Unicast Overlay Routing Overview

The overlay network is controlled by the Viptela Overlay Management Protocol (OMP), which is at the heart of Viptela overlay routing. This solution allows the building of scalable, dynamic, on-demand, and secure VPNs. The Viptela solution uses a centralized controller for easy orchestration, with full policy control that includes granular access control and a scalable secure data plane between all edge nodes.

The Viptela solution allows edge nodes to communicate directly over any type of transport network, whether public WAN, internet, metro Ethernet, MPLS, or anything else.

OMP Routing Protocol

The Viptela Overlay Management Protocol (OMP) is the protocol responsible for establishing and maintaining the Viptela control plane. It provides the following services:

- Orchestration of overlay network communication, including connectivity among network sites, service chaining, and VPN topologies
- Distribution of service-level routing information and related location mappings
- Distribution of data plane security parameters
- Central control and distribution of routing policy

OMP is the control protocol that is used to exchange routing, policy, and management information between the vSmart controllers and vEdge routers in the overlay network. It is enabled by default, so after you start up the vSmart controllers and vEdge routers, it is not necessary to explicitly configure or enable OMP. These devices automatically initiate OMP peering sessions between themselves, and the two IP end points of the OMP session are the system IP addresses of the two devices.

OMP is an all-encompassing information management and distribution protocol that enables the overlay network by separating services from transport. Services provided in a typical VPN setting are usually located within a VPN domain, and they are protected so that they are not visible outside the VPN. In such a traditional architecture, it is a challenge to extend VPN domains and service connectivity.

OMP addresses these scalability challenges by providing an efficient way to manage service traffic based on the location of logical transport end points. This method extends the data plane and control plane separation concept from within routers to across the network. OMP distributes control plane information, along with related policies. A central vSmart controller makes all decisions related to routing and access policies for the overlay routing domain. OMP is then used to propagate routing, security, services, and policies that are used by edge devices for data plane connectivity and transport.

OMP Route Advertisements

On vSmart controllers and vEdge routers, OMP advertises to its peers the routes and services that it has learned from
its local site, along with their corresponding transport location mappings, which are called TLOCs. These routes are called OMP routes or vRoutes, to distinguish them from standard IP routes. It is through OMP routes that the vSmart controllers learn the topology of the overlay network and the services available in the network.

OMP interacts with traditional routing at local sites in the overlay network. It imports information from traditional routing protocols, such as OSPF and BGP, and this routing information provides reachability within the local site. The importing of routing information from traditional routing protocols is subject to user-defined policies.

Because OMP operates in an overlay networking environment, the notion of routing peers is different from a traditional network environment. From a logical point of view, the overlay environment consists of a centralized controller and a number of edge devices. Each edge device advertises its imported routes to the centralized controller, and, based on policy decisions, this controller distributes the overlay routing information to other edge devices in the network. Edge devices never advertise routing information to each other, either using OMP or any other method. The OMP peering sessions between the centralized controller and the edge devices are used exclusively to exchange control plane traffic; they are never, in any situation, used for data traffic.

Registered edge devices automatically collect routes from directly connected networks, as well as static routes and routes learned from IGP protocols. The edge devices can also be configured to collect routes learned from BGP.

OMP performs path selection, loop avoidance, and policy implementation on each local device to decide which routes are installed in the local routing table of any edge device.

OMP advertises the following types of routes:

- OMP routes (also called vRoutes)—Prefixes that establish reachability between end points that use the OMP-orchestrated transport network. OMP routes can represent services in a central data center, services at a branch office, or collections of hosts and other end points in any location of the overlay network. OMP routes require and resolve into TLOCs for functional forwarding. In comparison with BGP, an OMP route is the equivalent of a prefix carried in any of the BGP AFI/SAFI fields.

- Transport locations (TLOCs)—Identifiers that tie an OMP route to a physical location. The TLOC is the only entity of the OMP routing domain that is visible to the underlying network, and it must be reachable via routing in the underlying network. A TLOC can be directly reachable via an entry in the routing table of the physical network, or it must be represented by a prefix residing on the outside of a NAT device and must be included in the routing table. In comparison with BGP, the TLOC acts as the next hop for OMP routes.

- Service routes—Identifiers that tie an OMP route to a service in the network, specifying the location of the service in the network. Services include firewalls, Intrusion Detection Systems (IDPs), and load balancers. Service route information is carried in both service and OMP routes.

(OMP also advertises policies configured on the vSmart controller that are executed on vEdge routers, including application-routing policy, cflowd flow templates, and data policy. For more information, see Policy Overview.)

The following figure illustrates the three types of OMP routes.
OMP Routes

Each vEdge router at a branch or local site advertises OMP routes to the vSmart controllers in its domain. These routes contain routing information that the vEdge router has learned from its site-local network.

A vEdge router can advertise one of the following types of site-local routes:

- Connected (also known as direct)
- Static
• BGP
• OSPF (inter-area, intra-area, and external)

OMP routes advertise the following attributes:

• TLOC—Transport location identifier of the next hop for the vRoute. It is similar to the BGP NEXT_HOP attribute. A TLOC consists of three components:
  ◦ System IP address of the OMP speaker that originates the OMP route
  ◦ Color to identify the link type
  ◦ Encapsulation type on the transport tunnel

• Origin—Source of the route, such as BGP, OSPF, connected, and static, and the metric associated with the original route.

• Originator—OMP identifier of the originator of the route, which is the IP address from which the route was learned.

• Preference—Degree of preference for an OMP route. A higher preference value is more preferred.

• Service—Network service associated with the OMP route.

• Site ID—Identifier of a site within the Viptela overlay network domain to which the OMP route belongs.

• Tag—Optional, transitive path attribute that an OMP speaker can use to control the routing information it accepts, prefers, or redistributes.

• VPN—VPN or network segment to which the OMP route belongs.

You configure some of the OMP route attribute values, including the system IP, color, encapsulation type, carrier, preference, service, site ID, and VPN. You can modify some of the OMP route attributes by provisioning control policy on the vSmart controller. See Centralized Control Policy.

---

TLOC Routes

TLOC routes identify transport locations. These are locations in the overlay network that connect to physical transport, such as the point at which a WAN interface connects to a carrier. A TLOC is denoted by a 3-tuple that consists of the system IP address of the OMP speaker, a color, and an encapsulation type. OMP advertises each TLOC separately.

TLOC routes advertise the following attributes:

• TLOC private address—Private IP address of the interface associated with the TLOC.

• TLOC public address—NAT-translated address of the TLOC.

• Carrier—An identifier of the carrier type, which is generally used to indicate whether the transport is public or private.
• Color—Identifies the link type.

• Encapsulation type—Tunnel encapsulation type.

• Preference—Degree of preference that is used to differentiate between TLOCs that advertise the same OMP route.

• Site ID—Identifier of a site within the Viptela overlay network domain to which the TLOC belongs.

• Tag—Optional, transitive path attribute that an OMP speaker can use to control the flow of routing information toward a TLOC. When an OMP route is advertised along with its TLOC, both or either can be distributed with a community TAG, to be used to decide how send traffic to or receive traffic from a group of TLOCs.

• Weight—Value that is used to discriminate among multiple entry points if an OMP route is reachable through two or more TLOCs.

The IP address used in the TLOC is the fixed system address of the vEdge router itself. The reason for not using an IP address or an interface IP address to denote a TLOC is that IP addresses can move or change; for example, they can be assigned by DHCP, or interface cards can be swapped. Using the system IP address to identify a TLOC ensures that a transport end point can always be identified regardless of IP addressing.

The link color represents the type of WAN interfaces on vEdge router. The Viptela solution offers predefined colors, which are assigned in the configuration of the vEdge routers. The color can be one of default, 3g, biz-internet, blue, bronze, custom1, custom2, custom3, gold, green, lte, metro-ethernet, mpls, private1, private2, public-internet, red, and silver.

The encapsulation is that used on the tunnel interface. It can be either IPsec or GRE.

The diagram to the right shows a vEdge router that has two WAN connections and hence two TLOCs. The system IP address of the router is 1.1.1.1. The TLOC on the left is uniquely identified by the system IP address 1.1.1.1, the color metro-ethernet, and the encapsulation IPsec, and it maps to the physical WAN interface with the IP address 184.168.0.69. The TLOC on the right is uniquely identified by the system IP address 1.1.1.1, the color biz-internet, and the encapsulation IPsec, and it maps to the WAN IP address 75.1.1.1.

You configure some of the TLOC attributes, including the system IP address, color, and encapsulation, and you can modify some of them by provisioning control policy on the vSmart controller. See Centralized Control Policy.

Service Routes

Service routes represent services that are connected to a vEdge router or to the local-site network in which the vEdge router resides. The vEdge router advertises these routes to vSmart controllers using service address family NLRI.
OMP Route Redistribution

OMP automatically redistributes the following types of routes that it learns either locally or from its routing peers:

- Connected
- Static
- OSPF intra-area routes
- OSPF inter-area routes

To avoid routing loops and less than optimal routing, redistribution of following types of routes requires explicit configuration:

- BGP
- OSPF external routes

To avoid propagating excessive routing information from the edge to the access portion of the network, the routes that vEdge routers receive via OMP are not automatically redistributed into the other routing protocols running on the routers. If you want to redistribute the routes received via OMP, you must enable this redistribution locally, on each vEdge router.

OMP sets the origin and sub-origin type in each OMP route to indicate the route’s origin (see the table below). When selecting routes, the vSmart controller and the vEdge routers take the origin type and subtype into consideration.

<table>
<thead>
<tr>
<th>OMP Route Origin Type</th>
<th>OMP Route Origin Subtype</th>
</tr>
</thead>
<tbody>
<tr>
<td>BGP</td>
<td>External</td>
</tr>
<tr>
<td></td>
<td>Internal</td>
</tr>
<tr>
<td>Connected</td>
<td>—</td>
</tr>
<tr>
<td>OSPF</td>
<td>External-1</td>
</tr>
<tr>
<td></td>
<td>External-2</td>
</tr>
<tr>
<td></td>
<td>Intra-area</td>
</tr>
<tr>
<td></td>
<td>Inter-area</td>
</tr>
<tr>
<td>Static</td>
<td>—</td>
</tr>
</tbody>
</table>

OMP also carries the metric of the original route. A metric of 0 indicates a connected route.
Administrative Distance

Administrative distance is the measure used to select the best path when there are two or more different routes to the same destination from multiple routing protocols. When the vSmart controller or vEdge router is selecting the OMP route to a destination, it prefers the one with the lower or lowest administrative distance value.

The following table lists the default administrative distances used by the Viptela devices:

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Administrative Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connected</td>
<td>0</td>
</tr>
<tr>
<td><strong>Static</strong></td>
<td>1</td>
</tr>
<tr>
<td>NAT (NAT and static routes cannot coexist in the same VPN; NAT overwrites static routes)</td>
<td>1</td>
</tr>
<tr>
<td>Learned from DHCP</td>
<td>1</td>
</tr>
<tr>
<td><strong>GRE</strong></td>
<td>5</td>
</tr>
<tr>
<td>EBGP</td>
<td>20</td>
</tr>
<tr>
<td>OSPF</td>
<td>110</td>
</tr>
<tr>
<td>IBGP</td>
<td>200</td>
</tr>
<tr>
<td>OMP</td>
<td>250</td>
</tr>
</tbody>
</table>

OMP Best-Path Algorithm and Loop Avoidance

vEdge routers advertise their local routes to the vSmart controller using OMP. Depending on the network topology, some routes might be advertised from multiple vEdge routers. Viptela devices use the following algorithm to choose the best route:

1. Check whether the OMP route is valid. If not, ignore it.
2. If the OMP route is valid and if it has been learned from the same Viptela device, select the OMP route with the lower administrative distance.
3. If the administrative distances are equal, select the OMP route with the higher OMP route preference value.
4. On vEdge routers only, if the OMP route preference values are equal, select the OMP route with the higher TLOC preference value.
5. If the TLOC preference values are equal, compare the origin type, and select one in the following order (select the first match):
   - Connected
   - Static
   - EBGP
   - OSPF intra-area
   - OSPF inter-area
   - OSPF external
   - IBGP
   - Unknown

6. If the origin type is the same, select the OMP route that has the lower origin metric.

7. On vEdge routers only, if the origin types are the same, select the OMP route the higher router ID.

8. If the router IDs are equal, a vEdge router selects the OMP route with the higher private IP address. If a vSmart controller receives the same prefix from two different sites and if all attributes are equal, the vSmart controller chooses both of them.

Here are some examples of choosing the best route:

- A vSmart controller receives a OMP route to 10.10.10.0/24 via OMP from a vEdge router with an origin code of OSPF, and it also receives the same route from another vSmart controller, also with an origin code of OSPF. If all other things are equal, the best-path algorithm chooses the route that came from the vEdge router.

- A vSmart controller learns the same OMP route, 10.10.10.0/24, from two vEdge routers in the same site. If all other parameters are the same, both routes are chosen and advertised to other OMP peers. By default, up to four equal-cost routes are selected and advertised.

A vEdge router installs an OMP route in its forwarding table (FIB) only if the TLOC to which it points is active. For a TLOC to be active, an active BFD session must be associated with that TLOC. BFD sessions are established by each vEdge router, which creates a separate BFD session with each of the remote TLOCs. If a BFD session becomes inactive, the vSmart controller removes from the forwarding table all the OMP routes that point to that TLOC.

A vSmart controller runs the best path algorithm on received paths and stores the best paths in the winning order. It then walks through this list and applies outbound policy on each path. If a path is not rejected by the outbound policy, it is advertised to its peers. A max of configured send-path-limit paths for each prefix are advertised to the peers.

Here are some examples of choosing the paths to advertise to peer:

- A vSmart controller receives 8 paths for prefix 99.99.99.0/24 and the send-path-limit is set to 4. When the outbound policy does not reject any paths the first 4 paths from the received best paths stored in the winning order is selected.

- A vSmart controller receives 8 paths for prefix 99.99.99.0/24 and the send-path-limit is set to 4. When the outbound policy rejects some of the paths, the first 4 non-rejected paths from the received best paths stored in the winning order is selected.
**Graceful Restart for OMP**

Graceful restart for OMP allows the data plane in the Viptela overlay network to continue functioning if the control plane stops functioning or becomes unavailable. With graceful restart, if the vSmart controller in the network goes down, or if multiple vSmart controllers go down simultaneously, the vEdge routers can continue forwarding data traffic. They do this using the last known good information that they received from the vSmart controller. When a vSmart controller is again available, its DTLS connection to the vEdge router is re-established, and the vEdge router then receives updated, current network information from the vSmart controller.

**Support for Traditional Unicast Routing Protocols**

The Viptela overlay network supports BGP and OSPF unicast routing protocols. These protocols can be configured on vEdge routers in any VPN except for VPN 0 and VPN 512 to provide reachability to networks at their local sites. vEdge routers can redistribute route information learned from BGP and OSPF into OMP so that OMP can better choose paths within the overlay network.

When the local site connects to a Layer 3 VPN MPLS WAN cloud, the vEdge router acts as an MPLS CE device and establishes a BGP peering session to connect to the PE router in the L3VPN MPLS cloud.

When the vEdge router or routers at a local site do not connect directly to the WAN cloud but are one or more hops from the WAN and connect indirectly through a non-Viptela hub router, standard routing must be enabled on the vEdge routers’ DTLS connections so that they can reach the WAN cloud. Either OSPF or BGP can be the routing protocol.

In both these types of topology, the BGP or OSPF session runs over a DTLS connection created on the loopback interface in VPN 0, which is the transport VPN that is responsible for carrying control traffic in the overlay network. The vBond orchestrator learns about this DTLS connection via the loopback interface and conveys this information to the vSmart controller so that it can track the TLOC-related information. In VPN 0, you also configure the physical interface that connects the vEdge router to its neighbor—either the PE router in the MPLS case or the hub or next-hop router in the local site—but you do not establish a DTLS tunnel connection on that physical interface.

**Additional Information**

- Configuring OMP
- Configuring Unicast Overlay Routing
- Routing Configuration Example