Cisco IOS-XE SD-WAN Installs OSPF External Route with DN-Bit

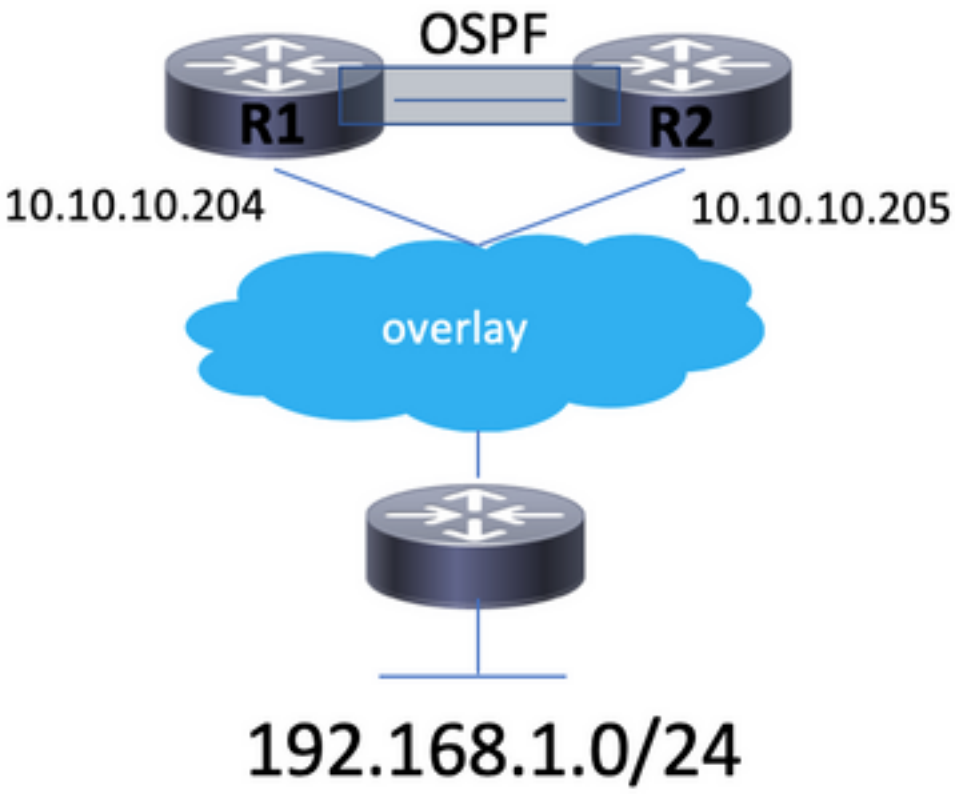
Contents

Introduction

This document describes the expected behavior of Cisco IOS®-XE SD-WAN software when Open Shortest Path First (OSPF) external routes are installed into the routing table.

Cisco IOS-XE SD-WAN Installs OSPF External Route with DN-Bit

The router that runs the Cisco IOS-XE SD-WAN software installs OSPF external routes (E1 or E2) into the routing table. For the purpose of the demonstration, consider this simple topology diagram:



Here is a pair of routers R1 and R2 that run Cisco IOS-XE SD-WAN software establish OSPF peering over service-side vpn (vrf 2 in this example). Routers have system-ip 10.10.10.204 and

10.10.10.205 correspondingly. System-ip is equal to OSPF router-id. Some other router advertises prefix 192.168.1.0/24 via Overlay Management Protocol (OMP) to this site.

Both routers are configured in a similar manner. The relevant configuration is provided here (the main point is that mutual redistribution between OSPF and OMP is done):

```
route-map omp2ospf permit 10
 set metric 1000
 set metric-type type-1
!
router ospf 2 vrf 2
 compatible rfc1583
 distance ospf external 110
 distance ospf inter-area 110
 distance ospf intra-area 110
 redistribute omp route-map omp2ospf
!
 omp
  no shutdown
  send-path-limit  4
  ecmp-limit       4
  graceful-restart
  no as-dot-notation
  timers
   holdtime               60
   advertisement-interval 1
   graceful-restart-timer 43200
   eor-timer              300
  exit
  address-family ipv4 vrf 2
   advertise ospf external
   advertise connected
   advertise static
  !
  address-family ipv4
   advertise connected
   advertise static
  !
  address-family ipv6
   advertise connected
   advertise static
  !
```

When normal condition routing table entry is done, 192.168.1.0/24 is installed into a Routing Information Base (RIB) from OMP and redistributed to OSPF. This entry looks like:

```
R1#sh ip route vrf 2 192.168.1.0 255.255.255.0

Routing Table: 2
Routing entry for 192.168.1.0/24
  Known via "omp", distance 251, metric 0, type omp
  Redistributing via ospf 2
  Advertised by ospf 2 subnets route-map omp2ospf
  Last update from 10.10.10.201 00:03:00 ago
  Routing Descriptor Blocks:
  * 10.10.10.201 (default), from 10.10.10.201, 00:03:00 ago
      Route metric is 0, traffic share count is 1

R1#show ip ospf database external 192.168.1.0
```

```
            OSPF Router with ID (172.16.1.204) (Process ID 2)

Type-5 AS External Link States

  LS age: 354
  Options: (No TOS-capability, DC, Downward)
  LS Type: AS External Link
  Link State ID: 192.168.1.0 (External Network Number )
  Advertising Router: 172.16.1.204
  LS Seq Number: 80000001
  Checksum: 0x25AE
  Length: 36
  Network Mask: /24
Metric Type: 1 (Comparable directly to link state metric)
MTID: 0
Metric: 1000
Forward Address: 0.0.0.0
External Route Tag: 0

  LS age: 355
  Options: (No TOS-capability, DC, Downward)
  LS Type: AS External Link
  Link State ID: 192.168.1.0 (External Network Number )
  Advertising Router: 172.16.1.205
  LS Seq Number: 80000001
  Checksum: 0x1FB3
  Length: 36
  Network Mask: /24
Metric Type: 1 (Comparable directly to link state metric)
MTID: 0
Metric: 1000
Forward Address: 0.0.0.0
External Route Tag: 0

R2#sh ip route vrf 2 192.168.1.0 255.255.255.0

Routing Table: 2
Routing entry for 192.168.1.0/24
  Known via "omp", distance 251, metric 0, type omp
  Redistributing via ospf 2
  Advertised by ospf 2 subnets route-map omp2ospf
  Last update from 10.10.10.201 00:04:13 ago
  Routing Descriptor Blocks:
  * 10.10.10.201 (default), from 10.10.10.201, 00:04:13 ago
      Route metric is 0, traffic share count is 1

R2#show ip ospf database external 192.168.1.0

            OSPF Router with ID (172.16.1.205) (Process ID 2)

Type-5 AS External Link States

  LS age: 317
  Options: (No TOS-capability, DC, Downward)
  LS Type: AS External Link
  Link State ID: 192.168.1.0 (External Network Number )
  Advertising Router: 172.16.1.204
  LS Seq Number: 80000001
  Checksum: 0x25AE
  Length: 36
  Network Mask: /24
Metric Type: 1 (Comparable directly to link state metric)
MTID: 0
```

```
Metric: 1000
Forward Address: 0.0.0.0
External Route Tag: 0

  LS age: 316
  Options: (No TOS-capability, DC, Downward)
  LS Type: AS External Link
  Link State ID: 192.168.1.0 (External Network Number )
  Advertising Router: 172.16.1.205
  LS Seq Number: 80000001
  Checksum: 0x1FB3
  Length: 36
  Network Mask: /24
Metric Type: 1 (Comparable directly to link state metric)
MTID: 0
Metric: 1000
Forward Address: 0.0.0.0
External Route Tag: 0
```

As you can see, both routers installed route into the RIB and then redistributed it into the OSPF. Both routers set DN-bit to external LSA type 5 and that should prevent these routes from being installed into the RIB as OSPF routes and hence redistributed back to the OMP, essentially preventing the loop. This is the same mechanism described in RFC 4576 and RFC 4577.

All routers have OMP peering established with vSmart controllers:

```
R1#show sdwan omp peers
R -> routes received
I -> routes installed
S -> routes sent

                        DOMAIN    OVERLAY   SITE
PEER            TYPE    ID        ID        ID        STATE    UPTIME           R/I/S
-----------------------------------------------------------------------------------
10.10.10.229    vsmart  1         1         1         up       1:19:35:34       30/12/5
10.10.10.230    vsmart  1         1         3         up       1:19:35:33       26/1/5

R2#show sdwan omp peers
R -> routes received
I -> routes installed
S -> routes sent

                        DOMAIN    OVERLAY   SITE
PEER            TYPE    ID        ID        ID        STATE    UPTIME           R/I/S
-----------------------------------------------------------------------------------
10.10.10.229    vsmart  1         1         1         up       0:01:38:48       30/10/6
10.10.10.230    vsmart  1         1         3         up       1:19:35:36       25/1/6
```

Now, R1 loses connectivity with both OMP peers:

```
Oct 11 12:53:57.777: %Cisco-SDWAN-Router-OMPD-3-ERRO-400002: R0/0: OMPD: vSmart peer
10.10.10.229 state changed to Init
Oct 11 12:53:57.777: %Cisco-SDWAN-Router-OMPD-6-INFO-400005: R0/0: OMPD: Number of vSmarts
connected : 1
Oct 11 12:53:58.777: %Cisco-SDWAN-Router-OMPD-3-ERRO-400002: R0/0: OMPD: vSmart peer
10.10.10.230 state changed to Init
Oct 11 12:53:58.777: %Cisco-SDWAN-Router-OMPD-6-INFO-400005: R0/0: OMPD: Number of vSmarts
connected : 0

R1#show sdwan omp peers
R -> routes received
```

```
I -> routes installed
S -> routes sent

                        DOMAIN    OVERLAY   SITE
PEER            TYPE    ID        ID        ID        STATE    UPTIME           R/I/S
-----------------------------------------------------------------------------------
10.10.10.229    vsmart  1         1         1         init-in-gr                30/12/0
10.10.10.230    vsmart  1         1         3         init-in-gr                26/1/0
```

R1 will mark the OMP route as stale (see OMP route state S), but continues keeping the route in the RIB installed by OMP protocol until graceful-restart-timer expired:

```
R1#show sdwan omp routes 192.168.1.0/24 | exclude not set
-------------------------------------------------
omp route entries for vpn 2 route 192.168.1.0/24
-------------------------------------------------
            RECEIVED FROM:
peer            10.10.10.229
path-id         1076
label           1002
status          C,I,R,S
    Attributes:
     originator       10.10.10.201
     type             installed
     tloc             10.10.10.201, biz-internet, ipsec
     overlay-id        1
     site-id          201207
     origin-proto     connected
     origin-metric    0
            RECEIVED FROM:
peer            10.10.10.230
path-id         775
label           1002
status          C,R,S
    Attributes:
     originator       10.10.10.201
     type             installed
     tloc             10.10.10.201, biz-internet, ipsec
     overlay-id        1
     site-id          201207
     origin-proto     connected
     origin-metric    0

R1#sh ip route vrf 2 192.168.1.0 255.255.255.0

Routing Table: 2
Routing entry for 192.168.1.0/24
  Known via "omp", distance 251, metric 0, type omp
  Redistributing via ospf 2
  Advertised by ospf 2 subnets route-map omp2ospf
  Last update from 10.10.10.201 00:23:35 ago
  Routing Descriptor Blocks:
  * 10.10.10.201 (default), from 10.10.10.201, 00:23:35 ago
      Route metric is 0, traffic share count is 1
```

The default graceful-restart-timer timer is 43,200 seconds (12 hours). Once it is expired, the route to 192.168.1.0/24 will still be there.

```
R1#sh ip route vrf 2 192.168.1.0 255.255.255.0

Routing Table: 2
```

```
Routing entry for 192.168.1.0/24
  Known via "ospf 2", distance 252, metric 1100, type extern 1
  Redistributing via omp
  Last update from 10.28.7.205 on Vlan2807, 00:04:11 ago
  Routing Descriptor Blocks:
  * 10.28.7.205, from 172.16.1.205, 00:04:11 ago, via Vlan2807
      SDWAN Down
      Route metric is 1100, traffic share count is 1

R1#show ip ospf database external 192.168.1.0

            OSPF Router with ID (172.16.1.204) (Process ID 2)

Type-5 AS External Link States

  LS age: 339
  Options: (No TOS-capability, DC, Downward)
  LS Type: AS External Link
  Link State ID: 192.168.1.0 (External Network Number )
  Advertising Router: 172.16.1.205
  LS Seq Number: 80000004
  Checksum: 0x19B6
  Length: 36
  Network Mask: /24
Metric Type: 1 (Comparable directly to link state metric)
MTID: 0
Metric: 1000
Forward Address: 0.0.0.0
External Route Tag: 0
```

It is installed as OSPF External Type 1 route now despite the fact that the OSPF LSA that corresponds has a DN-bit set.

Also, note that administrative distance (AD) is always 1 unit more than the AD of OMP (251 is the default for OMP, hence 252 in this case).

It is important to explain why the router installs this route with AD greater than the AD of OMP route. This is due to the fact that you try to prevent loop scenarios when OMP peering is reestablished again and reachability to the fabric is restored.

The process of route installation with AD=252 is also clearly seen if **debug ip routing** and **debug ip ospf rib redistribution** commands are enabled:

```
Oct 11 14:13:28.302: RT(2): del 192.168.1.0 via 10.10.10.201, omp metric [251/0]
Oct 11 14:13:28.303: RT(2): delete network route to 192.168.1.0/24
Oct 11 14:13:28.307: OSPF-2 REDIS: Notification to redistribute 192.168.1.0/24
Oct 11 14:13:28.307: RT(2): updating ospf 192.168.1.0/24 (0x2) [local lbl/ctx:1048577/0x0] omp-
tag:0  :
    via 10.28.7.205 Vl2807  0 1048578 0x100001
Oct 11 14:13:28.307: RT(2): add 192.168.1.0/24 via 10.28.7.205, ospf metric [252/1100]
```

This is expected behavior that was specifically introduced in Cisco IOS-XE SD-WAN software in order to avoid traffic blackhole scenarios when one of the routers is partitioned from the SD-WAN overlay. Blackhole might happen because a service side traffic is still load-balanced via both routers. This happens because two static routes point to both routers or some routes point to only one router that is partitioned.

In the case of ECMP (when R1 is partitioned from fabric) traffic follows two paths:

LAN -> R1 -> R2 -> remote router -> 192.168.1.0/24

LAN -> R2 -> remote router -> 192.168.1.0/24

Here, you can also see examples of outputs from R1 when R1 is partitioned from the fabric. As you can see, connectivity to LAN subnet 192.168.1.0/24 is still preserved via R2 (10.27.7.205 next-hop):

```
R1#ping vrf 2 192.168.1.1
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.1.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 4/33/44 ms
R1# traceroute vrf 2 192.168.1.1 numeric
Type escape sequence to abort.
Tracing the route to 192.168.1.1
VRF info: (vrf in name/id, vrf out name/id)
  1 10.28.7.205 4 msec 0 msec 0 msec
  2 192.168.1.1 4 msec *  0 msec
```