# Troubleshoot OMP Best Path Selection Peculiarities and Typical Confusions

#### **Contents**

**Introduction** 

**Prerequisites** 

Requirements

Components Used

**Background Information** 

Best Path Selection, Egress Policy and Send-Path Limit Order of Operations

vSmart Best Path Selection: Edge Router Sourced Routes vs Routes Received via Other vSmarts Case 1

vSmart Best Path Selection: Edge Router Sourced Routes vs Routes Received via Other vSmarts Case 2

**Related Information** 

#### Introduction

This document describes how to troubleshoot the OMP best-path selection and operation between the OMP, egress policy, and send-path-limit feature.

### **Prerequisites**

#### Requirements

Cisco recommends that you have knowledge of the Cisco Software Defined Wide Area Network (SDWAN) solution.

#### **Components Used**

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 demonstration, the lab was set up with three vSmart controllers and three Cisco IOS® XE routers with sites IDs 243, 244, and 245 advertising the same 172.16.1.0/24 prefix. There are also some other routers connected to the overlay (for example, with site-id 204). The last octet of any router system-ip is equal to site-id in this example (10.10.10.<site-id>). vSmarts are having system-ip 10.10.10.228, .229, and .230. In this example, each router has two transports (WAN interfaces) available hence two Transport Locators (TLOCs) with colors private1 and biz-internet. On private1, the circuit router has an IP address assigned in the form of 192.168.9.x and on biz-internet it has 192.168.10.x where x is a site-id.

Scenarios were tested with vSmarts running software versions 20.4.1 and 20.6.1.

# Best Path Selection, Egress Policy and Send-Path Limit Order of Operations

First of all, demonstrate the best path selection, egress policy, and send-path-limit order of operations. Routers with site-id 247 must receive prefix from routers with site-id 244 or 245, but not from 243.

Here is the policy to achieve this for reference:

```
policy
lists
 site-list site_247
  site-id 247
 site-list sites_244_245
  site-id 244
  site-id 245
 prefix-list ENK_PL
  ip-prefix 172.16.1.0/24
control-policy send_2_247
 sequence 10
  match route
   prefix-list ENK_PL
   site-list sites_244_245
  action accept
  !
  1
 sequence 20
  match route
    prefix-list ENK_PL
  action reject
 default-action accept
Ţ
!
apply-policy
site-list site_247
 control-policy send_2_247 out
1
Ţ
```

When you take a look a vSmart2, it has connectivity to two other vSmarts (site-id 1) and edge routers with site-id 243, 244, and 247. Site 245 is connected to some other vSmart controller and vSmart2 receives its prefix from them indirectly via other vSmart(s).

```
vsmart2# show omp peers
R -> routes received
I -> routes installed
S -> routes sent
```

PEER	TYPE	DOMAIN ID	OVERLAY ID	SITE ID	STATE	UPTIME	R/I/S
10.10.10.204	vedge	1	1	204	up	2:20:18:10	14/0/7
10.10.10.228	vsmart	1	1	1	up	2:20:18:06	247/0/9
10.10.10.230	vsmart	1	1	1	up	2:20:17:07	256/0/15
10.10.10.243	vedge	1	1	243	up	2:20:18:10	8/0/7
10.10.10.244	vedge	1	1	244	up	0:13:24:59	10/0/6
10.10.10.247	vedge	1	1	247	up	2:20:18:10	0/0/8

In the OMP table, you can notice that route is being received from two other vSmart controllers and also directly from sites 243 and 244:

```
vsmart2# show omp routes 172.16.1.0/24
Code:
   -> chosen
   -> installed
Red -> redistributed
Rej -> rejected
  -> looped
   -> resolved
   -> stale
Ext -> extranet
Inv -> invalid
Stg -> staged
IA -> On-demand inactive
  -> TLOC unresolved
```

VPN	PREFIX	FROM PEER	PATH ID	LABEL	STATUS	ATTRIBUTE TYPE	TLOC IP	COLOR
1	172.16.1.0/24	10.10.10.228 10.10.10.230 10.10.10.243 10.10.10.243 10.10.10.244	409 7187 69 81 68	1001 1002 1001 1001 1002	C,R C,R C,R C,R C,R	installed installed installed	10.10.10.243 10.10.10.244 10.10.10.243 10.10.10.243 10.10.10.244	publi biz-i publi priva biz-i
		10.10.10.244	81	1002	C,R	installed	10.10.10.244	priva

send-path-limit - in this demonstration is set to 1:

```
vsmart2# show running-config omp
omp
no shutdown
send-path-limit
no graceful-restart
```



Note: From all equal-cost multi-paths for given prefix selected as best-paths and accepted by outbound (egress) policy, not more than the number of paths specified in send-path-limit advertised. You can check which prefix is advertised to which peer. The route originated by site 243 has the lowest originator system-ip in the OMP route list. Because send-path-limit is set to 1, out of two available paths via TLOC private1 and biz-internet, the only route advertised to the routers with site-id 204 and 244 as well as to two other vSmart controllers (10.10.10.228, .230) is from the biz-internet TLOC because it has a highest private IP address (address assigned to the interface):

vsmart2# show omp tlocs ip 10.10.10.243 received | b PUBLIC

ADDRES	S					PSEUD	00	PUBLIC	
FAMILY	TLOC IP	COLOR	ENCAP	FROM PEER	STATUS	KEY	PUBLIC IP	PORT	PRIVATE
ipv4	10.10.10.243	B biz-internet	ipsec	10.10.10.228	C,R	1	192.168.10.243	12346	192.168.
				10.10.10.230	C,R	1	192.168.10.243	12346	192.168.
				10.10.10.243	C,I,R	1	192.168.10.243	12346	192.168.
	10.10.10.243	3 private1	ipsec	10.10.10.228	C,R	1	192.168.9.243	12346	192.168.
				10.10.10.230	C,R	1	192.168.9.243	12346	192.168.
				10.10.10.243	C,I,R	1	192.168.9.243	12346	192.168.

Site id 243 gets the next route from the list (from site 244) and it is also via biz-internet color because it has the highest TLOC private IP address. Site 243 does not get its own route because of the split-horizon rule though it has the lowest system-IP. Site 247 gets the route from site 244 as well because of the egress policy.

```
vsmart2# show omp routes 172.16.1.0/24 detail | nomore | exclude not\ set | b ADVERTISED | include peer
      10.10.10.204
                     10.10.10.243
    originator
                     10.10.10.243, biz-internet, ipsec
    tloc
peer
      10.10.10.228
                     10.10.10.243
    originator
                     10.10.10.243, biz-internet, ipsec
    tloc
      10.10.10.230
    originator
                     10.10.10.243
                     10.10.10.243, biz-internet, ipsec
      10.10.10.243
peer
    originator
                     10.10.10.244
    tloc
                     10.10.10.244, biz-internet, ipsec
       10.10.10.244
peer
    originator
                     10.10.10.243
                     10.10.10.243, biz-internet, ipsec
      10.10.10.247
peer
    originator
                     10.10.10.244
                     10.10.10.244, biz-internet, ipsec
    tloc
```

# vSmart Best Path Selection: Edge Router Sourced Routes vs Routes Received via Other vSmarts Case 1

In order to continue this demonstration, increase send-path-limit and set it to 16, enable debug omp policy prefix 172.16.1.0/24 level high, and observe the results. Now, vSmart2 also receives a route from site-id 245 via vSmart1 with system-ip 10.10.10.228 and vSmart3 with 10.10.10.230).

```
vsmart2# show omp routes 172.16.1.0/24
Code:
C  -> chosen
I  -> installed
Red -> redistributed
Rej -> rejected
L  -> looped
```

R -> resolved S -> stale Ext -> extranet Inv -> invalid

Stg -> staged

IA -> On-demand inactive
U -> TLOC unresolved

VPN	PREFIX	FROM PEER	PATH ID	LABEL	STATUS	ATTRIBUTE TYPE	TLOC IP	COLOR
1	172.16.1.0/24	10.10.10.228	10146	1001	C,R	installed	10.10.10.243	publi
		10.10.10.228	10448	1001	C,R	installed	10.10.10.243	priva
		10.10.10.228	10449	1002	C,R	installed	10.10.10.245	biz-i
		10.10.10.228	10450	1002	C,R	installed	10.10.10.245	priva
		10.10.10.230	10252	1002	C,R	installed	10.10.10.244	biz-i
		10.10.10.230	10577	1002	C,R	installed	10.10.10.244	priva
		10.10.10.230	10578	1002	C,R	installed	10.10.10.245	biz-i
		10.10.10.230	10579	1002	C,R	installed	10.10.10.245	priva
		10.10.10.243	69	1001	C,R	installed	10.10.10.243	publi
		10.10.10.243	81	1001	C,R	installed	10.10.10.243	priva
		10.10.10.244	68	1002	C,R	installed	10.10.10.244	biz-i
		10.10.10.244	81	1002	C,R	installed	10.10.10.244	priva

But vSmart2 advertises only routes from site 244, and not from 245, to site 247 now. This is a typical source of confusion because routes received directly from edge routers are preferred over routes received via vSmarts and not advertised to an Edge router and not sent to an Edge router, but only in case if vSmart found an OMP routing table entry for the same prefix from any other vSmart that Edge router is already connected to:

```
vsmart2# show omp routes 172.16.1.0/24 detail | nomore | exclude not\ set | b ADVERTISED | include peer
peer
        10.10.10.204
     originator
                      10.10.10.244
     originator
                      10.10.10.244
     originator
                      10.10.10.243
     originator
                      10.10.10.243
        10.10.10.228
peer
     originator
                      10.10.10.244
                      10.10.10.244
     originator
                      10.10.10.243
     originator
     originator
                      10.10.10.243
        10.10.10.230
     originator
                      10.10.10.244
                      10.10.10.244
     originator
                      10.10.10.243
     originator
     originator
                      10.10.10.243
peer
       10.10.10.243
     originator
                      10.10.10.244
     originator
                      10.10.10.244
       10.10.10.244
peer
```

10.10.10.243

originator

```
originator 10.10.10.243
peer 10.10.10.247
originator 10.10.10.244
originator 10.10.10.244
```

This is also confirmed from debug logs stored in /var/log/tmplog/vdebug where the reason for suppression is seen as vSmart Connectivity.

```
Oct 9 14:29:01 vsmart2 OMPD[1120]: omp_rib_out_process_entry[3792]: Peer: 10.10.10.247 NLRI: 1: 172.16
Oct 9 14:29:01 vsmart2 OMPD[1120]: omp_rib_out_process_entry[3792]: Peer: 10.10.10.247 NLRI: 1: 172.16
Oct 9 14:29:01 vsmart2 OMPD[1120]: omp_rib_out_process_entry[3792]: Peer: 10.10.10.247 NLRI: 1: 172.16
Oct 9 14:29:01 vsmart2 OMPD[1120]: omp_rib_out_process_entry[3792]: Peer: 10.10.10.247 NLRI: 1: 172.16
Oct 9 14:29:01 vsmart2 OMPD[1120]: omp_rib_out_process_entry[3792]: Peer: 10.10.10.247 NLRI: 1: 172.16
Oct 9 14:29:01 vsmart2 OMPD[1120]: omp_rib_out_process_entry[3792]: Peer: 10.10.10.247 NLRI: 1: 172.16
Oct 9 14:29:01 vsmart2 OMPD[1120]: omp_rib_out_process_entry[3792]: Peer: 10.10.10.247 NLRI: 1: 172.16
Oct 9 14:29:01 vsmart2 OMPD[1120]: omp_rib_out_process_entry[3792]: Peer: 10.10.10.247 NLRI: 1: 172.16
```

At the same time, keep in mind that site 247 receives both routes finally anyway, because by default it is connected to two vSmart controllers (max-control-connections 2) and vSmart3 advertises both routes to it because originators are directly connected to it:

Site-247#show sdwan omp routes 172.16.1.0/24 | begin PATH

			PAIN			ALIKIDULE		
VPN	PREFIX	FROM PEER	ID	LABEL	STATUS	TYPE	TLOC IP	COLOR
1	172.16.1.0/24	10.10.10.229	13	1002	C,I,R	installed	10.10.10.244	biz-i
		10.10.10.229	14	1002	C,I,R	installed	10.10.10.244	priva
		10.10.10.230	13	1002	C,R	installed	10.10.10.244	biz-i
		10.10.10.230	14	1002	C,R	installed	10.10.10.244	priva
		10.10.10.230	61	1002	C,I,R	installed	10.10.10.245	biz-i
		10.10.10.230	62	1002	C,I,R	installed	10.10.10.245	priva

ATTOTOLITE

A summarization of the best path selection and order of operations in included in this table.

```
    Prefer non-stale route over stale route
    Route resolvability

Next-hop TLOC is Reachable (data plane BFD session is here)
```

3. Prefer highest route preference
 4. Prefer highest TLOC preference
 5. Prefer best origin code (Connected, Static, eBGP, EIGRP Internal, OSPF Intra, OSPF Inter, OSPF External, EIGRP External iBGP, Unknown/Unset)
 6. Route Source Preference. On vSmart: prefer Edge router-sourced route over vSmart-sourced route
 7. Prefer OMP route with lowest origin metric
 8. Prefer route received from lowest System-IP
 9. Prefer route from highest Private TLOC IP-address originated from the same site-id
 10. outbound control policy
 11. send-path-limit

## vSmart Best Path Selection: Edge Router Sourced Routes vs Routes Received via Other vSmarts Case 2

This behavior can be seen in double-failure scenarios with controllers affinity configuration and outbound (egress) policy configuration that discriminates some routes from some sources over others based on some criteria as we do with policy in the previous scenarios. For the purpose of demonstration in this section, you need to increase route scale compared to the previous scenarios, so more sites with different site-ids are used. Consider typical deployment with three vSmart controllers and three geographical regions just like in the demonstration in the previous section. With help of affinity, each vSmart is assigned to the corresponding group 1, 2, or 3. **max-control-connections** is set to the default value of 2. vSmarts 1 and 2 are preferred for routers from region A. In region B vSmart 2 and 3 are preferred. For a region, C vSmart 3 and 1 are preferred.

Here is an example of configuration to assign vSmart controller to Group 1:

```
system controller-group-id 1
```

And also, an example of a configuration for the router from region A that prefers controllers from Groups 1 and 2. Controllers from Group 3 are used as a last resort to connect if none of the controllers from Groups 1 and 2 are available because max-control-connections is set to 2 by default:

```
system
controller-group-list 1 2 3
```

The same result can be achieved with the other configuration:

```
vpn 0
  interface ge0/0
  tunnel-interface
   exclude-controller-group-list 3
 !
!
!
```

**max-control-connections** is also set to a default value of 2 in this demonstration. send-path-limit set to value 16 on all routers and controllers.

Each region has 2 routers now originating prefix 10.0.0.0/8. Each of those routers has five transports (WAN interfaces) with TLOC colors from private1 to private5. cEdges originating this prefix are assigned to the regions as shown in this table. It also describes new system-ip addressing.

hostname / system-ip		vSmart1	vSmart2	vSmart3
		169.254.206.4	169.254.206.5	169.254.206.6
cEdge1	169.254.206.11	Region A	Region A	
cEdge2	169.254.206.12	Region A	Region A	
cEdge3	169.254.206.13		Region B	Region B
cEdge4	169.254.206.14		Region B	Region B
cEdge5	169.254.206.15	Region C		Region C
cEdge6	169.254.206.16	Region C		Region C

Such configuration and scale mean that each vSmart controller receives 20 paths from directly connected routers (4 router x 5 TLOCs) and in addition also 20 paths from each vSmart. In total, it gives 60 paths for the given prefix 10.0.0.0/8 in the OMP table of each vSmart controller in normal conditions. Some unimportant columns were removed from show omp route 10.0.0.0/8 vSmart1 output for brevity.

FROM PEER	STATUS	TLOC IP	COLOR	PREFERENCE
169.254.206.5	C.R	169.254.206.11	private1	-
169.254.206.5	C,R	169.254.206.11	private2	
169.254.206.5	C,R	169.254.206.11	private3	
169.254.206.5	C,R	169.254.206.11	private4	
169.254.206.5	C,R	169.254.206.11	private5	
169.254.206.5	C,R	169.254.206.12	private1	
169.254.206.5	C,R	169.254.206.12	private2	
169.254.206.5	C,R	169.254.206.12	private3	
169.254.206.5	C,R	169.254.206.12	private4	
169.254.206.5	C,R	169.254.206.12	private5	
169.254.206.5	C,R	169.254.206.13	private1	
169.254.206.5	C,R	169.254.206.13	private2	
169.254.206.5	C,R	169.254.206.13	private3	
169.254.206.5	C,R	169.254.206.13	private4	
169.254.206.5	C,R	169.254.206.13	private5	
169.254.206.5		169.254.206.14	private1	
169.254.206.5		169.254.206.14	private2	
169.254.206.5		169.254.206.14	private3	
169.254.206.5		169.254.206.14	private4	
169.254.206.5		169.254.206.14	private5	
169.254.206.6	C.R	169.254.206.13	private1	
169.254.206.6	Ć,R	169.254.206.13	private2	
169.254.206.6	Ć,R	169.254.206.13	private3	
169.254.206.6	C,R	169.254.206.13	private4	
169.254.206.6	C,R	169.254.206.13		
169.254.206.6	Ć,R	169.254.206.14	-	
169.254.206.6	Ć,R	169.254.206.14	•	
169.254.206.6	Ć,R	169.254.206.14	private3	
169.254.206.6	Ć,R	169.254.206.14	private4	
169.254.206.6	C,R	169.254.206.14		
169.254.206.6	C.R	169.254.206.15	private1	
169.254.206.6	C.R	169.254.206.15	private2	
169.254.206.6	C.R	169.254.206.15	private3	
169.254.206.6	C,R	169.254.206.15	private4	_
169.254.206.6	C,R	169.254.206.15	•	_
169.254.206.6	C,R	169.254.206.16	private1	
169.254.206.6	C,R	169.254.206.16	private2	_
169.254.206.6	C,R	169.254.206.16	private3	_
169.254.206.6	C,R	169.254.206.16	private4	_
169.254.206.6	C,R	169.254.206.16	private5	_
169.254.206.11		169.254.206.11	private1	_
169.254.206.11	C,R	169.254.206.11	private2	_
169.254.206.11	C,R	169.254.206.11	private3	_
169.254.206.11	C,R	169.254.206.11	private4	_
169.254.206.11	C,R	169.254.206.11	private5	_
169.254.206.12	C,R	169.254.206.12	private1	_
169.254.206.12		169.254.206.12	private2	_
169.254.206.12	C,R	169.254.206.12	private3	_
169.254.206.12	C,R	169.254.206.12	private4	_
169.254.206.12	C,R	169.254.206.12	private5	_
169.254.206.15	C,R	169.254.206.15	private1	_
169.254.206.15	C,R	169.254.206.15	private2	_
169.254.206.15	C,R	169.254.206.15	private3	_
169.254.206.15	C,R	169.254.206.15	private3	_
169.254.206.15	C,R	169.254.206.15	private4 private5	_
169.254.206.16	C,R	169.254.206.16	private1	_
169.254.206.16	C,R	169.254.206.16	private2	_
169.254.206.16	C,R	169.254.206.16	private2 private3	_
169.254.206.16		169.254.206.16	private3	_
169.254.206.16	C,R	169.254.206.16	private5	_
	-,		p. 114cc3	

Now let us discuss a failure scenario. Some spoke routers with site-id 20 that belong to Region A cannot connect to both controllers, for whatever reason, and are connected to only one controller vSmart3 which is the last resort vSmart for this region.

Site-20# show omp peers
R -> routes received
I -> routes installed
S -> routes sent

PEER	TYPE	DOMAIN ID	OVERLAY ID	SITE ID	STATE	UPTIME	R/I/S
169.254.206.6	vsmart	1	1	1	 up	0:00:26:31	10/4/0

If no control-policy is configured, this can lead to suboptimal routing for Site-20 from Region A because as per the best-path selection algorithm, vSmart3 advertises routes received from Edge routers first. They are more preferred than routes native to region A received via vSmart controllers vSmart1 and vSmart2:

vsmart3# show omp routes 10.0.0.0/8 advertised detail | nomore | b ADVERTISED | i originator\|peer\|\ t peer 192.168.206.20

```
originator
                 169.254.206.14
                 169.254.206.14, private2, ipsec
tloc
originator
                 169.254.206.14
                 169.254.206.14, private1, ipsec
tloc
                 169.254.206.14
originator
                 169.254.206.14, private3, ipsec
tloc
originator
                 169.254.206.14
tloc
                 169.254.206.14, private4, ipsec
                 169.254.206.14
originator
tloc
                 169.254.206.14, private5, ipsec
                 169.254.206.15
originator
                 169.254.206.15, private5, ipsec
tloc
                 169.254.206.15
originator
                 169.254.206.15, private2, ipsec
tloc
originator
                 169.254.206.15
                 169.254.206.15, private1, ipsec
tloc
originator
                 169.254.206.15
                 169.254.206.15, private3, ipsec
tloc
                 169.254.206.15
originator
                 169.254.206.15, private4, ipsec
tloc
originator
                 169.254.206.13
tloc
                 169.254.206.13, private5, ipsec
originator
                 169.254.206.13
                 169.254.206.13, private4, ipsec
tloc
originator
                 169.254.206.13
tloc
                 169.254.206.13, private3, ipsec
originator
                 169.254.206.13
                 169.254.206.13, private1, ipsec
tloc
originator
                 169.254.206.13
tloc
                 169.254.206.13, private2, ipsec
originator
                 169.254.206.16
tloc
                 169.254.206.16, private1, ipsec
```

In order to avoid suboptimal routing, vSmart must allow spokes to receive routes from the routers in the same region only. Here is an example of a control policy to achieve this result:

```
policy
lists
 site-list hubs_A
  site-id 11
  site-id 12
 site-list hubs_B
  site-id 13
  site-id 14
 site-list hubs_C
  site-id 15
  site-id 16
 site-list spokes_A
  site-id 20
 site-list spokes_B
  site-id 21
 site-list spokes_C
  site-id 10
 control-policy region_A
 sequence 10
  match route
   site-list hubs_A
  action accept
  ļ
 sequence 20
  match route
  action reject
 default-action accept
 control-policy region_B
 sequence 10
  match route
   site-list hubs_B
  action accept
  !
  sequence 20
  match route
  action reject
 default-action accept
 control-policy region_C
```

```
sequence 10
   match route
    site-list hubs_C
   action accept
  sequence 20
   match route
   action reject
  default-action accept
apply-policy
 site-list spokes_A
  control-policy region_A out
 site-list spokes_B
  control-policy region_B out
 site-list spokes_C
  control-policy region_C out
ļ
```

originator

originator

originator

tloc

169.254.206.11

169.254.206.12

169.254.206.12

169.254.206.11, private5, ipsec

169.254.206.12, private1, ipsec

But from the previous scenario, you know that Edge-sourced routes are preferred over routes received via vSmart controllers. Does it mean that Site-20 in current conditions cannot receive any routes?

Here is yet another important concept that is being missed frequently. Routes from cEdge1 and cEdge2 (system-ip 169.254.206.11 and 169.254.206.12) are though kept in vSmart3 OMP table even if they are less preferred and still marked as C ("chosen"). All steps in the best-path selection algorithm starting from step 8 (including) considered tie-breakers and routes are not removed from the OMP table, but sorted according to the described preference for the purpose of consequent processing by egress control policies and send-path-limit limitation.

Because vSmart3 cannot find an OMP routing table entry for the refix 10.0.0.0/8 from other vSmart that Edge router is already connected to (Site-20 connected to vSmart3 only), it advertises routes from site 11 and site 12 (cEdge1 and cEdge2 correspondingly) to the site 20 router:

```
vsmart3# show omp routes 10.0.0.0/8 advertised detail | nomore | b ADVERTISED | i originator\|peer\|\ t
peer
       192.168.206.20
    originator
                     169.254.206.11
                     169.254.206.11, private1, ipsec
    tloc
    originator
                     169.254.206.11
    tloc
                     169.254.206.11, private2, ipsec
    originator
                     169.254.206.11
    tloc
                     169.254.206.11, private3, ipsec
    originator
                     169.254.206.11
                     169.254.206.11, private4, ipsec
    tloc
```

```
tloc 169.254.206.12, private2, ipsec originator 169.254.206.12 tloc 169.254.206.12 tloc 169.254.206.12 tloc 169.254.206.12 tloc 169.254.206.12, private4, ipsec originator 169.254.206.12 tloc 169.254.206.12 tloc 169.254.206.12, private5, ipsec
```

# **Related Information**

- **OMP Documentation**
- Troubleshoot OMP Route Instability in Failover Scenario
- Technical Support & Documentation Cisco Systems