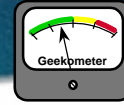


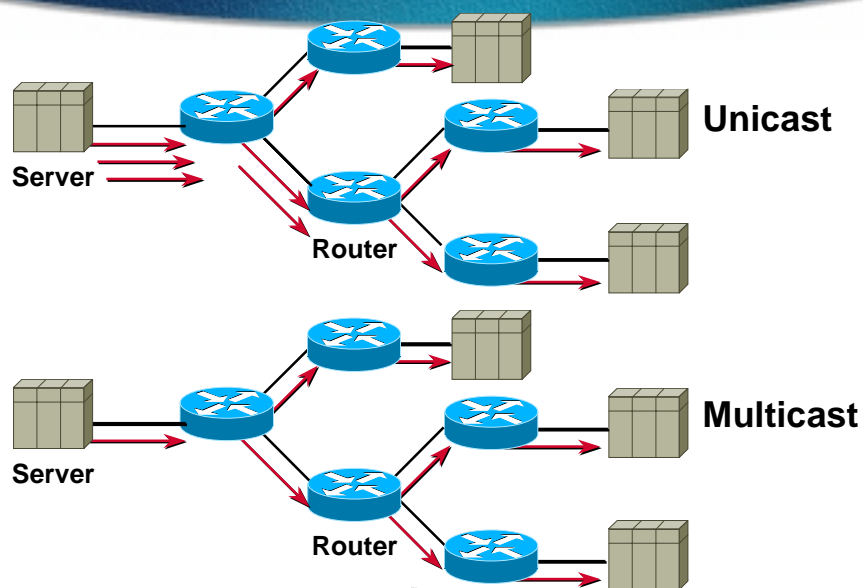
Agenda



- **Why Multicast?**
- **Multicast Addressing**
- **Host-Router Signaling: IGMP**
- **Multicast Distribution Trees**
- **Multicast Forwarding**
- **Multicast Routing Protocols**

www.cisco.com

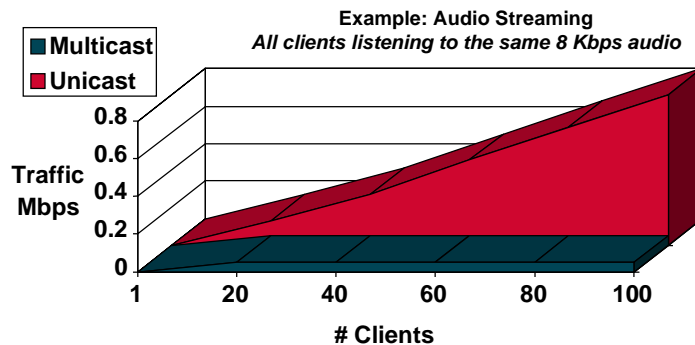
Unicast vs Multicast



www.cisco.com

Multicast Advantages

- **Enhanced Efficiency:** Controls network traffic and reduces server and CPU loads
- **Optimized Performance:** Eliminates traffic redundancy
- **Distributed Applications:** Makes multipoint applications possible



www.cisco.com

Multicast Disadvantages

Multicast is UDP Based!!!

- **Best Effort Delivery:** Drops are to be expected. Multicast applications should not expect reliable delivery of data and should be designed accordingly. Reliable Multicast is still an area for much research. Expect to see more developments in this area.
- **No Congestion Avoidance:** Lack of TCP windowing and “slow-start” mechanisms can result in network congestion. If possible, Multicast applications should attempt to detect and avoid congestion conditions.
- **Duplicates:** Some multicast protocol mechanisms (e.g. Asserts, Registers and SPT Transitions) result in the occasional generation of duplicate packets. Multicast applications should be designed to expect occasional duplicate packets.

www.cisco.com

Apps that Benefit from IP Multicast

- **Multimedia**
 - Streaming media
 - Training, corporate communications
 - Conferencing—video/audio
- **Data warehousing**
- **Financial applications**
- **Any one-to-many data push applications**

www.cisco.com

Agenda

- **Why Multicast?**
- **Multicast Addressing**
- **Host-Router Signaling: IGMP**
- **Multicast Distribution Trees**
- **Multicast Forwarding**
- **Multicast Routing Protocols**

www.cisco.com

Multicast Addressing

- **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

www.cisco.com

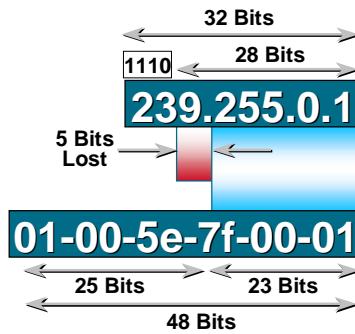
Multicast Addressing

- **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
 - Examples
 - Site-local scope: 239.253.0.0/16
 - Organization-local scope: 239.192.0.0/14

www.cisco.com

Multicast Addressing

IP Multicast MAC Address Mapping (FDDI and Ethernet)



www.cisco.com

Multicast Addressing

IP Multicast MAC Address Mapping (FDDI and Ethernet)

Be Aware of the 32:1 Address Overlap

32 - IP Multicast Addresses

224.1.1.1
224.129.1.1
225.1.1.1
225.129.1.1
⋮
238.1.1.1
238.129.1.1
239.1.1.1
239.129.1.1

1 - Multicast MAC Address
(FDDI and Ethernet)

0x0100.5E01.0101

www.cisco.com

Multicast Addressing

IP Multicast MAC Address Mapping (Token Ring)

A Layer 3 IPmc Address Maps to a single Token Ring Functional Address or the all ones' Broadcast address:

224.x.x.x

224.x.x.x

c0-00-00-04-00-00

(Shown in Token Ring, non-canonical format)

ff-ff-ff-ff-ff-ff

Results in high levels of unwanted interrupts for non-interested Hosts

www.cisco.com

Multicast Addressing

IP Multicast MAC Address Mapping (Token Ring)

Be Aware of the 268,435,200:1 Address Overlap

ALL 268,435,200 - IP Multicast Addresses

224.0.1.0
224.0.1.1
224.0.1.2
224.0.1.3
⋮
239.239.255.252
239.255.255.253
239.255.255.254
239.255.255.255

1 - Multicast MAC Address
(Token Ring)

0xFFFF.FFFF.FFFF

RUN AWAY!!!

www.cisco.com

Multicast Addressing

- **Dynamic Group Address Assignment**
 - Historically accomplished using SDR application
 - Sessions/groups announced over well-known multicast groups
 - Address collisions detected and resolved at session creation time
 - Has problems scaling
 - Future dynamic techniques under consideration
 - **Multicast Address Set-Claim (MASC)**
 - Hierarchical, dynamic address allocation scheme
 - Extremely complex garbage-collection problem.
 - Long ways off

www.cisco.com

Multicast Addressing

- **Static Group Address Assignment (new)**
 - 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
 - Defined in IETF draft
 - “draft-ietf-mboned-glop-addressing-00.txt”

www.cisco.com

Agenda

- Why Multicast?
- Multicast Addressing
- **Host-Router Signaling: IGMP**
- Multicast Distribution Trees
- Multicast Forwarding
- Multicast Routing Protocols

www.cisco.com

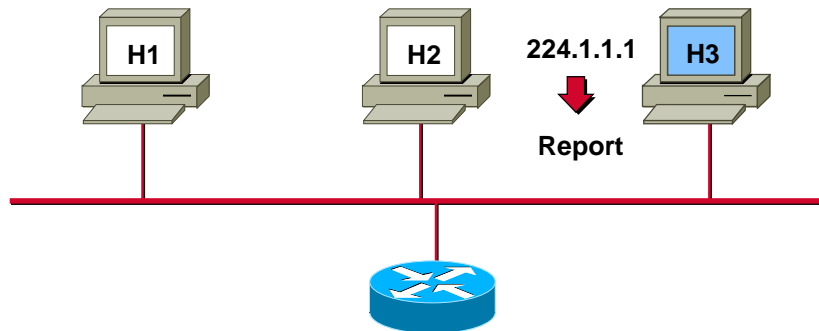
Host-Router Signaling: IGMP

- How hosts tell routers about group membership
- Routers solicit group membership from directly connected hosts
- RFC 1112 specifies first version of IGMP
- RFC 2236 specifies current version of IGMP
- Supported on UNIX systems, PCs, and MACs

www.cisco.com

Host-Router Signaling: IGMP

Joining a Group

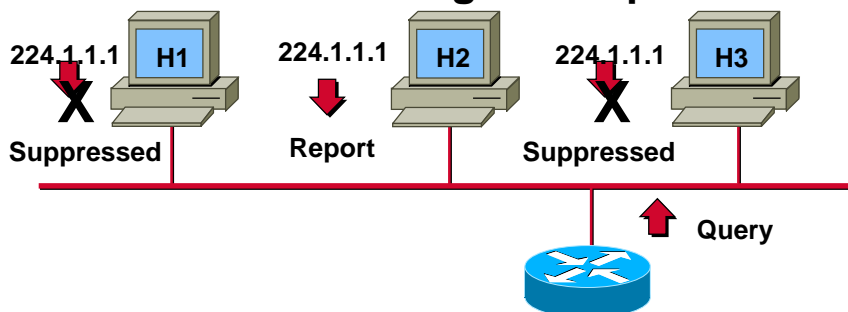


- Host sends IGMP Report to join group

www.cisco.com

Host-Router Signaling: IGMP

Maintaining a Group

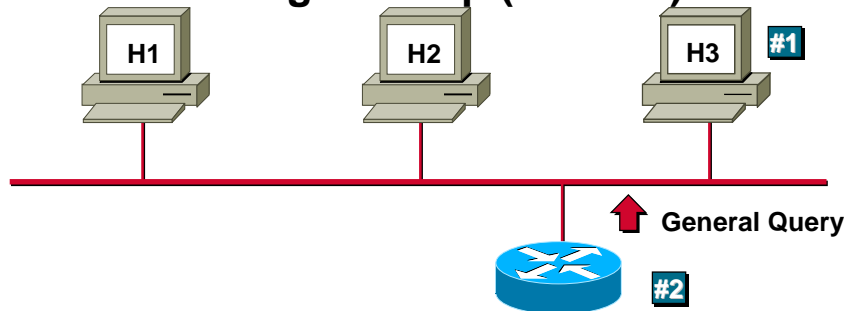


- Router sends periodic Queries to 224.0.0.1
- One member per group per subnet reports
- Other members suppress reports

www.cisco.com

Host-Router Signaling: IGMP

Leaving a Group (IGMPv1)

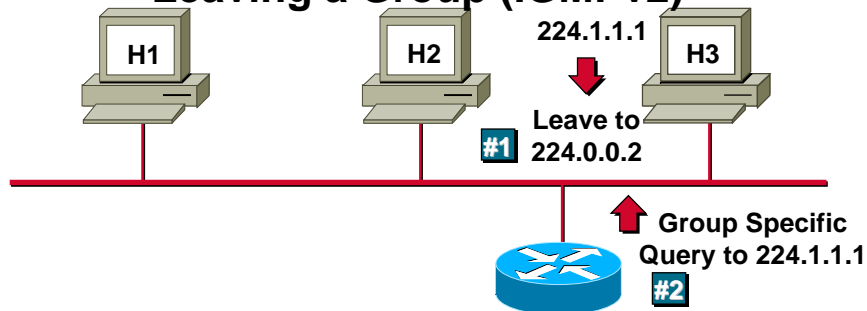


- Host quietly leaves group
- Router sends 3 General Queries (60 secs apart)
- No IGMP Report for the group is received
- Group times out (Worst case delay \sim 3 minutes)

www.cisco.com

Host-Router Signaling: IGMP

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 \sim 3 seconds
- Group 224.1.1.1 times out

www.cisco.com

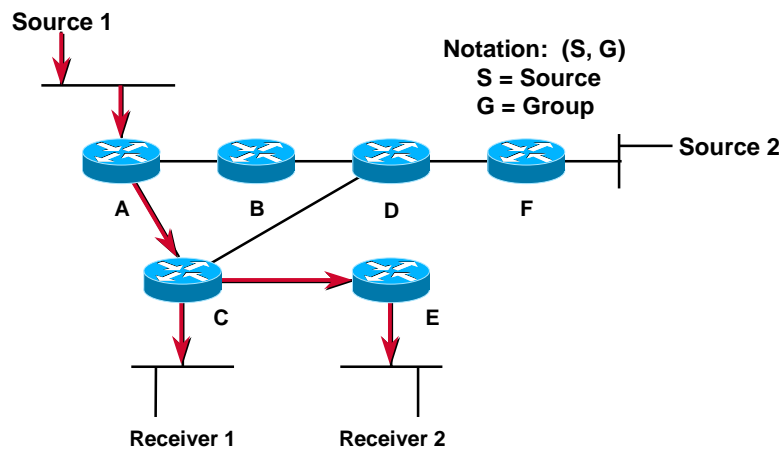
Agenda

- Why Multicast?
- Multicast Addressing
- Host-Router Signaling: IGMP
- **Multicast Distribution Trees**
- Multicast Forwarding
- Multicast Routing Protocols

www.cisco.com

Multicast Distribution Trees

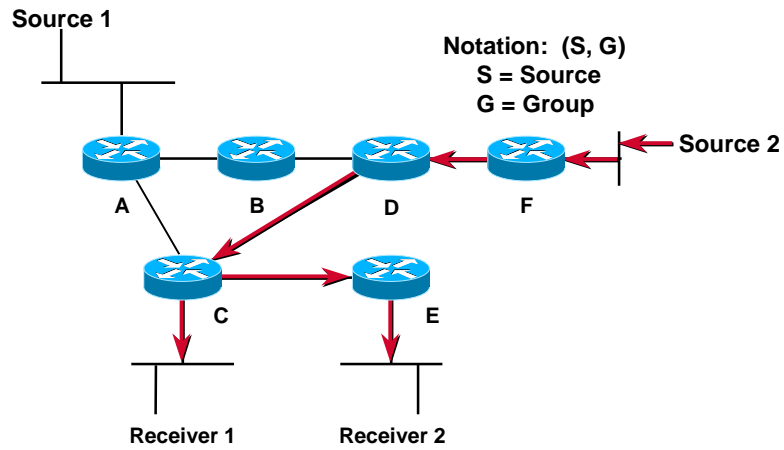
Shortest Path or Source Distribution Tree



www.cisco.com

Multicast Distribution Trees

Shortest Path or Source Distribution Tree

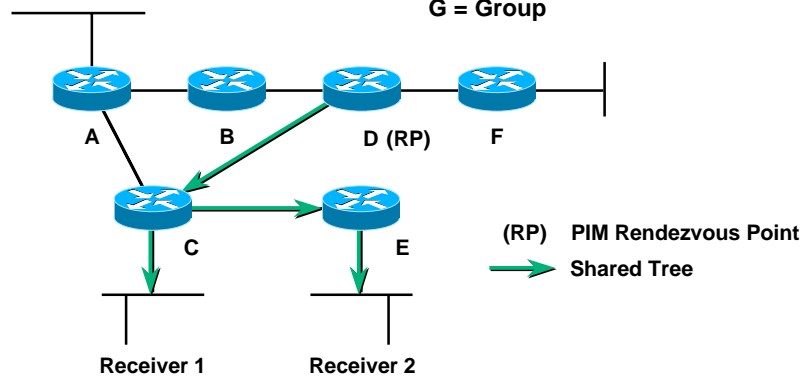


www.cisco.com

Multicast Distribution Trees

Shared Distribution Tree

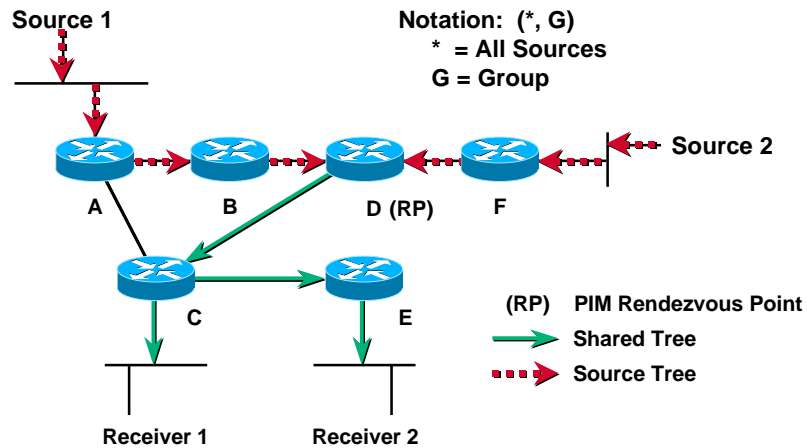
Notation: (*, G)
* = All Sources
G = Group



www.cisco.com

Multicast Distribution Trees

Shared Distribution Tree



www.cisco.com

Multicast Distribution Trees

Characteristics of Distribution Trees

- **Source or Shortest Path trees**
Uses more memory $O(S \times G)$ 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

www.cisco.com

Agenda

- Why Multicast?
- Multicast Addressing
- Host-Router Signaling: IGMP
- Multicast Distribution Trees
- **Multicast Forwarding**
- Multicast Routing Protocols

www.cisco.com

Multicast Forwarding

- **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”**

www.cisco.com

Multicast Forwarding

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.

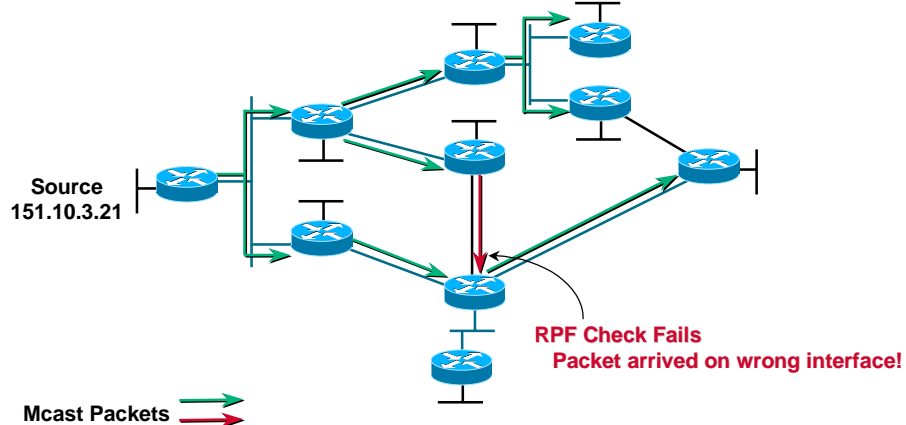
If the datagram arrived on the interface specified in the routing table for the source address; then the RPF check succeeds.

Otherwise, the RPF Check fails.

www.cisco.com

Multicast Forwarding

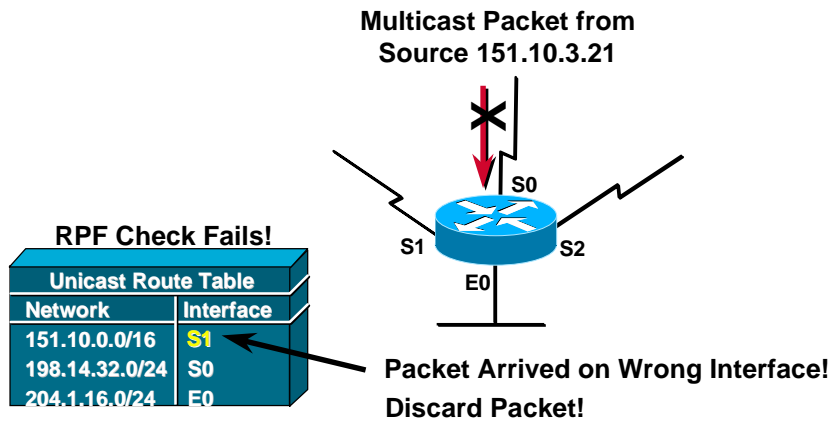
Example: RPF Checking



www.cisco.com

Multicast Forwarding

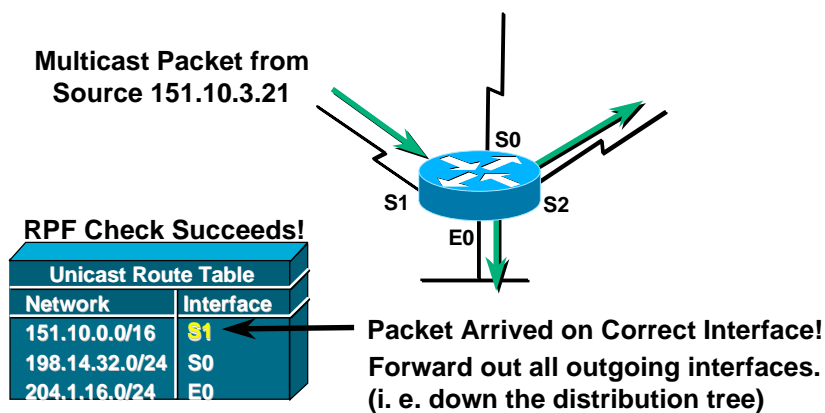
A closer look: RPF Check Fails



www.cisco.com

Multicast Forwarding

A closer look: RPF Check Succeeds



www.cisco.com

Agenda

- Why Multicast?
- Multicast Addressing
- Host-Router Signaling: IGMP
- Multicast Distribution Trees
- Multicast Forwarding
- **Multicast Routing Protocols**

www.cisco.com

Multicast vs. Unicast Routing

Multicast Routing is not unicast routing. You have to think of it differently. It is not like OSPF. It is not like RIP. It is not like anything you may be familiar with.

www.cisco.com

Types of Multicast Protocols

- **Dense-mode**
 - Uses “Push” Model
 - Traffic Flooded throughout network
 - Pruned back where it is unwanted
 - Flood & Prune behavior (typically every 3 minutes)
- **Sparse-mode**
 - Uses “Pull” Model
 - Traffic sent only to where it is requested
 - Explicit Join behavior

www.cisco.com

Multicast Protocol Overview

- **Currently, there are four multicast routing protocols:**
 - **DVMRPv3 (Internet-draft)**
 - DVMRPv1 (RFC 1075) is obsolete and was never used.
 - **MOSPF (RFC 1584)**
 - **PIM-DM (Internet-draft)**
 - **PIM-SM (RFC 2362)**
 - **Others (CBT, OCBT, QOSMIC, SM, etc.)**

www.cisco.com

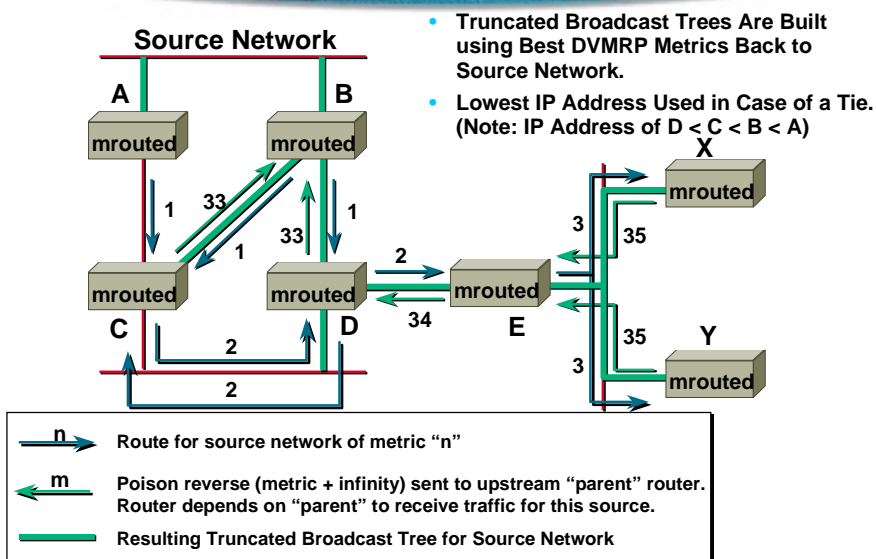
DVMRP Overview

- **Dense Mode Protocol**

- Distance vector-based
 - Similar to RIP
 - Infinity = 32 hops
 - Subnet masks in route advertisements
- DVMRP Routes used:
 - For RPF Check
 - To build Truncated Broadcast Trees (TBTs)
 - Uses special “Poison-Reverse” mechanism
- Uses Flood and Prune operation
 - Traffic initially flooded down TBT’s
 - TBT branches are pruned where traffic is unwanted.
 - Prunes periodically time-out causing reflooding.

www.cisco.com

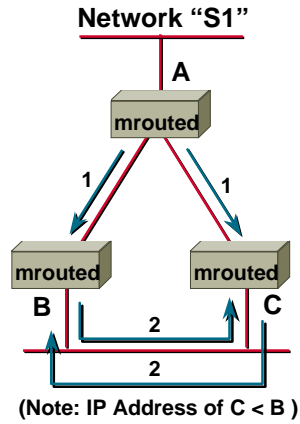
DVMRP—Source Trees



www.cisco.com

DVMRP—Source Trees

Forwarding onto Multi-Access Networks

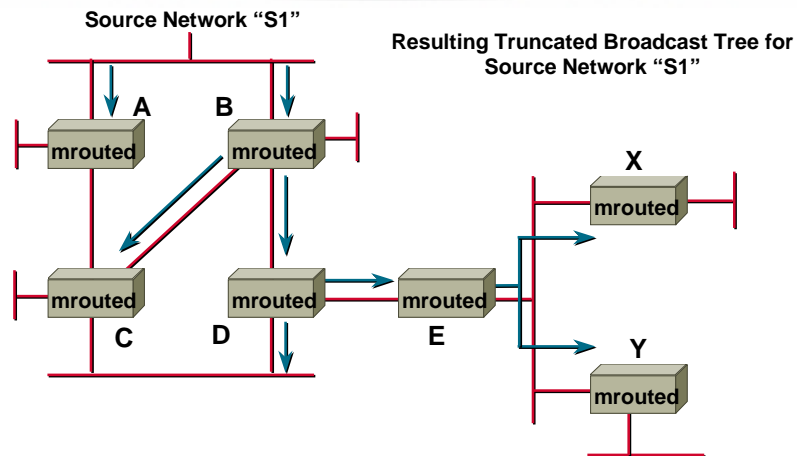


- Both B & C have routes to network S1.
- To avoid duplicates, only one router can be "Designated Forwarder" for network S1.
- Router with best metric is elected as the "Designated Forwarder".
- Lowest IP address used as tie-breaker.
- Router C wins in this example.

Route advertisement for network S1 of metric "n"

www.cisco.com

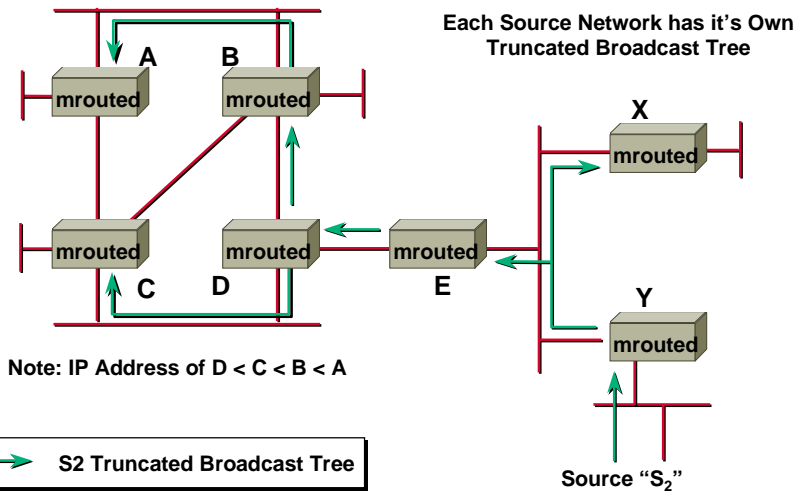
DVMRP — Source Trees



S1 Truncated Broadcast Tree

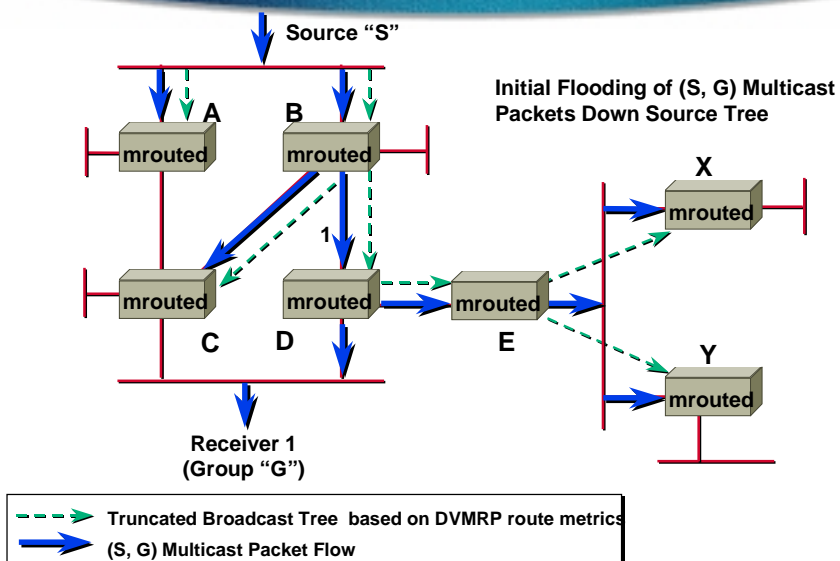
www.cisco.com

DVMRP — Source Trees



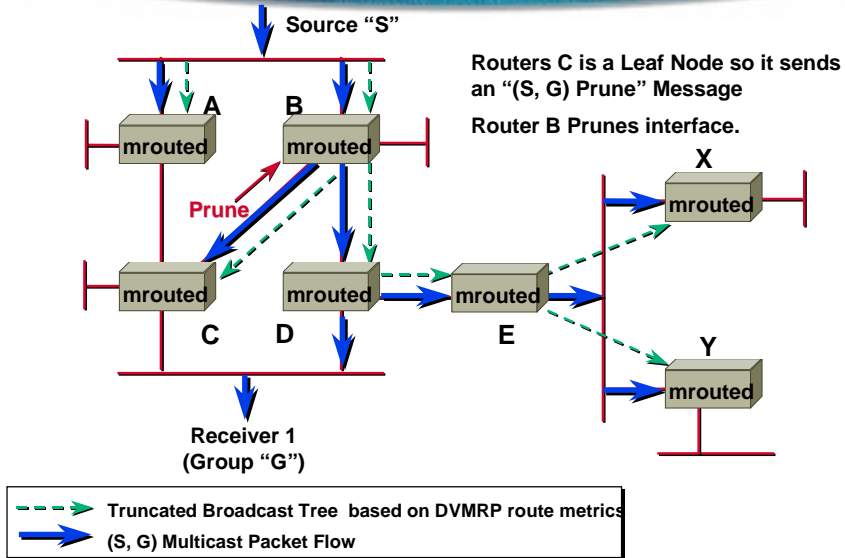
www.cisco.com

DVMRP—Flood and Prune

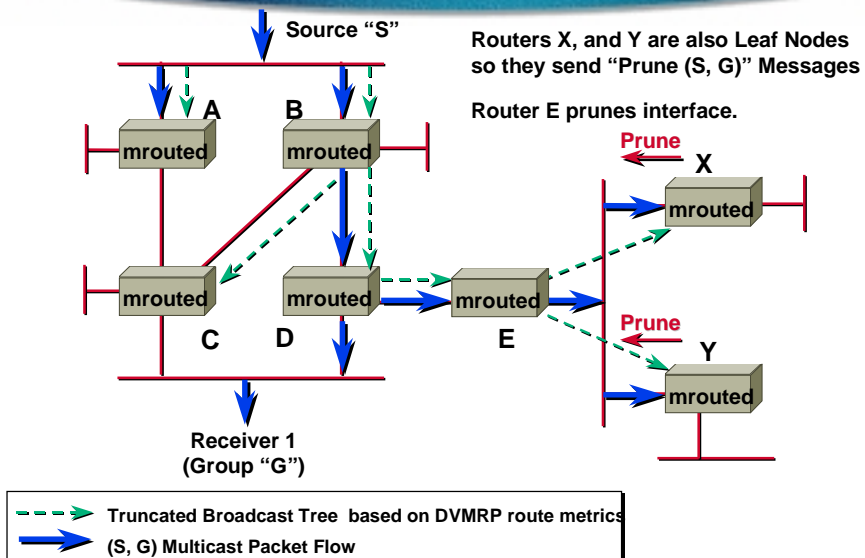


www.cisco.com

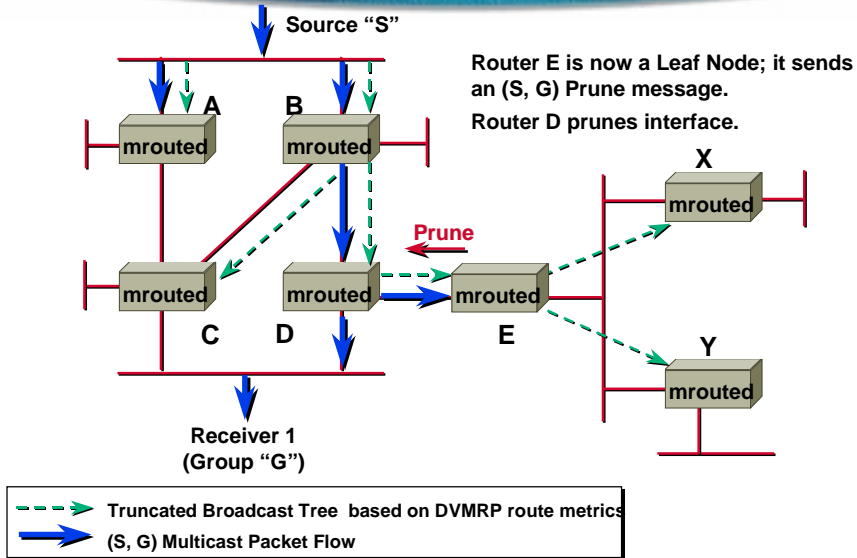
DVMRP—Flood and Prune



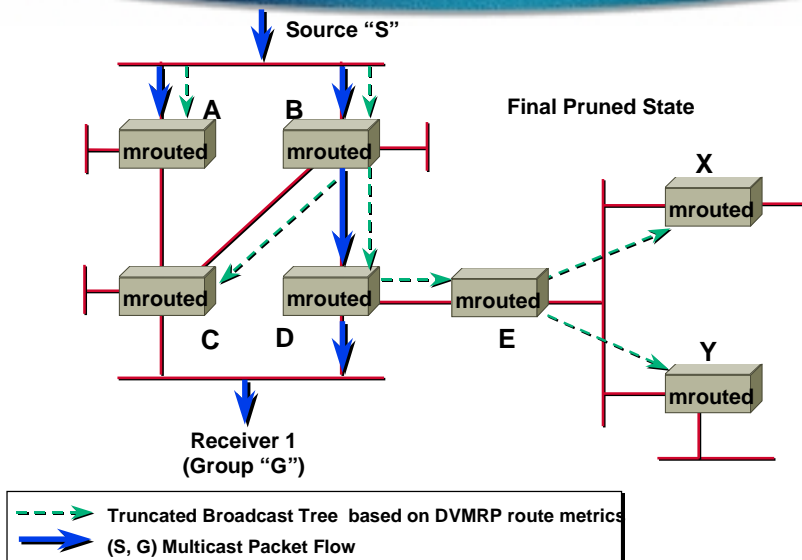
DVMRP—Flood and Prune



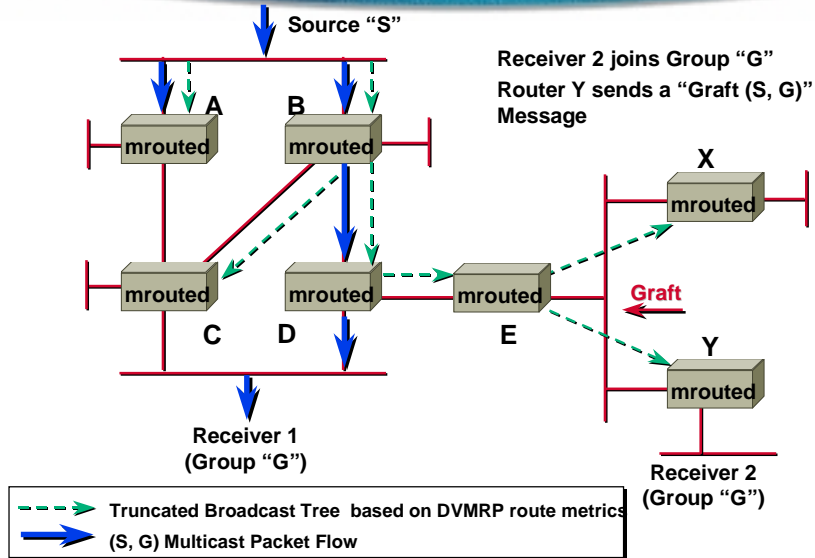
DVMRP—Flood and Prune



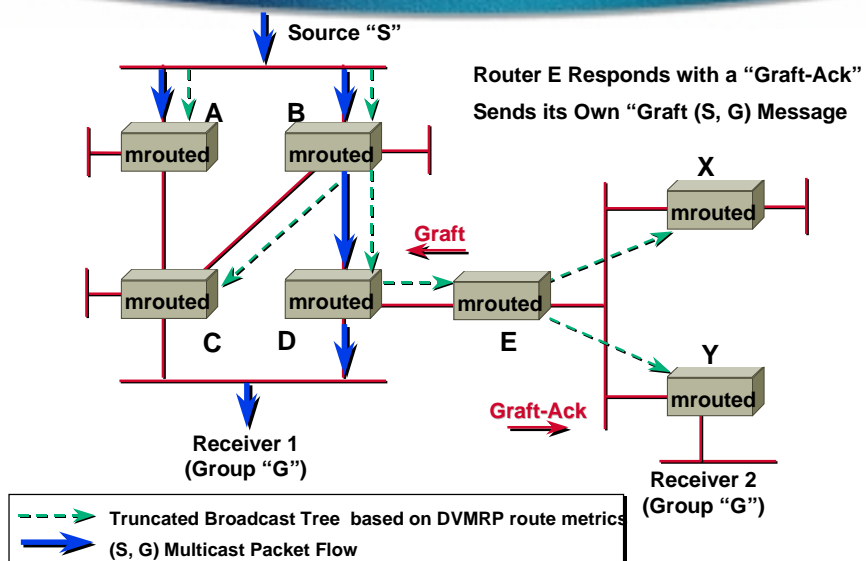
DVMRP—Flood and Prune



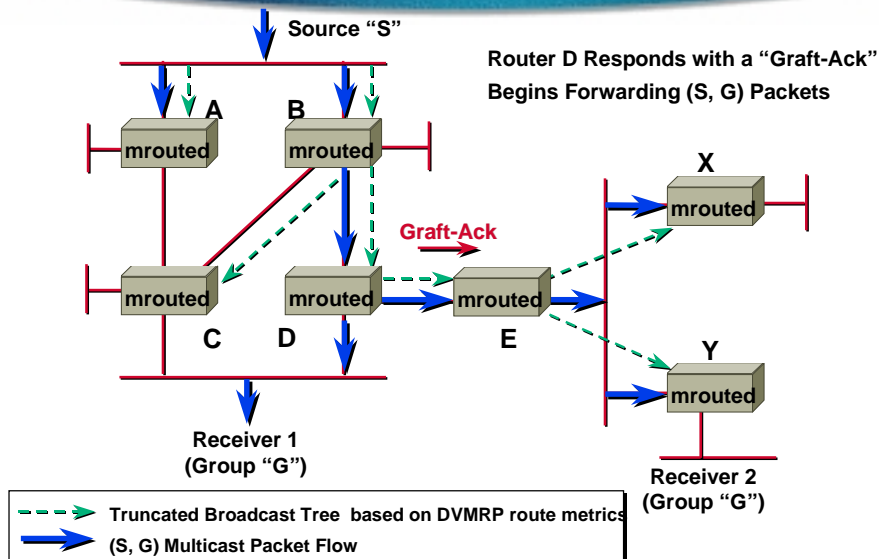
DVMRP—Grafting



DVMRP—Grafting



DVMRP—Grafting



DVMRP—Evaluation

- Widely used on the MBONE (being phased out)
- Significant scaling problems
 - Slow Convergence—RIP-like behavior
 - Significant amount of multicast routing state information stored in routers—(S,G) everywhere
 - No support for shared trees
 - Maximum number of hops < 32
- **Not appropriate** for large scale production networks
 - Due to flood and prune behavior
 - Due to its poor scalability

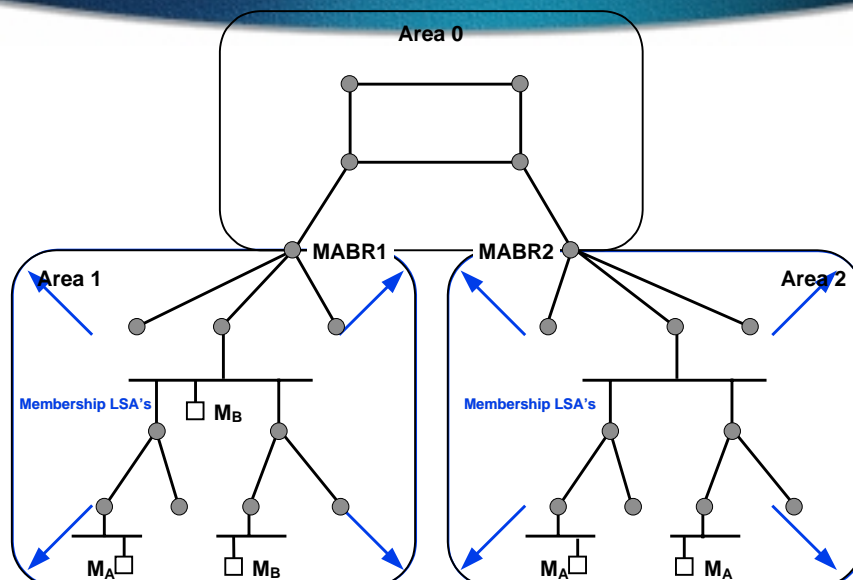
www.cisco.com

MOSPF (RFC 1584)

- **Extension to OSPF unicast routing protocol**
 - OSPF: Routers use link state advertisements to understand all available links in the network (route messages along least-cost paths)
 - MOSPF: Includes multicast information in OSPF link state advertisements to construct multicast distribution trees (each router maintains an up-to-date image of the topology of the entire network)
- **Group membership LSAs are flooded throughout the OSPF routing domain so MOSPF routers can compute outgoing interface lists**
- **Uses Dijkstra algorithm to compute shortest-path tree**
 - Separate calculation is required for each (S_{Net}, G) pair

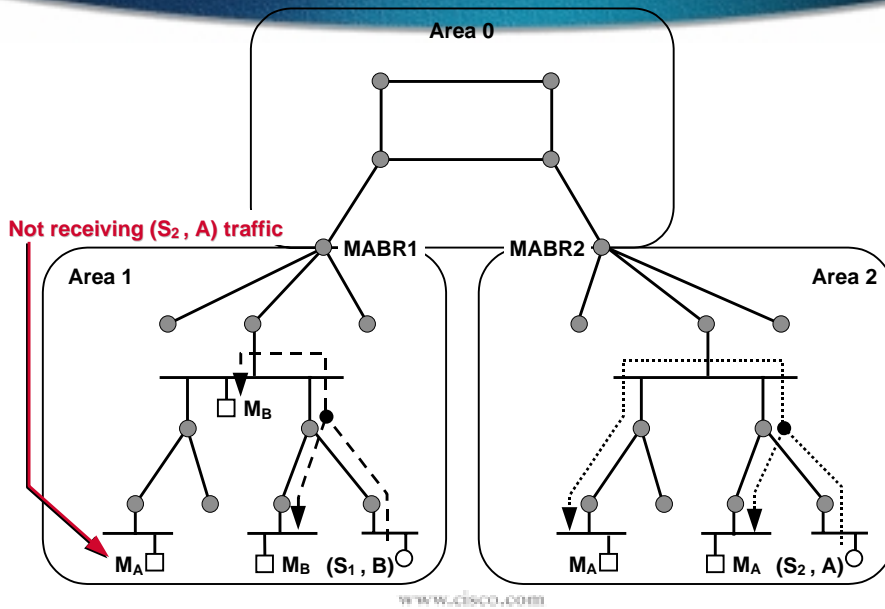
www.cisco.com

MOSPF Membership LSA's

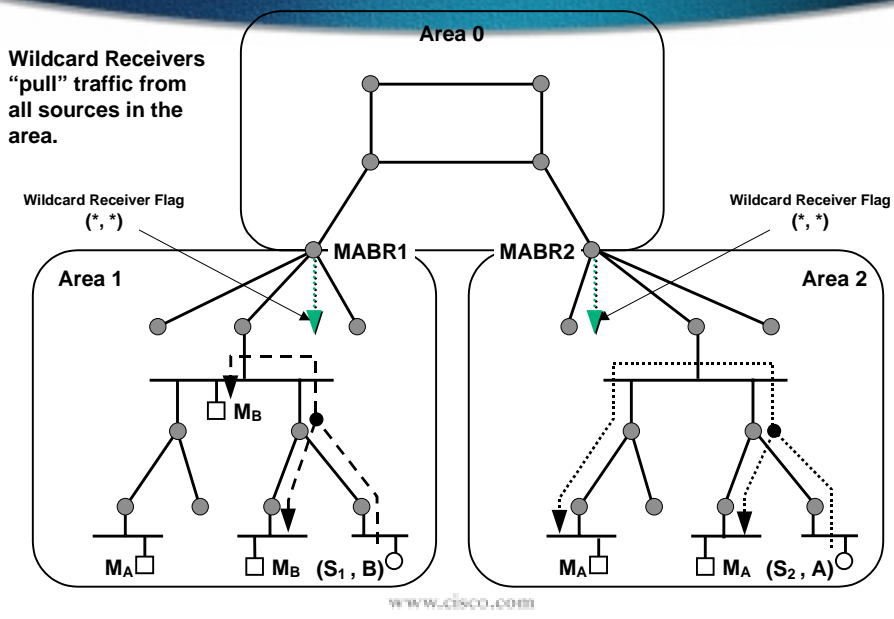


www.cisco.com

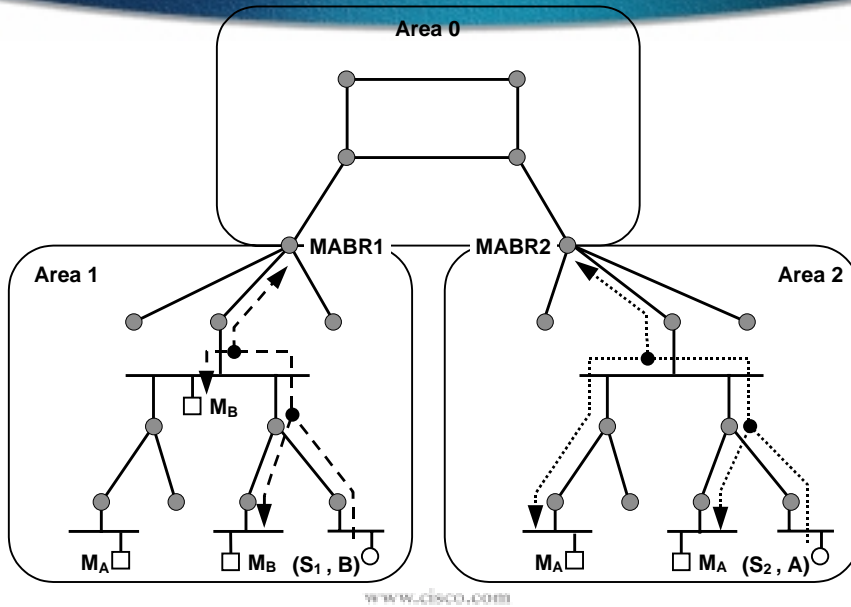
MOSPF Intra-Area Traffic



MOSPF Inter-Area Traffic

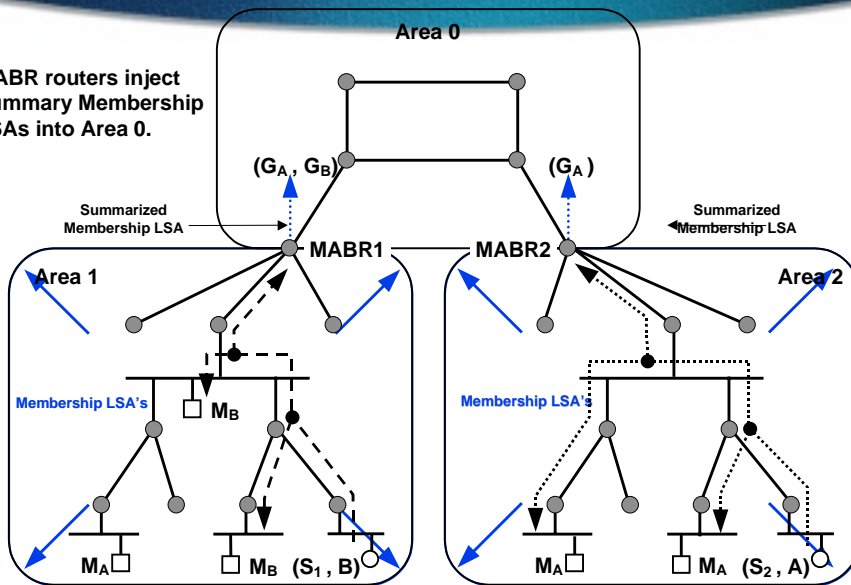


MOSPFP Inter-Area Traffic

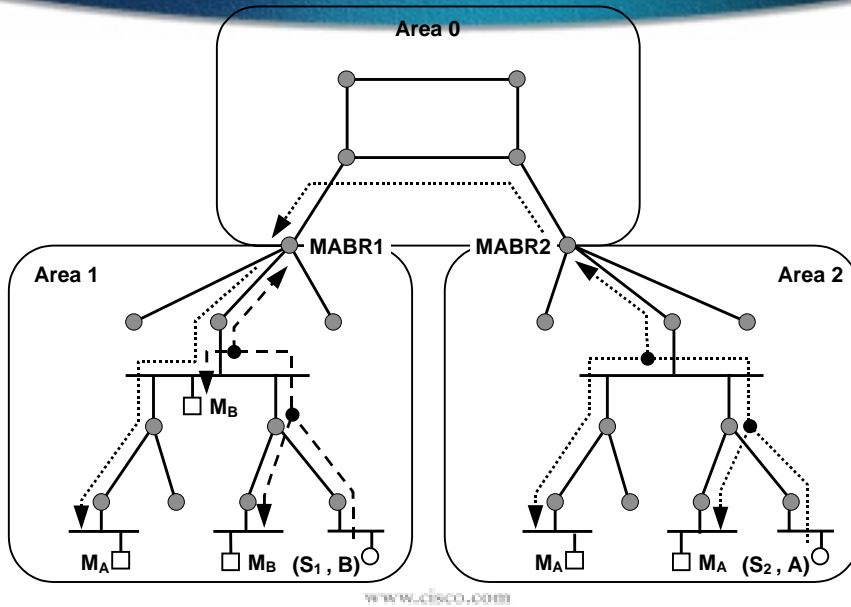


MOSPFP Inter-Area Traffic

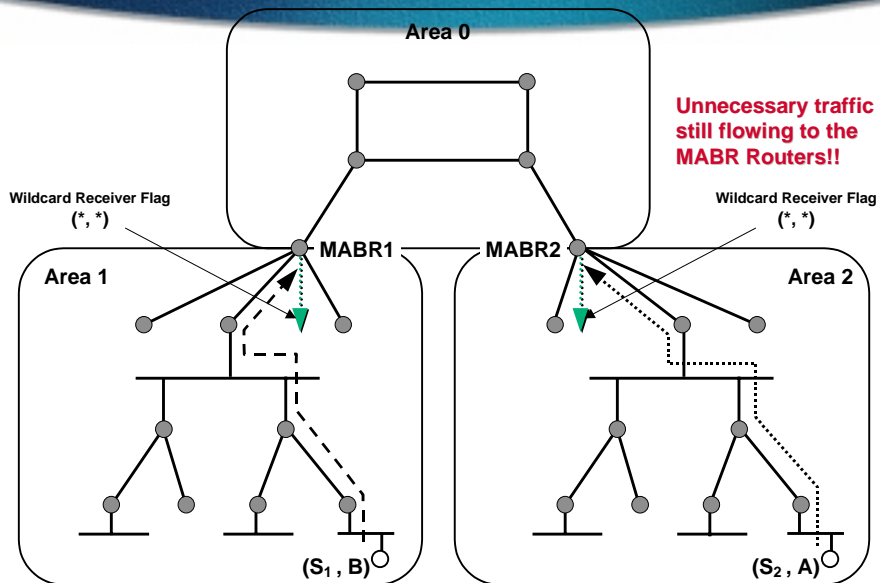
MABR routers inject Summary Membership LSAs into Area 0.



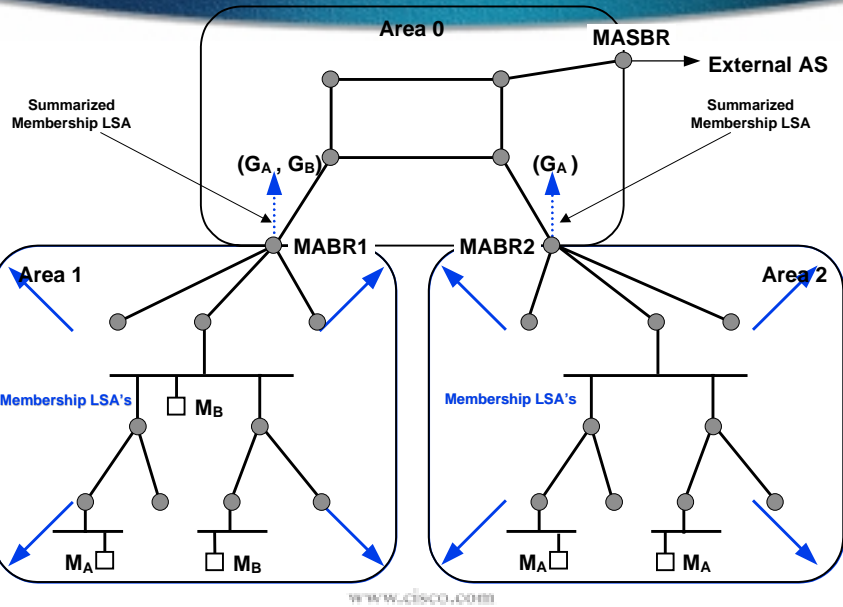
MOSPF Inter-Area Traffic



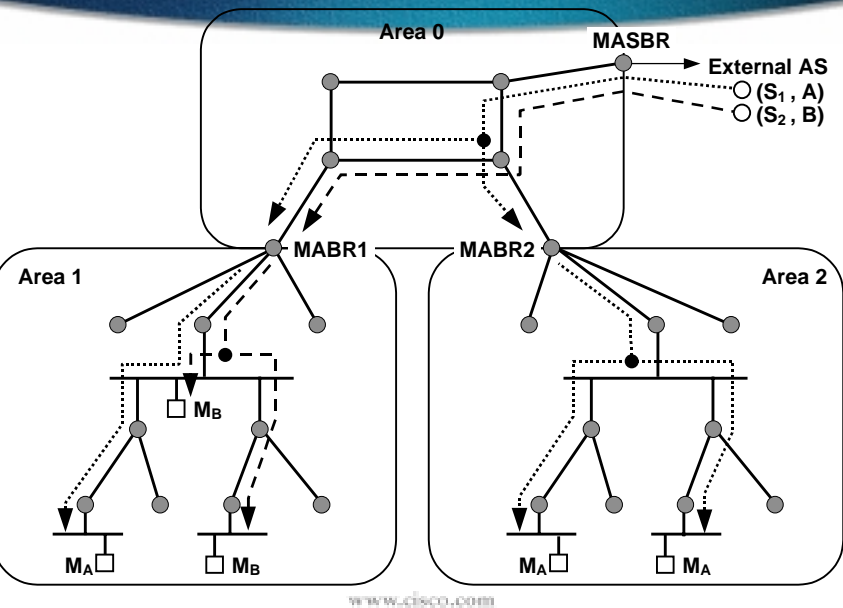
MOSPF Inter-Area Traffic



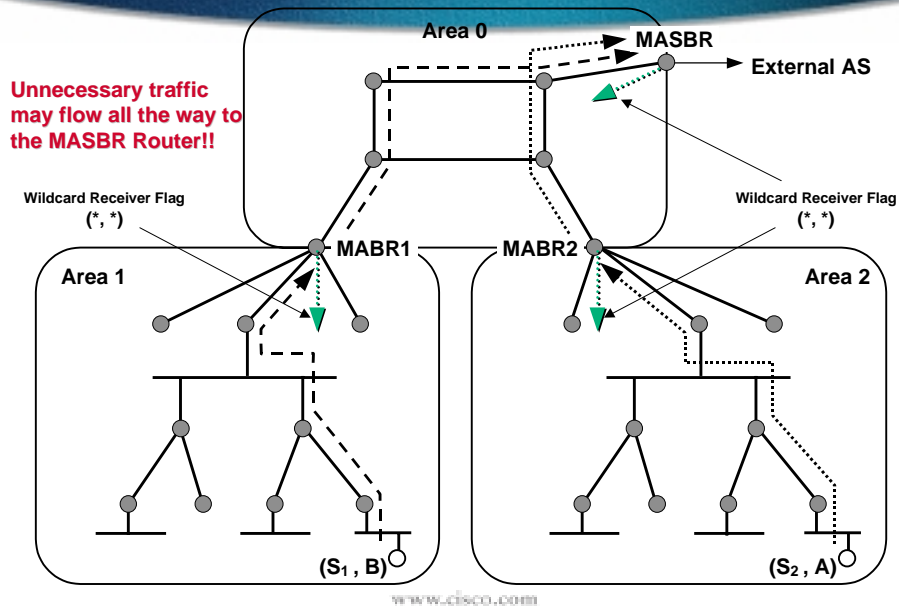
MOSPFP Inter-Domain Traffic



MOSPFP Inter-Domain Traffic



MOSPF Inter-Domain Traffic



MOSPF—Evaluation

- Does not flood multicast traffic everywhere to create state, Uses LSAs and the link-state database
 - Protocol dependent—works only in OSPF-based networks
 - Significant **scaling problems**
 - Dijkstra algorithm run for EVERY multicast (S_{Net}, G) pair!
 - Dijkstra algorithm rerun when:
 - Group Membership changes
 - Line-flaps
 - Does not support **shared-trees**
 - **Not appropriate for...**
 - General purpose multicast networks where the number of senders may be quite large.
 - IP/TV - (Every IP/TV client is a multicast source)
- WWW.CISCO.COM

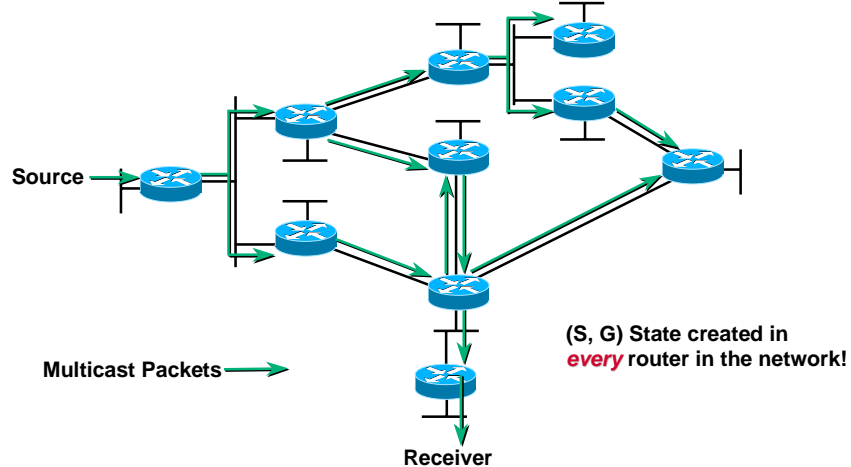
PIM-DM

- **Protocol Independent**
 - Supports all underlying unicast routing protocols including: static, RIP, IGRP, EIGRP, IS-IS, BGP, and OSPF
- **Uses reverse path forwarding**
 - **Floods** network and **prunes** back based on multicast group membership
 - Assert mechanism used to prune off redundant flows
- **Appropriate for...**
 - Smaller implementations and pilot networks

www.cisco.com

PIM-DM Flood and Prune

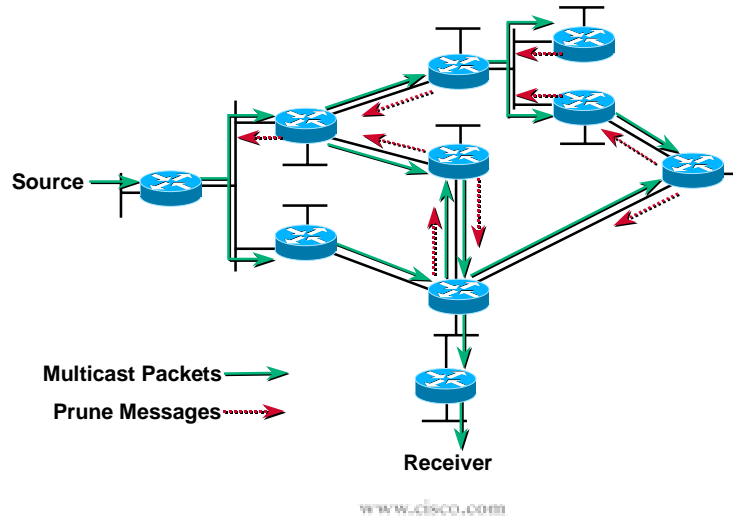
Initial Flooding



www.cisco.com

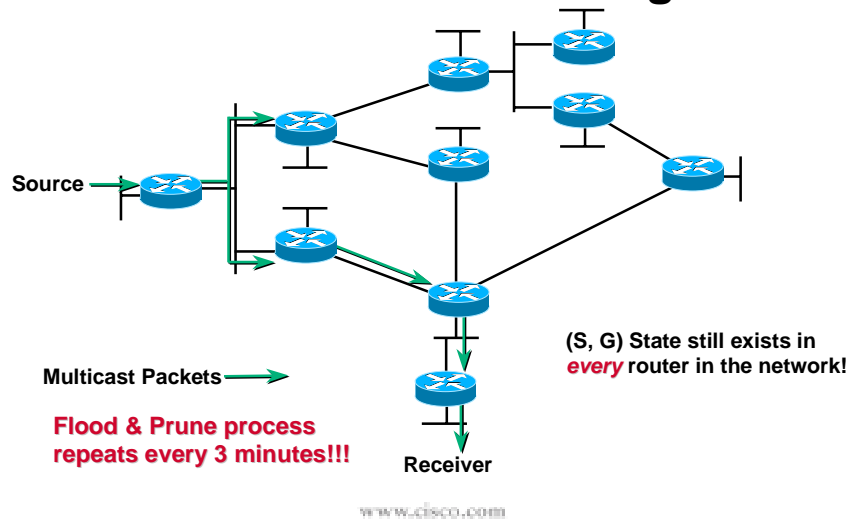
PIM-DM Flood and Prune

Pruning Unwanted Traffic

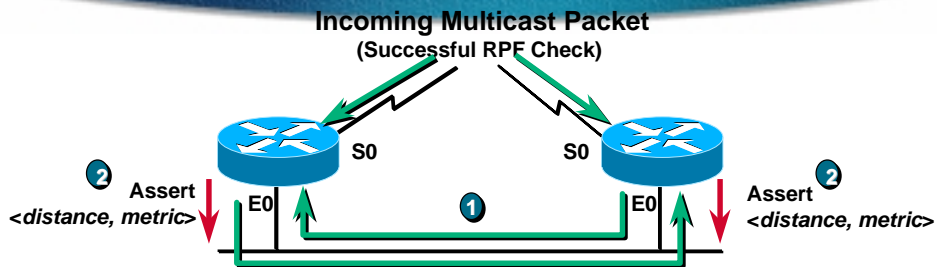


PIM-DM Flood and Prune

Results After Pruning



PIM-DM Assert Mechanism



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

www.cisco.com

PIM-DM—Evaluation

- Most effective for small pilot networks
- Advantages:
 - Easy to configure—two commands
 - Simple flood and prune mechanism
- Potential issues...
 - Inefficient flood and prune behavior
 - Complex Assert mechanism
 - Mixed control and data planes
 - Results in (S, G) state in every router in the network
 - Can result in non-deterministic topological behaviors
 - No support for shared trees

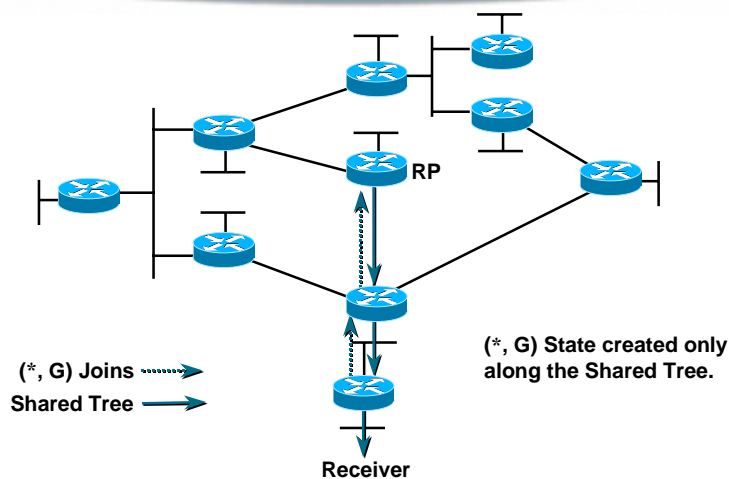
www.cisco.com

PIM-SM (RFC 2362)

- Supports both source and shared trees
 - Assumes no hosts want multicast traffic unless they specifically ask for it
- Uses a **Rendezvous Point (RP)**
 - Senders and Receivers “rendezvous” at this point to learn of each others existence.
 - Senders are “registered” with RP by their first-hop router.
 - Receivers are “joined” to the Shared Tree (rooted at the RP) by their local Designated Router (DR).
- Appropriate for...
 - Wide scale deployment for **both** densely and sparsely populated groups in the enterprise
 - Optimal choice for all production networks regardless of size and membership density.

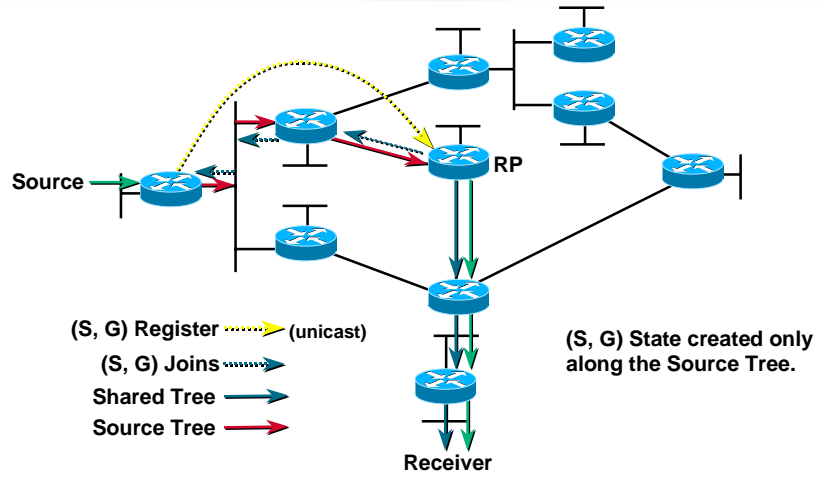
www.cisco.com

PIM-SM Shared Tree Joins



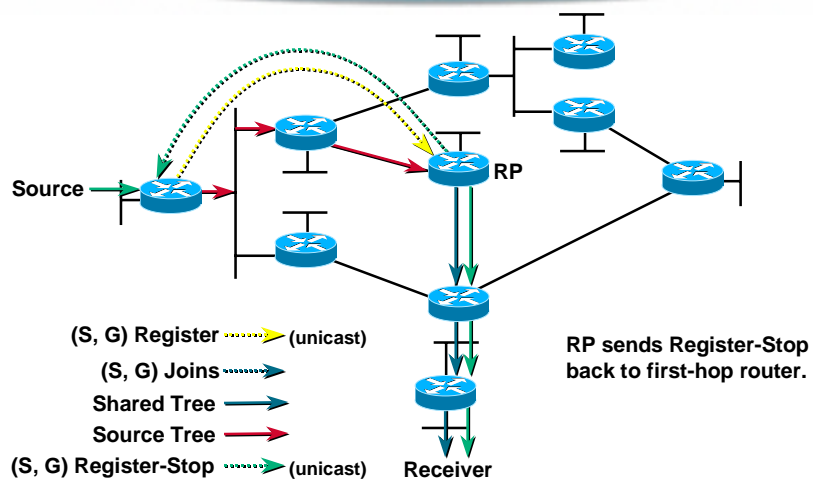
www.cisco.com

PIM-SM Sender Registration



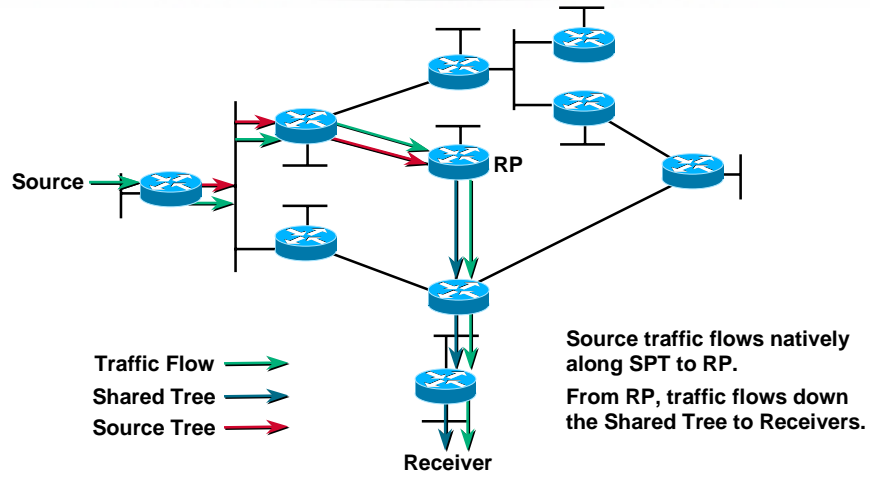
www.cisco.com

PIM-SM Sender Registration



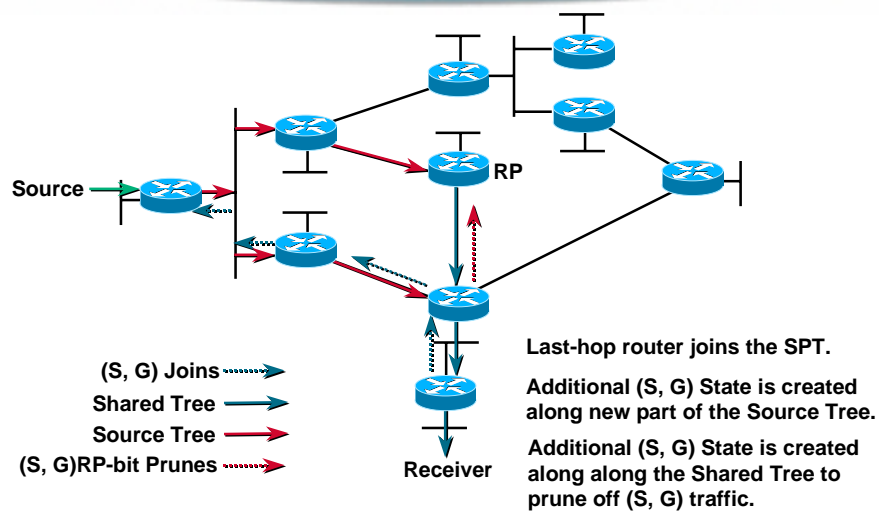
www.cisco.com

PIM-SM Sender Registration



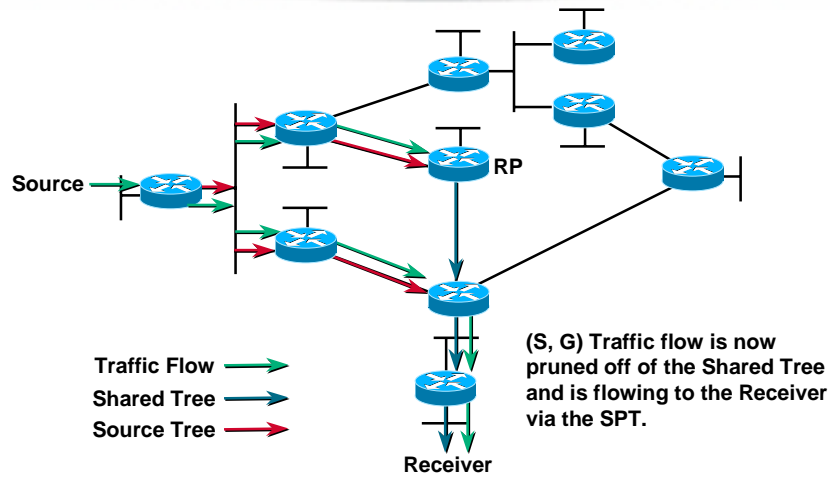
www.cisco.com

PIM-SM SPT Switchover



www.cisco.com

PIM-SM SPT Switchover



www.cisco.com

PIM-SM FFF

PIM-SM Frequently Forgotten Fact

The **default** behavior of PIM-SM is that routers with directly connected members will join the Shortest Path Tree as soon as they detect a new multicast source.

www.cisco.com

PIM-SM—Evaluation

- Effective for **sparse or dense** distribution of multicast receivers
- Advantages:
 - Traffic only sent down “joined” branches
 - Can switch to optimal source-trees for high traffic sources dynamically
 - Unicast routing protocol-independent
 - Basis for inter-domain multicast routing
 - When used with MBGP and MSDP

www.cisco.com

Other Multicast Protocols

- Academic works-in-progress
- Run primarily on MS Power Point

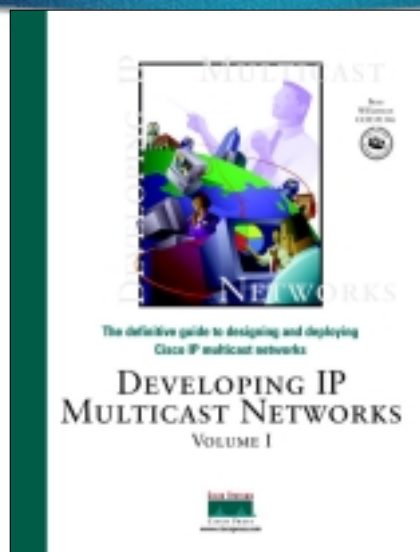
www.cisco.com

Documentation and Contact Info

- **EFT/Beta Site Web Page:**
<ftp://ftpeng.cisco.com/ipmulticast.html>
- **EFT/Beta Mailing List:**
multicast-support@cisco.com
- **Customer Support Mailing List:**
cs-ipmulticast@cisco.com

www.cisco.com

If All Else Fails—RTFB¹



¹ Read this fine book

www.cisco.com

