



IP Multicast: Advanced PIM Concepts

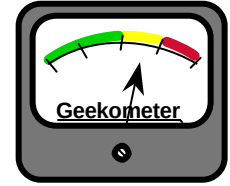




Beau Williamson
bwilliam@cisco.com



Agenda



- **PIM-DM Overview**
- **PIM-DM Advanced Concepts**
- **PIM-SM Overview**
- **PIM-SM Advanced Concepts**



PIM-DM Overview

- **Flood and Prune behavior**
- **Source trees created on demand based on RPF rule**
- **If the source goes inactive, the tree is torn down**
- **Branches that don't want data are pruned**



PIM-DM Overview

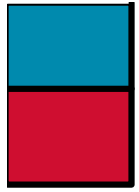
- **Grafts used to join existing source tree**
- **Uses Asserts to determine forwarder for multi-access LAN**
- **Prunes on non-RPF P2P links**
- **Rate-limited prunes on RPF P2P links**
- **PIM v1 packets ride inside IGMP packets**
- **PIM v2 is assigned new protocol 103**
 - v2 has some packet format extensions/changes



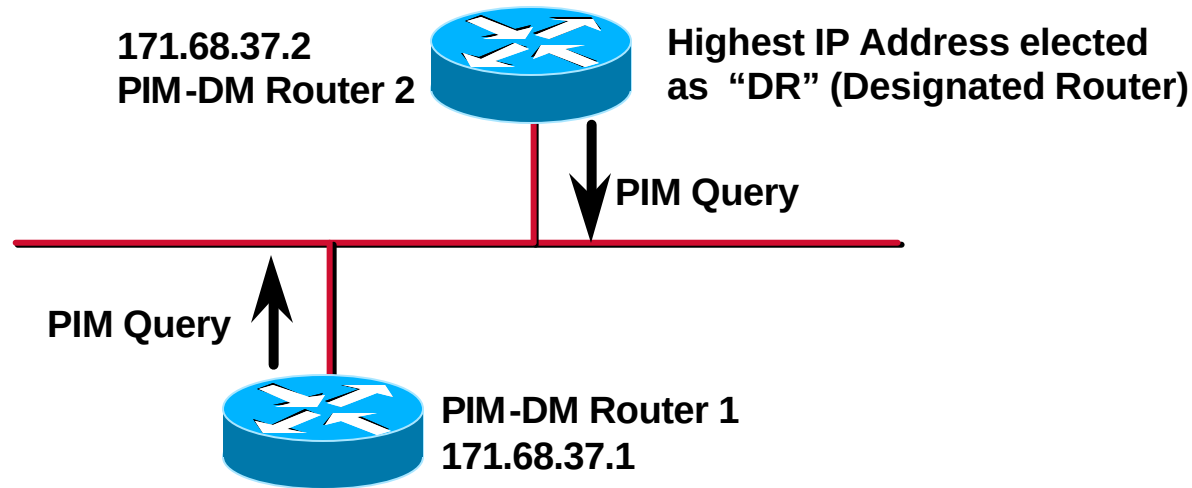


PIM-DM Advanced Concepts

- **PIM Neighbor Discovery**
- **PIM DM State**
- **PIM DM Forwarding**
- **PIM DM Pruning**
- **PIM DM Grafting**
- **PIM Assert Mechanism**
- **PIM DM State Maintenance**



PIM Neighbor Discovery



- PIM Queries are Multicast to the “All-Routers” (224.0.0.2) (with a TTL of 1) multicast group address periodically. (Default = 30 seconds)
- If the “DR” times-out, a new “DR” is elected.
- In PIM DM, interface is added to outgoing interface list for all groups when first neighbor is heard.





PIM Neighbor Discovery

```
wan-gw8>show ip pim neighbor
```

PIM Neighbor Table

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





PIM-DM Advanced Concepts

- PIM Neighbor Discovery
- **PIM DM State**
- PIM DM Forwarding
- PIM DM Pruning
- PIM DM Grafting
- PIM Assert Mechanism
- PIM DM State Maintenance



PIM-DM (*,G) State Rules

- **(*,G) created automatically**
 - When 1st (S,G) for group is created
 - (S,G)'s always have parent (*,G)
- **(*,G) reflect PIM neighbor adjacency**
 - IIF = NULL
 - OIL = all interfaces
 - with PIM-DM neighbors or
 - with directly connected hosts or
 - manually configured



PIM-DM (S,G) State Rules

- **(S,G) created by multicast data arrival**
 - Parent (*,G) created (if doesn't exist)
 - IIF = RPF Interface in direction of source
 - OIL = Copy of OIL from (*,G) minus IIF
- **Interfaces in OIL initially “Forward”**
 - Go to “Pruned” state when Prune rcvd
 - “Forward” intfct timers never expire
 - “Pruned” intfct timers expire in 3 minutes



PIM-DM OIL Rules

- **(*,G) OIL**
 - Reflects intfc's w/PIM neighbors or
 - Locally connected members or
 - Manually configured interfaces
- **(S,G) OIL**
 - Copy of (*,G) OIL less IIF
- **Interfaces in (S,G) OIL “pruned”**
 - When appropriate Prune received
 - Prune Expiration counter (3 min) started
 - Interface marked “Prune/Dense” (not removed)
 - Returns to “Forward/Dense” when Prune expires





PIM-DM State Flags

- **D = Dense 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
- **J = Join SPT**
 - Always on in (*,G) entry in PIM-DM
 - Basically meaningless in PIM-DM



PIM-DM State Example

```
(*, 224.1.1.1), 00:00:10/00:00:00, RP 0.0.0.0, flags: D
Incoming interface: Null, RPF nbr 0.0.0.0
Outgoing interface list:
  Serial0, Forward/Dense, 00:00:10/00:00:00
  Serial1, Forward/Dense, 00:00:10/00:00:00
  Serial3, Forward/Dense, 00:00:10/00:00:00

(128.9.160.43/32, 224.1.1.1), 00:00:10/00:02:49, flags: T
Incoming interface: Serial0, RPF nbr 198.92.1.129
Outgoing interface list:
  Serial1, Forward/Dense, 00:00:10/00:00:00
  Serial3, Prune/Dense, 00:00:05/00:02:55
```





PIM-DM Advanced Concepts

- PIM Neighbor Discovery
- PIM DM State
- **PIM DM Forwarding**
- PIM DM Pruning
- PIM DM Grafting
- PIM Assert Mechanism
- PIM DM State Maintenance

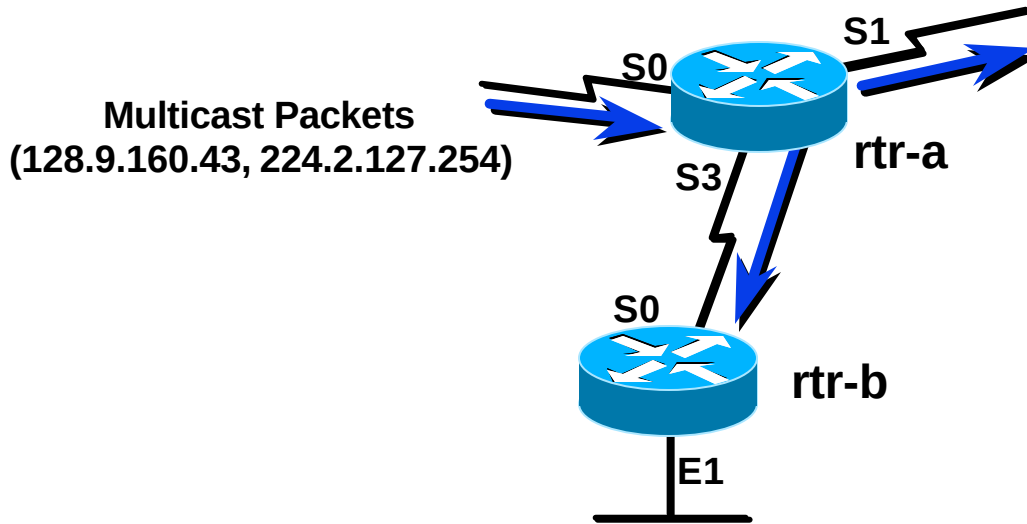


PIM-DM Forwarding Rules

- **Use longest match entry**
 - Use (S, G) entry if exists
 - Otherwise, use (*, G) entry
 - Effectively, only (S,G)'s used in PIM-DM
- **RPF check first**
 - If Packet didn't arrive via IIF, drop it.
- **Forward Packet (if RPF succeeded)**
 - Send out all “unpruned” interfaces in OIL



PIM DM Forwarding



- PIM Dense mode interfaces are placed on the “*oilist*” for a Multicast Group IF:
 - PIM neighbor heard on interface
 - Host on this interface has joined the group
 - Interface has been manually configured to join group.
- Packets are “flooded” out all interfaces in “*oilist*”.



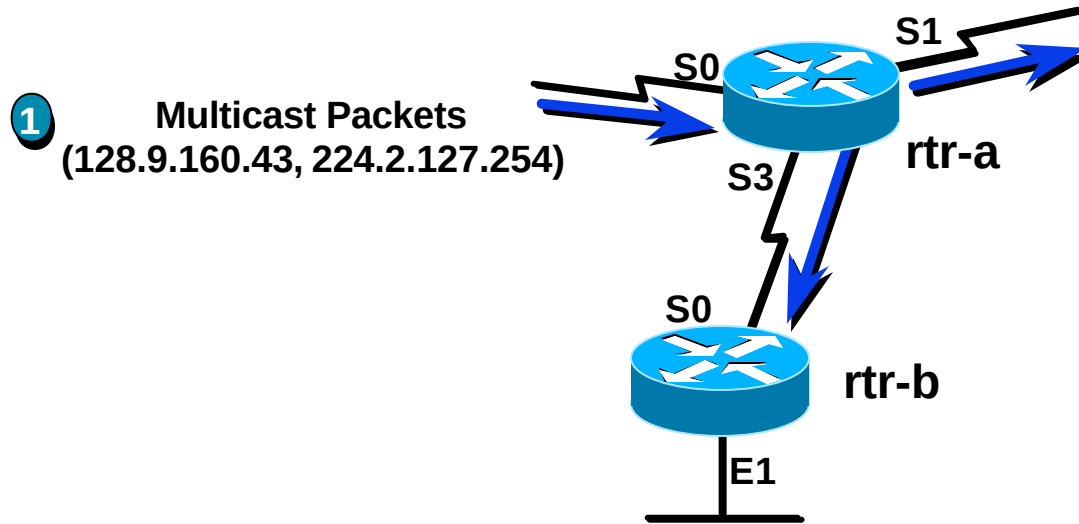


PIM-DM Advanced Concepts

- PIM Neighbor Discovery
- PIM DM State
- PIM DM Forwarding
- **PIM DM Pruning**
- PIM DM Grafting
- PIM Assert Mechanism
- PIM DM State Maintenance



PIM DM Pruning

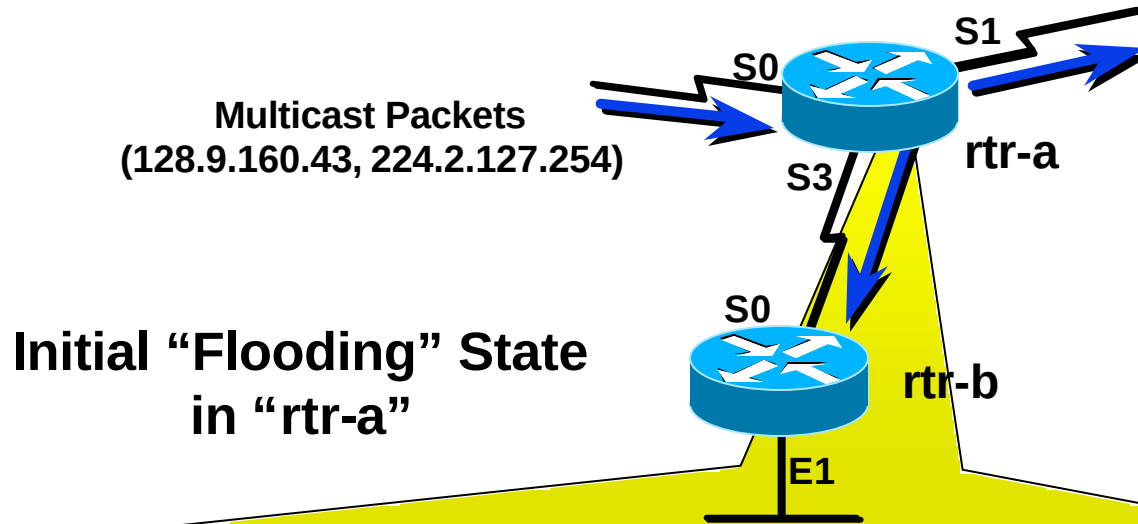


- 1 “rtr-a” initially floods (S, G) traffic out all interfaces in “*oilist*”.





PIM DM Pruning



Initial “Flooding” State
in “rtr-a”

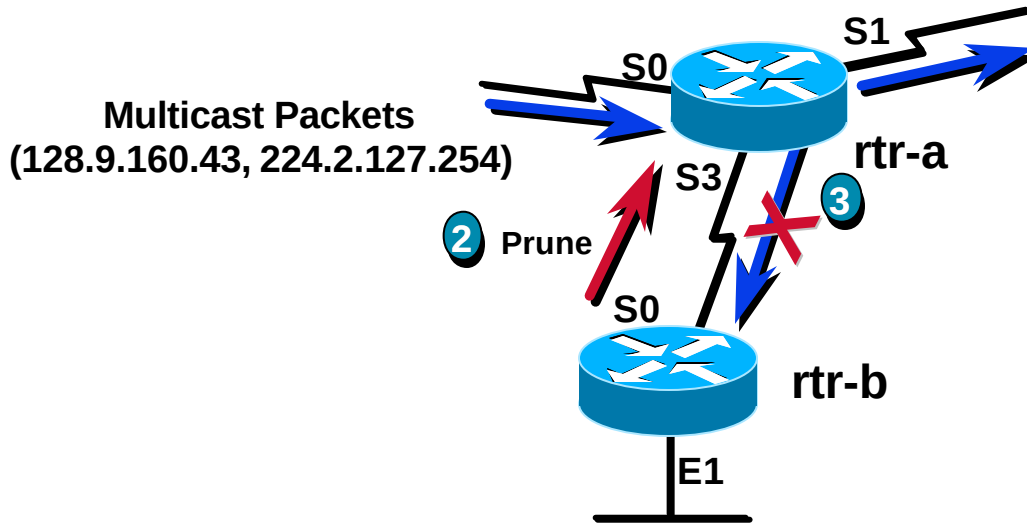
```
(*, 224.2.127.254), 00:00:10/00:00:00, RP 0.0.0.0, flags: D
Incoming interface: Null, RPF nbr 0.0.0.0
Outgoing interface list:
  Serial0, Forward/Dense, 00:00:10/00:00:00
  Serial1, Forward/Dense, 00:00:10/00:00:00
  Serial3, Forward/Dense, 00:00:10/00:00:00

(128.9.160.43/32, 224.2.127.254), 00:00:10/00:02:49, flags: T
Incoming interface: Serial0, RPF nbr 198.92.1.129
Outgoing interface list:
  Serial1, Forward/Dense, 00:00:10/00:00:00
  Serial3, Forward/Dense, 00:00:10/00:00:00
```





PIM DM Pruning

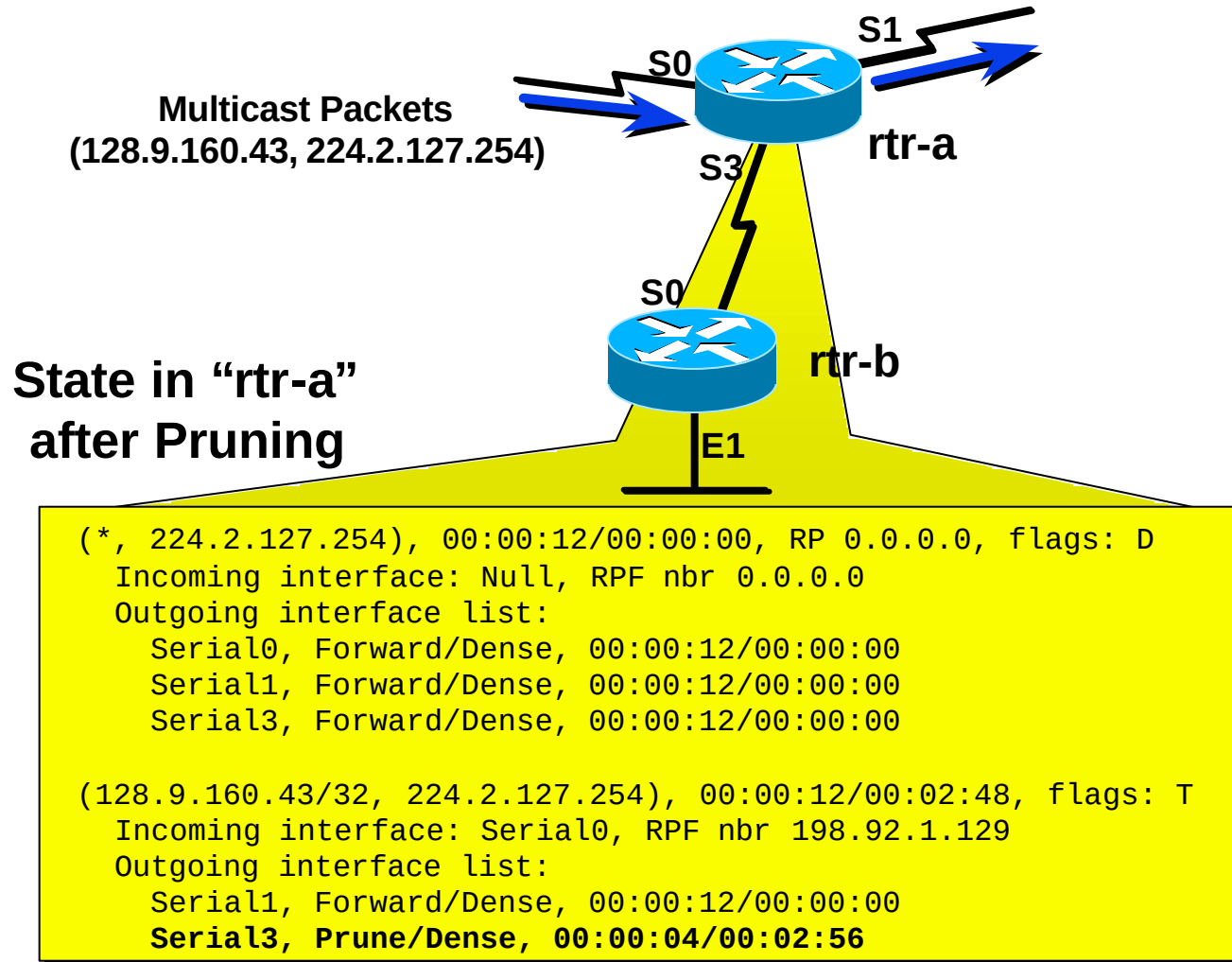


- 1 "rtr-a" initially floods (S, G) traffic out all interfaces in "olist".
- 2 "rtr-b" is a leaf node w/o receivers. Sends Prune for (S,G).
- 3 "rtr-a" Prunes interface for (S,G).



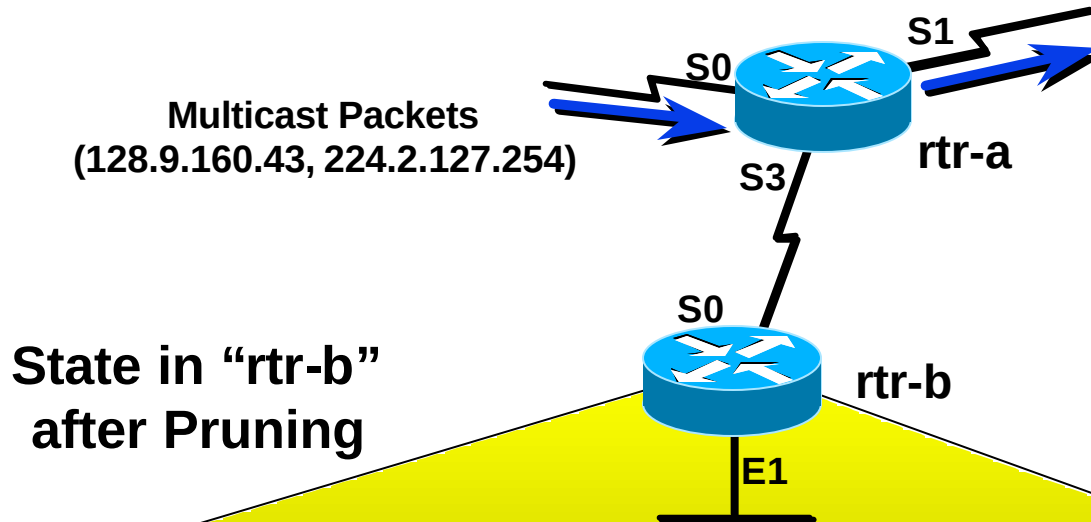


PIM DM Pruning





PIM DM Pruning

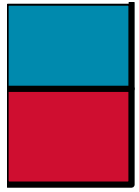


State in “rtr-b”
after Pruning

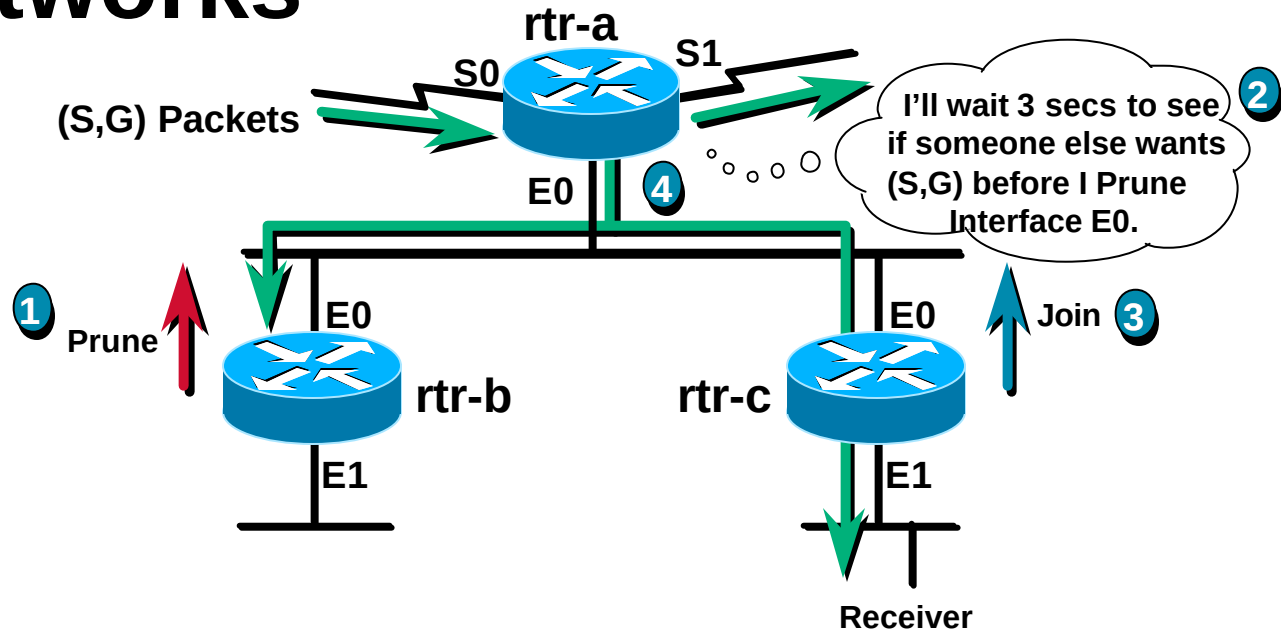
```
(*, 224.2.127.254), 00:00:12/00:00:00, RP 0.0.0.0, flags: D
Incoming interface: Null, RPF nbr 0.0.0.0
Outgoing interface list:
  Serial0, Forward/Dense, 00:00:12/00:00:00
  Ethernet0, Forward/Dense, 00:00:12/00:00:00
```

```
(128.9.160.43/32, 224.2.127.254), 00:00:12/00:02:48, flags: PT
Incoming interface: Serial0, RPF nbr 198.92.1.129
Outgoing interface list: Null
```



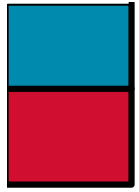


PIM Prune Delay on Multiaccess Networks



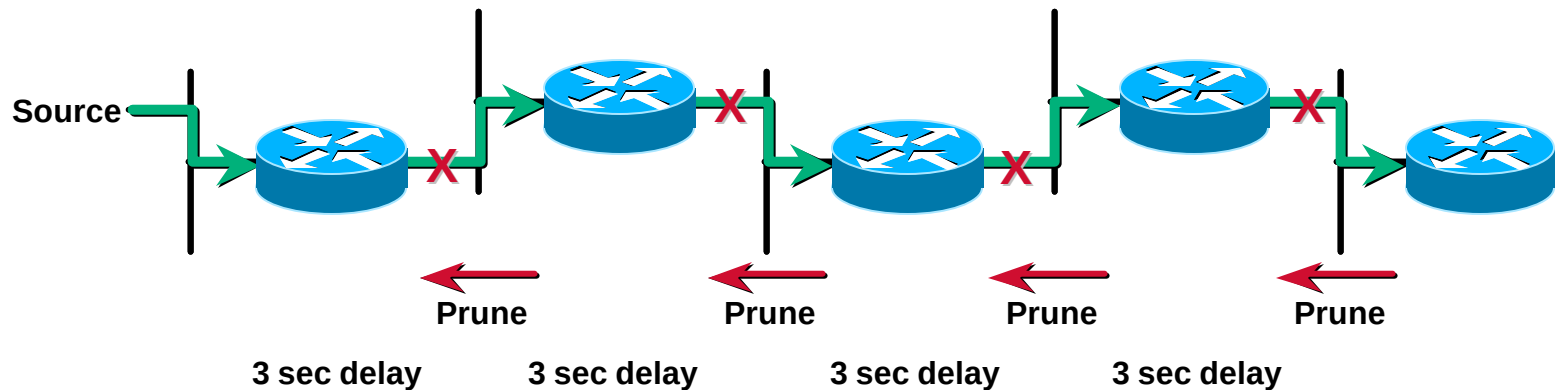
- 1 "rtr-b" is a leaf node w/o receivers. Sends Prune for (S,G).
- 2 "rtr-a" schedules a Prune for (S,G) to occur in 3 seconds.
- 3 "rtr-c" hears Prune from "rtr-b". Overrides with a Join.
- 4 "rtr-a" hears Join and cancels Prune for (S,G).





PIM Prune Delay on Multiaccess Networks

Watch out for the ripple affect of this Delay!!!



- Source begins sending traffic which is flooded everywhere.
- Leaf router has no receivers; sends prune which ripples up the tree.
- Total time to prune back to source = 12 seconds!
- Process repeats 3 minutes later when prunes timeout!



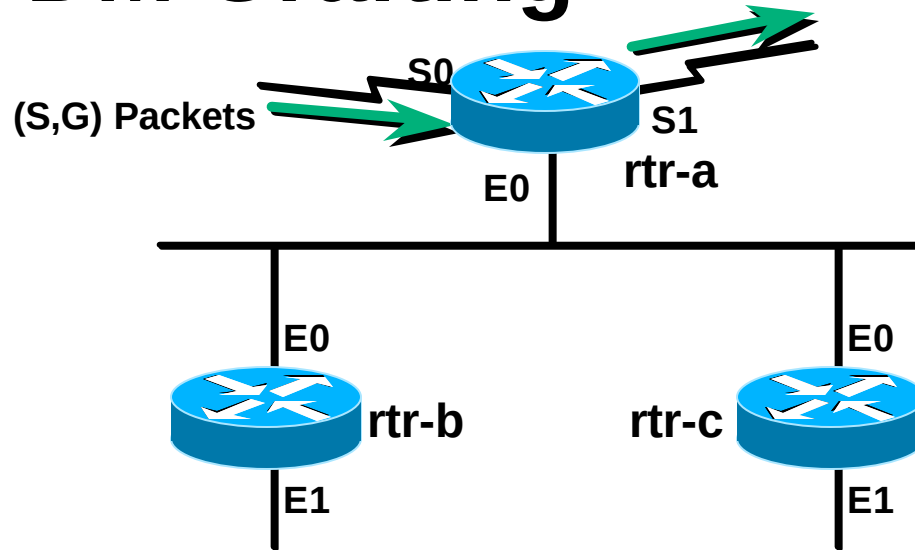


PIM-DM Advanced Concepts

- PIM Neighbor Discovery
- PIM DM State
- PIM DM Forwarding
- PIM DM Pruning
- **PIM DM Grafting**
- PIM Assert Mechanism
- PIM DM State Maintenance



PIM DM Grafting



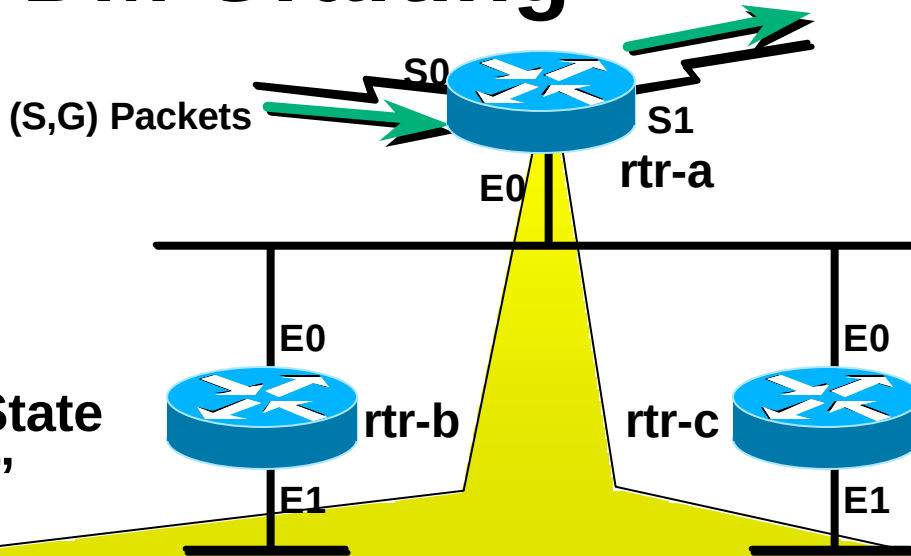
Beginning State

- “rtr-b” and “rtr-c” have previously Pruned (S,G) traffic.
- “rtr-a” is still forwarding traffic downstream via S1.





PIM DM Grafting



Beginning State
in "rtr-a"

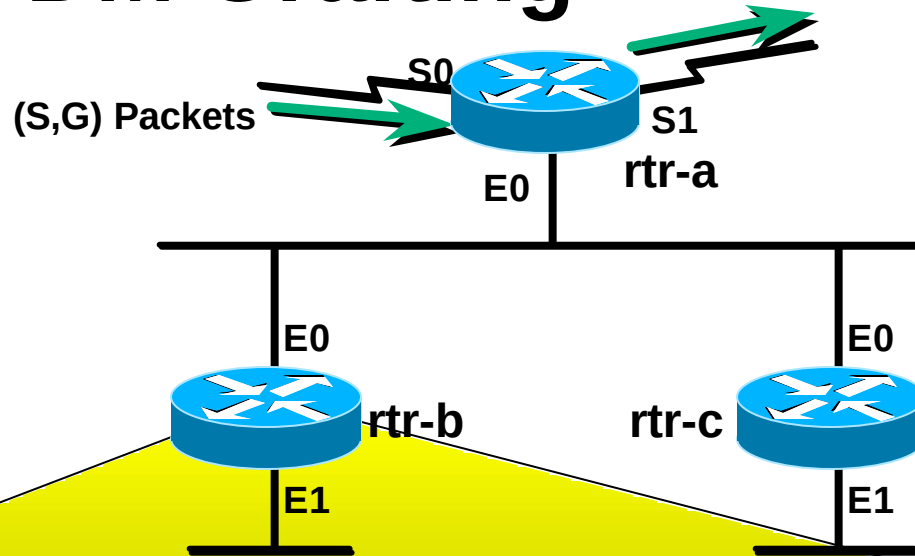
```
(*, 224.2.127.254), 00:04:10/00:00:00, RP 0.0.0.0, flags: D
Incoming interface: Null, RPF nbr 0.0.0.0
Outgoing interface list:
  Serial0, Forward/Dense, 00:04:10/00:00:00
  Serial1, Forward/Dense, 00:04:10/00:00:00
  Ethernet0, Forward/Dense, 00:04:10/00:00:00

(128.9.160.43/32, 224.2.127.254), 00:04:10/00:02:39, flags: T
Incoming interface: Serial0, RPF nbr 198.92.1.129
Outgoing interface list:
  Ethernet0, Prune/Dense, 00:01:29/00:01:30
  Serial1, Forward/Dense, 00:04:10/00:00:00
```





PIM DM Grafting



```
(*, 224.2.127.254), 00:04:10/00:00:00, RP 0.0.0.0, flags: D
Incoming interface: Null, RPF nbr 0.0.0.0
Outgoing interface list:
  Ethernet0, Forward/Dense, 00:04:10/00:00:00

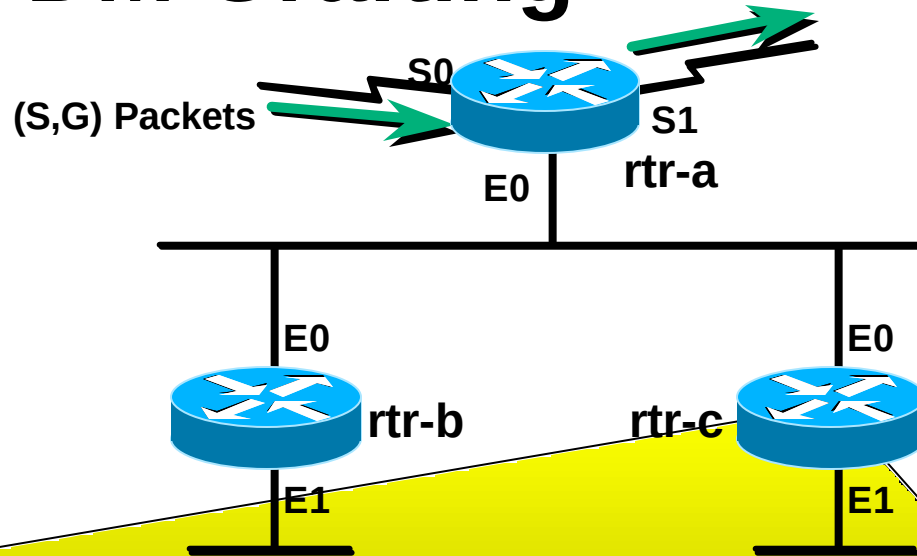
(128.9.160.43/32, 224.2.127.254), 00:04:10/00:02:39, flags: PT
Incoming interface: Ethernet0, RPF nbr 198.92.2.1
Outgoing interface list: Null
```

Beginning State in “rtr-b”





PIM DM Grafting

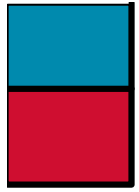


```
(*, 224.2.127.254), 00:04:10/00:00:00, RP 0.0.0.0, flags: D
Incoming interface: Null, RPF nbr 0.0.0.0
Outgoing interface list:
  Ethernet0, Forward/Dense, 00:04:10/00:00:00
```

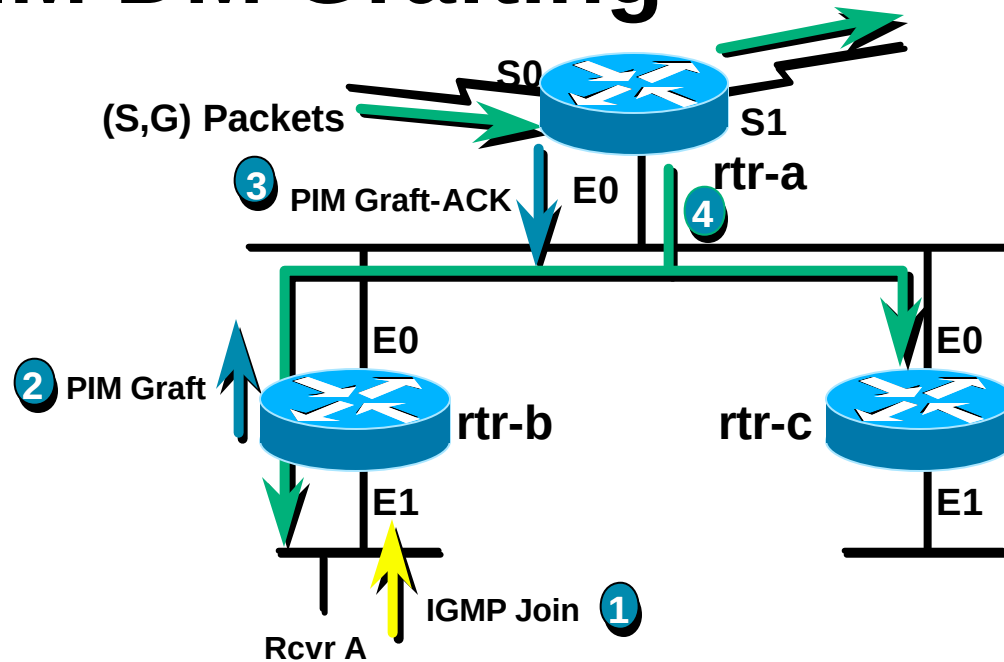
```
(128.9.160.43/32, 224.2.127.254), 00:04:10/00:02:39, flags: PT
Incoming interface: Ethernet0, RPF nbr 198.92.2.1
Outgoing interface list: Null
```

Beginning State in “rtr-c”





PIM DM Grafting

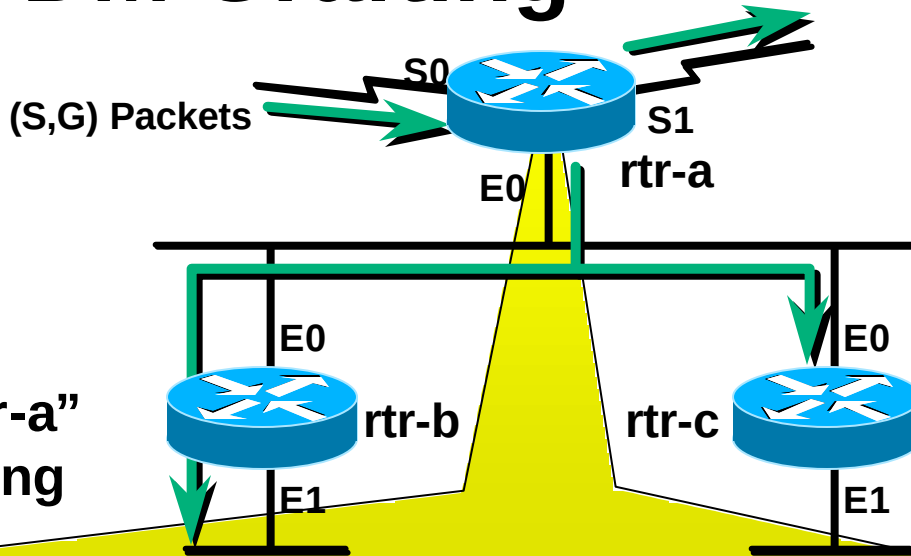


- 1 "Rcvr A" wishes to receive group G traffic. Sends IGMP Join for G.
- 2 "rtr-b" sends PIM Graft for Group (S,G).
- 3 "rtr-a" acknowledges with a PIM Graft-Ack.
- 4 "rtr-a" begins forwarding traffic for (S,G).





PIM DM Grafting



State in “rtr-a”
after Grafting

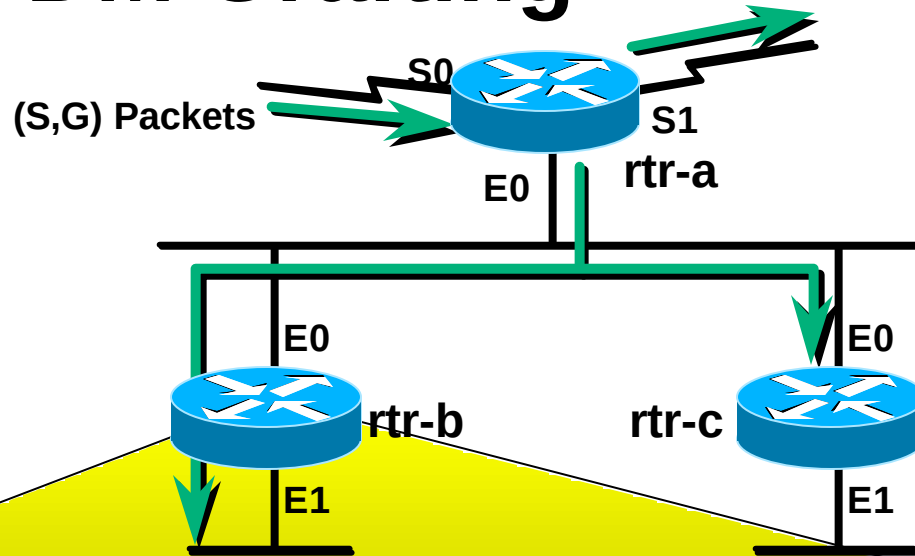
```
(*, 224.2.127.254), 00:04:10/00:00:00, RP 0.0.0.0, flags: D
Incoming interface: Null, RPF nbr 0.0.0.0
Outgoing interface list:
  Serial0, Forward/Dense, 00:04:10/00:00:00
  Serial1, Forward/Dense, 00:04:10/00:00:00
  Ethernet0, Forward/Dense, 00:04:10/00:00:00

(128.9.160.43/32, 224.2.127.254), 00:04:10/00:02:39, flags: T
Incoming interface: Serial0, RPF nbr 198.92.1.129
Outgoing interface list:
  Ethernet0, Forward/Dense, 00:00:25/00:00:00
  Serial1, Forward/Dense, 00:04:10/00:00:00
```





PIM DM Grafting

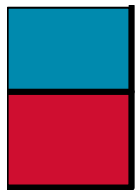


```
(*, 224.2.127.254), 00:04:10/00:00:00, RP 0.0.0.0, flags: D
Incoming interface: Null, RPF nbr 0.0.0.0
Outgoing interface list:
  Ethernet0, Forward/Dense, 00:04:10/00:00:00

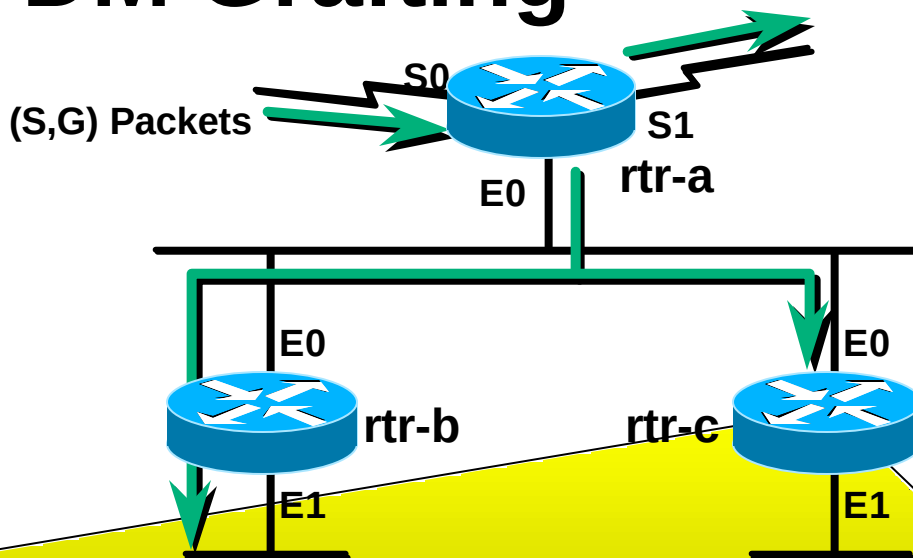
(128.9.160.43/32, 224.2.127.254), 00:04:10/00:02:39, flags: CT
Incoming interface: Ethernet0, RPF nbr 198.92.2.1
Outgoing interface list:
  Ethernet1, Forward/Dense, 00:00:26/00:00:00
```

State in "rtr-b" after Grafting





PIM DM Grafting



```
(*, 224.2.127.254), 00:04:10/00:00:00, RP 0.0.0.0, flags: D
Incoming interface: Null, RPF nbr 0.0.0.0
Outgoing interface list:
  Ethernet0, Forward/Dense, 00:04:10/00:00:00
```

```
(128.9.160.43/32, 224.2.127.254), 00:04:10/00:02:39, flags: PT
Incoming interface: Ethernet0, RPF nbr 198.92.2.1
Outgoing interface list: Null
```

State in “rtr-c” after Grafting



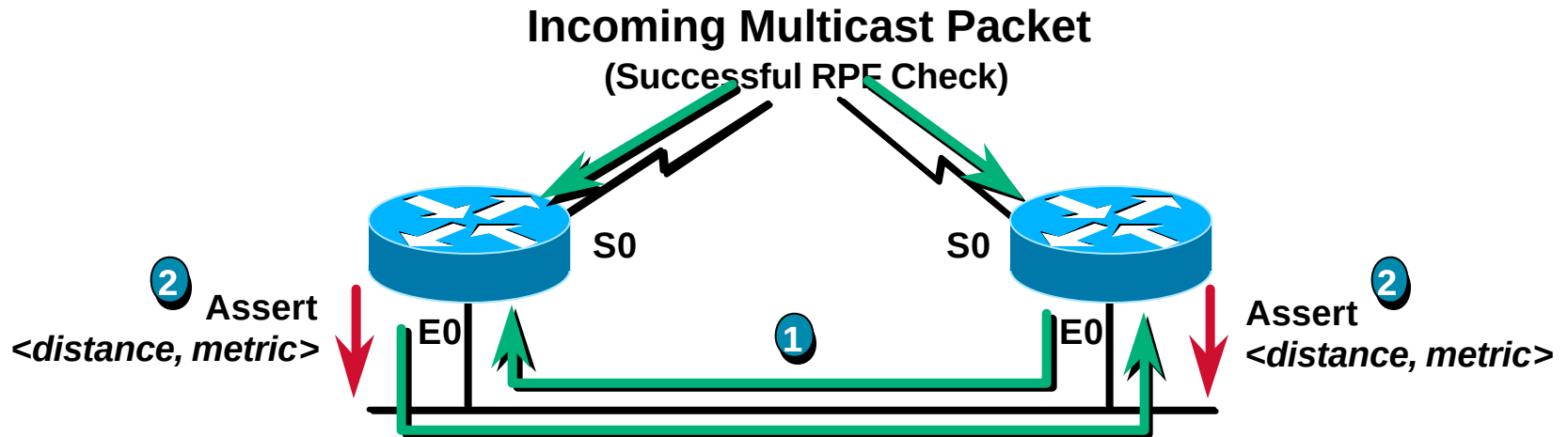


PIM-DM Advanced Concepts

- PIM Neighbor Discovery
- PIM DM State
- PIM DM Forwarding
- PIM DM Pruning
- PIM DM Grafting
- **PIM Assert Mechanism**
- PIM DM State Maintenance



PIM Assert Mechanism



- 1 Routers **receive** packet on an interface in their “*oiflist*”!!
 - Only one router should continue sending to avoid duplicate packets.
- 2 Routers send “PIM Assert” messages
 - Compare *distance* and *metric* values
 - Router with best route to source wins
 - If *metric* & *distance* equal, highest IP adr wins
 - Losing router stops sending (prunes interface)





PIM-DM Advanced Concepts

- PIM Neighbor Discovery
- PIM DM State
- PIM DM Forwarding
- PIM DM Pruning
- PIM DM Grafting
- PIM Assert Mechanism
- **PIM DM State Maintenance**



PIM DM State Maintenance

- **State is maintained by the “flood and prune” behavior of Dense mode.**
 - Received Multicast packets reset (S,G) entry “expiration” timers.
 - When (S, G) entry “expiration” timers count down to zero, the entry is deleted.
- **Interface prune state times out every 3 minutes causing periodic reflooding and pruning.**



PIM-SM Overview

- **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





PIM-SM Overview

- Only one RP is chosen for a particular group
- RP statically configured or dynamically learned (Auto-RP, PIM v2 candidate RP advertisements)
- Data forwarded based on the source state (S, G) if it exists, otherwise use the shared state (*, G)
- RFC 2117 - “PIM Sparse Mode Protocol Spec”





PIM-SM Advanced Concepts

- **PIM SM State**
- **PIM SM Forwarding**
- **PIM SM Joining**
- **PIM SM Registering**
- **PIM SM SPT-Switchover**
- **PIM SM Pruning**
- **PIM SM State Maintenance**





PIM-SM (*,G) State Rules

- **(*,G) creation**
 - Upon receipt of a (*,G) Join or
 - 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 hosts or
 - manually configured
- **(*,G) deletion**
 - When OIL = NULL and
 - no child (S,G) state exists



PIM-SM (S,G) State Rules

- **(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





PIM-SM OIL Rules

- **Interfaces in OIL added**
 - **By receipt of Join message**
 - Intfc's added to (*,G) are added to all (S,G)'s
- **Interfaces in OIL removed**
 - **By receipt of Prune message**
 - Intfc's 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





PIM-SM State Flags

- **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



PIM-SM State Flags (cont.)

- **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)





PIM-SM State Flags (cont.)

- **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





PIM-SM State Example

```
(* , 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

(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
```





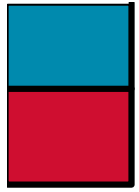
PIM-SM Advanced Concepts

- PIM SM State
- **PIM SM Forwarding**
- PIM SM Joining
- PIM SM Registering
- PIM SM SPT-Switchover
- PIM SM Pruning
- PIM SM State Maintenance

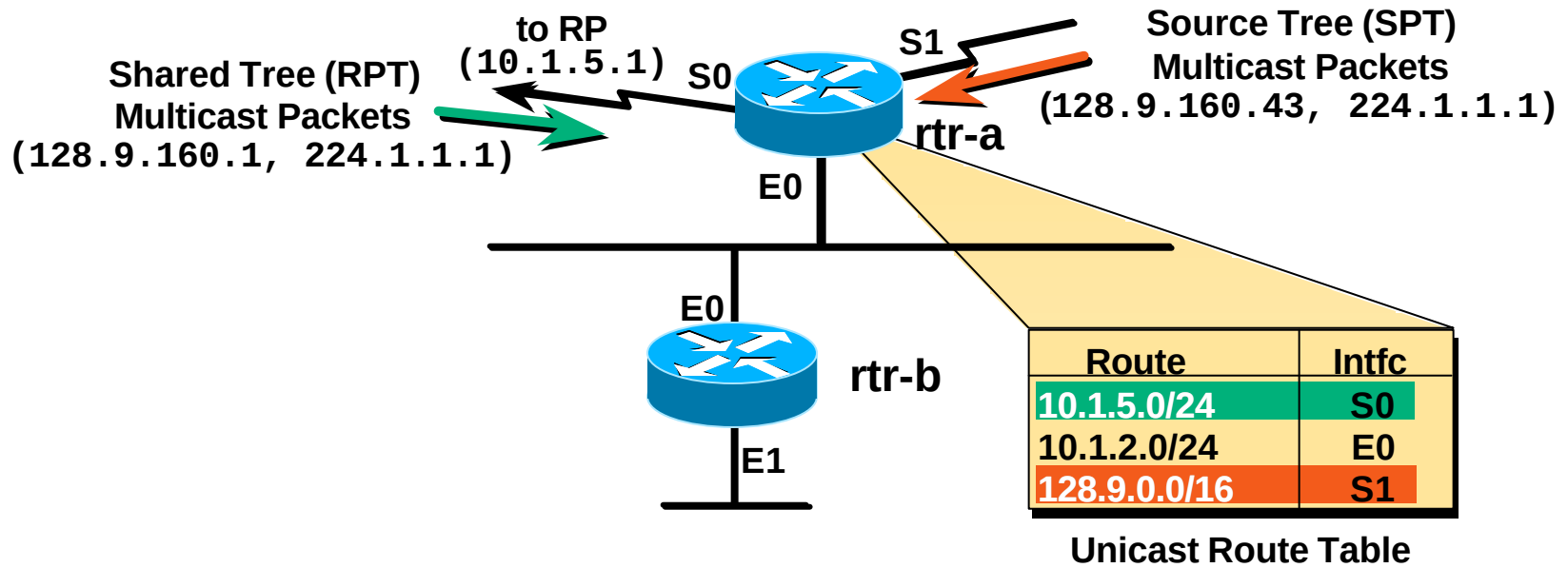


PIM-SM Forwarding Rules

- **Use longest match entry**
 - Use (S, G) entry if exists
 - Otherwise, use (*, G) entry
- **RPF check first**
 - If Packet didn't arrive via IIF, drop it.
- **Forward Packet (if RPF succeeded)**
 - Send out all “unpruned” interfaces in OIL

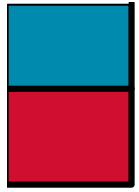


PIM SM Forwarding

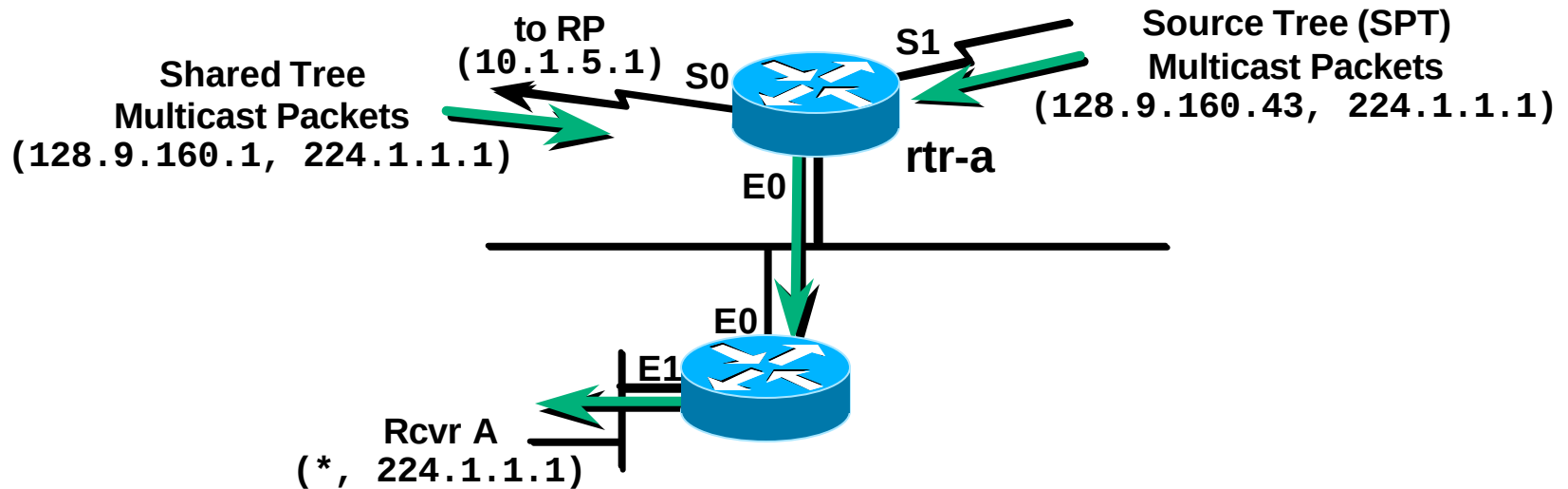


- Sparse mode RPF checks depends on tree type
 - For shared trees, uses RP address
 - For shortest-path trees, uses Source address





PIM SM Forwarding



- Packets are “forwarded” out all interfaces in “*olist*”.
- PIM Sparse mode interfaces are placed on the “*olist*” for a Multicast Group IF:
 - PIM neighbor Joins the group on this interface
 - Host on this interface has joined the group
 - Interface has been manually configured to join group.





PIM-SM Advanced Concepts

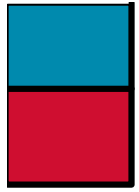
- PIM SM State
- PIM SM Forwarding
- **PIM SM Joining**
- PIM SM Registering
- PIM SM SPT-Switchover
- PIM SM Pruning
- PIM SM State Maintenance



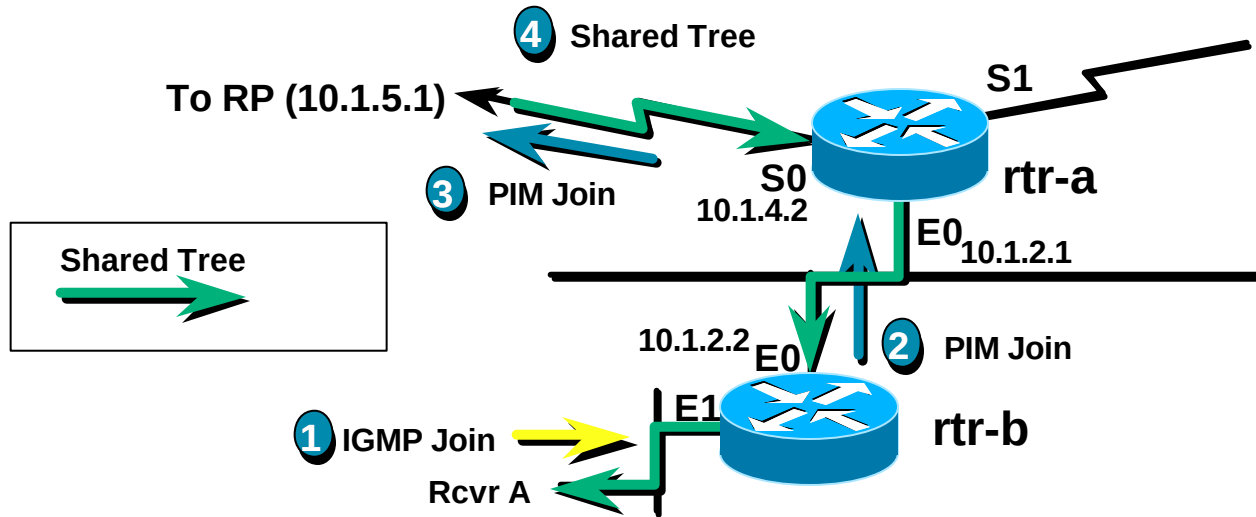
PIM SM Joining

- **Leaf routers send a (*,G) Join to toward RP**
 - Joins sent hop-by-hop along unicast path toward RP
 - Joins multicast to ALL-ROUTERS multicast address
(Except for NBMA-Mode Interfaces)
- **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 (*,G) tree.





PIM SM Joining (cont.)

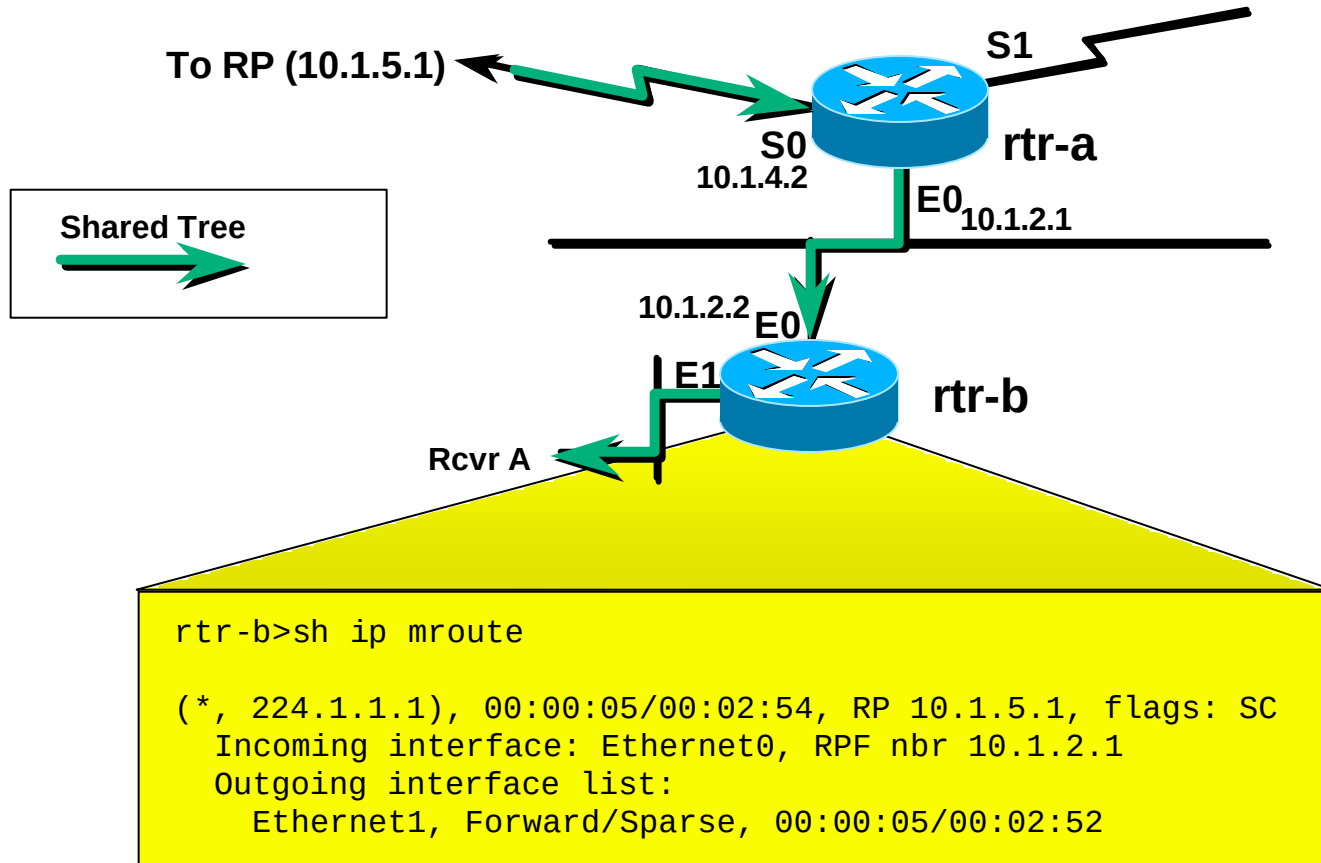


- 1** "Rcvr A" wishes to receive group G traffic. Sends IGMP Join for G.
- 2** "rtr-b" creates (*,G) state; sends (*,G) PIM Join towards RP.
- 3** "rtr-a" creates (*,G) state; sends (*,G) PIM Join towards RP.
- 4** Shared tree is built all the way back to the RP.





PIM SM Joining (cont.)

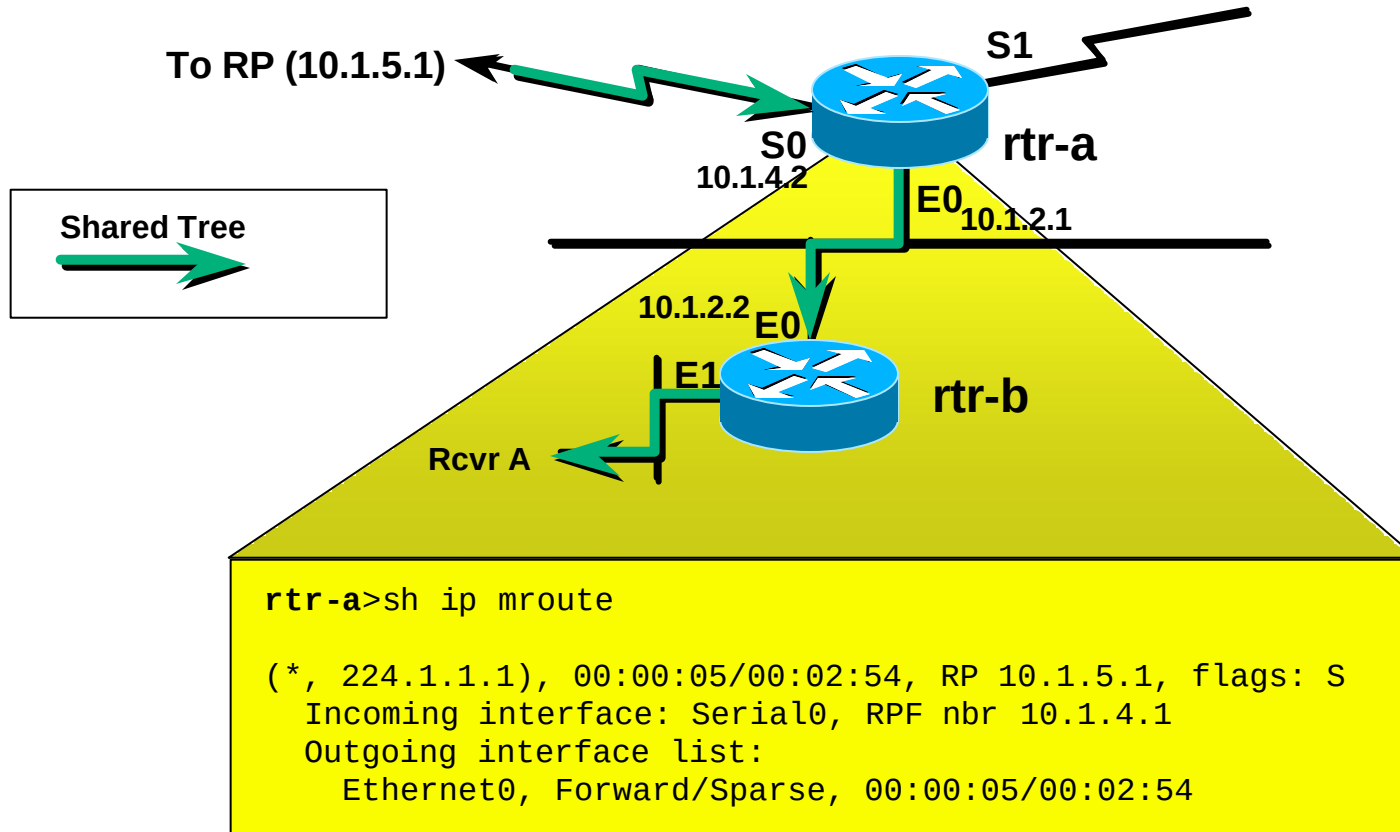


State in “rtr-b” after Joining (*, 224.1.1.1)





PIM SM Joining (cont.)



State in "rtr-a" after Joining (*, 224.1.1.1)





PIM-SM Advanced Concepts

- PIM SM State
- PIM SM Forwarding
- PIM SM Joining
- **PIM SM Registering**
- PIM SM SPT-Switchover
- PIM SM Pruning
- PIM SM State Maintenance



PIM SM Registering

- **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 the Mcast packet inside Register msg
 - Forwards Mcast packet down Shared Tree
 - Sends (S,G) Join toward Source / 1st-Hop router to build an (S,G) SPT between Source and RP





PIM SM Registering (cont.)

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



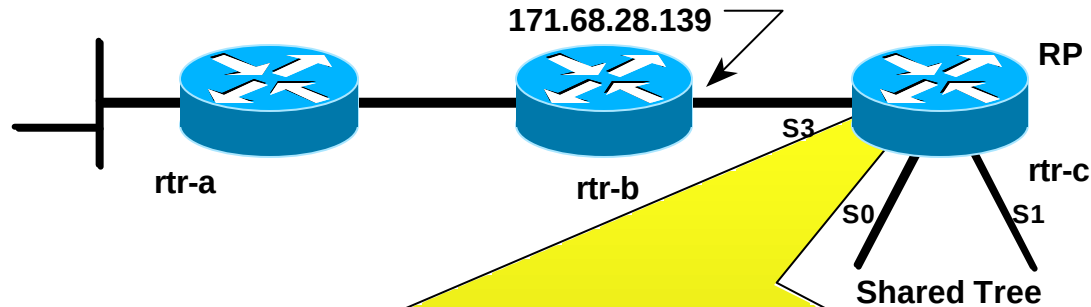


PIM SM Register Examples

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



PIM SM Registering — Receivers Join Group First



```
rtr-c>sh ip mroute 224.1.1.1
```

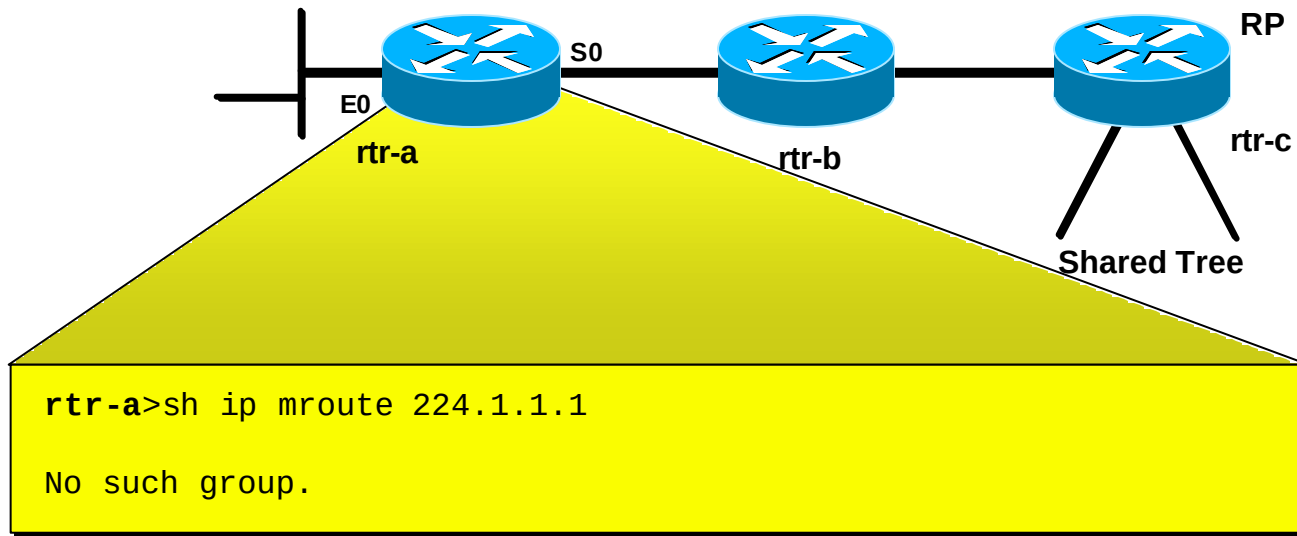
```
(*, 224.1.1.1), 00:00:03/00:02:56, RP 171.68.28.140, flags:S  
Incoming interface: Null, RPF nbr 0.0.0.0,  
Outgoing interface list:  
  Serial0, Forward/Sparse, 00:03:14/00:02:59  
  Serial1, Forward/Sparse, 00:03:14/00:02:59
```

State in “RP” before Registering
(with receivers on Shared Tree)





PIM SM Registering — Receivers Join Group First

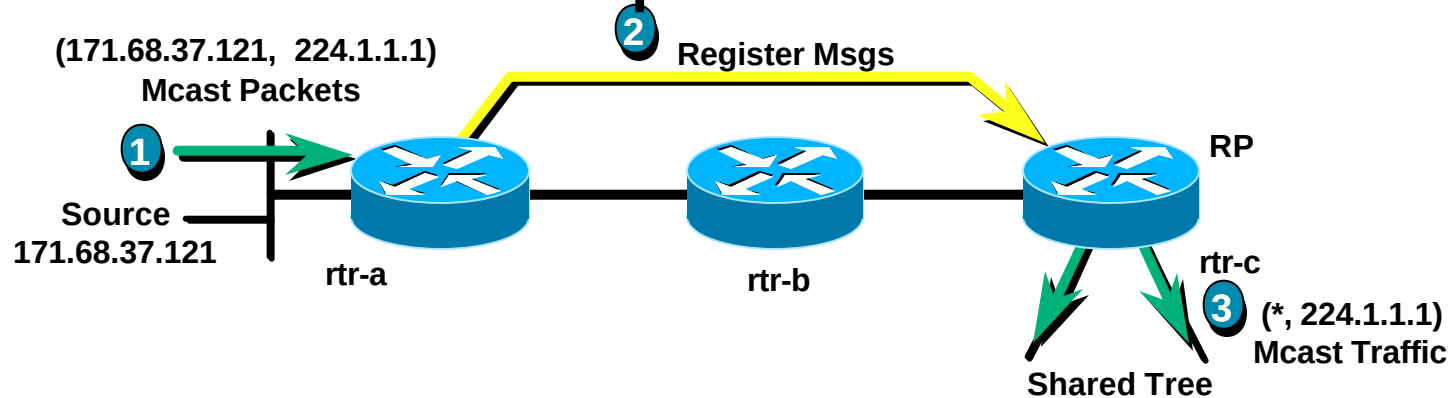


State in “rtr-a” before Registering
(with receivers on Shared Tree)





PIM SM Registering — Receivers Join Group First

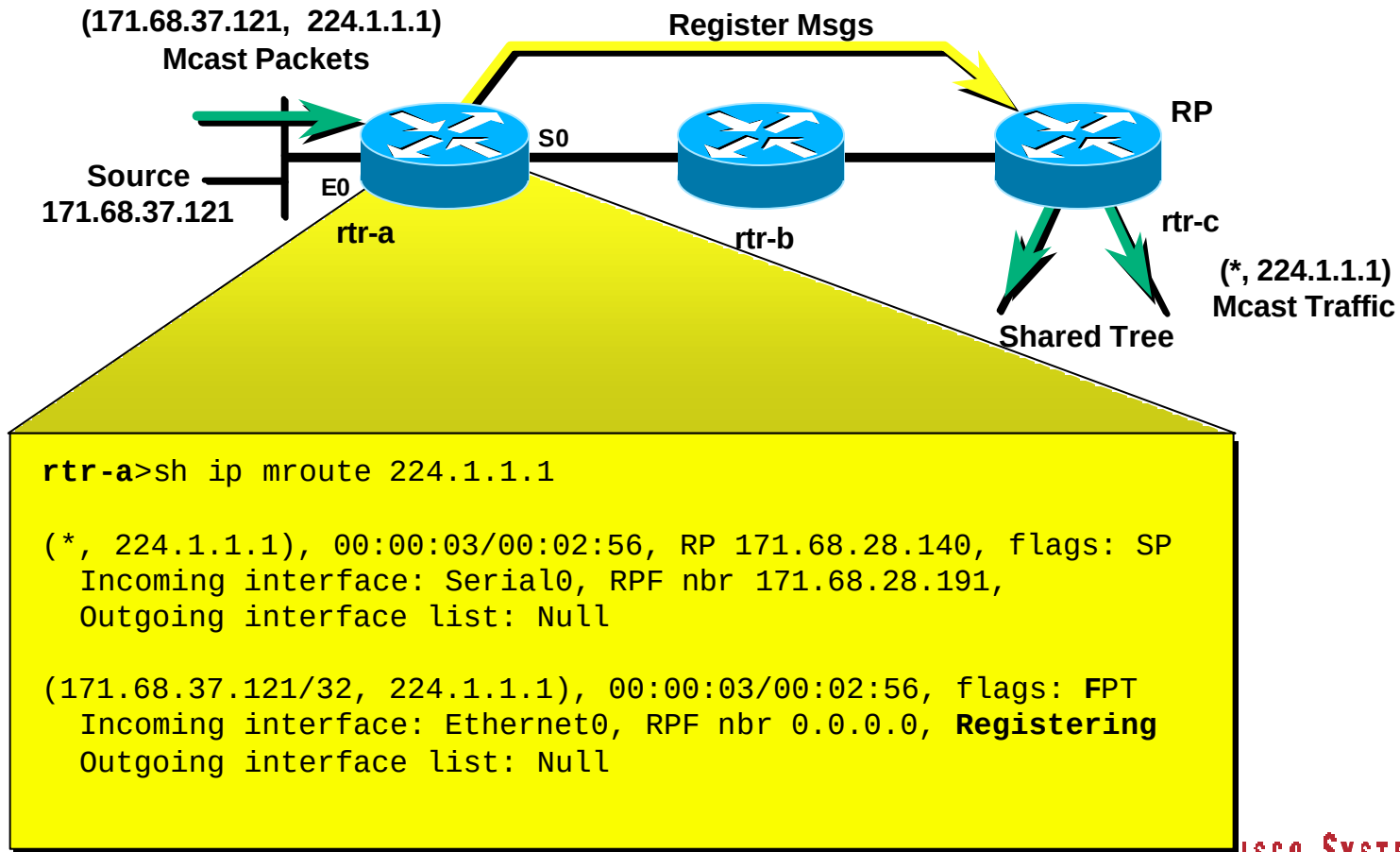


- 1 “Source” begins sending group G traffic.
- 2 “rtr-a” encapsulates packets in Registers; unicasts to RP.
- 3 “rtr-c” (RP) de-encapsulates packets; forwards down Shared tree.





PIM SM Registering — Receivers Join Group First

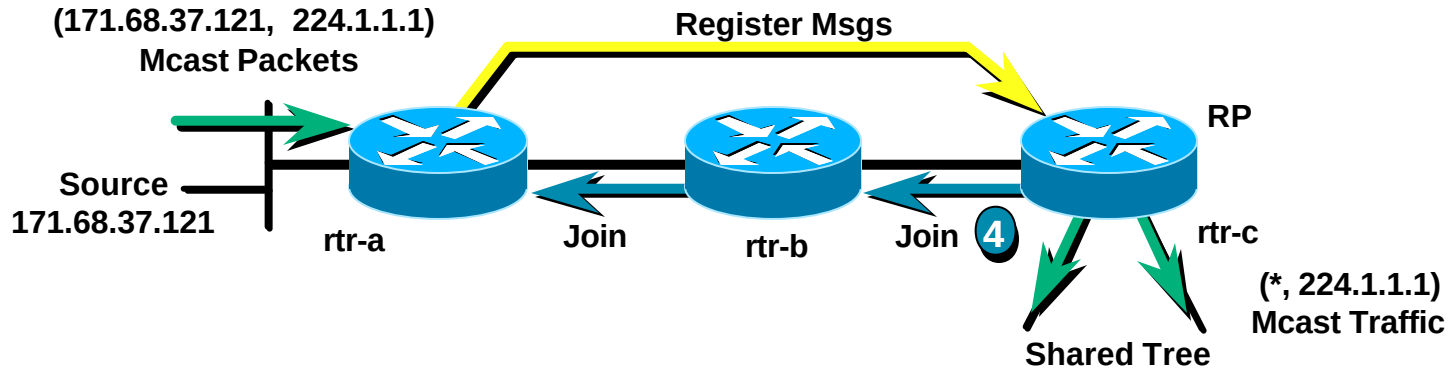


State in “rtr-a” while Registering





PIM SM Registering — Receivers Join Group First

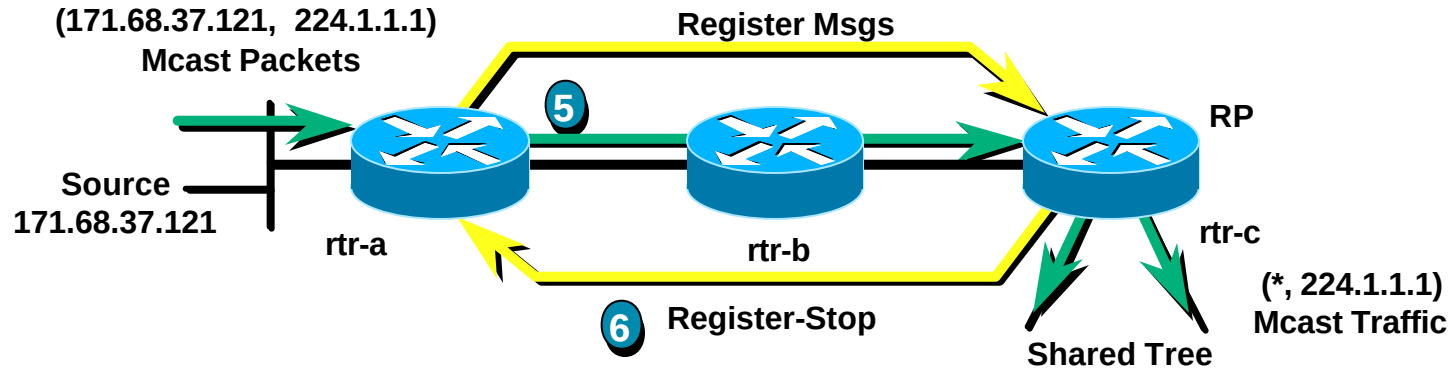


- 1 “Source” begins sending group G traffic.
- 2 “rtr-a” encapsulates packets in Registers; unicasts to RP.
- 3 RP (“rtr-c”) de-encapsulates packets; forwards down Shared tree.
- 4 RP sends (S,G) Join toward Source; builds SPT.





PIM SM Registering — Receivers Join Group First

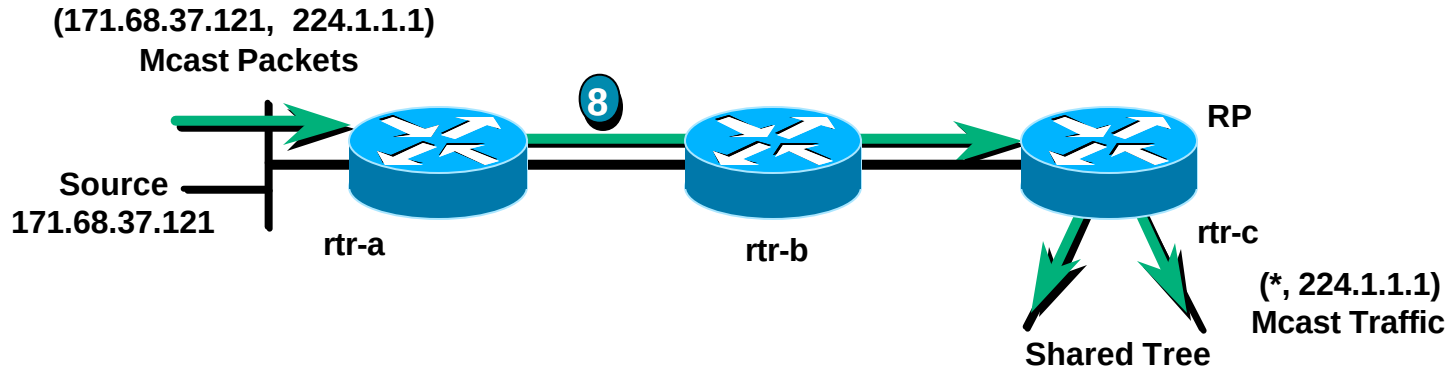


- 5** RP begins receiving (S,G) traffic down SPT.
- 6** RP sends “Register-Stop” to “rtr-a”.





PIM SM Registering — Receivers Join Group First

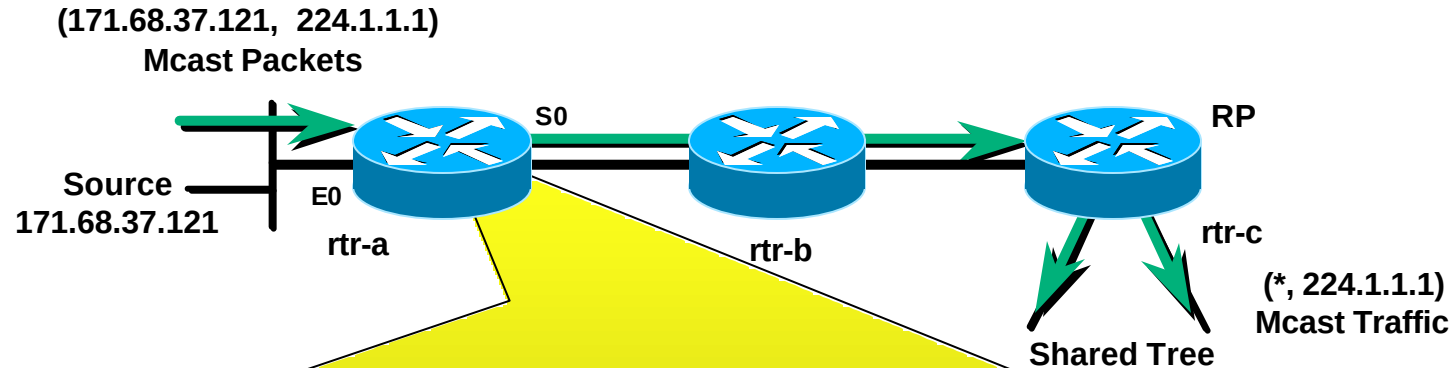


- 5 RP begins receiving (S,G) traffic down SPT.
- 6 RP sends "Register-Stop" to "rtr-a".
- 7 "rtr-a" stops encapsulating traffic in Register Messages.
- 8 (S,G) Traffic now flowing down a single path (SPT) to RP.





PIM SM Registering — Receivers Join Group First



```
rtr-a>sh ip mroute 224.1.1.1
```

```
(*, 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
```

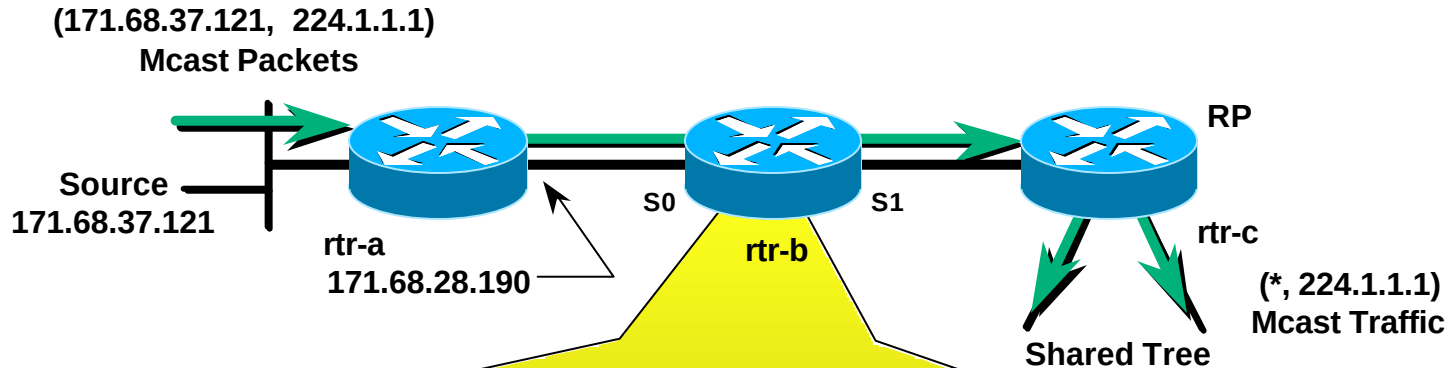
```
(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags: FT  
Incoming interface: Ethernet0, RPF nbr 0.0.0.0  
Outgoing interface list:  
Serial0, Forward/Sparse, 00:04:28/00:01:32
```

State in “rtr-a” after Registering





PIM SM Registering — Receivers Join Group First



```
rtr-b>sh ip mroute 224.1.1.1
```

```
(*, 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
```

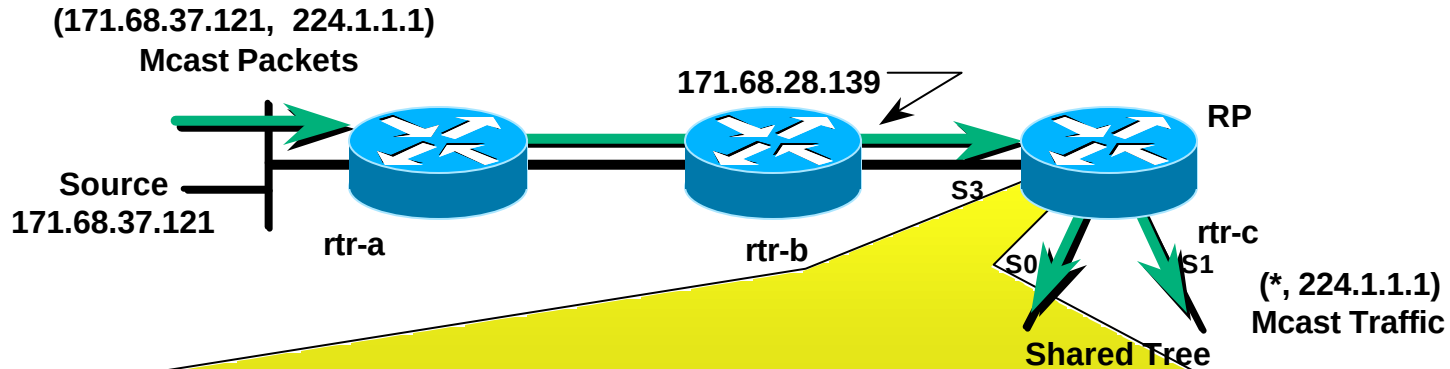
```
(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags: T  
Incoming interface: Serial0, RPF nbr 171.68.28.190  
Outgoing interface list:  
Serial1, Forward/Sparse, 00:04:28/00:01:32
```

**State in “rtr-b” after “rtr-a” Registers
(with receivers on Shared Tree)**





PIM SM Registering — Receiver Joins Group First



```
rtr-c>sh ip mroute 224.1.1.1
```

```
(*, 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
```

**State in “RP” after “rtr-a” Registers
(with receivers on Shared Tree)**



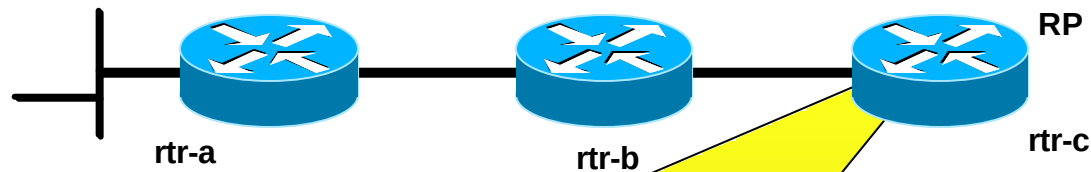


PIM SM Register Examples

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



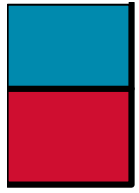
PIM SM Registering — Source Registers First



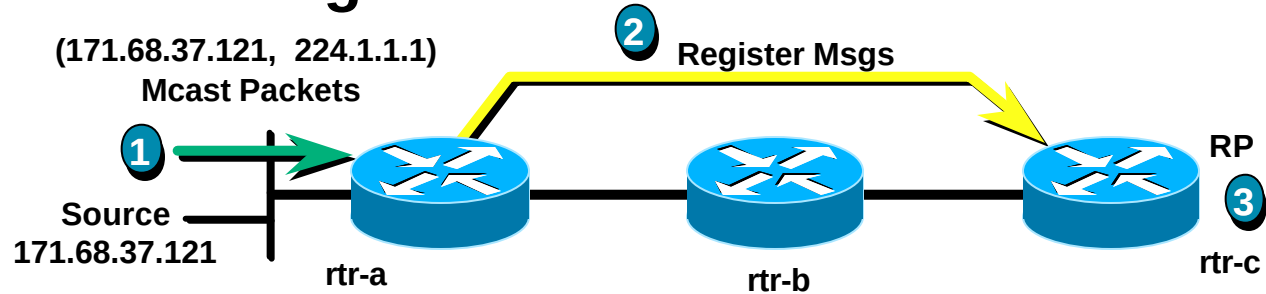
```
rtr-c>sh ip mroute 224.1.1.1  
Group 224.1.1.1 not found.
```

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





PIM SM Registering — Source Registers First

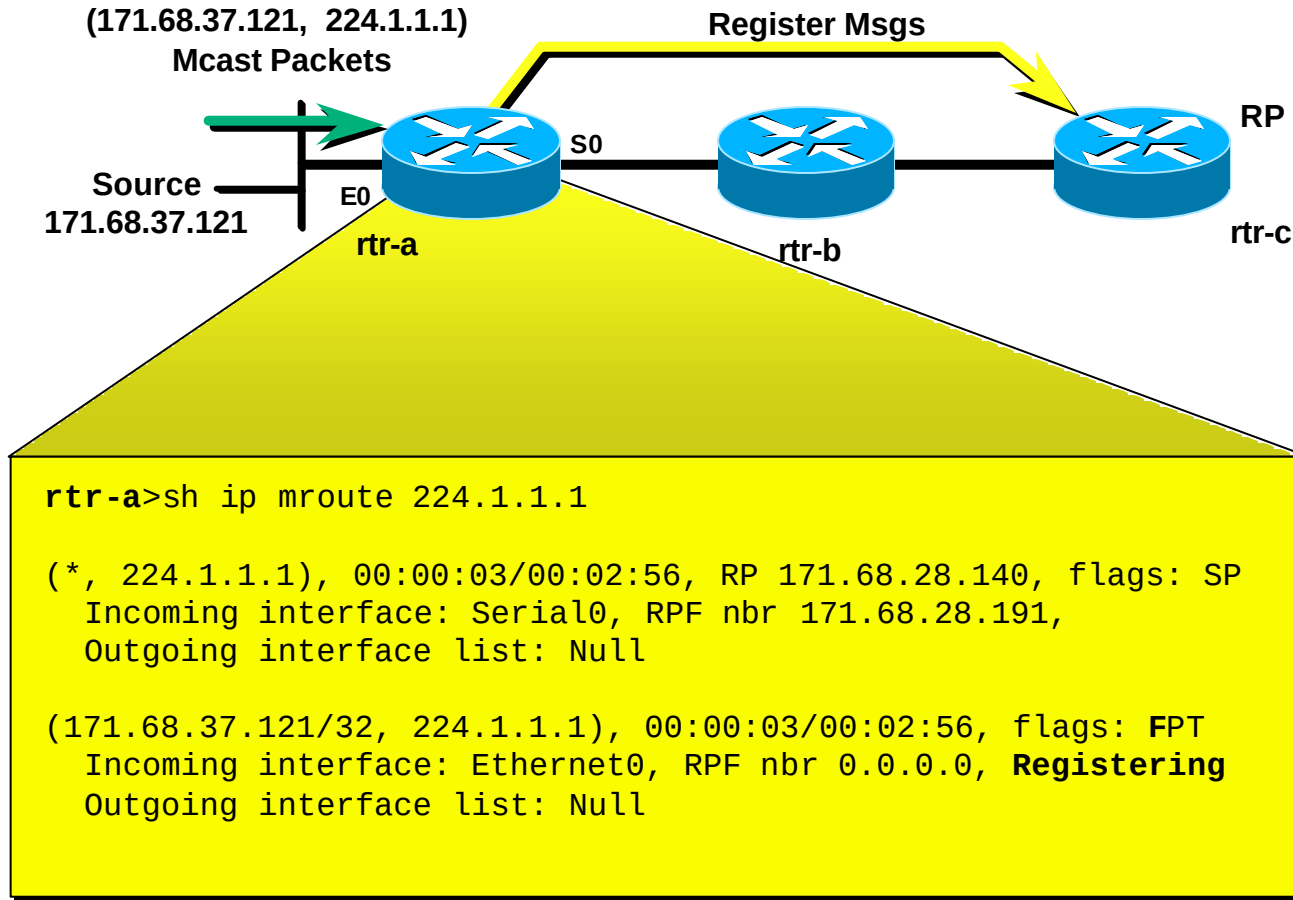


- 1 “Source” begins sending group G traffic.
- 2 “rtr-a” encapsulates packets in Registers; unicasts to RP.
- 3 “rtr-c” (RP) has no receivers on Shared Tree; discards packet.





PIM SM Registering — Source Registers First

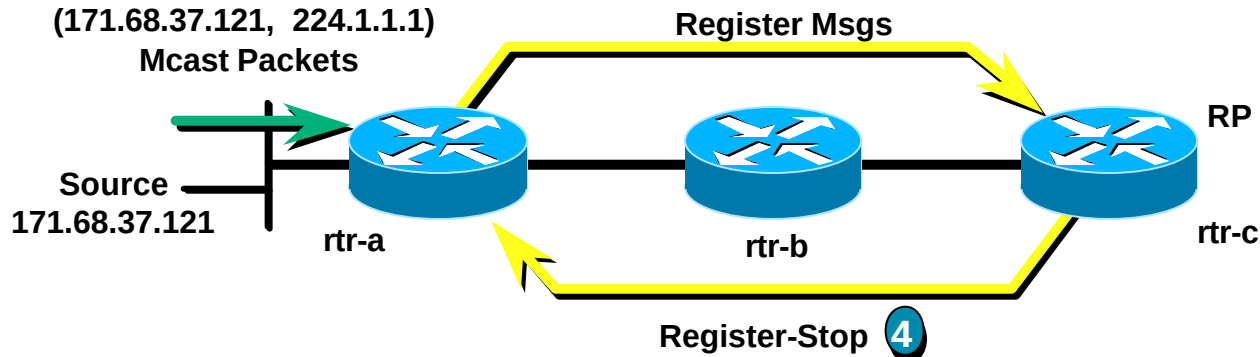


State in “rtr-a” while Registering





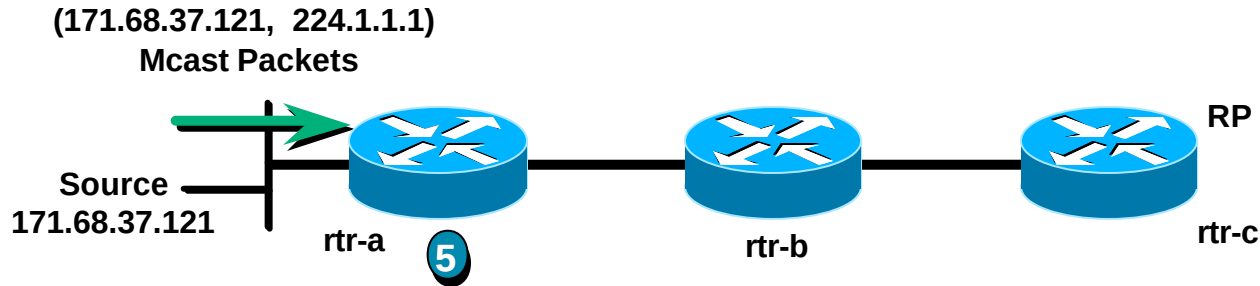
PIM SM Registering — Source Registers First



- 1 “Source” begins sending group G traffic.
- 2 “rtr-a” encapsulates packets in Registers; unicasts to RP.
- 3 RP (“rtr-c”) has no receivers on Shared Tree; discards packet.
- 4 RP sends “Register-Stop” to “rtr-a”.



PIM SM Registering — Source Registers First

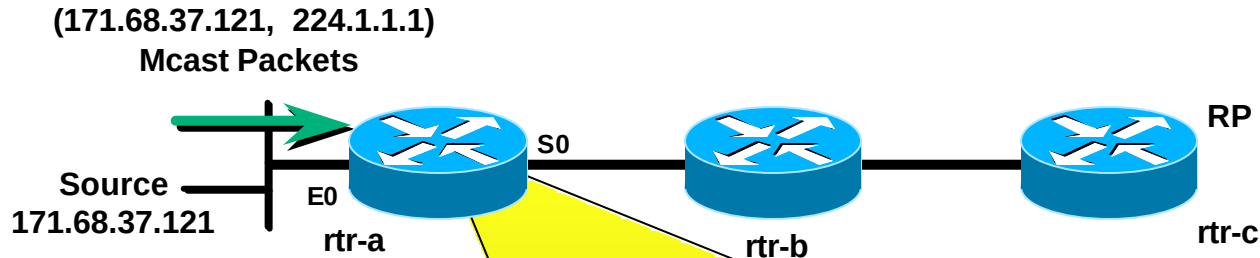


- 1 “Source” begins sending group G traffic.
- 2 “rtr-a” encapsulates packets in Registers; unicasts to RP.
- 3 RP (“rtr-c”) has no receivers on Shared Tree; discards packet.
- 4 RP sends “Register-Stop” to “rtr-a”.
- 5 “rtr-a” stops encapsulating traffic in Register Messages; drops packets from Source.





PIM SM Registering — Source Registers First



```
rtr-a>sh ip mroute 224.1.1.1
```

```
(*, 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
```

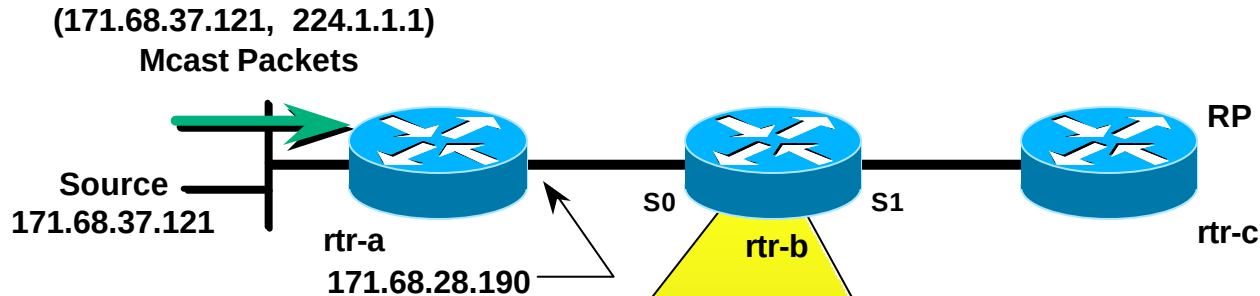
```
(171.68.37.121/32, 224.1.1.1), 00:01:28/00:01:32, flags: FPT  
Incoming interface: Ethernet0, RPF nbr 0.0.0.0  
Outgoing interface list: Null
```

**State in “rtr-a” after Registering
(without receivers on Shared Tree)**





PIM SM Registering — Source Registers First



```
rtr-b>sh ip mroute 224.1.1.1
```

```
(*, 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
```

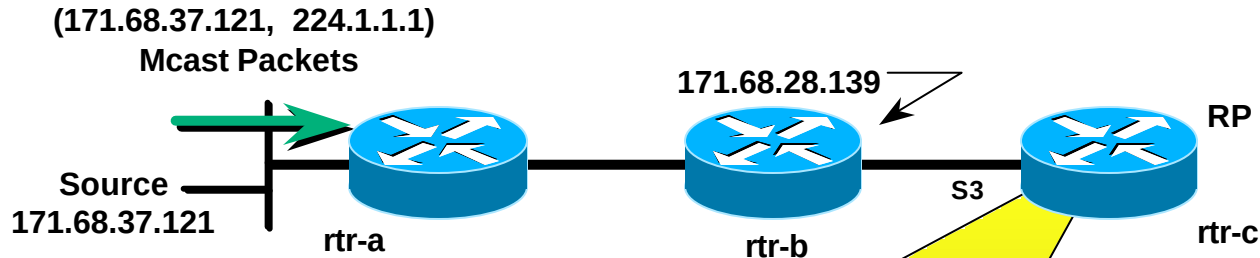
```
(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags: PT  
Incoming interface: Serial0, RPF nbr 171.68.28.190  
Outgoing interface list: Null
```

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





PIM SM Registering — Source Registers First



```
rtr-c>sh ip mroute 224.1.1.1
```

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

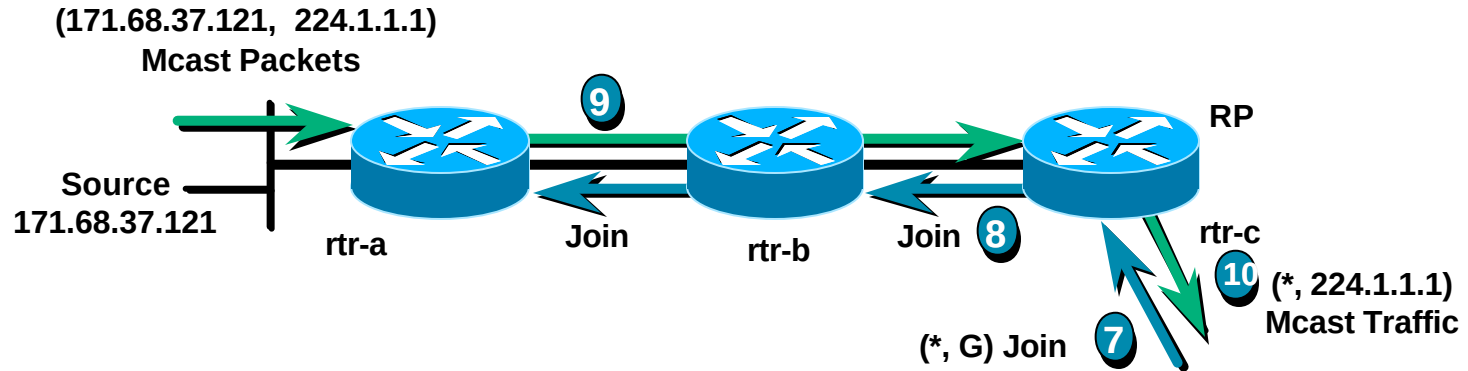
```
(171.68.37.121, 224.1.1.1, 00:01:15/00:01:45, flags: PT  
Incoming interface: Serial3, RPF nbr 171.68.28.139,  
Outgoing interface list: Null
```

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





PIM SM Registering — Source Registers First



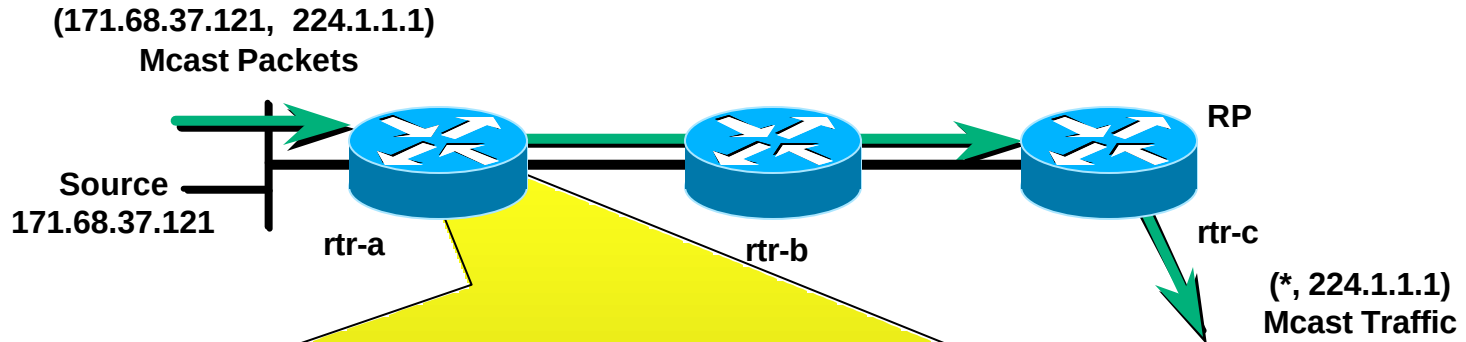
Receivers begin joining the Shared Tree

- 7 RP ("rtr-c") receives (*, G) Join from a receiver on Shared Tree.
- 8 RP sends (S,G) Joins for all known Sources in Group.
- 9 RP begins receiving (S,G) traffic down SPT.
- 10 RP forwards (S,G) traffic down Shared Tree to receivers.





PIM SM Registering — Source Registers First



```
rtr-a>sh ip mroute 224.1.1.1
```

```
(*, 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
```

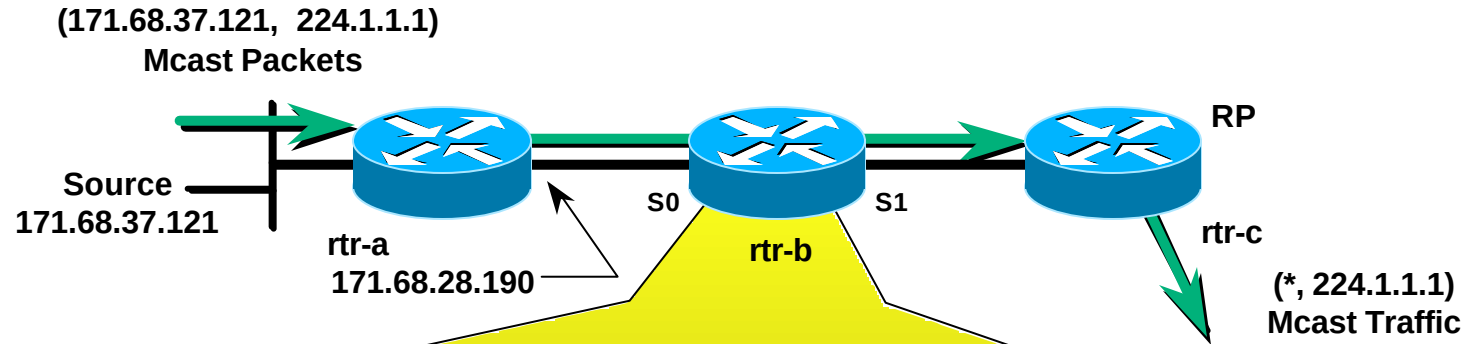
```
(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags: FT  
Incoming interface: Ethernet0, RPF nbr 0.0.0.0  
Outgoing interface list:  
Serial0, Forward/Sparse, 00:04:28/00:01:32
```

State in “rtr-a” after Receivers Join





PIM SM Registering — Source Registers First



```
rtr-b>sh ip mroute 224.1.1.1
```

```
(*, 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
```

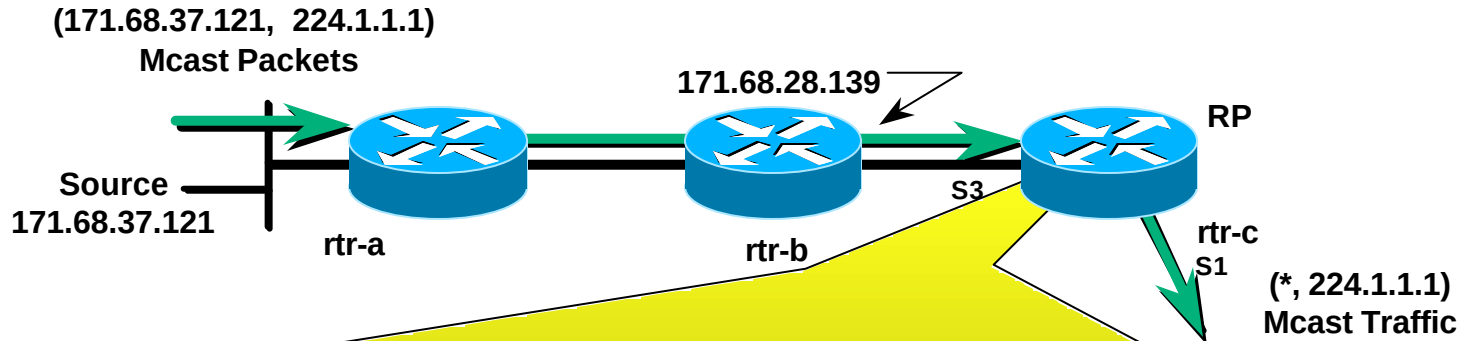
```
(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags: T  
Incoming interface: Serial0, RPF nbr 171.68.28.190  
Outgoing interface list:  
Serial1, Forward/Sparse, 00:04:28/00:01:32
```

State in “rtr-b” after Receivers Join





PIM SM Registering — Source Registers First



```
rtr-c>sh ip mroute 224.1.1.1
```

```
(*, 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, 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 Receivers Join





PIM SM Register Examples

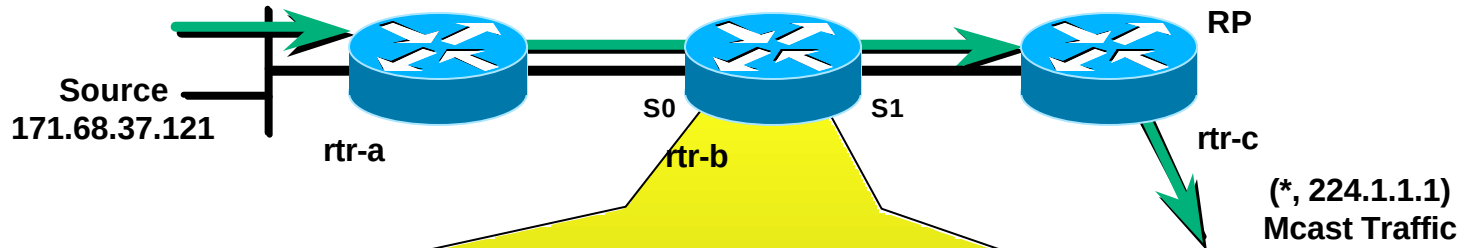
- **Receivers Join Group First**
- **Source Registers First**
- **Receivers along the SPT**



PIM SM Registering — Receivers along the SPT

(171.68.37.121, 224.1.1.1)

Mcast Packets



```
(*, 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

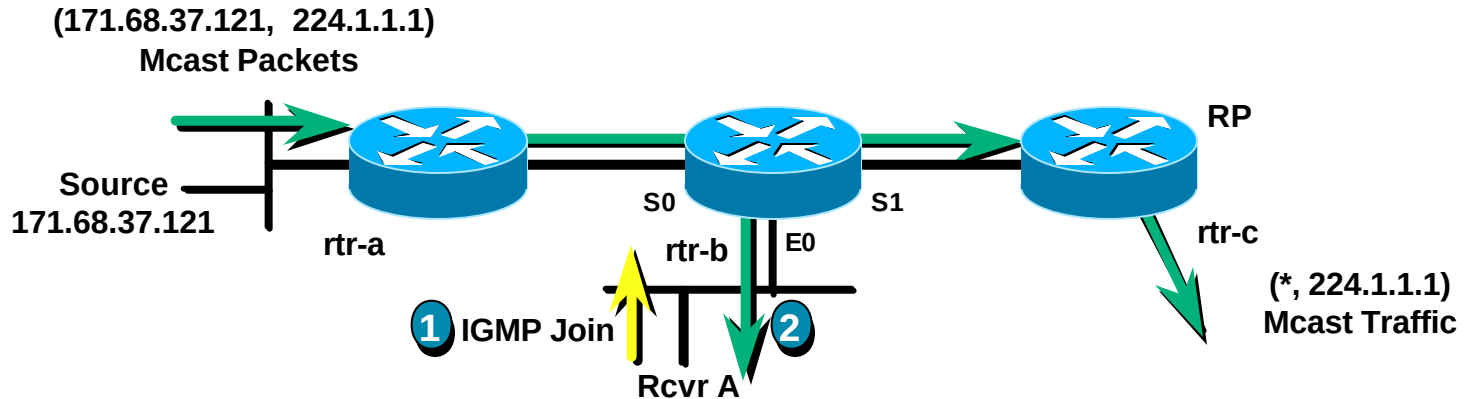
(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags: T
Incoming interface: Serial0, RPF nbr 171.68.28.190
Outgoing interface list:
  Serial1, Forward/Sparse, 00:04:28/00:01:32
```

Current state in “rtr-b”



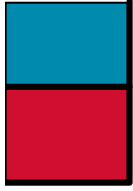


PIM SM Registering — Receivers along the SPT



- 1 “Rcvr A” wishes to receive group G traffic. Sends IGMP Join for G.
- 2 Group G traffic begins to flow to “Rcvr A”.

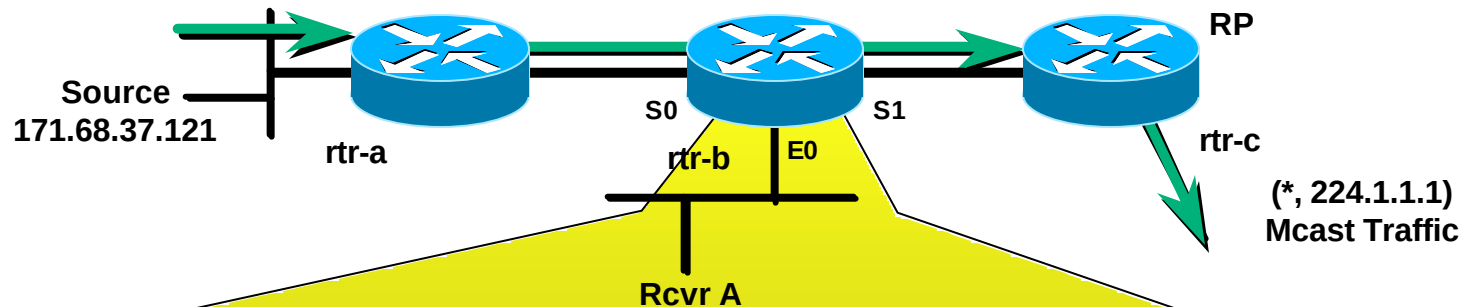




PIM SM Registering — Receivers along the SPT

(171.68.37.121, 224.1.1.1)

Mcast Packets



```
(*, 224.1.1.1), 00:04:28/00:01:32, RP 171.68.28.140, flags: SC  
Incoming interface: Serial1, RPF nbr 171.68.28.140,  
Outgoing interface list:  
  Ethernet0, Forward/Sparse, 00:00:30/00:02:30
```

```
(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags: CT  
Incoming interface: Serial0, RPF nbr 171.68.28.190  
Outgoing interface list:  
  Serial1, Forward/Sparse, 00:04:28/00:01:32  
  Ethernet0, Forward/Sparse, 00:00:30/00:02:30
```

Added
Interfaces

State in “rtr-b” after “Rcvr A” joins group
(Note: 171.68.37.121 traffic doesn't flow to RP then back down to rtr-b)





PIM-SM Advanced Concepts

- PIM SM State
- PIM SM Forwarding
- PIM SM Joining
- PIM SM Registering
- **PIM SM SPT-Switchover**
- PIM SM Pruning
- PIM SM State Maintenance



PIM SM SPT-Switchover

- **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.





PIM SM SPT-Switchover

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:

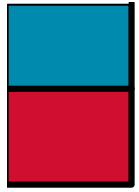
If “J” flag set in (*, G)

Join SPT for (S_i , G)

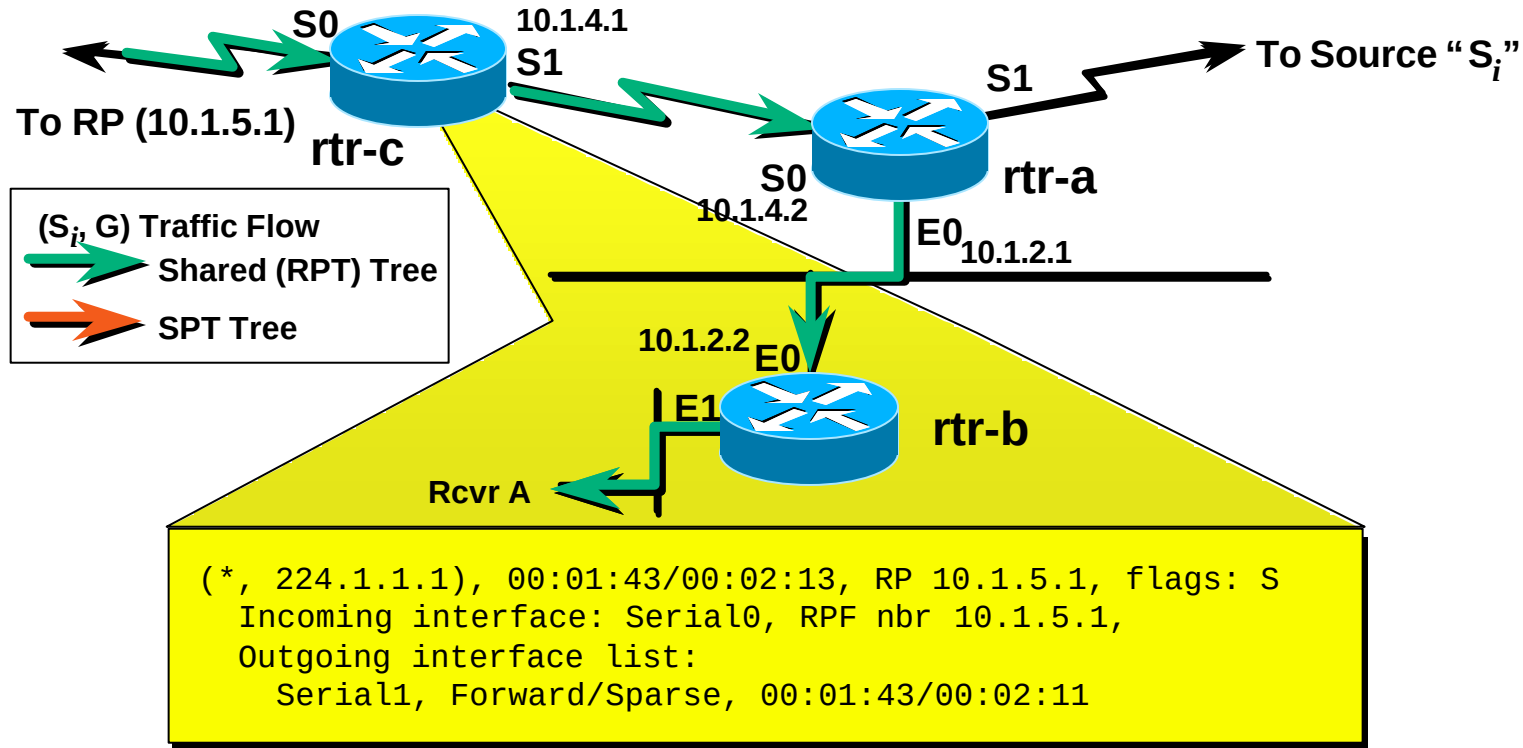
Mark (S_i , G) entry with “J” flag

Clear “J” flag in (*,G)



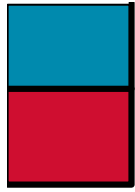


PIM SM SPT-Switchover

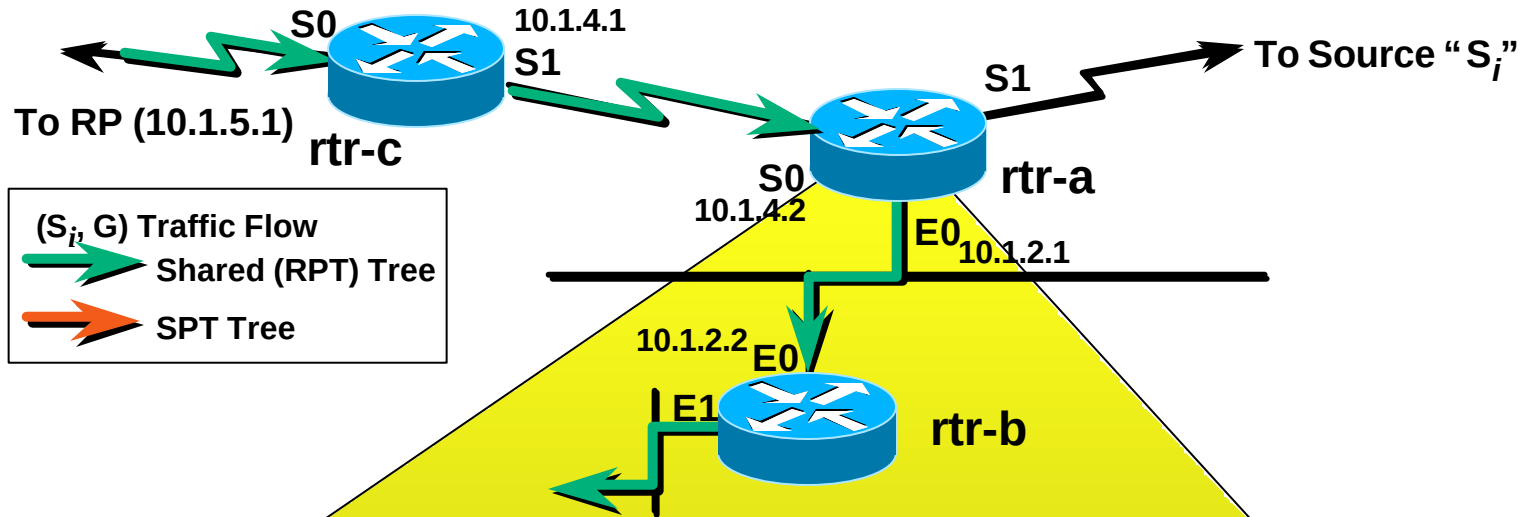


State in "rtr-c" before switch





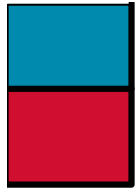
PIM SM SPT-Switchover



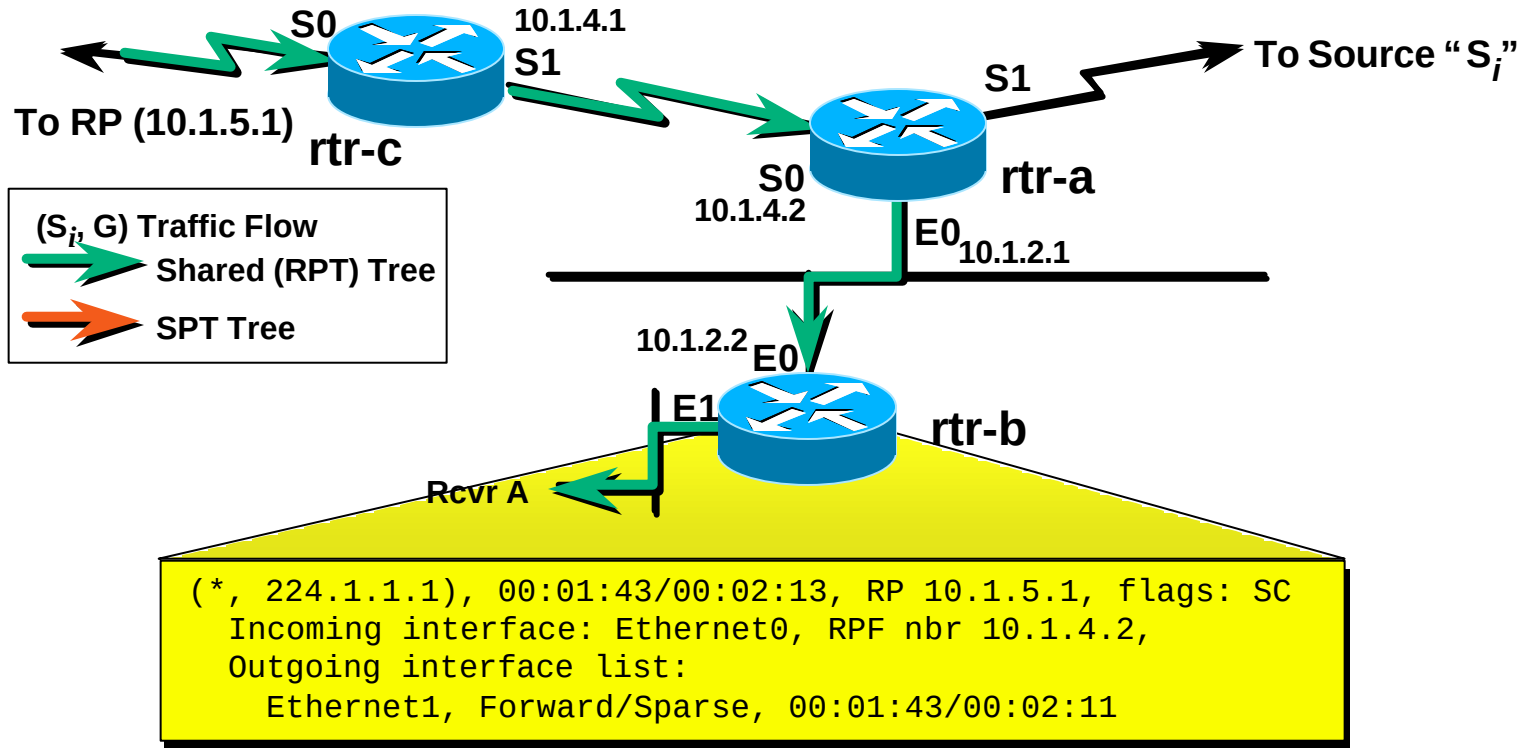
```
(*, 224.1.1.1), 00:01:43/00:02:13, RP 10.1.5.1, flags: S
Incoming interface: Serial0, RPF nbr 10.1.4.1,
Outgoing interface list:
  Ethernet0, Forward/Sparse, 00:01:43/00:02:11
```

State in "rtr-a" before switch



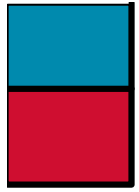


PIM SM SPT-Switchover

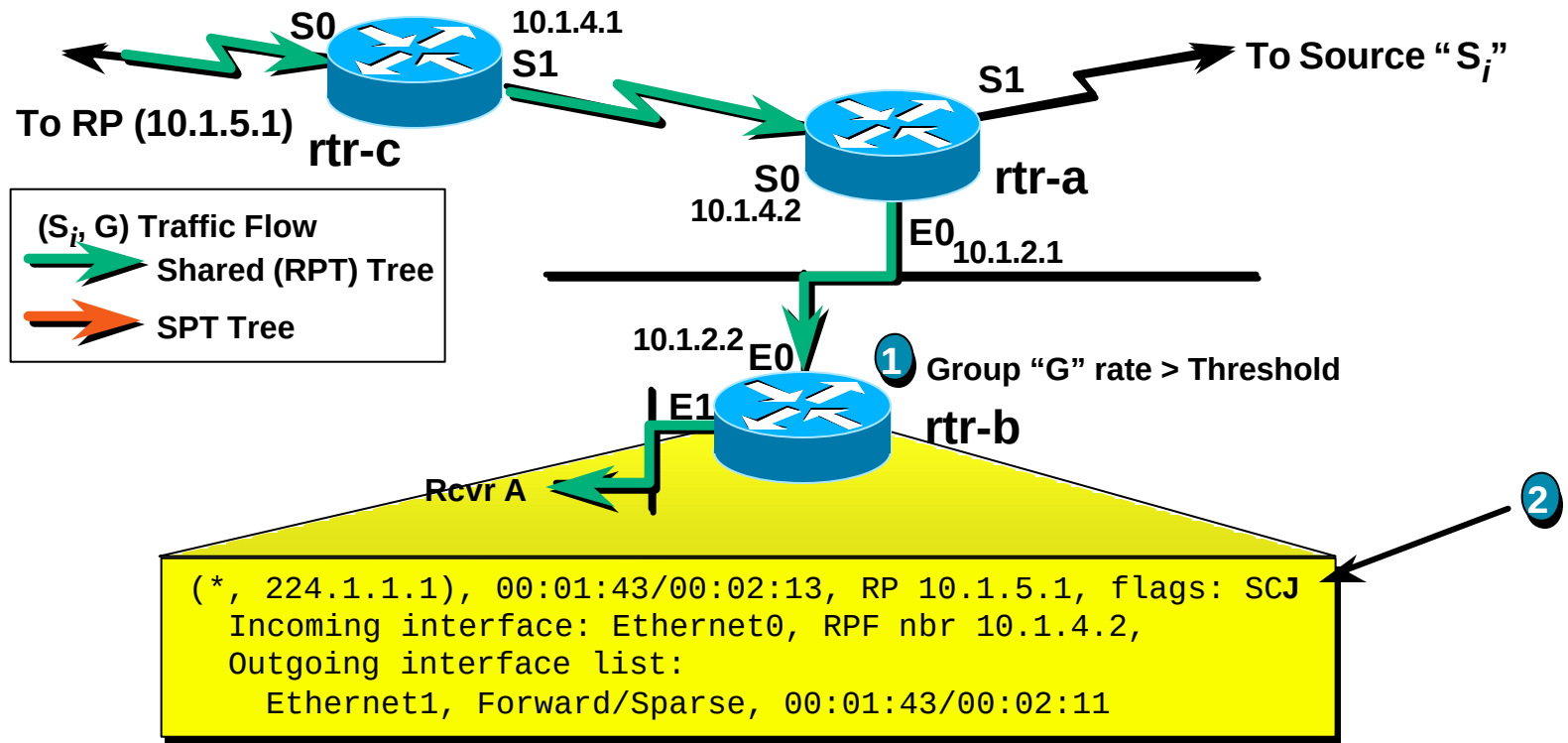


State in "rtr-b" before switch



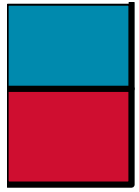


PIM SM SPT-Switchover

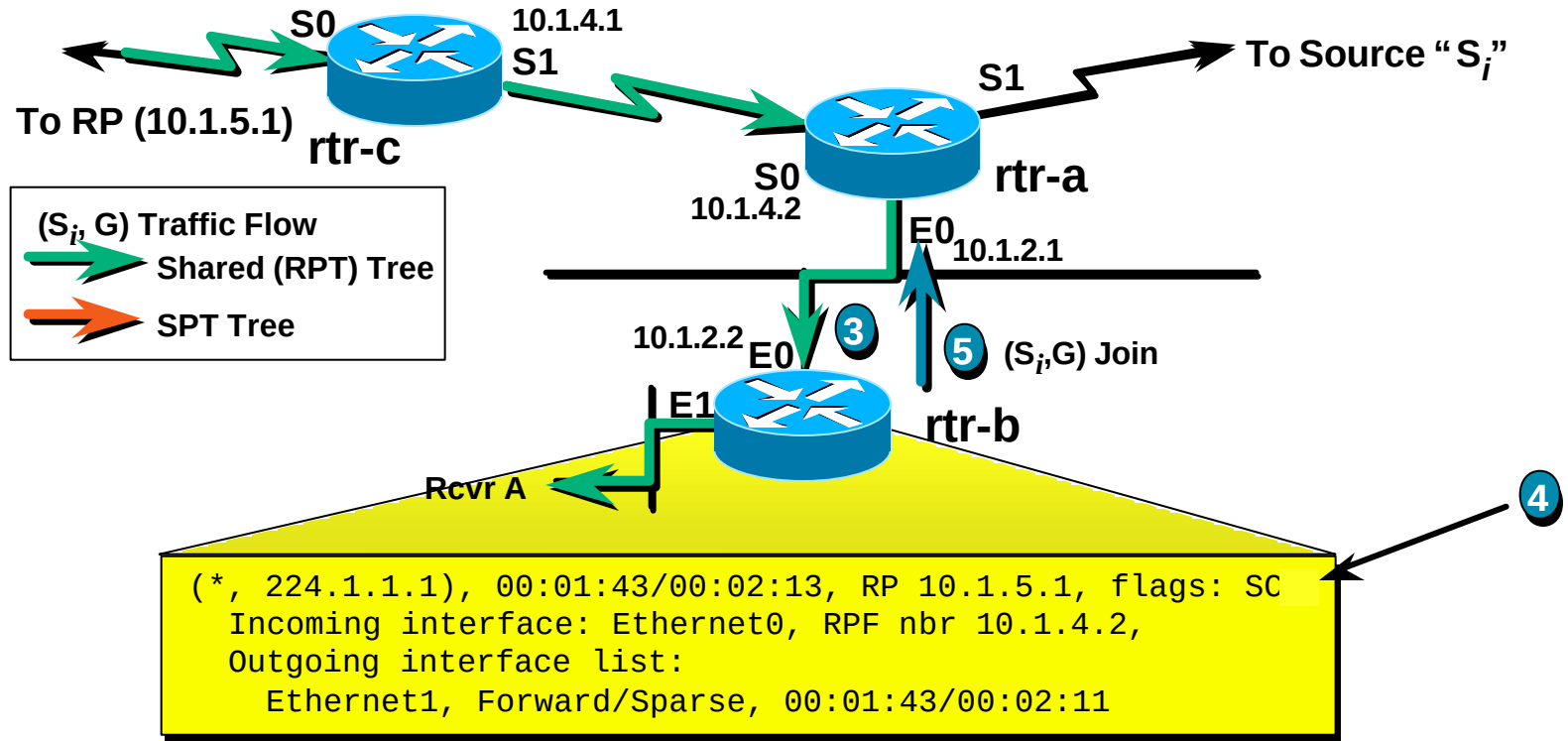


- 1 Group "G" rate exceeds SPT Threshold at "rtr-b";
- 2 Set J Flag in $(*, G)$ and wait for next (S_i, G) packet.



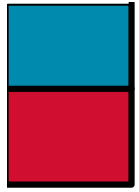


PIM SM SPT-Switchover

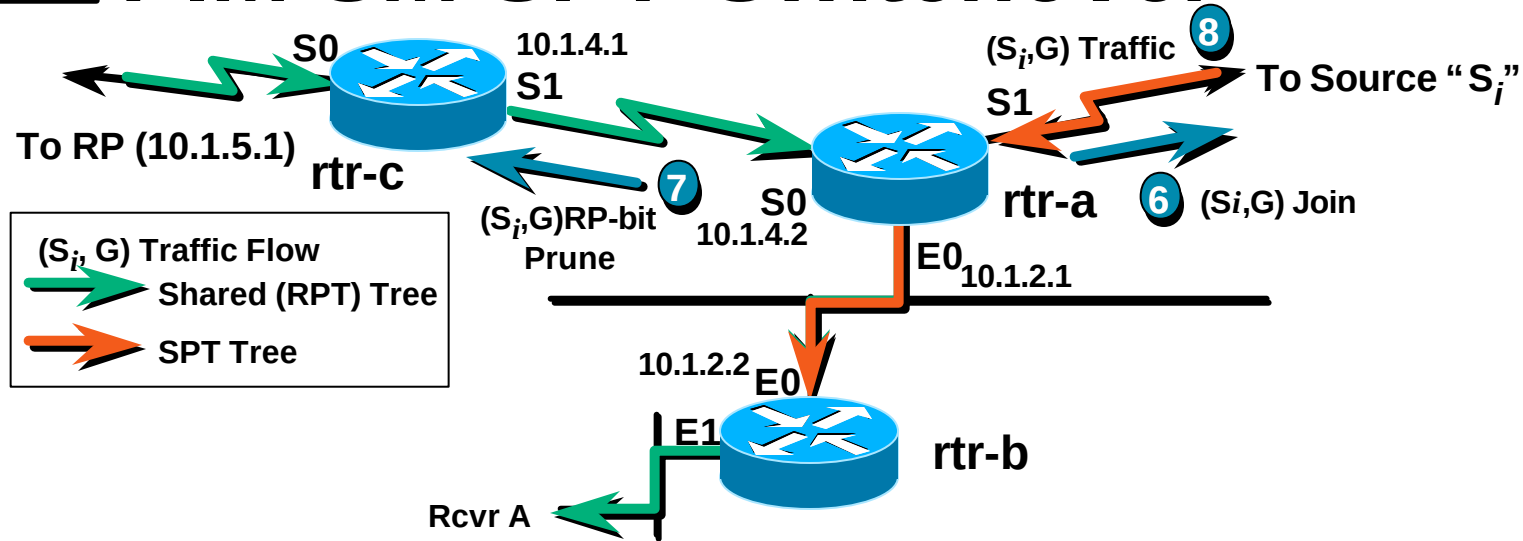


- 1 Group "G" rate exceeds SPT Threshold at "rtr-b";
- 2 Set J Flag in (*, G) and wait for next (S_i, G) packet.
- 3 (S_i, G) packet arrives down Shared tree.
- 4 Clear J Flag in (*, G) & 5 Send (S_i, G) Join towards S_i .

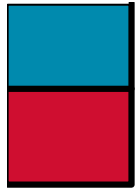




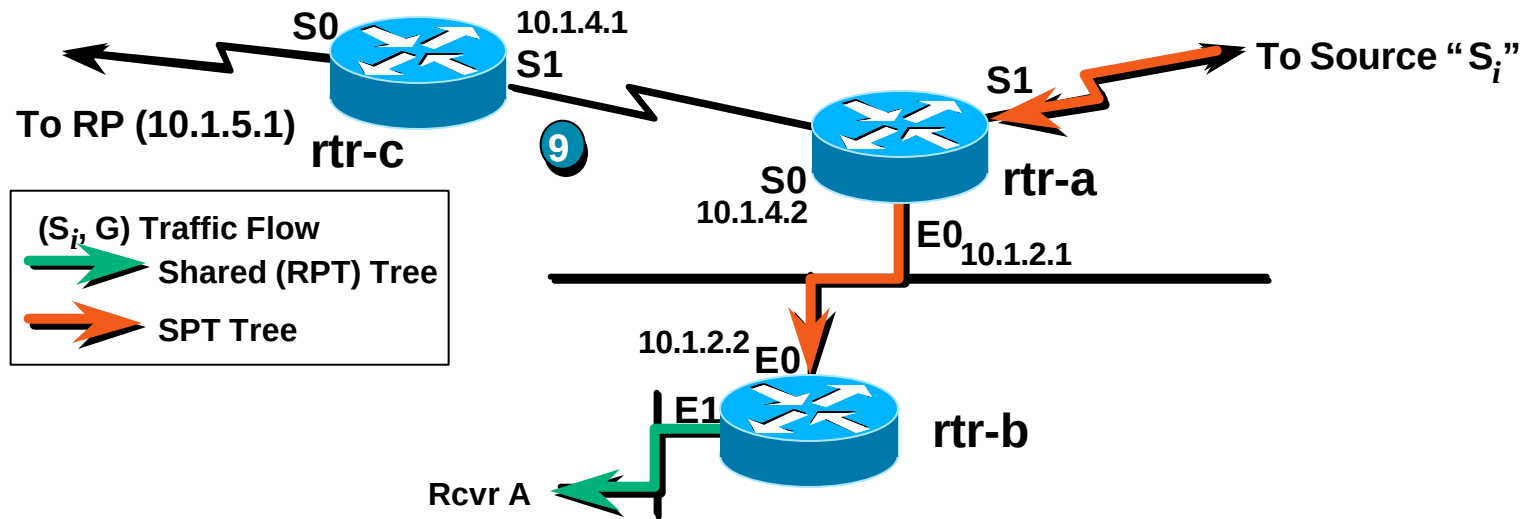
PIM SM SPT-Switchover



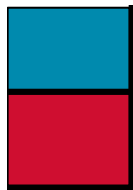
- 6 SPT & RPT diverge; "rtr-a" forwards (S_i, G) Join toward S_i .
- 7 "rtr-a" triggers (S_i, G) RP-bit Prune toward RP.
- 8 (S_i, G) traffic begins flowing down SPT tree.



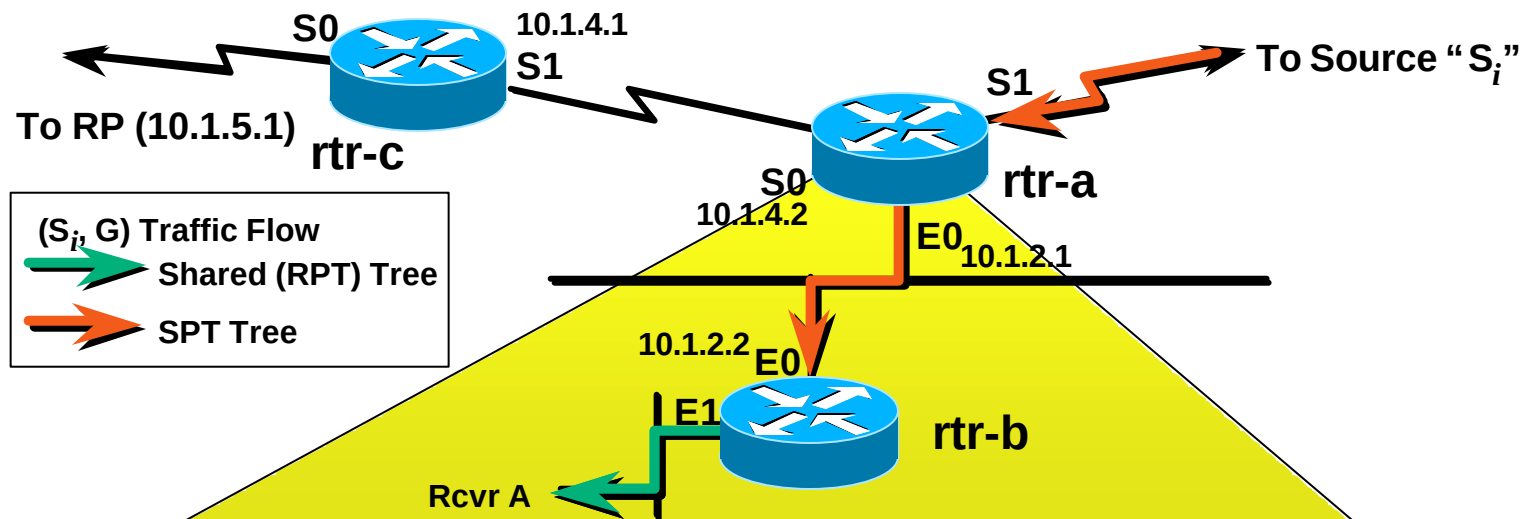
PIM SM SPT-Switchover



- ⑥ SPT & RPT diverge; “rtr-a” forwards (S_i, G) Join toward S_i .
- ⑦ “rtr-a” triggers (S_i, G) RP-bit Prune toward RP.
- ⑧ (S_i, G) traffic begins flowing down SPT tree.
- ⑨ (S_i, G) traffic ceases flowing down Shared tree.



PIM SM SPT-Switchover

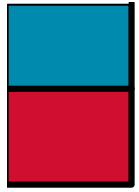


```
(*, 224.1.1.1), 00:01:43/00:02:13, RP 10.1.5.1, flags: S
Incoming interface: Serial0, RPF nbr 10.1.4.1,
Outgoing interface list:
  Ethernet0, Forward/Sparse, 00:01:43/00:02:11

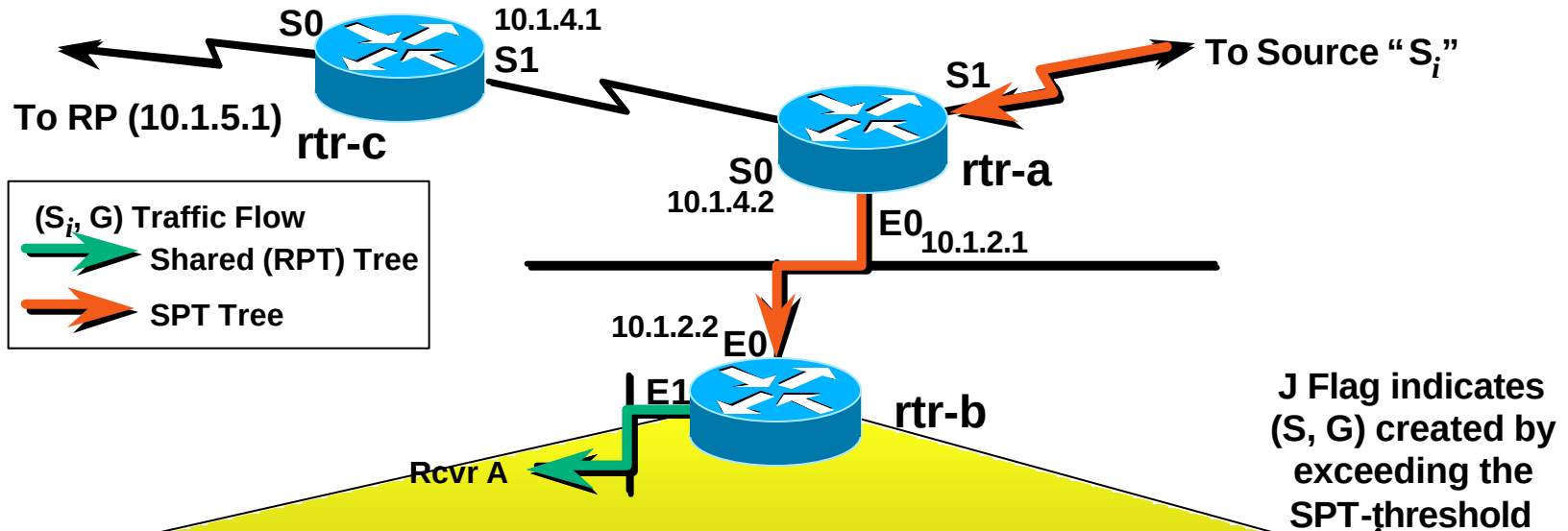
(171.68.37.121/32, 224.1.1.1), 00:13:28/00:02:53, flags: T
Incoming interface: Serial1, RPF nbr 10.1.9.2
Outgoing interface list:
  Ethernet0, Forward/Sparse, 00:13:25/00:02:30
```

State in "rtr-a" after switch





PIM SM SPT-Switchover

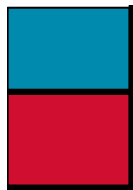


```
(*, 224.1.1.1), 00:01:43/00:02:13, RP 10.1.5.1, flags: SC
Incoming interface: Ethernet0, RPF nbr 10.1.2.1,
Outgoing interface list:
  Ethernet1, Forward/Sparse, 00:01:43/00:02:11

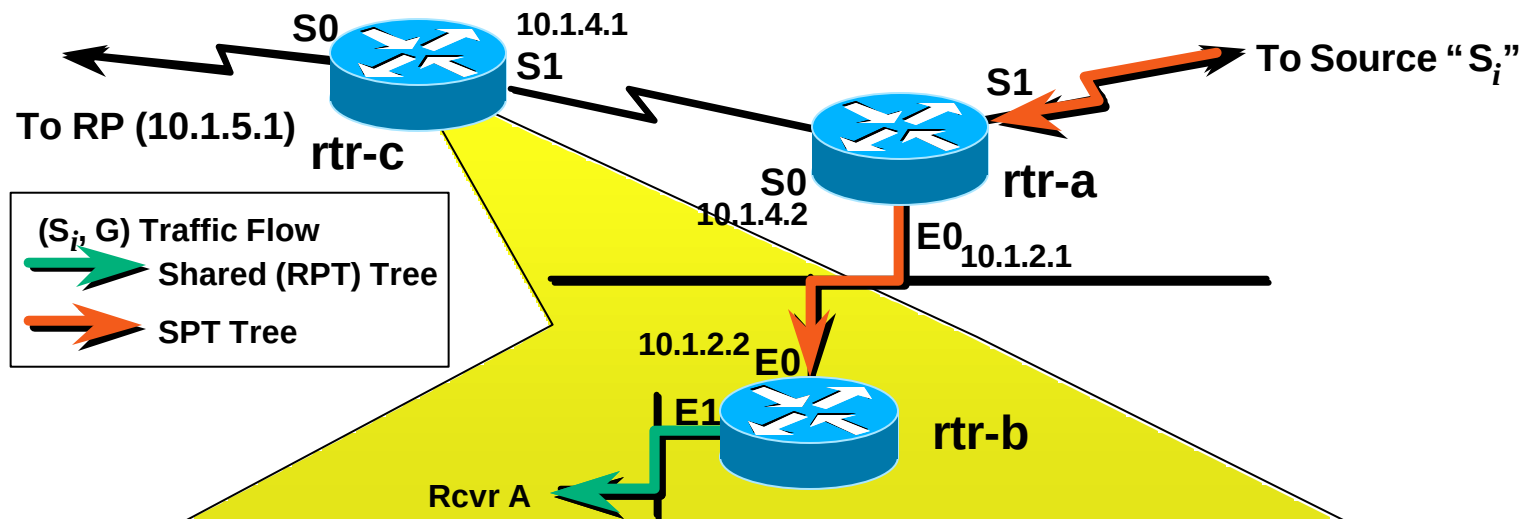
(171.68.37.121/32, 224.1.1.1), 00:13:28/00:02:53, flags: CJT
Incoming interface: Ethernet0, RPF nbr 10.1.2.1
Outgoing interface list:
  Ethernet1, Forward/Sparse, 00:13:28/00:02:53
```

State in "rtr-b" after switch





PIM SM SPT-Switchover



```
(*, 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
```

```
(171.68.37.121/32, 224.1.1.1), 00:13:28/00:02:53, flags: PR
Incoming interface: Serial0, RPF nbr 10.1.5.1
Outgoing interface list: Null
```

State in "rtr-c" after switch





PIM SM SPT-Switchover

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





PIM-SM Advanced Concepts

- PIM SM State
- PIM SM Forwarding
- PIM SM Joining
- PIM SM Registering
- PIM SM SPT-Switchover
- **PIM SM Pruning**
- PIM SM State Maintenance



PIM SM Pruning

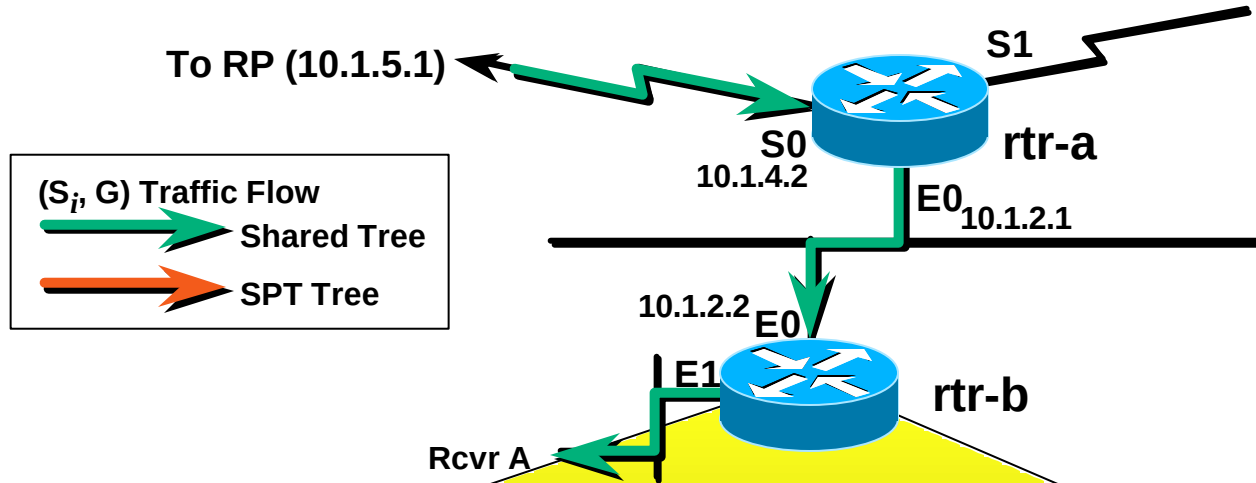
- IGMP group times out / last host sends Leave
- Interface removed from all (*,G) & (S,G) entries
 - IF all interfaces in “*oilist*” for (*,G) are pruned;
THEN send Prune up shared tree toward RP
 - Any (S, G) state allowed to time-out
- Each router along path “prunes” interface
 - IF all interfaces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





PIM SM Pruning —

Shared Tree Case



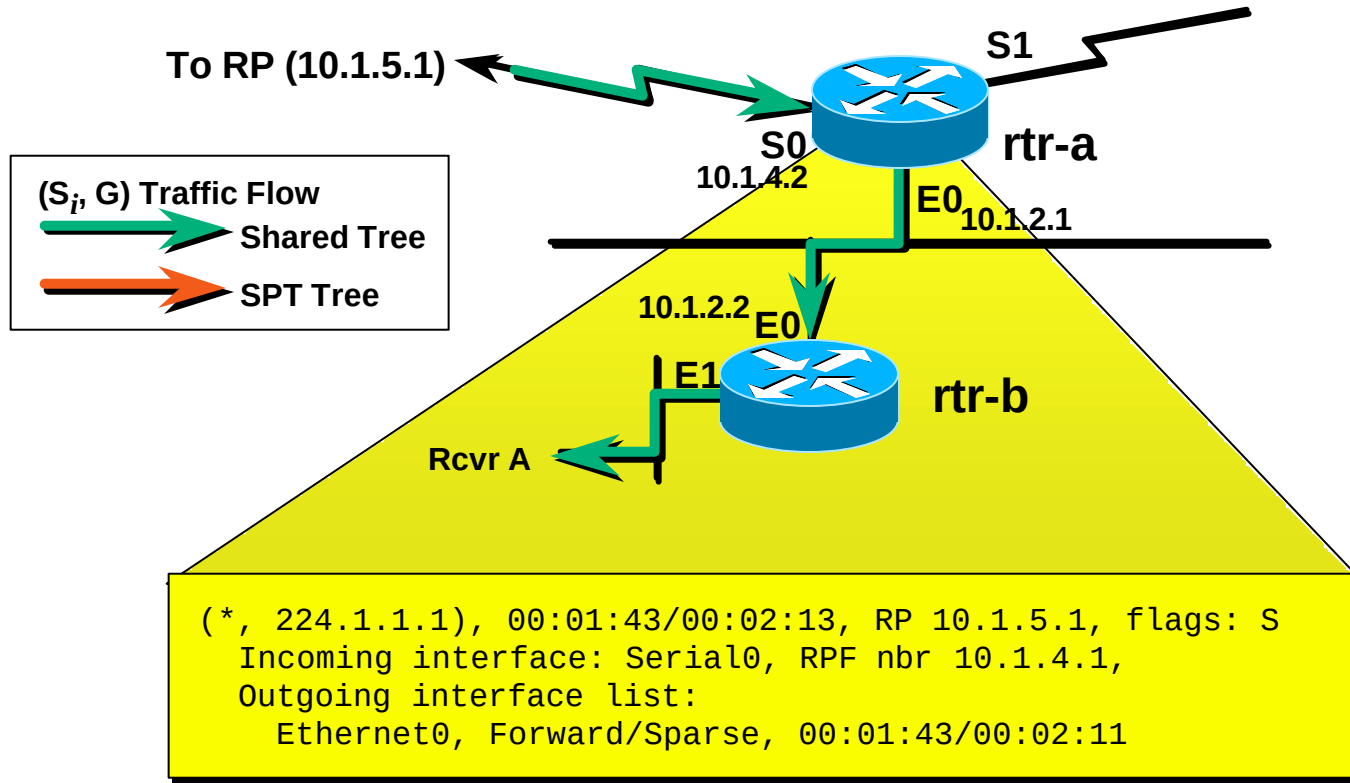
(*, 224.1.1.1), 00:01:43/00:02:13, RP 10.1.5.1, flags: SC
Incoming interface: Ethernet0, RPF nbr 10.1.2.1,
Outgoing interface list:
Ethernet1, Forward/Sparse, 00:01:43/00:02:11

State in "rtr-b" before Pruning





PIM SM Pruning — Shared Tree Case

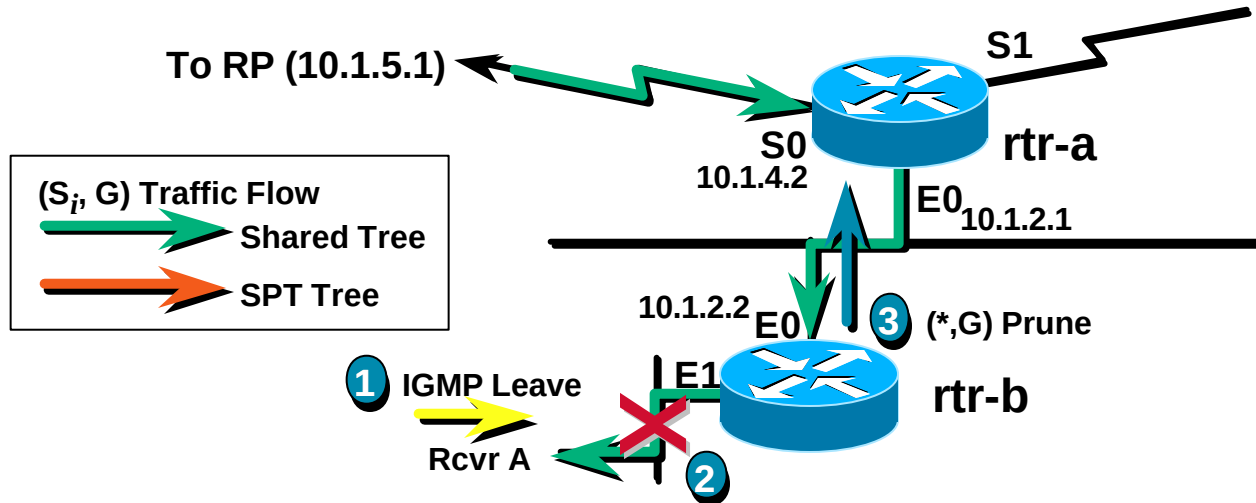


State in “rtr-a” before Pruning





PIM SM Pruning — Shared Tree Case

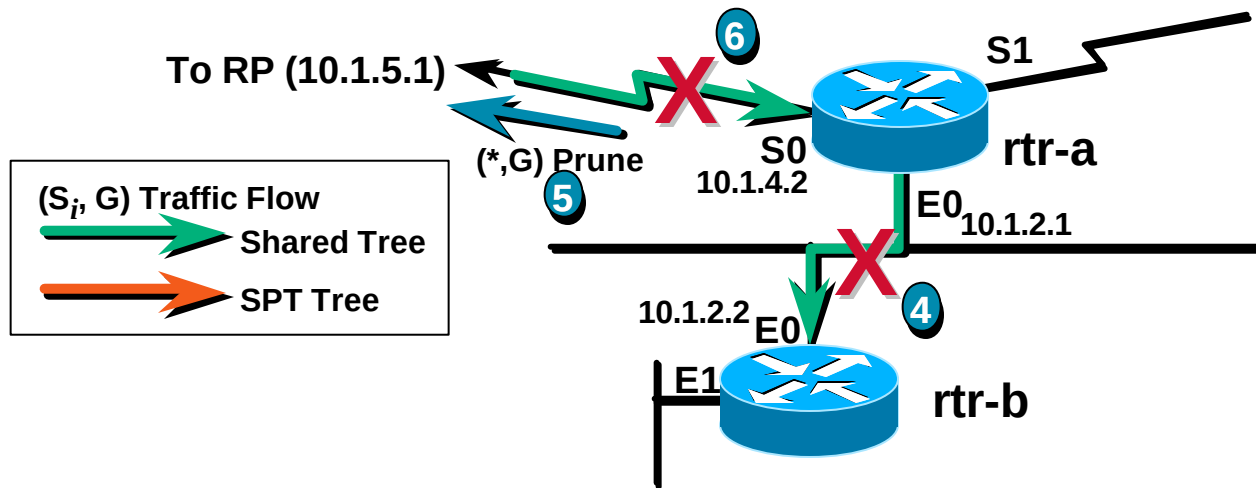


- 1 “rtr-b” is a Leaf router. Last host “Rcvr A”, leaves group G.
- 2 “rtr-b” removes E1 from $(*,G)$ and any (S_i,G) “oilists”.
- 3 “rtr-b” $(*,G)$ “oilist” now empty; sends $(*,G)$ Prune toward RP.





PIM SM Pruning — Shared Tree Case

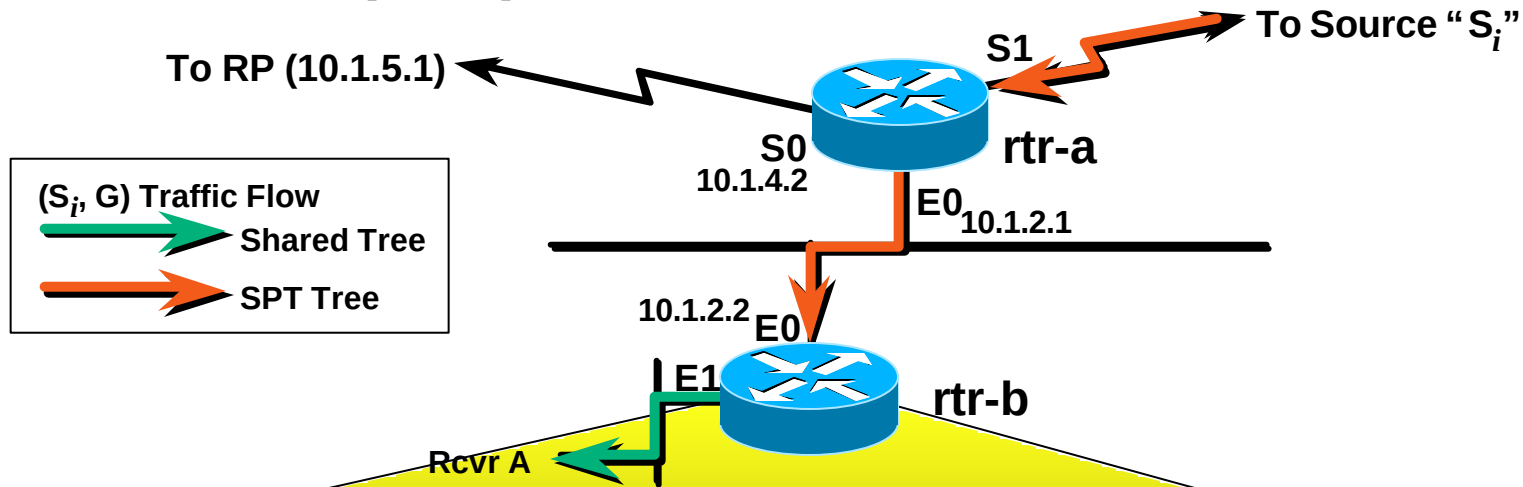


- ④ “rtr-a” receives Prune; removes E0 from $(*,G)$ “oilist”.
(After the 3 second Multi-access Network Prune delay.)
- ⑤ “rtr-a” $(*,G)$ “oilist” now empty; send $(*,G)$ Prune toward RP.
- ⑥ Pruning continues back toward RP.





PIM SM Pruning — Source (SPT) Case



```
(*, 224.1.1.1), 00:01:43/00:02:13, RP 10.1.5.1, flags: SC  
Incoming interface: Ethernet0, RPF nbr 10.1.2.1,  
Outgoing interface list:  
Ethernet1, Forward/Sparse, 00:01:43/00:02:11
```

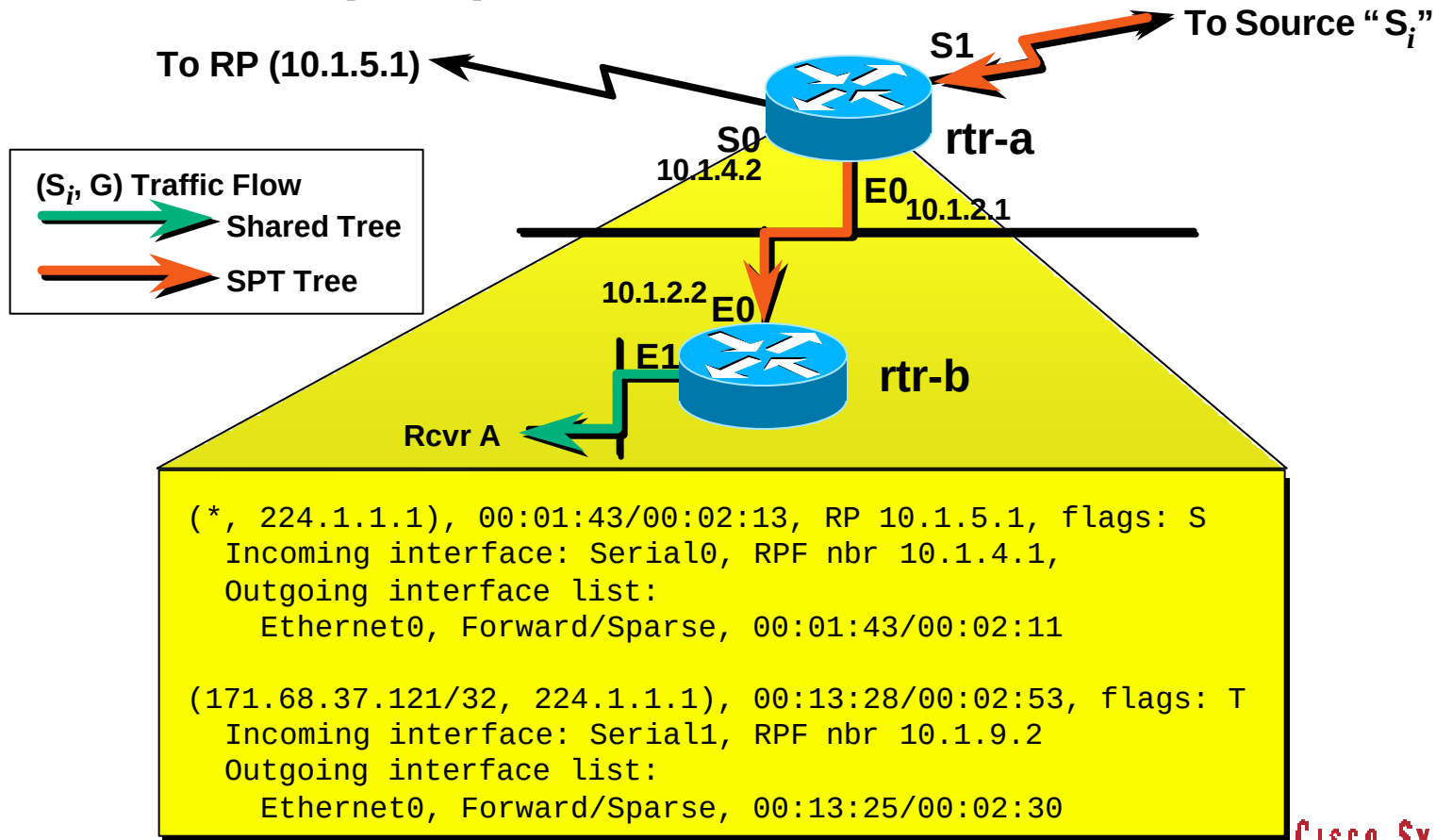
```
(171.68.37.121/32, 224.1.1.1), 00:13:28/00:02:53, flags: CJT  
Incoming interface: Ethernet0, RPF nbr 10.1.2.1  
Outgoing interface list:  
Ethernet1, Forward/Sparse, 00:13:25/00:02:30
```

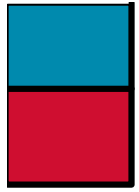
State in "rtr-b" before Pruning



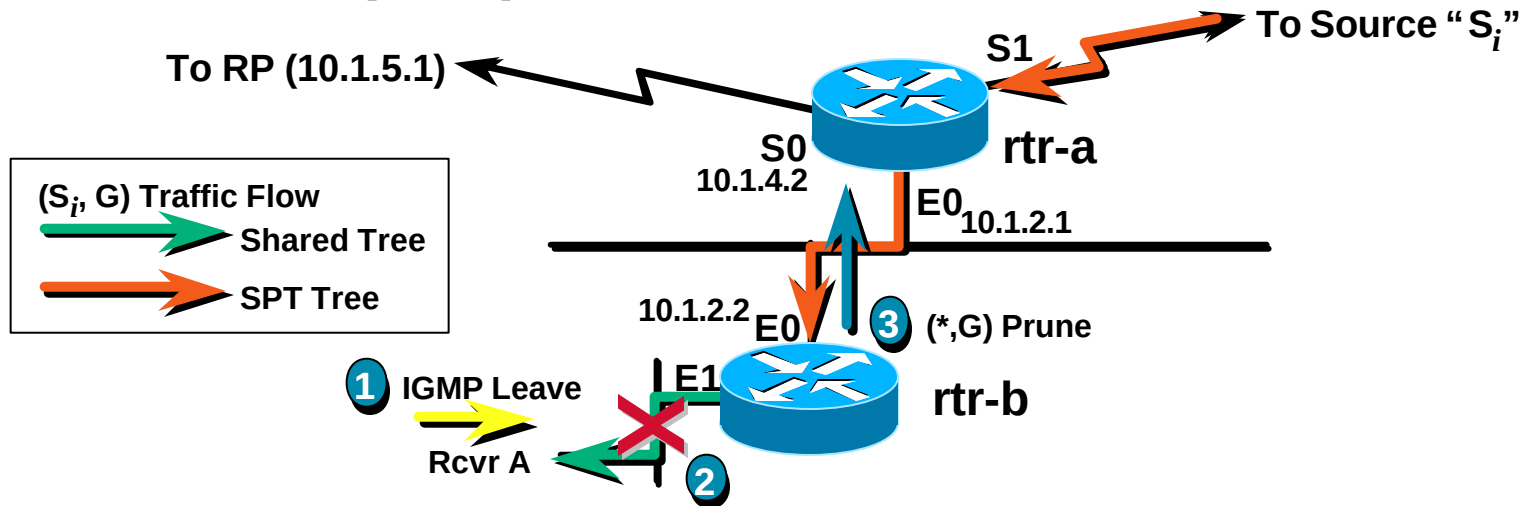


PIM SM Pruning — Source (SPT) Case





PIM SM Pruning — Source (SPT) Case

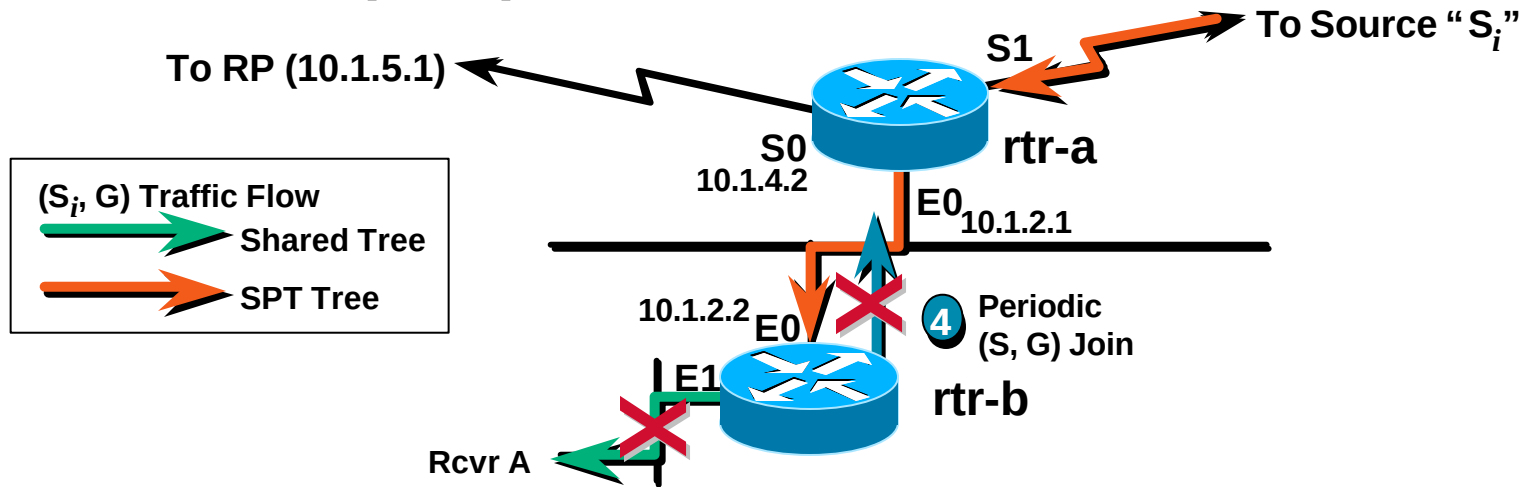


- 1 “rtr-b” is a Leaf router. Last host “Rcvr A”, leaves group G.
- 2 “rtr-b” removes E1 from (*,G) and any (S_i, G) “oilists”.
- 3 “rtr-b” (*,G) “oilist” now empty; sends (*,G) Prune toward RP.





PIM SM Pruning — Source (SPT) Case

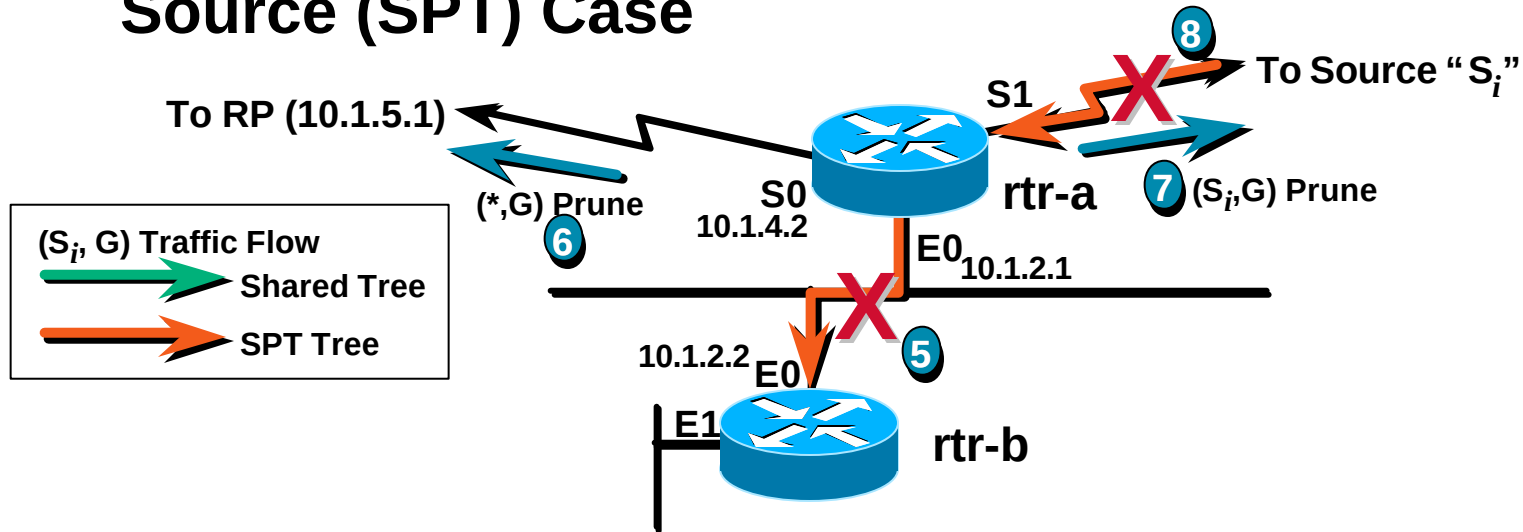


- 1 "rtr-b" is a Leaf router. Last host "Rcvr A", leaves group G.
- 2 "rtr-b" removes E1 from (*,G) and (S_i,G) "oilists".
- 3 "rtr-b" (*,G) "oilist" now empty; sends (*,G) Prune toward RP.
- 4 "rtr-b" stops sending periodic (S, G) joins.





PIM SM Pruning — Source (SPT) Case



- 5 "rtr-a" receives Prune; removes E0 from (*,G) "oilist".
(After the 3 second Multiaccess Network Prune delay.)
- 6 "rtr-a" (*,G) "oilist" now empty; sends (*,G) Prune toward RP.
- 7 (S_i,G) state times out. (S_i,G) Prune sent toward S_i.
- 8 (S_i,G) traffic ceases flowing down SPT.





PIM-SM Advanced Concepts

- PIM SM State
- PIM SM Forwarding
- PIM SM Joining
- PIM SM Registering
- PIM SM SPT-Switchover
- PIM SM Pruning
- **PIM SM State Maintenance**



PIM SM State Maintenance

- 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.

