



SDWAN Solution Design Guide

Version 1.2

May 2021

Table of Contents

1	INTRODUCTION	5
1.1	Intended Audience	5
1.2	Prerequisites.....	5
1.3	Disclaimer	5
1.4	Revision History.....	5
2	REFERENCE NETWORK ARCHITECTURE.....	6
3	DESIGN GUIDELINE FOR THE SD-WAN HEADEND	7
3.1	Design considerations of the headend architecture	7
3.1.1	Headend deployment options.....	9
3.1.2	Firewall dependencies for deployment Headend in Datacentre.....	9
3.1.3	Best Practice on headend hardening.....	10
3.2	Versa Analytics deployment options.....	11
3.2.1	Recommended Production Analytics Deployment.....	12
3.2.2	Role of the Log Forwarder	13
3.2.3	Best Practice for Versa Analytics Sizing	14
3.3	Best Practice Headend Design.....	15
4	BRANCH DEPLOYMENT	16
4.1	Single CPE Branch with Internet / MPLS Transport.....	16
4.1.1	Active-Active Dual CPE Branch with Internet / MPLS Transport	17
4.1.2	Best Practise Summary	17
4.2	Dual CPE Branch with Internet / MPLS & LTE Transport	18
4.2.1	Branch High-Availability.....	18
4.2.2	Cross-Connects in Active-Active HA	20
4.2.3	DIA on Active-Active HA	23
5	LTE TRANSPORT	24
5.1	Hot Standby Mode:	24
5.1.2	LTE as backup for Internet Traffic.....	29
5.1.3	Load Balance two wired WAN links plus LTE backup	29
5.1.4	Management Traffic Priority	30
5.2	Cold Standby Mode	31
5.2.1	Benefits of Cold Standby Mode	31
5.2.2	Limitations of Hot Standby Mode.....	31
5.2.3	Configure LTE Cold Standby.....	31
5.3	Comparison of LTE Hot Standby mode vs Cold Standby mode	32
6	SD-WAN OVERLAY.....	33
6.1	Overlay IP addressing	33
6.1.1	Best Practice IP overlay addressing	34
6.2	Encrypted and clear-text overlay	35
6.3	Static definition per WAN interface	36
6.3.1	Configure Static definition per WAN Interface	36
6.3.2	Dynamic Definition by SD-WAN Policy	36
7	SD-WAN TOPOLOGIES.....	38
7.1	Full Mesh	39
7.2	Hub-and-Spoke.....	41
7.2.1	Spoke-to-Hub Only	41
7.2.2	Spoke-To-Spoke via Hub.....	43

7.2.3	Spoke-To-Spoke Direct aka Partial Mesh.....	46
7.3	Regional Mesh aka Spoke-Hub-Hub-Spoke	50
7.4	Connecting sites over disjointed underlay networks	53
7.5	SD-WAN topologies for Geographical Isolated regions.....	54
7.6	Multi-VRF/Multi-Tenant Topologies.....	56
7.7	Best Practices Topologies	57
8	DIRECT INTERNET ACCESS	58
8.1	VOS Edge Device DIA architecture.....	59
8.1.1	Central or Regional Internet breakout	59
8.1.2	Breakout to an MPLS underlay for common SP services.....	61
8.1.3	Application based break-out	62
8.2	Best Practice: Remove the local breakout default route	66
8.3	Best Practices for DPI application recognition	67
8.4	Best Practice Summary.....	67
8.5	Performance based SaaS Optimisation	68
8.5.1	Setting up performance-based breakout configurations	68
8.5.2	Verification of Performance based breakout.	72
8.5.3	Different variants of performance-based breakout.	73
8.5.4	Performance based breakout best practices.....	76
8.6	Role of DNS proxy for different breakout scenarios	77
8.6.1	Verification of DNS proxy being used.	80
8.6.2	Best Practices for DNS Proxy	81
8.7	Importance Security services for internet breakout	82
8.7.1	Configuring NextGen Firewall security for DIA traffic	82
8.8	Role of Web Proxies with different breakout services	84
9	SD-WAN GATEWAY	86
9.1	MPLS L3VPN Interworking.....	86
9.1.1	Best Practices for SD-WAN Gateways.....	86
9.2	Options to interconnect legacy networks with SD-WAN Gateways.....	89
9.2.1	Option 1: BGP to exchange routes with the MPLS provider on the MPLS WAN interface.....	89
9.2.2	Option 2: BGP with the MPLS provider on the LAN interface	93
9.3	Gateway for Internet bound traffic.....	96
10	SD-WAN TRAFFIC OPTIMIZATION	97
10.1	Traffic-Steering	97
10.1.1	Best Practices for SD-WAN traffic Steering	98
10.2	Traffic Conditioning.....	101
10.2.1	Packet Replication	101
10.2.2	Best Practices Packet Replication	103
10.2.3	Forward Error Correction	103
10.2.4	Best Practices FEC.....	105
10.3	Business Intent traffic steering with Application-Steering-Templates.....	107
10.3.1	Best Practice Application-Steering-Templates	108
10.4	SD-WAN Path Policies.....	109
10.4.1	Best Practices Path Policies	110
11	ROUTING	111
11.1	Static Routing	111
11.1.1	Floating Static Route.....	111
11.1.2	Route Leaking with Static routes	112
11.2	Dynamic Routing	113
11.2.1	Fast Convergence – SD-WAN Side	113
11.2.2	Fast Convergence - LAN Side	115
11.3	Scalability.....	115
11.3.1	OSPF.....	115

11.3.2	BGP	116
11.4	Loop Prevention	117
11.5	Routing Security	120
11.5.1	IGP Security	120
11.5.2	BGP Security	120
11.6	Best Practice Routing	121
12	QOS	122
12.1	QoS Stages	122
12.2	Classification	122
12.2.1	Best Practices Classification	123
12.3	Use-Case Scenario	124
12.4	Policing/Rate-Limiting ingress traffic	125
12.4.1	Best Practices Policing	125
12.5	Access Control List (ACL)	126
12.5.1	Best Practice ACL	126
12.6	Hierarchical Shaper	127
12.6.1	Best Practices Shaping:	128
12.7	Per Tenant Shaper	128
12.8	QoS Rewrite and propagation	129
12.8.1	RW Policy	129
12.8.2	RW Options	131
12.8.3	Copy from Outer	131
12.8.4	Copy from Inner	132
12.9	QoS Propagation Policy on the Hub	132
13	VERSA ANALYTICS SCALING RECOMMENDATIONS	133
13.1	Usage Monitoring Logging Recommendation	133
13.2	Flow logging Recommendation	134
13.2.1	Limiting Flow Logging on the CPE	134
13.2.2	Flow Logging for Achieve Only	135
13.2.3	Throttling/disabling flow logs in log collectors	135
13.2.4	Database global limits and retention settings	135
13.3	Alarm Logging Recommendations	136
13.4	Historical Data Storage Recommendations	136
14	TRAFFIC TUNNELLING TO CLOUD SECURITY PROXY	137
14.1	Tunnelling Protocol Selection	137
14.2	Default Topology	137
14.2.1	Active Tunnel Failure	138
14.2.2	Application Performance Monitoring	138
14.3	Single Site Traffic Load Balancing	139
14.4	Local and Remote Tunnel	139
14.4.1	Remote Backup	140
14.4.2	SAAS Optimisation	141
14.5	High Availability Site	141
14.6	Proxy	142
14.6.1	DNS Proxy	142
14.6.2	HTTP Proxy	143
14.7	Best Practices	143

1 Introduction

The Versa SD-WAN solution is a highly robust and flexible platform that offers capabilities to address various use cases. This document addresses the most common use cases and describes the Versa recommendations and best practices for these deployments.

The objective is to help achieve a standardized approach to designing Versa SD-WAN solutions. This document is also intended as a source material for other solution design documents. However, it is not meant to cover detailed operational best practices nor act as a service management manual.

1.1 Intended Audience

This guide is targeted at network architects, engineers, administrators, and other technical audiences interested in designing, implementing and deploying Versa Networks SD-WAN solutions.

1.2 Prerequisites

Basic Versa Training is assumed from the audience to this document. Readers are expected to have a working knowledge of the Versa SD-WAN architecture as well as the wider Versa ecosystem.

1.3 Disclaimer

This document is intended as a generic guide. It does not cover every use case the versa secure SDWAN solution supports. The readers can use this document as a guide to explore use cases for their specific deployments. Users of this document are responsible for their own designs.

1.4 Revision History

Revision Date	Revision Version	Summary of Changes	Author
Aug-2020	1.0	Initial Issue	
Oct-2020	1.1	Section 3.2.1 – included log forwarder only Analytics disaster recovery. Updated broken links New section on Analytics Best practices New chapter 14 about Site2Site tunneling	Tayo Ogunseyinde, Michal Zakrzewksi
May 2021	1.2	Section 3.1 – Update to Inter DC Analytics Recommendation. Ammend Figure 2 Correct error on Section 12.8.3/4 update to definition of “Copy from inner” & “Copy from outer”	Tayo Ogunseyinde

2 Reference Network Architecture

The network diagram enclosed below represents a high-level blueprint of a typical network topology build on Versa Solutions.

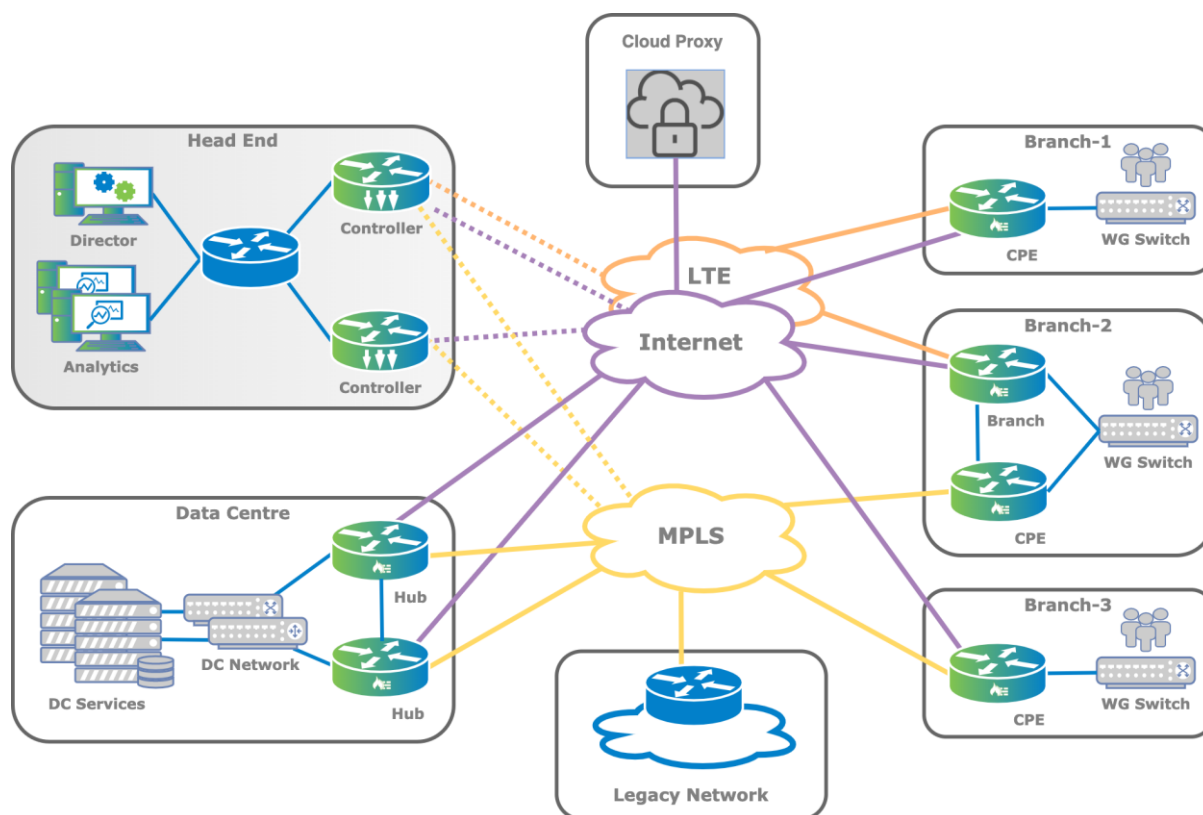


Figure 1 Reference Architecture

In this diagram, there is one Head End, one Data Centre and three Remote Branches. Additionally, there is also a pictogram of the Legacy Network and Cloud Proxy. The Transport Networks shown in the diagram are MPLS and Internet and LTE from one or many different Service Providers. The network orchestration, the Head-End, is reachable through Versa Controllers connected in all Transport Networks. Versa Director and Versa Analytics are hosted in the Head End.

All components depicted in above picture will be described in more details in this document.

The designs and best practices described in this document are based on 20.2.2 software version.

3 Design guideline for the SD-WAN headend

3.1 Design considerations of the headend architecture

A typical Versa SD-WAN headend architecture is deployed fully resilient with Geographical redundancy, or when deployed in public cloud, deployed in different availability zones. This requires that some infrastructure needs to exist to provide the necessary interconnectivity between the different headend components.

The SD-WAN headend topology has the following components:

- Northbound segment
- Control or southbound segment
- Service VNF router to interconnect datacenters

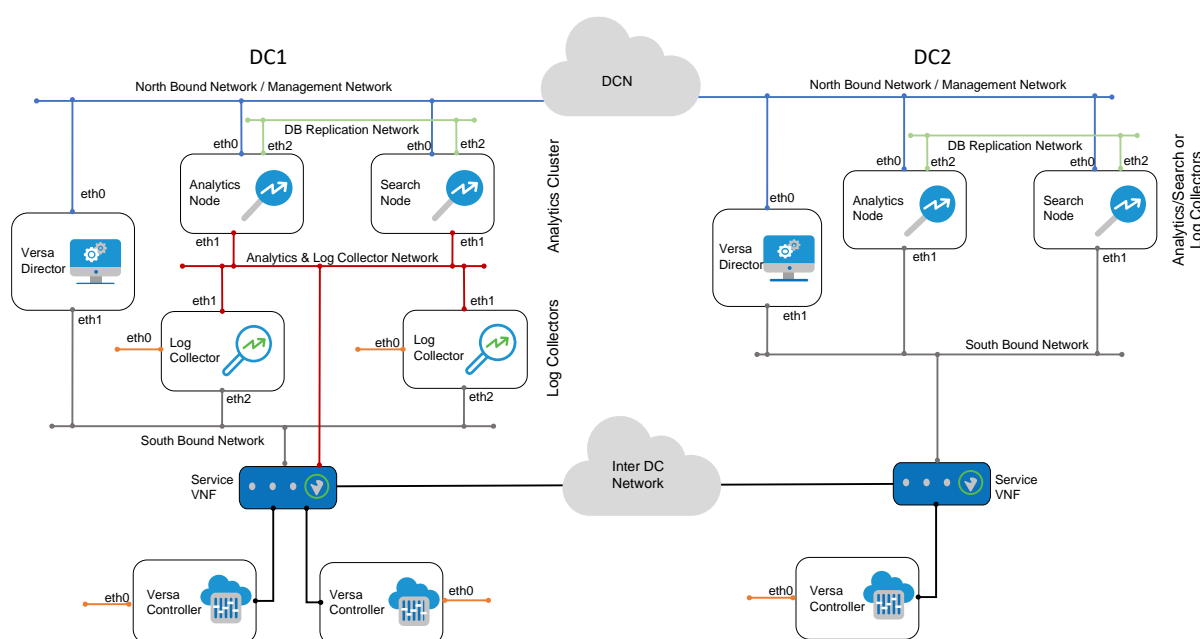


Figure 2 Typical Geo-Redundant Head-end design

The northbound network in the headend is typically how users or other OSS systems connect to the services, e.g. operators connecting to Director or Analytics GUI and OSS systems using the REST API.

The Versa Directors maintain Synchronization over the the northbound network. It must be noted that the SD-WAN controller eth0 are required to be connected to the northbound network as well. Versa Director uses the eth0/out-of-band network to configure the controllers. The southbound or control segment DC is used for connectivity to the SD-WAN overlay through the controllers. As Director and Analytics nodes are hosts (not routers) they require a router to be a gateway towards the redundant controllers. A dynamic routing protocol (OSPF or BGP) is required to provide accurate overlay IP prefix reachability status. BFD must be enabled towards the DC Service VNF device. This additional router might be an existing Datacenter router is the Datacenter network or a VOS Edge Device managed by Versa Director.

The solution assumes reachability between the DataCenters over the Versa Director North-Bound OSS Network. In the event an outage of Control Network Service router, the Versa

Director remain accessible via the Northbound Network allowing the administrator to gracefully failover to the standby Director.

Note: it is not mandatory to always connect the controllers to all Transport Domains. Technically if a SD-WAN branch device is connected to the controller over just one underlay, the SD-WAN will function. However, from a resiliency point of view, it is advised to connect the controller directly to all used Transport Domain.

Important Note !

Director North-bound OSS Network should not share fate with the Control Network to avoid Director Split Brain. Director Split Brain means that both the Directors are up but the network reachability between them is lost which makes both Directors assume that they are the Master.

Where the Versa Director must be reachable over the internet, we recommend appropriate security is in place to secure the North-Bound Traffic. However, in certain situations such as cloud hosted Versa Director, a third dedicated interface into a DMZ should be used. This allows separation between the link used for Director Sync and the internet access to the Director.

The below is simple single DC deployments scenario where service VNF routers are NOT required. This topology uses flat Layer 2 LANs for both the Northbound and the Control Network. VRRP is used on the Control LAN to for Gateway redundancy.

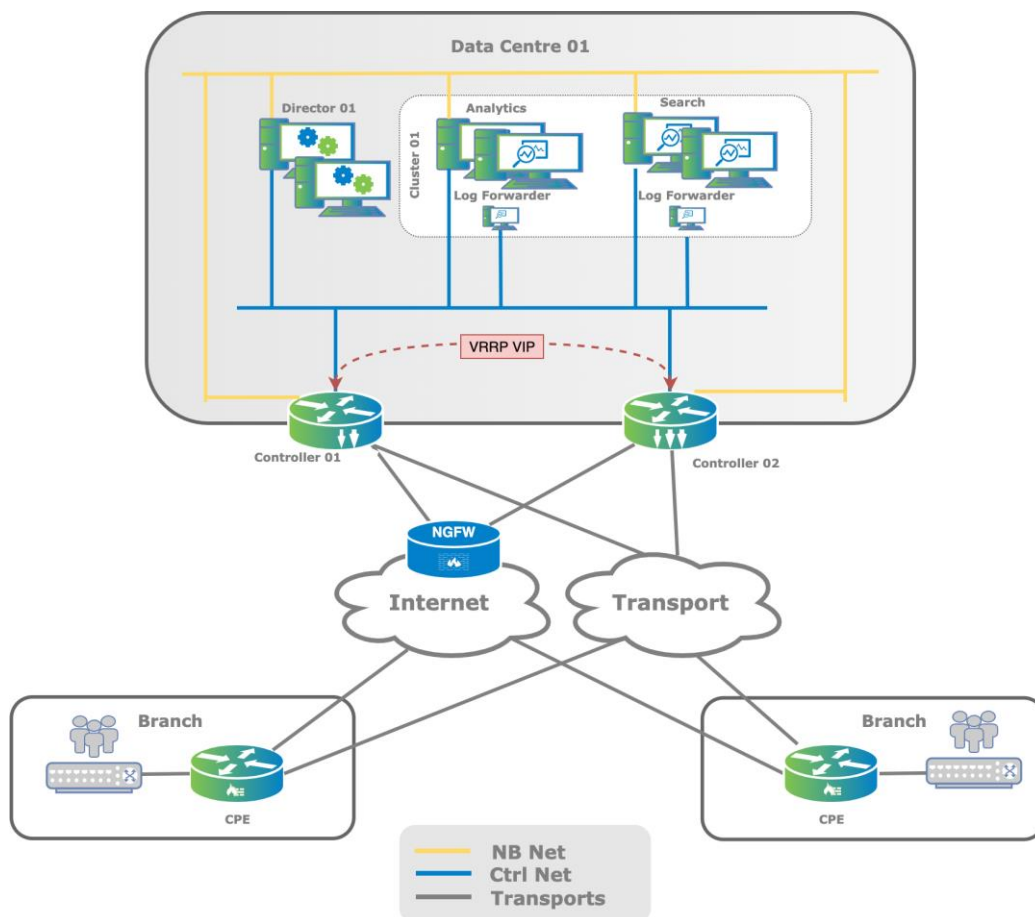


Figure 3 Non Geo-redundant Headend

This design however does not lend itself to creating Layer 3 Security Zones at the headend. If security is required, Firewalls will need to be in bridge mode.

3.1.1 Headend deployment options

The depicted typical headend architecture above can be deployed as host OS on physical hardware. We call this a bare metal deployment. Alternatively, this headend architecture can also be deployed on a virtualized architecture. For such virtualized architecture, Versa supports KVM, ESXi and the public clouds AWS or Azure.

It must be noted that a deployment on bare metal hardware is recommended as this gives better performance and/or scalability (approx. 20-30% over virtualized). However, when deployed on physical hardware, there might be some dependencies on the hardware (not all server hardware is supported, such as special RAID controllers).

For both bare metal and virtualized deployments, it is important to follow the hardware requirements as defined in the online documentation:

<https://docs.versa-networks.com/Getting Started/Deployment and Initial Configuration/Headend Deployment/Headend Basics/Hardware and Software Requirements for Headend>

3.1.2 Firewall dependencies for deployment Headend in Datacentre

When implementing the headend in an existing Datacenter, it must be noted that several protocols and functions are required to connect to it. This put requirements on the connectivity, reachability and security. Refer to the Versa documentation for details of the firewall requirement:

https://docs.versa-networks.com/Getting_Started/Deployment_and_Initial_Configuration/Firewall_Requirements/Firewall_Requirements

3.1.3 Best Practice on headend hardening.

Deploying Versa Directors which are exposed to Internet for external user access requires additional hardening with common IT hardening practices. For example, this includes installing OS patches in Ubuntu (provided by Versa, do not self-install OS patches directly from Ubuntu), but also other best practices such as installing official certificates (Versa ships with self-signed certificates) and changing the default passwords. Your Versa contact person can give you guidance on system hardening in more detail.

3.2 Versa Analytics deployment options

The minimum recommended Versa Analytics deployment for full redundancy requires four analytics nodes in a cluster. This design provides active-active HA by having two nodes in the Search personality and two nodes in 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 needs to be low for storage and query performance. Therefore, the recommendation is to allocate the nodes of the same cluster in the same datacenter or at least within the same availability zone.

Important Note !
 Cassandra database synchronization requires <10ms latency between the nodes in the same cluster.

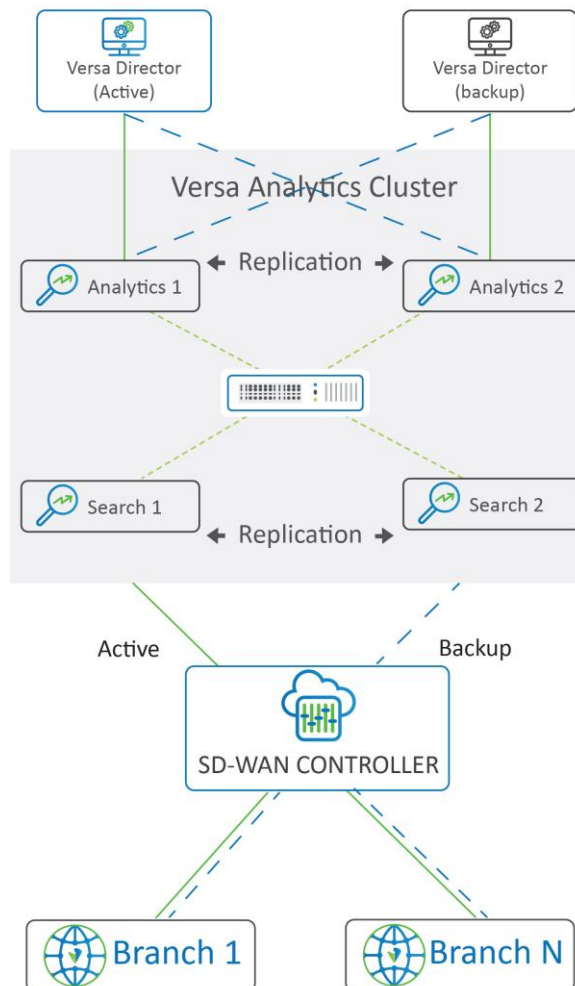


Figure 4 Typical Analytics Deployment

3.2.1 Recommended Production Analytics Deployment

Geographical resilience can be provided by adding another Versa Analytics cluster. This cluster can have the same number of nodes as the main cluster or can have a different number of Analytics and Search nodes. A smaller sized cluster is typically a disaster recovery site that is only activated in the event of the failure of the primary site.

Important Note !
Automatic Database replication is only applies within the same cluster

Versa recommendation is to use a backup cluster as disaster recovery. In the event of the loss of the active cluster, the backup will continue to process the logs and stats. This could later be sync back to the active cluster, once the later becomes available again.

Where the disaster recovery site is not required to visualize data in analytics dashboard during the failure of the primary site, then a log collector ONLY disaster recover site would be sufficient. This ensures that no logs are lost for the duration of the outage.

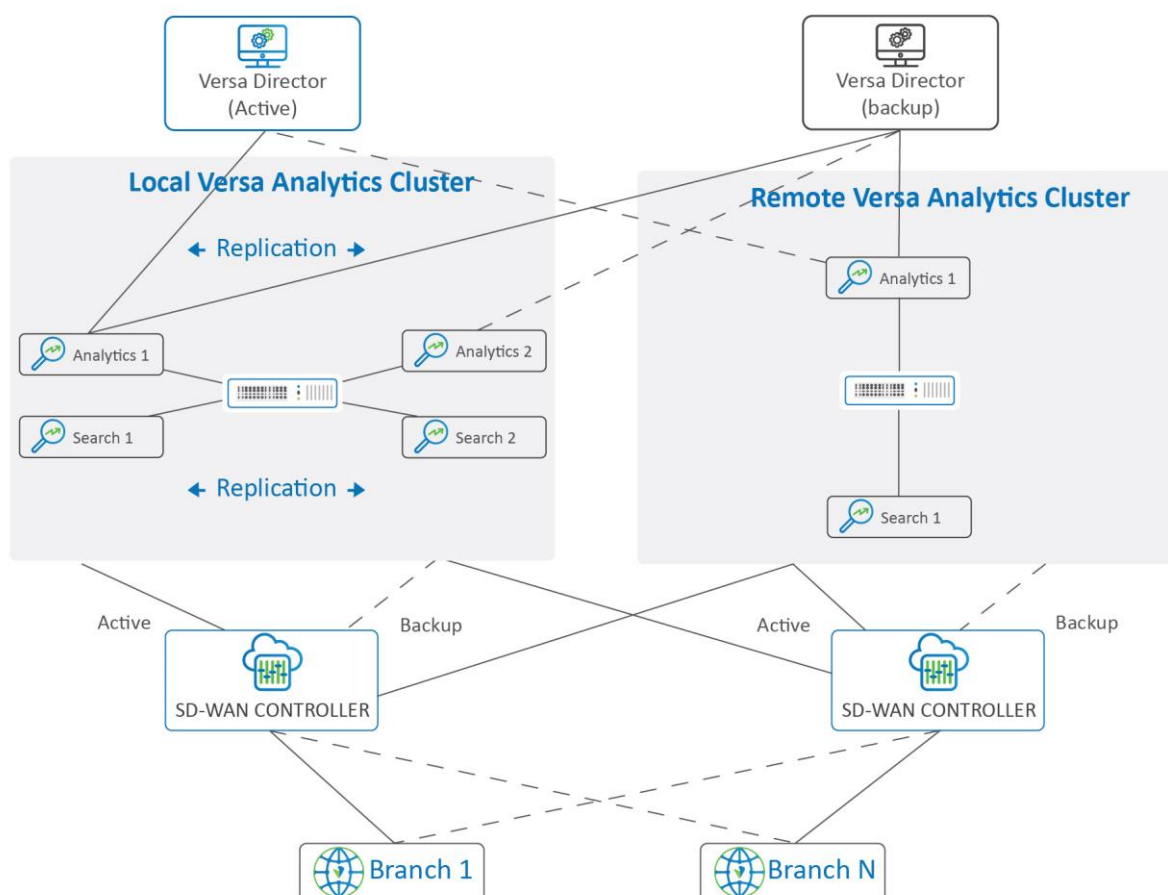


Figure 5 Recommendation for Geo HA

Important Note !

When different sized cluster are used. It is important to check that that the smaller Cluster can accommodate the number of LEF connections from the branches.

3.2.2 Role of the Log Forwarder

The VOS™ Edge Devices generate logs destined for the analytics database. These logs are carried in an IPFIX template and delivered to the Log forwarder. The Log forwarder will remove the IPFIX template overhead and stores the logs in clear text on the log collector disk.

The analytics database servers are then able to consume the logs from the log forwarder. Optionally logs can be forwarded to external 3rd party log forwarder in syslog format.

Log forwarder function is part of the Versa Analytics node. It can be run in Versa Analytics node itself (i.e. log forwarder process on the Analytics node) or as stand-alone log forwarder, where a separate VM or server is dedicated to this role. Separating the log forwarder function is recommended for larger deployments as this provides better scalability of the Analytics platform. When the log forwarder is separated, it typically is deployed south of the Analytics / Search nodes as illustrated in the diagram below.

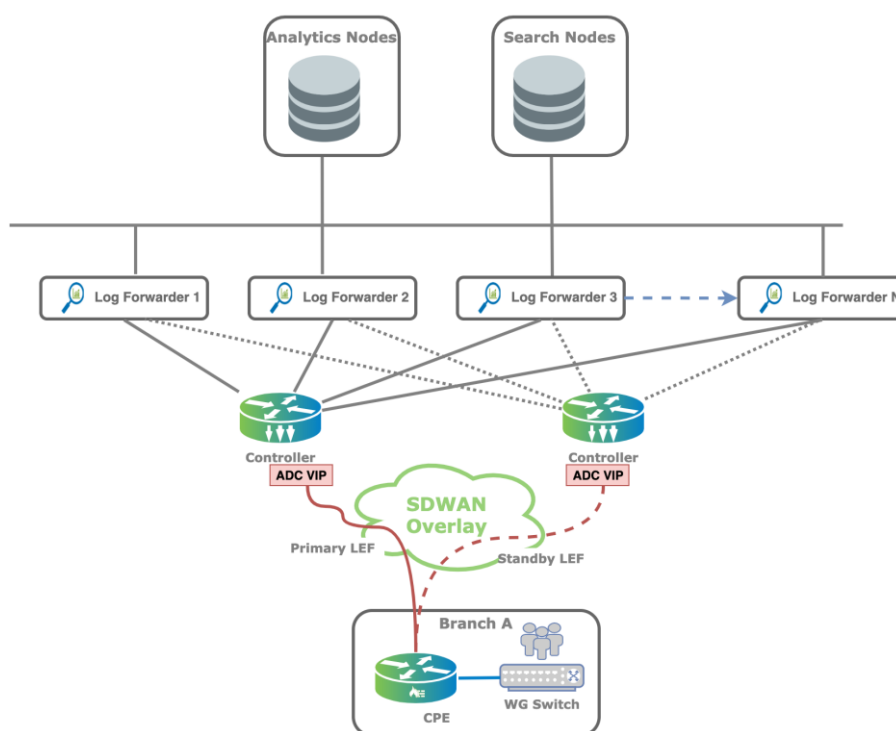


Figure 6 Analytics Log Forwarder

Important Note !

Versa recommends separate log forwarder for deployments larger than 1000 branches.

3.2.3 Best Practice for Versa Analytics Sizing

Versa Analytics scaling is driven by the logging configuration of the individual features enabled on the VOS Edge Devices. Most features include an optional logging configuration which may be enabled to log events related to a specific rule or profile. The number of features enabled with logging and the volume of logs each generates has a direct impact on the sizing / scalability of Versa Analytics.

Versa Analytics has two distinct database personalities i.e. Analytics Node and Search Node. The Analytics Node typically deliver the data for most of dashboards shown in the Analytics GUI. These is aggregated data and it is enabled by default. Therefore, the Dashboards are populated by even without enabling any configuration on the VOS Edge Devices. Analytics node scaling is determined by granularity of the reported data. Typically set between 5mins - 15mins. The chosen topology is also relevant to the scaling of the Dashboard function. A full mesh topology with multiple underlays will generate more SLA logging messages compared with a hub-spoke on 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 Analytics.

Search node are primarily tasked to store logging/event data. The search nodes drive the log sections as well as the log tables in Dashboard sections. These are typically heavily utilized as every event in the network triggers a log entry into the search node. In a minimal default configuration, the only information that is logged to these search nodes are alarm logs. Sizing of the search node is depending on the following:

- Number of features enabled with logging.
- Logging data volume and log rate.
- Retention period of logged data.
- Packet Capture¹

Feature like Firewall logging and Traffic Monitor logging (NetFlow) have significant impact to the overall scaling of the Analytics platform. Therefore, it is recommended to avoid creating wildcard logging rules.

Both Analytics and Search nodes may be scaled horizontally in response to increased analytics load. Versa Networks can provide a calculator tool to help sizing an Analytics cluster based on input information described above.

¹ Packet Capture and Traffic monitoring should only be enabled for troubleshooting specific traffic. Packet captures are note stored in the database but are stored as pcap files. These can quickly fill the disk space if not used correctly.

3.3 Best Practice Headend Design

A stable and fault-tolerant solution requires a proper headend design. Such a design factors in the expected utilization to dimension of compute resources with the ability to horizontally scale out when required. The headend must also have a taught through high-availability design, without a single point of failures. Finally, the headend must be secured and hardened to avoid any possible security vulnerabilities.

Versa recommends the following considerations are made when deploying a Versa headend.

- Ensure DNS is configured on Versa Director and Analytics. This is used for operations such as SPACK download as well as reverse lookup on Versa Analytics.
- Ensure NTP is configured and Synched across all nodes in the SD-WAN network. Versa recommends configuring all nodes into the same time zone e.g. UTC+1. This makes log correlation between various components practical.
- Platform hardening procedures such as signed SSL certificates, password hardening, SSL Banners for cli access.
- Ensure the latest OSPACK pack is installed in all components. Similarly, ensure that the latest SPACK is installed on the Director.
- Ensure that the appropriate ports are open on any intermediate firewalls.

Refer to the Headend hardening guide for further details.

Versa Networks Professional Services can assist in the design and validation of the headend before taken it into production.

4 Branch Deployment

Versa Networks SD-WAN solution offers flexible and comprehensive branch deployment options. The most common transport-based deployment scenarios are presented in this section.

One of the benefits with Versa SD-WAN branch configurations is that each Edge Device provides the same feature capabilities, regardless of whether it is in a hub configuration, spoke configuration or any other configuration. The Edge Device may have different configurations of topologies for different tenants on the same device at the same time.

4.1 Single CPE Branch with Internet / MPLS Transport

In the first scenario, a single device which has two different WAN links is located on the customer site. As depicted in the figure below a dedicated MPLS connection and an internet connection are terminated on a single CPE.

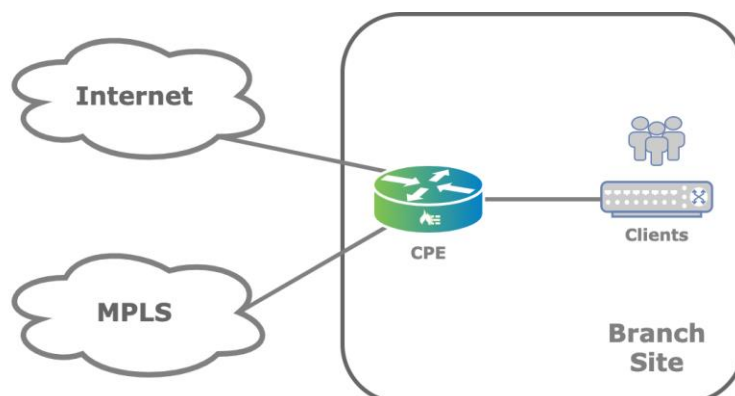


Figure 7 Single CPE with Dual WAN Link

In this topology, overlay tunnel between branches are formed through MPLS or internet underlay. This scenario be created easily using the Workflow function in Versa Director, even with extension using alternative underlays from different providers such as LTE connections.

A single CPE Branch configuration can be deployed for a single organization or as multi-tenant. The Edge Device provides full segregation between tenant or organizations.

Every organization in Edge Device configuration supports micro-segmentation using VRF's.

4.1.1 Active-Active Dual CPE Branch with Internet / MPLS Transport

In this scenario, the MPLS link is connected to one CPE-2 and the internet link is connected to CPE-1. The two paired CPEs devices are connected to each other using a cross connect link. The below is an illustration of the Dual CPE with WAN circuits terminated on individual CPEs.

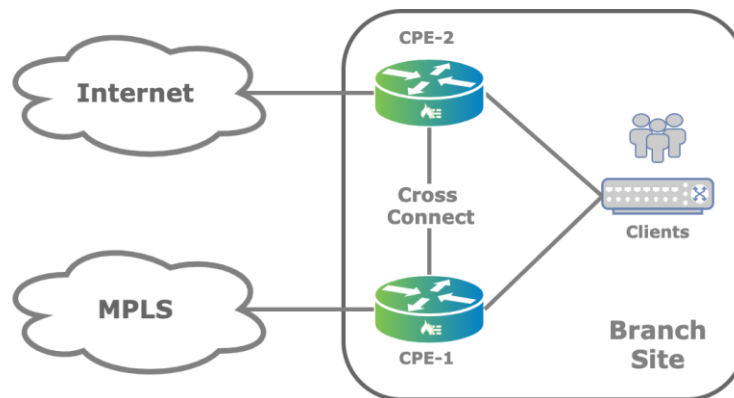


Figure 8 Dual CPE with Dual WAN Link

On the LAN side, VRRP protocol provides gateway redundancy. If there is no Layer2 reachability on the LAN, L3 routing protocol to existing LAN side routers may also be possible.

Each CPE of the HA pair maintains overlay connections to all other CPE's WAN transport. When the second underlay is physically attached to the other CPE, it is still logically represented on the local CPE using the cross-connect link. Using this approach each CPE can still be provisioned with SD-WAN policies leveraging all underlays.

This configuration is called Active-Active High Availability, since both CPE's are always carrying traffic to the underlay transport.

Important Note !

There is no state sync between Active-Active HA CPEs. This includes NAT and Session State information.

4.1.2 Best Practise Summary

- If security is required, then NGFW must be enabled on both CPEs. This is because the NGFW/SFW modules form part of the LAN service chain while the transport interfaces including the cross connect link belong to the Transport-VR.
- Always enable next-hop monitor to upstream WAN transport gateway alternatively, a dynamic routing protocol may also be enabled.
- Default workflow configuration is adequate.

4.2 Dual CPE Branch with Internet / MPLS & LTE Transport

The third scenario shows another form of two-CPE and multiple-wan branch deployment. As it is depicted in the Figure below, the CPE-2 is configured as multi-homed CPE which has the Internet and LTE connections. The CPE-1 is only connected to the MPLS transport domain. This is typically used where LTE is required as a backup to the fixed links.

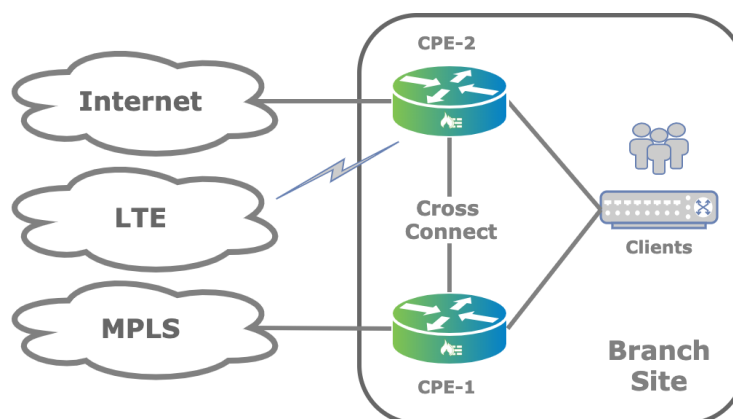


Figure 9 Dual CPE with Triple WAN Link

When the branch CPE devices are deployed in active-active mode then LAN side remains the same as [Section 4.1.1](#) above. This scenario is also valid for other combination of WAN links up to the Max of 8 WAN links per Tenant per CPE.

4.2.1 Branch High-Availability

Versa Branch High-Availability (HA) implementation requires two VOS Edge Devices to be present on the site. In order to provide branch redundancy Versa supports two modes of operation – Active-Active (stateless) and Active-Standby (stateless and stateful).

Most used HA is the Stateless Active-Active mode. The key benefit is that it is simple to deploy and can provides more on-site performance during standard operation due to the fact that both VOS Edge Devices are capable of processing traffic at all times.

State-less Active-Standby configuration may also be possible. In this configuration, each CPE has both underlays physically connected. The cross-connect link is not required. Each CPE can be configured as a standalone CPE with VRRP on the LAN side.

There are a couple of drawbacks;

1. Without the cross-connect each CPE cannot take advantage of the pair WAN transport, alternatively a layer 2 switch can be introduced on the WAN side to allow each CPE to have access to both WAN circuits.
2. Stateful connections are lost and sessions shall be re-established during a CPE failover (in most cases, end users hardly notice this TCP session re-establishment but this can be problematic for certain applications.).

Stateful Active-Standby HA on the other hand maintains a state sync between the Edge Devices. Only the active appliance can forward traffic. This type is useful where state is important such as NGFW and CGNAT traffic.

Please see comparison between HA modes:

	Stateless Active – Active	Stateless Active- Standby	Stateful Active - Standby
Overview	<p>Easy from configuration perspective using workflows.</p> <p>No additional links on WAN needed</p> <p>Cross-connect link between two CPE's</p>	<p>Easy from configuration perspective using workflows.</p> <p>Manual optimization configurations needed for VRRP tracking.</p> <p>All underlays physically connected on both CPE's.</p> <p>No cross-connect link needed.</p>	<p>Requires manual configuration and physical setup.</p> <p>Short flows and UTM/URL inspected flows are not re-evaluated – allow/drop after failover.</p> <p>Requires additional uplinks/switch (another active element lowering site MTBF)</p>
Use case	Stateless routed traffic	Stateless routed traffic	Use where State preservation is important such as FW, DNAT etc.
Complexity	Simple via Workflow	Simple via Workflow with some manual config extras	Manually configured
Available underlays	Local and remote uplinks via Crosslink	Local uplinks only	Local uplinks only
Performance	No impact by synchronization	No impact by synchronization	Requires sync of the Control and the Working Threads from Active to Standby. Unmeasured performance impact.
BGP/IPsec/SLA monitoring scalability and state	Minimum one SLA per circuit per device.	Minimal one SLA per circuit per device (double number SLA as all underlay circuits are connected to both CPE's)	Minimal one SLA per circuit per device (double number SLA as all underlay circuits are connected to both CPE's)
Convergence	<p>Upstream: 3 seconds by default (configurable VRRP, and other timers). Can be configured lower.</p> <p>Downstream from remote branch: SD-WAN</p>	<p>Upstream: 3 seconds by default (configurable VRRP, and other timers). Can be configured lower.</p> <p>Downstream from remote branch: SD-</p>	<p>Upstream: Few seconds by default (based on VRRP, Quorum Probes, BFD and other timers)</p> <p>Downstream from remote branch: SD-WAN Control</p>

	Control plane (MP BGP) + SLA Probes	WAN Control plane (MP BGP) + SLA Probes	plane (MP BGP) + SLA Probes
Traffic restoration	All session shall be restarted.	All session shall be restarted.	Long-term sessions are restored. Short-term sessions might be non-synched and restarted
Synchronized services between Primary and Secondary node	Not applicable	Not applicable	Data plane state incl. sessions Control plane state traffic steering table NAT bindings ADC persistency
None Synchronized services	All	All	AV, IDS/IPS, URL inspected flows after failover fail OR pass without security inspection if “sync flow allow” is set.

4.2.2 Cross-Connects in Active-Active HA

A typical layout of Active-Active branch is illustrated in the diagram below. This active-active HA topology has the benefit of using both WAN Transport underlays without having to dual-the underlays on the CPEs.

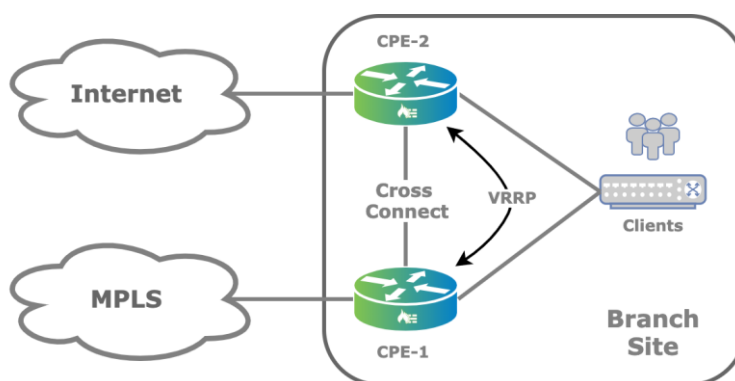


Figure 10 High Availability Active-Active

The Cross-Connect link is a physical connection between the redundant CPEs which emulates the missing transport domain on a given branch and provides redundancy to the attached clients.

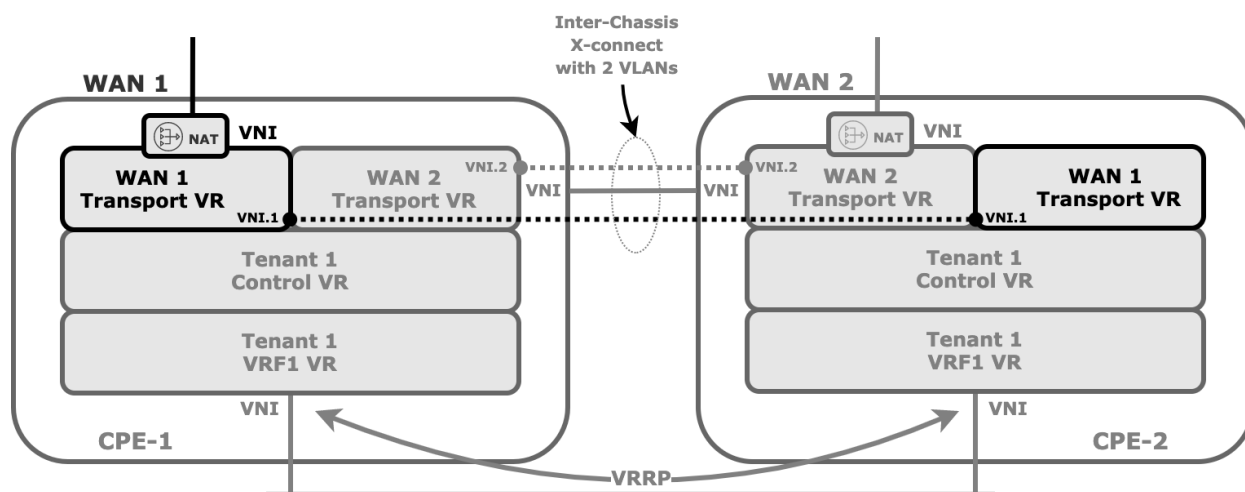


Figure 11 HA Pair in detail

The cross-connect link is VLAN tagged for each WAN Transport VR instance and IP addresses configured using workflow template.

As the WAN Transport VRs are completely distinct routing instances, it allows for IP addressing re-use.

When HA is enabled, by default the back-to-back logical interfaces on the cross-connects are assigned IP addresses from range 172.16.255.0/30, 172.16.255.1 for primary CPE and 172.16.255.2 for secondary CPE. The order of primary or secondary is determined by which device is using the primary device template configured in the workflows. The secondary device-template is auto-generated by Versa Director Workflow.

Below is default interface configuration for primary device (CPE-1):

Name	Description	Interfaces	IP Address/Prefix
<input type="checkbox"/> vni-0/0		vni-0/0.0	10.40.36.50/16
<input type="checkbox"/> vni-0/1		vni-0/1.1 vni-0/1.2	172.16.255.1/30
<input type="checkbox"/> vni-0/2		vni-0/2.0	192.168.1.2/24

Figure 13 : Default HA interface configuration of CPE-1

Below is default configuration for secondary device (CPE-2):

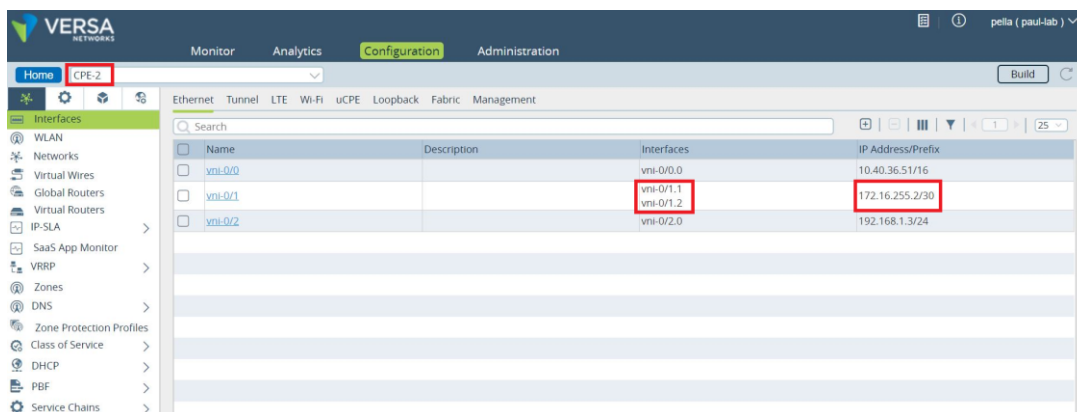


Figure 14: Default HA interface configuration of CPE-2

Static routes are configured within the transport VR of each CPE. Via the workflow, ICMP monitoring is configured within transport VR associated to cross-connect link, in order to direct traffic destined to WAN connection of the paired CPE.

In the event of any IP reachability issues over the cross-connect interface due to cross-connect interface going down or even other peer CPE device lost, the static route is withdrawn from the routing table of that corresponding WAN Transport VR.

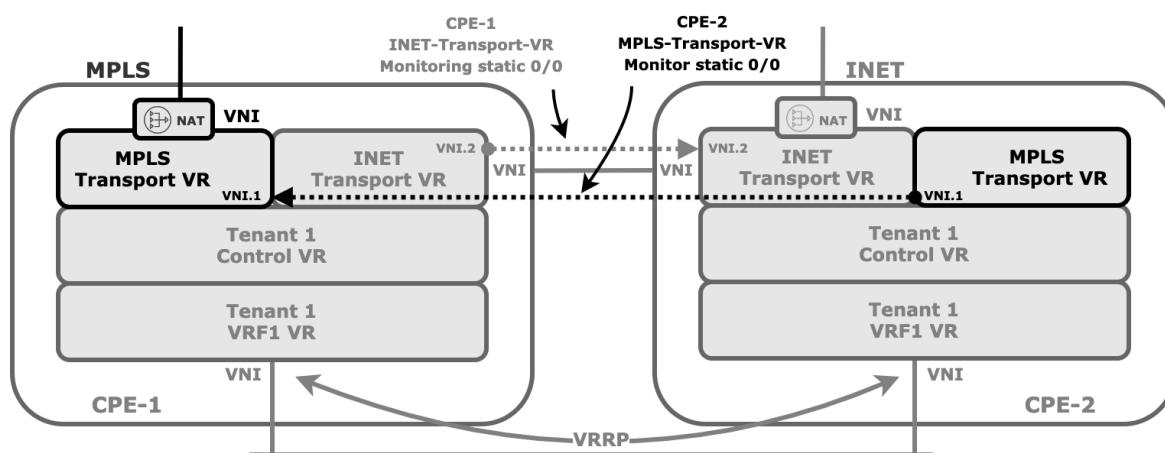


Figure 12 Static Route ICMP Monitoring on HA

ICMP Monitoring CPE-1 connecting MPLS transport-VR:

```
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
```

ICMP Monitoring CPE-2 connecting INTERNET transport-VR:

```
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, VRRP is used to elect a master/slave. The logical interface and Virtual IP address is used as the Next-Hop/gateway on the LAN.

4.2.3 DIA on Active-Active HA

Direct Internet Access (DIA) can also be configured in active-active HA scenario with specific considerations. When DIA is enabled, Versa Director workflow automates configuration of BGP peering between Transport-VR and LAN-VR to propagate default route.

The ICMP monitoring is between the cross-connect logical interfaces of CPE-1 INTERNET-Transport-VR to CPE-2 and therefore does not protect against INTERNET WAN link failure on CPE-2. This may result in local Internet blackhole scenario when INTERNET WAN link fails on CPE-2 as described in figure below.

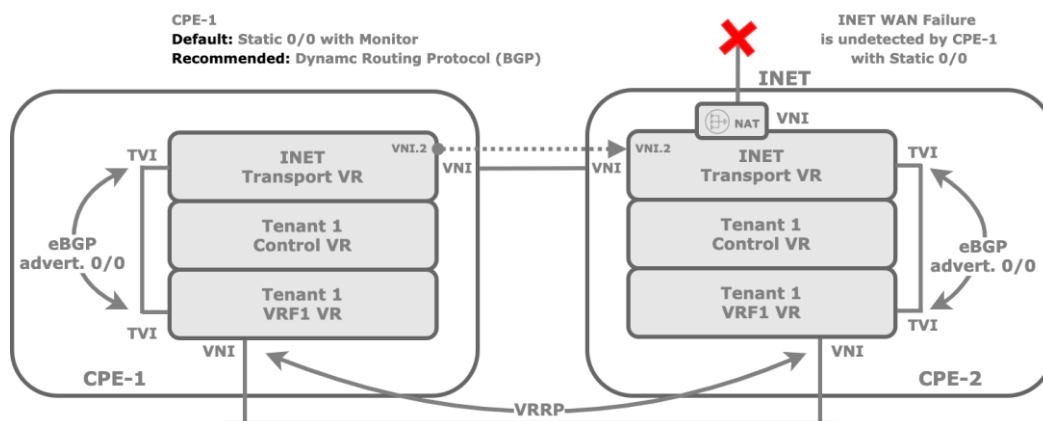


Figure 13 DIA on Active-Active HA

Additional measures must be configured to protect against this blackhole, such as Next Next-Hop (remote next-hop) monitor, at the expense of additional complexity as NAT may also be required to allow the Internet provider WAN interface to reply back to ICMP echo requests since ICMP will be sourced from 172.16.255.0/30 range and not necessarily routed back by Provider router.

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

For additional details on HA Active/Active topology please refer to:

<https://support.versa-networks.com/support/solutions/articles/23000020295-ha-active-active-deep-dive>

5 LTE Transport

The VOS Edge Devices supports LTE as a WAN link and may be used in just the same way as any other WAN link. However, in many use cases LTE is deployed as a backup path due to its relatively higher data costs.

The LTE as a backup transport link can be configured in the following modes:

- Hot Standby Mode
- Cold Standby Mode

5.1 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. In other words, the LTE link is used to send traffic only when the primary wired wan links are down or out of SLA compliance. Management plane traffic like LEF logging information or branch software uploads can be configured to avoid the LTE link if other wired wan links are available.

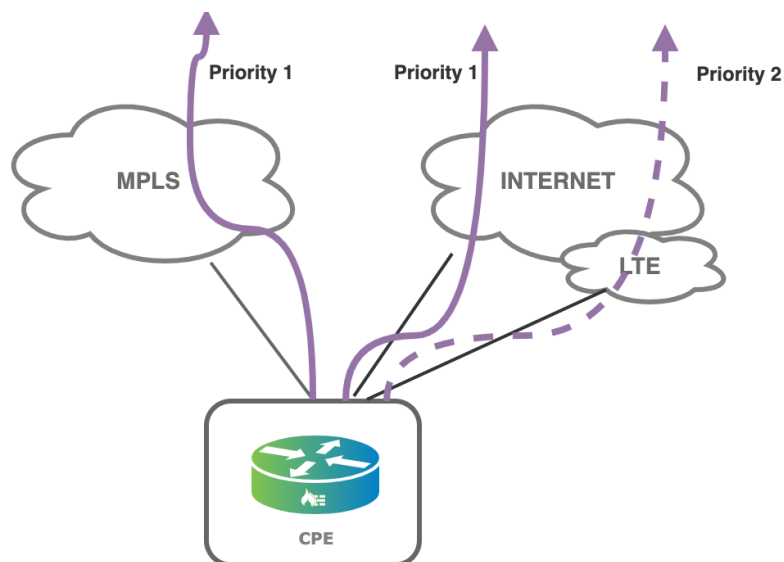


Figure 14 LTE Hot Standby

Benefits:

- Switchover to LTE is instantaneous when the primary wired WAN interfaces are down or out of SLA compliance.
- SLA based steering is possible because the LTE Link is actively monitored via SLA probes.
- The status and quality of the path is known at all times in standby mode, the LTE link can generate alarms if the path becomes unavailable prompting the administrator to take corrective action.

Limitations:

- The paths through the LTE link is actively monitored. Therefore, SLA traffic will still be sent over this link even in standby mode, consuming credits of the LTE data plan subscription.
- The amount of bandwidth consumed by the SLA traffic will vary depending on the number of sites in the network where to maintain an SLA peering with. Also, the SLA probe interval influence the amount of data generated over LTE.

Configuring LTE Hot Standby for SDWAN VPN Traffic (Site to Site)

The WAN links in Versa has a configuration attribute associated with it called “circuit-media”. The circuit-media can be any of the following

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

The following scenarios are described:

1. Use LTE as backup circuit on Local and Remote device.
2. Avoid LTE for non-Business and Scavenger traffic.

5.1.1.1 Use LTE as backup circuit on Local & Remote branch

The configuration describe in this section applies for scenarios where LTE is used. The described configuration should be applied to all devices in the networks (with LTE or without LTE interfaces), when Hot Standby LTE links is required.

The local branch decides what path to use to connect to the remote branch. This includes the WAN circuit on the local branch and the WAN circuit on the remote branch. In many networks, not all branches are deployed with LTE interfaces. This means that a non-LTE branch should not use the path to the remote LTE interface if the remote wired Internet interface is still available. The following section will cover the configuration needed to make the local and the remote branches use LTE only when needed.

The WAN link **Circuit-Media type** will be used to make the LTE link as hot-standby. First verify if the Circuit-Media type is correctly set for all the WAN links. In configuration, go to the *Services* → *SDWAN* → *System* → *Site Config* and edit the wan interfaces. Then ensure that the Media type is correctly set to Ethernet, DSL or LTE based on the type of WAN link. Note that Ethernet interfaces will be displayed as vni-0/0, vni-0/1 and LTE interfaces will be displayed as vni-0/100, vni-0/101.

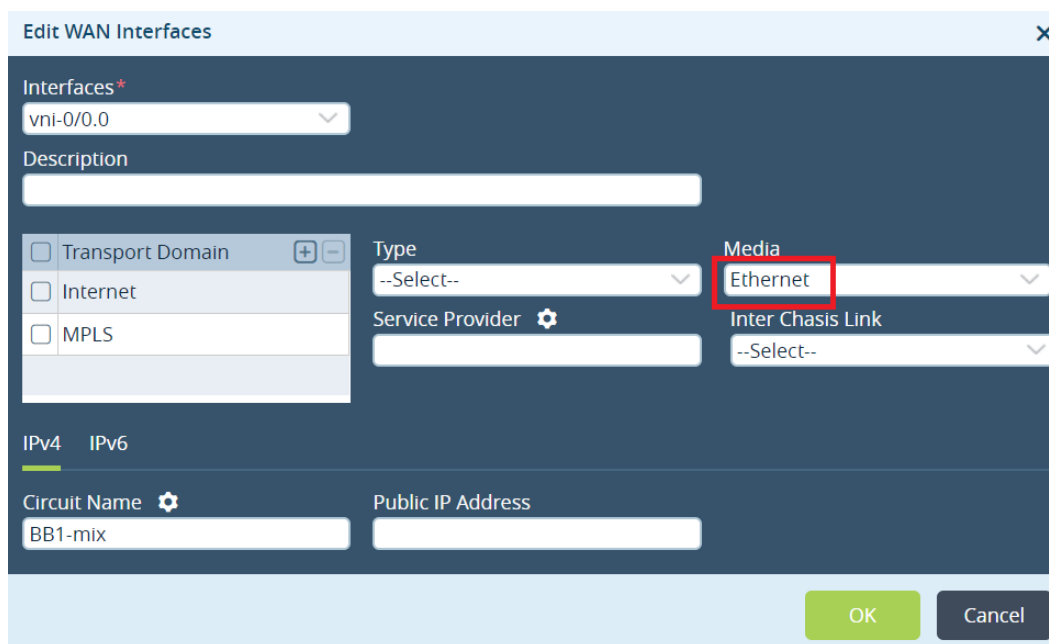


Figure 15 Configure Media typo for WAN link

The LTE interface is set as Hot-Standby on the local and remote branches using the circuit priority. Since the branch that initiates the traffic decides the path, it is important that the branch does not prioritize the LTE link on the remote branch to send traffic to. This configuration will apply also when the sending branch does not have an LTE circuit e.g. communication to a datacenter branch.

Priority	Local Circuit Media	Remote Circuit Media
1	Ethernet, DSL	Ethernet, DSL
2	LTE	Ethernet, DSL
3	Ethernet, DSL	LTE

In the SD-WAN Forwarding Profile, the Circuit Priorities need to be defined per above table. Typically, this is done in the Default-FP.

The Priority group 1 will cover the scenario when the Local and remote wired WAN links are up.

The Priority group 2 will cover the scenario when the wired WAN links on the branch are down and the only available wan link is LTE.

The Priority group 3 will cover the scenario when the wired wan links on the remote branch are down and LTE is the only available wan link.

Priority 4 does not have to be defined as it is implicit, but it will cover the scenario where LTE is the only wan link available on both the local and remote branch.

If the Local and Remote Circuit media has other type of WAN links like T1/E1s, then they can be mentioned alongside with Ethernet in the all the Priority groups.

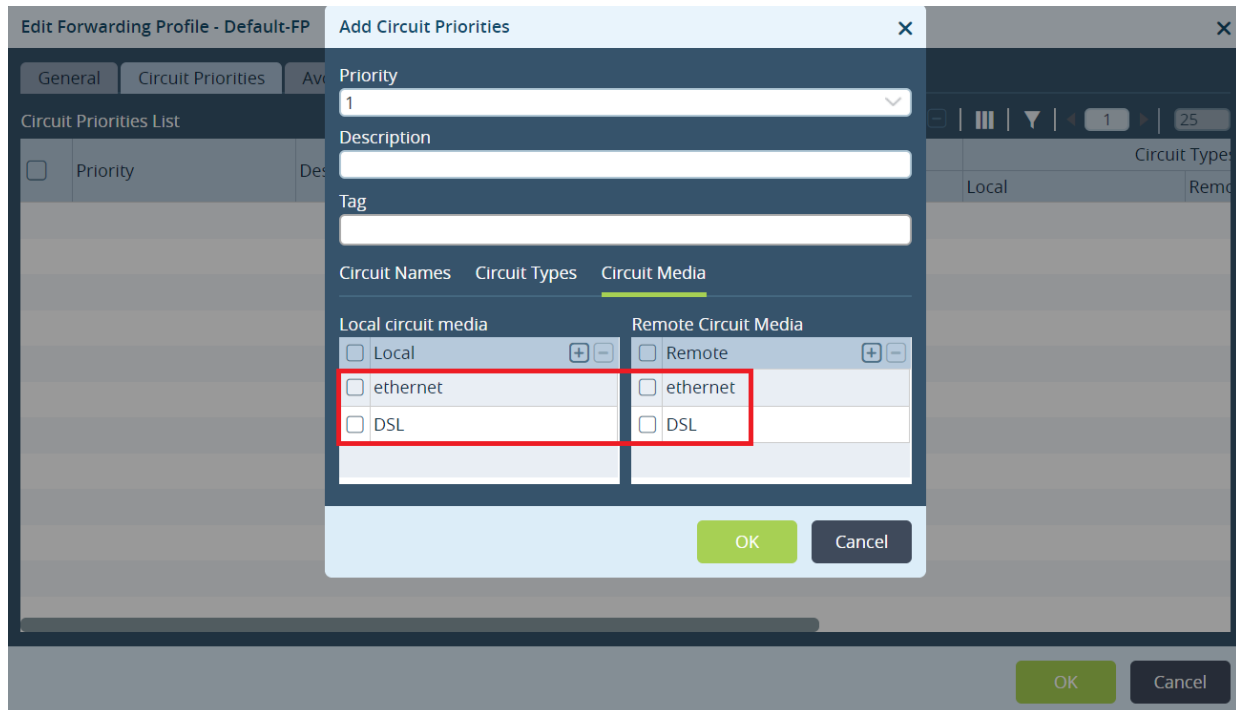


Figure 16 Configure Local and Remote Circuit Priorities

Then configure the SDWAN Policy Rule to match the traffic that should follow the wan link order of priority. We will create an SD-WAN Policy rule that will match all traffic (wildcard match). Go to *Services* → *SDWAN* → *Policies* and create a new Rule. For a wildcard catch all rule, keep all match conditions blank and in the Enforce option, set the SD-WAN Forwarding Profile Default-FP that was created earlier.

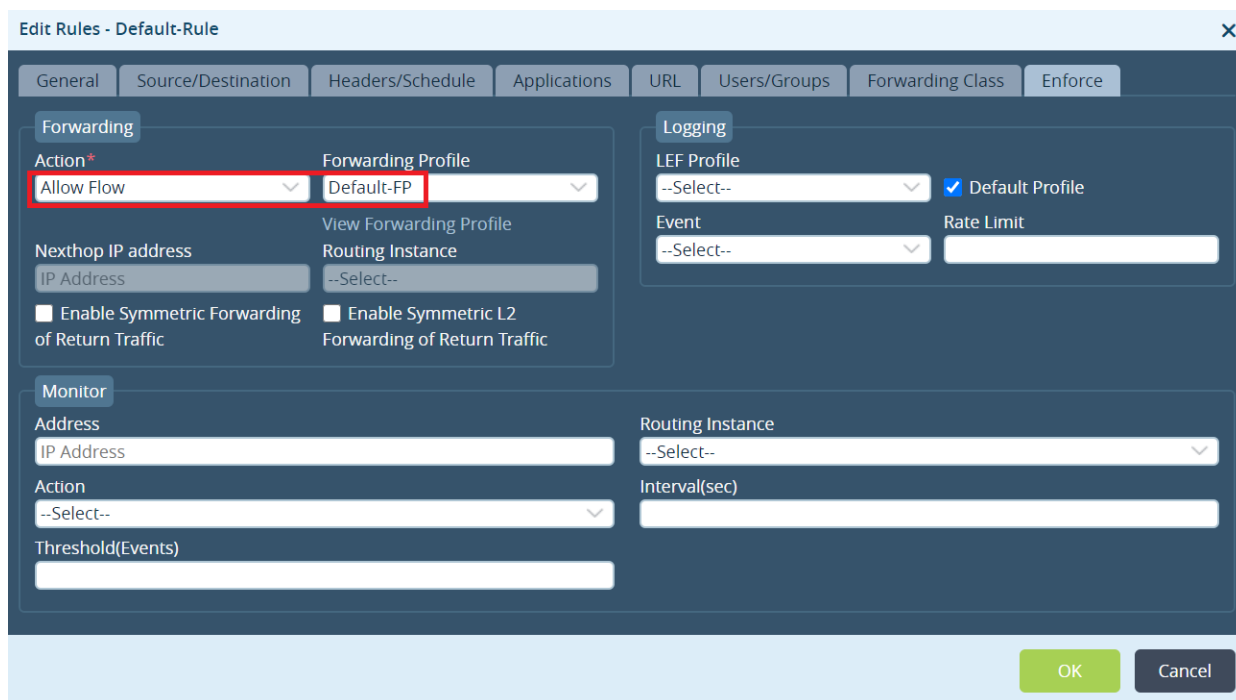


Figure 17 Configure SDWAN Policy Rule

5.1.1.2 Avoid LTE for non-Business and Scavenger traffic

The system can be configured to completely avoid the usage of LTE circuits for Non-business critical traffic and Scavenger traffic. The configuration can be done by using only the Local circuit media type or a combination of Local and Remote circuit media type as discussed in the previous section. The only difference is that this configuration will be done in the “Avoid Connections” option under the SDWAN Forwarding Profile. When a circuit is defined in “Avoid Connections” it will not be used even when this is the only path available.

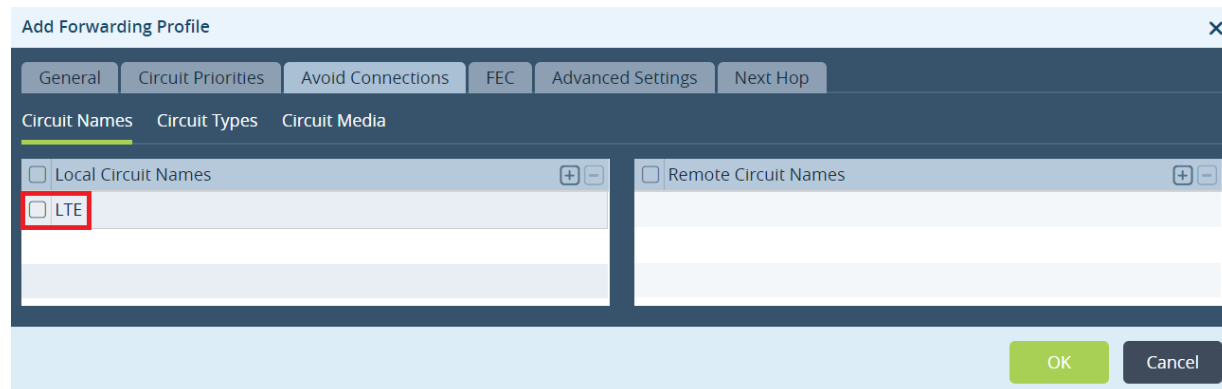


Figure: Block non-Business critical traffic over LTE in Forwarding Profile

Now a SD-WAN Policy rule should be created to classify the scavenger traffic and attach this to the Forwarding Policy that specifies to never use LTE.

Important Note !

Forwarding profile configuration affects egress traffic decisions. For effective implementation, this configuration must be applied uniformly across the network to prevent ingress traffic arriving on the LTE Link.

5.1.2 LTE as backup for Internet Traffic

This section describes the configuration for the LTE link as backup for local internet breakout (DIA) traffic. We will cover scenarios where there is one wired link and one LTE link or a scenario where there are more than one wired link and an LTE link.

To use LTE as Hot-Standby for Direct Internet Access (DIA) traffic, configure the Workflow Template Tunnels as shown below with the Load Balance option is unchecked .

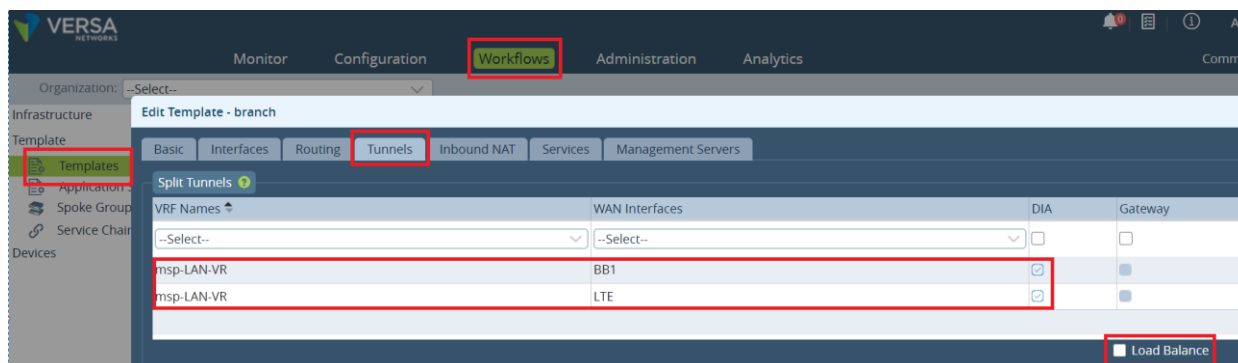


Figure: LTE in Hot Standby for DIA Traffic

The workflow creates two BGP sessions over a virtual interface pair that is created automatically between the respective “WAN-Transport-VR” and the “MSP-LAN-VR”. The default route is advertised from each of the WAN Transport VR’s to the LAN VR. The default route from the BB1 Transport VR comes with BGP Local Preference of 120 and the default route from LTE Transport VR comes with a BGP Local Preference value of 119. Based on this, the LAN VR will install the default route from BB1 Transport VR as the active default route making BB1 wan link as the primary wan link for local internet bound traffic.

5.1.3 Load Balance two wired WAN links plus LTE backup

Enable all Internet WAN links for DIA and check the load balance knob. This will set all WAN links to the same local preference.

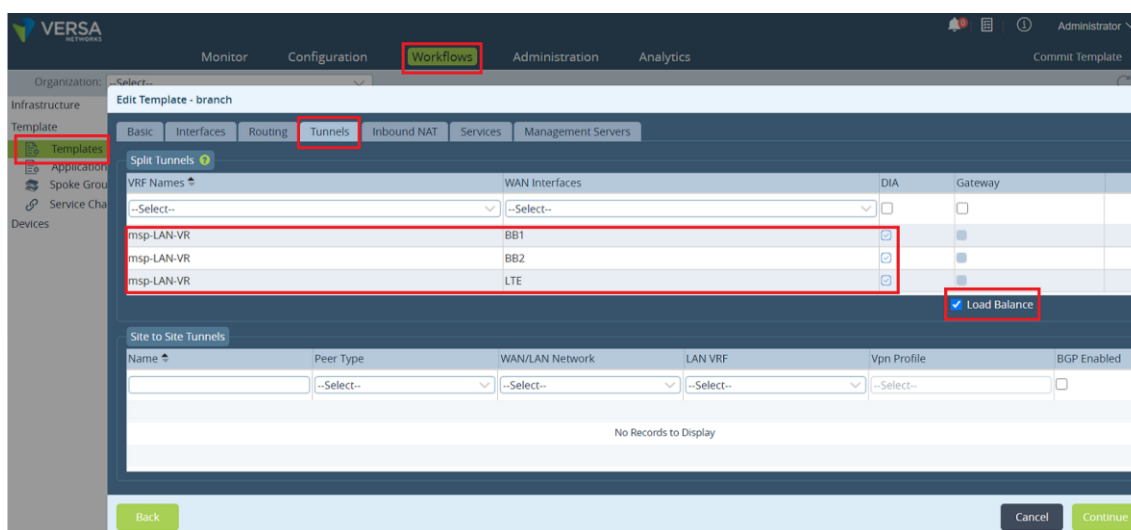


Figure 21: Load Balancing in Workflows

Since the Load Balance option is set, all the default routes from the BB1, BB2 and LTE Transport VRs will be advertised to the LAN VR with a BGP Local preference value of 120. This will cause the traffic to be load balanced across all the 3 WAN links. To make the LTE link as the backup and only active when the two wired WAN links are down, the BGP local preference of the advertised LTE-VR sourced default route to 119 needs to be modified.

Configuration path:

Services → *Virtual Routers* → *BGP* → *Peer/Peer Group* → *From_ST_LTE Policy* → *Color_ST_Route* → *Action*.

Change the Local Preference to 119 or any other value lower than 120.

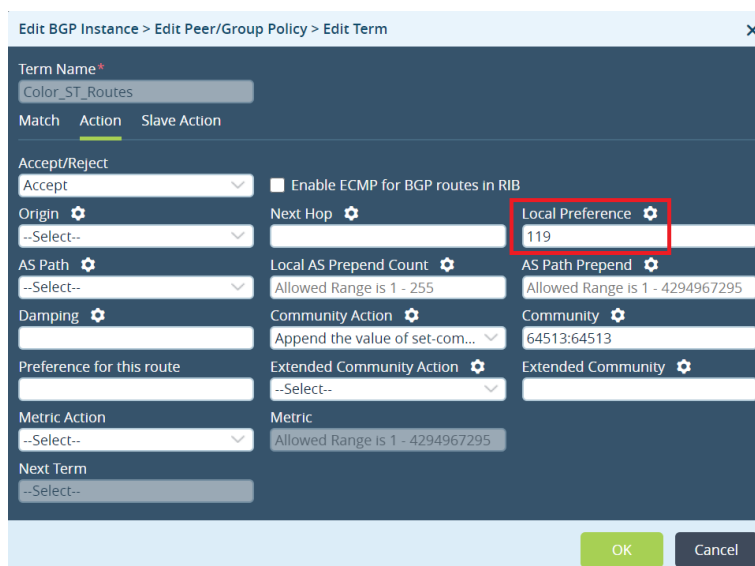


Figure 18 Load Balancing by modifying BGP Metrics

5.1.4 Management Traffic Priority

Like any other link, the LTE can also be used to provide connectivity to the controllers/headend. By default, the system will randomly assign an available link. To prevent unnecessary use the LTE data plan, Management traffic (all traffic towards the controllers/headend) like LEF logging information to Analytics or Software file upload should be preferred over the wired WAN links. To make this happen, the Management Priority of the LTE wan links should be set to be lower than the wired wan links (0-highest, 15-lowest). This can be configured under the site configuration under the SD-WAN section. Click on the option "Site" and select the wan interface where the Management priority can be set.

The screenshot shows the 'Edit WAN Interfaces' configuration window. At the top, there are dropdown menus for 'Interfaces*' (set to vni-0/100.0) and 'Encryption' (set to --Select--). Below these are three main sections: 'Shaping Rate', 'Management Traffic', and 'SLA Monitoring Policy'. The 'Management Traffic' section is highlighted with a red box, showing a 'Priority' field with the value '15'. The 'SLA Monitoring Policy' section shows 'SLA Monitoring' set to 'SLAM_Policy_vni-0/100.0'. The 'Bandwidth Monitoring Policy' section shows 'Bandwidth Monitoring' set to '--Select--'. At the bottom right, there are 'OK' and 'Cancel' buttons.

Figure 23: Configure Management traffic priority away from the LTE circuit

5.2 Cold Standby Mode

In this mode, the LTE link is configured but in an admin Down state link. The link state goes to UP only when all the Primary wired WAN interfaces are down.

5.2.1 Benefits of Cold Standby Mode

No data used on the LTE link till all the other primary WAN links are down.

5.2.2 Limitations of Hot Standby Mode

The VOS Edge Device does not know the actual status of the LTE connection till the primary WAN links are down and the system tries to bring up the LTE connection.

The traffic failover is also not instantaneous due to the lag in registering the SIM to the mobile network. Data transfer can only begin after the Mobile Data context is enabled on the SIM and the versa control plane is established on that path.

5.2.3 Configure LTE Cold Standby

In the LTE Cold standby mode, the LTE interface is been brought to the “admin down” state if the wired internet interface is active and working. An active monitor (IP-SLA Monitor) is configured to the next-hop of this wired internet interface. Once this monitor declares the next hop is no longer reachable, the LTE interface is activated automatically and will start building neighbor ships with all its peers in the transport domain.

For configuration of Cold Standby mode, please refer to the following Knowledge base article on the Versa support site:

<https://support.versa-networks.com/a/solutions/articles/23000017724>

5.3 Comparison of LTE Hot Standby mode vs Cold Standby mode

	Hot Standby Mode	Cold Standby Mode
LTE link status when in standby mode	UP	DOWN
Egress Data traffic in Standby mode	Only SLA traffic	None
Ingress Data traffic	SLA traffic but possible also data traffic per configuration remote devices	None
Traffic Failover to Standby mode?	Instantaneous	Not Instantaneous
Alarm generated when LTE link is down when in Standby mode	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

6 SD-WAN Overlay

The reference topology below describes a datacenter and two single homed remote branches. All sites and Headend (Director, Analytics and Controllers) are attached to available transport networks.

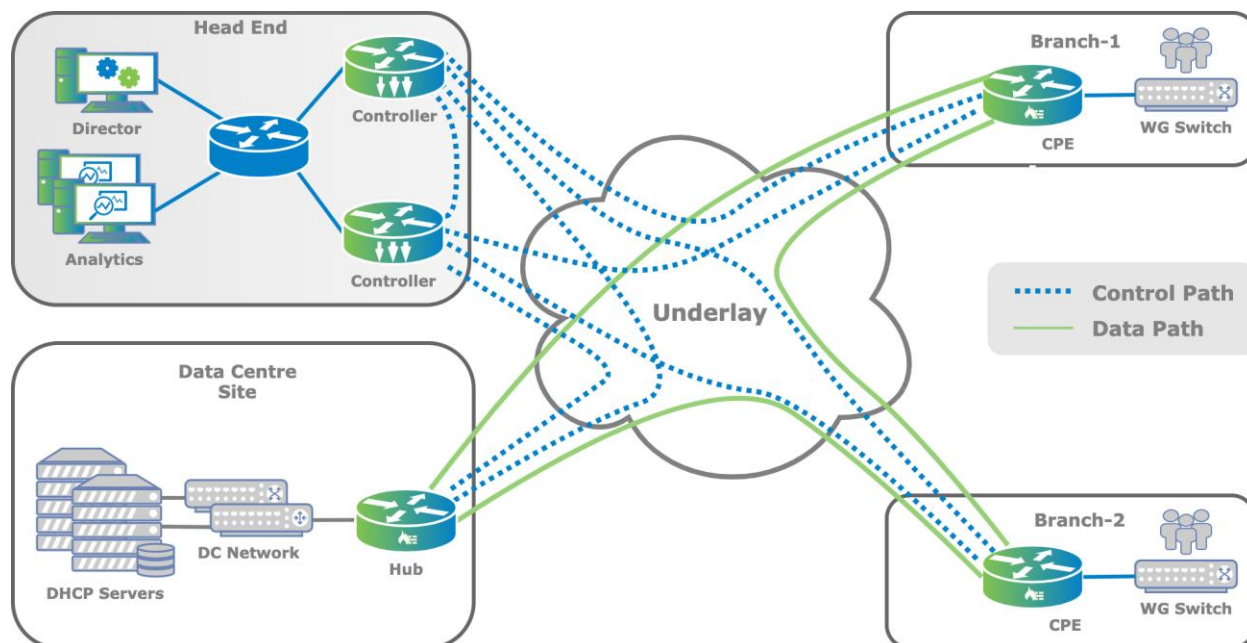


Figure 19 SD-WAN Network

Hubs and branches connect to the Versa Controller which serves as attachment point for 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 provides the flexibility to create multiple topologies through workflows without restrictions.

6.1 Overlay IP addressing

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

- encrypted overlay using IPsec tunnel
- plain-text overlay using VXLAN tunnel

For further reading refer to the document on the Versa docs portal:

[https://docs.versa-networks.com/Reference/Architecture/Secure Control and Data Overlay Tunnel Solution](https://docs.versa-networks.com/Reference/Architecture/Secure%20Control%20and%20Data%20Overlay%20Tunnel%20Solution)

The SD-WAN overlay tunnels are addressed using a specific overlay IP addressing scheme. In principle, the overlay network should only be routable in the SD-WAN control network i.e. between the controllers and the branches AND the control-network southbound of Versa Director (or northbound SD-WAN controller) as described below.

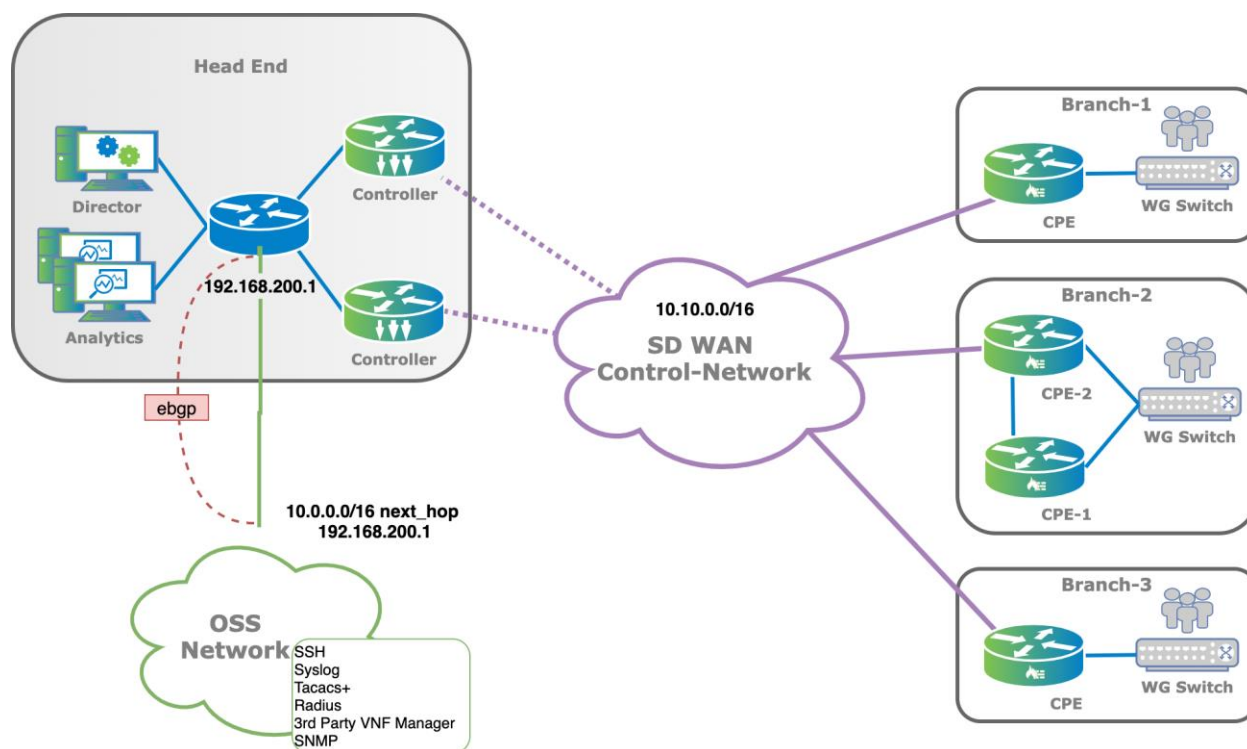


Figure 20 IP Overlay Addresses reachability

However, in certain deployments there are considerations to be made when choosing what an overlay IP address scheme should these include the Integration with OSS/BSS – the control network should have IP reachable between Edge Devices to services such as Tacacs, Radius, Syslog collectors, 3rd party VNF managers etc.

A knowledge base article on the Versa Support website provides an in-depth explanation of how the default overlay total address space carves out individual tenant address ranges.

<https://support.versa-networks.com/support/solutions/articles/23000010229-configuring-an-overlay-addressing-scheme-for-sd-wan>

6.1.1 Best Practice IP overlay addressing

The overlay IP addressing method and pool can to be configured during Versa Director initial setup and cannot be modified afterwards. Therefore, it is very important that the correct method and allocated subnet wide enough to cover the expected size of the deployment.

Versa recommends using “do not encode” option that optimizes the use of the IP space used for addressing.

6.2 Encrypted and clear-text overlay

This section considers use cases where it might be necessary to change the default tunnel used for data transport. By default, all traffic will always follow the encrypted overlay.

There are two options to achieve this:

- static definition per WAN interface
- dynamic definition per SD-WAN policy

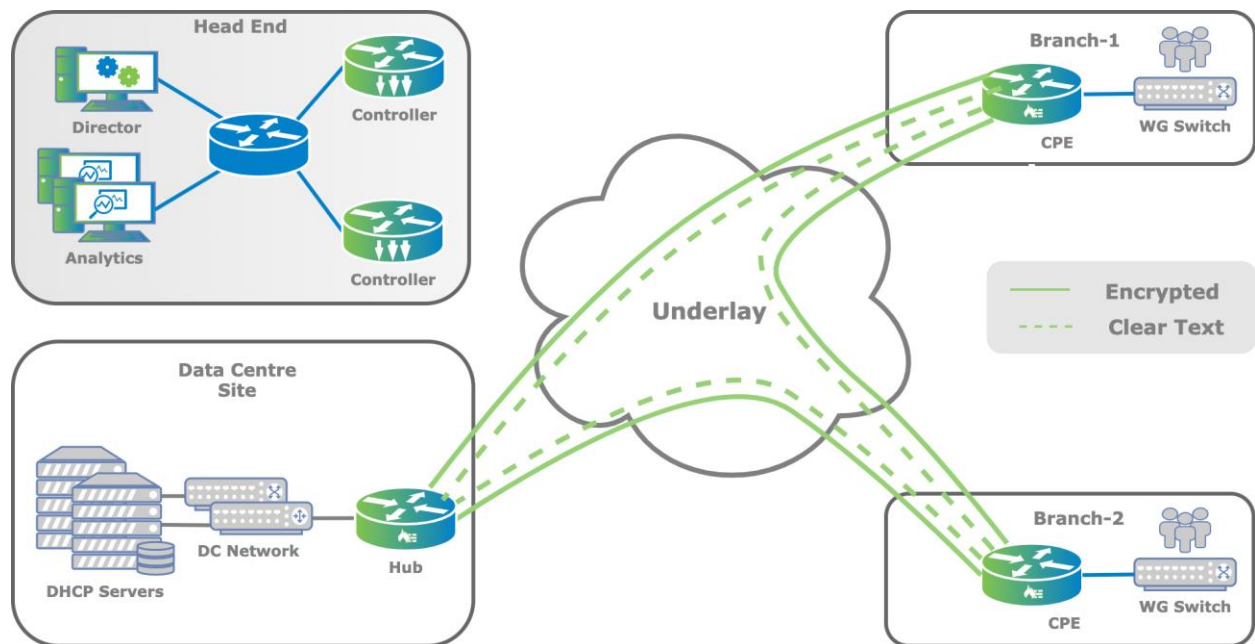


Figure 21 encrypted and clear text overlay

6.3 Static definition per WAN interface

This may be used if the underlay is a private/secured circuit such as MPLS service by a service provider L3VPN service. Traditionally, network administrators have considered these private L3VPN as secured and did not typically encrypt data over it. Consistently, MPLS L3VPN underlay may also be considered similarly and data can be transported via the clear text tunnel.

Benefit of Clear Text Transport is that it saves the platform IPsec overhead.

6.3.1 Configure Static definition per WAN Interface

Configuration path:

Configuration > Services > SD-WAN > Site > WAN Interface



Figure 22 Set Encryption per underlay

6.3.2 Dynamic Definition by SD-WAN Policy

Below are common reasons why encryption may be dynamically turned off for certain traffic;

- Traffic is already encrypted by application for example HTTPS or TLS/SSL secured application data.
- Traffic that is not of interest from a security point of view to the enterprise for example recreational traffic such as Facebook and YouTube.

6.3.2.1 Configure definition by SD-WAN policy

Configuration path:

Configuration > Services > SD-WAN > Forwarding Profiles

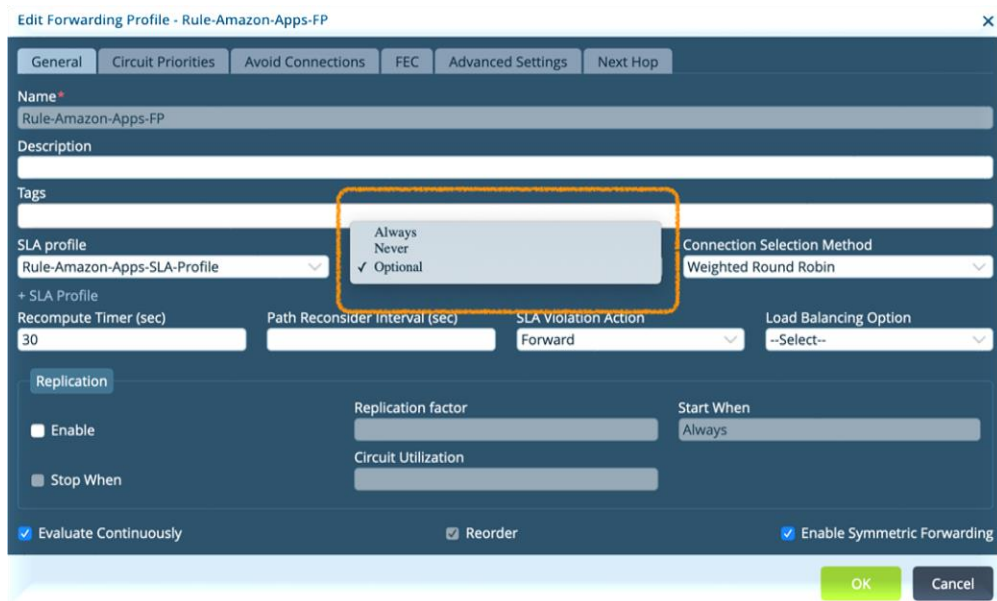


Figure 23 Set Encryption per protocol

SDWAN overlay traffic is classified with normal SDWAN policy rule and assigned to this SDWAN Forwarding profile. All traffic matching this SDWAN policy rule will follow the encryption settings for this Forwarding profile.

Important Note !
Where both WAN static definition and policy is configured, then SD-WAN Policy takes precedence.

7 SD-WAN Topologies

Versa solution supports the following SD-WAN overlay topologies:

- Full Mesh
- Hub-and-Spoke
- Regional Mesh
- Multi-VRF/Multi-Tenancy

The topologies are established by leveraging well-known routing techniques which have been used in MPLS-L3VPN networks for many years. Using MP-BGP communities, a fine grain control of route control can be achieved with flexible options to manipulate or fine tune by the network administrator. However, using workflows to create those topologies makes operations of such complex configuration simple.

Details of each supported topology are discussed in this chapter.

7.1 Full Mesh

A Full Mesh topology is used for any-to-any communication. In such topology branches will communicate directly using overlay tunnels without the need to transit through a Hub/centralized site.

The Full Mesh could be the preferred topology when branches need to communicate directly with each other, typically voice applications require a full mesh connectivity as compared to a Hub-and-Spoke topology where the hub side is usually distant from branches inducing delay. Another example where Full Mesh topology could be preferred is in a distributed security architecture where policy enforcement is performed at the branch. This removes the requirement to funnel traffic to hub sides for inspection.

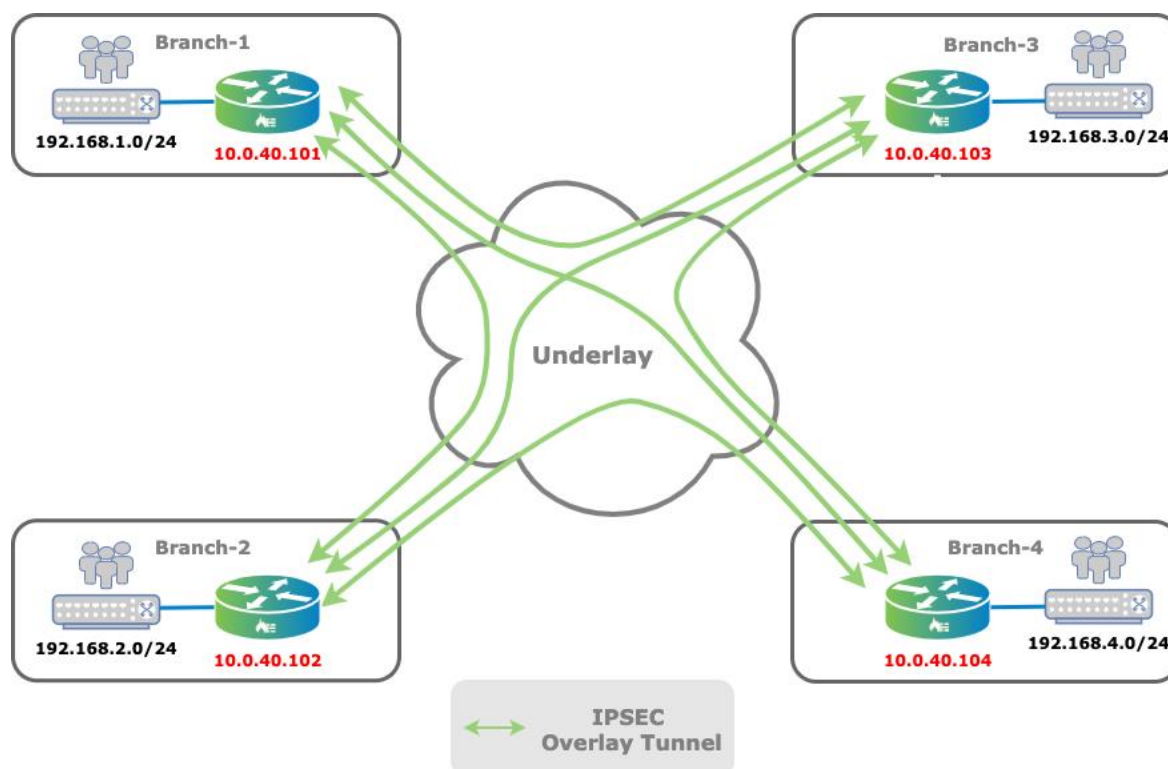


Figure 24 Full Mesh

In a Full Mesh topology probes SLA monitoring will be sent to every remote branch on every available transport.

```
admin@Branch-1-cli> show orgs org Tenant1 sd-wan sla-monitor status
```

SITE NAME	PATH HANDLE	FWD CLASS	LOCAL		REMOTE		ADAPTIVE MONITORING	DAMP STATE	DAMP FLAPS	CONN STATE	LAST FLAPS	LAST FLAPPED
			WAN LINK	WAN LINK	WAN LINK ID	WAN LINK ID						
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

Example above shows SLA monitoring view from Branch-1 with Internet and MPLS transport towards all branches and Controller.

SLA Monitoring probes are used to track reachability and measure link metrics per access-circuit towards any given remote site. SLA optimization features such as Adaptive SLA and Data Driven SLA are available to optimize the SLA load in large deployments.

In a Full Mesh topology proper scaling of the maximum number of branches needs to be assessed. In order to dimension the deployment, the number of variables need to be taken in consideration, amongst them:

- 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 1K branches where there is one tenant per site with two WAN links in different Transport Domains the SLA probe traffic will consumes 6.25Mbps of bandwidth at each site using Versa default SLA Monitoring configuration. Increasing the number of Branch CPEs will increase this bandwidth floor plus the CPU overhead required to run this task.

Versa solution can limit the SLA Monitoring traffic to keep link utilization low for instance on high cost links such as LTE connections.

Further reading on SLA optimization can be found at:

https://docs.versa-networks.com/Versa_Operating_System/VOS_SD-WAN_Configuration/Advanced_SD-WAN_Configuration/SD-WAN_Traffic_Steering_Engine

In a Full Mesh topology there is direct reachability to prefixes in remote branches. Traffic is routed to those prefixes by using the next hop of the remote branch loopback (TVI) interfaces. This can be seen in the routing table (below).

In case 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.

Branch-1 VRF routing table view:

```
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
```

Full mesh topologies are the default option in workflows.

7.2 Hub-and-Spoke

Versa SD-WAN solution supports different flavors of Hub-and-Spoke topologies:

- Spoke-To-Hub Only
- Spoke-To-Spoke via Hub
- Spoke-To-Spoke Direct
- Spoke-Hub-Hub-Spoke

7.2.1 Spoke-to-Hub Only

In Spoke-to-Hub Only, spokes routes are not re-advertised by the hub branch. The only prefixes advertised by default are hub routes. This topology should be used when spokes don't have to communicate with each other. A good example is ATM cash machines network where devices need to talk exclusively to resources in the customer DC.

In the diagram below, it is shown that spoke prefixes are only accepted by the hub and rejected by other spokes per BGP community configuration.

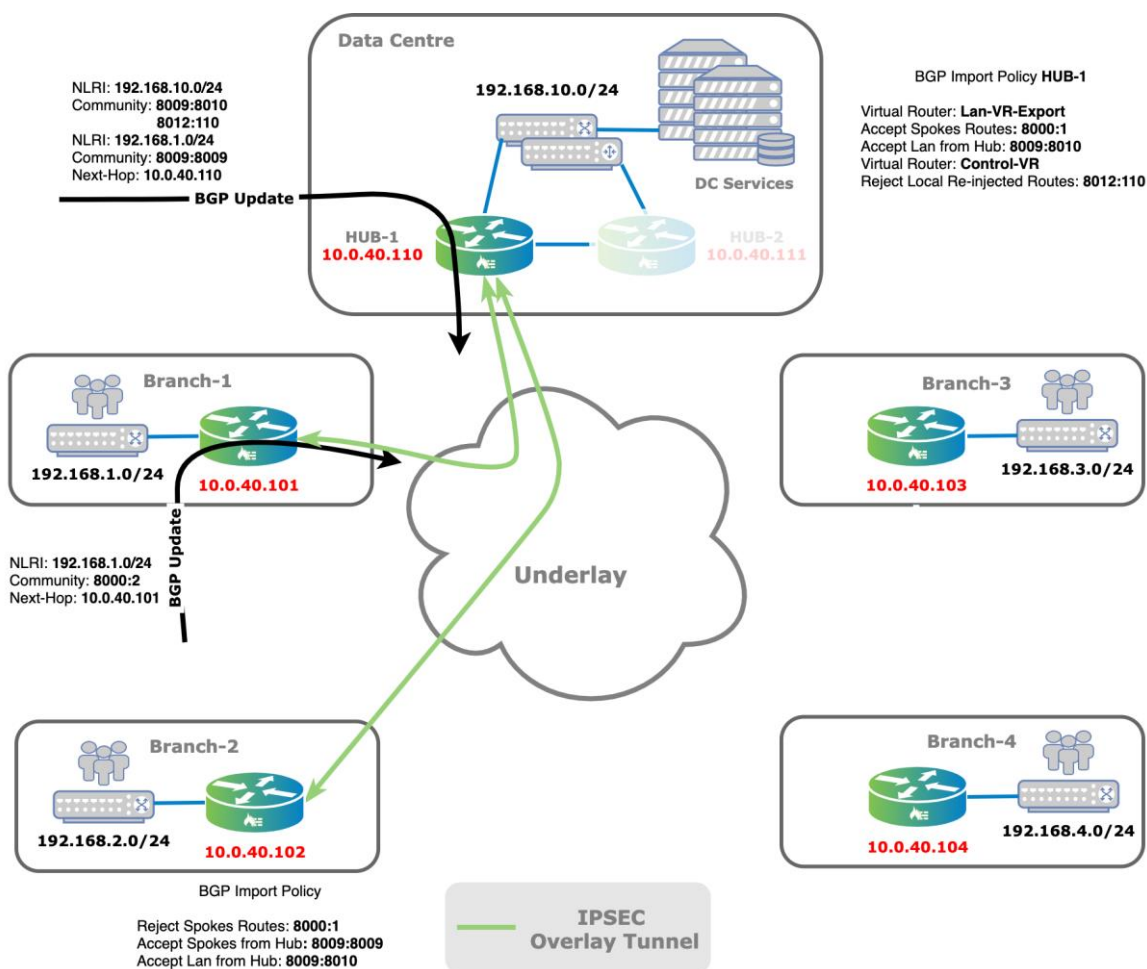


Figure 25 Spoke-to-Hub Only

The spoke Branch-1 VRF routing table only has routes 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
```

Spoke Branch-1 route table only shows destinations behind hubs again with Hub-1 being preferred, no spokes routes are present. Spoke Branch-1 advertised prefixes:

```
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 ]
```

BGP import policies are used to filter spoke routes. Spoke community **8000 : 2** is used to filtered out on hubs in the LAN-VR-Export and not advertised back to the spokes. Therefore, on the hub, all spoke prefixes are in the 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  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
```

Hubs have all spokes prefixes installed in the corresponding VRF. Redistribution policy could be implemented on hubs if needed to perform route summarization or to generate default static route for instance to attract traffic from spokes.

7.2.2 Spoke-To-Spoke via Hub

In Spoke-To-Spoke via Hub topology spoke sites are connected to each other via hub site. The data path between two spokes travel through the hub.

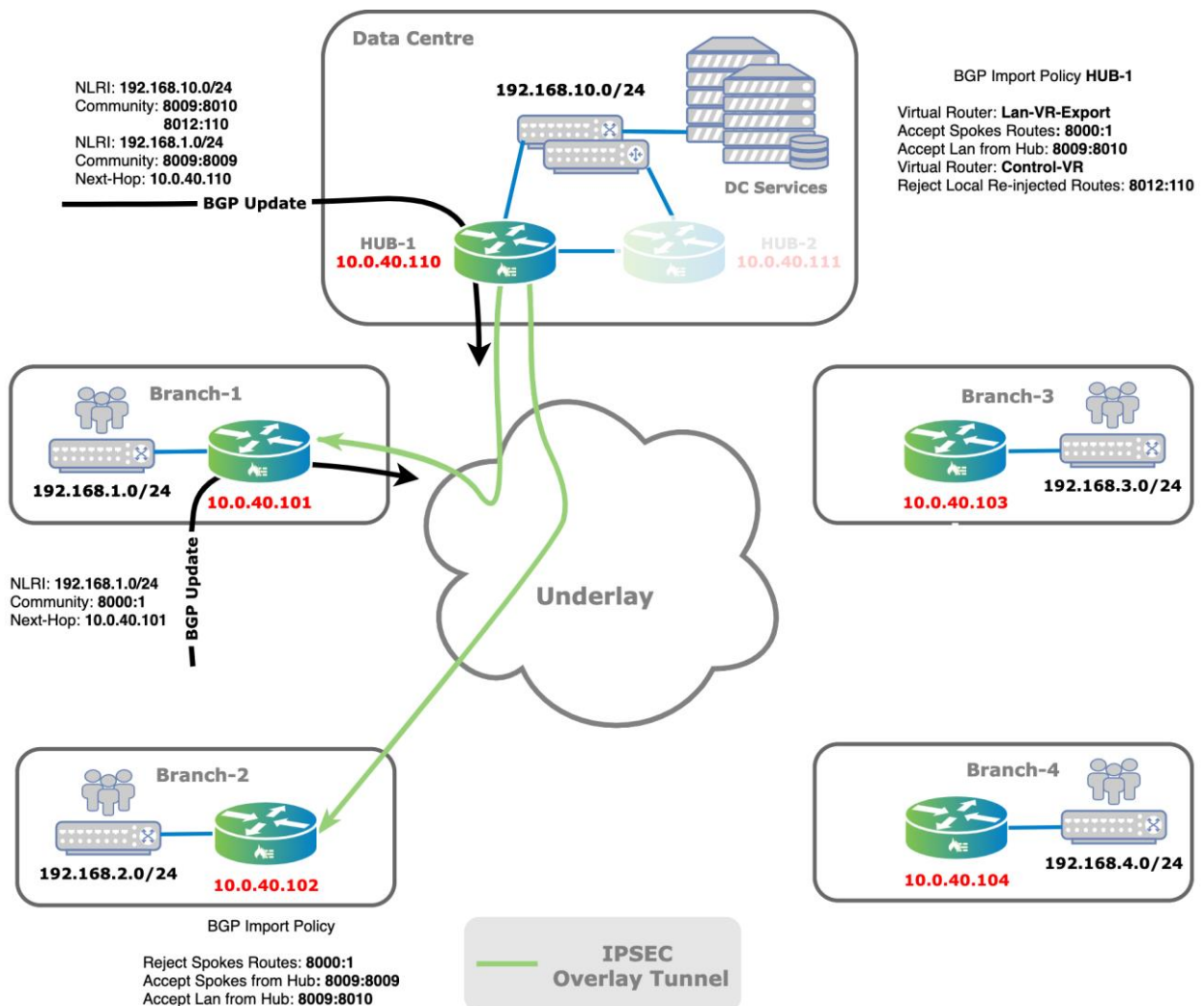


Figure 26 Spoke-to-Spoke via Hub

Hub-and-Spoke topology could be selected for instance when communication between branches is not required. Perhaps, security and other services are centralized at the hub side or WAN links cost of ownership dictates it.

Hub sites use the following architecture to manipulate VRF spoke routes:

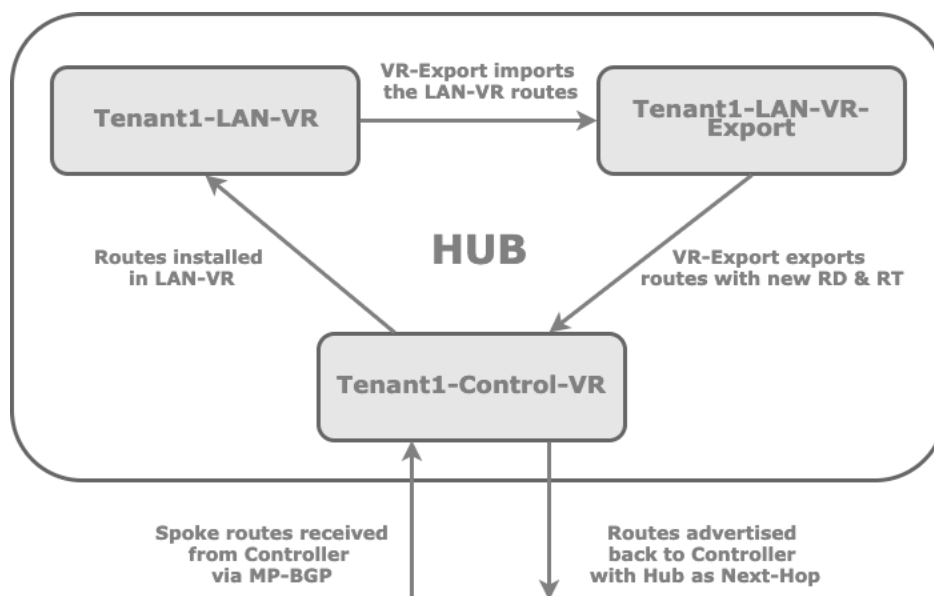


Figure 27 Hub internal routing policy

By changing route distinguisher on the hub for a set of VRF routes and advertising, the controllers will treat those routes as separated routes and will be accepted during BGP route selection. Original routes from spokes as well as hub advertised spoke routes will be present on the controller and reflected to the branches. Route Target filtering will do the rest of the routes selection on the Spokes by importing hub advertised spoke routes and as result selecting the hub as next-hop for the remote site routes.

In a Spoke-to-Spoke via Hub topology, branch will communicate only via the hub branch. Remote branches IP prefixes will have the hub as next-hop.

SLA monitoring is active only on paths towards hub sites and controllers, spoke sites are not being monitored, hence reducing the SLA probes traffic compared to a Full Mesh topology and addressing concern about scalability in a deployment with a high number of branches. SLA monitoring view from Branch-1:

```
admin@Branch-1-cli> show orgs org Tenant1 sd-wan sla-monitor status
```

SITE NAME	PATH HANDLE	FWD CLASS	LOCAL WAN LINK	REMOTE WAN LINK	LOCAL	REMOTE	ADAPTIVE MONITORING	DAMP STATE	DAMP FLAPS	CONN STATE	LAST FLAPS	LAST FLAPPED
					WAN LINK ID	WAN LINK ID						
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 spoke branch prefixes are re-advertised by the hub nodes, on the spoke branches all spoke prefixes are visible, but with the hub loopback TVI address as next-hop IP.

Branch-1 VRF routing table view show an example where 2 hubs are used. One hub is configured with a better priority and therefore owns 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
```

Example of Branch-1 advertised spoke routes:

```
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 ]
```

Community **8000:1** will color spokes routes, the BGP import-policy on spokes reject routes starting with this community. 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 ]
```

Community **8009:8009** colors spokes routes advertised by hubs, BGP import policy accepts those routes with hub as next-hop. Another community is used to color direct LAN route from hubs, this community is **8009:8010** which is also accepted by the BGP import policy.

In this topology, Hub-1 has been configured with a higher priority than Hub-2, that explains the dual entry in routing table for one prefix with Hub-1 as preferred next-hop but providing redundancy via Hub-2 when Hub-1 is not reachable.

The BGP import policy uses extended-community to accept routes from hubs and set a higher local-preference for Hub-1.

Extended **target:16002L:110** is derived from the site-id 110 which is Hub-1.

7.2.3 Spoke-To-Spoke Direct aka Partial Mesh

Partial Mesh topology is when some nodes are directly attached to each-other while other nodes are only attached to one or 2 nodes.

This type of topology could be selected when sites are geographically dispersed, and you want those to communicate directly with each other within the same region while inter-regional traffic will transit via Hubs branches or when high level of traffic are exchanged between specific sites but not with other.

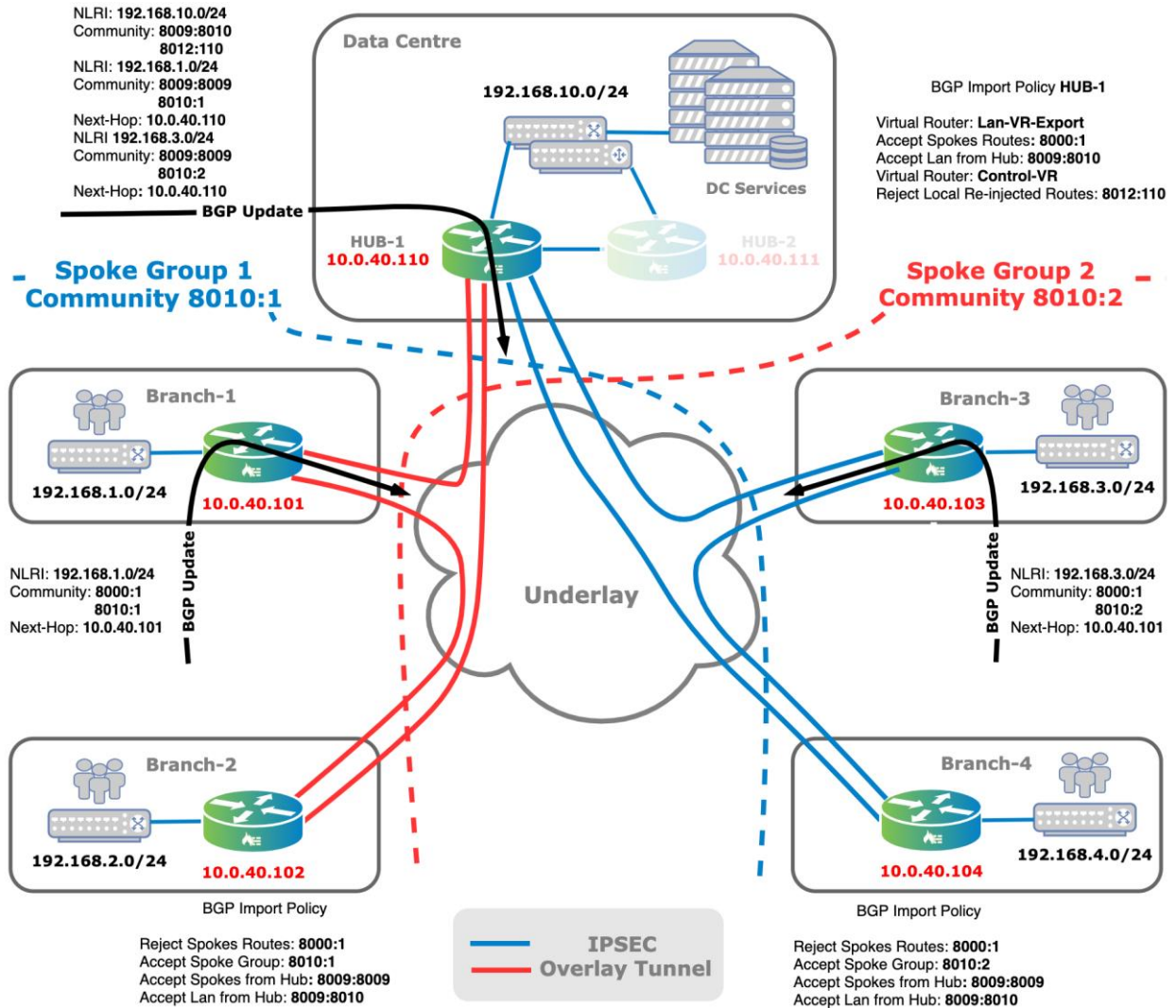


Figure 28 Partial Mesh

In a Spoke-to-Spoke Direct topology Versa solution uses Spoke Groups, within the same Spoke Group, branches will be able to communicate directly with each other and use the hubs to reach branches in different Spoke Group. For instance, in the above topology Branch-1 and Branch-2 will be direct but to reach Branch-3 next-hop will be Hub-1. Hubs will be connected using Full Mesh topology.

SLA monitoring view from Branch-1:

```
admin@Branch-1-cli> show orgs org Tenant1 sd-wan sla-monitor status
```

SITE NAME	PATH HANDLE	FWD CLASS	LOCAL		REMOTE		ADAPTIVE MONITORING	DAMP STATE	DAMP FLAPS	CONN STATE	FLAPS	LAST FLAPPED
			WAN LINK	LINK	WAN LINK	LINK						
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 will be performed on path towards Hub-1, Hub-2 and Branch-2 since those belong to the same Spoke group, branches in different Spoke Groups won't have SLA path monitoring active.

Branch-1 VRF routing 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. In the above output, it is shown that there are 3 entries in the routing table for 192.168.2.0/24, which is advertised by Branch-2 as well as Hub-1 and Hub-2. During underlay connectivity issues between Branch-1 and Branch-2 Hubs provide redundant path to reach Branch-2 prefixes.

Routes belonging to branches in different Spoke Groups are reachable only via hubs.

Important Note!
 Versa recommend deploying Spoke-To-Spoke Direct topology where feasible to provide redundancy and flexible meshing of branches using spoke groups.

BGP community and import policies will be used to accept or reject routes to create the expected topology.

Branch-2 advertised prefixes:

```

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 ]
    
```

Community **8010:1** corresponds to Spoke Group-1, during workflow configuration user are required to create a unique BGP Community for the associated spoke group. Import policy based on this community will be pushed to the spokes belonging to the same Spoke Group.

In the above topology Hub-1 has a higher priority which translates in an import policy rule to prefer the routes advertised by Hub-1 by manipulating the Local-Pref attribute. (Branch-2 > Hub-1 > Hub-2).

```

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
    
```


Branch-1 routing 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 will be directly reachable from hubs with a backup using remote hub(s), example above for prefixed from Branch-1 where direct route is active.

7.3 Regional Mesh aka Spoke-Hub-Hub-Spoke

In the Spoke-Hub-Hub-Spoke (SHHS) topology hub and branch devices are grouped per region. Within a particular region the topology could be anything (Full Mesh, Partial Mesh, Hub-and-Spoke).

Within a region branches will be communicating following the selected topology, when a branch in Region A wants to communicate to a branch in Region B it will be transiting through the regional hubs to reach the remote branch.

This topology can be selected based on geographical location when regional WAN transport networks are available and Hubs reachability between regions use the company backbone or high bandwidth WAN links.

Example on the next page shows 2 regional networks with different topology. Region A uses Spoke-To-Spoke Direct while region B uses Stoke-To-Spoke via hub. Communication between Branch-1 and Branch-3 uses Hub-1 and Hub-2 for the example sake but it could have been any regional hubs in the local region. This topology demonstrates how SHHS topology could be used with Versa SD-WAN solution applied to regional networks.

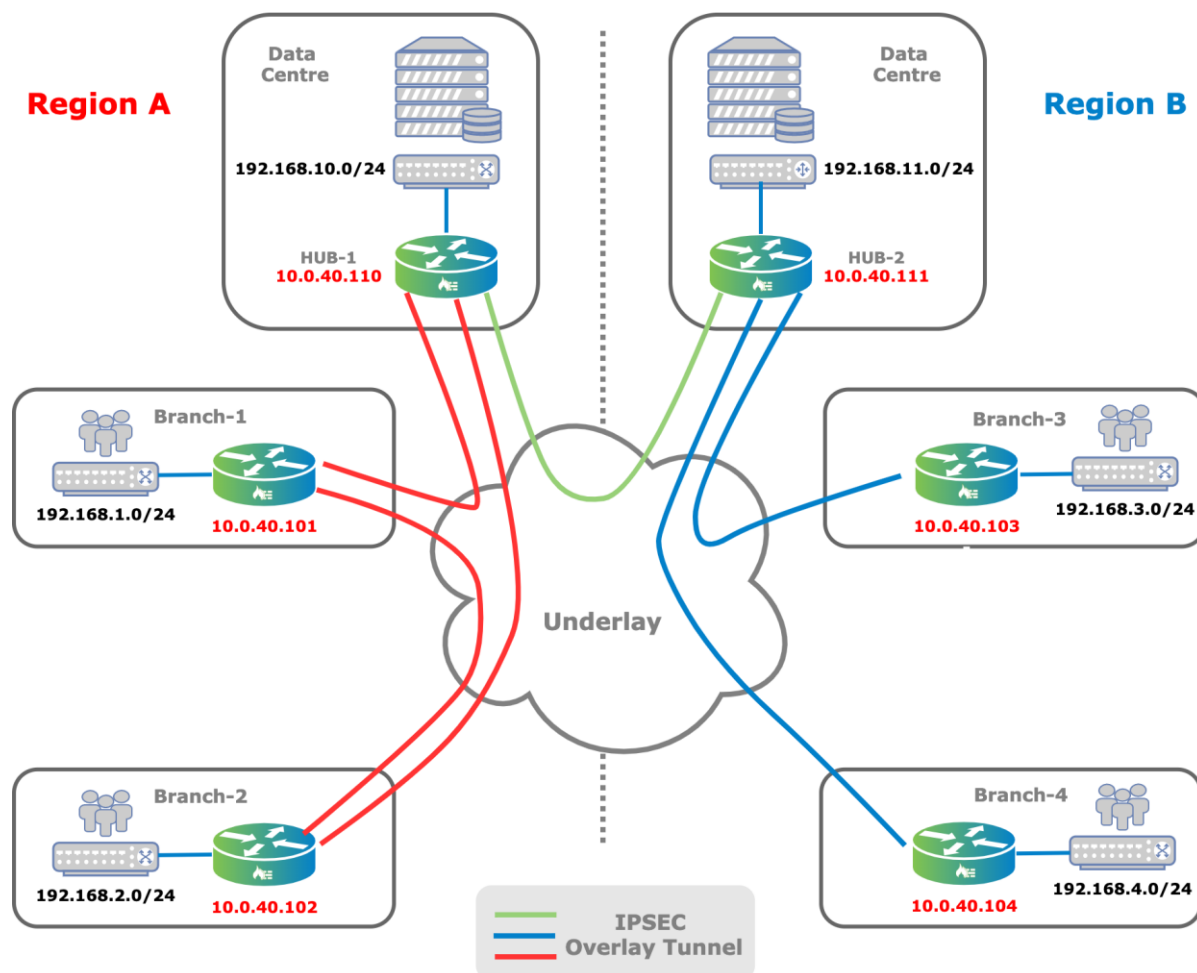


Figure 29 Regional Mesh (aka SHHS)

Branch-1 SD-WAN topology view:

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

Only regional Hubs and Branches can be seen since Region A is using Full Mesh.

Branch-3 SD-WAN topology view:

```
admin@Branch-3-cli> show orgs org Tenant1 sd-wan brief
SITE NAME      SITE ID  MANAGEMENT IP      TYPE    UP TIME      CONNECTIVITY STATUS  IS
-----
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
```

Branch-3 only is the Hub branches which is normal since Spoke-To-Spoke via Hub topology is used.

Branch-1 VRF routing 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  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 highlighted sections above are showing two prefixes in both regions.

192.168.2.0/24 prefix belongs to Branch-2, a direct route towards Branch-2 and a less preferred paths via Hub-1 for redundancy reason. Note the “+” next to the active route which is the direct route using Branch-2 next-hop.

192.168.3.0/24 prefix belongs to Branch-3, two entries are present and pointing to Hub-1 device.

Branch-3 VRF routing table:

```

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
    
```

The sections above are showing two prefixes in both regions and we can see no different treatment between regions. Spoke-To-Spoke via Hub topology is configured in Region B.

For clarity sakes **192.168.1.0/24** prefix is behind Branch-1 while **192.168.4.0/24** prefix is reachable via Branch-4.

7.4 Connecting sites over disjoint underlay networks

Gateways can be used to connect sites over disjoint underlay networks. Disjoined underlay means that the two sites do not have a common underlay that allow them to communicate directly. An example would be a site that has only Internet connectivity and a site that has only MPLS connectivity. There can also be failure scenarios where site that an internet and MPLS provisioned link is communicating with a site that has only an MPLS link. When the internet link on the first site is unavailable then without a gateway, the site will lose its connectivity to the MPLS only site. Another common scenario where disjoint underlay networks may occur is when NAT-Traversal issues don't allow two internet connected branches to connect directly. A third device, which is the Gateway, can interconnect the two branches.

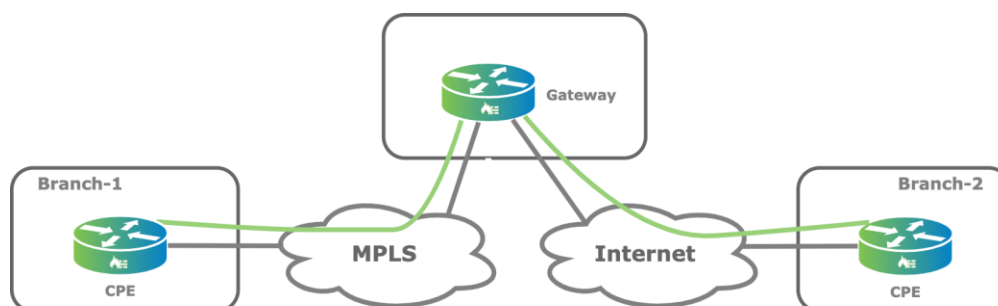


Figure 84 SDWAN Gateway to connect Branches over disjoint underlay

This configuration can be achieved in 3 ways on the Gateway:

1. Configure the branch as the Gateway and branches in the Spoke-to-Spoke-Direct topology in Workflow Template configuration.
2. Advertise a summary route from the Gateway. This can be any branch site and doesn't have to be configured as a Gateway in Workflow templates.
3. Advertise a default route from the Gateway. This can be any branch site and doesn't have to be configured as a Gateway in Workflow templates

7.5 SD-WAN topologies for Geographical Isolated regions

In certain region of the world governments placed restriction on encrypted (IPSEC) traffic to basically block it from going in and out of the country. Inside the country traffic is typically not impaired. This could cause SD-WAN regional network to become isolated.

With separated SD-WAN islands each SD-WAN domain operates within itself with no visibility or information about other domains.

The SD-WAN islands must be interconnected with IPSEC or other connection option which must be managed and provision outside the SD-WAN domain somehow, consider the stitching complexities on the boundary node that act like NNI points.

This separation prevents the control planes of each SD-WAN island from exchanging information with each other as result no connectivity between the domains.

Tenant creation and Service template across SD-WAN islands become an even more complex task.

The Versa SHHS solution tackles this problem, providing a solution that is highly scalable, compartmentalized, fully automated, and easy to deploy and operate.

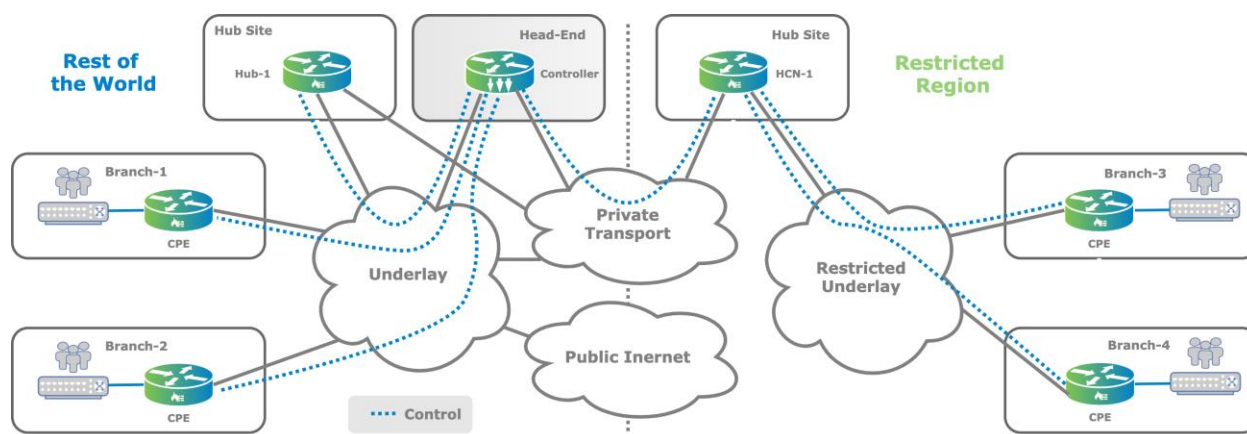


Figure 30 Restricted Region Control Plane

Above drawing shows an example of the Control Plane connection for an isolated region, an iBGP session is established between the Head-End controller and the Hub-Controller-Node (HCN-1) site in the Restricted Region to exchange SD-WAN route information over a suitable transport in this case a Private Transport which will allow IPSEC traffic.

Data plane flow between a branch in the restricted region and a branch with no restriction will traverse the Hub sites, using IPSEC tunnels. Note the Head-End controller is not part of the Data Plane communication.

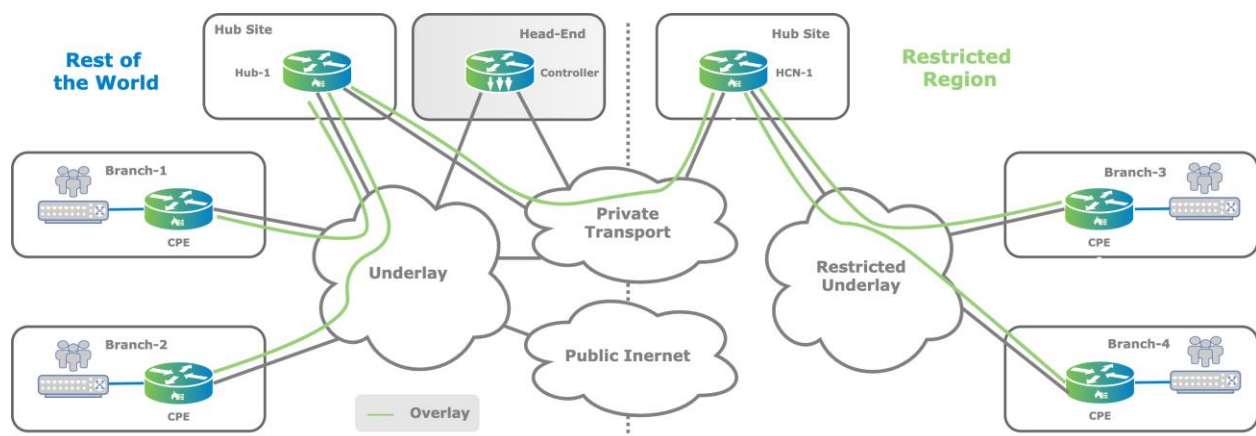


Figure 31 Restricted Region Data Plane

The above SD-WAN network topology could apply to restricted regions. Branches in this region will be create an SD-WAN island in the form of a restricted underlay with no connectivity with the rest of the world. Approved transport needs to be provided by a government approved Service Provider to allow IPSEC traffic, it could be a L2 service (L2VPN, VPLS, etc) or a L3 service. This service is represented by the Private Transport.

For such topology Versa has introduced a hierarchical controller structure called Hub-Controller node (HCN) Device Type. A HCN can interconnect the control planes of SD-WAN islands that may be using their own controller instances. MP-BGP-based Versa control plane assigns separate community values to represent each SD-WAN island. iBGP sessions are established between the regional HCNs and the Versa Headend Controller unifying the control plane.

HCNs have the ability to consolidate controller and hub functionalities onto one node to preserve resources in the control side of the network. HCN nodes exchange information with spokes, and they implement the data plane functions of hub nodes.

HCNs can be deployed in active/active for the control plane, maximizing uptime and ease of serviceability, multitenancy is preserved across the topology using the organization and suborganization structure. Data plane redundancy will be done by BGP next-hop route resolution.

The network administrator then provisions 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.

Once 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. 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, requiring no manual involvement from the network administrator. After the end-to-end data paths are set up, the users can start communicating between the spokes of separate SD-WAN islands.

This communication is seamless thanks to the Versa established advanced routing and SD-WAN capabilities.

7.6 Multi-VRF/Multi-Tenant Topologies

In the above examples it is assumed a single VRF/Single Tenant for clarity reasons and to explain differences between topologies.

Versa SD-WAN solution is highly flexible, it allows to define and establish different topologies at the VRF level or tenant level. The default Versa SD-WAN model provisions a full forwarding mesh through IP prefix advertisement in MP-BGP, you can configure the VPN topologies we discussed above using Versa Director workflows.

In the following diagram a SD-WAN topology with two VRFs within one organization is shown. A red colored VRF-A is configured for Spoke-To-Hub Only while the blue colored VRF-B uses a Full Mesh topology.

Same principle could apply to multi-tenant branches where topology could be different between organizations.

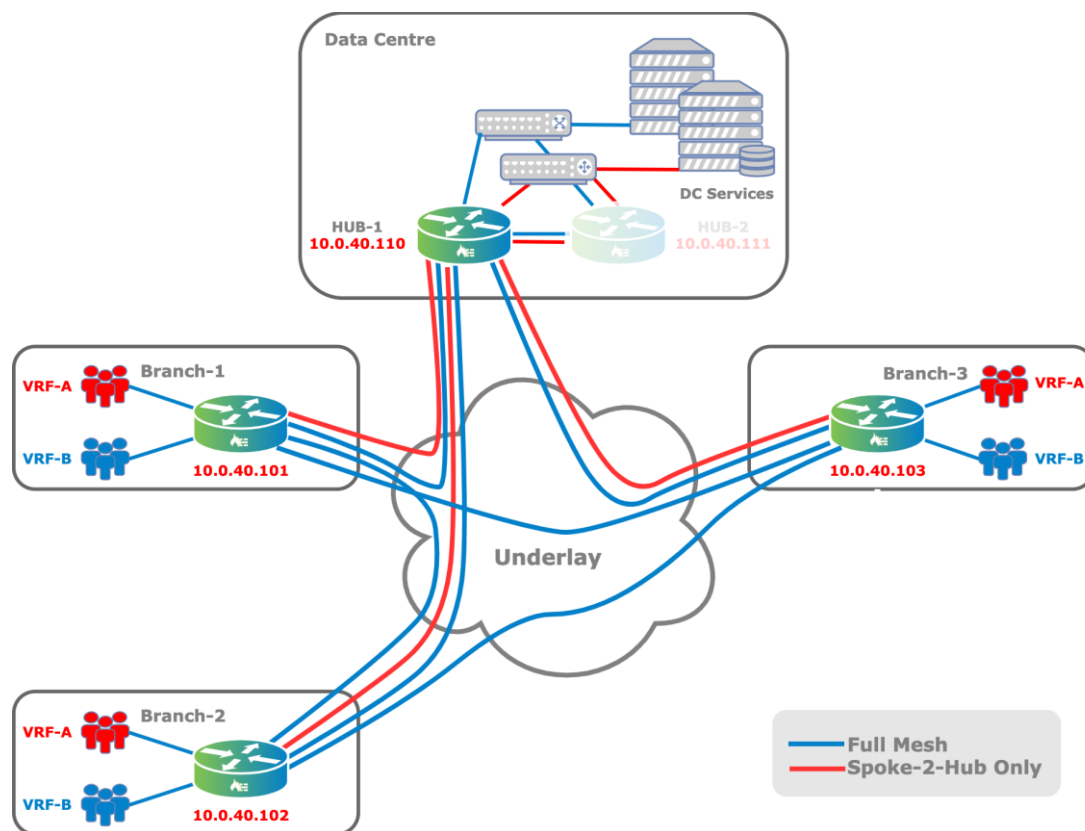


Figure 32 Different Topology per VRF

7.7 Best Practices Topologies

- Full Mesh topologies are the default option in Versa workflows. They can be deployed without restrictions to VPN's with less than 100 sites. If VPN's span with more sides, the SLA monitoring will consume significant bandwidth and therefore branches with low bandwidth connections are required to assign a relative high portion of their available bandwidth to SLA monitoring traffic only. Versa recommends low bandwidth branches to deploy in a Spoke-to-Spoke-via Hub topology with only SLA monitoring towards hub sites.
- Spoke-to-Spoke-Direct topologies are recommended over Full Mesh topologies. For many use-cases, having a hub node in the topology is desired, for example to avoid disjoint underlays situations caused by circuit failures or NAT-traversal issues. Also, the Spoke-to-Spoke-Direct topology give better support for automatically importing routes from LAN adjacent networks (avoiding the need to manually configure redistribution policies).
- Use a distinct WAN network Name on the Hub Sites.

8 Direct Internet Access

In a typical legacy Enterprise VPN topology, internet traffic from the branch is connected to a hub site over a private MPLS link from where it breaks to the internet out. The internet traffic typically goes through a web proxy which 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 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 like Office365 the need to get a better quality of experience for the user is important.

The primary advantages of DIA are:

- Reduced bandwidth requirements at headquarters
- Fewer network hops
- Reduced latency due for internet application to direct routing and better optimization.

The increased reliability of the Internet for WAN transport makes DIA desirable in branch deployments. Sending traffic directly from the branch to the Internet creates additional security challenges, special care should be taken to protect the branch in such scenario.

The DIA architecture is based on standard routing principles. Internally a logical connection is created between Tenant VRF and WAN transport VR, default route will be advertised with e-BGP in the Tenant VRF and will use the logical tunnel to resolve the next-hop. NAT rule will translate all LAN traffic into one public IP.

8.1 VOS Edge Device DIA architecture.

When configuring internet breakout on a branch, the configuration of the VOS Edge Device is enriched with 2 main components: internal connection between VRF and Transport WAN VR, supported by a eBGP connection to control route distribution and a CGNAT function to address translate all internet bound traffic to the public IP address typically attached to the Transport WAN VR interface.

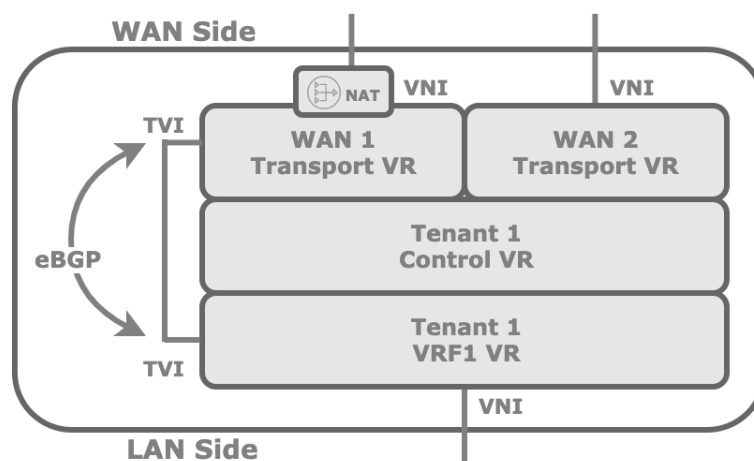


Figure 33 DIA Branch Internals

When multiple WAN are available DIA redundancy and DIA load balancing can be archived. Local-Pref will be programmed by the workflow to select the preferred WAN interfaces. In case of load balancing option is selected in the workflow, same Local-Pref value for both default routes will be advertised over the e-BGP towards the VRF.

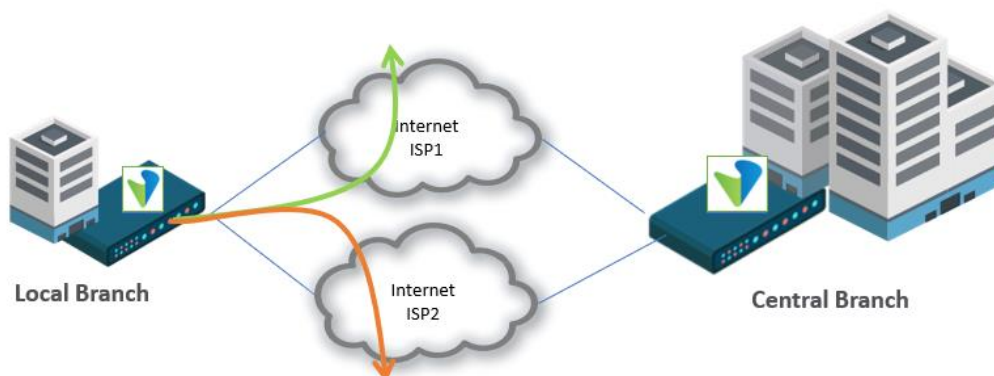


Figure 34 Multiple local internet breakout

DIA is configured using Workflow in Versa Director in the Tunnels tab. This will create configuration of all required infrastructure.

8.1.1 Central or Regional Internet breakout

In a typical deployment it is very common to provide default route redundancy. A default route pointing to a local DIA with a backup default route (less preferred metric) to the central DIA. In case the local branch internet circuit is malfunctional, the local sourced default route is withdrawn, and the central sourced default route will be activated.

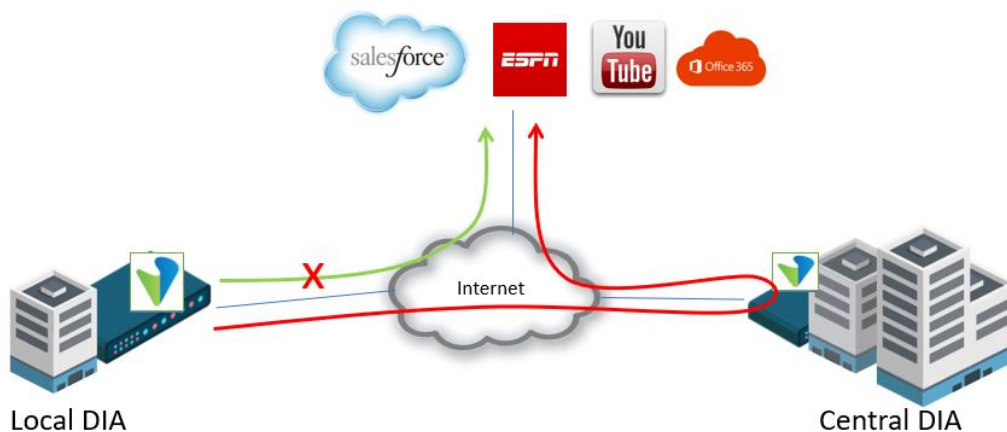


Figure 35 Local DIA with Central DIA backup

Using the configuration created by the workflow, a local internet breakout and central/regional/remote backup can easily be established. This can be achieved by using the Gateway field in the Workflow:

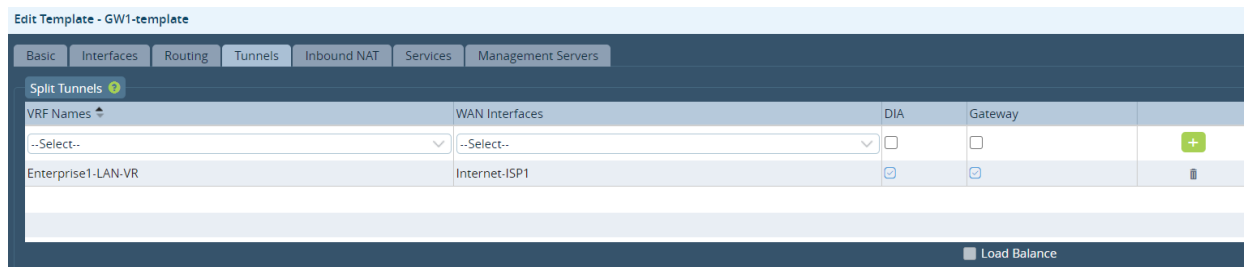


Figure 36 Configuring Internet breakout with workflow

To achieve this scenario, the local breakout Edge Device should NOT have the “Gateway” checkbox ticked, where the central Internet breakout Edge Device also has the “Gateway” box ticked. The GW checkbox results in having that Edge Device announce the default route to the rest of the SD-WAN. If the GW checkbox is not ticked, the default route sourced by the local DIA infrastructure, will not be made visible anywhere else in the SD-WAN VPN other than in the local Edge Device.

The local device routing table shows like as follows:

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

Above can be observed that the preferred route is the local breakout route (next-hop 169.254.0.2). Zooming in to the BGP table, we will find that it has a better distance (eBGP over iBGP):

```
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
    
```

By manipulating the BGP preference in the configuration, this behaviour can be altered.

Through this infrastructure advanced configurations can be created, using route-colouring with communities, where different regions prefer different regional default routes.

8.1.2 Breakout to an MPLS underlay for common SP services

Another use-case is a breakout service to the MPLS underlay. This breakout service may be used for reaching services offered by the MPLS provider from the underlay (such as VoIP services). This solution may also be used as gateway service during migration from legacy MPLS to SD-WAN.

The breakout to MPLS is achieved the same way as a local internet breakout service. In the workflow, a split tunnel can be created to the MPLS underlay. One important difference with an Internet breakout split tunnel is that the DIA checkbox should NOT be ticked. Marking the DIA checkbox would create a NATP instance, which is not needed in private MPLS underlay.

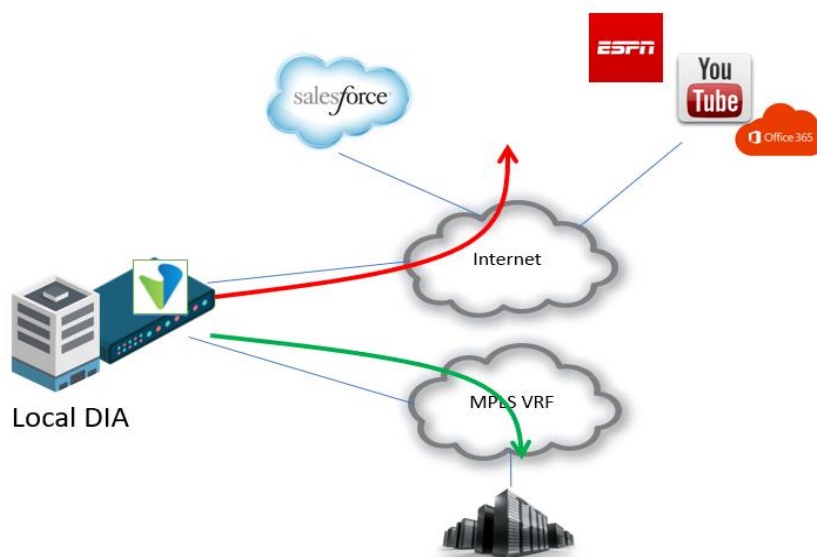


Figure 37 Split tunnel to both Internet and MPLS VRF

If you prefer the routes imported from MPLS to be advertised to the rest of the SD-WAN VPN, you can tick the “Gateway” checkbox in the workflow below.

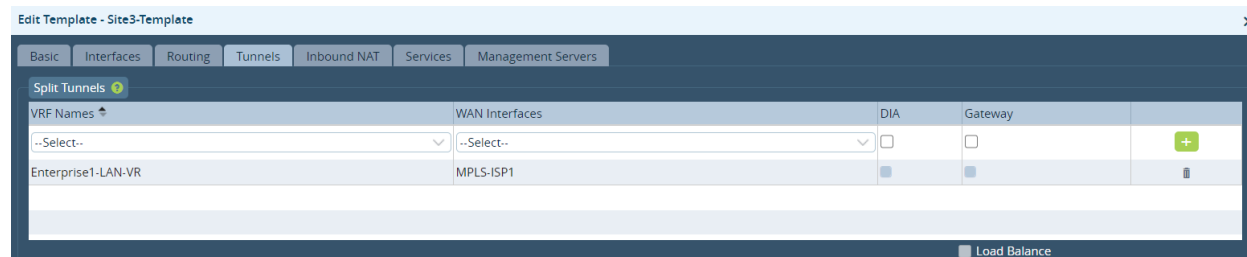
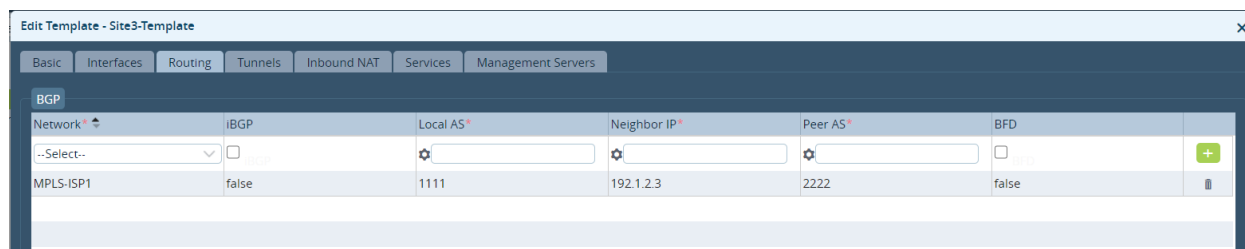


Figure 38 Adding Split tunnel to the MPLS underlay with workflow

The above infrastructure will create a similar internal connection between the LAN-VR and the MPLS-VR and run eBGP on top of this. This causes the routes in the MPLS-VR to be propagated into the LAN-VR.

The next step that is required is an import of the MPLS VRF routes in the SD-WAN Edge Device. This requires a dynamic routing connection between the SD-WAN Edge-Device and the MPLS provider infrastructure (typically the PE router). The easiest way to accomplish this, is to make the BGP connection configuration by the workflow:



39

Obviously, the (PE) router on the other side (owned by the MSP) should be configured with such dynamic routing configuration as well.

8.1.3 Application based break-out

A very common use-case is to provide a local breakout service for specific applications only, where the default route is pointing to some regional or central breakout point. This configuration is a little more involved. It will rely on the infrastructure created in section 9.2, where a local breakout infrastructure is present in the SD-WAN Edge Device, but also have an option to break out centrally.

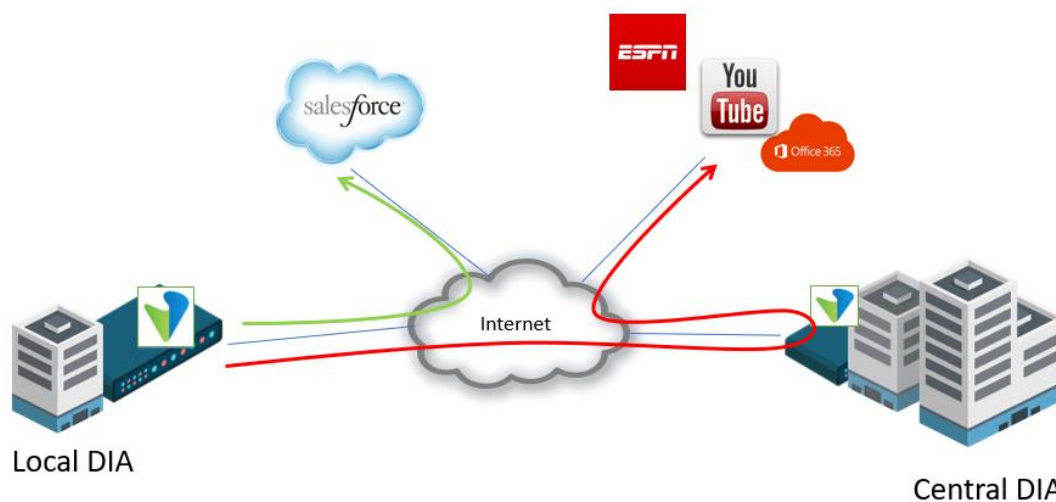


Figure 40 Application based break-out

As illustrated in the previous section, the infrastructure will prefer local breakout over central breakout. Since we want the default internet bound traffic to be centrally broken out, we need to alter the preference to something which is higher than the metric of 200. This is required because EBGP route AD (20) is better than iBGP route AD (200) by default.

We do this in the LAN-VR under BGP General, where we set the eBGP Preference to 220.

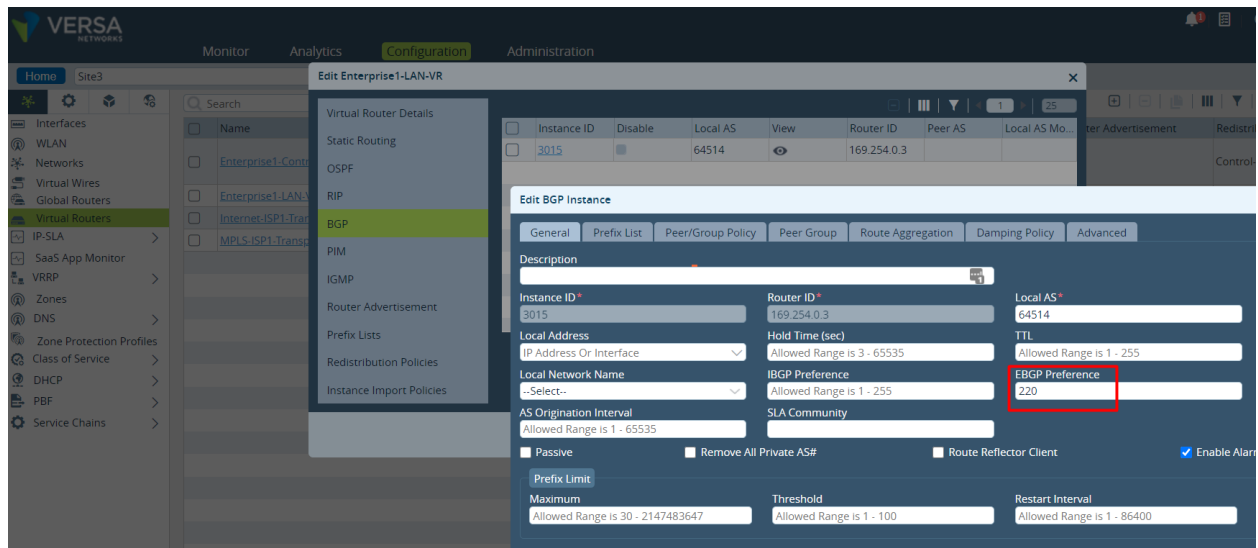


Figure 41 Setting route preference in BGP configuration

This will result the central default route to be 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 will break-out at the central breakout point. You can verify this by checking some internet sessions in the local branch whether they are SD-WAN yes/no:

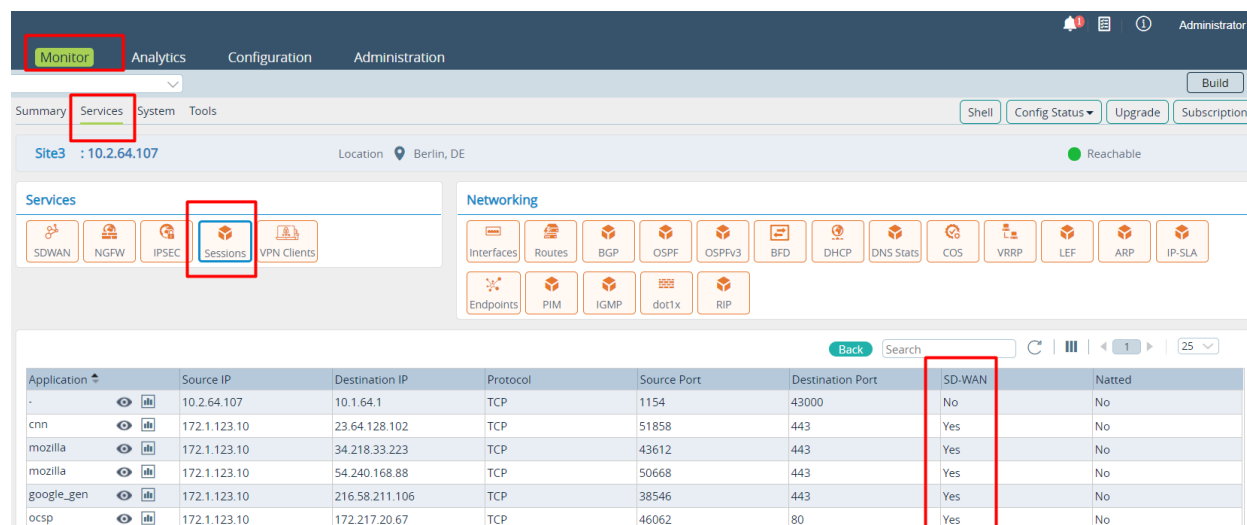


Figure 42 Intermediate verification of local Internet breakout with Monitor

Furthermore, the NAT rule needs to be modified to accommodate this use-case. The change is simple, but important:

The matching conditions of the Rule created by the workflow need to be modified to match on the source-zone (not the dest-zone). Therefore, the dest-zone entry needs to be removed and the source zone entry needs to be added. Select from the pull-down the zone “W-ST-
<tenant-name>-LAN-VR-<internet WAN>-VR”, as illustrated below.

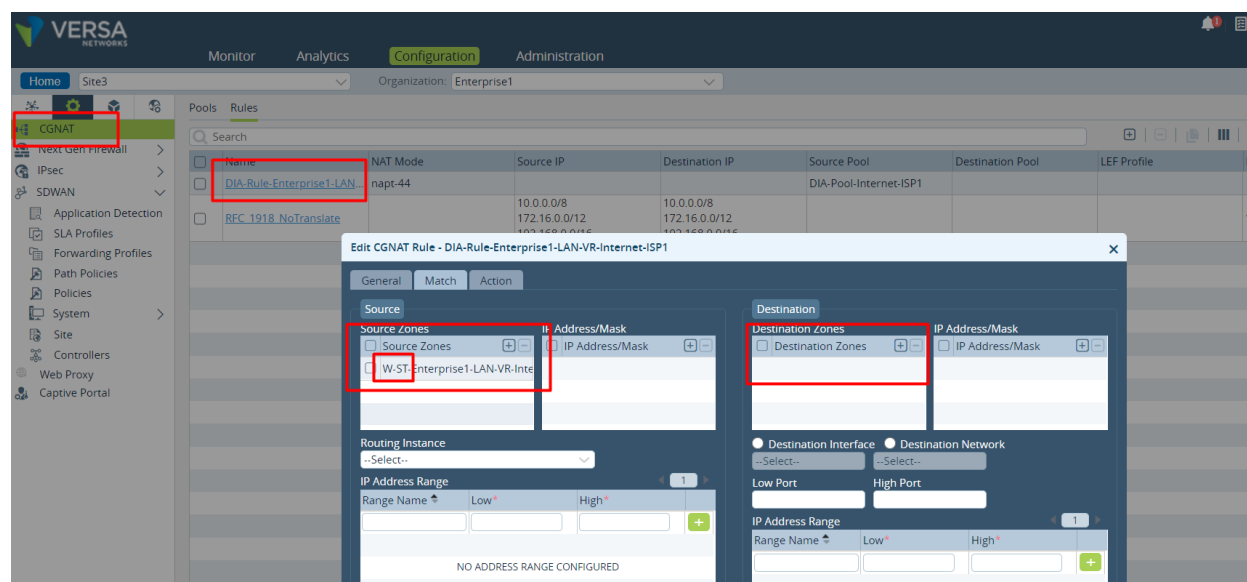
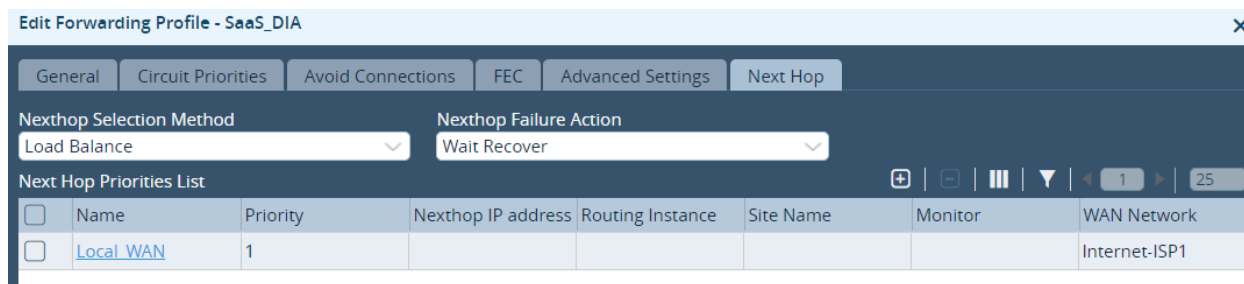


Figure 43 Modifying the NAT rule

To force a specific application breakout locally, where the remaining internet traffic breaks out at the centrally, an SD-WAN Policy is used.

In this example, a break-out of “Salesforce.com” traffic to the local DIA is achieved.

First a Forwarding profile should be created. In the forwarding profile we define the conditions of the next hop. In this simple example, only one next hop is the local DIA.



Next is the creation of a policy with matching criteria. This policy is then linked to the Forwarding profile created above.

The final policy will look like as follows:

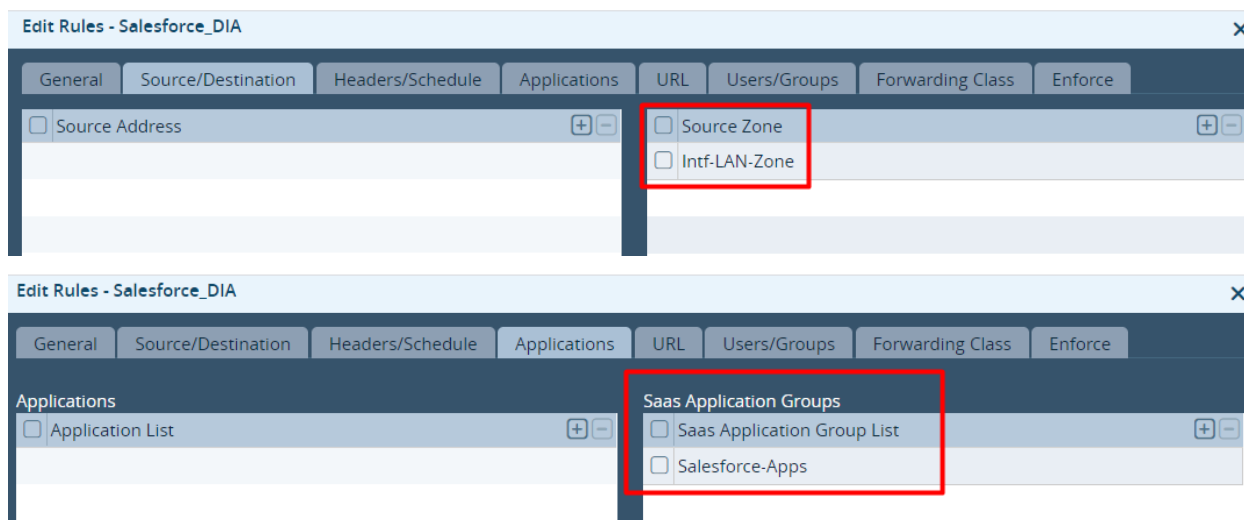


Figure 44 Specify the Application classification in SD-WAN policy

Where the policy is attached to the forwarding profile under “enforce”:

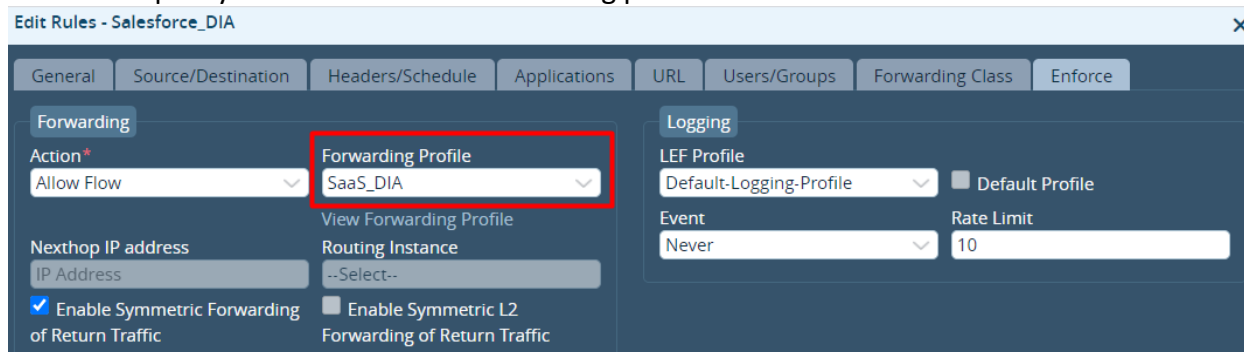


Figure 45 Enforce traffic to SaaS_DIA FP

For verification of the correct functioning of this configuration can be done by opening a browser to the application you selected (Salesforce in this example). There are multiple ways of verifying this. In this example we will inspect the session details and verify the egress in Versa Director Monitor. We select “sessions” and click on the filter icon:

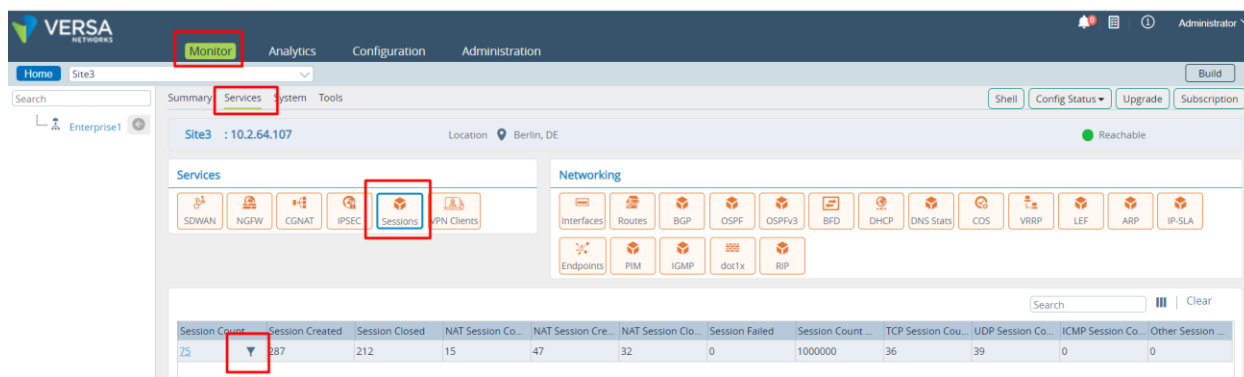


Figure 46 Verification of App based breakout with Monitor 1/2

Now we search for “salesforce” and click on the eye. If the breakout to the local DIA is correctly functioning, the Ingress and Egress circuit should show TVI-0/602 (the TVI with the same IP as you set in as Next-hop).

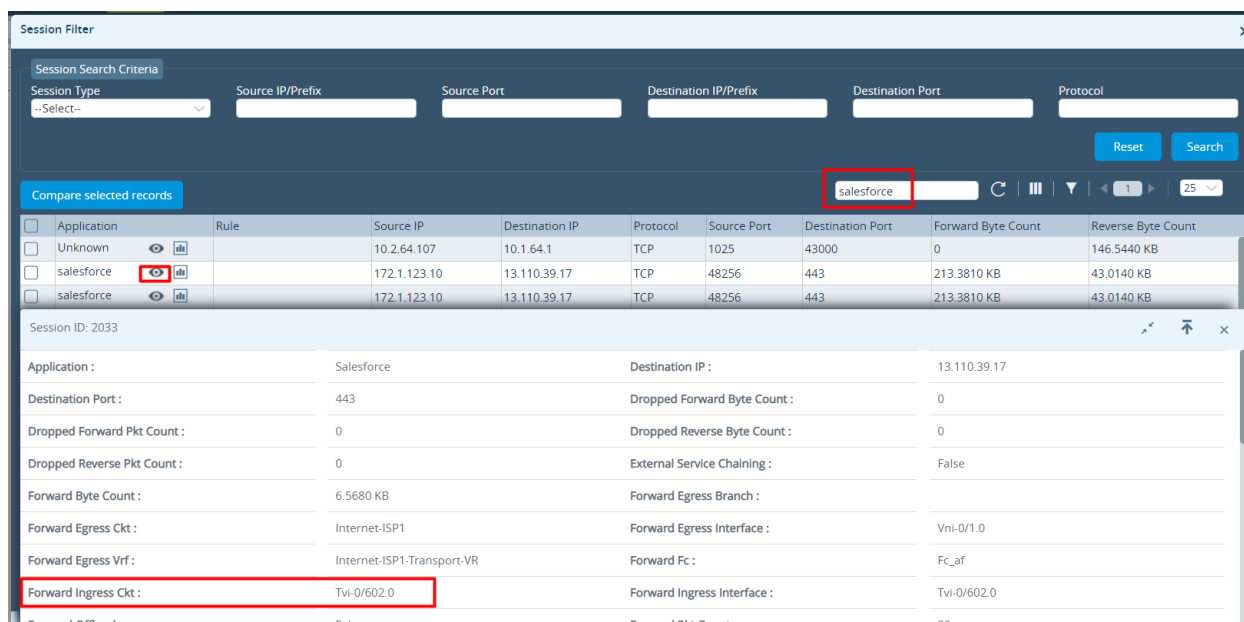


Figure 47 Verification of App based breakout with Monitor 2/2

You can also verify if the sessions of any other (non-Salesforce) traffic is NOT pointing to this TVI. In the session details you will find typically the “Tx Branch” with you central DIA Edge Device.

8.2 Best Practice: Remove the local breakout default route

The default route sourced by the local breakout DIA configuration can be suppressed as this is not needed for this configuration. If you like to keep this default-route, it will act as backup breakout point for the local site in case the central breakout fails. If steering traffic based on application use-case is used, this locally sourced default route is not needed. When the administrator may never want to breakout locally, except for the specific application, this default route should be removed. This can be done by filtering it with a prefix list or removing the redistribution of static routes under redistribution policies.

8.3 Best Practices for DPI application recognition

When testing the implementation describes in this chapter, you may notice that the traffic may not be locally broken out, while it is expected to do so. One very common reason for this is that DPI didn't manage to detect the application ID at the time the session is created. This is more often seen in lab setups and less in production deployments. In production deployments there is more "application cache" for the same application traffic (e.g. the system knows that traffic going to a specific destination IP is e.g. salesforce.com). Forcing the system with creating additional sessions (e.g. by refreshing the browser) will build up application cache.

8.4 Best Practice Summary

- Use DIA Gateway feature to enable central break-out to Internet, to ensure backup for local breakout.
- Use application-based local breakout to improve performance for one or a group of applications.
- For application-based traffic steering remove the default route populated by workflow template to always send other than one (or a group of) specific applications to central location.
- For application-based traffic steering testing and verification make sure to generate multiple traffic flows.

8.5 Performance based SaaS Optimisation

The Versa SD-WAN solution has the ability to select the best internet breakout to reach SaaS applications. For concluding on the application network performance, the platform is using the following methods of measurement:

1. **Passive/Inline Monitoring:** VOS Edge Device collects metrics for on-going TCP based SaaS application sessions transiting through 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 hence select the best path for the application.
2. **Active Monitoring:** Ability to pro-actively monitor SaaS locations using ICMP, TCP and HTTP probe, export collected metrics to remote sites and incorporate actively learnt metrics into the path selection decision. Metrics collected through active monitoring are Latency (RTT) and Packet Loss.

The data collected by the above measurement techniques are automatically combined with the already existing SLA measurement over the overlay. All those datapoints together conclude on a best path and the SD-WAN-policy will steer the traffic accordingly. As per the below diagram, the local DIA branch is measure a latency of 20ms to Salesforce.com, while from the same branch, reaching Salesforce.com via a central DIA, it will take 60ms. Therefore, the logic in the local DIA branch will steer the traffic to Internet using the local DIA.

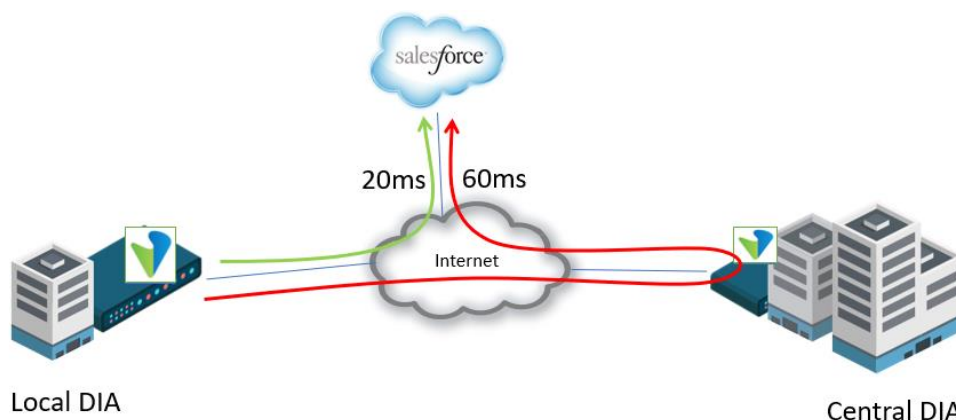


Figure 48 Performance based breakout overview

8.5.1 Setting up performance-based breakout configurations

This section describes performance-based application breakout based on Active Monitoring of the SaaS application.

Performance-based breakout is an add-on to the application-based breakout scenario which is described in [Section 8.5.4](#). It should leverage the infrastructure of local and remote DIA and the modification in the NAT rule. However, it relies on a different configuration of the SD-WAN policy.

For Performance-based breakout, the following additional building blocks need to be created:

- SaaS App Monitor in both local and remote breakout Edge Device
- 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 this design configuration example, we do like to break out the application “Salesforce” to the breakout point that provides the lowest latency to this SaaS. The latency is measured from the local Edge Device point of view. In this configuration example, there are two options:

- Salesforce traffic is locally broken out. The latency measurement is from the local Edge Device to Salesforce only over internet.
- Salesforce traffic is remotely broken out. The latency measurement is from the local Edge Device to the remote Edge Device (over the SD-WAN overlay) to Salesforce over internet connected to the remote Edge Device. This means that this configuration will “add” the internet latency with the overlay latency.

The configuration for this scenario is as follows:

For both the local Edge Device as well as the remote Edge Device, a **SaaS App Monitor** needs to be configured. This SaaS App Monitor intends to check the performance to the configured FQDN of the SaaS application. The below configuration is the minimal required configuration for the SaaS App Monitor. This configuration is taken from the LOCAL Edge Device.

Figure 49 Adding SaaS App Monitor

Next things that needs to be configured is the **SLA profile**. This configuration is only required on the local Edge Device. The minimum configuration needed is to attach the SaaS App Monitor to the SLA profile. In this configuration we have chosen to have the network choose the lowest latency or loss. Alternatively, you may also set a maximum latency or loss (all paths are considered as long as the maximum is not reached).

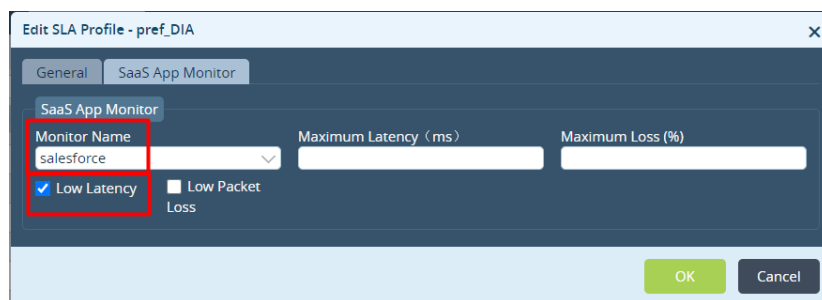


Figure 50 Adding the SaaS App Monitor to the SLA profile

Now the SLA profile is set, the **forwarding profile** should be configured. In the forwarding profile, we will reference to the SLA profile.

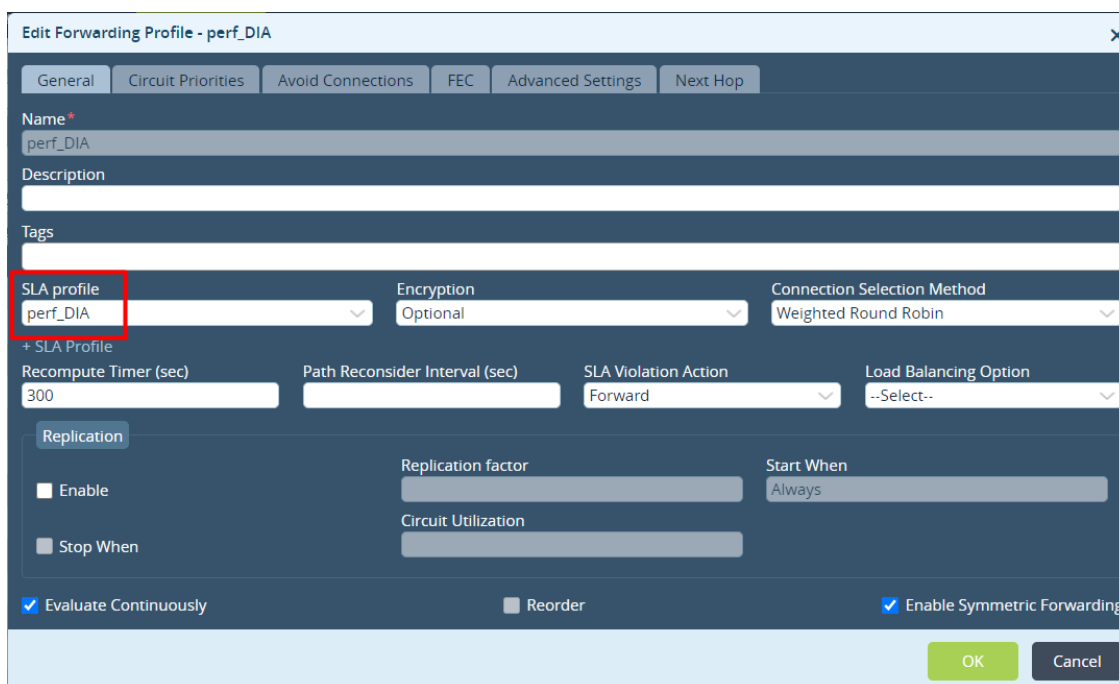


Figure 51 Adding SLA profile to the forwarding profile

In the forwarding profile we will also set the next hop tab. In this tab, we should specify the Next Hop options (breakout out options) for the traffic used by this forwarding profile. In this config example, we have a local breakout to internet, and we have a remote breakout to internet.

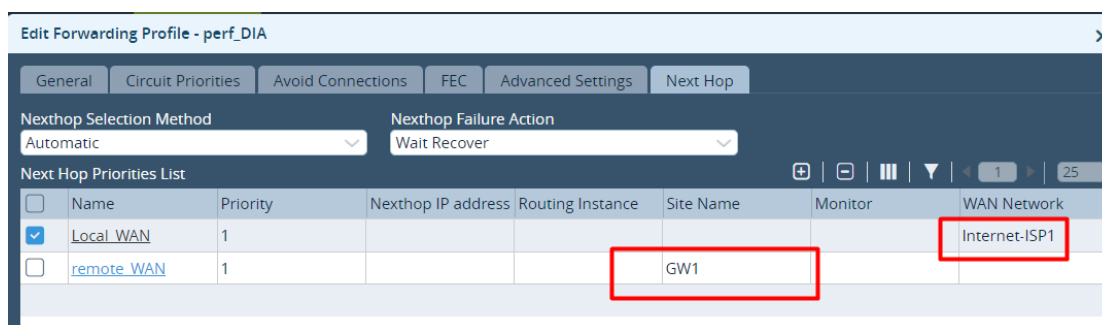


Figure 52 Setting the App breakout options

For the local internet breakout, we will select the local WAN interface.

For the remote internet breakout, we will select the remote Site where this remote breakout is available.

Now we have specified the next hop options for the traffic that will be using this Forwarding Profile, we need to specify the conditions what breakout to use. As we have chosen for the network to have automatically choose the breakout with the lowest latency (as per this example), the “Next-hop Selection” is set to “Automatic” and the priorities are the same (both priority 1).

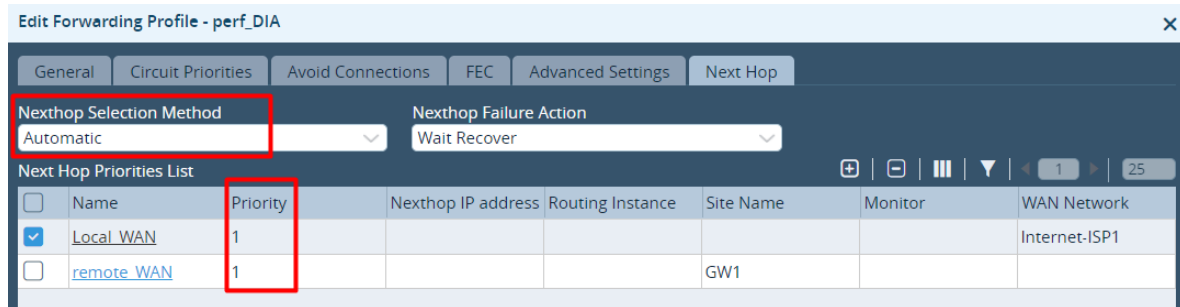


Figure 53 Automatic Next-hop priority list

Finally, an **SD-WAN Policy** should be created. This SD-WAN Policy defines the classification criteria for the traffic of this use-case. We have chosen to write this example for “Salesforce” traffic, sourced from the LAN-Zone. Those two criteria need to be defined in the SD-WAN policy.

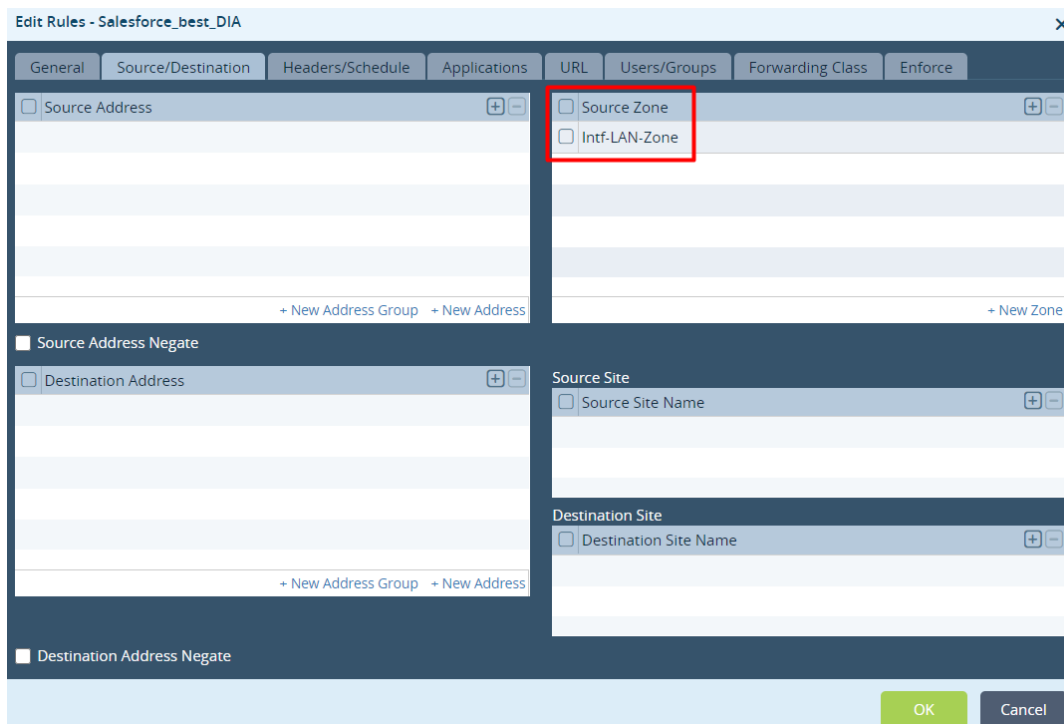


Figure 54 Setting up the SD-WAN policy

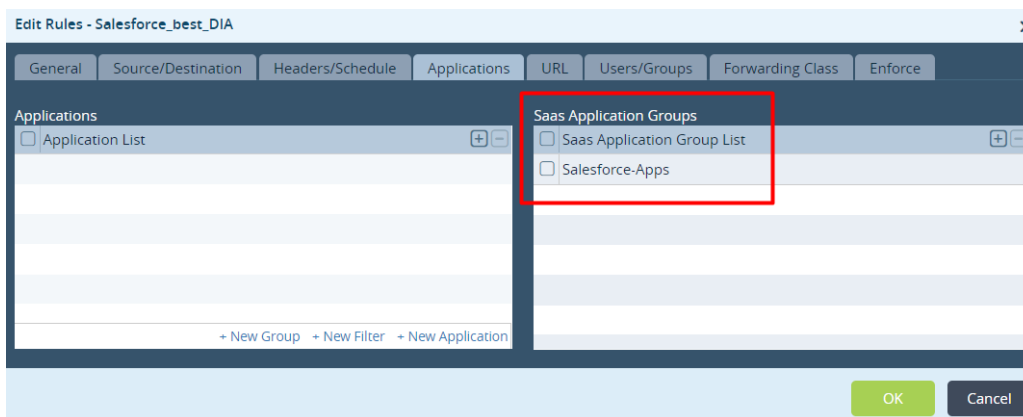


Figure 55 Classify with the application group

Note: This use-case only works with the SaaS Application Group classification.

Now we have defined the classifications, we can link this to the forwarding profile that was previously created:

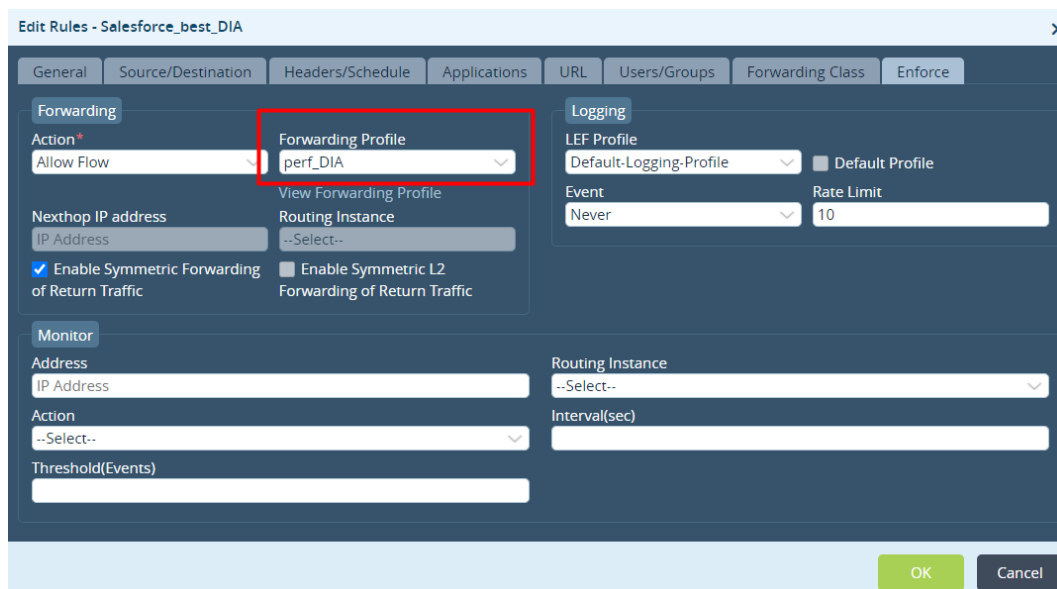


Figure 56 Enforce the forwarding profile

8.5.2 Verification of Performance based breakout.

Now the minimal configuration is in place, we should verify if this is working as expected. We will be using the CLI for this verification.

There are several verification points relevant. First, we will verify the correct operation of the SaaS App Monitor in both local breakout points as well as the remote breakout.

The local breakout point measures the latency to Salesforce of 225ms:

```
admin@Site3-cli> show application-monitor local detail
APPLICATION
MONITOR
NAME          ROUTING INSTANCE          TYPE  LOSS PERCENTAGE  LATENCY MILLISEC  LOCAL ORGANIZATION  EXPORT ORGANIZATION
-----
salesforce    Internet-ISPs1-Transport-VR  icmp  0.0             225.53            Enterprise1         Enterprise1
```

Where the remote breakout point measure a latency to Salesforce of 215ms:


```
admin@GW1-cli> show application-monitor local detail
```

APPLICATION MONITOR		LOSS	LATENCY	LOCAL	EXPORT
NAME	ROUTING INSTANCE	PERCENTAGE	MILLISEC	ORGANIZATION	ORGANIZATION
salesforce	Internet-ISPl-Transport-VR	icmp 0.0	215.75	Enterprise1	Enterprise1

For Salesforce traffic to reach the remote breakout point, we also need to add the latency of the overlay. We can verify this latency with this command:

```
admin@Site3-cli> show orgs org Enterprise1 sd-wan sla-monitor metrics last-1m GW1
```

SITE NAME	PATH HANDLE	FWD CLASS	LOCAL WAN LINK	REMOTE WAN LINK	LOCAL WAN	REMOTE WAN	TWO WAY	FWD DELAY	REV DELAY	PDU LOSS	FWD LOSS	REV LOSS
					LINK ID	LINK ID	DELAY VAR	DELAY VAR	RATIO	RATIO	RATIO	
GW1	6627844	fc_ef	Internet-ISPl	Internet-ISPl	2	2	171	0	1	0.0	0.0	0.0

Based on the above input data, the SD-WAN logic is computing the best path (in this example based on Latency) and will select the local breakout point as most optimal. This can be verified with this command:

```
admin@Site3-cli> show orgs org-services Enterprise1 sd-wan policies Default-Policy rules nexthop application-monitor detail
```

APPLICATION MONITOR NAME	NEXTHOP PRIORITY	NEXTHOP NAME	NEXTHOP STATUS	NEXTHOP ACTIVE	HIT COUNT	APPLICATION MONITOR	APPLICATION MONITOR	APPLICATION MONITOR	
						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

The final check is to verify if the Salesforce sessions are actually going to the local breakout:

```
admin@Site3-cli> show orgs org Enterprise1 sessions brief | select application salesforce
```

VSN ID	VSN VID	SESS ID	SOURCE IP	DESTINATION IP	SOURCE PORT	DESTINATION 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

The example described in this section is the minimal configuration needs to address this use-case. You might have different requirements or more complicated use-cases. The configuration options are very rich, and many alternative scenarios are possible (more local internet breakout, more different remote internet breakout places, different SLA criteria, etc..).

More details of all configurations required for this scenario can be found in the documentation:

https://docs.versa-networks.com/Versa_Operating_System/VOS_SD-WAN_Configuration/Advanced_SD-WAN_Configuration/Configure_SaaS_Application_Monitoring

8.5.3 Different variants of performance-based breakout.

In the previously described scenario, the example was based on performance-based breakout of a local breakout option vs a remote breakout option.

The same approach can be made also for the scenario where **multiple local breakout** possibilities are provided (WAN1 and WAN2). Each local WAN provided by a different ISP may

provide a different performance characteristic (ISP1 closer to the SaaS then ISP2). Another very common use-case is that the breakout is provided by a cloud security provider. In that case, an IPsec or GRE tunnel will be directing the traffic to the cloud security provider. For all those use-cases, the same logic can be applied.

For all those scenarios, the Forwarding profile selection method should be set to “Automatic”.

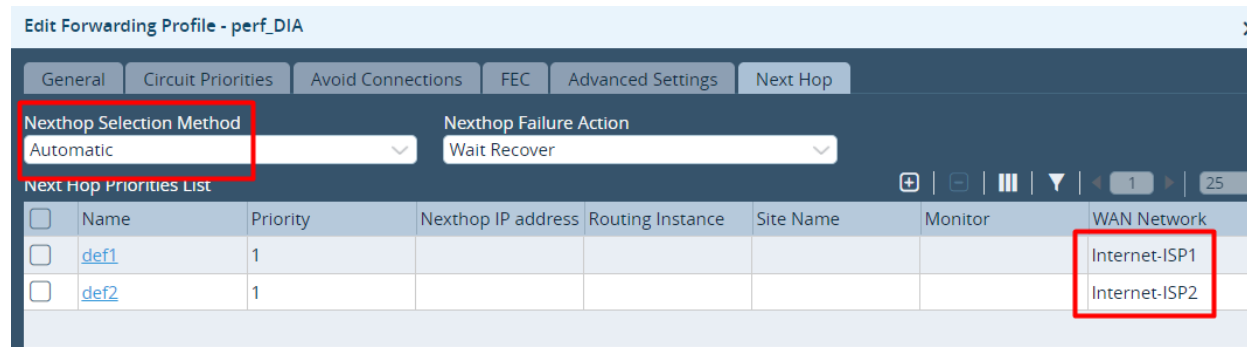


Figure 57 Enforce the best performing local breakout

Verification of best performing local breakout:

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

In the above CLI output, you can see different metrics for local circuit toward ISP1 and ISP2. For this test, the internet circuit to ISP1 is impaired with high loss and high latency, resulting in a bad link-score. Therefore the metric of ISP1 remains low, while the metric of ISP2 is increasing. This results ISP get more session (hitcount=5 vs 1).

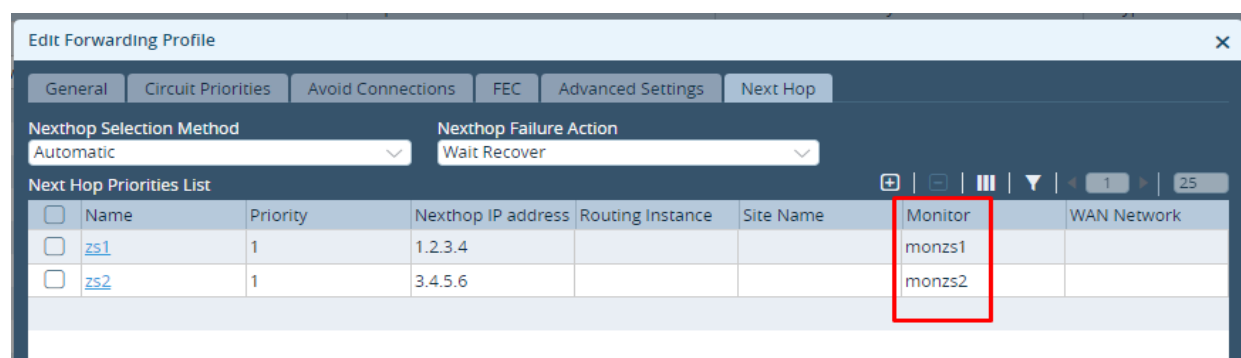


Figure 58 Example of performance-based breakout to an IPsec tunnel with monitor

Another use-case is **performance-based breakout to the best performing external cloud security provider**, the forwarding profile can set the next hop to a specific IP address (which is the IP of the remote tunnel IP endpoint).

For all scenarios, it is recommended to put 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 will no longer be considered.

GRE and IPsec are the most common cloud tunneling options. The basic principles of creating a Site-2-Site IPsec tunnel or GRE tunnel can be found in our documentation:

<https://docs.versa-networks.com/Getting Started/Deployment and Initial Configuration/Branch Deployment/Initial Configuration/05 Configure Site-to-Site Tunnels>

<https://docs.versa-networks.com/Versa FlexVNF/Versa FlexVNF Network and System Configuration/Configure Interfaces>

Once the tunnels are in place, a static route should be added to route traffic into such tunnel. It is recommended that the tunnel endpoint is monitored, so in case the tunnel fails, the static route is withdrawn. This can be done with an IP-SLA monitoring function as illustrated in the diagram below:

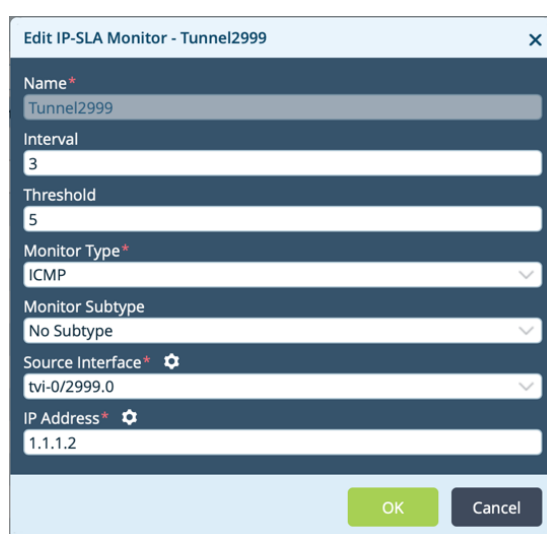


Figure 59 Setting up the monitoring for IPsec or GRE Tunnel

This monitoring function can be added to the Static route. If you desire an active/standby connection to the cloud security proxy, the preference or metric can influence which route to which tunnel will be active.

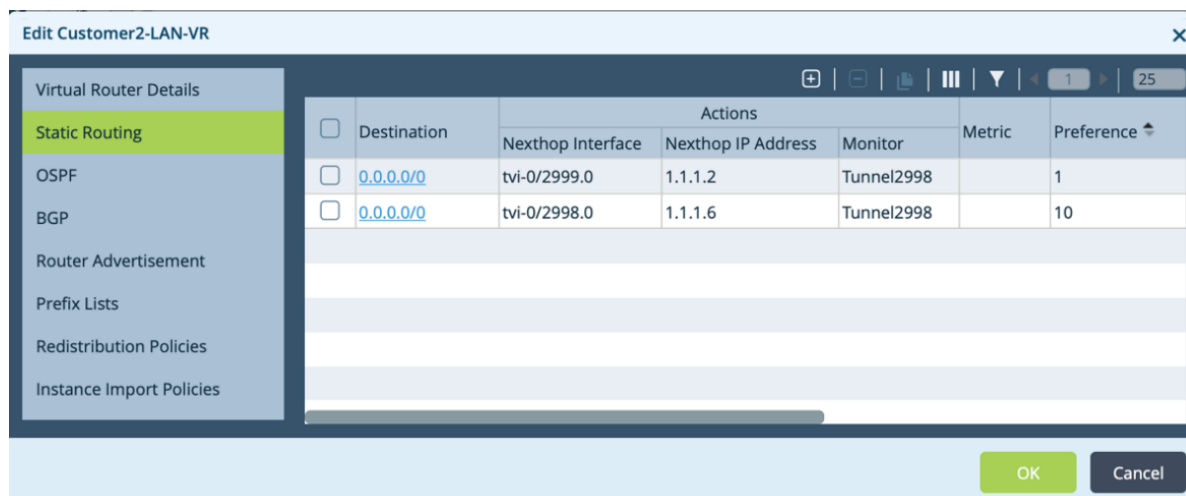


Figure 60 Map default static route to a next hop

8.5.4 Performance based breakout best practices

SaaS App Monitoring has many config options. Your SaaS application may not work for the basic measurement based on ICMP like is used in this example. It might require tuning of the SaaS App Monitor to get the best expected result. Consult your Versa contact point for advice when needed.

Check the next sections on the roles of DNS and Web Proxy. They may interfere with the correct functioning of this application. In these cases, the configuration needs to be enriched with DNS proxy (always recommended) and Web proxy chaining (required when a centralized Web proxy is used).

Important Note!

SaaS applications listed in the SaaS application-group list is suitable for SaaS optimisation. Applications not listed cannot leverage the Performance based breakout functionality.

8.6 Role of DNS proxy for different breakout scenarios

The end user application experience may be depending from where the application is broken out from. For example, with a global deployed Enterprise SD-WAN VPN is using a local internet breakout in Hongkong and a central breakout in London. Imagine the end-users are using an internal DNS server in London. If a user in Hongkong queries DNS from this London based DNS server for Salesforce.com, it will probably get the IP of the London based instance of Salesforce.com. However, in Hongkong there is also an instance of the Salesforce.com SaaS. If the network decides to breakout in Hongkong for Salesforce.com, it should query to the Hongkong DNS to get the IP of the local instance. This requires a **DNS proxy**. The configuration described below illustrates that a local DNS proxy policy rule matches Salesforce.com and based on where the network decides the best breakout point resides, the DNS is leveraged. This might be the Google DNS in Hongkong or the Google DNS in London.

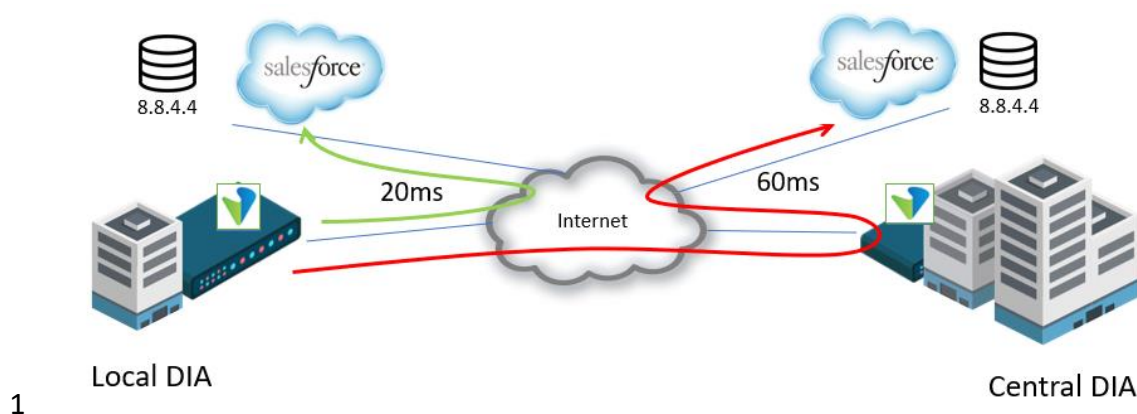


Figure 61 Best performing local SaaS with local DNS

The diagram above shows that internal metrics shows a better performance (in this example measured by latency) to the local SaaS. This local SaaS can only be contacted if the local instance of the global DNS service (Google DNS in Hongkong in this example) is leveraged. Imagine the remote global DNS is leveraged, this would have pointed to the remote SaaS. In that case, the traffic will still be broken out locally, but follow the internet to get to the remote SaaS. If the global DNS would have been leveraged, it would have point to the local SaaS.

The below configuration is described as extension to the Performance based SaaS optimization use-case described in section 8.5.

The DNS proxy configuration is only required on the local DIA Edge Device (no additional configuration required on the central DIA Edge Device).

For Performance based SaaS optimization, using the DNS of the local Internet service provider or the using a global DNS provider (such as Google DNS) is required. In this configuration example, Google DNS 8.8.4.4 should be used for Salesforce.com. First an **DNS Proxy Profile** is configured:

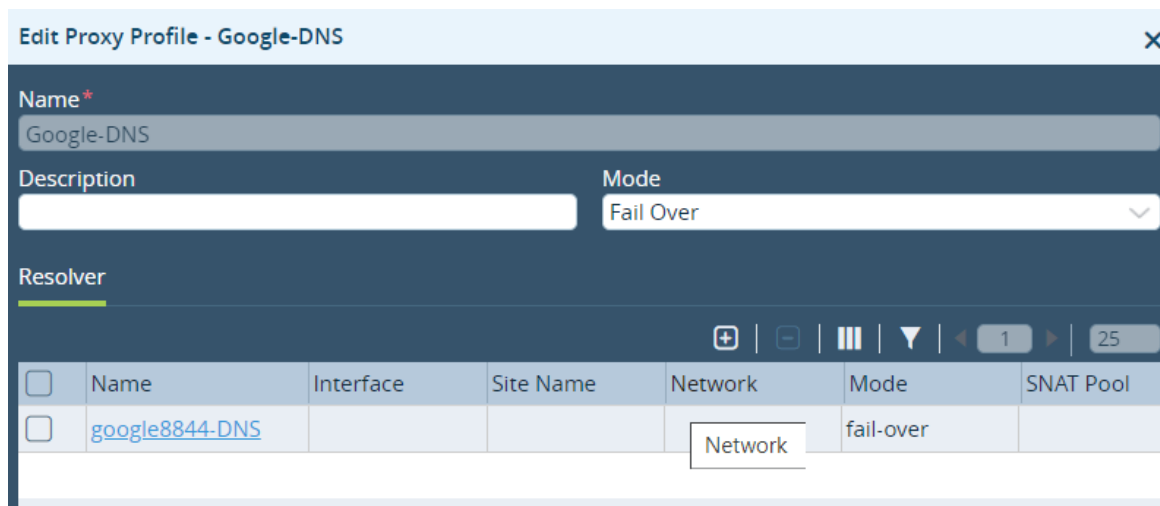


Figure 62 DNS Proxy

In this DNS Proxy Profile, the resolver of the Google DNS is used:

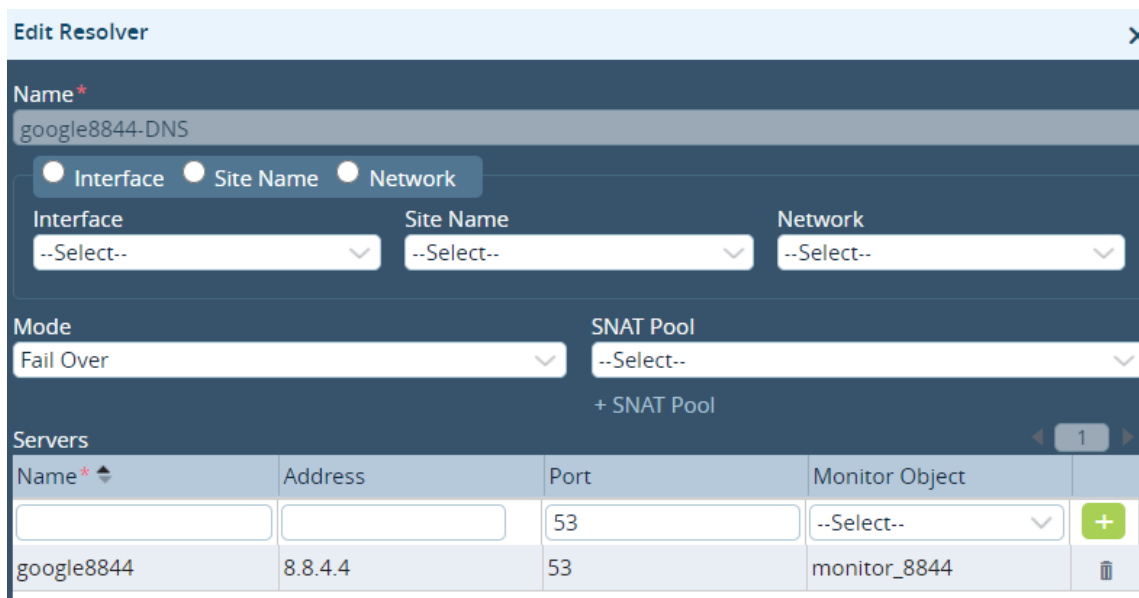
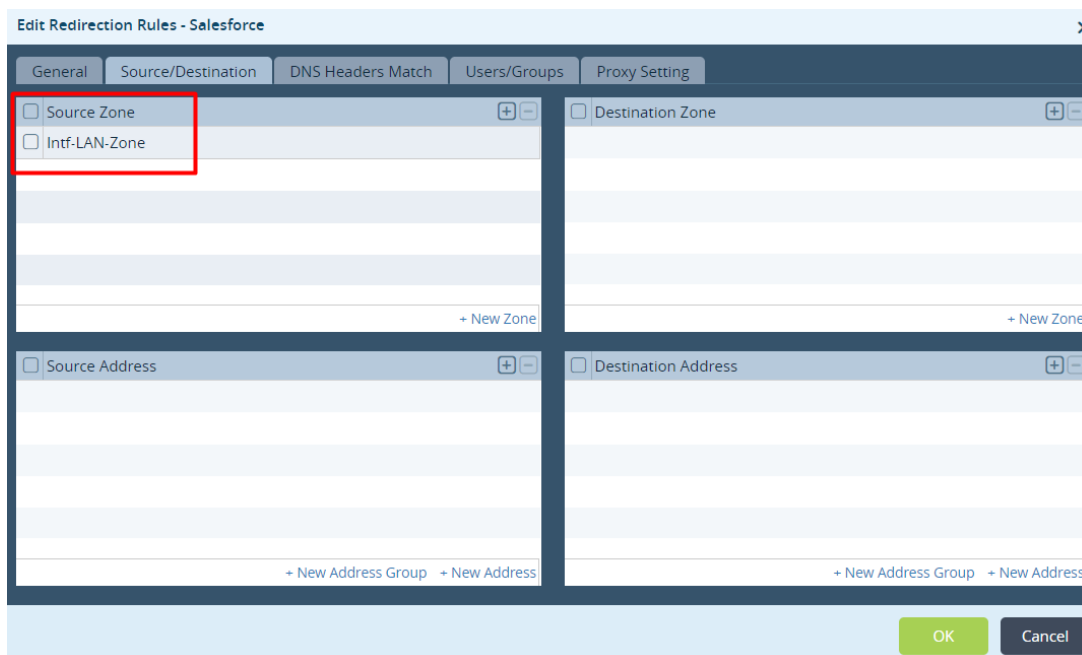


Figure 63 DNS proxy resolver config

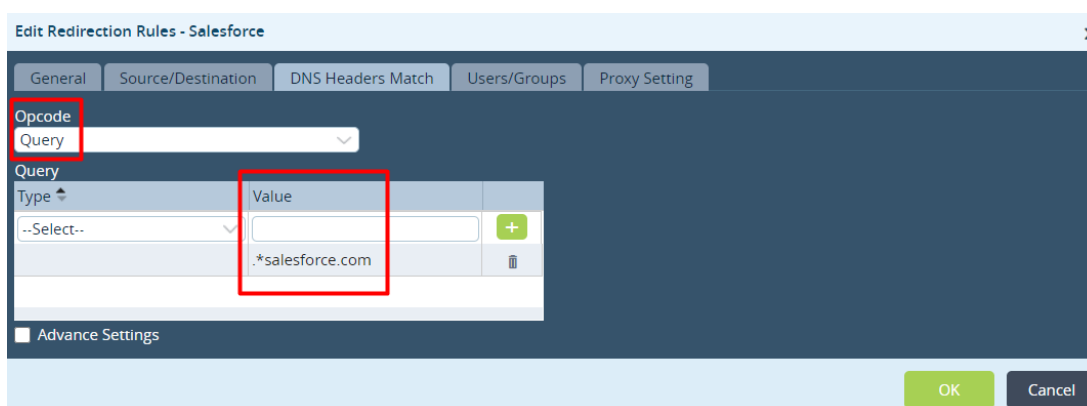
It is recommended to monitor the configured DNS server on reachability. In case the DNS service is no longer reachable, the monitor will fail, and the DNS proxy will stop forwarding DNS queries to this DNS server. The DNS server configured in the client is use instead.

Since the DNS server used is dynamically assigned, for this use-case we should not use any selection of the above part of this window. Only the specification of the DNS server is relevant. We have chosen to monitor the reachability of the DNS service by applying an IP-SLA monitoring object to this. In case the DNS server is not reachable, this DNS profile will not be used. This will avoid potentially blackholing DNS queries.

Next is the configuration of the **DNS proxy policy rule**. This rule specifies the matching conditions (where the DNS query comes from and for what FQDN it applies). The source zone should be listed:

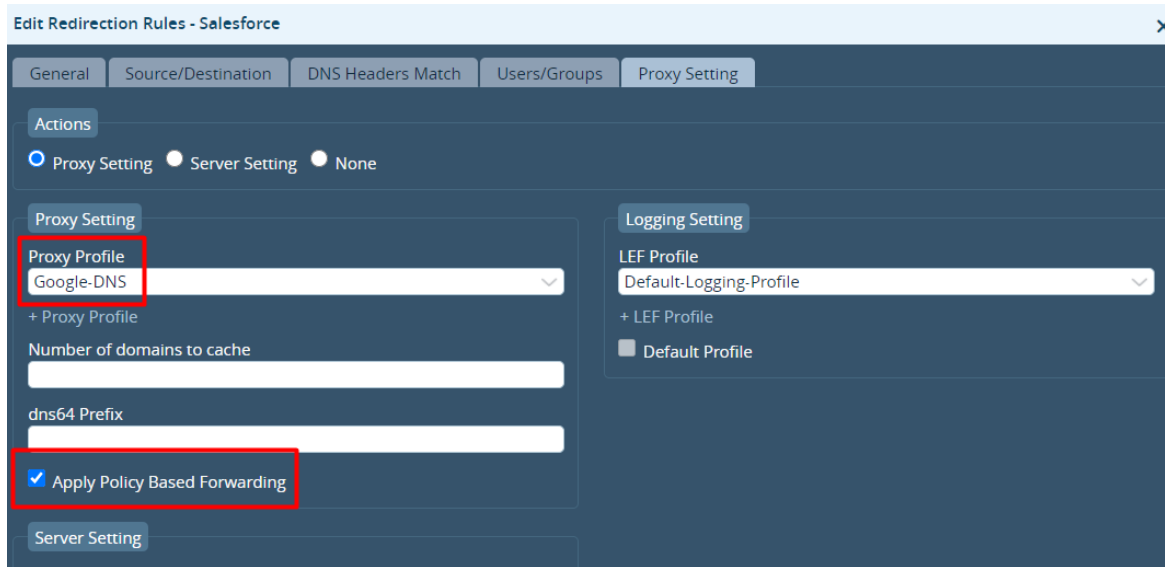


The DNS classifier should match the SaaS application relevant for this policy. The Opcode is **Query** and the Query value is the FQDN regex notation.



The exact value of regex notation is very important. Since we like to query for all traffic designated to the Salesforce.com domain, the regex notation is: `*.salesforce.com`.

Under Proxy Settings, the DNS-proxy-profile is referenced. Also, the configuration option “Apply Policy Based Forwarding” should be selected for this use-case. This enables the reference to SD-WAN policies which ensures that the SaaS application is broken out to the best performing DIA. Without this option set, the DNS proxy (in this specific configuration only) will just follow the routing table to get to the specified DNS server. Since the SD-WAN policy may dynamically decide whether there need to be broken out locally or centrally, the DNS-proxy needs to follow that decision.



We have also enabled a logging profile as it becomes useful in DNS query verification.

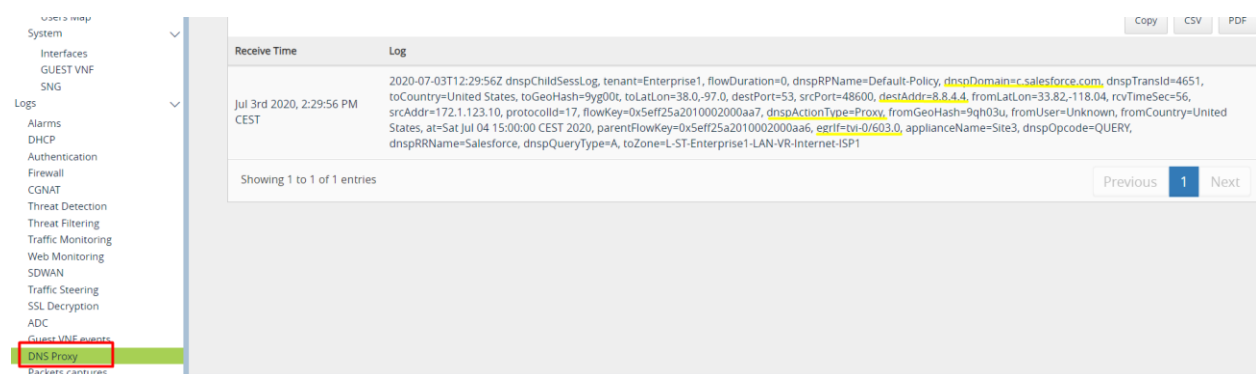
8.6.1 Verification of DNS proxy being used.

To verify if the above configuration functions correctly, it should first be verified what the current best performing path is to Salesforce.com:

```
admin@Site3-cli> show orgs org-services Enterprise1 sd-wan policies Default-Policy rules nexthop application-monitor detail
```

MONITOR NAME	NEXTHOP PRIORITY	NEXTHOP NAME	NEXTHOP STATUS	NEXTHOP ACTIVE	HIT COUNT	APPLICATION MONITOR NAME	APPLICATION MONITOR TYPE	APPLICATION MONITOR LATENCY	APPL 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

From the above command, we can see that the path to Salesforce.com from the local breakout gives the lowest latency (326ms vs 398ms) and therefore that is the active path. If the configuration functioning correctly, the DNS query should be sent out to Google DNS from the local DIA. This can be verified by logging to Analytics.



Zoom in to this log:


```

2020-07-03T12:29:56Z dnspChildSessLog, tenant=Enterprise1, flowDuration=0,
dnspRPName=Default-Policy, dnspDomain=c.salesforce.com, dnspTransId=4651,
toCountry=United States, toGeoHash=9yg00t, toLatLon=38.0,-97.0, destPort=53,
srcPort=48600, destAddr=8.8.4.4, fromLatLon=33.82,-118.04, rcvTimeSec=56,
srcAddr=172.1.123.10, protocolId=17, flowKey=0x5eff25a2010002000aa7,
dnspActionType=Proxy, fromGeoHash=9qh03u, fromUser=Unknown, fromCountry=United
States, at=Sat Jul 04 15:00:00 CEST 2020, parentFlowKey=0x5eff25a2010002000aa6,
egrIf=tvvi-0/603.0, applianceName=Site3, dnspOpcode=QUERY, dnspRRName=Salesforce,
dnspQueryType=A, toZone=L-ST-Enterprise1-LAN-VR-Internet-ISP1
    
```

From the above log details, you can see that for salesforce.com domain, DNS server 8.8.4.4 is used (default DNS for the client is some Enterprise internal DNS) and that the egress Interface is TVI-0/603. This is the internal interface for the DIA function (can be verified with “show interface brief”).

When the local DIA is impaired and the central DIA is preferred, the log should identify this:

```

admin@Site3-cli> show orgs org-services Enterprise1 sd-wan policies Default-Policy rules nexthop application-
monitor detail
    
```

MONITOR NAME	PRIORITY	NEXTHOP NAME	NEXTHOP STATUS	NEXTHOP ACTIVE	HIT COUNT	APPLICATION MONITOR NAME	APPLICATION MONITOR TYPE	APPLICATION MONITOR LATENCY	APPLICATION MONITOR 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

```

2020-07-03T12:42:47Z dnspChildSessLog, tenant=Enterprise1, flowDuration=0,
dnspRPName=Default-Policy, dnspDomain=c.salesforce.com, dnspTransId=18281,
toCountry=United States, toGeoHash=9yg00t, toLatLon=38.0,-97.0, destPort=53,
srcPort=48007, destAddr=8.8.4.4, fromLatLon=33.82,-118.04, rcvTimeSec=47,
srcAddr=172.1.123.10, protocolId=17, flowKey=0x5eff28a6010002000b59,
dnspActionType=Proxy, fromGeoHash=9qh03u, fromUser=Unknown, fromCountry=United
States, at=Sat Jul 04 15:00:00 CEST 2020, parentFlowKey=0x5eff28a6010002000b58,
egrIf=dtvi-0/51, applianceName=Site3, dnspOpcode=QUERY, dnspRRName=Salesforce,
dnspQueryType=A, toZone=ptvi
    
```

In this verification it shows that the DNS service 8.8.4.4 is used, but now the egress interface is DTVI-0/51 which is the overlay tunnel to the gateway (this can be verified with “show interface dynamic tunnels”).

8.6.2 Best Practices for DNS Proxy

In this example, it is described how the SD-WAN network can dynamically intercept the client DNS query and forwards this to the DIA point which give the best SaaS performance. The example describes the minimal configuration required for this use-case to work. More advanced scenarios of the use of DNS proxy can be imagined. For example, Enterprises often have their own internal DNS server. The IP endpoints may have a global DNS external service (e.g. Google DNS) configured or the internal DNS is configured, but local breakout to a global external DNS is preferred. For those use-cases DNS proxy will also provide a solution. The proxy intercepts DNS traffic for the internal domain (based on the internal FQDN), while the rest is served by the Global external DNS.

8.7 Importance Security services for internet breakout

Traditionally Enterprises follow strict security policies for connecting the LAN to Internet. In legacy networks, where Internet access is typically provided from a centralized Internet breakout, at one location security can be enforced easily. A centralized stack of firewall services, often acting as Web Proxy, will inspect the traffic coming from or going to Internet against the Enterprise security policies. However, now having access to local internet breakout services, Enterprises security policies should not change so scrubbing of the internet bound traffic is typically still desired. In summary, those different scenarios are observed in the field:

- No additional Security at local internet breakout (rely on NAT to secure your Enterprise LAN).
 - Very unsecure strategy and no controls to monitor or mitigate compromised networks.
- Enforce security policies through a nearby Cloud Security Proxy of some 3rd party Cloud Security Provider.
 - Secure alternative, but costly and Cloud Security Proxy Services may not always be provided locally, resulting in sub-optimal end user quality of application experiences.
- Enforce security policies through 3rd party firewall. All internet bound local breakout traffic is scrubbed by such 3rd party firewall. In case of the Versa solution, this 3rd party firewall is typically virtualized as guest VM on Versa uCPE.
 - Secure, but complex to manage. Separate security solution needs to be lifecycle managed.
- Enforce security policies using the Versa embedded NGFW with UTM.

Furthermore, enterprise security administrators may also make risk assessments on the type of application to be broken out locally (without full stack inspection). If local internet breakout is done only for e.g. Office365 (all other internet traffic is broken out centrally), Microsoft claims their O365 service is very secure by nature (Microsoft claims full stack security is done on their applications). For performance reasons they suggest not to enforce additional security policies. If the enterprise adopts this suggestion, they may can ignore additional implementation of security services.

Versa recommends for any breakout to internet, to deploy at least the IDS/IPS with the “Versa Recommended profile” and Anti-Virus. This will ensure that all traffic coming from an untrusted network gets inspected (regardless if the application provider claims to be secure or not). It is simple to add this recommendation to the configuration. Note that SSL decryption at the branch is also required to effectively investigate the SSL stream.

8.7.1 Configuring NextGen Firewall security for DIA traffic

It must be noted that all Firewall Policies rules which have an “allow” rule and addresses an internet bound application (and therefore logically will be internet broken out), an UTM security profile should be attached.

From the images shown below, the user should configure a Security Policy under Next Gen Firewall Service. In this security policy the Destination Zone must be the “L-ST-<org name>-

VR-<WAN name> zone. In this example, there are two local internet breakout points that can be used, so the matching zone should be on both. All other matching criteria can be left blank.

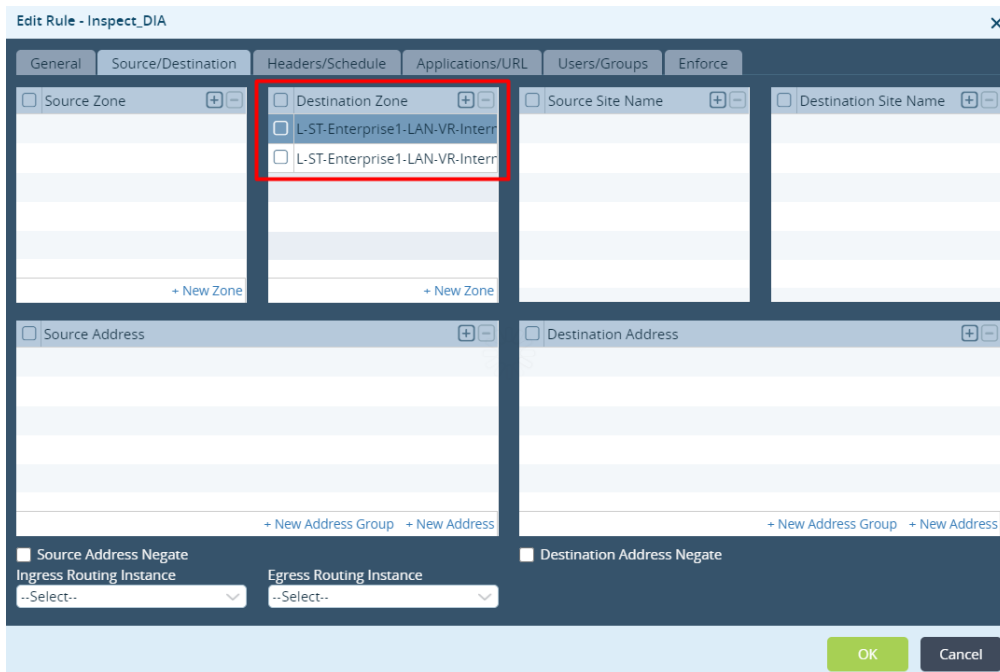


Figure 64 Security policy rule to match DIA traffic

In the Enforce Tab, the Security profile can be selected. In this example we would like to do Malware inspection by using the Anti-Virus module and IDS/IPS using the Vulnerability module.

A best practice with all security services is to enable logging to report possible vulnerabilities detected.

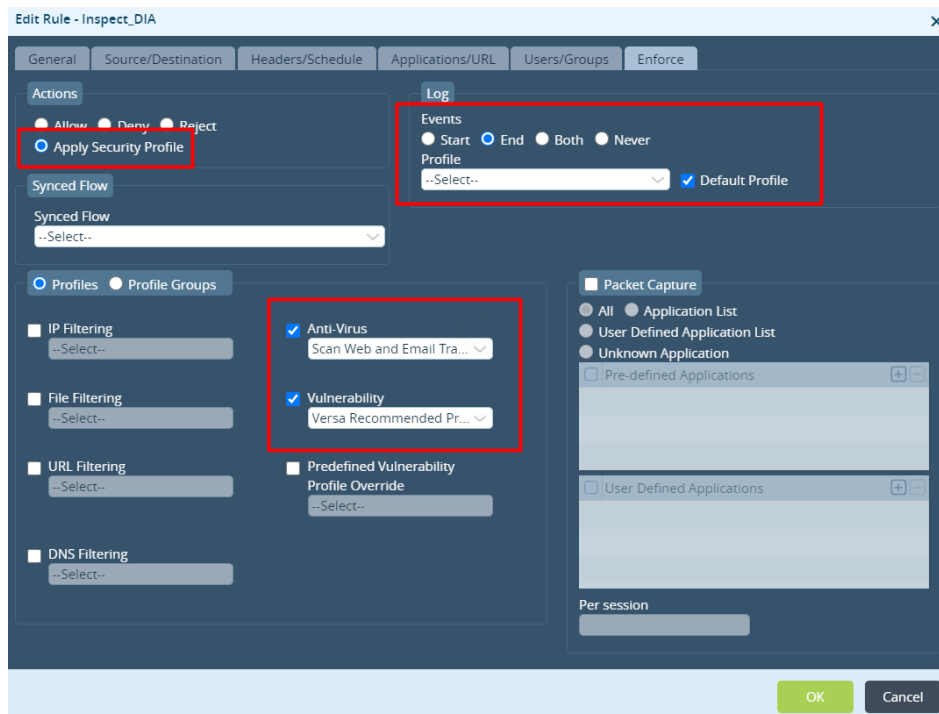


Figure 65 Enforcing Security Profile for DIA

8.8 Role of Web Proxies with different breakout services

Another common use-case for Internet Access with Enterprise security best practices is to use Web Proxies. The use of Web Proxies has a few benefits with enterprises as it provides more control of what traffic is going to internet (all internet bound traffic will be inspected by the web proxy). Another benefit is that with the use of web proxies, it is no longer needed to announce a default route in the enterprise LAN. Instead, endpoints reach internet by connecting to the proxy server as specified in the end-point configuration (possible by the use of a PAC).

However, if the Enterprise likes to leverage Quality of Experience improvements for specific SaaS applications as discussed in this section above, the connection to the centralized proxy service is required to be intercepted to allow local internet breakout. This local DIA point should intercept the proxy session and act as alternative web-proxy dynamically, without the end-point proxy configuration to be modified.

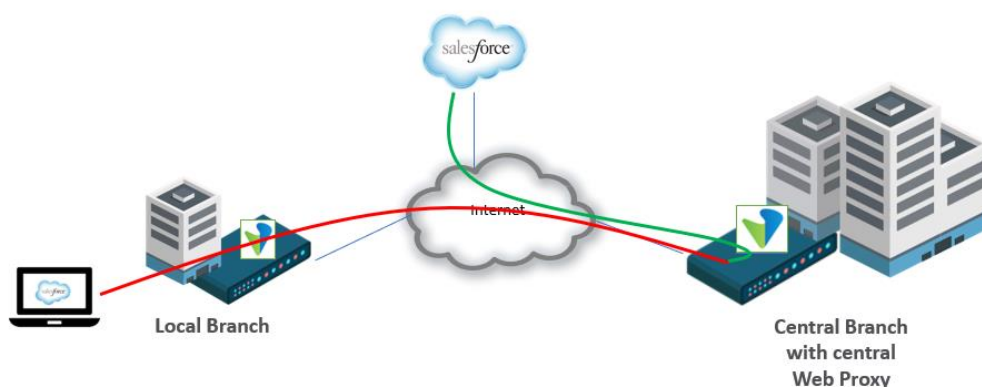


Figure 66 Centralized breakout through Web Proxy

Versa VOS Edge Device software can function as Web Proxy as illustrated in the diagram above. Alternatively, an existing 3rd party Web Proxy at central side can also be leveraged. When it is required for the local branch to do a local internet breakout, it will become difficult as the Client has configured the central Explicit Web Proxy endpoint in its configuration. Versa VOS Edge Device software Web Proxy is enhanced to also address this use-case.

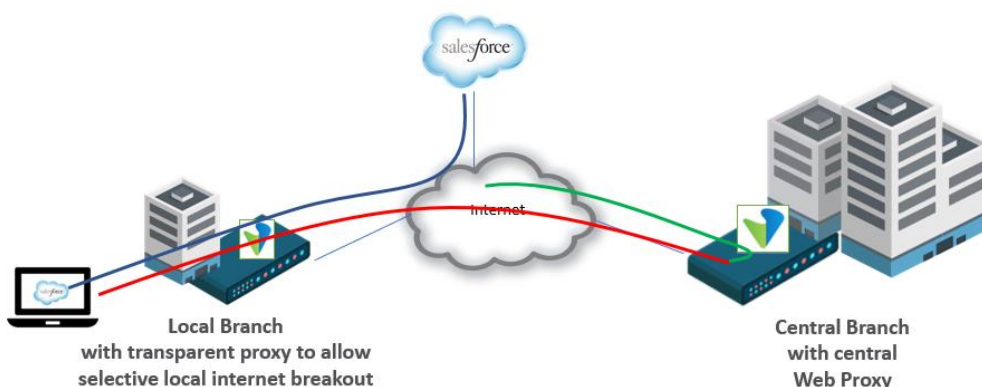


Figure 67 Local Proxy Breakout

In this scenario, the local branch is configured to provide Local DIA for Salesforce.com only. As the client computer sends all HTTPS traffic by default to the central explicit web proxy, the local branch will monitor on the explicit proxy port number for queries to Salesforce.com.

Once those are detected, the local branch can do internet breakout using a transparent web proxy and forward the traffic to internet. The previously discussed DNS proxy feature needs to be configured that DNS queries to Salesforce.com will be intercepted and redirected to the DNS service on the local internet service to retrieve the best DNS query for a local instance of Salesforce.com

A more advanced use-case is where the multiple Web proxies are in service. Besides the previously described scenario of redirecting local internet breakout for specific SaaS traffic, it is also possible to redirect specific traffic to an alternative proxy service. This is called proxy-chaining

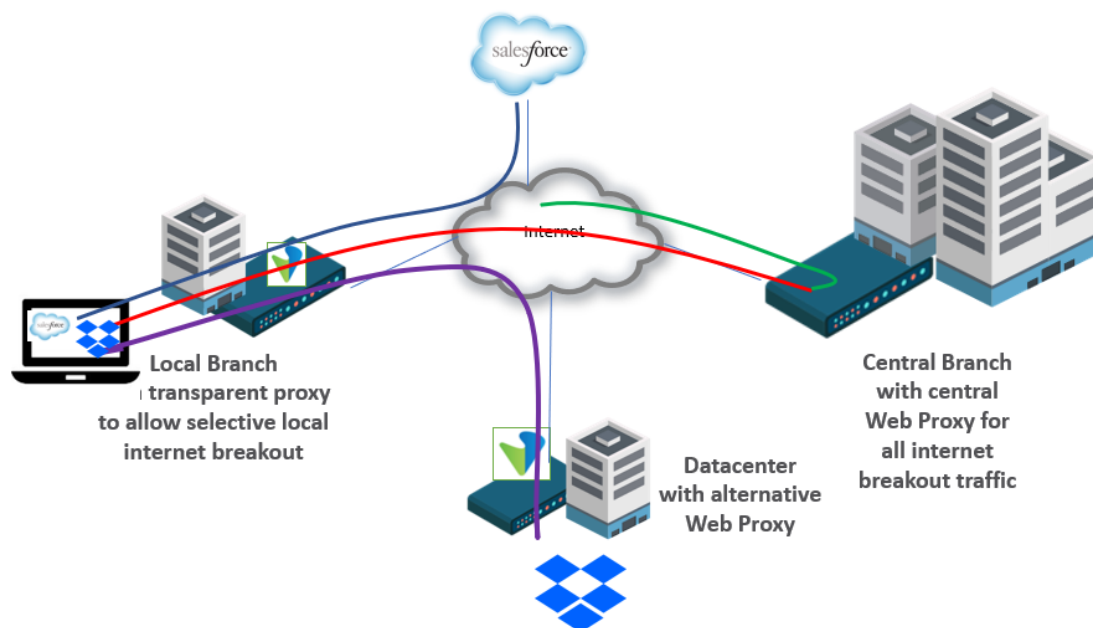


Figure 68 Proxy Chaining to Alternative Web Proxy

The diagram above describes that the Versa VOS Edge Device in the datacenter 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 breakout Salesforce.com directly to internet and proxy-chain dropbox.com traffic to the datacenter branch. All using their own DNS proxy configurations to ensure the correct DNS is used.

The above Web Proxy configuration will be defined in more configuration detail in the next release of this document.

9 SD-WAN Gateway

Versa SD-WAN Gateways (VOS Edge Device) are used to address three main use cases which are:

- Connecting to sites on an MPLS L3VPN Network
- Connecting sites over disjointed underlay networks
- Gateway for Internet bound traffic

9.1 MPLS L3VPN Interworking

The SD-WAN Gateway functionality allows sites that are connected to the SD-WAN VPN network to communicate with sites that are connected to the legacy MPLS VPN network. To enable this communication a Versa VOS Edge Device can be configured as gateway. It requires that there is an exchange of routing information between the MPLS underlay network and the SD-WAN VPN network. This exchange of routes is typically done using a dynamic routing protocol like BGP.

A gateway may be setup temporarily during SD-WAN migration when sites on the MPLS L3VPN network need to connect to sites that have migrated to SD-WAN. Once all the sites have migrated to an SD-WAN based VPN then the gateway is no longer required. A gateway may also be permanent when there are sites or services on the MPLS network that need to be accessed by SD-WAN enabled sites. An example of an MPLS VPN based service could be an Azure Express Route connection that need to be accessed by SDWAN enabled Branches.

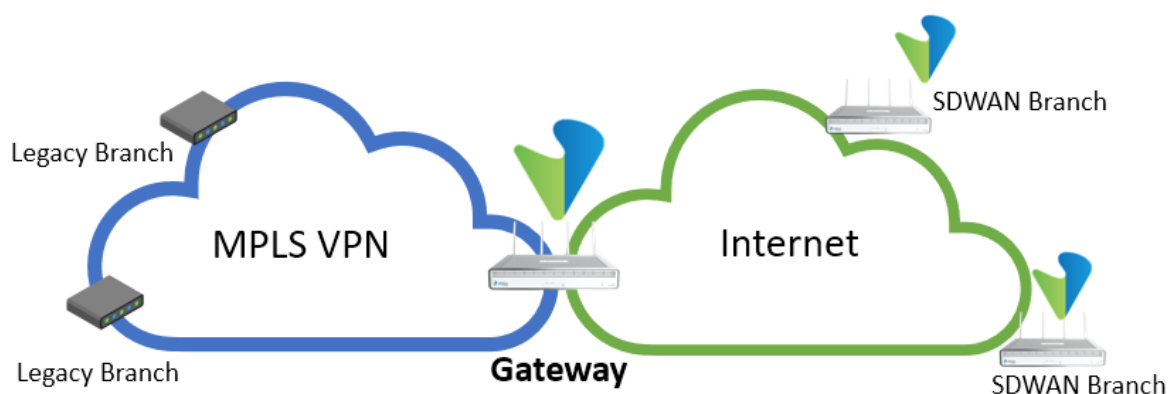


Figure 69 Gateway representation

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 multi-tenant and in a Service Provider environment it is common to find such multi-tenant Gateways that serve multiple Customer networks.

This can also be used for Non-SDWAN interconnects where the MPLS L3VPN can be replaced with any IP network.

9.1.1 Best Practices for SD-WAN Gateways

1. Use HA enabled sites as gateway for device level redundancy.
2. Use multiple gateways preferably across different availability zones or geo-regions for redundancy.

3. Keep traffic symmetric. When exchanging routes with the Provider, modify the BGP path attributes to ensure that bi-directional traffic uses the same Gateway device. Asymmetric routing may cause issues in the SD-WAN when SD-WAN traffic steering policies are required.
4. In the case of multiple Gateway sites where the peering is established to the same MPLS provider then preferably use the same BGP AS-number on the SD-WAN and provider side in all BGP neighborships. In addition, the default SD-WAN iBGP AS no is 64512. AS-PATH check ensures that routing loops are not formed when a route learnt from one gateway is inadvertently advertised back at another gateway. BGP AS-path check will take care of routing loops in case the same MPLS Provider learnt route gets re-advertised back at any Gateway site. Similarly, an SD-WAN route advertised from one gateway to the provider should not be re-advertised back by the Provider at another gateway site.
5. It is also advisable to color the routes using BGP communities when they are advertised and received at every BGP Gateway. An example could be coloring routes based on the region where they belong e.g. EMEA and America based routes can have different communities for identification. BGP Filters should be configured in the eBGP neighborhood with the MPLS Provider to only accept or advertise routes that are required. These filtering policies can be crafted based on the BGP community values.
6. An existing branch/DC/hub site can also double up as the gateway. Care should be taken to dimension the system to be able to handle the amount of traffic expected.
7. On gateways and hubs, it is always recommended to use unique network names for the WAN links. The WAN link names should be unique across the entire SD-WAN network.

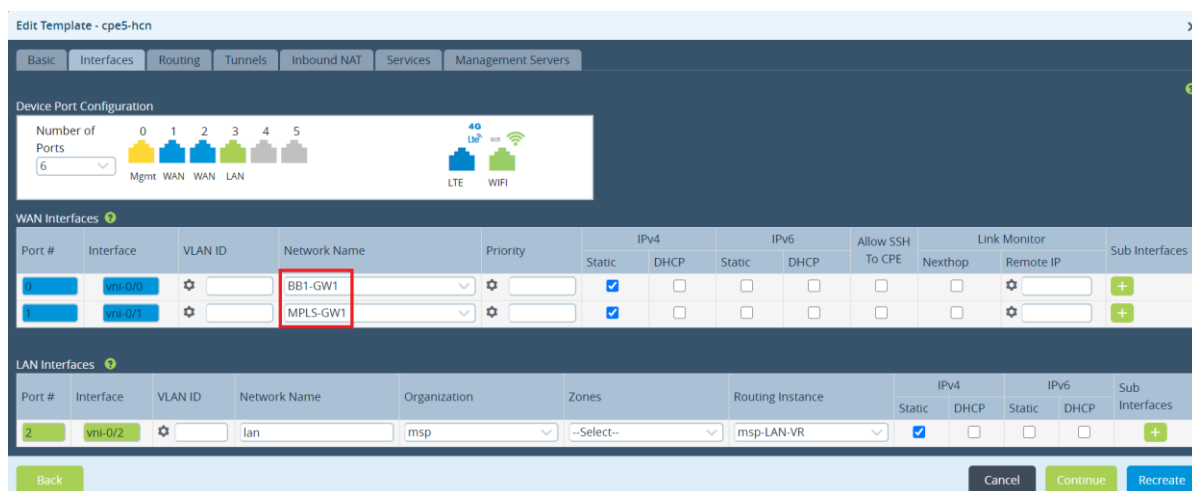


Figure 70 Workflow for unique WAN network names

This allows remote branches to use the remote circuit option to influence the traffic path in the SD-WAN policy. This is not specific to disjointed underlay networks but is a general recommendation for gateways and hubs. Such selection is done in the SD-WAN Forwarding Profile of the remote branch.

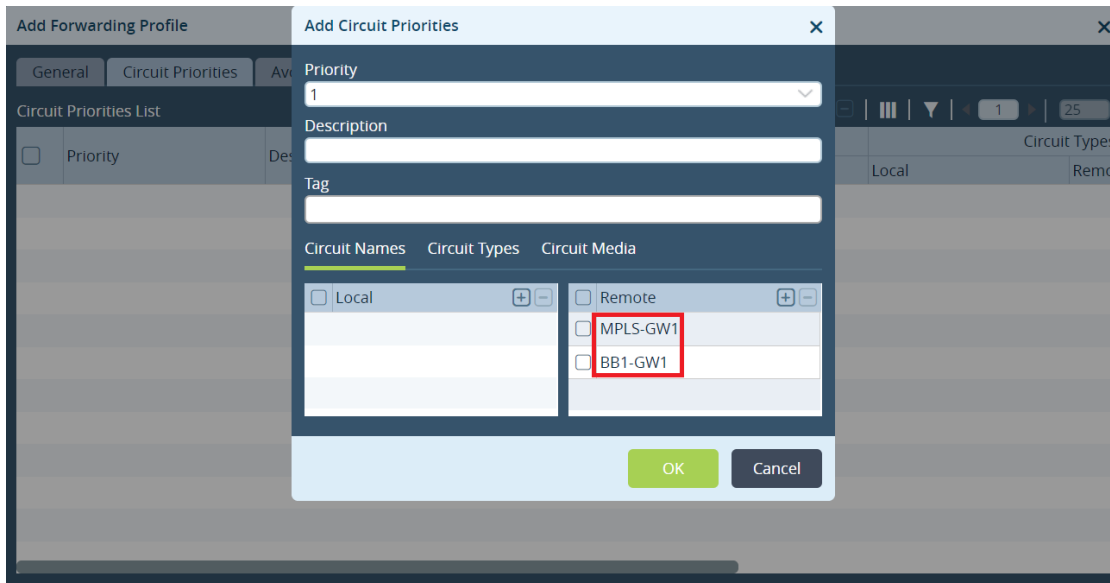


Figure 71 Forwarding Profile with Remote Circuit Name for Path Selection

9.2 Options to interconnect legacy networks with SD-WAN Gateways

There are two ways of configuring a gateway. This depends on how routes are exchanged with the MPLS Provider.

Option 1: BGP to exchange routes with the MPLS provider on the MPLS WAN interface

Option 2: BGP to exchange routes with the MPLS provider on the LAN interface

9.2.1 Option 1: BGP to exchange routes with the MPLS provider on the MPLS WAN interface

This method is most common for DIY Enterprise SD-WAN deployments.

In this case, the BGP neighborship to exchange routes from the MPLS Provider to the SD-WAN Network is configured in the MPLS-Transport-VR of the Gateway. This configuration can be automated through workflows.

The advantage of this approach is that you only need a single IP Interface to connect to the MPLS Provider over which the BGP neighborship is established. This is also seen commonly in Enterprise network designs where there is a single MPLS Provider that 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 only uses the respective WAN IP addresses for establishing connectivity to other SD-WAN branches and the Gateway. The SD-WAN Branch routes are exchanged over an encrypted overlay.

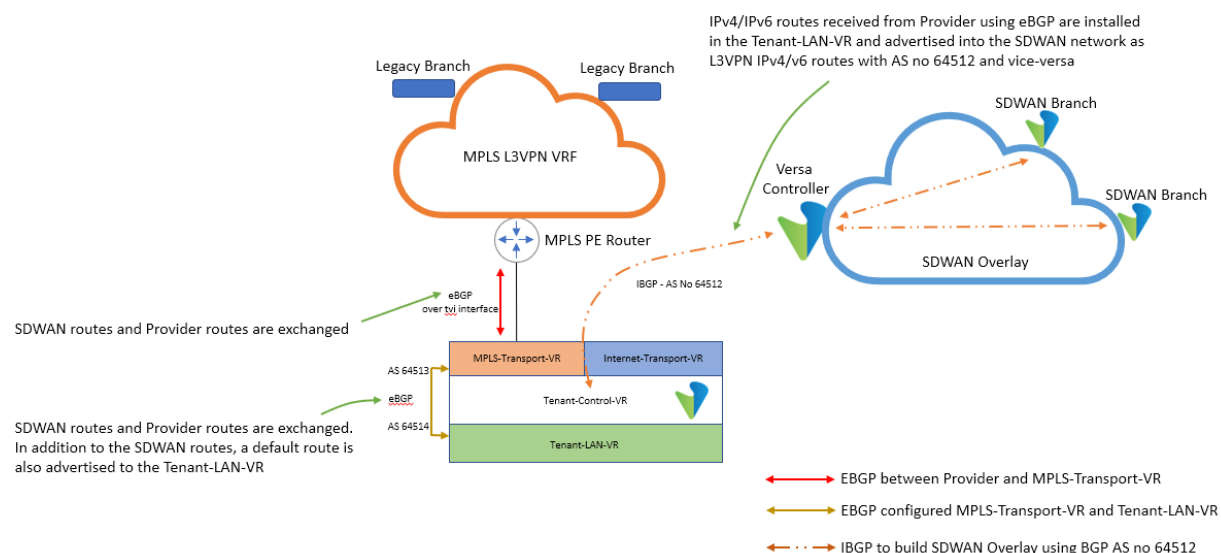


Figure 72 SD-WAN Gateway route exchange with MPLS Provider over WAN

The above picture shows how the gateway is configured internally. A virtual TVI interface pair is created (with workflow) in the MPLS-Transport-VR and Tenant-LAN-VR. Over this virtual interface an eBGP neighborship is established with AS-numbers 64513 & 64514. When the route is advertised into the SD-WAN network the default BGP AS-number 64512 of the SD-WAN overlay is appended to the BGP AS-Path list. If the SD-WAN route advertised by one gateway is inadvertently learnt by another gateway through the MPLS underlay provider then the route is automatically blocked due to the AS-Path check.

The configuration is as follows:

1. To create the eBGP session on the MPLS WAN link with the MPLS Provider PE router. For the VOS Edge Device, this can be done through the workflow.

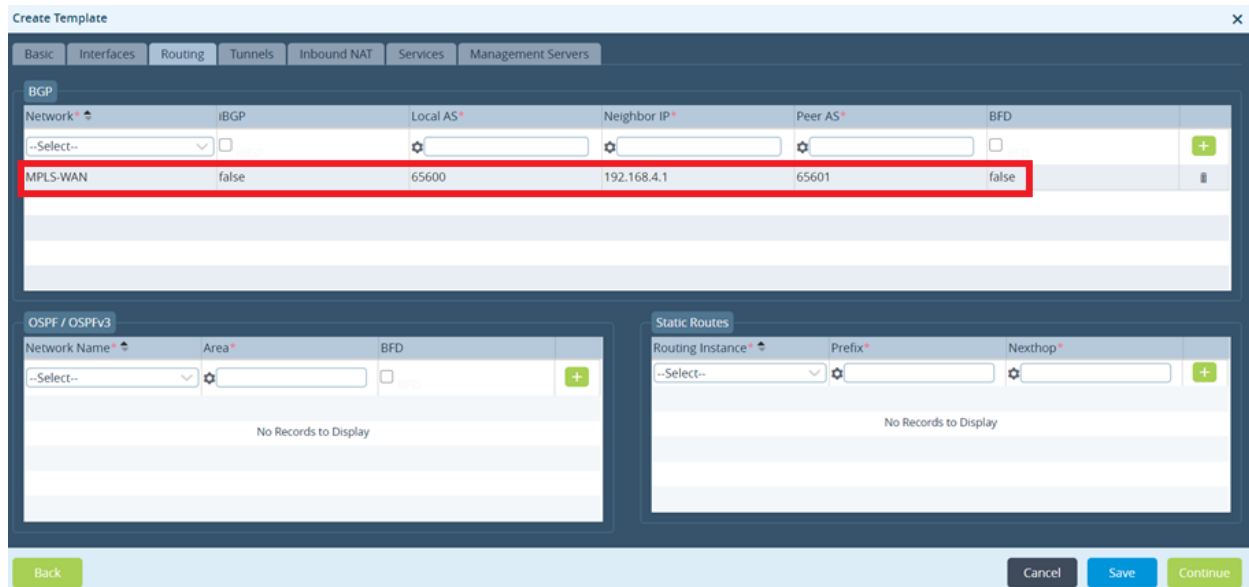


Figure 73 Workflow for eBGP towards PE router over WAN

2. To create the eBGP session between the MPLS-Transport-VR and Tenant-LAN-VR over a virtual interface pair (TVI interface).

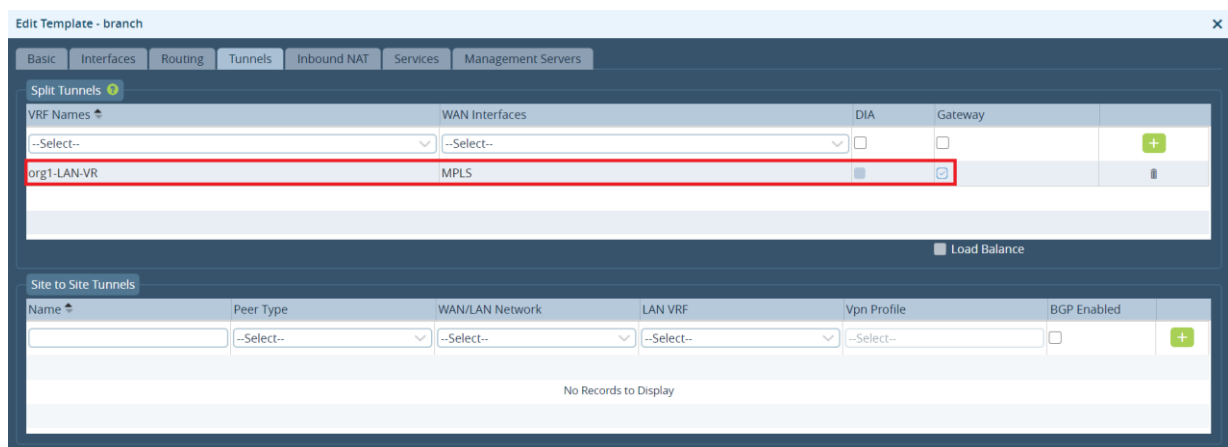


Figure 74 Workflow for eBGP between MPLS-Transport-VR to LAN-VR

9.2.1.1 High Availability for option 1

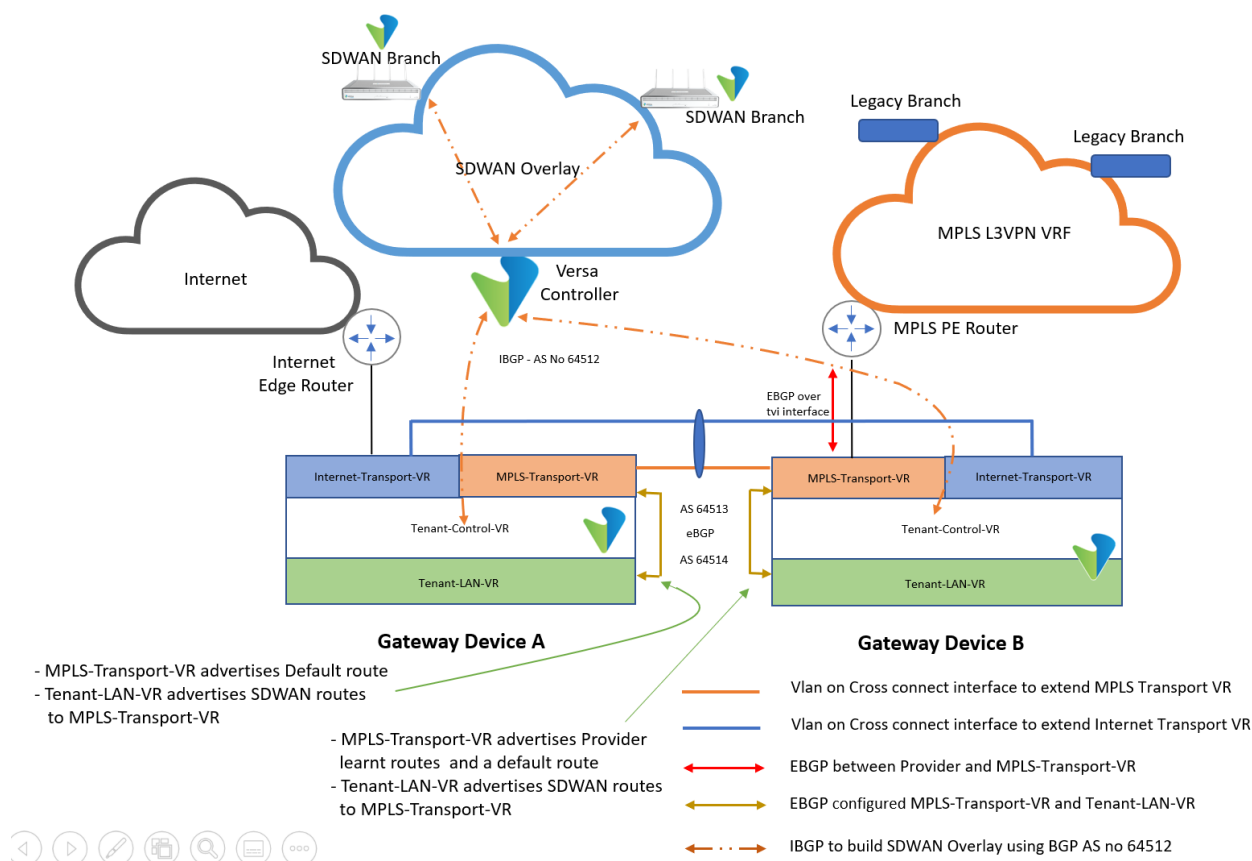


Figure 75 Gateway in HA Mode

When the branch is configured as Active-Active HA with the gateway option selected then there is a TVI interface created between the MPLS-Transport-VR and the Tenant-LAN-VR on both devices. Over this TVI interface an eBGP neighborhood is established with AS-Numbers 65413 and 65414. Between the two devices there is a cross-connect cable that extends the MPLS-Transport-VR to Device A and extends the Internet-Transport-VR to Device B.

In the above picture, on Device B where the MPLS link terminates, there is an eBGP session with the MPLS Provider over the MPLS wan link. Over this eBGP session, exchange of routes between the SD-WAN network and the legacy MPLS L3VPN VRF network happens. In addition to the MPLS L3VPN VRF routes, the MPLS-Transport-VR also advertises a default route to the Tenant-LAN-VR. When the routes from the MPLS provider are installed in the Tenant-LAN-VR, they are then advertised into the SD-WAN network.

Furthermore, the following HA design recommendation can be given:

- On Device A, the eBGP session between the MPLS-Transport-VR and Tenant-LAN-VR only advertises 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 of Device B using the cross-connect interface. The routes received from the MPLS Provider on Device B are learnt on Device A using the SD-WAN iBGP. The Tenant-LAN-VR will not receive the MPLS Provider routes over the eBGP session from the MPLS-Transport-VR. Hence Device B will always be the primary path for all traffic to and from the MPLS Provider. On Device A and Device B, the eBGP session between the MPLS-Transport-VR and Tenant-LAN-VR

advertises the SD-WAN routes to the MPLS-Transport-VR and a default route to the Tenant-LAN-VR. As mentioned in the last section if the default route is not required to be advertised then it can be rejected in the Redistribution Policy in the MPLS-Transport-VR on Device-A and Device-B.

- It is recommended to run a routing protocol like BGP or OSPF on the LAN to avoid any issues with asymmetric routing. But a failure on the LAN router connected to Device B can result in the LAN network only relying on the default route from. Instead of a default route the Device A can also be configured to advertise the specific MPLS Provider learnt routes. To achieve this, iBGP can be configured between Device A and Device B over the MPLS Cross connect link. See sample configuration below.
- If VRRP is used on the LAN for redundancy, then the active device in HA should terminate the MPLS transport link to maintain traffic symmetry.
- If the device is a dedicated gateway with just one Internet and MPLS link then there is little purpose in configuring it as HA in workflows because if one of the wan links or devices fail, the gateway is non-functional. If the device is not a dedicated Gateway i.e. it also functions as a regular site, hub or datacentre branch then it makes sense to configure HA using Workflows as the device is usable even when one of the WAN link or devices fails. If the device is a dedicated gateway and it is possible to replicate the wan links, then have 2x single gateway devices for redundancy.
- Since Device A in the picture above does not receive the routes from the MPLS Provider it relies on a default route which is advertised into the SD-WAN network and LAN network. In some scenarios it might be helpful that Device A also has a copy of the MPLS Provider routes rather than relying on the default route. To achieve this, iBGP peering can be configured between the MPLS-Transport-VR on both devices over the MPLS Cross-connect interface.

9.2.1.2 Best Practice for option 1

- There is a default route that is created in the MPLS-Transport-VR pointing to the Provider MPLS PE router as the next-hop. This default route is advertised into the SD-WAN network over the eBGP session between the MPLS-Transport-VR and Tenant-LAN-VR. If this default route is not required to be advertised into the SD-WAN network, then delete the term T1-STATIC from the Redistribution Policy ST-Policy in the MPLS-Transport-VR.

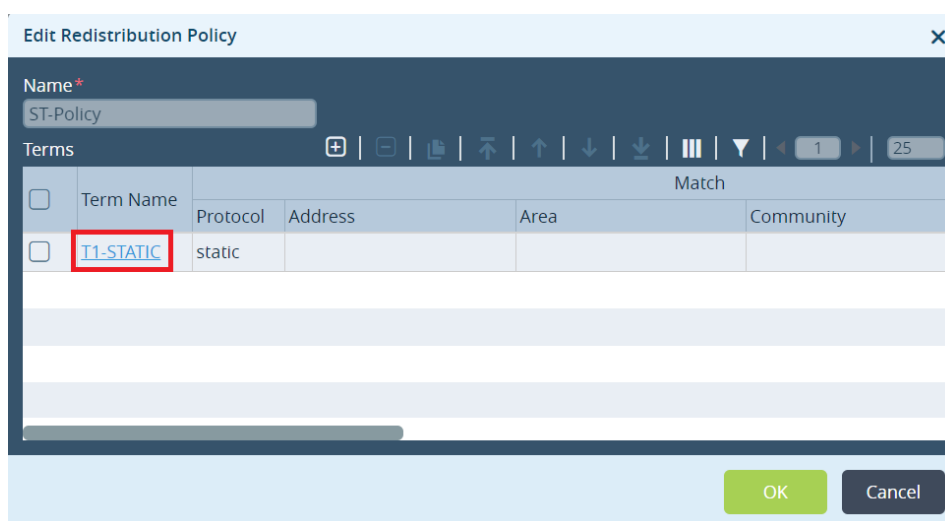


Figure 76 eBGP T1-STATIC redistribution Policy change

9.2.2 Option 2: BGP with the MPLS provider on the LAN interface

In this case, the BGP neighborhood to exchange routes with the MPLS Provider is configured through an IP interface in the Tenant-LAN-VR. This configuration can be easily done through Workflows.

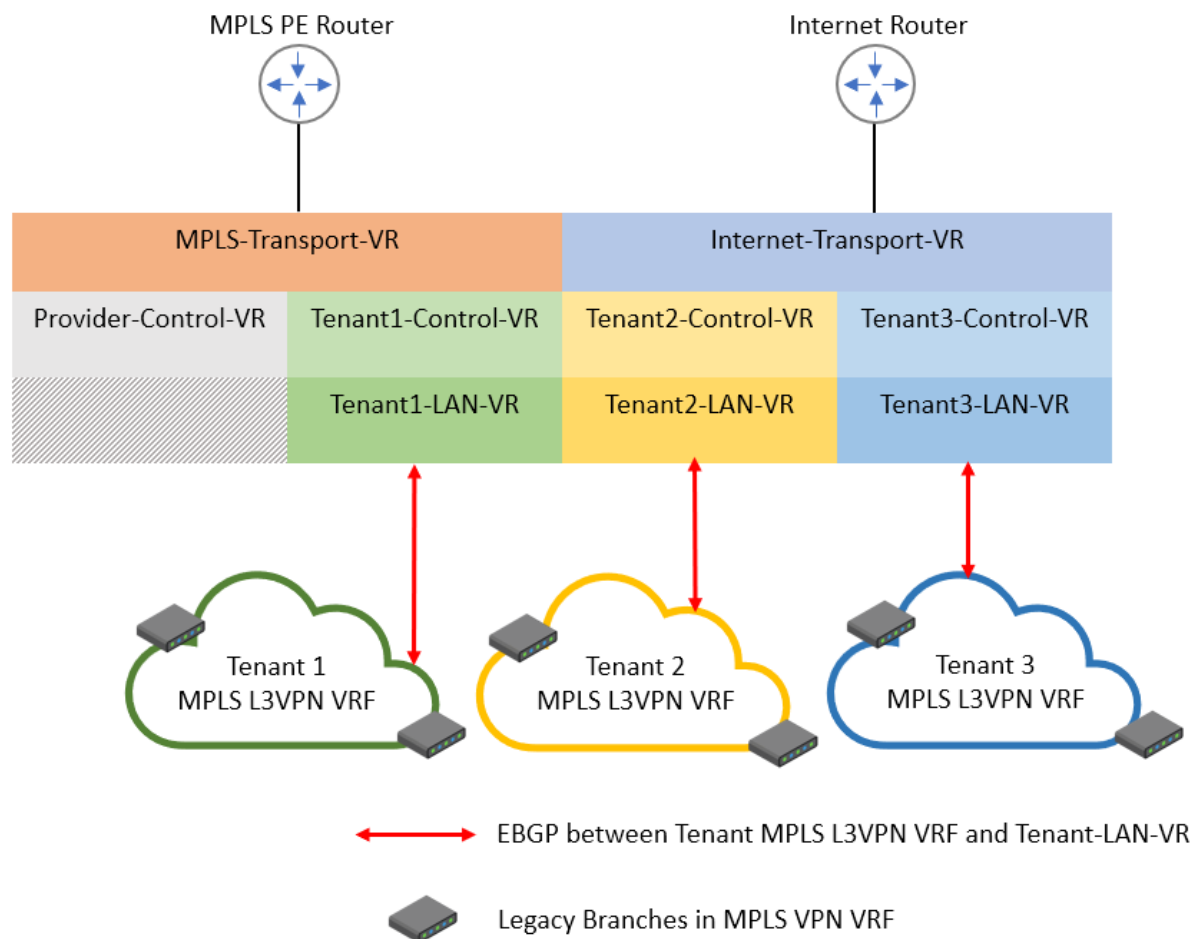


Figure 77 Gateway with Route Exchange with Provider in LAN VR

In the case of enterprises hosting their own gateways, they would need two interfaces or logical sub-interfaces from the MPLS Provider, one that terminates in the MPLS-Transport-VR which is the underlay to connect to the SD-WAN enabled sites and the other that terminates in the Tenant-LAN-VR and which is used to connect to non-SDWAN sites. Using BGP in the LAN-VR, the LAN & WAN routes of the non-SDWAN enabled sites are learnt. The LAN routes of the SDWAN enabled sites are also learnt since the Gateway is part of the SDWAN overlay established over the WAN link. This allows communication between the SDWAN enabled sites and the non-SDWAN enabled sites.

This configuration might be typically seen in a dedicated gateway service offered by a service provider who is also the MPLS service provider to their customers. This scenario is depicted in Figure 82. In such a scenario, the MPLS provider connects one interface or sub-interface of the Gateway into the the corresponding customers MPLS L3VPN VRF. On the Gateway, this interface is terminated in the customers LAN-VR. This enables the routing table of the

customers MPLS-L3VPN VRF to be learnt over BGP and installed in the corresponding LAN-VR of the Versa Gateway bringing in all the LAN and WAN routes of the customers non-SDWAN network into the Gateways LAN-VR routing table. Next, the LAN routes of the customers SDWAN enabled sites must be introduced in the LAN-VR of the Gateway to allow communication between the SDWAN enabled and non-SDWAN enabled sites.

To bring the LAN routes of the SDWAN network of each customer in the corresponding LAN-VR of the Gateway, a common WAN link in the MPLS-Transport-VR routing instance is used. This WAN link connects to a separate MPLS L3VPN VRF. This VRF has all the WAN routes of the SDWAN enabled branches and the WAN routes of the Gateway. This allows the Gateway to communicate directly with all the SDWAN branch CPE's of all customers.

Appropriate route leaking policies are setup by the provider to ensure that only the WAN IP route of the Gateway is available in the individual customers MPLS L3VPN network. This allow the Gateway to setup a connectivity path to the SDWAN branch CPE's but at the same time ensuring that the CPE's across customers cannot setup any direct communication between them. It is also required that the SDWAN CPE's have unique WAN IP addresses. In the case that the WAN IP addresses are not unique, they need to be source NAT'd.

An Internet WAN link in the Internet-Transport-VR can also be used for this purpose of connecting to the SDWAN branch CPE's.

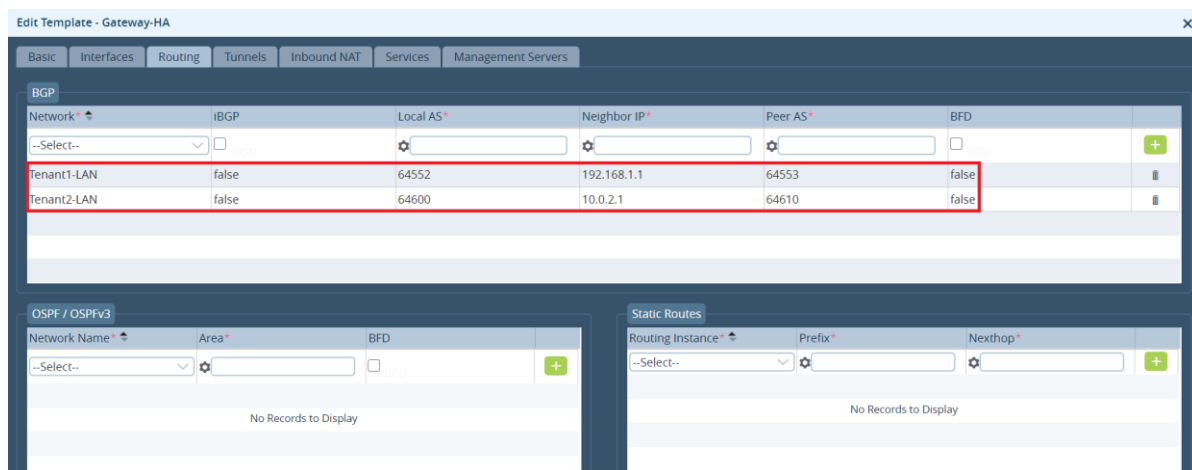


Figure 78 Gateway Configuration through Workflow Templates

Modification of the routing policy to filter sent/received routes can be done in the Device Template or directly under the Appliance. In the configuration, go to Networking and select the MPLS-Transport-VR and select BGP. Then select the BGP instance id and go to Peer/Group Policy.

9.2.2.1 High Availability for option 2

The gateways can be configured from the workflow with or without High Availability (HA). BGP path attributes like AS-path prepend, Local Preference or MED can be used to select which gateway device will serve as the primary gateway.

9.2.2.2 Best Practices for option 2

- When modifying route preference, care should be taken to avoid asymmetric routing.

- If a gateway device is chosen as Primary, then it should be the Primary device for traffic flow in both directions.
- Where the device is a dedicated gateway then there is little purpose in configuring it as HA in workflows because if one of the wan links or devices fail, the gateway becomes non-functional. However, where the device is not a dedicated gateway i.e. it also functions as a regular site, hub or datacenter branch then it makes sense to configure HA using Workflows as the device is usable even when one of the WAN link or devices fails. If the device is a dedicated gateway and it is possible to replicate the wan links and MPLS PE connection on the LAN, then have 2x single gateway devices for redundancy.

9.3 Gateway for Internet bound traffic

A Versa branch can be configured as a gateway to cover the following scenarios.

1. When the ability to apply SDWAN Path selection policies on Internet bound traffic with optional Traffic conditioning features like Forward Error Correction (FEC) or Packet Replication is required. This allows the mitigation of last mile wan link degradation using Versa SDWAN features.
2. When the ability to use MPLS circuit as secondary path to break out from a gateway in case local internet circuit is unavailable is required.

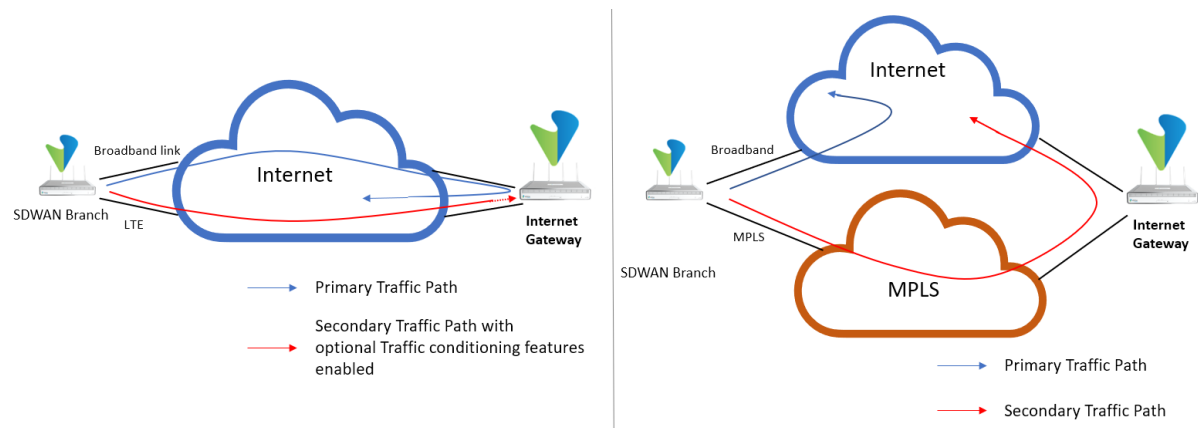


Figure 79 Internet Gateway

Refer Section 9 on how to configure a remote branch as an Internet Gateway.

10 SD-WAN Traffic Optimization

10.1 Traffic-Steering

SD-WAN optimizes traffic over available underlays (MPLS, broadband, LTE) to deliver traffic across the network in an optimal way to provide a better user experience. Versa VOS Edge Devices use SLA monitoring to gather performance metrics on all paths towards all peer branches (assuming for a full mesh topology). The metrics include latency, jitter, and packet loss. They determine whether a path is up, and if so, whether it meets user defined SLAs for a given application or set of applications. Versa SD-WAN can dynamically steer traffic to the best available link, and if the available links show any transmission issues, it will immediately apply remediation for delay, jitter and packet loss based on policies to ensure performance of the high-priority applications. The Versa SD-WAN platform provides powerful and flexible policy-based mechanisms for realizing a wide variety of wan path selection behaviors.

Traffic steering configuration is made up of three major components:

- **Policy rules:** Versa SD-WAN policy consists of one or more rules. A rule is used to identify traffic for which you want to specify path selection behavior. Traffic that does not match a specific rule is subject to default behavior. The traffic matching the rule conditions is attached to a Forwarding Profile.
- **Forwarding Profiles:** Forwarding Profile defines Circuit/Path priorities, connection method and load-balancing capabilities for traffic matching the policy attached to this Forwarding Profile
- **SLA Profiles:** SLA profiles define application/network thresholds which a path must meet for it to be SLA compliant in Forwarding Profile (optional).

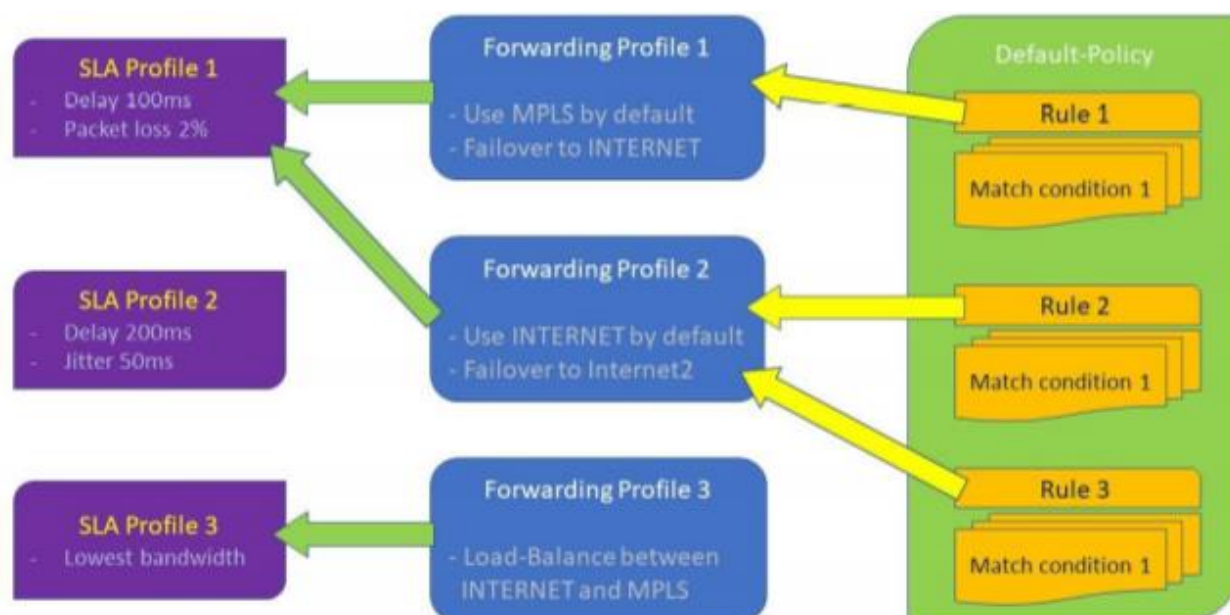


Figure 80 SD-WAN Policy Framework

10.1.1 Best Practices for SD-WAN traffic Steering

There are several best practices related to the SD-WAN traffic steering policy:

- Attention should be given to the match conditions to ensure that the SDWAN Policy matches the correct traffic. The following rule is true for all Policy configuration in VOS Edge Devices:
 - From the SD-WAN Policy Rule perspective, values within the same tab are processed as a logical OR function and values in different GUI tabs are processed as a logical AND function.

For example, where multiple addresses are included in the source address field, any one of the addresses can fulfill the match criteria for that field. If multiple source addresses are included, and a source zone is configured, then the traffic must match one of the source addresses AND one of the source zone parameters in order to result in a match. Both IPv4 and IPv6 addresses are supported.

Note: this rule classification evaluation is not specific to SD-WAN Policy Rules, but applicable for all policy rules in the Versa platform.

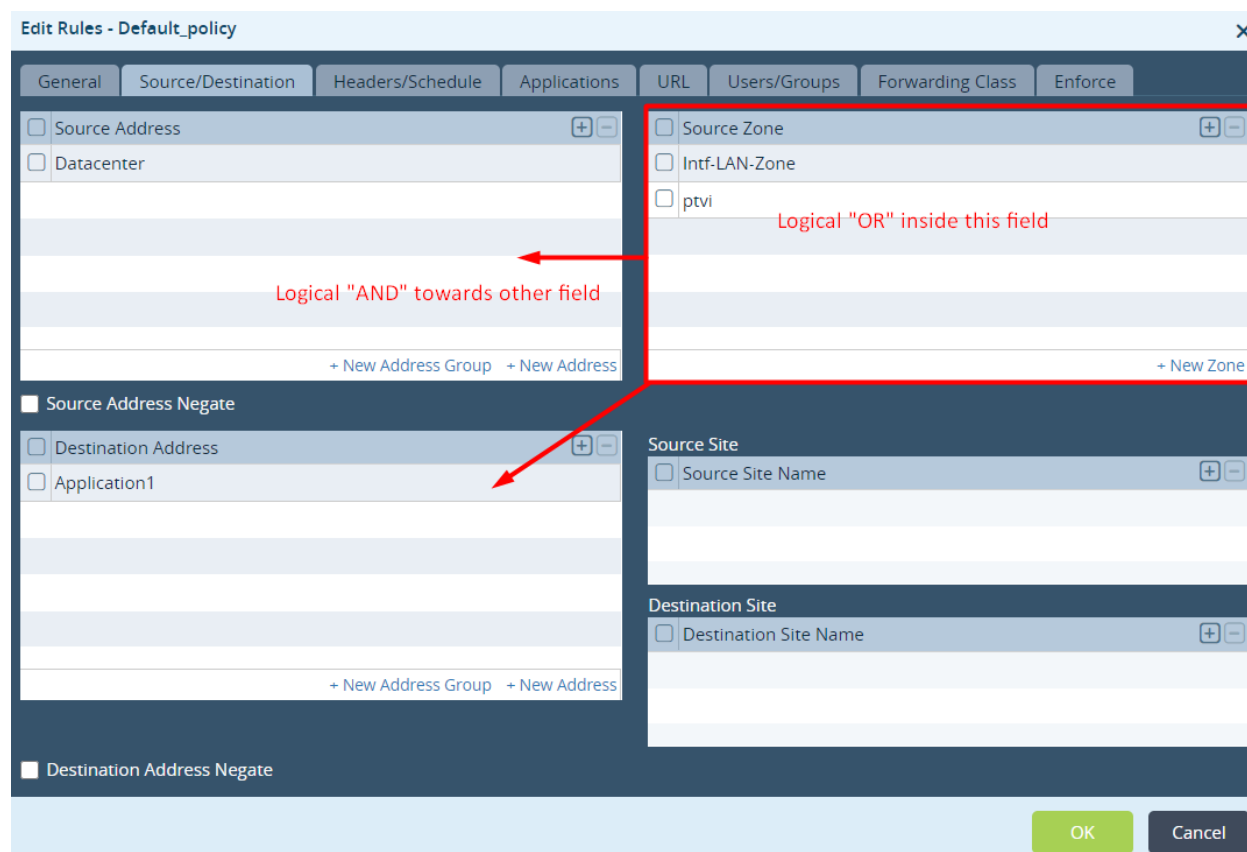


Figure 81 SD-WAN Policy Rule evaluation

- SD-WAN policies are rule order based: Have the matching criteria for the critical traffic in one of the top rules. SD-WAN rules are processed from top to bottom and the rule evaluation is exited at the first match.
- Match between tabs is an “AND” condition.
- SD-WAN Policies should be configured to classify applications with similar behavior (or characteristics) in order to apply the relevant forwarding profile.

- As a rule of thumb log only on “SLA Violated”. Limiting the rate of the log messages is also an interesting feature which could be used.
- Always set the bandwidth statement on the interface. This is used as the main input for the SD-WAN bandwidth related profiles.

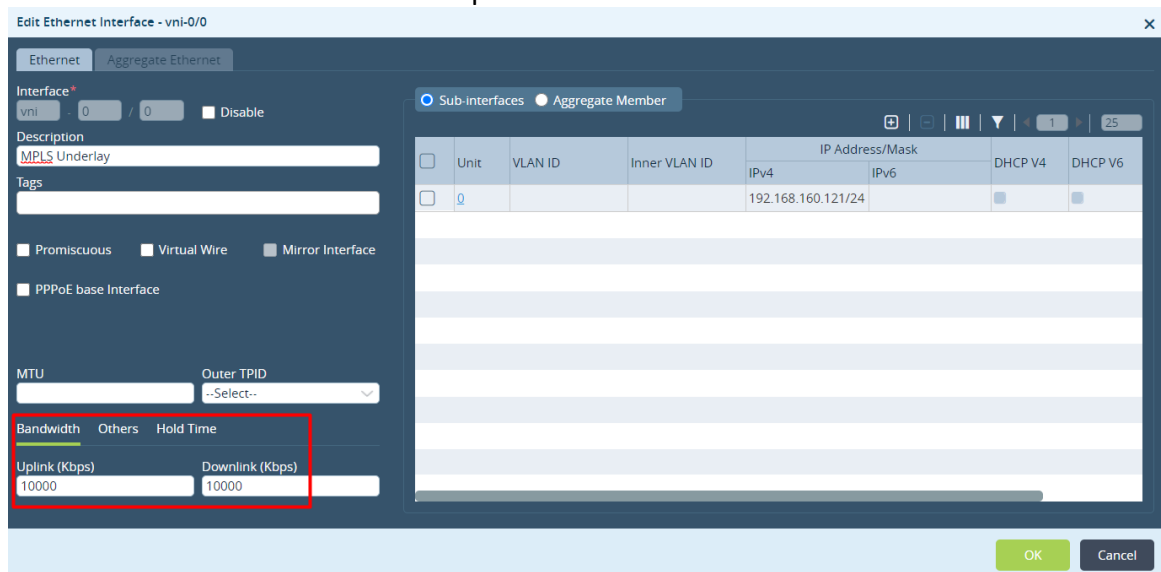


Figure 82 Bandwidth capacity configuration

- Set WAN Circuit priority to the same to enable load balancing between paths across the WAN circuits:
 - At least two circuits with equal priority in case load balancing is desired for a specific traffic.
 - When no circuits are specified, it is assumed all circuits have the same priority.
- Weighted-round-robin Connection Selection Method: Weights are assigned to each path based on the available bandwidth on the corresponding access circuit. This connection selection algorithm is recommended to be used in case the paths with equal priorities are different from provisioned bandwidth perspective.
- An elephant flow is a single large volume session such as FTP downloads or file copy. These present peculiar challenges for SD-WAN traffic steering; the following are best practices to optimize these flows:
 - Use Per Packet Load balancing where the traffic is no Jitter or loss sensitive.
 - Enable Gradual Migration, to prevent thunder heading of flows when a previously violated path becomes good again.
 - Set up relatively longer recompute timer, to prevent violating the path due to transient impairments.
 - Enable FEC to protect against loss especially where data integrity is important such as file copy. Set “Stop when” Circuit utilization hits a high watermark value to prevent FEC overhead from congesting the network.
- Reordering should be enabled globally on all branches in the network.
- Realtime Traffic such as voice and video UDP streams should be configured with the following best practices:
 - Enable “Evaluate Continuously” to allow Realtime flows to react better to changing network conditions.
 - Enable “Gradual Migration” (under Advanced Settings Tab in the Forwarding Profile).

- Enable Replication to mitigate against loss and jitter. Start when “SLA Violated” and Set “Stop when” circuit utilization hits a high watermark value (75% is recommended) to prevent replication overhead from congesting the network.
- Recompute Timer determines the number of SLA reports that should be accessed before a violation decision is made. It is not recommended to shorten the default recompute timer; instead, evaluate continuous should be enabled. However, a longer recompute timer could be set for relatively quality underlay networks, as well as for flows that can survive transient network conditions. The default of 300sec is recommended, but it might be convenient during lab testing to shorten this timer.
- SLA smoothing and SLA violation damping are disabled by default and should be enabled following an in-depth analysis. In case if the circuits/paths are not moving very often between complaint/non-complaint state the default config (disable) should be kept.
- If SLA smoothing is enabled, the Recompute timer should be also configured to a value lesser than the SLA smoothing interval.

10.2 Traffic Conditioning

Real-time applications have more stringent delay requirements than normal data transmissions. As a result, retransmission of the lost packets is generally not a valid option for such applications. In these cases, a better method to attempt recovery of information from packet loss is through Packet Replication and Forward Error Correction (FEC).

10.2.1 Packet Replication

Packet Replication is the most impactful option to choose for applications that are sensitive to latency and packet loss. Primarily used for real-time applications, packet replication takes an identified type of traffic and sends a copy of each packet across multiple paths. This eliminates the inherent latency anomalies and packet losses from negatively impacting these applications.

Important Note !

It must be understood that packet replication can have significant impact to the bandwidth utilization of the WAN circuits. Therefore, use this feature mainly for low volume UDP traffic which is sensitive to packet loss (such as VoIP traffic). Also, it is recommended to set both the “Start When SLA Violated” and “Stop when” knobs.

If one packet is lost in transit the copy of the packet is forwarded to the receiver. The receiving branch discard copies of the packet and forward only one packet to the receiver.

Reordering MUST be enabled on all branches to ensure that the receiving branch reorders the 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 with a single access links. This method works well when the amount of critical traffic that is being duplicated across the networks is far less than the capacity of the network.

The configuration path is:

Configuration ---> Templates (Device Templates/Devices) ---> <Name> ---> Services ---> SDWAN ---> Forwarding Profiles ---> General

Figure 83 Forwarding Profile Configuration

Available Options:

- **Enable:** Select the checkbox to enable packet replication. By default, replication is disabled.
- **Replication Factor:** Defines on how many paths the replication will be done. The default value is 2. If there are more than 2 paths between branches this parameter can be tweaked to a bigger value.
- **Start When:** if this parameter is not selected the forwarding profile will always replicate traffic for the defined rule. It can be specified “SLA Violation” option here, to enable packet replication only when all available paths have violated SLA. It is recommended to enable packet replication only when SLA Violation is true.
- **Stop When:** optional parameter used to prevent oversubscription of the links. When the utilization of any of the links over which we replicate data reaches “Circuit Utilization” percent, the replication will stop. And the flow will be passed only using one available link as it was before replication enabled.
- **Circuit Utilization:** if all WAN circuits of the device got utilized for more than value %, the replication stops. This applies for all circuits, even for those which are set to be avoided.

For additional details on how to configure SDWAN Packet Replication you can check the following link:

https://docs.versa-networks.com/Versa_Operating_System/Versa_Operating_System_SD-WAN_Configuration/Advanced_SD-WAN_Configuration/Configure_Replication_for_SD-WAN_Traffic_Steering

10.2.2 Best Practices Packet Replication

- Use packet replication for specific traffic only, such as VoIP.
 - Avoid enabling replication on wildcard traffic.
- Enable replication only on “SLA Violated” and always set the Stop when Circuit utilisation.²
- A replication factor of 2 is adequate for most cases. However, this can be increased to match the available paths for the flows being protected. E.g. 4 paths can benefit from a replication factor of 3
- Bandwidth utilisation increases by a factor of 100% for each unit increment in the replication factor. i.e. replication factor of 2 equals 100% increase in bandwidth, factor of 3 equals 200% increase etc.
- Enable replication for Realtime traffic and other business critical traffic where jitter and loss are important.

10.2.3 Forward Error Correction

Forward Error Correction (FEC) is a technology that is well known for its ability to correct bit errors at the physical layer. While FEC was traditionally used for this purpose, it has since been adapted to recover from packet loss at the network level. Packet-level FEC works by adding an additional loss recovery packet for every “N” packet that are sent. This additional loss recovery packet enables Versa 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 avoids transport-layer retransmissions and, in case of TCP, prevents TCP’s congestion avoidance mechanism from stepping in and 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. FEC can be turned on alongside with replication, at the sites having multiple paths, to provide maximum protection and correction. Or FEC can be used to recover packets independently where replication might not be useful, like at the sites having single path available for transporting traffic.

² Ensure that bandwidth values are set correctly on the interface for accurate circuit utilisation calculations.

The configuration path is:

Configuration ---> Templates (Device Templates/Devices) ---> <Name> ---> Services ---> SDWAN ---> Forwarding Profiles ---> FEC

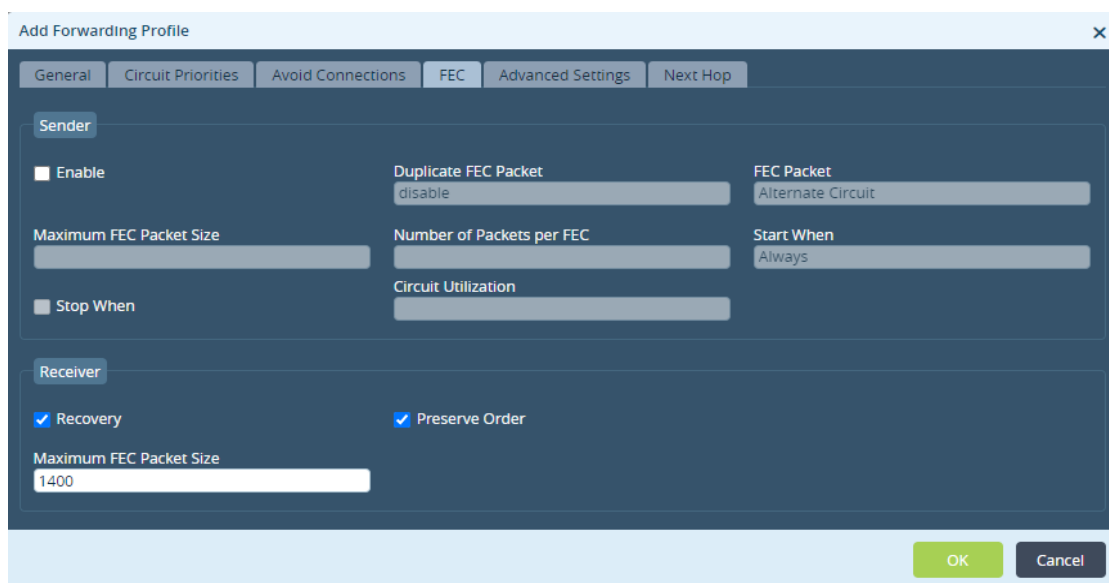


Figure 84 FEC configuration in Forwarding Profile

For FEC implementation is very important configure both sites: sender and receiver.

Available Options for Sender selection:

- Mode: set this value to enable for FEC to begin. (checkbox “Enable” in the GUI)
- Duplicate FEC Packet: 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.
- FEC Packet: Same options as the previous ones. The FEC packets could use the same WAN circuit or an alternate one.
- Maximum FEC Packet Size: specify the maximum size of the data in the packet to be protected. For example, voice packets are typically <100bytes therefore set this value to 100bytes to protect that length of packets. Packets larger than this are not protected.
- Number of Packets per FEC: this option controls after how many packets FEC parity data will be calculated. Default Value is 4, which means the system will generate FEC parity data after each 4 packets of the actual data being transmitted and affected by this Forwarding Profile.
- Start When: if this parameter is not selected the forwarding profile will always protect traffic for the defined rule. “SLA Violation” option can be specified here, to enable FEC only when all available paths have violated SLA.
- Stop When: optional parameter used to prevent oversubscription of the links. When the utilization of the link over which we send FEC data reaches “Circuit Utilization” percent, the FEC operations will stop.

Available Options for Receiver selection:

- Recovery: checkbox to enable packet recovery after receiving FEC packets. By default, receiver packet recovery is enabled.

- **Preserve Order:** This option is independent from “Reorder” option in the Forwarding Profile. Enabled by default. The system will place FEC packets into the original order when it receives them.

For additional details on how to configure SDWAN Forward Error Correction you can check the following link:

https://docs.versa-networks.com/Versa_Operating_System/Versa_Operating_System_SD-WAN_Configuration/Advanced_SD-WAN_Configuration/Configure_Forward_Error_Correction_for_SD-WAN_Traffic_Steering

10.2.4 Best Practices FEC

There are multiple ways of doing Layer 3 FEC, with pros and cons of each. So, there is no “one size fits all” solution. Below are some recommendations regarding in which situations Versa FEC feature should be used.

- **WAN circuits utilization.** When bandwidth is at a premium FEC is more efficient than packet replication. Only one parity packet is generated for a specified number of data carrying packets (range 1-32, default 4). The generated FEC parity packet can recover a packet on the peer branch only if there is one lost packet in the specified number of packets per FEC.
- **In case the packet corruption on a path has a specific pattern** (corrupting also the nth recovery packet and its replica) it is recommended to replicate the FEC packet on an alternate circuit. In this way at least a non-corrupted FEC packet will arrive at the remote site.
- **Small vs. large packets.** Layer 3 FEC works well on flows that use small packets, such as voice or a point of sale transaction. Large packet applications such as video and file transfers usually are not business critical applications and doing forward error correction on fragmented packets is difficult. The recommendation is to use FEC for packets less than 500 bytes in size.
- **LTE and FEC.** In most cases FEC should not be used on LTE interfaces. FEC increases the amount of traffic on a network that is already limited in capacity and it makes the situation worse. When packets are dropped in a wireless network, it is usually many packets in a row and for this kind of situation FEC does not help.
- **Adaptive Codecs vs 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. In this case tests should be performed to find out if the Layer 3 FEC feature provides any advantage if it is used 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 will arrive intact at a receiver. Denoted by the formula:

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

Where:

$P(FEC)$ = FEC recovery probability (%)

p_x = Packet loss (%)

w = number of packets per parity (npp)

Therefore, the FEC default settings of 4 (npp) FEC is only effective up to ~5% packet loss. Replication should be considered for higher loss percentages.

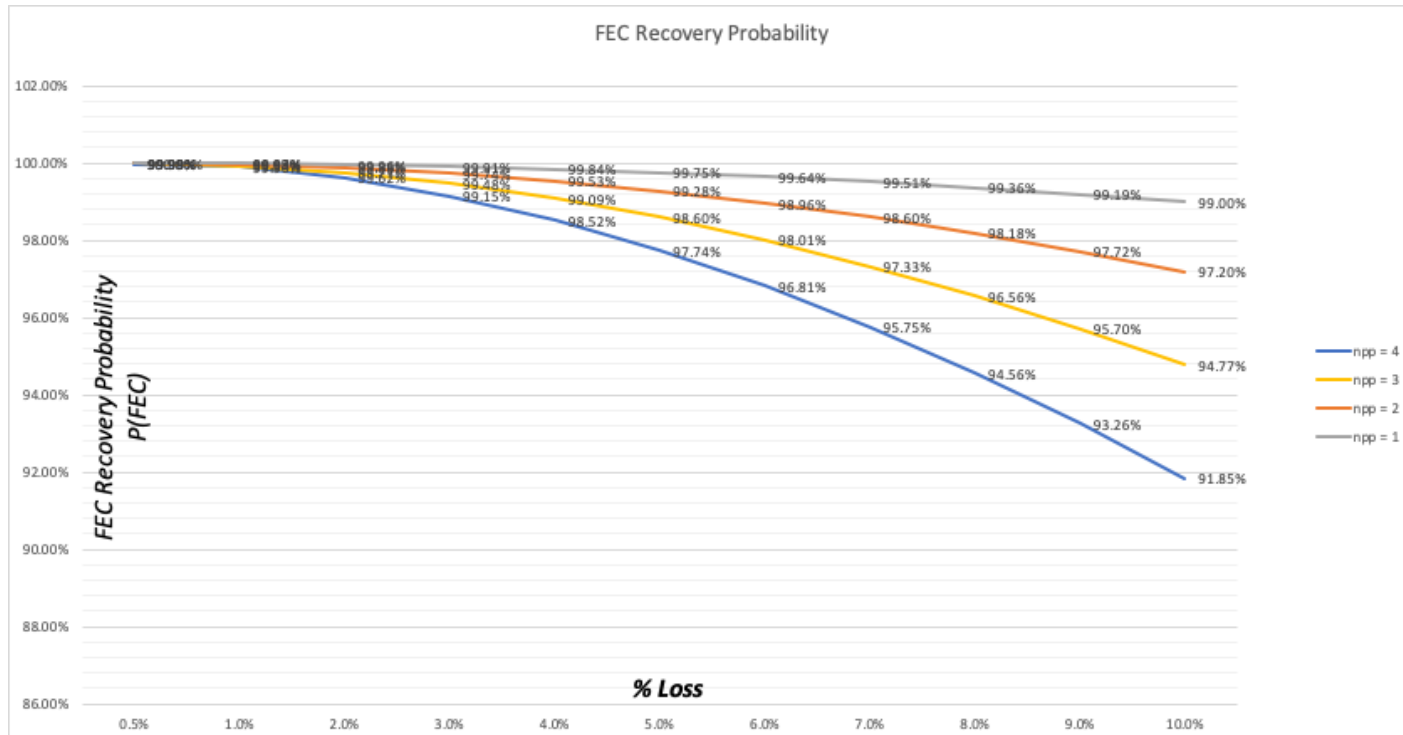


Figure 85 FEC Recovery Probability

- FEC performance increase the closer npp is to 1. However, there is a 25% increase in bandwidth utilization value for every unit decrease in $x - 1$ where $x = npp$. For example, the default npp is 4 reducing this to 3 has a 25% increase in the bandwidth utilization. When npp = 1 then there is a 100% increase in the bandwidth utilization.

10.3 Business Intent traffic steering with Application-Steering-Templates

Application Steering Templates automate Business Intent policies. It combines all business application and desired network characteristics into one configuration-template configuration. This includes classification, mapping into forwarding class (QoS) and attach SD-WAN policies, all in one template.

Application Steering Templates do not add new SD-WAN functionality, but it simplifies management of business applications in one template.

Application Steering Templates do not configure other QoS parameters than Forwarding Class assignment and Loss priority hence users will still have to define and apply QoS policies and rules to deliver predictable applications performance during congestion.

After the initial deployment phase is completed SD-WAN application templates can then be modified for detailed configuration. The Application Steering workflow is retained and used in similar ways as other workflow templates to create complex configuration modification.

A default application template per organization with 4 Traffic Categories is preconfigured, Real Time, Business Critical, Default and Low Priority which groups applications or family of applications under a particular Traffic Category (i.e. Voice Applications under Real Time, Office365-Apps under Business Critical and so on).

Note that the default application template is an example template to allow the administrator a quick start in writing application business intents. That means that the defaults should be evaluated and modified accordingly.

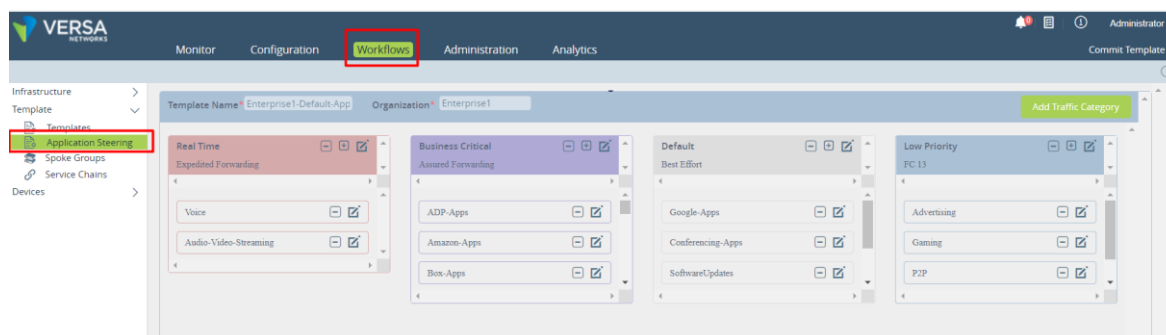


Figure 86 Application Steering Workflow Template

Default configuration like Forwarding Class and Circuit Selection Criteria are applied to each class, for instance Real Time Traffic Category:

The screenshot displays the 'Edit Traffic Category' interface. It includes a 'Name*' field with the value 'Real_Time'. The 'Forwarding Class*' dropdown is set to 'Forwarding Class 4 (Expedited-Forwarding)'. The 'Loss Priority' dropdown is set to 'Low'. The 'Traffic Conditioning' section has checkboxes for 'FEC' and 'Replication', and a 'Load Balance' dropdown set to '---Please Select---'. The 'Circuit Selection Criteria' section has checkboxes for 'Low Latency' (checked), 'Low Packet Loss', and 'Low Delay Variation' (checked). Below this is a table with one row: 'WAN Circuit' with a 'Priority' checkbox and a '+ -' icon. The text 'No Row Added' is centered below the table.

Figure 87 Traffic Category

Users can modify the Forwarding-Class and Loss Priority, set Traffic conditioning parameters (whether FEC and/or Replication), load balancing algorithm and dynamic or static Circuit path Selection Criteria that applies to the Applications configured under the Traffic Category.

Users can create Custom Traffic Categories too as well as modify predefined Categories.

10.3.1 Best Practice Application-Steering-Templates

- Use Application Steering Templates to automate the creation of SDWAN policies. Once service template is created, it should be customised to suite need.
- Application Steering templates only create CoS classifiers, ensure that schedulers are added to the appropriate interfaces.

10.4 SD-WAN Path Policies

Versa branches continuously monitor the performance of all paths towards all SD-WAN peer branches as well as towards controllers. A path is defined as any valid transport tunnel between the two branches.

For example, if two branches have two broadband links each, and all of them are in a single transport domain, there are four paths between the branches. This applies to path between branch and controllers too but have been left out of the drawing below for clarity sake:

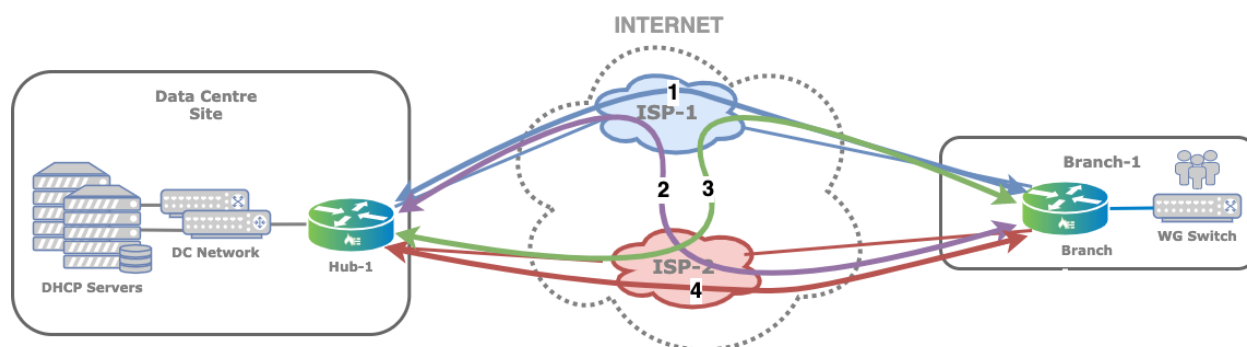


Figure 88 Possible Overlay Paths

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

SLA Monitoring (SLAM) may be configured for all 16 forwarding classes (network control through fc 15) using SLA Path Policies.

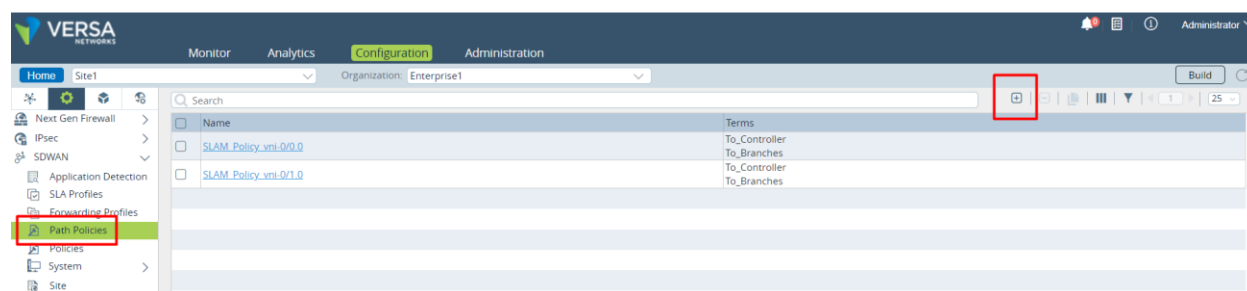


Figure 89 SD-WAN Path Policies

Above SLAM Path Policies are created by default on all WAN interfaces using EF Forwarding-Class at 2s interval towards remote Branches and 10s towards Controllers using NC Forwarding-Class. These policies can be modified, or Administrators can create new policies to suit their needs.

In a nutshell each SLAM Path Policy has a match condition on the remote site type whether it is a controller or a branch:

Configuration → Services → SDWAN → Path Policies → Edit Path Policy → Edit Terms

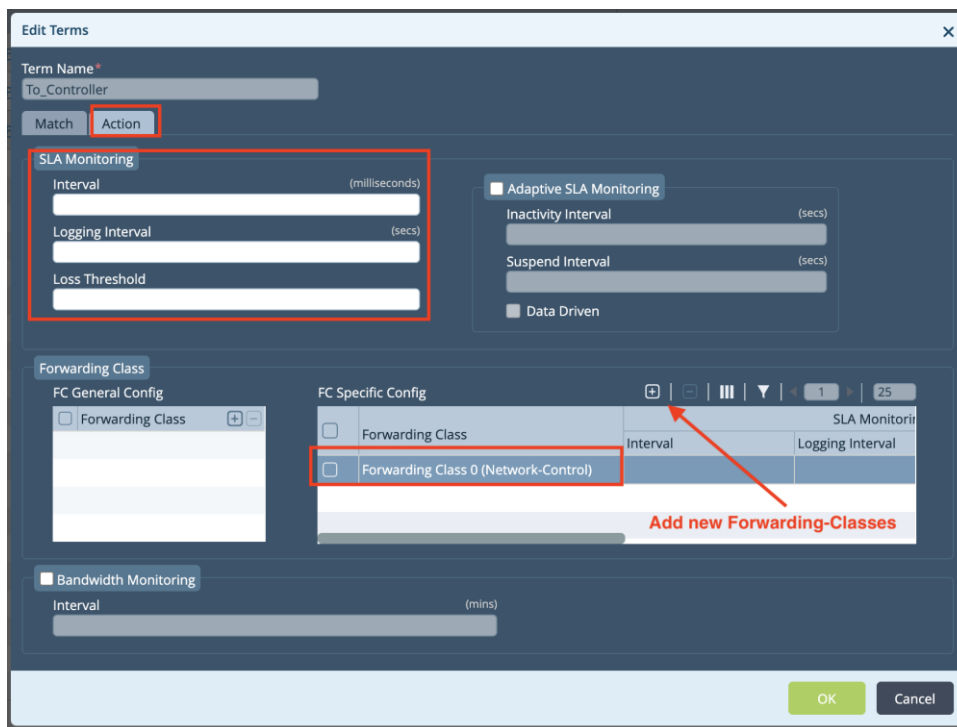


Figure 90 configuring path policies action details

For additional details on how to configure SLAM Path Policies you can check the following link:

https://docs.versa-networks.com/Versa_Operating_System/VOS_SD-WAN_Configuration/Advanced_SD-WAN_Configuration/Configure_SD-WAN_Path_Policies

10.4.1 Best Practices Path Policies

- The default SLA Path policy is generally enough for most use cases.
- If it is necessary to add additional SLA probes, then it is recommended to have no more than one SLA probe per Traffic class.
- Where granularity is required then four probes can be set for the four traffic classes - Network Control, Expedited Forwarding, Assured Forwarding and Best-Effort.
- This modification is only required for Branch to Branch Profile. Branch to Controller Profiles only require probes in the Network Control class.
- SLA Monitoring probes do come with a cost. On low bandwidth circuits configuring SLA Monitoring for several Forwarding-Classes with aggressive timer could use a considerable part of the available bandwidth. Hub-Spoke topology is recommended to minimize this overhead due to only having SLA probe to the hub.
- SLA optimization techniques such as Adaptive SLA and Data driven SLA can be enabled to reduce the SLA probe load particularly on low bandwidth links.
- Topology changes can be introduced to further optimize the network to prevent the full mesh of SLA and subsequently the SLA load. Please refer to [Chapter 7](#) SD-WAN Topologies for details.

11 Routing

The VOS Edge Device supports a range of routing protocols. This section will describe some of the common use cases about how these are used.

11.1 Static Routing

Versa Static routing is enriched with the following additional features;

- Next hop interface or IP.
- Attach Monitor Object.
- ICMP based monitors
- Change metric
- Change protocol preference
- Enable BFD
- No-install

11.1.1 Floating Static Route

Static routes can be conditionally withdrawn based on the state of another target IP address.

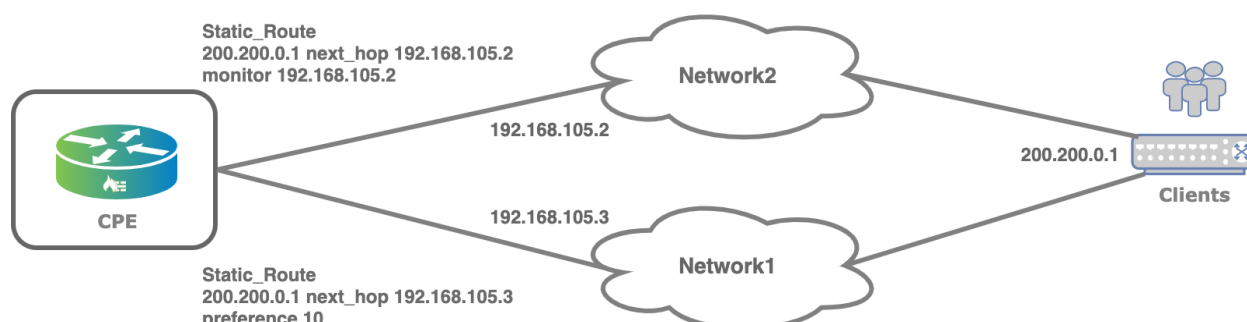


Figure 91 Floating Static Route

In the scenario below the next-hop IP address is monitored on the primary link, while the secondary static route is configured with a lower preference.

```
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
```

Output above shows the failure of the monitor probe caused the primary static route to be withdrawn. Monitor objects could be TCP, DNS or ICMP based.

11.1.2 Route Leaking with Static routes

Static routes next hop can be resolved to the same routing instance or to a different routing instance. Consider the scenario below;

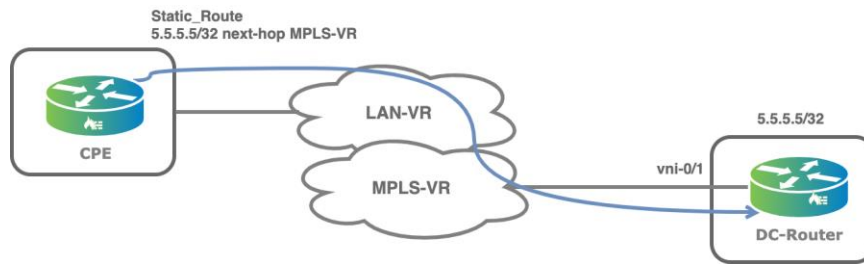


Figure 92 Route Leaking

Traffic is sourced from the LAN and the next hop is resolved to the transport VR.

Configuration stanza

```
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;
    }
  }
}
```

Routing Table output

```
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 output below captures with TCPdump the ping traffic at the DC-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
```


11.2 Dynamic Routing

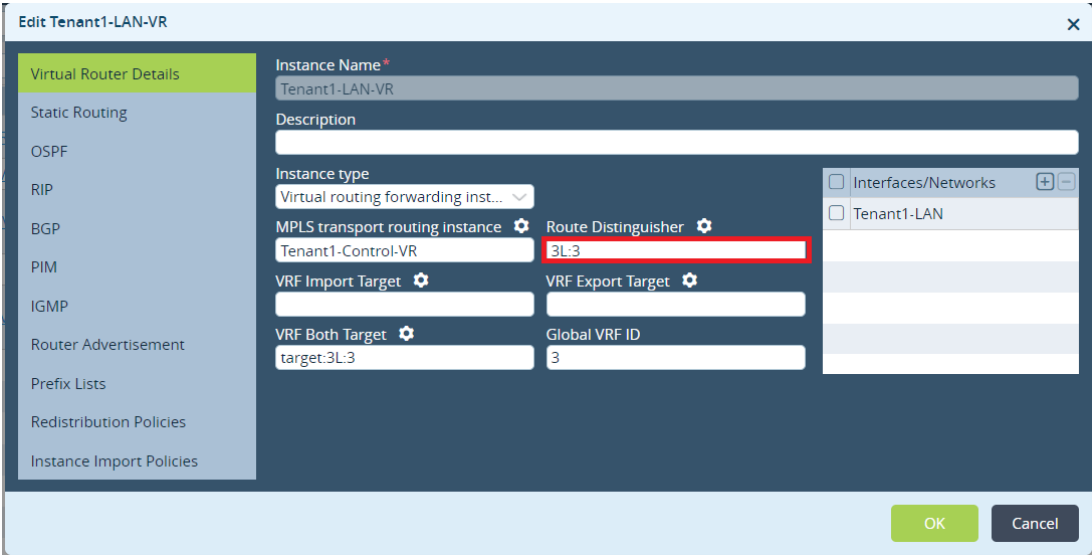
This chapter describes best practices for VOS Edge Device's dynamic routing for fast-convergence, scalability, loop prevention and security perspective.

11.2.1 Fast Convergence – SD-WAN Side

For dual-homed deployments a unique Route-Distinguisher should be configured per CPE LAN-VRF. Having unique Route Distinguisher per LAN-VRF will allow the dual-homed CPEs to advertise unique information for the same local prefix. Therefore, the Controllers (RRs) BGP best path computation would result in reflecting prefixes coming from both CPEs part of the dual-homed architecture. This improves convergence as it would not be necessary to resend a BGP update in a failure case. It also helps in troubleshooting as the distinctive RD would be much easier to map to the prefix origin (CPE). The workflows will typically configure the same RD for all nodes. Note that the workflow automatically set unique RD's when configuring an HA-pair.

Unique Routes Distinguishers are configured in:

Networking -> Virtual Router-> tenant-LAN-VR



The screenshot shows the 'Edit Tenant1-LAN-VR' configuration window. The left sidebar contains a menu with options: Virtual Router Details (selected), Static Routing, OSPF, RIP, BGP, PIM, IGMP, Router Advertisement, Prefix Lists, Redistribution Policies, and Instance Import Policies. The main configuration area includes the following fields:

- Instance Name*: Tenant1-LAN-VR
- Description: (empty)
- Instance type: Virtual routing forwarding inst...
- MPLS transport routing instance: Tenant1-Control-VR
- Route Distinguisher: 3L:3 (highlighted with a red box)
- VRF Import Target: (empty)
- VRF Export Target: (empty)
- VRF Both Target: target:3L:3
- Global VRF ID: 3

On the right side, there are two checkboxes: 'Interfaces/Networks' and 'Tenant1-LAN', both of which are unchecked. At the bottom right, there are 'OK' and 'Cancel' buttons.

Figure 93 RD configuration in LAN-VR

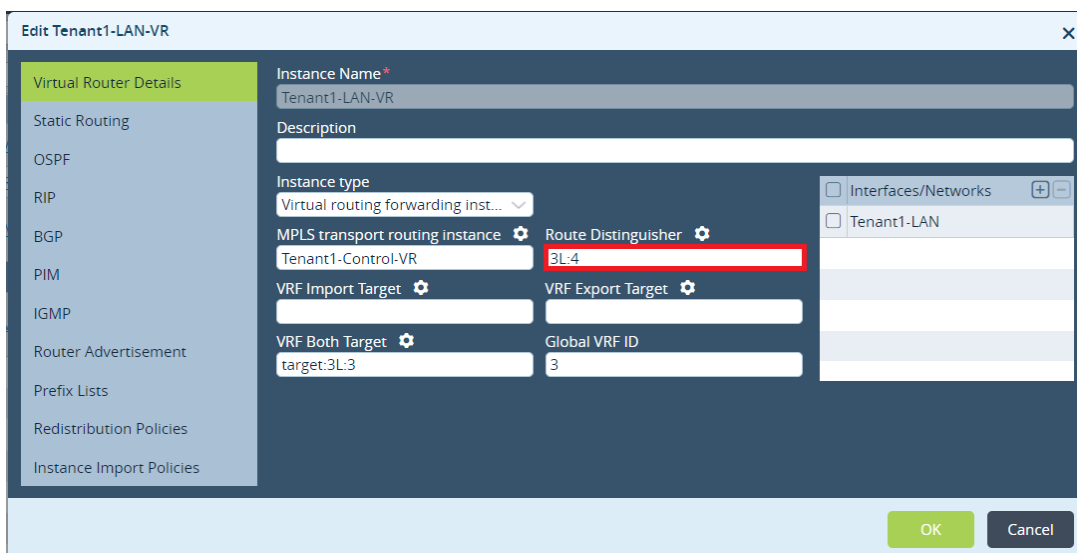


Figure 94 RD configuration from Branch1 and Branch1'

Controllers are reflecting the information coming from both CPEs, and as a result on the other sites two paths would be available and installed in the RIB.

```

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 ]

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
    
```

Other techniques such as BFD and adjusting GR timers are not desired in the SD-WAN overlay as it interferes with the headless operation of the Versa solution.

11.2.2 Fast Convergence - LAN Side

An important aspect of fast convergence is rapid failure detection. If the detection is not based on a link status failure, such as where neighbors connected via a Layer 2 network, then the alternate is a protocols-based timer-related detection. BFD is a high-speed protocol designed for fast failure detection as it can detect link failures in milliseconds. IGP/BGP are BFD clients and multiple routing protocol can piggyback a single BFD session. The recommendation is to enable BFD on the Versa CPE and establish BFD sessions with the customer owned L3 devices. VOS Edge Device supports BFD for RIP, OSPF, BGP and static.

In the below example it can be seen a BFD session established between two OSPF neighbors:

```
admin@branch8-cli> show ospf neighbor brief
State codes: atmtpt - 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
```

11.3 Scalability

VOS Edge Devices support a fully-fledged scalable and powerful routing suite. Below it will be presented some best practice configurations when a Versa CPE becomes part of an existing network.

11.3.1 OSPF

For Ethernet segments where there are only two OSPF routers, configure the network as OSPF point-to-point. This will prevent DR, BDR election on the respective ethernet segment and generation and flooding of an LSA type 2 representing the segment. As a result, CPU cycles will be saved.

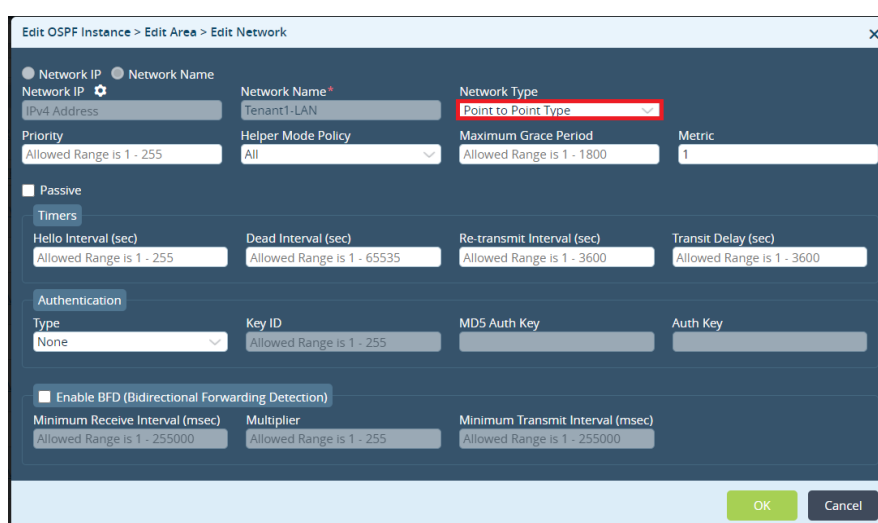


Figure 95 OSPF Point-to-Point configuration

11.3.2 BGP

BGP Peer Groups provide a mechanism for BGP peers that have the same outbound policies to be associated together. This feature has two major benefits: configuration reduction and the ability to replicate updates between peers.

A common reason for deploying peer groups is configuration reduction. The peer group is formed, and the common outbound policy is applied to the peer group. Each peer that requires the same outbound policy is assigned to the peer group, having as a result much less redundant configuration on routers with many BGP peers.

However, the key benefit of peer groups is the ability to replicate updates across peers. This is because for all peers the same outbound policy is used the BGP update messages send to them are the same. As a result, a BGP scalability enhancement can be derived: BGP update messages are generated once for each peer group and then are reused for all the peers' part of it. The best practice is to configure the BGP peers having the same outbound policy to be part of the same BGP peer group. Versa SDWAN controllers implement this best practice by default having all route reflector client's (CPE branches) part of the same BGP peer group.

In the below command output it can be seen also the outbound BGP policy which is applied for the peer group and all the group members:

```
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
```

11.4 Loop Prevention

In many cases routing loops appear because of mutual redistribution between two protocols at more than one point. As a best practice for dual homed topologies where mutual redistribution is performed between BGP and OSPF, Domain VPN Tag should be configured and Downward option to be enabled in the respective OSPF instance.

Edit OSPF Instance

Instance ID* 3014 Router ID* 6.6.6.6 Domain VPN Tag 2

Internal Admin Distance 31 External Admin Distance 111 Reference Bandwidth (Mbps)

Enable Alarms Disable DN Bit

Areas

Area ID	Type	Networks	Virtual Links
<input type="checkbox"/> 0.0.0.0	backbone	Tenant1-LAN	

OK Cancel

Figure 96 OSPF Loop Prevention

Let's have a short explanation of Down Bit and Domain Tag. The interaction between OSPF and BGP when these two protocols are used in an overlay network – MPLS (similarly for SD-WAN) is described in RFC 4577 - OSPF as the Provider/Customer Edge Protocol for BGP/MPLS IP Virtual Private Networks (VPNs). As a summary the Down Bit is a bit that is set in the Options field of an OSPF LSA type 3. It indicates the direction that the route has been advertised. If the OSPF route has been advertised from a PE (VOS Edge Device) router into an OSPF area, the Down Bit is set. Another PE (VOS Edge Device) router in the same area does not redistribute this route back in to the iBGP overlay network if this bit is set. The PE router does not even include the route in the SPF calculation. As such a possible routing loop is avoided if the site is multihomed. The Domain Tag serve the same purpose as the Down Bit, but for external routes (LSA type 5).

The below topology is used to explain how to check if DN bit and Domain Tag are set and possible issues in case if the Versa CPEs are not directly connected to the local LAN.

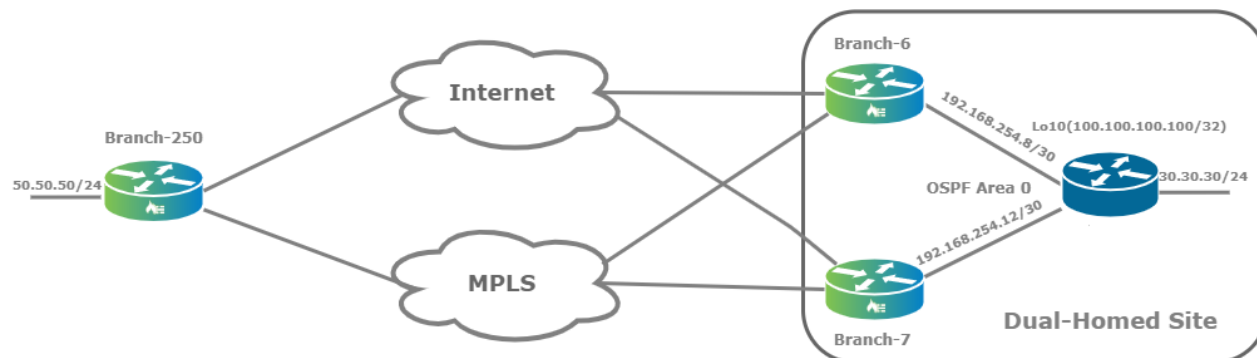


Figure 97 Loop Prevention Topology Diagram

On the single CPE site (Branch-250) the local LAN (50.50.50/24) is redistributed in MP-BGP.

On the dual-homed HA Active-Active site OSPF is running between Versa CPE devices (Branch-6 and Branch-7) and a local router.

To prevent routing loops both OSPF Down Bit and OSPF Domain VPN Tag are when MP-BGP to OSPF redistribution is performed. This can be seen for 50.50.50/24 prefix (coming from Branch-250) when it is redistributed from MP-BGP to OSPF on 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
```

These features are very helpful for routing loops prevention, but a problem arises when the Versa CPEs are not directly connected to the local LAN and there is another router providing Tenant based LAN connectivity and having OSPF neighborship with the Versa CPEs – per Tenant.

As you can see below a third router 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 of the Down Bit set in the LSA type 5 for 50.50.50/24 and the existing router is already VRF aware (acting as a PE) the LSA is not even taken in consideration for SPF calculation and as a result the prefix is not inserted in its RIB. This finally breaks the connectivity for the dual-homed site to the 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
```

But 50.50.50/24 is not present in the RIB:

```
IOU1#show ip route vrf Tenant1 50.50.50.0

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

The recommended solution is to check if the intermediary router (IOU1) has the capability to ignore the Down Bit set in the LSA option field and use the LSA in the SPF algorithm and finally insert the prefix in the RIB. Most vendors call this feature VRF-lite capability.

If the existing customer router does not support this capability the Down Bit setting should be disabled on the Versa appliances. In this situation special care should be taken when mutual redistribution between MP-BGP and OSPF is performed.

11.5 Routing Security

11.5.1 IGP Security

Authentication is the most important measure which can be taken to secure any routing protocol. In VOS both OSPF and RIP support authentication (clear text and HMAC-MD5 based), which simply is a mechanism by which two neighbors prove their identity to each other by using a shared secret. No protocol message is accepted from a neighbor unless the message is correctly authenticated. Authentication is also useful in certain circumstances like prevent routers to mistakenly join an OSPF domain.

Best practice is to configure IGP HMAC-MD5 authentication on all interfaces on which the IGP runs. Below is a screen to configure IGP authentication for OSPF and RIP respectively.

The screenshot shows the 'Edit OSPF Instance > Edit Area > Edit Network' configuration window. The 'Authentication' section is highlighted with a red box, showing the 'Type' dropdown set to 'None'. Other fields in this section include 'Key ID' (Allowed Range is 1 - 255), 'MD5 Auth Key', and 'Auth Key'. The 'Timers' section shows 'Hello Interval (sec)' (Allowed Range is 1 - 255), 'Dead Interval (sec)' (Allowed Range is 1 - 65535), 'Re-transmit Interval (sec)' (Allowed Range is 1 - 3600), and 'Transit Delay (sec)' (Allowed Range is 1 - 3600). The 'Enable BFD (Bidirectional Forwarding Detection)' section is also visible, with 'Minimum Receive Interval (msec)' (Allowed Range is 1 - 255000), 'Multiplier' (Allowed Range is 1 - 255), and 'Minimum Transmit Interval (msec)' (Allowed Range is 1 - 255000). The 'Network Name' is 'Tenant1-LAN' and 'Network Type' is 'Point to Point Type'. The 'Metric' is set to '1'.

Figure 98 Adding authentication for the OSPF protocol

11.5.2 BGP Security

VOS Edge Devices implements the BGP Dynamic Neighbors feature in case where many BGP peers need to be configured. This feature allows dynamic BGP peering to a group of remote neighbors that are defined by a range of IP addresses. Each range can be configured as a subnet IP address. The feature is configured using BGP peer groups and does not require manual neighbors' instantiation, just the subnet where the peers reside. As the peers are not configured individually measures should be put in place to prevent security issues. Together with BGP dynamic neighbors feature it is recommended to configure BGP authentication and TTL for the BGP session. This will prevent a rogue device to establish a BGP session by just using the right subnet (the one configured for the Dynamic Neighbors feature). This configuration is implemented as shown in the figure below:

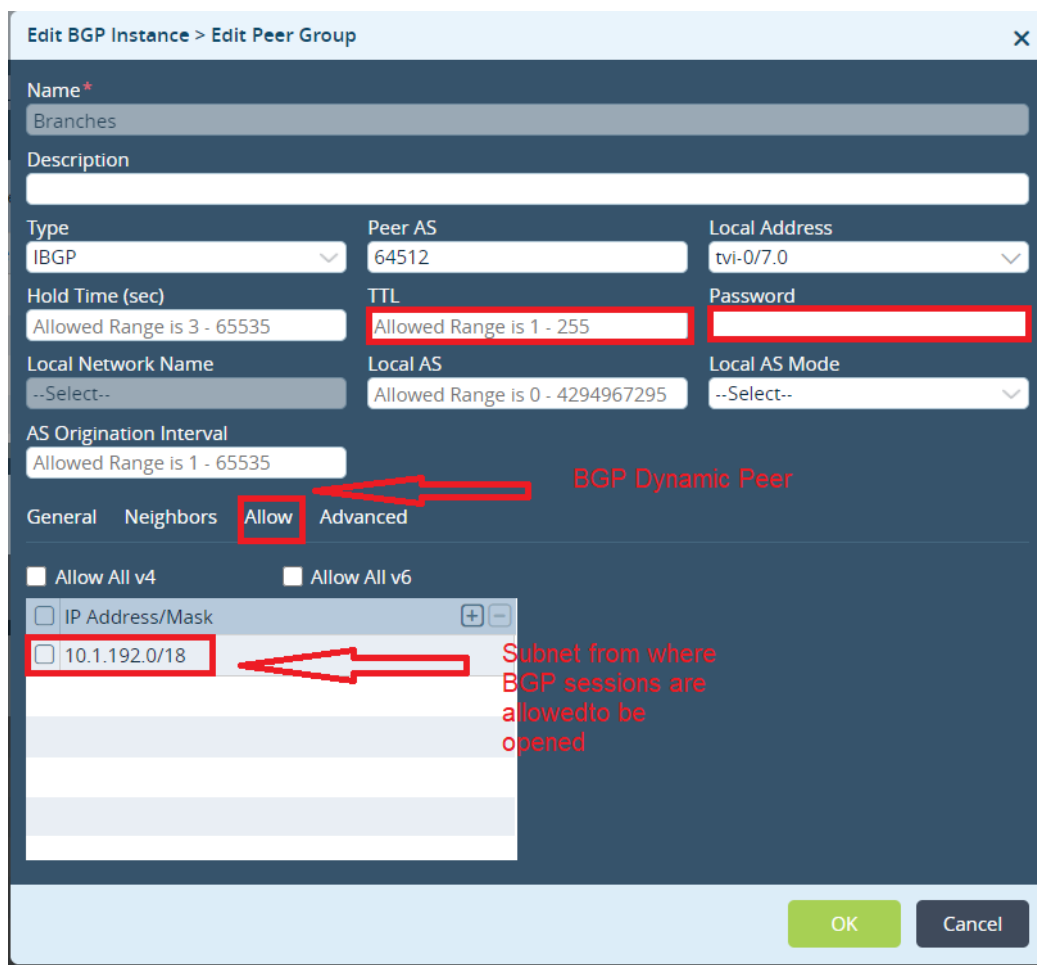


Figure 99 BGP security

11.6 Best Practice Routing

- For improved convergence in a dual homed branch, use unique Route Distinguisher per LAN-VRF to advertise unique information for the same local prefix.
- Use OSPF network as point-to-point where possible to prevent unnecessary CPU cycles.
- Where BGP Peers have the same outbound policy, configure them as part of the same BGP peer group.
- Enable HMAC-MD5 authentication on all interfaces on which the IGP runs.
- Configure BGP authentication and TTL for the BGP sessions.

12 QoS

During network congestion Quality of Service (QoS) configuration (also known as Class of Service or CoS) can be used to ensure business critical traffic is prioritized over less important traffic and is treated with higher priority by the network. Class of Service can also be used for other tasks like policing, shaping and marking of the QoS bits in the IPv4/IPv6 & VLAN header.

12.1 QoS Stages

The following picture depicts how a packet flows through VOS Edge Devices with regards to QoS function:

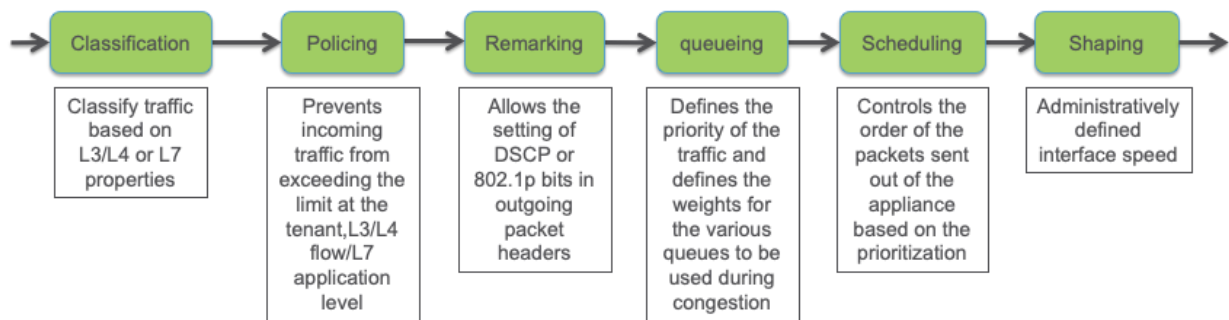


Figure 100 QoS Instruction Pipe Line

12.2 Classification

VOS Edge Devices supports a total 4 Traffic Classes with maximum 4 queues each and low/high drop probability (total 32 unique priorities):

- Network Control
- Expedited Forwarding
- Assured Forwarding
- Best Effort

There are 16 Forwarding Classes in total that map to the forwarding queues using a default mapping:

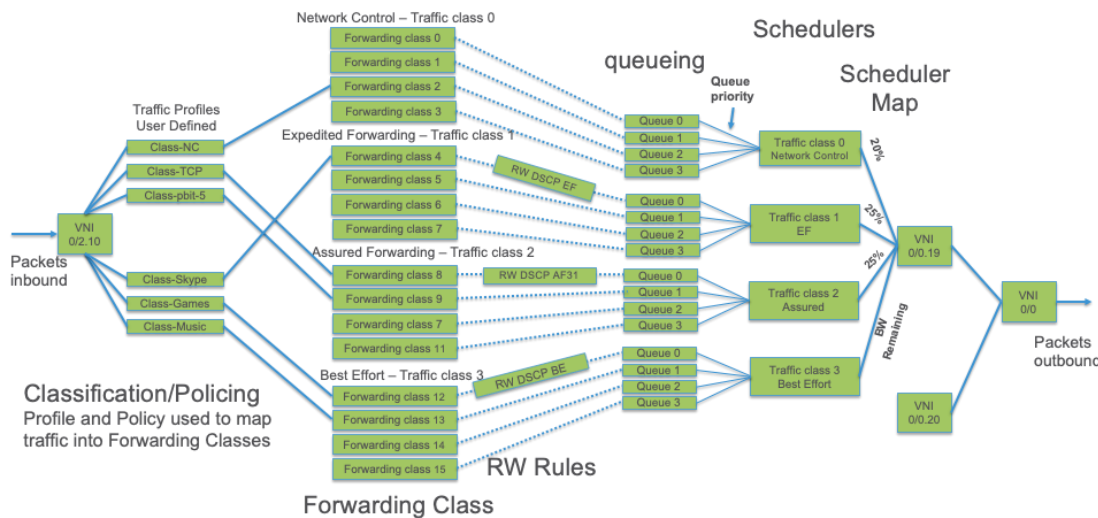


Figure 101 QoS Framework

During classification stage, ingress traffic is identified and associated to a forwarding class and loss priority. Classification can be achieved in 2 ways:

Using QoS-Policies (L3/L4 rules) which allow classification on the following fields:

- destination zone
- destination port
- DSCP
- ether-type
- ether-type-value
- ieee-802.1p
- IP-flags
- Time of Day
- source port
- destination port
- source zone
- source address
- TTL

Using App-QoS Policies (L3/L7 rules) which allow classification on the same fields as QoS-Policies(L3/L4) as well as classification on applications and URL categories.

12.2.1 Best Practices Classification

- Use the right match conditions especially when source/destination-based match conditions are defined in an App QoS policy.
- When a traffic is matched by both L3/L4 QoS Profile and App-QoS then App QoS rule will take precedence.
- Verify in the session details (CLI) if the session is attached to the expect QoS policy rule.

12.3 Use-Case Scenario

Assume a datacenter hosts an SAP application, Users in remote locations access this SAP application. The requirement is to set the outgoing DSCP value of the SAP traffic on the DC MPLS WAN interface to AF31.

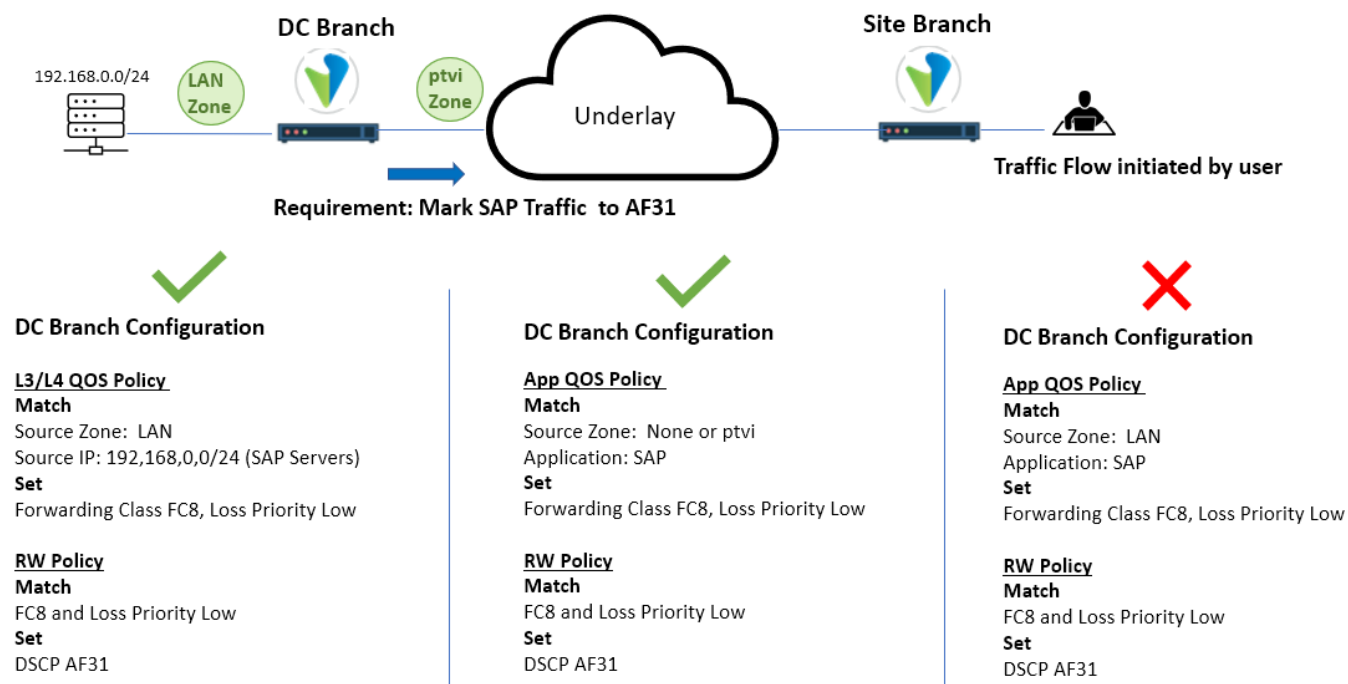


Figure 102 QoS Classification & DSCP Rewrite scenario

The following App-QoS policy is applied on a datacenter VOS Edge Device that matches the following conditions:

- *Source Zone - LAN*
- *Application - SAP*

The App QoS policy references a QoS Profile that puts this traffic in Forwarding Class - FC8. There is then a QoS Propagation Policy (RW Policy) applied on the MPLS WAN network that remarks all FC8 traffic to a DSCP value of af31.

The above configuration will work only if the traffic is originated by the LAN of the DC Appliance. When traffic is originated by the Remote branch clients accessing the SAP application the return traffic arrives from the DC LAN to the client, this traffic will not match the App QoS policy even though the Source Zone is set to LAN. This is because App QoS policy classifies traffic on a flow basis and since the traffic was originated by a remote branch, the first SAP application destined packet in the flow came from the ptvi zone (ptvi zone stands for traffic received over the SDWAN overlay network) it will not match the App QoS Policy which has the source zone as LAN. Therefore, although the QoS RW policy is applied to traffic outgoing into the MPLS network, the classification for this flow happens when the packets arrived on this branch from the remote site.

The proper configuration is to have a match condition either set the source zone to ptvi or leave the source zone as empty which means that the SAP application will be matched irrespective of whether it was initiated from a remote site or the local LAN.

If we must use an L3/L4 QoS policy for the same use case, then we must match based on the IP address of the SAP application or source port as L3/L4 QoS policy does not support Application based matching. For an L3/L4 policy, the match is only in one direction so the source LAN zone can be specified even if the traffic has originated from a remote branch.

12.4 Policing/Rate-Limiting ingress traffic

Policing provides the ability to avoid ingress traffic to overload an egress port or the appliance.

Policing is supported at the following level and this specific order:

- Globally at the Tenant level to enforce licensed bandwidth.
- At L3/L4 QoS-policy level, unidirectional policing ingress only.
- At L7 app QoS-policy level, bidirectional policing (ingress + egress).

Parameters used by the policing function are:

- Peak rate which defines the maximum transmission rate in pps or kbps.
- Burst size which defines number of bytes that are allowed over the configured peak rate. It is used to avoid retransmission during bursts of traffic.
- Loss priority to be used by congestion avoidance algorithms.

12.4.1 Best Practices Policing

- Burst size calculation could use the following formula:
 - Maximum Burst Size is 15000B, starting in release 16.1R2-S10 and above this is the default value used on policer. For earlier release you could use the following formula to calculate the Burst Size:
 - $\text{Policer rate in kbps} / 8$
 - Use Policer rather than Shaper for Real Time Traffic especially voice that is jitter sensitive.
- A Policer at the Tenant level may be configured to protect aggregated SD-WAN licensed bandwidth. Example 2 Tenants are configured on an appliance with a 200 Mbps licensed bandwidth, configure a 100Mbps policer for each of the organization to ensure fairness between the 2 tenants.
- Tenant level policer applies only to the inbound traffic arriving on the appliance from the LAN or WAN interface. Hence if you want to limit the download and upload for a particular organization to say 100Mbps, then a Tenant level policer can be used.

12.5 Access Control List (ACL)

QoS Policies can be used to emulate an Access Control List (ACL) function by denying traffic matching a particular L3/L4 rule in deployment where no security services are configured.

Configuration → Networking → Class Of Service → QoS Policies → Rules → Add/Edit

Configure the match condition, example using source zone:

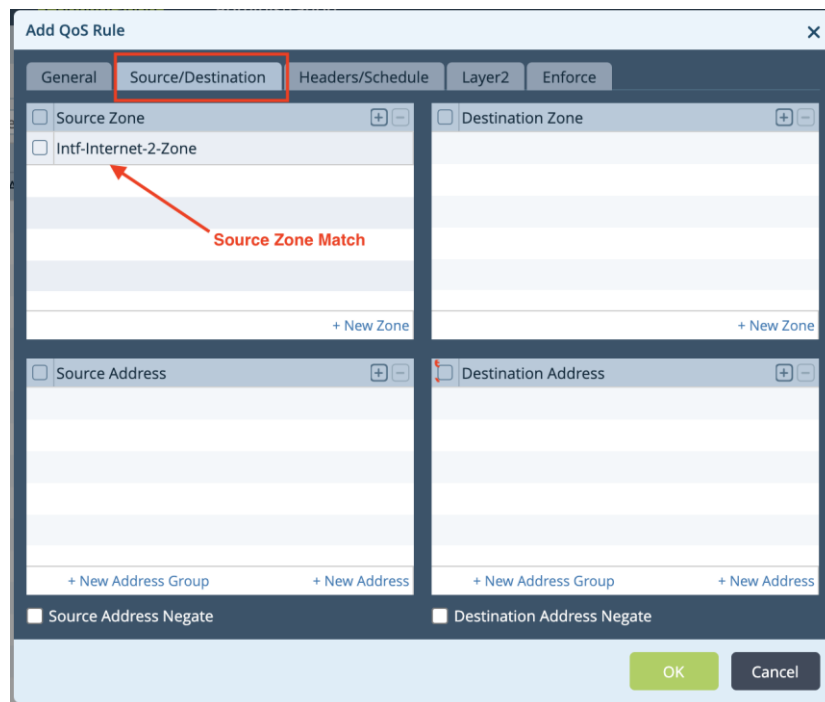


Figure 103 QoS Rule Configuration: Matching

And associated action Deny:

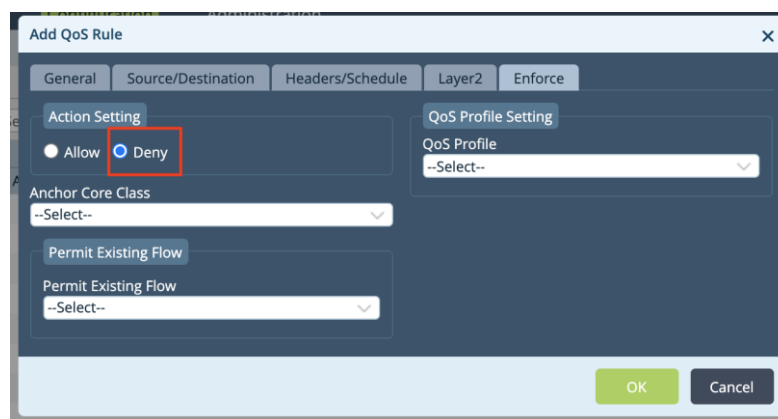


Figure 104 QoS Rule Configuration: Enforcement

12.5.1 Best Practice ACL

- Create a QoS Policy with a source zone matching the appliance interface you want to protect and associate the Deny action. This could be used to protect WAN interface to be pingable from Internet.
- This may be used as a stateless ACL to block certain traffic.

12.6 Hierarchical Shaper

The hierarchical scheduler block when configured is used on an egress interface to schedule and shape traffic. The following diagram displays the arrangement of hierarchical shapers in VOS Edge Devices:

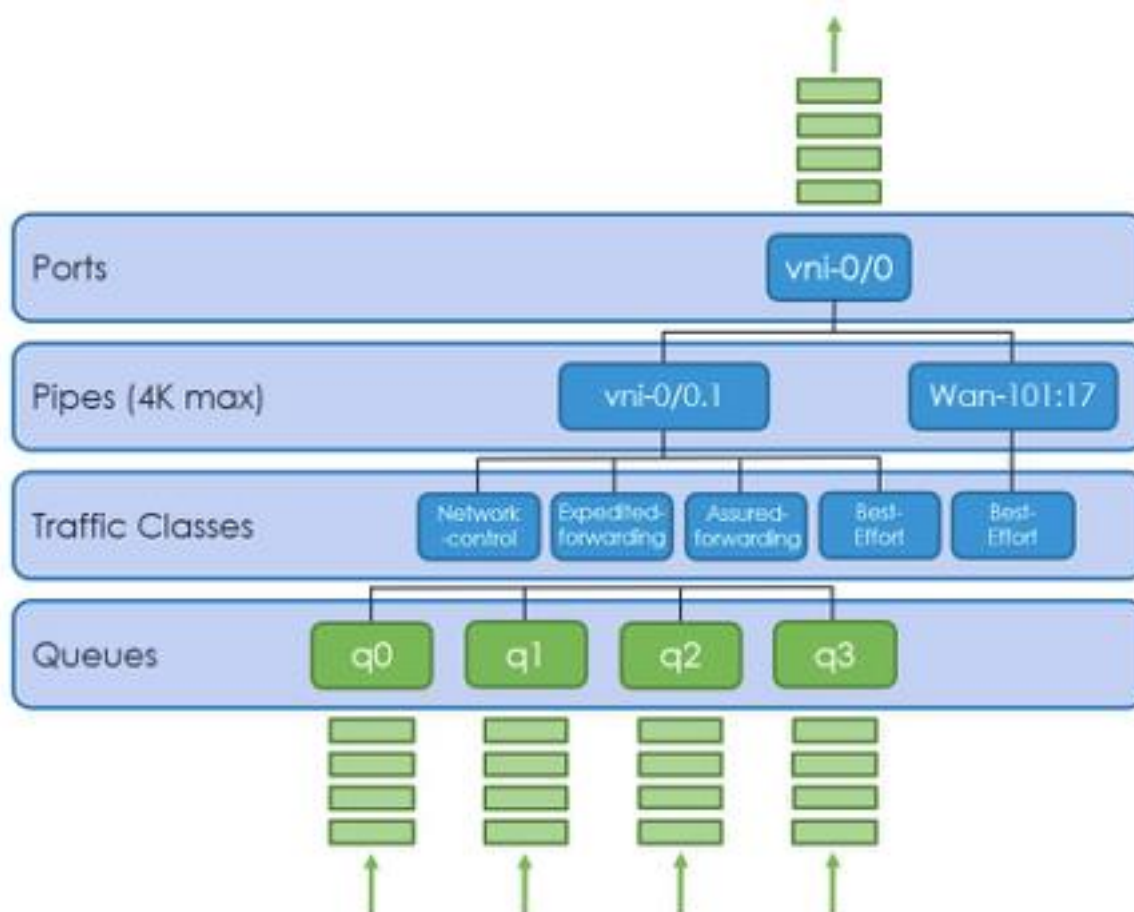


Figure 105 Shaping Architecture

Shaping can be performed using schedulers at the:

- port level
- pipes level
- traffic classes level

All ports have **equal** priority. All pipes within the ports have **equal** priority. These pipes represent either VLAN interfaces or a dynamically created IPsec path. Traffic classes within these pipes are handled in **strict priority order**. Network-control (tc0) > Expedited forwarding (tc1) > Assured Forwarding (tc2) > Best Effort(tc3). Each traffic class contains 4 queues. Each queue within these traffic classes are scheduled using **WRR**.

12.6.1 Best Practices Shaping:

- Apply a shaper at the port level, configure a scheduler per Traffic Class as a percentage of the port shaper rate or using an exact rate to make sure higher priority Traffic Class does not starve other Traffic Classes ensuring priority as well as fairness. For instance, a wrongly configured Network Control Traffic Class will starve Traffic Class 1,2 and 3 if no transmit rate is provided and the traffic in this class is using all available bandwidth on the port.
- For each scheduler users can associate a unique Drop Profile per loss priority for congestion avoidance mechanism using WRED (Weighted Random Early Detection) algorithm which keeps track of queue depth and when threshold is reached starts randomly dropping packets.
- Shapers burst size is automatically set to Interface Link Speed /8000 in Bytes, this is following Intel recommendation. Any higher Burst size than the recommended value might work but could cause intermittent hardware drops.

12.7 Per Tenant Shaper

VOS Edge Devices allows you to configure traffic shapers under a particular Provider Organization in order to allocate the amount of WAN facing bandwidth available on a per tenant basis.

Prerequisite, CoS should be already configured for the Provider Organization and applied to the WAN interface Administrators want to configure a per tenant shaper rate:

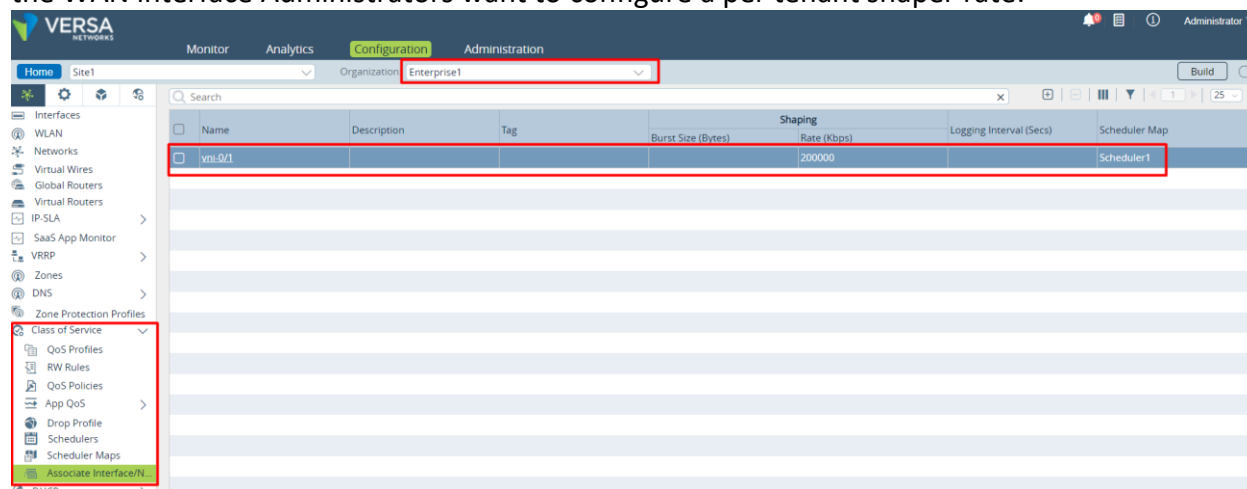


Figure 106 Tenant QoS Configuration

Once CoS is applied to the WAN interface under the Provider Org, per tenant shaper can be configured.

Configuration → Others → Organization → Limits → Edit Tenant

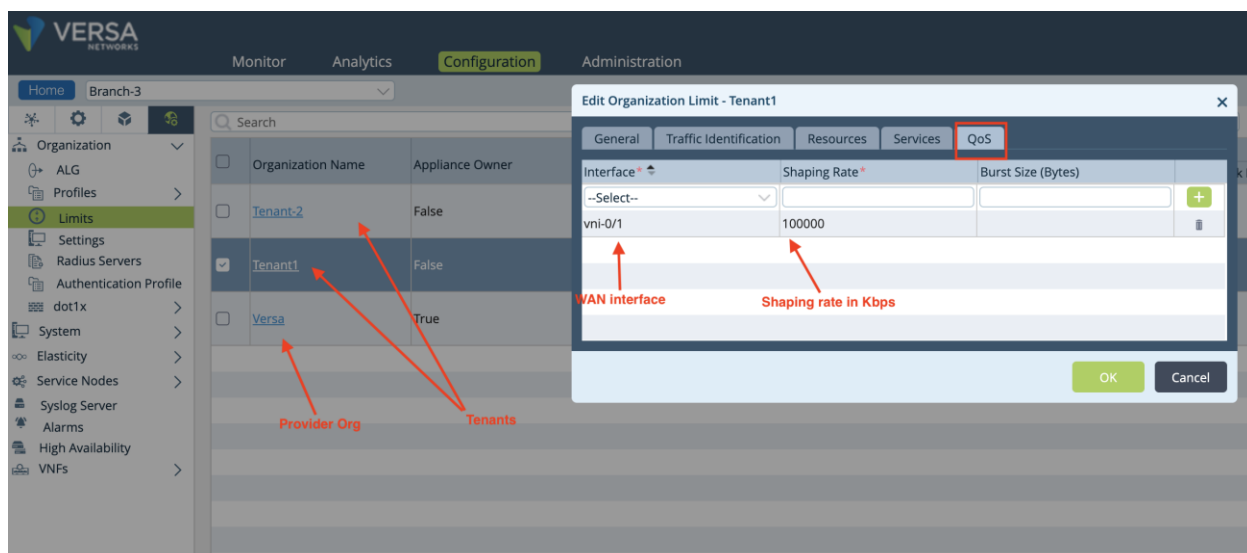


Figure 107 Tenant Shaper Configuration

12.7.1.1 Best Practice Per Tenant Shaper

- On a multi-tenant branch specify the amount of bandwidth each tenant is entitled to use. For example, on a branch with 2 tenants with each 100Mbps bandwidth configure the associated shaping rate at tenant level under the organization limit to enforce fairness between tenants.

12.8 QoS Rewrite and propagation

RW (Rewrite) rules provide the ability to rewrite packets QoS attributes when leaving the appliance to convey the importance of the packet. Downstream nodes can use the QoS attributes to classify traffic and take appropriate action during scheduling or in case of congestion by giving precedence to the mission critical traffic. A RW rule rewrites QoS bits for an existing Forwarding-Class. Following field can be rewritten by RW rules:

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

Versa uses an Overlay technology to transport packets from one branch to another. The packets are encapsulated in a VXLAN header and then transported to the remote branch. Hence there are two headers commonly referred to as the inner header and outer header.

There are two ways of changing the QoS markings on the Inner and Outer headers.

1. Using an RW Policy to classify traffic and set the QoS bits
2. RW Options - Copy the inner header markings to outer header and vice versa

12.8.1 RW Policy

A Re-Write Policy is used to remark the packets and frames based on the classification done by at ingress. The classification is done using an L3/L4 or App QoS Policy where the packets are assigned to a Forwarding Class (FC) and Loss Priority. The RW Policy uses these FC classes and Loss Priority information as match condition to set a particular QoS value.

A typical use-case is rewriting the LAN traffic passing through the CPE, whether it is destined for another LAN port, or egress somewhere on a remote branch. The DSCP of the traffic is modified and carried with the traffic elsewhere to the network.

There are 3 kinds of RW Policies

1. DSCP RW Policy
2. DSCP6 RW Policy
3. 802.1p RW Policy³

Depending on where the RW policy is applied, it modifies the QoS bits of either the inner or outer headers, but it does not propagate.

In a RW policy, the match condition is the Forwarding class (FC) and the Loss Priority. The Forwarding Class (FC) is set in the QoS Profile which is referenced in the L3/L4/App QoS Policy.

The QoS Profile has the following settings that need to be enabled depending on whether the RW Policy used is a DSCP/DSCP6 or 802.1p Policy.

- DSCP Rewrite
- Dot 1P Rewrite

Figure 108 Enabling RW policy in QoS Profile

If this setting is not enabled in the QoS Profile, then the RW Policy will not take effect. The use case of this setting is that there can be more than one QoS Policy that classifies traffic to the same Forwarding Class. But all the traffic may not be required to be remarked by the RW

³ In VOS Edge Device software version 20.2.3 and 20.1.1 onwards, the application of the 802.p RW policy will be moved from the Tunnel interface to the WAN Network.

policy and then this setting can help to differentiate which traffic within this Forwarding class will be evaluated by the RW Policy.

Another use-case would be in a multi-tenant appliance where the wan interfaces are owned by the Provider and has the RW Policy applied. There can be two customer tenants on the branch and each tenant can use the same Forwarding Class in their QoS Policies and depending on whether they want to have their traffic remarked, they can either set or unset the DSCP Rewrite and Dot1P Rewrite flag.

12.8.2 RW Options

The RW Options sets a flag to copy the markings between the Inner and Outer IP headers. This setting is global which means it affects all the traffic through the appliance and cannot be applied on a per interface or on a per traffic class basis.

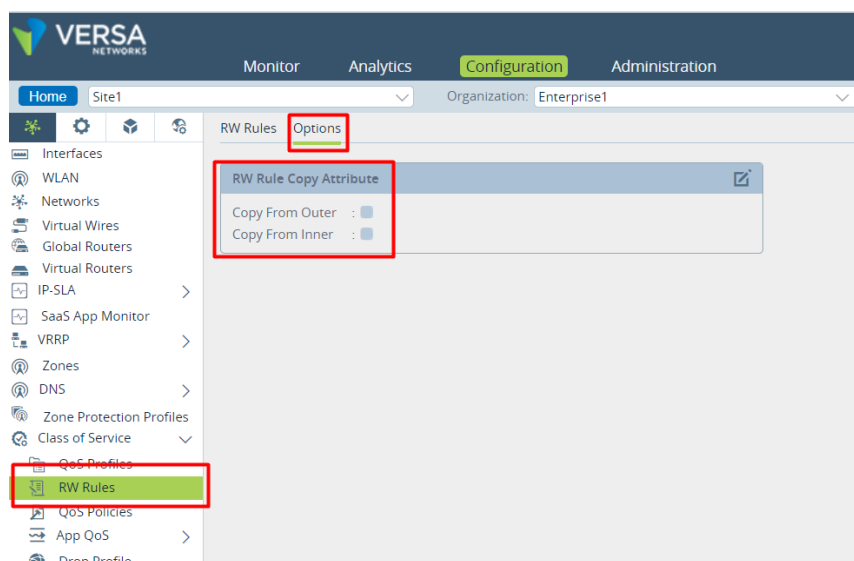


Figure 109 QoS RW Options

RW Option	Modifies
Copy from inner	Copies the marking of the inner IP header to the outer IP header when branch sends packets to a remote branch
Copy from outer	Copies the marking of the Outer IP header to the Inner IP header when packets are received from a remote branch

12.8.3 Copy from Outer

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 the LAN and, in the hub, the outgoing network can be a WAN interface.

“Copy from Outer” works on the model where the VOS Edge Device trusts the markings from the underlay. Outer markings should only be trusted when received on a private circuit like

MPLS but not from Internet. Hence this setting should only be used on a branch that only has an MPLS wan circuit.

12.8.4 Copy from Inner

It works on the model where the Appliance trusts the markings from the LAN. Since this setting remarks the outer IP header of the traffic coming from the LAN site, if there is an RW policy applied on the WAN network that also remarks the Outer IP header of the same packet then the marking made by the “Copy from Inner” setting will override the marking made by the RW policy.

12.9 QoS Propagation Policy on the Hub

On a Hub, the traffic arrives encrypted from a remote site, gets decrypted and then is encrypted again before sending it to another site. Since the packets are decrypted on the Hub, the QoS propagation policy is able to remark both the inner and outer header based on the classification made on the inner packet (actual application packet) by the QoS policy. Even the RW option flags like “Copy from Inner” and “Copy from Outer” can be used.

For Additional details on CoS configuration please refer to:

https://docs.versa-networks.com/Versa_Operating_System/Versa_Operating_System_Network_and_System_Configuration/Configure_CoS

13 Versa Analytics Scaling Recommendations

The following should be considered when scaling Versa analytics.

13.1 Usage Monitoring Logging Recommendation

When Log Export Functionality (LEF) statistics logging is enabled for firewall and SD-WAN, a VOS CPE device performs usage monitoring of firewall and SD-WAN traffic activity and exports various categories of usage statistics to Analytics nodes.

For the firewall service, the statistics for usage monitoring are aggregated for each unique source and destination IP address for each tenant. For the SD-WAN service, the statistics for usage monitoring are aggregated for each unique combination of tenant, application, source IP address, and access circuit.

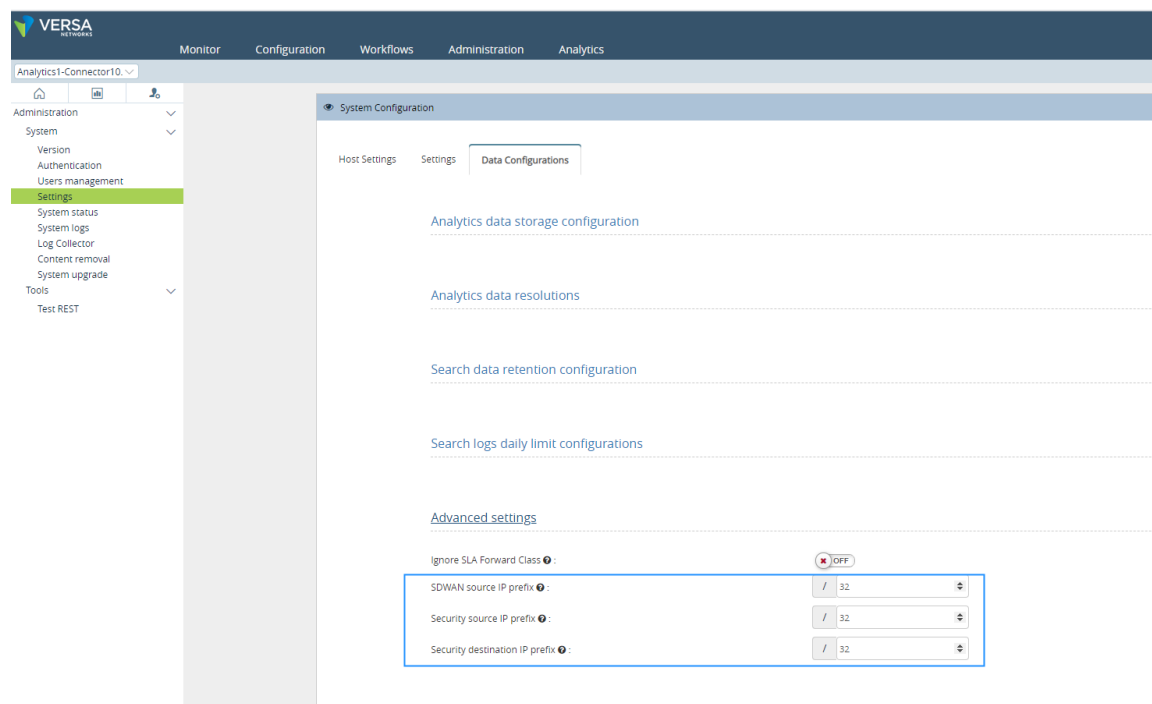
By default, a CPE exports all aggregated usage monitoring records up to a maximum of 16,384 (16K) records every 5 minutes.

To reduce the number of exported statistics log records, you specify the maximum number of log records to export per category or report type. For the busiest CPE devices (example hubs, back offices), it is recommended that you reduce the number of logs to a smaller number every 5 minutes.

Another way to reduce the storage is to use prefixes while storing these records instead of storing /32. The logs will be received with /32. However, while storing the data inside the database configured prefix is used.

Details of how to enable these configuration settings (20.2 and above releases) can be found in below link.

https://docs.versa-networks.com/Versa_Operating_System/VOS_Security_Configuration/Configure_Firewall_and_SD-WAN_Usage_Monitoring_Controls



13.2 Flow logging Recommendation

Versa analytics search engine provides ability to do live analysis of flow logs received for various services such as traffic monitoring, firewall logs, threat logs etc. Search engine indexes various fields of the logs to be able to issue queries with multiple filters.

Volume of flow logs can be huge if enabled by default for all the flows on a flexVNF device for various services. For example, a typical SDWAN branch can generate > 2 million access logs per day. These logs require huge amount of storage and compute to be able to retrieve the data in real time. So, these logs are kept for a short period of time in the database. Retention period per log type is set by default and can be changed using the GUI.

To avoid running out of resources on the search nodes, Versa recommendation is as follows:

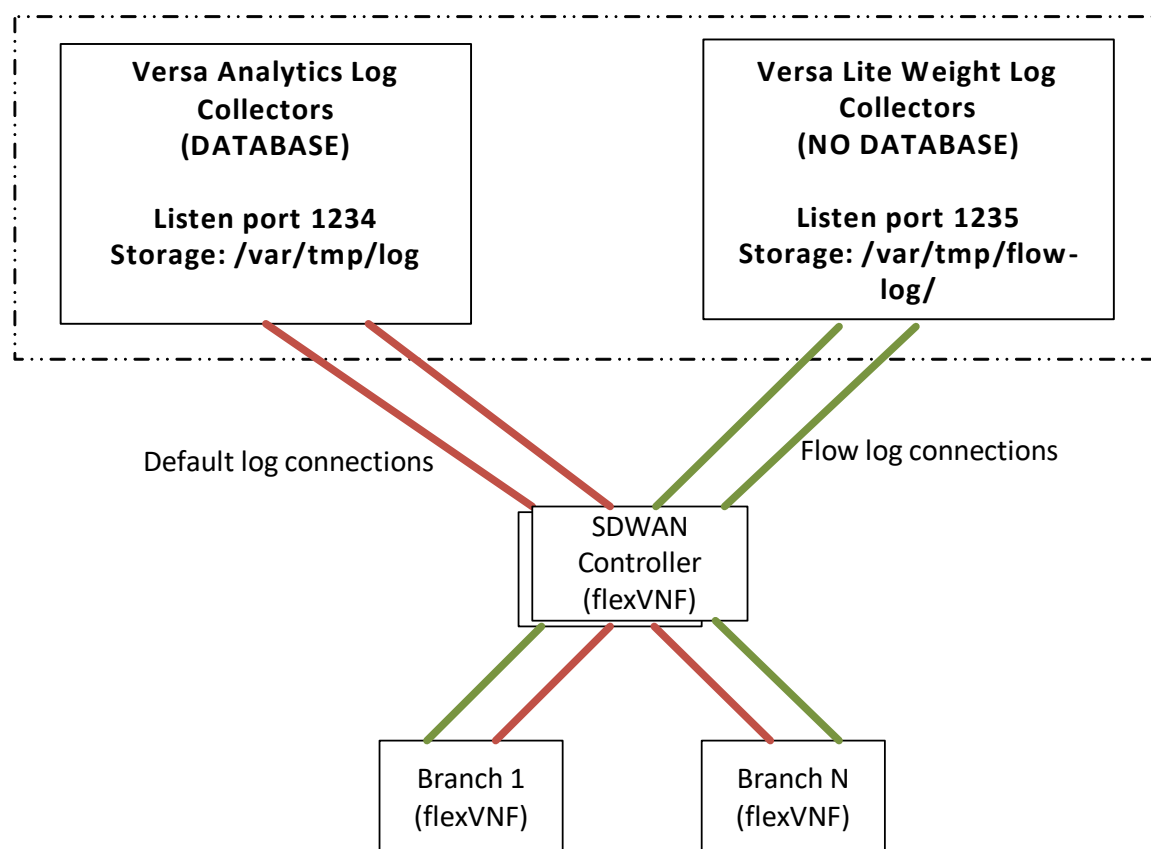
13.2.1 Limiting Flow Logging on the CPE

Enable flow logging on demand or for critical traffic such as threat logs or deny logs.

13.2.2 Flow Logging for Achieve Only

Configure flow log collector to collect and archive the logs without storing inside analytics database.

In this proposal, instead of ingesting into the database, logs are kept in files with timestamp as filenames under per tenant and per appliance directory. Here we will have 2 log collectors on an analytics node or alternately they can be on 2 different nodes. One log collector will listen on port for receiving logs to be consumed by analytics database and the other will listen on port for receiving flow logs that will be just kept for archival.



13.2.3 Throttling/disabling flow logs in log collectors

In some of the deployments, flow logging is enabled for all traffic for various reasons. The log collector exporter process parses these data and stores in files for database to consume. However, the database ingestion may not be able to keep up with the load resulting in delayed processing of the data and increased disk utilization. To avoid such backlogs, following configuration options can be enabled to throttle/disable flow log storage into disk.

13.2.4 Database global limits and retention settings

Set daily global limit on the cluster to ensure that flow logs stored in the database does not exceed configured limit.

Search logs daily limit configurations can be done under Administration->Settings->Data Configurations hierarchy. This will limit number logs ingested into the database in a day globally and/or at a per tenant level.

An alarm will be generated to indicate that the limit has reached. Limit value can be decided based on the cluster size, storage available per node, log retention policy etc. For a 4-node cluster, 20 million logs in a day would be a reasonable number.

13.3 Alarm Logging Recommendations

Alarm logs are sent to Versa Analytics nodes like any other logs from the devices in IPFIX format. These logs are stored in the analytics search nodes for any correlations. By default, these logs are kept for 7 days. The value can be modified by going to Administration>Settings>Data Configuration>Search data retention configuration hierarchy.

13.4 Historical Data Storage Recommendations

Various performance and fault monitoring data are stored in analytics nodes for historical reporting. Data are stored in 2 granularities: Hourly (5, 15, 30 or 60 mins) and daily.

Recommendation is to keep the hourly data for maximum 30 days and daily data for 3 months. Depending on storage available and cluster size, daily data can be kept for longer duration.

Cleanup cron scripts:

To delete aggregated analytics data from database, a cron job is automatically run every midnight to clean analytics aggregated data based on configuration under Administration>Settings>Data Configuration>Analytics aggregated data retention time.

For further details on this topic refer to;

https://docs.versa-networks.com/Versa_Analytics/Versa_Analytics_Configuration/Manage_Analytics_Logs

14 Traffic Tunneling to Cloud Security Proxy

Internet-bound traffic from the SDWAN network could be secured by an external, cloud-based, proxy and firewall services. In that scenario, the Internet traffic from a Branch location is routed, or more precisely tunnelled, through Cloud Security Proxy, short CSP, points of presence before it gets sent to the public Internet.

14.1 Tunneling Protocol Selection

There are two possible protocols commonly used for traffic tunneling to Cloud Security Proxy, GRE and IPsec. Using IPsec has the following advantages:

- There is no need for address provisioning on the CSP provider side. That is important for the case, where IP addresses for the Versa CPE are assigned dynamically
- NAT traversal: IPsec protocol traverses NAT, while GRE does not. This gives more flexibility for the customer who might decide to employ NAT services on Versa CPE devices
- Tunnel Monitoring: IPsec protocol maintains tunnel state, what is beneficial in monitoring these tunnels and acting upon network state changes impacting tunnel operation; for example, withdraw the route on the tunnel down state.

Versa recommends using the IPsec.

14.2 Default Topology

Versa allows for the topology to be built using the WorkFlow template starting from the release 20.x and onwards. This section describes the default topology created by the Workflow. Release 20.2.x supports IPsec Based Tunnels via the workflow

As depicted in the diagram below, the IPSEC/GRE terminates in the Internet Transport VR. An additional TVI interface is created and added to LAN VR to represent the tunnelled endpoint.

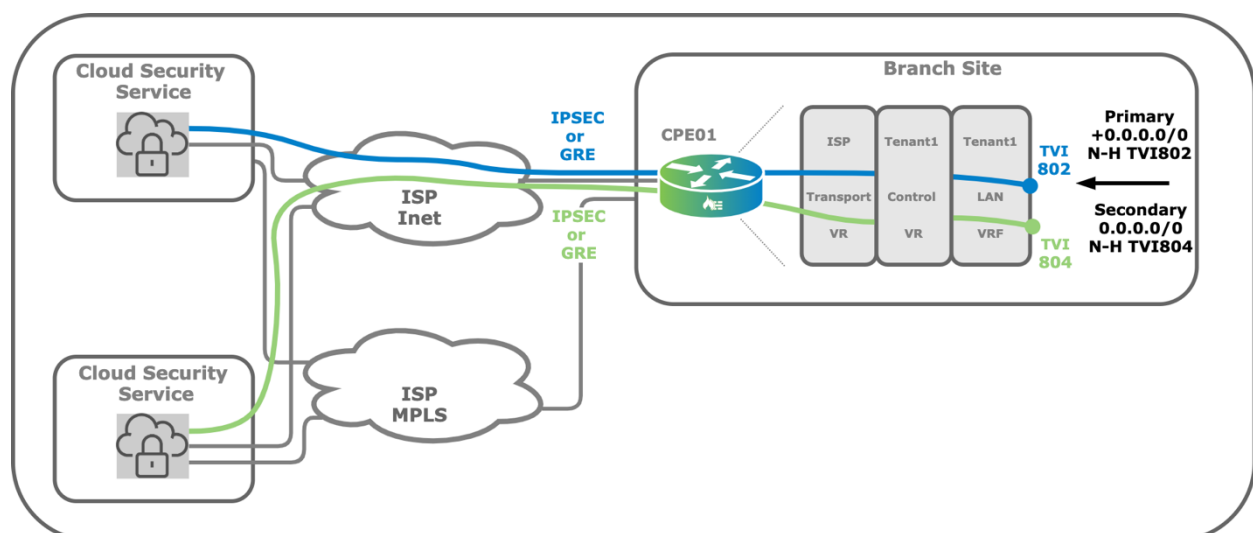


Figure 110 Site2Site Tunnel to Cloud Security Proxy

While VOS support use of the GRE and IPSEC protocols, the configuration of the tunnel parameters will differ for the each of them. For the IPSEC the VPN profile is created to enclose

all required IKE/IPSEC related parameters. While for the GRE tunnel an interface configuration will define what public IP address is selected as the source of the tunnel.

It is assumed the customer would want to have resiliency of this CSP service therefore the WorkFlow template is creating two tunnels to a different CSP endpoint. Tunnels are equal in importance, in fact, and only the routing policy will determine which one of them is going to send the user data actively making the other one standby.

Using the same WorkFlow template, the user can configure routing policy. For instance, a static default route pointing to the tunnel. Example of the static default route with higher priority to the blue tunnel is shown in the diagram above. Further, Versa VOS support use of the dynamic routing protocol, the BGP, if also supported by the CSP server. The use of the BGP allows for easy configuration for traffic load balancing configuring BGP Equal Cost Multipath (ECMP).

14.2.1 Active Tunnel Failure

Active tunnel failure should result in a tunnel change from Standby tunnel to become active. When using IPSEC, this will come naturally due to IPsec inbuilt DPD mechanism, while for the GRE it will not.

Important Note !
 For the GRE tunnelling Versa recommends setting up IP-SLA monitor for the tunnel state. The best practise would be to use public known IP address reachable through the tunnel, i.e. Google DNS IP. These monitors will help withdrawing the static route in an event of the tunnel failure.

Failure of the primary tunnel and route change is depicted in the diagram below.

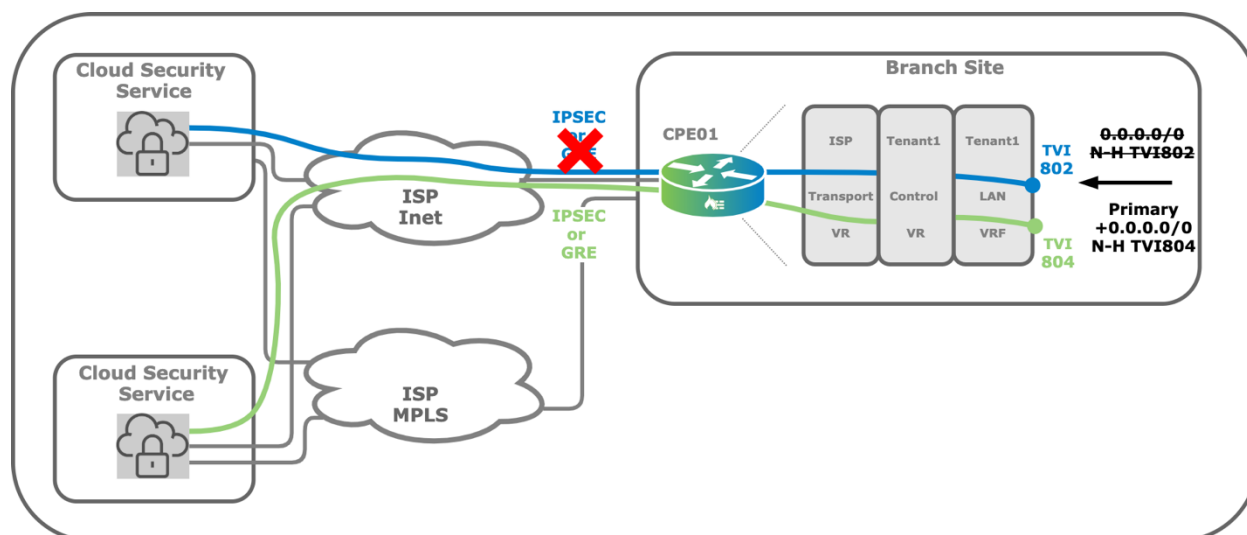


Figure 111 Failure of the primary tunnel

14.2.2 Application Performance Monitoring

Versa recommends considering the use of the Application Performance Monitoring (APM) for the most critical applications, that are tunnelled to CSP.

APM is based on TCP monitoring, and it performs passive application monitoring, or single-ended measurement, of TCP metrics for performance measurement, which include:

- connection setup time
- server connection reset rate
- application response time
- retransmission rate
- network round-trip time.

APM will provide detailed insight in the application performance helping the user to further optimize the network to achieve the best possible user experience.

14.3 Single Site Traffic Load Balancing

While using the default WorkFlow template the user will build a topology with Active and Standby tunnels as discussed in the paragraph above. An enterprise may want to alter this behaviour and use both tunnels simultaneously and load balance the traffic over both of tunnels actively.

Providing that Cloud Security Proxy server supports the BGP protocol, Versa recommends an option to install two Next Hop pointing to both tunnels and assign equal cost to them, what is known as the BGP ECMP.

However, in the situation where BGP is not supported by the CSP, Versa propose the below solution: creating an additional Transport VR per tunnel, configuring the split tunnelling from this Transport VR to the LAN VR and setting up the eBGP session on that connection. This would eventually provide desired traffic load balancing with the BGP ECMP.

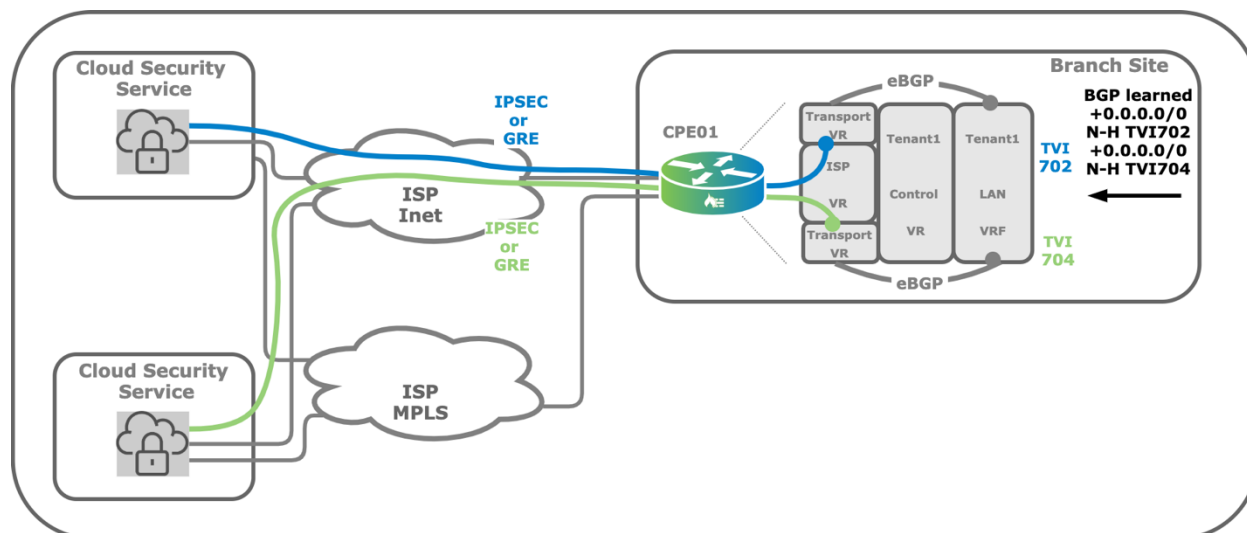


Figure 112 Traffic Load Balancing

14.4 Local and Remote Tunnel

In this section we would consider a network architecture with a centralised Hub located in the Data Centre. This centralised Hub would have an additional tunnel connectivity with the Cloud Security Proxy. The branch site CPE has an alternative transport connectivity; hence it can reach the centralised HUB location using either of the transport networks. See the diagram below.

This topology would give an opportunity to build an additional backup connectivity to the CSP using the tunnel on the remote location– the Hub.

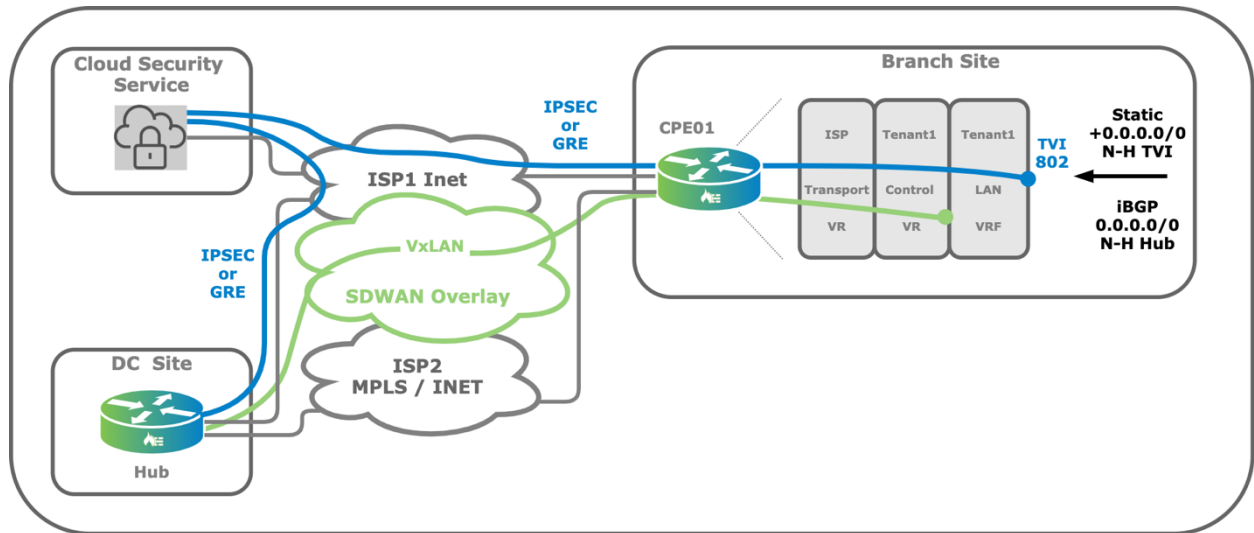


Figure 113 Local and Remote Tunnel

Technically the backup path would be created using an overlay to the central Hub.

The central Hub would be advertising the prefix pointing to the tunnel, in our case default route, to the remote locations. Hence the CPE routing table would have a primary path with Next Hop pointing to the local tunnel and remote backup path learned via iBGP with Next Hop being the remote Hub.

14.4.1 Remote Backup

In an event of the failure of the local Internet connection to the ISP, both local tunnels will be unavailable. As the consequence the CPE will revoke the static from the routing table (for the GRE, the IP-SLA monitor is the vehicle to achieve that) and use the route learned over iBGP.

Providing that the CPE has an alternative transport connectivity to the remote Hub the traffic will continue to flow over this remote tunnel. This is represented by the dotted line in the diagram below.

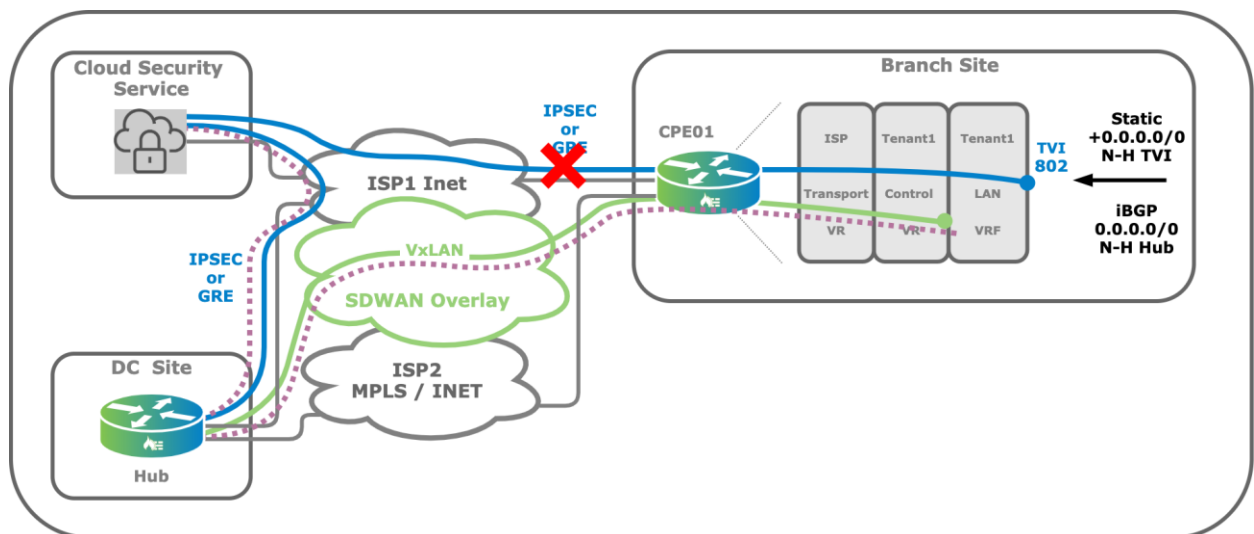


Figure 114 Failure of the local tunnel

14.4.2 SAAS Optimisation

In addition to the remote backup described in the section above the topology allows for sophisticated traffic steering using Versa SDWAN SaaS Optimisation functionality.

For the SaaS applications, the enterprise has an ability to automatically select the best breakout, either the local tunnel or the remote tunnel to reach these applications in the SaaS cloud.

Detailed discussion on the use of SaaS is part of the chapter “Performance based SaaS Optimisation”.

14.5 High Availability Site

Creating the High Availability (HA) site using the WorkFlow template will result in each CPE terminating one tunnel to Cloud Security Proxy. Similarly, to the single CPE site, the tunnel connection will terminate in the Internet transport VR, while an additional TVI interface is created and added to the LAN VR to represent the tunnelled endpoint.

The CPE that becomes master on the LAN, using the VRRP preference, will also become an active CPE tunnelling the user data to the Cloud Security Proxy. In our example above it is the CPE01 with traffic traversing blue tunnel.

The HA site topology has been visualised in the diagram below.

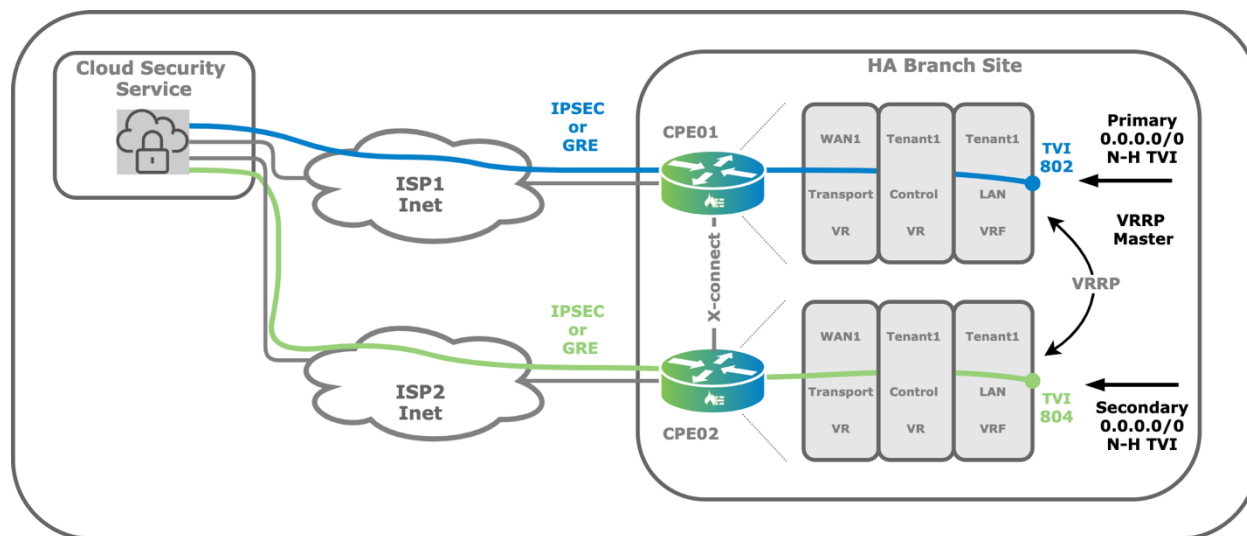


Figure 115 Tunneling on the High Availability Site

Handling the failure of the Internet connection on the primary CPE requires setting up IP SLA monitor and attaching this monitor to the VRRP tracking configuration. An event of the failure of the Internet Connection on the primary CPE the IP SLA monitor is going to fail lowering the VRRP priority of the primary CPE. This will enforce the VRRP master change from primary to secondary.

The traffic will flow from the secondary CPE to the Cloud Security Proxy site over the green tunnel. This scenario is depicted in the diagram below.

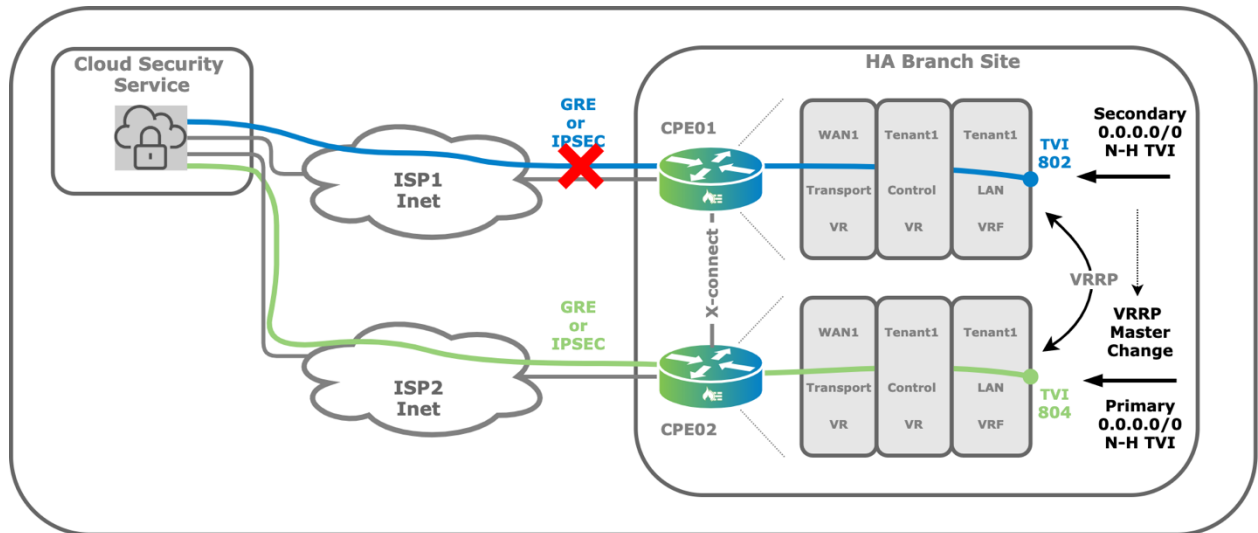


Figure 116 Failure of the Internet connection on primary CPE

14.6 Proxy

A key benefit of an on-prem SASE offering is the ability to selectively identify and breakout traffic. This is achieved using a combination of Versa OS application detection and various proxy functionalities.

14.6.1 DNS Proxy

The Versa solution can take advantage of the DNS proxy functionality to selectively split traffic between internal sites and external sites. For example, internal domains such as internal.company.net would not need to be resolved to an external DNS server. Therefore, does not need to be sent over the tunnel to the cloud proxy provider.

Refer to Section 8.6 for details on DNS proxy.

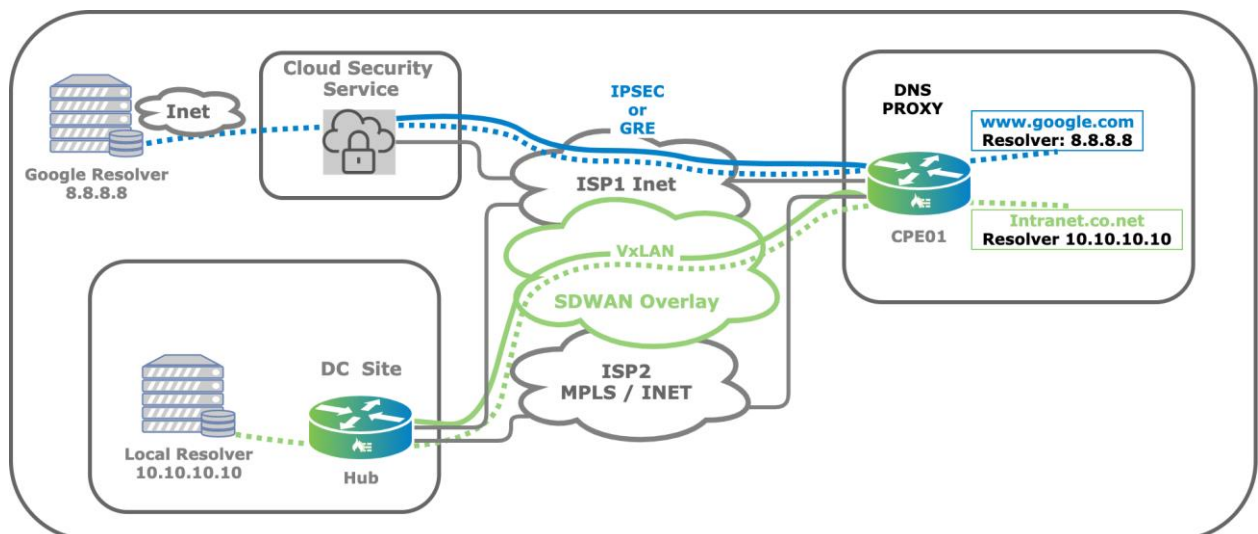


Figure 117 DNS Proxy

14.6.2 HTTP Proxy

Enterprises often require support for both internal and external proxies. Where internal proxies are used to provide centralise monitoring and filtering of internal web applications and external proxies(in this case Cloud Security Proxy providers) are similarly used for internet traffic. Versa OS provides support for HTTP/HTTPS proxies on the branch. This proxies LAN traffic and allows flexible policies to be applied fine grained traffic control.

The http transparent proxy has been shown in the following diagram:

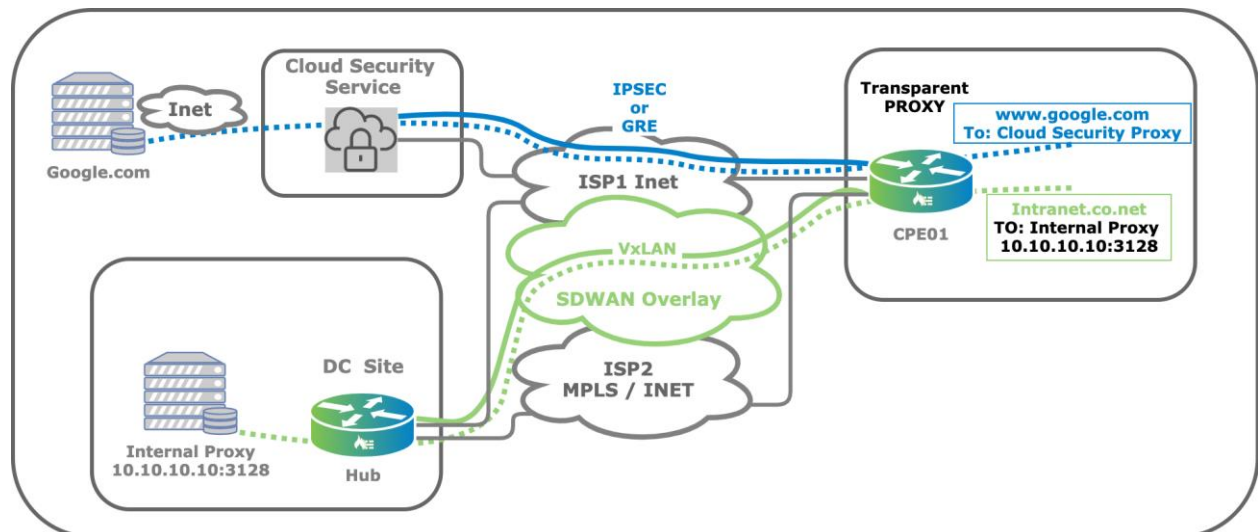


Figure 118 HTTP Proxy

Proxy chaining support is also important as it allows HTTP traffic to be proxies onwards to another proxy. Refer to Section 8.8 for further details.

14.7 Best Practices

In summary Versa recommends the following

- Use IPSEC rather than GRE
- Use dynamic routing protocol such as BGP, over the tunnel to benefit from dynamic route learning and ECMP
- When using GRE, BGP can also help dynamic convergence
 - Use BFD together with the routing protocol where supported by the CSP.
- App monitors provide ability to respond to brown-outs. This may be useful where IPSLA is not adequate to respond to network conditions.
- For the local/remote tunnel breakout with both tunnels active use SAAS optimisation
- While using GRE for statically configured routes (no BGP), use the IP SLA monitors as mechanism for route withdraw
- For the HA site use IP-SLA monitors for the VRRP tracking to enforce Master change

----- END OF THE DOCUMENT -----