使用DCNM構建Nexus 9000 VXLAN共用邊界多站點部署

目錄

簡介

本文檔旨在說明如何使用DCNM 11.2版共用邊界模型部署Cisco Nexus 9000 VXLAN多站點部署。

拓撲

SHARED-BORDER1
AS 65001
RID- 10.10.100.1
NVE- 192.168.100.2

RTR to
INET/external

Shared Loopback
192.168.100.1

SHARED-BORDER2
AS 65001
RID- 10.10.100.2
NVE-192.168.100.3

Anycast Multisite
Loopback-
192.168.200.1

DC1-BGW1
AS 65000
RID-10.10.10.1
NVE Loopback- 192.168.10.3

DC1-BGW2
AS 65000
RID- 10.10.10.2
NVE Loopback- 192.168.10.2

DC1-SPINE
AS 65000
RID- 10.10.10.4

DC1-VTEP
AS 65000
RID- 10.10.10.3
NVE Loopback- 192.168.10.1

DC1-Host1- 172.16.144.1- Vlan 144- 0000.0000.1441
DC1-Host2- 172.16.145.1- Vlan 145- 0000.0000.1451

Anycast Multisite
Loopback-
192.168.200.2

DC2-BGW1
AS 65002
RID- 10.10.20.3
NVE Loopback-192.168.20.3

DC2-BGW2
AS 65002
RID- 10.10.20.4
NVE Loopback-192.168.20.1

DC2-SPINE
AS 65002
RID- 10.10.20.6

DC2-VTEP
AS 65002
RID- 10.10.20.5
NVE Loopback-192.168.20.2

DC2-Host1- 172.16.144.2- Vlan 144- 0000.0000.1442
DC2-Host2- 172.16.145.2- Vlan 145- 0000.0000.1452

拓撲詳細資訊

DC1和DC2是運行vxlan的兩個資料中心位置；

DC1和DC2邊界網關具有與共用邊界的物理連線；

共用邊界具有外部連線(例如，網際網路);因此，VRF精簡連線將在共用邊界上終止，並且預設路由將由共用邊界注入到每個站點中的邊界網關

在vPC中配置共用邊框（使用DCNM部署交換矩陣時需要這樣做）

邊界網關在任播模式下配置

採用元件:

運行9.3(2)的Nexus 9ks

執行11.2版的DCNM

本文中的資訊是根據特定實驗室環境內的裝置所建立。文中使用到的所有裝置皆從已清除（預設）的組態來啟動。如果您的網路正在作用，請確保您已瞭解任何指令可能造成的影響。

高級步驟

1)考慮到本文檔基於兩個使用vxlan多站點功能的資料中心，因此必須建立兩個簡單結構

2)為共用邊界建立另一個簡單的交換矩陣

3)建立MSD並移動DC1和DC2

4)建立外部交換矩陣

5)建立多站點底層和重疊（適用於East/West）

6)在共用邊框上建立VRF擴展附件

第1步：為DC1建立簡易交換矩陣

- 登入到DCNM，從儀表板中選擇選項 — >「Fabric Builder」

- 選擇「建立交換矩陣」選項

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

- 接下來是提供交換矩陣名稱、模板,然後開啟多個頁籤,這些頁籤需要詳細資訊,如ASN、交換矩陣介面編號、任意鑄造網關MAC(AGM)

Add Fabric

* Fabric Name : DC1
* Fabric Template : Easy_Fabric_11_1 ▼

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

* 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*

#交換矩陣介面(即主幹/枝葉介面)可以是「未編號」或點對點;如果使用未編號,則所需的IP地址較少(因為IP地址是未編號環回地址)

AGM被交換矩陣中的主機用作預設網關MAC地址;所有作為預設網關的枝葉交換機上都是相同的

- 下一步是設定複製模式

Add Fabric

* Fabric Name : `DC1`

* Fabric Template : `Easy_Fabric_11_1` ▼

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

* 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*

#在此處選擇的複製模式可以是組播或IR入口複製；IR將以單播方式將vxlan vlan內的任何傳入BUM流量複製到其他VTEP（也稱為頭端複製），而多點傳送模式將傳送具有外部目的地IP位址的BUM流量（與為每個網路定義的多點傳送群組的BUM位址相同）至骨幹，且主幹會將基於外部目的地IP位址的OIL的多點傳送複製到其他VTEP

#組播組子網 — >複製BUM流量（如來自主機的ARP請求）時需要

#如果需要啟用TRM，請選中該覈取方塊並為TRM VRF提供MDT地址。

- 「vPC」的頁籤預設保留；如果備份SVI/VLAN需要任何更改，可以在此處定義這些更改
- 高級頁籤是下一部分

Add Fabric

* **Fabric Name :** DC1
* **Fabric Template :** Easy_Fabric_11_1

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

* **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** — ☐ ❓

#此處提到的站點ID自動填充到此DCNM版本中，該DCNM版本源自「常規」頁籤下定義的ASN

#填寫/修改其他相關欄位

* Resources頁籤是下一個需要環回、底層的IP編址方案的頁籤

Add Fabric

* Fabric Name : DC1

* Fabric Template : Easy_Fabric_11_1 ▼

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

Manual Underlay IP Address Allocation ☐ ❓ *Checking this will disable Dynamic Underlay IP Address Allocations*

Field	Value	Description
* 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)*

#第2層VXLAN VNI範圍 — >這些是稍後將對映到Vlan的VNID（將更低顯示它）

#第3層VXLAN VNI範圍 — >這些是第3層VNID，稍後還將對映到第3層VNI Vlan到Vn段

- **此處未顯示其他索引標籤；但是如果需要，請填滿其他頁籤；**

Add Fabric ✕

* Fabric Name : DC1

* Fabric Template : Easy_Fabric_11_1 ▼

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

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)*

[Save] [Cancel]

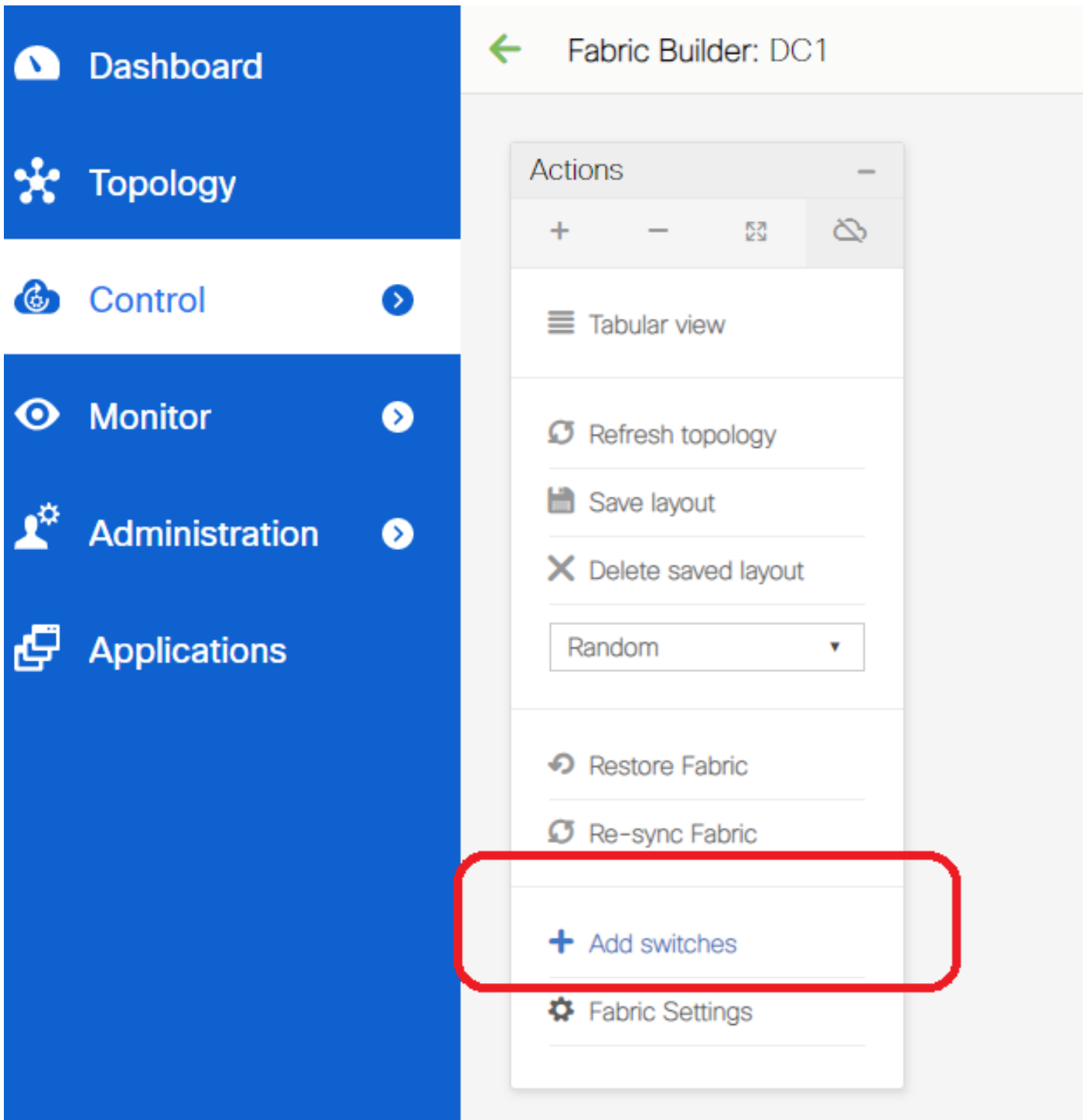- **儲存後，交換矩陣生成器頁面將顯示交換矩陣(來自DCNM-> Control-> Fabric Builder**

#此部分顯示每個結構的結構、 ASN和複製模式的完整清單

- 下一步是將交換機新增到DC1交換矩陣

第2步：將交換機新增到DC1交換矩陣

按一下上圖中的DC1，即可選擇新增交換機。

- 提供需要匯入到DC1交換矩陣的交換機的IP地址和憑證（根據本文檔開頭列出的拓撲，DC1-VTEP、DC1-SPINE、DC1-BGW1和DC1-BGW2是DC1的一部分）

#由於這是綠地部署，請注意，「preserve config」選項被選為「NO」；匯入時刪除框的所有配置，同時重新載入交換機

#選擇「開始發現」，以便DCNM開始根據「種子IP」列中提供的IP地址發現交換機

- DCNM完成交換機發現後，IP地址以及主機名將列在清單管理中

#選擇相關交換機，然後按一下「匯入交換矩陣」

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

OK   **Cancel**

Inventory Management ✕

Discover Existing Switches   PowerOn Auto Provisioning (POAP)

Discovery Information 〉 Scan Details 〉

← Back   *Note: Preserve Config selection is 'no'. Switch configuration will be erased.*   **Import into fabric**
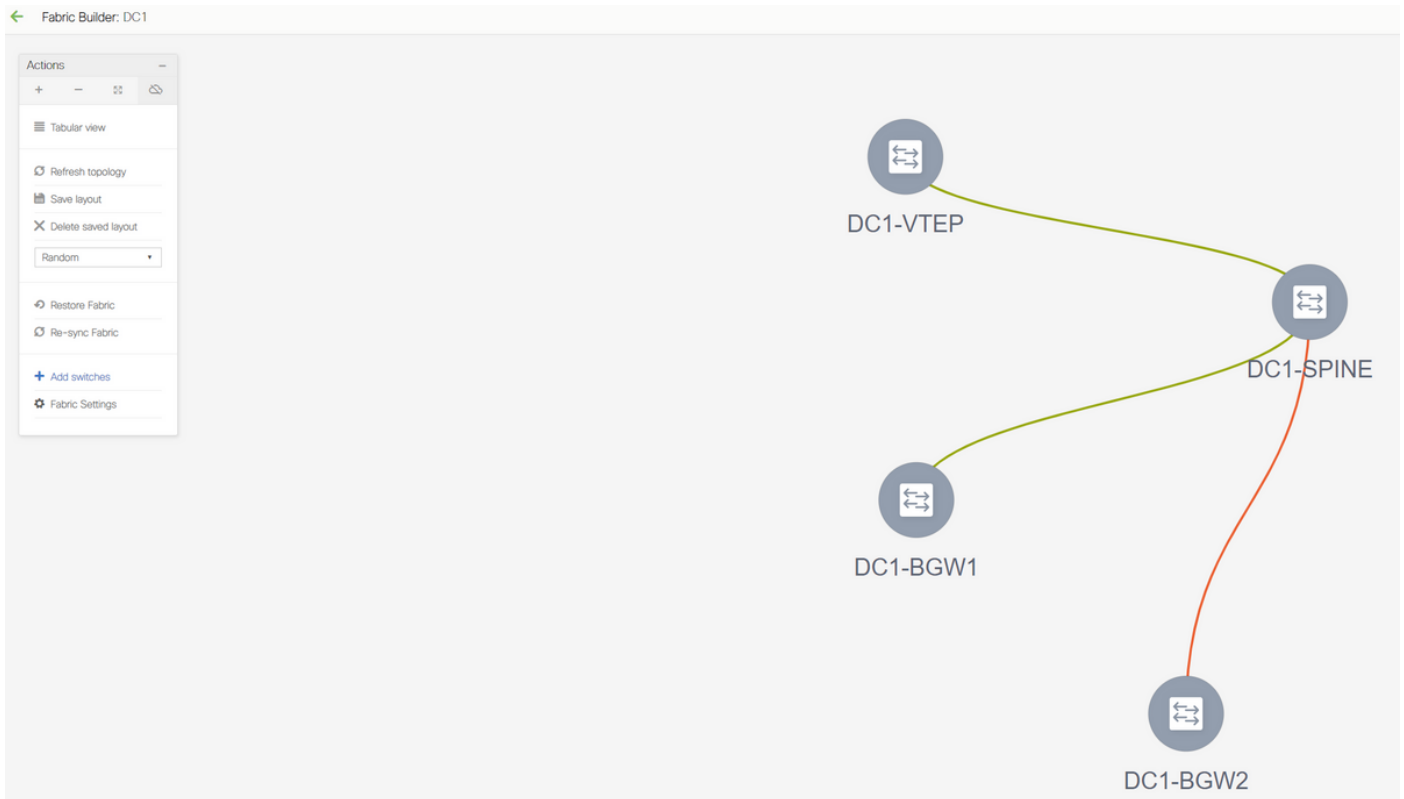
Show Quick Filter ▼ ▼

| ☐	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%

#匯入完成後，fabric builder下的拓撲可能如下所示；

#按一下一台交換機並將其與圖中的正確位置對齊，即可移動交換機
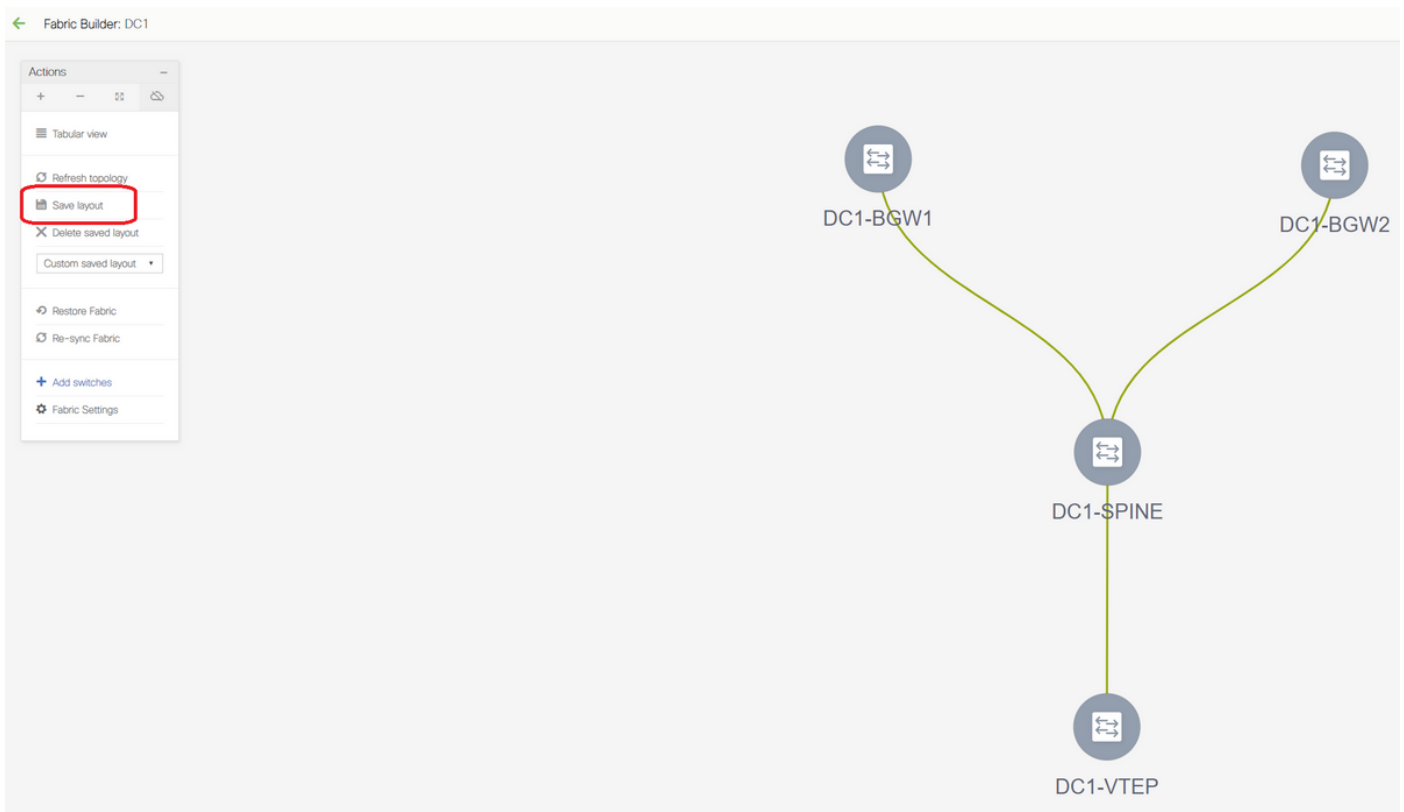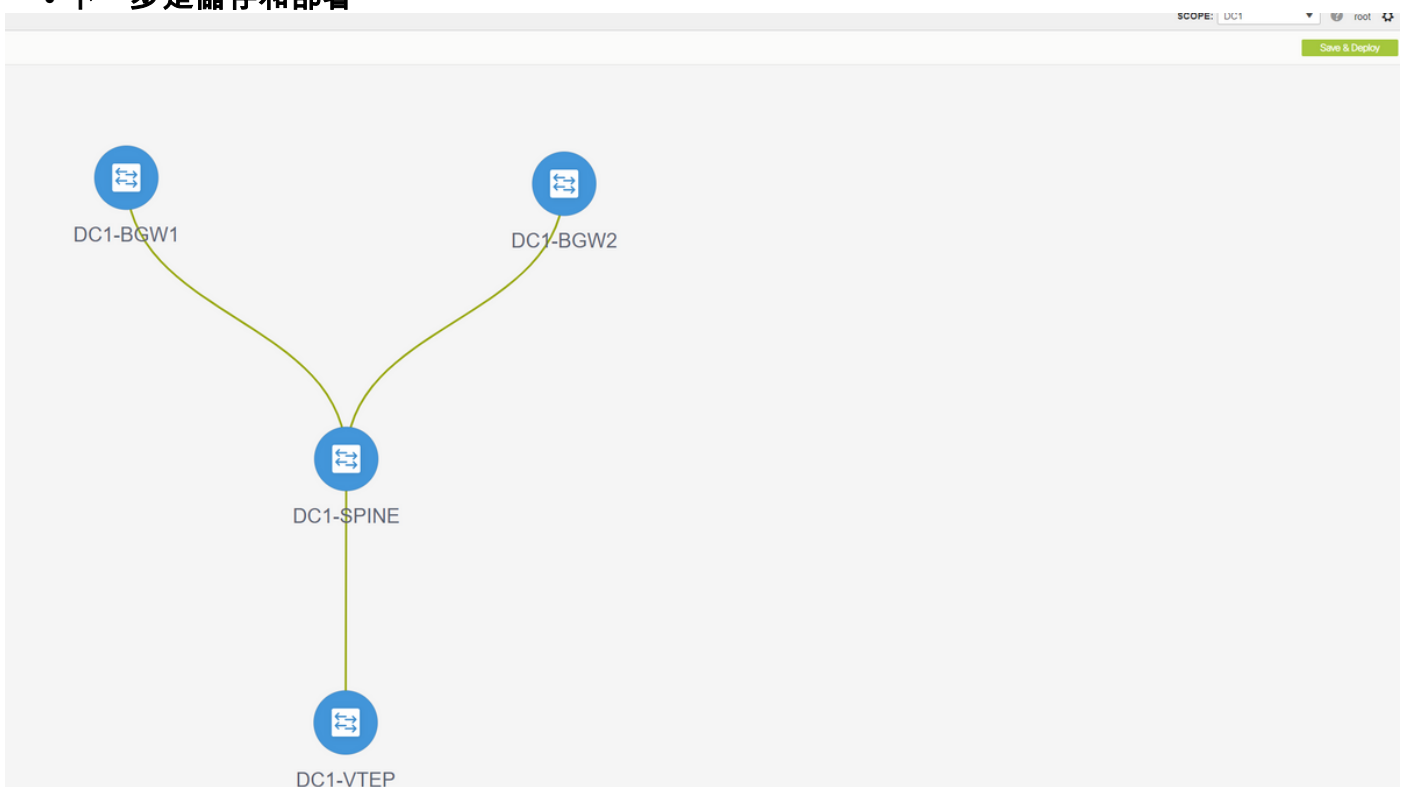

#按照需要佈局的順序重新排列交換機之後，選擇「儲存佈局」部分

- **為所有交換機設定角色**
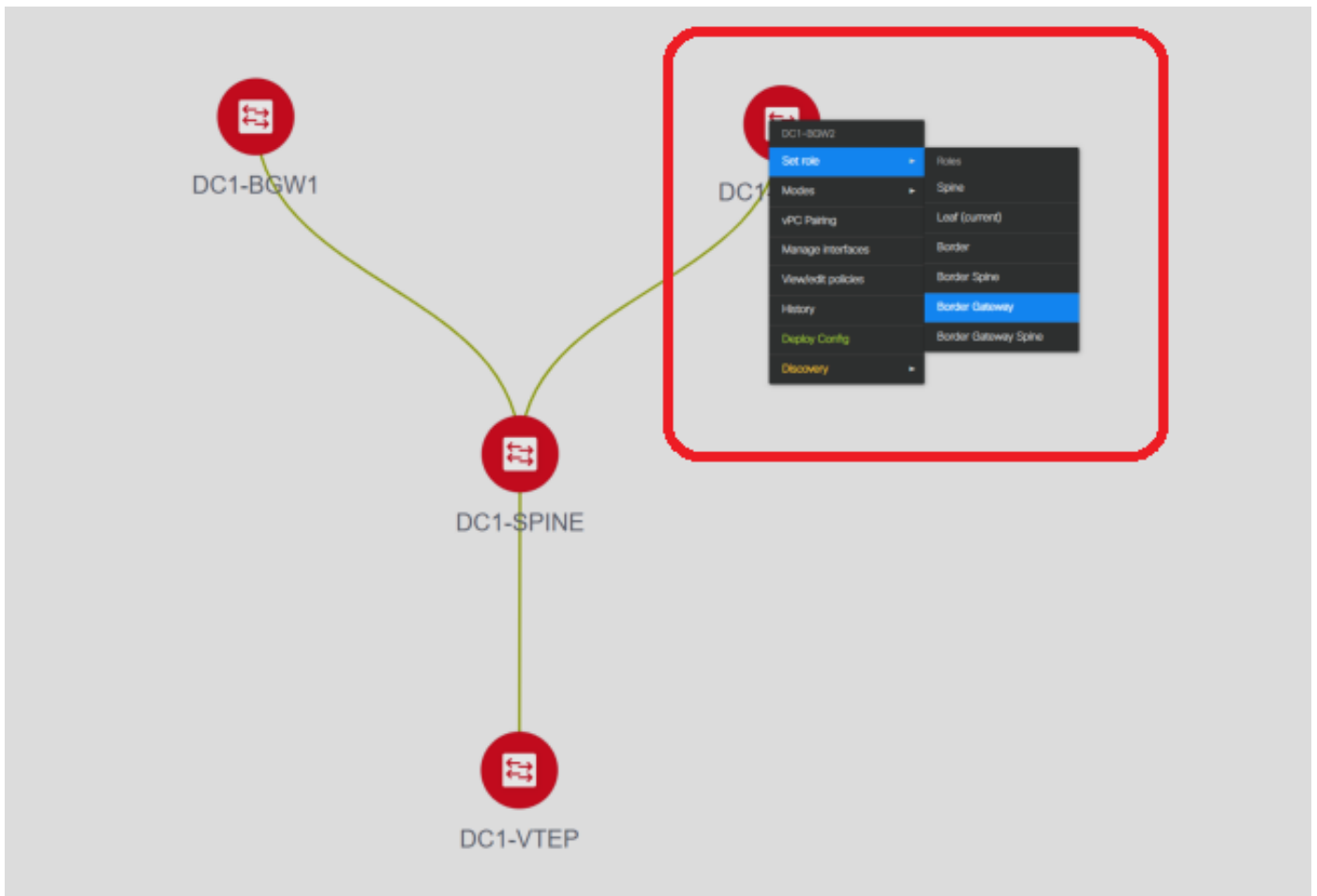
#按一下右鍵每台交換機並設定正確的角色；這裡，DC1-BGW1和DC1-BGW2是邊界網關

DC1-SPINE->將設定為role- Spine，DC1-VTEP->將設定為role-Leaf

- **下一步是儲存和部署**

DCNM現在將列出交換機，並且還將預覽DCNM將推送到所有交換機的配置。

Config Deployment

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

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%

Deploy Config

GW2

DC1-VTEP

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%

Close

GW2

DC1-VTEP

#一旦成功，狀態將反映，並且交換機將以綠色顯示

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	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%

Close
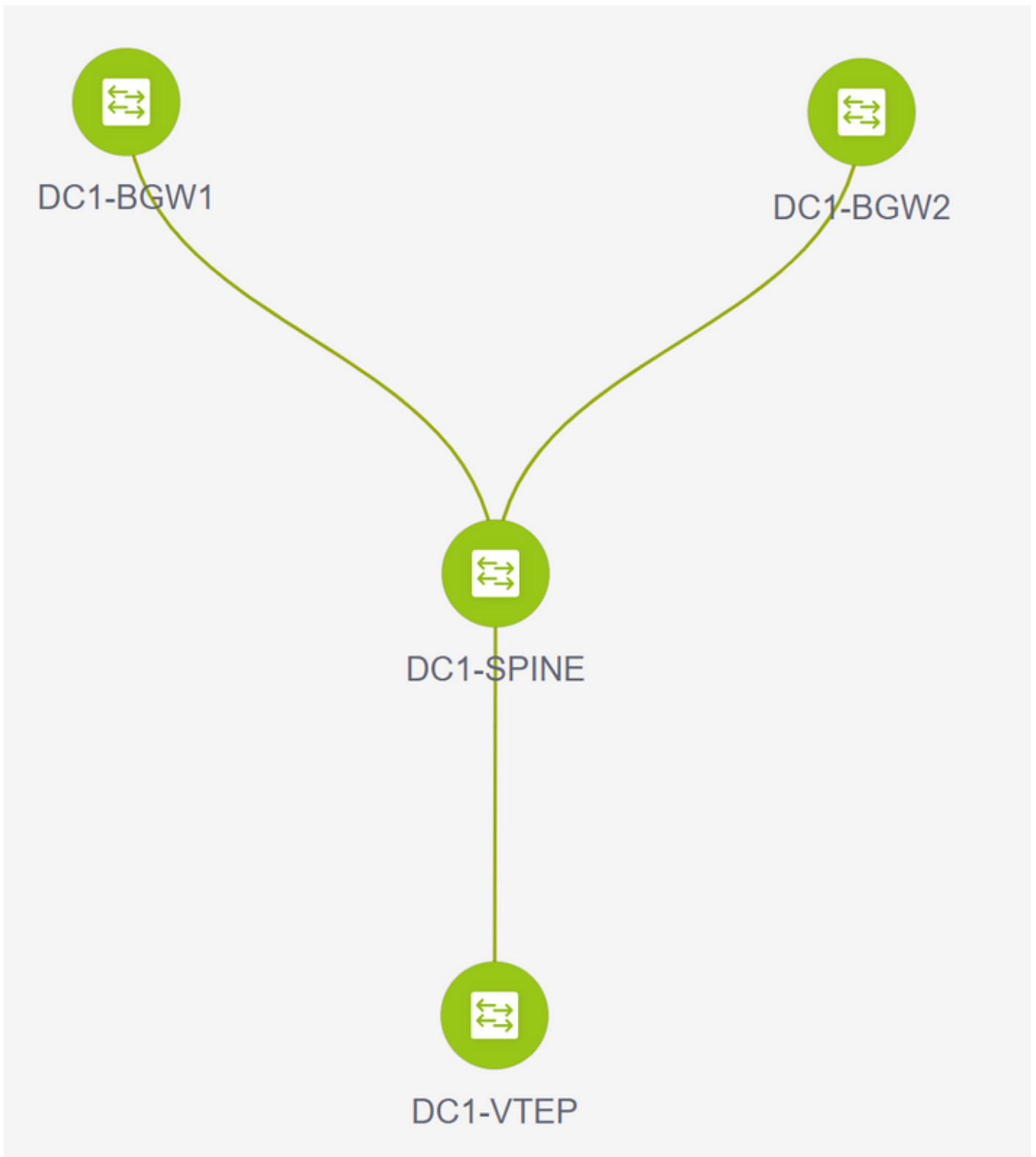
步驟3:網路/VRF的配置

- **網路/VRF的配置**

#選擇DC1 Fabric（從右上下拉選單中），Control > VRFs
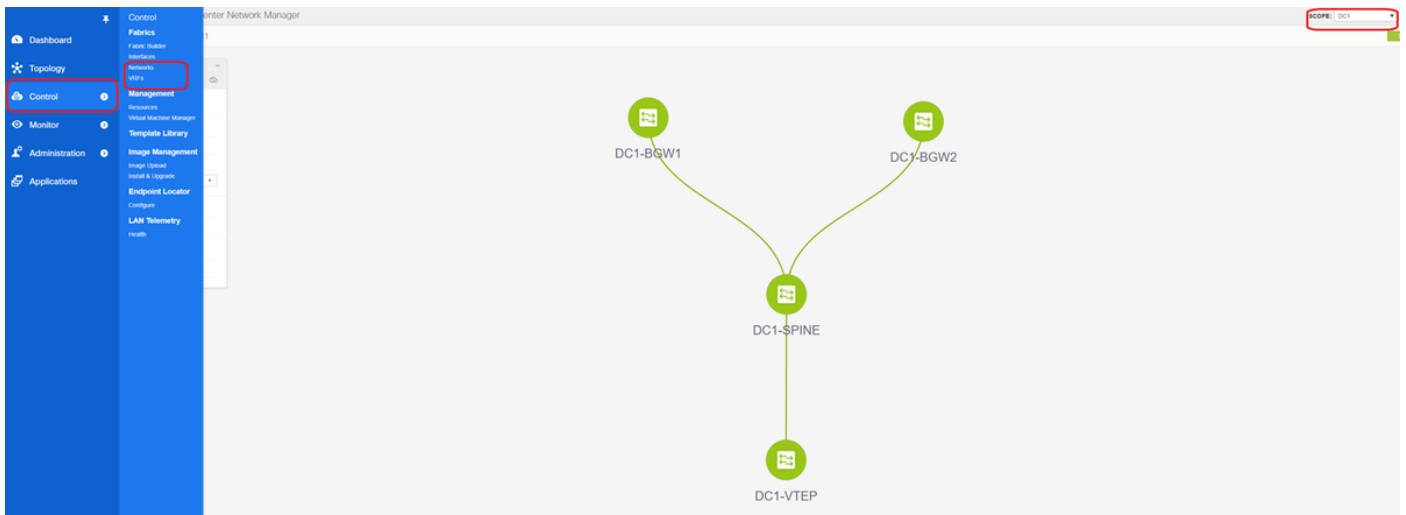
#接下來是建立VRF

11.2 DCNM版本自動填充VRF ID;如果為「Different」，請鍵入所需的名稱，然後選擇「Create VRF」

#此處使用的第3層VNID為1001445

- 下一步是建立網路

\#提供網路ID（即第2層Vlan的相應VNID）

\#提供SVI應加入的VRF;預設情況下，DCNM 11.2將VRF名稱填充為先前建立的名稱；根據需要進行更改

\# Vlan ID將是對映到此特定VNID的第2層VL

\# IPv4 Gateway->這是將在SVI上配置的任播網關IP地址，對交換矩陣內的所有VTEP都是相同的

- **Advanced Tab具有需要填充的額外行，例如；使用DHCP中繼；**

#填寫欄位後，按一下「建立網路」。

#建立需要成為此交換矩陣一部分的任何其他網路；

- 目前，VRF和網路在DCNM中剛剛定義；但並未從DCNM推送到交換矩陣中的交換機。可以使用以下命令進行驗證

#如果沒有部署到交換機，則狀態將為「不適用」。由於這是一個多站點並且涉及邊界網關，因此下面將討論網路/VRF的部署。

第4步：對DC2重複相同步驟

- 現在DC1已完全定義，也將對DC2執行相同的過程
- 完全定義DC2後，它如下所示
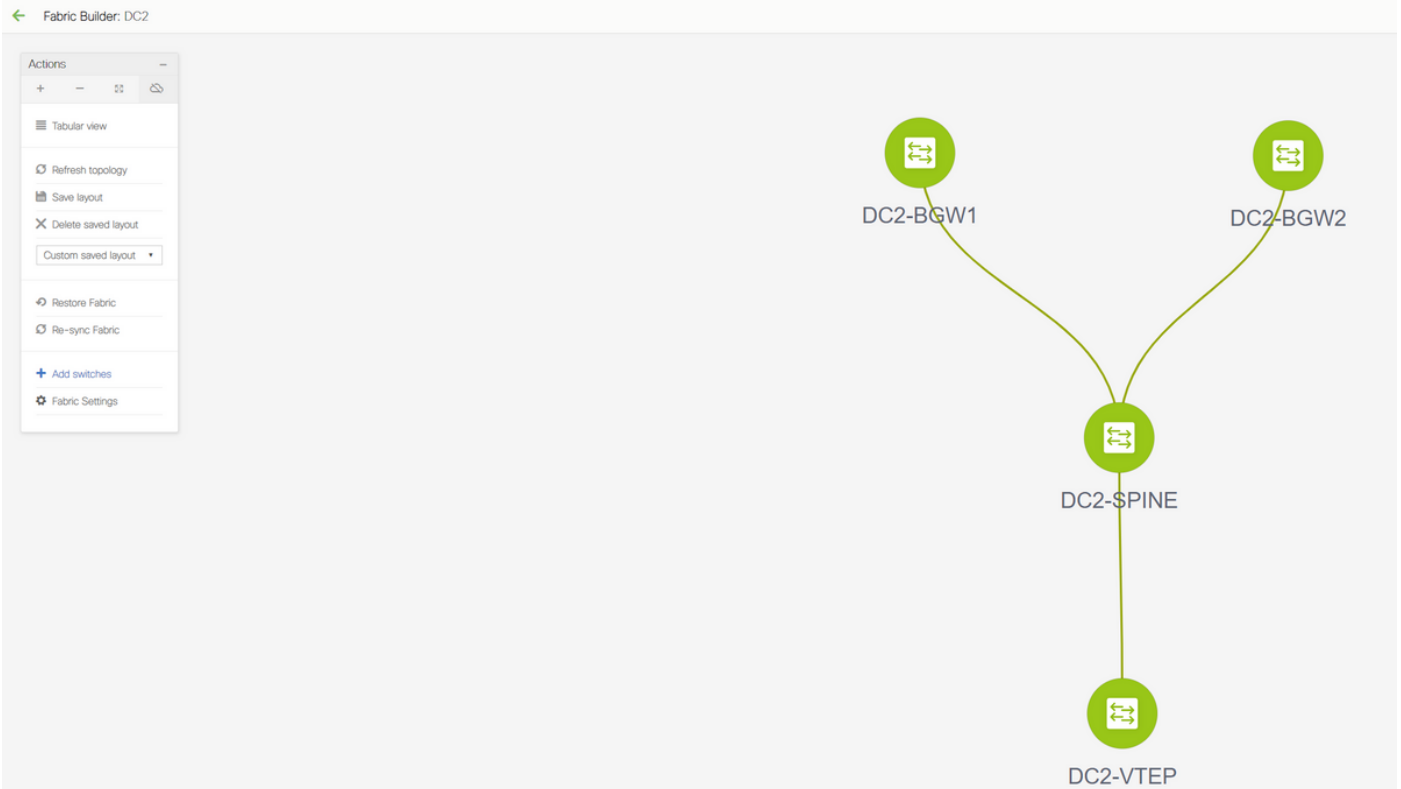
第5步：為共用邊界建立簡單的交換矩陣

- 在此建立另一個簡單交換矩陣，該交換矩陣將包括vPC中的共用邊框
- 請注意，通過DCNM部署時的共用邊界應配置為vPC，否則在DCNM上執行「重新同步」操作後，交換機間鏈路將關閉
- 共用邊界中的交換機將以「邊界」角色設定
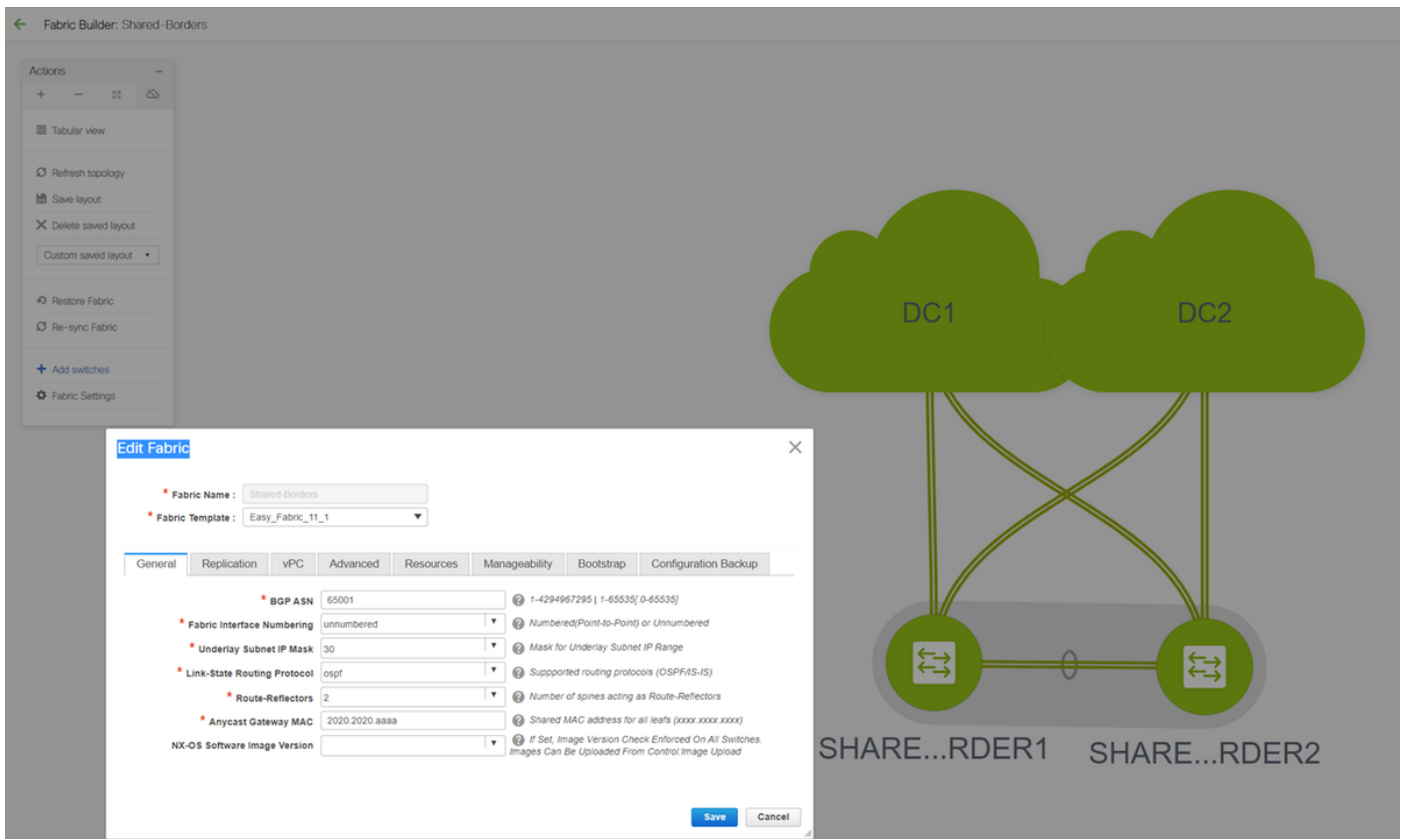
VRF數量也像以前一樣用於DC1和DC2交換矩陣

#共用邊界上不需要網路，因為共用邊界將沒有任何第2層Vlan/VNID;共用邊界不是從DC1到DC2的任何東/西流量的隧道終端；只有邊界網關在東/西DC1的VXLAN封裝/解除封裝方面起作用<>DC2流量

第6步 — 建立MSD和移動DC1和DC2交換矩陣

轉到Fabric Builder並建立新的交換矩陣並使用模板 — > MSD_Fabric_11_1
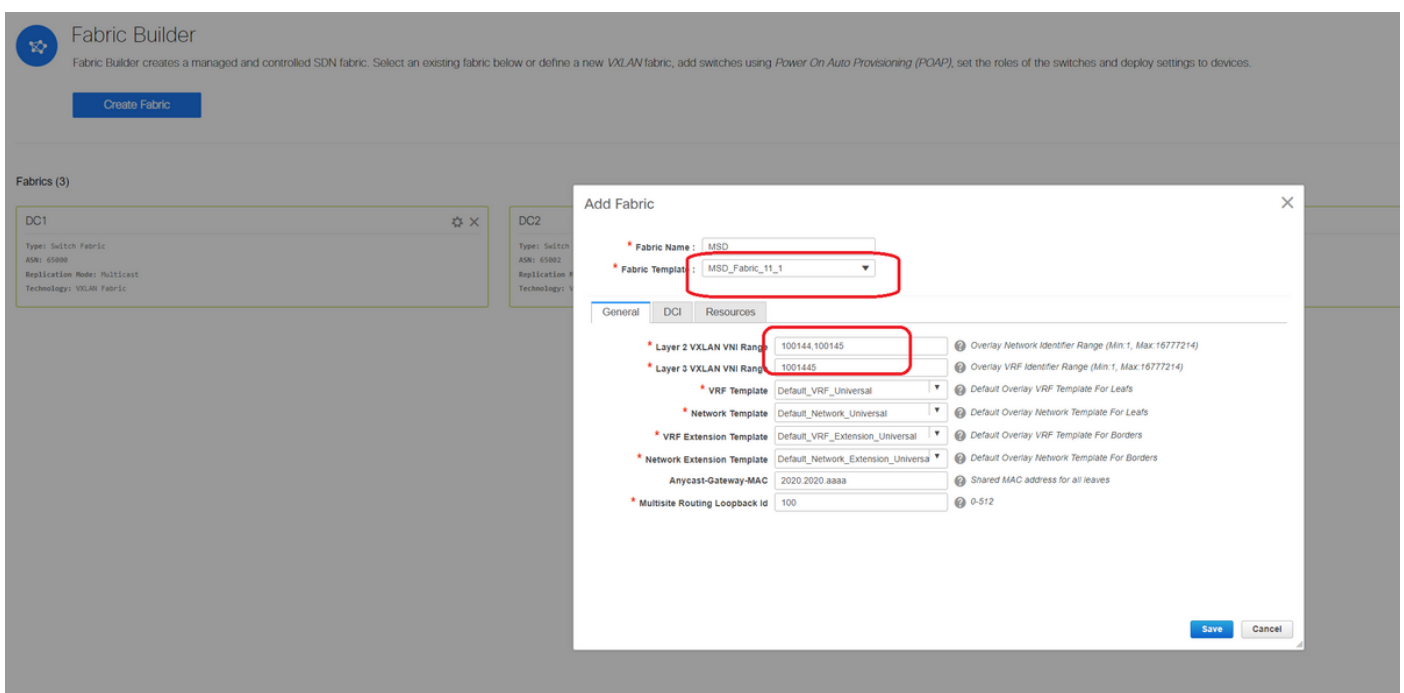
#請注意，多站點重疊IFC部署方法必須是「centralized_To_Route_Server」；在這裡，共用邊界被視為路由伺服器，因此從下拉選單使用此選項

#在「多站點**路由伺服器清單**」中；這裡，找出共用邊界上Loopback0（即路由環回）的環回IP地址，並填寫

ASN是位於共用邊界上的路由器（如需更多詳細資訊，請參閱本文檔頂部的圖表）；在本文檔中，共用邊界在同一個ASN中配置；相應地填寫

- 下一個頁籤是提供多站點環回IP範圍的位置，如下所示

#填寫所有欄位後，按一下「儲存」按鈕，將使用模板 — > MSD建立一個新的交換矩陣

#下一步是將DC1和DC2交換矩陣移至此MSD

#交換矩陣移動後，它如下所示
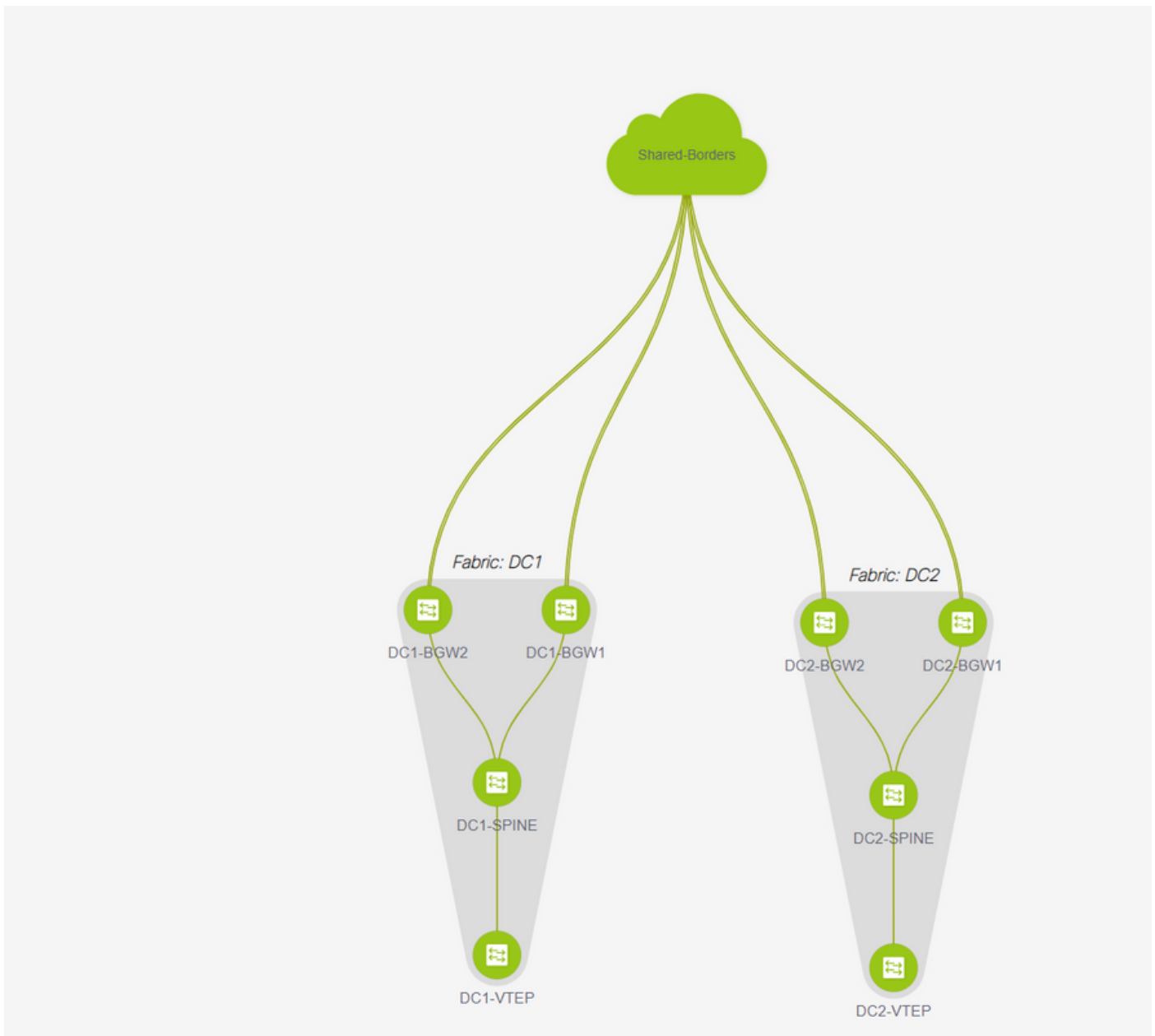
#完成後，點選「save&Deploy」按鈕，該按鈕會將所需配置推送到邊界網關，只要涉及多站點
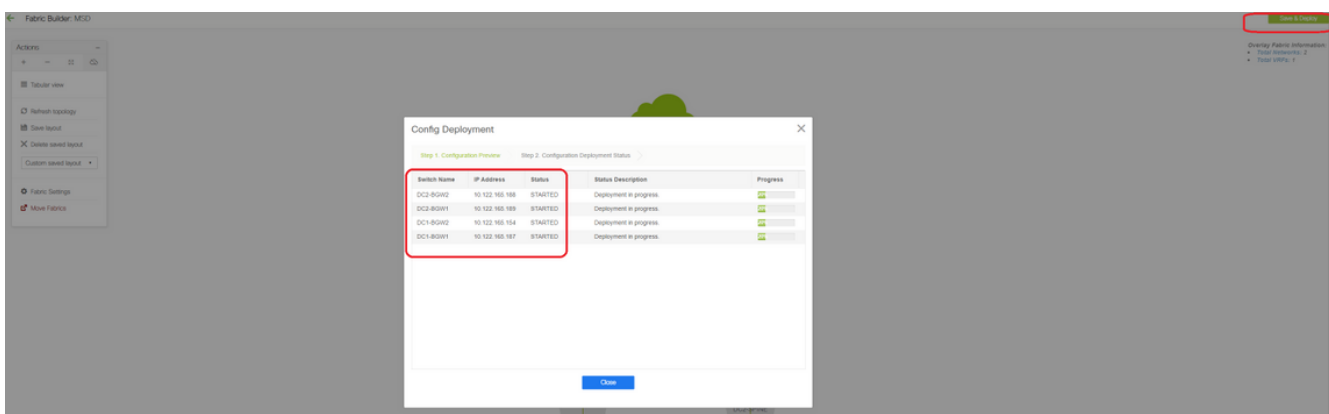
第7步：建立外部交換矩陣

#建立外部交換矩陣並將外部路由器新增到其中，如下所示；

Add Fabric

* Fabric Name : External

* Fabric Template : External_Fabric_11_1 ▼

General	Advanced	Resources	DCI	Configuration Backup	Bootstrap

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

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

#命名交換矩陣並使用模板 — > "External_Fabric_11_1";

#提供ASN

#最後，各種結構將如下所示
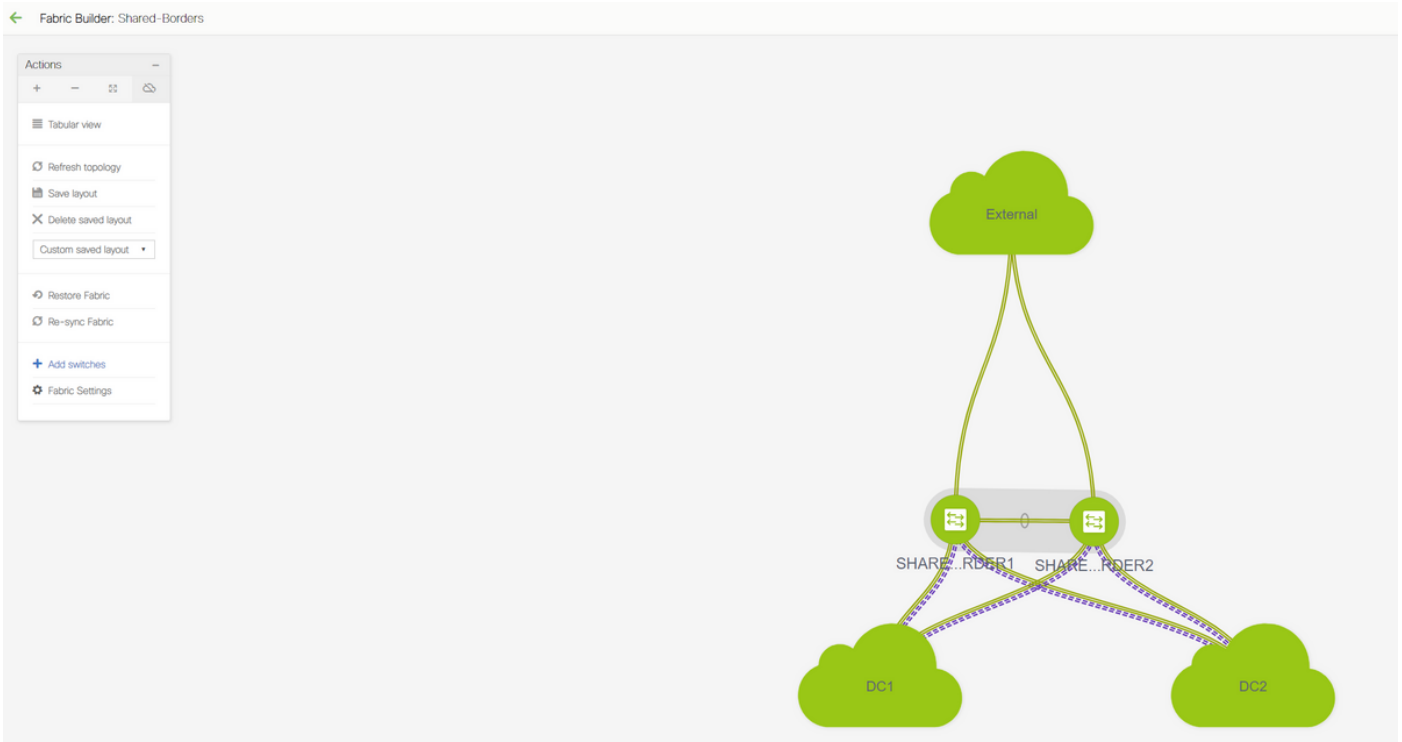
第8步：用於實現BGW之間環回可達性的eBGP底層協定（以及共用邊界之間的iBGP）

#共用邊界通過邊界網關和通往外部路由器的VRF-LITE連線運行eBGP l2vpn evpn

#使用環回形成eBGP l2vpn evpn之前，需要確保環回可以通過某種方法到達；在本示例中，我們使用從BGW到「共用」邊界的eBGP IPv4 AF，然後通告環回以進一步形成l2vpn evpn鄰居關係。
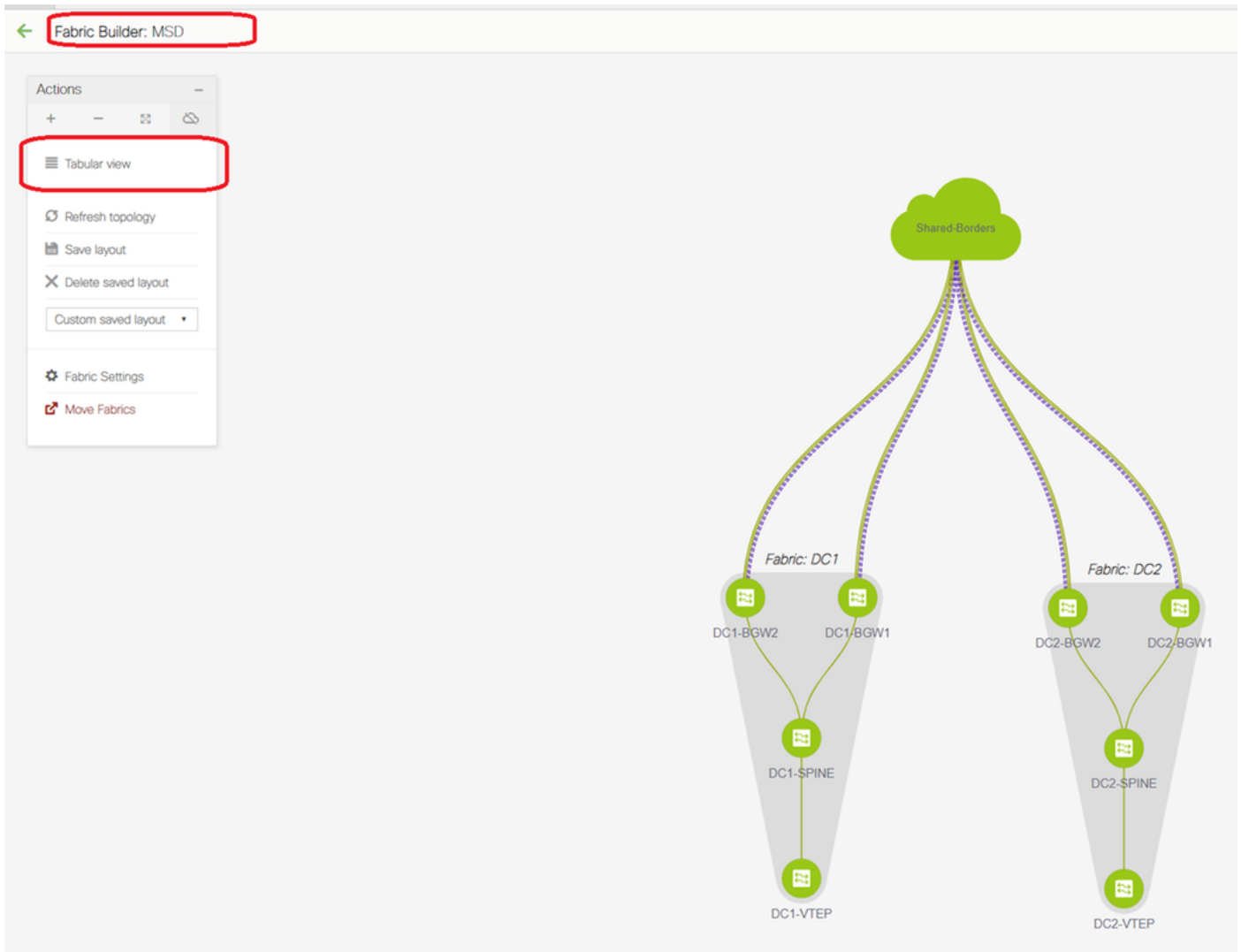
#選擇MSD交換矩陣後，切換到「表格檢視」
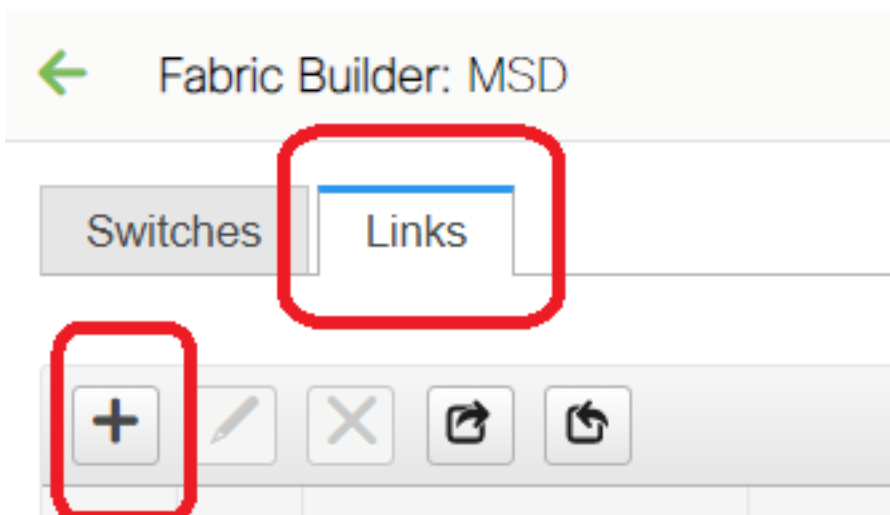
Link Management – Add Link   ✕

* 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

General
Advanced

* 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

Save

#選擇「交換矩陣間」並使用「Multisite_UNDERLAY」

#我們正在嘗試與共用邊界路由器建立IPv4 BGP鄰居關係;因此,請相應地選擇交換機和介面。

#請注意,如果CDP檢測到從DC1-BGW1到SB1的鄰居,則只需在此提供本節中的IP地址,並且執行「儲存和部署」後,將在相關介面上有效配置IP地址
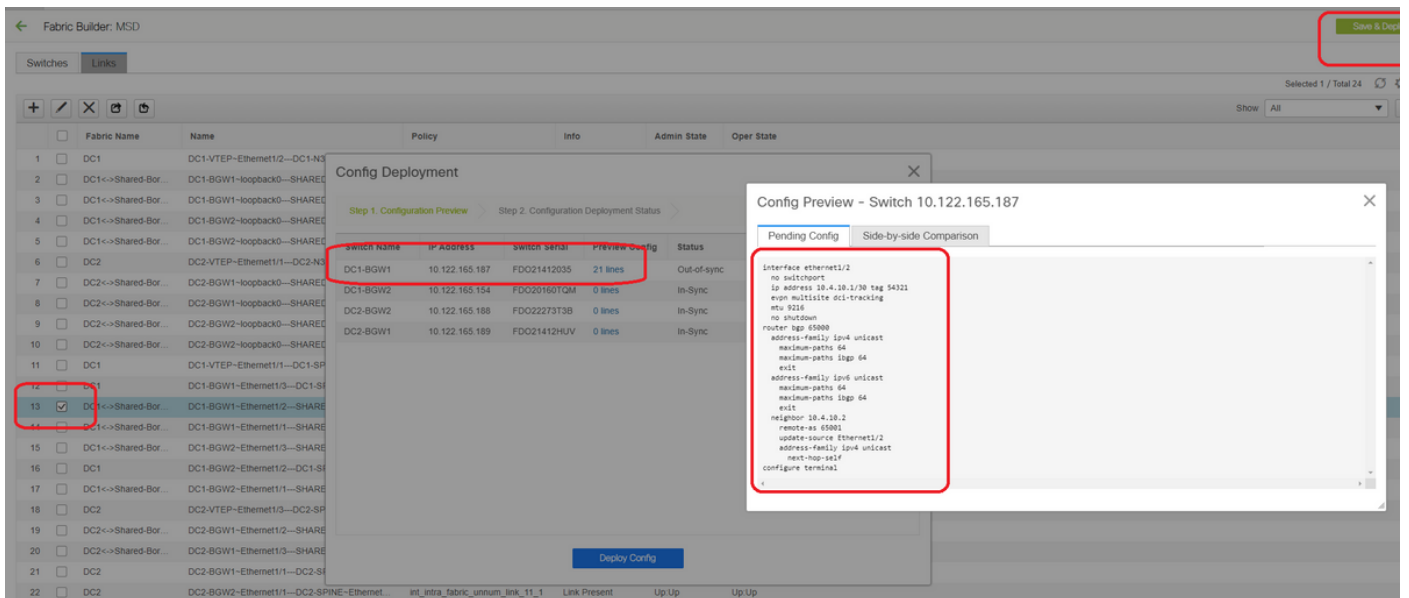
#選擇「儲存並部署」後，將為DC1-BGW1傳播所需的配置行；選擇「共用邊界」結構後，必須執行相同的步驟。
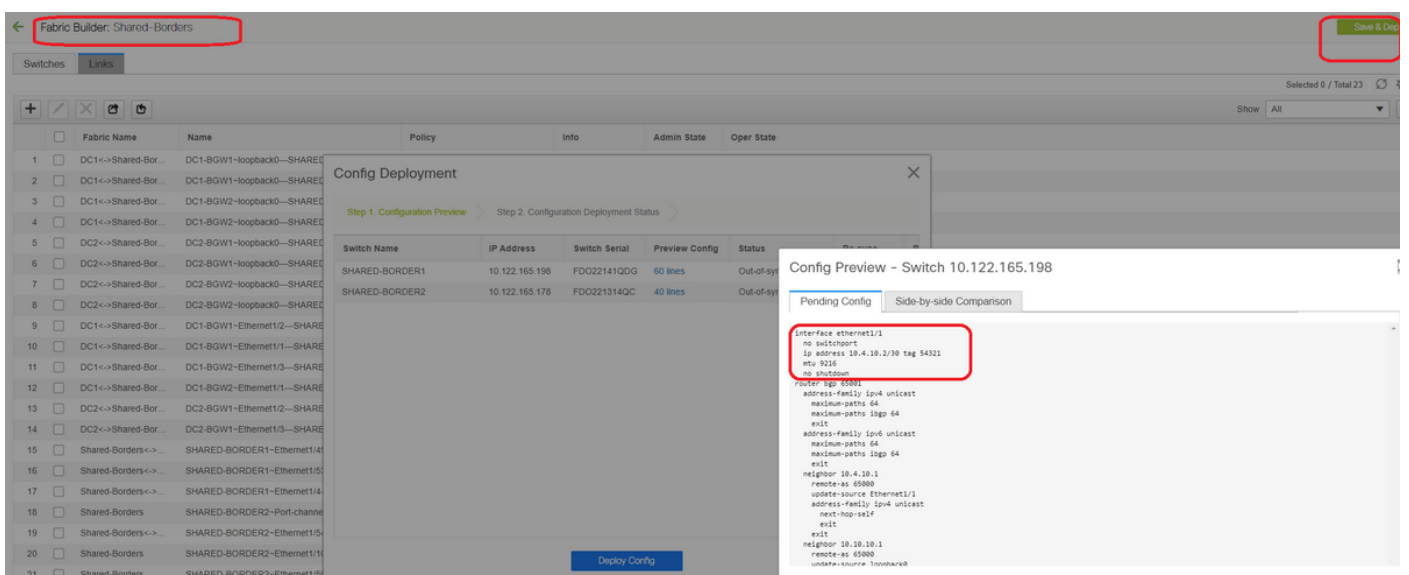
#在CLI中，可使用以下命令驗證相同內容；

```
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
```

#請注意，「save&Deploy」也必須在DC1交換矩陣上完成（選擇DC1的下拉框，然後執行同樣操作），以便相關IP編址、BGP配置傳播到DC1（邊界網關）中的交換機；

#此外，多站點底層必須從DC1-BGW、DC2-BGW建立到共用邊界；因此，必須同時執行上述相同步驟。

#最後，共用邊界將與DC1和DC2中的所有BGW建立eBGP IPv4 AF鄰居關係，如下所示；

```
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
```

#以上是建立從BGW到共用邊界的l2vpn EVPN鄰居關係的先決條件(請注意，它不是必須使用BGP;交換環回字首的任何其他機制都可以);最後，基本要求是所有BGW應可到達所有環回（共用邊界，BGW）

#請注意，需要在共用邊界之間建立iBGP IPv4 AF鄰居關係；截止到今天，DCNM沒有使用模板/下拉在共用邊界之間構建iBGP的選項；為此，必須完成自由形式配置，如下所示；
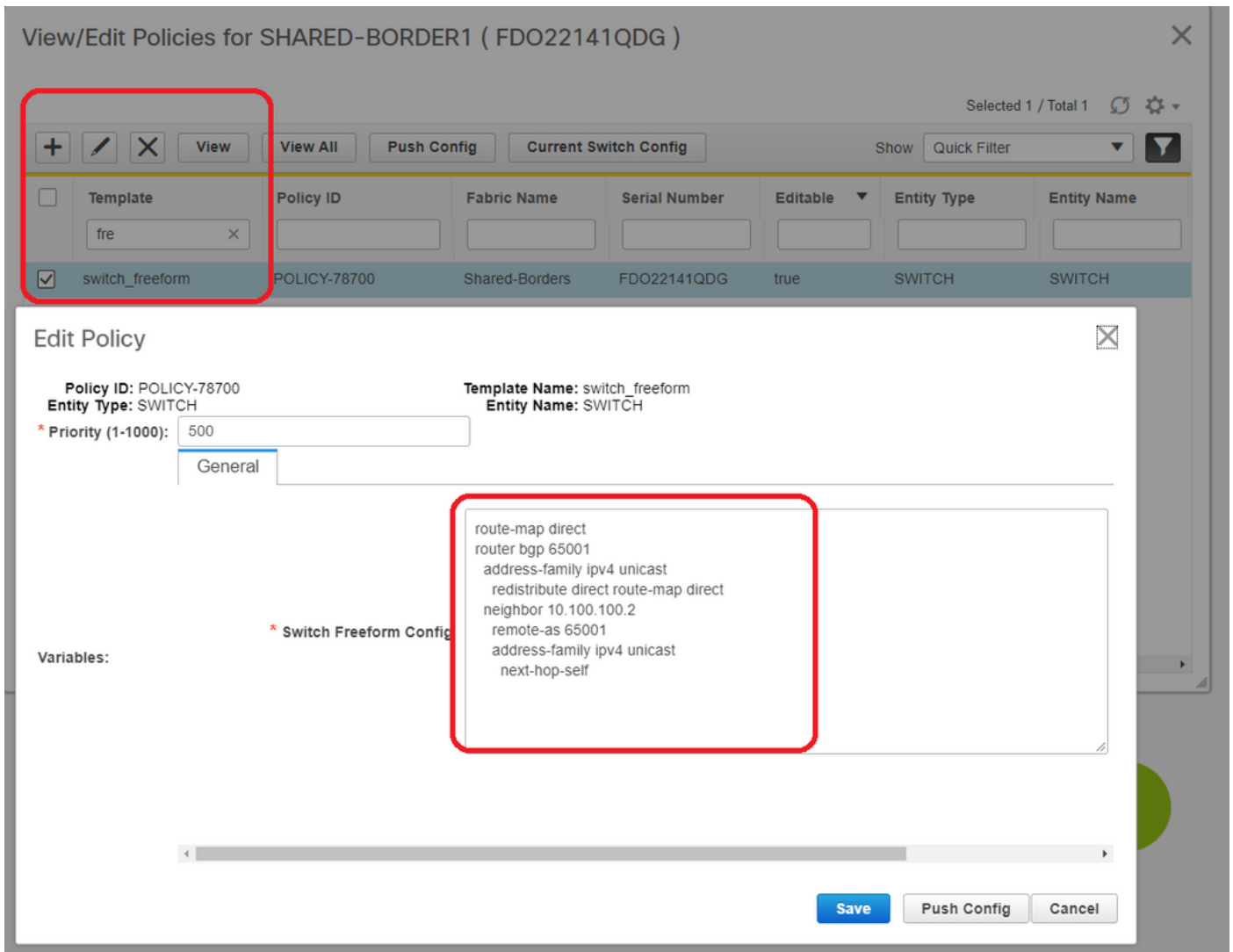
#查詢在共用邊界的備份SVI上配置的IP地址；如上所示，Shared-border1交換機上會新增freeform，並且指定的iBGP鄰居是Shared-border2(10.100.100.2)的iBGP鄰居

#請注意，在DCNM中提供自由形式的配置時，請在每個命令後提供正確的間距(保留偶數空格；即在路由器bgp65001後，提供兩個空格，然後發出neighbor <>命令等)

#還確保在BGP或其他形式中為直接路由（環迴路由）執行重分發直接通告以通告環回；在上方示例中，建立路由對映直接以匹配所有直接路由，然後在IPv4 AF BGP內完成redistribute direct

#從DCNM中「儲存和部署」配置後，iBGP鄰居關係的形成如下所示；

```
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
```
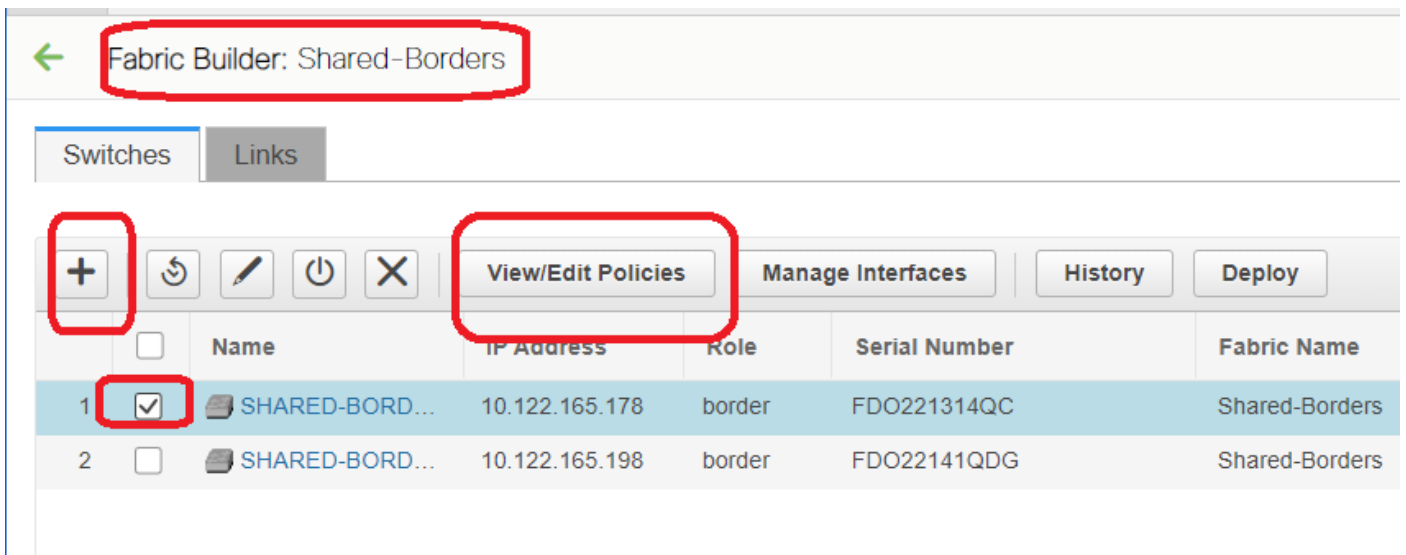
shared border1 to shared border2

#在以上步驟中，多站點底層已完全配置。

#下一步是構建多站點覆蓋；

第9步：構建從BGW到共用邊界的多站點重疊

#請注意，此處共用邊界也是路由伺服器

#選擇MSD，然後轉到「表格檢視」，可以在其中建立新連結；從這裡開始，必須建立一個新的多站點重疊鏈路，並且相關IP地址必須提供正確的ASN，如下所示；必須對所有l2vpn evpn鄰居（從每個BGW到每個共用邊界）執行此步驟
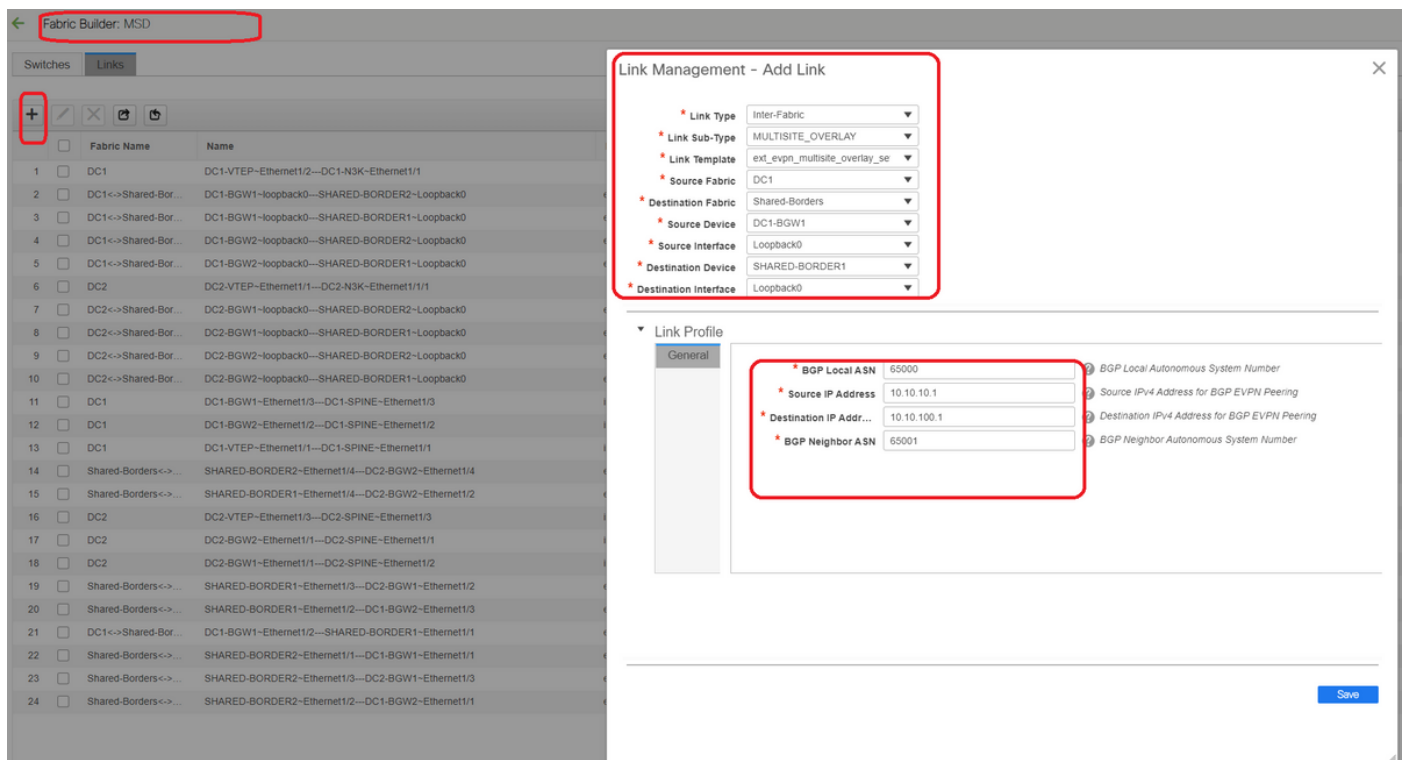
#以上是一個例子；對所有其他多站點重疊連結執行相同操作，最後CLI如下所示；

```
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
```

步驟10:在兩個站點上部署網路/VRF

#當我們完成多站點底層和重疊後，下一步是在所有裝置上部署網路/VRF;

#從交換矩陣上的VRF開始 — > DC1、DC2和共用邊界。
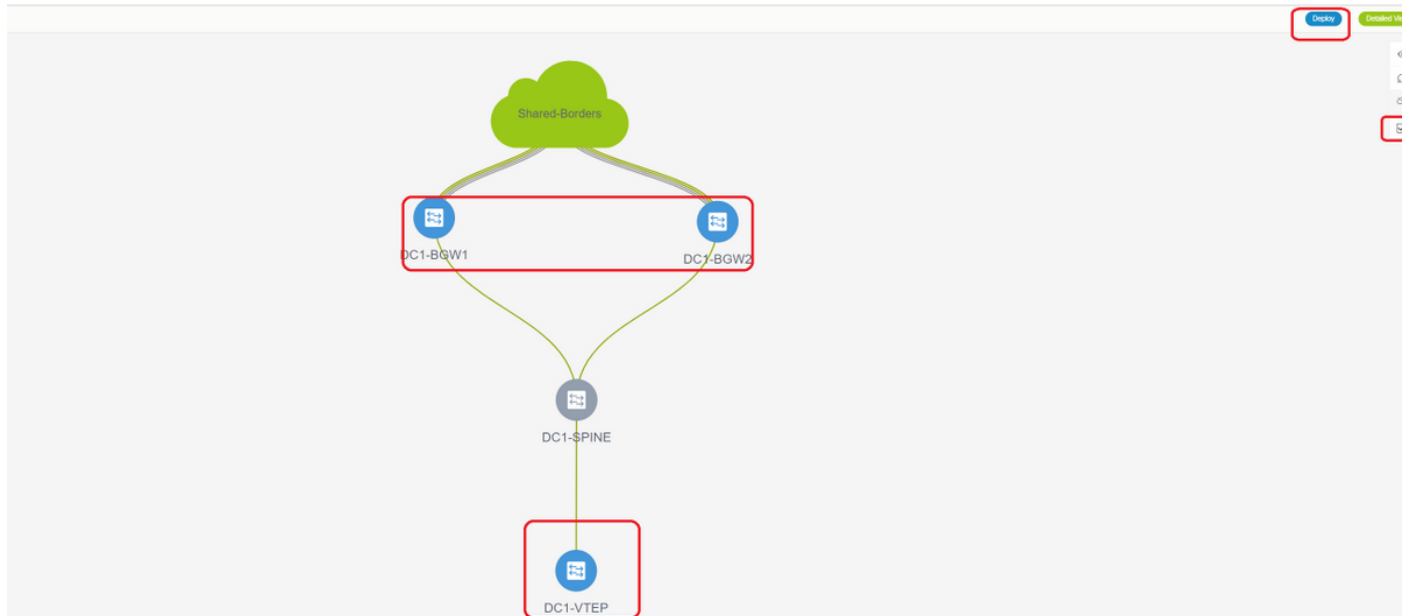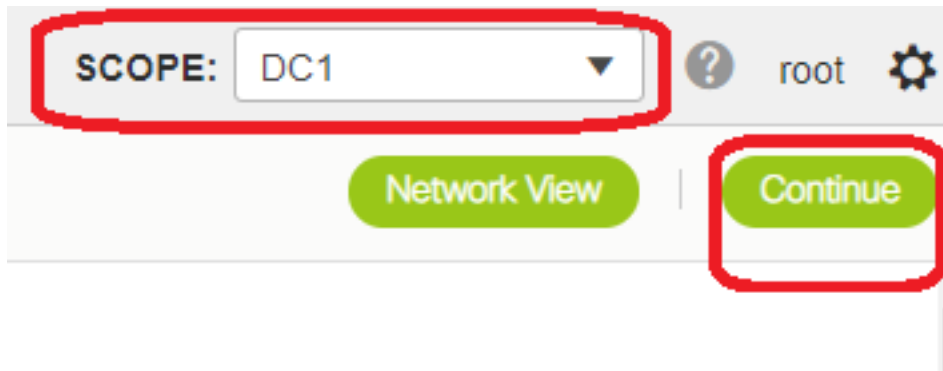
#選擇VRF檢視後，按一下「繼續」;這將列出拓撲中的裝置

#由於VRF必須部署到多個交換機（包括邊界網關和枝葉），請選中最右側的覈取方塊，然後選擇同時具有相同角色的交換機;如;一次可以選擇DC1-BGW1和DC1-BGW2，然後儲存兩台交換機;之後，選擇適用的枝葉交換機（此處是DC1-VTEP）
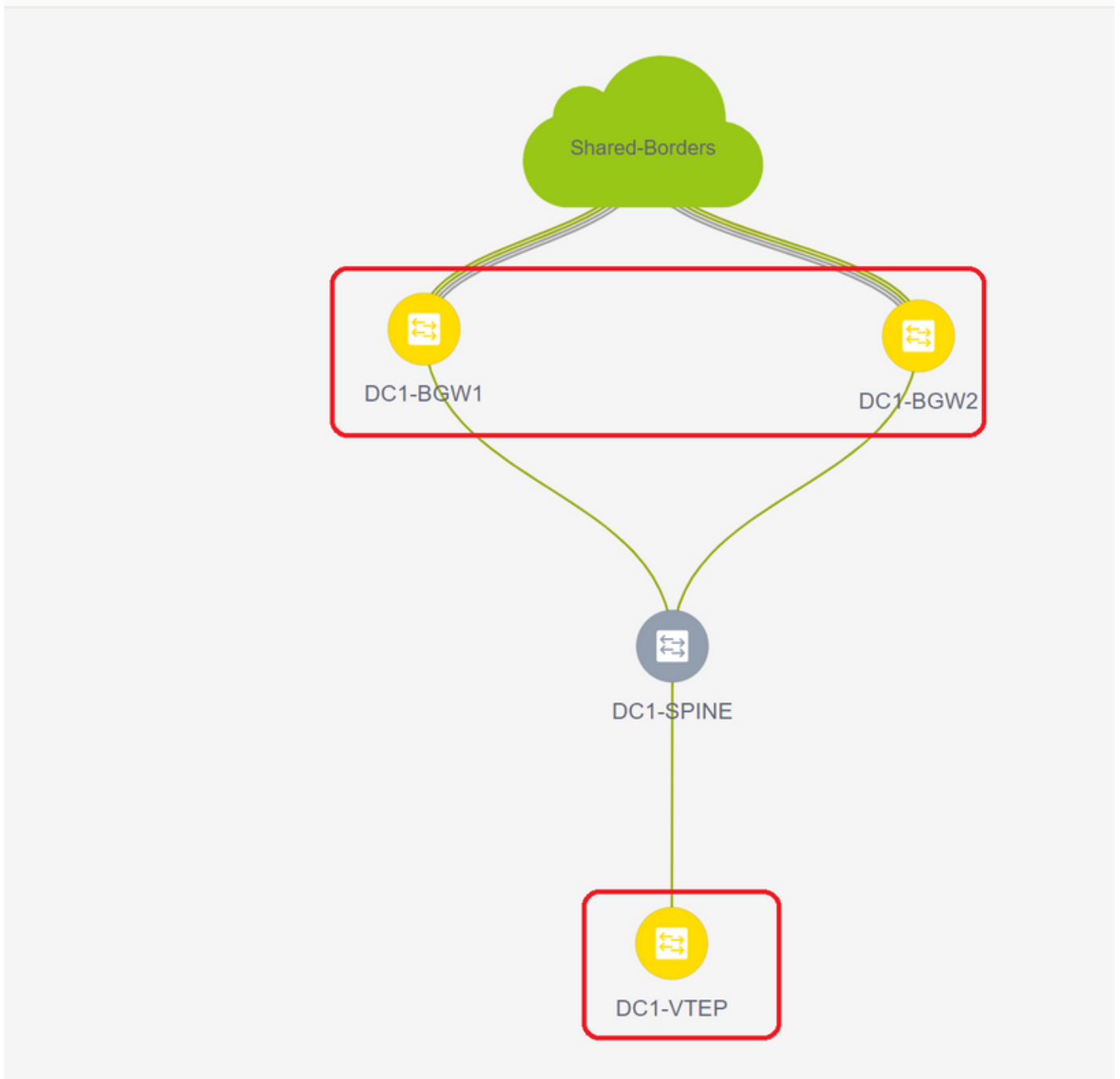
#如上所示，當選擇「部署」選項時，所有先前選擇的交換機都將開始部署，並且如果部署成功，最終變為綠色。

#部署網路必須執行相同的步驟；

#如果建立了多個網路，請記住在部署之前導航到後續頁籤以選擇網路
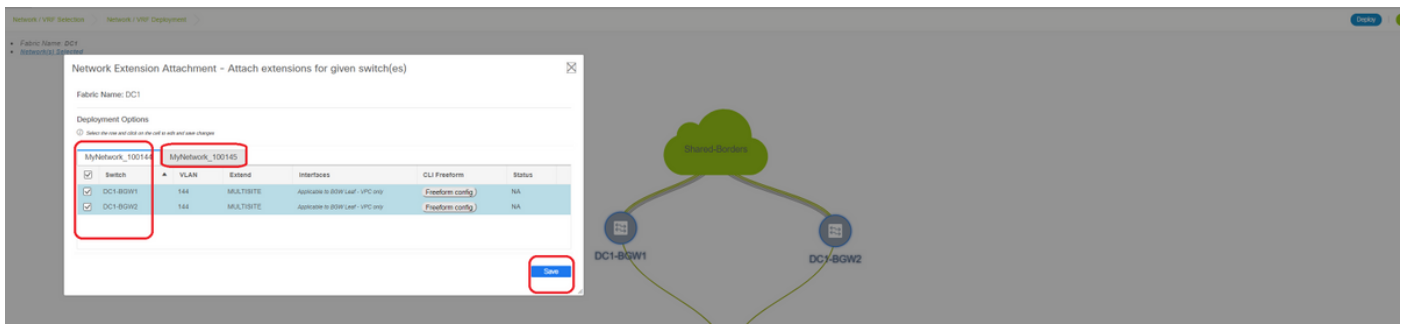
#狀態現在會從「NA」變為「DEPLOYED」，可以使用下面的交換機CLI驗證部署

```
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]
```
#以上也來自BGW;因此，簡而言之，我們在步驟早期選擇的所有交換機都將與網路和VRF一起部署

#交換矩陣DC2共用邊界也必須執行相同的步驟。請記住，共用邊界不需要任何網路或第2層VNID;僅需要L3 VRF。

步驟11:在枝葉交換機/VTEP上建立下游中繼/接入埠

#在此拓撲中，分別來自DC1-VTEP和DC2-VTEP的埠Eth1/2和Eth1/1連線到主機；因此在DCNM GUI中將這些埠作為中繼埠移動，如下所示

Shared-Borders

DC1-BGW1   DC1-BGW2

DC1-SP

DC1-VTEP

Set role   ▶

Modes   ▶

vPC Pairing

Manage interfaces

View/edit policies

History

Deploy Config

Discovery   ▶
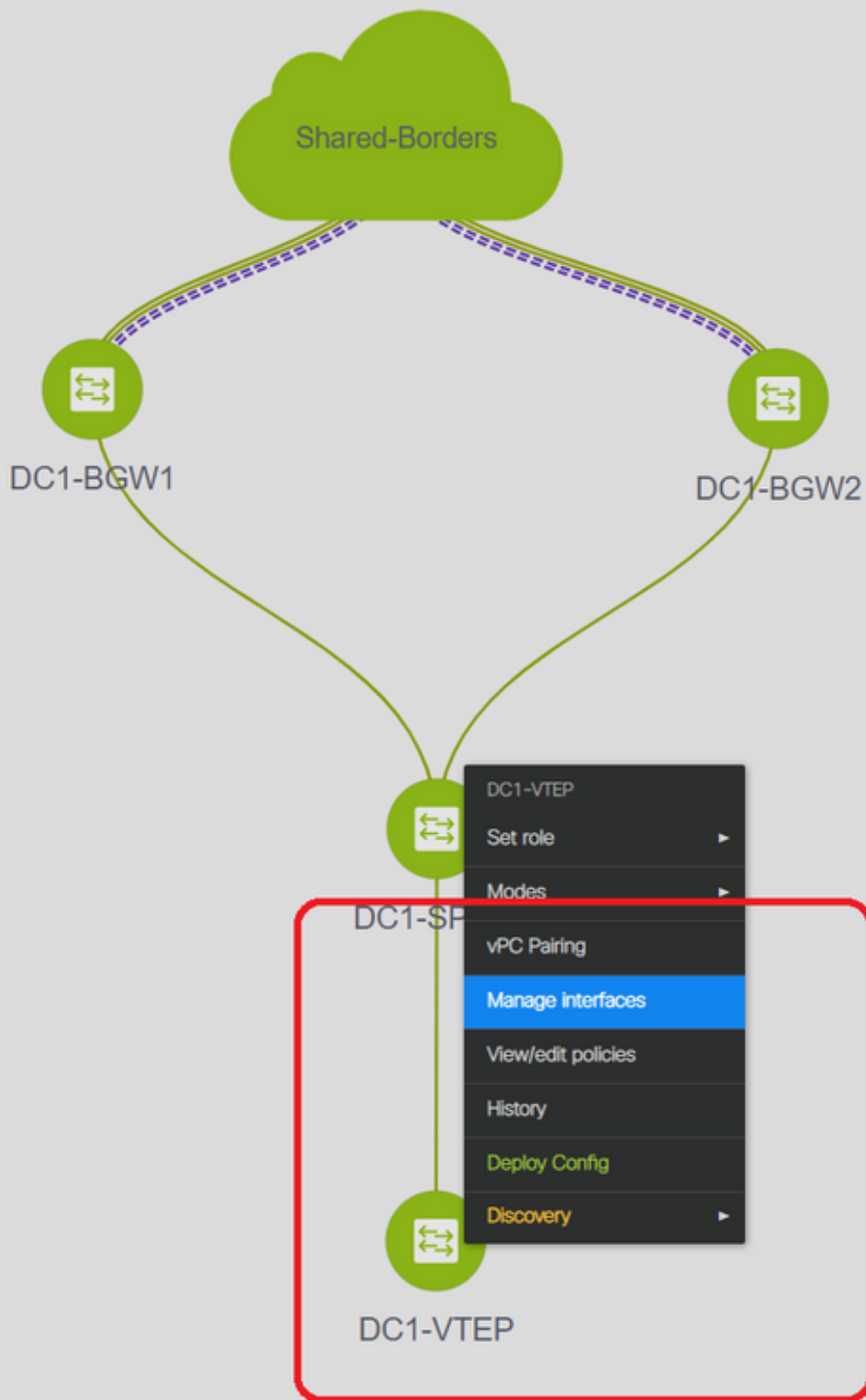
DC1-VTEP

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 sho
strictly match 'show run' o
with respect to case and
Any mismatches will yiel
unexpected diffs during a

#選擇相關介面並將「允許的vlan」從「無」更改為「全部」（或僅允許需要的vlan）

步驟12:共用邊框上所需的自由形式

#由於共用邊界交換機是路由伺服器，因此需要在BGP l2vpn evpn鄰居關係方面進行一些更改

#使用單播複製站點間BUM流量；表示在Vlan 144（例如）中到達BGW後的任何BUM流量；根據哪個BGW是指定轉發器(DF),DF將執行到遠端站點的單播複製；此複製在BGW從遠端BGW接收到第3類路由後實現；這裡，BGW正在形成l2vpn evpn對等僅與共用邊界對等；並且共用邊界不應具有任何第2層VNID（如果建立，將導致東/西流量被黑鎖）。 由於第2層VNID缺失，且路由型別3由每個VNID的BGW產生，因此共用邊界不會執行從BGW傳入的BGP更新；要解決此問題，請使用AF l2vpn evpn下的「retain route-target all」

#另一個要點是確保共用邊界不會更改下一跳（BGP DEFAULT會更改eBGP鄰居的下一跳）；在這裡，從站點1到2的單點傳播流量的站點間隧道應該從BGW到BGW（從dc1到dc2，反之亦然）；為此，必須為從共用邊界到每個BGW的每個l2vpn evpn鄰居建立並應用路由對映

#對於以上兩點，必須在共用邊框（如下所示）上使用自由形式

```
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
```
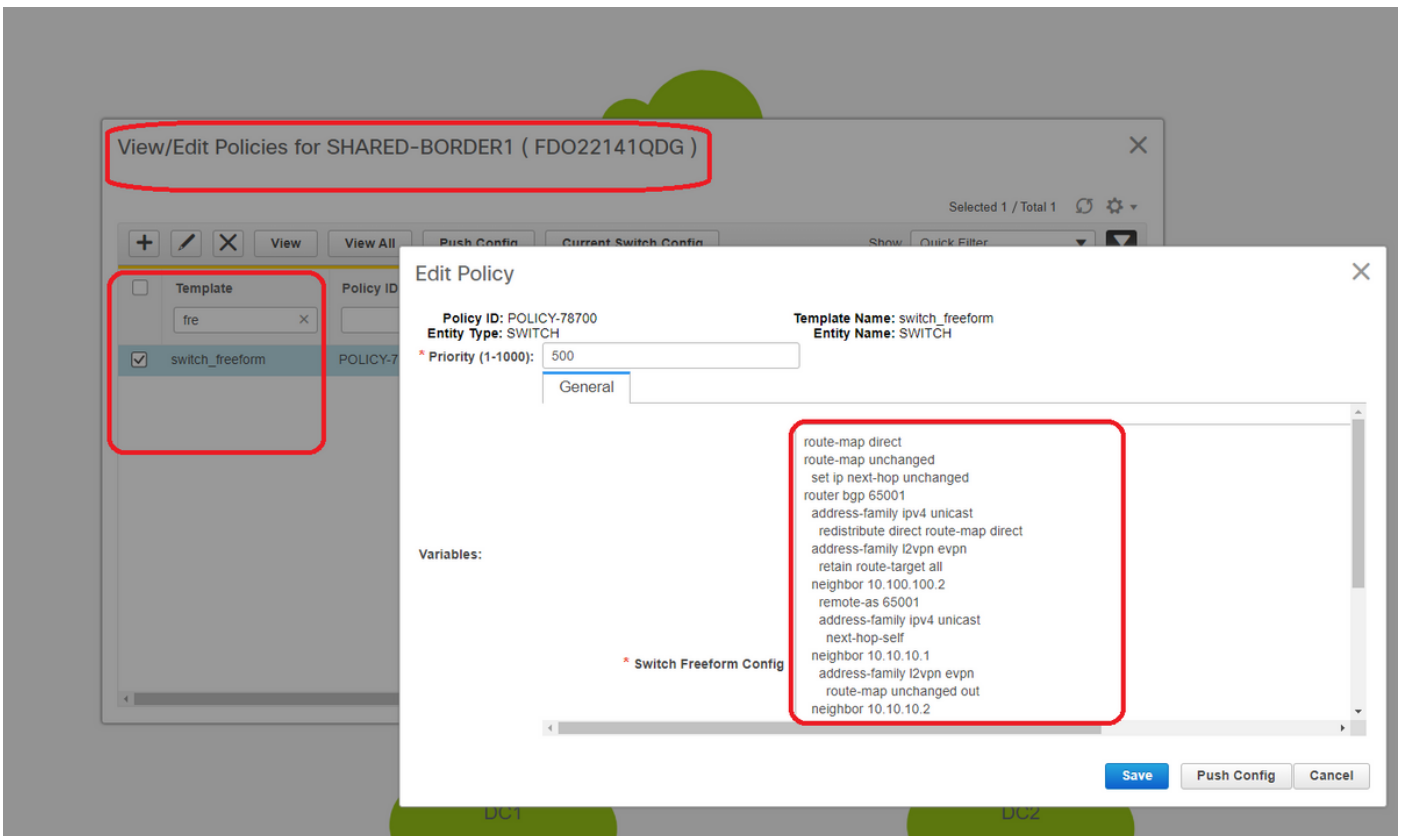

步驟13:BGW上租戶VRF內的環回

#對於來自枝葉交換機內所連線主機的北/南流量，BGW使用NVE Loopback1 IP地址的外部SRC IP;共用邊界僅在預設情況下形成與BGW的多站點環回Ip地址的NVE對等；因此，如果vxlan封包與BGW Loopback1的外部SRC IP位址進入共用邊界，則封包將由於SRCTEP未命中而遭捨棄；為避免這種情況，必須在每個BGW交換機上建立租戶 — VRF中的環回，然後通告到BGP，以便共用邊界接收此更新，然後與BGW Loopback1 IP地址形成NVE對等；

#最初NVE對等項在共用邊界上如下所示

```
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
```

Add Interface   ✕

	* Type:	Loopback	▼
	* Select a device	DC1-BGW2	▼
	* Loopback ID	2	
	* Policy:	int_loopback_11_1	▼

General

Interface VRF	tenant-1	❷ Interface VRF name, default VRF if not specified
* Loopback IP	172.17.10.2|	❷ IP address of the loopback
Route-Map TAG	12345	❷ Route-Map tag associated with interface IP
Interface Description		❷ Add description to the interface (Max Size 254)

Freeform Config

❷ Note ! All configs sho
strictly match 'show run' c
with respect to case and
Any mismatches will yiel
unexpected diffs during o

Save   Preview   Deploy

#如上所示，loopback2是從DCNM建立的，在tenant-1 VRF中配置，並給予12345標籤，因為此標籤是路由對映在進行通告時用於匹配環回的標籤

```
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:
```
#執行此步驟後，NVE對等項將顯示所有Loopback1 IP地址以及多站點環回IP地址。

```
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
```
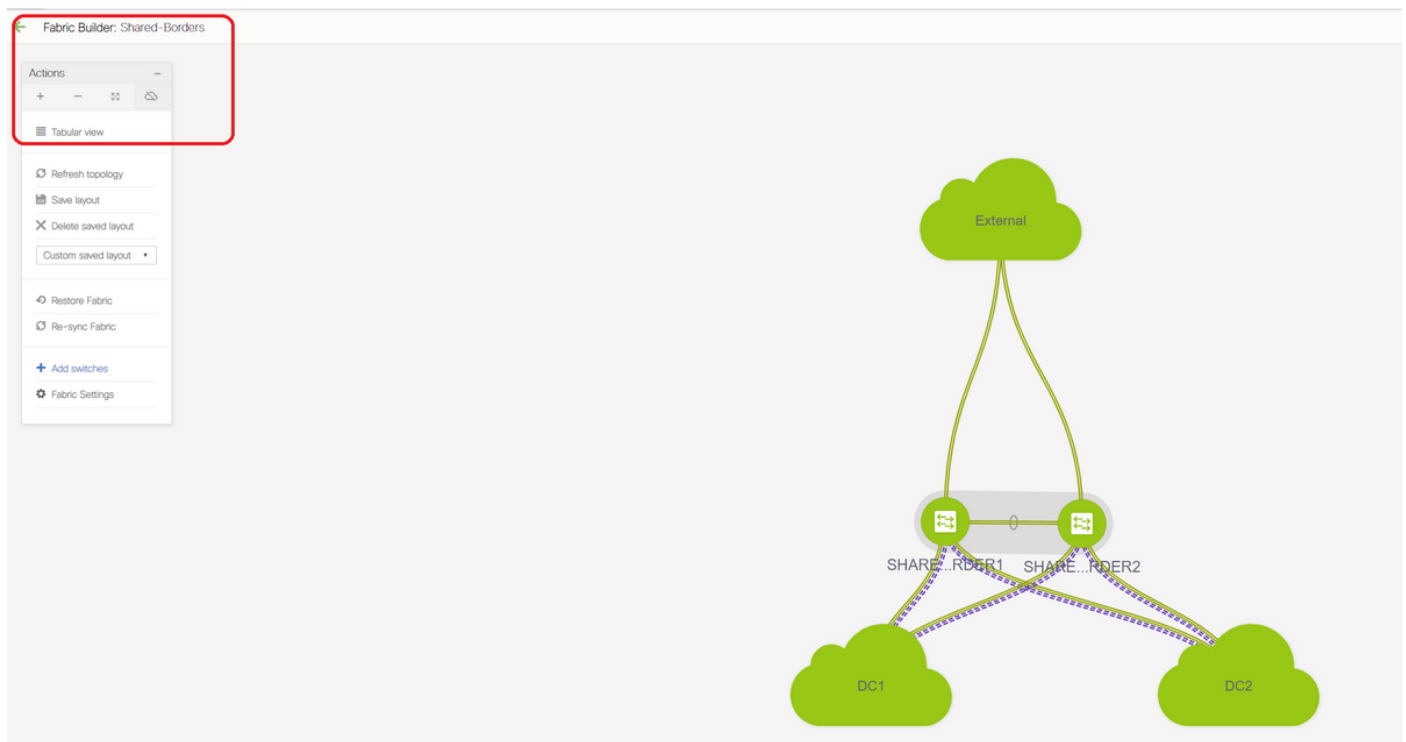#在此階段，應正確轉發東/西流量

步驟14:從共用邊界到外部路由器的VRFLITE擴展

#有時交換矩陣外部的主機必須與交換矩陣內的主機通訊。在此示例中，共用邊界使相同內容成為可能；

#位於DC1或DC2中的任何主機都將能夠通過共用邊界交換機與外部主機通訊。
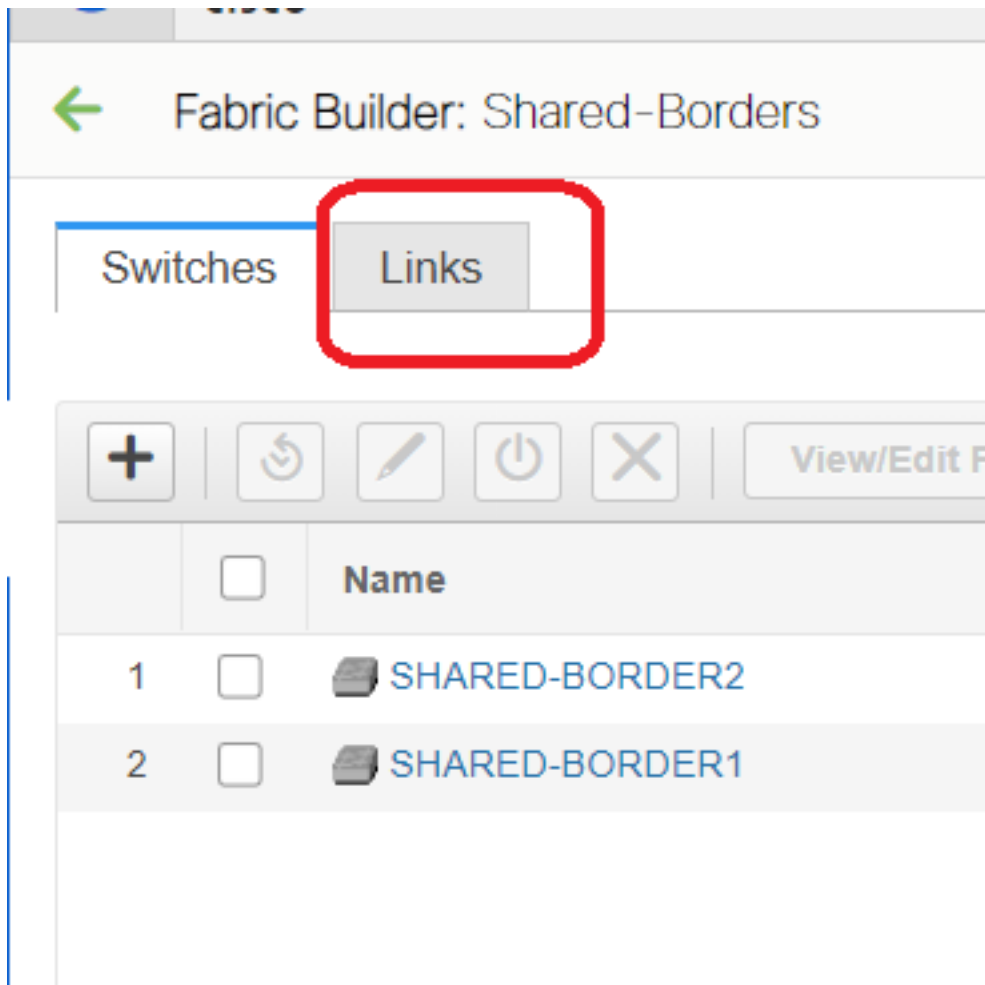
#為此目的，共用邊界將終止VRF Lite;在此示例中，eBGP從共用邊界運行到外部路由器，如圖開頭所示。

#要從DCNM配置此功能，需要新增**vrf擴展附件**。為實現這一目標，需要執行以下步驟。

a)將交換矩陣間鏈路從共用邊界新增到外部路由器

#將交換矩陣生成器範圍選擇為「共用邊框」並更改為表格檢視

Fabric Builder: Shared-Borders

Switches   Links

		Name
1		SHARED-BORDER2
2		SHARED-BORDER1

#選擇鏈路並新增「交換矩陣間」連結，如下所示

#必須從下拉選單中選擇VRF LITE子型別

#源交換矩陣為共用邊界，目標交換矩陣為外部，因為這將是從小型企業到外部的VRF LITE

#選擇流向外部路由器的相關介面

#提供IP地址和掩碼以及鄰居IP地址

ASN將自動填充。

#完成此操作後，按一下「儲存」

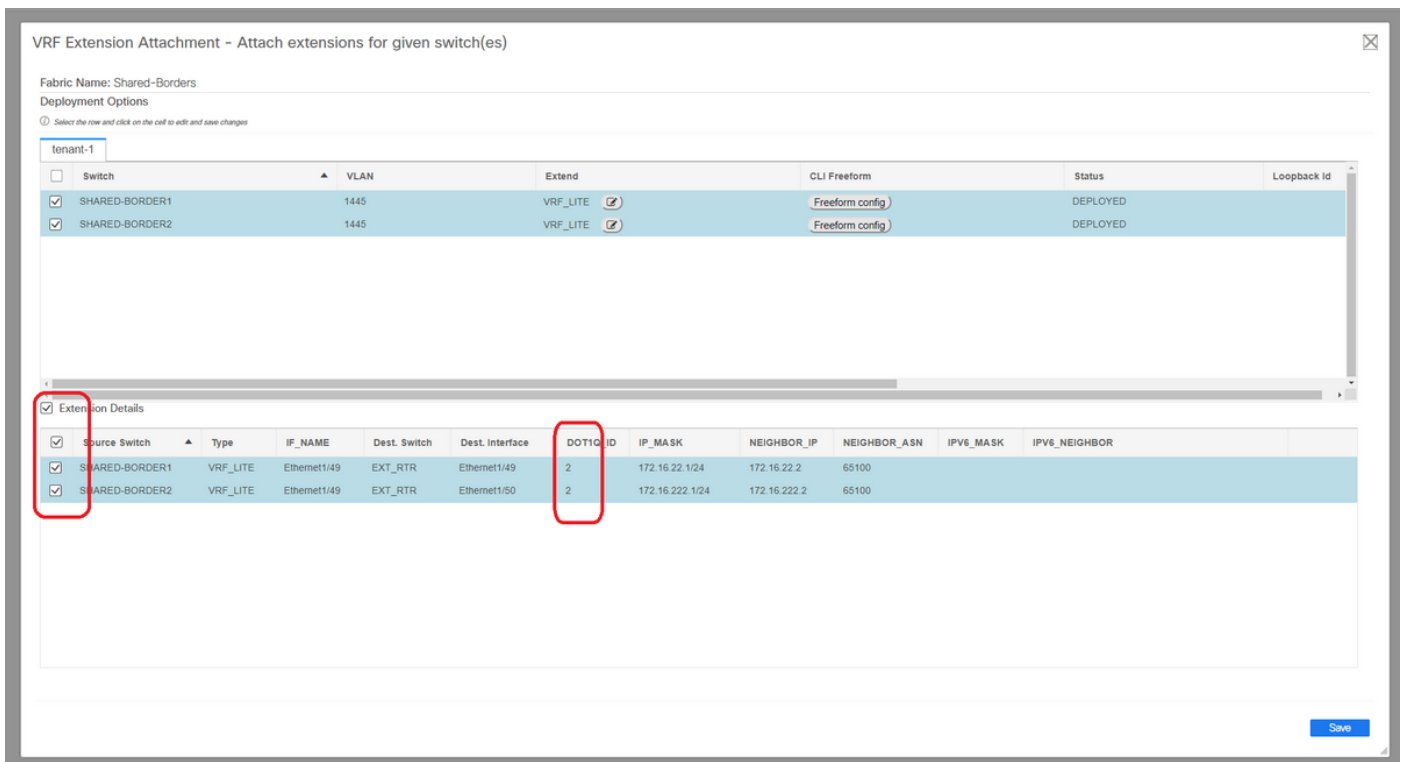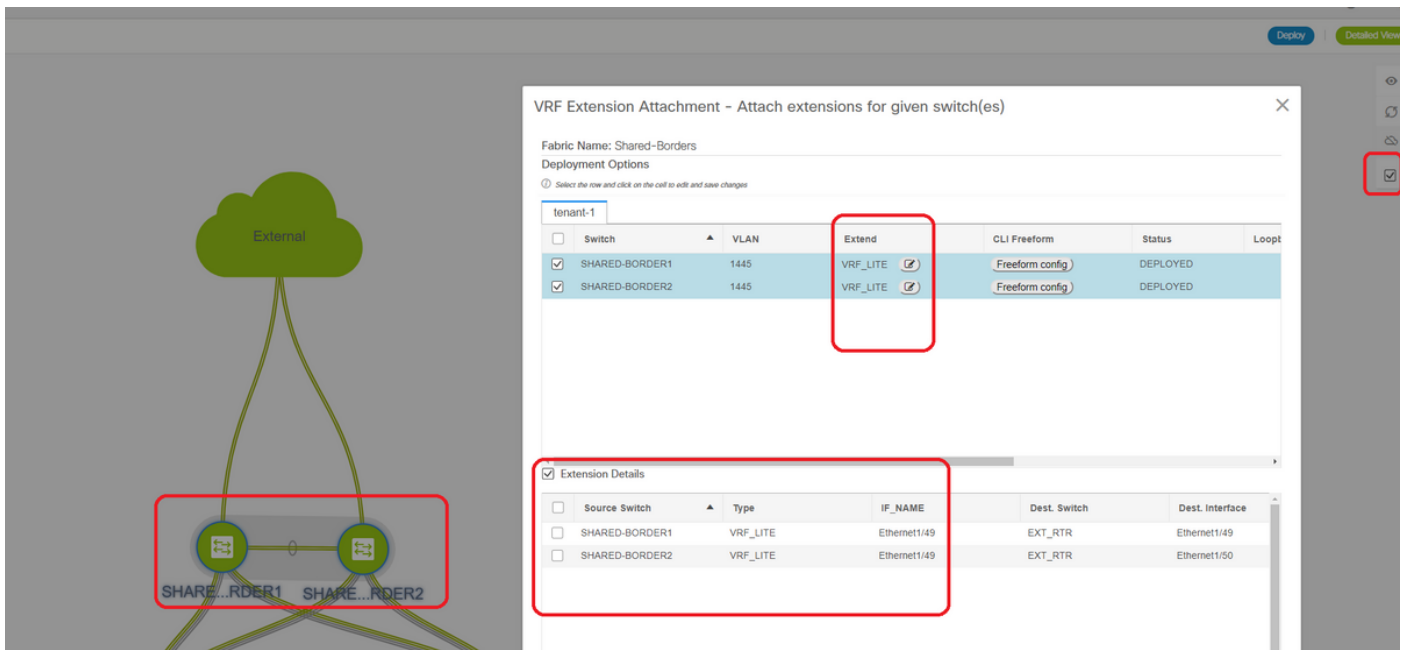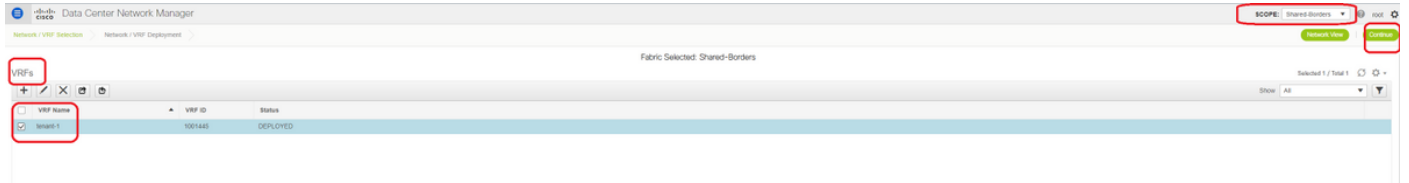#對共用邊界和VRFLITE中的所有外部第3層連線執行相同操作

b)新增VRF擴展

#轉至共用邊界VRF部分

VRF將處於部署狀態；選中右側的靈取方塊，以便可以選擇多個交換機

#選擇共用邊框，將會開啟「VRF EXtension attachment」視窗

#在「extend」下，從「None」更改為「VRFLITE」

#對兩個共用邊框執行相同操作

#完成後，「Extension Details」將填充先前在步驟a)中給出的VRF LITE介面。

DOT1Q ID自動填充為2

#其他欄位也會自動填充

#如果必須通過VRFLITE建立IPv6鄰居關係，則應對IPv6執行步驟a)

#現在按一下「儲存」

#最後，在網頁右上角執行「部署」。

#成功部署將導致將配置推送到共用邊界，其中包括在這些子介面上設定IP地址並與外部路由器建立 BGP IPv4鄰居

#請記住，在這種情況下，外部路由器配置（在子介面上設定IP地址和BGP鄰居關係語句）由CLI手動完成。

#可通過共用邊界上的以下命令進行CLI驗證；

```
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
```

#使用上述所有配置，將建立北/南連通性，如下所示（從外部路由器到交換矩陣中主機的ping）

```
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
```

Traceroute還指向資料包路徑中的正確裝置

```
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#
```