

# Building a Nexus 9000 VXLAN Multisite TRM using DCNM

## Contents

[Introduction](#)

[Topology](#)

[Details of the Topology](#)

[PIM/Multicast Details\(TRM Specific\)](#)

[Components Used](#)

[High Level Steps](#)

[Step 1: Creation of Easy Fabric for DC1](#)

[Step 2: Creation of Easy Fabric for DC2](#)

[Step 3: Creation of MSD For Multisite](#)

[Step 4: Moving DC1 and DC2 Fabrics into Multisite MSD](#)

[Step 5: Creation of VRFs](#)

[Step 6: Creation of Networks](#)

[Step 7: Creation of External Fabric for the DCI Switches](#)

[Step 8: Adding switches into each Fabric](#)

[Step 9: TRM settings for Individual Fabrics](#)

[Step 10: VRFLITE Configuration on Border Gateways](#)

[Step 11: Multisite Underlay Configuration on Border Gateways](#)

[Step 12: Multisite Overlay Settings for TRM](#)

[Step 13: Save/Deploy in MSD and Individual Fabrics](#)

[Step 14: VRF Extension attachments for MSD](#)

[Step 15: Pushing Network configurations to the Fabric from MSD](#)

[Step 16: Verifying VRF and Networks on all VRFs](#)

[Step 17: Deploying configurations on External Fabric](#)

[Step 18: Configuring iBGP Between DCI switches](#)

[Step 19: Verification of IGP/BGP neighborships](#)

[OSPF neighborships](#)

[BGP neighborships](#)

[BGP MVPN Neighborships for TRM](#)

[Step 20: Tenant VRF Loopback Creation on Border Gateway switches](#)

[Step 21: VRFLITE configurations on DCI switches](#)

[Unicast Verifications](#)

[East/West from DC1-Host1 to DC2-Host1](#)

[North/South from DC1-Host1 to PIM RP\(10.200.200.100\)](#)

[Multicast Verifications](#)

[Source in Non-vxlan\(behind Core Switch\), Receiver in DC2](#)

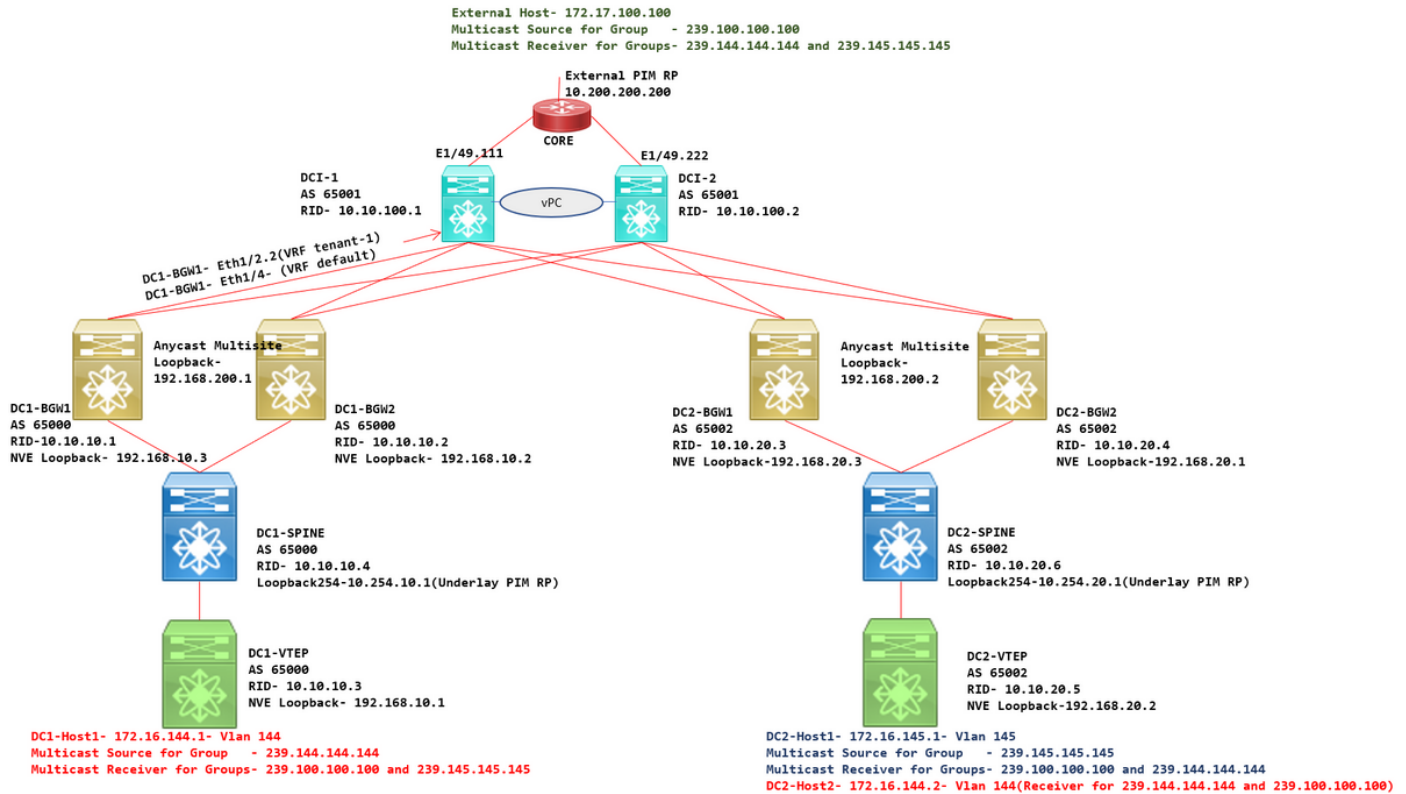
[Source in DC1, Receiver in DC2 as well as external](#)

[Source in DC2, Receiver in DC1 as well as external](#)

# Introduction

This document is to explain how to deploy a Cisco Nexus 9000 VXLAN Multisite TRM Fabric where Border Gateways are connected via DCI Switches

## Topology



## Details of the Topology

- DC1 and DC2 are two Datacenter Locations which are running VXLAN.
- DC1 and DC2 Border Gateways are connected to each other via DCI Switches.
- DCI switches do not run any VXLAN; Those are running eBGP for the underlay for reachability from DC1 to DC2 and Vice Versa. Also the DCI Switches are configured with the tenant vrf; In this example, it would be vrf- "tenant-1".
- DCI switches also connect to External Networks which are non-VXLAN.
- VRFLITE connections are terminated on Border Gateways(Support of Co-existence of VRFLITE and Border Gateway functions started from NXOS-9.3(3) and DCNM-11.3(1))
- Border Gateways are running in Anycast Mode; When running TRM(Tenant Routed Multicast) on this version, Border Gateways cannot be configured as vPC(refer Multisite TRM Configuration guide for other limitations)
- For this topology, All BGW switches will have Two Physical Connections towards each of the DCI switches; One link will be in default VRF(which will be used for the Inter-site Traffic) and other link will be in VRF tenant-1 which is used to extend VRFLITE out to the non-vxlan environment.

## PIM/Multicast Details(TRM Specific)

- Underlay PIM RP for both sites are the Spine switches and Loopback254 is configured for the same. Underlay PIM RP is used so that the VTEPs can send PIM Registers as well as PIM Joins to the Spines(for the Purposes of BUM traffic replication for various VNIDs)
- For TRM, RP can be specified by different means; Here for the purpose of the document, PIM RP is the core Router at the top of the topology which is external to the VXLAN Fabric.
- All VTEPs will have the Core router pointed as PIM RP configured in respective VRFs
- DC1-Host1 is sending multicast to the group- 239.144.144.144; DC2-Host1 is receiver for this group in DC2 and a Host External(172.17.100.100) to the vxlan is also subscribing to this group
- DC2-Host1 is sending multicast to the group- 239.145.145.145; DC1-Host1 is receiver for this group in DC1 and a Host External(172.17.100.100) to the vxlan is also subscribing to this group
- DC2-Host2 is in Vlan 144 and is receiver for Multicast groups- 239.144.144.144 and 239.100.100.100
- External Host(172.17.100.100) is sending traffic for which both DC1-Host1 and DC2-Host1 are receivers.
- This covers East/West Inter and Intra Vlan and North/South Multicast Traffic Flows

## Components Used

- Nexus 9k switches running 9.3(3)
- DCNM Running 11.3(1)

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 this Document is based on Two DCs utilizing VXLAN Multisite feature, Two Easy Fabrics have to be created
- 2) Create MSD and move DC1 and DC2
- 3) Create External Fabric and add DCI switches
- 4) Create Multisite Underlay and Overlay
- 5) Create VRF Extension attachments on Border Gateways
- 6) Verification of Unicast Traffic
- 7) Verification of Multicast Traffic

## Step 1: Creation of Easy Fabric for DC1

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

Good morning, admin!  
Let's get started.



**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.

Create Fabric

- Next is to provide the Fabric Name, Template and then under "General" Tab, Fill up the relevant ASN, fabric interface numbering, Any Cast Gateway MAC(AGM)

## Add Fabric

\* Fabric Name :

\* Fabric Template :

General | Replication | vPC | Protocols | Advanced | Resources | Manageability | Bootstrap | Configuration Backup

\* BGP ASN  ⓘ 1-4294967295 | 1-65535[ 0-65535]

Enable IPv6 Underlay  ⓘ

Enable IPv6 Link-Local Address  ⓘ

\* Fabric Interface Numbering  ⓘ Numbered(Point-to-Point) or Unnumbered

\* Underlay Subnet IP Mask  ⓘ Mask for Underlay Subnet IP Range

Underlay Subnet IPv6 Mask  ⓘ Mask for Underlay Subnet IPv6 Range

\* Link-State Routing Protocol  ⓘ Supported routing protocols (OSPF/IS-IS)

\* Route-Reflectors  ⓘ Number of spines acting as Route-Reflectors

\* Anycast Gateway MAC  ⓘ 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

# AGM is used by Hosts in the fabric as the Default Gateway MAC address. This will be the same on all leaf switches(as all Leaf switches within the fabric are running anycast Fabric Forwarding). Default Gateway IP address and MAC address are going to be the same on all Leaf switches

- Next is to set the Replication mode

## Add Fabric

\* Fabric Name : DC1

\* Fabric Template : Easy\_Fabric\_11\_1

General | **Replication** | vPC | Protocols | Advanced | Resources | Manageability | Bootstrap | Configuration Backup

\* Replication Mode : Multicast ? Replication Mode for BUM Traffic

\* Multicast Group Subnet : 239.1.1.0/24 ? 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 : 239.1.1.0 ? 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 ? (Min:0, Max: 1023)

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

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

Underlay Second Backup RP Loopback Id : ? Used for second Fallback Bidir-PIM Phantom RP (Min:0, Max:1023)

Underlay Third Backup RP Loopback Id : ? Used for third Fallback Bidir-PIM Phantom RP (Min:0, Max:1023)

# Replication mode for this document purpose is Multicast; Another option is to use the Ingress Replication(IR)

# Multicast group subnet will be the multicast group used by VTEPs to replicate BUM Traffic(Like ARP requests)

# Check box for "Enable Tenant Routed Multicast(TRM)" has to be enabled

# Populate other boxes as required.

- Tab for vPC is left untouched as the topology here is not using any vPC
- Next is to the Protocols tab

## Add Fabric

\* Fabric Name :

\* Fabric Template :

General	Replication	vPC	Protocols	Advanced	Resources	Manageability	Bootstrap	Configuration Backup
<p>* Underlay Routing Loopback Id <input type="text" value="0"/> ? (Min:0, Max:1023)</p> <p>* Underlay VTEP Loopback Id <input type="text" value="1"/> ? (Min:0, Max:1023)</p> <p>Underlay Anycast Loopback Id <input type="text"/> ? Used for vPC Peering in VXLANv6 Fabrics (Min:0, Max:1023)</p> <p>* Link-State Routing Protocol Tag <input type="text" value="UNDERLAY"/> ? Routing Process Tag (Max Size 20)</p> <p>* OSPF Area Id <input type="text" value="0.0.0.0"/> ? OSPF Area Id in IP address format</p> <p><b>Enable OSPF Authentication</b> <input type="checkbox"/> ?</p> <p>OSPF Authentication Key ID <input type="text"/> ? (Min:0, Max:255)</p> <p>OSPF Authentication Key <input type="text"/> ? 3DES Encrypted</p> <p>IS-IS Level <input type="text"/> ? Supported IS types: level-1, level-2</p> <p>Enable IS-IS Authentication <input type="checkbox"/> ?</p> <p>IS-IS Authentication Keychain Name <input type="text"/> ?</p> <p>IS-IS Authentication Key ID <input type="text"/> ? (Min:0, Max:65535)</p> <p>IS-IS Authentication Key <input type="text"/> ? Cisco Type 7 Encrypted</p> <p><b>Enable BGP Authentication</b> <input type="checkbox"/> ?</p> <p>BGP Authentication Key Encryption Type <input type="text"/> ? BGP Key Encryption Type: 3 - 3DES, 7 - Cisco</p> <p>BGP Authentication Key <input type="text"/> ? Encrypted BGP Authentication Key based on type</p> <p><b>Enable BFD</b> <input type="checkbox"/> ? Valid for IPv4 Underlay only</p> <p>Enable BFD For IBGP <input type="checkbox"/> ?</p> <p>Enable BFD For OSPF <input type="checkbox"/> ?</p> <p>Enable BFD For ISIS <input type="checkbox"/> ?</p> <p>Enable BFD For PIM <input type="checkbox"/> ?</p> <p>Enable BFD Authentication <input type="checkbox"/> ?</p> <p>BFD Authentication Key ID <input type="text"/> ?</p> <p>BFD Authentication Key <input type="text"/> ? Encrypted SHA1 secret value</p>								

# Modify the relevant boxes as needed.

- Next is Advanced tab

## Add Fabric

\* Fabric Name :

\* Fabric Template :

General	Replication	vPC	Protocols	Advanced	Resources	Manageability	Bootstrap	Configuration Backup
* VRF Template	<input type="text" value="Default_VRF_Universal"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
* Network Template	<input type="text" value="Default_Network_Universal"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
* VRF Extension Template	<input type="text" value="Default_VRF_Extension_Universal"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
* Network Extension Template	<input type="text" value="Default_Network_Extension_Universa"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Site Id	<input type="text" value="65000"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
* Intra Fabric Interface MTU	<input type="text" value="9216"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
* Layer 2 Host Interface MTU	<input type="text" value="9216"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
* Power Supply Mode	<input type="text" value="ps-redundant"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
* CoPP Profile	<input type="text" value="strict"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
VTEP HoldDown Time	<input type="text" value="180"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Brownfield Overlay Network Name Format	<input type="text" value="Auto_Net_VNISSVNISS_VLANSSVLAN_"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Enable VXLAN OAM	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Enable Tenant DHCP	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Enable NX-API	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Enable NX-API on HTTP	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Enable Policy-Based Routing (PBR)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Enable Strict Config Compliance	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Enable AAA IP Authorization	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Enable DCNM as Trap Host	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
* Greenfield Cleanup Option	<input type="text" value="Disable"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Enable Precision Time Protocol (PTP)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
PTP Source Loopback Id	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
PTP Domain Id	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Enable MPLS Handoff	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

# For this document purpose, all fields are left at default.

# ASN is auto populated from the one that was provided within General tab

- Next is to fill up the fields in "Resources" tab



## Add Fabric

\* Fabric Name :

\* Fabric Template :

---

General | Replication | vPC | Protocols | Advanced | **Resources** | Manageability | Bootstrap | Configuration Backup

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  ? Typically Loopback1 and Anycast Loopback IPv6 Address Range

Underlay Subnet IPv6 Range  ? IPv6 Address range to assign Numbered and Peer Link SVI IPs

BGP Router ID Range for IPv6 Underlay  ?

\* Layer 2 VXLAN VNI Range  ? Overlay Network Identifier Range (Min:1, Max:16777214)

\* Layer 3 VXLAN VNI Range  ? Overlay VRF Identifier Range (Min:1, Max:16777214)

\* Network VLAN Range  ? Per Switch Overlay Network VLAN Range (Min:2, Max:3967)

\* VRF VLAN Range  ? Per Switch Overlay VRF VLAN Range (Min:2, Max:3967)

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

\* VRF Lite Deployment  ? VRF Lite Inter-Fabric Connection Deployment Options

\* 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:3967)

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

# Underlay Routing Loopback IP range would be the ones used for protocols like BGP, OSPF

# Underlay VTEP loopback IP range are the ones that will be used for the NVE interface.

# Underlay RP is for the PIM RP that is used for BUM multicast groups.

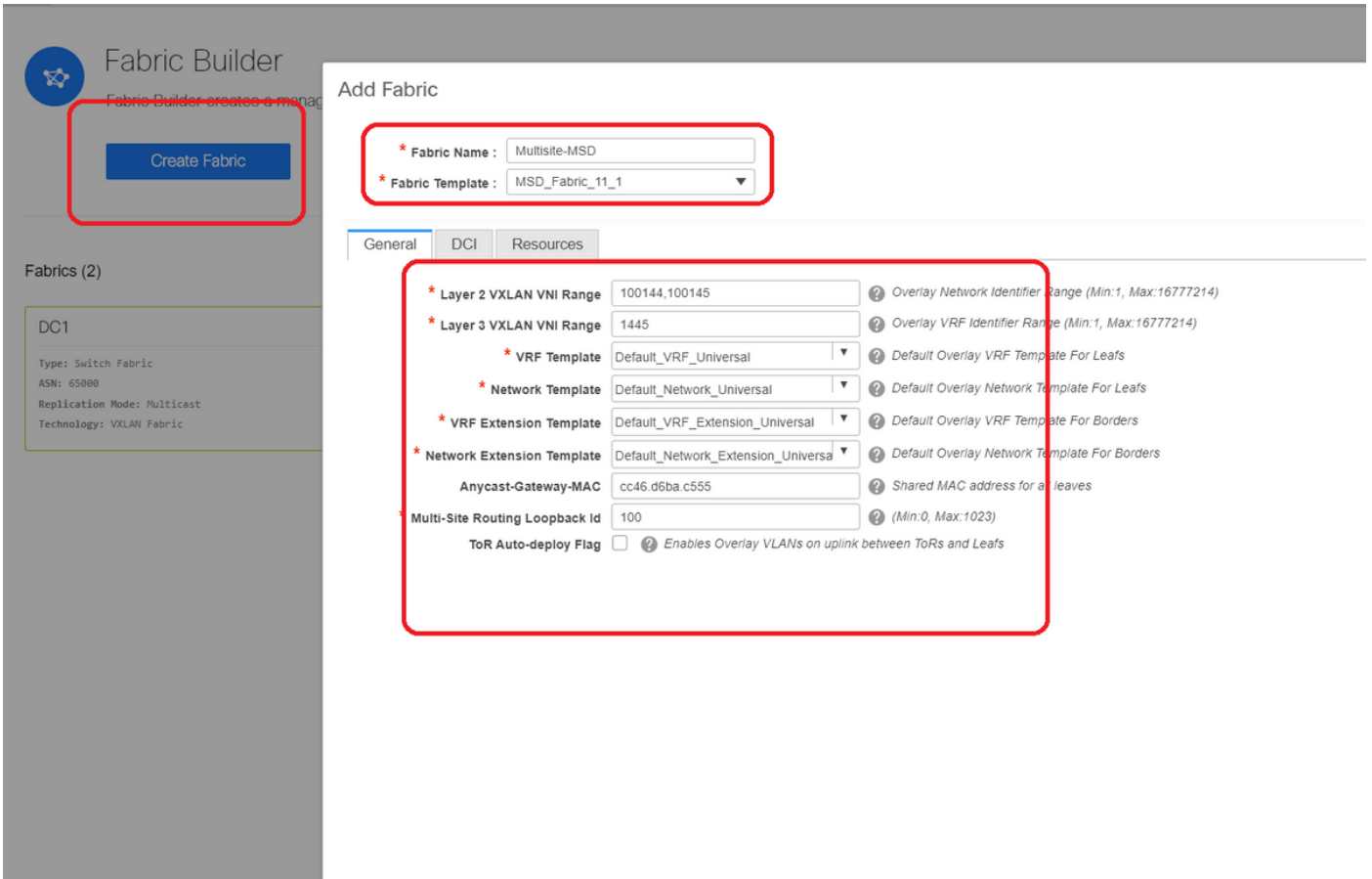
- Fill up other tabs with the relevant information and then "save"

## Step 2: Creation of Easy Fabric for DC2

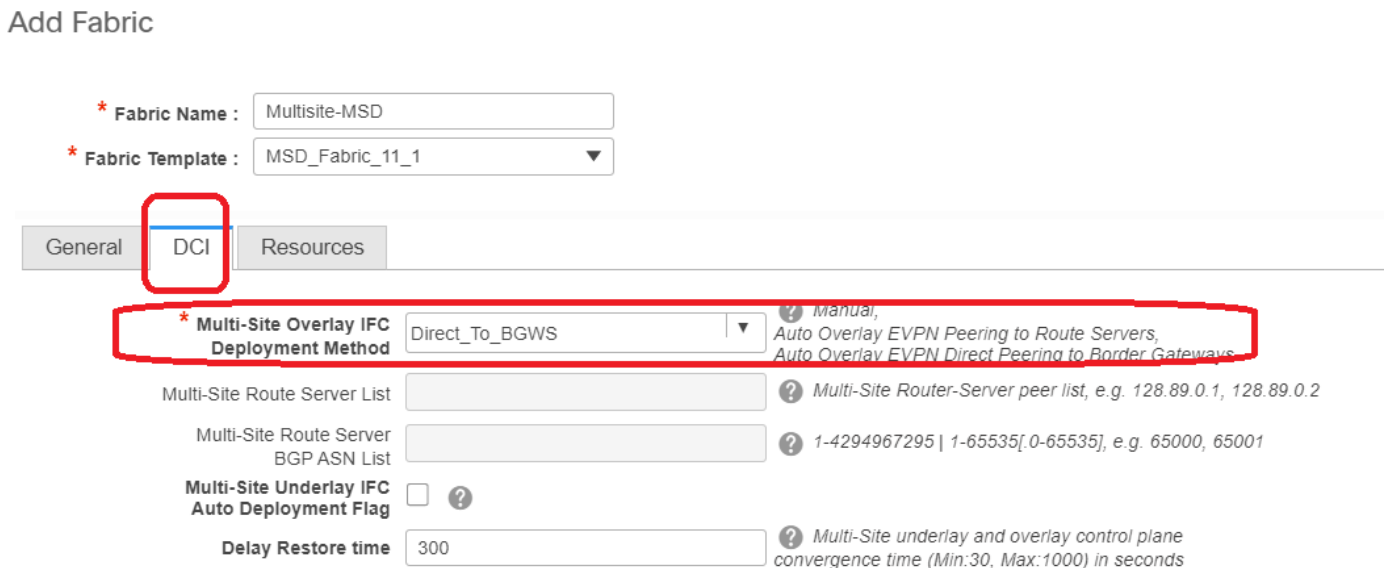
- Perform the same Task as in Step 1 to create an Easy Fabric for DC2
- Make Sure to provide different IP address block Under Resources for NVE and Routing Loopbacks and any other relevant areas
- ASNs should be different as well
- Layer 2 and Layer 2 VNIDs are same

## Step 3: Creation of MSD For Multisite

- An MSD fabric will have to be created as shown below.



- Fill up the DCI tab as well



# Multi-site Overlay IFC Deployment method is "Direct\_To\_BGWS" as here DC1-BGWs will form the Overlay connection with the DC2-BGWs. DCI switches shown in the topology are just transit layer 3 Devices(as well as VRFLITE)

- Next step is to mention the Multisite Loopback Range(This IP address will be used as the Multisite Loopback IP on DC1 and DC2 BGWs; DC1-BGW1 and DC1-BGW2 share the same

multisite Loopback IP; DC2-BGW1 and DC2-BGW2 share the same Multisite loopback IP but will be different from that of the DC1-BGWs

## Add Fabric

\* Fabric Name :

\* Fabric Template :

General DCI Resources

\* Multi-Site Routing Loopback IP Range  ? Typically Loopback100 IP Address Range

DCI Subnet IP Range  ? Address range to assign P2P DCI Links

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

# Once the fields are populated, Click the "save".

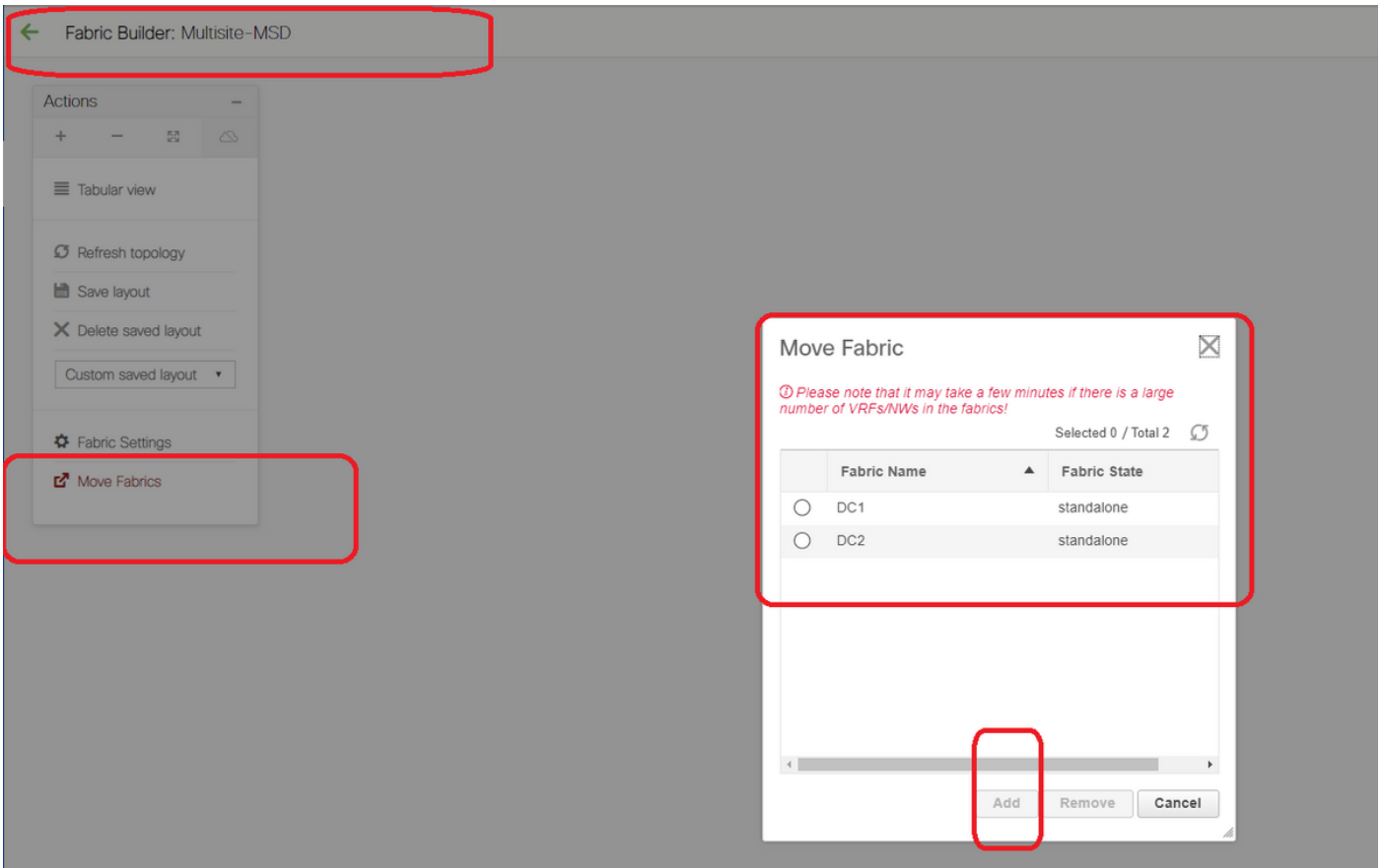
# Once steps 1 through 3 are done, the Fabric builder page will look like below.

Fabrics (3)

<b>DC1</b> <span>⚙️ ×</span> Type: Switch Fabric ASN: 65000 Replication Mode: Multicast Technology: VLAN Fabric	<b>DC2</b> <span>⚙️ ×</span> Type: Switch Fabric ASN: 65002 Replication Mode: Multicast Technology: VLAN Fabric	<b>Multisite-MSD</b> <span>⚙️ ×</span> Type: Multi-Fabric Domain Member Fabrics: None
---	---	---

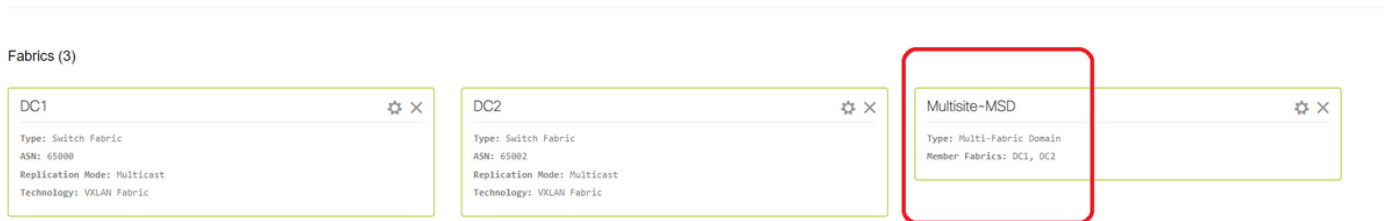
## Step 4: Moving DC1 and DC2 Fabrics into Multisite MSD

# In this step, the DC1 and DC2 fabrics are moved to Multisite-MSD which was created in Step 3. Below are the screenshots on how to achieve the same.



# Select the MSD, click on "move Fabrics" and then select DC1 and DC2 one by one and then "add".

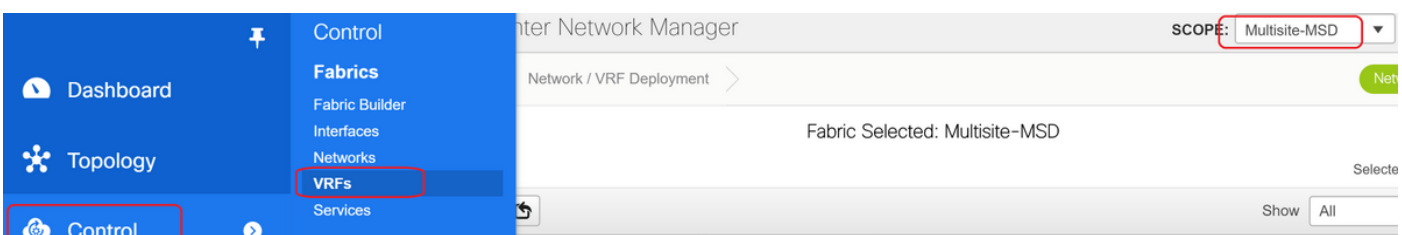
# Once both the fabrics are moved, the home page would look like below



# Multisite-MSD will show DC1 and DC2 as member fabrics

## Step 5: Creation of VRFs

# Creating VRFs can be done from MSD fabric which will be applicable for both the Fabrics. Below are the screenshots to achieve the same.



Network / VRF Selection

## Create VRF

**VRFs**

+ [edit] [delete]

VRF Name

No data available

**VRF Information**

\* VRF ID: 1445

\* VRF Name: tenant-1

\* VRF Template: Default\_VRF\_Universal

\* VRF Extension Template: Default\_VRF\_Extension\_Universal

VLAN ID: 1445 Propose VLAN ?

**VRF Profile**

General | Advanced

VRF Vlan Name:  ? if > 32 cha

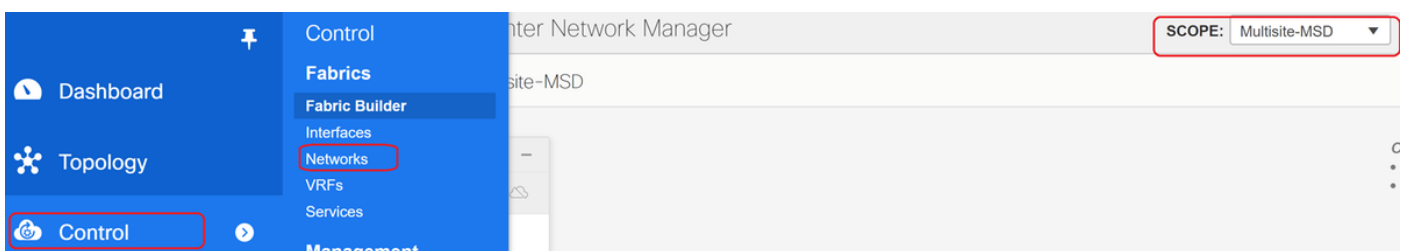
VRF Intf Description:  ?

VRF Description:  ?

# Fill in the advanced tab as well and then "create"

## Step 6: Creation of Networks

# Creating Vlans and corresponding VNIDs, SVIs can be done from MSD fabric which will be applicable for both Fabrics.



Network / VRF Sele

## Create Network

Networks

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

General

Advanced

- 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-name

Create Network

# In "advanced" tab, enable the checkbox if the BGWs are required to be the Gateway for the Networks

# Once all the fields are populated, Click "Create Network"

# Repeat the same steps for any other Vlans/Networks

## Step 7: Creation of External Fabric for the DCI Switches

# This example takes into consideration of DCI switches which are in the path of the packet from DC1 to DC2(as far as inter-site communication is concerned) which is commonly seen when there are more than 2 fabrics.

# External Fabric will include the Two DCI Switches that are at the top of the Topology shown in the beginning of this document

# Create the Fabri with the "external" template and specify the ASN

# Modify any other relevant fields for the deployment

**Fabric Builder**  
Fabric Builder creates a... using *Power On Auto P*

**Create Fabric**

**Fabrics (3)**

- DC1  
Type: Switch Fabric  
ASN: 65000  
Replication Mode: Multicast  
Technology: VXLAN Fabric

### Add Fabric

\* Fabric Name : DCI

\* Fabric Template : External\_Fabric\_11\_1

General | Advanced | Resources | Configuration Backup | Bootstrap

\* BGP AS # 65001 ? 1-4294967295 | 1-65535[0-6553

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

**Save**

## Step 8: Adding switches into each Fabric

# Here, all the switches per fabric will be added into the respective Fabric.

Procedure to add switches is shown in below screenshots.

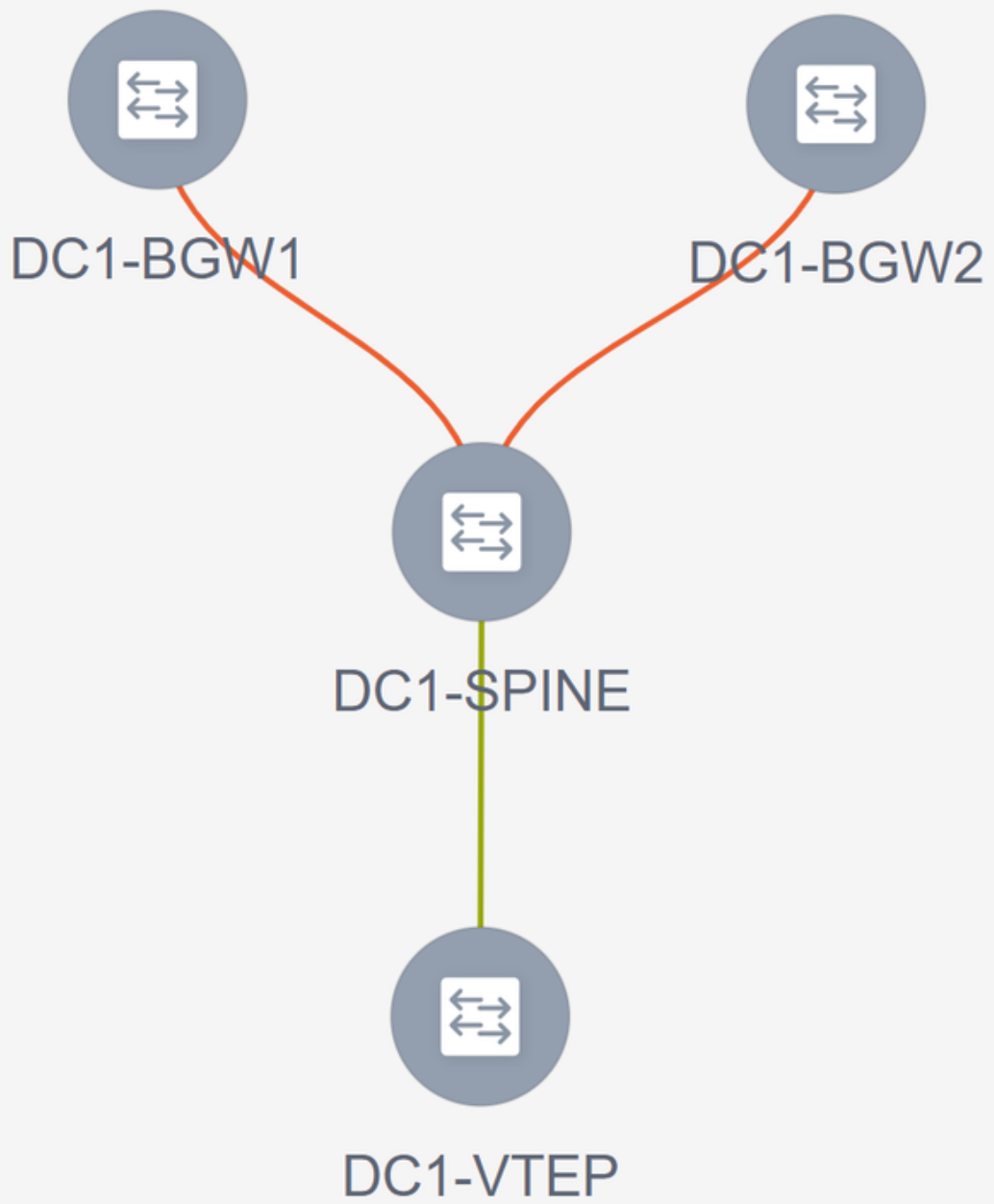
The screenshot shows the 'Fabric Builder: DC1' interface. On the left is a sidebar with 'Actions' including 'Add switches' (highlighted with a red box), 'Refresh topology', 'Save layout', 'Delete saved layout', 'Restore Fabric', 'Backup Now', 'Re-sync Fabric', and 'Fabric Settings'. The main area is titled 'Inventory Management' and has two tabs: 'Discover Existing Switches' (active) and 'PowerOn Auto Provisioning (POAP)'. Below the tabs are 'Discovery Information' and 'Scan Details' sections. The 'Discovery Information' section contains the following fields: 'Seed IP' (text input with '10.122.165.173,10.122.165.227,10' and an example below), 'Authentication Protocol' (dropdown menu set to 'MD5'), 'Username' (text input with 'admin'), 'Password' (password field with dots), 'Max Hops' (spinner set to '10'), and 'Preserve Config' (toggle switch set to 'no'). A note at the bottom of this section states: 'Selecting 'no' will clean up the configuration on switch(es)'. At the bottom of the main area is a blue 'Start discovery' button.

# If "Preseve Config" is "NO"; any switch configuration that is present will be erased; Exception is the hostname, Boot variable, MGMT0 IP address, Route in VRF Context Management

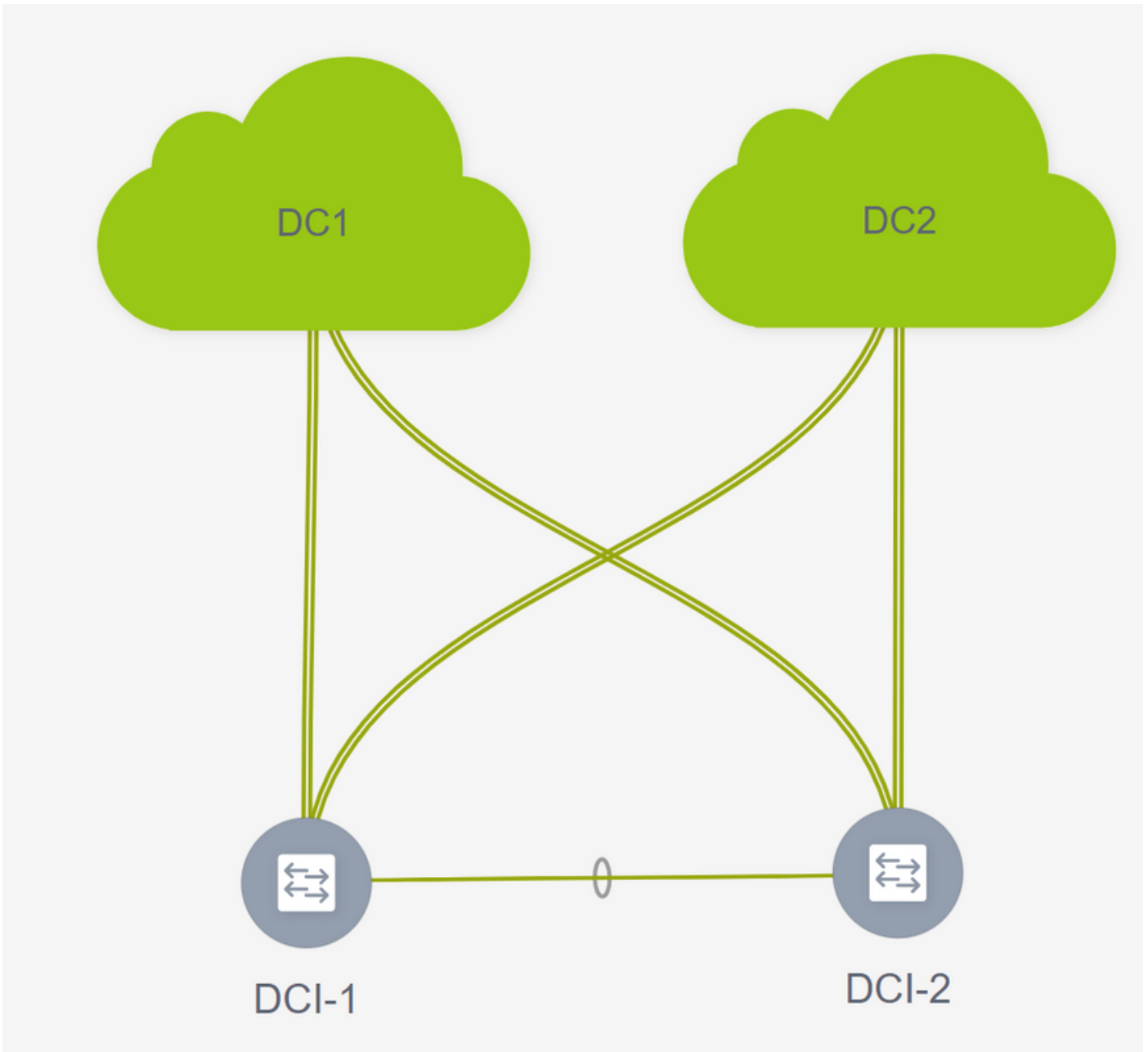
# Set the Roles on switches correctly(by Right click on switch, Set role and then relevant role

# Also arrange the layout of switches accordingly and then click "save layout"



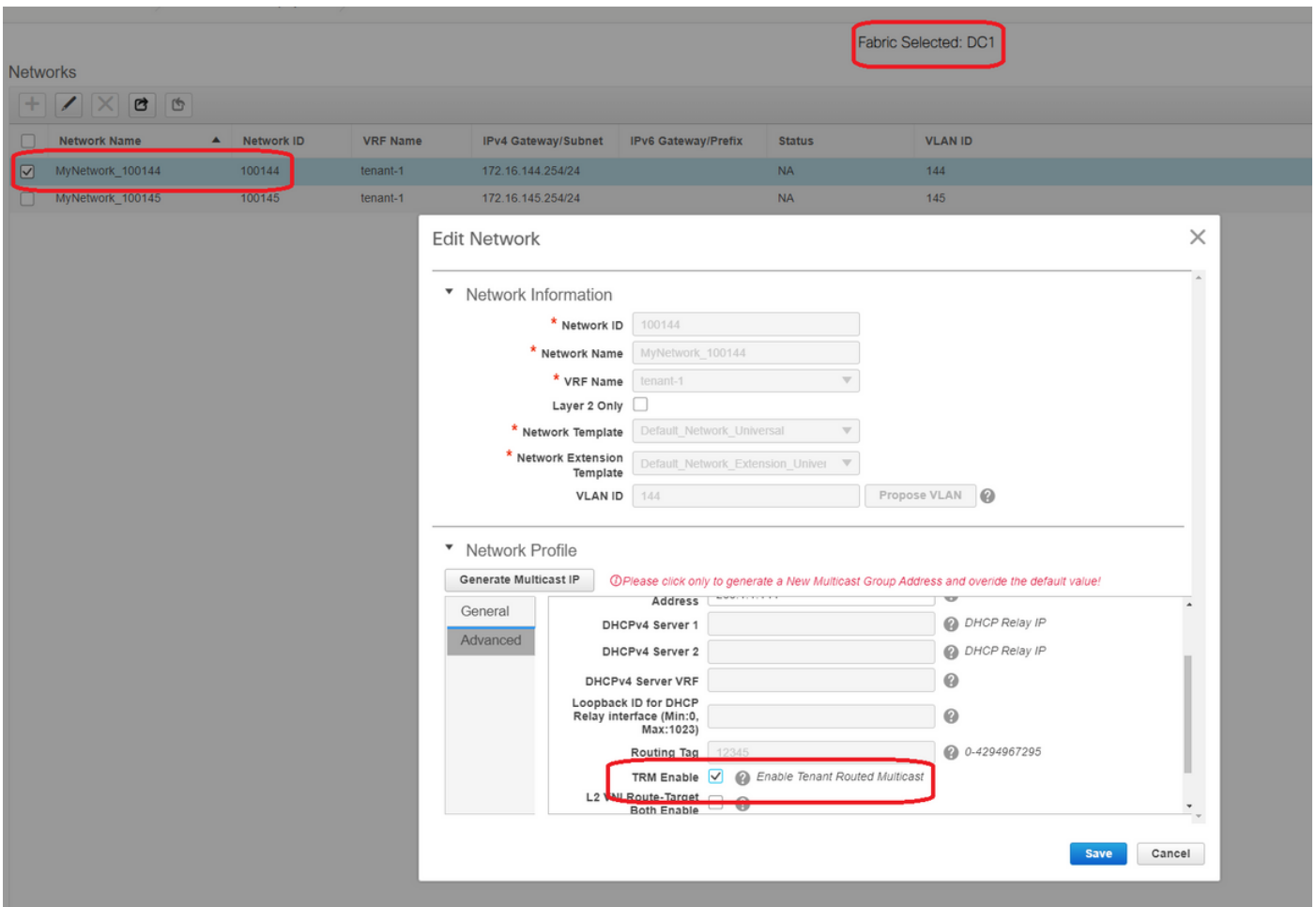






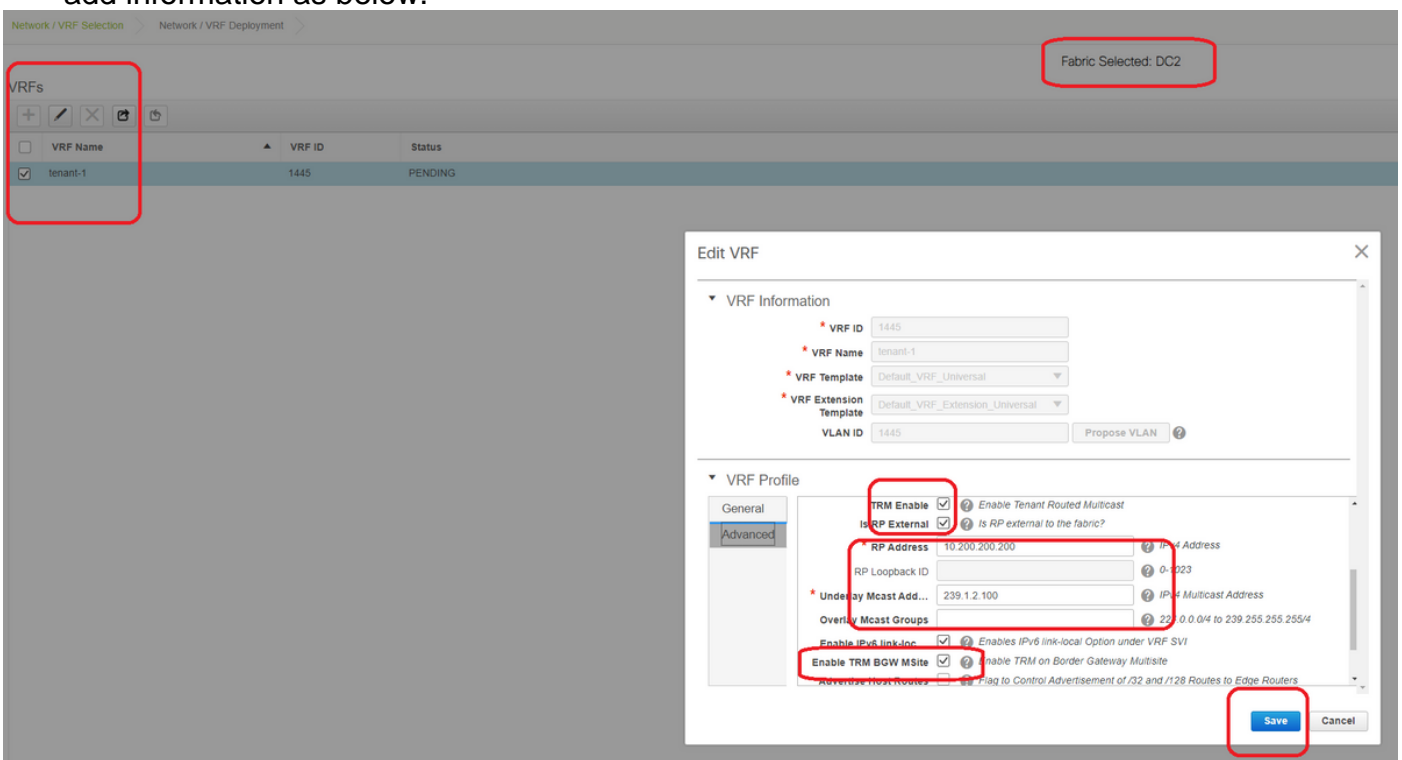
## Step 9: TRM settings for Individual Fabrics

- Next step is to enable TRM checkboxes on each Fabrics



# Perform this step for all Networks for all Fabrics.

- Once this is done, VRFs in individual fabrics are also required to make some changes and add information as below.



# This has to be done in DC1 and DC2 as well for the VRF section.

# Note that the multicast group for the VRF-> 239.1.2.100 was changed manually from the auto populated one; Best practice is to use different group for the Layer 3 VNI VRF and for any L2 VNI Vlans' BUM traffic multicast group

## **Step 10: VRFLITE Configuration on Border Gateways**

# Starting from NXOS 9.3(3) and DCNM 11.3(1), Border Gateways can act as Border Gateways and VRFLITE connectivity point(which will let the Border Gateway have a VRFLITE neighborship with an external router and so external devices can communicate with the devices in the fabric)

# For the purpose of this document, border Gateways are forming VRFLITE neighborship with the DCI router which are at the north of the topology shown above.

# One point to note is that; VRFLITE and Multisite Underlay Links cannot be the same physical Links. Separate links will have to be spun up to form the vrflite and multisite Underlay

# Screenshots below will illustrate how to achieve both VRF LITE and multisite extensions on Border Gateways.



Fabric Builder: Multisite-MSD

Actions



Tabular view



Refresh topology



Save layout



Delete saved layout

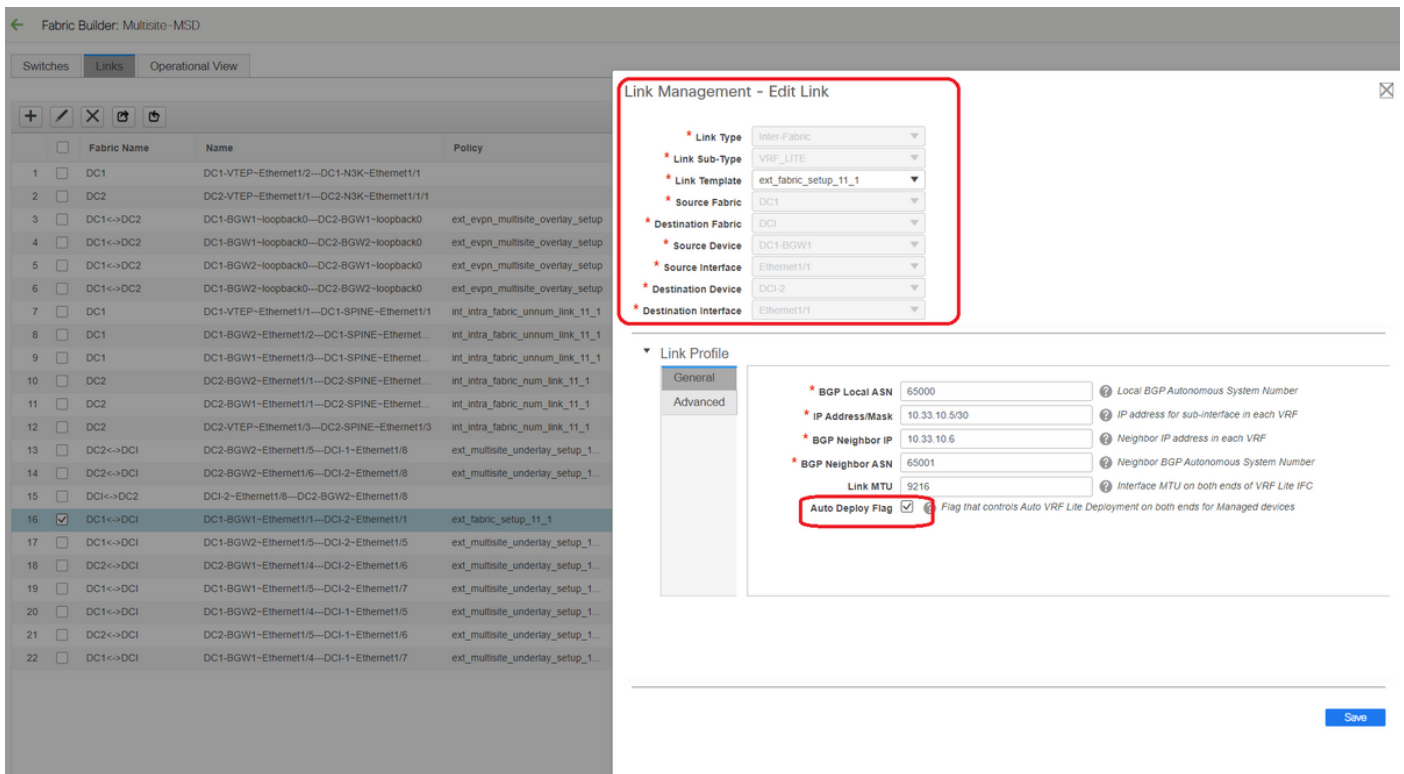
Custom saved layout



Fabric Settings



Move Fabrics



# Switch to "tabular view"

# Move to the tab "links" and then add an "inter-fabric VRFLITE" link and will have to specify the Source Fabric as DC1 and destination Fabric as DCI

# Select the right interface for source interface that leads to the correct DCI Switch

# under link profile, provide the local and remote IP addresses

# Also enable the check box- "auto deploy flag" so that the DCI switches' configuration for VRFLITE will be also auto populated(This is done in a future step)

# ASNs are auto populated

# Once all the fields are filled in with the correct information, Click the "save" button

- Above step will have to done for all BGW To DCI Connections on all 4 Border Gateways towards to the Two DCI Switches.
- Considering the topology of this document, there will be a total of 8 inter-fabric VRF LITE connections and it looks like below.

Switches Links Operational View							
+ ✎ ✕ ↺ ↻							
	<input type="checkbox"/>	Fabric Name	Name	Policy	Info	Admin State	Oper State
1	<input type="checkbox"/>	DC1	DC1-VTEP~Ethernet1/2---DC1-N3K~Ethernet1/1		Neighbor Present	Up:-	Up:-
2	<input type="checkbox"/>	DC2	DC2-VTEP~Ethernet1/1---DC2-N3K~Ethernet1/1/1		Neighbor Present	Up:-	Up:-
3	<input type="checkbox"/>	DC1	DC1-BGW2~Ethernet1/2---DC1-SPINE~Ethernet...	int_intra_fabric_unnum_link_11_1	Link Present	Up:Up	Up:Up
4	<input type="checkbox"/>	DC1	DC1-BGW1~Ethernet1/3---DC1-SPINE~Ethernet...	int_intra_fabric_unnum_link_11_1	Link Present	Up:Up	Up:Up
5	<input type="checkbox"/>	DC1	DC1-VTEP~Ethernet1/1---DC1-SPINE~Ethernet1/1	int_intra_fabric_unnum_link_11_1	Link Present	Up:Up	Up:Up
6	<input type="checkbox"/>	DC2	DC2-BGW2~Ethernet1/1---DC2-SPINE~Ethernet...		Link Present	Up:Up	Up:Up
7	<input type="checkbox"/>	DC2	DC2-VTEP~Ethernet1/3---DC2-SPINE~Ethernet1/3		Link Present	Up:Up	Up:Up
8	<input type="checkbox"/>	DC2	DC2-BGW1~Ethernet1/1---DC2-SPINE~Ethernet...		Link Present	Up:Up	Up:Up
9	<input type="checkbox"/>	DC2<->DC1	DC2-BGW2~Ethernet1/2---DC1-1~Ethernet1/4	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
10	<input type="checkbox"/>	DC2<->DC1	DC2-BGW2~Ethernet1/4---DC1-2~Ethernet1/4	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
11	<input type="checkbox"/>	DC1<->DC1	DC1-BGW1~Ethernet1/1---DC1-2~Ethernet1/1	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
12	<input type="checkbox"/>	DC1<->DC1	DC1-BGW2~Ethernet1/1---DC1-2~Ethernet1/2	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
13	<input type="checkbox"/>	DC2<->DC1	DC2-BGW1~Ethernet1/3---DC1-2~Ethernet1/3	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
14	<input type="checkbox"/>	DC2<->DC1	DC2-BGW1~Ethernet1/2---DC1-1~Ethernet1/3	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
15	<input type="checkbox"/>	DC1<->DC1	DC1-BGW1~Ethernet1/2---DC1-1~Ethernet1/1	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
16	<input type="checkbox"/>	DC1<->DC1	DC1-BGW2~Ethernet1/3---DC1-1~Ethernet1/2	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up

## Step 11: Multisite Underlay Configuration on Border Gateways

# Next step is to configure the Multisite Underlay on every Border Gateway in each Fabric.

# For this purpose, we will need separate physical links from BGWs to DCI switches. The links which were used for VRFLITE in step 10 cannot be used for multisite Overlay

# These interfaces will be part of "default vrf" unlike the previous one where the interfaces will be part of tenant vrf(this example, it is tenant-1)

# Below screenshots will help to walk through the steps to do this configuration.



Fabric Builder: Multisite-MSD

Switches Links Operational View

Link Management - Edit Link

\* Link Type: Inter-Fabric  
 \* Link Sub-Type: MULTISITE\_UNDERLAY  
 \* Link Template: ext\_multisite\_underlay\_setup\_...  
 \* Source Fabric: DC1  
 \* Destination Fabric: DC1  
 \* Source Device: DC1-BGW1  
 \* Source Interface: Ethernet1/4  
 \* Destination Device: DC1-1  
 \* Destination Interface: Ethernet1/7

\* BGP Local ASN: 65000  
 \* IP Address/Mask: 10.4.10.1/30  
 \* BGP Neighbor IP: 10.4.10.2  
 \* BGP Neighbor ASN: 65001  
 \* BGP Maximum Paths: 1  
 Routing TAG: 54321  
 Link MTU: 9216

Save

# The same step will have to be performed for all the connections from BGWs to DCI switches

# At the end, a total of 8 inter-fabric multisite underlay connections will be seen as below.

Fabric Builder: Multisite-MSD

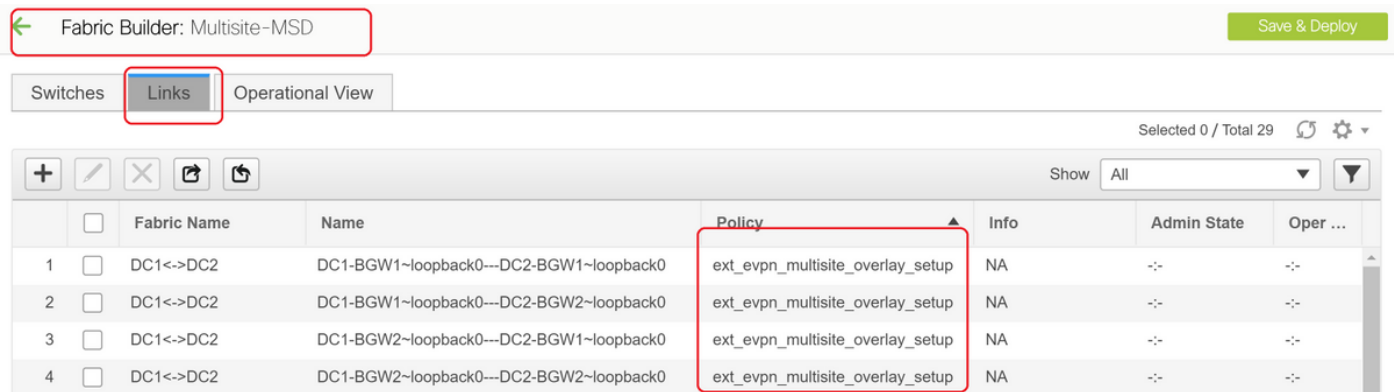
Switches Links Operational View

	<input type="checkbox"/>	Fabric Name	Name	Policy	Info	Admin State	Oper State
1	<input type="checkbox"/>	DC1<->DC2	DC1-BGW1~loopback0---DC2-BGW1~loopback0	ext_evpn_multisite_overlay_setup	NA	--	--
2	<input type="checkbox"/>	DC1<->DC2	DC1-BGW1~loopback0---DC2-BGW2~loopback0	ext_evpn_multisite_overlay_setup	NA	--	--
3	<input type="checkbox"/>	DC1<->DC2	DC1-BGW2~loopback0---DC2-BGW1~loopback0	ext_evpn_multisite_overlay_setup	NA	--	--
4	<input type="checkbox"/>	DC1<->DC2	DC1-BGW2~loopback0---DC2-BGW2~loopback0	ext_evpn_multisite_overlay_setup	NA	--	--
5	<input type="checkbox"/>	DC1<->DCI	DC1-BGW1~Ethernet1/1---DCI-2~Ethernet1/1	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
6	<input type="checkbox"/>	DC1<->DCI	DC1-BGW1~Ethernet1/2---DCI-1~Ethernet1/1	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
7	<input type="checkbox"/>	DC1<->DCI	DC1-BGW2~Ethernet1/1---DCI-2~Ethernet1/2	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
8	<input type="checkbox"/>	DC1<->DCI	DC1-BGW2~Ethernet1/3---DCI-1~Ethernet1/2	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
9	<input type="checkbox"/>	DC2<->DCI	DC2-BGW1~Ethernet1/2---DCI-1~Ethernet1/3	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
10	<input type="checkbox"/>	DC2<->DCI	DC2-BGW1~Ethernet1/3---DCI-2~Ethernet1/3	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
11	<input type="checkbox"/>	DC2<->DCI	DC2-BGW2~Ethernet1/4---DCI-2~Ethernet1/4	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
12	<input type="checkbox"/>	DC2<->DCI	DC2-BGW2~Ethernet1/2---DCI-1~Ethernet1/4	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
13	<input type="checkbox"/>	DC1<->DCI	DC1-BGW1~Ethernet1/4---DCI-1~Ethernet1/7	ext_multisite_underlay_setup_1...	Link Present	Up:Up	Up:Up
14	<input type="checkbox"/>	DC1<->DCI	DC1-BGW1~Ethernet1/5---DCI-2~Ethernet1/7	ext_multisite_underlay_setup_1...	Link Present	Up:Up	Up:Up
15	<input type="checkbox"/>	DC1<->DCI	DC1-BGW2~Ethernet1/4---DCI-1~Ethernet1/5	ext_multisite_underlay_setup_1...	Link Present	Up:Up	Up:Up
16	<input type="checkbox"/>	DC1<->DCI	DC1-BGW2~Ethernet1/5---DCI-2~Ethernet1/5	ext_multisite_underlay_setup_1...	Link Present	Up:Up	Up:Up
17	<input type="checkbox"/>	DC2<->DCI	DC2-BGW1~Ethernet1/4---DCI-2~Ethernet1/6	ext_multisite_underlay_setup_1...	Link Present	Up:Up	Up:Up
18	<input type="checkbox"/>	DC2<->DCI	DC2-BGW1~Ethernet1/5---DCI-1~Ethernet1/6	ext_multisite_underlay_setup_1...	Link Present	Up:Up	Up:Up
19	<input type="checkbox"/>	DC2<->DCI	DC2-BGW2~Ethernet1/6---DCI-2~Ethernet1/8	ext_multisite_underlay_setup_1...	Link Present	Up:Up	Up:Up
20	<input type="checkbox"/>	DC2<->DCI	DC2-BGW2~Ethernet1/5---DCI-1~Ethernet1/8	ext_multisite_underlay_setup_1...	Link Present	Up:Up	Up:Up

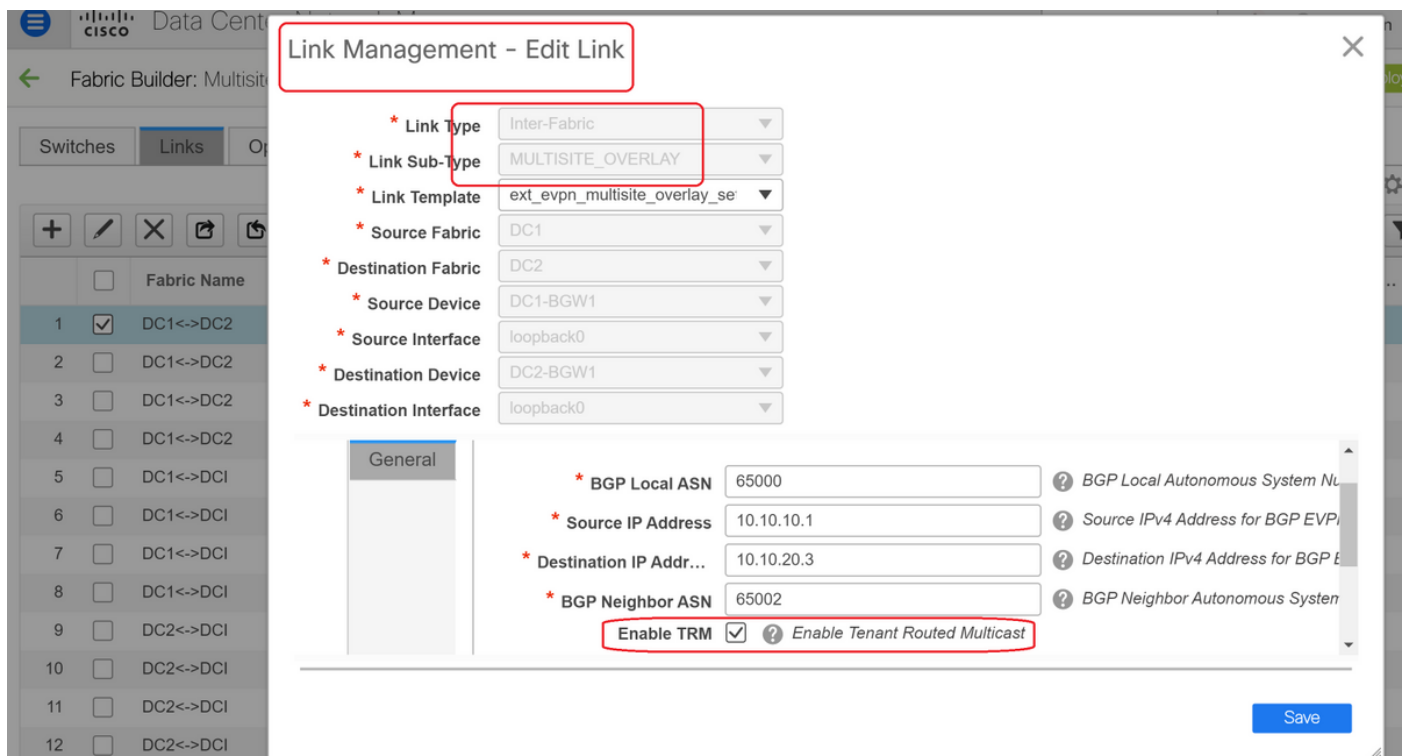
## Step 12: Multisite Overlay Settings for TRM

# When Multisite Underlay is completed, the multisite overlay interfaces/Links will be auto-populated and can be seen within the Tabular view under links within Multisite MSD fabric.

# By default, the Multisite Overlay will only form the bgp l2vpn evpn neighborhood from Each site BGWs to the other which is required for the unicast communication from one site to another. However, when Multicast is required to run between the sites(which are connected by teh vxlan multisite feature), it is required to enable the TRM checkbox as seen below for all the overlay interfaces within Multisite MSD Fabric. Screenshots will illustrate how to perform this.



	Fabric Name	Name	Policy	Info	Admin State	Oper ...
1	DC1<->DC2	DC1-BGW1~loopback0---DC2-BGW1~loopback0	ext_evpn_multisite_overlay_setup	NA	--	--
2	DC1<->DC2	DC1-BGW1~loopback0---DC2-BGW2~loopback0	ext_evpn_multisite_overlay_setup	NA	--	--
3	DC1<->DC2	DC1-BGW2~loopback0---DC2-BGW1~loopback0	ext_evpn_multisite_overlay_setup	NA	--	--
4	DC1<->DC2	DC1-BGW2~loopback0---DC2-BGW2~loopback0	ext_evpn_multisite_overlay_setup	NA	--	--



Link Management - Edit Link

\* Link Type: Inter-Fabric

\* Link Sub-Type: MULTISITE\_OVERLAY

\* Link Template: ext\_evpn\_multisite\_overlay\_se

\* Source Fabric: DC1

\* Destination Fabric: DC2

\* Source Device: DC1-BGW1

\* Source Interface: loopback0

\* Destination Device: DC2-BGW1

\* Destination Interface: loopback0

General

\* BGP Local ASN: 65000

\* Source IP Address: 10.10.10.1

\* Destination IP Addr...: 10.10.20.3

\* BGP Neighbor ASN: 65002

Enable TRM  Enable Tenant Routed Multicast

Save

## Step 13: Save/Deploy in MSD and Individual Fabrics

# Perform a save/deploy which will push relevant configurations as per the above steps that were done

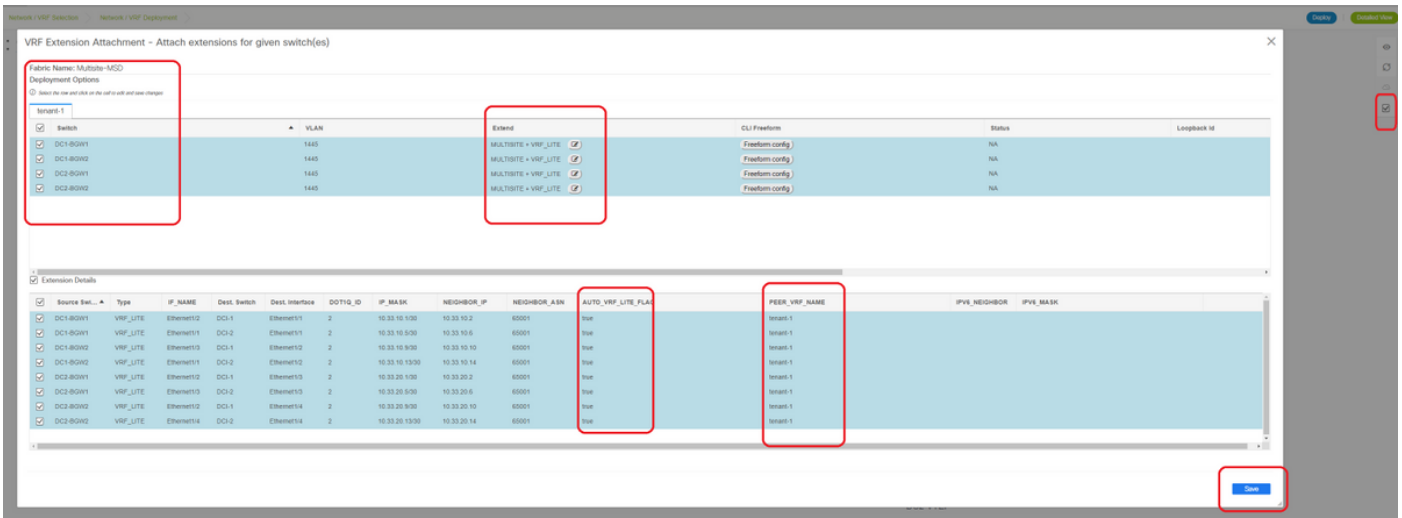
# When Selecting MSD, the configurations which will be pushed, will be only applicable for the Border Gateways.

# Hence it is required to save/deploy for the individual fabrics, which will push the relevant

configurations to all the regular Leaf switches/VTEPs

## Step 14: VRF Extension attachments for MSD

# Select the MSD and go to the VRF section

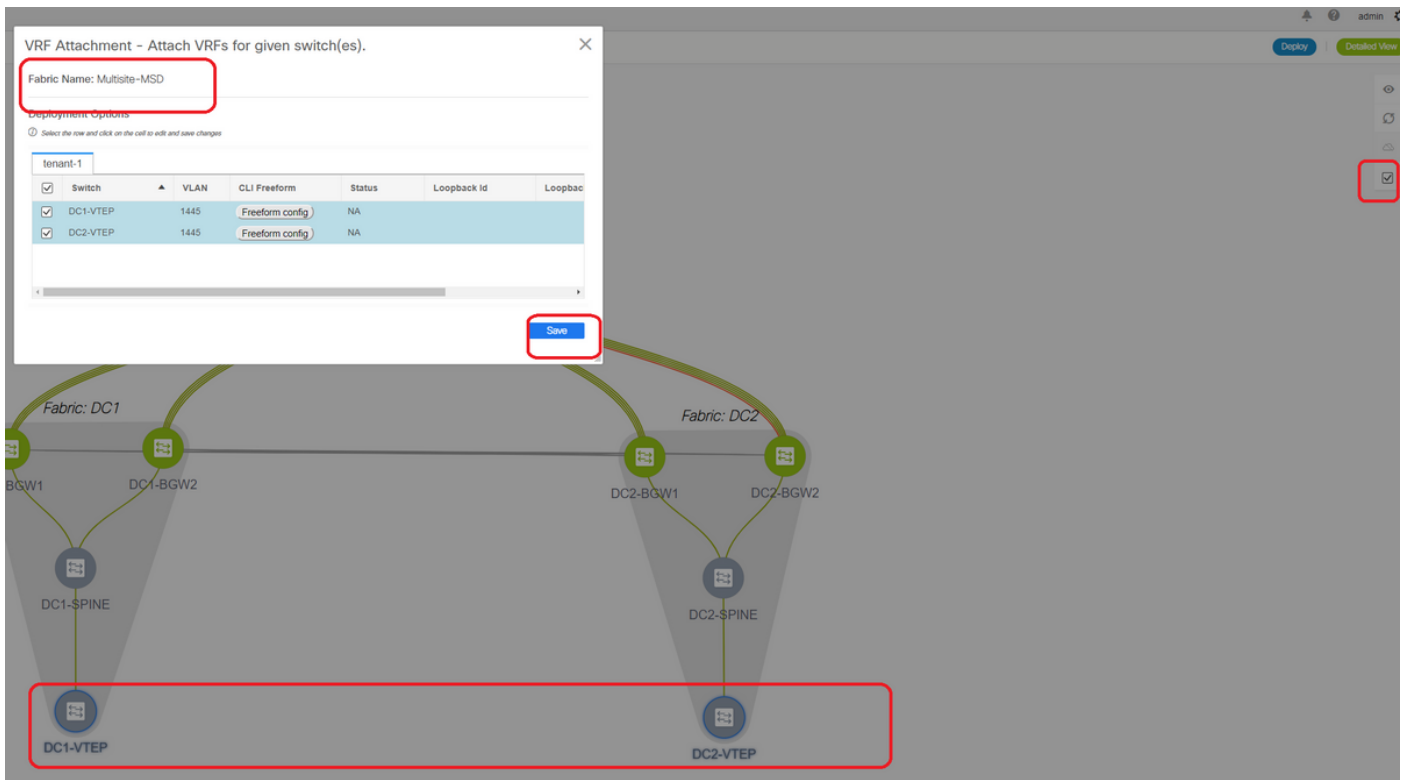


# Note that the Extend option has to be "MULTISITE+VRF\_LITE" as in this document, border Gateway functionality and the VRFLITE are integrated onto the Border Gateway switches.

# AUTO\_VRF\_LITE will be set to true

# PEER VRF NAME will have to be populated manually for all 8 as shown below from BGWs to DCI Switches(here, the example uses the same VRF NAME on DCI Switches)

# Once done, Click "save"



# While creating VRF Extensions, only the Border Gateways will have extra configurations towards the VRFLITE DCI switches

# Hence the regular leaf will have to be selected separately and then click on the "checkboxes" for each Tenant VRFs as shown above.

# Click on Deploy to push the configurations

## Step 15: Pushing Network configurations to the Fabric from MSD



# Select the relevant Networks within MSD fabric

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

Fabric Name: Multisite-MSD

Deployment Options

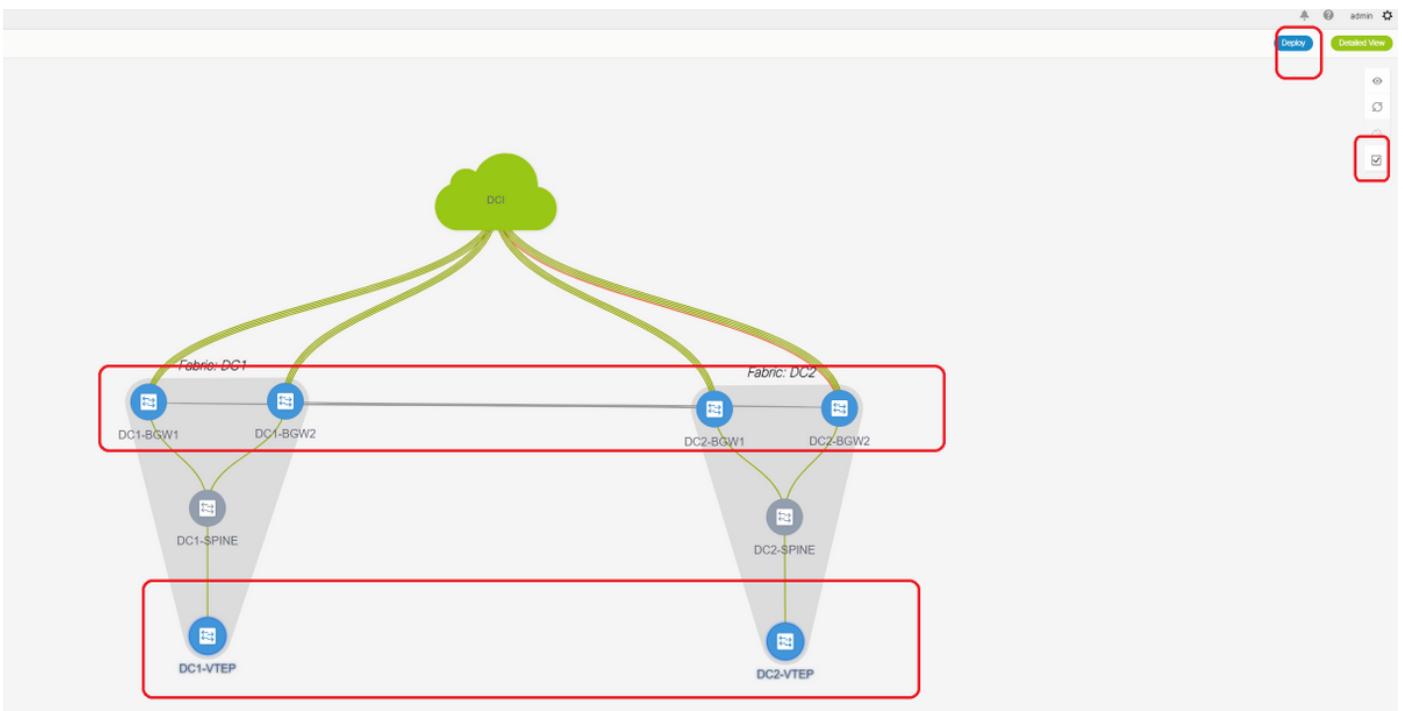
Select the row and click on the cell to add/edit configuration

Switch	VLAN	Extend	Interfaces	CLI Freeform	Status
<input checked="" type="checkbox"/> DC1-BGW1	144	MULTISITE	Applicable to BGW Leaf - VPC only	Freeform config	PENDING
<input checked="" type="checkbox"/> DC1-BGW2	144	MULTISITE	Applicable to BGW Leaf - VPC only	Freeform config	PENDING
<input checked="" type="checkbox"/> DC2-BGW1	144	MULTISITE	Applicable to BGW Leaf - VPC only	Freeform config	PENDING
<input checked="" type="checkbox"/> DC2-BGW2	144	MULTISITE	Applicable to BGW Leaf - VPC only	Freeform config	PENDING

Save

The diagram below shows two fabrics, Fabric: DC1 and Fabric: DC2. Fabric: DC1 includes switches DC1-BGW1, DC1-BGW2, DC1-SPINE, and DC1-VTEP. Fabric: DC2 includes switches DC2-BGW1, DC2-BGW2, DC2-SPINE, and DC2-VTEP. A red box highlights the four selected BGW switches in the diagram.

# Note that only the Border Gateways are selected at the moment; Perform the same and select the Regular Leaf switches/VTEPs-> DC1-VTEP and DC2-VTEP in this case.



# Once done, click the "deploy"(which will push configurations to all 6 switches above)

## Step 16: Verifying VRF and Networks on all VRFs

# This step is to verify if the VRF and Networks are shown as "Deployed" on all Fabrics; if its showing as pending, Make sure to "deploy" the configurations.

## Step 17: Deploying configurations on External Fabric

# This step is required so as to push all the relevant IP addressing, BGP, VRFLITE configurations to the DCI Switches.

# To do this, Select the External Fabric and click on "save & Deploy"

```
DCI-1# 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 173, IPv4 Unicast config peers 4, capable peers 4
22 network entries and 28 paths using 6000 bytes of memory
BGP attribute entries [3/504], 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	11	10	173	0	0	00:04:42	5
10.4.10.9	4	65000	11	10	173	0	0	00:04:46	5
10.4.20.37	4	65002	11	10	173	0	0	00:04:48	5
10.4.20.49	4	65002	11	10	173	0	0	00:04:44	5

```
DCI-1# sh ip bgp sum vrf tenant-1
```

```
BGP summary information for VRF tenant-1, address family IPv4 Unicast
BGP router identifier 10.33.10.2, local AS number 65001
BGP table version is 14, IPv4 Unicast config peers 4, capable peers 4
2 network entries and 8 paths using 1200 bytes of memory
BGP attribute entries [2/336], 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
10.33.10.1	4	65000	8	10	14	0	0	00:01:41	2
10.33.10.9	4	65000	10	11	14	0	0	00:03:16	2
10.33.20.1	4	65002	11	10	14	0	0	00:04:40	2
10.33.20.9	4	65002	11	10	14	0	0	00:04:39	2

```
DCI-2# 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 160, IPv4 Unicast config peers 4, capable peers 4
22 network entries and 28 paths using 6000 bytes of memory
BGP attribute entries [3/504], 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
10.4.10.5	4	65000	12	11	160	0	0	00:05:10	5
10.4.10.13	4	65000	12	11	160	0	0	00:05:11	5
10.4.20.45	4	65002	12	11	160	0	0	00:05:10	5
10.4.20.53	4	65002	12	11	160	0	0	00:05:07	5

```
DCI-2# sh ip bgp sum vrf tenant-1
```

```
BGP summary information for VRF tenant-1, address family IPv4 Unicast
BGP router identifier 10.33.10.6, local AS number 65001
BGP table version is 14, IPv4 Unicast config peers 4, capable peers 4
2 network entries and 8 paths using 1200 bytes of memory
BGP attribute entries [2/336], 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
10.33.10.5	4	65000	10	11	14	0	0	00:03:28	2
10.33.10.13	4	65000	11	11	14	0	0	00:04:30	2
10.33.20.5	4	65002	12	11	14	0	0	00:05:05	2
10.33.20.13	4	65002	12	11	14	0	0	00:05:03	2

# Once deployed, we will see 4 IPv4 BGP neighborships from Each DCI Switch to all BGWs and 4 IPv4 VRF BGP neighborships as well(which is for the tenant VRF EXTension)

## Step 18: Configuring iBGP Between DCI switches

# Considering that DCI switches are having links connected between each other, an iBGP IPv4 neighborship is ideal so that if any downstream connections go down on DCI-1 switch, the North to South traffic can still be forwarded via DCI-2

# For this, an iBGP IPv4 Neighborship is required between DCI switches and use next-hop-self as well on each side.

# A Freeform will have to be spun up on DCI switches to achieve this. The required lines of configurations are as below.

# DCI switches in the above topology are configured in vPC; so, the backup SVI can be used to build the iBGP Neighborships

# Select the DCI fabric and right click each switch and "view/edit policies"

The screenshot displays the 'View/Edit Policies for DCI-1(FDO22141QDG)' window. A table lists policies, with 'POLICY-477530' selected. The 'Edit Policy' dialog is open, showing the following details:

- Policy ID: POLICY-477530
- Template: switch\_freeform
- Priority (1-1000): 500
- Entity Type: SWITCH
- Entity Name: SWITCH
- Description: iBGP

The 'Switch Freeform Config' field contains the following configuration:

```
router bgp 65001
neighbor 10.10.8.2 remote-as 65001
address-family ipv4 unicast
next-hop-self
```

The 'Save' button is highlighted at the bottom of the dialog.

# Do the same change on DCI-2 switch and then "save&Deploy" to push the actual configurations to the DCI switches

# Once done, CLI verification can be done using the below command.

```
DCI-2# 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 187, IPv4 Unicast config peers 5, capable peers 5
24 network entries and 46 paths using 8400 bytes of memory
BGP attribute entries [6/1008], 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
10.4.10.5     4  65000   1206   1204    187   0    0  19:59:17  5
10.4.10.13    4  65000   1206   1204    187   0    0  19:59:19  5
10.4.20.45    4  65002   1206   1204    187   0    0  19:59:17  5
10.4.20.53    4  65002   1206   1204    187   0    0  19:59:14  5
10.10.8.1     4  65001     12     7     187   0    0  00:00:12  18 # iBGP neighborhood
from DCI-2 to DCI-1
```

## Step 19: Verification of IGP/BGP neighborships

### OSPF neighborships

# As all the Underlay IGP is OSPF in this example, All VTEPs will form OSPF neighborship with the spines and this includes the BGW switches in one site as well.

```
DC1-SPINE# show ip ospf neighbors
OSPF Process ID UNDERLAY VRF default
Total number of neighbors: 3
Neighbor ID      Pri State           Up Time  Address           Interface
10.10.10.3       1 FULL/ -         1d01h   10.10.10.3       Eth1/1 # DC1-Spine to DC1-
VTEP 10.10.10.2 1 FULL/ -         1d01h   10.10.10.2       Eth1/2 # DC1-Spine to DC1-BGW2 10.10.10.1 1 FULL/ -
1d01h 10.10.10.1 Eth1/3 # DC1-Spine to DC1-BGW1
```

# All loopbacks(BGP Router IDs, NVE loopbacks) are advertised in OSPF; Hence within a fabric, all Loopbacks are learnt via OSPF routing protocol which would help in further forming the l2vpn evpn neighborship

### BGP neighborships

# Within a fabric, This topology will have l2vpn evpn neighborships from Spines to the Regular VTEPs and also to Border Gateways.

```
DC1-SPINE# show bgp l2vpn evpn sum
BGP summary information for VRF default, address family L2VPN EVPN
BGP router identifier 10.10.10.4, local AS number 65000
BGP table version is 80, L2VPN EVPN config peers 3, capable peers 3
22 network entries and 22 paths using 5280 bytes of memory
BGP attribute entries [14/2352], BGP AS path entries [1/6]
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 1584 1560
80 0 0 1d01h 10 # DC1-Spine to DC1-BGW1 10.10.10.2 4 65000 1565 1555 80 0 0 1d01h 10 # DC1-Spine
to DC1-BGW2 10.10.10.3 4 65000 1550 1554 80 0 0 1d01h 2 # DC1-Spine to DC1-VTEP
```

# Considering that this is a multisite Deployment with Border Gateways peering from one site to other using eBGP l2vpn evpn, the same can be verified using below command on a Border Gateway switch.



```
DC1-BGW1# show bgp l2vpn evpn sum
BGP summary information for VRF default, address family L2VPN EVPN
BGP router identifier 10.10.10.1, local AS number 65000
BGP table version is 156, L2VPN EVPN config peers 3, capable peers 3
45 network entries and 60 paths using 9480 bytes of memory
BGP attribute entries [47/7896], BGP AS path entries [1/6]
BGP community entries [0/0], BGP clusterlist entries [2/8]
```

```
Neighbor      V    AS MsgRcvd MsgSent  TblVer  InQ OutQ Up/Down  State/PfxRcd
10.10.10.4 4 65000 1634 1560 156 0 0 1d01h 8 # DC1-BGW1 to DC1-SPINE 10.10.20.3 4 65002 1258
1218 156 0 0 20:08:03 9 # DC1-BGW1 to DC2-BGW1 10.10.20.4 4 65002 1258 1217 156 0 0 20:07:29 9 #
DC1-BGW1 to DC2-BGW2 Neighbor T AS PfxRcd Type-2 Type-3 Type-4 Type-5 10.10.10.4 I 65000 8 2 0 1
5 10.10.20.3 E 65002 9 4 2 0 3 10.10.20.4 E 65002 9 4 2 0 3
```

## BGP MVPN Neighborships for TRM

# With TRM Configurations in place, all the Leaf switches(including BGWs) Will form mvpn neighborhood with the spines

```
DC1-SPINE# show bgp ipv4 mvpn summary
BGP summary information for VRF default, address family IPv4 MVPN
BGP router identifier 10.10.10.4, local AS number 65000
BGP table version is 20, IPv4 MVPN config peers 3, capable peers 3
0 network entries and 0 paths using 0 bytes of memory
BGP attribute entries [0/0], 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   2596   2572     20    0  0   1d18h  0
10.10.10.2    4 65000   2577   2567     20    0  0   1d18h  0
10.10.10.3    4 65000   2562   2566     20    0  0   1d18h  0
```

# Also, the Border Gateways are required to form the mvpn neighborhood between each other so that the east/west multicast traffic will traverse correctly.

```
DC1-BGW1# show bgp ipv4 mvpn summary
BGP summary information for VRF default, address family IPv4 MVPN
BGP router identifier 10.10.10.1, local AS number 65000
BGP table version is 6, IPv4 MVPN config peers 3, capable peers 3
0 network entries and 0 paths using 0 bytes of memory
BGP attribute entries [0/0], BGP AS path entries [0/0]
BGP community entries [0/0], BGP clusterlist entries [2/8]
```

```
Neighbor      V    AS MsgRcvd MsgSent  TblVer  InQ OutQ Up/Down  State/PfxRcd
10.10.10.4    4 65000   2645   2571     6     0  0   1d18h  0
10.10.20.3    4 65002   2273   2233     6     0  0   1d12h  0
10.10.20.4    4 65002   2273   2232     6     0  0   1d12h  0
```

## Step 20: Tenant VRF Loopback Creation on Border Gateway switches

# Create Loopbacks in tenant VRF with unique IP addresses on All Border Gateways.

# For this purpose, Select DC1, right click on DC1-BGW1, Manage interfaces and then create loopback as shown below.

Add Interface ✕

\* Type: Loopback

\* Select a device: DC1-BGW1

\* Loopback ID: 2

\* Policy: int\_loopback\_11\_1

---

General

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

Loopback IP: 172.19.10.1 Loopback IP address for V4 underlay

Loopback IPv6 Address: Loopback IPv6 address for V6 underlay

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 should strictly match 'show run' output, with respect to case and newlines. Any mismatches will yield unexpected diffs during deploy.

Enable Interface  Uncheck to disable the interface

Save Preview Deploy

# Same step will have to be done on other 3 Border Gateways.

## Step 21: VRFLITE configurations on DCI switches

# In this topology, the DCI Switches are configured with VRFLITE towards the BGWs. VRFLITE is also configured towards the North Of DCI Switches(ie to the Core switches)

# For TRM Purposes, the PIM RP within the VRF tenant-1 is located in the Core Switch which is Connected via VRFLITE to the DCI switches

# This topology has IPv4 BGP neighborship from DCI switches to the Core Switch within VRF tenant-1 that is at the top of the diagram.

# For this purpose, Sub-interfaces are created and assigned with IP addresses and BGP neighborships are established as well(These are Done by CLI directly on the DCI and Core Switches)

```
DCI-1# sh ip bgp sum vrf tenant-1
BGP summary information for VRF tenant-1, address family IPv4 Unicast
BGP router identifier 10.33.10.2, local AS number 65001
BGP table version is 17, IPv4 Unicast config peers 5, capable peers 5
4 network entries and 10 paths using 1680 bytes of memory
BGP attribute entries [3/504], 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
```

```

10.33.10.1      4 65000    6366    6368      17    0    0    4d10h 2
10.33.10.9      4 65000    6368    6369      17    0    0    4d10h 2
10.33.20.1      4 65002    6369    6368      17    0    0    4d10h 2
10.33.20.9      4 65002    6369    6368      17    0    0    4d10h 2
172.16.111.2 4 65100 68 67 17 0 0 00:49:49 2 # This is towards the Core switch from DCI-1
# Above in red is the BGP neighbor towards the Core switch from DCI-1.

```

```

DCI-2# sh ip bgp sum vr tenant-1
BGP summary information for VRF tenant-1, address family IPv4 Unicast
BGP router identifier 10.33.10.6, local AS number 65001
BGP table version is 17, IPv4 Unicast config peers 5, capable peers 5
4 network entries and 10 paths using 1680 bytes of memory
BGP attribute entries [3/504], 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
10.33.10.5    4 65000    6368    6369     17    0    0    4d10h 2
10.33.10.13   4 65000    6369    6369     17    0    0    4d10h 2
10.33.20.5    4 65002    6370    6369     17    0    0    4d10h 2
10.33.20.13   4 65002    6370    6369     17    0    0    4d10h 2
172.16.222.2 4 65100 53 52 17 0 0 00:46:12 2 # This is towards the Core switch from DCI-2
# Respective BGP configurations are required on the Core switch as well(back to the DCI-1 and DCI-2)

```

## Unicast Verifications

### East/West from DC1-Host1 to DC2-Host1

# With all the above configurations pushed from DCNM and manual CLI(Steps 1 through 21), the unicast reachability should be working East/West

```

DC1-Host1# ping 172.16.144.2 source 172.16.144.1
PING 172.16.144.2 (172.16.144.2) from 172.16.144.1: 56 data bytes
64 bytes from 172.16.144.2: icmp_seq=0 ttl=254 time=0.858 ms
64 bytes from 172.16.144.2: icmp_seq=1 ttl=254 time=0.456 ms
64 bytes from 172.16.144.2: icmp_seq=2 ttl=254 time=0.431 ms
64 bytes from 172.16.144.2: icmp_seq=3 ttl=254 time=0.454 ms
64 bytes from 172.16.144.2: icmp_seq=4 ttl=254 time=0.446 ms

--- 172.16.144.2 ping statistics ---
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min/avg/max = 0.431/0.529/0.858 ms

```

### North/South from DC1-Host1 to PIM RP(10.200.200.100)

```

DC1-Host1# ping 10.200.200.100 source 172.16.144.1
PING 10.200.200.100 (10.200.200.100) from 172.16.144.1: 56 data bytes
64 bytes from 10.200.200.100: icmp_seq=0 ttl=250 time=0.879 ms
64 bytes from 10.200.200.100: icmp_seq=1 ttl=250 time=0.481 ms
64 bytes from 10.200.200.100: icmp_seq=2 ttl=250 time=0.483 ms
64 bytes from 10.200.200.100: icmp_seq=3 ttl=250 time=0.464 ms
64 bytes from 10.200.200.100: icmp_seq=4 ttl=250 time=0.485 ms

--- 10.200.200.100 ping statistics ---

```

5 packets transmitted, 5 packets received, 0.00% packet loss  
round-trip min/avg/max = 0.464/0.558/0.879 ms

## Multicast Verifications

For this document purpose, the PIM RP for the "tenant-1" VRF is configured and present external to the VXLAN Fabric; Per the topology, the PIM RP is configured on Core switch with the IP address-> 10.200.200.100

### Source in Non-vxlan(behind Core Switch), Receiver in DC2

Refer Topology which is shown at the beginning.

# North/South Multicast traffic sourced from Non-VXLAN host-> 172.17.100.100, Receiver is Present in both Datacenters; DC1-Host1-> 172.16.144.1 and DC2-Host1-> 172.16.144.2, Group -> 239.100.100.100

```
Legacy-SW#ping 239.100.100.100 source 172.17.100.100 rep 1
Type escape sequence to abort.
Sending 1, 100-byte ICMP Echos to 239.100.100.100, timeout is 2 seconds:
Packet sent with a source address of 172.17.100.100
```

```
Reply to request 0 from 172.16.144.1, 3 ms
Reply to request 0 from 172.16.144.1, 3 ms
Reply to request 0 from 172.16.144.2, 3 ms
Reply to request 0 from 172.16.144.2, 3 ms
```

### Source in DC1, Receiver in DC2 as well as external

```
DC1-Host1# ping multicast 239.144.144.144 interface vlan 144 vrf vlan144 cou 1
PING 239.144.144.144 (239.144.144.144): 56 data bytes
64 bytes from 172.16.144.2: icmp_seq=0 ttl=254 time=0.781 ms      # Receiver in DC2
64 bytes from 172.17.100.100: icmp_seq=0 ttl=249 time=2.355 ms  # External Receiver
```

```
--- 239.144.144.144 ping multicast statistics ---
1 packets transmitted,
From member 172.17.100.100: 1 packet received, 0.00% packet loss
From member 172.16.144.2: 1 packet received, 0.00% packet loss
--- in total, 2 group members responded ---
```

### Source in DC2, Receiver in DC1 as well as external

```
DC2-Host1# ping multicast 239.145.145.145 interface vlan 144 vrf vlan144 cou 1
PING 239.145.145.145 (239.145.145.145): 56 data bytes
64 bytes from 172.16.144.1: icmp_seq=0 ttl=254 time=0.821 ms      # Receiver in DC1
64 bytes from 172.17.100.100: icmp_seq=0 ttl=248 time=2.043 ms  # External Receiver
```

```
--- 239.145.145.145 ping multicast statistics ---
1 packets transmitted,
From member 172.17.100.100: 1 packet received, 0.00% packet loss
From member 172.16.144.1: 1 packet received, 0.00% packet loss
--- in total, 2 group members responded ---
```