

PIM Sparse Mode

Module 5

Module Objectives

Cisco.com

- **Identify and explain the basic mechanisms of PIM Sparse Mode.**
- **Configure and verify normal PIM SM operation.**

Module Agenda



Cisco.com

- **PIM SM Overview**
- **PIM SM Protocol Mechanics**
- **PIM SM Review**
- **Configuring PIM SM**

PIM Sparse Mode Overview

Cisco.com

- **Explicit join model**
 - Receivers join to the Rendezvous Point (RP)
 - Senders register with the RP
 - Data flows down the shared tree and goes only to places that need the data from the sources
 - Last hop routers can join source tree if the data rate warrants by sending joins to the source
- **RPF check depends on tree type**
 - For shared trees, uses RP address
 - For source trees, uses Source address

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

4

- **Explicit “Join” Model**

- Unlike PIM Dense mode, PIM Sparse mode uses the explicit join model whereby Receivers send PIM Join messages to a designated “Rendezvous Point” (RP). (The RP is the root of a shared distribution tree down which all multicast traffic flows.)
- In order to get multicast traffic to the RP for distribution down the shared tree, Senders send Register messages to the RP. Register messages cause the RP to send a “Join” towards the source so that multicast traffic can flow to the RP and hence down the shared tree.
- Last hop routers may be configured with an “SPT-Threshold” which, once exceeded, will cause the last hop router to join the “Shortest Path Tree” (SPT). This will result in the multicast traffic from the source to flow down the SPT from the source to the last hop router.

- **RPF Check depends on tree type**

- If traffic is flowing down the shared tree, the RPF check mechanism will use the IP address of the RP to perform the RPF check.
- If traffic is flowing down the SPT, the RPF check mechanism will use the IP address of the Source to perform the RPF check.

PIM Sparse Mode Overview

Cisco.com

- Only one RP is chosen for a particular group
- RP statically configured or dynamically learned (Auto-RP or PIM v2 BSR)
- Data forwarded based on the source state (S, G) if it exists, otherwise use the shared state (*, G)
- RFC 2362 - "PIM Sparse Mode Protocol Spec"

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

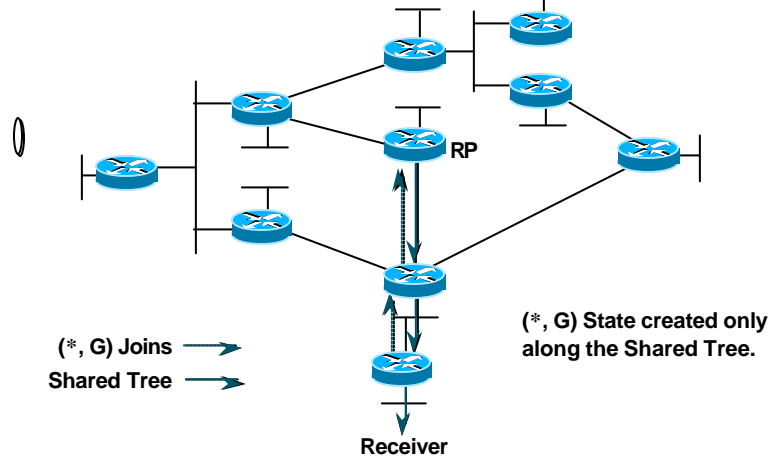
8/10/2001 2:37 PM

5

- **Only one RP for a group may be active at a time**
 - While it is usually the case that a single RP serves all groups, it is possible to configure different RP's for a range(s) group(s). This is accomplished via access-lists. This permits the RP's to be placed in different locations and can improve the traffic flow for the group if it is placed close to the Source(s).
- **RP Configuration**
 - RP's may be configured statically on each router (although they must all agree or your network will be broken!) in your network. However, a better solution is to use the Auto-RP or PIMv2 mechanisms to configure RP's.
- **Data Forwarding**
 - Multicast traffic forwarding In a PIM Sparse mode network is first attempted using any matching (S,G) entries in the Multicast Routing table. If no matching (S,G) state exists, then the traffic is forwarded using the matching (*,G) entry in the Multicast Routing table.
- **PIM Sparse Mode Spec**
 - PIM Sparse mode is now defined in RFC 2362.

PIM-SM Shared Tree Joins

Cisco.com



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

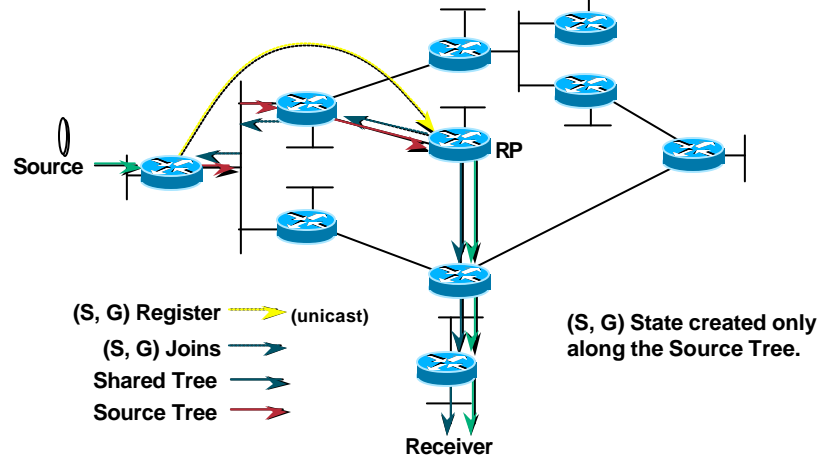
6

- PIM-SM Shared Tree Joins

- In this example, there is an active receiver (attached to leaf router at the bottom of the drawing) has joined multicast group “G”.
- The leaf router knows the IP address of the Rendezvous Point (RP) for group G and when it sends a (*,G) Join for this group towards the RP.
- This (*, G) Join travels hop-by-hop to the RP building a branch of the Shared Tree that extends from the RP to the last-hop router directly connected to the receiver.
- At this point, group “G” traffic can flow down the Shared Tree to the receiver.

PIM-SM Sender Registration

Cisco.com



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

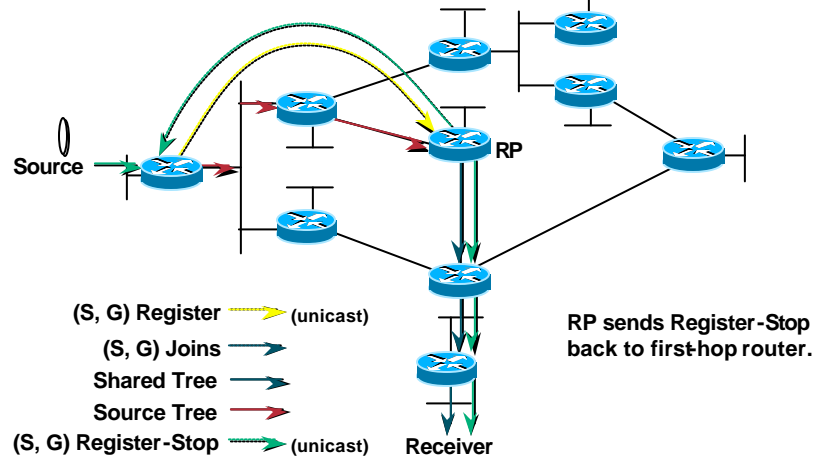
7

• PIM-SM Sender Registration

- As soon as an active source for group G sends a packet the leaf router that is attached to this source is responsible for “Registering” this source with the RP and requesting the RP to build a tree back to that router.
- The source router encapsulates the multicast data from the source in a special PIM SM message called the Register message and unicasts that data to the RP.
- When the RP receives the Register message it does two things
 - It de-encapsulates the multicast data packet inside of the Register message and forwards it down the Shared Tree.
 - The RP also sends an (S,G) Join back towards the source network S to create a branch of an (S, G) Shortest-Path Tree. This results in (S, G) state being created in all the router along the SPT, including the RP.

PIM-SM Sender Registration

Cisco.com



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

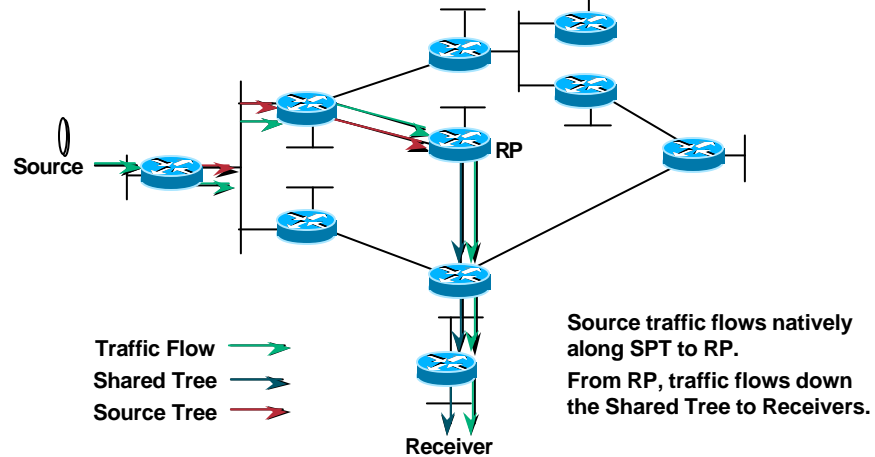
8

- **PIM-SM Sender Registration (cont.)**

- As soon as the SPT is built from the Source router to the RP, multicast traffic begins to flow natively from source S to the RP.
- Once the RP begins receiving data natively (i.e. down the SPT) from source S it sends a 'Register Stop' to the source's first hop router to inform it that it can stop sending the unicast Register messages.

PIM-SM Sender Registration

Cisco.com



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

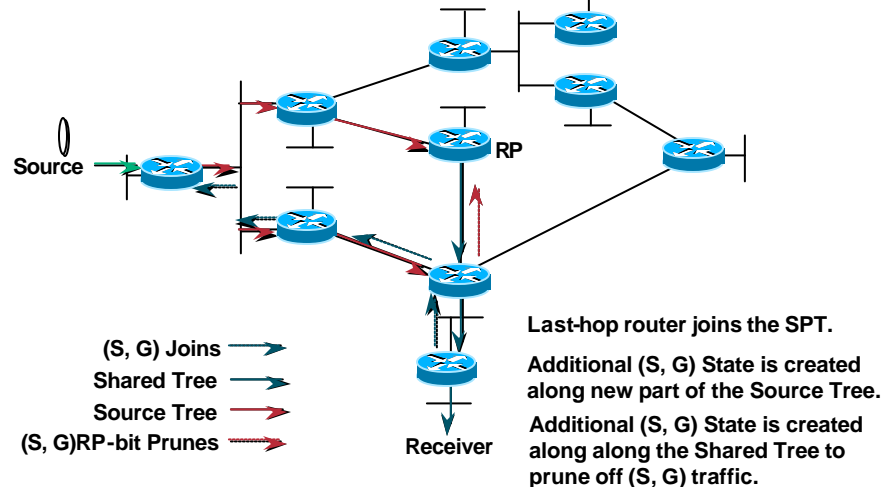
9

- **PIM-SM Sender Registration (cont.)**

- At this point, multicast traffic from the source is flowing down the SPT to the RP and from there, down the Shared Tree to the receiver.

PIM-SM SPT Switchover

Cisco.com



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

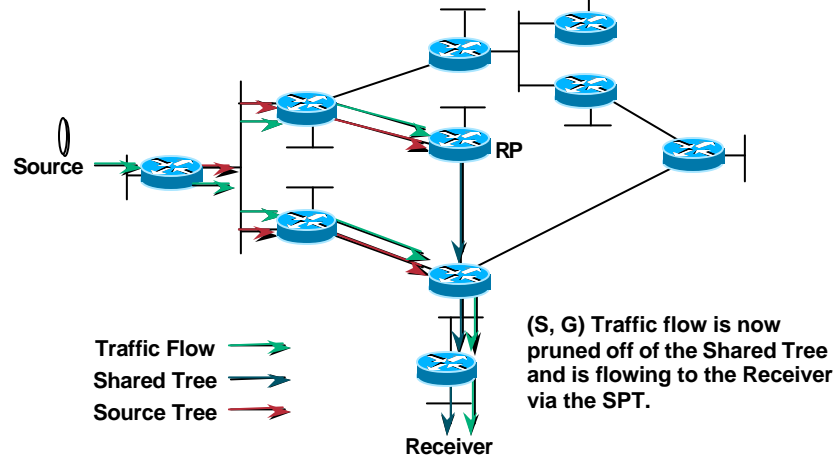
10

• PIM-SM Shortest-Path Tree Switchover

- PIM-SM has the capability for last-hop routers (i.e. routers with directly connected members) to switch to the Shortest-Path Tree and bypass the RP if the traffic rate is above a set threshold called the “SPT-Threshold”.
 - The default value of the SPT-Threshold” in Cisco routers is zero. This means that the default behaviour for PIM-SM leaf routers attached to active receivers is to immediately join the SPT to the source as soon as the first packet arrives via the (*,G) shared tree.
- In the above example, the last-hop router (at the bottom of the drawing) sends an (S, G) Join message toward the source to join the SPT and bypass the RP.
- This (S, G) Join messages travels hop-by-hop to the first-hop router (i.e. the router connected directly to the source) thereby creating another branch of the SPT. This also creates (S, G) state in all the routers along this branch of the SPT.
- Finally, special (S, G)RP-bit Prune messages are sent up the Shared Tree to prune off this (S,G) traffic from the Shared Tree.
 - If this were not done, (S, G) traffic would continue flowing down the Shared Tree resulting in duplicate (S, G) packets arriving at the receiver.

PIM-SM SPT Switchover

Cisco.com



Module5.ppt

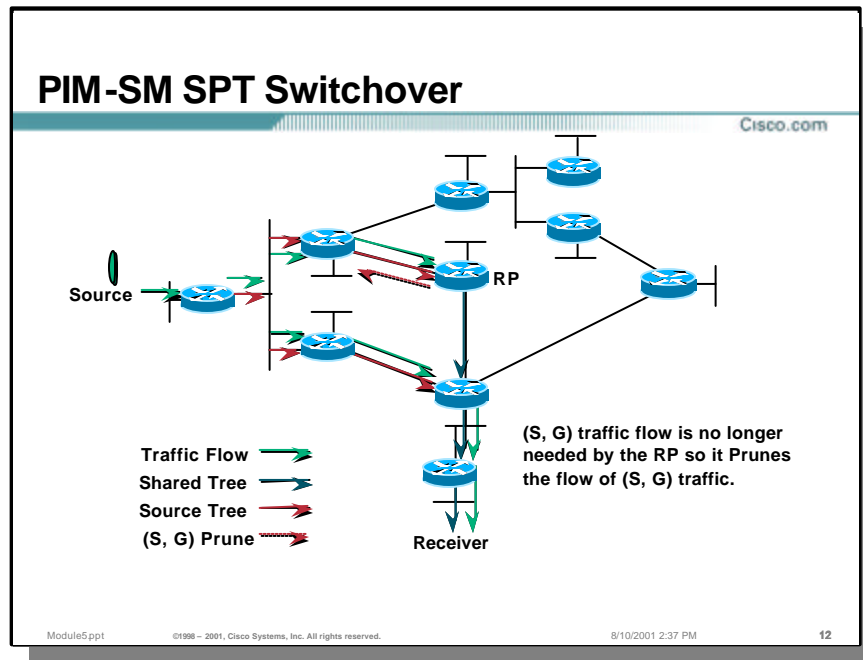
©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

11

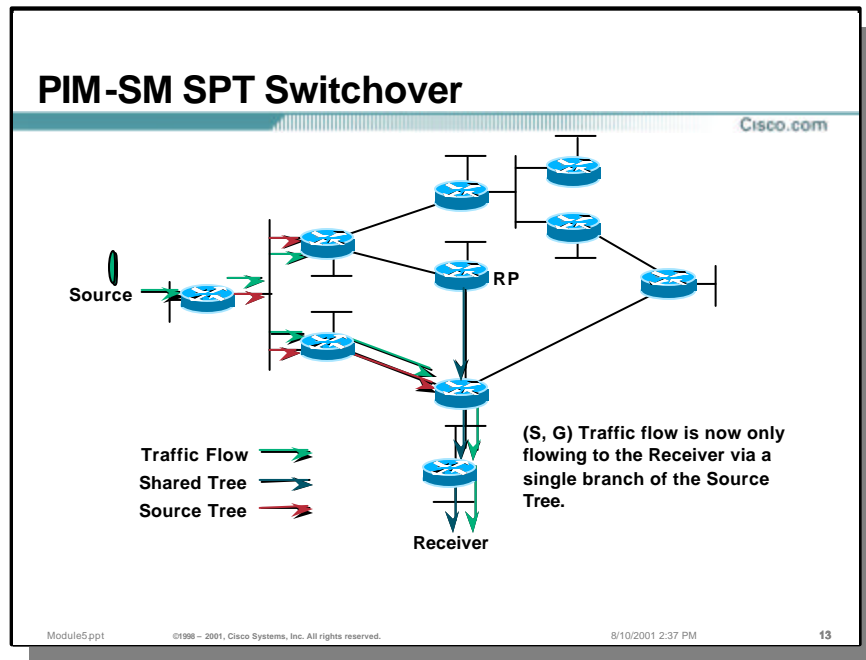
- **PIM-SM Shortest-Path Tree Switchover**

- At this point, (S, G) traffic is now flowing directly from the first-hop router to the last-hop router and from there to the receiver.
 - **Note:** The RP will normally send (S, G) Prunes back toward the source to shutoff the flow of now unnecessary (S, G) traffic to the RP *IFF* it has received an (S, G)RP-bit Prune on all interfaces on the Shared Tree. (This step has been omitted from the example above.)
- As a result of this SPT-Switchover mechanism, PIM SM also supports the construction and use of SPT (S,G) trees but in a much more economical fashion than PIM DM in terms of forwarding state.



- **PIM-SM Shortest-Path Tree Switchover**

- At this point, the RP no longer needs the flow of (S, G) traffic since all branches of the Shared Tree (in this case there is only one) have pruned off the flow of (S, G) traffic.
- As a result, the RP will send (S, G) Prunes back toward the source to shutoff the flow of the now unnecessary (S, G) traffic to the RP
 - Note: This will occur *IFF* the RP has received an (S, G)RP-bit Prune on all interfaces on the Shared Tree.



- **PIM-SM Shortest-Path Tree Switchover**

- As a result of the SPT-Switchover, (S, G) traffic is now only flowing from the first-hop router to the last-hop router and from there to the receiver. Notice that traffic is no longer flowing to the RP.
- As a result of this SPT-Switchover mechanism, it is clear that PIM SM also supports the construction and use of SPT (S,G) trees but in a much more economical fashion than PIM DM in terms of forwarding state.

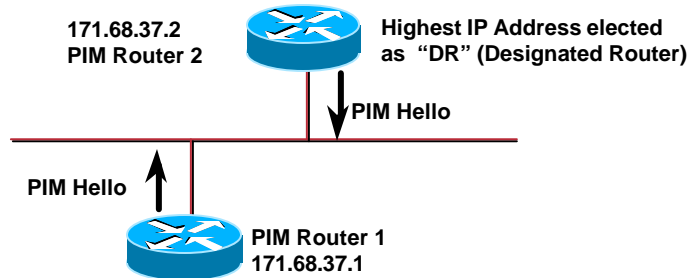
PIM-SM Protocol Mechanics

Cisco.com

- **PIM Neighbor Discovery**
- PIM State
- PIM SM Joining
- PIM SM Registering
- PIM SM SPT-Switchover
- PIM SM Pruning
- PIM SM State Maintenance

PIM Neighbor Discovery

Cisco.com



- PIMv2 Hellos are periodically multicast to the "All-PIM-Routers" (224.0.0.13) group address. (Default = 30 seconds)
 - Note: PIMv1 multicasts PIM Query messages to the "All-Routers" (224.0.0.2) group address.
- If the "DR" times-out, a new "DR" is elected.
- The "DR" is responsible for sending all Joins and Register messages for any receivers or senders on the network.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

15

• PIM Neighbor Discovery

- PIM Hellos are sent periodically to discover the existence of other PIM routers on the network and to elect the Designated Router.
- For Multi-Access networks (e.g. Ethernet), the PIM Hello messages are multicast to the "All-PIM-Routers" (224.0.0.13) multicast group address.

• Designated Router (DR)

- For multi-access networks, a Designated Router (DR) is elected. In PIM Sparse mode networks, the DR is responsible for sending Joins to the RP for members on the multi-access network and for sending Registers to the RP for sources on the multi-access network. For Dense mode, the DR has no meaning. The exception to this is when IGMPv1 is in use. In this case, the DR also functions as the IGMP Querier for the Multi-Access network.

• Designated Router (DR) Election

- To elect the DR, each PIM node on a multi-access network examines the received PIM Hello messages from its neighbors and compares the IP Address of its interface with the IP Address of its PIM Neighbors. The PIM Neighbor with the highest IP Address is elected the DR.
- If no PIM Hellos have been received from the elected DR after some period (configurable), the DR Election mechanism is run again to elect a new DR.

PIM Neighbor Discovery

Cisco.com

```
wan-gw8>show ip pim neighbor
PIM Neighbor Table
Neighbor Address  Interface      Uptime    Expires    Mode
171.68.0.70       FastEthernet0  2w1d     00:01:24   Sparse
171.68.0.91       FastEthernet0  2w6d     00:01:01   Sparse (DR)
171.68.0.82       FastEthernet0  7w0d     00:01:14   Sparse
171.68.0.86       FastEthernet0  7w0d     00:01:13   Sparse
171.68.0.80       FastEthernet0  7w0d     00:01:02   Sparse
171.68.28.70      Serial2.31     22:47:11 00:01:16   Sparse
171.68.28.50      Serial2.33     22:47:22 00:01:08   Sparse
171.68.27.74      Serial2.36     22:47:07 00:01:21   Sparse
171.68.28.170     Serial0.70     1d04h    00:01:06   Sparse
171.68.27.2       Serial1.51     1w4d     00:01:25   Sparse
171.68.28.110     Serial3.56     1d04h    00:01:20   Sparse
171.68.28.58      Serial3.102    12:53:25 00:01:03   Sparse
```

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

16

- **“show ip pim neighbor” command output**

- Neighbor Address - the IP address of the PIM Neighbor
- Interface - the interface where the PIM Hello of this neighbor was received.
- Uptime - the period of time that this PIM Neighbor has been active.
- Expires - the period of time after which this PIM Neighbor will no longer be considered as active. (Reset by the receipt of a another PIM Query.)
- Mode - PIM mode (Sparse, Dense, Sparse/Dense) that the PIM Neighbor is using.
- “(DR)” - Indicates that this PIM Neighbor is the Designated Router for the network.

PIM-SM Protocol Mechanics

Cisco.com

- PIM Neighbor Discovery
- **PIM State**
- PIM SM Joining
- PIM SM Registering
- PIM SM SPT-Switchover
- PIM SM Pruning
- PIM SM State Maintenance

PIM State

Cisco.com

- Describes the “state” of the multicast distribution trees as understood by the router at this point in the network.
- Represented by entries in the multicast routing (mroute) table
 - Used to make multicast traffic forwarding decisions
 - Composed of (*, G) and (S, G) entries
 - Each entry contains RPF information
 - Incoming (i.e. RPF) interface
 - RPF Neighbor (upstream)
 - Each entry contains an Outgoing Interface List (OIL)
 - OIL may be NULL

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

18

• PIM State

- In general, Multicast State basically describes the multicast distribution tree as it is understood by the router at this point in the network.
- However to be completely correct, “Multicast State” describes the multicast traffic “forwarding” state that is used by the router to forward multicast traffic.

• Multicast Routing (mroute) Table

- Multicast “state” is stored in the multicast routing (mroute) table and which can be displayed using the show ip mroute command.
- Entries in the mroute table are composed of (*, G) and (S, G) entries each of which contain:
 - RPF Information consisting of an Incoming (or RPF) interface and the IP address of the RPF (i.e. upstream) neighbor router in the direction of the source. (In the case of PIM-SM, this information in a (*, G) entry points toward the RP. PIM-SM will be discussed in a later module.)
 - Outgoing Interface List (OIL) which contains a list of interfaces that the multicast traffic is to be forwarded. (Multicast traffic must arrive on the Incoming interface before it will be forwarded out this interfaces. If multicast traffic does not arrive on the Incoming interface, it is simply discarded.)

PIM-SM State Example

Cisco.com

```
sj-mbone> show ip mroute
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, C - Connected, L - Local, P - Pruned
       R - RP-bit set, F - Register flag, T - SPT-bit set, J - Join SPT
       M - MSDP created entry, X - Proxy Join Timer Running
       A - Advertised via MSDP
Timers: Uptime/Expires
Interface state: Interface, Next-Hop or VCD, State/Mode

(*, 224.1.1.1), 00:13:28/00:02:59, RP 10.1.5.1, flags: SCJ
  Incoming interface: Ethernet0, RPF nbr 10.1.2.1,
  Outgoing interface list:
    Ethernet1, Forward/Sparse, 00:13:28/00:02:32
    Serial0, Forward/Sparse, 00:4:52/00:02:08

(171.68.37.121/32, 224.1.1.1), 00:01:43/00:02:59, flags: CJT
  Incoming interface: Serial0, RPF nbr 192.10.2.1
  Outgoing interface list:
    Ethernet1, Forward/Sparse, 00:01:43/00:02:11
    Ethernet0, forward/Sparse, 00:01:43/00:02:11
```

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

19

• PIM-SM State Example

- **(*, G) Entry** - The (*, 224.1.1.1) entry shown in sample output of the **show ip mroute** command is the (*, G) entry. If there is no matching entry for a particular (S, G) entry, this entry is used to forward traffic down the Shared Tree.
 - The **Expires** countdown timer in the first line of the (*, G) entry which shows when the entry will expire and be deleted. This entry will remain at roughly 3 minutes as long as there is an interface in the Outgoing Interface list.
 - The **Incoming interface** information is used to RPF check arriving (*, G) multicast traffic and is computed in the direction of the RP (in this case, 10.1.5.1.).
 - The **Outgoing Interface list** which reflects the interfaces where (*,G) Joins have been received or where directly connected members of group "G" reside. Traffic flowing down the Shared Tree are forwarded out these interfaces. The **Expires** countdown timers on these interfaces are reset to 3 minutes by the receipt of periodic (*, G) Joins. If the count ever reaches zero, the entry in the OIL is deleted.
- **(S, G) Entry** - The (128.9.160.43/32, 224.1.1.1) entry is an example of an (S, G) entry in the mroute table. This entry is used to forward any multicast traffic sent by source 128.9.160.43 to group 224.1.1.1. Notice the following:
 - The **Expires** countdown timer in the first line of the (S, G) entry which shows when the entry will expire and be deleted. This entry is reset to 3 minutes whenever an (S, G) multicast packet is forwarded.
 - The **Incoming interface** information is used to RPF check arriving (S, G) multicast traffic. If a packet does not arrive via this interface, the packet is discarded.
 - The **Outgoing Interface list** which reflects the interfaces where (S,G) packets are to be forwarded.

PIM-SM (*,G) State Rules

Cisco.com

- (*,G) creation
 - Receipt of a (*,G) Join or IGMP Report
 - Automatically if (S,G) must be created
- (*,G) reflects default group forwarding
 - IIF = RPF interface toward RP
 - OIL = interfaces
 - that received a (*,G) Join or
 - with directly connected members or
 - manually configured
- (*,G) deletion
 - When OIL = NULL and
 - no child (S,G) state exists

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

20

• PIM-SM (*,G) State Rules

- A (*, G) entry is created when a (*, G) Join or an IGMP Report is received
 - The later condition can be simulated by manually configuring the interface to join the group.
- (*, G) entries are also automatically created whenever an (S, G) entry for the group must be created.
 - The (*, G) entry is created first and then the (S, G) entry. The reason for this will become clear shortly.
- The IIF reflects the RPF interface/neighbor in the direction of the RP.
- The OIL of a PIM-SM (*, G) entry reflects interfaces that:
 - Have received a (*, G) Join or
 - Where a directly connected member has joined the group
 - The interface was manually configured to join the group. (Note: This may be accomplished using the **ip igmp static-group <group>** command.)
- (*, G) entries are deleted when its **Expires** timer counts down to zero. This will only occur when:
 - The OIL is Null and
 - No child (S, G) entry exists

PIM-SM (S,G) State Rules

Cisco.com

- **(S,G) creation**
 - By receipt of (S,G) Join or Prune or
 - By “Register” process
 - Parent (*,G) created (if doesn't exist)
- **(S,G) reflects forwarding of “S” to “G”**
 - IIF = RPF Interface normally toward source
 - RPF toward RP if “RP-bit” set
 - OIL = Initially, copy of (*,G) OIL minus IIF
- **(S,G) deletion**
 - By normal (S,G) entry timeout

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

24

• PIM-SM (S, G) Rules

- In PIM-SM, (S, G) state is created as a result of:
 - The receipt of an (S, G) Join or Prune or
 - The PIM-SM Register process which is triggered by a first-hop router receiving a packet from a directly connected source.
- When an (S, G) entry must be created, the following steps occur:
 - If a corresponding (*, G) entry does not exist, it is created first.
 - The RPF Information is computed for the source “S”. This information is stored in the (S, G) entry as the Incoming interface and the RPF neighbor (i.e. the PIM neighbor in the direction of the source).

The exception to this rule is if the RP-bit is set in the (S, G) entry, the RPF interface is pointed up the Shared Tree. This mechanism allows duplicate (S, G) traffic to be blocked from flowing down the Shared Tree after a downstream router has switched to the Shortest Path Tree. (More on this later.)
 - The OIL of the (S, G) entry is populated with a copy of the OIL from the parent (*, G) entry less the Incoming interface. (The Incoming interface must not appear in the OIL otherwise a multicast route loop could occur.)
- In PIM-SM, (S, G) entries are deleted when their **Expires** timer counts down to zero. The **Expires** timer is reset whenever an (S, G) packet is received and forwarded.

PIM-SM OIL Rules

Cisco.com

- **Interfaces in OIL added**
 - **By receipt of Join message**
 - Interfaces added to (*,G) are added to all (S,G)'s
- **Interfaces in OIL removed**
 - **By receipt of Prune message**
 - Interfaces removed from (*,G) are removed from all (S,G)'s
 - **Interface Expire timer counts down to zero**
 - Timer reset (to 3 min.) by receipt of periodic Join
or
 - By IGMP membership report

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

22

• PIM-SM Outgoing Interface List Rules

- Adding an interface
 - Interfaces are added to an (S, G) OIL when a (S, G) Join message is received on an interface.
 - Interfaces are added to the (*, G) OIL when a (*, G) Join message is received on an interface.
 - Anytime an interface is added to the (*, G) OIL, the interface is added to the OIL of all associated (S, G) OIL's. (Note: A check is always made to prevent the IIF from appearing in the OIL.)
- Removing an interface
 - Interfaces are removed from the OIL of a (*, G) or (S, G) entry if the interface's **Expires** timer counts down to zero.
Note: The interface **Expires** timer is reset to 3 minutes by the receipt of periodic Join messages sent by downstream routers once per minute or by an IGMP Report sent by a directly connected member on the interface.
 - Interfaces are removed from the OIL if an Prune message is received (and it is not overridden by another router if the interface is a multi-access network).
 - Interfaces removed from a (*, G) OIL, are removed from the OIL of all associated (S, G) OIL's.

PIM-SM OIL Rules

Cisco.com

- **Triggering Join/Prune Messages**
 - **(*;G) Joins are triggered when:**
 - The (*,G) OIL transitions from Null to non-Null
 - **(*;G) Prunes are triggered when:**
 - The (*,G) OIL transitions from non-Null to Null
 - **(S,G) Joins are triggered when:**
 - The (S,G) OIL transitions from Null to non-Null
 - **(S,G) Prunes are triggered when:**
 - The (S,G) OIL is Null AND
 - A packet is received on the incoming interface
 - **(S,G)RP-bit Prunes are triggered when:**
 - The (S,G) RPF info != the (*,G) RPF info

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

23

• PIM-SM Outgoing Interface List Rules

- Triggering Join/Prune Messages
 - (*,G) Joins are triggered whenever the (*,G) OIL is empty (Null) and an interface is added making the OIL non-Null.
 - (*,G) Prunes are triggered whenever the last interface is removed from the (*,G) OIL.
 - (S,G) Joins are triggered whenever the (S,G) OIL is empty (Null) and an interface is added making the OIL non-Null.
 - (S,G) Prunes are triggered whenever the (S,G) OIL is empty AND a packet is received on the incoming interface.

Note: This is an optimization that attempts to minimize the sending of (S,G) Prunes. Instead of sending the (S,G) Prune immediately when the last interface is removed, the state is just allowed to time out. However, if (S,G) traffic is still flowing, then the arrival of the next (S,G) packet will cause the prune to be sent.

- (S,G)RP-bit Prunes are sent whenever the (S,G) RPF information (incoming interface and RPF-neighbor) is not the same as the (*,G) RPF information. This indicates that the SPT and the Shared-Tree diverge at this point and that (S,G) traffic should be pruned from the Shared-Tree.

PIM-SM State Flags

Cisco.com

- **S = Sparse Mode**
- **C = Directly Connected Host**
- **L = Local (Router is member)**
- **P = Pruned (All intfcs in OIL = Prune)**
- **T = Forwarding via SPT**
 - Indicates at least one packet was forwarded

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

24

• PIM-SM State Flags

- “S” Flag ((*, G) entries only)
 - Indicates the group is operating in Sparse mode. (Appears only on (*, G) entries.)
- “C” Flag
 - Indicates that there is a member of the group directly connected to the router.
- “L” Flag
 - Indicates the router itself is a member of this group and is receiving the traffic. (This would be the case for the Auto-RP Discovery group 224.0.1.40 which all Cisco routers join automatically.)
- “P” Flag
 - Set whenever all interfaces in the outgoing interface list of an entry are Pruned (or the list is Null). This general means that the router will send Prune messages to the RPF neighbor to try to shutoff this traffic.)
- “T” Flag ((S, G) entries only)
 - Indicates that at least one packet was received via the SPT

PIM-SM State Flags (cont.)

Cisco.com

- **J = Join SPT**
 - In (*, G) entry
 - Indicates SPT-Threshold is being exceeded
 - Next (S,G) received will trigger join of SPT
 - In (S, G) entry
 - Indicates SPT joined due to SPT-Threshold
 - If rate < SPT-Threshold, switch back to Shared Tree
- **F = Register**
 - In (S,G) entry
 - “S” is a directly connected source
 - Triggers the Register Process
 - In (*, G) entry
 - Set when “F” set in at least one child (S,G)

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

25

• PIM-SM State Flags

- “J” Flag (Join SPT)
 - When this flag is set in a (*, G) entry, it indicates that the rate of traffic flowing down the Shared Tree is above the SPT-Threshold and will cause a switch to the SPT for the next packet received down the shared tree. (More on this later.)
 - When this flag is set in an (S, G) entry, it indicates that the (S, G) entry (and hence the SPT) was created as a result of the SPT-Threshold being exceeded. If the rate of this (S, G) traffic drops back below the SPT, the router will attempt to switch this traffic flow back to the Shared Tree.
- “F” Flag (Register)
 - This flag is set on an (S, G) entry when source “S” is directly connected to the router. This indicates that this router is a “first-hop” router and triggers it to send Register messages to the RP to inform the RP of this active source.
 - This flag can also be set for arriving (S, G) entries created at a border router such as a router that borders on a DVMRP or other dense mode cloud. This causes the router to perform a proxy-register operation and send (S, G) Register messages to the RP on behalf of the downstream DVMRP routers. This proxy-register operation follows the same rules as for directly connected sources.
 - The “F” flag is also set on a (*, G) entry if any associated (S, G) entries have the “F” flag set.

PIM-SM State Flags (cont.)

Cisco.com

- **R = RP bit**
 - (S, G) entries only
 - Set by (S,G)RP-bit Prune
 - Indicates info is applicable to Shared Tree
 - Used to prune (S,G) traffic from Shared Tree
 - Initiated by Last-hop router after switch to SPT
 - Modifies (S,G) forwarding behavior
 - IIF = RPF toward RP (i.e. up the Shared Tree)
 - OIL = Pruned accordingly

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

26

• PIM-SM Flags

- “R” Flag (RP-Bit)
 - This flag is set on (S, G) entries only and indicates that the (S, G) forwarding information in the entry is applicable to (S, G) traffic flowing down the Shared Tree.
 - The “R” flag is set on an (S, G) entry by the receipt of an (S, G)RP-bit Prune message. These messages are sent by downstream routers on the Shared Tree that are requesting that this specific (S, G) traffic flow be pruned off of the Shared Tree. This is done to eliminate duplicate (S, G) traffic after a downstream router has switched to the (S, G) Shortest-Path Tree.
 - Whenever the “R” flag is set on an (S, G) entry, the RPF information must be changed to point toward the RP instead of pointing at source “S”. This is done because the (S, G) entry is now applicable to (S, G) traffic arriving *down the Shared Tree*. As a result, the RPF information must point up the Shared Tree in order for arriving (S, G) packets to RPF correctly. (This should be made clear later.)

PIM-SM State Flags (cont.)

Cisco.com

- **X = Proxy Join Timer flag**
 - (S, G) entries only
 - Indicates Proxy Join Timer is running
 - Used to handle turn-around router case
 - More on this in another Module
 - When Proxy Join Timer is running
 - (S, G) Joins are sent toward the source
 - The sending of (S, G) Prunes are suppressed
 - Even if the OIL list is NULL

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

27

• PIM-SM Flags

- “X” Flag (Proxy Join Timer Running)
 - This flag is set on (S, G) entries only and is used to indicate that the “Proxy Join Timer” is running. When this timer is running, the router will continue to send (S, G) Joins in the direction of the source *even if the OIL is NULL*.
This is used to handle the special turn-around router situation which occurs when the SPT to the RP and the Shared Tree merge. (More on this special scenario will be presented in another module.)

PIM-SM State Flags (cont.)

Cisco.com

- **M = MSDP Created bit**
 - (S, G) entries only
 - Set when (S, G) learned via an MSDP SA msg
- **A = MSDP Advertise bit**
 - (S, G) entries only
 - (S, G) may be advertised in an MSDP SA msg
 - Presence of certain filters can affect this.
 - Indicates source is in local SM domain
 - Received a PIM (S,G) Register or
 - Source is directly connected or
 - (S,G) traffic was received on a DM interface
 - via the RPF interface

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

28

• PIM-SM Flags

- “M” Flag (MSDP Created)
 - This flag only appears on (S, G) entries and only on the router that is the active RP for group “G”.
 - The flag indicates that the RP has learned of this particular source via an MSDP “Source Active” message. (MSDP is addressed in more detail in another module.)
- “A” Flag (Advertise Flag)
 - This flag only appears on (S, G) entries and only on the router that is the active RP for group “G”.
 - The “A” flag indicates that this source is in the local PIM-SM domain and that it is a *candidate* for being announced to RP’s in other networks via MSDP “Source Active” messages.
 - A source is considered to be in the local domain if an (S, G) Register message was received for this source or the source is directly connected to the RP or the (S, G) traffic was received on a Dense mode interface that has been designated as a dense mode boundary interface.

PIM-SM Protocol Mechanics

Cisco.com

- PIM Neighbor Discovery
- PIM State
- **PIM SM Joining**
- PIM SM Registering
- PIM SM SPT-Switchover
- PIM SM Pruning
- PIM SM State Maintenance

PIM SM Joining

Cisco.com

- **Leaf routers send a (*,G) Join toward RP**
 - Joins sent hop-by-hop along path toward RP
- **Each router along path creates (*,G) state**
 - **IF no (*,G) state,**
 - Create it and send a Join toward RP
 - **ELSE**
 - Join process complete. Reached the Shared Tree.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

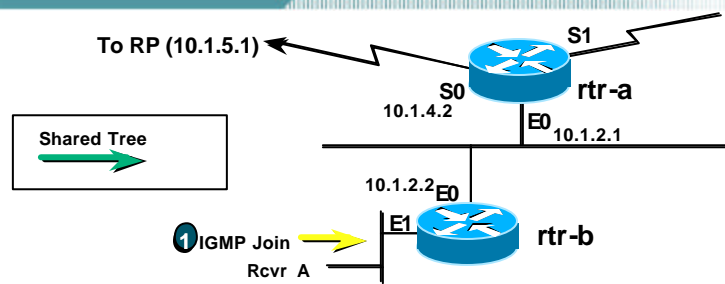
8/10/2001 2:37 PM

30

- **Leaf (last-hop) routers join the Shared Tree (RPT)**
 - When a last-hop router wishes to begin receiving multicast traffic for group “G”, it sends a PIM (*,G) Join message to its up-stream PIM Neighbor in the direction of the RP.
 - While the Join is multicast to the “All-Routers” (224.0.0.2) multicast address, the up-stream PIM Neighbor’s IP address is indicated in the body of the PIM Join Message. This permits all PIM routers on a Multi-Access network to be aware of the Join but only the indicated up-stream PIM Neighbor will perform the Join.
- **Routers up the Shared Tree (RPT) create (*,G) state**
 - When a PIM router receives a (*,G) Join for group “G” from one of its down-stream PIM Neighbors, it will check to see if any (*, G) state exists for group “G” in its Multicast Routing table.
 - If (*, G) state for group “G” already exists, then the interface from which the Join was received is placed on the (*,G) oolist.
 - If no (*, G) state for group “G” exists, a (*, G) entry is created, the interface from which the Join was received is placed in the (*, G) oolist and a (*, G) Join is sent towards the RP.
 - The end result of the above mechanism is to create (*, G) state all the way from the last-hop router to the RP so that group “G” multicast traffic will flow down the Shared Tree (RPT) to the last-hop router.

PIM SM Joining

Cisco.com



1 "Rcvr A" wishes to receive group G traffic. Sends IGMP Join for G.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

31

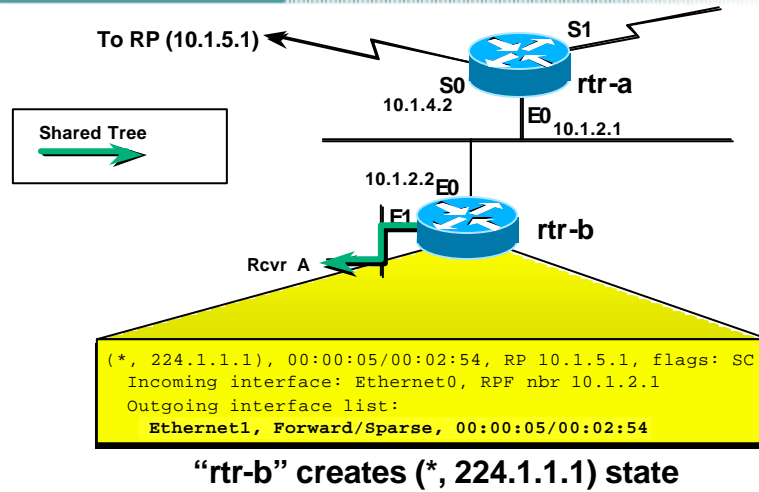
• PIM SM Joining Example

1) Receiver "A" wishes to receive group "G" multicast traffic and therefore sends an IGMP Host Membership message (sometimes loosely referred to as an IGMP Join) which is received by "rtr-b".

"rtr-b" has no existing (*, G) state for group "G" and therefore creates an entry. (See next slide.)

PIM SM Joining

Cisco.com



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

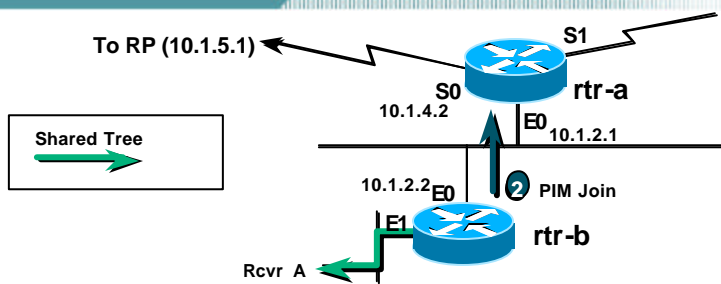
32

• State in “rtr-b” after Joining (*, 224.1.1.1)

- (*, 224.1.1.1)
 - indicates the (*, G) entry.
- 00:00:05/00:02:54
 - indicates that the entry has existed for 5 seconds and will expire in 2 minutes and 54 seconds.
- RP 10.1.5.1
 - is the IP Address of the Rendezvous Point for Group 224.1.1.1
- flags: SC
 - indicates that this is a Sparse mode group (S) and that there is a member of this group directly connected (C) to the router.
- Incoming interface: Ethernet0, RPF nbr 10.1.2.1
 - indicates the Incoming interface (up the Shared Tree toward RP) and
 - the RPF neighbor’s IP address (in the direction of the RP) is 10.1.2.1
- Outgoing interface list:
 - lists the interfaces that are in the outgoing interface list for this entry.
- Ethernet1, Forward/Sparse, 00:00:05/00:02:54
 - indicates Ethernet 1 is in the oilist; it’s in the “Forward” state; Sparse mode and that it has been in the list for 5 seconds and will expire in 2 minutes and 54 seconds if no further (*, G) Join or IGMP Report is received on this interface.

PIM SM Joining

Cisco.com



- 1 "Rcvr A" wishes to receive group G traffic. Sends IGMP Join for G.
- 2 "rtr-b" sends (*,G) Join towards RP.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

33

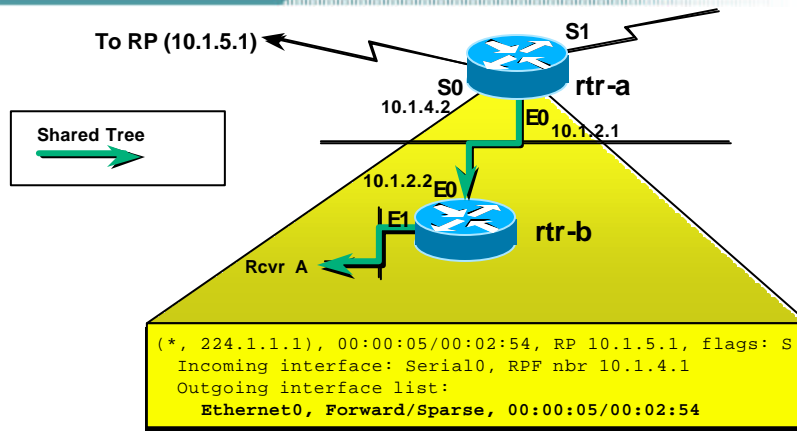
• PIM SM Joining Example

2) Because the OIL of the (*, G) transitioned from Null to non-Null (when "rtr-b" added Ethernet 1 to the OIL of the newly created entry), a PIM (*, G) Join is sent to rtr-b's up-stream PIM neighbor (rtr-a) in the direction of the RP.

When "rtr-a" receives the (*, G) Join it creates (*, G) state. (See next slide.)

PIM SM Joining

Cisco.com



“rtr-a” creates (*, 224.1.1.1) state.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

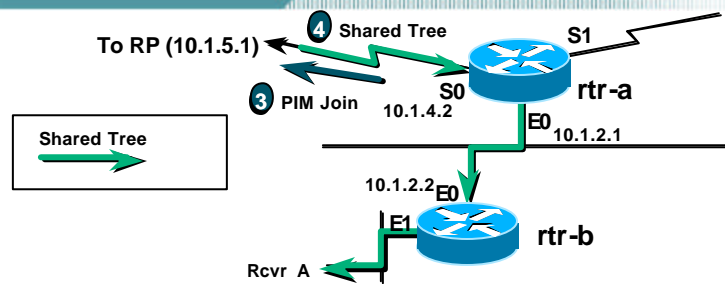
34

• State in “rtr-a” after Joining (*, 224.1.1.1)

- (*, 224.1.1.1)
 - indicates the (*, G) entry.
- 00:00:05/00:02:54
 - indicates that the entry has existed for 5 seconds and will expire in 2 minutes and 54 seconds.
- RP 10.1.5.1
 - is the IP Address of the Rendezvous Point for Group 224.1.1.1
- flags: S
 - indicates that this is a Sparse mode group (S).
- Incoming interface: Serial0, RPF nbr 10.1.4.1
 - indicates the Incoming interface (up the Shared Tree toward RP) and
 - the RPF neighbor’s IP address (in the direction of the RP) is 10.1.4.1
- Outgoing interface list:
 - lists the interfaces that are in the outgoing interface list for this entry.
- Ethernet0, Forward/Sparse, 00:00:05/00:02:54
 - indicates Ethernet 0 is in the oilist; it’s in the “Forward” state; Sparse mode and that it has been in the list for 5 seconds and will expire in 2 minutes and 54 seconds if no further (*, G) Join or IGMP Report is received on this interface.

PIM SM Joining

Cisco.com



- 1 "Rcvr A" wishes to receive group G traffic. Sends IGMP Join for G.
- 2 "rtr-b" sends (*,G) Join towards RP.
- 3 "rtr-a" sends (*,G) Join towards RP.
- 4 Shared tree is built all the way back to the RP.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

35

• PIM SM Joining Example

- 3) Because the OIL of the (*, G) transitioned from Null to non-Null (when "rtr-a" added Ethernet 0 to the OIL of the newly created entry), a PIM (*, G) Join is sent to rtr-a's up-stream PIM neighbor in the direction of the RP.

When the upstream router receives the (*, G) Join it too creates (*, G) state and creates a branch of the Shared Tree.

- 4) This process continues all the way back to the RP (or until a router is reached that is already on the Shared Tree and therefore already has a (*, G) entry.)

PIM-SM Protocol Mechanics

Cisco.com

- PIM Neighbor Discovery
- PIM State
- PIM SM Joining
- **PIM SM Registering**
- PIM SM SPT-Switchover
- PIM SM Pruning
- PIM SM State Maintenance

PIM SM Registering

Cisco.com

- **Senders begin sourcing Multicast Traffic**
 - Senders don't necessarily perform IGMP group joins.
- **1st-hop router unicasts "Registers" to RP**
 - A Mcast packet is encapsulated in each Register msg
 - Registers messages follow unicast path to RP
- **RP receives "Register" messages**
 - De-encapsulates Mcast packet inside Register msg
 - Forwards Mcast packet down Shared Tree
 - Sends (S,G) Join toward Source to build a SPT from the Source to the RP

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

37

- **All Senders are not necessarily Receivers and vice versa.**
 - It is not a requirement that all sources be receivers. In the case of a source-only host, it is permissible for the host to simply begin sending multicast traffic without ever joining the group via IGMP.
- **1st-hop router sends "Registers" to the RP**
 - In PIM Sparse mode, when a 1st-hop router receives the first multicast packet from directly connected source "S" for group "G", it creates (S, G) state and sets the "F" bit in the (S, G) entry to indicate that it is a directly connected "Source" and also sets the "Registering" flag to indicate that it's in the process of "Registering".
 - Next, the 1st-hop router encapsulates the original multicast packet in a PIM Register message and unicasts it to the RP. (Any subsequent multicast packets received from directly connected source "S" for group "G" are also encapsulated in a Register message and unicast to the RP. This continues until an (S, G) "Register-Stop" message is received from the RP.)
- **RP receives "Register messages"**
 - When the RP receives a "Register" message it will de-encapsulated the message. If this packet is to a Group for which the RP has (*, G) state, the RP will:
 - Forward the original packet out all interfaces in the the (*, G) entry's "oilst".
 - If it hasn't already done so, the RP creates (S, G) state and sends an (S, G) Join back towards the Source in order to join the Shortest-path Tree (SPT) to Source "S".

PIM SM Registering (cont.)

Cisco.com

- **1st-hop router receives (S,G) Join**
 - SPT between Source and RP is now built.
 - Begins forwarding traffic down SPT to RP
 - (S,G) Traffic temporarily flowing down 2 paths to RP
- **RP receives traffic down native SPT**
 - Sends a “Register-Stop” msg to the 1st-Hop router.
- **1st-hop router receives “Register-Stop”**
 - Stops encapsulating traffic in “Register” messages
 - (S,G) Traffic now flowing down single SPT to RP

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

38

- **1st-hop router receives (S, G) Join**
 - When the 1st-hop router receives the (S, G) Join (sent hop-by-hop from the RP), it processes it normally by adding the interface, from which the Join was received, to the “olist” of the existing (S, G) entry. (This entry was originally created when the 1st-hop router received the first multicast packet from directly connected Source “S”.) This completes the building of the Shortest-Path Tree (SPT) from the Source to the RP.
 - The 1st-hop router now begins forwarding Source “S” multicast traffic down the newly built Shortest-Path Tree (SPT) to the RP.
 - Note: (S, G) traffic temporarily flows to the RP via two methods; via Register messages (until a Register-Stop message is received) and the “native” Shortest-Path Tree (SPT).
- **RP begins receiving traffic down the (S, G) SPT.**
 - As soon as the RP begins receiving (S, G) traffic “natively” (i.e. not encapsulated in Register messages) down the SPT, the RP will set the “T” bit in the (S, G) entry to denote that traffic is successfully flowing down the Shortest-Path Tree (SPT).
 - Now when any (S, G) Register messages are received by the RP, it sees that the “T” bit is set in the (S, G) entry will respond by sending a PIM (S, G) Register-Stop message to the 1st-hop router. This notifies the 1st-hop router that traffic is now being received “natively” down the SPT.
- **1st-hop router receives “Register-Stop” message**
 - When the (S, G) Register-Stop message is received by the 1st-hop router, it clears the “Registering” flag in the (S, G) entry and stops encapsulating (S,G) traffic in Register messages. Traffic is now flowing only down the SPT to the RP.

PIM SM Register Examples

Cisco.com

- **Receivers Join Group First**
- **Source Registers First**
- **Receivers along the SPT**

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

39

- **PIM SM Register Examples**

- Depending on whether there are any existing Receivers for group “G” on the Shared Tree (RPT), the RP hands the Register process a little different.

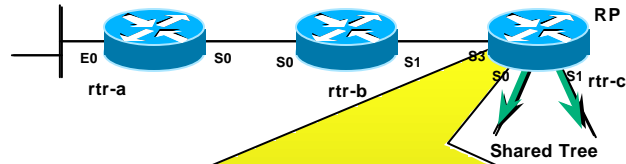
In the following examples we will consider the Register process for the cases when:

- Receivers join group “G” first;
- The Source Registers first.
- Receivers along the SPT.

PIM SM Registering

Receiver Joins Group First

Cisco.com



```
(*, 224.1.1.1), 00:03:14/00:02:59, RP 171.68.28.140, flags:S
Incoming interface: Null, RPF nbr 0.0.0.0,
Outgoing interface list:
  Serial0, Forward/Sparse, 00:03:14/00:02:45
  Serial1, Forward/Sparse, 00:03:14/00:02:45
```

State in "RP" before any source registers
(with receivers on Shared Tree)

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

40

- **State in RP before Registering (Rcvr's on Shared Tree)**

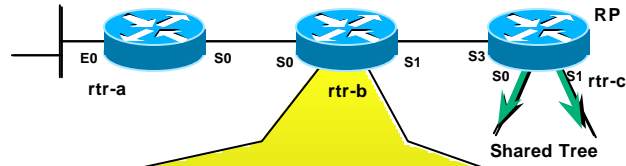
- Pay particular attention to the following in the (*, G) entry:

- The **"Incoming interface:"** is NULL and the **"RPF nbr"** is 0.0.0.0. This indicates that this router is the RP.
- The **"Outgoing interface list:"** contains **Serial0** and **Serial1** which are assumed to be the only two active branches of the Shared Tree (RPT).

PIM SM Registering

Receiver Joins Group First

Cisco.com



```
rtr-b>sh ip mroute 224.1.1.1  
No such group
```

State in "rtr-b" before any source registers
(with receivers on Shared Tree)

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

44

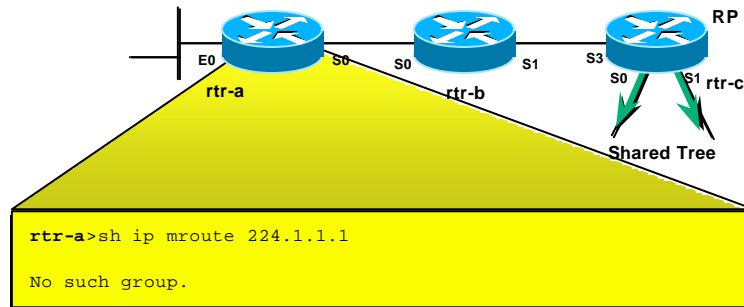
- **State in "rtr-b" before source registers**

- Note that there is no group state information for this Group yet.

PIM SM Registering

Receiver Joins Group First

Cisco.com



State in “rtr-a” before any source registers
(with receivers on Shared Tree)

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

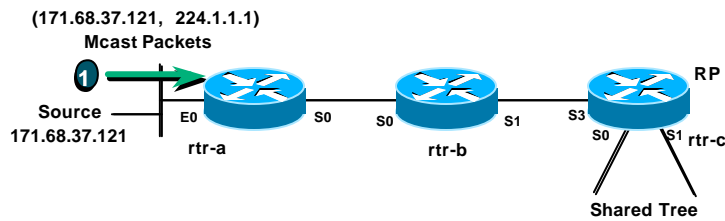
8/10/2001 2:37 PM

42

- **State in 1st-hop router (rtr-a) before source registers**
 - Note that there is no group state information for this Group yet.

PIM SM Registering Receiver Joins Group First

Cisco.com



- 1 "Source" begins sending group G traffic.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

43

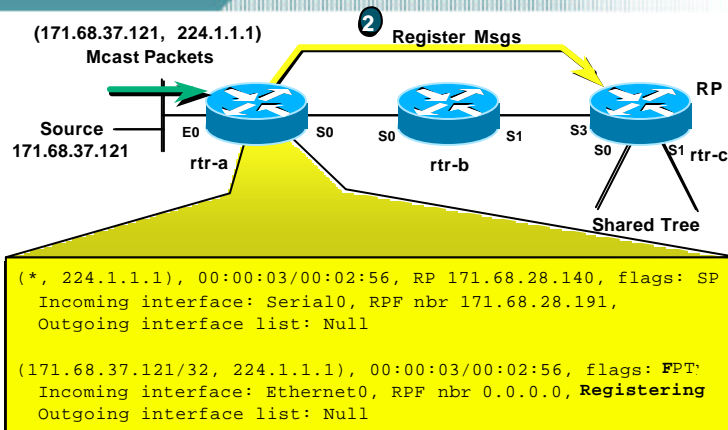
- **Receivers Join Group First Example**

- 1) Source "S" begins sending traffic to group "G".
- 2) 1st-hop router ("rtr-a") creates (*, G) and (S, G) state; encapsulates the multicast packets in PIM Register message(s) and unicasts it(them) to the RP.
- 3) The RP ("rtr-c") de-encapsulates the packets and sees that the packet is for group "G" for which it already has (*, G) state. It then forwards the packets down the Shared Tree.

PIM SM Registering

Receiver Joins Group First

Cisco.com



“rtr-a” creates (S, G) state for source
(After automatically creating a (*, G) entry)

- 1 “Source” begins sending group G traffic.
- 2 “rtr-a” encapsulates packets in Registers; unicasts to RP.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

44

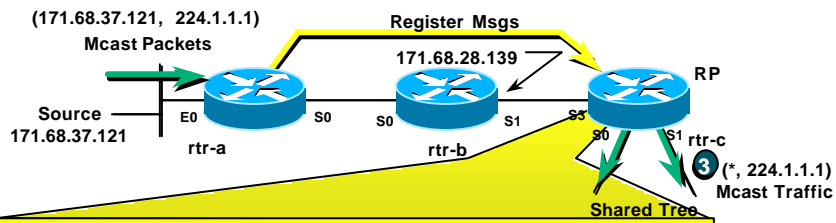
• 1st-hop router (rtr-a) creates (S, G) state

- A (*, G) entry must be created before the (S, G) entry can be created. Note that:
 - The RPF information for this entry points up the Shared Tree via **Serial0** with the RPF neighbor of **171.68.28.191**. (Serial 0 of “rtr-b”.)
 - Because in this example no members have joined the group (the sender is only sending), the OIL of the (*, G) entry is Null.
 - The “**P**” flag (Pruned) is set since the OIL is Null.
- The (S, G) entry is then created. Pay particular attention to the following:
 - The RPF information for this entry points towards the source via **Ethernet0**. The RPF neighbor is **0.0.0.0** because the source is directly connected.
 - The (S, G) OIL receives a copy of the (*, G) OIL. (Which is Null.)
 - The “**F**” flags are set in the (S, G) entry which indicates that this is a directly connected Source.
 - The “**Registering**” flag is set in the (S, G) entry which indicates that we are still sending Register messages to the RP for this Source.
 - The “**P**” flag (Pruned) is set since the OIL is Null.
- 2) The 1st-hop router encapsulates the multicast packets in PIM Register message(s) and unicasts them to the RP.

PIM SM Registering

Receiver Joins Group First

Cisco.com



```
(*, 224.1.1.1), 00:09:21/00:02:38, RP 171.68.28.140, flags: S
Incoming interface: Null, RPF nbr 0.0.0.0,
Outgoing interface list:
  Serial0, Forward/Sparse, 00:09:21/00:02:38
  Serial1, Forward/Sparse, 00:03:14/00:02:46

(171.68.37.121, 224.1.1.1, 00:01:15/00:02:46, flags:
Incoming interface: Serial3, RPF nbr 171.68.28.139,
Outgoing interface list:
  Serial0, Forward/Sparse, 00:00:49/00:02:11
  Serial1, Forward/Sparse, 00:00:49/00:02:11
```

“RP” processes Register; creates (S, G) state

③ “rtr-c” (RP) de-encapsulates packets; forwards down Shared tree.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

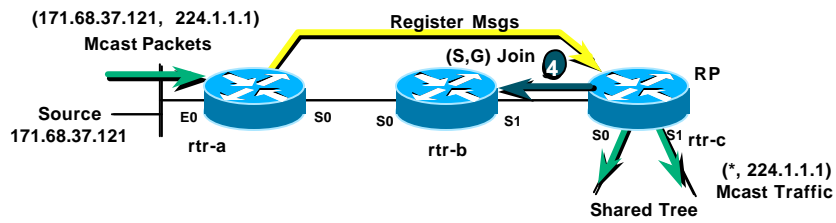
45

• The RP creates (S, G) state

- As a result of the Register message that was received from “rtr-a”, the RP creates (S, G) state as follows:
 - The RPF information is calculated using the source address contained in the multicast packet encapsulated inside of the register message. This results in an IIF of **Serial3** and an RPF neighbor of **171.68.28.139**.
 - Next, the OIL of the parent (*, G) entry is copied into the OIL of the new (S,G) entry. (An additional check is made to insure that the IIF does not appear in the OIL. If it does, it is removed to prevent a route loop.)
- Now the router is ready to forward the (S, G) packet that was encapsulated in the Register message using the newly created (S, G) state. (Note that traffic is always forwarded using the matching (S, G) entry if one exists. Otherwise, the (*, G) entry is used.) This is accomplished as follows:
 - Because this packet was received inside of a Register message, the RPF check is skipped.
 - Next, the router forwards a copy of the packet out all interfaces in the (S, G) OIL. In this case a copy is sent out **Serial0** and **Serial1** which corresponds to the two branches of the Shared Tree.
 - The “**T**” flag is not yet set in the (S, G) entry. However, when the first (S, G) packet is received natively (via the Incoming interface) and forwarded using this entry, the “**T**” flag will be set.

PIM SM Registering Receiver Joins Group First

Cisco.com



4 RP sends (S,G) Join toward Source to build SPT.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

46

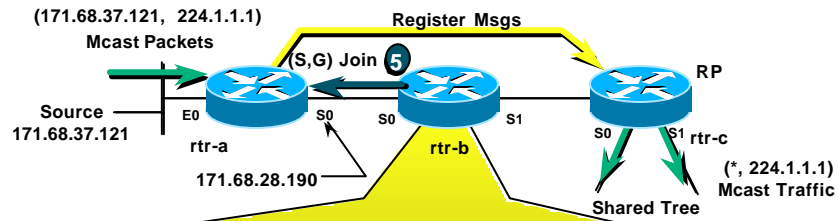
- **Receivers Join Group First Example (cont.)**

- 4) Because RP has existing (*, G) state (i.e. Receivers already waiting on the Shared Tree), it sends an (S, G) Join toward source “S” to build a Shortest-Path Tree (SPT) from source “S” to the RP.

PIM SM Registering

Receiver Joins Group First

Cisco.com



```
(* , 224.1.1.1), 00:04:28/00:01:32, RP 171.68.28.140, flags: SP
Incoming interface: Serial1, RPF nbr 171.68.28.140,
Outgoing interface list: Null

(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags:
Incoming interface: Serial0, RPF nbr 171.68.28.190
Outgoing interface list:
Serial1, Forward/Sparse, 00:04:28/00:01:32
```

“rtr-b” processes Join, creates (S, G) state
(After automatically creating the (*, G) entry)

5 “rtr-b” sends (S,G) Join toward Source to continue building SPT.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

47

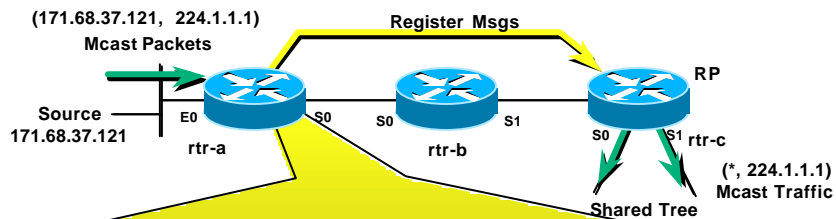
• “rtr-b” processes the (S, G) Join and creates state

- A (*, G) entry must be created before the (S, G) entry can be created. Note that:
 - The RPF information for the (*, G) entry points up the Shared Tree via **Serial1** with the RPF neighbor of **171.68.28.140**. (Serial 3 of the RP.)
 - Because in this example no members have joined the group, the OIL of the (*, G) entry is Null.
 - The “**P**” flag (Pruned) is set since the OIL is Null.
 - The (S, G) entry is then created. Pay particular attention to the following:
 - The RPF information for this entry points towards the source via **Serial0**. The RPF neighbor is **171.68.28.190**. (Serial 0 of “rtr-a”.)
 - The (S, G) OIL initially receives a copy of the (*, G) OIL. (Which is Null.)
 - Interface **Serial1** (which is the interface that received the (S, G) Join) is added to the (S, G) OIL.
 - The “**T**” flag is not yet set in the (S, G) entry. However, when the first (S, G) packet is forwarded using this entry, the flag will be “**T**” set.
- 5) Because the OIL of the (S, G) transitioned from Null to non-Null (when “rtr-b” added Serial1 to the OIL of the newly created entry), a PIM (S, G) Join is sent to rtr-a’s to continue the process of joining the SPT.

PIM SM Registering

Receiver Joins Group First

Cisco.com



```
(*, 224.1.1.1), 00:04:28/00:01:32, RP 171.68.28.140, flags: SP
Incoming interface: Serial0, RPF nbr 171.68.28.191,
Outgoing interface list: Null

(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags: FT
Incoming interface: Ethernet0, RPF nbr 0.0.0.0, Registering
Outgoing interface list:
Serial0, Forward/Sparse, 00:04:28/00:01:32
```

“rtr-a” processes the (S, G) Join; adds Serial 0 to OIL

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

48

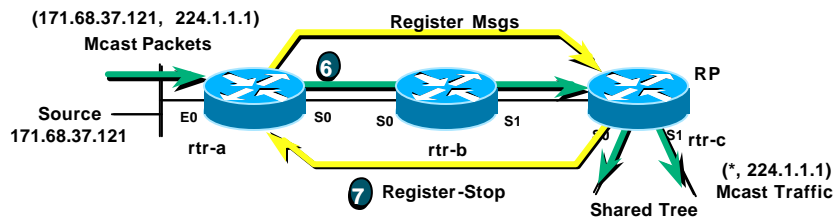
- “rtr-a” processes the (S, G) Join

- Because an (S, G) entry already existed, “rtr-a” simply added the interface on which it received the (S, G) join to the OIL. This results in the following:

- **Serial0** is now listed in the “**Outgoing interface list**” (OIL) since the RP joined the SPT via this interface.
 - The “**P**” flag (Pruned) is cleared since the OIL is no longer Null.

PIM SM Registering Receiver Joins Group First

Cisco.com



- 6 RP begins receiving (S,G) traffic down SPT.
- 7 RP sends "Register-Stop" to "rtr-a".

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

49

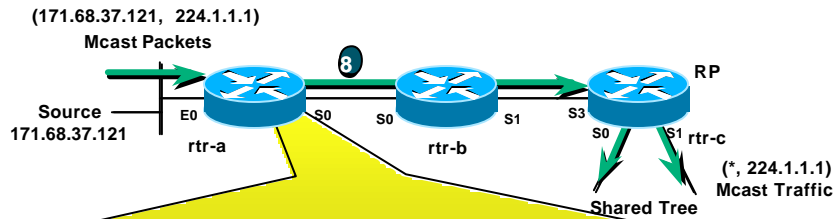
- **A branch of the (S,G) SPT has been built to the RP.**

- 6) Now that the SPT has been built from source "S" to the RP, traffic begins flowing down the Shortest-Path Tree (SPT). At this point, the RP is receiving the (S, G) traffic "natively" down the SPT. (This causes the "T" flags to be set in the (S, G) entries along this path including in the RP.)
- 7) The RP then sends an (S, G) "Register-Stop" to the 1st-hop router to inform it that the encapsulated group "G" Register messages from source "S" are no longer necessary.

PIM SM Registering

Receiver Joins Group First

Cisco.com



```
(*, 224.1.1.1), 00:04:28/00:01:32, RP 171.68.28.140, flags: SP
Incoming interface: Serial0, RPF nbr 171.68.28.191,
Outgoing interface list: Null

(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags: FT
Incoming interface: Ethernet0, RPF nbr 0.0.0.0,
Outgoing interface list:
Serial0, Forward/Sparse, 00:04:28/00:01:32
```

**“rtr-a” stops sending Register messages
(Final State in “rtr-a”)**

8) (S,G) Traffic now flowing down a single path (SPT) to RP.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

50

- **“rtr-a” stops sending Register messages**

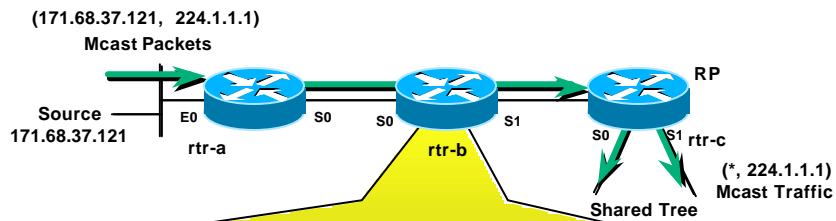
- When the 1st-hop router (rtr-a) receives the (S, G) Register-Stop message it ceases sending encapsulated Register messages for (S, G) traffic.
 - Notice that the “Registering” flag on the second line of the (S, G) entry is no longer being displayed indicating that “rtr-a” is not sending Registers.
 - This is the final state in “rtr-a” after the Registration process.

8) (S, G) traffic is now only flowing down the Shortest-Path Tree (SPT).

PIM SM Registering

Receiver Joins Group First

Cisco.com



```
(*, 224.1.1.1), 00:04:28/00:01:32, RP 171.68.28.140, flags: SP
Incoming interface: Serial1, RPF nbr 171.68.28.140,
Outgoing interface list: Null

(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags: T
Incoming interface: Serial0, RPF nbr 171.68.28.190
Outgoing interface list:
Serial1, Forward/Sparse, 00:04:28/00:01:32
```

Final state in "rtr-b"

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

54

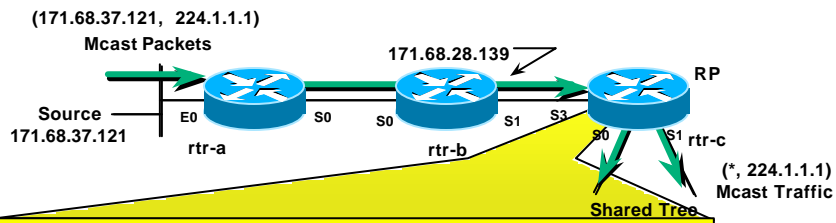
- **Final state in "rtr-b" after the Registration process**

- Pay particular attention to the following in the (S, G) entry:
 - The **"T"** flag is now set indicating that (S, G) traffic is flowing along this path.
- The (*, G) entry still has a Null OIL and the **"P"** flag is still set.
 - This is because there are no members that have joined the Shared Tree.

PIM SM Registering

Receiver Joins Group First

Cisco.com



```
(*, 224.1.1.1), 00:09:21/00:02:38, RP 171.68.28.140, flags: S
Incoming interface: Null, RPF nbr 0.0.0.0,
Outgoing interface list:
  Serial0, Forward/Sparse, 00:09:21/00:02:38
  Serial1, Forward/Sparse, 00:03:14/00:02:46

(171.68.37.121, 224.1.1.1, 00:01:15/00:02:46, flags: T
Incoming interface: Serial3, RPF nbr 171.68.28.139,
Outgoing interface list:
  Serial0, Forward/Sparse, 00:00:49/00:02:11
  Serial1, Forward/Sparse, 00:00:49/00:02:11
```

Final state in the "RP"
(with receivers on Shared Tree)

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

52

- **Final state in the RP after the Registration process**

- Pay particular attention to the following in the newly created (S, G) entry:
 - The **"T"** flag is now set indicating that (S, G) traffic is flowing along this path.

PIM SM Register Examples

Cisco.com

- Receivers Join Group First
- **Source Registers First**
- Receivers along the SPT

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

53

- **PIM SM Register Examples**

- Depending on whether there are any existing Receivers for group “G” on the Shared Tree (RPT), the RP hands the Register process a little different.

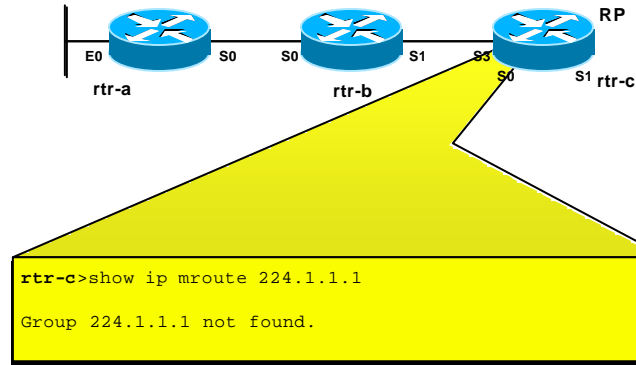
In the following examples we will consider the Register process for the cases when:

- Receivers join group “G” first;
- The Source Registers first.

PIM SM Registering

Source Registers First

Cisco.com



State in "RP" before Registering
(without receivers on Shared Tree)

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

54

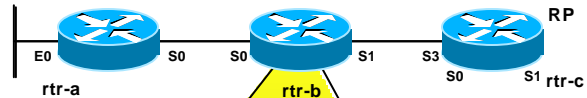
- **State in RP before Registering (w/o Rcvr's on Shared Tree)**

- Notice that no state for group "G" exists since there are no Receivers on the Shared Tree yet.

PIM SM Registering

Source Registers First

Cisco.com



```
rtr-b>show ip mroute 224.1.1.1  
Group 224.1.1.1 not found.
```

**State in "rtr-b" before any source registers
(with receivers on Shared Tree)**

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

55

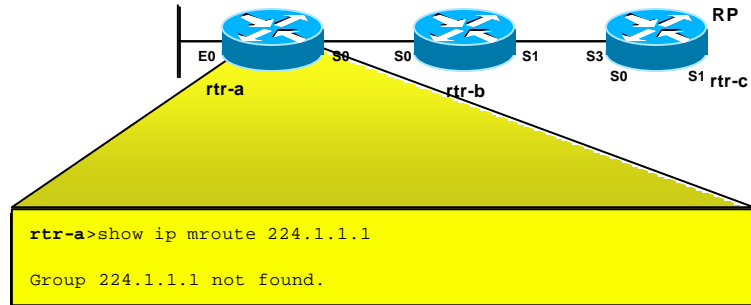
- **State in "rtr-b" before source registers**

- Note that there is no group state information for this Group yet.

PIM SM Registering

Source Registers First

Cisco.com



State in “rtr-a” before any source registers
(with receivers on Shared Tree)

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

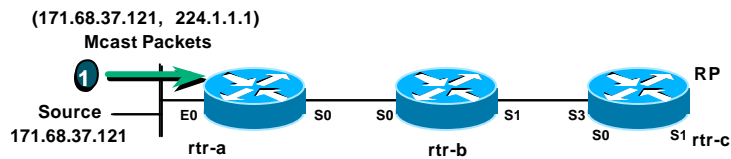
8/10/2001 2:37 PM

56

- **State in 1st-hop router (rtr-a) before source registers**
 - Note that there is no group state information for this Group yet.

PIM SM Registering Source Registers First

Cisco.com



1 "Source" begins sending group G traffic.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

57

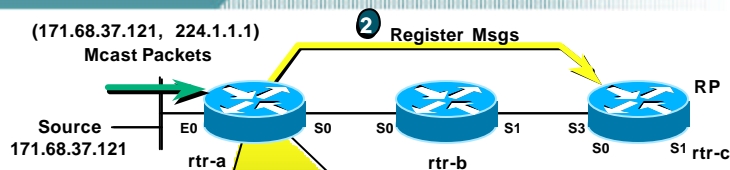
- **Source Registers First Example**

- 1) Source "S" begins sending traffic to group "G".
- 2) 1st-hop router ("rtr-a") creates (*, G) and (S, G) state; encapsulates the multicast packets in PIM Register message(s) and unicasts it(them) to the RP.
- 3) The RP ("rtr-c") de-encapsulates the (S, G) packet and creates (*, G) and (S, G) state. Since no one has joined the Shared Tree yet, the OIL's of these entries will be NULL.. Because the OIL of the (S, G) entry (just created) is NULL, the packet is discarded.

PIM SM Registering

Source Registers First

Cisco.com



```
(*, 224.1.1.1), 00:00:03/00:02:56, RP 171.68.28.140, flags: SP
Incoming interface: Serial0, RPF nbr 171.68.28.191,
Outgoing interface list: Null

(171.68.37.121/32, 224.1.1.1), 00:00:03/00:02:56, flags: FPT
Incoming interface: Ethernet0, RPF nbr 0.0.0.0, Registering
Outgoing interface list: Null
```

“rtr-a” creates (S, G) state for source
(After automatically creating a (*, G) entry)

- 1 “Source” begins sending group G traffic.
- 2 “rtr-a” encapsulates packets in Registers; unicasts to RP.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

58

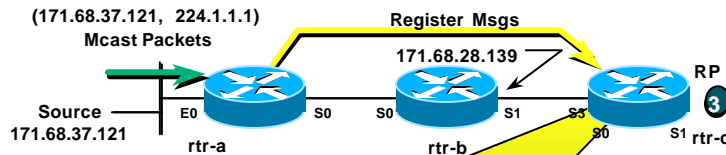
• 1st-hop router (rtr-a) creates (S, G) state

- A (*, G) entry must be created before the (S, G) entry can be created. Note that:
 - The RPF information for this entry points up the Shared Tree via **Serial0** with the RPF neighbor of **171.68.28.191**. (Serial 0 of “rtr-b”.)
 - Because in this example no members have joined the group (the sender is only sending), the OIL of the (*, G) entry is Null.
 - The “**P**” flag (Pruned) is set since the OIL is Null.
 - The (S, G) entry is then created. Pay particular attention to the following:
 - The RPF information for this entry points towards the source via **Ethernet0**. The RPF neighbor is **0.0.0.0** because the source is directly connected.
 - The (S, G) OIL receives a copy of the (*, G) OIL. (Which is Null.)
 - The “**F**” flags are set in the (S, G) entry which indicates that this is a directly connected Source.
 - The “**Registering**” flag is set in the (S, G) entry which indicates that we are still sending Register messages to the RP for this Source.
 - The “**P**” flag (Pruned) is set since the OIL is Null.
- 2) The 1st-hop router encapsulates the multicast packets in PIM Register message(s) and unicasts them to the RP.

PIM SM Registering

Source Registers First

Cisco.com



```
(*, 224.1.1.1), 00:01:15/00:01:45, RP 171.68.28.140, flags: SP
Incoming interface: Null, RPF nbr 0.0.0.0,
Outgoing interface list: Null

(171.68.37.121, 224.1.1.1), 00:01:15/00:01:45, flags: P
Incoming interface: Serial3, RPF nbr 171.68.28.139,
Outgoing interface list: Null
```

“RP” processes Register; creates (S, G) state
(After automatically creating the (*, G) entry)

③ **“rtr-c” (RP) has no receivers on Shared Tree; discards packet.**

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

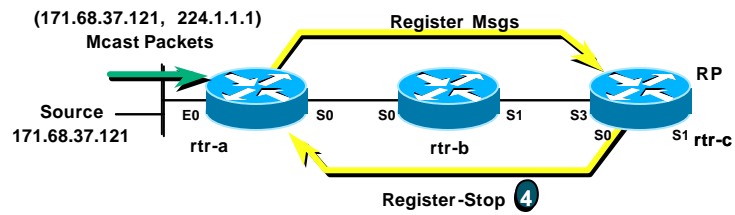
59

• The RP creates (S, G) state

- As a result of the Register message that was received from “rtr-a”, the RP creates (*, G) and (S, G) state. However, because no previous (*, G) state existed, it must be created before the (S,G) entry can be created.
 - This (*, G) entry is created as shown above. Notice that the (*, G) OIL is NULL. This is because the RP has not yet received any (*, G) Joins for this group. (Remember, in this example, the source registers first.)
- Next, the (S, G) entry can be created and is accomplished as follows:
 - The RPF information is calculated using the source address contained in the multicast packet encapsulated inside of the register message. This results in an IIF of **Serial3** and an RPF neighbor of **171.68.28.139**.
 - Next, the OIL of the parent (*, G) entry is copied into the OIL of the new (S,G) entry. Since the OIL of the (*, G) entry is NULL, this results in a NULL (S, G) OIL.
- Now the router is ready to forward the (S, G) packet that was encapsulated in the Register message using the newly created (S, G) state. This is accomplished as follows:
 - Because this packet was received inside of a Register message, the RPF check is skipped.
 - Next, the router forwards a copy of the packet out all interfaces in the matching (S, G) OIL. However, because the (S, G) OIL is NULL (i.e. there are no branches of the Shared Tree), the packet is simply discarded.

PIM SM Registering Source Registers First

Cisco.com



4 RP sends "Register-Stop" to "rtr-a".

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

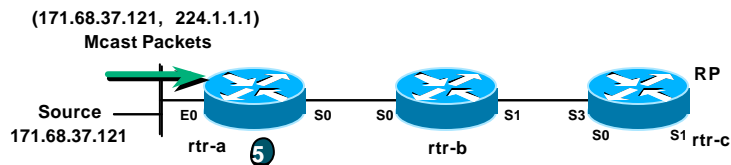
60

- **Source Registers First Example**

- 4) Since the RP has no (*, G) state and hence no receivers on the Shared Tree, it does not need the (S, G) traffic. Therefore the RP sends an (S, G) "Register-Stop" message to the 1st-hop router so it will stop sending Register messages.

PIM SM Registering Source Registers First

Cisco.com



- 5 “rtr-a” stops encapsulating traffic in Register Messages; drops packets from Source.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

61

- **Source Registers First Example**

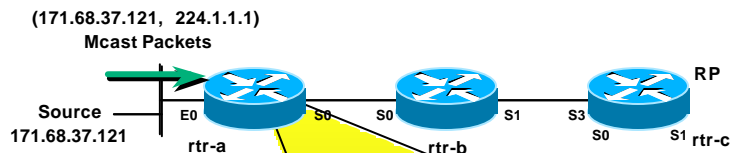
- 5) The 1st-hop router receives the (S, G) Register-Stop message and stops sending Register messages for (S, G) traffic.

Note: Eventually, the original (S, G) entry will time out (approx. 3 min.) and be deleted. The Register process will start over again when the 1st-hop router receives the next multicast packet from directly connected source “S”. The RP will again respond with a Register-Stop which will prevent the (S,G) traffic from flowing to the RP until it is needed.

PIM SM Registering

Source Registers First

Cisco.com



```
(*, 224.1.1.1), 00:01:28/00:01:32, RP 171.68.28.140, flags: SP
Incoming interface: Serial0, RPF nbr 171.68.28.191,
Outgoing interface list: Null

(171.68.37.121/32, 224.1.1.1), 00:01:28/00:01:32, flags: FPT
Incoming interface: Ethernet0, RPF nbr 0.0.0.0
Outgoing interface list: Null
```

State in "rtr-a" after Registering (without receivers on Shared Tree)

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

62

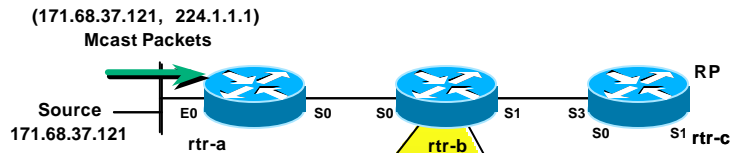
- **State in 1st-hop router after Registering (w/o Rcvr's on Shared Tree)**

- Pay particular attention to the following in the (S, G) entry:
 - The **"Registering"** flag is now cleared.
 - The **"Outgoing interface list"** is still Null since the RP did not join the SPT.
 - The **"P"** flag (Pruned) is still set since the olist is still Null.
 - The **"00:01:32"** Expiration time value will count down to zero at which time the (S, G) entry will be deleted. (The Register process will begin all over again when the next multicast packet is received from source "S".)

PIM SM Registering

Source Registers First

Cisco.com



```
rtr-b>show ip mroute 224.1.1.1  
Group 224.1.1.1 not found.
```

**State in "rtr-b" after "rtr-a" Registers
(without receivers on Shared Tree)**

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

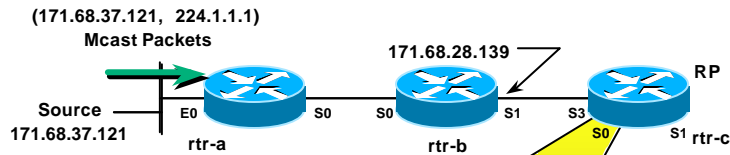
63

- **State in "rtr-b" after Registering (w/o Rcvr's on Shared Tree)**
 - Notice that no state exists in "rtr-b" at this point in time.

PIM SM Registering

Source Registers First

Cisco.com



```
(*, 224.1.1.1), 00:01:15/00:01:45, RP 171.68.28.140, flags: SP
Incoming interface: Null, RPF nbr 0.0.0.0,
Outgoing interface list: Null

(171.68.37.121, 224.1.1.1), 00:01:15/00:01:45, flags: P
Incoming interface: Serial3, RPF nbr 171.68.28.139,
Outgoing interface list: Null
```

State in "RP" after "rtr-a" Registers
(without receivers on Shared Tree)

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

64

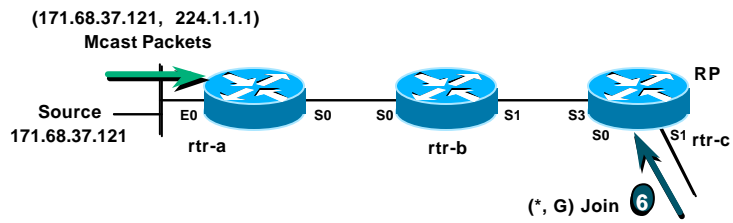
- **State in RP after Registering (w/o Rcvr's on Shared Tree)**

- Pay particular attention to the following in the newly created (S, G) entry:
 - The **"RPF nbr"** is the IP Address of "rtr-b".
 - The **"Incoming interface:"** is **Serial3** which is the RPF interface towards source "S" via "rtr-b".
 - The **"Outgoing interface list:"** is Null since the (*,G) OIL is also Null. (Indicates there are no Receivers on the Shared Tree yet.)
 - The **"P"** flag (Pruned) is set since the OIL is Null.
- The (S,G) state will remain in the RP as long as the source is still actively sending. This is accomplished by fact that the first-hop route will continue sending periodic Register messages to the RP as long as the first-hop router is receiving traffic from the source.

PIM SM Registering

Source Registers First

Cisco.com



Receivers begin joining the Shared Tree

6 RP ("rtr-c") receives (*, G) Join from a receiver on Shared Tree.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

65

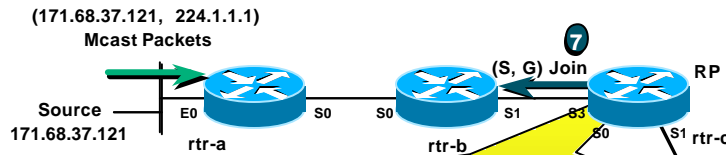
- **Source Registers First Example**

6) The RP now begins receiving (*, G) Joins from Last-hop routers with Receivers that wish to join the Shared Tree.

PIM SM Registering

Source Registers First

Cisco.com



```
(*, 224.1.1.1), 00:09:21/00:02:38, RP 171.68.28.140, flags: S
Incoming interface: Null, RPF nbr 0.0.0.0,
Outgoing interface list:
  Serial1, Forward/Sparse, 00:00:14/00:02:46

(171.68.37.121/32, 224.1.1.1, 00:01:15/00:02:46, flags: T
Incoming interface: Serial13, RPF nbr 171.68.28.139,
Outgoing interface list:
  Serial1, Forward/Sparse, 00:00:14/00:02:46
```

“RP” processes (*,G) Join
(Adds Serial1 to Outgoing Interface Lists)

7 RP sends (S,G) Joins for all known Sources in Group.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

66

• The RP process the (*, G) Join

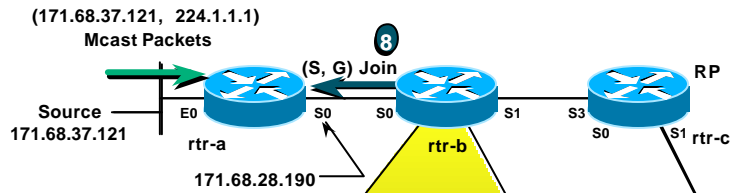
- In the (*, G) entry:
 - **Serial1** has been added to the (*, G) entry since a (*,G) Join was received on this interface which is the only active branch of the Shared Tree (RPT).
 - In the (S, G) entry:
 - **Serial1** has also been added to the (S, G) OIL because the OIL's of all (S,G) entries are always kept in sync with their parent (*, G).

Note: When the (S, G) OIL's are synchronized with the OIL of their parent (*,G) OIL, a check is made to insure that the IIF of the (S, G) does not appear in the OIL of the (S, G). This could result in a route loop.
- 7) The transitioning of the (*, G) OIL from Null to non-Null triggers the RP to scan its list of (S, G) entries for group “G” and send (S, G) Joins towards all sources. (This will cause a SPT to be built from each active source back to the RP which will eventually start the flow of (S, G) traffic to the RP.)

PIM SM Registering

Source Registers First

Cisco.com



```
(*, 224.1.1.1), 00:04:28/00:01:32, RP 171.68.28.140, flags: SP
Incoming interface: Serial1, RPF nbr 171.68.28.140,
Outgoing interface list: Null

(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags:
Incoming interface: Serial0, RPF nbr 171.68.28.190
Outgoing interface list:
Serial1, Forward/Sparse, 00:04:28/00:01:32
```

“rtr-b” processes Join, creates (S, G) state
(After automatically creating the (*, G) entry)

8 “rtr-b” sends (S,G) Join toward Source to continue building SPT.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

67

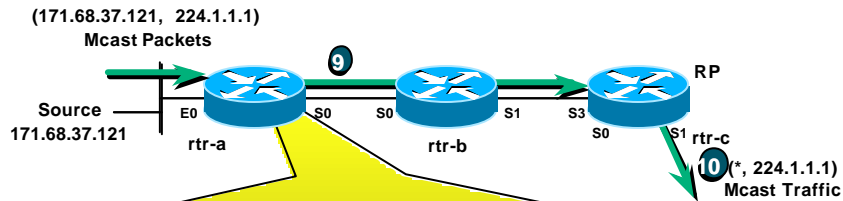
• “rtr-b” processes the (S, G) Join and creates state

- A (*, G) entry must be created before the (S, G) entry can be created. Note that:
 - The RPF information for the (*, G) entry points up the Shared Tree via **Serial1** with the RPF neighbor of **171.68.28.140**. (Serial 3 of the RP.)
 - Because in this example no members have joined the group, the OIL of the (*, G) entry is Null.
 - The **“P”** flag (Pruned) is set since the OIL is Null.
 - The (S, G) entry is then created. Pay particular attention to the following:
 - The RPF information for this entry points towards the source via **Serial0**. The RPF neighbor is **171.68.28.190**. (Serial 0 of “rtr-a”.)
 - The (S, G) OIL initially receives a copy of the (*, G) OIL. (Which is Null.)
 - Interface **Serial1** (which is the interface that received the (S, G) Join) is added to the (S, G) OIL.
 - The **“T”** flag is not yet set in the (S, G) entry. However, when the first (S, G) packet is forwarded using this entry, the flag will be **“T”** set.
- 8) Because the OIL of the (S, G) transitioned from Null to non-Null (when “rtr-b” added Serial1 to the OIL of the newly created entry), a PIM (S, G) Join is sent to rtr-a’s to continue the process of joining the SPT.

PIM SM Registering

Source Registers First

Cisco.com



```
(*, 224.1.1.1), 00:04:28/00:01:32, RP 171.68.28.140, flags: SP
Incoming interface: Serial0, RPF nbr 171.68.28.191,
Outgoing interface list: Null

(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags: FT
Incoming interface: Ethernet0, RPF nbr 0.0.0.0,
Outgoing interface list:
  Serial0, Forward/Sparse, 00:04:28/00:01:32
```

“rtr-a” processes the (S, G) Join; adds Serial0 to OIL

9 RP begins receiving (S,G) traffic down SPT.

10 RP forwards (S,G) traffic down Shared Tree to receivers.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

68

• 1st-hop router (rtr-a) processes the (S, G) Join

– The (S, G) Join is processed as follows:

- **Serial0** is added to the “**Outgoing interface list**” (OIL). (This is the interface on which the (S, G) Join arrived.)
- The “**P**” flag (Pruned) is cleared since the OIL is no longer Null.

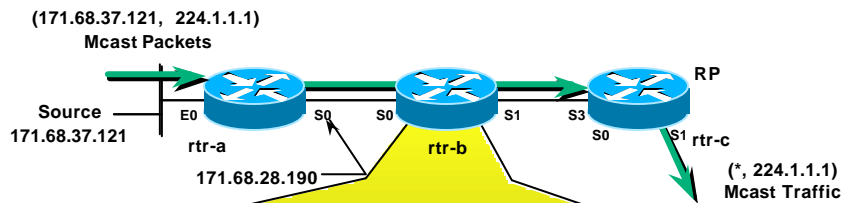
9) As a result of **Serial0** being added to the (S, G) OIL, traffic begins to flow down the SPT from the source to the RP.

10) The RP then forwards all incoming (S, G) traffic to the Receivers down the Shared Tree.

PIM SM Registering

Source Registers First

Cisco.com



```
(*, 224.1.1.1), 00:04:28/00:01:32, RP 171.68.28.140, flags: SP
Incoming interface: Serial11, RPF nbr 171.68.28.140,
Outgoing interface list: Null

(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags: T
Incoming interface: Serial10, RPF nbr 171.68.28.190
Outgoing interface list:
Serial11, Forward/Sparse, 00:04:28/00:01:32
```

Final state in "rtr-b" after Receivers Join

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

69

- State in "rtr-b" after Receivers Join

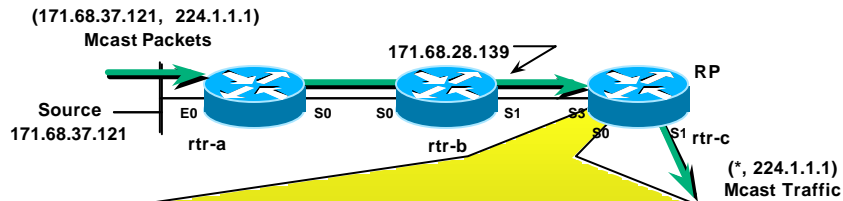
- Pay particular attention to the following:

- Both (*, G) and (S, G) state was created as a result of the (S, G) Join received from the RP.
 - The **"P"** flag set in the (*, G) entry since there are no receivers on the Shared Tree at this point in the network.
 - The **"T"** flag is set in the (S, G) entry indicating that traffic is flowing down the Shortest-Path Tree.
 - The **"RPF nbr"** is the IP Address of "rtr-a".
 - **Serial0** is the **"Incoming interface"** of the (S, G) entry since this is the RPF interface for source "S" via "rtr-a".
 - **Serial1** is listed in the **"Outgoing interface list"** of the (S, G) entry since the RP joined the SPT via this interface.

PIM SM Registering

Source Registers First

Cisco.com



```
(*, 224.1.1.1), 00:09:21/00:02:38, RP 171.68.28.140, flags: S
Incoming interface: Null, RPF nbr 0.0.0.0,
Outgoing interface list:
  Serial1, Forward/Sparse, 00:03:14/00:02:46

(171.68.37.121/32, 224.1.1.1, 00:01:15/00:02:46, flags: T
Incoming interface: Serial3, RPF nbr 171.68.28.139,
Outgoing interface list:
  Serial1, Forward/Sparse, 00:00:49/00:02:11
```

Final state in "RP" after Receivers Join

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

70

• State in RP after Receivers Join

- In the (*, G) entry:
 - **Serial1** has been added to the (*, G) entry since a (*,G) Join was received on this interface which is the only active branch of the Shared Tree (RPT).
- In the (S, G) entry:
 - **Serial1** has also been added to the (S, G) OIL because the OIL's of all (S,G) entries are always kept in sync with their parent (*, G).

Note: When the (S, G) OIL's are synchronized with the OIL of their parent (*, G) OIL, a check is made to insure that the IIF of the (S, G) does not appear in the OIL of the (S, G). This could result in a route loop.

PIM SM Register Examples

Cisco.com

- Receivers Join Group First
- Source Registers First
- Receivers along the SPT

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

74

- **PIM SM Register Examples**

- Depending on whether there are any existing Receivers for group “G” on the Shared Tree (RPT), the RP hands the Register process a little different.

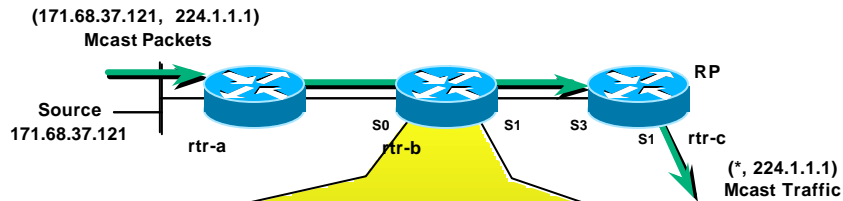
In the following examples we will consider the Register process for the cases when:

- Receivers join group “G” first;
- The Source Registers first.
- Receivers along the SPT.

PIM SM Registering

Receivers along the SPT

Cisco.com



```
(*, 224.1.1.1), 00:04:28/00:01:32, RP 171.68.28.140, flags: SP
Incoming interface: Serial11, RPF nbr 171.68.28.140,
Outgoing interface list: Null

(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags: T
Incoming interface: Serial10, RPF nbr 171.68.28.190
Outgoing interface list:
Serial11, Forward/Sparse, 00:04:28/00:01:32
```

Current state in "rtr-b"

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

72

- **State in "rtr-b" with traffic flowing on the SPT**

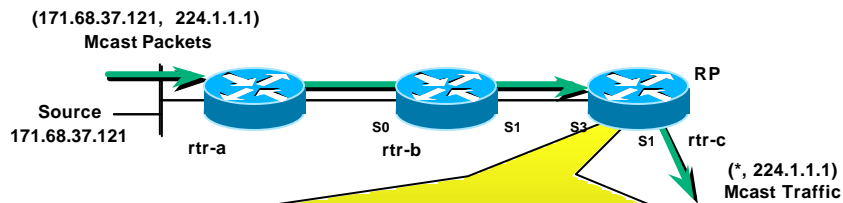
- Pay particular attention to the following:

- Both (*, G) and (S, G) state was created as a result of the (S, G) Join received from the RP.
 - The **"P"** flag set in the (*, G) entry since there are no receivers on the Shared Tree at this point in the network.
 - The **"T"** flag is set in the (S, G) entry indicating that traffic is flowing down the Shortest-Path Tree.
 - The **"RPF nbr"** is the IP Address of "rtr-a".
 - **Serial0** is the **"Incoming interface"** of the (S, G) entry since this is the RPF interface for source "S" via "rtr-a".
 - **Serial1** is listed in the **"Outgoing interface list"** of the (S, G) entry since the RP joined the SPT via this interface.

PIM SM Registering

Receivers along the SPT

Cisco.com



```
(*, 224.1.1.1), 00:09:21/00:02:38, RP 171.68.28.140, flags: S
Incoming interface: Null, RPF nbr 0.0.0.0,
Outgoing interface list:
  Serial1, Forward/Sparse, 00:03:14/00:02:46

(171.68.37.121/32, 224.1.1.1, 00:01:15/00:02:46, flags: T
Incoming interface: Serial3, RPF nbr 171.68.28.139,
Outgoing interface list:
  Serial1, Forward/Sparse, 00:00:49/00:02:11
```

Current state in the RP

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

73

- **State in the RP with traffic flowing on the SPT**

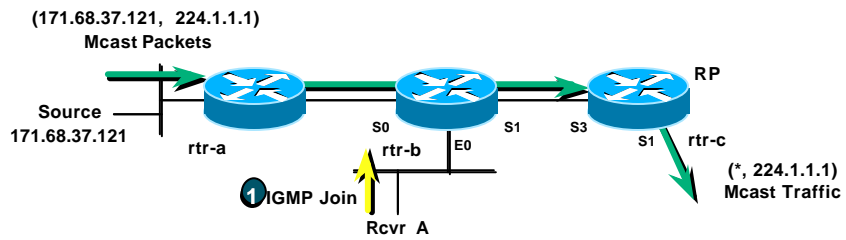
- Pay particular attention to the following:

- The (*, G) entry only has **Serial1** in its outgoing interface list.
 - In the (S, G) entry, **Serial0** is the **“Incoming interface”** since this is the RPF interface for source “S” via “rtr-b”.
 - **Serial1** is listed in the **“Outgoing interface list”** of the (S, G) entry because the OIL of the (S, G) entry is always kept in sync with the (*, G) OIL.

PIM SM Registering

Receivers along the SPT

Cisco.com



- 1 "Rcvr A" wishes to receive group G traffic. Sends IGMP Join for G.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

74

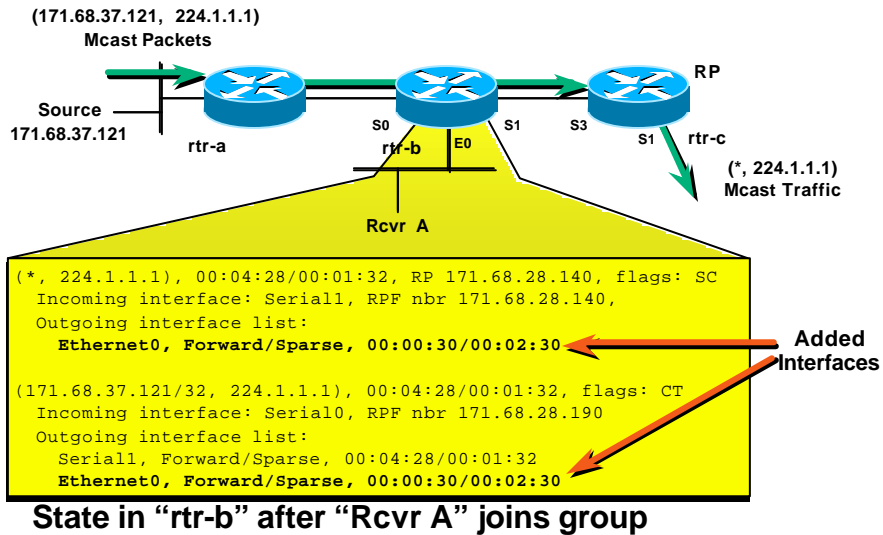
- **Receivers along the SPT**

- **Step 1:** A host directly connected to "rtr-b", Receiver "A", joins multicast group 224.1.1.1 by sending an IGMP Report.

PIM SM Registering

Receivers along the SPT

Cisco.com



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

75

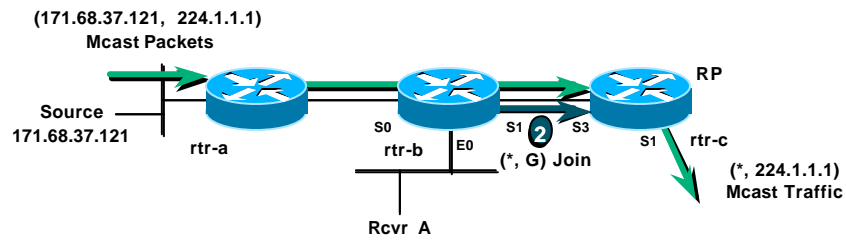
• Receivers along the SPT

– As a result of the IGMP Report sent by Receiver A for group 224.1.1.1, "rtr-b" updates its state for this group as follows.

- **Ethernet0** is added to the OIL of the (*, G) entry. This is done to permit any (*, 224.1.1.1) traffic flowing down the Shared Tree to be forwarded to Receiver "A".
- Next, the OIL's of all child (S, G) entries are synchronized with the OIL change just made to the OIL of the (*, G). This results in **Ethernet0** being added to the OIL of the (171.68.37.121/32, 224.1.1.1) entry. This permits traffic from this source to be "picked off" as it flows along the SPT through "rtr-b" on its way to the RP. (Note that this traffic does not flow to the RP and then back out the same interface to reach "rtr-b". This is a common misperception.)

PIM SM Registering Receivers along the SPT

Cisco.com



- 2 "rtr-b" triggers a (*,G) Join to join the Shared Tree

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

76

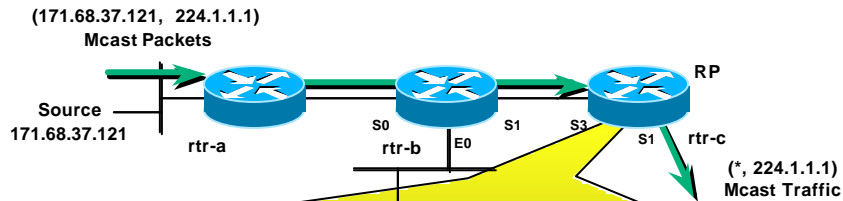
- Receivers along the SPT

- **Step 2:** Because the OIL of the (*, G) entry in "rtr-b" transitioned from NULL to non-Null (**Ethernet0** is now in the (*, G) OIL), a (*, G) Join message is triggered. This message is sent up the Shared Tree so that "rtr-b" will be placed on a branch of the Shared Tree.

PIM SM Registering

Receivers along the SPT

Cisco.com



```

(*, 224.1.1.1), 00:09:21/00:02:38, RP 171.68.28.140, flags: S
Incoming interface: Null, RPF nbr 0.0.0.0,
Outgoing interface list:
  Serial1, Forward/Sparse, 00:03:14/00:02:46
  Serial3, Forward/Sparse, 00:00:10/00:02:50

(171.68.37.121/32, 224.1.1.1, 00:01:15/00:02:46, flags: T
Incoming interface: Serial3, RPF nbr 171.68.28.139,
Outgoing interface list:
  Serial1, Forward/Sparse, 00:00:49/00:02:11
  
```

State in "RP" after "rtr-b" joins Shared Tree

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

77

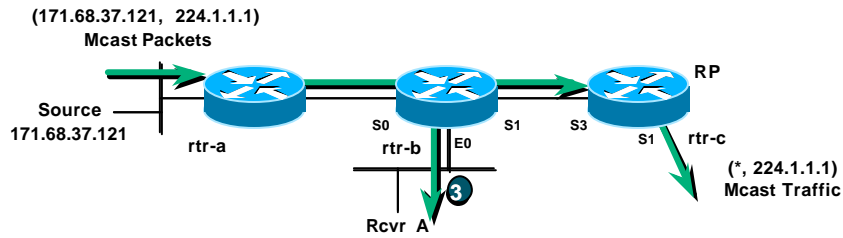
- **Receivers along the SPT**

- When the RP receives the (*, G) Join sent by "rtr-b", it adds **Serial3** to the (*, G) OIL.
- Next, the RP synchronizes the OIL's of all (S, G) entries by adding **Serial3** to each (S, G) OIL. However in this case, **Serial3** is the Incoming interface for the (171.68.37.121/32, 224.1.1.1) entry and is therefore not added to the OIL. (If it were, Serial3 would appear in both the incoming and outgoing interface list which could cause a route loop.)

PIM SM Registering

Receivers along the SPT

Cisco.com



3 Group G traffic begins to flow to "Rcvr A".

(Note: 171.68.37.121 traffic doesn't flow to RP then back down to rtr-b)

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

78

• Receivers along the SPT

- **Step 3:** Traffic from source 171.68.37.121 is now being "picked off" by "rtr-b" and forwarded out **Ethernet0** as the traffic flows down the SPT to the RP.
- Again, it is important to note that this source traffic does not flow to the RP and then turn around and come back out on the same interface that it arrived. (Refer to the state in the RP shown on the previous page.)

PIM-SM Protocol Mechanics

Cisco.com

- PIM Neighbor Discovery
- PIM State
- PIM SM Joining
- PIM SM Registering
- **PIM SM SPT-Switchover**
- PIM SM Pruning
- PIM SM State Maintenance

PIM SM SPT-Switchover

Cisco.com

- **SPT Thresholds may be set for any Group**
 - Access Lists may be used to specify which Groups
 - Default Threshold = 0Kbps (I.e. immediately join SPT)
 - Threshold = “infinity” means “never join SPT”.
- **Threshold triggers Join of Source Tree**
 - Sends an (S,G) Join up SPT for next “S” in “G” packet received.
- **Pros**
 - Reduces Network Latency
- **Cons**
 - More (S,G) state must be stored in the routers.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

80

• SPT Thresholds

- In PIM Sparse mode, SPT Thresholds may be configured to control when to switch to the Shortest-Path Tree (SPT).
- SPT Thresholds are specified in Kbps and can be used with Access List to specify to which Group(s) the Threshold applies.
- The default SPT-Threshold is 0Kbps. This means that any and all sources are immediately switched to the Shortest-Path Tree.
- If an SPT-Threshold of “Infinity” is specified for a group, the sources will not be switched to the Shortest-Path Tree (SPT) and will remain on the Shared Tree.

• Exceeding the Threshold

- When the Group’s SPT-Threshold is exceeded in a last-hop router, the next received packet for the group will cause an (S, G) Join to be sent toward the source of the packet. This builds a Shortest-Path Tree from the source “S” to the last-hop router.

• PROS

- By switching to the Shortest-Path Tree (SPT), the most optimal (usually) path is used to deliver the multicast traffic. Depending on the location of the source in relation to the RP, this switch to the SPT can reduce network latency substantially.

• CONS

- In networks with large numbers of senders (remember most multicast applications such as IP/TV Client, send RTCP multicast packets in the background and are therefore senders), an increased amount of state must be kept in the routers. In some cases, an Infinity threshold may be used to force certain groups to remain on the Shared Tree when latency is not an issue.

PIM SM SPT-Switchover

Cisco.com

SPT-Switchover Mechanism

Once each second

- Compute new $(*, G)$ traffic rate
- If threshold exceeded, set “J” flag in $(*, G)$

For each (S_i, G) packet received:

- If “J” flag set in $(*, G)$
 - Join SPT for (S_i, G)
 - Mark (S_i, G) entry with “J” flag
 - Clear “J” flag in $(*, G)$

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

8/1

• SPT-Threshold Myth

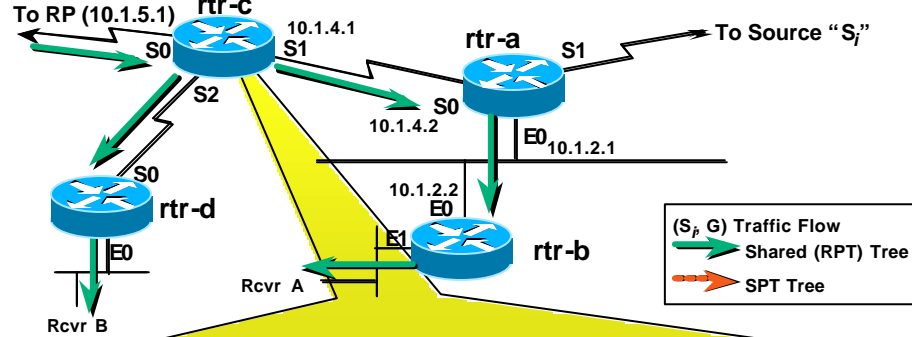
- This is a frequently misunderstood mechanism. Many people think that the traffic rates of the sources in the group are monitored and compared against the SPT-Threshold. THIS IS NOT THE CASE. Instead, the total aggregate rate of Group traffic *flowing down the Shared Tree (RPT)* is calculated once per second. If this total aggregate rate is exceeded, then the next Group packet received causes that source to be switched to the Shortest-Path Tree (SPT).

• SPT-Switchover Mechanism

- Once each second, the aggregate $(*, G)$ traffic rate is computed and checked against the SPT-Threshold. If the aggregate rate of all group traffic flowing down the Shared Tree (RPT) exceeds the threshold, then the “J” flag is set in the $(*, G)$ entry.
- As each multicast packet is received on the Shared Tree, the “J” bit is checked in the $(*, G)$ entry.
 - If the “J” flag is set, a new (S, G) entry is created for the source of the packet.
 - An (S, G) Join is sent towards the source in order to join the SPT.
 - The “J” flag is set in the (S, G) entry to denote that this entry was created as a result of the SPT-Threshold switchover.
 - The “J” flag in the $(*, G)$ is reset. (It will be set in one second if the aggregate rate on the Shared Tree is still over the SPT-Threshold.)
- This mechanism can sometimes result in low rate sources being switched to the SPT erroneously. However, the RPT-switchback mechanism will correct this situation and eventually only the high rate sources will be received via SPTs while low rate sources will remain on the Shared Tree.

PIM SM SPT-Switchover

Cisco.com



```
(*, 224.1.1.1), 00:01:43/00:02:13, RP 10.1.5.1, flags: S
Incoming interface: Serial0, RPF nbr 10.1.5.1,
Outgoing interface list:
Serial1, Forward/Sparse, 00:01:43/00:02:11
Serial2, Forward/Sparse, 00:00:32/00:02:28
```

State in "rtr-c" before switch

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

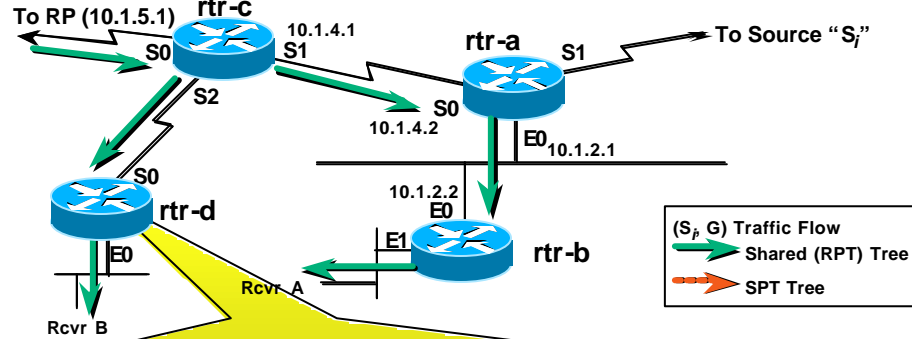
82

• PIM-SM SPT-Switchover Example

- Receivers A & B have joined multicast group 224.1.1.1 which has resulted in traffic flowing down the Shared Tree as shown by the solid arrows.
- The state in "rtr-c" prior to the switchover is as follows:
 - The IIF of the (*, G) entry points toward the RP via **Serial0**.
 - The OIL of the (*, G) entry contains **Serial1** and **Serial2** as a result of (*, G) Joins that were sent up the Shared Tree by "rtr-a" and "rtr-d", respectively.

PIM SM SPT-Switchover

Cisco.com



```
(*, 224.1.1.1), 00:01:43/00:02:13, RP 10.1.5.1, flags: SC
Incoming interface: Serial0, RPF nbr 10.1.4.8,
Outgoing interface list:
Ethernet0, Forward/Sparse, 00:01:43/00:02:11
```

State in "rtr-d" before switch

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

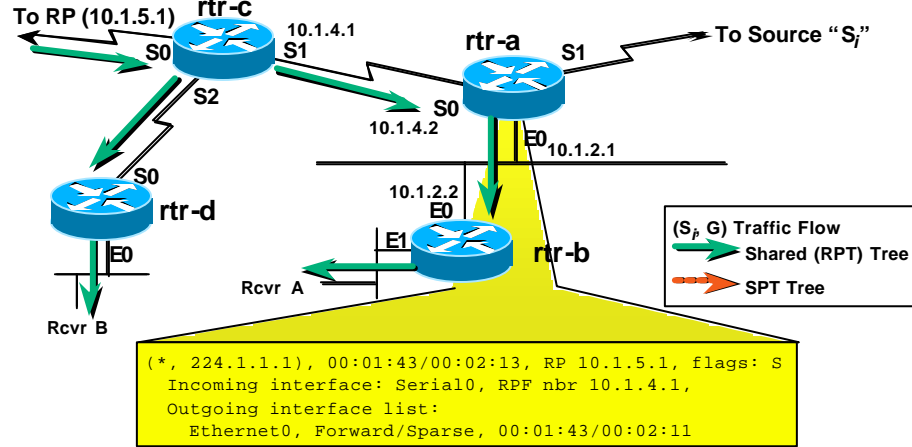
83

• PIM-SM SPT-Switchover Example

- The state in "rtr-d" prior to the switchover is as follows:
 - The IIF of the (*, G) entry points toward the RP via **Serial0**.
 - The OIL of the (*, G) entry contains **Ethernet0** as a result of the IGMP Reports for group 224.1.1.1 that are sent by Receiver "B".

PIM SM SPT-Switchover

Cisco.com



State in "rtr-a" before switch

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

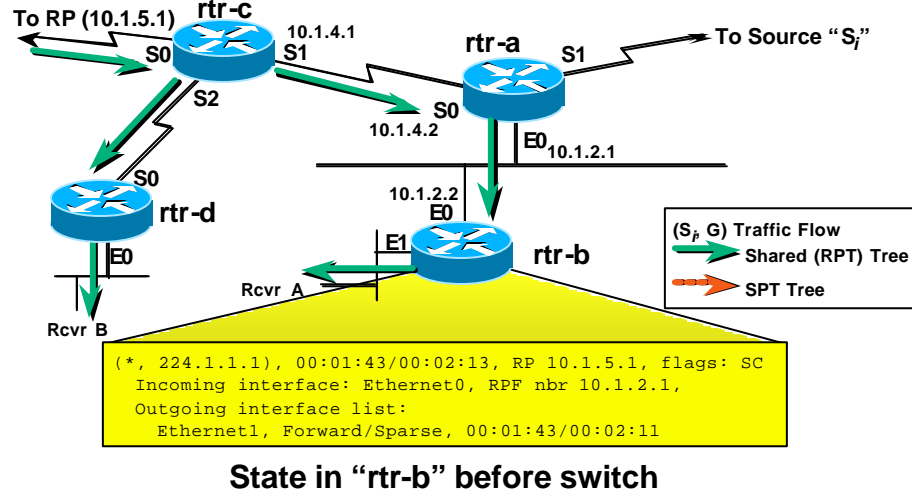
84

• PIM-SM SPT-Switchover Example

- The state in "rtr-a" prior to the switchover is as follows:
 - The IIF of the (*, G) entry points toward the RP via **Serial0**.
 - The OIL of the (*, G) entry contains **Ethernet0** as a result of (*, G) Joins that were sent up the Shared Tree by "rtr-b".

PIM SM SPT-Switchover

Cisco.com



Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

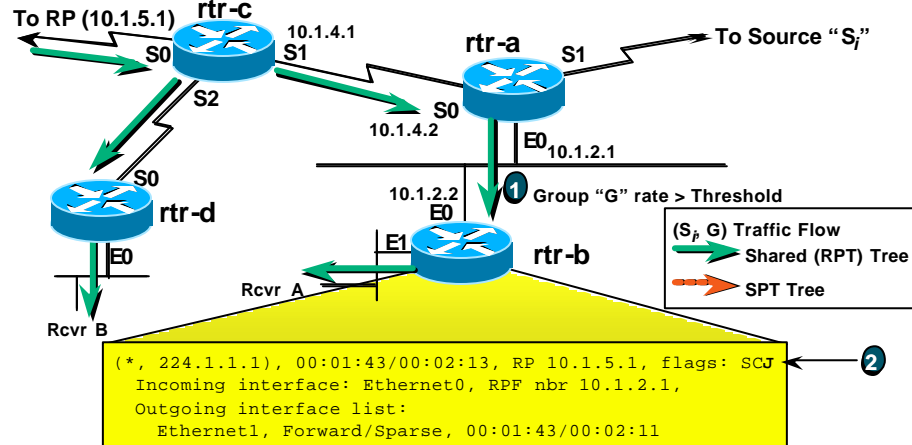
85

• PIM-SM SPT-Switchover Example

- The state in "rtr-b" prior to the switchover is as follows:
 - The IIF of the (*, G) entry points toward the RP via **Ethernet0**.
 - The OIL of the (*, G) entry contains **Ethernet1** as a result of the IGMP Reports for group 224.1.1.1 that are sent by Receiver "A".

PIM SM SPT-Switchover

Cisco.com



- 1 Group "G" rate exceeds SPT Threshold at "rtr-b";
- 2 Set J Flag in (*, G) and wait for next (S_p , G) packet.

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

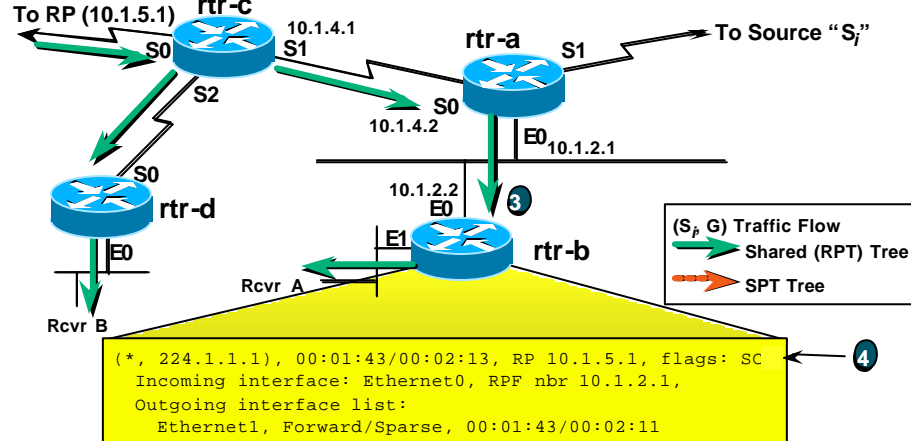
86

• PIM-SM SPT-Switchover Example

- Step 1: The total amount of all traffic flowing down the Shared Tree begins to exceed the SPT-Threshold configured at "rtr-b".
- Step 2: As a result, "rtr-b" sets the "J" flag in the (*, G) entry to denote that the rate is above the SPT-Threshold for this group.

PIM SM SPT-Switchover

Cisco.com



- ③ (S_pG) packet arrives down Shared tree.
- ④ Clear J Flag in the (*,G) & create (S_pG) state.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

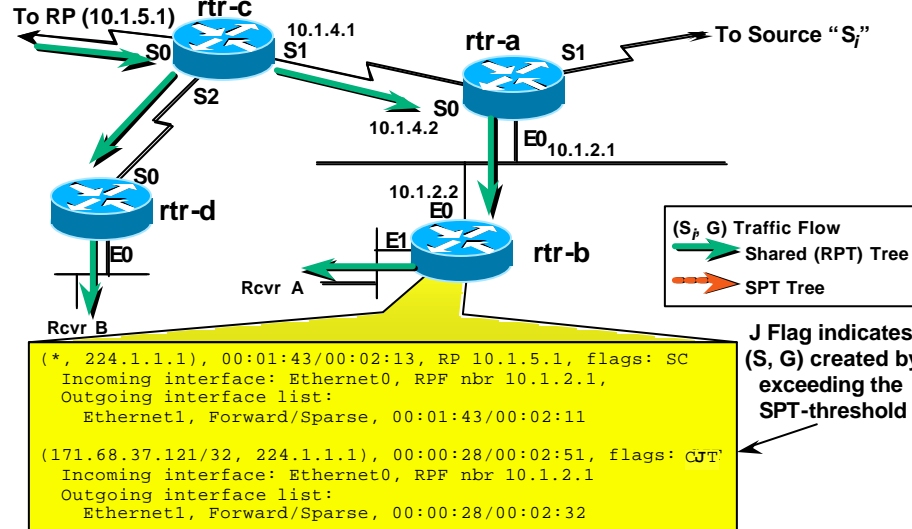
87

• PIM-SM SPT-Switchover Example

- Step 3: The very next packet to arrive via the Shared Tree happens to be from source (Si, G). Because there is a member directly connected to this router (denoted by the “C” flag) and the traffic rate is above the SPT-Threshold (denoted by the “J” flag), “rtr-b” initiates a switch to the SPT for (Si, G) traffic.
- Step 4: The “J” flag in the (*, G) entry is first cleared and a new traffic rate measurement interval (1 second) is started. Next, (Si, G) state is created for source “Si” sending to group “G”.

PIM SM SPT-Switchover

Cisco.com



New State in "rtr-b"

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

88

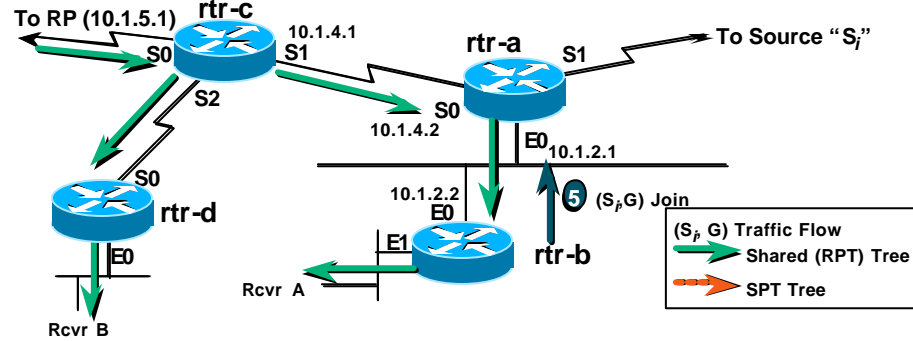
• PIM-SM SPT-Switchover Example

– The (171.68.37.121/32, 224.1.1.1) entry shown above is created as follows:

- To denote that this entry was created as a result of the SPT Switchover mechanism, the "J" flag is set on the (S, G) entry. (The "J" flag being set will cause "rtr-b" to monitor the rate of the (S, G) traffic *and* if the rate of this traffic drops below the SPT Threshold for over one minute, "rtr-b" will attempt to switch this traffic flow back to the Shared Tree.)
- The RPF information is calculated in the direction of source "Si". This results in an Incoming interface of **Ethernet0**, and an RPF neighbor address of "10.1.2.1.". (Note: That the RPF information for the (S, G) entry is the same as the (*, G) entry. This indicates that the Shared Tree and the SPT are following the same path at this point.)
- The OIL for the (S, G) entry is constructed by copying the OIL from the (*, G) entry and then removing the IIF from this list to prevent a possible route loop. This results in an (S, G) OIL containing only **Ethernet1**.

PIM SM SPT-Switchover

Cisco.com



5 Send (S_p, G) Join towards S_i .

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

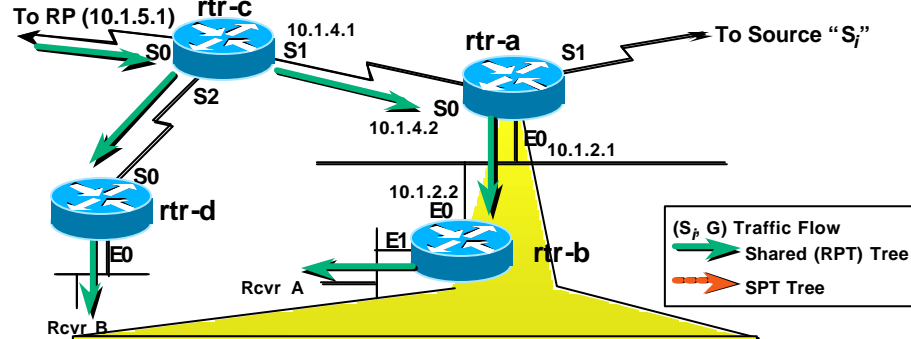
89

• PIM-SM SPT-Switchover Example

- Step 5: Once state has been created for (S_i, G), an (S, G) Join is sent toward source " S_i " to build a branch of the SPT to "rtr-b". These (S_i, G) Joins will continue to be sent periodically (once a minute) as long as the (S_i, G) entry is not Pruned (i.e. does not have a Null OIL).

PIM SM SPT-Switchover

Cisco.com



```

(*) , 224.1.1.1), 00:01:43/00:02:13, RP 10.1.5.1, flags: S
Incoming interface: Serial0, RPF nbr 10.1.4.1,
Outgoing interface list:
Ethernet0, Forward/Sparse, 00:01:43/00:02:11

(171.68.37.121/32, 224.1.1.1), 00:13:28/00:02:53, flags: T
Incoming interface: Serial1, RPF nbr 10.1.9.2
Outgoing interface list:
Ethernet0, Forward/Sparse, 00:13:25/00:02:30
    
```

New state in "rtr-a"

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

90

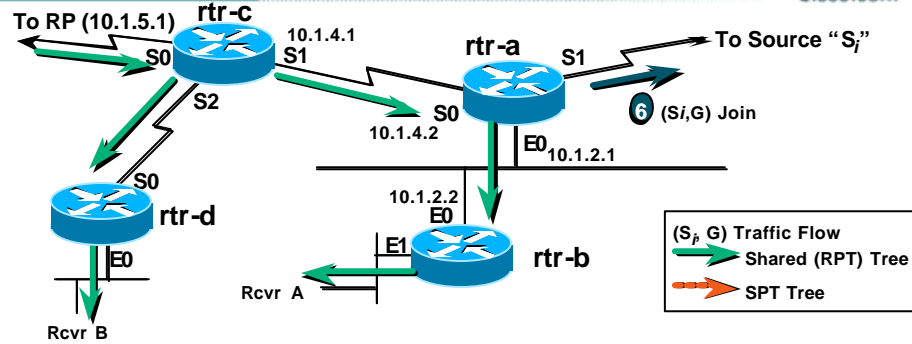
• PIM-SM SPT-Switchover Example

– When the (Si, G) Join is received by "rtr-a", the (171.68.37.121/32, 224.1.1.1) entry shown above is created as follows:

- The RPF information is calculated in the direction of source "Si". This results in an Incoming interface of **Serial1**, and an RPF neighbor address of "**10.1.9.1**". (Note: That the RPF information for the (S, G) entry is *not* the same as the (*, G) entry. This indicates that the paths of the Shared Tree and the SPT diverge at this point.)
- The OIL for the (S, G) entry is constructed by copying the OIL from the (*, G) entry and then removing the IIF from this list to prevent a possible route loop. This results in an (S, G) OIL containing only **Ethernet0**.

PIM SM SPT-Switchover

Cisco.com



⑥ “rtr-a” forwards (S_p, G) Join toward S_p

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

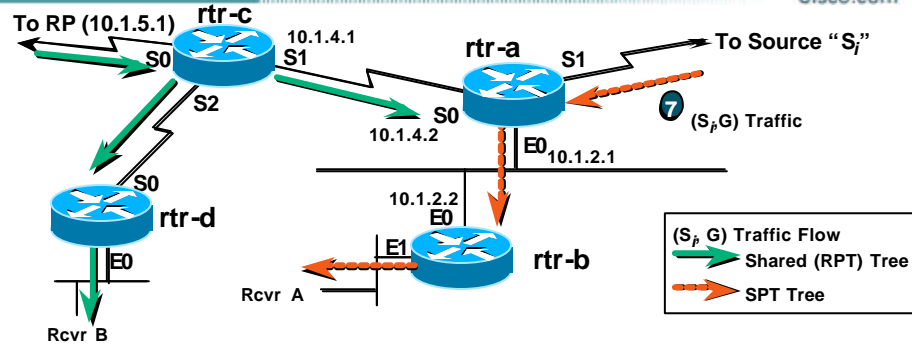
91

• PIM-SM SPT-Switchover Example

- Step 6: When the (S_i, G) state is created at “rtr-a”, an (S_i, G) Join is sent toward source “Si”. These (S_i, G) Joins will continue to be sent periodically (once a minute) as long as the (S_i, G) entry is not Pruned (i.e. does not have a Null OIL).

PIM SM SPT-Switchover

Cisco.com



- ⑥ "rtr-a" forwards (S_pG) Join toward S_p
- ⑦ (S_pG) traffic begins flowing down SPT tree.

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

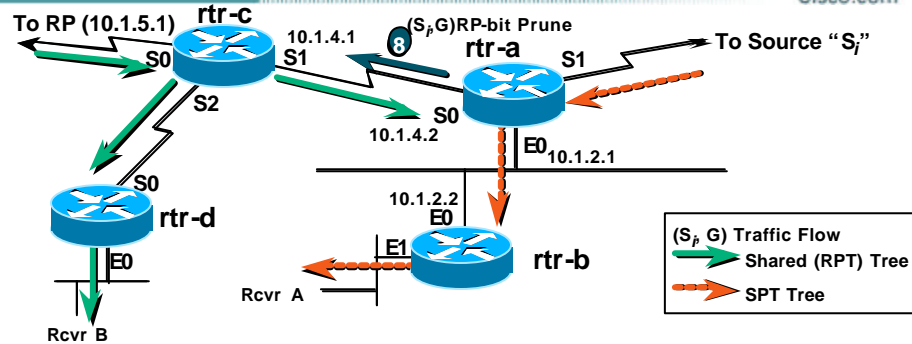
92

• PIM-SM SPT-Switchover Example

- Step 7: When the (Si, G) Joins reach the first-hop router directly connected to source "Si", a complete branch of the SPT has been built (shown by the dashed arrows). This permits (Si, G) traffic to flow via the SPT to "rtr-b" and receiver "A".

PIM SM SPT-Switchover

Cisco.com



- ⑥ "rtr-a" forwards (S_p, G) Join toward S_p
- ⑦ (S_p, G) traffic begins flowing down SPT tree.
- ⑧ SPT & RPT diverge, triggering (S_p, G) RP-bit Prunes toward RP.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

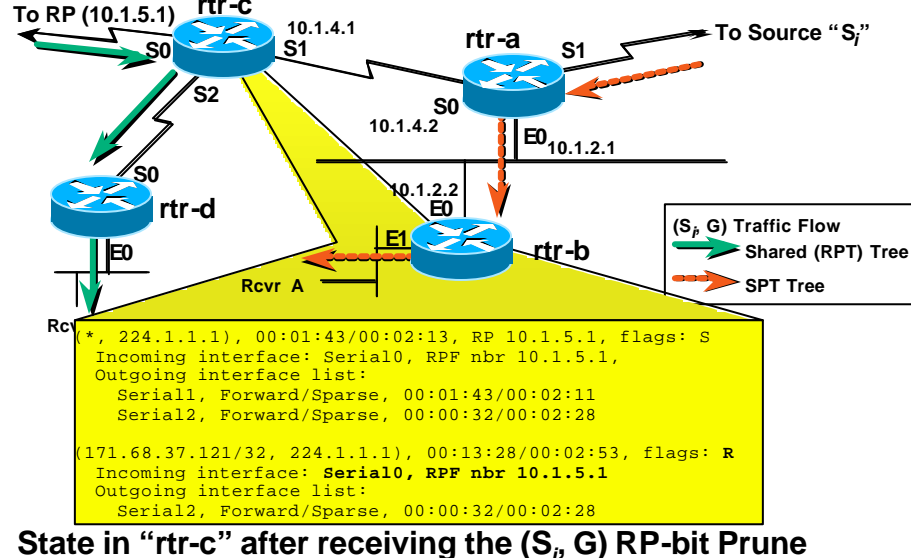
93

• PIM-SM SPT-Switchover Example

- Step 8: Because the paths of the Shared Tree and the SPT diverge at "rtr-a", (note the difference in RPF information on the previous page), this causes "rtr-a" to begin sending (S_i, G) RP-bit Prune messages up the Shared Tree to stop the flow of redundant (S_i, G) traffic down the Shared Tree. (Note: This step is delayed until traffic begins arriving via the SPT which is denoted by the "T" flag being set in the (S_i, G) entry in the mroute table.)

PIM SM SPT-Switchover

Cisco.com

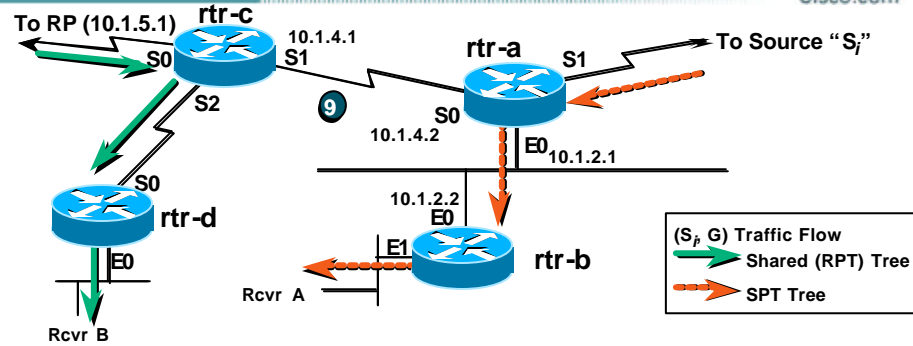


• PIM-SM SPT-Switchover Example

- When the (S_i, G)RP-bit Prune reaches "rtr-c", the (171.68.37.121/32, 224.1.1.1) entry shown above is created as follows:
 - Because this (S, G) entry was created as a result of the receipt of an (S,G)RP-bit Prune, the "R" bit is set to denote that this forwarding state is applicable to traffic flowing down the Shared Tree and not the Source Tree.
 - Because the "R" bit is set, the RPF information is calculated in the direction of the RP instead of source "S_i". (Remember, this entry is applicable to (S_i,G) traffic flowing down the *Shared Tree* and therefore the RPF information must point up the Shared Tree.) This results in an Incoming interface of **Serial0**, and an RPF neighbor address of "**10.1.5.1**".
 - The OIL for the (S, G) entry is constructed by copying the OIL from the (*, G) entry *minus the interface that the (S_i, G)RP-bit Prune was received*. Next, the IIF is removed from the OIL to prevent a possible route loop. These steps results in an (S, G) OIL containing only **Serial2**.
- At this point, (S_i, G) traffic flowing down the Shared Tree will be forwarded using the (S_i, G) entry. (S_i, G) traffic arriving at "rtr-a" will RPF correctly because the RPF information in the (S_i, G) entry is pointing up the Shared Tree (as a result of the "R" bit) and will then be forwarded out all interfaces in the (S_i, G) OIL. In this case, only **Serial2** remains in the (S_i, G) OIL and therefore (S_i, G) traffic will be sent to "rtr-d" *but not "rtr-a"*. This successfully prunes the redundant (S_i, G) traffic from the branch of the Shared Tree between "rtr-c" and "rtr-a".

PIM SM SPT-Switchover

Cisco.com



9 Unnecessary (S_p , G) traffic is pruned from the Shared tree.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

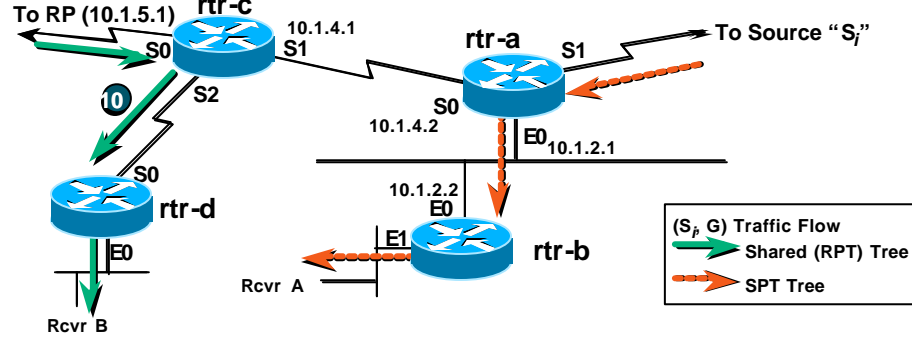
95

• PIM-SM SPT-Switchover Example

- Step 9: At this point, the redundant (S_i , G) traffic is pruned from the Shared Tree branch from “rtr-c” to “rtr-a”. (S_i , G) traffic is reaching receiver “A” via the SPT through “rtr-a” and “rtr-b”.

PIM SM SPT-Switchover

Cisco.com



- 9 Unnecessary (S_p, G) traffic is pruned from the Shared tree.
- 10 (S_p, G) traffic still flows via other branches of the Shared tree.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

96

• PIM-SM SPT-Switchover Example

- Step 10: (Si, G) traffic is still reaching receiver “B” via a branch of the Shared Tree through “rtr-c” and “rtr-d”. This is because the (Si, G) state in “rtr-c” still has Serial2 in its OIL.

Shared Tree Switchback Mechanism

- **Once each minute**
 - **If “J” flag set in (S_i, G) entry**
 - **Compute new (S_i, G) traffic rate**
 - **If rate < SPT-threshold**
 - **Rejoin $(*, G)$ Tree for (S_i, G) traffic**
 - **Send (S_i, G) prune up SPT toward S_i**
 - **Delete (S_i, G) entry**

• Shared Tree Switchback

- The Shared Tree Switchback (for lack of a better term) mechanism is used to switch sources back to the Shared Tree when their traffic rate falls below the SPT-Threshold.

• Switchback Algorithm

- The Switchback mechanism runs once a minute. (This helps prevent Sources from cycling between Shared Tree and Shortest-Path Tree too rapidly.)
- For each (S_i, G) entry in the Multicast Routing Table that has the “J” flag set, the mechanism computes the traffic rate for source S_i .
- If the rate has fallen below the SPT-Threshold, a switchback to the Shared Tree is initiated by the last-hop router by:
 - Sending a Join/Prune message that contains a $(*, G)$ Join without a (S_i, G) RP-bit Prune, up the Shared Tree (RPT). (This will cause the (S_i, G) Prune state along the RPT to be deleted which will permit (S_i, G) traffic to begin flowing down the RPT again.)
 - Deleting its (S_i, G) entry in the Multicast Routing Table.
 - Send (S_i, G) Prune up the Shortest-Path Tree (SPT) to stop traffic from flowing down the SPT.
- Note that this Switchback Algorithm is broken in older versions of IOS.

PIM-SM Protocol Mechanics

Cisco.com

- PIM Neighbor Discovery
- PIM State
- PIM SM Joining
- PIM SM Registering
- PIM SM SPT-Switchover
- **PIM SM Pruning**
- PIM SM State Maintenance

PIM SM Pruning

Cisco.com

- **IGMP group times out / last host sends Leave**
- **Interface removed from all (*,G) & (S,G) entries**
 - **IF all interfaces in “olist” for (*,G) are pruned;**
THEN send Prune up shared tree toward RP
 - **Any (S, G) state allowed to time-out**
- **Each router along path “prunes” interface**
 - **IF all interfaces in “olist” for (*,G) are pruned;**
THEN send Prune up shared tree toward RP
 - **Any (S, G) state allowed to time-out**

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

99

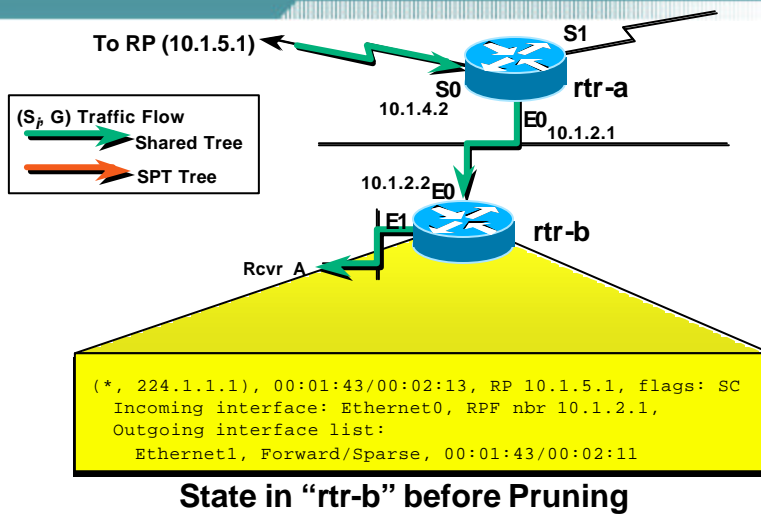
• SM Pruning

- Locally connected host sends an IGMP Leave (or IGMP state times out in the router) for group “G”.
- The interface is removed from the (*, G) and all (S, G) entries in the Multicast Routing Table.
 - If the (*, G) “Outgoing Interface list” is now Null, then send a (*, G) Prune up the Shared Tree (RPT) towards the RP.
 - Any remaining (S, G) entries are allowed to timeout and be deleted from the Multicast Routing Table.
- When the routers up the Shared Tree receive the (*, G) Prune, they remove the interface on which the Prune was received from their (*, G) “Outgoing interface list”.
 - If as a result of removing the interface the (*, G) “Outgoing Interface list” becomes Null, then forward a (*, G) Prune up the Shared Tree (RPT) towards the RP.
 - Any remaining (S, G) entries are allowed to timeout and be deleted from the Multicast Routing Table.

PIM SM Pruning

Shared Tree Case

Cisco.com



Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

100

• State in "rtr-b" before Pruning

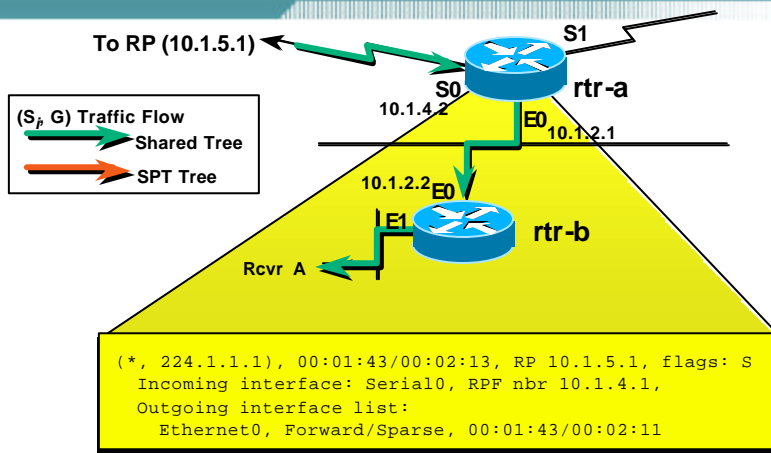
– Pay particular attention to the following:

- Traffic is flowing down the Shared Tree. (Denoted by the existence of only the (*, G) entry.)
- The **"Incoming interface"** is **Ethernet0**.
- The **"Outgoing interface list"** contains **Ethernet1**.
- The **"C"** flag is set in the (*, G) which denotes that there is a locally connected host for this group. (Rcvr A)

PIM SM Pruning

Shared Tree Case

Cisco.com



State in "rtr-a" before Pruning

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

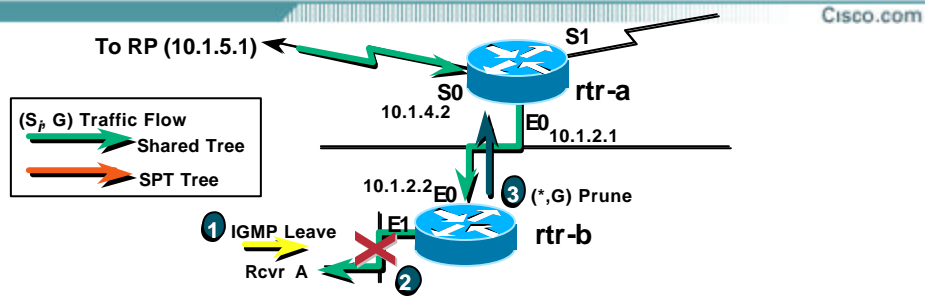
101

- State in "rtr-a" before Pruning — RPT Case

- Pay particular attention to the following:

- Traffic is flowing down the Shared Tree. (Denoted by the existence of only the (*, G) entry.)
 - The "Incoming interface" is Serial0.
 - The "Outgoing interface list" contains Ethernet0.

PIM SM Pruning Shared Tree Case



- 1 "rtr-b" is a Leaf router. Last host "Rcvr A", leaves group G.
- 2 "rtr-b" removes E1 from (*,G) and any (S_p,G) "oilists".
- 3 "rtr-b" (*,G) "oilist" now empty; sends (*,G) Prune toward RP.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

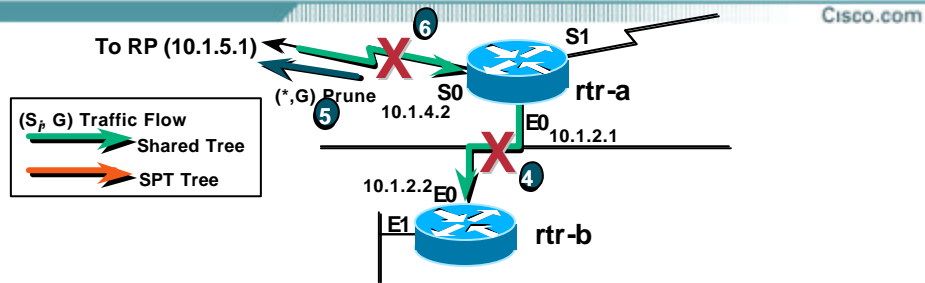
8/10/2001 2:37 PM

102

• PIM SM Pruning Example — RPT Case

- 1) The last-hop or Leaf router (rtr-b) receives an IGMP Group Leave message from "Rcvr A" for group "G". After performing the normal IGMP Leave processing and finding that "Rcvr A" was the last host to leave, the IGMP state for group "G" on interface "E1" is deleted.
- 2) This causes interface "E1" to be removed from the "Outgoing interface list" of the (*, G) entry and any (S_i, G) entries (in this case there are none) in the Multicast Routing Table. Because "E1" was the only interface in the (*, G) entry, it's outgoing interface list becomes Null.
- 3) Because the (*, G) "Outgoing interface list" is now Null, a (*, G) Prune is sent up the Shared Tree (RPT) via "E0" toward the RP.

PIM SM Pruning Shared Tree Case



- 4 "rtr-a" receives Prune; removes E0 from (*, G) "oilst".
(After the 3 second Multi-access Network Prune delay.)
- 5 "rtr-a" (*, G) "oilst" now empty; send (*, G) Prune toward RP.
- 6 Pruning continues back toward RP.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

103

• PIM SM Pruning Example — RPT Case (cont.)

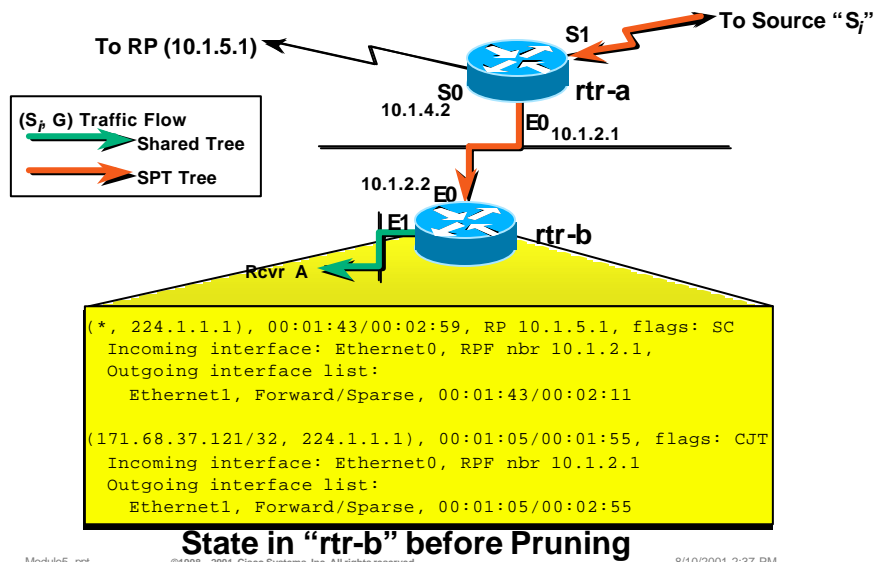
- 4) The (*, G) Prune is received by "rtr-a" which causes interface "E0" to be removed from the "Outgoing interface list" of the (*, G) entry in the Multicast Routing Table.

(Note: "rtr-a" delayed Pruning E0 from the (*, G) entry for 3 seconds since this is a Multi-Access network and it needed to wait for a possible overriding Join from another PIM neighbor. Since none was received, the interface was pruned.)

- 5) Because the (*, G) "Outgoing interface list" is now Null, a (*, G) Prune is forwarded on up the Shared Tree (RPT) via "S0" toward the RP.
- 6) This pruning continues back toward the RP or until a router is reached whose (*, G) "Outgoing interface list" doesn't go to Null as a result of the Prune.

PIM SM Pruning Source (SPT) Case

Cisco.com



• State in "rtr-b" before Pruning — SPT Case

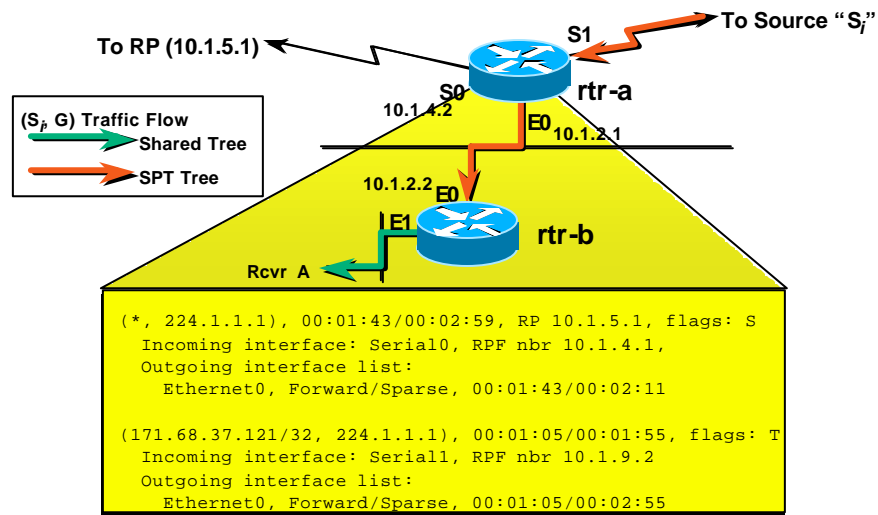
– Pay particular attention to the following:

- Both a (*, G) and (S, G) entries exist.
- The "J" flag is set in the (S, G) entry. This indicates that the (S, G) state was created as a result of the SPT-Threshold being exceeded.
- The "T" flag is set in the (S, G) entry. This indicates that (S, G) traffic is being successfully received via the Shortest-Path Tree (SPT).
- The "Incoming interface" is the same for the (*, G) and the (S, G) entry. This indicates that Shared Tree and the Shortest-Path tree are the same at this point.

PIM SM Pruning

Source (SPT) Case

Cisco.com



State in "rtr-a" before Pruning

Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

105

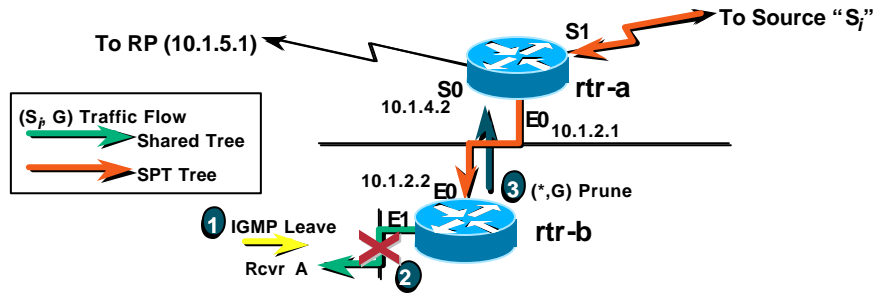
• State in "rtr-b" before Pruning — SPT Case

– Pay particular attention to the following:

- Both a (*, G) and (S, G) entries exist.
- The "T" flag is set in the (S, G) entry. This indicates that (S, G) traffic is being successfully received via the Shortest-Path Tree (SPT).
- The "Incoming interface" is different for the (*, G) and the (S, G) entry. This indicates that Shared Tree and the Shortest-Path tree diverge at this point.

PIM SM Pruning Source (SPT) Case

Cisco.com



- 1 "rtr-b" is a Leaf router. Last host "Rcvr A", leaves group G.
- 2 "rtr-b" removes E1 from (*,G) and any (S_p,G) "oilists".
- 3 "rtr-b" (*,G) "oilist" now empty; sends (*,G) Prune toward RP.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

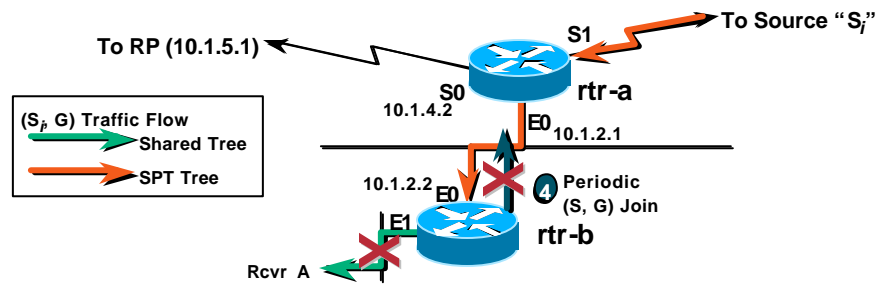
106

• PIM SM Pruning Example — SPT Case

- 1) The last-hop or Leaf router (rtr-b) receives an IGMP Group Leave message from "Rcvr A" for group "G". After performing the normal IGMP Leave processing and finding that "Rcvr A" was the last host to leave, the IGMP state for group "G" on interface "E1" is deleted.
- 2) This causes interface "E1" to be removed from the "Outgoing interface list" of the (*, G) entry and any (S_i, G) entries in the Multicast Routing Table. Because "E1" was the only interface in the (*, G) and the (S_i, G) entries, their outgoing interface lists become Null.
- 3) Because the (*, G) "Outgoing interface list" is now Null, a (*, G) Prune is sent up the Shared Tree (RPT) via "E0" toward the RP.

PIM SM Pruning Source (SPT) Case

Cisco.com



- ❶ “rtr-b” is a Leaf router. Last host “Rcvr A”, leaves group G.
- ❷ “rtr-b” removes E1 from (*,G) and any (S_i,G) “oilists”.
- ❸ “rtr-b” (*,G) “oilist” now empty; sends (*,G) Prune toward RP.
- ❹ “rtr-b” stops sending periodic (S, G) joins.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

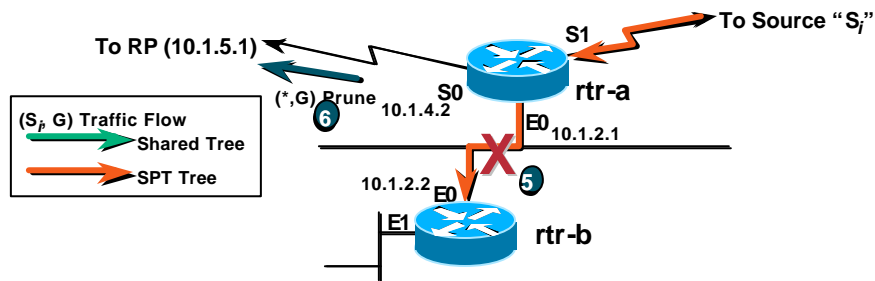
107

• PIM SM Pruning Example — SPT Case (cont.)

- 4) Because the (S_i, G) “Outgoing interface list” is now Null, “rtr-b” stops sending Periodic (S_i, G) Join messages up the Shortest-Path Tree (SPT).

PIM SM Pruning Source (SPT) Case

Cisco.com



- 5 "rtr-a" receives Prune; removes E0 from (*,G) "olist".
(After the 3 second Multiaccess Network Prune delay.)
- 6 "rtr-a" (*,G) "olist" now empty; sends (*,G) Prune toward RP.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

108

• PIM SM Pruning Example — SPT Case (cont.)

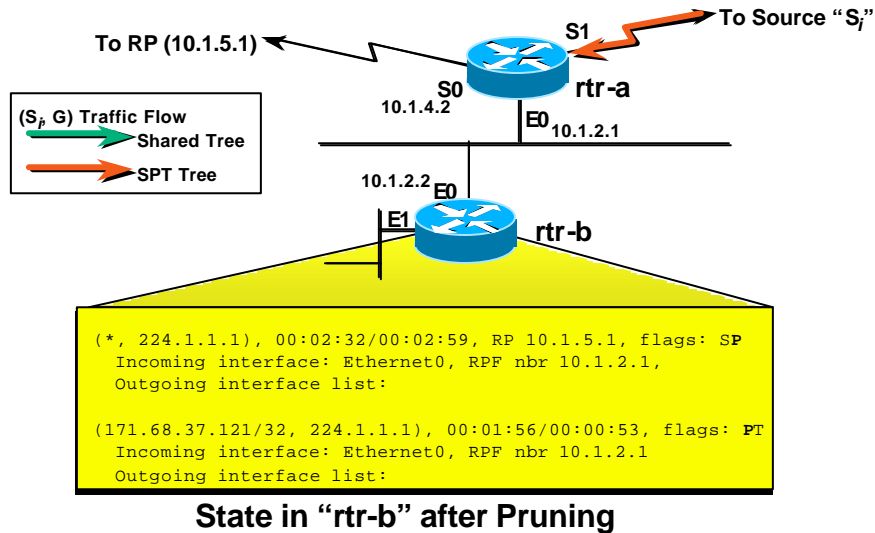
- 5) The (*, G) Prune is received by "rtr-a" which causes interface "E0" to be removed from the "Outgoing interface list" of the (*, G) entry in the Multicast Routing Table.

(Note: "rtr-a" delayed Pruning E0 from the (*, G) entry for 3 seconds since this is a Multi-Access network and it needed to wait for a possible overriding Join from another PIM neighbor. Since none was received, the interface was pruned.)

- 6) Because the (*, G) "Outgoing interface list" is now Null, a (*, G) Prune is forwarded on up the Shared Tree (RPT) via "S0" toward the RP.
- 7) Because "rtr-a" is no longer receiving (S_i, G) Join messages from "rtr-b", the (S_i, G) state eventually times out. This causes a (S_i, G) Prune to be sent up the Shortest-Path Tree (SPT) towards the source " S_i ".
- 8) Traffic stops flowing down the Shortest-Path Tree (SPT).

PIM SM Pruning Source (SPT) Case

Cisco.com



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

109

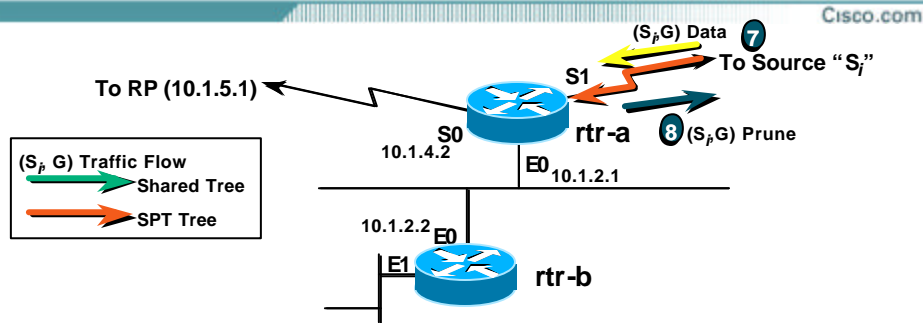
• PIM SM Pruning Example — SPT Case (cont.)

- 5) The (*, G) Prune is received by "rtr-a" which causes interface "E0" to be removed from the "Outgoing interface list" of the (*, G) entry in the Multicast Routing Table.

(Note: "rtr-a" delayed Pruning E0 from the (*, G) entry for 3 seconds since this is a Multi-Access network and it needed to wait for a possible overriding Join from another PIM neighbor. Since none was received, the interface was pruned.)

- 6) Because the (*, G) "Outgoing interface list" is now Null, a (*, G) Prune is forwarded on up the Shared Tree (RPT) via "S0" toward the RP.
- 7) Because "rtr-a" is no longer receiving (S_i, G) Join messages from "rtr-b", the (S_i, G) state eventually times out. This causes a (S_i, G) Prune to be sent up the Shortest-Path Tree (SPT) towards the source "S_i".
- 8) Traffic stops flowing down the Shortest-Path Tree (SPT).

PIM SM Pruning Source (SPT) Case



7 Another (S_p, G) data packet arrives via Serial1.

8 'rtr-a' responds by sending an (S_p, G) Prune toward source.

• PIM SM Pruning Example — SPT Case (cont.)

5) The $(*, G)$ Prune is received by "rtr-a" which causes interface "E0" to be removed from the "Outgoing interface list" of the $(*, G)$ entry in the Multicast Routing Table.

(Note: "rtr-a" delayed Pruning E0 from the $(*, G)$ entry for 3 seconds since this is a Multi-Access network and it needed to wait for a possible overriding Join from another PIM neighbor. Since none was received, the interface was pruned.)

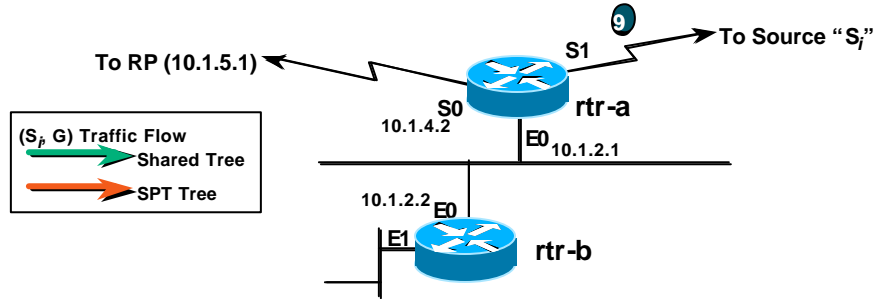
6) Because the $(*, G)$ "Outgoing interface list" is now Null, a $(*, G)$ Prune is forwarded on up the Shared Tree (RPT) via "S0" toward the RP.

7) Because "rtr-a" is no longer receiving (S_i, G) Join messages from "rtr-b", the (S_i, G) state eventually times out. This causes a (S_i, G) Prune to be sent up the Shortest-Path Tree (SPT) towards the source " S_i ".

8) Traffic stops flowing down the Shortest-Path Tree (SPT).

PIM SM Pruning Source (SPT) Case

Cisco.com



- 7 Another (S_p, G) data packet arrives via Serial1.
- 8 'rtr-a' responds by sending an (S_p, G) Prune toward source.
- 9 (S_p, G) traffic ceases flowing down SPT.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

142

• PIM SM Pruning Example — SPT Case (cont.)

- 5) The $(*, G)$ Prune is received by "rtr-a" which causes interface "E0" to be removed from the "Outgoing interface list" of the $(*, G)$ entry in the Multicast Routing Table.

(Note: "rtr-a" delayed Pruning E0 from the $(*, G)$ entry for 3 seconds since this is a Multi-Access network and it needed to wait for a possible overriding Join from another PIM neighbor. Since none was received, the interface was pruned.)

- 6) Because the $(*, G)$ "Outgoing interface list" is now Null, a $(*, G)$ Prune is forwarded on up the Shared Tree (RPT) via "S0" toward the RP.
- 7) Because "rtr-a" is no longer receiving (S_i, G) Join messages from "rtr-b", the (S_i, G) state eventually times out. This causes a (S_i, G) Prune to be sent up the Shortest-Path Tree (SPT) towards the source " S_i ".
- 8) Traffic stops flowing down the Shortest-Path Tree (SPT).

PIM-SM Protocol Mechanics

Cisco.com

- PIM Neighbor Discovery
- PIM State
- PIM SM Joining
- PIM SM Registering
- PIM SM SPT-Switchover
- PIM SM Pruning
- **PIM SM State Maintenance**

PIM SM State Maintenance

Cisco.com

- **Periodic Join/Prunes are sent to all PIM neighbors.**
- **Periodic Joins refresh interfaces in a PIM neighbor's "oilists".**
- **Periodic Prunes refresh prune state in a PIM neighbor.**
- **Received Multicast packets reset (S,G) entry "expiration" timers.**

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

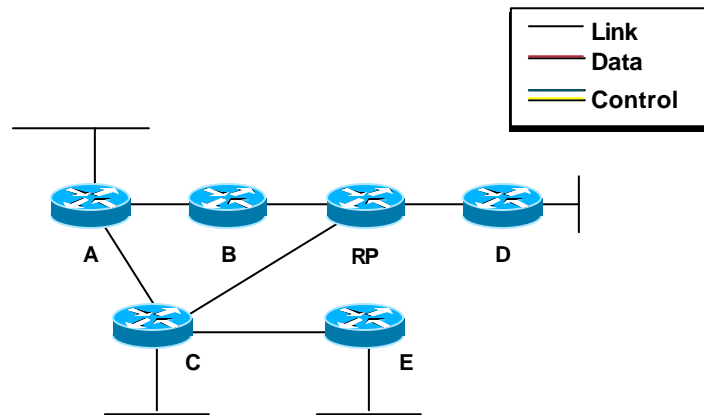
144

• PIM SM State Maintenance

- In PIM SM, Join/Prune state information has a normal expiration time of 3 minutes. If a periodic Join/Prune message is not received to refresh this state information, it automatically expires and is deleted. Therefore, a PIM router sends periodic Join/Prune messages to its PIM neighbors to maintain this state information.
- When a Join message is received from a PIM Neighbor, the expiration timer of the interface (in the Outgoing interface list) from which the Join was received is reset to 3 minutes. If the interface "expiration" timer goes to zero, the interface is removed from the "Outgoing interface list". (Note: This can trigger a Prune if the removal of the interface causes the "Outgoing interface list" to become Null.)
- When a Prune message is received in PIM Sparse mode, the interface on which the Prune was received is normally just removed from the Outgoing interface list. The exception to this is the special case of (S, G)RP-bit Prunes which are used to Prune (S, G) traffic from the Shared Tree. In this case, periodic (S, G)RP-bit Prunes must be sent in order to "refresh" the prune state in the upstream PIM Neighbor toward the RP.
- All (S, G) entries have entry expiration timers which are reset to 3 minutes by the receipt of an (S, G) packet received via the Shortest-Path Tree (SPT). If the source stops sending, this expiration timer goes to zero and the (S, G) entry is deleted.

PIM Sparse Mode Review

Cisco.com



Module5.ppt

©1998 - 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

145

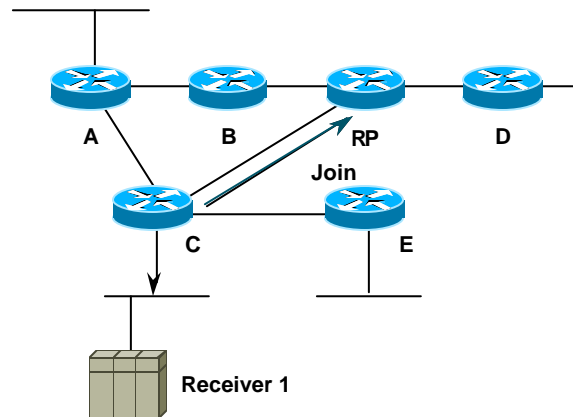
- **PIM Sparse Mode Review**

- The following slides will review all the major concepts previously present in a sample network situation.

PIM Sparse Mode Review

Cisco.com

Receiver 1 Joins Group G
C Creates (*, G) State, Sends
(*, G) Join to the RP



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

146

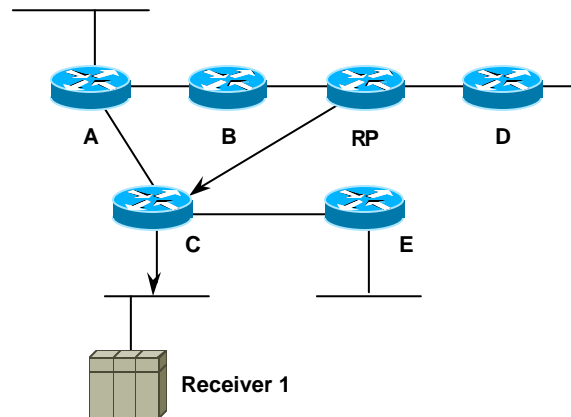
• PIM Sparse Mode Review (cont.)

- Receiver 1 joins group “G” by sending an IGMP Host message to “C”.
- “C” creates a (*, G) entry that has the interface towards Receiver 1 in the “Outgoing interface list”.
- “C” then sends a (*, G) Join up the Shared Tree toward the Rendezvous Point (RP).

PIM Sparse Mode Review

Cisco.com

RP Creates (*, G) State



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

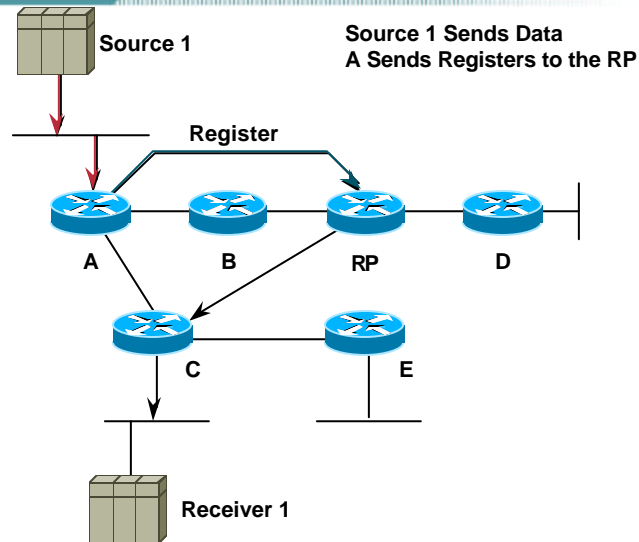
117

- **PIM Sparse Mode Review (cont.)**

- The Rendezvous Point (RP) receives the (*, G) Join from “C” and creates a (*, G) entry that has the interface towards “C” in the “Outgoing interface list”.
- The Shared Tree for group “G” has now been built (indicated by the arrows in the drawing) down to “Receiver 1”.

PIM Sparse Mode Review

Cisco.com



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

148

- **PIM Sparse Mode Review (cont.)**

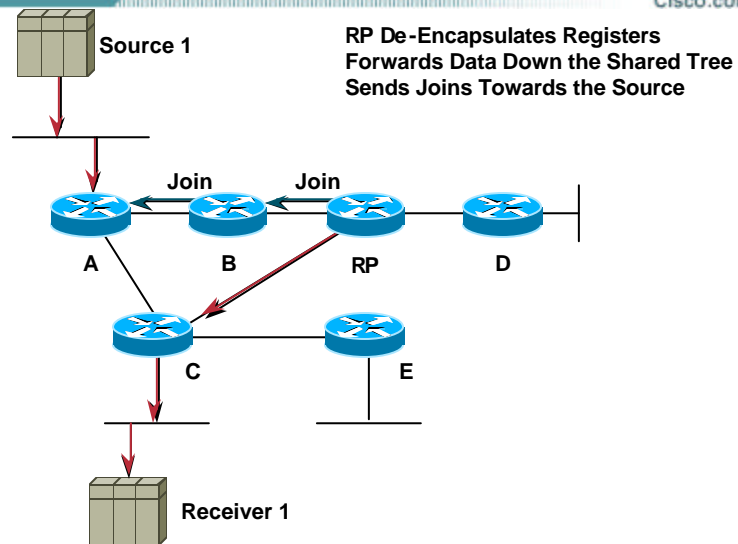
- “Source 1” begins sending data to group “G”.

(Note: It is not necessary for “Source 1” to join group “G” by sending an IGMP Host Membership message before sending to group “G”.)

- “A” encapsulates “Source 1” multicast packets in PIM Register messages and unicasts them to the Rendezvous Point (RP).

PIM Sparse Mode Review

Cisco.com



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

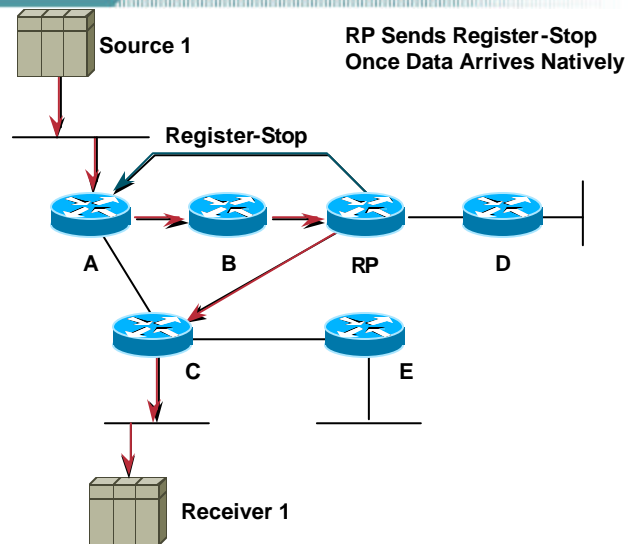
149

- **PIM Sparse Mode Review (cont.)**

- The Rendezvous Point (RP) receives and de-encapsulates the Register messages and finds that it contains a packet for group “G” for which it has (*, G) state.
- The RP forwards the de-encapsulated packet down the Shared Tree.
- The RP then sends a Join towards “Source 1” so that it can begin receiving “native” (I.e. unencapsulated) packets from the “Source 1”.

PIM Sparse Mode Review

Cisco.com



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

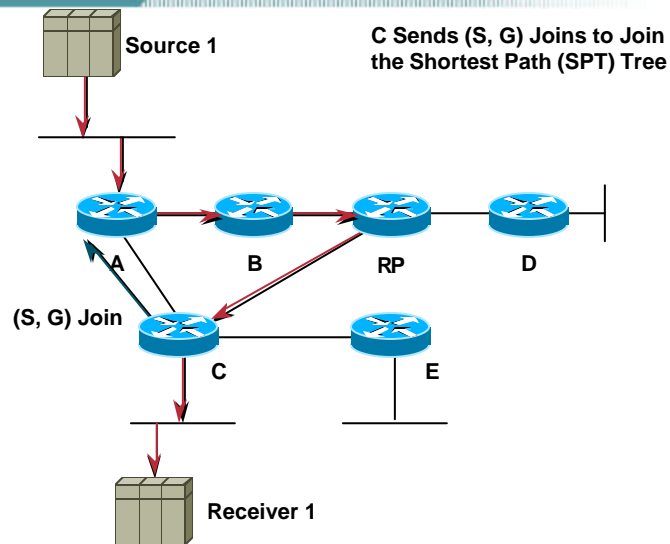
120

- **PIM Sparse Mode Review (cont.)**

- Once the data from “Source 1” begins arriving natively, the Rendezvous Point (RP) sends a “Register-Stop” message to notify “A” that it no longer needs to encapsulate traffic in Register messages.

PIM Sparse Mode Review

Cisco.com



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

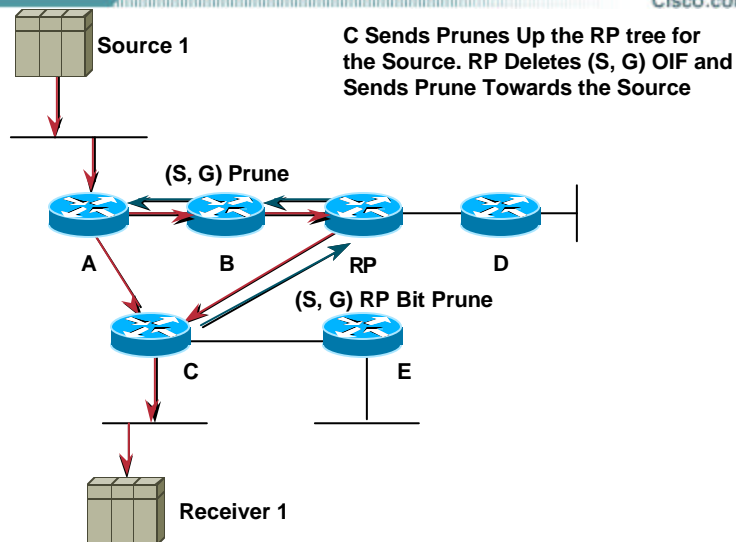
121

- **PIM Sparse Mode Review (cont.)**

- Traffic exceeds the SPT-Threshold and “C” begins the process of switching to the Shortest-Path Tree (SPT) for “Source 1” by sending an (S, G) Join to “A” up the SPT towards “Source 1”.

PIM Sparse Mode Review

Cisco.com



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

122

• PIM Sparse Mode Review (cont.)

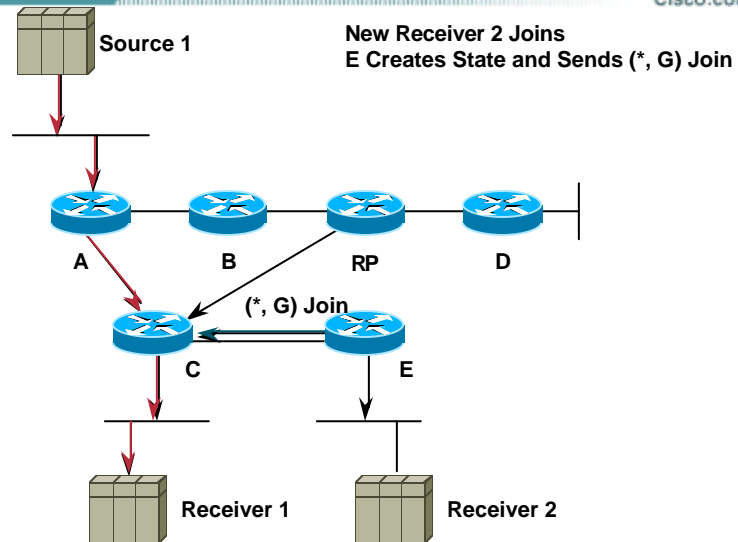
- In order to prune “Source 1” traffic from the Shared Tree (it’s now receiving this traffic via the Shortest-Path Tree), “C” sends an (S, G)RP-bit Prune up the Shared Tree.

(Note: The RP-bit indicates to up-stream routers that this prune should flow up the Shared Tree (RPT) to the RP.)

- The RP receives the (S, G)RP-bit Prune and removes the interface (towards “C”) from the (S, G) “olist”. This stops the flow of “Source 1” traffic down the Shared Tree to “C”.
- The RP’s (S, G) “olist” is now Null (i.e. there are no other branches on the Shared Tree that want “Source 1” traffic) and therefore no longer needs “Source 1” traffic. The RP responds by sending (S, G) Prunes towards “Source 1”. This stops the flow of “Source 1” traffic to the RP.

PIM Sparse Mode Review

Cisco.com



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

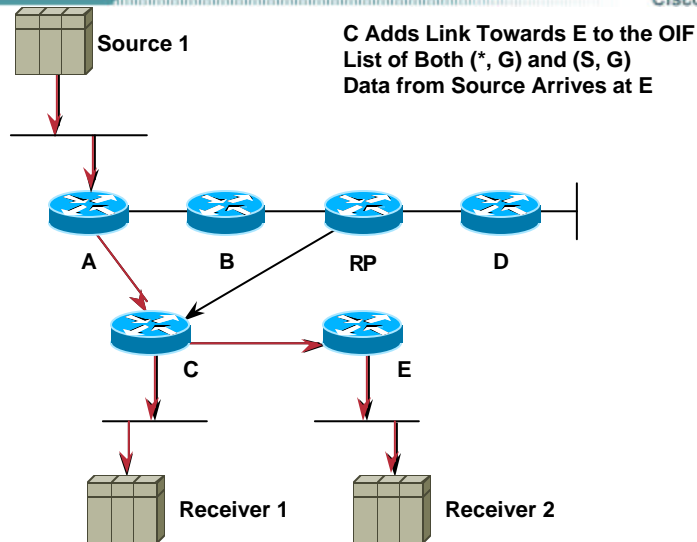
123

• PIM Sparse Mode Review (cont.)

- Receiver 2 also joins group “G” by sending an IGMP Host message to “E”.
- “E” creates a (*, G) entry that has the interface towards “Receiver 2” in the “Outgoing interface list”.
- “E” then sends a (*, G) Join up the Shared Tree toward the Rendezvous Point (RP).

PIM Sparse Mode Review

Cisco.com



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

124

• PIM Sparse Mode Review (cont.)

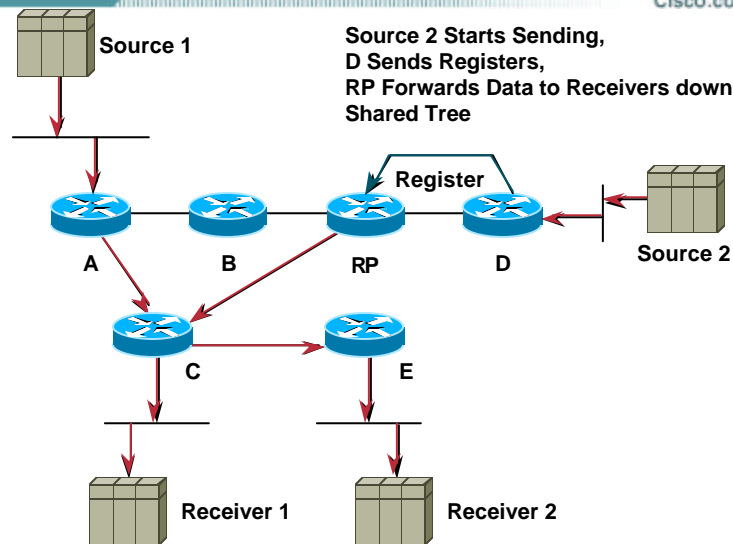
- “C” receives the (*, G) Join from “E” and adds the interface (towards “E”) to the “olist” in both the existing (*, G) and (S, G) entries in the Multicast Routing table.

(Note: Since “C” already had (*, G) state, it is not necessary to send a (*, G) join toward the RP again.)

- Group “G” traffic now begins flowing through “C” and “E” to “Receiver 2”.

PIM Sparse Mode Review

Cisco.com



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

125

• PIM Sparse Mode Review (cont.)

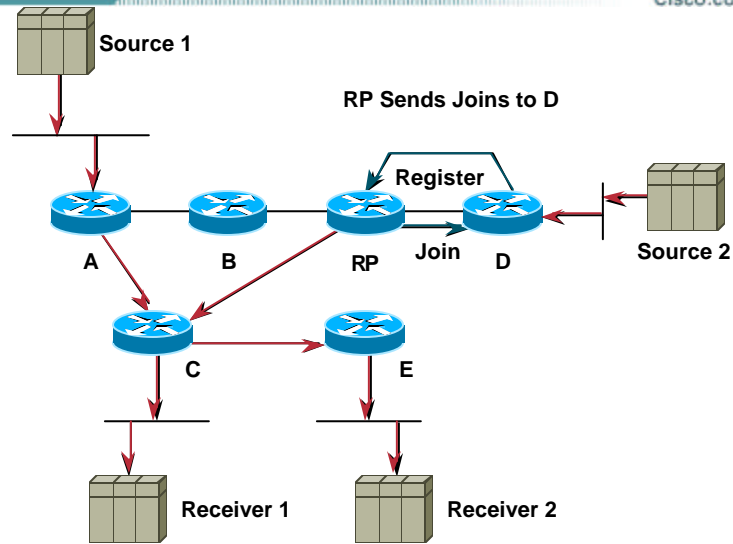
- “Source 2” now begins sending data to group “G”.

(Note: Again, it is not necessary for “Source 2” to join group “G” by sending an IGMP Host Membership message before sending to group “G”.)

- “D” encapsulates “Source 2” multicast packets in PIM Register messages and unicasts them to the Rendezvous Point (RP).
- The Rendezvous Point (RP) receives and de-encapsulates the Register messages and finds that it contains a packet for group “G” for which it has (*, G) state.
- The RP forwards the de-encapsulated packet down the Shared Tree.

PIM Sparse Mode Review

Cisco.com



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

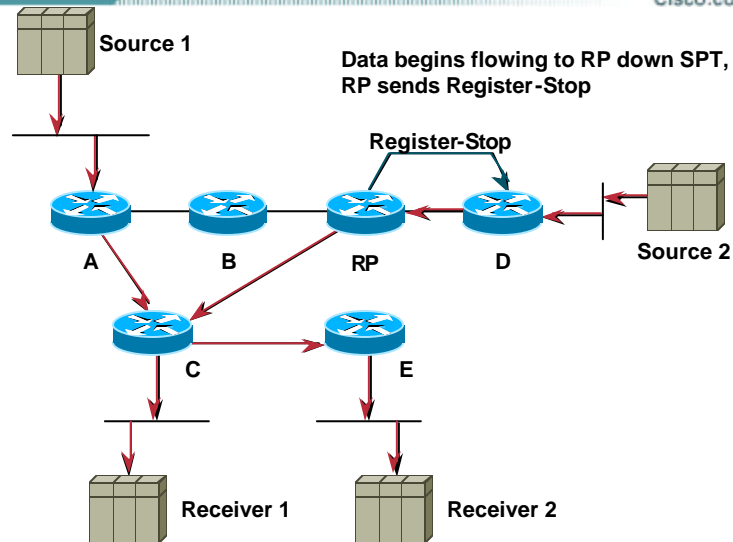
126

- **PIM Sparse Mode Review (cont.)**

- The RP then sends a Join towards “Source 2” so that it can begin receiving “native” (I.e. unencapsulated) packets from the “Source 2”.

PIM Sparse Mode Review

Cisco.com



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

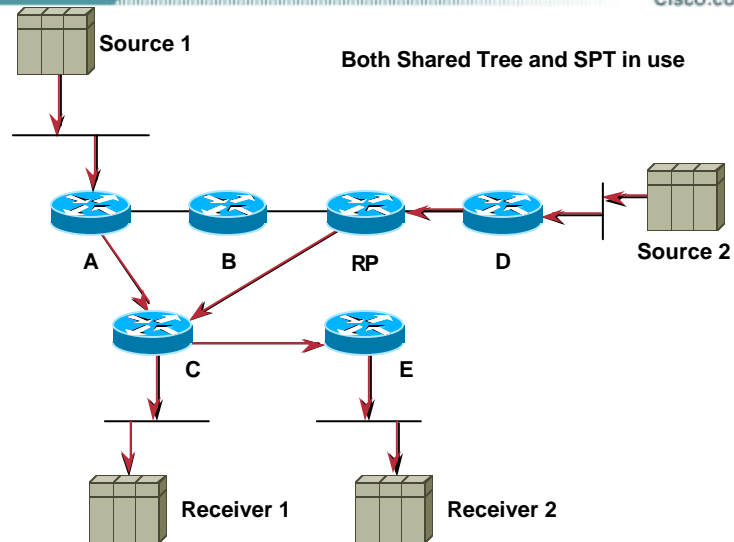
127

• PIM Sparse Mode Review (cont.)

- Data from “Source 2” begins arriving natively down the Shortest-Path Tree (SPT) via “D”.
- The Rendezvous Point (RP) sends a “Register-Stop” message to notify “D” that it no longer needs to encapsulate traffic in Register messages.

PIM Sparse Mode Review

Cisco.com



Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

128

- **PIM Sparse Mode Review (cont.)**

- At this point, both the Shared Tree (Source 2 traffic) and a Shortest-Path Tree (Source 1 traffic) are being used to deliver group “G” traffic to the Receivers.

PIM-SM Configuration Steps

Cisco.com

- **Enable Multicast Routing on every router**
- **Configure every interface for PIM**
- **Configure the RP**
 - **Using Auto-RP or BSR**
 - **Configure certain routers as Candidate RP(s)**
 - **Configure certain routers as Mapping Agents/C-BSR's**
 - **All other routers automatically learn elected RP**
 - **Anycast/Static RP addressing**
 - **RP address must be configured on every router**
 - **Note: Anycast RP requires MSDP**

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

129

- **Enable Multicast Routing (Forwarding) on every router.**

This insures that there will be no multicast black-holes caused by a non multicast router in the unicast RPF path back to the source or RP.

- **Configure every interface for PIM**

This insures that there will be no multicast black-holes caused by a non multicast interface in the unicast RPF path back to the source or the RP.

- **Configure the RP**

- Auto-RP (or BSR) are the simplest forms of RP configuration as they allow the routers in the network to automatically learn the address of the RP. This requires two additional command lines on one or more routers in the network that have been selected as candidate RP's.
 - Configure one or more routers as Candidate RP's using the appropriate Auto-RP or BSR command.
 - Configure one or more routers as Mapping Agents (Auto-RP) or Candidate BSR's (BSR).
- Anycast/Static RP addressing takes more work as the *single* RP address must be configured on every router in the network.
 - Anycast RP is a form of redundant static RP's which requires the use of the Multicast Source Discovery Protocol (MSDP) but provides rapid RP failover.

Enabling Multicast on the Router

Cisco.com

- **Global Configuration Command**

- `ip multicast-routing`

- Enables IP multicast forwarding in the router.
 - Configure on **EVERY** router in the network.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

130

- **Enabling Multicast on the Router**

- `ip multicast-routing`

- This command enables IP multicast forwarding on the router.
 - Be sure to configure this command on every router in the network to avoid blackholes caused by a non multicast appearing in the unicast RPF path to a source or the RP.

Enabling Multicast on an Interface

Cisco.com

- **Interface Configuration Commands**

- Enables multicast forwarding on the interface.
- Controls the interface's mode of operation.

`ip pim dense-mode`

- Interface mode is hard-wired to Dense mode operation.

`ip pim sparse-mode`

- Interface mode is hard-wired to Sparse mode operation.

`ip pim sparse-dense-mode`

- Interface mode is determined by the Group mode.
 - If Group is Dense, interface operates in Dense mode.
 - If Group is Sparse, interface operates in Sparse mode.
 - Decision is made on a packet-by-packet basis.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

131

- **Enabling Multicast on an Interface**

- The following commands enable multicast forwarding on an interface as well as determining the mode in which it operates.
- `ip pim dense-mode`
 - Causes the interface to be hard-wired into operating in dense mode for all multicast traffic flows.
- `ip pim sparse-mode`
 - Causes the interface to be hard-wired into operating in sparse mode for all multicast traffic flows.
- `ip pim sparse-dense-mode`
 - Causes the interface to dynamically determine the interface mode on a packet-by-packet basis depending on whether the destination group is Dense mode, or Sparse mode. (This mode is shown by the "D" or "S" flag on the (*,G) entry.)
 - If the destination group of a packet is in Dense mode, the interface uses dense mode operation to forward the packet. If the destination group of a packet is in Sparse mode, the interface uses sparse mode operation to forward the packet.

Group Mode vs. Interface Mode

Cisco.com

- **Avoid Group/Interface mode mismatches.**
 - Group and Interface mode should be the same.
 - Otherwise you may get unwanted/unpredictable results.
- **Sparse-Dense interfaces *always* match the Group mode.**
 - Should normally be used if running Auto-RP.
 - Permits Auto-RP groups to automatically run in Dense mode.
 - All other groups run in Sparse mode. (Assuming an RP is defined for all other groups.)
 - Can also be used for Sparse-only or Dense-only networks.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

132

- **Group Mode vs. Interface mode**

- Care should be taken to make sure that the Group mode always matches the interface mode. Otherwise, the forwarding mechanisms may not perform as desired/predicted.
- Group/Interface mismatches can be avoided by configuring

```
ip pim sparse-dense-mode
```

on an interface. This results in the interface always matching the Group mode.

- Sparse/Dense mode should normally be used if running Auto-RP. This allows the two Auto-RP groups (224.0.1.39 and 224.0.1.40) to operate in Dense mode. Once the routers in the network learn (via Auto-RP) the address of the elected RP, they will operate all other groups in Sparse mode. (By default, Auto-RP learned Group-to-RP ranges *never* include the Auto-RP groups.)
- Sparse/Dense mode can be also be used if pure Sparse mode or pure Dense mode operation is desired by either configuring or not configuring an RP address on each router, respectively.

Group Mode vs. Interface Mode

Cisco.com

Common Misconception

- **Interface Mode controls Group Mode.**

“If I set all interfaces to ‘ip pim sparse-mode’, the router will always operate in Sparse mode and never fall back into Dense mode.”

Bzzzztt!!! I’m sorry, but that’s the incorrect answer.

- **Group mode is independent of interface mode.**

- **Interface mode only controls how the interface operates.**

- ***Group mode is controlled by RP information!!!***

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

133

- **Common Misconception**

- “Interface Mode controls Group Mode.”
- This is a classic error often made by network administrators. They assume that, “If I set all interfaces to ‘ip pim sparse-mode’, the router will always operate in Sparse mode and never fall back into Dense mode.” Unfortunately, this is incorrect.
- Group mode is *solely* controlled by the existence of a valid RP. If a valid RP is learned/configured for a group range, those groups will operate in sparse mode and the (*,G) entry will be created with the “S” flag set. Otherwise, the groups will operate in Dense mode and the “D” flag will be set on the (*,G) entry.

Interface Mode Summary

Cisco.com

- **Let Group mode control Interface mode.**
 - Use 'ip pim sparse-dense-mode' command.
 - Allows maximum flexibility.
 - No need to ever change interface configuration.
- **Control Group mode with RP info.**
 - If RP info exists, Group = Sparse.
 - Therefore interface mode = Sparse for this group.
 - If RP info does not exist, Group = Dense.
 - Therefore interface mode = Dense for this group.

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

134

• Interface Mode Summary

- By configuring all interfaces with 'ip pim sparse-dense-mode' we allow the Group mode to determine the interface mode on a packet-by-packet basis. This allows maximum flexibility since we never have to reconfigure the interface to change mode.
- Control the Group mode (and hence the interface mode if 'ip pim sparse-dense-mode' is configured) by defining RP information in the network.
 - If a router has a valid RP address for a particular group, the group will be created in sparse mode; thereby causing the interface to operate in sparse mode when 'ip pim sparse-dense-mode' is configured.
 - If a router does *not* have a valid RP address for a particular group, the group will be created in dense mode; thereby causing the interface to operate in dense mode when 'ip pim sparse-dense-mode' is configured.

Avoiding DM Fallback

Cisco.com

“To always guarantee Sparse mode operation (and avoid falling back to Dense mode), make sure that every router *always* knows of an RP for every group.”

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

135

- **How to avoid falling back into Dense mode.**

It is often desired that the network *NEVER* fall back into Dense mode. Even if all primary and backup RP's fail, it is often better to have multicast forwarding stop instead of reverting back to dense mode.

- In order to prevent falling back into Dense mode, make sure that there is *always* an RP learned/configured for the entire multicast group range.

Avoiding DM Fallback

Cisco.com

- **Define an "RP-of-last-resort".**
 - **Configure as a Static RP on every router.**
 - Will only be used if all Candidate-RP's fail.
 - Can be a dummy address.
 - Recommendation: Use lowest priority C-RP address.
 - **Use ACL to avoid breaking Auto-RP**

```
ip pim rp-address <RP-of-last-resort> 10
access-list 10 deny 224.0.1.39
access-list 10 deny 224.0.1.40
access-list 10 permit any
```

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

136

• Avoiding DM Fallback

– In order to guarantee that the router will *never* fall back into dense mode, it is necessary to guarantee that the router will *never* lose RP information. This can be accomplished by defining a static, "RP-of-last-resort" in each router in the network. Since automatically learned RP's (Auto-RP or BSR) take precedence over statically defined RP's, the static entry will only be activated if all learned RP's timeout and/or fail.

- The recommendation is to define the lowest priority Candidate RP as the "RP-of-last-resort" by using a static RP definition pointing to this IP address. This locks the lowest priority RP into the bottom of the failover order. Even if this router fails (or its information times out), the static entry in each router will prevent a total loss of RP information.
- Special care must be taken if an "RP-of-last-resort" is defined when using Auto-RP. By default, a static RP definition that covers the Auto-RP group range *will* be interpreted as the RP for the two Auto-RP groups. (Unlike Auto-RP learned group ranges which have an implied *deny* for these two groups so that the two Auto-RP groups will default to using dense mode.)

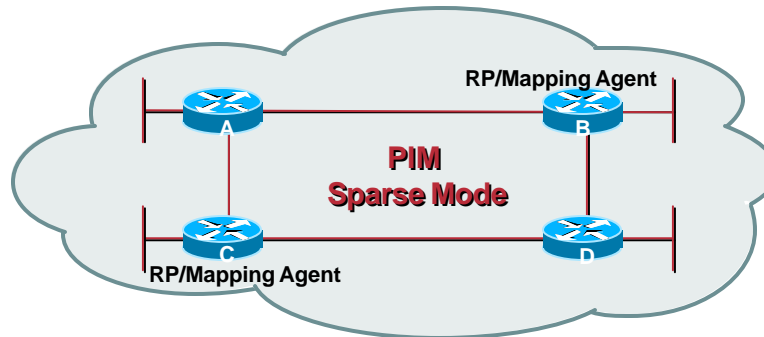
The following example shows how to configure an "RP-of-last-resort" so that the two Auto-RP groups do not accidentally switch to sparse mode:

```
ip pim rp-address <RP-of-last-resort> 10

access-list 10 deny 224.0.1.39
access-list 10 deny 224.0.1.40
access-list 10 permit any
```


Simple PIM-SM Configuration

Cisco.com



On every router: `ip multicast-routing`

On every interface: `ip pim sparse-dense-mode`

On routers B and C: `ip pim send-rp-announce loopback0 scope 16`
`ip pim send-rp-discovery loopback0 scope 16`

Module5.ppt

©1998 – 2001, Cisco Systems, Inc. All rights reserved.

8/10/2001 2:37 PM

137

• Example PIM-SM Configuration

- The above example network shows how to configure a network to run Sparse mode using Auto-RP using two Candidate-RP's/Mapping Agents.

Note: A common practice is to combine the function of Candidate-RP and Mapping Agent on the same router. This is done more as a configuration convenience than for any operational requirement.

- One every router in the network:
 - Configure the 'ip multicast-routing' global command to enable multicast on the router.
 - Configure the 'ip pim sparse-dense-mode' interface command on *EVERY* interface on each router. This allows the Auto-RP groups to function in Dense mode and all other groups to operate in Sparse or Dense mode depending on whether an RP has been configured for the group.
 - On the router(s) that are to function as Candidate-RP's, configure the 'ip pim send-rp-announce Loopback0 scope <ttl>' command. (Make sure the <ttl> value is sufficient to allow the message to reach all Mapping Agents in the network.)
 - On the router(s) that are to function as Mapping Agents, configure the 'ip pim send-rp-discovery Loopback0 scope <ttl>' command. (Make sure the <ttl> value is sufficient to allow the message to reach all routers in the network.)
- No additional configuration is generally necessary. The network is now completely enabled for PIM-SM IP Multicast!

