不恰当地使用“policy action set tloc-list”导致流量黑洞

目录

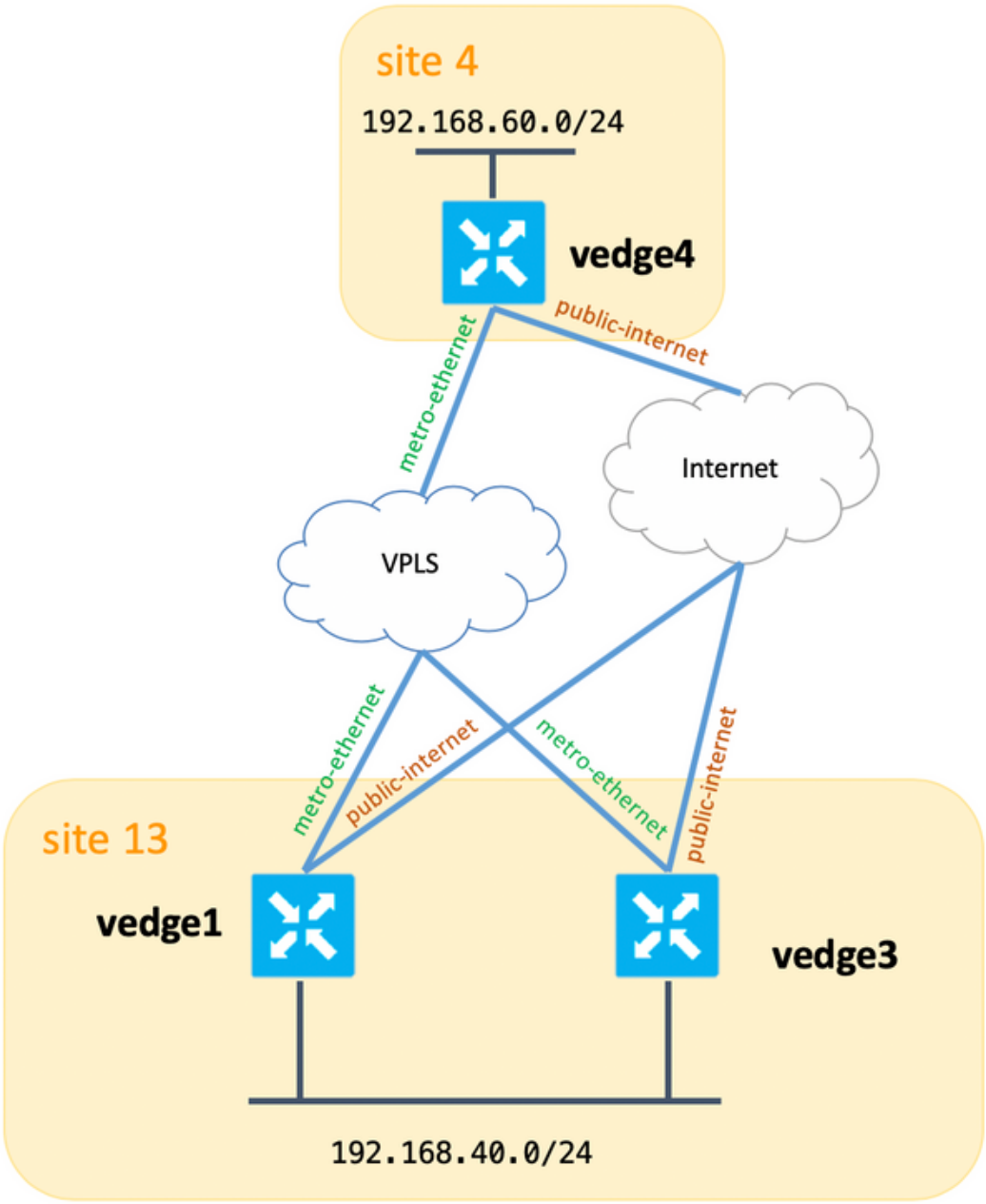
简介

本文档介绍**set tloc-list**操作的不适当的策略应用，当首选链路断开但备用路径仍然可用的某些情况下，该操作会导致流量黑洞。

> **注意**：本文档中显示的所有命令输出均来自vEdge路由器。但是，对于运行IOS®-XE SDWAN软件的路由器，故障排除方法仍然相同。使用**sdwan**关键字在IOS®-XE SDWAN软件上获取相同的输出。例如，**show sdwan omp routes**而不是**show omp routes**。

背景信息

为了便于演示和更好地理解后面描述的问题，请考虑以下拓扑图：



此外，下表汇总了系统设置：

主机名	站点ID	system-ip
vedge1	13 个	10.155.0.118
vedge3	13 个	10.155.0.120
vedge4	4	10.155.0.50
vsmart1	1	10.155.0.3

vEdge1和vEdge3都配置了指向服务端VPN中某个下一跳的静态路由：

```
vpn 40
 ip route 10.223.115.101/32 192.168.40.10
!
```

为了实现这些目标：

1.使vEdge1城域以太网链路成为进入“站点13”的入口流量的首选链路。

2. M将vEdge3城域以太网链路作为进入“站点13”的入口流量的第二个首选链路。

3.将vEdge1公共Internet链路作为进入“站点13”的入口流量的第三个首选链路。

4.使vEdge3公共 — internet链路成为进入“站点13”的入口流量的最低首选链路。

此vSmart控制策略配置如下：

```
policy
 lists
  tloc-list SITE13_TLOC_PREF
   tloc 10.155.0.118 color metro-ethernet encap ipsec preference 200
   tloc 10.155.0.118 color public-internet encap ipsec preference 100
   tloc 10.155.0.120 color metro-ethernet encap ipsec preference 150
   tloc 10.155.0.120 color public-internet encap ipsec preference 50
  !
  prefix-list SITE13_PREFIX
   ip-prefix 10.223.115.101/32
  !
  site-list site13
   site-id 13
  !
 control-policy TE_POLICY_2_SITE4
  sequence 10
   match route
    prefix-list SITE13_PREFIX
   !
   action accept
    set
     tloc-list SITE13_TLOC_PREF
    !
   !
  !
  default-action accept
 !
!
apply-policy
 site-list site4
  control-policy TE_POLICY_2_SITE4 out
 !
!
```

问题

正常条件

vSmart通过4个可能的TLOC作为下一跳获取这些路由：

```
vsmart1# show omp routes 10.223.115.101/32 | b PATH
                                          PATH                    ATTRIBUTE
VPN    PREFIX              FROM PEER      ID     LABEL    STATUS    TYPE       TLOC IP
COLOR            ENCAP  PREFERENCE
-----------------------------------------------------------------------------------------------
-------------------------------------
40     10.223.115.101/32   10.155.0.118   35     1002     C,R       installed  10.155.0.118
```

```
metro-ethernet   ipsec  -
                          10.155.0.118     37     1002     C,R       installed  10.155.0.118
public-internet  ipsec  -
                          10.155.0.120     35     1002     C,R       installed  10.155.0.120
metro-ethernet   ipsec  -
                          10.155.0.120     37     1002     C,R       installed  10.155.0.120
public-internet  ipsec  -
```

并相应地设置通告路由的优先级：

```
vsmart1# show omp routes 10.223.115.101/32 detail | nomore | b ADVERTISED | b "peer
10.155.0.50" | i Attributes\|originator\|\ tloc\|preference
    Attributes:
     originator       10.155.0.118
     tloc             10.155.0.120, public-internet, ipsec
     preference       50
    Attributes:
     originator       10.155.0.118
     tloc             10.155.0.120, metro-ethernet, ipsec
     preference       150
    Attributes:
     originator       10.155.0.118
     tloc             10.155.0.118, public-internet, ipsec
     preference       100
    Attributes:
     originator       10.155.0.118
     tloc             10.155.0.118, metro-ethernet, ipsec
     preference       200
```

vEdge4选择适当的TLOC并将此路由安装到路由表中：

```
vedge4# show ip routes 10.223.115.101/32 | b PROTOCOL
                                          PROTOCOL  NEXTHOP     NEXTHOP          NEXTHOP
VPN    PREFIX              PROTOCOL       SUB TYPE  IF NAME     ADDR             VPN      TLOC
IP          COLOR            ENCAP  STATUS
-----------------------------------------------------------------------------------------------
-------------------------------------------
40     10.223.115.101/32   omp            -         -           -                -
10.155.0.118     metro-ethernet   ipsec  F,S
```

流量转发按预期工作：

```
vedge4# traceroute vpn 40 10.223.115.101
Traceroute  10.223.115.101 in VPN 40
traceroute to 10.223.115.101 (10.223.115.101), 30 hops max, 60 byte packets
 1  192.168.40.4 (192.168.40.4)  0.835 ms  0.984 ms  1.097 ms
 2  192.168.40.10 (192.168.40.10)  2.955 ms  3.056 ms  3.218 ms
```

故障情况

最终，vEdge1上发生故障，面向服务端LAN的接口关闭（或者管理员关闭以进行测试，例如，结果将相同）：

```
vedge1# show interface vpn 40

                                       IF      IF      IF
TCP
                AF                     ADMIN   OPER    TRACKER  ENCAP  PORT
SPEED           MSS             RX      TX
VPN  INTERFACE  TYPE  IP ADDRESS       STATUS  STATUS  STATUS   TYPE   TYPE     MTU   HWADDR
MBPS   DUPLEX  ADJUST  UPTIME  PACKETS  PACKETS
-------------------------------------------------------------------------------------------
-----------------------------------------------------
40   ge0/4      ipv4  192.168.40.4/24  Up    Down    NA       null   service  1500
00:50:56:be:91:36  -      -       1420    -       129768   0
```

由于vEdge1没有10.223.115.101/32路由的有效下一跳，因此该路由会从路由和转发表中删除，并且不再将其通告给vSmart:

```
vedge1# show ip routes 10.223.115.101/32 | b PROTO
                                          PROTOCOL  NEXTHOP     NEXTHOP          NEXTHOP
VPN    PREFIX              PROTOCOL       SUB TYPE  IF NAME     ADDR             VPN      TLOC
IP          COLOR            ENCAP  STATUS
-------------------------------------------------------------------------------------------
-----------------------------------------
40     10.223.115.101/32   static           -         -           192.168.40.21    -        -
-                -      I

vedge1# show ip fib vpn 40 | i 10.223.115.101/32
vedge1#
vedge1# show omp routes 10.223.115.101/32 detail | nomore | b ADVERTISED
vedge1#
```

同时，vEdge3仍会通告此路由（这是预期结果）：

```
vedge3# show omp routes 10.223.115.101/32 detail | nomore | b ADVERTISED
            ADVERTISED TO:
peer    10.155.0.3
    Attributes:
     originator       10.155.0.120
     label            1002
     path-id          35
     tloc             10.155.0.120, metro-ethernet, ipsec
     ultimate-tloc    not set
     domain-id        not set
     site-id          13
     overlay-id        1
     preference       not set
     tag              not set
     origin-proto     static
     origin-metric    0
     as-path          not set
     unknown-attr-len not set
    Attributes:
     originator       10.155.0.120
     label            1002
     path-id          37
     tloc             10.155.0.120, public-internet, ipsec
     ultimate-tloc    not set
     domain-id        not set
     site-id          13
     overlay-id        1
```

```
     preference       not set
     tag              not set
     origin-proto     static
     origin-metric    0
     as-path          not set
     unknown-attr-len not set
```

vSmart现在可按预期从vEdge3获取2条路由：

```
vsmart1# show omp routes 10.223.115.101/32 | b PATH
                                          PATH                      ATTRIBUTE
VPN    PREFIX              FROM PEER      ID     LABEL    STATUS    TYPE       TLOC IP
COLOR            ENCAP  PREFERENCE
-----------------------------------------------------------------------------------------------
-------------------------------------
40     10.223.115.101/32   10.155.0.120   35     1002     C,R       installed  10.155.0.120
metro-ethernet   ipsec  -
                           10.155.0.120   37     1002     C,R       installed  10.155.0.120
public-internet  ipsec  -
```

但与此同时，vSmart继续宣传以下内容：

```
vsmart1# show omp routes 10.223.115.101/32 detail | nomore | b ADVERTISED | b "peer
10.155.0.50" | i Attributes\|originator\|\ tloc\|preference
    Attributes:
     originator       10.155.0.120
     tloc             10.155.0.120, public-internet, ipsec
     preference       50
    Attributes:
     originator       10.155.0.120
     tloc             10.155.0.120, metro-ethernet, ipsec
     preference       150
    Attributes:
     originator       10.155.0.120
     tloc             10.155.0.118, public-internet, ipsec
     preference       100
    Attributes:
     originator       10.155.0.120
     tloc             10.155.0.118, metro-ethernet, ipsec
     preference       200
```

如您所见，唯一的发起方已更改，这是预期行为，因为**tloc-list**操作与（大致来说）设置下一跳”操作类似，并且强制设置错误的TLOC，因此可达性丢失。

```
vedge4# ping vpn 40 10.223.115.101 count 5
Ping in VPN 40
PING 10.223.115.101 (10.223.115.101) 56(84) bytes of data.
^C
--- 10.223.115.101 ping statistics ---
5 packets transmitted, 0 received, 100% packet loss, time 3999ms

vedge4# traceroute vpn 40 10.223.115.101
Traceroute  10.223.115.101 in VPN 40
traceroute to 10.223.115.101 (10.223.115.101), 30 hops max, 60 byte packets
 1  * * *
 2  * * *
 3  * * *
```

```
 4  * * *
 5  * * *
```

解决方案

作为解决方案，为了避免设置错误的TLOC下一跳信息，提出了以下方法：

```
policy
 lists
  tloc-list vedge1-tlocs
   tloc 10.155.0.118 color metro-ethernet encap ipsec
   tloc 10.155.0.118 color public-internet encap ipsec
  !
  tloc-list vedge1-tlocs-preference
   tloc 10.155.0.118 color metro-ethernet encap ipsec preference 200
   tloc 10.155.0.118 color public-internet encap ipsec preference 100
  !
  tloc-list vedge3-tlocs
   tloc 10.155.0.120 color metro-ethernet encap ipsec
   tloc 10.155.0.120 color public-internet encap ipsec
  !
  tloc-list vedge3-tlocs-preference
   tloc 10.155.0.120 color metro-ethernet encap ipsec preference 150
   tloc 10.155.0.120 color public-internet encap ipsec preference 50
  !
 !
!
policy
 control-policy TE_POLICY_2_SITE4
  sequence 10
   match route
    prefix-list SITE13_PREFIX
    tloc-list   vedge1-tlocs
   !
   action accept
    set
     tloc-list vedge1-tlocs-preference
    !
   !
  !
  sequence 20
   match route
    prefix-list SITE13_PREFIX
    tloc-list   vedge3-tlocs
   !
   action accept
    set
     tloc-list vedge3-tlocs-preference
    !
   !
  !
  default-action accept
 !
!
```

此类策略可改善情况，防止通告具有错误TLOC下一跳的路由：

```
vsmart1# show omp routes 10.223.115.101/32 detail | nomore | b ADVERTISED | b "peer
10.155.0.50" | i Attributes\|originator\|\ tloc\|preference
    Attributes:
     originator       10.155.0.120
```

```
     tloc             10.155.0.120, public-internet, ipsec
     preference       50
    Attributes:
     originator       10.155.0.120
     tloc             10.155.0.120, metro-ethernet, ipsec
     preference       150
    Attributes:
     originator       10.155.0.120
     tloc             10.155.0.120, public-internet, ipsec
     preference       not set
```

因此，可保持整个故障场景的可达性：

```
vedge4# traceroute vpn 40 10.223.115.101
Traceroute  10.223.115.101 in VPN 40
traceroute to 10.223.115.101 (10.223.115.101), 30 hops max, 60 byte packets
 1  192.168.40.6 (192.168.40.6)  0.458 ms  0.507 ms  0.617 ms
 2  192.168.40.10 (192.168.40.10)  1.928 ms  1.976 ms  2.069 ms

vedge4# ping vpn 40 10.223.115.101
Ping in VPN 40
PING 10.223.115.101 (10.223.115.101) 56(84) bytes of data.
64 bytes from 10.223.115.101: icmp_seq=1 ttl=254 time=0.702 ms
64 bytes from 10.223.115.101: icmp_seq=2 ttl=254 time=0.645 ms
64 bytes from 10.223.115.101: icmp_seq=3 ttl=254 time=0.691 ms
64 bytes from 10.223.115.101: icmp_seq=4 ttl=254 time=0.715 ms
64 bytes from 10.223.115.101: icmp_seq=5 ttl=254 time=0.603 ms
^C
--- 10.223.115.101 ping statistics ---
5 packets transmitted, 5 received, 0% packet loss, time 4000ms
rtt min/avg/max/mdev = 0.603/0.671/0.715/0.044 ms
```

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）。