

Why Traffic is not Load Balanced over ECMP Paths From SD-WAN Hub Routers

Contents

[Introduction](#)

[Prerequisites](#)

[Requirements](#)

[Components Used](#)

[Background Information](#)

[Problem](#)

[Solution](#)

[Related Information](#)

Introduction

This document describes a typical issue with Equal-Cost Multipath (ECMP) routing in SD-WAN fabric when traffic from a spoke router is not load balanced over multiple hub routers that advertise the same prefix. It also explains how to solve this problem and how to use various troubleshooting commands including **show sdwan policy service-path** for troubleshooting of routing issues which was added in 17.2 Cisco IOS®-XE software.

Prerequisites

Requirements

Cisco recommends that you have knowledge of these topics:

- Basic understanding of Overlay Management Protocol (OMP)
- SD-WAN components and interaction between them

Components Used

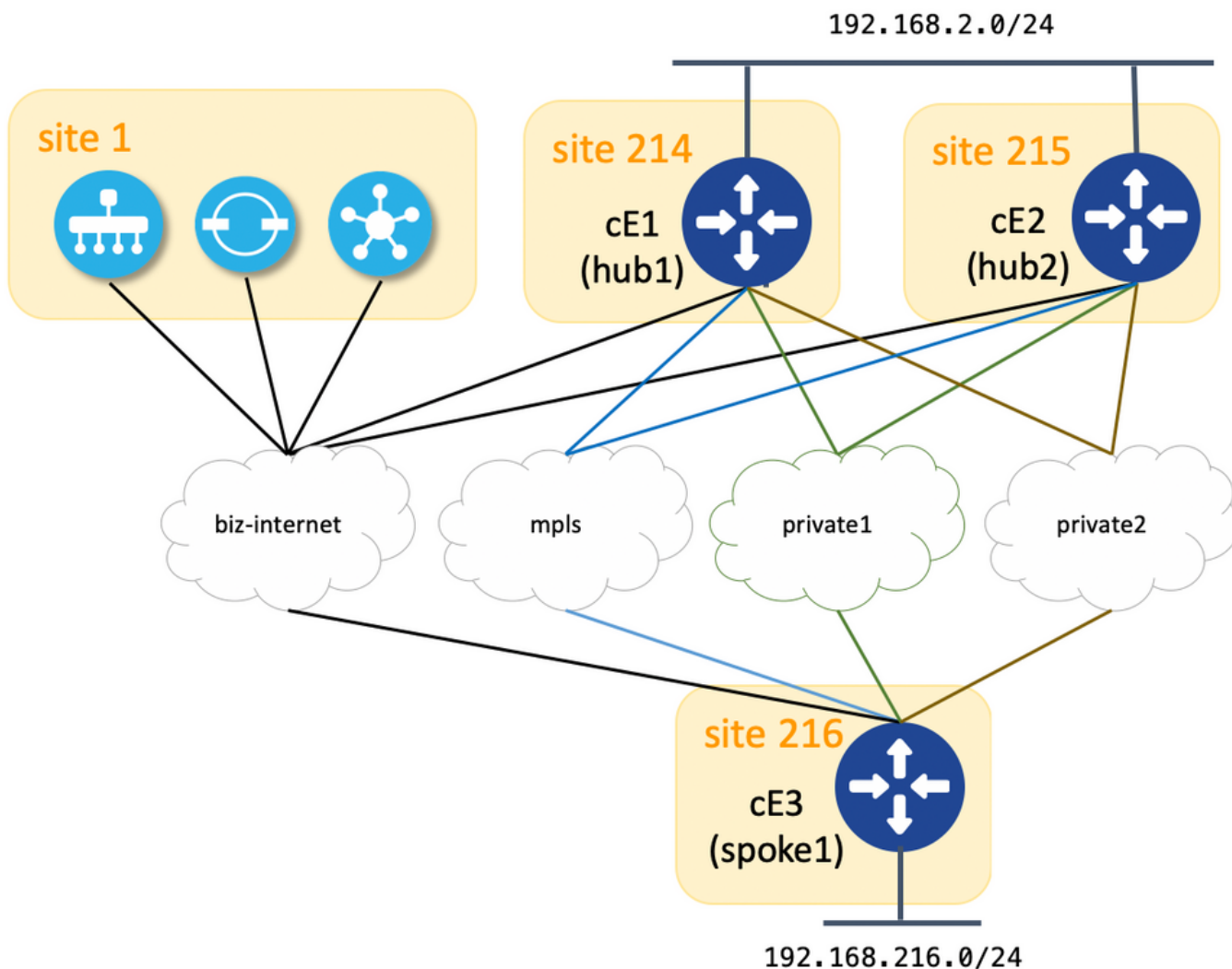
For the purpose of the demonstration, these software routers were used:

- 4 Cisco IOS-XE CSR1000v routers running 17.2.1v software version running in controller mode (SD-WAN)
- vSmart controller running 20.1.12 software version

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. If your network is live, ensure that you understand the potential impact of any command.

Background Information

For the purpose of this document, this lab topology is used:



Here you can find a summary of assigned site-id and system-ip parameters to each device in the SD-WAN fabric:

hostname	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

Each hub has 4 TLOCs (Transport location identifier) with colors assigned as per topology diagram and each hub advertises default route 0.0.0.0/0 to spoke (branch router cE3) together with 192.168.2.0/24 subnet. There is no policy configured on vSmart in order to prefer any path/device and all OMP settings are also set to default on all devices. Rest of the configuration is the standard minimal configuration for basic SD-WAN overlay functionality and hence not provided for the sake of brevity. You can expect active-active redundancy and egress traffic toward hub routers load-balanced across all available uplinks from the branch router.

Problem

Branch routers install default route and route to 192.168.2.0/24 subnet only via cE1 router (hub1):

```
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
```

This is because cE3 receives only 4 routes for default route 0.0.0.0/0 as well as for 192.168.2.0/24.

```
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 -
```

Although on vSmart, you can see that it receives all 8 routes (4 routes for each TLOC color on each hub):

```
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 -
```

If default route from cE1 (hub1) is lost, spoke routers install route from cE2 (hub2). Therefore there is no active-active redundancy and rather active-standby with cE1 acting as a primary router.

You can also check which egress path is taken for specific traffic flow with help of **show sdwan policy service-path** command as in the example here:

```
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
```

In order to see all available paths for specific traffic type, use **all** keyword:

```
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

So, this also confirms only 4 paths are available instead of 8 for router cE3 (spoke2).

If you check what exactly vSmart advertises, you see only 4 routes advertised toward cE3:

```
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
```

Based on this output, you can conclude that the issue is caused by the vSmart controller.

Solution

This behavior is caused by the default configuration of **send-path-limit** on vSmart controller. **send-path-limit** defines the maximum number of ECMP routes advertised from Edge router to vSmart controller and from vSmart controller to other Edge routers. The default value is 4 and usually, it is enough for Edge router (like in this topology with 4 uplinks on each hub router), but not enough for vSmart controller to send all available path to the other Edge routers. The maximum value that can be set for **send-path-limit** is 16 but in some extreme cases, this still can be not enough, though there is enhancement request [CSCvs89015](#) opened to increase maximum value to 128.

In order to solve this issue, you must reconfigure vSmart settings as in the example here:

```
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#
```

And then, all 8 routes are advertised by vSmart to branch routers and received by them:

```
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 -
```

Although the still branch routers install routes only via cE1 (hub1):

```
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 will confirm the same and hence the output is not provided for brevity.

The reason for this is also default configuration of another command **ecmp-limit** value. By default, Edge router installs only first 4 ECMP paths into the routing table, hence to fix this issue, you must reconfigure spoke routers as in the example here:

```
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 confirms both routes via both hubs are installed now:

```
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#
```

If you use vManage device templates based on feature templates, in order to achieve the same result you need to adjust your OMP feature template as on this screenshot (ECMP limit for OMP feature template used by routers and Number of Paths Advertised per Prefix for OMP feature template used by vSmart):

Basic Configuration

Timers

Advertise

BASIC CONFIGURATION

Graceful Restart for OMP

On Off

Overlay AS Number

Graceful Restart Timer (seconds)

Number of Paths Advertised per Prefix

ECMP Limit

Shutdown

Yes No

Related Information

- <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>
- [Technical Support & Documentation - Cisco Systems](#)