Secure Control and Data Overlay Tunnel Solution

For supported software information, click here.

The the Versa Operating System™ (VOS™) SD-WAN uses secure overlay tunnels, which are IPsec-over-VXLAN tunnels, for both the control plane and the forwarding plane.

Tunnel Encapsulation

Different tunnel encapsulation methods are used in the control plane, between SD-WAN CPE devices and Controller nodes, and in the forwarding plane (among CPE devices). In addition, the PKI framework differs between the control plane (CPE devices to Controller nodes) and forwarding plane (CPE devices to CPE devices). Using overlay technologies provides traffic segregation, isolation, and privacy.

The VXLAN encapsulation uses an additional 36 bytes (20 bytes for IPv4, 8 bytes for UDP, and 8 bytes for the VXLAN header), allowing administrators to define the most appropriate encapsulation to fulfill their requirements.

The VOS SD-WAN solution provides true multitenancy, which means you can configure the following encapsulations on a per-tenant basis.

- Traffic isolation:
  - Outer IP – UDP – VXLAN – Inner IP – Payload
  - Outer IP – UDP – VXLAN – Layer 2 - Inner IP – Payload
  - Outer IP – GRE – MPLS – Inner IP — Payload
  - Outer IP – UDP – VXLAN (GPE) – MPLS – Inner IP – Payload
- Traffic isolation and privacy:
  - Outer IP – UDP – VXLAN (GPE) – ESP – Inner IP – Payload
  - Outer IP – ESP (47) – GRE – MPLS – Inner IP — Payload
  - Outer IP – UDP – VXLAN – ESP (47) – GRE – MPLS – Inner IP — Payload

These encapsulation components provide the following:

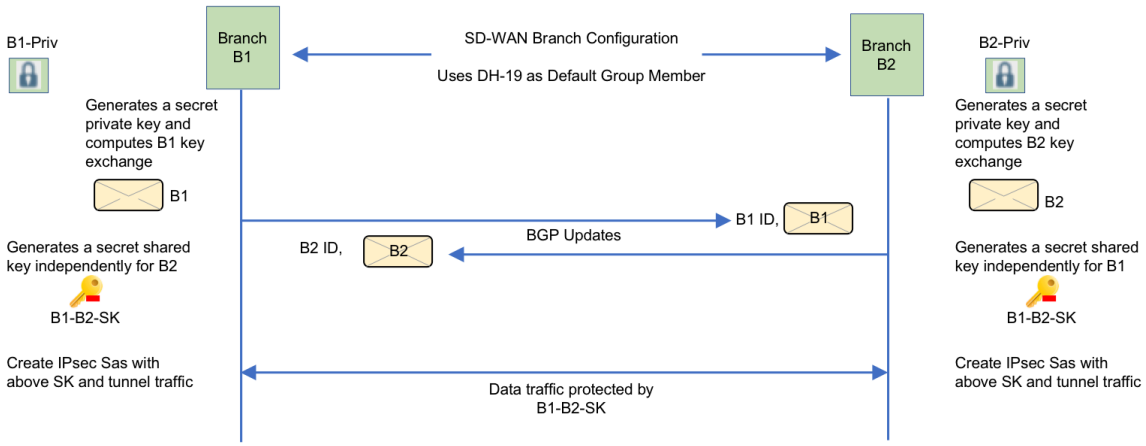
- VXLAN and generic routing encapsulation (GRE) isolate traffic.
- MPLS provides an encapsulation layer that is used to create multiple VPNs on a single tenant.
- IPsec with ESP encryption provides privacy.
- VXLAN generic protocol encapsulation (GPE) provides NAT traversal capabilities. NAT traversal is a key

requirement for SD-WANs. Because SD-WANs establish tunnels over any network. it is likely that NAT and firewall functions exist along the tunnel path. For example, for 4G access, NAT is generally enabled on the the NAT PDN gateway.

To establish a secure control plane, the Versa Networks SD-WAN solution leverages standard IPsec and IKE, along with the PKI framework, if required. For secure data and forwarding planes, however, a different framework is used. The following are the main highlights of this framework:

- Automatic key management with full security
  - Lightweight and scalable.
  - Variable length keys, from 768 through 8192 bits.
  - Strong Diffie-Hellman cryptographic algorithms, including elliptic curve.
  - Secret private keys are computed independently and are never seen on the wire. Instead, the security association (SA) information transits over the secure control overlay tunnels that each CPE device maintains with SD-WAN Controller nodes.
  - Every branch has a unique secret key for each peer, so there are only N-1 keys. This scheme avoids compromising the entire network when one CPE is compromised, thus allowing key revocation on a per-branch basis.
  - Rekeying is based on timers for key rotation per branch.
- Lightweight protocol
  - No overhead from exchanging many messages, unlike IKEv1 and IKEv2
  - Uses an existing secure IKE IPsec tunnel between Controller nodes to send key exchanges. This scheme avoids maintaining N-1 IKE control channels, thus increasing the solution scalability.
  - Versa's own protocol is exchanged through BGP updates. This extension to MP-BGP relating to key management removes the need to set up IKE-based IPsec tunnels between various branch sites.

The following figure shows the key management call flow between branches to establish a secure forwarding overlay tunnel. Branch-to-branch key exchanges are done through the secure control channel that each CPE maintains with SD-WAN Controller nodes using MP-BGP. Note that the control channel is already encrypted, as described earlier, using the standard IPsec/ IKE framework, making it impossible to see branch-to-branch key exchanges on the wire. In this figure, Diffie-Hellman group 19 is used only as an example for the IKE key exchange.



The following are the main features of the VOS IPsec implementation:

- Diffie-Hellman key exchange
- Diffie-Hellman Perfect Forwarding Secrecy
- Antireplay
- Support for both PSK and certificates (OCSP, CRL, chained certificates)
- IKE algorithms supported:
  - AES 128-bit encryption and MD5 hashing
  - AES 128-bit encryption and SHA-1 hashing
  - AES 128-bit encryption and SHA-256 hashing
  - AES 128-bit encryption and SHA-384 hashing
  - AES 128-bit encryption and SHA-512 hashing
  - AES 256-bit encryption and MD5 hashing
  - AES 256-bit encryption and SHA-1 hashing
  - AES 256-bit encryption and SHA-256 hashing
  - AES 256-bit encryption and SHA-384 hashing
  - AES 256-bit encryption and SHA-512 hashing
- IKE Diffie-Hellman groups supported:
  - Diffie-Hellman group 1—768-bit modulus
  - Diffie-Hellman group 2—1024-bit modulus
  - Diffie-Hellman group 5—1536-bit modulus
  - Diffie-Hellman group 14—2048-bit modulus
  - Diffie-Hellman group 19—256-bit elliptic curve
  - Diffie-Hellman group 20—384-bit elliptic curve
  - Diffie-Hellman group 21—521-bit elliptic curve
  - Diffie-Hellman group 25—192-bit elliptic curve

    - Diffie-Hellman group 26—224-bit elliptic curve
- IPsec algorithms supported:
    - ESP-AES128-CTR-SHA1
    - ESP-AES128-CTR-XCBC
    - ESP-AES128-GCM. This is the default.
    - ESP-AES128-MD5
    - ESP-AES128-SHA1
    - ESP-AES128-SHA256
    - ESP-AES128-SHA384
    - ESP-AES128-SHA512
    - ESP-AES256-GCM
    - ESP-AES256-MD5
    - ESP-AES256-SHA256
    - ESP-AES256-SHA384
    - ESP-AES256-SHA512
    - ESP-NULL-MD5

For more information, see Configure IPsec VPN Profiles.

VOS SD-WAN overlay secure, control-and-forwarding networks are constructed using multiple existing point-to-multipoint (P2MP) and point-to-point (P2P) tunnel technologies, which translate into multiple levels of data and control plane encapsulation, as shown in the following tables, which apply to Releases 16.1R2S10 and later and Releases 20.2 and later. This encapsulation is used for both the control and data plane forwarding. These releases add the VXLAN-GRP extension bits, which are used to carry additional metadata for optimizing and ensuring deterministic performance and QoS parameters on intermediate nodes and on the destination SD-WAN nodes.

IPv4/IPv6 (20/40)	**UDP (8)**	**VXLAN (8)**	**VXLAN-GPE Extension (8)**	**ESP (8)**	**IV (0/8/16)**	**MPLS (4)**	**GRE (4)**	**Payload**
Underlay Transport	VXLAN-GPE Port 4790	Base Overlay Header	Extended Metadata			MPLS Service Label VRF ID		

For Release 16.1R2S9 and earlier, the encapsulation is as follows:

IPv4/IPv6 (20/40)	**UDP (8)**	**VXLAN (8)**	**ESP (8)**	**IV (0/8/16)**	**MPLS (4)**	**GRE (4)**
Underlay Transport		P2MP Construct	P2P IPsec Tunnel		MPLS Service Label VRF ID	Data Payload

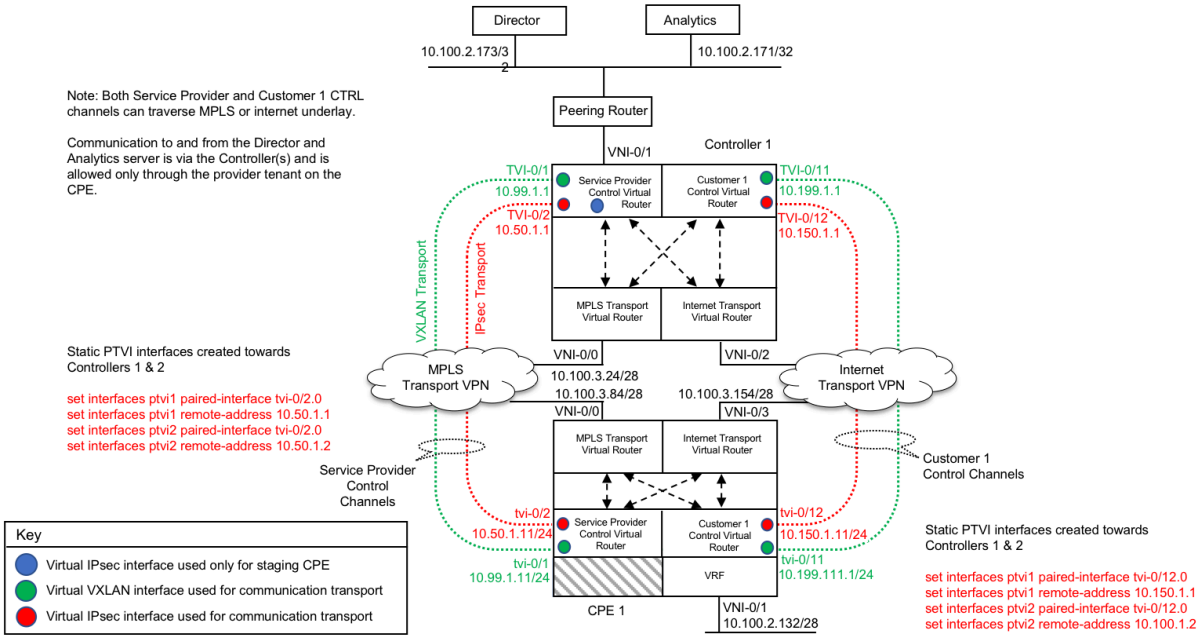
Trailer: Next Protocol (1) + Padding Length (1) + Padding (0-15) + HMAC (12-32)

VOS SD-WAN High-Level Packet Encapsulation

The figure above shows that VOS encapsulation uses several existing technologies:

- Transport underlay—The transport underlay forms the underlying connectivity transport for the SD-WAN VPN service. Depending on the underlay technology, the transport underlay can use IPv4, IPv6, or IPv4/6 with MPLS. The key point is that the underlay technology is irrelevant to the SD-WAN model as long as SD-WAN nodes can reach each other over IP. The underlay interfaces can be viewed as VNI interfaces (that is, physical interfaces) within the Versa interface construct.
- VXLAN—The VXLAN encapsulation provides point-to-multipoint (P2MP) connectivity on a per-tenant basis between the Controller nodes and all the CPE deevices configured for the tenant. VXLAN encapsulation uses an additional 36 bytes (20 bytes for IPv4, 8 bytes for UDP, and 8 bytes for the VXLAN header). The VXLAN interfaces are IP addressed and can be viewed as TVI interfaces within the Versa interface construct.
- ESP—ESP encapsulation provides ;point-to-point (P2P) IPsec tunnel encapsulation over the VXLAN P2MP architecture for P2P-encrypted tunnels between all Controller nodes and all a tenant's CPE devices. In the Versa construct, ESP tunnels can be seen as a combination of additional TVI interfaces to depict the local end of the tunnel and PTVI interfaces to the remote end.

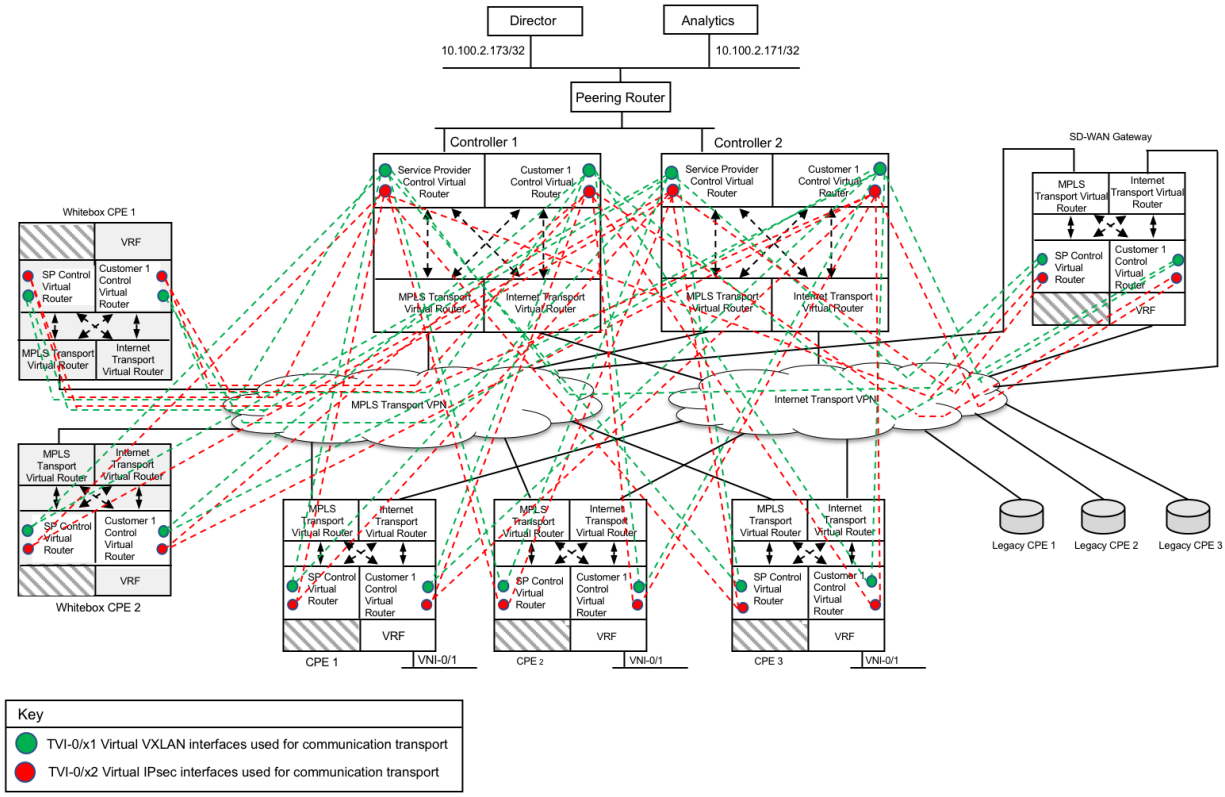
The following figure illustrates the relationship between a CPE device and a Controller node. This figure shows a single tenant (Customer1 and CPE1) in a service provider's multitenant setup, and it shows only control plane–related overlay tunnels. The figure shows the different TVI interfaces, also known as loopback interfaces (green is VXLAN while red is ESP), each porting one layer of the secure control overlay channel.



The default MTU size for a virtual network interface (physical or virtual NIC) is 1500 bytes, and the default MTU size for

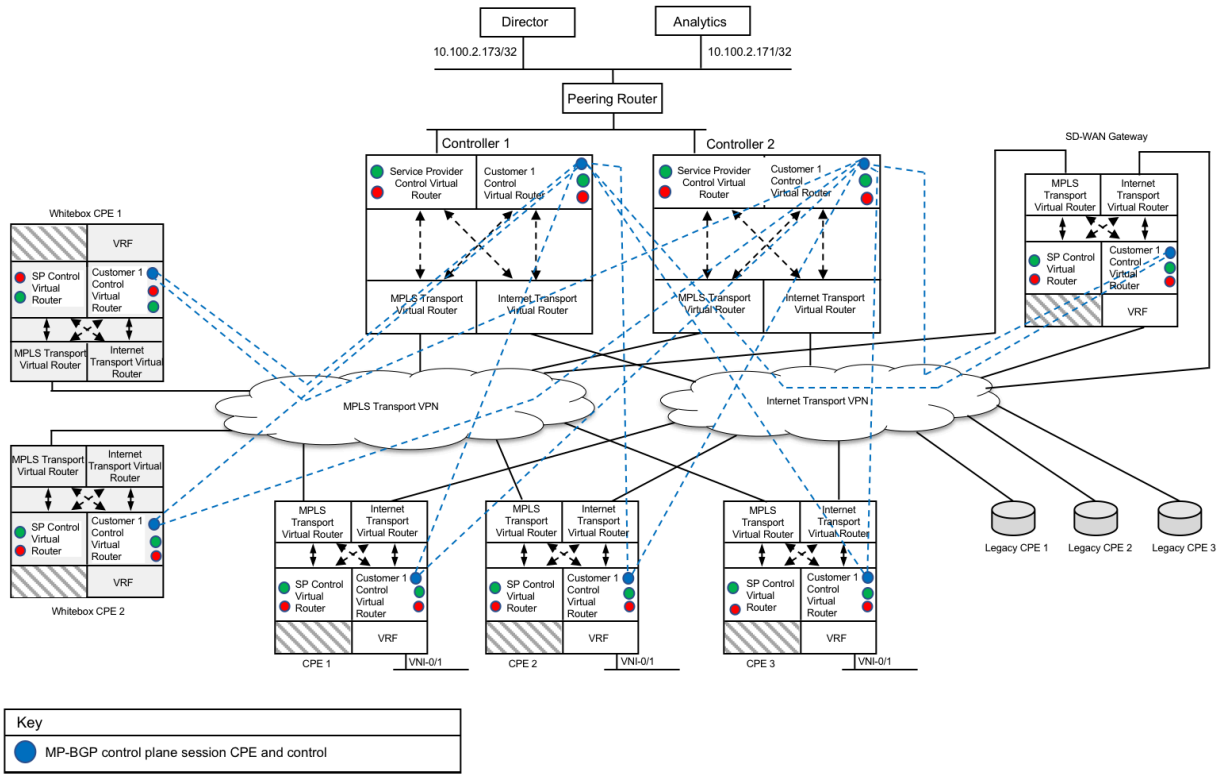
an IPsec tunnel virtual interface (loopback) is 1400 bytes. This means that two IP endpoints, which communicate using an SD-WAN overlay, can communicate with each other by sending IP packets whose size is 1400 bytes without requiring fragmentation. You can increase the VNI and TVI MTU sizes to 9216 bytes and 2500 bytes, respectively, and you can configure the IPsec TVI MTU to a maximum of 9000 bytes. However, note that fragmentation should be avoided as much as possible, because it can add a considerable amount of CPU overhead at CPE gateway level.

The following figure illustrates VOS SD-WAN control plane tunnels with five CPE devices and an SD-WAN gateway. The figure shows both the VXLAN construct (green) and the ESP tunnel construct (red).



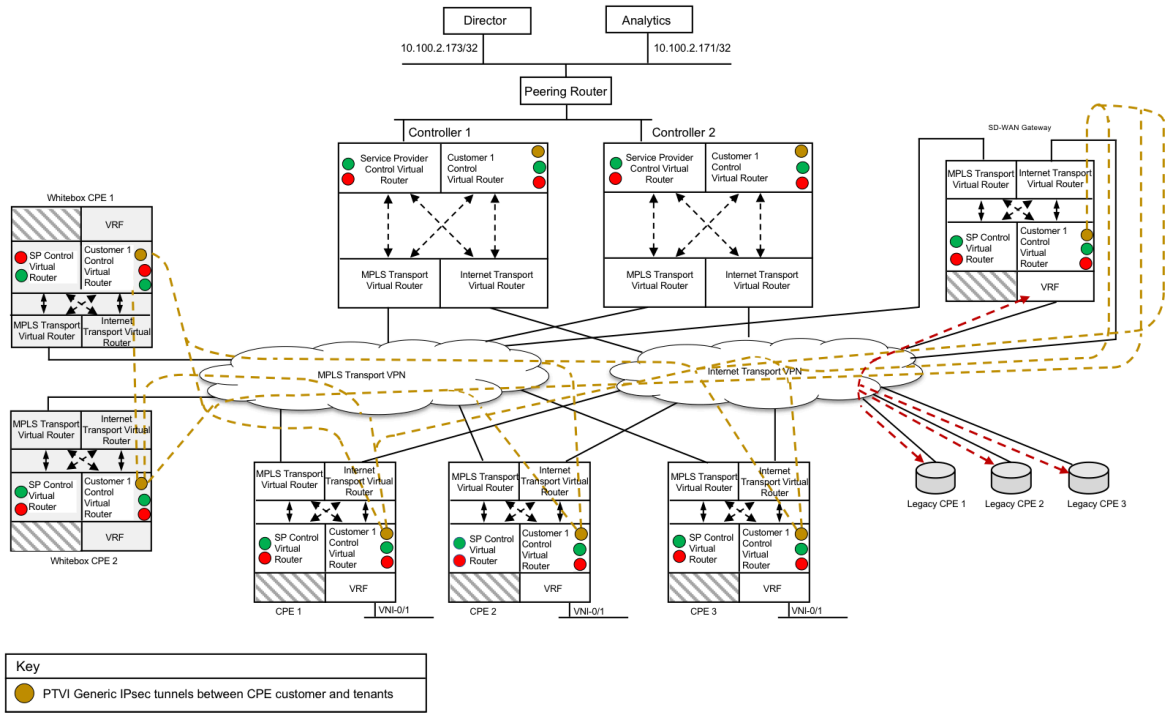
SD-WAN Tunnel Construct (VXLAN and ESP) Control Plane

The following figure shows the underlying SD-WAN VPN transport for the MP-BGP overlay required to service the IP prefix distribution for the tenant's Customer1. A full IBGP mesh is configured between the two Controller nodes that act as route reflectors. Although not shown here, the MP-BGP mesh is configured for each tenant. Therefore, two MP-BGP sessions exist per CPE device, one for each tenant (SP and CUST1).



MP-BGP Connectivity Model for Customer1

The default overlay forwarding plane, in which no redistribution policy is configured, results in a full-mesh topology based on IP prefix advertisement. As a reminder, IP prefixes for each CPE customer-side LAN are propagated towards every other CPE device (and SD-WAN gateway, also depicted here) thorugh the Controller nodes, which act as MP-BGP route reflectors. This full mesh of overlay forwarding/data tunnel topology created by customer prefix distribution results in the forwarding model illustrated in following figure.



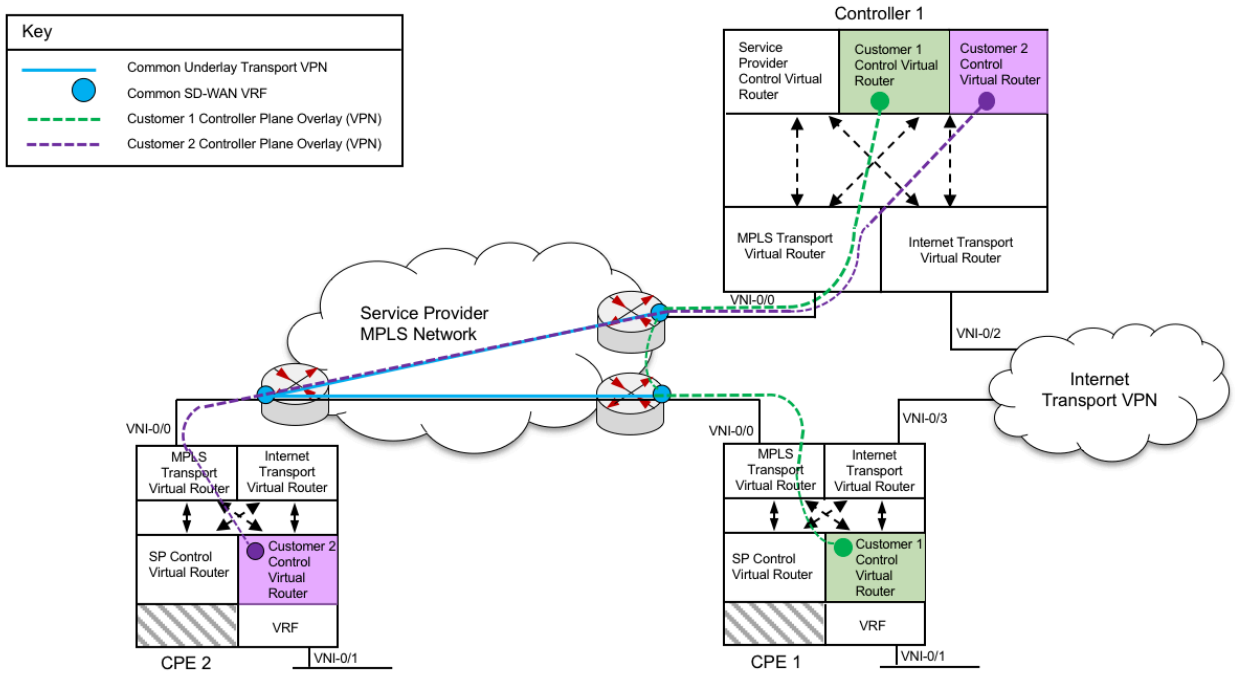
Default Full-Mesh Forwarding Model for Customer1

The VOSs SD-WAN solution is agnostic with regards to the underlying transport networks and, in most cases, gateways, CPE devices, and Controller nodes are deployed using a hybrid WAN model. In this model, a set of links leverages internet broadband access while another set of links is connected to MPLS VPN network access circuits. Using application-based link selection policies, the Versa Networks SD-WAN solution steers traffic on a per-application basis over the most appropriate transport networks. To enhance application-based link selection policies, you can use SLA profiles to look for the underlying transport network matching the requested SLAs. If required, you can configure up to six WAN interfaces per SD-WAN CPE, each leveraging a different transport WAN technology. The Versa Networks SD-WAN solution provides a secure over-the-top (overlay) VPN solution on a per-customer basis using any IP-capable or MPLS-capable network for the transport underlay. With that in mind, most SD-WAN deployments use either one or two underlay networks for transport. These underlay networks use either pure internet connectivity or hybrid connectivity, with internet connectivity and MPLS provided by a service provider.

From a service provider SD-WAN deployment perspective, the common best practice is to provision a single VPN for all SD-WAN data planes and control planes, because the SD-WAN overlay provides granular separation for all customers. The general consensus is that it serves no benefit to provide VPN separation in the service provider MPLS-VPN underlay when the SD-WAN overlay already provides it.

The common approach is to configure a global SD-WAN VRF on the underlay network to provide transport to the SD-WAN solution. This model provides the most scalable approach to an SD-WAN deployment from a service provider perspective.

The most scalable and best practice is to configure the SD-WAN Controller nodes with two transport virtual routers (which form the underlay) that have connections into an internet feed and a common VRF in the MPLS-VPN network. The following illustration shows this model from an SD-WAN control plane perspective.



Single VPN Underlay for SD-WAN Transport

It is recommended that you use a global SD-WAN VRF, because SD-WAN Controller nodes do not support an infinite number of WAN interfaces and because using a per-customer SD-WAN VRF approach tightly limits the total number of tenants that can be hosted by Controller nodes. However, it is still possible to greatly reduce the number of WAN interfaces used on SD-WAN Controller nodes if per-tenant SD-WAN VRFs become a mandatory requirement. However, doing so requires you to configure route leaking at the PE level, leading to additional provisioning overhead. As an example, a customer could deploy Controller nodes connected to their MPLS network using a single WAN interface. The Controller nodes are connected to all the customer's MPLS SD-WAN CPE devices, each having their own VRF using routing leaking on their PE router.

Supported Software Information

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

Additional Information

Configure IPsec VPN Profiles