Technical Guide

Protocol Independent Multicast - Sparse Mode (PIM-SM)

Feature Overview and Configuration Guide

Introduction

This guide provides information about two multicast protocols:

- Protocol Independent Multicast - Sparse Mode (PIM-SM) and
- Protocol Independent Multicast - Source Specific Multicast (PIM-SSM).

Products and software version that apply to this guide

This guide applies to AlliedWare Plus™ products that support PIM-SM and PIM-SSM, running version **5.4.4** or later.

To see whether your product supports PIM-SM and PIM-SSM, see the following documents:

- The product's Datasheet
- The product's Command Reference

These documents are available from the above links on our website at alliedtelesis.com.

Feature support may change in later software versions. For the latest information, see the above documents.

AlliedWare Plus™
OPERATING SYSTEM

Content

PIM-SM

Protocol Independent Multicast-Sparse Mode (PIM-SM) provides efficient communication between members of sparsely distributed multicast groups—the type of groups that are most common in wide-area internetworks.

Characteristics of PIM-SM

PIM Sparse Mode (PIM-SM) is designed on the principle that several hosts wishing to receive a multicast stream does not justify flooding the entire internetwork with periodic multicast traffic. PIM-SM is designed to limit multicast traffic so that only those switches interested in receiving traffic for a particular group receive the traffic.

Switches with directly attached or downstream members of a given group are required to join a Sparse Mode distribution tree by transmitting explicit join messages. If a switch does not become part of the distribution tree for a group, it does not receive multicast traffic addressed to the group. In contrast, Dense Mode multicast routing protocols assume downstream group membership and continue to forward multicast traffic on downstream links until explicit prune messages are received. The default forwarding action of a Sparse Mode multicast routing protocol is to block traffic unless it is explicitly requested, while the default action of the Dense Mode multicast routing protocols is to forward traffic, unless requested not to.

In PIM Sparse Mode, a router will not know the source IP of a stream is it not currently forwarding.

As a result, Sparse Mode routers need a way to find out the source IPs of the streams for which they receive downstream requests. To solve this, Sparse Mode introduces the concept of Rendezvous Points – specific designated routers that receive notification of all streams destined to specific ranges of multicast addresses (or, possibly, all multicast addresses).

In turn, the introduction of Rendezvous Points (RPs) into the protocol requires that the protocol also provide a way that routers find out the identities of the Rendezvous Points.

Hence Sparse Mode has the concept of a Bootstrap Router (BSR) that knows the identities of the RPs and provides this information to all other routers. In addition, the protocol needs a process by which the RPs are notified of new streams, and a process whereby a router, having learnt the source IP of a stream from the RP, then accesses the stream via a direct path from the source.

A Rendezvous Point (RP), is a point where receivers “meet” sources. For each multicast group, there is only a single active RP. Each receiver wishing to join a multicast group contacts its directly attached switch, which in turn joins the multicast distribution tree by sending an explicit join message to the group’s RP. The DR on the subnet containing a source, tunnels the stream to the RP to inform the RP that the stream is available. This model requires Sparse Mode switches to maintain some state information (the RP-list) prior to the arrival of data packets. In contrast, Dense Mode multicast routing protocols are data driven, since they do not define any state for a multicast group until the first data packet arrives.

Roles in PIM-SM

A multicast sender does not need to know the addresses of the members of the group in order to send to them, and the members of the group need not know the address of the sender. Group membership can change at any time. When PIM is enabled on the switch, and before the switch can route multicast traffic, it must establish which of the PIM routers in the network are performing some key roles:

- Designated Router
- Rendezvous Point
- Bootstrap Router

Designated Router

There must be one PIM Designated Router (DR) in each subnet in the network. Any PIM-SM interfaces on the subnet elect the DR with the highest DR priority. If there is more than one router with the same priority, or no priority, they choose the interface with the highest IP address. The DR performs all the PIM functionality for the subnet. If the current DR becomes unavailable, the remaining switches elect a new DR for the subnet by DR priority or IP address.

Rendezvous Point

Each multicast group must have a Rendezvous Point (RP). The DR on the subnet containing a multicast source sends multicast packets towards the RP. DRs with group members connected to them send join messages towards the group's RP. The RP for a given group or range of groups is found by an election process. A number of routers in the network may be configured to offer themselves as candidates for the role of RP for a Group or range of Groups. The RP candidate with the lowest priority is elected from all the RP candidates for a group or group range. If the RP becomes unavailable, a new RP is selected from among the remaining candidates.

Bootstrap Router

Again, the Bootstrap Router for a network is chosen by election. A number of routers in the network can be configured to be candidates for the BSR role. Each PIM-SM network must have at least one Bootstrap Router (BSR) candidate, unless all switches in the domain are configured statically with information about all RPs in the domain. Every switch that is a BSR candidate periodically sends a Bootstrap Candidate Advertisement message to advertise that it is available as a BSR candidate. The BSR candidates in the network elect the switch with the highest preference value to be the BSR. The elected BSR listens to PIM Candidate RP Advertisement messages specifying RP candidates for multicast groups. It maintains a list of RP candidates and sends a bootstrap message every BSM interval, specifying all the multicast groups in the PIM network, and their RP candidates. Each switch uses this information and a standardized selection algorithm to determine the RP for each group.

In summary:

- Each multicast group must have at least one RP candidate
- Each PIM-SM domain must have at least one BSR candidate, unless all routers in the domain are configured statically with information about all RPs in the domain
- Each subnet must have at least one DR candidate.

Operation of PIM-SM

To understand how paths are established through a PIM network, from sources to listeners, it is useful to look at the process from the two ends. First, let us look how the request from the listener is delivered to the RP. Then let us look at how the stream from the source is registered with the RP. Then, finally, we will see how, once the requesting host and the stream have met up with each other at the RP, the network then can discover the most direct route between each other.

Requests for streams

Requests can come in the form of IGMP Reports from directly connected hosts or in the form of PIM Joins from downstream neighbors. The router acts the same way in both cases.

If the router receives a new request for a group that it is already forwarding to one or more interfaces, then it simply adds a new downstream interface to the list of egress interfaces. There is no need to signal anything upstream or to send an ACK to the requester; it simply starts forwarding the stream out the port the request arrived on.

If the router was not already forwarding the requested stream, then there is rather more involved:

- The router does not know what source device (if any) might be transmitting the requested stream, so it cannot simply send a request towards that device to ask for a copy of the stream. Instead, it has to rely on an RP to provide the stream. By looking up its Group-to-RP mapping (that it has learnt from the BSR) the router knows the address of the RP that should be able to provide the requested stream.
- The router sends a Join request off towards the RP. The request is not unicast to the RP, it is sent as a multicast towards the neighbor that is the next hop on the path towards the RP. This neighbor, in turn, will forward the request to its next-hop neighbor on the path towards the RP. The request will continue to be forwarded from neighbor to neighbor in this way until either it arrives at a router that is already receiving the requested stream, or it reaches the RP.
- If the request reaches a router that is already receiving the stream in question, before getting to the RP, then this router will start sending the stream down in the direction from which the Join arrived, and will not forward the Join request any further upstream.
- If the request ends up going all the way to the RP, then the RP will start forwarding the stream down in the direction from which the Join request arrived. If the RP has never actually received the requested stream, then it will simply do nothing—there is no mechanism in PIM for a router to say “sorry, that stream does not exist”.

Once the requesting router receives the stream, it will see the source IP of the stream. At that point, it can take the opportunity to send a request directly towards the source, rather than having to receive the stream via the RP. The RFC does not specify how long after starting to receive the stream that a router should start sending Join requests directly towards the source. However, AlliedWare Plus, like most implementations, starts immediately. As soon as it receives a packet in the stream, it learns the source IP of the stream, and then starts sending requests directly towards the source.

Soon enough, the stream will start arriving along a path directly from the source. At that point, the router prunes itself from the stream that is arriving from the RP. At that point, the router's forwarding of the stream has reached its final steady state.

What happens to a stream?

Now, let's cut over to an overview of how a stream gets treated by PIM-SM.

The treatment of a stream

A source simply sends a stream. It has no idea which devices (if any) will receive the stream. It is the routers in the network that pick up the stream, and forward it to where it needs to go. The router nearest the source, known as the first-hop router, has the responsibility of making sure that the RP for a given stream gets a copy of that stream.

The first-hop router, like all routers in the network, needs to learn a group-to-RP mapping from the BSR. Having learnt that mapping, it knows the address of the RP to which it must forward any given stream.

Of course, the first-hop router can't forward the stream to the RP by multicast, as the routers in between will not know to forward the stream. Instead, the first-hop router has to encapsulate the packets of the stream with a unicast header, and unicast them to the RP.

This is a process known as Registering the stream with the RP.

Once the RP starts receiving the stream by this unicast tunneling method, it has two choices of how to proceed:

1. If it has had requests for this stream, then it will unencapsulate the packets, and forward them out the interface(s) on which it has received the Join requests. At the same time, it will send PIM Join requests upstream towards the source of the stream, to establish a path that can deliver the stream by multicast. Once the Joins have arrived at the first-hop router, and that router sends the stream by multicast, the RP requests that the first-hop router stop tunneling the stream in unicast register packets.
2. If the RP has no currently active requests for the stream, it simply signals to the first-hop router to stop sending the register packets for the stream.

Multi-Access LANs

If the PIM-SM network includes multi-access LAN links for transit, as well as point-to-point links, then a mechanism is needed to prevent multiple trees forwarding the same data to the same group member. Two or more switches on a LAN may have different information about how to reach the RP or the multicast sender. They could each send a join message to two different switches closer to the RP for an RPT or the sender for an SPT. This could potentially cause two copies of all the multicast traffic towards the receiver.

When PIM switches notice duplicate data packets on the LAN, they elect a single switch to forward the data packets, by each sending PIM Assert messages. If one of the upstream switches is on an SPT and the other is on an RPT, the switch on the SPT has the shortest path to the sender, and wins the Assert election. If both switches are on RPTs the switch with the shortest path to the RP (the lowest sum of metrics to the RP) wins the Assert. If both switches are on an SPT, then the switch with the shortest path to the sender (the lowest sum of metrics to the sender's DR) wins the Assert.

The switch that won the Assert election forwards these data packets, and acts as the local Designated Router for any IGMP members on the LAN. The downstream switches on the LAN also receive the Assert messages, and send all their join messages to the Assert winner. The result of an Assert election will timeout after the Assert Time. As long as the situation causing the duplication remains unchanged, the Assert winner sends an Assert message at a the Assert time interval, before the previous Assert messages time out. When the last downstream switch leaves the SPT, the Assert winner sends an Assert Cancel message saying that it is about to stop forwarding data on the SPT. Any RPT downstream switches then switch back to the RP tree.

PIM-SM Terminology

PIM-SM hello messages

When PIM-SM is enabled on a switch, it sends out a PIM-SM Hello message on all its PIM-SM enabled interfaces, and listens for Hello messages from its PIM-SM neighbors. When a switch receives a Hello message, it records the interface, IP address, priority for becoming a DR, and the timeout for the neighbor's information. The switch sends Hello messages regularly at the Hello Time interval.

Multicast Routing Information Base (MRIB)

The MRIB is a multicast topology table derived from the unicast routing table. In PIMSM, the MRIB decides where to send Join/Prune messages. It also provides routing metrics for destination addresses. These metrics are used when sending and processing Assert messages.

Tree Information Base (TIB)

The TIB is the collection of states at a PIM-SM router storing the state of all multicast distribution trees at that PIM-SM router. It is created by receiving Join/Prune messages, Assert messages, and IGMP/MLD information from local hosts.

Upstream

Upstream specifies when traffic is going towards the root of the tree. The root of the tree may be either the Source or the RP (Rendezvous Point).

Downstream

Downstream specifies when traffic is going away from the root of the tree. The root of the tree may be either the Source or the RP (Rendezvous Point).

Source-Based Trees

In the Source-Based Trees concept, the forwarding paths are based on the shortest unicast path to the source. If the unicast routing metric is the hop count, then the branches of the multicast Source-Based Trees are the minimum hop. If the routing metric is the delay, then the branches of the multicast Source-Based Trees are the minimum delay.

A corresponding multicast tree directly connects the source to all receivers for every multicast source. All traffic to the members of an associated group passes along the tree made for their source.

Shared Trees Shared Trees, or RP trees (RPT), emanate from the Rendezvous Point (RP) that receives all traffic from the sources, and forwards that traffic to the receivers.

There is a single tree for each multicast group, regardless of the number of sources. Only the routers on the tree know about the group, and information is only sent to interested receivers. With an RP, receivers have a place to join, even if no source exists. The shared tree is unidirectional, and information flows only from the RP to the receivers. If a host other than the RP has to send data on the tree, then the data must first be tunneled to the RP, then multicast to the members. This means that even if a receiver is also a source, it can only use the tree to receive packets from the RP, and to send packets to the RP (unless the source is located between the RP and the receivers).

(*,G) A ‘star G entry’ is a PIM-SM or IGMP/MLD join message that is requesting to join group G (e.g. ff0e::1) from any (*) source IP address.

(S,G) An ‘S G entry’ is a PIM-SM or MLD join message that is requesting to join group G (e.g ff0e::1) from a specified source IP address (e.g. 2001::1), where S is the IP address that is generating the multicast data for G. Note that PIM-SM supports (S,G) entries, but IGMPv2/MLDv1 do not support (S,G) entries.

BSM Boot Strap Messages, as issued by the BSR (Boot Strap Router), which is an elected router that distributes information about the RP (Rendezvous Point), where an RP is a router in a multicast network domain that acts as a shared route for a multicast shared tree.

MLD Multicast Listener Discovery. There are two versions: MLDv1 and MLDv2. MLDv1 is used by an IPv6 router to discover the presence of multicast listeners. MLDv2 provides additional features such as the ability to specify a source IPv6 address when sending a join.

PIM-SM Configuration

This section firstly provides three PIM-SM configuration examples:

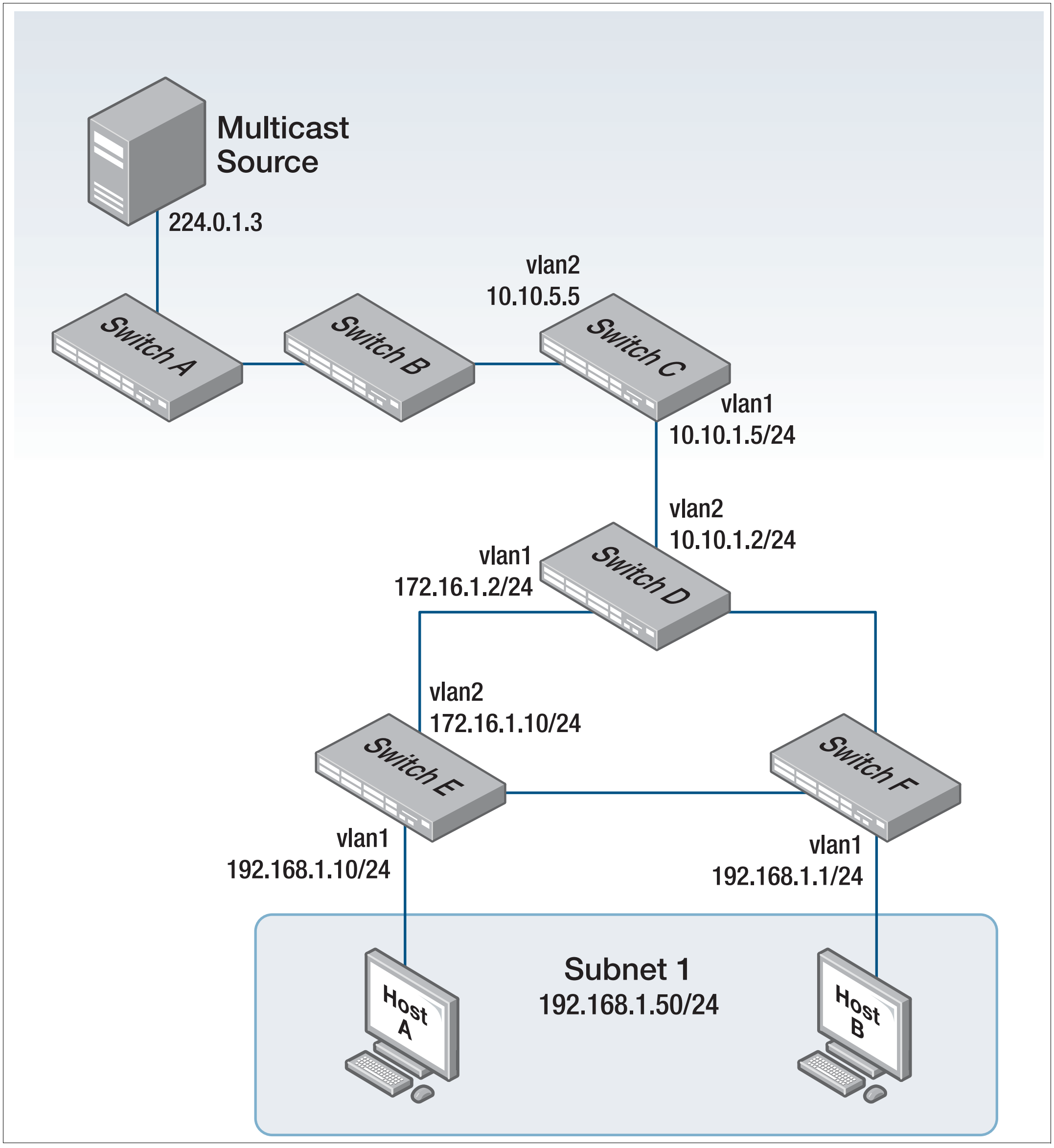
- "Static Rendezvous Point configuration" on page 10
- "Dynamic Rendezvous Point configuration" on page 13
- "Bootstrap Router configuration" on page 14

Then it discusses commands that help when you have a large multicast network on SBx908 GEN2 or SBx8100 Series switches:

- "Large multicast networks" on page 17

Both Rendezvous Point (RP) configuration examples refer to the network topology in the following graphic and use Allied Telesis managed Layer 3 Switches as the PIM routers.

Figure 2: PIM-SM Rendezvous Point configuration example



Static Rendezvous Point configuration

In this example using the above network topology, Switch C is the Rendezvous Point (RP) and all switches are statically configured with RP information. Host A and Host B join group **224.0.1.3** for all the sources. They send the IGMP membership reports to Subnet 1. Two switches are attached to Subnet 1, Switch E and Switch F. Both of these switches have default Designated Router (DR) priority on **vlan1**. Because Switch E has a higher IP address on **vlan1**, Switch E becomes the DR and is responsible for sending Join messages to the RP (Switch C).

While configuring the RP, ensure that:

- Every switch includes the **ip pim rp-address 10.10.1.5** statement, even if it does not have any source or group member attached to it.
- There is only one RP address for the whole multicast group.
- All interfaces running PIM-SM must have sparse-mode enabled. In the configuration sample output below, both **vlan1** and **vlan2** are pim sparse-mode enabled.

See the following configuration output for **Switch D**:

```
hostname Switch D
!!
interface vlan1
 ip pim sparse-mode
!
interface vlan2
 ip pim sparse-mode
!
ip multicast-routing
ip pim rp-address 10.10.1.5
!
```

Configure all the switches with the same **ip pim rp-address 10.10.1.5** statement as shown above.

Verifying configuration

Use the following commands to verify the RP configuration, interface details, and the multicast routing table.

RP details

For Switch D, the **show ip pim sparse-mode rp mapping** command shows that **10.10.1.5** is the RP for all multicast groups **224.0.0.0/4**, and is statically configured. All other switches will have a similar output.

```
awplus#show ip pim sparse-mode rp mapping
PIM Group-to-RP Mappings
Group(s): 224.0.0.0/4, Static
    RP: 10.10.1.5
         Uptime: 5d02h15m
```

For Switch D, the **show ip pim sparse-mode rp-hash** command displays the selected RP for the specified group, in this example **224.0.1.3**.

```
awplus#show ip pim sparse-mode rp-hash 224.0.1.3
    RP: 10.10.1.5
```

Interface details

For Switch E, the **show ip pim sparse-mode interface** command displays the interface details and shows that Switch E is the DR on Subnet 1.

```
awplus#show ip pim sparse-mode interface
Total configured interfaces: 16   Maximum allowed: 31
Total active interfaces:     12

Address     Interface VIFindex Ver/   Nbr      DR   DR
                                      Mode    Count  Prior
192.168.1.10    vlan2   0       v2/S   1         1   192.168.1.10
172.16.1.10     vlan3   2       v2/S   1         1   172.16.1.10
```

IP multicast routing table

Note that the multicast routing table displayed for an RP switch is different to that displayed for other switches. For Switch C, because this switch is the RP and the root of this multicast tree, the **show ip pim sparse-mode mroute** command shows **RPF nbr** (next-hop to reach RP) as **0.0.0.0** and **RPF idx** (incoming interface for this (*, G) state) as **None**.

```
awplus#show ip pim sparse-mode mroute
IP Multicast Routing Table

(*,*,RP) Entries: 0
(*,G) Entries: 1
(S,G) Entries: 0
(S,G,rpt) Entries: 0
(*, 224.0.1.3)
RP: 10.10.1.5
RPF nbr: 0.0.0.0
RPF idx: None
Upstream State: JOINED
 Local     ................................
 Joined    j...............................
 Asserted  ................................
 Outgoing  o...............................
```

For Switch E, the **show ip pim sparse-mode mroute** command displays the IP multicast routing table.

```
awplus#show ip pim sparse-mode mroute
IP Multicast Routing Table

(*,*,RP) Entries: 0
(*,G) Entries: 1
(S,G) Entries: 0
(S,G,rpt) Entries: 0
(*, 224.0.1.3)
RP: 10.10.1.5
RPF nbr: 172.16.1.2
RPF idx: port1.0.2
Upstream State: JOINED
 Local     ................................
 Joined    j...............................
 Asserted  ................................
 Outgoing  o...............................
```

On Switch E, **port1.0.2** is the incoming interface of the (*, G) entry, and **port1.0.1** is on the outgoing interface list of the (*, G) entry. This means that there is a group member through **port1.0.1**, and RP is reachable through **port1.0.2**.

Multiple static RPs

The AlliedWare Plus PIM-SM implementation supports multiple static RPs. It also supports usage of static RP and the BSR (Bootstrap Router) mechanism simultaneously.

Use the **ip pim rp-address** command to statically configure the RP address for multicast groups:

```
ip pim rp-address <ip-address> group-list <group-prefix> [override]
```

where:

- ***<group-prefix>*** - is the multicast group IP prefix address of the RP.
- **override** - enables statically defined RPs to override dynamically learned RPs.

You need to understand the following information before using this command.

If the RP address configured by the BSR, and the statically configured RP address are both available for a group range, then the RP address configured through the BSR is chosen over the statically configured RP address, unless the ‘override’ parameter is specified, in which case, the static RP will be chosen.

After configuration, the RP address is inserted into a static RP group tree based on the configured group ranges. For each group range, multiple static RPs are maintained in a linked list. This list is sorted in a descending order of IP addresses. When selecting static RPs for a group range, the first element (which is the static RP with highest IP address) is chosen.

RP address deletion is handled by removing the static RP from all the existing group ranges and recalculating the RPs for existing TIB states if required.

Note: A unique RP address may only be specified once as a static RP.

Example

```
awplus# configure terminal
```

```
awplus(config)# ip pim rp-address 192.0.2.10 group-list 233.252.0.0/24
override
```

```
AR-3050-SP-m4#sh ip pim sp rp mapping
PIM Group-to-RP Mappings
Group(s): 233.252.0.0/24, Static
    RP: 192.0.2.10
         Uptime: 00:00:17
```

Note: Previously this feature was available using ACLs on AlliedWare Plus switches only. From software version 5.4.8-0.5 it is available on all AlliedWare Plus devices.

Dynamic Rendezvous Point configuration

A static RP configuration works for a small, stable PIM domain. However, it is not practical for a large and not so stable internetwork. In such a network, if the RP fails, the network administrator may have to change the static configurations on all PIM switches. An additional reason for choosing dynamic configuration is because changes in routing traffic levels might require a change in the RP.

The Bootstrap Router (BSR) mechanism is used to dynamically maintain the RP information. To configure the RP dynamically in the above network topology, Switch C on **port1.0.1** and Switch D on **vlan1** are configured as RP candidates using the **ip pim rp-candidate** command. Switch D on **vlan1** is also configured as the BSR candidate. Since no other device has been configured as a BSR candidate, Switch D becomes the BSR router and is responsible for sending group-to-RP mapping information to all other PIM switches in this PIM domain.

The following output displays the complete configuration at Switch C.

```
awplus#show run
!
interface vlan1
 ip pim sparse-mode
!
interface vlan2
 ip pim sparse-mode
!
ip multicast-routing
ip pim rp-candidate vlan1
```

The following output displays the complete configuration at Switch D.

```
awplus#show run
!
interface vlan1
 ip pim sparse-mode
!
interface vlan2
 ip pim sparse-mode
!
ip multicast-routing
ip pim bsr-candidate vlan1
ip pim rp-candidate vlan1 priority 2
!
```

The highest priority switch is chosen as the RP. If two or more switches have the same priority, a hash function in the BSR mechanism is used to choose the RP to make sure that all devices in the PIM domain have the same RP for the same multicast group.

Use the <***interface***> [priority <***priority***>] parameters of the **ip pim rp-candidate** command to change the default priority of any RP candidate.

PIM group-to-RP mappings

The **show ip pim sparse-mode rp mapping** command displays the group-to-RP mapping details. The output shows information about RP candidates. There are two RP candidates for the group range **224.0.0.0/4**. RP candidate **10.10.1.5** has a default priority of **192**, whereas RP candidate **172.16.1.2** has been configured to have a priority of **2**. Since RP candidate **172.16.1.2** has a higher priority, it is selected as the RP for the multicast group **224.0.0.0/4**.

See the following configuration output for Switch D.

```
awplus#show ip pim sparse-mode rp mapping
This system is the Bootstrap Router (v2)
Group(s): 224.0.0.0/4
  RP: 10.10.1.5
    Info source: 172.16.1.2, via bootstrap, priority 192
         Uptime: 00:00:13, expires: 00:02:29
  RP: 172.16.1.2
    Info source: 172.16.1.2, via bootstrap, priority 2
         Uptime: 00:34:42, expires: 00:01:49
```

RP details

The **show ip pim sparse-mode rp-hash** command displays information about the RP router for a particular group. See the following configuration output for Switch D.

This output shows that **172.16.1.2** has been chosen as the RP for the multicast group **224.0.1.3**.

```
awplus#show ip pim sparse-mode rp-hash 224.0.1.3
Group(s): 224.0.0.0/4
    RP: 172.16.1.2
    Info source: 172.16.1.2, via bootstrap
```

After RP information reaches all PIM switches in the domain, various state machines maintain all routing states as the result of Join/Prune messages from members of the multicast group.

Bootstrap Router configuration

In a PIM network, every PIM multicast group needs to be associated with the IP address of a Rendezvous Point (RP). This address is used as the root of a group-specific distribution tree, whose branches extend to all nodes in the domain that want to receive traffic sent to the group. For all senders to reach all receivers, all devices in the domain use the same mappings of group addresses to RP addresses. In order to determine the RP for a multicast group, a PIM device maintains a collection of group-to-RP mappings, called the RP-Set.

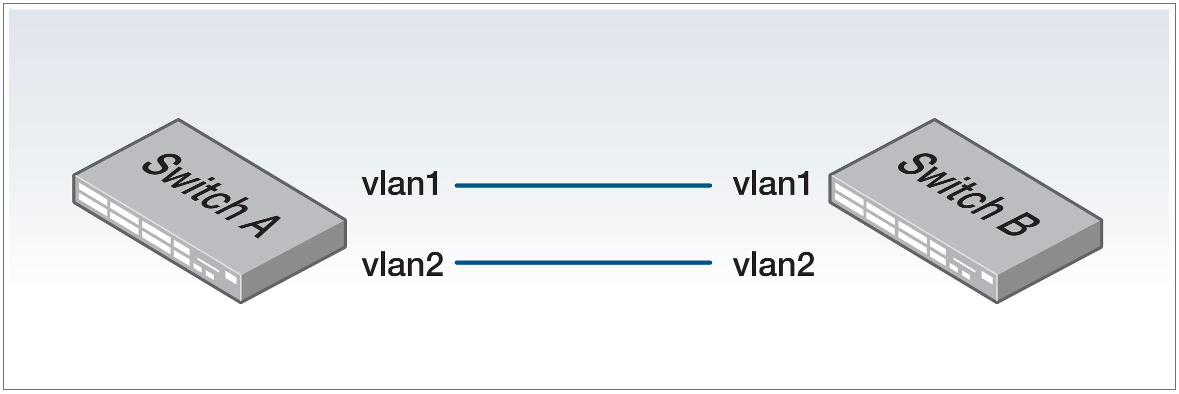
The Bootstrap Router (BSR) mechanism is the standard way that a multicast router can learn the set of group-to-RP mappings required in order to function.

Some of the PIM devices within a PIM domain are configured as RP candidates. A subset of the RP candidates will eventually be used as the actual RPs for the domain. An RP configured with a lower value in the priority field has higher priority.

Some of the PIM devices in the domain are configured to be BSR candidates. One of these BSR candidates is elected to be the BSR for the domain, and all PIM devices in the domain learn the result of this election through Bootstrap messages (BSM). The BSR candidate with highest value in the BSR priority field is the elected BSR.

The RP candidates then report their candidacy to the elected BSR, which chooses a subset of the RP candidates, and distributes corresponding group-to-RP mappings to all the devices in the domain through Bootstrap messages.

Figure 3: Bootstrap router configuration



Switch A

Enter the following commands to configure vlan1 on Switch A as the BSR candidate. The default priority is 64.

```
awplus# configure terminal
awplus(config)# ip pim bsr-candidate vlan1
awplus(config)# exit
```

Switch B

Enter the following commands to configure vlan1 on Switch B as the BSR candidate with a hash mask length of 10 and a priority of 25 and to configure vlan1 as the RP candidate with a priority of 0.

```
awplus# configure terminal
awplus(config)# ip pim bsr-candidate vlan1 10 25
awplus(config)# ip pim rp-candidate vlan1 priority 0
awplus(config)# exit
```

Validation commands

Use **show ip pim sparse-mode bsr-router** to verify the BSR candidate state on Switch A.

```
awplus#show ip pim sparse-mode bsr-router
PIMv2 Bootstrap information
This system is the Bootstrap Router (BSR)
  BSR address: 20.0.1.21
  Uptime:      00:37:12, BSR Priority: 64, Hash mask length: 10
  Expires:     00:01:32
  Role: Candidate BSR
  State: Elected BSR
```

Use **show ip pim sparse-mode bsr-router** to verify the BSR candidate state on Switch B. The initial state of the BSR candidate is pending before transitioning to BSR candidate.

```
awplus#show ip pim sparse-mode bsr-router
PIMv2 Bootstrap information
  BSR address: 20.0.1.21
  Uptime:      00:02:39, BSR Priority: 64, Hash mask length: 10
  Expires:     00:00:03
  Role: Candidate BSR
  State: Pending BSR

awplus#show ip pim sparse-mode bsr-router
PIMv2 Bootstrap information
 BSR address: 20.0.1.21
 Uptime:      00:40:20, BSR Priority: 64, Hash mask length: 10
 Expires:     00:02:07
 Role: Candidate BSR
 State: Candidate BSR
```

Use **show ip pim sparse-mode rp mapping** to verify RP-set information on Switch A.

```
awplus#show ip pim sparse-mode rp mapping
PIM Group-to-RP Mappings
This system is the Bootstrap Router (v2)
Group(s): 224.0.0.0/4
 RP: 20.0.1.11
 Info source: 20.0.1.11, via bootstrap, priority 0
 Uptime: 00:00:30, expires: 00:02:04
```

Use **show ip pim sparse-mode rp mapping** to verify RP-set information on Switch B.

```
awplus#show ip pim sparse-mode rp mapping
PIM Group-to-RP Mappings
Group(s): 224.0.0.0/4
 RP: 20.0.1.11
 Info source: 20.0.1.21, via bootstrap, priority 0
 Uptime: 00:00:12, expires: 00:02:18
```

Large multicast networks

From software version 5.4.8-1.2 onwards:

- SBx908 GEN2 switches support PIM-SM networks of up to 1024 interfaces
- SBx908 GEN2 system and SBx8100 systems containing SBx81CFC960 and SBx81XLEM cards support 8k (8192) multicast groups

From software version 5.4.9-0.1 onwards:

- SBx908 GEN2 switches support 32k (32,768) multicast groups
- x950 Series switches support PIM-SM networks of up to 1024 interfaces

This section discusses commands that help increase performance and simplify management of large multicast networks on these switches.

IGMP Querier on SBx908 GEN2 switches

If the SwitchBlade x908 GEN2 is acting as an IGMP querier, and you have 8192 or more multicast groups, you must turn off IGMP report suppression on all VLANs requiring IGMP group joins. To do this, use the command:

```
awplus(config-if)#no ip igmp snooping report-suppression
```

Silicon profile 3 on SBx8100 Series switches

To support 8192 multicast groups on the SwitchBlade x8100, you also need to change to silicon-profile 3 with a routing ratio weighting of multicast. Use the commands:

```
awplus(config)#platform silicon-profile profile3
awplus(config)#platform routingratio ipv4andipv6 weighting multicast
```

Create only (S,G) entries in hardware

When there are many multicast groups with many downstream interfaces, it can be helpful to prevent a large number of (*,G) entries from being created in hardware, using the command:

```
awplus(config)#ip multicast-routing [vrf <vrf-name>] ssm-only-hw
```

This command suppresses the creation of the (*,G) entries in hardware, but does not suppress the (S,G) entries. This improves the performance. (*,G) entries will still be created as needed in the CPU.

However, using this command may cause multicast data to be briefly flooded on the incoming interface if it comes from interfaces other than the “correct” incoming interface. The “correct” incoming interface is determined by Unicast Reverse Path Forwarding (uRPF).

Reduction of traffic to CPU

As part of software version 5.4.8-2.3, AlliedWare Plus now has the ability to block packets from coming up to the CPU when the multicast packet is seen on the wrong interface. By default this option is off and can be turned on using the following command.

```
awplus(config)#ip pim sparse-mode wrong-vif suppression
```

This will turn on the ability to block multicast packets from coming up to the CPU when they are on the wrong interface. This is done by creating a hardware entry for this multicast stream and dropping the packet instead of bringing it up to the CPU. NOTE: For each multicast stream seen on the wrong interface it will consume an entry normally used for multicast routing. For example if the network has 10K multicast streams on the wrong interface, this leaves only 22K hardware entries left for normal multicast traffic on a SBx908 GEN2.

Join/Prune message batching

From software version 5.4.8-2.6 onwards, IPv4 PIM-SM can be configured to attempt to batch Join and Prune messages for multiple groups in one PIM message packet. By default this is disabled and can be enabled with the following command.

```
awplus(config)#ip pim sparse-mode join-prune-batching
```

This will enable batching of PIM-SM Join and Prune messages, reducing the number of packets sent by the PIM-SM daemon when many groups are in use, it is recommended to enable this option if more than 4K streams are in use.

Filtering show command output

If you want to see information about a single interface or group, you can use the | symbol and filter the output. There are two particularly useful options: **begin** and **include**.

| begin This skips all the output before the first line that has the specified text. For example, to see the PIM interface details of vlan555 (and onwards), use the command:

```
awplus#show ip pim sparse-mode interface detail | begin vlan555
```

This gives the following output.

```
awplus#show ip pim sparse-mode interface detail | begin vlan555
...skipping
vlan555 (vif 555):
  Address 192.168.1.149, DR 192.168.1.149
  Hello period 30 seconds, Next Hello in 15 seconds
  Triggered Hello period 5 seconds
  Neighbors:
   192.168.1.22
```

| include This only displays lines of output that contain the specified text. For example, to see which VLAN has the address 10.2.104.10, use the command:

```
awplus#show ip pim sparse-mode interface detail | include 10.2.104.10
```

This gives the following output.

```
awplus#show ip pim sparse-mode interface | include 10.2.104.10
10.2.104.10      vlan1004   8        v2/S   0      1
10.2.104.10
```

Viewing the maximum number of multicast groups supported in hardware

You can see the maximum number of multicast groups by checking the route limit/route threshold setting in the **show ip mroute count** command:

```
awplus#show ip mroute count

IP Multicast Statistics
Total 8 routes using 1408 bytes memory
Route limit/Route threshold: 32768/32768
Total NOCACHE/WRONGVIF/WHOLEPKT recv from fwd: 398/0/0
Total NOCACHE/WRONGVIF/WHOLEPKT sent to clients: 398/0/0
Immediate/Timed stat updates sent to clients: 0/80
Reg ACK recv/Reg NACK recv/Reg pkt sent: 0/0/0
Next stats poll: 00:01:04

Forwarding Counts: Pkt count/Byte count, Other Counts: Wrong If pkts
Fwd msg counts: WRONGVIF/WHOLEPKT recv
Client msg counts: WRONGVIF/WHOLEPKT/Imm Stat/Timed Stat sent
Reg pkt counts: Reg ACK recv/Reg NACK recv/Reg pkt sent

(192.168.2.121, 224.6.1.1), Forwarding: 11/10, Other: 0
  Fwd msg: 0/0, Client msg: 0/0/0/10,  Reg: 0/0/0

(192.168.2.121, 224.6.1.2), Forwarding: 11/10, Other: 0
  Fwd msg: 0/0, Client msg: 0/0/0/10,  Reg: 0/0/0

(192.168.2.121, 224.6.1.3), Forwarding: 11/10, Other: 0
  Fwd msg: 0/0, Client msg: 0/0/0/10,  Reg: 0/0/0

(192.168.2.121, 224.6.1.4), Forwarding: 11/10, Other: 0
  Fwd msg: 0/0, Client msg: 0/0/0/10,  Reg: 0/0/0

(192.168.2.121, 224.6.1.5), Forwarding: 11/10, Other: 0
  Fwd msg: 0/0, Client msg: 0/0/0/10,  Reg: 0/0/0

(192.168.2.121, 224.6.1.6), Forwarding: 11/10, Other: 0
  Fwd msg: 0/0, Client msg: 0/0/0/10,  Reg: 0/0/0

(192.168.2.121, 224.6.1.7), Forwarding: 11/10, Other: 0
  Fwd msg: 0/0, Client msg: 0/0/0/10,  Reg: 0/0/0

(192.168.2.121, 224.6.1.8), Forwarding: 11/10, Other: 0
  Fwd msg: 0/0, Client msg: 0/0/0/10,  Reg: 0/0/0
```

When 32,000 entries are learnt the show command displays can get very long and take a long time to complete. Often administrator are only interested in the total number of entries learnt, so using the ‘brief’ version of show commands can be useful:

To display the L2 IGMP groups learnt and the number of interfaces configured use the command:

```
awplus# show ip igmp groups brief
```

```
awplus#show ip igmp groups brief
IGMP Groups Brief

  IGMP Configured Interfaces             98
  IGMP Configured Interfaces (Up)        98
  IGMP Configured Interfaces (Down)      0

  IGMP Stopped Groups                    32000
  IGMP Dynamic Groups                    32000
  IGMP Static Groups                     0
```

To display L3 PIM sparse-mode multicast entries learnt, use the command:

```
awplus# show ip pim sparse-mode mroute brief
```

```
awplus#show ip pim sparse-mode mroute brief
IP Multicast Routing Table

(*,*,RP) Entries: 0
(*,G) Entries: 0
(S,G) Entries: 32000
(S,G,rpt) Entries: 32000
FCR Entries: 0
MRIB Msg Cache Hit: 0
```

To see the number of multicast entries learnt in hardware and which downstream interfaces are used, use the command:

```
awplus# show platform table ipmulti-brief.
```

Stack members should contain the same number of entries across all members. The **Hit** bit parameter shows the route is still active. Normally, you would only expect it to be active on one stack member, but there are cases where it should be present on both stack members. If there are no port counts there are no downstream members and no traffic will be forwarded. The **show platform table ipmulti** command can be used to debug this further.

```
awplus# sh platform table ipmulti-brief

Stack member 1:

[Instance 4]
Source Ip         Multicast         VLAN   L2 Port   L3 Port   Hit
Addr              IP Addr           ID     Count     Count     Bit
------------------------------------------------------------------
10.36.20.1        239.255.1.5       10     2         1         yes
10.36.20.1        239.255.1.4       10     2         1         yes
10.36.20.1        239.255.1.3       10     2         1         yes
10.36.20.1        239.255.1.2       10     2         1         yes
10.36.20.1        239.255.1.1       10     2         1         yes
2001:db8:ffff::1  ff08::5           10     3         1          no
2001:db8:ffff::1  ff08::4           10     3         1          no
2001:db8:ffff::1  ff08::2           10     3         1          no
2001:db8:ffff::1  ff08::3           10     3         1          no


Stack member 2:

[Instance 8]
Source Ip         Multicast         VLAN   L2 Port   L3 Port   Hit
Addr              IP Addr           ID     Count     Count     Bit
------------------------------------------------------------------
10.36.20.1        239.255.1.5       10     2         1          no
10.36.20.1        239.255.1.4       10     2         1          no
10.36.20.1        239.255.1.3       10     2         1          no
10.36.20.1        239.255.1.2       10     2         1          no
10.36.20.1        239.255.1.1       10     2         1          no
2001:db8:ffff::1  ff08::5           10     3         1         yes
2001:db8:ffff::1  ff08::4           10     3         1         yes
2001:db8:ffff::1  ff08::2           10     3         1         yes
2001:db8:ffff::1  ff08::3           10     3         1         yes
```

PIM-SM Show Commands

show ip pim sparse-mode bsr-router

```
awplus-1#show ip pim sparse-mode bsr-router
PIMv2 Bootstrap information
BSR address: 192.168.6.1
Uptime: 02:19:26, BSR Priority: 64, Hash mask length: 10
Expires: 00:01:27
Role: Candidate BSR
State: Candidate BSR
Candidate RP: 192.168.1.1(vlan1)
Advertisement interval 60 seconds
Next C-RP advertisement in 00:00:09
```

The following list describes the meaning of the output above:

- **BSR address: 192.168.6.1**

  The BSR is switch awplus-4—192.168.6.1

- **Uptime: 02d19h26m, BSR Priority: 64, Hash mask length: 10**

  The BSR has been an active candidate for more than 2 days and has a Priority of 64, and is telling routers to use a mask length of 10 in the hash algorithm they use for choosing RPs.

- **Expires: 00:01:27**

  This timer will count down until it is refreshed by a Bootstrap message.

- **Role: Candidate BSR**
  State: Candidate BSR

  If this switch (that is, the switch on which the show command has been executed) is the BSR it will show Elected BSR, otherwise it will show Candidate BSR.

- **Candidate RP: 192.168.1.1(vlan1)**

  This shows that we have configured the switch to use the IP address of VLAN1 as its RP Candidate address. The command to configure its candidacy was: **ip pim rp-candidate vlan1**

- **Advertisement interval 60 seconds**
  Next C-RP advertisement in 00:00:09

  This switch will send its next RP Candidate Advertisement message in 9 seconds.

show ip pim sparse-mode interface

```
Total configured interfaces: 100   Maximum allowed: 100
Total active interfaces:     100

Address           Interface VIFindex Ver/   Nbr    DR          DR
                                     Mode   Count  Prior
10.1.100.4        vlan100   4        v2/S   2      1     10.1.100.6
10.2.101.10       vlan1001  5        v2/S   0      1     10.2.101.10
10.2.102.10       vlan1002  6        v2/S   0      1     10.2.102.10
```

Note that this command displays the DR Priority and which IP address is the DR for each interface.

show ip pim sparse-mode mroute

```
awplus-2#show ip pim sparse-mode mroute
IP Multicast Routing Table
(*,*,RP) Entries: 0
(*,G) Entries: 3
(S,G) Entries: 1
(S,G,rpt) Entries: 1
FCR Entries: 0
(*, 225.1.1.1)
RP: 192.168.1.1
RPF nbr: 192.168.2.1
RPF idx: vlan2
Upstream State: JOINED
Local
.....................................................................
.................................
Joined
.....j...............................................................
.................................
Asserted
.....................................................................
.................................
FCR:
Interop listener rx-data flags (ES,EDW,RXD,DAJ,EOE)
0x00000000 0x00000000 0x00000001
(192.168.1.50, 225.1.1.1)
RPF nbr: 192.168.2.1
RPF idx: vlan2
SPT bit: 1
Upstream State: JOINED
Local
.....................................................................
.................................
Joined
.....j...............................................................
.................................
Asserted
.....................................................................
.................................
Outgoing
.....o...............................................................
.................................
Interop listener rx-data flags (ES,EDW,RXD,DAJ,EOE)
0x00000000 0x00000000 0x00000001
(192.168.1.50, 225.1.1.1, rpt)
RP: 192.168.1.1
RPF nbr: 192.168.2.1
RPF idx: vlan2
Upstream State: NOT PRUNED
Local
.....................................................................
.................................
Pruned
.....................................................................
.................................
Outgoing
.....o...............................................................
.................................
Interop listener rx-data flags (ES,EDW,RXD,DAJ,EOE)

0x00000000 0x00000000 0x00000001
```

We can see that there are actually three entries for the Group 225.1.1.1, as shown in red above.

(*,G) entry

First we have the (*,G) entry:

- **(*, 225.1.1.1)**

  This entry is for a tree that is set up to receive a stream (any stream) for group 225.1.1.1 from the RP for that group.

- **RP: 192.168.1.1**

  The RP for this group is 192.168.1.1—switch awplus-1.

- **RPF nbr: 192.168.2.1**

  This shows the RPF (Reverse Path Forwarding) neighbor. This is the next-hop for the route back to the RP of 192.168.1.1.

- **RPF idx: vlan2**

  This is the interface over which the source network is reached.

- **Upstream State: JOINED**

  There is a Listener that wants this group. A PIM Join has been sent to the RP for this group.

- **Local ...i...**.

  An 'i' is shown here if there is a there is a Listener connected to this switch that wants this group.

  There are no other options.

  The position of the letter amongst the dots indicates which interface the Listener is on (corresponding to the **show ip mvif** output), so if you have:

  Local ...i..

  We can see that the 'i' is next to the third dot, which indicates mvif 3:

```
awplus-1#show ip mvif
Interface Vif Owner TTL Local Remote Uptime
Idx Module Address Address
vlan1 2 PIM-DM 1 192.168.1.1 0.0.0.0 04:17:43
vlan2 3 PIM-DM 1 192.168.2.1 0.0.0.0 04:17:43
vlan3 4 PIM-DM 1 192.168.3.1 0.0.0.0 04:17:43
```

This means that the local listener is attached to vlan2 (index 3)

- **Joined .....j...**

  The 'j' shown here means that there has been a PIM Join from a downstream interface for this Group. There are no other options.

- **Asserted ....**

  Indicates whether the outgoing interface is involved in an assert process or not, and what role—Assert Winner ('W') or Assert Loser ('L')— it has. In this case, the router is not asserting on any interfaces.

- **FCR: Forwarding Cache Register (the forwarding table):**

  Interop listener rx-data flags (ES,EDW,RXD,DAJ,EOE)

  0x00000000 0x00000000 0x00000001

  Listener—flags indicating which modules listen to events for this FCR entry.

  rx-data—flags indicating which modules wants to receive data corresponding to this FCR entry.

  The hex values are as follows:

  0x00000001—PIM-SM

  0x00000002—PIM-DM

  The flags are:

  ES—indicating that data corresponding to this FCR is from an external source (an interface belonging to another module).

  EDW—indicating this FCR is waiting to be deleted.

  RXD—indicating at least one other module listens to data for this FCR.

  DAJ—indicating at least one other module wants to be alerted if data hits this FCR.

  EOE— indicating that the external olist is empty (there are no interfaces belonging to other modules in the olist).

  The hex value of the flags are:

  0x00000001 EOE

  0x00000002 DAJ

  0x00000004 RXD

  0x00000008 EDW

  0x00000010 ES

(S,G) entry Then we have the (S,G) entry:

- **(192.168.1.50, 225.1.1.1)**

  This is a distribution tree for the stream from source 192.168.1.50, destined to group 225.1.1.1. So, any Tree Information Base entry in the (S,G) category tracks the router's current involvement in such a distribution tree

- RPF nbr: 192.168.2.1
  RPF idx: vlan2
  SPT bit: 1

  “SPT” stands for “Source Path Tree”. Having this bit set (i.e. value 1) means that the this is a distribution tree for forwarding streams direct from their source (rather than via the RP).
  Upstream State: JOINED
  Local ....
  Joined .....j...
  Asserted ....
  Outgoing .....o..

  The 'o' here means that there is a downstream interface that this group is going to.

- Interop listener rx-data flags (ES,EDW,RXD,DAJ,EOE)

  0x00000000 0x00000000 0x00000001

(S,G,rpt) entry

And finally the (S,G,rpt) entry:

- **(192.168.1.50, 225.1.1.1, rpt)**

  This is more a type of ‘negative’ distribution tree, telling routers what NOT to forward. We can see below that the Upstream State is NOT PRUNED, i.e Forwarding.

- RP: 192.168.1.1

  RPF nbr: 192.168.2.1

  RPF idx: vlan2

  Upstream State: NOT PRUNED
  Local ....
  Pruned ....
  Outgoing .....o...
  Interop   listener   rx-data   flags (ES,EDW,RXD,DAJ,EOE)
        0x00000000 0x00000000 0x00000001

  The significance of this (S,G,rpt) entry is that it is effectively an 'exception' to the (*,225.1.1.1) entry. The (*,255.1.1.1) says 'send any stream for 225.1.1.1, from any source, to me'. But, actually we do not want the RP to send the (192.168.1.50,225.1.1.1) stream to us, because we are now getting that stream directly from the source. So, the (192.168.1.50,225.1.1.1,rpt) entry indicates that we have said to the RP 'send streams for 225.1.1.1 to me, from any source except 192.168.1.50'.

show ip pim sparse-mode rp mapping

```
awplus-1#show ip pim sparse-mode rp mapping
PIM Group-to-RP Mappings
Group(s): 224.0.0.0/4
RP: 192.168.1.1
Info source: 192.168.6.1, via bootstrap, priority 3
Uptime: 00:06:06, expires: 00:01:34
```

- **Group(s): 224.0.0.0/4**

  The RP 192.168.1.1 shown below is the RP for all Multicast Groups (224.0.0.0/4). With a different configuration, it could be just for one or several groups—see **ip pim sparse-mode rp-candidate** and **ip pim sparse-mode rp mapping** commands.

- **RP: 192.168.1.1**

  Info source: 192.168.6.1, via bootstrap, priority 192

  This information is provided by the BSR 192.168.6.1.

- **Uptime: 00:06:06, expires: 00:01:34**

  The **Uptime** shows the time elapsed since the switches agreed that 192.168.1.1 is the RP for the group or groups—in this case all Groups.

  The **expires** time shows the time left on the RP's holdtime, advertised by the BSR.

show ip pim sparse-mode rp-hash

```
show ip pim sparse-mode rp-hash <group-addr>
```

```
awplus-1#show ip pim sparse-mode rp-hash 225.1.1.1
RP: 192.168.1.1
Info source: 192.168.6.1, via bootstrap
```

This command shows which RP a particular multicast group is mapped to.

show ip rpf

This command displays the Reverse Path Forwarding (RPF) information for a specified address somewhere in the network.

```
awplus-2#sh ip rpf 192.168.1.50
RPF information for 192.168.1.50
RPF interface: vlan2
RPF neighbor: 192.168.2.1
RPF route: 192.168.1.0/24
RPF type: unicast (ospf)
RPF recursion count: 0
Doing distance-preferred lookups across tables
Distance: 110
Metric: 20
```

This command, shown on switch awplus-2 above, gives us the following RPF information.

When a multicast packet enters one of the switch’s interfaces, it will check the source IP address against the networks it knows about via that interface. If the switch finds a matching routing entry for the source IP address of the multicast packet, the RPF check passes and the packet is forwarded to all other interfaces that are participating in multicast for this group.

The routing information provided in this command is taken from the unicast IP routing table (**show ip route**).

- **RPF information for 192.168.1.50**
  RPF interface: vlan2

  The interface via which the source 192.168.1.50 is reached.

- **RPF neighbor: 192.168.2.1**

  The IP address of the neighbor switch’s interface via which the source is reached— the next-hop.

- **RPF route: 192.168.1.0/24**

  This is the network in the unicast routing table that the source is a member of.

- **RPF type: unicast (ospf)**

  We see here that the unicast route to the source's network was learned via OSPF.

- **RPF recursion count: 0**

  Whether the RPF nexthop is a recursive or not. Despite the name implying a count, the only possible values it can have are: 0—not recursive nexthop, or 1—recursive nexthop.

- **Doing distance-preferred lookups across tables**

  All of the Unicast and Multicast routing tables in the switch (i.e Unicast Routing Table, Static Mroute table) are searched for a match for the source network.

  If there are different routes for the same source network from different routing protocol tables, the Distance of the routes will be used to decide the route that is to be used—the route with the lowest Distance will be chosen, not the route with the longest mask.

- **Distance: 110**

  The Administrative Distance or Preference of this route to the source network. It was learned via OSPF, so it has a Distance of 110.

- **Metric: 20**

  The Metric of this route to the source network.

The IP route table entry from which this information is derived is:

```
awplus#show ip route 192.168.1.50
Routing entry for 192.168.1.0/24
Known via "ospf", distance 110, metric 20, best
Last update 1d01h48m ago
* 192.168.2.1, via vlan2
```

show ip mroute

The IP mroute table shows all active multicast forwarding entries in the AlliedWare plus routing engine.

```
awplus#show ip mroute
IP Multicast Routing Table
Flags: I - Immediate Stat, T - Timed Stat, F - Forwarder installed
Timers: Uptime/Stat Expiry
Interface State: Interface (TTL)
(192.168.1.50, 225.1.1.1), uptime 00:00:52, stat expires 00:02:38
Owner PIM-SM, Flags: TF
Incoming interface: vlan1
Outgoing interface list:
vlan2 (1)
```

Let's understand the meaning of the information in this mroute entry.

- **(192.168.1.50, 225.1.1.1), uptime 00:00:52, stat expires 00:02:38**

  For the group 225.1.1.1 from source 192.168.1.50

- **Owner PIM-DM, Flags: TF**

  The PIM version used here is PIM Dense Mode.

  We can see from the Flags TF that Timed Stat (T) is being used and that this switch is the Forwarder (F) for this group.

  There are two possible values of the 'stat' type. Timed Stat and Immediate Stat refer to the way traffic statistics are collected for the given mroute by the routing protocol via the routing engine. Timed Stat means that the mroute entry is checked for traffic hitting the entry to refresh the PIM keep alive timer at a specific time; Immediate Stat signifies that the mroute can be queried at anytime by the protocol module for traffic hitting the mroute.

  The "forwarder" flag is relevant to the case where more than one PIM router could be forwarding onto a given downstream LAN. To avoid duplication of shared traffic, PIM routers connected to a shared segment will elect a single Forwarder for that particular segment. PIM relies on a process called the PIM Assert Mechanism to make this determination. The Assert winner becomes the forwarder, and denotes this fact by displaying the "F" flag on the multicast route table entry.

- **Incoming interface: vlan1**

  This group is arriving on the switch’s vlan1 interface.

- **Outgoing interface list:**

  vlan2 (1)

  The group is being sent out of the switch's vlan2 interface, as there have not been any Prunes—there are Listeners that want to receive this group via vlan2.

show ip mroute count

```
awplus#show ip mroute count
IP Multicast Statistics
Total 5 routes using 620 bytes memory
Route limit/Route threshold: 2048/2048
Total NOCACHE/WRONGVIF/WHOLEPKT recv from fwd: 330/0/0
Total NOCACHE/WRONGVIF/WHOLEPKT sent to clients: 330/0/0
Immediate/Timed stat updates sent to clients: 0/326
Reg ACK recv/Reg NACK recv/Reg pkt sent: 0/0/0
Next stats poll: 00:01:39
Forwarding Counts: Pkt count/Byte count, Other Counts: Wrong If pkts
Fwd msg counts: WRONGVIF/WHOLEPKT recv
Client msg counts: WRONGVIF/WHOLEPKT/Imm Stat/Timed Stat sent
Reg pkt counts: Reg ACK recv/Reg NACK recv/Reg pkt sent
(192.168.1.50, 225.1.1.1), Forwarding: 26/25, Other: 0
Fwd msg: 0/0, Client msg: 0/0/0/0, Reg: 0/0/0
```

- IP Multicast Statistics

  Total 5 routes using 620 bytes memory

  Route limit/Route threshold: 2048/2048

  Total NOCACHE/WRONGVIF/WHOLEPKT recv from fwd: 330/0/0

  This is the number of packets which the routing engine has received from the kernel (referred to as 'fwd').

- **Total NOCACHE/WRONGVIF/WHOLEPKT sent to clients: 330/0/0**

  This is the number of packets which the routing engine has sent to routing protocol processes like PIMd (clients).

  NOCACHE—kernel notification for packets with no matching multicast route in the forwarder.

  WRONGVIF—kernel notification for packets arriving on the incorrect incoming interface (from the RPF point of view).

  WHOLEPKT—kernel notification for packets that have been encapsulated (register packets, PIM-SM only).

- **Immediate/Timed stat updates sent to clients: 0/326**

  The number of Immediate or Timed Stats updates sent to the client routing process (PIMd)—in this case 326 Timed updates.

- **Reg ACK recv/Reg NACK recv/Reg pkt sent: 0/0/0**

  Reg ACK recv—Acknowledgement for Register requests from the routing protocol client (PIMd) to the kernel

  Reg NACK recv—Negative Acknowledgement for Register request from the routing protocol client (PIMd) to the kernel

  Reg pkt sent—Register packets sent from the kernel to the RP.

- **Next stats poll: 00:01:39**

  The length of time until the next Timed Status update.

- Forwarding Counts: Pkt count/Byte count, Other Counts: Wrong If pkts

  Fwd msg counts: WRONGVIF/WHOLEPKT recv

  Client msg counts: WRONGVIF/WHOLEPKT/Imm Stat/Timed Stat sent

  Reg pkt counts: Reg ACK recv/Reg NACK recv/Reg pkt sent

  (192.168.1.50, 225.1.1.1), Forwarding: 119/118, Other: 0

  For this Source and Group, the number of packets/bytes.

- **Fwd msg: 0/0, Client msg: 0/0/0/0, Reg: 0/0/0**

  These counters are explained below as Fwd msg: A/B, Client msg: C/D/E/F, Reg: G/H/I

  A—'wrong VIF' messages received from forwarder for this Source and Group. Kernel notification for packets arriving on the incorrect incoming interface.

  B—register request messages received from forwarder kernel notification for packets which have been encapsulated (register packets).

  C—'wrong VIF' messages sent to clients (routing protocol modules) for this Source and Group. Number of packets from the routing protocol client (PIMd) arriving on the incorrect incoming interface.

  D—register request messages sent to clients (routing protocol modules)—packets which have been encapsulated (register packets).

  E—'immediate stat' messages sent to clients (routing protocol modules)—number of Immediate Update Status messages sent.

  F—'timed stat' messages sent to clients (routing protocol modules)—Number of Timed Update status messages sent.

  G—Register acknowledgements received from client (protocol module process—currently only PIM-SM) for this Source and Group. The number of Acknowledgements for Register requests from the client (PIM) to the kernel.

  H—Register negative acknowledgements received from client (protocol module process—currently only PIM-SM). The number of Negative Acknowledgements for Register requests from the client (PIM) to the kernel.

  I—Register packets sent to the forwarder. The number of Register packets sent from the kernel to the RP.

PIM-SSM

Protocol Independent Multicast - Source Specific Multicast (PIM-SSM) is derived from Protocol Independent Multicast - Sparse Mode (PIM-SM) and is a simplified version of PIM-SM.

For details of the commands used to configure PIM-SSM, see the PIM-SM and IGMP chapters of your switch’s Command Reference. The Command Reference is available on our website at alliedtelesis.com.

Characteristics of PIM-SSM

One of the significant characteristics of PIM Sparse Mode and PIM Dense Mode is the fact that hosts, and most of the routers, in the network do not know the source address of the multicast groups they wish to join.

Keeping the network in the dark about the source addresses of the groups makes the network management a bit simpler.

It means that you:

- Don’t need a process by which hosts are told the source addresses of the streams in advance (although, that is only a small saving of hassle, as you still need a process to tell the hosts the group addresses of the streams in advance).
- Can change the multicast servers around as much as you like, without having to go telling all the hosts that the server address has changed.

But, it has quite big **disadvantages**:

- It makes the multicast routing protocol more complicated. A lot of the functionality within PIM Sparse Mode and PIM Dense Mode is necessitated by the fact that the hosts and routers do not know the source of a group being requested. The whole business of Rendezvous Points in Sparse Mode and State Refreshes in Dense Mode are ways that those protocols deal with the fact that stream’s source addresses are not known to the requesting hosts.
- If you are receiving multicast feeds from multiple external content providers, then you need to be careful that the group addresses to which these providers are sending do not overlap with each other.
- You are somewhat vulnerable to multicast DOS attacks. If an attacker knows the group address of a stream in use in your network, then they can simply send in multicast packets destined to that group address. These packets will interfere with the genuine stream, as they will be forwarded to hosts listening to that group, irrespective of what source IP they come from.

In light of these disadvantages, a variant of Multicast routing, called **Source Specific Multicast** (**SSM**), has been defined.

In SSM routing:

- The hosts requesting streams need to know the source address of the stream they are requesting, and must specify the source in their request
- Routers differentiate between streams to the same group address, but from different source addresses. If they have been requested to send a stream (S1,G), but not a stream to the same group, from a different source (S2,G), they will forward (S1,G), but not (S2,G).

A version of PIM Sparse Mode has been created that supports SSM. Unsurprisingly, it has been named PIM SSM.

PIM-SSM IP address range

The Internet Assigned Numbers Authority (IANA) has reserved the address range 232.0.0.0 through 232.255.255.255 for SSM applications. Although PIM-SSM can technically be configured to use the entire 224/4 multicast address range, PIM-SSM operation is guaranteed only in the 232.0.0.0/8 range, except 232.0.0.0/24, which is reserved.

IGMPv3 and SSM-mapping

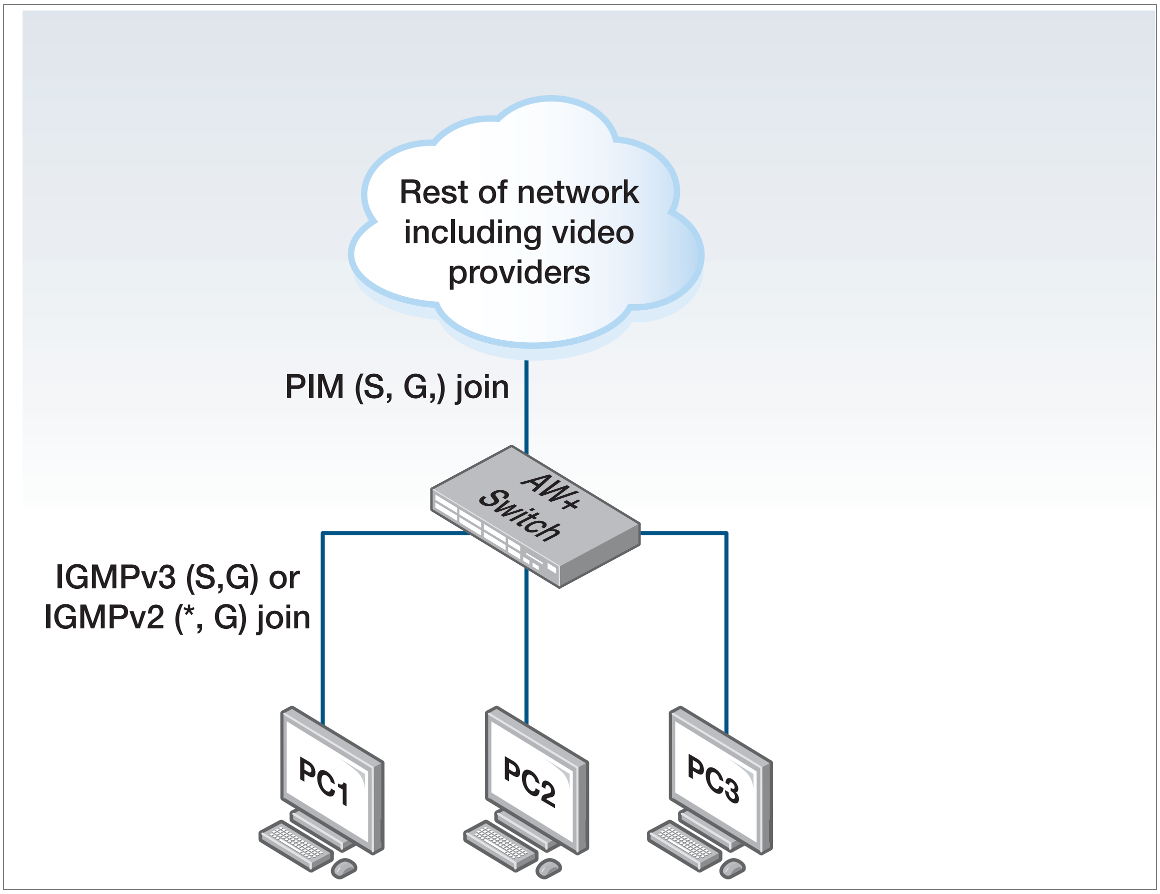
A restriction of PIM-SSM is that it requires a “Source, Group” (S,G) join and only IGMPv3/MLDv2 support this. Source Specific Multicast Mapping is required to use PIM-SSM when you have older multicast client devices that do not support IGMPv3. This additional feature allows you to statically map IGMPv1/v2 (*,G) joins into PIM (S,G) joins, which in turn allows the device to talk to an upstream PIM-SSM network.

Note that IGMPv3 (*,G) joins cannot be mapped by SSM-Mapping.

How PIM-SSM works

To join multicast group **232.1.1.1** each PC must send an IGMPv3 join with the source IP address specified. The join will be a (S,G) join, for example (**85.1.1.1,232.1.1.1**). The router will receive the IGMP join and check if the group address is in the SSM range. Then:

- If the group address is in the SSM range, the router will verify that a specific source or sources have been included in the IGMP join.
- If a specific source or sources has been included in the IGMP join, then the router will forward a PIM (S,G) join towards the source IP address.
- If the source IP address is not specified, then the router will discard the IGMP join and the PC will not join the group.
- If an IGMPv2 join is received for the SSM range then by default the join is discarded because no source IP address is specified. IGMP joins for group addresses that are not in the SSM range do not need to specify a specific source IP address.


Rest of network including video providers
PIM (S, G,) join
AW+ Switch
IGMPv3 (S,G) or IGMPv2 (*, G) join
PC1
PC2
PC3

How IGMP-SSM mapping works

In the example above ("How PIM-SSM works"), if an IGMPv2 join is sent it is discarded because IGMPv2 only supports (*,G) joins. To resolve this issue, IGMP SSM-Mapping allows the router to be statically configured with source IP addresses for each group address or range of group addresses. This allows the router to receive a (*,G) join, match the group address via a software ACL, and based on this, insert the matching source IP address. The router then treats the join as a normal (S,G) join. If no match is found then the (*,G) is used. If the group address is in the SSM range then the join is discarded.

Configuring PIM-SSM

Table 6: General configuration procedure for PIM-SSM

To enable SSM on the device	

`awplus#`
`configure terminal`
	Enter Global Configuration mode.

`awplus(config)#`
`ip igmp ssm-map enable`
	This command applies to VLAN interfaces configured for IGMP.

To specify the static mode of defining SSM	

`awplus#`
`configure terminal`
	Enter Global Configuration mode.

`awplus(config)#` access-list {*<1-99>*|*<1300-1999>*} permit *<source>* OR access-list standard *<standard-access-list-name>* permit *<source>*	Configure either a Standard Numbered ACL, Expanded Numbered ACL, or Standard Named ACL. Specify a multicast group address range and wildcard mask with the ***source*** parameter.
`awplus(config)#` ip igmp ssm-map static {*<access-list-number>*|*<access-list-number-expanded>*|*<access-list-name>*} *<ip-address>*	This command applies to VLAN interfaces configured for IGMP. SSM statically assigns sources to IGMPv1 and IGMPv2 groups to translate such (*,G) groups' memberships to (S,G) memberships for use with PIM-SSM.
To define a non-default SSM range of IP multicast addresses in IGMP	

`awplus#`
`configure terminal`
	Enter Global Configuration mode.

`awplus(config)#` access-list {*<1-99>*} permit *<source>* OR access-list standard *<standard-access-list-name>* permit *<source>*	Configure either a Standard Numbered ACL or Standard Named ACL.
`awplus(config)#` ip igmp ssm range {*<access-list-number>*|*<access-list-name>*}	Incoming IGMPv1 and IGMPv2 join requests are ignored if the multicast IP address is in the SSM range and no SSM mapping is configured for these addresses. By default, the SSM range is 232/8. To define the SSM range to be other than the default, specify either an access-list name or and access-list number.
To define the Source Specific Multicast (SSM) range of IP multicast addresses	
`awplus#`	Enter Global Configuration mode.

Table 6: General configuration procedure for PIM-SSM (continued)

`awplus(config)#`
`ip pim ssm default`
	The default keyword defines the SSM range as 232/8.
	OR

`awplus(config)#`
`ip pim ssm range {<access-list>|<named-access-list>}`
	To define the SSM range to be other than the default, use the named-access-list parameter option.