

CISCO *Live!*

GO BEYOND

#CiscoLiveAPJC



Design, Automate, and Manage VXLAN Multi-Site with NDFC

Parth Patel, Sr. Technical Leader, TME
Data Center Networking BU
BRKDCN-2128

Cisco Webex App

Questions?

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

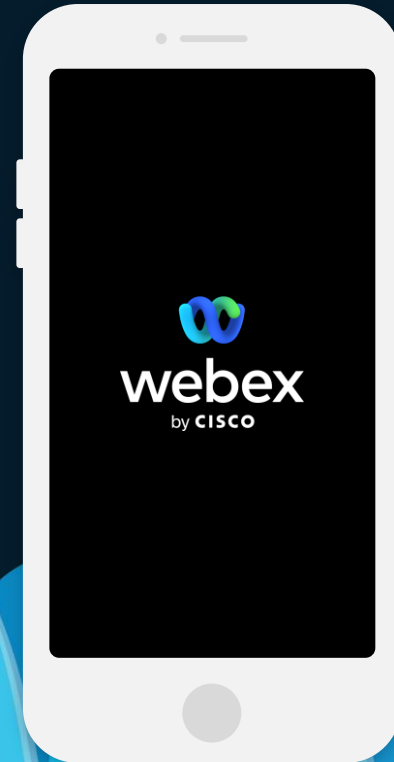
How

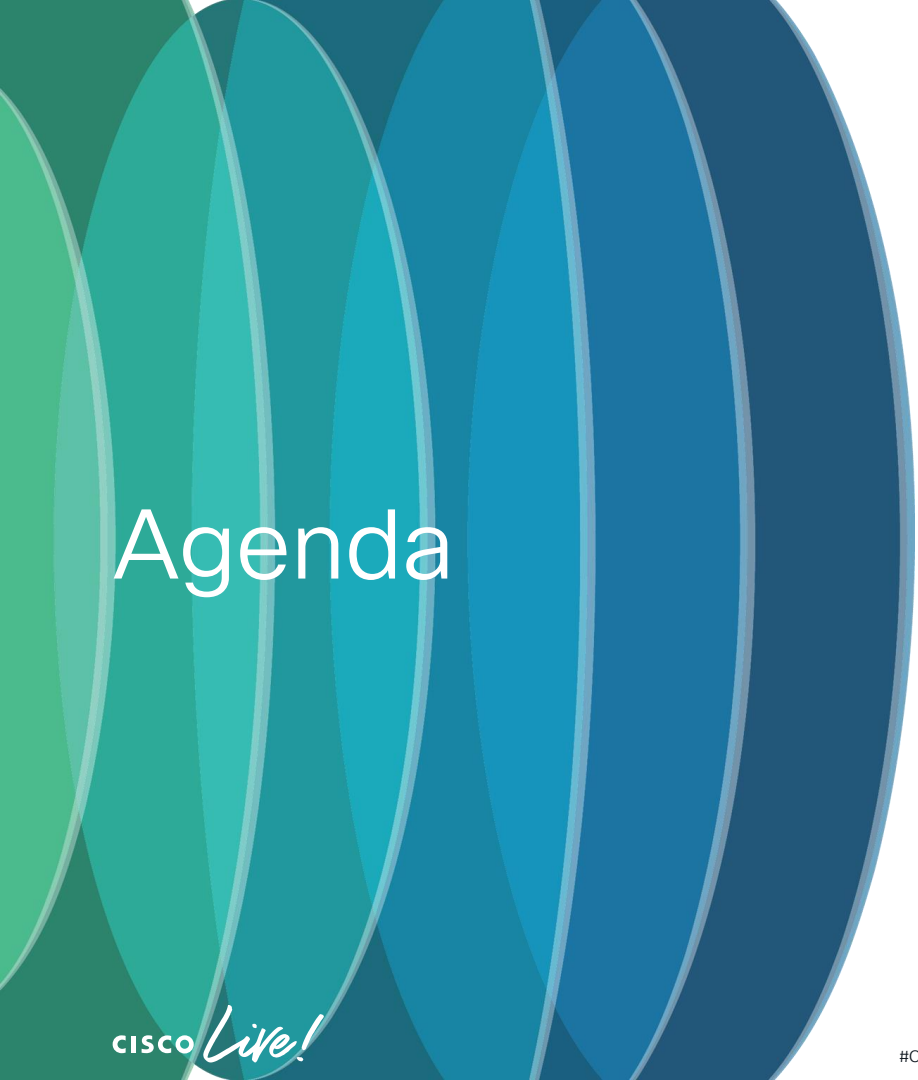
- 1 Find this session in the Cisco Live Mobile App
- 2 Click “Join the Discussion”
- 3 Install the Webex App or go directly to the Webex space
- 4 Enter messages/questions in the Webex space

Webex spaces will be moderated by the speaker until November 15, 2024.



<https://ciscolive.ciscoevents.com/ciscolivebot/#BRKDCN-2128>





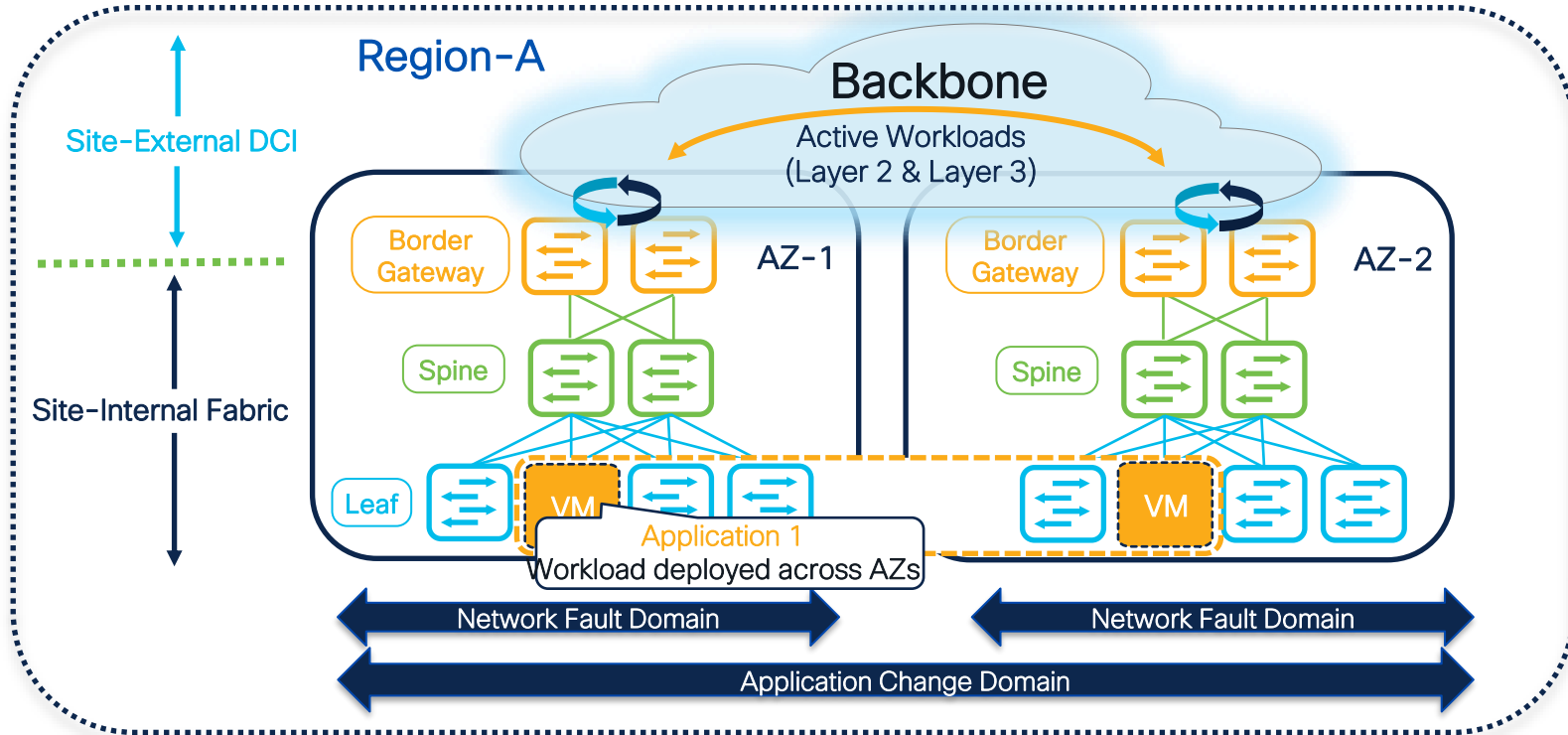
Agenda

- Design VXLAN EVPN Multi-Site with NDFC
- Introduction to NDFC Federation
- Automate VXLAN EVPN Multi-Site with NDFC Federation
- Conclusion

VXLAN Multi-Site Design Options

VXLAN EVPN Multi-Site Overview

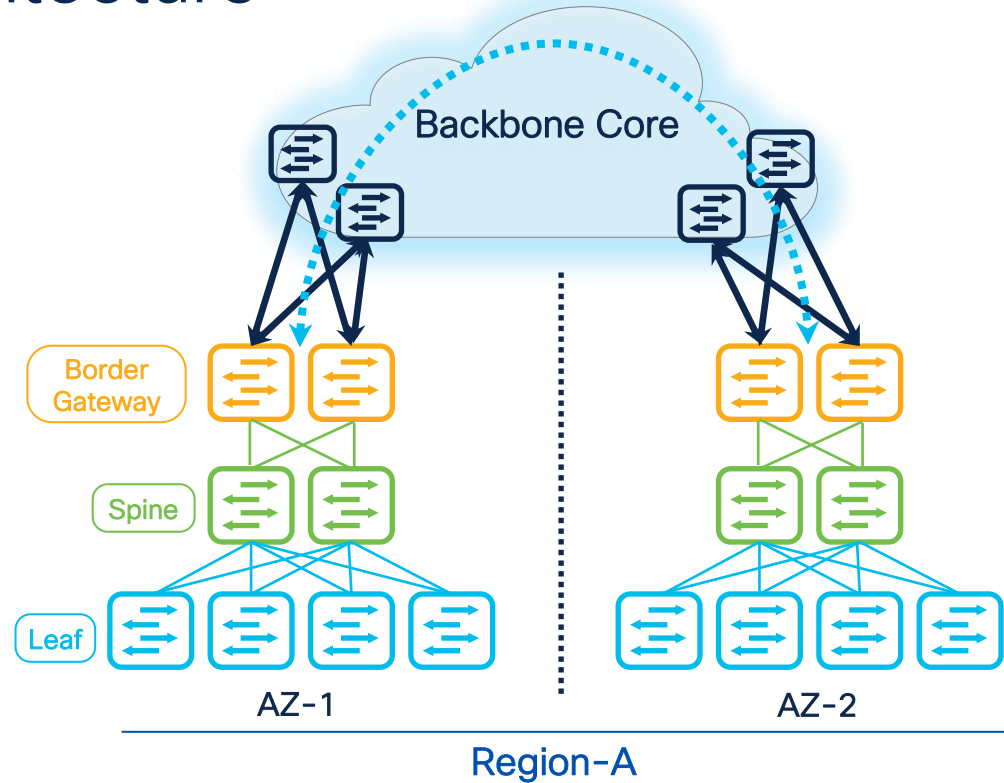
Functional Components



VXLAN Multi-Site Architecture

DCI- BGW to Cloud

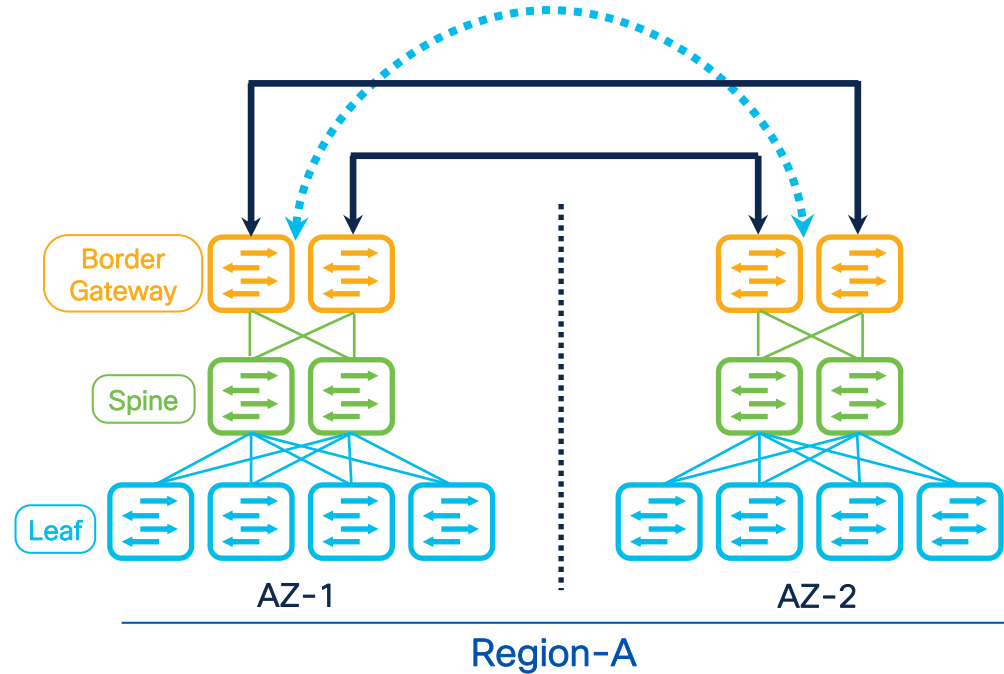
- Scalable design option with horizontal scale-out within and across multiple sites.
- Backbone/Cloud can be any routed service. (flat L3, MPLS-L3VPN)
- Multi-Site Underlay: eBGP IPv4 Unicast** ↔
 - Advertise **Lo0 (Overlay Control Plane)**, **Lo1 (BUM, External networks)**, and **Lo100 (Multi-Site Inter-Site Transit communication)**
 - Site-External DCI BUM: Ingress-Replication or Multicast supported. *Currently NDFC supports only Ingress-Replication.
 - Site-Internal Fabric BUM: Ingress-Replication or Multicast supported independently at each site.
- Multi-Site Overlay: eBGP EVPN Overlay** ↔·····
 - Full-mesh BGP EVPN peering across all BGWs.
- Ensure that PIP / VIP of all BGWs are known by every BGW and MTU must accommodate VXLAN encapsulated traffic.



VXLAN Multi-Site Architecture

DCI- BGW Back-to-Back

- Available option for connecting 2 or 3 sites.
- Minimum topology is the square. An enhanced option is to add links between site local BGWs for improved ECMP and Failure scenarios.
- **Multi-Site Underlay: eBGP IPv4 Unicast** ↔
 - Advertise **Lo0 (Overlay Control Plane)**, **Lo1 (BUM, External networks)**, and **Lo100 (Multi-Site Inter-Site Transit communication)**
 - Site-External DCI BUM: Ingress-Replication or Multicast supported. *Currently NDFC supports only Ingress-Replication.
 - Site-Internal Fabric BUM: Ingress-Replication or Multicast supported independently at each site.
- **Multi-Site Overlay: eBGP EVPN Overlay** ↔.....→
 - Full-mesh BGP EVPN peering across all BGWs.
- Ensure that PIP / VIP of all BGWs are known by every BGW and MTU must accommodate VXLAN encapsulated traffic.



Multi-Destination Traffic Forwarding

EVPN Route-Type 4 (Ethernet Segment Route)

- BGW advertises BGP EVPN Type-4 update to all of their BGP EVPN peers (Site-Internal and Site-External)
- BGW will use this information for DF (Designated Forwarded) election on per VLAN/VNI
- EVPN Type-4 update will also carry ES Import Route-Target Extended Community
- Remote site BGWs will not install the route due to different RT and Site ID
- Command to check: **show bgp l2vpn evpn route-type 4**
 - BGP routing table entry for [4]:[0300.0000.00fd.e900.0309]:[32]:[10.1.1.1]/136
 - [4] = EVPN Route-Type 4
 - [03] = ESI Type MAC-based
 - [0000.0000.fde9] = ESI system MAC for Site-ID and RT {fde9 = 65001}
 - [000309] = ESI local discriminator
 - [32] = IP Address Length
 - [10.1.1.1] = Sender (BGW Primary IP) used for DF election process

evpn multisite border-gateway 65001
delay-restore time 300



Multi-Destination Traffic Forwarding

BGW Designated Forwarded (DF) Election

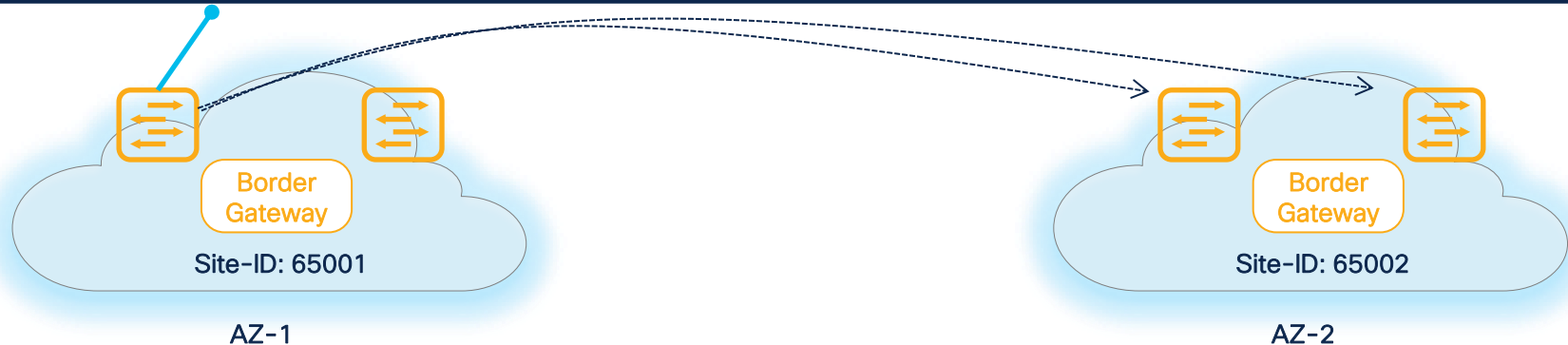
- BGW learns about other BGWs originator IPs on the same site from exchanging EVPN Type 4 routes. The originator IP address used is the BGW NVE source interface's primary IP address.
- Each BGW creates an ordinal list of originator IP in numerical order from lowest to highest. Every BGW is then given an ordinal value based on its position in the ordinal list starting from position 0. The BGW with the lowest originator IP would get an ordinal value of 0. The ordinal value decides which BGW will be the DF for a VLAN/VNI.
- DF election Formula uses “mod” operation:
 - $I = V \bmod N$ {I = ordinal value, V = VLAN #, N = # BGWs in a site}
- If there are 4 BGWs in one site. The ordinal list of the BGWs will be arranged as shown below.
- Ordinal List = 10.1.1.1, 10.1.1.2, 10.1.1.3, 10.1.1.4 {I=0, I=1, I=2, I=3}
- The DF for VLAN 48 will be: “ $I = 48 \bmod 4 = 0 = 10.1.1.1$ ”
- The DF for VLAN 49 will be: “ $I = 49 \bmod 4 = 1 = 10.1.1.2$ ”
- The DF for VLAN 50 will be: “ $I = 50 \bmod 4 = 2 = 10.1.1.3$ ”
- The DF for VLAN 51 will be: “ $I = 51 \bmod 4 = 3 = 10.1.1.4$ ”

Multi-Destination Traffic Forwarding

EVPN Route-Type 3 (Inclusive Multicast Ethernet Tag Route)

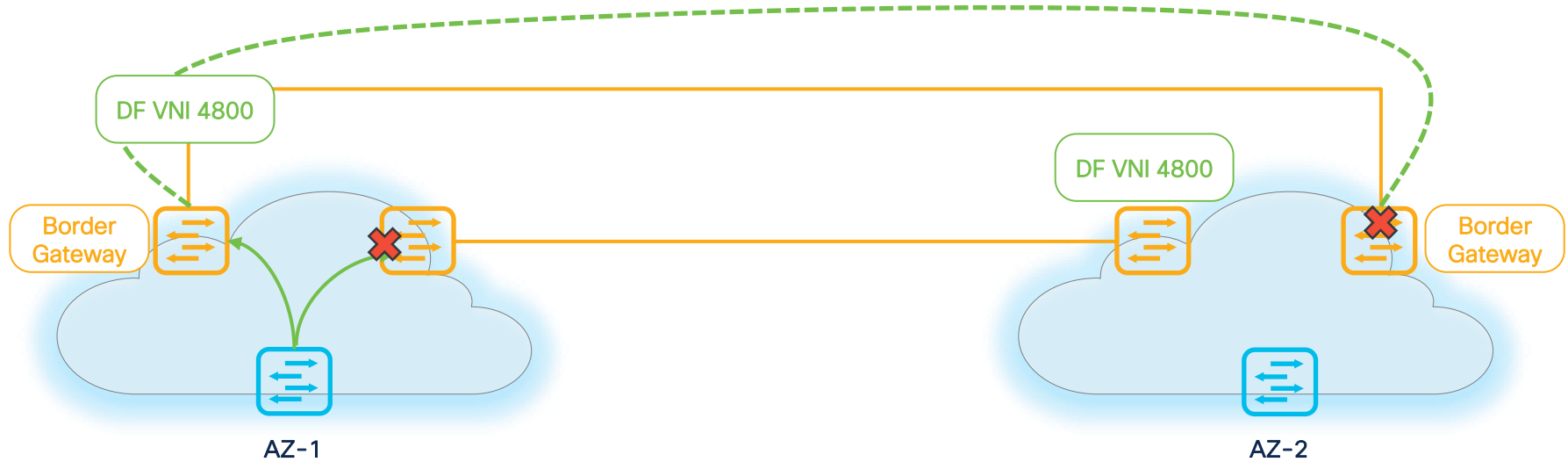
- BGW advertises BGP EVPN Type-3 only towards Site-External DCI
- Receiving BGW switch does not forward it to Site-Internal Fabric
- Command to check: **show bgp l2vpn evpn route-type 3,**
- **show l2route evpn imet all detail**
- **Local: Advertising from Site 1 BGW to Site 2 BGW**
BGP routing table entry for [3]:[0]:[32]:[10.3.0.4]/88
AS-Path: NONE, path locally originated
Extcommunity: RT:65001:10000 ENCAP:8
PMSI Tunnel Attribute:
 flags: 0x00, Tunnel type: Ingress Replication
 Label: 10000, Tunnel Id: 10.3.0.4
Path-id 1 advertised to peers: 20.2.0.4

- **Remote: Learning from Site 2 BGW to Site 1 BGW**
BGP routing table entry for [3]:[0]:[32]:[20.3.0.4]/88
Path type: external, path is valid, is best path
Imported to 1 destination(s)
Imported paths list: L2-10000
AS-Path: 65002 , path sourced external to AS
Extcommunity: RT:65001:10000 ENCAP:8
PMSI Tunnel Attribute:
 flags: 0x00, Tunnel type: Ingress Replication
 Label: 10000, Tunnel Id: 20.3.0.4



VXLAN Multi-Site Architecture

DCI- BGW Back-to-Back BUM (w/o Local L3 link)



Region-A



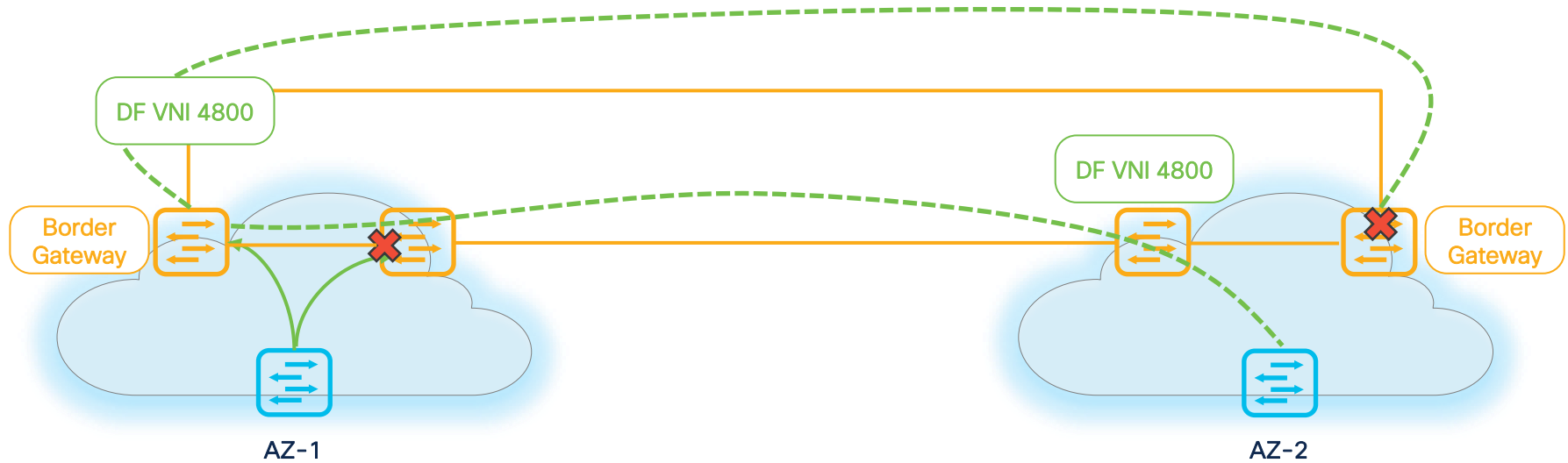
Drop due to Designated Forwarder (DF) rule

Physical connection

BUM packet

VXLAN Multi-Site Architecture

DCI- BGW Back-to-Back BUM (with Local L3 link)



Region-A



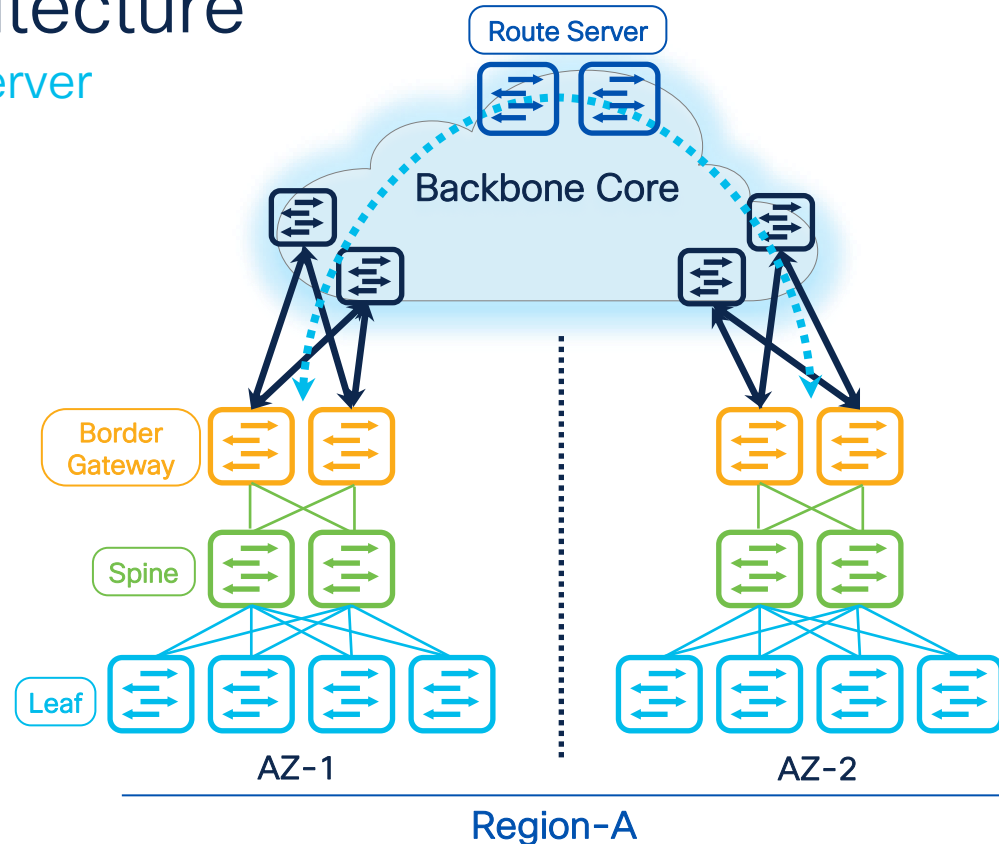
Drop due to Designated Forwarder (DF) rule

Physical connection

BUM packet

VXLAN Multi-Site Architecture

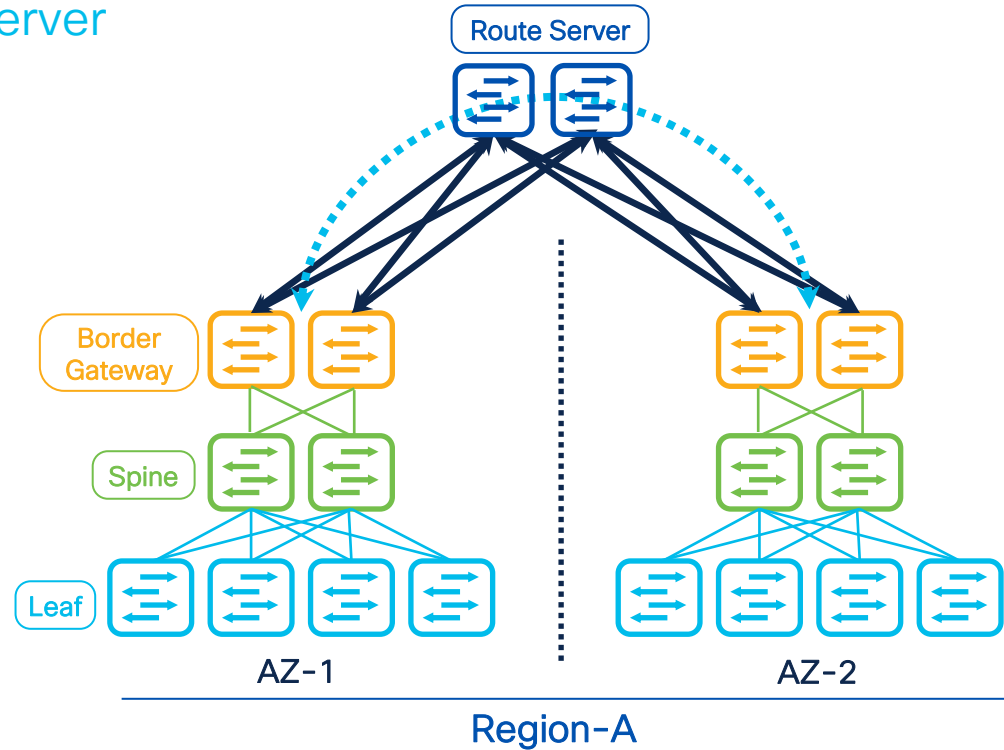
- BGP Route-Server (RS) function can be used in a highly scalable design to act as the central EVPN peering point for the Multi-Site.
- RS does not need to be on the data path.
- **Multi-Site Underlay: eBGP IPv4 Unicast** 
 - **Advertise Lo0 (Overlay Control Plane), Lo1 (BUM, External networks), and Lo100 (Multi-Site Inter-Site Transit communication)**
 - Site-External DCI BUM: Ingress-Replication or Multicast supported.
*Currently NDFC supports only Ingress-Replication.
 - Site-Internal Fabric BUM: Ingress-Replication or Multicast supported independently at each site.
- **Multi-Site Overlay: eBGP EVPN Overlay** 
 - BGP EVPN peering only between BGWs and RS.
- **Route-Server (RS) must support...**
 - EVPN AFI and Router-Server function per RFC 7947
 - Next-hop-unchanged, Retain RT, and RT Rewrite function
- **Ensure that PIP / VIP of all BGWs are known by every BGW and MTU must accommodate VXLAN encapsulated traffic.**



VXLAN Multi-Site Architecture

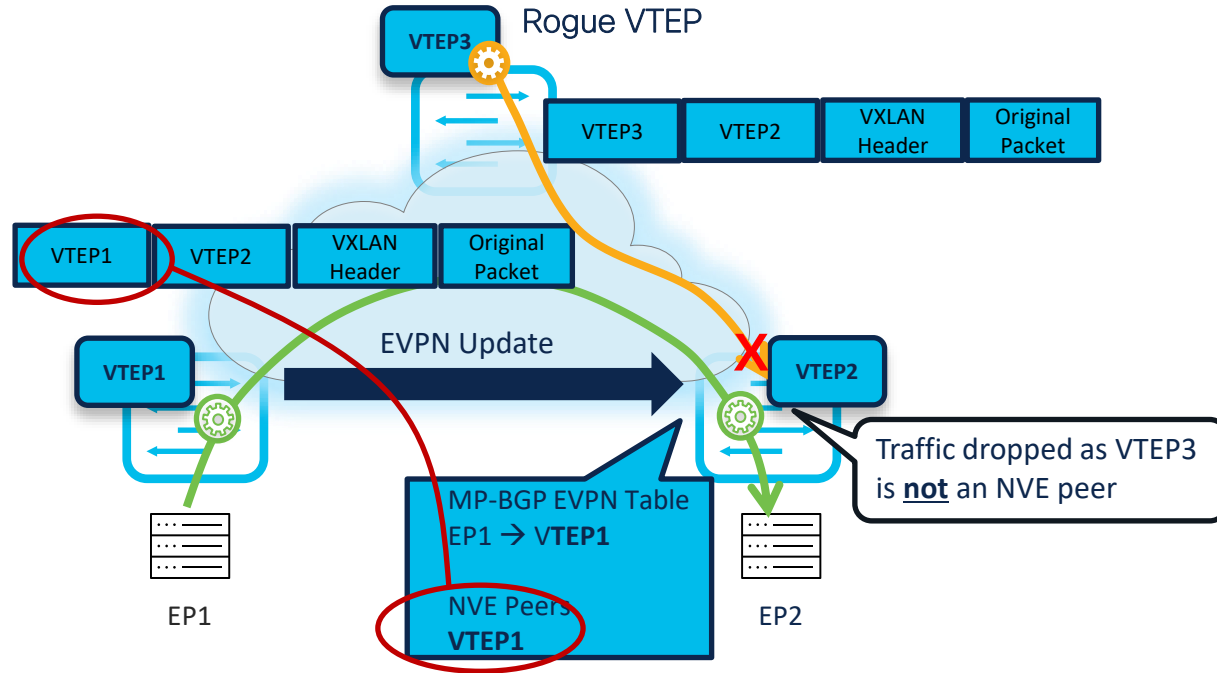
DCI- BGW to Centralized Router-Server

- BGP Route-Server (RS) function can be used in a highly scalable design to act as the central EVPN peering point for the Multi-Site.
- RS can also be on the data path.
- **Multi-Site Underlay: eBGP IPv4 Unicast** ↔
 - Site-External DCI BUM: Ingress-Replication or Multicast supported.
*Currently NDFC supports only Ingress-Replication.
 - Site-Internal Fabric BUM: Ingress-Replication or Multicast supported independently at each site.
- **Multi-Site Overlay: eBGP EVPN Overlay** ←.....→
 - BGP EVPN peering only between BGWs and RS.
- **Route-Server (RS) must support...**
 - EVPN AFI and Router-Server function per RFC 7947
 - Next-hop-unchanged, Retain RT, and RT Rewrite function
- Ensure that PIP / VIP of all BGWs are known by every BGW and MTU must accommodate VXLAN encapsulated traffic.



Special Considerations for Layer-3 extension across Multi-Site

VXLAN Native Data Plane Security



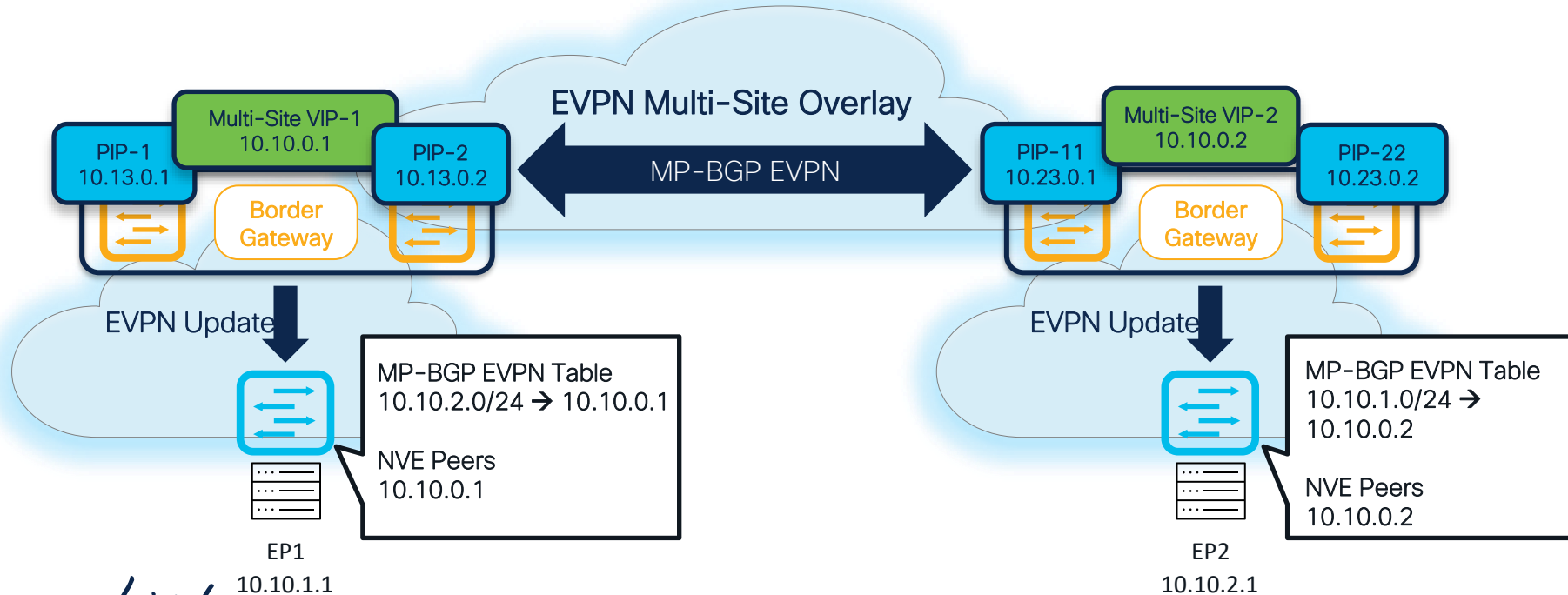
- VXLAN EVPN offers a native data plane security functionality
 - VXLAN traffic originated from a remote VTEP is only accepted when sourced from a TEP address that is an “NVE peer”
 - An NVE peer’s address is added to the local table based on the reception of MP-BGP EVPN updates carrying that specific address as next-hop
- Prevents the insertion of rogue VTEPs in a VXLAN EVPN fabric

Why can this become an issue for
Layer-3 communication across sites?

Inter-Site Layer 3 Traffic – Control Plane

Inter-Site Type-2 and Type-5 EVPN updates always carry the local Multi-Site VIP as next-hop address

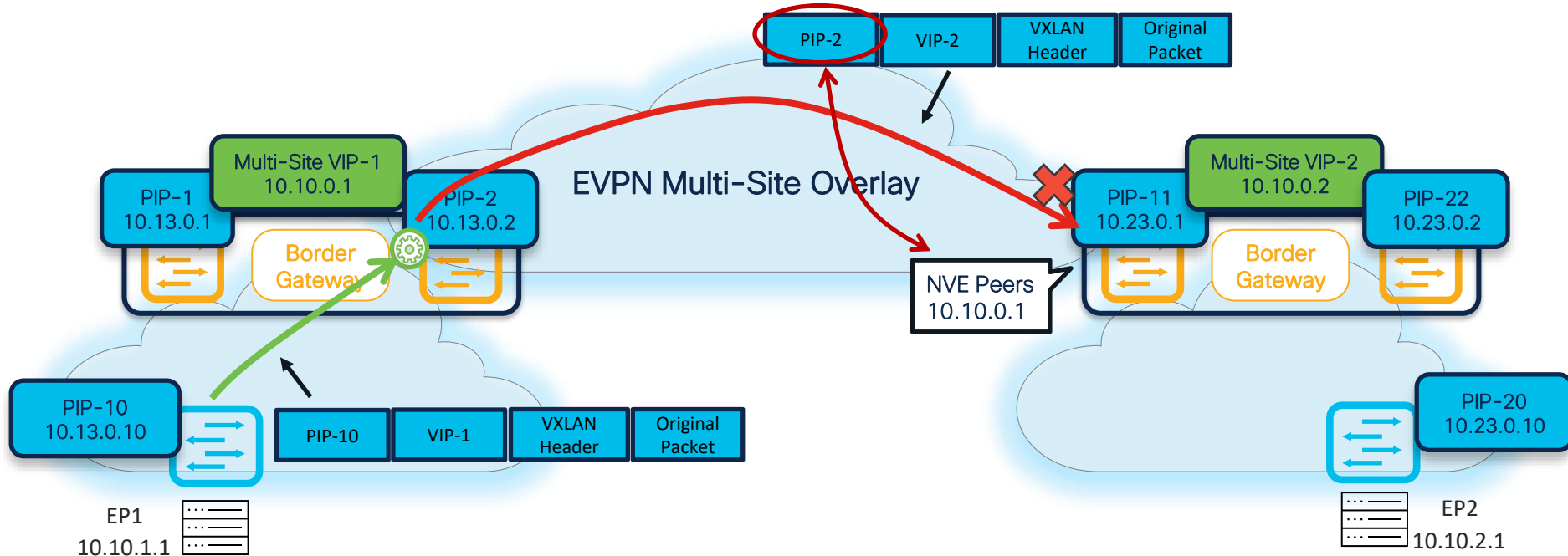
- Only exception are Type-5 updates for L3 networks locally connected to the BGWs



Inter-Site Layer 3 Traffic – Data Plane

Inter-Site traffic is always sourced by local BGWs from their specific PIP address

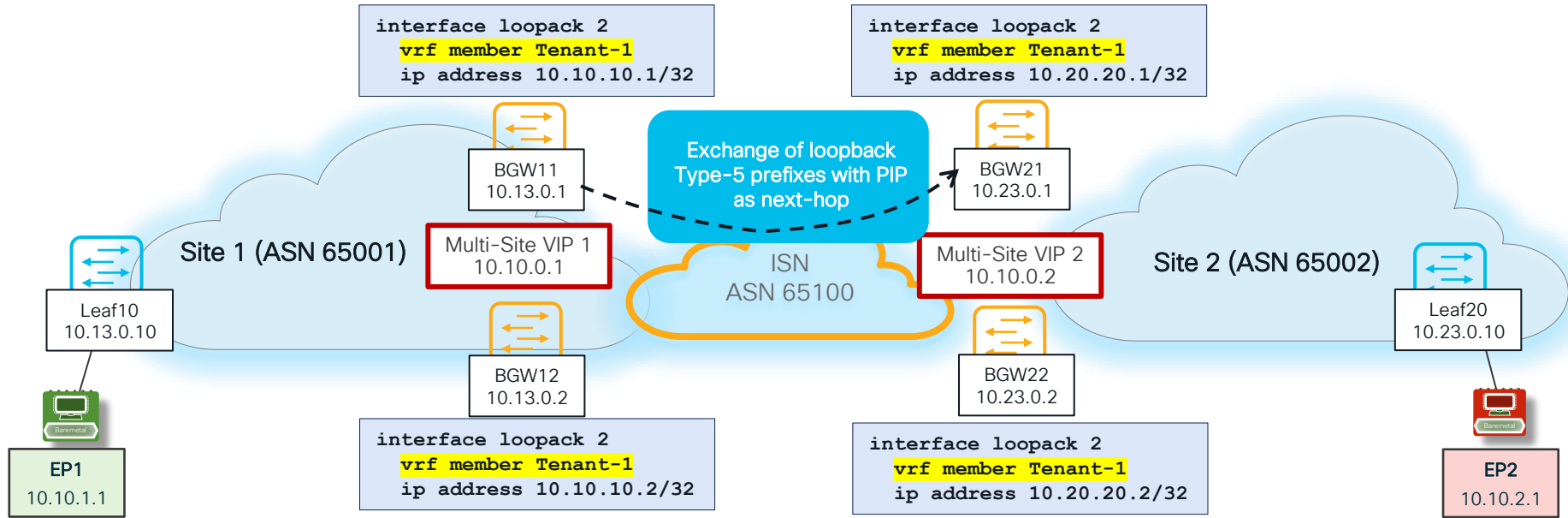
Same applies to Intra-Site traffic between local BGW and leaf nodes



How to Prevent this issue?

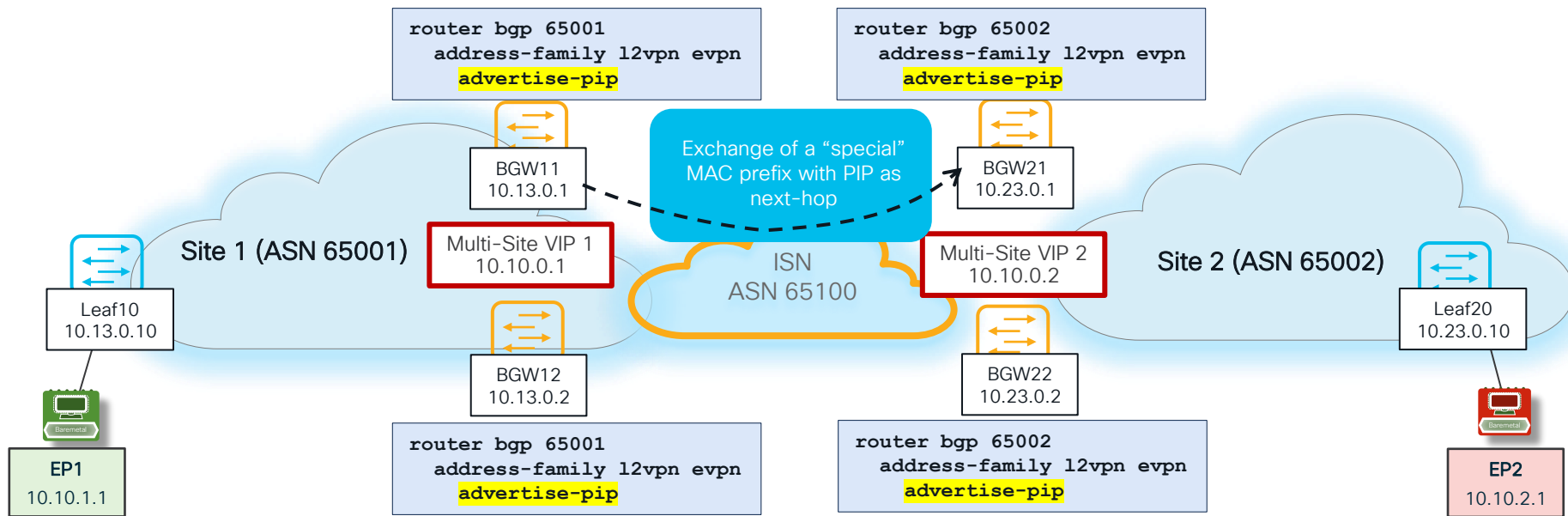
Layer-3 Only Connectivity across Sites

Solution 1: Definition of a Loopback Interface in a Tenant VRF



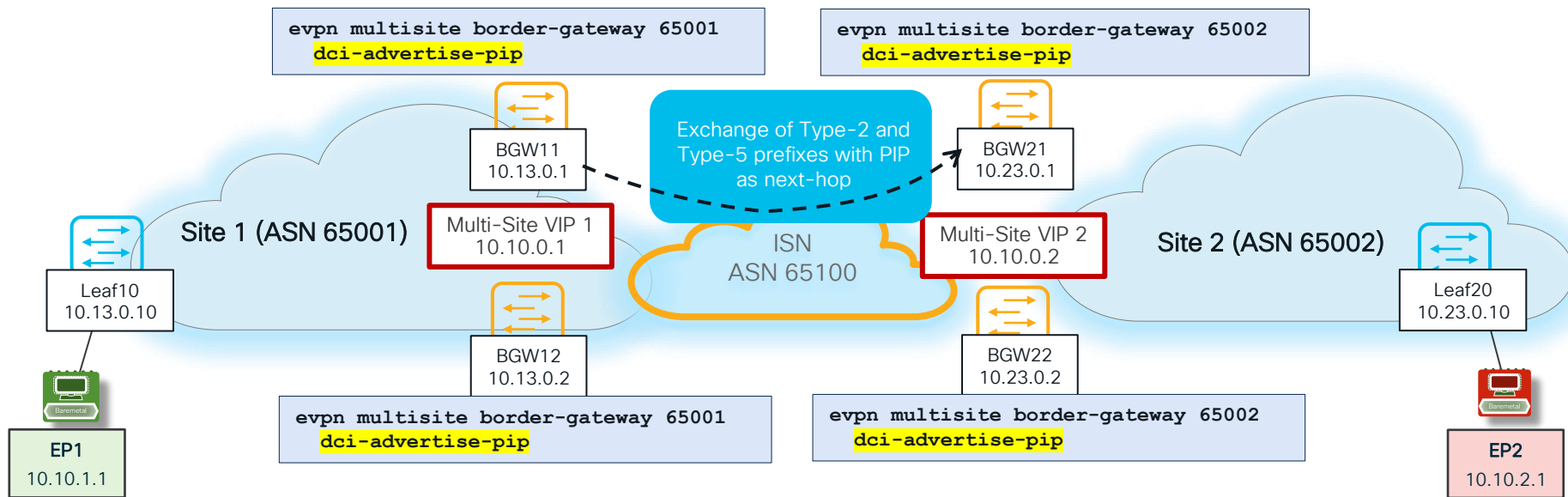
Layer-3 Only Connectivity across Sites

Solution 2: Enable “advertise-pip” on the BGWs



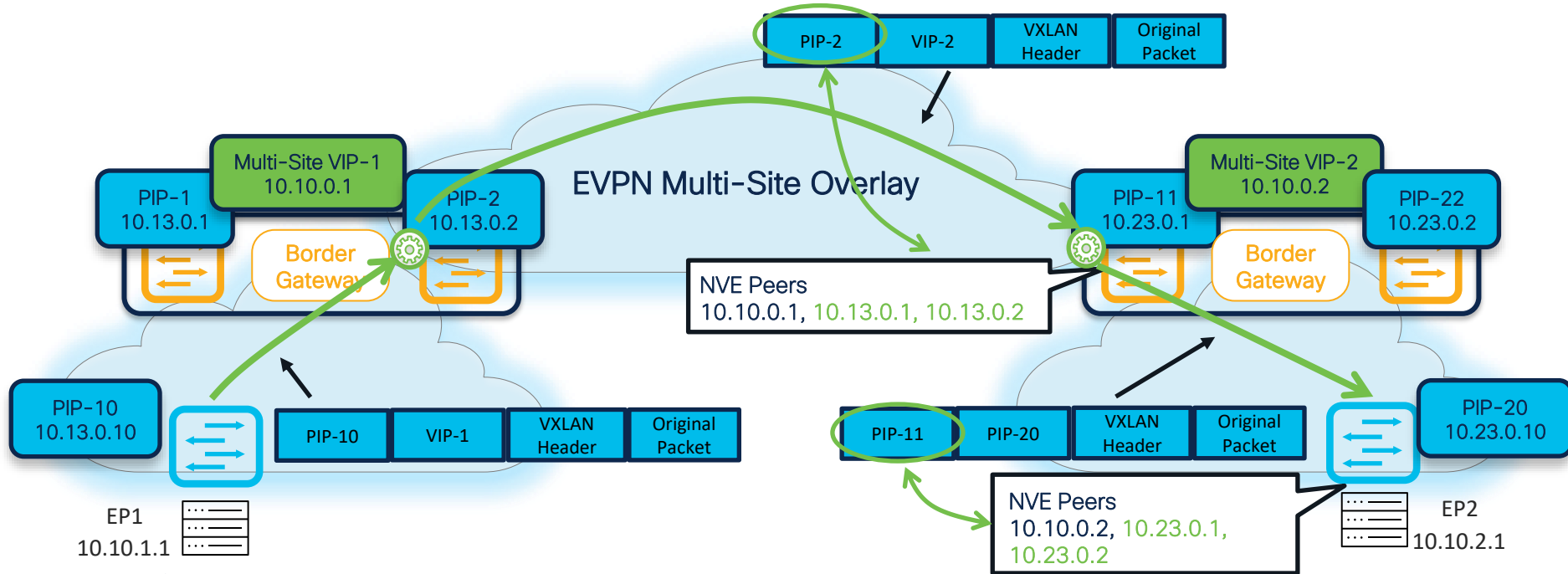
Layer-3 Only Connectivity across Sites

Solution 3: Enable “dci-advertise-pip” on the BGWs



Installing PIP Addresses as NVE Peers

Adopting any of the 3 solutions described in the previous slide ensures that the PIP addresses of the BGWs can be installed as NVE peers in the local leaf nodes and in the remote BGW nodes

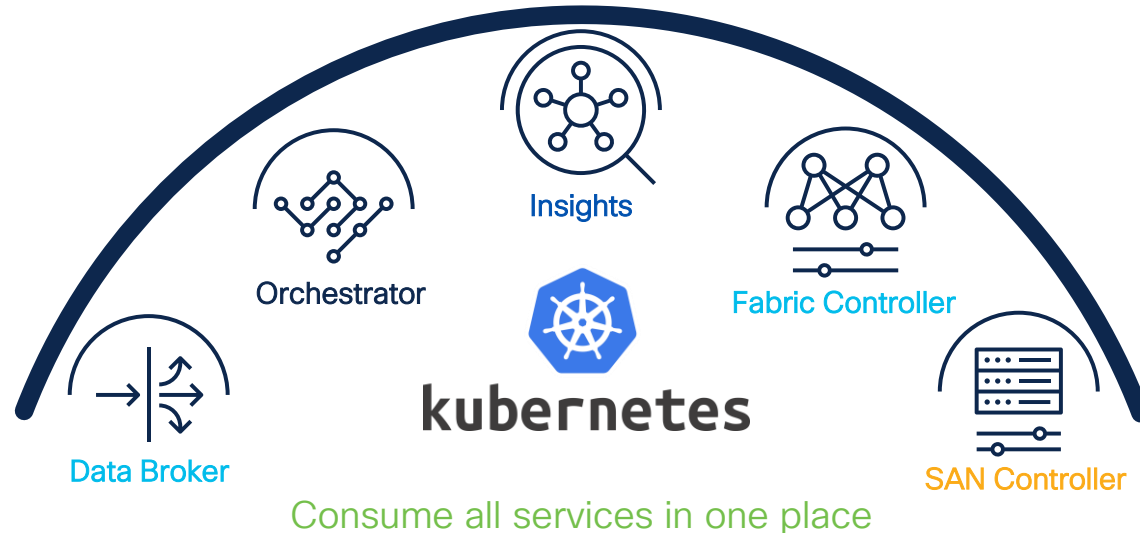


What is Nexus Dashboard Fabric Controller?

Cisco Nexus Dashboard

Simple to automate, simple to consume

Powering automation
Unified agile platform



Private cloud

| Third-party Apps

Public cloud



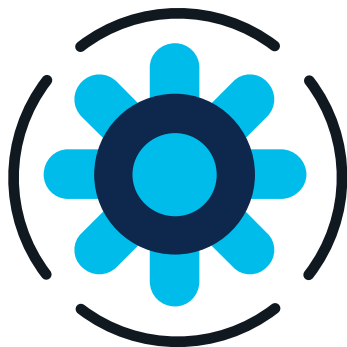
aws



Google Cloud



Cisco Nexus Dashboard Fabric Controller



Automation

Accelerate provisioning
and simplify deployments



Management

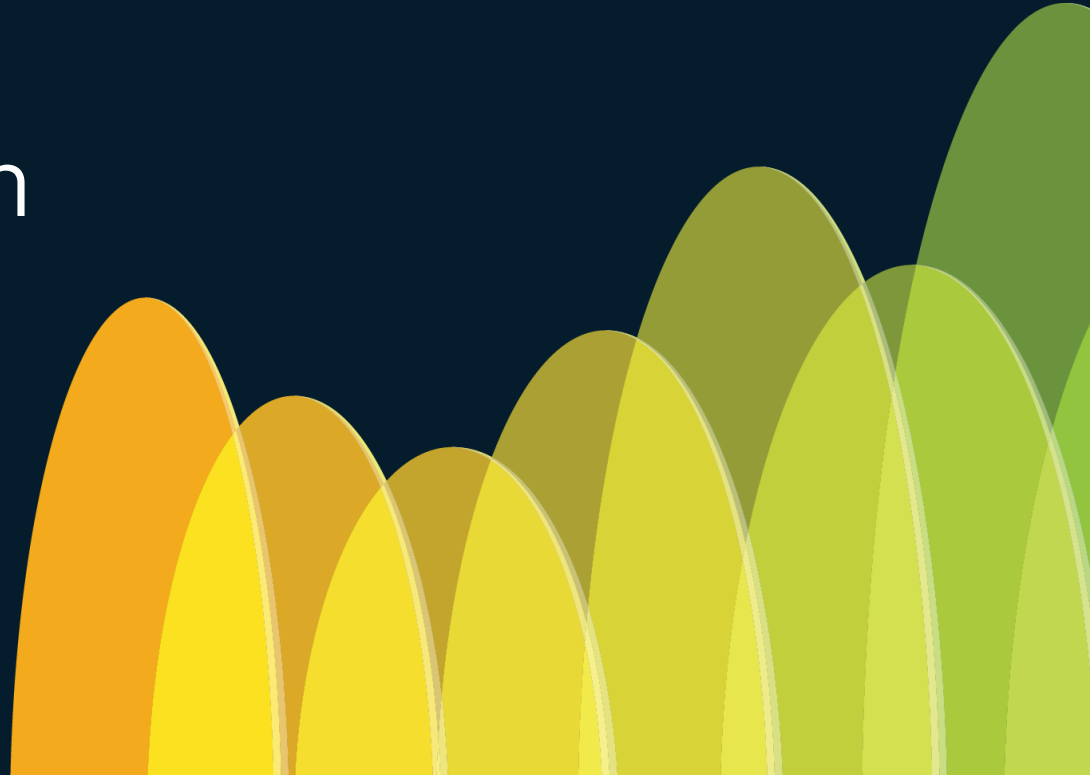
In depth Management
and control for all
network deployments



Visibility

Get Centralized Visibility
and Monitoring views

NDFC Federation One-Manage

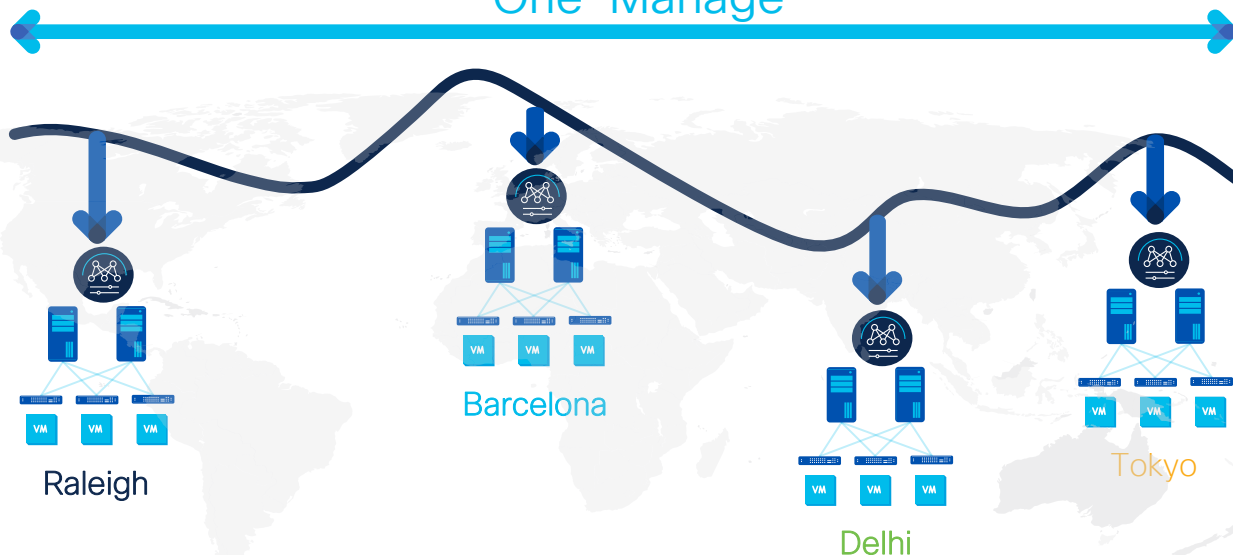


Multi-Cluster NDFC Deployments



Cisco NDFC 12.2(2)

One-Manage



Single sign-on for clusters
(With remote authentication)



One view dashboard to see
summary of fabrics, switch
details etc.



VXLAN EVPN Fabric

- Unified view and automation workflows for Multi-Site deployments
- Federation of MSDs
- L2/L3 Network stretch



No. of clusters in a
federation: 4
(12.2(2) scale)

Benefits

End-to-End connectivity and policy between loosely coupled DC sites

Disaster Recovery, IP Mobility and Active-Active use cases

Network policy consistency

Change control domains separation

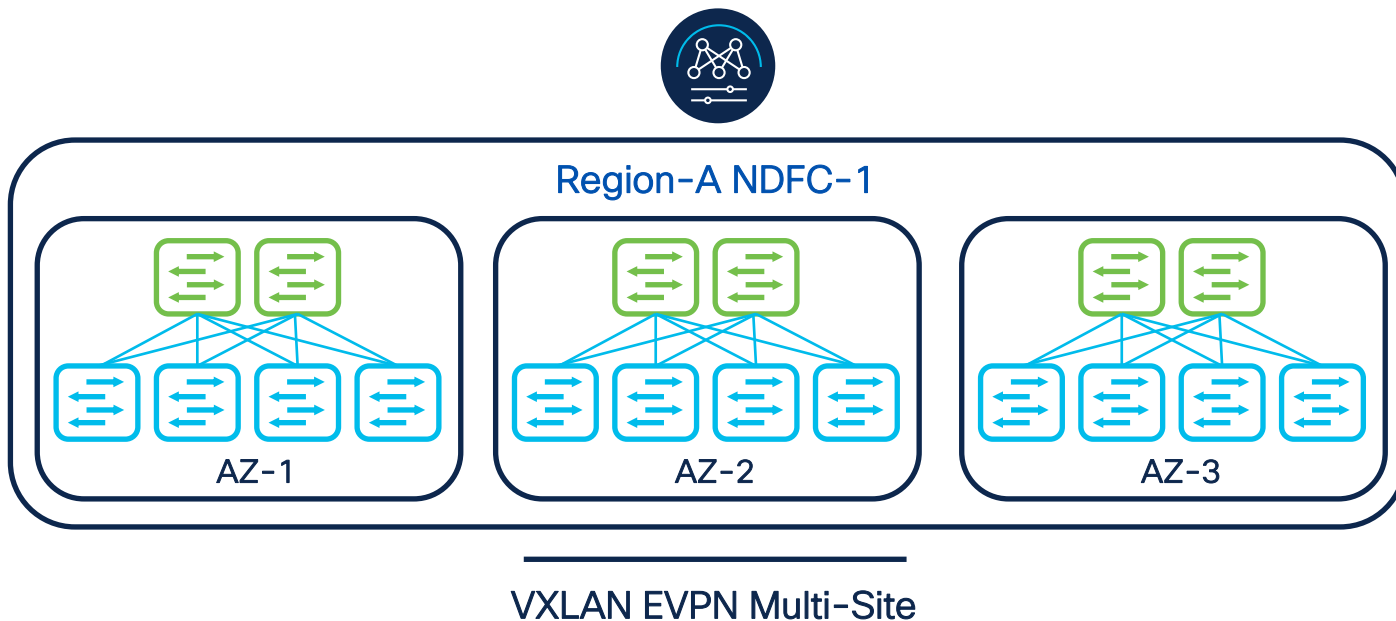
Flexible geo-redundancy for NDFC

Single point of Orchestration

NDFC-managed VXLAN Multi-Site



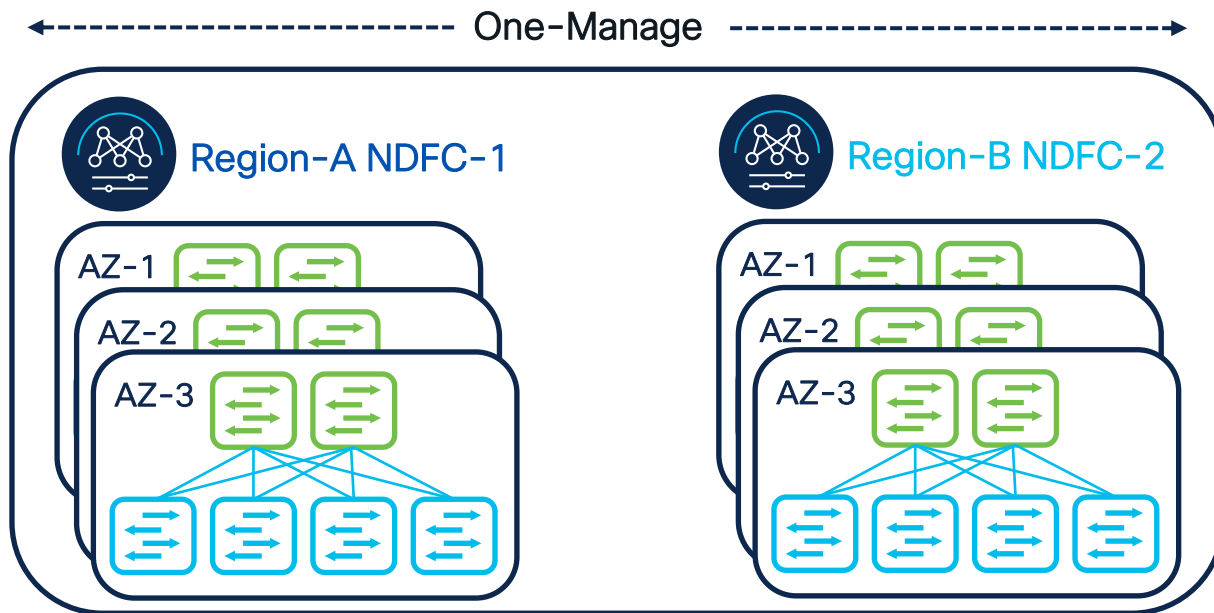
Use-case: Managing multiple DCs in a single instance of NDFC



NDFC-managed VXLAN Multi-Site

One-Manage

Use-case 1: Managing multiple DCs with multiple instances of NDFC

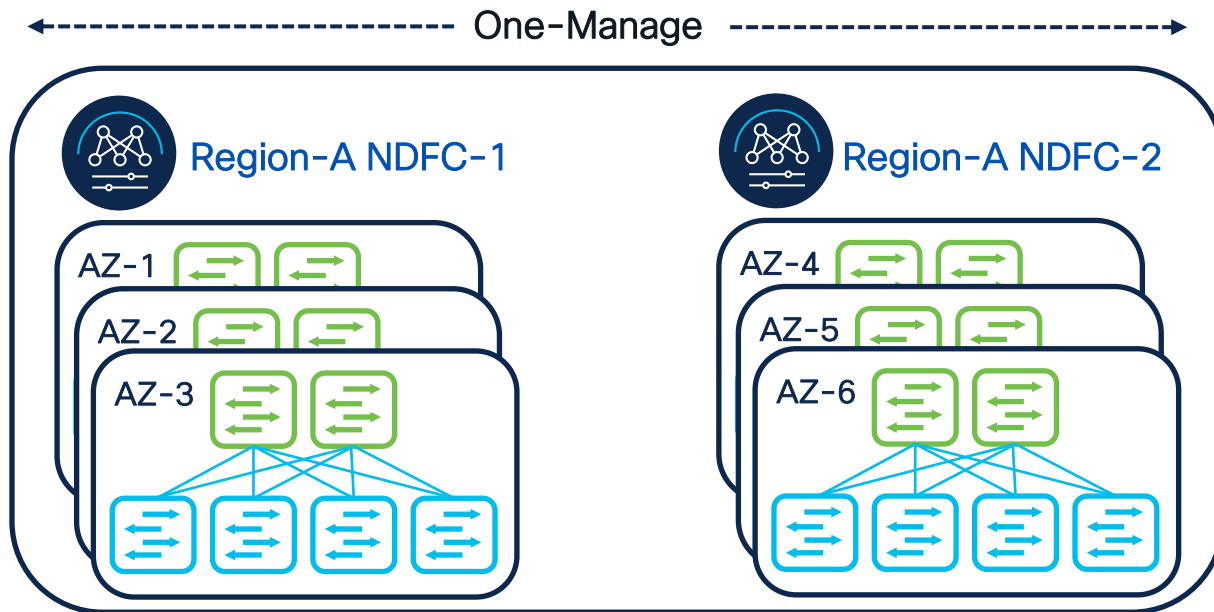


VXLAN EVPN Multi-Site

NDFC-managed VXLAN Multi-Site

One-Manage

Use-case 2: Managing single large-scale DC with 500+ switches



Automate VXLAN Multi-Site with NDFC Federation

One-Manage: Configuration Steps

Step 1 – Create Individual ND clusters and fabrics

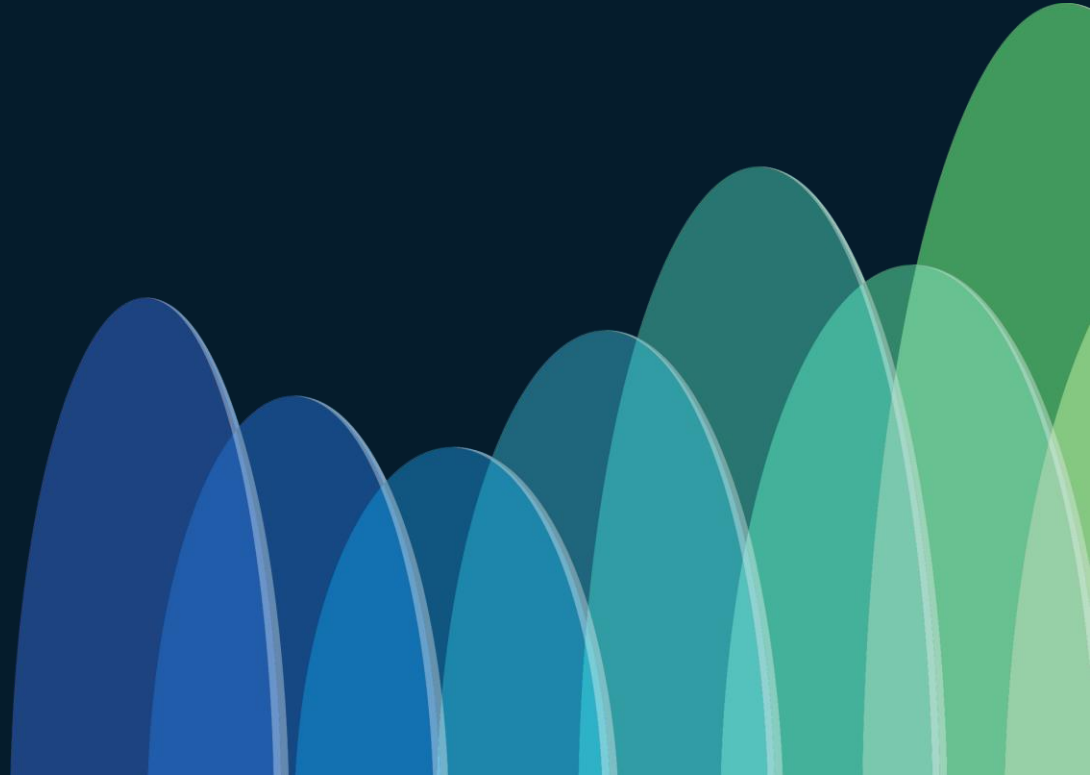
Step 2 – Create ND Federation across Multiple ND clusters

Step 3 – Create Multi-Cluster Fabric from ND Federation and Add Member ND/Fabrics

Step 4 – Recalculate and Deploy Multi-Site DCI Configs across BGWs/RS

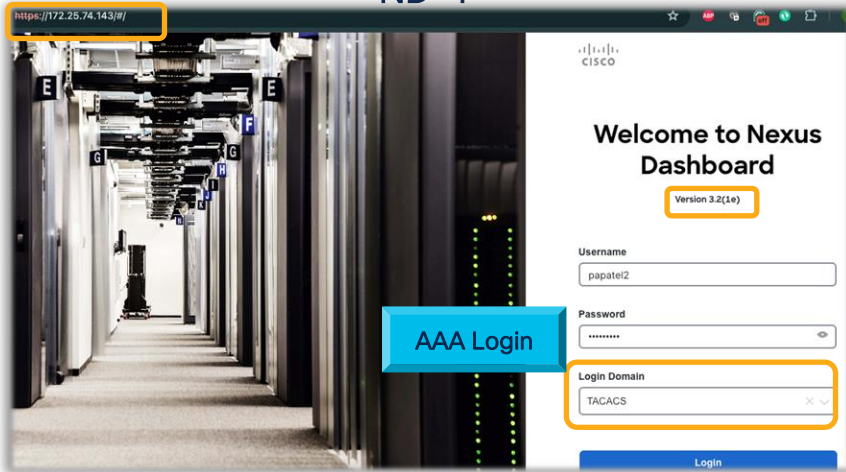
Step 5 – Create, Attach, and Deploy VXLAN Overlays across Multi-Cluster Fabric

NDFC Day-0: Multi-Site DCI

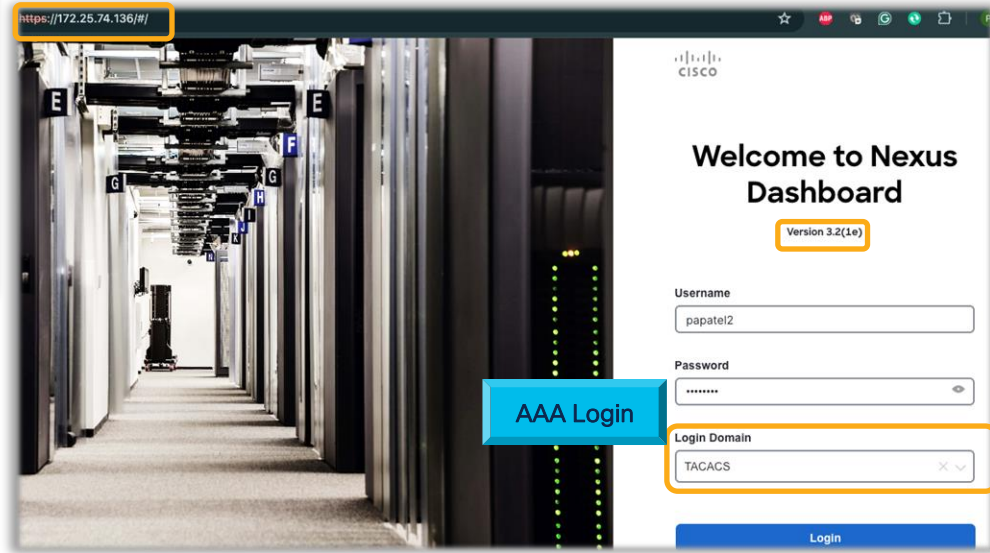


Step 1 -> Create Individual ND clusters and fabrics

ND-1

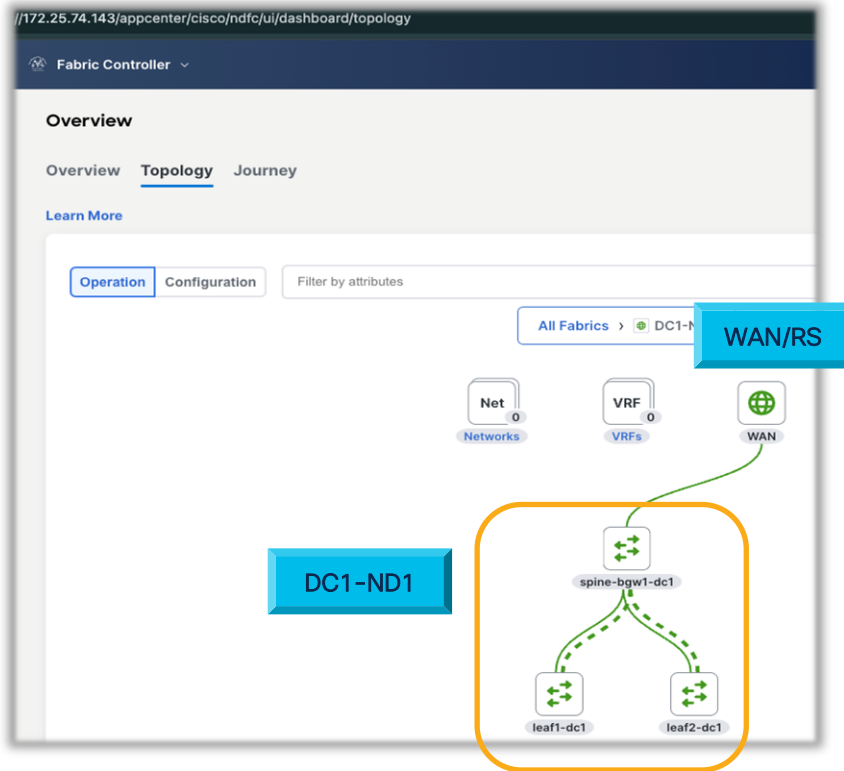


ND-2

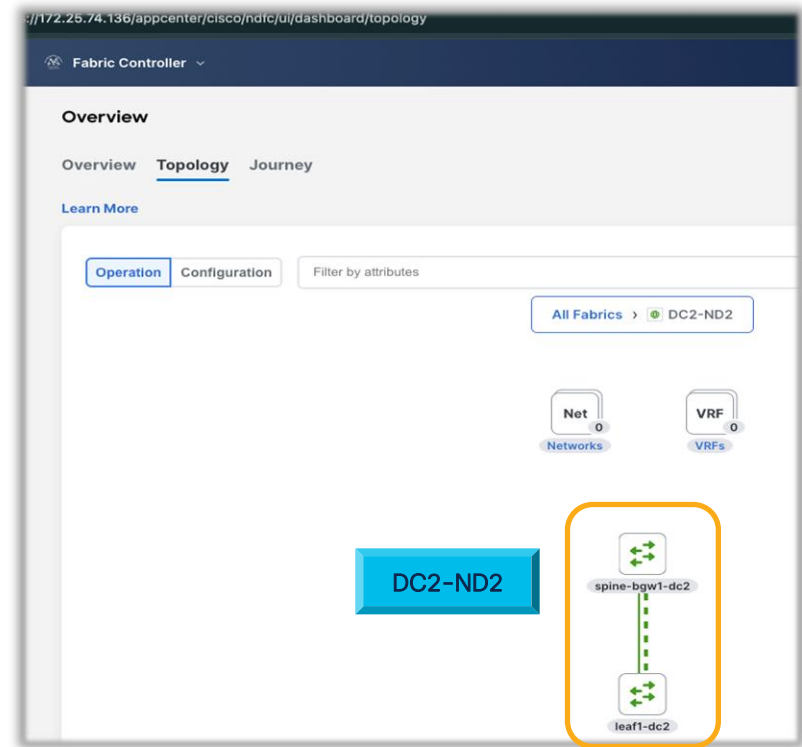


Step 1 -> Create Individual ND clusters and fabrics

ND-1

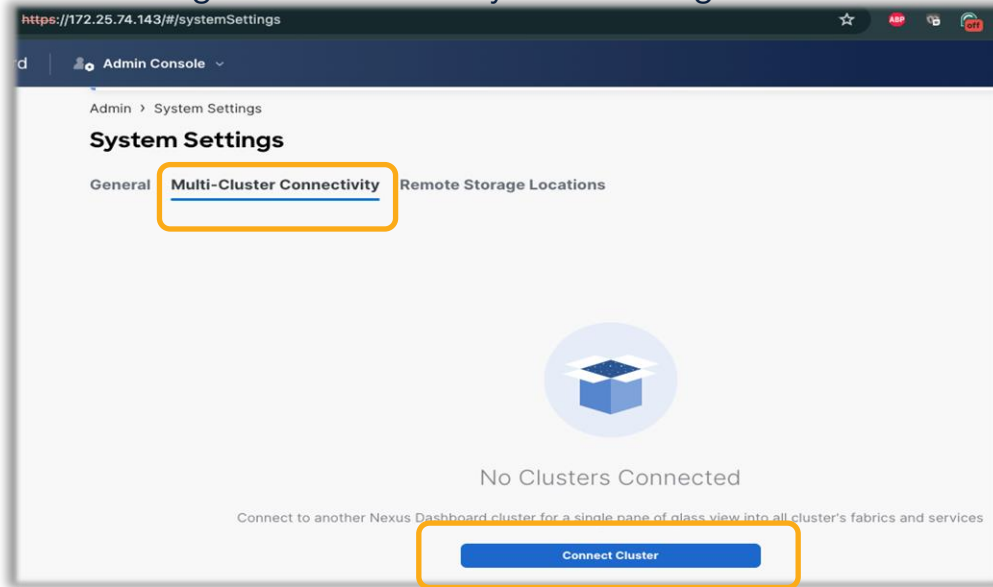


ND-2



Step 2 -> Create ND Federation

- Login to either ND-1 or ND-2 admin console to start ND Federation.
- In this case, we will use ND-1 as the Federation Manager/Primary.
- Navigate to Admin > System Settings > Multi-Cluster Connectivity.



The 'Connect Cluster' dialog box is shown, featuring a blue button at the top right labeled 'Access details of ND2'. The form contains the following fields: 'Hostname/IP Address' with the value '172.25.74.136', 'Username' with the value 'papatel2', 'Password' with masked characters '*****', and 'Login Domain' with a dropdown menu set to 'TACACS'. At the bottom right, there are 'Cancel' and 'Save' buttons, with the 'Save' button highlighted by an orange box. A vertical 'Give feedback' button is located on the right edge of the dialog.

Step 2 -> Create ND Federation



Federation of ND1 and ND2

System Settings

General Multi-Cluster Connectivity Remote Storage Locations

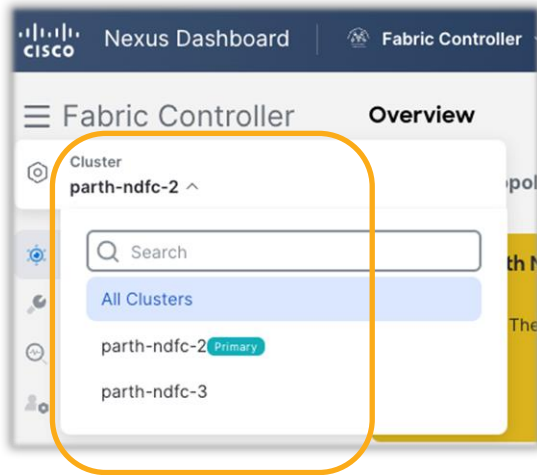
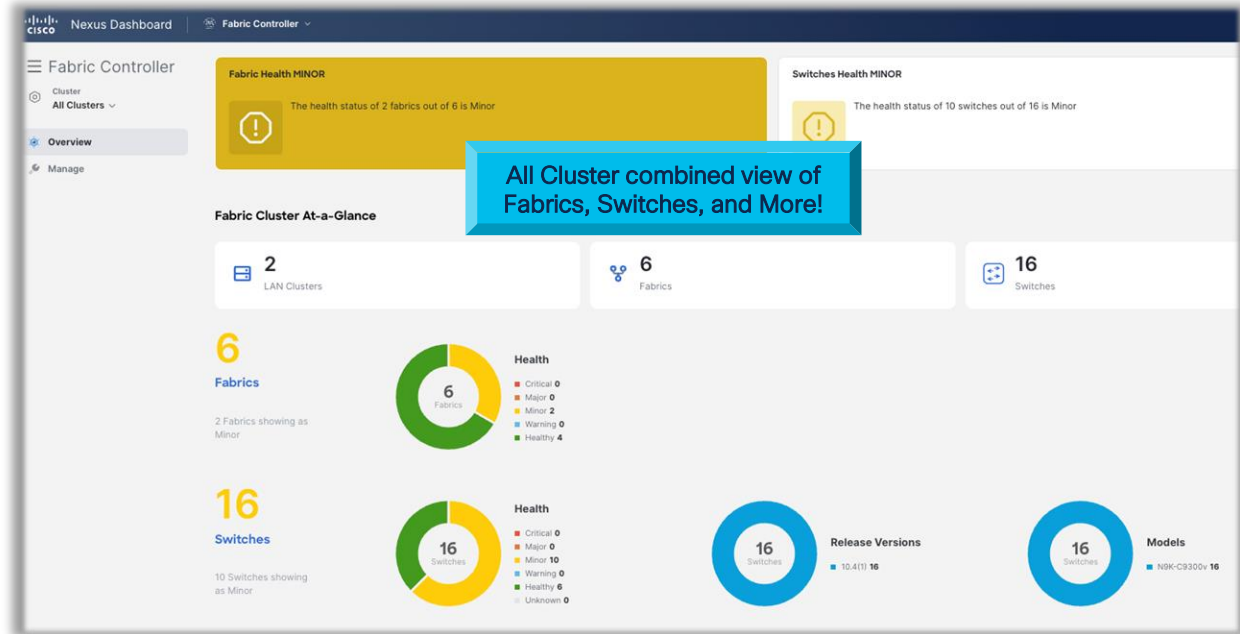
Filter by attributes Connect Cluster Delete Federation

Connectivity Status	Name	URL
↑ Up	parth-ndfc-2 Primary Local	172.25.74.143
↑ Up	parth-ndfc-3	172.25.74.136

2 items found Rows per page 10 < 1 >

Step 3 -> Create Multi-Cluster Fabric

- Login to NDFC-1 from ND-1 and navigate to “All Clusters”



Step 3 -> Create Multi-Cluster Fabric

- Click on Manage > Fabric Groups

The image shows a sequence of three screenshots from the Cisco Nexus Dashboard Fabric Controller interface, illustrating the steps to create a multi-cluster fabric.

First Screenshot: The main dashboard is visible. The 'Manage' menu is open, and 'Fabric Groups' is highlighted. A yellow box highlights the 'Fabric Groups' option in the 'Manage' menu.

Second Screenshot: A close-up of the 'Actions' dropdown menu for 'Fabric Groups'. The options are 'Create Fabric Group', 'Edit Fabric Group', and 'Delete Fabric Group'. A yellow box highlights the 'Create Fabric Group' option.

Third Screenshot: The 'Create Fabric Group' wizard is shown. The first step, 'Select Fabric Group Type', is active. The 'VXLAN' option is selected, indicated by a blue radio button. The description for VXLAN is: 'Domain that can contain multiple VXLAN BGP EVPN Fabrics and External Fabrics.'

Step 3 -> Create Multi-Cluster Fabric

- Select DC1-ND1 Fabric from ND-1 to be imported as first member fabric:controller into One-Manage Multi-Cluster Fabric

The screenshot shows the 'Create Fabric Group' interface. On the left, a sidebar lists the steps: 1. Select Fabric Group Type, 2. Select Child Fabric (current step), 3. Settings, 4. Summary, and 5. Result. The main area is titled 'Select Child Fabric' and contains a table of available fabrics. The table has columns for Fabric Name, Cluster Name, Fabric Technology, Fabric Type, ASN, and Fabric Health. The row for 'DC1-ND1' is selected, indicated by a blue dot and a blue background. An orange box highlights the table and the 'Next' button at the bottom right.

Fabric Name	Cluster Name	Fabric Technology	Fabric Type	ASN	Fabric Health
☐ EXT	parth-ndfc-2	Custom	External Connectivity Network	99	Healthy
☒ DC1-ND1	parth-ndfc-2	VXLAN EVPN	Data Center VXLAN EVPN	65100	Healthy
☐ WAN	parth-ndfc-2	Custom	Multi-Site External Network	65099	Healthy
☐ Fab1	parth-ndfc-2	VXLAN EVPN	Data Center VXLAN EVPN	65001	Minor
☐ DC2-ND2	parth-ndfc-3	VXLAN EVPN	Data Center VXLAN EVPN	65200	Healthy

5 items found

Rows per page: 10

Back Next

Step 3 -> Create Multi-Cluster Fabric

Fabric Group Name

Fabric Group Type
VXLAN EVPN Multi-Site

General Parameters | DCI | Security | Resources | Configuration Backup

Layer 2 VXLAN VNI Range*
 L2VNI Label (for stretched overlays)

Layer 3 VXLAN VNI Range*
 L3VNI Label (for stretched overlays)

Anycast-Gateway-MAC
 Distributed Anycast Gateway (for EP Mobility)

Multi-Site VTEP VIP Loopback Id*
 Multi-Site BGP NH for VXLAN EVPN on BGWs

Border Gateway IP TAG
 Routing tag associated with IP address of loopback and DCI interfaces. Tag value 0-4294967295

Step 3 -> Create Multi-Cluster Fabric

VXLAN Multi-Site
BGW Design Model

BGP EVPN Router-Server
Loopback IP for EVPN
peering

Auto Deploy Multi-Site
Underlay Configs

Enable BGP send-community

General Parameters **DCI** Security Resources Configuration Backup

Multi-Site Overlay IFC Deployment Method*
Centralized_To_Route_Server
Manual, Auto Overlay EVPN Peering to Route Servers, Auto Overlay EVPN Direct Peering to Border Gateways

Multi-Site Route Server List*
9.9.9.9
Multi-Site Route-Server peer list (typically loopback IP address on Route-Server for Multi-Site EVPN peering with BGWs), e.g. 128.89.0.1, 128.89.0.2. Provide a list of IPv6 addresses for IPv6 underlay, e.g. fd00::b10:1, fd00::b10:2

Multi-Site Route Server BGP ASN
65099
BGP ASN of Route-Server
5535[0-65535], e.g. 65000, 65001

Enable 'redistribute direct' on Route Servers
☐
For auto-created Multi-Site overlay IFCs in Route Servers. Applicable only when Multi-Site Overlay IFC Deployment Method is Centralized_To_Route_Server

Route Server IP TAG

Routing tag associated with Route Server IP for redistribute direct. This is the IP used in eBGP EVPN peering. Tag value 0-4294967295

Multi-Site Underlay IFC Auto Deployment Flag
☒

BGP Send-community on Multi-Site Underlay IFC
☒
For auto-created Multi-Site Underlay IFCs

Step 3 -> Create Multi-Cluster Fabric

The screenshot shows the 'Resources' tab in a configuration interface. It contains three input fields with associated callouts:

- Multi-Site VTEP VIP Loopback IP Range***: The input is `10.10.0.0/24`. A blue callout box states: "Multi-Site Loopback IP range allocation on BGWs". A note to the right says: "Typically Loopback100 IP Address Range".
- DCI Subnet IP Range***: The input is `10.10.1.0/24`. A blue callout box states: "DCI Underlay P2P Subnet". A note to the right says: "Address range to assign P2P DCI Links".
- Subnet Target Mask***: The input is `30`. A blue callout box states: "DCI Underlay P2P Subnet Mask". A note to the right says: "Target Mask for Subnet Range (Min:8, Max:31)".

Step 3 -> Create Multi-Cluster Fabric

- Select Multi-Cluster Fabric Group and Add more Fabrics:Controller
- Add WAN-ND1 and DC2-ND2

Add Child Fabric ×

Filter by attributes

Fabric Name	Cluster Name	Fabric Technology	Fabric Type	ASN	Fabric Health
☐ EXT	parth-ndfc-2	Custom	External Connectivity Network	99	 Healthy
☒ WAN	parth-ndfc-2	Custom	Multi-Site External Network	65099	 Healthy
☐ Fab1	parth-ndfc-2	VXLAN EVPN	Data Center VXLAN EVPN	65001	 Minor
☐ DC2-ND2	 parth-ndfc-3	VXLAN EVPN	Data Center VXLAN EVPN	65200	 Healthy

Step 4 -> Perform R&D for DCI Configs

Multi-Cluster Fabric Group - One-Manage-DCI

Overview Child Fabrics Switches Links Networks VRFs Event Analytics History Resources

Filter by attributes

	Fabric Name	Fabric Technology	Fabric Type	ASN	Cluster	MSD Name	Fabric Health
☐	DC1-ND1	VXLAN EVPN	Data Center VXLAN EVPN	65100	parth-ndfc-2	One-Manage-DCI	 Healthy
☐	WAN	Custom	Multi-Site External Network	65099	parth-ndfc-2	One-Manage-DCI	 Healthy
☐	DC2-ND2	VXLAN EVPN	Data Center VXLAN EVPN	65200	parth-ndfc-3	One-Manage-DCI	 Healthy

Actions ^

- Edit Fabric
- Add Child Fabric
- Recalculate and Deploy
- More >

Step 4 -> Perform R&D for DCI Configs

Pending Config - One-Manage-DCI - route-server1

EVPN Overlay
Feature

Side-by-Side Comparison

```
nv overlay evpn
feature nv overlay
interface ethernet1/1
no switchport
ip address 10.10.1.5/30 tag 54321
mtu 9216
no shutdown
```

DCI Underlay
Routed Port
towards BGW

Retain RT coming from
BGW EVPN NLRI

EVPN NH unchanged on RS for
BGW-to-BGW VTEP termination

```
router bgp 65099
address-family ipv4 unicast
maximum-paths 1
maximum-paths ibgp 1
network 9.9.9.9/32
exit
address-family ipv6 unicast
maximum-paths 1
maximum-paths ibgp 1
exit
address-family l2vpn evpn
retain route-target all
exit
template peer OVERLAY-PEERING
update-source loopback0
ebgp-multihop 5
address-family l2vpn evpn
route-map unchanged out
send-community both
exit
exit
route-map unchanged permit 10
set ip next-hop unchanged
configure terminal
```

eBGP IPv4 Unicast (default) DCI
Underlay peering towards BGW

```
neighbor 10.10.1.1
remote-as 65100
update-source Ethernet1/2
address-family ipv4 unicast
next-hop-self
send-community both
exit
```

eBGP EVPN DCI Overlay
peering towards BGW

```
neighbor 10.2.0.3
remote-as 65100
inherit peer OVERLAY-PEERING
address-family l2vpn evpn
rewrite-evpn-rt-asn
ex
```

Rewrite ASN portion of Auto RT

Step 4 -> Perform R&D for DCI Configs

Pending Config - One-Managed-DCI - spine-bgw1-dc1

Pending Config Side-by-Side Comparison

```
route-map rmap-redirect-direct permit 10
  match tag 54321
evpn multisite border-gateway 65100
  delay-restore time 300
router bgp 65100
  address-family ipv4 unicast
    redistribute direct route-map rmap-redirect-direct
    maximum-paths 64
    maximum-paths ibgp 64
  exit
```

BGW Interface
advertisement for DCI
reachability

BGW Site ID

```
neighbor 10.10.1.2
  remote-as 65099
  update-source Ethernet1/3
  address-family ipv4 unicast
    next-hop-self
    send-community both
  exit
```

eBGP IPv4 Unicast
(default) DCI Underlay
peering towards RS

```
neighbor 9.9.9.9
  remote-as 65099
  update-source loopback0
  ebgp-multihop 5
  peer-type fabric-external
  address-family l2vpn evpn
    send-community both
    rewrite-evpn-rt-asn
```

Rewrite and Re-origination
function on BGW

eBGP EVPN
DCI Overlay
peering
towards RS

Rewrite ASN
portion of Auto
RT

BGW VTEP Interface
Lo1 = NVE PIP
Lo100 = NVE Multi-Site
Anycast VIP

```
interface nve1
  multisite border-gateway interface loopback100
interface loopback100
  ip address 10.10.0.1/32 tag 54321
  ip router ospf UNDERLAY area 0.0.0.0
  ip pim sparse-mode
  no shutdown
interface ethernet1/3
  no switchport
  ip address 10.10.1.1/30 tag 54321
  evpn multisite dci-tracking
  mtu 9216
  no shutdown
interface ethernet1/1
  no switchport
  ip address 10.4.0.2/30
  evpn multisite fabric-tracking
  description connected-to-leaf1-dc1-Ethernet1/1
  mtu 9216
  ip router ospf UNDERLAY area 0.0.0.0
  ip ospf network point-to-point
  ip pim sparse-mode
  no shutdown
```

BGW DCI-tracking towards
Site-External underlay

BGW Fabric-tracking towards
Site-Internal underlay

Step 5 -> Create, Attach, and Deploy Overlays

- Navigate to Multi-Cluster Fabric Group > Networks > Create

The screenshot shows the 'Create Network' form with several blue callout boxes highlighting specific fields and buttons:

- Auto Generated Name**: Points to the 'Network Name' field containing 'MyNetwork_30000'.
- L2 or L3 Network**: Points to the 'Layer 2 Only' checkbox, which is currently unchecked.
- VXLAN Overlay VRF**: Points to the 'Create VRF' button next to the 'VRF Name' field.
- Auto Generated L2VNI Label**: Points to the 'Network ID' field containing '30000'.
- Network SVI**: Points to the 'IPv4 Gateway/NetMask' field containing '10.10.10.1/24'.

Other visible fields and buttons include:

- VRF Name**: 'MyVRF_50000' with a 'Create VRF' button.
- Network ID**: '30000'.
- VLAN ID**: An empty field with a 'Propose VLAN' button.
- Network Template**: 'Default_Network_Universal >'.
- Network Extension Template**: 'Default_Network_Extension_Universal >'.
- General Parameters** and **Advanced** tabs.
- Close** and **Create** buttons at the bottom right.

Step 5 -> Create, Attach, and Deploy Overlays

OverviewChild FabricsSwitchesLinksNetworksVRFsEvent AnalyticsHistoryResources

Select Network

☒	Network Name	Network ID	VRF Name	IPv4 Gateway/Prefix	IPv6 Gateway/Prefix	Network Status	VLAN ID	VLAN N
☒	MyNetwork_30000	30000	MyVRF_50000	10.10.10.1/24		NA		

Create

Edit

Multi-Attach

Multi-Detach

Deploy

Select Switches to attach all Selected Networks (1)

Total No. of Attachment : 5

Filter by attributes

☒	Cluster Name	Fabric Name	Switch	IP Address	Serial Number	Model Number	Role	VPC Peer
☒	parth-ndfc-2	DC1-ND1	leaf1-dc1	192.168.100.71	9HHS3S8BIO7	N9K-C9300v	leaf	
☒	parth-ndfc-2	DC1-ND1	leaf2-dc1	192.168.100.72	9K4FDX6S4UU	N9K-C9300v	leaf	
☒	parth-ndfc-2	DC1-ND1	spine-bgw1-dc1	192.168.100.73	9NFAK9M6CZ5	N9K-C9300v	border gateway spine	
☒	parth-ndfc-3	DC2-ND2	leaf1-dc2	192.168.100.76	923KAU3N5OS	N9K-C9300v	leaf	
☒	parth-ndfc-3	DC2-ND2	spine-bgw1-dc2	192.168.100.75	9N19GSLKTYD	N9K-C9300v	border gateway spine	

Select Interfaces

Filter by attributes

Bulk Paste

☐	Network Name	Switch Name	Peer Switch Name	ToR Switches	Interfaces List	Action
☐	MyNetwork_30000	leaf1-dc1			eth1/8-9	View Interfaces
☐	MyNetwork_30000	leaf2-dc1			eth1/8-9	View Interfaces
☐	MyNetwork_30000	spine-bgw1-dc1				View Interfaces
☐	MyNetwork_30000	spine-bgw1-dc2				View Interfaces
☐	MyNetwork_30000	leaf1-dc2			eth1/8-9	View Interfaces

CancelPreviousNext

Step 5 -> Create, Attach, and Deploy Overlays

Summary

Networks Selected
1

Devices Selected
5

Network Attachments
5

Device Interface Association
6

Device Interface De-Association
0

☐ Deploy Later

☒ Proceed to Full Switch Deploy(Recommended)

☐ Proceed to Individual Network Deploy

Filter by attributes

Resync All

Cluster Name	Fabric Name	Switch Name	IP Address	Role	Serial Number	Fabric Status	Pending Config	Status Description	Progress	Resync Switch
parth-ndfc-2	DC1-ND1	spine-bgw1-dc1	192.168.100.73	border gateway spine	9NFAK9M6CZ5	Out-Of-Sync	48 Lines	Out-of-Sync	<div></div>	Resync
parth-ndfc-2	DC1-ND1	leaf2-dc1	192.168.100.72	leaf	9K4FDX6S4UU	Out-Of-Sync	67 Lines	Out-of-Sync	<div></div>	Resync
parth-ndfc-2	DC1-ND1	leaf1-dc1	192.168.100.71	leaf	9HHS3S8BIO7	Out-Of-Sync	67 Lines	Out-of-Sync	<div></div>	Resync
parth-ndfc-3	DC2-ND2	leaf1-dc2	192.168.100.76	leaf	923KAU3N5OS	Out-Of-Sync	67 Lines	Out-of-Sync	<div></div>	Resync
parth-ndfc-3	DC2-ND2	spine-bgw1-dc2	192.168.100.75	border gateway spine	9N19GSLKTYD	Out-Of-Sync	48 Lines	Out-of-Sync	<div></div>	Resync

Step 5 -> Create, Attach, and Deploy Overlays

Pending Config - One-Managed-DCI - spine-bgw1-dc1

Pending Config Side-by-Side

L3VNI VRF Configs

```
vlan 2000
  vn-segment 50000
  configure terminal
  vrf context myvrf_50000
    vni 50000
    rd auto
    address-family ipv4 unicast
      route-target both auto
    address-family ipv6 unicast
      route-target both auto
  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
```

VXLAN Tenant

```
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 50000 associate-vrf
  member vni 30000
  multisite ingress-replication
  mcast-group 239.1.1.1
```

L2VNI Network Configs

```
vlan 2300
  vn-segment 30000
  configure terminal
  evpn
    vni 30000 l2
    rd auto
    route-target import auto
    route-target export auto
```

NVE Tunnel Configs

Step 5 -> Create, Attach, and Deploy Overlays

Pending Config - One-Managed-DCI - leaf1-dc1

Pending Config Side-by-Side Comparison

```
vlan 2000
  vn-segment 50000
  configure terminal
  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
  exit
configure terminal
interface nve1
  member vni 50000 associate-vrf
  member vni 30000
  mcast-group 239.1.1.1
```

VXLAN Tenant

NVE Tunnel Configs

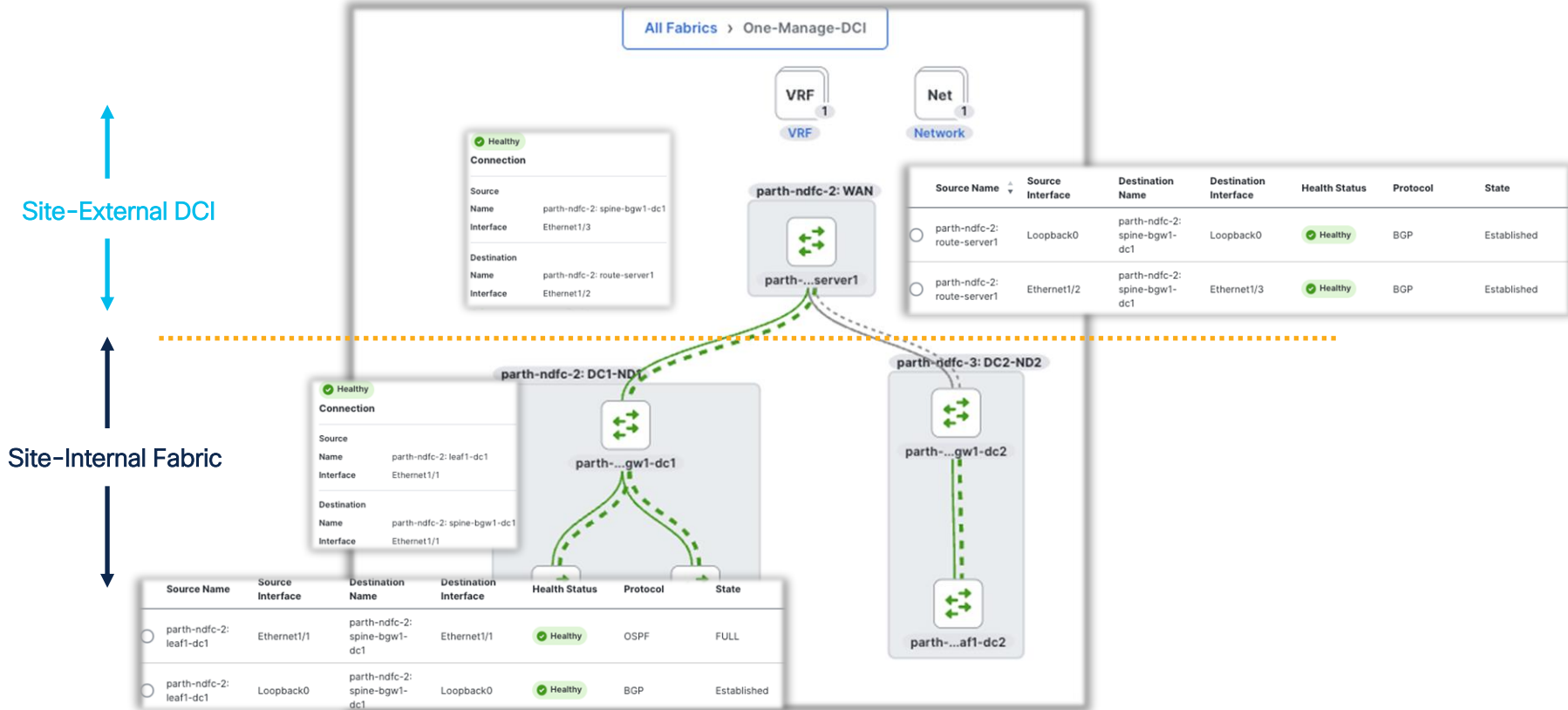
```
vlan 2300
  vn-segment 30000
  configure terminal
  interface Vlan2300
    vrf member myvrf_50000
    ip address 10.10.10.1/24 tag 12345
    fabric forwarding mode anycast-gateway
    no shutdown
  exit
  evpn
    vni 30000 l2
    rd auto
    route-target import auto
    route-target export auto
```

L3 SVI Configs

L2VNI Network Configs

```
interface ethernet1/8
  switchport
  switchport mode trunk
  mtu 9216
  spanning-tree port type edge trunk
  no shutdown
  switchport trunk allowed vlan 2300
interface ethernet1/9
  switchport
  switchport mode trunk
  mtu 9216
  spanning-tree port type edge trunk
  no shutdown
  switchport trunk allowed vlan 2300
```

One-Manage Multi-Site Topology View



Verification and Validation with NDFC

Verification through NDFC

Keeping you away from CLI



Step 1

Verify Network and VRF attachments

Job Execution Status:

Network Status **Deployed**

VRF Status **Deployed**



Step 2

Deployment History

Configuration Execution Status:

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



Step 3

Show commands

Service / features status
(CLI through NDFC)

Attachment deployment status

Job execution perspective

1

Config Preview

2

Deploy Progress

Filter by attributes

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

Success or Failure deployment details

Multi-Stage Preview and Deployment

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

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	
Edge-router	Vlan100	INTERFACE	UNDERLAY	Detailed History	● FAILED	91UP4O2KI8R	
Edge-router	SWITCH	SWITCH	UNDERLAY	Detailed History	● SUCCESS	91UP4O2KI8R	

CLI response messages for easier troubleshooting

Command Execution Details for Edge-router(91UP4O2KI8R)

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

Successfully deployed admin May 02 2023, 12:20:20

CLI command #7 hsrp version 2 failed due to error invalid command

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

Switches

Fabric Name == DC1 X

Switch	IP Address	Role	Serial
BGW1-DC1	192.168.101.25	Gateway	9592
LEAF1-DC1	192.168.101.21	Leaf	97EV
LEAF2-DC1	192.168.101.22	Leaf	988K
☒ LEAF3-DC1	192.168.101.23	Leaf	9V7C

50 Rows

Actions ^

- Change Mode
- Provision RMA
- Change Serial Number
- Copy Run Start
- Reload
- Restore Switch
- Show Commands
- Exec Commands
- Delete Switch(es)

More

Commands*

- show
- current_running_config
- show
- show_bgp_evpn_neighbors
- show_bgp_l2vpn_evpn_summary
- show_bgp_sessions
- show_capture_elam
- show_epld_boot_info
- show_mac_and_arp
- show_sys_state

NDFC pre-built
commands or user
commands

```
1 #show ip ospf nei
2
3 OSPF Process ID UNDERLAY VRF default
4 Total number of neighbors: 2
5 Neighbor ID    Pri State          Up Time  Address      Interface
6 10.2.0.6       1 FULL/ -       2d22h    10.4.0.29    Eth1/2
7 10.2.0.1       1 FULL/ -       2d22h    10.4.0.21    Eth1/3
8
9 #show bgp l2vpn evpn su
10 BGP summary information for VRF default, address family L2VPN EVPN
11 BGP router identifier 10.2.0.3, local AS number 65001
12 BGP table version is 29, L2VPN EVPN config peers 2, capable peers 2
13 9 network entries and 16 paths using 2436 bytes of memory
14 BGP attribute entries [16/2752], BGP AS path entries [1/10]
15 BGP community entries [0/0], BGP clusterlist entries [6/24]
16
17 Neighbor      V  AS MsgRcvd MsgSent  TblVer  InQ OutQ Up/Down State/PfxRcd
18 10.2.0.1      4 65001  4252   4240    29     0   0   2d22h  5
19 10.2.0.6      4 65001  4252   4239    29     0   0   2d22h  5
20
21 #show int nve1
22 nve1 is up
23 admin state is up, Hardware: NVE
24 MTU 9216 bytes
25 Encapsulation VXLAN
```

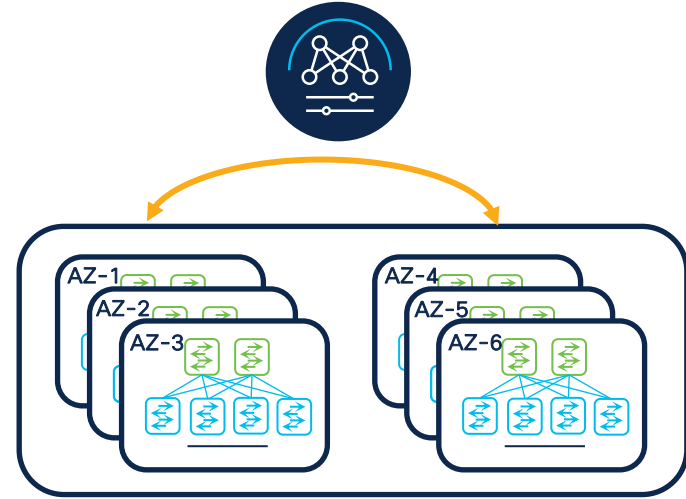
Appendix

White Paper and Document References

- VXLAN BGP EVPN
 - <https://www.cisco.com/c/en/us/products/switches/nexus-9000-series-switches/white-paper-listing.html>
- NDFC
 - <https://www.cisco.com/c/en/us/products/cloud-systems-management/prime-data-center-network-manager/white-paper-listing.html>

Conclusion

Key points to remember



- VXLAN EVPN Multi-Site maintains clear change and fault domain separation to deploy large-scale and highly available DC architectures.
- NDFC provides simplified mechanism to extend and provide end-to-end network and policy consistency across regions, all with a single point of orchestration.
- 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 claim a **Cisco Live T-Shirt**.



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



Continue your education



- Visit the Cisco Stand for related demos
- Book your one-on-one Meet the Expert 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



Thank you

CISCO *Live!*

#CiscoLiveAPJC

CISCO *Live!*

GO BEYOND

#CiscoLiveAPJC