为什么流量不通过SD-WAN中心路由器的ECMP路径进行负载均衡

目录

简介

本文描述SD-WAN交换矩阵中等价多路径(ECMP)路由的典型问题，当来自分支路由器的流量未在通告相同前缀的多个中心路由器上进行负载均衡时。它还说明了如何解决此问题以及如何使用各种故障排除命令(包括**show sdwan policy service-path**)来排除17.2 Cisco IOS®-XE软件中添加的路由问题。

先决条件

要求

Cisco 建议您了解以下主题：

- 对重叠管理协议(OMP)的基本了解
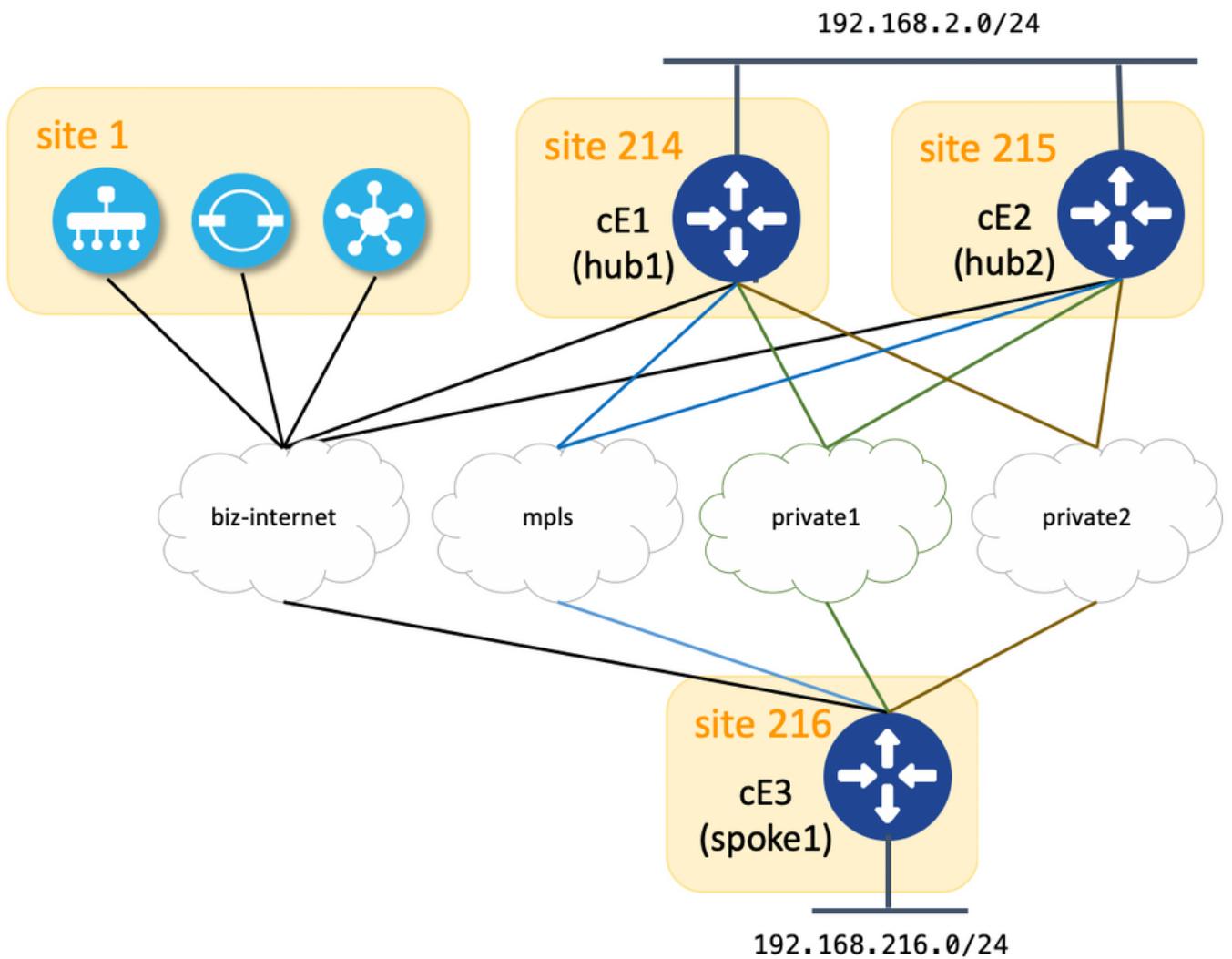- SD-WAN组件及其之间的交互

使用的组件

为便于演示，使用了以下软件路由器：

- 4台运行17.2.1v软件版本的Cisco IOS-XE CSR1000v路由器，在控制器模式(SD-WAN)下运行
- 运行20.1.12软件版本的vSmart控制器

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始（默认）配置。如果您的网络处于活动状态，请确保您了解所有命令的潜在影响。

背景信息

在本文中，我们使用以下实验拓扑：

192.168.2.0/24

site 1

site 214
cE1
(hub1)

site 215
cE2
(hub2)

biz-internet   mpls   private1   private2

site 216
cE3
(spoke1)

192.168.216.0/24

在此，您可以找到SD-WAN交换矩阵中每台设备的已分配站点ID和系统IP参数的摘要：

主机名	system-ip	站点ID
cE1（集线器1）	192.168.30.214	214
cE2（集线器2）	192.168.30.215	215
cE3(spoke1)	192.168.30.216	216
vSmart	192.168.30.113	1

每个集线器有4个TLOC(传输位置标识符)，根据拓扑图分配颜色，每个集线器将默认路由
0.0.0.0/0与192.168.2.0/24子网一起通告给辐条（分支路由器cE3）。vSmart上未配置策略以首选任
何路径/设备，并且所有OMP设置也设置为所有设备的默认设置。其余配置是基本SD-WAN重叠功
能的标准最小配置，因此为简便起见不提供。您可以期望从分支机构路由器到中心路由器的所有可
用上行链路实现负载均衡的主用 — 主用冗余和出口流量。

问题

分支路由器仅通过cE1路由器(hub1)安装默认路由和到192.168.2.0/24子网的路由：

```
ce3#show ip route vrf 2 | b Gateway Gateway of last resort is 192.168.30.214 to network 0.0.0.0
m* 0.0.0.0/0 [251/0] via 192.168.30.214, 00:08:30, sdwan_system_ip m 192.168.2.0/24 [251/0] via
```

```
192.168.30.214, 00:10:01, sdwan_system_ip 192.168.216.0/24 is variably subnetted, 2 subnets, 2
masks C 192.168.216.0/24 is directly connected, Loopback2 L 192.168.216.216/32 is directly
connected, Loopback2
```

这是因为cE3仅接收4条默认路由0.0.0.0/0和192.168.2.0/24的路由。

```
ce3#show sdwan omp routes vpn 2 | begin PATH PATH ATTRIBUTE VPN PREFIX FROM PEER ID LABEL STATUS
TYPE TLOC IP COLOR ENCAP PREFERENCE --------------------------------------------------------
----------------------------------------------------------------- 2 0.0.0.0/0
192.168.30.113 61614 1003 C,I,R installed 192.168.30.214 mpls ipsec - 192.168.30.113 61615 1003
C,I,R installed 192.168.30.214 biz-internet ipsec - 192.168.30.113 61616 1003 C,I,R installed
192.168.30.214 private1 ipsec - 192.168.30.113 61617 1003 C,I,R installed 192.168.30.214
private2 ipsec - 2 192.168.2.0/24 192.168.30.113 61610 1003 C,I,R installed 192.168.30.214 mpls
ipsec - 192.168.30.113 61611 1003 C,I,R installed 192.168.30.214 biz-internet ipsec -
192.168.30.113 61612 1003 C,I,R installed 192.168.30.214 private1 ipsec - 192.168.30.113 61613
1003 C,I,R installed 192.168.30.214 private2 ipsec - 2 192.168.216.0/24 0.0.0.0 68 1003 C,Red,R
installed 192.168.30.216 biz-internet ipsec - 0.0.0.0 81 1003 C,Red,R installed 192.168.30.216
private1 ipsec - 0.0.0.0 82 1003 C,Red,R installed 192.168.30.216 private2 ipsec -
```

虽然在vSmart上，您可以看到它收到所有8条路由（每个集线器上每个TLOC颜色有4条路由）：

```
vsmart1# show omp routes vpn 2 | b PATH PATH ATTRIBUTE VPN PREFIX FROM PEER ID LABEL STATUS TYPE
TLOC IP COLOR ENCAP PREFERENCE ------------------------------------------------------------
----------------------------------------------------------------- 2 0.0.0.0/0 192.168.30.214
66 1003 C,R installed 192.168.30.214 mpls ipsec - 192.168.30.214 68 1003 C,R installed
192.168.30.214 biz-internet ipsec - 192.168.30.214 81 1003 C,R installed 192.168.30.214 private1
ipsec - 192.168.30.214 82 1003 C,R installed 192.168.30.214 private2 ipsec - 192.168.30.215 66
1003 C,R installed 192.168.30.215 mpls ipsec - 192.168.30.215 68 1003 C,R installed
192.168.30.215 biz-internet ipsec - 192.168.30.215 81 1003 C,R installed 192.168.30.215 private1
ipsec - 192.168.30.215 82 1003 C,R installed 192.168.30.215 private2 ipsec - 2 192.168.2.0/24
192.168.30.214 66 1003 C,R installed 192.168.30.214 mpls ipsec - 192.168.30.214 68 1003 C,R
installed 192.168.30.214 biz-internet ipsec - 192.168.30.214 81 1003 C,R installed
192.168.30.214 private1 ipsec - 192.168.30.214 82 1003 C,R installed 192.168.30.214 private2
ipsec - 192.168.30.215 66 1003 C,R installed 192.168.30.215 mpls ipsec - 192.168.30.215 68 1003
C,R installed 192.168.30.215 biz-internet ipsec - 192.168.30.215 81 1003 C,R installed
192.168.30.215 private1 ipsec - 192.168.30.215 82 1003 C,R installed 192.168.30.215 private2
ipsec -
```

如果来自cE1(hub1)的默认路由丢失，分支路由器会安装来自cE2(hub2)的路由。 因此，在cE1充当
主路由器的情况下，不存在主用 — 主用冗余和主用 — 备用冗余。

您还可以通过show sdwan policy service-path命令检查为特定流量采用哪个出口路径，如以下示例
所示：

```
ce3#show sdwan policy service-path vpn 2 interface Loopback2 source-ip 192.168.216.216 dest-ip
192.168.2.1 protocol 6 source-port 53453 dest-port 22 dscp 48 app ssh Next Hop: IPsec Source:
192.168.109.216 12347 Destination: 192.168.110.214 12427 Local Color: biz-internet Remote Color:
mpls Remote System IP: 192.168.30.214
```

要查看特定流量类型的所有可用路径，请使用**all关键字**：

```
ce3#show sdwan policy service-path vpn 2 interface Loopback2 source-ip 192.168.216.216 dest-ip
192.168.2.1 protocol 6 source-port 53453 dest-port 22 dscp 48 app ssh all Number of possible
next hops: 4 Next Hop: IPsec Source: 192.168.109.216 12347 Destination: 192.168.110.214 12427
Local Color: biz-internet Remote Color: mpls Remote System IP: 192.168.30.214 Next Hop: IPsec
Source: 192.168.108.216 12367 Destination: 192.168.108.214 12407 Local Color: private2 Remote
Color: private2 Remote System IP: 192.168.30.214 Next Hop: IPsec Source: 192.168.107.216 12367
Destination: 192.168.107.214 12407 Local Color: private1 Remote Color: private1 Remote System
IP: 192.168.30.214 Next Hop: IPsec Source: 192.168.109.216 12347 Destination: 192.168.109.214
12387 Local Color: biz-internet Remote Color: biz-internet Remote System IP: 192.168.30.214
```

因此，这也确认只有4条路径可用，而不是8条路由器cE3(spoke2)可用。

如果您检查vSmart到底通告什么，您只会看到4条通告到cE3的路由：

```
vsmart1# show omp routes vpn 2 0.0.0.0/0 detail | nomore | exclude not\ set | b ADVERTISED\ TO:
| b peer\ \ \ \ 192.168.30.216 peer 192.168.30.216 Attributes: originator 192.168.30.214 label
1003 path-id 61629 tloc 192.168.30.214, private2, ipsec site-id 214 overlay-id 1 origin-proto
static origin-metric 0 Attributes: originator 192.168.30.214 label 1003 path-id 61626 tloc
192.168.30.214, mpls, ipsec site-id 214 overlay-id 1 origin-proto static origin-metric 0
Attributes: originator 192.168.30.214 label 1003 path-id 61628 tloc 192.168.30.214, private1,
ipsec site-id 214 overlay-id 1 origin-proto static origin-metric 0 Attributes: originator
192.168.30.214 label 1003 path-id 61627 tloc 192.168.30.214, biz-internet, ipsec site-id 214
overlay-id 1 origin-proto static origin-metric 0
```
根据此输出，您可以断定问题是由vSmart控制器引起的。

解决方案

此行为是由vSmart控制器上**send-path-limit**的默认配置导致的。**send-path-limit**定义从边缘路由器通告给vSmart控制器以及从vSmart控制器通告给其他边缘路由器的ECMP路由的最大数量。默认值为4，通常情况下，对于边缘路由器（如此拓扑中每个中心路由器上有4个上行链路）而言，它已足够，但对于vSmart控制器而言，它不足以将所有可用路径发送到其他边缘路由器。可为**send-path-limit**设置的最大值为16，但在某些极端情况下，这仍然不够，尽管已打开增强请求CSCvs89015将最大值增加到128。

要解决此问题，必须重新配置vSmart设置，如下例所示：

```
vsmart1# conf t Entering configuration mode terminal vsmart1(config)# omp vsmart1(config-omp)#
send-path-limit 8 vsmart1(config-omp)# commit Commit complete. vsmart1(config-omp)# end vsmart1#
show run omp omp no shutdown send-path-limit 8 graceful-restart ! vsmart1#
```
然后，所有8条路由都由vSmart通告给分支路由器，并由它们接收：

```
ce3#show sdwan omp routes vpn 2 | begin PATH PATH ATTRIBUTE VPN PREFIX FROM PEER ID LABEL STATUS
TYPE TLOC IP COLOR ENCAP PREFERENCE -------------------------------------------------------------
------------------------------------------------------------------------ 2 0.0.0.0/0
192.168.30.113 61626 1003 C,I,R installed 192.168.30.214 mpls ipsec - 192.168.30.113 61627 1003
C,I,R installed 192.168.30.214 biz-internet ipsec - 192.168.30.113 61628 1003 C,I,R installed
192.168.30.214 private1 ipsec - 192.168.30.113 61629 1003 C,I,R installed 192.168.30.214
private2 ipsec - 192.168.30.113 61637 1003 C,R installed 192.168.30.215 mpls ipsec -
192.168.30.113 61638 1003 C,R installed 192.168.30.215 biz-internet ipsec - 192.168.30.113 61639
1003 C,R installed 192.168.30.215 private1 ipsec - 192.168.30.113 61640 1003 C,R installed
192.168.30.215 private2 ipsec - 2 192.168.2.0/24 192.168.30.113 61610 1003 C,I,R installed
192.168.30.214 mpls ipsec - 192.168.30.113 61611 1003 C,I,R installed 192.168.30.214 biz-
internet ipsec - 192.168.30.113 61612 1003 C,I,R installed 192.168.30.214 private1 ipsec -
192.168.30.113 61613 1003 C,I,R installed 192.168.30.214 private2 ipsec - 192.168.30.113 61633
1003 C,R installed 192.168.30.215 mpls ipsec - 192.168.30.113 61634 1003 C,R installed
192.168.30.215 biz-internet ipsec - 192.168.30.113 61635 1003 C,R installed 192.168.30.215
private1 ipsec - 192.168.30.113 61636 1003 C,R installed 192.168.30.215 private2 ipsec - 2
192.168.216.0/24 0.0.0.0 68 1003 C,Red,R installed 192.168.30.216 biz-internet ipsec - 0.0.0.0
81 1003 C,Red,R installed 192.168.30.216 private1 ipsec - 0.0.0.0 82 1003 C,Red,R installed
192.168.30.216 private2 ipsec -
```
虽然仍然是分支的路由器只通过cE1(hub1)安装路由：

```
ce3#sh ip route vrf 2 0.0.0.0 Routing Table: 2 Routing entry for 0.0.0.0/0, supernet Known via
```

```
"omp", distance 251, metric 0, candidate default path, type omp Last update from 192.168.30.214
on sdwan_system_ip, 01:11:26 ago Routing Descriptor Blocks: * 192.168.30.214 (default), from
192.168.30.214, 01:11:26 ago, via sdwan_system_ip Route metric is 0, traffic share count is 1
ce3#sh ip route vrf 2 192.168.2.0 Routing Table: 2 Routing entry for 192.168.2.0/24 Known via
"omp", distance 251, metric 0, type omp Last update from 192.168.30.214 on sdwan_system_ip,
01:33:56 ago Routing Descriptor Blocks: * 192.168.30.214 (default), from 192.168.30.214,
01:33:56 ago, via sdwan_system_ip Route metric is 0, traffic share count is 1 ce3#
```

show sdwan policy service-path将确认相同，因此不提供输出以简化。

此原因也是另一个命令ecmp-limit值的**默认配置**。默认情况下，Edge路由器仅将前4条ECMP路径安装到路由表中，因此，要解决此问题，您必须重新配置分支路由器，如下例所示：

```
ce3#config-t admin connected from 127.0.0.1 using console on ce3 ce3(config)# sdwan ce3(config-
sdwan)# omp ce3(config-omp)# ecmp-limit 8 ce3(config-omp)# commit Commit complete.
```

show ip route确认通过两个集线器的两个路由现在都已安装：

```
ce3#sh ip ro vrf 2 | b Gateway Gateway of last resort is 192.168.30.215 to network 0.0.0.0 m*
0.0.0.0/0 [251/0] via 192.168.30.215, 00:00:37, sdwan_system_ip [251/0] via 192.168.30.214,
00:00:37, sdwan_system_ip m 192.168.2.0/24 [251/0] via 192.168.30.215, 00:00:37, sdwan_system_ip
[251/0] via 192.168.30.214, 00:00:37, sdwan_system_ip 192.168.216.0/24 is variably subnetted, 2
subnets, 2 masks C 192.168.216.0/24 is directly connected, Loopback2 L 192.168.216.216/32 is
directly connected, Loopback2 ce3#
```

如果使用基于功能模板的vManage设备模板，为了获得与此屏幕截图相同的结果，您需要调整OMP功能模板（路由器使用的OMP功能模板的ECMP限制和vSmart使用的OMP功能模板的每前缀通告的路径数）：

相关信息

- https://www.cisco.com/c/en/us/td/docs/routers/sdwan/command/sdwan-cr-book/config-cmd.html#wp3085259372
- https://www.cisco.com/c/en/us/td/docs/routers/sdwan/command/sdwan-cr-book/config-cmd.html#wp2570227565
- https://www.cisco.com/c/en/us/td/docs/routers/sdwan/command/sdwan-cr-book/operational-cmd.html#wp5579365410
- 技术支持和文档 - Cisco Systems