Consolidated SD-WAN Design Guide

Introduction to SD-WAN Design

For supported software information, click here.

The Versa Networks SD-WAN solution is a highly robust and flexible platform that offers capabilities to address various SD-WAN use cases.

This series of SD-WAN design articles addresses the most common SD-WAN use cases, and they describe the Versa Networks recommendations and best practices for SD-WAN deployments. The objective is to help achieve a standardized approach to designing Versa Networks SD-WAN solutions. Note that these articles are not meant to cover detailed operational best practices nor to act as a service management manual.

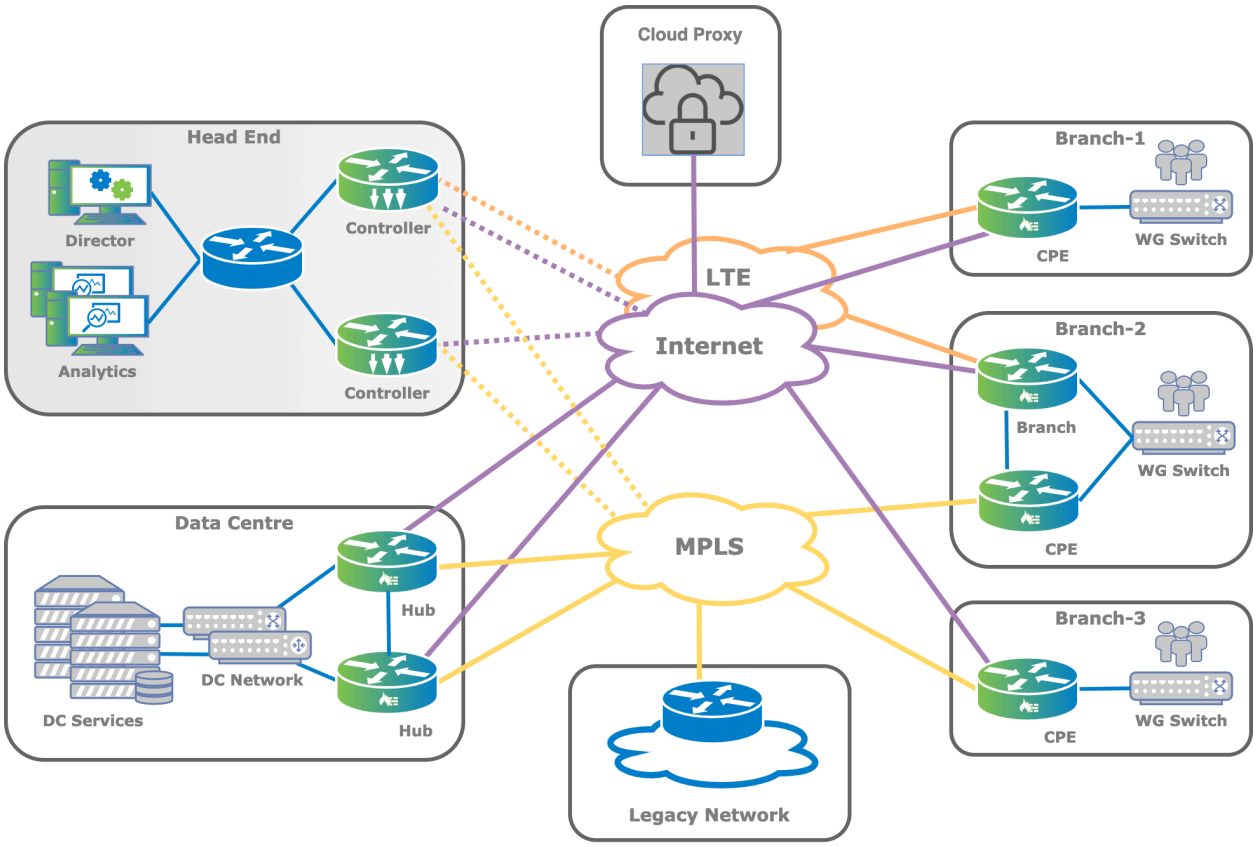
These articles are targeted at network architects, engineers, administrators, and other technical audiences interested in designing, implementing, and deploying Versa Networks SD-WAN solutions. These articles assume that you are familiar with the basics of Versa Networks products and that you have a working knowledge of the Versa Network SD-WAN architecture as well as the wider Versa Networks ecosystem.

These articles are intended to be only a generic guide so that you can use them to explore use cases for your specific deployment. The articles do not cover every use case that the Versa Networks secure SD-WAN solution supports.

The designs and best practices described in these articles are based on the Release 20.2.2 software. While you can use these designs in networks running Release 16.1R2 software, not all features are available in the earlier software release.

Reference Network Architecture

The designs in these SD-WAN articles are based on the reference network architecture shown in the following figure. The figure shows a high-level blueprint of a typical network topology built using the Versa Networks SD-WAN solution. This topology has one headend, one data center, and three remote branches. The headend consists of a Versa Analytics cluster, a Versa Director, and two Versa Controller nodes. Other components of the topology are a legacy network and a cloud proxy. There are three transport networks in the figure, MPLS, Internet, and LTE, and they are provided by one or more service providers. The network orchestration provided by the headend is reachable through Versa Networks Controller nodes, which are connected to all transport networks. Versa Director and Versa Analytics are hosted in the headend.



SD-WAN Design Guide Articles

This SD-WAN design guide includes the following articles:

- SD-WAN Headend Design Guidelines—Provides architectural and deployment guidelines for the Versa Networks Versa Analytics cluster, a Versa Director, and a Versa Controller headend components.
- Branch Deployment Options—Describes the most common transport-based branch deployment scenarios.
- LTE Transport Modes—In many scenarios, LTE is deployed as a backup path because its data costs are higher. This article discusses the LTE hot standby and cold standby mode.
- SD-WAN Overlay Networks—Discusses the IP addressing scheme for the overlay network and whether to have data traffic to follow the encrypted overlay,
- SD-WAN Topologies—Discusses the SD-WAN overlay topologies, which are full mesh, hub and spoke, regional mesh, and multi-VRF (also called multitenancy).
- VOS Edge Device Direct Internet Access—Describes the breaking out of traffic to the internet at the branch, which is called direct internet access, or DIA.
- SD-WAN Gateway Use Cases—Describes the three main use cases for VOS edge devices that act as SD-WAN gateways, which are connecting to sites on an MPS Layer 2 VPN network, connecting to sites over disjointed underlay networks, and acting as a gateway for internet-bound traffic.
- SD-WAN Traffic Optimization—Discusses how to use traffic steering, traffic conditioning, and SD-WAN path policies to optimize SD-WAN traffic flow.
- VOS Edge Routing Protocols—Describes common use cases for static and dynamic routing protocols.

- QoS—Discusses how to configure quality of service (QoS), also known as class of service, or CoS, to ensure that the network prioritizes business critical traffic over less important traffic and treat this traffic with higher priority. You can also use QoS for other tasks, such as policing, shaping, and remarking the QoS bits in the IPv4/IPv6 and VLAN headers.

Supported Software Information

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

Additional Information

Branch Deployment Options
LTE Transport Modes
QoS
SD-WAN Gateway Use Cases
SD-WAN Headend Design Guidelines
SD-WAN Overlay Networks
SD-WAN Topologies
SD-WAN Traffic Optimization
VOS Edge Device Direct Internet Access
VOS Edge Routing Protocols

SD-WAN Headend Design Guidelines

For supported software information, click here.

This article provides architectural and deployment guidelines for the Versa Networks headend components. The headend consists of a Versa Analytics cluster, a Versa Director, and a Versa Controller.

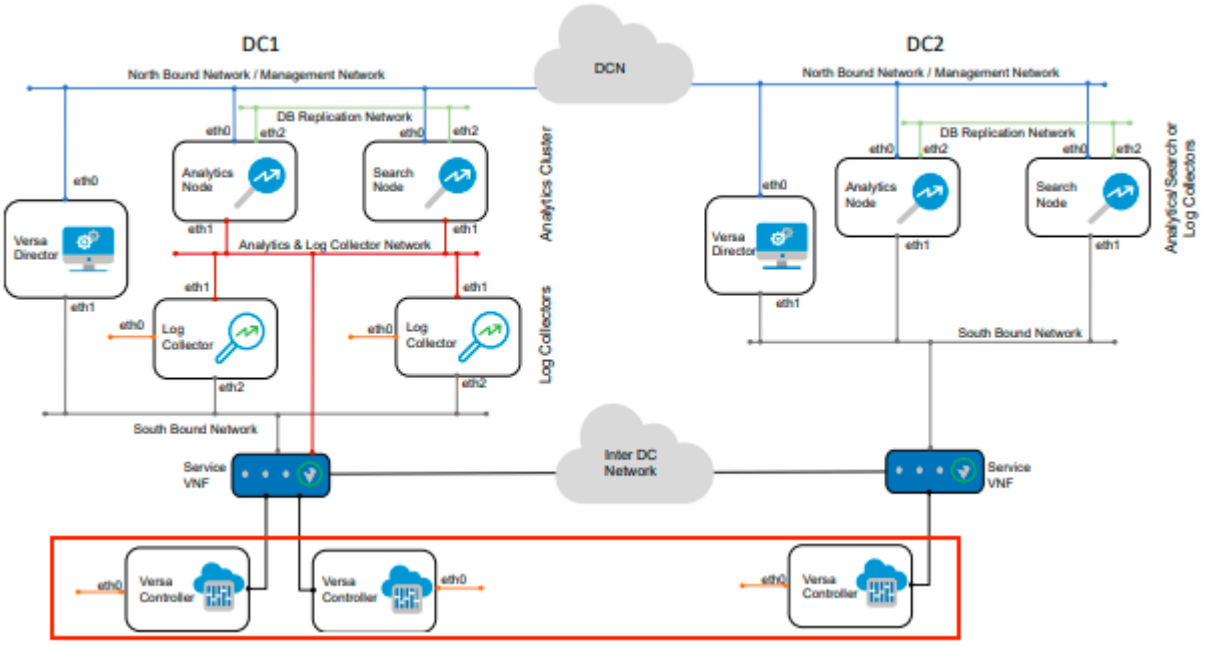
Design Considerations for Headend Architecture

A typical Versa SD-WAN headend architecture is fully resilient, providing geographical redundancy. When you deploy the headend in a public cloud, you deploy it in different availability zones. For these types deployments, the infrastructure must provide the capability for interconnectivity among the headend components.

The SD-WAN headend topology has the following components:

- Northbound segment, for management traffic
- Southbound segment, for control traffic
- Service VNF router to interconnect data centers

The following figure illustrates the headend components, which shows a topology that has headend components in two geographically separated data centers.



Administrators and other OSS systems use the northbound network to access the Versa Director, Analytics GUI, and REST APIs. The northbound network is also used to maintain synchronization between Versa Directors in a high availability cluster. Controller nodes require the northbound network to connect to Versa Director over the out-of-band (eth0) interface so that the Versa Director can provision the Controller nodes.

The southbound network, or control segment, connects to the SD-WAN overlay through the Controller nodes. Because the Director and Analytics nodes are hosts and not routers, the topology requires an additional router to provide a redundant connection to both Controller nodes. A dynamic routing protocol (either OSPF or BGP) must be enabled on this router to provide accurate overlay IP prefix reachability status, and BFD must be enabled towards the service VNF device in the data center. This additional router can be an existing router in the data center network, or it can be a VOS edge device managed by Versa Director.

The solution assumes reachability between the data center networks over the Director northbound operations support system (OSS) network. If the control network service router experiences an outage, the Director node can be accessible using the northbound network, and and the administrator can gracefully fail over to the standby Director node.

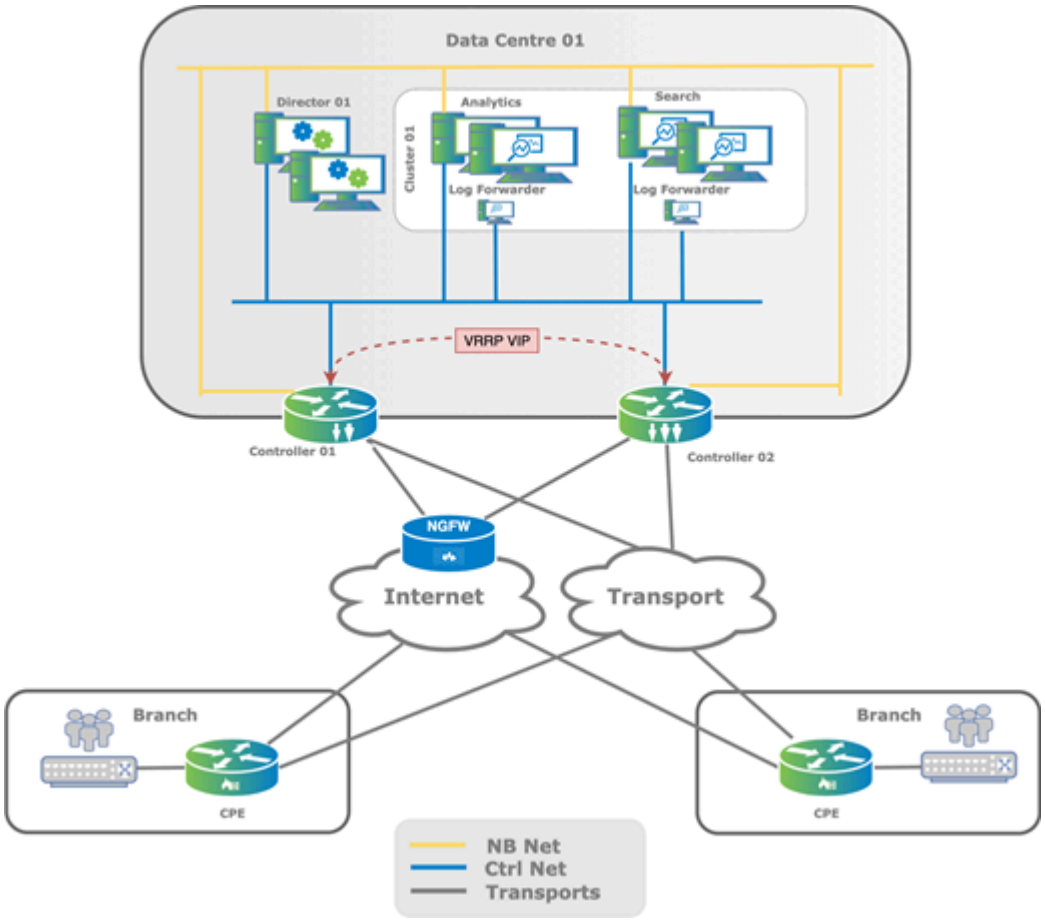
Note: It is not mandatory to always have the Controller nodes connect to all transport domains. Technically, the SD-WAN network can function if an SD-WAN branch device is connected to the Controller node over just one underlay. However, from a resiliency point of view, it is recommended that the Controller node connect directly to all the transport domains that you use.

Note: The Director northbound OSS network must be separate from the control network to avoid the possibility that a split-brain state arises between the Director nodes. In a split-brain scenario, both Director nodes are up but network

reachability between them is lost. The results is that both Director nodes assume that they are the active nodes.

If internet access is enabled to the Versa Director then you must secure the northbound traffic with appropriate security. Where users are able to access Versa Director directly from the internet such as the cloud-hosted Versa Directors, you must use a third dedicated interface into a DMZ. This allows separation between the link used for Director Sync and the user internet access to the Director.

The following figure shows a simple single data center deployment that is not geographically dispersed and in which service VNF routers are not required. This topology uses flat Layer 2 LANs for both the northbound and control networks, and it uses VRRP in the control LAN for gateway redundancy. However, this design is not suitable for creating Layer 3 security zones at the headend. If security is required, you must configure firewalls in bridge mode.



Headend Deployment Options

You can deploy the typical headend architectures shown above as host OSs on physical hardware, which is called a bare-metal deployment. You can also deploy these headend architectures in a virtualized architecture, for which Versa Networks supports ESXi and KVM, and the AWS, Azure, and Google Cloud Platform public clouds.

Deploying the headend on bare-metal hardware provides better performance and scalability, typically, 20 to 30 percent

over a virtualized architecture. However, deploying on physical hardware has hardware dependencies. For example, not all server hardware, such as special RAID controllers, is supported.

For both bare-metal and virtualized deployments, it is important to follow the hardware requirements described in Hardware and Software Requirements for Headend.

Firewall Dependencies for Headend Deployment in a Data Center

When you implement the headend in an existing data center, several protocols and functions are required to connect to the data center. You must follow the requirements for connectivity, reachability, and security. For more information on firewall requirements, see Firewall Requirements.

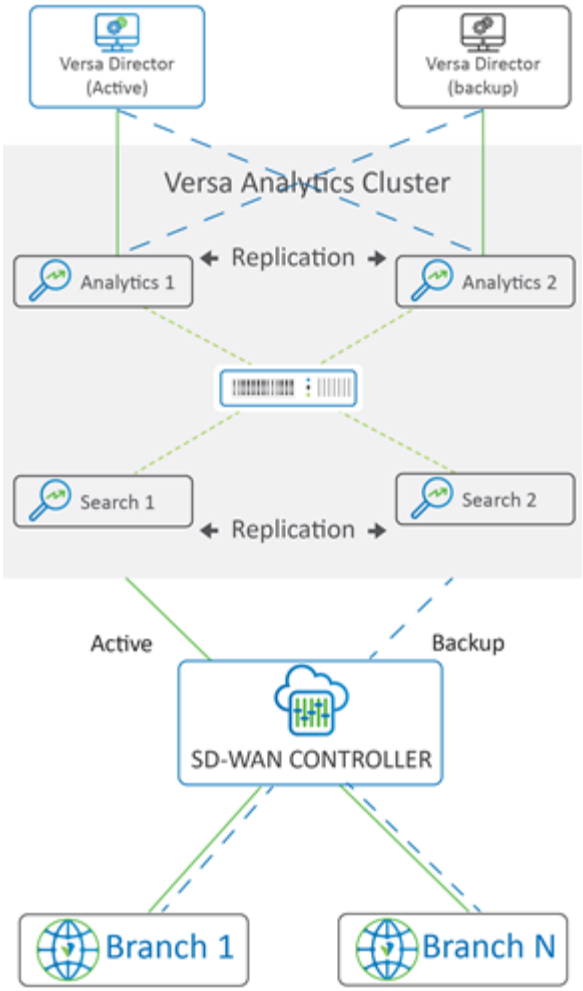
Best Practices for Hardening the Headend

If you deploy Versa Directors that are exposed to the internet so that external users can access them, you must apply additional by using common IT hardening practices. This includes installing Ubuntu OS patches, installing official certificates (Versa Networks software ships with self-signed certificates), and changing the default passwords. Note that you must use the OS patches provided by Versa and not install OS patches directly from Ubuntu.

For more information, see the system hardening articles.

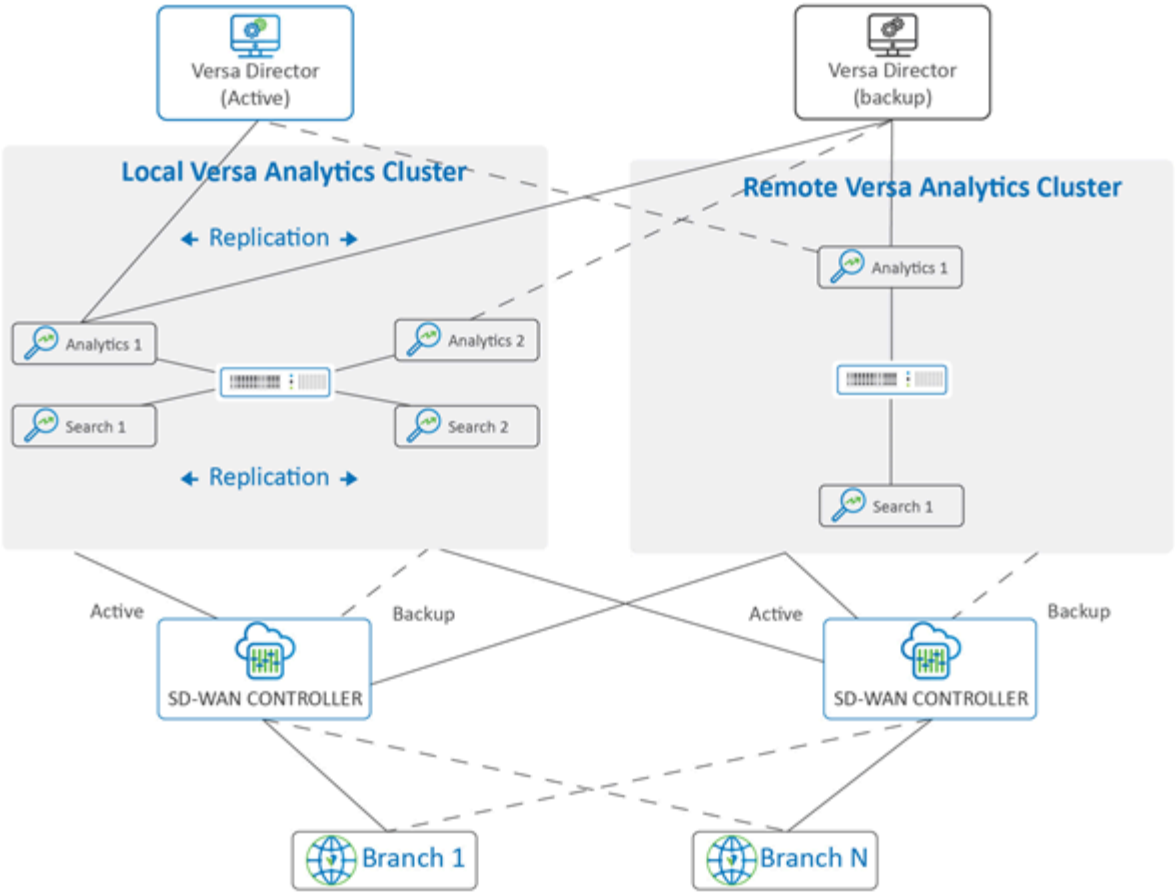
Versa Analytics Deployment Options

The minimum recommended Versa Analytics deployment for full redundancy requires a cluster of four Analytics nodes, as illustrated in the figure below. This design provides active-active high availability (HA), with two nodes having the Search personality and two nodes having the Analytics personality. Redundancy is achieved by replicating the data between one or more nodes of the cluster. Because there is a large amount of data movement between the nodes of the cluster, the network latency must be low for the best storage and query performance. Therefore, it is recommended that you allocate the nodes of the same cluster in the same data center, or at least within the same availability zone. Note that synchronization of the Cassandra database requires less than 10 milliseconds of latency between the nodes in the same cluster.



Recommended Production Analytics Deployment

To provide geographical resilience and also to provide disaster recovery, you can add a second Versa Analytics cluster, as illustrated in the following figure below. The second cluster can have either the same number of nodes as the main cluster or a different number of analytics and search nodes. For disaster recovery, you typically use a smaller Analytics cluster that is activated only if the primary site fails. Note that when you use clusters of different sizes, it is important to check that that the smaller cluster can accommodate the number of LEF connections from the branches.



It is recommended that your topology have a backup Analytics cluster for disaster recovery. If the active Analytics cluster goes down, the backup continues to process the logs and statistics. When the active cluster becomes available again, the logs and statistics from the back can then be synchronized to the active cluster.

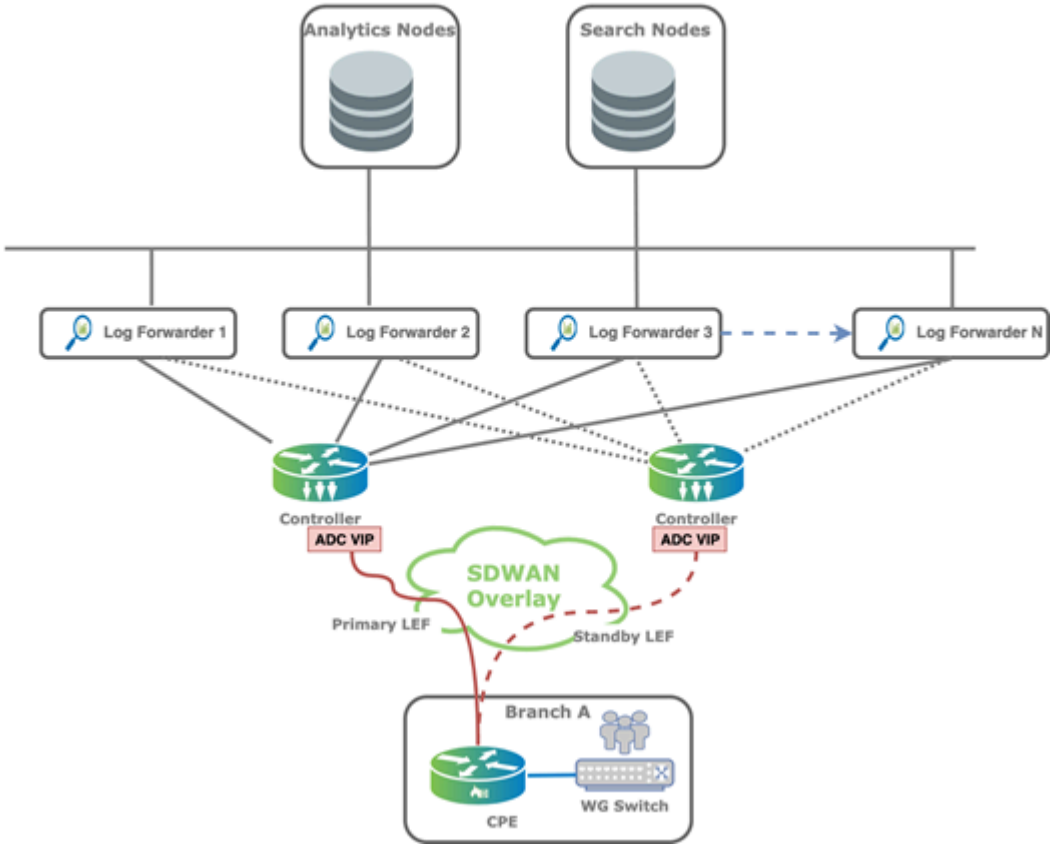
If the topology has multiple Analytics clusters, database replication happens automatically only within the same cluster.

Role of the Log Forwarder

The Versa Operating System™ (VOS™) edge devices generate logs destined for the Analytics database. These logs are carried in an IPFIX template and are delivered to the log forwarder. The log forwarder removes the IPFIX template overhead and stores the logs in clear text on log collector disk. The Analytics database servers can then access the logs from the log forwarder. Optionally, logs can be forwarded in syslog format to third-party log forwarders.

The log forwarder function is part of the Versa Analytics node. It can run on the Versa Analytics node itself or as a standalone log forwarder, where a separate VM or server is dedicated to this role. For larger deployments, especially for deployments with 1000 or more branches, it is recommended that you separate the log forwarder function, because it

provides better scalability for the Analytics platform. When the log forwarder runs on a separate device, you typically deploy it south of the Analytics or Search nodes, as shown in the following figure.



Best Practices for Versa Analytics Sizing

The scaling of a Versa Analytics cluster is driven by the logging configuration for the individual features enabled on VOS edge devices. Most features include an optional logging configuration to log events related to a specific rule or profile. The number of features for which you enable logging and the volume of logs that each feature generates has a direct impact on the sizing and scalability of the Versa Analytics cluster.

Versa Analytics has two distinct database personalities, Analytics node and Search node. The Analytics node delivers the data for most of the dashboards shown in the Analytics GUI. These data are aggregated data, and the data aggregation is enabled by default. Therefore, the dashboards are populated without configuration on the VOS edge devices. Analytics node scaling is determined by granularity of the reported data, which is usually set between 5 and 15 minutes. The topology also affects the scaling of the Dashboard feature. A full-mesh topology with multiple underlays generates more SLA logging messages compared to a hub-spoke topology with a single underlay. This is because SLA monitoring between branches in the underlay contributes significantly to the log volume as it reports SLA measurements to the Analytics node.

The primary tasks of the Search node are to store logging and event data, and to drive the log sections and log tables on the Dashboard. These functions are heavily utilized, because every event in the network triggers a log entry on the Search node. In a minimal default configuration, the only information that is logged to the search nodes are the alarm logs. The sizing of the search node depends on the following:

- Number of features enabled with logging
- Logging data volume and log rate
- Retention period of logged data
- Packet capture—You should enable packet capture and traffic monitoring only for traffic related to specific troubleshooting. Packet captures are not stored in the database, but rather as stored as PCAP files. If these files are not processed correctly, they can quickly fill up the disk space.

Feature such as firewall logging and traffic monitor logging (Netflow) have a significant impact on the overall scaling of the Analytics cluster. Therefore, it is recommended that you avoid creating wildcard logging rules. In response to an increases analytics load, you can scale both the Analytics and Search nodes horizontally. Versa Networks has a calculator tool that can help you size an Analytics cluster based on the information described in the next section.

Best Practices for Headend Design

A stable and fault-tolerant solution requires proper design of the headend. A proper design factors in the expected utilization in order to dimension the compute resources so that they can be horizontally scaled out when required. The headend design must be well thought out high availability design, to ensure that there is no single point of failures. Finally, the headend must be secured and hardened to avoid possible security vulnerabilities.

It is recommended that you consider the following when deploying a Versa Networks headend:

- Ensure that you configure DNS on the Director and Analytics nodes. DNS is used for operations such as downloading security packs (SPacks) an, on Analytics devices, for reverse lookup.
- Ensure that you configure NTP and synchronize it across all nodes in the SD-WAN network. It is recommended that you configure all nodes to be in the same time zone, to make log correlation between various components practical.
- Perform platform hardening procedures such as signed SSL certificates, password hardening, and SSL banners for CLI access.
- Ensure that you have installed the latest OS security pack (OS SPack) on all components and that you have installed the latest SPack on the Director node
- Ensure that the appropriate ports on any intermediate firewalls are open.

For more information, see the hardening articles.

For assistance with the design and validation of the headend before you move it into production, contact Versa Networks Professional Services.

Supported Software Information

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

Additional Information

Firewall Requirements
Hardware and Software Requirements for Headend
Scalability and Performance

Branch Deployment Options

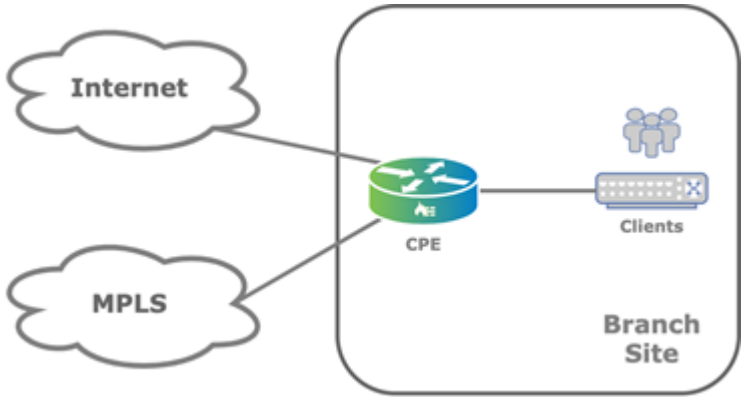
For supported software information, click here.

The Versa Networks SD-WAN solution offers flexible and comprehensive branch deployment options. This article describes the most common transport-based branch deployment scenarios.

One of the benefits with Versa SD-WAN branch configuration is that each Versa Operating System™ (VOS™) edge device provides the same feature capabilities, regardless of whether it is in a hub, spoke, or any other configuration. You can configure different topologies for different tenants on the same edge device, and at the same time.

Branch with a Single CPE Device and Dual Transports

In the first branch deployment scenario, a customer branch site has a single CPE device that has two different WAN transport links. The following figure illustrates this scenario, showing that a dedicated MPLS connection and an internet connection terminate on the customer's CPE device.

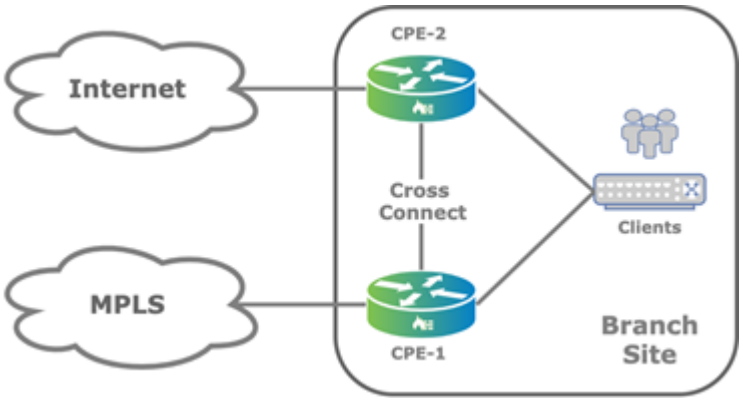


In this topology, overlay tunnels between branches are formed through the MPLS or internet underlay. You configure this functionality using a Director Workflow function. You can use alternate underlays from different providers, such as LTE connections.

You can configure a single CPE branch to support a single organization (also called a tenant) or, as multitenant, to support multiple tenants. The edge device provides full segregation between organizations or tenants. To provide addition segmentation, each organization on the edge device uses virtual routing and forwarding (VRF).

Branch with Active–Active Dual CPE Devices and Dual Transports

A variant of the scenario described in the previous section is a customer branch site that has two CPE devices in an active–active setup, instead of having a single CPE device, and that has two different WAN transport links, Here, the internet link connects to one of the CPE devices (CPE-1 in the figure below) and the MPLS link connects to the other (CPE-2 in the figure). The two paired CPE devices provide high availability (HA) and connect to each other using a cross-connect link.



On LAN side of the CPE device, the Virtual Router Redundancy Protocol (VRRP) provides gateway redundancy. If Layer 2 reachability on the LAN is not available, you can use Layer 3 routing protocols to reach the existing LAN side routers.

Each CPE device in the active–active HA pair maintains overlay connections to the WAN transports on the other CPE device. When the second underlay is physically attached to the other CPE device, it is logically represented on the local CPE device using a cross-connect link. In this deployment model, you can provision each CPE device with SD-WAN policies that leverage all the underlays.

This type of configuration is called active–active HA because both CPE devices always carry traffic to the underlay transport.

Note that state information, including NAT and session state, is not synchronized between the active–active HA CPE devices.

Best Practices for Single or Dual CPE Devices and Dual Transports

The following are best practices for branch deployments that consist of a single CPE device or two CPE devices in an active–active setup, and two WAN transports:

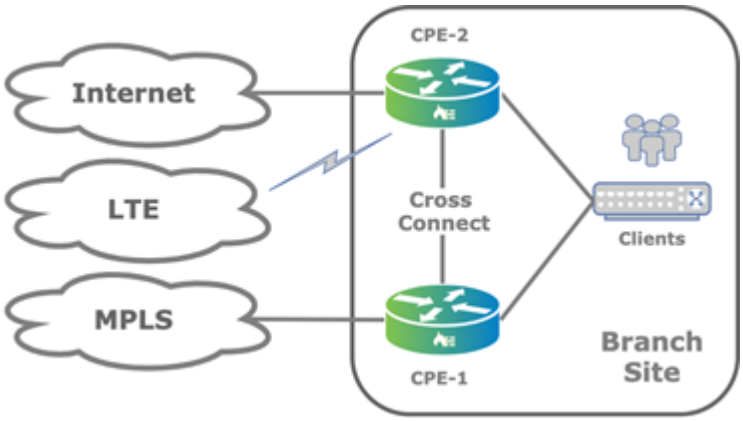
- For security, you should enable next-generation firewall (NGFW) on both CPE devices, because the NGFW and stateful firewall (SFW) software part of the LAN service chain, while the transport interfaces, including the cross-connect link, belong to the transport-VR domain.
- You should enable next-hop monitoring to upstream WAN transport gateway. Alternatively, you can enable a dynamic routing protocol.

- For this scenario, you can use the default Director Workflow for configuration.

Branch with Two CPE Devices and Three Transports

A third branch deployment scenario is a customer branch site has two CPE devices and three WAN links, as illustrated in the figure below. This is another form of the previous scenario, which has two CPE devices and two WAN transports.

The figure shows that the CPE-2 device is configured as a multihomed CPE device that has both internet and LTE connections. The second CPE device, CPE-1, connects only to the MPLS transport domain. You typically use this type of scenario when the LTE connection is required as a backup for the fixed transports.



When you deploy the branch CPE devices in active–active HA mode, the LAN side remains the same as described in Branch with Active–Active Dual CPE Devices and Dual Transports. This scenario also works for other combinations of WAN links, up to the maximum of 15 WAN links (for Releases 22.1.1 and later) or 8 WAN links (for Releases 21.2 and earlier) per tenant per device.

Note: Before you add more than 8 WAN links on any one VOS node that you are upgrading to Release 22.1.1, you must upgrade to Release 22.1.1 all the VOS nodes that communicate directly with the one running Release 22.1.1, including the Controller nodes.

Branch HA

To implement HA at a customer branch site, the site must have two VOS edge devices. To provide branch redundancy, VOS devices support two modes of operation: active–active mode and active–standby mode.

Active–active mode, which is stateless, is the more commonly used HA mode. It is simple to deploy and provides better performance during standard operations, because both the VOS edge devices are able to process traffic at all times.

For active–standby mode, which is both stateless and stateful, both underlays for each CPE device are physically connected, and so the cross-connect link is not required. You can configure each CPE device as a standalone CPE device and use VRRP on the LAN side.

There are a few drawbacks to active–standby mode:

- Without the cross-connect, each CPE device cannot take advantage of both WAN transports. As a workaround, you can introduce a Layer 2 switch on the WAN side to allow each CPE device to have access to both WAN circuits.
- During a CPE device failover, stateful connections are lost and TCP sessions are re-established. While the user may not notice the re-establishment of the TCP session, the functioning of some devices may be affected.

In stateful active–standby HA mode, only the active CPE device can forward traffic.

Stateful active–standby HA mode maintains a stateful synchronization between the edge devices. You use this mode when state is important, such as for NGFW and CGNAT traffic.

The following table compares the HA modes.

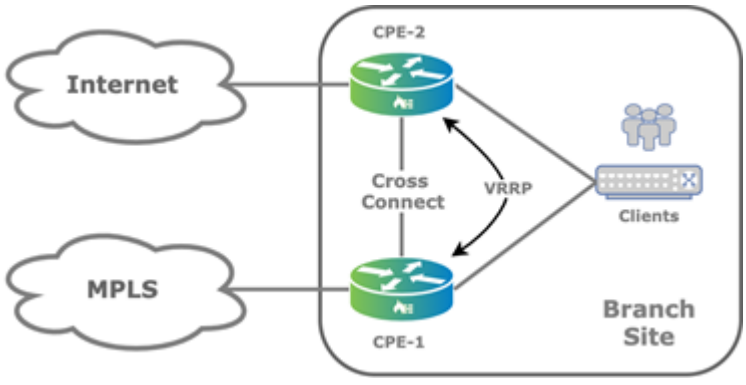
	Stateless Active–Active Mode	**Stateless Active–Standby Mode**	**Stateful Active–Standby Mode**
Overview	• Easy to configure using Director Workflows • Requires no additional WAN links • Uses cross-connect link between two CPE devices	• Easy to configure using Director Workflows • Manual optimization configurations required for VRRP tracking • All underlays are physically connected on both CPE devices • No cross-connect link is needed	• Requires manual configuration and physical setup • Short flows and flows inspected for UTM and URLs are not re-evaluated; after failover action is allow or drop • Requires additional uplinks or a switch, adding another active element that lowers site MTBF
Use cases	• Stateless routed traffic	• Stateless routed traffic	• Use where state preservation is important, such as firewalls and destination NAT
Complexity	• Easy to configure using Director Workflows	• Easy to configure using Director Workflows and some manual configuration	• Requires manual configuration

	Stateless Active–Active Mode	Stateless Active–Standby Mode	Stateful Active–Standby Mode
Available underlays	• Local and remote uplinks through cross-connnect link	• Local uplinks only	• Local uplinks only
Performance	• Not impacted by synchronization	• Not impacted by synchronization	• Requires synchronization of the control place and the working threads after transition from active to standby. • Unmeasured performance impact
BGP, IPsec, SLA monitoring scalability and state	• Minimum of one SLA per circuit per device	• Minimum of one SLA per circuit per device, but twice the number of SLAa because all underlay circuits are connected to both CPE devices	• Minimum of one SLA per circuit per device, but twice the number of SLA because all underlay circuits are connected to both CPE devices
Convergence			
• Upstream	• 3 seconds by default • Configure using VRRP and other timers • Can be configure to a shorter time	• 3 seconds by default • Configure using VRRP and other timers • Can be configure to a shorter time	• A few seconds by default, based on VRRP, quorum probes, BFD, and other timers
• Downstream	• From remote branch: SD-WAN control plane (MP-BGP) and SLA probes	• From remote branch: SD-WAN control plane (MP-BGP) and SLA probes	• From remote branch: SD-WAN control plane (MP-BGP) and SLA probes
Traffic restoration	• All session are	• All sessions are	• Long-term sessions

	Stateless Active–Active Mode	Stateless Active–Standby Mode	Stateful Active–Standby Mode
	restarted	restarted	are restored • Short-term sessions are restarted if they are out of sync
Synchronized services between primary and secondary nodes	• Not applicable	• Not applicable	• Data plane state, including sessions • Control plane state • Traffic-steering table • NAT bindings • ADC persistency
Node synchronized services	• All	• All	• Inspected antivirus, IDS/IPS, URL flows, after failover fail or pass and without security inspection if the synchronized flow is set to allow

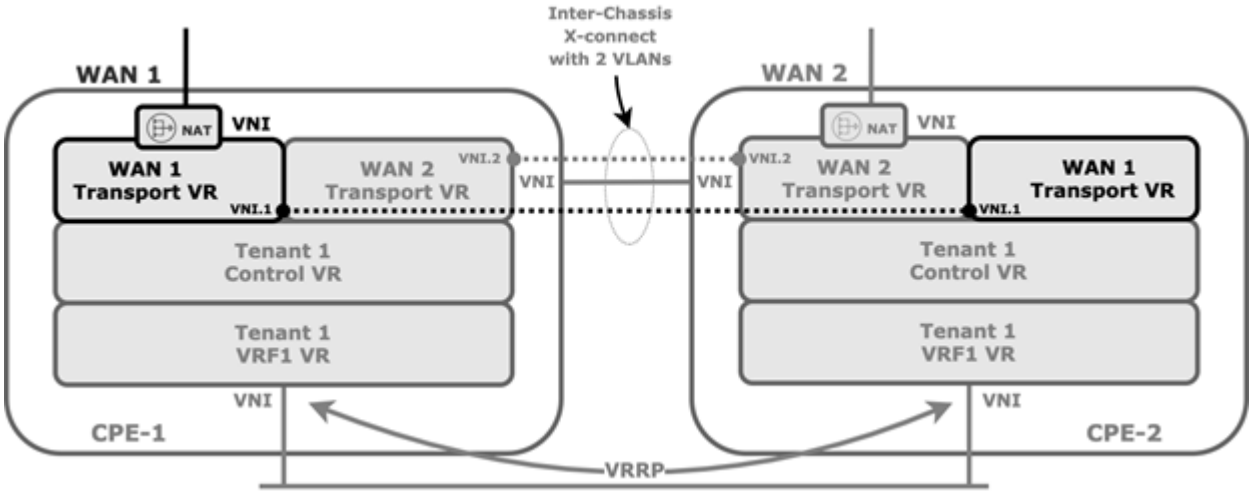
Cross-Connects in an Active–Active HA Topology

The following figure shows an active–active branch HA topology that uses two WAN transport underlays but does not have the dual underlays on the CPE devices.



The cross-connect link is a physical connection between the redundant CPE devices that emulates the missing transport

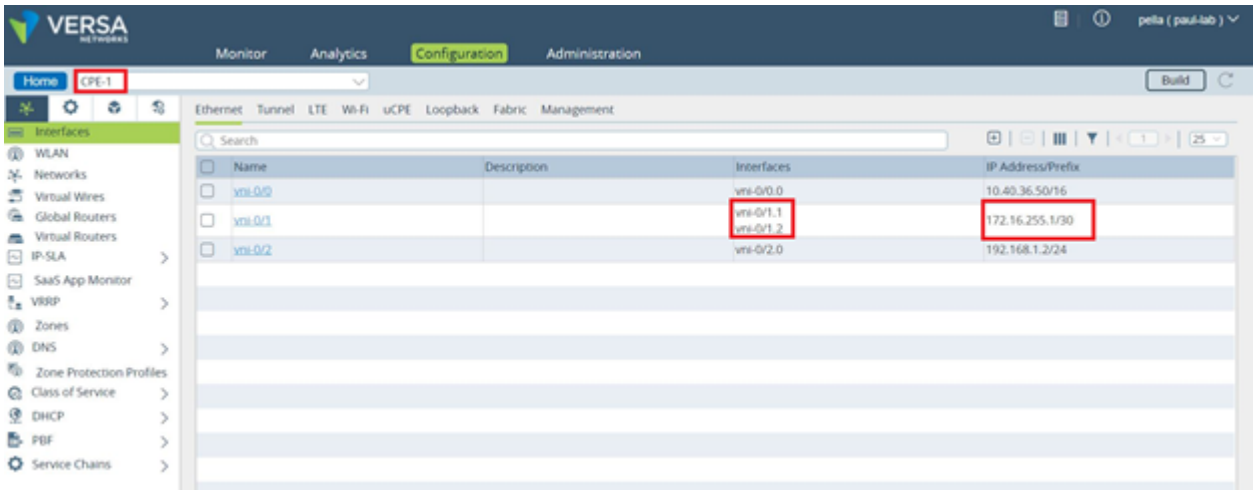
domain in a branch and provides redundancy to the attached clients, as illustrated in the following figure.



For the cross-connect link, you configure VLAN tagging for each WAN transport virtual router (VR) instance and you configure IP addresses configured using Workflow templates. Because the WAN transport VRs are distinct routing instances, they allow for reuse of IP addresses.

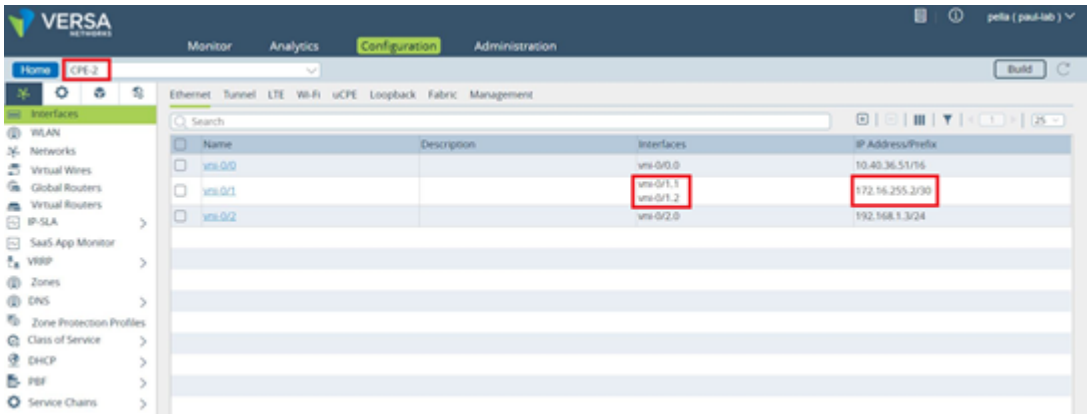
When you enable HA, by default, the back-to-back logical interfaces on the cross-connects are assigned IP addresses from the address range 172.16.255.0/30. The primary CPE device is assigned the address 172.16.255.1, and the second CPE device is assigned 172.16.255.2. Which CPE device is the primary or secondary is determined by which device uses the primary device template that is configured in the Director Workflow. The template for the secondary device is automatically generated by the Workflow.

The following screenshot shows a default HA interface configuration for the CPE-1 device. Notice that the assigned IP address is 172.16.255.1.

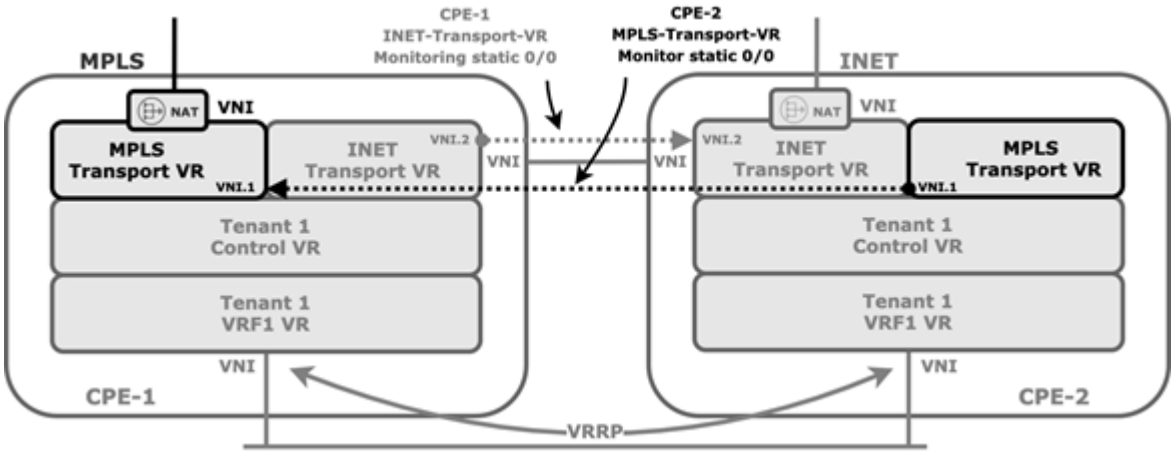


The following screenshot shows a default HA interface configuration for the CPE-2 device. Here, the assigned IP

address is 172.16.255.2.



You configure static routes in the transport VR of each CPE device, and you use the Workflow to configure ICMP monitoring in the transport VR associated with the cross-connect link to direct traffic destined to the WAN connection of the paired CPE device. The following figure show static route ICMP monitoring with HA.



If the cross-connect interface fails over, ir the peer CPE device goes down, or if there is any other IP reachability issue over the cross-connect interface, the static route is withdrawn from the routing table of the corresponding WAN transport VR.

To configure ICMP monitoring on the CPE-1 device that connects to the MPLS transport VR from the CLI:

```
admin@CPE-1-cli> show configuration | display set | match icmp
set routing-instances INET-Transport-VR routing-options static route 0.0.0.0/0 172.16.255.1 none icmp
set routing-instances INET-Transport-VR routing-options static route 0.0.0.0/0 172.16.255.1 none icmp interval 5
set routing-instances INET-Transport-VR routing-options static route 0.0.0.0/0 172.16.255.1 none icmp
threshold 6
```

To configure ICMP monitoring on the CPE-2 device that connects to the internet transport-VR from the CLI:

```
admin@CPE-2-cli> show configuration | display set | match icmp
```

```
set routing-instances MPLS-Transport-VR routing-options static route 0.0.0.0/0 172.16.255.2 none icmp
set routing-instances MPLS-Transport-VR routing-options static route 0.0.0.0/0 172.16.255.2 none icmp interval 5
set routing-instances MPLS-Transport-VR routing-options static route 0.0.0.0/0 172.16.255.2 none icmp threshold 6
```

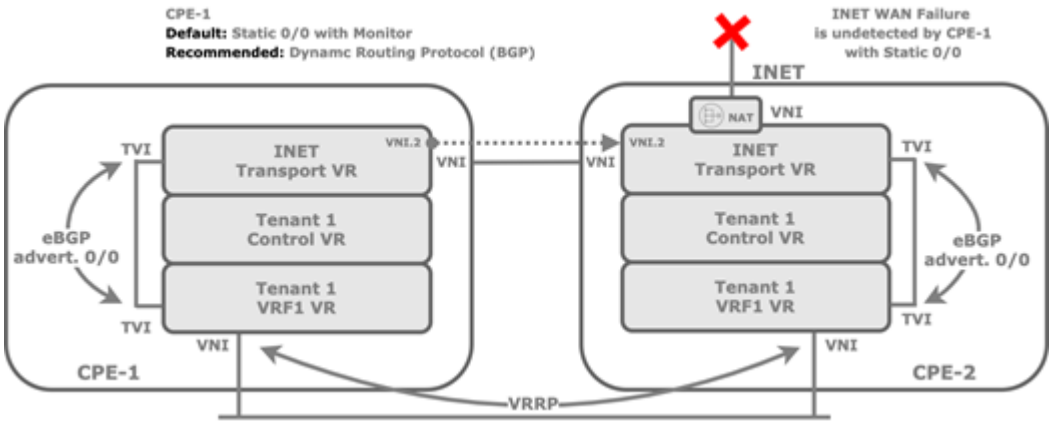
On the LAN side, you use VRRP to elect an active node and a standby node. The logical interface and its virtual IP address is used as the next hop or gateway on the LAN.

DIA in an Active–Active HA Topology

When you configure direct internet access (DIA) in an active–active HA scenario, there are a few things to note.

When you enable DIA, the Director Workflow automates the configuration of BGP peering between the transport VR and the LAN-VR, which is needed to propagate default route.

The ICMP monitoring is between the cross-connect logical interfaces of the CPE-1 internet transport VR and the CPE-2 device and therefore does not protect against internet WAN link failure on the CPE-2 device. This may result in a local internet black hole scenario if the internet WAN link fails on CPE-2, as show in the following figure.



To protect against the local black hole, you must take additional measures, such as monitoring the next next-hop (that is the remote next hop). Doing this may add complexity to the network design, because you may need to use NAT to allow the internet provider WAN interface to reply to ICMP echo requests. The NAT would be necessary because ICMP requests are sourced from the 172.16.255.0/30 prefix range and are not necessarily routed back by the provider router.

The recommended solution is to use dynamic routing over the cross-connect interface between transport VRs to propagate routes and the default route from main transport-VR of each CPE device.

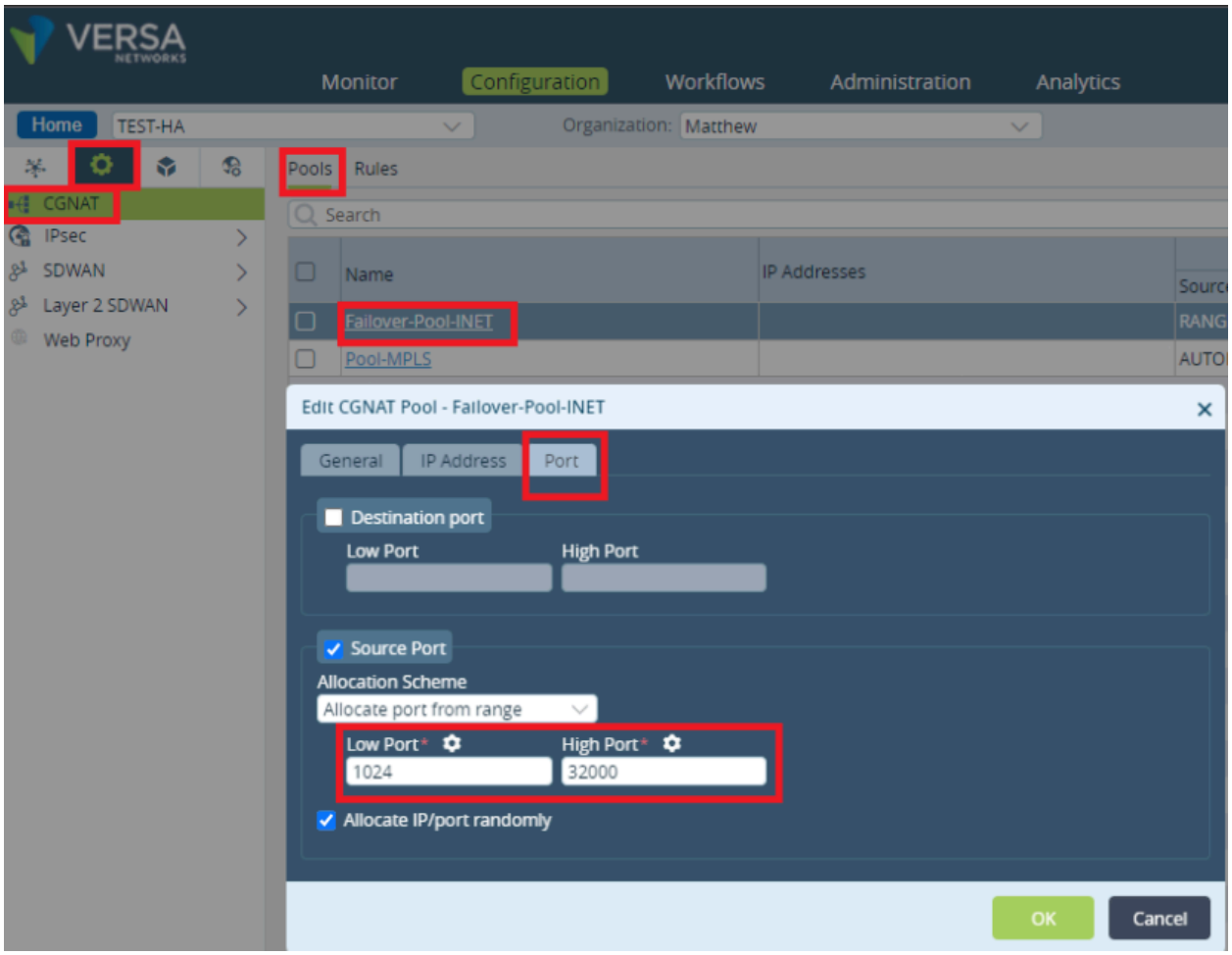
Ports Used in a Branch Active–Active HA Topology

By default, each VOS branch device tries to send SD-WAN traffic using UDP port 4790 as both the source and destination port. However, for an active–active topology, this port cannot be used by the VOS device that is reachable

over cross-connect links because active devices cannot both use the same port. However, when the traffic passes through the cross-connect, the source port is NATed to a random port in the range 1024 through 32000.

To change the range of ports used for NATing:

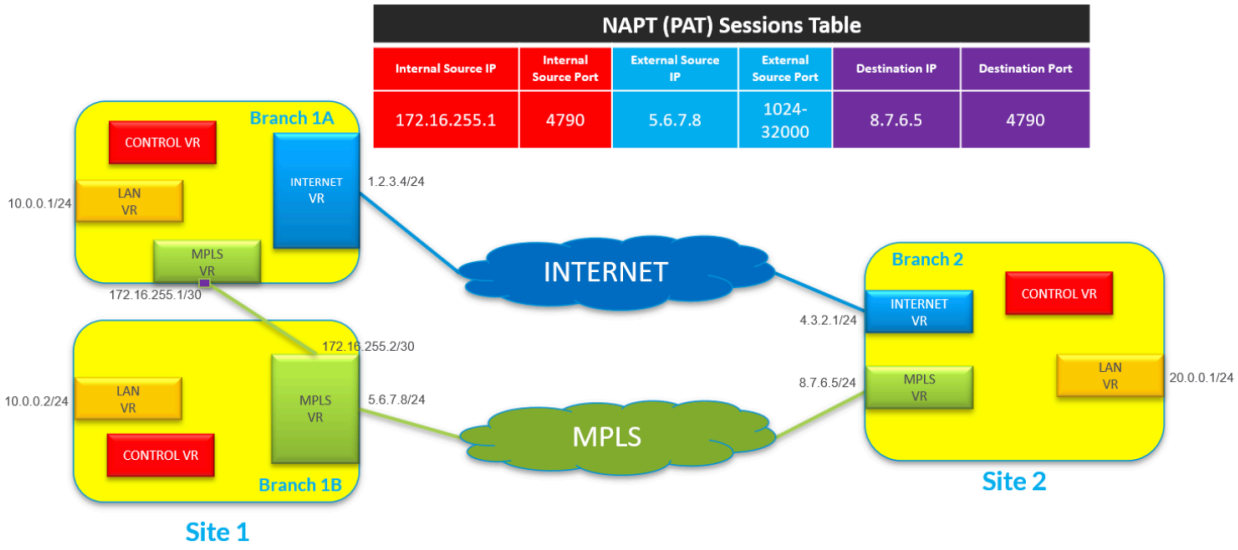
1. In Director view:
   a. Select the Configuration tab in the top menu bar.
   b. Select Devices > Devices in the left menu bar.
   c. Select an organization in the left menu bar.
   d. Select a device in the main pane. The view changes to Appliance view.
2. Select the Configuration tab.
3. Select Services > CGNAT in the left menu bar, select the Pools tab in the horizontal menu bar, and select the CGNAT pool.



4. In the Edit CGNAT Pool popup window, select the Port tab.
5. In the Allocation Scheme field, select Allocate Port from Range, and then enter the lowest and highest port numbers.

6. Click OK.

The following figure illustrates how and where port translation occurs. In this example, for the Branch1A to reach the MPLS transport of the Branch2, it originates SD-WAN traffic from the source address 172.16.255.1:4790 and sends it to the destination address 8.7.6.5:4790. However, because Branch1B is already using 5.6.7.8:4790 as the source address for connections to Branch2, the VOS device performs a NAT translation of the traffic from Branch1A from 172.16.255.1:4790 to 5.6.7.8:1024 (or to a random port number in the range 1024 through 32000).



Supported Software Information

Releases 20.2 and later support all content described in this article, except:

- Release 22.1.1 supports up to a maximum of 15 WAN links per tenant per CPE device.

Additional Information

Configure Virtual Routers
Overview of Configuration Templates

LTE Transport Modes

For supported software information, click here.

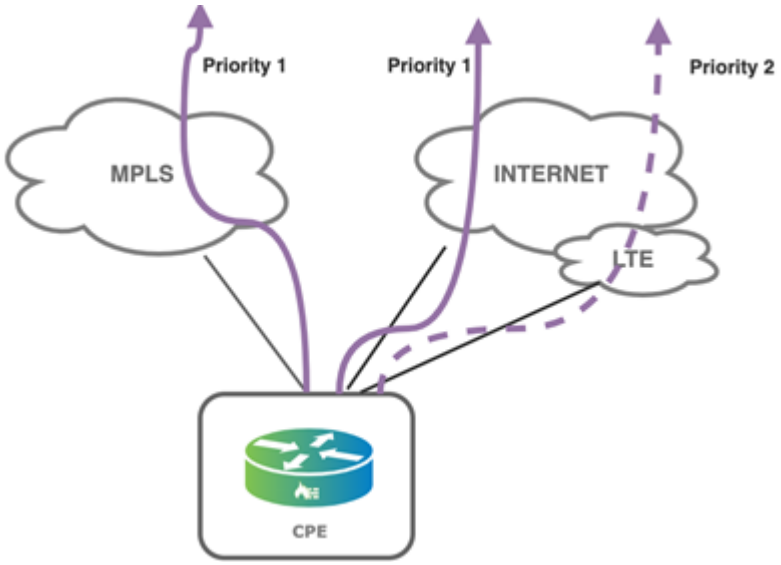
On Versa Operating System™ (VOS™) edge devices, you can use LTE as a WAN link, and it acts just like any other WAN link. However, in many scenarios, LTE is deployed as a backup path because its data costs are higher. When you use LTE as a backup transport link, you can configure it in one of the following modes:

- Hot standby mode—In this mode, the LTE interface is up and SLA packets are sent over this link to remote sites to determine the path metrics, but the LTE link is not actively used for sending traffic.
- Cold standby mode—In this mode, you configure the LTE link and place it in an administratively Down state. The link state goes to Up only when all the primary wired WAN interfaces are down.

This article discusses the LTE hot standby and cold standby modes.

Hot Standby Mode

In hot standby mode, the LTE interface is up, and SLA packets are sent over this link to remote sites to determine the path metrics. However, the LTE link is used to send traffic only when the primary wired WAN links are down or out of SLA compliance. You can configure management traffic, such as LEF logging information or branch software uploads, to avoid the LTE link when other wired WAN links are available. The following figure illustrates the traffic priories in hot standby mode.



The following are the benefits of hot standby mode:

- When the primary wired WAN interfaces go down or are out of SLA compliance, the switchover to LTE occurs instantly.
- SLA-based steering is possible because the LTE link is actively monitored through SLA probes.
- The status and quality of the path is known at all times in standby mode. The LTE link generates alarms if the path becomes unavailable, prompting the administrator to take corrective action.

The following are the limitations of hot standby mode:

- The path through the LTE link is actively monitored. This means that SLA traffic is sent over the LTE link when it is in standby mode, thus consuming credits of the LTE data plan subscription.
- The amount of bandwidth used by the SLA traffic varies, depending on the number of network sites with which SLA peering is maintained. Also, the SLA probe interval influences the amount of data sent over the LTE link.

This section describes the following LTE hot standby configuration scenarios:

- Use LTE hot standby for SD-WAN VPN (site-to-site) traffic
  - Use the LTE link as the backup circuit on local and remote devices
  - Avoid the LTE link for non-business and scavenger traffic
- Use the LTE link as a backup for internet traffic
- Load-balance two wired WAN links, with an LTE backup
- Provide connectivity for management traffic

Configure LTE Hot Standby for SD-WAN VPN Traffic

This section discusses two scenarios for using LTE hot standby for SD-WAN VPN (site-to-site) traffic:

- Use LTE as backup circuit on local and remove device
- Avoid LTE for non-business and scavenger traffic

When you configure WAN links, a circuit media attribute is associated with each link. When you configure LTE hot standby for SD-WAN VPN (site-to-site) traffic, the circuit media can be any of the following:

- Cable
- DSL
- Ethernet
- LTE
- T1
- T3

Use LTE as Backup Circuit on the Local and Remote Branches

This section describes a scenario that requires the LTE links to be in hot standby mode. You should apply this configuration to all devices in the network, both those with LTE interfaces and those that do not have LTE interfaces.

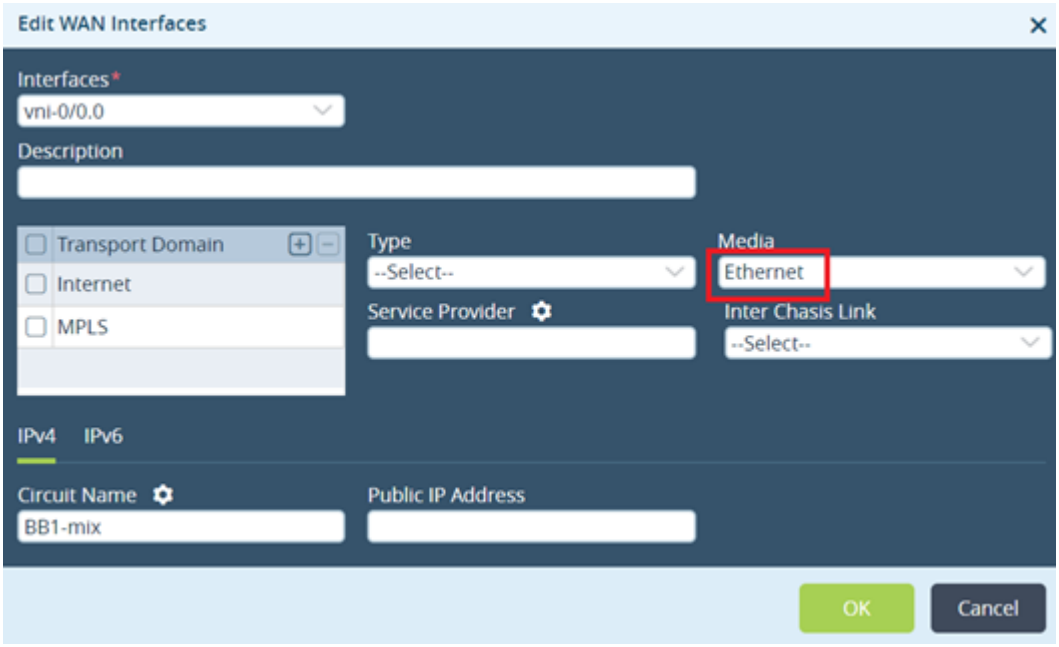
In this scenario, the local branch decides which path to use to connect to the remote branch. This decision includes choosing the WAN circuits on the local and remote branches. In many networks, not all branches are deployed with LTE interfaces. A non-LTE branch must not use the path to the remote LTE interface if the remote wired internet interface is available.

The configuration shown here allows the local and the remote branches to use LTE only when needed.

The WAN link circuit media type is used to set the LTE link to be in hot standby mode. To verify whether the circuit media type is correct for all the WAN links:

1. In Appliance view, select the Configuration tab in the top menu bar.

2. Select Services > SD-WAN > System > Site Configuration in the left menu bar.
3. Click the Edit icon. The Edit WAN Interfaces popup window displays.

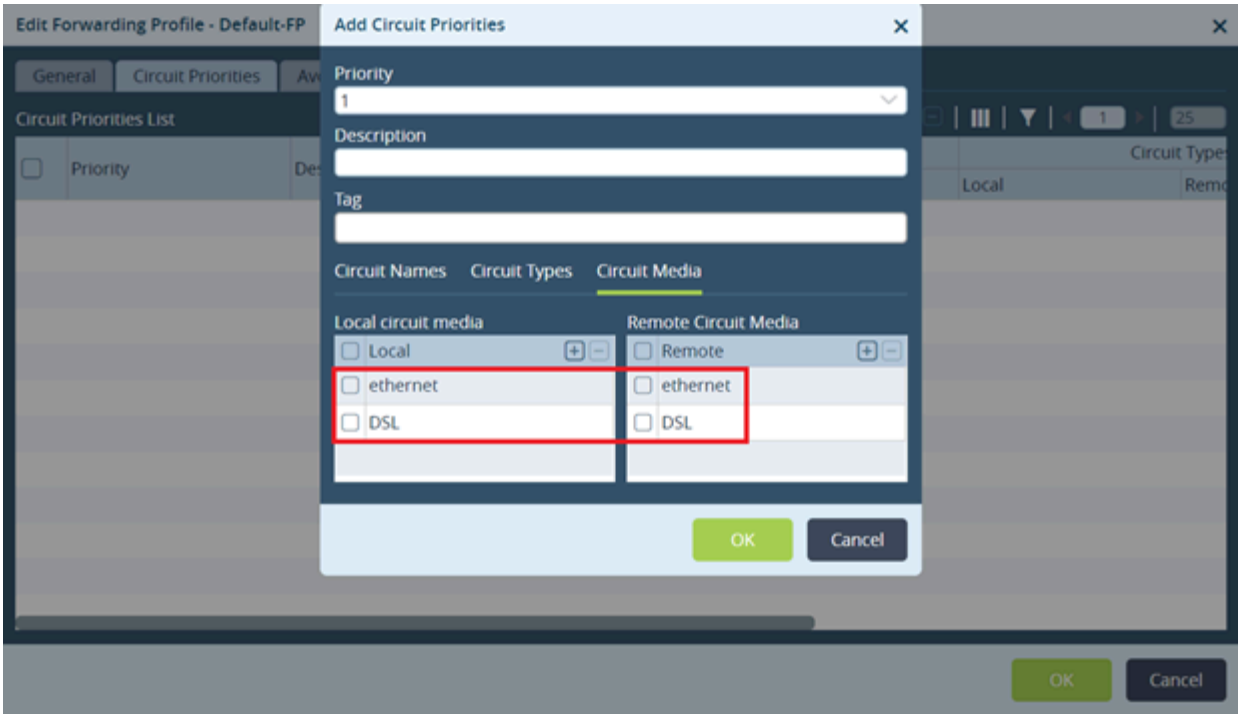


In the Media field, check that the media type is set to either Ethernet, DSL, or LTE, depending on the type of WAN link. The Ethernet interfaces are displayed as vni-0/0 and vni-0/1, and the LTE interfaces are displayed as vni-0/100 and vni-0/101.

To configure the LTE interface to be in hot standby mode, you set the circuit priority on the local and remote branches. Because the branch that initiates the traffic chooses the path, it is important that the local branch not prioritize the LTE link on the remote branch. This configuration applies when the sending branch does not have an LTE circuit, for example, to allow communication with a branch in a data center.

You configure the circuit priorities in the default SD-WAN forwarding profile. To configure the circuit priorities:

1. In Appliance view, select the Configuration tab in the top menu bar.
2. Select Configuration > Services > SD-WAN > Forwarding Profiles.
3. Click the Add icon. The Add Forwarding Profile popup window displays.
4. Select the Circuit Priorities tab.
5. Click the Add icon. The Add Circuit Priorities popup window displays.

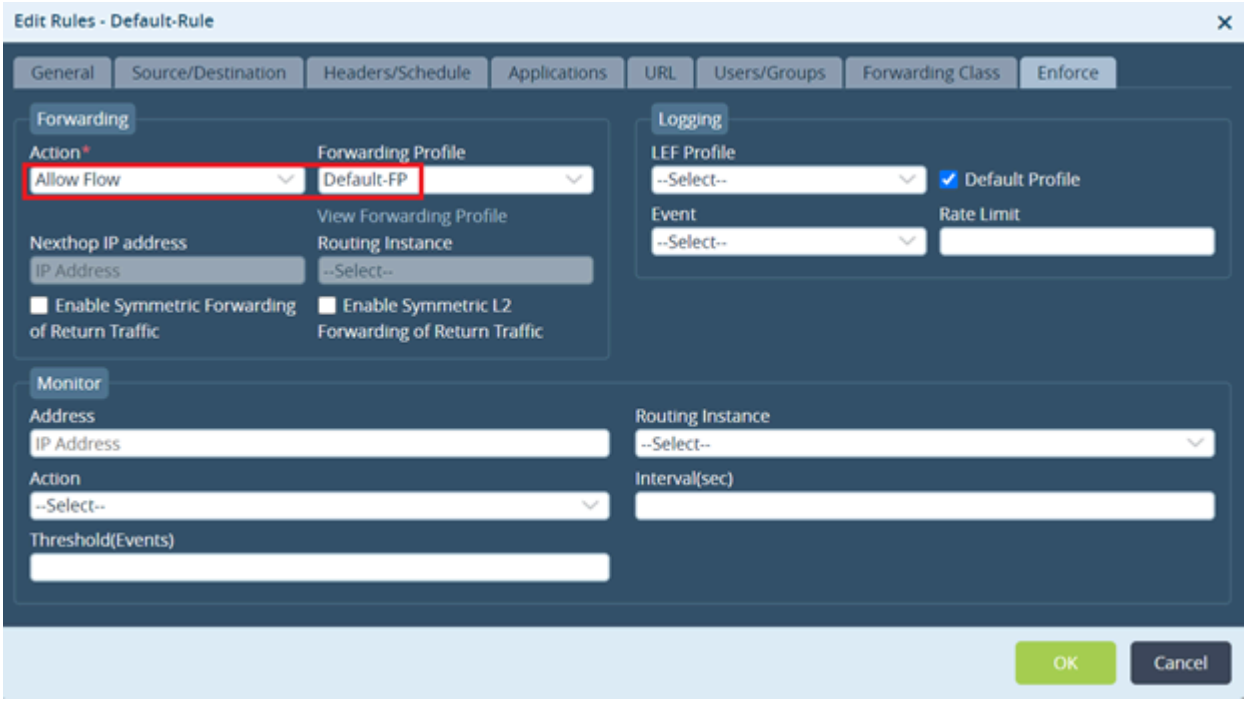


Set the circuit priorities as shown in the following table. Note that if the local and remote circuit media are other types of WAN links, such as E1 or T1, you can include them in the priority groups that contain Ethernet media. For circuit priorities, a lower value indicates a more preferred link and a higher value indicates a less preferred link.

Priority	Purpose	Local Circuit Media	Remote Circuit Media
1	Use this circuit when local and remote wired WAN links are up	Ethernet, DSL	Ethernet, DSL
2	Use this circuit when the wired WAN links on the local branch are down and LTE is the only available WAN link	LTE	Ethernet, DSL
3	Use this circuit when the wired WAN links on the remote branch are down and LTE is the only available WAN link	Ethernet, DSL	LTE
4	Use this circuit when LTE is the only WAN link available on both the local and remote branches; you do not need to configure this priority		

Then you configure an SD-WAN policy rule to match the traffic and have it follow the WAN link priority order. To create an SD-WAN policy rule that matches all traffic (a wildcard match):

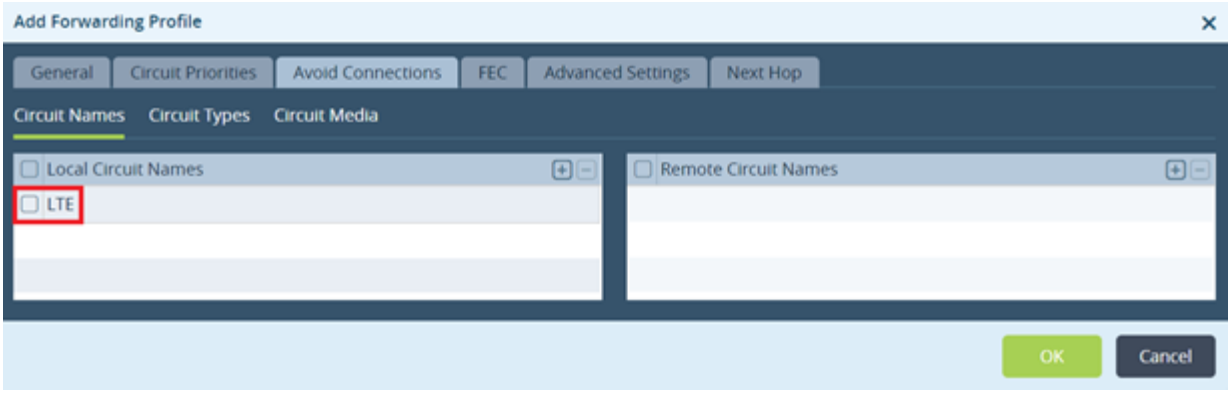
1. In Appliance view, select the Configuration tab in the top menu bar.
2. Select Services > SD-WAN > Policies in the left menu bar.
3. Select the Rules tab to create a new rule. For a wildcard catch all rule, do not configure any match conditions. Select the Enforce tab, in the Forwarding Profile field, select the SD-WAN forwarding profile that you created above, in this example called Default-FP.



Avoid LTE for Non-Business and Scavenger Traffic

You can configure the local circuit media or a combination of local and remote circuit media to completely avoid using LTE circuits for non-business critical traffic and scavenger traffic. (Scavenger traffic is a QoS category that includes suspect traffic that may be dangerous to the network.) To do this, when you configure an SD-WAN forwarding profile, you define the circuits to avoid.

This configuration is the same as that shown in the previous section. However, in the Avoid connections tab, you define the circuits to avoid. These circuits are never used even if the path is the only one that is available.



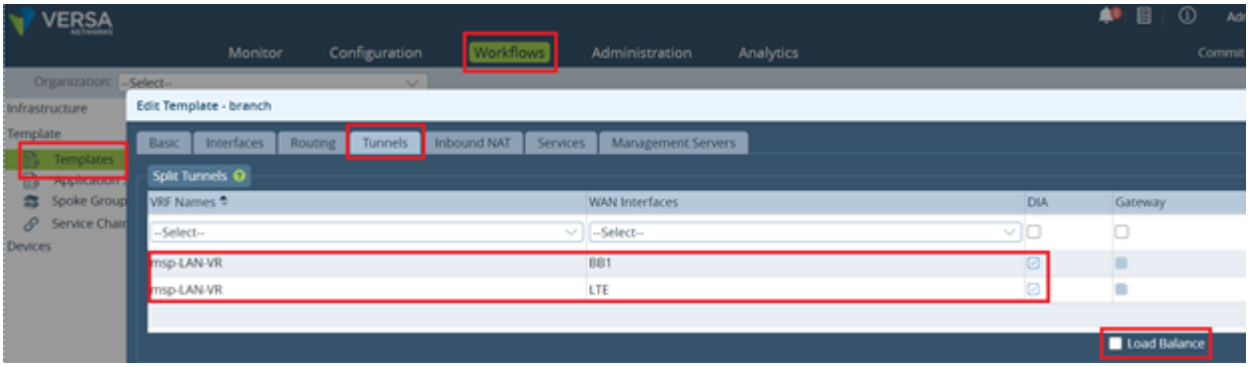
You then create an SD-WAN policy rule to classify the scavenger traffic and attach the rule to the forwarding policy that specifies to never use LTE.

Note that the forwarding profile configuration affects egress traffic decisions. For an effective implementation, you must apply this configuration uniformly across the network to prevent ingress traffic from arriving on the LTE link.

Configure LTE as a Backup for Internet Traffic

This section describes how to configure an LTE link to be a backup for local internet breakout (DIA) traffic. The scenarios in this section have either one wired link and one LTE link, or more than one wired link and one LTE link.

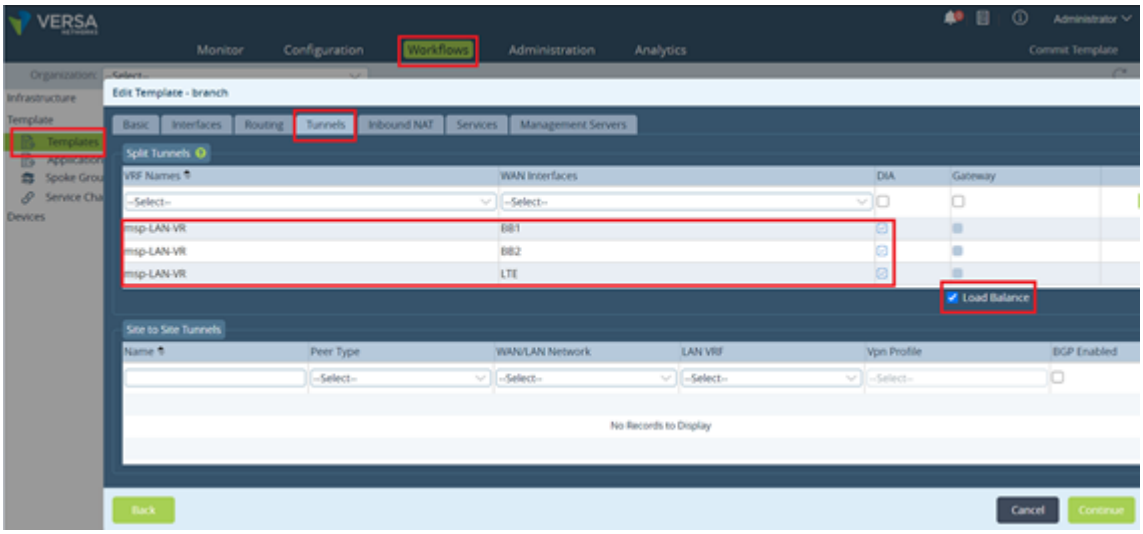
To use LTE as a hot standby for direct internet access (DIA) traffic, in the Workflows template, in the Tunnels tab, do not click the Load Balance option.



The Workflows template automatically creates two BGP sessions over a virtual interface pair that run between the WAN transport VR and the MSP LAN VR. The default route is advertised from each of the WAN transport VRs to the LAN VR. The default route from the BB1 transport VR has a BGP local preference value of 120 by default, and the default route from the LTE transport VR has a BGP local preference value of 119. Based on the local preference values, the LAN VR installs the default route from the BB1 transport VR as the active default route, thus making BB1 WAN link the primary WAN link for local internet-bound traffic.

To load-balance between two wired WAN links for DIA, while having an LTE link as a hot-standby backup, you enable

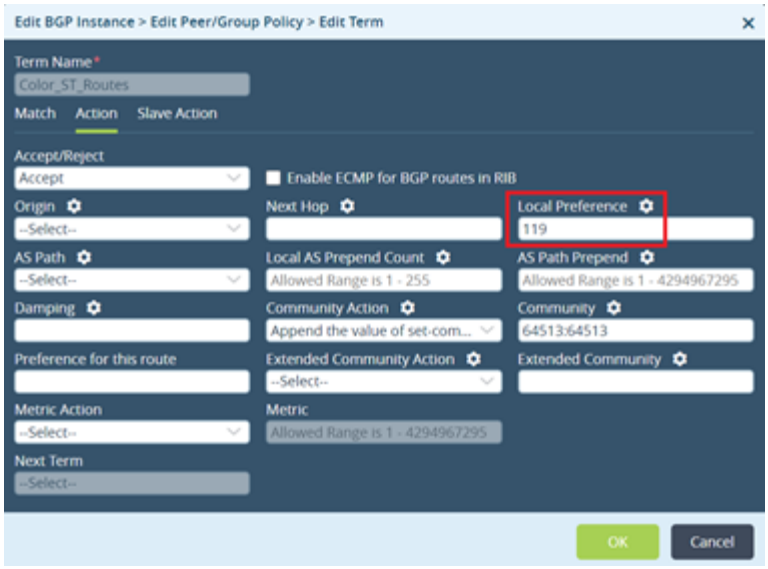
DIA on all internet WAN and check the load balance knob. This configuration sets all WAN links to the same local preference.



Because you select the load balance option, all the default routes from the BB1, BB2, and LTE transport VRs are advertised to the LAN VR with a BGP local preference value of 120. The result is that traffic is load-balanced across all the three WAN links. To have the LTE link be the backup and be active only when the two wired WAN links are down, modify the BGP local preference of the default route advertised by the LTE VR to a value of 119 or lower.

To modify the local preference value:

1. In Appliance view, select the Configuration tab in the top menu bar.
2. Select Networking > Virtual Routers in the left menu bar, and select a virtual router instance. The Virtual Router popup window displays.
3. Select the BGP tab in the left menu bar. The main pane displays a list of the BGP instances that are already configured.
4. Select the BGP instance for the LTE link. The Edit BGP Instance popup window displays.
5. Select the Peer/Group Policy tab. The Add Add Peer/Group Policy popup window displays.
6. Select the Action tab, and change the local preference value to a value of 119 or lower.



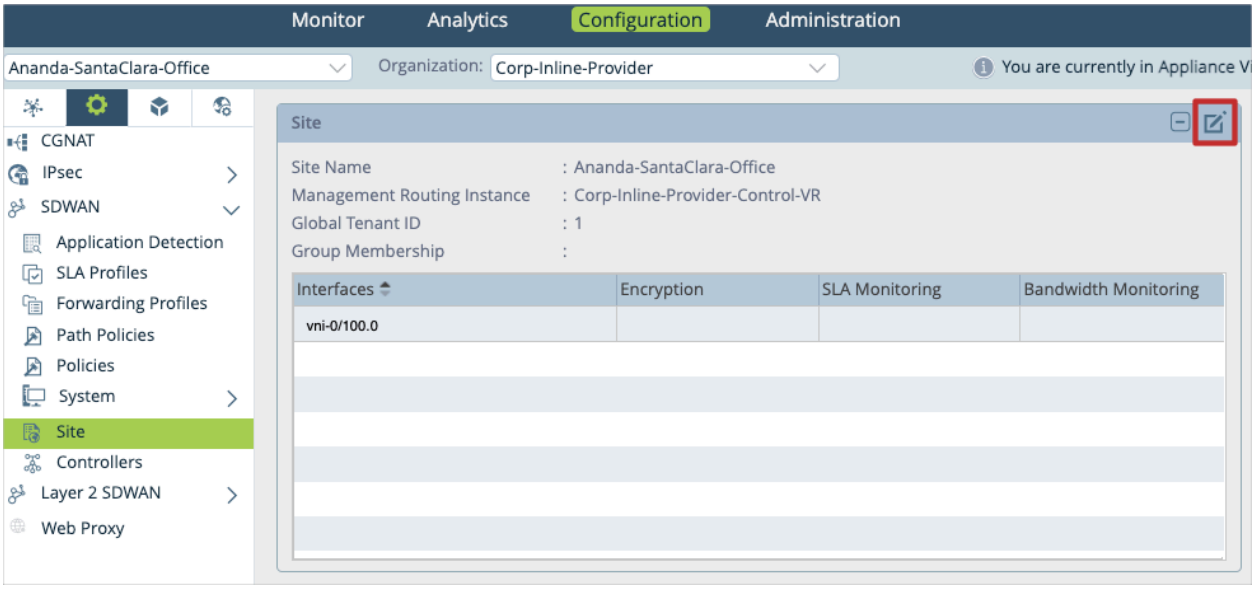
Minimize the Sending of Management Traffic on LTE Links

By default, the VOS device randomly assigns an available link to provide connectivity to the Controller nodes and Versa Networks headend devices. This means that the LTE link could be used to send management traffic. Management traffic is all traffic towards the Controller and headend devices, including LEF logging information, which is sent to the Analytics node, or uploading software image files.

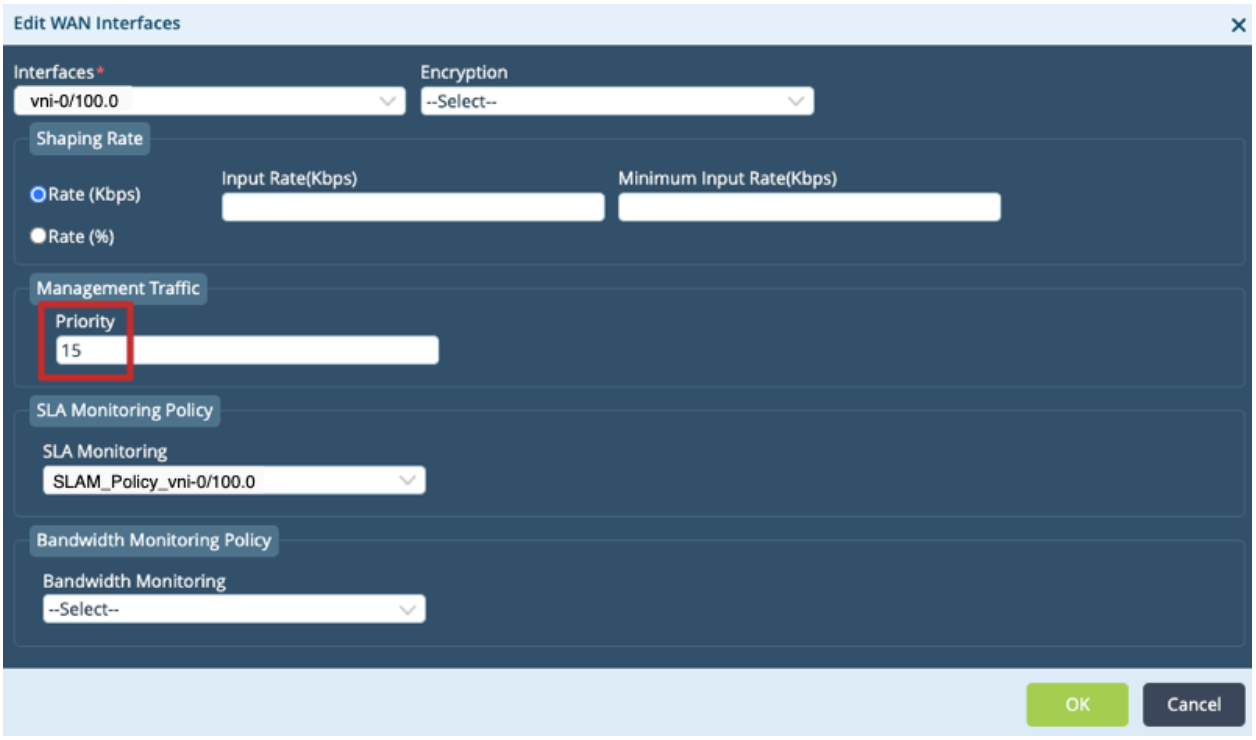
To prevent the unnecessary use of the LTE data plan, you configure management traffic so that the preferred links are the wired WAN links. To do this, you set the management priority of the LTE WAN links to a value that is lower than that of the wired WAN links. For management priorities, a value 0 indicates the highest priority and a value 15 is the lowest.

To set the management traffic priority to minimize the use of LTE links for management traffic:

1. In Appliance view, select the Configuration tab in the top menu bar.
2. Select Services > SD-WAN > Site in the left menu bar. The following screen displays.



3. Click the Edit icon in the main pane. The Edit Site popup window displays.
4. In the WAN Interfaces table, select the LTE interface. The Edit WAN Interfaces popup window displays.



5. In the Management Traffic > Priority field, enter a value of 15. Note that for management priorities, a value 0 indicates the highest priority and a value 15 is the lowest.
6. Click OK.

Cold Standby Mode

In cold standby mode, you configure the LTE link in an administratively Down state. Only when all the primary wired WAN interfaces are down does the LTE link state change to Up.

The primary benefits of cold standby mode is that no data is used on the LTE link until all the primary WAN links are down.

The following are the limitations of cold standby mode:

- The VOS edge device does not know the actual status of the LTE connection until the primary WAN links go down and the VOS devices tries to bring up the LTE connection.
- The traffic failover is not instantaneous, because there is lag time while registering the SIM to the mobile network. Data transfer can begin only after the mobile data context is enabled on the SIM and the Versa Networks control plane is established on that path.

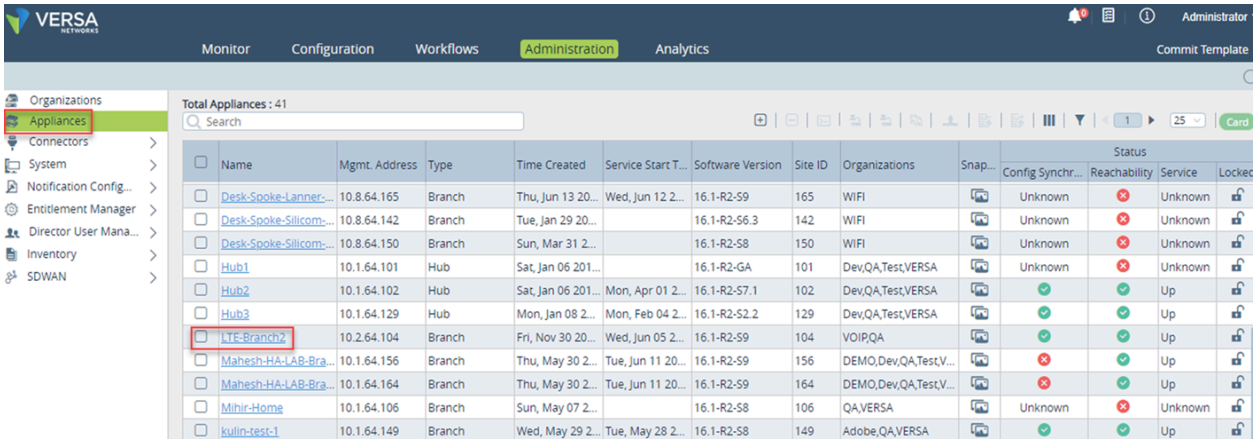
To configure LTE cold standby mode, you configure the LTE interface to be in an administratively Down state, and you configure an active wired interface that has an active monitor (an IP SLA monitor) to the active interface's next hop. When the IP SLA monitor determines that the next hop is no longer reachable, the LTE link is activated automatically and starts building neighbor sessions with all its peers in the transport domain.

To configure cold standby mode, you do the following:

1. Create a monitor for the primary WAN circuit.
2. Create monitor-group to check the status of the primary links.
3. Associate the monitor group to the LTE interface as a standby.

To create a primary WAN interface:

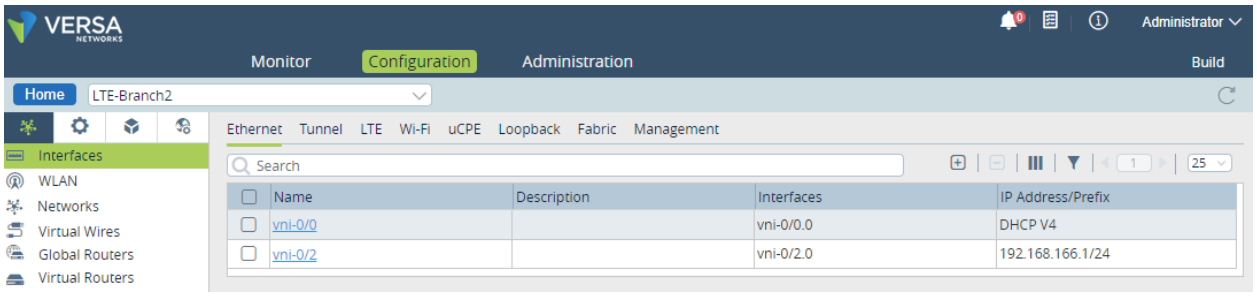
1. In Director view, select the Administration tab in the top menu bar.
2. Select Appliances in the left menu bar.
3. Select a device in the main pane. In the following screenshot the device is LTE-Branch2.



4. Select the Configuration tab in the top menu bar.

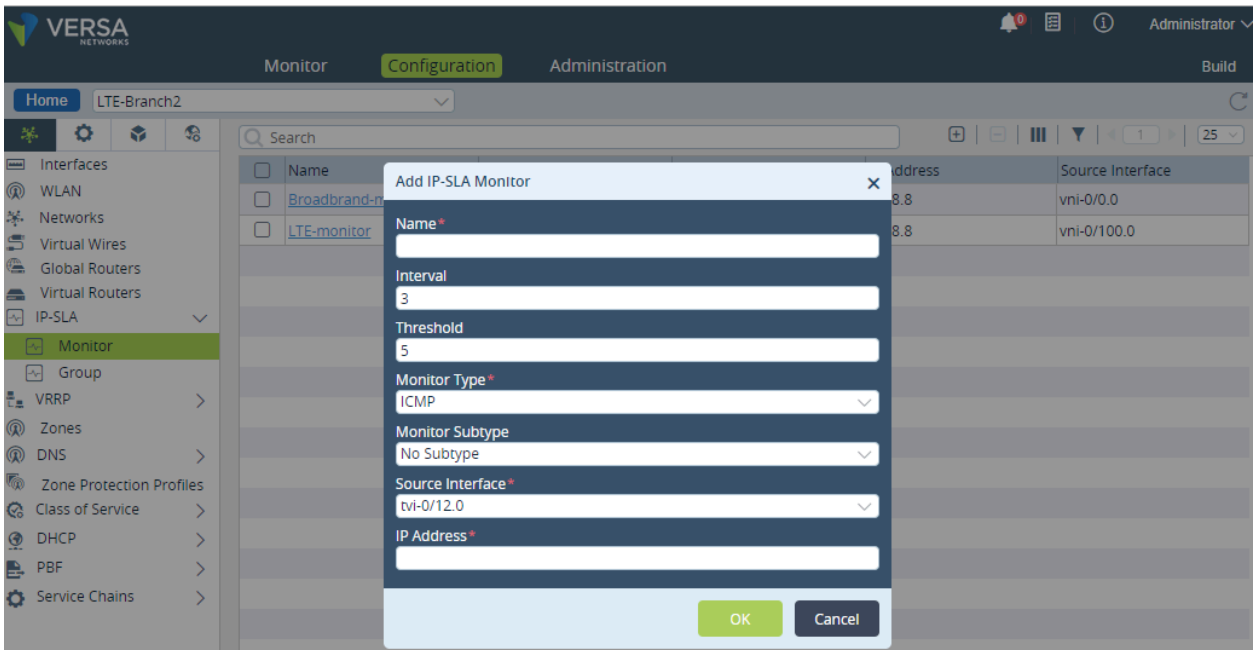
5. Select Network > Interfaces in the left menu bar.

6. Click the Add icon to create an interface. This example configures vni-0/00 as the primary link.



7. Select Network > IP SLA > Monitor in the left menu bar.

8. Click the Add icon to create a monitor for the primary WAN interface. The Add IP SLA Monitor popup window displays. Enter information for the following fields.



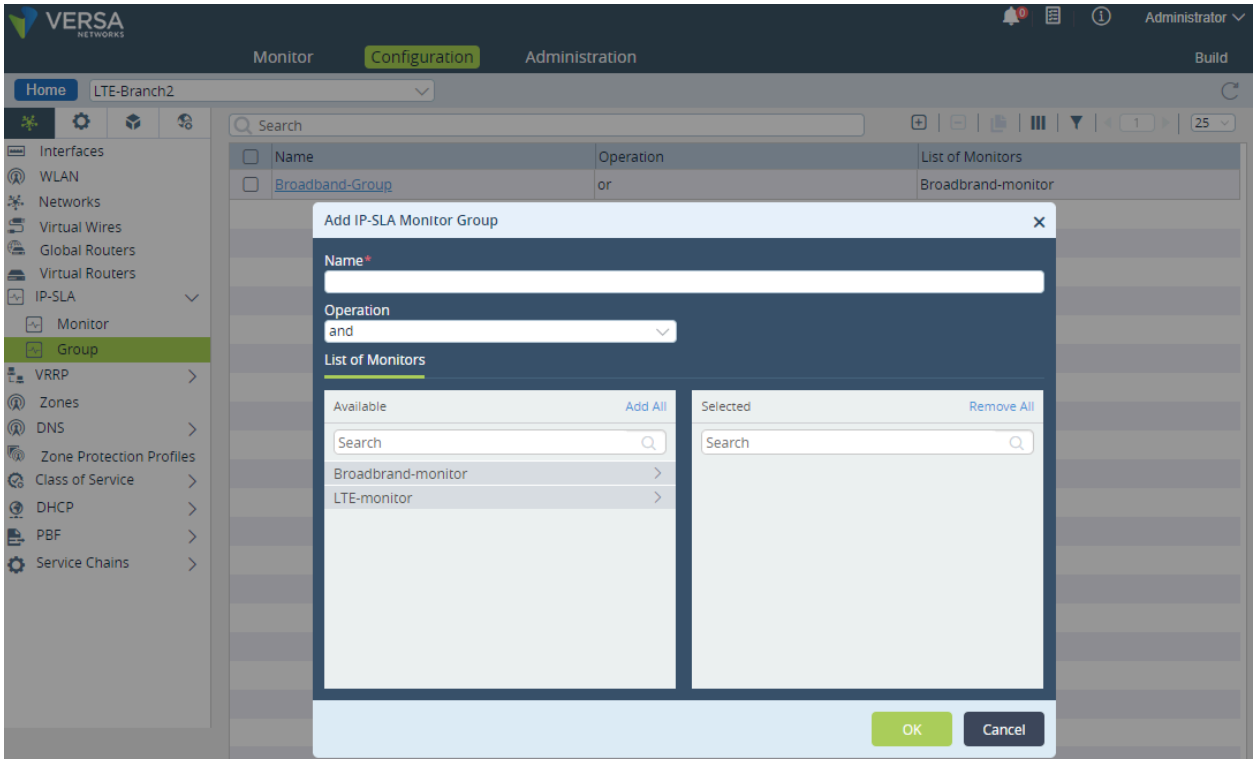
Field	Description
Name	Enter a name for the IP SLA monitor object. This example uses the name Broadband-monitor.
Interval	Click, and enter the frequency, in seconds, at which to send ICMP packets to the IP address. *Range*: 1 through 60 seconds *Default*: 3 seconds
Threshold	Enter the maximum number of ICMP packets to send to the IP address. If the IP address does not respond after this number of packets, the monitor object, and hence the IP address, is marked as down. *Range*: 1 through 60 *Default*: 5
Monitor Type	Select the type of packets sent to the IP address. The available options are DNS, ICMP, or TCP.
Monitor Subtype	Select the subtype: ◦ HA probe type—Select to avoid interchassis HA split brain. For more information, see Configure Interchassis HA. ◦ Layer 2 loopback type—Select to monitor an external service node configured as a Layer 2 loopback (virtual wire). ◦ No subtype—Do not use a monitor subtype. This is the default. *Default*: No subtype
Source Interface	Select the source interface on which to send the probe packets. This interface determines the routing instance through which to send the probe packets. This routing instance is the target routing instance for the probe packets.
IP Address	Enter the IP address to monitor.

9. Click OK.

To create a monitor group and add the monitor object:

1. Continuing from the previous procedure, select Network > IP SLA > Group in the left menu bar.

2. Click the Add icon to create a monitor group. The Add IP SLA Monitor Group popup window displays. Enter information for the following fields.

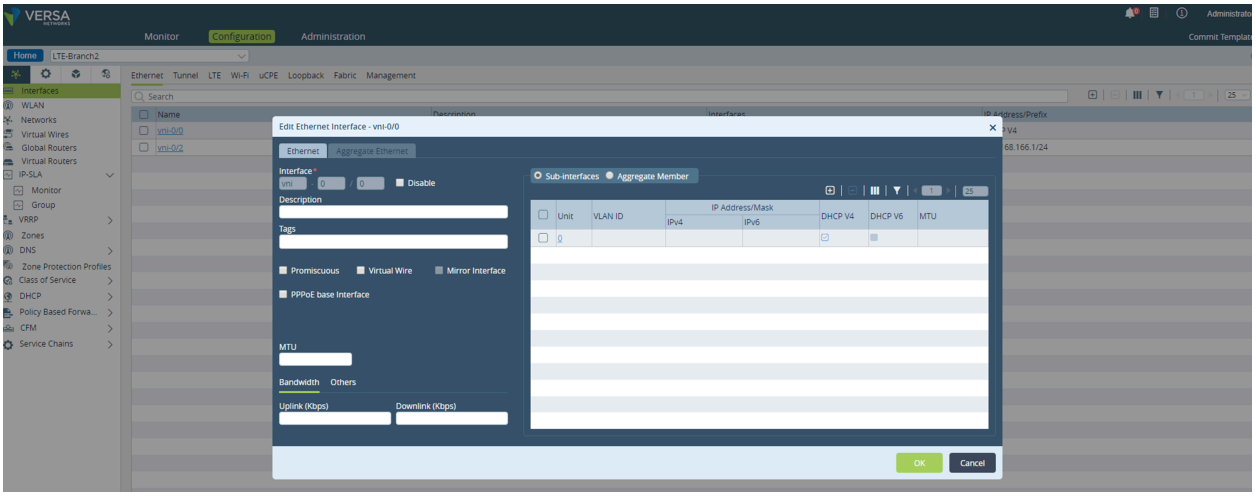


Field	Description
Name	Enter a name for the IP SLA monitor group. This example uses the name Broadband-Group.
Operation	Select the boolean operation to perform on the monitors: ◦ AND—In an AND operation, the monitor group result is Up only if all monitors are Up. Otherwise, the monitor group result is Down. ◦ OR—In an OR operation, the monitor group result is Up if at least one of the monitors is Up. It is down only if all monitors are Down.
List of Monitors (Table)	
◦ Available	Displays the list of available monitors for this appliance. Select and click on the monitor that you want to add to the group, here, Broadband-monitor.
◦ Selected	Displays the monitor that you added to the group.

3. Click OK.

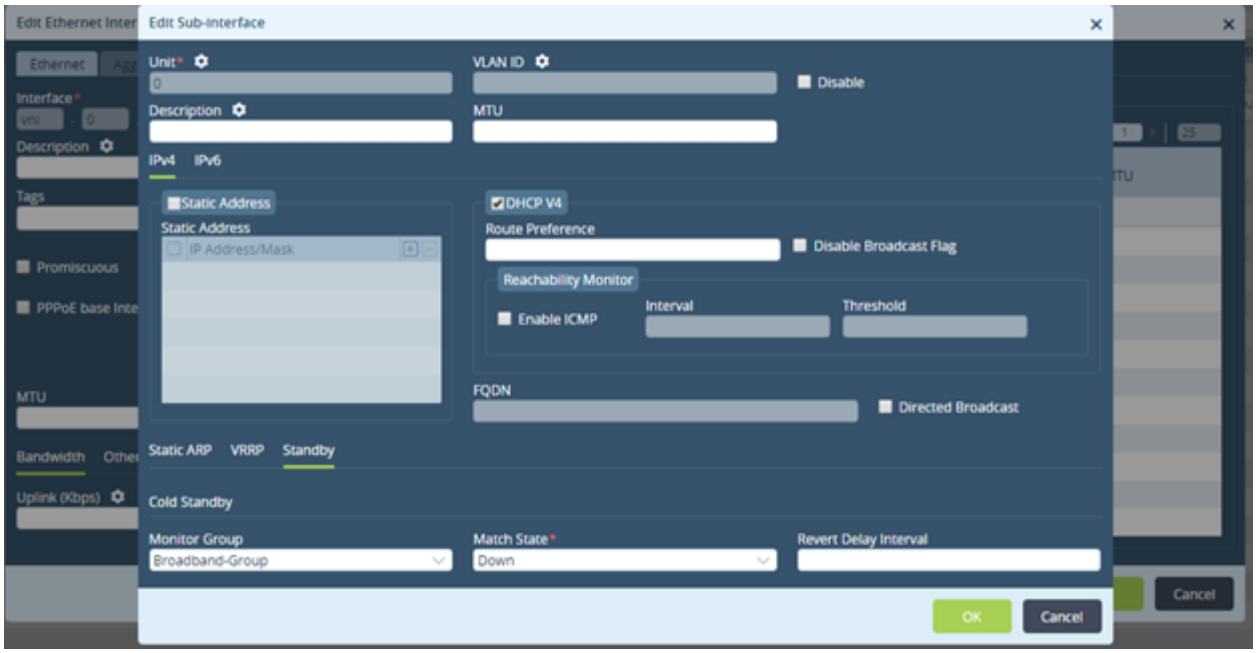
Next, you associate the monitor group (Broadband-Group) with an LTE interface as the standby option, with the match state configured as "down". To configure this use case scenario:

1. Continuing from the previous procedure, select Network > Interfaces in the left menu bar.
2. Select the LTE link (vni-0/100) in the main pane. The Edit Ethernet Interface popup window displays.



3. Select Subinterfaces

4. Select the Standby tab and enter information for the following fields.

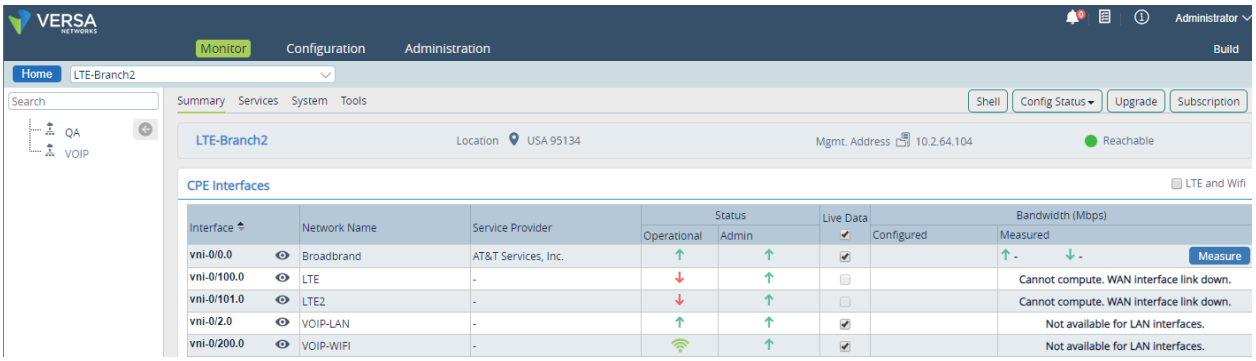


Field	Description
Monitor Group	Select the monitor group, here, Broadband-Group.
Match State	Select "down."

5. Click OK.

To verify the interface status:

1. In Director view, select the Administration tab in the top menu bar.
2. Select Appliances in the left menu bar.
3. Select the LTE-Branch2 device in the main pane. The view changes to Appliance view.
4. Select the Monitor tab in the top menu bar. The main pane displays the monitor dashboard for the LTE-Branch2 device.
5. Check the operational status of the branch's interfaces. If all the primary links are Up, the backup LTE interface should show a Down. status. If all the primary links are down, the backup LTE link status should show as up.



Comparison of LTE Hot Standby Mode and Cold Standby Mode

The following table compares hot standby mode and cold standby mode.

	Hot Standby Mode	**Cold Standby Mode**
LTE link status	Up	Down
Egress data traffic	Only SLA traffic	None
Ingress data traffic	SLA traffic but possibly also data traffic depending on the remote device configuration	None
Traffic failover to standby mode	Instantaneous	Not Instantaneous
Alarm generated when LTE link is down	Yes	No
Can be used for traffic when primary path SLA's are violated	Yes	No
Can use LTE as standby for selective traffic classes	Yes	No

Supported Software Information

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

Additional Information

Configure IP SLA Monitor Objects
Configure LTE
Configure SD-WAN Traffic Steering

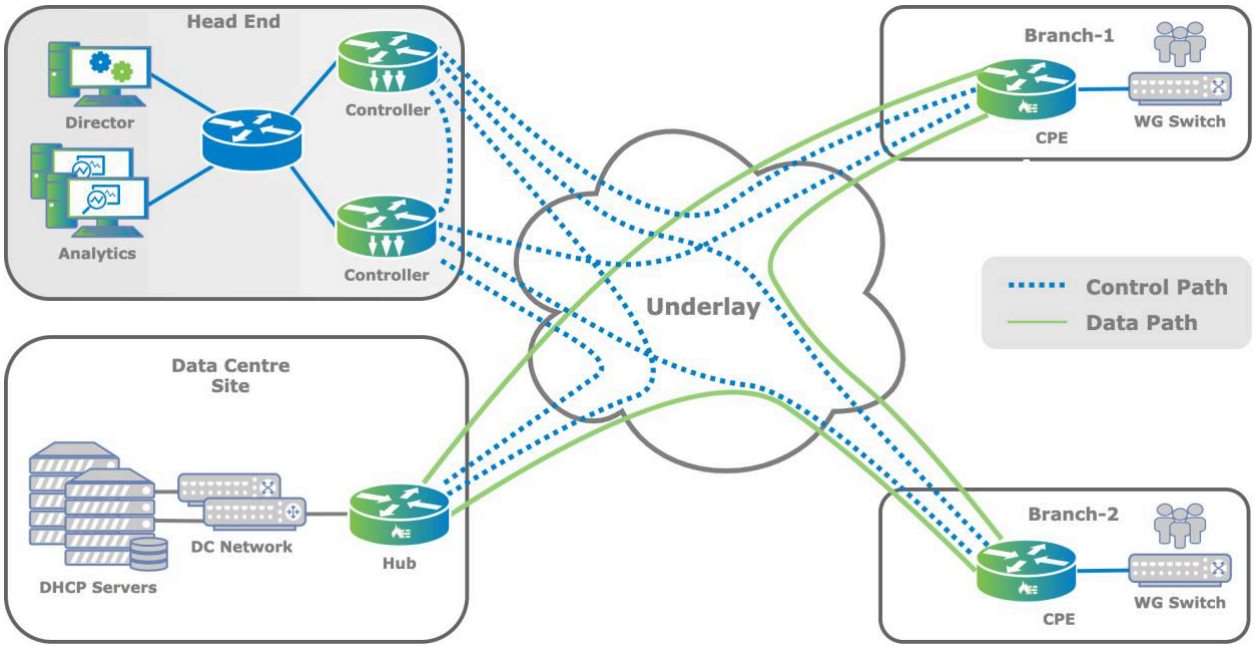
Configure Virtual Routers

SD-WAN Overlay Networks

For supported software information, click here.

The Versa Networks SD-WAN is based on overlay tunnels, and all traffic traveling through the tunnels is encrypted. In the SD-WAN overlay design you need to consider the IP addressing scheme for the overlay network and whether you want all data traffic to follow the encrypted overlay.

The following figure shows the topology of an SD-WAN overlay network, to illustrate the overlay and underlay networks. The network consists of a data center two single-homed remote branches that are managed by a Versa Networks headend, which consists of Director, Analytics and Controller node. All sites and the headend are connected to all the available transport networks.



Hubs and branches connect to the Controller node, which serves as an attachment point for the management plane and control plane. SD-WAN overlay topologies are built through the exchange of MP-BGP NLRI communities in combination with import and export policies, and they provide the flexibility to create multiple topologies, using Director Workflows, without any restrictions.

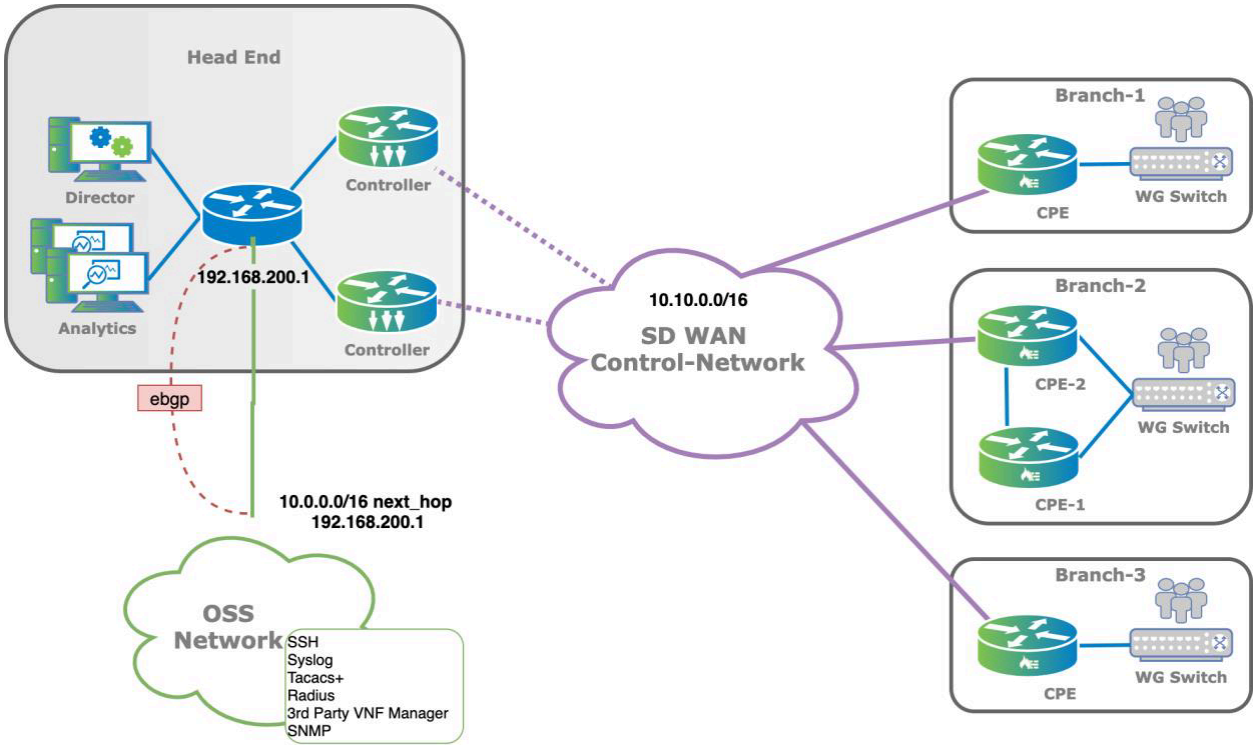
Overlay IP Addressing

The Versa Networks SD-WAN is based on overlay tunnels, which are used to abstract the underlay networks. By default, two overlay networks are built between the branches:

- Encrypted overlay, which uses an IPsec tunnel
- Plain-text overlay, which uses a VXLAN tunnel

For more information about the tunnels used for overlay networks, see Secure Control and Data Overlay Tunnel Solution.

The addresses for the SD-WAN overlay tunnels follow a specific overlay IP addressing scheme. In principle, the overlay network is routable in the SD-WAN control network, that is, between both the Controller nodes and the branches, and the control network southbound of the Director node (or northbound of the SD-WAN Controller node), as illustrated in the following figure.



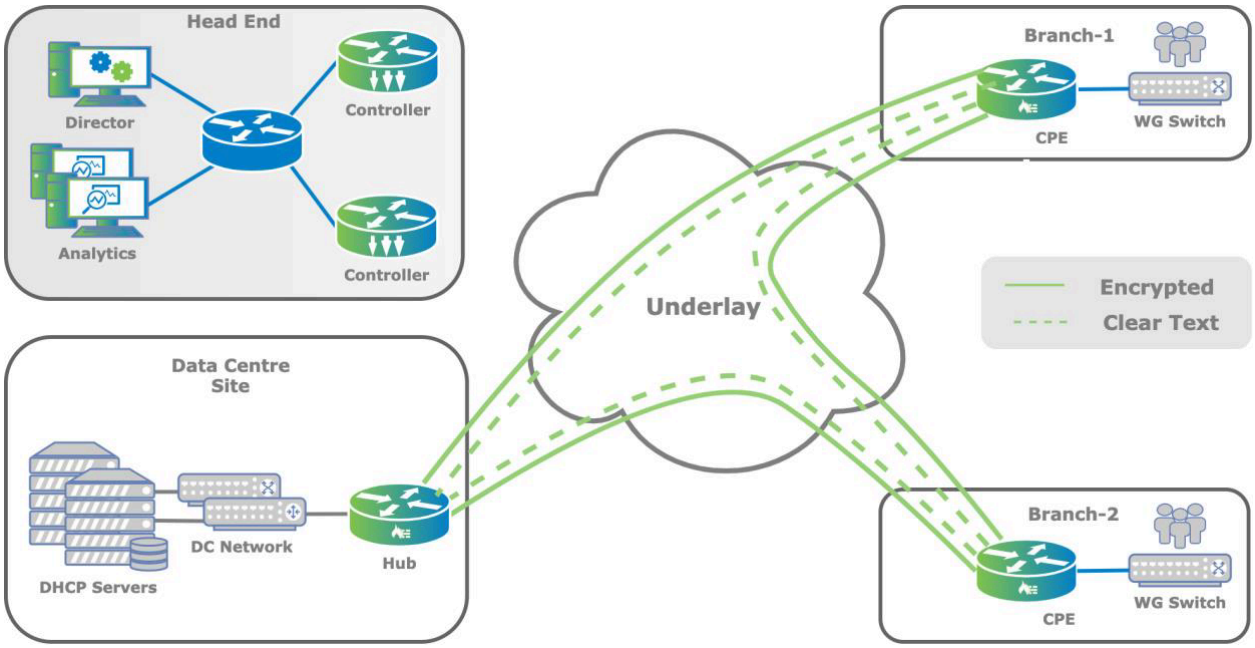
However, in some deployments, you must choose what to include in the overlay IP addressing scheme when you are integration with an OSS/BSS. For example, the control network must have IP address reachability between VOS edge devices and services such as TACACS+, RADIUS, syslog collectors, and third-party VNF managers. For more information, see Configure the Overlay Addressing Scheme

The following are best practices for overlay IP addressing:

- Configure the overlay IP addressing method and pool when you initially set up the Versa Director. You cannot modify the method later. Ensure that you choose the correct method and that you allocate a large enough subnet to cover the expected size of the deployment.
- It is recommended that you use the “do not encode” option to optimizes the use of IP addressing space.

Encrypted and Clear-Text Overlay

By default, all data traffic follows the encrypted overlay, as illustrated in the following figure.



If you need to change the default tunnel used for data transport, you can do so in one of the following ways:

- Statically configuring encrypted or clear-text transmission of data per WAN interface
- Dynamically setting the transmission mode by configuring an SD-WAN policy

Note that if you configure both WAN interface static definition and SD-WAN policy, the SD-WAN policy takes precedence.

Define Per-Interface Encryption Statically

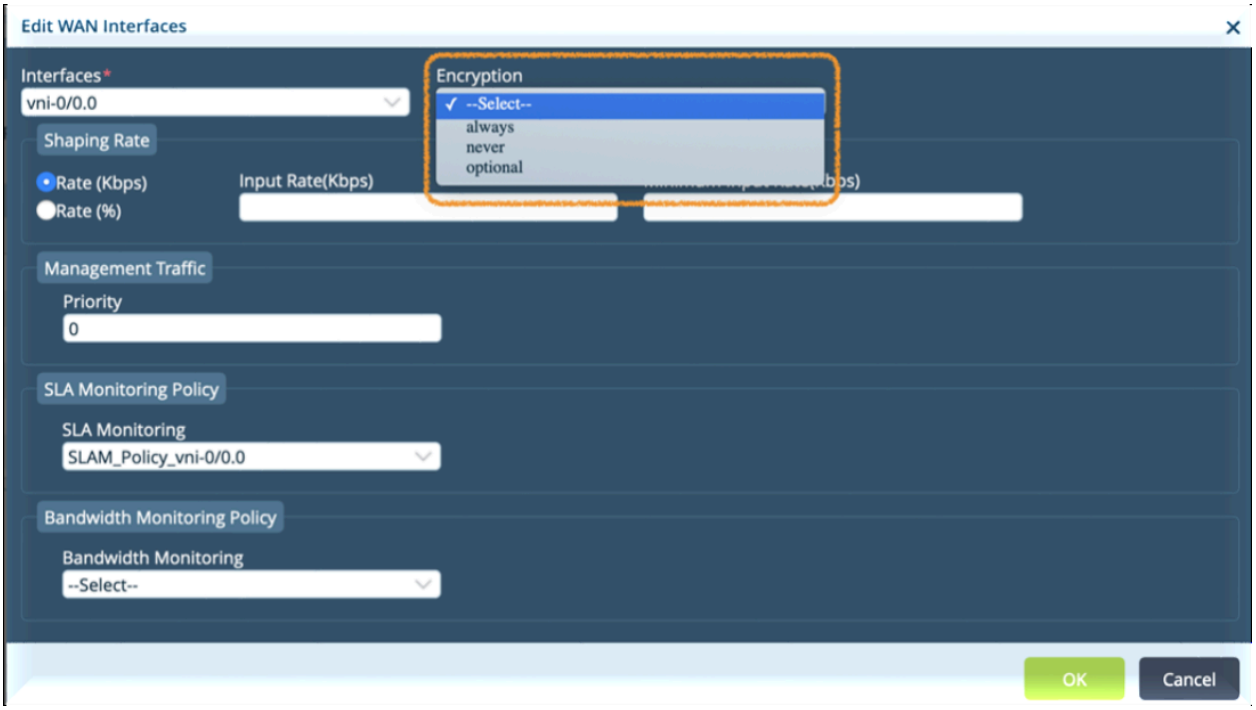
You can statically define the encryption method per WAN interface if the underlay is a private or secured circuit such as an MPLS service provided as part of a service provider Layer 3 VPN service. Traditionally, network administrators have considered these private Layer 3 VPNs as secured and did not typically encrypt data over it. In a similar manner, you can consider the MPLS Layer 3 VPN as secured, and so data can be transported using the clear-text tunnel.

A benefit of clear-text transport it that it does not use IPsec overhead on the platform

To configure static definition per WAN interface:

1. In Appliance view, select the Configuration tab in the top menu bar.
2. Select Services > SD-WAN > Site in the left menu bar.

3. In the Site pane, click the Edit icon.
4. In the WAN Interfaces tab, select a WAN interface. The Edit WAN Interfaces popup window displays.



5. In the Encryption field, select the desired encryption for the interface:
   - Always—Encrypt all traffic.
   - Never—Do not encrypt traffic.
   - Optional—Encryption is optional.
6. Click OK.

For more information, see Configure Encryption on WAN Interfaces.

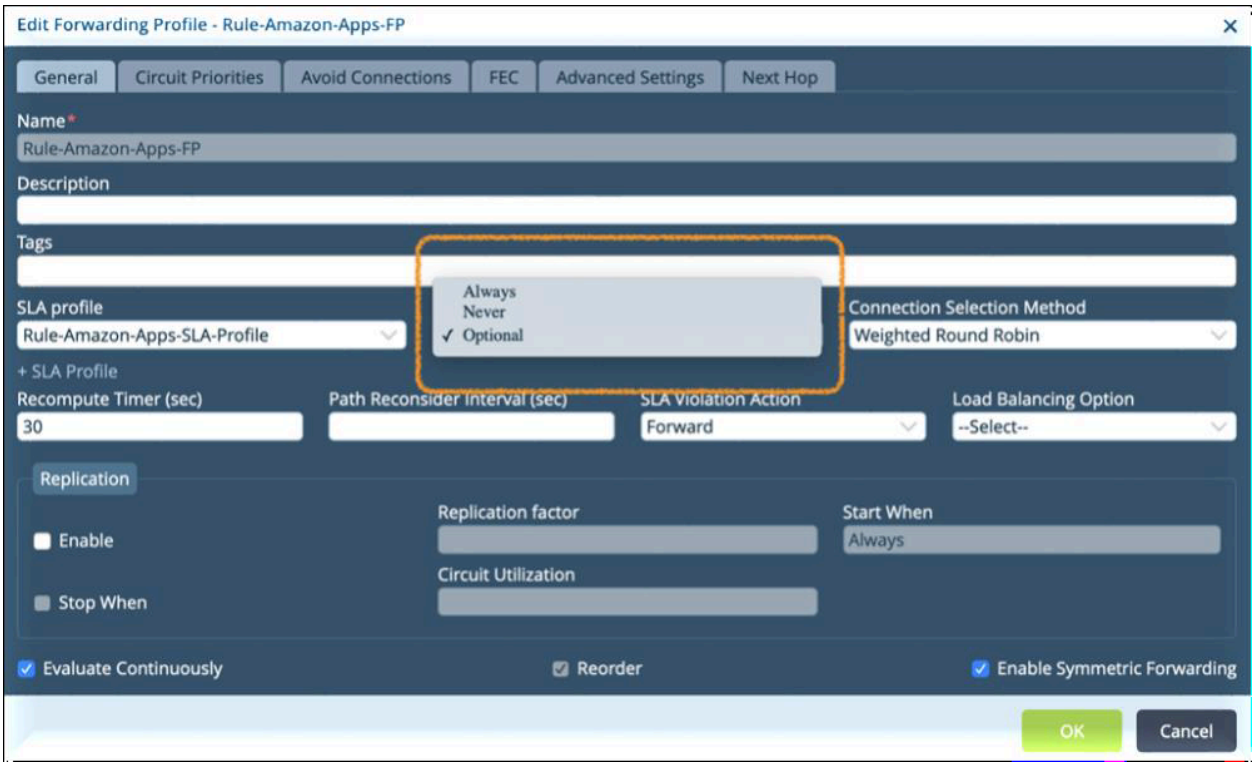
Use SD-WAN Policy To Dynamically Define Encryption

In some scenarios, you need to dynamically turn off encryption for some traffic, for example, for traffic that is already encrypted by an application (such as HTTPS or TLS/SSL secured application data) and traffic that is of no interest, from a security point of view, to the enterprise (such as recreational traffic from Facebook and YouTube). To turn off encryption for these types of traffic, you configure an SD-WAN forwarding profile, in which you set the desired encryption, and then you associate the forwarding with an SD-WAN policy that identifies the traffic to match.

To use SD-WAN policy to dynamically define encryption:

1. In Appliance view, select the Configuration tab in the top menu bar.
2. Select Services > SD-WAN > Forwarding Profiles in the left menu bar.

3. Click the Add icon. The Edit Forwarding Profile popup window displays.



4. In the Encryption field, select the desired encryption.

For more information, see Configure SD-WAN Traffic-Steering.

Supported Software Information

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

Additional Information

Configure Encryption on WAN Interfaces
Configure SD-WAN Traffic-Steering

SD-WAN Topologies

For supported software information, click here.

The Versa Networks solution supports the following SD-WAN overlay topologies:

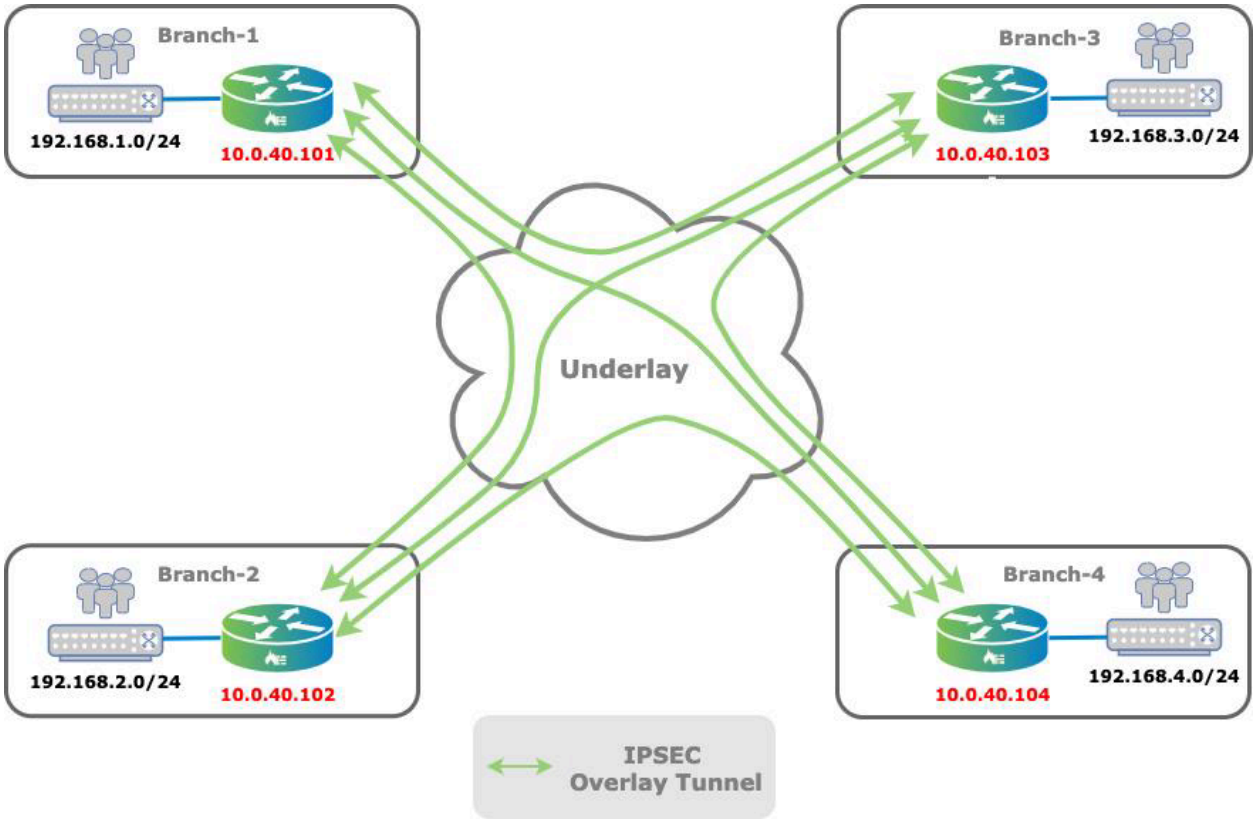
- Full mesh

- Hub and spoke
- Regional mesh
- Multi-VRF, or multitenancy

These topologies are established by using that well-known routing techniques that have been used for a long time in MPLS Layer 3 VPN networks. They use MP-BGP communities to achieve fine-grained route control and to provide flexible options for manipulating and fine-tuning routes. You can use Director Workflows to create these topologies, thus simplifying these complex configurations.

Full-Mesh Topology

You use a full-mesh topology for any-to-any communication. In this type of topology, branches communicate directly using overlay tunnels, and traffic does not need to transit through a hub or centralized site. The following figure illustrates a full-mesh topology.



A full-mesh topology is generally the preferred topology when branches must communicate directly with each other. Typically, you choose a full-mesh topology over a hub-and-spoke topology for voice applications, because a hub-and-spoke topology introduces delay when the hub is distant from the branches. Another instance in which a full-mesh topology is preferred over hub and spoke is a distributed security architecture, where policy enforcement is performed at the branch. Here, the full-mesh topology avoids the need to funnel traffic to hub sites for inspection.

In Director Workflows, the full-mesh topology is the default option.

In a full-mesh topology SLA monitoring probes are sent to every remote branch on every available transport. SLA monitoring probes are used to track reachability and to measure link metrics for each access circuit towards any given remote site. You can use SLA optimization features such as adaptive SLA and data-driven SLA to optimize the SLA load in large deployments.To verify SLA monitoring status, issue the following CLI command:

```
admin@Branch-1-cli> show orgs org Tenant1 sd-wan sla-monitor status
                                    LOCAL  REMOTE
                                    WAN    WAN
          PATH     FWD    LOCAL     REMOTE    LINK   LINK    ADAPTIVE     DAMP     DAMP   CONN
LAST
SITE NAME     HANDLE    CLASS  WAN LINK   WAN LINK ID     ID      MONITORING   STATE    FLAPS
STATE  FLAPS  FLAPPED
-------------------------------------------------------------------------------------------------------------------
Branch-2     6689028   fc_ef   MPLS      MPLS      1      1       active      disable  0      up     1      00:06:26
          6693380   fc_ef   Internet  Internet  2      2       active      disable  0      up     1      00:06:26
Branch-3     6754564   fc_ef   MPLS      MPLS      1      1       active      disable  0      up     1      00:06:32
          6758916   fc_ef   Internet  Internet  2      2       active      disable  0      up     1      00:06:31
Branch-4     6820100   fc_ef   MPLS      MPLS      1      1       active      disable  0      up     2      00:05:45
          6824452   fc_ef   Internet  Internet  2      2       active      disable  0      up     2      00:05:44
Controller-1 69888     fc_nc   MPLS      MPLS      1      1       disable     disable  0      up     1      00:16:38
          74240     fc_nc   Internet  Internet  2      2       disable     disable  0      up     1      00:16:38
```

The output above shows the SLA monitoring view from Branch-1, which has internet and MPLS transports towards all branches and towards the Controller node.

In a full-mesh topology, you must determine the proper scaling of the maximum number of branches. To dimension the deployment, you must consider many variables, including the following:

- Number of WAN links
- Number of tenants
- Forwarding classes being monitored
- SLA monitor interval
- Branch hardware
- Bandwidth to the branch

For example, in a full-mesh topology with 1000 branches that have one tenant per site and two WAN links in different transport domains, if you use the Versa Operating System™ (VOS™) device default SLA monitoring configuration, the SLA probe traffic consumes 6.25 Mbps of bandwidth at each site. Increasing the number of branch CPE devices increases both the bandwidth usage and the CPU overhead to perform SLA monitoring. You can limit the SLA monitoring traffic to lower link utilization, for instance, on high-cost links such as LTE connections. For more information, see Configure SLA Monitoring for SD-WAN Traffic Steering.

In a full-mesh topology, there is direct reachability to prefixes in remote branches. Traffic is routed to those prefixes using the next hop of the remote branch loopback (TVI) interfaces. If there is an underlay cut or the SLA probing cannot declare the remote branch to be reachable, the SLA monitoring session is down and therefore the next hop is not

reachable. The result is that this prefix is withdrawn from the routing table and thus it is not displayed in the output of the **show route** command. Note that if the route is not directly reachable, perhaps because it is reachable through a hub, the route has the interface name "indirect." The route table below illustrates that for the Branch-1 VRF, three of the are indirect routes.

```
admin@Branch-1-cli> show route routing-instance Tenant1-LAN-VR
Routes for Routing instance : Tenant1-LAN-VR AFI: ipv4 SAFI: unicast
Codes: E1 - OSPF external type 1, E2 - OSPF external type 2
IA - inter area, iA - intra area,
L1 - IS-IS level-1, L2 - IS-IS level-2
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
RTI - Learnt from another routing-instance
+ - Active Route
Prot    Type     Dest Address/Mask   Next-hop       Age      Interface name
----    ----     -----------------   --------       ---      --------------
BGP     N/A      +0.0.0.0/0          169.254.0.2    1w6d20h  tvi-0/603.0
conn    N/A      +169.254.0.2/31     0.0.0.0        1w6d20h  tvi-0/603.0
local   N/A      +169.254.0.3/32     0.0.0.0        1w6d20h  directly connected
conn    N/A      +192.168.1.0/24     0.0.0.0        1w6d20h  vni-0/2.0
local   N/A      +192.168.1.1/32     0.0.0.0        1w6d20h  directly connected
BGP     N/A      +192.168.2.0/24     10.0.40.102    1w6d20h  Indirect
BGP     N/A      +192.168.3.0/24     10.0.40.103    1w6d20h  Indirect
BGP     N/A      +192.168.4.0/24     10.0.40.104    00:05:58 Indirect
```

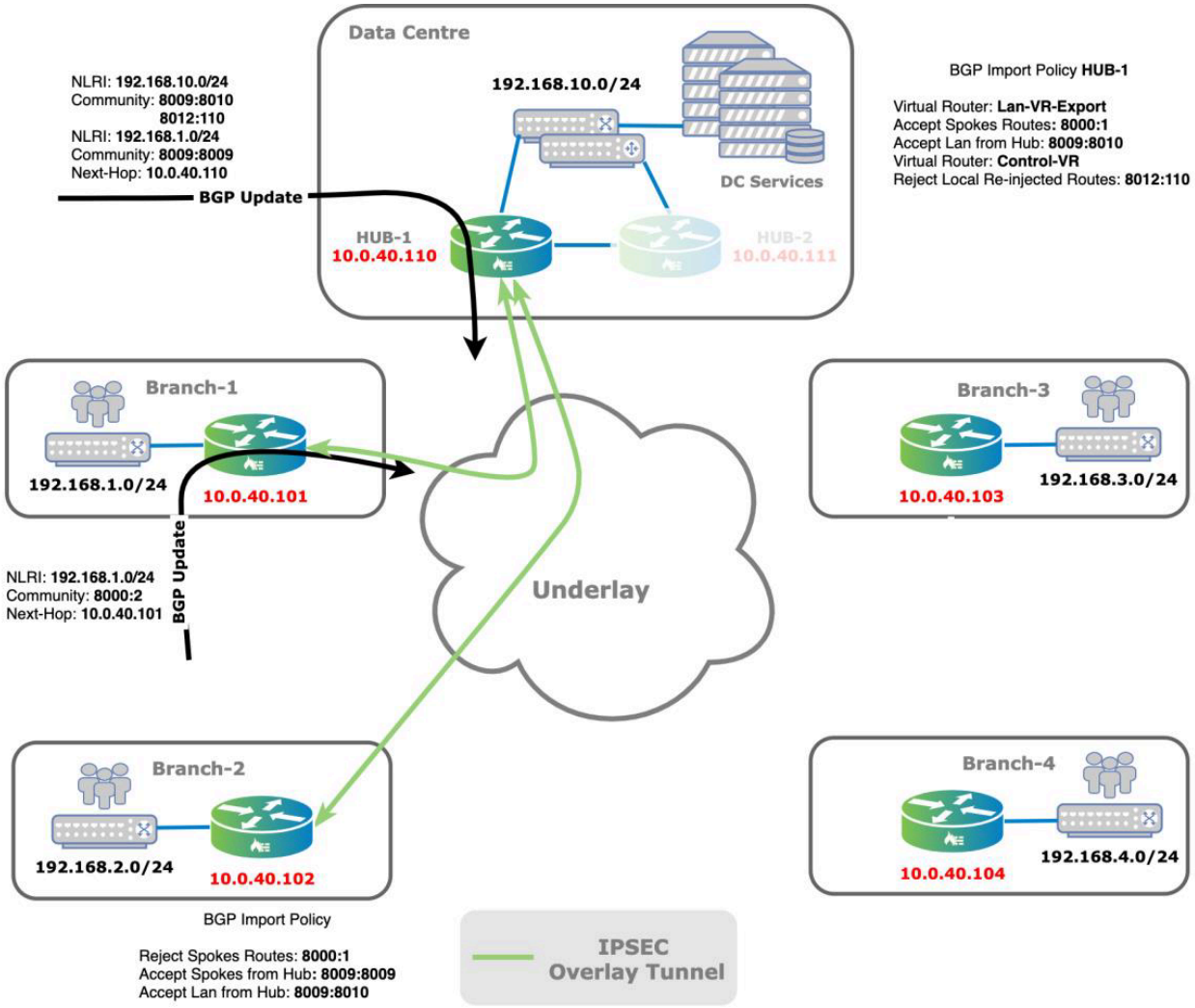
Hub-and-Spoke Topology

The Versa Networks SD-WAN solution supports different types of hub-and-spoke topologies:

- Spoke to hub only
- Spoke to spoke through a hub (spoke to spoke through another SD-WAN edge device)
- Spoke to spoke direct
- Spoke-hub-hub-spoke

Spoke-to-Hub Only

In a spoke-to-hub-only topology, the only prefixes advertised, by default, are hub routes, and spokes routes are not re-advertised by the hub branch. You use this topology when spokes do not have to communicate with each other. A good example is a network of ATM cash machines in which devices communicate exclusively with resources in the customer data center. The following figures shows that spoke prefixes are accepted only by the hub and that they are rejected by other spokes based on the BGP community configuration.



The following CLI output shows that the spoke Branch-1 VRF route table contains routes only from the hub:

```
admin@Branch-1-cli> show route routing-instance Tenant1-LAN-VR
Routes for Routing instance : Tenant1-LAN-VR AFI: ipv4 SAFI: unicast
Codes: E1 - OSPF external type 1, E2 - OSPF external type 2
IA - inter area, iA - intra area,
L1 - IS-IS level-1, L2 - IS-IS level-2
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
RTI - Learnt from another routing-instance
+ - Active Route

Prot   Type   Dest Address/Mask   Next-hop       Age      Interface name
----   ----   -----------------   --------       ---      --------------
conn   N/A   +192.168.1.0/24      0.0.0.0        2d20h04m  vni-0/2.0
local  N/A   +192.168.1.1/32      0.0.0.0        2d20h04m  directly connected
BGP    N/A   +192.168.10.0/24     10.0.40.110    2d20h04m  Indirect
BGP    N/A    192.168.10.0/24     10.0.40.111    2d20h04m  Indirect
```

The route table on spoke Branch-1 shows only destinations behind hubs, again with Hub-1 being preferred, and the

table shows no spokes routes. The following output shows the prefixes advertised by spoke Branch-1:

```
admin@Branch-1-cli> show route table l3vpn.ipv4.unicast advertising-protocol bgp
Routes for Routing instance : Tenant1-Control-VR AFI: ipv4 SAFI: unicast

Routing entry for 192.168.1.0/24
   Peer Address          : 10.0.40.1
   Route Distinguisher   : 2L:2
   Next-hop              : 10.0.40.101
   VPN Label             : 24704
   Local Preference      : 110
   AS Path               : N/A
   Origin                : Igp
   MED                   : 0
   Community             : [ 8000:2 8001:110 8002:111 ]
   Extended community    : [ target:2L:2 ]
```

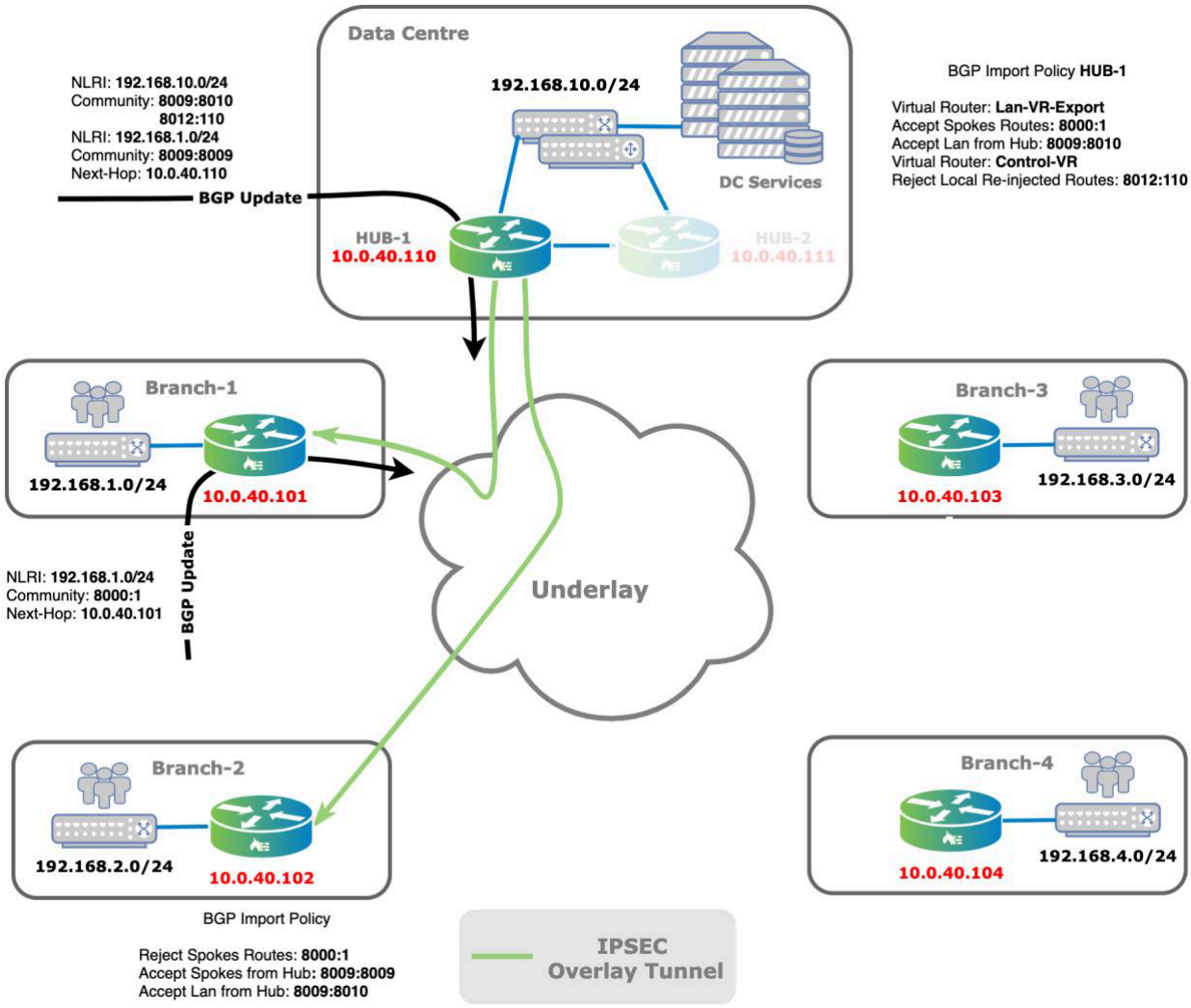
You use BGP import policies to filter spoke routes. For example, using spoke community 8000:2 filters out spoke routes on hubs in the LAN-VR-Export VR, and these routes are not advertised back to the spokes. Therefore, the hub route tables contains all spoke prefixes:

```
admin@Hub-1-cli> show route routing-instance Tenant1-LAN-VR
Routes for Routing instance : Tenant1-LAN-VR AFI: ipv4 SAFI: unicast
Codes: E1 - OSPF external type 1, E2 - OSPF external type 2
IA - inter area, iA - intra area,
L1 - IS-IS level-1, L2 - IS-IS level-2
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
RTI - Learnt from another routing-instance
+ - Active Route

Prot   Type  Dest Address/Mask   Next-hop       Age       Interface name
----   ----  -----------------   --------       ---       --------------
BGP    N/A  +192.168.1.0/24      10.0.40.101     00:21:24  Indirect
BGP    N/A  +192.168.2.0/24      10.0.40.102     00:21:27  Indirect
BGP    N/A  +192.168.3.0/24      10.0.40.103     00:21:23  Indirect
BGP    N/A  +192.168.4.0/24      10.0.40.104     00:21:26  Indirect
BGP    N/A   192.168.10.0/24     10.0.40.111     00:36:45  Indirect
conn   N/A  +192.168.10.0/24     0.0.0.0         00:51:13  vni-0/2.0
local  N/A  +192.168.10.1/32     0.0.0.0         00:51:13  directly connected
```

On hubs, all spokes prefixes are installed in the corresponding VRF. You can implement redistribution policy on hubs to perform route summarization or to generate default a static route, for instance, to attract traffic from spokes.

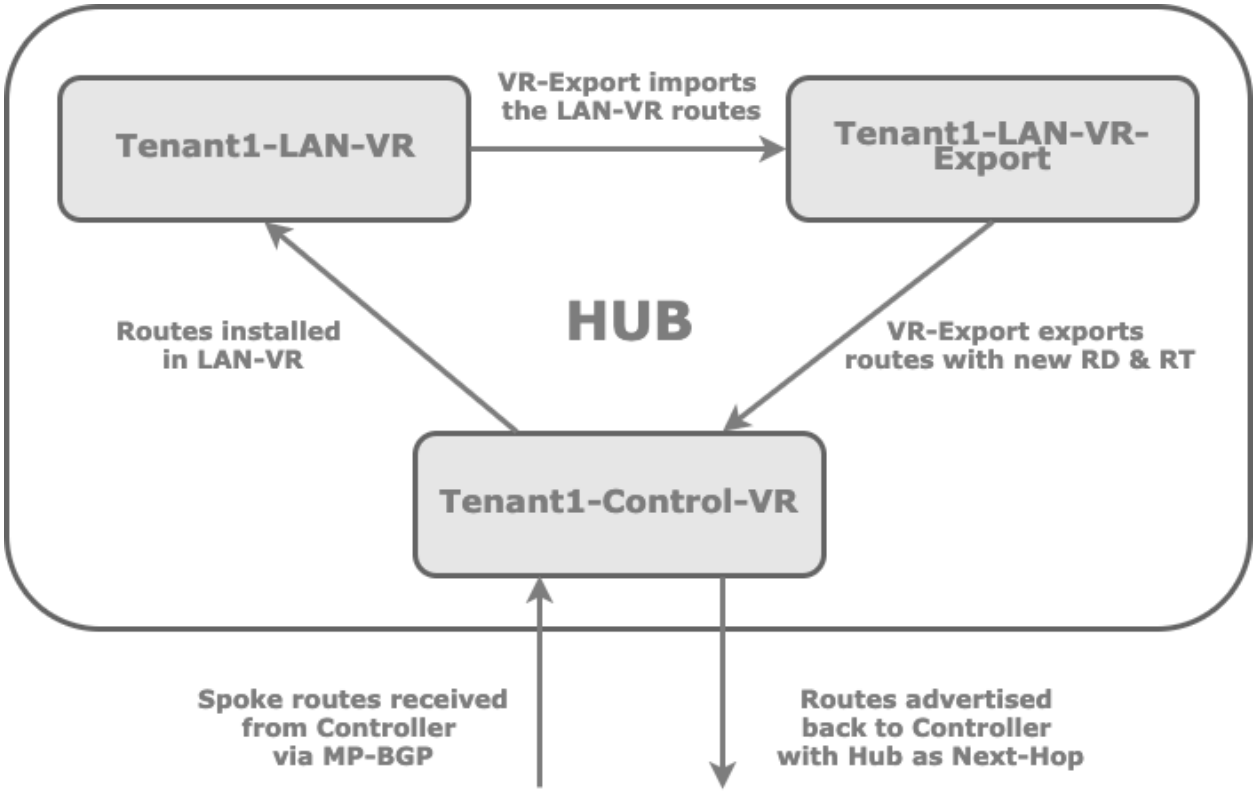
Spoke-to-Spoke via Hub (Spoke-to-Spoke Through Another SD-WAN Edge Device)

In a spoke-to-spoke via hub topology (spoke-to-spoke through another SD-WAN edge device), spoke sites are connected to each other through hub site. The data path between two spokes travels through the hub. The following figure illustrates this topology.



You can use the spoke-to-spoke through a hub topology when communication between branches is not required, for example, when security and other services are centralized at the hub site or when the cost of ownership of WAN links dictates it.

The following figure shows the method that hub sites use manipulate VRF spoke routes. With this method, you change the route distinguisher on the hub for a set of VRF routes that you are advertising so that the Controller nodes can separate the routes and accept them during BGP route selection. The Controller nodes have the original routes from the spokes and the spoke routes advertised by the hubs, and they reflects them to the branches. You use route-target filtering on the spokes to perform the remainder of the route selection. With route-target filter, you import the hub-advertised spoke routes, and the Controller nodes use these routes to select the hub as the next hop towards the remote sites.



In a spoke-to-spoke via hub topology, the branches communicate only through the hub. The IP prefixes of remote branches always have the hub as the next hop.

SLA monitoring is active only on paths towards hub sites and Controller nodes, and spoke sites are not monitored. This reduces the amount of SLA probe traffic compared to a full-mesh topology and addresses the concerns of scalability in deployments that have a large number of branches.

The following CLI output shows the SLA monitoring view on Branch-1:

```
admin@Branch-1-cli> show orgs org Tenant1 sd-wan sla-monitor status
                                  LOCAL  REMOTE
                                  WAN    WAN
            PATH    FWD    LOCAL    REMOTE    LINK   LINK   ADAPTIVE    DAMP    DAMP   CONN
LAST
SITE NAME   HANDLE  CLASS  WAN LINK  WAN LINK  ID     ID     MONITORING  STATE   FLAPS
STATE  FLAPS  FLAPPED
-------------------------------------------------------------------------------------------------------------------------------------------
Controller-1  69888   fc_nc  MPLS      MPLS      1      1      disable     disable  0      up     1      3d01h39m
            74240   fc_nc  Internet  Internet  2      2      disable     disable  0      up     1      3d01h39m
Hub-1         7213316 fc_ef  MPLS      MPLS      1      1      suspend     disable  0      up     1      2d21h02m
            7217668 fc_ef  Internet  Internet  2      2      suspend     disable  0      up     1      2d21h02m
Hub-2         7278852 fc_ef  MPLS      MPLS      1      1      suspend     disable  0      up     1      2d21h00m
            7283204 fc_ef  Internet  Internet  2      2      suspend     disable  0      up     1      2d21h00m
```

Because the hub nodes re-advertise the spoke branch prefixes, the spoke branches learn all the spoke prefixes. The

next-hop IP address for the spoke branches is the hub's loopback TVI address.

The following output from the Branch-1 VRF routes table shows a deployment with two hubs. The hub that you configure with a higher priority is the one that maintains the active route

```
admin@Branch-1-cli> show route routing-instance Tenant1-LAN-VR
Routes for Routing instance : Tenant1-LAN-VR AFI: ipv4 SAFI: unicast
Codes: E1 - OSPF external type 1, E2 - OSPF external type 2
IA - inter area, iA - intra area,
L1 - IS-IS level-1, L2 - IS-IS level-2
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
RTI - Learnt from another routing-instance
+ - Active Route

Prot   Type   Dest Address/Mask   Next-hop     Age        Interface name
----   ----   -----------------   --------     ---        --------------
conn   N/A   +192.168.1.0/24      0.0.0.0      1d23h16m   vni-0/2.0
local  N/A   +192.168.1.1/32      0.0.0.0      1d23h16m   directly connected
BGP    N/A   +192.168.2.0/24      10.0.40.110  03:26:28   Indirect
BGP    N/A    192.168.2.0/24      10.0.40.111  03:26:28   Indirect
BGP    N/A   +192.168.3.0/24      10.0.40.110  1d23h16m   Indirect
BGP    N/A    192.168.3.0/24      10.0.40.111  1d23h16m   Indirect
BGP    N/A   +192.168.4.0/24      10.0.40.110  1d23h16m   Indirect
BGP    N/A    192.168.4.0/24      10.0.40.111  1d23h16m   Indirect
BGP    N/A   +192.168.10.0/24     10.0.40.110  1d23h16m   Indirect
BGP    N/A    192.168.10.0/24     10.0.40.111  1d23h16m   Indirect
```

The following CLI output shows an example of the spoke routes advertised by Branch-1:

```
admin@Branch-1-cli> show route table l3vpn.ipv4.unicast advertising-protocol bgp
Routes for Routing instance : Tenant1-Control-VR AFI: ipv4 SAFI: unicast

Routing entry for 192.168.1.0/24
   Peer Address         : 10.0.40.1
   Route Distinguisher  : 2L:2
   Next-hop             : 10.0.40.101
   VPN Label            : 24704
   Local Preference     : 110
   AS Path              : N/A
   Origin               : Igp
   MED                  : 0
   Community            : [ 8000:1 8001:110 8002:111 ]
   Extended community   : [ target:2L:2 ]
```

The community string 8000:1 marks the spokes routes so that the BGP import policy on the spokes can identify them, and in this case, it rejects routes starting with this community string. The following CLI output is an example of a spoke route advertised by Hub-1:

```
admin@Hub-1-cli> show route table l3vpn.ipv4.unicast advertising-protocol bgp

Routing entry for 192.168.2.0/24
   Peer Address         : 10.0.40.1
```

```
Route Distinguisher  : 16002L:110
Next-hop             : 10.0.40.110
VPN Label            : 24705
Local Preference     : 100
AS Path              : N/A
Origin               : Incomplete
MED                  : 0
Community            : [ 8000:0 8000:1 8001:110 8002:111 8009:8009 ]
Extended community   : [ target:16002L:0 target:16002L:110 ]
```

The community string 8000:0 marks the spokes routes advertised by hubs so that the BGP import policy can identify them, and in this case, it accepts these routes, which have the hub as the next hop.

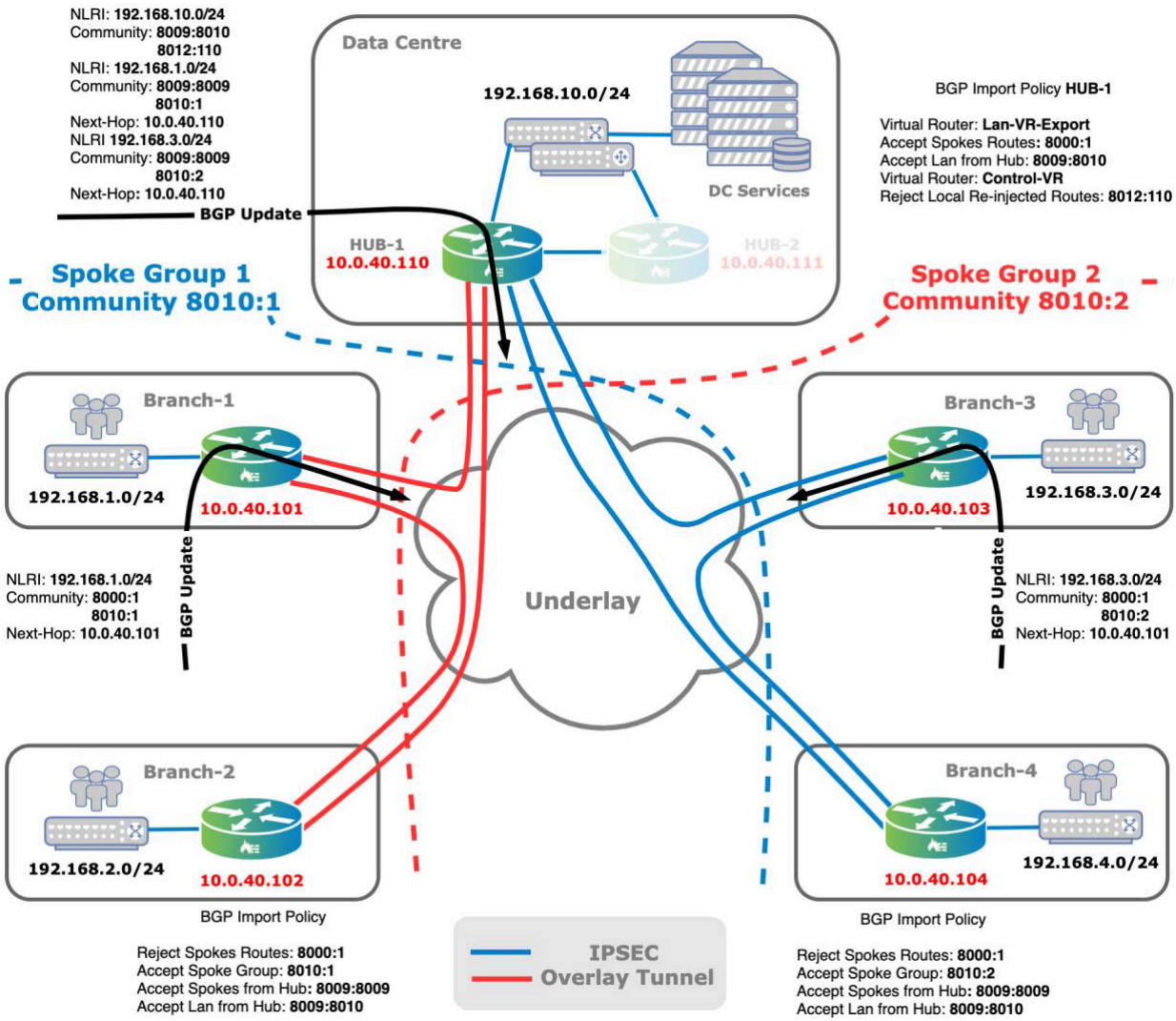
The community string 8009:8010, which you can see in the diagram above for Hub-1 (192.168.10.0/24), marks the direct LAN route from hubs, which the BGP import policy also accepts.

In this topology, Hub-1 is configured with a higher priority than Hub-2. This configuration explains why there are two entries in the route table for each prefix and why Hub-1 is the preferred next hop. Having two hubs provides redundancy, because Hub-2 is used when Hub-1 is not reachable.

The BGP import policy uses the extended-community attribute to accept routes from the hubs and to set a higher local preference for Hub-1. The extended target string 16002L:110 is derived from site ID 110, which is Hub-1.

Spoke-to-Spoke Direct and Partial Mesh

In a partial-mesh topology, some nodes are directly attached to each other, while other nodes are attached only to one or two nodes. You can select this topology when there are geographically dispersed sites in the same region that you want to communicate directly with each other, and when you want inter-regional traffic to transit through hub branches or when there is a high level of traffic exchanged between specific sites. The following figure illustrates a partial-mesh topology.



For a spoke-to-spoke–direct topology, you use spoke groups. Branches within the same spoke group can communicate directly with each other, and they use hubs to reach branches in different spoke groups. In the topology shown above, Branch-1 and Branch-2 communicate with each other directly, but to reach Branch-3 the next hop is Hub-1. The hubs are connected using a full-mesh topology.

It is recommended that you deploy a spoke-to-spoke–direct topology whenever feasible, so that you can use spoke groups to provide redundancy and flexible meshing of branches.

The following CLI output shows the SLA monitoring view on Branch-1:

```
admin@Branch-1-cli> show orgs org Tenant1 sd-wan sla-monitor status
                                    LOCAL REMOTE
                                    WAN   WAN
          PATH    FWD    LOCAL     REMOTE    LINK  LINK    ADAPTIVE    DAMP     DAMP   CONN
LAST
SITE NAME     HANDLE   CLASS  WAN LINK  WAN LINK  ID    ID      MONITORING  STATE    FLAPS
STATE  FLAPS  FLAPPED
```

```
-------------------------------------------------------------------------------------------------------------------
Branch-2     6689028 fc_ef MPLS     MPLS     1    1       active    disable 0     up     1     00:04:59
             6693380 fc_ef Internet Internet 2    2       active    disable 0     up     1     00:04:59
Controller-1 69888   fc_nc MPLS     MPLS     1    1       disable   disable 0     up     5     1d22h41m
             74240   fc_nc Internet Internet 2    2       disable   disable 0     up     1     1d22h45m
Hub-1        7213316 fc_ef MPLS     MPLS     1    1       suspend   disable 0     up     7     02:48:59
             7217668 fc_ef Internet Internet 2    2       suspend   disable 0     up     3     02:48:58
Hub-2        7278852 fc_ef MPLS     MPLS     1    1       suspend   disable 0     up     5     02:48:41
             7283204 fc_ef Internet Internet 2    2       suspend   disable 0     up     3     02:48:41
```

SLA monitoring is performed on the paths towards Hub-1, Hub-2, and Branch-2, because these belong to the same spoke group. SLA monitoring is not performed on branches in different spoke groups.

The following CLI output shows the Branch-1 VRF route table:

```
admin@Branch-1-cli> show route routing-instance Tenant1-LAN-VR
Routes for Routing instance : Tenant1-LAN-VR AFI: ipv4 SAFI: unicast
Codes: E1 - OSPF external type 1, E2 - OSPF external type 2
IA - inter area, iA - intra area,
L1 - IS-IS level-1, L2 - IS-IS level-2
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
RTI - Learnt from another routing-instance
+ - Active Route

Prot   Type  Dest Address/Mask   Next-hop        Age     Interface name
----   ----  -----------------   --------        ---     --------------
conn   N/A  +192.168.1.0/24      0.0.0.0         3d18h59m vni-0/2.0
local  N/A  +192.168.1.1/32      0.0.0.0         3d18h59m directly connected
BGP    N/A  +192.168.2.0/24      10.0.40.102     00:25:29 Indirect
BGP    N/A   192.168.2.0/24      10.0.40.110     00:25:30 Indirect
BGP    N/A   192.168.2.0/24      10.0.40.111     00:25:30 Indirect
BGP    N/A  +192.168.3.0/24      10.0.40.110     00:20:49 Indirect
BGP    N/A   192.168.3.0/24      10.0.40.111     00:20:48 Indirect
BGP    N/A  +192.168.4.0/24      10.0.40.110     00:24:56 Indirect
BGP    N/A   192.168.4.0/24      10.0.40.111     00:24:56 Indirect
BGP    N/A  +192.168.10.0/24     10.0.40.110     03:09:26 Indirect
BGP    N/A   192.168.10.0/24     10.0.40.111     02:55:29 Indirect
```

A spoke-to-spoke direct topology incorporates additional redundancy. The CLI output above shows three route table entries for 192.168.2.0/24. These routes are advertised by Branch-2 as well as by Hub-1 and Hub-2. When there are underlay connectivity issues between Branch-1 and Branch-2, the hubs provide a redundant path to reach the Branch-2 prefixes.

Routes for branches in different spoke groups are only reachable through the hubs.

To create the expected topology, you use BGP community strings and import policies to accept or reject routes.

The following CLI output shows the prefixes advertised by Branch-2:

```
admin@Branch-2-cli> show route table l3vpn.ipv4.unicast advertising-protocol bgp

Routes for Routing instance : Internet-Transport-VR AFI: ipv4 SAFI: unicast
```

```
Routes for Routing instance : MPLS-Transport-VR AFI: ipv4 SAFI: unicast
Routes for Routing instance : Tenant1-Control-VR AFI: ipv4 SAFI: unicast

Routing entry for 192.168.2.0/24
   Peer Address         : 10.0.40.1
   Route Distinguisher  : 2L:2
   Next-hop             : 10.0.40.102
   VPN Label            : 24704
   Local Preference     : 110
   AS Path              : N/A
   Origin               : Igp
   MED                  : 0
   Community            : [ 8000:1 8001:110 8002:111 8010:1 ]
   Extended community   : [ target:2L:2 ]
```

Here, the community string 8010:1 corresponds to spoke Group-1. The community string is a unique BGP community that is associated with the spoke group and that you assign during the Workflow configuration. Based on this community string, the import policy based is pushed to the spokes that are in the same spoke group.

In the topology shown in the figure above, Hub-1 has a higher priority. You configure an import policy that manipulates the Local-Pref attribute (Branch-2 > Hub-1 > Hub-2) to prefer the routes advertised by Hub-1. The following output shows the Local Preference configuration on Branch-1:

```
admin@Branch-1-cli> show route table l3vpn.ipv4.unicast receive-protocol bgp 192.168.2.0

Routes for Routing instance : Internet-Transport-VR AFI: ipv4 SAFI: unicast

Routes for Routing instance : MPLS-Transport-VR AFI: ipv4 SAFI: unicast

Routes for Routing instance : Tenant1-Control-VR AFI: ipv4 SAFI: unicast

Routing entry for 192.168.2.0/24
   Peer Address         : 10.0.40.1
   Route Distinguisher  : 2L:2
   Next-hop             : 10.0.40.102
   VPN Label            : 24704
   Local Preference     : 110
   AS Path              : N/A
   Origin               : Igp
   MED                  : 0
   Community            : [ 8000:1 8001:110 8002:111 8009:8009 8010:1 ]
   Extended community   : [ target:2L:2 ]
   Preference           : Default

Routing entry for 192.168.2.0/24
   Peer Address         : 10.0.40.1
   Route Distinguisher  : 16002L:110
   Next-hop             : 10.0.40.110
   VPN Label            : 24705
   Local Preference     : 102
   AS Path              : N/A
```

```
    Origin              : Incomplete
    MED                 : 0
    Community           : [ 8000:0 8000:1 8001:110 8002:111 8009:8009 8009:8009 8010:1 ]
    Extended community  : [ target:16002L:0 target:16002L:110 ]
    Preference          : Default

Routing entry for 192.168.2.0/24
    Peer Address        : 10.0.40.1
    Route Distinguisher : 16002L:111
    Next-hop            : 10.0.40.111
    VPN Label           : 24705
    Local Preference    : 101
    AS Path             : N/A
    Origin              : Incomplete
    MED                 : 0
    Community           : [ 8000:0 8000:1 8001:110 8002:111 8009:8009 8009:8009 8010:1 ]
    Extended community  : [ target:16002L:0 target:16002L:111 ]
    Preference          : Default
```

The following CLI output shows the entries in the Branch-1 route table:

```
admin@Hub-1-cli> show route routing-instance Tenant1-LAN-VR

Routes for Routing instance : Tenant1-LAN-VR AFI: ipv4 SAFI: unicast

Codes: E1 - OSPF external type 1, E2 - OSPF external type 2
IA - inter area, iA - intra area,
L1 - IS-IS level-1, L2 - IS-IS level-2
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
RTI - Learnt from another routing-instance
+ - Active Route

Prot   Type  Dest Address/Mask  Next-hop       Age        Interface name
----   ----  -----------------  --------       ---        --------------
BGP    N/A  +192.168.1.0/24     10.0.40.101     1d03h20m   Indirect
BGP    N/A   192.168.1.0/24     10.0.40.111     1d03h20m   Indirect
BGP    N/A  +192.168.2.0/24     10.0.40.102     1d03h20m   Indirect
BGP    N/A   192.168.2.0/24     10.0.40.111     1d03h20m   Indirect
BGP    N/A  +192.168.3.0/24     10.0.40.103     1d03h19m   Indirect
BGP    N/A   192.168.3.0/24     10.0.40.111     1d03h19m   Indirect
BGP    N/A  +192.168.4.0/24     10.0.40.104     1d03h19m   Indirect
BGP    N/A   192.168.4.0/24     10.0.40.111     1d03h19m   Indirect
BGP    N/A   192.168.10.0/24    10.0.40.111     1w4d03h    Indirect
conn   N/A  +192.168.10.0/24    0.0.0.0         1w4d03h    vni-0/2.0
local  N/A  +192.168.10.1/32    0.0.0.0         1w4d03h    directly connected
```

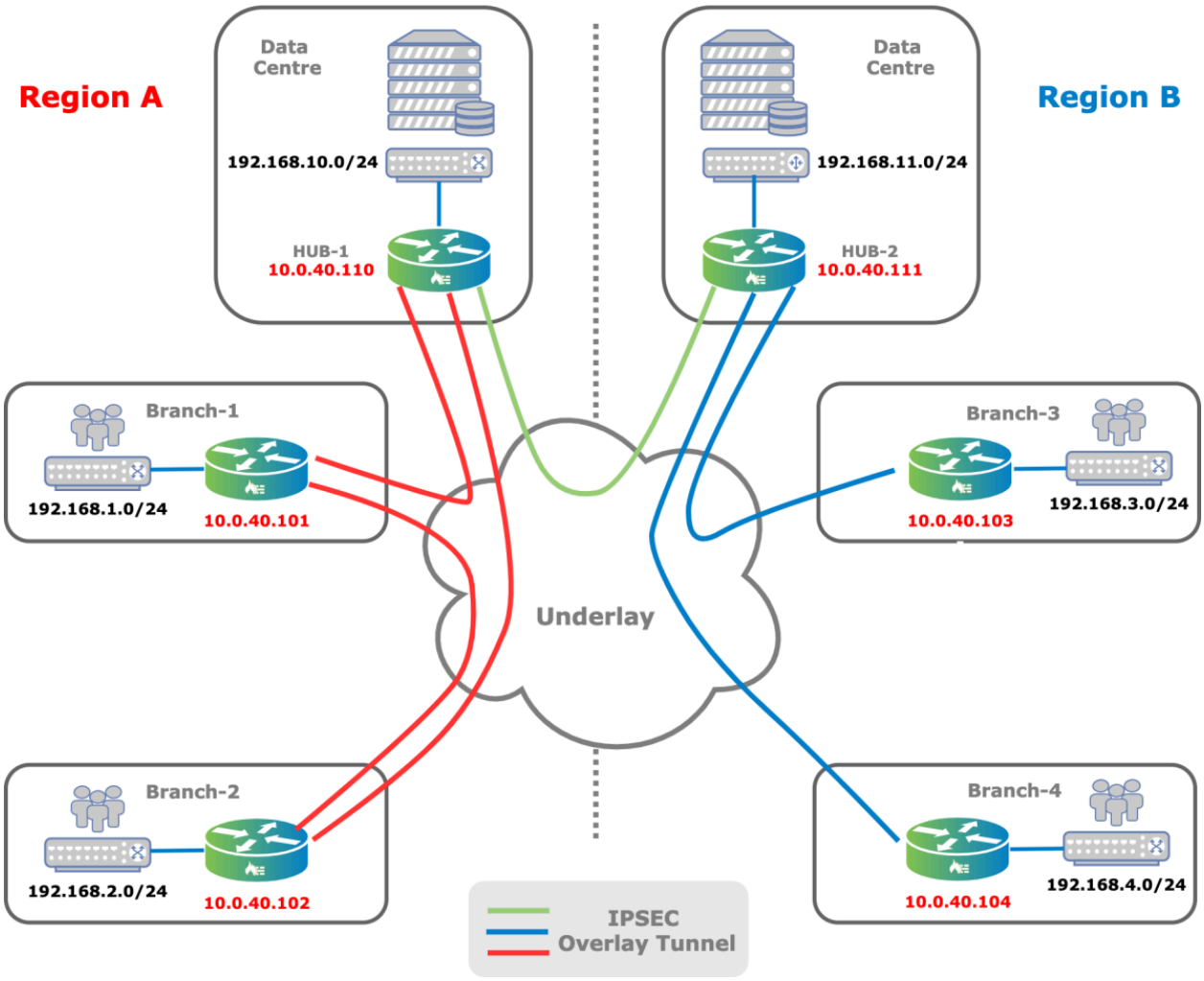
Spokes routes are directly reachable from hubs, and the backup is to use the remote hubs. For example, in the output above, the prefixes learned from Branch-1 (the first two entries in the roue table) show that the direct route is active.

Spoke-Hub-Hub-Spoke (Regional-Mesh) Topology

In a spoke-hub-hub-spoke (SHHS) topology, also called a regional-mesh topology, you group hub and branch devices

by region. Within a particular region, the topology can be anything—full mesh, partial mesh, or hub and spoke—and branches communicate based on the selected topology. When a branch in one region wants to communicate with a branch in another region, the communication transits through regional hubs.

You can select this topology when there is geographical separation and when regional WAN transport networks are available and hubs between regions use the company backbone or high-bandwidth WAN links. The following figures shows two regional networks with different topologies: Region A uses a spoke-to-spoke–direct topology, and Region B uses a spoke-to-spoke topology through a hub. In this example, communication between Branch-1 and Branch-3 uses Hub-1 and Hub-2, but the branches can use any regional hubs in the local region. This topology demonstrates how you can use the SHHS topology in regional networks.



The following CLI output shows the SD-WAN topology view from Branch-1. Because Region A uses a full-mesh topology, the output shows only regional hubs and branches.

```
admin@Branch-1-cli> show orgs org Tenant1 sd-wan brief
          SITE  MANAGEMENT                    CONNECTIVITY  IS
SITE NAME    ID   IP       TYPE    UP TIME       STATUS      CTRLR
```

```
--------------------------------------------------------------------------------
Branch-1      101   10.0.40.101 local   12d:19h:48m:37s  -            no
Branch-2      102   10.0.40.102 remote  22m:34s          Connected    no
Controller-1  1     10.0.40.1   remote  6d:21h:54m:14s   Connected    yes
Hub-1         110   10.0.40.110 remote  12d:19h:47m:28s  Connected    no
```

The following CLI output shows the SD-WAN topology view from Branch-3. Branch-3 is the only hub branch, which is normal in a spoke-to-spoke through a hub topology:

```
admin@Branch-3-cli> show orgs org Tenant1 sd-wan brief
                SITE        MANAGEMENT                      CONNECTIVITY     IS
SITE NAME       ID          IP          TYPE     UP TIME        STATUS        CTRLR
--------------------------------------------------------------------------------
Branch-3        103         10.0.40.103 local    12d:22h:41m:34s -              no
Controller-1    1           10.0.40.1   remote   7d:21h:59m:35s  Connected      yes
Hub-2           111         10.0.40.111 remote   3h:14m:8s       Connected
```

The following CLI output shows the entries in the Branch-1 VRF route table. The two prefixes in both regions are highlighted. The prefix 192.168.2.0/24, for Branch-2, is a direct route towards Branch-2 through Hub-1 for redundancy, and it is the less preferred path. The preferred route is the direct route that uses Branch-2 as the next hop and that is the active route, as indicated by the plus sign (+). The prefix 192.168.3.0/24 is for Branch-3, and the route table entries points to Hub-1.

```
admin@Branch-1-cli> show route routing-instance Tenant1-LAN-VR

Routes for Routing instance : Tenant1-LAN-VR AFI: ipv4 SAFI: unicast
Codes: E1 - OSPF external type 1, E2 - OSPF external type 2
IA - inter area, iA - intra area,
L1 - IS-IS level-1, L2 - IS-IS level-2
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
RTI - Learnt from another routing-instance
+ - Active Route

Prot    Type    Dest Address/Mask     Next-hop     Age     Interface name
----    ----    -----------------     --------     ---     --------------
conn    N/A     +192.168.1.0/24       0.0.0.0      1w5d22h     vni-0/2.0
local   N/A     +192.168.1.1/32       0.0.0.0      1w5d22h     directly connected
BGP     N/A     +192.168.2.0/24       10.0.40.102 03:15:57     Indirect
BGP     N/A      192.168.2.0/24       10.0.40.110 03:08:36     Indirect
BGP     N/A     +192.168.3.0/24       10.0.40.110 00:01:06     Indirect
BGP     N/A     +192.168.4.0/24       10.0.40.110 00:01:13     Indirect
BGP     N/A     +192.168.10.0/24      10.0.40.110 03:08:35     Indirect
BGP     N/A     +192.168.11.0/24      10.0.40.110 03:08:36     Indirect
```

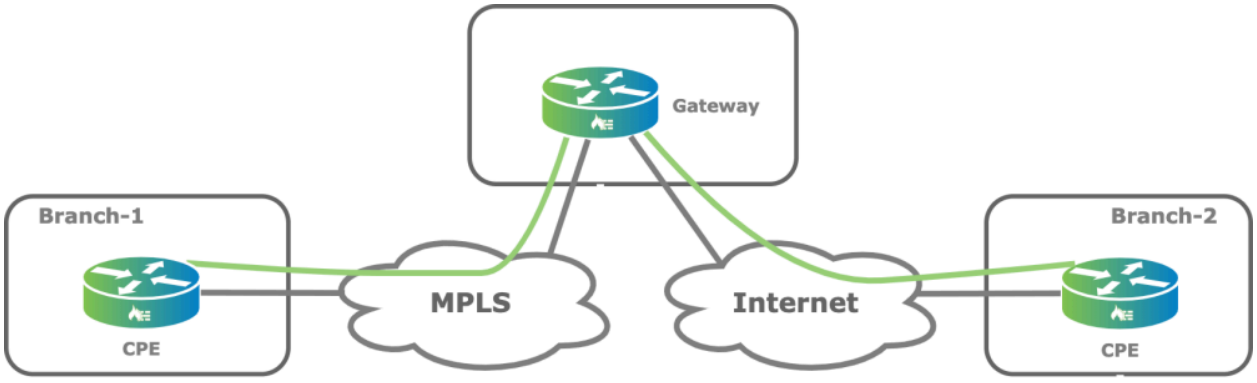
The following CLI output shows the entries in the Branch-3 VRF route table. The highlighted output shows two prefixes in both regions and that there is no difference in the treatment between regions. Region B uses a spoke-to-spoke through a hub topology. The prefix 192.168.1.0/24 is behind Branch-1, and the prefix 192.168.4.0/24 prefix is behind Branch-4.

```
admin@Branch-3-cli> show route routing-instance Tenant1-LAN-VR
```

```
Routes for Routing instance : Tenant1-LAN-VR AFI: ipv4 SAFI: unicast

Codes: E1 - OSPF external type 1, E2 - OSPF external type 2
IA - inter area, iA - intra area,
L1 - IS-IS level-1, L2 - IS-IS level-2
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
RTI - Learnt from another routing-instance
+ - Active Route

Prot   Type  Dest Address/Mask   Next-hop        Age     Interface name
----   ----  ----------------    --------        ---     --------------
BGP    N/A  +192.168.1.0/24      10.0.40.111     02:46:42 Indirect
BGP    N/A  +192.168.2.0/24      10.0.40.111     02:46:42 Indirect
conn   N/A  +192.168.3.0/24      0.0.0.0         1w5d22h  vni-0/2.0
local  N/A  +192.168.3.1/32      0.0.0.0         1w5d22h  directly connected
BGP    N/A  +192.168.4.0/24      10.0.40.111     00:04:31 Indirect
BGP    N/A  +192.168.10.0/24     10.0.40.111     02:46:43 Indirect
BGP    N/A  +192.168.11.0/24     10.0.40.111     02:46:43 Indirect
```

Connecting Sites over Disjointed Underlay Networks

You can use gateways to connect to sites over disjointed underlay networks. In a disjointed underlay, two sites do not have a common underlay that allows them to communicate directly. An example is one site that has only Internet connectivity and a seconnd site that has only MPLS connectivity. Another example is one in which a site provisioned with internet and MPLS links communicates with a site that has only an MPLS link. When the internet link on the first site becomes unavailable, then without a gateway, that site loses connectivity to the MPLS-only site. Another common scenario for using a disjointed underlay network is when NAT traversal issues do not allow two internet-connected branches to connect directly. A third device, which is the gateway, can interconnect the two branches, as shown in the following figure.



You can connect the sites by configuring one of the following in on the gateway device:

- Configure the branch as the gateway and configure the branches in the spoke-to-spoke-direct topology in the

Workflow templates

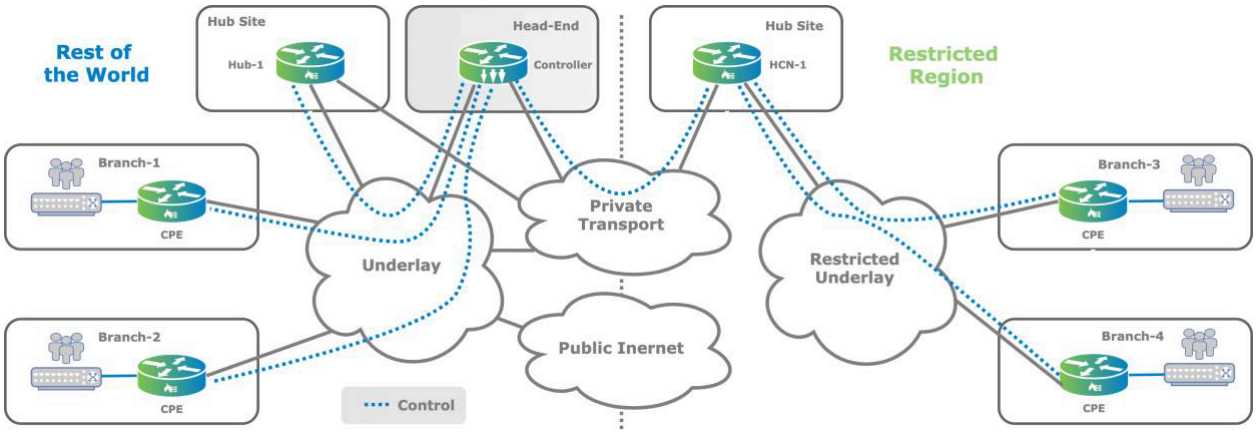
- Advertise a summary route from the gateway. The gateway can be any branch site, and you do not need to configure the branch as a gateway in the Workflow templates.
- Advertise a default route from the gateway. The gateway can be any branch site, and you do not need to configure the branch as a gateway in the Workflow templates.

SD-WAN Topologies for Geographically Isolated Regions

In some regions of the world, government place restriction on encrypted (IPsec) traffic to block it from going in and out of the country. Inside the country, traffic is typically not affected, but blocking the IPsec traffic can isolate the SD-WAN regional network. For these situations, you can create separated SD-WAN islands in which each SD-WAN domain operates within itself and has no visibility into or information about other domains. To interconnect these SD-WAN islands, you use IPsec or other connection options, which you must provision and manage outside the SD-WAN domain. You must consider the stitching complexities on the boundary node, which acts like a network-to-network interface (NNI) point.

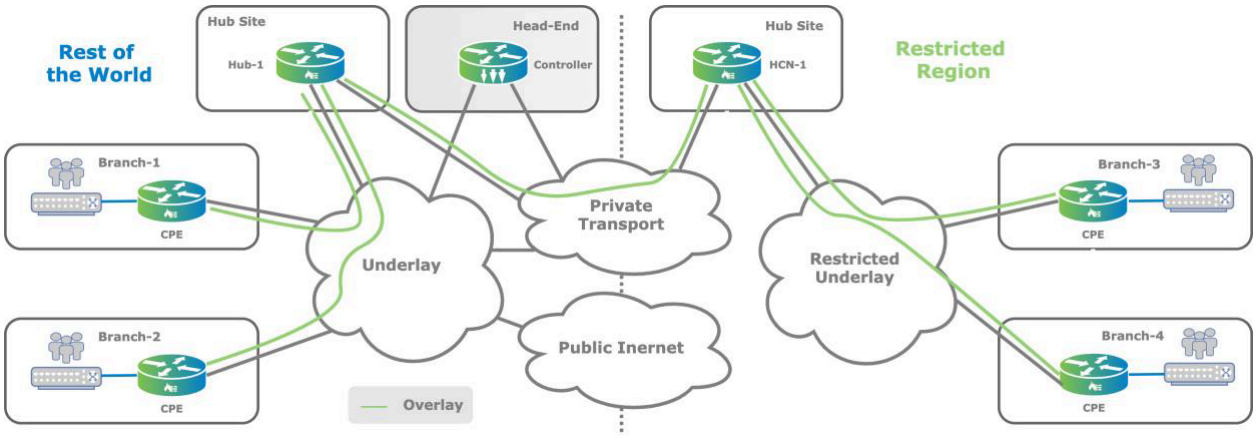
The separation prevents the control planes of each SD-WAN island from exchanging information with each other, and there is no connectivity between domains. Creating tenants and service templates across SD-WAN islands become a complex task. The SHHS topology provides a solution that is highly scalable, compartmentalized, fully automated, and easy to deploy and operate.

The following figure shows a control plane connection for an isolated region. An IBGP session is established between the headend Controller node and the hub Controller node (HCN-1) site in the restricted region and is used to exchange SD-WAN route information over a suitable transport, in this case a private transport, that allows IPsec traffic. The data plane flow between a branch in the restricted region and a branch with no restriction traverses the hub sites using IPsec tunnels. Note that the headend Controller node is not part of the data plane communication.



You can apply the SD-WAN network topology shown in the following figure to restricted regions. Branches in this region create an SD-WAN island in the form of a restricted underlay that has no connectivity with the rest of the world. An approved transport must be provided by a government-approved service provider to allow IPsec traffic, which can be a

Layer 2 service (such as Layer 2 VPN and VPLS) or a Layer 3 service. This service is represented by the private transport.



For such a topology, the VOS software provides a hierarchical controller structure called a hub-controller node (HCN) device type. An HCN can interconnect the control planes of SD-WAN islands that may be using their own Controller instances. The MP-BGP–based control plane assigns separate community values to represent each SD-WAN island, and IBGP sessions are established between the regional HCNs and the Versa headend Controller node, unifying the control plane. HCNs can consolidate Controller and hub functionalities into one node to preserve resources on the control side of the network. HCN nodes exchange information with spokes and implement the data plane functions of hub nodes.

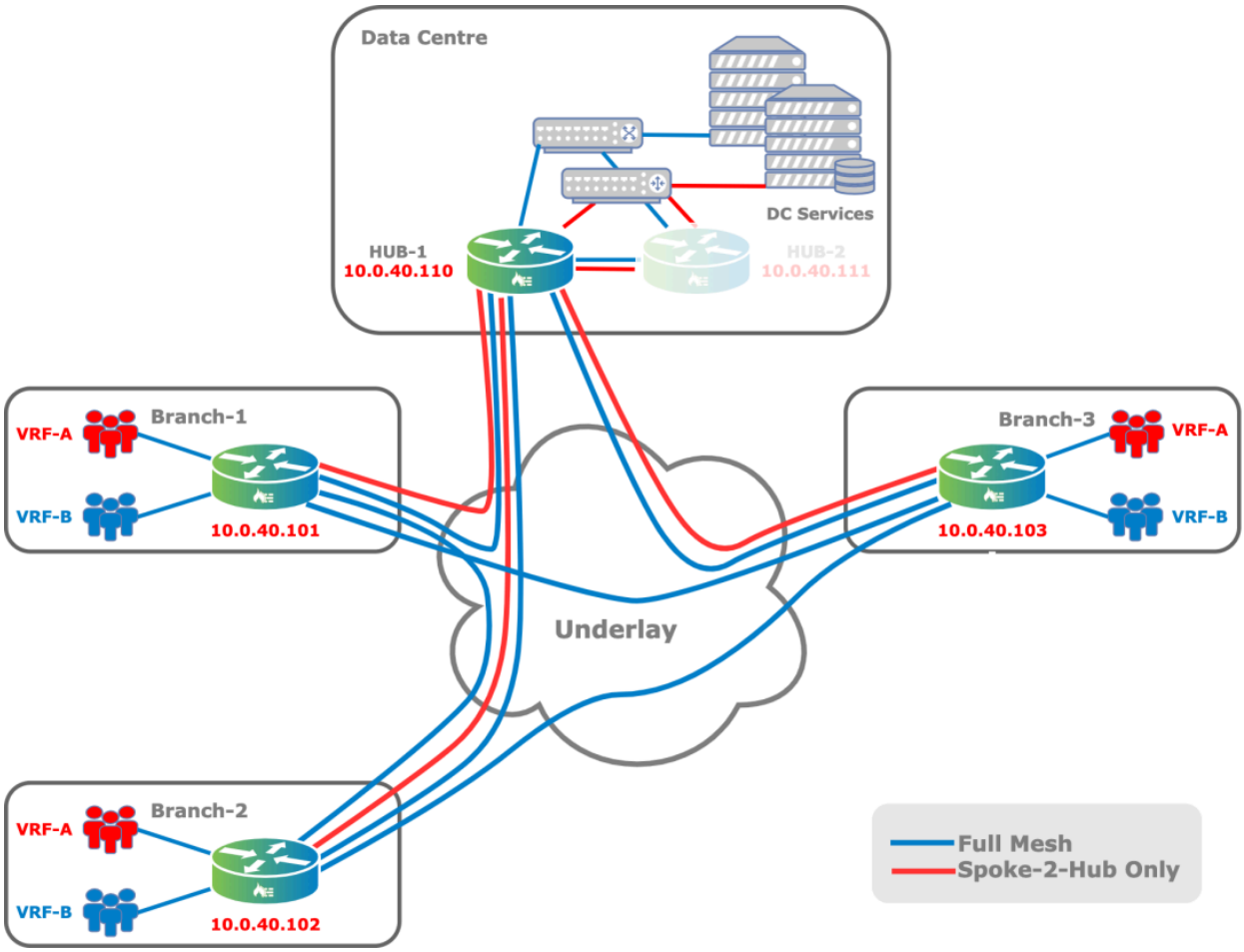
You can deploy HCNs in active–active mode for the control plane to maximize uptime and ease of serviceability. In active–active mode, multitenancy is preserved across the topology using the organization and sub-organization structure. Data plane redundancy is provided by BGP next-hop route resolution. You then provision tenants and services, starting from the spokes of one SD-WAN island to the other SD-WAN island, using the expanded templates for SHHS deployment. After the control plane is fully functional, SD-WAN paths are automatically established between Branch-1 and Hub-1, Hub-1 and HCN-1, and HCN-1 and Branch-3 in a hop-by-hop manner, in both directions, as shown in the figure above. These paths can be separate SD-WAN tunnels or shared SD-WAN tunnels that may already be present between spokes and hub routers. All data plane operations are handled automatically and require no manual configuration. After the end-to-end data paths are set up, users can communicate seamlessly between the spokes of separate SD-WAN islands.

Multi-VRF (Multitenant) Topologies

The examples discussed thus far in this article use a single VRF (a single tenant) to explain the differences between topologies. The Versa Networks SD-WAN solution is highly flexible and allows you to define and establish different topologies at the VRF, or tenant, level. The default Versa SD-WAN model provisions a full forwarding mesh using IP prefix advertisement in MP-BGP. You can configure the VPN topologies discussed in previous sections using Director Workflows.

The following figure shows an SD-WAN topology with two VRFs in one organization. VRF-A (red) is configured for

spoke-to-hub only, and VRF-B (blue) uses a full-mesh topology. This principle can apply to multitenant branches in which each organization's topology may be different.



Best Practices for SD-WAN Topologies

The following are a few best practices for various SD-WAN topologies:

- The full-mesh topology is the default option in Director Workflows, and you can deploy a full mesh without restrictions for VPNs with up to 100 sites. If there are more sites, SLA monitoring consumes a significant amount bandwidth. This means that branches with low-bandwidth connections must assign a relatively high proportion of their available bandwidth for SLA monitoring traffic. For low-bandwidth branches, it is recommended that you deploy a spoke-to-spoke through a hub topology so that SLA monitoring is performed only towards the hubs.
- Spoke-to-spoke-direct topologies are recommended over full-mesh topologies. For many use cases, having a hub node is preferable in the topology, for example, to avoid disjointed underlays caused by circuit failures or NAT traversal issues. Also, spoke-to-spoke–direct topology provides better support for automatically importing routes from LAN adjacent networks, which avoids manual configuration of redistribution policies.
- Use a distinct WAN network name on the hub sites.

Supported Software Information

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

Additional Information

Configure Regional Hub-and-Controller Nodes for SHHS Topologies
Configure SLA Monitoring for SD-WAN Traffic Steering

VOS Edge Device Direct Internet Access

For supported software information, click here.

In a typical legacy enterprise VPN topology, internet traffic from the branch is directed to a hub site over a private MPLS link, and from the hub, it is sent do, or breaks out to, the internet. The internet-bound traffic typically passes through a web proxy that is placed in the hub, where it may be inspected by a firewall. With the advent of high-speed and more reliable internet circuits, it is more economical for enterprises to break traffic out to the internet directly at the branch office. Also, as enterprises move some of their enterprise workloads to the public cloud or use SaaS-based services such as Office 365, a better user experience is important. The breaking out of traffic to the internet at the branch is called direct internet access, or DIA.

The following are the primary advantages of DIA:

- Reduces the bandwidth requirements at headquarters
- Fewer network hops
- Reduces the latency and offers better optimization for internet applications because of direct routing

The increased reliability of the internet for WAN transport makes DIA desirable in branch deployments. However, sending traffic directly from the branch to the internet creates additional security challenges and the branch requires extra protection.

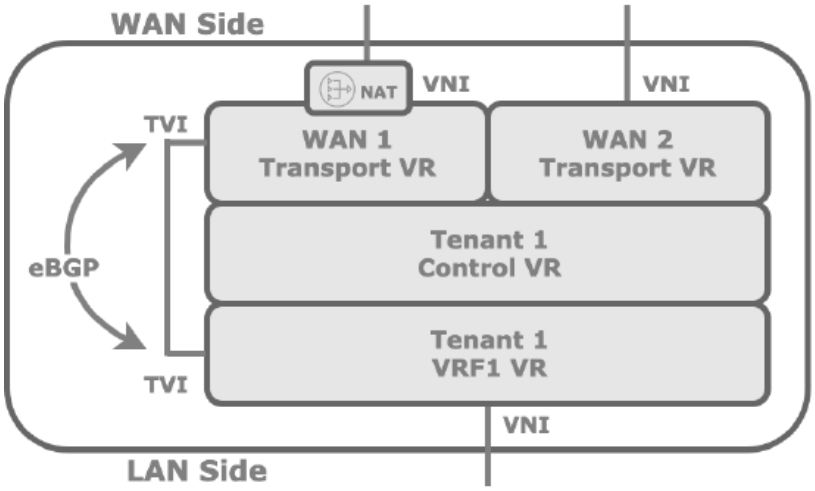
The DIA architecture discussed here is based on standard routing principles. Internally, a logical connection is created between a tenant VRF and the WAN transport VR. The tenant VRF uses EBGP to advertise the default route, and it uses the logical tunnel to resolve the next hop. A NAPT rule translates all LAN traffic into one public IP address space.

VOS Edge Device DIA Architecture

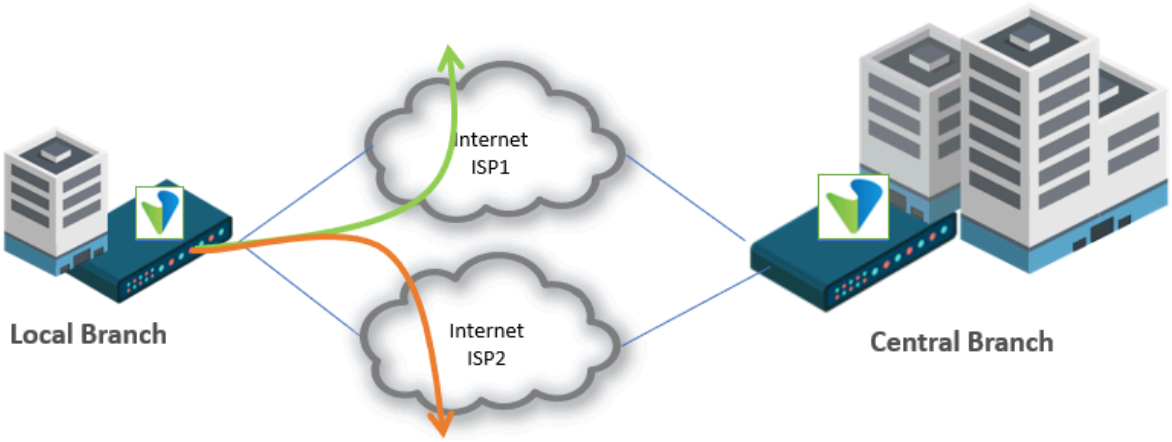
When you configure internet breakout on a branch, the configuration of the VOS edge device has the following main components, which are illustrated in the following figure:

- Internal connection between the VRF and the transport WAN VR, which is supported by an EBGP connection, to control route distribution

- CGNAT function, to translate the IP addresses of all internet-bound traffic to the public IP address that is typically attached to the the transport WAN VR interface



When multiple WANs are available, you can achieve DIA redundancy and DIA load balancing, as shown in the following figure. To do this, you set the BGP local preference in the Director Workflow to select the preferred WAN interfaces. If you select the load-balancing option is selected, the same local preference value is used for both that default routes that are advertised over EBGP towards the VRF.



You configure DIA using Director Workflows, when you configure tunnels. For more information, see Configure Direct Breakout to the Internet.

Central or Regional Internet Breakout

In a typical deployment, it is common to provide redundancy for the default route. For the two default routes, one points to a local DIA and the backup default route, with a less preferred metric, points to the central DIA. If the local branch internet circuit malfunctions, the locally sourced default route is withdrawn and the centrally sourced default route is activated.

You use the Director Workflow to create a local internet breakout and establish the central, regional, or remote backup in the Create/Edit Template window > Tunnels tab, as shown in the following screenshot.

To configure this use case, ensure that you do not click the Gateway field for the local breakout edge device if you click the Gateway field for the central internet breakout edge device. When you click the Gateway field, the edge device announces the default route to the rest of the SD-WAN. If you do not click the Gateway field, the default route sourced by the local DIA infrastructure is visible only on the local edge device and is not visible anywhere else in the SD-WAN VPN. For more information, see Configure Device Templates and Configure Site-to-Site Tunnels.

To view the route table on the local device, issue the **show route routing-instance** CLI command. For example:

```
admin@Site3-cli> show route routing-instance
Prot   Type  Dest Address/Mask   Next-hop       Age      Interface name
----   ----  ----------------   --------       ---      --------------
BGP    N/A   0.0.0.0/0          10.2.64.101    00:31:23 Indirect
BGP    N/A  +0.0.0.0/0          169.254.0.2    00:00:18 tvi-0/603.0
conn   N/A  +169.254.0.2/31      0.0.0.0        00:00:20 tvi-0/603.0
local  N/A  +169.254.0.3/32      0.0.0.0        00:00:20 directly connected
BGP    N/A   172.1.118.0/24      10.2.64.101    00:31:23 Indirect
BGP    N/A  +172.1.118.0/24      10.2.64.102    00:31:23 Indirect
```

In the route table output above, the preferred route is the local breakout route, the route whose next hop is 169.254.0.2. You can check the BGP route table to verify that this route has a better (lower) distance, favoring EBGP over IBGP. For example:

```
admin@Site3-cli> show route routing-instance Enterprise1-LAN-VR 0.0.0.0/0

Routes for Routing instance : Enterprise1-LAN-VR  AFI: ipv4  SAFI: unicast
[+] - Active Route

Routing entry for 0.0.0.0 (mask 0.0.0.0)
Known via 'BGP', distance 200,
  Redistributing via BGP
  Last update from 10.2.64.101 00:34:20 ago
Routing Descriptor Blocks:
* 10.2.64.101 , via Indirect 00:34:20 ago

Routing entry for 0.0.0.0 (mask 0.0.0.0) [+]
Known via 'BGP', distance 20,
  Redistributing via BGP
  Last update from 169.254.0.2 00:03:15 ago
Routing Descriptor Blocks:
* 169.254.0.2 , via Indirect 00:03:15 ago
```

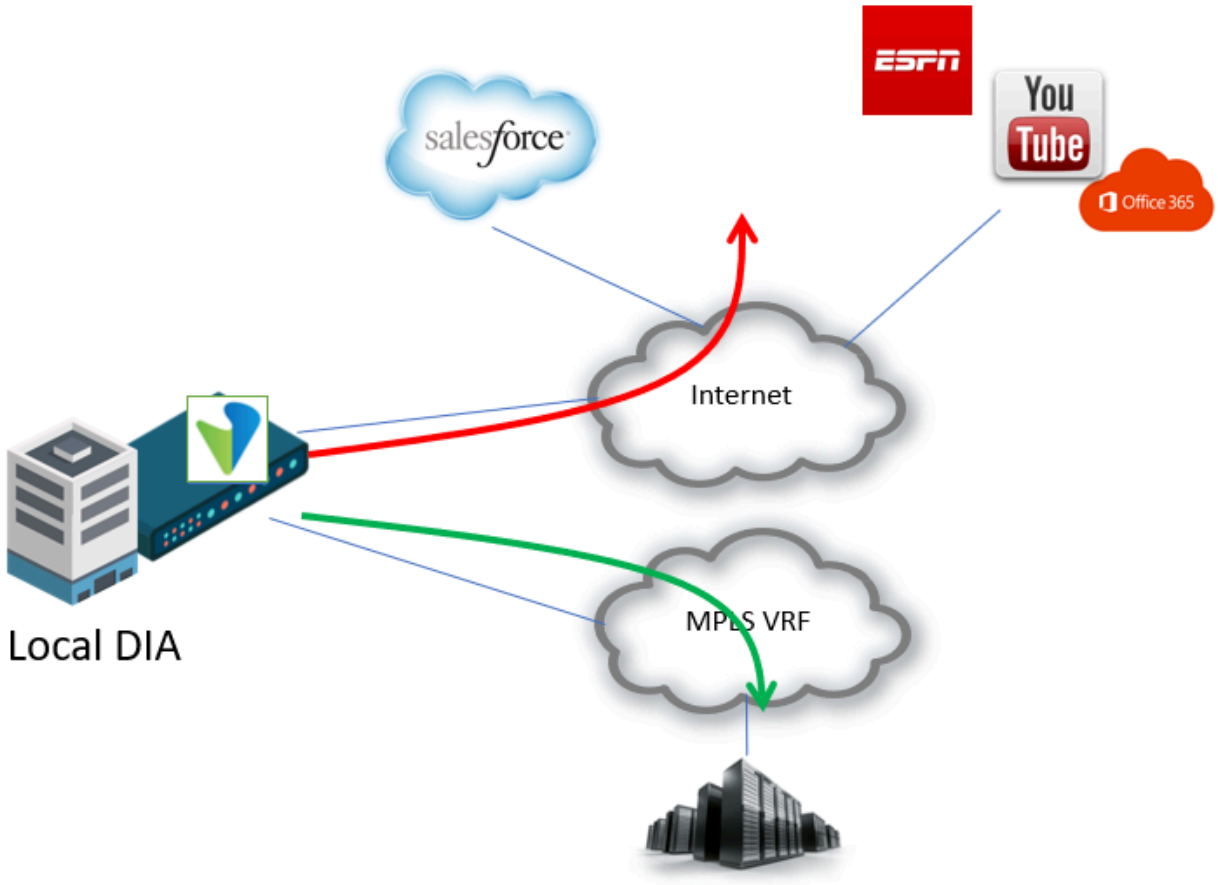
You can alter this behavior by manipulating the BGP preference in the configuration.

With this infrastructure, you can create advanced configurations by using route coloring with communities, where different regions prefer different regional default routes.

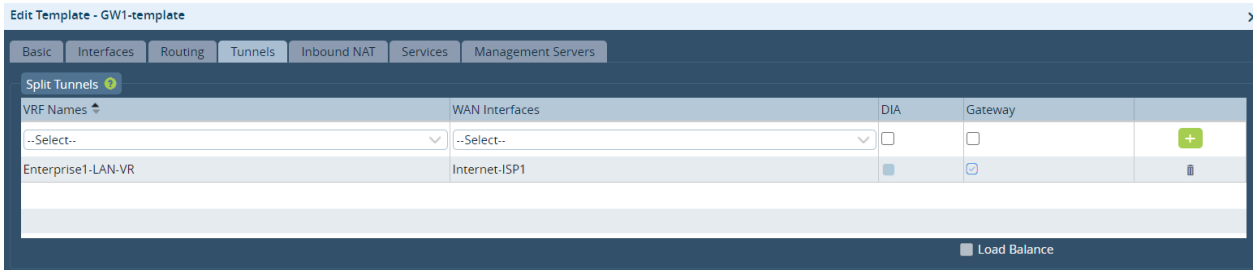
Breakout to an MPLS Underlay for Common Services

Another use case for DIA is a breakout service to the MPLS underlay. You can use this type of breakout to reach services offered by an MPLS provider from the underlay, such as VoIP services. You can also use this solution to provide gateway services during migration from a legacy MPLS network to an SD-WAN network.

You achieve the breakout to MPLS the same way as a local internet breakout service. In the Director Workflow, you create a split tunnel to the MPLS underlay, as illustrated below.

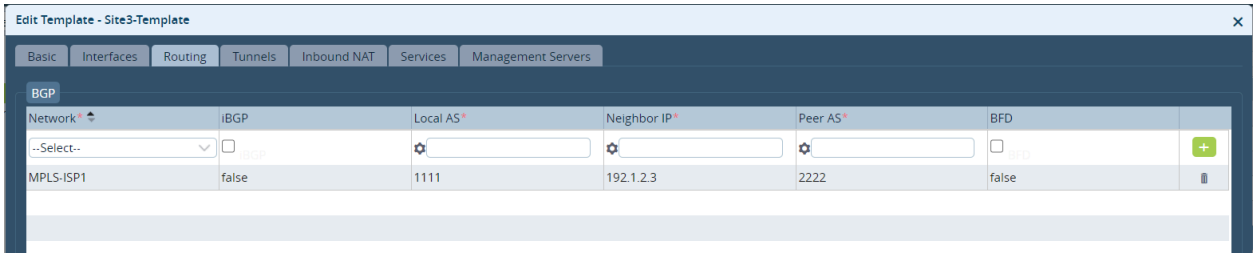


One important difference from an Internet breakout split tunnel is that you must not click the DIA field on the Tunnels tab in the Create or Edit Template window. Clicking this fields creates an NAPT instance, which is not required in a private MPLS underlay. If you prefer to advertise the routes imported from MPLS to the rest of the SD-WAN VPN, click the Gateway field on the Tunnels tab, as shown here:



The infrastructure shown in in the screenshot above also creates an internal connection between the LAN-VR and the MPLS-VR, and it runs EBGP over the connection. In this way, the routes in the MPLS-VR are propagated to the LAN-VR.

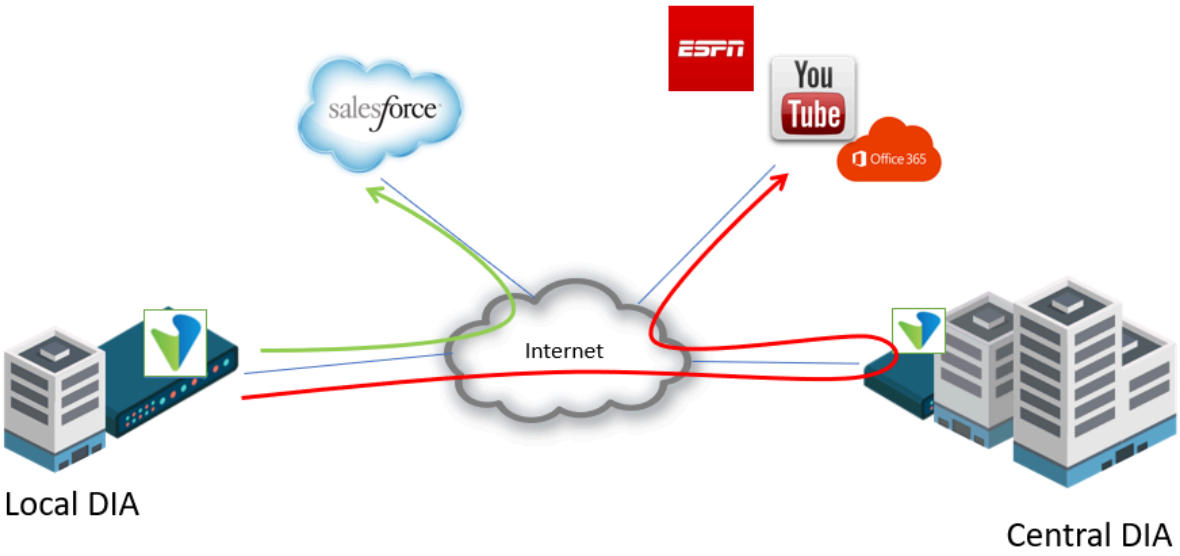
Next, you must import the MPLS VRF routes into the SD-WAN edge device. To do this requires a dynamic routing connection between the SD-WAN edge device and the MPLS provider infrastructure, which is typically the PE router. The easiest way is to configure the BGP connection on the Routing tab of the Create/Edit template:



Note that the remote PE router, which is operated by the managed service provider (MSP), must also be configured with similar dynamic routing.

Application-Based Breakout

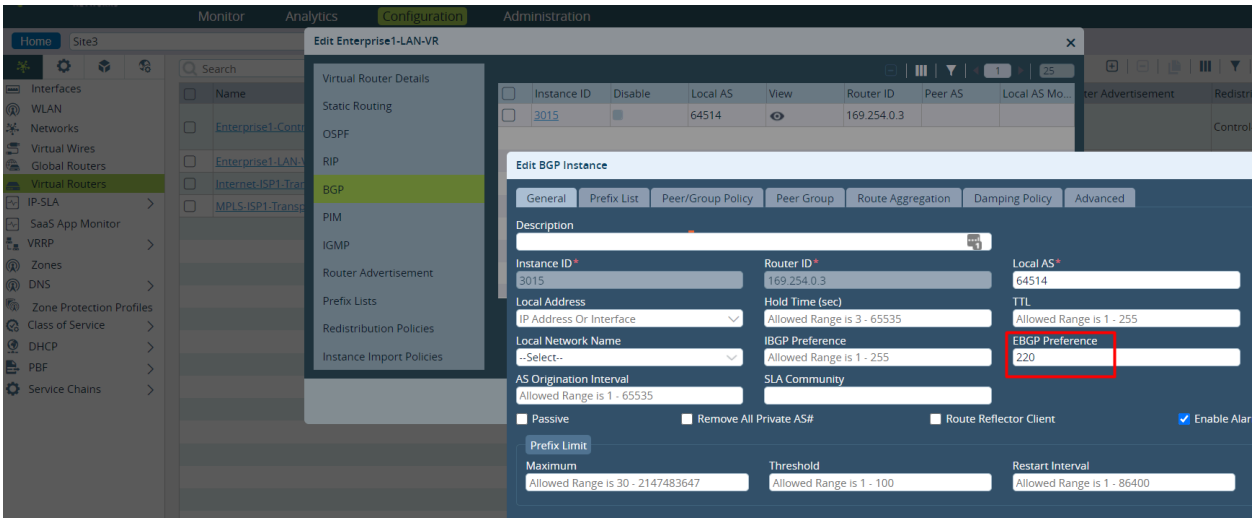
A common DIA use case is to provide a local breakout service for specific applications, in which the default route points a regional or central breakout point. This configuration is a little more involved and relies on the infrastructure shown in Option 1: BGP to Exchange Routes with MPLS Provider on MPLS WAN Interface and Option 2: BGP with MPLS Provider on LAN Interface. For these usage case, the SD-WAN edge device has a local breakout infrastructure and also has an option to break out centrally. The following figure illustrates this use cae.



As described in Breakout to MPLS Underlay for Common Services above, this design prefer sa local breakout over a central breakout. For the default internet-bound traffic to break out centrally, you must alter the preference to one that is higher than 200, because, by default, the EBGP route administrative distance (20) is better than that for the IBGP (200).

To set the EBGP preference to 220 for the LAN-VR, to make the router to central default router be the preferred route:

1. In Director view:
   a. Select the Configuration tab in the top menu bar.
   b. Select Templates > Device Templates in the horizontal menu bar.
   c. Select an organization in the left menu bar.
   d. Select a post-staging template in the main pane. The view changes to Appliance view./li>
2. Select the Configuration tab in the top menu bar.
3. Select Networking > Virtual Routers in the left menu bar.
4. Click the Add icon. The Configure Virtual Router popup window displays.
5. Select the BGP tab in the left menu bar. The main pane displays a list of the BGP instances that are already configured.
6. Click the Add icon. The Add BGP Instance popup window displays.
7. Select the General tab.
8. In the EBGP Preference field, enter 220.



9. For information about configuring the other fields, see Configure BGP.
10. Click OK.

Now, all internet traffic breaks out at the central breakout point.

To verify that the central default route is preferred:

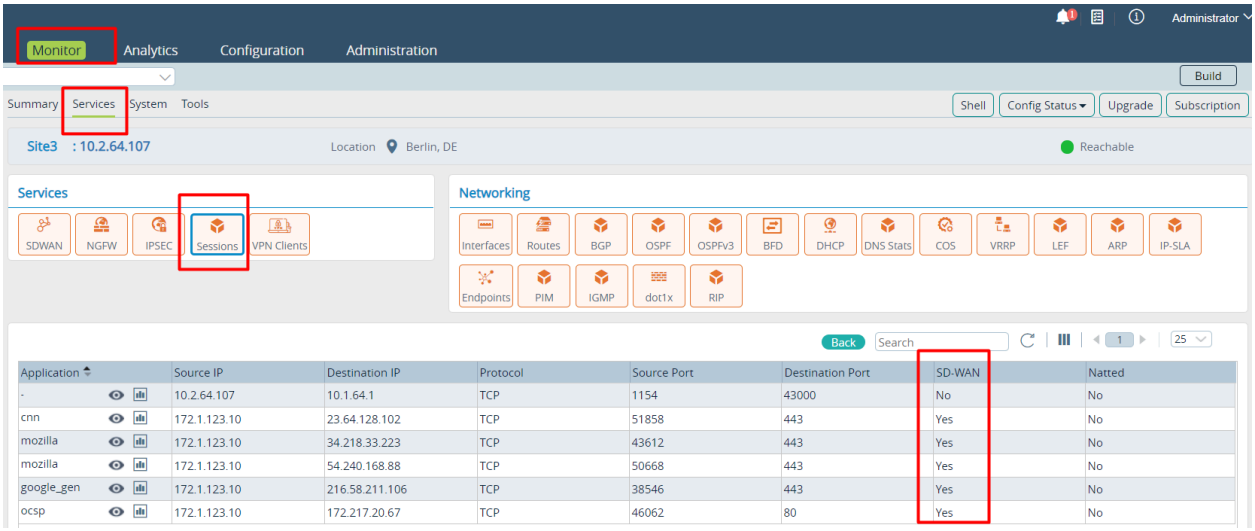
```
admin@Site3-cli> show route routing-instance Enterprise1-LAN-VR 0.0.0.0/0

Routes for Routing instance : Enterprise1-LAN-VR  AFI: ipv4  SAFI: unicast
[+] - Active Route

Routing entry for 0.0.0.0 (mask 0.0.0.0) [+]
Known via 'BGP', distance 200,
  Redistributing via BGP
  Last update from 10.2.64.101 01:16:57 ago
Routing Descriptor Blocks:
* 10.2.64.101 , via Indirect 01:16:57 ago

Routing entry for 0.0.0.0 (mask 0.0.0.0)
Known via 'BGP', distance 220,
  Redistributing via BGP
  Last update from 169.254.0.2 00:00:06 ago
Routing Descriptor Blocks:
* 169.254.0.2 , via Indirect 00:00:06 ago
```

Now, all internet traffic breaks out at the central breakout point.

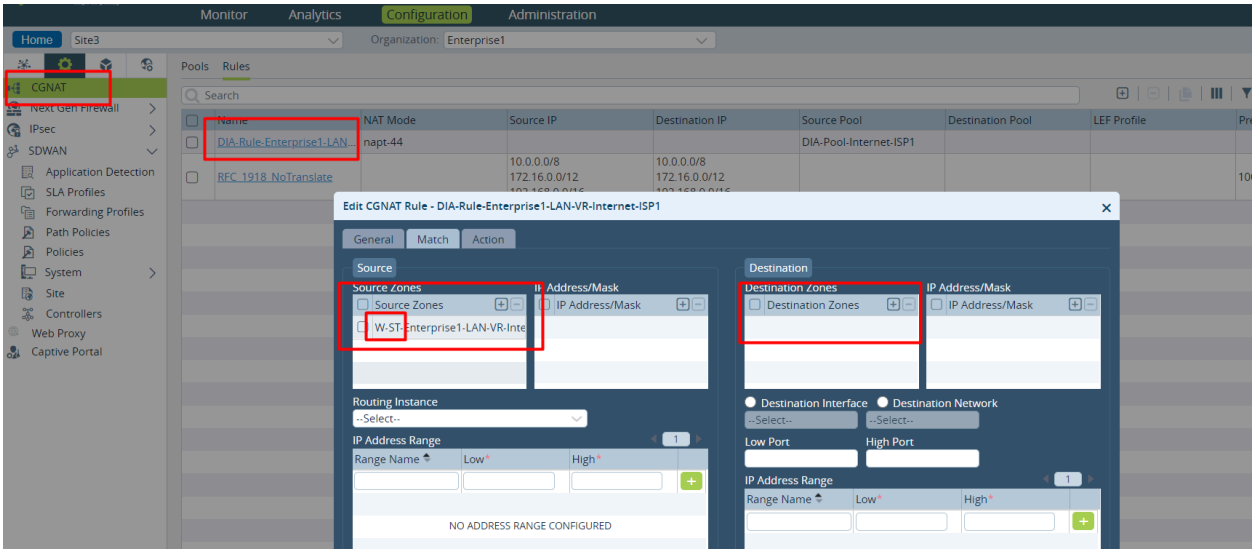
You can also verify that the central default route is preferred by checking whether some internet sessions on the local branch are SD-WAN sessions:

1. In the Appliance view, select the Monitor tab in the top menu bar.
2. Select a provider organization.

3. Select the Services tab.
4. Select Sessions. Check whether the SD-WAN column displays Yes or No.

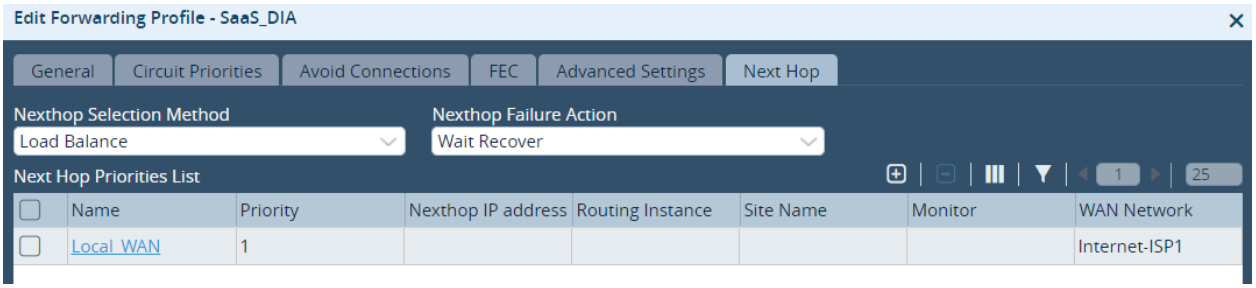


For this use case, you must also modify the CGNAT rule. The change is simple, but important, you must modify the match condition of the rule to match the source zone, not the destination zone. To do this, you remove the destination zone and add a source zone in the Edit CGNAT Rule popup window. In the example here, you select the zone W-ST-*tenant-name*-LAN-VR-*internet-WAN*-VR. For more information, see Configure CGNAT Rules.



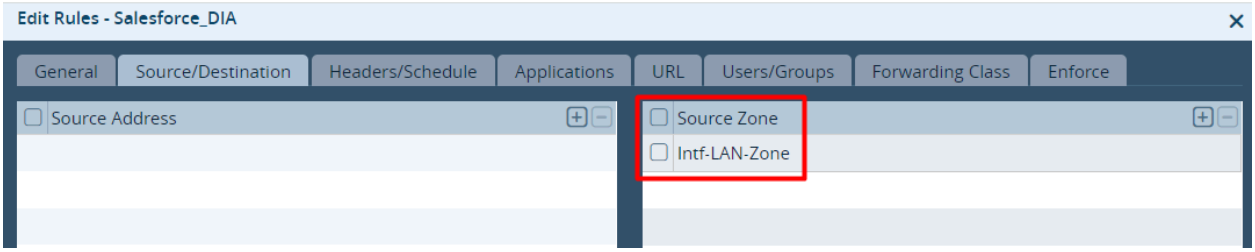
To force a specific application to break out locally and to have the remaining internet-bound traffic break out centrally, you create an SD-WAN policy. This example locally breaks out the traffic to Salesforce.com:

1. Create a forwarding profile, here called SaaS_DIA. On the Next Hop tab, select the local DIA as the next hop. For more information, see Configure SD-WAN Traffic-Steering.

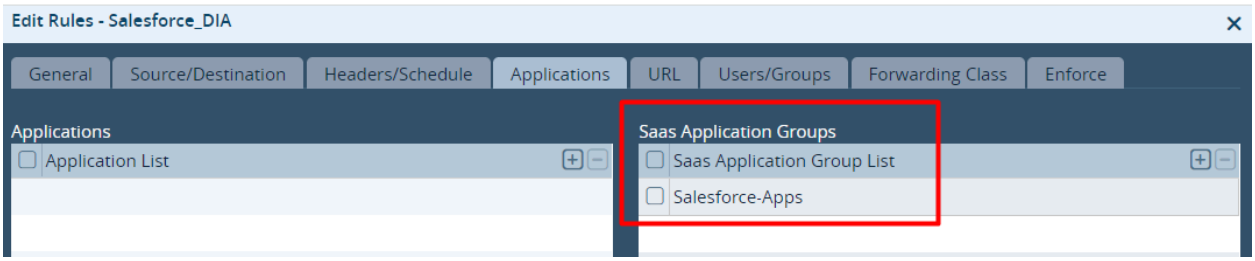


2. Create a forwarding profile rule to associate with the forwarding profile. For more information, see Configure SD-WAN Traffic Steering Policy.

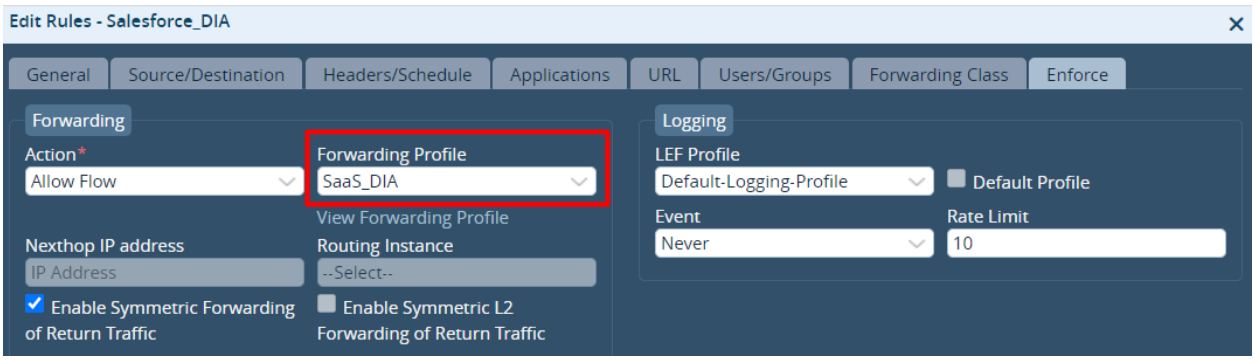
   a. Select the Source/Destination tab, and select the source zone.



   b. Select the Applications tab, and in the SaaS Application Groups table, select Salesforce-Apps.



   c. Select the Enforce tab, and in the Forwarding Profile field, select the forwarding profile you created in Step 1.



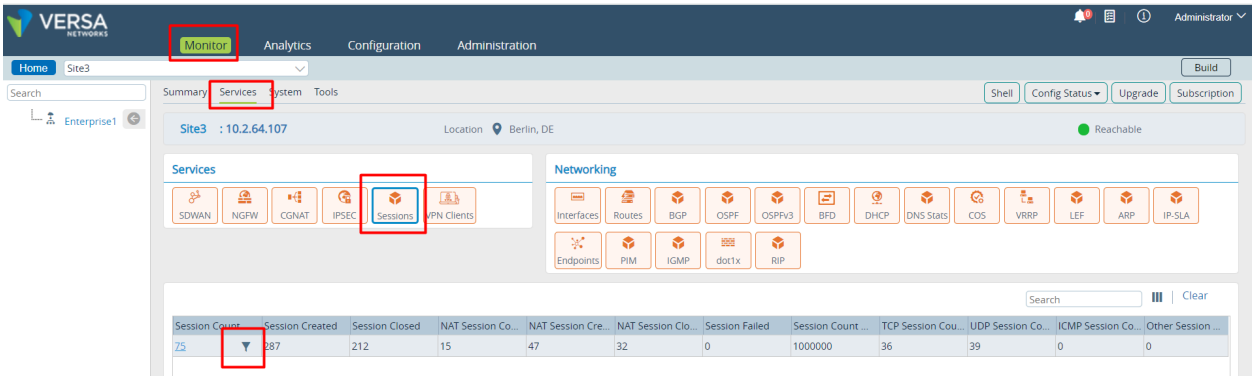
   d. Click OK.

To verify the configuration, open a browser and navigate to the website or application, Salesforce in this example.

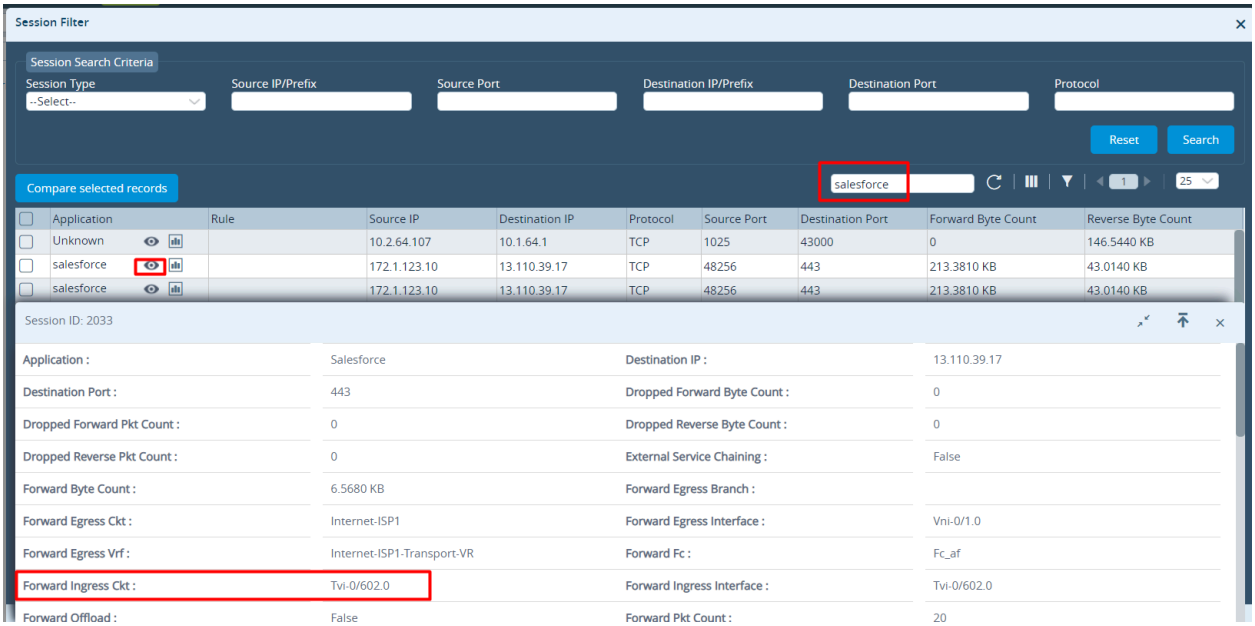
You can verify the operation in a number of ways.

To verify the session details and the egress:

1. In the Appliance view, select the Monitor tab in the top menu bar.
2. Select a provider organization.
3. Select the Services tab.
4. Select the Sessions tab, and search for salesforce.



5. Click the Eye icon to view details. If the breakout to the local DIA is functioning correctly, the Ingress and Egress Circuit fields show the interface TVI-0/602, which is the interface that has the IP address that you set as the next hop.



6. To verify that sessions for other (non-Salesforce) traffic are not pointing to this TVI, select that application and click the ; Eye icon to view details. Typically, you see that the Tx Branch is your central DIA edge device.

Best Practices

The following are best practices for using DIA:

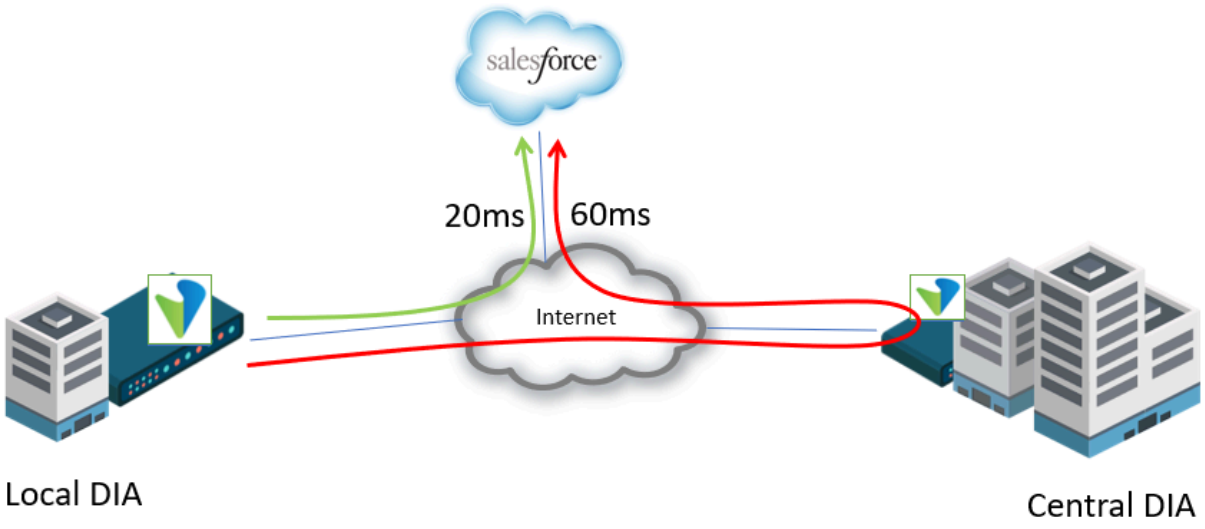
- Use the DIA gateway feature to enable central breakout to internet to ensure backup for local breakout.
- Use application-based local breakout to improve performance for one application or a group of applications.
- For application-based traffic steering, remove the default route populated by the workflow template to the central location, except for one or a group of applications.
- To test and verify application-based traffic steering, generate multiple traffic flows.
- Remove the local breakout default route—You can suppress the default route sourced by the local breakout DIA configuration, because it is not required for the configuration. If you retain it, the default route acts as a backup breakout point for the local site if the central breakout fails. If you steer traffic based on the application use case, the locally sourced default route is not required. If you do not need a local breakout except for specific applications, you must remove this default route. You can do this by filtering the default route with a prefix list or by removing the redistribution of static routes under redistribution policies.
- Deep packet inspection (DPI) application recognition—When you test the implementation, the traffic may not locally break out as expected. A common reason is that the DPI does not detect the application ID when the session is created. This is more common in lab setups than in production deployments, because production deployments have more application cache for the same application traffic. For example, the system is aware of the traffic that goes to a specific destination IP address, such as salesforce.com. To build up the application cache, create additional sessions, for example, by refreshing the browser.

Performance-Based SaaS Optimization for DIA

The Versa Networks SD-WAN solution selects the best internet breakout to reach SaaS applications by performing monitoring to measure application network performance. The following types of monitoring are available:

- Passive (inline monitoring)—VOS edge devices collect metrics for active TCP-based SaaS application sessions transiting the edge device. These metrics include the network and server response times and packet loss estimates in each direction. These metrics are used to assess application quality on a particular path and then to select the best path for the application.
- Active monitoring—Active monitoring proactively monitors SaaS locations using HTTP, ICMP, and TCP probes, exports the collected metrics to remote sites, and incorporates actively learned metrics in path selection. The metrics collected through active monitoring are latency (RTT) and packet loss.

The data collected by these measurement techniques are automatically combined with the existing SLA measurement over the overlay to select the best path, and the SD-WAN policy steers the traffic accordingly. The following figure illustrates how this works. Here, the local DIA branch measures a latency of 20 milliseconds to Salesforce.com, while from the same branch, it takes 60 milliseconds to reach Salesforce.com through the central DIA. Therefore, the logic in the local DIA branch steers the traffic to internet using the local DIA.



Basic Performance-Based Breakout Configuration

This section describes how to configure performance-based application breakout that is based on active monitoring of an SaaS application, and it build on the application-based breakout scenario. Performance-based breakout leverages the infrastructure of having local and remote DIA, and it requires that you modify the NAT rule. However, for performance-based breakout, you configure the SD-WAN policy differently.

For performance-based breakout, you create the following additional building blocks:

- SaaS application monitor on both the local and remote breakout edge devices
- SD-WAN SLA profile for the local edge device
- SD-WAN forwarding profile for the local edge device
- SD-WAN policy for the local edge device

In the design configuration example here, the Salesforce application breaks out using the breakout point that provides the lowest latency to this SaaS. The latency is measured from the point of view of the local edge device. This configuration example has two breakout options:

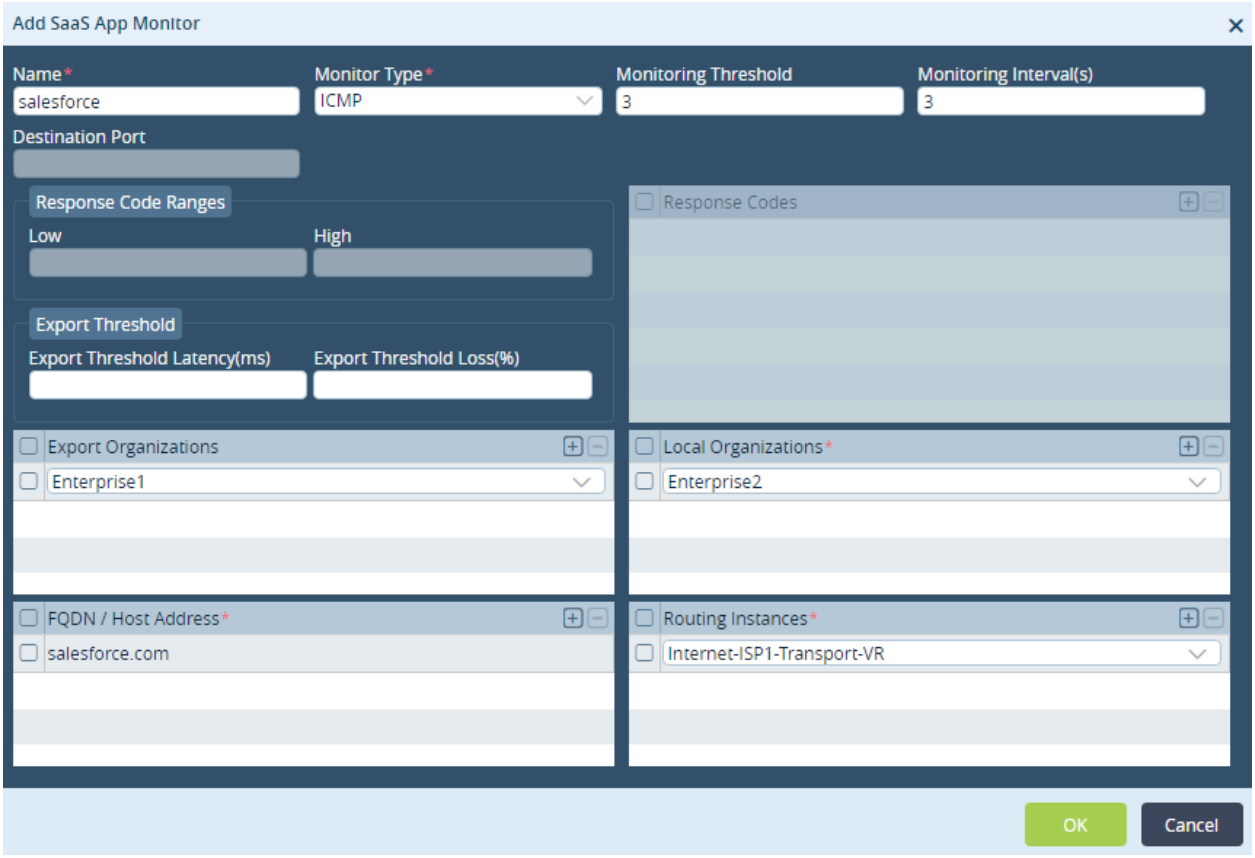
- Break out Salesforce traffic locally—The latency measurement is from the local edge device to Salesforce only, over the internet.
- Break out Salesforce traffic remotely—The latency measurement is from the local edge device to the remote edge device over the SD-WAN overlay, and then to Salesforce over the internet connection to the remote edge device. This configuration adds together the internet latency and overlay latency.

For both the local and remote edge devices, you configure a SaaS application monitor that checks the performance to a specific SaaS application FQDN. Note that the example here shows only the minimum required configuration for the SaaS application monitor, and it shows the configuration of the local edge device only. You might have different requirements or more complicated use cases. Many configuration options are available, and many alternative scenarios are possible, such as more local internet breakouts, more remote internet breakout places, and different SLA criteria.

For more information, see Configure SaaS Application Monitoring, Configure SLA Profiles for SD-WAN Traffic Steering, and Configure SD-WAN Traffic Steering.

To configure performance-based monitoring for the Salesforce SaaS application:

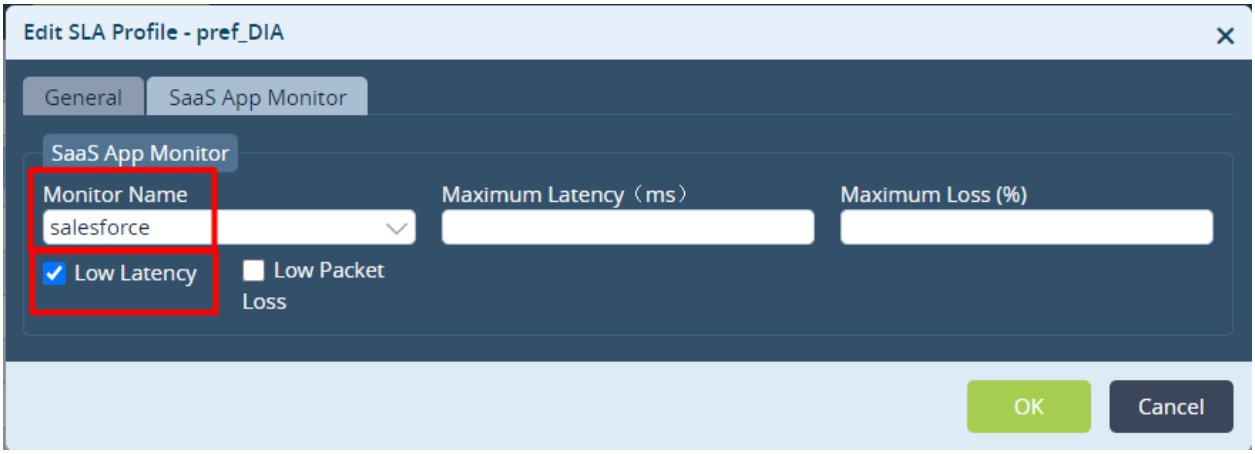
1. Add a Saas application monitor named salesforce:
   a. In Appliance view, select the Configuration tab in the top menu bar.
   b. Select Networking > SaaS App Monitor in the left menu bar.
   c. Click the Add icon to add a SaaS application monitor.



   d. In the Name field, enter "salesforce".
   e. In the FQDN/Host Address table, click the Add icon to select the FQDN salesforce.com.
   f. Click OK.
2. Configure an SLA profile that references the SaaS application monitor you configured in Step 1:
   a. In Appliance view, select the Configuration tab in the top menu bar.
   b. Select Services > SD-WAN > SLA Profiles in the left menu bar.
   c. Click the Add icon to add an SLA profile, or select an existing SLA profile, as shown here. The Create

SLA Profile or Edit SLA Profile popup window displays. (Note that the screenshot shows the incorrect SLA profile name. It should be "perf_DIA".)

d. Select the SaaS App Monitor tab, and in the Monitor Name field, select salesforce. This example uses Low Latency, to have the network to choose the lowest latency or loss. Alternatively, you can set values for the maximum latency or maximum loss, to have all paths be considered as long as the maximum value is not reached.



e. Click OK.

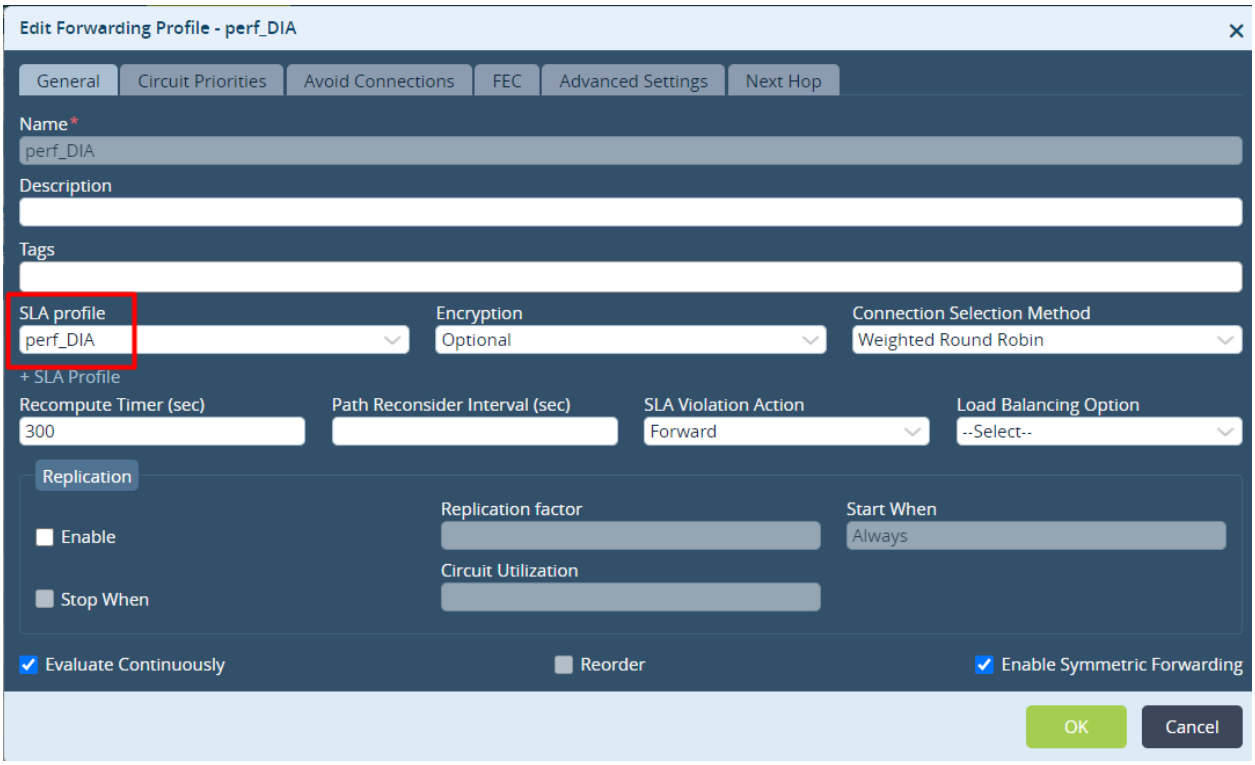
3. Configure a forwarding profile that references the SLA profile you configured in Step 2:

a. In Appliance view, select the Configuration tab in the top menu bar.

b. Select Services > SD-WAN > Forwarding Profiles in the left menu bar. The Add Forwarding Profile or Edit Forwarding Profile popup window displays.

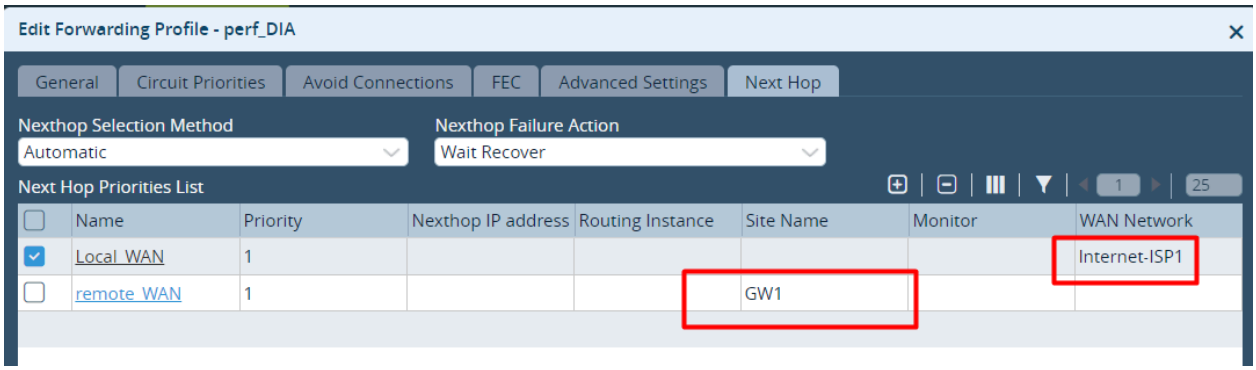
c. Click the Add icon to add a forwarding profile, or select an existing forwarding profile, as shown here.

d. Select the General tab, and in the SLA Profile field, select the SLA profile.



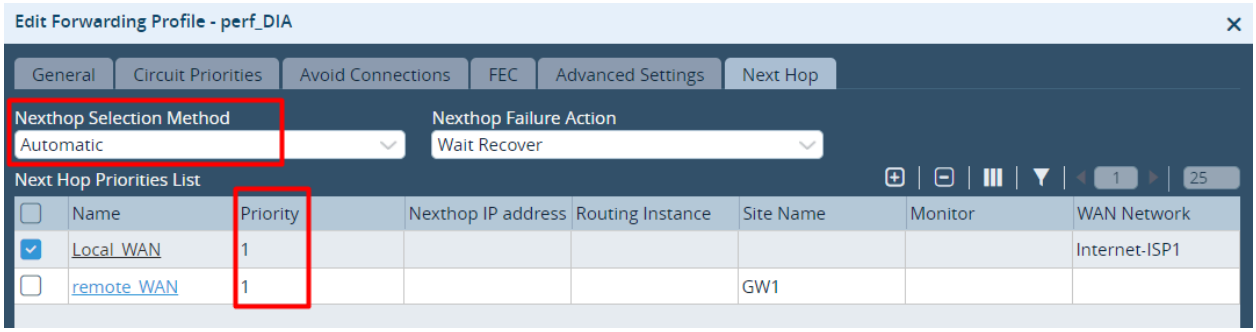
e. Select the Next Hop tab and specify the breakout options for the traffic used by this forwarding profile. In this example, there is a local breakout and a remote breakout to the internet:

i. For the local internet breakout, select the local WAN interface, and for the remote internet breakout, select the remote site where the remote breakout is available.



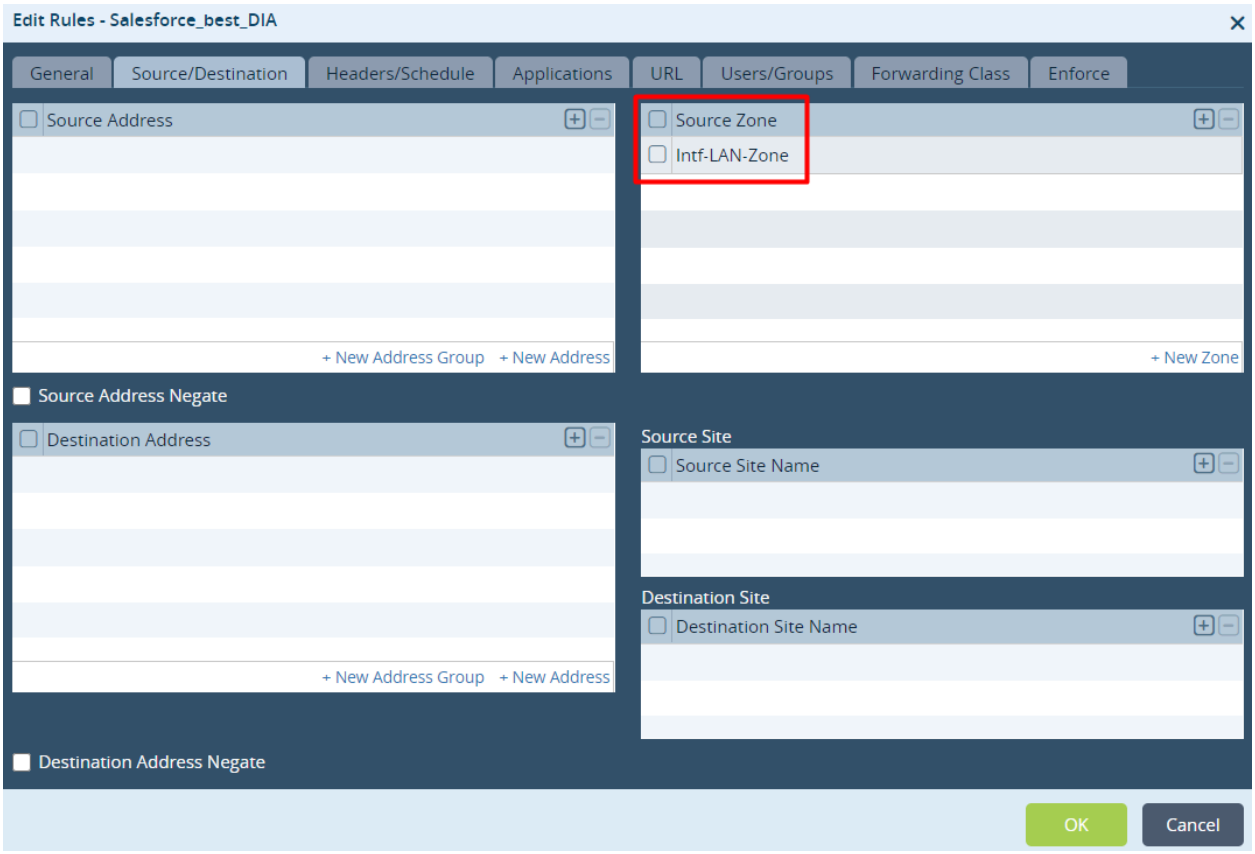
ii. Select the conditions for the breakout to use. In this example, because we have chosen to have the network automatically chooses the breakout that has the lowest latency, we select Automatic in the Next-Hop Selection Method field.

iii. Set the same priority value (here, 1) for both breakouts.



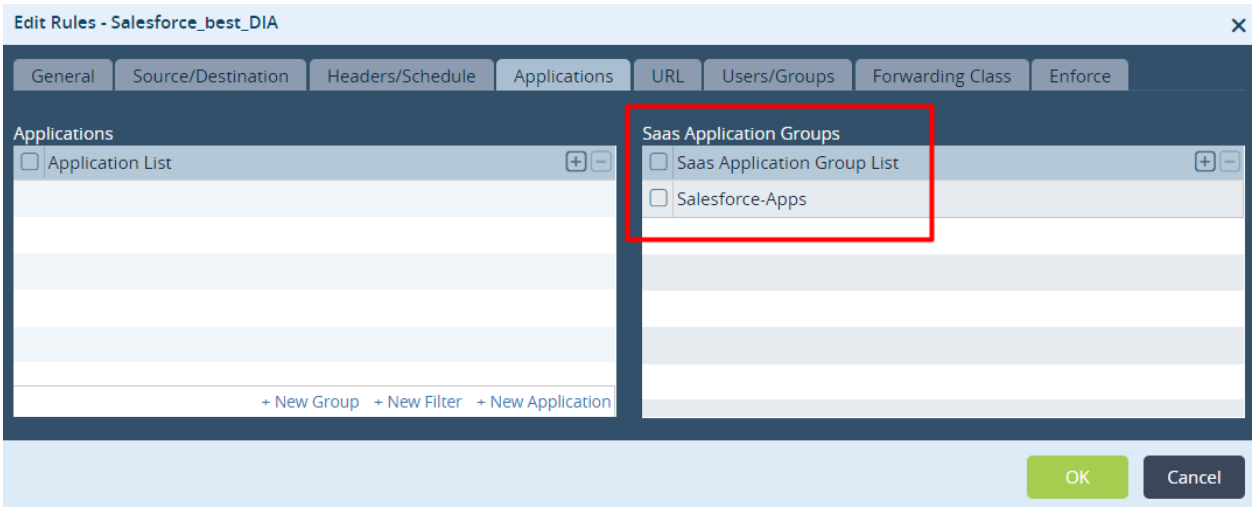
f. Click OK.

4. Configure an SD-WAN policy that defines the traffic classification criteria for this configuration example.
   a. In Appliance view, select a post-staging template.
   b. Select the Configuration tab in the top menu bar.
   c. Select Configuration > Services > SD-WAN > Policies > Rules in the left menu bar
   d. Click the Add icon to add a rule, or select an existing SD-WAN policy rule, as shown here.
   e. Select the Source/Destination tab, and in the Source Zone table, select the source zone. In this example, the source for the Saleforce traffic is a LAN zone.

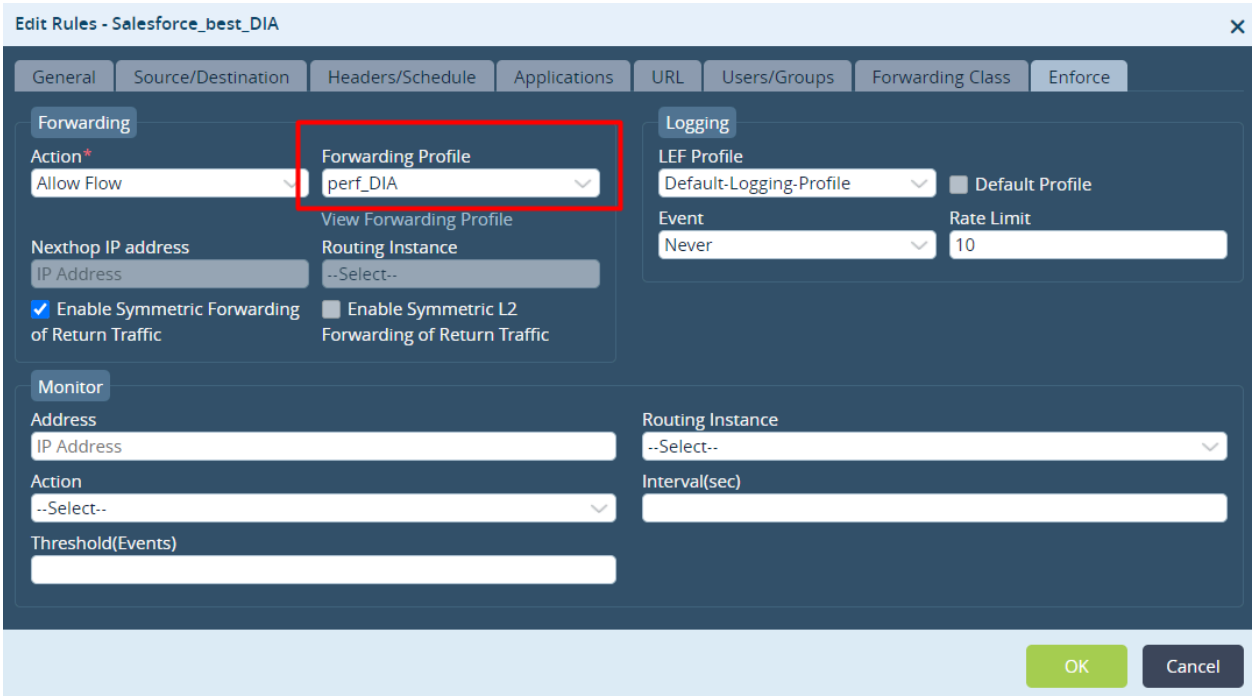


f. Select the Applications tab, and in the SaaS Application Groups table, select Salesforce-Apps. Note that this

example works only with the SaaS Application Group classification.



g. Select the Enforce tab, and in the Forwarding Profile, select the forwarding profile you configured in Step 3.



h. Click OK.

To verify the performance-based breakout configuration:

1. Verify the operation of the SaaS application monitor at the local and remote breakout points by issuing the **show application-monitor local detail** CLI command.

   The following output shows that the local breakout point measures the latency to Salesforce as 225 ms:

   ```
   admin@Site3-cli> show application-monitor local detail
   ```

```
APPLICATION
MONITOR                              LOSS      LATENCY   LOCAL       EXPORT
NAME       ROUTING INSTANCE          TYPE  PERCENTAGE  MILLISEC  ORGANIZATION
ORGANIZATION
-------------------------------------------------------------------------------------------
salesforce   Internet-ISP1-Transport-VR  icmp  0.0       225.53    Enterprise1   Enterprise1
```

The output at the remote breakout point measures the latency to Salesforce as 215 ms:

```
admin@GW1-cli> show application-monitor local detail
APPLICATION
MONITOR                              LOSS      LATENCY   LOCAL       EXPORT
NAME       ROUTING INSTANCE          TYPE  PERCENTAGE  MILLISEC  ORGANIZATION
ORGANIZATION
-------------------------------------------------------------------------------------------
salesforce   Internet-ISP1-Transport-VR  icmp  0.0       215.75    Enterprise1   Enterprise1
```

2. For the Salesforce traffic to reach the remote breakout point, you must add latency of the overlay. To verify this latency, issue the following CLI command:

```
admin@Site3-cli> show orgs org Enterprise1 sd-wan sla-monitor metrics last-1m GW1
                                 LOCAL  REMOTE
                                 WAN    WAN    TWO   FWD   REV   PDU   FWD
REV
SITE  PATH    FWD   LOCAL WAN    REMOTE WAN    LINK  LINK  WAY   DELAY DELAY LOSS
LOSS  LOSS
NAME  HANDLE  CLASS LINK         LINK          ID    ID    DELAY VAR   VAR   RATIO RATIO
RATIO
-------------------------------------------------------------------------------------------------
GW1   6627844 fc_ef Internet-ISP1 Internet-ISP1 2     2     171   0     1     0.0   0.0   0.0
```

3. The command output in Steps 1 and 2 shows that the SD-WAN logic is computing the best path (in this example, based on latency) and will select the local breakout point because it is better. To verify this:

```
admin@Site3-cli> show orgs org-services Enterprise1 sd-wan policies Default-Policy rules nexthop
application-monitor detail
                                          APPLICATION  APPLICATION  APPLICATION
APPLICATION
              NEXTHOP   NEXTHOP    NEXTHOP  NEXTHOP  HIT   MONITOR     MONITOR
MONITOR     MONITOR
NAME              PRIORITY  NAME       STATUS   ACTIVE   COUNT  NAME        TYPE
LATENCY     LOSS
-------------------------------------------------------------------------------------------------------------------
Salesforce_best_DIA   1       Local_WAN   up      yes     127    salesforce   icmp       226.87      0.
0
              1       remote_WAN  up      -       88     salesforce   icmp       389.7       0.0
```

4. Verify whether Salesforce sessions are actually going to the local breakout:

```
admin@Site3-cli> show orgs org Enterprise1 sessions brief | select application salesforce
VSN  VSN  SESS            DESTINATION    SOURCE  DESTINATION
ID   VID  ID    SOURCE IP    IP          PORT    PORT       PROTOCOL  NATTED  SD-WAN
APPLICATION
-------------------------------------------------------------------------------------------
```

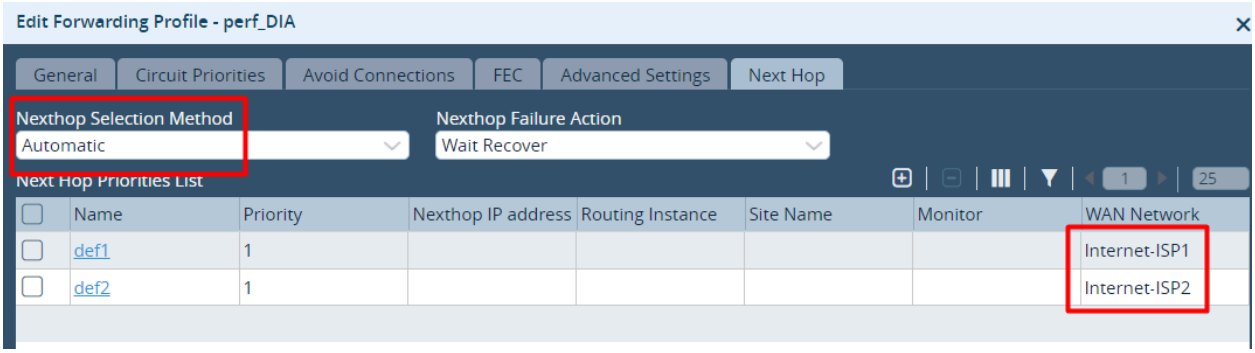
```
0   2   955   172.1.123.10   13.110.37.145   40880   443   6   No    No   salesforce
0   2   956   172.1.123.10   13.110.37.145   40880   443   6   Yes   No   salesforce
0   2   965   172.1.123.10   13.110.37.145   40886   443   6   No    No   salesforce
0   2   966   172.1.123.10   13.110.37.145   40886   443   6   Yes   No   salesforce
0   2   969   172.1.123.10   13.110.37.145   40890   443   6   No    No   salesforce
0   2   970   172.1.123.10   13.110.37.145   40890   443   6   Yes   No   salesforce
```

Variants of the Basic Performance-Based Breakout Configuration

The configuration example above is based on a performance-based breakout that has a local breakout option and a remote breakout option. You can use the same approach when there are multiple local breakout possibilities, for example, two WAN transports. If these local WANs are provided by different ISPs, each may have different performance characteristics. For example, the WAN transport provided by ISP1 might be closer to the SaaS application server than ISP2. Another common use case is when a cloud security provider provides the breakout. In this case, an IPsec or GRE tunnel directs the traffic to the cloud security provider and you can apply the same logic for all such use cases.

For these scenarios, you configure the next-hop selection method to Automatic in the forwarding profile, as follows:

1. Configure a forwarding profile with two ISPs as next-hop options. For more information, see Configure SD-WAN Traffic-Steering.
2. In Appliance view, select the Configuration tab in the top menu bar.
3. Select Services > SD-WAN > Forwarding Profiles in the left menu bar.
4. Click the Add icon, or select an existing forwarding profile, as shown here.
5. Select the Next Hop tab.
6. In the Next-Hop Selection Method, select Automatic.
7. In the Next-Hope Failure Action field, select Wait Recover.



8. Click OK.

To verify which local breakout is the best performing one, issue the following CLI command:

```
admin@Site2-cli> show orgs org-services Enterprise1 sd-wan application-metrics brief
                 METRIC                          REMOTE   REMOTE           HIT
APPLICATION      TYPE    DESTINATION IP          LOCAL CIRCUIT  CIRCUIT  SITE    METRIC  TTL   COUNT
```

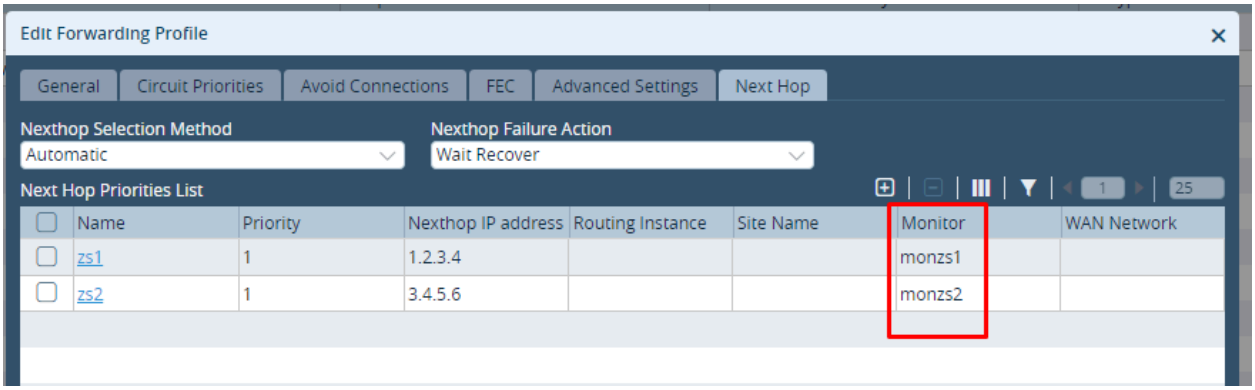
```
-------------------------------------------------------------------------------------------
Concur-Apps  VLR     0.0.0.0/0         Internet-ISP1  -       -      11     2996  2
                                       Internet-ISP2  -       -      59     2997  1
                     23.38.19.188/32   Internet-ISP1  -       -      7      2822  0
                                       Internet-ISP2  -       -      4      2996  1
                     92.123.164.163/32 Internet-ISP1  -       -      11     2996  1
                                       Internet-ISP2  -       -      61     2997  5
```

The CLI output above shows the different metrics for the local circuits toward ISP1 and ISP2. For this test, the internet circuit to ISP1 is impaired by high loss and high latency, resulting in a bad link score. Therefore, for the last entry in the output, the metric for ISP1 remains low, while the metric for ISP2 increases. The result is that ISP2 gets more sessions (a hit count of 5 compared to a hit count of 1 for ISP1).

Another example is a performance-based breakout to the best performing external cloud security provider. In this use case, you set the forwarding profile next hop to the IP address of the remote tunnel IP endpoint. For all scenarios, it is recommended that you add a monitor to the next hop. In this case, the monitor is verifying the IP endpoint of the tunnel. If the IP endpoint is not reachable, this next hop is not considered.

To configure the forwarding profile next hop to the IP address of the remote tunnel IP endpoint:

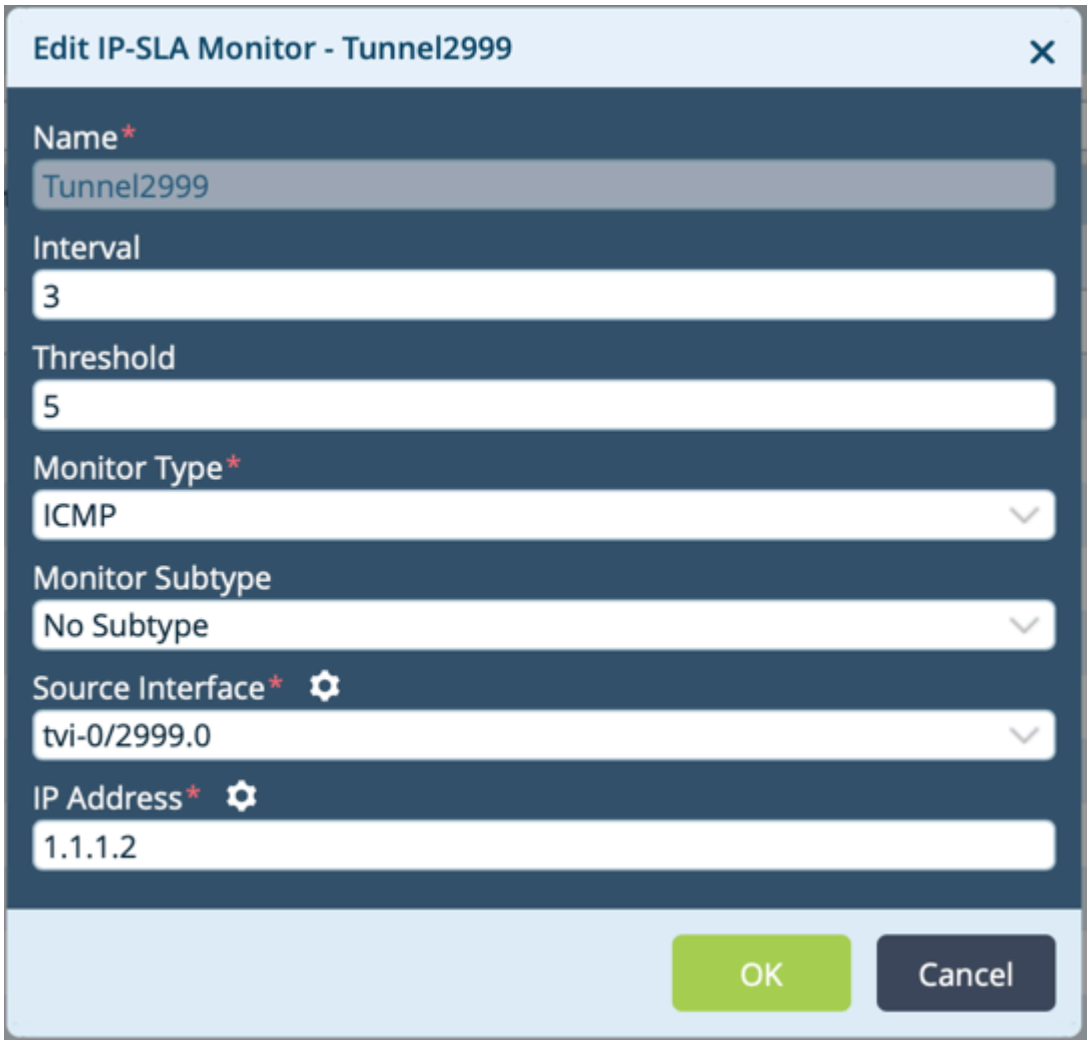
1. Configure a forwarding profile with two ISPs as next hop options. For more information, see Configure SD-WAN Traffic-Steering.
2. In Appliance view, select the Configuration tab in the top menu bar.
3. Select Services > SD-WAN > Forwarding Profiles in the left menu bar.
4. Click the Add icon, or select an existing forwarding profile, as shown here. The Add Forwarding Profile or Edit Forwarding Profile popup window displays.
5. Select the Next Hop tab.
6. In the Next-Hop Select Method field, select Automatic.
7. In the Next-Hop Failure Action field, select Wait Recover.
8. Click the Add icon, and in the Add Next-Hop Priorities popup window, specify the IP address for next hop to the remote tunnel IP end point. GRE and IPsec are the most common tunnel options. For more information, see Configure Site-to-Site Tunnels and Configure Interfaces.



9. In the Monitor Field, select a monitor to use for the tunnel.
10. Click OK twice.

After you configure the tunnels, add a static route to route traffic to them. It is recommended that you monitor the tunnel endpoint is monitored so that if the tunnel fails, the static route is withdrawn. You do this by configuring an IP SLA monitor, as follows:

1. In Appliance view, select the Configuration tab in the top menu bar.
2. Select Networking > IP SLA > Monitor in the left menu bar. The main pane displays the IP SLA monitor objects that are already configured. For more information, see Configure IP SLA Monitor Objects.
3. Click the Add icon, or select an existing IP SLA monitor, as shown below. The Add IP SLA Monitor or Edit IP SLA Monitor popup window displays.



4. Associate this IP SLA monitor with the static route. If you want an active/standby connection to the cloud security proxy, the preference or metric can influence which route to which tunnel is active. For more information, see Associate an IP SLA Monitor Object with a Static Route.

a. In Appliance view, select the Configuration tab in the top menu bar.

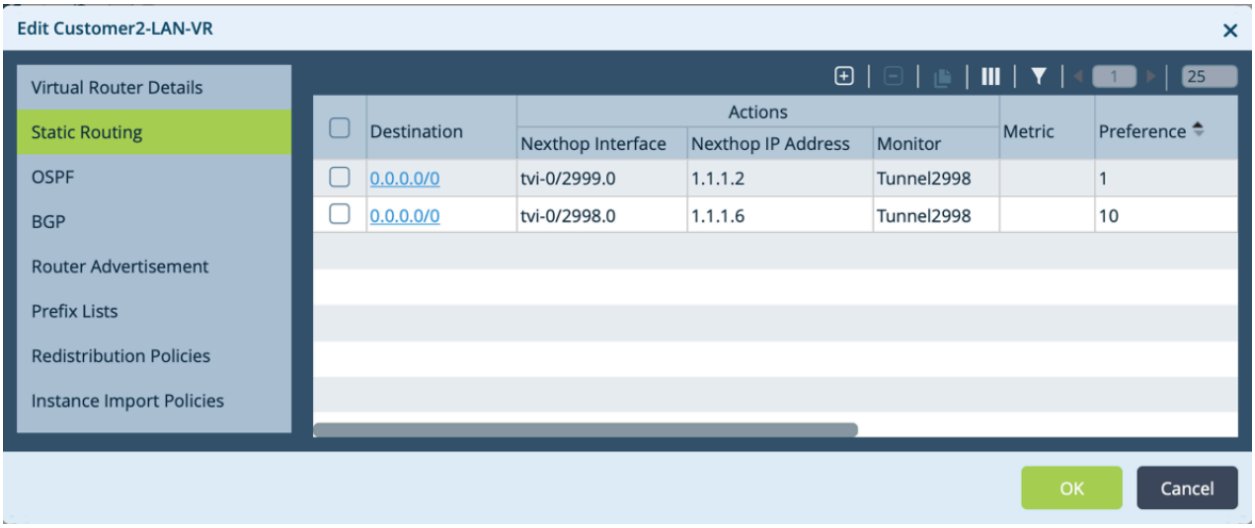
b. Select Networking > Virtual Routers in the left menu bar

c. Select a virtual router instance.

d. In the Edit VR popup window, select the Static Routing tab.

e. Select the IPv4/IPv6 Unicast tab in the horizontal menu bar.

f. Click the Add icon to add a static route to associate with the monitor object. If you are adding an IP SLA monitor to an existing static route, click the address of the static route in the Destination column.



Best Practices for Performance-Based Breakout

SaaS application monitoring has many configuration options. Your SaaS application may not work for the basic measurement based on ICMP, as shown in this example, and for best results you may to ture the SaaS application monitor.

DNS proxy and web proxy may interfere with the functioning of the SaaS application, so check the following sections to understand the roles of these two types of proxies. If the functioning is affected, enhance the configuration with DNS proxy (always recommended) and web proxy chaining (required when you use a centralized web proxy).

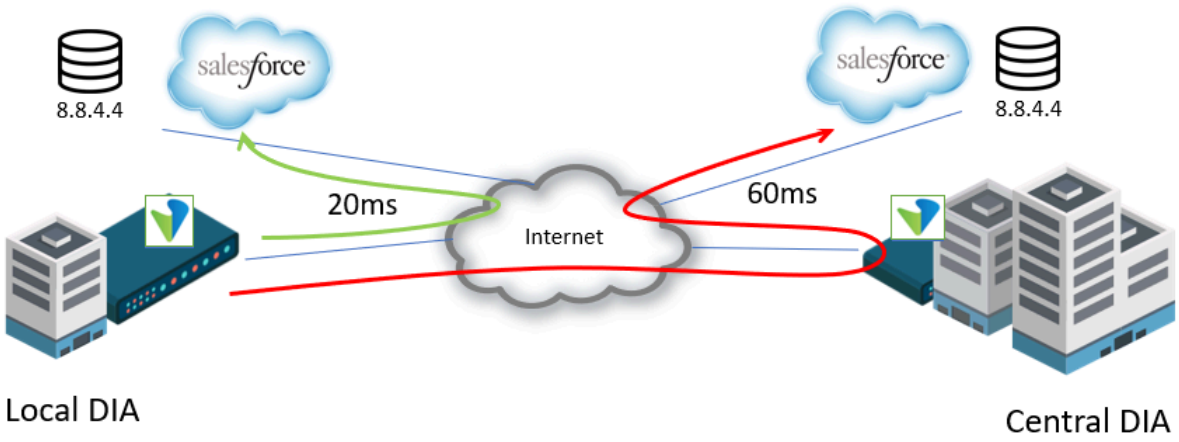
Note that the you can optimize the SaaS applications listed in the SaaS application group list. Applications that are not listed cannot leverage the performance-based breakout functionality.

DNS Proxy in Breakout Scenarios

The end-user application experience may depend on where the application breaks out from. For example, suppose a globally deployed enterprise SD-WAN VPN uses a local internet breakout in Hong Kong and a central breakout in London. The end users use an internal DNS server in London. If a user in Hong Kong does a DNS query for

Salesforce.com from the London DNS server, the query likely fetches the IP address of the London-based instance of Salesforce.com. However, in Hong Kong there is also an instance of the Salesforce.com SaaS. If the network decides to breakout in Hong Kong for Salesforce.com, it must query the Hong Kong DNS server to receive the IP address of the local instance. This scenario requires a DNS proxy. The configuration in this section shows that a local DNS proxy policy rule matches Salesforce.com, and based on where in the network the policy decides that the best breakout point is. that DNS server is used. The DNS server might be the Google DNS server in Hong Kong or the Google DNS server in London.

The following shows that internal metrics report a better performance, measured hereby latency, to the local SaaS. This local SaaS can be contacted only if the local instance of the global DNS service (here, the Google DNS server in Hong Kong) is used. If the remote global DNS is used, the pointer is to the remote SaaS. In that case, traffic still breaks out locally, but it follows the internet to get to the remote SaaS.



DNS Proxy Configuration

The DNS proxy configuration described here is an extension to the use case described in Performance-Based SaaS Optimization.

You configure the DNS proxy only on the local DIA edge device. No additional configuration is required on the central DIA edge device.

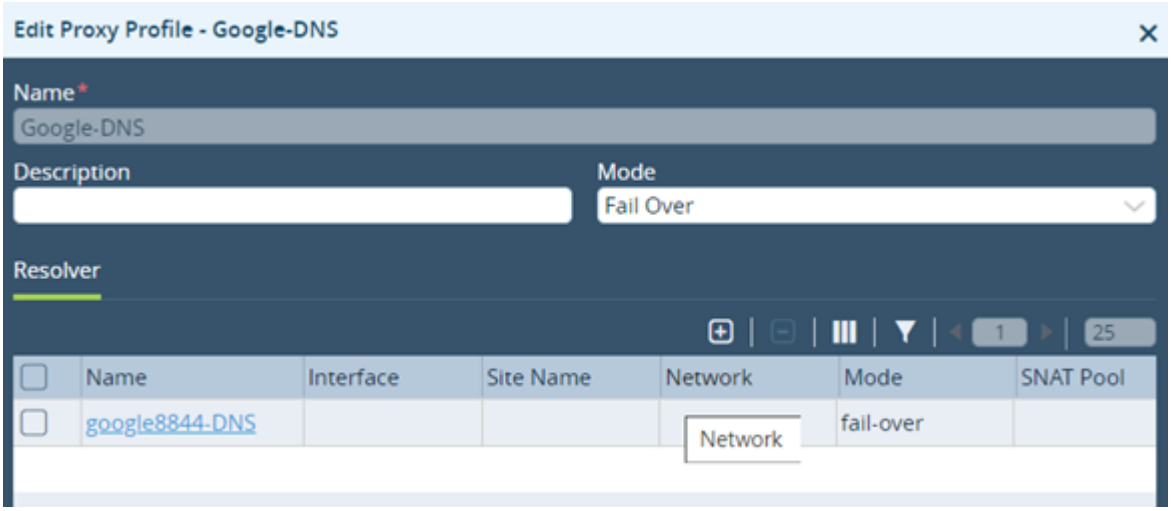
For performance-based SaaS optimization, you must use the DNS server of the local ISP or a global DNS provider, such as a Google DNS server. This configuration example uses the Google DNS server 8.8.4.4 for Salesforce.com.

To configure the DNS proxy:

1. Configure a DNS proxy profile. For more information, see Configure DNS Proxy Profiles.
   a. In Appliance view, select the Configuration tab in the top menu bar.
   b. Select Networking > DNS > Proxy Profiles in the left menu bar.

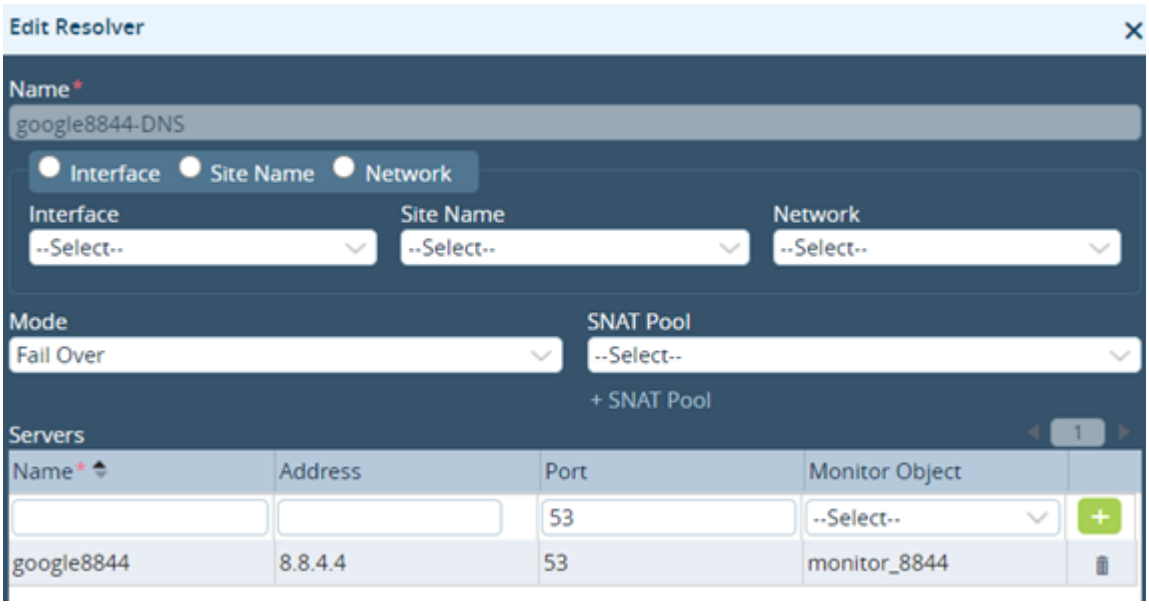
c. Click the Add icon, or select an existing proxy profile, as shown here. The Add Proxy Profile or Edit Proxy Profile popup window displays.

d. In the Name file, enter a name for the DNS proxy profile. Here, the name is Google-DNS.



e. Select the Resolver tab to configure a DNS resolver.

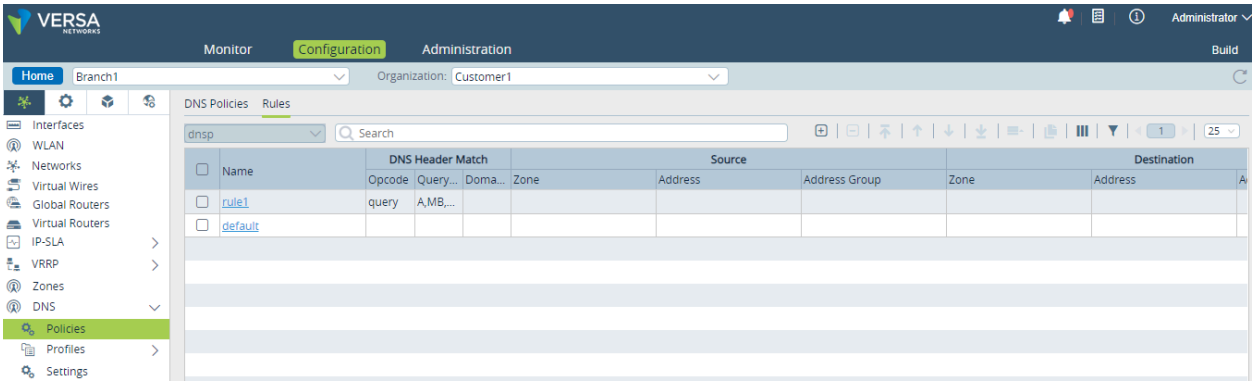
f. Click the Add icon, or select an existing DNS resolver, as shown here. The Add Resolver or Edit Resolver popup window displays.



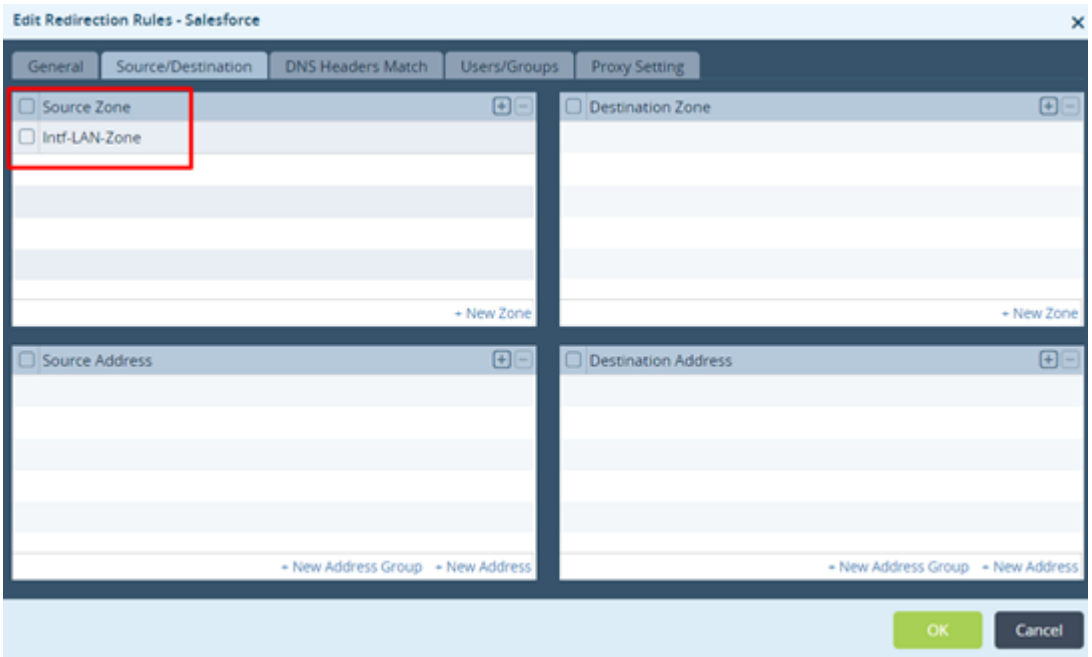
It is recommended that you monitor the reachability of the configured DNS server. If the DNS server is not reachable, the monitor fails, the DNS proxy stops forwarding DNS queries to that DNS server, and instead, the DNS server configured one the client is used. Because the DNS server is dynamically assigned, for this example, only the specification of the DNS server is relevant. This example chooses to monitor the

reachability of the DNS server by applying an IP SLA monitoring object. If the DNS server is not reachable, this DNS profile is not used, thus& avoiding the potential blackholing of DNS queries.

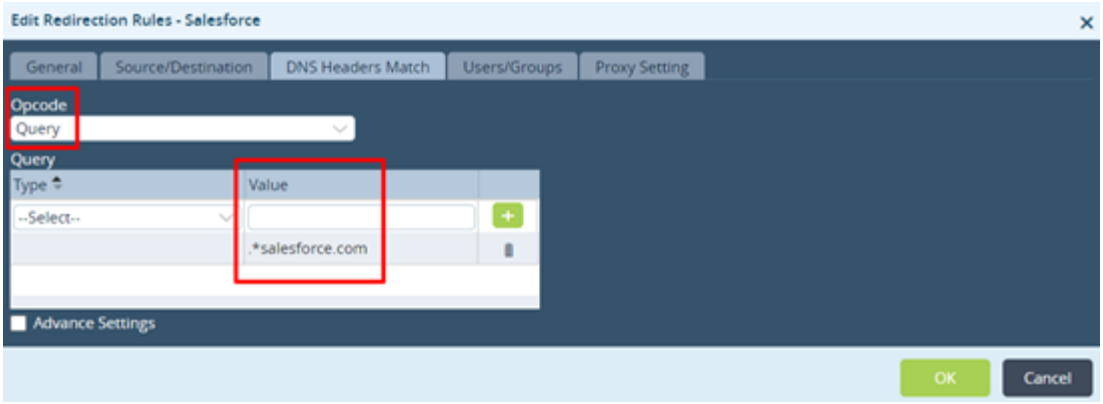
2. Configure DNS proxy policy rule to specify matching conditions, that is, where the DNS query originates from and to which FQDN it applies). For more information, see Configure DNS Redirection Rules.
   a. In Appliance view, select the Configuration tab in the top menu bar.
   b. Select Networking > DNS > Policies in the left menu bar.
   c. Select the Rules tab in the horizontal menu.



   d. Click the Add icon. The Add Redirection Rules popup window displays.
   e. Select the Source/Destination tab, and in the Source Zone table, select the source zone.



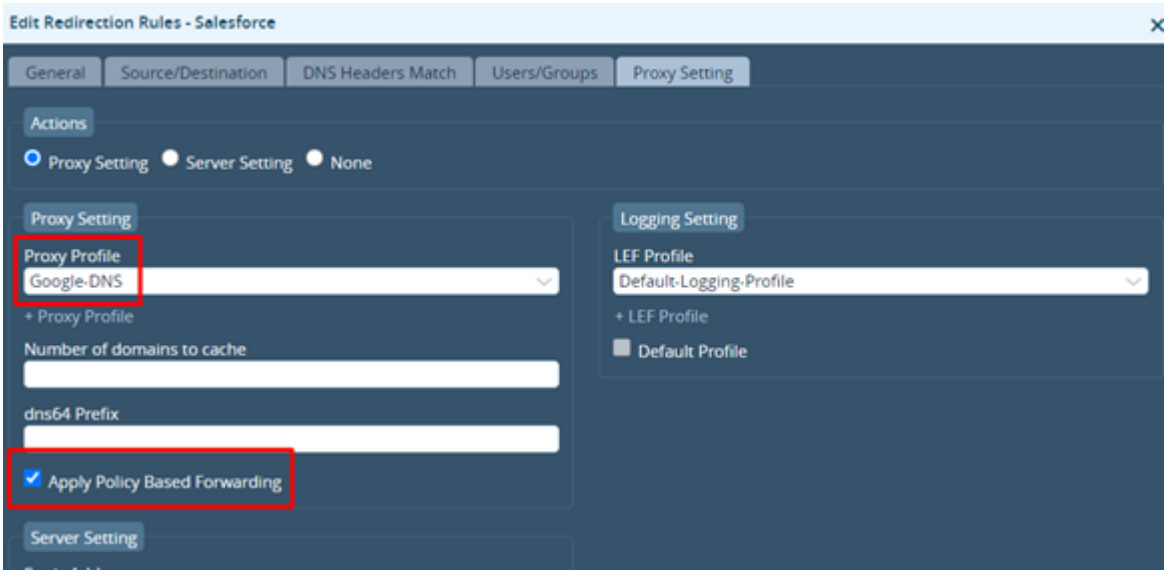
   f. Select the DNS Headers Match tab and enter the following information:



g. In the Opcode field, select the value Query.

h. In the Query Value field, select the FQDN regex notation. The exact value of regex notation is important. Because we want the query for be for all traffic designated to the Salesforce.com domain, use the regex notation ".*salesforce.com".

i. Select the Proxy setting tab.



j. In the Actions field, click Proxy Setting.

k. In the Proxy Profile field, select the proxy profile you created in Step 1, here, Google-DNS.

l. Click Apply Policy-Based Forwarding. This setting enables the looking up of SD-WAN policies to ensure that the SaaS application breaks out to the best performing DIA. If you do not select this option, the DNS server 8.8.4.4 is used, but the egress interface is now DTVI-0/51, which is the overlay tunnel to the gateway. You can also verify whether the proxy, here, follows the routing table to get to the specified DNS server. Because the SD-WAN policy may dynamically decide whether the break out is local or central, the DNS proxy must follow that decision.

m. In the LEF Profile field, select the log profile to use, which is useful when you want to verify DNS queries.

3. Click OK.

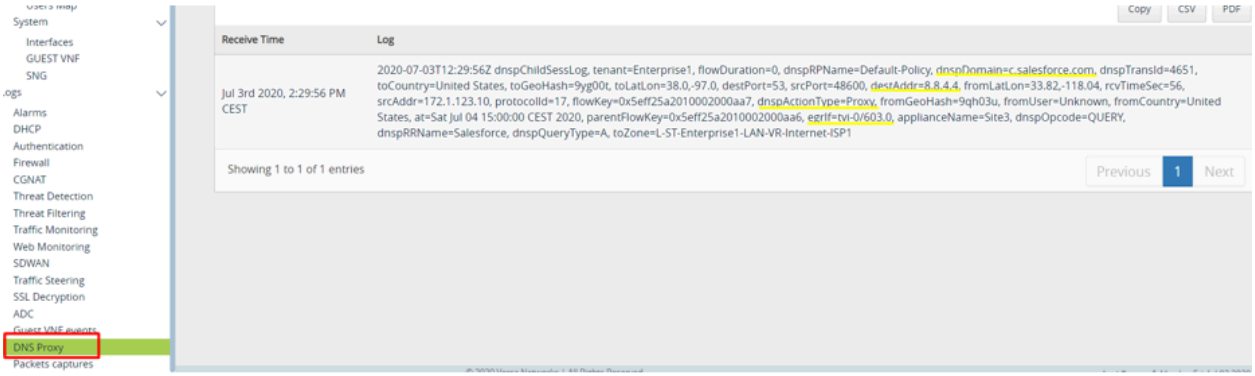
To verify the DNS proxy configuration:

1. To verify that the current path to Salesforce.com is the best performing path, issue the following CLI command:

```
admin@Site3-cli> show orgs org-services Enterprise1 sd-wan policies Default-Policy rules nexthop
application-monitor detail
                                                                                    APPLICATION  APPLICATION  APPLICATION  APPL
                NEXTHOP  NEXTHOP   NEXTHOP  NEXTHOP  HIT  MONITOR      MONITOR      MONITOR      MONITOR
NAME            PRIORITY NAME      STATUS   ACTIVE   COUNT NAME        TYPE         LATENCY      LOSS
-------------------------------------------------------------------------------------------------------------
Salesforce_best_DIA  1        Local_WAN  up       yes      79    salesforce   icmp         326.42       0.0
                     1        remote_WAN up       -        108   salesforce   icmp         398.2        0.0
```

The output shows that the path to Salesforce.com from the local breakout has the lowest latency (326 ms compared to 398 ms) and is the active path.

2. To verify that the configuration is functioning correctly and that the DNS query is sent to the Google DNS server from the local DIA:
   a. Select Analytics > Home > Logs > DNS Proxy.



   b. Zoom in to view details:

The output above shows that for salesforce.com, the DNS server 8.8.4.4 is used (the default DNS for the client is an enterprise internal DNS) and that the egress interface is TVI-0/603, which is the internal interface for DIA. (Check the dnspDomain, destPort, destAddr, and egrIf fields in the output.)You can also verify the DIA interface by issuing the **show interface brief** CLI command.

3. To whether the local DIA might be impaired and so the central DIA is preferred, issue the following CLI command:

```
admin@Site3-cli> show orgs org-services Enterprise1 sd-wan policies Default-Policy rules nexthop
application-monitor detail
                                                                                    APPLICATION  APPLICATION  APPLICATION  APPL
                NEXTHOP  NEXTHOP   NEXTHOP  NEXTHOP  HIT  MONITOR      MONITOR      MONITOR      MONITOR
NAME            PRIORITY NAME      STATUS   ACTIVE   COUNT NAME        TYPE         LATENCY      LOSS
-------------------------------------------------------------------------------------------------------------
Salesforce_best_DIA  1        Local_WAN  up       -        191   salesforce   icmp         329.41       0.0
```

1	remote_WAN	up	yes	108	salesforce	icmp	188.72	0.0

In Analytics, you can display the details:

This verification shows that the DNS server 8.8.4.4 is used, but the egress interface is now DTVI-0/51, which is the overlay tunnel to the gateway. You can also verify the DIA interface by issuing the **show interface dynamic tunnels** CLI command.

Best Practices for DNS Proxy

The configuration example here describes how the SD-WAN network can dynamically intercept the client DNS query and forwards the query to the DIA point that provides the best SaaS performance. The example shows the minimum configuration required for this use case to work.

One can image more advanced uses of DNS proxy. For example, enterprises often use their own internal DNS server. You might configure a global DNS external service, such as Google DNS, and an internal DNS for the IP endpoints, but you prefer local breakout to a global external DNS. For this type of scenarios, a DNS proxy provides a solution, intercepting DNS traffic for the internal domain (based on the internal FQDN) and having the rest of the traffic be served by the global external DNS.

Security Services for Internet Breakout

Traditionally, enterprises follow strict security policies to connectLANs to the internet. In legacy networks, in which internet access is typically provided from a centralized internet breakout, you can easily enforce security at one location. A centralized stack of firewall services, often acting as web proxy, inspect the traffic coming from or going to the internet against enterprise security policies. However, now, with the introduction of local internet breakout services, you must ensure that enterprise security policies do not change, because you still want to scrub internet-bound traffic.

The following list summarizes some of the security scenarios observed in the field:

- There is no additional security at local internet breakout. Rather, you rely on NAT to secure the enterprise LAN. This is an unsecured strategy, and there are no controls to monitor or mitigate compromised networks.
- Security policies are enforced through a nearby cloud security proxy of a third-party cloud security provider. This is a secure alternative, but costly cloud security proxy services may not always be provided locally, and this can result in less than optimal application experiences for end users.
- Security policies are enforced through a third-party firewall, and all internet-bound local breakout traffic is scrubbed by that firewall. In the case of the Versa Networks solution, this third-party firewall is typically virtualized as a guest VM on a Versa uCPE. This is a secure method, but it is complex to manage, and you have to manage the lifecycle of the security solution separately.
- Security policies are enforced using the Versa networks next-generation firewall (NGFW) with unified threat management (UTM) features.

Enterprise security administrators may also make risk assessments about which types of application to break out locally, without performing a full stack inspection. If you do local internet breakout for a particular application, for example, Office 365, and you break out all other internet traffic centrally, Microsoft claims that the Office 365 service is extremely secure.

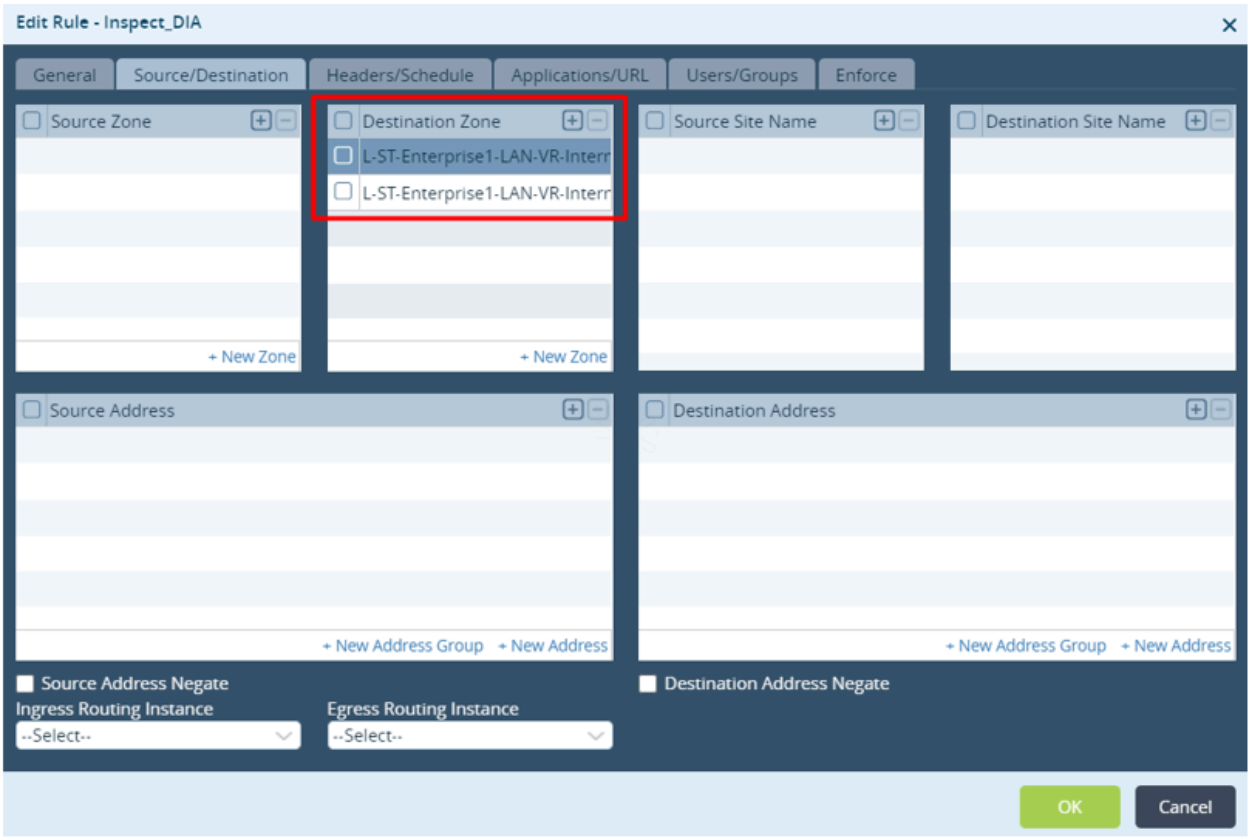
(Microsoft claims that full-stack security is performed on their applications.) For performance reasons they suggest that you not enforce additional security policies, and that if you adopt this suggestion, you do not need to implement additional security services.

For any breakout to the internet, Versa Networks recommends that you deploy, at a minimum, IDS/IPS with the Versa-recommended profile and antivirus. These two features ensures that all traffic from untrusted networks is inspected, regardless of whether the application provider claims to be secure. Note that you must also implement SSL decryption at the branch to effectively investigate the SSL stream.

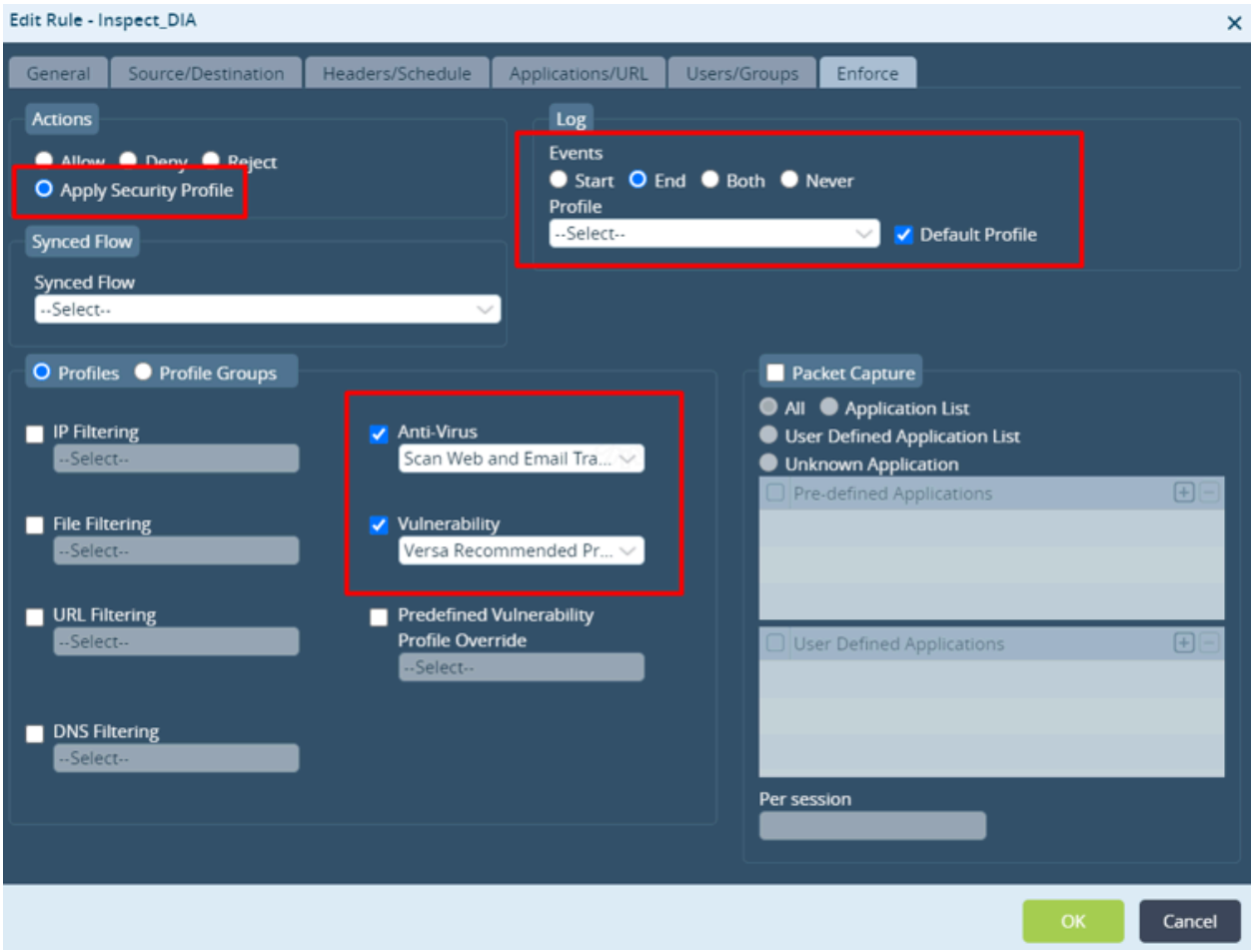
The following example shows how to configure NGFW security for DIA traffic. In this scenario, for all firewall policy rules that have an allow rule and that apply to an internet-bound application, and therefore are logically are broken out to the internet, you must attach a UTM security profile to the firewall policy rule.

To configure NGFW security for DIA traffic, configure the NGFW access policy rule:

1. In Appliance view, select the Configuration tab in the top menu bar.
2. Select Services > Next Gen Firewall > Security > Policies in the left menu bar, and select the Rules tab.
3. Click the Add icon to define rules for the policy, or select an existing rule, as shown here.
4. Select the Source/Destination tab:



5. In the Destination Zone table, select L-ST-*organization-name*-VR-*WAN-name* as the destination zone. In this example, there are two local internet breakout points, so you must select the matching zone for both. You can leave all other matching criteria blank.
6. Select the Enforce tab:



7. In the Actions field, click Apply Security Profile.
8. In the Log field, click Default Profile.
9. Click Antivirus and select an option. Here, the option selected is Scan Web and Email Traffic.
10. Click Vulnerability and select an option. Here, the option selected is Versa Recommended Profile, which performs malware inspection using the VOS antivirus module and IDS/IPS using the VOS vulnerability module.
11. Click OK.

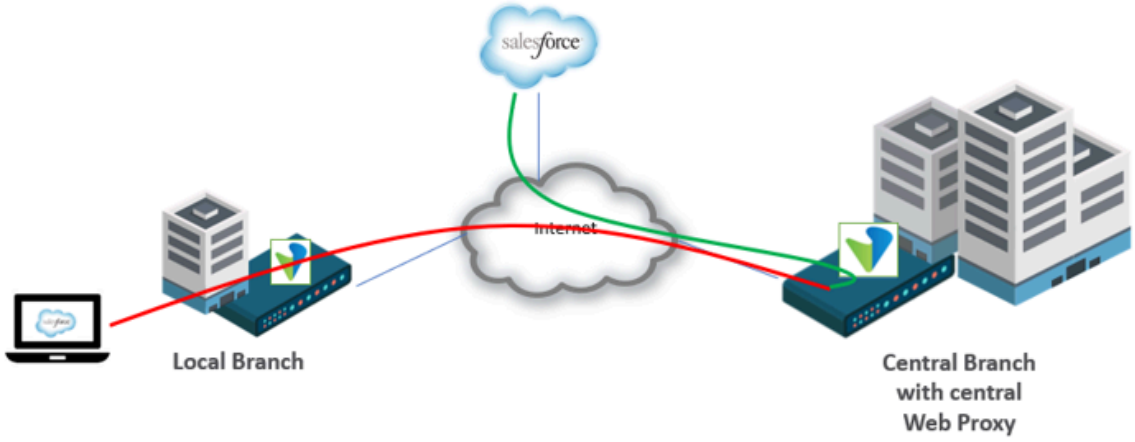
Web Proxies with Different Breakout Services

Using web proxies to provide enterprise security for internet access is a common use case and a best practice. Using web proxies provides you with more control over the traffic that goes to the internet, because the web proxy inspects all the internet-bound traffic. Another benefit is that with web proxies, you do not have to announce a default route in the

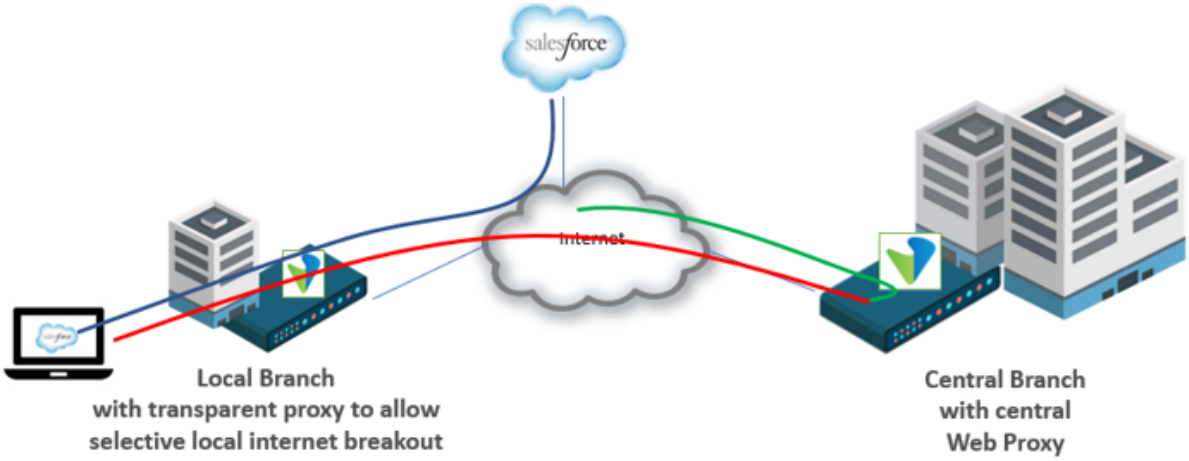
enterprise LAN. Instead, endpoints reach the internet by connecting to the proxy server, as specified in the endpoint configuration and a proxy autoconfiguration (PAC) file.

However, if you prefer to leverage quality-of-experience improvements for specific SaaS applications, as discussed in DNS proxy section, above, you must intercept the connection to the centralized proxy service to allow local internet breakout. The local DIA point must intercept the proxy session and act as an alternative web proxy dynamically, without modifying the endpoint proxy configuration.

The following figure shows how the Versa VOS edge device software can function as web proxy. Alternatively, you can use an existing third-party web proxy at the central location. However, it becomes difficult for the local branch to do a local internet breakout, because the client configuration includes an explicit central web proxy endpoint. The VOS web proxy software web proxy addresses this scenario.

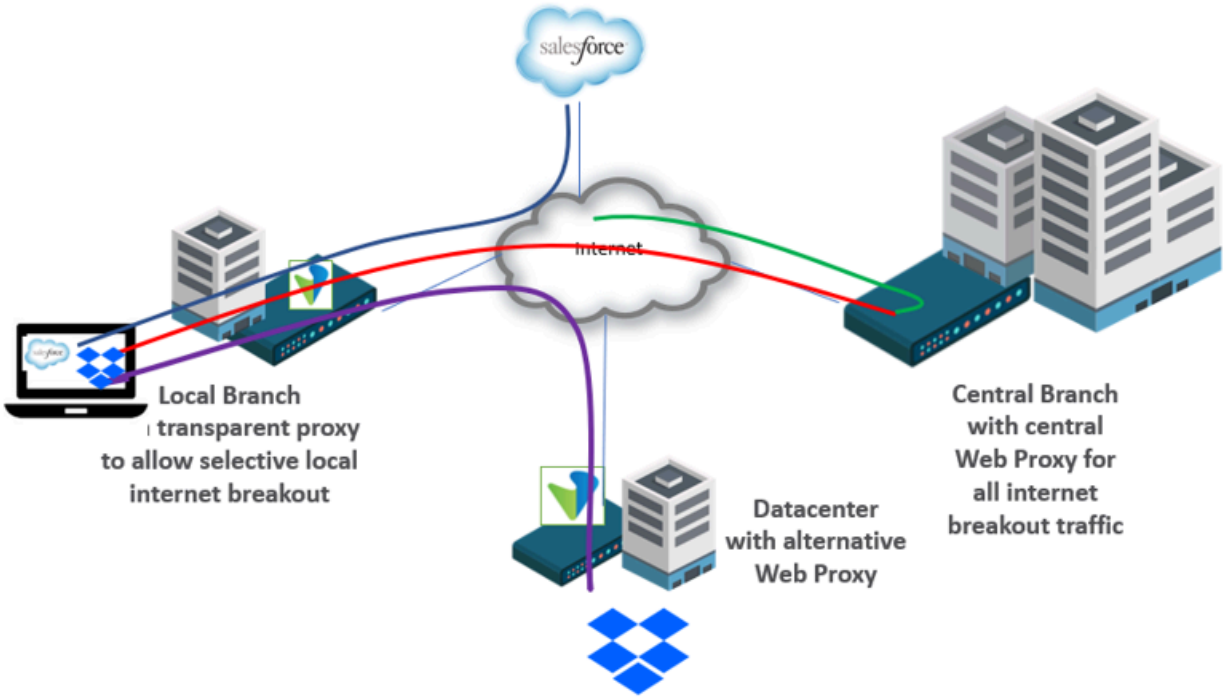


In the following figure, the local branch is configured to provide local DIA only for Salesforce.com. Because the client computer, by default, sends all HTTPS traffic to the central explicit web proxy, the local branch monitors the explicit proxy port number for queries to Salesforce.com. When it detects these queries, the local branch can use a transparent web proxy to break out to the internet and then forward the traffic to internet. You must also configure the DNS proxy feature so that DNS queries to Salesforce.com are intercepted.



A more advanced use case is to use multiple web proxies. In addition to redirecting local internet breakout for specific SaaS traffic, using multiple web proxies allows you to redirect specific traffic to an alternative proxy service, a process that is called proxy chaining.

The following figure shows that the VOS edge devices in the data center and in the central branch are both configured for explicit web proxy. The client in the local branch is configured with a proxy in the central branch. However, the local branch can break out Salesforce.com directly to the internet and can proxy-chain dropbox.com traffic to the data center branch. You use DNS proxy configurations to select the correct DNS server.



Supported Software Information

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

Additional Information

Configure Basic Features
Configure Direct Breakout to the Internet
Configure DNS Proxy Profiles
Configure DNS Redirection Rules
Configure IP SLA Monitor Objects
Configure SaaS Application Monitoring
Configure SD-WAN Traffic Steering
Configure SLA Profiles for SD-WAN Traffic Steering
Configure Virtual Routers
Performance-Based SaaS Optimization
SD-WAN Gateway Use Cases
VOS Edge Device DIA Architecture and Best Practices

SD-WAN Gateway Use Cases

For supported software information, click here.

There are three main use cases for VOS edge devices that act as SD-WAN gateways:

- Connect to sites on an MPS Layer 2 VPN network
- Connect to sites over disjointed underlay networks
- Act as a gateway for internet-bound traffic

This chapter discusses each of these use cases.

Connect to Sites on an MPLS Layer 3 VPN Network

An SD-WAN gateway allows sites connected to the SD-WAN VPN network to communicate with sites that are connected to a legacy MPLS VPN network. To enable this communication, you configure a VOS edge device as a gateway. This gateway facilitates the routing of information between the MPLS underlay network and the SD-WAN VPN network. The exchange of routes is typically done using a dynamic routing protocol such as BGP.

When you are migrating to SD-WAN, you can set up a gateway temporarily so that sites in the MPLS Layer 3 VPN network can continue to connect to sites that have already migrated to SD-WAN. After you have migrated all the sites to they SD-WAN-based VPN, the gateway is no longer required. You can also use a permanent gateway when the MPLS

network has sites or services that must be accessed by SD-WAN–enabled sites, for example, if the SD-WAN–enabled sites need to access Azure Express Route connection in the MPLS network.

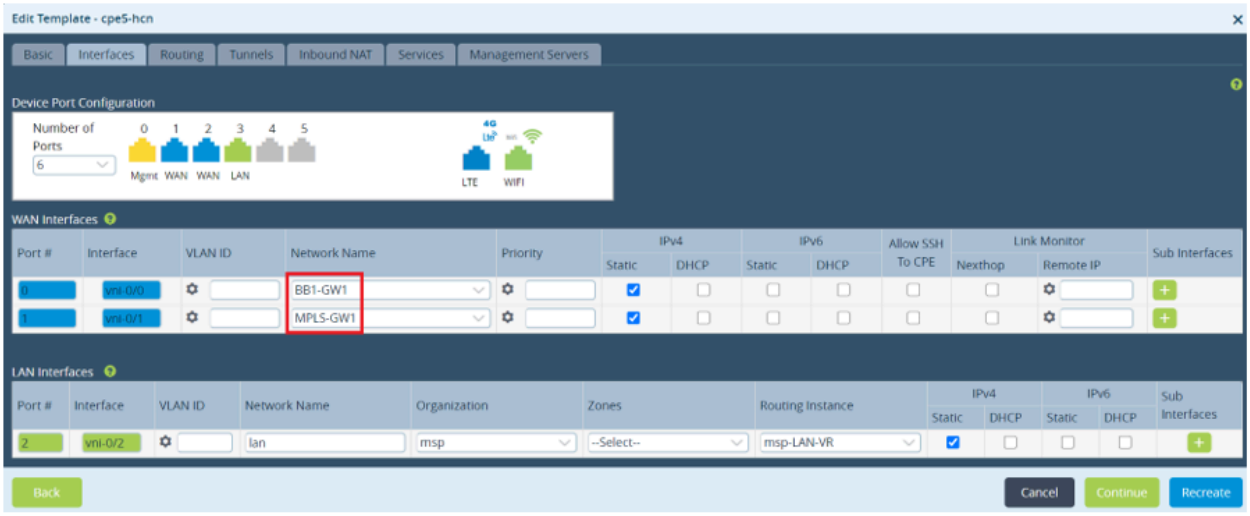
The following figure illustrates an SD-WAN gateway.

The SD-WAN gateway uses the same VOS edge device software that is used in a regular SD-WAN branch, which means that a regular branch can also serve as a gateway. SD-WAN gateways can be multitenant, and in a service provider environment, it is common to find such multitenant gateways that serve multiple customer networks. An SD-WAN gateway can also be used for non-SD-WAN interconnects, in which you can replace the MPLS Layer 3 VPN with any IP network.

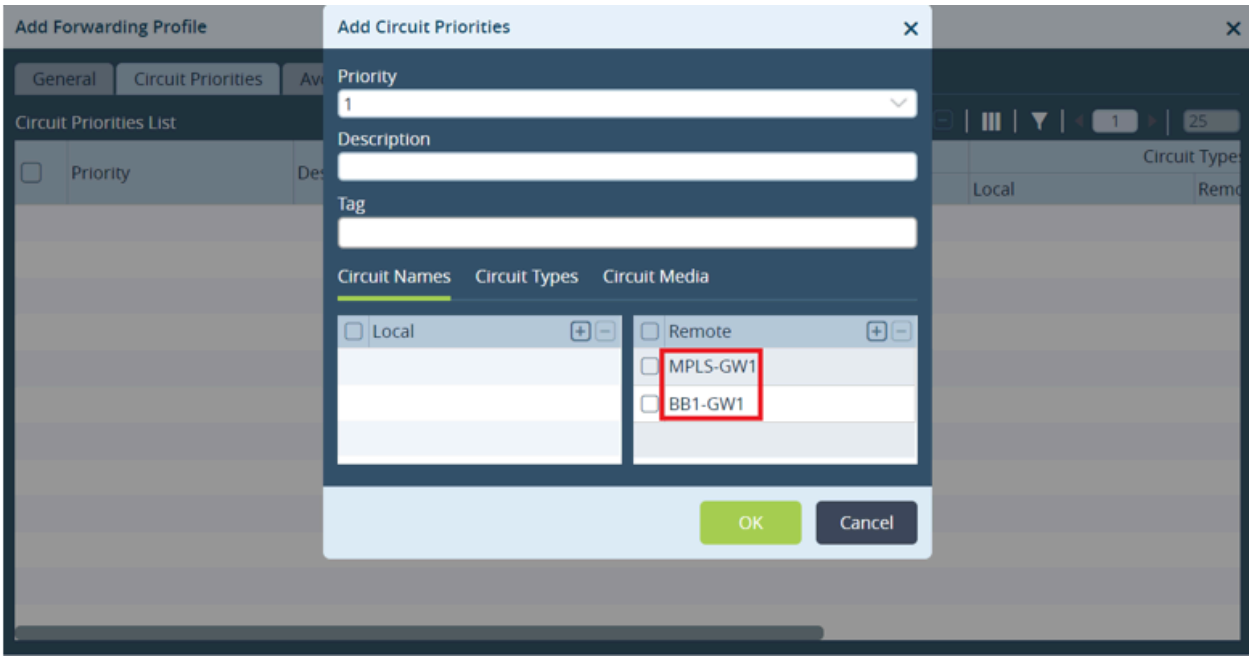
The following are best practices for SD-WAN gateways:

- For device-level redundancy, use HA-enabled sites as gateway.
- For redundancy, use multiple gateways across different availability zones or geographic regions.
- Keep traffic symmetric. For routes that the gateway exchanges with the provider, modify the BGP path attributes to ensure that bidirectional traffic uses the same gateway device. Asymmetric routing may cause issues in the SD-WAN when you use SD-WAN traffic-steering policies.
- If you have multiple gateway sites that establish peering to the same MPLS provider, use the same BGP AS number for the SD-WAN and the provider side of the BGP session. (Note that the default SD-WAN IBGP AS number is 64512.) The BGP AS path check ensures that routing loops are not formed when a route learned from one gateway is inadvertently advertised back to another gateway. The BGP AS path check also takes care of routing loops if the same route learned from an MPLS provider is advertised back to any gateway site. Similarly, an SD-WAN route advertised from one gateway to the provider must not be advertised back by the provider at another gateway site.
- It is recommended that you use BGP communities to color the routes when they are advertised and received at each BGP gateway. An example is to color routes based on region. For example, you can use different communities to identify routes from EMEA and North America. You can configure BGP filters in the EBGP peering session with the MPLS provider to accept or advertise only routes that are required. You can craft these filtering policies based on BGP community values.
- An existing branch, data center, or hub site can also double as the gateway. Make sure that you dimension the system to handle the expected amount of traffic.
- On gateways and hubs, it is recommended that you always use unique network names for the WAN links, as

illustrated in the following screenshot. Also, the WAN link names must be unique across the entire SD-WAN network as shown in Figure 2 in the Add/Edit Template > Interfaces tab. For more information, see Create Device Templates.



This naming strategy allows remote branches to use the remote circuit option to influence the traffic path in the SD-WAN policy. This behavior is not specific to disjoined underlay networks, but rather it is a general recommendation for gateways and hubs. You can make this selection in the SD-WAN forwarding profile on the remote branch. For more information, see Configure SD-WAN Traffic-Steering Forwarding Profiles.



Interconnect Legacy Networks with SD-WAN Gateways

This sections describes two methods for interconnecting legacy networks with SD-WAN gateways. They differ in how the routes are exchanged with the MPLS provider. The two methods are:

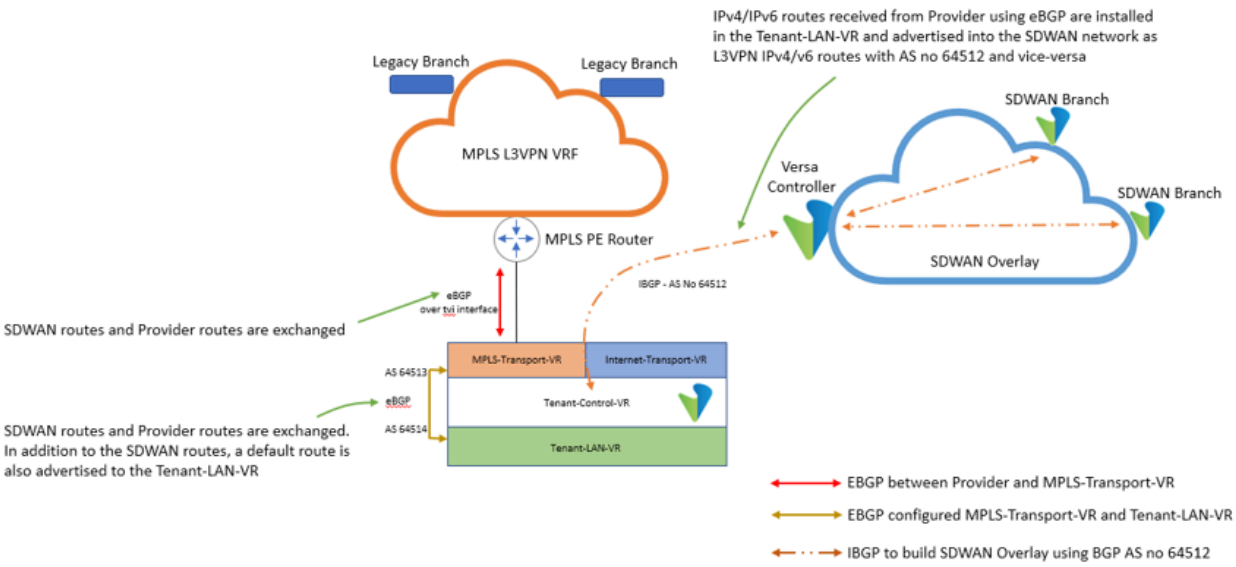
- Use BGP to exchange routes with the MPLS provider on the MPLS WAN interface.
- Establish a BGP session with the MPLS provider on a LAN interface.

Use BGP to Exchange Routes with the MPLS Provider on the MPLS WAN Interface

Using BGP to exchange routes with the MPLS provider on the MPLS WAN interface is the most common method for DIY enterprise SD-WAN deployments. For this method, you configure a BGP peering session on the MPLS-Transport-VR on the gateway to exchange routes from the MPLS Provider to the SD-WAN network. You can use Director Workflows to automate this configuration.

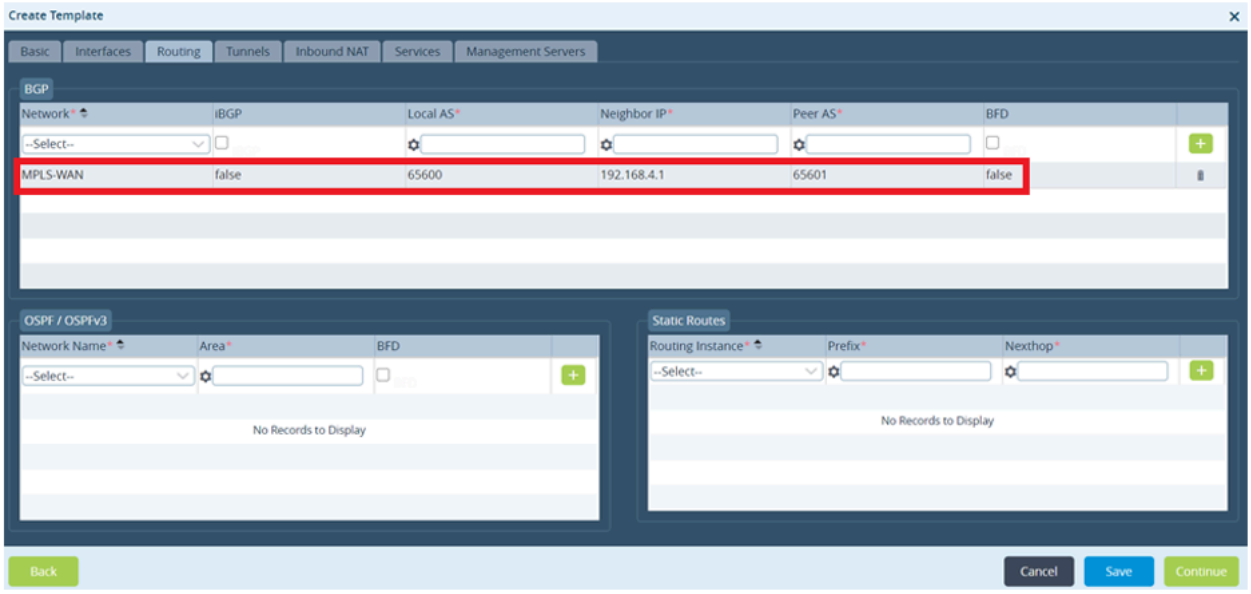
The advantage of this approach is that you need only a single IP interface to connect to the MPLS provider, and this is the interface over which you establish the BGP peering session. Having a singe IP interface is commonly seen in enterprise network designs in which a single MPLS provider offers MPLS VPN underlay services. This MPLS VPN network has connections to the legacy branches and provides MPLS underlay connectivity to SD-WAN branches. Whether an SD-WAN branch has an MPLS or internet underlay, the SD-WAN VPN network uses only the WAN IP addresses to establish connectivity with other SD-WAN branches and the gateway. The SD-WAN branch routes are exchanged over an encrypted overlay.

The following figure shows how the gateway is configured internally. A virtual TVI interface pair is created (using a Director Workflow) in the MPLS-Transport-VR and Tenant-LAN-VR. Over this virtual interface, an EBGP neighbor session is established with AS numbers 64513 and 64514. When the route is advertised to the SD-WAN network, the default BGP AS number of the SD-WAN overlay, 64512, is appended to the BGP AS path list. If the SD-WAN route advertised by one gateway is inadvertently learned by another gateway through the MPLS underlay provider, the route is automatically blocked as a result of the AS path check.

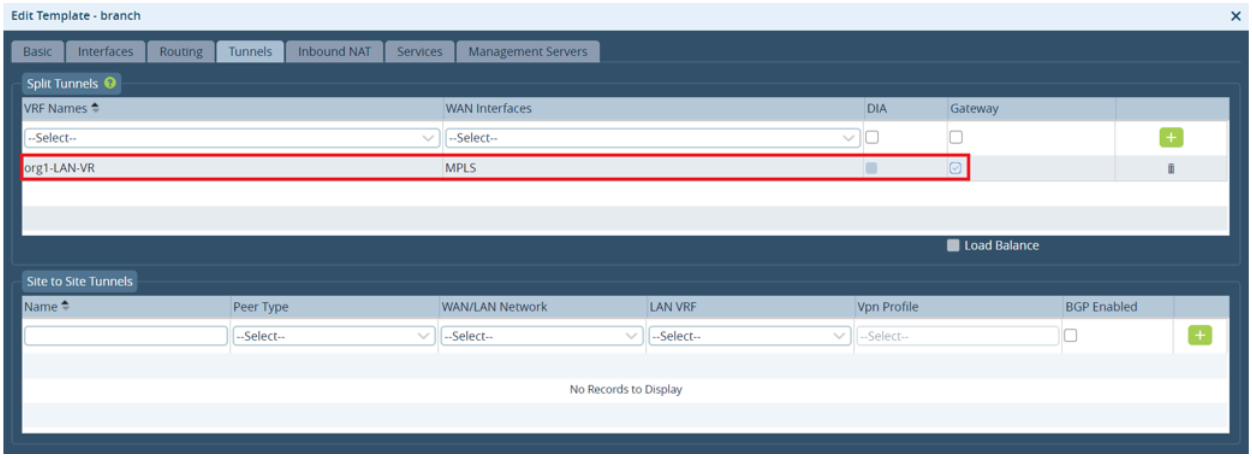


To configure an EBGP session on the MPLS WAN link to the MPLS PE router:

1. In Director view, select the Workflows tab in the top menu bar.
2. Select Template > Templates in the left menu bar.
3. Click the Add icon, or select an existing template. The Create/Edit Template window displays. For more information, see Create Device Templates.



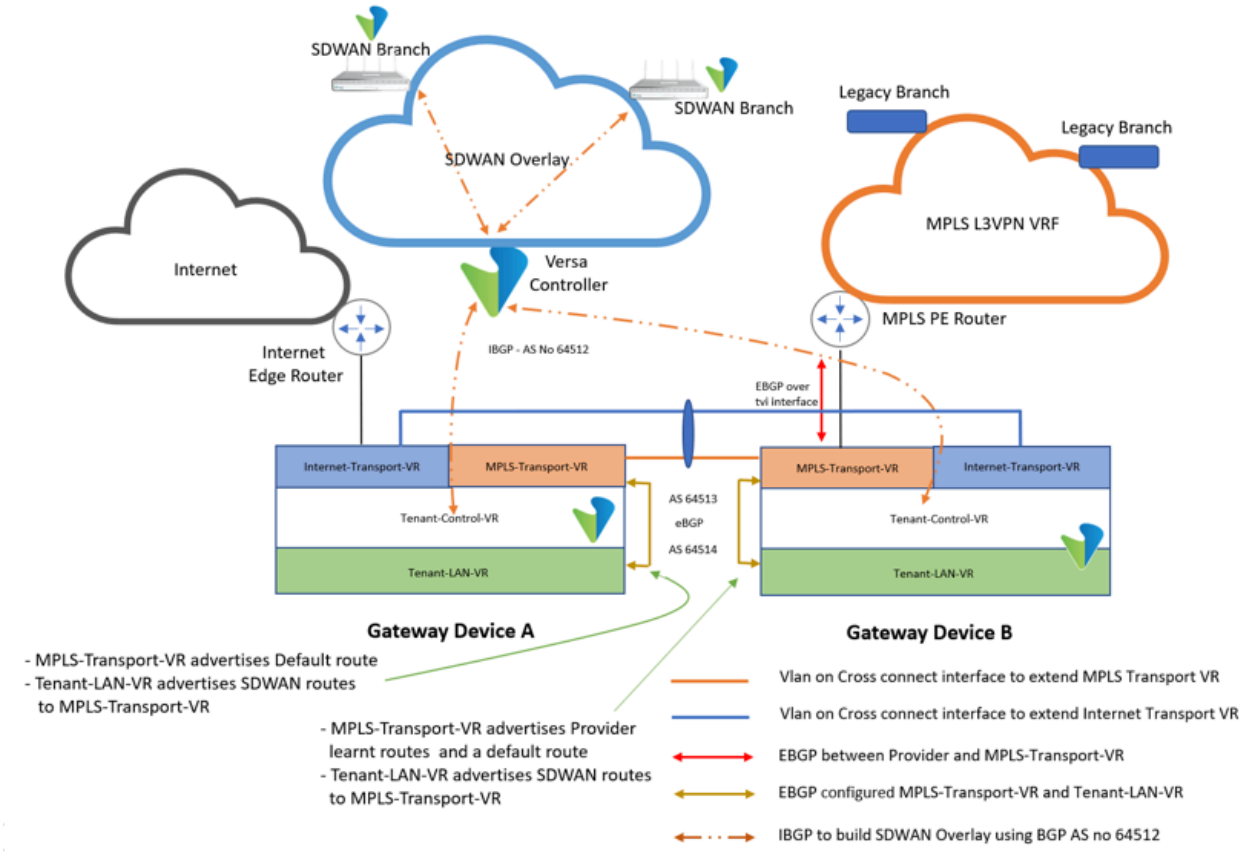
4. Select the Routing tab, and configure the MPLS WAN link.
5. Select the Tunnels tab, and create the BGP session between the MPLS-Transport-VRE and Tenant-LAN-VR over a virtual interface pair (TVI interface) as shown below.



6. Click OK.

High Availability

The following figure illustrates how to achieve high availability (HA) when you use BGP to exchange routes with the MPLS provider on the MPLS WAN interface.



When you configure the branch with active-active (HA) and with the gateway option selected, a TVI interface is created

between the MPLS-Transport-VR and the Tenant-LAN-VR on both devices. Over this TVI interface, an EBGP peering session is established between the ASs 65413 and 65414. A cross-connect cable between the two devices extends the MPLS-Transport-VR to Device A and extends the Internet-Transport-VR to Device B.

For Device B, the figure shows that where the MPLS link terminates, there is an EBGP session with the MPLS provider over the MPLS WAN link. Over this EBGP session, routes are exchanged between the SD-WAN network and the legacy MPLS Layer 3 VPN VRF network. In addition, the MPLS-Transport-VR advertises a default route to the Tenant-LAN-VR. When routes from the MPLS provider are installed in the Tenant-LAN-VR, they are advertised to the SD-WAN network.

Based on the figure above, the following are HA design recommendations:

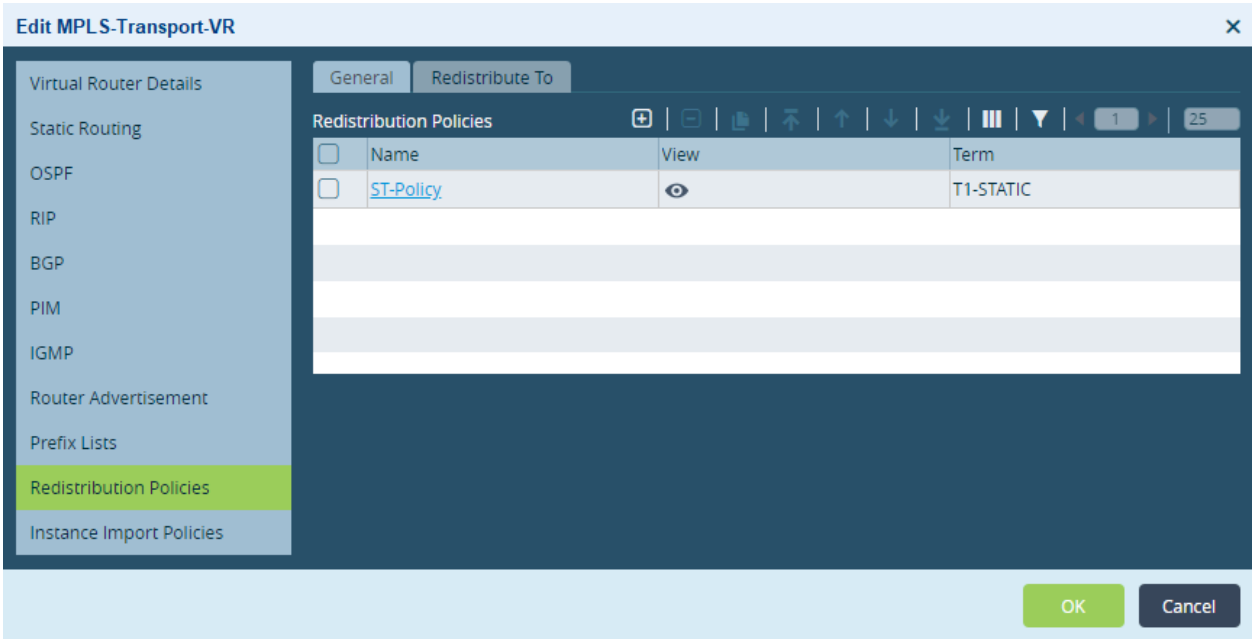
- On Device A, the EBGP session between the MPLS-Transport-VR and the Tenant-LAN-VR advertises only a default route to the Tenant-LAN-VR. The SD-WAN routes from the Tenant-LAN-VR are advertised into the MPLS-Transport-VR. The MPLS-Transport-VR has a static default route pointing to the MPLS-Transport-VR on Device B using the cross-connect interface. Device A learns the routes received from the MPLS provider on Device B using the SD-WAN multiprotocol IBGP session, specifically through the Tenant-Control-VR using a route reflector through the Controller node. The Tenant-LAN-VR on Device A does not receive the MPLS provider routes over the EBGP session from the MPLS-Transport-VR. As a result, Device B is always the primary path for all traffic to and from the MPLS provider. On both Devices A and B, the EBGP session between the MPLS-Transport-VR and the Tenant-LAN-VR advertises the SD-WAN routes to the MPLS-Transport-VR, and it advertises a default route to the Tenant-LAN-VR. If you do not need or want to advertise the default route, you can configure the redistribution policy for the MPLS-Transport-VR on both devices not to advertise the default route. In the example in the Best Practices policy, you configure this in the redistribution policy named ST-Policy.
- It is recommended that you run a routing protocol, such as BGP or OSPF, on the LAN to avoid any issues with asymmetric routing. However, if the LAN router connected to Device B fails, the LAN network is left relying only on the default route. In addition to configuring a default route, you can configure Device A to advertise the specific routes learned from the MPLS provider. To do this, you configure IBGP can be configured between Device A and Device B over the MPLS cross-connect link. For more information see the configuration example in the section Establish a BGP Session with the MPLS Provider on a LAN Interface, below.
- If you use VRRP on the LAN for redundancy, the active device in HA must terminate the MPLS transport link to maintain symmetric traffic.
- If the device is a dedicated gateway with only one internet and MPLS link, configuring the device as HA in Director Workflows is not useful, because if one of the WAN links or devices fail, the gateway becomes non-functional. However, if the device is not a dedicated gateway, that is, if it also functions as a regular site, hub, or data center branch, it makes sense to configure HA using Director Workflows, because the device is usable even if one of the WAN link or devices fails. If the device is a dedicated gateway and if you can replicate the WAN links, it is recommended that, for redundancy, you have twice the number of single gateway devices.
- Because Device A in the figure above does not receive routes from the MPLS provider, it relies on a default route that is advertised into the SD-WAN and LAN networks. In some scenarios, it may be helpful for Device A to also have a copy of the MPLS provider routes so that it does not have to rely on the default route. To do this, you can configure iBGP peering between the MPLS-Transport-VR on both devices over the MPLS cross-connect interface.

Best Practices

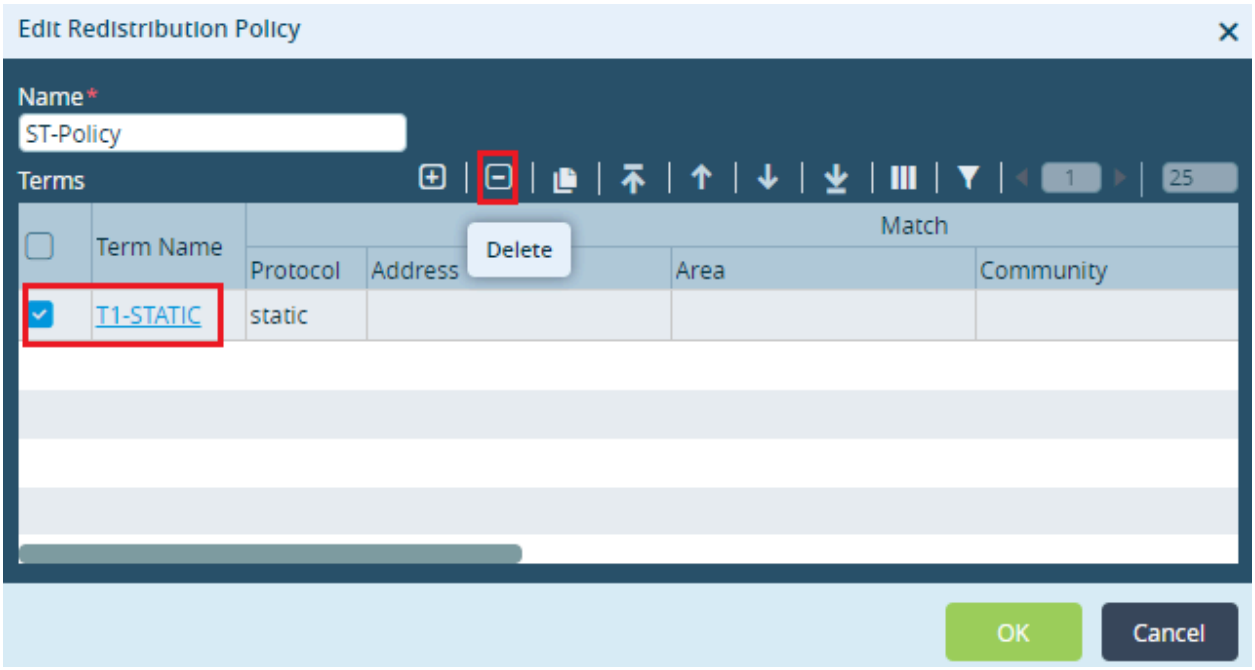
When you use BGP to exchange routes with the MPLS provider on the MPLS WAN interface, a default route is created in the MPLS-Transport-VR that points to the Provider MPLS PE router as the next hop. This default route is advertised to the SD-WAN network over the EBGP session between the MPLS-Transport-VR and Tenant-LAN-VR. If you do not want this default route to be advertised to the SD-WAN network, use a redistribution policy to delete it. For example, to

delete the term T1-STATIC from the Redistribution Policy ST-Policy in the MPLS-Transport-VR:

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 or controller 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 virtual router instance, here, MPLS-Transport-VR.
5. In the Edit VR popup window, select the Redistribution Policies tab.



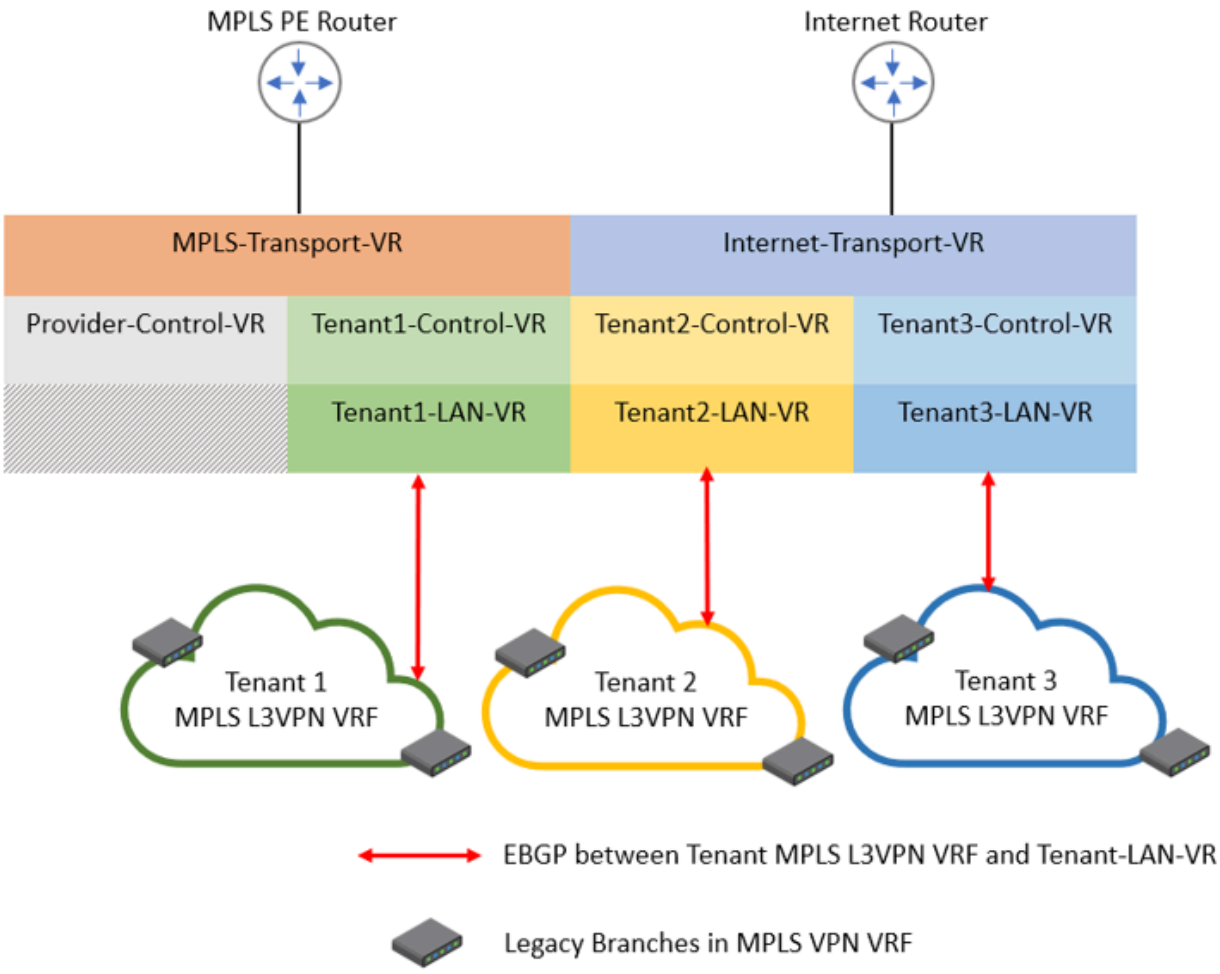
6. Click a term name, here, ST-Policy. The Edit Redistribution Policy window displays.



7. Select the term to delete, here, T1-STATIC, and click the Delete icon.
8. Click OK.

Establish a BGP Session with the MPLS Provider on a LAN Interface

In a scenario in which you establish a BGP session with the MPLS provider on a LAN interface, you use Director Workflows to configure a BGP peering session to exchange routes with the MPLS provider through an IP interface in the Tenant-LAN-VR, as illustrated below.

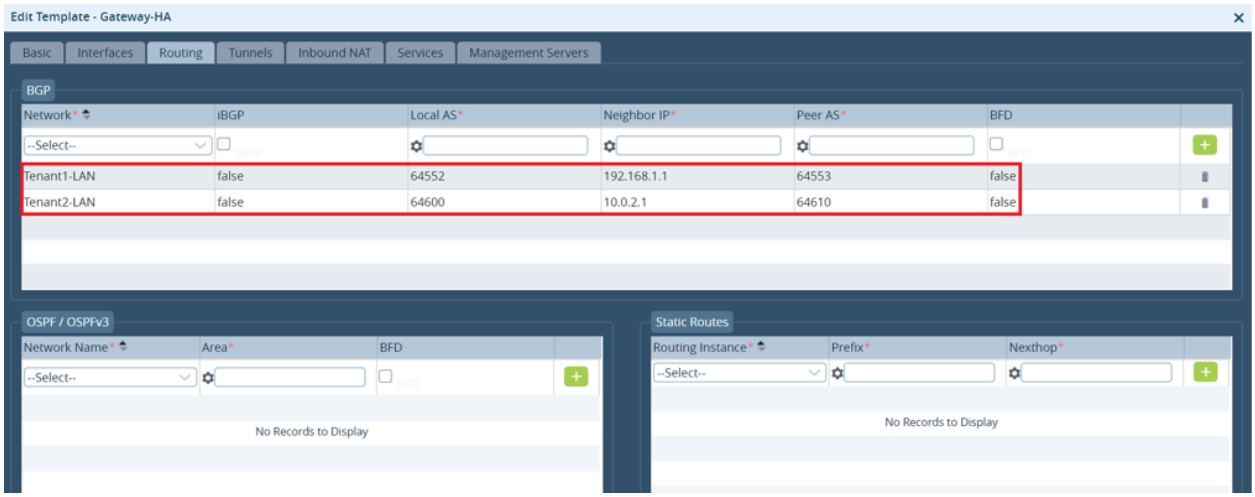


Enterprises hosting their own gateways need two interfaces or logical subinterfaces from the MPLS provider, one that terminates on the MPLS-Transport-VR and that is the underlay to connect to other SD-WAN-enabled sites on MPLS, and a second that terminates in the Tenant-LAN-VR and that is used to connect to non-SD-WAN sites on the MPLS network.

This configuration is commonly seen when a service provider offers a dedicated gateway service their customers and is also the MPLS service provider for their customers. In this scenario, the MPLS provider has one interface or subinterface in the customer VRF that terminates in the customer's Tenant-LAN-VR. The interface that connects to the WAN-Transport-VR is used only to connect to SD-WAN–enabled sites. For any two sites to connect over SD-WAN, only the branch WAN IP connectivity is required. Hence, the peering with the provider on the MPLS-Transport-VR can be a common VRF into which the branch WAN IP addresses of customer VRFs are placed, but the individual customer VRFs have only the WAN IP address of the gateway branch. The EBGP peering session in the Tenant-LAN-VR is with an interface on the provider's MPLS PE router that is part of the customer's VRF in which the legacy branch routes are placed. This design allows each site in the customers SD-WAN network to set up an SLA to the gateway using its underlay IP address, but sites in two different customer networks cannot establish an SLA or any communication path between them. To achieve this, the customer SD-WAN branch MPLS interface WAN IP addresses must be unique or source-NATed.

To configure this type of gateway:

1. In Director view, select the Workflows tab in the top menu bar.
2. Select Template > Templates in the left menu bar.
3. Click the Add icon to create a new template, or select an existing template, as shown here. For more information, see Create Device Templates.
4. Select the Routing tab, and configure the tenant LANs, as shown below:



5. Click OK.

You can modify the routing policy to filter sent and received routes in the device template or directly for the VOS device. In the configuration, in the Appliance view, select Networking in the left menu bar, select the MPLS-Transport-VR, and select BGP. Then select the BGP instance ID and select Peer/Group Policy. For more information, see Configure Peer and Peer Group Policy.

High Availability

You can use Director Workflows to configure gateways with or without HA. You can use BGP path attributes, such as AS-path prepend, local preference, or MED, to select the gateway device that you want to serve as the primary gateway.

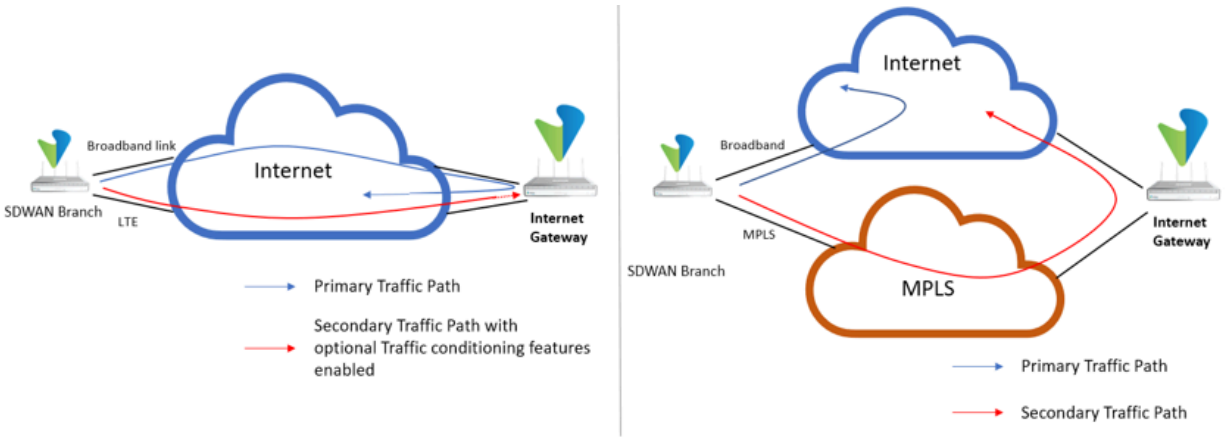
Best Practices

- When you modify the route preference, ensure that you avoid asymmetric routing.
- If you chose a gateway device as the primary, this device must be the primary device for traffic flow in both directions.
- If the device is a dedicated gateway, configuring the device as HA in Director Workflows is not useful, because if one of the WAN links or devices fail, the gateway becomes non-functional. However, if the device is not a dedicated gateway, that is, if it also functions as a regular site, hub, or data center branch, it makes sense to configure HA using Director Workflows, because the device is usable even if one of the WAN link or devices fails. If the device is a dedicated gateway and if you can replicate the WAN links, it is recommended that, for redundancy, you have twice the number of single gateway devices.

Gateway for Internet-Bound Traffic

You can configure a VOS branch device as a gateway in the following scenarios:

- To apply SD-WAN path selection policies on internet-bound traffic with optional traffic conditioning features such as forward error correction (FEC) or packet replication. This configuration can mitigate last-mile WAN link degradation.
- To use an MPLS circuit as secondary path to break out from a gateway if the local internet circuit is unavailable, as shown in the following figure.



Supported Software Information

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

Additional Information

Configure Basic Features
Configure Virtual Routers

SD-WAN Traffic Optimization

For supported software information, click here.

This article discusses how to use traffic steering, traffic conditioning, and SD-WAN path policies to optimize SD-WAN traffic flow.

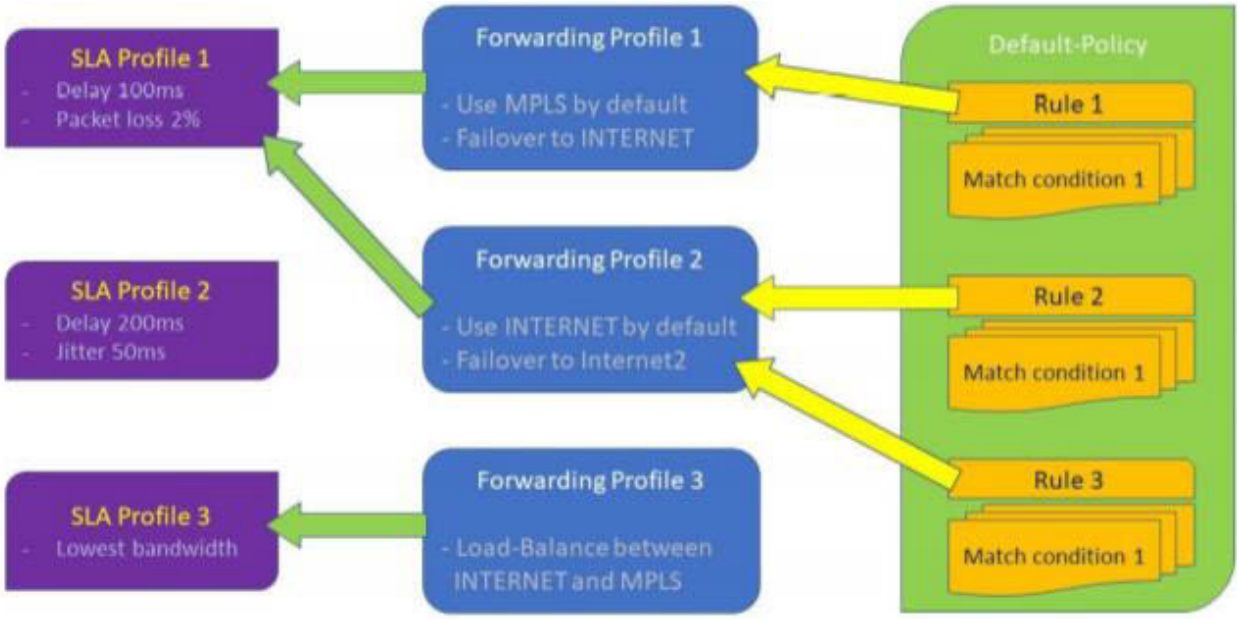
Traffic Steering

An SD-WAN optimizes traffic over available underlays—MPLS, broadband, and LTE—to deliver traffic across the network in an optimal way. Versa Networks VOS edge devices use SLA monitoring to gather performance metrics about

all paths towards all peer branches (assuming a full-mesh topology). The metrics include latency, jitter, and packet loss, and they are used to determine whether a path is up, and if so, whether it meets the user-defined SLAs for an application or a set of applications. The Versa Networks SD-WAN can dynamically steer traffic to the best available link, and if the available links show any transmission issues, the SD-WAN immediately applies remediation for the delay, jitter, and packet loss based on policies, to ensure the continued performance of high-priority applications. The Versa Networks SD-WAN platform provides powerful and flexible policy-based mechanisms for a wide variety of WAN path selection behavior.

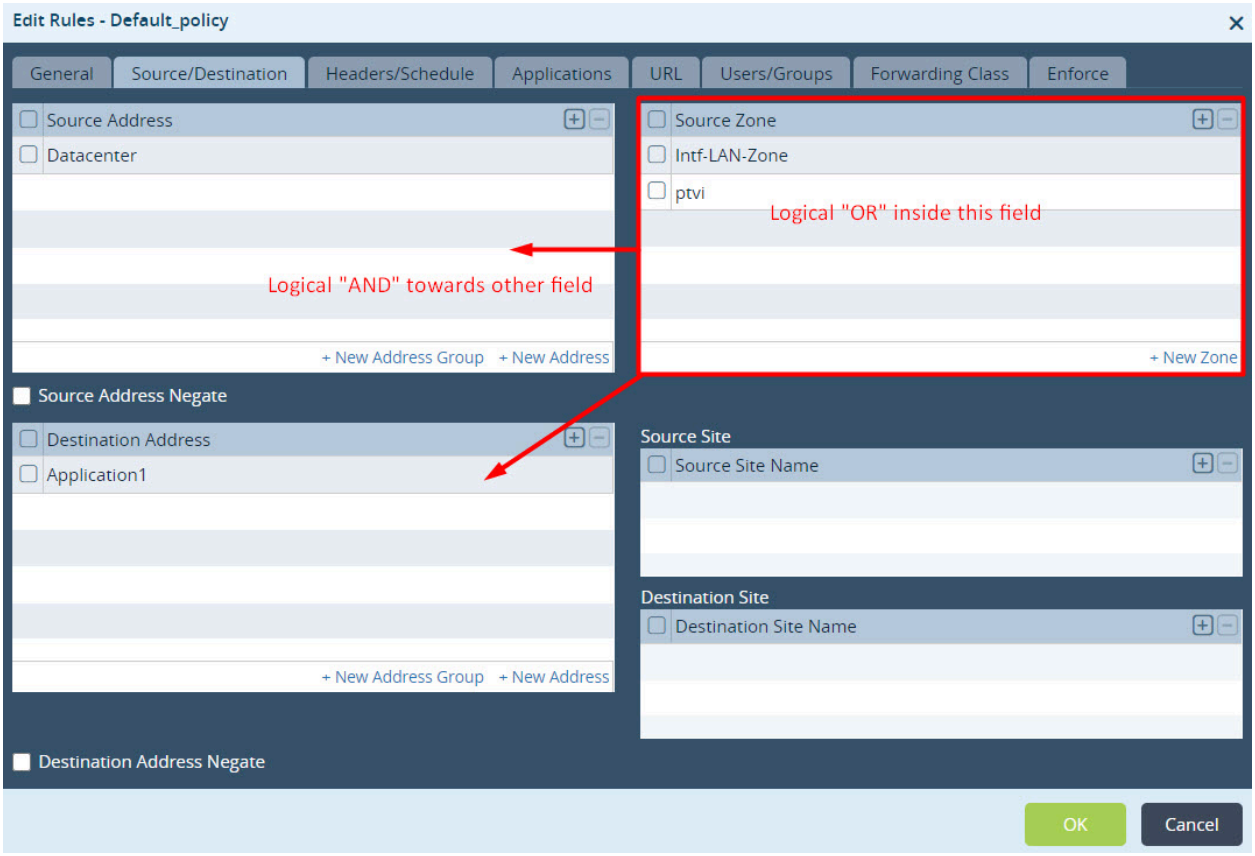
Traffic steering configuration consists of three major components, which are illustrated in the following figure:

- Policy rules—A VOS SD-WAN policy consists of one or more rules. A rule identifies traffic for which you want to specify path selection behavior. Traffic that does not match a specific rule is subject to default behavior. You associate traffic-matching rules with a forwarding profile.
- Forwarding profiles—Forwarding profiles define circuit or path priorities, connection methods, and load-balancing capabilities for traffic that matches the policy associated the forwarding profile.
- SLA profiles—SLA profiles, which are optional and which you associate with a forwarding profile, define application or network thresholds for a path to meet SLA compliance.

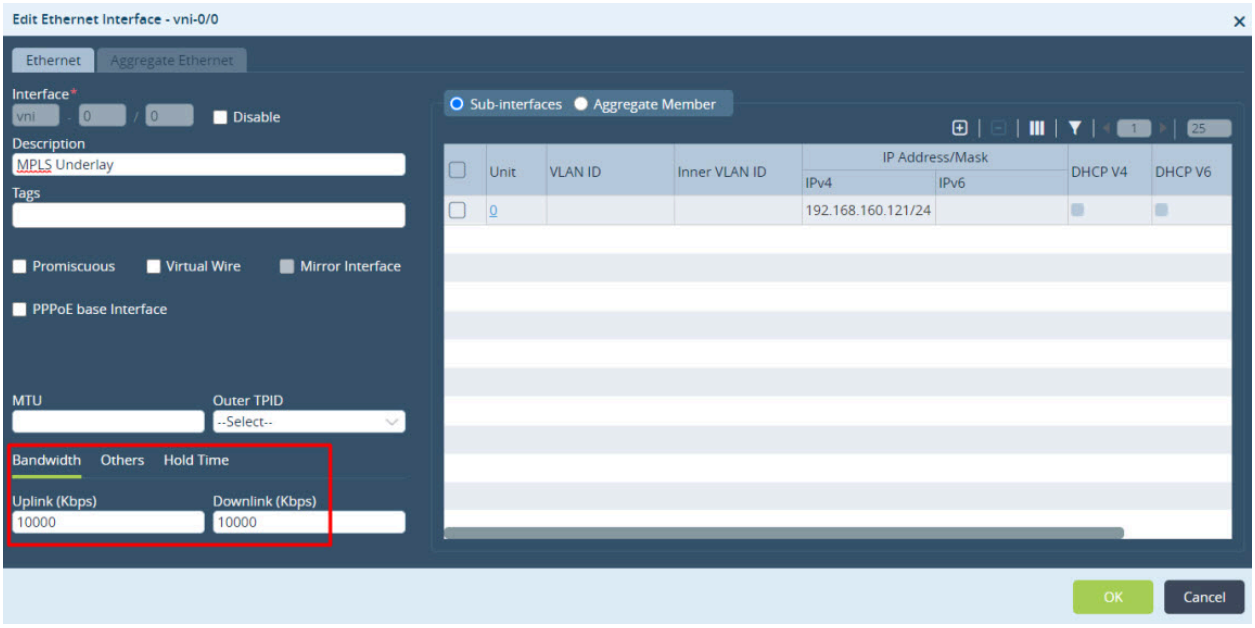


The following are best practices related to SD-WAN traffic-steering policy:

- Ensure that the match conditions in the SD-WAN policy match the correct traffic. For SD-WAN policy, and also for all policy configurations on VOS edge devices, all rule values that you configure on the same GUI tab are processed as a logical OR function, and rule values that you configure on different GUI tabs are processed as a logical AND function. For example, if you include multiple addresses in the source address field, any one of the addresses can fulfill the match criteria for that field. If you include multiple source addresses and if you also configure a source zone, the traffic must match one of the source addresses AND one of the source zone parameters. See the screenshot below. Note that the VOS device supports IPv4 and IPv6 addresses. For more information, see Configure SD-WAN Policy.



- SD-WAN policy rules are evaluated in the order that you include them in the policy. Policy rules are processed starting with the first (top) rule, and the rule evaluation is exited at the first match. So make sure that you configure the matching criteria for critical traffic in one of the initial rules.
- When you want to apply the same forwarding profile to different applications because the applications have similar characteristics because because you want the application traffic to be handled similarly, design and configure the SD-WAN policies so that you can associate the them with the same forwarding profile.
- As a rule of thumb, log only on SLA Violated to reduce the number of logsx. Another possibility is to limit the rate of the log messages.
- Always set the bandwidth on the interface. This parameter is used as the main input for the SD-WAN bandwidth-related profiles.



- To enable load balancing between paths across the WAN circuits, set the WAN circuit priority to the same value:
  - If you want to do load balancing for specific traffic, configure at least two circuits with equal priority.
  - If you specify no circuits, all circuits are assumed to have the same priority.
- With the weighted round-robin (WRR) connection selection method, weights are assigned to each path based on the available bandwidth on the access circuit. It is recommended that you use WRR if paths with equal priorities are provisioned with different bandwidths.
- An elephant flow, which is a single large volume session such as FTP downloads or file copy, presents peculiar challenges for SD-WAN traffic steering. The following are best practices to optimize these flows:
  - Use per-packet load balancing if the traffic is not sensitive to jitter or loss.
  - Enable the Gradual Migration option in the traffic-steering forwarding profile, to prevent a thundering herd of flows when a previously violated path becomes good again.
  - Set up a relatively longer recompute timer, to prevent violating the path because of transient impairments.
  - Enable FEC to protect against loss especially where data integrity, such as in file copying, is important . Configure a "stop when" circuit utilization value for when circuit utilization reaches hits a high watermark value, to prevent FEC overhead from congesting the network.
- Enable reordering globally on all branches in the network.
- The following are best practices when configuring real-time traffic, such as voice and video UDP streams:
  - Enable the Evaluate Continuously option, to allow real-time flows to react better to changing network conditions.
  - Enable the Gradual Migration option in the traffic-steering forwarding profile.
  - Enable replication to mitigate against loss and jitter. Set the start to be SLA Violated and set the stop to Stop When circuit utilization hits a high watermark value (75% is recommended) to prevent replication overhead from congesting the network.
- The recomputation timer determines the number of SLA reports to access before a violation decision is made. It is recommended that you not shorten the default recompute timer, which is 300 seconds. (However, it might be convenient to lower the value during lab testing.) Rather, enable the Evaluate Continuously option in the forwarding profile. However, you can set a longer recompute timer for relatively high-quality underlay networks, as well as for

flows that can survive transient network conditions.

- SLA smoothing and SLA violation damping are disabled by default, and if you choose to enable them, you should do so only after an in-depth analysis. If the circuits or paths do not often most between the compliant and noncompliant states, you should use the default configuration
- If you enable SLA smoothing, configure the recomputation timer to a value that is less than the SLA smoothing interval.

Traffic Conditioning

Real-time applications have more stringent delay requirements than normal data transmissions. As a result, retransmission of lost packets is generally not a valid option for these types applications. In these cases, better methods to use when attempting to recover information from packet loss are packet replication and forward error correction (FEC).

Packet Replication

Packet replication is the most effective option for applications that are sensitive to latency and packet loss. Packet replication, which is primarily used for real-time applications, takes an identified type of application traffic and sends a copy of each packet across multiple paths. This mechanism prevents inherent latency anomalies and packet losses from negatively impacting the applications.

One drawback of packet replication is that it can have a significant impact on the bandwidth utilization of WAN circuits. Therefore, you should use this feature mainly for low-volume UDP traffic that is sensitive to packet loss, such as VoIP traffic. Also, it is recommended that you configure both Start When SLA Violated and Stop When options in the forwarding profile that you associate with packet replication.

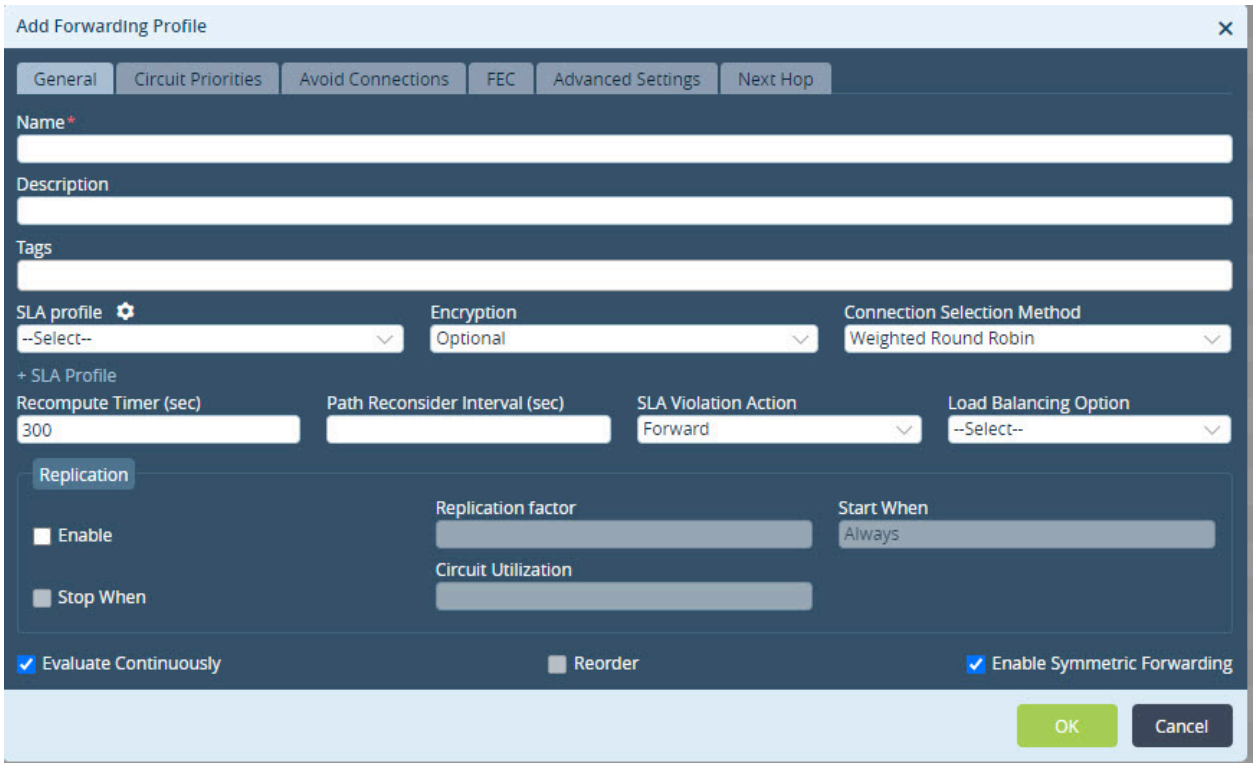
Packet replication works as follows: If a transmitted packet is lost in transit, a copy of the packet is forwarded to the receiver. The receiving branch discards the copies of the packet and forwards only one packet to the receiver. You must enable reordering on all branches to ensure that the receiving branch reorders packets before forwarding them to the receiver.

Packet replication is suitable for branches with multiple SD-WAN paths between branches, while FEC is also effective on a branch that has only a single access links. Packet replication works well when the amount of critical traffic that is being duplicated across the networks is far less than the capacity of the network.

By default, packet replication is disabled. To enable replication, you configure it when you configure an SD-WAN traffic-steering forwarding profile:

1. In Director view:
    a. Select the Configuration tab in the top menu bar.
    b. Select Templates > Device Templates in the horizontal menu bar.
    c. Select an organization in the left menu bar.
    d. Select a template in the main pane. The view changes to Appliance view.

2. Select the Configuration tab in the top menu bar.
3. Select Services > SD-WAN > Forwarding Profiles in the left menu bar.
4. Click the Add icon. The Add Forwarding Profile popup window displays. Enter the following information in the Replication group of fields. For more information, see Configure Replication for SD-WAN Traffic Steering.



5. To enable packet replication, click Enable. By default, replication is disabled.
6. In the Replication Factor field, enter the number of egress packets to send for each ingress packet. For example, if you configure a replication factor of 2, for each ingress packet, two egress packets (the original and one copy) are forwarded to the next hop. The default replication factor is 2. If there are more than two paths between branches, you can increase the value.
7. In the Start When field, select SLA Violation, to enable packet replication only when all available paths have violated the SLA. Otherwise, replication is enabled on all flows that use the forwarding profile. It is recommended that when you enable packet replication, you always select the SLA Violation option.
8. Optionally, click Stop When to set a circuit utilization threshold value to use to stop packet replication and thus prevent oversubscription of links. When the utilization of any of the links used for replication reaches the percentage set in the Circuit Utilization option, replication stops, and the flow passes using only one available link, as was the case before you enabled replication.
9. When you click Stop When, in the Circuit Utilization field, enter the circuit utilization threshold at which replication stops automatically if all the device's WAN circuits exceed this threshold. Specify this as a percentage of the total circuit bandwidth. This threshold applies to all circuits, even those that are set as to be avoided.
10. Click OK.

The following are best practices for packet replication:

- Use packet replication only for specific traffic, such as VoIP. Avoid enabling replication on wildcard traffic.
- Enable replication only when Start When is SLA Violated, and always set the Stop When and Circuit Utilization. For the circuit utilization calculations to be accurate, ensure that the bandwidth values are set correctly on the interface.
- A replication factor of 2 is adequate for most cases. However, you can increase the value to match the available paths for the flows that are protected. For example, four paths can benefit from a replication factor of 3.
- Bandwidth utilization increases by 100% for each unit increment in the replication factor. That is, a replication factor of 2 equates to a 100% increase in bandwidth, a factor of 3 equates to a 200% increase, and so on.
- Enable replication for real-time traffic and other business-critical traffic for which jitter and loss are important.

Forward Error Correction

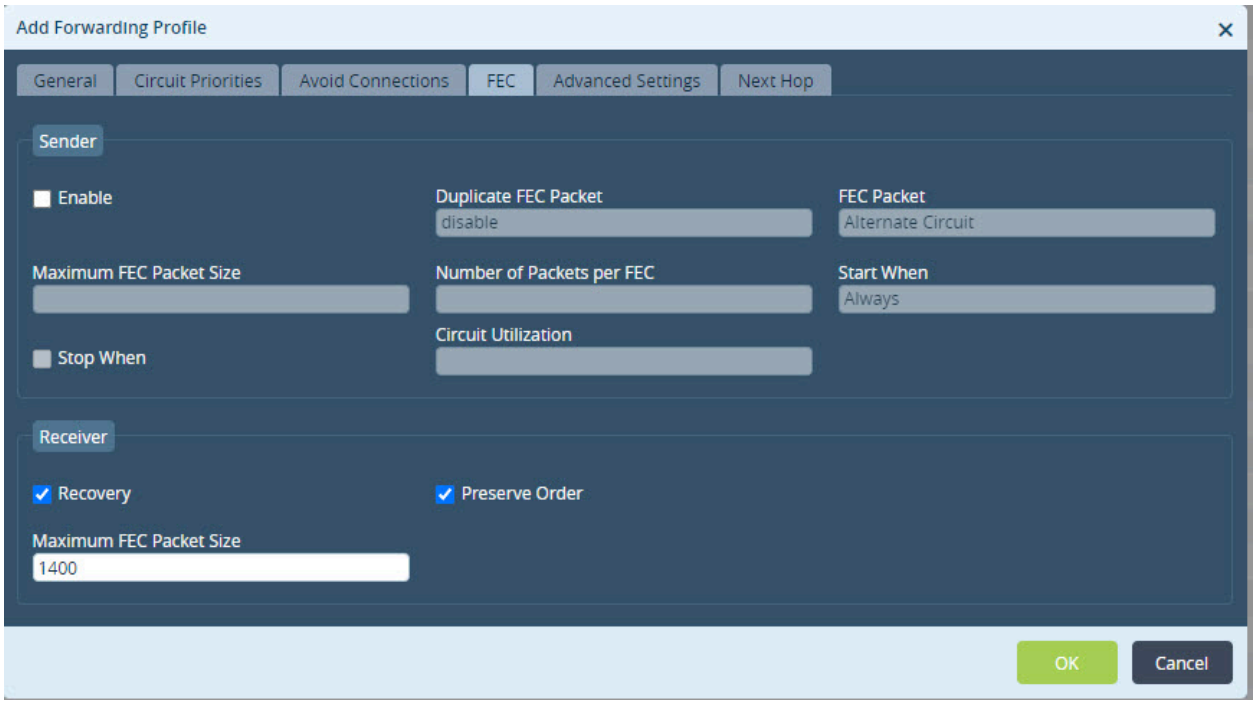
FEC is a mechanism to correct bit errors at the physical layer. While FEC has traditionally been used for this purpose, it has been adapted so that it can be used to recover from packet loss at the network level.

Packet-level FEC works by adding an additional loss recovery packet for every *n*th packet that is sent. This additional loss recovery packet enables a VOS edge device to reconstitute lost packets at the far end of a WAN link, before the packets are delivered to TCP or other transport layers. This mechanism avoids transport layer retransmission and, in the case of TCP, prevents the TCP congestion avoidance mechanism from lowering the throughput available to the application. For the modest overhead of an additional loss recovery packet, FEC reduces packet loss dramatically, enabling applications to benefit from the maximum throughput that the WAN link can support.

You can turn on FEC along with replication at the sites that have multiple paths, to provide maximum protection and correction. You can also use FEC to recover packets independently when replication might not be useful, for example, at sites with a single path for transport traffic.

To configure FEC:

1. In Appliance view, select the Configuration tab in the top menu bar.
2. Select Services > SD-WAN > Forwarding Profiles in the left menu bar.
3. Click the Add icon. The Add Forwarding Profile popup window displays. Enter the following information in the FEC tab. For more information, see Configure SD-WAN Traffic-Steering.



4. For FEC on the sender, configure the following options in the Sender group of fields:
   a. To enable FEC, click Enable.
   b. In the Duplicate FEC Packet field, if extra protection is required, additional FEC parity data packet can be sent. The FEC parity packets could be transmitted over the same WAN link or over another available WAN circuit. By default, duplicate FEC packets are not sent.
   c. In the FEC Packet field, select the circuit on which to send the FEC packets. The default if the alternate circuit. If you configure Duplicate FEC Packets, select the same circuit that you use for the duplicate packets.
   d. In the Maximum FEC Packet Size field, enter the maximum size of the data in the packet to protect. For example, voice packets are typically less than 100 bytes, so set the maximum size to 100bytes to protect the traffic.
   e. In the Number of Packets per FEC field, nter the number of data packets after which an FECe packet is generated and sent to the peer branch. The default value is 4, which means the system generates FEC parity data after each four packets of actual data transmitted according to this forwarding profile.
   f. In the Start When field, if you do not select any value, the forwarding profile always protects traffic for the defined rule. If you select SLA Violation, the FEC is enabled only when all available paths have a violated SLA.
   g. Click the Stop When field to optionally prevent oversubscription of the links. When the utilization of the link over the FEC data is sent reaches the percentage configured in the Circuit Utilization field, the FEC operations stop.
5. For FEC on the receiver, configure the following options in the Receiver group of fields:
   a. Click the Recovery field to enable packet recovery after the receiver receives FEC packets. By default, receiver packet recovery is enabled.
   b. Click the Preserve Order to reorder out-of-order packets and forward them in their original order. By default, reordering is enabled. Note that this option is independent from Reorder option in the Forwarding Profile.
6. Click OK.

There are multiple ways of configuring Layer 3 FEC, and the following are some recommendations:

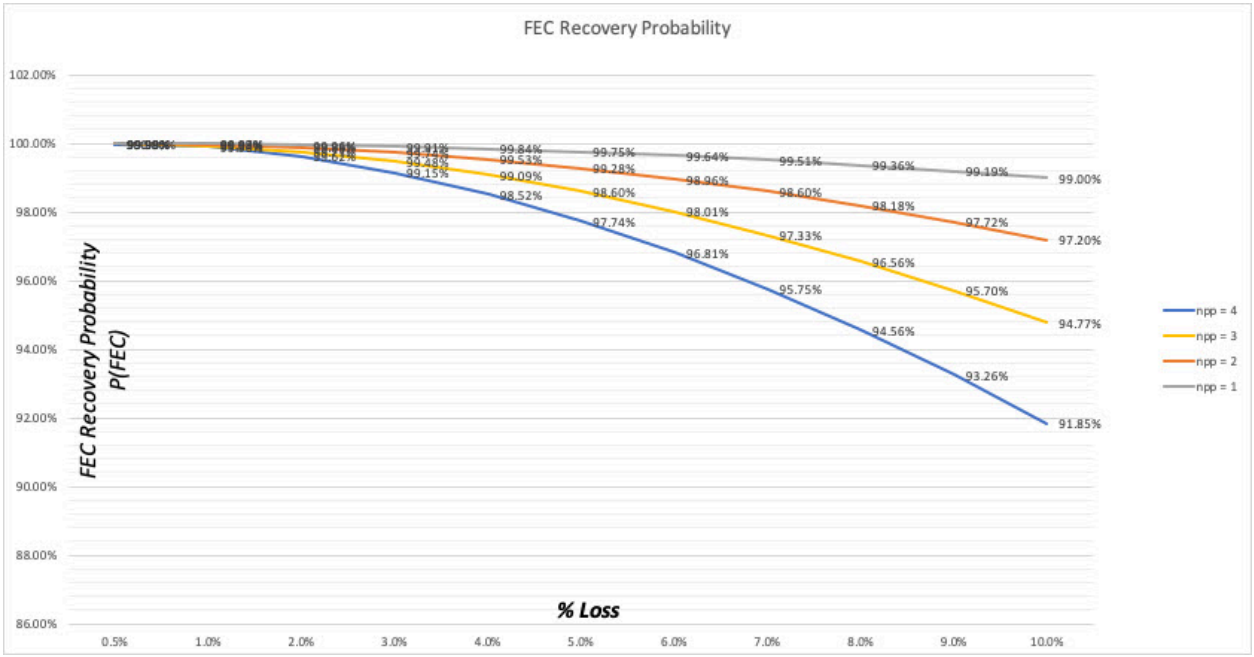
- WAN circuit utilization—When bandwidth is at a premium, FEC is more efficient than packet replication, and it generates only one parity packet for a specified number of data-carrying packets (range 1 through 32, default 4). The generated FEC parity packet can recover a packet on the peer branch only if one packet is lost in the specified number of packets per FEC.
- If packet corruption on a path has a specific pattern (corrupting the *n*th recovery packet and its replica), it is recommended that you replicate the FEC packet on an alternate circuit so that non-corrupted FEC packets can reach the remote site.
- Small versus large packets—Layer 3 FEC works well on flows of small packets, such as voice or point-of-sale transactions. Applications that have large packets, such as video and file transfers, are usually not business-critical applications, and performing forward error correction on fragmented packets is difficult. It is recommended that you use FEC for packets less than 500 bytes.
- LTE and FEC—In most cases, you should not use FEC on LTE interfaces, because FEC increases the amount of traffic on a network that is already limited in capacity. When packets are dropped in a wireless network, it is usually many packets in a row, and FEC is not useful in such a scenario.
- Adaptive codecs versus FEC—Many of the latest voice and video codecs support FEC within the application codec. Usually, these adaptive codecs are more efficient than Layer 3 FEC. For these cases, you must perform tests to find out whether the Layer 3 FEC feature provides any advantage if you use it on top of the voice or video codec FEC feature.
- Versa implements single-dimensional FEC, which is most effective against bit error and single packet loss within a protected block. FEC improves the probability that a stream arrives intact at a receiver. You can calculate the probability with the following formula:

  $$P(FEC = (1 - p_x)^{w+1} = ((w + 1)(1 - p_x)^{w})p_x$$

  Where:

  - $P(FEC)$ is the FEC recovery probability, as a percentage
  - p_x is the packet loss, as a percentage
  - w is the number of packets per parity (npp)

  This formula indicates that the FEC default settings of 4 (npp) is effective only up about a 5 percent packet loss, as illustrated with the following graph. You should consider replication for higher loss percentages.



- FEC performance increases when the npp is closer to 1. There is, however, a 25 percent increase in the bandwidth utilization value for every unit decrease in *x* – 1, where *x* = npp. For example, if the default npp is 4, reducing it to 3 increases bandwidth utilization by 25 percent. When npp = 1, bandwidth utilization increases by 100 percent.

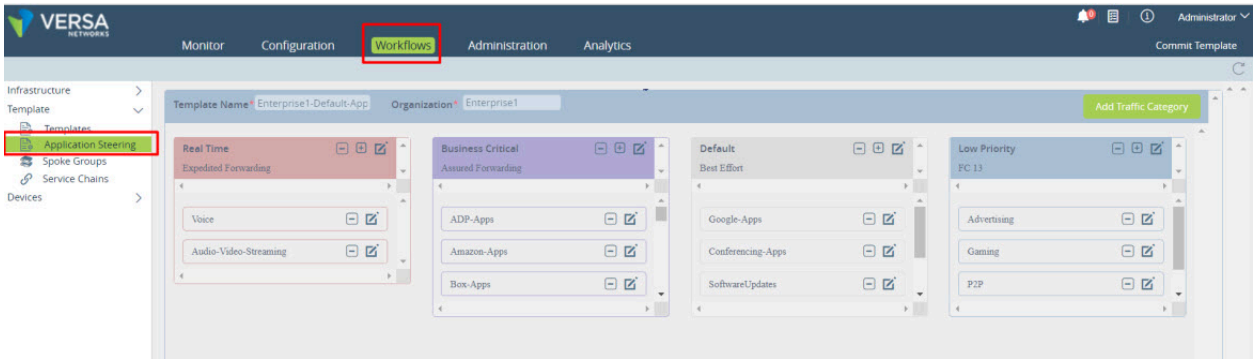
Business-Intent Traffic Steering with Application-Steering Templates

Application-steering templates automate business-intent policies and combine all business applications and network characteristics into a single configuration template. These characteristics include classification, mapping into forwarding classes (QoS), and SD-WAN policies.

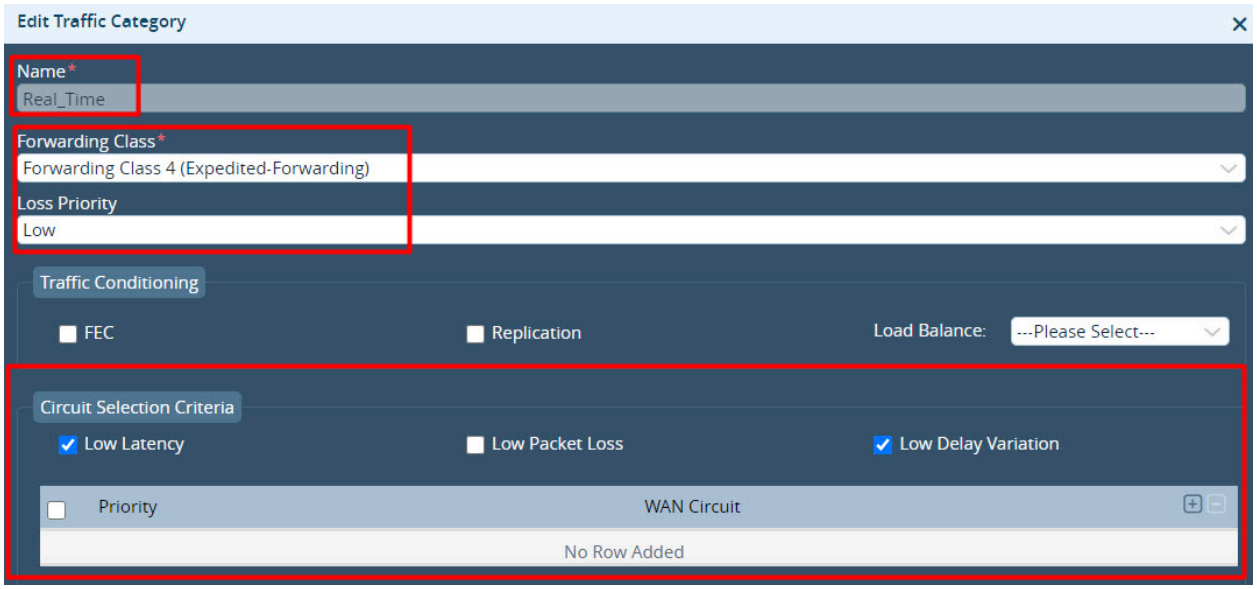
Application-steering templates do not add new SD-WAN functionality, but rather they simplify management of business applications in one template. These templates do not configure other QoS parameters other than forwarding class assignment and loss priority, so you still have to define and apply QoS policies and rules to deliver predictable application performance during congestion.

After the initial deployment is complete, you can modify SD-WAN application-steering templates for detailed configurations. The application-steering workflow is retained and is used in similar ways as other Workflow templates, to create complex configuration modifications. For more information, see Create Application-Steering Templates.

A default application-steering template is preconfigured for each organization, with four traffic categories—Real Time, Business Critical, Default and Low Priority, as shown in the screenshot below These categories group applications or family of applications under a particular traffic category. For example, voice applications are grouped under Real Time, and Office 365 applications are grouped under Business Critical. Note that the default application template is an example template that is designed to allow the administrator a quick start in writing application business intents. You can evaluate it and then modify it as needed.



For each traffic category, default criteria are applied. For example, for the Real Time category, a default configuration for Forwarding Class and Circuit Selection Criteria is applied, as shown in the following screenshot. For more information, see Add a Traffic Category.



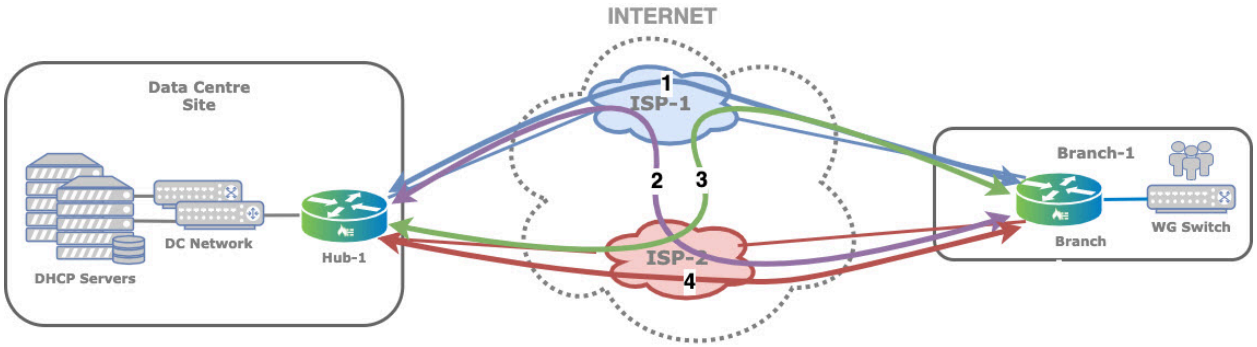
You can modify the values in Forwarding Class and Loss Priority fields and set traffic conditioning parameters (such as FEC and replication), load-balancing algorithm, and dynamic or static circuit path selection criteria that apply to the applications configured for the traffic category. You can create custom traffic categories or modify predefined categories.

The following are best practices for application-steering templates:

- Use application-steering templates to automate the creation of SD-WAN policies. You can customize the service template to suit your requirements.
- Application-steering templates create CoS classifiers and ensure that schedulers are added to the appropriate interfaces.

SD-WAN Path Policies

BOS branches continuously monitor the performance of all paths towards all SD-WAN peer branches and Controller nodes. A path is defined as any valid transport tunnel between the two branches. For example, if two branches have two broadband links each, and each of them is in a single transport domain, there are four paths between the branches, as illustrated in the following. This applies to the paths between branches and Controller nodes as well but is not represented in Figure 9.

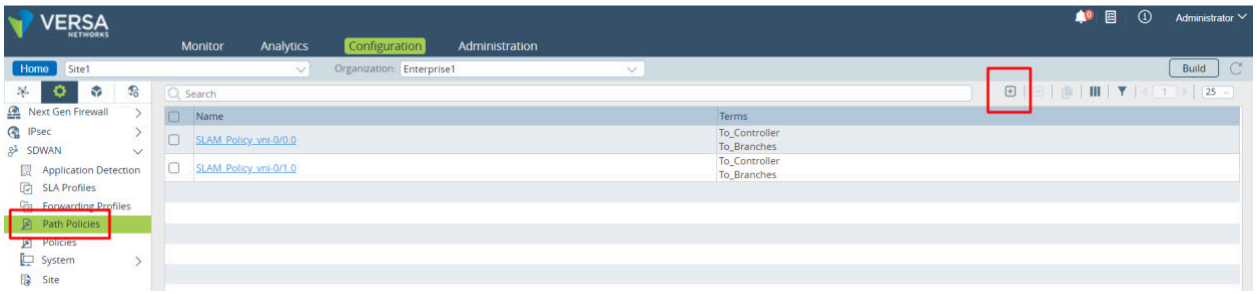


Each path is monitored by sending request-response style SLA probes at a configured interval. Because a network may impose differentiated treatment for different forwarding classes, SLA probes are sent for each forwarding class. The metrics computed are delay, and forward and reverse delay variation and loss (statistical and actual traffic loss in forward and reverse directions).

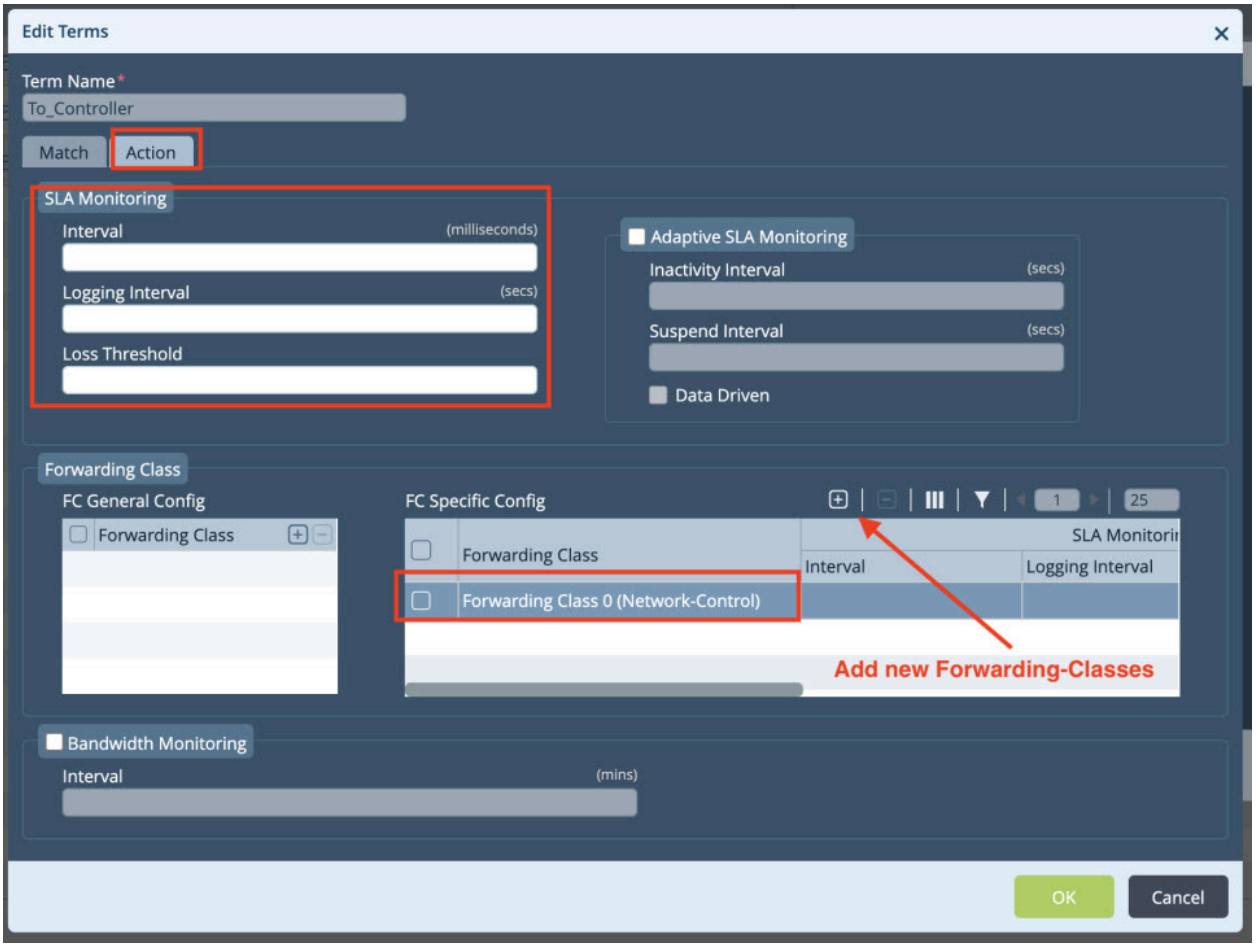
You can configure SLA monitoring (SLAM) for all 16 forwarding classes (network control through FC 15) using SLA path policies. These SLAM path policies are created, by default, on all WAN interfaces using the EF forwarding class at 2-seconds intervals towards remote branches, and using the NC forwarding class at 10-second intervals towards Controller nodes. You can modify these polices or create new one to suit your requirements. Each SLAM path policy has a match condition on the remote site type, that is, whether it is a Controller node or a branch.

To configure a path policy:

1. In Director view:
   a. Select the Configuration tab in the top menu bar.
   b. Select Templates > Device Templates in the left menu bar.
   c. Select an organization in the left menu bar.
   d. Select a post-staging template in the main pane. The view changes to Appliance view.
2. Select the Configuration tab in the top menu bar.
3. Select Services > SD-WAN > Path Policies. The main pane displays a list of the path policies that are already configured.



4. Click the Add icon. The Add Path Policy popup window displays. Select the Action tab and enter the following information fields. For more information, see Configure SD-WAN Path Policies.



The following are best practices for path policies:

- The default SLA path policy is generally sufficient for most use cases.
- If you want to configure additional SLA probes, it is recommended that you not have more than one SLA probe for each traffic class.
- For granularity, you can use four probes for the four traffic classes—Network Control, Expedited Forwarding, Assured Forwarding, and Best-Effort. This modification is required only for a branch-to branch-profile. Branch-to-

Controller profiles require probes only for the Network Control class.

- On low-bandwidth circuits, configuring SLA monitoring for several forwarding classes with aggressive timer can use a considerable portion of the available bandwidth. A hub-and-spoke topology is recommended to minimize this overhead, rather than having only SLA probe to the hub.
- You can enable SLA optimization techniques such as adaptive SLA and data-driven SLA to reduce the SLA probe load, especially on low-bandwidth links.
- Gradually introduce topology changes to further optimize the network to prevent a full mesh of SLA probes and the subsequent SLA load. For more information, see SD-WAN Topologies.

Supported Software Information

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

Additional Information

Configure SD-WAN Path Policies
Configure SD-WAN Traffic Steering
Create Application-Steering Templates
SD-WAN Topologies

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.

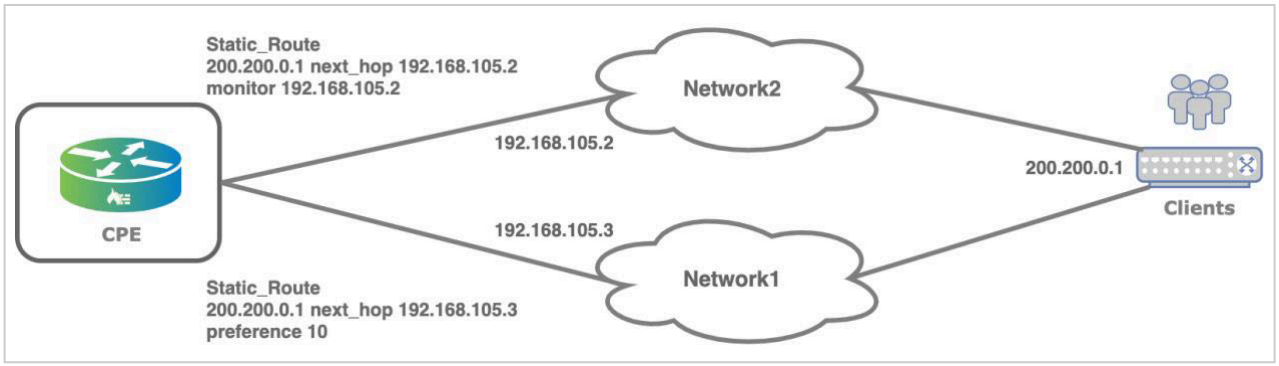
Static Routing

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.

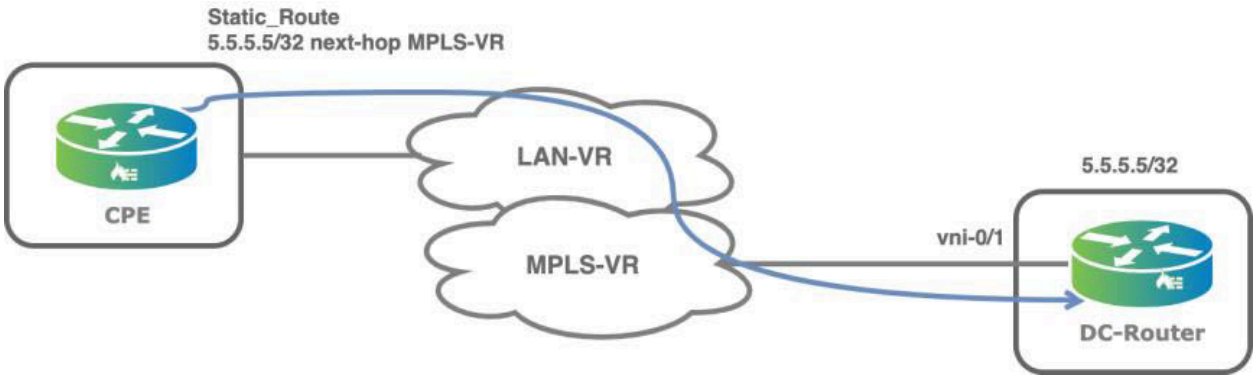


The following output shows a scenario in which the next-hop IP address is monitored on the primary link, while the second static route is configured with a lower preference. The output shows that the primary static route is withdrawn because the monitor probe fails. Monitor objects can be based DNS,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

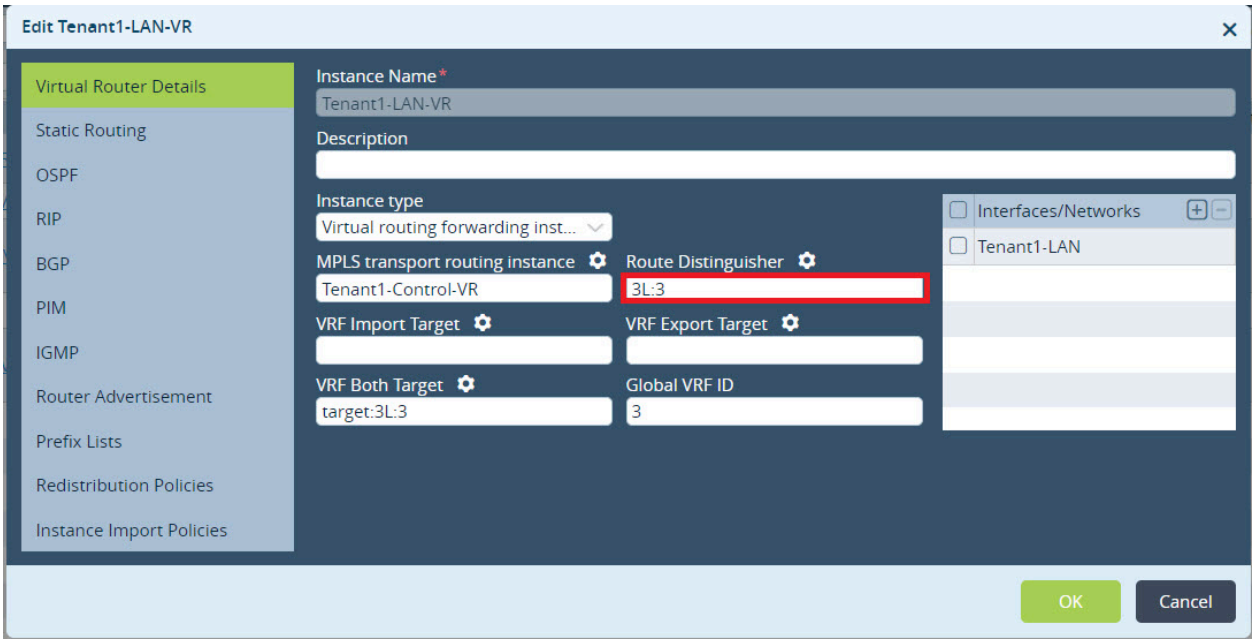
Fast Convergence on the SD-WAN Side

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.

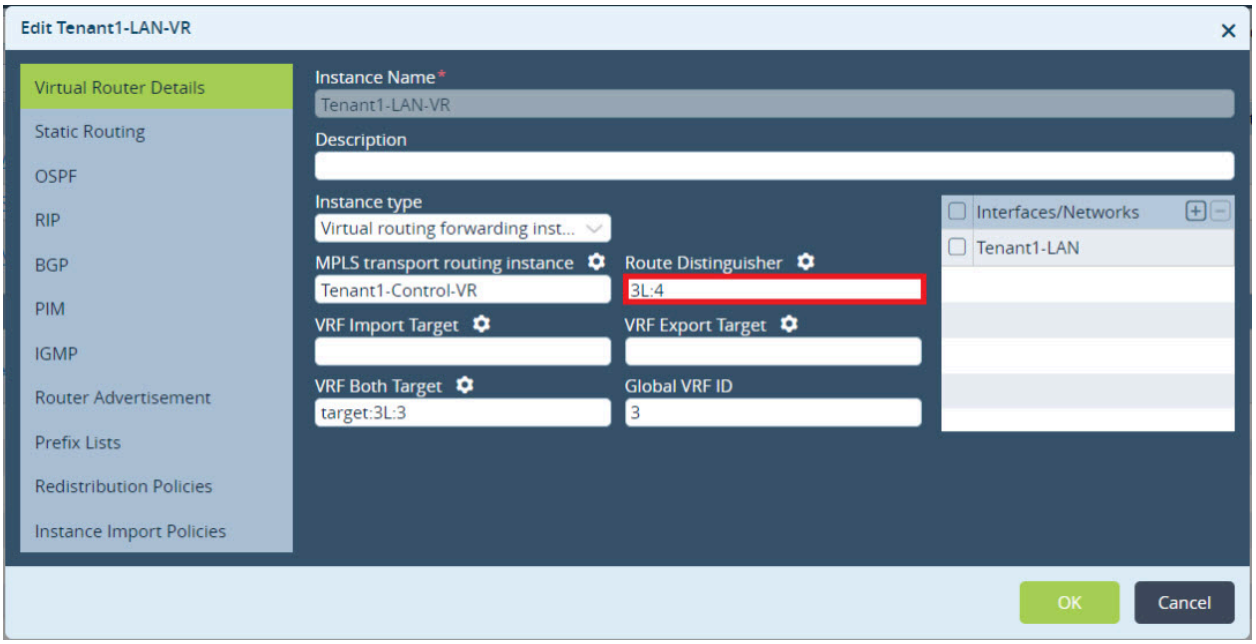
Note: In an EVPN multihoming context, the BGP EVPN route distinguisher should be different for VTEP endpoints that have the same ESI values.

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.



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



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:

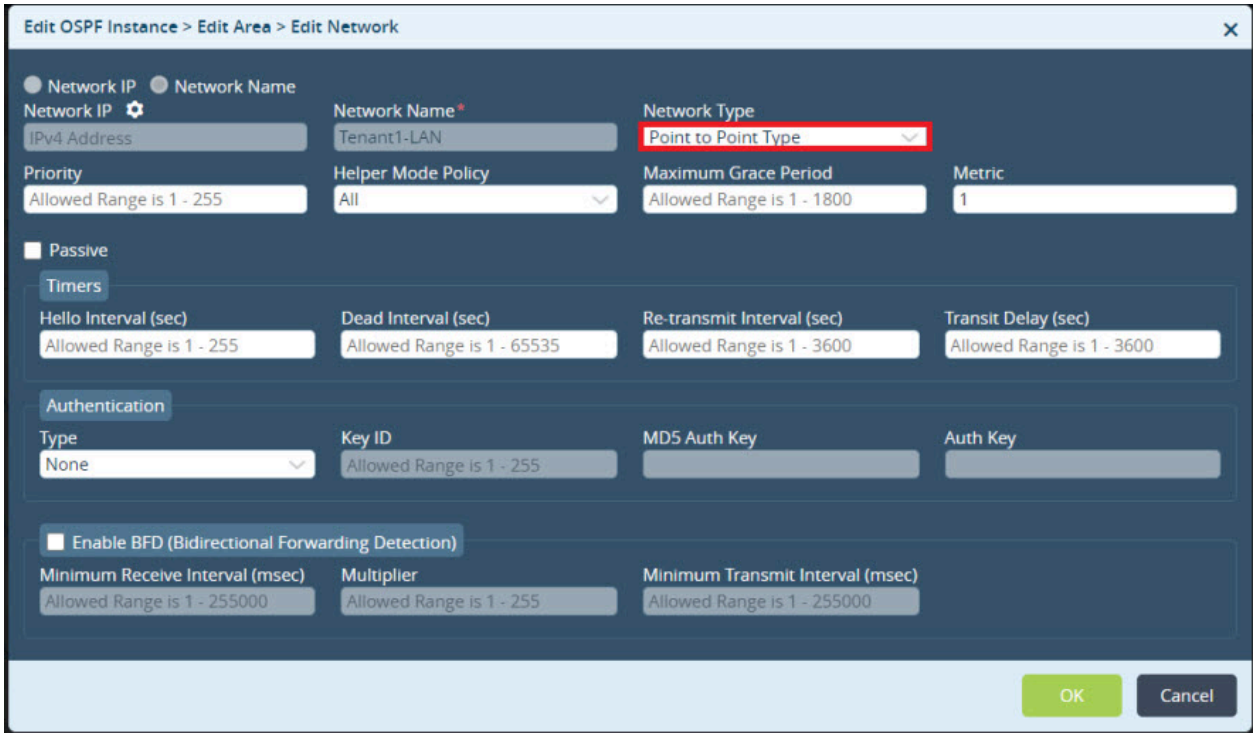
```
admin@branch8-cli> show ospf neighbor brief
State codes: atmpt - attempt, exchg - exchange, exst - exchange start,
             load - loading, 2-way - two-way, full - full
Op codes:    gdown - going down, gup - going up

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.



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.

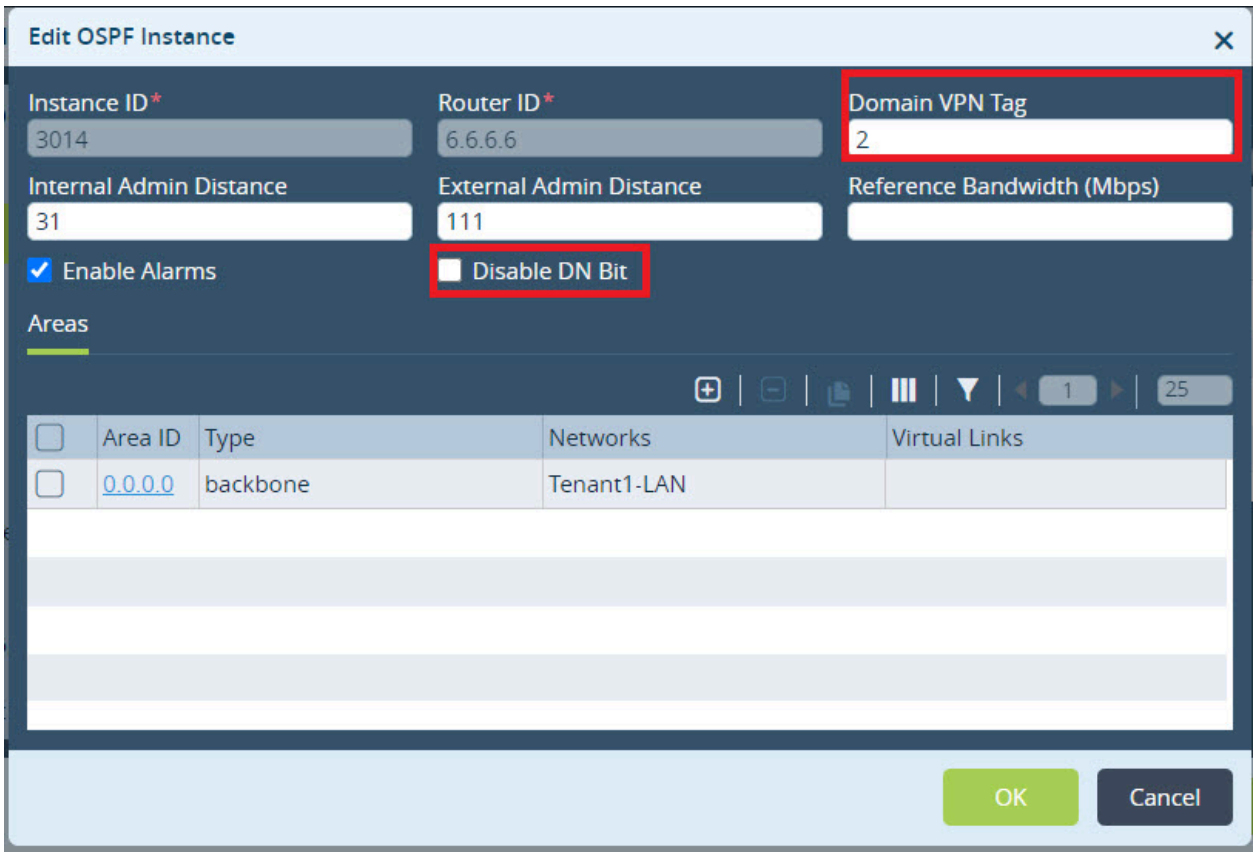
The following sample output shows the outbound BGP policy that is applied for the peer group and hence to all peers in the group:

```
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.

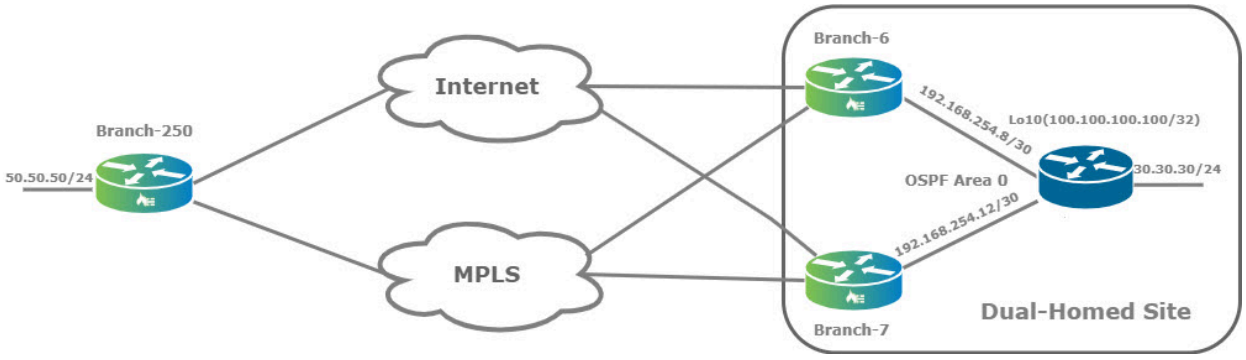


Let's talk more about the Down bit and the domain tag to understand how they work in regards to loop prevention. When

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
```

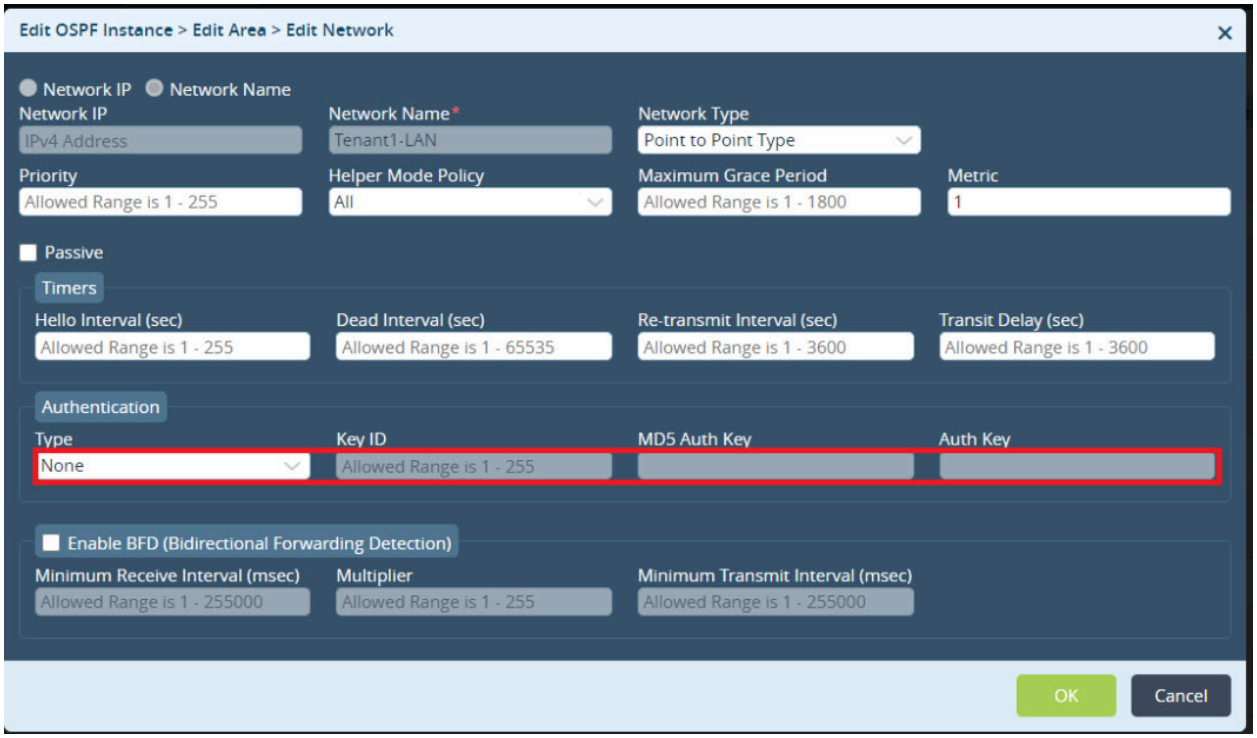
The recommended solution is to check whether the intermediary router (here, IOU1) has the capability to ignore the Down bit set in the LSA option field and can instead use the LSA in the SPF algorithm, so that it can add the prefix to its route table. Most vendors call this feature VRF-lite. If the existing 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

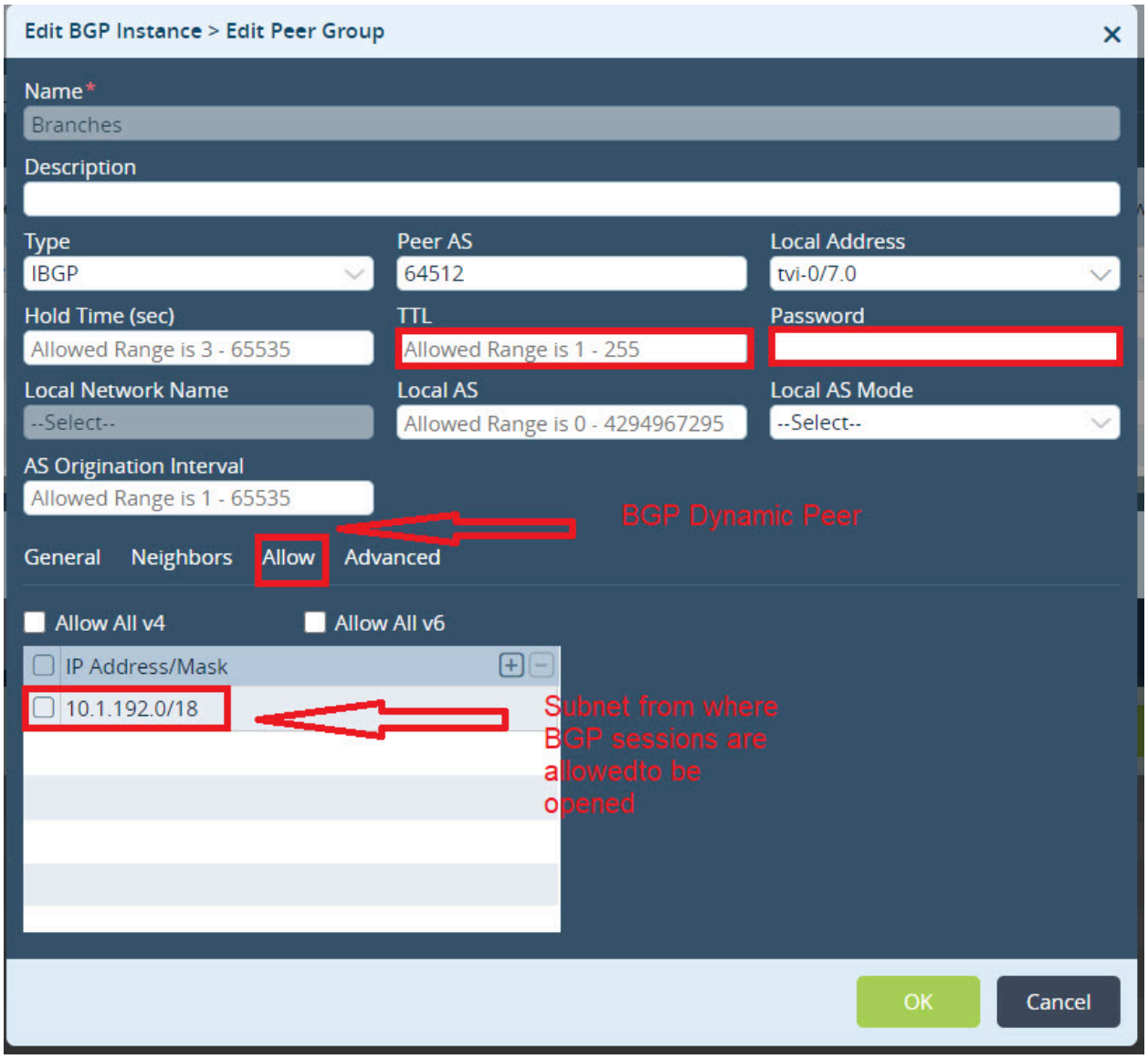
Authentication is the most important measure you can take to secure any routing protocol. The VOS OSPF and RIP protocols support clear-text and HMAC-MD5 authentication, which are simply mechanisms by which two neighbors prove their identity to each other by using a shared secret. No protocol messages 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.



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:



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

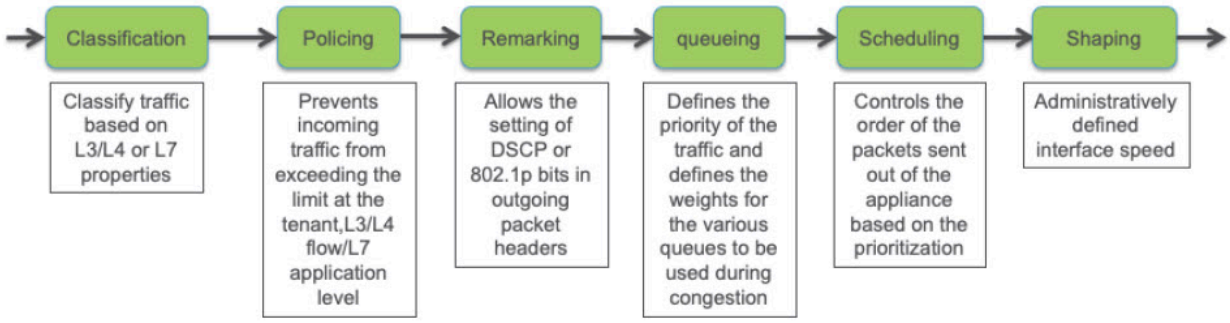
Configure Virtual Routers

QoS

For supported software information, click here.

To handle periods of network congestion, you can configure quality of service (QoS), also known as class of service, or CoS, to ensure that the network prioritizes business critical traffic over less important traffic and treat this traffic with higher priority. You can also use QoS for other tasks, such as policing, shaping, and remarking the QoS bits in the IPv4/IPv6 and VLAN headers.

The following figure shows how a packet that is processed by QoS functions flows through a Versa Operating System™ (VOS™) edge device. Each green box is a stage in the VOS QoS processing. The figure illustrates the order in which the QoS stages are performed by the VOS software.



This article discusses some of the QoS functions and best practices for using them.

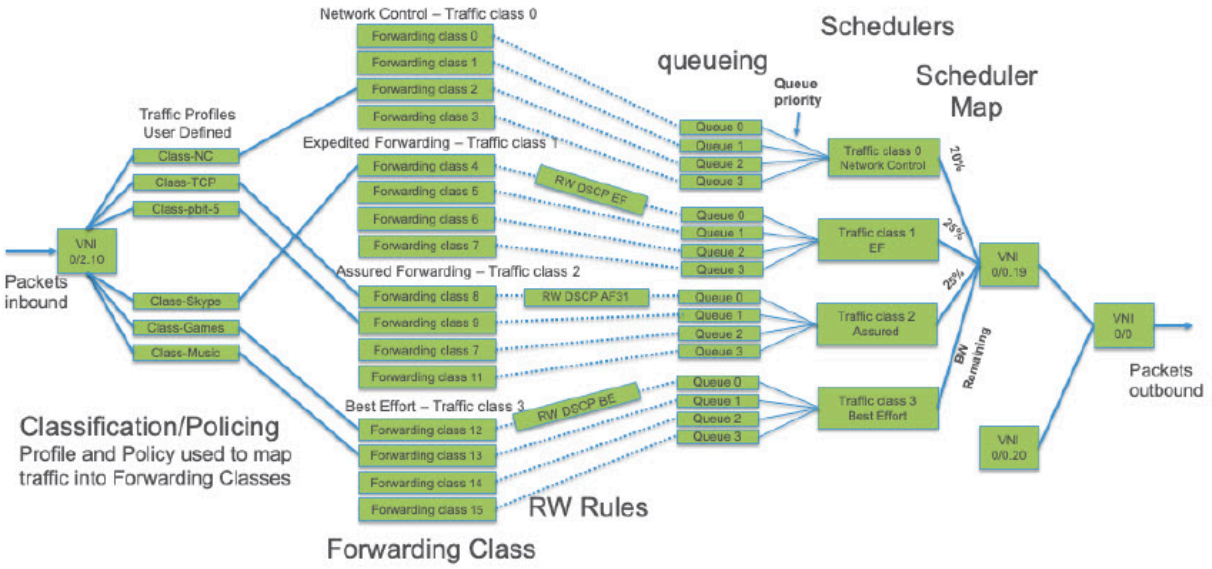
Classification

VOS edge devices support four traffic classes:

- Network control
- Expedited forwarding
- Assured forwarding
- Best effort

Each traffic class can use a maximum of four queues, and each queue can have a low or a high drop probability. The result is that 32 unique classification priorities are available.

If you use the default queue mapping, there are 16 forwarding classes that map to the forwarding queues, as illustrated in the following figure.

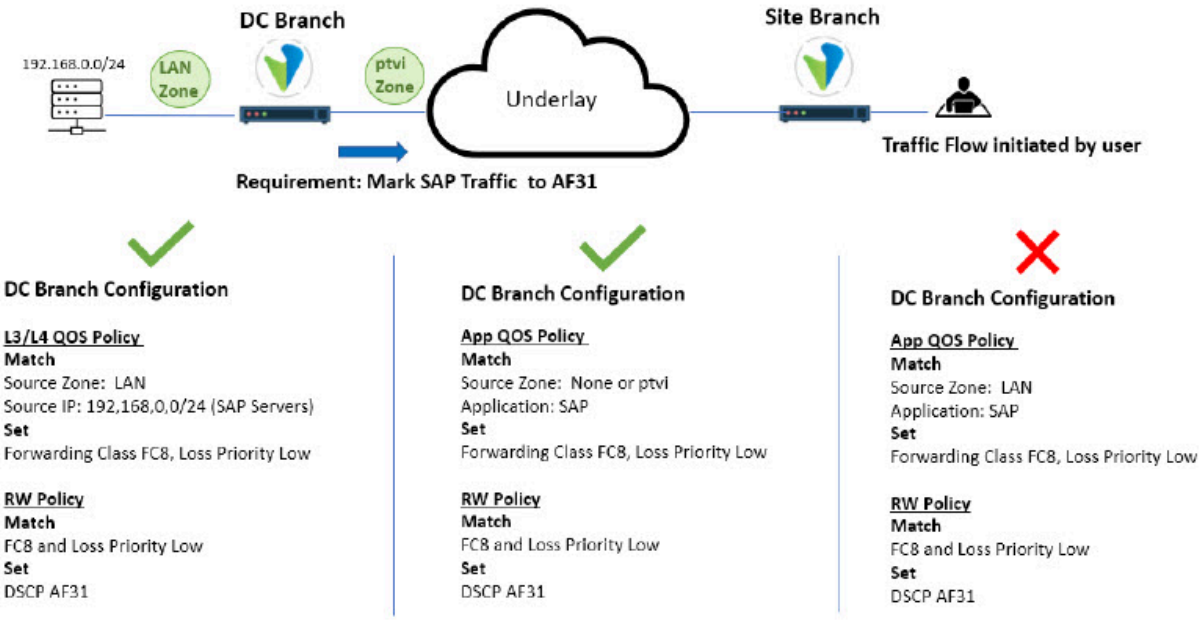


During the classification stage, ingress traffic is identified and associated with a forwarding class and loss priority. Classification can be done in two ways:

- Using QoS policies (Layer 3 and Layer 4 rules), which allow classification based on the following fields in the packet and packet header:
  - Destination port
  - Destination zone
  - DSCP
  - Ether-type
  - Ether-type-value
  - IEEE-802.1p
  - IP flags
  - Source IP address
  - Source port
  - Source zone
  - Time of day
  - TTL
- Using Application QoS (App QoS) policies (Layer 3 and Layer 7 rules), which allow classification based on the same fields as QoS policies classification as well as on applications and URL categories.

The following example illustrates how QoS classification works. Here, we have data center which hosts a SAP

application that users access from remote locations, and we want to set the outgoing DSCP value of the SAP traffic on the data center MPLS WAN interface to AF31. The following figure illustrates the policies required to effect this scenario. In addition to Layer 3/Layer 4 QoS policies and Layer 4 through Layer 7 App QoS policies, the configuration uses rewrite (RW) policies



On the VOS edge device in the data center, you apply an App QoS policy applies that matches the following conditions:

- Source zone—LAN
- Application—SAP

This App QoS policy references a QoS profile that places this traffic in the forwarding class FC8. Then, you apply a QoS propagation policy (that is, a rewrite policy) on the MPLS WAN network that remarks all FC8 traffic to a DSCP value of AF31.

This configuration works only if the traffic originates on a LAN connected to the data center VOS device. When the traffic originates from remote branch clients that are accessing the SAP application, the return traffic arrives from the DC LAN to the client. The traffic does not match the App QoS policy even though the source zone is set to LAN, because the App QoS policy classifies traffic based on a flow. Because the traffic originated at the remote branch, the first packet in the flow came from the PTVI zone (the PTVI zone is for traffic received over the SD-WAN overlay network), and so it does not match the App QoS policy whose source zone is set to LAN. Therefore, although the QoS rewrite policy is applied to traffic outgoing towards the MPLS network, the classification for the flow is performed when the packets arrive on this branch from the remote site.

The proper configuration is to have a match condition that either sets the source zone to PTVI or leaves the source zone empty, so that the SAP application matches regardless of whether it was initiated from a remote site or the local LAN.

The flow must be classified twice separately for both session directions: forward bidirectional flow and opposite bidirectional flow. For example:

- Forward-Initiated-App-QoS-Policy Match—Source LAN, destination PTVI will match the session for both session directions, forward and reverse: LAN-to-PTVI and PTVI-to-LAN.
- Reverse-Initiated-App-QoS-Policy Match—Source PTVI, Destination LAN will match the session for both session directions, forward and reverse: PTVI-to-LAN and LAN-to-PTVI.

If you must use a Layer 3/Layer 4 QoS policy in this use case, you need. to match based on the IP address of the SAP application or or on the source port, because a Layer 3/Layer 4 QoS policy does not support application-based matching. For a Layer 3/Layer 4 policy, the match is only in one direction, so you can specify the source LAN zone even if the traffic originates from a remote branch.

The flow must be classified twice separately for both flow directions: forward session direction and reverse session direction. For example:

- Forward-QoS-Policy Match—Source LAN, destination PTVI.
- Reverse-QoS-Policy Match—Source PTVI, destination LAN.

The following are best practices for QoS classification:

- Ensure that you use the correct match conditions, especially when you define source- or destination-based match conditions in an App QoS policy.
- When a traffic matches both a Layer 3 or Layer 4 QoS profile and an App QoS policy, the App QoS policy takes precedence.
- Verify the session details from the CLI to determine whether the session is attached to the expected QoS policy rule.

Policing Ingress Traffic

Policing allows you to rate-limit ingress traffic, providing a mechanism to prevent ingress traffic from overloading an egress port or the VOS device itself.

You can configure policing at the following levels. The policing configured at a lower level takes precedence over the policing configured at a higher level.

- Globally, at the tenant level, to enforce licensed bandwidth
- At the Layer 3/Layer 4 QoS policy level, for unidirectional policing on ingress traffic only
- At the Layer 7 App QoS policy level, for bidirectional policing on ingress and egress traffic

Policing uses the following parameters:

- Peak rate, which defines the maximum transmission rate, in pps or Kbps.
- Burst size, which defines the number of bytes that are allowed beyond the configured peak rate. You set the burst size to avoid retransmission during bursts of traffic.

- Loss priority to be used by congestion-avoidance algorithms.

The following are best practices for policing:

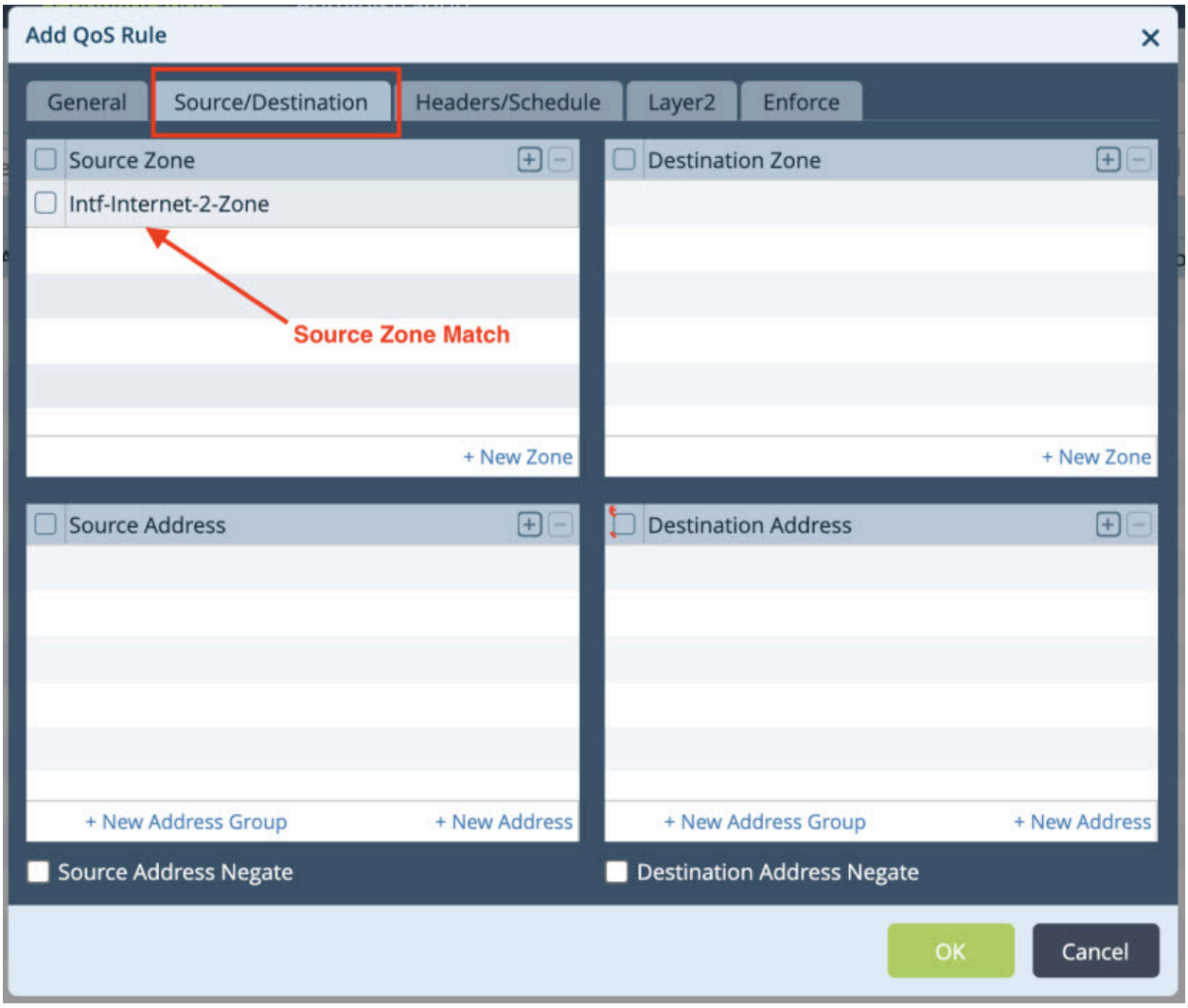
- The default maximum burst size for a policer is 15000 bytes.
- Use a policer rather than a shaper for real-time traffic especially for traffic that is sensitive to jitter, such as voice traffic.
- You can configure a policer at the tenant level to protect aggregated SD-WAN licensed bandwidth. For example, if two tenants are configured on VOS device that has a licensed bandwidth of 200 Mbps, you. can configure a 100-Mbps policer for each organization to ensure fairness between them.
- A tenant-level policer applies only to inbound traffic arriving on the VOS device from a LAN or WAN interface. Hence, you can use a tenant-level policer to limit the download and upload bandwidth for a particular organization.

Use QoS Policy as an Access Control List

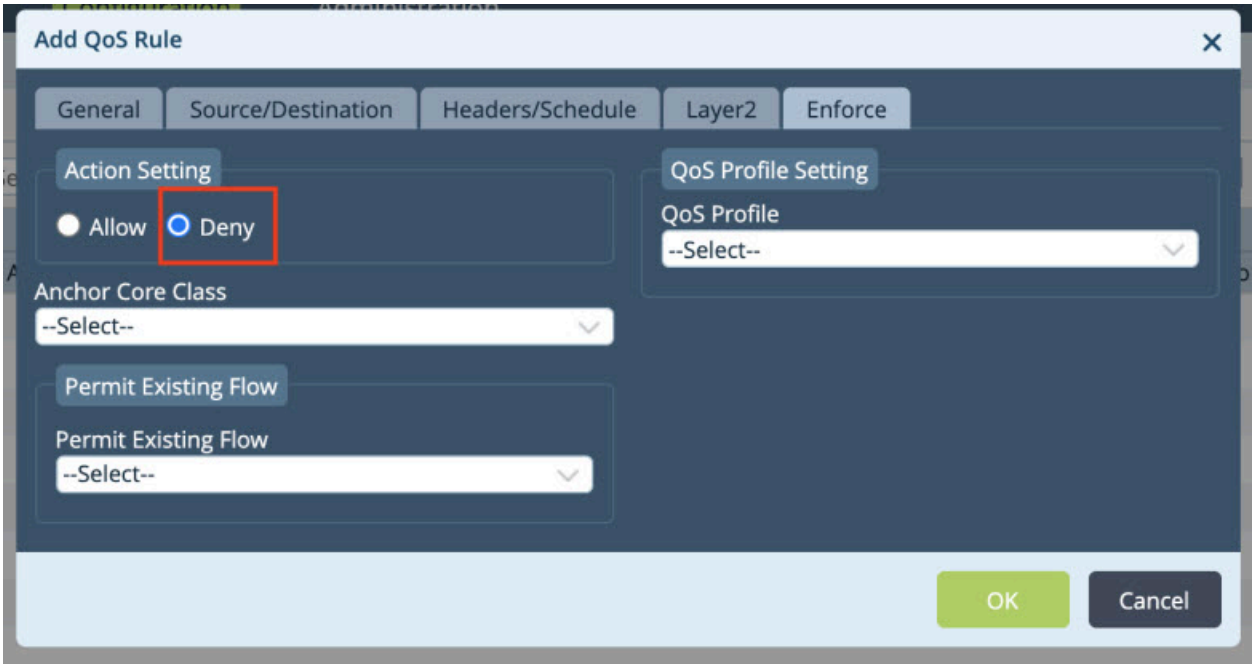
You can use QoS policies to emulate the function of an access control list (ACL) in a deployment in which no security services are configured and you want to deny traffic that matches a particular Layer 3 or Layer 4 rule. To do this, you create a QoS policy rule that has a match condition and a Deny action. Note that you can apply QoS policy to deny traffic only for through traffic or for external traffic destined for the VOS device. You cannot use QoS policy rules to control traffic generated by the VOS device itself.

To configure a policy rule to deny traffic:

1. In Director view:
   a. Select the Administration tab in the top menu bar.
   b. Select Appliances in the left menu bar.
   c. Select an appliance in the main pane. The view changes to Appliance View.
2. Select the Configuration tab in the top menu bar.
3. In the Networking tab in the left menu bar, select Class of Service > QoS Policies.
4. Select the Rules tab in the horizontal menu bar to define the matching criteria to select the incoming packets to which to apply the QoS policy.
5. Click the Add icon to add the rule.
6. To have the match condition be a source zone, select the Source/Destination tab and then select the source zone.



7. Select the Enforce tab to configure the Deny action.



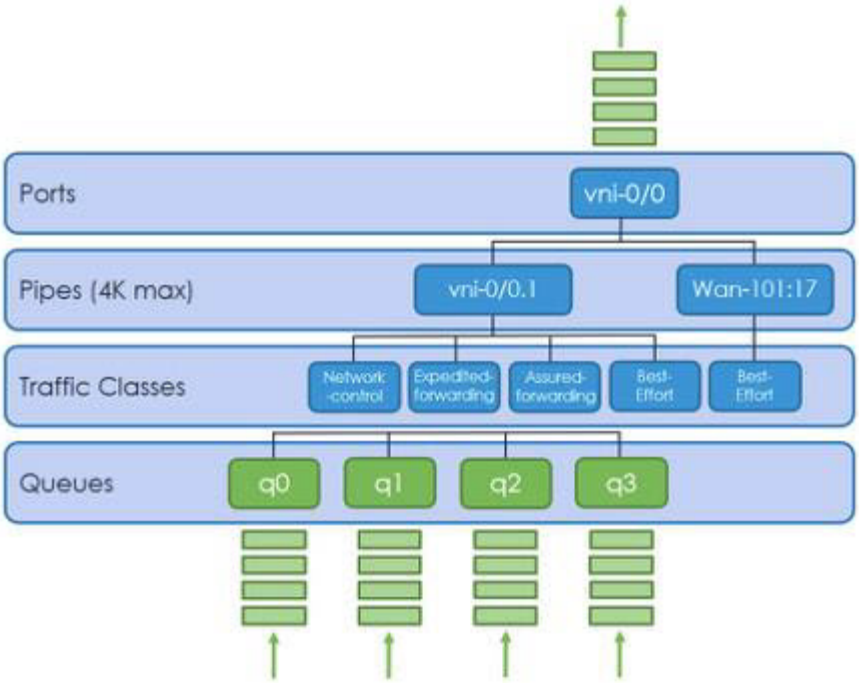
8. Click OK.

The following are best practices for using QoS Policy to emulate access lists:

- Create a QoS policy with a source zone that matches the VOS device interface you want to protect and associate the Deny action with the zone. You could use this configuration, for example, to prevent a WAN interface from being ping-able from the internet.
- Use this style of QoS policy as a stateless access list to block certain traffic.

Hierarchical Shaping

You can configure a hierarchical scheduler block on an egress interface to schedule and shape traffic. The following figures illustrates the arrangement of hierarchical shapers on VOS Edge devices.



You can use schedulers to perform shaping at the following levels:

- Port
- Pipe
- Traffic class

All ports have equal priority.

All pipes with a port have equal priority. A pipe represent seither a VLAN interface or a dynamically created IPsec path.

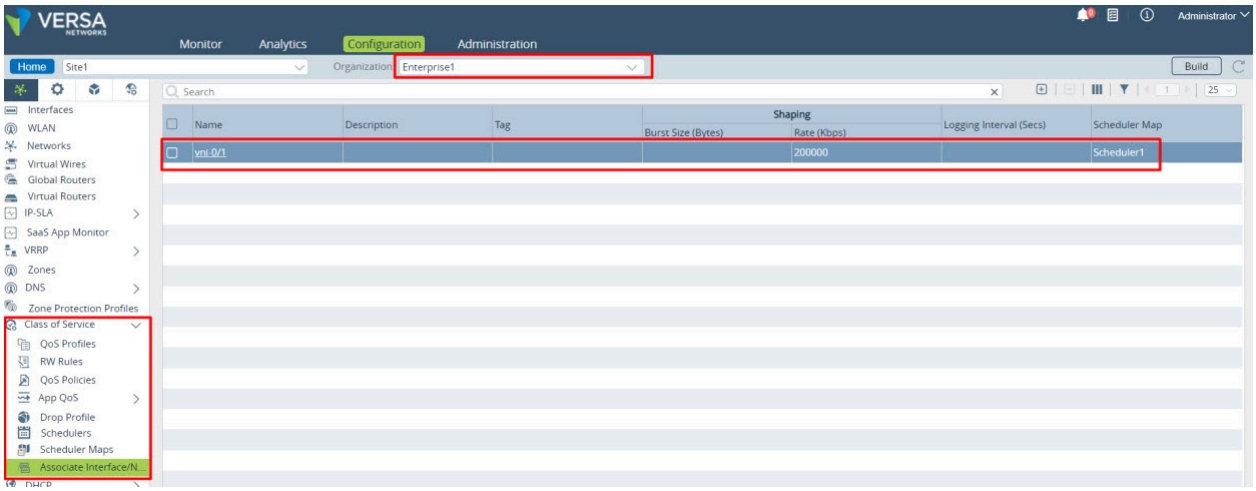
Traffic classes a pipe are handled in a strict priority order, which is, from highest to lowest, network control (tc0), expedited forwarding (tc1), assured forwarding (tc2), and best effort (tc3). Each traffic class has four queues that are scheduled using weighted round-robin (WRR).

The following are best practices for hierarchical shapers:

- To ensure that a higher priority traffic class does not starve the other traffic classes, thus ensuring both priority and fairness, apply a shaper at the port level. and configure a scheduler for each traffic class as a percentage of the port shaper rate or using an exact rate. For instance, an incorrectly configured network control traffic class can starve traffic classes 1, 2, and 3 if no transmit rate is configured and the network control traffic is using all available bandwidth on the port.
- For each scheduler, you can associate a unique drop profile for each loss priority as a congestion avoidance mechanism, that uses the. weighted random early detection (WRED) algorithm to keep track of queue depth and then, when the threshold is reached, starts randomly dropping packets.
- The shapers burst size is automatically set to the interface link speed divided 8000 bytes (burst size = interface link speed/8000 bytes) as per the Intel recommendation. If you configure a higher burst size, it may work but it may cause intermittent drops as a result of hardware limitations.

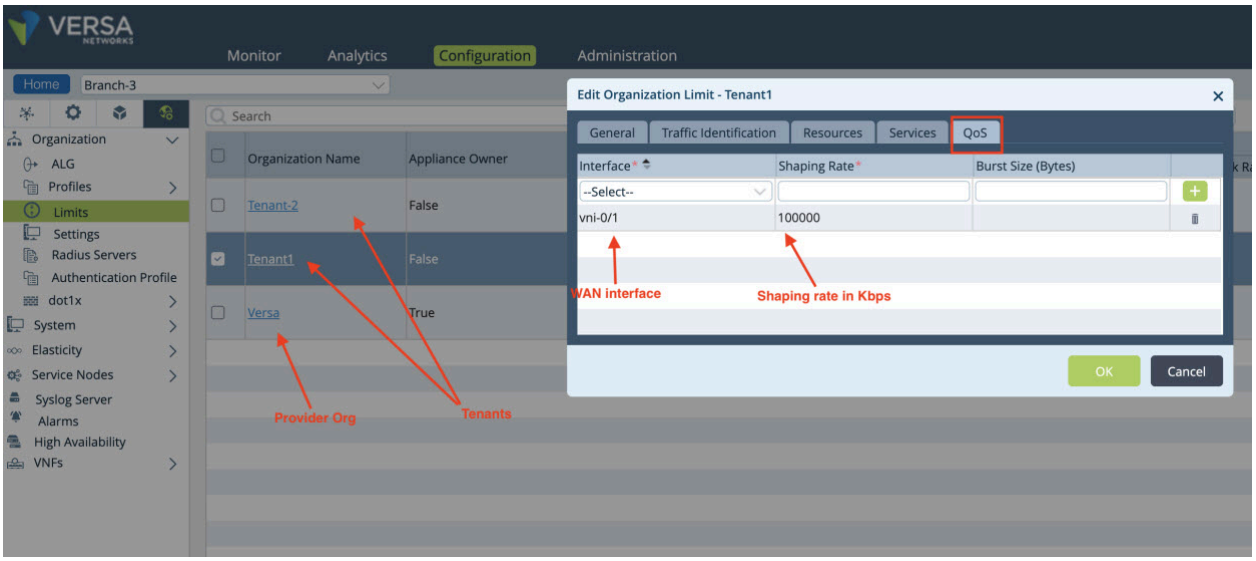
Per-Tenant Shapers

On VOS edge devices, you can configure traffic shapers for individual provider organizations to allocate the amount of WAN-facing bandwidth available on a per-tenant basis. The prerequisite for doing this is that you must have already configured CoS for the provider organization and already applied it to the WAN interface on which you want to configure a per-tenant shaper, as illustrated in the following screenshot.



Then configure the per-tenant traffic shaper for the tenant or tenants:

1. In Director view:
   a. Select the Administration tab in the top menu bar.
   b. Select Appliances in the left menu bar.
   c. Select an appliance in the main pane. The view changes to Appliance View.
2. Select the Configuration tab in the top menu bar.
3. Select Others > Organization > Limits in the left menu bar.
4. Select the tenant in the main panel. The Edit Organization Limit popup window displays.
5. Select the QoS tab. In the Interface, select the WAN interface, and in the Shaping Rate field, enter the shaping rate, in Kbps.



6. Click OK.

The following is a best practice for per-tenant shapers:

- On a multitenant branch, configure the amount of bandwidth each tenant is entitled to use. For example, on a branch with two tenants that are allowed 100 Mbps of bandwidth, configure the shaping rate for each tenant to enforce fairness between tenants.

QoS Rewrite and Propagation

Rewrite (RW) rules allow you to rewrite the attributes of packets QoS attributes as they are leaving the VOS device so that you can convey the importance of the packet. Downstream nodes can use the QoS attributes to classify traffic and then take the appropriate scheduling action, or in case of congestion can give precedence to critical traffic. A rewrite rule rewrites QoS bits for an existing forwarding class.

With rewrite rules, you can rewrite the following fields:

- IEEE 802.1p bits in the VLAN header
- TOS bits in the IPv4 header
- Traffic class bits in the IPv6 header

The Versa Networks technology uses an overlay to transport packets from one branch to another. The packets are encapsulated in a VXLAN header and then transported to the remote branch. Hence, the packets have two headers, referred to as the inner header and the outer header. You can use two methods to change the QoS markings on the inner and outer headers:

- Use a rewrite policy to classify traffic and set the QoS bits.
- Use rewrite options to copy the inner header markings to the outer header, and vice versa.

Rewrite Policy

You can use a rewrite policy to remark packets and frames based on the classification done on ingress. The classification is done using a Layer 3/Layer 4 or App QoS policy that assigns the packets to a forwarding class and loss priority. The rewrite policy uses the forwarding class and loss priority information as match conditions to set a QoS value.

A typical use case is to rewrite the LAN traffic passing through a CPE device, whether the traffic is destined for another LAN port or for egress somewhere else on a remote branch. The DSCP of the traffic is modified and is carried, or propagated, with the traffic as it moves through the network.

There are three types of rewrite policies:

- DSCP rewrite policy
- DSCP6 rewrite policy
- 802.1p rewrite policy

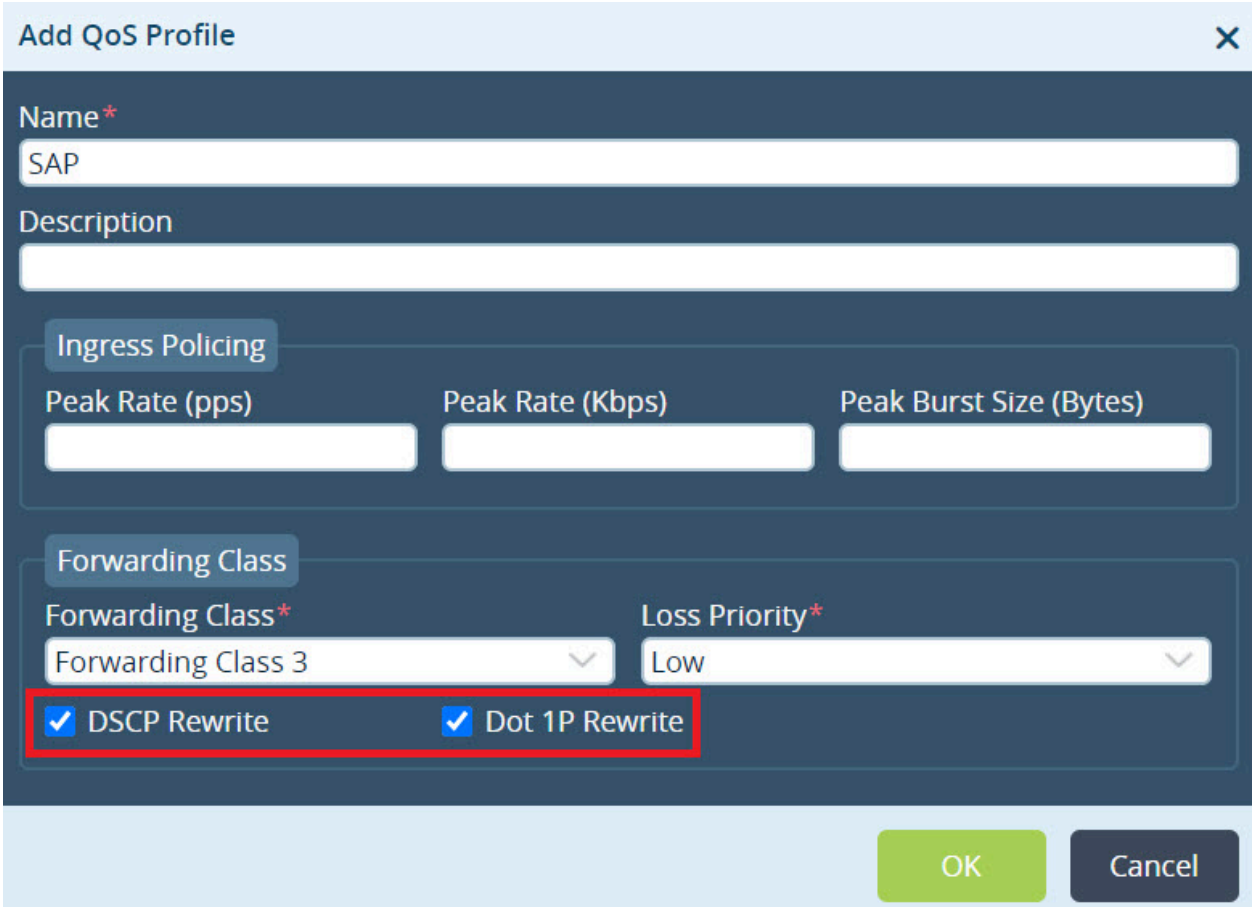
Depending on where you apply the rewrite policy, it modifies the QoS bits of either the inner or outer header, but it does not propagate.

In a rewrite policy, the match condition is the forwarding class and the loss priority. You configure the forwarding class in the QoS profile, which is referenced in the Layer 3/Layer 4 or App QoS policy.

You enable the following options in the QoS profile depending on whether the type of rewrite policy:

- DSCP Rewrite
- Dot 1P Rewrite

If you do not select either the DSCP rewrite or Dot 1P rewrite option in the QoS profile, the rewrite policy does not take effect.

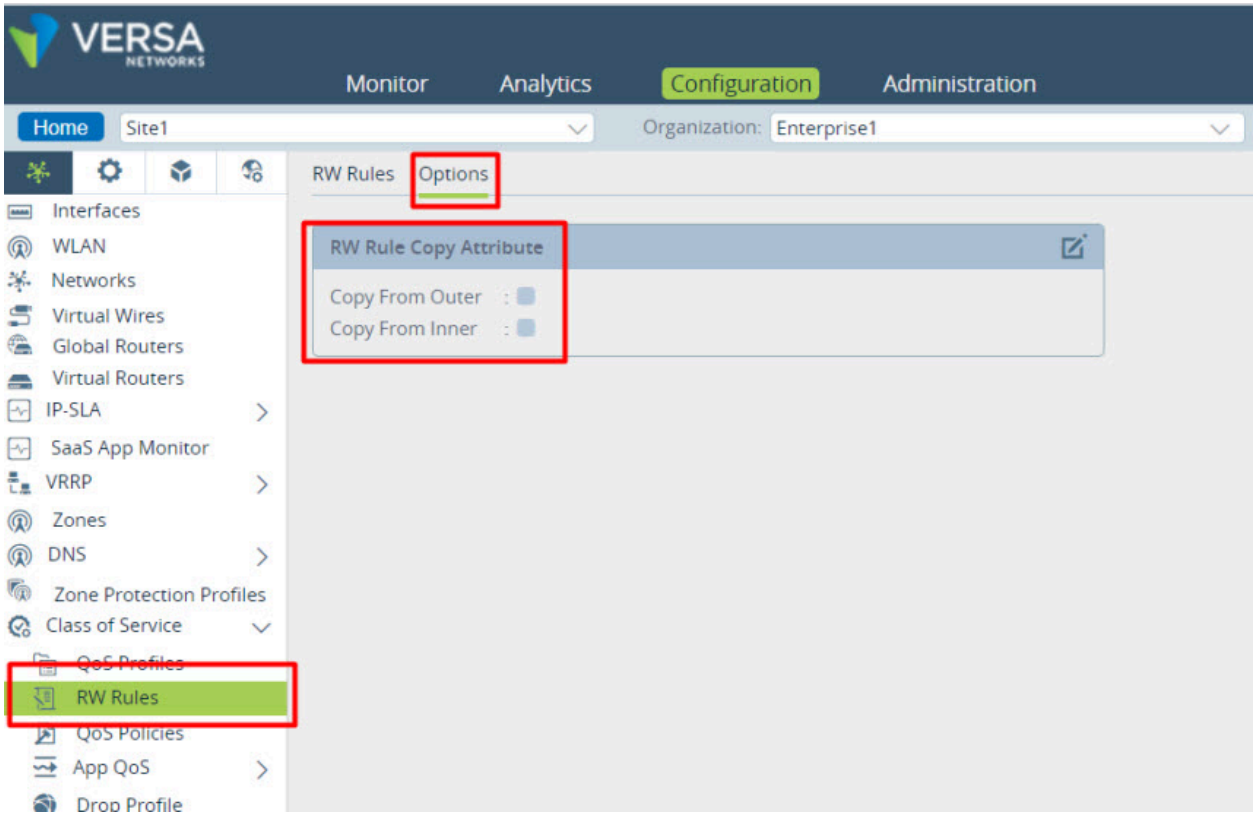


One use case for this setting is when you have more than one QoS policy that classifies traffic to the same forwarding class. However, if you do not need to remark all the traffic with the rewrite policy, you can use this setting to help differentiate which traffic is forwarding class is evaluated by the rewrite policy.

A second use case is on a multitenant VOS device. on which the rewrite policy is applied on the WAN interfaces that are owned by the provider. If there are two customer tenants on the branch and each tenant uses the same forwarding class in their QoS policies, they can choose whether to set the DSCP Rewrite and Dot1P Rewrite option, depending on whether they want their traffic to be remarked..

Rewrite Options

The rewrite options set a flag to copy the markings between the inner and outer IP headers. This is a global setting, so it affects all the traffic passing through the VOS device. You cannot apply in on a per-interface or per-traffic class basis.



You can configure the following options:

- Copy From Outer—Copy the marking of the outer IP header to the inner IP header when a branch sends packets to a remote branch.
- Copy From Inner—Copy the marking of the inner IP header to the outer IP header when a branch receives packets from a remote branch.

The Copy From Outer setting remarks the inner IP header of traffic coming from a remote site. If the outgoing network applies a rewrite policy that also remarks the inner IP header of the same packet, the marking in the rewrite policy overwrites the marking made by the Copy From Outer option. For example, on a regular branch the outgoing network could be a LAN and, on the hub the outgoing network could be a WAN interface.

With the Copy From Outer setting, the VOS edge device trusts the markings from the underlay. You should trust outer markings on packets received on a private circuit such as MPLS, but you should not trust them on packets received not from the internet. Therefore, you should use this setting only on a branch that has only an MPLS WAN circuit.

The Copy From Inner settings works on a model whereby the VOS device trusts the markings from the LAN. Because this setting remarks the outer IP header of the traffic coming from the LAN site, if you apply a rewrite policy on the WAN network that also remarks the outer IP header of the same packet, the marking made by the Copy From Inner setting overrides the marking made by the rewrite policy.

QoS Propagation Policy on Hubs

On a hub, traffic that arrives encrypted from a remote site, is decrypted, and is encrypted again before it is sent it to another site. Because the packets are decrypted on the hub, the QoS propagation policy can remark both the inner and outer headers based on the classification made on the inner packet (that is, the actual application packet). The hub can also use the rewrite option flags like Copy From Outer and Copy From Inner.

Supported Software Information

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

Additional Information

Configure Adaptive Shaping
Configure CoS
Configure Policy-Based Forwarding
Configure Schedule Objects