The bridge to possible

Advanced SD-WAN routing troubleshooting

Lessons learned from the field

Eugene Khabarov, BU Escalation Engineer, CCIE #51348
@ekhabaro

About me

and why I'm the right person to talk about OMP troubleshooting

Eugene Khabarov

Career:

- 15+ years in IT as support engineer, network engineer, consulting engineer, network architect
- Cisco TAC engineer since 2017 at Cisco Systems Belgium
- EMEA SD-WAN TAC Team Lead since 2019
- Enterprise BU (SD-WAN) escalation engineer since 2021

Key distinctives:

- Software engineer who writes (almost) no code for Cisco's products
- Translation proxy between customers, TAC and Software Engineering (a.k.a BU)
- Software bugs hunter, firefighter with (the worst) SD-WAN networks failures

Baseline and Objectives

- This is advanced level session, so basics are not covered, Cisco SD-WAN basic level knowledge is a must
- The session main objectives are:
  - share experience about some typical failures seen in the field to help you avoid them in your network
  - demonstrate some well-known SD-WAN features from different angle (not how to use them, but rather what problems misuse causes)
- Consider this session as a ”cookbook” for SD-WAN routing failures, but not a “Tour de Force”
- Not an exhaustive guide, there are always more…
- Based on real-life TAC cases and escalations
- The session is OMP protocol oriented mainly, no multicast routing, centralized control (almost), data or AAR policies discussed
- Main topics touched:
  - Implicit ACL and underlay routing
  - OMP tuning/features
  - OMP path selection issues
  - OMP interaction with service-side routing protocols:
    - OSPF
    - EIGRP
    - BGP
- Heavily CLI based, old-school classic ☺

Cisco Webex App

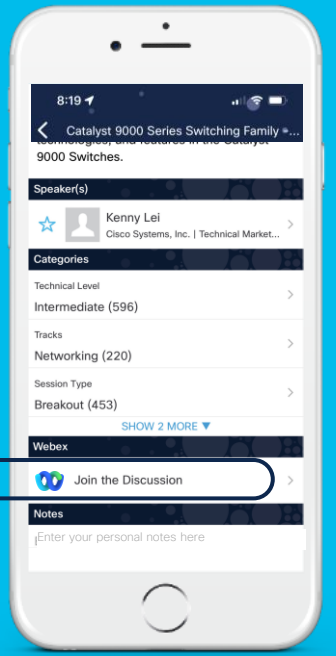
Questions?

Use Cisco Webex App to chat
with the speaker after the session

How

1. Find this session in the Cisco Live Mobile App
2. Click “Join the Discussion”
3. Install the Webex App or go directly to the Webex space
4. Enter messages/questions in the Webex space

Webex spaces will be moderated until February 24, 2023.



“It’s good to learn from your mistakes. It’s better to learn from other people’s mistakes.”

Warren Buffett

Agenda

- Part 1: SD-WAN Routing Troubleshooting basics
- Part 2: Issues seen in the field
  - 2.1 VPN 0 (GRT) Routing Cases
    - Case 1. Can not establish BGP peering with my ISP in the underlay
    - Case 2. BGP session established while it should not
  - 2.2 OMP Features Cases
    - Case 3. vSmart does not advertise any OMP routes
    - Case 4. Traffic is not load-balanced over ECMP
    - Case 5. OMP Path selection and global scalability
    - Case 6. OMP double failure scenario

Agenda

- 2.3 Service-side routing and OMP
  - Case 7. OSPF and DN-bit in SD-WAN
  - Case 8. WAN Edge does not install route from type3 LSA from area 0 into the RIB
  - Case 9. OMP to EIGRP redistribution
  - Case 10. OMP-BGP routing loop
  - Case 11. propagate-aspath and overlay-as
  - Case 12. Temporary blackholing on redundancy recovery with BGP

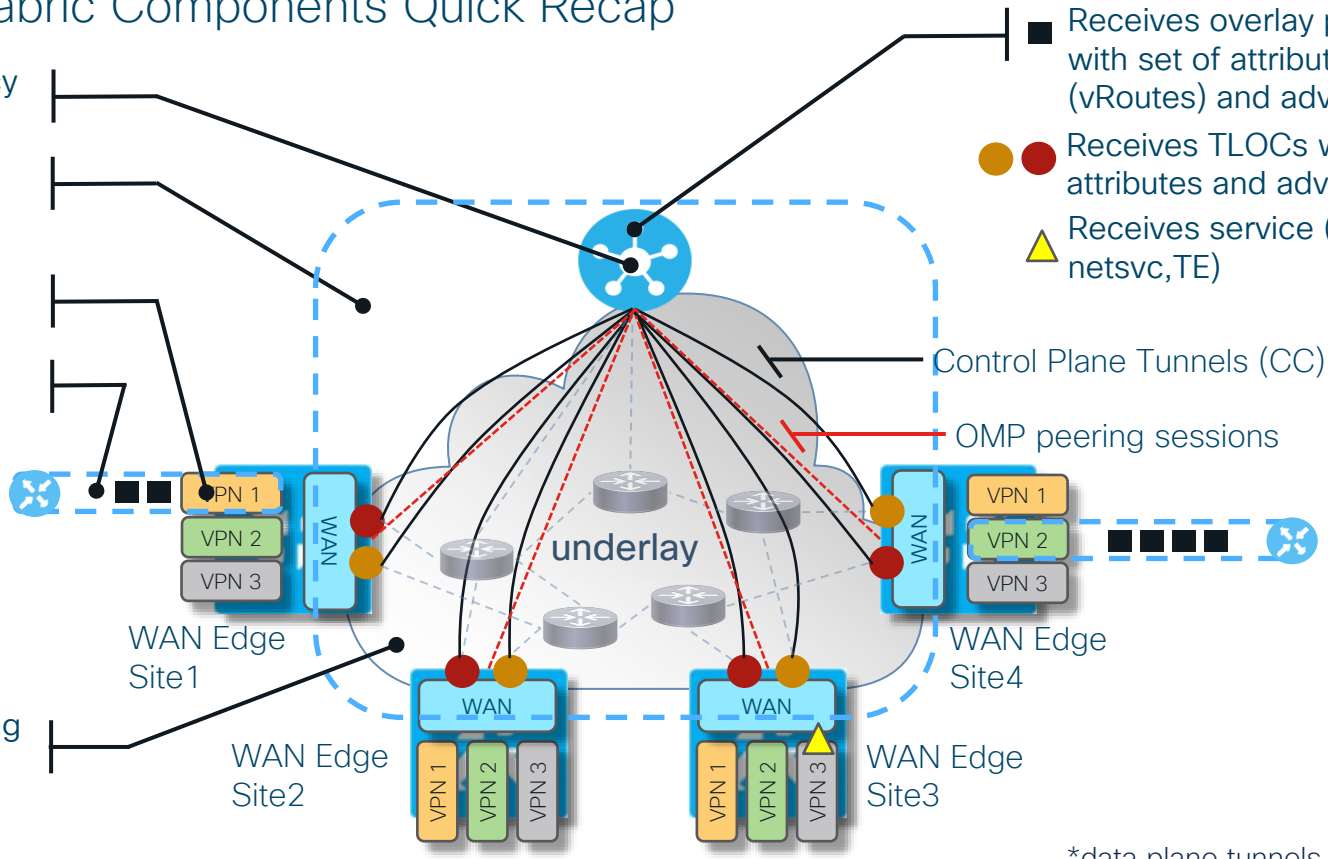
Q&A

Part 1: SD-WAN Routing Troubleshooting Basics

What should you know to troubleshoot various issues with routing and forwarding

Cisco SD-WAN Routing

SD-WAN Fabric Components Quick Recap



*data plane tunnels are not shown

Tshoot 101: OMP Peering

To exchange routing information, first of all peering with and between vSmart controllers must be established

```
cE1_BR1#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
-----------------------------------------------------------------------
10.0.0.101  vsmart  1       1       101    up       3:23:27:57 2086/65/2024
```

Tshoot 101: OMP Peering

If there is no OMP peering between WAN Edge and vSmart

- Check control connections and transport -> VPN 0 underlay routing tshoot
- Check for possible duplicate System ID
- Check OMP protocol status (**`show [sdwan] omp summary`**)

```
cE1_BR1#show sdwan omp peer
                        DOMAIN     OVERLAY    SITE
PEER             TYPE   ID         ID         ID         STATE     UPTIME        R/I/S
-------------------------------------------------------------------------------------
10.0.0.101       vsmart 1          1          2          init-in-gr             80/0/0
```

Tshoot 101: Graceful Restart Timer

If all vSmart controller connections are lost, the WAN Edge router continues to operate with the latest control plane information (vRoutes, TLOCs, IPSec keys, centralized policies, cflow template,) for the duration of configured OMP **`graceful-restart`** timer (12h default, 7d max).

```
vsmart# show omp peer
                    DOMAIN   OVERLAY   SITE
PEER        TYPE    ID       ID        ID         STATE      UPTIME     R/I/S
-------------------------------------------------------------------------------
10.0.0.1    vedge   1        1         2111002    down-in-gr            4/0/0
```

```
cE1_BR1# show sdwan omp peer
                      DOMAIN   OVERLAY   SITE
PEER          TYPE    ID       ID        ID       STATE      UPTIME     R/I/S
-------------------------------------------------------------------------------
10.0.0.101    vsmart  1        1         2        init-in-gr            82/82/0
```

*remember **`security ipsec rekey`** 2x times **`graceful-restart`**

Few other OMP timers to remember

`advertisement-interval` - the time between OMP Update packets, default is 1s.

`holdtime` – how long to wait before closing the OMP connection to a peer. If the peer does not receive 3 consecutive keepalive messages within the hold time, the OMP connection to the peer is closed. Default is 60 sec. (hello sent every 1/3 of holdtime)

`eor-timer` - how long to wait after an OMP session has gone down and then come back up before flushing stale routes which were not refreshed/updated. Default is 300 sec.

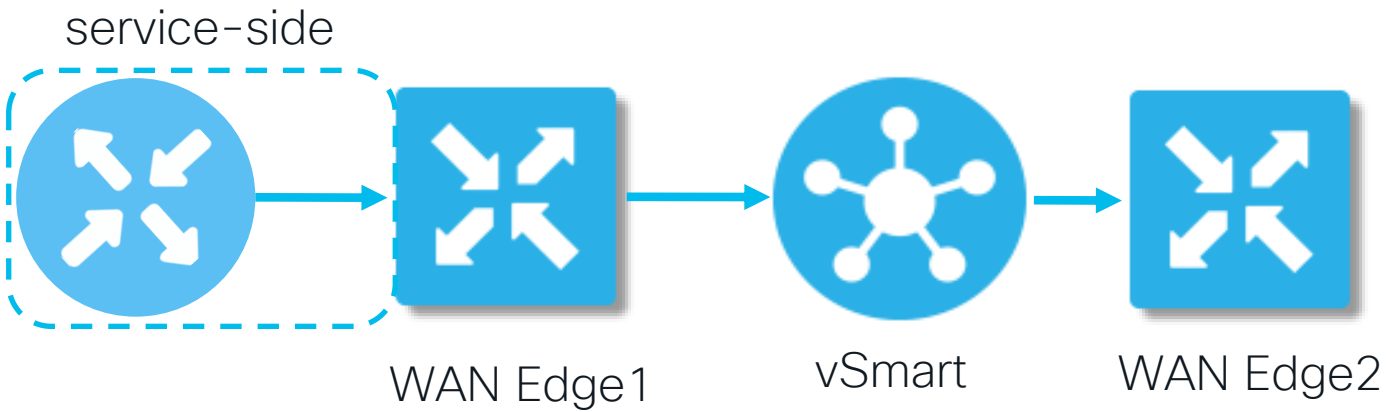
Tshoot 101: OMP Summary – What to pay attention for?

```
cE1_BR1#show sdwan omp summary
oper-state               UP
admin-state              UP
personality              vedge
omp-uptime               9:19:13:19
routes-received          4106
routes-installed         65
routes-sent              2022
tlocs-received           45
tlocs-installed          43
tlocs-sent               2
services-received        16
services-installed       0
services-sent            16
mcast-routes-received    0
mcast-routes-installed   0
mcast-routes-sent        0
hello-sent               42345
hello-received           42292
handshake-sent           12
handshake-received       12
alert-sent               9
alert-received           2
inform-sent              64
inform-received          63
update-sent              67492
update-received          63560
policy-sent              0
policy-received          2
total-packets-sent       109922
total-packets-received   105931
vsmart-peers             1
```

OMP routing basics

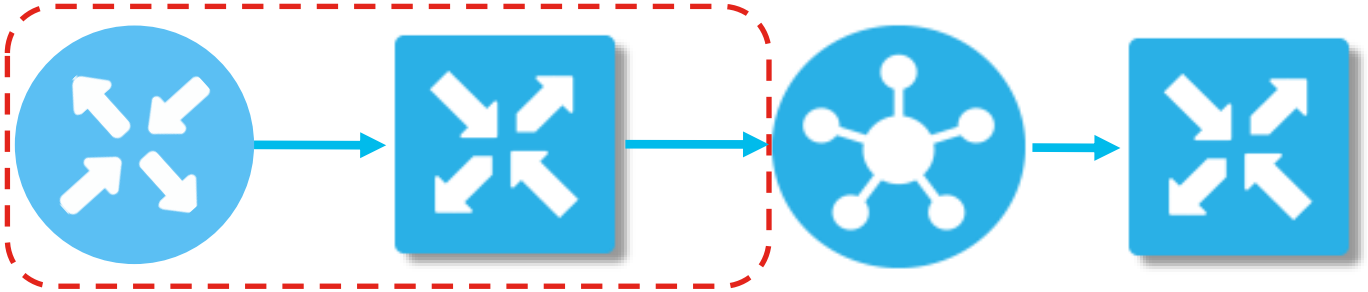
OMP Routing Basics - Topology

It will serve us as an exhibit for further discussion



Routing info announcement direction

OMP Routing Basics - WAN Edge
From service-side and out to vSmart



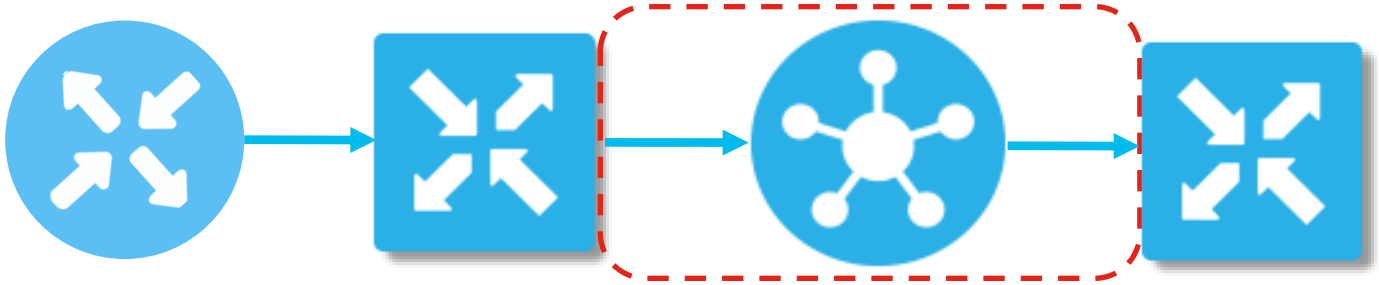
- By default, static, connected, and OSPF internal routes are automatically redistributed into OMP.

  All other (BGP, EIGRP, OSPF external) redistribution must be configured (into OMP and from OMP into another routing protocol)

  *redistribution can be controlled with localized policy starting from 20.5/17.5

- 4 best equal-cost paths are advertised to vSmart by default (configurable via **`send-path-limit [1-16]`**)
- Only the best paths are advertised to vSmart (routes installed into FIB) (non-best paths still can be sent if **`send-backup-paths`** configured)

OMP Routing Basics - vSmart

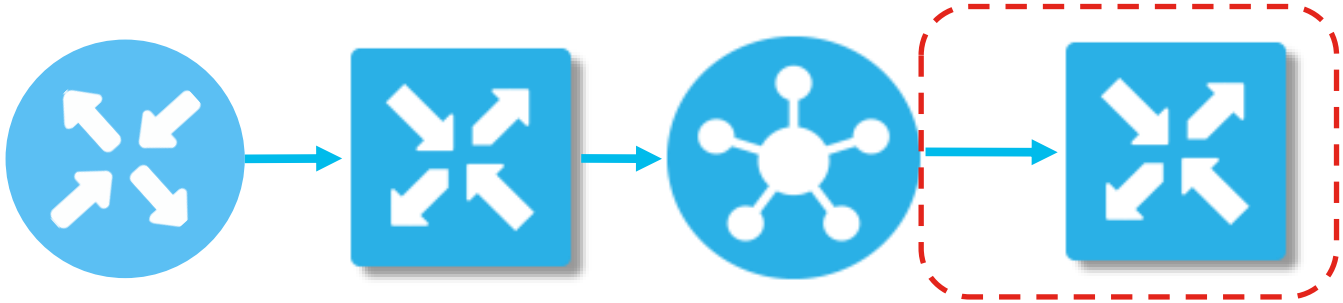


1. Applies incoming control policy
2. Performs best paths selection (TLOC preference and TLOC reachability is not considered on vSmart)
3. Stores best paths in an ordered (sorted) list
4. Applies outgoing control policy and advertises best paths to WAN Edges

- By default, only 4 best* equal-cost paths are advertised (configurable via **`send-path-limit [1-16]`** and **`controller-send-path-limit [4-128]`**)

*non-best paths can be sent also if **`send-backup-paths`** configured; 16 ECMP paths are sent by default between controller starting from 20.5 (behavior change)

OMP Routing Basics - WAN Edge
From vSmart



To decide which OMP routes are installed into the routing (RIB) and forwarding (FIB) tables:

- OMP performs loop avoidance (based on system-ip) and best path selection
- Implements localized policy

Up to 4 ECMP paths can be installed by default (configurable via **`ecmp-limit [1-16]`**)

- Installs route into RIB/FIB table if TLOC is active and there is a BFD session associated with it in the “up” state (route resolved).

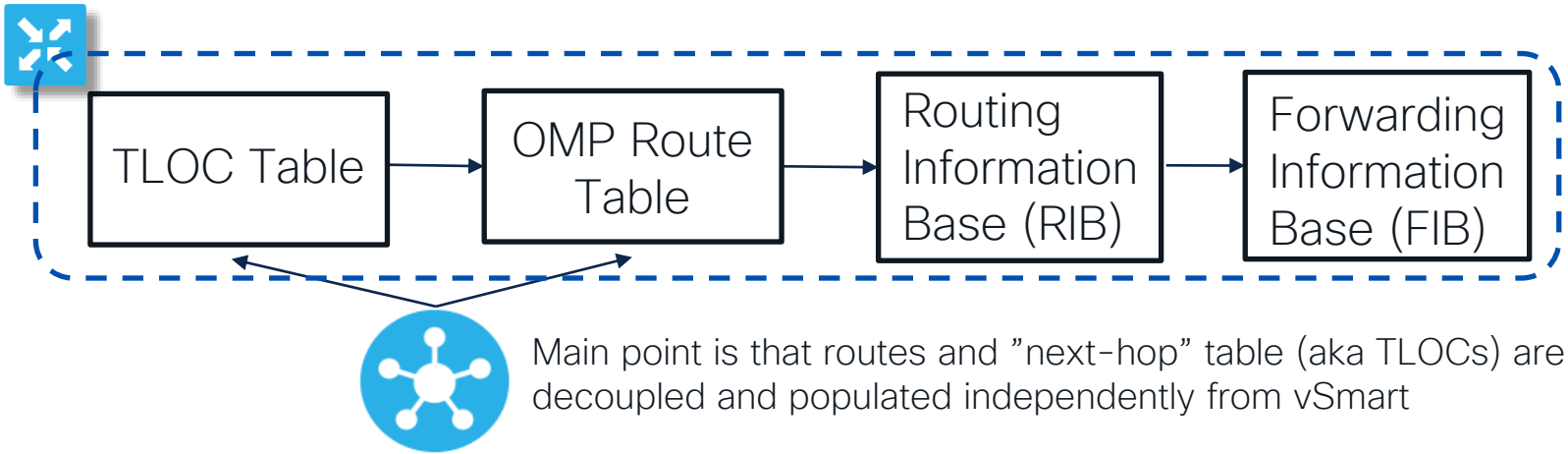
*Default OMP Administrative Distance is 250/251 for vEdge/cEdge and cannot be changed

Loop avoidance

- OMP loop avoidance is based on originator System-IP, not Site-ID value
- Native built-in loop prevention mechanisms when OMP interacts with EIGRP, OSPF and BGP:
- OSPF uses “Down Bit” (RFC 4577). Set when route redistributed from OMP to OSPF, on WAN Edge. When LSA distributed through service-side network gets to the other WAN Edge, route is not installed into RIB because DN bit is set*
- BGP uses SoO, extended community which value is set to the OMP site ID. When the other WAN Edge receives the BGP update from the service-side network and there SoO community matches its own site ID, then route will not be installed into RIB (and hence won’t be readvertised to OMP to prevent loops)*. BGP peers at site must send BGP extended communities and have the same site ID
- EIGRP uses “External Protocol” ID field. It is set to a value of “OMP-Agent”. When the other WAN Edge on the same site receives such update, it installs the route into the EIGRP topology table, sets “SDWAN-Down” flag and then install into the RIB with AD to 252. This, in turn, makes OMP the preferred route because it has an AD of 251

*- such BGP/OSPF routes still can be installed into the RIB with AD = OMP AD +1 (252 for cEdge) if no corresponding OMP route exists that can pre-empt. The same mechanism of “SDWAN-Down” flag used there as for EIGRP.

Routing/Forwarding Information Base (RIB/FIB)



- RIB: Control plane view on what WAN Edge router decided to install from all routing protocol tables like OMP and use for unicast forwarding
- FIB: Forwarding plane version of the RIB that includes fully resolved “next-hop” information, implemented in hardware on some platforms

OMP Routing Basics – Best Path Selection

Route Resolvability
WAN Edge: Next-hop TLOC is Reachable (data plane BFD session formed)

Route Preference
Prefer Route with highest route preference

MRF?

- Prefer access-region paths [intra-region] over core-region paths [inter-region]
- Enforce the transport-gateway router (TR) path check (user-config dependent):
- Prefer TR-sourced paths, or, prefer 'direct' paths' (if available) or ECMP between all paths
- Compare the affinity in the paths based on the affinity preference list configuration

TLOC Preference
WAN Edge: Prefer route with highest TLOC preference

Origin Type
Prefer route with best origin (Connected, Static, eBGP, EIGRP Internal, OSPF Intra, OSPF Inter, OSPF External, EIGRP External iBGP, Unknown/Unset

Origin Metric
Prefer OMP Route with lowest origin metric

Route Source Preference*
On vSmart: prefer Edge router-sourced route over vSmart-sourced route

Tiebreaker*
Prefer route received from lowest System-IP

Final Tiebreaker*
Prefer route from **highest** Private TLOC IP-address originated from the same site-id

Inbound control policy

Incoming update

Installed into OMP table, sorted

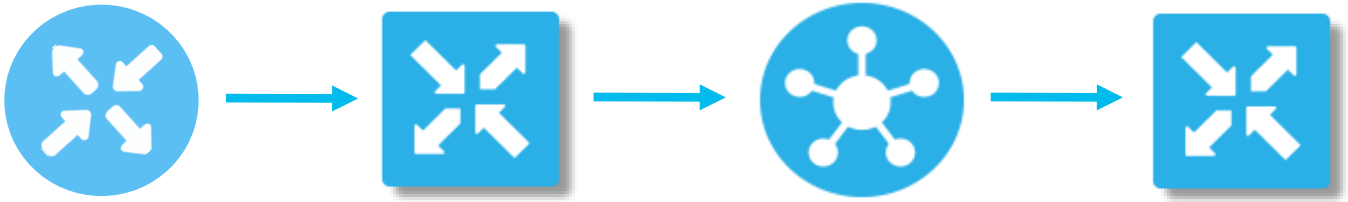
Outbound control policy

`send-backup-paths`

`send-path-limit`

Outgoing announce

Missing Route(s) troubleshooting algorithm



Check on WAN Edge:

1. RIB/FIB (**`show ip route/show [sdwan] ip fib`**)
2. OMP table if route is not in RIB (**`show [sdwan] omp route`**)
3. TLOC information presented (**`show [sdwan] omp tloc`**)
4. BFD session with remote TLOC (**`show [sdwan] bfd sessions`**) -> troubleshoot data plane tunnels
5. Local policy filtering on redistribution to/from OMP table (**`show run policy/show run route-map`**)

Check on vSmart:

1. OMP peering, control connections (**`show omp peer, show control connections`**)
2. OMP route and TLOC tables on vSmart (**`show omp route, show omp tloc`**)
3. Centralized control policies filters (**`show run policy, show run apply-policy, test policy`**)

RIB and FIB Tables (vEdge)

```
vEdge1# show ip route vpn 1 10.4.3.0/24

                                PROTOCOL  NEXTHOP  NEXTHOP NEXTHOP
VPN   PREFIX       PROTOCOL     SUB TYPE  IF NAME  ADDR    VPN  TLOC IP        COLOR        ENCAP  STATUS
-------------------------------------------------------------------------------------------------------
1     10.4.3.0/24  omp            -         -        -       -   10.255.241.101 mpls         ipsec  F,S
1     10.4.3.0/24  omp            -         -        -       -   10.255.241.101 biz-internet ipsec  F,S
```

```
vEdge1# show ip fib vpn 1 10.4.3.0/24

                           NEXTHOP     NEXTHOP          NEXTHOP  SA
VPN  PREFIX                IF NAME     ADDR             LABEL    INDEX    TLOC IP          COLOR
-----------------------------------------------------------------------------------------------
1    10.4.3.0/24           ipsec       10.4.1.2         1003     531      10.255.241.101   mpls
1    10.4.3.0/24           ipsec       64.100.1.23      1003     532      10.255.241.101   biz-internet
```

Tshoot 101: RIB and FIB Tables (IOS-XE)

```
cEdge1#show ip route vrf 1 10.4.3.0 255.255.255.0
Routing Table: 1
Routing entry for 10.4.3.0/24
  Known via "omp", distance 251, metric 0, type omp
  Last update from 10.255.241.102 05:42:07 ago
  Routing Descriptor Blocks:
   10.255.241.102 (default), from 10.255.241.102, 05:42:07 ago
      Route metric is 0, traffic share count is 1
 * 10.255.241.101 (default), from 10.255.241.101, 05:42:07 ago
      Route metric is 0, traffic share count is 1
```

```
cEdge1#show sdwan ip fib vpn 1 | include 10.4.3.0
1    10.4.3.0/24         ipsec       10.4.1.2        1003     1385     10.255.241.101   mpls
1    10.4.3.0/24         ipsec       10.4.2.2        1006     1390     10.255.241.102   mpls
1    10.4.3.0/24         ipsec       64.100.1.23     1003     1404     10.255.241.101   biz-internet
1    10.4.3.0/24         ipsec       64.100.1.24     1006     1405     10.255.241.102   biz-internet
```

`show sdwan ip fib` – shows only SDWAN part of FIB on IOS-XE, non-SDWAN routes are in IOS CEF table (**`show ip cef`**). OMP routes are never placed into IOS CEF!

Tshoot 101: OMP Routes and Associated TLOCs

- TLOCs are uniquely identified with 3 parameters: System ID, color, encapsulation

```
cE1_BR1#show sdwan omp routes 172.16.144.0/24 | b PATH
                                        PATH                   ATTRIBUTE
VPN    PREFIX           FROM PEER    ID     LABEL  STATUS     TYPE       TLOC IP    COLOR          ENCAP   PREFERENCE
-----------------------------------------------------------------------------------------------------------------------
1      172.16.144.0/24  10.0.0.101   3870   1003   C,I,R      installed  10.0.0.2   mpls           ipsec   -
                        10.0.0.101   3871   1003   C,I,R      installed  10.0.0.2   biz-internet   ipsec   -
```

```
cE1_BR1#show sdwan omp tlocs table
                                                                                                                        PUBLIC   PRIVATE
ADDRESS                                                   PSEUDO                 PUBLIC                   PRIVATE  PUBLIC  IPV6    PRIVATE  IPV6     BFD
FAMILY   TLOC IP   COLOR         ENCAP  FROM PEER   STATUS  KEY     PUBLIC IP     PORT    PRIVATE IP     PORT     IPV6    PORT    IPV6     PORT     STATUS
------------------------------------------------------------------------------------------------------------------------------------------------------------
         10.0.0.2  mpls          ipsec  10.0.0.101  C,I,R   1       192.168.9.12  12426   192.168.9.12   12426    ::      0       ::       0        up
         10.0.0.2  biz-internet  ipsec  10.0.0.101  C,I,R   1       192.168.10.12 12346   192.168.10.12  12346    ::      0       ::       0        up
```

vSmart/vEdge:

show omp route
show omp tlocs

Tshoot 101: Check TLOC advertisements

Confirm TLOC advertisement from WAN Edge

```
cE1_BR1#show sdwan omp tlocs table | include COLOR|0\.0\.0\.0
FAMILY    TLOC IP     COLOR         ENCAP  FROM PEER     STATUS     KEY  PUBLIC IP       PORT    PRIVATE IP      PORT    IPV6  PORT  IPV6  PORT  STATUS
ipv4      10.0.0.1    mpls          ipsec  0.0.0.0       C,Red,R    1    192.168.9.11    12386   192.168.9.11    12386   ::    0     ::    0     up
          10.0.0.1    biz-internet  ipsec  0.0.0.0       C,Red,R    1    192.168.10.11   12426   192.168.10.11   12426   ::    0     ::    0     up
```

Confirm TLOC received by vSmart and advertised to remote WAN Edge

```
vsmart1# show omp tlocs ip 10.0.0.1 color mpls advertised detail | nomore | begin peer.*10.0.0.2 | exclude "not set"
peer     10.0.0.2
    Attributes:
     encap-proto        0
     encap-spi          336
     encap-auth         sha1-hmac,ah-sha1-hmac
     encap-encrypt      aes256
     public-ip          192.168.9.11
     public-port        12386
     private-ip         192.168.9.11
     private-port       12386
     public-ip          ::
     public-port        0
     private-ip         ::
     private-port       0
     site-id            1
     preference         0
     weight             1
     version            3
     gen-id             0x80000015
     carrier            default
     restrict           1
     on-demand          0
     groups             [ 0 ]
     bandwidth          0
     qos-group          default-group
```

Verify that control policy does not filter TLOCs

```
policy
 lists
  tloc-list SITE1
   tloc 10.0.0.1 color mpls encap ipsec
  !
 !
!
policy
 lists
  site-list SITE1
   site-id 1
  !
  site-list SITE2
   site-id 2
  !
 !
 control-policy RESTRICT_MPLS
  sequence 10
   match tloc
    tloc-list SITE1
   !
   action reject
   !
  !
  default-action accept
 !
!
apply-policy
 site-list SITE1
  control-policy RESTRICT_MPLS in
 !
 site-list SITE2
  control-policy RESTRICT_MPLS out
 !
!
```

Notes:

- Check **`show omp tloc received`**.
- If a TLOC is Rejected (Rej) or Invalid (Inv), it won't be advertised to the other WAN Edge.
- Ensure that a control policy doesn't filter the TLOC when it's advertised from the vSmart.
- You can see that the TLOC is received on the vSmart but you won't see it on remote WAN Edge.

<<<=== the policy is applied to routing updates coming IN the vSmart, it will filter tlocs before adding them to the OMP table

OR

<<<=== the policy is applied to routing updates coming OUT the vSmart, it will filter tlocs after adding them to the OMP table

OMP Routing Typical Issue – TLOC resolvability

Missing Routes from 2nd DC in WAN Edge RIB. Expected 4 default routes, but only 2 installed into FIB.

```
cE1_BR1#sh ip route vrf 3 0.0.0.0

Routing Table: 3
Routing entry for 0.0.0.0/0, supernet
  Known via "omp", distance 251, metric 0, candidate default path, type omp
  Last update from 10.0.0.11 on Sdwan-system-intf, 00:14:06 ago
  Routing Descriptor Blocks:
  * 10.0.0.11 (default), from 10.0.0.11, 00:14:06 ago, via Sdwan-system-intf
      Route metric is 0, traffic share count is 1
```

```
cE1_BR1#show sdwan ip fib vpn 3
                                          NEXTHOP     NEXTHOP          NEXTHOP  SA
VPN  PREFIX                               IF NAME     ADDR             LABEL    INDEX    TLOC IP          COLOR
-------------------------------------------------------------------------------------------------------------------
3    0.0.0.0/0                            ipsec       192.168.9.13     1008     2189     10.0.0.11        mpls
3    0.0.0.0/0                            ipsec       192.168.10.13    1008     2199     10.0.0.11        biz-internet
```

```
cE1_BR1#show sdwan omp routes vpn 3 0.0.0.0/0 | begin PATH
            PATH                    PSEUDO
FROM PEER   ID     LABEL    STATUS  KEY     TLOC IP     COLOR          ENCAP  PREFERENCE
-------------------------------------------------------------------------------------
10.0.0.101  16844  1008     C,I,R   1       10.0.0.11   mpls           ipsec  -
10.0.0.101  16845  1008     C,I,R   1       10.0.0.11   biz-internet   ipsec  -
10.0.0.101  18924  1004     Inv,U   1       10.0.0.12   mpls           ipsec  -
10.0.0.101  18925  1004     Inv,U   1       10.0.0.12   biz-internet   ipsec  -
```

OMP Routing Typical Issue – TLOC resolvability (2)

```
cE1_BR1#show sdwan omp tlocs table

ADDRESS                                                    PSEUDO            PUBLIC                  PRIVATE PUBLIC  IPV6 PRIVATE IPV6  BFD
FAMILY  TLOC IP    COLOR         ENCAP  FROM PEER   STATUS KEY   PUBLIC IP     PORT    PRIVATE IP    PORT    IPV6    PORT IPV6    PORT
STATUS
----------------------------------------------------------------------------------------------------------------------------------------------
ipv4    10.0.0.1   mpls          ipsec  0.0.0.0     C,Red,R 1    192.168.9.11  12386   192.168.9.11  12386   ::      0    ::      0     up
        10.0.0.1   biz-internet  ipsec  0.0.0.0     C,Red,R 1    192.168.10.11 12426   192.168.10.11 12426   ::      0    ::      0     up
        10.0.0.11  mpls          ipsec  10.0.0.101  C,I,R  1     192.168.9.13  12406   192.168.9.13  12406   ::      0    ::      0     up
        10.0.0.11  biz-internet  ipsec  10.0.0.101  C,I,R  1     192.168.10.13 12346   192.168.10.13 12346   ::      0    ::      0     up
        10.0.0.12  mpls          ipsec  10.0.0.101  C,I,R  1     192.168.9.14  12406   192.168.9.14  12406   ::      0    ::      0     down
        10.0.0.12  biz-internet  ipsec  10.0.0.101  C,I,R  1     192.168.10.14 12426   192.168.10.14 12426   ::      0    ::      0     down
```

```
cE1_BR1#show sdwan bfd sessions "system-ip 10.0.0.12"
                                SOURCE TLOC    REMOTE TLOC                 DST PUBLIC    DST PUBLIC          DETECT      TX
SYSTEM IP  SITE ID  STATE       COLOR          COLOR          SOURCE IP    IP            PORT        ENCAP   MULTIPLIER  INTERVAL(msec) UPTIME   TRANSITIONS
------------------------------------------------------------------------------------------------------------------------------------------------------
10.0.0.12  34       down        mpls           mpls           192.168.9.11  192.168.9.14  12406      ipsec   7           1000           NA       0
10.0.0.12  34       down        biz-internet   biz-internet   192.168.10.11 192.168.10.14 12426      ipsec   7           1000           NA       1
```

OMP Routing Typical Issue – TLOC resolvability

There is new and better command available starting from 17.8.1/20.8.1 to do the same in one go - **`show [sdwan] omp verify-routes`**

```
cEdge1# show sdwan omp verify-routes vpn 1 0.0.0.0/0
                 PATH                    ATTRIBUTE                                                 TLOC                 BFD     RIB
FROM PEER        ID     LABEL    STATUS  TYPE       TLOC IP          COLOR            ENCAP  STATUS    PREFERENCE  STATUS  STATUS
-------------------------------------------------------------------------------------------------------------------------------
169.254.206.4    70     1003     C,I,R   installed  169.254.206.12   mpls             ipsec  C,I,R     -           up      F,S
169.254.206.4    71     1003     Inv,U   installed  169.254.206.12   biz-internet     ipsec  C,I,R     -           down    -
169.254.206.4    72     1003     C,I,R   installed  169.254.206.11   mpls             ipsec  C,I,R     -           up      F,S
169.254.206.4    73     1003     Inv,U   installed  169.254.206.11   biz-internet     ipsec  C,I,R     -           down    -
```

Useful Commands and Debugs

IOS XE SD-WAN:

```
show ip route vrf <>

show sdwan ip fib vpn <>
show sdwan omp tlocs
show sdwan omp routes
show sdwan omp peers
show sdwan omp summary

*debug platform software sdwan omp best-path
*debug platform software sdwan omp packets
*debug platform software sdwan omp events
```

* Replaced with `set platform software trace` from version 17.10

```
show logging process ompd internal <>
```

vSmart:

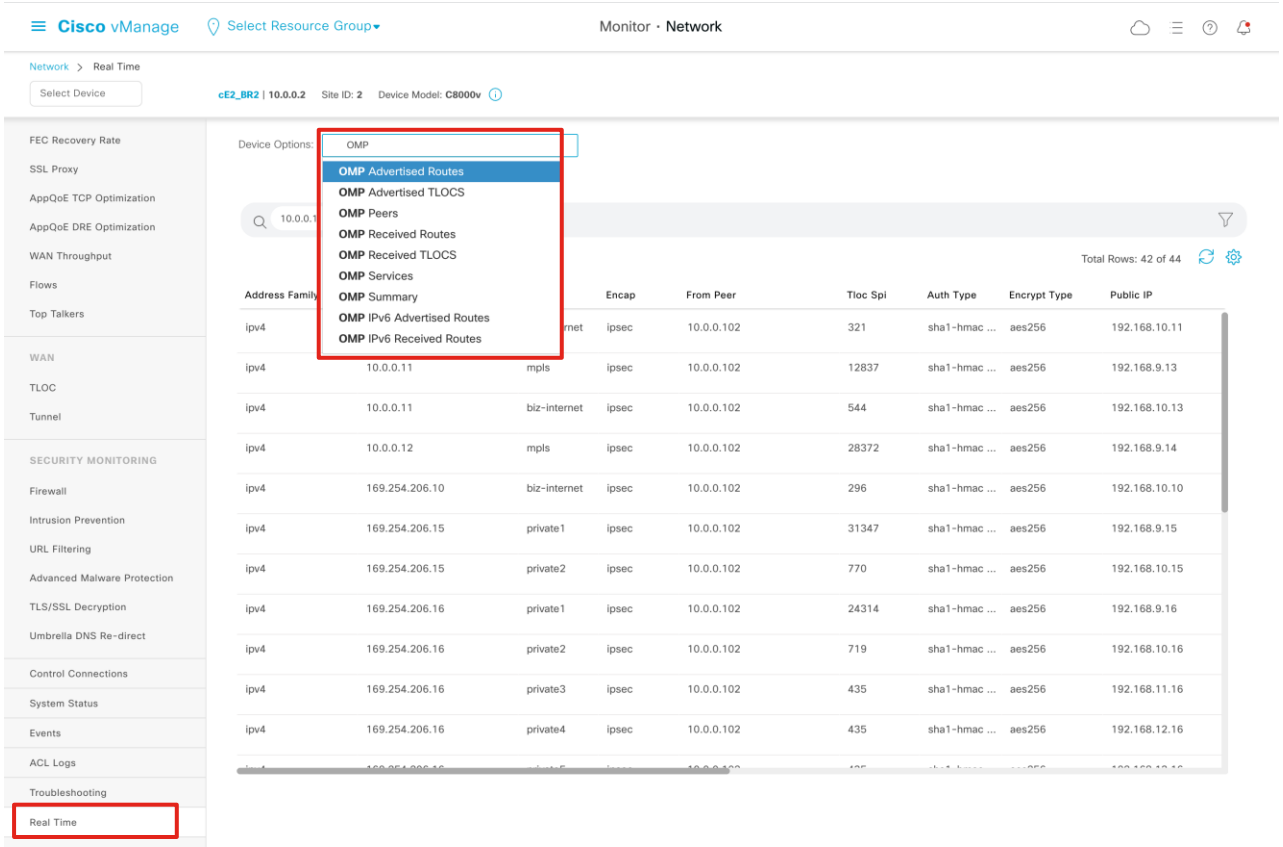
```
show omp tlocs <> [detail]
show omp routes <> [detail]
show omp peers <> [detail]
test policy match control-policy
```

vEdge:

```
show ip route vpn <>
show ip fib vpn <>

show omp tlocs
show omp routes
show omp peers
show omp summary
show omp verify-routes

debug omp best-path
debug omp packets
debug omp events
debug omp policy

show log /var/log/tmplog/vdebug tail -f
```

vManage – Monitor OMP routing

Monitor>Network>*[Device]*>Real Time



Useful Control Policy Debugging Commands

- Debugging:
  - **debug omp policy [direction | peer-address | vpn | prefix | level]**
  - **debug omp best-path [level]**
  - **debug omp packets [direction | peer-address | packet-type]**

  - **debug omp events [level]**
  - Logs will be stored in **/var/log/tmplog/vdebug**
  - Recommended approach to check them is either:
    - enter **vshell** and use **tail -f /var/log/tmplog/vdebug**
    - or **show log /var/log/tmplog/vdebug tail -f**
- **show support omp peer peer-ip <system-ip>** can be used to find policies applied to a

```
vsmart1# show support omp peer peer-ip 10.0.0.1 | include -pol
        site-pol: BR1 route-pol-in: TAG route-pol-out: None data-pol-in: None
           data-pol-out: None pfr-pol: None mem-pol: None cflowd:None
```

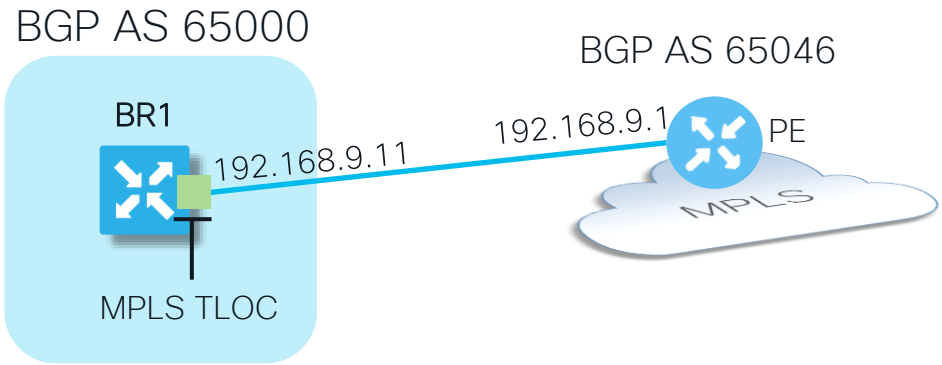
- starting from 20.8.1 **test policy match control-policy** can be used to find matching sequence in a policy on vSmart control policy

Part 2: Issues Seen in the Field

VPN 0 (GRT) Routing Cases underlay routing issues

Case 1. Can not establish BGP peering with my ISP in the underlay

Case 1. Can not establish BGP peering with my ISP in the underlay. Topology.



Case 1. Can not establish BGP peering with my ISP in the underlay

Symptoms. BGP peer bouncing between ”Idle” and “Active”

```
BR1#show ip bgp summary
BGP router identifier 10.0.0.11, local AS number 65000
BGP table version is 1, main routing table version 1

Neighbor        V           AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
192.168.9.1     4        65046       0       0        1    0    0 00:01:56 Idle

cE1_BR1#show ip bgp summary
BGP router identifier 10.0.0.11, local AS number 65000
BGP table version is 1, main routing table version 1

Neighbor        V           AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
192.168.9.1     4        65046       0       0        1    0    0 00:02:49 Active
```

BGP peer is reachable via icmp

```
BR1#ping 192.168.9.1
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 192.168.9.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/1/1 ms
```

But can not connect to tcp/179:

```
BR1#telnet 192.168.9.1 179
Trying 192.168.9.1, 179 ...
% Connection timed out; remote host not responding
```

Your ISP denies problems on their side…

Case 1. Can not establish BGP peering with my ISP in the underlay

Form a theory. What would we check further?

Facts:

- It's not transport issue, peer is reachable via icmp
- It's not BGP misconfiguration, we are trying to connect to the right peer
- It's not ISP issue, we trust in our ISP

Case 1. Can not establish BGP peering with my ISP in the underlay

packet-trace tool is our universal troubleshooting helper:

```
BR1#debug platform packet-trace packet 128
 Please remember to turn on 'debug platform condition start' for packet-trace to work
BR1#debug platform condition ipv4 192.168.9.1/32 both
BR1#debug platform condition start
BR1#show platform packet-trace summary
Pkt    Input                     Output                    State  Reason
0      INJ.2                     Gi2                       FWD
1      Gi2                       Gi2                       DROP   479 (SdwanImplicitAclDrop)
2      INJ.2                     Gi2                       FWD
```

Case 1. Can not establish BGP peering with my ISP in the underlay

We will find out that it was BGP packet from ISP router blocked by Implicit ACL:

```
BR1#show platform packet-trace packet 1
Packet: 1           CBUG ID: 1
Summary
  Input     : GigabitEthernet2
  Output    : GigabitEthernet2
  State     : DROP 479 (SdwanImplicitAclDrop)
  Timestamp
    Start   : 3038482805441123 ns (12/30/2022 13:03:48.251693 UTC)
    Stop    : 3038482805473473 ns (12/30/2022 13:03:48.251725 UTC)
Path Trace
  Feature: IPV4(Input)
    Input       : GigabitEthernet2
    Output      : <unknown>
    Source      : 192.168.9.1
    Destination : 192.168.9.11
    Protocol    : 6 (TCP)
      SrcPort   : 179
      DstPort   : 22575
  Feature: SDWAN Implicit ACL
    Action : DISALLOW
    Reason : SDWAN_SERV_BGP
```

Case 1. Can not establish BGP peering with my ISP in the underlay

`implicit-acl-logging` could be also enabled to identify drops:

```
BR1(config)# policy
BR1(config-policy)# implicit-acl-logging
BR1(config-policy)# commit
```

And then drops will be logged:

```
BR1#show logging last 100 | include Implicit-ACL
Dec 30 13:07:59.212: %Cisco-SDWAN-cE1-FTMD-5-NTCE-1000026: FLOW LOG vpn-0 src: 192.168.9.1/179 dst: 192.168.9.11/35747
proto: 6 tos: 192 inbound-acl, Implicit-ACL, Result: denyPkt SDWAN_SERV_BGP count: 1 bytes: 58 Ingress-Intf:
GigabitEthernet2 Egress-intf: GigabitEthernet2
```

Case 1. Can not establish BGP peering with my ISP in the underlay

Solution 1: allow BGP in implicit ACL:

```
sdwan
 interface GigabitEthernet2
  tunnel-interface
   allow-service bgp

BR1#show ip bgp summary
BGP router identifier 10.0.0.11, local AS number 65000
BGP table version is 1, main routing table version 1

Neighbor        V           AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
192.168.9.1     4        65046       4       4        1    0    0 00:00:24        1
```

Case 1. Can not establish BGP peering with my ISP in the underlay

Solution 2: configure explicit ACL as in the example:

Implicit ACL	SD-WAN Explicit ACL	SD-WAN Explicit ACL (default)	Result
✓	X		X
✓		X	✓
X	✓		✓
X		✓	X

```
sdwan
 policy
 access-list ALLOW_BGP_PEER
   sequence 10
    match
     source-ip         192.168.9.1/32
     destination-port 179
    !
    action accept
    !
   !
   sequence 20
    match
     source-ip    192.168.9.1/32
     source-port 179
    !
    action accept
    !
   !
   default-action drop
  !
!
sdwan
 interface GigabitEthernet0/0/0
 access-list ALLOW_BGP_PEER in
 exit
!
```

- Control plane tunnels are not affected
- Data plane tunnels (“BFD”) are not affected
- Overlay traffic is not affected

Case 2. Why BGP established even if “no allow-service bgp” configured?

Case 2. Why BGP established even if “no allow-service bgp” configured?

Symptoms. BGP peer has established session:

```
BR1#show ip bgp summary | b Neighbor
Neighbor        V           AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
192.168.9.1     4        65000      12      11        2    0    0 00:04:21        1
```

But BGP is not allowed under **tunnel-interface** section

```
BR1#show sdwan running-config "sdwan interface GigabitEthernet2 tunnel-interface allow-service"
sdwan
 interface GigabitEthernet2
  tunnel-interface
   no allow-service all
   no allow-service bgp
   no allow-service dhcp
   allow-service dns
   allow-service icmp
   no allow-service sshd
   no allow-service netconf
   no allow-service ntp
   no allow-service ospf
   no allow-service stun
   no allow-service https
   no allow-service snmp
  exit
 exit
!
```

Case 2. Why BGP established even if “no allow-service bgp” configured?

If session is cleared, it got established again:

```
BR1#clear ip bgp *

BR1#show ip bgp summary | b Neighbor
Neighbor        V           AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
192.168.9.1     4        65000       0       0        1    0    0 00:00:13 Idle

BR1#show ip bgp summary | b Neighbor
Neighbor        V           AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
192.168.9.1     4        65000       6       2        1    0    0 00:00:02        1
```

With packet-trace you can even see few “SdwanImplicitAclDrop” like in the Case 1

```
BR1#debug platform condition ipv4 192.168.9.1/32 both
BR1#debug platform packet-trace packet 128
Please remember to turn on 'debug platform condition start' for packet-trace to work
BR1#show platform packet-trace summary | b SdwanImplicitAclDrop
1     Gi2                         Gi2                 DROP   479 (SdwanImplicitAclDrop)
2     INJ.2                       Gi2                 FWD
3     Gi2                         Gi2                 DROP   479 (SdwanImplicitAclDrop)
4     Gi2                         internal0/0/rp:0    PUNT   11  (For-us data)
5     INJ.2                       Gi2                 FWD
6     Gi2                         internal0/0/rp:0    PUNT   11  (For-us data)
…
```

Case 2. Why BGP established even if “no allow-service bgp” configured?

And this is indeed BGP packet from the neighbor dropped due to Implicit ACL not allowing BGP service:

```
BR1#show platform packet-trace packet 3 | begin Summary
Summary
  Input     : GigabitEthernet2
  Output    : GigabitEthernet2
  State     : DROP 479 (SdwanImplicitAclDrop)
State     : DROP 479 (SdwanImplicitAclDrop)
Timestamp
    Start   : 965430525669342 ns (05/23/2022 16:08:35.119054 UTC)
    Stop    : 965430525683208 ns (05/23/2022 16:08:35.119068 UTC)
Path Trace
  Feature: IPV4(Input)
    Input       : GigabitEthernet2
    Output      : <unknown>
    Source      : 192.168.9.1
    Destination : 192.168.9.11
    Protocol    : 6 (TCP)
      SrcPort   : 179
      DstPort   : 53399
  Feature: SDWAN Implicit ACL
    Action : DISALLOW
    Reason : SDWAN_SERV_BGP
```

Case 2. Why BGP established even if “no allow-service bgp” configured?

Form a theory. What would we check further?

Facts:

- It’s not BGP misconfiguration
- There is no explicit ACL configured to override implicit ACL
- Issue is reproducible every time and consistent across the versions and platfroms (unlikely a bug)
- Only some packets are dropped, not all of them

Case 2. Why BGP established even if “no allow-service bgp” configured?

Let’s check next BGP packet (nr.5) that was allowed.

Compared to previous packets, this one is self-generated by cE1 (injected)

Destination is BGP peer

```
show platform packet-trace packet 5
Packet: 5           CBUG ID: 13887
Summary
  Input     : INJ.2
  Output    : GigabitEthernet2
  State     : FWD
Timestamp
    Start   : 965431957078314 ns (05/23/2022 16:08:36.550463 UTC)
    Stop    : 965431957550849 ns (05/23/2022 16:08:36.550936 UTC)
Path Trace
  Feature: IPV4(Input)
    Input       : internal0/0/rp:0
    Output      : <unknown>
    Source      : 192.168.9.11
    Destination : 192.168.9.1
    Protocol    : 6 (TCP)
      SrcPort   : 11600
      DstPort   : 179
  Feature: SDWAN Internal Intf
    VRF ID      : 0 (Global VPN 0)
    Encap Type  : unknown
    IP DSCP     : 48
    IP Version  : 4
    IP Protocol : 6
    Dst Port    : 179
    Is Marked High Priority         : NO
    Is SDWAN Control Tunnel Traffic : NO
    Set HIGH_QUEUE                  : NO (NOT marked high priority, NOT SDWAN
control tunnel traffic)
    Skip SDWAN Policy               : TRUE
```

Case 2. Why BGP established even if “no allow-service bgp” configured?

Let’s check one more packet (nr.6) that was allowed:

```
BR1#show platform packet-trace packet 6
Packet: 6           CBUG ID: 14142
Summary
  Input     : GigabitEthernet2
  Output    : internal0/0/rp:0
  State     : PUNT 11  (For-us data)
  Timestamp
    Start   : 965431958123058 ns (05/23/2022 16:08:36.551508 UTC)
    Stop    : 965431958153208 ns (05/23/2022 16:08:36.551538 UTC)
Path Trace
  Feature: IPV4(Input)
    Input       : GigabitEthernet2
    Output      : <unknown>
    Source      : 192.168.9.1
    Destination : 192.168.9.11
    Protocol    : 6 (TCP)
      SrcPort   : 179
      DstPort   : 5062
  Feature: SDWAN Implicit ACL
    Action : ALLOW
    Reason : SDWAN_NAT_DIA
```

This packet is from the BGP peer and intended for local router (punted to CPU) and the reason to allow the packet is shown as SDWAN NAT DIA

Case 2. Why BGP established even if “no allow-service bgp” configured?

In this case router with ”implicit ACL” acts like a stateful firewall, egress packet nr.5 resulted in return (ingress) packet nr.6 to be allowed.

And it works that way only if NAT enabled (e.g. for DIA):

```
ip nat route vrf 1 0.0.0.0 0.0.0.0 global
ip nat inside source list nat-dia-vpn-hop-access-list interface GigabitEthernet2 overload
!
interface GigabitEthernet2
 ip nat outside
```

Or simply said, NAT entry has a precedence over implicit ACL. Explicit SD-WAN ACL can still override this.

Case 2. Why BGP established even if “no allow-service bgp” configured?

Egress self-generated packet (nr.5) creates NAT table entry for BGP session and return traffic is allowed by this entry:

```
BR1#show ip nat translations
Pro  Inside global         Inside local          Outside local         Outside global
tcp  192.168.9.11:5063     192.168.9.11:60456    192.168.9.1:179       192.168.9.1:179
Total number of translations: 1

BR1#show sdwan nat-fwd ip-nat-translation-verbose | b 179
nat-fwd ip-nat-translation-verbose 192.168.9.11 192.168.9.1 60456 179 0 6
 inside-global-addr  192.168.9.11
 outside-global-addr 192.168.9.1
 inside-global-port  5063
 outside-global-port 179
 flags               2113536
 application-type    0
 entry-id            0x3279dbd0
 in_mapping_id       1
 out_mapping_id      0
 create_time         "Mon May 23 16:08:49 2022"
 last_used_time      "Mon May 23 16:22:34 2022"
 pkts_in             40
 pkts_out            37
 timeout             "86381 seconds"
 usecount            1
 input-idb           internal0/0/rp:0
 output-idb          GigabitEthernet2
 bytes_in            1211
 bytes_out           1304
```

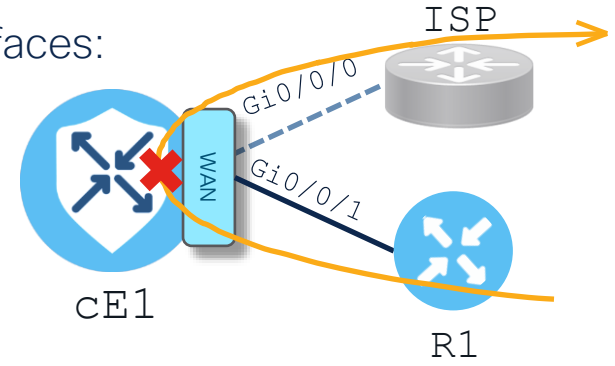
Preceding packets are dropped legitimately when remote peer tries to init a session

Case 2 ½. Forwarding in global routing table (VPN 0) is broken

Case 2 ½. Forwarding in global routing table (GRT/vpn 0) is broken

Objective and configuration – forwarding between GRT interfaces:

- cEdge 1 (cE1) establishes eBGP peering with ISP
- ISP advertises default route
- cE1 advertises cE1-R1 subnet via BGP
- R1 connected to the interface of cE1 in vpn 0
- R1 use static default route (or peer with cE1 via eBGP)
- R1 should be able to reach the Internet



Problem:

- R1 can not reach to the Internet
- traceroute shows packets are not passing through cE1, but connectivity to R1 is fine:

```
BR1#ping 203.0.113.22 source 198.51.100.1
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 203.0.113.22, timeout is 2
seconds:
Packet sent with a source address of 198.51.100.1
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/1/1 ms
```

```
router bgp 65128
 neighbor 203.0.113.22 remote-as 64537
 !
 address-family ipv4
  network 198.51.100.0
  neighbor 203.0.113.22 activate
!
ip route 198.51.100.0 255.255.255.0 Null0 254
!
interface GigabitEthernet0/0/0
 description "To ISP in VPN 0/GRT"
 ip address 203.0.113.23 255.255.255.252
!
interface GigabitEthernet0/0/1
 description "To R1 in VPN 0/GRT"
 ip address 198.51.100.1 255.255.255.252
!
```

Case 2 ½. Forwarding in global routing table (GRT/vpn 0) is broken

Form a theory. What would we check further?

Facts:

- IPv4 forwarding in global routing table is very basic functionality enabled by default
- It's not BGP misconfiguration of any router of any sort
- It's not R1 static routing misconfig or any other issues with R1

What would be the the difference between router running in controler mode (SD-WAN) vs autonomous mode ("normal" IOS-XE") because configuration mentioned before supposed to work?

It SD-WAN tunnel-enabled tranport interface (Gi0/0/0).

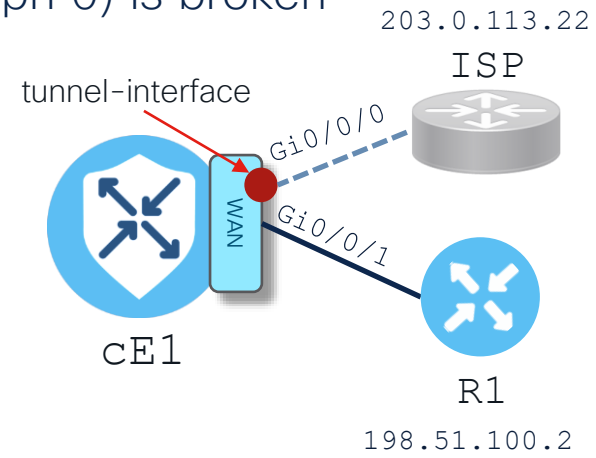
Case 2 ½. Forwarding in global routing table (GRT/vpn 0) is broken

Let's check why cE1 dropping traffic:

```
cE1#debug platform condition ipv4 198.51.100.2/32 both
cE1#debug platform packet-trace packet 1024 fia-trace
cE1#debug platform condition start
```

```
R1#ping 203.0.113.22 source 198.51.100.2
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 203.0.113.22, timeout is 2 seconds:
Packet sent with a source address of 198.51.100.2
.....
Success rate is 0 percent (0/5)
```

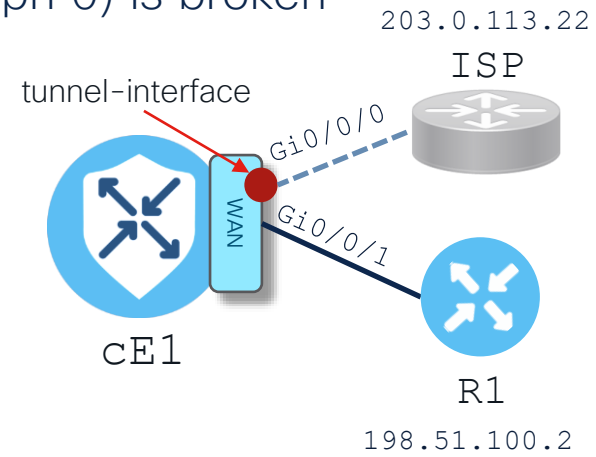
```
cE1#show platform packet-trace summary
Pkt    Input                Output               State  Reason
0      Gi0/0/1              Gi0/0/0              FWD
1      Gi0/0/0              Gi0/0/0              DROP   480 (SdwanImplicitAclDrop)
2      Gi0/0/1              Gi0/0/0              FWD
3      Gi0/0/0              Gi0/0/0              DROP   480 (SdwanImplicitAclDrop)
4      Gi0/0/1              Gi0/0/0              FWD
```



Case 2 ½. Forwarding in global routing table (GRT/vpn 0) is broken

Then check packet details:

```
BR1#show platform packet-trace packet 1
Packet: 1           CBUG ID: 1
Summary
  Input      : GigabitEthernet0/0/0
  Output     : GigabitEthernet0/0/0
  State      : DROP 480 (SdwanImplicitAclDrop)
  Timestamp
    Start    : 8849923204756 ns (09/30/2020 11:07:10.23093
UTC)
    Stop     : 8849923236231 ns (09/30/2020 11:07:10.23124
UTC)
Path Trace
  Feature: IPV4(Input)
    Input        : GigabitEthernet0/0/0
    Output       : <unknown>
    Source       : 198.51.100.2
    Destination  : 203.0.113.22
    Protocol     : 1 (ICMP)
<skipped>
  Feature: SDWAN Implicit ACL
    Action : DISALLOW
    Reason : SDWAN_IMPL_ACL_DEFAULT
<skipped>
  Feature: IPV4_SDWAN_IMPLICIT_ACL
    Entry        : Input - 0x8120ff50
    Input        : GigabitEthernet0/0/0
    Output       : <unknown>
    Lapsed time  : 27311 ns
```



Case 2 ½. Forwarding in global routing table (GRT/vpn 0) is broken

How to solve this? Wrong ways.

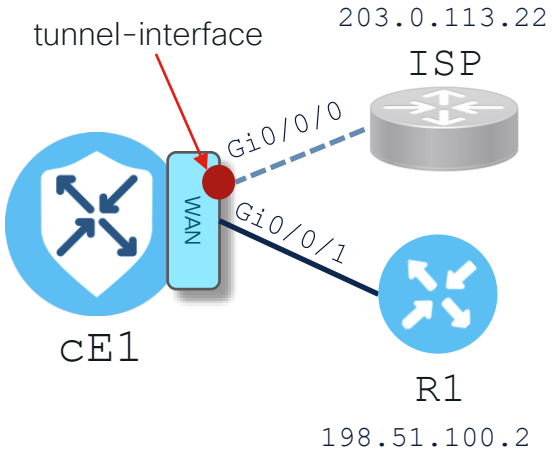
Even if we allow all in the implicit-acl:

```
cE1(config-interface-GigabitEthernet0/0/0)# tunnel-interface
cE1(config-tunnel-interface)# allow-service all
```

Or configure explicit IOS-ACL:

```
cE1(config)# ip access-list extended ALLOW_FWD
cE1(config-ext-nacl)# 10 permit ip any any
cE1(config-ext-nacl)# interface GigabitEthernet0/0/0
cE1(config-if)# ip access-group ALLOW_FWD in
cE1(config-if)# commit
```

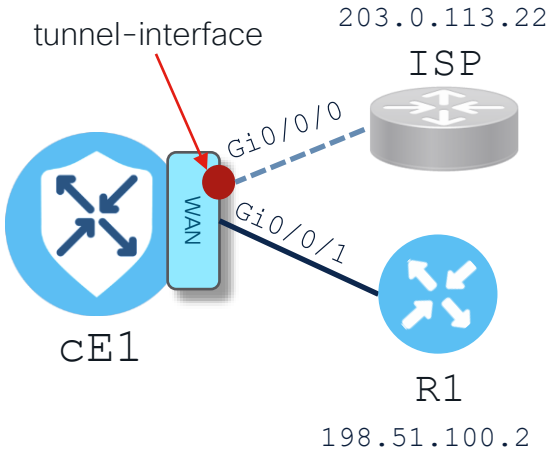
Result is the same, “SdwanImplicitAclDrop”



Case 2 ½. Forwarding in global routing table (GRT/vpn 0) is broken

Solution 1. Configure explicit SD-WAN ACL (localized policy), example:

```
policy
 access-list ALLOW_FWD
  sequence 10
   match
    destination-ip 198.51.100.0/24
   !
   action accept
   !
 !
!
sdwan
 interface GigabitEthernet0/0/0
  access-list ALLOW_FWD in
 exit
!
```



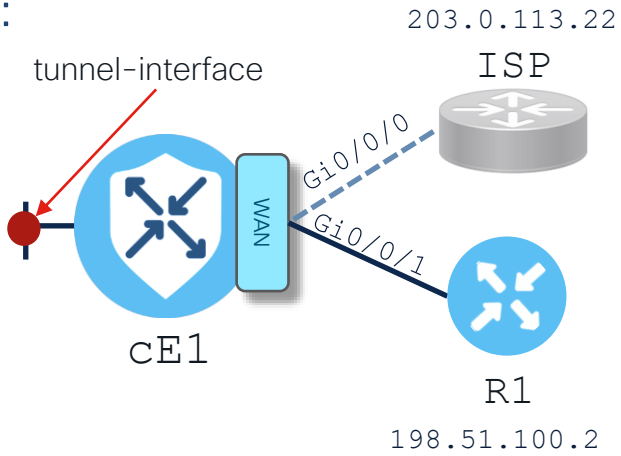
Mind typical confusion – SD-WAN explicit/implicit ACL vs IOS-XE ACL

* `allow-service all` didn't help because transit traffic is not in the list of "services"

Case 2 ½. Forwarding in global routing table (GRT/vpn 0) is broken

Solution 2. Move tunnel interface (TLOC) to Loopback:

```
interface Loopback0
 no shutdown
 ip address 198.51.100.255 255.255.255.255
exit
interface Tunnel1
 ip unnumbered Loopback0
 ipv6 unnumbered Loopback0
 tunnel source Loopback0
exit
sdwan
 interface Loopback0
  tunnel-interface
   encapsulation ipsec
   color biz-internet
   bind GigabitEthernet0/0/0
  exit
 exit
 interface GigabitEthernet0/0/0
  no tunnel-interface
 exit
```



"SdwanImplicitAclDrop" is not seen anymore, FWD only:

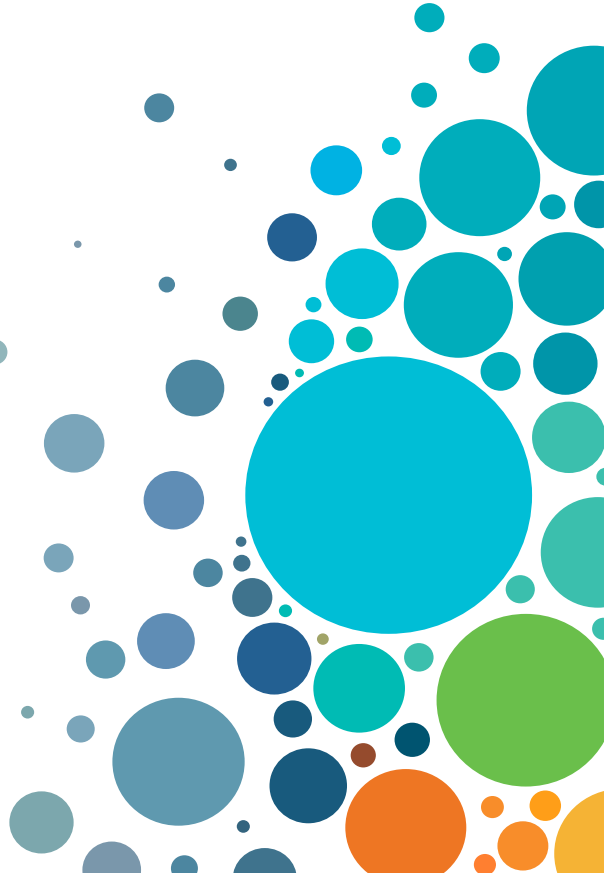
```
R1#ping 203.0.113.22 source 198.51.100.2
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 203.0.113.22, timeout is 2 seconds:
Packet sent with a source address of 198.51.100.2
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/1/1 ms
```

*May not work in 17.6.1, 17.6.2

OMP Features Cases

Overlay routing issues

Case 3. vSmart does not advertising any OMP routes



Case 3. Why my vSmart does not advertise any routes?

Symptoms. In the OMP table we can see **only** locally-originated routes (0.0.0.0):

```
BR1#show sdwan omp routes | count C,I,R
Number of lines which match regexp = 0
BR1#show sdwan omp routes | b PATH
                                          PATH                        ATTRIBUTE
VPN     PREFIX                FROM PEER   ID     LABEL    STATUS      TYPE       TLOC IP    COLOR            ENCAP  PREFERENCE
-------------------------------------------------------------------------------------------------------------------------------
1       192.168.1.0/24        0.0.0.0     66     1002     C,Red,R     installed  10.0.0.1   mpls             ipsec  -
                              0.0.0.0     68     1002     C,Red,R     installed  10.0.0.1   biz-internet     ipsec  -
```

Control Connections (CC) and Peering with vSmart is up, but nothing received:

```
BR1#show sdwan control connections | i vsmart
vsmart  dtls 10.0.0.100  1      1      192.168.20.213  12346 192.168.20.213   12346 biz-internet    No     up     0:00:11:49  10
vsmart  dtls 10.0.0.100  1      1      192.168.20.213  12346 192.168.20.213   12346 mpls            No     up     0:00:11:49  10

BR1#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
-----------------------------------------------------------------------------------
10.0.0.100    vsmart    1         1         1         up       0:00:02:35       0/0/2
```

Case 3. Why my vSmart does not advertise any routes?

Form a theory. What would we check further?

Facts:

- It's not underlay transport or control connections issue, OMP peering is up and stable
- OMP has zero-line config requirements for basic operations
- There is no any policy applied anywhere to filter routing information

Case 3. Why my vSmart does not advertise any routes?

Next let's check on vSmart, peer is "up", and we got 2 routes:

```
vsmart1# show omp peers 10.0.0.1
R -> routes received
I -> routes installed
S -> routes sent

                        DOMAIN    OVERLAY   SITE
PEER       TYPE    ID     ID        ID        STATE    UPTIME           R/I/S
-------------------------------------------------------------------------------
10.0.0.1   vedge   1      1         101       up       0:00:25:16       2/0/0
```

But they are rejected:

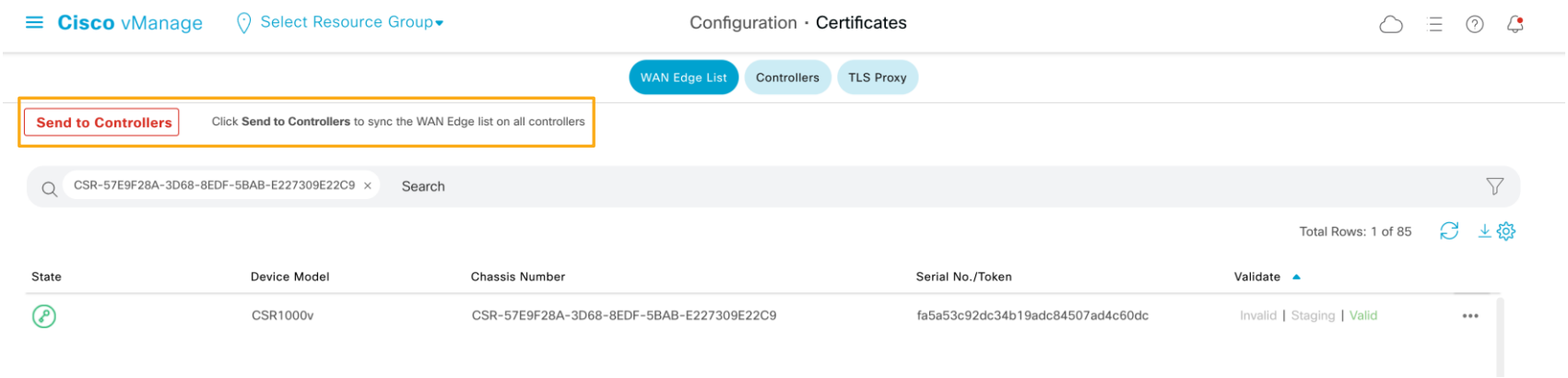
```
vsmart1# show omp routes received | include 10.0.0.1
2      192.168.1.0/24      10.0.0.1   66     1003     Rej,Inv,U,Stg installed  10.0.0.1   mpls            ipsec  -
                           10.0.0.1   68     1003     Rej,Inv,U,Stg installed  10.0.0.1   biz-internet    ipsec  -
```

And that router is in "staging" mode:

```
vsmart1# show control valid-vedges | i CSR-57E9F28A-3D68-8EDF-5BAB-E227309E22C9
CSR-57E9F28A-3D68-8EDF-5BAB-E227309E22C9 B6159EFC                          staging   CALO - 100589  N/A
```

Case 3. Why my vSmart does not advertise any routes?

Let's check on vManage, certificate is Valid, but…



Operator forgot to click "Send to Controllers" after onboarding

Case 3 ½ . vSmart does not advertising any OMP routes

Case 3 ½. Why my vSmart does not advertise any routes?

In the OMP table we don't see any routes:

```
BR1#show sdwan omp routes | count C,I,R
Number of lines which match regexp = 0
BR1#show sdwan omp routes

BR1#
```

Control connections (CC) with vSmart is up but no OMP peering:

```
BR1#show sdwan control connections | i vsmart
vsmart  dtls 10.0.0.100  1      1      192.168.20.213  12346 192.168.20.213   12346 biz-internet    No     up      0:00:11:49  10
vsmart  dtls 10.0.0.100  1      1      192.168.20.213  12346 192.168.20.213   12346 mpls            No     up      0:00:11:49  10

BR1#show sdwan omp peers

BR1#
```

Case 3 ½. Why my vSmart does not advertise any routes?

Form a theory. What would we check further?

Facts

- It's not transport or control connections issue as control connections are up and stable
- OMP has zero-line config requirements for basic operations, but OMP peering is not up

Case 3 ½. Why my vSmart does not advertise any routes?

Let's check OMP status:

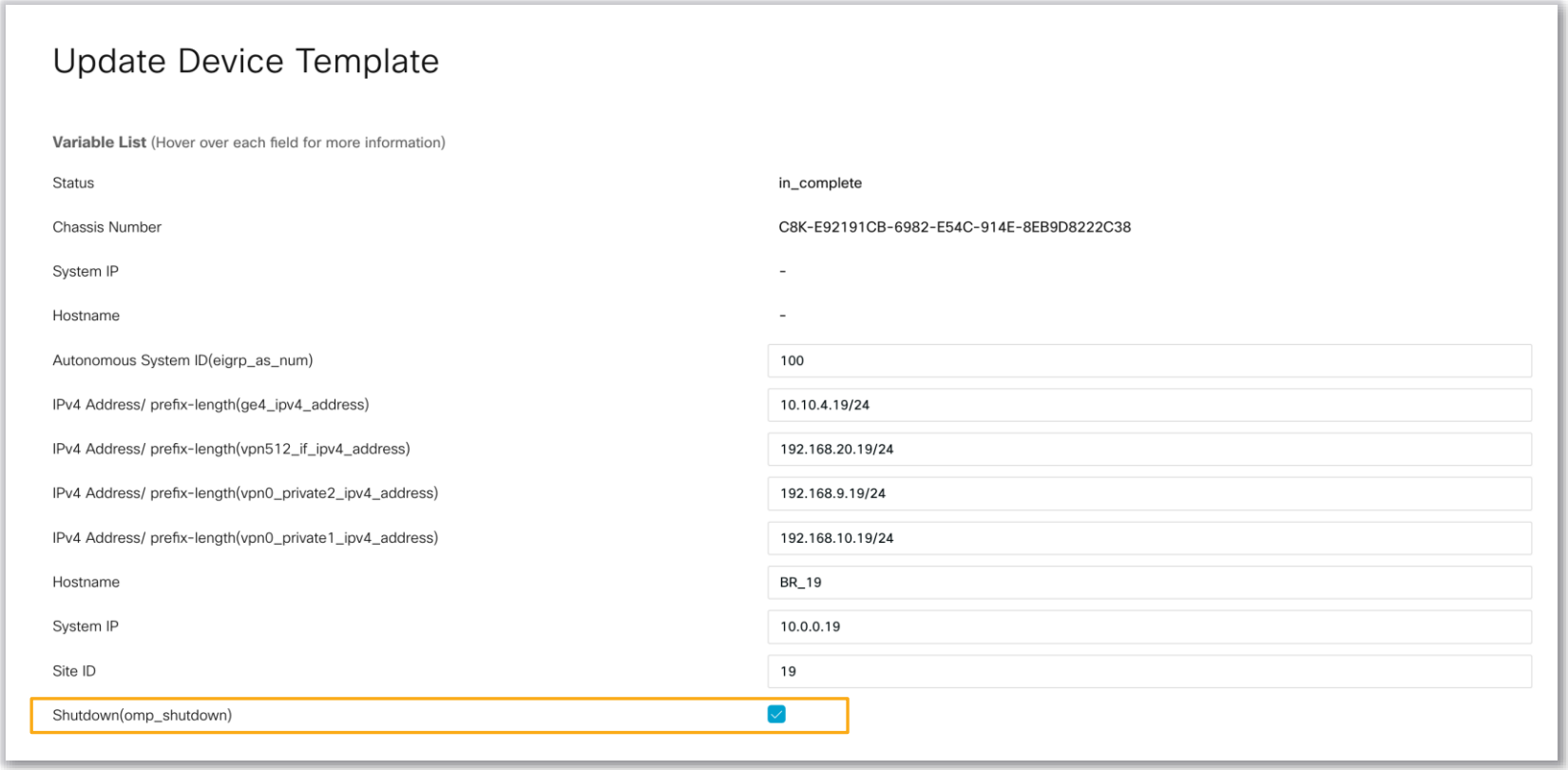
```
BR1#show sdwan omp summary
oper-state                DOWN
admin-state               DOWN
personality               vedge
omp-uptime                5:01:28:29
routes-received
routes-installed
routes-sent
tlocs-received
tlocs-installed
tlocs-sent
services-received
services-installed
services-sent
mcast-routes-received
mcast-routes-installed
mcast-routes-sent
hello-sent
hello-received
handshake-sent
handshake-received
alert-sent
alert-received
inform-sent
inform-received
update-sent
update-received
policy-sent
policy-received
total-packets-sent
total-packets-received
vsmart-peers
```

Case 3 ½. Why my vSmart does not advertise any routes?

Admin state “DOWN” means the same as for an interface status:

```
BR1#show sdwan running-config | sec omp
omp
  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
   advertise connected
   advertise static
  !
  address-family ipv6
   advertise connected
   advertise static
  !
```

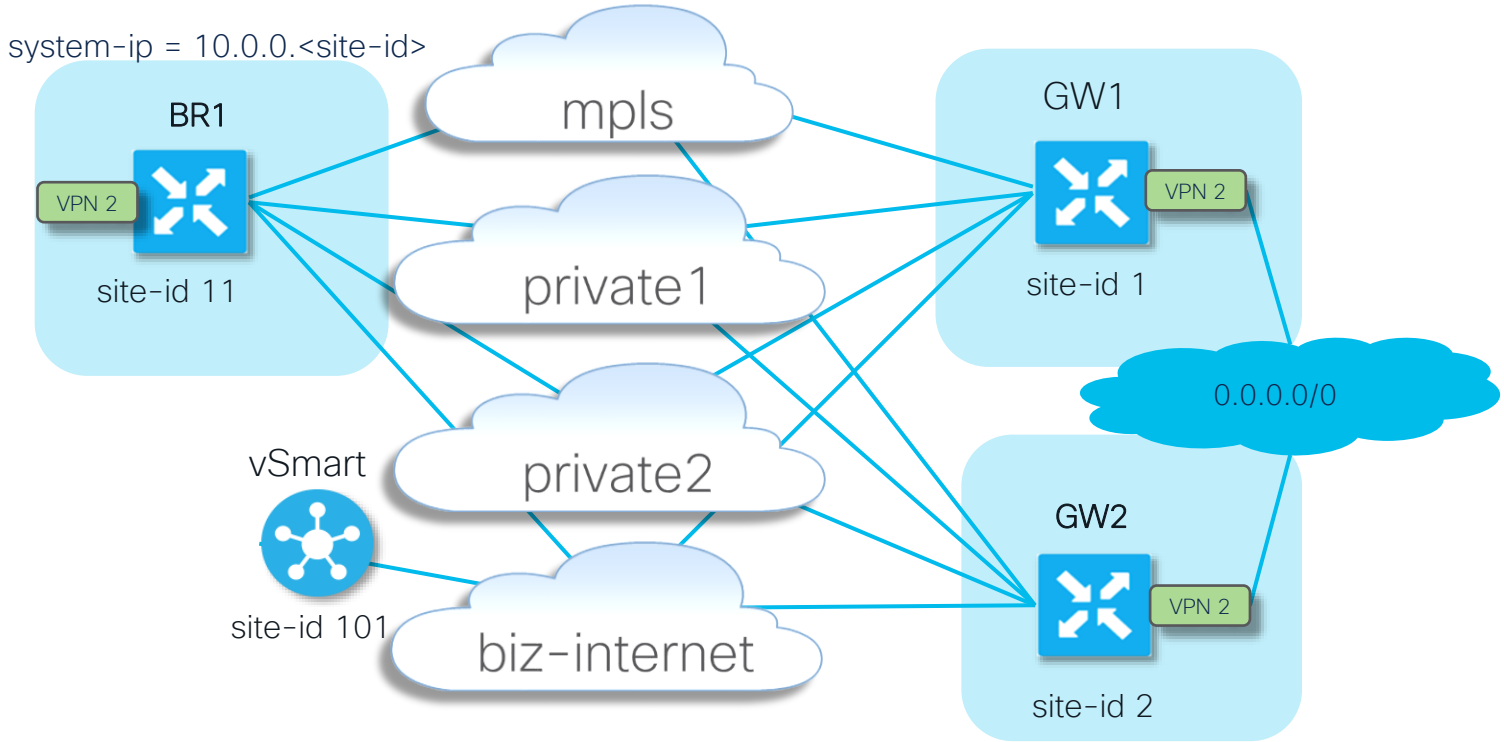
Case 3 ½. Why my vSmart does not advertise any routes?



Operator confused omp shutdown/no shutdown checkmark

Case 4. Traffic is not load-balanced over ECMP paths

Case 4. Traffic is not load-balanced over ECMP (active-active hubs redundancy)



Case 4. Traffic is not load-balanced over ECMP (active-active hubs redundancy)

Symptoms. Branch router R1 installs default route only via GW1:

```
BR1#show ip route vrf 2 | begin Gateway
Gateway of last resort is 10.0.0.1 to network 0.0.0.0

m*    0.0.0.0/0 [251/0] via 10.0.0.1, 00:08:30, sdwan_system_ip
```

This is because BR1 receives only 4 paths 0.0.0.0/0 and all resolvable via the same TLOC IP 10.0.0.1

```
BR1#show sdwan omp routes vpn 2 | begin PATH | exclude C,Red,R
                                    PATH                        ATTRIBUTE
VPN     PREFIX              FROM PEER     ID      LABEL   STATUS    TYPE       TLOC IP     COLOR            ENCAP  PREFERENCE
-------------------------------------------------------------------------------------------------------------------------
2       0.0.0.0/0           10.0.0.101    61614   1003    C,I,R     installed  10.0.0.1    mpls             ipsec  -
                            10.0.0.101    61615   1003    C,I,R     installed  10.0.0.1    biz-internet     ipsec  -
                            10.0.0.101    61616   1003    C,I,R     installed  10.0.0.1    private1         ipsec  -
                            10.0.0.101    61617   1003    C,I,R     installed  10.0.0.1    private2         ipsec  -
```

4 paths

Case 4. Traffic is not load-balanced over ECMP (active-active hubs redundancy)

At the same time vSmart has all 8 routes (4 routes for each TLOC color from 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           10.0.0.1    66      1003     C,R        installed   10.0.0.1    mpls             ipsec  -
                            10.0.0.1    68      1003     C,R        installed   10.0.0.1    biz-internet     ipsec  -
                            10.0.0.1    81      1003     C,R        installed   10.0.0.1    private1         ipsec  -
                            10.0.0.1    82      1003     C,R        installed   10.0.0.1    private2         ipsec  -
                            10.0.0.2    66      1003     C,R        installed   10.0.0.2    mpls             ipsec  -
                            10.0.0.2    68      1003     C,R        installed   10.0.0.2    biz-internet     ipsec  -
                            10.0.0.2    81      1003     C,R        installed   10.0.0.2    private1         ipsec  -
                            10.0.0.2    82      1003     C,R        installed   10.0.0.2    private2         ipsec  -
```

8 paths

Case 4. Traffic is not load-balanced over ECMP (active-active hubs redundancy)

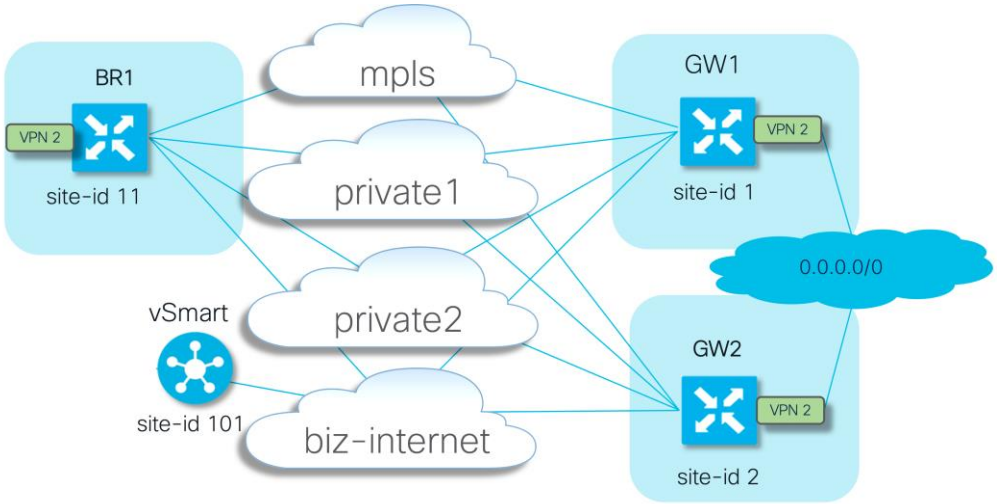
- Result: If default route from GW1 lost, spoke router installs route from GW2, hence there is no active-active redundancy and rather active-standby with GW1 acting as a primary router.
- We can also check which egress path is taken for specific traffic flow:

```
BR1#show sdwan policy service-path vpn 2  interface Loopback2 source-ip 192.168.1.1 dest-ip 192.168.12.1 protocol 6 source-port 53453
dest-port 22 dscp 48 app ssh
Next Hop: IPsec
  Source: 192.168.9.3 12347 Destination: 192.168.10.1 12427 Local Color: biz-internet Remote Color: mpls Remote System IP: 10.0.0.1
```

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

```
BR1#show sdwan policy service-path vpn 2  interface Loopback2 source-ip 192.168.1.1 dest-ip 192.168.12.1 protocol 6 source-port 53453
dest-port 22 dscp 48 app ssh all
Number of possible next hops: 4
Next Hop: IPsec
  Source: 192.168.9.3 12347 Destination: 192.168.10.1 12427 Local Color: biz-internet Remote Color: mpls Remote System IP: 10.0.0.1
Next Hop: IPsec
  Source: 192.168.8.3 12367 Destination: 192.168.8.1 12407 Local Color: private2 Remote Color: private2 Remote System IP: 10.0.0.1
Next Hop: IPsec
  Source: 192.168.7.3 12367 Destination: 192.168.7..1 12407 Local Color: private1 Remote Color: private1 Remote System IP: 10.0.0.11
Next Hop: IPsec
  Source: 192.168.9.3 12347 Destination: 192.168.9.1 12387 Local Color: biz-internet Remote Color: biz-internet Remote System IP:
10.0.0.1
```

Case 4. Traffic is not load-balanced over ECMP (active-active hub redundancy)



Form a theory. What would we check further?

Facts:

- vSmart is a control plane “brain” and route distribution control/filtering/enforcement point
- There is no centralized control policy applied on vSmart to filter something in our case
- OMP has zero-line config requirements for basic operations (and hence has some default settings)

Case 4. Traffic is not load-balanced over ECMP (active-active hub redundancy)

If you check what exactly vSmart advertises (unimportant attributes excluded):

```
vsmart1# show omp routes vpn 2 0.0.0.0/0 detail | nomore | exclude not\ set | b ADVERTISED\ TO: | b "peer    10.0.0.1" | exclude
label|path-id|overlay|origin
peer    10.0.0.1
    Attributes:
     originator       10.0.0.1
     tloc             10.0.0.1, private2, ipsec
     site-id          1
    Attributes:
     originator       10.0.0.1
     tloc             10.0.0.1, mpls, ipsec
     site-id          1
    Attributes:
     originator       10.0.0.1
     tloc             10.0.0.1, private1, ipsec
     site-id          11
    Attributes:
     originator       10.0.0.1
     tloc             10.0.0.1, biz-internet, ipsec
     site-id          1
```

4 total

You will find only 4 routes advertised toward branch (BR1) and all of them are from GW1 only

Case 4. Traffic is not load-balanced over ECMP (active-active hub redundancy). Solution?

As you know or remember already, by default vSmart sends only 4 routes because of **`send-path-limit`**. Let's fix this:

```
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#
```

Case 4. Traffic is not load-balanced over ECMP (active-active hub redundancy). Solution?

Then, all 8 routes will be advertised by vSmart and received on the branch router:

```
BR1#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           10.0.0.101  61626   1003    C,I,R     installed 10.0.0.1    mpls             ipsec  -
                            10.0.0.101  61627   1003    C,I,R     installed 10.0.0.1    biz-internet     ipsec  -
                            10.0.0.101  61628   1003    C,I,R     installed 10.0.0.1    private1         ipsec  -
                            10.0.0.101  61629   1003    C,I,R     installed 10.0.0.1    private2         ipsec  -
                            10.0.0.101  61637   1003    C,R       installed 10.0.0.2    mpls             ipsec  -
                            10.0.0.101  61638   1003    C,R       installed 10.0.0.2    biz-internet     ipsec  -
                            10.0.0.101  61639   1003    C,R       installed 10.0.0.2    private1         ipsec  -
                            10.0.0.101  61640   1003    C,R       installed 10.0.0.2    private2         ipsec  -
```

8 total

But still branch routers install routes only via GW1:

```
BR1#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 10.0.0.1 on sdwan_system_ip, 01:11:26 ago
  Routing Descriptor Blocks:
  * 10.0.0.1 (default), from 10.0.0.1, 01:11:26 ago, via sdwan_system_ip
      Route metric is 0, traffic share count is 1
```

`show sdwan policy service-path` will confirm the same and hence the output is not provided for brevity

Case 4. Traffic is not load-balanced over ECMP (active-active hub redundancy). Final solution.

Reason - there are also default settings.
WAN Edge router installs only first 4 equal paths into the routing table because of **`ecmp-limit`**.
Let's change this:

```
BR1#config-t

admin connected from 127.0.0.1 using console on ce3
R1(config)# sdwan
R1(config-sdwan)# omp
R1(config-omp)# ecmp-limit 8
R1(config-omp)# commit
Commit complete.
```

`show ip route` confirms both routes via both hubs are installed now:

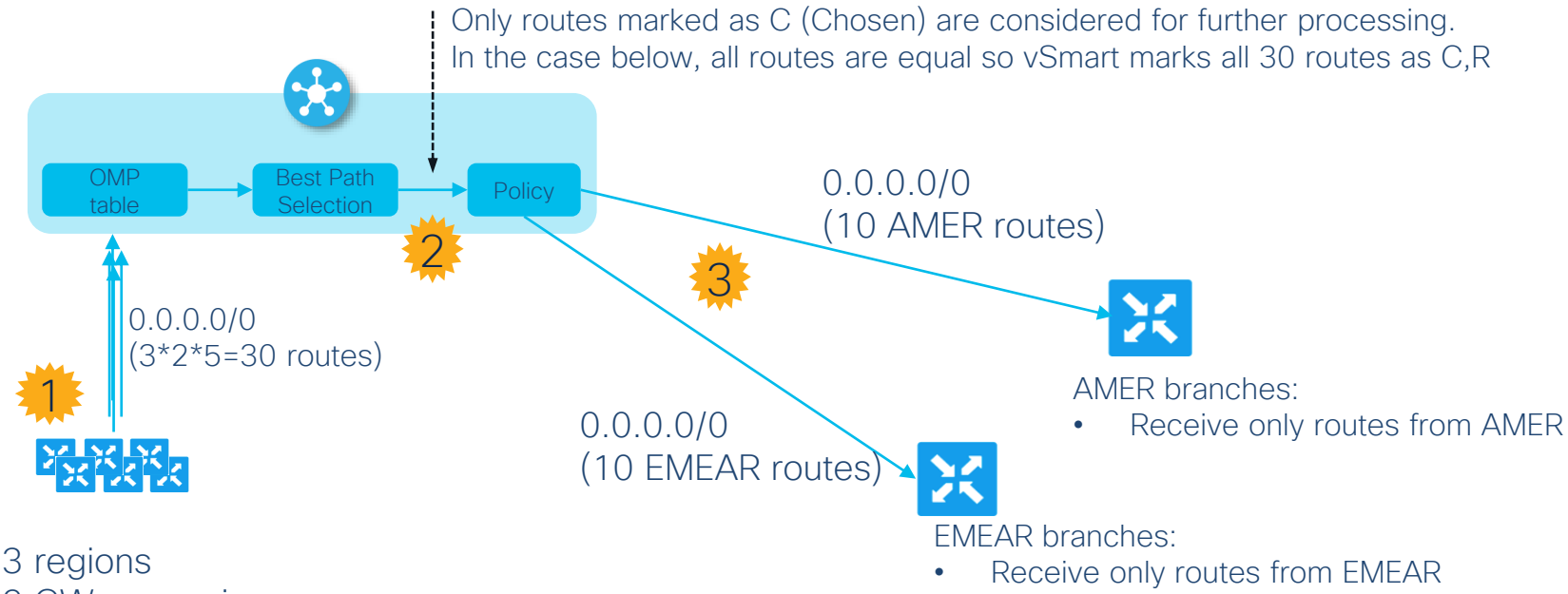
```
BR1#sh ip route vrf 2 | b Gateway
Gateway of last resort is 10.0.0.2 to network 0.0.0.0

m*    0.0.0.0/0 [251/0] via 10.0.0.2, 00:00:37, sdwan_system_ip
                [251/0] via 10.0.0.1, 00:00:37, sdwan_system_ip
```

Conclusion: mind OMP configuration defaults

Case 5. OMP Path selection and global scalability

Case 5. OMP Path selection and global scalability

Case 5. OMP Path selection and global scalability. Objective.



- 3 regions
- 2 GW per region
- each GW with 5 TLOCs
- each advertising 0.0.0.0/0

Case 5. OMP Path selection and global scalability. Output from vSmart.

VPN	PREFIX	FROM PEER	ID	LABEL	STATUS	TYPE	TLOC IP	COLOR	ENCAP	PREFERENCE
1	0.0.0.0/0	10.0.0.10	81	1003	C,R	installed	10.0.0.10	private1	ipsec	-
		10.0.0.10	82	1003	C,R	installed	10.0.0.10	private2	ipsec	-
		10.0.0.10	83	1003	C,R	installed	10.0.0.10	private3	ipsec	-
		10.0.0.10	84	1003	C,R	installed	10.0.0.10	private4	ipsec	-
		10.0.0.10	85	1003	C,R	installed	10.0.0.10	private5	ipsec	-
		10.0.0.11	81	1003	C,R	installed	10.0.0.11	private1	ipsec	-
		10.0.0.11	82	1003	C,R	installed	10.0.0.11	private2	ipsec	-
		10.0.0.11	83	1003	C,R	installed	10.0.0.11	private3	ipsec	-
		10.0.0.11	84	1003	C,R	installed	10.0.0.11	private4	ipsec	-
		10.0.0.11	85	1003	C,R	installed	10.0.0.11	private5	ipsec	-
		10.0.0.12	81	1003	C,R	installed	10.0.0.12	private1	ipsec	-
		10.0.0.12	82	1003	C,R	installed	10.0.0.12	private2	ipsec	-
		10.0.0.12	83	1003	C,R	installed	10.0.0.12	private3	ipsec	-
		10.0.0.12	84	1003	C,R	installed	10.0.0.12	private4	ipsec	-
		10.0.0.12	85	1003	C,R	installed	10.0.0.12	private5	ipsec	-
		10.0.0.13	81	1003	C,R	installed	10.0.0.13	private1	ipsec	-
		10.0.0.13	82	1003	C,R	installed	10.0.0.13	private2	ipsec	-
		10.0.0.13	83	1003	C,R	installed	10.0.0.13	private3	ipsec	-
		10.0.0.13	84	1003	C,R	installed	10.0.0.13	private4	ipsec	-
		10.0.0.13	85	1003	C,R	installed	10.0.0.13	private5	ipsec	-
		10.0.0.14	81	1003	C,R	installed	10.0.0.14	private1	ipsec	-
		10.0.0.14	82	1003	C,R	installed	10.0.0.14	private2	ipsec	-
		10.0.0.14	83	1003	C,R	installed	10.0.0.14	private3	ipsec	-
		10.0.0.14	84	1003	C,R	installed	10.0.0.14	private4	ipsec	-
		10.0.0.14	85	1003	C,R	installed	10.0.0.14	private5	ipsec	-
		10.0.0.15	81	1003	C,R	installed	10.0.0.15	private1	ipsec	-
		10.0.0.15	82	1003	C,R	installed	10.0.0.15	private2	ipsec	-
		10.0.0.15	83	1003	C,R	installed	10.0.0.15	private3	ipsec	-
		10.0.0.15	84	1003	C,R	installed	10.0.0.15	private4	ipsec	-
		10.0.0.15	85	1003	C,R	installed	10.0.0.15	private5	ipsec	-

EMEAR

APAC

AMER

Case 5. OMP Path selection and global scalability. Outputs from Edge router.

`omp ecmp-limit 16` configured to install all 16 equal paths into RIB.

On EMEAR branch router:

VPN	PREFIX	FROM PEER	ID	LABEL	STATUS	TYPE	TLOC IP	COLOR	ENCAP	PREFERENCE
1	0.0.0.0/0	10.0.0.100	1	1003	C,I,R	installed	10.0.0.10	private1	ipsec	-
		10.0.0.100	2	1003	C,I,R	installed	10.0.0.10	private2	ipsec	-
		10.0.0.100	3	1003	C,I,R	installed	10.0.0.10	private3	ipsec	-
		10.0.0.100	4	1003	C,I,R	installed	10.0.0.10	private4	ipsec	-
		10.0.0.100	5	1003	C,R	installed	10.0.0.10	private5	ipsec	-
		10.0.0.100	6	1003	C,R	installed	10.0.0.11	private1	ipsec	-
		10.0.0.100	7	1003	C,R	installed	10.0.0.11	private2	ipsec	-
		10.0.0.100	8	1003	C,R	installed	10.0.0.11	private3	ipsec	-
		10.0.0.100	9	1003	C,R	installed	10.0.0.11	private4	ipsec	-
		10.0.0.100	10	1003	C,R	installed	10.0.0.11	private5	ipsec	-

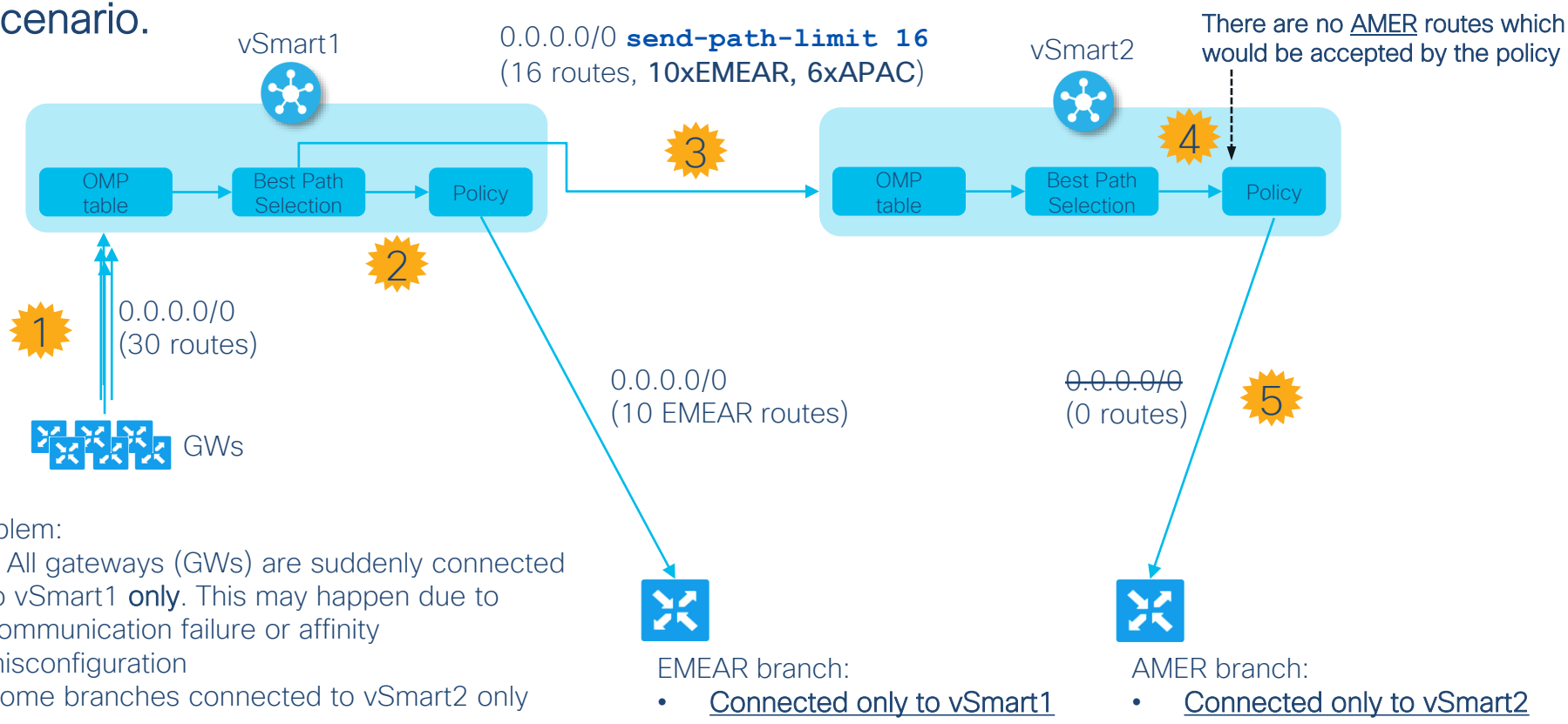
From EMEAR GW

On AMER branch router:

VPN	PREFIX	FROM PEER	ID	LABEL	STATUS	TYPE	TLOC IP	COLOR	ENCAP	PREFERENCE
1	0.0.0.0/0	10.0.0.100	1	1003	C,I,R	installed	10.0.0.14	private1	ipsec	-
		10.0.0.100	2	1003	C,I,R	installed	10.0.0.14	private2	ipsec	-
		10.0.0.100	3	1003	C,I,R	installed	10.0.0.14	private3	ipsec	-
		10.0.0.100	4	1003	C,I,R	installed	10.0.0.14	private4	ipsec	-
		10.0.0.100	5	1003	C,R	installed	10.0.0.14	private5	ipsec	-
		10.0.0.100	6	1003	C,R	installed	10.0.0.15	private1	ipsec	-
		10.0.0.100	7	1003	C,R	installed	10.0.0.15	private2	ipsec	-
		10.0.0.100	8	1003	C,R	installed	10.0.0.15	private3	ipsec	-
		10.0.0.100	9	1003	C,R	installed	10.0.0.15	private4	ipsec	-
		10.0.0.100	10	1003	C,R	installed	10.0.0.15	private5	ipsec	-

From AMER GW

Case 5. OMP Path selection and scalability. Multiple vSmarts – Failure scenario.



Problem:

- All gateways (GWs) are suddenly connected to vSmart1 **only**. This may happen due to communication failure or affinity misconfiguration
- Some branches connected to vSmart2 only

EMEAR branch:

- Connected only to vSmart1

AMER branch:

- Connected only to vSmart2

* vSmart1/2 can be also groups of vSmarts, e.g. group 1 and group 2.

Case 5. OMP Path selection and scalability. Solution.

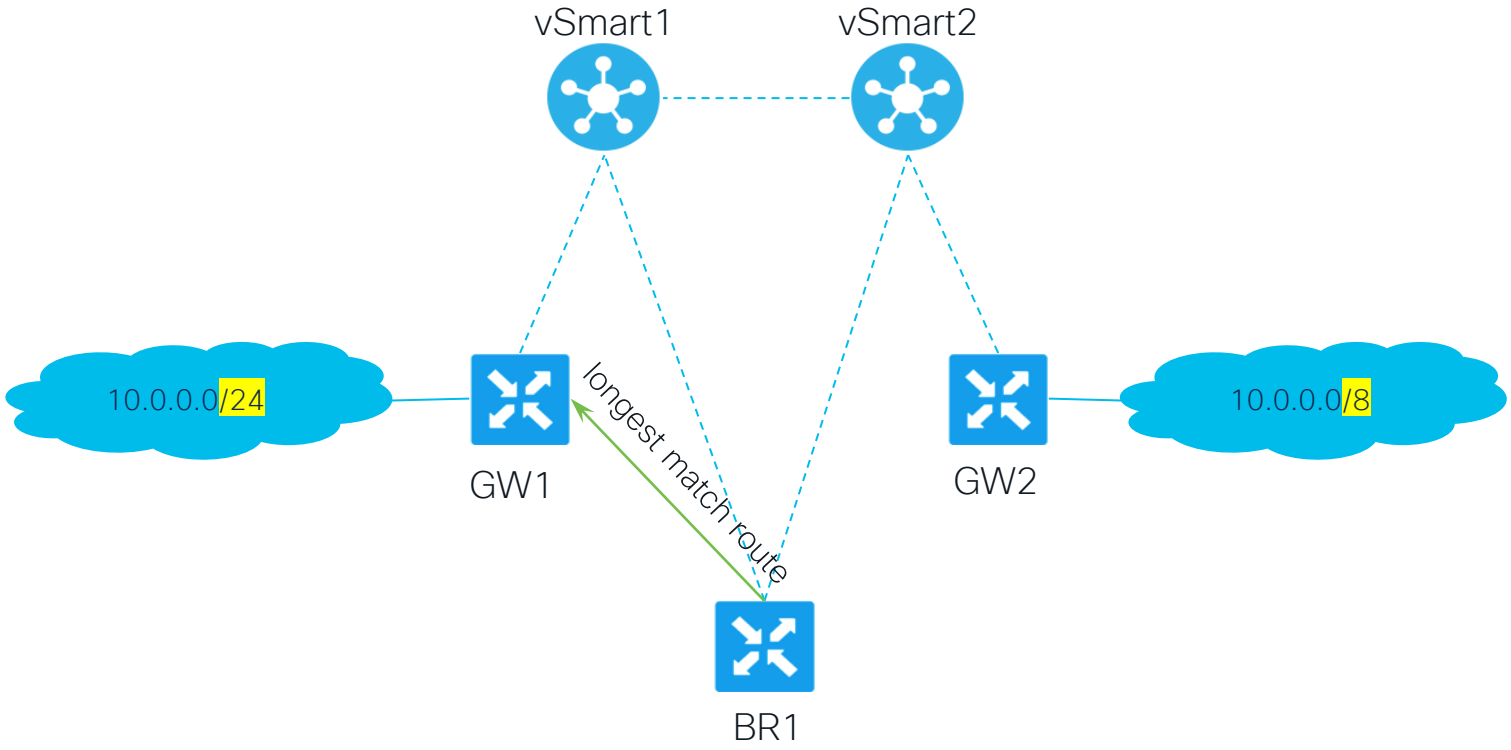
- Properly plan affinity groups and redundancy (i.e. always only 1 control plane “hop” between route source and destination)
- Increase **`max-control-connections`** (default is 2) to peer with all vSmarts. 🤔 Questionable, impact on scale
- Increase **`controller-send-path-limit`** (available starting from 20.5). Best option

```
omp
  controller-send-path-limit [4-128]
```

Case 6 vSmart double failure scenario

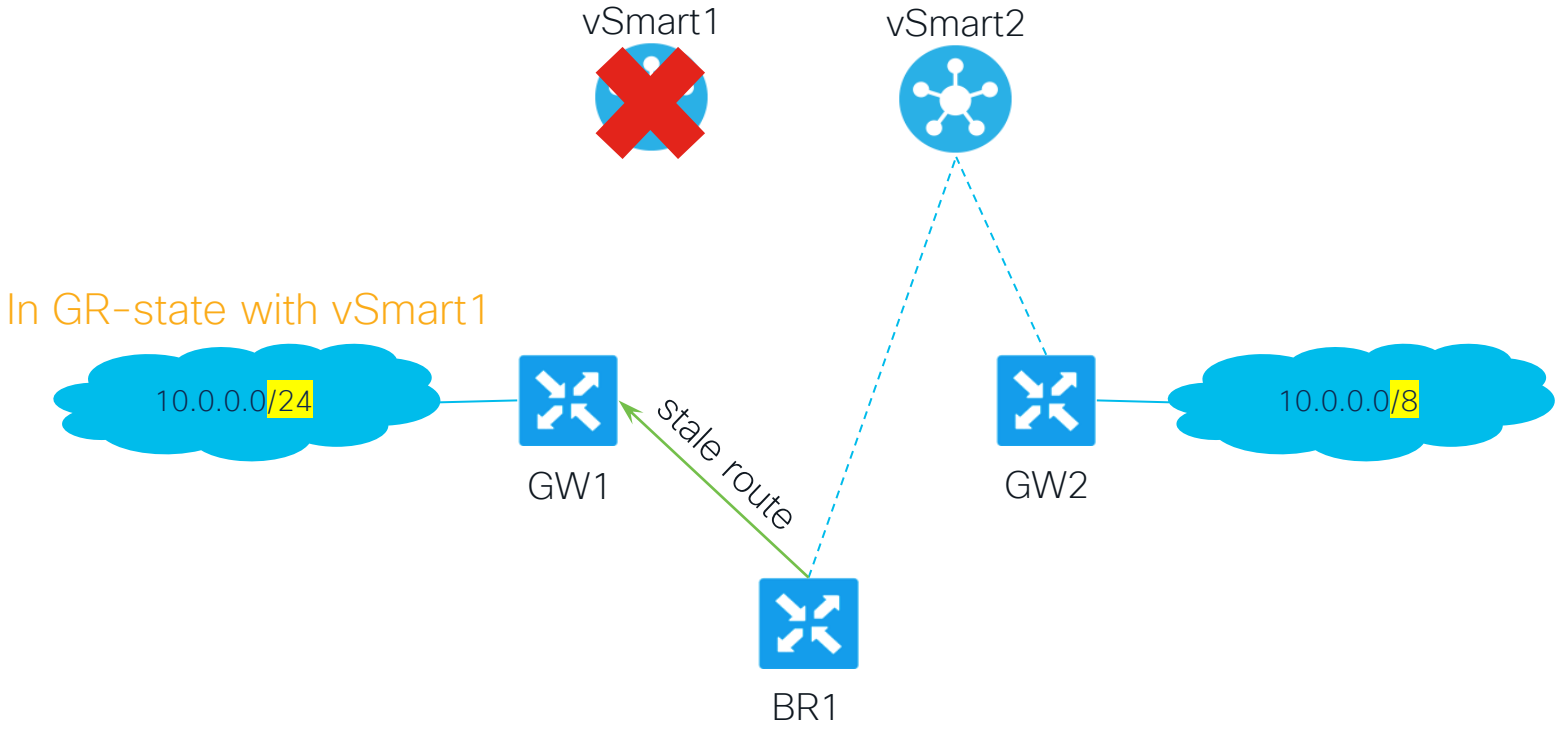
or why you may not want to enable graceful-restart

Case 6. vSmart double failure scenario



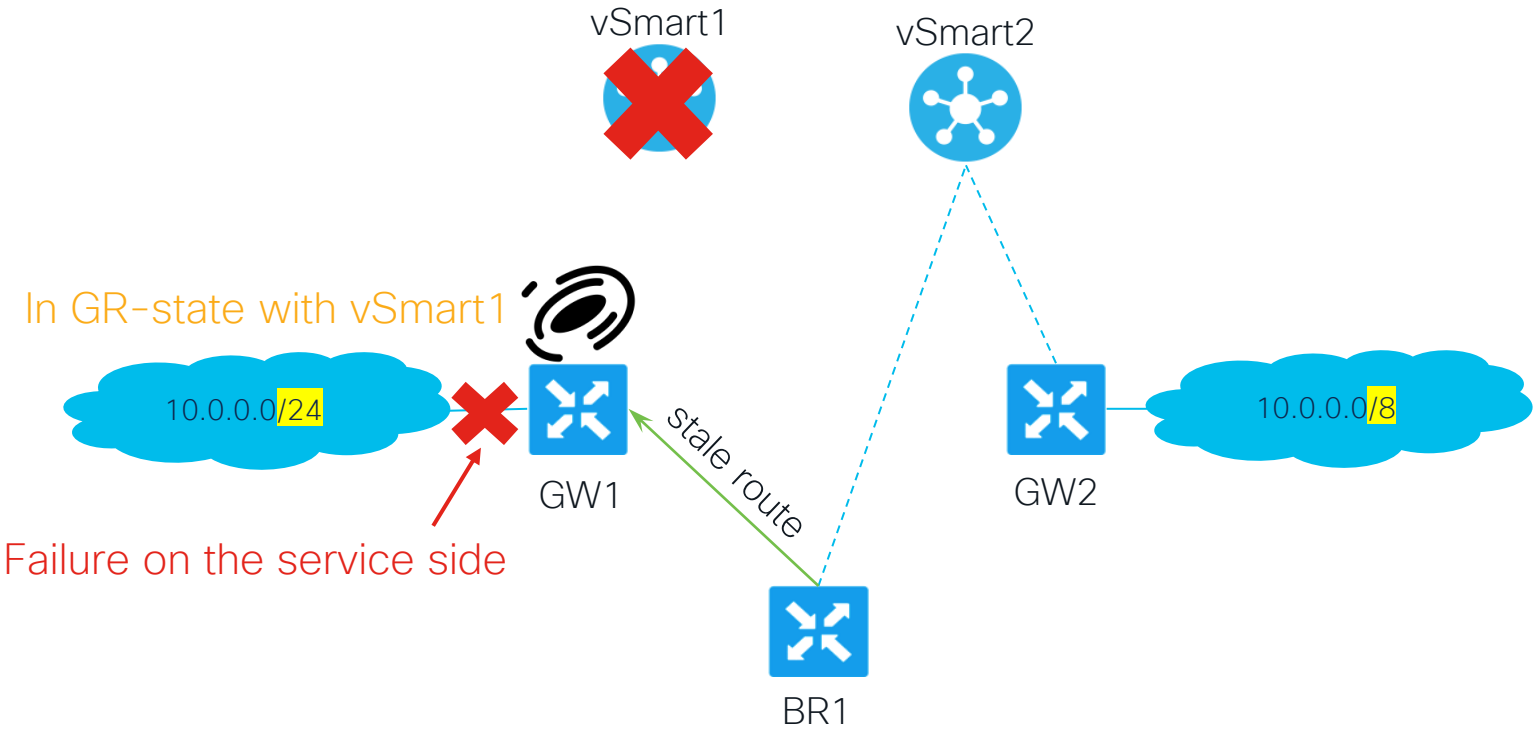
* vSmart1/2 can be also groups of vSmarts, e.g. group 1 and group 2.

Case 6. vSmart double failure scenario



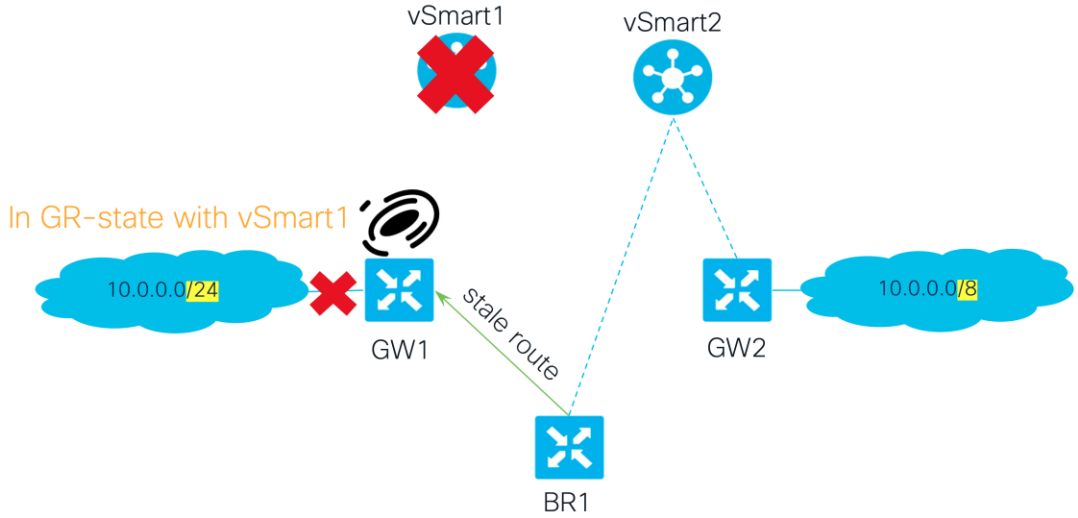
* BR1 prefers longest match stale route over less specific route via GW2

Case 6. vSmart double failure scenario



SD-WAN can't handle double failure scenarios with GR configured, this is expected

Case 6. vSmart double failure scenario. Solutions.

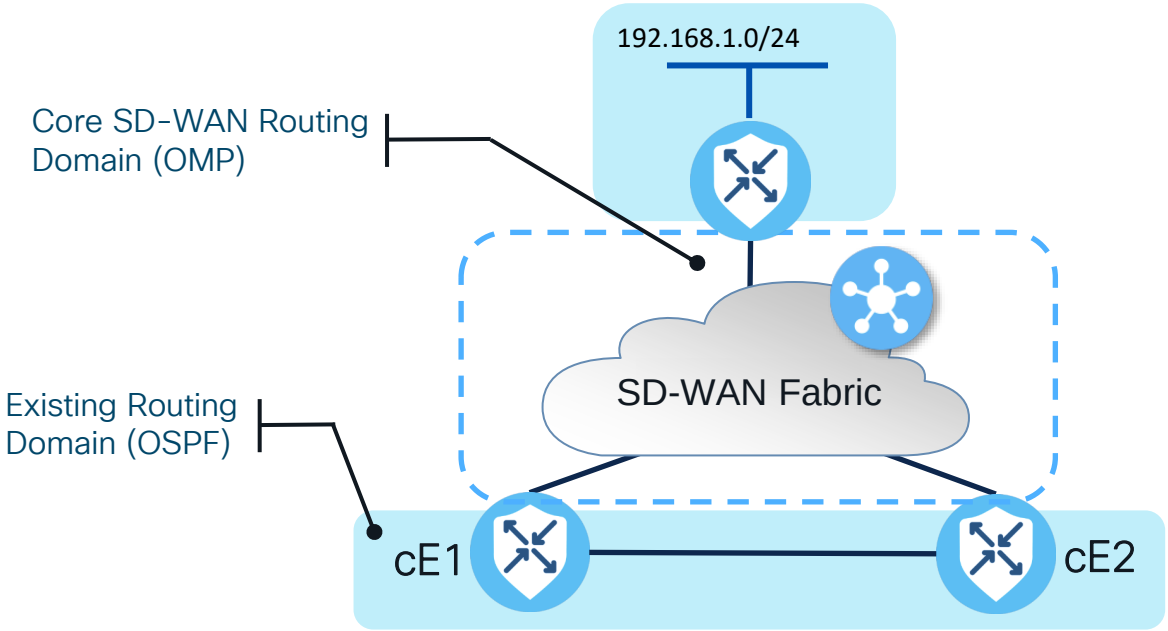


1. Be careful with summarization.
2. Ensure vSmart redundancy, geo-reservation and underlay paths diversity.
3. Properly plan contollers affinity if you use it, mind (2) also.
4. Don’t use GR 🤔 Questionable.

Service-side routing protocols and OMP

Case 7. OSPF and DN-bit in SD-WAN

Case 7. OSPF and DN-bit in SD-WAN

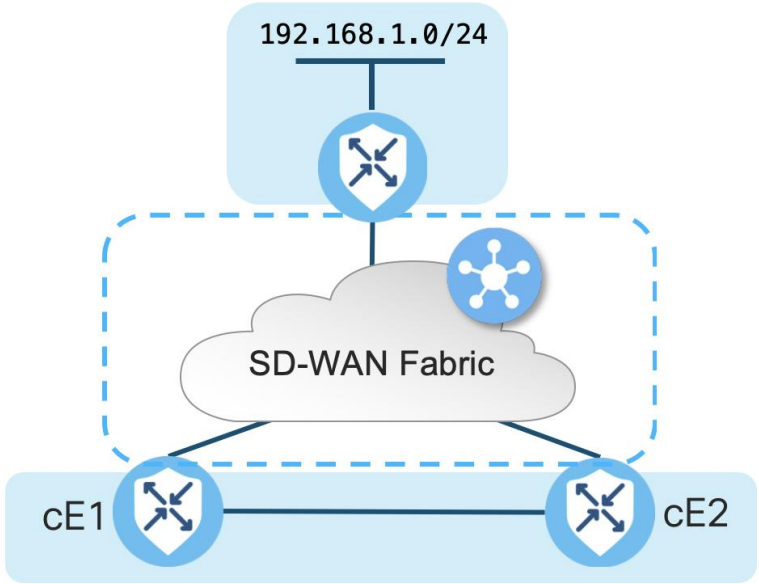


*This is not a break-fix troubleshooting case, rather demonstration of expected behavior

Case 7. OSPF and DN-bit in SD-WAN

No tricks, very simple config on all 3 routers (only relevant part):

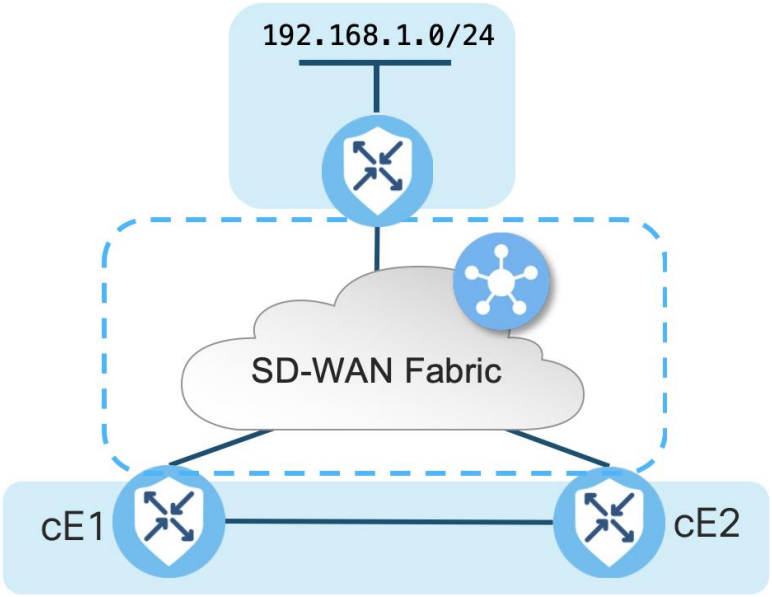
```
route-map omp2ospf permit 10
 set metric 1000
 set metric-type type-1
!
router ospf 2 vrf 2
 redistribute omp route-map omp2ospf
!
 omp
  no shutdown
  send-path-limit  4
  ecmp-limit       4
  graceful-restart
  no as-dot-notation
  timers
   holdtime               60
   advertisement-interval 1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



Case 7. OSPF and DN-bit in SD-WAN

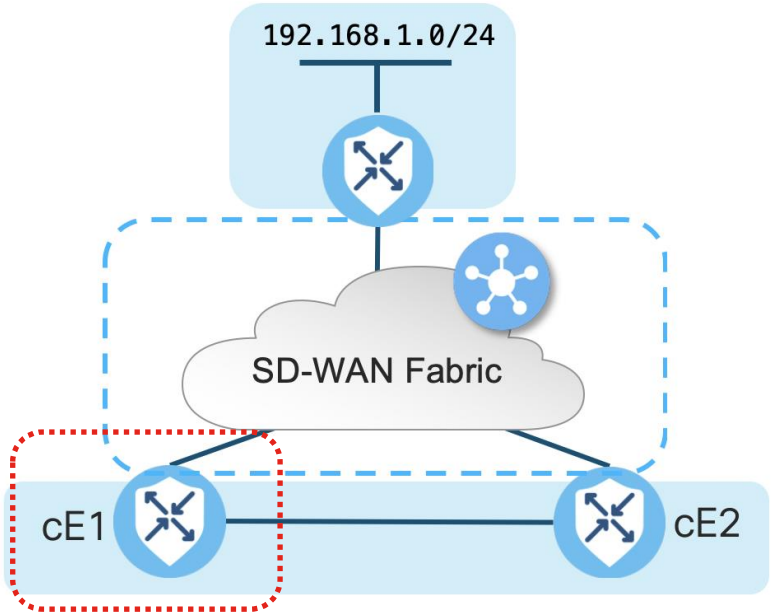
In a normal conditions, both cE1 and cE2 prefers OMP route to 192.168.1.0/24 vs OSPF:

```
cE1#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 10.0.0.3 00:03:00 ago
  Routing Descriptor Blocks:
  * 10.0.0.3 (default), from 10.0.0.3, 00:03:00 ago
      Route metric is 0, traffic share count is 1

cE2#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 10.0.0.3 00:04:13 ago
  Routing Descriptor Blocks:
  * 10.0.0.3 (default), from 10.0.0.3, 00:04:13 ago
      Route metric is 0, traffic share count is 1
```



Case 7. OSPF and DN-bit in SD-WAN

This is because we can see DN-bit set for LSAs generated by both routers

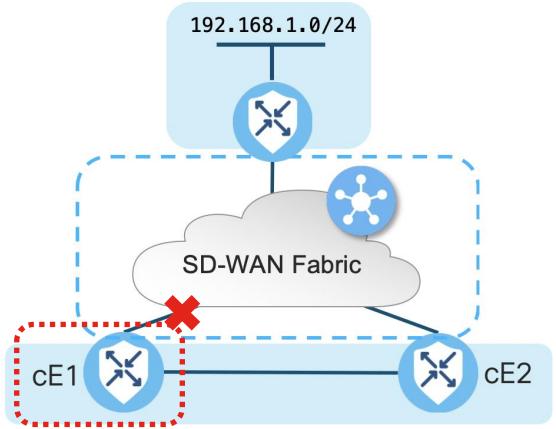



```
cE1#show ip ospf database external 192.168.1.0

            OSPF Router with ID (10.0.0.1) (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: 10.0.0.1
  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: 10.0.0.2
  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
```

Case 7. OSPF and DN-bit in SD-WAN

What if cE1 disconnected from the fabric?

```
Oct 11 12:53:58.777: %Cisco-SDWAN-Router-OMPD-3-ERRO-400002: R0/0: OMPD: vSmart peer 10.0.0.100 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

cE1#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
-------------------------------------------------------------------------------------
10.0.0.100    vsmart  1         1         3         init-in-gr                 26/1/0
```

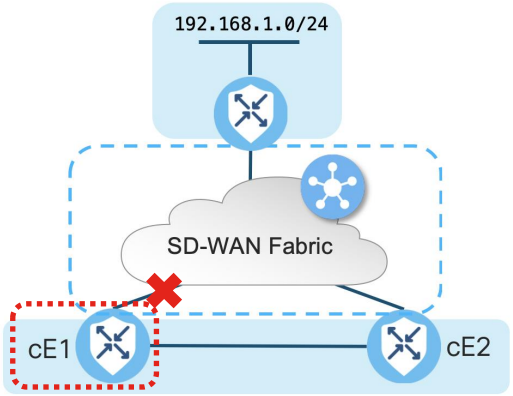


Case 7. OSPF and DN-bit in SD-WAN

cE1 marks the OMP route as stale (see OMP route state S), but keeps the route in the RIB installed by OMP protocol until **`graceful-restart-timer`** expired:

```
cE1#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            10.0.0.100
path-id         1076
label           1002
status          C,I,R,S
    Attributes:
     originator       10.0.0.3
     type             installed
     tloc             10.0.0.3, biz-internet, ipsec
     overlay-id        1
     site-id          201207
     origin-proto     connected
     origin-metric    0
```

```
cE1#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 10.0.0.3 00:23:35 ago
  Routing Descriptor Blocks:
  * 10.0.0.3 (default), from 10.0.0.3, 00:23:35 ago
      Route metric is 0, traffic share count is 1
```



Case 7. OSPF and DN-bit in SD-WAN

Once **`graceful-restart-timer`** timer expires, route to 192.168.1.0/24 will be still there.

```
cE1#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 10.0.0.2, 00:04:11 ago, via Vlan2807
      SDWAN Down
      Route metric is 1100, traffic share count is 1
```

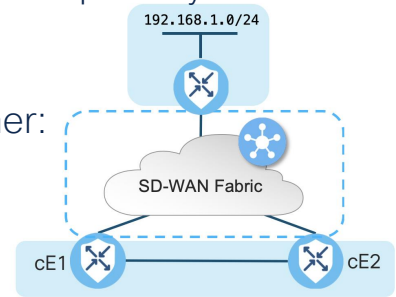
- OSPF route with AD 252 from LSA with DN-bit
- “SDWAN down” flag is set to the route

```
cE1#show ip ospf database external 192.168.1.0

            OSPF Router with ID (10.0.0.1) (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: 10.0.0.2
  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
```

Why OSPF route was installed into RIB while LSA has a DN-bit set?

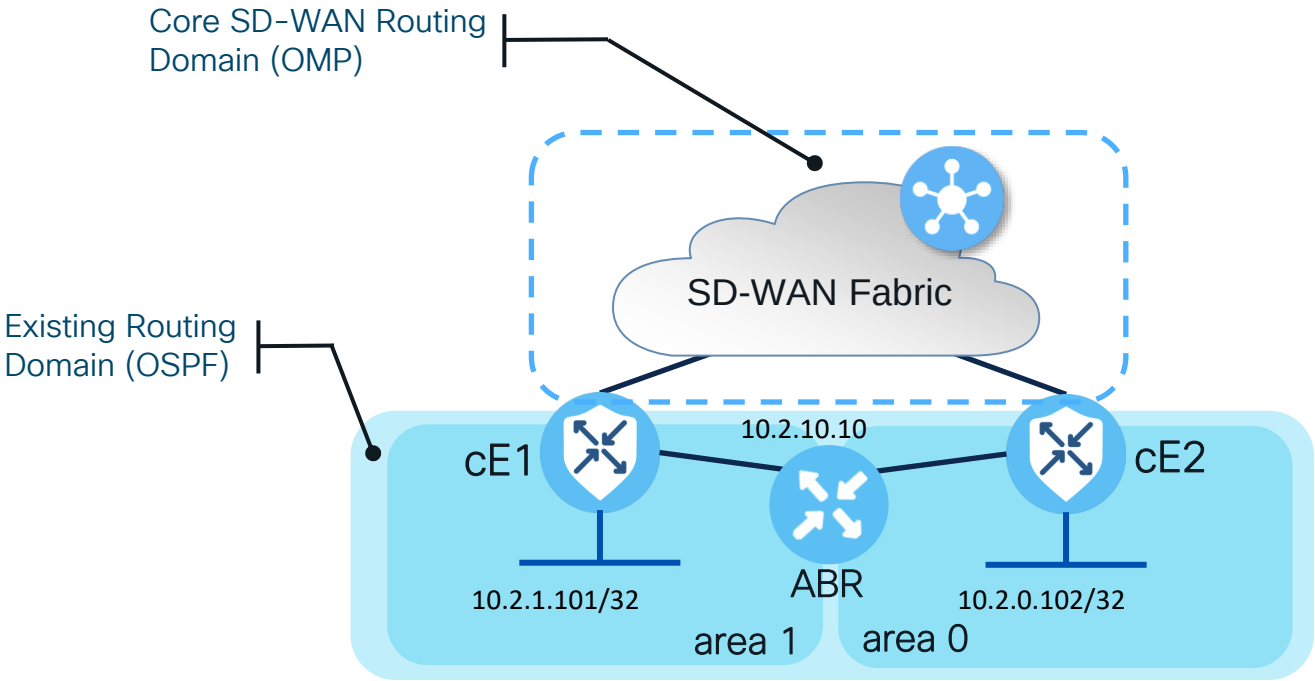
Case 7. Why OSPF routes with DN-bit installed into RIB?

- It is installed as OSPF External Type 1 route now despite the fact that corresponding OSPF LSA has a DN-bit set.
- administrative distance (AD) is always 1 unit more than the AD of OMP
- If OMP comes up, OMP route with AD 251 will instantly pre-empt OSPF route with AD 252
- expected behaviour to avoid traffic blackhole scenarios when one of the routers is partitioned from the SD-WAN overlay
- Why? Fool-proof design. Without this mechanism, blackhole might happen if service side traffic is still load-balanced via both routers e.g. because two static routes pointing to both routers or some routes pointing to only one router that is partitioned (e.g. still FHRP primary router for whatever reason)
- In case of ECMP (when cE1 is partitioned from fabric) egress traffic follows either:
  - LAN -> cE1 -> cE2 -> remote router -> 192.168.1.0/24
  - LAN -> cE2 -> remote router -> 192.168.1.0/24



Case 8. WAN Edge does not install route from type3 LSA from backbone area into the RIB

Case 8. WAN Edge does not install route from type3 LSA from area 0 into the RIB



Problem: Route from cE1 loopback is not getting installed into RIB of cE2, while route from cE2 loopback is installed into RIB of cE1 successfully

Case 8. WAN Edge does not install route from type3 LSA from area 0 into the RIB

No tricks, very simple config here:

```
ce1#show running-config vrf 2
vrf definition 2
 rd 1:2
 !
 address-family ipv4
  route-target export 1:2
  route-target import 1:2
 exit-address-family
 !
 address-family ipv6
 exit-address-family
!
interface GigabitEthernet7
 vrf forwarding 2
 ip address 192.168.70.101 255.255.255.0
 ip ospf network point-to-point
 ip ospf 2 area 1
!
interface Loopback1
 vrf forwarding 2
 ip address 10.2.1.101 255.255.255.255
 ip ospf 2 area 1
!
router ospf 2 vrf 2
!
end
```

```
ABR#show running-config
router ospf 2
 router-id 10.2.10.10
!
interface GigabitEthernet5
 ip address 192.168.70.10 255.255.255.0
 ip ospf network point-to-point
 ip ospf 2 area 1
!
interface GigabitEthernet6
 ip address 192.168.80.10 255.255.255.0
 ip ospf network point-to-point
 ip ospf 2 area 0
!
interface Loopback1
 ip address 10.2.0.10 255.255.255.255
 ip ospf 2 area 0
!
interface Loopback2
 ip address 10.2.1.10 255.255.255.255
 ip ospf 2 area 1
end
```

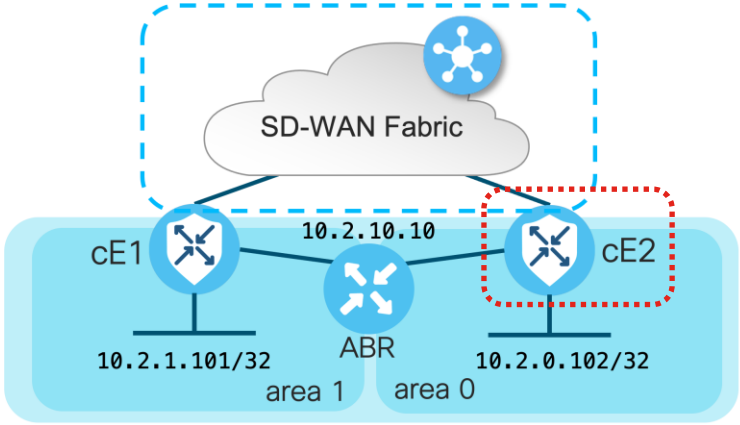
```
ce2#show running-config vrf 2
vrf definition 2
 rd 1:2
 !
 address-family ipv4
  route-target export 1:2
  route-target import 1:2
 exit-address-family
 !
 address-family ipv6
 exit-address-family
!
interface GigabitEthernet8
 vrf forwarding 2
 ip address 192.168.80.102 255.255.255.0
 ip ospf network point-to-point
 ip ospf 2 area 0
!
interface Loopback2
 vrf forwarding 2
 ip address 10.2.0.102 255.255.255.255
 ip ospf 2 area 0
!
router ospf 2 vrf 2
!
end
```

Case 8. WAN Edge does not install route from type3 LSA from area 0 into the RIB

On cE2, route for cE1 loopback (from area 1) is installed into the RIB:

```
cE2#show ip ospf database

            OSPF Router with ID (10.2.0.102) (Process ID 2)

                Router Link States (Area 0)

Link ID         ADV Router      Age         Seq#       Checksum Link count
10.2.0.102      10.2.0.102      563         0x800000A1 0x005603 3
10.2.10.10      10.2.10.10      118         0x800000B3 0x00C1AF 2

                Summary Net Link States (Area 0)

Link ID         ADV Router      Age         Seq#       Checksum
10.2.1.10       10.2.10.10      579         0x80000001 0x009C68
10.2.1.101      10.2.10.10      569         0x80000003 0x00A292
192.168.70.0    10.2.10.10      26          0x80000003 0x00EB7E

cE2#show ip ospf rib | b Codes
Codes: * - Best, > - Installed in global RIB

*>  10.2.0.10/32, Intra, cost 2, area 0
      via 192.168.80.10, GigabitEthernet8
*   10.2.0.102/32, Intra, cost 1, area 0, Connected
      via 10.2.0.102, Loopback2
*>  10.2.1.10/32, Inter, cost 2, area 0
      via 192.168.80.10, GigabitEthernet8
*>  10.2.1.101/32, Inter, cost 3, area 0
      via 192.168.80.10, GigabitEthernet8
*>  192.168.70.0/24, Inter, cost 2, area 0
      via 192.168.80.10, GigabitEthernet8
*   192.168.80.0/24, Intra, cost 1, area 0, Connected
      via 192.168.80.102, GigabitEthernet8
```

```
cE2#sh ip route vrf 2 ospf | b Gate
Gateway of last resort is not set

      10.0.0.0/32 is subnetted, 4 subnets
O        10.2.0.10 [110/2] via 192.168.80.10, 00:19:04,
GigabitEthernet8
O IA     10.2.1.10 [110/2] via 192.168.80.10, 00:19:04,
GigabitEthernet8
O IA     10.2.1.101 [110/3] via 192.168.80.10, 00:19:04,
GigabitEthernet8
```



Case 8. WAN Edge does not install route from type3 LSA from area 0 into the RIB

On ABR, routes from both cEdges are installed as well:

```
ABR#show ip ospf database

            OSPF Router with ID (10.2.10.10) (Process ID 2)

                Router Link States (Area 0)

Link ID         ADV Router      Age         Seq#       Checksum Link count
10.2.0.102      10.2.0.102      186         0x800000A2 0x005404 3
10.2.10.10      10.2.10.10      228         0x800000B4 0x004461 3

                Summary Net Link States (Area 0)

Link ID         ADV Router      Age         Seq#       Checksum
10.2.1.10       10.2.10.10      308         0x80000002 0x009A69
10.2.1.101      10.2.10.10      52          0x80000004 0x00A093
192.168.70.0    10.2.10.10      231         0x80000003 0x00EB7E

                Router Link States (Area 1)

Link ID         ADV Router      Age         Seq#       Checksum Link count
10.2.1.101      10.2.1.101      60          0x800000AB 0x00DE85 3
10.2.10.10      10.2.10.10      231         0x800000A8 0x001EA6 3

                Summary Net Link States (Area 1)

Link ID         ADV Router      Age         Seq#       Checksum
10.2.0.10       10.2.10.10      308         0x80000002 0x00A55F
10.2.0.102      10.2.10.10      52          0x80000004 0x00A192
192.168.80.0    10.2.10.10      228         0x80000003 0x007DE2
```

```
ABR#show ip ospf rib | b Codes
Codes: * - Best, > - Installed in global RIB

*    10.2.0.10/32, Intra, cost 1, area 0, Connected
       via 10.2.0.10, Loopback1
*>   10.2.0.102/32, Intra, cost 2, area 0
       via 192.168.80.102, GigabitEthernet6
*    10.2.1.10/32, Intra, cost 1, area 1, Connected
       via 10.2.1.10, Loopback2
*>   10.2.1.101/32, Intra, cost 2, area 1
       via 192.168.70.101, GigabitEthernet5
*    192.168.70.0/24, Intra, cost 1, area 1, Connected
       via 192.168.70.10, GigabitEthernet5
*    192.168.80.0/24, Intra, cost 1, area 0, Connected
       via 192.168.80.10, GigabitEthernet6

ABR#sh ip route ospf | b Gate
Gateway of last resort is not set

      10.0.0.0/8 is variably subnetted, 6 subnets, 3 masks
O        10.2.0.102/32 [110/2] via 192.168.80.102,
00:35:38, GigabitEthernet6
O        10.2.1.101/32 [110/2] via 192.168.70.101,
00:35:44, GigabitEthernet5
```

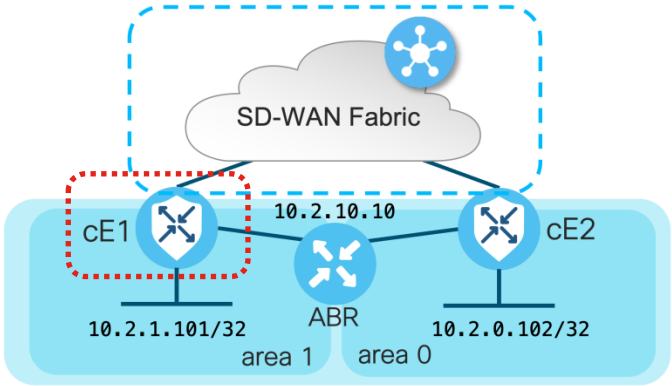
Case 8. WAN Edge does not install route from type3 LSA from area 0 into the RIB

But the problem is that on cE1, routes from cE2 and ABR routes from area 0 are not installed:

```
cE1#sh ip ospf database

            OSPF Router with ID (10.2.1.101) (Process ID 2)

                Router Link States (Area 1)

Link ID         ADV Router      Age         Seq#       Checksum Link count
10.2.1.101      10.2.1.101      238         0x800000AB 0x00DE85 3
10.2.10.10      10.2.10.10      411         0x800000A8 0x001EA6 3

                Summary Net Link States (Area 1)

Link ID         ADV Router      Age         Seq#       Checksum
10.2.0.10       10.2.10.10      488         0x80000002 0x00A55F
10.2.0.102      10.2.10.10      232         0x80000004 0x00A192
192.168.80.0    10.2.10.10      408         0x80000003 0x007DE2

ce1#show ip ospf rib  | b Codes
Codes: * - Best, > - Installed in global RIB

*>  10.2.1.10/32, Intra, cost 2, area 1
      via 192.168.70.10, GigabitEthernet7
*   10.2.1.101/32, Intra, cost 1, area 1, Connected
      via 10.2.1.101, Loopback1
*   192.168.70.0/24, Intra, cost 1, area 1, Connected
      via 192.168.70.101, GigabitEthernet7
```

```
ce1#show ip ospf database  summary 10.2.0.102

            OSPF Router with ID (10.2.1.101) (Process ID 2)

                Summary Net Link States (Area 1)

  LS age: 437
  Options: (No TOS-capability, DC, Upward)
  LS Type: Summary Links(Network)
  Link State ID: 10.2.0.102 (summary Network Number)
  Advertising Router: 10.2.10.10
  LS Seq Number: 80000004
  Checksum: 0xA192
  Length: 28
  Network Mask: /32
        MTID: 0         Metric: 2
```

No it’s not about DN bit…

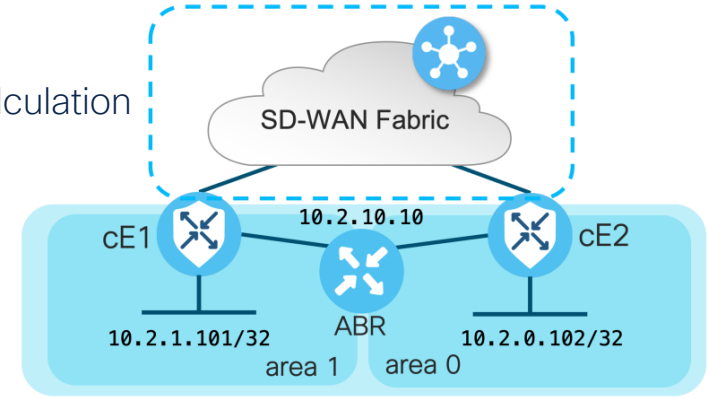


Case 8. WAN Edge does not install route from type3 LSA from area 0 into the RIB

Form a theory. What would we check further?

Keep in mind:

- There is not about DN-bit, this is Type 3 routes generated on the serivce side
- This route didn't pass any OMP-> OSPF redistribution
- If route didn't get into RIB from LSDB, it's result of SPF calculation



Case 8. WAN Edge does not install route from type3 LSA from area 0 into the RIB

If it is about SPF, let's enable SPF debugs on cE1 and on ABR:

```
ce1#debug ip ospf spf inter
OSPF SPF inter debugging is on

ABR#clear ip ospf process
Reset ALL OSPF processes? [no]: yes


ce1#term mon
Sep 21 2020 21:28:29.519 UTC: OSPF-2 INTER: Process partial spfQ: LSA 3/10.2.0.10/10.2.10.10, age 1, seq 0x80000003, area 1
Sep 21 2020 21:28:29.519 UTC: OSPF-2 INTER: Partial SPF for prefix 10.2.0.10/32, LSA 3/10.2.0.10/10.2.10.10
Sep 21 2020 21:28:29.519 UTC: OSPF-2 INTER: Start partial processing: type 3, LSID 10.2.0.10, mask 255.255.255.255,
Sep 21 2020 21:28:29.519 UTC: OSPF-2 INTER:   adv_rtr 10.2.10.10, age 1, seq 0x80000003, area 1
Sep 21 2020 21:28:29.519 UTC: OSPF-2 INTER: Downward bit set/Non-backbone LSA
Sep 21 2020 21:28:29.519 UTC: OSPF-2 INTER: Process partial spfQ: LSA 3/192.168.80.0/10.2.10.10, age 1, seq 0x80000004, area 1
Sep 21 2020 21:28:29.519 UTC: OSPF-2 INTER: Partial SPF for prefix 192.168.80.0/24, LSA 3/192.168.80.0/10.2.10.10
Sep 21 2020 21:28:29.519 UTC: OSPF-2 INTER: Start partial processing: type 3, LSID 192.168.80.0, mask 255.255.255.0,
Sep 21 2020 21:28:29.519 UTC: OSPF-2 INTER:   adv_rtr 10.2.10.10, age 1, seq 0x80000004, area 1
Sep 21 2020 21:28:29.519 UTC: OSPF-2 INTER: Downward bit set/Non-backbone LSA
Sep 21 2020 21:28:29.519 UTC: OSPF-2 EXTER: Process partial external spf queue
Sep 21 2020 21:28:29.519 UTC: OSPF-2 EXTER: Process partial nssa spf queue
Sep 21 2020 21:28:29.920 UTC: OSPF-2 SPF  : SPF due to Non-MAXAGE in lsa 3, LS ID 10.2.0.102, from 10.2.10.10
Sep 21 2020 21:28:29.920 UTC: OSPF-2 SPF  : Detect generic change in LSA type 3, LSID 10.2.0.102, from 10.2.10.10 area 1
Sep 21 2020 21:28:29.920 UTC: OSPF-2 SPF  : Schedule partial SPF type 3, LSID 10.2.0.102, adv_rtr 10.2.10.10 area 1
Sep 21 2020 21:28:29.920 UTC: OSPF-2 SPF  : Service partial SPF, spf instance 98, 1/0/0/0
Sep 21 2020 21:28:29.920 UTC: OSPF-2 EXTER: Process partial Opaque LSA queue
Sep 21 2020 21:28:29.920 UTC: OSPF-2 INTER: Process partial summary spf queue
Sep 21 2020 21:28:29.920 UTC: OSPF-2 INTER: Process partial spfQ: LSA 3/10.2.0.102/10.2.10.10, age 1, seq 0x80000005, area 1
Sep 21 2020 21:28:29.920 UTC: OSPF-2 INTER: Partial SPF for prefix 10.2.0.102/32, LSA 3/10.2.0.102/10.2.10.10
Sep 21 2020 21:28:29.920 UTC: OSPF-2 INTER: Start partial processing: type 3, LSID 10.2.0.102, mask 255.255.255.255,
Sep 21 2020 21:28:29.920 UTC: OSPF-2 INTER:   adv_rtr 10.2.10.10, age 1, seq 0x80000005, area 1
Sep 21 2020 21:28:29.920 UTC: OSPF-2 INTER: Downward bit set/Non-backbone LSA
```

Case 8. WAN Edge does not install route from type3 LSA from area 0 into the RIB

Analyzing results. Clearly there is no DN bit set in the corresponding type 3 LSA. So why do we ignore routes coming from “non-backbone” LSA while they are from backbone Area 0?!

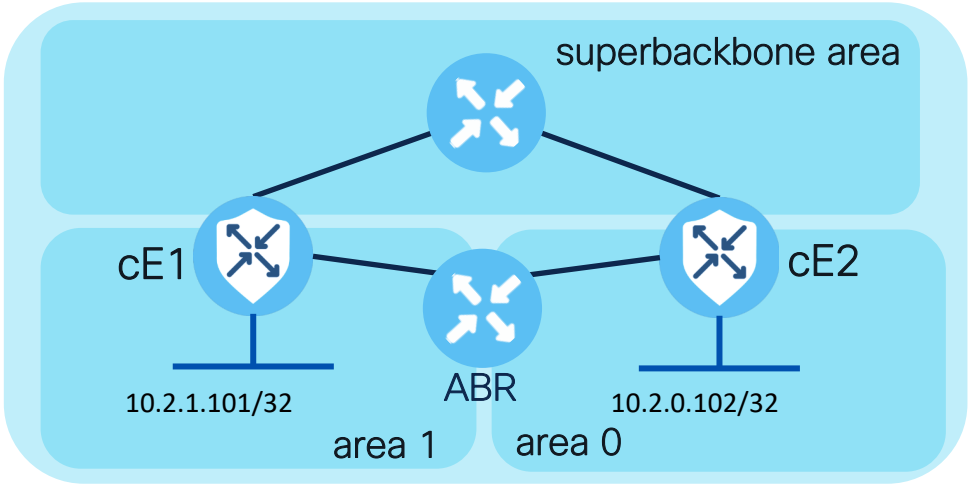
- This scenario and topology would match MPLS VPN case 100%
- cEdge behaves like a PE device in MPLS L3 VPN, we can confirm this also:

```
ce1#show ip ospf | i MPLS
  Connected to MPLS VPN Superbackbone, VRF 2

ce2#show ip ospf | i MPLS
  Connected to MPLS VPN Superbackbone, VRF 2
```

- SD-WAN fabric (OMP) is considered MPLS VPN Superbackbone here
- PE should ignore summary LSA generated by another PE originated from VPN and process area 0 summaries.
- As per RFC4577 4.1.4:
  - Two sites that are not in the same OSPF area will see the VPN backbone as being an integral part of the OSPF backbone. However, if there are area 0 routers that are NOT PE routers, then the VPN backbone actually functions as a sort of higher-level backbone, providing a third level of hierarchy above area 0

Case 8. WAN Edge does not install route from type3 LSA from area 0 into the RIB

Ephemeral topology to illustrate the issue, ABR is not really ABR anymore:



4.2.3. OSPF Areas

If a PE has a link that belongs to a non-zero area, the PE functions as an Area Border Router (ABR) for that area.

...

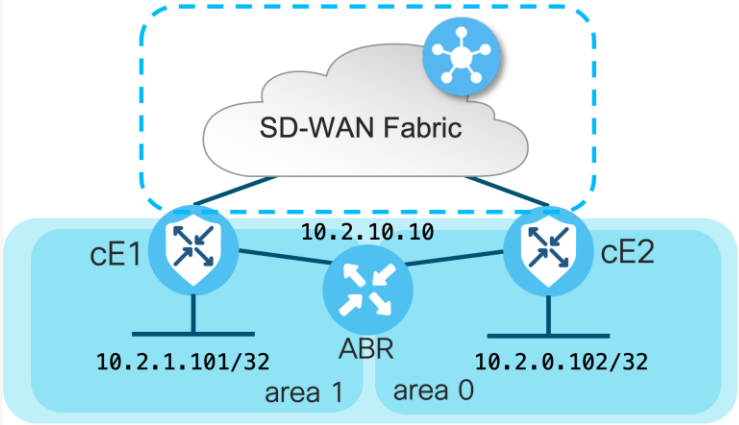
If the router has active attachments to multiple areas, only backbone summary-LSAs are examined (When we say that an ABR processes only backbone summary-LSAs, we are saying that the router will process only LSA-3 received from adjacencies in Area 0, see RFC 2328 Section 16.2)

Case 8. WAN Edge does not install route from type3 LSA from area 0 into the RIB

Solution 1. (required IOS-XE 17.3.2, 17.4+):

- **`capability vrf-lite`**

```
cE1(config)# router ospf 2 vrf 2
cE1(config-router)# capability vrf-lite
cE1(config-router)# commit
Commit complete.

ce1#show ip ospf rib 10.2.0.102 | b Codes
Codes: * - Best, > - Installed in global RIB
LSA: type/LSID/originator

*>  10.2.0.102/32, Inter, cost 3, area 1
     SPF Instance 100, age 00:01:00
      contributing LSA: 3/10.2.0.102/10.2.10.10 (area 1)
     Flags: RIB, HiPrio
      via 192.168.70.10, GigabitEthernet7, label 1048578, strict label 1048578
       Flags: RIB
       LSA: 3/10.2.0.102/10.2.10.10
       Source: 10.2.10.10 (area 1)
ce1#sh ip route vrf 2 ospf | i 10.2.0.102
O IA     10.2.0.102 [110/3] via 192.168.70.10, 00:01:03, GigabitEthernet7
```



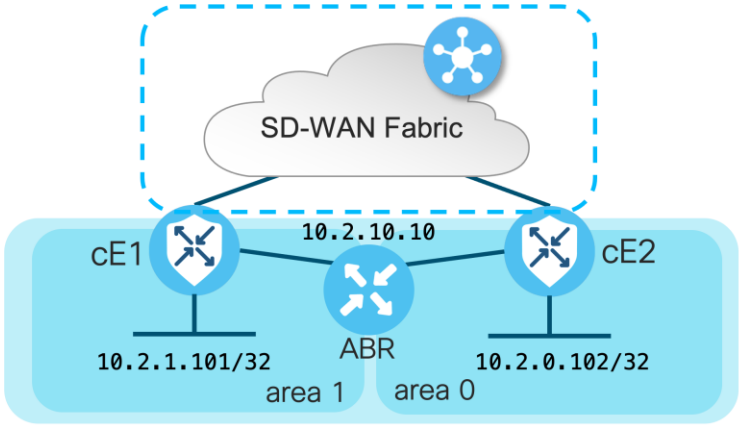
- Dangerous! Prone to loops for other routes OMP→OSPF→OMP
- Not required for vEdges, vEdge will install such routes without issues

Case 8. WAN Edge does not install route from type3 LSA from area 0 into the RIB

Solution 2.

- Divide OSPF domain into two separate domains and perform mutual redistribution between them on former ABR non-SDWAN router

```
ABR#sh ip ospf interface brief
Interface    PID   Area            IP Address/Mask    Cost  State Nbrs F/C
Lo1          1     0               10.2.0.10/32       1     LOOP  0/0
Gi6          1     0               192.168.80.10/24   1     P2P   1/1
Lo2          2     1               10.2.1.10/32       1     LOOP  0/0
Gi5          2     1               192.168.70.10/24   1     P2P   1/1

ABR#sh run | s r o
router ospf 2
 router-id 10.2.10.10
 redistribute ospf 1 subnets match internal external 1 external 2
router ospf 1
 router-id 10.1.10.10
 redistribute ospf 2 subnets match internal external 1 external 2

cE1#show ip route vrf 2 ospf | b Gate
Gateway of last resort is not set

      10.0.0.0/32 is subnetted, 4 subnets
O E2     10.2.0.10 [110/1] via 192.168.70.10, 00:04:02, GigabitEthernet7
O E2     10.2.0.102 [110/2] via 192.168.70.10, 00:04:02, GigabitEthernet7
O        10.2.1.10 [110/2] via 192.168.70.10, 00:05:46, GigabitEthernet7
O E2  192.168.80.0/24 [110/1] via 192.168.70.10, 00:04:02, GigabitEthernet7
```

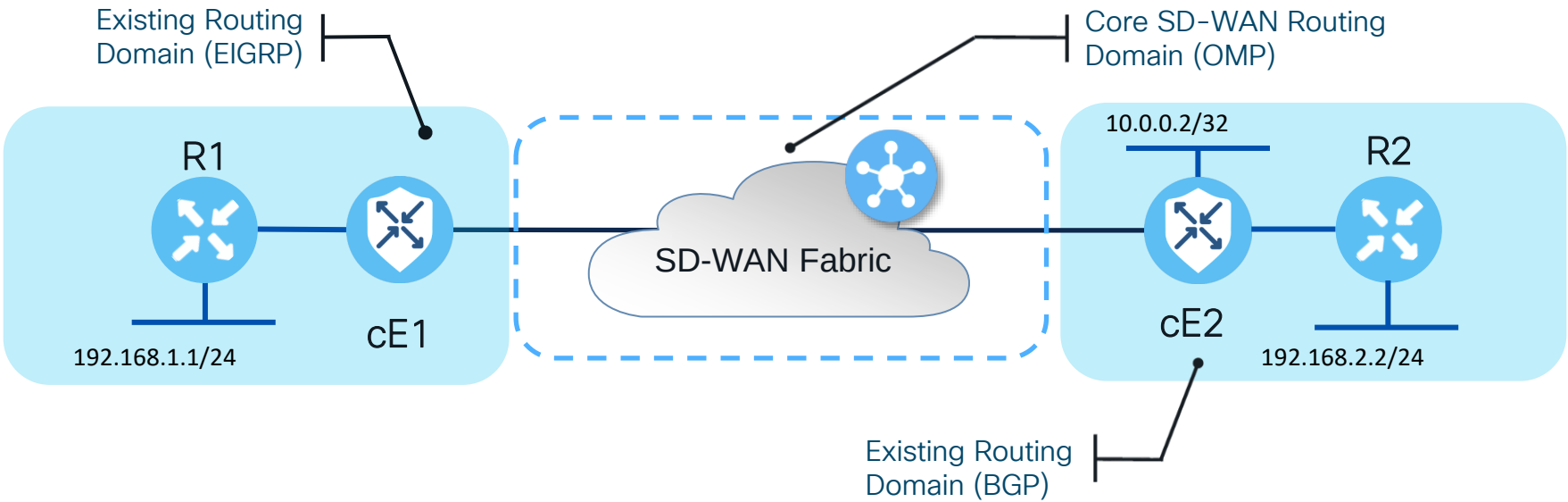


- Dangerous! Prone to loops OMP→OSPF→OSPF→OMP
- ABR is rather ASBR now

Case 9. OMP to EIGRP redistribution

Case 9. OMP to EIGRP redistribution

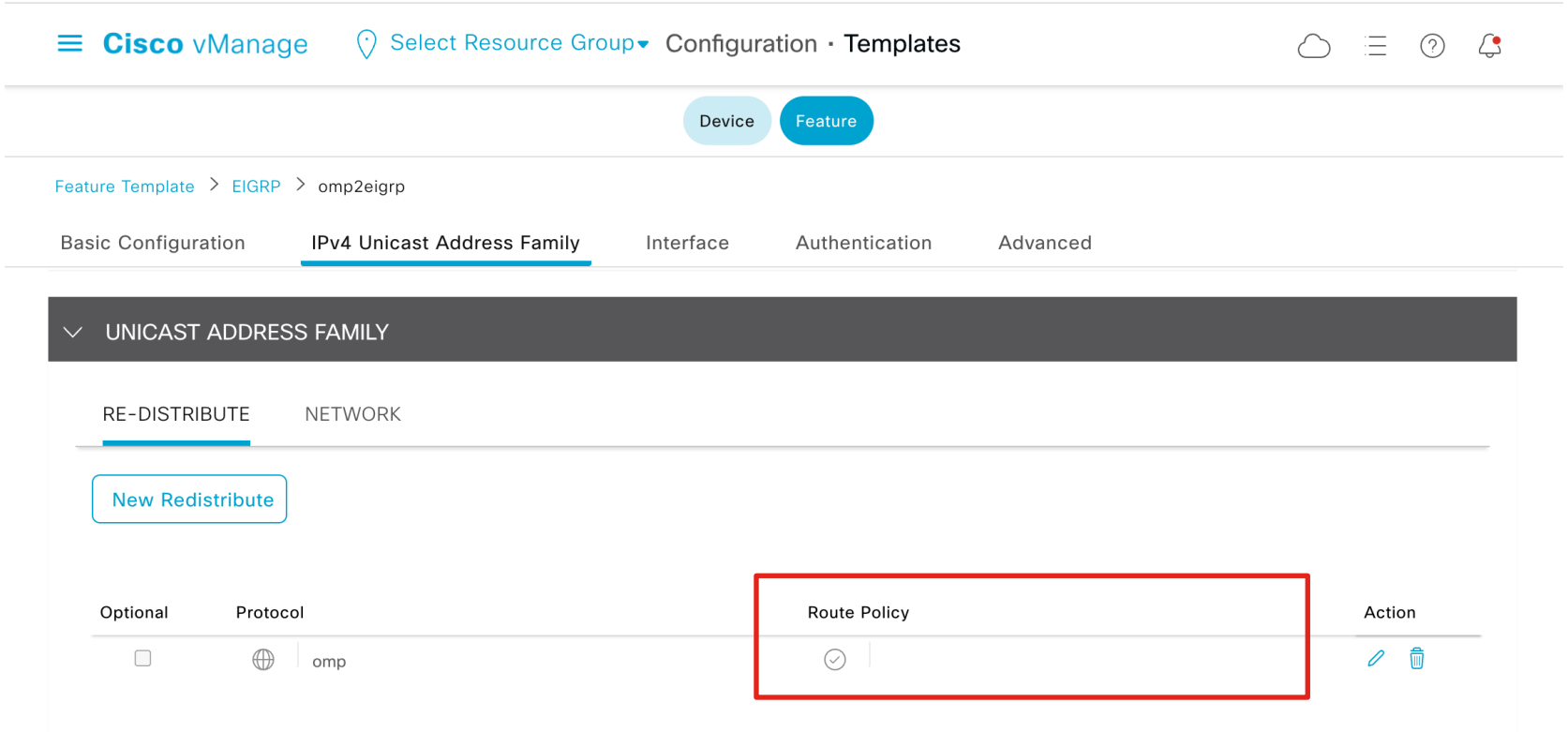
```
ce2(config-vrf-1)# show config
sdwan
 omp
  address-family ipv4 vrf 1
   advertise bgp

 !
```



*This is not a break-fix troubleshooting case, rather demonstration of expected behavior

Case 9. OMP to EIGRP redistribution

vManage feature template for cE1 – EIGRP:



Case 9. OMP to EIGRP redistribution

vManage template for cE1 resulted in the following configuration:

```
478  router eigrp eigrp-name
479   address-family ipv4 vrf 30 autonomous-system 100
480    af-interface GigabitEthernet4
481     no dampening-change
482     no dampening-interval
483     hello-interval 5
484     hold-time      15
485     split-horizon
486     exit-af-interface
487    !
488    network 10.0.0.0 0.0.0.255
489    topology base
490     redistribute omp
491     exit-af-topology
492    !
493    exit-address-family
494   !
495  !
```

We get used to the fact that EIGRP has default seed metric of infinity during the redistribution process. So will redistribution work here if no seed metric defined?

Case 9. OMP to EIGRP redistribution

Let's check routing on cE1:

```
cE1#sh ip ro vrf 3 192.168.2.2

Routing Table: 3
Routing entry for 192.168.2.0/24
  Known via "omp", distance 251, metric 0, type omp
  Redistributing via eigrp 100
  Advertised by eigrp 100
  Last update from 10.0.0.2 on Sdwan-system-intf, 00:08:36 ago
  Routing Descriptor Blocks:
  * 10.0.0.2 (default), from 10.0.0.2, 00:08:36 ago, via Sdwan-system-intf
      Route metric is 0, traffic share count is 1
```

Route is received from OMP and redistributed into EIGRP, so far all fine.

Case 9. OMP to EIGRP redistribution

And then topology table:

```
cE1#show ip eigrp vrf 3 topology 192.168.2.0/24
EIGRP-IPv4 VR(eigrp-name) Topology Entry for AS(100)/ID(192.168.70.101)
          Topology(base) TID(0) VRF(3)
EIGRP-IPv4(100): Topology base(0) entry for 192.168.2.0/24
  State is Passive, Query origin flag is 1, 1 Successor(s), FD is 1
  Descriptor Blocks:
  10.0.0.2, from Redistributed, Send flag is 0x0
      Composite metric is (1/0), route is External
      Vector metric:
        Minimum bandwidth is 0 Kbit
        Total delay is 0 picoseconds
        Reliability is 0/255
        Load is 0/255
        Minimum MTU is 0
        Hop count is 0
        Originating router is 192.168.70.101
      External data:
        AS number of route is 0
        External protocol is OMP-Agent, external metric is 4294967294
        Administrator tag is 0 (0x00000000)
```

Shouldn't it be declined by neighbor once we advertise it? Keep in mind: 4294967294 = 0xFFFFFFFE

Case 9. OMP to EIGRP redistribution

But this route will be installed by R1, surprise surprise!

```
R1#sh ip route eigrp | i 192.168.2.0/24
D EX   192.168.2.0/24 [170/257] via 192.168.70.101, 00:04:14, GigabitEthernet5

R1#sh ip eigrp topology 192.168.2.0/24
EIGRP-IPv4 Topology Entry for AS(100)/ID(192.168.1.1) for 192.168.2.0/24
  State is Passive, Query origin flag is 1, 1 Successor(s), FD is 257
  Descriptor Blocks:
  192.168.70.101 (GigabitEthernet5), from 192.168.70.101, Send flag is 0x0
      Composite metric is (257/1), route is External
      Vector metric:
        Minimum bandwidth is 0 Kbit
        Total delay is 10 microseconds
        Reliability is 0/255
        Load is 1/255
        Minimum MTU is 1500
        Hop count is 1
        Originating router is 192.168.70.101
      External data:
        AS number of route is 0
        External protocol is Unknown protocol, external metric is 4294967294
        Administrator tag is 0 (0x00000000)
```

Case 9. OMP to EIGRP redistribution

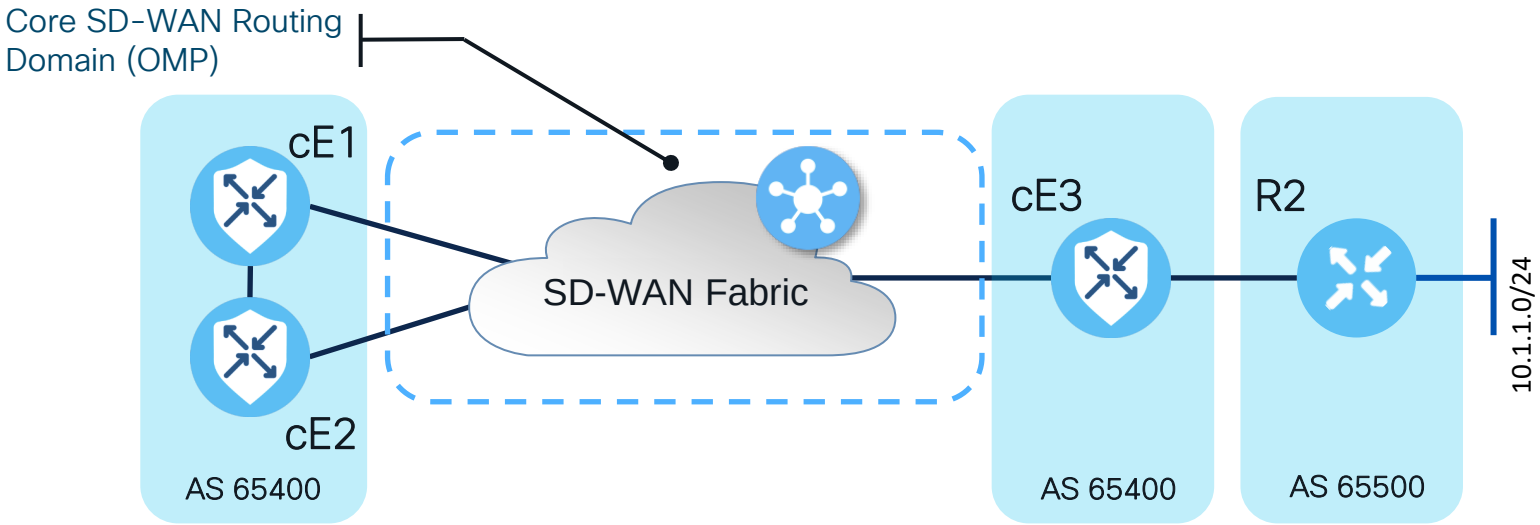
Analyzing results. Why was it installed? Key points:

- Scaled-Bw = (((10^7)/Min BW in kbps)*256) when BW>0
  - Scaled-BW = 1, when BW=0 because you can not divide by zero ;)
- Scaled-Delay = ((delay in picoseconds*256)/(10^7))
- Default K-values metric: 256*(Scaled BW + Scaled Delay)
- Delay and bandwidth can not be zero – if they are, then values will be set to the defined minimum value of one "1"
- No need for seed metric (for most of the simple scenarios)
- Can lead to suboptimal routing in legacy network (CSCvv76258, CSCvp89135 enhancements to improve this and make metrics configurable per prefix with route-map)

Case 10. OMP–BGP routing loop

Case 10. OMP-BGP routing loop

.



SD-WAN router	site-id	system-ip
cE1	1	10.0.0.1
cE2	2	10.0.0.2
cE3	3	10.0.0.3

*This is not a break-fix troubleshooting case, but demonstration of loop prevention and OMP features

Case 10. OMP-BGP routing loop

Bidirectional redistribution and **propagate-aspath** configured on all routers

```
cE1#
router bgp 65400
 address-family ipv4 vrf 1
  redistribute omp
  propagate-aspath
  neighbor 192.168.160.102 remote-as
65400
  neighbor 192.168.160.102 activate
 exit-address-family
 !
sdwan
 omp
  address-family ipv4 vrf 1
   advertise bgp
  !
 !
!
```

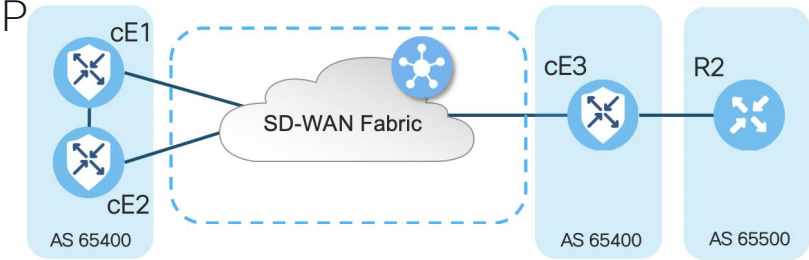
```
cE2#
router bgp 65400
 address-family ipv4 vrf 1
  redistribute omp
  propagate-aspath
  neighbor 192.168.160.101 remote-as
65400
  neighbor 192.168.160.101 activate
  neighbor 192.168.160.101 send-community
both
 exit-address-family
 !
sdwan
 omp
  address-family ipv4 vrf 1
   advertise bgp
  !
 !
!
```

```
cE3#
router bgp 65400
 address-family ipv4 vrf 1
  redistribute omp
  propagate-aspath
  neighbor 192.168.60.11 remote-as 65500
  neighbor 192.168.60.11 activate
 exit-address-family
 !
sdwan
 omp
  address-family ipv4 vrf 1
   advertise bgp
  !
 !
!
```

propagate-aspath - Carry the BGP AS path into OMP

Note that cE1 does not send BGP communities to cE2

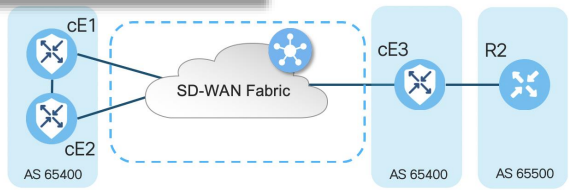
Let's see effects caused by that misconfig and denonstrate how loop prevention mechanisms work



Case 10. OMP-BGP routing loop

In the initial state, the route is redistributed by cE3 and learnt by cE1 and cE2 via OMP, both redistribute route to BGP and advertise it to each other

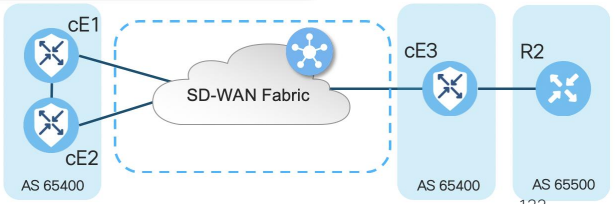
```
cE1#show bgp vpnv4 unicast vrf 1 10.1.1.1/24
BGP routing table entry for 1:1:10.1.1.1/24, version 342041
Paths: (2 available, best #2, table 1)
  Advertised to update-groups:
     4           5
  Refresh Epoch 1
  65500
    192.168.160.102 (via vrf 1) from 192.168.160.102 (192.168.109.102)
      Origin incomplete, metric 1000, localpref 50, valid, internal
      Extended Community: SoO:0:2 RT:1:1
      rx pathid: 0, tx pathid: 0
      Updated on Aug 21 2020 11:23:32 GMT
  Refresh Epoch 1
  65500
    10.0.0.3 (via default) from 0.0.0.0 (192.168.109.101)
      Origin incomplete, metric 1000, localpref 50, valid, sourced, best
      Extended Community: SoO:0:1 RT:1:1
      rx pathid: 0, tx pathid: 0x0
      Updated on Aug 21 2020 11:23:32 GMT
```



Case 10. OMP-BGP routing loop

Note that route advertised by cE1 to cE2 has no SoO set (because it was not configured):

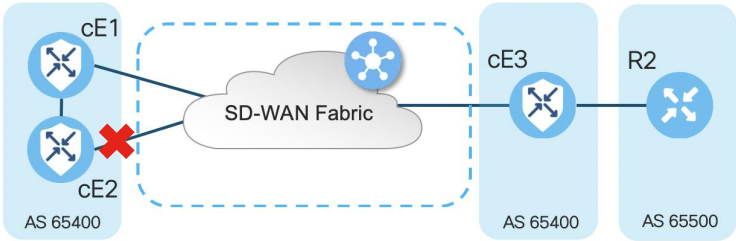
```
cE2#show bgp vpnv4 unicast vrf 1 10.1.1.1/24
BGP routing table entry for 1:1:10.1.1.1/24, version 327810
Paths: (2 available, best #2, table 1)
  Advertised to update-groups:
     5           6
  Refresh Epoch 1
  65500
    192.168.160.101 (via vrf 1) from 192.168.160.101 (192.168.109.101)
      Origin incomplete, metric 1000, localpref 50, valid, internal
      Extended Community: RT:1:1
      rx pathid: 0, tx pathid: 0
      Updated on Aug 21 2020 11:23:32 GMT
  Refresh Epoch 1
  65500
    10.0.0.3 (via default) from 0.0.0.0 (192.168.109.102)
      Origin incomplete, metric 1000, localpref 50, valid, sourced, best
      Extended Community: SoO:0:2 RT:1:1
      rx pathid: 0, tx pathid: 0x0
      Updated on Aug 21 2020 11:23:32 GMT
```



Case 10. OMP-BGP routing loop

Let's simulate failure now, cE2 is disconnected from the SD-WAN fabric

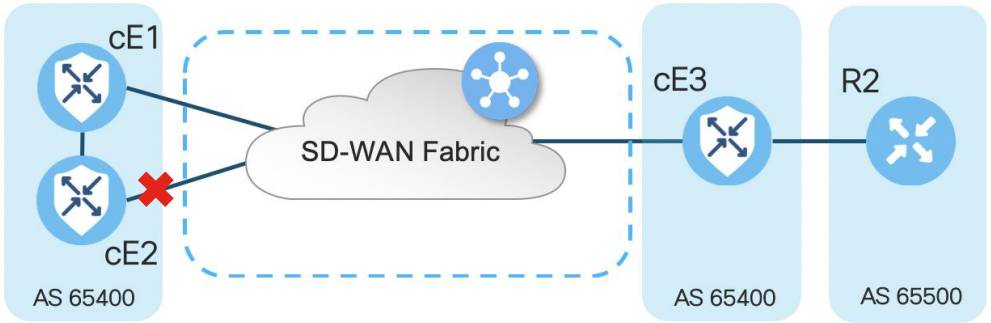
```
ce2(config)# interface GigabitEthernet 2
ce2(config-if)# shutdown
ce2(config-if)# end
Uncommitted changes found, commit them? [yes/no/CANCEL] yes
Commit complete.
ce2#show bgp vpnv4 unicast vrf 1 10.1.1.1/24
BGP routing table entry for 1:1:10.1.1.1/24, version 345276
Paths: (1 available, best #1, table 1)
  Advertised to update-groups:
     6
  Refresh Epoch 1
  65500
    192.168.160.101 (via vrf 1) from 192.168.160.101 (192.168.109.101)
      Origin incomplete, metric 1000, localpref 50, valid, internal, best
      Extended Community: SoO:0:1 RT:1:1
      rx pathid: 0, tx pathid: 0x0
      Updated on Aug 21 2020 11:23:32 GMT
```



Case 10. OMP-BGP routing loop

cE1 still prefers route via OMP (this is the only route remains) originated by cE3:

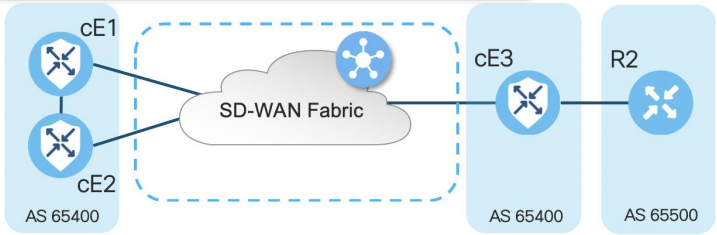
```
ce1#show bgp vpnv4 unicast vrf 1 10.1.1.1/24
BGP routing table entry for 1:1:10.1.1.1/24, version 342041
Paths: (1 available, best #1, table 1)
  Advertised to update-groups:
     4           5
  Refresh Epoch 1
  65500
    10.0.0.3 (via default) from 0.0.0.0 (192.168.109.101)
      Origin incomplete, metric 1000, localpref 50, valid, sourced, best
      Extended Community: SoO:0:1 RT:1:1
      rx pathid: 0, tx pathid: 0x0
      Updated on Aug 21 2020 11:23:32 GMT
```



Case 10. OMP-BGP routing loop

Next, connectivity on WAN interface of cE2 restored, cE2 prefers route from cE1 via iBGP (because of better AD)

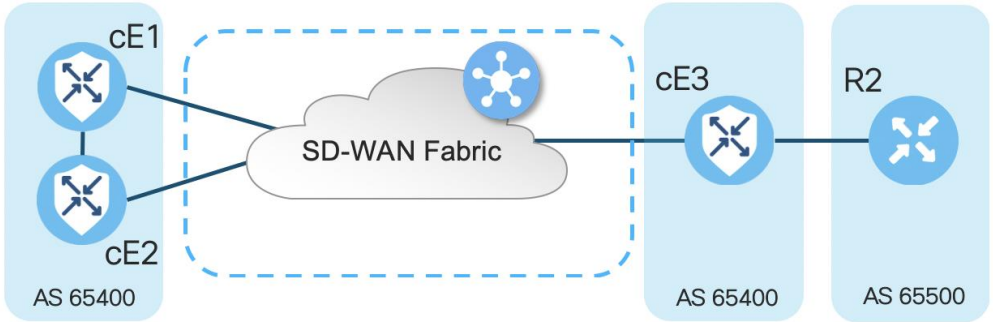
```
ce2(config)# interface GigabitEthernet 2
ce2(config-if)# no shutdown
ce2(config-if)# end
Uncommitted changes found, commit them? [yes/no/CANCEL] yes
Commit complete.
ce2#show bgp vpnv4 unicast vrf 1 10.1.1.1/24
BGP routing table entry for 1:1:10.1.1.1/24, version 345276
Paths: (1 available, best #1, table 1)
  Advertised to update-groups:
     6
  Refresh Epoch 1
  65500
    192.168.160.101 (via vrf 1) from 192.168.160.101 (192.168.109.101)
      Origin incomplete, metric 1000, localpref 50, valid, internal, best
      Extended Community: RT:1:1
      rx pathid: 0, tx pathid: 0x0
      Updated on Aug 21 2020 11:23:32 GMT
```



Case 10. OMP-BGP routing loop

cE1 still prefers route via OMP originated by cE3. Keep in mind that cE1 redistributes OMP into BGP:

```
ce1#show bgp vpnv4 unicast vrf 1 10.1.1.1/24
BGP routing table entry for 1:1:10.1.1.1/24, version 569358
Paths: (1 available, best #1, table 1)
  Advertised to update-groups:
     4           5
  Refresh Epoch 1
  65500
    10.0.0.3 (via default) from 0.0.0.0 (192.168.109.101)
      Origin incomplete, metric 1000, localpref 50, valid, sourced, best
      Extended Community: SoO:0:1 RT:1:1
      rx pathid: 0, tx pathid: 0x0
      Updated on Aug 21 2020 15:13:09 GMT
```



Case 10. OMP-BGP routing loop

Now something happens with cE3 connectivity to R2. For testing, the interface is shut down, and R2 BGP peer is lost:

```
ce3(config)# interface GigabitEthernet 6
ce3(config-if)# shutdown
ce3(config-if)# commit
```

As a result, the routing loop is formed between cE1 and cE2 (cE2 redistributes route from OMP and advertise to cE1 via BGP, cE1 redistributes BGP to OMP and advertise to cE2):

```
ce1#show bgp vpnv4 unicast vrf 1 10.1.1.1/24
BGP routing table entry for 1:1:10.1.1.1/24, version
732548
Paths: (1 available, best #1, table 1)
  Advertised to update-groups:
     5
  Refresh Epoch 1
  65500
    192.168.160.102 (via vrf 1) from 192.168.160.102
(192.168.109.102)
      Origin incomplete, metric 1000, localpref 50,
valid, internal, best
      Extended Community: SoO:0:2 RT:1:1
      rx pathid: 0, tx pathid: 0x0
      Updated on Aug 21 2020 15:38:47 GMT
```

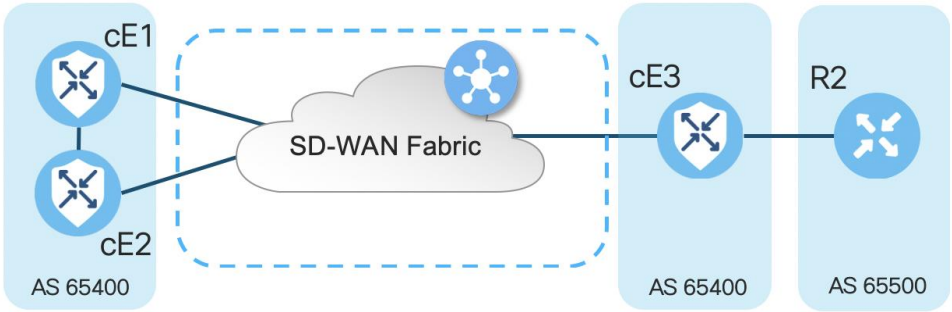
```
ce2#show bgp vpnv4 unicast vrf 1 10.1.1.1/24
BGP routing table entry for 1:1:10.1.1.1/24, version
639650
Paths: (1 available, best #1, table 1)
  Advertised to update-groups:
     5          6
  Refresh Epoch 1
  65500
    10.0.0.1 (via default) from 0.0.0.0
(192.168.109.102)
      Origin incomplete, metric 1000, localpref 50,
valid, sourced, best
      Extended Community: SoO:0:2 RT:1:1
      rx pathid: 1, tx pathid: 0x0
      Updated on Aug 21 2020 15:38:47 GMT
```

Case 10. OMP-BGP routing loop

Routing loop formed as a result of misconfiguration making impossible SoO mechanism to prevent the loop. What are the other options?

Keep in mind:

- cE1 does not send BGP communities to cE2, this is intentional (fixing that would be obvious)
- AS-PATH is a loop prevention mechanism for eBGP only, hence didn't help us within AS 65400
- **`propagate-aspath`** has been already configured

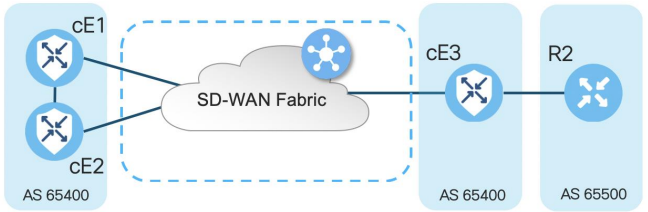


Case 10. OMP-BGP routing loop

Solution 1. Configure **overlay-as**

```
config-transaction
 sdwan
  omp
   overlay-as 64512
   exit
```

- By default, OMP is transparent to BGP even if **propagate-aspath** configured
- **overlay-as** prepends AS specified as a parameter of this command to BGP AS_PATH of routes exported from OMP to BGP
- If you configure the same overlay AS number on multiple devices in the overlay network, all these devices are considered to be part of the same AS, and as a result, they do not forward any routes that contain the overlay AS number, hence routing loop will be prevented.

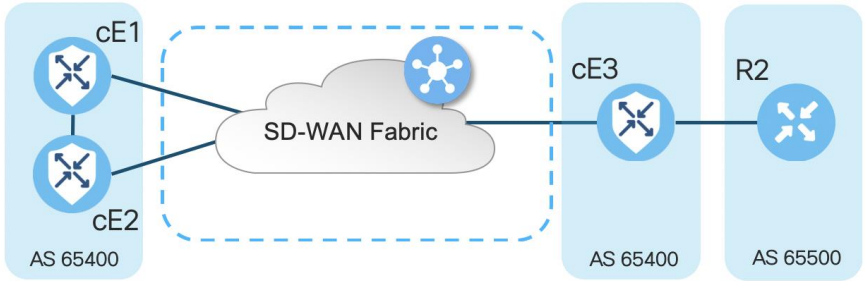


Case 10. OMP-BGP routing loop

Or simply change site-id to match on both cE1 and cE2 **and** configure **`send-community extended`** on cE1 as well, that would be obvious fix:

```
cE1#
router bgp 65400
 address-family ipv4 vrf 1
 neighbor 192.168.160.102 send-
community both
```

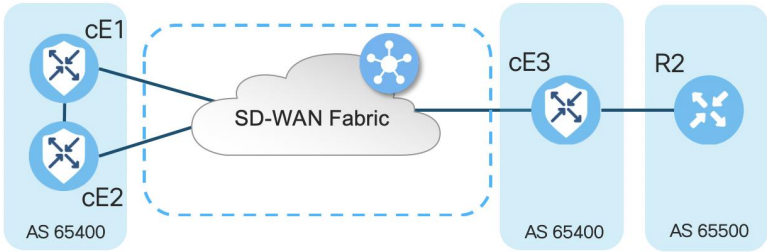
- site-id preserved as BGP SoO extended community attribute (you can notice SoO:0:<site-id>). That is used to identify routes that have originated from the same site so that the re-advertisement of that prefix back to OMP can be prevented (route installed into the RIB with AD 252 only if OMP is down)
- For SoO to function, BGP peers at site must send BGP extended communities



Case 10. OMP-BGP routing loop

If you change just site-id to match on both cE1 and cE2 and **`send-community extended`** is not configured cE1:

- vSmart advertises routes back to the router with the same site-id as in the route itself, since the originator attribute of the route is different, loop prevention will not be triggered, but control plane routing loop will not form because the OMP route will not be installed into the RIB/FIB.
- This is because the OMP route will stay in the Inv,U (Invalid,Unresolved) state. By default, data plane tunnels can not be established between sites with the same site-id unless **`allow-same-site-tunnels`** is configured. If the data plane tunnel BFD session is in the down state, TLOC will remain unresolved and hence route can not be resolved as well.
- Still can lead to a loop and traffic blackhole in a corner case if vSmart controller rewrites tloc-list with control policy (e.g in a hub'n'spoke topologies where originator will be equal to a hub). Enhancement opened CSCwa16188 to handle this as well.



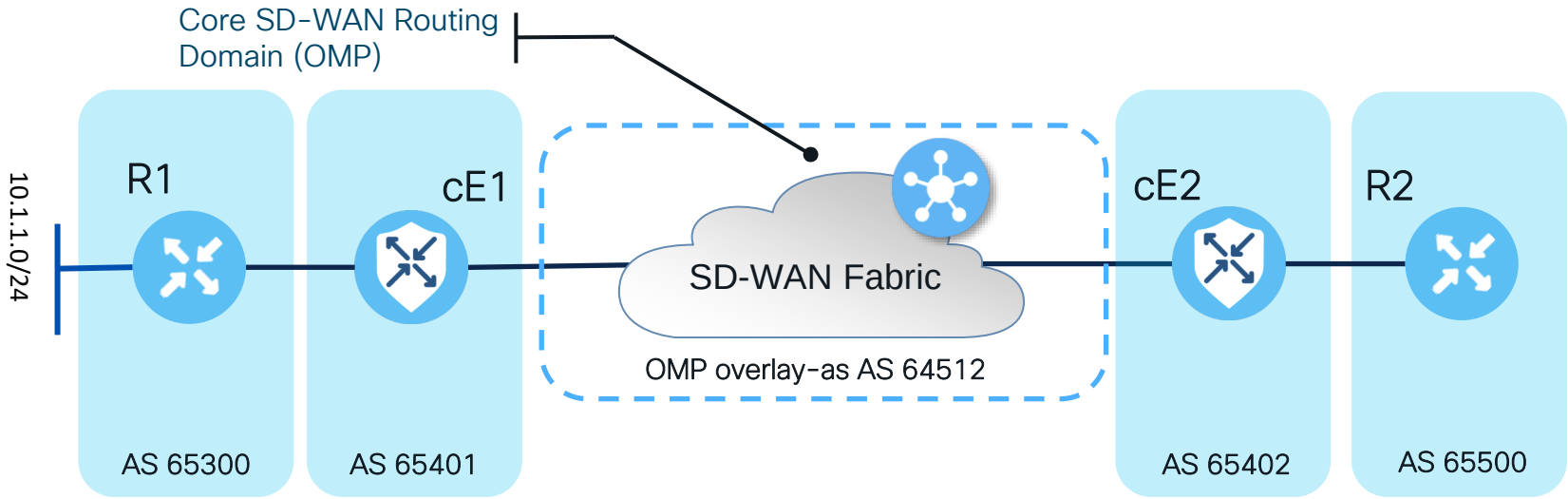
Case 11. propagate-aspath and overlay-as

or “Why Edge Does Not Advertise Its Own AS When BGP Routes Are Advertised Into OMP”

cisco Live!

Case 11. propagate-aspath and overlay-as

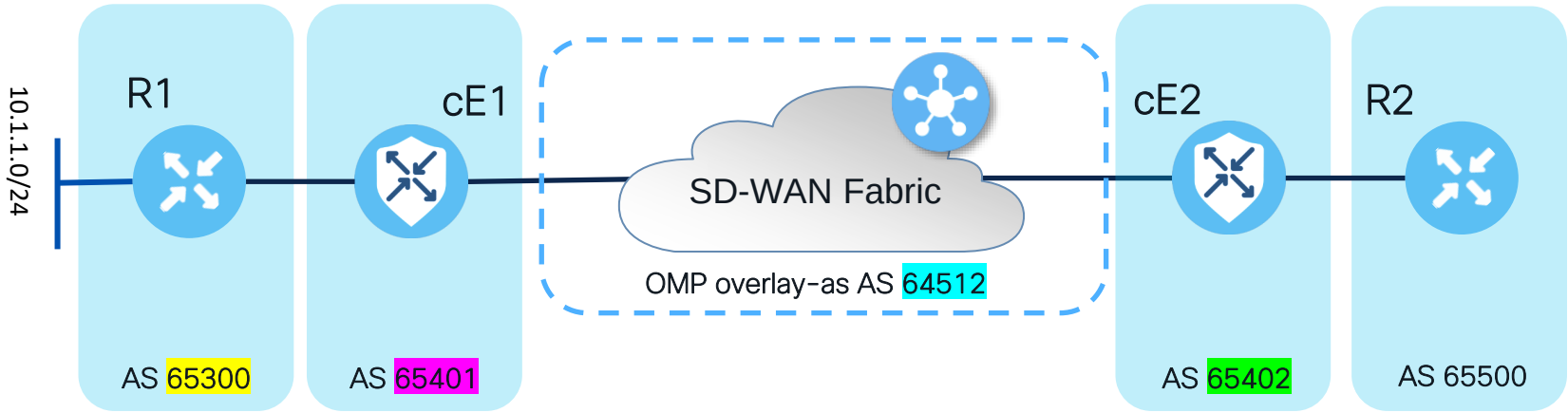
·How features work?



- **`overlay-as`** and **`propagate-aspath`** are dependent on each other
- **`overlay-as`** is export (egress) feature for OMP to BGP redistribution

Case 11. propagate-aspath and overlay-as

For better visual comprehension I highlighted each AS with it's own colour:



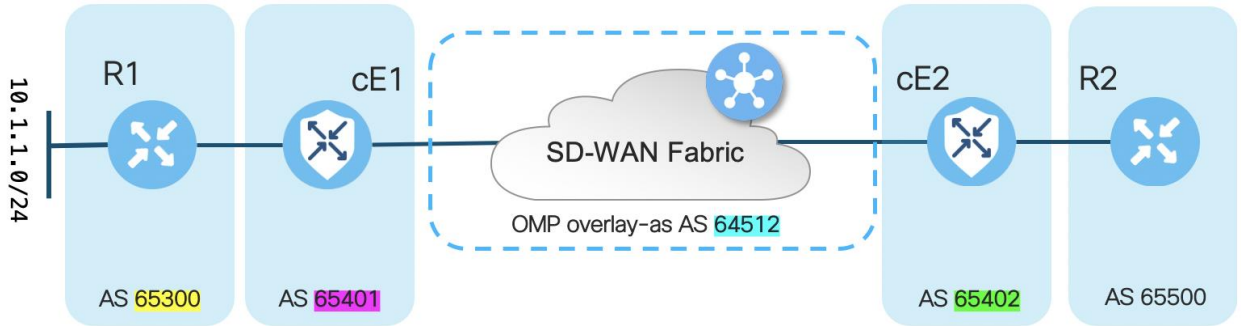
Case 11. propagate-aspath and overlay-as. Variant 1.

feature/router	cE1	cE2
propagate-aspath	?	x
overlay-as	?	?

✓ - configured
? – does not matter
x – not configured

cE2 receives route and advertises it into BGP towards R2 with its own AS only and ignoring AS_PATH (normal eBGP behaviour).

```
R2#show ip bgp | i 10.1.1.0/24
 *>   10.1.1.0/24 192.168.60.103         1000             0 65402 ?
```



Case 11. propagate-aspath and overlay-as. Variant 2 (pre-17.6 old behaviour).

feature/router	cE1	cE2
propagate-aspath	?	x
overlay-as	?	✓

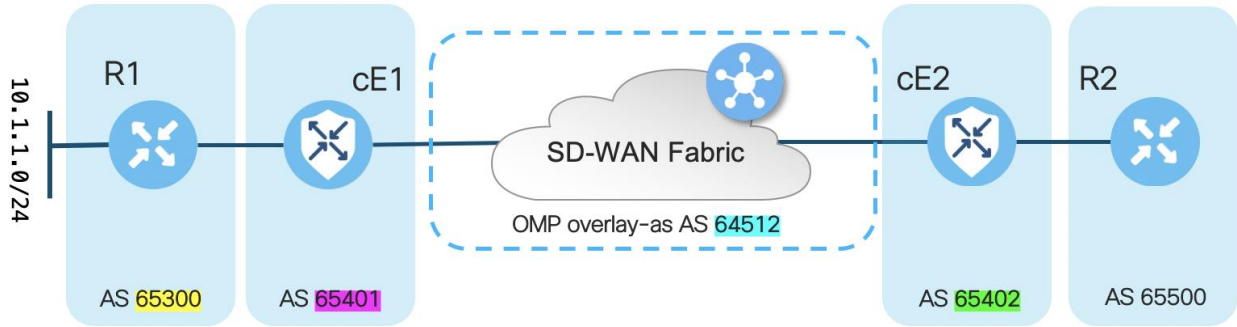
✓ - configured
? – does not matter
x – not configured

cE2:

```
sdwan
 omp
  overlay-as 64512
 !
!
```

cE2 receives omp route, ignores AS-PATH attribute in it and advertises the route into BGP towards R2 adding Overlay-AS and its own BGP AS:

```
R2#show ip bgp | i 10.1.1.0/24
 *>   10.1.1.0/24 192.168.60.103         1000             0 65402 64512 ?
```



Case 11. propagate-aspath and overlay-as. Variant 2 (new behaviour from 17.6).

feature/router	cE1	cE2
propagate-aspath	?	x
overlay-as	?	✓

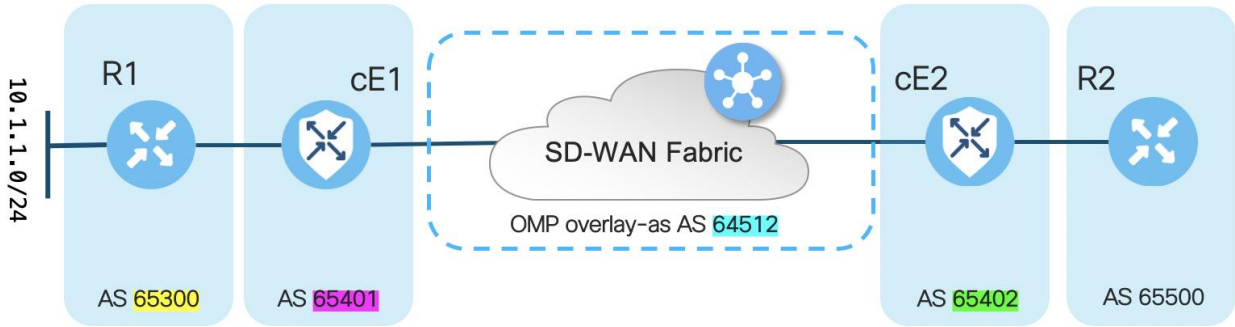
✓ - configured
? – does not matter
x – not configured

cE2:

```
sdwan
 omp
  overlay-as 64512
 !
!
```

cE2 receives omp route, ignores AS_PATH attribute and advertises it into BGP towards R2 adding own BGP AS only, Overlay-AS won't be added because **`propagate-aspath`** is not configured on cE2

```
R2#show ip bgp | i 10.1.1.0/24
 *>   10.1.1.0/24 192.168.60.103         1000             0 65402 ?
```



Case 11. propagate-aspath and overlay-as. Variant 3.

feature/router	cE1	cE2
propagate-aspath	x	✓
overlay-as	?	✓

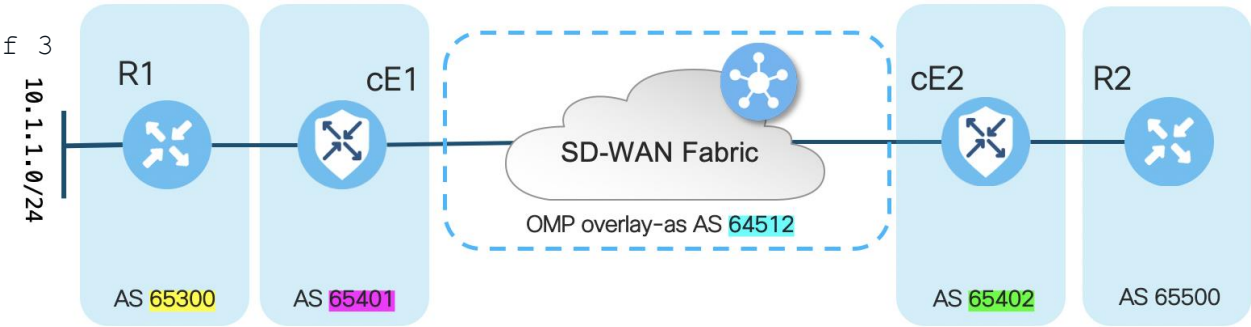
✓ - configured
? – does not matter
x – not configured

cE2:

```
router bgp 65402
 address-family ipv4 unicast vrf 3
  propagate-aspath
  exit-address-family
 !
!
sdwan
 omp
  overlay-as 64512
 !
!
```

cE2 receives omp route without AS_PATH from cE1 (**propagate-aspath** is not configured on cE1) and advertises the route into BGP towards R2 adding Overlay-AS and its own BGP AS:

```
R2#show ip bgp | i 10.1.1.0/24
 *>   10.1.1.0/24 192.168.60.103         1000             0 65402 64512 ?
```



Case 11. propagate-aspath and overlay-as. Variant 4.

feature/router	cE1	cE2
propagate-aspath	✓	✓
overlay-as	?	x

✓ - configured
? – does not matter
x – not configured

cE1:

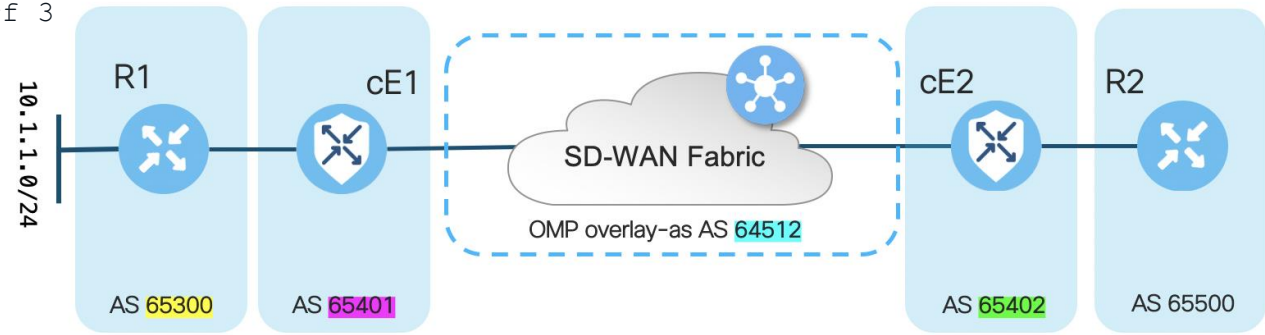
```
router bgp 65401
 address-family ipv4 unicast vrf 3
  propagate-aspath
  exit-address-family
 !
!
```

cE2:

```
sdwan
 omp
  no overlay-as
 !
!
```

cE2 receives omp route and advertises it into BGP, pre-pending the received AS_PATH with its own BGP AS:

```
R2#show ip bgp | i 10.1.1.0/24
 *>   10.1.1.0/24 192.168.60.103         1000             0 65402 65300 ?
```



Case 11. propagate-aspath and overlay-as. Variant 5.

feature/router	cE1	cE2
propagate-aspath	✓	✓
overlay-as	?	✓

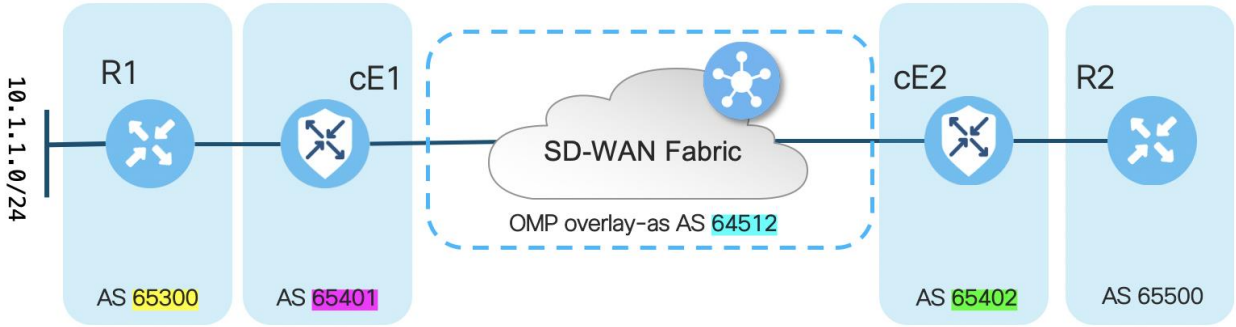
✓ - configured
? – does not matter
x – not configured

cE2:

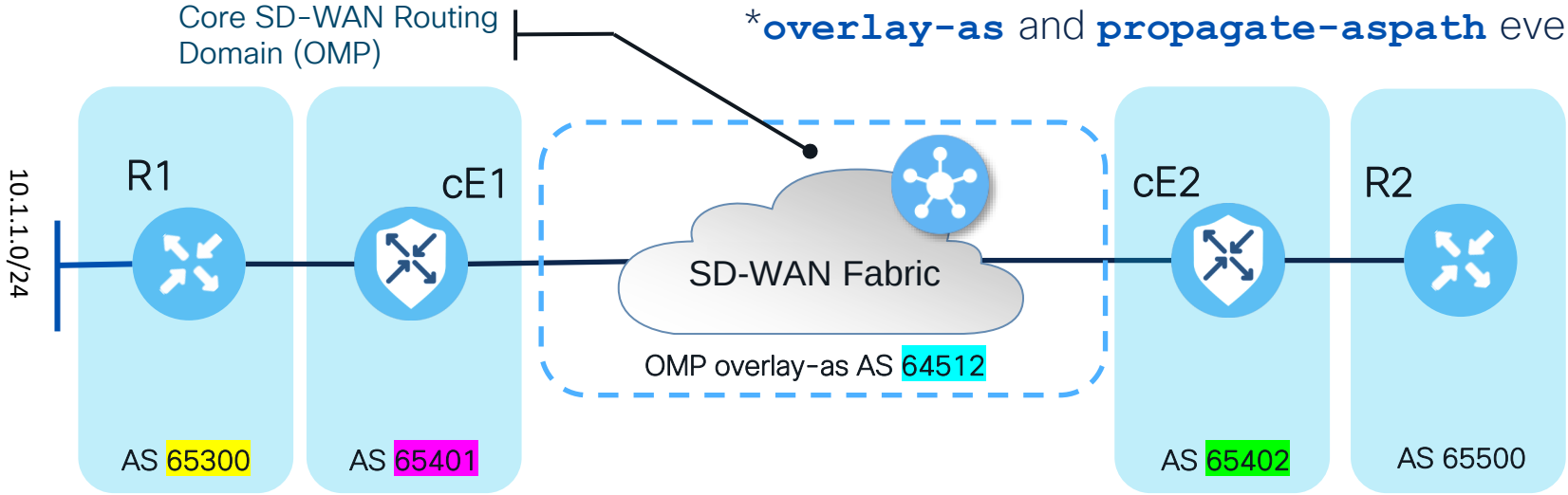
```
sdwan
 omp
  overlay-as 64512
 !
!
```

cE2 receives omp route and advertises it into BGP towards R2 pre-pending the AS_PATH with Overlay-AS followed by its own BGP AS:

```
R2#show ip bgp | i 10.1.1.0/24
 *>   10.1.1.0/24 192.168.60.103         1000             0 65402 64512 65300 ?
```



Case 11. propagate-aspath and overlay-as. Common confusion: Why Edge Does Not Advertise Its Own AS When BGP Routes Are Advertised Into OMP?



User's (wrong) expectations:

```
R2#show ip bgp | i 10.1.1.0/24
 *>   10.1.1.0/24 192.168.60.103         1000             0 65402 64512 65401 65300 ?
```

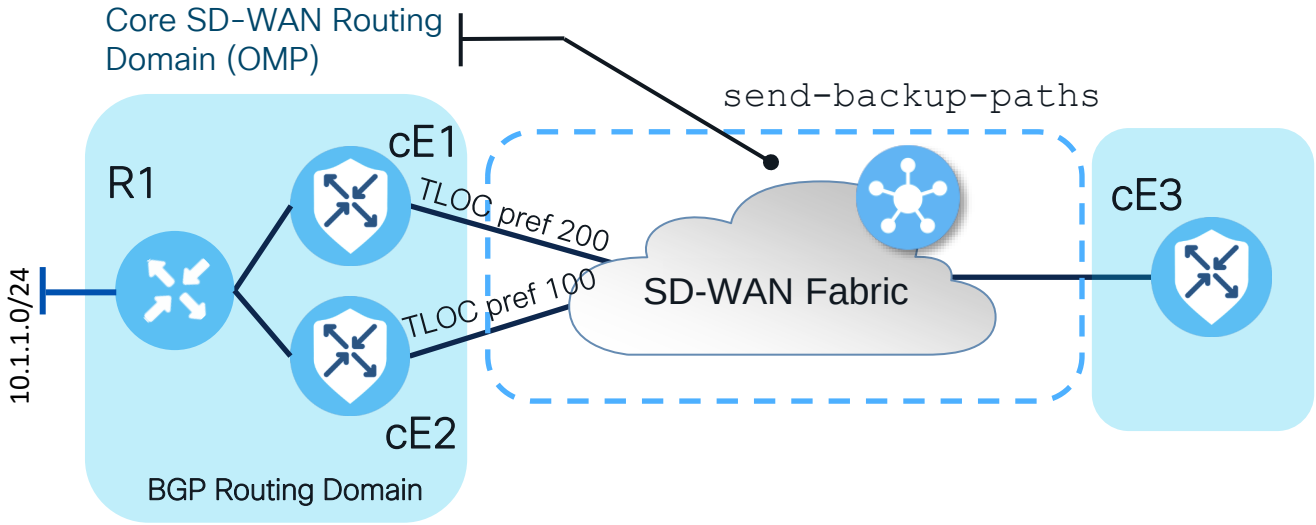
Reality (expected behaviour):

```
R2#show ip bgp | i 10.1.1.0/24
 *>   10.1.1.0/24 192.168.60.103         1000             0 65402 64512 65300 ?
```

Simple explanation: OMP is not BGP. It's redistribution between protocols despite the command (**advertise**)!

Case 12. Temporary blackholing on redundancy recovery

Case 12. Temporary blackholing on redundancy recovery

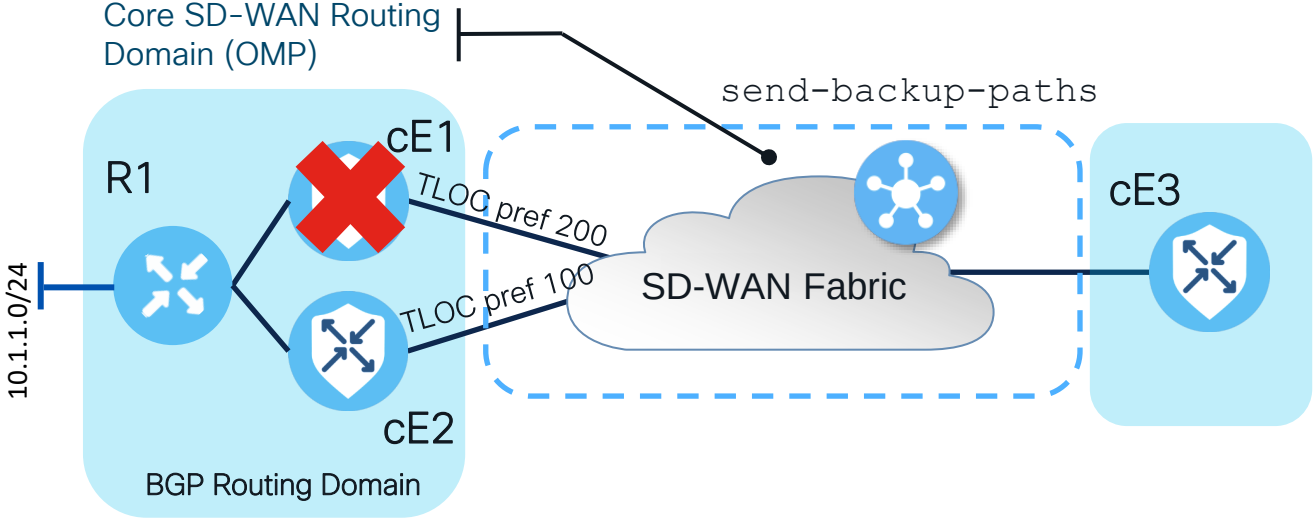


OMP table on cE3:
10.1.1.0/24 via **10.0.0.1** (C,I,R)
10.1.1.0/24 via **10.0.0.2** (R)

SD-WAN router	site-id	system-ip
cE1	12	10.0.0.1
cE2	12	10.0.0.2
cE3	3	10.0.0.3

Case 12. Temporary blackholing on redundancy recovery

cE1 rebooted – switchover to backup as soon as BFD session down

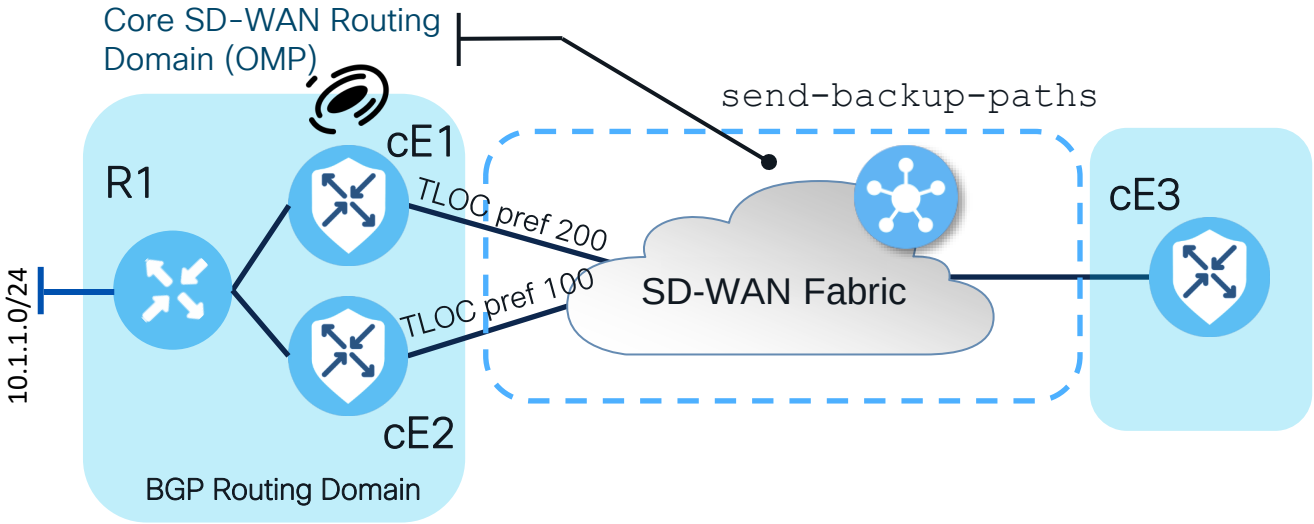


OMP table on vSmart:
10.1.1.0/24 via **10.0.0.1** (C,R,S)
10.1.1.0/24 via **10.0.0.2** (C,R)

OMP table on cE3:
10.1.1.0/24 via **10.0.0.1** (Inv,U)
10.1.1.0/24 via **10.0.0.2** (C,I,R)

Case 12. Temporary blackholing on redundancy recovery

cE1 restarted – switchover back to primary path cE1 as soon as BFD session up - blackhole



OMP table on vSmart:
10.1.1.0/24 via **10.0.0.1** (C,R,S)*
10.1.1.0/24 via **10.0.0.2** (C,R)

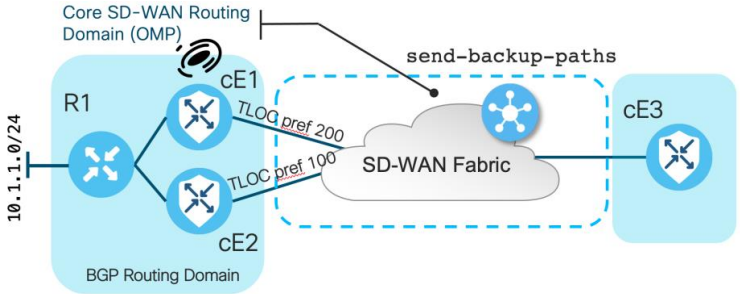
OMP table on cE3:
10.1.1.0/24 via **10.0.0.1** (C,I,R)
10.1.1.0/24 via **10.0.0.2** (R)

*S because eor-timer hasn't expired yet

Case 12. Temporary blackholing on redundancy recovery

Outputs from cE1 taken after reload once BGP and OMP peering re-established:

```
cEdge1#show bgp vpnv4 unicast vrf 1
BGP table version is 1, local router ID is 172.16.2.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
              t secondary path, L long-lived-stale,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

     Network          Next Hop            Metric LocPrf Weight Path
Route Distinguisher: 1:1 (default for vrf 1)
 *    10.1.1.0        192.168.1.1              0    100      0 65001 i
```

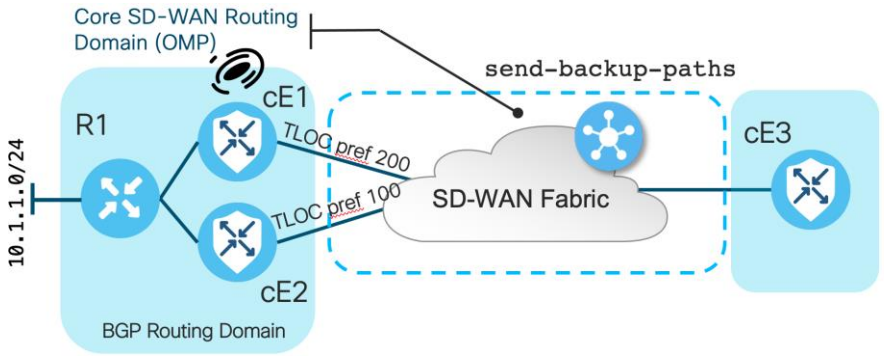


Nov 17 14:59:32.957: %BGP-5-ADJCHANGE: neighbor 192.168.1.1 vpn vrf 1 Up

...

Nov 17 14:59:50.759: %Cisco-SDWAN-cEdge2-OMPD-5-NTCE-400002: R0/0: OMPD: vSmart peer 10.10.10.1 state changed to Up

Case 12. Temporary blackholing on redundancy recovery

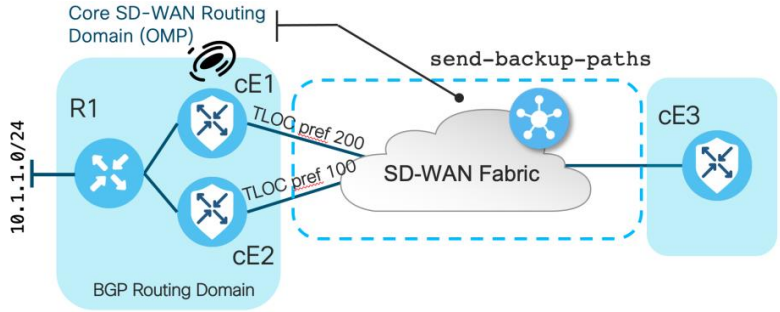


Why traffic is getting blackholed?

- Stale route via cE1 still advertised to cE3 from vSmart controller because of **`send-backup-paths`** and has better preference over route via cE2
- cE3 installs 10.1.1.0/24 into the routing table via cE1 as soon as TLOC advertisement reaches to cE3 and data plane tunnel re-established
- cE1 does not have BGP route to 10.1.1.0/24 installed into the RIB and FIB for about 60 seconds

Case 12. Temporary blackholing on redundancy recovery

Solution to blackholing



- Expected behaviour for BGP and not a SD-WAN failure
- BGP has a initial update (best-path) timer which is 120 sec by default before initiating the best path selection
- Can be reduced under bgp process as below:

```
cEdge1(config-router)# bgp update-delay ?
Possible completions:
  <0..65535, 1 .. 3600>[120]
```

Q&A

References
and recommended resources

- Cisco Troubleshooting Tech Notes: https://www.cisco.com/c/en/us/support/routers/sd-wan/products-tech-notes-list.html

Complete your Session Survey

- Please complete your session survey after each session. Your feedback is important.
- Complete a minimum of 4 session surveys and the Overall Conference survey (open from Thursday) to receive your Cisco Live t-shirt.

- All surveys can be taken in the Cisco Events Mobile App or by logging in to the Session Catalog and clicking the "Attendee Dashboard" at https://www.ciscolive.com/emea/learn/sessions/session-catalog.html

Continue Your Education

Visit the Cisco Showcase for related demos.

Book your one-on-one Meet the Engineer meeting.

Attend any of the related sessions at the DevNet, Capture the Flag, and Walk-in Labs zones.

Visit the On-Demand Library for more sessions at ciscolive.com/on-demand.


cisco
The bridge to possible
Thank you
cisco Live!


CISCO Live!
ALL IN