ADDC5.0 Underlay自动化失败问题的经验案例

ADDC解决方案 Underlay自动化 王鹤1 2020-03-25 发表

组网及说明

ADDC5.0分布式网关方案Underlay自动化配置，发现设备自动化失败，EVPN邻居无法自动建立

问题描述

无

过程分析

ADDC5.0分布式网关方案Underlay自动化配置，Underlay自动化配置是ADDC5.0方案顺利开局的基础前提，而交换机Underlay配置自动化成功的标志就是Spine交换机与Leaf交换机的EVPN邻居正能够常建立，ADDC5.0分布式网关方案交换机Underlay自动化失败问题排查前提：当前环境的SNA Center、SeerEngine和vDHCP等软件均已部署完毕，网络设备已按规划上架并连线。排查思路如下：

1）**步骤1**：Underlay协议是否为OSPF，SeerEngine支持三种Underlay IGP协议，分别为OSPF、ISIS和BGP，如果是OSPF协议则进行步骤3进行排查，如果是ISIS/BGP协议则请检查相关配置。

2）**步骤2**：查看交换机管理口能否正常获取管理网段IP地址。如果未获取到管理网段IP地址，则需要排查Fabric Director上添加的DHCP服务器配置，以及管理网能否正常收发DHCP报文，如果交换机管理口已经正常获取到IP地址，则进行步骤4进行排查。

3）**步骤4**：查看交换机能否正常下载对应设备的启动模板文件。如果未下载到正确启动模板，则进行步骤5进行排查，如果已经下载到正确的启动模板文件，则进行步骤8进行排查。

4）**步骤5**：查看交换机角色是否配置正确，如果当前交换机角色与规划角色不一致，则需要手动更改交换机角色后重新进行自动化操作，如果当前交换机角色正确，则进行步骤6进行排查。

5）**步骤6**：查看DHCP服务器是否指定正确的模板文件。如果指定模板文件错误且设备解析错误，则进行修改，如果指定了正确的模板文件设备解析错误，则进行步骤7进行排查。

6）**步骤7**：查看TFTP服务是否正常。如果服务都正常并且基础网络正常，仍然无法下载正常配置模板，请拨打热线电话400-810-0504寻求帮助。如果服务异常，则修改对应配置再次检查，如果检查后服务依然异常，请拨打热线电话400-810-0504寻求帮助。

7）**步骤8**：查看当前交换机是否堆叠，如果交换机不采用堆叠部署，则进行步骤11进行排查，如果交换机采用堆叠部署，则进行步骤9进行排查。

8）**步骤9**：查看交换机角色是否一致。如果不一致，则需要修改成一致，如果一致且自动堆叠无法正常进行则需要排查交换机能否手动堆叠，如果可以手动堆叠自动化无法正常堆叠，请拨打热线电话400-810-0504寻求帮助。

10）**步骤10**：查看Spine Master交换机是否获取环回口IP地址，如果获取到环回口地址，则进行步骤13进行排查，如果Spine Master交换机获取不到IP地址，则进行步骤12进行排查。

11）**步骤11**：查看模板中配置的MASTER SPINE MAC是否填写的小写MAC。如果填写的是小写MAC地址且Spine Master无法获取环回口地址，请拨打热线电话400-810-0504寻求帮助。如果填写是大写MAC地址，则需要进行修改配置并重新进行自动化操作。

12）**步骤12**：查看Leaf交换机是否获取环回口IP地址，如果获取到环回口地址，则进行步骤13进行排查，如果不能获取到对应环回口地址，则需要排查Spine Master交换机的NETCONF用户名密码与Leaf是否一致，并检查管理网能否通信。

13）**步骤13**：查看交换机间EVPN邻居是否建立。如果EVPN邻居无法建立则排查交换机互联口的连接情况，如果EVPN邻居正常建立，但是缺失其他相关配置，请拨打热线电话400-810-0504寻求帮助。

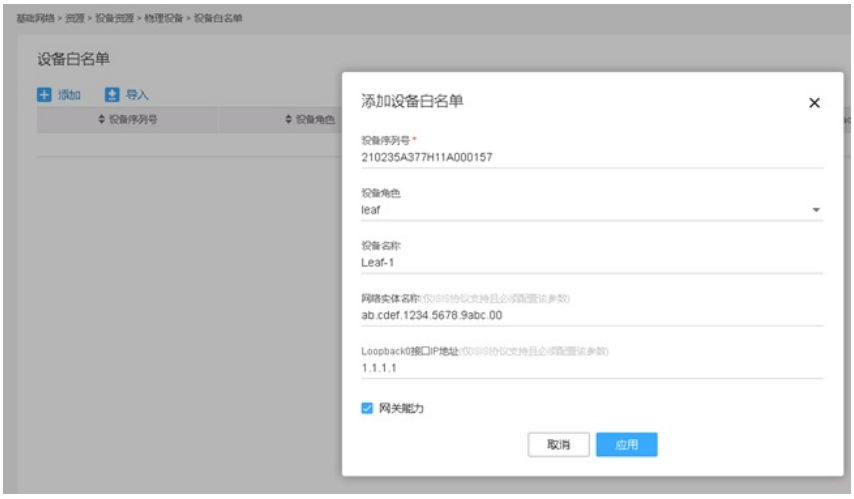
解决方法

1.判断Underlay协议是否为OSPF。

当前Fabric Director在交换机Underlay协议方面支持OSPF/ISIS/BGP三种IGP协议，如下图：

如果Underlay协议选择了ISIS，请在【基础网络/资源/设备资源/物理设备/设备白名单】路径配置设备

曰名单功能，并检查设备曰名单是否填写了网络实体名称和LoopBack0接口IP地址。 如下图：



如果Underlay协议选择了BGP，请配置【Underlay互通网络】，用于给设备LoopBack1接口下发IP地址互通。

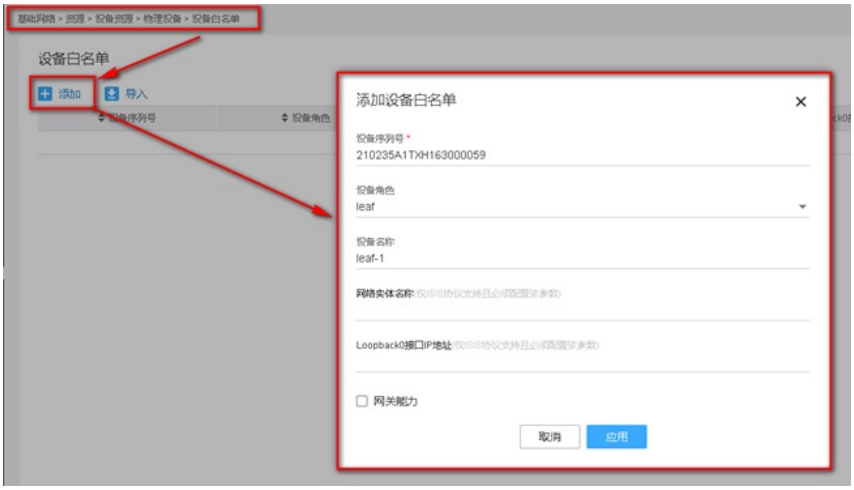
对于Underlay IGP协议选择OSPF协议的场景请直接进行步骤3进行排查

2. 查看该Fabric是否为设备白名单上线

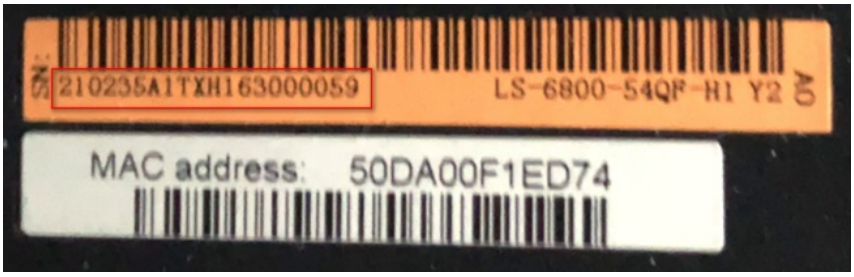
查看该Fabric是否为设备白名单上线，如下图，在SeerEngine的【基础网络/网络/Fabric/FabricName/设置】路径中，确认设备纳管方式是否为设备白名单功能。

如果设备纳管方式为强制纳管，则无需配置设备白名单，所有设备都可以通过Underlay自动化功能自动上线。

如果设备纳管方式为设备白名单，则需要在SeerEngine的【基础网络/资源/设备资源/物理设备/设备白名单】路径中，依据设备的SN码配置对应的设备白名单，如下图：



设备SN码第一种查看方式，直接获取交换机机箱上黄色标签记录的SN数值，如下图：



设备SN码第二种查看方式，登录交换机命令行界面，使用“display device manuinfo”命令查看，槽位1 CPU 0 的“DEVICE_SERIAL_NUMBER”，下图标红部分所示：

```
<leaf-1.12>display device manuinfo
 Slot 1 CPU 0:
DEVICE_NAME         : S6800-54QF
DEVICE_SERIAL_NUMBER : 210235A1TXH163000059
MAC_ADDRESS         : 50DA-00F1-ED74
MANUFACTURING_DATE   : 2016-03-13
VENDOR_NAME         : H3C
PRODUCT ID          : LS-6800-54QF-H1
```

. 查看交换机管理口能否正常获取管理网段IP地址

查看交换机管理口能否正常获取管理网段IP地址。交换机设备在没有配置文件的启动过程中首先会从VLAN 1 接口发起DHCP DISCOVER请求，如果获取不到IP地址，则会继续从管理接口发起DHCP DISCOVER请求，如果交换机管理口能够获取DHCP服务器分配的正确IP地址，则进行步骤4进行排查，如果交换机管理口无法获取DHCP服务器上分配的IP地址段，则需要按照下文步骤排查DHCP服务器配置以及交换机所在管理口的管理网是否正常传递DHCP报文。

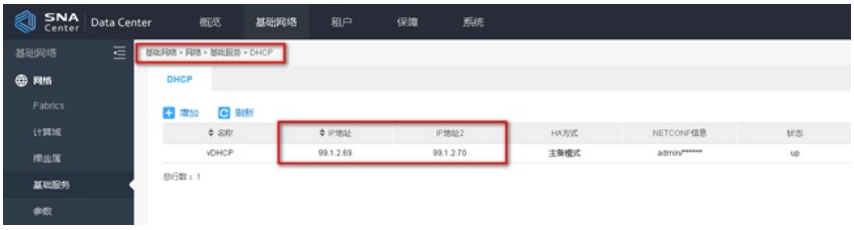
交换机启动过程中，管理口获取到IP地址的正常日志显示如下，启动过程中请勿使用CTRL_C or CTRL_D中断交换机自动化过程。

```
System is starting...
Cryptographic algorithms tests passed.
Startup configuration file does not exist.
Performing automatic configuration... Press CTRL_C or CTRL_D to break.
Automatic configuration attempt: 1.
Not ready for automatic configuration: no interface available.
Waiting for the next...
Automatic configuration is running, press CTRL_C or CTRL_D to break.
Automatic configuration attempt: 2.
Interface used: M-GigabitEthernet0/0/0.
Enable DHCP client on M-GigabitEthernet0/0/0.
Obtained an IP address for M-GigabitEthernet0/0/0: 100.1.1.102.
```

如果交换机管理口无法正常获取管理网IP地址，请排查管理网络是否通畅。

如果是二层组网，SNA Center与交换机VLAN1接口/M接口在同一网段，此时保证连线正确，VLAN放通即可。

如果是三层组网，需要在管理交换机上创建管理网地址池的网关，并在该网关的接口上配置DHCP Relay，将交换机发送的DHCP Discover报文中继给vDHCP地址。可以在SeerEngine的【基础网络/网络/基础服务/DHCP】中添加vDHCP或查看vDHCP地址。



如果上述步骤都正常，则进行步骤4进行排查。如果检查后服务依然异常，请拨打热线电话400-810-0504寻求帮助。

4.查看设备角色是否自动生成错误

如果Underlay自动化完成，交换机会以设定的角色自动加入到SeerEngine中，如下图，在SeerEngine的【基础网络/网络/Fabrics/FabricName/交换设备】路径中查看：

SeerEngine上的设备角色由交换机上配置的fabric-role和设备是否具有网关能力共同决定。如果设备角色与预期的不一致，请依据如下标准检查配置：

A. 设备角色为Spine && 网关能力为True -> 边界设备

B. 设备角色为Spine && 网关能力为False -> Underlay物理设备

C. 设备角色为Leaf && 网关能力为True -> 边界设备

D. 设备角色为Leaf && 网关能力为False -> 接入设备

设备角色可以在设备命令行输入命令“display vcf fabric-role”查看，如下：

```
[leaf-1.11]display vcf-fabric role
Default role: leaf
Current role: leaf
[leaf-1.11]
```

网关能力请在设备白名单中查看，如果设备未配置白名单，则默认网关能力为FALSE。

5.查看交换机能否正常下载对应设备的启动模板文件或软件版本

交换机管理口获取到IP地址后，由于DHCP服务器回复的DHCP OFFER报文中携带了Option 67记录了模板文件的名称，Option 66记录了TFTP地址。设备收到DHCP OFFER报文，自动化脚本会自动向TFTP服务器请求名称为模板名_设备角色.template的文件，用作实际自动化的模板文件。下面以Leaf启动为例：

```
Interface used: M-GigabitEthernet0/0/0.
Enable DHCP client on M-GigabitEthernet0/0/0.
Obtained an IP address for M-GigabitEthernet0/0/0: 100.1.1.102.
Obtained configuration file name fabric3.template and TFTP server name 100.1.1.10.
Resolved the TFTP server name to 100.1.1.10.
INFO: Not found tag file device_tag.csv.
Successfully downloaded file fabric3_leaf.template.
Executing the configuration file. Please wait...
Automatic configuration successfully completed.
```

交换机设备出厂默认会设置VCF角色，例如：S12500、S9800系列交换机默认出厂为Spine角色，而S6800系列交换机默认出厂为Leaf角色。例如，S6800交换机做Leaf设备时，下载的模板文件名称应该是模板名_leaf.template。

如果下载对应模板错误、无法下载模板、软件版本下载错误或者软件版本无法下载可能会导致自动化失败，则进行步骤5进行排查。如果下载模板和软件版本正确则进行步骤8进行排查。如果下载软件版本成功但无法升级，请检查该软件版本是否能用于目的设备的升级。

6.查看交换机角色是否正确

对于需要使用到非默认交换机角色的场景，例如，使用S6800交换机做为Spine交换机，需要首先将对应交换机开机更改默认VCF角色才能确保该交换机获取对应正确自动化模板，登录交换机命令行界面，分别使用“system-view”“vcf-fabric role spine”命令即可将当前S6800交换机更改成Spine模式，如下所示：

```
<spine-1.11>system-view
[spine-1.11]vcf-fabric role spine
```

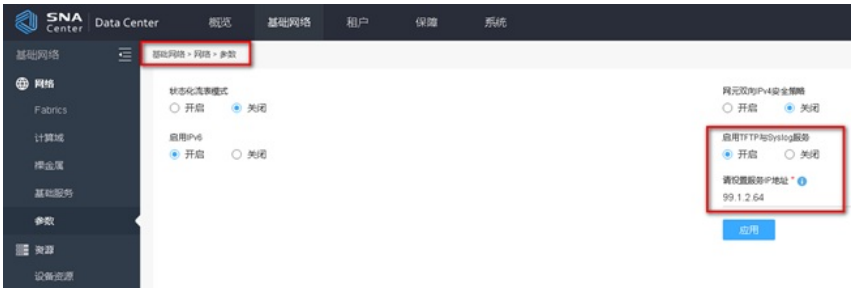
交换机修改对应角色后无需保存，修改完毕交换机角色后在对应flash中新增rolename文件，文件记录对应交换机新角色，可以登录交换机命令行界面，使用“more flash:/rolename”，或者“display vcf-fabric role”命令查看当前交换机VCF 角色，如下所示：

```
<spine-1.11>more flash:/rolename
spine
<spine-1.11>display vcf-fabric role
Default role: leaf
Current role: spine
```

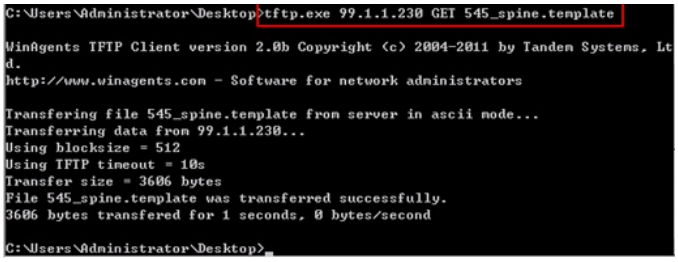
如果交换机角色正确，则进行步骤6进行排查，如果交换机角色不正确则需要进行修改。

7.查看TFTP服务是否正常。

首先查看SeerEngine路径【基础网络/网络/参数】中，是否开启了TFTP服务以及配置了TFTP地址，如下图：



然后使用tftp client测试能否下载正确的模板文件，如下图：



如果可以下载对应模板文件，但是交换机无法获取对应模板文件，请拨打热线电话400-810-0504寻求帮助。如果tftp client无法获取对应模板文件，但是TFTP 服务正常，请检查管理网中间网络问题，例如，是否存在防火墙阻断等问题。如果仍旧无法解决，请拨打热线电话400-810-0504寻求帮助。

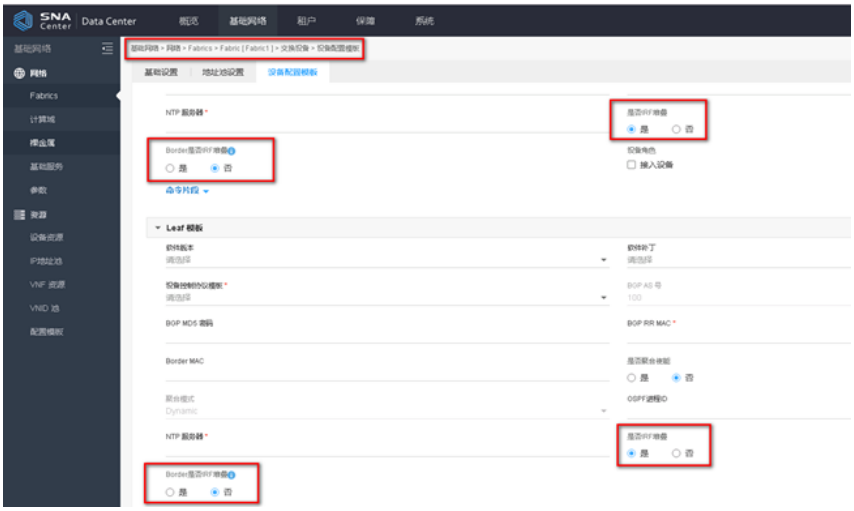
8.交换机是否堆叠。

如果交换机采用堆叠方式部署，则进行步骤9进行排查，如果交换机采用单机部署则进行步骤11进行排查。

9.堆叠交换机角色是否一致。

无论是Spine交换机还是Leaf交换机，如果采用堆叠方式部署，需要保证如下两个条件：
第一：堆叠设备角色一致。
第二：设备之间存在至少一根互联堆叠线缆。

同时，需要在SeerEngine的【基础网络/网络/Fabrics/FabricName/交换设备/设备配置模板】路径中确认是否勾选了IRF堆叠，如下图：



如果预期堆叠设备是Border设备，请将“Border是否IRF堆叠”选项选为“是”，同时填写Border Mac。如果Border设备的fabric-role为leaf，请在Leaf模板中配置，如果Border设备的fabric-role为spine，请在spine模板中配置。

判断设备角色方法参照步骤5，如果设备角色一致且存在连线仍然出现自动化失败的问题，请尝试手动配置两台设备堆叠，如果手动无法堆叠，则需要排查交换机自身堆叠失败原因，例如：千兆光转电模块无法用作堆叠口，解决后继续进行步骤10进行排查。如果设备能够手动堆叠但是自动化堆叠失败，请拨打热线电话400-810-0504寻求帮助。

10.查看Spine Master交换机是否已给自己分配环回口IP地址。

交换机进行自动化过程中Spine Master交换机会给自身和全网能够发现的其他Spine或者Leaf交换机下发环回口IP地址。如果在SeerEngine上选择了控制器分配，则是由SDN控制器通过管理网给Spine/Leaf分配业务环回口地址。请登录SeerEngine前台【基础网络/网络/Fabrics/Fabric/交换设备/基础设置】查看“VTEP IP分配方式”，如下图：

正确的Spine Master自动化完成后接口地址情况，登录交换机命令行界面，使用“display ip interface brief”命令确认Spine交换机环回口Loop0是否获取到IP地址，如下所示：

```
<spine-1.11>display ip interface brief
*down: administratively down
(s): spoofing  (l): loopback
Interface          Physical Protocol IP address      VPN instance Description
Loop0              up       up(s)    101.1.1.254     --           --
MGE0/0/0           up       up       99.1.1.11       mgmt         --
MGE0/0/1           down     down     --              --           --
Vlan1              up       up       --              --           --
```

如果Spine Master交换机没有分配环回口，则请检查Spine Master获取到的template文件中BGP RR MAC是否为小写，如下图：

如果Spine Master交换机已经给自己分配环回口地址，则进行步骤12进行排查，

12.查看Leaf交换机是否获取环回口IP地址。

查看Leaf交换机是否获取到环回口IP地址以及环回口IP地址分配方式，查看命令参考步骤10。当选用的是Spine Master分配IP地址方式时，Leaf交换机启动后，参与Spine Master交换机的计

算Fabric拓扑过程，Spine Master通过LLDP识别Leaf的管理口即PeerIP，使用当前Spine Master设备的用户名密码通过管理网登录使用NETCONF协议将对应环回口地址下发到对应Spine/Leaf交换机上。

Spine Master上记录的拓扑信息，登录交换机依次使用“system-view”, “probe”, “display system internal vcf-fabric topology”命令查看拓扑信息,如下所示：

```
<spine-1.11>system-view
[spine-1.11] probe
[spine-1.11-probe]display system internal vcf-fabric topology
                          Topology Information
-------------------------------------------------------------------------------
 Ten-GigabitEthernet1/0/1:
    PeerIP : 99.1.1.13
    sysName : leaf-1.13
    sysdesc : h3c
    mac : 84d9-3191-546b
    PeerIntf : Ten-GigabitEthernet1/0/1
    role : leaf
    LoopbackIP : 101.1.1.252
 Ten-GigabitEthernet1/0/2:
    PeerIP : 99.1.1.12
    sysName : leaf-1.12
    sysdesc : h3c
    mac : 50da-00f1-e6d5
    PeerIntf : Ten-GigabitEthernet5/0/2
    role : leaf
    LoopbackIP : 101.1.1.253
```

如果Leaf交换机获取到环回口IP地址，则进行步骤13进行排查。如果Leaf交换机没有获取到环回口地址，则需要登录到Spine Master上检查对应NETCONF用户名密码与Leaf上是否一致，并且测试管理网能够登录。排查步骤如下：

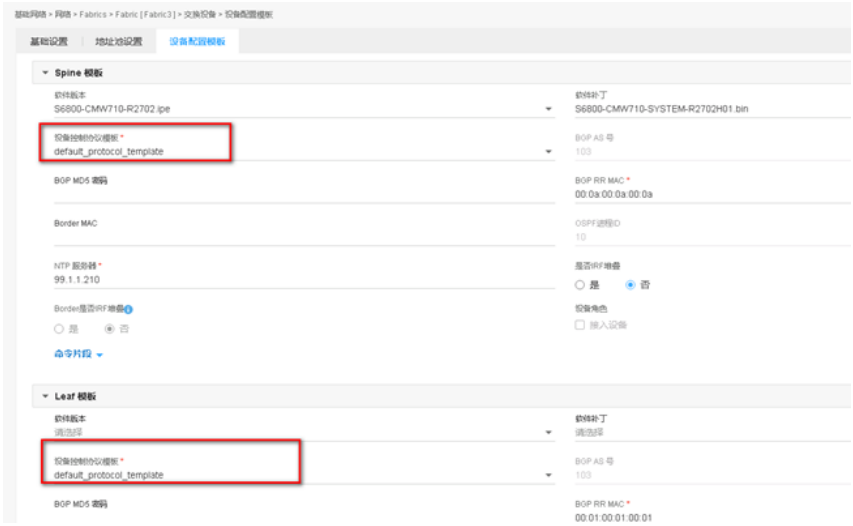
（1）登录Spine Master交换机命令行界面，使用“display local-user”命令查看当前Local-user是否存在对应用户，以及对用用户的权限，如下：

```
<spine-1.11>display local-user
Device management user admin:
  State:                    Active
  Service type:             SSH/Telnet/HTTP/HTTPS
  User group:               system
  Bind attributes:
  Authorization attributes:
    Work directory:         flash:
    User role list:         network-admin, network-operator
Total 1 local users matched.
```

（2） 登录Spine交换机命令行界面，使用“ssh2 99.1.1.12 vpn-instance mgmt”使用该admin用户远程登录Leaf设备配置对应环回口地址，命令如下：

```
<spine-1.11>ssh2 99.1.1.12 vpn-instance mgmt
Username: admin
Press CTRL+C to abort.
Connecting to 99.1.1.12 port 22.
The server is not authenticated. Continue? [Y/N]:y
Do you want to save the server public key? [Y/N]:n
admin@99.1.1.12's password:
Enter a character ~ and a dot to abort.
******************************************************************************
* Copyright (c) 2004-2017 New H3C Technologies Co., Ltd. All rights reserved.*
* Without the owner's prior written consent,                                 *
* no decompiling or reverse-engineering shall be allowed.                    *
******************************************************************************
<leaf-1.12>
```

如果使用ssh客户端无法登录，或者ssh登陆到Leaf交换机之缺失network-admin权限，则需要登录到SeerEngine上进行修改，将Spine模板和Leaf模板的“设备控制协议模板”配置成相同账号密码的模板，如下图：



修改完成后，将Leaf交换机清空配置，重新启动，并进行步骤13进行排查。

如果可以登录，权限也正确，但是Leaf没有获取环回口地址，请拨打热线电话400-810-0504寻求帮助。

13.查看交换机EVPN邻居是否建立。

ADDC5.0方案中，Leaf交换机获取到环回口IP地址后，Spine与Leaf的互联口会借用对应环回口配置OSPF/ISIS/BGP，包括EVPN作为建立邻居使用。登录交换机命令行界面，使用“display bgp peer l2vpn evpn”命令查看对应EVPN邻居是否正常建立，如下所示：

```
<spine-1.11>display bgp peer l2vpn evpn
 BGP local router ID: 101.1.1.254
 Local AS number: 100
 Total number of peers: 2          Peers in established state: 2
  * - Dynamically created peer
 Peer            AS  MsgRcvd  MsgSent OutQ PrefRcv Up/Down  State
 101.1.1.252     100       52       44    0       0 00:38:56 Established
 101.1.1.253     100       34       34    0       0 00:30:26 Established
```

如果查看之后没有邻居，继续使用“display current-configuration configuration bgp”命令查看是否存在对应配置，确保对应BGP配置中包含对应peer配置，如下所示：

```
<spine-1.11>display current-configuration configuration bgp
#
bgp 100
 non-stop-routing
 group evpn internal
 peer evpn connect-interface LoopBack0
 peer 101.1.1.252 group evpn
 peer 101.1.1.253 group evpn
 #
 address-family l2vpn evpn
  undo policy vpn-target
  peer evpn enable
  peer evpn reflect-client
#
return
<spine-1.11>
```

如果配置中不存在对应peer配置，可以尝试将互联端口shutdown之后再undo shutdown。手动触发接口变化事件，再次查看EVPN邻居，如果邻居仍然无法建立，请检查Spine与Leaf交换机之间的互联线的连通性，如果邻居仍然无法建立，请拨打热线电话400-810-0504寻求帮助。

本案例中Underlay自动化失败，EVPN邻居无法建立的原因为Leaf到Spine的连线错误，从而无法通过LLDP学习邻居，故而定位。