如果使用NAT，为什么vEdge无法建立IPSec隧道？

目录

简介

本文档介绍当vEdge路由器对数据平面隧道使用IPSec封装，并且一台设备在网络地址转换(NAT)设备后执行对称NAT(RFC3489)或地址相关映射(RFC4787)，而另一台设备具有直接互联网接入(DIA)或某些设备时可能出现的问题在传输端接口上配置的其他类型的NAT。

背景信息

> **注意：**本文仅适用于vEdge路由器，并基于vEdge软件18.4.1和19.1.0中的行为编写。在较新版本中，行为可能不同。如有疑问，请查阅文档或联系思科技术支持中心(TAC)。

为了进行演示，问题在SD-WAN TAC实验中重现。设备设置在下表中汇总：

主机名	站点ID	system-ip	专用IP	公共IP
vedge1	232	10.10.10.232	192.168.10.232	198.51.100.232
vedge2	233	10.10.10.233	192.168.9.233	192.168.9.233
vsmart	1	10.10.10.228	192.168.0.228	192.168.0.228
vbond	1	10.10.10.231	192.168.0.231	192.168.0.231

两台设备上的传输端配置相当通用。以下是vEdge1的配置：

```
vpn 0
 interface ge0/0
  ip address 192.168.10.232/24
```

```
  !
  tunnel-interface
   encapsulation ipsec
   color biz-internet
   no allow-service bgp
   no allow-service dhcp
   allow-service dns
   allow-service icmp
   no allow-service sshd
   no allow-service netconf
   no allow-service ntp
   no allow-service ospf
   no allow-service stun
   allow-service https
  !
  no shutdown
 !
 ip route 0.0.0.0/0 192.168.10.11
!
```

vEdge2:

```
interface ge0/1
  ip address 192.168.9.233/24
  !
  tunnel-interface
   encapsulation ipsec
   color biz-internet
   no allow-service bgp
   no allow-service dhcp
   allow-service dns
   allow-service icmp
   no allow-service sshd
   no allow-service netconf
   no allow-service ntp
   no allow-service ospf
   no allow-service stun
   allow-service https
  !
  no shutdown
 !
 ip route 0.0.0.0/0 192.168.9.1
```

为了演示本文档中的问题，虚拟自适应安全设备(ASAv)防火墙驻留在两个vEdge路由器之间。ASAv正在根据以下规则进行地址转换：

- 如果来自vEdge1的流量用于控制器，则源端口12346-12426将转换为52346-52426
- 如果来自vEdge1的流量用于到其他站点的数据平面连接，则源端口12346-12426将转换为42346-42426
- 来自vEdge1的所有其他流量也映射到同一公有地址(198.51.100.232)

以下是供参考的ASAv NAT配置：

```
object network VE1
 host 192.168.10.232
object network CONTROLLERS
 subnet 192.168.0.0 255.255.255.0
object network VE1_NAT
 host 198.51.100.232
object service CONTROL
 service udp source range 12346 12445 destination range 12346 12445
```

```
object service CC_NAT_CONTROLLERS
 service udp source range 52346 52445 destination range 12346 12445
object service CC_NAT_OTHER
 service udp source range 42346 42445 destination range 12346 12445
object network ALL
 subnet 0.0.0.0 0.0.0.0
nat (ve1-iface,ve2-iface) source static VE1 VE1_NAT destination static CONTROLLERS CONTROLLERS
service CONTROL CC_NAT_CONTROLLERS
nat (ve1-iface,ve2-iface) source static VE1 VE1_NAT destination static ALL ALL service CONTROL
CC_NAT_OTHER
nat (ve1-iface,ve2-iface) source dynamic VE1 VE1_NAT
```

问题

工作场景

在正常状态下，我们可以观察到数据平面隧道已建立，双向转发检测(BFD)处于up状态。

请注意vEdge1设备(52366)上用于与控制器建立控制连接的公共端口：

```
vEdge1# show control local-properties wan-interface-list

 NAT TYPE: E -- indicates End-point independent mapping
           A -- indicates Address-port dependent mapping
           N -- indicates Not learned
           Note: Requires minimum two vbonds to learn the NAT type

                         PUBLIC          PUBLIC PRIVATE         PRIVATE
PRIVATE                                  MAX   RESTRICT/           LAST         SPI TIME    NAT  VM
INTERFACE                IPv4            PORT   IPv4             IPv6
PORT    VS/VM COLOR            STATE CNTRL CONTROL/     LR/LB  CONNECTION   REMAINING   TYPE CON

STUN                                            PRF
-------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------
-----------
ge0/0                    198.51.100.232  52366  192.168.10.232  ::
12366    2/1  biz-internet     up     2      no/yes/no   No/No  0:00:00:28   0:11:59:17  N    5
```

在vEdge2上，未使用NAT，因此私有地址和端口相同：

```
vEdge2# show control local-properties wan-interface-list

 NAT TYPE: E -- indicates End-point independent mapping
           A -- indicates Address-port dependent mapping
           N -- indicates Not learned
           Note: Requires minimum two vbonds to learn the NAT type

                         PUBLIC          PUBLIC PRIVATE         PRIVATE
PRIVATE                                  MAX   RESTRICT/           LAST         SPI TIME    NAT  VM
INTERFACE                IPv4            PORT   IPv4             IPv6
PORT    VS/VM COLOR            STATE CNTRL CONTROL/     LR/LB  CONNECTION   REMAINING   TYPE CON

STUN                                            PRF
-------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------
-----------
ge0/1                    192.168.9.233   12366  192.168.9.233   ::
```

```
12366    2/1  biz-internet     up     2      no/yes/no   No/No  0:00:00:48   0:11:58:53  N    5
```

在show tunnel statistics**sfrom vEdge1中**，我们可以看到tx/rx计数器正在递增：

```
vEdge1# show tunnel statistics dest-ip 192.168.9.233

TCP
TUNNEL                                    SOURCE  DEST
TUNNEL                                            MSS
PROTOCOL  SOURCE IP       DEST IP         PORT    PORT   SYSTEM IP     LOCAL COLOR   REMOTE COLOR
MTU     tx-pkts  tx-octets  rx-pkts  rx-octets  ADJUST
-------------------------------------------------------------------------------------------------
-----------------------------------------------------------
ipsec     192.168.10.232  192.168.9.233  12366   12366  10.10.10.233  biz-internet  biz-internet
1441    223      81163      179      40201      1202
```

从vEdge2的相同输出中，您可以看到rx/rx数据包计数器正在递增。请注意，目的端口(42366)与用于建立控制连接的端口(52366)不同：

```
vEdge2# show tunnel statistics dest-ip 198.51.100.232

TCP
TUNNEL                                    SOURCE  DEST
TUNNEL                                            MSS
PROTOCOL  SOURCE IP      DEST IP         PORT    PORT   SYSTEM IP     LOCAL COLOR   REMOTE COLOR
MTU     tx-pkts  tx-octets  rx-pkts  rx-octets  ADJUST
-------------------------------------------------------------------------------------------------
-----------------------------------------------------------
ipsec     192.168.9.233  198.51.100.232  12366   42366  10.10.10.232  biz-internet  biz-internet
1441    296      88669      261      44638      1201
```

但是，两台设备上的BFD会话仍处于工作状态：

```
vEdge1# show bfd sessions site-id 233 | tab

                                     SRC    DST                  SITE
DETECT      TX
SRC IP          DST IP         PROTO  PORT   PORT   SYSTEM IP     ID    LOCAL COLOR   COLOR
STATE  MULTIPLIER  INTERVAL  UPTIME      TRANSITIONS
-------------------------------------------------------------------------------------------------
-----------------------------------------------------------
192.168.10.232  192.168.9.233  ipsec  12366  12366  10.10.10.233  233   biz-internet  biz-
internet  up     7           1000      0:00:02:42  0



vEdge2# show bfd sessions site-id 232 | tab

                                     SRC    DST                  SITE
DETECT      TX
SRC IP          DST IP          PROTO  PORT   PORT   SYSTEM IP     ID    LOCAL COLOR   COLOR
STATE  MULTIPLIER  INTERVAL  UPTIME      TRANSITIONS
```

```
------------------------------------------------------------------------------------------------
-------------------------------------------------------
192.168.9.233  198.51.100.232  ipsec  12366  52366  10.10.10.232  232   biz-internet  biz-
internet  up     7           1000      0:00:03:00  0
```

用于控制和数据平面连接的不同端口不会导致任何问题，因此连接就位。

故障场景

用户希望在vEdge2路由器上启用直接互联网接入(DIA)。为此，此配置已应用于vEdge2:

```
vpn 0
 interface ge0/1
  nat
   respond-to-ping
  !
 !
!
vpn 1
 ip route 0.0.0.0/0 vpn 0
!
```

BFD会话意外关闭，而且仍处于关闭状态。清除隧道统计信息后，您可以看到RX计数器在show tunnel statistics输**出中不**增加：

```
vEdge2# show tunnel statistics dest-ip 198.51.100.232


TCP
TUNNEL                                    SOURCE  DEST
TUNNEL                                            MSS
PROTOCOL  SOURCE IP      DEST IP         PORT    PORT   SYSTEM IP     LOCAL COLOR   REMOTE COLOR
MTU     tx-pkts  tx-octets  rx-pkts  rx-octets  ADJUST
------------------------------------------------------------------------------------------------
-------------------------------------------------------
ipsec     192.168.9.233  198.51.100.232  12346   52366  10.10.10.232  biz-internet  biz-internet
1442    282      48222      0        0          1368

vEdge2# show bfd sessions site-id 232
                                      SOURCE TLOC      REMOTE TLOC
DST PUBLIC                        DST PUBLIC         DETECT      TX
SYSTEM IP        SITE ID  STATE       COLOR            COLOR           SOURCE IP
IP                                PORT        ENCAP  MULTIPLIER  INTERVAL(msec) UPTIME
TRANSITIONS
------------------------------------------------------------------------------------------------
------------------------------------------------------------------------------------------------
-------------
10.10.10.232     232      down        biz-internet     biz-internet     192.168.9.233
198.51.100.232                   52366       ipsec  7           1000           NA              0

vEdge2# show tunnel statistics dest-ip 198.51.100.232


TCP
TUNNEL                                    SOURCE  DEST
TUNNEL                                            MSS
PROTOCOL  SOURCE IP      DEST IP         PORT    PORT   SYSTEM IP     LOCAL COLOR   REMOTE COLOR
MTU     tx-pkts  tx-octets  rx-pkts  rx-octets  ADJUST
```

```
-------------------------------------------------------------------------------------------------
-------------------------------------------------------
ipsec     192.168.9.233  198.51.100.232  12346   52366  10.10.10.232  biz-internet  biz-internet
1442    285      48735      0        0          1368
```

最初，客户怀疑该问题与隧道MTU有关。如果将上述输出与“工作场景”部分的输出进行比较，您会注意到在工作场景中，隧道MTU为1441，而失败场景中为1442。根据文档，隧道MTU应为1442（隧道开销的默认接口MTU为1500 - 58字节），但BFD为向上，隧道MTU降低1字节。对于您的参考，在BFD处于**down状态时，**将显示隧**道统计信**息以及下面提供的**show tunnel statistics bfd的输**出：

```
vEdge1# show tunnel statistics dest-ip 192.168.9.233 ; show tunnel statistics bfd dest-ip
192.168.9.233


TCP
TUNNEL                                    SOURCE  DEST
TUNNEL                                            MSS
PROTOCOL  SOURCE IP       DEST IP         PORT    PORT   SYSTEM IP     LOCAL COLOR   REMOTE COLOR
MTU     tx-pkts  tx-octets  rx-pkts  rx-octets  ADJUST
-------------------------------------------------------------------------------------------------
-------------------------------------------------------
ipsec     192.168.10.232  192.168.9.233  12346   12346  10.10.10.233  biz-internet  biz-internet
1442    133      22743      0        0          1362


                                                          BFD   BFD   BFD     BFD     BFD   BFD
BFD     BFD
                                                          ECHO  ECHO  ECHO    ECHO    PMTU  PMTU
PMTU    PMTU
TUNNEL                                    SOURCE  DEST    TX    RX    TX      RX      TX    RX
TX      RX
PROTOCOL  SOURCE IP       DEST IP         PORT    PORT    PKTS  PKTS  OCTETS  OCTETS  PKTS  PKTS
OCTETS  OCTETS
-------------------------------------------------------------------------------------------------
----------------
ipsec     192.168.10.232  192.168.9.233  12346   12346  133   0     22743   0       0     0
0       0

vEdge1# show tunnel statistics dest-ip 192.168.9.233 ; show tunnel statistics bfd dest-ip
192.168.9.233


TCP
TUNNEL                                    SOURCE  DEST
TUNNEL                                            MSS
PROTOCOL  SOURCE IP       DEST IP         PORT    PORT   SYSTEM IP     LOCAL COLOR   REMOTE COLOR
MTU     tx-pkts  tx-octets  rx-pkts  rx-octets  ADJUST
-------------------------------------------------------------------------------------------------
-------------------------------------------------------
ipsec     192.168.10.232  192.168.9.233  12346   12346  10.10.10.233  biz-internet  biz-internet
1442    134      22914      0        0          1362


                                                          BFD   BFD   BFD     BFD     BFD   BFD
BFD     BFD
                                                          ECHO  ECHO  ECHO    ECHO    PMTU  PMTU
PMTU    PMTU
TUNNEL                                    SOURCE  DEST    TX    RX    TX      RX      TX    RX
TX      RX
```

```
PROTOCOL  SOURCE IP       DEST IP        PORT    PORT   PKTS  PKTS  OCTETS  OCTETS  PKTS  PKTS
OCTETS  OCTETS
-------------------------------------------------------------------------------------------
----------------
ipsec     192.168.10.232  192.168.9.233  12346   12346  134   0     22914   0       0     0
0       0
```

如果BFD处于up状态：

```
vEdge1# show tunnel statistics dest-ip 192.168.9.233 ; show tunnel statistics bfd dest-ip
192.168.9.233 ;


TCP
TUNNEL                                  SOURCE  DEST
TUNNEL                                          MSS
PROTOCOL  SOURCE IP       DEST IP        PORT    PORT   SYSTEM IP     LOCAL COLOR   REMOTE COLOR
MTU     tx-pkts  tx-octets  rx-pkts  rx-octets  ADJUST
-------------------------------------------------------------------------------------------
-----------------------------------------------------------
ipsec     192.168.10.232  192.168.9.233  12346   12346  10.10.10.233  biz-internet  biz-internet
1441    3541     610133     3504     592907     1361


                                                        BFD   BFD   BFD     BFD     BFD   BFD
BFD     BFD
                                                        ECHO  ECHO  ECHO    ECHO    PMTU  PMTU
PMTU    PMTU
TUNNEL                                  SOURCE  DEST    TX    RX    TX      RX      TX    RX
TX      RX
PROTOCOL  SOURCE IP       DEST IP        PORT    PORT   PKTS  PKTS  OCTETS  OCTETS  PKTS  PKTS
OCTETS  OCTETS
-------------------------------------------------------------------------------------------
----------------
ipsec     192.168.10.232  192.168.9.233  12346   12346  3522  3491  589970  584816  19    13
20163   8091

vEdge1# show tunnel statistics dest-ip 192.168.9.233 ; show tunnel statistics bfd dest-ip
192.168.9.233 ;


TCP
TUNNEL                                  SOURCE  DEST
TUNNEL                                          MSS
PROTOCOL  SOURCE IP       DEST IP        PORT    PORT   SYSTEM IP     LOCAL COLOR   REMOTE COLOR
MTU     tx-pkts  tx-octets  rx-pkts  rx-octets  ADJUST
-------------------------------------------------------------------------------------------
-----------------------------------------------------------
ipsec     192.168.10.232  192.168.9.233  12346   12346  10.10.10.233  biz-internet  biz-internet
1441    3542     610297     3505     593078     1361


                                                        BFD   BFD   BFD     BFD     BFD   BFD
BFD     BFD
                                                        ECHO  ECHO  ECHO    ECHO    PMTU  PMTU
PMTU    PMTU
TUNNEL                                  SOURCE  DEST    TX    RX    TX      RX      TX    RX
TX      RX
PROTOCOL  SOURCE IP       DEST IP        PORT    PORT   PKTS  PKTS  OCTETS  OCTETS  PKTS  PKTS
OCTETS  OCTETS
-------------------------------------------------------------------------------------------
```

```
---------------
ipsec     192.168.10.232  192.168.9.233  12346   12346  3523  3492  590134  584987  19    13
20163   8091
```

注意：顺便说一下，我们可以通过查看上述输出来确定BFD数据包大小和封装。请注意，在两个输出之间只收到一个BFD数据包，因此提交BFD Echo RX八位字节值584987 - 584816将给我们171字节的结果。它可用于精确计算BFD自身使用的带宽。

BFD陷入关闭状**态**的原因不是MTU，而是NAT配置。这是工作方案和失败方案**之间唯**一的**更改**。您可以在此看到，由于DIA配置，NAT静态映射由vEdge2在转换表中自动创建，以允许数据平面IPSec流量绕行：

```
vEdge2# show ip nat filter nat-vpn 0 nat-ifname ge0/1 vpn 0 protocol udp 192.168.9.233
198.51.100.232

                            PRIVATE                          PRIVATE  PRIVATE
PUBLIC  PUBLIC
NAT  NAT                    SOURCE         PRIVATE DEST    SOURCE   DEST     PUBLIC SOURCE
PUBLIC DEST     SOURCE  DEST    FILTER       IDLE         OUTBOUND  OUTBOUND  INBOUND  INBOUND
VPN  IFNAME  VPN  PROTOCOL  ADDRESS        ADDRESS         PORT     PORT     ADDRESS
ADDRESS         PORT    PORT    STATE        TIMEOUT     PACKETS   OCTETS    PACKETS  OCTETS
DIRECTION
-------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------
------
0    ge0/1   0    udp       192.168.9.233  198.51.100.232  12346    52366    192.168.9.233
198.51.100.232  12346   52366   established  0:00:00:59  53        8321      0        0        -
```

如您所见，使用的是端口52366而不是42366。这是因为vEdge2需要52366个端口，并从vSmart通告的OMP TLOC中获知了该端口：

```
vEdge2# show omp tlocs ip 10.10.10.232 | b PUBLIC

PUBLIC           PRIVATE
ADDRESS                                                                    PSEUDO
PUBLIC                   PRIVATE  PUBLIC  IPV6    PRIVATE  IPV6     BFD
FAMILY   TLOC IP          COLOR            ENCAP  FROM PEER        STATUS    KEY     PUBLIC IP
PORT    PRIVATE IP       PORT     IPV6    PORT    IPV6     PORT     STATUS
-------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------
ipv4     10.10.10.232     biz-internet     ipsec  10.10.10.228     C,I,R     1
198.51.100.232   52366   192.168.10.232   12346    ::      0       ::       0        down
```

解决方案

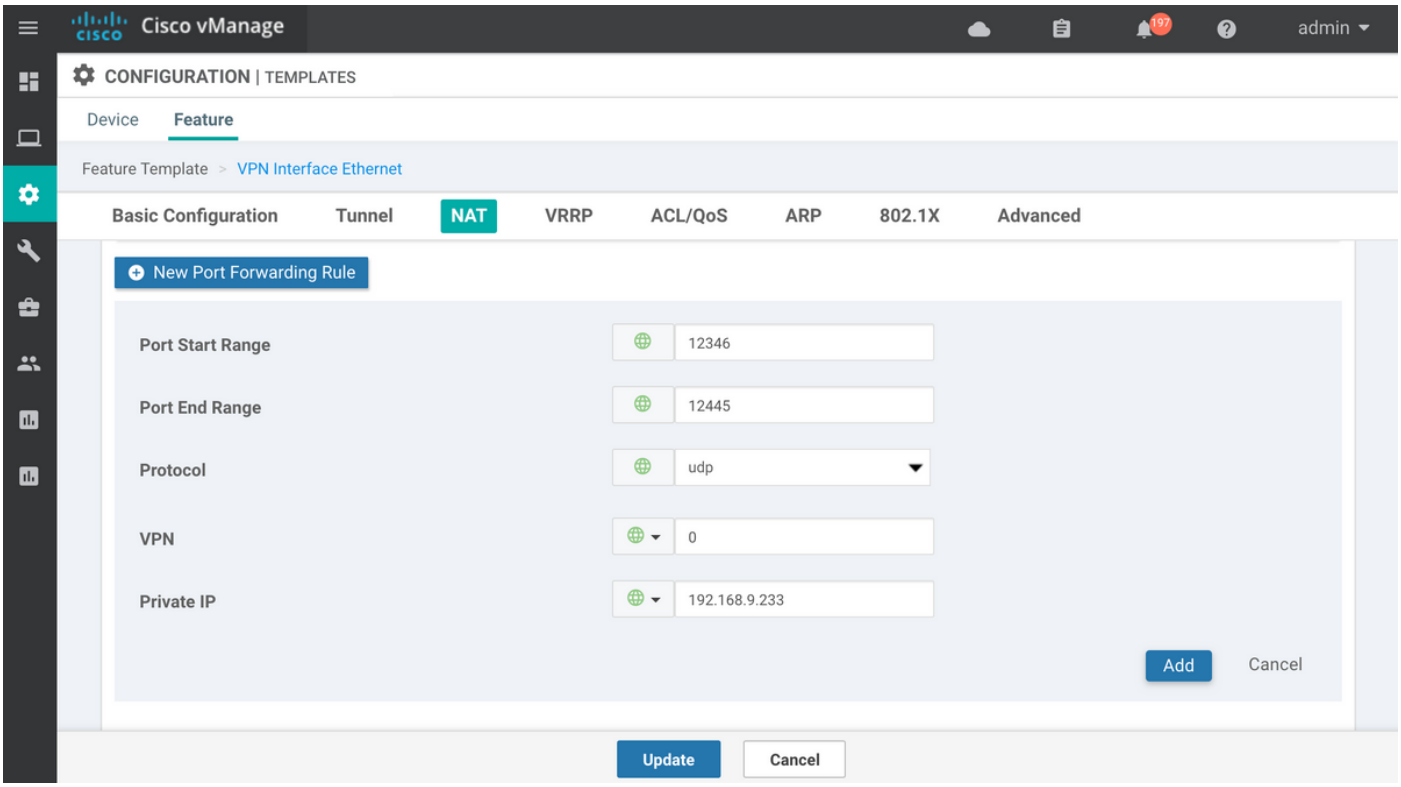
NAT端口转发

乍一看，解决此类问题的方法很简单。您可以在vEdge2传输接口上配置静态NAT免除端口转发，以强制绕过来自任何来源的数据平面连接过滤：

```
vpn 0
 interface ge0/1
  nat
   respond-to-ping
   port-forward port-start 12346 port-end 12445 proto udp
    private-vpn        0
    private-ip-address 192.168.9.233
   !
  !
 !
!
```

此范围12346至12446可支持所有可能的初始端口(12346、12366、12386、12406和12426加端口偏移)。 有关详细信息，请参阅“Viptela部署的防火墙端口”。

如果使用的是设备功能模板而不是CLI模板，则要实现此目的，我们需要更新或添加新的VPN以太网功能模板，以使用新端口转发规则(VPN 0)**的相应传输(VPN 0)接口**，如图所示：



显式ACL

此外，还可以使用另一个显式ACL的解决方案。如果**在策略部分**下配置了**implicit-acl-logging**，则您可能会在**/var/log/tmplog/vdebug文件中注意到以下消**息：

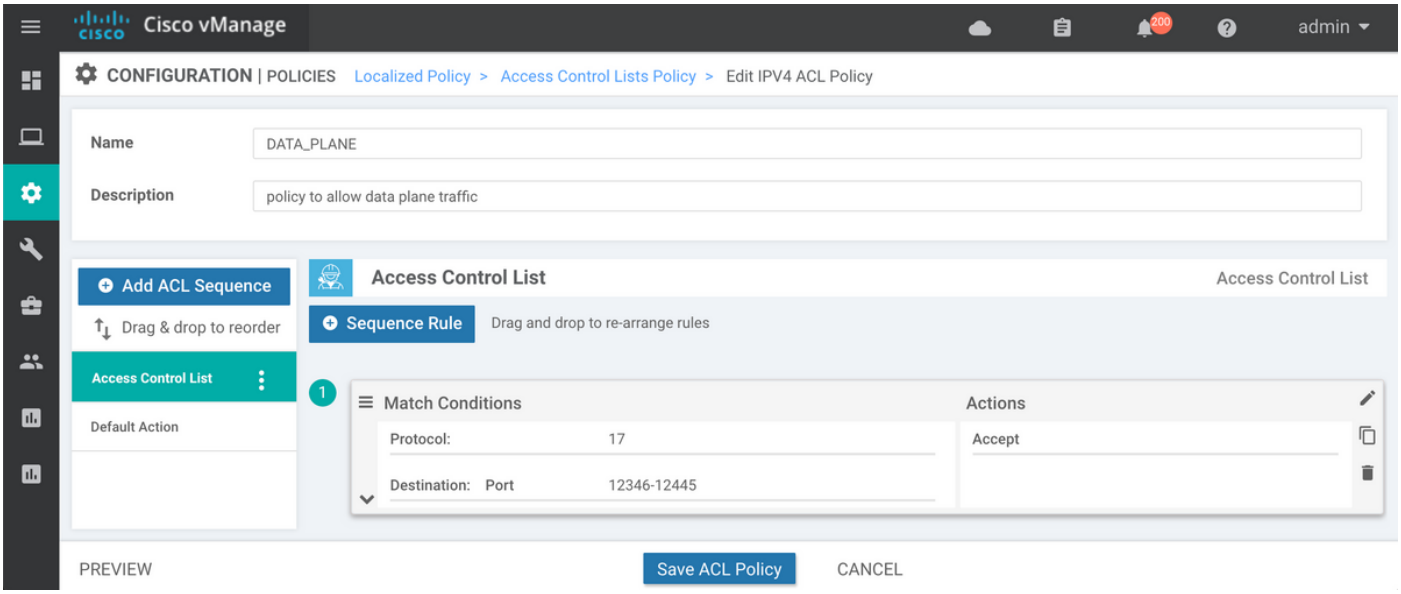
```
local7.notice: Jun  8 17:53:29 vEdge2 FTMD[980]: %Viptela-vEdge2-FTMD-5-NTCE-1000026: FLOW LOG
vpn-0 198.51.100.232/42346 192.168.9.233/12346 udp: tos: 192  inbound-acl, Implicit-ACL, Result:
denyPkt count 2: Byte count 342 Ingress-Intf ge0/1 Egress-intf cpu
```

它解释了根本原因，因此，您需要明确允许vEdge2上访问控制列表(ACL)中的传入数据平面数据包，如下所示：

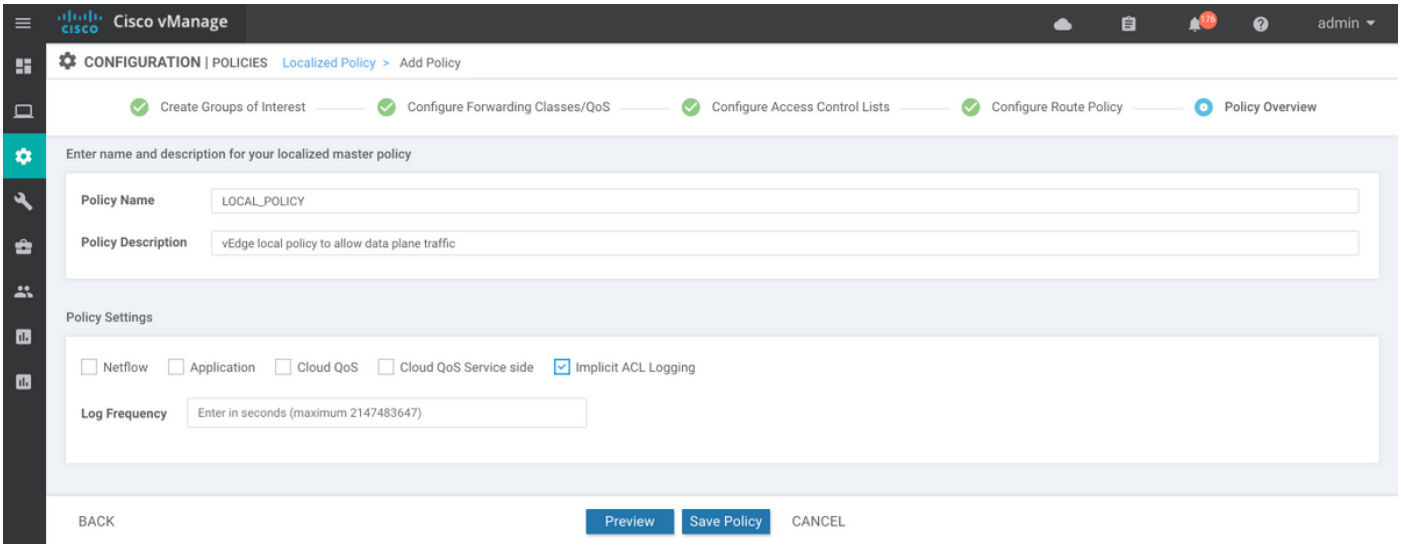
```
vpn 0
```

```
 interface ge0/1
  ip address 192.168.9.233/24
  nat
   respond-to-ping
  !
  tunnel-interface
   encapsulation ipsec
   color biz-internet
   no allow-service bgp
   no allow-service dhcp
   allow-service dns
   allow-service icmp
   no allow-service sshd
   no allow-service netconf
   no allow-service ntp
   no allow-service ospf
   no allow-service stun
   allow-service https
  !
  mtu      1506
  no shutdown
  access-list DATA_PLANE in
 !
!
policy
 implicit-acl-logging
 access-list DATA_PLANE
  sequence 10
   match
destination-port 12346 12445 protocol 17 ! action accept ! ! default-action drop ! !
```

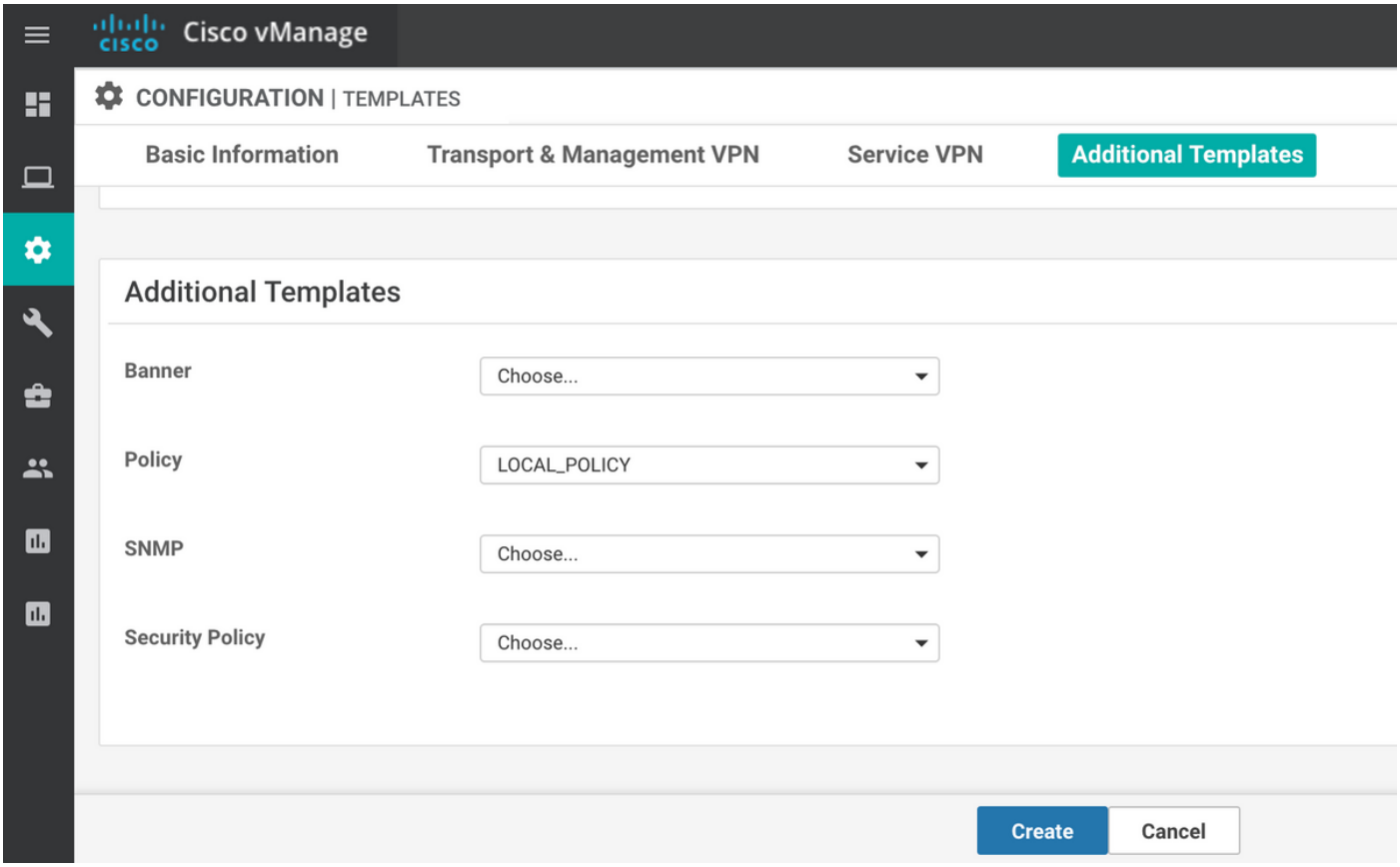
如果正在使用设备功能模板，则需要创建本地化策略并在配置访问控制列表向**导步骤上配**置ACL:



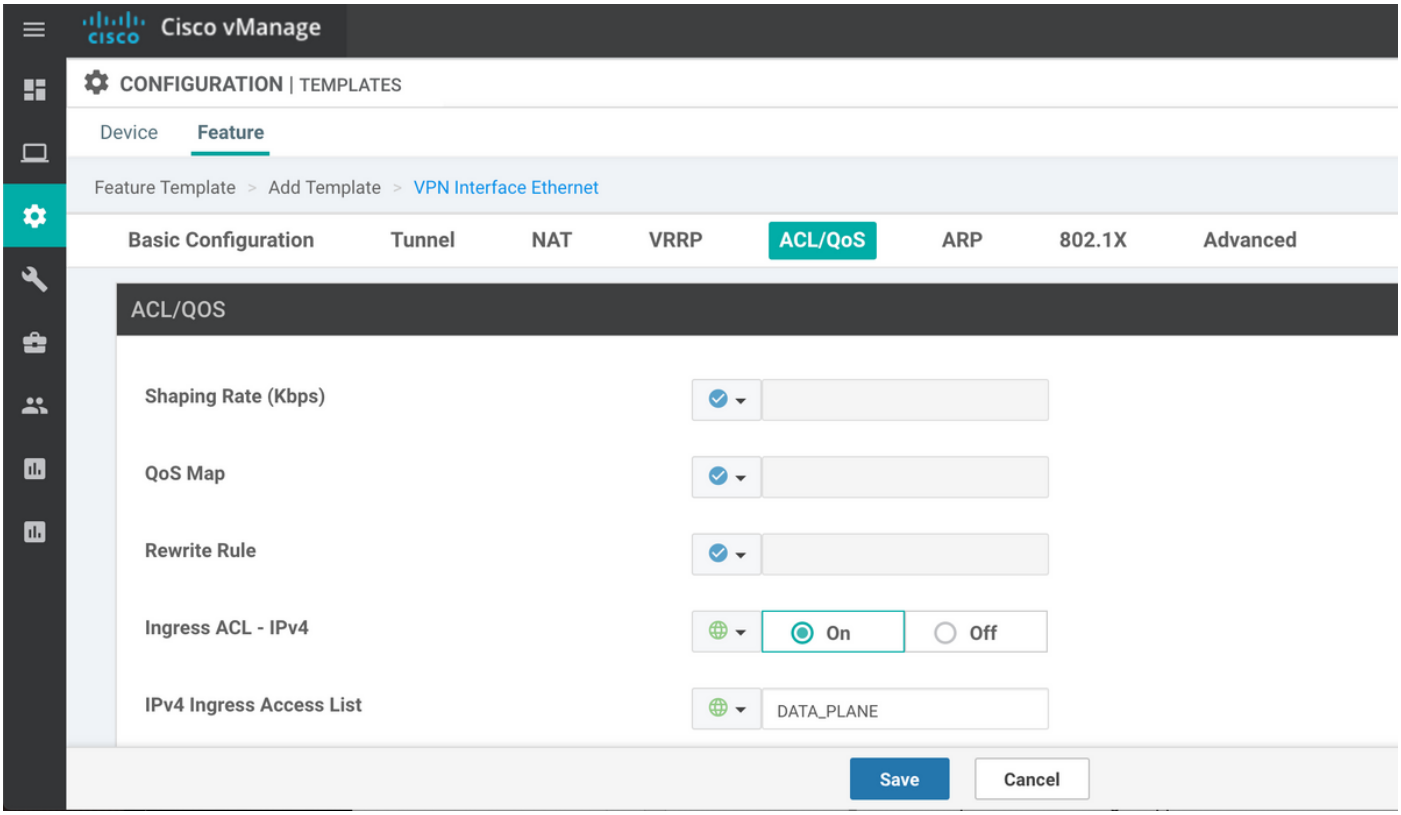
如果**尚未启用**implicit-acl-logging，则最好在最后一步中启用它，然后单击“保存策**略”按**钮：



本地化策略(**在本例中**命名为LOCAL_POLICY)应在设备模板中引用：



然后，应在VPN接口以太**网功能模**板下的入口(in)方向应用ACL（在本例中为命名DATA_PLANE）：



一旦配置ACL并将其应用到接口以绕过数据平面流量，BFD会话就会再次**进入**up状态：

```
vEdge2# show tunnel statistics dest-ip 198.51.100.232 ; show bfd sessions site-id 232


TCP
TUNNEL                                      SOURCE  DEST
TUNNEL                                              MSS
PROTOCOL  SOURCE IP      DEST IP         PORT    PORT    SYSTEM IP     LOCAL COLOR    REMOTE COLOR
MTU     tx-pkts  tx-octets  rx-pkts  rx-octets  ADJUST
-------------------------------------------------------------------------------------------------
-----------------------------------------------------------
ipsec     192.168.9.233  198.51.100.232  12346   42346  10.10.10.232  biz-internet  biz-internet
1441    1768     304503     1768     304433     1361

                                        SOURCE TLOC      REMOTE TLOC
DST PUBLIC                         DST PUBLIC         DETECT      TX
SYSTEM IP        SITE ID  STATE       COLOR            COLOR            SOURCE IP
IP                                  PORT        ENCAP  MULTIPLIER  INTERVAL(msec) UPTIME
TRANSITIONS
-------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------
-------------
10.10.10.232     232      up          biz-internet     biz-internet     192.168.9.233
198.51.100.232                  52346       ipsec  7           1000          0:00:14:36      0
```

其他注意事项

请注意，使用ACL的解决方法比NAT端口转发更实用，因为您还可以根据远程站点的源地址进行匹配，以提高安全性并防止对设备的DDoS攻击，例如：

```
 access-list DATA_PLANE
  sequence 10
```

```
   match
    source-ip        198.51.100.232/32
    destination-port 12346 12445
    protocol         17
   !
   action accept
   !
  !
```

另请注意，对于任何其他传入流量(未使用**allowed-services**指定)，例如，对于默认**iperf** 端口5001显式ACL **seq 20**，与本示例中的流量相比，这不会产生任何影响：

```
policy
 access-list DATA_PLANE
  sequence 10
   match
    source-ip        198.51.100.232/32
    destination-port 12346 12445
    protocol         17
   !
   action accept
   !
  !
  sequence 20
   match
    destination-port 5001
    protocol         6
   !
   action accept
   !
  !
```

您仍需要NAT端口转发免除规则才能**使**用iperf:

```
vEdgeCloud2# show running-config vpn 0 interface ge0/1 nat
vpn 0
 interface ge0/1
  nat
   respond-to-ping
   port-forward port-start 5001 port-end 5001 proto tcp
    private-vpn        0
    private-ip-address 192.168.9.233
   !
  !
 !
!
```

结论

这是由NAT软件设计细节导致的、无法避免的vEdge路由器上的预期行为。