排除ACI VMM集成故障

目录

简介

本文档介绍了解ACI虚拟机管理器集成(VMM)并对其进行故障排除的步骤。

背景信息

本文档中的内容摘自思科以应用为中心的基础设施故障排除，第二版书籍，特别是**VMM集成 — 概述、VMM集成 — vCenter连接、VMM集成 — 主机动态发现**和**VMM集成 — 虚拟机监控程序上行链路负载平衡**章节。

Virtual Machine Manager概述

ACI控制器能够与第三方虚拟机管理器(VMM)集成。

这是ACI的主要功能之一，因为它可以简化和自动化交换矩阵和与之连接的工作负载的端到端网络配置操作。ACI提供单一重叠策略模型，可跨多个工作负载类型（即虚拟机、裸机服务器和容器）扩展。

本章将重点介绍与VMware vCenter VMM集成相关的一些典型故障排除场景。

读者将浏览以下内容：

- 调查vCenter通信故障。
- 主机和VM动态发现流程和故障场景。
- 虚拟机监控程序负载平衡算法。

vCenter连接

基于角色的访问控制(RBAC)

APIC能够与vCenter控制器连接的机制取决于与给定VMM域关联的用户帐户。概述与VMM域关联的vCenter用户的具体要求，以确保APIC能够在vCenter上成功执行操作，无论其是推送和检索库存和配置，还是监控和侦听托管库存相关事件。

消除此类要求担忧的最简单方法是使用具有完全访问权限的管理员vCenter帐户；但是，ACI管理员并不总是可以获得这种自由。

从ACI版本4.2开始，自定义用户帐户的最低权限如下：

- **警报** APIC在文件夹上创建两个警报。一个用于DVS，另一个用于端口组。在APIC上删除EPG或VMM域策略时，会引发警报，但是vCenter无法删除相应的端口组或DVS，因为有虚拟机连接到它。
- **分布式交换机**
- **dvPort组**
- **文件夹**
- **网络** APIC管理网络设置，例如添加或删除端口组、设置主机/DVS MTU、LLDP/CDP、LACP等。
- **主机** 如果除了上述功能外还使用AVS，用户需要数据中心的Host权限，APIC将在此创建DVS。**Host.Configuration.Advanced设置主机。本地操作。重新配置虚拟机Host.Configuration.Network configuration**AVS和虚拟第4层到第7层服务VM的自动放置功能需要此功能。对于AVS，APIC创建VMK接口，并将其置于用于OpFlex的VTEP端口组中。
- **虚拟机** 如果使用服务图，则还需要虚拟设备的虚拟机权限。**虚拟机。配置。修改设备设置虚拟机。配置。设置**

排除RBAC相关问题

RBAC问题最常在VMM域的初始设置期间遇到，但是，如果vCenter管理员要在初始设置完成后修改与VMM域关联的用户帐户的权限，则可能会遇到。

症状可以通过以下方式表现出来：

- 部分或完全无法部署新服务（DVS创建、端口组创建，某些对象已成功部署，但不是全部）。
- ACI管理员视图中未填写或缺少运行资产。
- 因不受支持的vCenter操作或以上任何场景（例如端口组部署故障）而引发故障。
- vCenter控制器报告为脱机，故障表明存在连接或与凭证相关的问题。

RBAC相关问题的解决方案

验证是否已向VMM域中配置的vCenter用户授予上述所有权限。

另一种方法是使用VMM域配置中定义的凭证直接登录vCenter，并尝试类似的操作（端口组创建等）。 如果用户在直接登录vCenter时无法执行这些操作，显然不会向用户授予正确的权限。

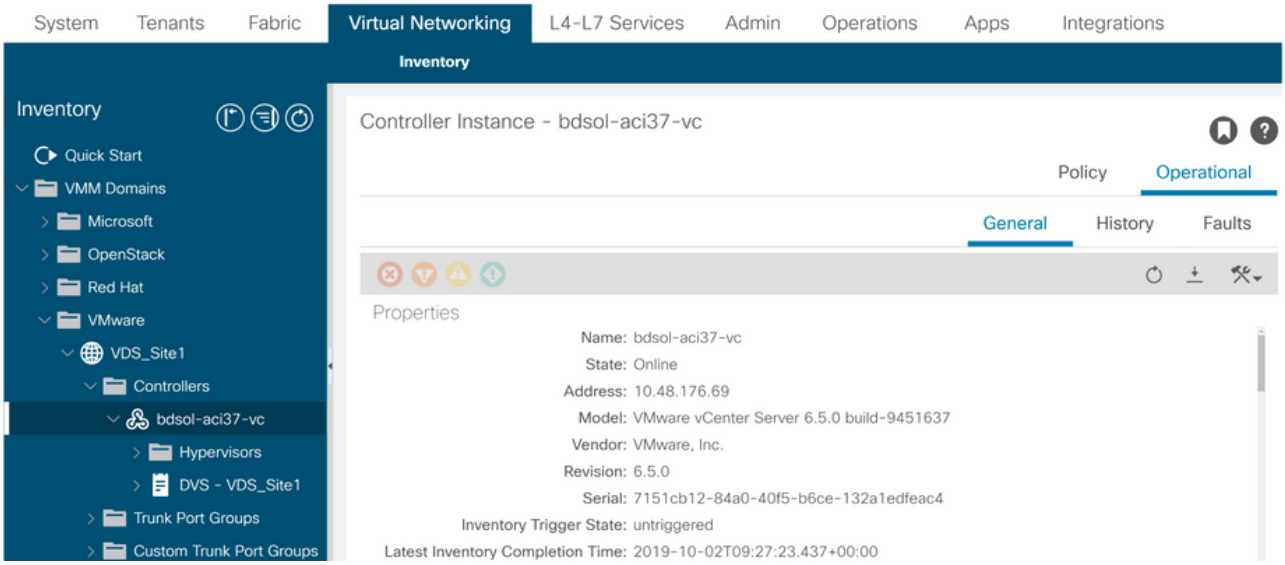
连接故障排除

在对VMM连接相关问题进行故障排除时，必须注意ACI如何与vCenter通信的一些基本行为。

第一个也是最相关的行为是集群中只有一个APIC在任何指定点发送配置和收集资产。此APIC称为此VMM域的**分片领导**。但是，多个APIC正在侦听**vCenter Events**，以便解释分片领导因任何原因错过了事件的场景。按照相同的APIC分布式架构，给定VMM域将有一个处理主数据和功能的APIC（在本例中为分片领导）和两个副本(在VMM中，它们称为跟**随**者)。 要跨APIC分配VMM通信和功能的处理，任何两个VMM域可以具有相同或不同的共享领导。

通过导航到GUI中感兴趣的VMM控制器或使用下面列出的CLI命令可以找到vCenter连接状态。

VMWare VMM域 — vCenter连接状态




```
apic2# show vmware domain name VDS_Site1 vcenter 10.48.176.69
Name                         : bdsol-aci37-vc
```

```
Type                          : vCenter
Hostname or IP                : 10.48.176.69
Datacenter                    : Site1
DVS Version                   : 6.0
Status                        : online
Last Inventory Sync           : 2019-10-02 09:27:23
Last Event Seen               : 1970-01-01 00:00:00
Username                      : administrator@vsphere.local
Number of ESX Servers         : 2
Number of VMs                 : 2
Faults by Severity            : 0, 0, 0, 0
Leader                        : bdsol-aci37-apic1

Managed Hosts:
ESX              VMs       Adjacency   Interfaces
---------------  --------  ----------  ------------------------------------------------
10.48.176.66     1         Direct      leaf-101 eth1/11, leaf-102 eth1/11
10.48.176.67     1         Direct      leaf-301 eth1/11, leaf-302 eth1/11
```

如果VMM控制器指示为脱机，将引发类似于以下的故障：

```
Fault fltCompCtrlrConnectFailed
Rule ID:130
Explanation:
This fault is raised when the VMM Controller is marked offline. Recovery is in process.
Code: F0130
Message: Connection to VMM controller: hostOrIp with name name in datacenter rootContName in
domain: domName is failing repeatedly with error: [remoteErrMsg]. Please verify network
connectivity of VMM controller hostOrIp and check VMM controller user credentials are valid.
```

以下步骤可用于排除VC和APIC之间的连接问题。

1.确定分片领导者

排除APIC和vCenter之间的连接问题的第一步是了解哪个APIC是给定VMM域的共享领导。确定此信息的最简单方法是在任何APIC上运行“show vmware domain name <domain>”命令。

```
apic1# show vmware domain name VDS_Site1
Domain Name                         : VDS_Site1
Virtual Switch Mode                 : VMware Distributed Switch
Vlan Domain                         : VDS_Site1 (1001-1100)
Physical Interfaces                 : leaf-102 eth1/11, leaf-301 eth1/11, leaf-302 eth1/11,
                                     leaf-101 eth1/11
Number of EPGs                      : 2
Faults by Severity                  : 0, 0, 0, 0
LLDP override                       : RX: enabled, TX: enabled
CDP override                        : no
Channel Mode override               : mac-pinning
NetFlow Exporter Policy             : no
Health Monitoring                   : no

vCenters:
Faults: Grouped by severity (Critical, Major, Minor, Warning)
vCenter               Type      Datacenter            Status    ESXs   VMs    Faults
--------------------  --------  --------------------  --------  -----  -----  ---------------
10.48.176.69          vCenter   Site1                 online    2      2      0,0,0,0

APIC Owner:
Controller    APIC      Ownership
------------  --------  ---------------
bdsol-        apic1     Leader
```

```
aci37-vc
bdsol-        apic2     NonLeader
aci37-vc
bdsol-        apic3     NonLeader
aci37-vc
```

2.检验与vCenter的连接

确定与vCenter主动通信的APIC后，使用ping等工具检验IP连接。

```
apic1# ping 10.48.176.69
PING 10.48.176.69 (10.48.176.69) 56(84) bytes of data.
64 bytes from 10.48.176.69: icmp_seq=1 ttl=64 time=0.217 ms
64 bytes from 10.48.176.69: icmp_seq=2 ttl=64 time=0.274 ms
64 bytes from 10.48.176.69: icmp_seq=3 ttl=64 time=0.346 ms
64 bytes from 10.48.176.69: icmp_seq=4 ttl=64 time=0.264 ms
64 bytes from 10.48.176.69: icmp_seq=5 ttl=64 time=0.350 ms
^C
--- 10.48.176.69 ping statistics ---
5 packets transmitted, 5 received, 0% packet loss, time 4084ms
rtt min/avg/max/mdev = 0.217/0.290/0.350/0.052 ms
```

如果使用FQDN而不是IP地址配置vCenter，可以使用nslookup命令验证名称解析。

```
apic1:~> nslookup bdsol-aci37-vc
Server: 10.48.37.150
Address: 10.48.37.150#53
Non-authoritative answer:
Name: bdsol-aci37-vc.cisco.com
Address: 10.48.176.69
```

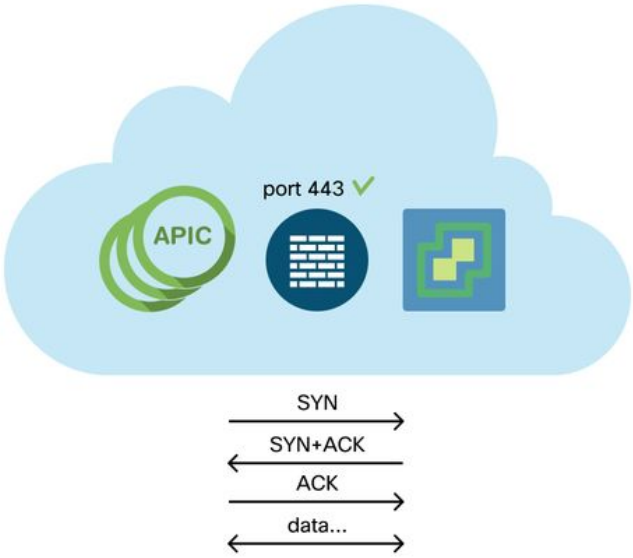
3.检查是否使用了OOB或INB

检查APIC路由表，验证连接是否首选带外或带内，以及使用的网关：

```
apic1# bash
admin@apic1:~> route
Kernel IP routing table
Destination     Gateway         Genmask         Flags Metric Ref    Use Iface
default         10.48.176.1     0.0.0.0         UG    16     0        0 oobmgmt
```

4.确保所有APIC和vCenter（包括通信路径中的任何防火墙）之间允许使用端口443。

vCenter <-> APIC - HTTPS（TCP端口443）— 通信



可以使用卷曲测试从APIC到vCenter的常规HTTPS可达性：

```
apic2# curl -v -k https://10.48.176.69
* Rebuilt URL to: https://10.48.176.69/*   Trying 10.48.176.69...
* TCP_NODELAY set
* Connected to 10.48.176.69 (10.48.176.69) port 443 (#0)
...
```

使用netstat命令验证分片报头在端口443上已建立TCP连接。

```
apic1:~> netstat -tulaen | grep 10.48.176.69
tcp 0 0 10.48.176.57:40806 10.48.176.69:443 ESTABLISHED 600 13062800
```

5.执行数据包捕获

如果可能，请沿着分片领导者和vCenter之间的路径执行数据包捕获，以识别任一设备是否正在发送和接收流量。

VMware资产

下表显示了VMWare VDS参数的列表，并指定这些参数是否可由APIC配置。

由APIC管理的VMware VDS参数

VMware VDS	默认值	是否可使用思科APIC策略进行配置？
名称	VMM域名	是（派生自域）
描述	'APIC虚拟交换机'	无
文件夹名称	VMM域名	是（派生自域）
version	vCenter支持的最高	Yes
发现协议	LLDP	Yes
上行链路端口和上行链路名称	8	是(来自思科APIC版本4.2(1))
上行链路名称前缀	上行链路	是(来自思科APIC版本4.2(1))

最大 MTU	9000	Yes
LACP策略	禁用	Yes
端口镜像	0个会话	Yes
警报	在文件夹级别添加了2个警报	无

下表显示了VMWare VDS端口组参数的列表，并指定这些参数是否可由APIC配置。

VMWare VDS端口组参数由APIC管理

VMware VDS端口组	默认值	可使用APIC策略进行配置
名称	租户名称 |应用配置文件名称 | EPG名称	是（源自EPG）
端口绑定	静态绑定	无
VLAN	从VLAN池中选择	Yes
负载均衡算法	根据APIC上的端口通道策略派生	Yes
混杂模式	禁用	Yes
伪传输	禁用	Yes
MAC更改	禁用	Yes
阻止所有端口	错误	无

VMware资产故障排除

发生资产同步事件以确保APIC了解可能需要APIC动态更新策略的vCenter事件。vCenter和APIC之间可以发生两种类型的资产同步事件；完全资产同步和基于事件的资产同步。APIC和vCenter之间的完全库存同步的默认计划是每24小时一次，但是也可以手动触发这些计划。基于事件的资产同步通常与触发任务（如vMotion）相关联。在此场景中，如果虚拟机从一个主机移动到另一个主机，并且这些主机连接到两个不同的枝叶交换机，则APIC将侦听VM迁移事件，在按需部署即时性场景中，取消源枝叶上的EPG编程，并在目标枝叶上编程EPG。

根据与VMM域关联的EPG的部署即时性，无法从vCenter提取资产可能会导致不良后果。在资产未能完成或部分完成的情况下，始终会引发错误，指示导致故障的一个或多个对象。

场景1 — 支持无效的虚拟机：

如果虚拟机从一个vCenter移动到另一个vCenter，或者确定虚拟机具有无效的支持（例如，连接到旧/已删除DVS的端口组），vNIC将报告存在操作问题。

```
Fault fltCompVNicOperationalIssues
Rule ID:2842
Explanation:
This fault is raised when ACI controller failed to update the properties of a VNIC (e.g., it can
not find the EPG that the VNIC attached to).
Code: F2842
Message: Operational issues detected for VNic name on VM name in VMM controller: hostOrIp with
name name in datacenter rootContName in domain: domName due to error: issues.
Resolution:
Remediate the virtual machines indicated in the fault by assigning a valid port group on the
affected vNIC of the VM.
```

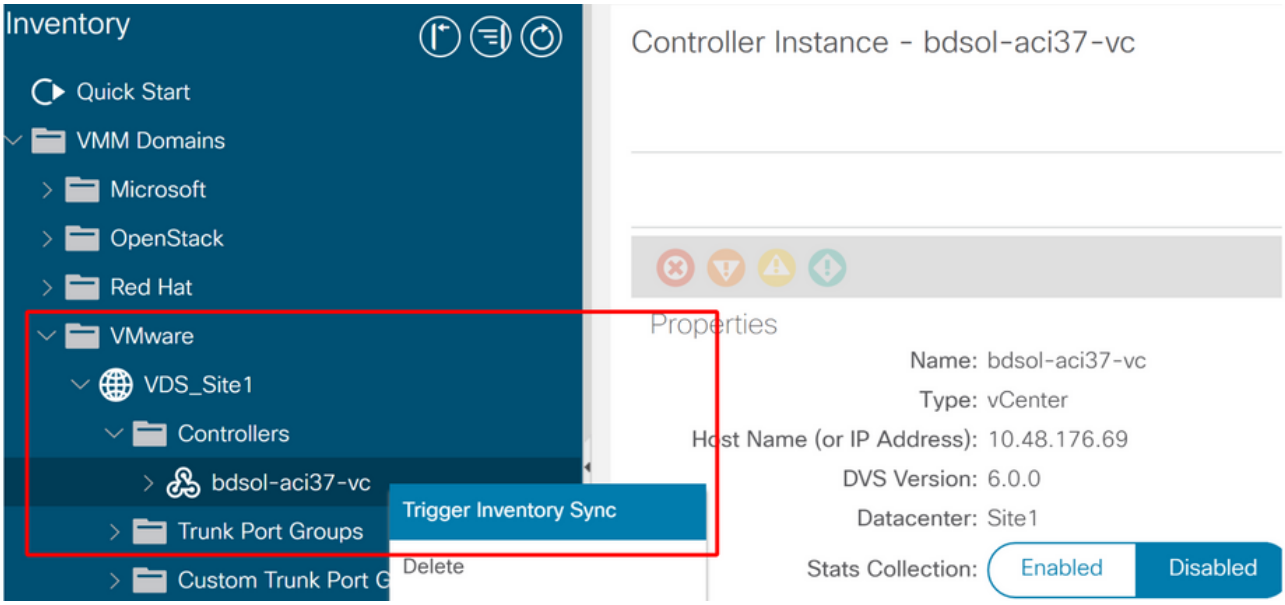
场景2 — vCenter管理员修改了vCenter上的VMM托管对象：

不支持从vCenter修改APIC管理的对象。如果在vCenter上执行了不受支持的操作，则会出现以下故

障。

```
Fault fltCompCtrlrUnsupportedOperation
Rule ID:133
Explanation:
This fault is raised when deployment of given configuration fails for a Controller.
Code: F0133
Message: Unsupported remote operation on controller: hostOrIp with name name in datacenter
rootContName in domain domName detected, error: [deployIssues]
Resolution:
If this scenario is encountered, try to undo the unsupported change in vCenter and then trigger
an 'inventory sync' manually.
```

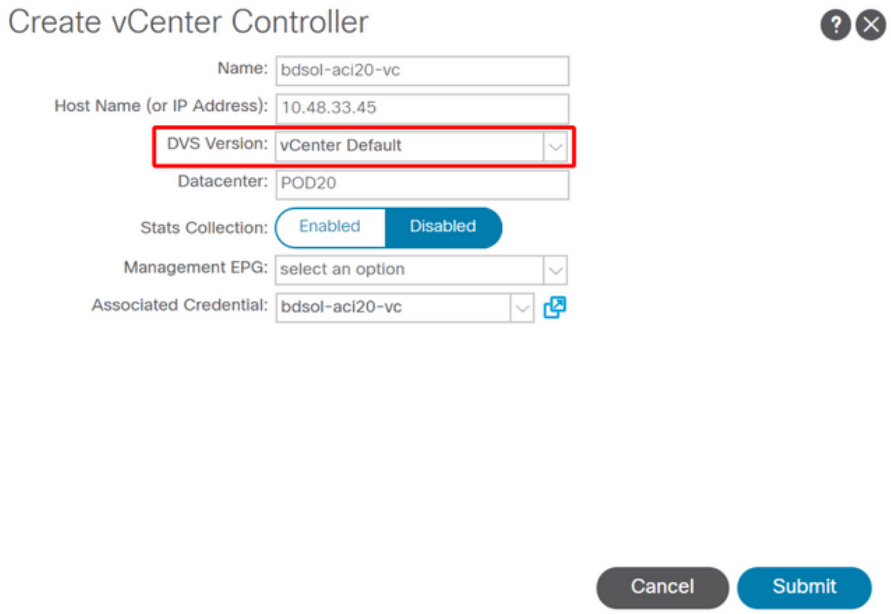
VMWare VMM域 — vCenter控制器 — 触发资产同步



VMware DVS版本

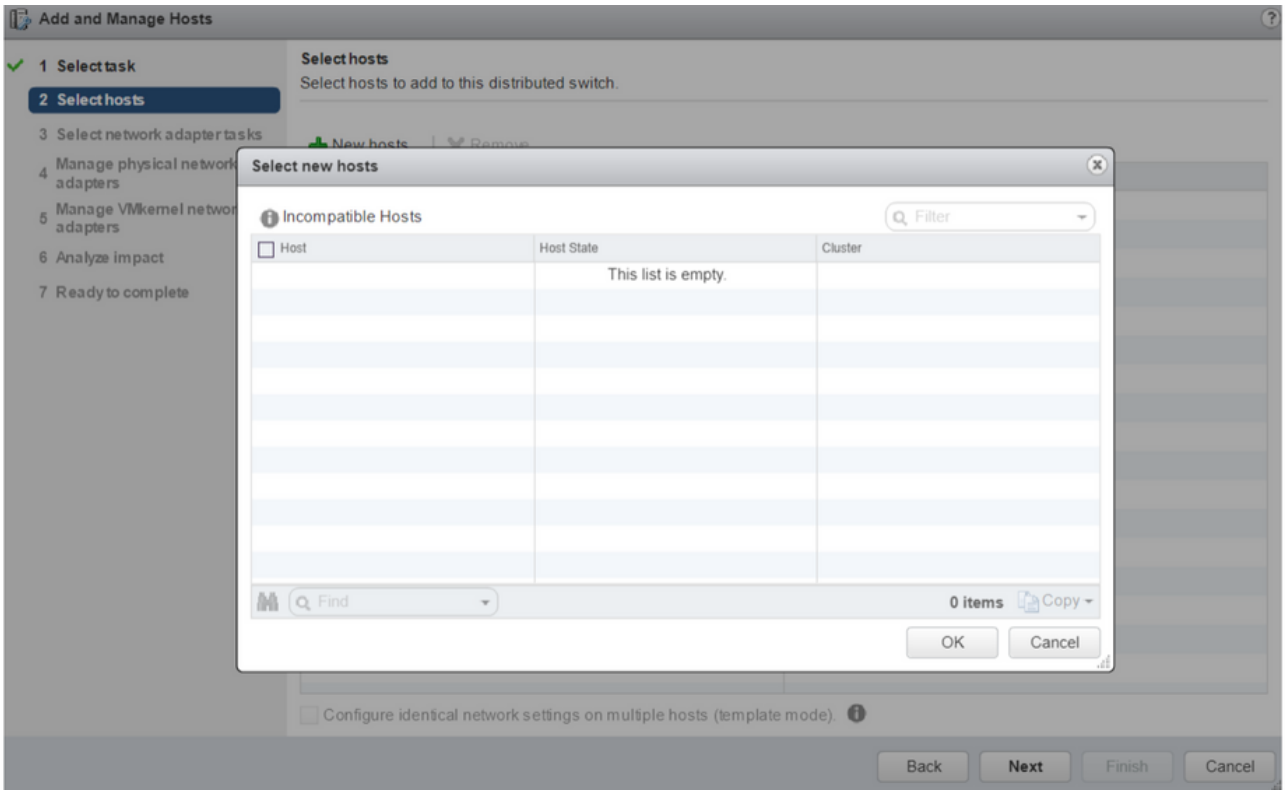
当创建作为VMM域一部分的新vCenter控制器时，DVS版本的默认设置将是使用“vCenter Default”。选择此项时，将使用vCenter版本创建DVS版本。

VMWare VMM域 — vCenter控制器创建

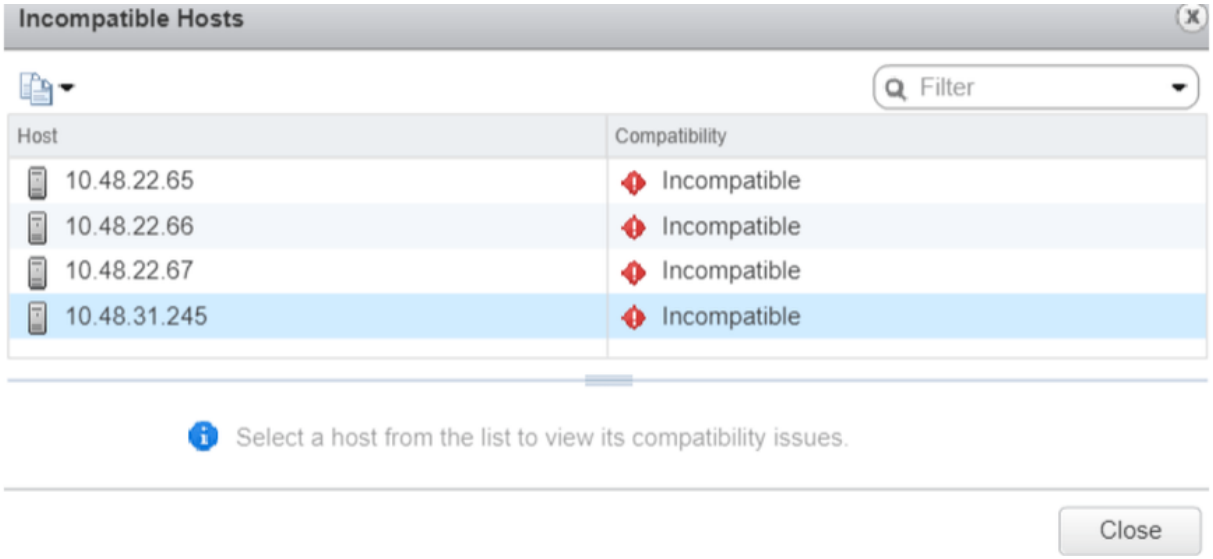


这意味着，在运行6.5的vCenter和运行6.0的ESXi服务器的示例中，APIC将创建版本为6.5的DVS，因此vCenter管理员无法将运行6.0的ESXi服务器添加到ACI DVS中。

APIC托管DVS - vCenter主机添加 — 空列表



APIC托管DVS - vCenter主机添加 — 不兼容的主机



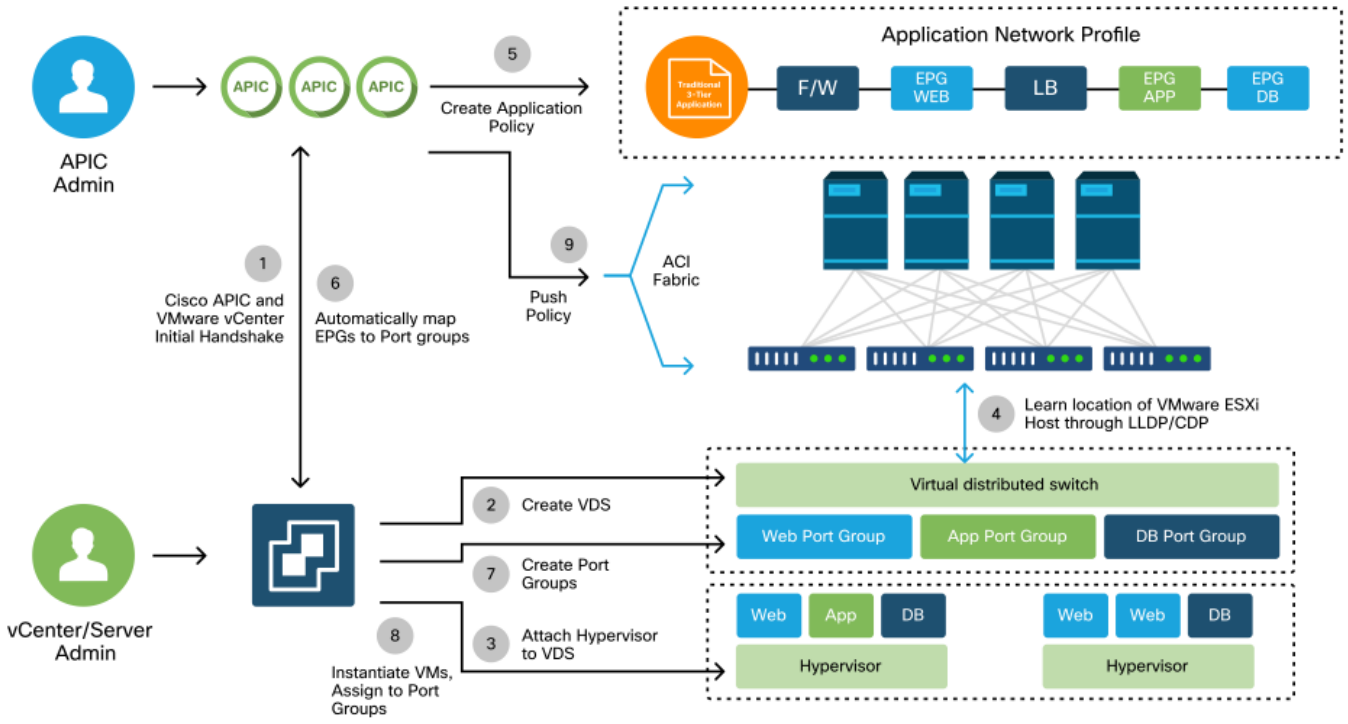
因此，在创建VMM域时，请确保选择正确的“DVS版本”，以便将必要的ESXi服务器添加到DVS。

主机动态发现

主机/虚拟机发现过程

ACI中的VMM集成将自身与手动调配区分开来，因为交换矩阵可以动态发现主机和适用虚拟机所连接的位置，从而高效地部署策略。通过此动态过程，ACI可以优化枝叶交换机上硬件资源的利用率，因为VLAN、SVI、分区规则等只有在连接了需要策略的终端时才部署在节点上。从易用性角度来看，网络管理员的优势在于ACI将调配VLAN/策略，使虚拟机以自动化方式连接。为了确定必须在何处部署策略，APIC将使用来自多个源的信息。下图概述使用基于DVS的VMM域时主机发现过程的基本步骤。

VMWare VMM域 — 部署工作流



简而言之，当出现以下情况时，需要执行以下关键步骤：

- 虚拟机监控程序和枝叶交换机之间交换LLDP或CDP。
- 主机向vCenter报告邻接信息。
- vCenter向APIC通知邻接信息： APIC通过资产同步了解主机。
- APIC将策略推送到枝叶端口： 请复习本部分中的“解决即时性”子部分，进一步了解这些情况。
- 如果vCenter邻接信息丢失，APIC可以删除策略。

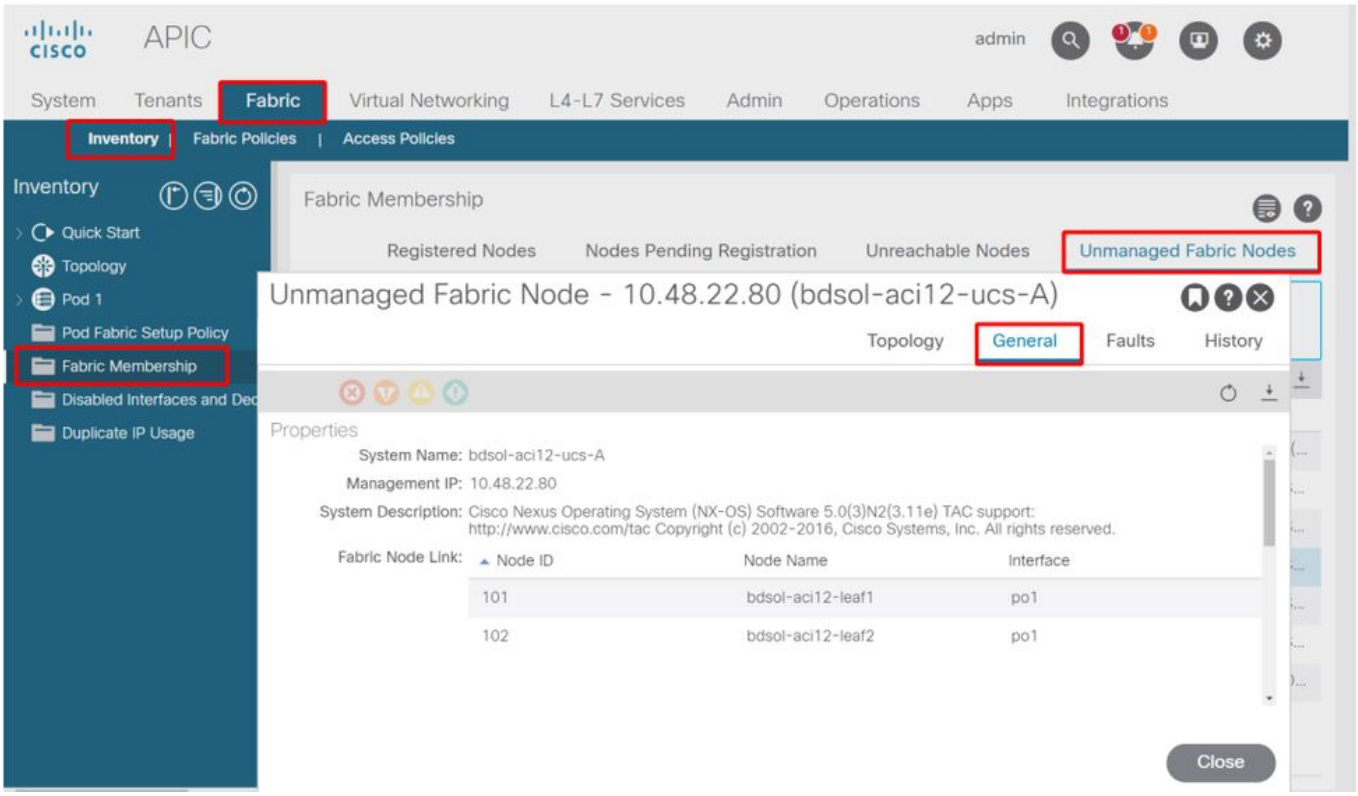
可以看到，CDP/LLDP在发现过程中起着关键作用，必须确保配置正确，并且两端都使用相同的协议。

交换矩阵松动节点/中间交换机 — 使用案例

在使用刀片机箱并在枝叶交换机和虚拟机监控程序之间使用中间交换机的部署中，APIC需要“缝合”邻接关系。在此场景中，可以使用多个发现协议作为中间交换机，这些协议要求可能与主机不同。

在使用刀片服务器和中间交换机（即刀片机箱交换机）的设置中，ACI应检测中间交换机并映射其后的虚拟机监控程序。中间交换机在ACI中称为“松散节点”或“非托管交换矩阵节点”。 检测到的LooseNodes可在“Fabric > Inventory > Fabric Membership > Unmanaged Fabric Nodes”下查看。通过在GUI中导航至这些类型的服务器之一，用户可以查看从枝叶到中间交换机再到主机的路径。

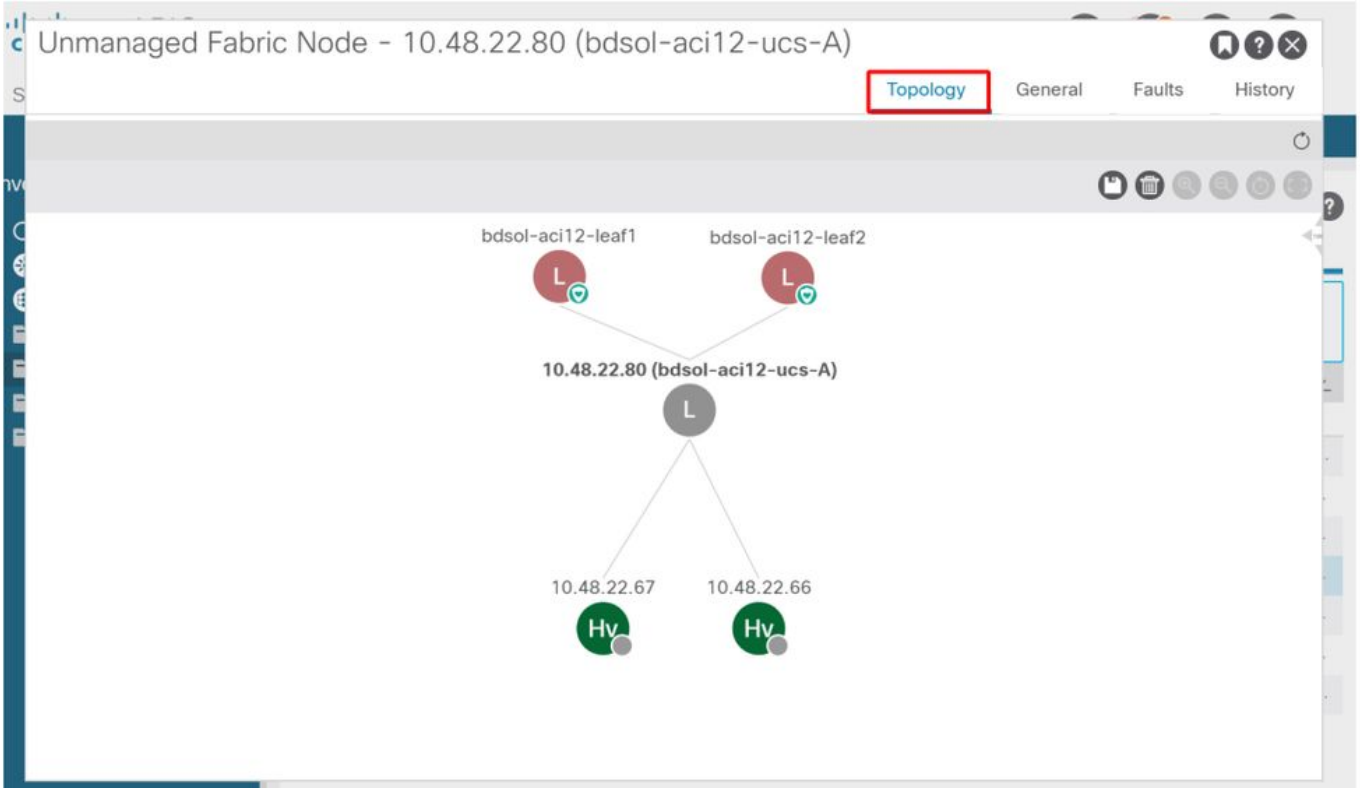
APIC UI — 非托管交换矩阵节点(LooseNodes)



如果部署了LLDP或CDP发现，ACI可以确定此类LooseNodes的拓扑，前提是中间交换机下游的虚拟机监控程序通过VMM集成进行管理，并且枝叶自身与下游的中间交换机具有邻接关系。

下图说明了这一概念。

APIC UI — 非托管交换矩阵节点路径

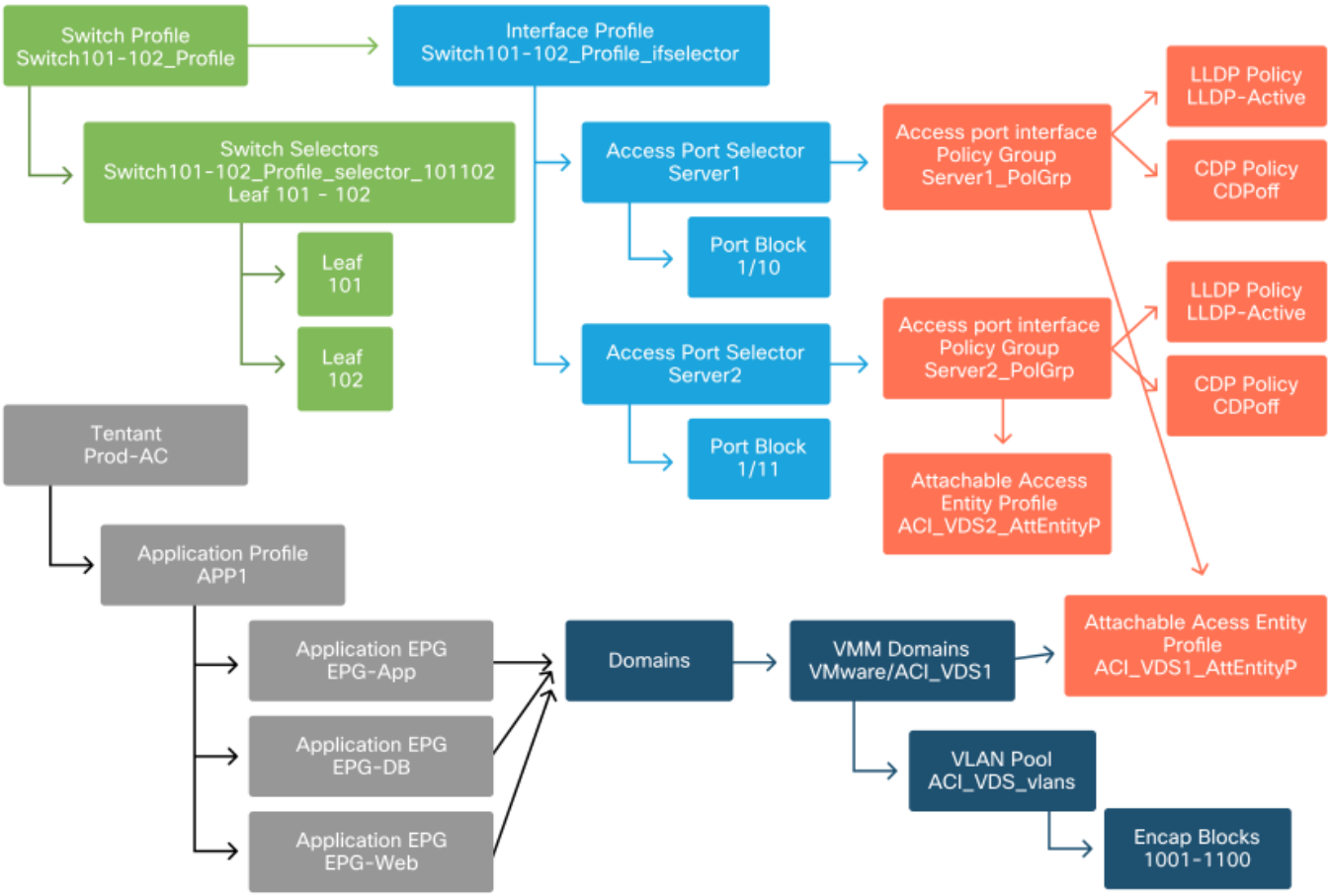


解决即时性

在关键服务利用VMM集成DVS的情形中，例如与vCenter/ESXi的管理连接，谨慎使用预调配解决方案即时性。通过此设置，动态主机发现机制将被删除，而是在面向主机的接口上静态编程策略/VLAN。在此配置中，VMM VLAN将始终部署到与VMM域引用的AEP关联的所有接口。这消除了由于发现协议相关的邻接事件而从端口删除关键VLAN（例如管理）的可能性。

请参阅下图：

预调配部署示例



如果为ACI_VDS1 VMM域中的EPG设置了预调配，则将在服务器1的链路上部署VLAN，但不在服务器2的链路上部署VLAN，因为服务器2的AEP不包括ACI_VDS1 VMM域。

要汇总分辨率立即设置，请执行以下操作：

- 按需 — 在枝叶和主机以及连接到端口组的VM之间建立邻接关系时部署策略。
- 立即 — 在枝叶和主机之间建立邻接关系时部署策略。
- 预调配 — 使用包含VMM域的AEP将策略部署到所有端口，无需邻接关系。

故障排除情况

VM无法为其默认网关解析ARP

在此方案中，已配置VMM集成，并且已将DVS添加到虚拟机监控程序，但VM无法解析ACI中其网关的ARP。为使虚拟机具有网络连接，请验证已建立邻接关系并部署了VLAN。

首先，用户可以在枝叶上使用“show lldp neighbors”或“show cdp neighbors”检查枝叶是否检测到主机，具体取决于所选协议。

```
Leaf101# show lldp neighbors
Capability codes:
 (R) Router, (B) Bridge, (T) Telephone, (C) DOCSIS Cable Device
 (W) WLAN Access Point, (P) Repeater, (S) Station, (O) Other
Device ID            Local Intf      Hold-time  Capability  Port ID
bdsol-aci37-apic1     Eth1/1          120                    eth2-1
bdsol-aci37-apic2     Eth1/2          120                    eth2-1
```

```
bdsol-aci37-os1       Eth1/11         180        B           0050.565a.55a7
S1P1-Spine201         Eth1/49         120        BR          Eth1/1
S1P1-Spine202         Eth1/50         120        BR          Eth1/1
Total entries displayed: 5
```

从故障排除的角度来看，如果需要，可以从ESXi端在CLI和GUI上进行验证：

```
[root@host:~] esxcli network vswitch dvs vmware list
VDS_Site1
  Name: VDS_Site1
 ...
  Uplinks: vmnic7, vmnic6
  VMware Branded: true
  DVPort:
        Client: vmnic6
        DVPortgroup ID: dvportgroup-122
        In Use: true
        Port ID: 0

        Client: vmnic7
        DVPortgroup ID: dvportgroup-122
        In Use: true
        Port ID: 1

[root@host:~] esxcfg-nics  -l
Name    PCI          Driver      Link Speed     Duplex MAC Address       MTU    Description
vmnic6  0000:09:00.0 enic        Up   10000Mbps Full   4c:77:6d:49:cf:30 9000   Cisco Systems
Inc Cisco VIC Ethernet NIC
vmnic7  0000:0a:00.0 enic        Up   10000Mbps Full   4c:77:6d:49:cf:31 9000   Cisco Systems
Inc Cisco VIC Ethernet NIC

[root@host:~] vim-cmd hostsvc/net/query_networkhint --pnic-name=vmnic6 | grep -A2 "System Name"
              key = "System Name",
              value = "Leaf101"
           }
```

vCenter Web客户端 — 主机 — vmnic LLDP/CDP邻接详细信息

vmnic6

All | Properties | CDP | **LLDP**

Link Layer Discovery Protocol

Chassis ID	00:3a:9c:45:12:6b
Port ID	Eth1/11
Time to live	109
TimeOut	60
Samples	437068
Management Address	10.48.176.70
Port Description	topology/pod-1/paths-101/pathep-[eth1/11]
System Description	topology/pod-1/node-101
System Name	S1P1-Leaf101

Peer device capability

Router	Enabled
Transparent bridge	Enabled
Source route bridge	Disabled
Network switch	Disabled
Host	Disabled
IGMP	Disabled
Repeater	Disabled

如果从ESXi主机看不到枝叶LLDP邻接关系，这通常是由使用配置为生成LLDPDU而不是ESXi OS的网络适配器引起的。确保验证网络适配器是否启用了LLDP，从而消耗所有LLDP信息。如果是这种情况，请务必禁用适配器本身上的LLDP，以便通过vSwitch策略对其进行控制。

另一个原因可能是枝叶和ESXi虚拟机监控程序之间使用的发现协议之间不一致。确保在两端使用相同的发现协议。

要检查ACI和APIC UI中的DVS之间是否对CDP/LLDP设置进行了调整，请导航到“虚拟网络> VMM域> VMWare >策略> vSwitch策略”。 确保仅启用LLDP或CDP策略，因为它们是互斥的。

APIC UI - VMWare VMM域 — vSwitch策略

Properties

Port Channel Policy: VDS_lacpLagPol

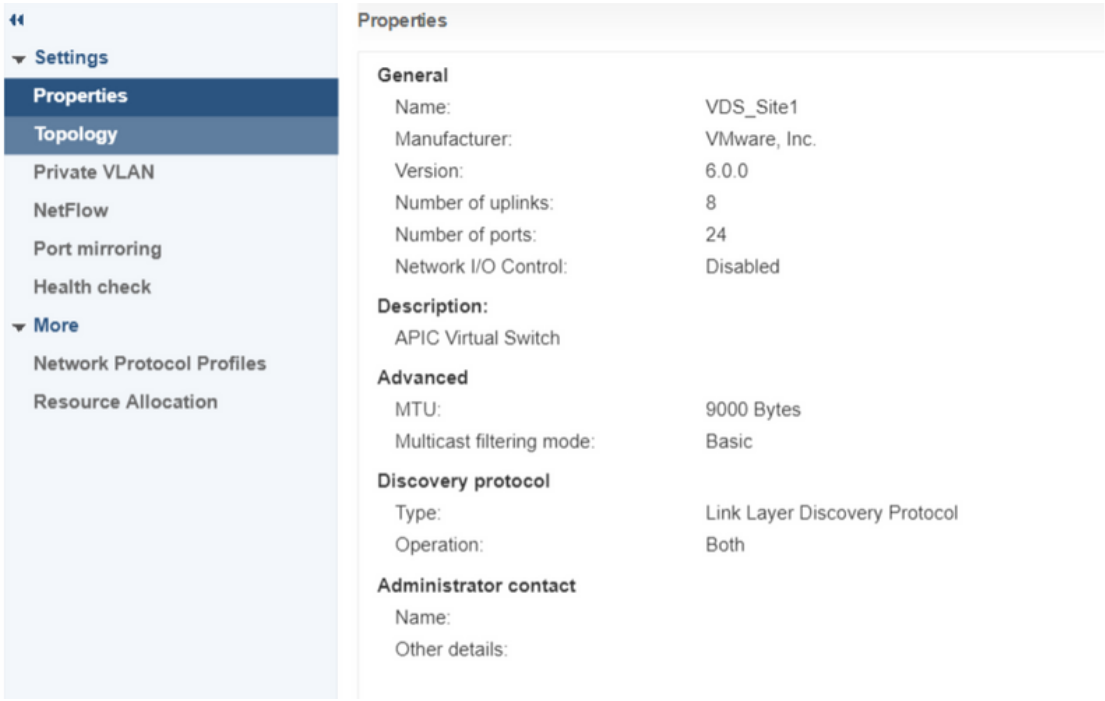
LLDP Policy: LLDP_enabled

CDP Policy: CDP_disabled

NetFlow Exporter Policy: select an option

在vCenter中，转到：'网络> VDS >配置'。

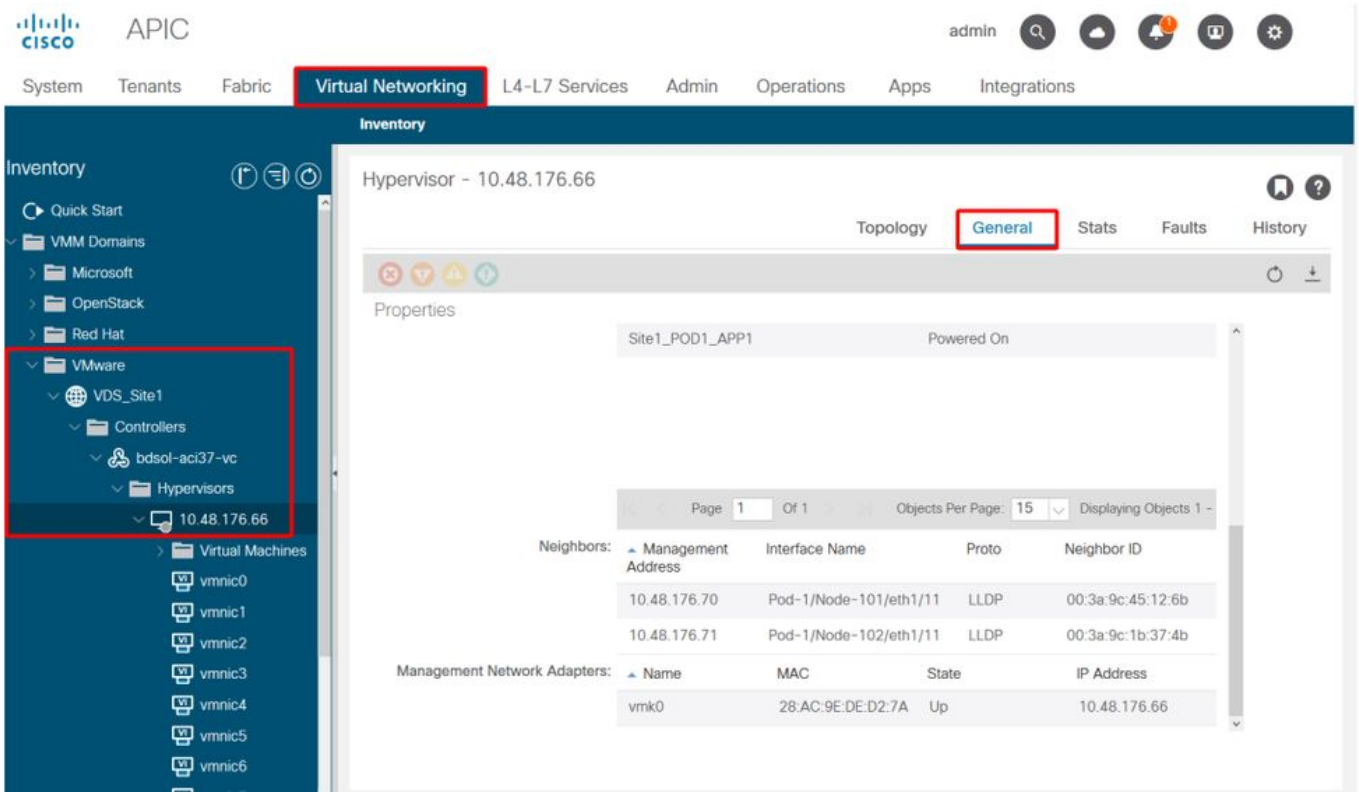
vCenter Web客户端UI - VDS属性



如果需要，请更正LLDP/CDP设置。

然后验证APIC是否在UI中的“Virtual Networking > VMM Domains > VMWare > Policy > Controller > Hypervisor > General”下针对枝叶交换机观察到ESXi主机的LLDP/CDP邻居关系。

APIC UI - VMWare VMM域 — 虚拟机监控程序详细信息



如果显示预期值，则用户可以验证指向主机的端口上是否存在VLAN。

```
S1P1-Leaf101# show vlan encap-id 1035
```

```
VLAN Name                              Status    Ports
---- --------------------------------- --------- -------------------------------
12   Ecommerce:Electronics:APP         active    Eth1/11

VLAN Type  Vlan-mode
---- ----- ----------
12   enet  CE
```

vCenter/ESXi管理VMK连接到APIC推送的DVS

在vCenter或ESXi管理流量需要使用VMM集成DVS的场景中，必须格外小心，以避免在激活动态邻接和激活所需的VLAN时陷入僵局。

对于vCenter（通常在配置VMM集成之前构建），必须使用物理域和静态路径确保vCenter VM的封装VLAN始终在枝叶交换机上编程，以便在VMM集成完全设置之前使用。即使在设置VMM集成后，建议将此静态路径保留到位，以确保此EPG的可用性。

对于ESXi虚拟机监控程序，根据Cisco.com上的“思科ACI虚拟化指南”，迁移到vDS时，务必确保部署将连接VMK接口的EPG，且分辨率即时性设置为“预调配”。这将确保VLAN始终在枝叶交换机上编程，而不依赖ESXi主机的LLDP/CDP发现。

未在LooseNode后发现主机邻接关系

LooseNode发现问题的典型原因如下：

- CDP/LLDP未启用 中间交换机、枝叶交换机和ESXi主机之间必须交换CDP/LLDP对于Cisco UCS，这通过vNIC上的网络控制策略来实现
- LLDP/CDP邻居的管理IP发生更改会中断连接 vCenter将在LLDP/CDP邻接中看到新的管理IP，但不会更新APIC触发手动库存同步修复
- VMM VLAN未添加到中间交换机 APIC不编程第三方刀片/中间交换机。4.1(1)版本中提供Cisco UCSM集成应用(ExternalSwitch)。必须配置VLAN并将其中继到连接到ACI枝叶节点的上行链路和连接到主机的下行链路

F606391 — 主机上的物理适配器缺少邻接关系

当看到以下故障时：

```
Affected Object: comp/prov-VMware/ctrlr-[DVS-DC1-ACI-LAB]-DVS1/hv-host-104
Fault delegate: [FSM:FAILED]: Get LLDP/CDP adjacency information for the physical adapters on
the host: bdsol-aci20-os3 (TASK:ifc:vmmmgr:CompHvGetHpNicAdj)
```

请查看“VM cannot resolve ARP for its default gateway”部分的工作流程，因为这意味着缺少CDP/LLDP邻接关系。应端到端检验这些邻接关系。

虚拟机监控程序上行链路负载均衡

当将虚拟机监控程序（如ESXi）连接到ACI交换矩阵时，它们通常使用多个上行链路进行连接。事实上，建议将ESXi主机连接到至少两个枝叶交换机。这将最大限度地降低故障场景或升级的影响。

为了优化在虚拟机监控程序上运行的工作负载使用上行链路的方式，VMware vCenter配置允许为虚拟机生成的流量配置多个负载均衡算法，以流向虚拟机监控程序上行链路。

要使所有虚拟机监控程序和ACI交换矩阵与相同的负载均衡算法配置保持一致，以确保建立正确的连接，这一点至关重要。否则可能会导致ACI交换矩阵中的流量间歇性丢弃和终端移动。

在ACI交换矩阵中，可以通过过度警报来查看这一点，例如：

```
F3083 fault
ACI has detected multiple MACs using the same IP address 172.16.202.237.
MACs: Context: 2981888. fvCEps:
uni/tn-BSE_PROD/ap-202_Voice/epg-VLAN202_Voice/cep-00:50:56:9D:55:B2;
uni/tn-BSE_PROD/ap-202_Voice/epg-VLAN202_Voice/cep-00:50:56:9D:B7:01;
or
[F1197][raised][bd-limits-exceeded][major][sys/ctx-[vxlan-2818048]/bd-[vxlan-16252885]/fault-
F1197]
Learning is disabled on BD Ecommerce:BD01
```

本章将介绍VMWare ESXi主机到ACI的连接，但它适用于大多数虚拟机监控程序。

机架式服务器

在查看ESXi主机可以连接到ACI交换矩阵的各种方式时，它们分为两组，即与交换机相关的负载平衡算法和独立于交换机的负载平衡算法。

交换机独立负载均衡算法是一种无需特定交换机配置的连接方式。对于与交换机相关的负载均衡，需要特定于交换机的配置。

确保根据下表验证vSwitch策略是否符合“ACI访问策略组”要求。

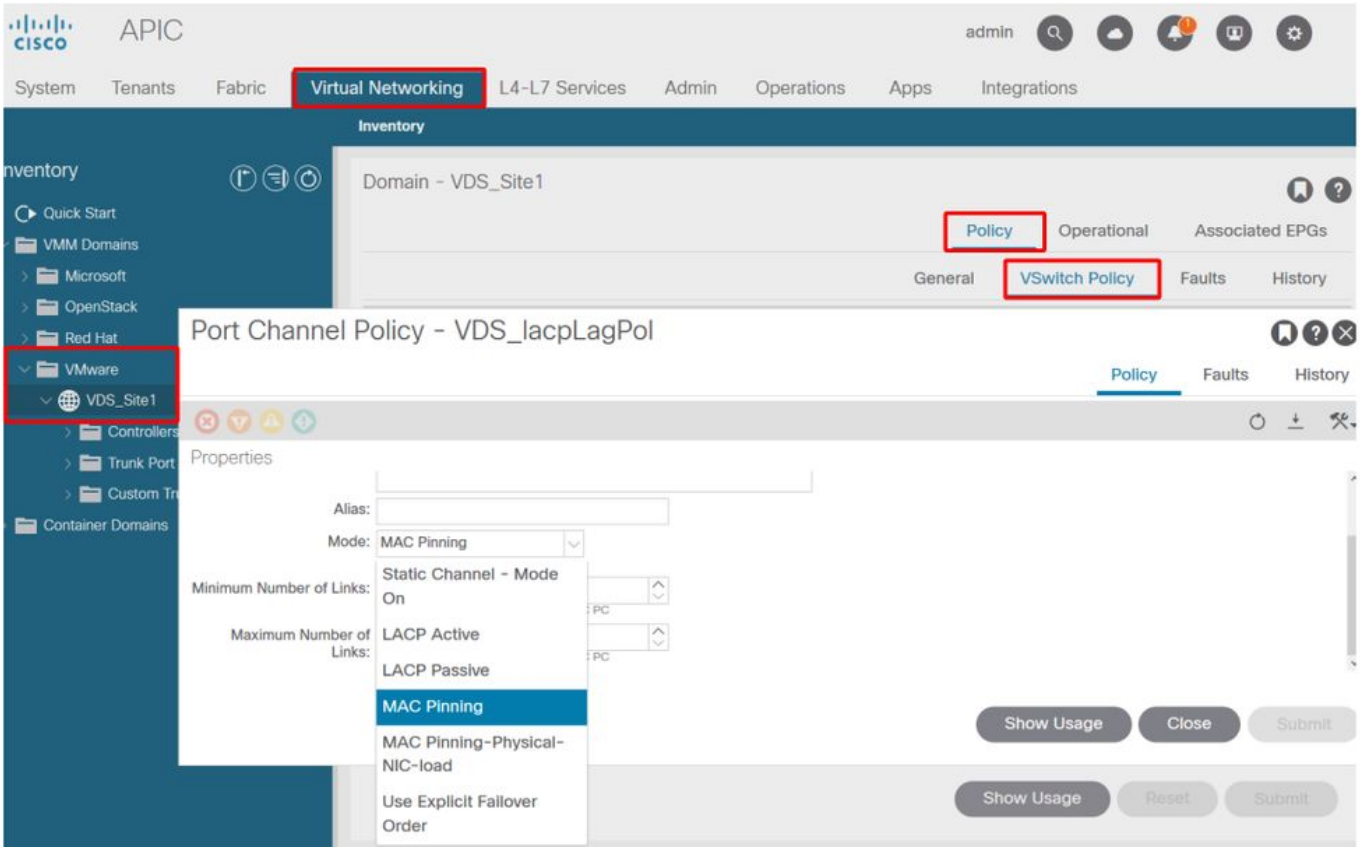
团队和ACI vSwitch策略

VMware组合和故障切换模式	ACI vSwitch策略	描述	ACI访问策略组 — 需要端口通道
基于源虚拟端口的路由	MAC定位	根据交换机上的虚拟端口ID选择上行链路。在虚拟交换机为虚拟机或VMKernel适配器选择上行链路后，它始终通过与此虚拟机或VMKernel适配器相同的上行链路转发流量。	无
基于源MAC散列的路由	不适用	根据源MAC地址的哈希值选择上行链路	不适用
显式故障切换顺序	使用显式故障切换模式	从活动适配器列表中，始终使用通过故障切换检测标准的最高级别上行链路。使用此选项不会执行任何实际的负载均衡。	无
链路聚合(LAG) — 基于IP散列	静态通道 — 模式打开	根据每个数据包的源IP地址和目标IP地址的散列值选择上行链路。对于非IP数据包，交换机使用这些字段的数据计算散列。基于IP的组合要求在ACI端端口通道/VPC配置为“mode on”。	是（信道模式设置为“开”）
链路聚合(LAG)- LACP	LACP主动/被动	根据选定的散列值选择上行链路（20个不同的可用散列值选项）。基于LACP的分组要求在ACI端端口通道/VPC上配置启用LACP。确保在ACI中创建增强型Lag策略并将其应用于VSwitch策略。	是（通道模式设置为“LACP主动/被动”）
基于物理网卡负载(LBT)的路	MAC固定 — 物理NIC负载	可用于分布式端口组或分布式端口。根据连接到端口组或端口的物理网络适配器的	无

由	当前负载选择上行链路。如果上行链路在30秒内以75%或更高的速率保持忙碌状态，主机的vSwitch会将部分虚拟机流量移动到具有可用容量的物理适配器。

请参阅以下屏幕截图，了解如何验证端口通道策略，以作为vSwitch策略的一部分。

ACI vSwitch策略 — 端口通道策略



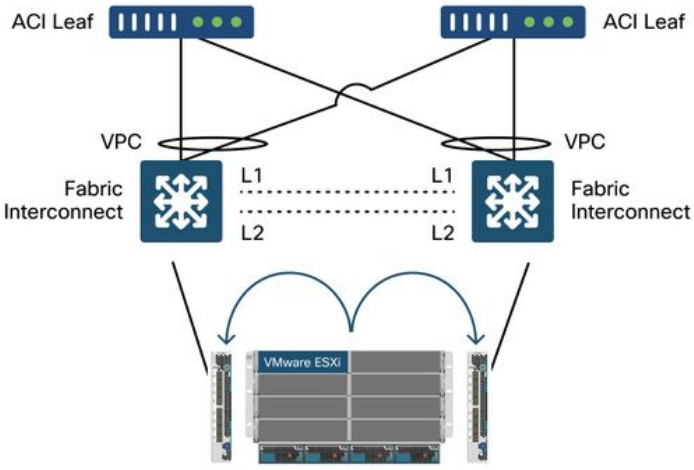
注意：有关VMware网络功能的更详细说明，请参阅vSphere网络，网址为 https://docs.vmware.com/en/VMware-vSphere/6.5/com.vmware.vsphere.networking.doc/GUID-D34B1ADD-B8A7-43CD-AA7E-2832A0F7EE76.html

Cisco UCS B系列使用案例

使用Cisco UCS B系列服务器时，请务必注意它们在机箱内连接到没有统一数据平面的UCS交换矩阵互联(FI)。此使用案例同样适用于采用类似拓扑的其他供应商。因此，从ACI枝叶交换机端和vSwitch端使用的负载均衡方法可能会有所不同。

以下是带ACI的UCS FI拓扑：

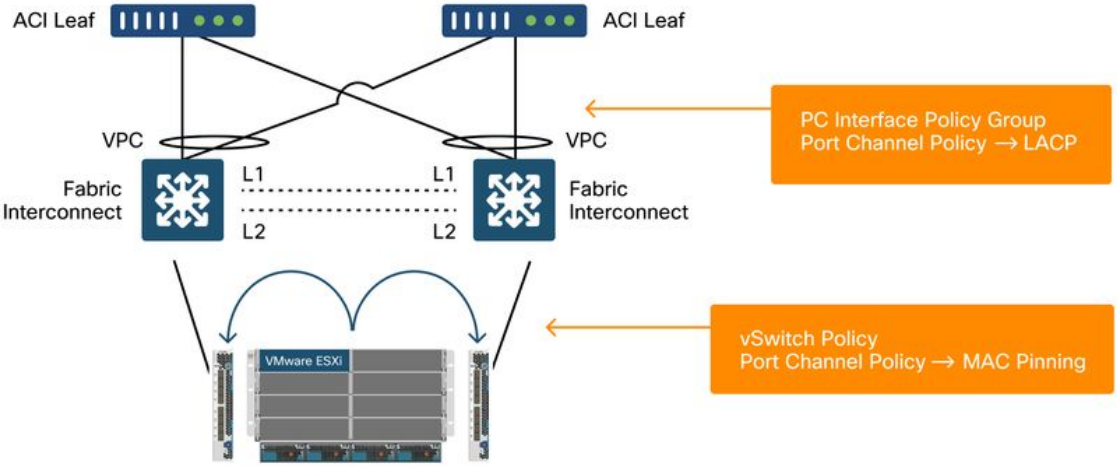
带ACI枝叶交换机的Cisco UCS FI — 拓扑



需要注意的关键事项：

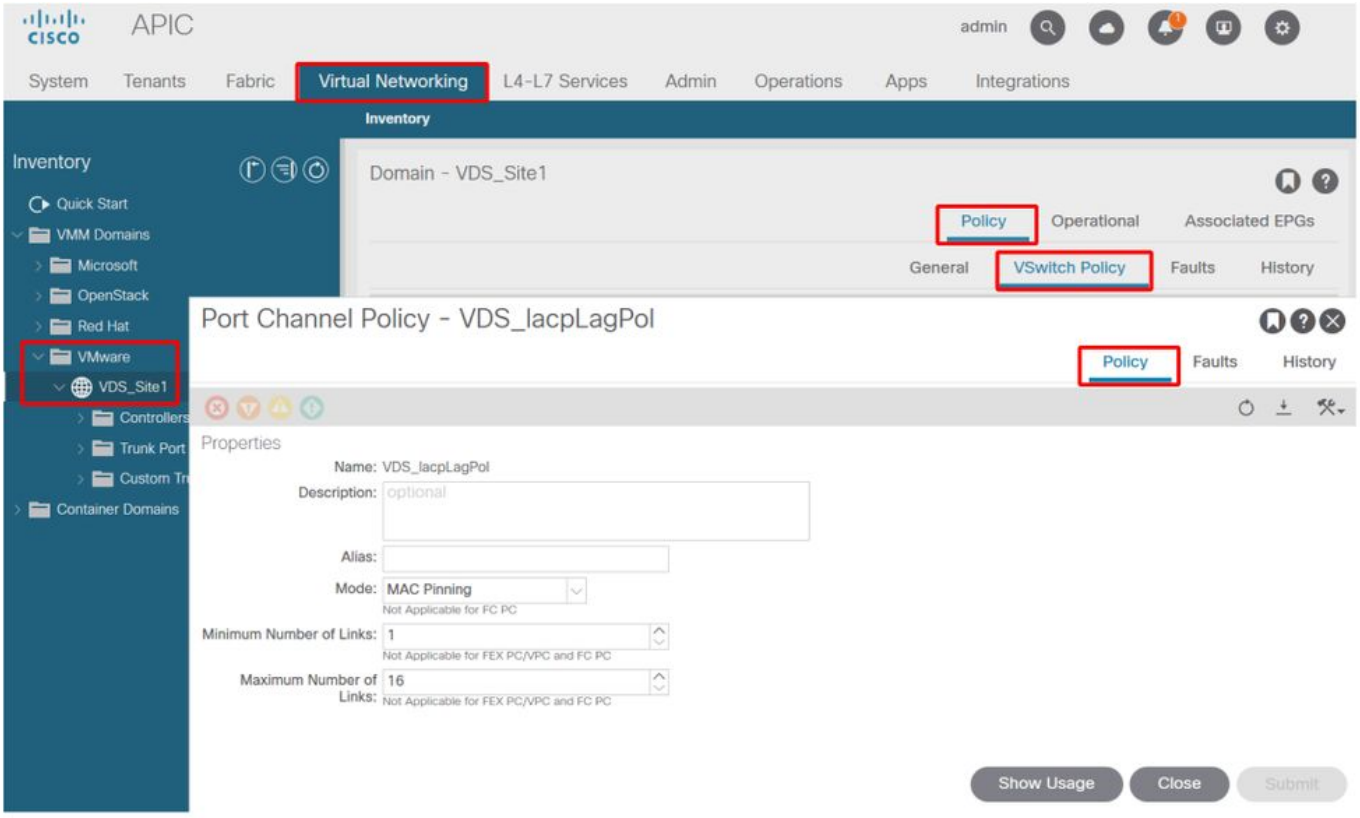
- 每个Cisco UCS FI都有一个指向ACI枝叶交换机的端口通道。
- UCS FI仅出于心跳目的直接互连（不用于数据平面）。
- 每个刀片服务器的vNIC都固定到特定UCS FI或使用UCS交换矩阵故障切换（主用 — 备用）使用通往其中一个FI的路径。
- 在ESXi主机的vSwitch上使用IP哈希算法将导致UCS FI上的MAC摆动。

要正确配置此设置，请执行以下操作：



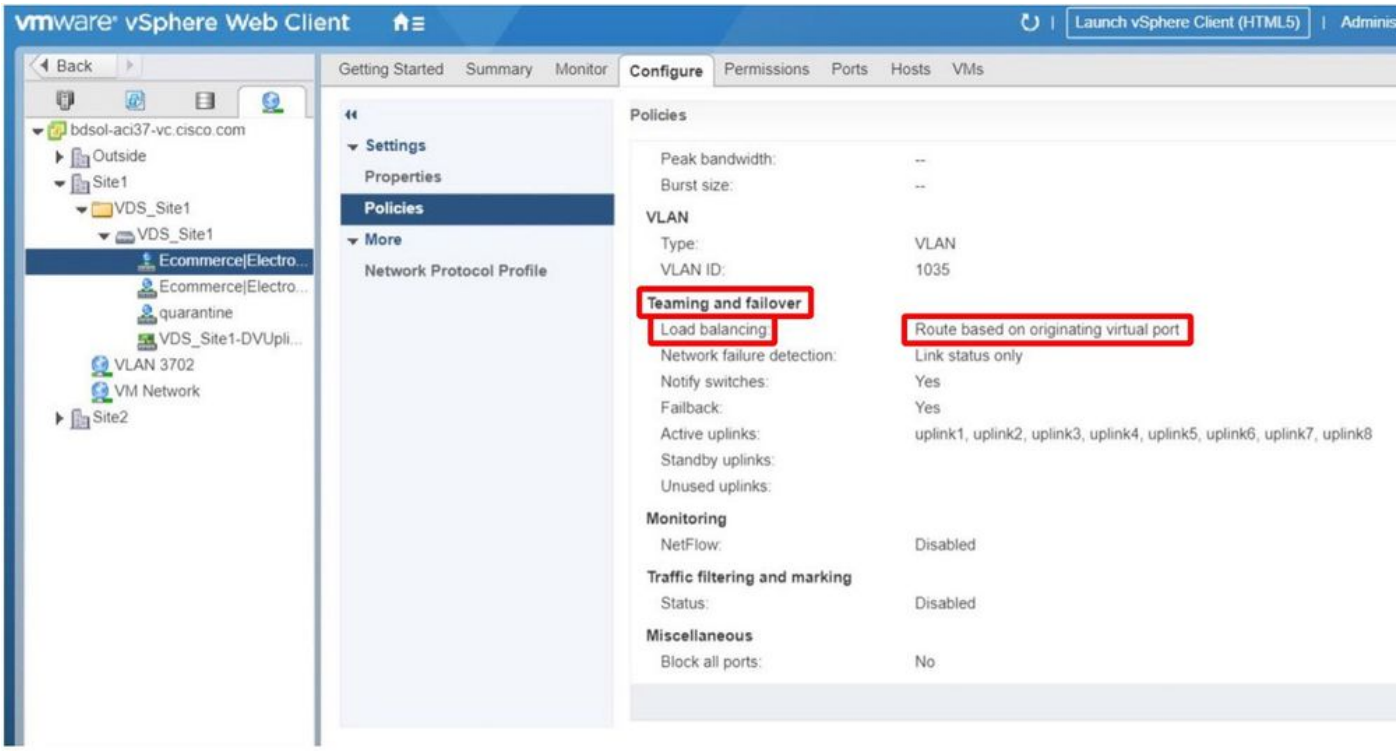
当MAC Pinning作为ACI中的vSwitch策略的一部分在端口通道策略上配置时，将显示为VDS上端口组的“基于源虚拟端口的路由”分组配置。

ACI — 端口通道策略作为vSwitch策略的一部分



上面示例中使用的端口通道策略由向导自动命名，因此它称为“CDS_lacpLagPol”，尽管我们使用模式“MAC Pinning”。

VMWare vCenter — ACI VDS — 端口组 — 负载均衡设置



关于此翻译
思科采用人工翻译与机器翻译相结合的方式将此文档翻译成不同语言，希望全球的用户都能通过各自的语言得到支持性的内容。

请注意：即使是最好的机器翻译，其准确度也不及专业翻译人员的水平。

Cisco Systems, Inc. 对于翻译的准确性不承担任何责任，并建议您总是参考英文原始文档（已提供链接）。