*
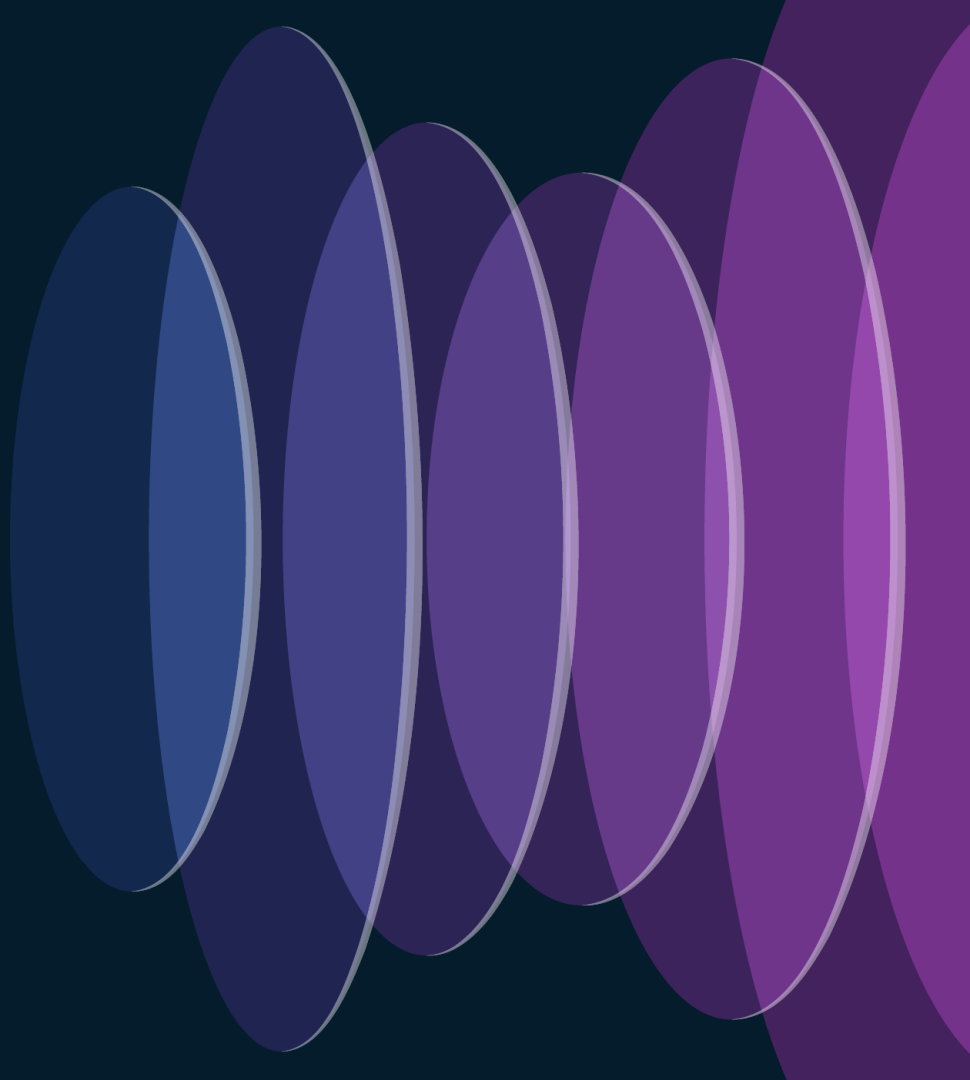
VXLAN BGP EVPN Greenfield

Not on VXLAN EVPN Today?

NDFC Fabric Controller Mode	Build VXLAN fabric in few minutes	Templates already embed best practices	IP addresses, overlay pool, routing profiles, replication attributes –all taken care by NDFC

Step 2
Discover
Import switches with POAP or Day-0 config
Define switch Roles (Border, Leaf, Spine, etc)
[Optional] Create vPC pairs

Step 1
Create
Define fabric settings (Underlay, Overlay) - AS#, Replication Mode, IGP, IP Pools, etc.

Step 3
Recalculate and Deploy
Generates config based on intent
Preview side by side diffs

Cisco best practice implemented
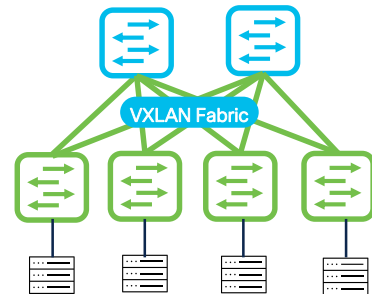
VXLAN Fabric

Fast, automated process

NDFC Day-0:
VXLAN EVPN
Underlay
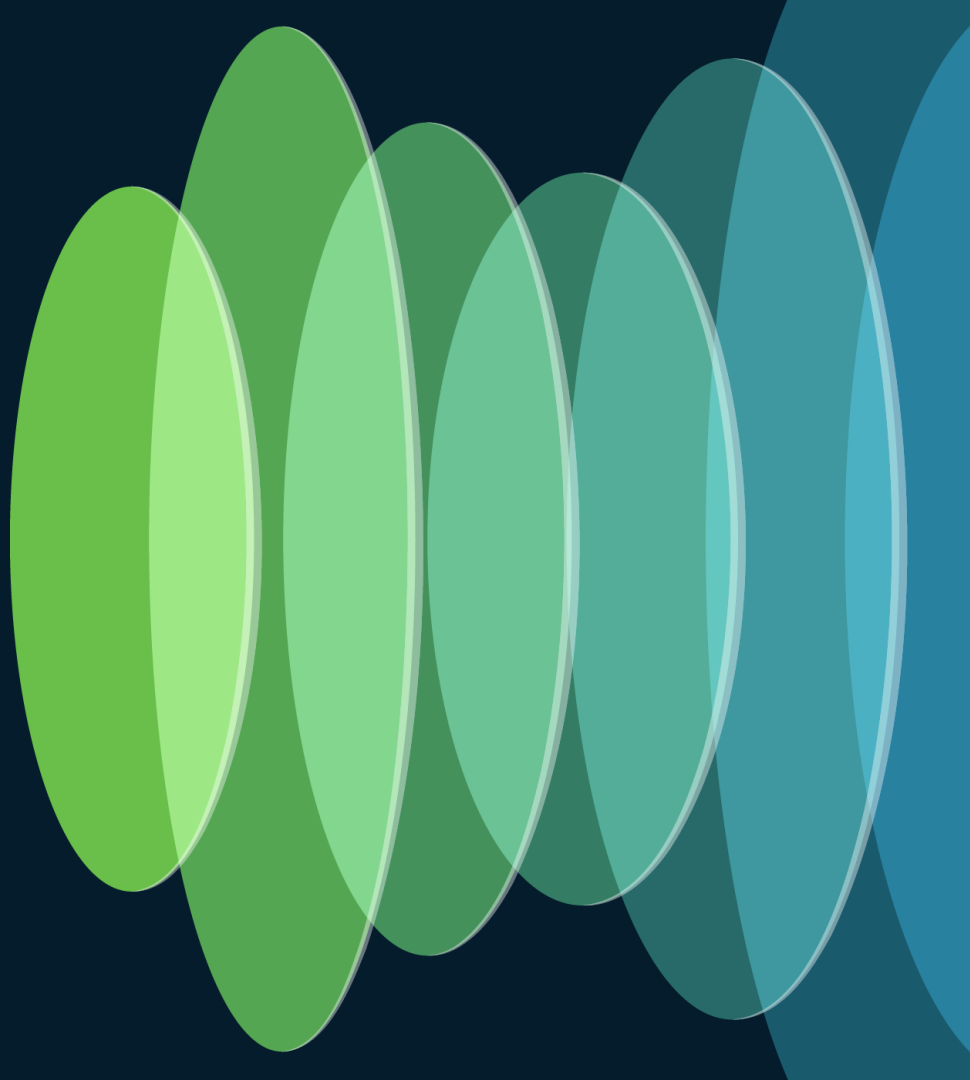
Step1 -> Create a Fabric
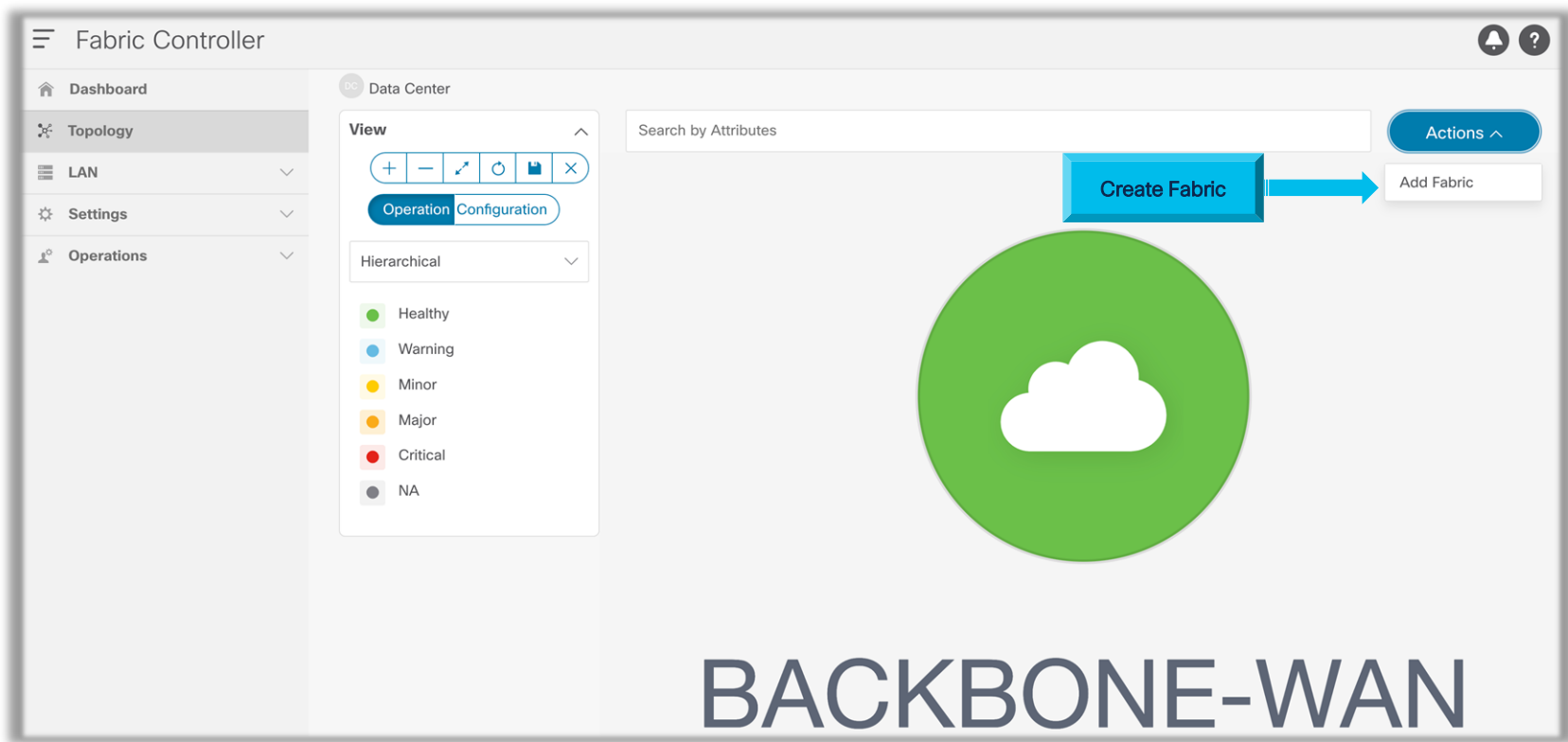
Step1 -> Create a Fabric
(continued)

Pick Fabric

Data Center VXLAN EVPN >

General Parameters   Replication   VPC   Protocols   Advanced   Resources   Manageability   Bootstrap

BGP ASN

	1-4294967295 | 1-65535[.0-65535] It is a good practice to have a unique ASN for each Fabric.

BGP ASN
(arrow) ➔ **BGP ASN is required**

Enable IPv6 Underlay
☐ If not enabled, IPv4 underlay is used

VXLANv4
or
VXLANv6
(arrow) ➔
Enable IPv6 Link-Local Address
☐ If not enabled, Spine-Leaf interfaces will use global IPv6 addresses

Fabric Interface Numbering*

p2p   Numbered(Point-to-Point) or Unnumbered

Underlay Subnet IP Mask*

30   (arrow) ⬅   **Underlay IP > /30 or /31**

Underlay Subnet IPv6 Mask

Select an Option   Mask for Underlay Subnet IPv6 Range

Underlay Routing Protocol*

ospf   (arrow) ⬅   **IGP > OSPF or ISIS**

Route-Reflectors*

RR > 2 or 4   (arrow) ➔
2   Number of spines acting as Route-Reflectors

Anycast Gateway MAC*

2020.0000.00aa   (arrow) ⬅   **Distributed Anycast Gateway**

Enable Performance Monitoring
☐   (arrow) ⬅   **PM Metrics**

Step1 -> Create a Fabric
(continued)

**BUM >
Multicast or
Ingress
Replication**

**L2VNI
Multicast
Group**

RP > 2 or 4

General Parameters   **Replication**   VPC   Protocols   Advanced   Resources   Manageability   Bootstrap

Replication Mode*

Multicast

Replication Mode for BUM Traffic

Multicast Group Subnet*

239.1.1.0/25

Multicast pool prefix between 8 to 30. A multicast group IP from this pool is used for BUM traffic for each overlay network.

Enable Tenant Routed Multicast (TRM)

☐   For Overlay Multicast Support In VXLAN Fabrics

Default MDT Address for TRM VRFs

Default Underlay Multicast group IP assigned for every overlay VRF.

Rendezvous-Points*

2

Number of spines acting as Rendezvous-Point (RP)

RP Mode*

asm

Multicast RP Mode

Underlay RP Loopback Id*

254

(Min:0, Max:1023)

Step1 -> Create a Fabric
(continued)

Cisco's Best Practice Configuration Templates

VXLAN Overlay Mode > CLI or Config-Profile

General Parameters **Replication** **VPC** **Protocols** **Advanced** **Resources** **Manageability** **Bootstrap**

VRF Template*
Default_VRF_Universal   Default Overlay VRF Template For Leafs

Network Template*
Default_Network_Universal   Default Overlay Network Template For Leafs

VRF Extension Template*
Default_VRF_Extension_Universal   Default Overlay VRF Template For Borders

Network Extension Template*
Default_Network_Extension_Universal   Default Overlay Network Template For Borders

Overlay Mode
cli   VRF/Network configuration using config-profile or CLI, default is config-profile

config-profile

cli ✓   Enable PVLAN on switches except spines and super spines

PVLAN Secondary Network Template
Select an Option   Default PVLAN Secondary Network Template

Site Id
65009   For EVPN Multi-Site Support (Min:1, Max: 281474976710655). Defaults to Fabric ASN

Intra Fabric Interface MTU*
9216   (Min:576, Max:9216). Must be an even number

Layer 2 Host Interface MTU*
9216   (Min:1500, Max:9216). Must be an even number

Unshut Host Interfaces by Default
☑

Step1 -> Create a Fabric
(continued)

General Parameters | Replication | VPC | Protocols | Advanced | **Resources** | Manageability | Bootstrap

Manual Underlay IP Address Allocation

☐ Checking this will disable Dynamic Underlay IP Address Allocations

Underlay Routing Loopback IP Range*
10.2.0.0/22 ← **Router ID**

Underlay VTEP Loopback IP Range*
10.3.0.0/22 ← **VTEP IP**

Underlay RP Loopback IP Range*
10.254.254.0/24 ← **RP IP**

Underlay Subnet IP Range*
10.4.0.0/16 ← **P2P Underlay IP**

L2VNI Label →

Layer 2 VXLAN VNI Range*
30000-49000 Overlay Network Identifier Range (Min:1, Max:16777214)

L3VNI Label →

Layer 3 VXLAN VNI Range*
50000-59000 Overlay VRF Identifier Range (Min:1, Max:16777214)

Network VLAN Range*
2300-2999 Per Switch Overlay Network VLAN Range (Min:2, Max:4094)

VRF VLAN Range*
2000-2299 Per Switch Overlay VRF VLAN Range (Min:2, Max:4094)

Subinterface Dot1q Range*
2-511 Per Border Dot1q Range For VRF Lite Connectivity (Min:2, Max:4093)

VRF_LITE Handoff →

VRF Lite Deployment*
Back2Back&ToExternal VRF Lite Inter-Fabric Connection Deployment Options. If 'Back2Back&ToExternal' is selected, VRF Lite IFCs are auto created between border devices of two Easy Fabrics, and between border devices in Easy Fabric and edge routers in External Fabric. The IP address is taken from the 'VRF Lite Subnet IP Range' pool.

Auto Deploy for Peer
☑ Whether to auto generate VRF LITE sub-interface and BGP peering configuration on managed neighbor devices. If set, auto created VRF Lite IFC links will have 'Auto Deploy for Peer' enabled.

VRF_LITE IP Range →

VRF Lite Subnet IP Range*
10.33.0.0/16 Address range to assign P2P Interfabric Connections

VRF Lite Subnet Mask*
30 (Min:8, Max:31)

L4-L7 Service Network →

Service Network VLAN Range*
3000-3199 Per Switch Overlay Service Network VLAN Range (Min:2, Max:4094)

Route Map Sequence Number Range*
1-65534 (Min:1, Max:65534)

Step1 -> Create a Fabric
(continued)

| General Parameters | Replication | VPC | Protocols | Advanced | Resources | Manageability | **Bootstrap** |

Enable Bootstrap
☑
Automatic IP Assignment For POAP

Enable Local DHCP Server
☑
Automatic IP Assignment For POAP From Local DHCP Server

DHCP Version

DHCPv4 ⌄

DHCP Scope Start Address*

192.168.101.81
Start Address For Switch POAP

DHCP Scope End Address*

192.168.101.91
End Address For Switch POAP

Switch Mgmt Default Gateway*

192.168.101.254
Default Gateway For Management VRF On The Switch

Switch Mgmt IP Subnet Prefix*

24
(Min:8, Max:30)

NDFC Built-In Bootstrap POAP Services. Supports OOB and In band POAP

Step2 -> Add Switches

Step2 -> Add Switches
(continued)

Switch Addition Mechanism*
- ⦿ Discover   ◯ Bootstrap(POAP)   ◯ Pre-provision

Seed Switch Details

Seed IP*

192.168.101.24

Ex: "2.2.2.20" or "10.10.10.40-60" or "2.2.2.20, 2.2.2.21"

← **Switch Mgmt0 IP**

Authentication Protocol*

MD5 ⌄

Username*

admin

Password*

•••••••• 👁

← **Switch Discovery Credentials**

Max Hops*

2

← **Switch Hops based on CDP**

Preserve Config

☐

← **VXLAN Greenfield or Brownfield**

Unchecking this will clean up the configuration on switch(es)

Discover Switches

Close   **Discover Switches**

Step2 -> Add Switches
(continued)

Seed Switch Details

Fabric	Switch	Authentication Protocol	Username
DC1	192.168.1	MD5	admin

Password	Max Hops		Preserve config
● Set	2	Switch Inventory Management	● Disabled

← Back

Discovery Results

Filter by attributes

	Switch Name	Serial Number	IP Address	Model	Version	Status	Progress
☑	LEAF3-DC1	9E485UOMX0B	192.168.101.23	N9K-C9300v	9.3(8)	● Manageable	
☑	BORDER-DC1	9C4QW0VHSLO	192.168.101.26	N9K-C9300v	9.3(8)	● Manageable	
☑	SPINE1-DC1	9PHNC61N225	192.168.101.24	N9K-C9300v	9.3(8)	● Manageable	Add Switches
☑	LEAF1-DC1	9FGRDPPVWNX	192.168.101.21	N9K-C9300v	9.3(8)	● Manageable	
☐	BACKBONE-ROUTER	9WD2A8WRKVQ	192.168.101.27	N9K-C9300v	9.3(8)	● Already Managed In BACKBONE-WAN	

Close   Add Switches

Step3 -> Set Role

Step4 -> VPC Pairing
(optional)

Hierarchical Topology

NET — Networks (1)

VRF — VRFs (1)

BACKBONE-WAN

BORDER-DC1

SPINE1-DC1   SPINE2-DC1

LEAF3-DC1   LEAF1-DC1   LEAF2-DC1

Leaf VPC Pairing

LEAF1-DC1

Detailed View

Preview Config

Deploy Config

Discovery ▸

Set Role

Manage Interfaces

Manage Policies

vPC Pairing

ToR/Access Pairing

More ▸
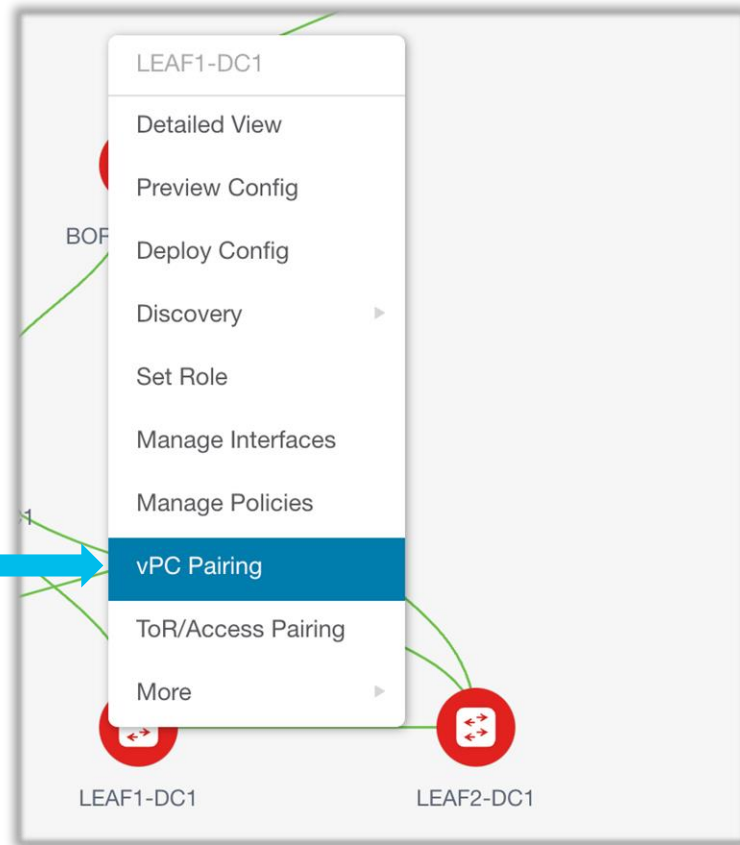
LEAF1-DC1   LEAF2-DC1

Step4 -> VPC Pairing
(optional)

Select vPC Peer for LEAF1-DC1

Virtual Peerlink

Filter by attributes

	Device	Recommended	Reason	Serial Number	IP Address
○	SPINE2-DC1	False	Switches have different roles	922ANP25GML	192.168.101.25
◉	LEAF2-DC1	False	N9K-C9300v doesn't support Virtual Fabric Peering	988KWTIDPZ2	192.168.101.22
○	SPINE1-DC1	False			
○	BORDER-DC1	False			
○	LEAF3-DC1	False			

Select vPC Peer for LEAF1-DC1

Virtual Peerlink

Filter by attributes

	Device	Recommended	Reason	Serial Number	IP Address
◉	LEAF2-DC1	True	Switches are connected and have same role	988KWTIDPZ2	192.168.101.22
○	SPINE2-DC1	False	Switches have different roles	922ANP25GML	192.168.101.25
○	SPINE1-DC1	False	Switches have different roles	9PHNC61N225	192.168.101.24
○	BORDER-DC1	False	Switches have different roles	9C4QW0VHSLO	192.168.101.26
○	LEAF3-DC1	False	Switches are not connected	9E485UOMX0B	192.168.101.23

Cancel   Save

Step5 -> Recalculate and Deploy
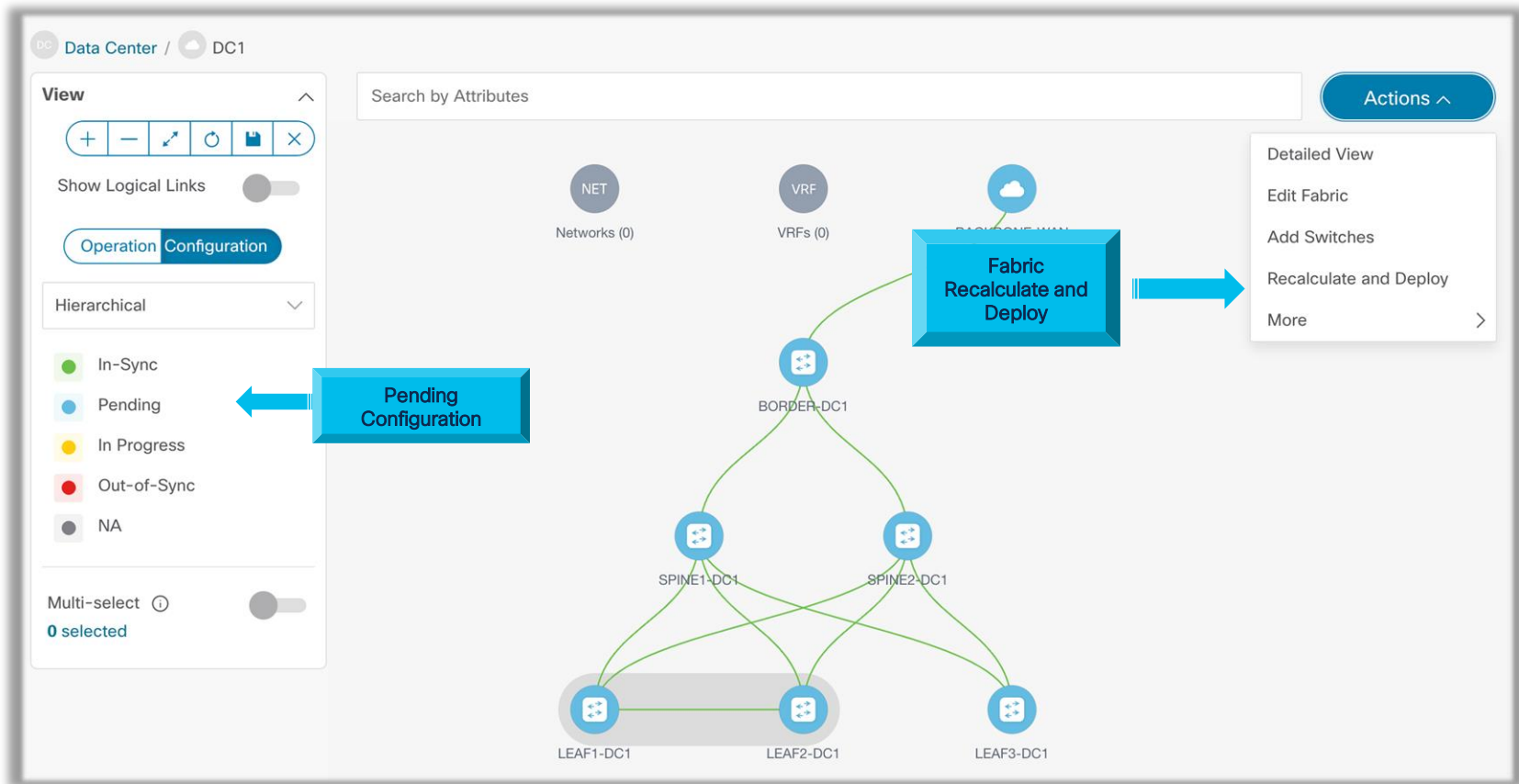
Step5 -> Recalculate and Deploy
(continued)

Deploy Configuration - DC1

1 Config Preview   2 Deploy Progress

Filter by attributes   Resync All

Switch Name	IP Address		Number	Fabric Status	Pending Config	Status Description	Progress	Resync Switch
BORDER-DC1	192.168.101.26		W0VHSLO	● Out-Of-Sync	357 Lines	Out-of-Sync	————	Resync
LEAF3-DC1	192.168.101.23	leaf	9E485UOMX0B	● Out-Of-Sync	523 Lines	Out-of-Sync	————	Resync
LEAF2-DC1	192.168.101.22	leaf	988KWTIDPZ2	● Out-Of-Sync	549 Lines		————	Resync
LEAF1-DC1	192.168.101.21	leaf	9FGRDPPVWNX	● Out-Of-Sync	549 Lines	Out-of-Sync	————	Resync
SPINE2-DC1	192.168.101.25	spine	922ANP25GML	● Out-Of-Sync	341 Lines	Out-of-Sync	————	Resync
SPINE1-DC1	192.168.101.24	spine	9PHNC61N225	● Out-Of-Sync	349 Lines	Out-of-Sync	————	Resync

Fabric Configuration status

Pending Configuration

Close   Deploy All

Step5 -> Recalculate and Deploy
(continued)

Pending Config - DC1 - SPINE2-DC1

Pending Config   Side-by-Side Comparison

```
feature ngoam
feature nxapi
feature ospf
feature pim
ipv6 switch-packets lla
nv overlay evpn
feature lldp
feature bgp
feature nv overlay
ip pim anycast-rp 10.254.254.1 10.2.0.2
ip pim anycast-rp 10.254.254.1 10.2.0.6
ip pim rp-address 10.254.254.1 group-list 239.1.1.0/25
ip pim ssm range 232.0.0.0/8
ngoam install acl
nxapi http port 80
nxapi https port 443
snmp-server host 10.124.124.250 traps version 2c public udp-port 2162
router bgp 65100
  router-id 10.2.0.2
  neighbor 10.2.0.1
    remote-as 65100
    update-source loopback0
    address-family l2vpn evpn
      send-community both
      route-reflector-client
      exit
      ..
```

Spine related features

Spine RR/RP function

Spine EVPN RR Client

Pending Config - DC1 - SPINE2-DC1

Pending Config   Side-by-Side Comparison

```
interface ethernet1/1
  no switchport
  ip address 10.4.0.9/30
  description connected-to-LEAF3-DC1-Ethernet1/2
  mtu 9216
  ip router ospf UNDERLAY area 0.0.0.0
  ip ospf network point-to-point
  ip pim sparse-mode
  no shutdown
interface ethernet1/2
  no switchport
  ip address 10.4.0.13/30
  description connected-to-BORDER-DC1-Ethernet1/3
  mtu 9216
  ip router ospf UNDERLAY area 0.0.0.0
  ip ospf network point-to-point
  ip pim sparse-mode
  no shutdown
```

CDP Link & IGP configs

Step5 -> Recalculate and Deploy
(continued)

Pending Config - DC1 - LEAF1-DC1

Pending Config   Side-by-Side Comparison

```
vpc domain 1
  ip arp synchronize
  peer-gateway
  peer-switch
  delay restore 150
  peer-keepalive destination 192.168.101.22 source 192.168.101.21
  auto-recovery reload-delay 360
  ipv6 nd synchronize
interface port-channel500
  switchport
  switchport mode trunk
  spanning-tree port type network
  description "vpc-peer-link LEAF1-DC1--LEAF2-DC1"
  no shutdown
  vpc peer-link
interface ethernet1/2
  channel-group 500 force mode active
  description "PO 500 (vpc-peer-link) member LEAF1-DC1-Ethernet1/2 t
  no shutdown
```

VPC Best Practice Configs

Pending Config - DC1 - SPINE2-DC1

Pending Config   Side-by-Side Comparison

Running Config

```
12  hostname SPINE2-DC1
13  icam monitor scale
14  interface ethernet1/1



23  interface ethernet1/10
```

Switch running config v NDFC Intent

Expected Config

```
 8  feature nv overlay
 9  feature nxapi
10  feature ospf
11  feature pim
12  hostname SPINE2-DC1

14  interface ethernet1/1
15    description connected-to-LEAF3-DC1-Ethernet1/2
16    ip address 10.4.0.9/30
17    ip ospf network point-to-point
18    ip pim sparse-mode
19    ip router ospf UNDERLAY area 0.0.0.0
20    mtu 9216
21    no shutdown
22    no switchport
23  interface ethernet1/10
24    mtu 9216
25    no switchport
26    shutdown
```

NDFC VXLAN EVPN Topology View
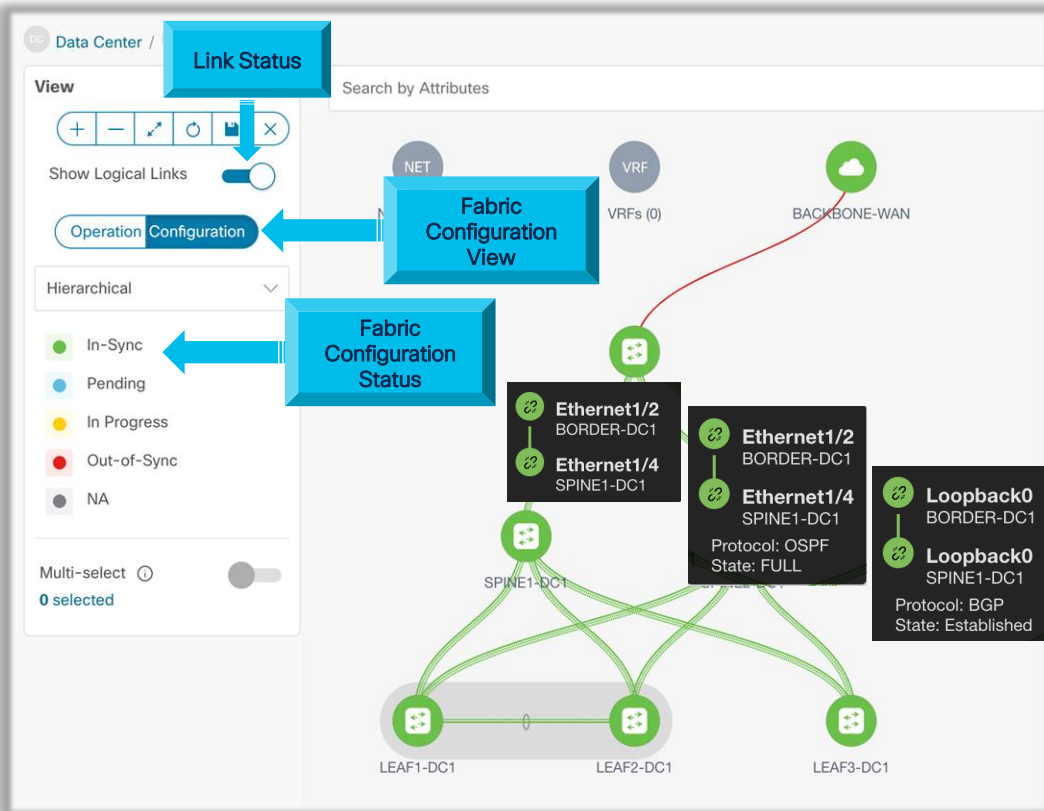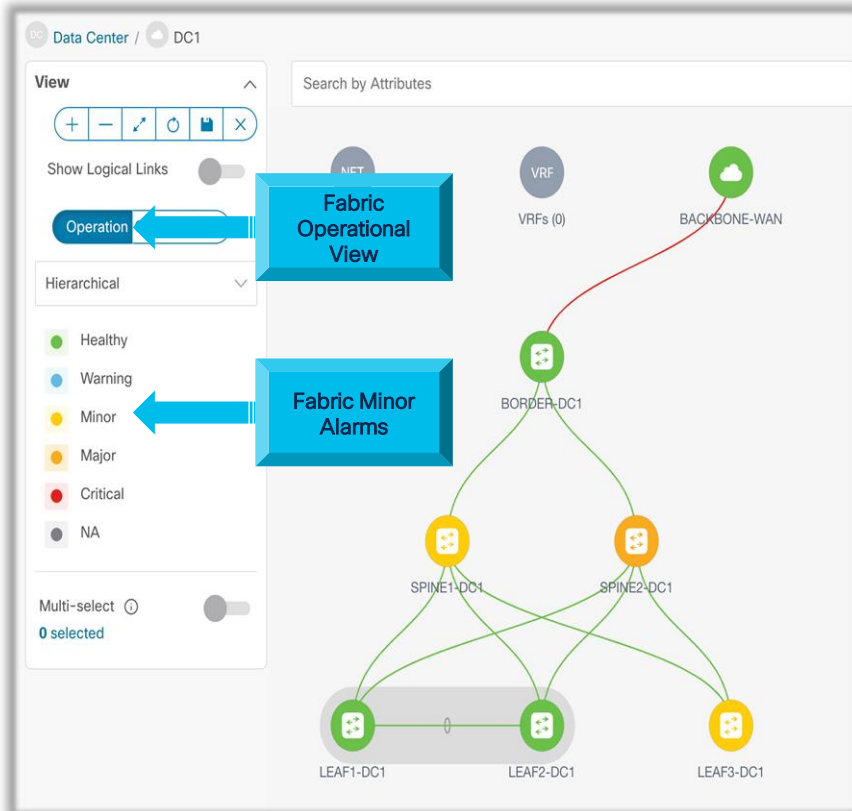
NDFC Day-1:
VXLAN EVPN
Overlay
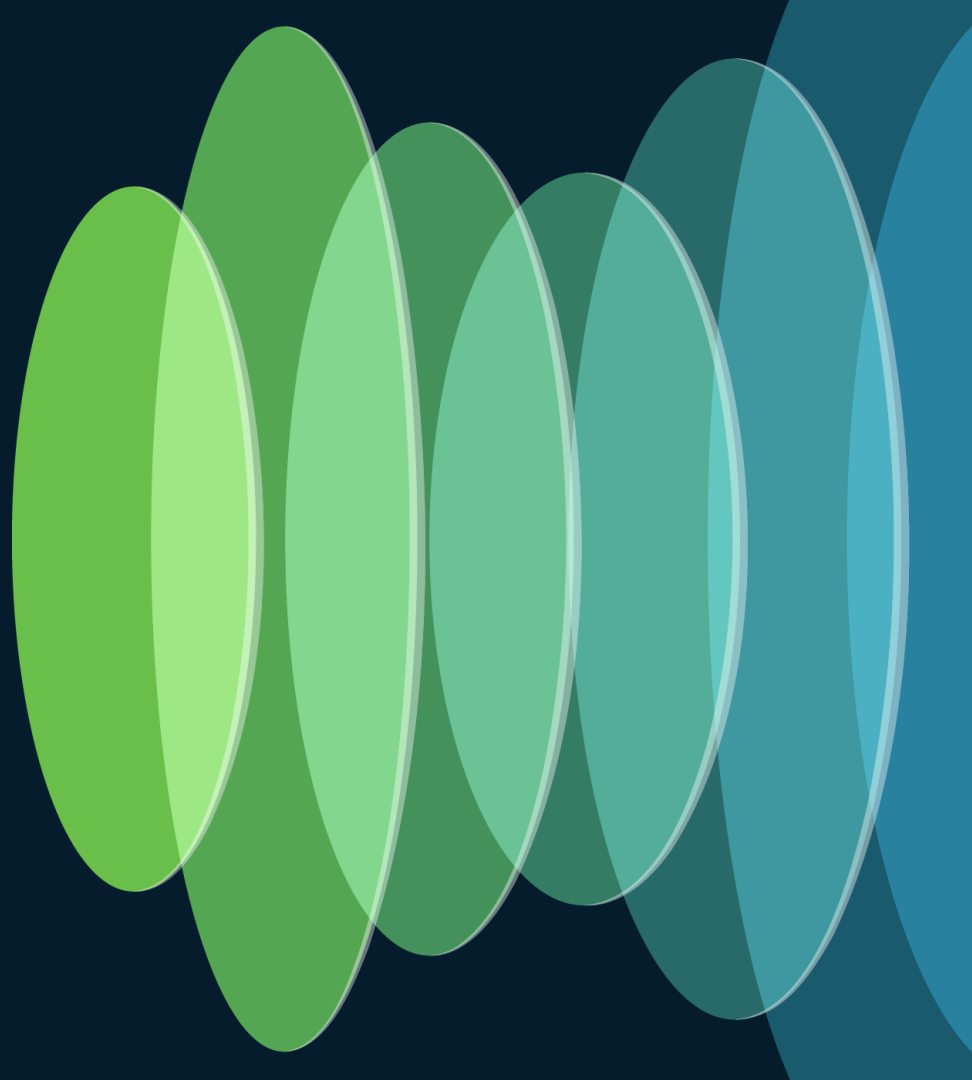
Deploy Network and VRF
in Data Center VXLAN EVPN

Have VXLAN EVPN fabric ready

Create Network and VRF from VXLAN EVPN topology or fabric overview

Attach Network and VRF to switches

(optional) Preview the attached configuration

Deploy the configuration

Step1 -> Create Network and VRF

Step2 -> Attach Network and VRF

Step2 -> Attach Network and VRF
(continued)

Network and VRF Configs

Pending Config – DC1 – LEAF3-DC1

Pending Config

```
vlan 2000
  vn-segment 50000
vrf context myvrf_50000
  vni 50000
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn
  address-family ipv6 unicast
    route-target both auto
    route-target both auto evpn
exit
interface Vlan2000
  vrf member myvrf_50000
  ip forward
  ipv6 address use-link-local-only
  no ip redirects
  no ipv6 redirects
  mtu 9216
  no shutdown
```

L3VNI VRF Configs

```
router bgp 65100
  vrf myvrf_50000
    address-family ipv4 unicast
      advertise l2vpn evpn
      redistribute direct route-map fabric-rmap-redist-subnet
      maximum-paths ibgp 2
      exit
    address-family ipv6 unicast
      advertise l2vpn evpn
      redistribute direct route-map fabric-rmap-redist-subnet
      maximum-paths ibgp 2
configure terminal
interface nve1
  member vni 30000
    mcast-group 239.1.1.1
  member vni 50000 associate-vrf
vlan 2300
  vn-segment 30000
interface Vlan2300
  vrf member myvrf_50000
  ip address 10.10.10.1/24 tag 12345
  fabric forwarding mode anycast-gateway
  no shutdown
exit
```

VXLAN Tenant

NVE Tunnel Configs

L3 SVI Configs

```
evpn
  vni 30000 l2
    rd auto
    route-target import auto
    route-target export auto
configure terminal
interface ethernet1/3
  switchport trunk allowed vlan add 2300
interface ethernet1/4
  switchport trunk allowed vlan add 2300
interface ethernet1/5
  switchport trunk allowed vlan add 2300
```
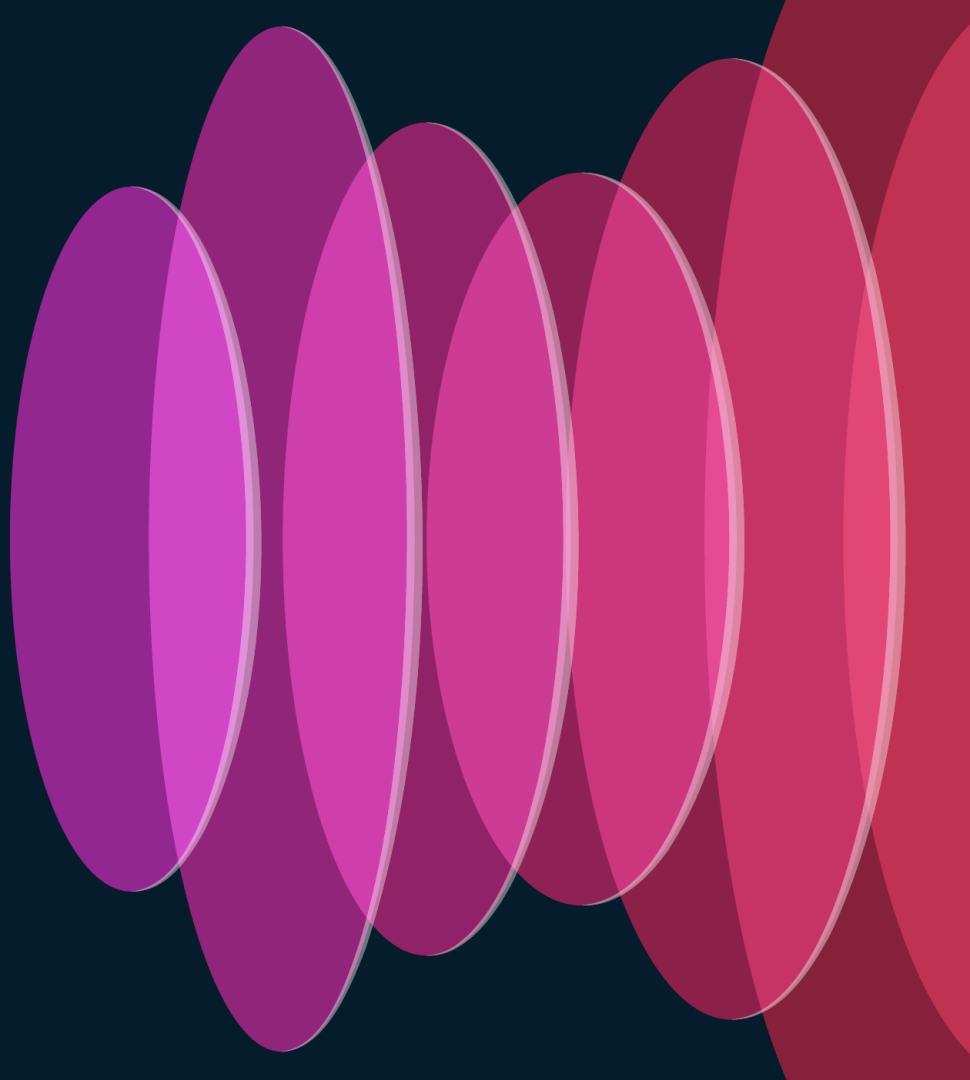
L2VNI Network Configs

Network and VRF Deployment Status

In VXLAN EVPN Fabric
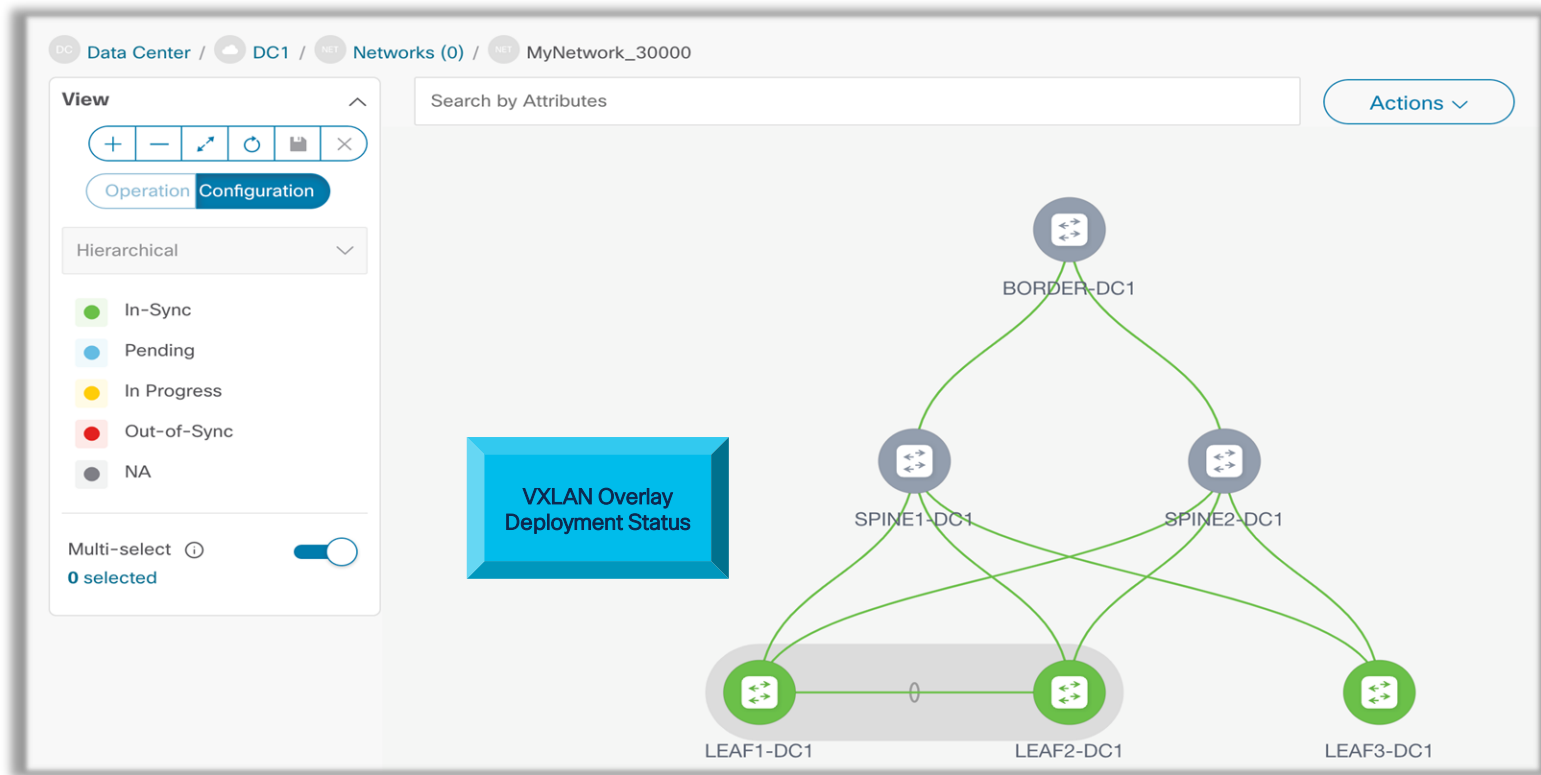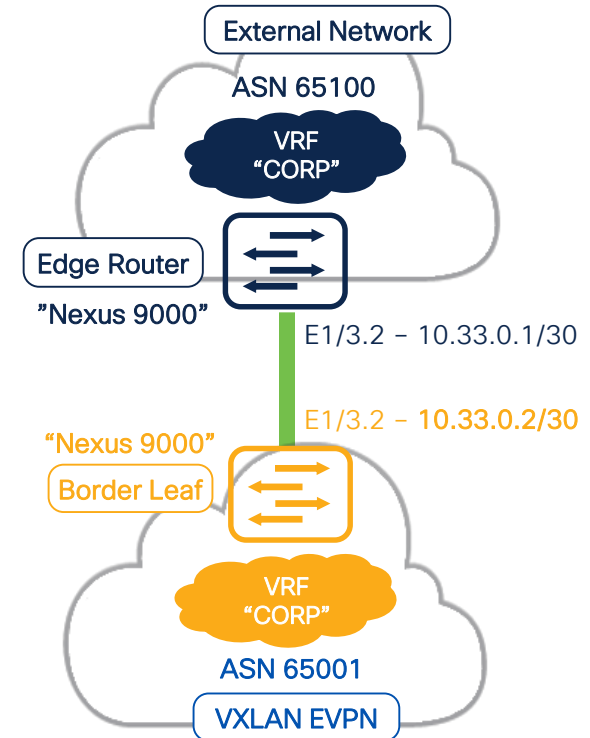
Manage and Deploy
External IP Handoff
with NDFC

cisco Live!

VRF-LITE: Border to Nexus Edge
Topology and IFC considerations

- IFC Automated

- Advertise Host (disabled)

- Advertise Default-Route (enabled)

- Config Static 0/0 Route (enabled)

External Network
ASN 65100
VRF "CORP"

Edge Router
"Nexus 9000"

E1/3.2 – 10.33.0.1/30

E1/3.2 – 10.33.0.2/30

"Nexus 9000"
Border Leaf

VRF "CORP"

ASN 65001

VXLAN EVPN

Managing Edge Devices
part of "External Connectivity Network" Fabric

- Create Fabric, Discover Switches, Set Role, and Recalculate & Deploy

VRF-LITE: Border to Nexus Edge
Defining IFC deployment type

- Review Fabric Settings for **VRF-Lite IFC deployment type:**

LAN > Fabrics > Select (your DC VXLAN Fabric) > Actions > Edit Fabric > **Resources Tab**

Fabric Name

DC1

Pick Fabric

Data Center VXLAN EVPN >

General Parameters   Replication   VPC   Protocols   Advanced   Resources

Per VRF Per Dotq1 association

Select Deployment Type

Subinterface Dot1q Range*

2-511

Per Border Dot1q Range For VRF Lite Connectivity (Min:2, Max:4093)

VRF Lite Deployment*

Back2Back&ToExternal

VRF Lite Inter-Fabric Connection Deployment Options. If 'Back2Back&ToExternal' is selected, VRF Lite IFCs are auto created between border devices of two Easy Fabrics, and between border devices in Easy Fabric and edge routers in External Fabric. The IP address is taken from the 'VRF Lite Subnet IP Range' pool.

Auto Deploy for Peer

Deploy VRF-LITE if Edge device is Nexus and managed by NDFC

Whether to auto generate VRF LITE sub-interface and BGP peering configuration on managed neighbor devices. If set, auto created VRF Lite IFC links will have 'Auto Deploy for Peer' enabled.

eBGP peering subnet details

VRF Lite Subnet IP Range*

10.33.0.0/16

Address range to assign P2P Interfabric Connections

VRF Lite Subnet Mask*

30

(Min:8, Max:31)

Close   Save

VRF-LITE: Border to Nexus Edge
Defining IFC Link on physical interface

- IFC link has been defined
  - Policy should be **ext_fabric_setup**
    - IPs auto selected from **VRF-Lite Subnet IP Range**

VRF-LITE: Border to Nexus Edge
Defining VRF extensions on Border

- Verify DC VXLAN Fabric VRFs were created and customize if needed
  - LAN > Fabrics > **Double click** (your DC VXLAN Fabric) > **VRFs**

Edit VRF if you want to modify route advertisement or other VRF specifics

VRF-LITE: Border to Nexus Edge

Defining VRF extensions on Border

Edit VRF

CORP

VRF ID*
9999

VLAN ID
99   Propose VLAN

VRF Template*
Default_VRF_Universal >

VRF Extension Template*
Default_VRF_Extension_Universal >

General Parameters   Advanced   Route Target

VRF VLAN Name
CORP

General Parameters   **Advanced**

Advertise Host Routes
☐   Flag to Control Advertisement of /32 and /128 Routes to Edge Routers

Advertise Default Route
☑   ...of Default Route Internally

Config Static 0/0 Route
☑   Flag to Control Static Default Route Configuration

Define according to your use case

BGP Neighbor Password
                                     VRF Lite BGP neighbor password (Hex String)

BGP Password Key Encryption Type
Select an Option   ˅   VRF Lite BGP Key Encryption Type: 3 - 3DES, 7 - Cisco

VRF-LITE: Border to Nexus Edge
Defining VRF extensions on Border

- **Attach** VRF to Border and extend through **VRF-Lite**

VRF-LITE: Border to Nexus Edge
Defining VRF extensions on Border

- **Attach** VRF to Border and extend through **VRF-Lite**

Extend Options for Border

| VRF_LITE | ✓ |
| NONE | |

Extend Options for Border Gateway

MULTISITE	
MULTISITE + VRF_LITE	
VRF_LITE	✓
NONE	

Edit VRF Attachment - CORP

BORDER-DC1(92VPGGIF4HB)

Detach ⬤ **Attach**

VLAN*
99

Extend*
VRF_LITE   ✕ ∨

VRF-LITE: Border to Nexus Edge

Defining VRF extensions on Border

Extension

Filter by attributes   Attach-All   Detach-All

Action	Attached	Source Switch	Type	IF_NAME	Dest. Switch	Dest. Interface	DOT1Q_ID	IP_MASK	IP_TAG	NEIGHBOR_IP	NEIGHBOR_ASN
Edit	🔴 Detached	BORDER-DC1	VRF_LITE	Ethernet1/3	EDGE-ROUTER	Ethernet1/3	2	10.33.0.1/30		10.33.0.2	65100

Extension Type

Destination Edge router details

Physical Interface MTU

MTU
9216

ENABLE_IFC_NETFLOW
☐

AUTO_VRF_LITE_FLAG
☑

Provide VRF name for External Edge. By default, NDFC uses the same name as Border VRF

PEER_VRF_NAME*
CORP

ROUTE_MAP_IN
                                        IPv4 Inbound Route Map. No route-map if blank

Provide Route-MAP if different from NDFC default

ROUTE_MAP_OUT
                                        IPv4 Outbound Route Map. If blank, use 'EXTCON-RMAP-FILTER' or 'EXTCON-RMAP-FILTER-ALLOW-HOST'

IPV6_ROUTE_MAP_IN
                                        IPv6 Inbound Route Map. No route-map if blank

IPV6_ROUTE_MAP_OUT
                                        IPv6 Outbound Route Map. If blank, use 'EXTCON-RMAP-FILTER-V6' or 'EXTCON-RMAP-FILTER-V6-ALLOW-HOST'

Cancel   Save

VRF-LITE: Border to Nexus Edge

Preview and Deploy VRF extensions on Border

Pending Config – DC1 – BORDER-DC1

Pending Config   Side-by-Side Comparison

```
vlan 99
  name CORP
  vn-segment 9999
vrf context corp
  description CORP
  vni 9999
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto evpn
  ip route 0.0.0.0/0 10.33.0.2
  address-family ipv6 unicast
    route-target both auto
    route-target both auto evpn
exit
interface Vlan99
  description CORP
  vrf member corp
  ip forward
  ipv6 address use-link-local-only
  no ip redirects
  no ipv6 redirects
  mtu 9216
  no shutdown
```

L3VNI VRF Configs

Default route static towards External Edge

```
router bgp 65001
  vrf corp
    address-family ipv4 unicast
      advertise l2vpn evpn
      redistribute direct route-map fabric-rmap-redist-subnet
      maximum-paths ibgp 2
      network 0.0.0.0/0
      exit
    address-family ipv6 unicast
      advertise l2vpn evpn
      redistribute direct route-map fabric-rmap-redist-subnet
      maximum-paths ibgp 2
      exit
    neighbor 10.33.0.2
      remote-as 65100
      address-family ipv4 unicast
        send-community both
        route-map extcon-rmap-filter out
```

L3 Tenant CORP

Advertisement of Default route towards site internal VTEP

eBGP External Edge neighbor

```
interface ethernet1/3.2
  encapsulation dot1q 2
  mtu 9216
  vrf member corp
  ip address 10.33.0.1/30
  no shutdown
interface nve1
  source-interface loopback1
  host-reachability protocol bgp
  no shutdown
  member vni 9999 associate-vrf
```

Border Physical Interface Config

VRF-LITE: Border to Nexus Edge
Preview and Deploy VRF extensions on External Edge

- Once configurations are deployed on Border Leaf (DC VXLAN EVPN), navigate to **External Network Fabric** and **perform Recalculate and Deploy**

Pending Config - External-Network - EDGE-ROUTER

Pending Config   Side-by-Side Comparison

```
vrf context corp
  address-family ipv4 unicast
exit
router bgp 65100
  vrf corp
    address-family ipv4 unicast
    neighbor 10.33.0.1
      remote-as 65001
      address-family ipv4 unicast
        send-community both
configure terminal
interface ethernet1/3.2
  mtu 9216
  vrf member corp
  encapsulation dot1q 2
  ip address 10.33.0.2/30
  no shutdown
```

VRF Configs

eBGP Border Leaf neighbor

Edge Physical Interface Config

Enabling VRF-Lite: Manual
Border Leaf to Non-Nexus Edge Router

Step 2

IFC prototypes

Define IFC Type: "Manual"

Define IFC link on physical interface

Recalculate config and deploy

Step 4

External Network Fabric VRF-Lite Extension

Define sub-interfaces, invoke BGP policies

NDFC will generate Sub-int and eBGP peering on Edge router

Deploy on External Network Fabric

Step 1

Create and Import

Define VXLAN EVPN and External Fabric. Set respective roles (e.g. Border, BGW, Edge)

Uncheck Fabric monitor mode in External Fabric if the Edge router is in managed mode

For Non-Nexus ensure SNMP configs for discovery

Step 3

DC VXLAN EVPN Fabric VRF-Lite Extension

Define Individual VRF extension on the Border leaf

NDFC will generate Sub-int and eBGP peering on Border leaf

Deploy on Data Center VXLAN EVPN Fabric

VRF-LITE: Border to Non-Nexus Edge

Topology and IFC considerations

- IFC Manual

- Advertise Host (disabled)

- Advertise Default-Route (enabled)

- Config Static 0/0 Route (enabled)

External Network

ASN 65099

VRF "CORP"

Edge Router

"ASR 9000"

Gig0/0/0/0.2 – 10.10.10.1/31

E1/1.2 – **10.10.10.0/31**

"Nexus 9000"

Border Leaf

VRF "CORP"

ASN 10

VXLAN EVPN

VRF-LITE: Border to Non-Nexus Edge

Defining IFC deployment type

- Review Fabric Settings for **VRF-Lite IFC deployment type:**

  LAN > Fabrics > Select (your DC VXLAN Fabric) > Actions > Edit Fabric > **Resources Tab**

Fabric Name

VXLAN

Pick Fabric

Data Center VXLAN EVPN >

| General Parameters | Replication | VPC | Protocols | Advanced | **Resources** |

Per VRF Per Dotq1 association

Select Deployment Type

Subinterface Dot1q Range*

2-511

Per Border Dot1q Range For VRF Lite Connectivity (Min:2, Max:4093)

VRF Lite Deployment*

Manual

VRF Lite Inter-Fabric Connection Deployment Options. If 'Back2Back&ToExternal' is selected, VRF Lite IFCs are auto created between border devices of two Easy Fabrics, and between border devices in Easy Fabric and edge routers in External Fabric. The IP address is taken from the 'VRF Lite Subnet IP Range' pool.

eBGP peering subnet details

VRF Lite Subnet IP Range*

10.33.0.0/16

Address range to assign P2P Interfabric Connections

VRF Lite Subnet Mask*

30

(Min:8, Max:31)

Close   Save

VRF-LITE: Border to Non-Nexus Edge
Defining IFC Link on physical interface

- Verify link from DC VXLAN Fabric to External Fabric was discovered

    LAN > Fabrics > **Double Click** (your DC VXLAN Fabric) > **Links**

Use filters when having multiple fabrics

If no policy exists, or neighbor not found. A link can be created manually

VRF-LITE: Border to Non-Nexus Edge

Defining IFC Link on physical interface

- If Link discovered, values are pre-filled
- When not, make sure:
  - Link Type is **Inter-Fabric**
  - Link Sub-Type is **VRF_Lite**
  - Link Template: **ext_fabric_setup**
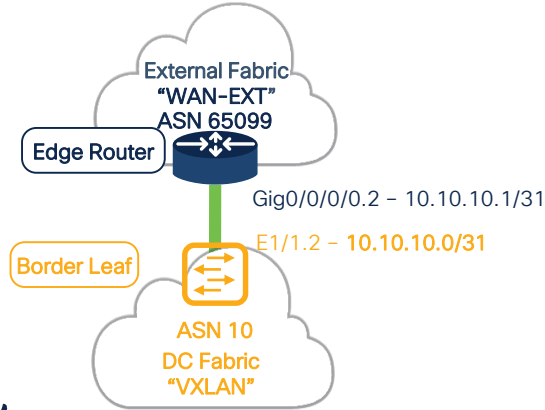
External Fabric
"WAN-EXT"
ASN 65099

Edge Router

Gig0/0/0/0.2 – 10.10.10.1/31

E1/1.2 – **10.10.10.0/31**

Border Leaf

ASN 10
DC Fabric
"VXLAN"

Link Type
Inter-Fabric

Link Sub-Type*
VRF_LITE

Link Template*
ext_fabric_setup

VRF_LITE ✓
MULTISITE_UNDERLAY
MULTISITE_OVERLAY
BGP_OVER_IPSEC

Source Fabric
VXLAN

Destination Fabric
WAN-EXT

Source Device*
vxlan-border

Destination Device*
Backbone-Edge

Source Interface*
Ethernet1/1

Destination Interface*
GigabitEthernet0/0/0/0

General Parameters Advanced

Source BGP ASN*
10
BGP Autonomous System Number in Source Fabric

Source IP Address/Mask*
10.10.10.0/31
IP address for sub-interface in each VRF in Source Fabric

Destination IP*
10.10.10.1
IP address for sub-interface in each VRF in Destination Fabric

Destination BGP ASN*
65099
BGP Autonomous System Number in Destination Fabirc

Link MTU
9216
Interface MTU on both ends of VRF Lite IFC

VRF-LITE: Border to Non-Nexus Edge
Defining VRF extensions on Border

- Verify DC VXLAN Fabric VRFs were created and customize if needed
  - LAN > Fabrics > **Double click** (your DC VXLAN Fabric) > **VRFs**

Overview	Switches	Links	Interfaces	Interface Groups	Policies	Networks	VRFs	Services	Event Analytics	History	Resources

	VRF Name	VRF Status	VRF ID	
				Create
✓	CORP	● NA	9999	Edit
				Deploy

Filter by attributes   Actions ⌄

Edit VRF if you want to modify route advertisement or other VRF specifics

VRF-LITE: Border to Non-Nexus Edge
Defining VRF extensions on Border

VRF-LITE: Border to Non-Nexus Edge
Defining VRF extensions on Border

- **Attach** VRF to Border and extend through **VRF-Lite**

VRF-LITE: Border to Non-Nexus Edge
Defining VRF extensions on Border

- **Attach** VRF to Border and extend through **VRF-Lite**

Extend Options for Border

VRF_LITE	✓
NONE	

Extend Options for Border Gateway

MULTISITE	
MULTISITE + VRF_LITE	
VRF_LITE	✓
NONE	

vxlan-border(9CZ07PR08CT)

Detach ⬤ **Attach**

VLAN*

2000

Extend*

VRF_LITE ✕ ⌄

CLI Freeform Config

Edit ›

All configs should strictly match the 'show run' output, including cases and new line
Any mismatches will yield unexpected diffs during deploy

Cancel   Save

VRF-LITE: Border to Non-Nexus Edge

Preview and Deploy VRF extensions on Border

Pending Config - VXLAN - vxlan-border

Pending Config

```
vrf context corp
  ip route 0.0.0.0/0 10.10.10.1
exit
router bgp 10
  vrf corp
    address-family ipv4 unicast
      network 0.0.0.0/0
      exit
    neighbor 10.10.10.1
      remote-as 65099
      address-family ipv4 unicast
        send-community both
        route-map extcon-rmap-filter out
configure terminal
interface ethernet1/1.2
  encapsulation dot1q 2
  mtu 9216
  vrf member corp
  ip address 10.10.10.0/31
  no shutdown
interface loopback10
  vrf member corp
  ip address 2.2.2.2/32 tag 12345
configure terminal
```

- Default route static towards External Edge
- Advertisement Default route towards site internal
- eBGP External Edge neighbor
- Border Physical Interface Config

VRF-LITE: Border to Non-Nexus Edge
Defining the policies on the Edge Router

- After completing the configurations on the VXLAN Fabric (Border Leaf), Navigate to **External Fabric (Edge Router)** and apply the following **policies**
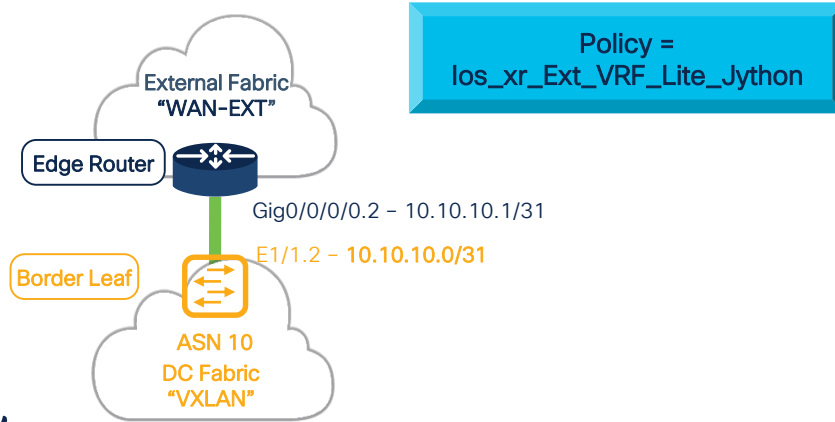 - **ios_xr_base_bgp**

Policy = Ios_xr_base_bgp

Switch List:
Backbone-Edge

Priority*
500
1-1000

Description

Template Name
ios_xr_base_bgp

BGP ASN*
65099

BGP Router id*
1.1.1.1

VRF-LITE: Border to Non-Nexus Edge

Defining the policies on the Edge Router

- After completing the configurations on the VXLAN Fabric (Border Leaf), Navigate to **External Fabric (Edge Router)** and apply the following **policies**
  - ios_xr_base_bgp
  - ios_xr_Ext_VRF_Lite_Jython

Policy =
Ios_xr_Ext_VRF_Lite_Jython

External Fabric
"WAN-EXT"

Edge Router

Gig0/0/0/0.2 – 10.10.10.1/31

E1/1.2 – **10.10.10.0/31**

Border Leaf

ASN 10
DC Fabric
"VXLAN"

Field	Value	Description
Layer-3 Interface*	GigabitEthernet0/0/0/0	Subinterface Parent interface (e.g. te0/2/0/3)
Encapsulation dot1q VLAN ID*	2	1-4094
VRF Name	CORP	Subinterface VRF name, default VRF if not specified
Subinterface IPv4 Address/Netmask*	10.10.10.1/31	For IPv4 VRF Lite peering (e.g. 10.33.1.0/30)
Subinterface MTU*	9216	Layer-3 MTU (Min:64, Max:12000)
Neighbor IPv4 Address*	10.10.10.0	BGP Peer IPv4 Address
Neighbor ASN*	10	BGP ASN of IPv4/IPv6 Neighbor
Subinterface IPv6 Address/Prefix Length		For IPv6 VRF Lite peering (e.g. 1::1/112)
Neighbor IPv6 Address		BGP Peer IPv6 Address
Local ASN		The fabric ASN will be used if not specified
Route Distinguisher*	65099:4	Specify the route distinguiher used for vrf
Route Policy Name*	ALLOW_ALL	Specify the route policy used for vrf

VRF-LITE: Border to Non-Nexus Edge

Deploy configs on the Edge Router

Pending Config - WAN-EXT - Backbone-Edge

Pending Config   Side-by-Side Comparison

```
route-policy ALLOW_ALL
  pass
end-policy
router bgp 65099
 bgp router-id 1.1.1.1
 address-family ipv4 unicast
 address-family vpnv4 unicast
 address-family ipv6 unicast
 address-family vpnv6 unicast
 vrf CORP
  rd 65099:4
  address-family ipv4 unicast
   maximum-paths ebgp 4
   redistribute connected
   exit
  neighbor 10.10.10.0
   remote-as 10
   address-family ipv4 unicast
    send-community-ebgp
    route-policy ALLOW_ALL in
    route-policy ALLOW_ALL out
    send-extended-community-ebgp
    exit
   exit
  exit
```
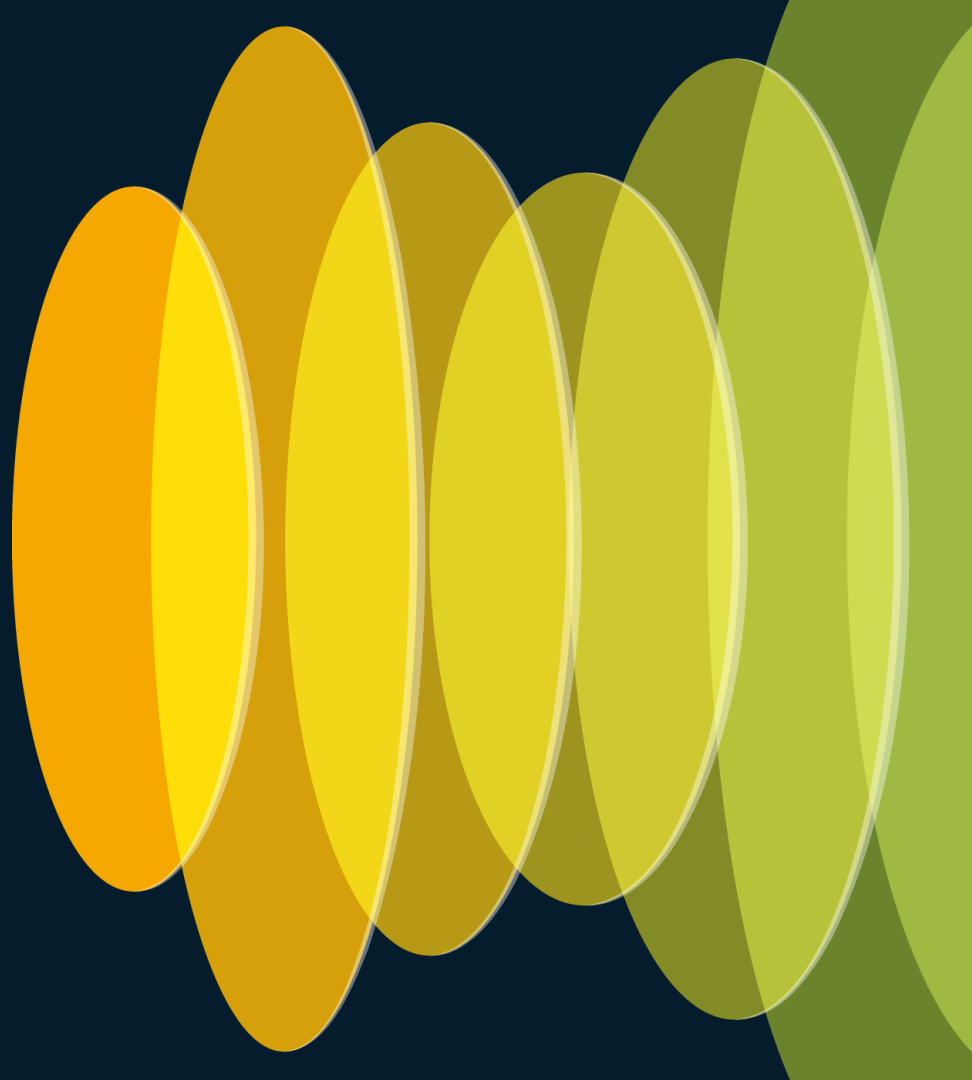
IOS-XR eBGP Policy for allowing routes

IOS-XR VRF Definition

eBGP Border Leaf neighbor

```
vrf CORP
 address-family ipv4 unicast
interface GigabitEthernet0/0/0.2
 encapsulation dot1q 2
 vrf CORP
 ipv4 address 10.10.10.1 255.255.255.254
 mtu 9216
 exit
```

IOS-XR Physical Interface Config

Verification through NDFC
Keeping you away from CLI

Step 2
Deployment History

Configuration Execution Status:

Verify Deployment History Status **Success** for Underlay, Overlay, Interfaces, and more

Step 1
Verify Network and VRF attachments

Job Execution Status:

Network Status **Deployed**

VRF Status **Deployed**

Step 3
Show commands

Service / features status
(CLI through NDFC)

Attachment deployment status

Job execution perspective

Success or Failure deployment details

Multi-Stage Preview and Deployment

Switch Name	IP Address	Status	Status Description	Progress
LEAF3-DC1	192.168.101.23	● SUCCESS	Deployment completed.	Executed 5 / 5
SPINE1-DC1	192.168.101.24	● COMPLETED	No Commands to execute.	
SPINE2-DC1	192.168.101.25	● COMPLETED	No Commands to execute.	

Switch Name	IP Address	Status	Status Description	Progress
Edge-router	192.168.101.99	● FAILED	Deployment failed. Check deployment history for more information.	Executed 1 / 12
Edge-Catalyst	192.168.101.101	● COMPLETED	No Commands to execute.	

Role	VRF Status	Status Description	Progress
border gateway	● Deployment In-Pr	Adding diff to deployment queue	

Role	VRF Status	Status Description	Progress
border gateway	● In-Sync	Config compliance sync completed	

Deployment History Tool
Commands execution perspective

Command Execution Details for Edge-router(91UP4O2KI8R)

Config	Status	CLI Response
vrf member eng	● SUCCESS	Warning: Deleted all L3 config on interface Vlan100 VRF eng does not exist. Create vrf to make interface Vlan100 operational
ip address 4.4.4.1/24	● SUCCESS	
no ip redirects	● SUCCESS	
no ipv6 redirects	● SUCCESS	
hsrp version 2	● FAILED	Failed to configure hsrp version 2 on Vlan100 due to invalid command.
hsrp 44	● FAILED	Failed to execute this command, since command#7 (hsrp version 2) failed.
ip 4.4.4.2	● FAILED	Failed to execute this command, since command#7 (hsrp version 2) failed.

Fabric Overview - External-Edge

Overview Switches Links Interfaces Policies Event Analytics History Resources

Deployment History

Policy Change History

Filter by attributes

HostName	Entity Name	Entity Type	Source	Commands	Status	Serial Number	Status
Edge-router	Vlan100	INTERFACE	UNDERLAY	Detailed History	● SUCCESS	91UP4O2KI8R	Successfully deployed admin May 02 2023, 12:20:20
Edge-router	Vlan100	INTERFACE	UNDERLAY	Detailed History	● FAILED	91UP4O2KI8R	CLI command ' hsrp version 2' fa error inval
Edge-router	SWITCH	SWITCH	UNDERLAY	Detailed History	● SUCCESS	91UP4O2KI8R	Suc

CLI response messages for easier troubleshooting

Command Execution Details for Edge-router(91UP4O2KI8R)

Config	Status	CLI Response
router bgp 65100	● SUCCESS	
address-family ipv4 unicast	● SUCCESS	
network 100.1.1.0/24	● SUCCESS	

Show Commands Tool

Switch Config perspective

NDFC pre-built commands or user commands

Conclusion

Key points to remember

1 NDFC simplifies automation and management of VXLAN EVPN fabrics using Cisco's best practices

2 NDFC provides flexible design options alongside automation, consistency, compliance, and management for VXLAN EVPN and Multi-Site

3 NDFC provides a single plane of glass solution to automate and manage Nexus and Non-Nexus devices

Complete Your Session Evaluations

Complete a minimum of 4 session surveys and the Overall Event Survey to be entered in a drawing to **win 1 of 5 full conference passes** to Cisco Live 2025.

Earn 100 points per survey completed and compete on the Cisco Live Challenge leaderboard.

Level up and earn **exclusive prizes!**

Complete your surveys in the **Cisco Live mobile app.**

Continue your education

- Visit the Cisco Showcase for related demos

- Book your one-on-one Meet the Engineer meeting

- Attend the interactive education with DevNet, Capture the Flag, and Walk-in Labs

- Visit the On-Demand Library for more sessions at www.CiscoLive.com/on-demand

Contact me at: Cisco Webex

Thank you

CISCO
The bridge to possible

CISCO Live!

#CiscoLive