排除ACI管理和核心服务故障 — Pod策略

目录

简介

本文档介绍了解ACI Pod策略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，尤其是管理和核心服务 — **POD策略 — BGP RR/日期和时间/SNMP** 第章。

Pod策略概述

使用Pod策略组在系统上应用BGP RR、日期和时间以及SNMP等管理服务。Pod策略组管理一组与ACI交换矩阵的基本功能相关的Pod策略。这些Pod策略与以下组件相关，其中很多组件默认在ACI交换矩阵中调配。

Pod策略

Pod策略	需要手动配置
日期和时间	Yes
BGP路由反射器	Yes
SNMP（服务器网络管理协议）	Yes
ISIS	无
COOP	无
管理访问	无
MAC安全	Yes

即使在单个ACI交换矩阵中，也需要配置Pod策略组和Pod配置文件。这并不特定于多Pod甚至多站点部署。此要求适用于所有ACI部署类型。

本章重点介绍这些基本的Pod策略，以及如何验证它们是否正确应用。

日期和时间策略

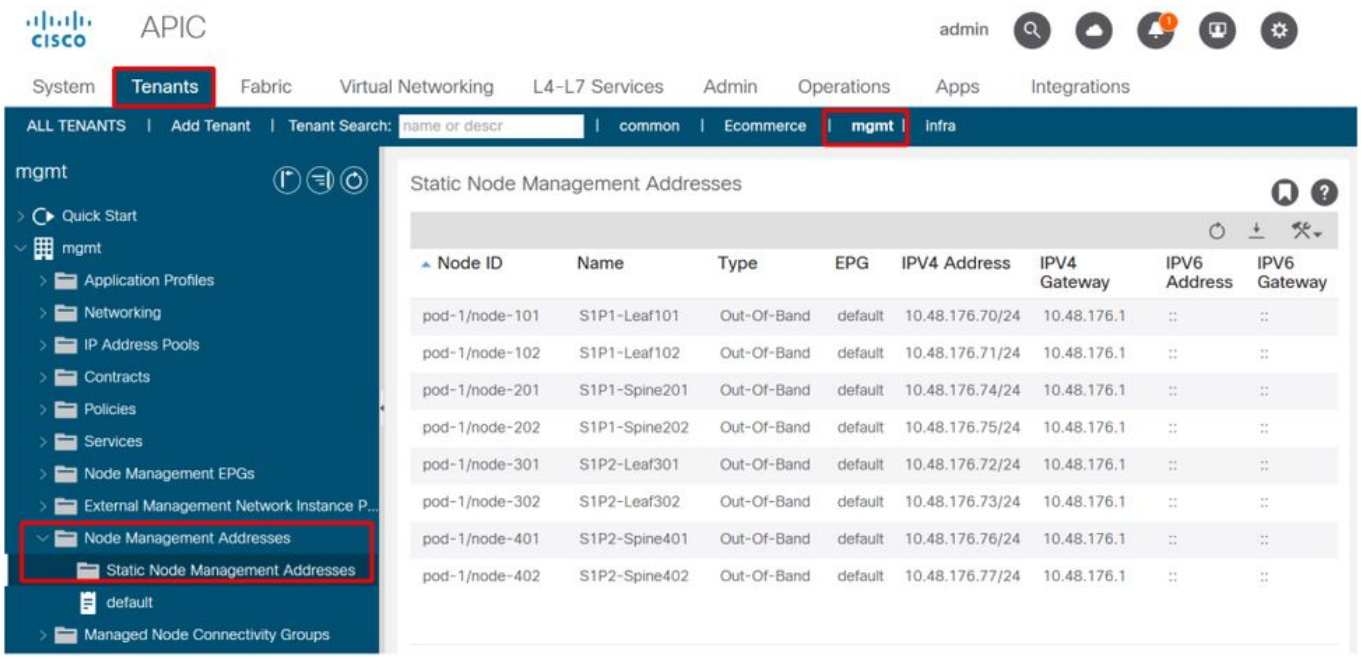
时间同步在ACI交换矩阵中起着关键作用。从验证证书，到保持APIC和交换机中的日志时间戳一致，最佳实践是使用NTP将ACI交换矩阵中的节点同步到一个或多个可靠的时间源。

为了使节点正确同步到NTP服务器提供程序，需要依赖关系为节点分配管理地址。这可以在管理租户下使用静态节点管理地址或管理节点连接组完成。

故障排除工作流程

1.验证是否已将节点管理地址分配给所有节点

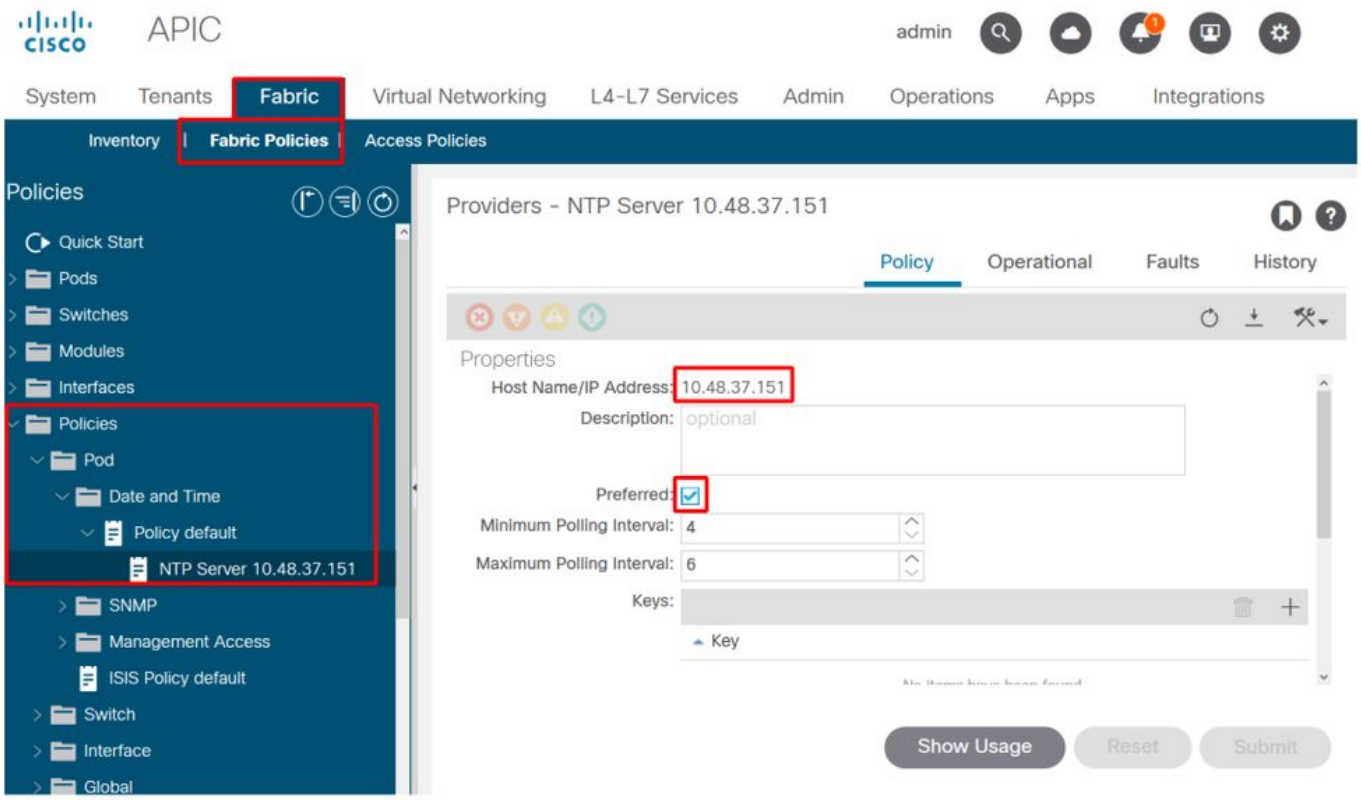
管理租户 — 节点管理地址



2.验证NTP服务器是否已配置为NTP提供程序

如果有多个NTP提供程序，则使用“首选”复选框将其中至少一个提供程序标记为首选时间源，如下图所示。

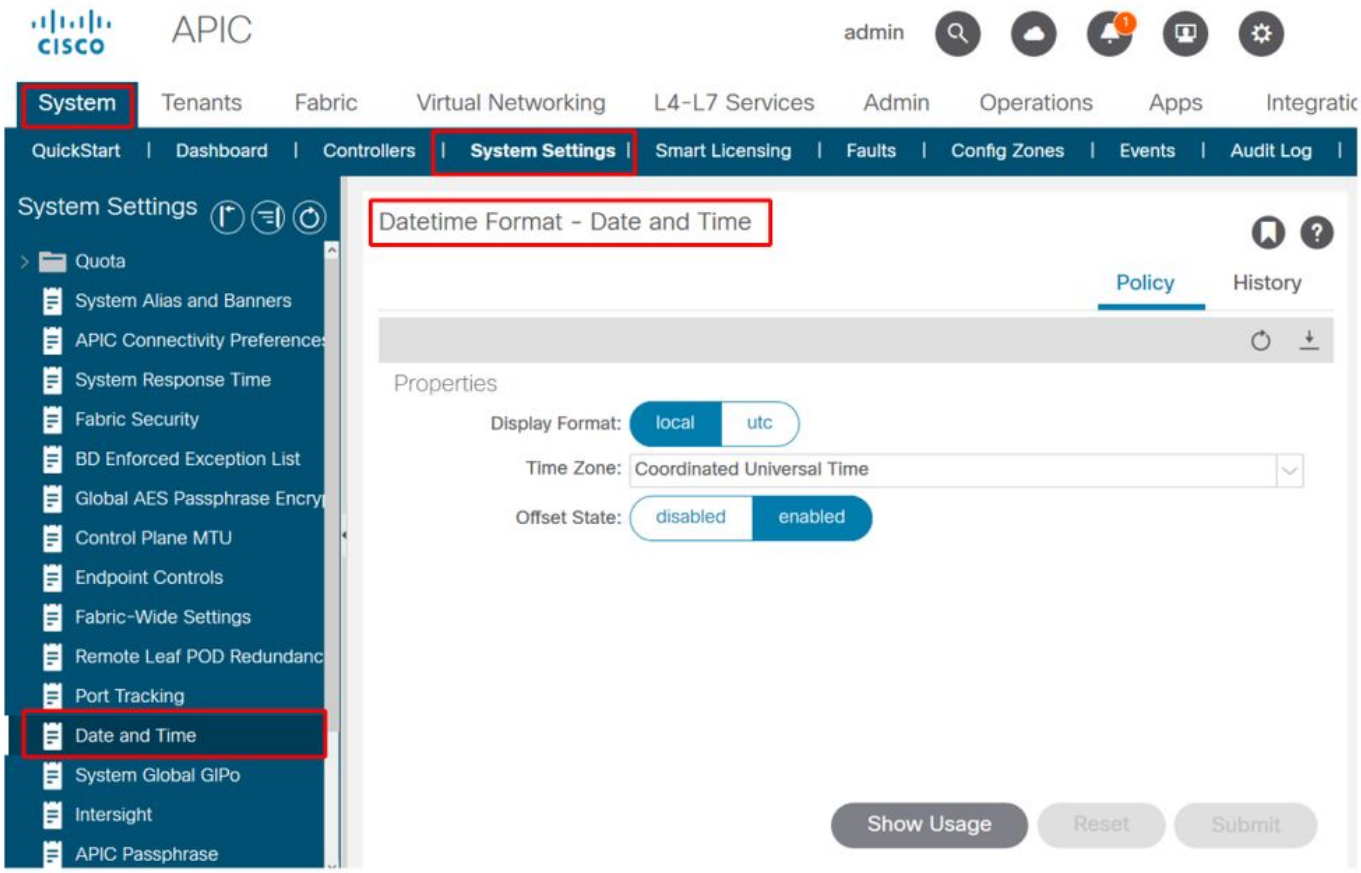
NTP提供商/服务器在日期和时间Pod策略下



3.在“系统设置”下验证日期和时间格式

下图显示了一个示例，其中日期和时间格式已设置为UTC。

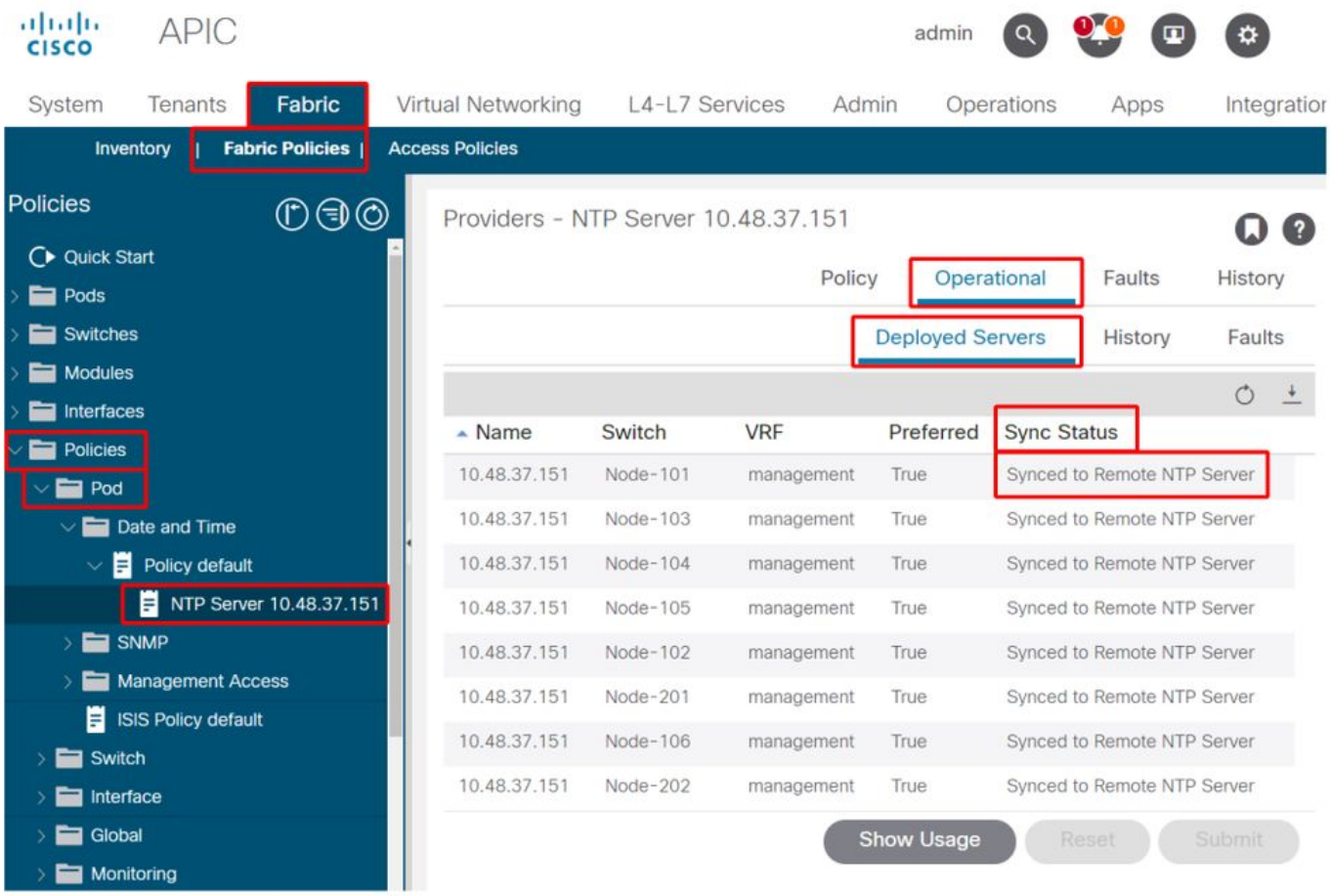
“系统设置”下的日期和时间设置



4.验证所有节点的NTP提供程序的运行同步状态

如下图所示，“同步状态”(Sync Status)列应显示“已同步到远程NTP服务器”(Synced to Remote NTP Server)。 请注意，同步状态正确收敛到.Synced to Remote NTP Server可能需要几分钟时间。状态
.

NTP提供程序/服务器同步状态



或者，可以在APIC和交换机上使用CLI方法验证与NTP服务器的正确时间同步。

APIC - NX-OS CLI

下面的“refId”列根据层级显示NTP服务器下次源。

```
apic1# show ntpq
nodeid     remote                           refid                      st        t   when
poll      reach     auth  delay     offset    jitter
------  -  ------------------------------  -------------------------  --------  --  --------  -
-------  --------  ----  --------  --------  --------
1       *  10.48.37.151                     192.168.1.115              2         u   25
64        377       none  0.214     -0.118    0.025
2       *  10.48.37.151                     192.168.1.115              2         u   62
64        377       none  0.207     -0.085    0.043
3       *  10.48.37.151                     192.168.1.115              2         u   43
64        377       none  0.109     -0.072    0.030


apic1# show clock
Time : 17:38:05.814 UTC Wed Oct 02 2019
```

APIC - Bash

```
apic1# bash
admin@apic1:~> date
Wed Oct  2 17:38:45 UTC 2019
```

交换机

使用“show ntp peers”命令确保NTP提供程序配置已正确推送到交换机。

```
leaf1# show ntp peers
-------------------------------------------------------------------------------
   Peer IP Address                 Serv/Peer Prefer KeyId   Vrf
-------------------------------------------------------------------------------
   10.48.37.151                    Server    yes    None    management

leaf1# show ntp peer-status
Total peers : 1
* - selected for sync, + - peer mode(active),
- - peer mode(passive), = - polled in client mode
   remote                        local           st poll reach delay vrf
-------------------------------------------------------------------------------
*10.48.37.151                     0.0.0.0          2  64   377   0.000  management
```

此处的“*”字符非常重要，因为它控制着NTP服务器是否实际用于同步。

在以下命令中验证发送/接收的数据包数量，以确保ACI节点可连接到NTP服务器。

```
leaf1# show ntp statistics peer ipaddr 10.48.37.151
...
packets sent:         256
packets received:     256
...
```

BGP路由反射器策略

ACI交换矩阵在枝叶和主干节点之间使用多协议BGP(MP-BGP)，更具体地说，使用iBGP VPNv4交换从外部路由器（连接到L3Outs）接收的租户路由。 为了避免全网状iBGP对等拓扑，主干节点将从枝叶收到的VPNv4前缀反射到交换矩阵中的其他枝叶节点。

如果没有BGP路由反射器(BGP RR)策略，将不会在交换机上创建BGP实例，并且不会建立BGP VPNv4会话。在多Pod部署中，每个Pod至少需要一个配置为BGP RR的主干，并且实际上需要多个主干以实现冗余。

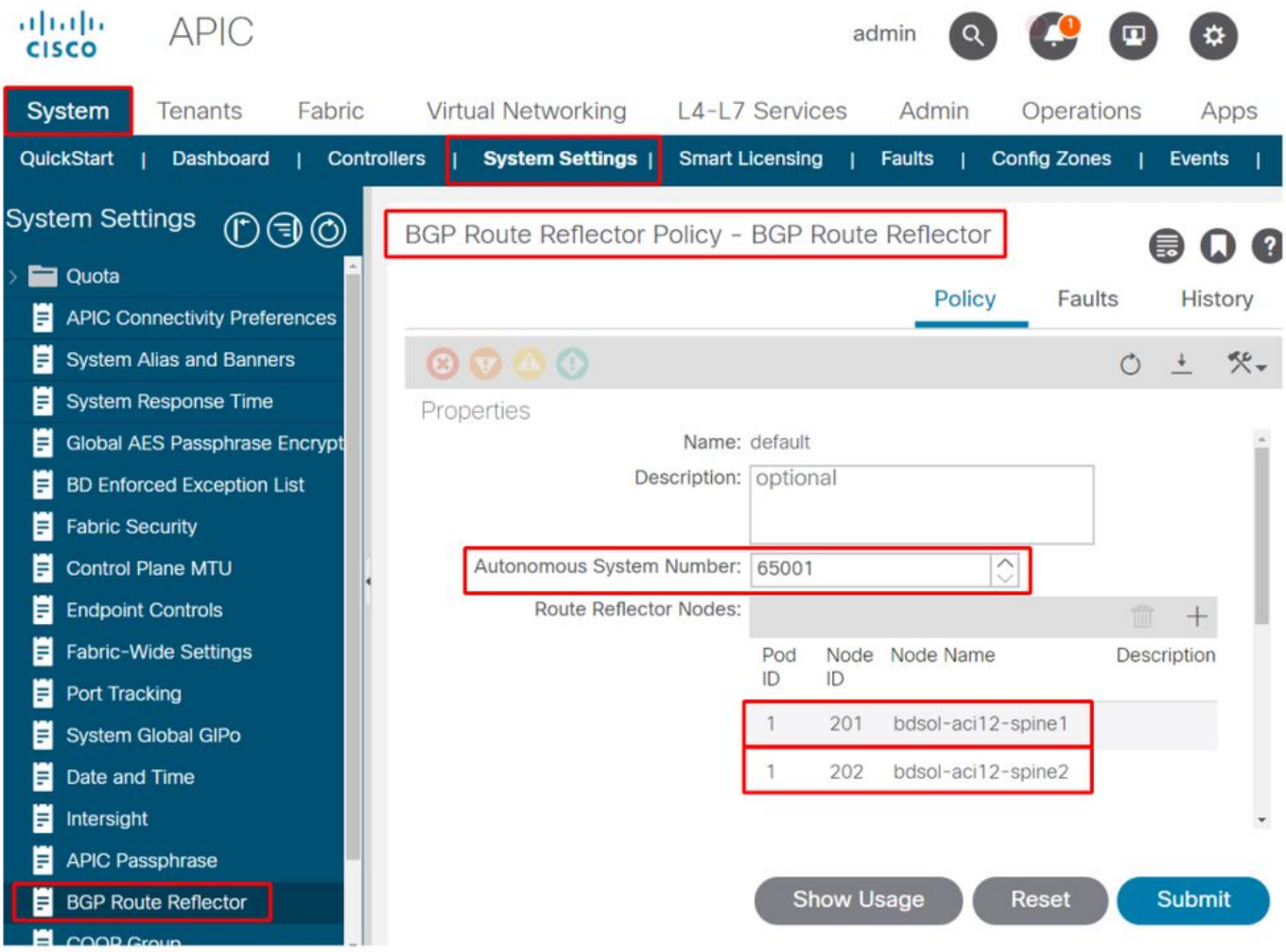
因此，BGP RR策略是每个ACI交换矩阵中一个必要的配置。BGP RR策略还包含ACI交换矩阵用于每台交换机上的BGP进程的ASN。

故障排除工作流程

1.验证BGP RR策略是否配置了ASN和至少一个主干

以下示例涉及单个Pod部署。

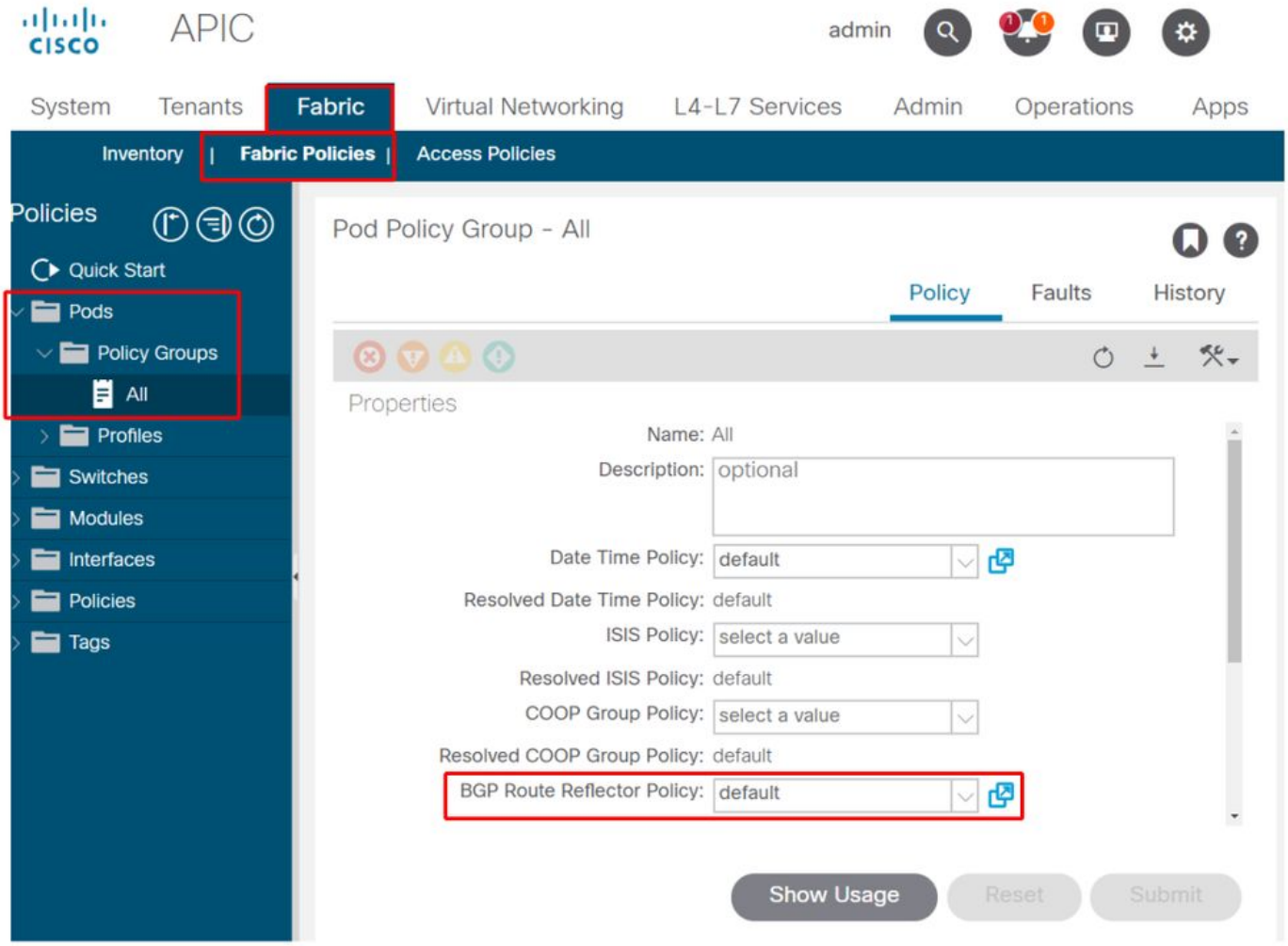
系统设置下的BGP路由反射器策略



2.验证BGP RR策略是否应用在Pod策略组下

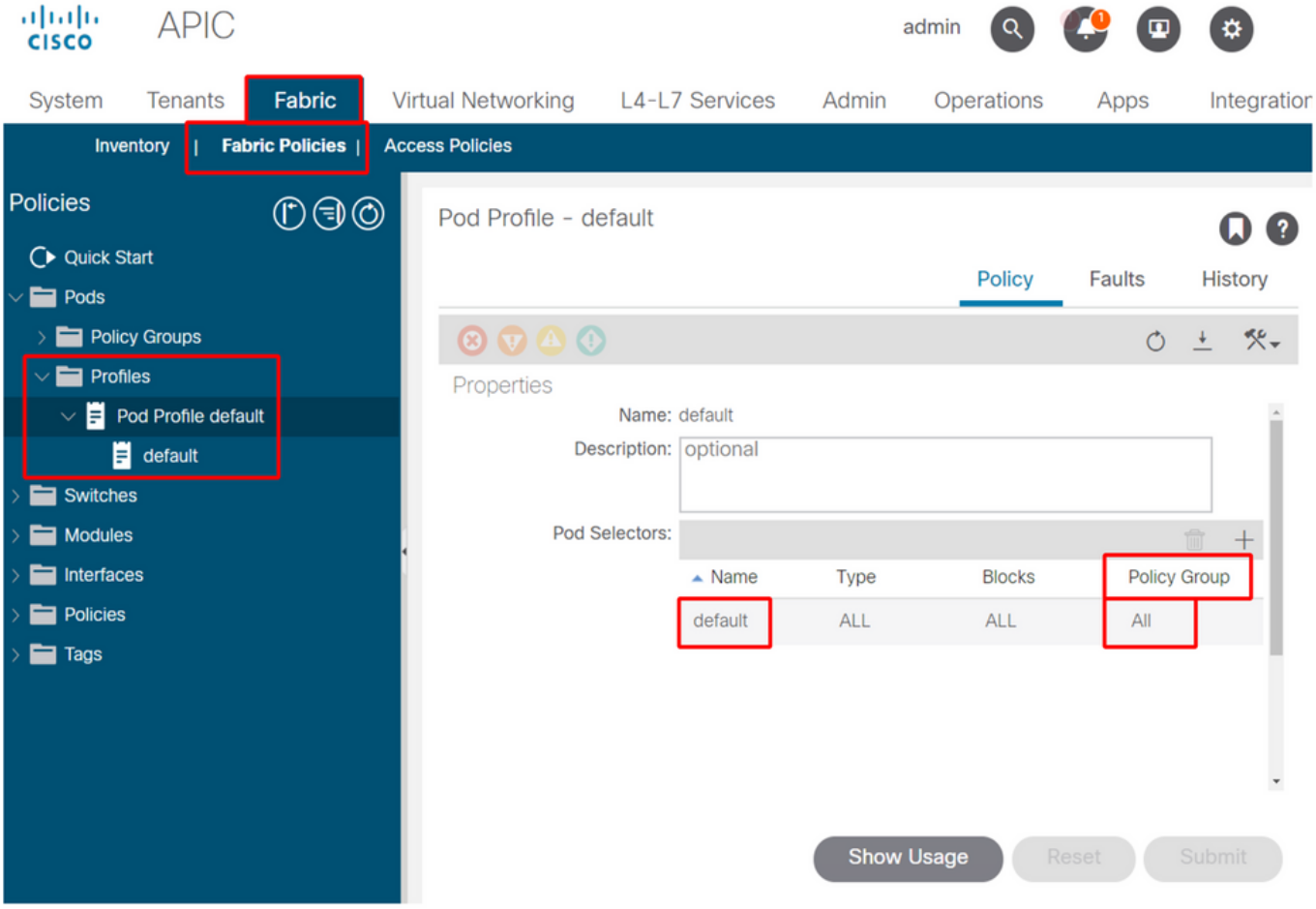
在Pod策略组下应用默认BGP RR策略。即使条目为空，默认BGP RR策略也将作为Pod策略组的一部分应用。

在Pod策略组下应用的BGP路由反射器策略



3.验证Pod策略组是否应用在Pod配置文件下

Pod策略组在Pod配置文件下应用



4.登录到主干，并验证BGP进程是否正在使用已建立的VPN4对等会话运行

```
spine1# show bgp process vrf overlay-1

BGP Process Information
BGP Process ID                 : 26660
BGP Protocol Started, reason:  : configuration
BGP Protocol Tag               : 65001
BGP Protocol State             : Running
BGP Memory State               : OK
BGP asformat                   : asplain
Fabric SOO                     : SOO:65001:33554415
Multisite SOO                  : SOO:65001:16777199
Pod SOO                        : SOO:1:1
...
   Information for address family VPNv4 Unicast in VRF overlay-1
   Table Id                    : 4
   Table state                 : UP
   Table refcount              : 9
   Peers      Active-peers    Routes     Paths      Networks    Aggregates
     7          6               0          0          0           0

   Redistribution
       None

   Wait for IGP convergence is not configured
   Additional Paths Selection route-map interleak_rtmap_golf_rtmap_path_advertise_all
   Is a Route-reflector

   Nexthop trigger-delay
```

```
       critical 500 ms
       non-critical 5000 ms

   Information for address family VPNv6 Unicast in VRF overlay-1
   Table Id                  : 80000004
   Table state               : UP
   Table refcount            : 9
   Peers      Active-peers    Routes     Paths      Networks   Aggregates
   7          6               0          0          0          0

   Redistribution
       None

   Wait for IGP convergence is not configured
   Additional Paths Selection route-map interleak_rtmap_golf_rtmap_path_advertise_all
   Is a Route-reflector

   Nexthop trigger-delay
       critical 500 ms
       non-critical 5000 ms
...
   Wait for IGP convergence is not configured
   Is a Route-reflector

   Nexthop trigger-delay
       critical 500 ms
       non-critical 5000 ms
```

如上所示，枝叶和主干节点之间的MP-BGP仅承载VPNv4和VPNv6地址系列。IPv4地址系列仅用于枝叶节点上的MP-BGP。

也可以使用以下命令轻松观察主干和枝叶节点之间的BGP VPNv4和VPNv6会话。

```
spine1# show bgp vpnv4 unicast summary vrf overlay-1
BGP summary information for VRF overlay-1, address family VPNv4 Unicast
BGP router identifier 10.0.136.65, local AS number 65001
BGP table version is 15, VPNv4 Unicast config peers 7, capable peers 6
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
10.0.136.64     4   65001     162     156       15    0    0 02:26:00 0
10.0.136.67     4   65001     154     154       15    0    0 02:26:01 0
10.0.136.68     4   65001     152     154       15    0    0 02:26:00 0
10.0.136.69     4   65001     154     154       15    0    0 02:26:01 0
10.0.136.70     4   65001     154     154       15    0    0 02:26:00 0
10.0.136.71     4   65001     154     154       15    0    0 02:26:01 0


spine1# show bgp vpnv6 unicast summary vrf overlay-1
BGP summary information for VRF overlay-1, address family VPNv6 Unicast
BGP router identifier 10.0.136.65, local AS number 65001
BGP table version is 15, VPNv6 Unicast config peers 7, capable peers 6
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
10.0.136.64     4   65001     162     156       15    0    0 02:26:11 0
10.0.136.67     4   65001     155     155       15    0    0 02:26:12 0
10.0.136.68     4   65001     153     155       15    0    0 02:26:11 0
```

```
10.0.136.69     4   65001     155     155        15    0    0 02:26:12 0
10.0.136.70     4   65001     155     155        15    0    0 02:26:11 0
10.0.136.71     4   65001     155     155        15    0    0 02:26:12 0
```

注意上述输出中的“Up/Down”列。它应列出一个表示建立BGP会话的持续时间。另请注意，在示例中，“PfxRcd”列显示每个BGP VPNv4/VPNv6对等体为0，因为此ACI交换矩阵尚未配置L3Outs，因此枝叶和主干节点之间没有交换外部路由/前缀。

5.登录枝叶并验证BGP进程是否正在运行已建立的VPN4对等会话

```
leaf1# show bgp process vrf overlay-1

BGP Process Information
BGP Process ID                 : 43242
BGP Protocol Started, reason:  : configuration
BGP Protocol Tag               : 65001
BGP Protocol State             : Running
...

leaf1#  show bgp vpnv4 unicast summary vrf overlay-1
BGP summary information for VRF overlay-1, address family VPNv4 Unicast
BGP router identifier 10.0.136.64, local AS number 65001
BGP table version is 7, VPNv4 Unicast config peers 2, capable peers 2
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
10.0.136.65     4   65001     165     171        7    0    0 02:35:52 0
10.0.136.66     4   65001     167     171        7    0    0 02:35:53 0
```

上述命令输出显示的BGP VPNv4会话数量等于ACI交换矩阵中存在的脊柱节点数量。这与主干节点不同，因为它们建立到每个枝叶和其他路由反射器主干节点的会话。

SNMP

务必从一开始就明确本部分涉及的SNMP功能的特定子集。ACI交换矩阵中的SNMP功能与SNMP Walk功能或SNMP Trap功能相关。这里的重要区别是SNMP Walk控制UDP端口161上的**入口**SNMP流量，而SNMP Trap控制**传出**SNMP流量，SNMP Trap服务器在UDP端口162上侦听。

ACI节点上的入口管理流量需要节点管理EPG（带内或带外）提供必要的合同，以允许流量流动。因此，这也适用于入口SNMP流量。

本节将介绍进入ACI节点（APIC和交换机）的入口SNMP流量（SNMP漫游）。 它将不包括出口SNMP流量（SNMP陷阱），因为这会将本部分的范围扩展到监控策略和监控策略依赖关系（例如，监控策略范围、监控包等）。

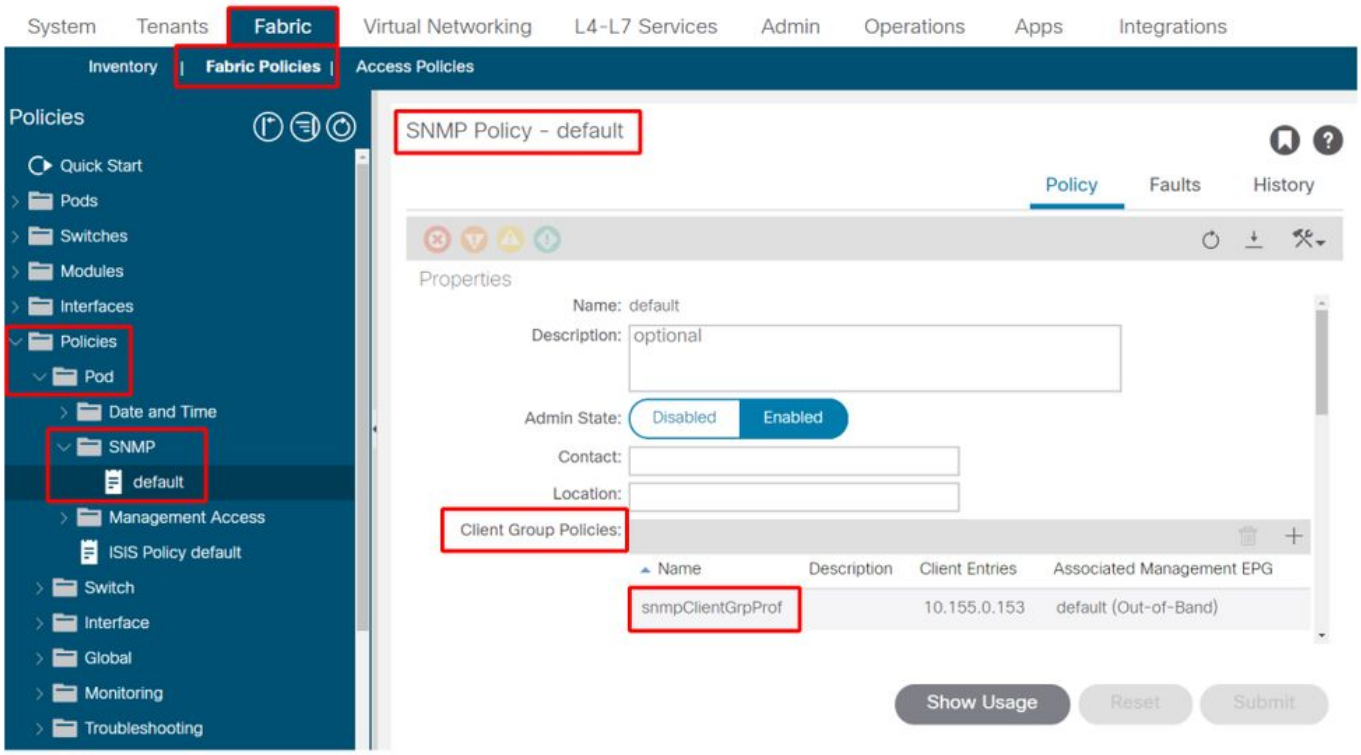
本节还不会介绍ACI支持哪些SNMP MIB。此信息可在思科CCO网站以下链接中找到：https://www.cisco.com/c/dam/en/us/td/docs/switches/datacenter/aci/apic/sw/1-x/mib/list/mib-support.html

故障排除工作流程

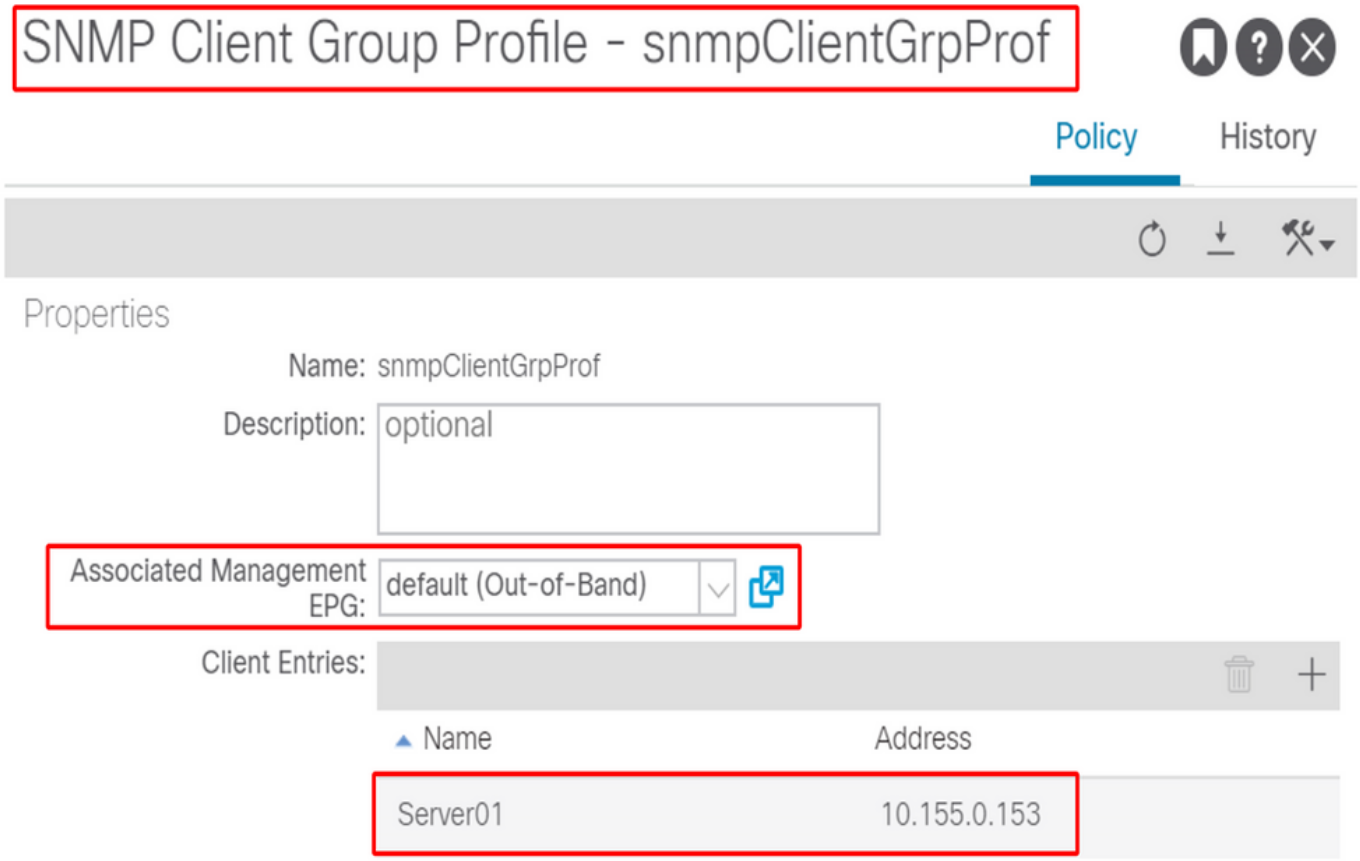
1. SNMP Pod策略 — 验证是否配置了客户端组策略

确保至少有一个SNMP客户端配置为客户端组策略的一部分，如下面的屏幕截图所示。

Pod策略 — SNMP策略 — 客户端组策略

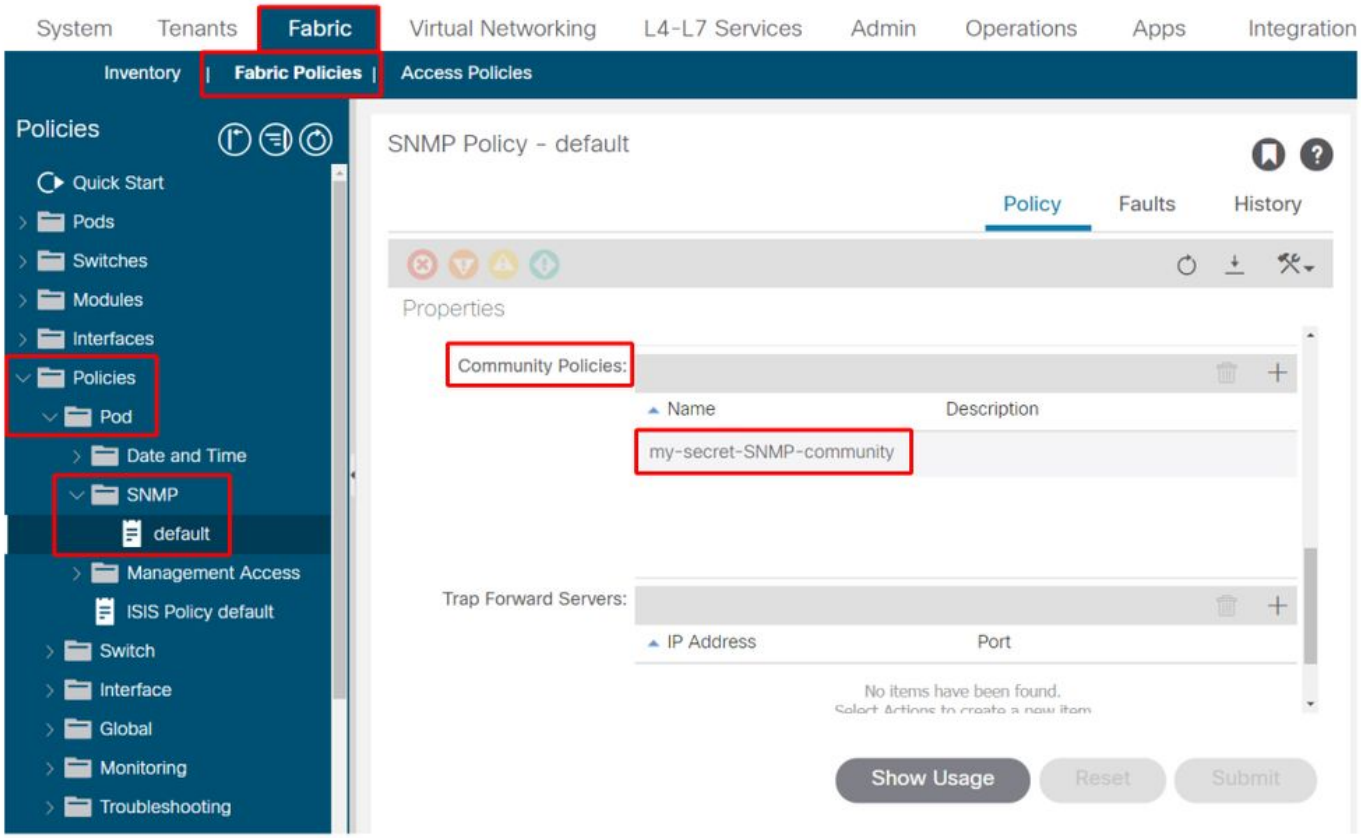


Pod策略 — SNMP策略 — 客户端组策略

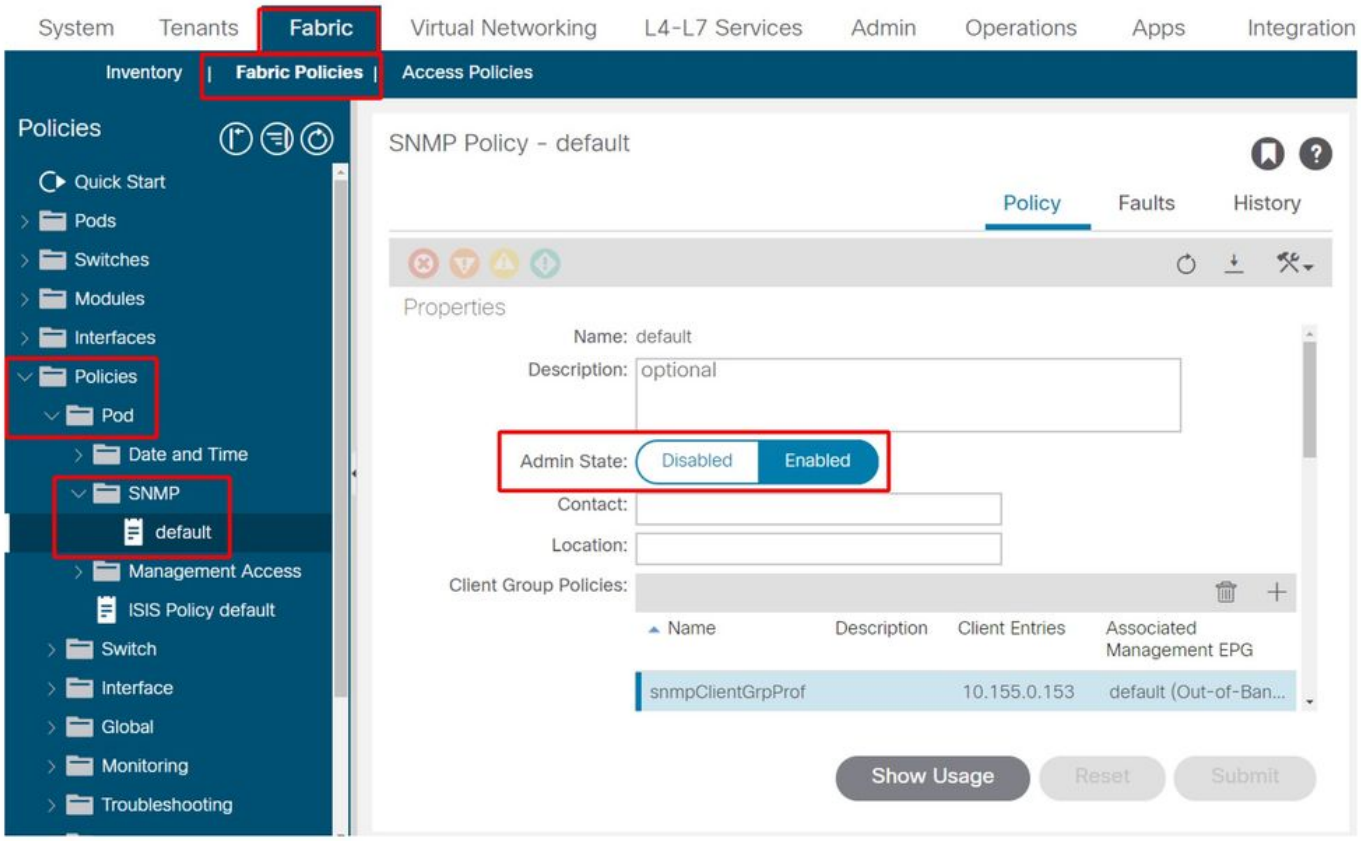


2. SNMP Pod策略 — 验证是否至少配置了一个社区策略

Pod策略 — SNMP策略 — 社区策略



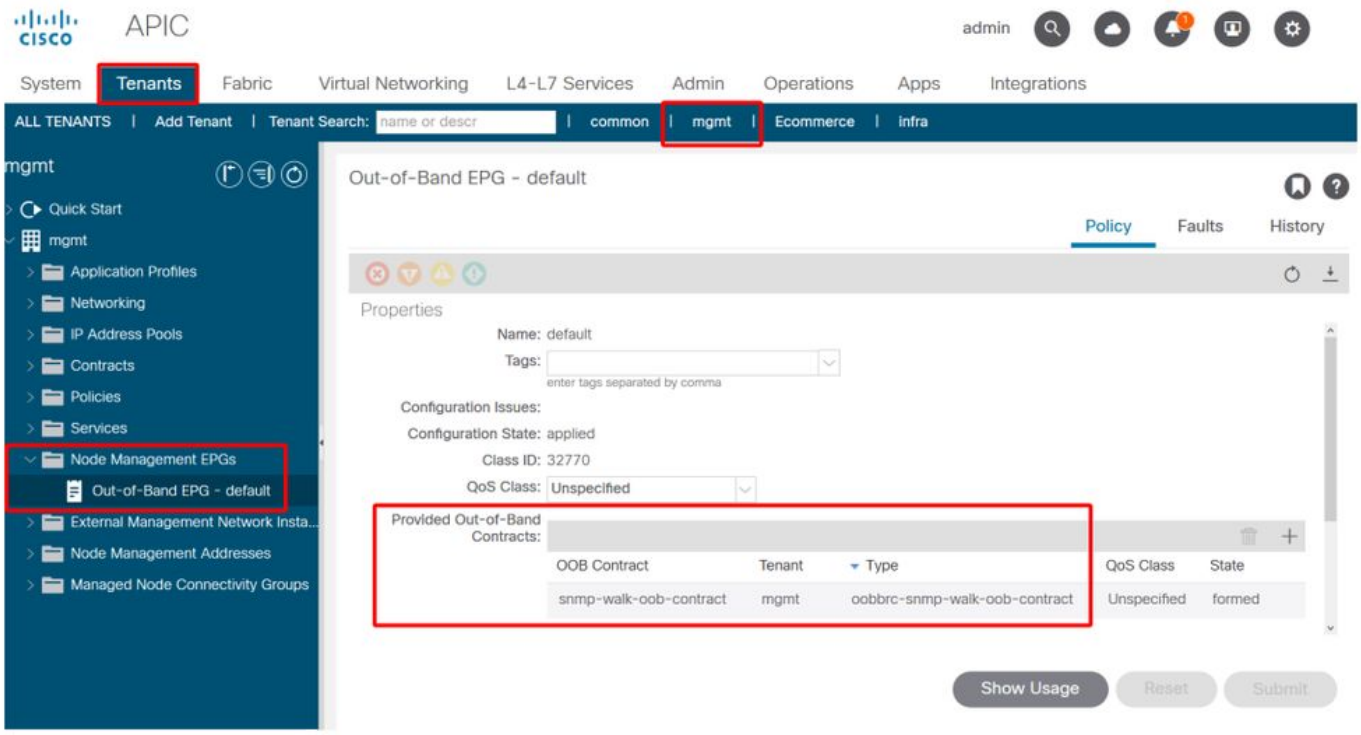
3. SNMP Pod策略 — 验证管理状态是否设置为“启用”



4.管理租户 — 验证OOB EPG是否提供允许UDP端口161的OOB合同

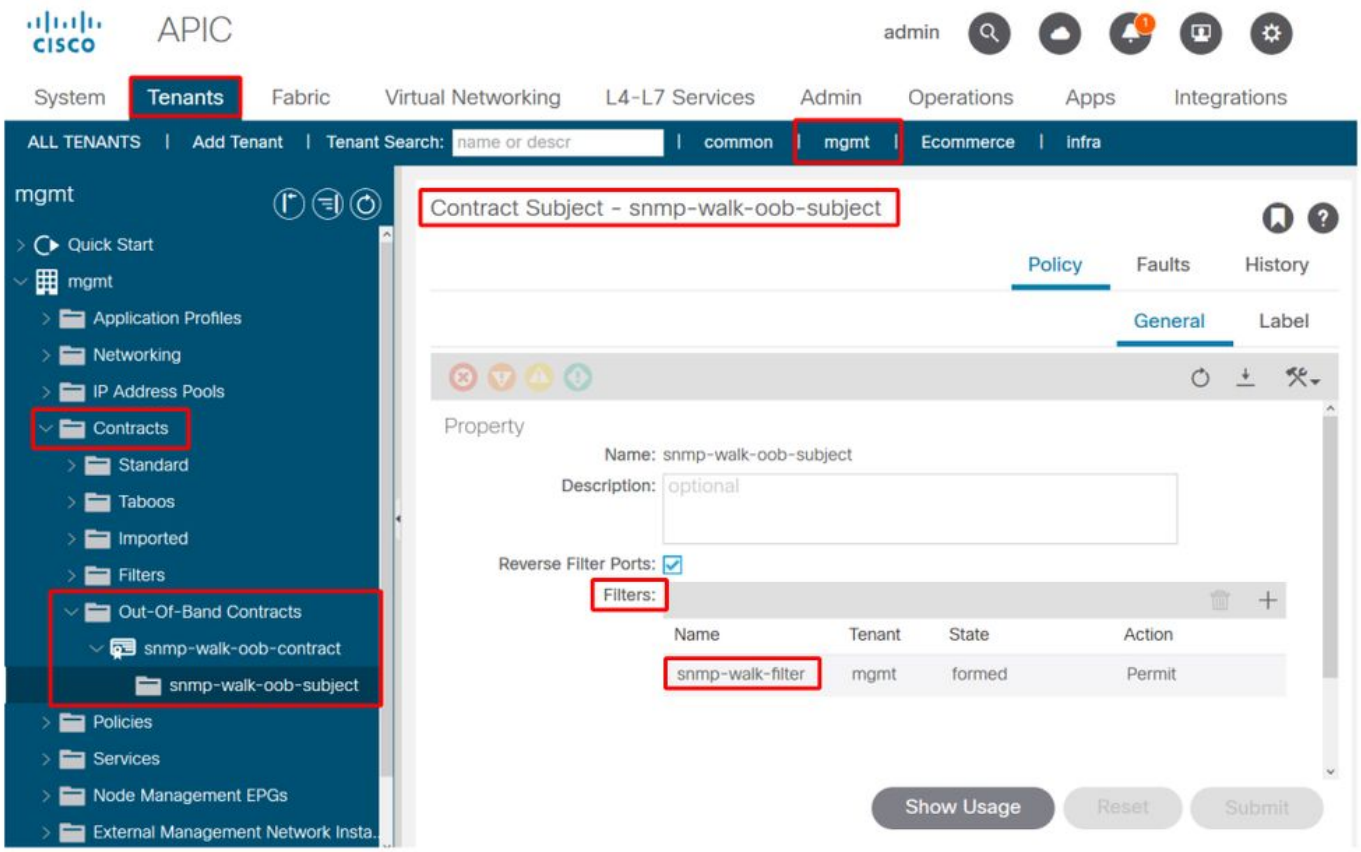
OOB EPG管理到APIC和交换机OOB管理端口的连接。因此，它会影响所有进入OOB端口的流量。

确保此处提供的合同包含所有必要的管理服务，而不仅仅是SNMP。例如：它还需要至少包含SSH（TCP端口22）。 否则，将无法使用SSH登录交换机。请注意，这不适用于APIC，因为它们具有允许SSH、HTTP和HTTPS的机制，可防止用户完全锁定。



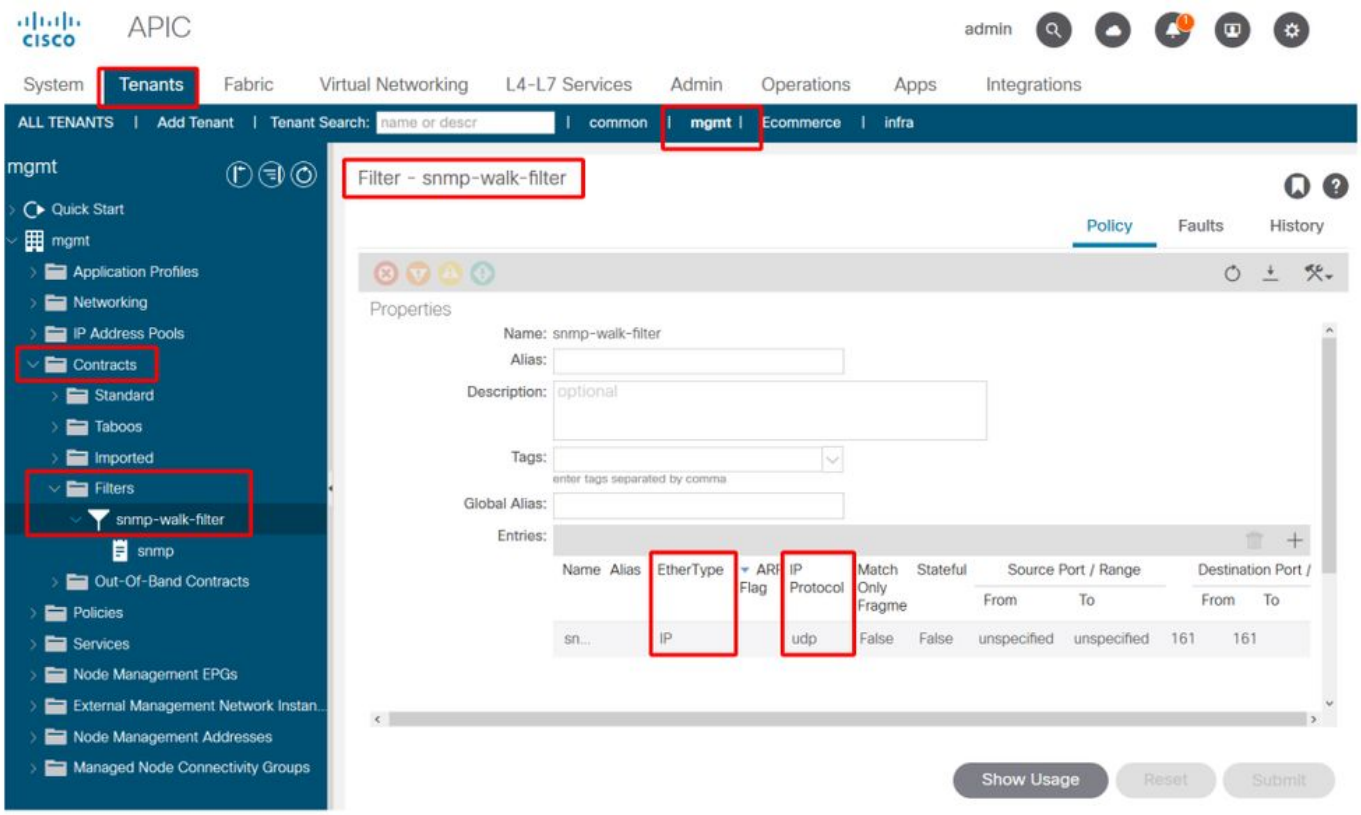
5.管理租户 — 验证OOB合同是否存在，并且是否具有允许UDP端口161的过滤器

管理租户 — OOB EPG — 提供的OOB合同



在下图中，并非必须仅允许UDP端口161。具有允许以任何方式使用UDP端口161的过滤器的合同是

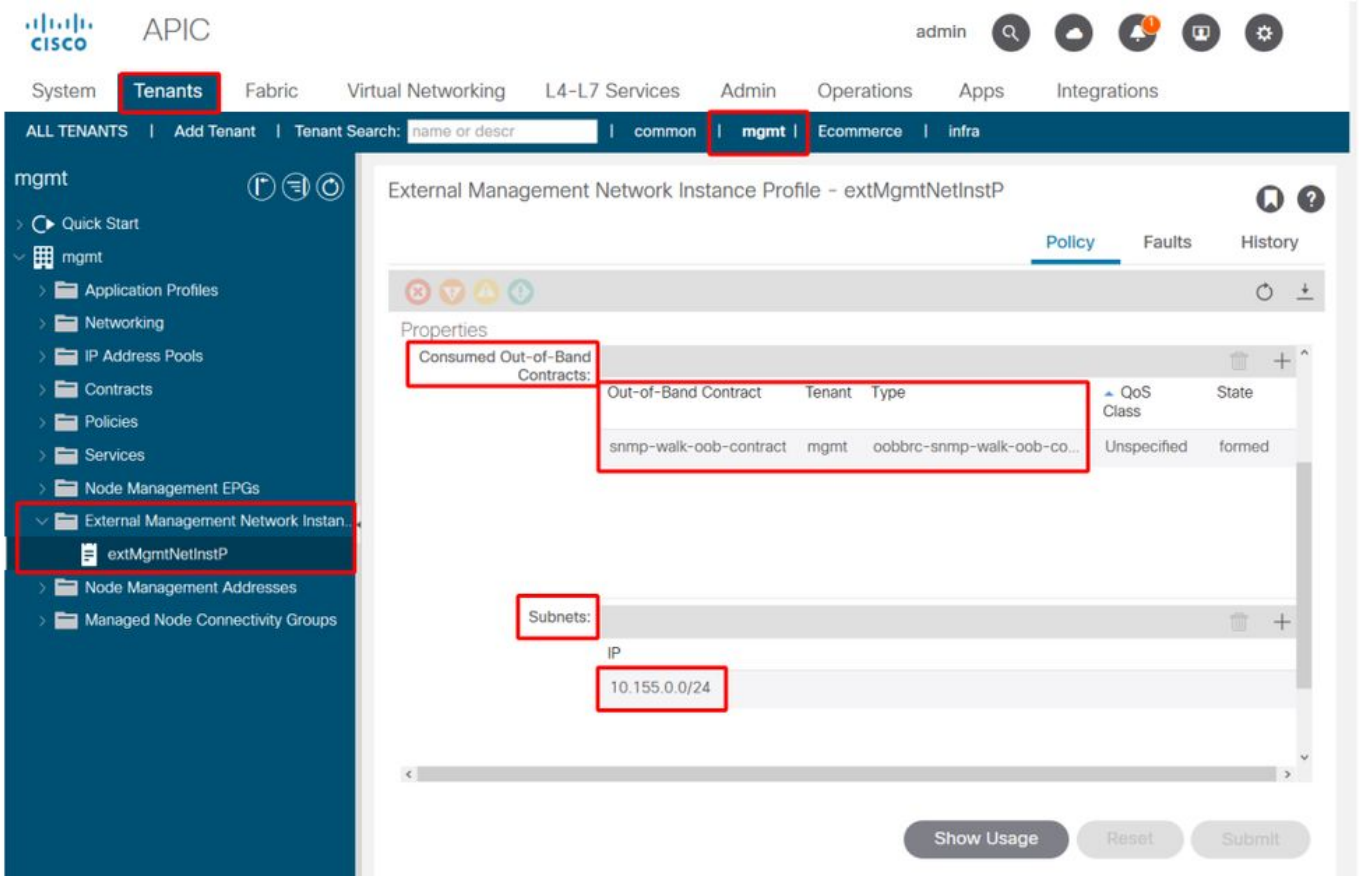
正确的。这甚至可以是具有来自公共租户的默认过滤器的合同主题。在我们的示例中，为清楚起见，仅针对UDP端口161配置了特定过滤器。



6.管理租户 — 验证外部管理网络实例配置文件是否具有使用OOB合同的有效子网

外部管理网络实例配置文件(ExtMgmtNetInstP)表示由其中的“子网”定义的外部源，需要使用通过OOB EPG可到达的服务。因此，ExtMgmtNetInstP使用由OOB EPG提供的同一OOB合同。这是允许UDP端口161的合同。此外，ExtMgmtNetInstP还指定可能使用OOB EPG提供的服务的允许子网范围。

管理租户 — ExtMgmtNetInstP，带消耗的OOB合同和子网



如上图所示，需要基于CIDR的子网表示法。图中显示了特定的/24子网。要求子网条目包括SNMP Pod策略中配置的SNMP客户端条目（请参阅图Pod策略 — SNMP策略 — 客户端组策略）。

如前所述，请注意包括所有所需的外部子网，以防止其他必要的管理服务被锁定。

7.登录交换机并执行tcpdump以观察是否观察到SNMP Walk数据包 — UDP端口161

如果SNMP Walk数据包通过OOB端口进入交换机，这意味着所有必要的SNMP和基于OOB的策略/参数都已正确配置。因此，这是一种适当的验证方法。

枝叶节点上的Tcpdump利用其Linux shell和Linux网络设备。因此，有必要按照以下示例捕获接口“eth0”上的数据包。在本示例中，SNMP客户端正在对OID .1.0.802.1.1.2.1.1.1.0执行SNMP Get请求。

```
leaf1# ip addr show eth0
2: eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc mq state UP qlen 1000
   link/ether f4:cf:e2:28:fc:ac brd ff:ff:ff:ff:ff:ff
   inet 10.48.22.77/24 brd 10.48.22.255 scope global eth0
       valid_lft forever preferred_lft forever
   inet6 fe80::f6cf:e2ff:fe28:fcac/64 scope link
       valid_lft forever preferred_lft forever


leaf1# tcpdump -i eth0 udp port 161
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on eth0, link-type EN10MB (Ethernet), capture size 65535 bytes
22:18:10.204011 IP 10.155.0.153.63392 > 10.48.22.77.snmp:  C=my-snmp-community
GetNextRequest(28)  .iso.0.8802.1.1.2.1.1.1.0
22:18:10.204558 IP 10.48.22.77.snmp > 10.155.0.153.63392:  C=my-snmp-community GetResponse(29)
.iso.0.8802.1.1.2.1.1.2.0=4
```

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）。