

OMP Path Selection Problem when Egress Path is Enforced on vEdge

Contents

[Introduction](#)

[Topology](#)

[Configuration](#)

[Problem](#)

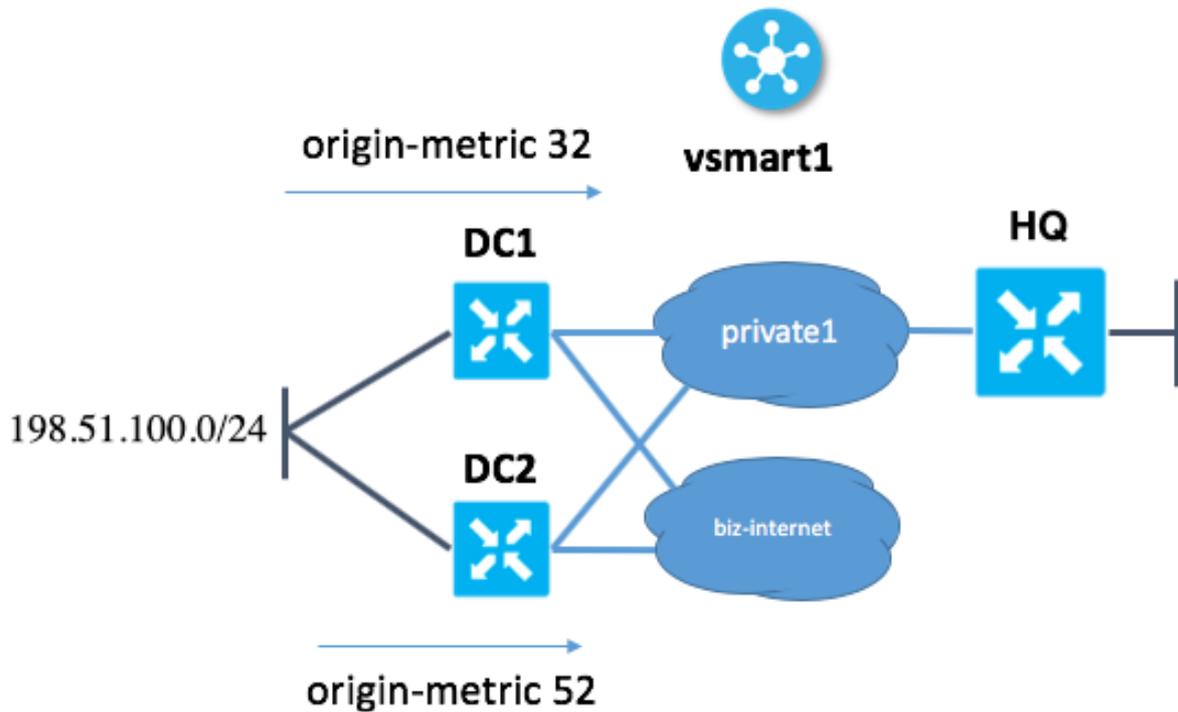
[Solution](#)

Introduction

This document describes the problem that occurs with a design of the redundancy when Overlay Management Protocol (OMP) path selection is enforced on a vEdge device and not on the vSmart controller that causes unwanted results and loss of reachability to the remote site in case of link failure even if the backup path is available. This problem is also alternatively known as "vSmart doesn't take into account TLOC state on remote vEdge".

Topology

In order to understand the problem better, here is a simple topology diagram that depicts the setup:



Configuration

Here you can find the brief description of the configuration.

- Site DC1 has TLOC colors "**private1**" and "**biz-internet**"
- Site DC2 has TLOC colors "**private1**" and "**biz-internet**"
- Site HQ has TLOC color "**private1**" only
- In DC1 and DC2 both colors are used for control connection to vSmart

Both DC sites (DC1 and DC2) advertise the same network, 198.51.100.0/24.

In each site, vEdge learns the router from the DC via some kind of dynamic routing protocol, e.g. Border Gateway Protocol (BGP).

Each DC site tags the prefix with a different metric:

At site DC1 vEdge set origin-metric 32

At site DC2 vEdge set origin-metric 52

```
hostname site-id system-ip
DC1      21    10.100.0.21
DC2      41    10.100.0.41
HQ       100   10.100.0.100
vSmart   100   10.100.0.20
```

Problem

At the time of normal operation:

1. vSmart receives 198.51.100.0/24 from both DC1 and DC2.

```
vsmart1# show omp routes 198.51.100.0/24
Code:
C -> chosen
I -> installed
Red -> redistributed
Rej -> rejected
L -> looped
R -> resolved
S -> stale
Ext -> extranet
Inv -> invalid
Stg -> staged
U -> TLOC unresolved
```

VPN	PREFIX	FROM PEER	PATH	STATUS	ATTRIBUTE	TLOC IP		
COLOR	ENCAP	PREFERENCE	ID LABEL		TYPE			
3	198.51.100.0/24	10.100.0.21	36 1003	C,R	installed	10.100.0.21		
biz-internet	ipsec -	<=====	METRIC 32 (PREFERRED)	10.100.0.21	49 1003	C,R		
installed	10.100.0.21	privatel ipsec -	<=====	METRIC 32 (PREFERRED)	10.100.0.41	36 1003	R	
installed	10.100.0.41	biz-internet ipsec -	<=====	METRIC 52	10.100.0.41	49 1003	R	installed
10.100.0.41	privatel ipsec -	<=====	METRIC 52					

2. vSmart advertises to HQ the route with destination DC1 (via private1 and biz-internet) because it has the lowest origin-metric as per [OMP route selection criteria](#).

```
vsmart1# show omp routes 198.51.100.0/24 vpn 3 detail
```

```
omp route entries for vpn 3 route 198.51.100.0/24
```

```
RECEIVED FROM: <===== RECEIVED FROM vEdge in DC1 in "biz-internet" color peer
10.100.0.21 path-id 36 label 1003 status C,R loss-reason not set lost-to-peer not set lost-to-
path-id not set Attributes: originator 10.100.0.21 type installed tloc 10.100.0.21, biz-
internet, ipsec ultimate-tloc not set domain-id not set overlay-id 1 site-id 21 preference not
set tag 1000030021 origin-proto eBGP origin-metric 32 as-path "65001 65001 65001" unknown-attr-
len not set RECEIVED FROM: <===== RECEIVED FROM vEdge in DC1 in "privatel" color
peer 10.100.0.21 path-id 49 label 1003 status C,R loss-reason not set lost-to-peer not set lost-
to-path-id not set Attributes: originator 10.100.0.21 type installed tloc 10.100.0.21, privatel,
ipsec ultimate-tloc not set domain-id not set overlay-id 1 site-id 21 preference not set tag
1000030021 origin-proto eBGP origin-metric 32 as-path "65001 65001 65001" unknown-attr-len not
set RECEIVED FROM: <===== RECEIVED FROM vEdge in DC2 in "biz-internet" color peer
10.100.0.41 path-id 36 label 1003 status R loss-reason origin-metric lost-to-peer 10.100.0.21
lost-to-path-id 49 Attributes: originator 10.100.0.41 type installed tloc 10.100.0.41, biz-
internet, ipsec ultimate-tloc not set domain-id not set overlay-id 1 site-id 41 preference not
set tag 1000030041 origin-proto eBGP origin-metric 52 as-path "65001 65001 65001 65001 65001"
unknown-attr-len not set RECEIVED FROM: <===== RECEIVED FROM vEdge in DC2 in
"privatel" color peer 10.100.0.41 path-id 49 label 1003 status R loss-reason tloc-id lost-to-
peer 10.100.0.41 lost-to-path-id 36 Attributes: originator 10.100.0.41 type installed tloc
10.100.0.41, privatel, ipsec ultimate-tloc not set domain-id not set overlay-id 1 site-id 41
preference not set tag 1000030041 origin-proto eBGP origin-metric 52 as-path "65001 65001 65001
65001 65001" unknown-attr-len not set ADVERTISED TO: <===== WE ADVERTISE TO HQ vEdge
ONLY BEST ROUTES WITH METRIC 32 peer 10.100.0.100 Attributes: originator 10.100.0.21 label 1003
path-id 4410 tloc 10.100.0.21, biz-internet, ipsec ultimate-tloc not set domain-id not set site-
id 21 overlay-id 1 preference not set tag 1000030021 origin-proto eBGP origin-metric 32 as-path
```

```
"65001 65001 65001" unknown-attr-len not set Attributes: originator 10.100.0.21 label 1003 path-id 4439 tloc 10.100.0.21, private1, ipsec ultimate-tloc not set domain-id not set site-id 21 overlay-id 1 preference not set tag 1000030021 origin-proto eBGP origin-metric 32 as-path "65001 65001 65001" unknown-attr-len not set
```

3. HQ vEdge flags the route with TLOC "**biz-internet**" as "**Inv,U**" because this vEdge does not have TLOC biz-internet.

4. HQ vEdge flags the route with TLOC "**private1**" as "**C,I,R**" and installs the route.

DC1 failure scenario:

1. In failure scenario, DC1 vEdge uplink in color "**private1**" fails (interface goes in down state) while "**biz-internet**" stays up.

2. vSmart receives 198.51.100.0/24 from DC1 (reachable only via **biz-internet**) and DC2 (**biz-internet** and **private1**).

3. vSmart advertises to HQ vEdge routes to DC1 (via **biz-internet**) because DC1 has the lowest metric.

```
vsmart1# show omp routes 198.51.100.0/24 detail
```

```
-----  
omp route entries for vpn 3 route 198.51.100.0/24  
-----
```

```
                RECEIVED FROM:  
peer            10.100.0.21  
path-id         36  
label           1003  
status          C,R  
loss-reason     not set  
lost-to-peer    not set  
lost-to-path-id not set  
Attributes:  
  originator    10.100.0.21  
  type          installed  
  tloc          10.100.0.21, biz-internet, ipsec  
  ultimate-tloc not set  
  domain-id     not set  
  overlay-id    1  
  site-id       21  
  preference    not set  
  tag           1000030021  
  origin-proto  eBGP  
  origin-metric 32  
  as-path       "65001 65001 65001"  
  unknown-attr-len not set  
                RECEIVED FROM:  
peer            10.100.0.41  
path-id         36  
label           1003  
status          R  
loss-reason     origin-metric  
lost-to-peer    10.100.0.21  
lost-to-path-id 36  
Attributes:  
  originator    10.100.0.41
```

type installed
tloc 10.100.0.41, biz-internet, ipsec
ultimate-tloc not set
domain-id not set
overlay-id 1
site-id 41
preference not set
tag 1000030041
origin-proto eBGP
origin-metric 52
as-path "65001 65001 65001 65001 65001"
unknown-attr-len not set

RECEIVED FROM:

peer 10.100.0.41
path-id 49
label 1003
status R
loss-reason tloc-id
lost-to-peer 10.100.0.41
lost-to-path-id 36

Attributes:

originator 10.100.0.41
type installed
tloc 10.100.0.41, privatel, ipsec
ultimate-tloc not set
domain-id not set
overlay-id 1
site-id 41
preference not set
tag 1000030041
origin-proto eBGP
origin-metric 52
as-path "65001 65001 65001 65001 65001"
unknown-attr-len not set

ADVERTISED TO:

peer 10.100.0.31

Attributes:

originator 10.100.0.21
label 1003
path-id 5906
tloc 10.100.0.21, biz-internet, ipsec
ultimate-tloc not set
domain-id not set
site-id 21
overlay-id 1
preference not set
tag 1000030021
origin-proto eBGP
origin-metric 32
as-path "65001 65001 65001"
unknown-attr-len not set

ADVERTISED TO:

peer 10.100.0.41

Attributes:

originator 10.100.0.21
label 1003
path-id 7689
tloc 10.100.0.21, biz-internet, ipsec
ultimate-tloc not set
domain-id not set
site-id 21
overlay-id 1
preference not set
tag 1000030021

```

origin-proto    eBGP
origin-metric   32
as-path         "65001 65001 65001"
unknown-attr-len not set

```

ADVERTISED TO: <===== THIS IS WHAT WE ADVERTISE TO HQ SITE peer 10.100.0.100 Attributes:
 originator 10.100.0.21 label 1003 path-id 4410 tloc 10.100.0.21, biz-internet, ipsec ultimate-
 tloc not set domain-id not set site-id 21 overlay-id 1 preference not set tag 1000030021 origin-
 proto eBGP origin-metric 32 as-path "65001 65001 65001" unknown-attr-len not set

4. HQ vEdge flags the route with TLOC "**biz-internet**" as "**Inv,U**" because this vEdge does not have TLOC **biz-internet**.

The result is that HQ vEdge cannot reach 198.51.100.0/24.

Solution

vSmart could have sent the routes towards DC2 (with less preferred higher metric) and in that case HQ vEdge would still reach the destination with the use of the "private1" TLOC via DC2, which is still up:

```

VEDGE-HQ-1# show bfd sessions site-id 41

```

DST PUBLIC	SYSTEM IP	SITE ID	STATE	SOURCE TLOC	DST PUBLIC	Detect	TX	REMOTE TLOC	COLOR	SOURCE IP	TRANSITIONS	
IP	IP			COLOR	PORT	ENCAP	MULTIPLIER	INTERVAL(msec)		UPTIME		
10.100.0.41	192.168.41.1	41	up	private1	12406	ipsec	7	1000	private1	192.168.11.1	12:04:02:25	0

But there is no route via "**private1**" TLOC via DC2 on HQ vEdge installed because vSmart has already selected biz-internet route with lower metric as the best path. vSmart does not advertise OMP routes with different metrics by default, hence does not let receiving vEdge device decide which path to take (and take into account available TLOCs and its states). vSmart does not take into account the TLOC colors available on the remote device (HQ vEdge in our case) to which you advertise the route and does not take into account its state because there is no such mechanism to control this.

This is the OMP corner case that can be seen in similar topology with iBGP route reflector and peering on physical interfaces addresses.

The first solution option is to use add path like functionality (RFC7911) available in OMP and called "[send-backup-paths](#)" on vSmart:

```

omp
  send-backup-paths

```

It advertises all available paths, so remote HQ vEdge chooses the path based on TLOC availability.

The second solution option here is to remove route-policy action "set metric" for the corresponding prefix on DC1 and DC2 vEdges and then perform centralized route selection enforcement via vSmart control-policy as shown here for example:

```

policy
lists
site-list site_11
site-id 11
!
prefix-list PREFIX
ip-prefix 198.51.100.0/24
!
control-policy SET_PREF
sequence 10
match route
prefix-list PREFIX
site-id 21
!
action accept
set
preference 200
!
!
!
sequence 20
match route
prefix-list PREFIX
site-id 41
!
action accept
set
preference 100
!
!
!
default-action accept
!
apply-policy
site-list site_11
control-policy SET_PREF out
!

```

Here, site-id 11 is the HQ vEdge and prefix-list PREFIX contains prefixes which you want to be preferred over one TLOC color or another. Since both OMP routes are on HQ vEdge, once vEdge can not reach biz-internet anymore, it will install a route via private1 in the Routing Information Base (RIB) from it's OMP routes table.