Multicast Overlay Routing Overview

The Viptela overlay multicast implementation extends native multicast by creating a secure optimized multicast tree that runs on top of the overlay network.

Overlay Multicast Routing Overview

The Viptela overlay multicast software uses PIM Sparse Mode (PIM-SM) for multicasting traffic on the overlay network. PIM-SM builds unidirectional shared trees rooted at a rendezvous point (RP), and each multicast group has one shared tree that is rooted at a single RP. Once a shared tree has been built such that a last-hop router learns the IP address for the multicast source, the router engages in a switchover from the shared tree to initiate the construction of a source (or shortest-path) tree. The source tree uses the lowest metric path between the source and last-hop router, which may be entirely, partially, or not at all congruent with the shared tree.

The Viptela design optimizes multicast packet distribution throughout the overlay network by eliminating packet replication on the ingress router, that is, on the router connected to a multicast source. Instead, the ingress router forwards multicast streams to a vEdge router that is designated to be a replicator, and it is this router that forwards streams to multicast receivers. This design saves bandwidth and computational resources on the ingress router.

The figure below illustrates the Viptela design. Here, the ingress router vEdge-2 forwards the multicast streams from the source at its site to vEdge-1, which is a vEdge router that is designated to be a replicator. vEdge-1 replicates the stream and forwards it to the receiver, which is located behind vEdge-3 at Site 3.
Multicast Protocols

Viptela overlay multicast supports the PIM routing protocol and IGMP.

PIM

Viptela overlay multicast supports PIM version 2 (defined in RFC 4601), with some restrictions.

On the service side, the Viptela software supports native multicast. A vEdge router appears as a native PIM router and establishes PIM neighborship with other PIM routers at a local site. To properly extend multicast trees into the overlay network, a vEdge router may require other supporting routers in a local site. If a PIM-SM RP is required at a site, that function must be provided by a non-Viptela router, because the vEdge router currently has no native support for the rendezvous point functionality. Receivers residing downstream of a vEdge router can join multicast streams by exchanging IGMP membership reports directly with the device, and no other routers are required. This applies only to...
sites that have no requirement for supporting local sources or PIM SM rendezvous points.

On the transport side, PIM-enabled vEdge routers originate multicast service routes (called multicast autodiscover routes), sending them via OMP to the vSmart controllers. The multicast autodiscover routes indicate whether the router has PIM enabled and whether it is a replicator. If the router is a replicator and the load threshold has been configured, this information is also included in the multicast autodiscover routes. Each PIM router also conveys information learned from the PIM join messages sent by local-site multicast-enabled routers, including multicast group state, source information, and RPs. These routes assist vEdge routers in performing optimized joins across the overlay when joining existing multicast sources.

vEdge routers support PIM source-specific mode (SSM), which allows a multicast source to be directly connected to the router.

**IGMP**

Viptela overlay multicast routing supports the Internet Group Management Protocol (IGMP) version 2 (defined in RFC 2236). vEdge routers use IGMP to process receiver membership reports for the hosts in a particular VPN and to determine, for a given group, whether multicast traffic should be forwarded and state should be maintained. vEdge routers listen for both IGMPv1 and IGMPv2 group membership reports.

**Rendezvous Points**

The root of a PIM multicast shared tree resides on a router configured to be a rendezvous point (RP). Each RP acts as the RP and the root of a shared tree (or trees) for specific multicast group ranges. In the Viptela overlay network, RPs are non-Viptela routers that reside in the local-site network. The RP function is typically assigned to one or two locations in the network; it is not required at every site. vEdge routers do not currently support the RP functionality, so non-Viptela routers must provide this function in the applicable sites.

The Viptela software supports the auto-RP protocol for distributing RP-to-group mapping information to local-site PIM routers. With this information, each PIM router has the ability to forward joins to the correct RP for the group that a downstream IGMP client is attempting to join. Auto-RP updates are propagated to downstream PIM routers if such routers are present in the local site.

**Replicators**

For efficient use of WAN bandwidth, strategic vEdge routers can be deployed and configured as replicators throughout the overlay network. Replicators mitigate the requirement for an ingress router to replicate a multicast stream once for each receiver.

As discussed above, replicators advertise themselves, via OMP multicast-autodiscover routes, to the vSmart controllers in the overlay network. The controllers then forward the replicator location information to the PIM-enabled vEdge routers that are in the same VPN as the replicator.

A replicator vEdge router receives streams from multicast sources, replicates them, and forwards them to multicast receivers. The details of the replication process are discussed below, in the section Multicast Traffic Flow through the Overlay Network.
A replicator is typically vEdge router located at a colo site or another site with a higher-speed, or a high-speed, connection to the WAN transport network.

**Multicast Service Routes**

vEdge routers send multicast service routes to the vSmart controller via OMP. From these routes, the controller processes and forwards joins for requested multicast groups towards the source address as specified in the original PIM join message that helped originate the OMP multicast service route. The source address can be either the IP address of an RP if the originating router is attempting to join the shared tree or the IP address of the actual source of the multicast stream if the originating router is attempting to join the source tree.

**Multicast Traffic Flow through the Overlay Network**

Let’s look at a high-level topology of the Viptela overlay network multicast solution to illustrate how traffic from multicast sources is delivered to multicast receivers. The topology contains five vEdge routers:
• vEdge router vEdge-3 is located at a site with two multicast sources, Source-1 and Source-2. This site also has a non-Viptela router that functions as a PIM-SM RP. Even though the vEdge-3 router is the ingress router for streams from these two multicast sources, it performs no packet replication. Instead, it forwards the multicast streams to replicators in the overlay network. The vEdge-3 router has learned the addresses of the replicators via OMP from a vSmart controller.

• vEdge routers vEdge-1 and vEdge-2 are two multicast replicators in the overlay network. Their job as replicators is to receive streams from multicast sources, replicate the streams, and then forward them to receivers. In this topology, the vEdge-3 router forwards the multicast streams from the two multicast sources in its local network to vEdge-1 or vEdge-2, or both, and these routers then replicate and forward the streams to the receivers located in the local sites behind vEdge routers vEdge-4 and vEdge-5. Which replicator receives a stream depends on the group address, the identity of the vEdge routers that joins that given group, and the current load of the replicator. The typical situation is that only a single replicator is replicating traffic for a given group, but this may vary depending on the physical scope of the given group.

• vEdge router vEdge-4 is located at a site that has one multicast receiver, Receiver-3, which receives streams from Source-1 and Source-2.

• vEdge router vEdge-5 is located at another site with one multicast receiver, Receiver-4. This receiver gets streams only from one source, Source-1.

Now, let’s examine how multicast traffic flows from the sources to the receivers.

The two multicast sources, Source-1 and Source-2, send their multicast streams (the blue stream from Source-1 and the green stream from Source-2) to the RP. Because the destination IP addresses for both streams are at remote sites, the RP forwards them to vEdge-3 for transmission onto the transport/WAN network. vEdge-3 has learned from the vSmart controller that the network has two replicators, vEdge-1 and vEdge-2, and so forwards the two multicast streams to them, without first replicating the streams.

The two replicators have learned from a vSmart controller the locations of multicast receivers for the two streams. The vEdge-1 replicator makes one copy of the green stream and forwards it to vEdge-4, which in turns forwards it to the Receiver-3. The vEdge-2 replicator makes one copy of the green stream, which it forwards to vEdge-5 (from which it goes on to Receiver-4), and it makes two copies of the blue stream, which it forward to vEdge-4 and vEdge-5 (and which they then forward to the two receivers).
Now, let’s look at the multicast configurations on the five vEdge routers:

- vEdge router vEdge-1 is a PIM replicator for a particular VPN. If we assume that no multicast sources, receivers, or RPs are located in its local network, the configuration of this router is simple: In the VPN, enable the replicator functionality, with the `router multicast-replicator local` command, and enable PIM, with the `router pim` command.

- vEdge router vEdge-2 also acts only as a replicator in the same VPN as vEdge-1, and you configure it with the same commands, `router multicast-replicator local` and `router pim`, when configuring the VPN. Each replicator can accept a maximum number of new PIM joins, and when this threshold value is reached, all new joins are sent to the second replicator. (If there is only one replicator, new joins exceeding the threshold are dropped.)

- vEdge router vEdge-4 runs PIM. You enable PIM explicitly on the service side within a VPN, specifying the service-side interface that connects to the multicast domain in the local network. So within the VPN, you include the `router pim interface` command. You can also enable auto-RP with the `router pim auto-rp` command.

On the transport side, no explicit configuration is required. The vEdge router automatically directs multicast traffic—both OMP control plane messages and multicast streams—to VPN 0, which is the WAN transport VPN.
vEdge router vEdge-5 is also configured to run PIM in the same way as vEdge-4: You configure the service-side interface name and RP information.

On all five of these vEdge routers, PIM must be enabled in the same VPN so that the multicast streams can be transmitted and received.

Additional Information

Configuring Multicast Overlay Routing