Cisco Systems

IP Multicast

Overview on what we need to know !

Agenda

Cisco.com

- **Multicast Addressing**
- **IGMP**
- **Multicast Distribution Trees**
- **PIM Modes**
- **Anycast RP**

IP Multicast Addressing

Cisco.com

- **Multicast Group Concept**
  - **Multicast Group == arbitrary group of receivers**
    - receivers == host with interest in receiving multicast data stream
  - **Hosts signal interest to "join" a multicast group with IGMP**
    - receivers join or leave a group dynamically
- **IP Multicast Addresses == IP class D address**
  - 1110.....

Link Local Address (reserved)	**224.0.0.0 /24**
Globally Scoped Addresses	**224.0.1.0 - 238.255.255.255**
- Source Specific MCast **- GLOP address**	**232.0.0.0 /8** **233.0.0.0 /8**
Limited Scoped Addresses	**239.0.0.0 – 239.255.255.255**

Layer-2 Multicast Addressing

Cisco.com

- **Why is it needed ?**
- **Implementation depending on different layer-2 protocols**

Token Ring

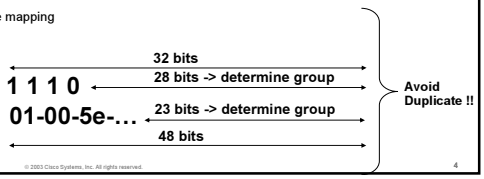
all mcast groups map into all 1's broadcast or in single functional mcast address

RPR

a single bit in the frame distinguishs multi/broadcast from unicast

Ethernet / FDDI

Incomplete mapping



Ethernet/FDDI Mcast Address

Cisco.com

2^5 (32) IP multicast addresses
Map into the same layer-2 mcast

224.1.1.1
224.129.1.1
225.1.1.1
225.129.1.1
:
:
:
239.1.1.1
239.129.1.1

→ 01-00-5e-01-01-01

Agenda

Cisco.com

- **Multicast Addressing**
- **IGMP**
- **Multicast Distribution Trees**
- **PIM Modes**
- **Anycast RP**

IGMP
Internet Group Management Protocol

Cisco.com

- IGMP version 1, 2 and 3

IGMP v1

ver	type	unused	checksum
Group Address			

Membership query
Membership report

IGMP v2

Ver/type	Max response Time	checksum
Group Address		

Membership Query
Version1 Membership Report
Version2 membership Report
Leave Group Message

IGMPv3 Source Specific Joins

Multicast in Layer-2 switching environment

Cisco.com

Problem:

Mcast router sees LAN – sends mcast to LAN

➔Both host get mcast traffic – only one wants it

Solution:

IGMP snooping

(intelligent switchs intercepts IGMP)

CGMP

(router switch cooperation protocol, where router instructs switch to open/close specific ports)

Agenda

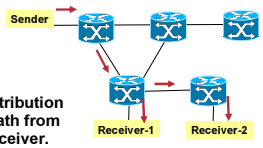
Cisco.com

- Multicast Addressing
- IGMP
- Multicast Distribution Trees // RPF
- PIM Modes
- Anycast RP

Multicast Distribution Trees

Cisco.com

Source Trees

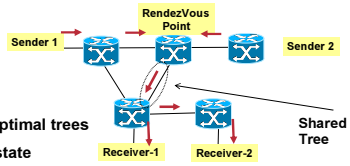


- Builds optimal distribution tree on shortest path from source to every receiver.
- Every Router has source, group specific forwarding entries (S,G)
  - ➔100 sources one group means 100 S,G states

Multicast Distribution Trees

Cisco.com

Shared Trees



- Might build sub-optimal trees
- Less forwarding state consumed
  - ➔ (RP,G) versus (S,G)
  - ➔100 sources one group means 1 RP,G states

Multicast Forwarding /

Cisco.com

- Reverse Path Forwarding

"learn to stand on your head"

The source not just the destination becomes crucial for the forwarding decision

Multicast Forwarding Decision // RPF

Cisco.com

- When a multicast packet arrives a router performs a RPF check.

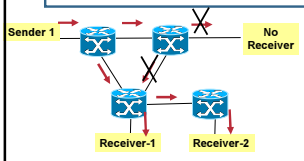
  If the packet arrives on the correct RPF interface

  ➔Packet is forwarded to all neighboring mcast routers

  ➔ Packet is forwarded to LANs with Receivers

  If the packet does NOT arrive on RPF interface

  ➔the packet is forwarded, otherwise it is dropped



Definition:
RPF interface is the interface, "closest" to the source.
==
The RPF interface is the interface to which the route to the source points

RPF – Multicast Routing

Cisco.com

- Which routes are used for the RPF ?
  1. Can be a specific multicast Routing Protocol

     DVMRP, mBGP

     Useful if Unicast and Multicast is to be kept apart

     ~ example MBONE as a subset of the internet
  2. Can be the normal unicast Routing Protocols

     PIM (Protocol Independent Multicast) uses the Unicast Routing Table ➔ PIM is based on unicast routing
  3. PIM uses distance preferred route lookup across tables

     Distance prefered lookup comes into play, when there is routing information besides unicast routes

     (i.e. static mroute, mBGP, DVMRP)

Agenda

Cisco.com

- Multicast Addressing
- IGMP
- Multicast Distribution Trees
- PIM Modes
- Anycast RP

PIM – Different Operating Modes

Cisco.com

PIM is always a soft-state protocol

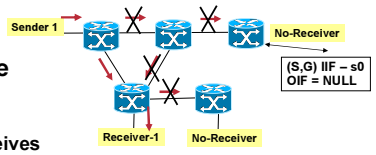
➔States time out after a while, need to be refreshed ➔considerable state maintenance

- PIM Dense Mode
  - Multicast Trees are built data-driven
  - Easy to start with, but not scaling. Should NEVER be used in production networks
- PIM Sparse Mode
  - Multicast trees, built control-driven and independent of actual traffic
  - RP needed so that receivers can "meet" sources.
  - Actual data traffic can flow on the shortest path
  - Scalable as state is only built along actual distribution trees
- PIM Sparse-Dense Mode
  - All groups without RP (RendezVous Point) are handled as Dense
- PIM Source Specific Multicast
  - If Sources are known beforehand, no needfor a RP.

PIM Dense Mode

Cisco.com

Flood & Prune



Every router receives and floods the multicast to every other router towards every LAN segment.

If there is no interest the last-hop router signals this upstream to "prune" the multicast tree

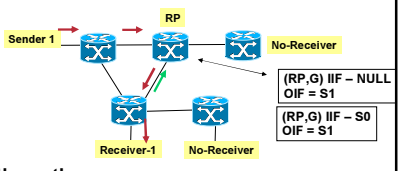
Problem:

Every router now has Multicast state for every Group !!!

PIM Sparse Mode

Cisco.com

Control Driven



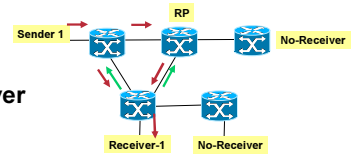
Only Router along the path from Source to RP or from RP to Receiver will create (S,G) or (RP,G) state

Problem:

Traffic not on optimal == shortest path from Source to receiver

PIM Sparse Mode cont.

Cisco.com



SPT Switch Over

Once last Hop Router receives traffic
➔Knows the source initiates direct JOIN to the Source

Once it receives the traffic on the shortest path tree the RP tree is "pruned"

NO Problem: ☺

Agenda

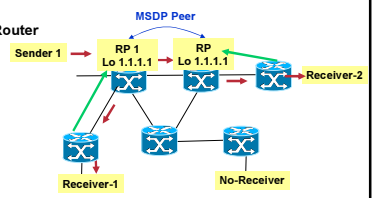
Cisco.com

- Multicast Addressing
- IGMP
- Multicast Distribution Trees
- PIM Modes
- RP Redundancy // Anycast RP

PIM Sparse Mode -- RP Redundancy

Cisco.com

- Different Methods available
  - ➔Auto RP
  - ➔ PIM Bootsrap Router
  - ➔Anycast RP



Anycast RP

Last Hop Router
Joins "closest RP

RPs communicate via MSDP as soon as source becomes active

Last Hop Router receives traffic first via "closest" RP and then JOINs the source directly to switches to the SPT

Any questions ?

Cisco Systems

Deeper Dive Review of Multicast Fundamentals

Agenda

- Multicast Addressing
- Internet Group Management Procotol (IGMP)
- Multicast Forwarding
- PIM-Dense Mode (DM)
- PIM-Sparse Mode (SM)
- PIM-Source Specific Mode (SSM)
- PIM-Bidirectional (BIDIR)

Multicast Addressing

Cisco.com

- IP Multicast Group Addresses
  224.0.0.0 – 239.255.255.255
  Class "D" Address Space
    High order bits of 1st Octet = "1110"
- Reserved Link-local Addresses
  224.0.0.0 – 224.0.0.255
  Transmitted with TTL = 1
  Examples:

224.0.0.1	All systems on this subnet
224.0.0.2	All routers on this subnet
224.0.0.4	DVMRP routers
224.0.0.5	OSPF routers
224.0.0.13	PIMv2 Routers
224.0.0.22	IGMPv3

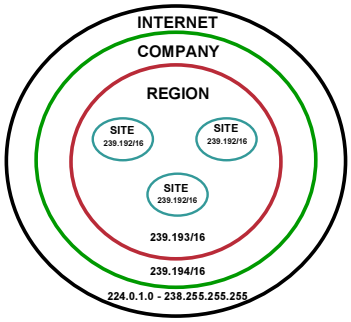
Multicast Addressing

Cisco.com

- Administratively Scoped Addresses
  239.0.0.0 – 239.255.255.255
  Private address space
    Similar to RFC1918 unicast addresses
    Not used for global Internet traffic
    Used to limit "scope" of multicast traffic
    Same addresses may be in use at different locations for different multicast sessions

IP Multicast Address Scoping

Cisco.com



How are Multicast Addresses Assigned ?

Cisco.com

- Dynamic Address Assignment
  - Session Directory Tool (SDR)
    - Historically used to announce session/group information on a well-known multicast group
    - Has scaling problems
  - Other proposed methods (MADCAP) ….

How are Multicast Addresses Assigned ? (contd.)

Cisco.com

- Static Global Group Address Assignment
  - Temporary method to meet immediate needs
  - Group range: 233.0.0.0 – 233.255.255.255
    - Your AS number is inserted in middle two octets
    - Remaining low-order octet used for group assignment
    - Example: Hexadecimal value of AS#5662 is 161E. 16 hex is 22 decimal and 1E hex is 30 decimal. 233.22.30.0/24.
  - Defined in RFC 2770
    - "GLOP Addressing in 233/8"
    - http://www.ogig.net/glop/ will figure it out for you.
- Manual Address Allocation by the Admin !!
  - Is still the most common practice in Enterprises

Agenda

Cisco.com

- Multicast Addressing
- Internet Group Management Protocol (IGMP)
- Multicast Forwarding
- PIM-Sparse Mode (SM)
- PIM-Source Specific Mode (SSM)
- PIM-Bidirectional (BIDIR)

Host-Router Signaling: IGMP

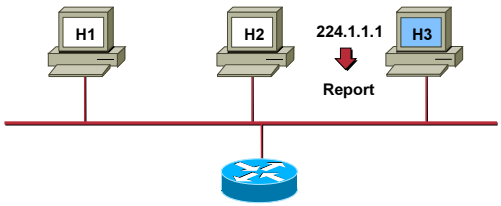
Cisco.com

- How hosts tell routers about group membership
- Routers solicit group membership from directly connected hosts
- RFC 1112 specifies version 1 of IGMP
  - Supported on Windows 95
- RFC 2236 specifies version 2 of IGMP
  - Supported on latest service pack for Windows and most UNIX systems
- RFC 3376 specifies version 3 of IGMP
  - Windows XP, FreeBSD, Linux have it.
  - ftp://ftpeng.cisco.com/ipmulticast/ssm/index.html#Stacks

Host-Router Signaling: IGMP

Cisco.com

Joining a Group

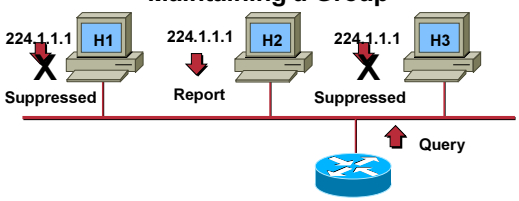


- Host sends IGMP Report to join group

Host-Router Signaling: IGMP

Cisco.com

Maintaining a Group

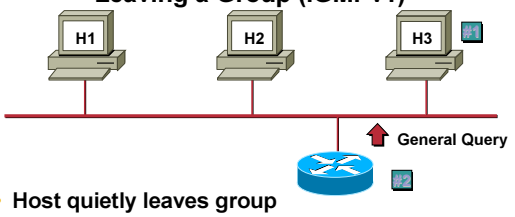


- Router sends periodic Queries to 224.0.0.1
- One member per group per subnet reports
- Other members suppress reports

Host-Router Signaling: IGMP

Cisco.com

Leaving a Group (IGMPv1)

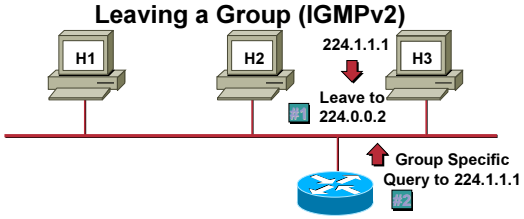


- Host quietly leaves group
- Router sends 3 General Queries (60 secs apart)
- No IGMP Report for the group is received
- Group times out (Worst case delay ~= 3 minutes)

Host-Router Signaling: IGMP

Cisco.com

Leaving a Group (IGMPv2)



- Host sends Leave message to 224.0.0.2
- Router sends Group specific query to 224.1.1.1
- No IGMP Report is received within ~3 seconds
- Group 224.1.1.1 times out

Host-Router Signaling: IGMPv3

Cisco.com

- Adds Include/Exclude Source Lists
  - Enables hosts to listen only to a specified subset of the hosts sending to the group
  - Requires new 'IPMulticastListen' API
    - New IGMPv3 stack required in the O/S.
  - Apps must be rewritten to use IGMPv3 Include/Exclude features

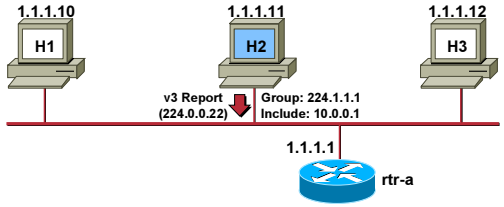
Host-Router Signaling: IGMPv3

Cisco.com

- New Membership Report address
  - 224.0.0.22 (IGMPv3 Routers)
    - All IGMPv3 Hosts send reports to this address
      - Instead of the target group address as in IGMPv1/v2
    - All IGMPv3 Routers listen to this address
    - Hosts do not listen or respond to this address
  - No Report Suppression
    - All Hosts on wire respond to Queries
      - Host's complete IGMP state sent in single response
    - Response Interval may be tuned over broad range
      - Useful when large numbers of hosts reside on subnet

IGMPv3—Joining specific Source(s)

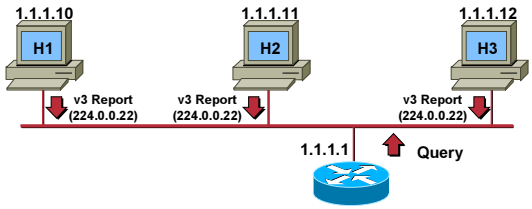
Cisco.com



- IGMPv3 Report contains desired source(s) in the Include list.
- Only "Included" source(s) are joined.

IGMPv3—Maintaining State

Cisco.com



- Router sends periodic queries
- All IGMPv3 members respond
  - Reports contain multiple Group state records

Agenda

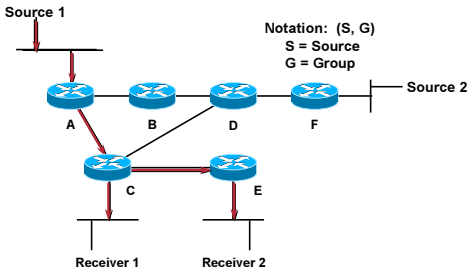
Cisco.com

- Multicast Addressing
- Internet Group Management Protocol (IGMP)
- Multicast Forwarding
- PIM-Sparse Mode (SM)
- PIM-Source Specific Mode (SSM)
- PIM-Bidirection (BIDIR)

Multicast Distribution Trees

Cisco.com

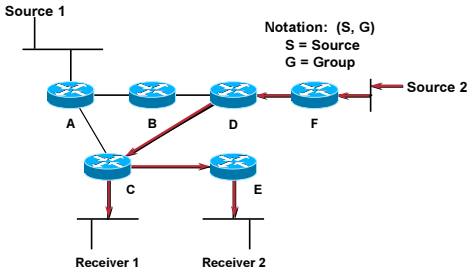
Shortest Path or Source Distribution Tree



Multicast Distribution Trees

Cisco.com

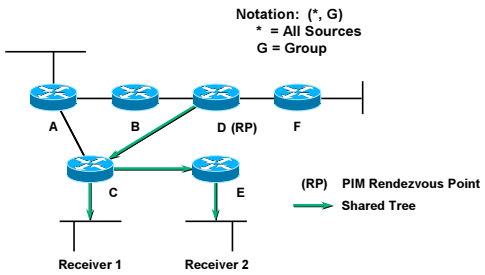
Shortest Path or Source Distribution Tree



Multicast Distribution Trees

Cisco.com

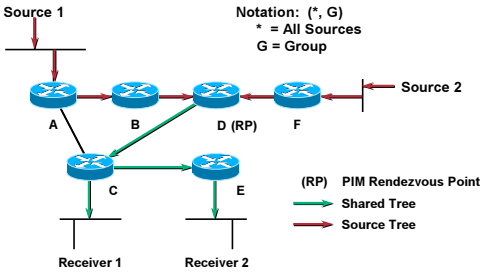
Shared Distribution Tree



Multicast Distribution Trees

Cisco.com

Shared Distribution Tree



Multicast Distribution Trees

Cisco.com

Characteristics of Distribution Trees

- Source or Shortest Path trees

  Uses more memory O(S x G) but you get optimal paths from source to all receivers; minimizes delay

- Shared trees

  Uses less memory O(G) but you may get sub-optimal paths from source to all receivers; may introduce extra delay

Multicast Forwarding

Cisco.com

- Multicast Routing is backwards from Unicast Routing

  Unicast Routing is concerned about where the packet is going.

  Multicast Routing is concerned about where the packet came from.

- Multicast Routing uses “Reverse Path Forwarding”

Multicast Forwarding

Cisco.com

Reverse Path Forwarding (RPF)

- **What is RPF?**

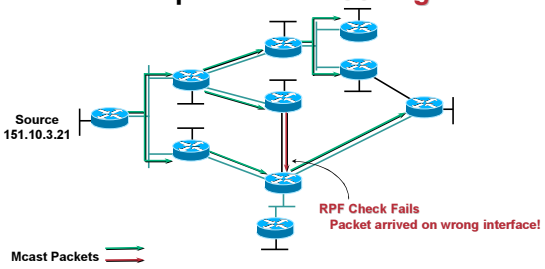
  A router forwards a multicast datagram only if received on the up stream interface to the source (i.e. it follows the distribution tree).

- **The RPF Check**
  - The routing table used for multicasting is checked against the “source” IP address in the packet.
  - If the datagram arrived on the interface specified in the routing table for the source address; then the RPF check succeeds.
  - Otherwise, the RPF Check fails.

Multicast Forwarding

Cisco.com

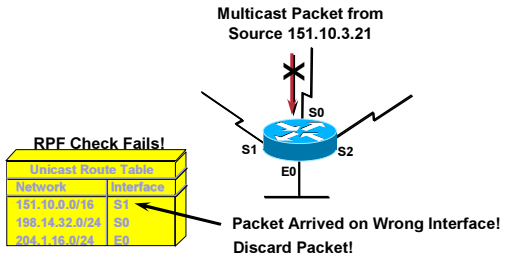
Example: RPF Checking



Multicast Forwarding

Cisco.com

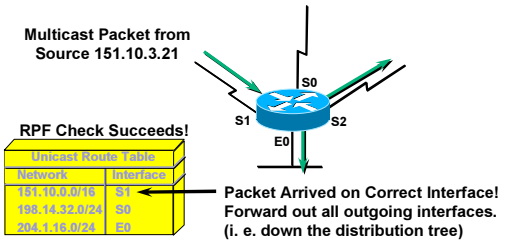
A closer look: RPF Check Fails



Multicast Forwarding

Cisco.com

A closer look: RPF Check Succeeds



Multicast Forwarding

Cisco.com

```
Go_Raiders#show ip rpf 151.10.3.21
RPF information for ? (151.10.3.21)
  RPF interface: Serial1
  RPF neighbor: ? (192.10.2.1)
  RPF route/mask: 151.10.0.0/16
  RPF type: unicast (eigrp 1)
  RPF recursion count: 0
  Doing distance-preferred lookups across tables
```

Multicast Forwarding

Cisco.com

```
Beat_Buc's> 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:
    Serial0, Forward/Sparse, 00:13:28/00:02:32
    Serial1, Forward/Sparse, 00:4:52/00:02:08

(151.10.3.21/32, 224.1.1.1), 00:01:43/00:02:59, flags: CJT
  Incoming interface: Serial1, RPF nbr 192.10.2.1
  Outgoing interface list:
    Ethernet0, Forward/Sparse, 00:01:43/00:02:11
    Serial2, forward/Sparse, 00:01:43/00:02:11
```


Agenda

Cisco.com

- Multicast Addressing
- Internet Group Management Protocol (IGMP)
- Multicast Forwarding
- PIM-Dense Mode (DM)
- PIM-Sparse Mode (SM)
- PIM-Spource Specific Mode (SSM)
- PIM-Bidirectional (BIDIR)

Agenda

- PIM-DM Overview
- PIM-DM Protocol Mechanics

PIM Dense Mode Overview

Cisco.com

- Uses “Push” Model
  - Traffic is initially flooded to all PIM neighbors
  - Branches that don’t want data are pruned
- Multicast forwarding state is created by the arrival of data
- If the source goes inactive, the tree is torn down

PIM Dense Mode Overview

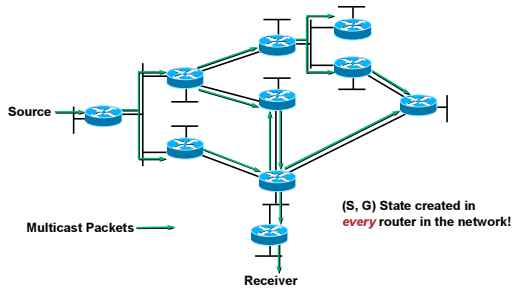
Cisco.com

- Grafts are used to join existing source tree
- Asserts are used to determine forwarder for multi-access LAN
- Prunes are sent on non-RPF P2P links
  - Asserts are sent on non-RPF multi-access links
- Rate-limited prunes are sent on all P2P links

PIM Dense Mode Overview

Cisco.com

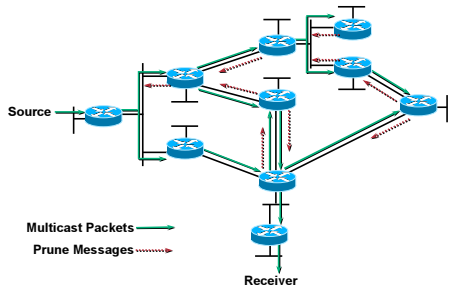
Initial Flooding



PIM Dense Mode Overview

Cisco.com

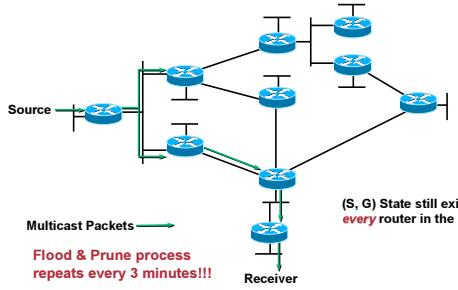
Pruning Unwanted Traffic



PIM Dense Mode Overview

Cisco.com

Results After Pruning



(S, G) State still exists in *every* router in the network!

Flood & Prune process repeats every 3 minutes!!!

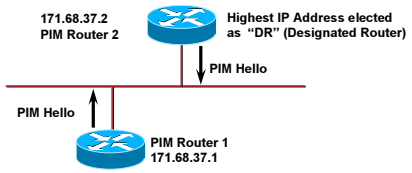
PIM-DM Protocol Mechanics

Cisco.com

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

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.

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  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 Protocol Mechanics

Cisco.com

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

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

PIM-DM 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), 00:00:10/00:00:00, 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 (*,G) State Rules

Cisco.com

- (*,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

Cisco.com

- (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" intfc timers never expire
  - "Pruned" intfc timers expire in 3 minutes

PIM-DM OIL Rules

Cisco.com

- (*,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 Protocol Mechanics

Cisco.com

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

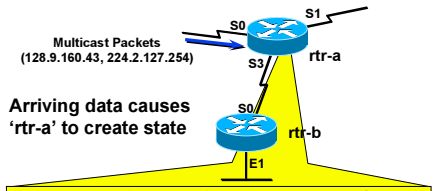
PIM-DM Forwarding Rules

Cisco.com

- 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

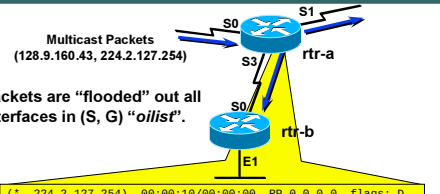
Cisco.com




```
(*, 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 Forwarding

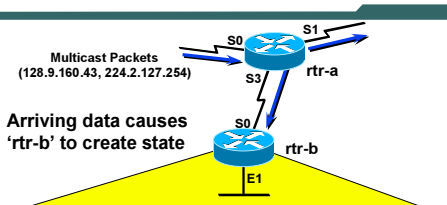
Cisco.com




```
(*, 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 Forwarding

Cisco.com




```
(*, 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

(128.9.160.43/32, 224.2.127.254), 00:00:12/00:02:48, flags: PT
  Incoming interface: Serial0, RPF nbr 198.92.2.31
  Outgoing interface list: Null
```

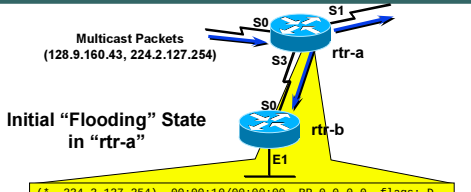

PIM-DM Protocol Mechanics

Cisco.com

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

PIM DM Pruning

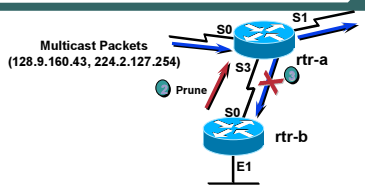
Cisco.com




```
(*, 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

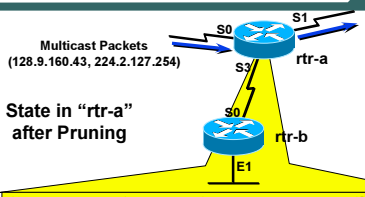
Cisco.com



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

PIM DM Pruning

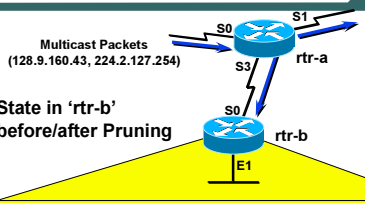
Cisco.com




```
(*, 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
    Serial1, Forward/Dense, 00:00:12/00:00:00
    Serial3, Forward/Dense, 00:00:12/00:00:00

(128.9.160.43/32, 224.2.127.254), 00:00:12/00:02:48, flags: T
  Incoming interface: Serial0, RPF nbr 198.92.1.129
  Outgoing interface list:
    Serial1, Forward/Dense, 00:00:12/00:00:00
    Serial3, Prune/Dense, 00:00:12/00:02:56
```


PIM DM Pruning

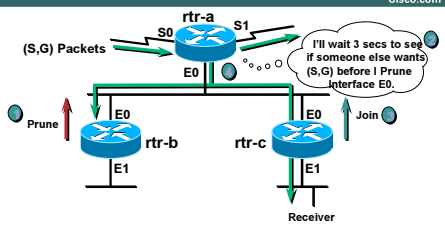
Cisco.com




```
(*, 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

(128.9.160.43/32, 224.2.127.254), 00:00:12/00:02:48, flags: PT
  Incoming interface: Serial0, RPF nbr 198.92.2.31
  Outgoing interface list: Null
```

Prune Delay on Multiaccess Networks

Cisco.com



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

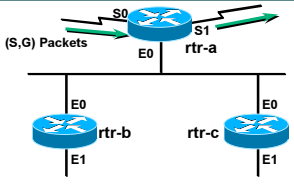
PIM-DM Protocol Mechanics

Cisco.com

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

PIM DM Grafting

Cisco.com

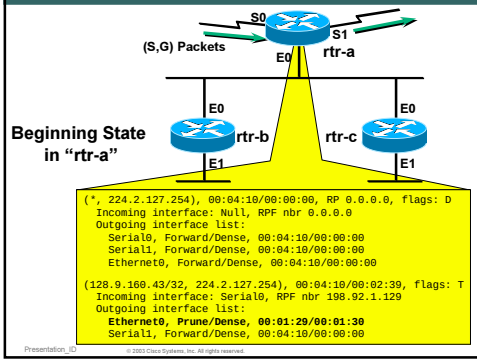


Beginning State

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

PIM DM Grafting

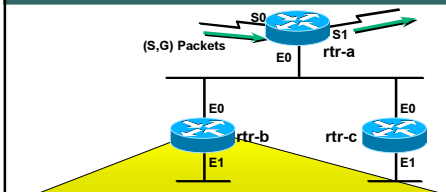
Cisco.com




```
(*, 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

Cisco.com

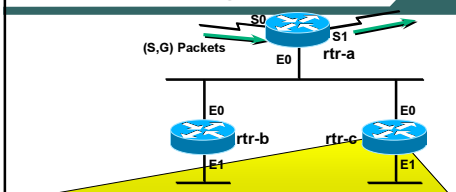



```
(*, 224.2.127.254), 00:04:10/00:00:00, RP 0.0.0.0, flags: D
  Incoming interface: Null, RPF nbr 0.0.0.0
  Outgoing interface list:
    Ethernet0, Forward/Dense, 00:04:10/00:00:00

(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

Cisco.com

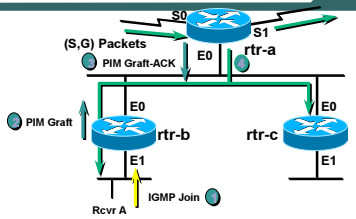



```
(*, 224.2.127.254), 00:04:10/00:00:00, RP 0.0.0.0, flags: D
  Incoming interface: Null, RPF nbr 0.0.0.0
  Outgoing interface list:
    Ethernet0, Forward/Dense, 00:04:10/00:00:00

(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

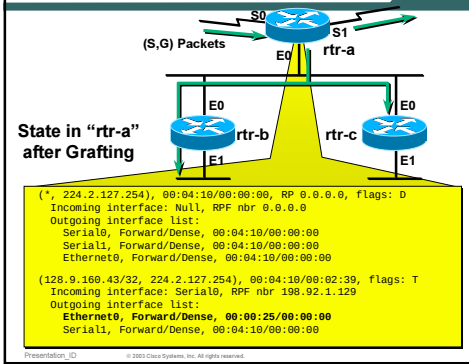
Cisco.com



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

PIM DM Grafting

Cisco.com




```
(*, 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
```

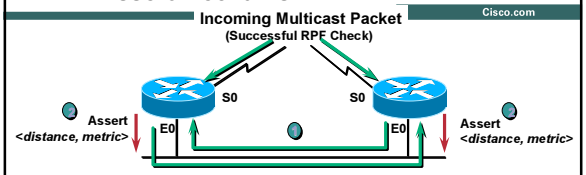

PIM-DM Protocol Mechanics

Cisco.com

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

PIM Assert Mechanism

Cisco.com



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

Agenda

Cisco.com

- Multicast Addressing
- Internet Group Management Protocol (IGMP)
- Multicast Forwarding
- PIM-Dense Mode (DM)
- PIM-Sparse Mode (SM)
- PIM-Spource Specific Mode (SSM)
- PIM-Bidirectional (BIDIR)

PIM-SM 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

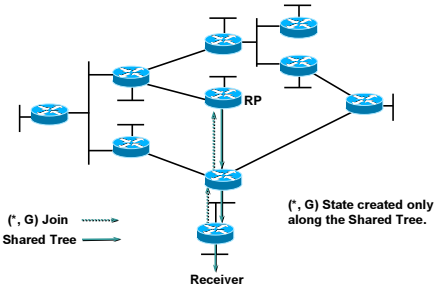
PIM-SM Overview

Cisco.com

- 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 2326 - "PIM Sparse Mode Protocol Spec"

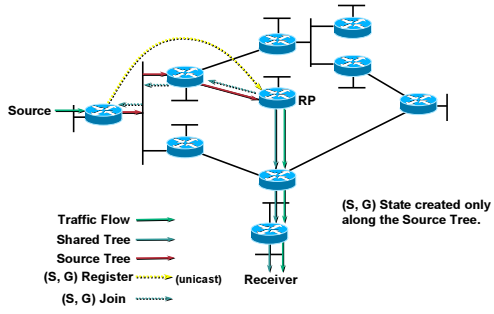
PIM-SM Shared Tree Join

Cisco.com



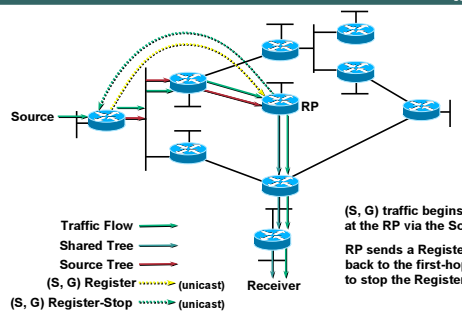
PIM-SM Sender Registration

Cisco.com



PIM-SM Sender Registration

Cisco.com

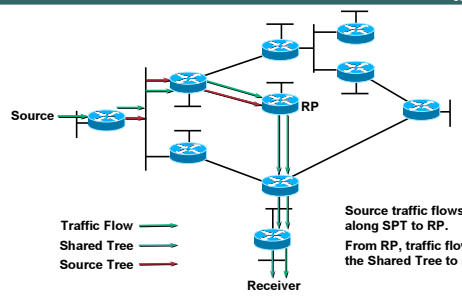


(S, G) traffic begins arriving at the RP via the Source tree.

RP sends a Register-Stop back to the first-hop router to stop the Register process.

PIM-SM Sender Registration

Cisco.com

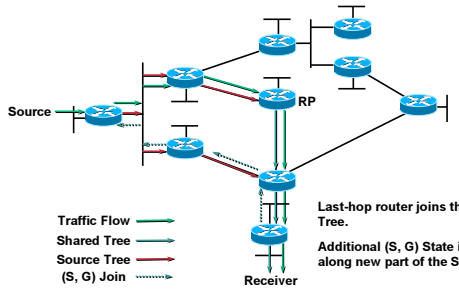


Source traffic flows natively along SPT to RP.

From RP, traffic flows down the Shared Tree to Receivers.

PIM-SM SPT Switchover

Cisco.com

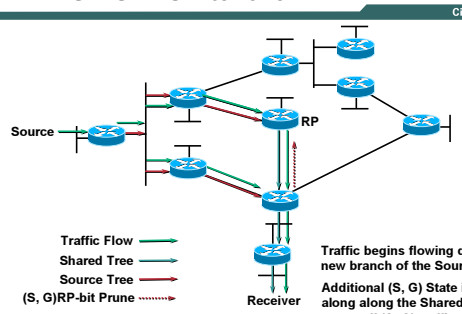


Last-hop router joins the Source Tree.

Additional (S, G) State is created along new part of the Source Tree.

PIM-SM SPT Switchover

Cisco.com



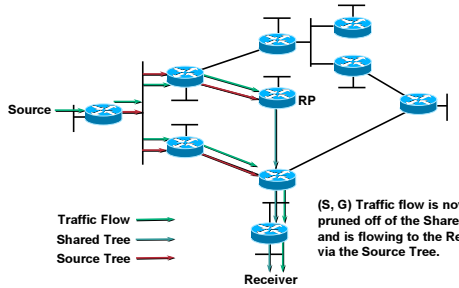
Traffic Flow
Shared Tree
Source Tree
(S, G)RP-bit Prune

Traffic begins flowing down the new branch of the Source Tree.

Additional (S, G) State is created along along the Shared Tree to prune off (S, G) traffic.

PIM-SM SPT Switchover

Cisco.com

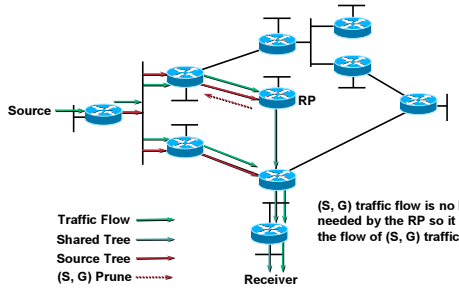


Traffic Flow
Shared Tree
Source Tree

(S, G) Traffic flow is now pruned off of the Shared Tree and is flowing to the Receiver via the Source Tree.

PIM-SM SPT Switchover

Cisco.com

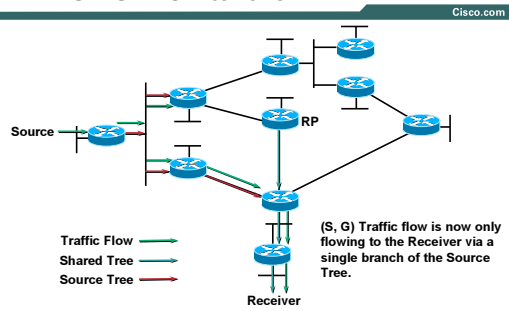


Traffic Flow
Shared Tree
Source Tree
(S, G) Prune

(S, G) traffic flow is no longer needed by the RP so it Prunes the flow of (S, G) traffic.

PIM-SM SPT Switchover

Cisco.com



(S, G) Traffic flow is now only flowing to the Receiver via a single branch of the Source Tree.

PIM-SM Protocol Mechanics

Cisco.com

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

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
```

PIM-SM (*,G) State Rules

Cisco.com

- (*,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

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

PIM-SM OIL Rules

Cisco.com

- 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 Protocol Mechanics

Cisco.com

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

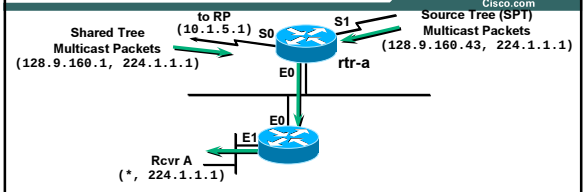
PIM-SM Forwarding Rules

Cisco.com

- 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

Cisco.com



- Packets are "forwarded" out all interfaces in "*oilist*".
- PIM Sparse mode interfaces are placed on the "*oilist*" 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 Protocol Mechanics

Cisco.com

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

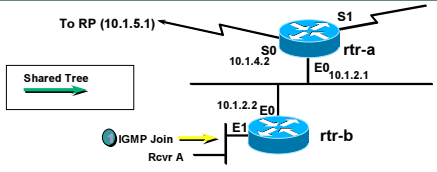
PIM SM Joining

Cisco.com

- Leaf routers send a (*,G) Join to toward RP
  - Joins sent hop-by-hop via unicast path toward RP
- Each router along path creates (*,G) state
  - IF no (*,G) state, create it & send a Join toward RP
  - ELSE Join process complete. Reached the (*,G) tree.

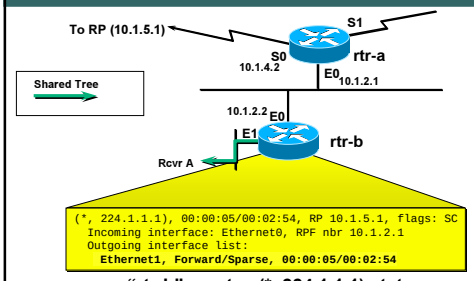
PIM SM Joining

Cisco.com



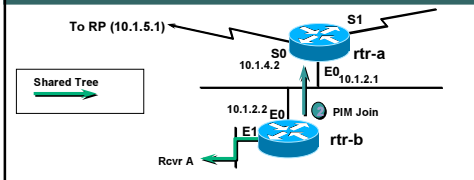
- "Rcvr A" wishes to receive group G traffic. Sends IGMP Join for G.

PIM SM Joining



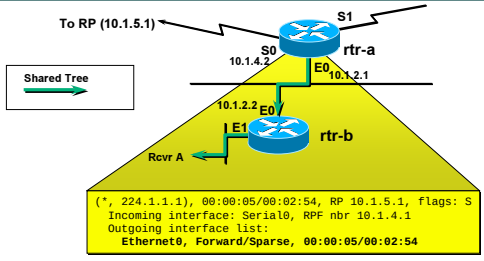
"rtr-b" creates (*, 224.1.1.1) state

PIM SM Joining



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

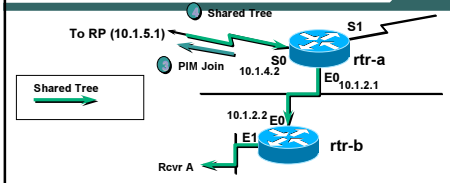
PIM SM Joining



"rtr-a" creates (*, 224.1.1.1) state.

PIM SM Joining

Cisco.com



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

PIM-SM Protocol Mechanics

Cisco.com

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

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 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

Cisco.com

- 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

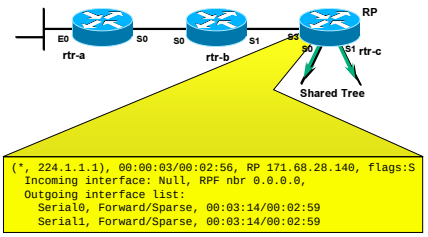
Cisco.com

- Receivers Join Group First
- Source Registers First

PIM SM Registering

Receiver Joins Group First

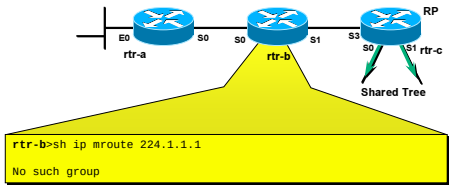
Cisco.com



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

PIM SM Registering
Receiver Joins Group First

Cisco.com

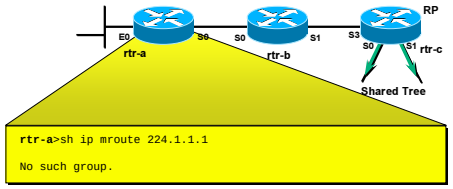



```
rtr-b>sh ip mroute 224.1.1.1

No such group
```

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

PIM SM Registering
Receiver Joins Group First

Cisco.com

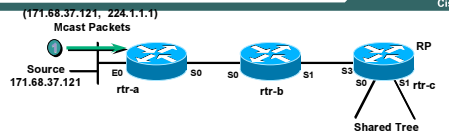



```
rtr-a>sh ip mroute 224.1.1.1

No such group.
```

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

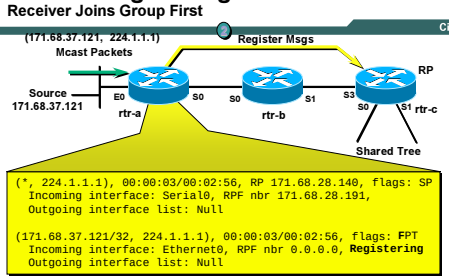
PIM SM Registering
Receiver Joins Group First

Cisco.com



- “Source” begins sending group G traffic.

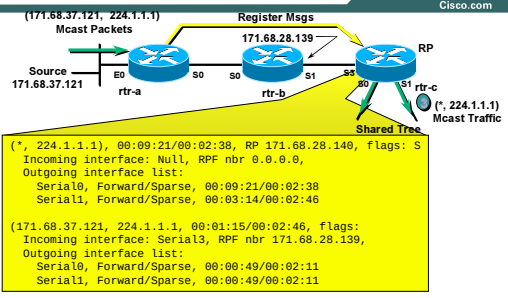
PIM SM Registering
Receiver Joins Group First



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

- "Source" begins sending group G traffic.
- "rtr-a" encapsulates packets in Registers; unicasts to RP.

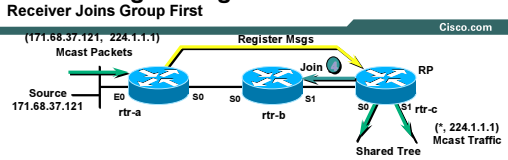
PIM SM Registering
Receiver Joins Group First



"RP" processes Register; creates (S, G) state

- "rtr-c" (RP) de-encapsulates packets; forwards down Shared tree.

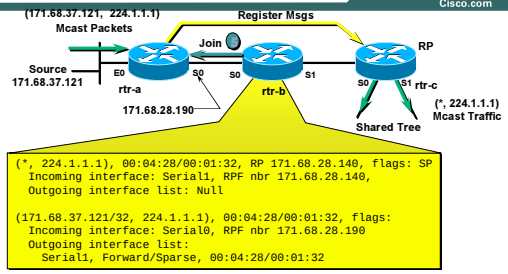
PIM SM Registering
Receiver Joins Group First



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

PIM SM Registering
Receiver Joins Group First

Cisco.com



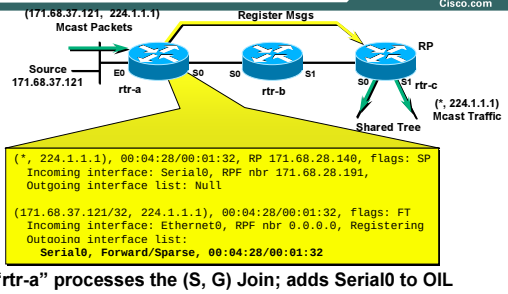

```
(*, 224.1.1.1), 00:04:28/00:01:32, RP 171.68.28.140, flags: SP
  Incoming interface: Serial1, RPF nbr 171.68.28.140,
  Outgoing interface list: Null

(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags:
  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)

- RP sends (S,G) Join toward Source to build SPT.
- "rtr-b" sends (S,G) Join toward Source to continue building SPT.

PIM SM Registering
Receiver Joins Group First

Cisco.com

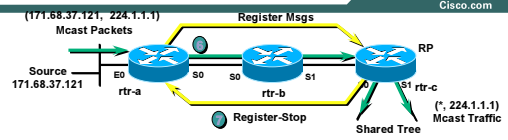



```
(*, 224.1.1.1), 00:04:28/00:01:32, RP 171.68.28.140, flags: SP
  Incoming interface: Serial0, RPF nbr 171.68.28.191,
  Outgoing interface list: Null

(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 Serial0 to OIL

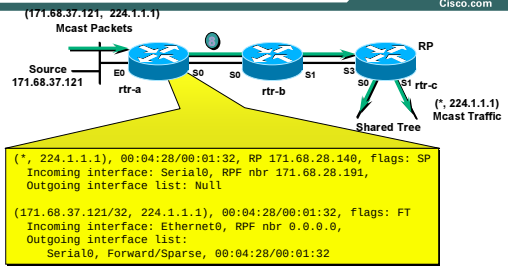
PIM SM Registering
Receiver Joins Group First

Cisco.com



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

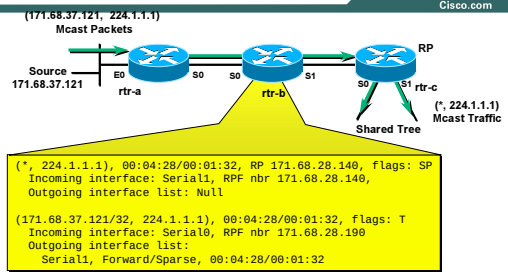
PIM SM Registering
Receiver Joins Group First



"rtr-a" stops sending Register messages
(Final State in "rtr-a")

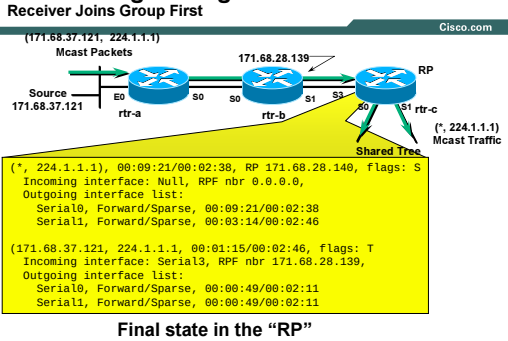
- **(S,G) Traffic now flowing down a single path (SPT) to RP.**

PIM SM Registering
Receiver Joins Group First



Final state in "rtr-b"

PIM SM Registering
Receiver Joins Group First



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

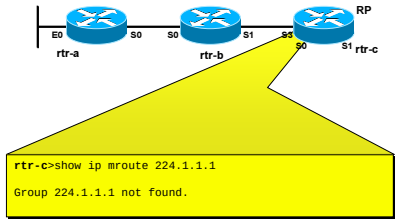
PIM SM Register Examples

Cisco.com

- Receivers Join Group First
- Source Registers First

PIM SM Registering
Source Registers First

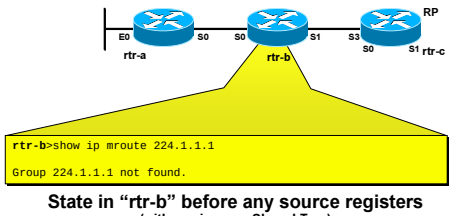
Cisco.com



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

PIM SM Registering
Source Registers First

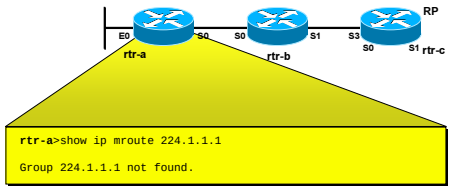
Cisco.com



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

PIM SM Registering
Source Registers First

Cisco.com

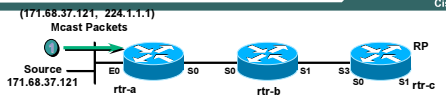



```
rtr-a>show ip mroute 224.1.1.1

Group 224.1.1.1 not found.
```

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

PIM SM Registering
Source Registers First

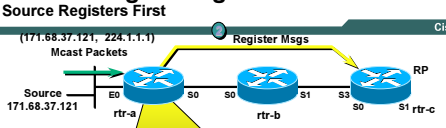
Cisco.com



- "Source" begins sending group G traffic.

PIM SM Registering
Source Registers First

Cisco.com

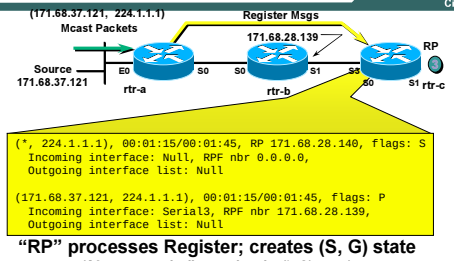



```
(*, 224.1.1.1), 00:00:03/00:02:56, RP 171.68.28.140, flags: SP
  Incoming interface: Serial0, RPF nbr 171.68.28.191,
  Outgoing interface list: Null

(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)

- "Source" begins sending group G traffic.
- "rtr-a" encapsulates packets in Registers; unicasts to RP.

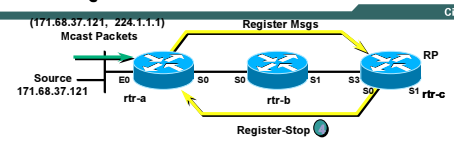
PIM SM Registering
Source Registers First



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

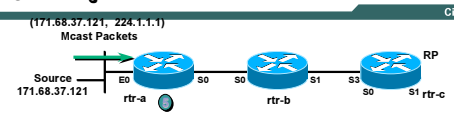
- "rtr-c" (RP) has no receivers on Shared Tree; discards packet.

PIM SM Registering
Source Registers First



- "rtr-c" (RP) has no receivers on Shared Tree; discards packet.
- RP sends "Register-Stop" to "rtr-a".

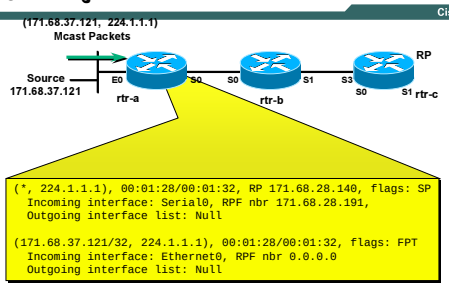
PIM SM Registering
Source Registers First



- "rtr-c" (RP) has no receivers on Shared Tree; discards packet.
- RP sends "Register-Stop" to "rtr-a".
- "rtr-a" stops encapsulating traffic in Register Messages; drops packets from Source.

PIM SM Registering
Source Registers First

Cisco.com

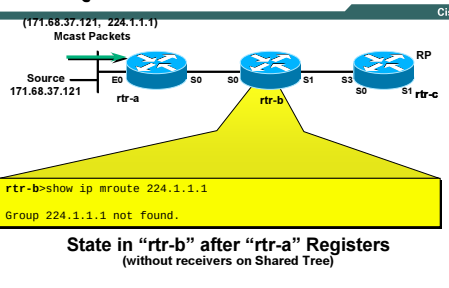



```
(*, 224.1.1.1), 00:01:28/00:01:32, RP 171.68.28.140, flags: SP
  Incoming interface: Serial0, RPF nbr 171.68.28.191,
  Outgoing interface list: Null

(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)

PIM SM Registering
Source Registers First

Cisco.com

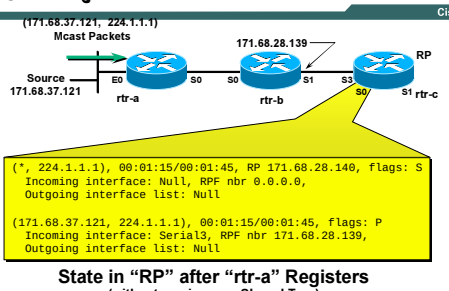



```
rtr-b>show ip mroute 224.1.1.1

Group 224.1.1.1 not found.
```

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

PIM SM Registering
Source Registers First

Cisco.com

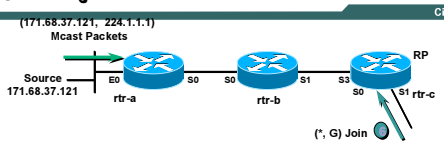



```
(*, 224.1.1.1), 00:01:15/00:01:45, RP 171.68.28.140, flags: S
  Incoming interface: Null, RPF nbr 0.0.0.0,
  Outgoing interface list: Null

(171.68.37.121, 224.1.1.1), 00:01:15/00:01:45, flags: 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)

PIM SM Registering
Source Registers First

Cisco.com

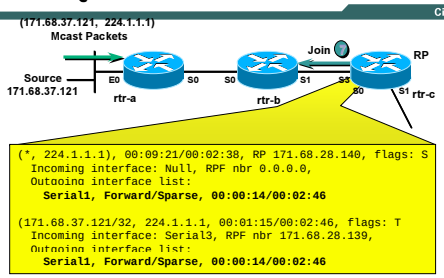


Receivers begin joining the Shared Tree

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

PIM SM Registering
Source Registers First

Cisco.com

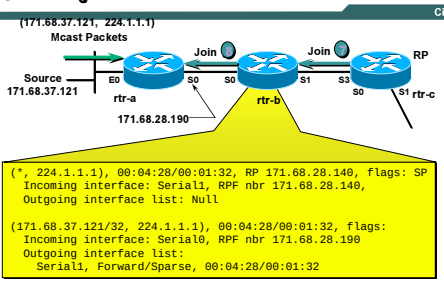


"RP" processes (*,G) Join
(Adds Serial1 to Outgoing Interface Lists)

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

PIM SM Registering
Source Registers First

Cisco.com

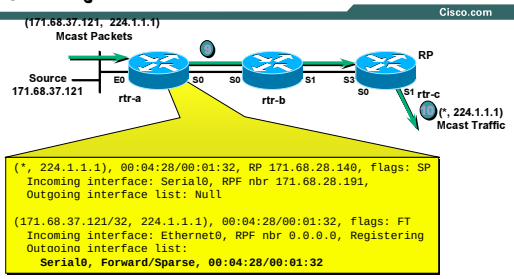


"rtr-b" processes Join, creates (S, G) state
(After automatically creating the (*, G) entry)

- RP sends (S,G) Joins for all known Sources in Group.
- "rtr-b" sends (S,G) Join toward Source to continue building SPT.

PIM SM Registering
Source Registers First

Cisco.com



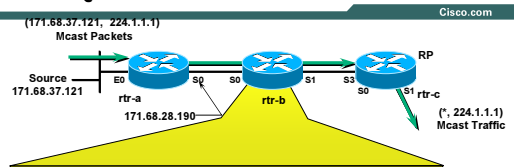

```
(*, 224.1.1.1), 00:04:28/00:01:32, RP 171.68.28.140, flags: SP
  Incoming interface: Serial0, RPF nbr 171.68.28.191,
  Outgoing interface list: Null

(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, Registering
  Outgoing interface list:
    Serial0, Forward/Sparse, 00:04:28/00:01:32
```

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

- RP begins receiving (S,G) traffic down SPT.
- RP forwards (S,G) traffic down Shared Tree to receivers.

PIM SM Registering
Source Registers First

Cisco.com

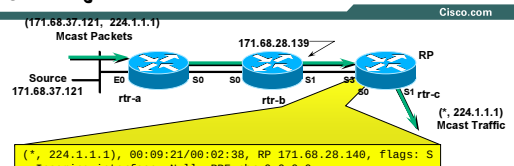



```
(*, 224.1.1.1), 00:04:28/00:01:32, RP 171.68.28.140, flags: SP
  Incoming interface: Serial1, RPF nbr 171.68.28.140,
  Outgoing interface list: Null

(171.68.37.121/32, 224.1.1.1), 00:04:28/00:01:32, flags: 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" after Receivers Join

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

PIM-SM Protocol Mechanics

Cisco.com

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

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.

PIM SM SPT-Switchover

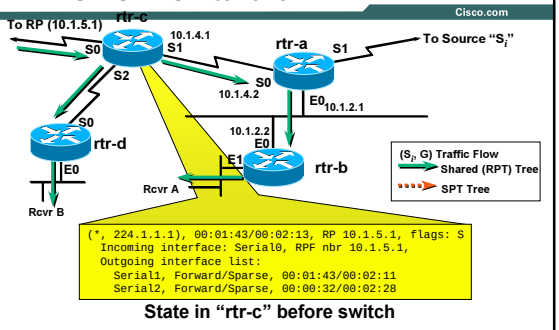
Cisco.com

SPT-Switchover Mechanism

Once each second
  Compute new (*, G) traffic rate
  If threshold exceeded, set "J" flag in (*, G)

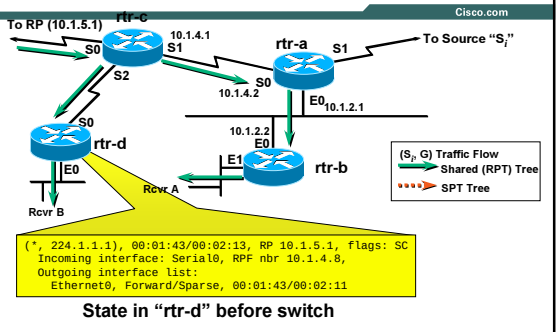
For each (S_i, G) packet received:
  If "J" flag set in (*, G)
    Join SPT for (S_i, G)
    Mark (S_i, G) entry with "J" flag
    Clear "J" flag in (*,G)

PIM SM SPT-Switchover



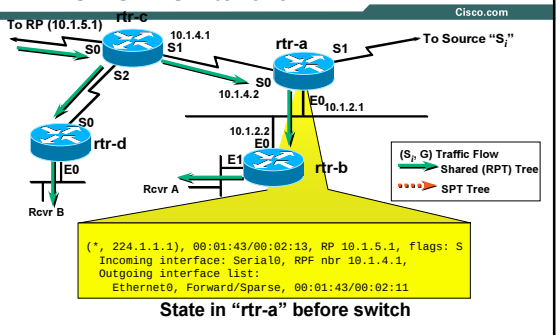
State in "rtr-c" before switch

PIM SM SPT-Switchover



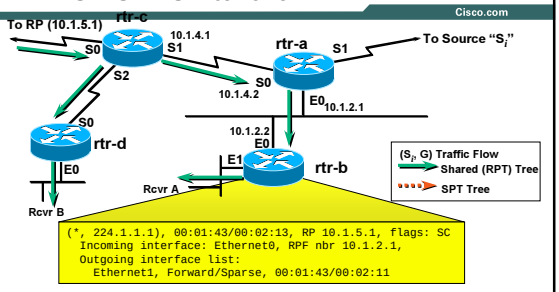
State in "rtr-d" before switch

PIM SM SPT-Switchover



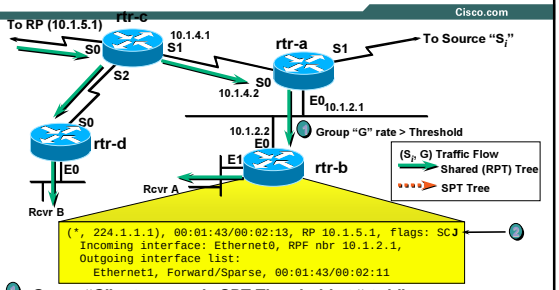
State in "rtr-a" before switch

PIM SM SPT-Switchover



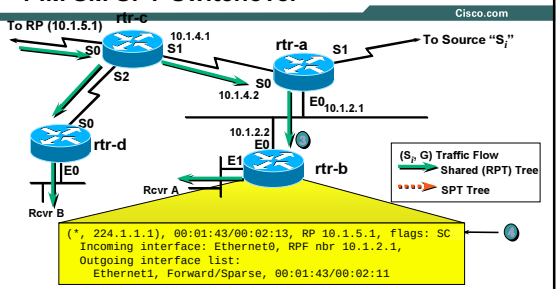
State in "rtr-b" before switch

PIM SM SPT-Switchover



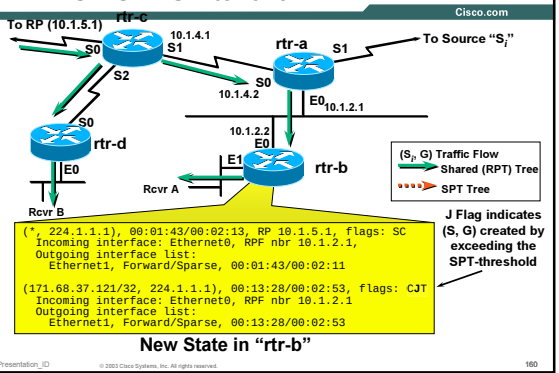
- Group "G" rate exceeds SPT Threshold at "rtr-b";
- Set J Flag in (*, G) and wait for next (S_i,G) packet.

PIM SM SPT-Switchover

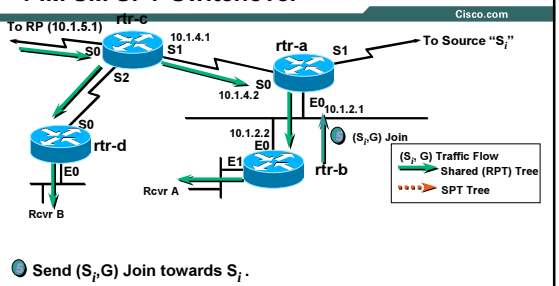


- (S_i,G) packet arrives down Shared tree.
- Clear J Flag in the (*,G) & create (S_i,G) state.

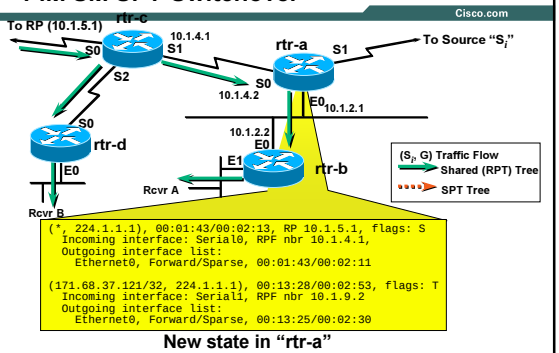
PIM SM SPT-Switchover



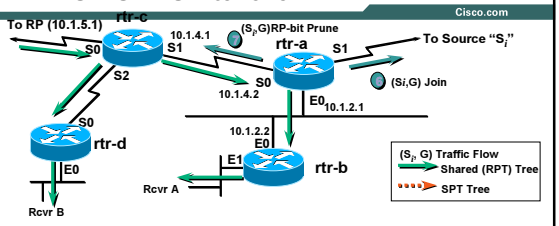
PIM SM SPT-Switchover



PIM SM SPT-Switchover

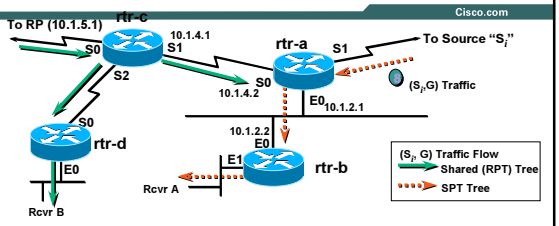


PIM SM SPT-Switchover



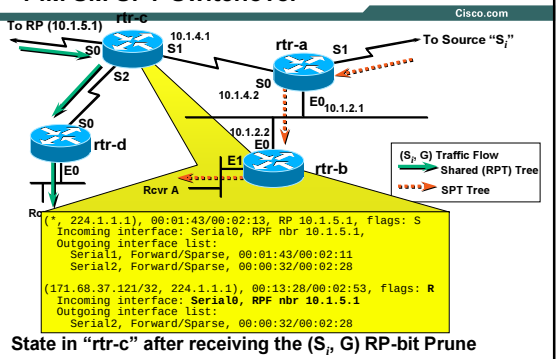
- "rtr-a" forwards (S_i,G) Join toward S_i.
- SPT & RPT diverge, triggering (S_i,G)RP-bit Prunes toward RP.

PIM SM SPT-Switchover



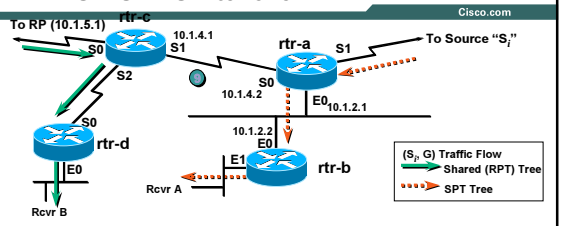
- (S_i, G) traffic begins flowing down SPT tree.

PIM SM SPT-Switchover



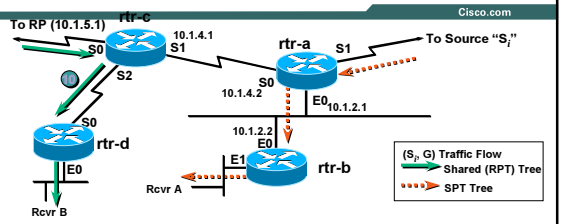
State in "rtr-c" after receiving the (S_i, G) RP-bit Prune

PIM SM SPT-Switchover



- Unnecessary (S_i, G) traffic is pruned from the Shared tree.

PIM SM SPT-Switchover



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

PIM SM SPT-Switchover

Shared Tree Switchback Mechanism

- Once each second when the (S,G) state is older than 1 minute
  - If "J" flag set in (Si , G) entry
    - Compute new (Si , G) traffic rate
    - If rate < SPT-threshold
      - Rejoin (*, G) Tree for (Si , G) traffic
      - Send (Si , G) prune up SPT toward Si
      - Delete (Si , G) entry

PIM-SM Protocol Mechanics

Cisco.com

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

PIM SM Pruning

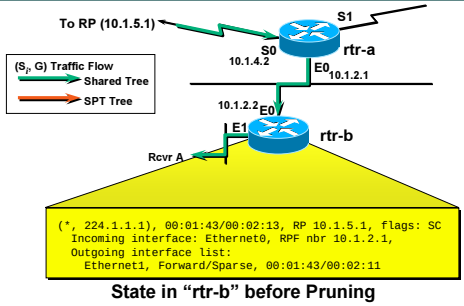
Cisco.com

- IGMP group times out / last host sends Leave
- Interface removed from all (*,G) and (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

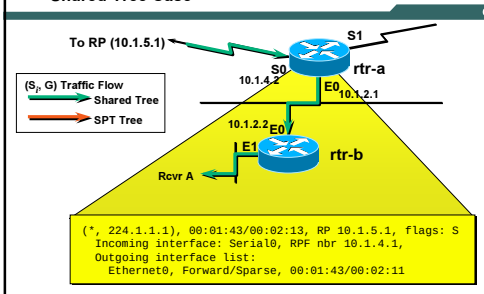
Cisco.com



State in "rtr-b" before Pruning

PIM SM Pruning
Shared Tree Case

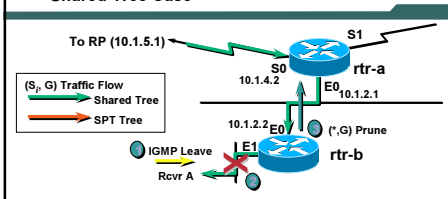
Cisco.com



State in "rtr-a" before Pruning

PIM SM Pruning
Shared Tree Case

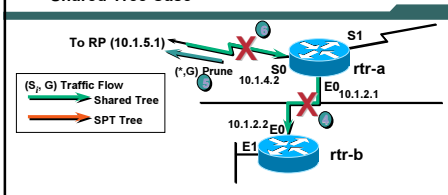
Cisco.com



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

PIM SM Pruning
Shared Tree Case

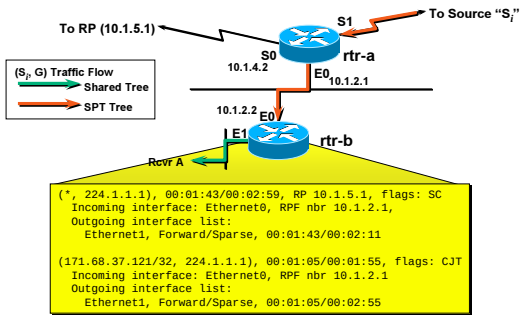
Cisco.com



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

PIM SM Pruning
Source (SPT) Case

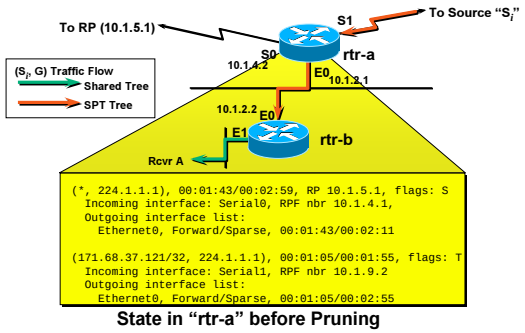
Cisco.com



State in "rtr-b" before Pruning

PIM SM Pruning
Source (SPT) Case

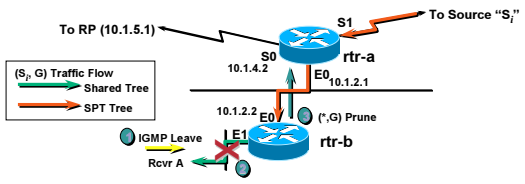
Cisco.com



State in "rtr-a" before Pruning

PIM SM Pruning
Source (SPT) Case

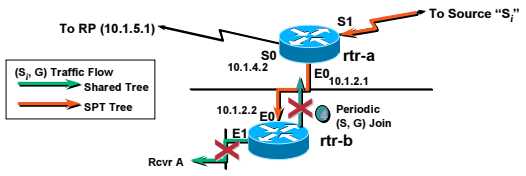
Cisco.com



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

PIM SM Pruning
Source (SPT) Case

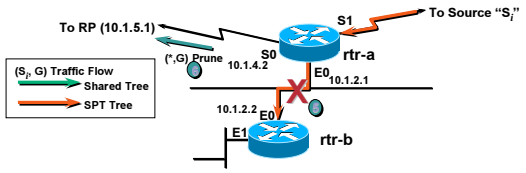
Cisco.com



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

PIM SM Pruning
Source (SPT) Case

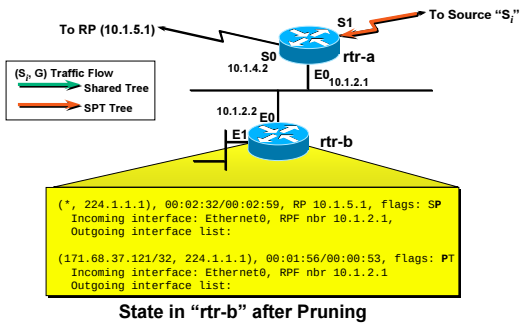
Cisco.com



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

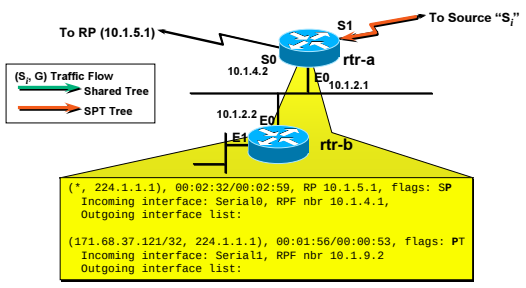
PIM SM Pruning
Source (SPT) Case

Cisco.com



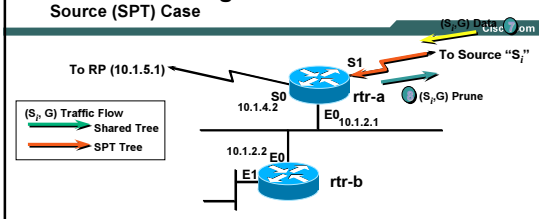
State in "rtr-b" after Pruning

PIM SM Pruning
Source (SPT) Case



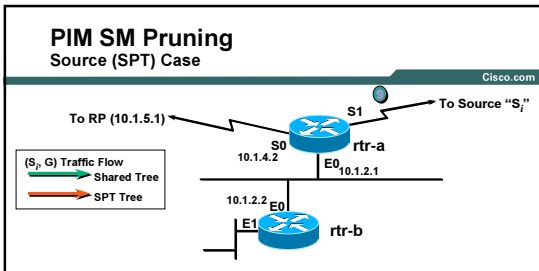
State in "rtr-a" after Pruning

PIM SM Pruning
Source (SPT) Case



- Another (S_i,G) data packet arrives via Serial1.
- 'rtr-a' responds by sending an (S_i,G) Prune toward source.

PIM SM Pruning
Source (SPT) Case



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

How does the network know about the RP ?

Cisco.com

- Static configuration
- AutoRP
- BSR

Static RP's

Cisco.com

- Hard-coded RP address
  - When used, must be configured on every router
  - All routers must have the same RP address
  - RP fail-over not possible
    - Exception: If Anycast RPs are used.
- Command
  - `ip pim rp-address <address> [group-list <acl>] [override]`
  - Optional group list specifies group range
    - Default: Range = 224.0.0.0/4
  - Override keyword "overrides" Auto-RP information
    - Default: Auto-RP learned info takes precedence

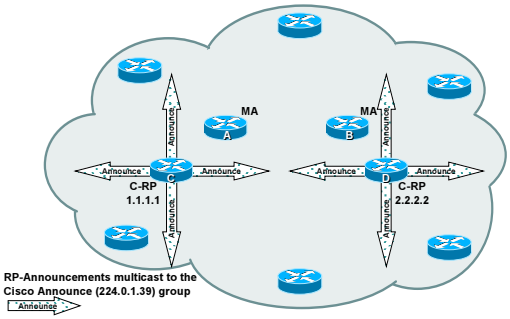
Auto-RP Overview

Cisco.com

- All routers automatically learn RP address
  - No configuration necessary except on:
    - Candidate RPs (ip pim send-rp-announce)
    - Mapping Agents (ip pim send-rp-discovery)
- Makes use of Multicast to distribute info
  - Two specially IANA assigned Groups used
    - Cisco-Announce - 224.0.1.39
    - Cisco-Discovery - 224.0.1.40
  - These groups normally operate in Dense mode
- Permits backup RP's to be configured
- Can be used with Admin-Scoping

Auto-RP—From 10,000 Feet

Cisco.com

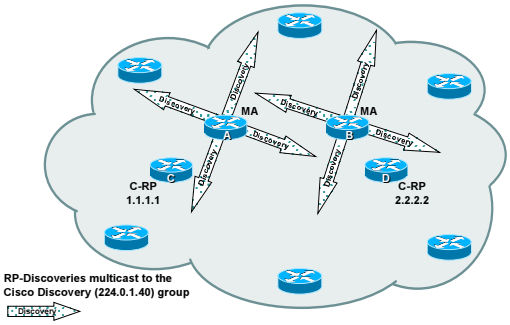


RP-Announcements multicast to the Cisco Announce (224.0.1.39) group

Announce

Auto-RP—From 10,000 Feet

Cisco.com



RP-Discoveries multicast to the Cisco Discovery (224.0.1.40) group

Discovery

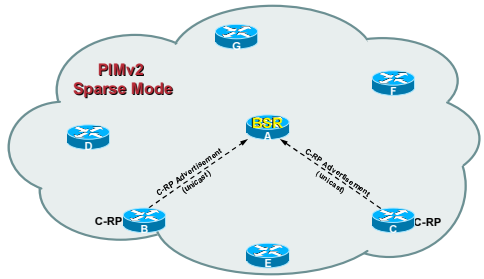
BSR Overview

Cisco.com

- A single Bootstrap Router (BSR) is elected
  - Multiple Candidate BSR's (C-BSR) can be configured
    - 'ip pim bsr-candidate'
  - C-RP's send C-RP announcements to the BSR
    - 'ip pim rp-candidate'
    - C-RP announcements are sent via unicast
    - BSR stores *ALL* C-RP announcements in the "RP-set"
  - BSR periodically sends BSR messages to all routers
    - BSR Messages contain entire RP-set and IP address of BSR
    - Messages are flooded hop-by-hop throughout the network away from the BSR
  - All routers select the RP from the RP-set
    - All routers use the same selection algorithm; select same RP
- BSR cannot be used with Admin-Scoping

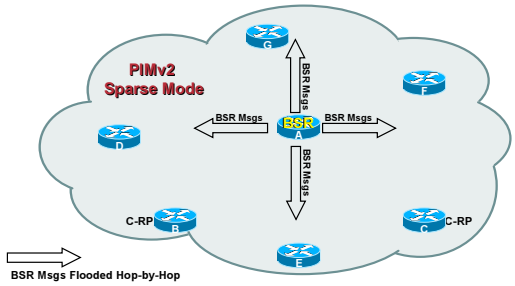
BSR—From 10,000 feet

Cisco.com



BSR—From 10,000 feet

Cisco.com



Agenda

Cisco.com

- Multicast Addressing
- Internet Group Management Protocol (IGMP)
- Multicast Forwarding
- PIM-Dense Mode (DM)
- PIM-Sparse Mode (SM)
- PIM-Spource Specific Mode (SSM)
- PIM-Bidirectional (BIDIR)

Source Specific Multicast

Cisco.com

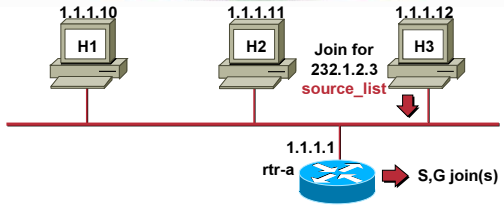
- Assume a One-to-Many Multicast Model.
  - Example: Video/Audio broadcasts, Stock Market data
- Why does PIM-SM need a Shared Tree?
  - So that hosts and 1st hop routers can learn who the active source is for the group.
- What if this was already known?
  - Hosts could use IGMPv3 to signal *exactly* which (S,G) SPT to join.
  - The Shared Tree & RP wouldn't be necessary.
  - Different sources could share the same Group address and not interfere with each other.
- Result: Source Specific Multicast (SSM)

Source Specific Multicast

Cisco.com

- SSM Advantages
  - Allows immediate use of shortest forwarding path to a specific source, without need to create shared tree.
  - Eliminates dependence on MSDP for finding sources.
  - Simplifies address allocation for global, single source groups when combined with elimination of shared trees.

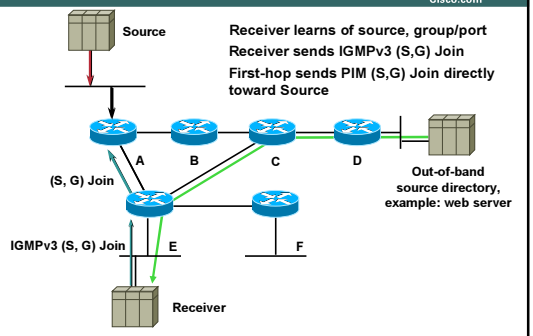
IGMPv3



- Host sends IGMPv3 join for group which can specify a list of sources to explicity include.
- Router adds membership
- Router sends S,G join directly to sources in the source_list, and is not required to send *,G join to RP (and must not in 232/8).

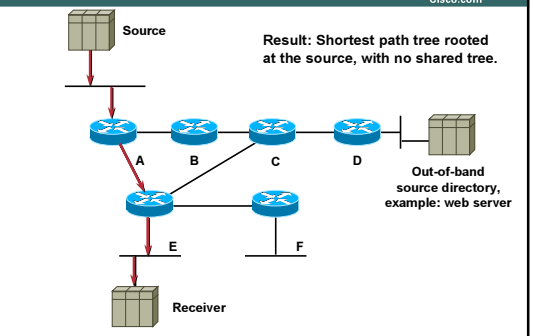
PIM Source Specific Mode

Cisco.com



PIM Source Specific Mode

Cisco.com



SSM Configuration

Cisco.com

- Global command
  - `ip pim ssm {default | <acl>}`
  - Defines SSM address range
    - Default range = 232.0.0.0/8
    - Use ACL for other ranges
  - Prevents Shared Tree Creation
    - (*, G) Joins never sent or processed
    - PIM Registers never sent or processed
  - Available beginning in IOS versions
    - 12.1(5)T, 12.2, 12.0(15)S, 12.1(8)E

SSM – Summary

Cisco.com

- Uses Source Trees only.
  - Hosts are responsible for source & group discovery.
  - Hosts must signal router which (S,G) to join.
- Solves multicast address allocation problems.
  - Flows differentiated by ***both*** source and group.
  - Content providers can use same group ranges.
    - Since each (S,G) flow is unique.
- Helps prevent certain DoS attacks
  - "Bogus" source traffic:
    - Can't consume network bandwidth.
    - Not received by host application.

Host Signalling
Bootstrap solutions (1)

Cisco.com

- Cisco IOS value added IP SSM bootstrap solutions

  URD: (URL Rendezvous Directory)
  Enable existing receiver applications for IP SSM via the web.

  IGMP v3lite:
  Provide for a partial IGMPv3 API on IGMPv1/v2 hosts.
  Enable to write and run IP SSM applications NOW
- Common idea of URD and IGMP v3lite:

  Pre:
  No IGMPv3 kernel support, an application CAN ONLY
  use IGMPv1/v2 group memberships or it will not get any traffic.

  Idea:
  Generate a (S,G) channel subscription (somehow) in addition to the IGMPv1/v2 membership that MUST already come from the kernel of the applications host.
  … and let the router figure out the right thing to do …

Host Signalling
How URD works

Cisco.com

Host Signalling
How URD works

Cisco.com

Users Desktop

Users favourite Browser

Please select you TV program:

Click here for the Movie

Click

Click here for the Sports

Click here for the News

Http:/www.broadercast.com/sports.htm

Host Signalling
How URD works

Cisco.com

Users Desktop

Users favourite Browser

Thank you for choosing this Sports channel

Currently showing Euro 2000 Soccer live from Brussels

England : Germany 3 : 1

Min 89:00

Host Signalling
How URD works

Cisco.com

Users Desktop

Old streaming video receiver application. Does IP Multicast, but not IP SSM

Works fine if we don't try to run it in the SSM-Range

Users favourite Browser

Thank you for choosing this Sports channel

Currently showing Euro 2000 Soccer live from Brussels

England : Germany 3 : 1

Min 89:00

Host Signalling
How URD works

Cisco.com

Users Desktop

Old streaming video receiver application. Does IP Multicast, but not IP SSM

That is..
Unless some unwanted traffic disturbs the reception, maybe some DoS attack...

Users favourite Browser

Thank you for choosing this Sports channel

Currently showing
Euro 2000 Soccer
live from Brussels

England : Germany
3 : 1

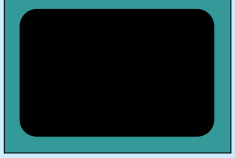
Min 89:00

Host Signalling
How URD works

Cisco.com

Users Desktop

Old streaming video receiver application. Does IP Multicast, but not IP SSM



Running the application on an SSM-Range alone does not help: The application will receive nothing!

Users favourite Browser

Thank you for choosing this Sports channel

Currently showing
Euro 2000 Soccer
live from Brussels

England : Germany
3 : 1

Min 89:00

Host Signalling
How URD works

Cisco.com

Users Desktop

Old streaming video receiver application. Does IP Multicast, but not IP SSM

But thanks to URD, the old application can run on an address in the SSM-Range and will only receive traffic from the right source!

Users favourite Browser

Thank you for choosing this Sports channel

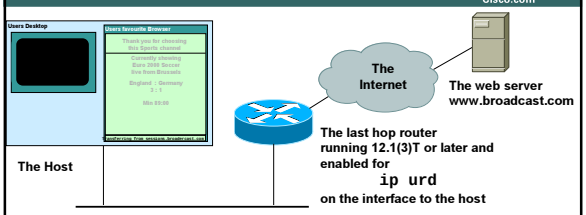
Currently showing
Euro 2000 Soccer
live from Brussels

England : Germany
3 : 1

Min 89:00

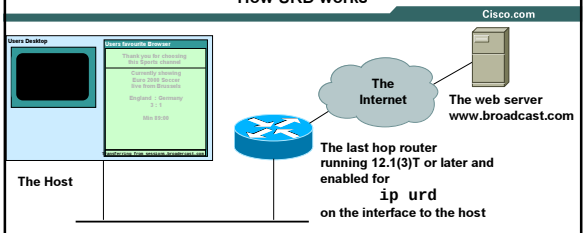
Retrieved URL String successfully

Host Signalling
How URD works



Let's zoom out a bit...

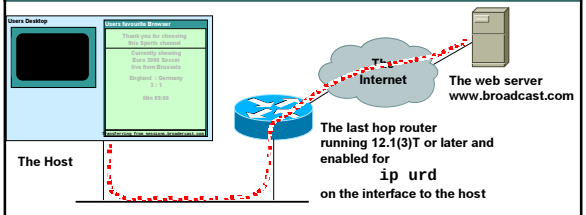
Host Signalling
How URD works



If the browser tries to retrieve the URL

```
http://www.broadercast.com:659/urd-
helper?group=232.3.4.5&source=192.44.81.5
```


Host Signalling
How URD works

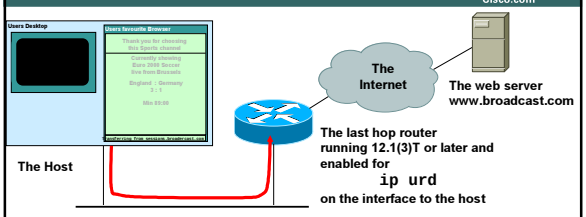


If the browser tries to retrieve the URL

```
http://www.broadercast.com:659/urd-
helper?group=232.3.4.5&source=192.44.81.5
```

Then it wants to open a TCP connection to www.broadercast.com, port 659

Host Signalling
How URD works

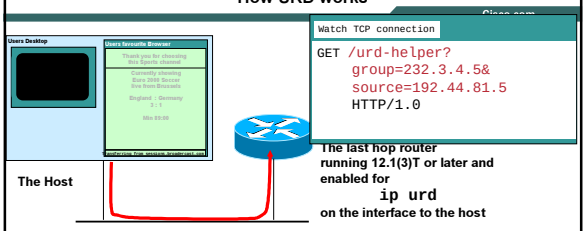


If the browser tries to retrieve the URL

```
http://www.broadercast.com:659/urd-
helper?group=232.3.4.5&source=192.44.81.5
```

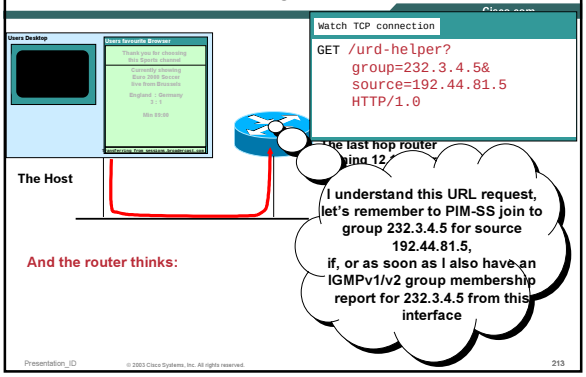
Then it wants to open a TCP connection to www.broadercast.com, port 659
But it only gets up to the first-hop router, who intercepts all TCP connections to port 659, whatever destination address they are for !

Host Signalling
How URD works



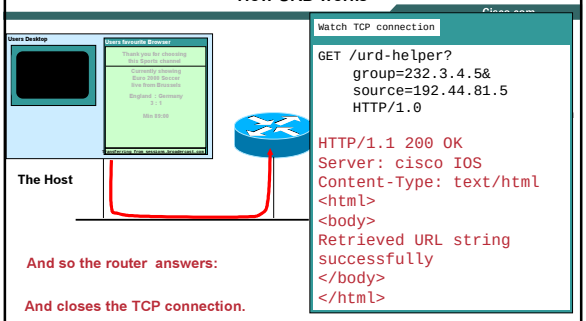
The router disgueses itself as a web server and listens to what the host want to have.

Host Signalling
How URD works



And the router thinks:

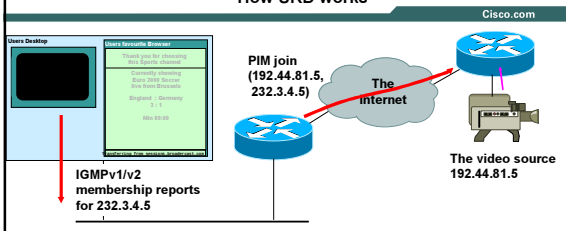
Host Signalling
How URD works



And so the router answers:

And closes the TCP connection.

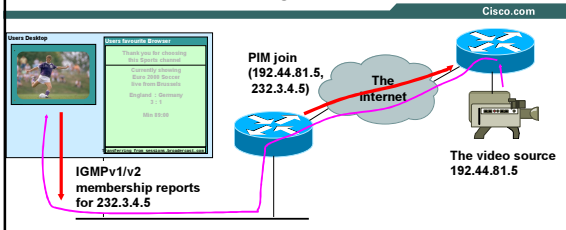
Host Signalling
How URD works



And once it sees the first IGMPv1/v2 report for the group (from the application), the router will join to the source via PIM-SS and continue as long as the IGMPv1/v2 group reports come in.

Note: The URL request from the browser and the first IGMPv1/2 report from the application may arrive in any order within ~ 1 minute

Host Signalling
How URD works



And finally the picture arrives and is being forwarded as long as the application runs and sends the IGMPv1/v2 membership reports

Host Signalling
How URD works

Cisco.com

Users Desktop

Old streaming video receiver application. Does IP Multicast, but not IP SSM

And all the user could notice, is the string returned by the router (may be hidden)!

Users favourite Browser

Thank you for choosing this Sports channel

Currently showing Euro 2000 Soccer live from Brussels

England : Germany 3 : 1

Min 89:00

<Retrieved URL String successfully>

Host Signalling
URD (1)

Cisco.com

- Enable IP SSM for existing applications
- Works with every browser that supports frames (or one click more for those without)
- No plugins required
  - Complete host platform independence
- Nothing to configure on the host
- 0.00 changes to the host
- URL easily added to WWW server HTML pages
  - No additional CGI scripts required.

Host Signalling
URD (2)

Cisco.com

- Supported in IOS 12.1(3)T and later.
- Supported in the process, fast and CEF paths
- Intercepting solely based on TCP port 659
  - If first hop router is not URD enabled, www-server may want to reply to HTTP on that port too (error discovery)
- Port 659 assigned by IANA for Cisco.
- URD - URL Rendezvous Directory
  - Name still carries the idea that it is also quite simple to write a CGI-Script to completely emulate an RP, I.e.: add web pages, where you would click onto if you are a source, and the script would then create the URD command URLs for the receivers.

Agenda

Cisco.com

- Multicast Addressing
- Internet Group Management Protocol (IGMP)
- Multicast Forwarding
- PIM-Dense Mode (DM)
- PIM-Sparse Mode (SM)
- PIM-Spource Specific Mode (SSM)
- PIM-Bidirectional (BIDIR)

Many-to-Any State Problem

Cisco.com

- Creates huge amounts of (S,G) state
  - –State maintenance workloads skyrocket
    - •High OIL fanouts make the problem worse
  - –Router performance begins to suffer
- Using Shared-Trees only
  - –Provides some (S,G) state reduction
    - •Results in (S,G) state only along SPT to RP
    - •Frequently still too much (S,G) state
    - •Need a solution that only uses (*,G) state

Eliminating (S,G) State

Cisco.com

- Bidirectional Shared-Trees
  - –Allows data to flow up the Shared Tree
    - •Source traffic follows Shared Tree to get to the RP and all other receivers on the Shared Tree
  - –Cannot use current (*,G) RPF rules
    - •Care must be taken to avoid multicast loops
  - –Requires a Designated Forwarder (DF)
    - •Responsible for forwarding traffic up Shared Tree
      - –DF's will accept data on the interfaces in their OIL.
      - –Then send it out all other interfaces. (Including the IIF.)

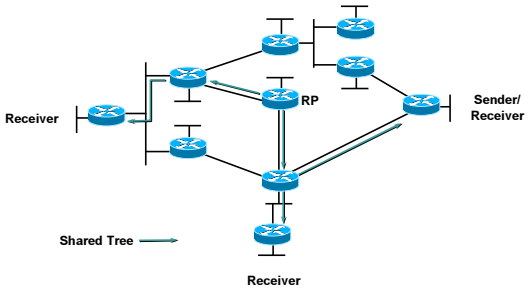
Bidirectional (Bidir) PIM

Cisco.com

- Idea:
  - Use the same tree for traffic from sources towards RP and from RP to receivers
- Benefits:
  - Less state in routers
    - Only (*, G) state is used
    - Source traffic follows the Shared Tree
      - Flows up the Shared Tree to reach the RP.
      - Flows down the Shared Tree to reach all other receivers.

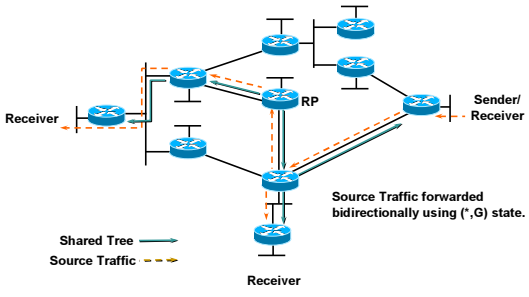
Bidirectional PIM – Overview

Cisco.com



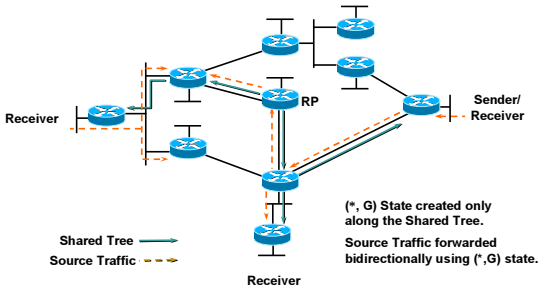
Bidirectional PIM – Overview

Cisco.com



Bidirectional PIM – Overview

Cisco.com



(*, G) State created only along the Shared Tree.

Source Traffic forwarded bidirectionally using (*,G) state.

PIM Modifications for Bidir Operation

Cisco.com

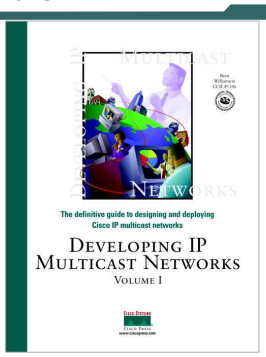
- Designated Forwarders (DF)
  - On each link the router with the best path to the RP is elected to be the DF.
    - Note: Designated Routers (DR) are not used for bidir groups.
  - The DF is responsible for forwarding traffic upstream towards the RP.
  - No special treatment is required for local sources.

Bidir PIM–Evaluation

Cisco.com

- Ideal for Many to Many applications
- Drastically reduces network mroute state.
  - Eliminates *ALL* (S,G) state in the network.
    - SPT's between sources to RP eliminated.
    - Source traffic flows both up and down Shared Tree.
  - Allows Many-to-Any applications to scale.
    - Permits virtually an unlimited number of sources.

If All Else Fails—RTFB[1]



[1] Read this fine book

CISCO SYSTEMS

EMPOWERING THE
INTERNET GENERATION℠