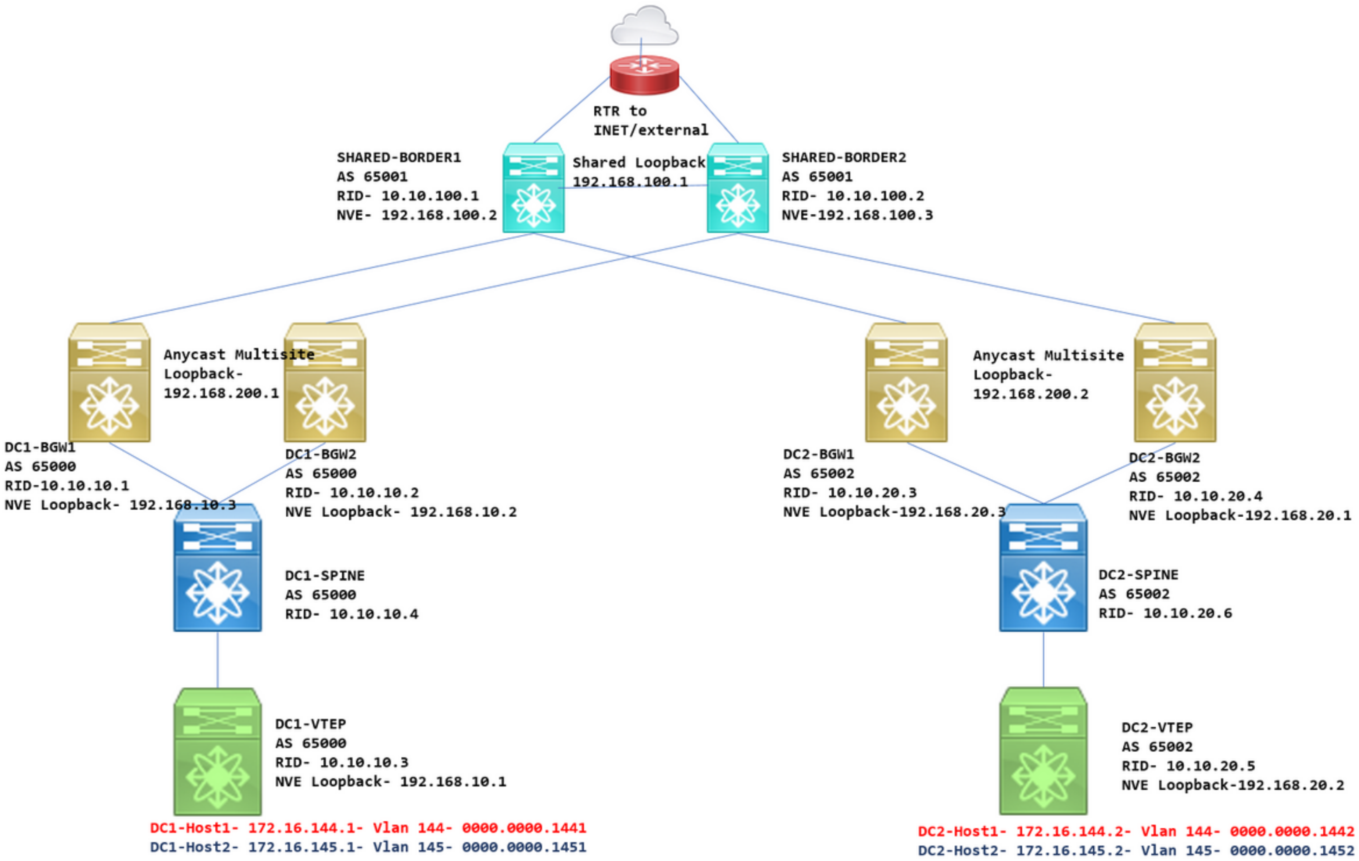
Building a Nexus 9000 VXLAN Shared Border Multisite Deployment using DCNM

Contents

Introduction

This document is to explain how to deploy a Cisco Nexus 9000 VXLAN Multisite Deployment using shared border model using DCNM 11.2 version.

Topology



Details of The topology

DC1 and DC2 are two datacenter locations which are running vxlan;

DC1 and DC2 Border Gateways are having physical connections to the shared borders;

Shared borders have the External connectivity(eg; Internet); so the VRF lite connections are terminated on shared borders and a default route is injected by the shared borders to Border Gateways in each site

Shared borders are configured in vPC(This is a requirement when the fabric is deployed using DCNM)

Border Gateways are configured in Anycast mode

Components Used:

Nexus 9ks running 9.3(2)

DCNM running 11.2 Version

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. If your network is live, make sure that you understand the potential impact of any command.

High Level Steps

1) Considering that this document is based on Two Datacenters utilizing vxlan multisite feature, Two Easy Fabrics have to be created

2) Create Another easy Fabric for the Shared Border

3) Create MSD and move DC1 and DC2

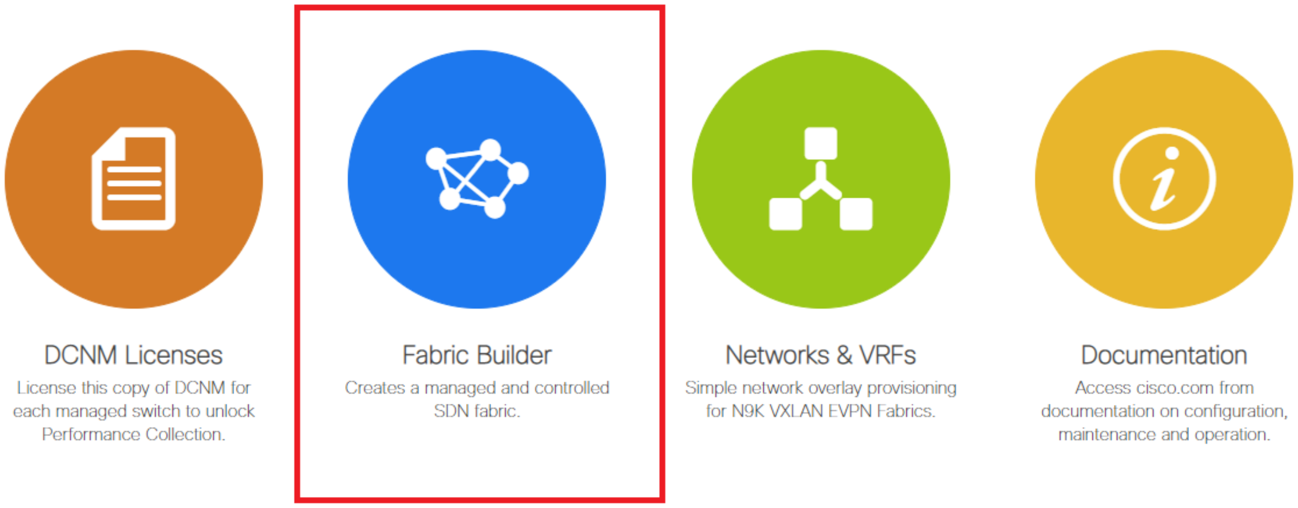
4) Create External Fabric

5) Create Multisite Underlay and Overlay(For East/West)

6) Create VRF Extension Attachments on Shared borders

Step 1: Creation of Easy Fabric for DC1

- **Login to the DCNM and from the Dashboard, Select the option-> "Fabric Builder"**



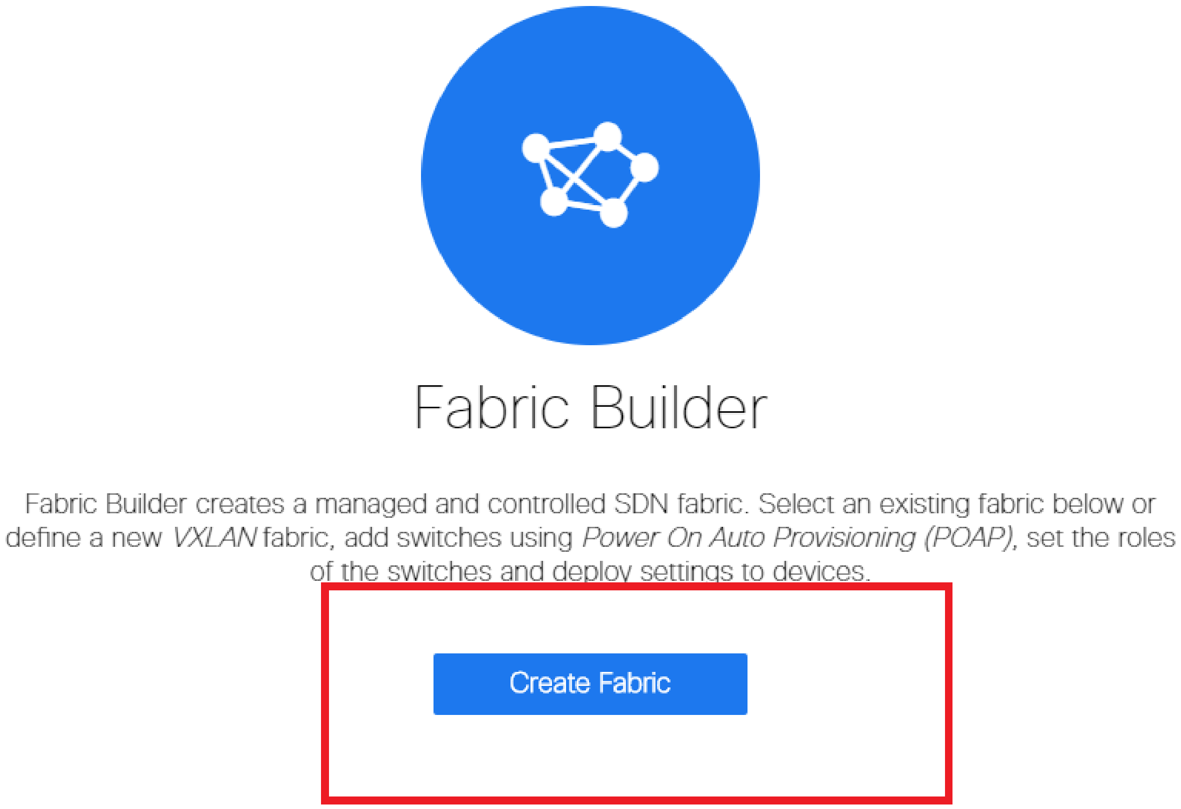
DCNM Licenses
License this copy of DCNM for each managed switch to unlock Performance Collection.

Fabric Builder
Creates a managed and controlled SDN fabric.

Networks & VRFs
Simple network overlay provisioning for N9K VXLAN EVPN Fabrics.

Documentation
Access cisco.com from documentation on configuration, maintenance and operation.

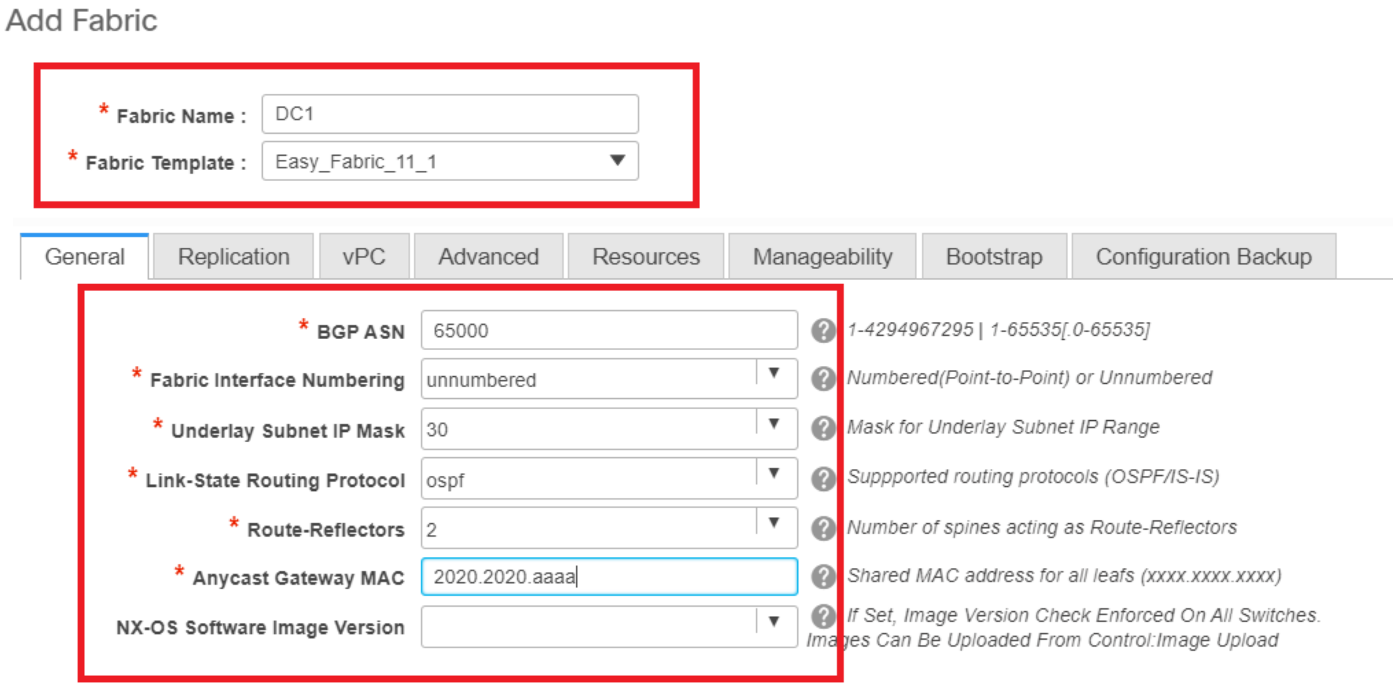
- **Select the "create fabric" option**



Fabric Builder

Fabric Builder creates a managed and controlled SDN fabric. Select an existing fabric below or define a new *VXLAN* fabric, add switches using *Power On Auto Provisioning (POAP)*, set the roles of the switches and deploy settings to devices.

- **Next is to provide the fabric name, template and then multiple tabs will open up which will need details like ASN, Fabric Interface Numbering, Any Cast Gateway MAC(AGM)**



Add Fabric

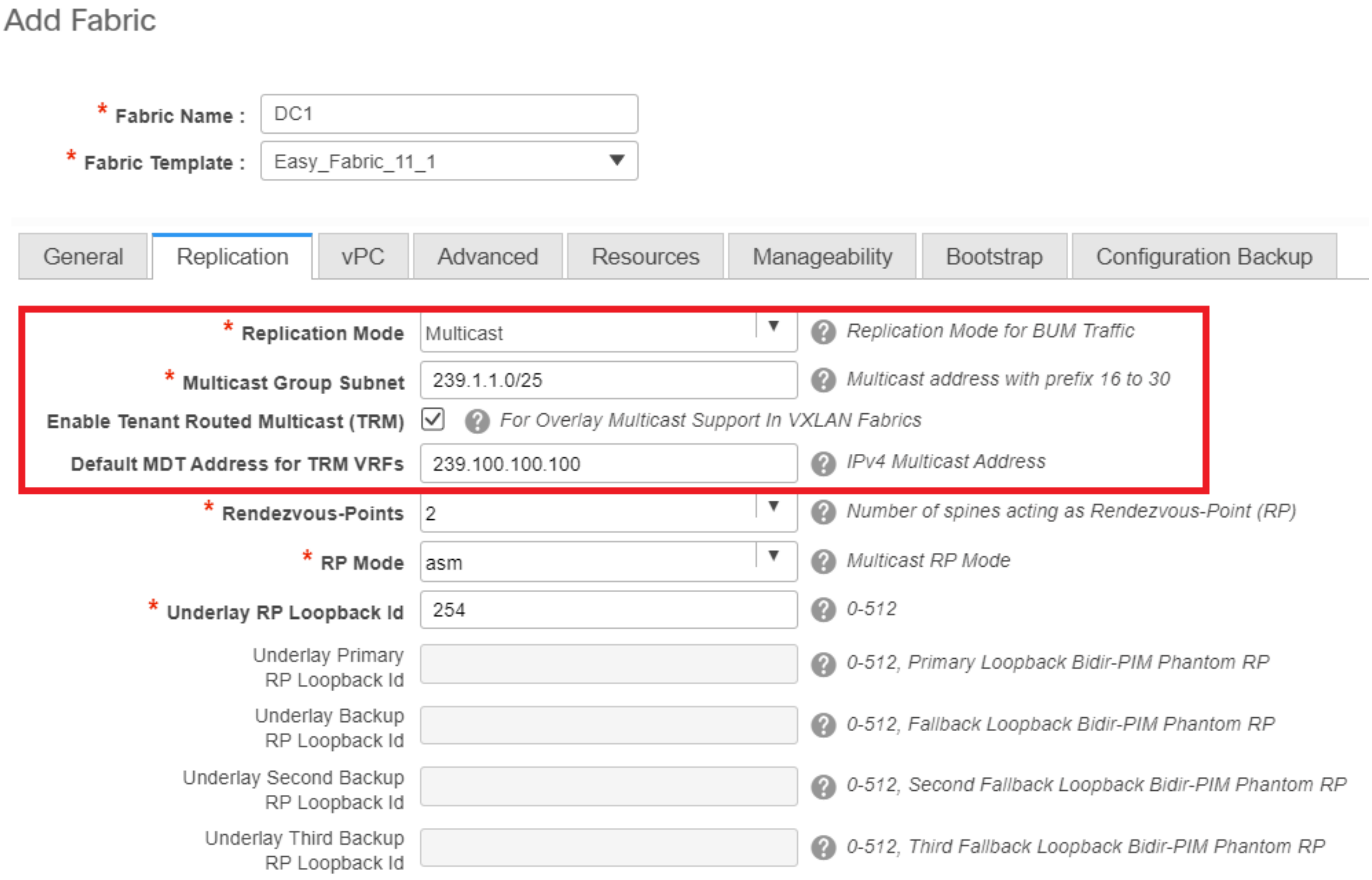
* Fabric Name :	DC1
* Fabric Template :	Easy_Fabric_11_1 ▼

* BGP ASN	65000	1-4294967295 | 1-65535[.0-65535]
* Fabric Interface Numbering	unnumbered ▼	Numbered(Point-to-Point) or Unnumbered
* Underlay Subnet IP Mask	30 ▼	Mask for Underlay Subnet IP Range
* Link-State Routing Protocol	ospf ▼	Suppported routing protocols (OSPF/IS-IS)
* Route-Reflectors	2 ▼	Number of spines acting as Route-Reflectors
* Anycast Gateway MAC	2020.2020.aaaa	Shared MAC address for all leafs (xxxx.xxxx.xxxx)
NX-OS Software Image Version	▼	If Set, Image Version Check Enforced On All Switches. Images Can Be Uploaded From Control:Image Upload

# Fabric Interfaces(which are the Spine/Leaf interfaces) can be "unnumbered" or point to point; If unnumbered is used, the IP addresses required are less(as the IP address is that of the unnumbered loopback)

# AGM is used by the Hosts in the Fabric as the Default Gateway MAC address; This will be the same on all Leaf switches which are the Default Gateways

- **Next is to set the Replication mode**



Add Fabric

* Fabric Name : DC1

* Fabric Template : Easy_Fabric_11_1

* Replication Mode: Multicast — Replication Mode for BUM Traffic

* Multicast Group Subnet: 239.1.1.0/25 — Multicast address with prefix 16 to 30

Enable Tenant Routed Multicast (TRM): ☑ — For Overlay Multicast Support In VXLAN Fabrics

Default MDT Address for TRM VRFs: 239.100.100.100 — IPv4 Multicast Address

* Rendezvous-Points: 2 — Number of spines acting as Rendezvous-Point (RP)

* RP Mode: asm — Multicast RP Mode

* Underlay RP Loopback Id: 254 — 0-512

Underlay Primary RP Loopback Id: — 0-512, Primary Loopback Bidir-PIM Phantom RP

Underlay Backup RP Loopback Id: — 0-512, Fallback Loopback Bidir-PIM Phantom RP

Underlay Second Backup RP Loopback Id: — 0-512, Second Fallback Loopback Bidir-PIM Phantom RP

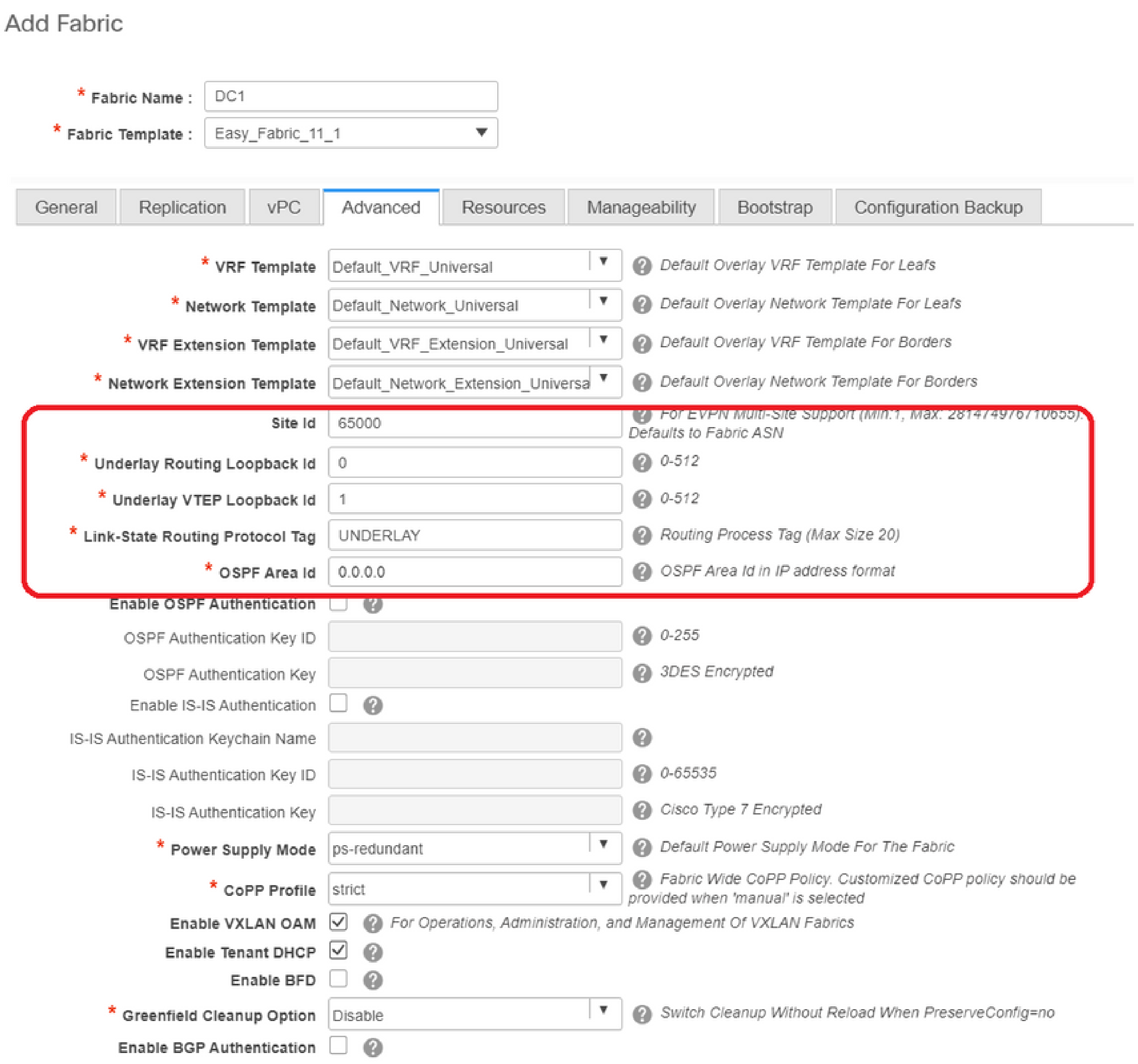
Underlay Third Backup RP Loopback Id: — 0-512, Third Fallback Loopback Bidir-PIM Phantom RP

# Replication mode selected here can be either multicast or IR-Ingress Replication; IR will replicate any incoming BUM traffic within a vxlan vlan in a unicast fashion to other VTEPs which is also called Head end replication whereas Multicast mode will send the BUM traffic with an outer destination IP address as that of the Multicast group defined for each Networks up to the Spine and Spines will do the multicast replication based off of the Outer destination IP address' OIL to other VTEPs

# Multicast Group subnet-> Required to replicate the BUM traffic(like ARP request from a host)

# If TRM is required to be enabled, select the check box against the same and provide the MDT address for the TRM VRFs.

- **Tab for "vPC" is left at default; If any changes are required for the backup SVI/VLAN, those can be defined here**
- **Advanced tab is the Next Section**



Add Fabric

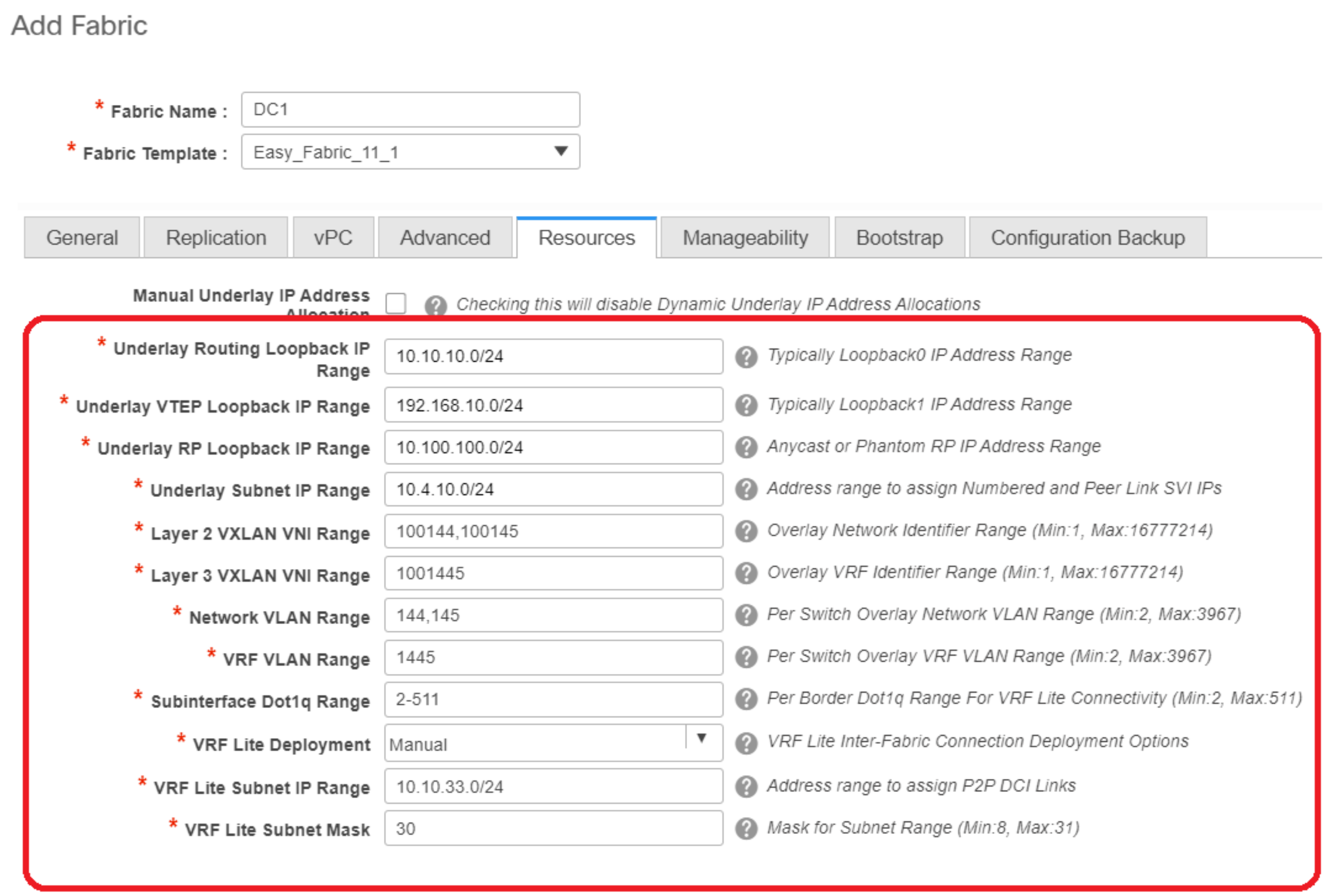
* Fabric Name :	DC1
* Fabric Template :	Easy_Fabric_11_1 ▼

* 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_Universa ▼	Default Overlay Network Template For Borders
Site Id	65000	For EVPN Multi-Site Support (Min:1, Max: 281474976710655). Defaults to Fabric ASN
* Underlay Routing Loopback Id	0	0-512
* Underlay VTEP Loopback Id	1	0-512
* Link-State Routing Protocol Tag	UNDERLAY	Routing Process Tag (Max Size 20)
* OSPF Area Id	0.0.0.0	OSPF Area Id in IP address format
Enable OSPF Authentication	□	
OSPF Authentication Key ID		0-255
OSPF Authentication Key		3DES Encrypted
Enable IS-IS Authentication	□	
IS-IS Authentication Keychain Name		
IS-IS Authentication Key ID		0-65535
IS-IS Authentication Key		Cisco Type 7 Encrypted
* Power Supply Mode	ps-redundant ▼	Default Power Supply Mode For The Fabric
* CoPP Profile	strict ▼	Fabric Wide CoPP Policy. Customized CoPP policy should be provided when 'manual' is selected
Enable VXLAN OAM	☑	For Operations, Administration, and Management Of VXLAN Fabrics
Enable Tenant DHCP	☑	
Enable BFD	□	
* Greenfield Cleanup Option	Disable ▼	Switch Cleanup Without Reload When PreserveConfig=no
Enable BGP Authentication	□	

# Site ID mentioned here is auto-populated on this DCNM Version which is derived from the ASN that is defined underneath the "General" tab

# Fill up/Modify other fields which are relevant

- **Resources tab is the Next one which would need the IP addressing Scheme for Loopbacks, Underlays**



Add Fabric

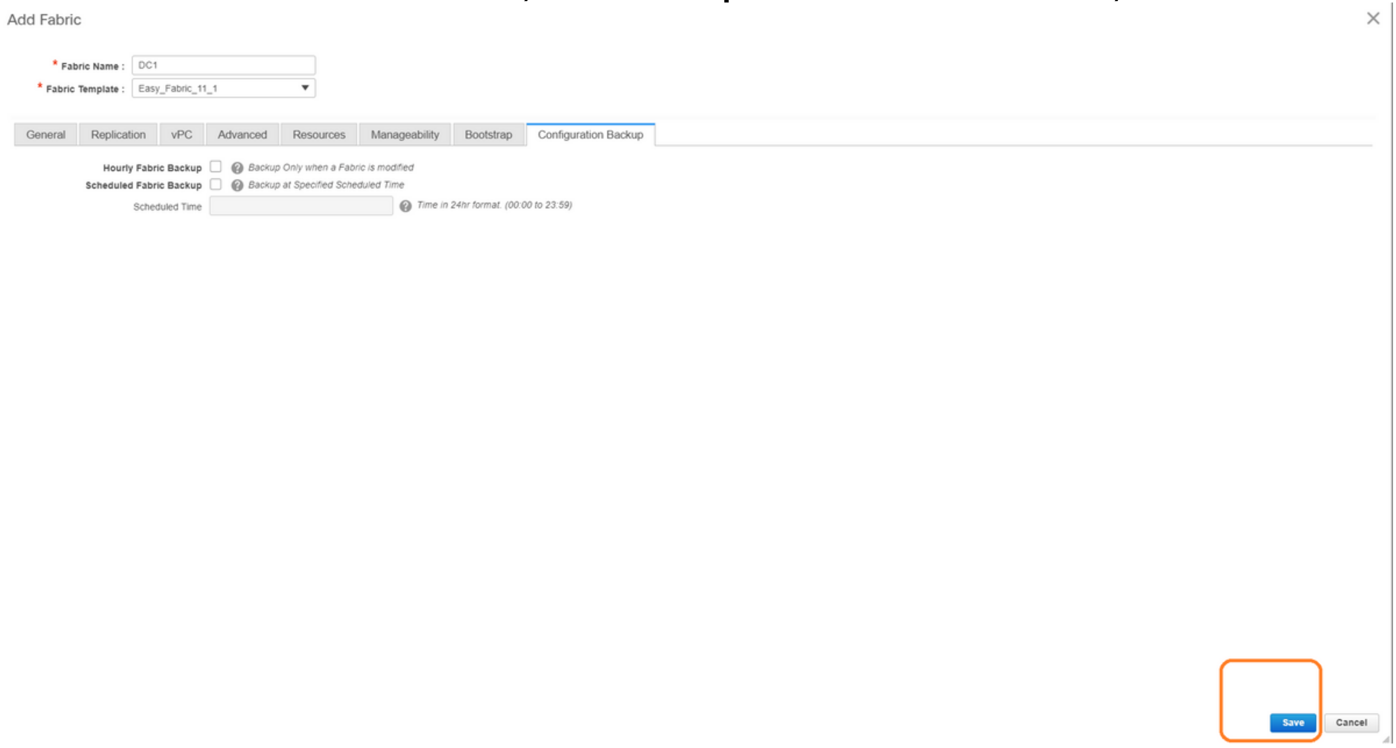
* Fabric Name :	DC1
* Fabric Template :	Easy_Fabric_11_1 ▼

Manual Underlay IP Address Allocation	□	Checking this will disable Dynamic Underlay IP Address Allocations
* Underlay Routing Loopback IP Range	10.10.10.0/24	Typically Loopback0 IP Address Range
* Underlay VTEP Loopback IP Range	192.168.10.0/24	Typically Loopback1 IP Address Range
* Underlay RP Loopback IP Range	10.100.100.0/24	Anycast or Phantom RP IP Address Range
* Underlay Subnet IP Range	10.4.10.0/24	Address range to assign Numbered and Peer Link SVI IPs
* Layer 2 VXLAN VNI Range	100144,100145	Overlay Network Identifier Range (Min:1, Max:16777214)
* Layer 3 VXLAN VNI Range	1001445	Overlay VRF Identifier Range (Min:1, Max:16777214)
* Network VLAN Range	144,145	Per Switch Overlay Network VLAN Range (Min:2, Max:3967)
* VRF VLAN Range	1445	Per Switch Overlay VRF VLAN Range (Min:2, Max:3967)
* Subinterface Dot1q Range	2-511	Per Border Dot1q Range For VRF Lite Connectivity (Min:2, Max:511)
* VRF Lite Deployment	Manual ▼	VRF Lite Inter-Fabric Connection Deployment Options
* VRF Lite Subnet IP Range	10.10.33.0/24	Address range to assign P2P DCI Links
* VRF Lite Subnet Mask	30	Mask for Subnet Range (Min:8, Max:31)

# Layer 2 VXLAN VNI Range-> These are the VNIDs which will be later mapped to Vlans(Will show it further down)

# Layer 3 VXLAN VNI Range-> These are the Layer 3 VNIDs that will also be later mapped to layer 3 VNI Vlan to Vn-segment

- **Other tabs are not shown here; but do Fill up the other tabs if needed;**

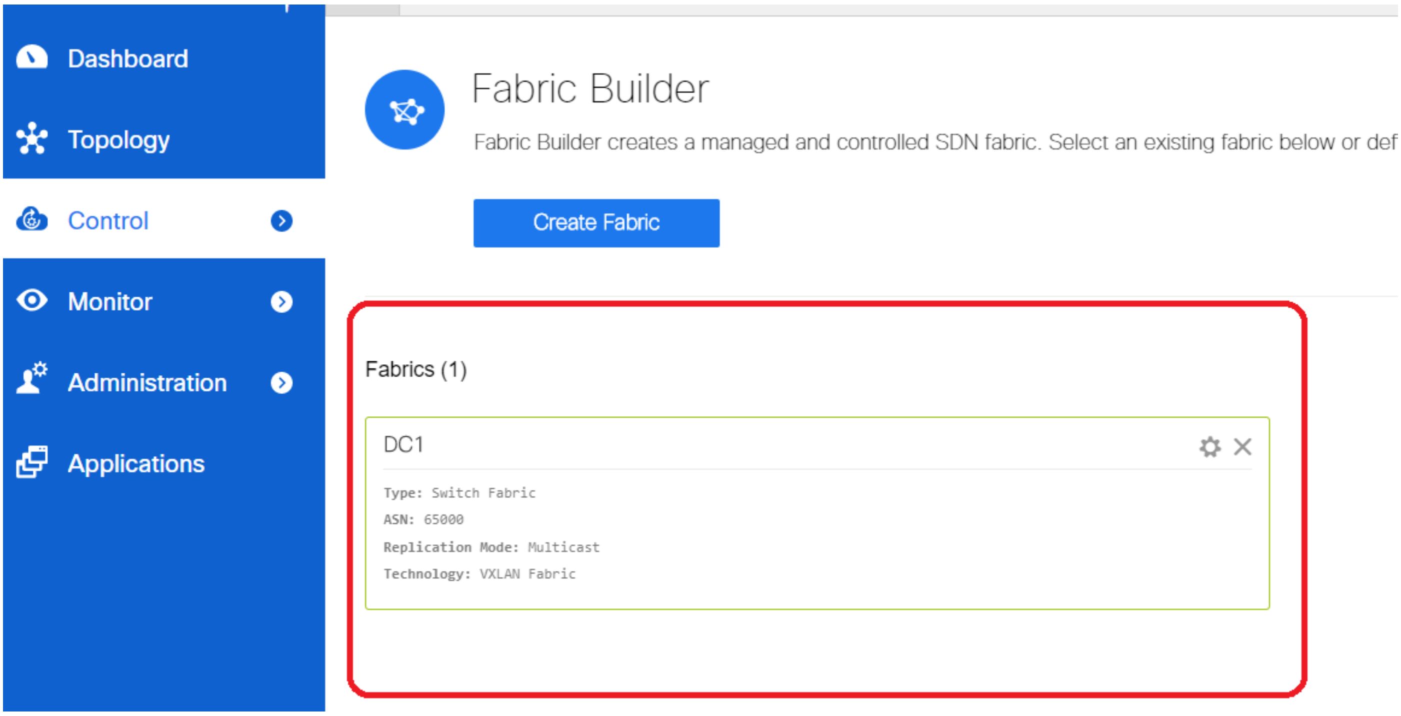


Add Fabric

* Fabric Name :	DC1
* Fabric Template :	Easy_Fabric_11_1 ▼

Hourly Fabric Backup	□	Backup Only when a Fabric is modified
Scheduled Fabric Backup	□	Backup at Specified Scheduled Time
Scheduled Time		Time in 24hr format. (00:00 to 23:59)

- **Once you save, the fabric builder page will show the Fabric(From DCNM-> Control-> Fabric Builder**



Fabric Builder

Fabric Builder creates a managed and controlled SDN fabric. Select an existing fabric below or def

Fabrics (1)

DC1

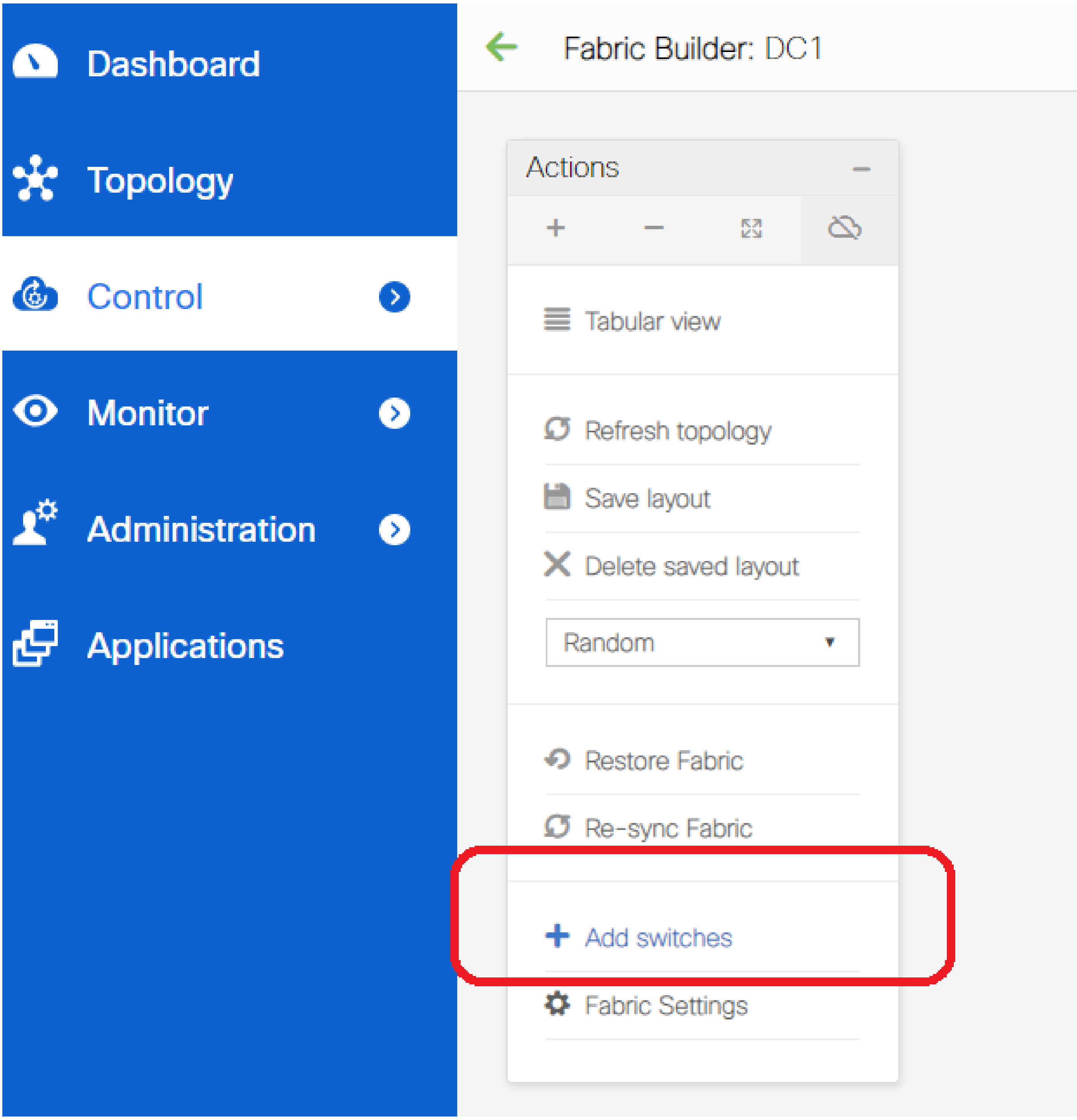
Type: Switch Fabric
ASN: 65000
Replication Mode: Multicast
Technology: VXLAN Fabric

# This section shows the full list of Fabrics, ASN, Replication modes for each of the Fabrics

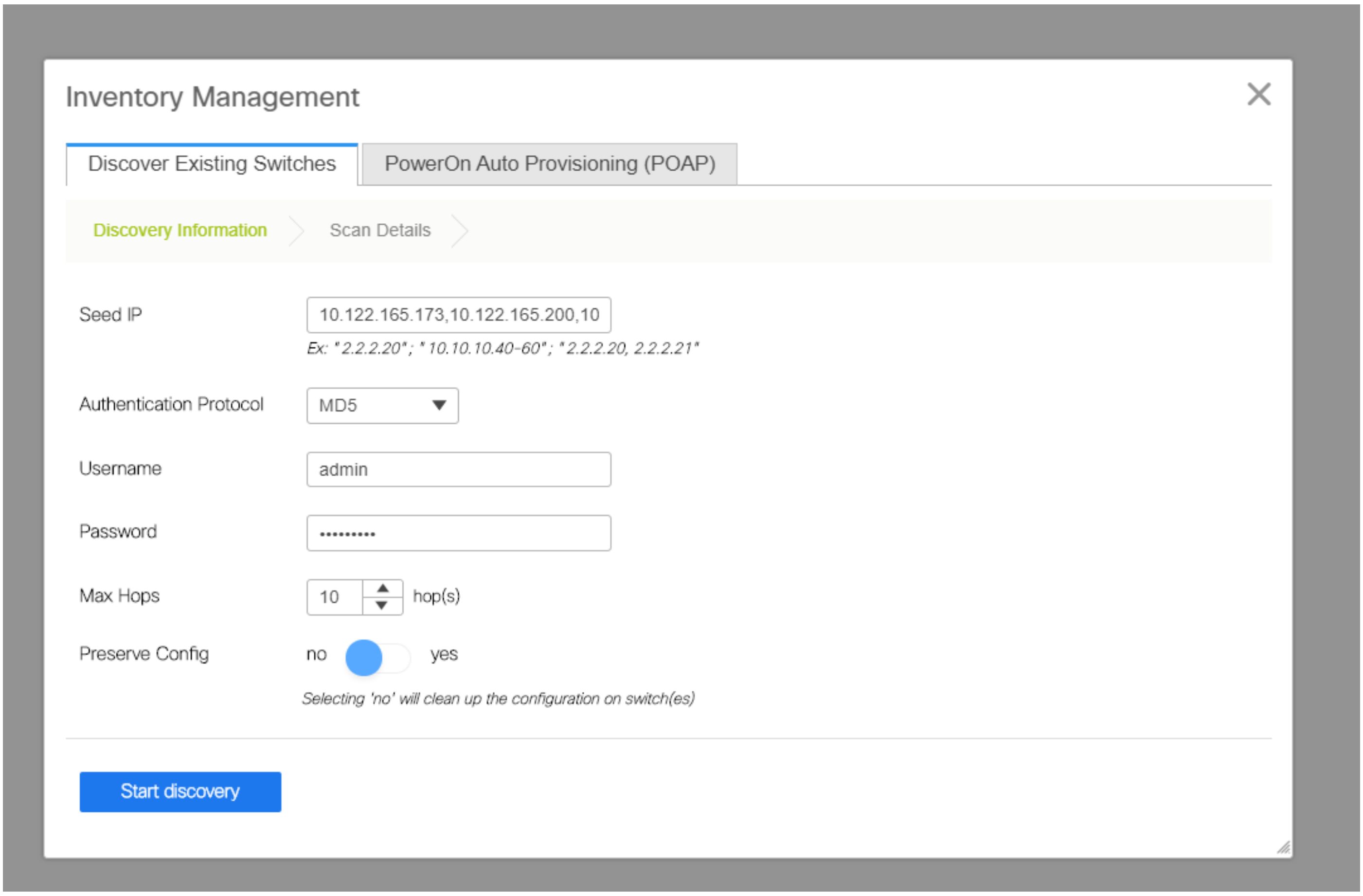
- **Next step is to Add switches onto the DC1 Fabric**

Step 2: Add switches into the DC1 Fabric

Click on DC1 in the diagram above and that would give the option to add switches.



- **Provide IP addresses and credentials of the switches that needs to be imported to DC1 Fabric(Per topology listed in the beginning of this document, DC1-VTEP, DC1-SPINE, DC1-BGW1 & DC1-BGW2 are part of DC1)**



Inventory Management

Discover Existing Switches | PowerOn Auto Provisioning (POAP)

Discovery Information > Scan Details >

Seed IP: 10.122.165.173,10.122.165.200,10

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

Authentication Protocol: MD5

Username: admin

Password: •••••••••

Max Hops: 10 hop(s)

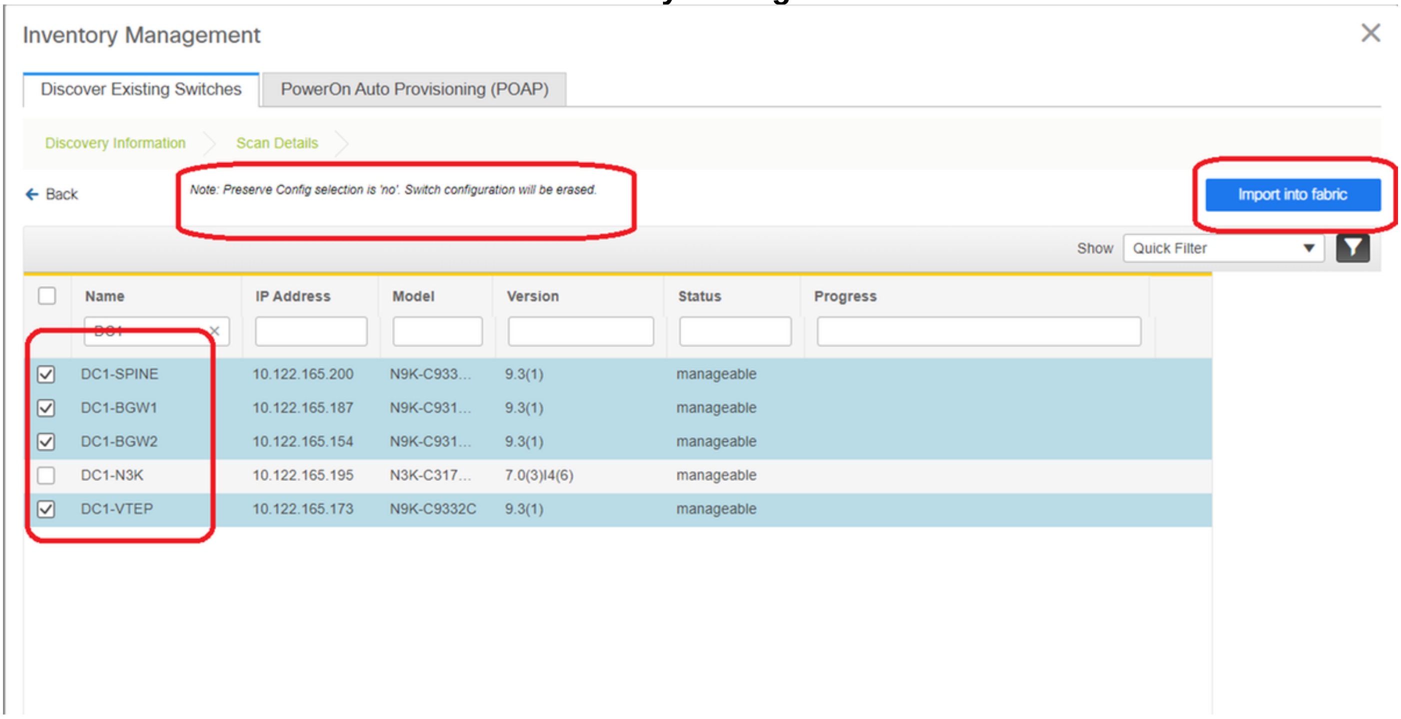
Preserve Config: no ● yes

Selecting 'no' will clean up the configuration on switch(es)

# Since this is a Greenfield deployment, Note that the "preserve config" option is selected as "NO"; which will delete all the configurations of the boxes while doing the import and also will reload the switches

# Select the "Start discovery" so that DCNM will start discovering the switches based on the IP addresses provided in the "seed IP" column

- **Once the DCNM finishes discovering the switches, the IP addresses along with the hostnames will be listed in the Inventory management**



Inventory Management

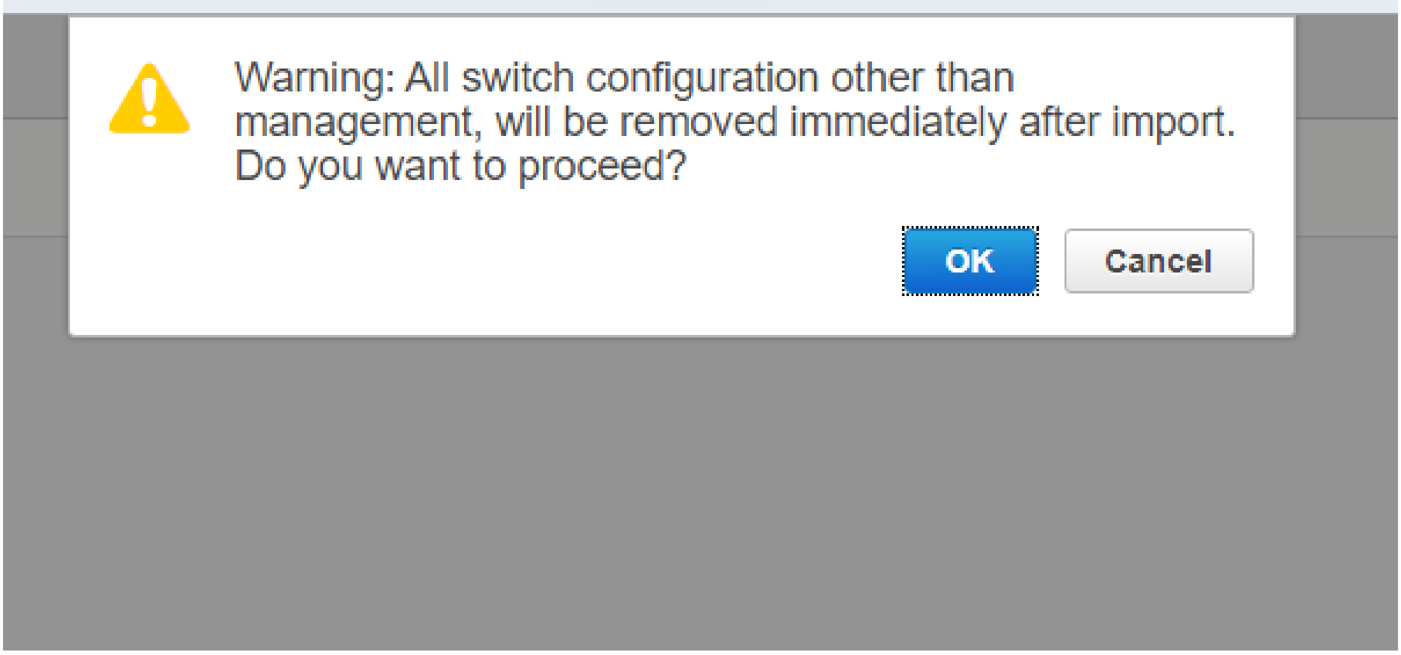
Discover Existing Switches | PowerOn Auto Provisioning (POAP)

Discovery Information > Scan Details >

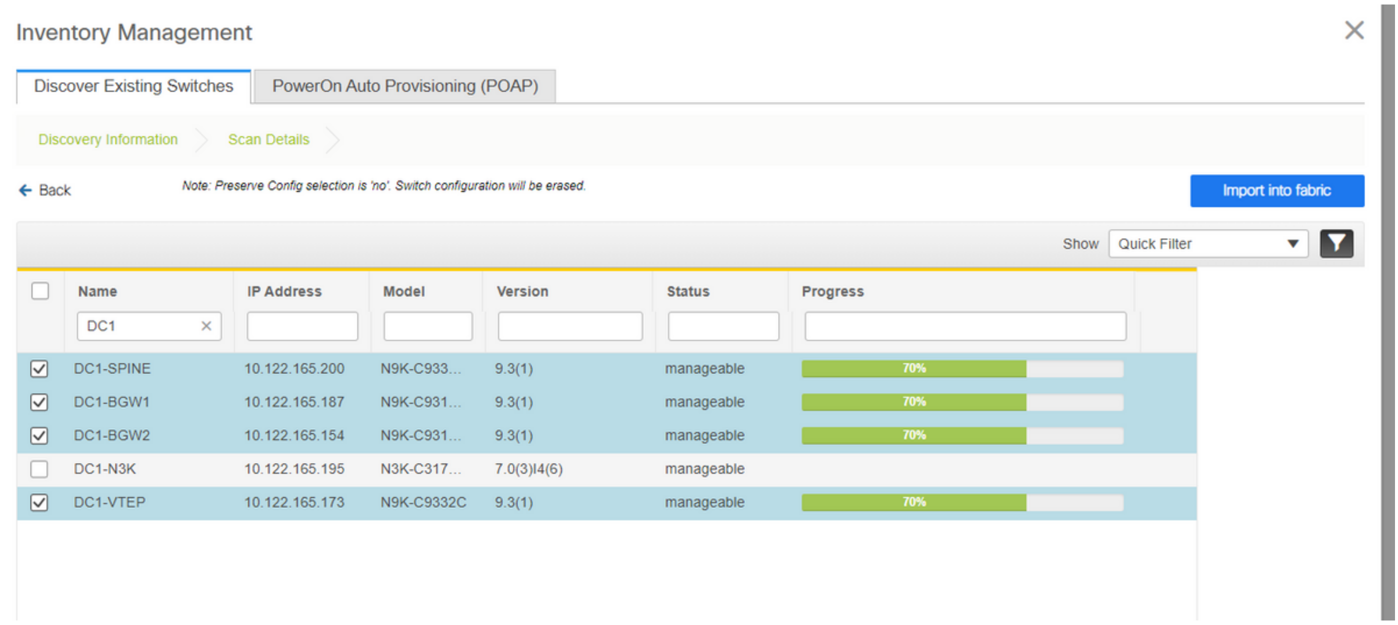
Note: Preserve Config selection is 'no'. Switch configuration will be erased.

□	Name	IP Address	Model	Version	Status	Progress
	DC1					
☑	DC1-SPINE	10.122.165.200	N9K-C933...	9.3(1)	manageable	
☑	DC1-BGW1	10.122.165.187	N9K-C931...	9.3(1)	manageable	
☑	DC1-BGW2	10.122.165.154	N9K-C931...	9.3(1)	manageable	
□	DC1-N3K	10.122.165.195	N3K-C317...	7.0(3)I4(6)	manageable	
☑	DC1-VTEP	10.122.165.173	N9K-C9332C	9.3(1)	manageable	

# Select the relevant switches and then Click on "Import into fabric"



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

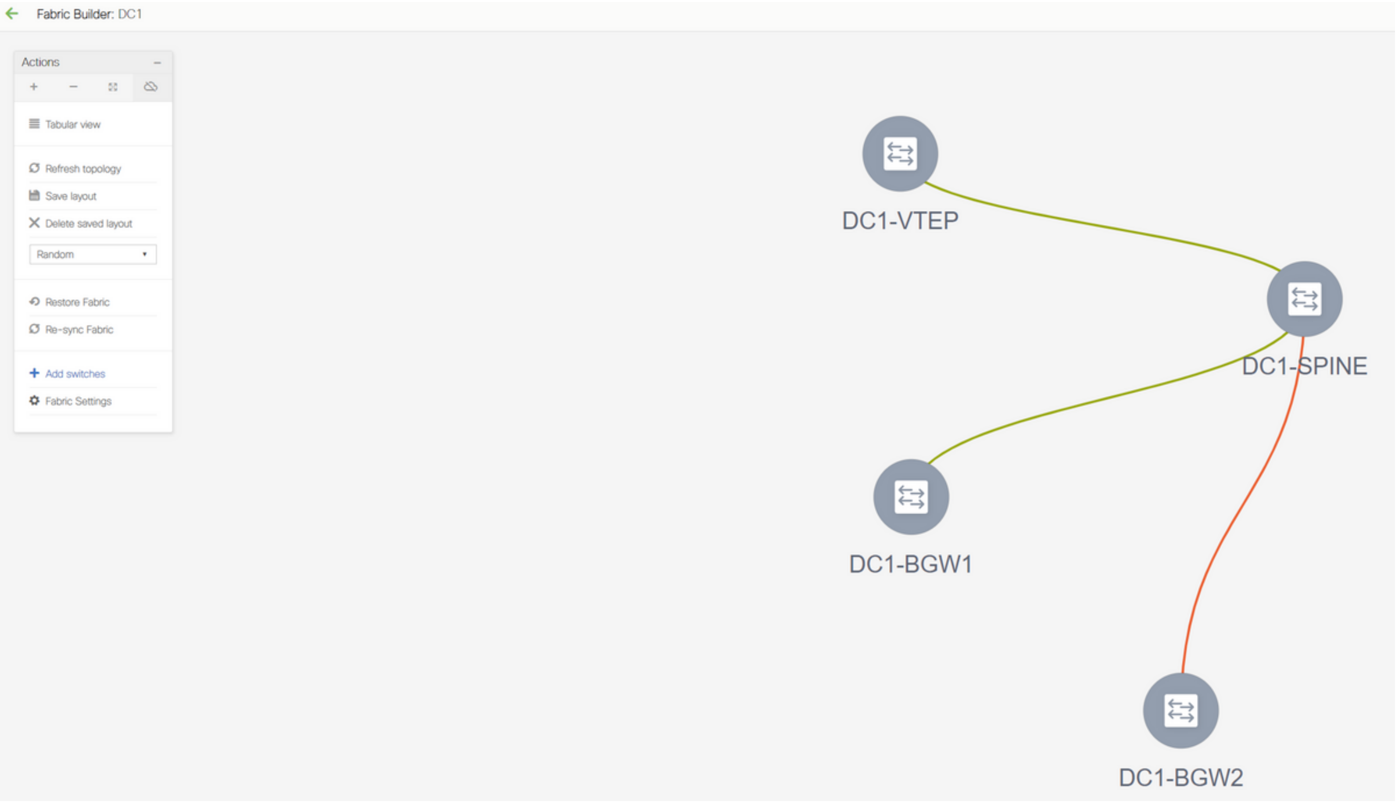


Inventory Management

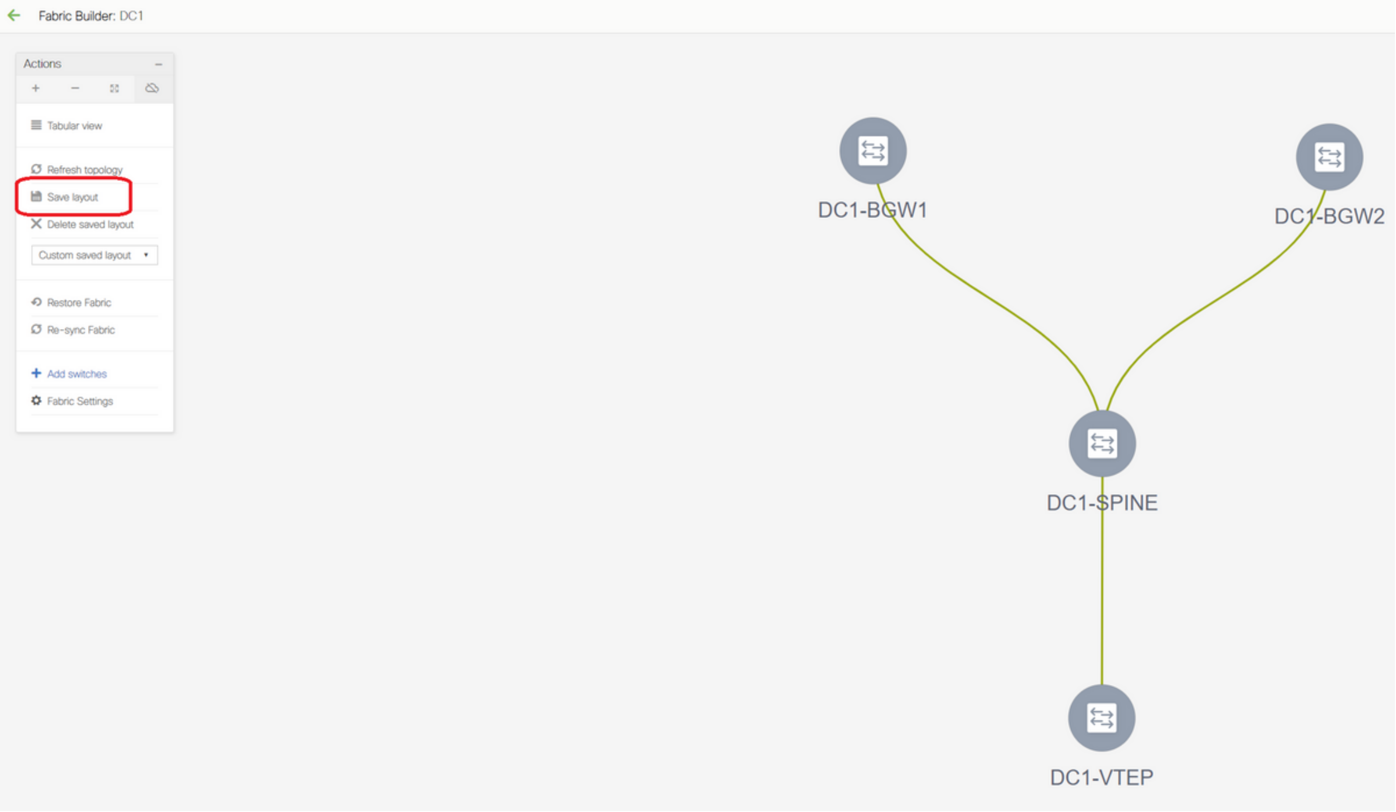
Note: Preserve Config selection is 'no'. Switch configuration will be erased.

□	Name	IP Address	Model	Version	Status	Progress
	DC1					
☑	DC1-SPINE	10.122.165.200	N9K-C933...	9.3(1)	manageable	70%
☑	DC1-BGW1	10.122.165.187	N9K-C931...	9.3(1)	manageable	70%
☑	DC1-BGW2	10.122.165.154	N9K-C931...	9.3(1)	manageable	70%
□	DC1-N3K	10.122.165.195	N3K-C317...	7.0(3)I4(6)	manageable	
☑	DC1-VTEP	10.122.165.173	N9K-C9332C	9.3(1)	manageable	70%

# Once the import is done, the topology under fabric builder may look like below;

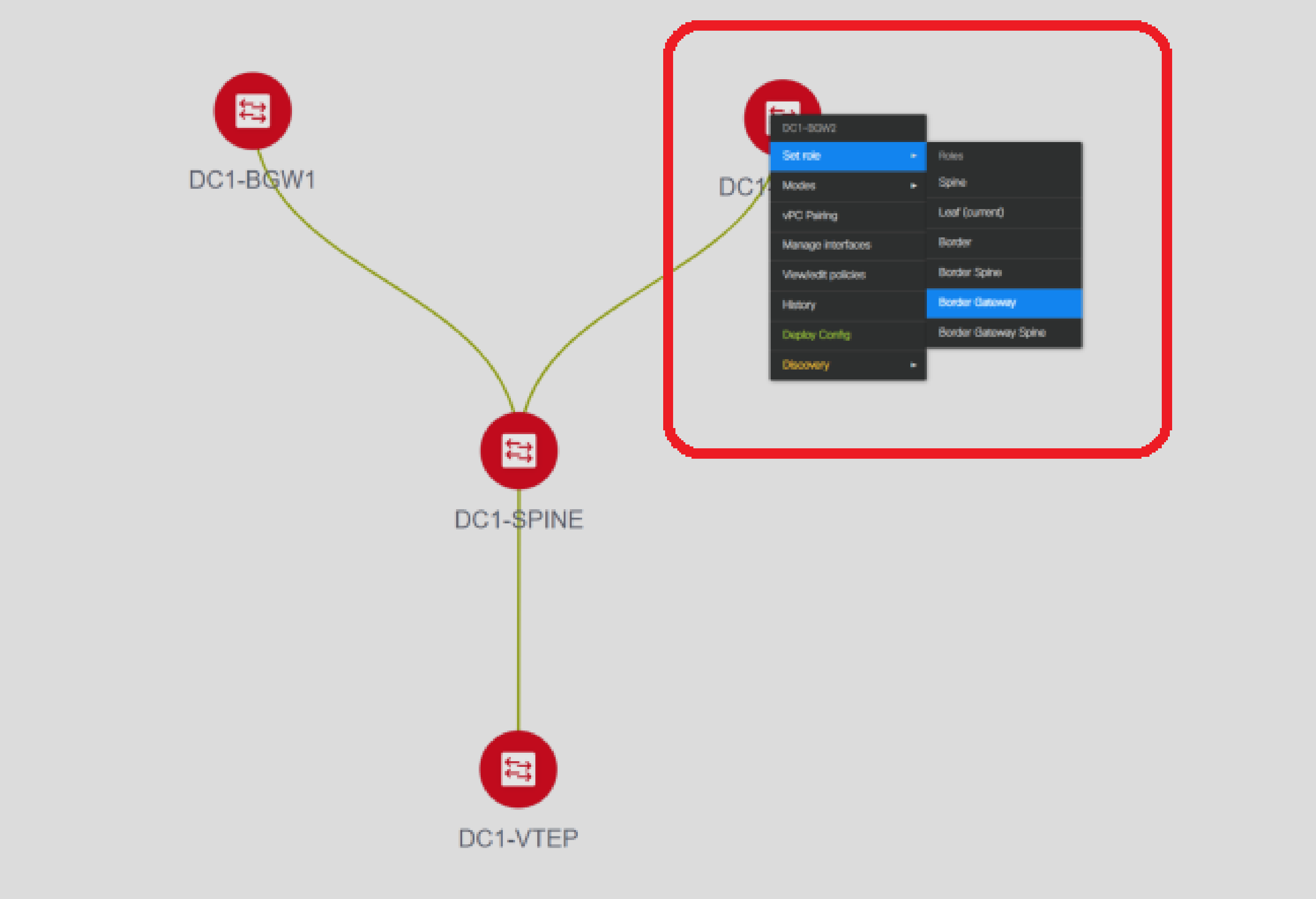


# The switches can be moved around by clicking on one switch and aligning it to the right location within the diagram



# Select the "save layout" section after rearranging the switches in the order the layout is needed

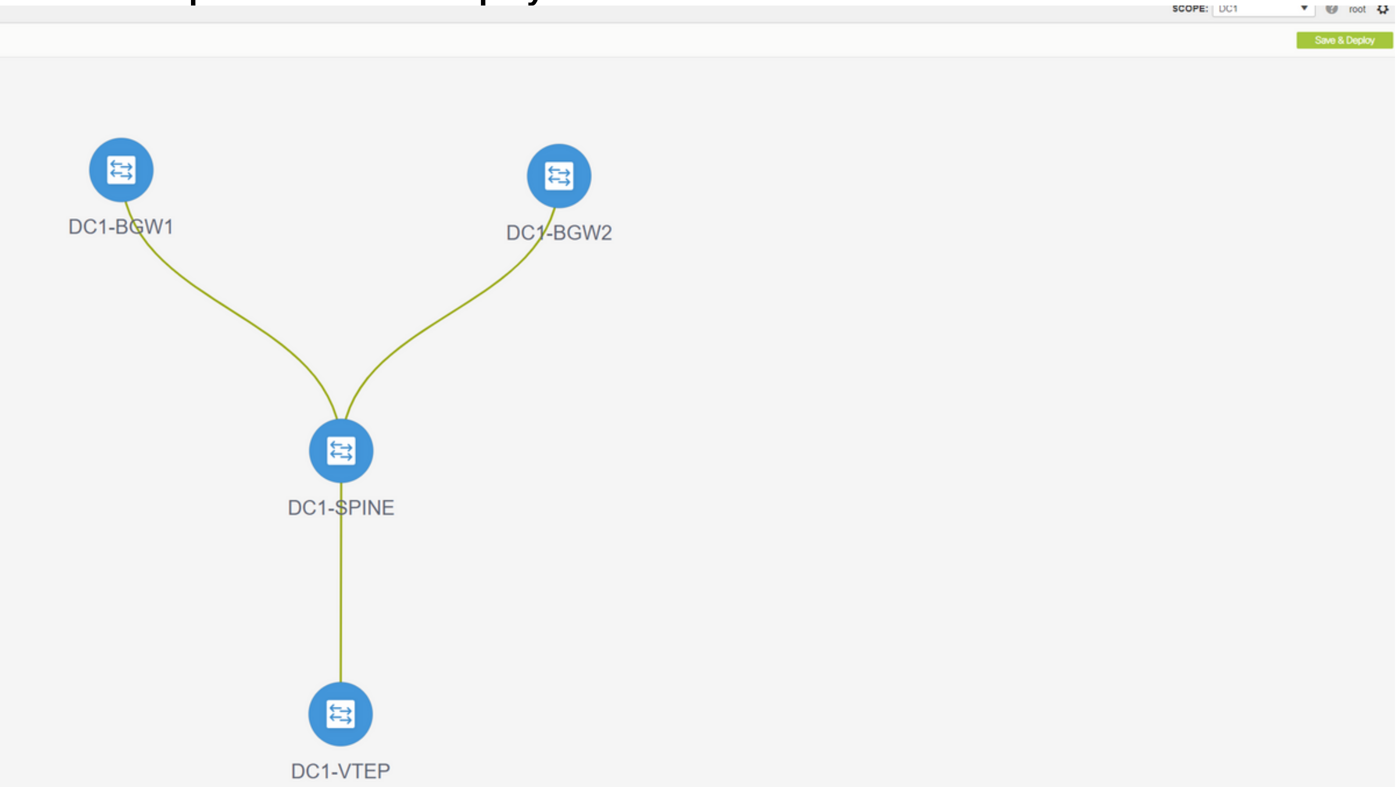
- **Setting roles for all switches**



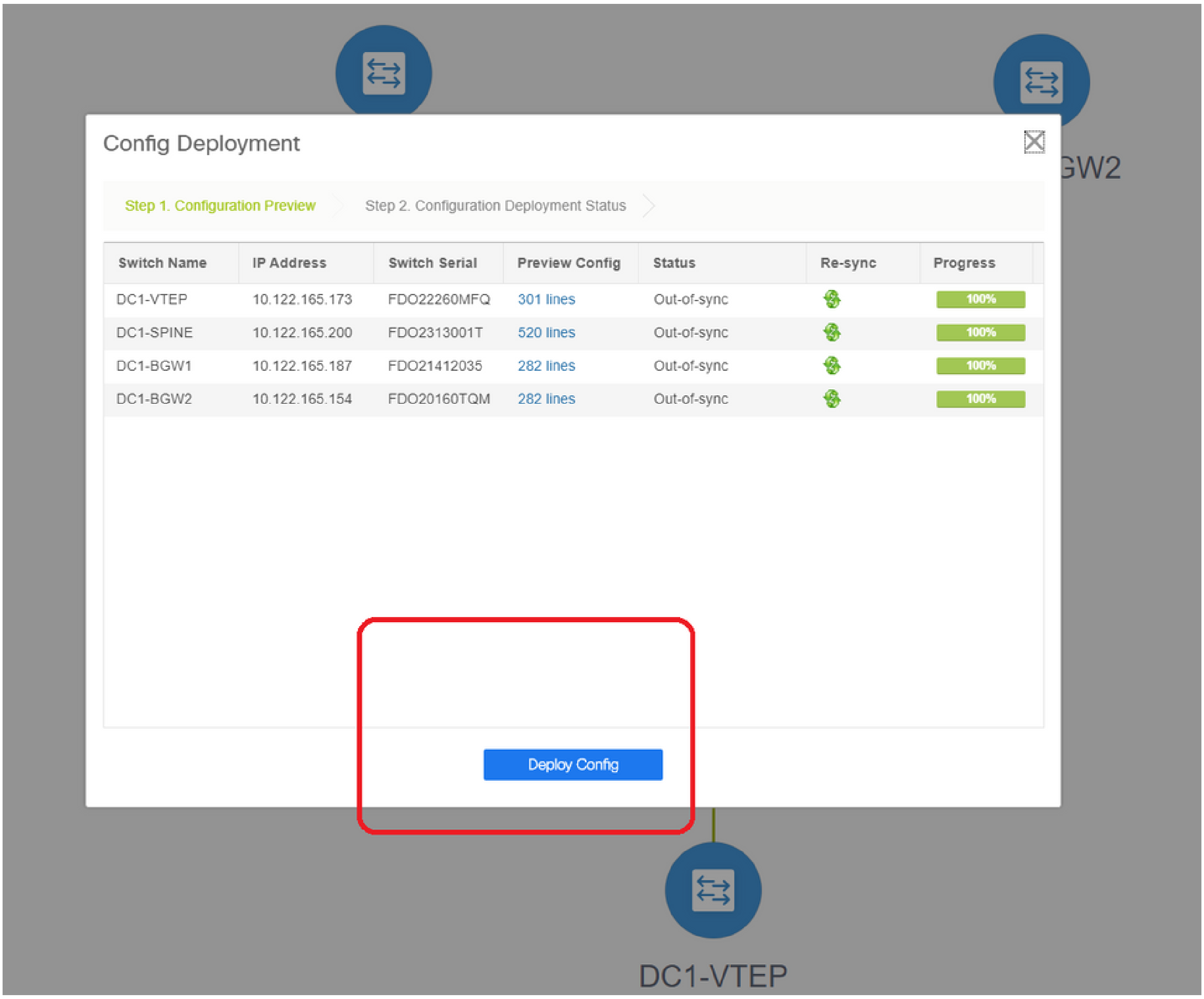
# Right Click each of the switches and set the right role; Here, DC1-BGW1 & DC1-BGW2 are the border gateways

# DC1-SPINE-> Will be set to role- Spine, DC1-VTEP-> Will be set to role-Leaf

- **Next Step is to save and Deploy**



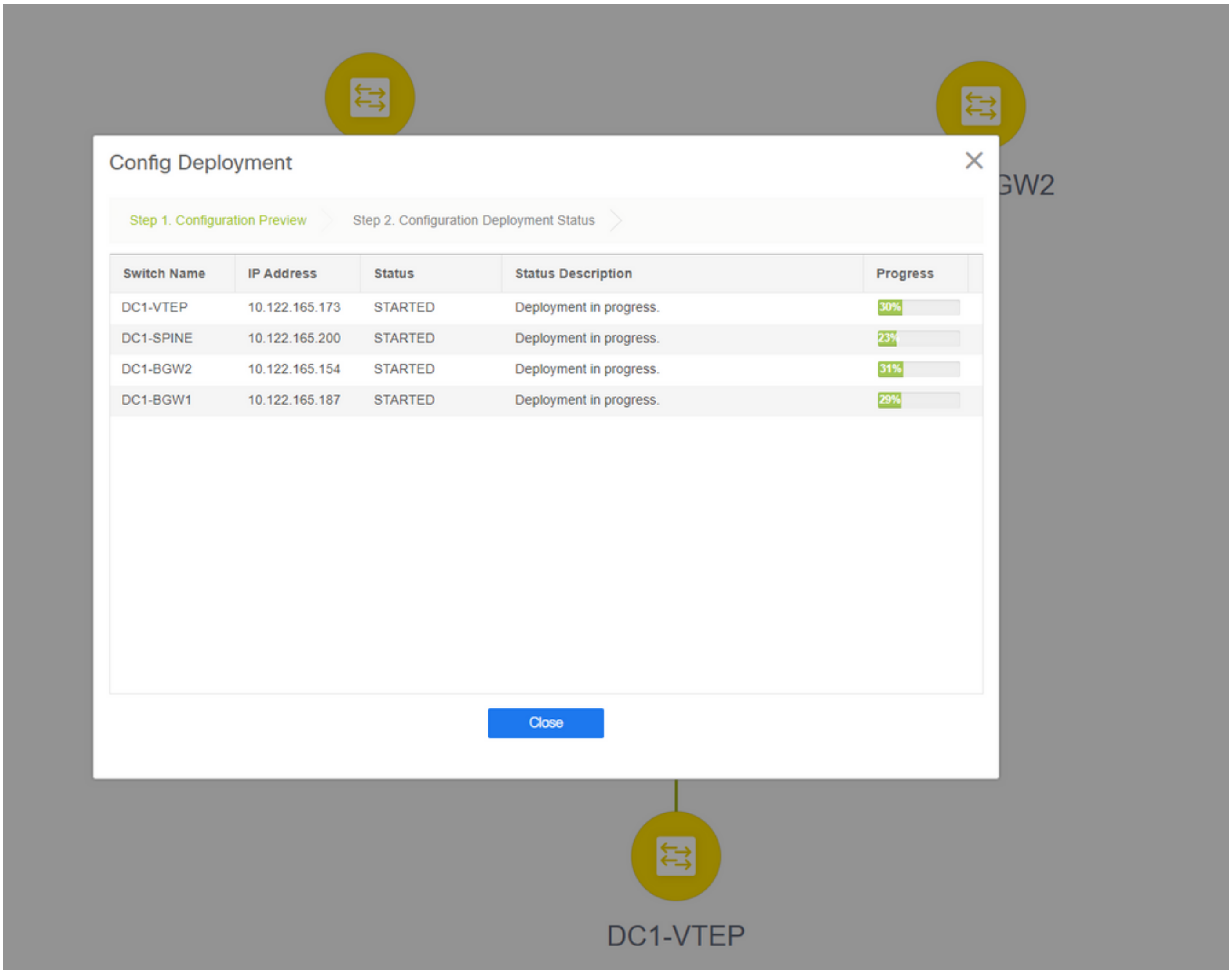
DCNM Will now list the switches and also will have the preview of the configurations that DCNM Is going to push to all switches.



Config Deployment

Step 1. Configuration Preview
Step 2. Configuration Deployment Status
Deploy Config
GW2
DC1-VTEP

Switch Name	IP Address	Switch Serial	Preview Config	Status	Re-sync	Progress
DC1-VTEP	10.122.165.173	FDO22260MFQ	301 lines	Out-of-sync	♻	100%
DC1-SPINE	10.122.165.200	FDO2313001T	520 lines	Out-of-sync	♻	100%
DC1-BGW1	10.122.165.187	FDO21412035	282 lines	Out-of-sync	♻	100%
DC1-BGW2	10.122.165.154	FDO20160TQM	282 lines	Out-of-sync	♻	100%

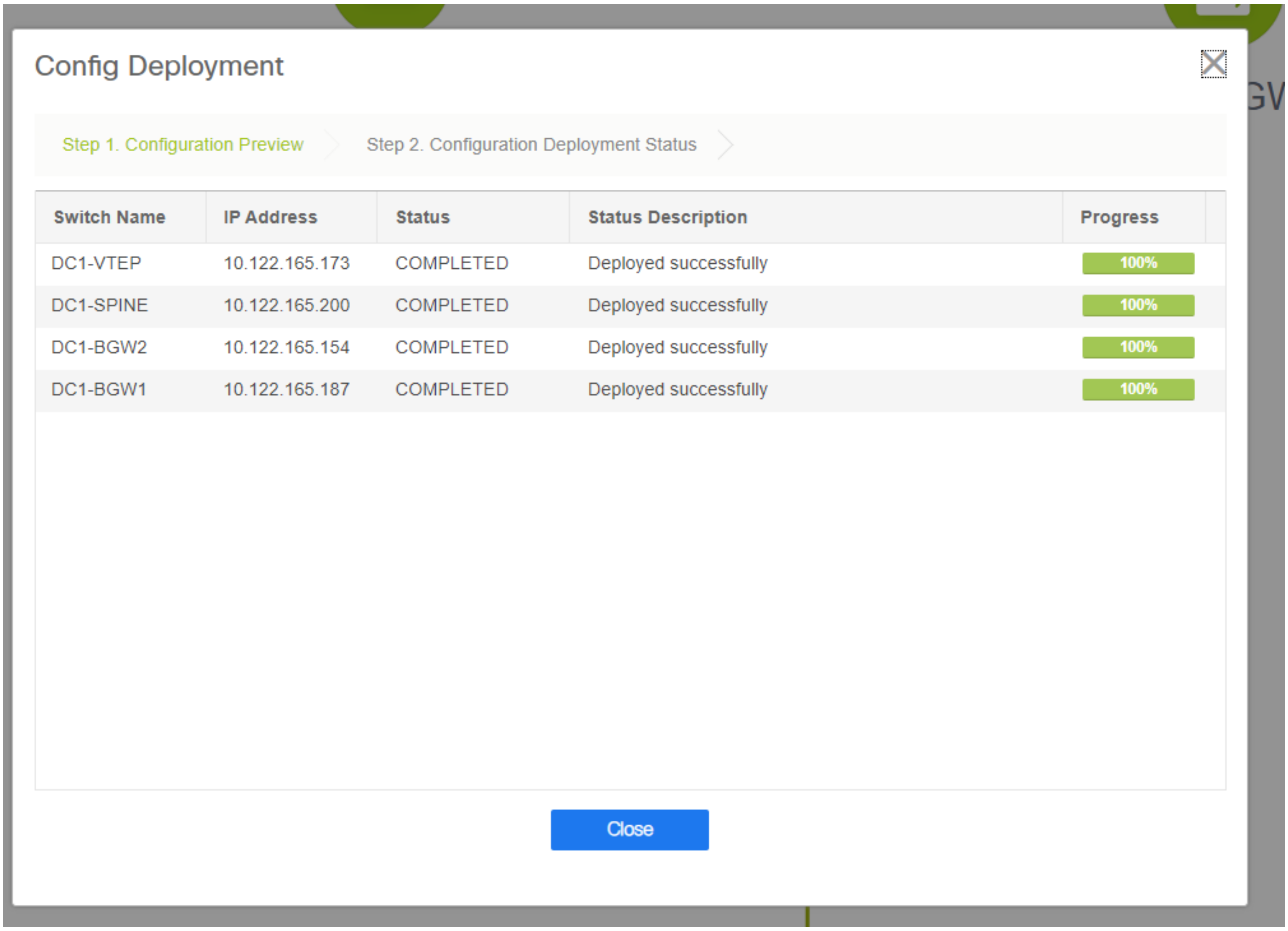


Config Deployment

Step 1. Configuration Preview   Step 2. Configuration Deployment Status

Switch Name	IP Address	Status	Status Description	Progress
DC1-VTEP	10.122.165.173	STARTED	Deployment in progress.	30%
DC1-SPINE	10.122.165.200	STARTED	Deployment in progress.	23%
DC1-BGW2	10.122.165.154	STARTED	Deployment in progress.	31%
DC1-BGW1	10.122.165.187	STARTED	Deployment in progress.	29%

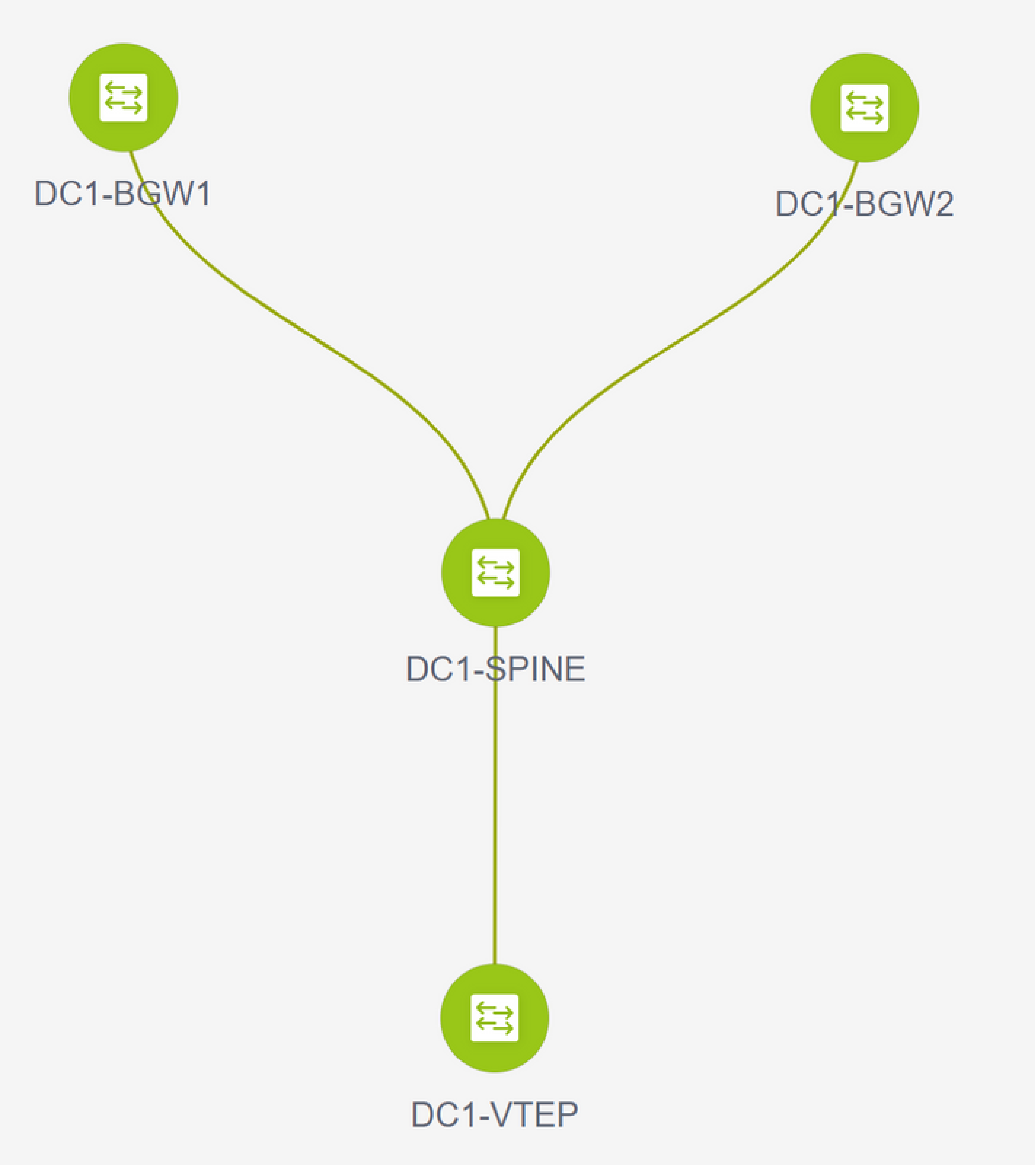
# Once it is successful, the status will reflect and also the switches will be shown in Green



Config Deployment

Step 1. Configuration Preview
Step 2. Configuration Deployment Status
Close

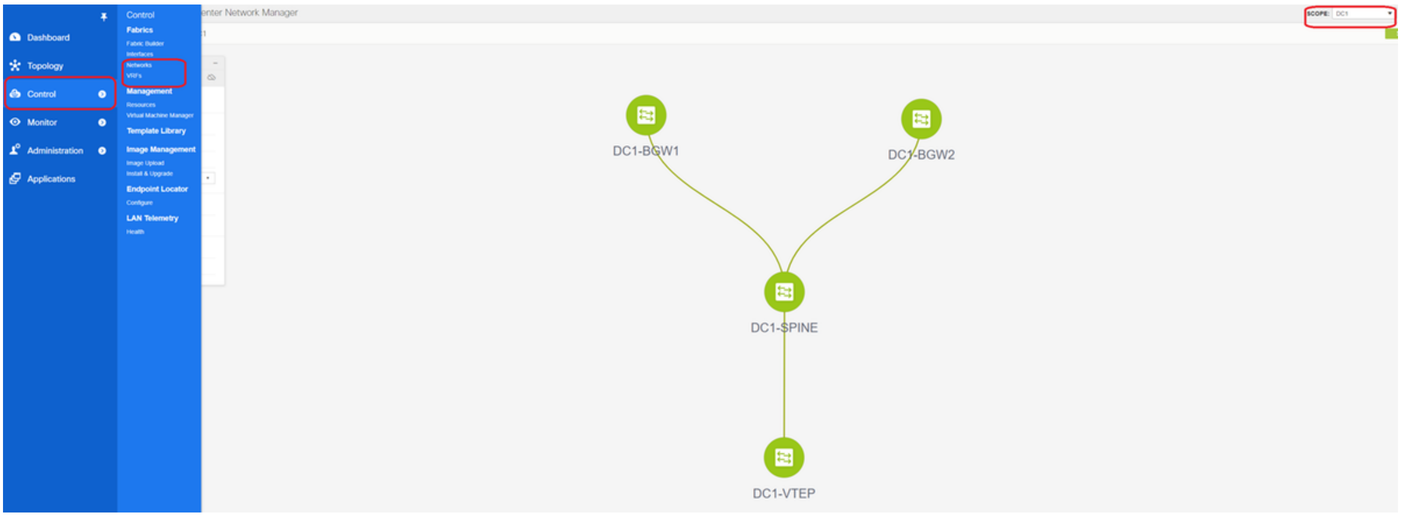
Switch Name	IP Address	Status	Status Description	Progress
DC1-VTEP	10.122.165.173	COMPLETED	Deployed successfully	100%
DC1-SPINE	10.122.165.200	COMPLETED	Deployed successfully	100%
DC1-BGW2	10.122.165.154	COMPLETED	Deployed successfully	100%
DC1-BGW1	10.122.165.187	COMPLETED	Deployed successfully	100%



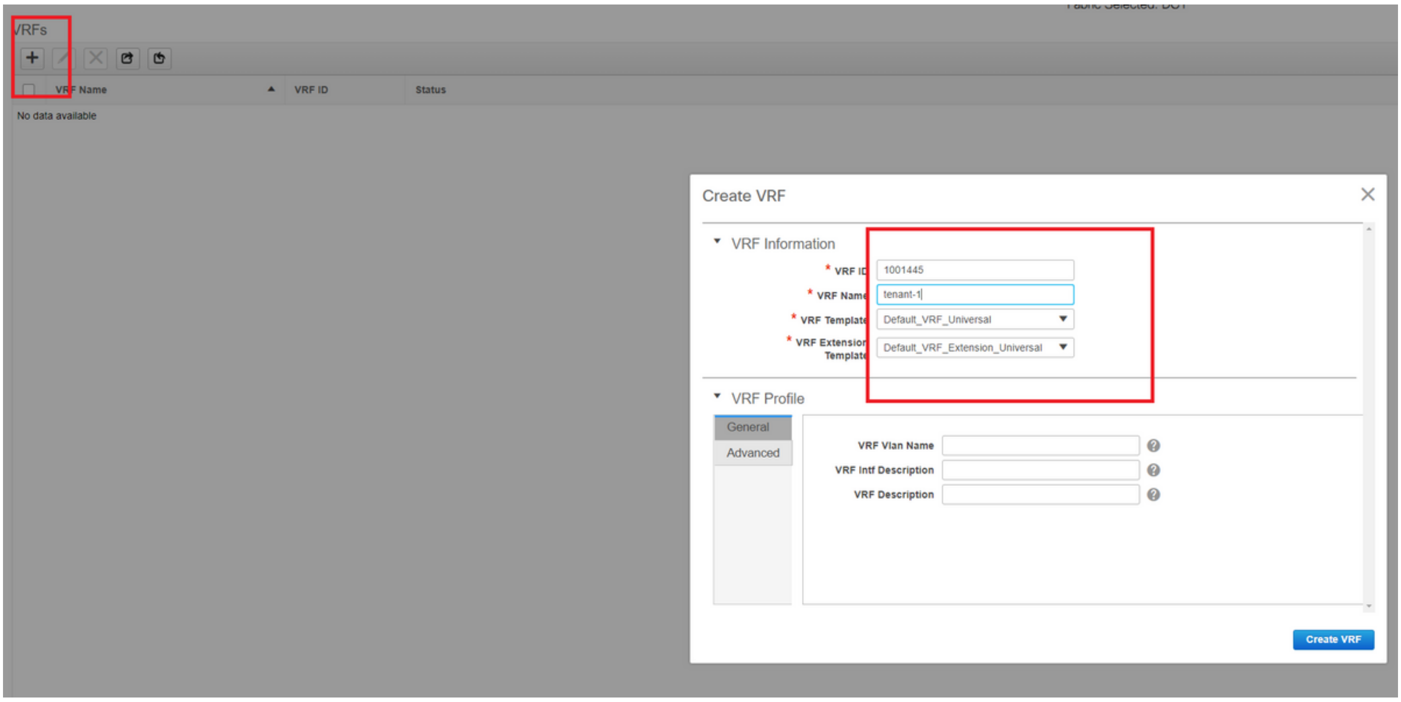
Step 3: Configuration of Networks/VRFs

- **Configuration of Networks/VRFs**

# Select DC1 Fabric(from the top right drop down), Control > VRFs



# Next is to Create VRF



* VRF ID: 1001445

* VRF Name: tenant-1

* VRF Template: Default_VRF_Universal

* VRF Extension Template: Default_VRF_Extension_Universal

VRF Vlan Name:

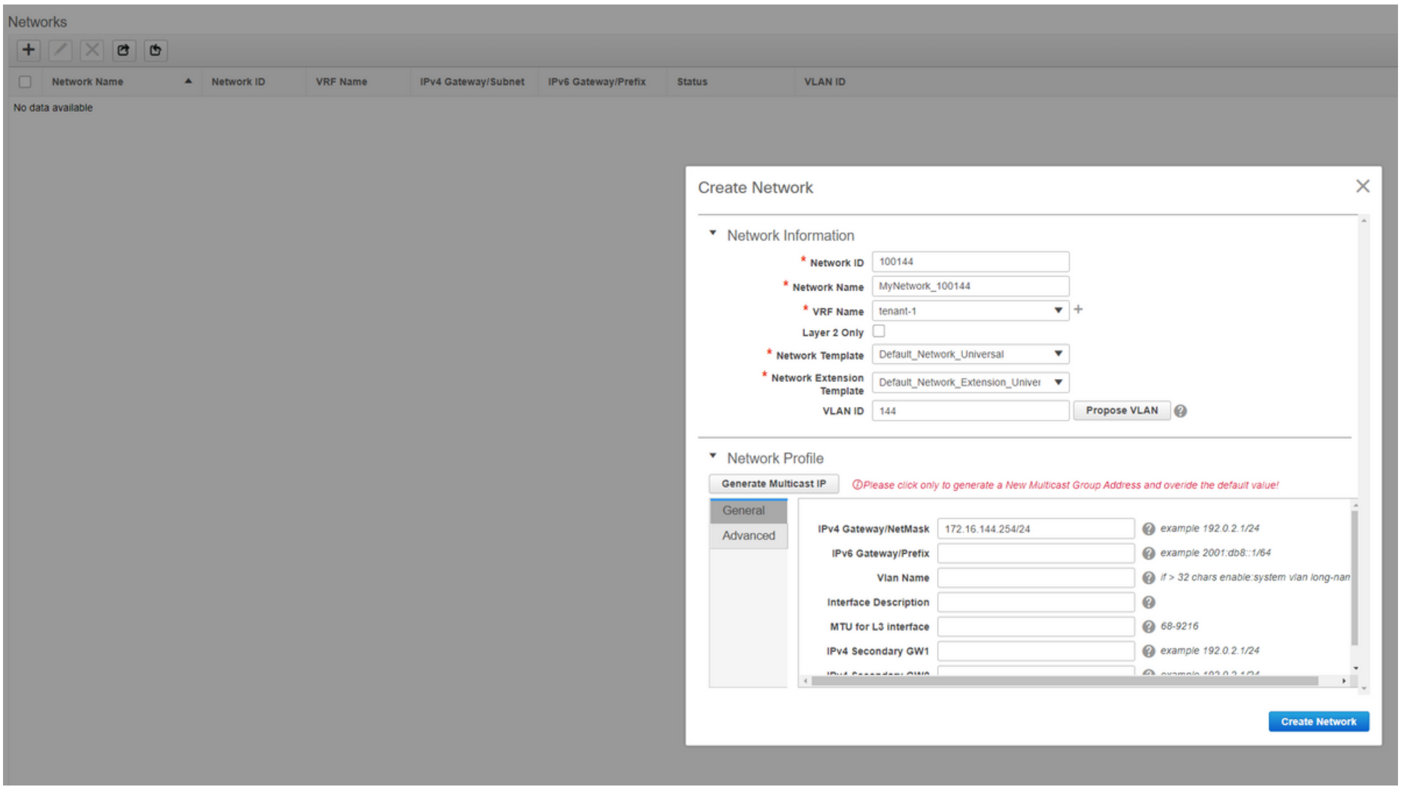
VRF Intf Description:

VRF Description:

# 11.2 DCNM Version is auto-populating the VRF ID; If its Different, type in the one that you need and select the "Create VRF"

# Here, the Layer 3 VNID used is 1001445

- **Next step is to create the Networks**



Networks

☐	Network Name ▲	Network ID	VRF Name	IPv4 Gateway/Subnet	IPv6 Gateway/Prefix	Status	VLAN ID

No data available

Create Network

▼ Network Information

* Network ID	100144
* Network Name	MyNetwork_100144
* VRF Name	tenant-1
Layer 2 Only	☐
* Network Template	Default_Network_Universal
* Network Extension Template	Default_Network_Extension_Univer
VLAN ID	144

▼ Network Profile

ⓘPlease click only to generate a New Multicast Group Address and overide the default value!

IPv4 Gateway/NetMask	172.16.144.254/24	example 192.0.2.1/24
IPv6 Gateway/Prefix		example 2001:db8::1/64
Vlan Name		if > 32 chars enable:system vlan long-nam
Interface Description		
MTU for L3 interface		68-9216
IPv4 Secondary GW1		example 192.0.2.1/24

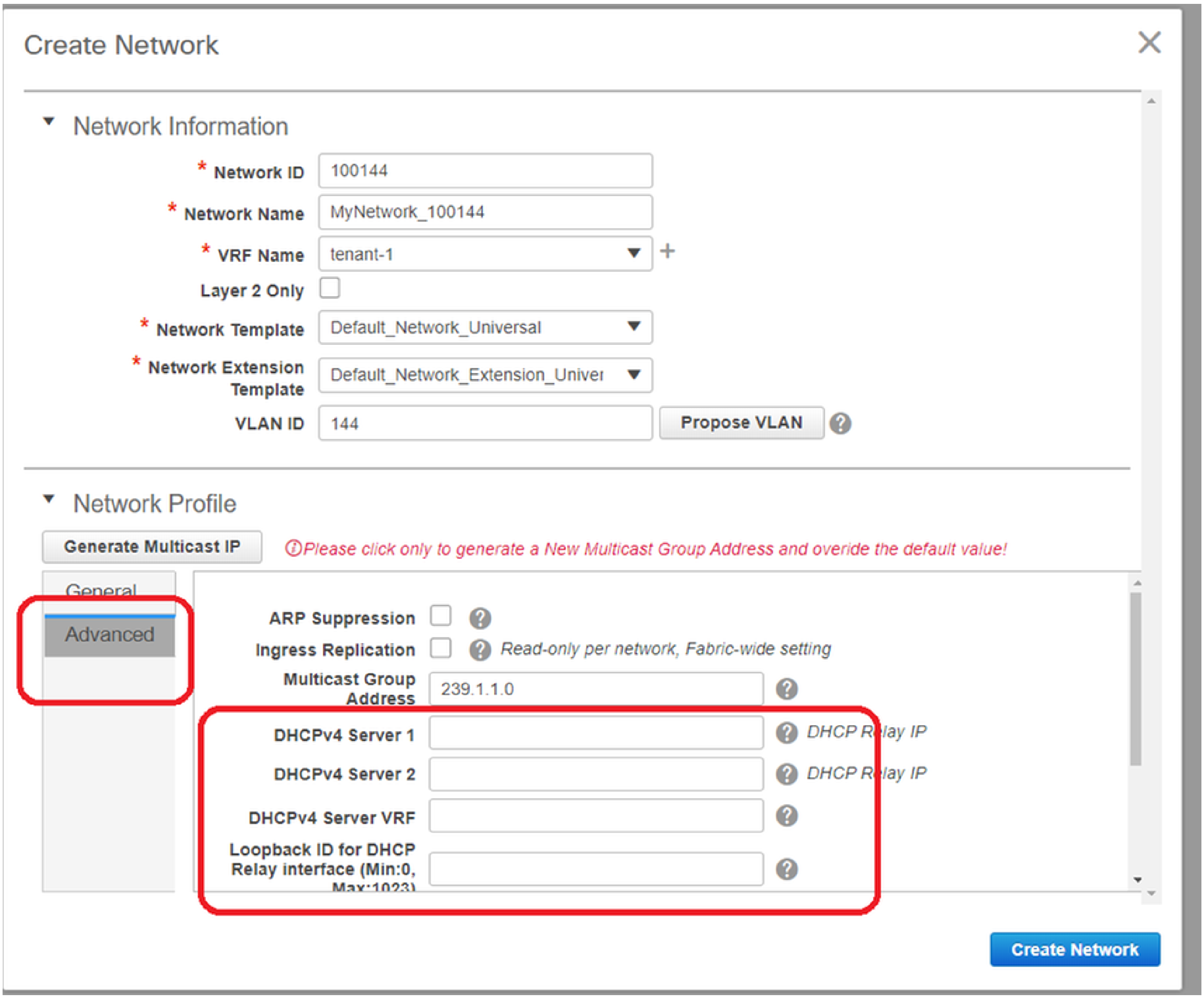
# Provide the Network ID(Which is the Layer 2 Vlans' corresponding VNID

# Provide the VRF that the SVI should be part of; By default, DCNM 11.2 populates the VRF Name to the previously created one; Change as needed

# Vlan ID will be Layer 2 VLan that is mapped to this particular VNID

# IPv4 Gateway-> This is the Anycast Gateway IP address which will be configured on the SVI and will be the same for all the VTEPs within the fabric

- **Advanced Tab has extra rows that need to be filled if eg; DHCP Relay is using;**



Create Network

▼ Network Information

* Network ID	100144
* Network Name	MyNetwork_100144
* VRF Name	tenant-1 ▼ +
Layer 2 Only	□
* Network Template	Default_Network_Universal ▼
* Network Extension Template	Default_Network_Extension_Univer ▼
VLAN ID	144 Propose VLAN ?

▼ Network Profile

Generate Multicast IP ⓘPlease click only to generate a New Multicast Group Address and overide the default value!

General

Advanced

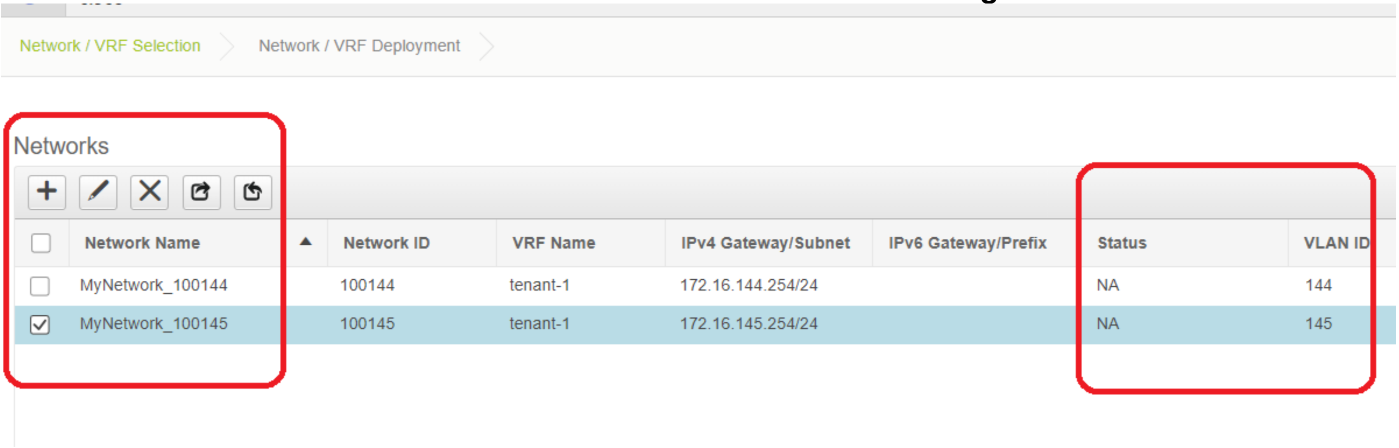
ARP Suppression	□ ?
Ingress Replication	□ ? Read-only per network, Fabric-wide setting
Multicast Group Address	239.1.1.0 ?
DHCPv4 Server 1	? DHCP Relay IP
DHCPv4 Server 2	? DHCP Relay IP
DHCPv4 Server VRF	?
Loopback ID for DHCP Relay interface (Min:0, Max:1023)	?

Create Network

# Once the fields are populated, click on "create Network".

# Create any other Networks that are required to be part of this fabric;

- **At this moment, VRF and Networks are just defined in DCNM; but not pushed from DCNM to the switches in the fabric. This can be verified using below**



Network / VRF Selection 〉 Network / VRF Deployment 〉

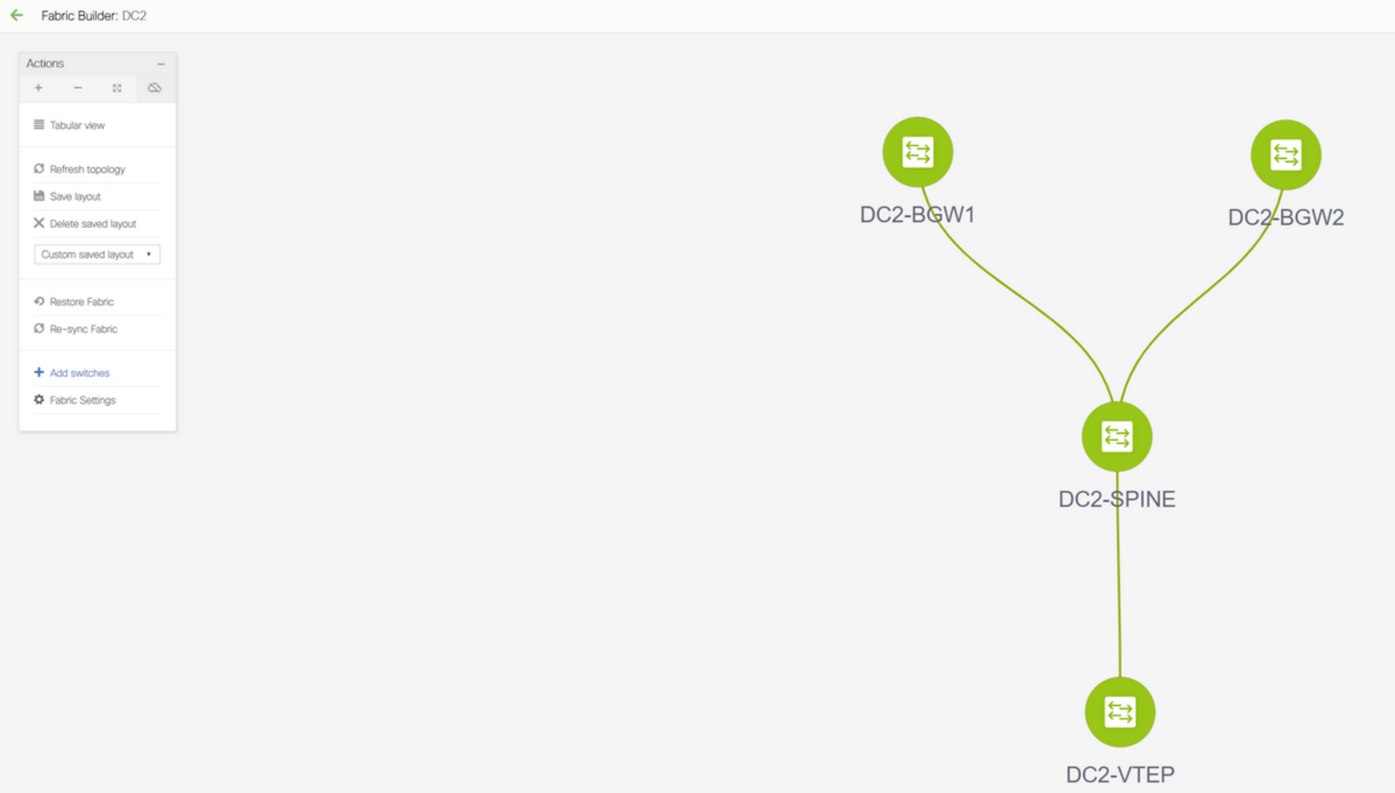
Networks

□	Network Name ▲	Network ID	VRF Name	IPv4 Gateway/Subnet	IPv6 Gateway/Prefix	Status	VLAN ID
□	MyNetwork_100144	100144	tenant-1	172.16.144.254/24		NA	144
☑	MyNetwork_100145	100145	tenant-1	172.16.145.254/24		NA	145

# Status will be in "NA" if this is NOT Deployed to the switches. Since this is a multisite and involves Border Gateways, Deployment of Networks/VRFs will be discussed further down.

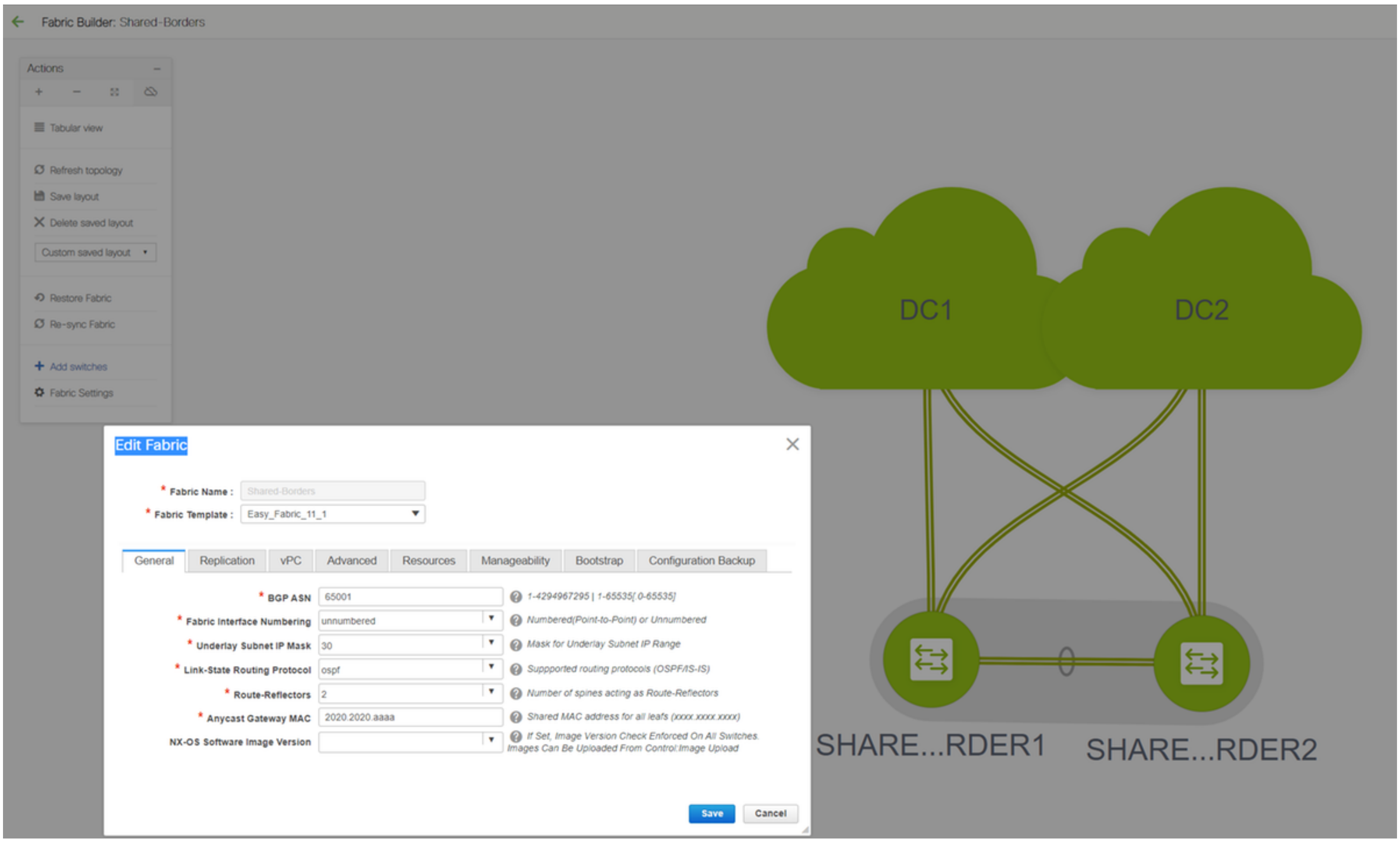
Step 4: Repeat the same steps for DC2

- Now that the DC1 is fully defined, will carry out the same procedure for DC2 as well
- Once DC2 is fully defined, it will look like below



Step 5: Creation of easy fabric for Shared borders

- This is where another easy fabric is created which will include the shared borders which are in vPC
- Note that the Shared borders while deploying via DCNM Should be configured as vPC as otherwise, the inter-switch links will be shut down after a "re-sync" operation is performed on DCNM
- The Switches in Shared borders are to be set with the role of "Border"



Edit Fabric

* Fabric Name : Shared-Borders

* Fabric Template : Easy_Fabric_11_1

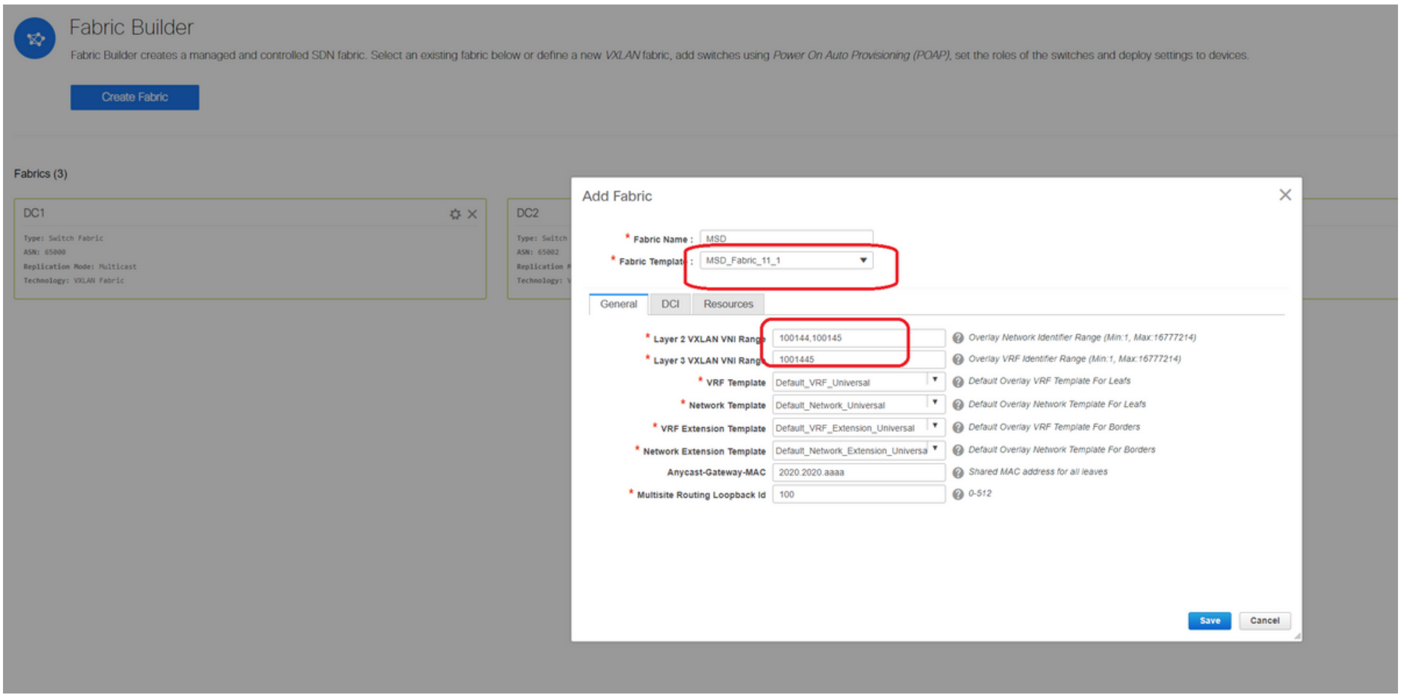
* BGP ASN	65001	1-4294967295 | 1-65535[.0-65535]
* Fabric Interface Numbering	unnumbered	Numbered(Point-to-Point) or Unnumbered
* Underlay Subnet IP Mask	30	Mask for Underlay Subnet IP Range
* Link-State Routing Protocol	ospf	Suppported routing protocols (OSPF/IS-IS)
* Route-Reflectors	2	Number of spines acting as Route-Reflectors
* Anycast Gateway MAC	2020.2020.aaaa	Shared MAC address for all leafs (xxxx.xxxx.xxxx)
NX-OS Software Image Version		If Set, Image Version Check Enforced On All Switches. Images Can Be Uploaded From Control:Image Upload

# VRFs are also created like it was for DC1 and DC2 fabrics

# Networks are not required on a shared border as the Shared border will not have any Layer 2 Vlans/VNIDs; Shared borders are not a tunnel termination for any East/West Traffic from DC1 to DC2; Only the Border Gateways would play a role in terms of vxlan encapsulation/decapsulation for East/West DC1<>DC2 traffic

Step 6- Creation of MSD and Moving DC1 and DC2 Fabrics

Go to Fabric builder and create new Fabric and use the template -> MSD_Fabric_11_1



Fabric Builder creates a managed and controlled SDN fabric. Select an existing fabric below or define a new *VXLAN* fabric, add switches using *Power On Auto Provisioning (POAP)*, set the roles of the switches and deploy settings to devices.

DC1

Type: Switch Fabric
ASN: 65000
Replication Mode: Multicast
Technology: VXLAN Fabric

DC2

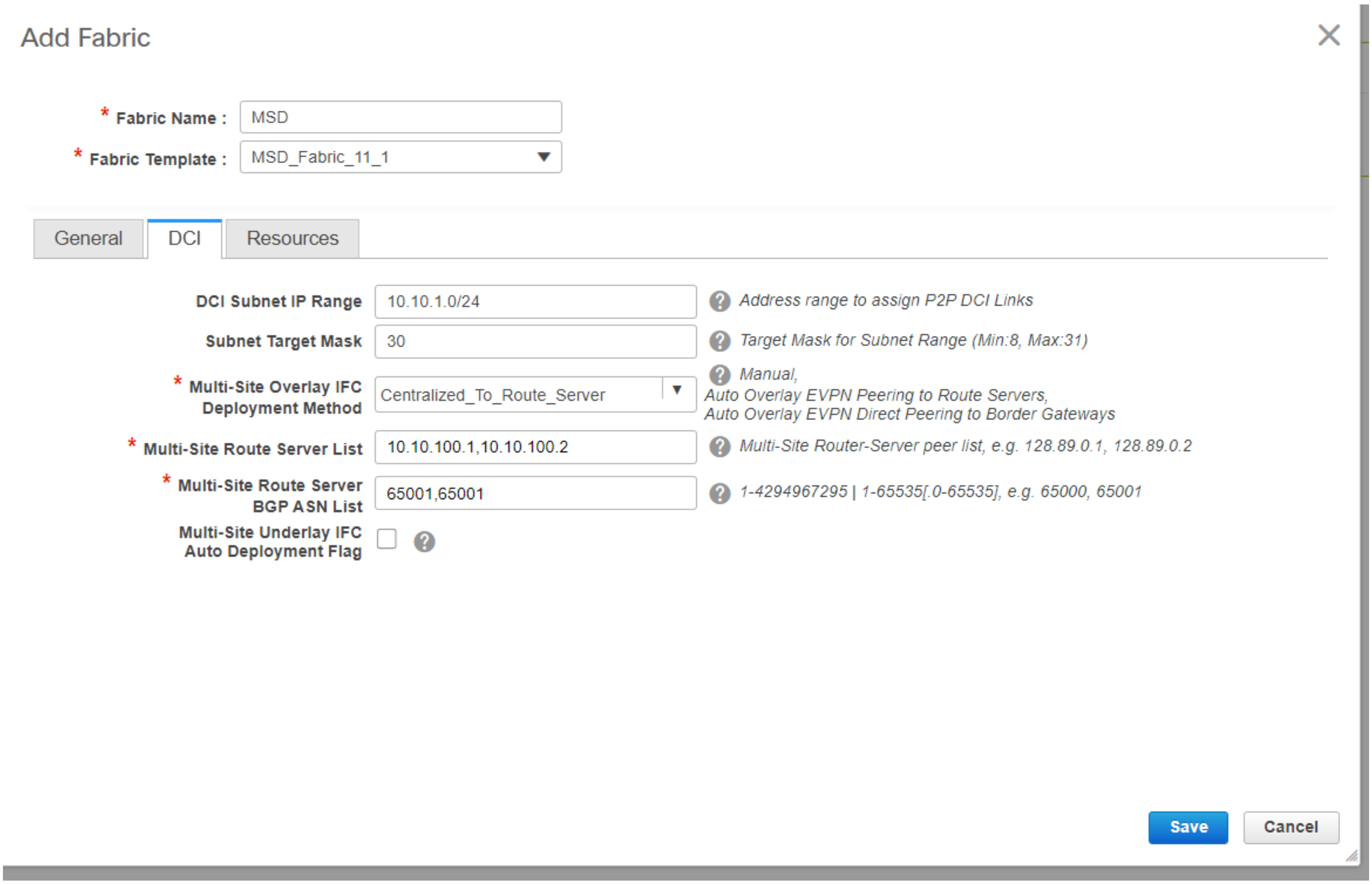
Type: Switch
ASN: 65002
Replication
Technology: V

Add Fabric

* Fabric Name : MSD

* Fabric Template : MSD_Fabric_11_1

* Layer 2 VXLAN VNI Range	100144,100145	Overlay Network Identifier Range (Min:1, Max:16777214)
* Layer 3 VXLAN VNI Range	1001445	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_Universa	Default Overlay Network Template For Borders
Anycast-Gateway-MAC	2020.2020.aaaa	Shared MAC address for all leaves
* Multisite Routing Loopback Id	100	0-512



Add Fabric

* Fabric Name : MSD

* Fabric Template : MSD_Fabric_11_1

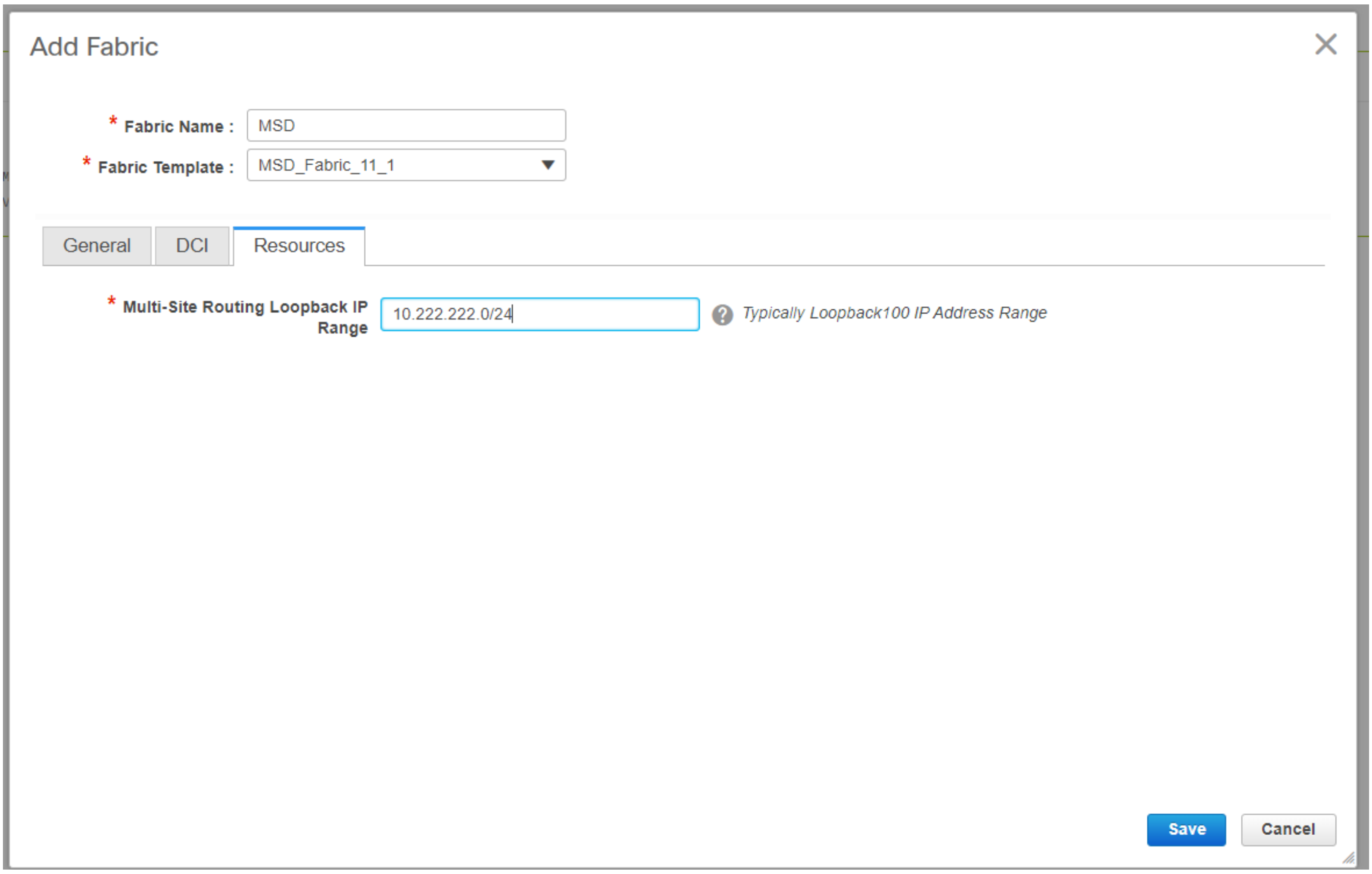
DCI Subnet IP Range	10.10.1.0/24	*Address range to assign P2P DCI Links*
Subnet Target Mask	30	*Target Mask for Subnet Range (Min:8, Max:31)*
* Multi-Site Overlay IFC Deployment Method	Centralized_To_Route_Server	*Manual, Auto Overlay EVPN Peering to Route Servers, Auto Overlay EVPN Direct Peering to Border Gateways*
* Multi-Site Route Server List	10.10.100.1,10.10.100.2	*Multi-Site Router-Server peer list, e.g. 128.89.0.1, 128.89.0.2*
* Multi-Site Route Server BGP ASN List	65001,65001	*1-4294967295 | 1-65535[.0-65535], e.g. 65000, 65001*
Multi-Site Underlay IFC Auto Deployment Flag	□	

# Note that Multi-site Overlay IFC Deployment Method has to be "**centralized_To_Route_Server**"; Here, the Shared borders are considered as Route servers and so this option is used from the drop down

# Within the "**Multisite Route Server List**"; Here, find out the Loopback IP addresses of Loopback0(which is the Routing loopback) on shared border and fill it up

# **ASN** is the one on shared border(Refer diagram on top of this document for more details); For the purpose of this document, both the shared borders are configured in the same ASN; Fill in accordingly

- **Next tab is where the Multisite Loopback IP range is provided as shown below**



Add Fabric

* Fabric Name : MSD

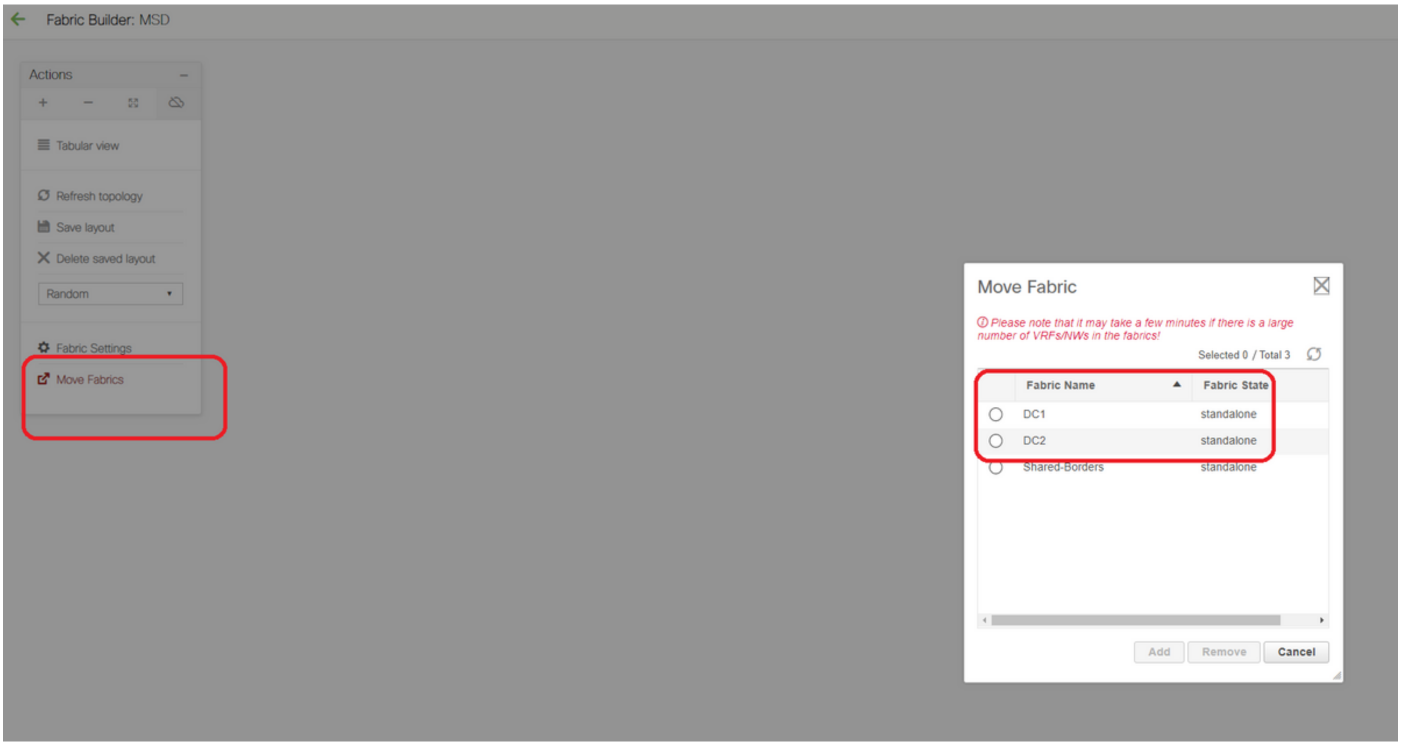
* Fabric Template : MSD_Fabric_11_1

* Multi-Site Routing Loopback IP Range: 10.222.222.0/24

Typically Loopback100 IP Address Range

# Once all the fields are populated, click the "save" button and a new Fabric will be created with the template-> MSD

# Next is to move the DC1 and DC2 fabrics to this MSD



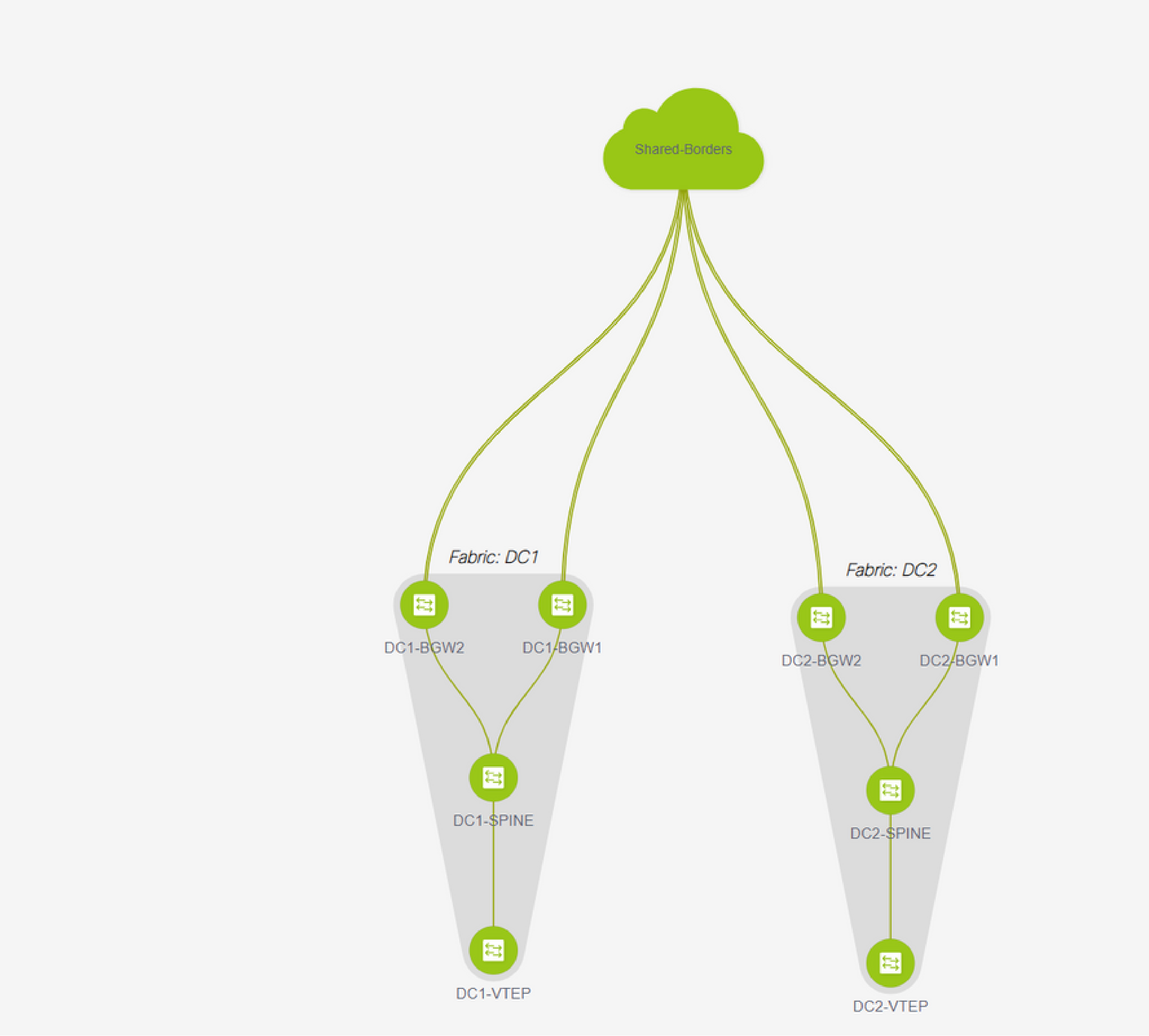
Move Fabric

Please note that it may take a few minutes if there is a large number of VRFs/NWs in the fabrics!

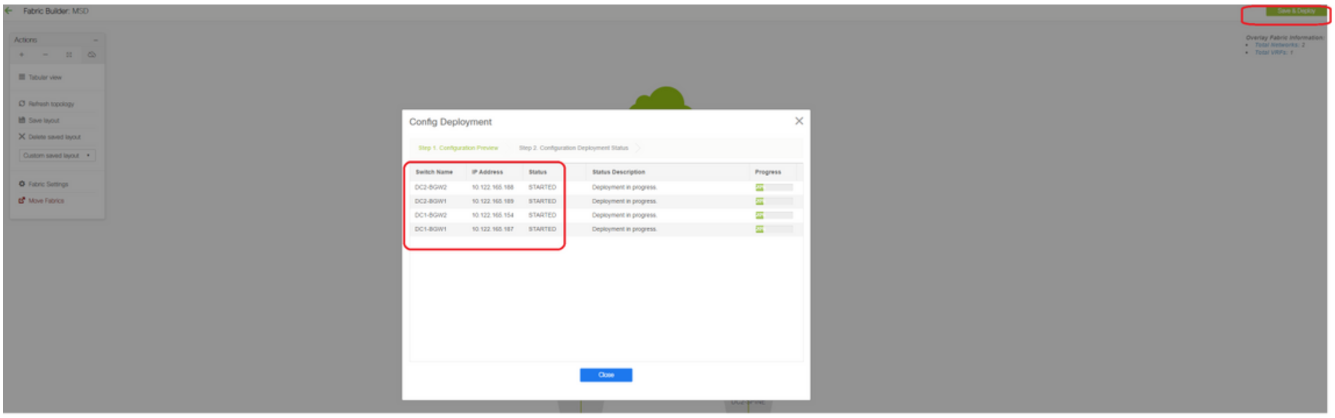
Selected 0 / Total 3

	Fabric Name	Fabric State
○	DC1	standalone
○	DC2	standalone
○	Shared-Borders	standalone

# After the fabric move, it looks like below



# Once done, click on the "save&Deploy" button which will push the required configurations as far as multisite is concerned to the border Gateways



Config Deployment

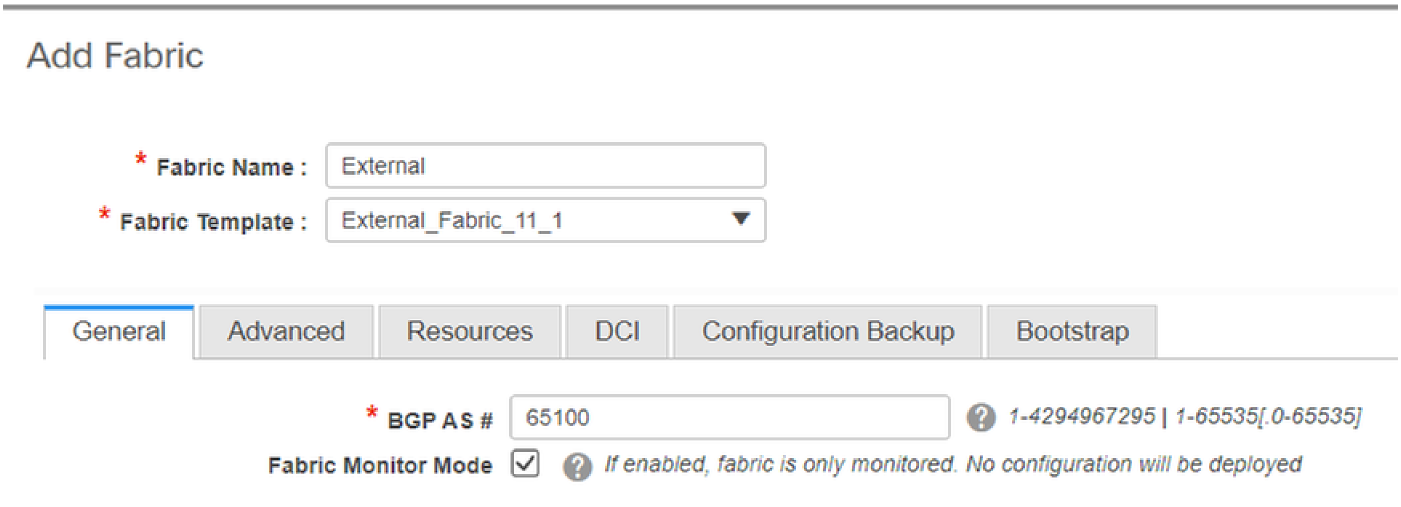
Step 1. Configuration Preview

Step 2. Configuration Deployment Status

Switch Name	IP Address	Status	Status Description	Progress
DC2-BGW2	10.122.165.188	STARTED	Deployment in progress.	
DC2-BGW1	10.122.165.189	STARTED	Deployment in progress.	
DC1-BGW2	10.122.165.154	STARTED	Deployment in progress.	
DC1-BGW1	10.122.165.187	STARTED	Deployment in progress.	

Step 7: Creation of External Fabric

# Create External Fabric and add the external router to it as shown below;



Add Fabric

* Fabric Name : External

* Fabric Template : External_Fabric_11_1

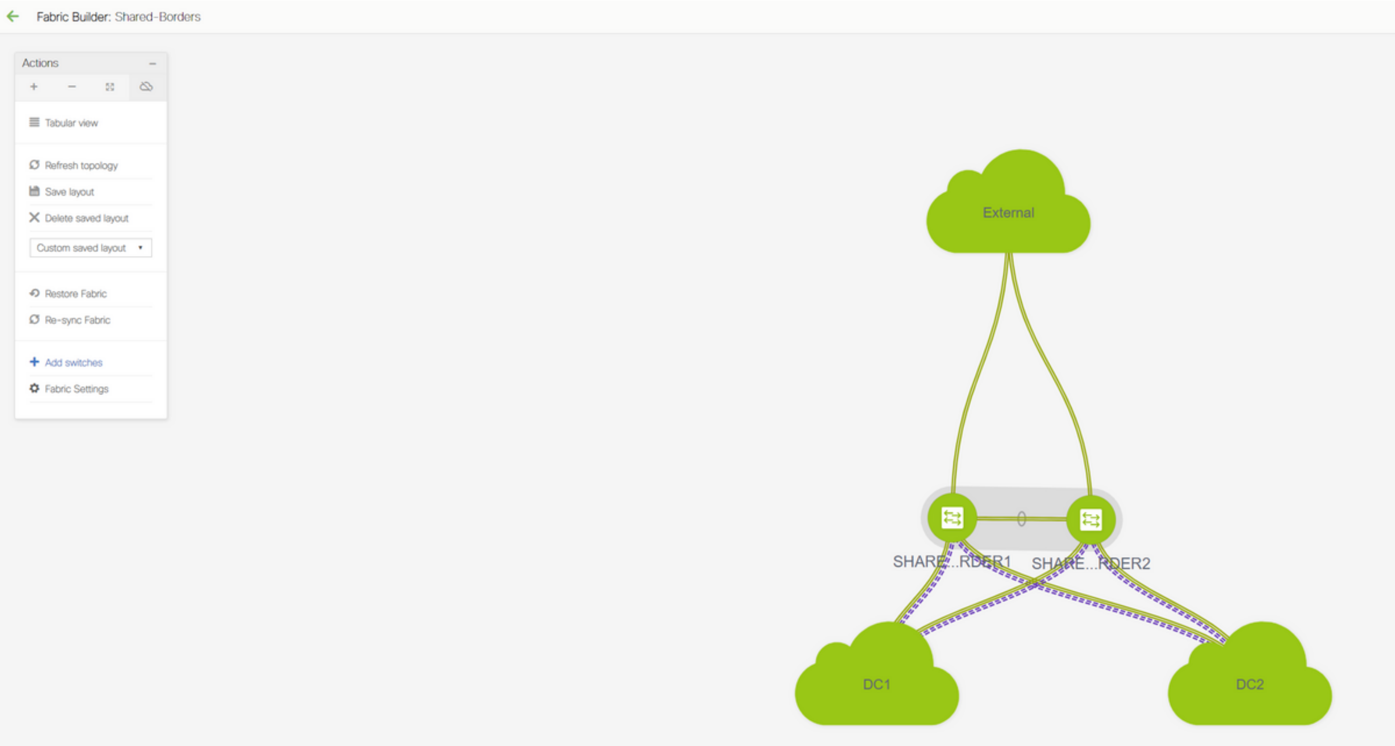
* BGP AS # 65100 *1-4294967295 | 1-65535[.0-65535]*

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

# Name the fabric and use the template-> "External_Fabric_11_1";

# Provide the ASN

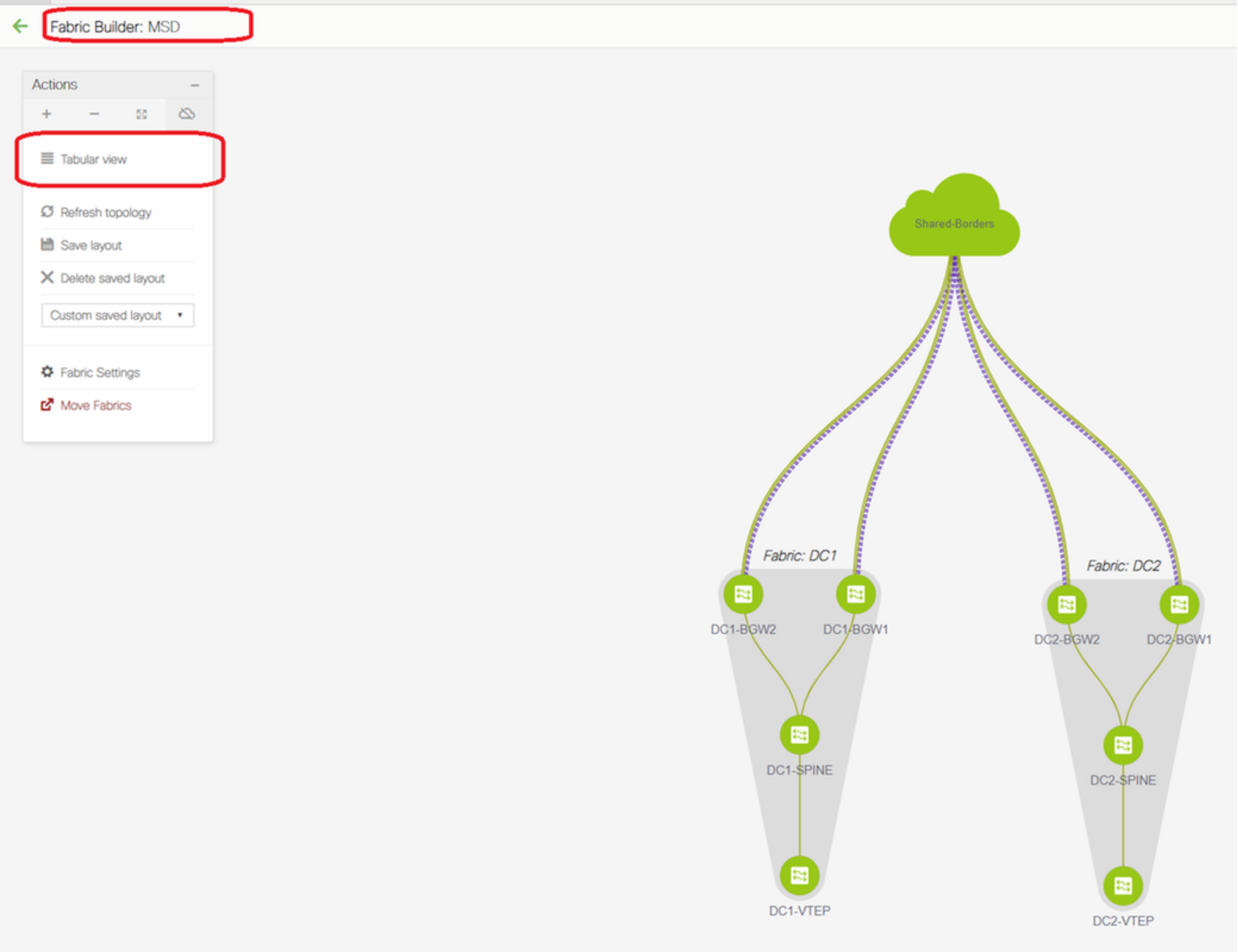
# At the end, the various fabrics will look like below



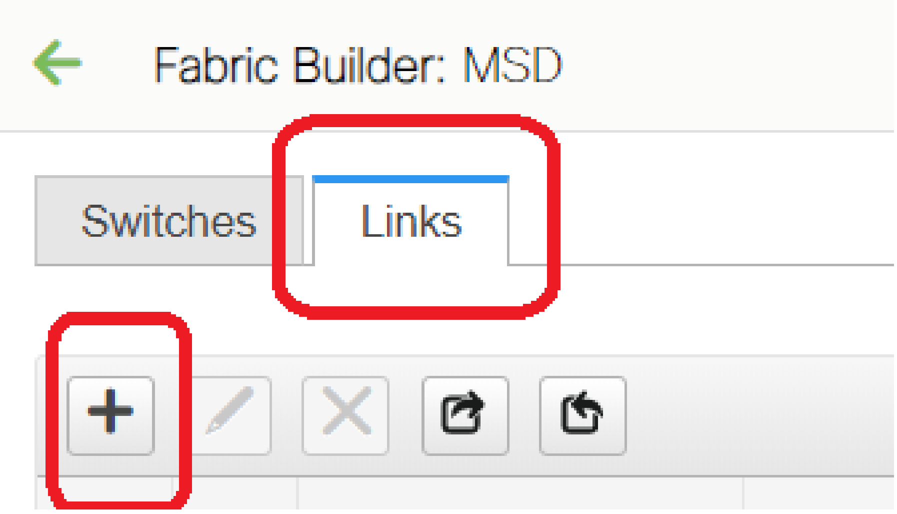
Step 8: eBGP Underlay for loopback reachability between BGWs(iBGP Between Shared borders as well)

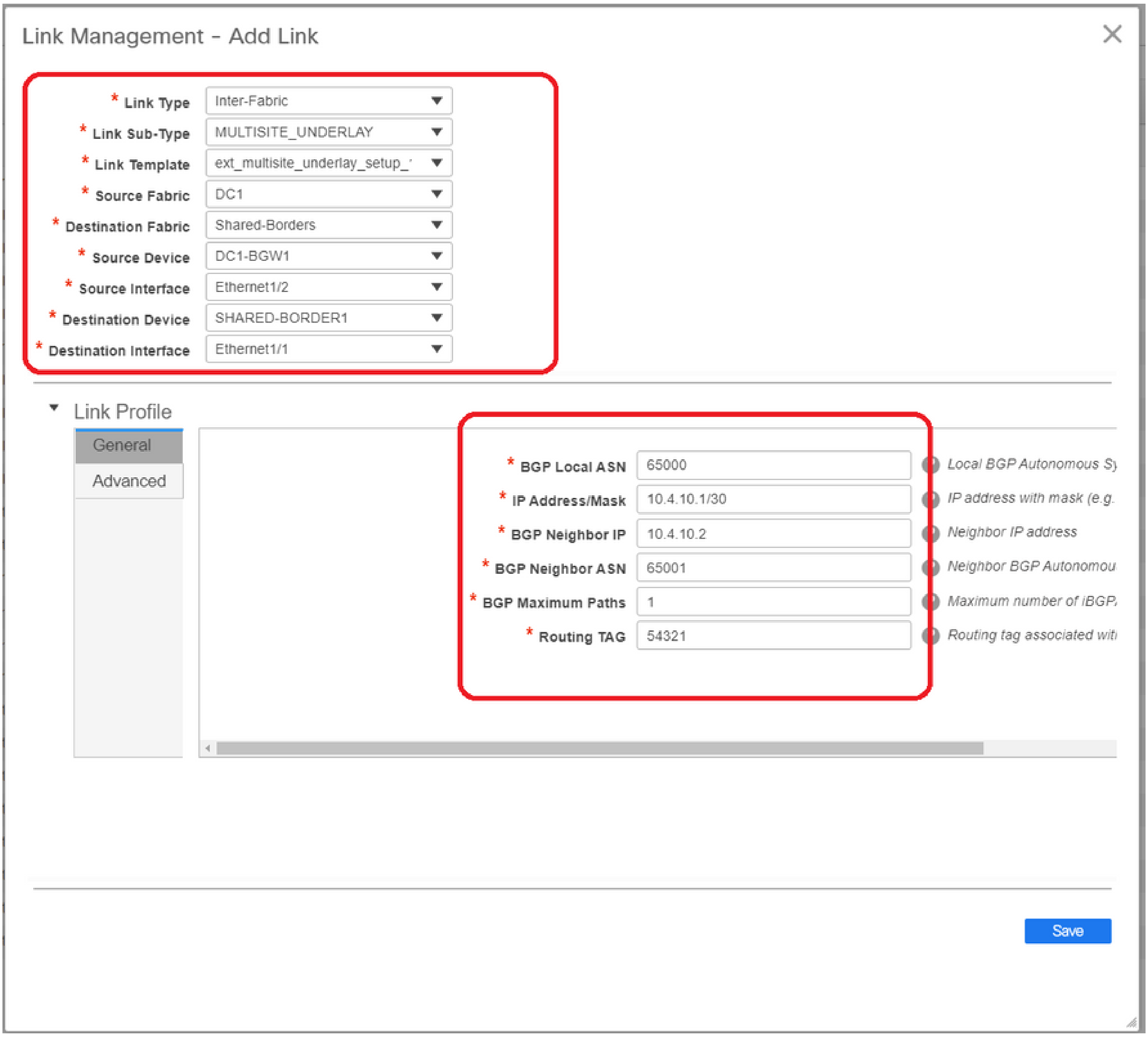
# Shared borders run eBGP l2vpn evpn with the Border Gateways and VRF-LITE connections towards the external router

# Before Forming eBGP l2vpn evpn with the loopbacks, it is required to make sure that the loopbacks are reachable via some method; In this example, we are using eBGP IPv4 AF from BGWs to Shared borders and then advertise the loopbacks to further form the l2vpn evpn neighborship.



# Once the MSD fabric is selected, switch to "tabular view"





Link Management - Add Link

* Link Type	Inter-Fabric
* Link Sub-Type	MULTISITE_UNDERLAY
* Link Template	ext_multisite_underlay_setup_'
* Source Fabric	DC1
* Destination Fabric	Shared-Borders
* Source Device	DC1-BGW1
* Source Interface	Ethernet1/2
* Destination Device	SHARED-BORDER1
* Destination Interface	Ethernet1/1

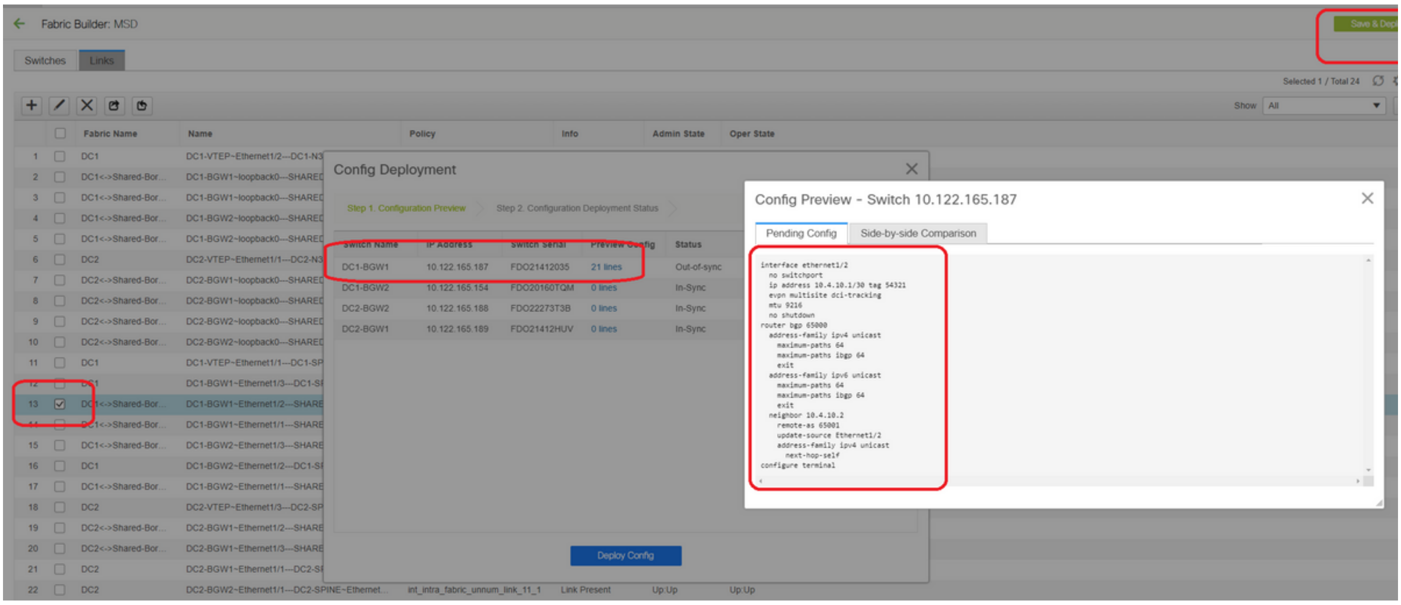
Link Profile

* BGP Local ASN	65000	Local BGP Autonomous Sy
* IP Address/Mask	10.4.10.1/30	IP address with mask (e.g.
* BGP Neighbor IP	10.4.10.2	Neighbor IP address
* BGP Neighbor ASN	65001	Neighbor BGP Autonomou
* BGP Maximum Paths	1	Maximum number of iBGP/
* Routing TAG	54321	Routing tag associated wit

# Select the "inter-fabric" and use the "Multisite_UNDERLAY"

# We are here trying to form an IPv4 BGP Neighborship with the Shared border router; So select the switches and interfaces accordingly.

# Note that If CDP is detecting the neighbor from DC1-BGW1 to SB1, it is only required to provide the IP addresses here in this section and that will effectively configure the IP addresses on the relevant interfaces after performing "save & Deploy"



Fabric Builder: MSD

Config Deployment

Step 1. Configuration Preview — Step 2. Configuration Deployment Status

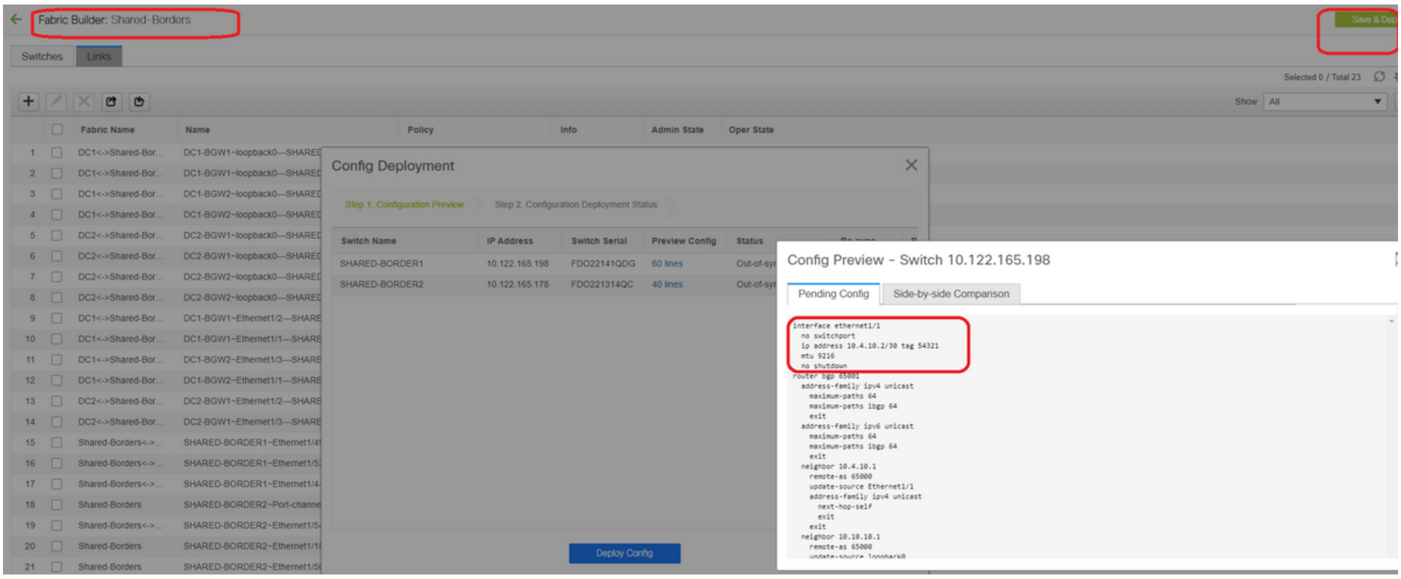
Switch Name	IP Address	Switch Serial	Preview Config	Status
DC1-BGW1	10.122.165.187	FDO21412035	21 lines	Out-of-sync
DC1-BGW2	10.122.165.154	FDO20160TQM	0 lines	In-Sync
DC2-BGW2	10.122.165.188	FDO22273T3B	0 lines	In-Sync
DC2-BGW1	10.122.165.189	FDO21412HUV	0 lines	In-Sync

Config Preview - Switch 10.122.165.187

Pending Config | Side-by-side Comparison

```
interface ethernet1/2
  no switchport
  ip address 10.4.10.1/30 tag 54321
  evpn multisite dci-tracking
  mtu 9216
  no shutdown
router bgp 65000
  address-family ipv4 unicast
    maximum-paths 64
    maximum-paths ibgp 64
    exit
  address-family ipv6 unicast
    maximum-paths 64
    maximum-paths ibgp 64
    exit
  neighbor 10.4.10.2
    remote-as 65001
    update-source Ethernet1/2
    address-family ipv4 unicast
      next-hop-self
configure terminal
```

# Once Save and deploy is selected, the required configuration lines are propagated for DC1-BGW1; Same step will have to be performed after selecting the "Shared border" fabric too.



Fabric Builder: Shared-Borders

Config Deployment

Step 1. Configuration Preview — Step 2. Configuration Deployment Status

Switch Name	IP Address	Switch Serial	Preview Config	Status
SHARED-BORDER1	10.122.165.198	FDO22141QDG	60 lines	Out-of-syn
SHARED-BORDER2	10.122.165.178	FDO221314QC	40 lines	Out-of-syn

Config Preview - Switch 10.122.165.198

Pending Config | Side-by-side Comparison

```
interface ethernet1/1
  no switchport
  ip address 10.4.10.2/30 tag 54321
  mtu 9216
  no shutdown
router bgp 65001
  address-family ipv4 unicast
    maximum-paths 64
    maximum-paths ibgp 64
    exit
  address-family ipv6 unicast
    maximum-paths 64
    maximum-paths ibgp 64
    exit
  neighbor 10.4.10.1
    remote-as 65000
    update-source Ethernet1/1
    address-family ipv4 unicast
      next-hop-self
      exit
    exit
  neighbor 10.10.10.1
    remote-as 65000
    update-source loopback0
```

# From CLI, the same can be verified using the command below;

```
DC1-BGW1# show ip bgp sum
BGP summary information for VRF default, address family IPv4 Unicast
BGP router identifier 10.10.10.1, local AS number 65000
BGP table version is 11, IPv4 Unicast config peers 1, capable peers 1
2 network entries and 2 paths using 480 bytes of memory
BGP attribute entries [1/164], BGP AS path entries [0/0]
BGP community entries [0/0], BGP clusterlist entries [0/0]

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.4.10.2       4 65001       6       7       11    0    0 00:00:52 0
```

# Note that the "save&Deploy" has to be done on the DC1 fabric as well(Select the drop down for DC1 and then perform the same) so that the Relevant IP addressing, BGP configurations are propagated to the switches in DC1(which are the Border Gateways);

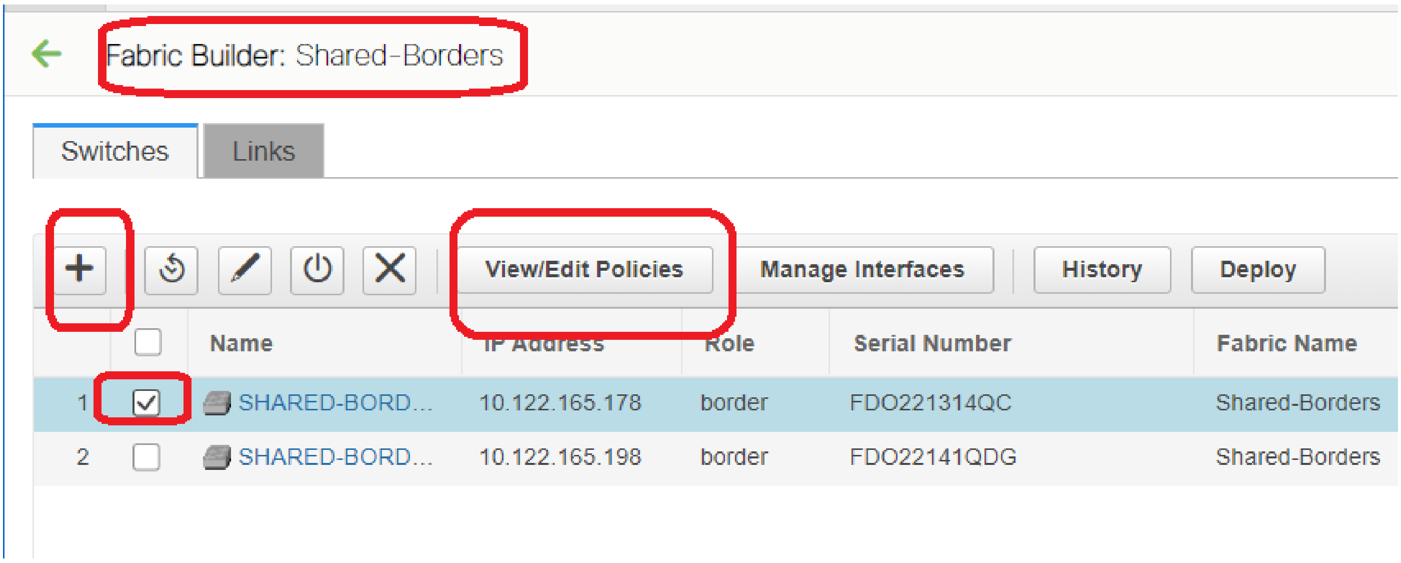
# Also, the multisite underlay has to be created from DC1-BGWs, DC2-BGWs to Shared borders; so, same steps as above have to be done for the same too.

# At the end, Shared borders will have eBGP IPv4 AF neighborship with all BGWs in DC1 and DC2 as below;

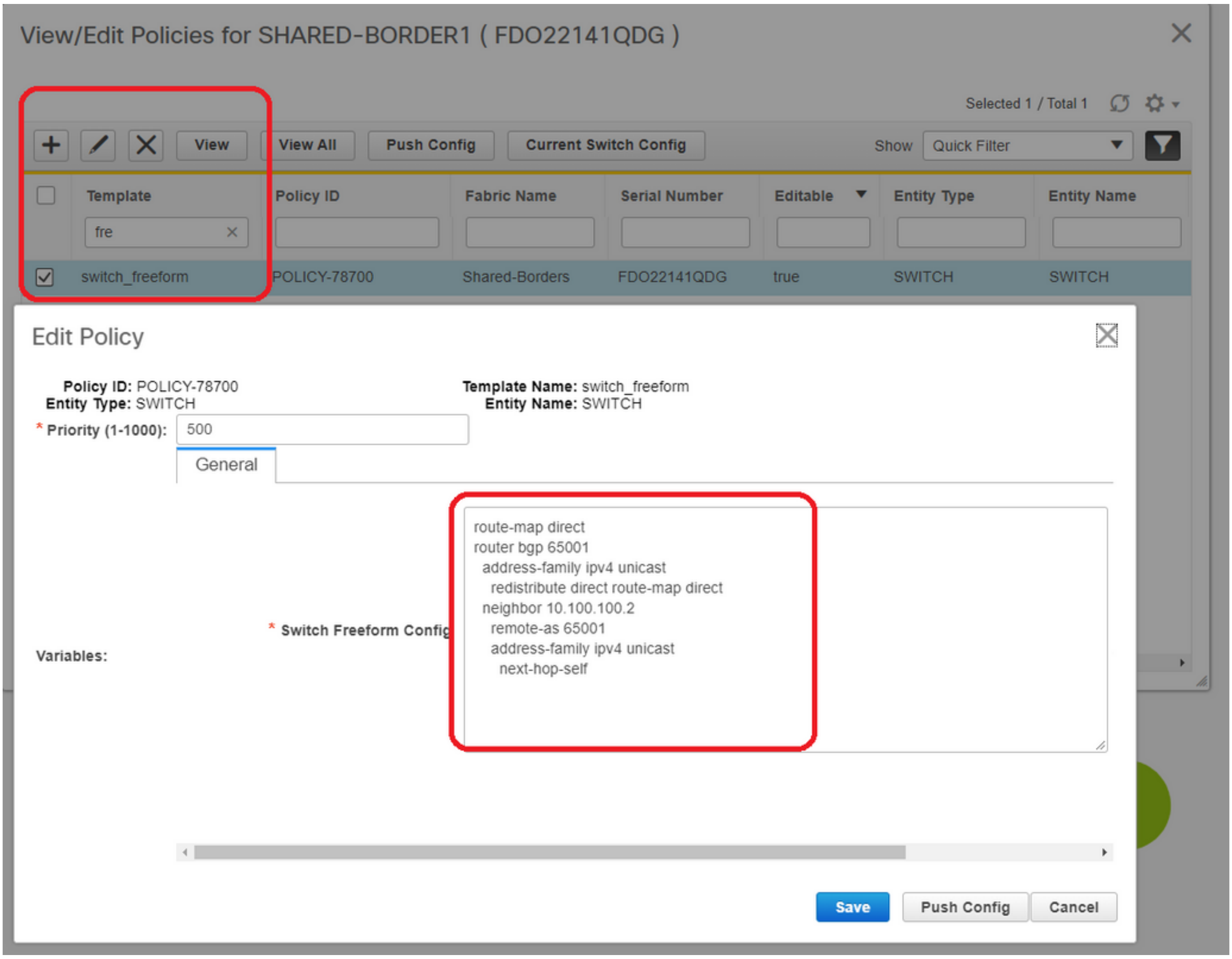
```
SHARED-BORDER1# sh ip bgp sum
BGP summary information for VRF default, address family IPv4 Unicast
BGP router identifier 10.10.100.1, local AS number 65001
BGP table version is 38, IPv4 Unicast config peers 4, capable peers 4
18 network entries and 20 paths using 4560 bytes of memory
BGP attribute entries [2/328], BGP AS path entries [2/12]
BGP community entries [0/0], BGP clusterlist entries [0/0]

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.4.10.1       4 65000    1715    1708       38    0    0    1d03h 5
10.4.10.6       4 65000    1461    1458       38    0    0    1d00h 5
10.4.10.18      4 65002    1459    1457       38    0    0    1d00h 5
10.4.10.22      4 65002    1459    1457       38    0    0    1d00h 5

SHARED-BORDER2# sh ip bgp sum
BGP summary information for VRF default, address family IPv4 Unicast
BGP router identifier 10.10.100.2, local AS number 65001
BGP table version is 26, IPv4 Unicast config peers 4, capable peers 4
18 network entries and 20 paths using 4560 bytes of memory
BGP attribute entries [2/328], BGP AS path entries [2/12]
BGP community entries [0/0], BGP clusterlist entries [0/0]

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.4.10.10      4 65000    1459    1458       26    0    0    1d00h 5
10.4.10.14      4 65000    1461    1458       26    0    0    1d00h 5
10.4.10.26      4 65002    1459    1457       26    0    0    1d00h 5
10.4.10.30      4 65002    1459    1457       26    0    0    1d00h 5
```

# Above is the pre-requisite prior building the l2vpn evpn neighborship from BGWs to Shared borders(Note that its not mandatory to use BGP; any other mechanism to exchange loopback prefixes would do); At the end, the base requirement is that all loopbacks(of Shared borders, BGWs) should be reachable from all BGWs

# Please also note that an iBGP IPv4 AF neighborship needs to be established between Shared borders; As of Today, DCNM do not have an option to build an iBGP between shared borders using a template/drop down; For that, a Freeform configuration has to be done which is shown below;



	☐	Name	IP Address	Role	Serial Number	Fabric Name
1	☑	SHARED-BORD...	10.122.165.178	border	FDO221314QC	Shared-Borders
2	☐	SHARED-BORD...	10.122.165.198	border	FDO22141QDG	Shared-Borders



View/Edit Policies for SHARED-BORDER1 (FDO22141QDG)

	Template	Policy ID	Fabric Name	Serial Number	Editable	Entity Type	Entity Name
□	fre						
☑	switch_freeform	POLICY-78700	Shared-Borders	FDO22141QDG	true	SWITCH	SWITCH

Edit Policy

Policy ID: POLICY-78700
Entity Type: SWITCH
Template Name: switch_freeform
Entity Name: SWITCH
* **Priority (1-1000):** 500

Variables:

* **Switch Freeform Config**

```
route-map direct
router bgp 65001
  address-family ipv4 unicast
    redistribute direct route-map direct
  neighbor 10.100.100.2
    remote-as 65001
    address-family ipv4 unicast
      next-hop-self
```

# Find the IP addresses which are configured on the Backup SVI of shared borders; As shown above, freeform is added on Shared-border1 switch and the iBGP neighbor specified is that of the Shared-border2(10.100.100.2)

# Note that while providing the configurations within the freeform in DCNM, provide the correct spacing after each commands(leave even number of spaces; meaning, after router bgp 65001, provide two spaces and then give the neighbor <> command and so on)

# Also make sure to perform a redistribute direct for the direct routes(loopback routes) in BGP or some other form to advertise the loopbacks; in the example above, a route-map direct is created to match all direct routes and then redistribute direct is done within the IPv4 AF BGP

# Once the configuration is "saved and deployed" from DCNM, iBGP neighborship is formed as shown below;

```
SHARED-BORDER1# sh ip bgp sum
BGP summary information for VRF default, address family IPv4 Unicast
BGP router identifier 10.10.100.1, local AS number 65001
BGP table version is 57, IPv4 Unicast config peers 5, capable peers 5
18 network entries and 38 paths using 6720 bytes of memory
BGP attribute entries [4/656], BGP AS path entries [2/12]
BGP community entries [0/0], BGP clusterlist entries [0/0]
```

```
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.4.10.1       4 65000    1745    1739       57    0    0    1d04h 5
10.4.10.6       4 65000    1491    1489       57    0    0    1d00h 5
10.4.10.18      4 65002    1490    1487       57    0    0    1d00h 5
10.4.10.22      4 65002    1490    1487       57    0    0    1d00h 5
10.100.100.2    4 65001      14       6       57    0    0 00:00:16 18 # iBGP neighborship from
shared border1 to shared border2
```

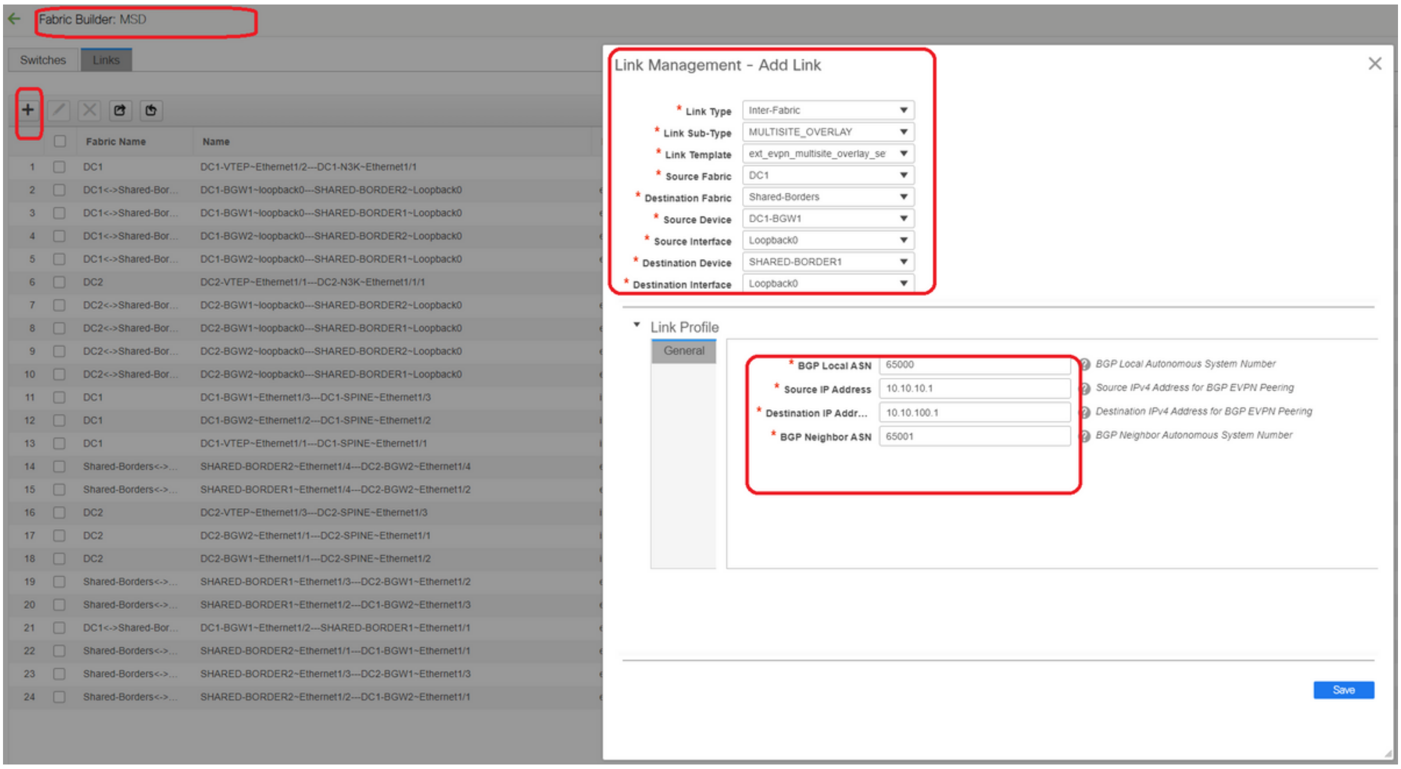
# With above step, the multisite underlay is fully configured.

# Next Step is to build the multisite overlay;

Step 9: Building Multisite Overlay from BGWs to Shared borders

# Note that, here Shared borders are also the route servers

# Select the MSD and then go to the "Tabular view" where a new link can be created; From there, a new multisite overlay link has to be created and the relevant IP addresses will have to provided with the correct ASN as below; This step has to be done for all the l2vpn evpn neighbors(which is from every BGW to Every Shared border)



	☐	Fabric Name	Name
1	☐	DC1	DC1-VTEP~Ethernet1/2---DC1-N3K~Ethernet1/1
2	☐	DC1<->Shared-Bor...	DC1-BGW1~loopback0---SHARED-BORDER2~Loopback0
3	☐	DC1<->Shared-Bor...	DC1-BGW1~loopback0---SHARED-BORDER1~Loopback0
4	☐	DC1<->Shared-Bor...	DC1-BGW2~loopback0---SHARED-BORDER2~Loopback0
5	☐	DC1<->Shared-Bor...	DC1-BGW2~loopback0---SHARED-BORDER1~Loopback0
6	☐	DC2	DC2-VTEP~Ethernet1/1---DC2-N3K~Ethernet1/1/1
7	☐	DC2<->Shared-Bor...	DC2-BGW1~loopback0---SHARED-BORDER2~Loopback0
8	☐	DC2<->Shared-Bor...	DC2-BGW1~loopback0---SHARED-BORDER1~Loopback0
9	☐	DC2<->Shared-Bor...	DC2-BGW2~loopback0---SHARED-BORDER2~Loopback0
10	☐	DC2<->Shared-Bor...	DC2-BGW2~loopback0---SHARED-BORDER1~Loopback0
11	☐	DC1	DC1-BGW1~Ethernet1/3---DC1-SPINE~Ethernet1/3
12	☐	DC1	DC1-BGW2~Ethernet1/2---DC1-SPINE~Ethernet1/2
13	☐	DC1	DC1-VTEP~Ethernet1/1---DC1-SPINE~Ethernet1/1
14	☐	Shared-Borders<->...	SHARED-BORDER2~Ethernet1/4---DC2-BGW2~Ethernet1/4
15	☐	Shared-Borders<->...	SHARED-BORDER1~Ethernet1/4---DC2-BGW2~Ethernet1/2
16	☐	DC2	DC2-VTEP~Ethernet1/3---DC2-SPINE~Ethernet1/3
17	☐	DC2	DC2-BGW2~Ethernet1/1---DC2-SPINE~Ethernet1/1
18	☐	DC2	DC2-BGW1~Ethernet1/1---DC2-SPINE~Ethernet1/2
19	☐	Shared-Borders<->...	SHARED-BORDER1~Ethernet1/3---DC2-BGW1~Ethernet1/2
20	☐	Shared-Borders<->...	SHARED-BORDER1~Ethernet1/2---DC1-BGW2~Ethernet1/3
21	☐	DC1<->Shared-Bor...	DC1-BGW1~Ethernet1/2---SHARED-BORDER1~Ethernet1/1
22	☐	Shared-Borders<->...	SHARED-BORDER2~Ethernet1/1---DC1-BGW1~Ethernet1/1
23	☐	Shared-Borders<->...	SHARED-BORDER2~Ethernet1/3---DC2-BGW1~Ethernet1/3
24	☐	Shared-Borders<->...	SHARED-BORDER2~Ethernet1/2---DC1-BGW2~Ethernet1/1

Link Management - Add Link

* Link Type	Inter-Fabric
* Link Sub-Type	MULTISITE_OVERLAY
* Link Template	ext_evpn_multisite_overlay_se
* Source Fabric	DC1
* Destination Fabric	Shared-Borders
* Source Device	DC1-BGW1
* Source Interface	Loopback0
* Destination Device	SHARED-BORDER1
* Destination Interface	Loopback0

Link Profile

General

* BGP Local ASN	65000	BGP Local Autonomous System Number
* Source IP Address	10.10.10.1	Source IPv4 Address for BGP EVPN Peering
* Destination IP Addr...	10.10.100.1	Destination IPv4 Address for BGP EVPN Peering
* BGP Neighbor ASN	65001	BGP Neighbor Autonomous System Number

# Above is one example; Perform the same for all other multisite Overlay Links and at the end, the CLI will look like below;

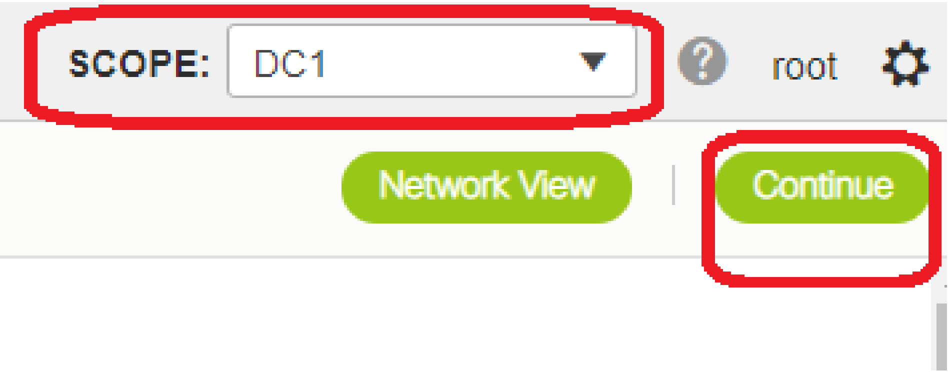
```
SHARED-BORDER1# sh bgp l2vpn evpn summary
BGP summary information for VRF default, address family L2VPN EVPN
BGP router identifier 10.10.100.1, local AS number 65001
BGP table version is 8, L2VPN EVPN config peers 4, capable peers 4
1 network entries and 1 paths using 240 bytes of memory
BGP attribute entries [1/164], BGP AS path entries [0/0]
BGP community entries [0/0], BGP clusterlist entries [0/0]
```

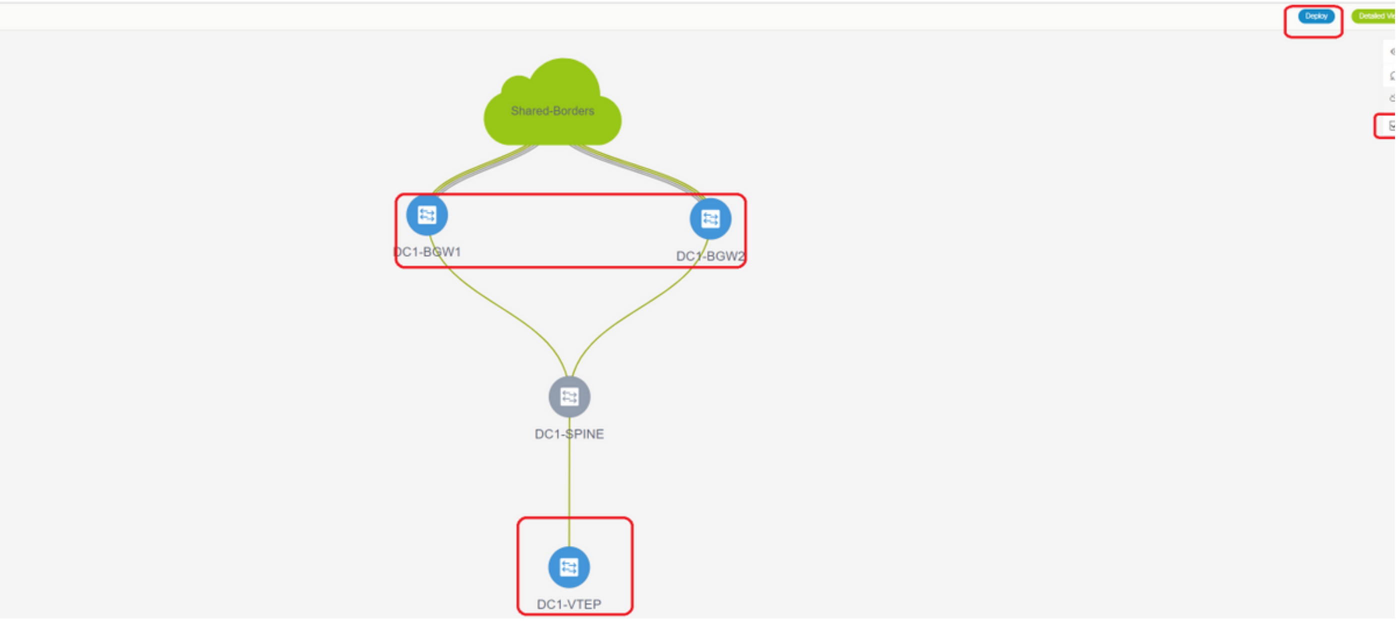
```
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.10.10.1      4 65000      21      19        8    0    0 00:13:52 0
10.10.10.2      4 65000      22      20        8    0    0 00:14:14 0
10.10.20.1      4 65002      21      19        8    0    0 00:13:56 0
10.10.20.2      4 65002      21      19        8    0    0 00:13:39 0

SHARED-BORDER2# sh bgp l2vpn evpn summary
BGP summary information for VRF default, address family L2VPN EVPN
BGP router identifier 10.10.100.2, local AS number 65001
BGP table version is 8, L2VPN EVPN config peers 4, capable peers 4
1 network entries and 1 paths using 240 bytes of memory
BGP attribute entries [1/164], BGP AS path entries [0/0]
BGP community entries [0/0], BGP clusterlist entries [0/0]

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.10.10.1      4 65000      22      20        8    0    0 00:14:11 0
10.10.10.2      4 65000      21      19        8    0    0 00:13:42 0
10.10.20.1      4 65002      21      19        8    0    0 00:13:45 0
10.10.20.2      4 65002      22      20        8    0    0 00:14:15 0
```

Step 10: Deploying Networks/VRFs on both sites

# As we have finished the multisite Underlay and Overlay, Next step is to deploy the Networks/VRFs on all devices;

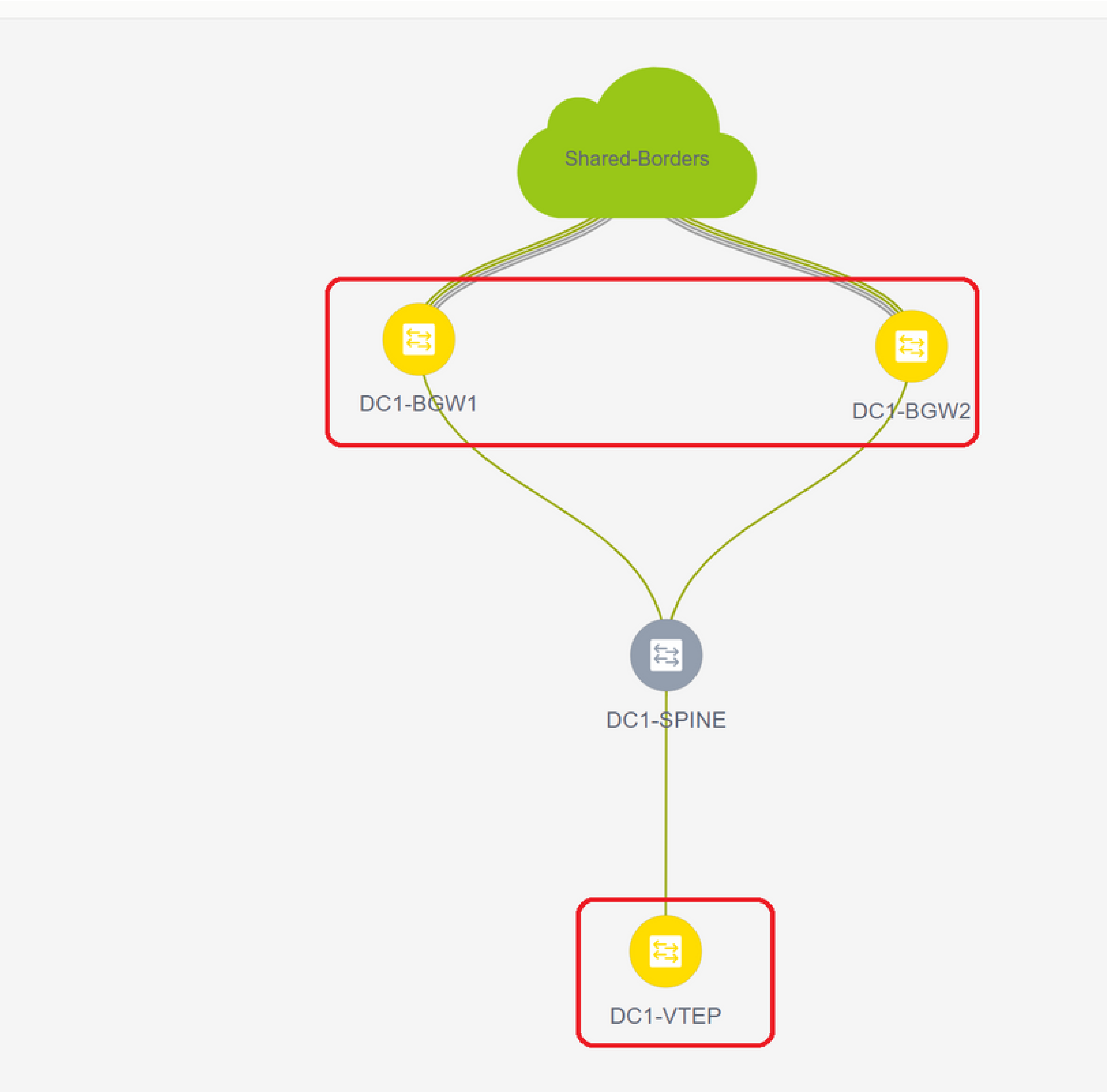
# Starting with VRFs on Fabrics-> DC1, DC2 and Shared borders.


SCOPE: DC1



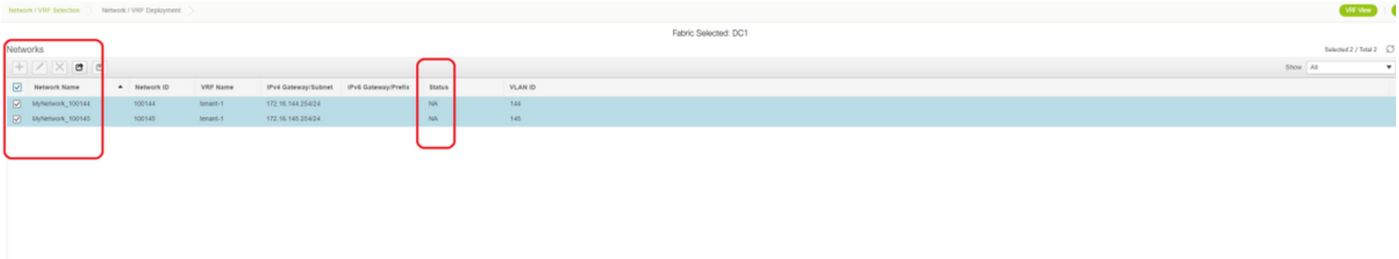
# Once the VRF View is selected, click on "continue"; This will list out the devices in the topology

# Since the VRF has to be deployed to multiple switches(including Border Gateways and Leaf), select the Checkbox at the far right and then select the switches which has the same role at one time; eg; DC1-BGW1 and DC1-BGW2 can be selected at one time and then save both the switches; After this, select the leaf switches which are applicable(here it would be DC1-VTEP)



# As seen above, When the "Deploy" option is selected, all the switches which were previously selected will start the deployment and will finally turn green if the deploy was successful.

# Same steps will have to be performed for deploying Networks;



Fabric Selected: DC1

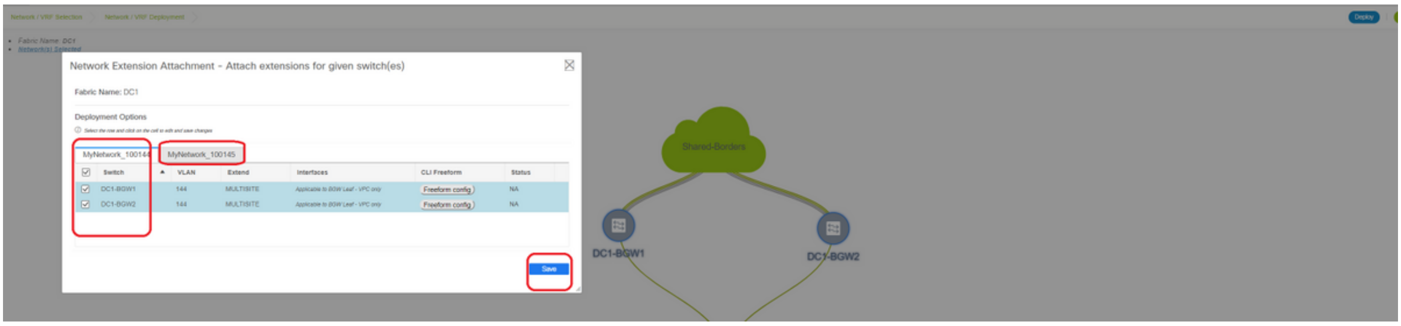
Networks

Selected 2 / Total 2

Show: All

☑	Network Name	Network ID	VRF Name	IPv4 Gateway/Subnet	IPv6 Gateway/Prefix	Status	VLAN ID
☑	MyNetwork_100144	100144	tenant-1	172.16.144.254/24		NA	144
☑	MyNetwork_100145	100145	tenant-1	172.16.145.254/24		NA	145

# If multiple networks are created, keep in mind to navigate to the subsequent tabs to select the networks before deploying



Network / VRF Selection > Network / VRF Deployment

- Fabric Name: DC1
- Network(s) Selected

Network Extension Attachment - Attach extensions for given switch(es)

Fabric Name: DC1

Deployment Options

Select the row and click on the cell to edit and save changes

MyNetwork_100144 | MyNetwork_100145

☑	Switch	VLAN	Extend	Interfaces	CLI Freeform	Status
☑	DC1-BGW1	144	MULTISITE	Applicable to BGW Leaf - VPC only	Freeform config	NA
☑	DC1-BGW2	144	MULTISITE	Applicable to BGW Leaf - VPC only	Freeform config	NA

Save

# The status will now turn to "DEPLOYED" from "NA" and the switch' CLI below can be used in order to verify the deployments

```
DC1-VTEP# sh nve vni
Codes: CP - Control Plane         DP - Data Plane
       UC - Unconfigured          SA - Suppress ARP
       SU - Suppress Unknown Unicast
       Xconn - Crossconnect
       MS-IR - Multisite Ingress Replication

Interface VNI      Multicast-group   State Mode Type [BD/VRF]      Flags
--------- -------- ----------------- ----- ---- ------------------ -----
nve1      100144   239.1.1.144       Up    CP   L2 [144]                # Network1 which is VLan
144 mapped to VNID 100144
nve1      100145   239.1.1.145       Up    CP   L2 [145]                # Network2 Which is Vlan
145 mapped to VNID 100145
nve1      1001445  239.100.100.100   Up    CP   L3 [tenant-1]           # VRF- tenant1 which is
mapped to VNID 1001445


DC1-BGW1# sh nve vni
Codes: CP - Control Plane         DP - Data Plane
       UC - Unconfigured          SA - Suppress ARP
       SU - Suppress Unknown Unicast
       Xconn - Crossconnect
       MS-IR - Multisite Ingress Replication

Interface VNI      Multicast-group   State Mode Type [BD/VRF]      Flags
--------- -------- ----------------- ----- ---- ------------------ -----
nve1      100144   239.1.1.144       Up    CP   L2 [144]           MS-IR
nve1      100145   239.1.1.145       Up    CP   L2 [145]           MS-IR
nve1      1001445  239.100.100.100   Up    CP   L3 [tenant-1]
```

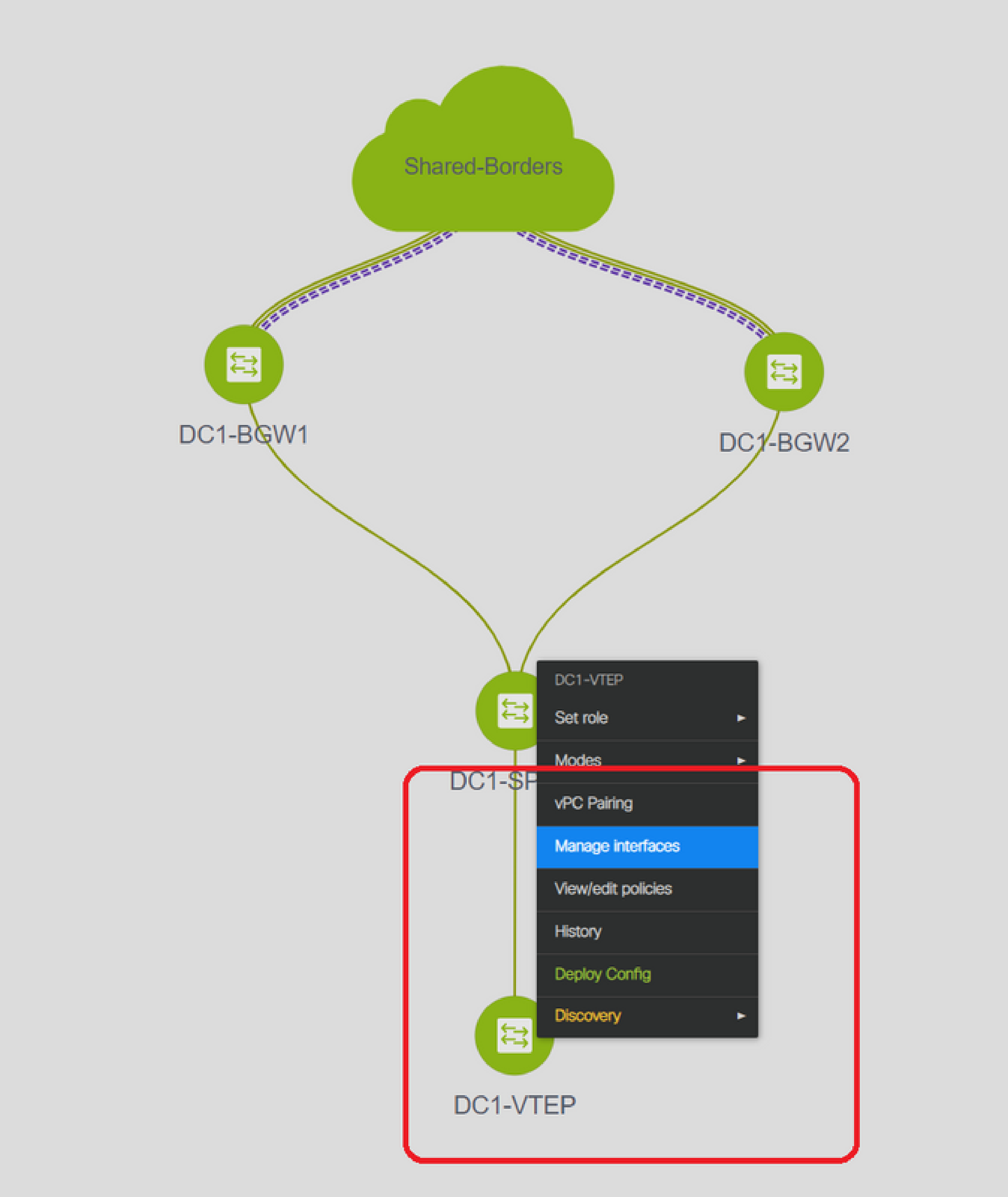
# Above is from BGW as well; so in short, all switches which we had selected earlier in the step will be deployed with the Networks and VRF

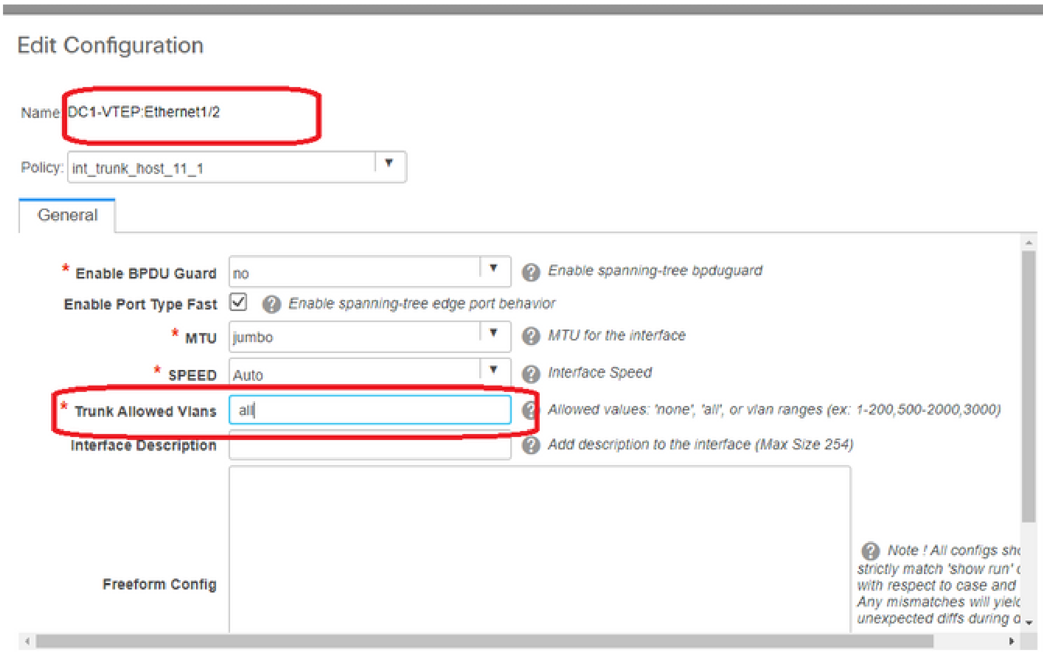
# Same steps have to be carried out for the Fabric DC2, Shared border as well. Keep in mind that the Shared borders DO NOT need any networks or layer 2 VNIDs; only L3 VRF is required.

Step 11: Creating downstream Trunk/Access ports on Leaf Switches/VTEP

# In this Topology, Ports Eth1/2 and Eth1/1 from DC1-VTEP and DC2-VTEP respectively are

connected to the hosts; so Moving those as trunk ports in DCNM GUI as shown below





Edit Configuration

Name DC1-VTEP:Ethernet1/2

Policy: int_trunk_host_11_1

General

* Enable BPDU Guard: no — Enable spanning-tree bpduguard

Enable Port Type Fast: ☑ — Enable spanning-tree edge port behavior

* MTU: jumbo — MTU for the interface

* SPEED: Auto — Interface Speed

* Trunk Allowed Vlans: all — Allowed values: 'none', 'all', or vlan ranges (ex: 1-200,500-2000,3000)

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

Freeform Config — Note ! All configs sh strictly match 'show run' with respect to case and Any mismatches will yiel unexpected diffs during d

# Select the relevant interface and change the "allowed vlans" from none to "all"(or only the vlans that are needing to be allowed)

Step 12: Freeforms required on Shared border

# Since Shared border switches are the route-servers, it is required to make some changes in terms of the BGP l2vpn evpn neighborships

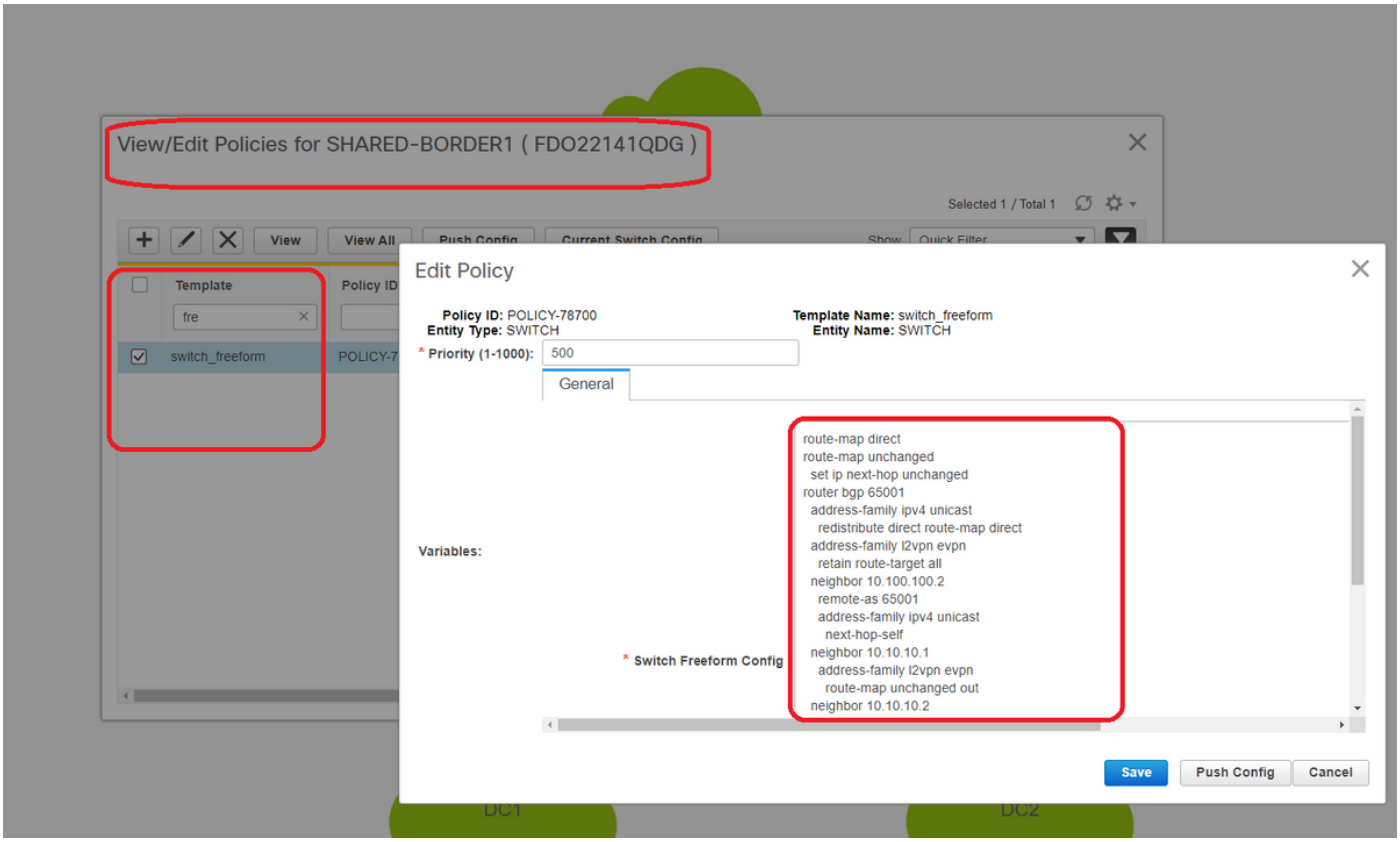
# inter-site BUM traffic is replicated using Unicast; Means, any BUM traffic in Vlan 144(eg) after it arrives on the BGWs; depending upon which BGW is the Designated forwarder(DF), DF will perform a unicast replication to remote site; This replication is achieved after the BGW receives a type 3 route from the remote BGW; Here, the BGWs are forming l2vpn evpn peering only with shared borders; and Shared borders should not have any layer 2 VNIDs(if created, this will result in blackholing of East/West Traffic). Since layer 2 VNIDs are missing and the route-type 3 is originated by BGWs per VNID, the Shared borders will not honor the BGP update coming in from BGWs; To fix this, Use the "retain route-target all" under the AF l2vpn evpn

# Another point is to make sure that the Shared borders do not change the Next HOP(BGP BY default changes the next hop for eBGP neighborships); Here, the inter site tunnel for unicast traffic from site 1 to 2 and vice versa should be from BGW to BGW(from dc1 to dc2 and vice versa); To achieve this, a route-map has to be created and applied for every l2vpn evpn neighborships from shared border to each BGWs

# For both the above points, a freeform has to be used on shared borders like below

```
route-map direct
```

```
route-map unchanged
  set ip next-hop unchanged
router bgp 65001
  address-family ipv4 unicast
    redistribute direct route-map direct
  address-family l2vpn evpn
    retain route-target all
  neighbor 10.100.100.2
    remote-as 65001
    address-family ipv4 unicast
      next-hop-self
  neighbor 10.10.10.1
    address-family l2vpn evpn
      route-map unchanged out
  neighbor 10.10.10.2
    address-family l2vpn evpn
      route-map unchanged out
  neighbor 10.10.20.1
    address-family l2vpn evpn
      route-map unchanged out
  neighbor 10.10.20.2
    address-family l2vpn evpn
      route-map unchanged out
```



View/Edit Policies for SHARED-BORDER1 (FDO22141QDG)

Selected 1 / Total 1

☐	Template	Policy ID
	fre	
☑	switch_freeform	POLICY-7

Edit Policy

Policy ID: POLICY-78700
Entity Type: SWITCH
Template Name: switch_freeform
Entity Name: SWITCH
* **Priority (1-1000):** 500

General

Variables:

* **Switch Freeform Config**

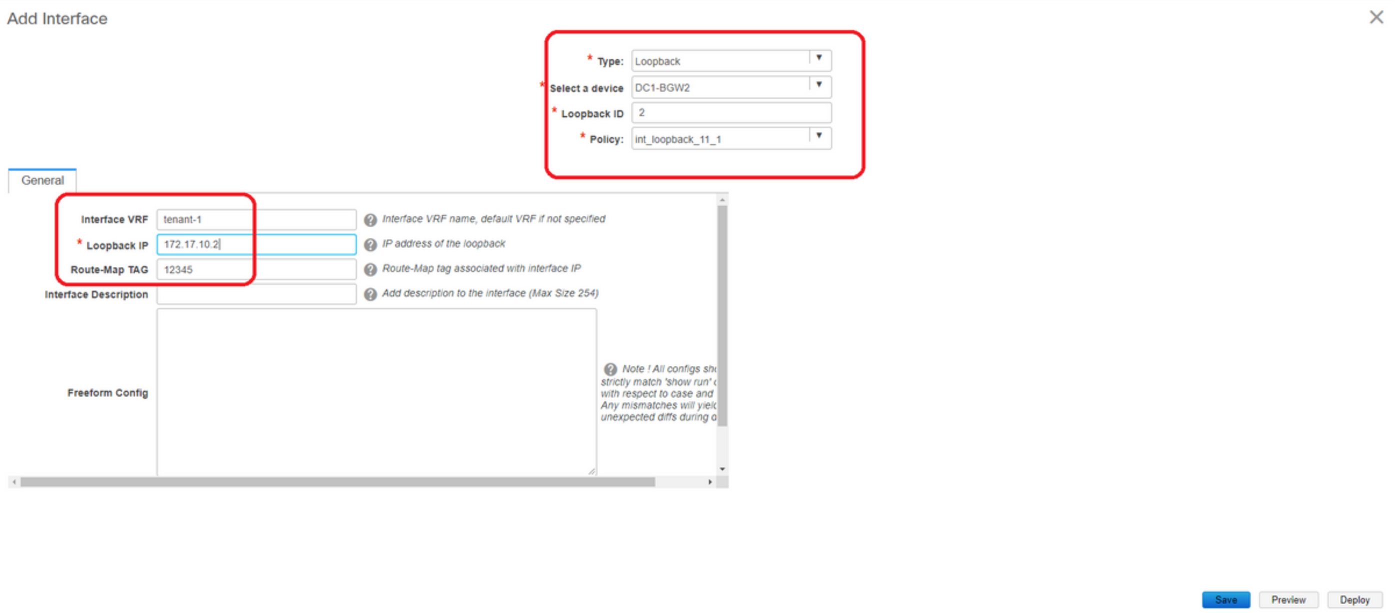
```
route-map direct
route-map unchanged
 set ip next-hop unchanged
router bgp 65001
 address-family ipv4 unicast
  redistribute direct route-map direct
 address-family l2vpn evpn
  retain route-target all
 neighbor 10.100.100.2
  remote-as 65001
  address-family ipv4 unicast
   next-hop-self
 neighbor 10.10.10.1
  address-family l2vpn evpn
   route-map unchanged out
 neighbor 10.10.10.2
```

Step 13: Loopback within tenant VRFs on BGWs

# for North/South traffic from hosts connected within the leaf switches, the BGWs use the Outer SRC IP of the NVE Loopback1 IP address; Shared borders will only by default form the NVE Peering with the Multisite Loopback Ip address of BGWs; so if a vxlan packet comes to the shared border with an outer SRC IP address of the BGW Loopback1, the packet will be dropped due to the SRCTEP Miss; To avoid this, a loopback in tenant-VRF has to be created on every BGW switch and then advertise to the BGP so that the Shared borders receive this update and then form the NVE Peering with the BGW Loopback1 IP address ;

# Initially the NVE Peering will look like below on shared borders

```
SHARED-BORDER1# sh nve pee
Interface Peer-IP                                 State LearnType Uptime   Router-Mac
--------- --------------------------------------  ----- --------- -------- -----------------
nve1      10.222.222.1                            Up    CP        01:20:09 0200.0ade.de01    #
Multisite Loopback 100 IP address of DC1-BGWs
nve1      10.222.222.2                            Up    CP        01:17:43 0200.0ade.de02    #
Multisite Loopback 100 IP address of DC2-BGWs
```



Add Interface

* Type: Loopback

* Select a device DC1-BGW2

* Loopback ID 2

* Policy: int_loopback_11_1

General

Interface VRF tenant-1 — Interface VRF name, default VRF if not specified

* Loopback IP 172.17.10.2 — IP address of the loopback

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

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

Freeform Config — Note ! All configs sh strictly match 'show run' with respect to case and Any mismatches will yiel unexpected diffs during o

# As shown above, the loopback2 is created from DCNM and is configured in tenant-1 VRF and is given the Tag of 12345 as this is the tag which the route-map uses to match the loopback while making the advertisement

```
DC1-BGW1# sh run vrf tenant-1

!Command: show running-config vrf tenant-1
!Running configuration last done at: Tue Dec 10 17:21:29 2019
!Time: Tue Dec 10 17:24:53 2019

version 9.3(2) Bios:version 07.66

interface Vlan1445
  vrf member tenant-1

interface loopback2
  vrf member tenant-1
vrf context tenant-1
  vni 1001445
  ip pim rp-address 10.49.3.100 group-list 224.0.0.0/4
  ip pim ssm range 232.0.0.0/8
  rd auto
  address-family ipv4 unicast
    route-target both auto
    route-target both auto mvpn
    route-target both auto evpn
  address-family ipv6 unicast
    route-target both auto
    route-target both auto evpn
```

```
router bgp 65000
  vrf tenant-1
    address-family ipv4 unicast
      advertise l2vpn evpn
      redistribute direct route-map fabric-rmap-redist-subnet
      maximum-paths ibgp 2
    address-family ipv6 unicast
      advertise l2vpn evpn
      redistribute direct route-map fabric-rmap-redist-subnet
      maximum-paths ibgp 2
```

```
DC1-BGW1# sh route-map fabric-rmap-redist-subnet
route-map fabric-rmap-redist-subnet, permit, sequence 10
  Match clauses:
    tag: 12345
  Set clauses:
```

# After this step, the NVE peerings will show for all the Loopback1 Ip addresses along with the multisite loopback IP address.

```
SHARED-BORDER1# sh nve pee
Interface Peer-IP                                 State LearnType Uptime   Router-Mac
--------- --------------------------------------  ----- --------- -------- -----------------
nve1      192.168.20.1                            Up    CP        00:00:01 b08b.cfdc.2fd7
nve1      10.222.222.1                            Up    CP        01:27:44 0200.0ade.de01
nve1      192.168.10.2                            Up    CP        00:01:00 e00e.daa2.f7d9
nve1      10.222.222.2                            Up    CP        01:25:19 0200.0ade.de02
nve1      192.168.10.3                            Up    CP        00:01:43 6cb2.aeee.0187
nve1      192.168.20.3                            Up    CP        00:00:28 005d.7307.8767
```

# At this stage, the East/West traffic should be forwarded correctly

Step 14: VRFLITE Extensions from Shared borders to the External Routers

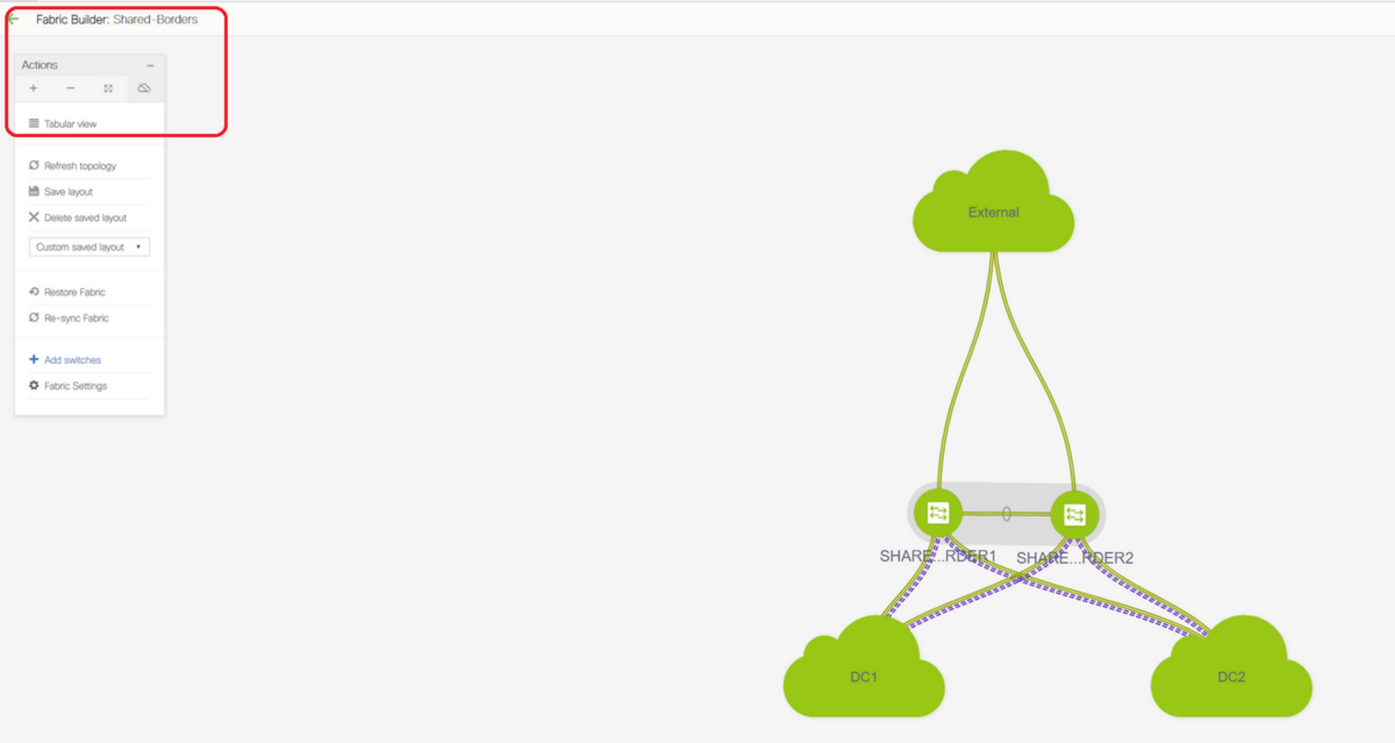
# There will be situations when hosts outside the fabric will have to talk to the Hosts within the fabric. In this example, the same is made possible by the shared borders;

# Any host that is living in DC1 or DC2 will be able to talk to external hosts via the shared border switches.

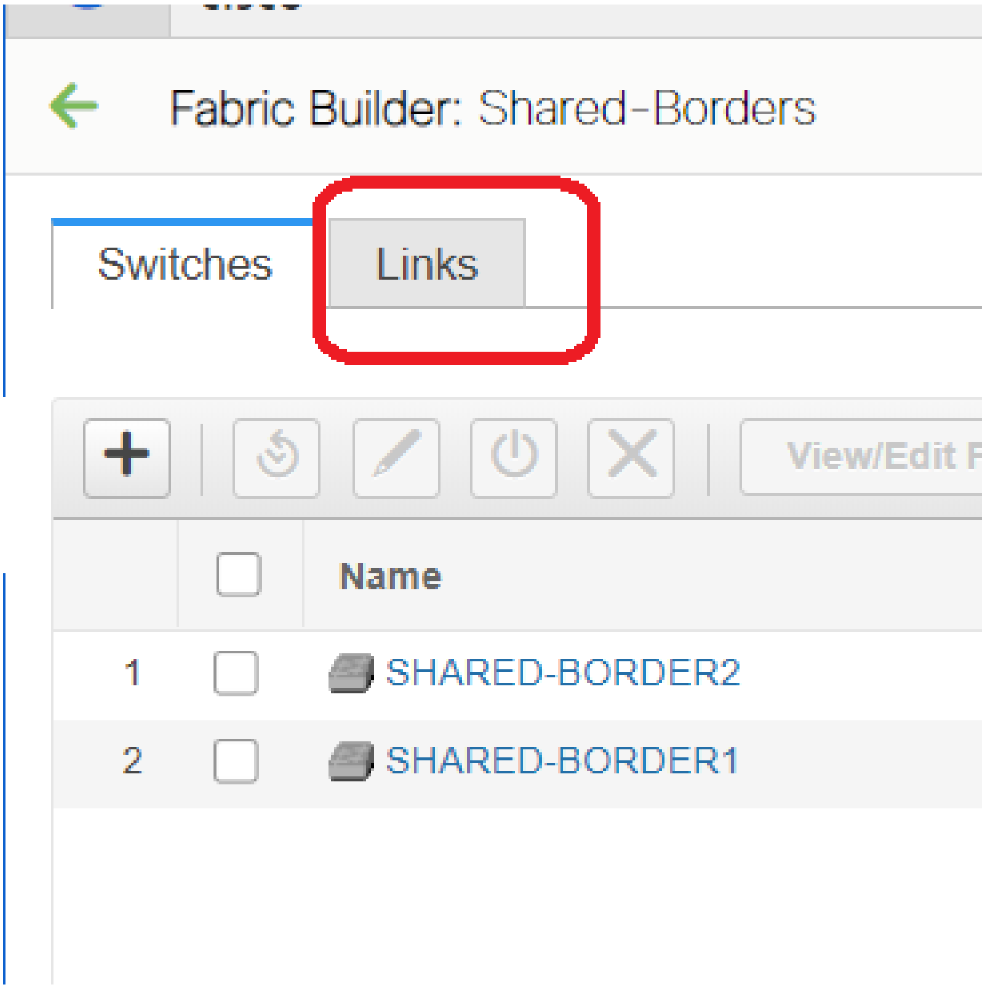
# For that purpose, Shared borders are terminating the VRF Lite; Here in this example eBGP is running from Shared borders to the External routers as shown in the diagram in the beginning.

# For configuring this from DCNM, it is required to **add vrf extension attachments**. Below steps are to be done for achieving the same.

a) Adding Inter-Fabric links from shared borders to External Routers

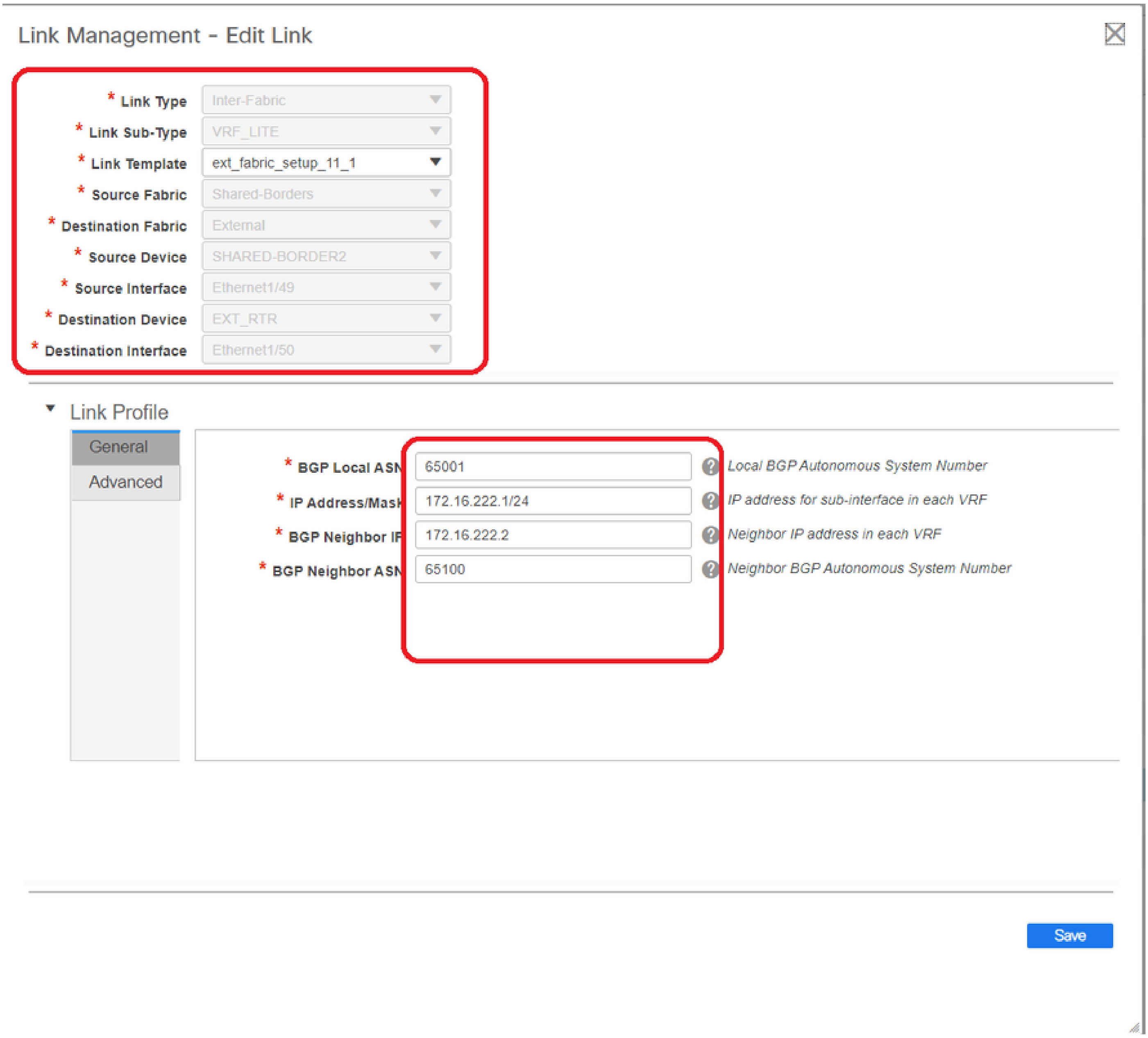


# Select the Fabric builder scope to "shared border" and Change to Tabular View



	□	Name
1	□	SHARED-BORDER2
2	□	SHARED-BORDER1

# Select the links and add an "Inter-Fabric" link as shown below



Link Management - Edit Link

* Link Type: Inter-Fabric

* Link Sub-Type: VRF_LITE

* Link Template: ext_fabric_setup_11_1

* Source Fabric: Shared-Borders

* Destination Fabric: External

* Source Device: SHARED-BORDER2

* Source Interface: Ethernet1/49

* Destination Device: EXT_RTR

* Destination Interface: Ethernet1/50

Link Profile

General | Advanced

* BGP Local ASN: 65001 — Local BGP Autonomous System Number

* IP Address/Mask: 172.16.222.1/24 — IP address for sub-interface in each VRF

* BGP Neighbor IP: 172.16.222.2 — Neighbor IP address in each VRF

* BGP Neighbor ASN: 65100 — Neighbor BGP Autonomous System Number

# A VRF LITE sub-type has to be selected from the drop down

# Source Fabric is Shared borders and destination Fabric is External as this is going to be a VRF LITE from SB to External

# Select the relevant interfaces that are going towards the external router

# Provide the IP address and mask and the neighbor IP address

# ASN Will be auto-populated.

# Once this is done, click the Save

# Perform the same for both the Shared borders and for all the external layer 3 connections that are in VRFLITE

b) Adding VRF Extensions

# Go to the Shared Border VRF section

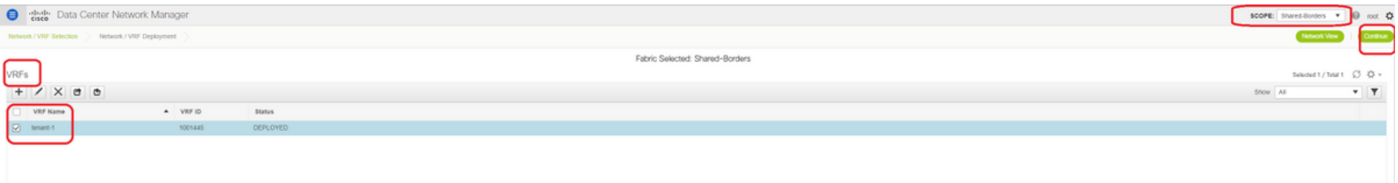
# VRF will be in deployed status; Select the checkbox at the right so that multiple switches can be selected

# Select the Shared borders and the "VRF EXtension attachment" window will open up

# Under "extend", change from "None" to "VRFLITE"

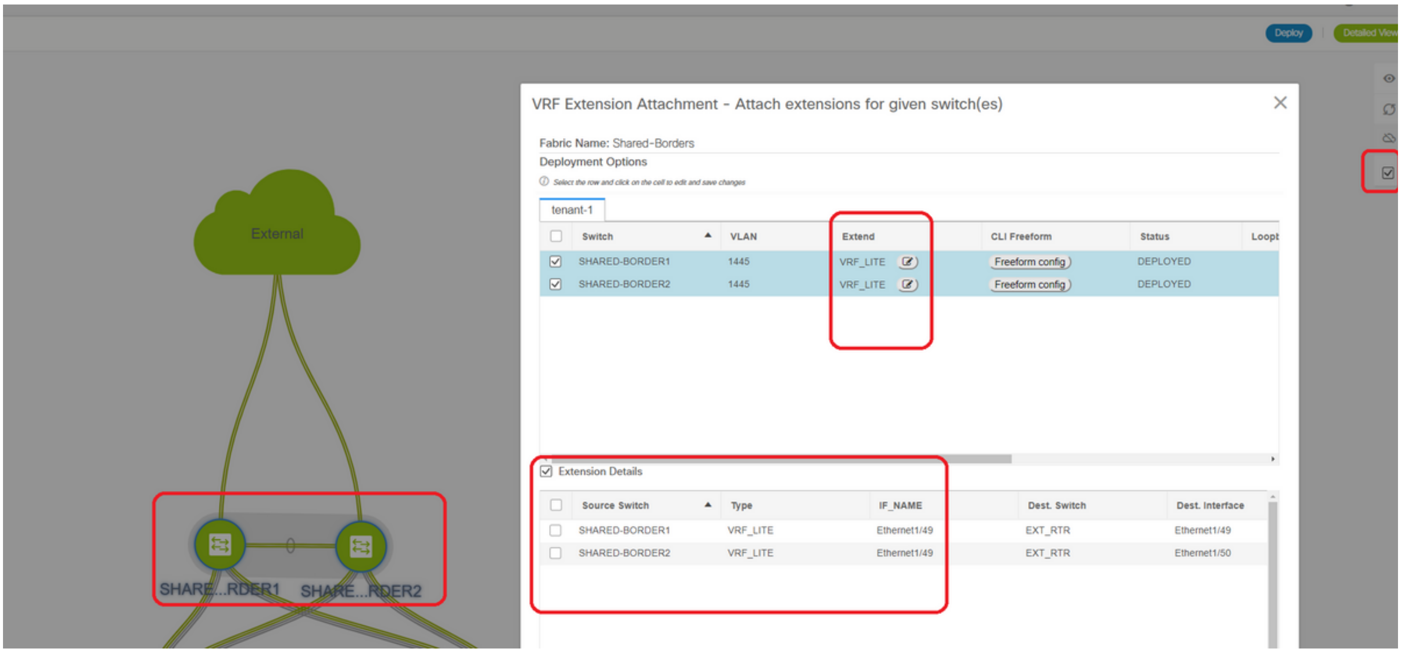
# Do the same for both shared borders

# Once that is done, "Extension Details" will populate the VRF LITE interfaces which were previously given in step a) above.



Fabric Selected: Shared-Borders

	VRF Name	VRF ID	Status
☑	tenant-1	1001445	DEPLOYED



VRF Extension Attachment - Attach extensions for given switch(es)

Fabric Name: Shared-Borders

Deployment Options

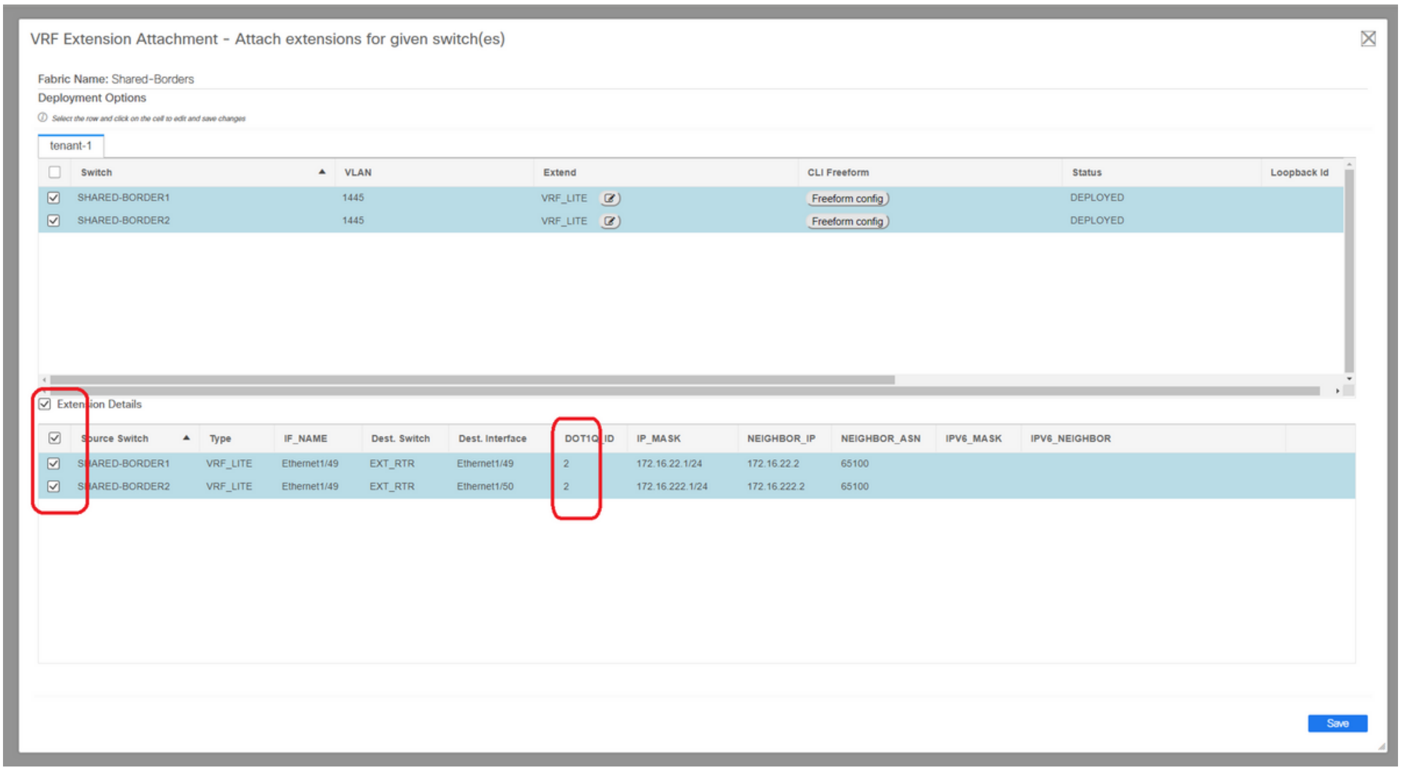
Select the row and click on the cell to edit and save changes

tenant-1

	Switch	VLAN	Extend	CLI Freeform	Status	Loopt
☑	SHARED-BORDER1	1445	VRF_LITE	Freeform config	DEPLOYED	
☑	SHARED-BORDER2	1445	VRF_LITE	Freeform config	DEPLOYED	

☑ Extension Details

	Source Switch	Type	IF_NAME	Dest. Switch	Dest. Interface
□	SHARED-BORDER1	VRF_LITE	Ethernet1/49	EXT_RTR	Ethernet1/49
□	SHARED-BORDER2	VRF_LITE	Ethernet1/49	EXT_RTR	Ethernet1/50



VRF Extension Attachment - Attach extensions for given switch(es)

Fabric Name: Shared-Borders

Deployment Options

ⓘ Select the row and click on the cell to edit and save changes

tenant-1

□	Switch ▲	VLAN	Extend	CLI Freeform	Status	Loopback Id
☑	SHARED-BORDER1	1445	VRF_LITE	Freeform config	DEPLOYED	
☑	SHARED-BORDER2	1445	VRF_LITE	Freeform config	DEPLOYED	

☑ Extension Details

☑	Source Switch ▲	Type	IF_NAME	Dest. Switch	Dest. Interface	DOT1Q_ID	IP_MASK	NEIGHBOR_IP	NEIGHBOR_ASN	IPV6_MASK	IPV6_NEIGHBOR
☑	SHARED-BORDER1	VRF_LITE	Ethernet1/49	EXT_RTR	Ethernet1/49	2	172.16.22.1/24	172.16.22.2	65100		
☑	SHARED-BORDER2	VRF_LITE	Ethernet1/49	EXT_RTR	Ethernet1/50	2	172.16.222.1/24	172.16.222.2	65100		

# DOT1Q ID is auto populated to 2

# Other fields are also auto populated

# If IPv6 neighborship has to be established via VRFLITE, step a) should be done for IPv6

# Now click the Save

# Finally, do the "Deploy" at the top right of the web page.

# A successful deploy will result in pushing configurations to the Shared borders which includes setting IP addresses on those sub-interfaces and establishing BGP IPv4 Neighborships with the external routers

# Keep in mind that the external router configurations(setting IP addresses on sub-interfaces and BGP Neighborship statements) are done manually by CLI in this case.

# CLI Verifications can be done by below commands on both the shared borders;

```
SHARED-BORDER1# sh ip bgp sum vr tenant-1
BGP summary information for VRF tenant-1, address family IPv4 Unicast
BGP router identifier 172.16.22.1, local AS number 65001
BGP table version is 18, IPv4 Unicast config peers 1, capable peers 1
9 network entries and 11 paths using 1320 bytes of memory
BGP attribute entries [9/1476], BGP AS path entries [3/18]
BGP community entries [0/0], BGP clusterlist entries [0/0]

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
172.16.22.2     4 65100      20      20       18    0    0 00:07:59 1

SHARED-BORDER2# sh ip bgp sum vr tenant-1
BGP summary information for VRF tenant-1, address family IPv4 Unicast
BGP router identifier 172.16.222.1, local AS number 65001
BGP table version is 20, IPv4 Unicast config peers 1, capable peers 1
```

```
9 network entries and 11 paths using 1320 bytes of memory
BGP attribute entries [9/1476], BGP AS path entries [3/18]
BGP community entries [0/0], BGP clusterlist entries [0/0]

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
172.16.222.2    4 65100      21      21       20    0    0 00:08:02 1
```

# With all the above Configurations, North/South reachability will be established too as shown below(pings from the External Router to Hosts in Fabric)

```
EXT_RTR# ping 172.16.144.1                                   # 172.16.144.1 is Host in DC1
Fabric
PING 172.16.144.1 (172.16.144.1): 56 data bytes
64 bytes from 172.16.144.1: icmp_seq=0 ttl=251 time=0.95 ms
64 bytes from 172.16.144.1: icmp_seq=1 ttl=251 time=0.605 ms
64 bytes from 172.16.144.1: icmp_seq=2 ttl=251 time=0.598 ms
64 bytes from 172.16.144.1: icmp_seq=3 ttl=251 time=0.568 ms
64 bytes from 172.16.144.1: icmp_seq=4 ttl=251 time=0.66 ms
^[[A^[[A
--- 172.16.144.1 ping statistics ---
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min/avg/max = 0.568/0.676/0.95 ms
EXT_RTR# ping 172.16.144.2 # 172.16.144.2 is Host in DC2 Fabric
PING 172.16.144.2 (172.16.144.2): 56 data bytes
64 bytes from 172.16.144.2: icmp_seq=0 ttl=251 time=1.043 ms
64 bytes from 172.16.144.2: icmp_seq=1 ttl=251 time=6.125 ms
64 bytes from 172.16.144.2: icmp_seq=2 ttl=251 time=0.716 ms
64 bytes from 172.16.144.2: icmp_seq=3 ttl=251 time=3.45 ms
64 bytes from 172.16.144.2: icmp_seq=4 ttl=251 time=1.785 ms

--- 172.16.144.2 ping statistics ---
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min/avg/max = 0.716/2.623/6.125 ms
```

# Traceroutes also point to the right devices in the path of the packet

```
EXT_RTR# traceroute 172.16.144.1
traceroute to 172.16.144.1 (172.16.144.1), 30 hops max, 40 byte packets
 1  SHARED-BORDER1 (172.16.22.1)  0.914 ms  0.805 ms  0.685 ms
 2  DC1-BGW2 (172.17.10.2)  1.155 ms DC1-BGW1 (172.17.10.1)  1.06 ms  0.9 ms
 3  ANYCAST-VLAN144-IP (172.16.144.254) (AS 65000)  0.874 ms  0.712 ms  0.776 ms
 4  DC1-HOST (172.16.144.1) (AS 65000)  0.605 ms  0.578 ms  0.468 ms
EXT_RTR# traceroute 172.16.144.2 traceroute to 172.16.144.2 (172.16.144.2), 30 hops max, 40 byte
packets 1 SHARED-BORDER2 (172.16.222.1) 1.137 ms 0.68 ms 0.66 ms 2 DC2-BGW2 (172.17.20.2) 1.196
ms DC2-BGW1 (172.17.20.1) 1.193 ms 0.903 ms 3 ANYCAST-VLAN144-IP (172.16.144.254) (AS 65000)
1.186 ms 0.988 ms 0.966 ms 4 172.16.144.2 (172.16.144.2) (AS 65000) 0.774 ms 0.563 ms 0.583 ms
EXT_RTR#
```