

為什麼流量無法通過SD-WAN中心路由器的ECMP路徑進行負載均衡

目錄

[簡介](#)

[必要條件](#)

[需求](#)

[採用元件](#)

[背景資訊](#)

[問題](#)

[解決方案](#)

[相關資訊](#)

簡介

本檔案將說明SD-WAN交換矩陣中的等價多重路徑(ECMP)路由的一個典型問題，即來自分支路由器的流量無法通過通告相同首碼的多個集線器路由器進行負載平衡。本文還說明如何解決此問題，以及如何使用各種疑難排解指令(包括show sdwan policy service-path)來疑難排解在17.2 Cisco IOS®-XE軟體中新增的路由問題。

必要條件

需求

思科建議您瞭解以下主題：

- 對重疊管理通訊協定(OMP)的基本瞭解
- SD-WAN元件和它們之間的互動

採用元件

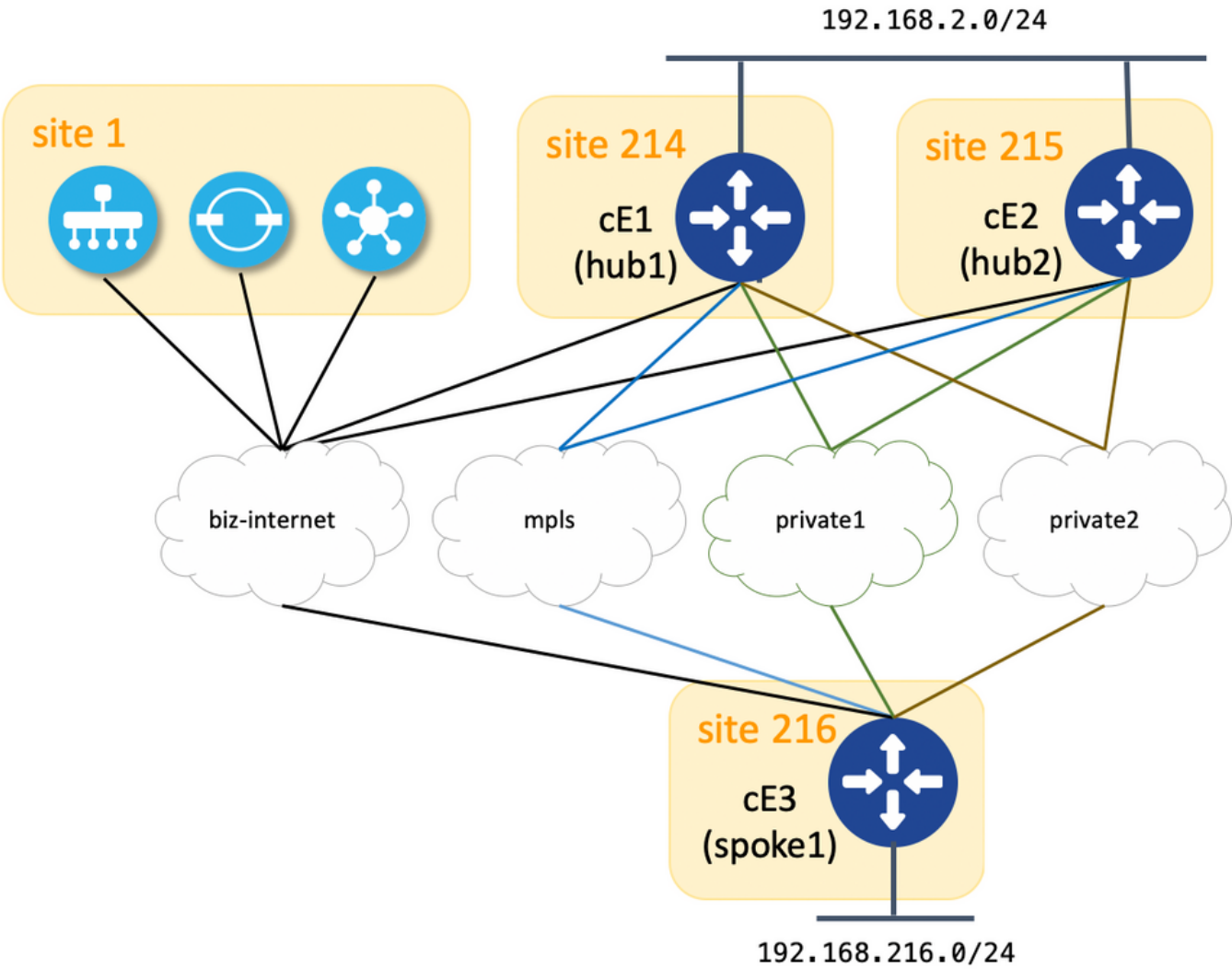
為了演示目的，使用了以下軟體路由器：

- 4台運行17.2.1v軟體版本、在控制器模式(SD-WAN)下運行的Cisco IOS-XE CSR1000v路由器
- 運行20.1.12軟體版本的vSmart控制器

本文中的資訊是根據特定實驗室環境內的裝置所建立。文中使用到的所有裝置皆從已清除（預設）的組態來啟動。如果您的網路運作中，請確保您瞭解任何指令可能造成的影響。

背景資訊

為了本文檔的目的，使用以下實驗拓撲：



您可以在此處找到為SD-WAN交換矩陣中的每台裝置分配的site-id和system-ip引數的摘要：

主機名	system-ip	site-id
cE1(hub1)	192.168.30.214	214
cE2(hub2)	192.168.30.215	215
cE3(spoke1)	192.168.30.216	216
vSmart	192.168.30.113	1

每個集線器具有4個TLOC(傳輸位置識別符號)，並根據拓撲圖分配顏色；每個集線器將0.0.0.0/0預設路由與192.168.2.0/24子網一起通告給分支（分支路由器cE3）。vSmart上未配置策略以優先使用任何路徑/裝置，並且所有裝置上的所有OMP設定也均設定為預設值。其餘配置是基本的SD-WAN重疊功能的標準最小配置，因此不提供這些配置以簡化操作。您可以期望在分支路由器的所有可用上行鏈路上進行負載均衡的主用 — 主用冗餘和流向中心路由器的輸出流量。

問題

分支機構路由器安裝預設路由並僅通過cE1路由器(hub1)路由到192.168.2.0/24子網：

```
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
```

這是因為對於192.168.2.0/24的預設路由0.0.0.0/0,cE3只收到4條路由。

```
ce3#show sdwan omp routes vpn 2 | begin PATH
```

VPN	PREFIX	FROM PEER	PATH ID	LABEL	STATUS	ATTRIBUTE TYPE	TLOC IP	COLOR
2	0.0.0.0/0	192.168.30.113	61614	1003	C,I,R	installed	192.168.30.214	mpls
		192.168.30.113	61615	1003	C,I,R	installed	192.168.30.214	biz-i
		192.168.30.113	61616	1003	C,I,R	installed	192.168.30.214	priva
		192.168.30.113	61617	1003	C,I,R	installed	192.168.30.214	priva
2	192.168.2.0/24	192.168.30.113	61610	1003	C,I,R	installed	192.168.30.214	mpls
		192.168.30.113	61611	1003	C,I,R	installed	192.168.30.214	biz-i
		192.168.30.113	61612	1003	C,I,R	installed	192.168.30.214	priva
		192.168.30.113	61613	1003	C,I,R	installed	192.168.30.214	priva
2	192.168.216.0/24	0.0.0.0	68	1003	C,Red,R	installed	192.168.30.216	biz-i
		0.0.0.0	81	1003	C,Red,R	installed	192.168.30.216	priva
		0.0.0.0	82	1003	C,Red,R	installed	192.168.30.216	priva

雖然在vSmart上，您可以看到它接收所有8條路由（每個集線器上每個TLOC顏色對應的4條路由）：

```
vsmart1# show omp routes vpn 2 | b PATH
```

VPN	PREFIX	FROM PEER	PATH ID	LABEL	STATUS	ATTRIBUTE TYPE	TLOC IP	COLOR
2	0.0.0.0/0	192.168.30.214	66	1003	C,R	installed	192.168.30.214	mpls
		192.168.30.214	68	1003	C,R	installed	192.168.30.214	biz-i
		192.168.30.214	81	1003	C,R	installed	192.168.30.214	priva
		192.168.30.214	82	1003	C,R	installed	192.168.30.214	priva
		192.168.30.215	66	1003	C,R	installed	192.168.30.215	mpls
		192.168.30.215	68	1003	C,R	installed	192.168.30.215	biz-i
		192.168.30.215	81	1003	C,R	installed	192.168.30.215	priva
		192.168.30.215	82	1003	C,R	installed	192.168.30.215	priva
2	192.168.2.0/24	192.168.30.214	66	1003	C,R	installed	192.168.30.214	mpls
		192.168.30.214	68	1003	C,R	installed	192.168.30.214	biz-i
		192.168.30.214	81	1003	C,R	installed	192.168.30.214	priva
		192.168.30.214	82	1003	C,R	installed	192.168.30.214	priva
		192.168.30.215	66	1003	C,R	installed	192.168.30.215	mpls
		192.168.30.215	68	1003	C,R	installed	192.168.30.215	biz-i

192.168.30.215	81	1003	C,R	installed	192.168.30.215	private
192.168.30.215	82	1003	C,R	installed	192.168.30.215	private

如果來自cE1(hub1)的預設路由丟失，則分支路由器安裝來自cE2(hub2)的路由。 因此，當cE1充當主路由器時，沒有主用 — 主用冗餘，而是主用 — 備用。

您還可以使用show sdwan policy service-path 命令檢查針對特定流量採用的輸出路徑，如以下範例所示：

```
ce3#show sdwan policy service-path vpn 2 interface Loopback2 source-ip 192.168.216.216 dest-ip 192.168.216.216
Next Hop: IPsec
Source: 192.168.109.216 12347 Destination: 192.168.110.214 12427 Local Color: biz-internet Remote Color: biz-internet
```

若要檢視特定流量型別的所有可用路徑，請使用all 關鍵字：

```
ce3#show sdwan policy service-path vpn 2 interface Loopback2 source-ip 192.168.216.216 dest-ip 192.168.216.216
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: biz-internet
Next Hop: IPsec
Source: 192.168.108.216 12367 Destination: 192.168.108.214 12407 Local Color: private2 Remote Color: private2
Next Hop: IPsec
Source: 192.168.107.216 12367 Destination: 192.168.107.214 12407 Local Color: private1 Remote Color: private1
Next Hop: IPsec
Source: 192.168.109.216 12347 Destination: 192.168.109.214 12387 Local Color: biz-internet Remote Color: biz-internet
```

因此，這也確認路由器cE3(spoke2)只有4條可用路徑，而不是8條可用路徑。

如果您檢查vSmart通告的具體內容，您只能看到向cE3通告的4條路由：

```
vsmart1# show omp routes vpn 2 0.0.0.0/0 detail | nomore | exclude not\ set | b ADVERTISED\ TO: | b peer
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

```

根據此輸出，您可以斷定問題是由vSmart控制器引起的。

解決方案

此行為是由在vSmart控制器上預設配置send-path-limit引起的。send-path-limit定義從邊緣路由器通告到vSmart控制器以及從vSmart控制器通告到其他邊緣路由器的ECMP路由的最大數量。預設值為4，通常情況下，它對於邊緣路由器（如本拓撲中每個集線器路由器上有4個上行鏈路）足夠，但對於虛擬智慧控制器而言，它不足以將所有可用路徑傳送到其他邊緣路由器。可以為send-path-limit設定的最大值為16，但在某些極端情況下，這仍然不夠，儘管已開啟增強請求[CSCvs89015](#)，以將最大值增加至128。

為了解決此問題，您必須重新設定vSmart設定，如以下範例所示：

```

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#

```

然後，vSmart會向分支機構路由器通告所有8條路由，並接收這些路由：

```
ce3#show sdwan omp routes vpn 2 | begin PATH
```

VPN	PREFIX	FROM PEER	PATH ID	LABEL	STATUS	ATTRIBUTE TYPE	TLOC IP	COLOR
2	0.0.0.0/0	192.168.30.113	61626	1003	C,I,R	installed	192.168.30.214	mpls
		192.168.30.113	61627	1003	C,I,R	installed	192.168.30.214	biz-i
		192.168.30.113	61628	1003	C,I,R	installed	192.168.30.214	priva
		192.168.30.113	61629	1003	C,I,R	installed	192.168.30.214	priva
		192.168.30.113	61637	1003	C,R	installed	192.168.30.215	mpls
		192.168.30.113	61638	1003	C,R	installed	192.168.30.215	biz-i
		192.168.30.113	61639	1003	C,R	installed	192.168.30.215	priva
		192.168.30.113	61640	1003	C,R	installed	192.168.30.215	priva
2	192.168.2.0/24	192.168.30.113	61610	1003	C,I,R	installed	192.168.30.214	mpls
		192.168.30.113	61611	1003	C,I,R	installed	192.168.30.214	biz-i
		192.168.30.113	61612	1003	C,I,R	installed	192.168.30.214	priva
		192.168.30.113	61613	1003	C,I,R	installed	192.168.30.214	priva
		192.168.30.113	61633	1003	C,R	installed	192.168.30.215	mpls
		192.168.30.113	61634	1003	C,R	installed	192.168.30.215	biz-i
		192.168.30.113	61635	1003	C,R	installed	192.168.30.215	priva
		192.168.30.113	61636	1003	C,R	installed	192.168.30.215	priva
2	192.168.216.0/24	0.0.0.0	68	1003	C,Red,R	installed	192.168.30.216	biz-i
		0.0.0.0	81	1003	C,Red,R	installed	192.168.30.216	priva
		0.0.0.0	82	1003	C,Red,R	installed	192.168.30.216	priva

雖然仍舊的分支路由器僅通過cE1(hub1)安裝路由：

```
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將確認相同，因此不提供輸出以簡化操作。

此原因也是其他命令ecmp-limit值的預設設定。預設情況下，邊緣路由器僅將前4個ECMP路徑安裝到路由表中，因此要解決此問題，必須重新配置分支路由器，如以下示例所示：

```
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 確認現在已安裝通過兩個集線器的兩個路由：

```
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#
```

如果您使用基於功能模板的vManage裝置模板，則為了獲得與本螢幕截圖相同的結果，您需要調整OMP功能模板（路由器使用的OMP功能模板的ECMP限制，以及vSmart使用的OMP功能模板的「每個字首通告的路徑數」）：

Device **Feature**

Feature Template > OMP

Basic Configuration

Timers

Advertise

BASIC CONFIGURATION

Graceful Restart for OMP

☒ ▾

☒ On

☐ Off

Overlay AS Number

☒ ▾

Graceful Restart Timer (seconds)

☒ ▾

43200

Number of Paths Advertised per Prefix

☐ ▾

8

ECMP Limit

☐ ▾

8

Shutdown

☒ ▾

☐ Yes

☒ No

相關資訊

- <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>
- [技術支援與文件 - Cisco Systems](#)

關於此翻譯

思科已使用電腦和人工技術翻譯本文件，讓全世界的使用者能夠以自己的語言理解支援內容。請注意，即使是最佳機器翻譯，也不如專業譯者翻譯的內容準確。Cisco Systems, Inc. 對這些翻譯的準確度概不負責，並建議一律查看原始英文文件（提供連結）。