VOS Edge Routing Protocols

For supported software information, click here.

The Versa Operating System™ (VOS™) Edge device supports a range of routing protocols. This articles describes some common use cases for the routing protocols.

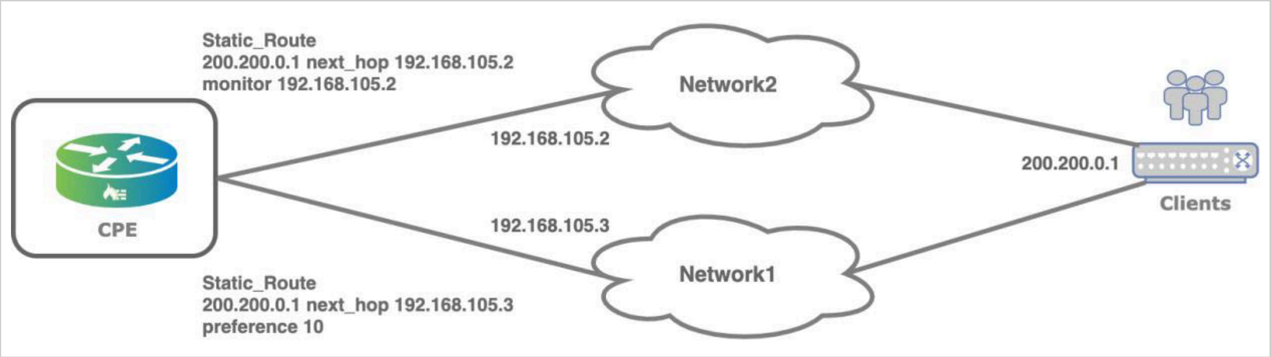
Static Routing

The VOS software provides the following additional features for actions to take when handling static routes:

- Next-hop interface or IP address
- Attach monitor object
- ICMP-based monitors
- Change metric
- Change protocol preference
- Enable BFD
- No-install

Floating Static Routes

Static routes can be conditionally withdrawn based on the state of another target IP address, as illustrated in the following figure.

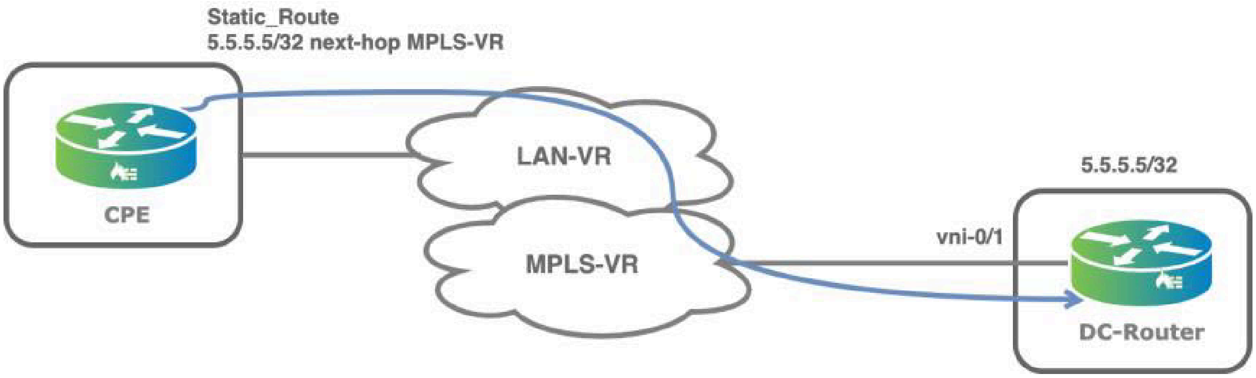


The following output shows a scenario in which the next-hop IP address is monitored on the primary link, while the second static route is configured with a lower preference. The output shows that the primary static route is withdrawn because the monitor probe fails. Monitor objects can be based DNS, ICMP, or TCP.

```
admin@Router-cli> show route routing-instance Router-DC | grep 200.200.0.1/32
static N/A +200.200.0.1/32 192.168.105.2 00:01:42 vni-0/2.0
static N/A 200.200.0.1/32 192.168.105.3 00:04:00 vni-0/2.0
[ok][2020-06-15 07:46:04]

admin@Router-cli> show monitor brief
NAME            ADDRESS       VRF       TENANT       STATE  TYPE
-------------------------------------------------------------------
Monitor-Network1 192.168.105.2   Router-DC  Provider-Org  Up     icmp

[ok][2020-06-15 07:46:07]
admin@Router-cli> show monitor brief
NAME             ADDRESS      VRF       TENANT       STATE  TYPE
--------------------------------------------------------------------
Monitor-Network1  192.168.105.2  Router-DC  Provider-Org  Down   icmp

[ok][2020-06-15 07:46:09]
admin@Router-cli> show route routing-instance Router-DC | grep 200.200.0.1/32
static N/A +200.200.0.1/32 192.168.105.3 00:04:09 vni-0/2.0
```

Route Leaking

The static route next hop can be resolved to the same routing instance or to a different routing instance, as illustrated by the following figure, which shows that for traffic originating from the LAN, the next hop is resolved to the transport VR.



The following snippet shows an example of how to configure route leaking with static routes:

```
admin@PocBr1-cli> show configuration routing-instances PoC-Org-LAN-VR routing-options
static {
    route {
        5.5.5.5/32 MPLS-Transport-VR none {
            preference 1;
        }
    }
}
```

The route table shows the following static route:

```
admin@PocBr1-cli> show route routing-instance PoC-Org-LAN-VR | grep 5.5.5.5/32
static N/A +5.5.5.5/32 0.0.0.0 00:20:45 Indirect
```

The following sample output uses the **tcpdump** command to capture the ping traffic on the data center router:

```
admin@Router-cli> tcpdump vni-0/1 filter "host 5.5.5.5"
Starting capture on vni-0/1
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on _vni_0_1, link-type EN10MB (Ethernet), capture size 262144 bytes
03:23:06.084351 52:54:00:f1:38:4c > 52:54:00:e7:c3:9c, ethertype IPv4 (0x0800), length 98: 169.254.0.202 > 5.5.5.5: ICMP echo request, id 38, seq 75, length 64
03:23:07.088364 52:54:00:f1:38:4c > 52:54:00:e7:c3:9c, ethertype IPv4 (0x0800), length 98: 169.254.0.202 > 5.5.5.5: ICMP echo request, id 38, seq 76, length 64
03:23:08.088510 52:54:00:f1:38:4c > 52:54:00:e7:c3:9c, ethertype IPv4 (0x0800), length 98: 169.254.0.202 > 5.5.5.5: ICMP echo request, id 38, seq 77, length 64
03:23:09.092360 52:54:00:f1:38:4c > 52:54:00:e7:c3:9c, ethertype IPv4 (0x0800), length 98: 169.254.0.202 > 5.5.5.5: ICMP echo request, id 38, seq 78, length 64
03:23:10.096368 52:54:00:f1:38:4c > 52:54:00:e7:c3:9c, ethertype IPv4 (0x0800), length 98: 169.254.0.202 > 5.5.5.5: ICMP echo request, id 38, seq 79, length 64
^C
```

Dynamic Routing

This section describes best practices for using dynamic routing on VOS edge devices for fast convergence, scalability, loop prevention, and security.

Fast Convergence on the SD-WAN Side

For dual-homed deployments, you should configure a unique route distinguisher (RD) for each CPE LAN-VRF, to allow the dual-homed CPEs to advertise unique information for the same local prefix. In this way, the BGP best-path computation done on the Controller nodes reflects the prefixes from both CPEs in the dual-homed architecture. Having both sets of prefixes improves convergence, because it is not necessary to resend a BGP update when a failure occurs. It also helps in troubleshooting, because having a unique route distinguisher makes it much easier to map to the CPE from which the prefix originated. Director Workflows typically configure the same route distinguisher for all nodes. However, note that the Workflows automatically set unique route distinguishers when you configure an HA pair.

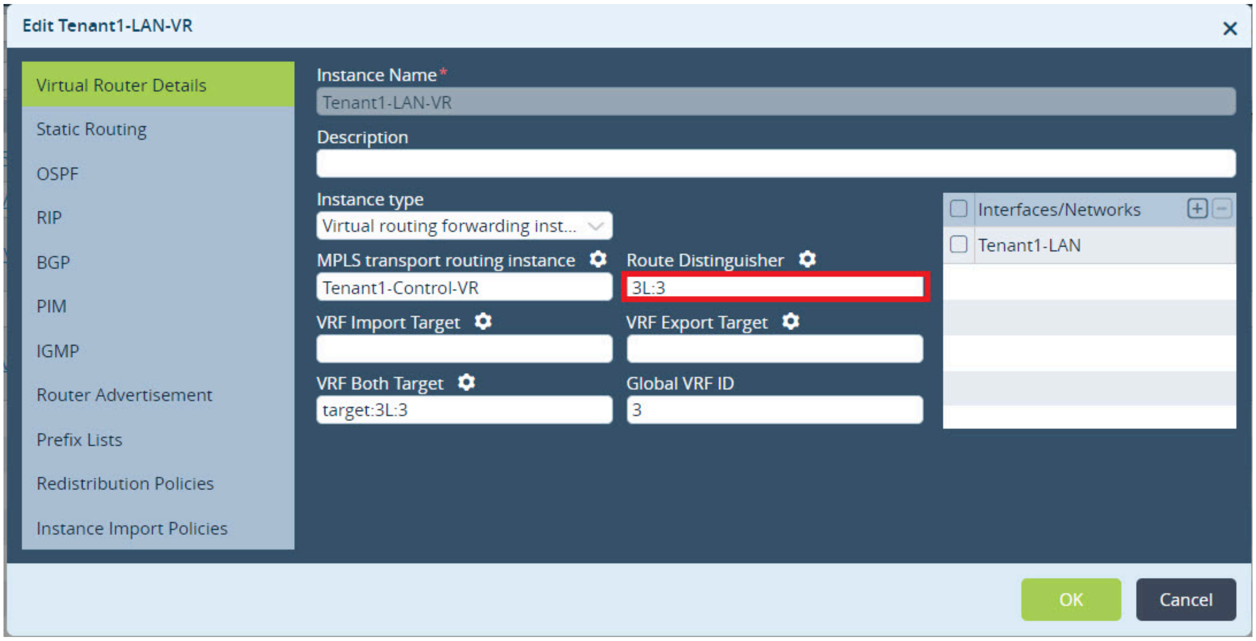
Note: In an EVPN multihoming context, the BGP EVPN route distinguisher should be different for VTEP endpoints that have the same ESI values.

To configure route distinguishers:

1. In the Director view:
    a. Select the Configuration tab in the top menu bar.

b. Select Devices > Devices in the horizontal menu bar.
c. Select an organization in the left menu bar.
d. Select a branch in the main pane. The view changes to Appliance view.

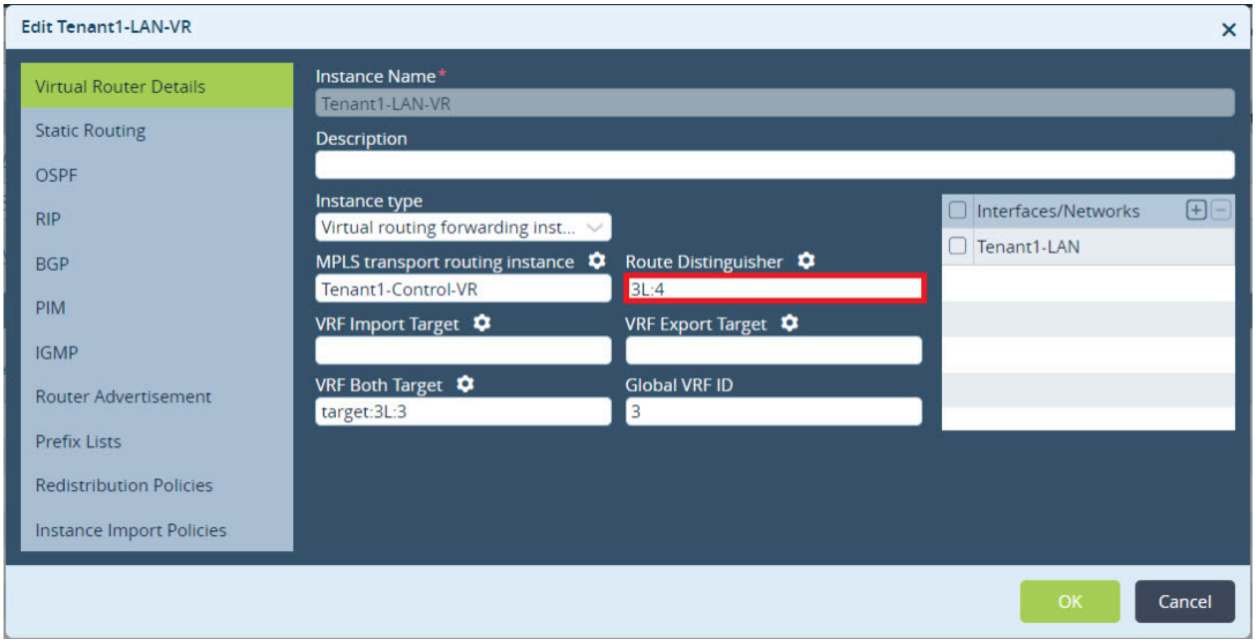
2. Select the Configuration tab in the top menu bar.
3. Select Networking > Virtual Routers in the left menu bar.
4. Select a LAN-VR in the main pane. The Edit LAN-VR popup window displays, and the Virtual Router Details tab is selected.
5. In the Route Distinguisher field, enter an RD value for the first CPE.



Instance Name*: Tenant1-LAN-VR
Description:
Instance type: Virtual routing forwarding inst...
MPLS transport routing instance: Tenant1-Control-VR
Route Distinguisher: 3L:3
VRF Import Target:
VRF Export Target:
VRF Both Target: target:3L:3
Global VRF ID: 3

□ Interfaces/Networks
□ Tenant1-LAN

6. In the Route Distinguisher field, enter an RD value for the second CPE.



Instance Name* Tenant1-LAN-VR
Description
Instance type: Virtual routing forwarding inst...
MPLS transport routing instance: Tenant1-Control-VR
Route Distinguisher: 3L:4
VRF Import Target:
VRF Export Target:
VRF Both Target: target:3L:3
Global VRF ID: 3

□ Interfaces/Networks
□ Tenant1-LAN

The following output shows that the Controller nodes reflect the route information that comes from both CPEs.

```
admin@Controller-RR-cli> show route table l3vpn.ipv4.unicast advertising-protocol bgp neighbor-address
10.0.160.101
...
Routing entry for 192.168.4.0/24
   Peer Address : 10.0.160.101
   Route Distinguisher: 3L:3
   Next-hop : 10.0.160.103
   VPN Label : 24704
   Local Preference : 110
   AS Path : N/A
   Origin : Igp
   MED : 0
   Community : [ N/A ]
   Extended community : [ target:3L:3 ]

Routing entry for 192.168.4.0/24
   Peer Address : 10.0.160.101
   Route Distinguisher: 3L:4
   Next-hop : 10.0.160.104
   VPN Label : 24704
   Local Preference : 109
   AS Path : N/A
   Origin : Igp
   MED : 0
   Community : [ N/A ]
   Extended community : [ target:3L:3 ]
```

As a result, on the other sites, two paths are available and installed in the route table. For example:

```
admin@Branch1-cli> show route routing-instance Tenant1-LAN-VR 192.168.4.0
```

```
Routes for Routing instance : Tenant1-LAN-VR AFI: ipv4 SAFI: unicast
[+] - Active Route

Routing entry for 192.168.4.0 (mask 255.255.255.0) [+]
Known via 'BGP', distance 200,
   Redistributing via BGP
   Last update from 10.0.160.103 00:01:17 ago
Routing Descriptor Blocks:
* 10.0.160.103 , via Indirect 00:01:17 ago

Routing entry for 192.168.4.0 (mask 255.255.255.0)
Known via 'BGP', distance 200,
   Redistributing via BGP
   Last update from 10.0.160.104 00:01:17 ago
Routing Descriptor Blocks:
* 10.0.160.104 , via Indirect 00:01:17 ago
```

It is recommended that you not use other techniques in the SD-WAN overlay, such using BFD and adjusting graceful restart timers, because they interfere with the headless operation of the Versa solution.

Fast Convergence on the LAN Side

An important aspect of fast convergence is rapid failure detection. The detection can be based on a link status failure, such as when neighbors are connected through a Layer 2 network. An alternate detection approach is to use BFD, which is a protocol that offers timer-related detection. BFD is a high-speed protocol designed to detect fast link failures in milliseconds. IGPs and BGP are BFD clients, and multiple routing protocols can piggyback a single BFD session. It is recommended that you enable BFD on the VOS CPE devices. and establish BFD sessions with customer-owned Layer 3 devices. On VOS edge device, you can use BFD with BGP, OSPF, RIP, and static routes.

The following example command output shows a BFD session that is established between two OSPF neighbors:

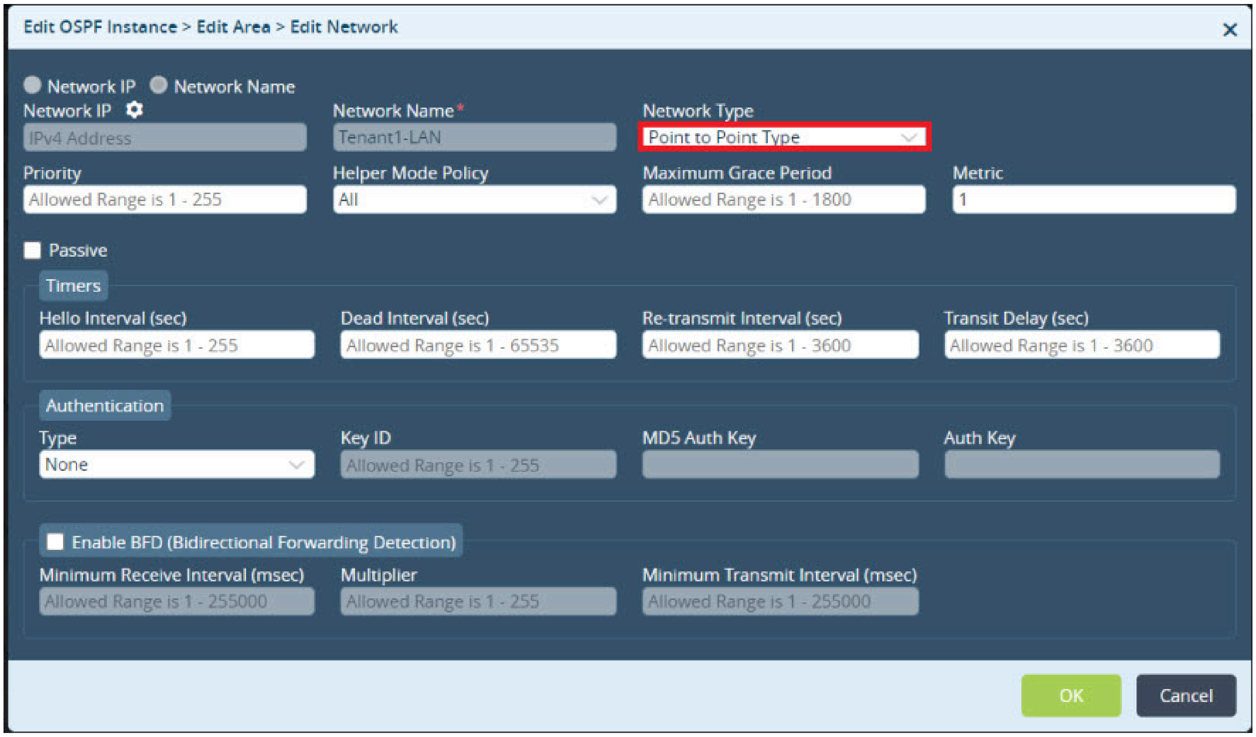
```
admin@branch8-cli> show ospf neighbor brief
State codes: atmpt - attempt, exchg - exchange, exst - exchange start,
          load - loading, 2-way - two-way, full - full
Op codes:    gdown - going down, gup - going up

Intf address    Interface    State    Neighbor ID    Pri    Op
------------    ---------    -----    -----------    ---    --
192.168.254.2    vni-0/4.10    full     1.1.1.1        1      up

admin@branch8-cli> show bfd session org Tenant1 routing-instance-name Tenant1-LAN-VR session-
summary
Instance        Address        State    RxPkts        TxPkts
Tenant1-LAN-VR   192.168.254.2 up       8005          16966
```

Scalability

VOS edge devices support a fully fledged scalable and powerful routing software suite. This section discusses some best practices for configuring BGP and OSPF for when you are. adding. a Versa CPE to an existing network.

OSPF

For Ethernet segments that have only two OSPF routers, configure the network as OSPF as a point-to-point network. Doing so prevents the election of a designated router (DR) and a backup designated router (BDR) on the Ethernet segment and thus prevents the generation and flooding of Type 2 LSAs on the segment. As a result, fewer CPU cycles are used.



Edit OSPF Instance > Edit Area > Edit Network

○ Network IP ○ Network Name

Network IP: IPv4 Address
Network Name*: Tenant1-LAN
Network Type: Point to Point Type

Priority: Allowed Range is 1 - 255
Helper Mode Policy: All
Maximum Grace Period: Allowed Range is 1 - 1800
Metric: 1

□ Passive

Timers
Hello Interval (sec): Allowed Range is 1 - 255
Dead Interval (sec): Allowed Range is 1 - 65535
Re-transmit Interval (sec): Allowed Range is 1 - 3600
Transit Delay (sec): Allowed Range is 1 - 3600

Authentication
Type: None
Key ID: Allowed Range is 1 - 255
MD5 Auth Key:
Auth Key:

□ Enable BFD (Bidirectional Forwarding Detection)
Minimum Receive Interval (msec): Allowed Range is 1 - 255000
Multiplier: Allowed Range is 1 - 255
Minimum Transmit Interval (msec): Allowed Range is 1 - 255000

BGP

BGP peer groups provide a mechanism to associate together BGP peers that have the same outbound policies. The two major benefits of deploying peer groups are a reduction in the required configuration and the ability to replicate updates across peers.

A common reason for using peer groups is to reduce the configuration effort. For peers that require the same outbound policy, you create a peer group. You then apply the common outbound policy to the peer group and you assign each BGP peer to the peer group. For routers that have many BGP peers, using peer groups significantly reduces the configuration required.

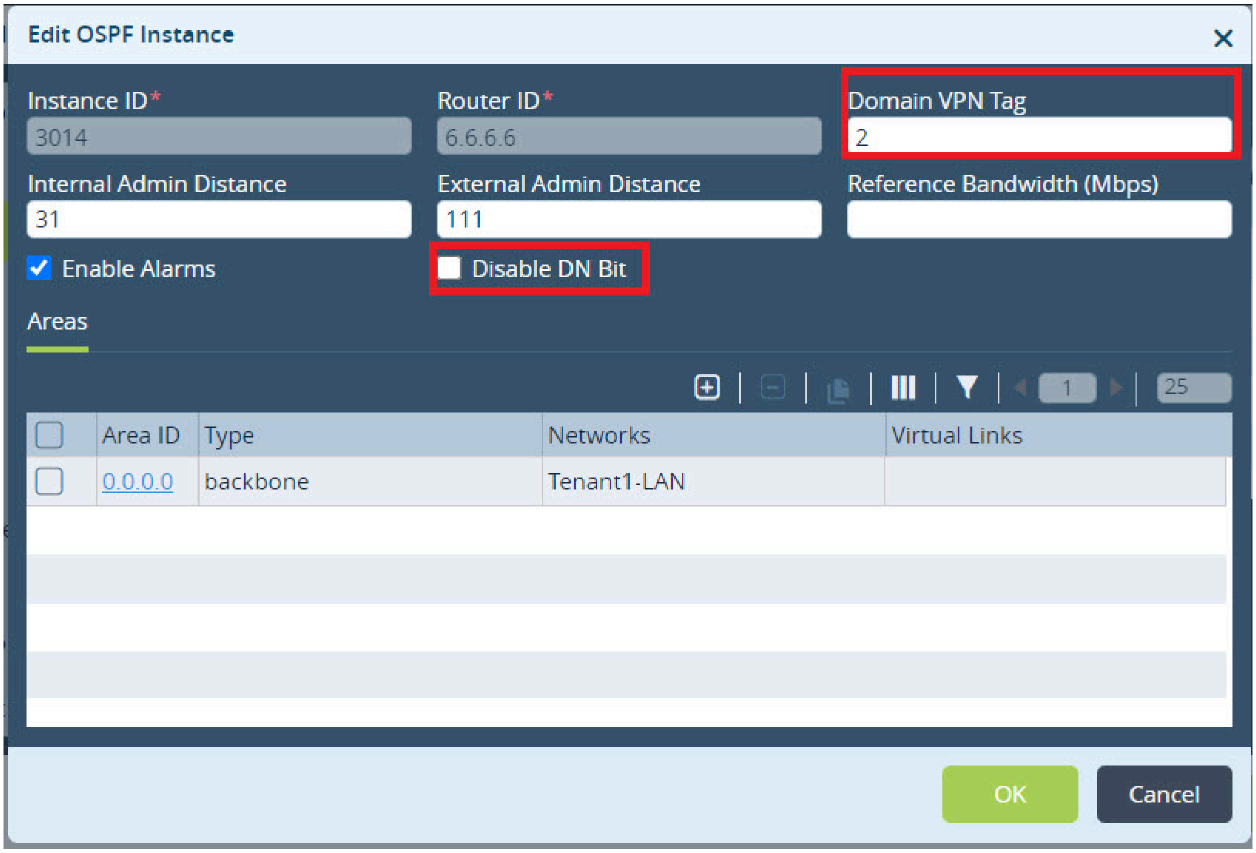
However, the key benefit of using peer groups is the ability to replicate updates across peers. For all peers in the group, the same outbound policy is used, and so the BGP update messages sent to the peers are the same. This mechanism improves BGP scalability, because BGP update messages are generated once for each peer group and then are reused for all the peers in the group. A best practice is to place the BGP peers having the same outbound policy in the same BGP peer group. Versa SD-WAN Controller nodes implement this best practice by default, by having all route reflector clients (CPE branches) be part of the same BGP peer group.

The following sample output shows the outbound BGP policy that is applied for the peer group and hence to all peers in the group:

```
admin@Controller-1-cli> show bgp group brief

routing-instance: Provider-Control-VR
BGP instance: 3
Group Type: Internal                    Local AS: 64512
   Name: Branches
   Options: < PEER-AS EXPORT-POLICY >
   peer-as:
   export-policy: TO_SDWAN
   Total peers: 7        Established: 7

   10.1.192.101+37024
   10.1.192.102+43934
   10.1.192.103+44734
   10.1.192.104+41982
   10.1.192.105+34777
   10.1.192.106+36272
   10.1.192.107+33559
```

Loop Prevention

In many cases, routing loops appear because of mutual route redistribution between two protocols at more than one point. For dual-homed topologies in which mutual redistribution is performed between BGP and OSPF, you should configure a domain VPN tag and enable the DN bit (Down bit) option in the OSPF instance.



Edit OSPF Instance

Instance ID*: 3014

Router ID*: 6.6.6.6

Domain VPN Tag: 2

Internal Admin Distance: 31

External Admin Distance: 111

Reference Bandwidth (Mbps):

☑ Enable Alarms

□ Disable DN Bit

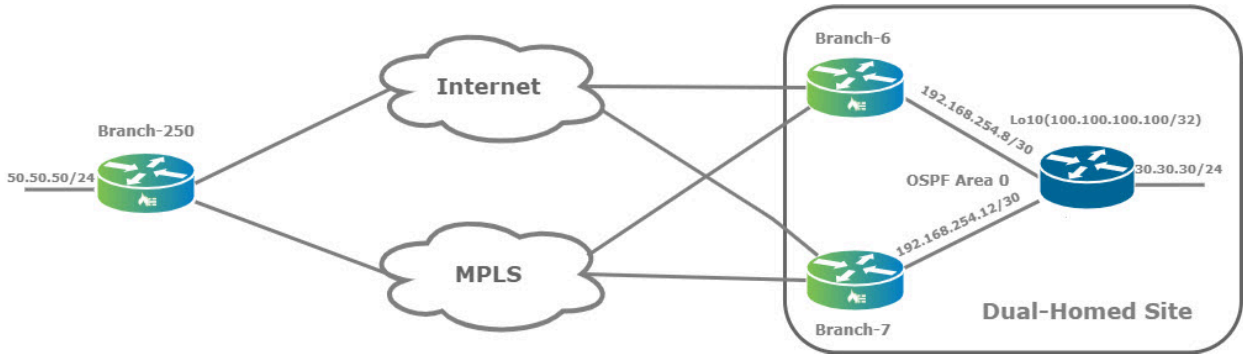
Areas

□	Area ID	Type	Networks	Virtual Links
□	0.0.0.0	backbone	Tenant1-LAN	

Let's talk more about the Down bit and the domain tag to understand how they work in regards to loop prevention. When you use OSPF and BGP in an SD-WAN overlay network, the interaction between the two protocols is similar to that between OSPF and a BGP/MPLS network, as described in RFC 4577. As a summary:

- If the OSPF route has been advertised from a PE router (that is, a VOS edge device) into an OSPF area, the Down (DN) bit is set in the Options field of an OSPF LSA header Type 3, Type 5, or Type 7 LSA to ensure that the route is ignored by any other PE routers that receive them.
- If a particular VRF in a PE is associated with an instance of OSPF, the VPN route tag, domain VPN tag, or OSPF tag is required . In this case, the tag is configured with a special OSPF route tag value that is set in the OSPF Type 5 LSA. For the route tag value, the VOS device uses the VRF instance number of the virtual router (that is, of Tenant-LAN-VR), which is assigned randomly by the Director node and which is the same value across the SD-WAN network. So when a PE router receives the route after it has traversed multiple CE routers, if the PE router has the same VRF instance number, it ignore the routes because it has the same route tag value. Configuring and including the VPN route tag is required for backwards-compatibility with deployed implementations that do not set the DN bit in Type 5 LSAs.

The topology in the following figure explains how to check whether the Down bit and domain tag are set and shows the possible issues that might occur if the VOS CPEs are not directly connected to the local LAN.



In this topology, on the single CPE site (Branch-250), the prefix of the local LAN (50.50.50/24) is redistributed into MP-BGP. On the dual-homed HA active-active site, OSPF runs between the VOS CPE devices (Branch-6 and Branch-7) and a local router.

To prevent routing loops, both the OSPF Down bit and the OSPF domain VPN tag are set when routes are redistributed from MP-BGP into OSPF redistribution. You can see this for the prefix 50.50.50/24 (which comes from Branch-250) when it is redistributed from MP-BGP into OSPF on the Branch-6 and Branch-7 routers:

```
admin@branch6-cli> show ospf database routing-instance Tenant1-LAN-VR external detail
50.50.50.0        6.6.6.6        0x80000001        15        0x0000A795
   Age: 15 secs; Sequence number: 0x80000001; Checksum: 0x0000A795
   Options: DN, E
   Type: E2
   Metric: 100
   Forwarding address: 0.0.0.0
   External Route Tag: 2
50.50.50.0        7.7.7.7        0x80000001        18        0x000089AF
   Age: 18 secs; Sequence number: 0x80000001; Checksum: 0x000089AF
   Options: DN, E
   Type: E2
   Metric: 100
   Forwarding address: 0.0.0.0
   External Route Tag: 2
```

The Down bit and VPN tag are helpful for preventing routing loops prevention. However, a problem arises when the VOS CPEs are not directly connected to the local LAN and another router provides tenant-based LAN connectivity and is an OSPF neighbor with a tenant on the VOS CPE. The following sample output shows a third router that has Branch-6 and Branch-7 as OSPF neighbors in VRF Tenant1:

```
router ospf 1 vrf Tenant1
 router-id 1.1.1.1
 redistribute connected subnets route-map C2OSPF
 passive-interface Ethernet1/3.10

IOU1# show ip ospf 1 neighbor

Neighbor ID      Pri    State        Dead Time       Address         Interface
7.7.7.7           0     FULL/  -     00:00:38        192.168.254.13  Ethernet0/2.10
```

```
6.6.6.6          0      FULL/  -      00:00:30        192.168.254.9    Ethernet0/1.10
```

Because the Down bit is set in the Type 5 LSA for 50.50.50/24 and the existing router is already VRF aware (it is acting as a PE), the LSA is not considered for SPF calculation and the prefix is not added to the router's route table. This breaks the connectivity for the dual-homed site to remote locations:

```
IOU1# show ip ospf 1 database external 50.50.50.0

            OSPF Router with ID (1.1.1.1) (Process ID 1)

                Type-5 AS External Link States
    LS age: 1524
    Options: (No TOS-capability, No DC, Downward)
    LS Type: AS External Link
    Link State ID: 50.50.50.0 (External Network Number )
    Advertising Router: 6.6.6.6
    LS Seq Number: 80000001
    Checksum: 0xA795
    Length: 36
    Network Mask: /24
        Metric Type: 2 (Larger than any link state path)
        MTID: 0
        Metric: 100
        Forward Address: 0.0.0.0
        External Route Tag: 2

    LS age: 1525
    Options: (No TOS-capability, No DC, Downward)
    LS Type: AS External Link
    Link State ID: 50.50.50.0 (External Network Number )
    Advertising Router: 7.7.7.7
    LS Seq Number: 80000001
    Checksum: 0x89AF
    Length: 36
    Network Mask: /24
        Metric Type: 2 (Larger than any link state path)
        MTID: 0
        Metric: 100
        Forward Address: 0.0.0.0
        External Route Tag: 2
```

However, the prefix 50.50.50/24 is not present in the route table:

```
IOU1# show ip route vrf Tenant1 50.50.50.0

Routing Table: Tenant1
% Network not in table
```

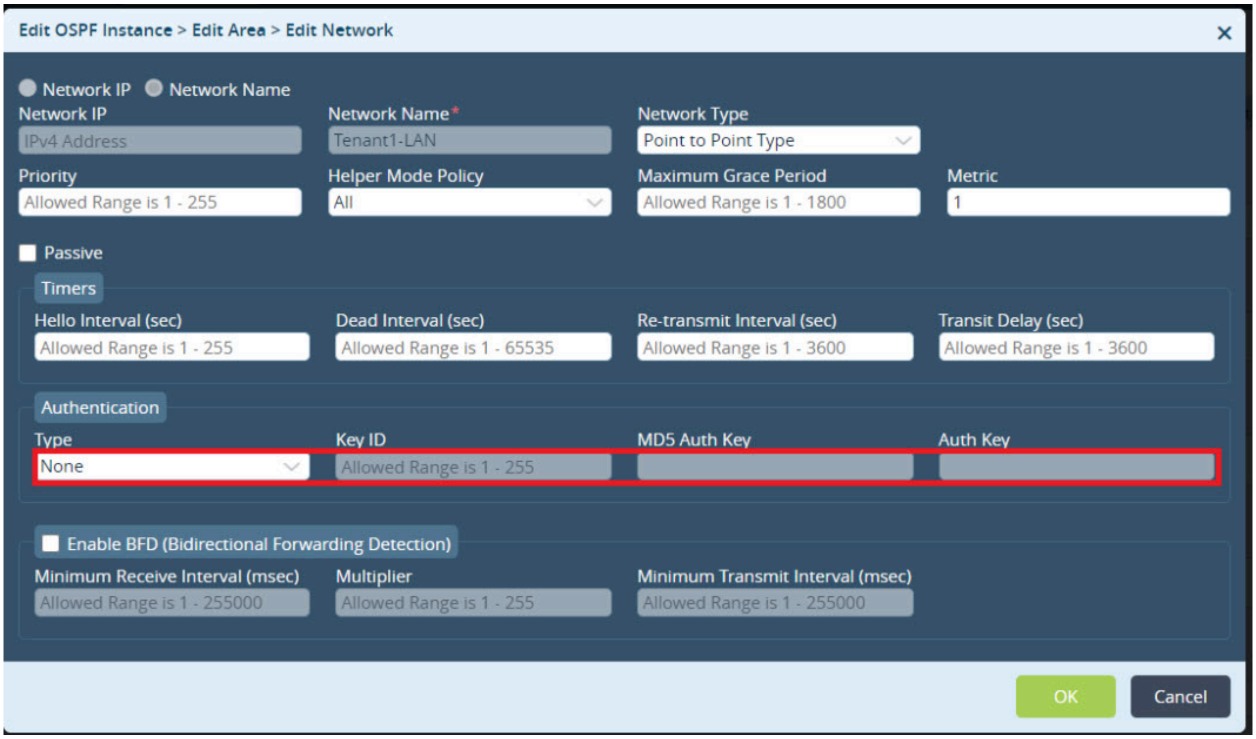
The recommended solution is to check whether the intermediary router (here, IOU1) has the capability to ignore the Down bit set in the LSA option field and can instead use the LSA in the SPF algorithm, so that it can add the prefix to its route table. Most vendors call this feature VRF-lite. If the existing customer router does not support this capability, you should disable the Down bit setting on VOS devices. In this situation, you must be especially careful when mutual route redistribution between MP-BGP and OSPF is performed.

Routing Security

IGP Security

Authentication is the most important measure you can take to secure any routing protocol. The VOS OSPF and RIP protocols support clear-text and HMAC-MD5 authentication, which are simply mechanisms by which two neighbors prove their identity to each other by using a shared secret. No protocol messages are not accepted from a neighbor unless the message is correctly authenticated. Authentication is also useful in other situations, for example, to prevent routers from mistakenly joining an OSPF domain.

A best practice is to configure IGP HMAC-MD5 authentication on all interfaces on which the IGP runs. The following screen shows how to configure authentication for OSPF.



Edit OSPF Instance > Edit Area > Edit Network ×

○ Network IP ○ Network Name

Network IP	Network Name*	Network Type	
IPv4 Address	Tenant1-LAN	Point to Point Type	
Priority	Helper Mode Policy	Maximum Grace Period	Metric
Allowed Range is 1 - 255	All	Allowed Range is 1 - 1800	1

□ Passive

Timers

Hello Interval (sec)	Dead Interval (sec)	Re-transmit Interval (sec)	Transit Delay (sec)
Allowed Range is 1 - 255	Allowed Range is 1 - 65535	Allowed Range is 1 - 3600	Allowed Range is 1 - 3600

Authentication

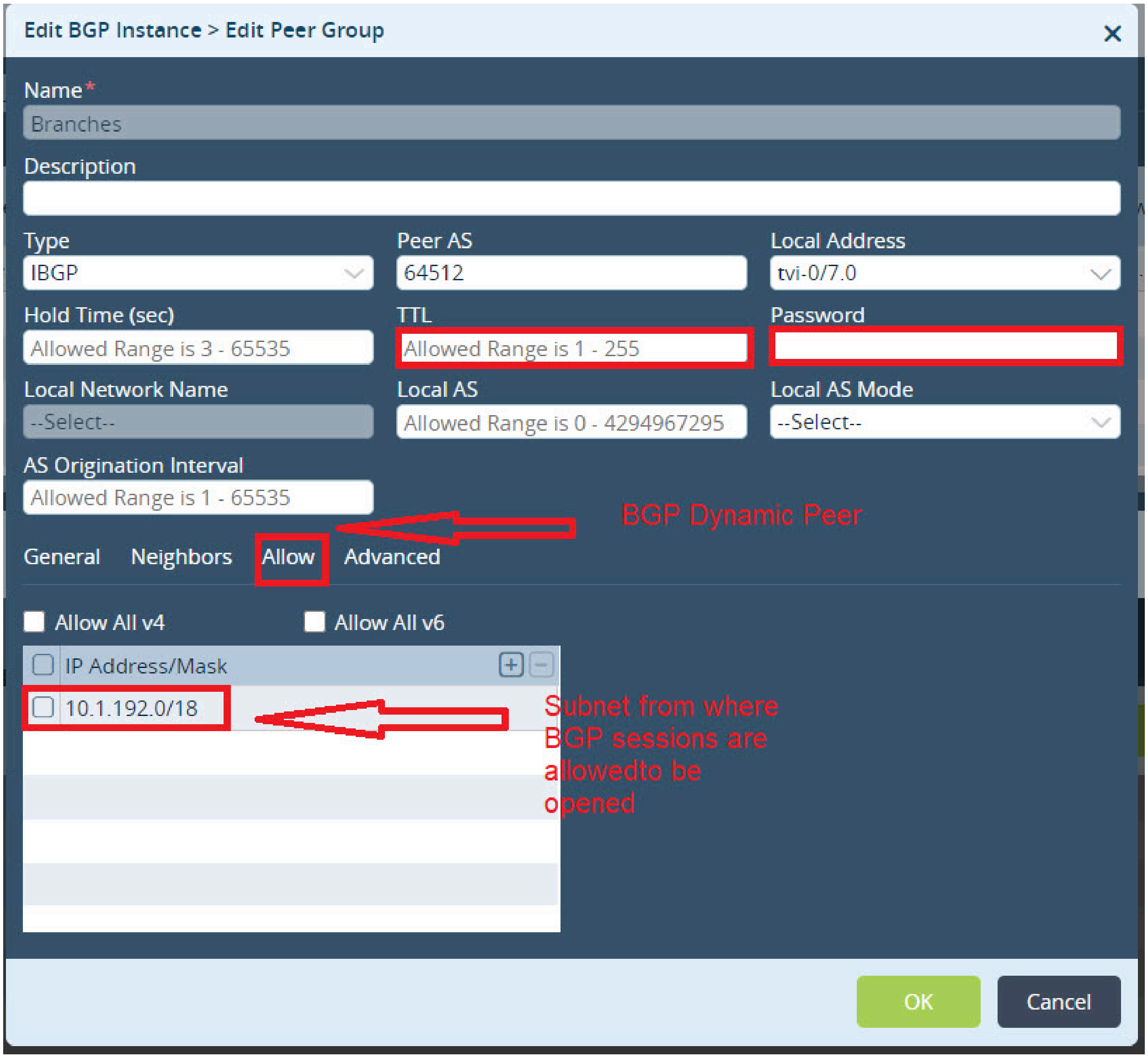
Type	Key ID	MD5 Auth Key	Auth Key
None	Allowed Range is 1 - 255		

□ Enable BFD (Bidirectional Forwarding Detection)

Minimum Receive Interval (msec)	Multiplier	Minimum Transmit Interval (msec)
Allowed Range is 1 - 255000	Allowed Range is 1 - 255	Allowed Range is 1 - 255000

BGP Security

The VOS edge device software implement the BGP dynamic neighbors feature, which is useful when you need to configure many BGP peers. This feature allows dynamic BGP peering to a group of remote neighbors that are defined by a range of IP addresses. You configure each range as a subnet IP address. You configure the BGP dynamic neighbors in BGP peer groups, so you do not need to configure the peers individually except for configuring their subnets. However, because you do not configure the peers individually, you must take measures to prevent security issues. When you use BGP dynamic neighbors, it is recommended that you configure BGP authentication and TTL for the BGP session, to prevents a rogue device from establishing a BGP session by using the subnet configured for the dynamic neighbor. The following screen illustrates the configuration:



Edit BGP Instance > Edit Peer Group

Name*		
Branches		
Description		
Type	Peer AS	Local Address
IBGP	64512	tvi-0/7.0
Hold Time (sec)	TTL	Password
Allowed Range is 3 - 65535	Allowed Range is 1 - 255	
Local Network Name	Local AS	Local AS Mode
--Select--	Allowed Range is 0 - 4294967295	--Select--
AS Origination Interval		
Allowed Range is 1 - 65535		

BGP Dynamic Peer

General | Neighbors | Allow | Advanced

☐ Allow All v4 ☐ Allow All v6

☐	IP Address/Mask
☐	10.1.192.0/18

Subnet from where BGP sessions are allowedto be opened

Best Practices for Dynamic Routing

- For improved convergence in a dual-homed branch, use a unique route distinguisher for each LAN-VRF to advertise unique information for the same local prefix.
- Configure the OSPF network as a point-to-point network to prevent the use of unnecessary CPU cycles.
- Configure BGP peers as part of the same BGP peer group so that the peers use the same outbound policy.
- Enable HMAC-MD5 authentication on all interfaces on which an IGP runs.
- Configure BGP authentication and TTL for BGP sessions.

Supported Software Information

Releases 20.2 and later support all content described in this article.

Additional Information

Configure Virtual Routers