

Deploying Shared-Border in VXLAN Multi-Site with Cisco NDFC

Introduction

The Shared-Border is a common external connectivity point for multiple VXLAN BGP EVPN fabrics interconnected with EVPN Multi-Site architecture. Unlike the BGW (Border Gateway), the Shared-Border does not have any specific requirement other than “normal” VXLAN EVPN support; it is solely a Shared-Border node topologically outside one or more sites. The Shared-Border operates like a traditional VTEP, but unlike the Site-Internal VTEPs, the Shared-Border is a Site-External VTEP. In the case of external connectivity, the Shared-Border operates solely in Layer 3 mode; hence, no BUM replication between the BGW nodes of the VXLAN EVPN fabrics and Shared-Border nodes is necessary. We must configure the VXLAN BGP EVPN VTEP on the Shared-border, and it must be present in a different autonomous system than the one that includes the BGWs.

Depending on hardware and software capabilities, the Shared-Border can enable external connectivity with various Layer 3 technologies. Some examples are Cisco Nexus 9000 Series switches (VRF Lite and MPLS L3VPN), Cisco Nexus 7000 Series Switches (VRF Lite, MPLS L3VPN, and LISP), Cisco ASR 9000 Series Aggregation Services Routers (VRF Lite and MPLS L3VPN), and Cisco ASR 1000 Series routers (VRF Lite and MPLS L3VPN).

Shared-Border Use-Cases

Flexible integration in a scalable Multi-Site architecture

Today, large Enterprises and Service Providers deploy scalable data centers with upwards of 1000 racks within the data center's physical location. To simplify operations and limit the fault domain, the data center should logically be segmented into smaller fabrics, able to extend any VRF and network anywhere within the data center.

For example, assume that we need to design a large VXLAN EVPN data center in New York with 500 switches while considering future growth, availability, and scalability. Today, NX-OS supports 512 VTEPs in a single fabric. A VTEP is a Nexus 9000 switch acting as a VXLAN Tunnel End Point to encapsulate Layer 2 and Layer 3 VXLAN Overlay traffic over a generic IP-routed fabric.

One way to accomplish this is to design a large spine-leaf VXLAN fabric with 500 switches. However, this approach can introduce challenges such as a common underlay plane, common overlay plane, faith sharing, single point of change, admin, and fault domain.

Instead, another approach is to implement VXLAN EVPN Multi-Site to address all the shortcomings of a single fabric option. Some of the key advantages of Multi-Site are the following:

1. Multiple underlay domains - Isolated
2. Multiple replicate domains for BUM - Interconnected and controlled
3. Multiple overlay domains - Interconnected and controlled
4. Multiple overlay control plane domains - Interconnected and controlled
5. Multiple VNI admin domains - Downstream VNI
6. Flexible Layer 2 and Layer 3 DCI services
7. VXLAN to IP handoff
8. Layer 4 to Layer 7 service insertion and redirection
9. Integration with legacy networks (vPC, FabricPath)

10. VXLAN Layer 3 extension to Public Cloud

Hence, splitting a single 500-switch fabric into smaller fabrics with a Multi-Site extension is strongly recommended. We can create five individual VXLAN EVPN fabrics with 100 switches or ten fabrics with 50 switches and interconnect them to extend any VRF and network anywhere between these fabrics. This approach allows us to deploy horizontal scale-out architecture while maintaining the overall VTEP scale and other attributes. But, in such a design, we still need to decide on the north-to-south ingress/egress point, service nodes perimeter point, and more.

For example, we need to address the following:

- Where do we connect the DMZ/perimeter firewall?
- Where do we connect Internet/WAN links?
- How can we optimize the traffic paths and minimize the hair pinning?

In this design approach, we can place a Shared-Border plane centrally as a deterministic point for any Layer 3 north-to-south or service insertion use cases. The Shared-Border belongs to an independent fabric serving as a common entry and exit point for a given data center.

Flexible Hardware and Software requirements

Shared-Border is independent of any VXLAN EVPN Multi-Site software or hardware requirements; it is solely a border leaf node. The Shared-Border is also independent of a BGW (Border Gateway) from a functionality and licensing point of view. The minimum licensing requirement for Shared-Border is Network Essentials.

Flexible IP Handoff options

Shared-Border can terminate and handoff VXLAN EVPN traffic to external networks using VRF Lite (VXLAN to Native IP/IPv6 and vice versa) or MPLS VPN (VXLAN to MPLS-LDP/MPLS-SR and vice versa). Hence, Shared-Border can be utilized in a two-box or a one-box handoff solution.

Note: The support for VPN handoff is dependent on the specific hardware and software versions.

Service Node Insertion and Redirection

Shared-Border can be implemented as a set of standalone VTEPs or as a pair of VTEPs that are part of a vPC domain (vPC with Peer-Link or vPC Fabric-Peering). Hence, it simplifies the interconnection with Layer 4 to Layer 7 service nodes. Typically, the Shared-Border operates in Layer 3 mode. But if there are specific DMZ use cases, such as applications having their default gateway on a firewall cluster and the cluster is connected to the Shared-Border, we can extend the Layer 2 VNI across VXLAN EVPN Multi-Site and Shared-Border fabrics.

Centralized VRF Route-Leaking

The Shared-Border approach allows network admins to implement a centralized route-leaking option to simplify configurations, operations, troubleshooting, security domains, and more. The individual VXLAN EVPN fabrics rely on the Shared-Border as the inter-VRF leaking point.

Shared-Border Design

Availability Zones and Regions

When describing data center deployment architectures, a geographical location is often referred to as a "site." At the same time, the term "site" may also refer to a specific VXLAN EVPN fabric part of a Multi-Site architecture, and this may lead to confusion because multiple fabrics may be deployed in a given "site" geographical location. Hence, it is helpful to introduce terms like "Availability Zone" and "Region" to differentiate deployment scenarios.

An Availability Zone (AZ) refers to a set of network components representing a specific infrastructure fault domain. For VXLAN EVPN deployments, an AZ corresponds to a fabric part of a particular NDFC MSD construct. The geographic placement of AZs depends on the use case; for scaling-out network designs, for example, it is possible to deploy multiple AZs in the same physical (and geographic) data center location.

A Region is a collection of one or more AZs representing a single change and resource domain; a region typically includes AZs deployed in one or more geographic data center locations. In terms of a VXLAN EVPN deployment with NDFC, a Region represents a single fabric or multiple fabrics managed through a single NDFC controller (and hence part of the same NDFC MSD construct). So a controller's scope is that of managing all the data centers (or AZs) within the region.

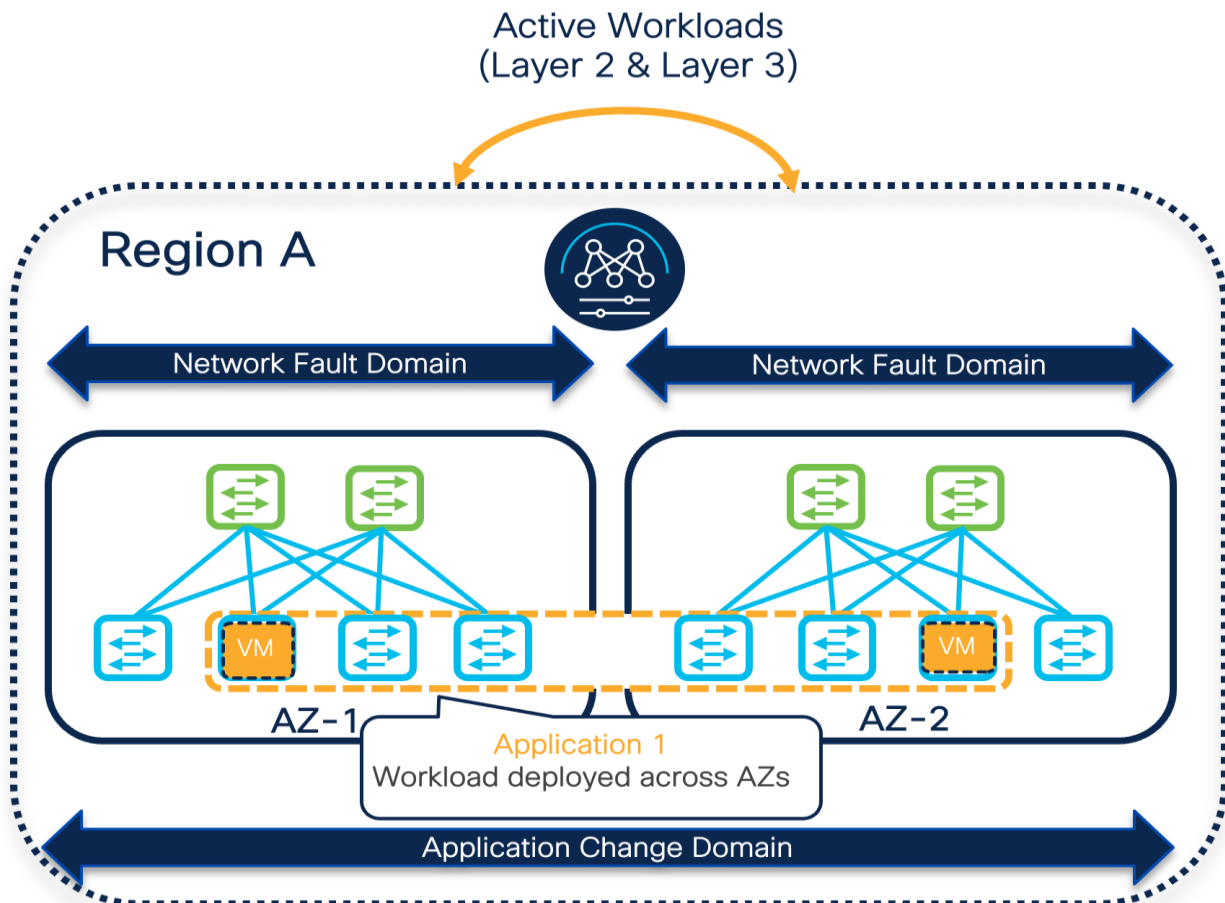


Figure 1. Availability Zones and Regions

Design Option 1

DCI- BGW to Cloud

The BGW to Cloud deployment model provides scalable design options within and across multiple sites. The Backbone/Cloud/IP-Core can be any routed service such as IP Layer 3 or MPLS-L3VPN network. The IP-Core is responsible for advertising and exchanging the loopback information between BGWs and the Shared-Border. In this approach, the BGWs in a given AZ peer full-mesh with the BGWs deployed in other AZs. The Shared-Border acts as an external VTEP and participates in EVPN overlay sessions with the BGWs. We must ensure that the Primary IP and Virtual IP (typically Lo0 for the EVPN control plane, and Lo1 and Lo100 for the VXLAN data plane) of all BGWs and Shared-Border are known to each other, and the MTU must accommodate VXLAN encapsulated traffic.

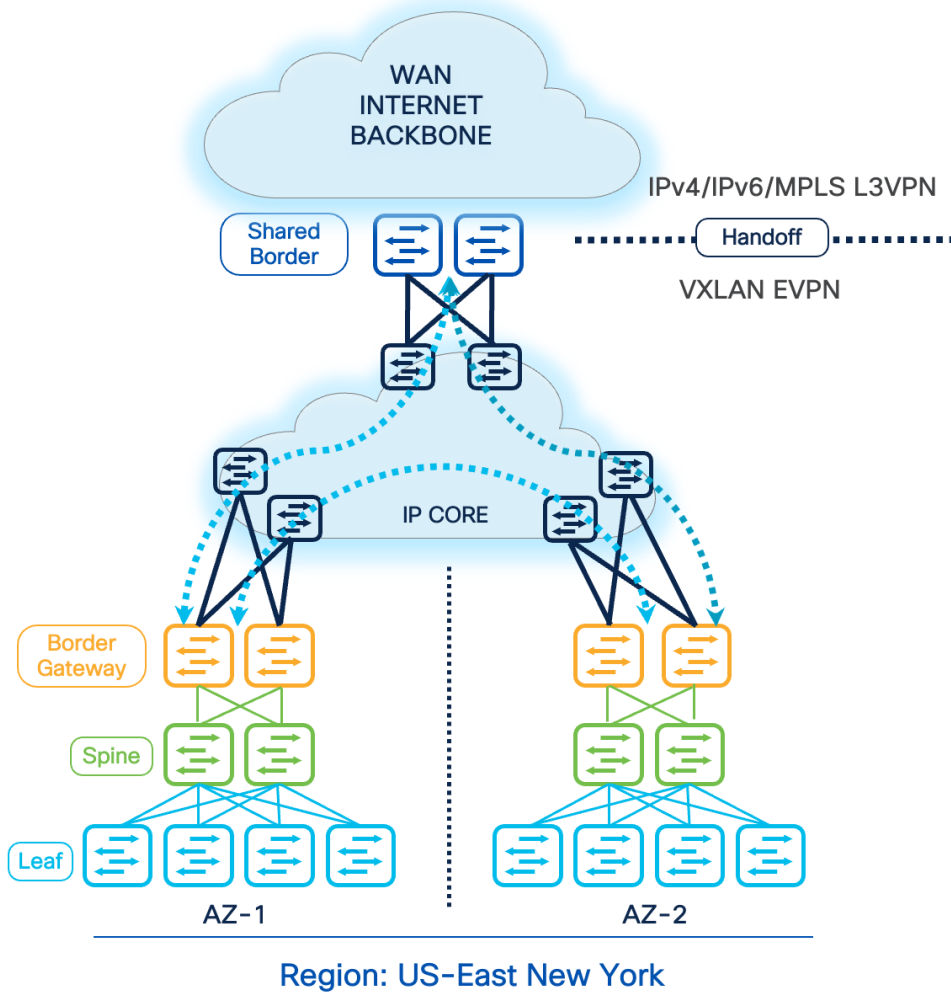


Figure 2. DCI- BGW to Cloud

Connectivity Key:

Multi-Site Underlay: eBGP IPv4 Unicast

- Site-External DCI BUM: Ingress-Replication or Multicast supported. At this time, Cisco NDFC supports only Ingress-Replication.
- Site-Internal Fabric BUM: Ingress-Replication or Multicast supported independently at each site.

-
- The eBGP IPv4 Unicast is used to exchange the IP reachability across BGWs and Shared-Border. Furthermore, if Shared-Border is running as a Layer 3 only VTEP, the BUM functionality and L2VNI definition can be skipped on the Shared-Border device.

Multi-Site Overlay: eBGP EVPN Overlay

- Full-Mesh BGP EVPN peering across all BGWs and Shared-Border.

Design Option 2

DCI- BGW Back-to-Back

Another option is to connect VXLAN EVPN AZs using the BGW Back-to-Back deployment model. In this approach, the BGWs and Shared-Border are directly connected. Hence, considering the cable availability, physical restrictions, geographic locations, and other dependencies, this model is limited and recommended for connecting a maximum of two sites. As a best practice design principle, connecting every BGW and Shared-Border is recommended. Still, due to certain restrictions, if this is not possible, the minimum topology for Back-to-Back is the square topology. The square connectivity mandates the deployment of a local Layer 3 connection between BGWs of a given site to ensure seamless and improved ECMP, BUM, data plane traffic, and failure scenarios.

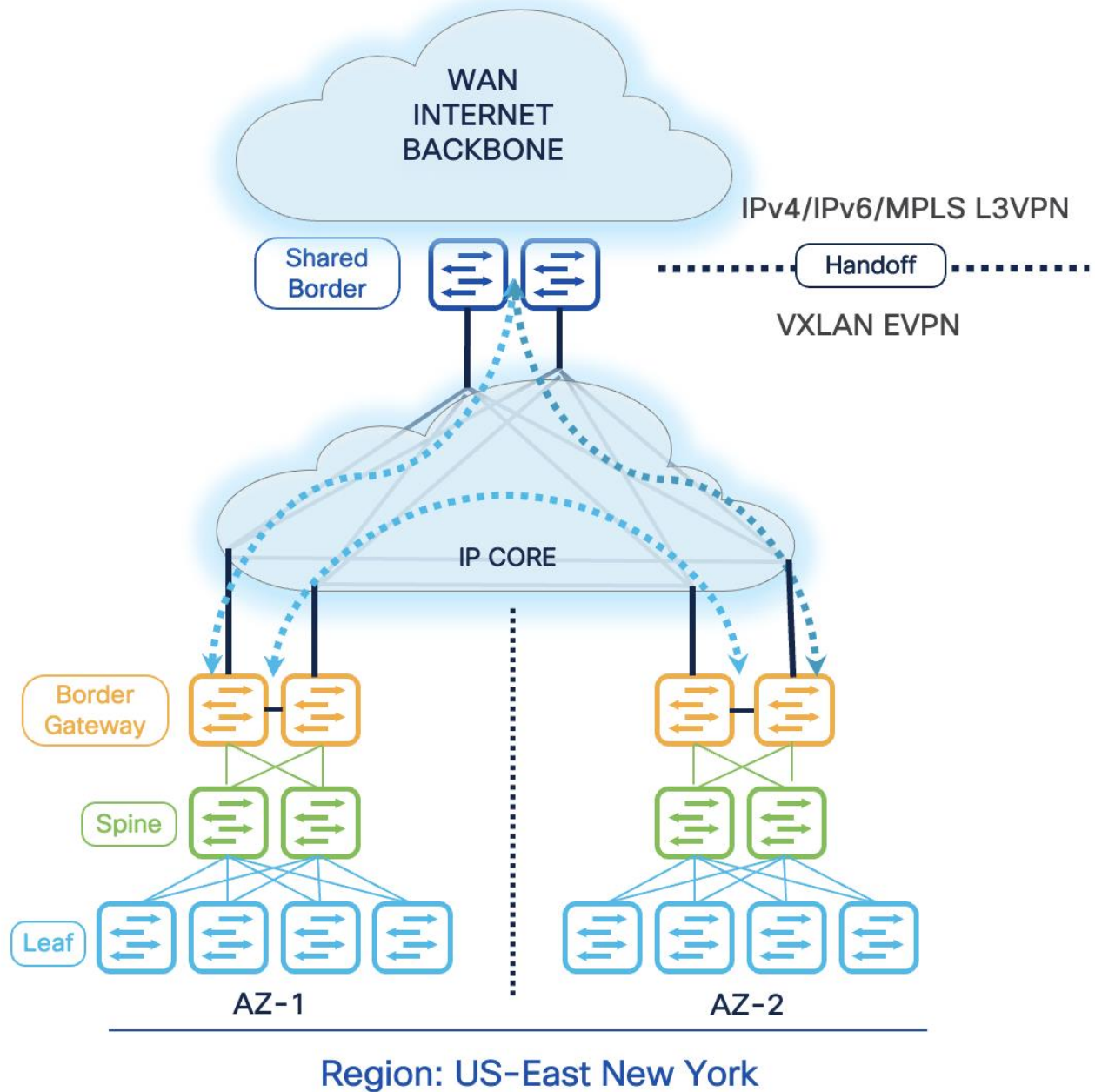


Figure 3. DCI- BGW Back-to-Back

Connectivity Key:

Multi-Site Underlay: eBGP IPv4 Unicast

- Site-External DCI BUM: Ingress-Replication or Multicast supported. At this time, Cisco NDFC supports only Ingress-Replication.
- Site-Internal Fabric BUM: Ingress-Replication or Multicast supported independently at each site.

-
- The eBGP IPv4 Unicast is used to exchange the IP reachability across BGWs and Shared-Border. Furthermore, if Shared-Border is running as a Layer 3 only VTEP, the BUM functionality and L2VNI definition can be skipped on the Shared-Border device.

Multi-Site Overlay: eBGP EVPN Overlay

- Full-Mesh BGP EVPN peering across all BGWs and Shared-Border.

Design Option 3

DCI- BGW to Centralized Route Server

The previous design options require us to implement a Full-Mesh configuration of EVPN sessions across all participating BGWs across all available sites. The EVPN Full-Mesh peering and adjacencies can significantly increase as we grow horizontally. The Full-Mesh option may introduce challenges from a physical cabling, configuration, management, and troubleshooting point of view. Therefore, for multiple VXLAN Sites, it is recommended that you leverage the BGP EVPN Route Server model. This model helps contain the overall connectivity, configurations, management, and more.

The Route Server model allows administrators to place a switch or router capable of running certain functionality and peer directly with the BGWs. It is essentially like a RR (Route Reflector) for eBGP EVPN sessions. Therefore, all the BGWs peer directly or indirectly with the Route Servers. The Route Server can be Nexus or Non-Nexus devices that comply with RFC 7947 and support EVPN AFI and BGP extensions, such as next-hop-unchanged, retain RTs, and RT rewrite functions.

Furthermore, the Route Server does not need to be on the data plane path. Therefore, we can place a set of devices acting as the Route Server in the backbone WAN and establish eBGP EVPN Multi-Hop peering with the BGWs. Another approach is to physically connect every BGW to the Route Server and establish the peering. Thus, depending on the overall physical and logical network connectivity, the Route Server may or may not be part of the data plane.

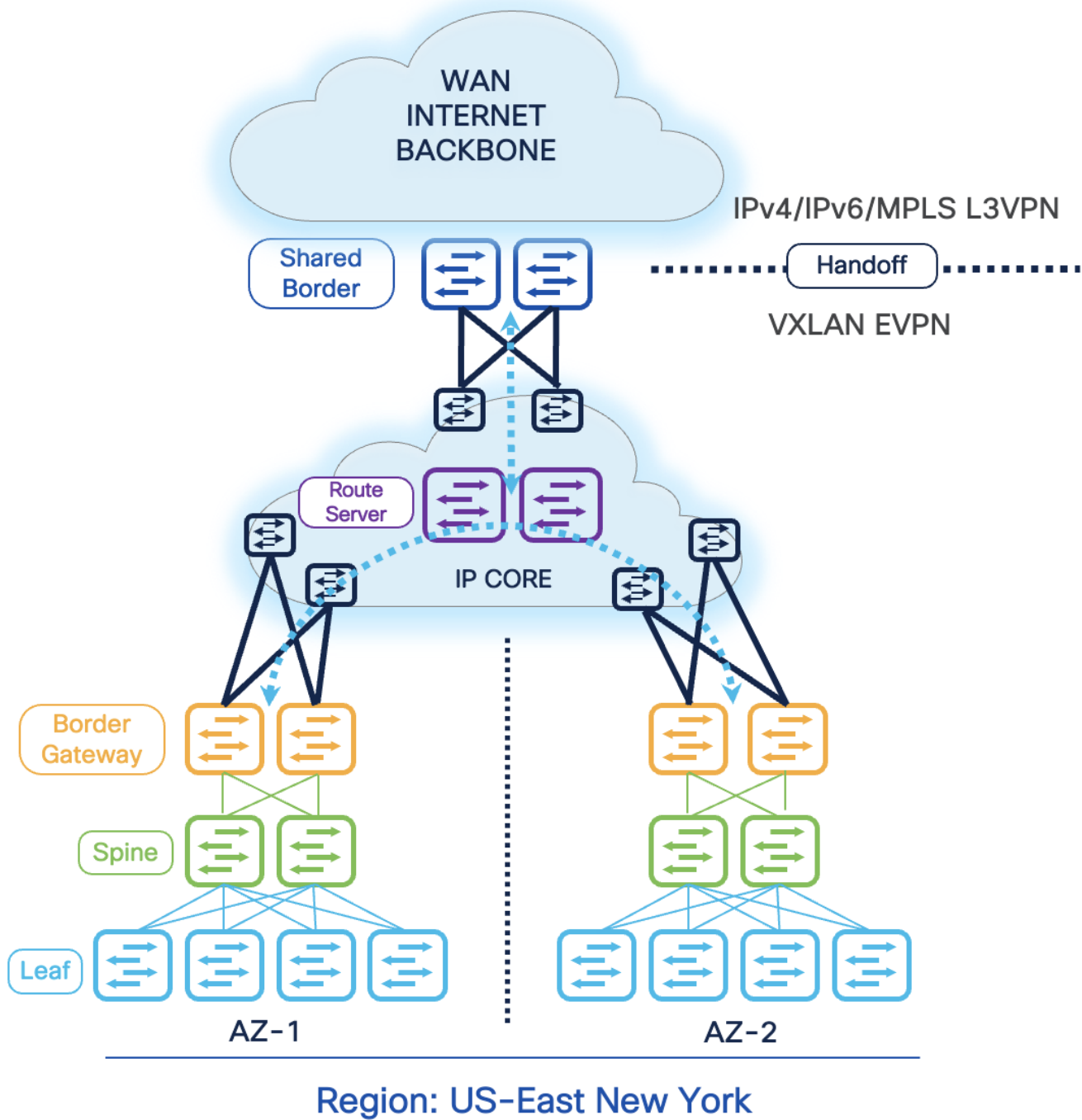


Figure 4. DCI- BGW to Centralized Route Server

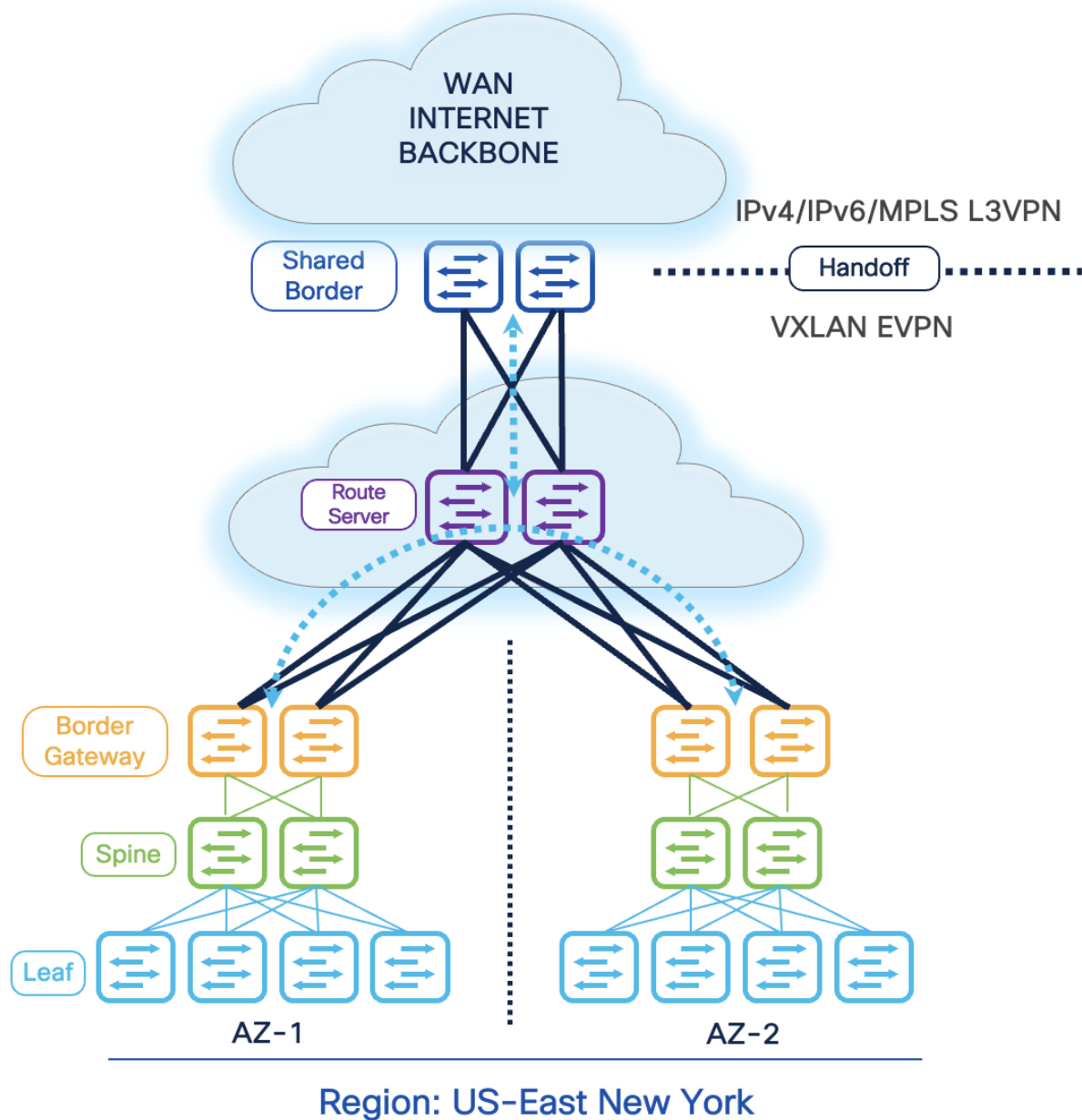


Figure 5. DCI- BGW to Centralized Route Server (Route Server in Data Path)

Connectivity Key:

Multi-Site Underlay: eBGP IPv4 Unicast

- Site-External DCI BUM: Ingress-Replication or Multicast supported. At this time, Cisco NDFC supports only Ingress-Replication.
- Site-Internal Fabric BUM: Ingress-Replication or Multicast supported independently at each site.
- The eBGP IPv4 Unicast is used to exchange the IP reachability across BGWs and Shared-Border. Furthermore, if Shared-Border is running as a Layer 3 only VTEP, the BUM functionality and L2VNI definition can be skipped on the Shared-Border device.

Multi-Site Overlay: eBGP EVPN Overlay ←-----→

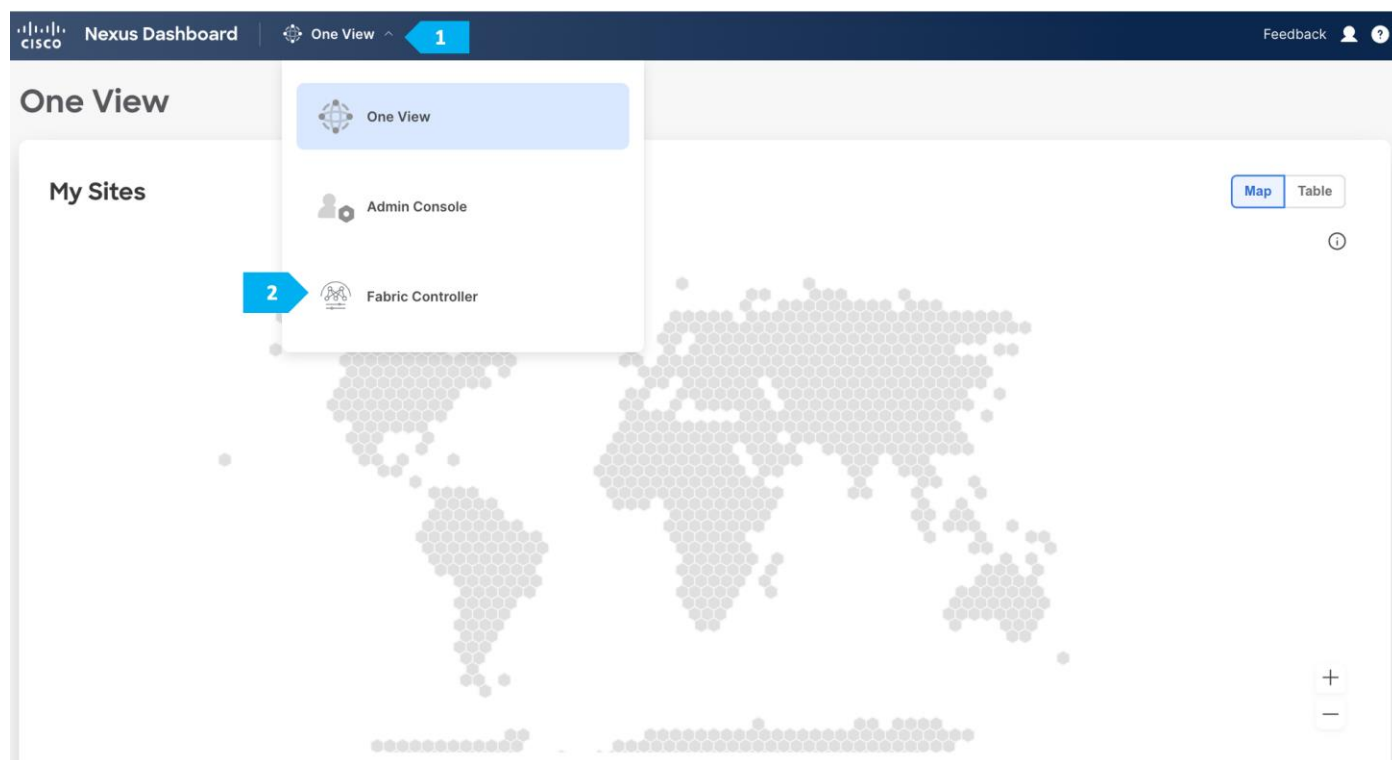
- BGP EVPN peering across all BGWs and Shared-Border via the Route Server.

Automation and Management

In the next steps we will start building the DCI-BGW to Centralized Route Server topology using NDFC. We will build the following network components.

- AZ1-New-York
- AZ2-New-York
- Backbone
- Shared border
- New-York Multi-Site Domain (MSD)

To create all the fabrics above please login to the ND cluster and choose **Fabric Controller**.



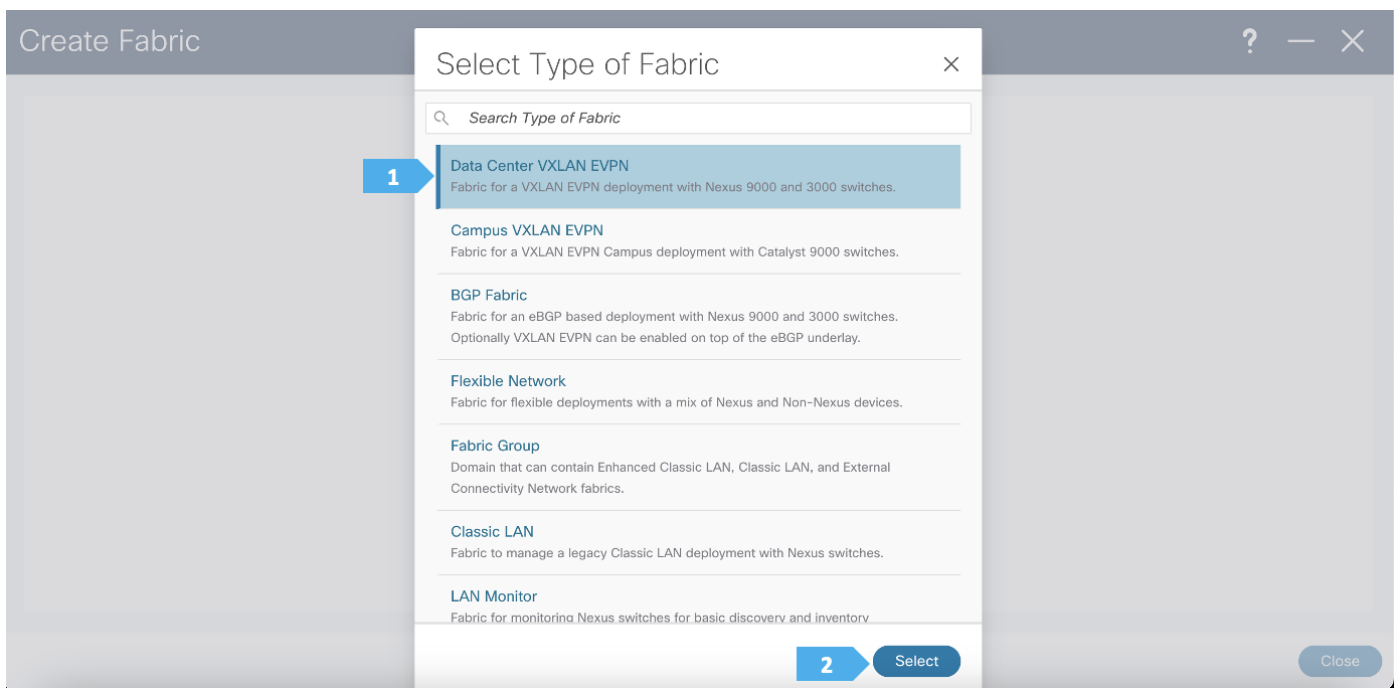
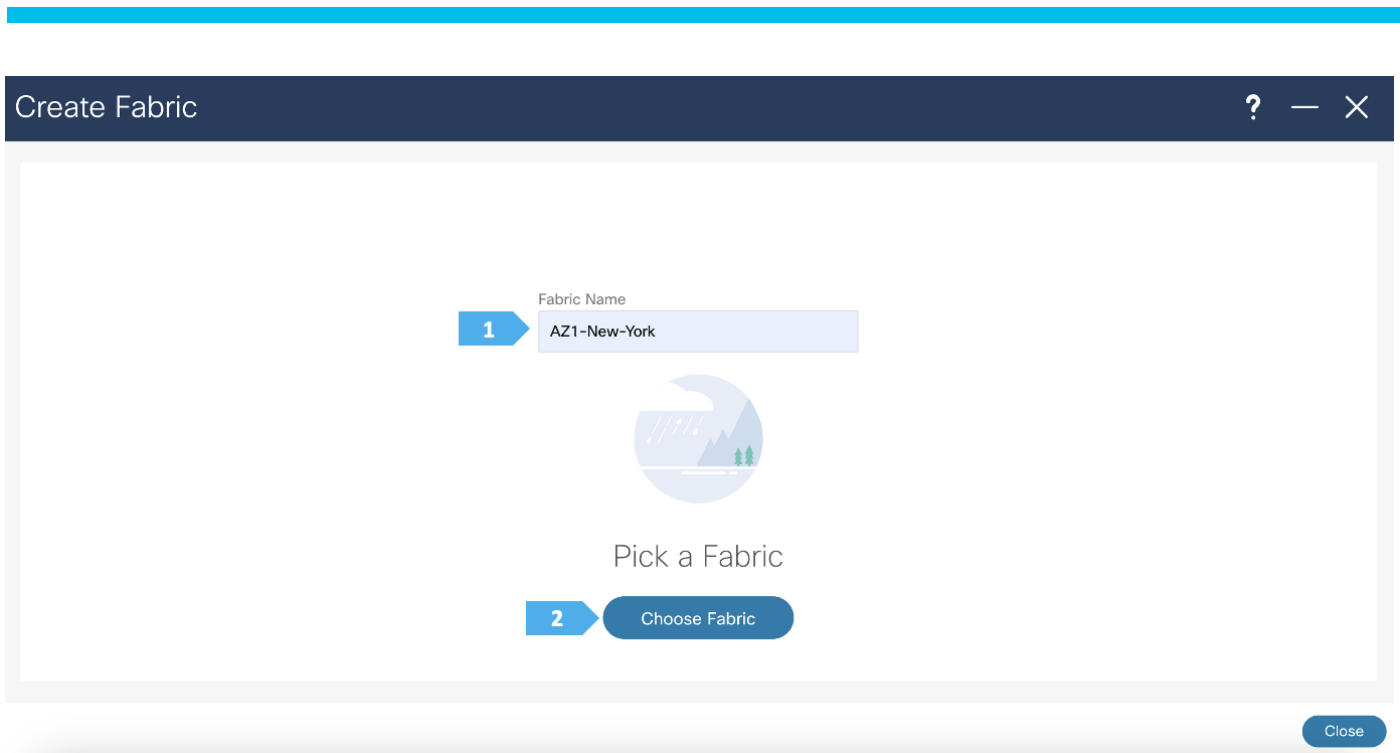
Creating AZ1-New-York Fabric

Step 1. Creating the fabric and choosing the template

The first fabric that we will be creating is AZ1-New-York, which will be a VXLAN EVPN fabric. It will contain Leaf-101 and Leaf-102 as leaf nodes. For this fabric to be part of VXLAN EVPN Multi-Site, it must have BGWs (Border Gateways) so that it can exchange network and endpoint reachability information using the MP-BGP EVPN overlay control plan with other fabrics. In this fabric, we will show how to use the BGW function using the BGW Spine role by using BGWS-201 and BGWS-202.

AZ1-New-York will use the Data Center VXLAN EVPN fabric template, which is a fabric for a VXLAN EVPN deployment with Nexus 9000 and 3000 switches.

The screenshot displays the Cisco Fabric Controller web interface. The top navigation bar includes the Cisco logo, 'Nexus Dashboard', 'Fabric Controller', and 'Feedback' with user and help icons. The main interface is titled 'Fabric Controller' and features a left-hand navigation menu with options: Dashboard, Topology (highlighted with a blue arrow and the number '1'), LAN, Virtual Management, Settings, and Operations. The main content area is divided into a 'View' panel on the left and a search area on the right. The 'View' panel includes a search bar, a 'Search by Attributes' field, and a 'View' dropdown menu with options for 'Operation' and 'Configuration'. Below this is a 'Custom Saved' dropdown and a legend for health status: Healthy (green), Warning (blue), Minor (yellow), Major (orange), Critical (red), and NA (grey). The search area on the right has a search bar and an 'Actions' button (highlighted with a blue arrow and the number '2'). A dropdown menu for 'Actions' is open, showing 'Add Fabric' (highlighted with a blue arrow and the number '3') and 'Resync vCenters'.



After clicking select we will be presented with a screen with multiple tabs. The overlay and underlay network parameters are included in these tabs.

Please note that the parameters displayed are the minimum to get the fabric up and running and to make it part of a multi-site setup. Please refer to the following link and choose the configuration guide based on the software version being used to understand what each parameter does and to modify the settings based on the specifics of your deployment:

<https://www.cisco.com/c/en/us/support/cloud-systems-management/prime-data-center-network-manager/products-installation-and-configuration-guides-list.html>

Filling in the parameters in the “General Parameters” tab

In this tab, we will be filling in only the BGP ASN field. Enter the BGP AS number that the fabric is associated with. In this example, will be using 65001 as the BGP ASN.

The screenshot shows the configuration page for fabric AZ1-New-York. The 'General Parameters' tab is active. The BGP ASN field is highlighted with a blue box and the number 2. The field contains the value 65001. Other fields include Fabric Name (AZ1-New-York), Enable IPv6 Underlay, Enable IPv6 Link-Local Address, Fabric Interface Numbering (p2p), Underlay Subnet IP Mask (30), Underlay Subnet IPv6 Mask (Select an Option), and Underlay Routing Protocol (ospf). The page has a 'Close' button and a 'Save' button at the bottom right.

Filling in the parameters in the “Replication” tab

Replication Mode: The mode of replication that is used in the fabric for BUM (Broadcast, Unknown Unicast, Multicast) traffic. The choices are Ingress Replication or Multicast. We will be using the Multicast replication mode.

Create Fabric ? — ✕

Fabric Name
AZ1-New-York

Pick Fabric
Data Center VXLAN EVPN >

General Pa **1** Replication VPC Protocols Advanced Resources Manageability Bootstrap Configuration Backup Flow Monitor

2 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*

Close Save

Filling in the parameters in the “vPC” tab

In the **AZ1-New-York** fabric, we will be using fabric vPC peering which provides an enhanced dual-homing access solution without the overhead of wasting physical ports for vPC Peer Link. This feature preserves all the characteristics of a traditional vPC. We will use all defaults and select only “**Enable QoS for Fabric vPC-Peering**” to enable QoS on spine switches for guaranteed delivery of fabric vPC peering communication. Please refer to the appropriate configuration guide for guidelines on using QoS for fabric vPC peering.

We can see that all the parameters are automatically populated by NDFC.

Create Fabric



Fabric Name
AZ1-New-York

Pick Fabric
Data Center VXLAN EVPN >

General Parameters Resources **1** VPC Protocols Advanced Resources Manageability Bootstrap Configuration Backup Flow Monitor

vPC Peer Link VLAN*
3600 VLAN for vPC Peer Link SVI (Min:2, Max:4094)

Make vPC Peer Link VLAN as Native VLAN

vPC Peer Keep Alive option*
management Use vPC Peer Keep Alive with Loopback or Management

vPC Auto Recovery Time (In Seconds)*
360 (Min:240, Max:3600)

vPC Delay Restore Time (In Seconds)*

Close Save

Create Fabric



vPC Peer Link Port Channel ID
500 (Min:1, Max:4096)

vPC IPv6 ND Synchronize
 Enable IPv6 ND synchronization between vPC peers

vPC advertise-pip
 For Primary VTEP IP Advertisement As Next-Hop Of Prefix Routes

Enable the same vPC Domain Id for all vPC Pairs
 (Not Recommended)

vPC Domain Id
 vPC Domain Id to be used on all vPC pairs

vPC Domain Id Range
1-1000 vPC Domain Id range to use for new pairings

1 Enable Qos for Fabric vPC-Peering
 Qos on spines for guaranteed delivery of vPC Fabric Peering communication

Qos Policy Name*
spine_qos_for_fabric_vpc_peering Qos Policy name should be same on all spines

Close Save

Filling in the parameters in the “Protocols” tab

The Protocol tab is mostly for the parameters used in the underlay. Most of the parameters are automatically generated. For the purpose of this setup, we will leave everything as default.

Fabric Name

AZ1-New-York

Pick Fabric

Data Center VXLAN EVPN >

General Parameters Replication **1** Protocols Advanced Resources Manageability Bootstrap Configuration Backup Flow Monitor

Underlay Routing Loopback Id*

0

(Min:0, Max:1023)

Underlay VTEP Loopback Id*

1

(Min:0, Max:1023)

Underlay Anycast Loopback Id

Used for vPC Peering in VXLANv6 Fabrics (Min:0, Max:1023)

Underlay Routing Protocol Tag*

UNDERLAY

Underlay Routing Process Tag

Close

Save

Filling in the parameters in the “Advanced” tab

In the Advanced tab, everything is automatically populated. We will only change the Overlay mode parameters.

Overlay Mode: We can create a VRF or network in CLI or config-profile mode at the fabric level. For the purpose of this setup, we will be using CLI.

Note: Starting with NDFC release 12.1.3b, the default Overlay option for new deployments of the Data Center VXLAN EVPN fabric type is “CLI”.

Fabric Name

AZ1-New-York

Pick Fabric

Data Center VXLAN EVPN >

General Parameters Replication VPC **1** Advanced Resources Manageability Bootstrap Configuration Backup Flow Monitor

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

2

Close

Save

Filling in the parameters in the “Resources” tab

By default, Nexus Dashboard Fabric Controller allocates the underlay IP address resources (for loopbacks, fabric interfaces, and so on) dynamically from the defined pools. It's good practice to enter unique values for the Underlay Routing Loopback IP Range and Overlay VTEP Loopback IP Range fields to proactively avoid duplicate IDs across individual fabrics once we connect them through multi-site.

Edit Fabric : AZ1-New-York

Fabric Name
AZ1-New-York

Pick Fabric
Data Center VXLAN EVPN >

General Parameters Replication VPC Protocols **1** Resources Manageability Bootstrap Configuration Backup Flow Monitor

Manual Underlay IP Address Allocation
 Checking this will disable Dynamic Underlay IP Address Allocations

2 Underlay Routing Loopback IP Range*
10.11.0.0/22 Typically Loopback0 IP Address Range

3 Underlay VTEP Loopback IP Range*
10.12.0.0/22 Typically Loopback1 IP Address Range

4 Underlay RP Loopback IP Range*
10.254.10.0/24 Anycast or Phantom RP IP Address Range

5 Underlay Subnet IP Range*
10.13.0.0/16 Address range to assign Numbered and Peer Link SVI IPs

Underlay MPLS Loopback IP Range
Used for VXLAN to MPLS SR/LDP Handoff

Underlay Routing Loopback IPv6 Range
Typically Loopback0 IPv6 Address Range

Underlay VTEP Loopback IPv6 Range

Close Save

<input type="checkbox"/>	Auto Deploy Default VRF	Whether to auto generate Default VRF interface and BGP peering configuration on VRF LITE IFC auto deployment. If set, auto created VRF Lite IFC links will have 'Auto Deploy Default VRF' enabled.
<input type="checkbox"/>	Auto Deploy Default VRF for Peer	Whether to auto generate Default VRF interface and BGP peering configuration on managed neighbor devices. If set, auto created VRF Lite IFC links will have 'Auto Deploy Default VRF for Peer' enabled.
	Redistribute BGP Route-map Name	Route Map used to redistribute BGP routes to IGP in default vrf in auto created VRF Lite IFC links
1	VRF Lite Subnet IP Range*	Address range to assign P2P Interfabric Connections
	VRF Lite Subnet Mask*	(Min:8, Max:31)
	Service Network VLAN Range*	Per Switch Overlay Service Network VLAN Range (Min:2, Max:4094)
	Route Map Sequence Number Range*	(Min:1, Max:65534)

[Close](#) [Save](#)

Filling in the parameters in the “Manageability”, “Bootstrap”, “Configuration Backup” and “Flow Monitor” tabs

We will use the defaults for all these tabs so all what we need to do is to click Save.

Step 2. Adding switches to the AZ1-New-York Fabric

The screenshot shows the Cisco Nexus Dashboard Fabric Controller interface. The main content area displays the configuration for the 'AZ1-New-York' fabric. A large green circle with a white cloud icon and the number '1' is overlaid on the page. A context menu is open over this icon, listing several actions: 'AZ1-New-York', 'Detailed View', 'Edit Fabric', 'Add Switches' (which is highlighted in blue), 'Recalculate and Deploy', and 'More'. The left sidebar contains a navigation menu with items: Dashboard, Topology, LAN, Virtual Management, Settings, and Operations. The top navigation bar shows 'Nexus Dashboard Fabric Controller' and a 'Feedback' button.

Use seed IP address to discover the switches. We will be using the admin user and password to discover switches. Uncheck preserve config to clear existing switch configurations and reload the devices. Max hop count allows the discovery of connected switches by the number of hops.

Add Switches - Fabric: AZ1-New-York

Switch Addition Mechanism*
 Discover

Seed Switch Details

Seed IP*
100.64.254.101
Ex: "2.2.2.20" or "10.10.10.40-60" or "2.2.2.20, 2.2.2.21"

Authentication Protocol*
MD5

Username*
admin

Password*
.....

Max Hops*
2

Preserve Config

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

5 Discover Switches

Add Switches - Fabric: AZ1-New-York

Warning

All switch configuration other than management, will be removed immediately after import. Do you want to proceed?

Cancel Confirm 1

Switch Addition Mechanism*
 Discover

Seed Switch Details

Seed IP*
100.64.254.101
Ex: "2.2.2.20" or "10.10.10.40-60" or "2.2.2.20,

Authentication Protocol*
MD5

Username*
admin

Password*
.....

Max Hops*
2

Preserve Config

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

After the switches are discovered, add these switches as part of the AZ1-New-York fabric and click "Add Switches".

Add Switches - Fabric: AZ1-New-York



Seed Switch Details

Fabric: AZ1-New-York
 Switch: 100.64.254.101
 Authentication Protocol: MD5
 Username: admin
 Password: ● Set
 Max Hops: 2
 Preserve config: ● Disabled

[← Back](#)

Discovery Results

Filter by attributes

<input type="checkbox"/>	Switch Name	Serial Number	IP Address	Model	Version	Status	Progress
<input type="checkbox"/>	RS-10	9W9A4AM8HLH	100.64.254.10	N9K-C9300v	10.2(5)	● Manageable	
<input checked="" type="checkbox"/>	BGWS-201	9AOZRKA9IY1	100.64.254.201	N9K-C9300v	10.2(5)	● Manageable	
<input checked="" type="checkbox"/>	BGWS-202	9046ZFS3G8	100.64.254.202	N9K-C9300v	10.2(5)	● Manageable	
<input type="checkbox"/>	RS-11	9AB4MSSB0XQ	100.64.254.11	N9K-C9300v	10.2(5)	● Manageable	
<input checked="" type="checkbox"/>	Leaf-101	9ZEA13L749S	100.64.254.101	N9K-C9300v	10.2(5)	● Manageable	
<input checked="" type="checkbox"/>	Leaf-102	99KJ3DPI53G	100.64.254.102	N9K-C9300v	10.2(5)	● Manageable	

1

2

Add Switches

Please wait until the Progress for all switches being added is green, then click Close.

Add Switches - Fabric: AZ1-New-York



Fabric: AZ1-New-York
 Switch: 100.64.254.101
 Authentication Protocol: MD5
 Username: admin
 Password: ● Set
 Max Hops: 2
 Preserve config: ● Disabled

[← Back](#)

Discovery Results

Filter by attributes

<input type="checkbox"/>	Switch Name	Serial Number	IP Address	Model	Version	Status	Progress
<input type="checkbox"/>	RS-10	9W9A4AM8HLH	100.64.254.10	N9K-C9300v	10.2(5)	● Manageable	
<input type="checkbox"/>	BGWS-201	9AOZRKA9IY1	100.64.254.201	N9K-C9300v	10.2(5)	● Switch Added	<div style="width: 100%; height: 10px; background-color: green;"></div>
<input type="checkbox"/>	BGWS-202	9046ZFS3G8	100.64.254.202	N9K-C9300v	10.2(5)	● Switch Added	<div style="width: 100%; height: 10px; background-color: green;"></div>
<input type="checkbox"/>	RS-11	9AB4MSSB0XQ	100.64.254.11	N9K-C9300v	10.2(5)	● Manageable	
<input type="checkbox"/>	Leaf-101	9ZEA13L749S	100.64.254.101	N9K-C9300v	10.2(5)	● Switch Added	<div style="width: 100%; height: 10px; background-color: green;"></div>
<input type="checkbox"/>	Leaf-102	99KJ3DPI53G	100.64.254.102	N9K-C9300v	10.2(5)	● Switch Added	<div style="width: 100%; height: 10px; background-color: green;"></div>

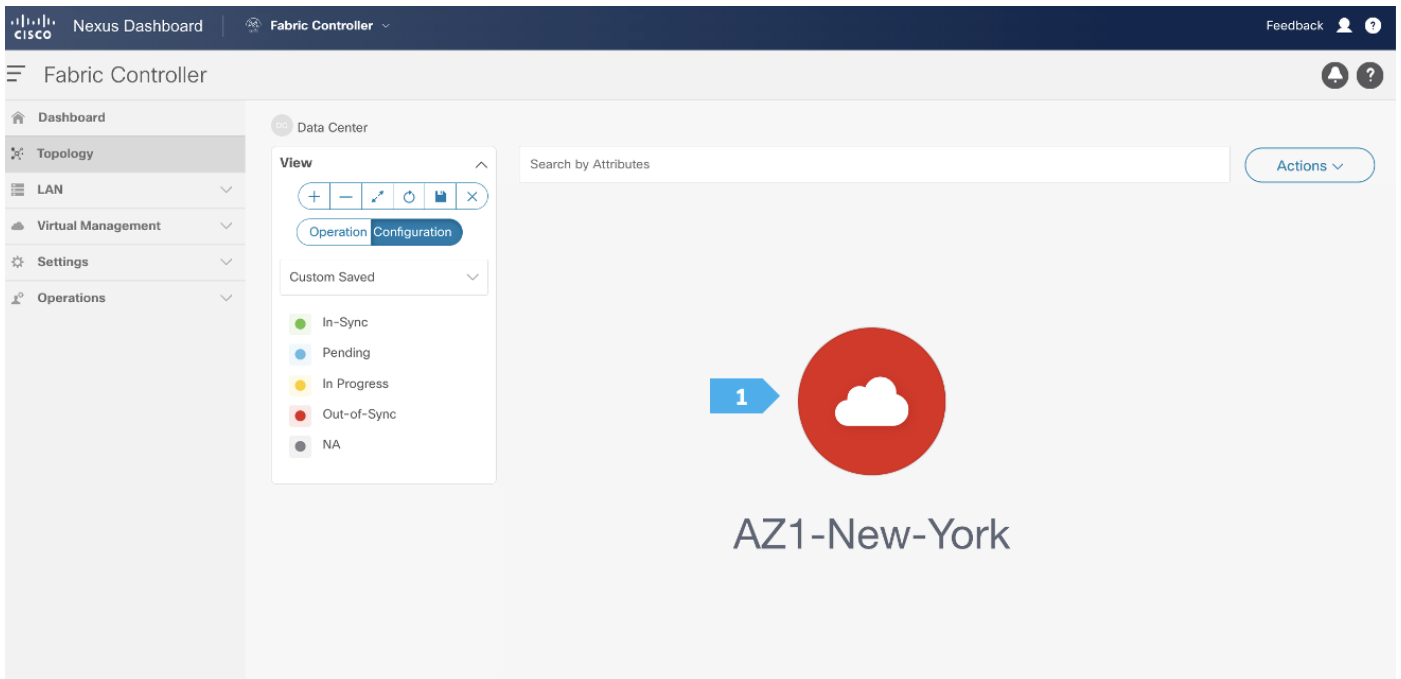
1

Close

Add Switches

Step 3. Changing the devices' roles

After the devices are added to the AZ1-New-York fabric, they will be assigned a default role depending on the platform. BGWS-210 and BGWS-202 will get the Border Gateway Spine role and Leaf-101 and Leaf-102 will get the Leaf roles, and the relevant configurations will be pushed to the respective devices. We can do these steps after we double-click on the AZ1-New-York fabric as shown in the next screen.



We see the fabric color is red, which means that it is out of sync because the intended configuration that we want is not yet pushed to the switches.

- Toggle the Multi-select option.
- Press Ctrl click and hold anywhere in the whitespace and drag the cursor up, down, left, or right to highlight the BGWS-201 and BGW-202.
- Release the modifier key “ctrl” before releasing the mouse drag to end the switch selection.
- Right-click and choose Set Role.

Nexus Dashboard | Fabric Controller | Feedback

Fabric Controller

Dashboard | Topology | LAN | Virtual Management | Settings | Operations

Data Center / AZ1-New-York

Search by Attributes

View

Show Logical Links

Operation Configuration

Custom Saved

- In-Sync
- Pending
- In Progress
- Out-of-Sync

Enable "Multi-select" then press ctrl-click and drag to select multiple nodes or shift-click to select individual nodes

Multi-select 0 selected

1

NET Networks (0) VRF VRFs (0)

Nexus Dashboard | Fabric Controller | Feedback

Fabric Controller

Dashboard | Topology | LAN | Virtual Management | Settings | Operations

Data Center / AZ1-New-York

Search by Attributes

View

Show Logical Links

Operation Configuration

Custom Saved

- In-Sync
- Pending
- In Progress
- Out-of-Sync
- NA

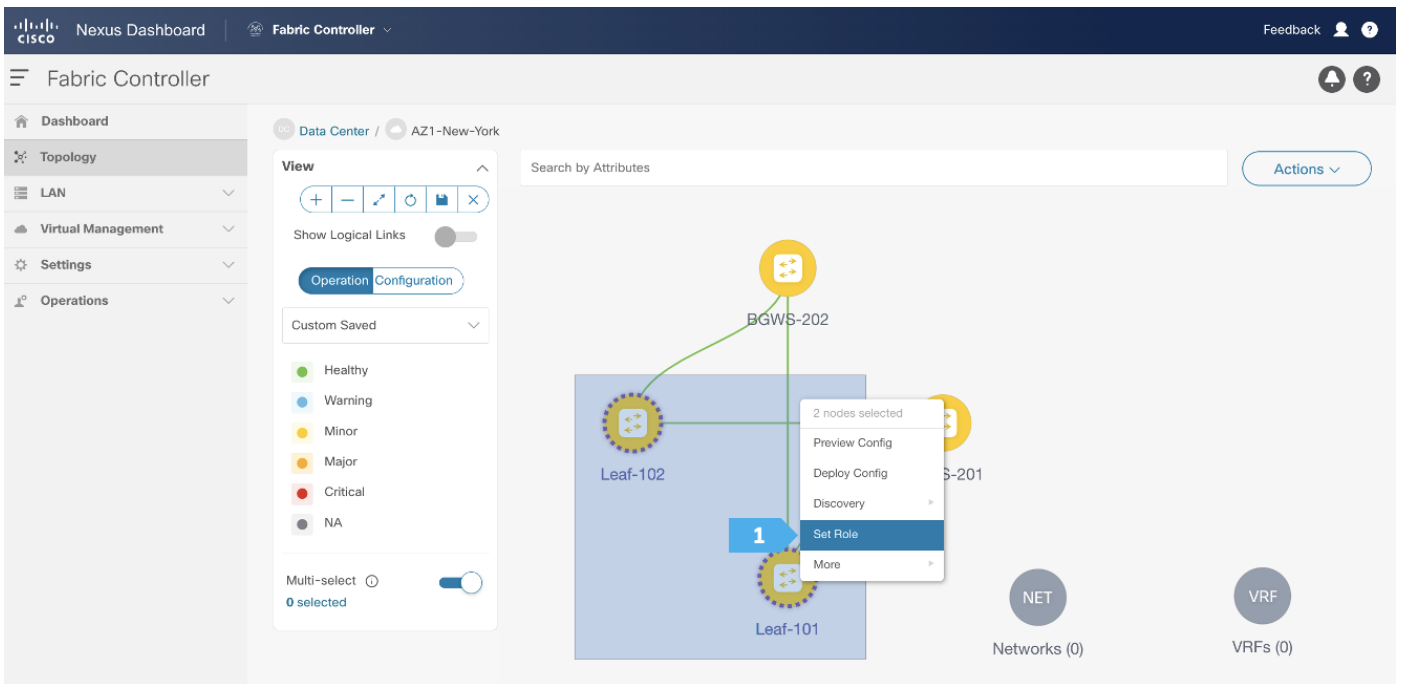
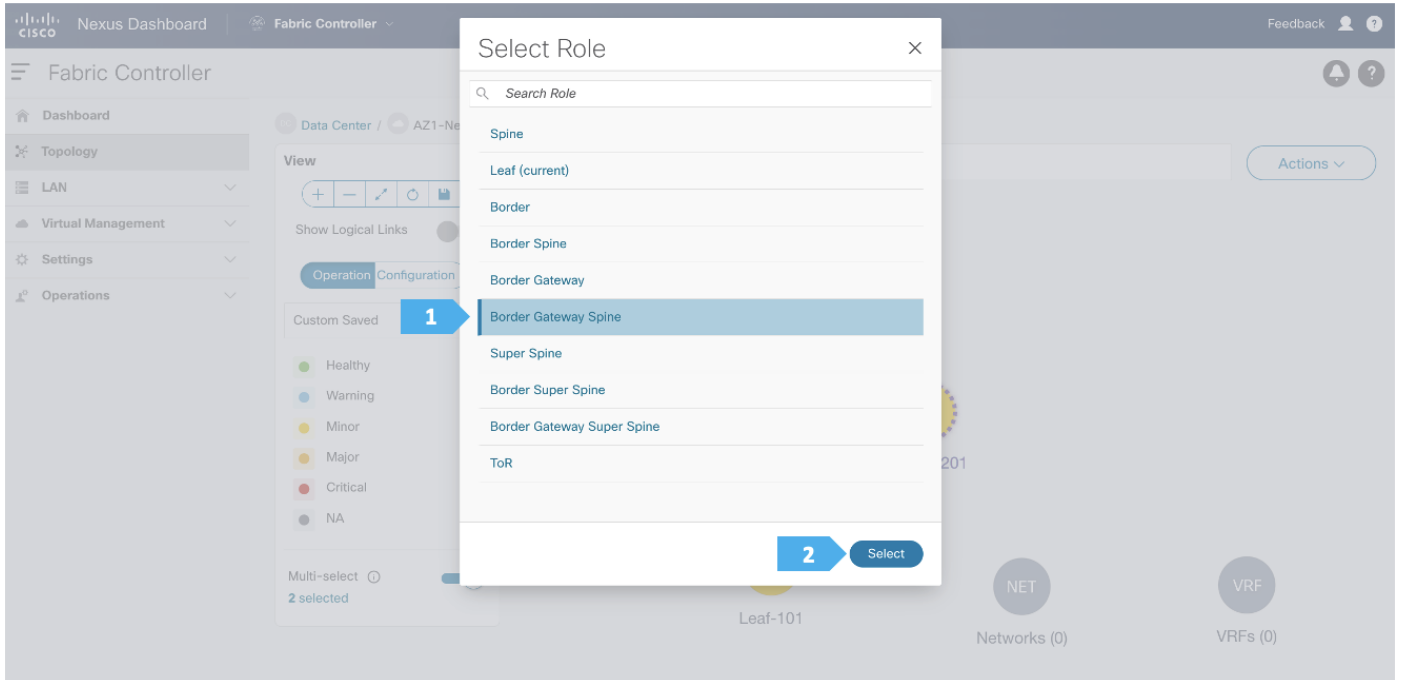
Multi-select 0 selected

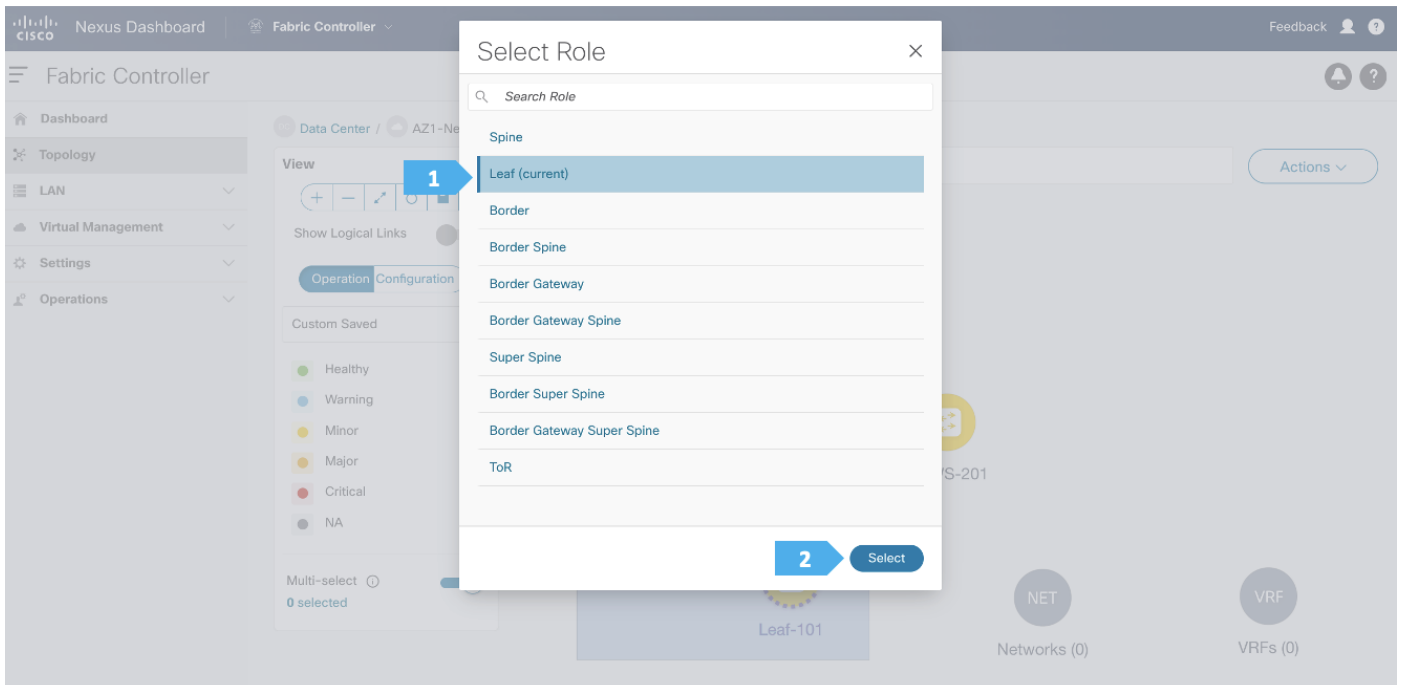
1

2 nodes selected

- Preview Config
- Deploy Config
- Discovery
- Set Role
- More

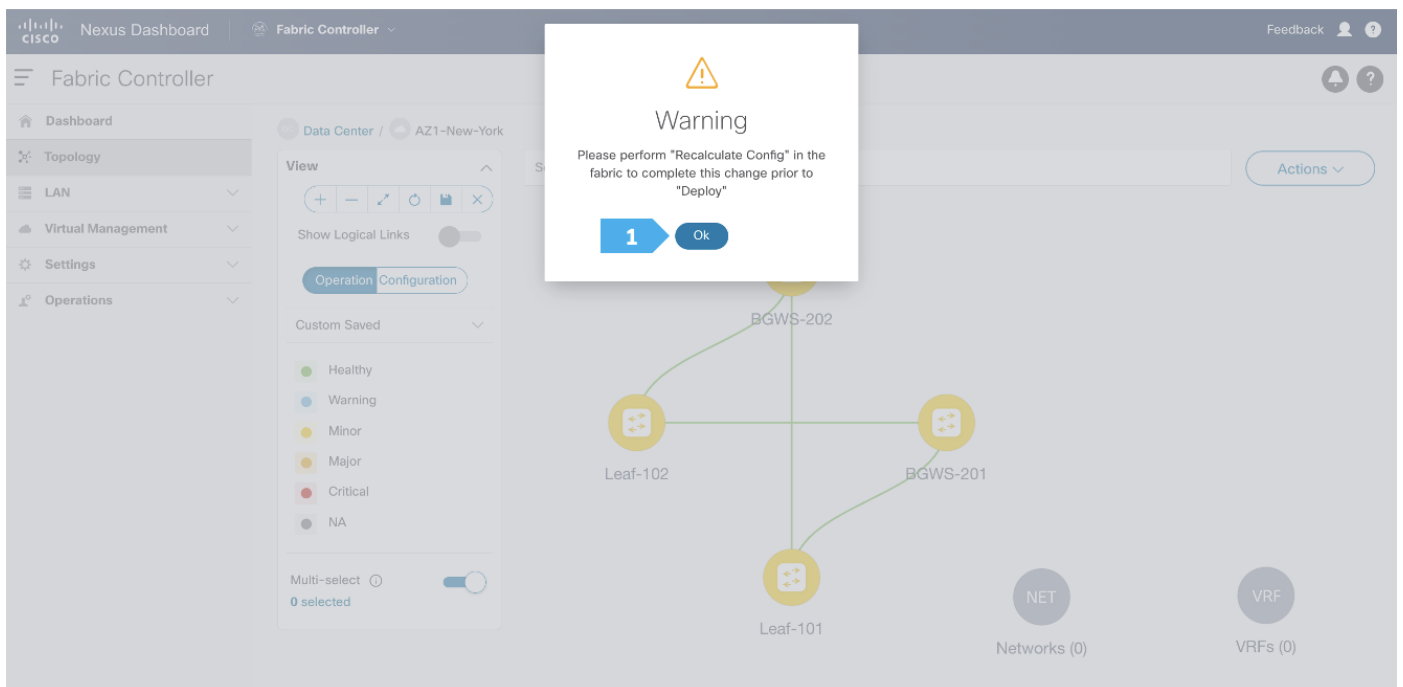
NET Networks (0) VRF VRFs (0)





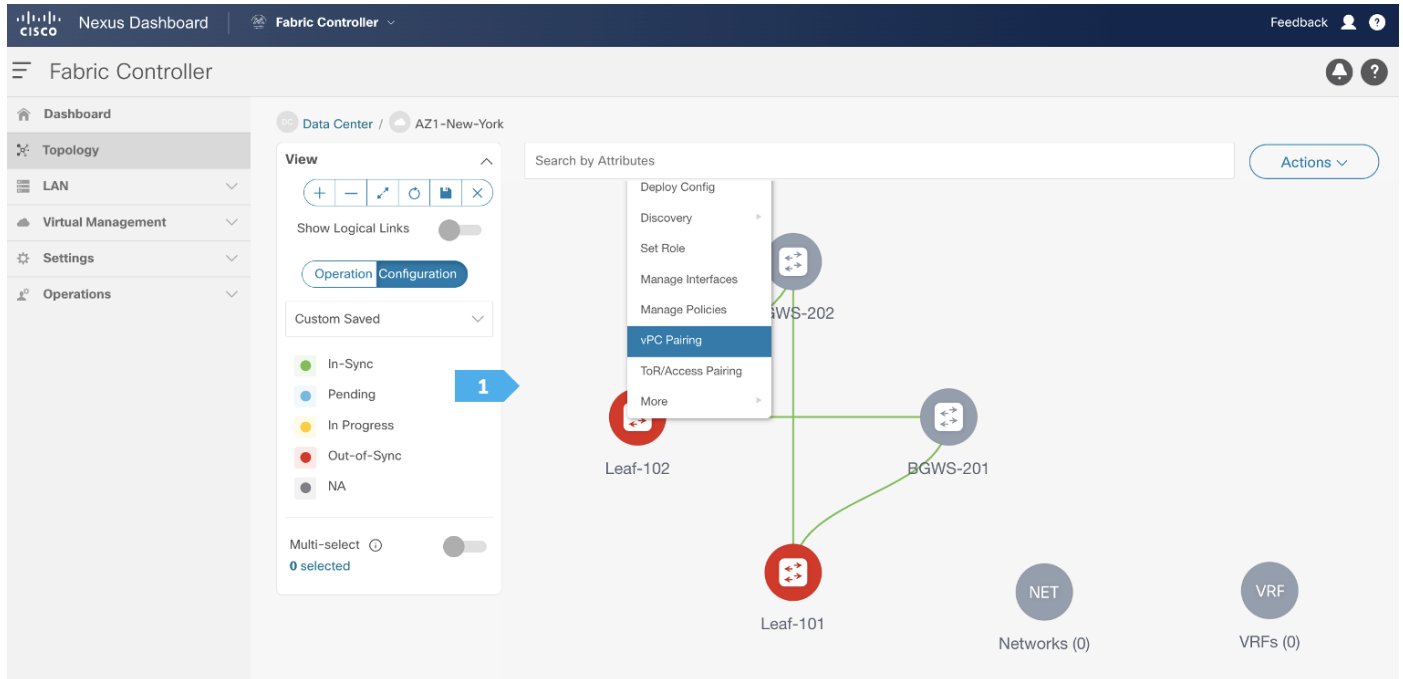
Click Ok in the warning window that appears.

The warning window tells us to perform a Recalculate and Deploy action; however, we will create additional configuration policies described in the next steps before performing the Recalculate and Deploy action.

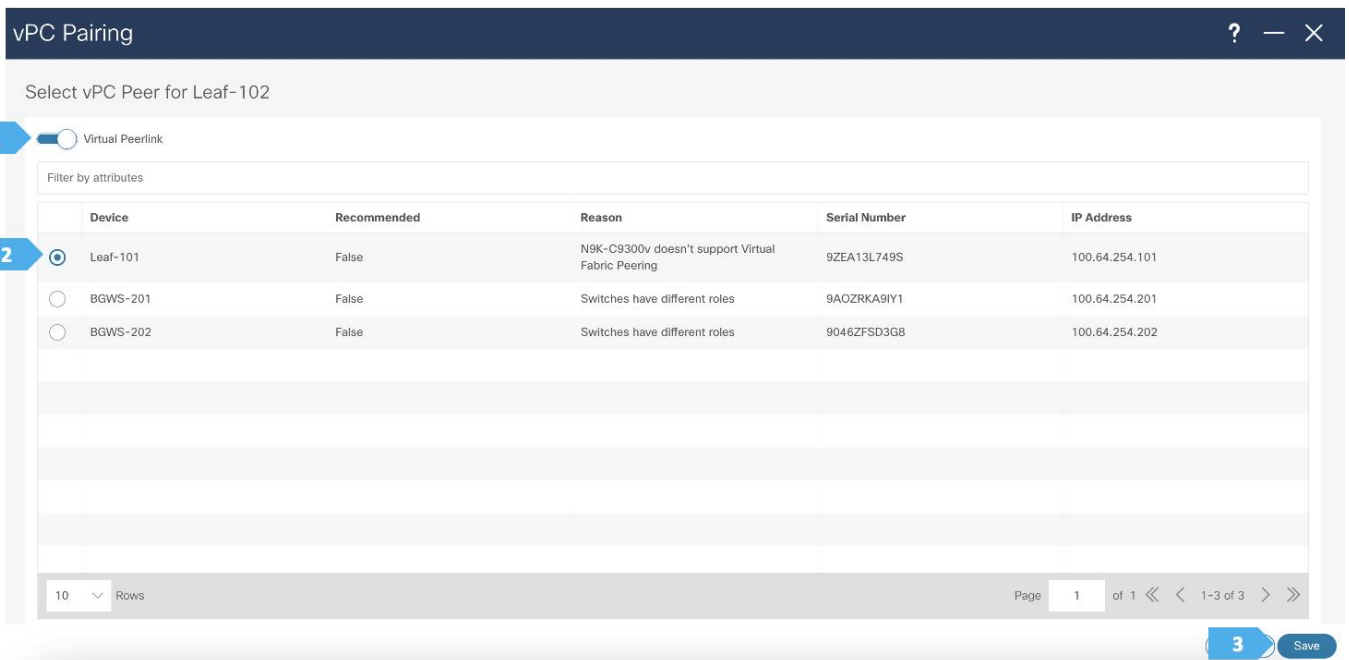


Step 4. Configuring vPC between leaf switches

To configure Leaf-101 and Leaf-102 as vPC Peers, right click on one of the leaf switches and select vPC Pairing.



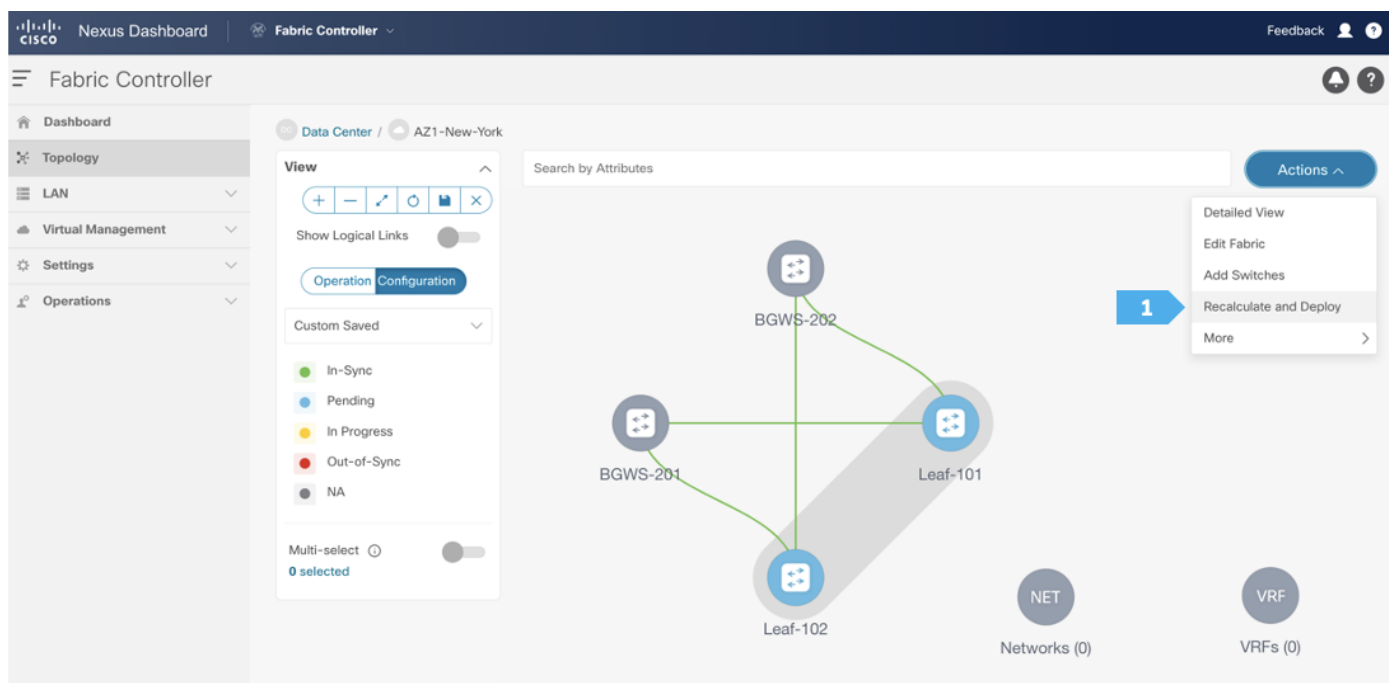
Select the peer switch to form vPC. In **AZ1-New-York** we don't have a direct link between the leaf switches, so fabric peering can be configured by selecting the "Virtual Peerlink".



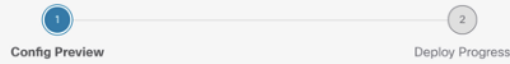
NDFC performs additional checks, such as whether vPC Fabric Peering is supported on the selected device and verifying the minimum NX-OS version and hardware requirement for the feature to be operational. Furthermore, NDFC recommends vPC pairing based on the overall requirement of the feature, thus saving operating time for network admins.

Step 5. Recalculating and deploying to the fabric

At this point, we are ready to push the configuration to the AZ1-New-York fabric. Choose “Recalculate and Deploy” as shown in the next screen.



We can click on the “Pending config” for each switch to view the configuration that will be provisioned before clicking “Deploy All”.



Filter by attributes

Resync All

Switch Name	IP Address	Role	Serial Number	Fabric Status	Pending Config	Status Description	Progress	Resync Switch
Leaf-102	100.64.254.102	leaf	99KJ3DPI53G	● Out-Of-Sync	538 Lines	Out-of-Sync	<div style="width: 100%;"></div>	Resync
Leaf-101	100.64.254.101	leaf	9ZEA13L749S	● Out-Of-Sync	538 Lines	Out-of-Sync	<div style="width: 100%;"></div>	Resync
BGWS-202	100.64.254.202	border gateway spine	9046ZFS3G8	● Out-Of-Sync	347 Lines	Out-of-Sync	<div style="width: 100%;"></div>	Resync
BGWS-201	100.64.254.201	border gateway spine	9AOZRKA9IY1	● Out-Of-Sync	347 Lines	Out-of-Sync	<div style="width: 100%;"></div>	Resync

1 Deploy All

Wait until the “Progress” for all the switches are green before clicking “Close”.



Filter by attributes

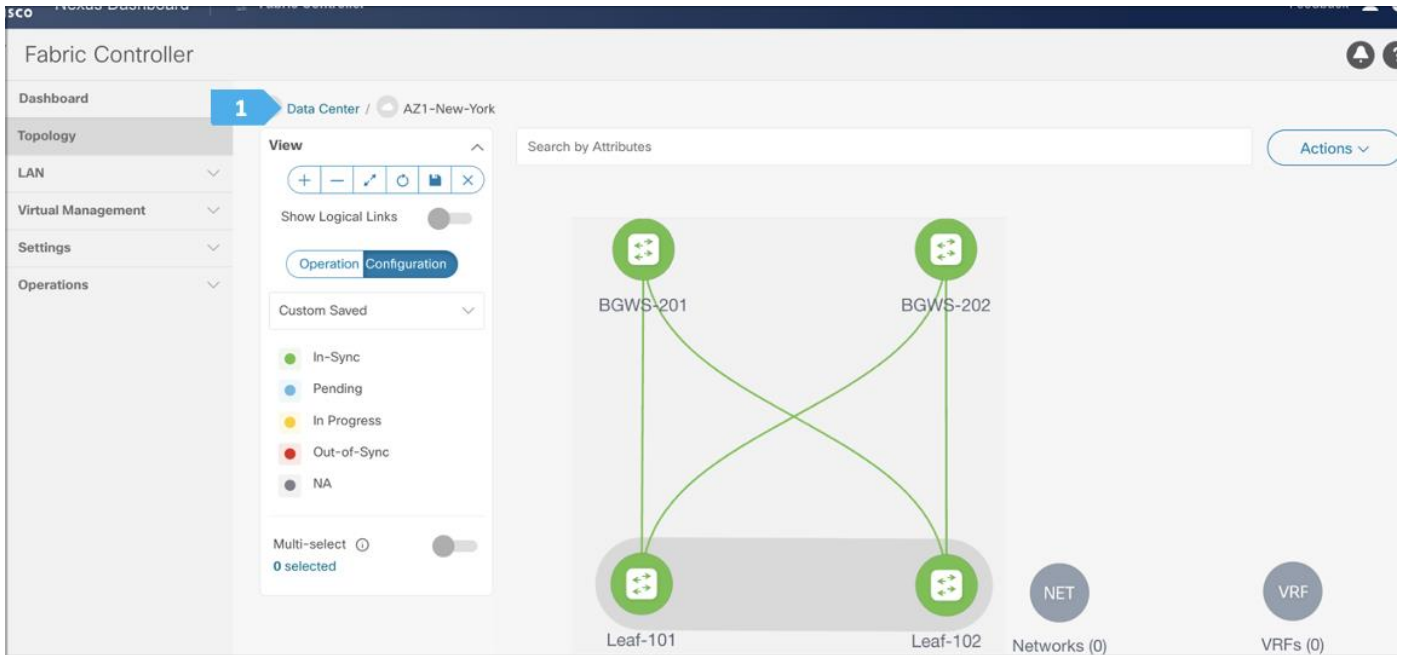
Switch Name	IP Address	Status	Status Description	Progress
Leaf-102	100.64.254.102	● SUCCESS	Deployment completed.	<div style="width: 100%;"><div style="width: 100%;"></div></div> Executed 538 / 538
Leaf-101	100.64.254.101	● SUCCESS	Deployment completed.	<div style="width: 100%;"><div style="width: 100%;"></div></div> Executed 538 / 538
BGWS-202	100.64.254.202	● SUCCESS	Deployment completed.	<div style="width: 100%;"><div style="width: 100%;"></div></div> Executed 347 / 347
BGWS-201	100.64.254.201	● SUCCESS	Deployment completed.	<div style="width: 100%;"><div style="width: 100%;"></div></div> Executed 347 / 347

1 Close

Fabric AZ1-New-York is deployed.

Now all the switches in AZ1-New-York fabric are green, meaning they are “In-Sync”.

Click on Data Center to go to the data center view.

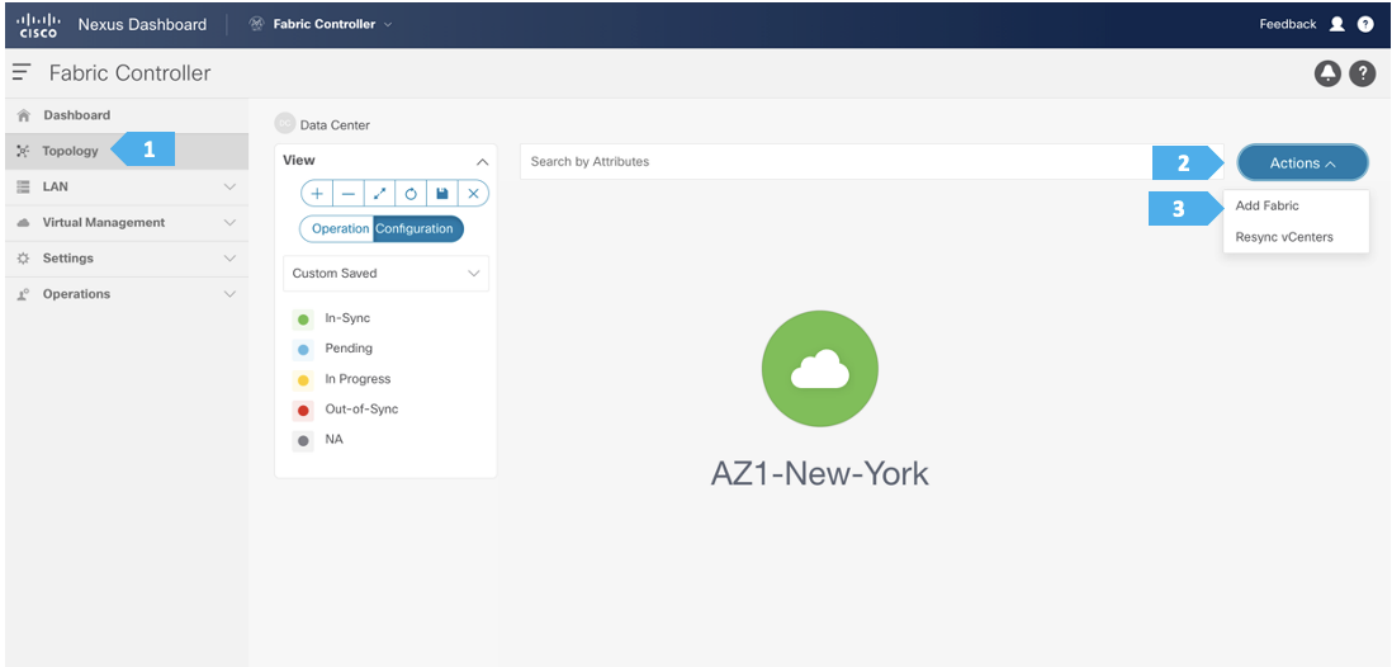


Creating AZ2-New-York Fabric

Creating the fabric and choosing the template

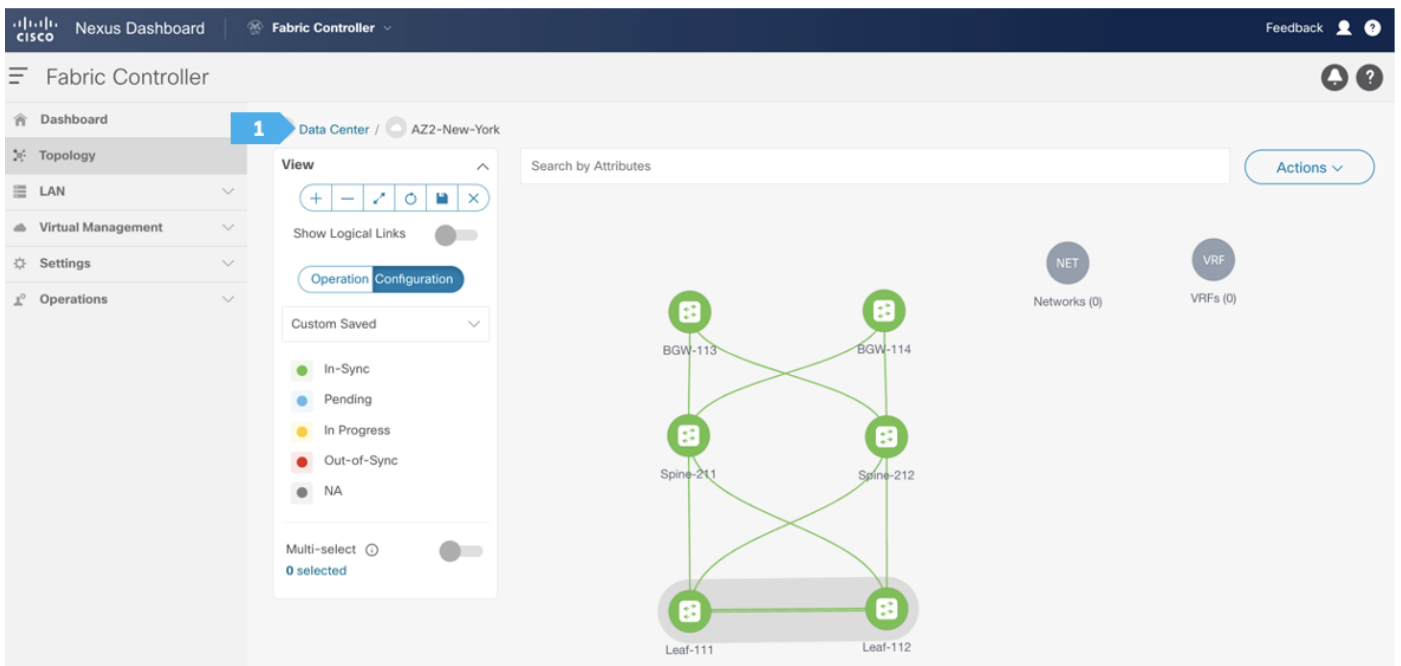
The second fabric that we will be creating is AZ2-New-York, which is a VXLAN EVPN fabri. It will contain Leaf-111 and Leaf-112 as leaf nodes, and Spine-211 and Spine-212 as spine nodes. For this fabric to be part of VXLAN EVPN Multi-Site, it must have BGWs (Border Gateways) so that it can exchange network and endpoint reachability information using the MP-BGP EVPN overlay control plan to other fabrics. In this fabric, we will show how to use the BGW function using dedicated BGW nodes BGW-113 and BGW-114.

AZ2-New-York will also use the Data Center VXLAN EVPN fabric template, which is the option for a VXLAN EVPN deployment with Nexus 9000 and 3000 switches. Repeat all the steps done for AZ1-New-York, where you make sure to choose unique BGP AS number, IP subnet, etc.



Fabric AZ2-New-York is deployed.

After finishing all the steps, the switches in the AZ2-New-York fabric should become green, meaning they are “In-Sync”, and we should have a Topology such as the screen below. Click on “Data Center” to go back to main Topology.

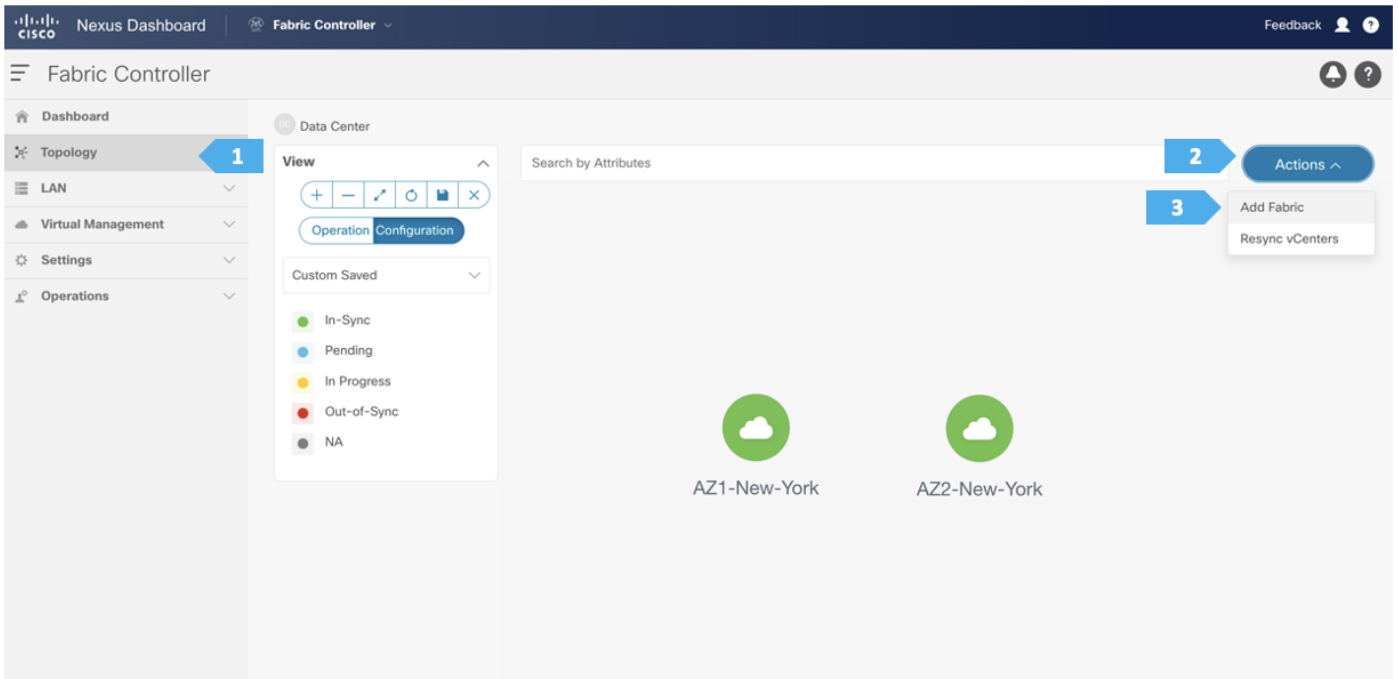


Creating Backbone Fabric

Step 1. Creating the fabric and choosing the template

The third network component that we will be creating is the Backbone. Because this fabric is a Multi-Site Interconnect Network, we will use Route Server (Centralized EVPN peering). With this option, all the BGW nodes deployed in different sites will peer with the same pair of Route Server devices, usually deployed in the Inter-Site Network (ISN).

In the Backbone fabric, we will have RS-10 and RS-11 with the role of “Core Router”. Loopback IP addresses are required on Route Servers to establish BGP EVPN full-mesh peering with the BGW nodes that are associated with different fabrics in the Multi-Site domain.



1 Fabric Name
Backbone



Pick a Fabric

2 Choose Fabric

Close

Select Type of Fabric

Search Type of Fabric

- Fabric Group**
Domain that can contain Enhanced Classic LAN, Classic LAN, and External Connectivity Network fabrics.
- Classic LAN**
Fabric to manage a legacy Classic LAN deployment with Nexus switches.
- LAN Monitor**
Fabric for monitoring Nexus switches for basic discovery and inventory management.
- VXLAN EVPN Multi-Site**
Domain that can contain multiple VXLAN EVPN Fabrics with Layer-2/Layer-3 Overlay Extensions and other Fabric Types.
- Multi-Site Interconnect Network**
Fabric to interconnect VXLAN EVPN fabrics for Multi-Site deployments with a mix of Nexus and Non-Nexus devices.
- External Connectivity Network**
Fabric for Core and Edge router deployments with a mix of Nexus and Non-Nexus devices.

2 Select

Choose Fabric

Close

After clicking Select, we will be presented with a screen with multiple tabs. This type of fabric only needs one parameter, which is the BGP AS number; the rest of the parameters in all tabs are automatically populated.

Please note that the parameters displayed are the minimum to get the fabric up and running, and to make it part of a Multi-Site setup. Please refer to the following link and choose the configuration guide based on the software version being used to understand what each parameter does and to modify the settings based on the specifics of your deployment:

<https://www.cisco.com/c/en/us/support/cloud-systems-management/prime-data-center-network-manager/products-installation-and-configuration-guides-list.html>

Filling in the parameters in the “General Parameters” tab

In this tab, we will be filling in only the BGP ASN. Enter the BGP AS number that is associated with the fabric. In this example, we will be using 65003 as the BGP ASN number.

Note: Please uncheck the “Fabric Monitor Mode” option since NDFC will be managing the devices that belongs to the Multi-Site Interconnect Network.

Create Fabric

Fabric Name
Backbone

Pick Fabric
Multi-Site Interconnect Network >

1 General Parameters Advanced Resources Configuration Backup Bootstrap Flow Monitor

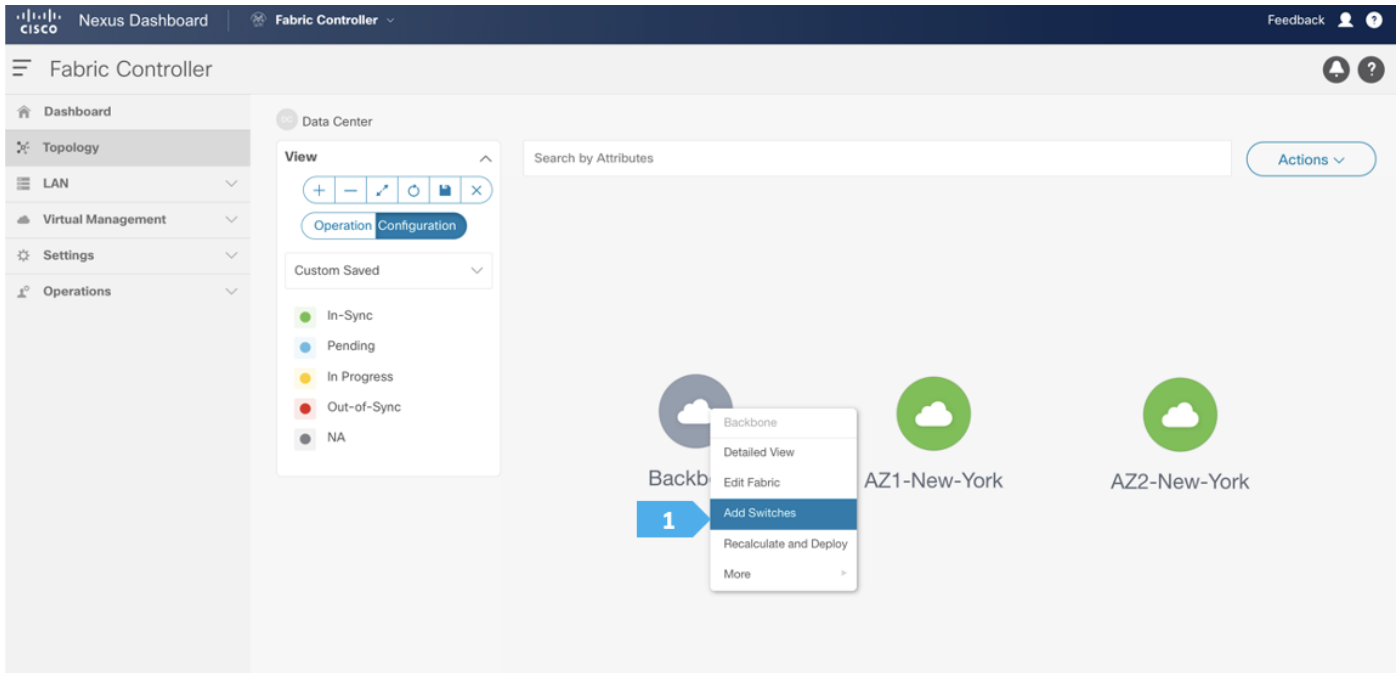
2 BGP AS #* 65003
1-4294967295 | 1-65535[0-65535] It is a good practice to have a unique ASN for each Fabric.

3 Fabric Monitor Mode
If enabled, fabric is only monitored. No configuration will be deployed

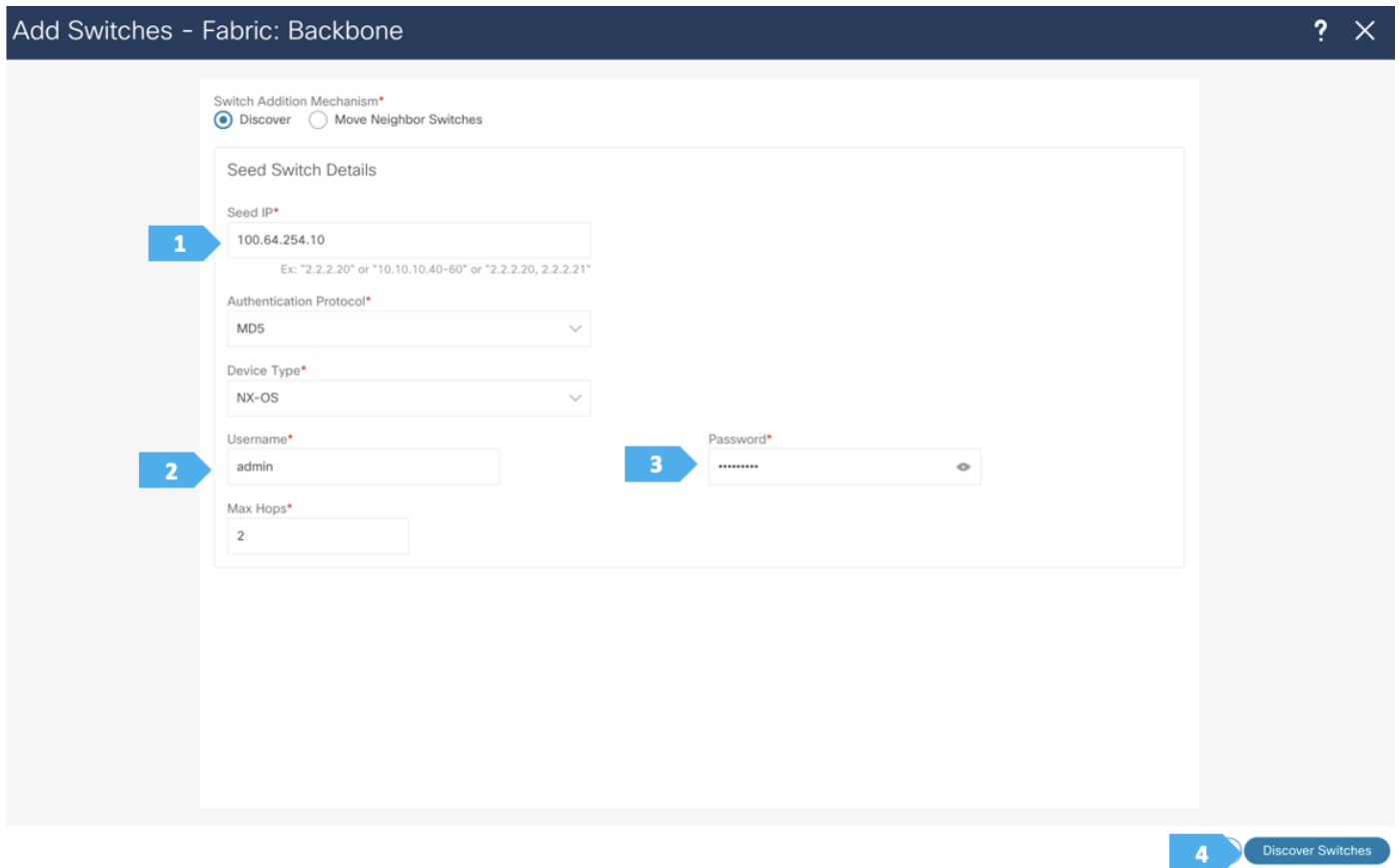
Enable Performance Monitoring (For NX-OS Switches Only)

4 Save

Step 2. Adding switches to the Backbone Fabric



Use a seed IP address to discover the switches. The max hop count allows the discovery of connected switches by the number of hops.



After the switches are discovered, choose the switches to be part of the Backbone fabric and click “Add Switches”.

Add Switches - Fabric: Backbone

Seed Switch Details

Fabric	Switch	Authentication Protocol	Username
Backbone	100.64.254.10	MD5	admin
Password	Max Hops	Preserve config	
Set	2	Enabled	

Discovery Results

Filter by attributes

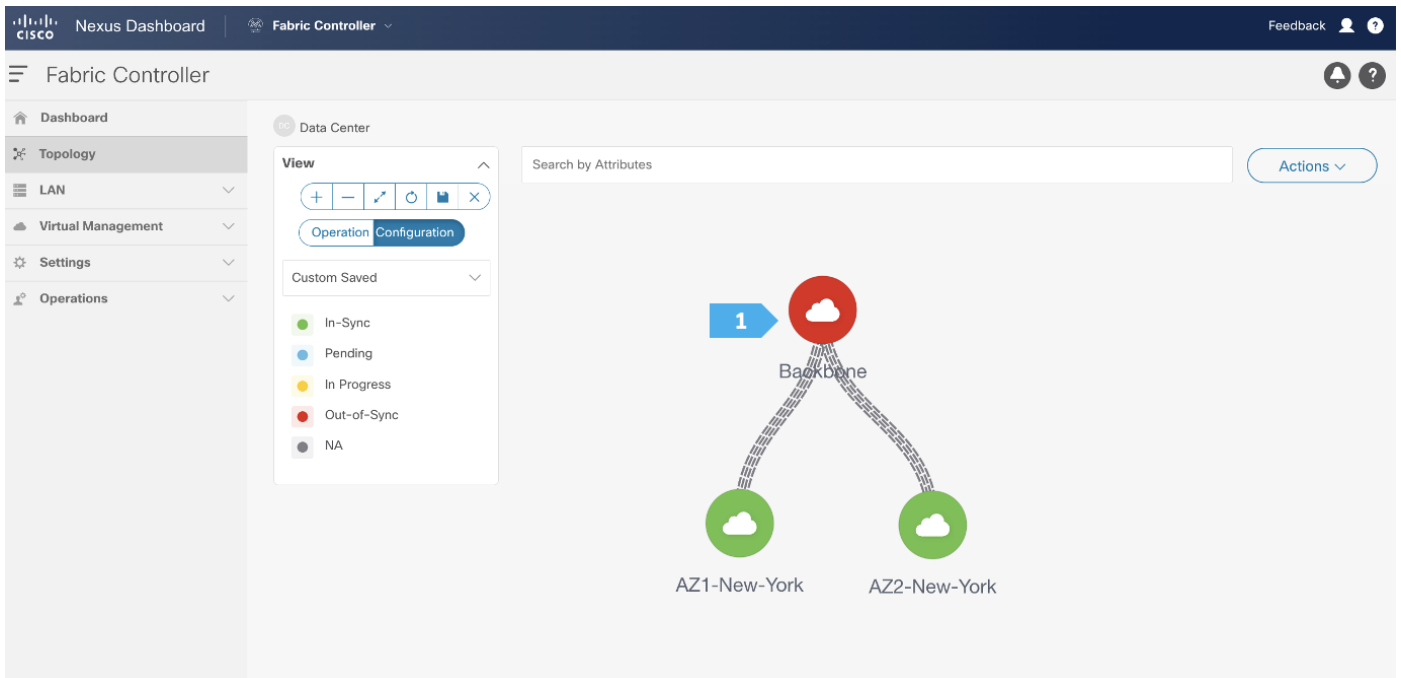
<input type="checkbox"/>	Switch Name	Serial Number	IP Address	Model	Version	Status	Progress
<input checked="" type="checkbox"/>	RS-10	9W9A4AM8HLH	100.64.254.10	N9K-C9300v	10.2(5)	Manageable	
<input type="checkbox"/>	SB-21	96T9O5DS3BJ	100.64.254.21	N9K-C9300v	10.2(5)	Manageable	
<input type="checkbox"/>	BGWS-201	9AOZRKA9IY1	100.64.254.201	N9K-C9300v	10.2(5)	Already Managed In AZ1-	
<input type="checkbox"/>	BGWS-202	9046ZFS3G8	100.64.254.202	N9K-C9300v	10.2(5)	Already Managed In AZ1-	
<input checked="" type="checkbox"/>	RS-11	9AB4MSSB0XQ	100.64.254.11	N9K-C9300v	10.2(5)	Manageable	
<input type="checkbox"/>	BGW-114	9UGXZDIWVVV	100.64.254.114	N9K-C9300v	10.2(5)	Already Managed In AZ2-	
<input type="checkbox"/>	Leaf-101	9ZEA13L749S	100.64.254.101	N9K-C9300v	10.2(5)	Already Managed In AZ1-	
<input type="checkbox"/>	Spine-211	9LE10D1ZXIZ	100.64.254.211	N9K-C9300v	10.2(5)	Already Managed In AZ2-	
<input type="checkbox"/>	SB-20	9K1BU3YG7MC	100.64.254.20	N9K-C9300v	10.2(5)	Manageable	
<input type="checkbox"/>	BGW-113	9FG3KP3OV6	100.64.254.113	N9K-C9300v	10.2(5)	Already Managed In AZ2-	

2 Add Switches

Please wait until the Progress for all of the switches being added is green, then click “Close”.

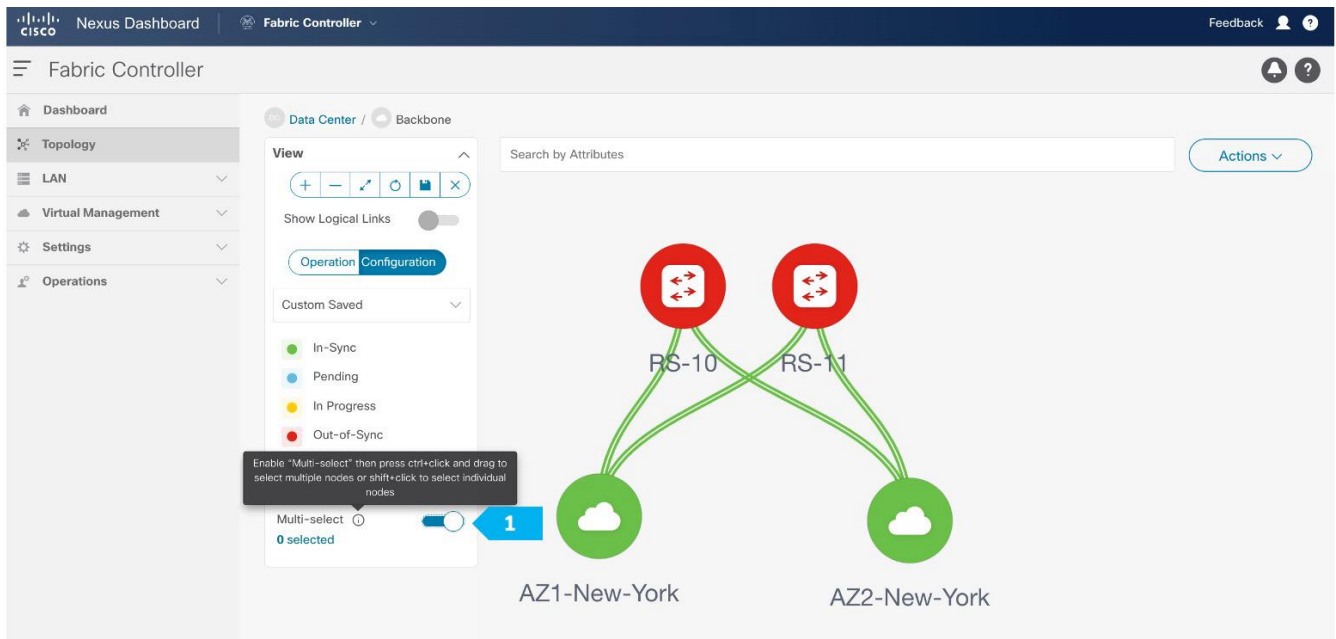
Step 3. Changing the devices’ role

After the devices are added to the Backbone fabric, they will be assigned a default role depending on the platform. In this example configuration, we will assign RS-10 and RS-11 the “Core Router” role. Assigning this role will push the relevant configurations to the respective devices. We can assign this role after we double-click on the Backbone fabric, as shown in the next screen.



We see the fabric color is red, which means that it is out of sync and the intended configuration that we want is not yet pushed to the switches.

Enable the Multi-select option as shown below, then press Ctrl + click and drag your mouse to select RS-10 and RS-11.



We must release the modifier keys “ctrl” before releasing mouse drag to end the switch selection.

Nexus Dashboard | Fabric Controller | Feedback

Fabric Controller

Dashboard | Topology | LAN | Virtual Management | Settings | Operations

Data Center / Backbone

Search by Attributes

View

Show Logical Links

Operation Configuration

Custom Saved

- In-Sync
- Pending
- In Progress
- Out-of-Sync
- NA

Multi-select 0 selected

RS-10

AZ1-New-York

AZ2-New-York

2 nodes selected

- Preview Config
- Deploy Config
- Discovery
- Set Role
- More

Nexus Dashboard | Fabric Controller | Feedback

Fabric Controller

Dashboard | Topology | LAN | Virtual Management | Settings | Operations

Data Center / Backbone

Search by Attributes

View

Show Logical Links

Operation Configuration

Custom Saved

- In-Sync
- Pending
- In Progress
- Out-of-Sync
- NA

Multi-select 0 selected

Select Role

Search Role

- Border Spine
- Border Gateway
- Border Gateway Spine
- Super Spine
- Border Super Spine
- Border Gateway Super Spine
- Access
- Aggregation
- Edge Router
- Core Router
- ToR

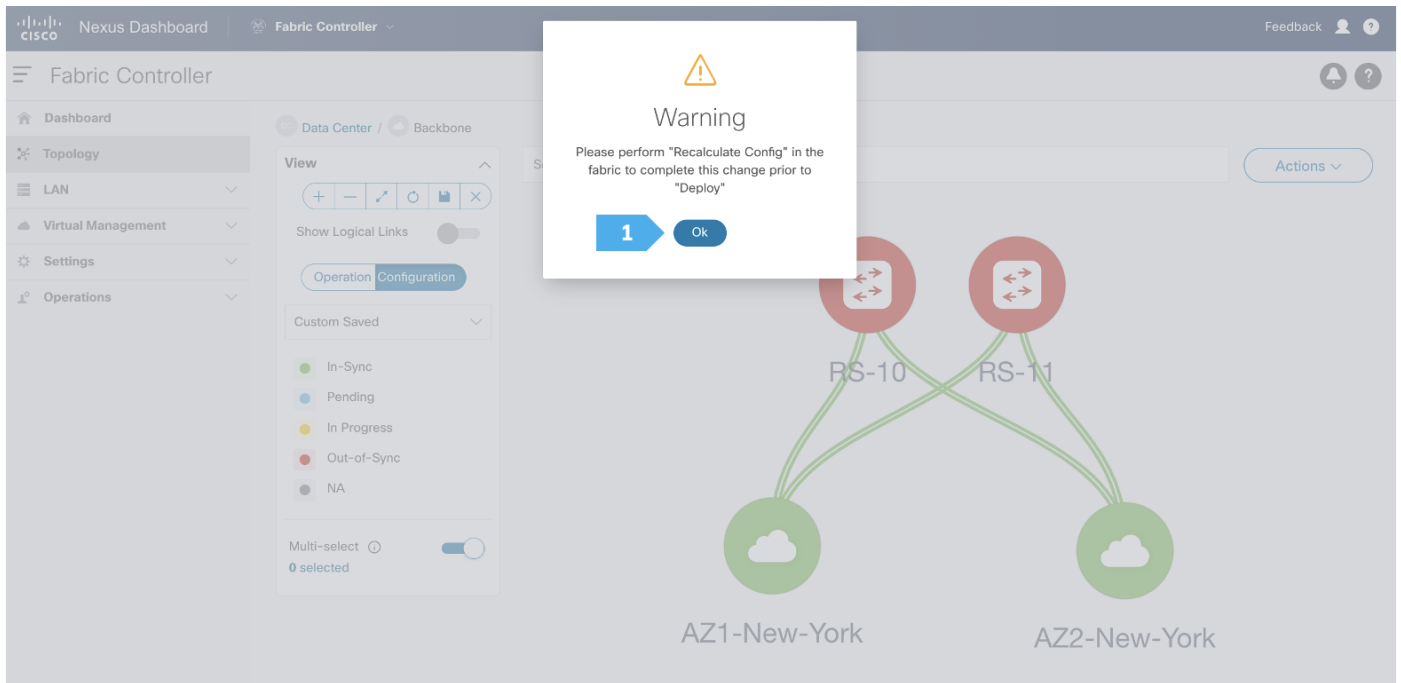
AZ1-New-York

AZ2-New-York

1

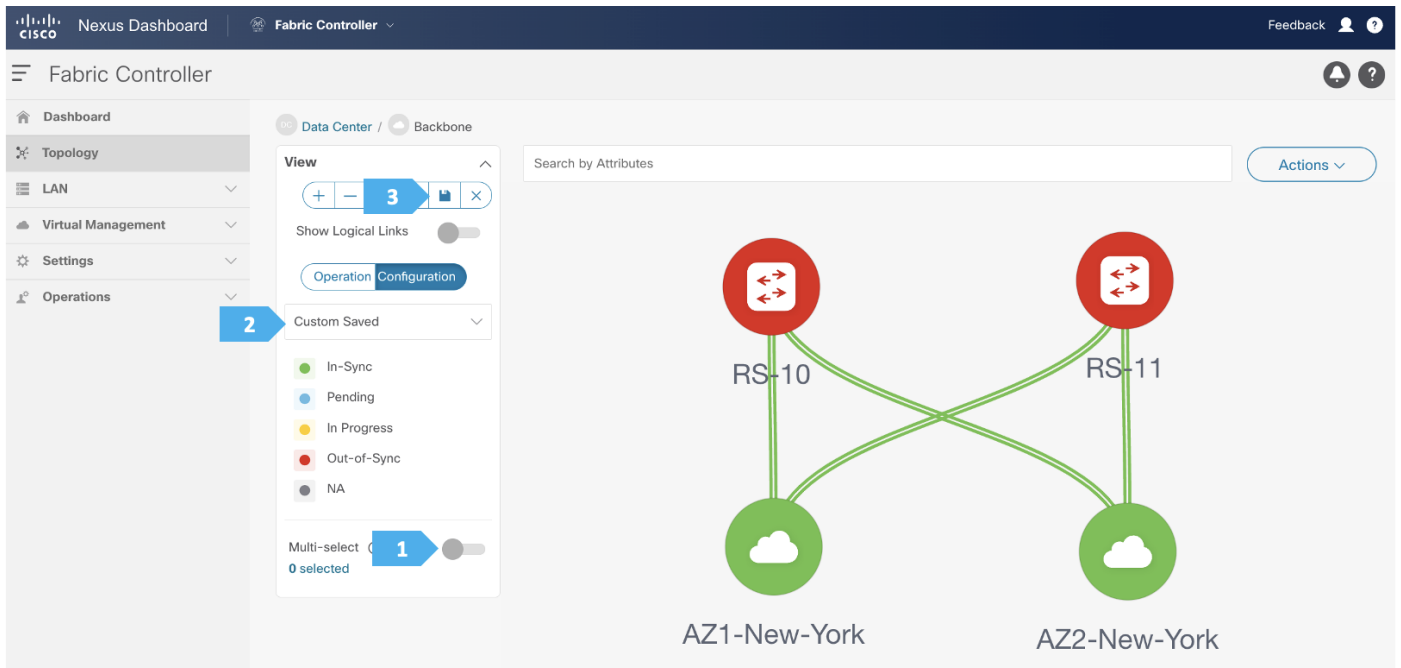
2

Select



After setting the device role, toggle the Multi-select option to disable the multi-select function.

To adjust the topology to look like the screenshot below, choose “Custom Saved” and move the switches around to update the topology like below, then click the save icon as shown.



Step 4. Create a loopback interface per each route server

We need to create a loopback on RS-10 and RS-11, following the steps shown in the next screen shots. Loopback IP addresses are required on Route Servers to establish BGP EVPN full-mesh peering with the BGWs that are associated with different fabrics in the Multi-Site domain. Each AZ will deploy dedicated BGWs that will peer with the Route Servers.

Create Interface

1

Type*

Port Channel ^

Port Channel ✓

virtual Port Channel (vPC)

Straight-through (ST) FEX

Active-Active (AA) FEX

2

Loopback

Tunnel

Ethernet

Switch Virtual Interface (SVI)

We need to choose one of the Route Server nodes from the drop down menu as show. The loopback must be provisioned in the “default” VRF. Repeat for both RS-10 and RS-11.

Create Interface



1

Type*
Loopback

Select a device*

- RS-10
- RS-10**
- RS-11

Policy*
int_loopback >
Policy Options

2

Interface VRF
default Interface VRF name, default VRF if not specified

3

Loopback IP
10.254.254.10 Configured if VRF is non-default. For default VRF configured only if underlay is V4, add config to freeform if underlay is V6.

Loopback IPv6 Address
Configured if VRF is non-default. For default VRF configured only if underlay is V6, add config to freeform if underlay is V4.

Route-Map TAG
12345 Route-Map tag associated with interface IP

Interface Description
Add description to the interface (Max Size 254)

Freeform Config

4

Save Preview Deploy

Create Interface



Type*
Loopback

Select a device*
RS-10

Loopback ID*
0

Policy*
int_loopback
Policy Options

Interface VRF
default Interface VRF name, default VRF if not specified

Loopback IP
10.254.254.10 Configured if VRF is non-default. For default VRF configured only if underlay is V4, add config to freeform if underlay is V6.

Loopback IPv6 Address
Configured if VRF is non-default. For default VRF configured only if underlay is V6, add config to freeform if underlay is V4.

Route-Map TAG
12345 Route-Map tag associated with interface IP

Interface Description
Add description to the interface (Max Size 254)

Freeform Config

Save 1 Deploy

Please repeat the the same steps for RS-11.

Create Interface

Type*
Loopback

1 Select a device*
RS-11
RS-10
RS-11

Policy*
int_loopback >

Policy Options

2 Interface VRF
default

3 Loopback IP
10.254.254.11

Loopback IPv6 Address

Route-Map TAG
12345

Interface Description

Freeform Config

4 Save Preview Deploy

Create Interface



Type*
Loopback

Select a device*
RS-11

Loopback ID*
0

Policy*
int_loopback

Policy Options

Interface VRF
default Interface VRF name, default VRF if not specified

Loopback IP
10.254.254.11 Configured if VRF is non-default. For default VRF configured only if underlay is V4, add config to freeform if underlay is V5.

Loopback IPv6 Address
Configured if VRF is non-default. For default VRF configured only if underlay is V5, add config to freeform if underlay is V4.

Route-Map TAG
12345 Route-Map tag associated with interface IP

Interface Description
Add description to the interface (Max Size 254)

Freeform Config

Save **1** Deploy

Step 5. Recalculate and Deploy to the fabric

The screenshot shows the Cisco Nexus Dashboard Fabric Controller interface. The top navigation bar includes the Cisco logo, 'Nexus Dashboard', and 'Fabric Controller'. The left sidebar contains navigation options: Dashboard, Topology, LAN, Virtual Management, Settings, and Operations. The main area displays a network topology diagram with two switches, RS-10 and RS-11, connected to two access zones, AZ1-New-York and AZ2-New-York. The diagram shows a mesh-like connection between the switches and the access zones. On the right side, an 'Actions' menu is open, showing options: Detailed View, Edit Fabric, Add Switches, Recalculate and Deploy (highlighted with a blue arrow and the number 2), and More. A blue arrow with the number 1 points to the 'Actions' button. The interface also includes a 'View' panel on the left with various controls like zoom, pan, and filters, and a 'Search by Attributes' field at the top right.

Deploy Configuration - Backbone

1 Config Preview 2 Deploy Progress

Filter by attributes Resync All

Switch Name	IP Address	Role	Serial Number	Fabric Status	Pending Config	Status Description	Progress	Resync Switch
RS-10	100.64.254.10	core router	9W9A4AM8HLH	● Out-Of-Sync	5 Lines	Out-of-Sync	<div style="width: 100%;"></div>	Resync
RS-11	100.64.254.11	core router	9AB4MSSB0XQ	● Out-Of-Sync	5 Lines	Out-of-Sync	<div style="width: 100%;"></div>	Resync

1 Deploy All

We can click on the “Pending config” for each switch to view the configuration before clicking “Deploy All”.

Deploy Configuration - Backbone

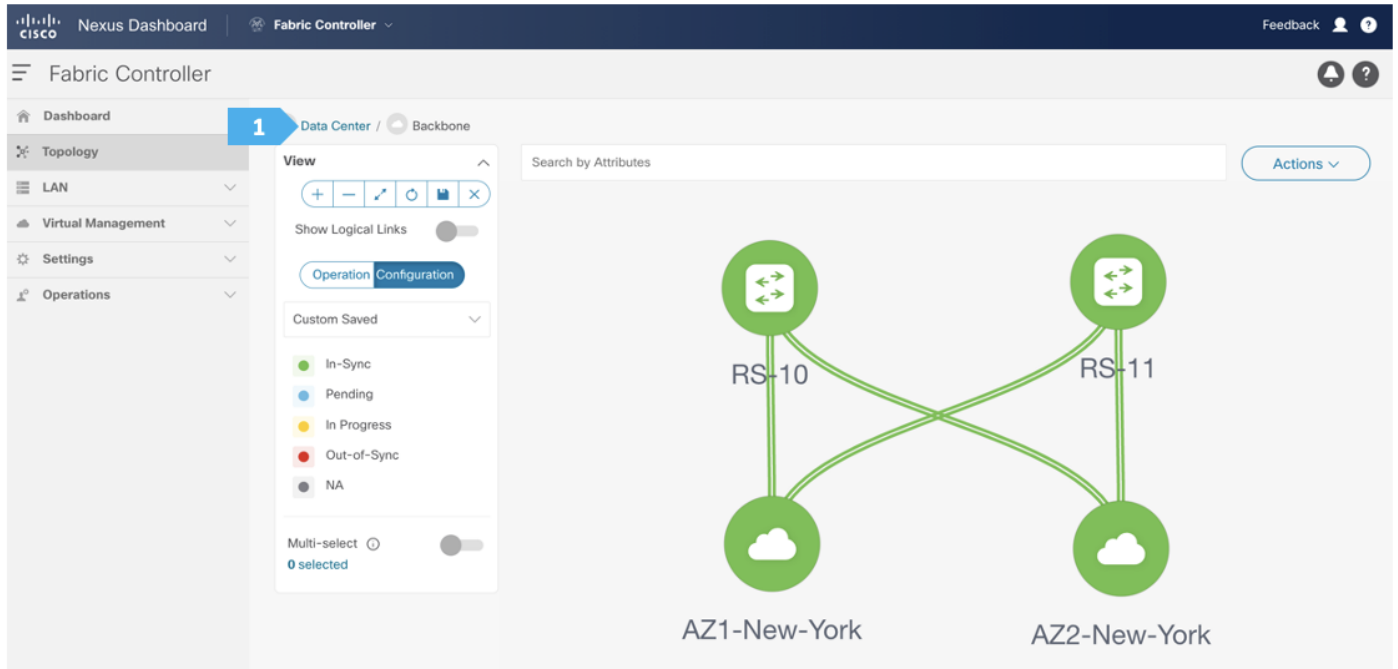
✓ Config Preview 2 Deploy Progress

Filter by attributes

Switch Name	IP Address	Status	Status Description	Progress
RS-10	100.64.254.10	● SUCCESS	Deployment completed.	<div style="width: 100%;"><small>Executed 5 / 5</small></div>
RS-11	100.64.254.11	● SUCCESS	Deployment completed.	<div style="width: 100%;"><small>Executed 5 / 5</small></div>

Backbone Fabric is deployed.

Now all of the switches in the “Backbone” fabric are green, meaning they are “In-Sync” with the intended configuration on NDFC. Click on “Data Center” to go back to main Topology.

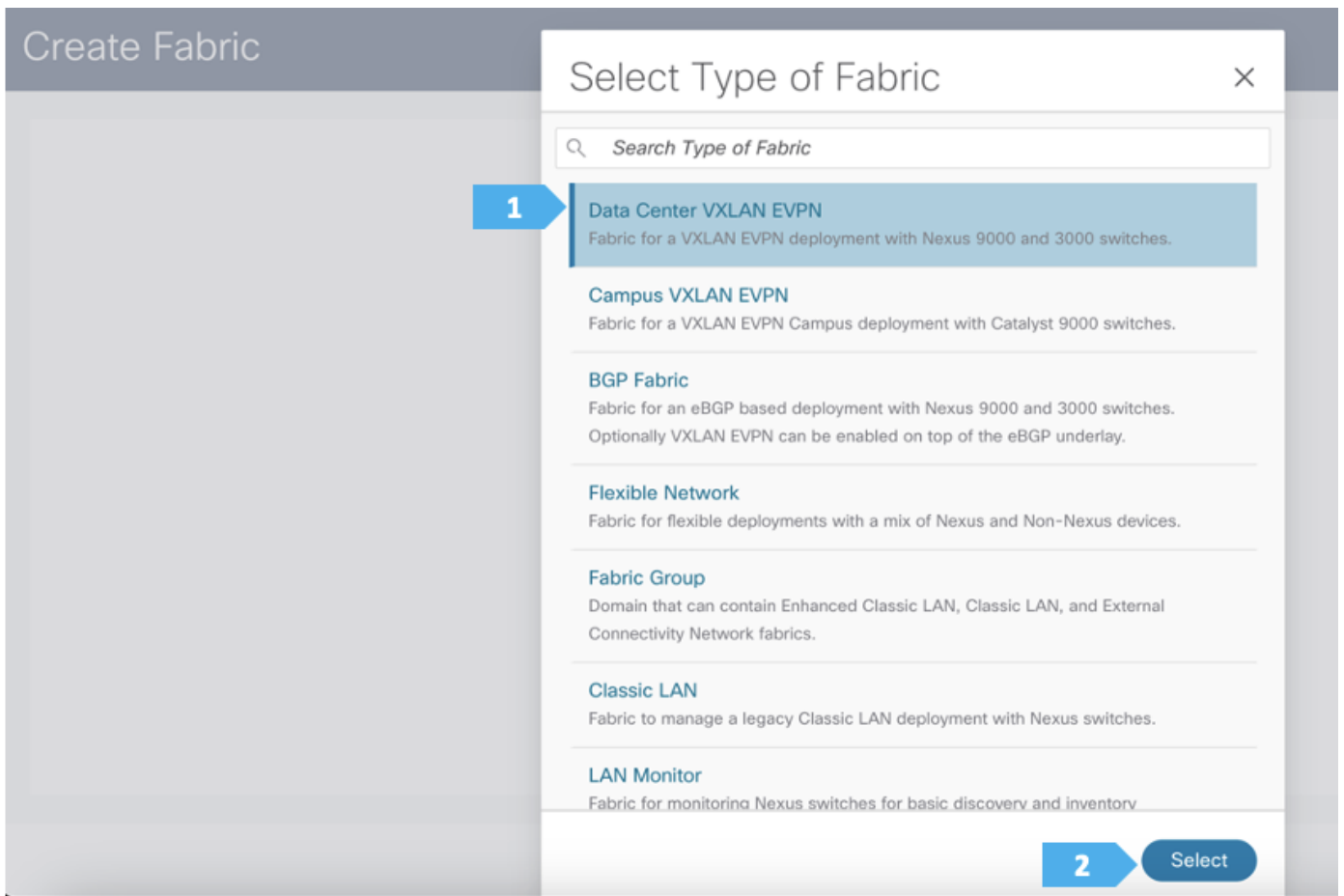
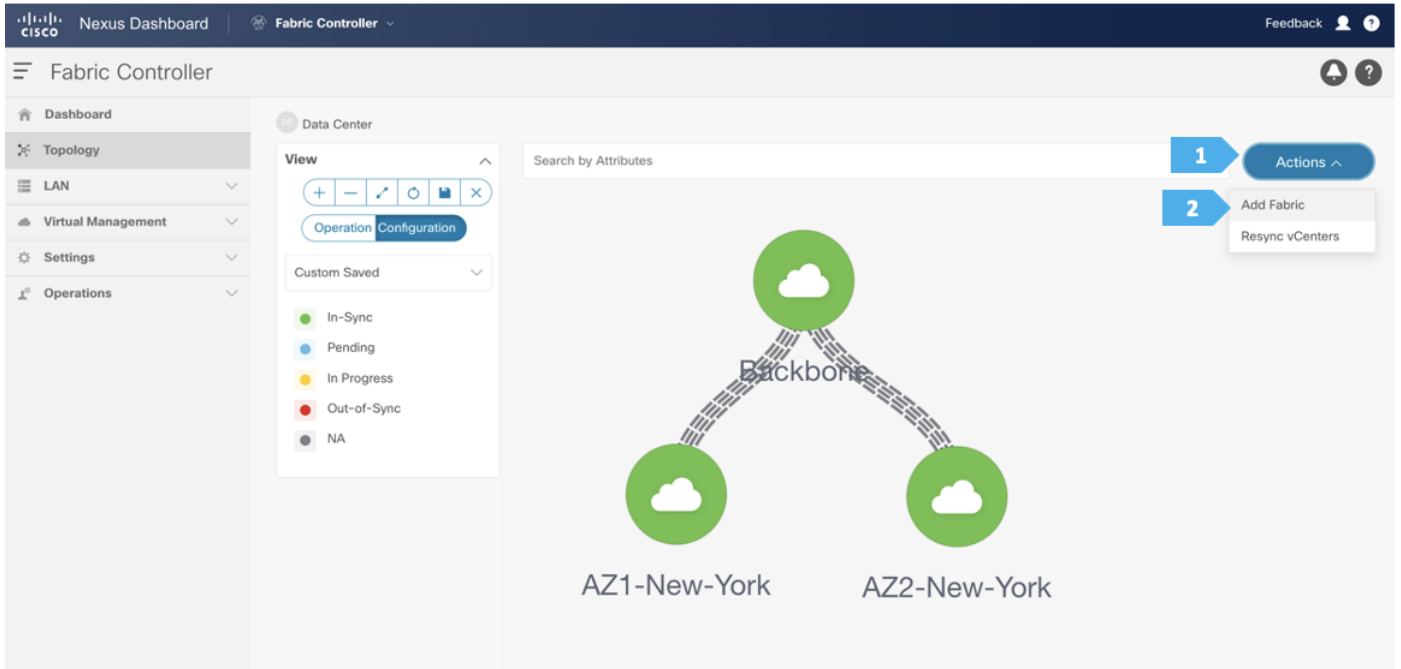


Creating Shared-Border Fabric

Step 1. Creating the fabric and choosing the template

The fourth fabric that we will be creating is Shared-Border, which is a VXLAN EVPN fabric containing SB-20 and SB-21 with the role of “Border”.

The shared border acts as a common external connectivity point for multiple VXLAN BGP EVPN fabrics that are part of the same EVPN Multi-Site architecture. Unlike the BGW, the shared border is completely independent of any VXLAN EVPN Multi-Site software or hardware requirements. It is solely a border node topologically residing outside of a VXLAN EVPN fabric. The shared border operates like a traditional VTEP, but unlike the site-internal VTEPs discussed previously, the shared border is a site-external VTEP.



After clicking Select, we will be presented with a screen with multiple tabs. The overlay and underlay network parameters are included in these tabs.

Please note that the parameters displayed are the minimum to get the fabric up and running and to make it part of a multi-site setup. Please refer to the following link and choose the configuration guide based on the version that you will be using to understand what each parameter does and to make changes based on our design:

<https://www.cisco.com/c/en/us/support/cloud-systems-management/prime-data-center-network-manager/products-installation-and-configuration-guides-list.html>

Filling in the parameters in the “General Parameters” tab

In this tab, we will be entering information only in the **BGP ASN** field. Enter the BGP AS number that the fabric is associated with. In this example, will be using **65004** as the BGP ASN.

Fabric Name
Shared-Border

Pick Fabric
Data Center VXLAN EVPN >

1 General Parameters Replication VPC Protocols Advanced Resources Manageability Bootstrap Configuration Backup Flow Monitor

2 BGP ASN*
65004
1-4294967295 | 1-65535[0-65535] It is a good practice to have a unique ASN for each Fabric.

Enable IPv6 Underlay
 If not enabled, IPv4 underlay is used

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
Mask for Underlay Subnet IP Range

Underlay Subnet IPv6 Mask
Select an Option
Mask for Underlay Subnet IPv6 Range

Underlay Routing Protocol*
ospf
Used for Spine-Leaf Connectivity

Route-Reflectors*
2
Number of spines acting as Route-Reflectors

Filling in the parameters in the “Replication” tab

Replication Mode: This is the mode of replication that is used in the fabric for BUM (Broadcast, Unknown Unicast, Multicast) traffic. The choices are Ingress Replication or Multicast. We will be using the **Multicast** replication mode.

Create Fabric

Fabric Name
Shared-Border

Pick Fabric
Data Center VXLAN EVPN >

General Parameters 1 Replication VPC Protocols Advanced Resources Manageability Bootstrap Configuration Backup Flow Monitor

2 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)

Underlay Primary RP Loopback Id
Used for Bidir-PIM Phantom RP (Min:0, Max:1023)

Filling in the parameters in the “Protocols” tab

The Protocol tab is mostly for the parameters used in the underlay. Most of the parameters are automatically generated. For the purpose of this setup, we will leave everything with the default settings.

Create Fabric

Fabric Name
Shared-Border

Pick Fabric
Data Center VXLAN EVPN >

General Parameters Replicat **1** Protocols Advanced Resources Manageability Bootstrap Configuration Backup Flow Monitor

Underlay Routing Loopback Id*
0 (Min:0, Max:1023)

Underlay VTEP Loopback Id*
1 (Min:0, Max:1023)

Underlay Anycast Loopback Id
Used for vPC Peering in VXLANv6 Fabrics (Min:0, Max:1023)

Underlay Routing Protocol Tag*
UNDERLAY Underlay Routing Process Tag

OSPF Area Id*
0.0.0.0 OSPF Area Id in IP address format

Enable OSPF Authentication

OSPF Authentication Key ID
(Min:0, Max:255)

OSPF Authentication Key
3DES Encrypted

Filling in the parameters in the “Advanced” tab

In the Advanced tab, everything is automatically populated. We will only change the setting in the Overlay mode field.

Overlay Mode: We can create a VRF or network in CLI or config-profile mode at the fabric level. We will be using CLI for this example configuration.

Create Fabric

Fabric Name
Shared-Border

Pick Fabric
Data Center VXLAN EVPN >

General Parameters Replication VPC **1** Advanced Resources Manageability Bootstrap Configuration Backup Flow Monitor

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
2 cli Enable PVLAN on switches except spines and super spines

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

Site Id
65004 For EVPN Multi-Site Support (Min:1, Max: 281474976710655). Defaults to Fabric ASN

Filling in the parameters in the “Resources” tab

By default, Nexus Dashboard Fabric Controller dynamically allocates the underlay IP address resources (for loopbacks, fabric interfaces, and so on) from the defined pools. Please make sure to choose a unique pool per fabric.

Edit Fabric : Shared-Border

Fabric Name

Shared-Border

Pick Fabric

Data Center VXLAN EVPN >

General Parameters Replication VPC Protocols **1** Resources Manageability Bootstrap Configuration Backup Flow Monitor

Manual Underlay IP Address Allocation

Checking this will disable Dynamic Underlay IP Address Allocations

Underlay Routing Loopback IP Range*

10.41.0.0/22

Typically Loopback0 IP Address Range

Underlay VTEP Loopback IP Range*

10.42.0.0/22

Typically Loopback1 IP Address Range

Underlay RP Loopback IP Range*

10.254.40.0/24

Anycast or Phantom RP IP Address Range

Underlay Subnet IP Range*

10.43.0.0/16

Address range to assign Numbered and Peer Link SVI IPs

Edit Fabric : Shared-Border

Auto Deploy Default VRF

Whether to auto generate Default VRF interface and BGP peering configuration on VRF LITE IFC auto deployment. If set, auto created VRF Lite IFC links will have 'Auto Deploy Default VRF' enabled.

Auto Deploy Default VRF for Peer

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

Redistribute BGP Route-map Name

Route Map used to redistribute BGP routes to IGP in default vrf in auto created VRF Lite IFC links

VRF Lite Subnet IP Range*

10.44.0.0/16

Address range to assign P2P Interfabric Connections

VRF Lite Subnet Mask*

30

(Min:8, Max:31)

Service Network VLAN Range*

3000-3199

Per Switch Overlay Service Network VLAN Range (Min:2, Max:4094)

Route Map Sequence Number Range*

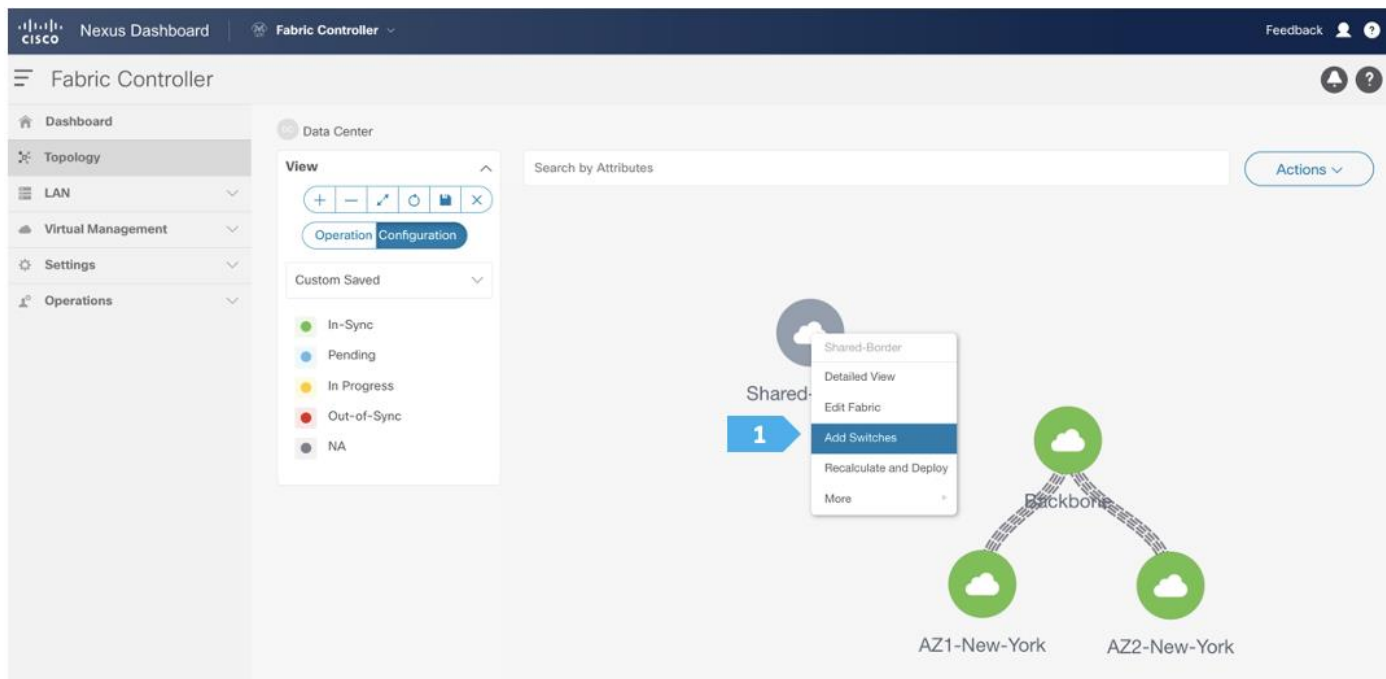
1-65534

(Min:1, Max:65534)

Filling in the parameters in the “Manageability”, “Bootstrap”, “Configuration Backup” and “Flow Monitor” tabs

We will use the defaults for all these tabs, so all what we need to do is to click Save in each window.

Step 2. Adding switches to the Shared-Border Fabric



Use seed IP address to discover the switches. We will use the admin username and password to discover the switches. Uncheck the “Preserve Config” option to clear the switch configuration and to reload the devices. Max hop count allows for the discovery of connected switches by the number of hops.

Switch Addition Mechanism*
 Discover

Seed Switch Details

1 Seed IP*
100.64.254.20
Ex: "2.2.2.20" or "10.10.10.40-60" or "2.2.2.20, 2.2.2.21"

Authentication Protocol*
MD5

2 Username*
admin

3 Password*

Max Hops*
2

4 Preserve Config

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

5 Discover Switches

Switch Addition Mechanism*
 Discover

Seed Switch Details

Seed IP*
100.64.254.20
Ex: "2.2.2.20" or "10.10.10.40-60" or "2.2.2.20,

Authentication Protocol*
MD5


Username*
admin

Password*

Max Hops*
2

Preserve Config

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


Warning
All switch configuration other than management, will be removed immediately after import. Do you want to proceed?
Cancel **Confirm** 1

After the switches are discovered, add these switches to be part of the Shared-Border fabric, then click "Add Switches".

Add Switches - Fabric: Shared-Border



Discover

Seed Switch Details

Fabric: Shared-Border
 Switch: 100.64.254.20
 Authentication Protocol: MD5
 Username: admin
 Password: ● Set
 Max Hops: 2
 Preserve config: ● Disabled

[← Back](#)

Discovery Results

Filter by attributes

<input type="checkbox"/>	Switch Name	Serial Number	IP Address	Model	Version	Status	Progress
<input type="checkbox"/>	RS-10	9W9A4AM8HLH	100.64.254.10	N9K-C9300v	10.2(5)	● Already Managed In Back	
<input checked="" type="checkbox"/>	SB-21	96T9O5DS3BJ	100.64.254.21	N9K-C9300v	10.2(5)	● Manageable	
<input type="checkbox"/>	RS-11	9AB4MSSB0XQ	100.64.254.11	N9K-C9300v	10.2(5)	● Already Managed In Back	
<input type="checkbox"/>	BGWS-201	9AOZRKA9IY1	100.64.254.201	N9K-C9300v	10.2(5)	● Already Managed In AZ1-	
<input type="checkbox"/>	BGWS-202	9046ZFS3G8	100.64.254.202	N9K-C9300v	10.2(5)	● Already Managed In AZ1-	
<input type="checkbox"/>	BGW-114	9UGXZDIWVW	100.64.254.114	N9K-C9300v	10.2(5)	● Already Managed In AZ2-	
<input checked="" type="checkbox"/>	SB-20	9K1BU3YG7MC	100.64.254.20	N9K-C9300v	10.2(5)	● Manageable	
<input type="checkbox"/>	BGW-113	9GFG3KP3OV6	100.64.254.113	N9K-C9300v	10.2(5)	● Already Managed In AZ2-	

➔ Add Switches

1

Please wait until the Progress for all switches being added is green, then click Close.

Add Switches - Fabric: Shared-Border



Discover

Seed Switch Details

Fabric: Shared-Border
 Switch: 100.64.254.20
 Authentication Protocol: MD5
 Username: admin
 Password: ● Set
 Max Hops: 2
 Preserve config: ● Disabled

[← Back](#)

Discovery Results

Filter by attributes

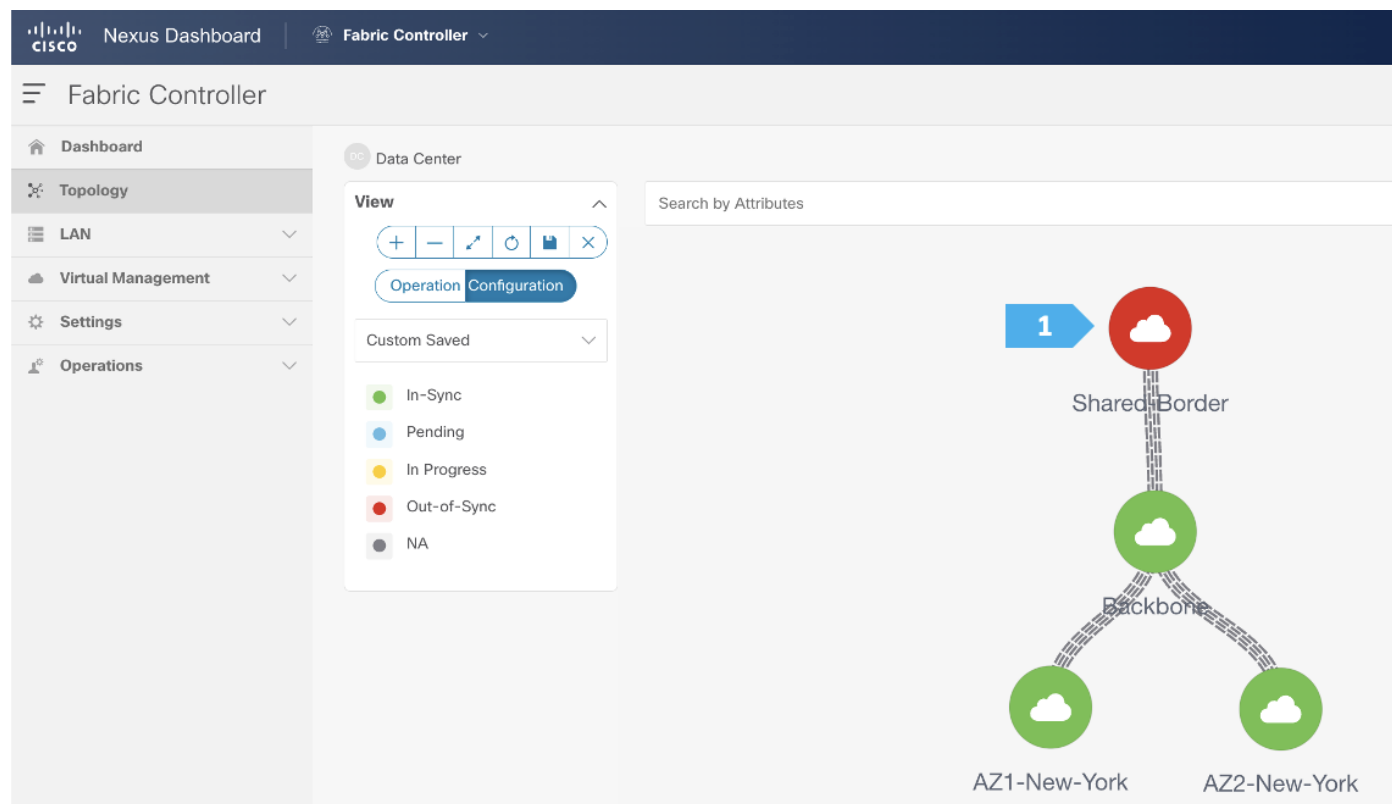
<input type="checkbox"/>	Switch Name	Serial Number	IP Address	Model	Version	Status	Progress
<input type="checkbox"/>	RS-10	9W9A4AM8HLH	100.64.254.10	N9K-C9300v	10.2(5)	● Already Managed In Back	
<input type="checkbox"/>	SB-21	96T9O5DS3BJ	100.64.254.21	N9K-C9300v	10.2(5)	● Switch Added	<div style="width: 100%; height: 10px; background-color: green;"></div>
<input type="checkbox"/>	RS-11	9AB4MSSB0XQ	100.64.254.11	N9K-C9300v	10.2(5)	● Already Managed In Back	
<input type="checkbox"/>	BGWS-201	9AOZRKA9IY1	100.64.254.201	N9K-C9300v	10.2(5)	● Already Managed In AZ1-	
<input type="checkbox"/>	BGWS-202	9046ZFS3G8	100.64.254.202	N9K-C9300v	10.2(5)	● Already Managed In AZ1-	
<input type="checkbox"/>	BGW-114	9UGXZDIWVW	100.64.254.114	N9K-C9300v	10.2(5)	● Already Managed In AZ2-	
<input type="checkbox"/>	SB-20	9K1BU3YG7MC	100.64.254.20	N9K-C9300v	10.2(5)	● Switch Added	<div style="width: 100%; height: 10px; background-color: green;"></div>

➔ Add Switches

1

Step 3. Changing the devices' role

After the devices are added to the Shared-Border fabric, they will be assigned a default role depending on the platform. SB-20 and SB-21 will get the “Border” role, which will push the relevant configuration to the respective devices. We can assign this role after we double-click on the Shared-Border fabric as shown in the next screen.



We see the fabric color is red, which means that it is out of sync and the intended configuration that we want is not yet pushed to the switches.

Enable Multi-Select as shown and press Ctrl + click, then drag the mouse to select SB-20 and SB-21. You must release the modifier key “ctrl” before releasing the mouse drag to end the switch selection.

Nexus Dashboard | Fabric Controller

Fabric Controller

- Dashboard
- Topology
- LAN
- Virtual Management
- Settings
- Operations

DC Data Center / Shared-Border

View

Show Logical Links

Operation Configuration

Custom Saved

- In-Sync
- Pending
- In Progress
- Out-of-Sync

Enable "Multi-select" then press ctrl+click and drag to select multiple nodes or shift+click to select individual nodes

Multi-select 0 selected

1

Search by Attributes

Nexus Dashboard | Fabric Controller

Fabric Controller

- Dashboard
- Topology
- LAN
- Virtual Management
- Settings
- Operations

DC Data Center / Shared-Border

View

Show Logical Links

Operation Configuration

Custom Saved

- In-Sync
- Pending
- In Progress
- Out-of-Sync
- NA

Multi-select 0 selected

1

Search by Attributes

NET Networks (0) VRF VRFs (0)

Nexus Dashboard | Fabric Controller

Fabric Controller

- Dashboard
- Topology
- LAN
- Virtual Management
- Settings
- Operations

Data Center / Shared-

View

Show Logical Links

Operation Configuration

Custom Saved

- In-Sync
- Pending
- In Progress
- Out-of-Sync
- NA

Multi-select 2 selected

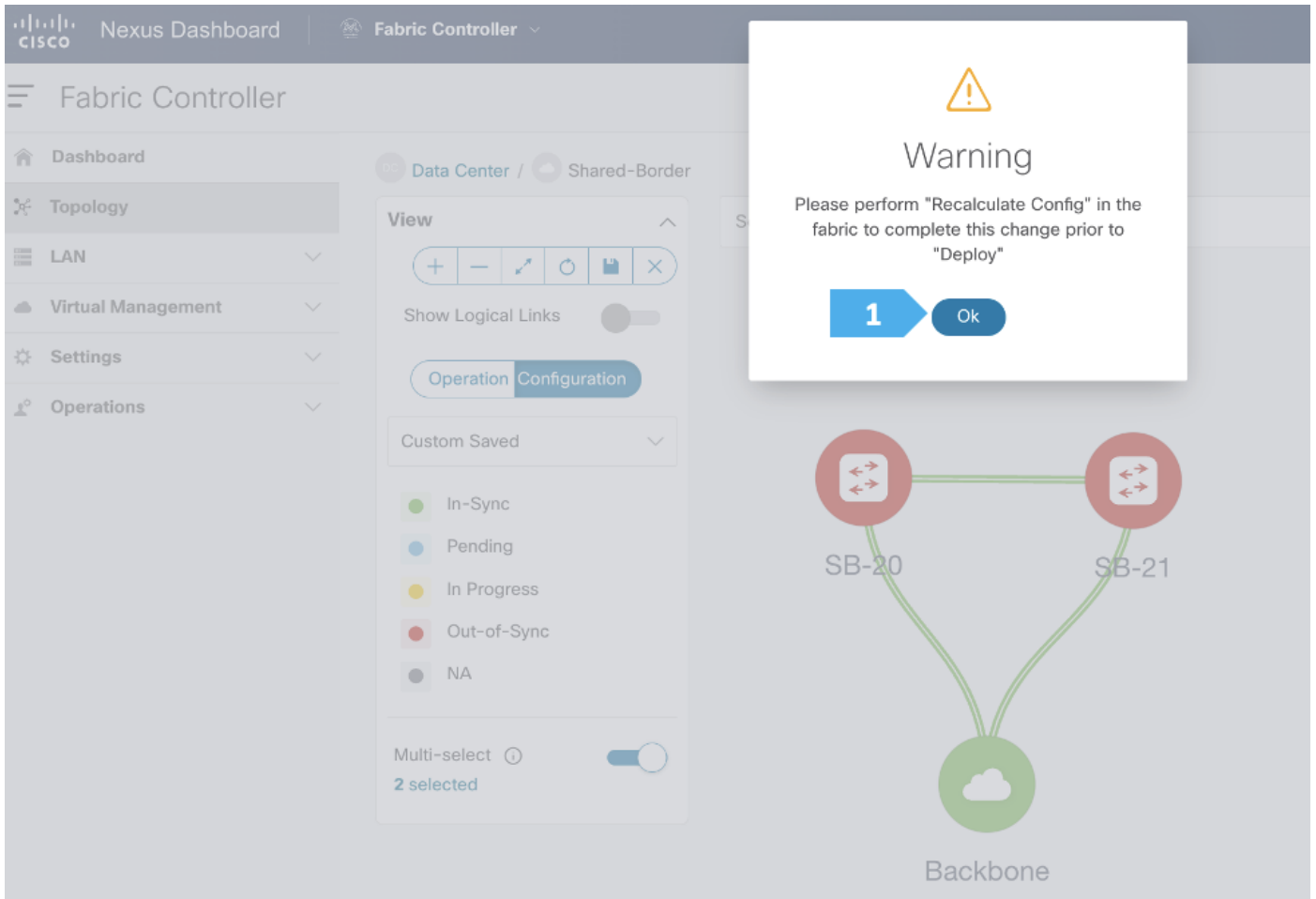
Select Role

Search Role

- Spine
- Leaf (current)
- Border**
- Border Spine
- Border Gateway
- Border Gateway Spine
- Super Spine
- Border Super Spine
- Border Gateway Super Spine
- ToR

2 Select

Backbone



After setting the role, toggle the Multi-select option to disable the multi-select function.

Step 4. Configure vPC between borders

To configure SB-20 and SB-21 as vPC Peers, click on one of the leaf switches and select **vPC Pairing**.

Note: The Shared-Borders are independent Layer 3 VTEPs. The vPC pairing is optional and only required for connecting Site-External service nodes, such as firewall, load balancer, TCP Optimizers, and so on.

Nexus Dashboard | Fabric Controller

Fabric Controller

- Dashboard
- Topology
- LAN
- Virtual Management
- Settings
- Operations

DC Data Center / Shared-Border

Search by Attributes

View

+ - ↗ ↻ 📄 ✕

Show Logical Links

Operation Configuration

Custom Saved

- In-Sync
- Pending
- In Progress
- Out-of-Sync
- NA

Multi-select 0 selected

- SB-20
- Detailed View
- Preview Config
- Deploy Config
- Discovery
- Set Role
- Manage Interfaces
- Manage Policies
- vPC Pairing
- More

Select the peer switch and click Save.

We can click on the “Pending config” for each switch to view the configuration before clicking “Deploy All”.

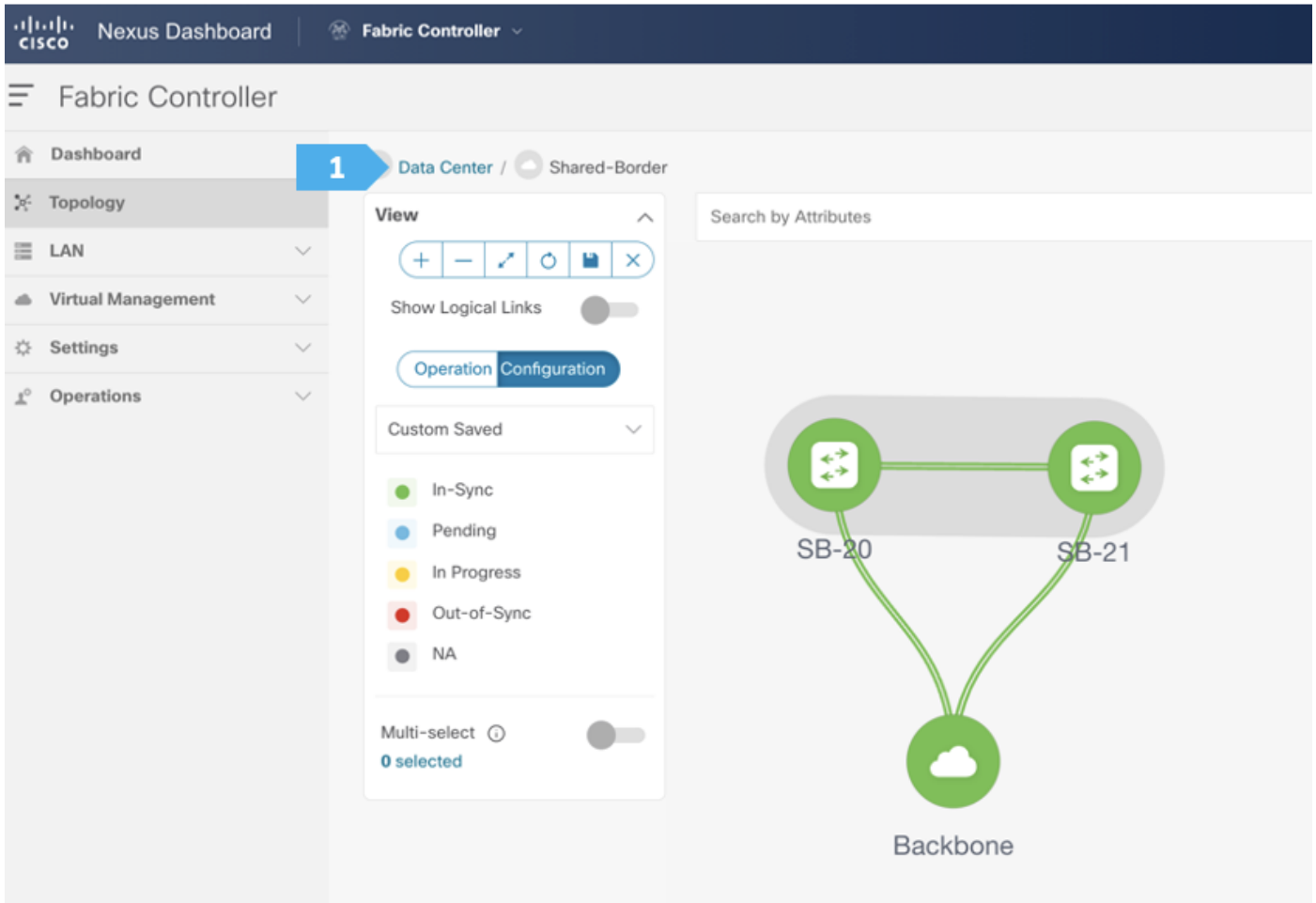
Switch Name	IP Address	Role	Serial Number	Fabric Status	Pending Config	Status Description	Progress	Resync Switch
SB-21	100.64.254.21	border	96T9O5DS3BJ	● Out-Of-Sync	355 Lines	Out-of-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync
SB-20	100.64.254.20	border	9K1BU3YG7MC	● Out-Of-Sync	355 Lines	Out-of-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync

Wait until the “Progress” for each of the switches shows green before clicking “Close”.

Shared-Border Fabric is deployed.

Now all of the switches in the Shared-Border fabric are green, meaning they are “In-Sync”.

Click on “Data Center” to go back to the main topology view.



Creating New-York MSD Fabric

Step 1. Creating the fabric and choosing the template

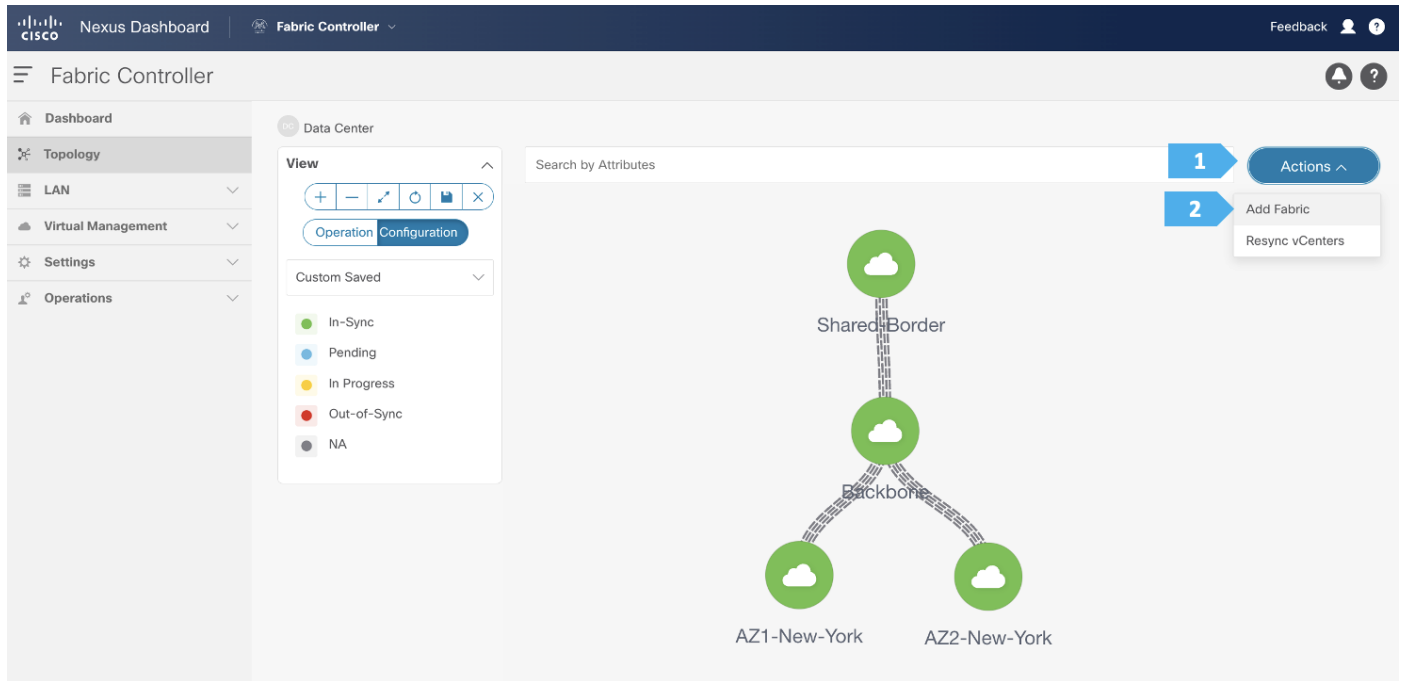
The third fabric that we will be creating is the **New-York** Multi-Site Domain (MSD) fabric. A Multi-Site Domain is a multi fabric container that is created to manage multiple member fabrics.

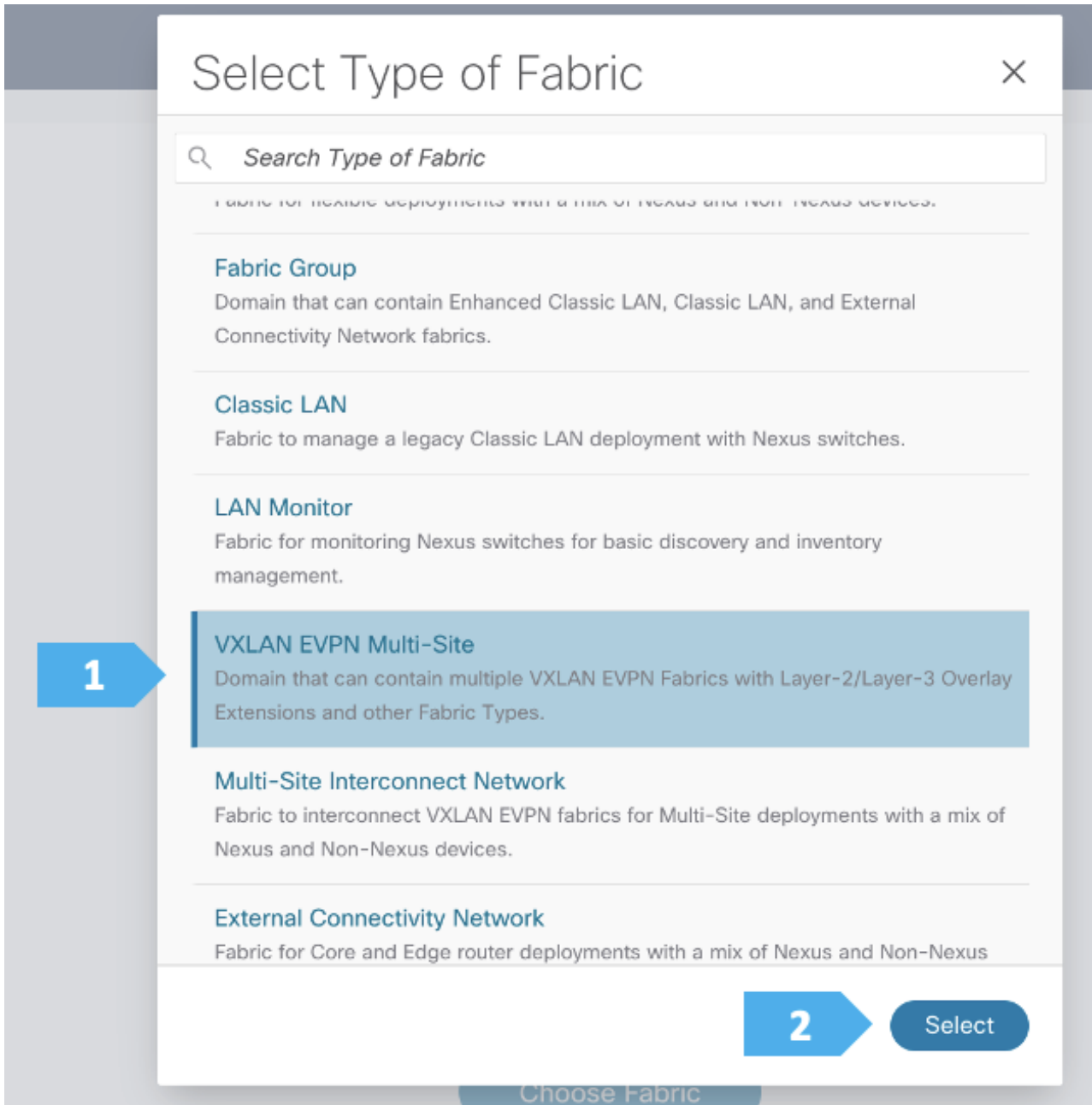
An MSD is a single point of control for the definition of overlay networks and VRFs that are shared across member fabrics. When we move fabrics that are designated to be part of the Multi-Site overlay network domain under the MSD as member fabrics, these member fabrics share the networks and VRFs created at the MSD level. This way, we can consistently provision network and VRFs for different fabrics at one go. It significantly reduces the time and complexity involving multiple fabric provisioning.

As server networks and VRFs are shared across the member fabrics as one stretched network, the provisioning function for the new networks and VRFs is provided at the MSD fabric level. The creation of any new network or VRF is only allowed in the MSD. All member fabrics inherit any new network and VRF created for the MSD.

The topology view for the MSD fabric displays all member fabrics and how they are connected to each other in one view. We can deploy overlay networks and VRFs on member fabrics from a single

topology deployment screen instead of visiting and deploying from each member fabric deployment screen separately.





Filling in the parameters in the “General Parameters” tab

All the parameters in the General Parameters tab will be automatically populated.

Create Fabric

Fabric Name

New-York

Pick Fabric

VXLAN EVPN Multi-Site >

1

General Parameters

DCI

Resources

Configuration Backup

Layer 2 VXLAN VNI Range*

30000-49000

Overlay Network Identifier Range (Min:1, Max:16777214)

Layer 3 VXLAN VNI Range*

50000-59000

Overlay VRF Identifier Range (Min:1, Max:16777214)

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

Enable Private VLAN (PVLAN)

Enable PVLAN on MSD and its child fabrics

PVI AN Secondary Network Template

Filling in the parameters in the “DCI” tab

Since we will be employing the Route Server design using RS-10 and RS-11, we need to change the “**Multi-Site Overlay IFC Deployment Method**” to the “**Centralized_To_Route_Server**” option. We need to supply the Loopback IP addresses as well as the BGP AS number for RS-10 and RS-11 that we created in the Backbone fabric as shown in the next screen shot.

Edit Fabric : New-York

Fabric Name

New-York

Pick Fabric

VXLAN EVPN Multi-Site >

General P **1** DCI Resources Configuration Backup

Multi-Site Overlay IFC Deployment Method*

2

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*

3

10.254.254.10, 10.254.254.11

Multi-Site Router-Server peer list, e.g. 128.89.0.1, 128.89.0.2

Multi-Site Route Server BGP ASN List*

4

65003,65003

1-4294967295 | 1-65535[,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.

5

Multi-Site Underlay IFC Auto Deployment Flag

Filling in the parameters in the “Resources” tab

In the Resources tab, we need to supply the Multi-Site routing loopback IP range and the DCI subnet IP range.

Fabric Name
New-York

Pick Fabric
VXLAN EVPN Multi-Site >

General Parameters 1 Resources Configuration Backup

2 Multi-Site Routing Loopback IP Range*
10.254.0.0/24 Typically Loopback100 IP Address Range.

3 DCI Subnet IP Range*
10.254.1.0/24 Address range to assign P2P DCI Links

Subnet Target Mask*
30 Target Mask for Subnet Range (Min:8, Max:31)

4 Save

Step 2. Moving Fabrics Under the MSD Fabric as a Member

Double-click on the New-York MSD fabric, then click **Actions** > **Add Child Fabric** and start adding all the fabrics as member fabrics.

We can also click **Detailed View** > **Actions** > **Add Child Fabrics** to add member fabrics to the MSD. A list of child fabrics that are not part of any MSD appears. Member fabrics of other MSD container fabrics are not displayed here.

As AZ1-New-York fabric is to be associated with the New-York MSD fabric, select the AZ1-New-York fabric, and click Select.

We can see that AZ1-New-York is now added to the MSD fabric and is displayed in the Child Fabrics in the Fabrics list table.



Fabric Controller

- Dashboard
- Topology
- LAN
- Virtual Management
- Settings
- Operations

Data Center

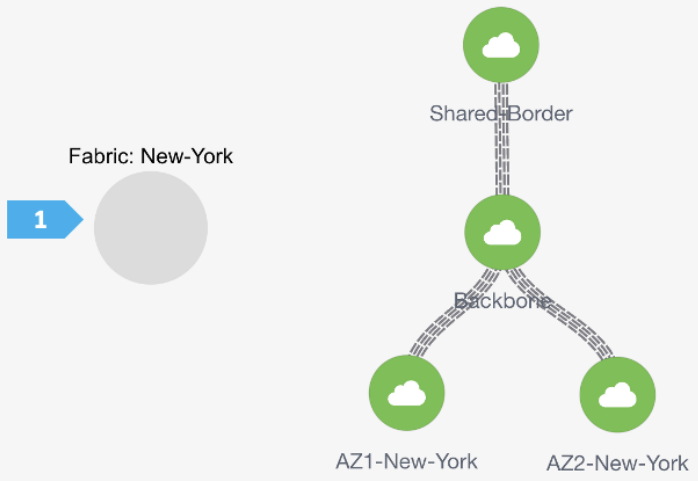
View

Operation Configuration

Custom Saved

- In-Sync
- Pending
- In Progress
- Out-of-Sync
- NA

Search by Attributes



Data Center / New-York

View

Show Logical Links

Operation Configuration

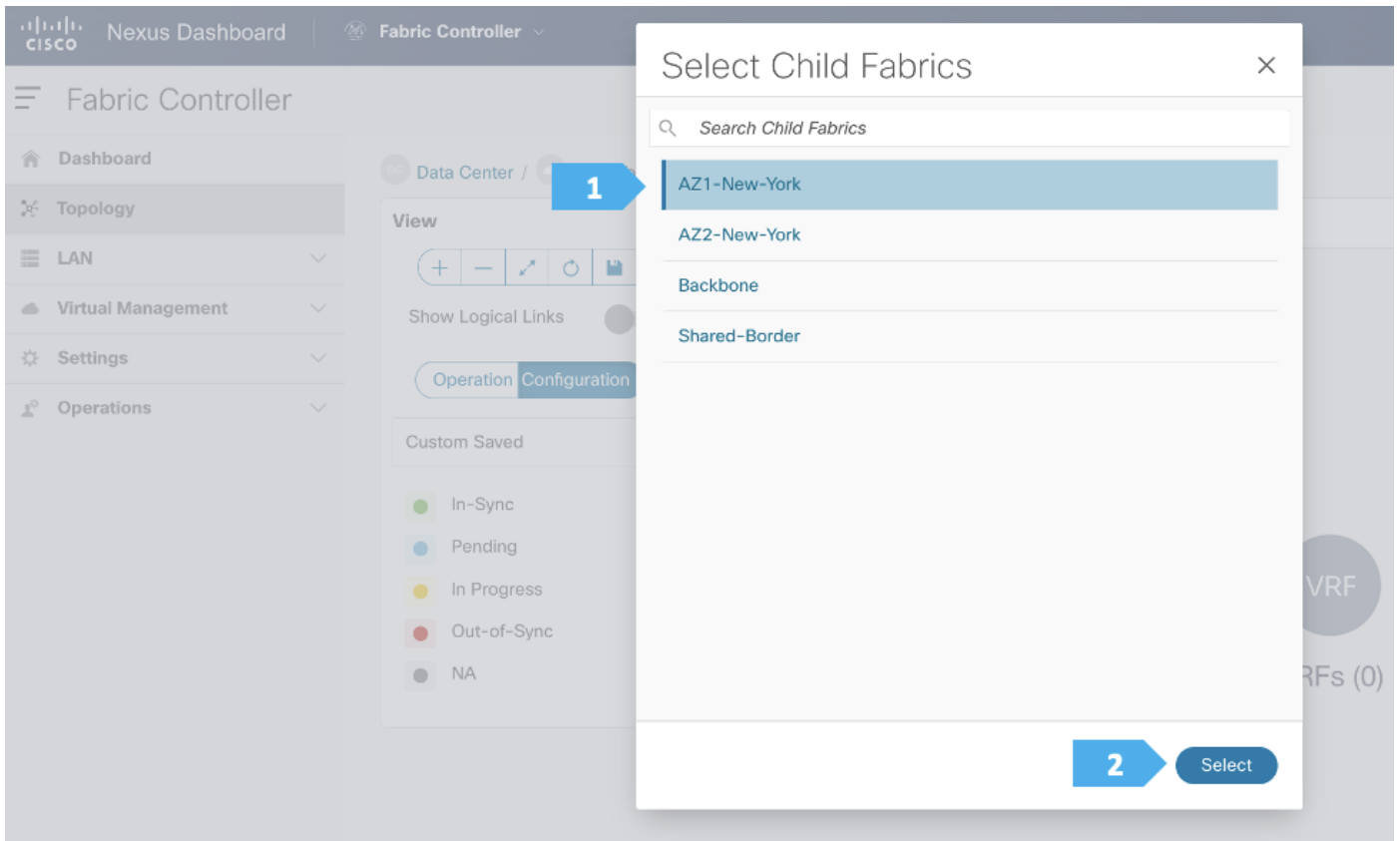
Custom Saved

- In-Sync
- Pending

Search by Attributes

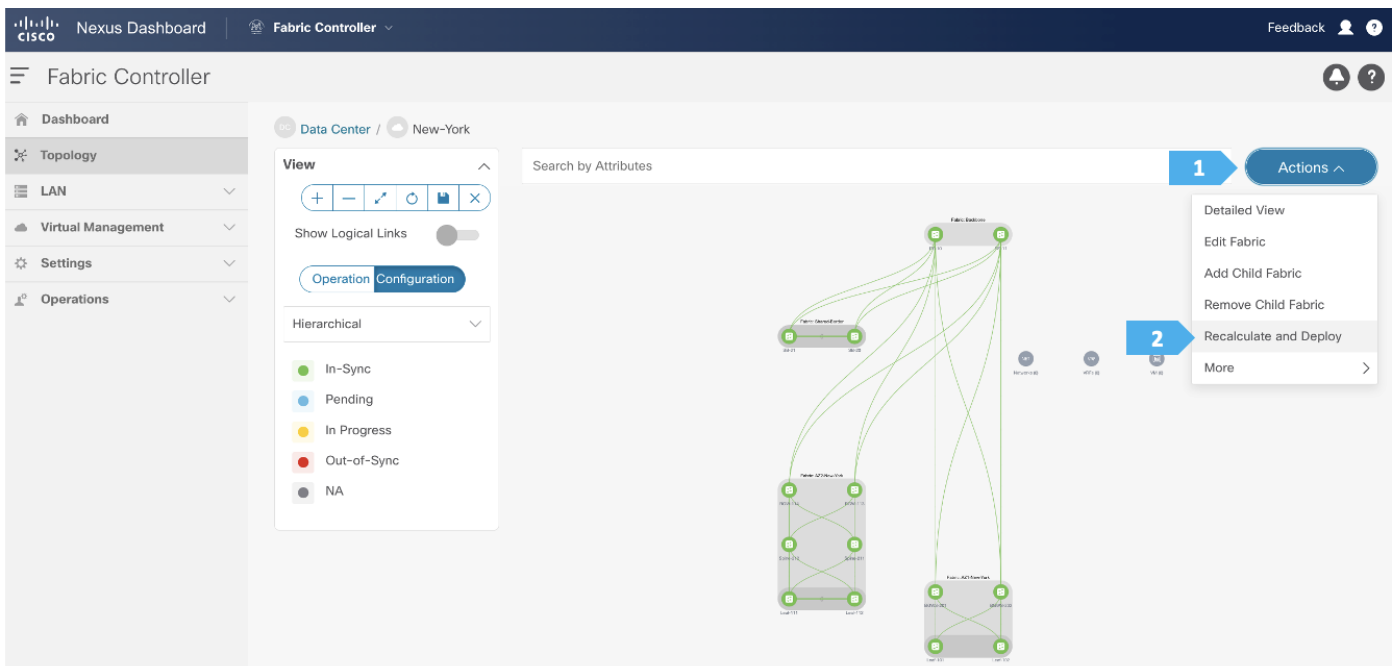
1 Actions

- Detailed View
- Edit Fabric
- 2 Add Child Fabric
- Remove Child Fabric
- Recalculate and Deploy
- More



Repeat these steps for all the fabrics until all the fabrics are part of the New-York MSD fabric. Click on Hierarchical view and click Save. We can also drag and move the fabrics around with the mouse cursor to achieve the view that we want.

Step 3. Recalculate and Deploy to the fabric



1 Config Preview 2 Deploy Progress

Filter by attributes Resync All

Switch Name	IP Address	Role	Serial Number	Fabric Status	Pending Config	Status Description	Progress	Resync Switch
RS-10	100.64.254.10	core router	9W9A4AM8HLH	● Out-Of-Sync	101 Lines	Out-of-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync
RS-11	100.64.254.11	core router	9AB4MSSB0XQ	● Out-Of-Sync	101 Lines	Out-of-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync
BGW-114	100.64.254.114	border gateway	9UGXZDIWIVW	● Out-Of-Sync	90 Lines	Out-of-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync
BGW-113	100.64.254.113	border gateway	9GFG3KP3OV6	● Out-Of-Sync	90 Lines	Out-of-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync
BGWS-202	100.64.254.202	border gateway spine	9046ZFS3G8	● Out-Of-Sync	92 Lines	Out-of-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync
BGWS-201	100.64.254.201	border gateway spine	9AOZRKA9IY1	● Out-Of-Sync	92 Lines	Out-of-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync

2 Deploy All

In the next few screens, we will go through a sample CLI on RS-10 and BGW-114.

RS-10 Sample CLI (Route Server)

```
nv overlay evpn
feature nv overlay
```

Extend the capability of VXLAN with EVPN (nv overlay evpn).

```
interface ethernet1/3
  no switchport
  ip address 10.254.1.2/30 tag 54321
  mtu 9216
  no shutdown

interface ethernet1/4
  no switchport
  ip address 10.254.1.14/30 tag 54321
  mtu 9216
  no shutdown

interface ethernet1/5
  no switchport
  ip address 10.254.1.18/30 tag 54321
  mtu 9216
  no shutdown

interface ethernet1/6
  no switchport
  ip address 10.254.1.30/30 tag 54321
  mtu 9216
  no shutdown
```

Assigning IPv4 addresses to interfaces between RS-10 and remote devices.

```
router bgp 65003
  address-family ipv4 unicast
    maximum-paths 64
    maximum-paths ibgp 64
    network 10.254.254.10/32
  exit
  address-family ipv6 unicast
    maximum-paths 64
    maximum-paths ibgp 64
  exit
  address-family Layer 2vpn evpn
```

```
    retain route-target all
```

Retain and advertise all EVPN routes when there are no local VNI configured with matching import route targets

```
exit
```

```
  template peer OVERLAY-PEERING
    update-source loopback0
    ebgp-multihop 5
    address-family Layer 2vpn evpn
```

```
    route-map unchanged out
```

```
    send-community both
```

```
  exit
```

```
exit
```

```
neighbor 10.254.1.1
  remote-as 65001
  update-source Ethernet1/3
  address-family ipv4 unicast
    next-hop-self
  exit
exit
```

```
neighbor 10.254.1.13
  remote-as 65001
  update-source Ethernet1/4
  address-family ipv4 unicast
    next-hop-self
  exit
exit
```

```
neighbor 10.254.1.17
  remote-as 65002
  update-source Ethernet1/5
  address-family ipv4 unicast
```

The route map enforces the policy to leave the overlay next hop unchanged when the route server is used

```

    next-hop-self
    exit
  exit
neighbor 10.254.1.29
  remote-as 65002
  update-source Ethernet1/6
  address-family ipv4 unicast
    next-hop-self
    exit
  exit
neighbor 10.11.0.3
  remote-as 65001
  inherit peer OVERLAY-PEERING
  address-family Layer 2vpn evpn
    rewrite-evpn-rt-asn
    exit
  exit
neighbor 10.11.0.4
  remote-as 65001
  inherit peer OVERLAY-PEERING
  address-family Layer 2vpn evpn
    rewrite-evpn-rt-asn
    exit
  exit
neighbor 10.21.0.3
  remote-as 65002
  inherit peer OVERLAY-PEERING
  address-family Layer 2vpn evpn
    rewrite-evpn-rt-asn
    exit
  exit
neighbor 10.21.0.5
  remote-as 65002
  inherit peer OVERLAY-PEERING
  address-family Layer 2vpn evpn
    rewrite-evpn-rt-asn

```

```
configure terminal
```

```

route-map unchanged permit 10
  set ip next-hop unchanged

```

The autonomous system portion of the automated route target (ASN:VNI) will be rewritten upon receipt from the site-external network (rewrite-evpn-rt-asn) without modification of any configuration on the site-internal VTEPs. If a route server stands in between the BGWs of the individual sites, an additional rewrite to the destination autonomous system is performed. The route-target rewrite helps ensure that the ASN portion of the automated route target matches the destination autonomous system.

The route map enforces the policy to leave the overlay next hop unchanged when the route server is used. The route server is not a VTEP or BGW and hence should not have the next hop pointing to itself.

BGW-114 Sample CLI (Border Gateway)

```
route-map rmap-redirect-direct permit 10
  match tag 54321
```

```
evpn multisite border-gateway 65002
  delay-restore time 300
```

```
router bgp 65002
  address-family ipv4 unicast
    redistribute direct route-map rmap-redirect-direct
    maximum-paths 64
    maximum-paths ibgp 64
  exit
```

```
maximum-paths 64
  maximum-paths ibgp 64
  exit
```

```
neighbor 10.254.1.25
  remote-as 65003
  update-source Ethernet1/2
  address-family ipv4 unicast
    next-hop-self
  exit
  exit
```

```
neighbor 10.254.1.30
  remote-as 65003
  update-source Ethernet1/1
  address-family ipv4 unicast
    next-hop-self
  exit
  exit
```

```
neighbor 10.254.254.10
  remote-as 65003
  update-source loopback0
  ebgp-multihop 5
```

```
peer-type fabric-external
```

Rewrite RMAC to BGW to enable Rewrite and Reorigination functions on BGW

```
address-family Layer 2vpn evpn
  send-community both
  rewrite-evpn-rt-asn
```

```
        exit
    exit
neighbor 10.254.254.11
    remote-as 65003
    update-source loopback0
    ebgp-multihop 5
    peer-type fabric-external
    address-family Layer 2vpn evpn
        send-community both
        rewrite-evpn-rt-asn
configure terminal
```

```
interface nve1
```

```
    multisite border-gateway interface loopback100
    source-interface loopback1
    host-reachability protocol bgp
    no shutdown
```

Define the loopback100 interface as the EVPN Multi-Site source interface (anycast and virtual IP VTEP).

```
interface loopback100
```

```
    ip address 10.254.0.3/32 tag 54321
    ip router ospf UNDERLAY area 0.0.0.0
    ip pim sparse-mode
    no shutdown
```

```
interface ethernet1/1
```

```
    no switchport
    ip address 10.254.1.29/30 tag 54321
    evpn multisite dci-tracking
    mtu 9216
    no shutdown
```

EVPN Multi-Site interface tracking is used for the site-external underlay (**evpn multisite dci-tracking**). This command is mandatory to enable the Multi-Site virtual IP address on the BGW. At least one of the physical interfaces that are configured with DCI tracking must be up to enable the Multi-Site BGW function.

```
interface ethernet1/2
```

```
    no switchport
    ip address 10.254.1.26/30 tag 54321
    evpn multisite dci-tracking
    mtu 9216
    no shutdown
```

```
interface ethernet1/3
```

```
    no switchport
    ip address 10.23.0.10/30
    evpn multisite fabric-tracking
```

EVPN Multi-Site interface tracking for the site-internal underlay (**evpn multisite fabric-tracking**). This command is mandatory to enable the Multi-Site virtual IP address on the BGW. At least one of the physical interfaces that are configured with fabric tracking must be up to enable the Multi-Site BGW function (keeping the virtual IP VTEP address active).

```
description connected-to-Spine-211-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 ethernet1/4
no switchport
ip address 10.23.0.14/30
evpn multisite fabric-tracking
ip router ospf UNDERLAY area 0.0.0.0
ip ospf network point-to-point
ip pim sparse-mode
no shutdown
description connected-to-Spine-212-Ethernet1/2
mtu 9216
configure terminal
```

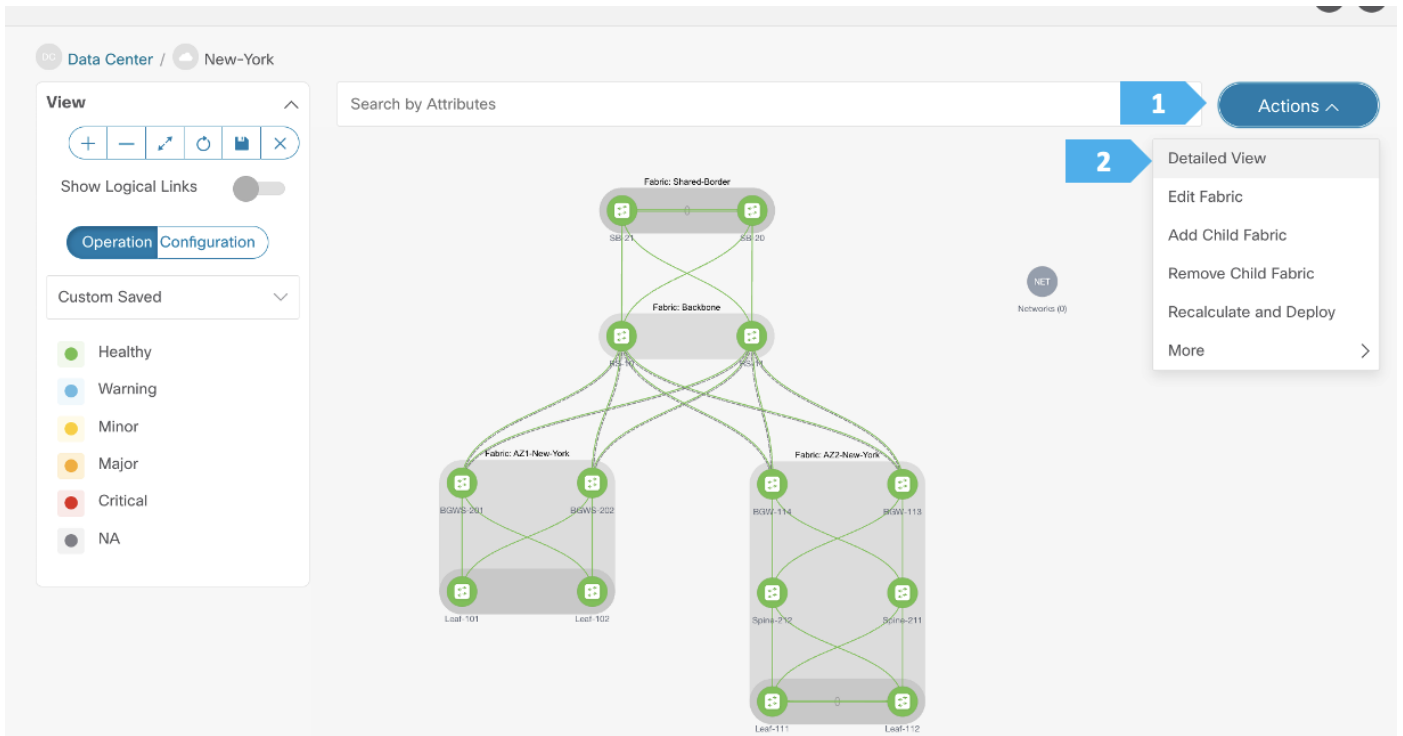
Note: For more information on Multi-Site designs and configurations, please see the following link: <https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/white-paper-c11-739942.html>

Step 4. Add the necessary policy to allow NDFC to deploy the VXLAN EVPN Multi-Site configuration on the shared border switches

By default, NDFC deploys the VXLAN EVPN Multi-Site configuration on switches with the role of border gateway or core router. NDFC does not deploy the configuration on any switch that does not have a role of border gateway or core router, even if those devices are part of the Multi-Site domain.

In this Shared-Border use case, we want to make sure that NDFC automates the VXLAN EVPN Multi-Site underlay and overlay configuration along with the rest of the devices. This step adds the necessary policy so that NDFC deploys the VXLAN EVPN Multi-Site configuration on the Shared Border switches.

Note: The Shared-Border is a normal “Border” VTEP and is independent of the VXLAN Multi-Site capabilities of BGW (Border Gateway). The configurations shown in the subsequent steps are necessary to enable EVPN Control Plane peering to receive the Type-2 and Type-5 routes from the respective BGWs.



Fabric Overview - New-York

Overview Child Fabrics Switches Links In **1** Policies Networks VRFs Event Analytics History Resources

Filter by attributes

2 Actions ^

3 type Actions ^

<input type="checkbox"/>	Policy ID	Switch	IP Address	Template	Descripti...	Entity Name	Entity Type	Source	Priority	type
<input type="checkbox"/>	POLICY-235590	BGW-113	100.64.254.113	bgp_lb_id		SWITCH	SWITCH		10	PYTHON
<input type="checkbox"/>	POLICY-244350	BGW-113	100.64.254.113	nve_lb_id		SWITCH	SWITCH		10	PYTHON
<input type="checkbox"/>	POLICY-244360	BGW-113	100.64.254.113	switch_role_s		SWITCH	SWITCH		10	PYTHON
<input type="checkbox"/>	POLICY-247050	BGW-114	100.64.254.114	nve_lb_id		SWITCH	SWITCH		10	PYTHON 9UGXZDIWIW true
<input type="checkbox"/>	POLICY-247060	BGW-114	100.64.254.114	switch_role_s		SWITCH	SWITCH		10	PYTHON 9UGXZDIWIW true

- Add Policy
- Edit Policy
- Delete Policy
- Generated Config
- Push Config



1 Switch List:



Pick a Template

Choose Template

Select Switches



Search Switches

Select All

Show Selected

Leaf-112

9WWE533TNY4

100.64.254.112

leaf

RS-10

9W9A4AM8HLH

100.64.254.10

core router

RS-11

9AB4MSSB0XQ

100.64.254.11

core router

SB-20

9K1BU3YG7MC

100.64.254.20

border

SB-21

96T9O5DS3BJ

100.64.254.21

border

Spine-211

9LE10D1ZXIZ

100.64.254.211

spine

Spine-212

915CB85DBTP

100.64.254.212

spine

1

2

Select (2)

Switch List:

SB-20 SB-21 >



Pick a Template

1

Choose Template

Select Policy Template



shared

ext_base_shared_border
N9K

1

shared_border_state
N9K

2

Select

Create Policy

Switch List: SB-20 SB-21

Priority* 500 (1-1000)

Description SB-20 and SB-21 Shared Border State

Template Name shared_border_state

1 Save

Fabric Overview - New-York

Overview Child Fabrics Switches Links Interfaces Policies Networks VRFs Event Analytics History Resources

Filter by attributes

2 Actions ^

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

Policy ID	Switch	IP Address	Template	Descripti...	Entity Name	Entity Type	Source	Priority	Content Type	Serial Number	Editabl
<input type="checkbox"/> POLICY-235590	BGW-113	100.64.254.113	bgp_lb_id		SWITCH	SWITCH		10	PYTHON	9GFG3KP30V	true
<input type="checkbox"/> POLICY-244350	BGW-113	100.64.254.113	nve_lb_id		SWITCH	SWITCH		10	PYTHON	9GFG3KP30V	true
<input type="checkbox"/> POLICY-244360	BGW-113	100.64.254.113	switch_role_s		SWITCH	SWITCH		10	PYTHON	9GFG3KP30V	true

Actions v

We can click on “Pending Config” to see the cli that will get pushed out. However, what is going to happen is:

- The RS-10 and RS-11 interfaces facing SB-201 and SB-21 will get an IP address and vice versa.
- The eBGP between RS-10 and RS-11, and between SB-20 and SB-21, will also be added.

1 Config Preview 2 Deploy Progress

Filter by attributes Resync All

Switch Name	IP Address	Role	Serial Number	Fabric Status	Pending Config	Status Description	Progress	Resync Switch
RS-11	100.64.254.11	core router	9AB4MSSB0XQ	● Out-Of-Sync	38 Lines	Out-of-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync
RS-10	100.64.254.10	core router	9W9A4AM8HLH	● Out-Of-Sync	38 Lines	Out-of-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync
SB-21	100.64.254.21	border	96T9O5DS3BJ	● Out-Of-Sync	54 Lines	Out-of-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync
SB-20	100.64.254.20	border	9K1BU3YG7MC	● Out-Of-Sync	54 Lines	Out-of-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync
BGW-113	100.64.254.113	border gateway	9GFG3KP3OV6	● In-Sync	0 Lines	In-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync
BGW-114	100.64.254.114	border gateway	9UGXZDIWIVV	● In-Sync	0 Lines	In-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync
BGWS-202	100.64.254.202	border gateway spine	9046ZFSD3G8	● In-Sync	0 Lines	In-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync
BGWS-201	100.64.254.201	border gateway spine	9AOZRKA9IY1	● In-Sync	0 Lines	In-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync

2 Deploy All

We can click on “Pending Config” to see the cli that will get pushed out, as shown in the next few screen shots.

SB-21 Sample CLI (Shared-Border)

```
route-map rmap-redirect-direct permit 10
  match tag 54321

router bgp 65004
  address-family ipv4 unicast
  redistribute direct route-map rmap-redirect-direct
  maximum-paths 64
  maximum-paths ibgp 64
  exit
  address-family ipv6 unicast
  maximum-paths 64
  maximum-paths ibgp 64
  exit
  neighbor 10.254.1.42
  remote-as 65003
  update-source Ethernet1/2
  address-family ipv4 unicast
  next-hop-self
```

```
        exit
    exit
neighbor 10.254.1.46
    remote-as 65003
    update-source Ethernet1/1
    address-family ipv4 unicast
        next-hop-self
    exit
    exit
neighbor 10.254.254.10
    remote-as 65003
    update-source loopback0
    ebgp-multihop 5
    address-family Layer 2vpn evpn
        send-community both
        rewrite-evpn-rt-asn
    exit
    exit
neighbor 10.254.254.11
    remote-as 65003
    update-source loopback0
    ebgp-multihop 5
    address-family Layer 2vpn evpn
        send-community both
        rewrite-evpn-rt-asn
configure terminal
interface ethernet1/1
    no switchport
    ip address 10.254.1.45/30 tag 54321
    mtu 9216
    no shutdown
interface ethernet1/2
    no switchport
    ip address 10.254.1.41/30 tag 54321
    mtu 9216
    no shutdown
configure terminal
```

RS-10 Sample CLI (Route Server)

```
interface ethernet1/1
    no switchport
```

```
ip address 10.254.1.37/30 tag 54321
mtu 9216
no shutdown
interface ethernet1/2
no switchport
ip address 10.254.1.42/30 tag 54321
mtu 9216
no shutdown
router bgp 65003
neighbor 10.254.1.38
remote-as 65004
update-source Ethernet1/1
address-family ipv4 unicast
next-hop-self
exit
exit
neighbor 10.254.1.41
remote-as 65004
update-source Ethernet1/2
address-family ipv4 unicast
next-hop-self
exit
exit
neighbor 10.41.0.1
remote-as 65004
inherit peer OVERLAY-PEERING
address-family Layer 2vpn evpn
rewrite-evpn-rt-asn
exit
exit
neighbor 10.41.0.2
remote-as 65004
inherit peer OVERLAY-PEERING
address-family Layer 2vpn evpn
rewrite-evpn-rt-asn
configure terminal
```

Creating vPCs, VRFs, and Networks

We will be attaching four hosts to the leaf switches as follows:

- Host-1011 and Host-1031 will be connected using a vPC in the fabric AZ1-New-York to Leaf-101 and Leaf-102

- Host-1021 and Host-1032 will be connected using a vPC in the fabric AZ2-New-York to Leaf-111 and Leaf-112

We will show how to create one vPC in the following steps. Please create the remaining vPCs based on the same procedures.

Create Interface

1 Type*
virtual Port Channel (vPC) ▾

2 Select a vPC pair*
Leaf-101---Leaf-102 ▾

3 vPC ID*
5

Policy*
[int_vpc_access_host](#) >

Policy Options

Peer-1 Port-Channel ID*
5 Peer-1 VPC port-channel number (Min:1, Max:4096)

Peer-2 Port-Channel ID*
5 Peer-2 VPC port-channel number (Min:1, Max:4096)

3 Enable Config Mirroring If enabled, Peer-1 config will be copied to Peer-2

4 Peer-1 Member Interfaces
e1/5 A list of member interfaces for Peer-1 [e.g. e1/5,eth1/7-9]

Peer-2 Member Interfaces
e1/5 A list of member interfaces for Peer-2 [e.g. e1/5,eth1/7-9]

5 Port Channel Mode*
active Channel mode options: on, active and passive

6 Save

After you have created the required vPC, perform a “Recalculate and Deploy” in each fabric. We will show how to do this for the fabric AZ1-New-York; repeat the same steps for the fabric AZ2-New-York.

Fabric Overview - AZ1-New-York

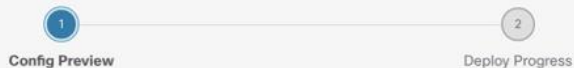
1 Actions ^

2 Recalculate and Deploy

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

Filter by attributes

<input type="checkbox"/>	Device Name	Interface	Admin Status	Oper. Status	Reason	Policies	Overlay Network	Sync Status
<input type="checkbox"/>	Leaf-102	Port-channel6		Not discovered		int_vpc_access_po_11_1	NA	● NA
<input type="checkbox"/>	Leaf-102	Port-channel5		Not discovered		int_vpc_access_po_11_1	NA	● NA
<input type="checkbox"/>	Leaf-101-Leaf-102	vPC6		Not discovered		int_vpc_access_host	NA	● NA



Filter by attributes

Resync All

Switch Name	IP Address	Role	Serial Number	Fabric Status	Pending Config	Status Description	Progress	Resync Switch
BGWS-201	100.64.254.201	border gateway spine	9AOZRKA9IY1	In-Sync	0 Lines	In-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync
BGWS-202	100.64.254.202	border gateway spine	9046ZFSD3G8	In-Sync	0 Lines	In-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync
Leaf-101	100.64.254.101	leaf	9ZEA13L749S	Out-Of-Sync	27 Lines	Out-of-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync
Leaf-102	100.64.254.102	leaf	99KJ3DPI53G	Out-Of-Sync	27 Lines	Out-of-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync

1 Deploy All

In this section, we will create vPCs, VRFs and networks. The following procedures are just an example to demonstrate the concept. Feel free to choose the VRF names and IP addresses based on our setup.

VRF CORP (Internal private networks):

- Network 192.168.101.0/24 will contain Host-1011
- Network 192.168.102.0/24 will contain Host-1021

VRF DMZ (App that requires internet or SaaS apps):

- Network 192.168.103.0/24 will contain Host-1031, Host-1032

Cisco Nexus Dashboard Fabric Controller

Fabric Controller

LAN Fabrics

Filter by attributes

Fabric Name	Fabric Technology	Fabric Type	ASN
New-York	VXLAN Fabric	Multi-Fabric Domain	NA
AZ1-New-York	VXLAN Fabric	Switch Fabric	65001
AZ2-New-York	VXLAN Fabric	Switch Fabric	65002
Backbone	External	External	65003
Shared-Border	VXLAN Fabric	Switch Fabric	65004

10 Rows

Fabric New-York

Warning

Alarms(1)

CRITICAL	MAJOR	MINOR	WARNING
0	0	0	1

Child Fabrics

- AZ1-New-York Minor
- AZ2-New-York Minor
- Backbone Minor
- Shared-Border Minor

Fabric Info

ASN: NA
Fabric Technology: VXLAN Fabric
Fabric Type: Multi-Fabric Domain
Deployment Status: Enabled

Inventory

Switch Configuration

14 Switches In-Sync (14)

Note: We can also double-click on the “New-York” MSD fabric to go directly to the next page.

Fabric Overview - New-York

Overview Child Fabrics Switches Links Interfaces **1** Networks VRFs Event Analytics History Resources

Filter by attributes

Network Name	Network ID	VRF Name	IPv4 Gateway/Suffix	IPv6 Gateway/Prefix	Network Status	Actions
No rows found						<ul style="list-style-type: none"> Create Edit Multi-Attach Multi-Detach Deploy Import Export Delete Add to Interface Group Remove from Interface Group

2

Create Network

1

Layer 2 Only

VRF Name* 5

2

3

Network Template*
[Default_Network_Universal >](#)

Network Extension Template*
[Default_Network_Extension_Universal >](#)

General Parameters **Advanced**

4 example 192.0.2.1/24

Note: We will be showing how to create VRFs from the Create Network tab. If you prefer to create VRFs first, then create VRFs from the VRF tab; the VRF will then be available to select when we create the network in this case.

Create VRF



1

VRF Name*
CORP

VRF ID*
50000

VLAN ID
2000

Propose VLAN

2

VRF Template*
[Default_VRF_Universal >](#)

VRF Extension Template*
[Default_VRF_Extension_Universal >](#)

General Parameters Advanced Route Target

3

VRF VLAN Name
CORP if > 32 chars enable:system vlan long-name

4

VRF Interface Description
Internal private networks

5

VRF Description
Internal private networks

6

Create

Create Network



Network Name*
MyNetwork_30101

Layer 2 Only

VRF Name*
CORP

Create VRF

Network ID*
30101

VLAN ID
101

Propose VLAN

Network Template*
[Default_Network_Universal >](#)

Network Extension Template*
[Default_Network_Extension_Universal >](#)

General Parameters Advanced

IPv4 Gateway/NetMask
192.168.101.254/24 example 192.0.2.1/24

IPv6 Gateway/Prefix List
example 2001:db8::1/64,2001:db9::1/64

1

Create

Fabric Overview - New-York

Actions ↕ ↻ ? — ×

Overview Child Fabrics Switches Links Interfaces Policies **Networks** VRFs Event Analytics History Resources

Filter by attributes

1 Actions ^

<input type="checkbox"/>	Network Name	Network ID	VRF Name	IPv4 Gateway/Suffix	IPv6 Gateway/Prefix	Network Status	VLAN
<input type="checkbox"/>	MyNetwork_30101	30101	CORP	192.168.101.254/24		● NA	101
<input type="checkbox"/>							
<input type="checkbox"/>							
<input type="checkbox"/>							
<input type="checkbox"/>							
<input type="checkbox"/>							
<input type="checkbox"/>							
<input type="checkbox"/>							
<input type="checkbox"/>							

- 2 Create
- Edit
- Multi-Attach
- Multi-Detach
- Deploy
- Import
- Export
- Delete
- Add to Interface Group
- Remove from Interface Group

Create Network

? — ×

1 Network Name*
MyNetwork_30102

Layer 2 Only

2 VRF Name*
CORP × ▼ Create VRF

3 Network ID*
30102

4 VLAN ID
102 Propose VLAN

Network Template*
[Default_Network_Universal >](#)

Network Extension Template*
[Default_Network_Extension_Universal >](#)

5 General Parameters Advanced
IPv4 Gateway/NetMask
192.168.102.254/24 example 192.0.2.1/24

IPv6 Gateway/Prefix List
 example 2001:db8::1/64,2001:db9::1/64

6 Create

Filter by attributes

1

Actions ↕

<input type="checkbox"/>	Network Name	Network ID	VRF Name	IPv4 Gateway/Suffix	IPv6 Gateway/Prefix	Network Status	VLAN
<input type="checkbox"/>	MyNetwork_30101	30101	CORP	192.168.101.254/24		● NA	101
<input type="checkbox"/>	MyNetwork_30102	30102	CORP	192.168.102.254/24		● NA	102

2

- Create
- Edit
- Multi-Attach
- Multi-Detach
- Deploy
- Import
- Export
- Delete
- Add to Interface Group
- Remove from Interface Group

Create Network

1

Network Name*

MyNetwork_30103

Layer 2 Only

VRF Name*

CORP



Create VRF

5

2

Network ID*

30103

3

VLAN ID

103

Propose VLAN

Network Template*

[Default_Network_Universal >](#)

Network Extension Template*

[Default_Network_Extension_Universal >](#)

4

General Parameters **Advanced**

IPv4 Gateway/NetMask

192.168.103.254/24

example 192.0.2.1/24

Create VRF



1 VRF Name*
DMZ

VRF ID*
50001

VLAN ID
2001

Propose VLAN 2

VRF Template*
Default_VRF_Universal >

VRF Extension Template*
Default_VRF_Extension_Universal >

General Parameters Advanced Route Target

3 VRF VLAN Name
DMZ # > 32 chars enable:system vlan long-name

4 VRF Interface Description
Apps that requires internet access

5 VRF Description
Apps that requires internet access

6 Create

Create Network



Network Name*
MyNetwork_30103

Layer 2 Only

VRF Name*
DMZ Create VRF

Network ID*
30103

VLAN ID
103 Propose VLAN

Network Template*
Default_Network_Universal >

Network Extension Template*
Default_Network_Extension_Universal >

General Parameters Advanced

IPv4 Gateway/NetMask
192.168.103.254/24 example 192.0.2.1/24

IPv6 Gateway/Prefix List
example 2001:db8::1/64,2001:db9::1/64

6 Create

Fabric Overview - New-York

Actions ? - X

Overview Child Fabrics Switches Links Interfaces Policies **Networks** VRFs Event Analytics History Resources

Filter by attributes

Actions

<input type="checkbox"/>	Network Name	Network ID	VRF Name	IPv4 Gateway/Suffix	IPv6 Gateway/Prefix	Network Status	VLAN ID	Interface Group
<input type="checkbox"/>	MyNetwork_30101	30101	CORP	192.168.101.254/24		● NA	101	
<input type="checkbox"/>	MyNetwork_30102	30102	CORP	192.168.102.254/24		● NA	102	
<input type="checkbox"/>	MyNetwork_30103	30103	DMZ	192.168.103.254/24		● NA	103	

Now we will start attaching networks to interfaces as per the lab setup.

Fabric Overview - New-York

Network
MyNetwork_30101

2 X

Overview Child Fabrics Switches Links Interfaces Policies **Networks** VRFs Event Analytics History Resources

Filter by attributes

<input type="checkbox"/>	Network Name	Network ID	VRF Name	IPv4 Gateway/Suffix	IPv6 Gateway/Prefix	Network Status
1 <input type="checkbox"/>	MyNetwork_30101	30101	CORP	192.168.101.254/24		● NA
<input type="checkbox"/>	MyNetwork_30102	30102	CORP	192.168.102.254/24		● NA
<input type="checkbox"/>	MyNetwork_30103	30103	DMZ	192.168.103.254/24		● NA

Network Info

Network ID 30101	VRF Name CORP
IPv4 Gateway 192.168.101.254/24	IPv6 Gateway NA
Status ● NA	VLAN ID 101
Network Template Default_Network_Univers al	Network Extension Template Default_Network_Extensi on_Universal
Interface Group NA	Mcast Group NA

Networks

Network Status



Switch Roles Association



50 Rows

1 Network Attachments

Filter by attributes

3 Actions ^

<input type="checkbox"/>	Network Name	Network ID	VLAN ID	Switch	Ports	Status	Attachment	Switch Role	
<input type="checkbox"/>	MyNetwork_30101	30101		BGWS-202	NA	● NA	Detached	border gatew spine	
<input checked="" type="checkbox"/>	MyNetwork_30101	30101		Leaf-101	NA	● NA	Detached	leaf	
<input type="checkbox"/>	MyNetwork_30101	30101		Leaf-102	NA	● NA	Detached	leaf	
<input type="checkbox"/>	MyNetwork_30101	30101		BGWS-201	NA	● NA	Detached	border gateway spine	
<input type="checkbox"/>	MyNetwork_30101	30101		Leaf-111	NA	● NA	Detached	leaf	
<input type="checkbox"/>	MyNetwork_30101	30101		Leaf-112	NA	● NA	Detached	leaf	AZ2-New-York
<input type="checkbox"/>	MyNetwork_30101	30101		BGW-113	NA	● NA	Detached	border gateway	AZ2-New-York
<input type="checkbox"/>	MyNetwork_30101	30101		BGW-114	NA	● NA	Detached	border gateway	AZ2-New-York
<input type="checkbox"/>	MyNetwork_30101	30101		SB-21	NA	● NA	Detached	border	Shared-Border
<input type="checkbox"/>	MyNetwork_30101	30101		SB-20	NA	● NA	Detached	border	Shared-Border

- History
- 4 Edit
- Preview
- Deploy
- Import
- Export
- Quick Attach
- Quick Detach

Note: We can also double-click on “MyNetwork_30101” to go directly to the next page.

Leaf-111 (9XYGQULY6H4) - Leaf-112 (9WWE533TNY4)

Detach Attach 1

VLAN*

'Interface Attachment(s)'

Filter by attributes

<input type="checkbox"/>	Interface/Ports	Switch	Status	Port Type	Port Description	Neighbor Info
<input checked="" type="checkbox"/>	Port-channel5	Leaf-111	false	access		
<input type="checkbox"/>	Port-channel6	Leaf-111	false	access		
<input checked="" type="checkbox"/>	Port-channel5	Leaf-112	false	access		
<input type="checkbox"/>	Port-channel6	Leaf-112	false	access		
<input type="checkbox"/>	Ethernet1/7	Leaf-111	false	trunk		
<input type="checkbox"/>	Ethernet1/7	Leaf-112	false	trunk		
<input type="checkbox"/>	Ethernet1/8	Leaf-111	false	trunk		
<input type="checkbox"/>	Ethernet1/8	Leaf-112	false	trunk		
<input type="checkbox"/>	Ethernet1/9	Leaf-111	false	trunk		
<input type="checkbox"/>	Ethernet1/9	Leaf-112	false	trunk		
<input type="checkbox"/>	Ethernet1/10	Leaf-111	false	trunk		
<input type="checkbox"/>	Ethernet1/10	Leaf-112	false	trunk		
<input type="checkbox"/>	Ethernet1/11	Leaf-111	false	trunk		
<input type="checkbox"/>	Ethernet1/11	Leaf-112	false	trunk		

4 Save

Fabric Overview - New-York

Overview Child Fabrics Switches Links Interfaces Policies **1** VRFs Event Analytics History Resources

Filter by attributes

<input type="checkbox"/>	VRF Name	VRF Status	VRF ID
<input type="checkbox"/>	DMZ	DEPLOYED	50001
2 <input checked="" type="checkbox"/>	CORP	DEPLOYED	50000

50 Rows

VRF CORP **3**

VRF Info

VRF ID: 50000 | VLAN ID: 2000

VRF Template: Default_VRF_Universal | VRF Extension Template: Default_VRF_Extension_Universal

Status: DEPLOYED | VRF Description: Internal private networks

L3VniMcastGroup: NA

VRFs

VRF Status: 10 (4 DEPLOYED, 6 NA)

Switch Roles Association: 10 (4 leaf, 2 border gateway, 2 border gateway spine)

Note: We can double-click on the VRF CORP to go to the next page.

VRF Overview - CORP

VRF Attachments Networks

Filter by attributes **2** Actions

<input type="checkbox"/>	VRF Name	VRF ID	VLAN ID	Switch	Status	Attachment	Switch Role	Fabric Name	Loopback	History
<input checked="" type="checkbox"/>	CORP	50000		BGW-113	NA	Detached	border gateway	AZ2-New-York		Edit
<input checked="" type="checkbox"/>	CORP	50000		BGW-114	NA	Detached	border gateway	AZ2-New-York		Preview
<input checked="" type="checkbox"/>	CORP	50000		BGWS-201	NA	Detached	border gateway spine	AZ1-New-York		Deploy
<input checked="" type="checkbox"/>	CORP	50000		BGWS-202	NA	Detached	border gateway spine	AZ1-New-York		Import
<input type="checkbox"/>	CORP	50000	2000	Leaf-101	DEPLOYED	Attached	leaf	AZ1-New-York		Export
<input type="checkbox"/>	CORP	50000	2000	Leaf-102	DEPLOYED	Attached	leaf	AZ1-New-York		Quick Attach
<input type="checkbox"/>	CORP	50000	2000	Leaf-111	DEPLOYED	Attached	leaf	AZ2-New-York		Quick Detach
<input type="checkbox"/>	CORP	50000	2000	Leaf-112	DEPLOYED	Attached	leaf	AZ2-New-York		

50 Rows Page 1 of 1 1-10 of 10

VRF Overview - CORP

Actions  

Overview **VRF Attachments** Networks

Filter by attributes

2 Actions ^

<input type="checkbox"/>	VRF Name	VRF ID	VLAN ID	Switch	Status	Attachment	Switch Role	Fabric Name	Loopback	History
<input checked="" type="checkbox"/>	CORP	50000	2000	BGW-113	PENDING	Attached	border gateway	AZ2-New-York		Edit
<input checked="" type="checkbox"/>	CORP	50000	2000	BGW-114	PENDING	Attached	border gateway	AZ2-New-York		Preview
<input checked="" type="checkbox"/>	CORP	50000	2000	BGWS-201	PENDING	Attached	border gateway spine	AZ1-New-York		Deploy
<input checked="" type="checkbox"/>	CORP	50000	2000	BGWS-202	PENDING	Attached	border gateway spine	AZ1-New-York		Import
<input type="checkbox"/>	CORP	50000	2000	Leaf-101	DEPLOYED	Attached	leaf	AZ1-New-York		Export
<input type="checkbox"/>	CORP	50000	2000	Leaf-102	DEPLOYED	Attached	leaf	AZ1-New-York		Quick Attach
<input type="checkbox"/>	CORP	50000	2000	Leaf-111	DEPLOYED	Attached	leaf	AZ2-New-York		Quick Detach
<input type="checkbox"/>	CORP	50000	2000	Leaf-112	DEPLOYED	Attached	leaf	AZ2-New-York		

50 Rows

Page 1 of 1 << < 1-10 of 10 > >>

For the VRF DMZ, Host-1031 and Host-1032 are in same VLAN, so to be able to extend Layer 2 and send/receive TYPE-2 routes, we need to attach the network 192.168.103/24 to the BGWs in the AZ1-New-York and AZ2-New-York fabrics.

Filter by attributes

3 Actions ^

2

<input type="checkbox"/>	Network Name	Network ID	VLAN ID	Switch	Ports	Status	Attachment	Switch Role	
<input checked="" type="checkbox"/>	MyNetwork_30103	30103	103	BGWS-202	NA	PENDING	Attached	border gateway spine	History Edit Preview Deploy Import Export Quick Attach Quick Detach
<input checked="" type="checkbox"/>	MyNetwork_30103	30103	103	BGW-113	NA	PENDING	Attached	border gate	
<input checked="" type="checkbox"/>	MyNetwork_30103	30103	103	BGW-114	NA	PENDING	Attached	border gateway	
<input checked="" type="checkbox"/>	MyNetwork_30103	30103	103	BGWS-201	NA	PENDING	Attached	border gateway spine	
<input type="checkbox"/>	MyNetwork_30103	30103		SB-21	NA	NA	Detached	border	Shared-Border
<input type="checkbox"/>	MyNetwork_30103	30103		SB-20	NA	NA	Detached	border	
<input type="checkbox"/>	MyNetwork_30103	30103	103	Leaf-111	Port-channel6	DEPLOYED	Attached	leaf	AZ2-New-York
<input type="checkbox"/>	MyNetwork_30103	30103	103	Leaf-112	Port-channel6	DEPLOYED	Attached	leaf	AZ2-New-York
<input type="checkbox"/>	MyNetwork_30103	30103	103	Leaf-101	Port-channel6	DEPLOYED	Attached	leaf	AZ1-New-York
<input type="checkbox"/>	MyNetwork_30103	30103	103	Leaf-102	Port-channel6	DEPLOYED	Attached	leaf	AZ1-New-York

4

50 Rows

Page 1 of 1 << >> 1-10 of 10 >>>

The following diagram shows a VXLAN EVPN fabric with two legitimate VTEPs and one rogue VTEP. An endpoint connected to VTEP-1 is being advertised as EVPN Type-2 (MAC and IP) to a remote VTEP. Therefore, VTEP-2 updates its NVE Peer table by listing the VTEP-1 IP address as a legitimate peer IP. At the same time, the rogue VTEP is trying to establish a VXLAN data plane tunnel towards VTEP-2, but as VTEP-2 does not recognize the VTEP-3 IP address in the NVE Peer list, it will drop the traffic. Hence, in the Nexus 9000 VTEP, we implement the SRC_TEP_MISS check for validating data plane security. This prevents the insertion of rogue VTEPs in a VXLAN EVPN fabric.

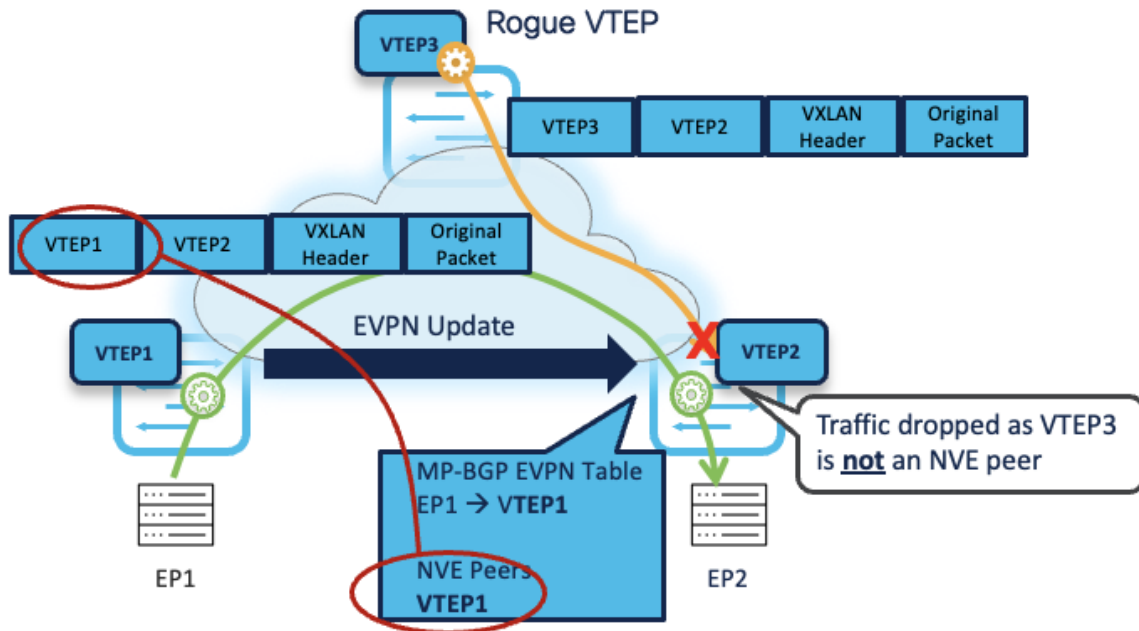


Figure 6. Native VXLAN Data Plane Security

Now, let's see what special considerations we must ensure to comply with the above implementation.

Inter-Site Layer 3 Traffic- Control Plane

By default, the Inter-Site EVPN Type-2 (MAC only, MAC + IP) and Type-5 (Prefix) updates always carry the local Multi-Site VIP as the Next-Hop address. The exceptions are Type-5 updates for Layer 3 networks and prefixes locally connected to the BGWs.

It is important to note that Multi-Site VIP (Virtual IP address) only applies to devices running with Border Gateway (BGW) roles. Hence, the Shared-Border (Border role) does not carry the Multi-Site VIP. It instead uses a regular VTEP IP (typically Primary Lo1 in the case of a standalone Border, or Shared vPC secondary IP in the case of a vPC Border).

The following diagram shows that from the Control Plane perspective, the BGWs act as the next hop, and the appropriate NVE Peer IP list is also updated with the Multi-Site VIP.

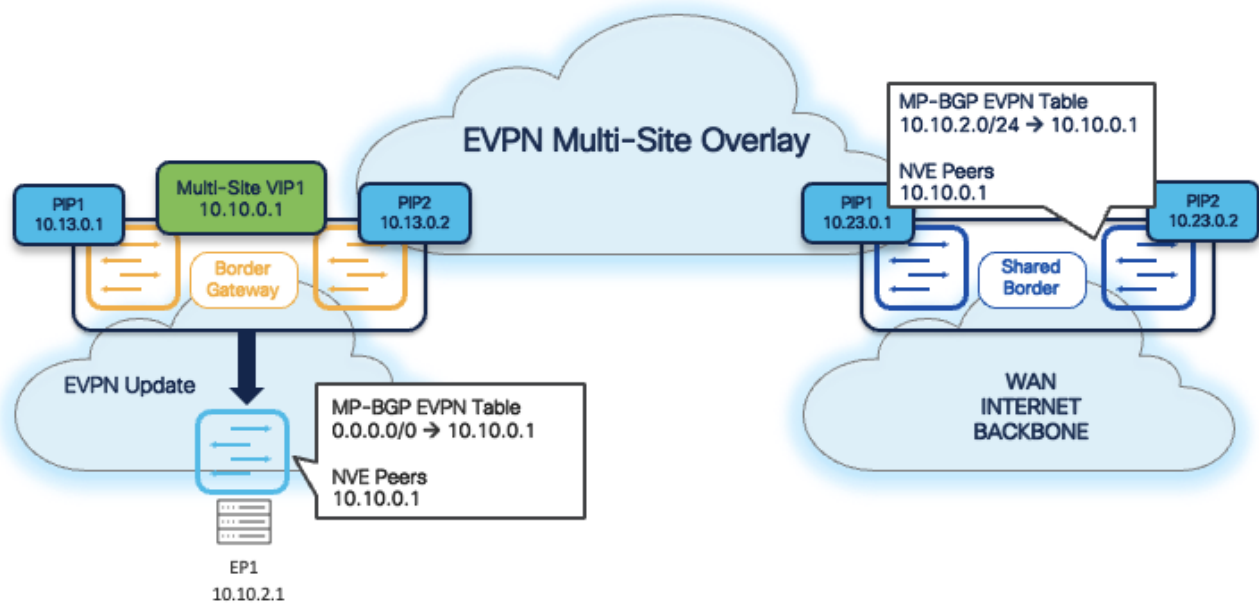


Figure 7. Layer 3 Traffic - Control Plane

Inter-Site Layer 3 Traffic- Data Plane

From a data plane point of view, the Inter-Site Layer 3 traffic is always sourced by the local BGWs using their specific PIP address. This also applies to Shared-Border architecture, where we extend Layer 3 services on the Border devices for North-to-South traffic.

As shown in the following diagram, the BGWs use the Outer SRC IP of the NVE IP address, while, by default, the Shared-Borders will only learn and form NVE Peering with the Multi-Site VIP of the BGWs. Hence, if a VXLAN packet comes to the Shared-Border with an Outer SRC IP address of the BGW, the packet will be dropped due to the SRC_TEP_MISS check.

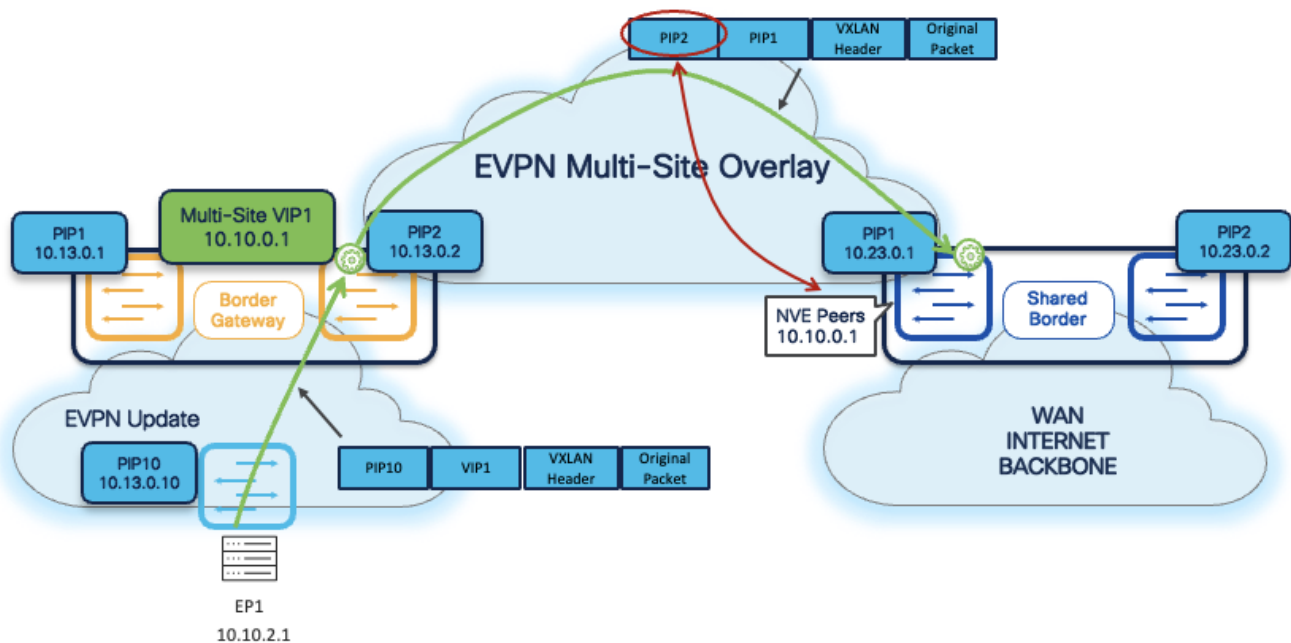


Figure 8. Layer 3 Traffic - Data Plane

To address the above situation in a Shared-Border architecture and deployments, the most common approach is to define and advertise a loopback in a Tenant-VRF (VXLAN VRF L3VNI) on every BGW, and then advertise to Shared-Borders as part of BGP EVPN updates. Once the EVPN update arrives at the Shared-Border, it will form the NVE Peering with the BGW Primary IP address.

Starting with NDFC release 12.1.3b, a new feature is introduced to simplify the configuration to address the special handling of Layer 3 communication between BGWs and Shared-Border. Following are the steps required to enable this feature:

- Step 1. Navigate to the respective Data Center VXLAN EVPN fabric settings.
- Step 2. Under the Resources tab, enable the flag for **Per VRF Per VTEP Loopback Auto-Provisioning**. Once the flag is enabled, NDFC proposes the IP subnet pool.
- Step 3. Save the fabric setting and perform a **Recalculate and Deploy**.
- Step 4. Navigate to VRF Attachments and select the VRF.
- Step 5. Click **Actions > Quick Attach**.
- Step 6. Click **Deploy**.

Note: The VRF is already attached and deployed on the BGWs, but you must perform Quick Attach one more time for the Resource Manager to assign and allocate unique IP addresses on the devices.

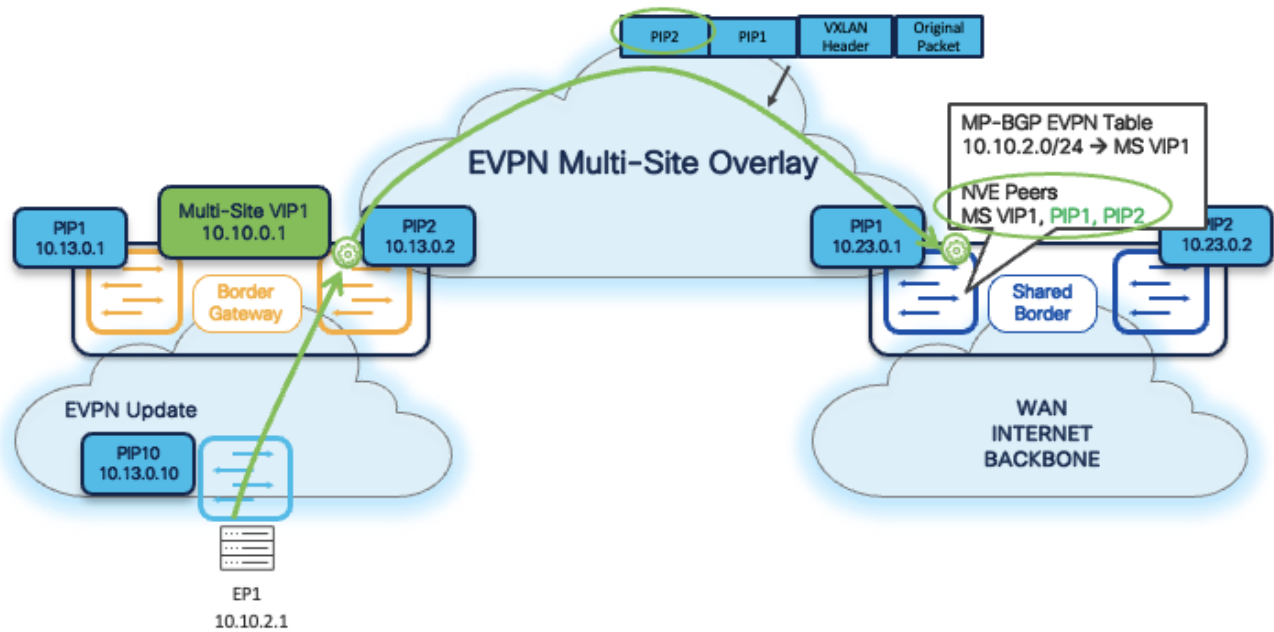


Figure 9. Installing PIP Addresses as NVE Peers

External Network IP Handoff Use-Cases

Typically, workloads often require communication with services outside of the Data Center domain in a Data Center fabric. This also includes users accessing applications and services from the Internet and WAN. The VXLAN EVPN Border devices are considered a handoff point for North-to-South communication. The Shared Border is a Site External VTEP to perform VXLAN EVPN to IP handoff. The Shared Border optimizes the traffic flows and reduces natural traffic hair-pinning. Also, it provides a deterministic handoff point in a VXLAN Multi-Site environment where multiple sites can rely on these Shared Borders to communicate with the External network such as WAN, Backbone, and Internet.

From the connectivity point of view, the Shared Borders supports Inter-AS option A (VRF Lite) and seamless VXLAN-MPLS gateway (Border-PE). Thus, a network admin can adopt different options based on the overall scale, configuration management, and operations.

Furthermore, NDFC fully supports the VRF Lite provisioning of Nexus and non-Nexus devices using a single plane of glass solution.

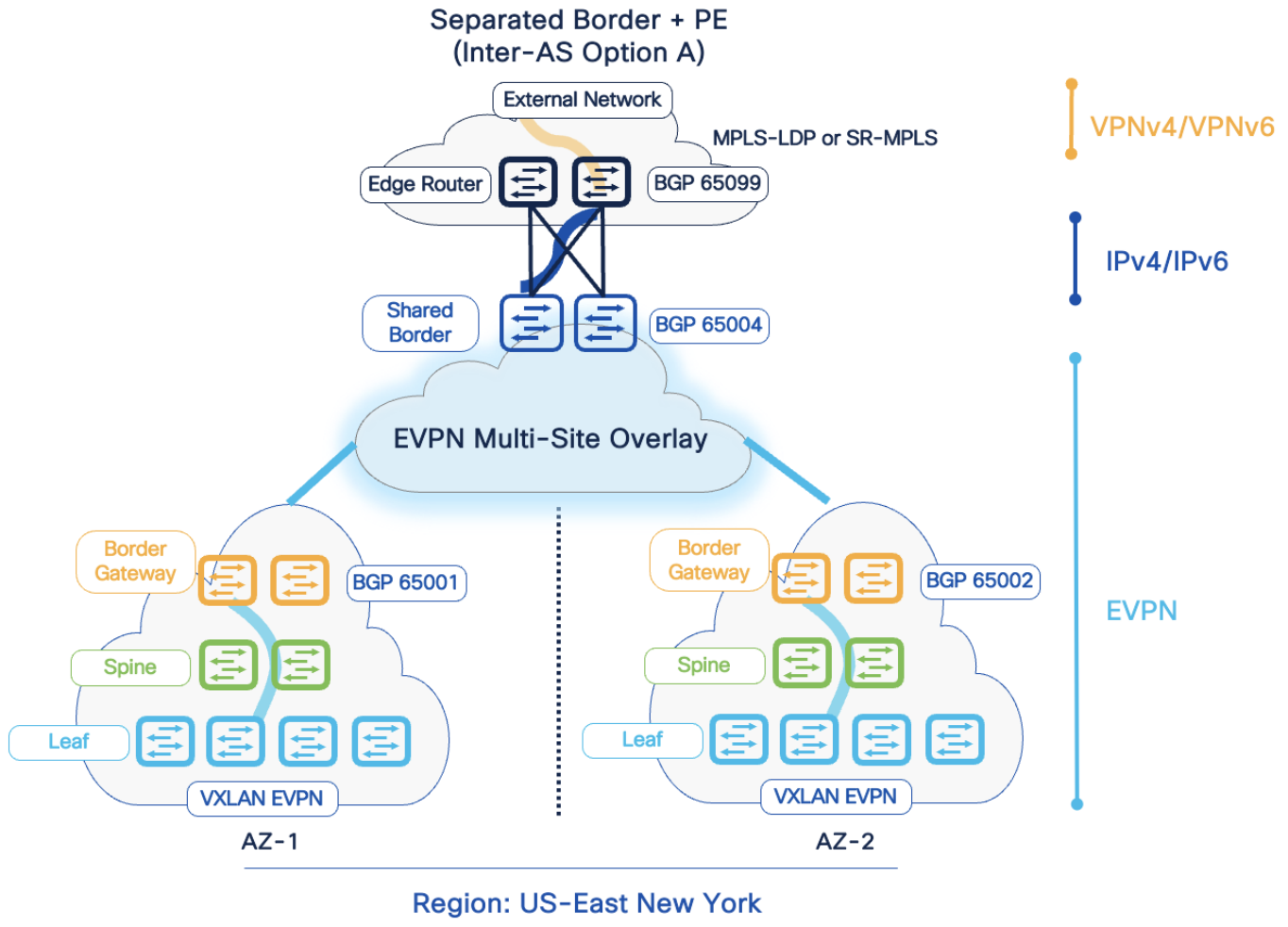


Figure 10. IP handoff using VRF Lite (2-box solution)

VRF-LITE Handoff

-
- ✓ 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**
 - 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
 - ✓ **Peering Type = L2 Trunk Interfaces with SVIs**
 - Physical interface configured as 802.1Q trunk port
 - SVI peers per VRF basis
 - SVI serves as next hop routing

Figure 11. Considerations for VRF Lite Handoff

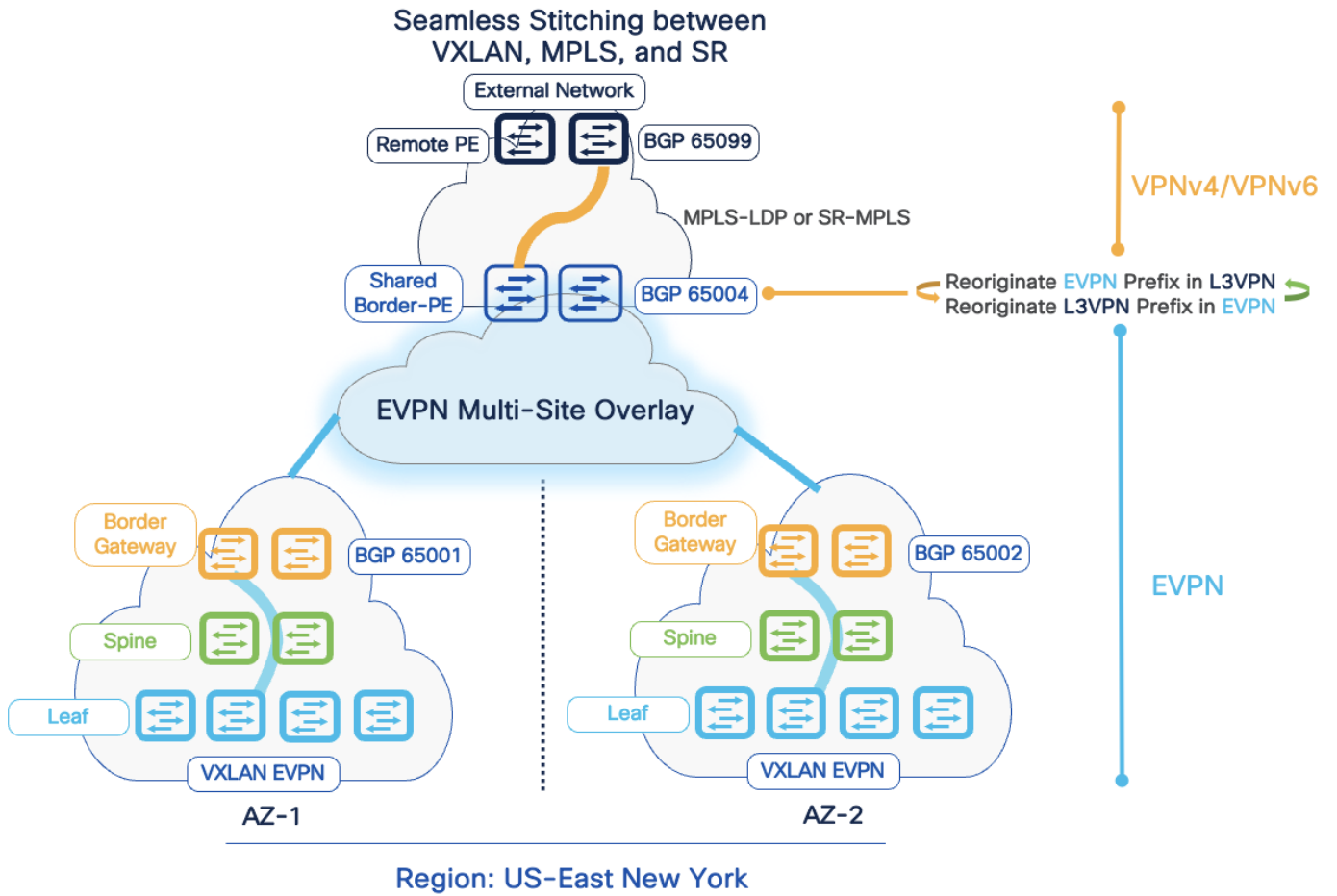


Figure 12. IP handoff using VPN (Single-box solution)

VPN Handoff

- ✓ 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

Figure 13. Considerations for VPN Handoff

Edge Router Placement Option

Nexus and Non-Nexus support

TIP Supported Hardware and Software might vary depending on NDFC version
[Check compatibility matrix 12.1.3b](#)

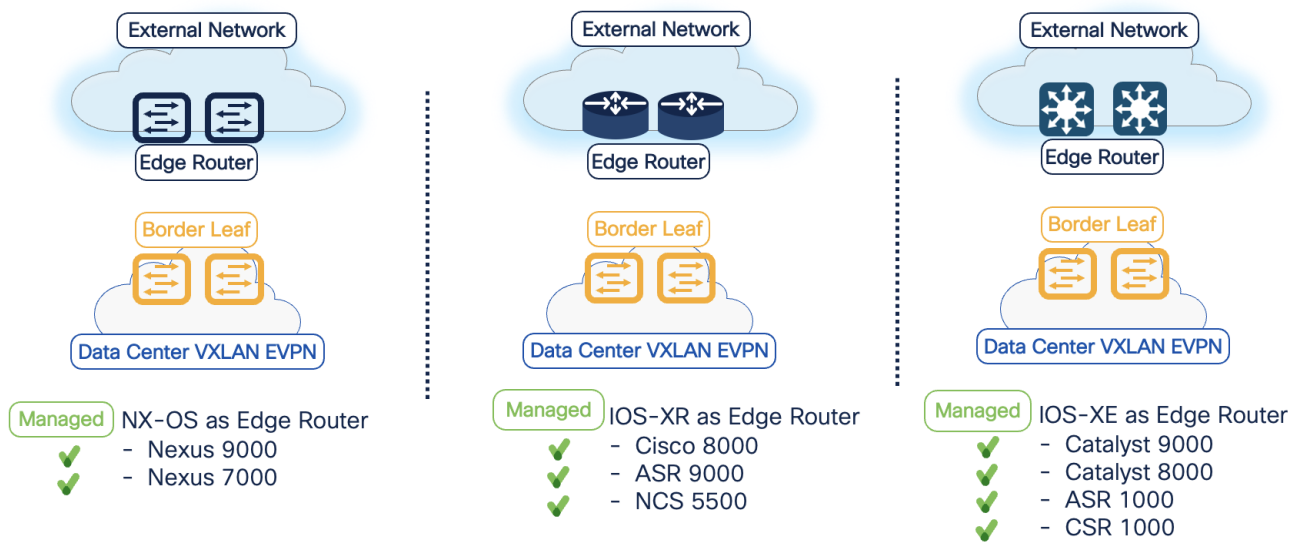


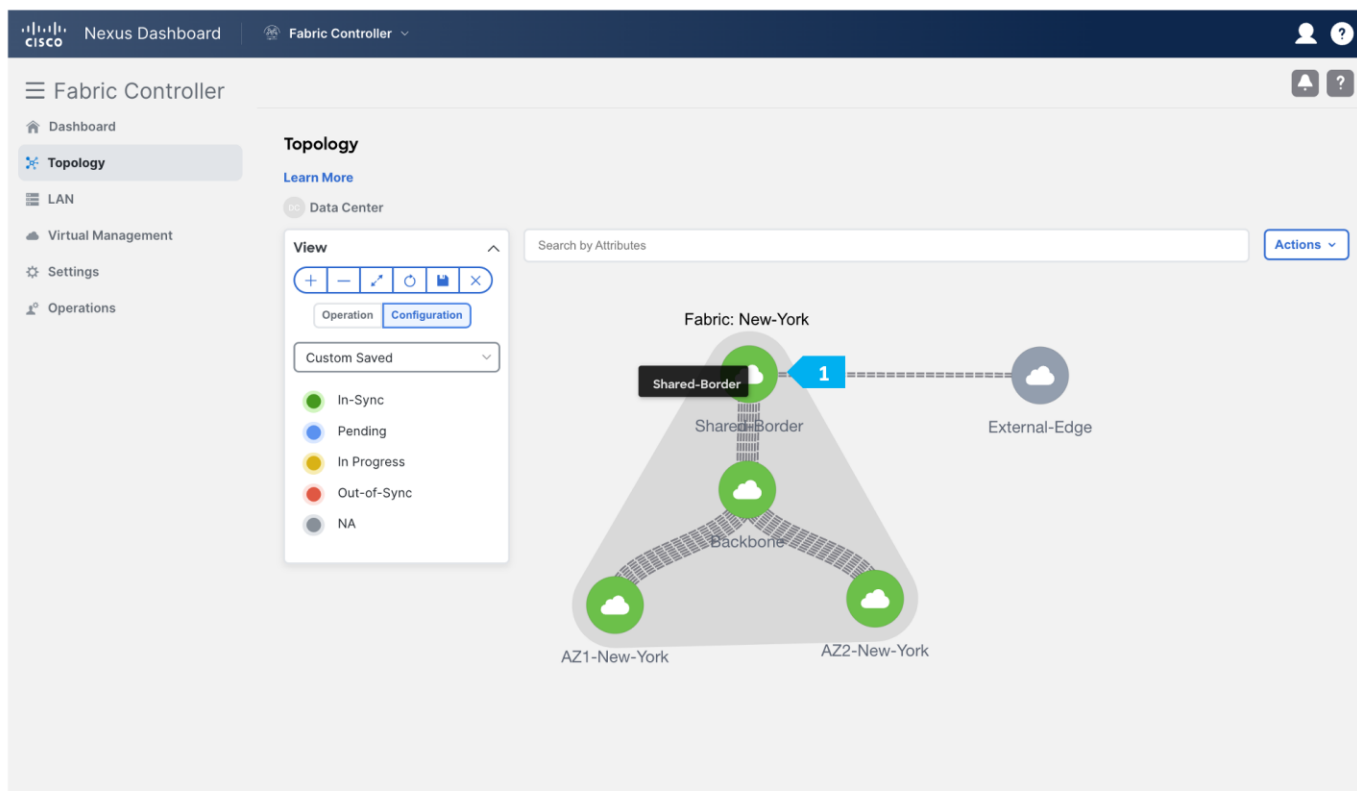
Figure 14. Edge Router options with NDFC

VRF Lite to External Edge Router

Step 1. Change the VRF Lite settings in Shared-Border Fabric

In this step, we will specify the VRF Lite method for extending inter fabric connections. The VRF Lite Subnet IP Range field specifies resources reserved for IP addresses used for VRF Lite when VRF Lite IFCs are auto created. When we select Back2Back&ToExternal, then VRF Lite IFCs are automatically created.

The **Auto Deploy for Peer** check box is applicable for VRF Lite deployments. When we select this checkbox, auto-created VRF Lite IFCs will have the Auto Generate Configuration for Peer field in the VRF Lite tab set.



Nexus Dashboard | Fabric Controller

Fabric Controller

- Dashboard
- Topology**
- LAN
- Virtual Management
- Settings
- Operations

Topology
Learn More

Data Center / New-York / Shared-Border

View

Search by Attributes

1 Actions

2

- Detailed View
- Edit Fabric
- Add Switches
- Recalculate and Deploy
- More

External-Edge

SB-20 SB-21

NET

Networks (5) VRFs (2)

Backbone

View controls: Show Logical Links, Operation/Configuration, Custom Saved, Legend (In-Sync, Pending, In Progress, Out-of-Sync, NA), Multi-select.

Edit Fabric : Shared-Border

? — X

Fabric Name: Shared-Border

Pick Fabric: Data Center VXLAN EVPN >

General Parameters | Replication | vPC | Protocols | **1 Resources** | Manageability | Bootstrap | Configuration Backup | Flow Monitor

Manual Underlay IP Address Allocation Checking this will disable Dynamic Underlay IP Address Allocations

Underlay Routing Loopback IP Range* Typically Loopback0 IP Address Range

Underlay VTEP Loopback IP Range* Typically Loopback1 IP Address Range

Underlay RP Loopback IP Range* Anycast or Phantom RP IP Address Range

Underlay Subnet IP Range* Address range to assign Numbered and Peer Link SVI IPs

Underlay MPLS Loopback IP Range Used for VXLAN to MPLS SR/LDP Handoff

Underlay Routing Loopback IPv6 Range Typically Loopback0 IPv6 Address Range

Underlay VTEP Loopback IPv6 Range

Close Save

Edit Fabric : Shared-Border

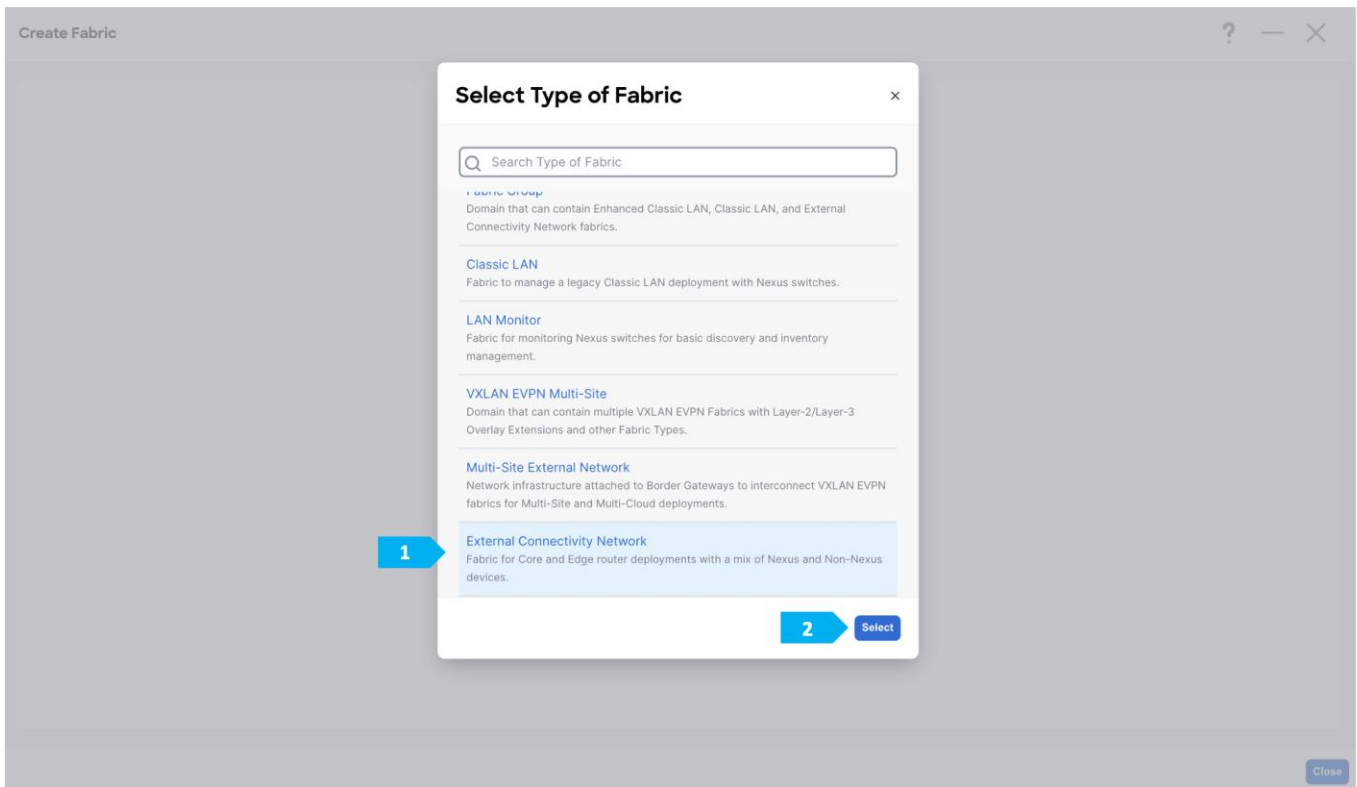
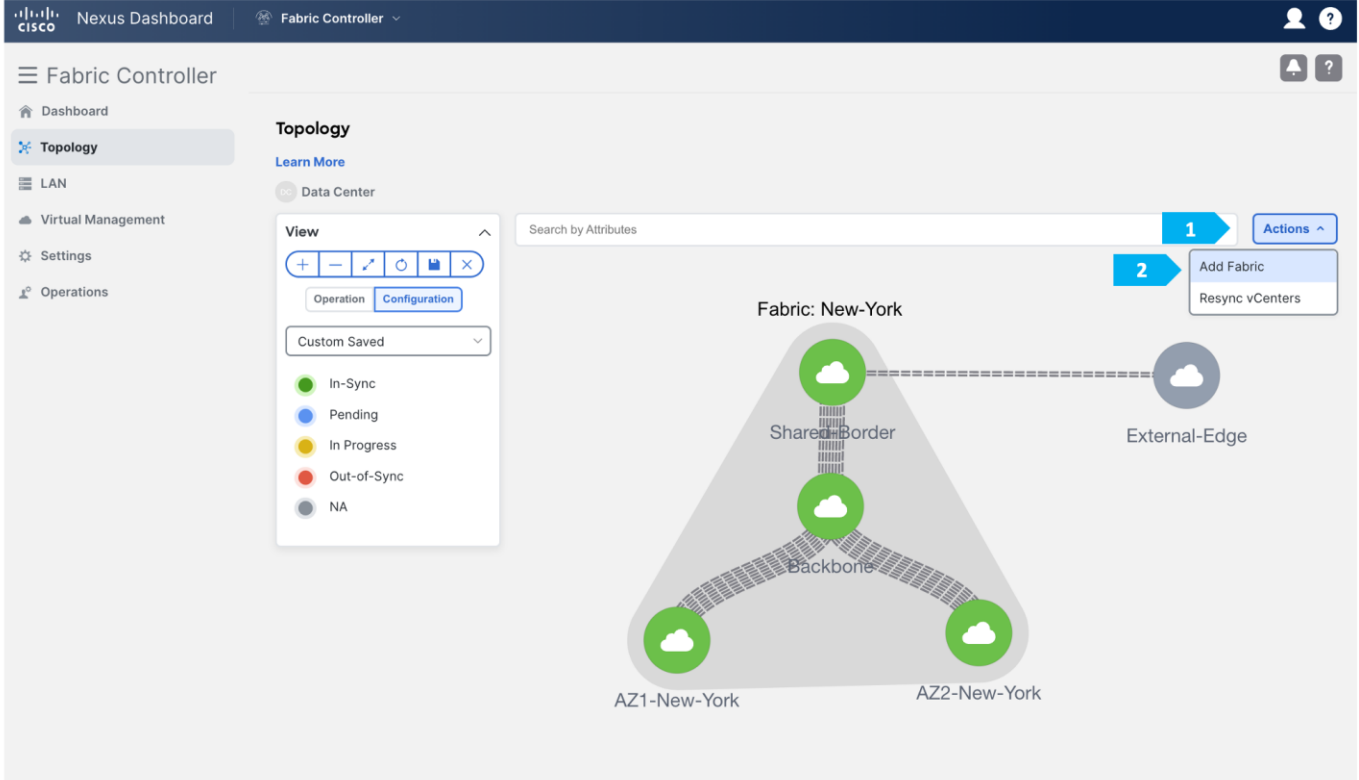


Layer 2 VXLAN VNI Range*	30000-49000	Overlay Network Identifier Range (Min:1, Max:16777214)
Layer 3 VXLAN VNI Range*	50000-59000	Overlay VRF Identifier Range (Min:1, Max:16777214)
Network VLAN Range*	2300-2999	Per Switch Overlay Network VLAN Range (Min:2, Max:4094)
VRF VLAN Range*	2000-2299	Per Switch 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 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	<input checked="" type="checkbox"/>	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.
Auto Deploy Default VRF	<input type="checkbox"/>	Whether to auto generate Default VRF interface and BGP peering configuration on VRF LITE IFC auto deployment. If set, auto created VRF Lite IFC links will have 'Auto Deploy Default VRF' enabled.
Auto Deploy Default VRF for Peer	<input type="checkbox"/>	Whether to auto generate Default VRF interface and BGP peering configuration on managed neighbor devices. If set, auto created VRF Lite IFC links will have 'Auto Deploy Default VRF for Peer' enabled.

Close Save 3

The screenshot shows the Cisco Nexus Dashboard Fabric Controller interface. The main view is the 'Topology' section, displaying a network diagram with a 'Backbone' node and several 'NET' (Network) and 'VRF' (Virtual Routing and Forwarding) nodes. A warning dialog box is overlaid on the topology, with a yellow warning icon and the text: 'Warning: Please perform "Recalculate and Deploy", if there are any switches in the fabric prior to "Deploy"'. The dialog box has an 'OK' button and a blue arrow icon with the number '1' next to it. The interface includes a sidebar with navigation options like 'Dashboard', 'Topology', 'LAN', 'Virtual Management', 'Settings', and 'Operations'. The top navigation bar shows 'Nexus Dashboard' and 'Fabric Controller'.

Step 2. Creating the External fabric and adding the devices



Create Fabric

? — ✕

Fabric Name

External-Core

Pick Fabric

[External Connectivity Network >](#)

1 General Parameters Advanced Resources Configuration Backup Bootstrap Flow Monitor

BGP AS #*

65099

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

Fabric Monitor Mode

If enabled, fabric is only monitored. No configuration will be deployed

Enable Performance Monitoring (For NX-OS and IOS XE Switches Only)

Close Save

3

After clicking Save, double-click on the “External-Core” fabric.

The screenshot shows the Cisco Fabric Controller interface. The top navigation bar includes "Nexus Dashboard" and "Fabric Controller". The left sidebar lists navigation options: Dashboard, Topology (selected), LAN, Virtual Management, Settings, and Operations. The main content area is titled "Topology" and includes a "Learn More" link and a "Data Center" filter. A "View" panel on the left shows zoom controls and a "Custom Saved" dropdown. The main topology diagram, titled "Fabric: New-York", shows a central "Shared Border" and "Backbone" area with three green cloud icons. To the left is an "External-Core" icon with a blue arrow labeled "1" pointing to it. To the right is an "External-Edge" icon. At the bottom are "AZ1-New-York" and "AZ2-New-York" icons. A search bar and "Actions" dropdown are at the top right of the diagram area.

CISCO Nexus Dashboard Fabric Controller

Fabric Controller

Dashboard

Topology

LAN

Virtual Management

Settings

Operations

Topology

Learn More

Data Center / External-Core

View

Search by Attributes

1 Actions

2

- Detailed View
- Edit Fabric
- Add Switches
- Recalculate and Deploy
- More

Show Logical Links

Operation Configuration

Custom Saved

- In-Sync
- Pending
- In Progress
- Out-of-Sync
- NA

Multi-select

0 selected

Add Switches - Fabric: External-Core

Switch Addition Mechanism*

Discover Move Neighbor Switches

Seed Switch Details

1 Seed IP*

100.64.254.31

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

Authentication Protocol*

MDS

Device Type*

NX-OS

2 Username*

admin

3 Password*

Set as individual device write credential

Max Hops*

2

Close Discover Switches 4

Switch Addition Mechanism*

Discover Move Neighbor Switches

Seed Switch Details

Fabric	Switch	Authentication Protocol	Username
External-Core	100.64.254.31	MD5	admin
Password	Max Hops	Preserve config	
<input type="checkbox"/> Set as individual device write credential	2	● Enabled	

[← Back](#)

Discovery Results

Filter by attributes

Switch Name	Serial Number	IP Address	Model	Version	Status	Progress
<input checked="" type="checkbox"/> Core-31	9IKJ1WYRPGM	100.64.254.31	N9K-C9300v	10.3(3)	● Manageable	
<input type="checkbox"/> SB-21	9DNMFIOD0IX	100.64.254.21	N9K-C9300v	10.3(3)	● Already Managed In Shared-	
<input type="checkbox"/> SB-20	9NL6XB9DR3X	100.64.254.20	N9K-C9300v	10.3(3)	● Already Managed In Shared-	
<input type="checkbox"/> RS-10	9HPQ7WQ5H60	100.64.254.10	N9K-C9300v	10.3(3)	● Already Managed In Backbo-	
<input type="checkbox"/> RS-11	9QJ5I4EIVPR	100.64.254.11	N9K-C9300v	10.3(3)	● Already Managed In Backbo-	

[Close](#) [Add Switches](#) **2**

After clicking “Add Switches”, wait until the progress bar is green, then click Close.

We will be using a Nexus device for this configuration, so right-click on “Core-31” and choose **Set Role**, then select “Edge Router” as the role for this device.

Nexus Dashboard Fabric Controller

Fabric Controller

Dashboard Topology LAN Virtual Management Settings Operations

Topology

Learn More

Data Center / External-Core

Search by Attributes [Actions](#)

View

+ - [Icons]

Show Logical Links

Operation Configuration

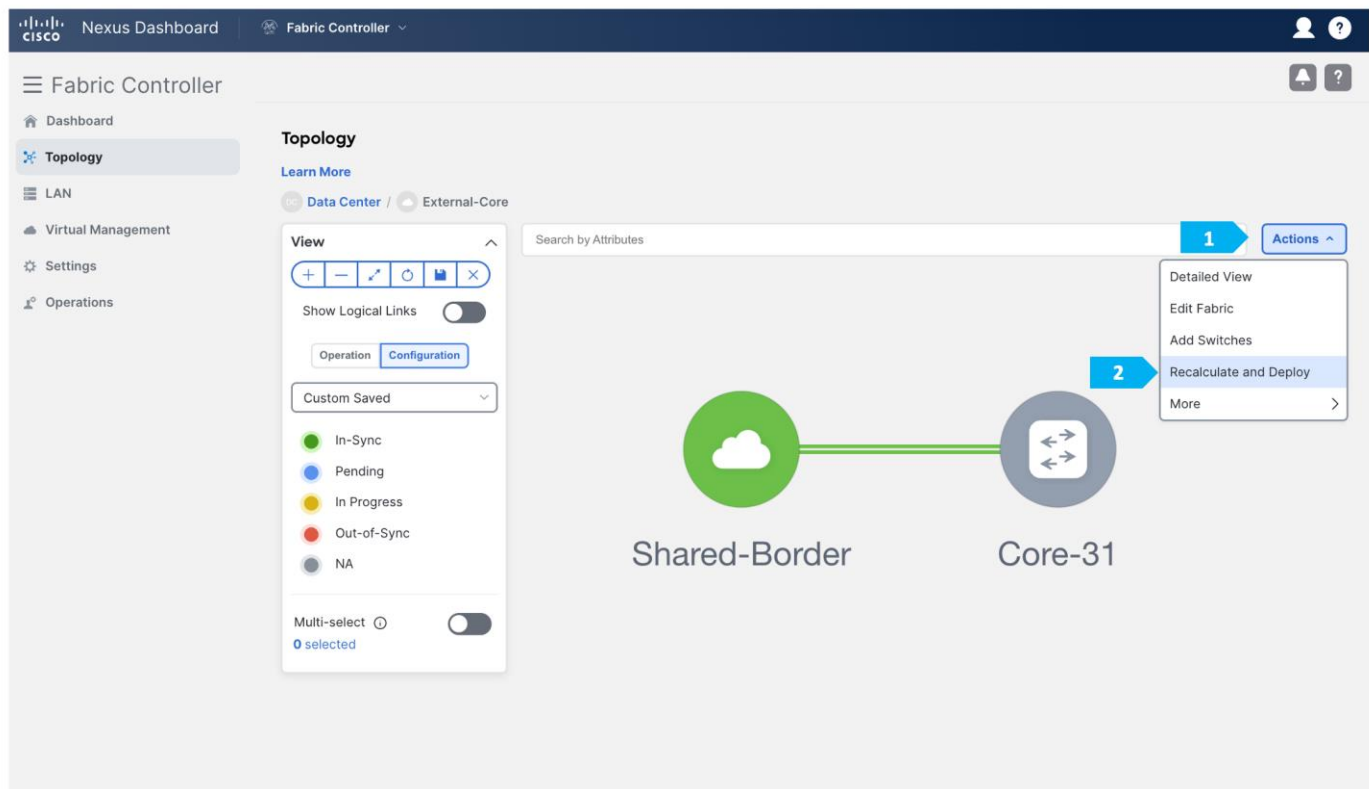
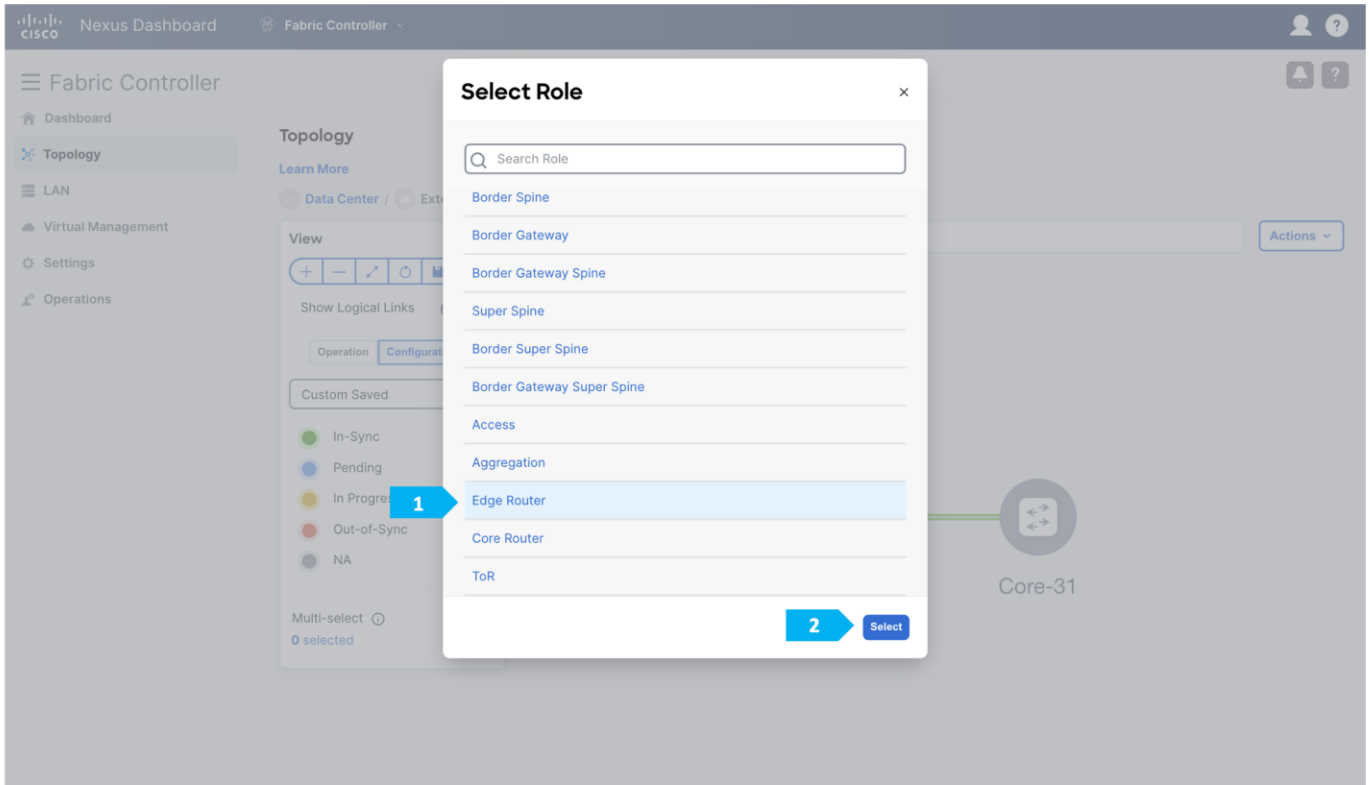
Custom Saved

- In-Sync
- Pending
- In Progress
- Out-of-Sync
- NA

Multi-select 0 selected

Shared-Border Core-31

- Core-31
- Detailed View
- Preview Config
- Deploy Config
- Discovery
- Set Role**
- vPC Pairing
- More



1 Config Preview 2 Deploy Progress

Filter by attributes Resync All

Switch Name	IP Address	Role	Serial Number	Fabric Status	Pending Config	Status Description	Progress	Resync Switch
Core-31	100.64.254.31	edge router	9IKJ1WYRPGM	Out-Of-Sync	5 Lines	Out-of-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync

Close Deploy All 1

Step 3. Checking links between Shared-Borders and Edge Router

Nexus Dashboard Fabric Controller

Fabric Controller

Dashboard Topology LAN Virtual Management Settings Operations

Topology Learn More

Data Center / External-Core

Search by Attributes Actions

View

Show Logical Links

Operation Configuration

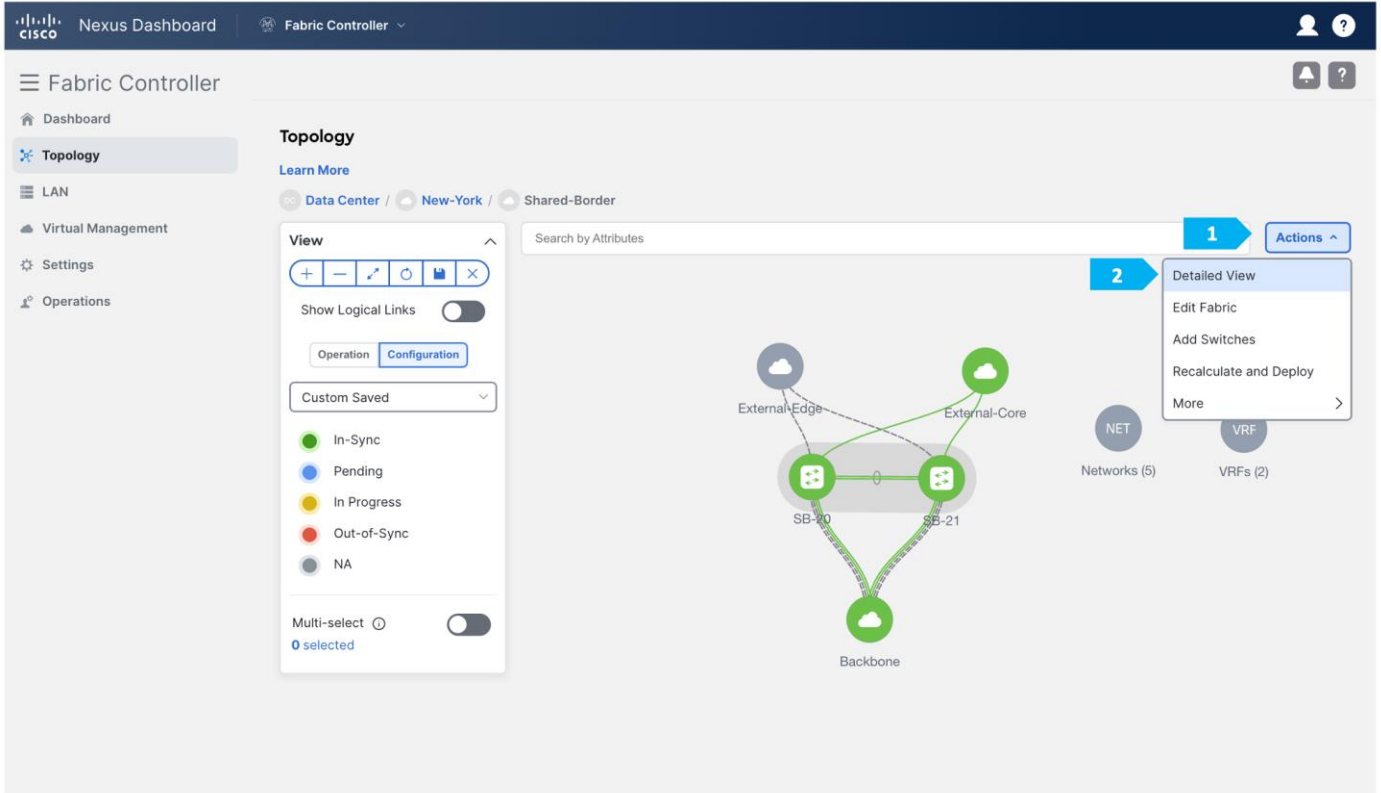
Custom Saved

- In-Sync
- Pending
- In Progress
- Out-of-Sync
- NA

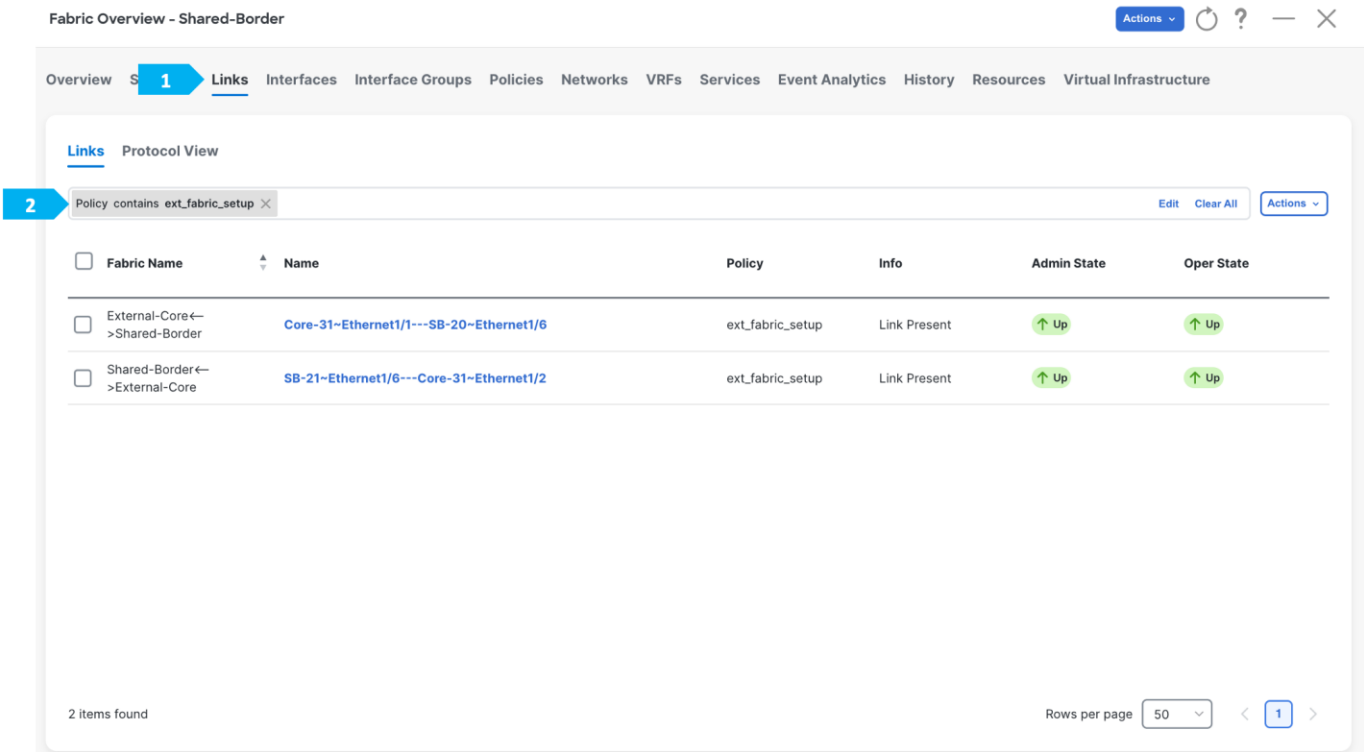
Multi-select 0 selected

1 Shared-Border Core-31

Double-click on the “Shared-Border” fabric.



The “Links” tab will show the Core-31-to-SB-20 link and the Core-31-to-SB-21 link, and the policy should be “ext_fabric_setup” as shown below. We can refine the search using “Policy contains ext_fabric_setup”.



Links Protocol View

Policy contains ext_fabric_setup

Edit Clear All

Actions

Fabric Name	Name	Policy	Info	Admin State	
<input checked="" type="checkbox"/> External-Core ->Shared-Border	Core-31~Ethernet1/1---SB-20~Ethernet1/6	ext_fabric_setup	Link Present	↑ Up	↑ Up
<input type="checkbox"/> Shared-Border ->External-Core	SB-21~Ethernet1/6---Core-31~Ethernet1/2	ext_fabric_setup	Link Present	↑ Up	↑ Up

- Create
- Edit
- Delete
- Import
- Export

2 items found

Rows per page 50 < 1 >

Repeat the same step for SB-21 to Core-31.

Source Fabric External-Core	Destination Fabric Shared-Border
Source Device* Core-31	Destination Device* SB-20
Source Interface* Ethernet1/1	Destination Interface* Ethernet1/6

General Parameters Advanced Default VRF

Source BGP ASN*
65099 BGP Autonomous System Number in Source Fabric

Source IP Address/Mask
10.44.0.1/30 IP address for sub-interface in each VRF in Source Fabric

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

Source IPv6 Address/Mask
IPv6 address for sub-interface in each VRF in Source Fabric

Destination IPv6 Address
IPv6 address for sub-interface in each VRF in Destination Fabric

Destination BGP ASN*
65004 BGP Autonomous System Number in Destination Fabric

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

Auto Generate Configuration for Peer
 If enabled, auto generate VRF Lite configuration for managed NX-OS neighbor devices

1 Cancel Save

Step 4. Attach VRF Extension

Note: Before doing this step, please make sure that the interfaces between Shared borders SB-20, SB-21, and Core-31 are routed ports and not trunk ports.

Go to the “VRF” tab and double-click on the “CORP” VRF.

Fabric Overview - Shared-Border

Actions ↕ ↻ ? — ×

Overview Switches Links Interfaces Interface Groups Policies **1** VRFs Services Event Analytics History Resources Virtual Infrastructure

Filter by attributes

Actions ↕

<input type="checkbox"/>	VRF Name	VRF Status	VRF ID
<input type="checkbox"/>	DMZ	DEPLOYED	50001
<input checked="" type="checkbox"/>	CORP	DEPLOYED	50000

2

2 items found

Rows per page 50 < 1 >

VRF Overview - CORP

Actions ↕ Refresh — ×

Overview **VRF Attachments** Networks

Filter by attributes

Actions ↕

<input type="checkbox"/>	VRF Name	VRF ID	VLAN ID	Switch	Status	Attachment	Switch Role	Fabric Name	Loopback ID	Loopback IPV4 Address	Loopback IPV6 Address
<input checked="" type="checkbox"/>	CORP	50000	2000	SB-20	DEPLOYED	Attached	border	Shared-Border			
<input type="checkbox"/>	CORP	50000	2000	SB-21	DEPLOYED	Attached	border	Shared-Border			

1

2

3

- History
- Edit
- Preview
- Deploy
- Import
- Export
- Quick Attach
- Quick Detach

2 items found

Rows per page 50 < 1 >

SB-20(9NL6XB9DR3X) - SB-21(9DNMFIO0IX)

Detach Attach

VLAN*

Extend*

SB-20(9NL6XB9DR3X)

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

Loopback Id

Loopback IPv4 Address

Loopback IPv6 Address

Import EVPN Route Target

Export EVPN Route Target

SB-21(9DNMFIO0IX)

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

Loopback Id

Loopback IPv4 Address

Loopback IPv6 Address

Import EVPN Route Target

Export EVPN Route Target

Extension

Filter by attributes 1 [Attach-All](#) [Detach-All](#)

Action	Attached	Source Switch	Type	IF_NAME	Dest. Switch	Dest. Interface	DOT1Q_ID	IP_MASK	IP_TAG	NEIGHBO...	NEIGHBO...	IPV6_MA...	IPV6
Edit	● Detached	SB-20	VRF_LITE	Ethernet1/6	Core-31	Ethernet1/1	2	10.44.0.2/30		10.44.0.1	65099		
Edit	● Detached	SB-21	VRF_LITE	Ethernet1/6	Core-31	Ethernet1/2	2	10.44.0.5/30		10.44.0.6	65099		

[Cancel](#) [Save](#) 2

Nexus Dashboard Fabric Controller

Fabric Controller

- Dashboard
- Topology
- LAN
- Virtual Management
- Settings
- Operations

Topology

Learn More

Data Center / New-York / Shared-Border

View

Show Logical Links

[Operation](#) [Configuration](#)

Custom Saved

- Healthy
- Warning
- Minor
- Major
- Critical
- NA

Multi-select

0 selected

Search by Attributes

1 [Actions](#)

- Detailed View
- Edit Fabric
- Add Switches
- Recalculate and Deploy
- More >

2

NET

Networks (5)

VRFs (2)

© 2022 Cisco and/or its affiliates. All rights reserved.

Page 122 of 135

1
2

Config Preview
Deploy Progress

Resync All

Switch Name	IP Address	Role	Serial Number	Fabric Status	Pending Config	Status Description	Progress	Resync Switch
SB-20	100.64.254.20	border	9NL6XB9DR3X	● Out-Of-Sync	21 Lines	Out-of-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync
SB-21	100.64.254.21	border	9DNMFI0D0IX	● Out-Of-Sync	21 Lines	Out-of-Sync	<div style="width: 100%; height: 10px; background-color: green;"></div>	Resync

Close
Deploy All
1

After we finish the deployment in the “Shared-Border” fabric, double-click on the “External-Core” fabric and perform a “Recalculate and Deploy”.

Nexus Dashboard
Fabric Controller

Fabric Controller

- Dashboard
- Topology
- LAN
- Virtual Management
- Settings
- Operations

Topology

Learn More

Data Center / New-York / Shared-Border

View

+ - 🔍 🗨️ ✕

Show Logical Links

Operation Configuration

Custom Saved

- In-Sync
- Pending
- In Progress
- Out-of-Sync
- NA

Multi-select 0 selected

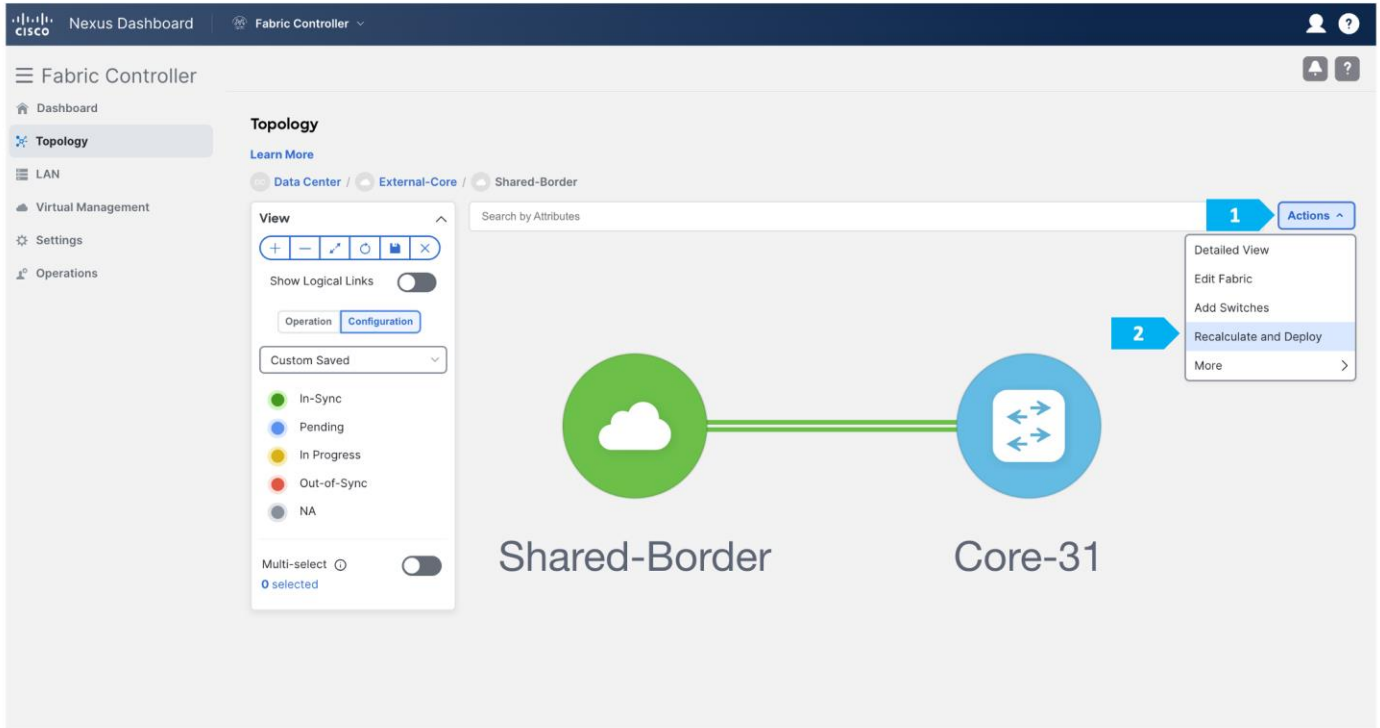
Actions

NET

Networks (5)

VRF

VRFs (2)



By now we have implemented VRF Lite between Shared-Border and External-Core for the CORP VRF.

Note: The VRF Lite deployment option shown above gives an example of one way to extend IP handoff services between VXLAN EVPN and external networks. The deployment of Shared-Borders can be Layer 3 independent devices (no vPC) or part of a vPC domain. By default, the Shared-Borders extend Layer 3 services across different routing domains. The VXLAN EVPN traffic behavior changes based on the deployment model. For example, Shared-Borders running as Layer 3 independent devices use its Primary VTEP IP as the BGP NH (next-hop) to advertise the Site-External prefixes to VXLAN EVPN fabrics. However, Shared-Borders that are part of a vPC domain will use the Secondary known as VIP VTEP IP as the BGP NH to advertise the Site-External prefixes to VXLAN EVPN fabrics.

To handle specific traffic and link failure scenarios, the following is recommended:

- Use “Advertise-PIP” of vPC Border devices when doing VXLAN EVPN to IP handoff. For more information, see:

https://www.cisco.com/c/en/us/td/docs/dcn/nx-os/nexus9000/104x/configuration/vxlan/cisco-nexus-9000-series-nx-os-vxlan-configuration-guide-release-104x/m_configuring_vpc_multihoming.html

- Starting with the NDFC 12.1.3b release, the “Advertise-PIP” option is enabled by default for vPC Border devices.
- In unique failure scenarios, such as a Zig-Zag failure where the Shared-Border-1 loses all of its links towards Edge and the Shared-Border-2 loses all of its links towards the VXLAN EVPN fabric, special considerations must be accounted for, such as the deployment of Layer 3 Underlay link across the Shared-Border for continuous connectivity across VXLAN EVPN and External IP networks, and a per-VRF iBGP session for handling locally attached EPs, service nodes, or external devices.

Service Node Peering Use-Cases

In earlier sections, we discussed how Shared-Border can be implemented in the vPC domain to connect with Layer 4 to Layer 7 service nodes, such as firewalls, load balancers, TCP Optimizers, and more.

While this document does not cover details about Layer 4 to Layer 7 design, best practices, and use cases, it is important to highlight two common use cases with Shared Border as follows:

Layer 2 Extension for DMZ:

Typically, data center applications such as SaaS and other critical customer-facing applications require Internet connectivity. In the data center, the network admin deploys a perimeter firewall for traffic inspection, especially for traffic traversing between untrust and trust zones. Therefore, service nodes such as firewalls host network gateway services for these applications. In such circumstances, the VXLAN EVPN fabric acts as a Layer 2 bridging domain between endpoint applications and the firewall.

In a Shared-Border architecture, the placement and connectivity of Layer 4 to Layer 7 services become crucial to avoid traffic hair pinning and to achieve deterministic traffic flows. When we have multiple Availability Zones (AZs), the Shared-Border becomes a natural choice to connect with the service nodes.

It is also important to note that Layer 2 BUM and bridging traffic must flow across these fabrics. The site/AZ-specific Border Gateway (BGW) is responsible for distributing the Layer 2 information of endpoints within and across the fabrics. At this time, Cisco NDFC supports Ingress-Replication (IR) as the replication method for DCI (VXLAN Multi-Site). The BGWs advertise EVPN Type-3 (IMET) routes to form an IR table with the L2VNI and the VTEP information. Therefore, we must ensure that Layer 2 VXLAN traffic arrives at the Shared Border to process and forward to the service nodes. Hence, the replication method for the Shared-Border fabric must be set to Ingress-Replication during the Day-0 fabric configuration using NDFC. From a configuration point of view, we must create and deploy Layer 2 only VNI across Leaf, BGW, Shared-Border, and the interface connecting between the Shared-Border and the service node.

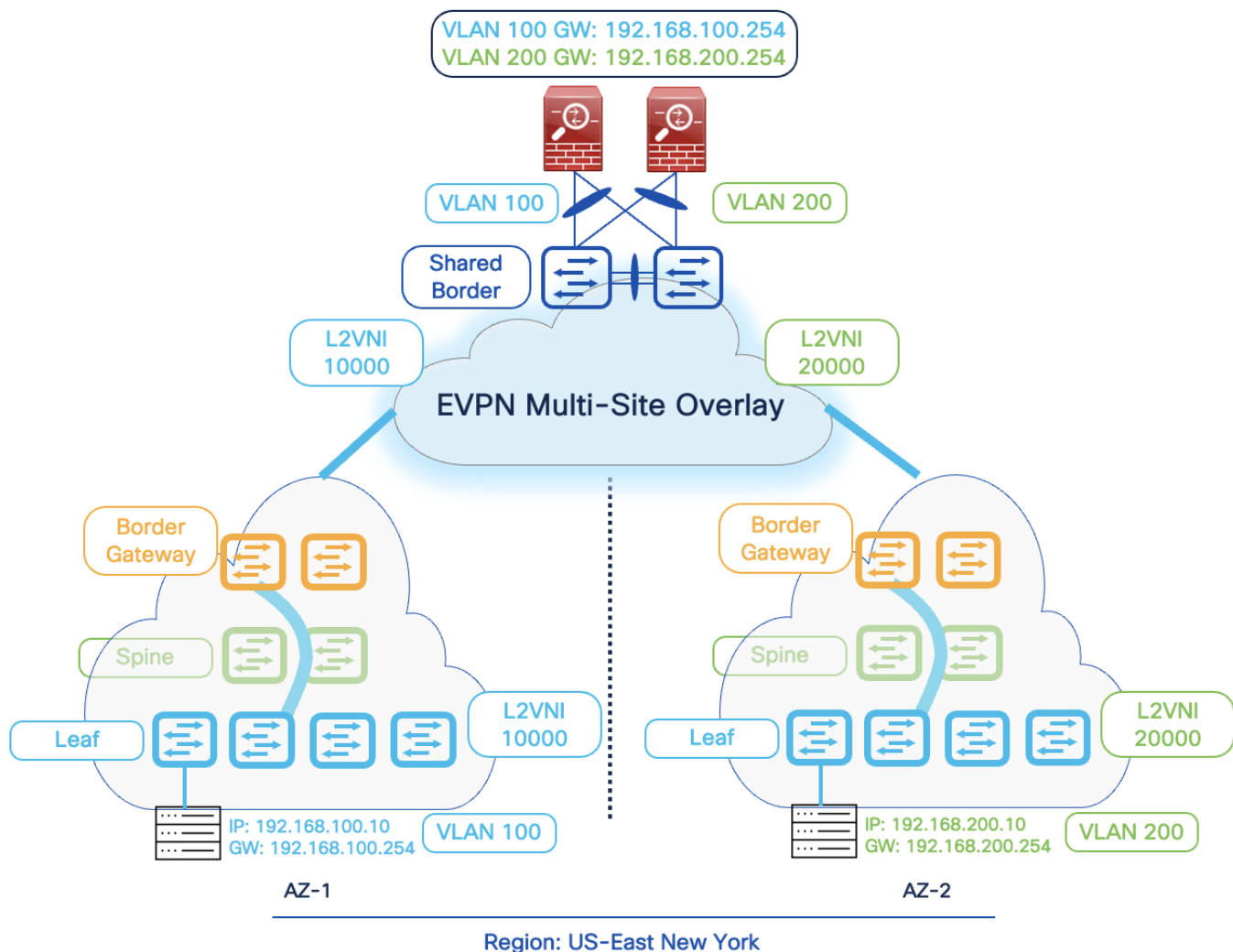


Figure 15. Layer 4-7 Use-Case for DMZ

Inter-Tenant VRF (VRF Fusion)

Another popular use-case for firewall peering in a VXLAN EVPN environment is implementing Inter-Tenant VRF connectivity for Layer 3 communications across different VRFs. By default, a VRF signifies unique and separate control and data plane functionality on a VTEP. One of the advantages of the VXLAN EVPN environment is to achieve Secure Multi-Tenancy and Mobility at scale. Hence, if an endpoint is part of VRF X, the same endpoint can't communicate with another endpoint that is part of VRF Y.

Due to different data center use cases such as migration, mergers, and inter-domain connectivity, traffic is expected to leak across other tenants. While various methods such as EVPN RT import/export, Downstream VNI, and Centralized Route Leaking are available to perform the route leaking on Cisco Nexus 9000 and NX-OS devices, one of the other standard methods is to rely on an external service node to inspect and perform these additional functionalities.

Therefore, a service node such as a firewall acts as a fusion stitching point to enable communication between VRF X and VRF Y. From a configuration point of view, Cisco NDFC supports static routing or

dynamic routing using BGP between the Shared-Border and the service node. The example in this document is based on static routing, but the same can be implemented using BGP.

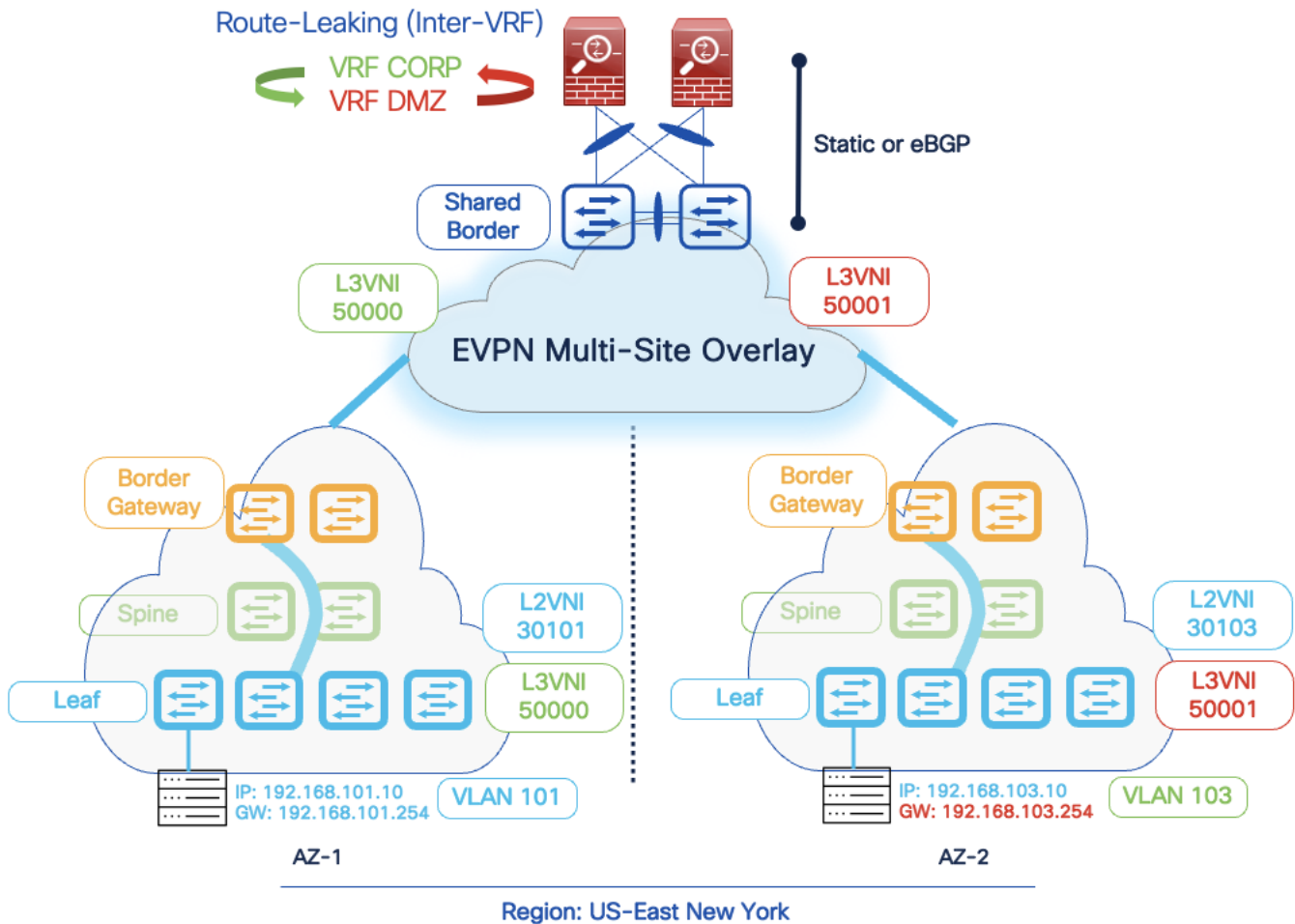


Figure 16. Layer 4-7 Use-Case for Inter-Tenant

In this example, we need to connect the firewall to SB-20 and SB-21 using vPC.



Nexus Dashboard | Fabric Controller

Fabric Controller

- Dashboard
- Topology
- LAN
- 1** Fabrics
- Switches
- Interfaces
- Services
- Virtual Management
- Settings
- Operations

LAN Fabrics

Filter by attributes

Fabric Name	Fabric Technology	Fabric Type	ASN
<input type="radio"/> New-York Hide child Fabrics	VXLAN Fabric	Multi-Fabric Domain	NA
<input type="radio"/> AZ1-New-York	VXLAN Fabric	Switch Fabric	65001
<input type="radio"/> AZ2-New-York	VXLAN Fabric	Switch Fabric	65002
<input type="radio"/> Backbone	External	External	65003
2 <input type="radio"/> Shared-Border	VXLAN Fabric	Switch Fabric	65004

10 Rows

Fabric Shared-Border **3**

Minor

Alarms(98)

CRITICAL	MAJOR	MINOR	WARNING
0	0	98	0

Fabric Info

ASN: 65004
Fabric Technology: VXLAN Fabric
Fabric Type: Switch Fabric
Deployment Status: Enabled

Inventory

Switch Configuration

2 Switches
In-Sync (2)

Switch Health

2 Switches
Minor (2)

Fabric Overview - Shared-Border

Actions

- Overview
- Switches
- 1** Interfaces
- Interface Groups
- Policies
- Networks
- VRFs
- Services
- Event Analytics
- History
- Resources
- Virtual Infrastructure

Filter by attributes

Device Name	Interface	Admin Status	Oper. Status	Reason	Policies	Overlay Network	3
<input type="checkbox"/> SB-20	mgmt0	↑ Up	↑ Up	ok	int_mgmt	NA	● In-Sync
<input type="checkbox"/> SB-20	Vlan1	↓ Down	↓ Down	Administratively down	NA	NA	● NA
<input type="checkbox"/> SB-20	Vlan2000	↑ Up	↑ Up	ok	NA	CORP	● NA
<input type="checkbox"/> SB-20	Vlan2001	↑ Up	↑ Up	ok	NA	DMZ	● NA
<input type="checkbox"/> SB-20	Vlan3600	↑ Up	↑ Up	ok	int_fabric_vlan_11_1	NA	● In-Sync
<input type="checkbox"/> SB-20	Loopback0	↑ Up	↑ Up	ok	int_fabric_loopback_11_1	NA	● In-Sync
<input type="checkbox"/> SB-20	Loopback1	↑ Up	↑ Up	ok	int_fabric_loopback_11_1	NA	● In-Sync
<input type="checkbox"/> SB-20	Loopback2	↑ Up	↑ Up	ok	int_loopback	NA	● In-Sync
<input type="checkbox"/> SB-20	Loopback3	↑ Up	↑ Up	ok	int_loopback	NA	● In-Sync
<input type="checkbox"/> SB-20	Port-channel500	↑ Up	↑ Up	ok	int_vpn_peer_link_00	NA	● In-Sync

50 Rows

Page 1 of 3

2 Actions

- Create Interface
- Create Subinterface
- Edit
- Preview
- Deploy
- No Shutdown
- Shutdown
- Add to Interface Group
- Remove from Interface Group
- Breakout
- UnBreakout
- More

Create Interface



1 Type*
virtual Port Channel (vPC) ▾

2 Select a vPC pair*
SB-20---SB-21 ▾

3 vPC ID*
5

4 Policy*
int_vpc_trunk_host >

Policy Options

5 Peer-1 Port-Channel ID*
5 Peer-1 VPC port-channel number (Min:1, Max:4096)

Peer-2 Port-Channel ID*
5 Peer-2 VPC port-channel number (Min:1, Max:4096)

6 Enable Config Mirroring
 If enabled, Peer-1 config will be copied to Peer-2

Peer-1 Member Interfaces
e1/5 A list of member interfaces for Peer-1 [e.g. e1/5,eth1/7-9]

Peer-2 Member Interfaces
e1/5 A list of member interfaces for Peer-2 [e.g. e1/5,eth1/7-9]

Port Channel Mode*
active ▾ Channel mode options: on, active and passive

7 Save Preview Deploy

Create Interface



Type*
virtual Port Channel (vPC) ▾

Select a vPC pair*
SB-20---SB-21 ▾

vPC ID*
5

Policy*
int_vpc_trunk_host

Policy Options

Peer-1 Port-Channel ID
5 Peer-1 VPC port-channel number (Min:1, Max:4096)

Peer-2 Port-Channel ID
5 Peer-2 VPC port-channel number (Min:1, Max:4096)

Enable Config Mirroring
 If enabled, Peer-1 config will be copied to Peer-2

Peer-1 Member Interfaces
e1/5 A list of member interfaces for Peer-1 [e.g. e1/5,eth1/7-9]

Peer-2 Member Interfaces
e1/5 A list of member interfaces for Peer-2 [e.g. e1/5,eth1/7-9]

Save **1** Deploy

We must create an external fabric and specify that a service node resides in that external fabric during the service node creation. NDFC does not auto-detect or discover any service nodes. We must also specify the service node name, type, and form factor. The name of the service node must be unique within a fabric. NDFC does not define a new switch role for a service node.

NDFC manages the switches that are attached to a service node. It also manages the interfaces of these attached switches. Ensure that the interfaces that the service node is attached to are in trunk mode and do not belong to any Interface Group. When the attached switches are forming a vPC pair, the name of the attached switch is a combination of both switches.

Note: Navigate to Data Center VXLAN EVPN fabric overview (in our case Shared-Border) and the Services tab to make these configurations.

The screenshot displays the 'Create New Service Node' configuration interface. At the top, a progress bar indicates three steps: 1. Create Service Node (active), 2. Create Route Peering, and 3. Create Service Policy. The main configuration area includes the following fields:

- 1. Service Node Name*: External-Firewall
- 2. Service Node Type*: Firewall
- 3. Form Factor*: Physical
- 4. External Fabric*: External-Edge
- 5. Service Node Interface*: Gig0/0
- 6. Attached Fabric*: Shared-Border
- 7. Attached Switch*: SB-20 - SB-21
- 8. Attached Switch Interface*: vPC5
- Link Template*: service_link_vpc

A 'Save' button is located at the bottom right of the configuration area.

We will need to enter information for the inside network and the outside network.

Create Route Peering



1 Detach Attach

2 Peering Name*
Route-Peering

3 Deployment*
Inter-Tenant Firewall

4 Peering Option*
Static Peering

Inside Network

5 VRF*
CORP

6 Network Type*
Inside Network

7 Service Network*
CORP-INSIDE

Outside Network

VRF*
DMZ

Network Type*
Outside Network

Service Network*
DMZ-OUTSIDE

Cancel Save

Note: We need to type the name in the **Network Type** field.

Create Route Peering



1 VLAN ID*
3000

Propose

Network ID*
30000

Service Network Template*
Service_Network_Universal

General Parameters Advanced

2 IPv4 Gateway/NetMask*
20.0.0.1/24

IPv6 Gateway/Prefix
example 2001:db8:1/64

3 VLAN Name
CORP-INSIDE-FW

Interface Description

Peering Template*
service_static_route

4 Static Routes
192.168.103.0/24, 20.0.0.254

One Static Route per line, example 1.2.3.0/24, 1.2.2.2

VLAN ID*
3001

Propose

Network ID*
30001

Service Network Template*
Service_Network_Universal

General Parameters Advanced

2 IPv4 Gateway/NetMask*
20.0.1.1/24

IPv6 Gateway/Prefix
example 2001:db8:1/64

3 VLAN Name
DMZ-OUTSIDE-FW

Interface Description

Peering Template*
service_static_route

4 Static Routes
192.168.101.0/24, 20.0.1.254
192.168.102.0/24, 20.0.1.254

One Static Route per line, example 1.2.3.0/24, 1.2.2.2

Cancel Save

External-Firewall Detail



1 **Route Peering** Service Policy

Filter by attributes

	Peering Name	Deploym...	Peering Option	Status	Attachm... Status	Service Network One			Service Network Two			Reverse Next Hop IP	Reverse Next Hop IPv6	La Up
						VRF	Network Name	Gateway IP	VRF	Network Name	Gateway IP			
2 <input checked="" type="checkbox"/>	Route-Peering	InterTenantFW	StaticPeering	Pending	Attached	CORP	CORP-INSIDE	20.0.0.1/24	DMZ	DMZ-OUTSIDE	20.0.1.1/24	-	06	17

10 Rows Page 1 of 1 << < 1-1 of 1 > >>

3 **Actions**

- Add
- Edit
- Attach
- Detach
- Preview
- 4 **Deploy**
- Import
- Export
- Delete

External-Firewall Detail



Overview **Route Peering** Service Policy

Filter by attributes

Deploy Route Peering

Deploying Route Peering,

Route-Peering

Proceed by clicking Deploy ..

1 **Deploy**

	Peering Name	Deploym...	Peering Option	Status	Attachm... Status	VRF	Network Name	Gateway IP	VRF	Network Name	Gateway IP	Next Hop IP	Reverse Next Hop IP	Reverse Next Hop IPv6	La Up	
<input checked="" type="checkbox"/>	Route-Peering	InterTenantFW	StaticPeering	Pending	Attached	CORP	CORP-INSIDE	20.0.0.1/24	DMZ	DMZ-OUTSIDE	20.0.1.1/24	-	-	-	06	17

10 Rows Page 1 of 1 << < 1-1 of 1 > >>

External-Firewall Detail



Overview **Route Peering** Service Policy

Filter by attributes

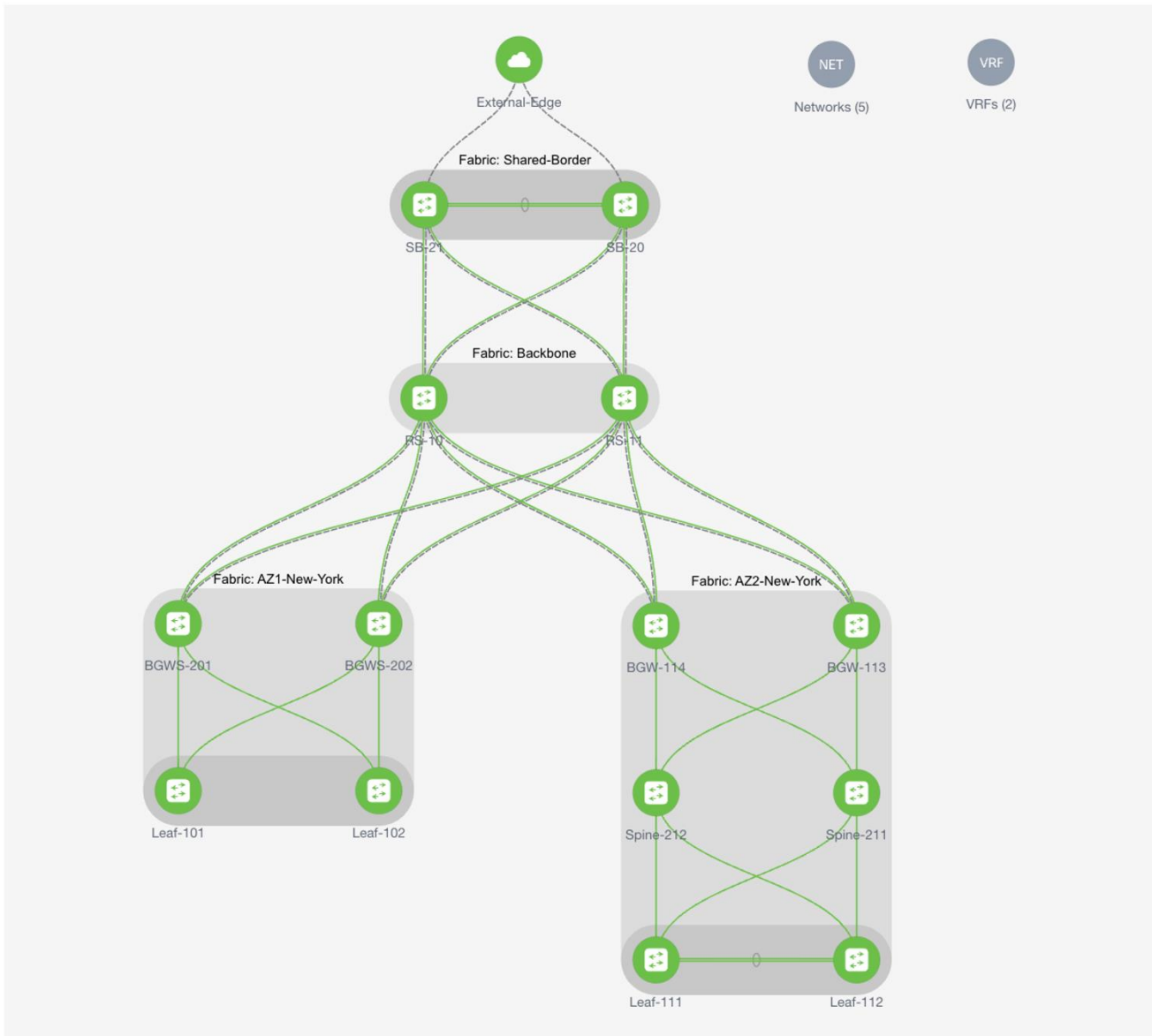
Actions

	Peering Name	Deploym...	Peering Option	Status	Attachm... Status	Service Network One			Service Network Two			Next Hop IP	Reverse Next Hop IP	Next Hop IPv6	Reverse Next Hop IPv6	Last Updated
						VRF	Network Name	Gateway IP	VRF	Network Name	Gateway IP					
<input type="checkbox"/>	Route-Peering	InterTenantFW	StaticPeering	In-Sync	Attached	CORP	CORP-INSIDE	20.0.0.1/24	DMZ	DMZ-OUTSIDE	20.0.1.1/24	-	-	-	-	06/05/2023, 17:22:50

10 Rows

Page 1 of 1 << < 1-1 of 1 > >>

Final NDFC Topology



Conclusion

Shared-Border, which is a site external VTEP, interconnects VXLAN EVPN Multi-Site domains to provide a deterministic connectivity point for Layer 3 IP services and handoff. Flexible deployment models and architecture of Shared-Border allows a network admin to optimize Layer 2 and Layer 3 DCI traffic flows by interconnecting various Availability Zones and extending the connectivity to shared services.

Additional Information

Configuration Guides and White Papers

<https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/white-paper-c11-739942.html#Verificationandshowcommands>

<https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/whitepaper-c11-742114.html>

<https://www.cisco.com/c/en/us/products/cloud-systems-management/prime-data-center-network-manager/white-paper-listing.html>

<https://www.cisco.com/c/en/us/products/switches/nexus-9000-series-switches/white-paper-listing.html>

<https://www.cisco.com/c/en/us/support/cloud-systems-management/prime-data-center-network-manager/products-installation-and-configuration-guides-list.html>