当流量遵循非对称路径时，TCP连接无法建立

目录

简介

本文档介绍在SD-WAN交换矩阵中使用非对称路径进行流量转发时出现的问题。

问题

无法从host1(hostname - edgelin1)建立到host2(hostname - edgeclien2)的安全外壳(SSH)连接，但同时SSH在相反方向工作正常。

```
[root@edgeclient2 user]# ssh user@192.168.40.21
user@192.168.40.21's password:
Last login: Sun Feb 10 13:26:32 2019 from 192.168.60.20
[user@edgeclient1 ~]$
```

```
[root@edgeclient1 user]# ssh user@192.168.60.20
<nothing happens after that>
```

或

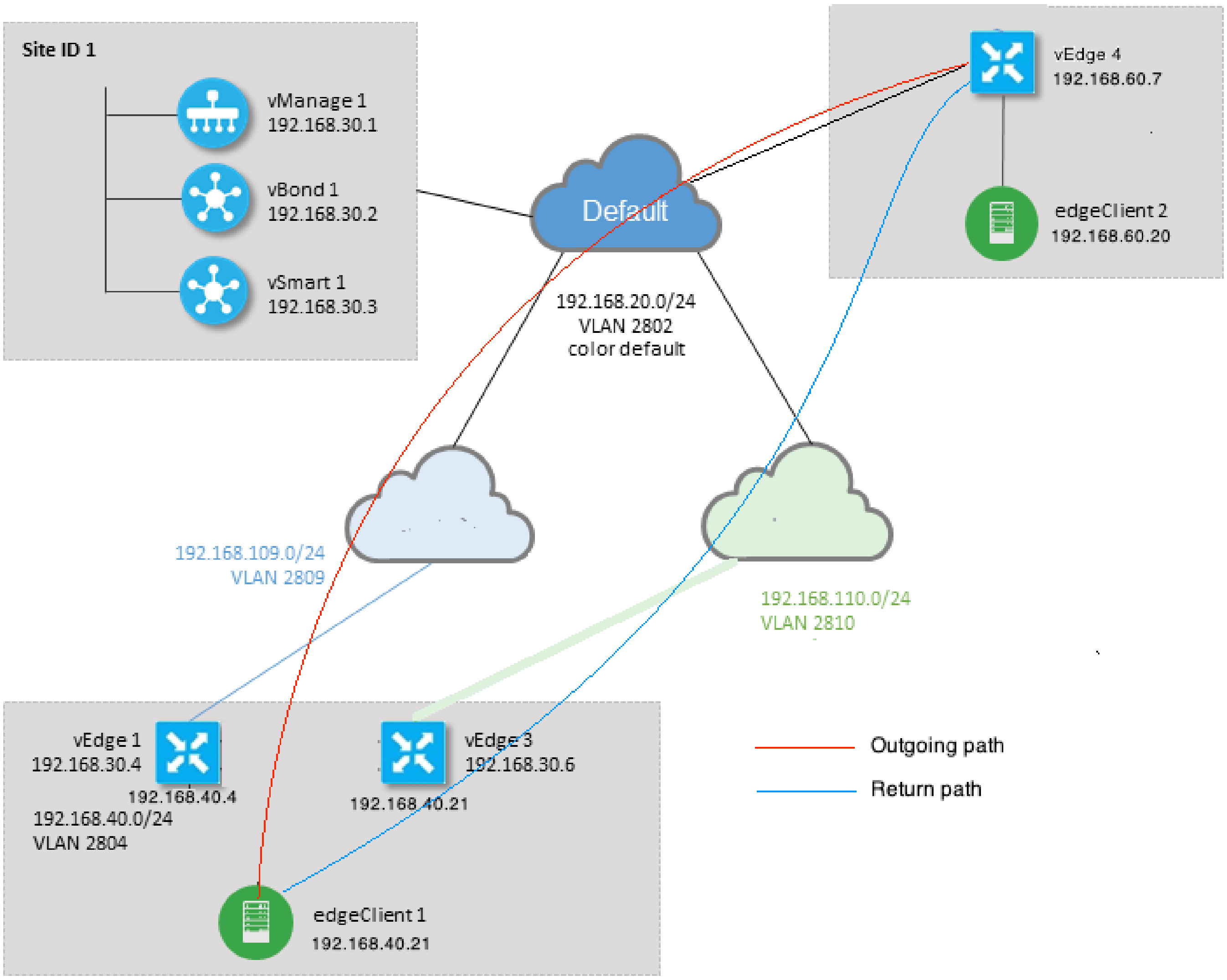
```
[user@edgeclient1 ~]$ ssh user@192.168.60.20
ssh_exchange_identification: Connection closed by remote host
```

edgeclient1和edgeclient2 SSH守护程序和客户端都具有已知的良好配置，并且可以从本地LAN网段成功建立连接：

```
vedge4# request execute vpn 40 ssh user@192.168.60.20
user@192.168.60.20's password:
Last login: Sun Feb 10 13:28:23 2019 from 192.168.60.7
[user@edgeclient2 ~]$
```

所有其他传输控制协议(TCP)应用都有类似的问题。

拓扑图



诊断

在vEdge1和vEdge3的服务端接口上，在相应方向上配置并应用了此访问控制列表(ACL):

```
policy
 access-list SSH_IN
  sequence 10
   match
    source-ip      192.168.40.21/32
    destination-ip 192.168.60.20/32
   !
   action accept
    count SSH_IN
   !
  !
  default-action accept
 !
 access-list SSH_OUT
  sequence 10
   match
    source-ip      192.168.60.20/32
    destination-ip 192.168.40.21/32
   !
   action accept
```

```
    count SSH_OUT
   !
  !
  default-action accept
 !
!
```

已在vEdge4上应用镜像ACL:

```
policy
 access-list SSH_IN
  sequence 10
   match
    source-ip      192.168.60.20/32
    destination-ip 192.168.40.21/32
   !
   action accept
    count SSH_IN
   !
  !
  default-action accept
 !
 access-list SSH_OUT
  sequence 10
   match
    source-ip      192.168.40.21/32
    destination-ip 192.168.60.20/32
   !
   action accept
    count SSH_OUT
   !
  !
  default-action accept
 !
!
```

此外，所有vEdge路由器上都启用了应用可视性，并且在SSH连接建立阶段检查了流：

```
vedge1# show app cflowd flows | tab ; show policy access-list-counters

                                                             TCP
TIME    EGRESS  INGRESS
                                   SRC    DEST        IP     CNTRL  ICMP                   TOTAL
TOTAL  MIN  MAX                            TO      INTF    INTF
VPN  SRC IP         DEST IP        PORT   PORT  DSCP  PROTO  BITS   OPCODE  NHOP IP        PKTS
BYTES  LEN  LEN  START TIME                EXPIRE  NAME    NAME
------------------------------------------------------------------------------------------------
----------------------------------------------------------------------
40   192.168.40.21  192.168.60.20  47866  22    0     6      24     0       192.168.109.7  3
227    66   87   Sun Feb 17 14:13:25 2019  34      ge0/0   ge0/1


         COUNTER
NAME     NAME     PACKETS  BYTES
----------------------------------
SSH_IN   SSH_IN   3        227
SSH_OUT  SSH_OUT  2        140

vedge3# show app cflowd flows | tab ; show policy access-list-counters
```

```
                                                             TCP
TIME    EGRESS  INGRESS
                                   SRC   DEST         IP     CNTRL  ICMP                   TOTAL
TOTAL  MIN  MAX                            TO      INTF    INTF
VPN  SRC IP         DEST IP        PORT  PORT   DSCP  PROTO  BITS   OPCODE  NHOP IP        PKTS
BYTES  LEN  LEN  START TIME                EXPIRE  NAME    NAME
------------------------------------------------------------------------------------------------
----------------------------------------------------------------------
40   192.168.60.20  192.168.40.21  22    47866  0     6      18     0       192.168.40.21  8
480    60   60   Sun Feb 17 14:14:08 2019  51      ge0/1   ge0/0


         COUNTER
NAME     NAME     PACKETS  BYTES
----------------------------------
SSH_IN   SSH_IN   0        0
SSH_OUT  SSH_OUT  7        420

vedge4# show app cflowd flows | tab ; show policy access-list-counters

                                                              TCP
TIME    EGRESS  INGRESS
                                   SRC    DEST         IP     CNTRL  ICMP
TOTAL  TOTAL  MIN  MAX                            TO      INTF    INTF
VPN  SRC IP         DEST IP        PORT   PORT   DSCP  PROTO  BITS   OPCODE  NHOP IP        PKTS
BYTES  LEN  LEN  START TIME                EXPIRE  NAME    NAME
------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------
40   192.168.40.21  192.168.60.20  47866  22     0     6      2      0       192.168.60.20  4
240    60   60   Sun Feb 17 14:17:44 2019  37      ge0/2   ge0/0
40   192.168.60.20  192.168.40.21  22     47866  0     6      18     0       192.168.110.6  8
592    74   74   Sun Feb 17 14:17:44 2019  49      ge0/0   ge0/2


         COUNTER
NAME     NAME     PACKETS  BYTES
----------------------------------
SSH_IN   SSH_IN   8        592
SSH_OUT  SSH_OUT  4        240
```

从这些输出中可以看到，入站和出站流是非对称的。edgeclient1(192.168.40.21)正在尝试与edgeclient2(192.168.60.20)建立SSH会话，传入流量通过vEdge1返回，返回流量通过vEdge3返回。从ACL计数器中，您还可以看到该传入和传出的数量vEdge4上的数据包与vEdge1和vEdge3上相应方向的和不匹配。同时，使用ping测试时没有丢**包**:

```
[root@edgeclient1 user]# ping -f 192.168.60.20 -c 10000
PING 192.168.60.20 (192.168.60.20) 56(84) bytes of data.

--- 192.168.60.20 ping statistics ---
10000 packets transmitted, 10000 received, 0% packet loss, time 3076ms
rtt min/avg/max/mdev = 0.128/0.291/6.607/0.623 ms, ipg/ewma 0.307/0.170 ms

[root@edgeclient2 user]# ping -f 192.168.40.21 -c 10000
PING 192.168.40.21 (192.168.40.21) 56(84) bytes of data.

--- 192.168.40.21 ping statistics ---
10000 packets transmitted, 10000 received, 0% packet loss, time 3402ms
rtt min/avg/max/mdev = 0.212/0.318/2.766/0.136 ms, ipg/ewma 0.340/0.327 ms
```

另外，我们还回顾了SSH在反向运行良好，并且文件也可以通过scp/sftp复制，而不会出现任何问题。

解决方案

最初怀疑存在某些深度数据包检测(DPI)配置或数据策略，但没有激活这些配置或数据策略：

```
vedge3# show policy from-vsmart
% No entries found.

vedge1# show policy from-vsmart
% No entries found.
```

但最终发现TCP优化已启用：

```
vedge1# show app tcp-opt active-flows

                                                                          EGRESS  INGRESS
                                   SRC    DEST                            INTF    INTF     TX
RX                  UNOPT   PROXY
VPN  SRC IP         DEST IP        PORT   PORT  START TIME                NAME    NAME     BYTES
BYTES  TCP STATE    REASON  IDENTITY
------------------------------------------------------------------------------------------------
--------------------------------------------
40   192.168.40.21  192.168.60.20  47868  22    Sun Feb 17 14:18:13 2019  ge0_0   ge0_1    314
0      In-progress  -       Client-Proxy

vedge1# show app tcp-opt expired-flows

                                                  SRC    DEST
TX     RX                UNOPT   PROXY
TIMESTAMP      VPN  SRC IP         DEST IP        PORT   PORT   START TIME                END
TIME                  BYTES  BYTES  TCP STATE  REASON  IDENTITY      DELETE REASON
------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------
1549819969608  40   192.168.40.21  192.168.60.7   22     56612  Sun Feb 10 18:32:49 2019  Sun
Feb 10 18:36:03 2019  5649   4405   Optimized  -       Server-Proxy  CLOSED
1549820055487  40   192.168.40.21  192.168.60.7   22     56613  Sun Feb 10 18:34:15 2019  Sun
Feb 10 19:07:46 2019  5719   4669   Optimized  -       Server-Proxy  CLOSED
1550408210511  40   192.168.40.21  192.168.60.20  47862  22     Sun Feb 17 13:56:50 2019  Sun
Feb 17 13:56:58 2019  401    0      Optimized  -       Client-Proxy  STATE-TIMEOUT
1550408981634  40   192.168.40.21  192.168.60.20  47864  22     Sun Feb 17 14:09:41 2019  Sun
Feb 17 14:09:49 2019  401    0      Optimized  -       Client-Proxy  STATE-TIMEOUT
1550409205399  40   192.168.40.21  192.168.60.20  47866  22     Sun Feb 17 14:13:25 2019  Sun
Feb 17 14:13:33 2019  227    0      Optimized  -       Client-Proxy  STATE-TIMEOUT
1550409493042  40   192.168.40.21  192.168.60.20  47868  22     Sun Feb 17 14:18:13 2019  Sun
Feb 17 14:18:21 2019  401    0      Optimized  -       Client-Proxy  STATE-TIMEOUT
```

此外，在debugs ftm tcpopt **CONN_TEARDOWN消**息中可以看到。

```
vedge1# show log /var/log/tmplog/vdebug tail "-f"
local7.debug: Feb 17 13:56:50 vedge1 FTMD[662]: ftm_tcpopt_flow_add[268]: Created new tcpflow :-
vrid-3 192.168.40.21/47862 192.168.60.20/22
local7.debug: Feb 17 13:56:58 vedge1 FTMD[662]: ftm_tcpd_send_conn_tear_down[388]: Trying to
pack and send the following message to TCPD
local7.debug: Feb 17 13:56:58 vedge1 FTMD[662]: ftm_tcpd_send_conn_tear_down[408]: Sending
following CONN_TD msg
local7.debug: Feb 17 13:56:58 vedge1 FTMD[662]: ftm_tcpd_send_conn_tear_down[413]:
192.168.40.21:47862->192.168.60.20:22; vpn:40; syn_seq_num:4172167164; identity:0; cport_prime:0
local7.debug: Feb 17 13:56:58 vedge1 FTMD[662]: ftm_tcpd_msgq_tx[354]: Transfering size = 66
```

```
bytes data
local7.debug: Feb 17 13:56:58 vedge1 FTMD[662]: ftm_tcpd_send_conn_tear_down[416]: Successfully
sent conn_td msg to TCPD
local7.debug: Feb 17 13:56:58 vedge1 FTMD[662]: ftm_tcpopt_propagate_tear_down[1038]: Sent
CONN_TEARDOWN msg to tcpd for existing tcpflow :- vrid-3 192.168.40.21/47862 192.168.60.20/22 ;
identity:CLIENT_SIDE_PROXY . Send Successful !
local7.debug: Feb 17 13:56:58 vedge1 FTMD[662]: ftm_tcpopt_append_expired_err_flow_tbl[958]:
Appending flow vrid-3 192.168.40.21/47862 192.168.60.20/22  to the expired flow table at Sun Feb
17 13:56:58 2019
local7.debug: Feb 17 13:56:58 vedge1 FTMD[662]: ftm_tcpopt_append_expired_err_flow_tbl[980]:
Appending flow vrid-3 192.168.40.21/47862 192.168.60.20/22  to the error flow table at Sun Feb
17 13:56:58 2019
local7.debug: Feb 17 13:56:58 vedge1 FTMD[662]: ftm_tcpopt_flow_delete[293]: Removing tcpflow :-
vrid-3 192.168.40.21/47862 192.168.60.20/22
local7.debug: Feb 17 13:56:58 vedge1 TCPD[670]: handle_upstream_connect[538]: Error - BP NULL
local7.debug: Feb 17 13:56:58 vedge1 FTMD[662]: ftm_tcpd_msg_decode[254]: FTM-TCPD: Received
FTM_TCPD__PB_FTM_TCPD_MSG__E_MSG_TYPE__CONN_CLOSED msg
local7.debug: Feb 17 13:56:58 vedge1 FTMD[662]: ftm_tcpd_handle_conn_closed[139]: FTM-TCPD:
Received CONN_CLOSED for following C->S
local7.debug: Feb 17 13:56:58 vedge1 FTMD[662]: ftm_tcpd_handle_conn_closed[150]:
192.168.40.21:47862->192.168.60.20:22; vpn:40; syn_seq_num:4172167164; identity:0;
cport_prime:47862; bind_port:0
local7.debug: Feb 17 13:56:58 vedge1 FTMD[662]: ftm_tcpd_handle_conn_closed[184]: FTM-TCPD:
Could not find entry in FT for following flow
local7.debug: Feb 17 13:56:58 vedge1 FTMD[662]: ftm_tcpd_handle_conn_closed[185]: vrid-3
192.168.40.21/47862 192.168.60.20/22
```

此处您可以看到TCP优化工作正常时的示例（可以看到CONN_EST消息）：

```
vedge3# show log /var/log/tmplog/vdebug tail "-f -n 0"
local7.debug: Feb 17 15:41:13 vedge3 FTMD[657]: ftm_tcpd_msg_decode[254]: FTM-TCPD: Received
FTM_TCPD__PB_FTM_TCPD_MSG__E_MSG_TYPE__CONN_CLOSED msg
local7.debug: Feb 17 15:41:13 vedge3 FTMD[657]: ftm_tcpd_handle_conn_closed[139]: FTM-TCPD:
Received CONN_CLOSED for following C->S
local7.debug: Feb 17 15:41:13 vedge3 FTMD[657]: ftm_tcpd_handle_conn_closed[150]:
192.168.40.21:47876->192.168.60.20:22; vpn:40; syn_seq_num:2779178897; identity:0;
cport_prime:47876; bind_port:0
local7.debug: Feb 17 15:41:15 vedge3 FTMD[657]: ftm_tcpd_msg_decode[258]: FTM-TCPD: Received
FTM_TCPD__PB_FTM_TCPD_MSG__E_MSG_TYPE__CONN_EST msg
local7.debug: Feb 17 15:41:15 vedge3 FTMD[657]: ftm_tcpd_handle_conn_est[202]: FTM-TCPD:
Received CONN_EST for following C->S
local7.debug: Feb 17 15:41:15 vedge3 FTMD[657]: ftm_tcpd_handle_conn_est[213]:
192.168.40.21:47878->192.168.60.20:22; vpn:40; syn_seq_num:2690847868; identity:0;
cport_prime:47878; bind_port:0
local7.debug: Feb 17 15:41:15 vedge3 FTMD[657]: ftm_tcpopt_flow_add[268]: Created new tcpflow :-
vrid-3 192.168.40.21/47878 192.168.60.20/22
```

结论

TCP优化要求流对称，因此要解决此问题，必须禁用TCP优化**(no vpn 40 tcp优化)或必须创建数据策略，以强制TCP流在两个方向上采用相同路径。**有关此信息，请参阅SD-WAN设计指南第23页的Traffic Symmetry for DPI。