CHAPTER 7

Multicast VPN

Multicast is a popular feature used mainly by IP-networks of Enterprise customers. Multicast allows the efficient distribution of information between a single multicast source and multiple receivers. An example of a multicast source in a corporate network would be a financial information server provided by a third-party company such as Bloomberg's or Reuters. The receivers would be individual PCs scattered around the network all receiving the same financial information from the server. The multicast feature allows a single stream of information to be transmitted from a source device, regardless of how many receivers are active for the information from that source device. The routers automatically replicate a single copy of the stream to each interface where multicast receivers can be reached. Therefore, multicast significantly reduces the amount of traffic required to distribute information to many interested parties.

This chapter describes in detail how an MPLS VPN service provider can provide multicast services between multiple sites of a customer VPN that has an existing multicast network or is intending to deploy the multicast feature within their network. This feature is known as *multicast VPN* (mVPN) and is available from Cisco IOS 12.2(13)T onward. This chapter includes an introduction to general IP Multicast concepts, an overall description of the mVPN feature and architecture, a detailed description of each IP Multicast component modified to support the mVPN feature, and a case study that shows how you can implement mVPN in an MPLS VPN backbone.

Introduction to IP Multicast

IP multicast is an efficient mechanism for transmitting data from a single source to many receivers in a network. The destination address of a multicast packet is always a multicast group address. This address comes from the IANA block 224.0.0.0–239.255.255.255. (Before the concept of classless interdomain routing, or CIDR, existed, this range was referred to as the D-class.) A source transmits a multicast packet by using a multicast group address, while many receivers "listen" for traffic from that same group address.

Examples of applications that would use multicast are audio/video services such as IPTV, Windows Media Player, conferencing services such as NetMeeting or stock tickers, and financial information such as those that TIBCO and Reuters provide.

NOTE If you want to gain a more complete or detailed understanding of IP multicast, then read the Cisco Press book titled *Developing IP Multicast Networks* (ISBN 1-57870-077-9) or any other book that provides an overview of multicast technologies. You can obtain further information on advanced multicast topics from http://www.cisco.com/go/ipmulticast.

Multicast packets are forwarded through the network by using a *multicast distribution tree*. The network is responsible for replicating the same packet at each *bifurcation point* (the point at which the branches fork) in the tree. This means that only one copy of the packet travels over any particular link in the network, making multicast trees extremely efficient for distributing the same information to many receivers.

There are two types of distribution trees: source trees and shared trees.

Source Trees

A *source tree* is the simplest form of distribution tree. The source host of the multicast traffic is located at the root of the tree, and the receivers are located at the ends of the branches. Multicast traffic travels from the source host down the tree toward the receivers. The forwarding decision on which interface a multicast packet should be transmitted out is based on the multicast forwarding table. This table consists of a series of multicast state entries that are cached in the router. State entries for a source tree use the notation (S, G) pronounced *S comma G*. The letter *S* represents the IP address of the source, and *G* represents the group address.

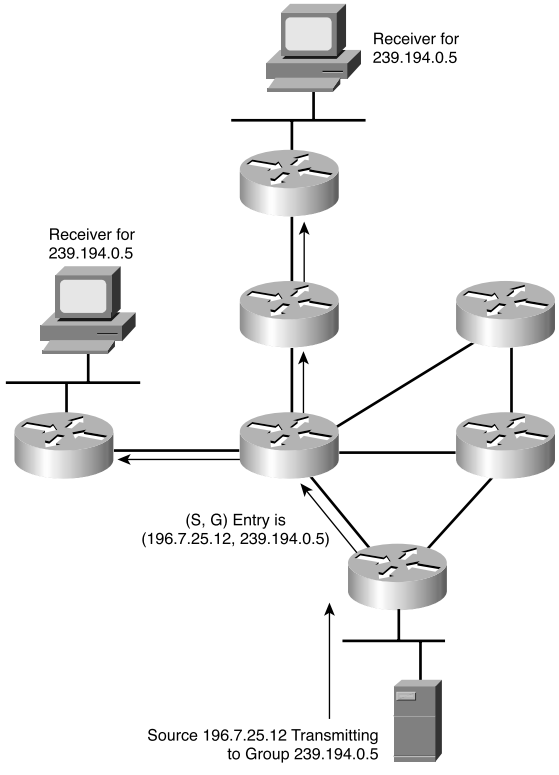
NOTE The notion of direction is used for packets that are traveling along a distribution tree. When a packet travels from a source (or root) toward a receiver, it is deemed to be traveling *down the tree.* If a packet is traveling from the receiver toward the source (such as a control packet), it is deemed to be traveling *up the tree.*

A source tree is depicted in Figure 7-1. The host 196.7.25.12 at the root of the tree is transmitting multicast packets to the destination group 239.194.0.5, of which there are two interested receivers. The forwarding cache entry for this multicast stream is (196.7.25.12, 239.194.0.5).

A source tree implies that the route between the multicast source and receivers is the shortest available path; therefore, source trees are also referred to as *shortest path trees* (SPTs). A separate source tree exists for every source that is transmitting multicast packets, even if those sources are transmitting data to the same group. This means that there will be an (S, G) forwarding state entry for every active source in the network. Referring to our

earlier example, if another source, such as 196.7.25.18, became active that was also transmitting to group 239.194.0.5, then an additional state entry (and a different SPT) would be created as (196.7.25.18, 239.194.0.5). Therefore, source trees or SPTs provide optimal routing at the cost of additional multicast state information in the network.

Figure 7-1 *Source Distribution Tree*



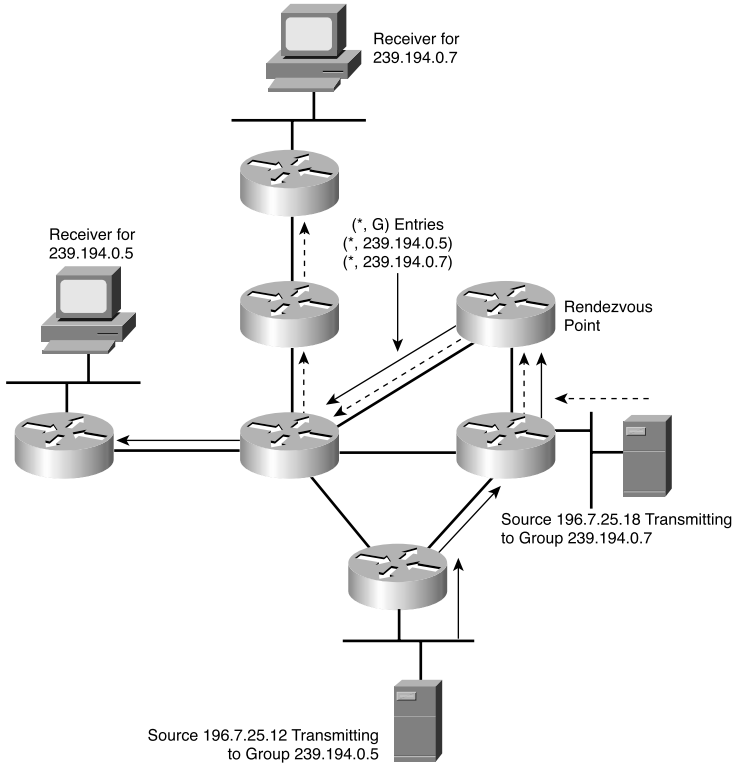
The important thing to remember about source trees is that the receiving end can only join the source tree if it has knowledge of the IP address of the source that is transmitting the group in which it is interested. In other words, to join a source tree, an explicit (S, G) join must be issued from the receiving end. (This explicit [S, G] join is issued by the last hop router, not the receiving host. The receiving host makes the last hop router aware that it wants to receive data from a particular group, and the last hop router figures out the rest.)

Shared Trees

Shared trees differ from source trees in that the root of the tree is a common point somewhere in the network. This common point is referred to as the *rendezvous point (RP)*. The RP is the point at which receivers join to learn of active sources. Multicast sources must transmit their traffic to the RP. When receivers join a multicast group on a shared tree, the root of the tree is always the RP, and multicast traffic is transmitted from the RP down toward the receivers. Therefore, the RP acts as a go-between for the sources and receivers. An RP can be the root for all multicast groups in the network, or different ranges of multicast groups can be associated with different RPs.

Multicast forwarding entries for a shared tree use the notation (*, G), which is pronounced *star comma G*. This is because all sources for a particular group share the same tree. (The multicast groups go to the same RP.) Therefore, the * or wildcard represents all sources. A shared tree is depicted in Figure 7-2. In this example, multicast traffic from the source host 196.7.25.18 and 196.7.25.12 travel to the RP and then down the tree toward the two receivers. There are two routing entries, one for each of the multicast groups that share the tree: (*, 239.194.0.5) and (*, 239.194.0.7). In a shared tree, if more sources become active for either of these two groups, there will still be only two routing entries due to the wildcard representing all sources for that group.

Figure 7-2 *Shared Distribution Tree*



Shared trees are not as optimal in their routing as source trees because all traffic from sources must travel to the RP and then follow the same (*, G) path to receivers. However, the amount of multicast routing state information required is less than that of a source tree. Therefore, there is a trade-off between optimal routing versus the amount of state information that must be kept.

Shared trees allow the receiving end to obtain data from a multicast group without having to know the IP address of the source. The only IP address that needs to be known is that of the RP. This can be configured statically on each router or learned dynamically by mechanisms such as Auto-RP or Bootstrap Router (BSR).

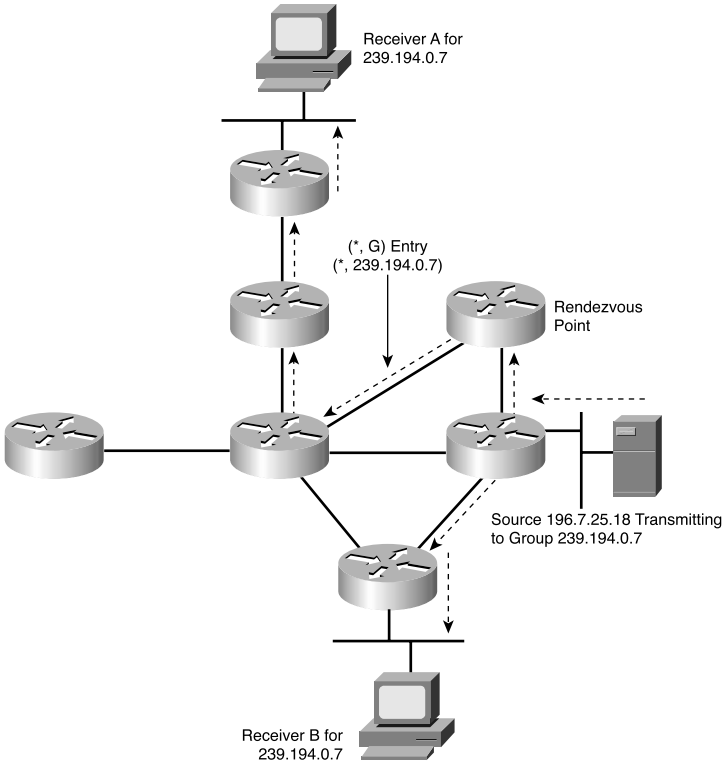
Shared trees can be categorized into two types: unidirectional and bidirectional. Unidirectional trees are essentially what has already been discussed; sources transmit to the RP, which then forwards the multicast traffic down the tree toward the receivers.

In a bidirectional shared tree, multicast traffic can travel up and down the tree to reach receivers. Bidirectional shared trees are useful in an any-to-any environment, where many sources and receivers are evenly distributed throughout the network. Figure 7-3 shows a bidirectional tree. Source 196.7.25.18 is transmitting to two receivers A and B for group 239.194.0.7. The multicast traffic from the source host is forwarded in both directions as follows:

- Up the tree toward the root (RP). When the traffic arrives at the RP, it is then transmitted down the tree toward receiver A.
- Down the tree toward receiver B. (It does not need to pass the RP.)

Bidirectional trees offer improved routing optimality over unidirectional shared trees by being able to forward data in both directions while retaining a minimum amount of state information. (Remember, state information refers to the amount of (S, G) or (*, G) entries that a router must hold.)

Figure 7-3 *Bidirectional Shared Tree*



Multicast Forwarding

Packet forwarding in a router can be divided into two types: unicast forwarding and multicast forwarding. The difference between unicast forwarding and multicast forwarding can be summarized as follows:

- Unicast forwarding is concerned with where the packet is going.
- Multicast forwarding is concerned with where the packet came from.

In unicast routing, the forwarding decision is based on the destination address of the packet. At each router along the path, you can derive the next-hop for the destination by finding the longest match entry for that destination in the unicast routing table. The unicast packet is then forwarded out the interface that is associated with the next-hop.

Forwarding of multicast packets cannot be done in the same manner because the destination is a multicast group address that you will most likely need to forward out multiple interfaces. Multicast group addresses do not appear in the unicast routing table; therefore,

forwarding of multicast packets requires a different process. This process is called *Reverse Path Forwarding* (RPF), and it is the basis for forwarding multicast packets in most multicast routing protocols. In particular, RPF is used with Protocol Independent Multicast (PIM), which is the protocol used and described throughout this chapter.

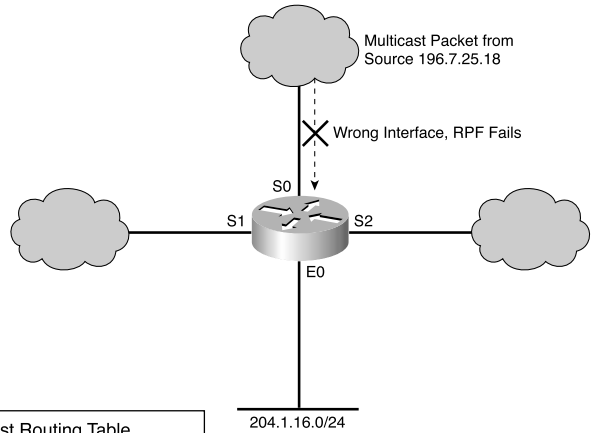
RPF

Every multicast packet received on an interface at a router is subject to an RPF check. The RPF check determines whether the packet is forwarded or dropped and prevents looping of packets in the network. RPF operates like this:

- When a multicast packet arrives at the router, the *source* address of that packet is checked to make sure that the incoming interface indeed leads back to the source. (In other words, it is on the reverse path.)
- If the check passes, then the multicast packet is forwarded out the relevant interfaces (but not the RPF interface).
- If the RPF check fails, the packet is discarded.

The interface used for the RPF check is referred to as the *RPF interface*. The way that this interface is determined depends on the multicast routing protocol that is in use. This chapter is concerned only with PIM, which is the most widely used protocol in Enterprise networks. PIM is discussed in the next section. PIM uses the information in the unicast routing table to determine the RPF interface. Figure 7-4 shows the process of an RPF check for a packet that arrives on the wrong interface. A multicast packet from the source 196.7.25.18 arrives on interface S0. A check of the multicast routing table shows that network 196.7.25.0 is reachable on interface S1, not S0; therefore, the RPF check fails and the packet is dropped.

Figure 7-4 *RPF Check Fails*

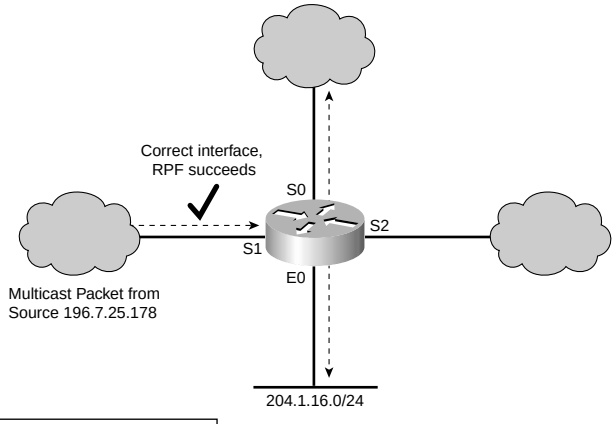


Unicast Routing Table	
Network	Interface
196.7.25.0/24	S1
196.7.26.0/24	S0
204.1.16.0/24	E0

Packet arrived on wrong interface, therefore packet should be dropped.

Figure 7-5 shows the RPF check for a multicast packet that arrives on the correct interface. The multicast packet with source arrives on interface S1, which matches the interface for this network in the unicast routing table. Therefore, the RPF check passes and the multicast packet is replicated to the interfaces in the outgoing interface list (called the *olist*) for the multicast group.

Figure 7-5 *RPF Check Succeeds*



Unicast Routing Table	
Network	Interface
196.7.25.0/24	S1
196.7.26.0/24	S0
204.1.16.0/24	E0

Packet arrived on correct interface, therefore packet should be replicated.

If the RPF check has to refer to the unicast routing table for each arriving multicast packet, this will have a detrimental affect on router performance. Instead, the RPF interface is cached as part of the (S, G) or (*, G) multicast forwarding entry. When the multicast forwarding entry is created, the RPF interface is set to the interface that leads to the source network in the unicast routing table. If the unicast routing table changes, then the RPF interface is updated automatically to reflect the change.

Example 7-1 shows a multicast forwarding entry for (194.22.15.2, 239.192.20.16). You can also refer to this entry as a *multicast routing table entry.* The presence of the source in the (S, G) notation indicates that this entry is associated with a source tree or shortest path tree. The incoming interface is the RPF interface, which has been set to POS3/0. This setting matches the next-hop interface shown in the OSPF routing entry for the source 194.22.15.2. There are two interfaces in the outgoing olist: Serial4/0 and Serial4/2. The outgoing interface list provides the interfaces that the multicast packet should be replicated out. Therefore, packets that pass the RPF check from source 194.22.15.2 (they must come in on

POS3/0) that are destined to group 239.192.20.16 are replicated out interface Serial4/0 and Serial4/2.

Example 7-1 *Source Tree Multicast Forwarding Entry*

```
(194.22.15.2, 239.192.20.16), 00:03:30/00:03:27, flags: sT
  Incoming interface: POS3/0, RPF nbr 194.22.15.17
  Outgoing interface list:
    Serial4/0, Forward/Sparse-Dense, 00:03:30/00:02:55
    Serial4/2, Forward/Sparse-Dense, 00:02:45/00:02:05

Routing entry for 194.22.15.2/32
  Known via "ospf 1", distance 110, metric 2, type intra area
  Last update from 194.22.15.17 on POS3/0, 1w5d ago
  Routing Descriptor Blocks:
  * 194.22.15.17, from 194.22.15.2, 1w5d ago, via POS3/0
      Route metric is 2, traffic share count is 1
```

For completeness, a shared tree routing entry is shown in Example 7-2. This entry represents all sources transmitting to group 239.255.0.20. The RPF interface is shown to be FastEthernet0/1, which is the next-hop interface to the RP 196.7.25.1. Remember that the root of a shared tree are always the RP; therefore, the RPF interface for a shared tree is the reverse path back to the RP.

Example 7-2 *Shared Tree Multicast Forwarding Entry*

```
(*, 239.255.0.20), 2w5d/00:00:00, RP 196.7.25.1, flags: SJCL
  Incoming interface: FastEthernet0/1, RPF nbr 192.168.2.34
  Outgoing interface list:
    FastEthernet0/0, Forward/Sparse, 00:03:29/00:02:54
```

The outgoing interface lists in the preceding examples are determined by the particular multicast protocol in use.

PIM

Over the years, various multicast protocols have been developed, such as Distance Vector Multicast Routing Protocol (DVMRP), Multicast Open Shortest Path First (MOSPF), and Core Base Trees (CBT). The characteristic that these protocols have in common is that they create a multicast routing table based on their own discovery mechanisms. The RPF check does not use the information already available in the unicast routing table.

The protocol that is the most widely deployed and relevant to this chapter is PIM. As discussed previously, PIM uses the unicast routing table to discover whether the multicast packet has arrived on the correct interface. The RPF check is independent because it does not rely on a particular protocol; it bases its decisions on the contents of the unicast routing table.

Several PIM modes are available: dense mode (PIM DM), sparse mode (PIM SM), Bidirectional PIM (PIM Bi-Dir), and a recent addition known as Source Specific Multicast (SSM).

PIM DM

The deployment of PIM DM is diminishing because it has been proven to be inefficient in comparison to PIM SM. PIM DM is based on the assumption that for every subnet in the network, at least one receiver exists for every (S, G) multicast stream. Therefore, all multicast packets are pushed or flooded to every part of the network. Routers that do not want to receive the multicast traffic because they do not have a receiver for that (S, G) send a prune message back up the tree. Branches that do not have receivers are pruned off, the result being a source distribution tree with branches that have receivers. Periodically, the prune message times out, and multicast traffic begins to flood through the network again until another prune is received.

PIM SM

PIM SM is more efficient than PIM DM in that it does not use flooding to distribute traffic. PIM SM employs the pull model, in which traffic is distributed only where is it requested. Multicast traffic is distributed to a branch only if an explicit join message has been received for that multicast group. Initially, receivers in a PIM SM network join the shared tree (rooted at the RP). If the traffic on the shared tree reaches a certain bandwidth threshold, the last hop router (that is, the one to which the receiver is connected) can choose to join a shortest-path tree to the source. This puts the receiver on a more optimal path to the source.

PIM Bi-Dir

PIM Bi-Dir creates a two-way forwarding tree, as shown in Figure 7-3. All multicast routing entries for bidirectional groups are on a (*, G) shared tree. Because traffic can travel in both directions, the amount of state information is kept to a minimum. Routing optimality is improved because traffic does not have to travel unnecessarily toward the RP. Source trees are never built for bidirectional multicast groups. Bidirectional trees in the service provider network are covered in the section “Case Study of mVPN Operation in SuperCom” later in this chapter.

SSM

SSM implies that the IP address of the source for a particular group is known before a join is issued. SSM in Cisco IOS is implemented in addition to PIM SM and co-exists with IP Multicast networks based on PIM SM. SSM always builds a source tree between the receivers and the source. The source is learned through an out-of-band mechanism. Because the source is known, an explicit (S, G) join can be issued for the source tree that obviates the need for shared trees and RPs. Because no RPs are required, optimal routing is assured; traffic travels the most direct path between source and receiver. SSM is a recent innovation in multicast networks and is recommended for new deployments, particularly in the service provider core for an mVPN environment. A practical deployment of SSM is discussed in the section, “Case Study of mVPN Operation in SuperCom” later in this chapter.

Multicast is a powerful feature that allows the efficient one-to-many distribution of information. Multicast uses the concept of distribution trees, where the source is the root of the tree and the receivers are at the leaves of the tree. The routers replicate packets at each branch of the tree, known as the bifurcation point. The tree is represented as a series of multicast state entries in each router, and packets are forwarded down this tree (toward the leaves) by using RPF. There are various modes of multicast operation in networks with the most popular one being PIM SM.

Enterprise Multicast in a Service Provider Environment

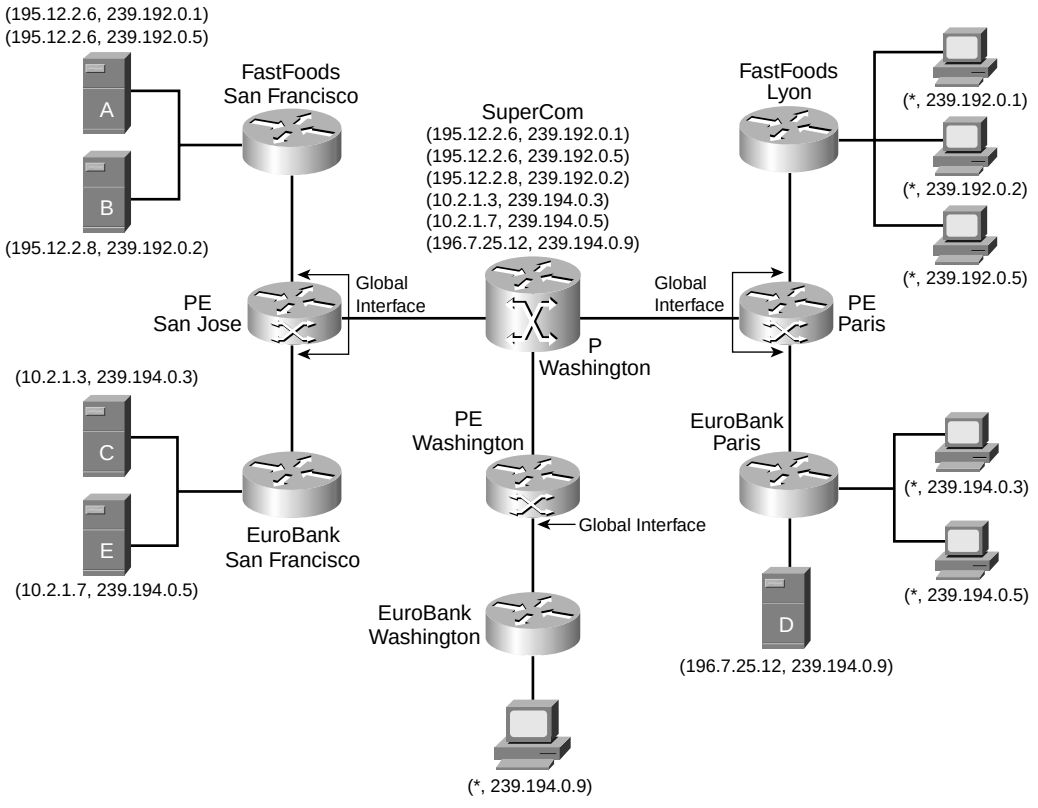
The fundamental problem that service providers face today when offering native multicast services to end customers is the amount of multicast distribution information (that is [S, G] or [*, G] states) that needs to be maintained to provide the most optimal multicast traffic distribution. When a multicast source becomes active within a particular customer site, the multicast traffic must travel through the service provider network to reach all PE routers that have receivers connected to CE routers for that multicast group. To prevent unnecessary traffic delivery, the service provider must avoid sending traffic to PE routers that have no interested receivers. To accomplish this goal and achieve optimal routing, each P router in the network must maintain state information for all active customer distribution trees.

However, a problem arises in that the service provider has no visibility into how its end customers manage multicast within their enterprise. In addition, the service provider does not have control over the distribution of sources and receivers or the number of groups that the end customer chooses to use. In this situation, the P routers are required to support an unbounded amount of state information based on the enterprise customer's application of multicast.

Figure 7-6 illustrates this scenario in the SuperCom network. (This chapter uses SuperCom as the example network.) As shown in the figure, SuperCom provides native multicast services to VPN customers FastFoods and EuroBank. In this example, native multicast means that the SuperCom network provides both customers with multicast services via the global multicast routing table by using standard multicast procedures. To obtain multicast services, each EuroBank or FastFoods site must ultimately connect to a SuperCom global interface (that is, one with no VRF defined). Multicast traffic travels across the SuperCom network using standard IP multicast; no tunnels or encapsulations are used. The FastFoods organization has three active distribution trees rooted at two sources (A and B). Similarly, EuroBank has three active distribution trees rooted at three sources (C, D, and E). Each of these trees has at least one receiver that is connected to a CE somewhere in the global multicast network.

To provide optimal multicast traffic distribution, the Washington P router must hold the state information for all six trees. This applies equally to any other P and PE routers that are in the path of the distribution trees. Because all multicast routing operates in the global SuperCom table, it is possible that multicast groups that different customers use will conflict (as would be the case with multiple customers using the same RFC 1918 addressing in a unicast network). To avoid this situation, SuperCom must allocate each VPN a unique range of multicast groups.

Figure 7-6 *Supporting Native Enterprise Multicast*



The total amount of state information that the SuperCom network must hold is determined by the way the customer deploys multicast in his network. For each unique customer source, a separate state entry exists in the global table for each multicast group that source is serving. Deploying features such as bidirectional trees reduces the amount of multicast state information required, although traffic distribution is not as optimal. Given that the amount of state information is unbounded (cannot be limited) and the service provider must allocate and manage multicast groups, the deployment of native multicast services in this manner is not recommended from a scaling and provisioning standpoint.

A common way to provide multicast over a service provider IP or MPLS VPN network is to overlay generic routing encapsulation (GRE) tunnels between CE routers. This eliminates the need for any state information to be kept in the P routers because all multicast packets between VPN sites are encapsulated by using GRE within the service provider network. This solution also allows different enterprise customers to use overlapping multicast groups. However, the disadvantage of this solution is that unless the customer implements a full mesh of GRE tunnels between CE routers, optimal multicast routing is

not achieved. In fact, more bandwidth can be wasted by multicast traffic backtracking over different GRE tunnels across the P network. Further to this, Multicast over GRE is inherently unscalable due to the potential number of tunnels required and the amount of operational and management overhead.

A more scalable model for providing multicast within a VPN can be derived from the way optimal unicast routing is achieved in an MPLS VPN.

In an MPLS VPN

- A P router maintains routing information and labels for the global routing table only. It does not hold routing or state information for customer VPNs.
- A CE router maintains a routing adjacency with its PE router neighbor only. CE routers do not peer with other CE routers but still have the ability to access other CE routers in their VPN through the most optimal route that the P network provides.

As you will see, the mVPN solution that is implemented in Cisco IOS provides a scalable and efficient method of transporting multicast traffic between sites of a VPN customer and has similar characteristics mentioned in the previous bullet points.

In a service provider network that is enabled with mVPN

- A P router maintains multicast state entries for the global routing table only. It does not hold multicast state entries for customer VPNs.
- A CE router maintains a multicast PIM adjacency with its PE router neighbor only. CE routers do not have multicast peerings with other CE routers, but they can exchange multicast information with other CE routers in the same VPN.

The following sections describe the mVPN architecture as implemented by Cisco IOS.

mVPN Architecture

The mVPN solution discussed in this chapter is based on section 2 of *Multicast in MPLS/BGP VPNs* Internet draft (draft-rosen-vpn-mcast).

Section 2 of this Internet draft describes the concept of *multicast domains* in which CE routers maintain a PIM adjacency with their local PE router only, and not with other CE routers. As mentioned previously, this adjacency characteristic is identical to that used in MPLS VPNs. Enterprise customers can maintain their existing multicast configurations, such as PIM SM/PIM DM and any RP discovery mechanisms, and they can transition to an mVPN service by using multicast domains without configuration changes. P routers do not hold state information for individual customer source trees; instead, they can hold as little as a single state entry for each VPN (assuming that PIM Bi-Dir is deployed) regardless of the number of multicast groups within that VPN.

If a service provider is using PIM SM in the core (instead of PIM Bi-Dir), then the greatest amount of state information that would be required in a P router would be roughly equivalent to the number of PE routers in the backbone multiplied by the number of VPNs

defined on those PE routers. This should be significantly less than the number of potential customer multicast groups. Although you can reduce the amount of P-network state information, the real point to note here is that with multicast domains, regardless of which multicast mode the service provider is using (PIM SM, Bi-Dir, SSM), the amount of state information in the core is deterministic. The core information does not depend on the customer's multicast deployment.

Customer networks are also free to use whatever multicast groups they need without the possibility of overlap with other VPNs. These groups are invisible to the P router network, in the same manner that VPN unicast routes are invisible to P routers in an MPLS VPN network.

Multicast Domain Overview

A multicast domain is a set of multicast-enabled virtual routing and forwarding instances (VRFs) that can send multicast traffic to each other. These multicast VRFs are referred to as mVRFs. Multicast domains map all of a customer's multicast groups that exist in a particular VPN to a single unique global multicast group in the P-network. This is achieved by encapsulating the original customer multicast packets within a provider packet by using GRE. The destination address of the GRE packet is the unique multicast group that the service provider has allocated for that multicast domain. The source address of the GRE packet is the BGP peering address of the originating PE router. A different global multicast group address is required for every multicast domain. Therefore, the set of all customer multicast states (*, G^1)…(*, G^N) can be mapped to a single (S, G) or (*, G) in the service provider network.

NOTE The use of GRE in a multicast domain is not the same as the overlay solution in which point-to-point GRE tunnels are used between CE routers. The GRE tunnels used here are between PE routers in a multicast configuration. The tunnels can be considered point-to-multipoint connections if PIM SM is deployed or even multipoint-to-multipoint if using PIM Bi-Dir. Therefore, the use of GRE for multicast domains is inherently more efficient than GRE overlay.

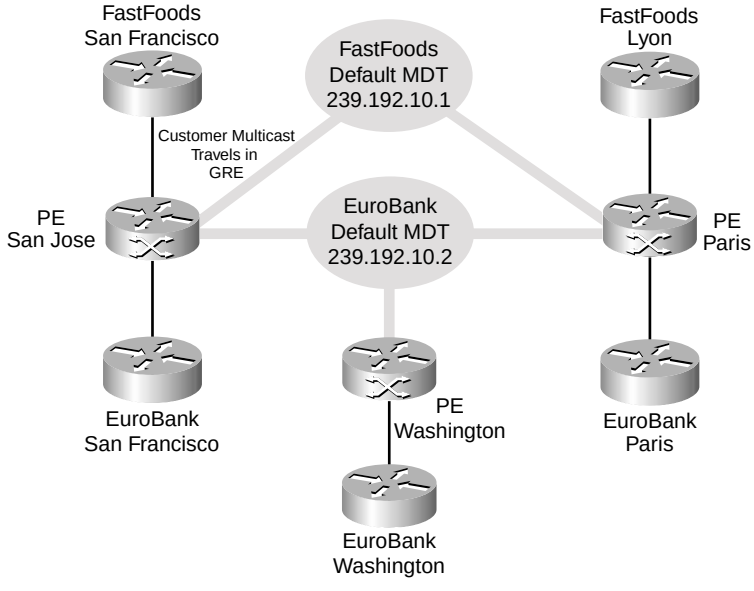
Each PE router that is supporting an mVPN customer is part of the multicast domain for that customer. Multiple end customers can attach to a particular PE router, which means that the PE router can be a member of many multicast domains—one for each mVPN customer who is connected to it.

One of the major attractions of the multicast domain solution is that the P routers do not need a software upgrade to enable new multicast features to support mVPNs. Only native multicast is required in the core network to support multicast domains. The advantage of this is that native multicast is a mature technology in Cisco IOS; therefore, the operational risk is minimized in the service provider network when deploying multicast domains.

The P-network builds a default multicast distribution tree (Default-MDT) between PE routers for each multicast domain by using a unique multicast group address that the service provider allocates. These unique multicast groups are referred to as *MDT-Groups.* Each mVRF belongs to a default MDT. Therefore, the amount of state information that a P router must hold is not a function of the number of customer multicast groups in the network; instead, it is the number of VPNs. This considerably reduces the amount of state information required in a P router. If the MDT is configured as a bidirectional tree, then it is possible to have a single (*, G) multicast state entry for each VPN.

Figure 7-7 shows the concept of multicast domains in the SuperCom network. The FastFoods and EuroBank VPNs belong to separate multicast domains. The SuperCom core creates a Default-MDT for each of these multicast domains by using the MDT-group addresses 239.192.10.1 for FastFoods and 239.192.10.2 for EuroBank. The PE routers at San Jose and Paris join both Default-MDTs as they are connected to the FastFoods and EuroBank sites. The Washington PE router only needs to connect to the Default-MDT for the EuroBank VPN.

Figure 7-7 *Multicast Domains*



Any EuroBank or FastFoods packets that travel along these Default-MDTs are encapsulated by using GRE. The source of the outer packet is the Multiprotocol BGP peering address of the sending PE router, and the destination is the appropriate MDT-group address. GRE essentially hides the customer multicast packet from the P-network and allows us to map many multicast groups in a VPN to a single provider multicast group. The SuperCom P routers only see the source and destination of the outer IP header that SuperCom allocates. This source and destination appear as an (S, G) state entry in the SuperCom global multicast table.

Assuming that the SuperCom network has been configured with PIM Bi-Dir, only two (*, G) states are required in each P router: (*, 239.192.10.1) and (*, 239.192.10.2). This compares favorably with the six states required in the native multicast network described earlier in Figure 7-6. Also note in our example that the amount of state information in the P-network is always bounded to two entries regardless of how many new sources and groups FastFoods or EuroBank introduce.

NOTE A P router is only aware of the PE router source addresses and the MDT-Group addresses that form the MDTs. CE router traffic traveling along an MDT is forwarded in a GRE-encapsulated packet (P-packet) using the MDT-group address as the destination (more on this in the later section, "MDTs"). The GRE P-packet uses IP only, and no MPLS label headers are applied to MDT traffic**.** Only pure IP multicast exists in the core.

NOTE mVPN will be supported from IOS versions 12.2(13)T and 12.0(23)S for Cisco 7200 and 7500 series routers. Support for Cisco 10000 series routers will be available from IOS version 12.0(23)SX, Cisco 12000 series is suppored in 12.0(26)S. The initial release will permit a VPN to participate only in a single multicast domain; access to Internet multicast or other multicast domains will not be permissible. However, it is expected that this limitation will be removed in future versions of IOS.

PIM SM or SSM are the only multicast modes supported in the P-network for mVPN.

To summarize, the goals of the multicast domain solution are as follows:

- To deliver Enterprise Multicast to customers who subscribe to an MPLS VPN service
- To minimize the amount of state information in the P-network (the service provider core) while providing optimal routing
- To allow customers the freedom to choose their own multicast groups, multicast operations mode, RP placement, and so on
- To allow multicast in the P-network to be completely separated from the operation of multicast in the customer network.

The various components used to deliver multicast domains are explained in the following sections.

Multicast VRF

On a PE router, each VRF can have an associated multicast routing and forwarding table configured, referred to as a multicast VRF (mVRF). The mVRF is the PE router's view into the enterprise VPN multicast network. The mVRF contains all the multicast routing information

for that VPN. This information includes the state entries for distribution trees or RP-to-group mappings (if PIM SM is being used). When a PE router receives multicast data or control packets from a CE router interface in a VRF, multicast routing such as RPF checks and forwarding will be performed on the associated mVRF.

The PE router also can configure multicast features or protocols in the context of the mVRF. For example, if the customer network were using static RP configurations (that is, it was not using Auto-RP to distribute RP information), then the PE router would need to configure the same static RP entry information that was being used in the C-network. The multicast routing protocols in Cisco IOS such as PIM, IGMP, and MSDP have been modified to operate in the context of an mVRF and as such only modify data structures and states within that mVRF.

Example 7-3 shows the PIM and MSDP commands available in the context of a VRF.

Example 7-3 *VRF-Aware Multicast Configuration Commands*

```
SuperCom_Paris(config)#ip pim vrf EuroBank ?
accept-register      Registers accept filter
accept-rp            RP accept filter
bsr-candidate        Candidate bootstrap router (candidate BSR)
register-rate-limit  Rate limit for PIM data registers
register-source      Source address for PIM Register
rp-address           PIM RP-address (Rendezvous Point)
rp-announce-filter   Auto-RP announce message filter
rp-candidate         To be a PIMv2 RP candidate
send-rp-announce     Auto-RP send RP announcement
send-rp-discovery    Auto-RP send RP discovery message (as RP-mapping agent)
spt-threshold        Source-tree switching threshold
ssm                  Configure Source Specific Multicast
state-refresh        PIM DM State-Refresh configuration

SuperCom_Paris(config)#ip msdp vrf EuroBank ?
  default-peer       Default MSDP peer to accept SA messages from
  description        Peer specific description
  filter-sa-request  Filter SA-Requests from peer
  mesh-group         Configure an MSDP mesh-group
  originator-id      Configure MSDP Originator ID
  peer               Configure an MSDP peer
  redistribute       Inject multicast route entries into MSDP
  sa-filter          Filter SA messages from peer
  sa-limit           Configure SA limit for a peer
  sa-request         Configure peer to send SA-Request messages to
  shutdown           Administratively shutdown MSDP peer
  timer              MSDP timer
  ttl-threshold      Configure TTL Thresold for MSDP Peer
```

In addition to the commands in the previous example, there are several multicast **show** commands that support VRF contexts. These are shown in Example 7-4.

Example 7-4 *VRF-Aware Multicast show Commands*

```
SuperCom_Paris#show ip pim vrf EuroBank ?
  autorp      Global AutoRP information
  bsr-router  Bootstrap router (v2)
  interface   PIM interface information
  mdt         Multicast tunnel information
  neighbor    PIM neighbor information
  rp          PIM Rendezvous Point (RP) information
  rp-hash     RP to be chosen based on group selected

SuperCom_Paris#show ip msdp vrf EuroBank ?
  count     SA count per AS
  peer      MSDP Peer Status
  sa-cache  MSDP Source-Active Cache
  summary   MSDP Peer Summary

SuperCom_Paris#show ip igmp vrf EuroBank ?
  groups      IGMP group membership information
  interface   IGMP interface information
  membership  IGMP membership information for forwarding
  tracking    IGMP Explicit Tracking information
  udlr        IGMP undirectional link multicast routing information
```

Example 7-5 shows the commands to enable multicast for the EuroBank VRF. The **ip multicast-routing vrf** enables multicast routing on the associated EuroBank VRF. In addition, any multicast interfaces in the EuroBank VRF will also require PIM to be enabled, as shown with the **ip pim sparse** command. The various PIM adjacencies that can exist are discussed in the following section.

Example 7-5 *Enabling Multicast in a VRF*

```
ip multicast-routing vrf EuroBank
!
interface Serial0/0
 ip vrf forwarding EuroBank
 ip address 192.168.2.26 255.255.255.252
 ip pim sparse
```

NOTE If the **ip vrf forwarding** command is removed from the PE router configuration, not only is the **ip address** command removed from any associated VRFs, but the **ip pim sparse** command is also removed.

PIM Adjacencies

Each VRF that has multicast routing enabled has a single PIM instance created on the PE router. This VRF-specific PIM instance forms a PIM adjacency with each PIM-enabled CE router in that mVRF. The customer multicast routing entries that each PIM instance creates are specific to the corresponding mVRF.

In addition to the CE router PIM adjacency, the PE router forms two other types of PIM adjacencies. The first is a PIM adjacency with other PE routers that have mVRFs in the same multicast domain. This PE router PIM adjacency is accessible through the *multicast tunnel interface* (MTI) and is used to transport multicast information between mVRFs (through a MDT) across the backbone. MDTs and MTIs are described later in this chapter. The PE router PIM adjacencies are maintained by using the same PIM instance that is used between the PE router and CE router for the associated mVRF.

The second type of PIM adjacency is created by the global PIM instance. The PE router maintains global PIM adjacencies with each of its IGP neighbors, which will be P routers, or directly connected PE routers (that are also providing a P router function). The global PIM instance is used to create the multicast distribution trees (MDTs) that connect the mVRFs.

NOTE CE routers do not form PIM adjacencies with each other, nor does a CE router form an adjacency with a PE router by using the global PIM instance.

Figure 7-8 shows the different types of PIM adjacencies in the SuperCom network for the FastFoods VPN. A PIM adjacency exists between the San Francisco FastFoods CE router and San Jose PE router, as well as between the Lyon FastFoods CE router and the Paris PE router. Because the FastFoods mVRFs are part of the same multicast domain, a PIM adjacency is active between the San Jose and Paris PE routers. Both San Jose and Paris PE routers have separate PIM adjacency in the global table to the Washington P router.

Figure 7-8 *PIM Adjacencies*