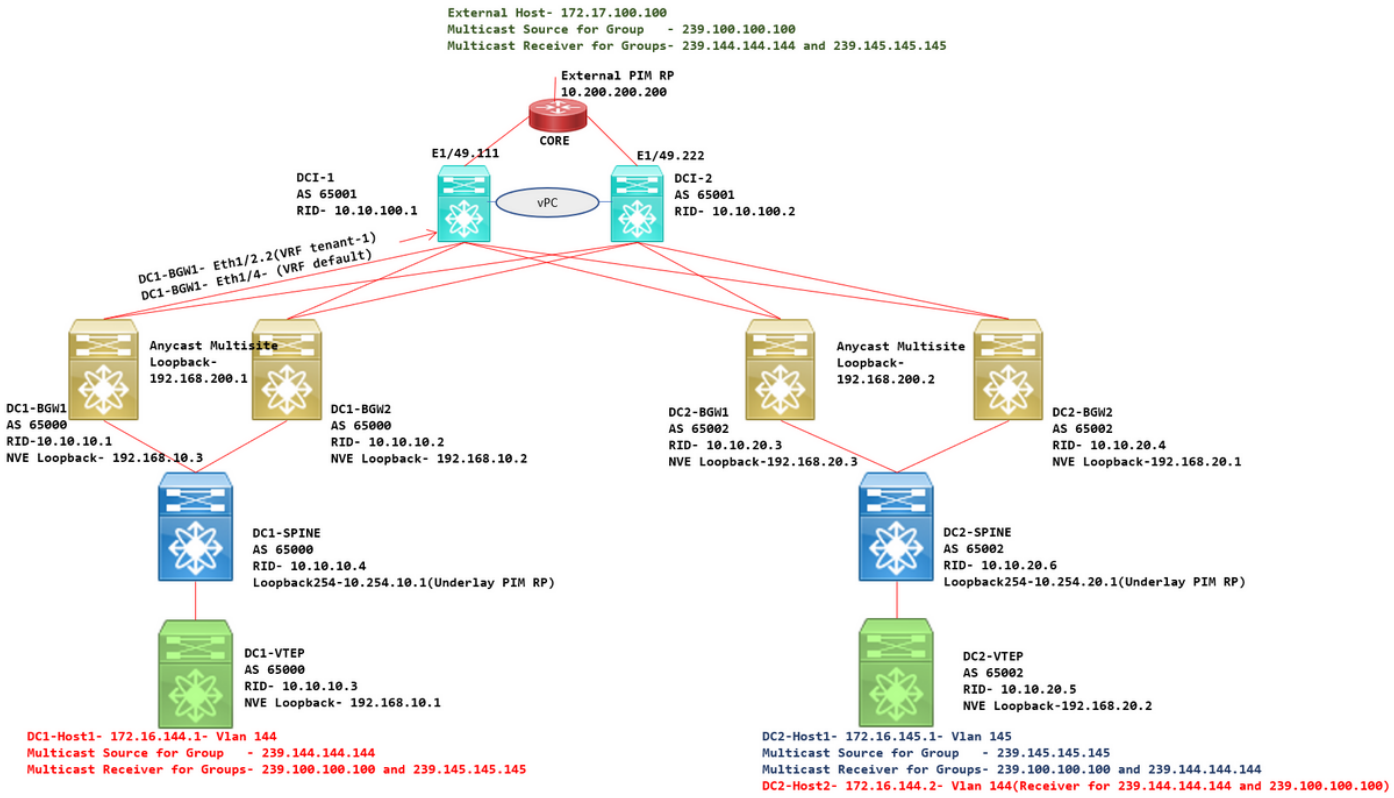
使用DCNM构建Nexus 9000 VXLAN多站点TRM

目录

简介

本文档将介绍如何部署Cisco Nexus 9000 VXLAN多站点TRM交换矩阵，其中边界网关通过DCI交换机连接

拓扑



拓扑详细信息

- DC1和DC2是运行VXLAN的两个数据中心位置。
- DC1和DC2边界网关通过DCI交换机相互连接。
- DCI交换机不运行任何VXLAN;这些设备正在为底层运行eBGP，以实现从DC1到DC2的可达性，反之亦然。此外，DCI交换机还配置了租户vrf;在本例中，它为vrf- "tenant-1"。
- DCI交换机还连接到非VXLAN的外部网络。
- VRFLITE连接在边界网关上终止(支持从NXOS-9.3(3)和DCNM-11.3(1)开始的VRFLITE和边界网关功能共存)
- 边界网关在任播模式下运行；在此版本上运行TRM（租户路由组播）时，边界网关不能配置为vPC（请参阅多站点TRM配置指南了解其他限制）
- 对于此拓扑，所有BGW交换机都有两个到每个DCI交换机的物理连接；一个链路将处于默认VRF中（将用于站点间流量），而另一个链路将处于VRF租户–1中，该租户用于将VRFLITE扩展到非vxlan环境。

PIM/组播详细信息（TRM特定）

- 两个站点的底层PIM RP是主干交换机，环回254配置为相同。使用底层PIM RP，以便VTEP可以将PIM寄存器和PIM加入发送到主干（用于各种VNID的BUM流量复制）
- 对于TRM，RP可以通过不同的方式指定；在本文中，PIM RP是位于拓扑顶部的核心路由器，位于VXLAN交换矩阵外部。
- 所有VTEP的核心路由器将指定为在各自VRF中配置的PIM RP

- DC1-Host1正在向组239.144.144.144发送组播；DC2-Host1是DC2中此组的接收方，而vxlan的外部主机(172.17.100.100)也订阅此组
- DC2-Host1正在向组239.145.145.145发送组播；DC1-Host1是DC1中此组的接收方，而vxlan的外部主机(172.17.100.100)也订阅此组
- DC2-Host2位于Vlan 144中，是组播组239.144.144.144和239.100.100.100的接收方
- 外部主机(172.17.100.100)正在发送DC1-Host1和DC2-Host1都是接收方的流量。
- 这包括东/西VLAN间和VLAN内以及北/南组播流量

使用的组件

- 运行9.3(3)的Nexus 9k交换机
- DCNM运行11.3(1)

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始（默认）配置。如果您使用的是真实网络，请确保您已经了解所有命令的潜在影响。

高级步骤

1)考虑到本文档基于两个使用VXLAN多站点功能的数据中心，必须创建两个简易交换矩阵

2)创建MSD并移动DC1和DC2

3)创建外部交换矩阵并添加DCI交换机

4)创建多站点底层和重叠

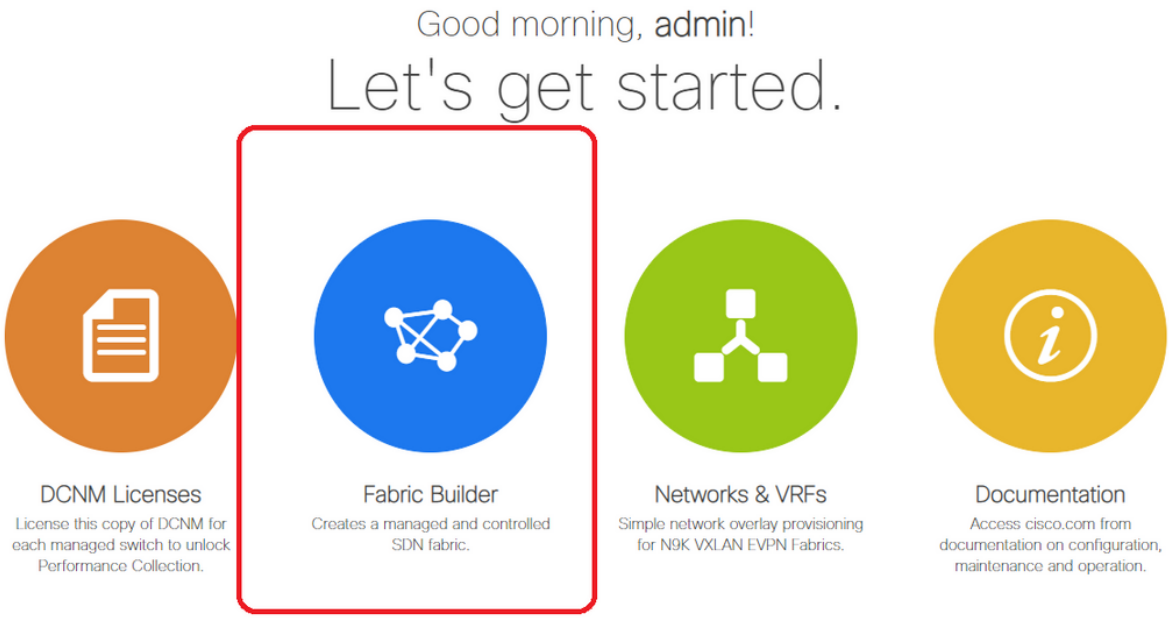
5)在边界网关上创建VRF扩展附件

6)单播流量验证

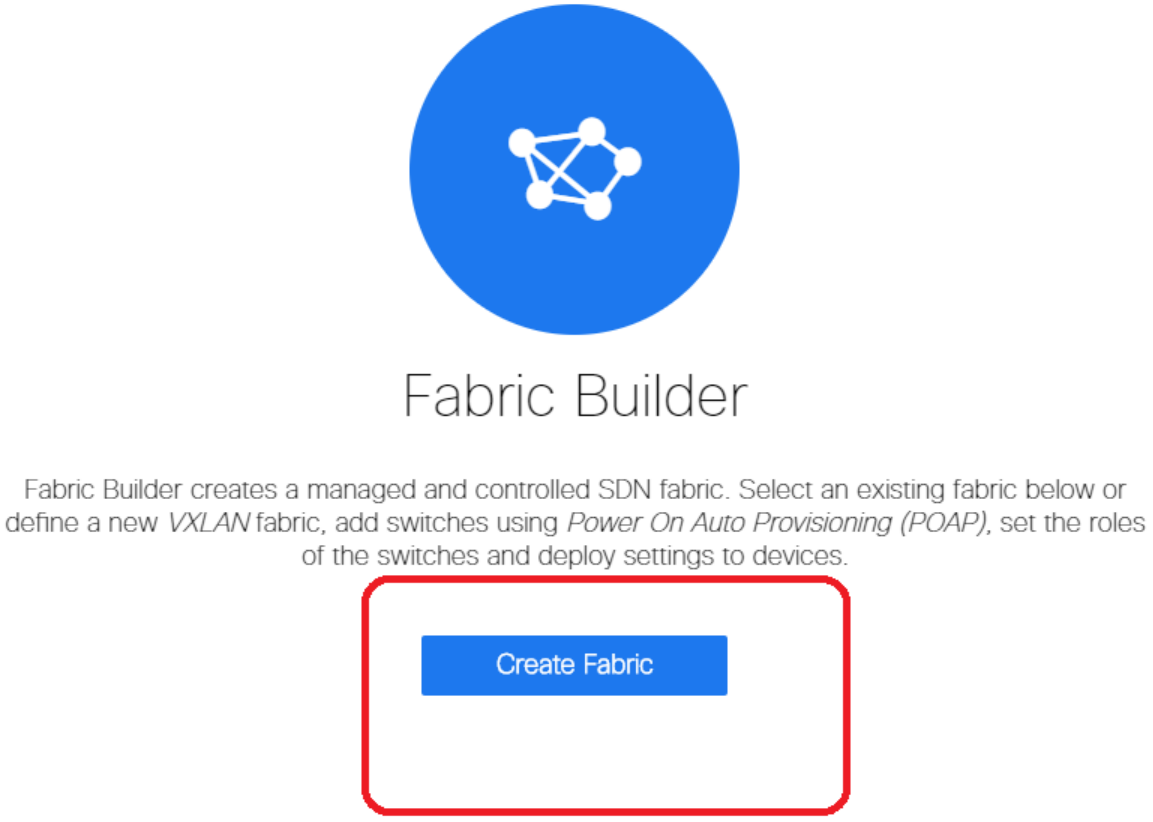
7)组播流量验证

步骤 1：为DC1创建简易交换矩阵

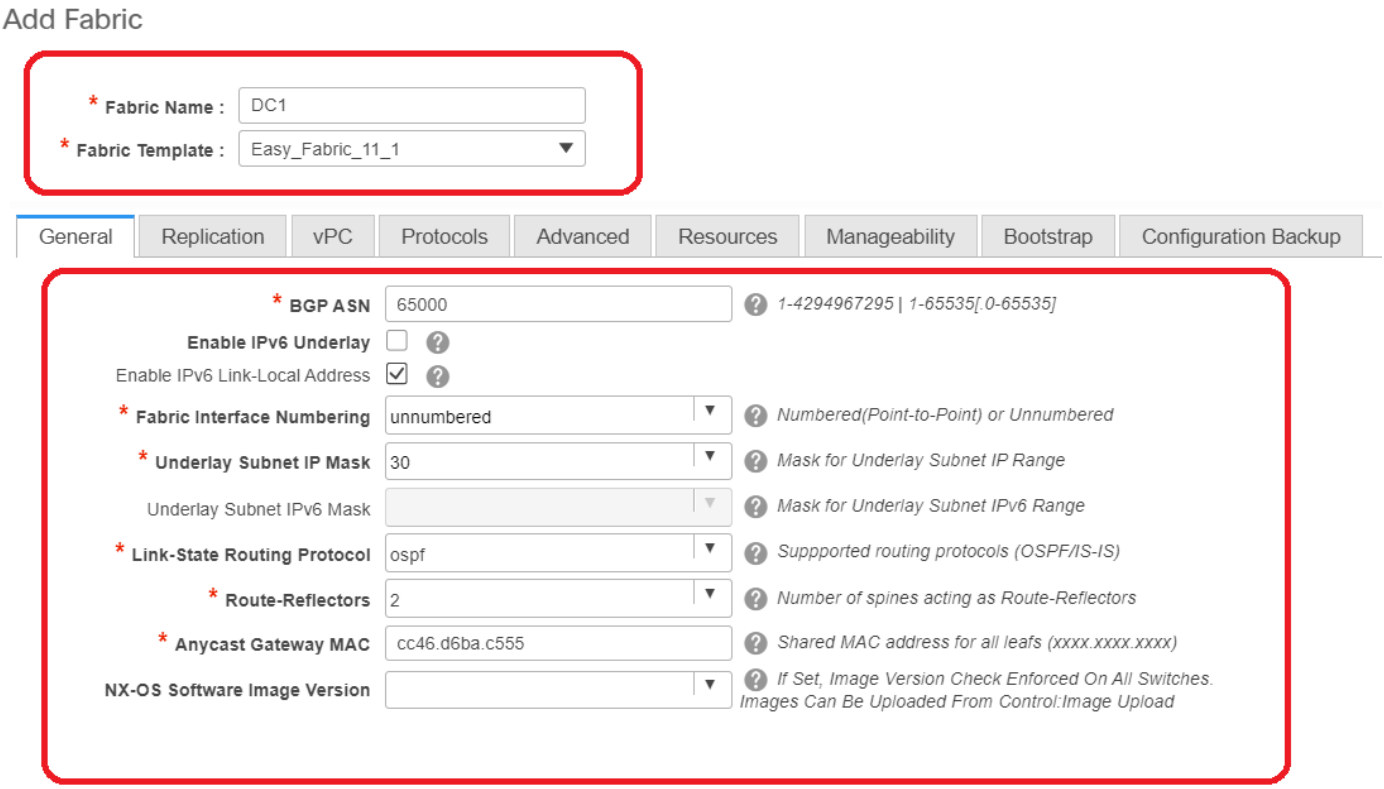
- 登录DCNM，从控制面板中选择选项 — > "Fabric Builder"



- 选择“创建交换矩阵”选项

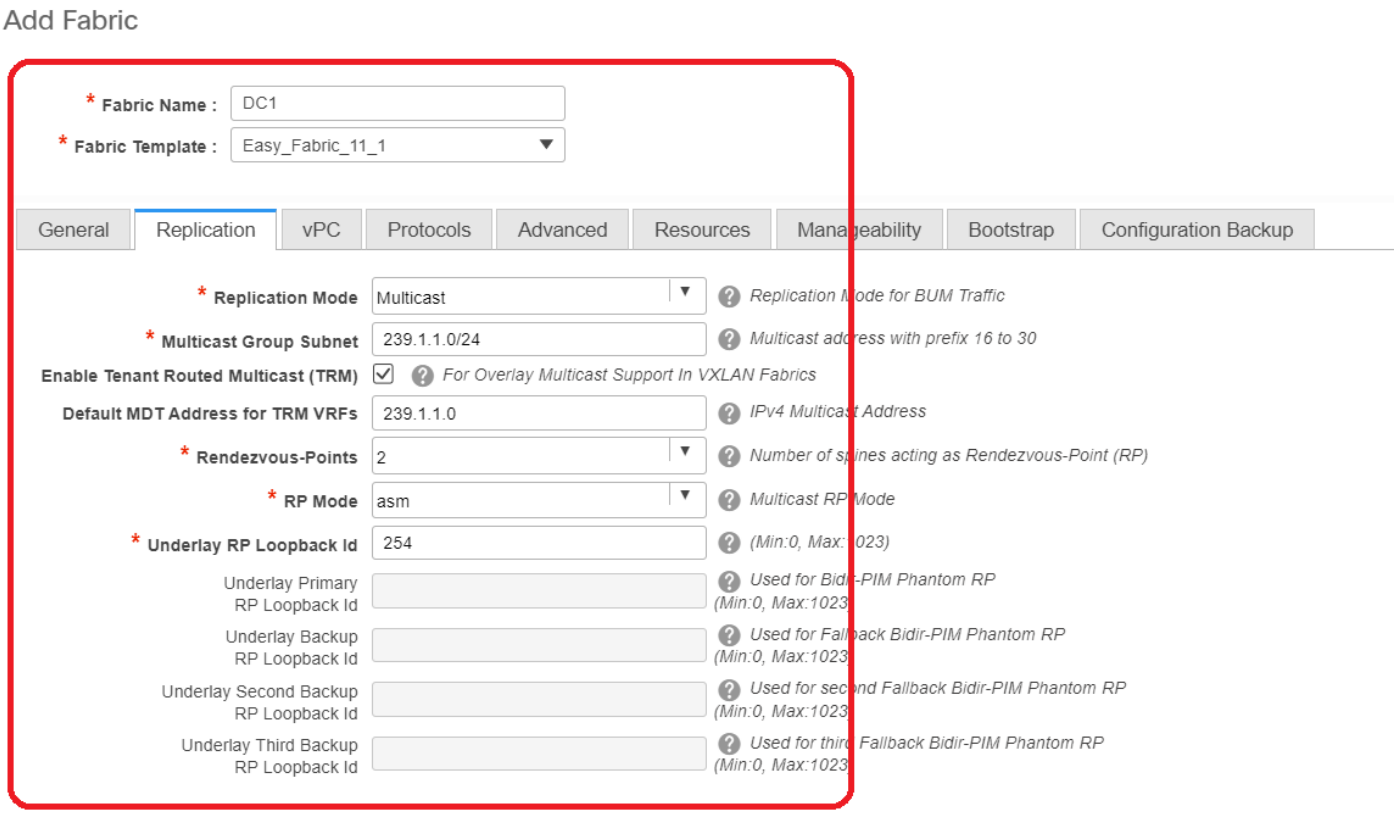


- 接下来，提供交换矩阵名称、模板，然后在“常规”选项卡下填写相关ASN、交换矩阵接口编号、任何广播网关MAC(AGM)



#交换矩阵中的主机将AGM用作默认网关MAC地址。所有枝叶交换机上的情况相同（因为交换矩阵内的所有枝叶交换机都运行任播交换矩阵转发）。 所有枝叶交换机上的默认网关IP地址和MAC地址将相同

- **接下来是设置复制模式**



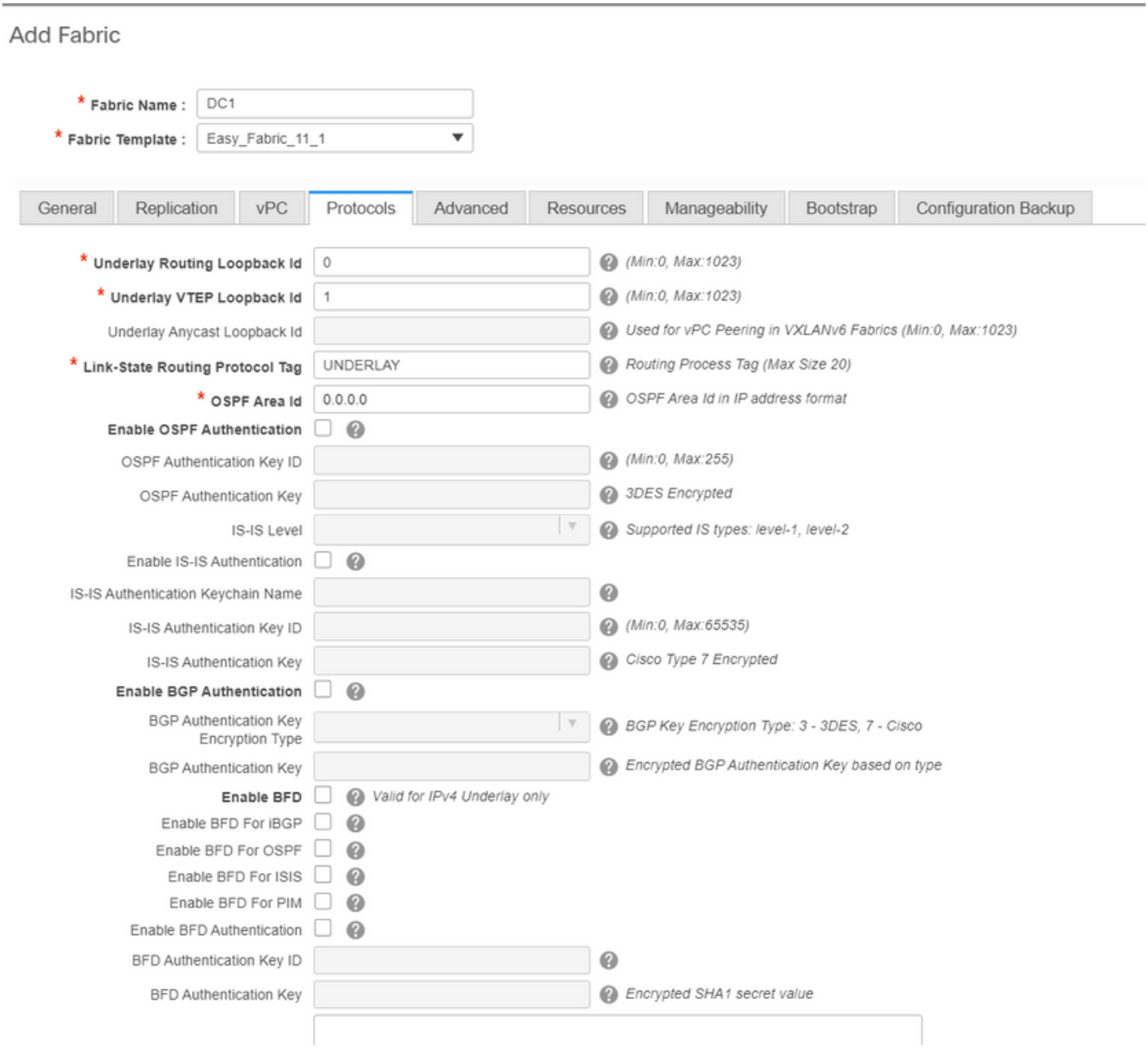
#本文档目的的复制模式是组播；另一个选项是使用入口复制(IR)

#组播组子网是VTEP用于复制BUM流量（如ARP请求）的组播组

#必须启用“启用租户路由组播(TRM)”复选框

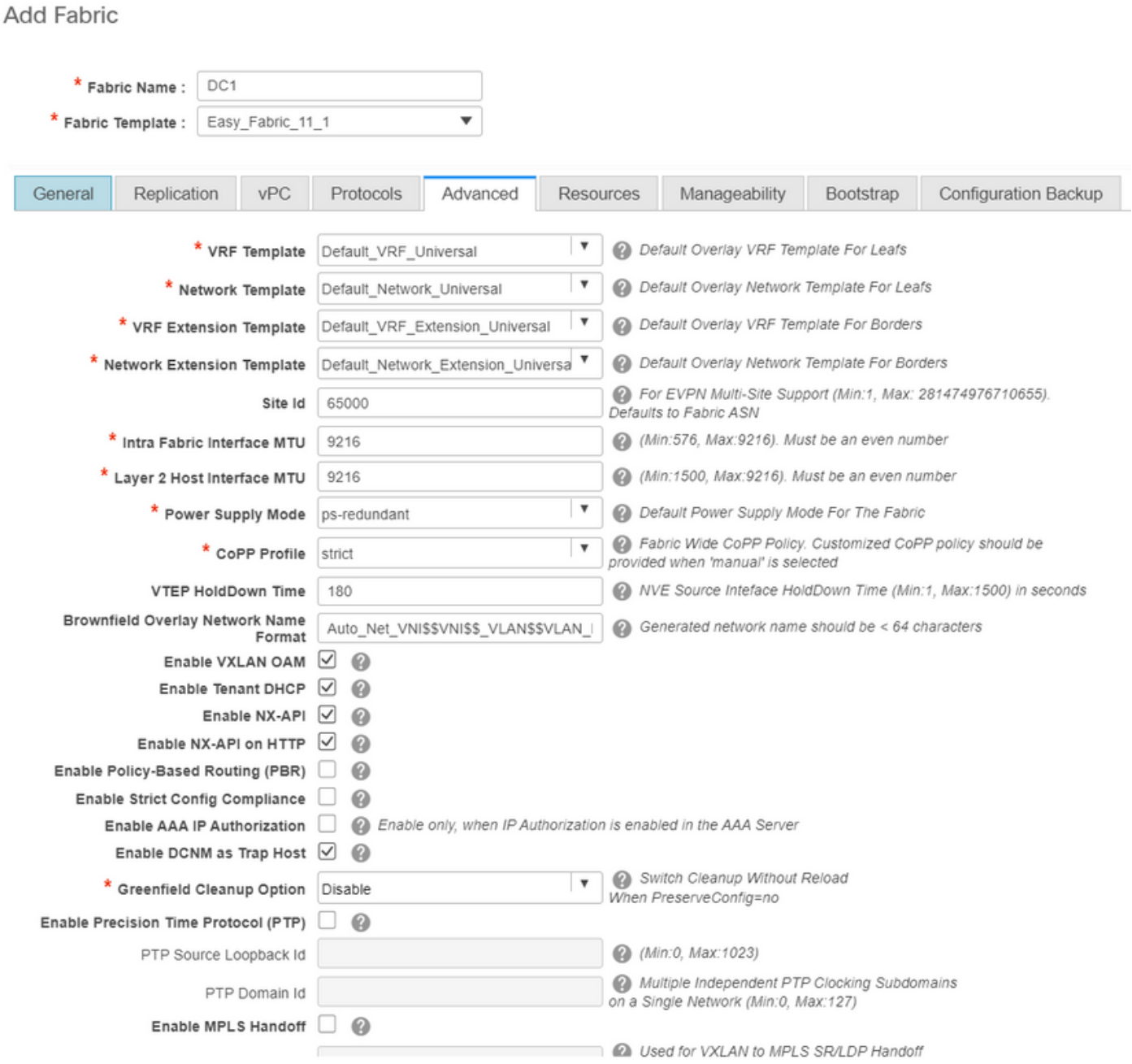
#根据需要填写其他框。

- vPC的选项卡保持不变，因为此处的拓扑未使用任何vPC
- 接下来是“协议”选项卡



#根据需要修改相关框。

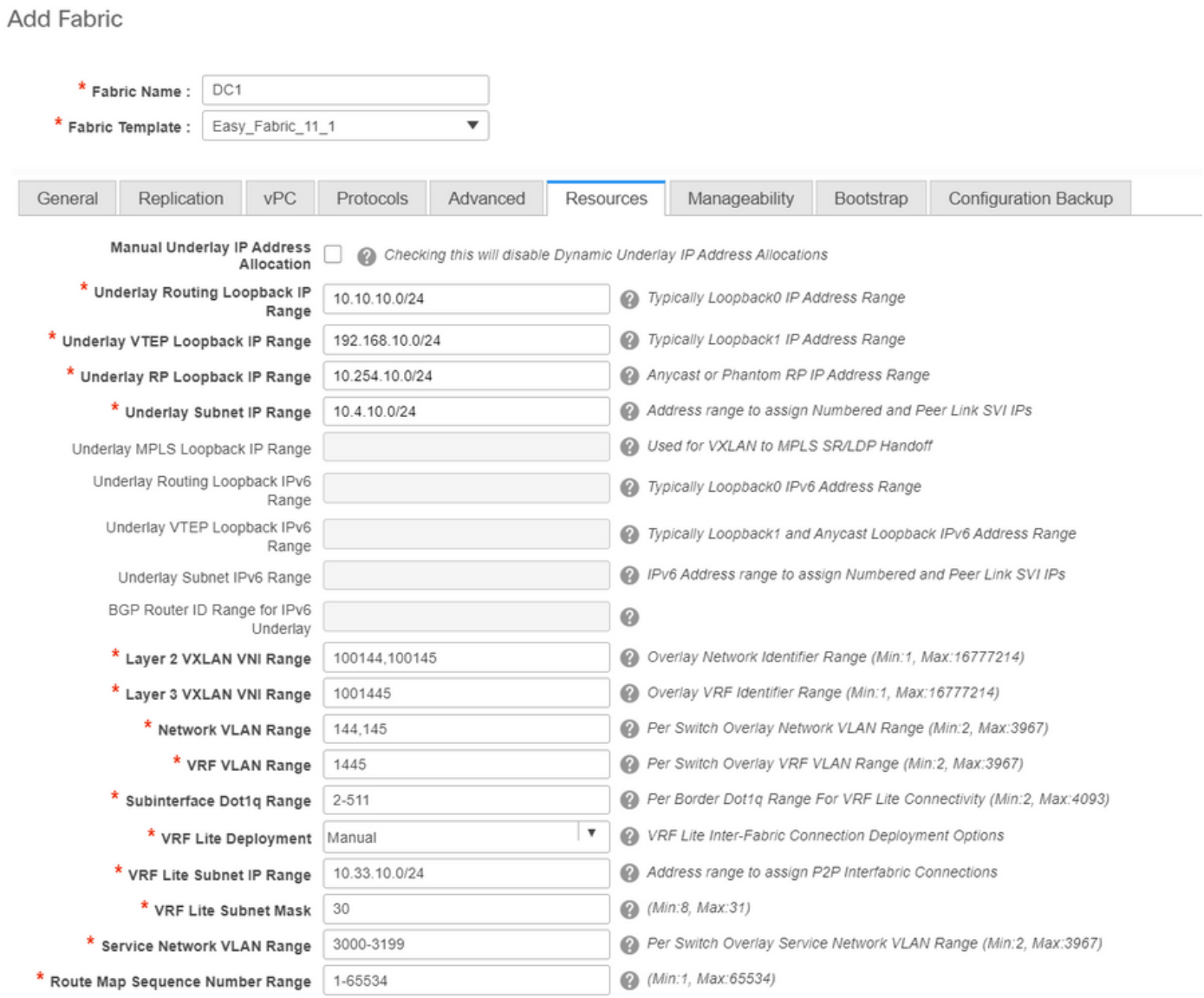
- 接下来是“高级”选项卡



#在本文档中，所有字段都保留为默认值。

ASN自动填充自“常规”选项卡中提供的ASN

- 接下来是填写“资源”选项卡中的字段



#底层路由环回IP范围是用于BGP、OSPF等协议的范围

#底层VTEP环回IP范围是将用于NVE接口的IP范围。

#底层RP用于BUM组播组的PIM RP。

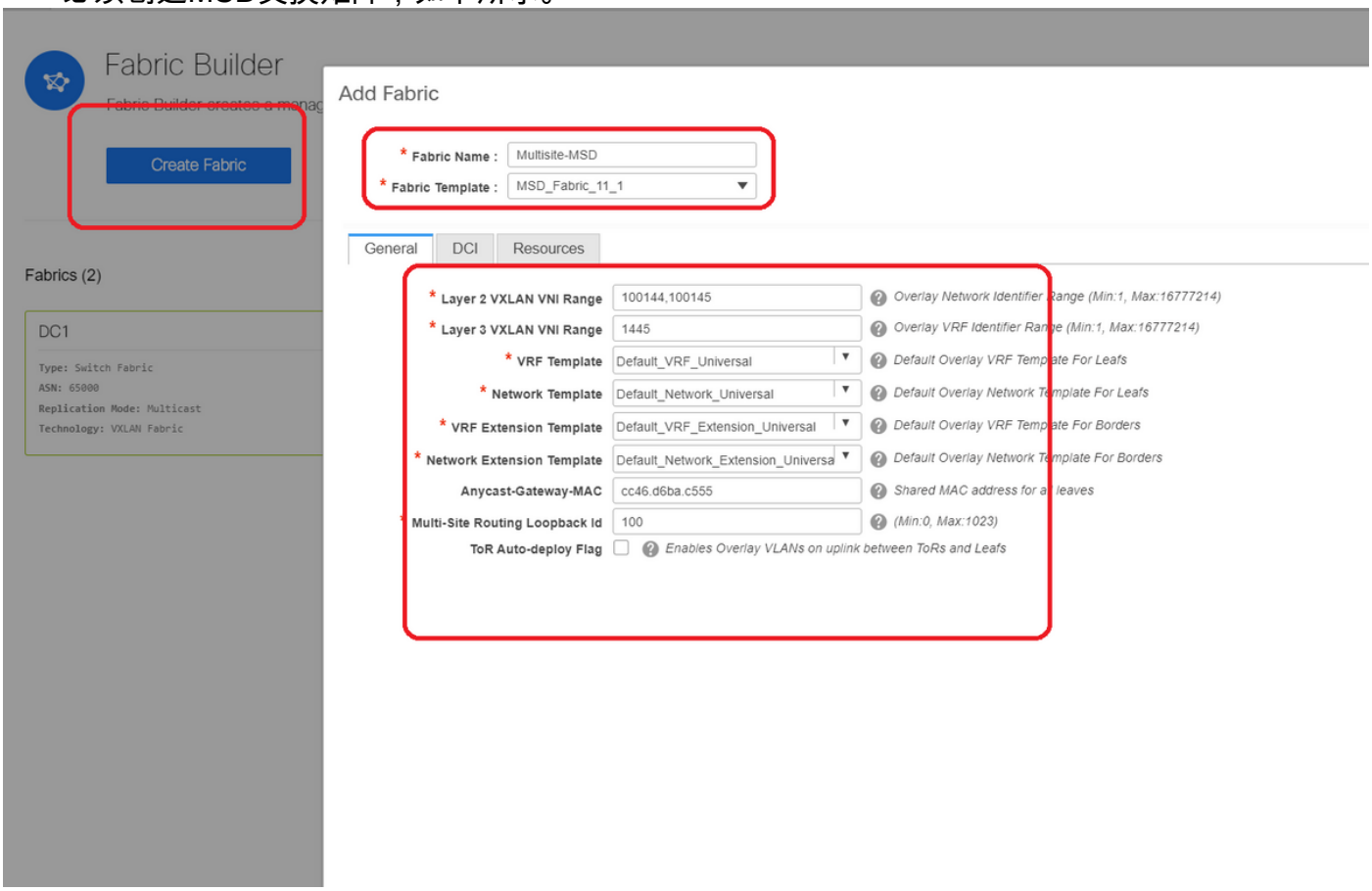
- 用相关信息填写其他选项卡，然后“保存”

步骤 2：为DC2创建简易交换矩阵

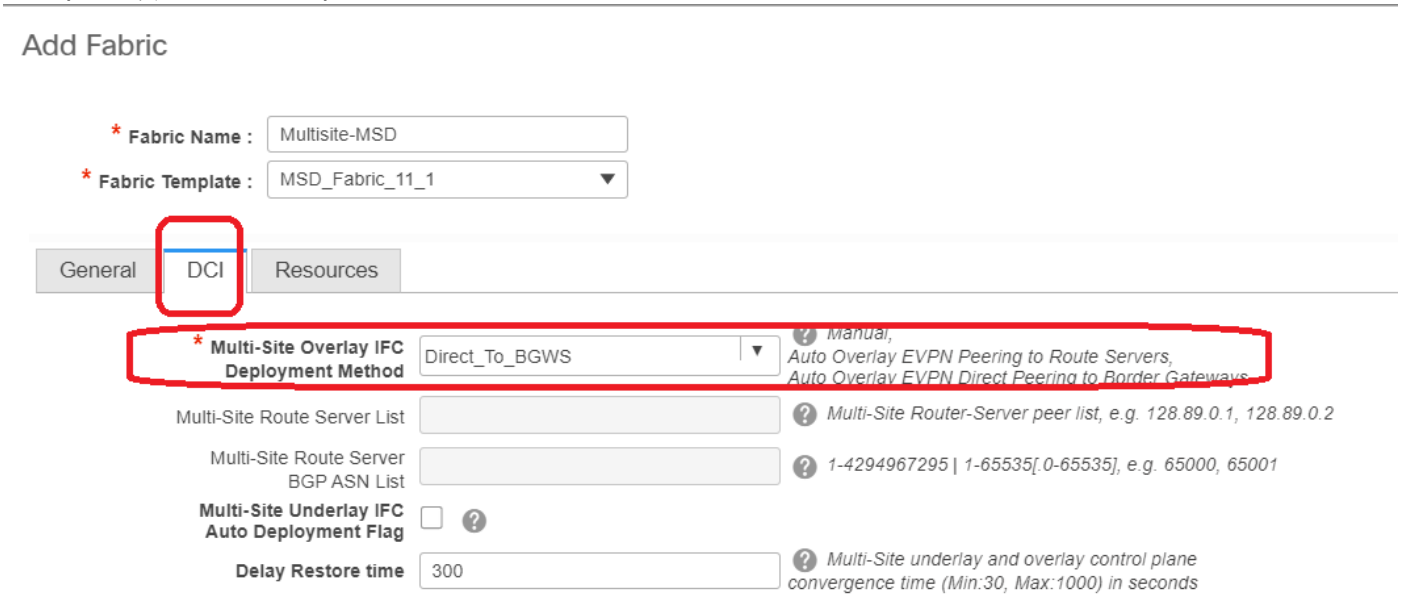
- 执行与步骤1相同的任务，为DC2创建Easy Fabric
- 确保在Resources for NVE and Routing Loopbacks（NVE和路由环回）和任何其他相关区域下提供不同的IP地址块
- ASN也应不同
- 第2层和第2层VNID相同

步骤 3：为多站点创建MSD

- 必须创建MSD交换矩阵，如下所示。

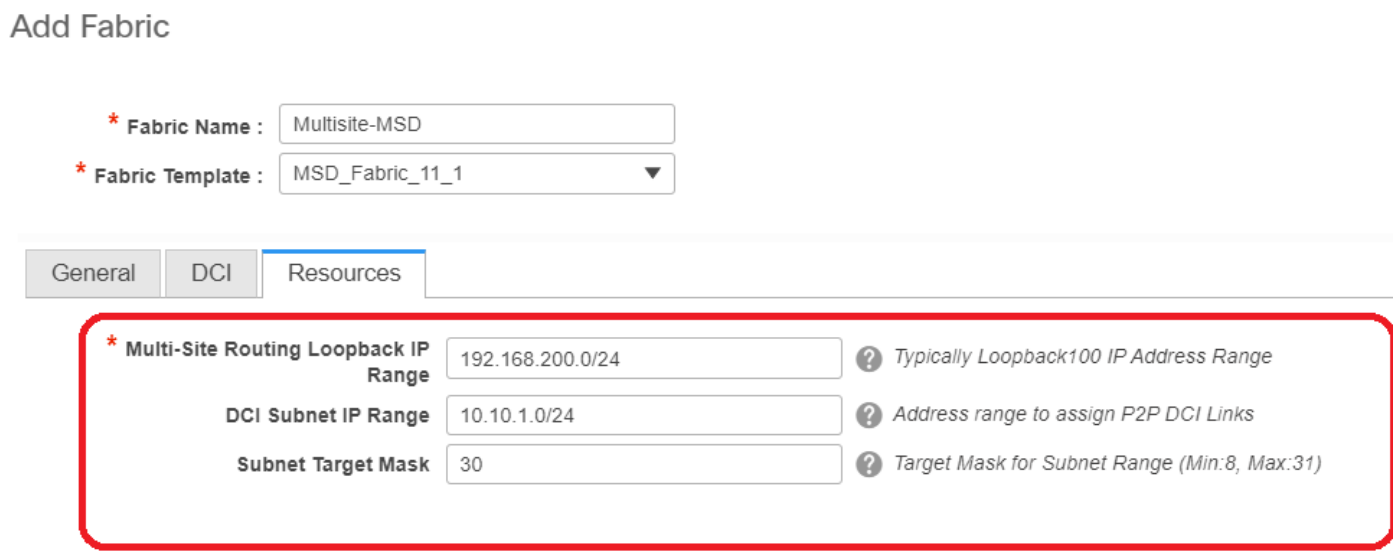


- 同时填写DCI选项卡



#多站点重叠IFC部署方法为“Direct_To_BGWS”，因为DC1-BGW将与DC2-BGW形成重叠连接。拓扑中显示的DCI交换机只是中转第3层设备（以及VRFLITE）

- 下一步是提及多站点环回范围(此IP地址将用作DC1和DC2 BGW上的多站点环回IP;DC1-BGW1和DC1-BGW2共享相同的多站点环回IP;DC2-BGW1和DC2-BGW2共享相同的多站点环回IP，但与DC1-BGW不同

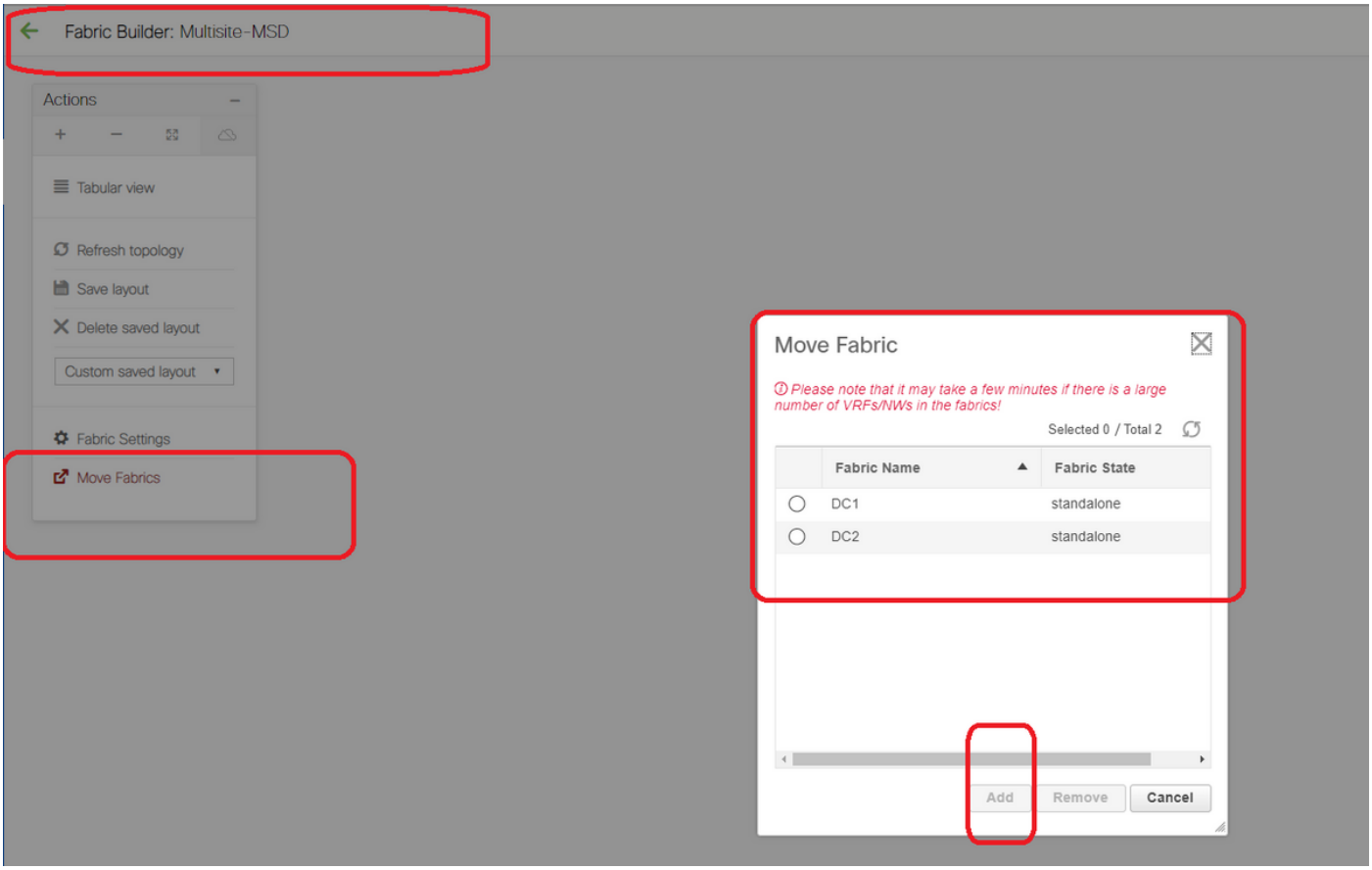


#填写字段后，点击“保存”。

#完成步骤1至3后，交换矩阵构建器页面如下所示。

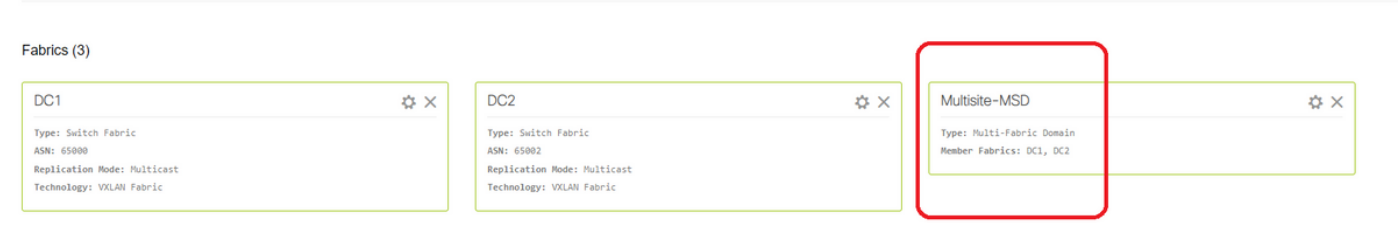
步骤 4：将DC1和DC2交换矩阵移入多站点MSD

#在此步骤中，DC1和DC2交换矩阵移至步骤3中创建的多站点MSD。以下是如何实现相同的屏幕截图。



#选择MSD，点击“移动交换矩阵”，然后依次选择DC1和DC2，然后选择“添加”。

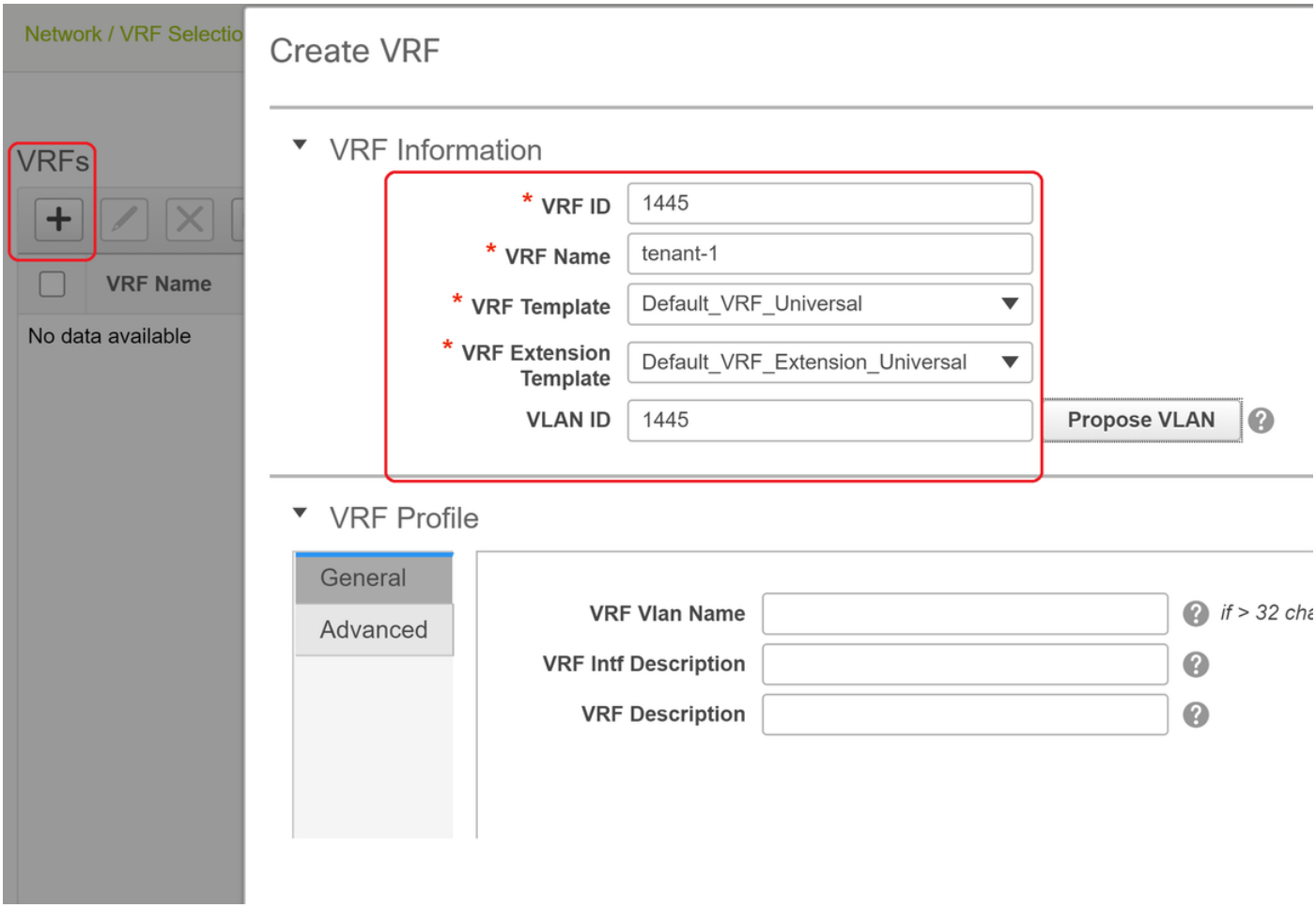
#移动两个结构后，主页将如下所示



#多站点 — MSD将DC1和DC2显示为成员结构

步骤 5：创建VRF

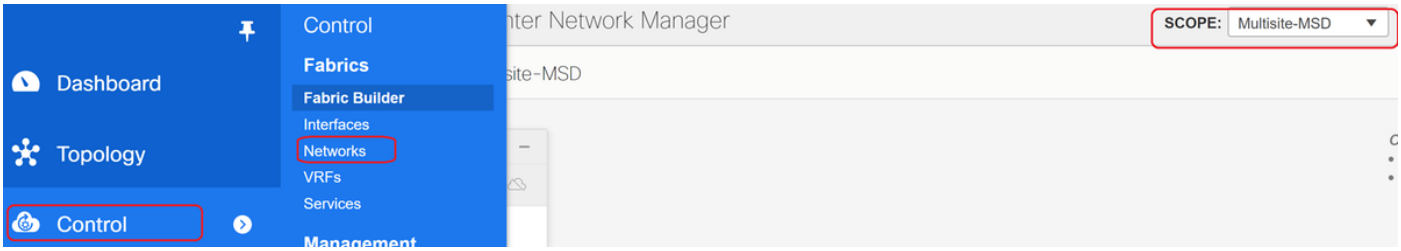
#可以从MSD交换矩阵创建VRF，这将适用于两个交换矩阵。下面是实现相同效果的屏幕截图。

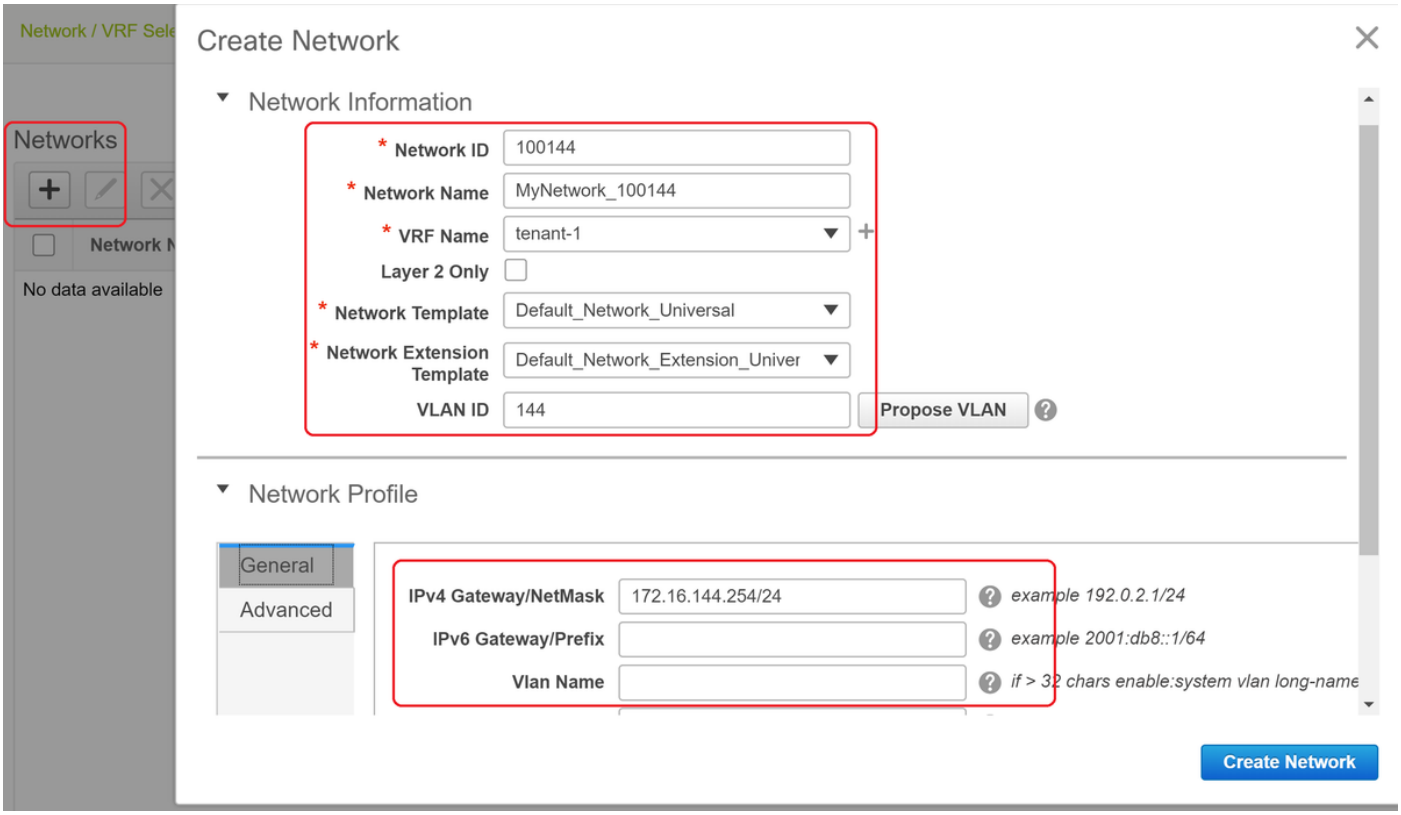


#填写高级选项卡，然后“创建”

步骤 6：创建网络

#创建VLAN和相应的VNID时，SVI可以从MSD交换矩阵中完成，这两个交换矩阵都适用。





#在“高级”选项卡中，如果BGW需要成为网络网关，请启用复选框

#填写所有字段后，点击“创建网络”

#对任何其他VLAN/网络重复相同步骤

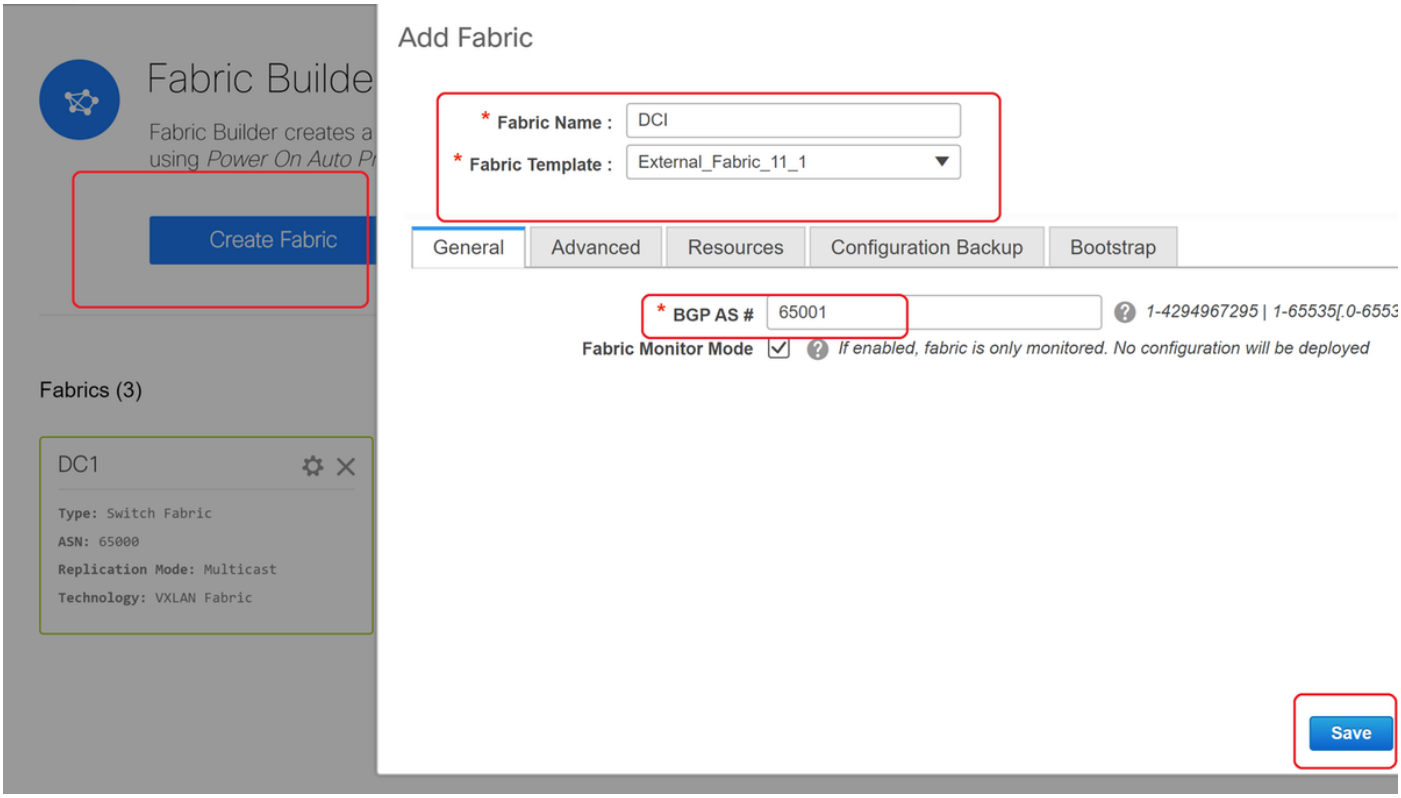
步骤 7：为DCI交换机创建外部交换矩阵

#本示例考虑了数据包从DC1到DC2（就站点间通信而言）路径中的DCI交换机，当有2个以上交换矩阵时，通常会看到这些交换机。

#外部交换矩阵将包括位于本文档开头所示拓扑顶部的两个DCI交换机

#使用“外部”模板创建交换矩阵并指定ASN

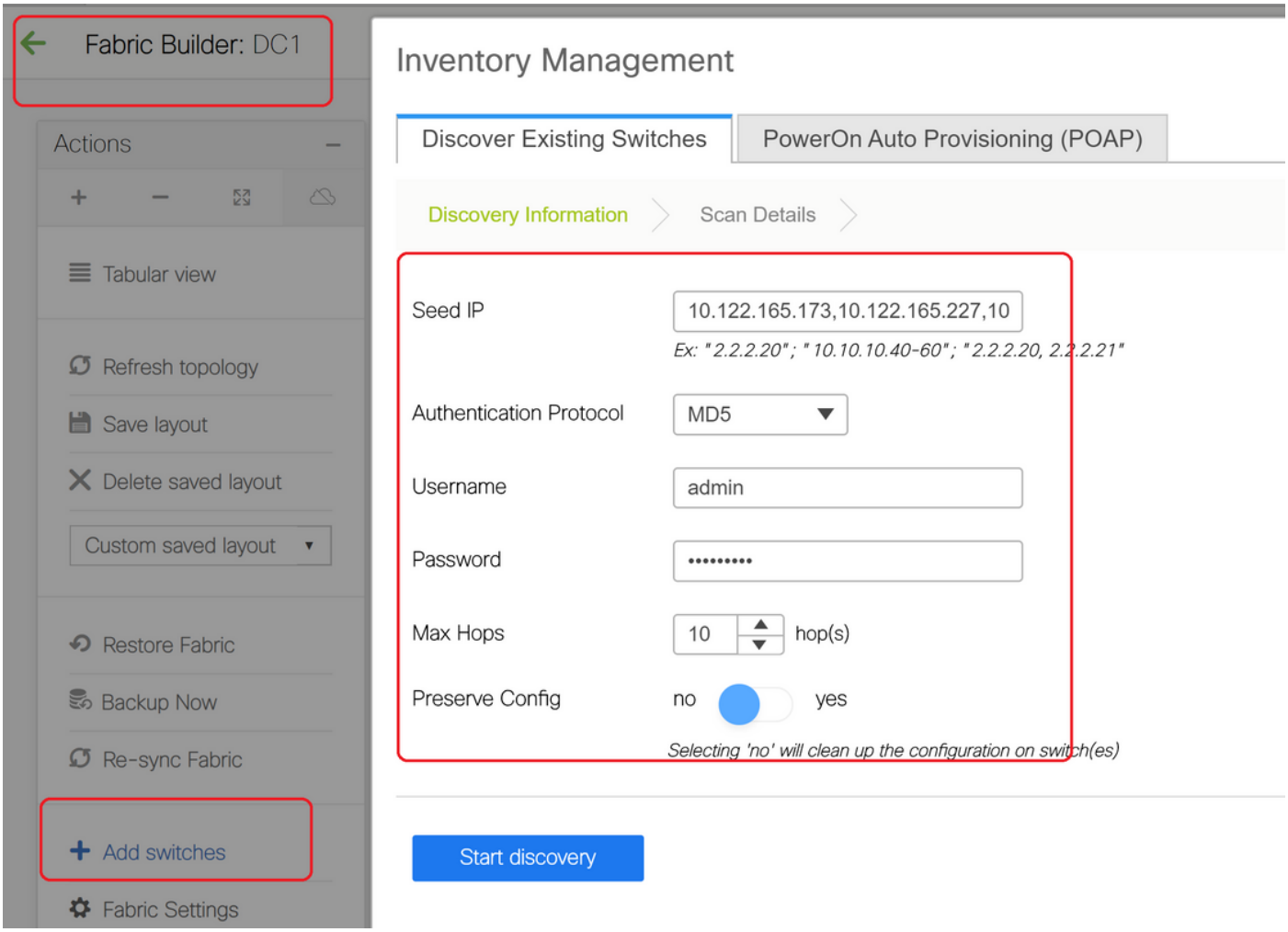
#修改部署的任何其他相关字段



步骤 8::将交换机添加到每个交换矩阵

#在此，每个交换矩阵的所有交换机都将添加到相应交换矩阵。

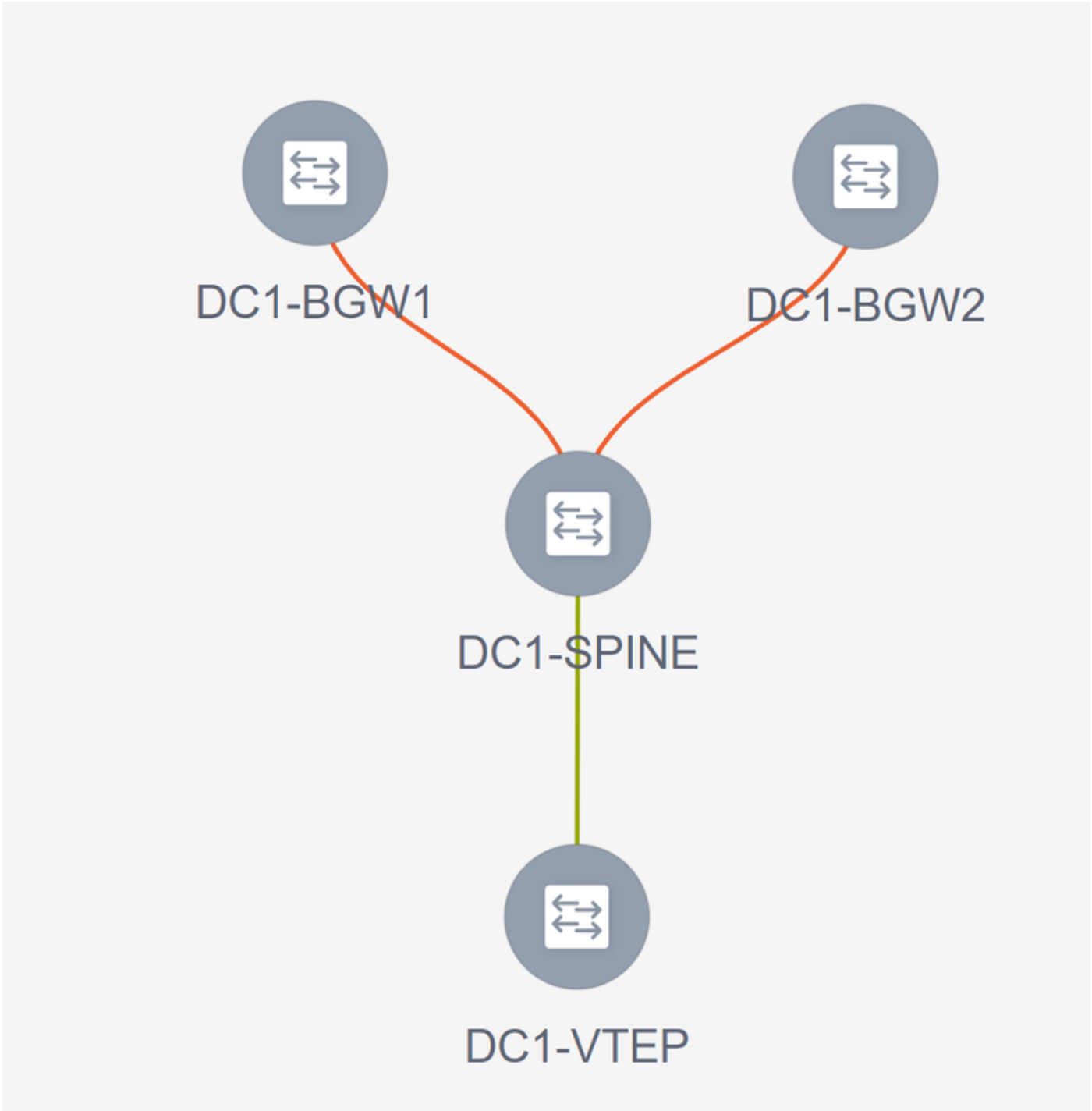
添加交换机的步骤如下屏幕截图所示。

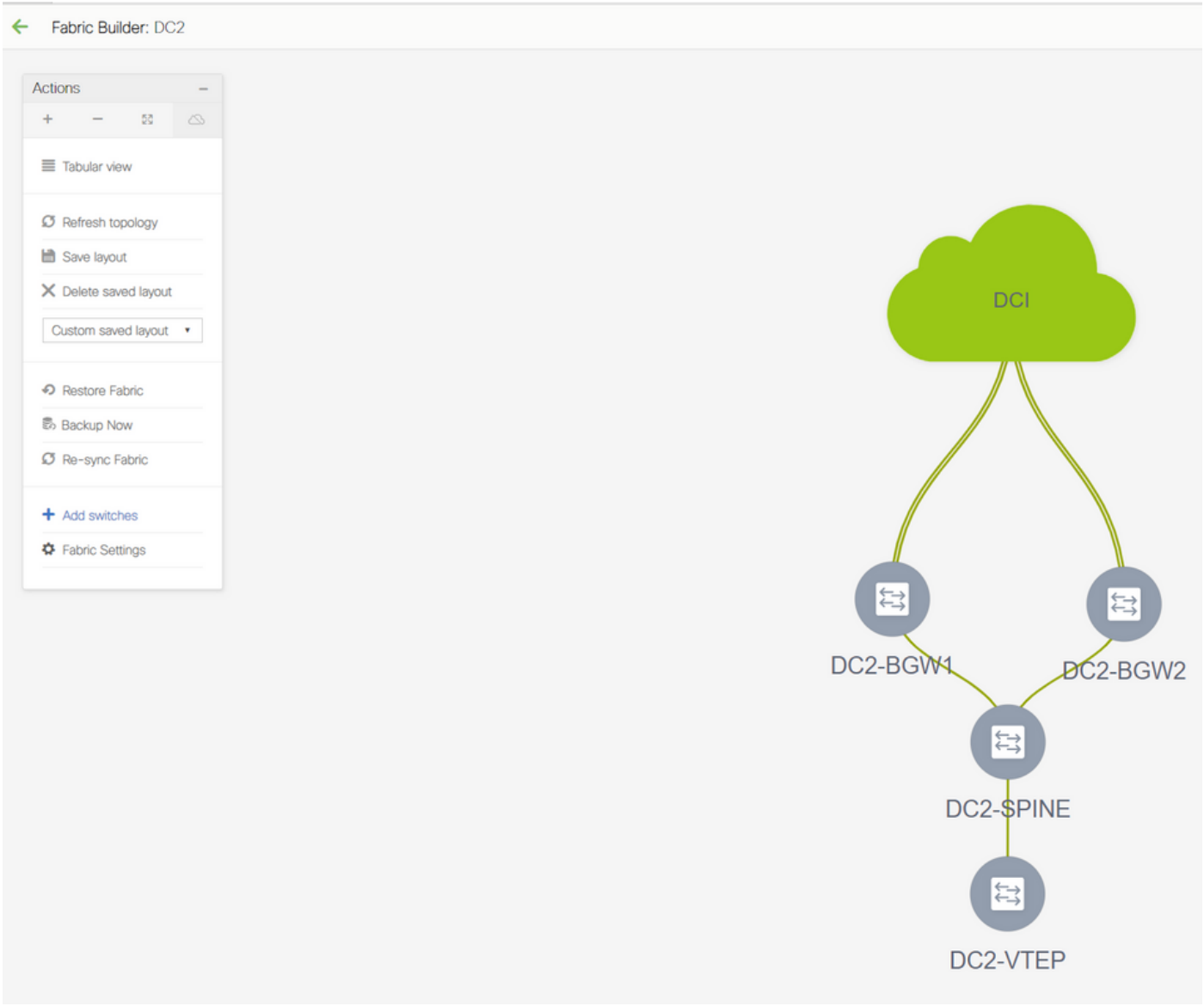


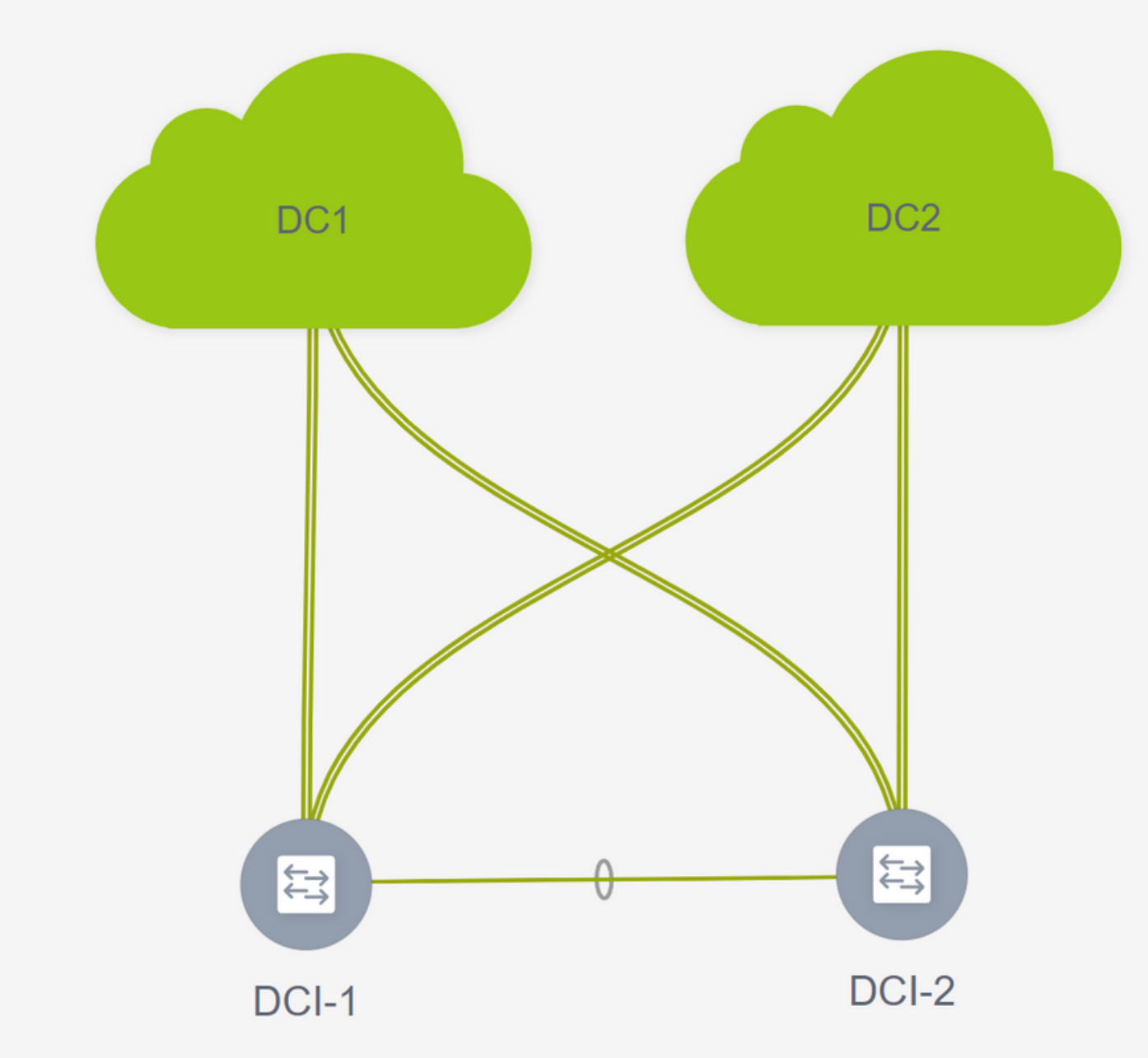
#如果“保留配置”为“否”；现有的任何交换机配置都将被清除；VRF环境管理中的主机名、引导变量、MGMT0 IP地址、路由是例外

#正确设置交换机上的角色(通过右键单击交换机、设置角色和相关角色

#还会相应地排列交换机的布局，然后单击“保存布局”

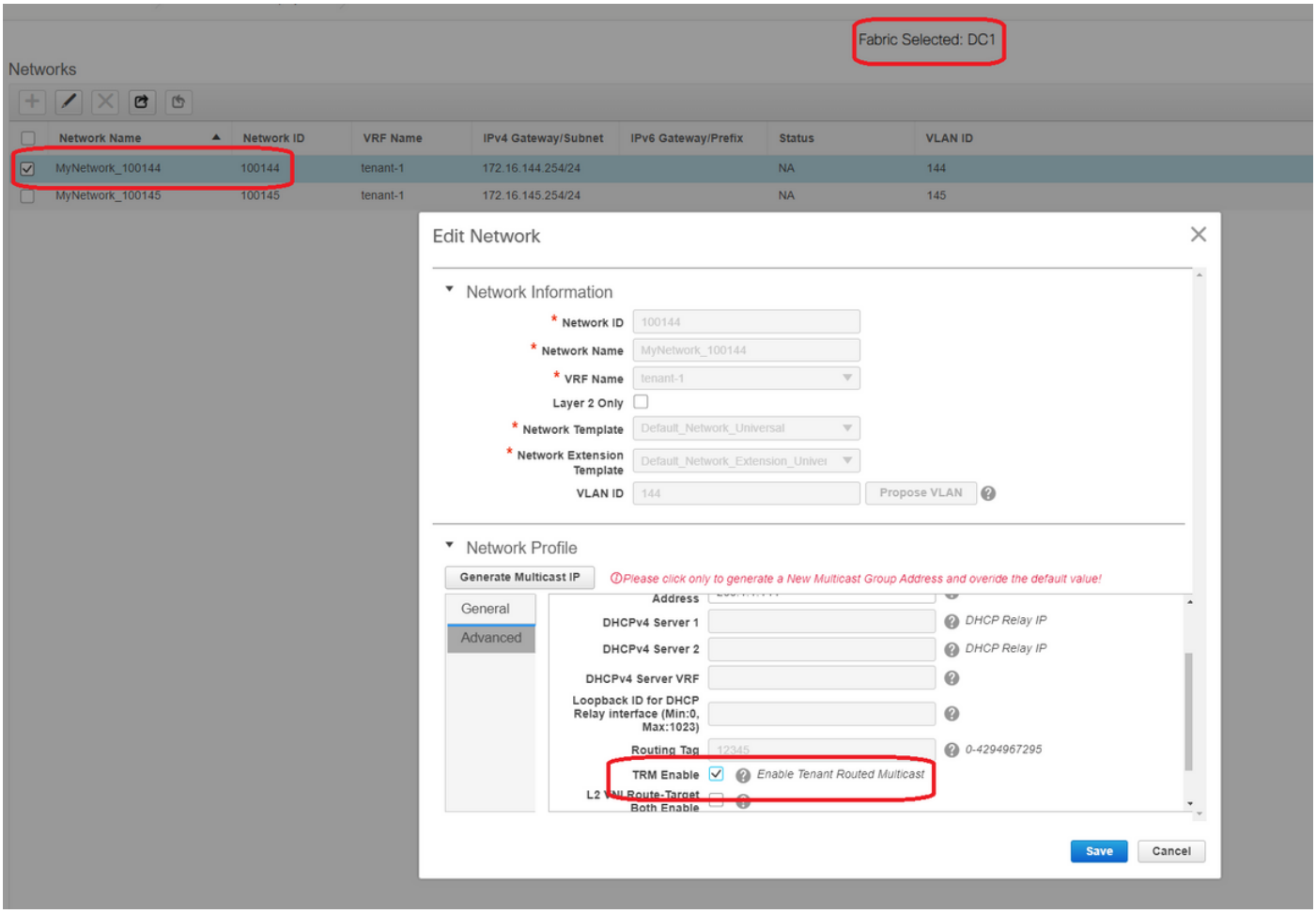

DC1-BGW1
DC1-BGW2
DC1-SPINE
DC1-VTEP


Fabric Builder: DC2
Actions
Tabular view
Refresh topology
Save layout
Delete saved layout
Custom saved layout
Restore Fabric
Backup Now
Re-sync Fabric
Add switches
Fabric Settings
DCI
DC2-BGW1
DC2-BGW2
DC2-SPINE
DC2-VTEP



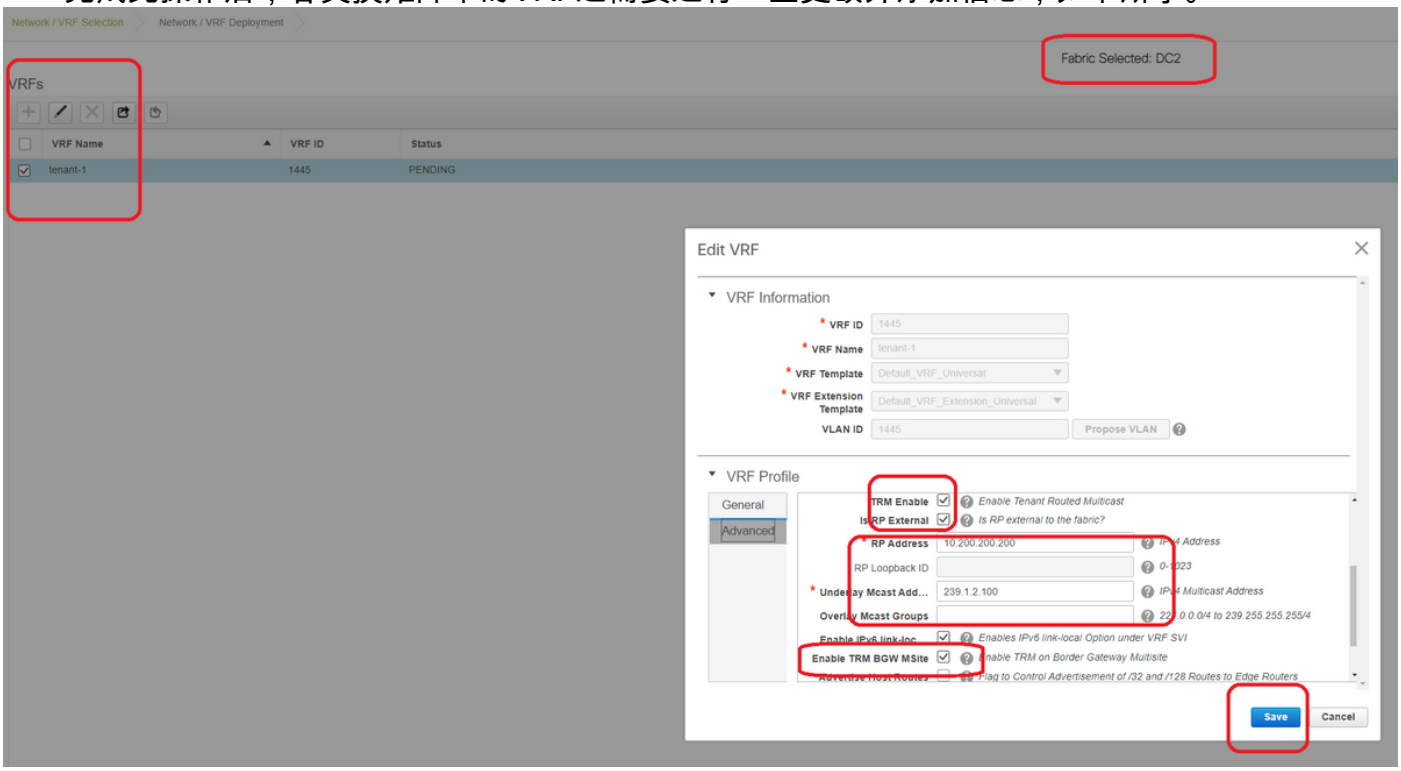
步骤 9：单个交换矩阵的TRM设置

- 下一步是在每个交换矩阵上启用TRM复选框



#对所有交换矩阵的所有网络执行此步骤。

- 完成此操作后，各交换矩阵中的VRF还需要进行一些更改并添加信息，如下所示。



#这必须在DC1和DC2以及VRF部分中完成。

#请注意，VRF-> 239.1.2.100的组播组已手动从自动填充的组更改；最佳实践是为第3层VNI VRF和

任何第2层VNI VLAN的BUM流量组播组使用不同的组

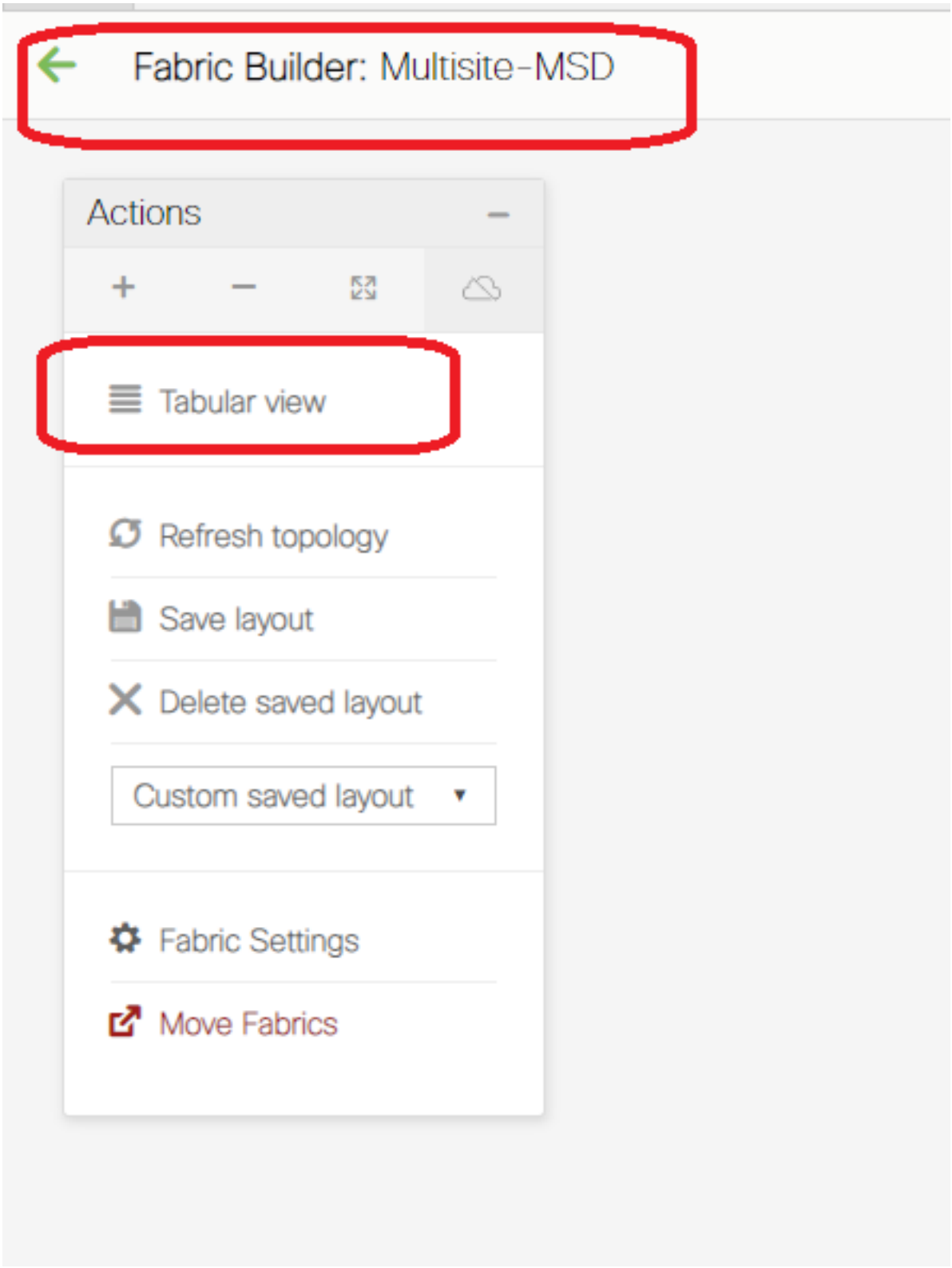
步骤 10：边界网关上的VRFLITE配置

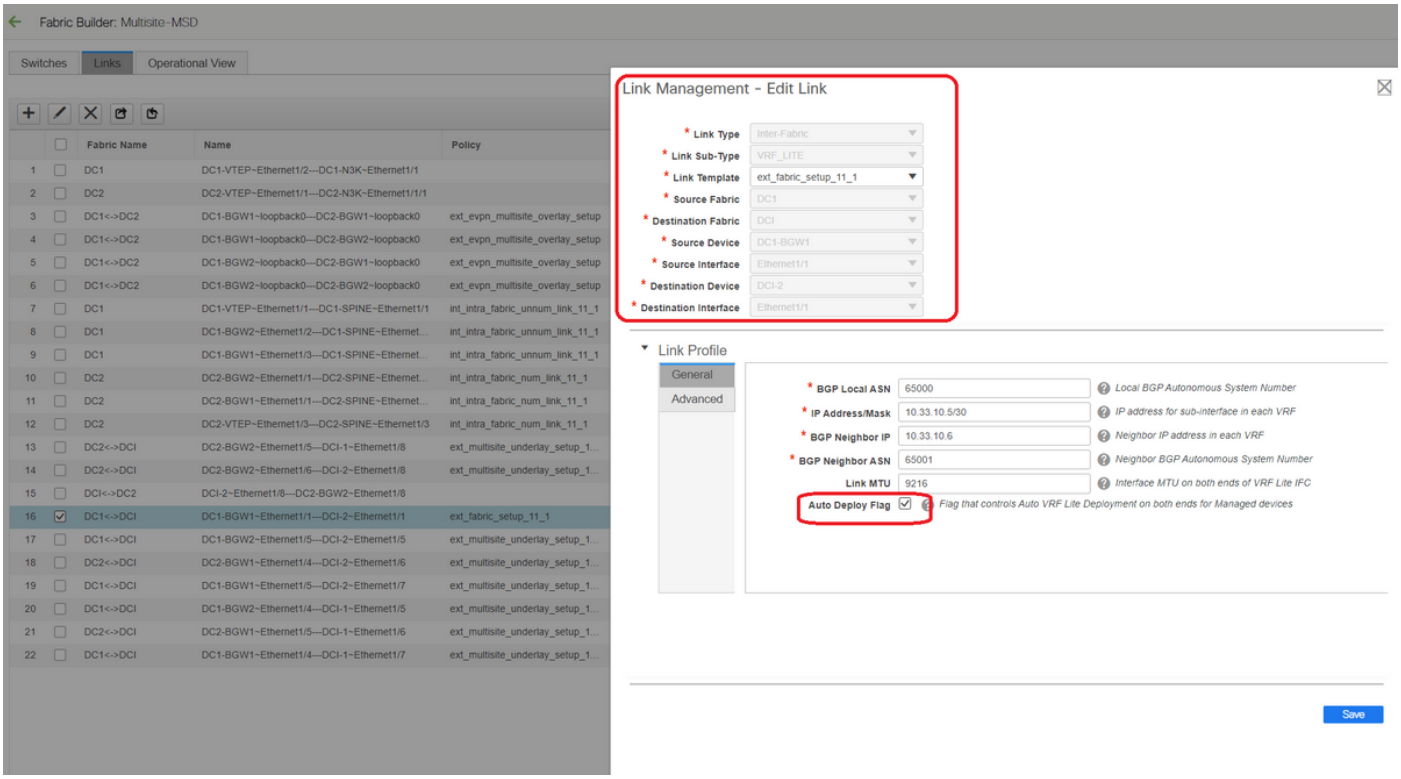
#从NXOS 9.3(3)和DCNM 11.3(1)开始，边界网关可以充当边界网关和VRFLITE连接点（这将使边界网关与外部路由器具有VRFLITE邻居关系，因此外部设备可以与交换矩阵中的设备通信）

#在本文档中，边界网关与DCI路由器（位于上图所示拓扑的北部）形成VRFLITE邻居关系。

#有一点需要注意：VRFLITE和多站点底层链路不能是同一物理链路。必须将单独的链路向上旋转，以形成虚拟站点和多站点底层

#下面的屏幕截图将说明如何在边界网关上实现VRF LITE和多站点扩展。


Fabric Builder: Multisite-MSD
Actions
Tabular view
Refresh topology
Save layout
Delete saved layout
Custom saved layout
Fabric Settings
Move Fabrics



#切换到“表格视图”

#移至选项卡“links”，然后添加“交换矩阵间VRFLITE”链接，并且必须将源交换矩阵指定为DC1，将目标交换矩阵指定为DCI

#为指向正确DCI交换机的源接口选择正确的接口

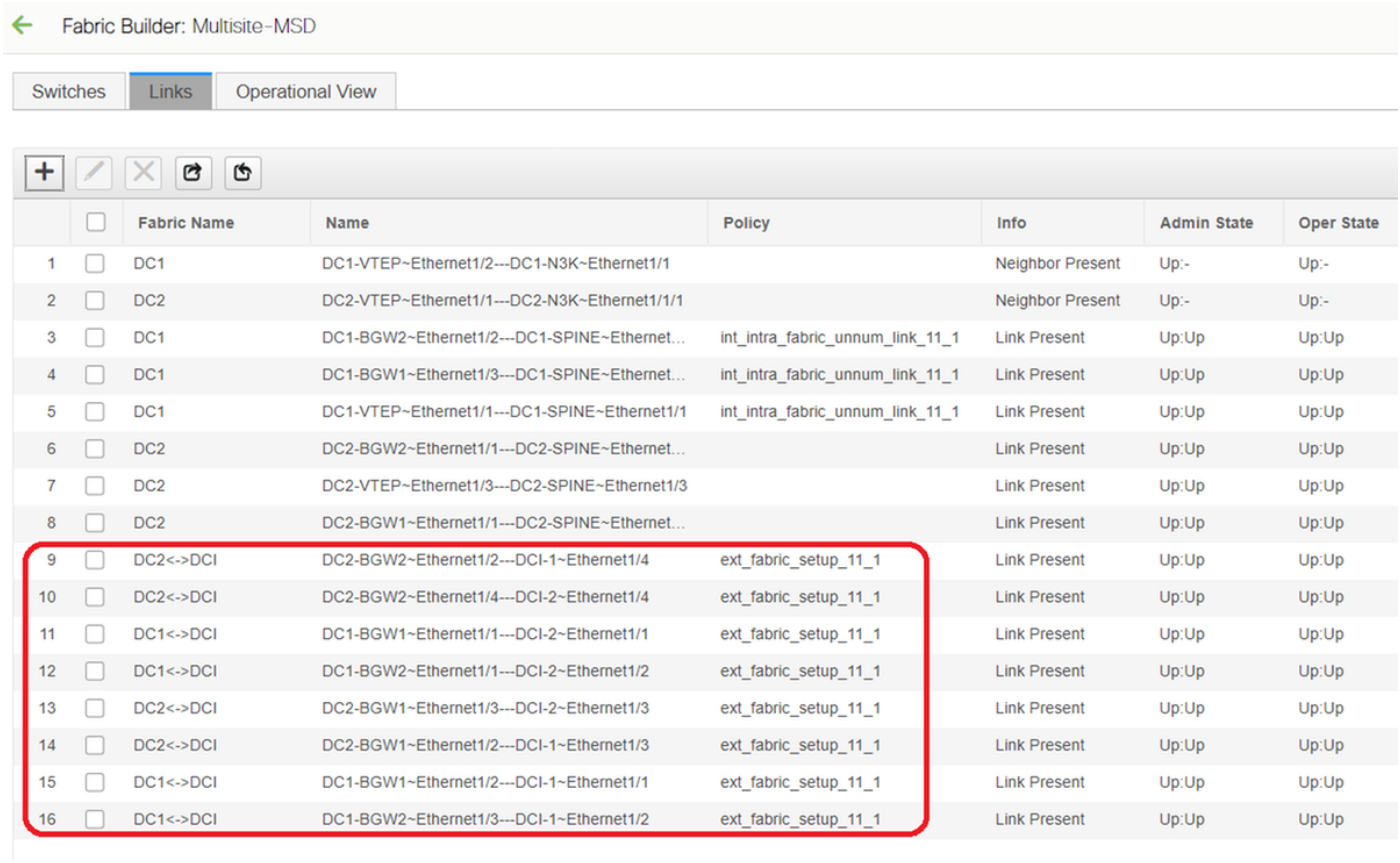
#在链路配置文件下，提供本地和远程IP地址

#还启用复选框 — “自动部署标志”，以便VRFLITE的DCI交换机配置也自动填充（这在以后的步骤中完成）

自动填充ASN数量

#填写所有字段并填写正确信息后，单击“保存”按钮

- 上述步骤必须对指向两台DCI交换机的所有4个边界网关上的所有BGW到DCI连接执行。
- 考虑到本文档的拓扑，共有8个交换矩阵间VRF LITE连接，如下所示。



		Fabric Name	Name	Policy	Info	Admin State	Oper State
1		DC1	DC1-VTEP~Ethernet1/2---DC1-N3K~Ethernet1/1		Neighbor Present	Up:-	Up:-
2		DC2	DC2-VTEP~Ethernet1/1---DC2-N3K~Ethernet1/1/1		Neighbor Present	Up:-	Up:-
3		DC1	DC1-BGW2~Ethernet1/2---DC1-SPINE~Ethernet...	int_intra_fabric_unnum_link_11_1	Link Present	Up:Up	Up:Up
4		DC1	DC1-BGW1~Ethernet1/3---DC1-SPINE~Ethernet...	int_intra_fabric_unnum_link_11_1	Link Present	Up:Up	Up:Up
5		DC1	DC1-VTEP~Ethernet1/1---DC1-SPINE~Ethernet1/1	int_intra_fabric_unnum_link_11_1	Link Present	Up:Up	Up:Up
6		DC2	DC2-BGW2~Ethernet1/1---DC2-SPINE~Ethernet...		Link Present	Up:Up	Up:Up
7		DC2	DC2-VTEP~Ethernet1/3---DC2-SPINE~Ethernet1/3		Link Present	Up:Up	Up:Up
8		DC2	DC2-BGW1~Ethernet1/1---DC2-SPINE~Ethernet...		Link Present	Up:Up	Up:Up
9		DC2<->DCI	DC2-BGW2~Ethernet1/2---DCI-1~Ethernet1/4	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
10		DC2<->DCI	DC2-BGW2~Ethernet1/4---DCI-2~Ethernet1/4	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
11		DC1<->DCI	DC1-BGW1~Ethernet1/1---DCI-2~Ethernet1/1	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
12		DC1<->DCI	DC1-BGW2~Ethernet1/1---DCI-2~Ethernet1/2	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
13		DC2<->DCI	DC2-BGW1~Ethernet1/3---DCI-2~Ethernet1/3	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
14		DC2<->DCI	DC2-BGW1~Ethernet1/2---DCI-1~Ethernet1/3	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
15		DC1<->DCI	DC1-BGW1~Ethernet1/2---DCI-1~Ethernet1/1	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up
16		DC1<->DCI	DC1-BGW2~Ethernet1/3---DCI-1~Ethernet1/2	ext_fabric_setup_11_1	Link Present	Up:Up	Up:Up

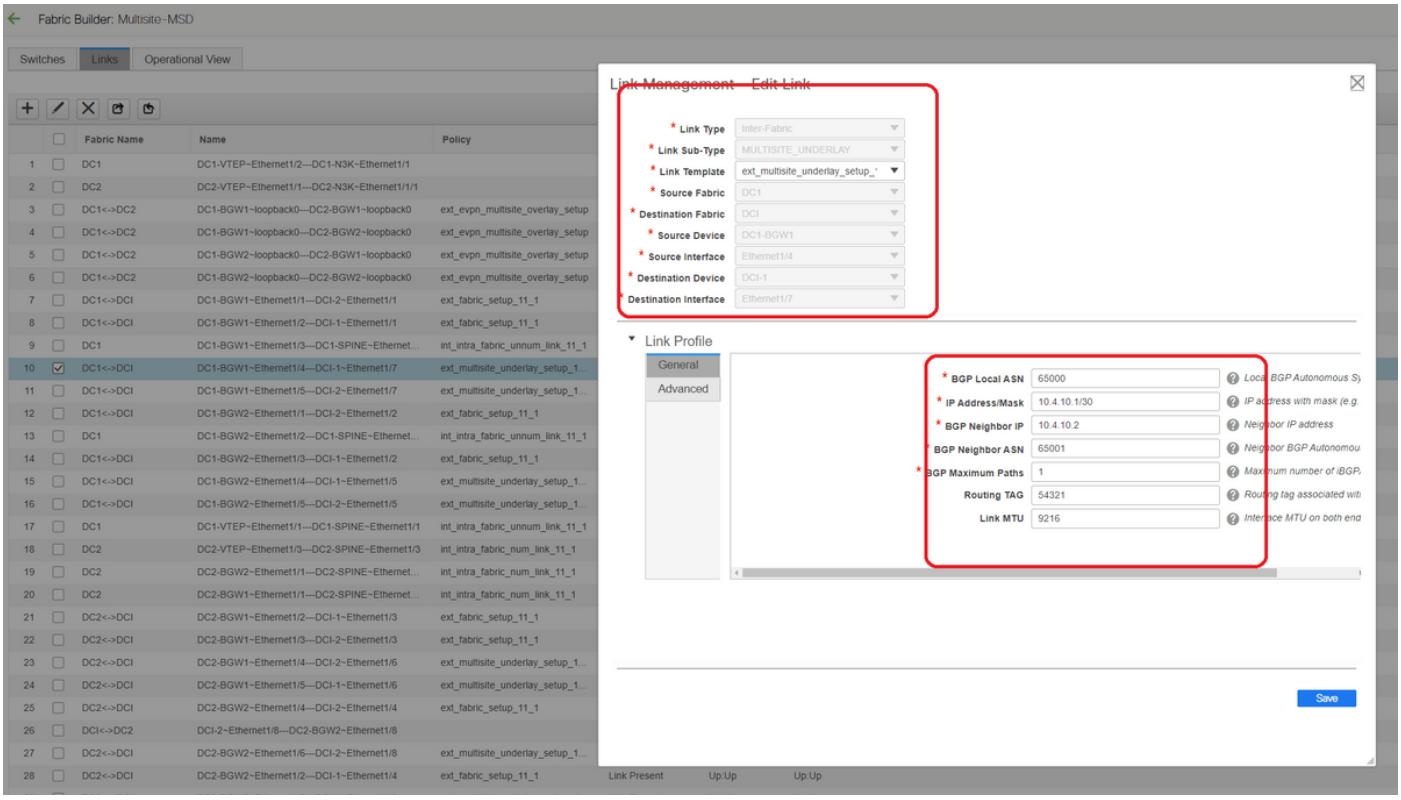
步骤 11：边界网关上的多站点底层配置

#下一步是在每个交换矩阵的每个边界网关上配置多站点底层。

#为此，我们需要从BGW到DCI交换机的单独物理链路。第10步中用于VRFLITE的链路不能用于多站点重叠

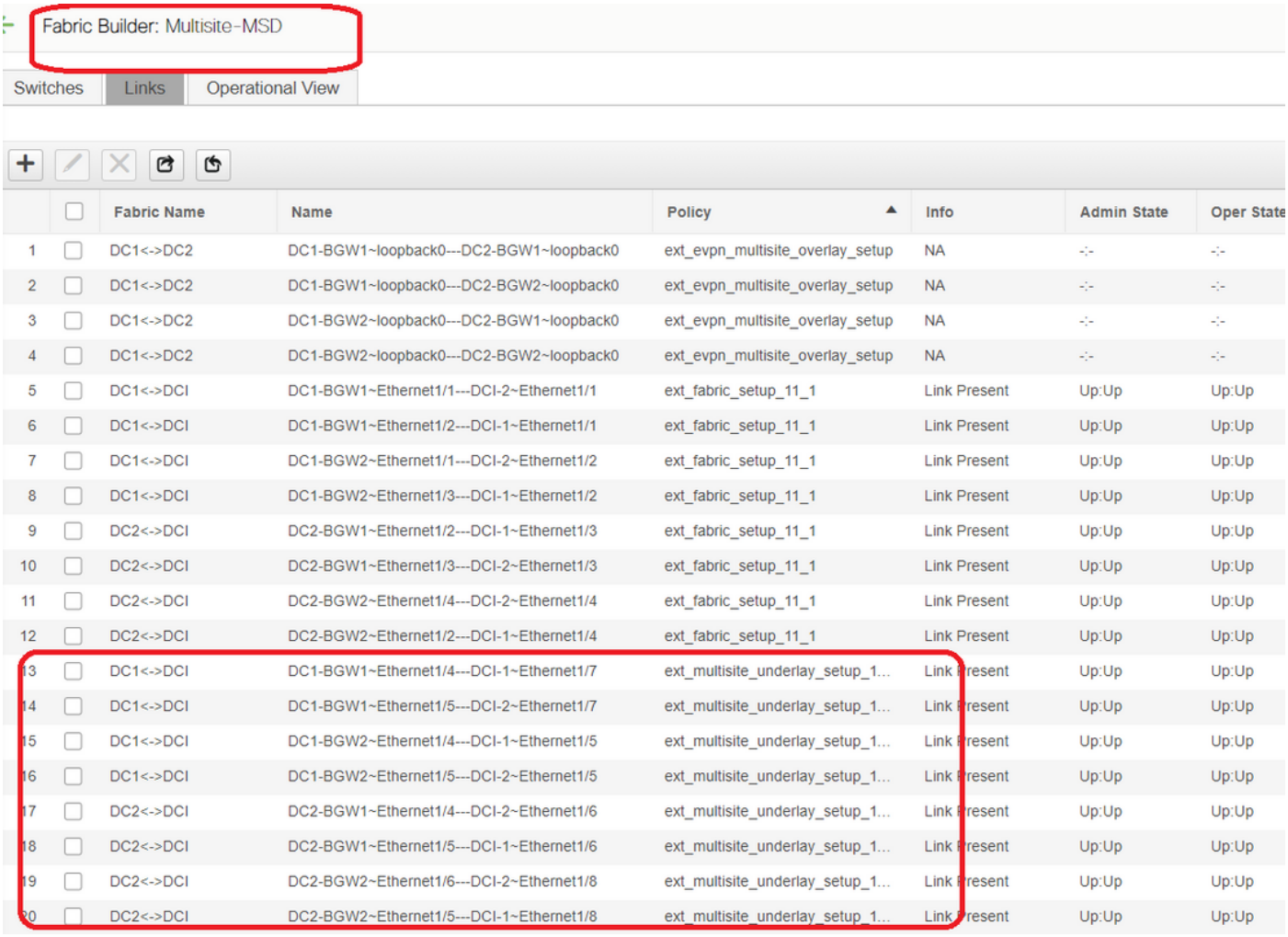
#这些接口将是“默认vrf”的一部分，而不是前一个接口将是租户vrf的一部分（本例中为tenant-1）

#下面的屏幕截图将帮助您完成执行此配置的步骤。



#必须对从BGW到DCI交换机的所有连接执行相同的步骤

#最后，共8个交换矩阵间多站点底层连接将如下所示。

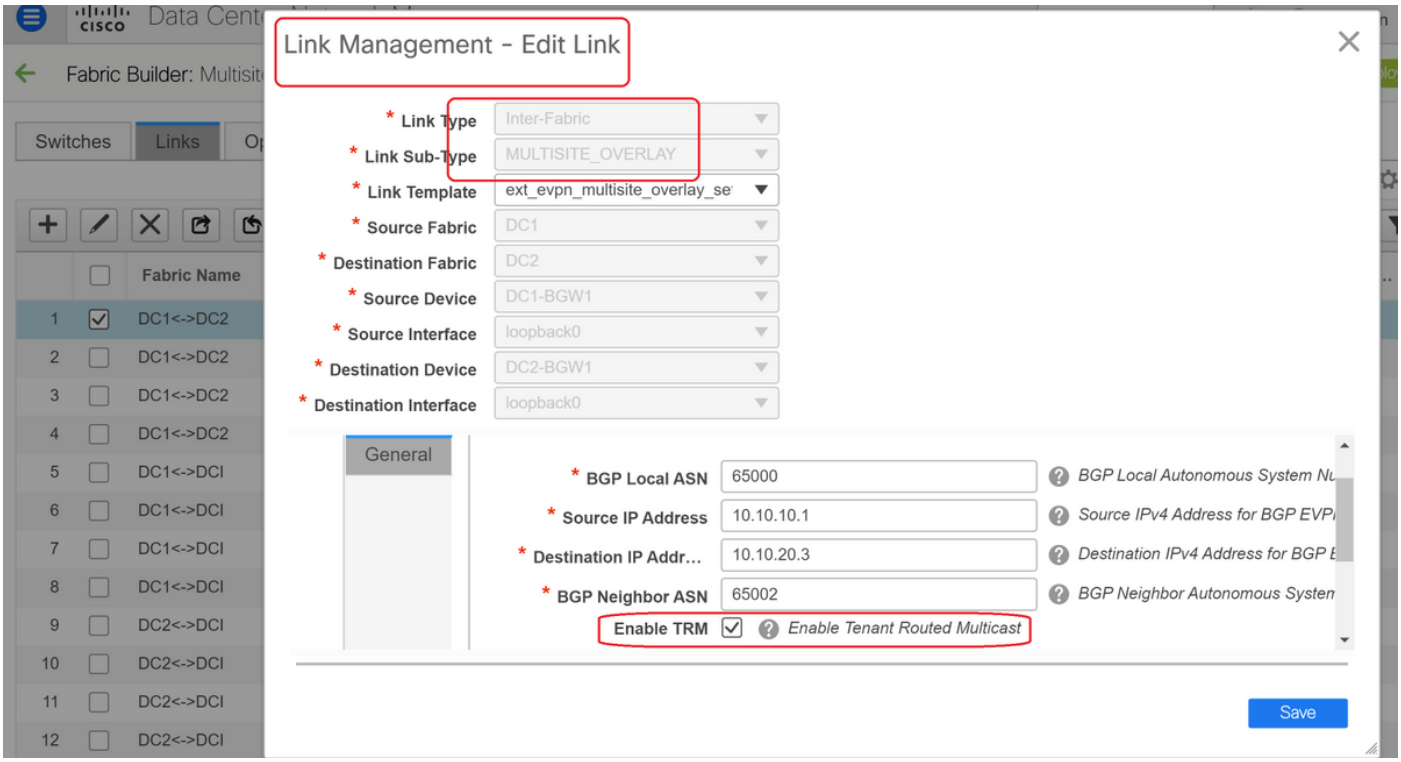


		Fabric Name	Name	Policy	Info	Admin State	Oper State
1		DC1<->DC2	DC1-BGW1~loopback0---DC2-BGW1~loopback0	ext_evpn_multisite_overlay_setup	NA	-:-	-:-
2		DC1<->DC2	DC1-BGW1~loopback0---DC2-BGW2~loopback0	ext_evpn_multisite_overlay_setup	NA	-:-	-:-
3		DC1<->DC2	DC1-BGW2~loopback0---DC2-BGW1~loopback0	ext_evpn_multisite_overlay_setup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

步骤 12：TRM的多站点重叠设置

#当多站点底层完成后，多站点重叠接口/链路将自动填充，并可在多站点MSD交换矩阵内的链路下的表格视图中看到。

#默认情况下，多站点重叠将仅形成从每个站点BGW到另一个站点的bgp l2vpn evpn邻居关系，这是从一个站点到另一个站点的单播通信所必需的。但是，当需要在站点（通过vxlan多站点功能连接）之间运行组播时，需要启用TRM复选框，如下所示，适用于多站点MSD交换矩阵内的所有重叠接口。屏幕截图将说明如何执行此操作。



步骤 13：在MSD和单个交换矩阵中保存/部署

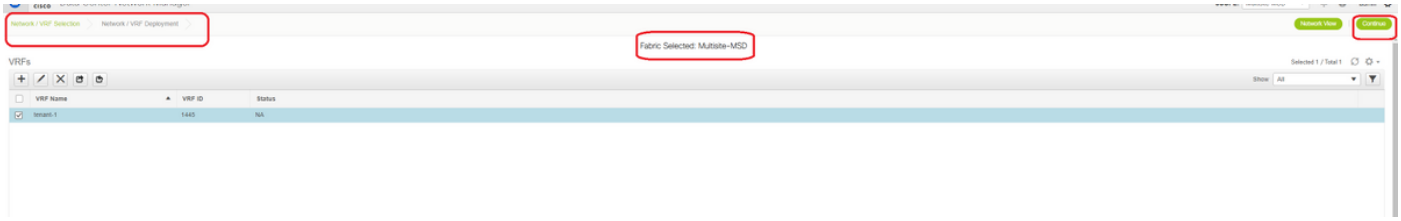
#执行保存/部署，按照上述步骤推送相关配置

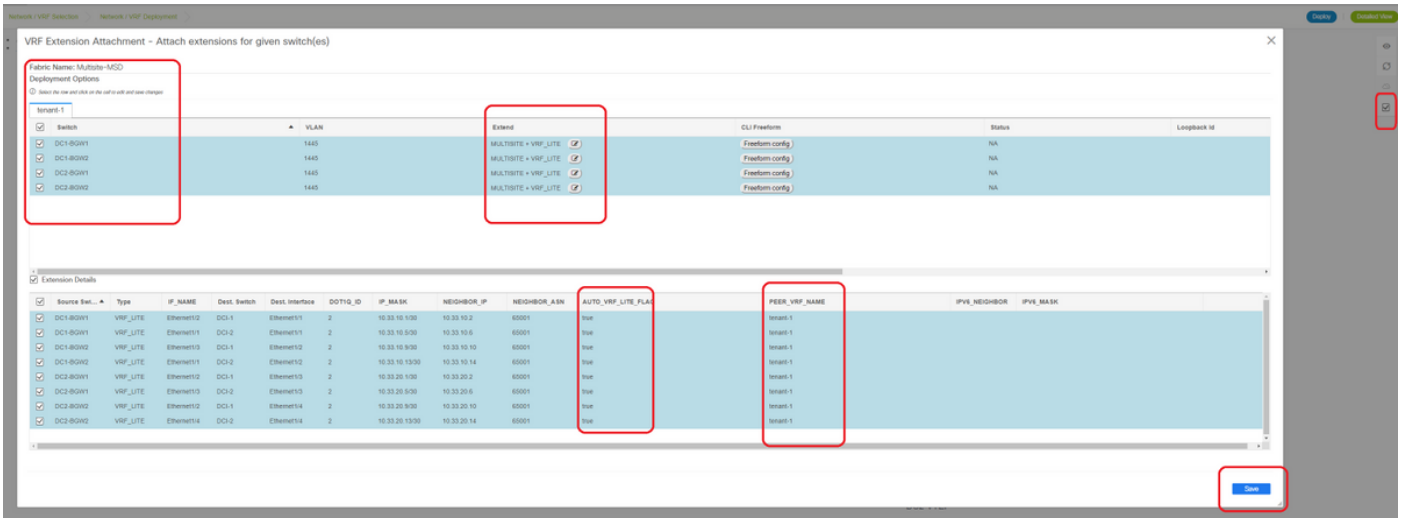
#选择MSD时，将推送的配置将仅适用于边界网关。

#因此，需要为各个交换矩阵保存/部署，这会将相关配置推送到所有常规枝叶交换机/VTEP

步骤 14：MSD的VRF扩展附件

#选择MSD并转至VRF部分



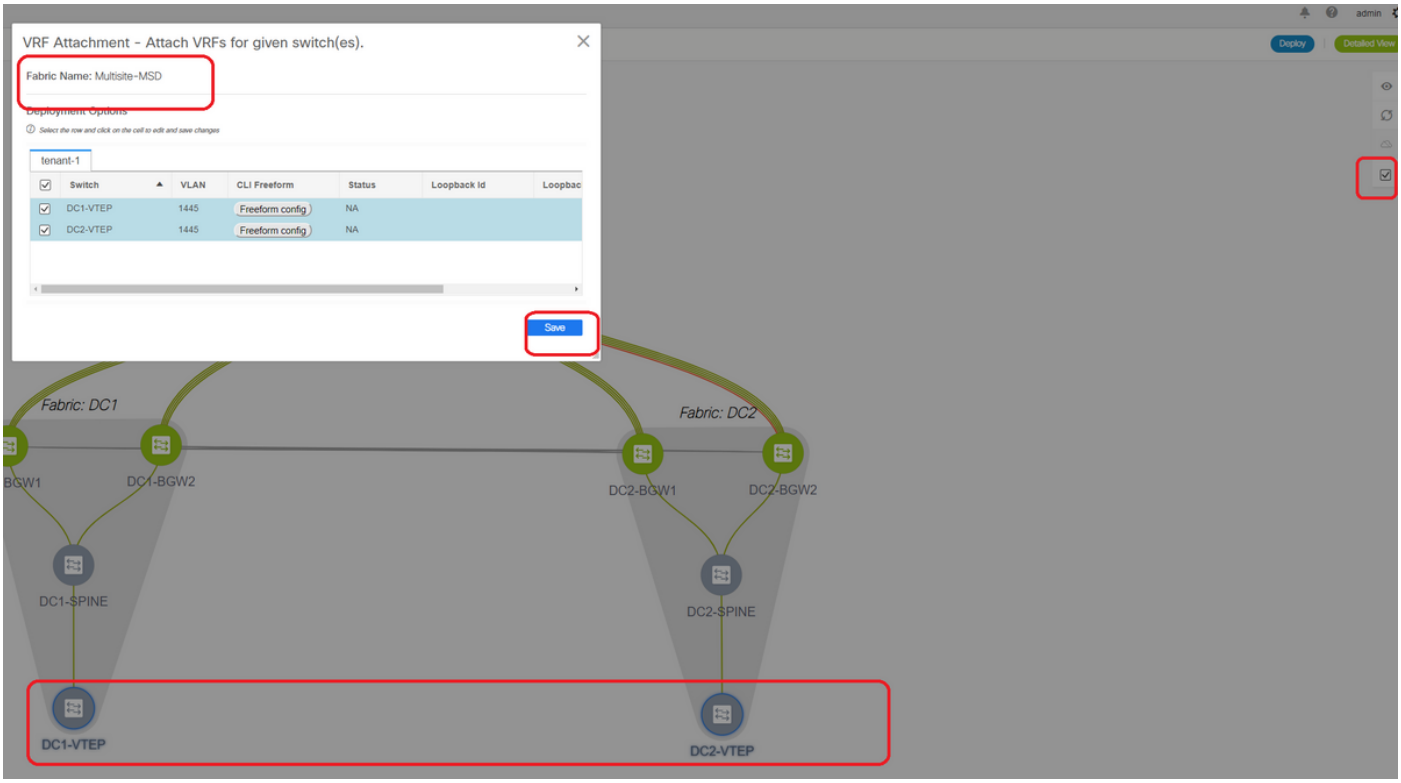


#请注意，扩展选项必须是“MULTISITE+VRF_LITE”，如本文档所示，边界网关功能和VRFLITE已集成到边界网关交换机。

AUTO_VRF_LITE将设置为true

#从BGW到DCI交换机的所有8个VRF名称必须手动填充（此处，示例在DCI交换机上使用相同的VRF名称），如下所示

#完成后，点击“保存”



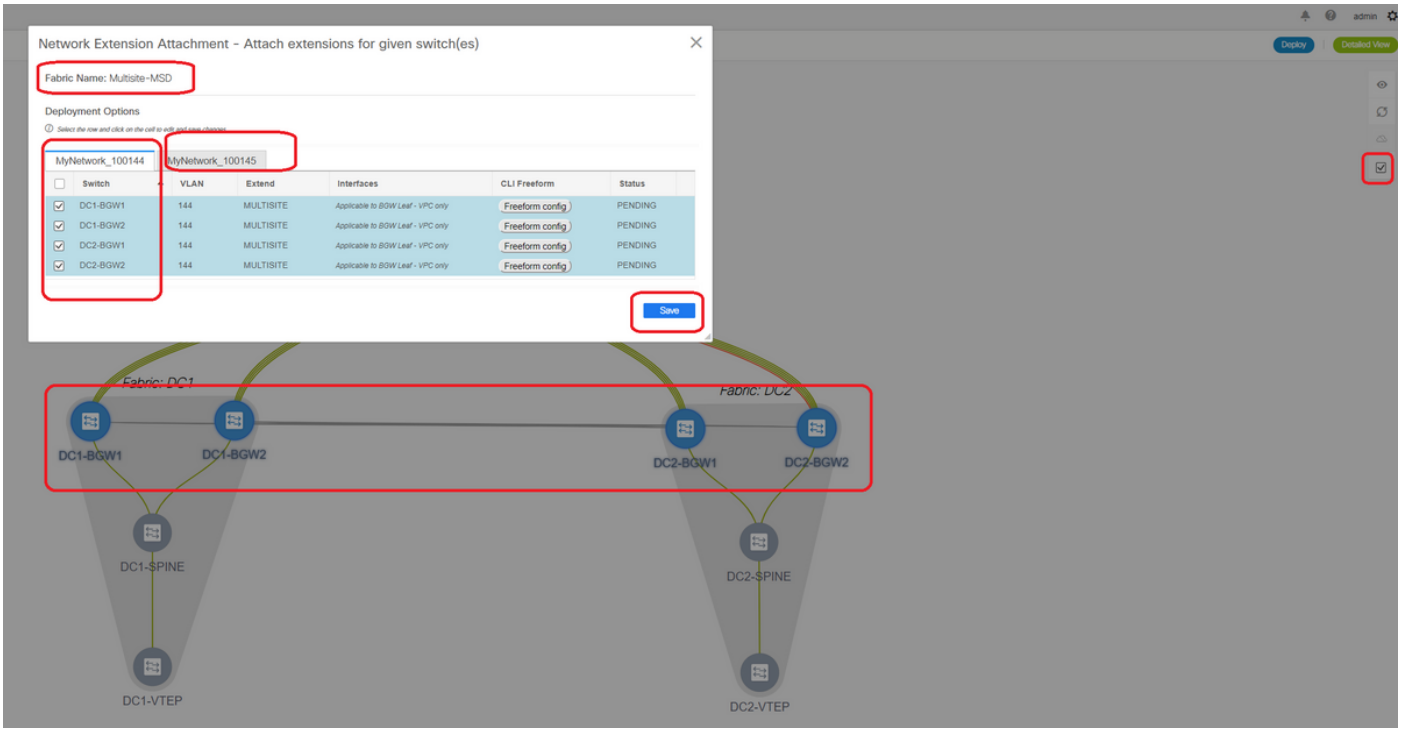
#在创建VRF扩展时，只有边界网关将具有针对VRFLITE DCI交换机的额外配置

#因此，必须单独选择常规枝叶，然后点击每个租户VRF的“复选框”，如上所示。

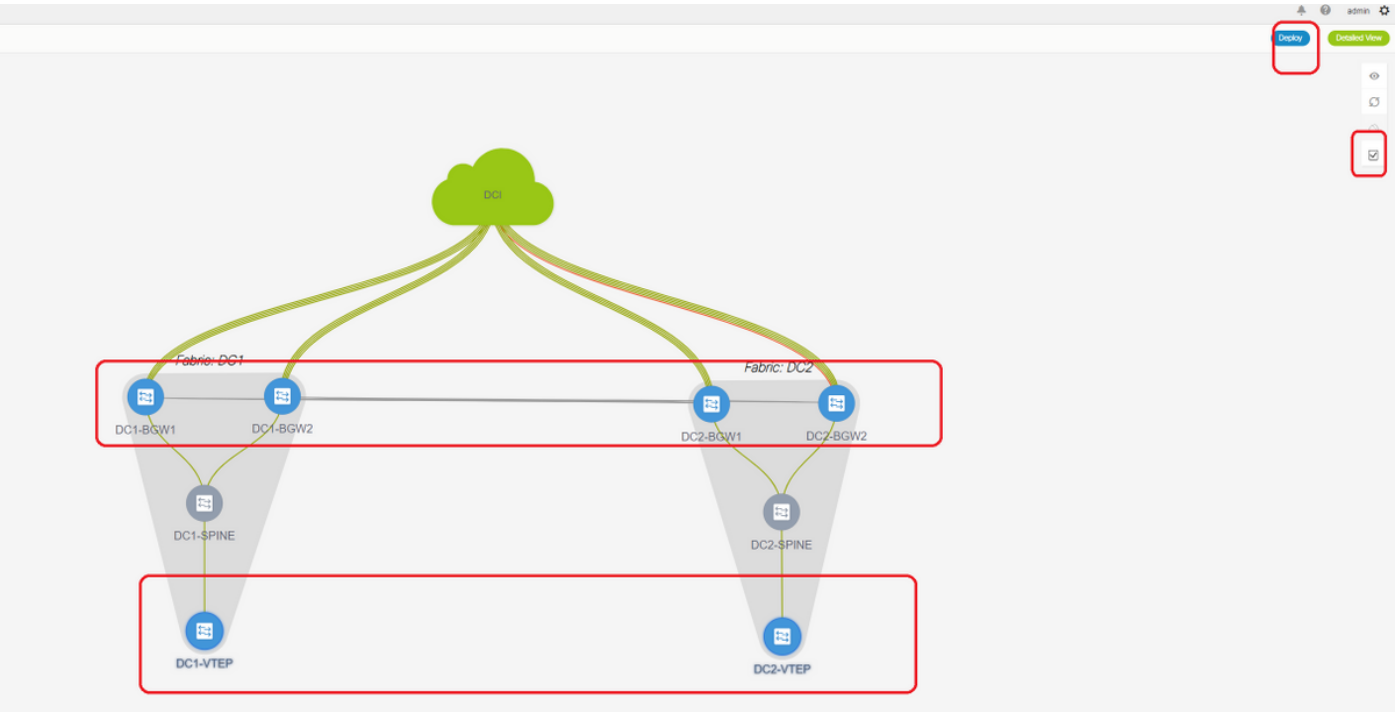
#点击部署(Deploy)推送配置

步骤 15：将网络配置从MSD推送到交换矩阵

#选择MSD交换矩阵中的相关网络



#请注意，此时仅选择边界网关；执行相同操作，并在本例中选择Regular Leaf switches/VTEPs-> DC1-VTEP和DC2-VTEP。



#完成后，点击“部署”（将配置推送到上述所有6台交换机）

步骤 16：检验所有VRF上的VRF和网络

#此步骤是验证VRF和网络是否在所有交换矩阵上显示为“已部署”；如果其显示为挂起，请确保“部署”配置。

步骤 17：在外部交换矩阵上部署配置

#要将所有相关IP编址、BGP、VRFLITE配置推送到DCI交换机，需要执行此步骤。

#要执行此操作，请选择外部交换矩阵并点击“保存和部署”

```
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
```

#部署后，我们将看到4个IPv4 BGP邻居关系从每个DCI交换机到所有BGW，4个IPv4 VRF BGP邻居关系（适用于租户VRF EXtension）

第 18 步：在DCI交换机之间配置iBGP

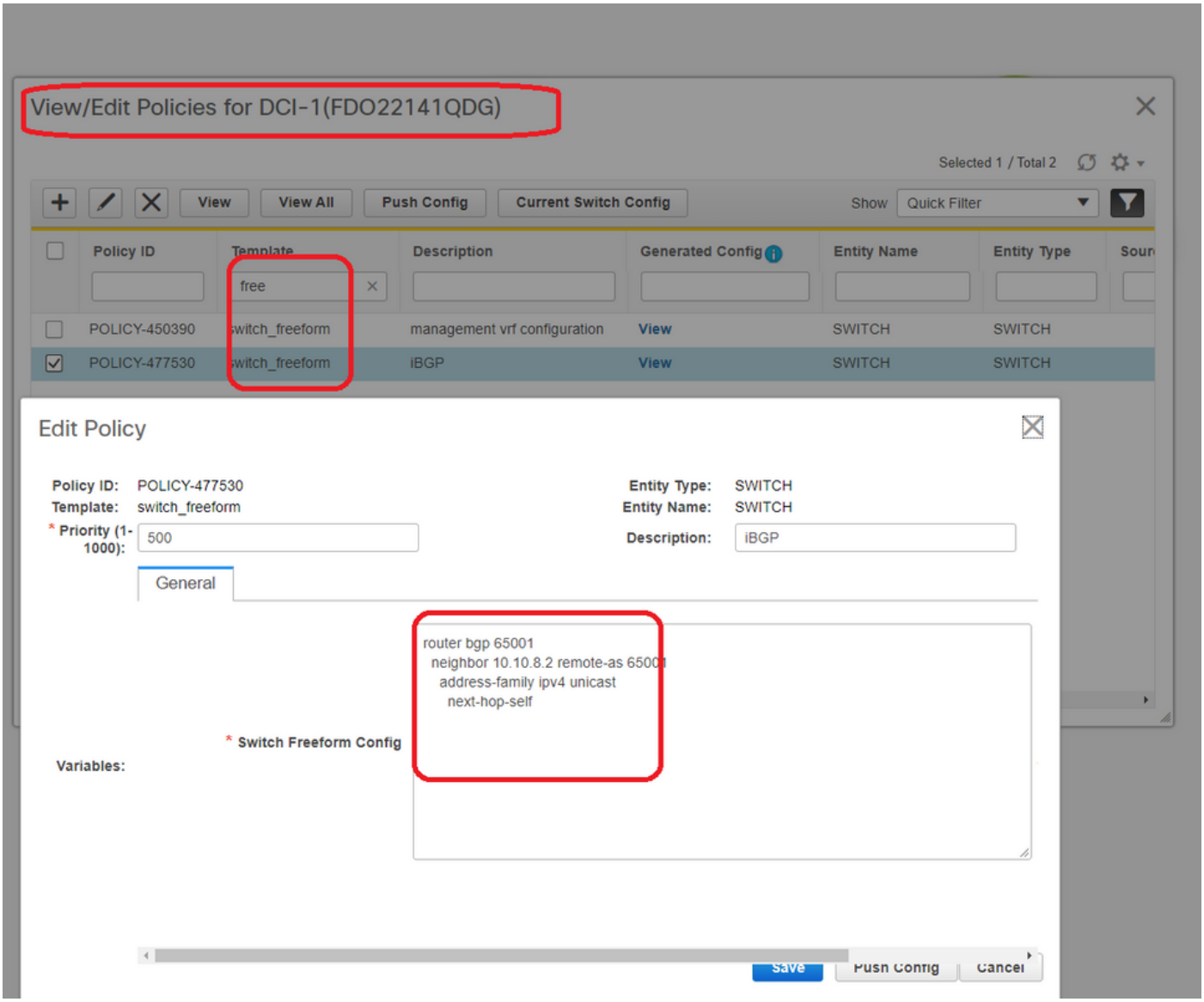
#考虑到DCI交换机之间有相互连接的链路，iBGP IPv4邻居关系是理想的，因此，如果DCI-1交换机上的任何下行连接中断，北向南流量仍可通过DCI-2转发

#为此，DCI交换机之间需要iBGP IPv4邻居关系，并且在每端使用next-hop-self。

#必须在DCI交换机上启动自由形式，才能实现此目的。所需的配置行如下所示。

#上述拓扑中的DCI交换机在vPC中配置；因此，备份SVI可用于构建iBGP邻居关系

#选择DCI交换矩阵，右键点击每台交换机并“查看/编辑策略”



#在DCI-2交换机上执行相同的更改，然后“保存并部署”，将实际配置推送到DCI交换机

#完成后，可以使用以下命令完成CLI验证。

```
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 neighborship
from DCI-2 to DCI-1
```

第 19 步：验证IGP/BGP邻居关系

OSPF邻居关系

#由于本示例中所有底层IGP都是OSPF，所有VTEP将与主干形成OSPF邻居关系，这还包括一个站点中的BGW交换机。

```
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
```

#所有环回（BGP路由器ID、NVE环回）在OSPF中通告；因此，在交换矩阵中，所有环回都通过OSPF路由协议获取，这有助于进一步形成l2vpn evpn邻居关系

BGP邻居

#在交换矩阵中，此拓扑将具有从主干到常规VTEP以及到边界网关的l2vpn evpn邻居关系。

```
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
```

#考虑到这是使用eBGP l2vpn evpn从一个站点到另一个站点对等的具有边界网关的多站点部署，可以在边界网关交换机上使用以下命令来验证相同情况。

```
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
```

TRM的BGP MVPN邻居关系

TRM配置到位后，所有枝叶交换机（包括BGW）将与主干形成mvpn邻居关系

```
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
```

#此外，边界网关需要在彼此之间形成mvpn邻居关系，以便东/西组播流量能够正确传输。

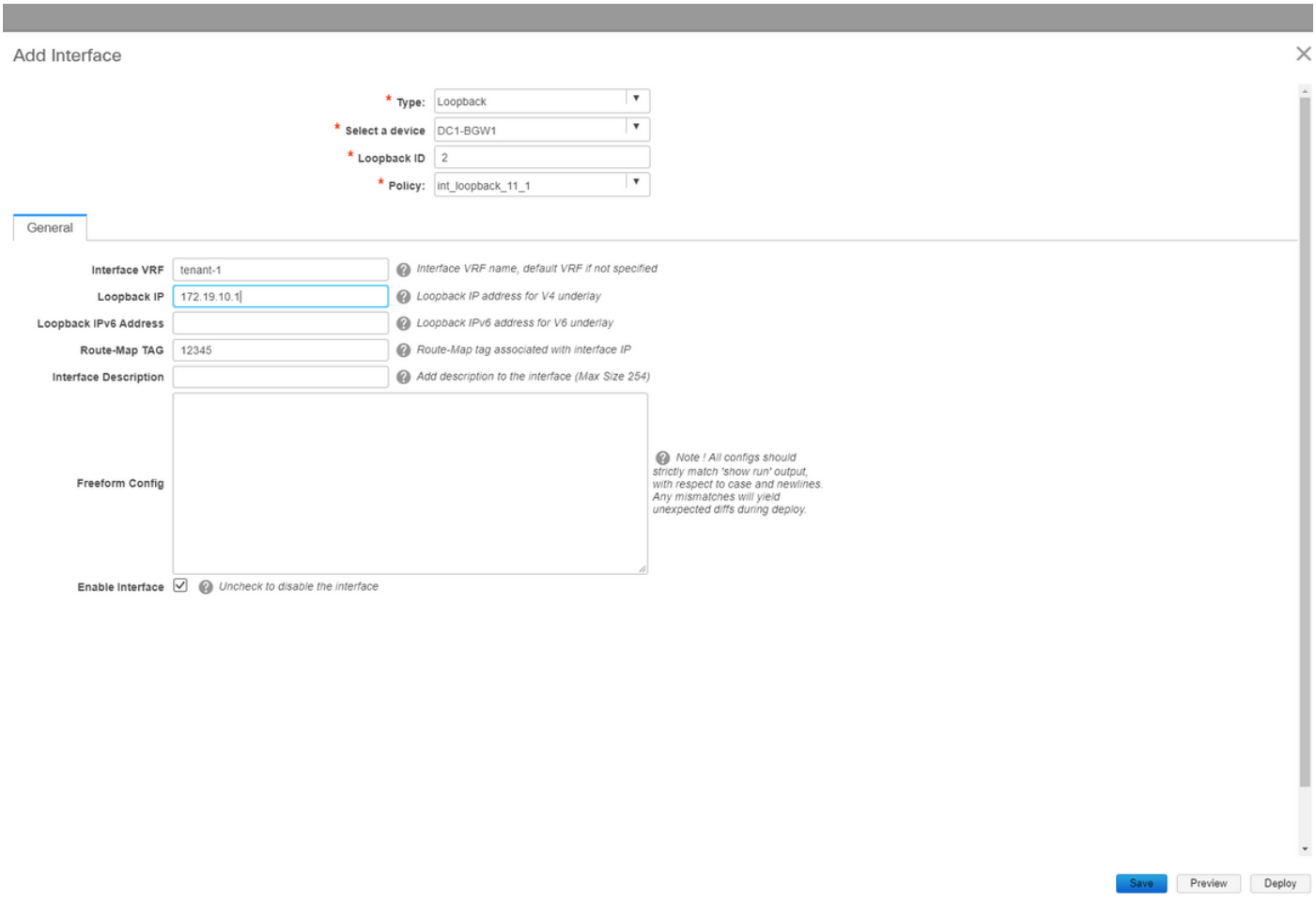
```
DC1-BGW1# show bgp ipv4 mvpn summary
BGP summary information for VRF default, address family IPv4 MVPN
BGP router identifier 10.10.10.1, local AS number 65000
BGP table version is 6, IPv4 MVPN config peers 3, capable peers 3
0 network entries and 0 paths using 0 bytes of memory
BGP attribute entries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

第 20 步：在边界网关交换机上创建租户VRF环回

#在租户VRF中创建环回，在所有边界网关上使用唯一的IP地址。

#为此，选择DC1，右键单击DC1-BGW1，管理接口，然后创建环回，如下所示。



#必须在其他3个边界网关上执行相同步骤。

第 21 步：DCI交换机上的VRFLITE配置

#在此拓扑中，DCI交换机配置了指向BGW的VRFLITE。VRFLITE也配置为DCI交换机的北部（即核心交换机）

#出于TRM目的，VRF租户–1中的PIM RP位于通过VRFLITE连接到DCI交换机的核心交换机中

#此拓扑具有从DCI交换机到位于图顶部的VRF租户–1内核心交换机的IPv4 BGP邻居关系。

#为此，创建子接口并为其分配IP地址，同时建立BGP邻居关系（这些由CLI直接在DCI和核心交换机上完成）

```
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
```

#以红色表示从DCI-1到核心交换机的BGP邻居。

```
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
```

#核心交换机（返回DCI-1和DCI-2）上也需要相应的BGP配置

单播验证

从DC1-Host1到DC2-Host1的East/West

#从DCNM和手动CLI推送上述所有配置（步骤1至21）后，单播可达性应在East/West运行

```
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
```

从DC1-Host1向北/向南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
```

组播验证

在本文档中，“租户–1” VRF的PIM RP已配置并呈现在VXLAN交换矩阵外部；根据拓扑，PIM RP在核心交换机上配置IP地址 — > 10.200.200.100

非vxlan（核心交换机后）的源，DC2的接收器

参考拓扑，如开头所示。

#源自非VXLAN主机的北/南组播流量 — > 172.17.100.100，接收器存在于两个数据中心；DC1-Host1-> 172.16.144.1和DC2-Host1-> 172.16.144.2，组 — > 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
```

DC1中的源，DC2中的接收器以及外部

```
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 ---
```

DC2中的源，DC1中的接收器以及外部

```
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 ---
```