NETWORKERS 2004

INTRODUCTION TO IP MULTICAST

tkramer@cisco.com

RST-1701
9783_05_2004_X2

Agenda

Cisco.com

- Why Multicast?
- Multicast Fundamentals
- PIM Protocols
- RP choices
- Multicast at Layer 2
- Interdomain IP Multicast

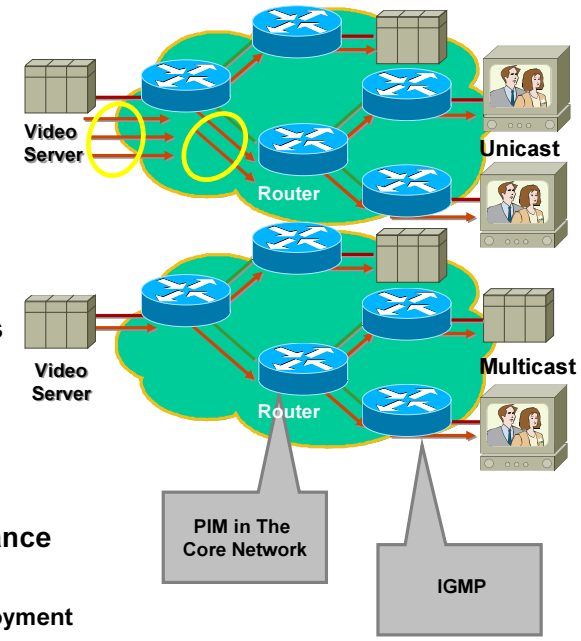
WHY MULTICAST?

RST-1701
9783_05_2004_X2

Unicast vs. Multicast
Scalability == IP Multicast

Cisco.com

- **Host to Network technologies**
  - IGMP v2 and v3
  - Filtering
  - Access Lists (IP, MAC)
- **Network to Network technologies**
  - PIM advanced features
- **Controls Network Traffic**
  - Reduces Network traffic
  - Reduces Servers and CPU load
- **Resiliency & Optimised Performance**
  - Eliminates traffic redundancy
  - Robust and resilient multicast deployment

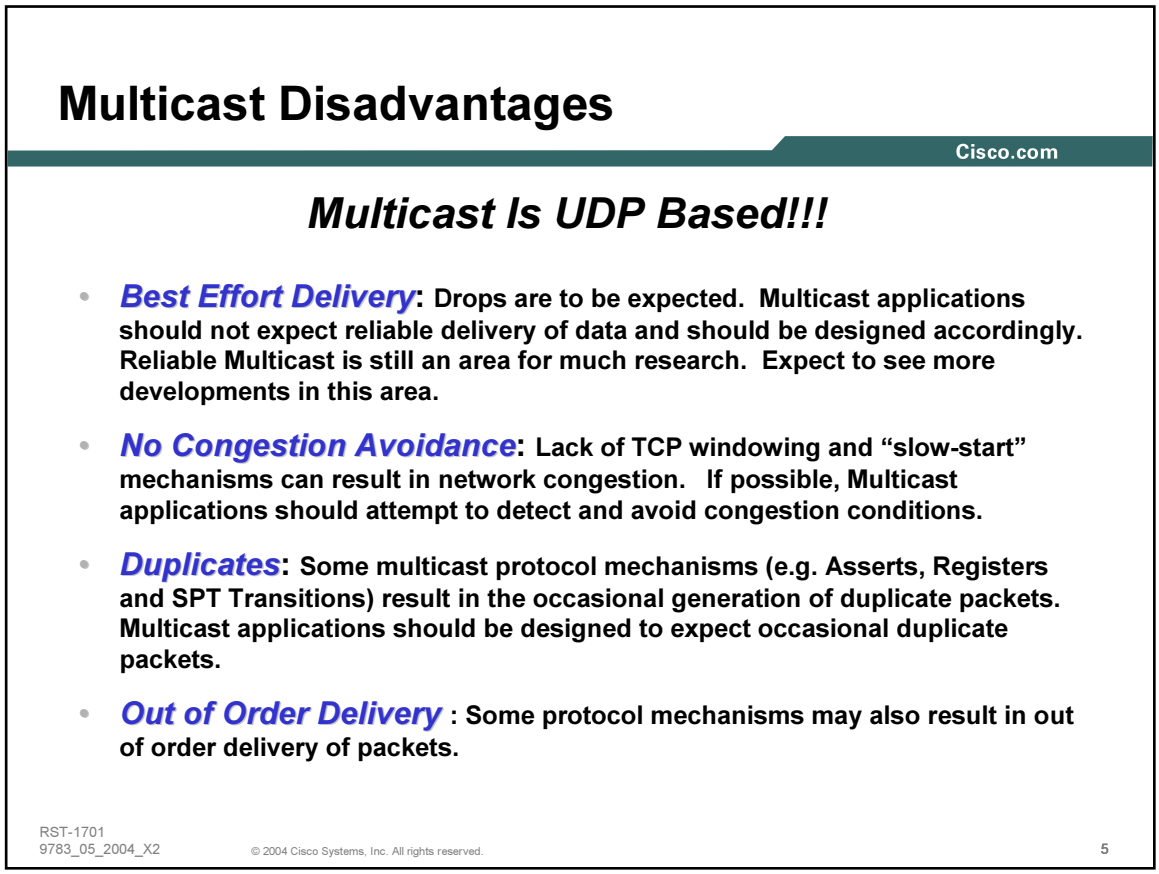


RST-1701
9783_05_2004_X2

4

- STBs are multicast enabled.
- IGMPv2 in last Windows (XP) versions

- Unicast transmission sends multiple copies of data, one copy for each receiver
  - Ex: host transmits 3 copies of data and network forwards each to 3 separate receivers
  - Ex: host can only send to one receiver at a time
- Multicast transmission sends a single copy of data to multiple receivers
  - Ex: host transmits 1 copy of data and network replicates at last possible hop for each receiver, each packet exists only one time on any given network
  - Ex: host can send to multiple receivers simultaneously



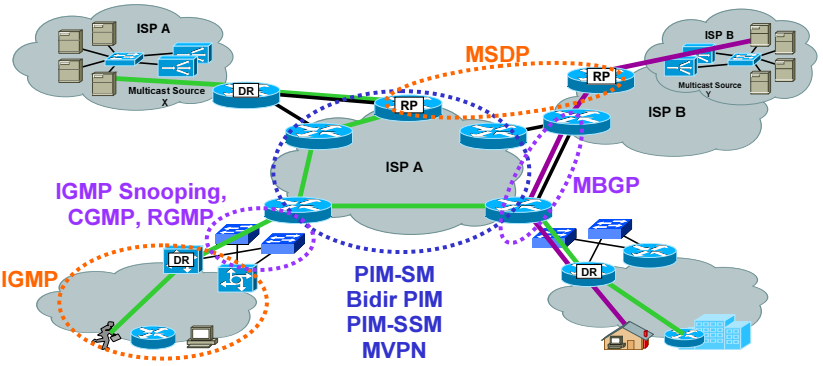
- Multicast Disadvantages
  - Most Multicast Applications are UDP based. This results in some undesirable side-effects when compared to similar unicast, TCP applications.
  - Best Effort Delivery results in occasional packet drops. Many multicast applications that operate in real-time (e.g. Video, Audio) can be impacted by these losses. Also, requesting retransmission of the lost data at the application layer in these sort of real-time applications is not feasible.
    - Heavy drops on Voice applications result in jerky, missed speech patterns that can make the content unintelligable when the drop rate gets high enough.
    - Moderate to Heavy drops in Video is sometimes better tolerated by the human eye and appear as unusual "artifacts" on the picture. However, some compression algorithms can be severally impacted by even low drop rates; causing the picture to become jerky or freeze for several seconds while the decompression algorithm recovers.
  - No Congestion Control can result in overall Network Degradation as the popularity of UDP based Multicast applications grow.
  - Duplicate packets can occasionally be generated as multicast network topologies change.

MULTICAST FUNDAMENTALS

Multicast Components

Cisco End-to-End Architecture

Cisco.com



Campus Multicast

- End Stations (hosts-to-routers):
  - IGMP
- Switches (Layer 2 Optimization):
  - CGMP, IGMP Snooping or RGMP
- Routers (Multicast Forwarding Protocol):
  - PIM Sparse Mode or Bidirectional PIM

Interdomain Multicast

- Multicast routing across domains
  - MBGP
- Multicast Source Discovery
  - MSDP with PIM-SM
- Source Specific Multicast
  - PIM-SSM

RST-1701
9783_05_2004_X2

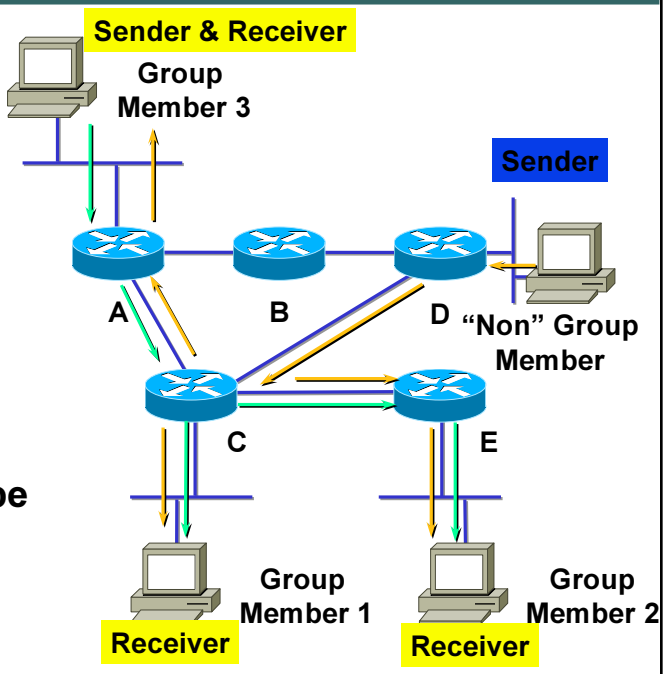
7

- IP Multicast can be packaged into two solutions from a functional standpoint.
  - The first of these solutions is Campus multicast or Intradomain Multicast:
  - The first is end station management or host to router protocols (IGMP) which request admission to join or leave multicast groups.
  - The second component of this solution is switch membership management via CGMP or IGMP snooping. This ensures that the switch intelligently handles multicast forwarding without flooding the network.
  - The third component is PIM Sparse Mode which ensures that the network has appropriate information necessary to forward multicast packets down a multicast distribution tree from source to destination.
  - The second of these is Internet Multicast which includes:
  - MBGP for AS to AS multicast routing information
  - and
  - MSDP for third party source discovery across PIM Sparse Mode Clouds.

IP Multicast Group Concept

Cisco.com

1. You **MUST BE** a "member" of a group to receive it data
2. If you send to group address, all members receive it
3. You **DO NOT** have to be a member of a group to send to a group



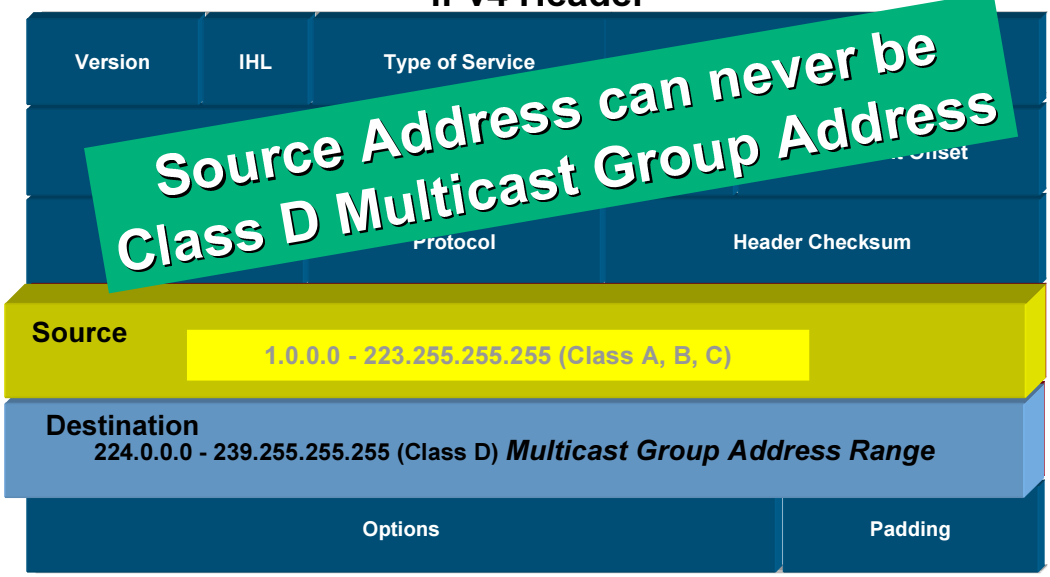
RST-1701
9783_05_2004_X2

8

- Recognize distinction between sender behavior and source behavior.

Multicast Addressing

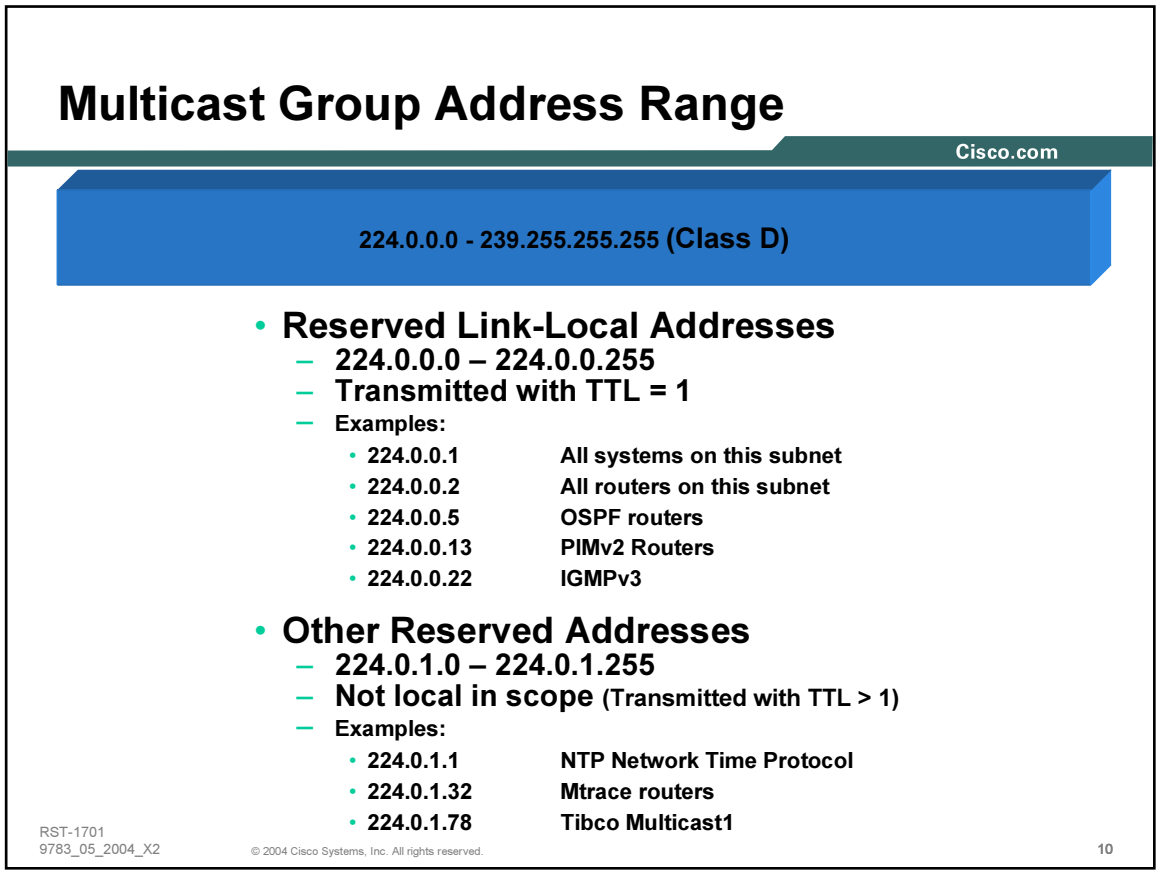
Cisco.com



RST-1701
9783_05_2004_X2

9

9783_05_2004_X2.scr



- IANA Reserved Addresses

  - IANA is the responsible Authority for the assignment of reserved class D addresses. Other interesting reserved addresses are:
  - 224.0.0.2 - PIMv1 (ALL-ROUTERS - due to transport in IGMPv1)
  - 224.0.0.5 - OSPF ALL ROUTERS (RFC1583)
  - 224.0.0.6 - OSPF DESIGNATED ROUTERS (RFC1583)
  - 224.0.0.9 - RIP2 Routers
  - 224.0.0.13 - PIMv2
  - 224.0.1.39 - CISCO-RP-ANNOUNCE (Auto-RP)
  - 224.0.1.40 - CISCO-RP-DISCOVERY (Auto-RP)

  "ftp://ftp.isi.edu/in-notes/iana/assignments/multicast-addresses" is the authoritative source for reserved multicast addresses.

- Additional Information

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**
  - **Examples**
    - **Site-local scope: 239.255.0.0/16**
    - **Organization-local scope: 239.192.0.0/14**
- **SSM (Source Specific Multicast) Range**
  - **232.0.0.0 – 232.255.255.255**
  - **Primarily targeted for Internet style Broadcast**

RST-1701
9783_05_2004_X2

11

- IANA Reserved Addresses

  - IANA is the responsible Authority for the assignment of reserved class D addresses. Other interesting reserved addresses are:
  - 224.0.0.2 - PIMv1 (ALL-ROUTERS - due to transport in IGMPv1)
  - 224.0.0.5 - OSPF ALL ROUTERS (RFC1583)
  - 224.0.0.6 - OSPF DESIGNATED ROUTERS (RFC1583)
  - 224.0.0.9 - RIP2 Routers
  - 224.0.0.13 - PIMv2
  - 224.0.1.39 - CISCO-RP-ANNOUNCE (Auto-RP)
  - 224.0.1.40 - CISCO-RP-DISCOVERY (Auto-RP)

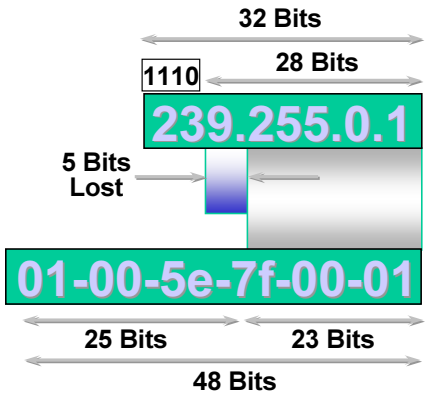
  "ftp://ftp.isi.edu/in-notes/iana/assignments/multicast-addresses" is the authoritative source for reserved multicast addresses.

- Additional Information

Multicast Addressing

Cisco.com

IP Multicast MAC Address Mapping
(FDDI and Ethernet)



RST-1701
9783_05_2004_X2

12

- Ethernet & FDDI Multicast Addresses
  - The low order bit (0x01) in the first octet indicates that this packet is a Layer 2 multicast packet. Furthermore, the “0x01005e” prefix has been reserved for use in mapping L3 IPmc addresses into L2 MAC addresses.

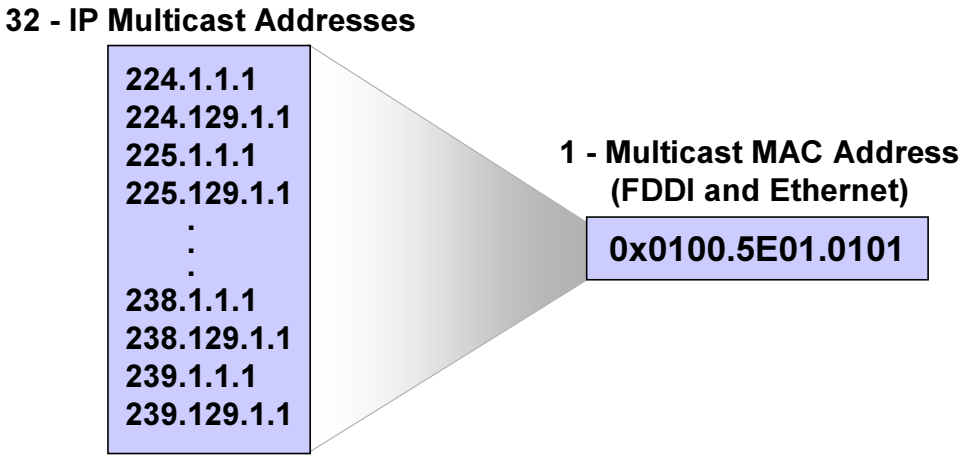
    When mapping L3 to L2 addresses, the low order 23 bits of the L3 IPmc address are mapped into the low order 23 bits of the IEEE mac address
- L2/L3 Multicast Address Overlap
  - Since there are 28 bits of unique address space for an IPmc address (32 minus the first 4 bits containing the 1110 Class D prefix) and there are only 23 bits plugged into the IEEE MAC address - there are 5 bits of overlap or 28-23 = 5. 2**5 = 32 therefore there is a 32:1 overlap of L3 addresses to L2 addresses - so beware several L3 addresses can map to the same L2 multicast address!
  - For example, all of the following IPmc addresses map to the same L2 multicast of 01-00-5e-0a-00-01:
  - 224.10.0.1, 225.10.0.1, 226.10.0.1, 227.10.0.1
  - 228.10.0.1, 229.10.0.1, 230.10.0.1, 231.10.0.1
  - 232.10.0.1, 233.10.0.1, 234.10.0.1, 235.10.0.1

Multicast Addressing

Cisco.com

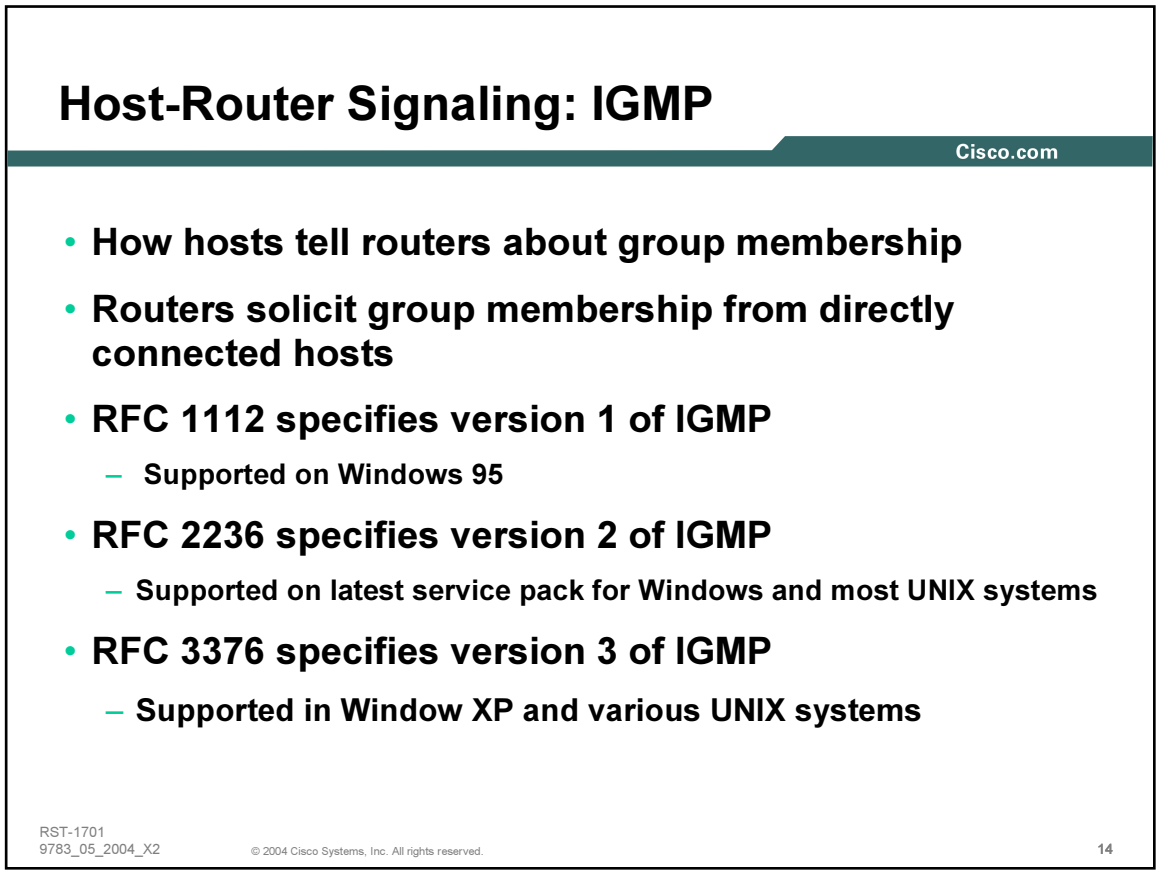
IP Multicast MAC Address Mapping
(FDDI & Ethernet)

Be Aware of the 32:1 Address Overlap



RST-1701
9783_05_2004_X2

13



- IGMP
  - The primary purpose of IGMP is to permit hosts to commincate their desire to receive multicast traffic to the IP Multicast router(s) on the local network. This, in turn, permits the IP Multicast router(s) to “Join” the specified multicast group and to begin forwarding the multicast traffic onto the network segment.

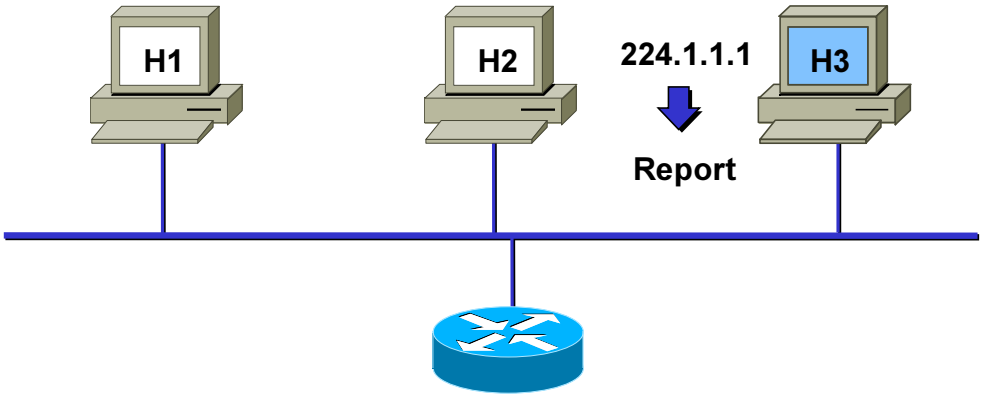
    The initial specification for IGMP (v1) was documented in RFC 1112, “Host Extensions for IP Multicasting”. Since that time, many problems and limitations with IGMPv1 have been discovered. This has lead to the development of the IGMPv2 specification which was ratified in November, 1997 as RFC 2236.

    Even before IGMPv2 had been ratified, work on the next generation of the IGMP protocol, IGMPv3, had already begun. However, the IGMPv3 specification is still in the working stage and has not been implemented by any vendors.

Host-Router Signaling: IGMP

Cisco.com

Joining a Group

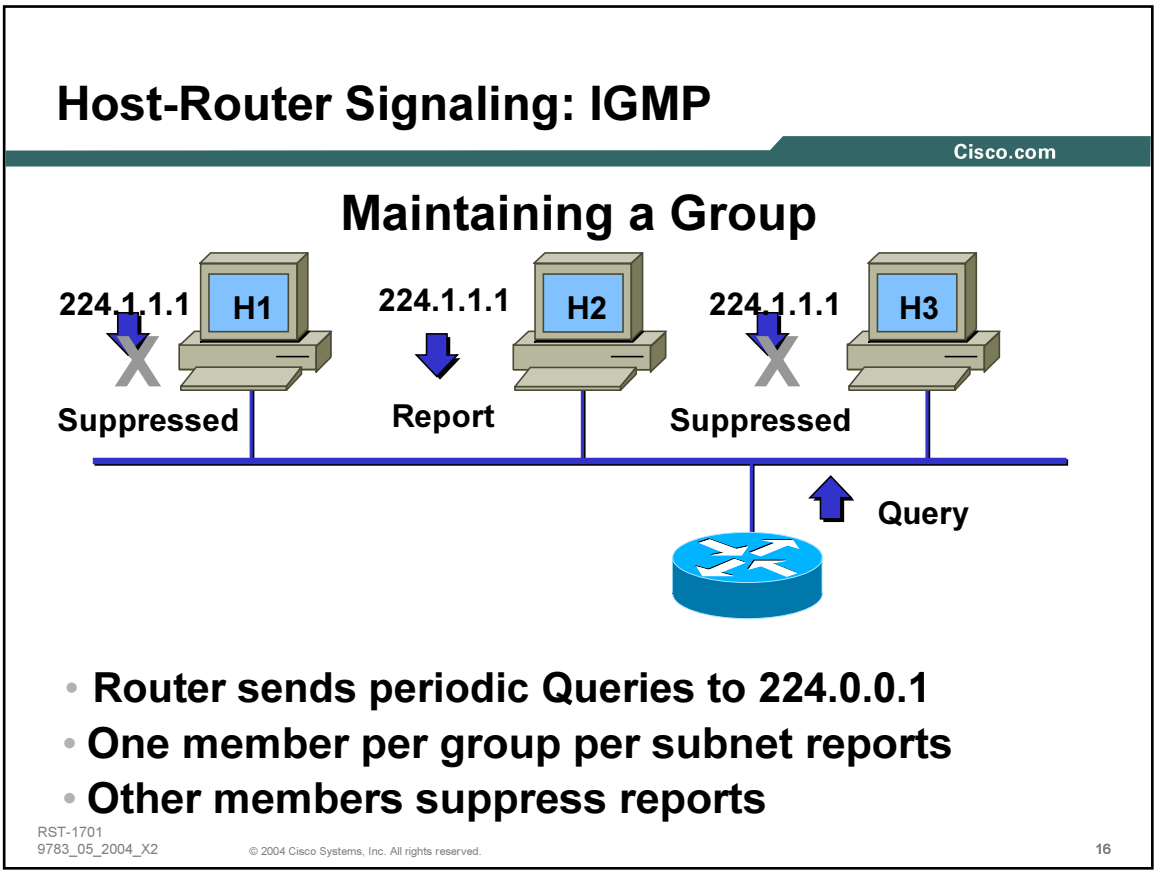


- **Host sends IGMP Report to join group**

RST-1701
9783_05_2004_X2

15

- Asynchronous Joins
  - Members joining a group do not have to waited for a query to join; they send in an unsolictied report indicating their interest. This reduces join latency for the end system joining if no other members are present.



- Query-Response Process
  - The router multicasts periodic IGMPv1 Membership Queries to the “All-Hosts” (224.0.0.1) group address.

    Only one member per group responds with a report to a query. This is to save bandwidth on the subnet network and processing by the hosts. This is process is called “Response Suppression”. (See section below.)

- Response Suppression Mechanism
  - The “Report Suppression” mechanism is accomplished as follows:

    When a host receives the Query, it starts a count-down timer for each multicast group of which it is a member. The count-down timers are each initialized to a random count within a given time range. (In IGMPv1 this was a fixed range of 10 seconds. Therefore the count-down timers were randomly set to some value between 0 and 10 seconds.)

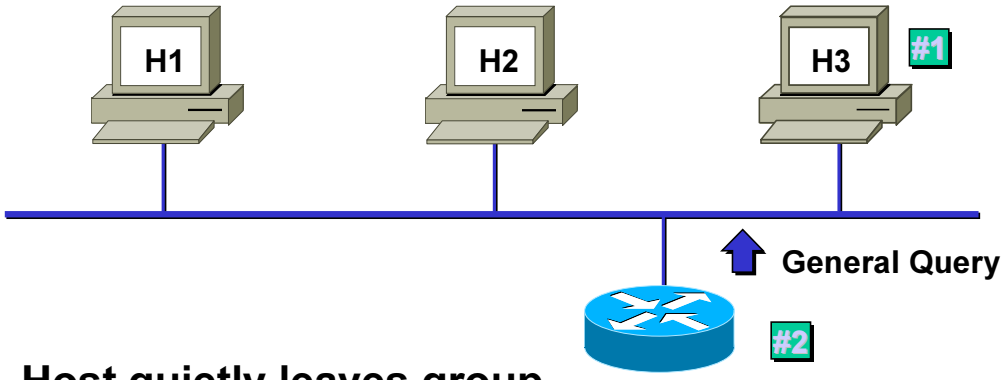
    When a count-down timer reaches zero, the host sends a Membership Report for the group associated with the count-down timer to notify the router that the group is still active.

    However, if a host receives a Membership Report before its

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)

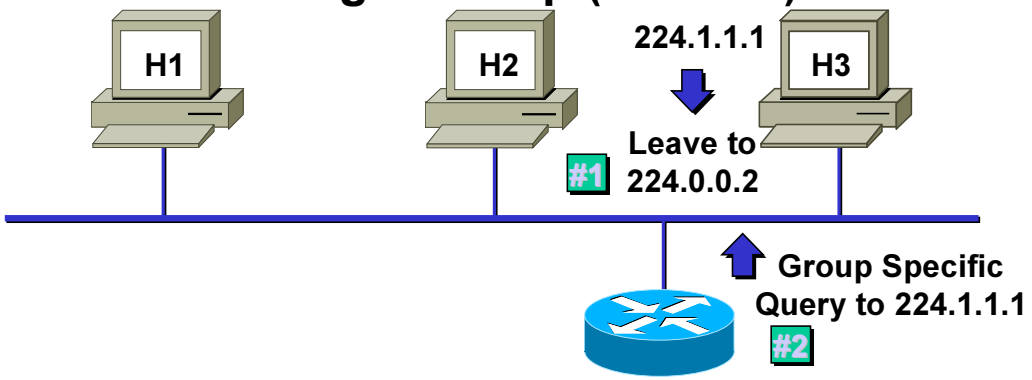
RST-1701
9783_05_2004_X2

17

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

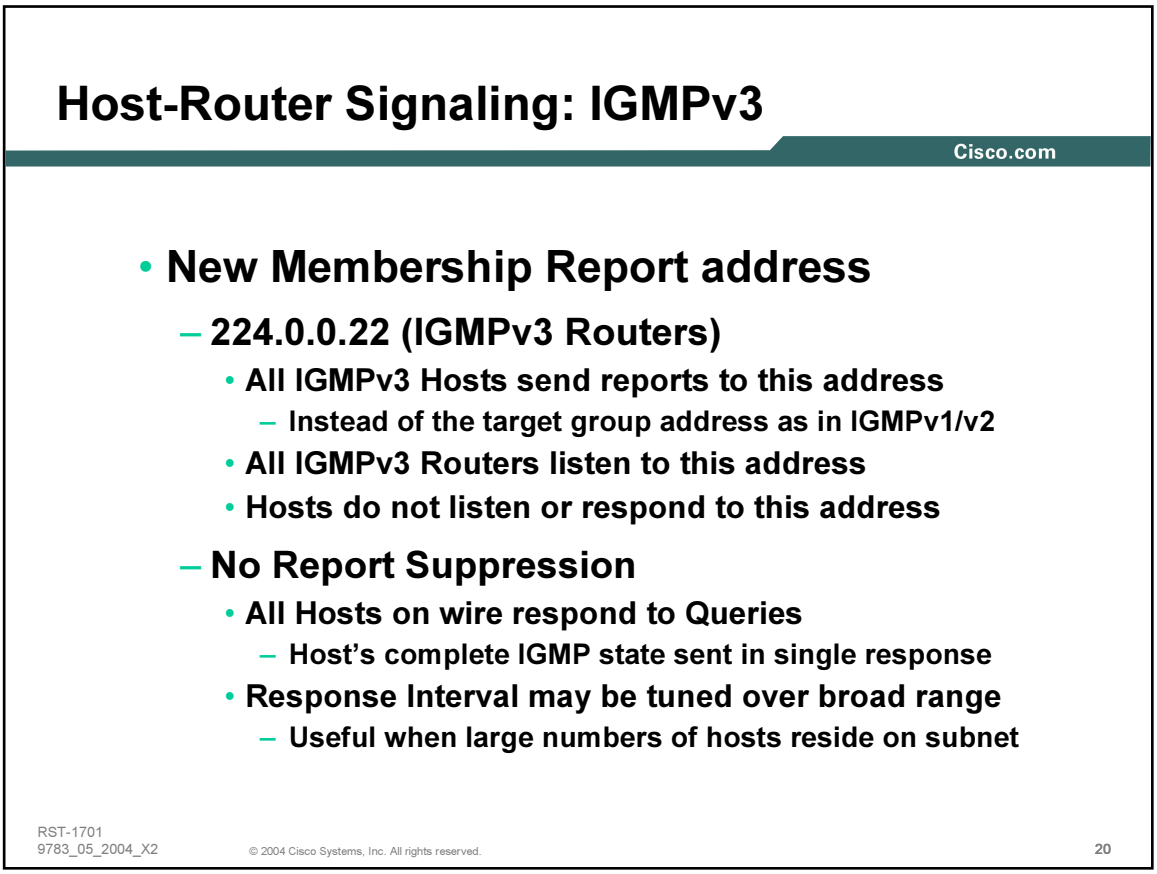
Cisco.com

- **RFC 3376**
  - **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**

RST-1701
9783_05_2004_X2

19

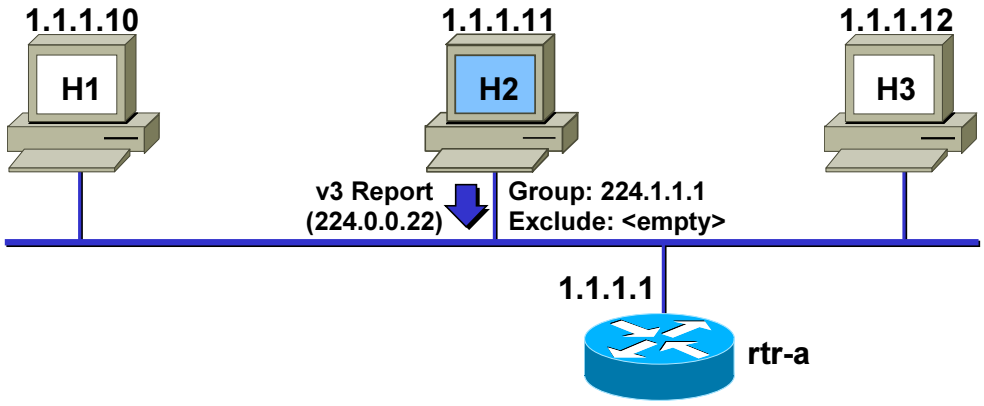
- IGMPv3
  - The IDMR is completing work on IGMPv3.
  - The key change in IGMPv3 is the addition of Group records each containing a list of sources to Include or Exclude. This permits a host to signal which set of hosts that they wish to receive group traffic.
  - IGMPv3 requires that the 'IPMulticastListen' API be changed to accommodate the Include/Exclude filter list. This means that the IGMP stack in the OS will have to be updated to support IGMPv3.
  - In order to take advantage of the benefits of IGMPv3, applications must be (re)written to support the new API.



- IGMPv3
  - IGMPv3 is assigned its own "All IGMPv3 Routers" link-local multicast group address, 224.0.0.22.
    - IGMPv3 hosts no longer send their reports to the target multicast group address. Instead, they send their IGMPv3 Membership Reports to the "All IGMPv3 Routers" multicast address.
    - Routers listen to the 224.0.0.22 address in order to receive and maintain IGMP membership state for every member on the subnet! This is a radical change over the behavior in IGMPv1/v2 where the routers only maintained group state on a subnet basis.
    - Hosts do not listen to 224.0.0.22 and therefore do not hear other hosts' IGMPv3 membership reports.
  - IGMPv3 drops the Report Suppression mechanism that was used in IGMPv1/v2.
    - All IGMPv3 hosts on the wire respond to Queries by sending and IGMPv3 membership reports containing their total IGMP state for all groups in the report.
    - In order to prevent huge bursts of IGMPv3 Reports, the Response Interval may now be tuned over a much greater range than before. This permits the network engineer to

IGMPv3—Joining a Group

Cisco.com



- **Joining member sends IGMPv3 Report to 224.0.0.22 immediately upon joining**

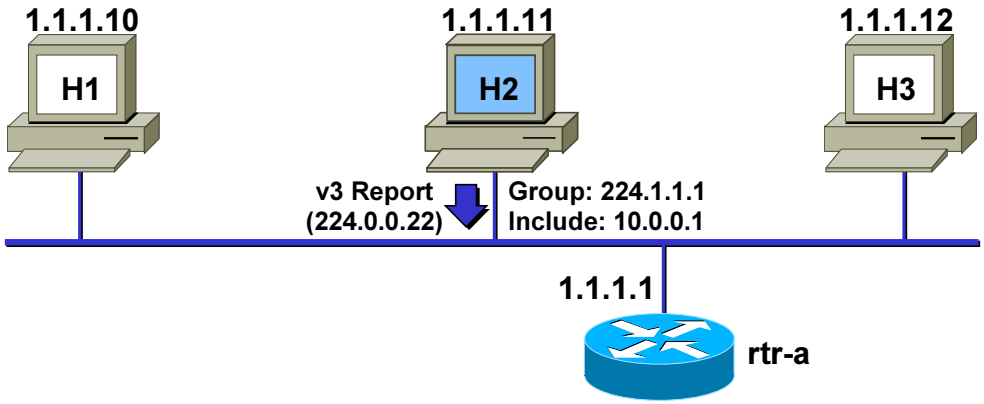
RST-1701
9783_05_2004_X2

21

- Asynchronous Joins
  - Members joining a group do not have to waited for a query to join; they send in an unsolicited IGMPv3 Membership Report indicating their interest. This reduces join latency for the end system joining if no other members are present.
    - In the example above, Host 2 is joining multicast group 224.1.1.1 and is willing to receive any and all sources in this group.

IGMPv3—Joining Specific Source(s)

Cisco.com



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

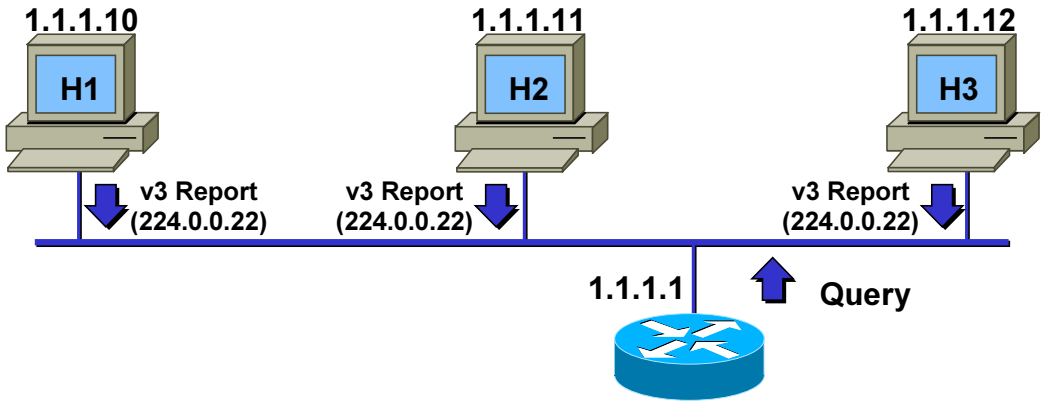
RST-1701
9783_05_2004_X2

22

- Joining only specific Source(s)
  - Hosts may signal the router that it wishes to receive only a specific set of sources sending to the group. This is done by using an "Include" list in the Group record of the Report. When an Include list is in use, only the specific sources listed in the Include list are joined.
    - In the example above, Host 2 is joining multicast group 224.1.1.1 and only wants to receive source 10.0.0.1 sending to the group.

IGMPv3—Maintaining State

Cisco.com



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

RST-1701
9783_05_2004_X2

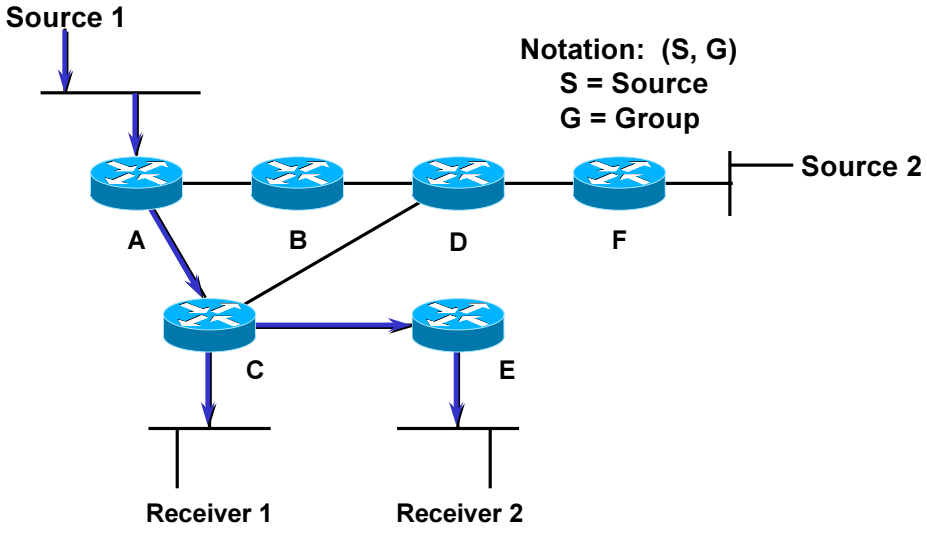
23

- Query-Response Process
  - The router multicasts periodic Membership Queries to the “All-Hosts” (224.0.0.1) group address.
  - All hosts on the wire respond by sending back an IGMPv3 Membership Report that contains their complete IGMP Group state for the interface.

Multicast Distribution Trees

Cisco.com

Shortest Path or Source Distribution Tree



RST-1701
9783_05_2004_X2

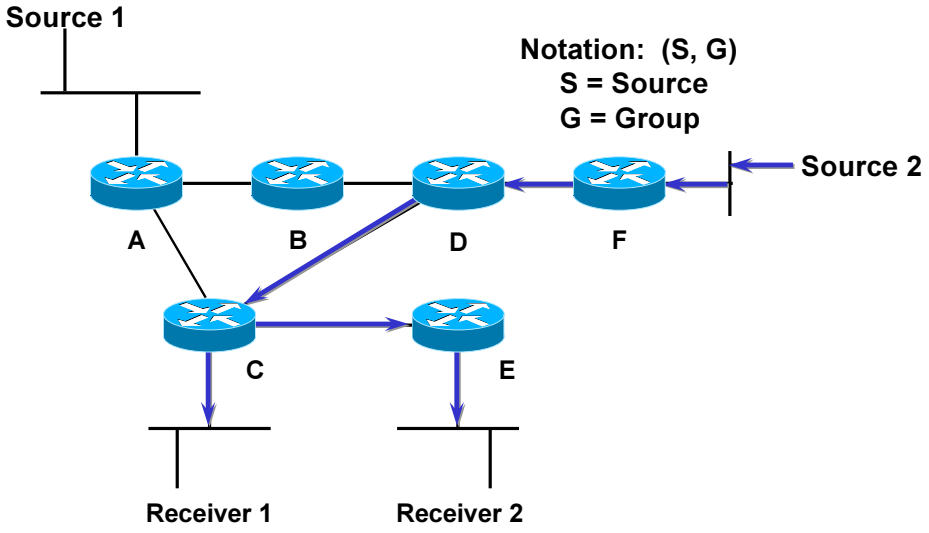
24

- Shortest path or source distribution tree is the network path with the least hops
  - Ex: shortest path between Source and Receiver 1 is via Routers A and C, and shortest path to Receiver 2 is one additional hop via Router E

Multicast Distribution Trees

Cisco.com

Shortest Path or Source Distribution Tree



RST-1701
9783_05_2004_X2

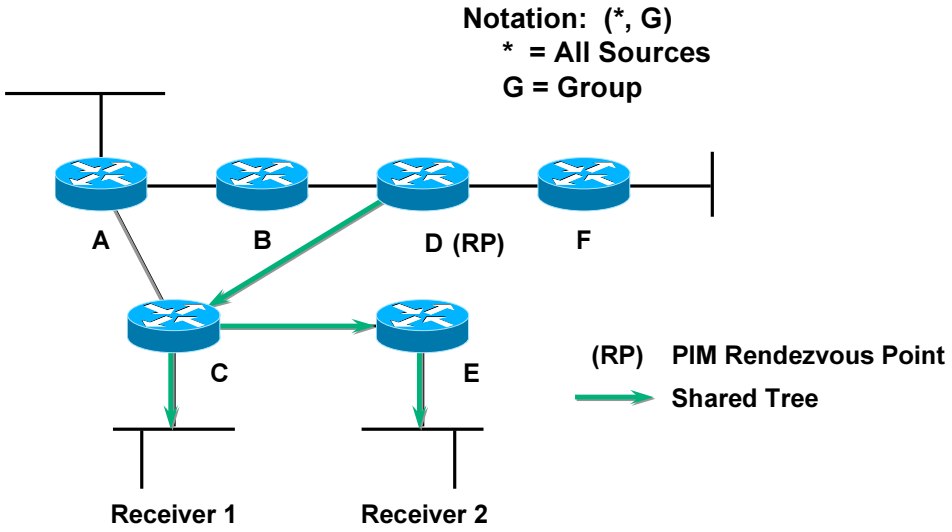
25

- Shortest path or source distribution tree is the network path with the least hops
  - Ex: shortest path between Source and Receiver 1 is via Routers A and C, and shortest path to Receiver 2 is one additional hop via Router E

Multicast Distribution Trees

Cisco.com

Shared Distribution Tree



RST-1701
9783_05_2004_X2

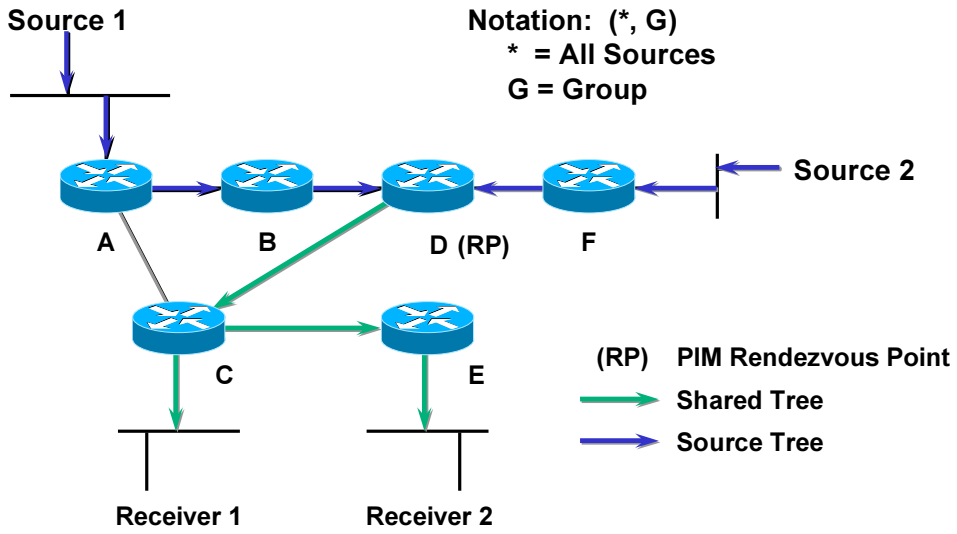
26

- Shared distribution tree is the path derived from a shared point through which sources and receivers must send/receive multicast data
  - Ex: regardless of location/number of receivers, senders register with Shared Root (Router D) and always send a single copy of multicast data through the Shared Root to registered receivers
  - Ex: regardless of location/number of sources, group members always receive forwarded multicast data from Shared Root (Router D)

Multicast Distribution Trees

Cisco.com

Shared Distribution Tree



RST-1701
9783_05_2004_X2

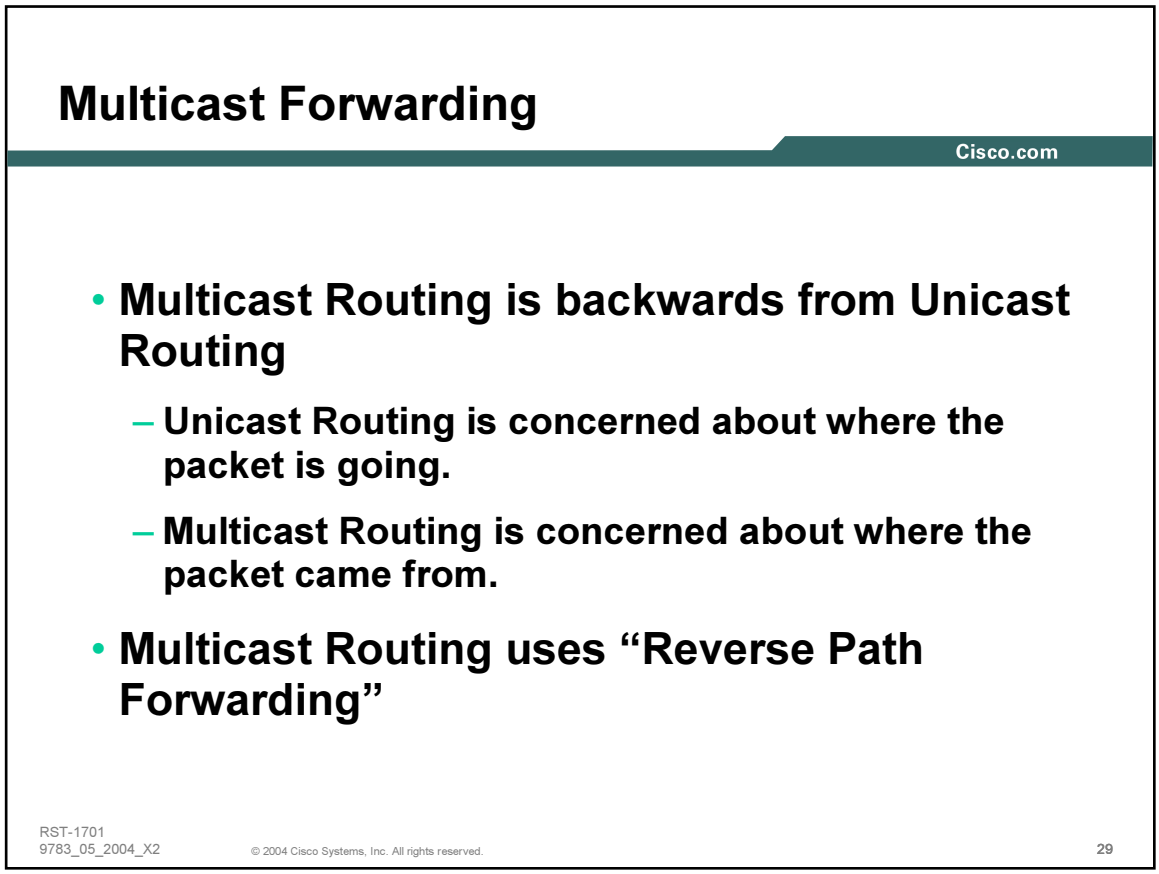
27

- Shared distribution tree is the path derived from a shared point through which sources and receivers must send/receive multicast data
  - Ex: regardless of location/number of receivers, senders register with Shared Root (Router D) and always send a single copy of multicast data through the Shared Root to registered receivers
  - Ex: regardless of location/number of sources, group members always receive forwarded multicast data from Shared Root (Router D)

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**

- Source or Shortest Path Tree Characteristics
  - Provides optimal path (shortest distance and minimized delay) from source to all receivers, but requires more memory to maintain
- Shared Tree Characteristics
  - Provides suboptimal path (may not me shortest distance and may introduce extra delay) from source to all receivers, but requires less memory to maintain



- Multicast Forwarding
  - Routers must know packet origination, rather than destination (opposite of unicast)
    - ... origination IP address denotes known source
    - ... destination IP address denotes unknown group of receivers
  - Multicast routing utilizes Reverse Path Forwarding (RPF)
    - ... Broadcast: floods packets out all interfaces except incoming from source; initially assuming every host on network is part of multicast group
    - ... Prune: eliminates tree branches without multicast group members; cuts off transmission to LANs without interested receivers
    - ... Selective Forwarding: requires its own integrated unicast routing protocol

Multicast Forwarding

Cisco.com

Reverse Path Forwarding (RPF)

- **What is RPF?**

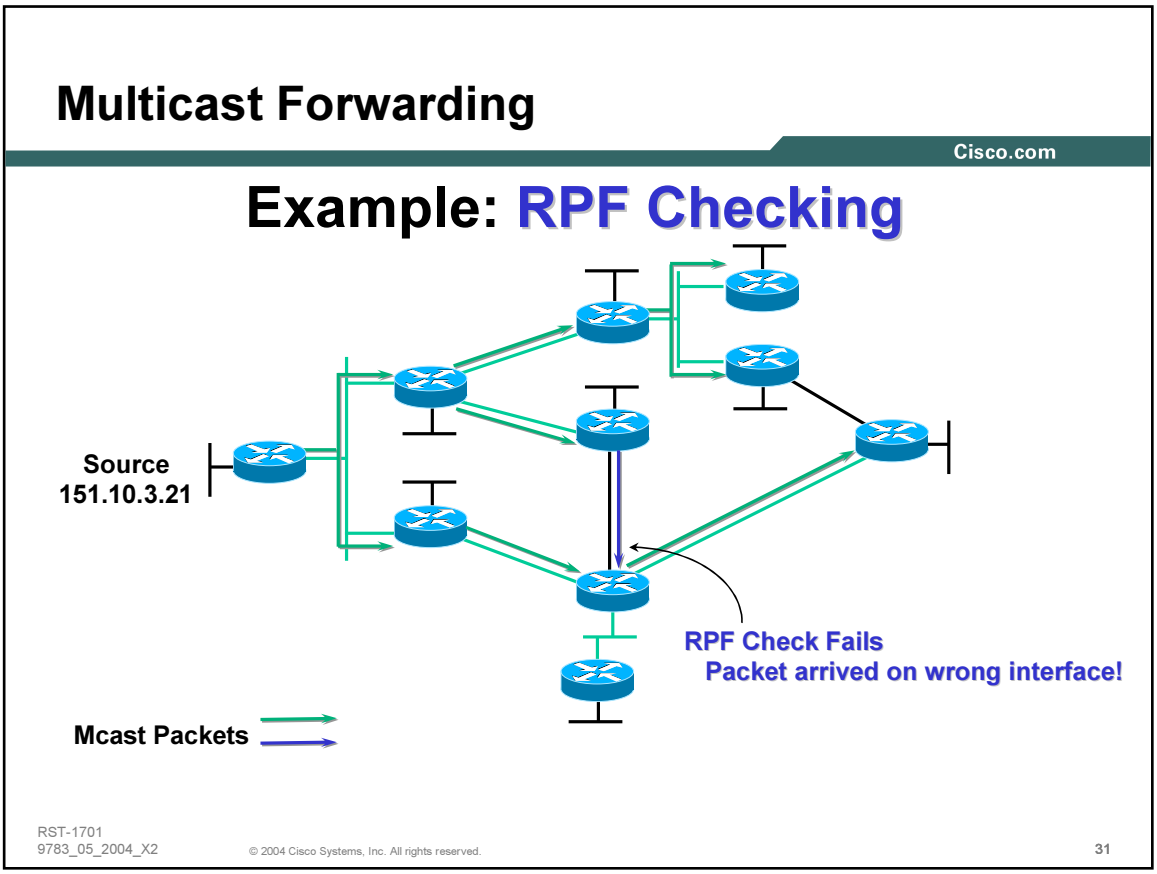
  A router forwards a multicast datagram only if received on the up stream interface to the source (i.e. it follows the distribution tree).

- **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.**

RST-1701
9783_05_2004_X2

30

- Reverse Path Forwarding
  - Routers forward multicast datagrams received from incoming interface on distribution tree leading to source
  - Routers check the source IP address against their multicast routing tables (RPF check); ensure that the multicast datagram was received on the specified incoming interface

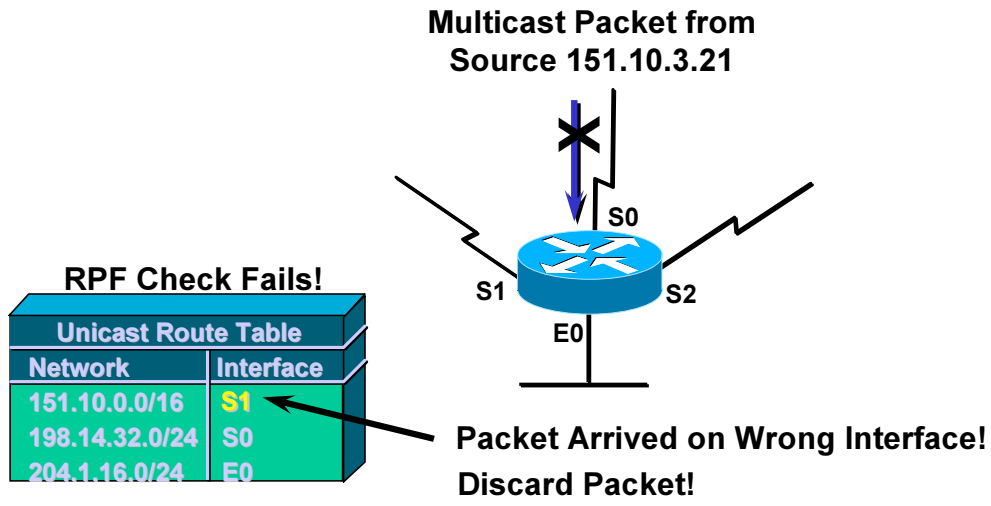


- Multicast Forwarding: RPF Checking
  - Source floods network with multicast data
  - Each router has a designated incoming interface (RPF interface) on which multicast data can be received from a given source
  - Each router receives multicast data on one or more interfaces, but performs RPF check to prevent duplicate forwarding
- Example: Router receives multicast data on two interfaces
  - 1) performs RPF Check on multicast data received on interface E0; RPF Check succeeds because data was received on specified incoming interface from source 151.10.3.21; data forwarded through all outgoing interfaces on the multicast distribution tree
  - 2) performs RPF Check on multicast data received on interface E1; RPF Check fails because data was not received on specified incoming interface from source 151.10.3.21; data silently dropped

Multicast Forwarding

Cisco.com

A closer look: RPF Check Fails



RST-1701
9783_05_2004_X2

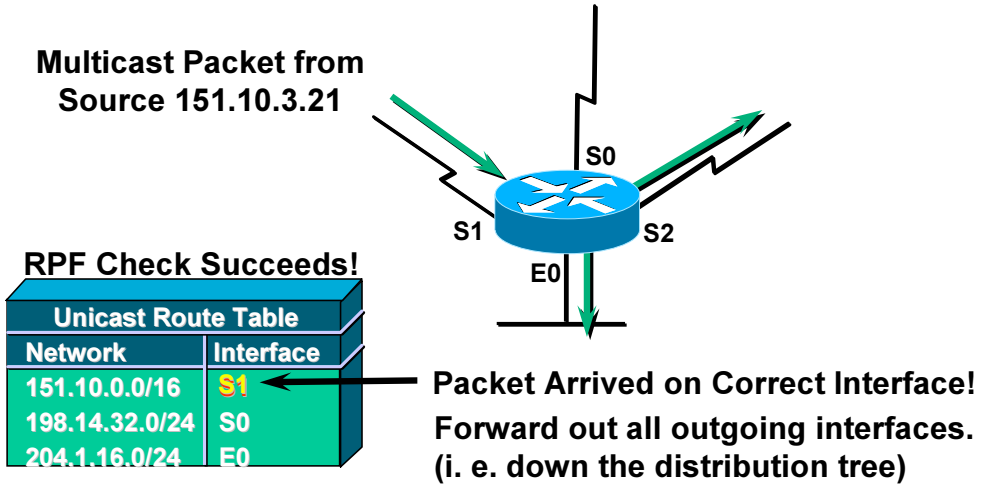
32

- Multicast Forwarding: RPF Check Fails
  - Ex: Router can only accept multicast data from Source 151.10.3.21 on interface S1
    - ... multicast data is silently dropped because it arrived on an interface not specified in the RPF check (S0)

Multicast Forwarding

Cisco.com

A closer look: RPF Check Succeeds



RST-1701
9783_05_2004_X2

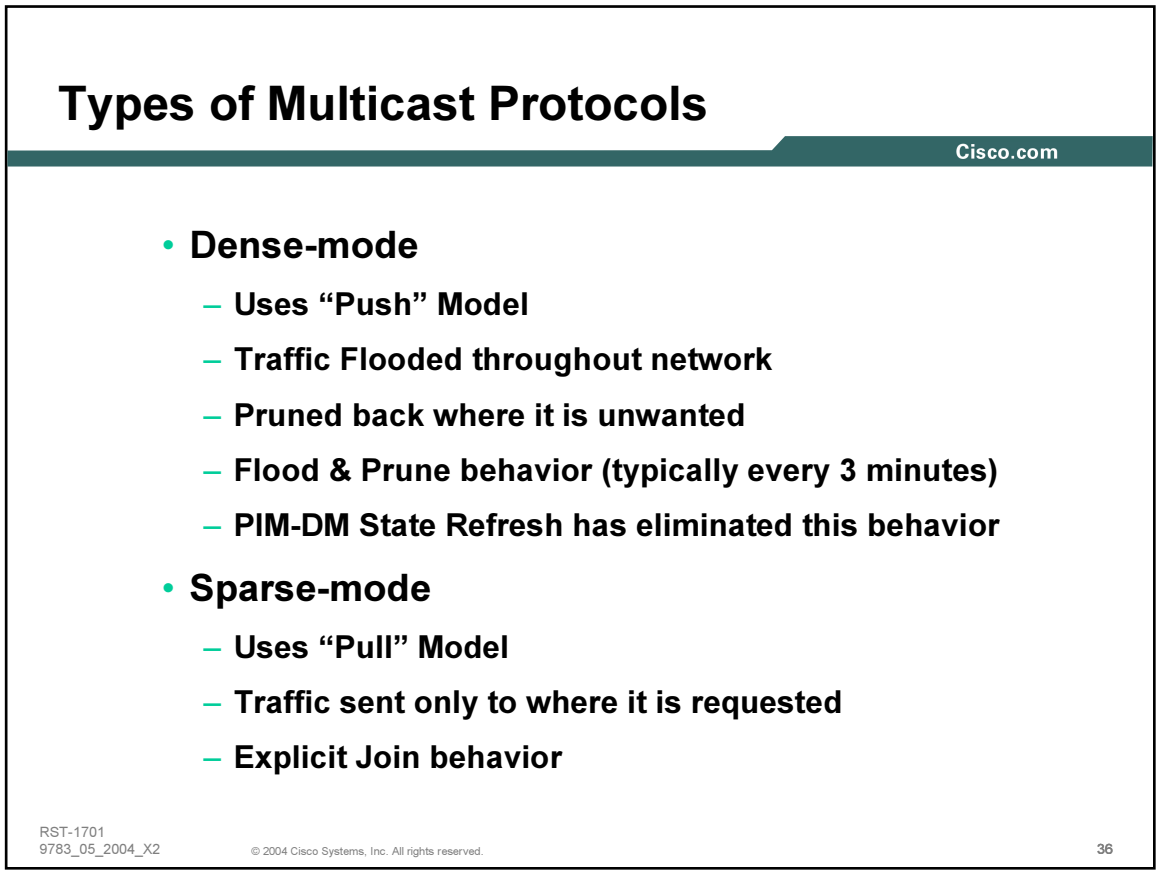
33

- Multicast Forwarding: RPF Check Succeeds
  - Ex: Router can only accept multicast data from Source 151.10.3.21 on interface S1
    - ... multicast data is forwarded out all outgoing on the distribution tree because it arrive on the incoming interface specified in the RPF check (S1)

“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.”

Multicast vs. Unicast Routing

PIM PROTOCOLS



- Dense-mode multicast protocols
  - Initially flood/broadcast multicast data to entire network, then prune back paths that don't have interested receivers
- Sparse-mode multicast protocols
  - Assumes no receivers are interested unless they explicitly ask for it
- Sparse-dense mode multicast protocols
  - Behaves in manner that is appropriate for distribution of receiver group (sparse or dense, or any combination thereof)

PIM-DM

Cisco.com

- **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...**
  - **Lab work and router performance testing**

RST-1701
9783_05_2004_X2

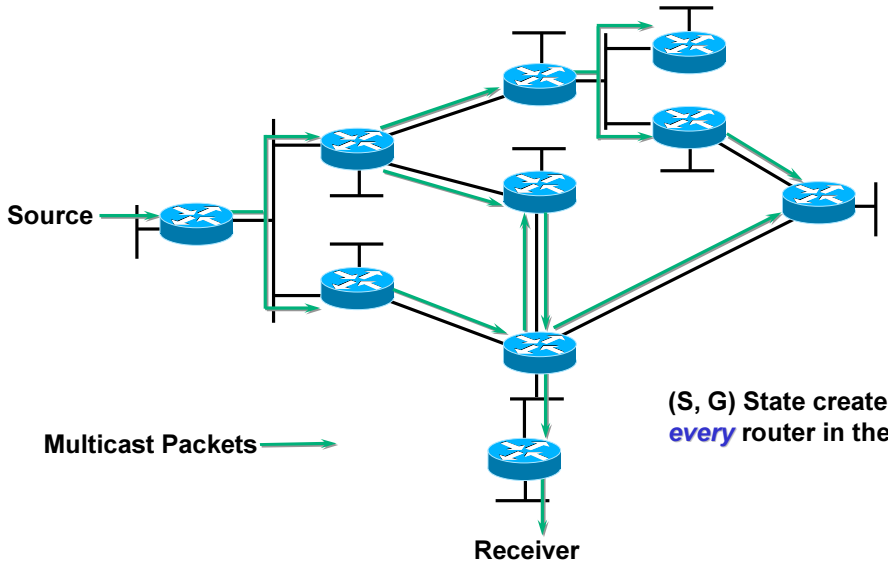
37

- Protocol Independent Multicast (PIM) Dense-mode (Internet-draft)
  - Uses Reverse Path Forwarding (RPF) to flood the network with multicast data, then prune back paths based on uninterested receivers
    - Interoperates with DVMRP
- Appropriate for
  - Densely distributed receivers located in close proximity to source
  - Few senders -to- many receivers (due to frequent flooding)
  - High volume of multicast traffic
  - Constant stream of traffic

PIM-DM Flood and Prune

Cisco.com

Initial Flooding



RST-1701
9783_05_2004_X2

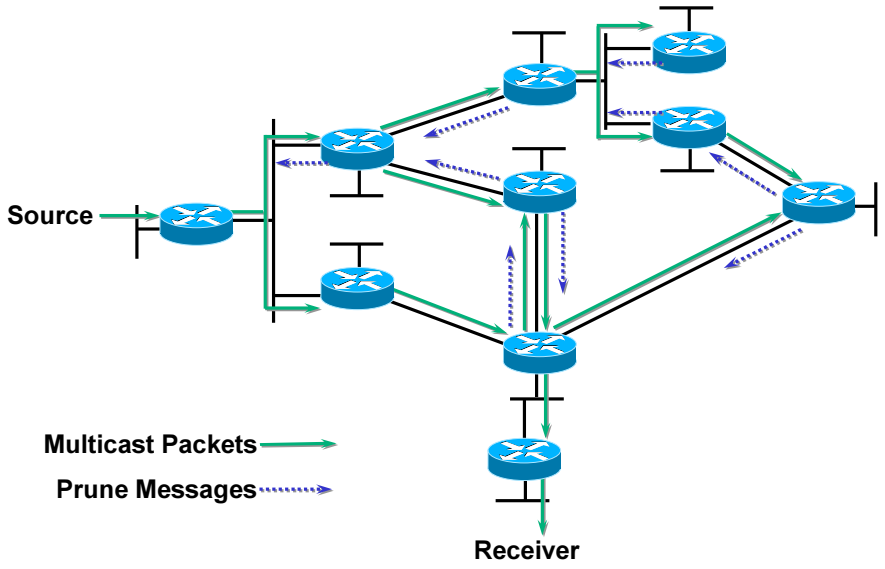
38

- PIM-DM Initial Flooding
  - PIM-DM is similar to DVMRP in that it initially floods multicast traffic to all parts of the network.
  - However unlike DVMRP, which pre-builds a “Truncated Broadcast Tree” that is used for initial flooding, PIM-DM initially floods traffic out ALL non RPF interfaces where there is:
    - Another PIM-DM neighbor or
    - A directly connected member of the group
  - The reason that PIM-DM does not use “Truncated Broadcast Trees” to pre-build a spanning tree for each source network is that this would require running a separate routing protocol as does DVMRP. (At the very least, some sort of Poison-Reverse messages would have to be sent to build the TBT.) Instead, PIM-DM uses other mechanisms to prune back the traffic flows and build Source Trees.
- Initial Flooding Example
  - In this example, multicast traffic being sent by the source is flooded throughout the entire network.
  - As each router receives the multicast traffic via its RPF interface (the interface in the direction of the source), it forwards the multicast traffic to all of its PIM-DM neighbors.

PIM-DM Flood and Prune

Cisco.com

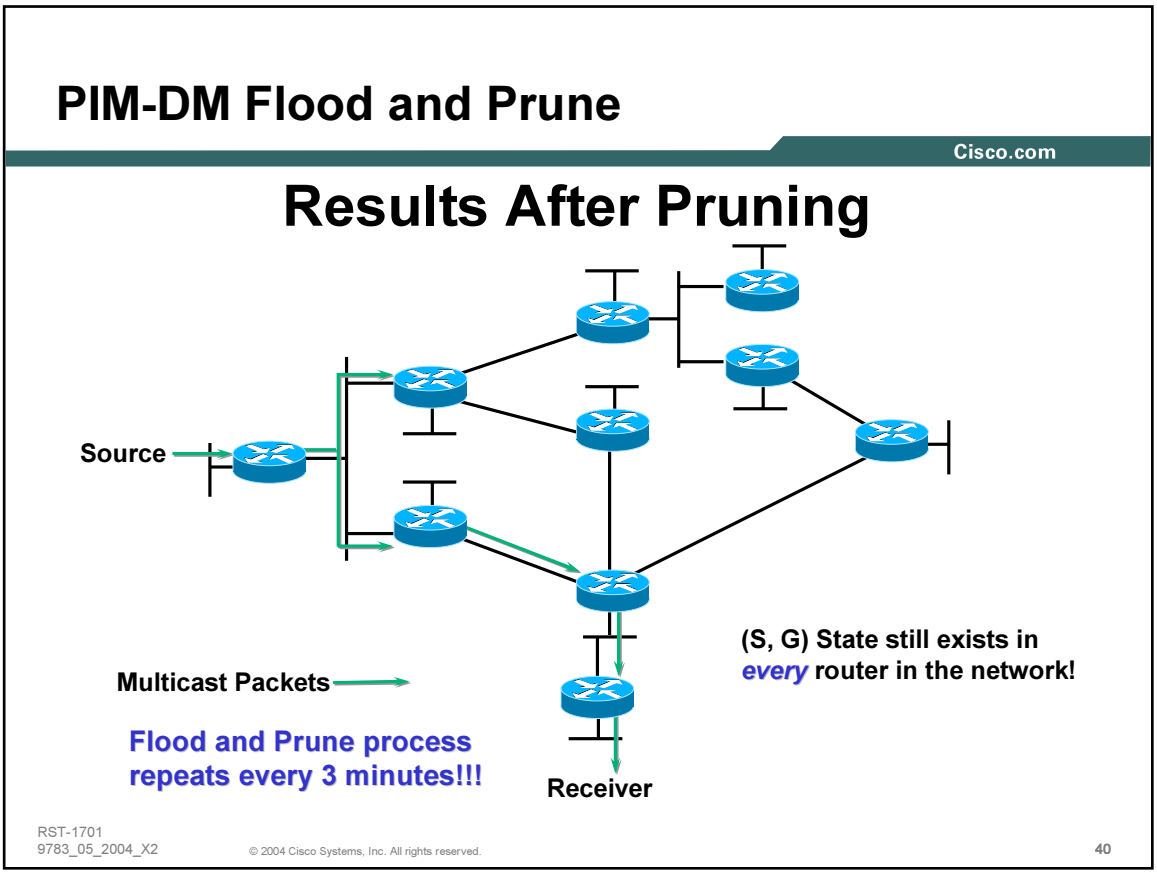
Pruning Unwanted Traffic



RST-1701
9783_05_2004_X2

39

- Pruning unwanted traffic
  - In the example above, PIM Prunes (denoted by the dashed arrows) are sent to stop the flow of unwanted traffic.
  - Prunes are sent on the RPF interface when the router has no downstream members that need the multicast traffic.
  - Prunes are also sent on non-RPF interfaces to shutoff the flow of multicast traffic that is arriving via the wrong interface (i.e. traffic arriving via an interface that is not in the shortest path to the source.)
    - An example of this can be seen at the second router from the receiver near the center of the drawing. Multicast traffic is arriving via a non-RPF interface from the router above (in the center of the network) which results in a Prune message.



- Results after Pruning
  - In the final drawing in our example shown above, multicast traffic has been pruned off of all links except where it is necessary. This results in a Shortest Path Tree (SPT) being built from the Source to the Receiver.
  - Even though the flow of multicast traffic is no longer reaching most of the routers in the network, (S, G) state still remains in ALL routers in the network. This (S, G) state will remain until the source stops transmitting.
  - In PIM-DM, Prunes expire after three minutes. This causes the multicast traffic to be re-flooded to all routers just as was done in the “Initial Flooding” drawing. This periodic (every 3 minutes) “Flood and Prune” behavior is normal and must be taken into account when the network is designed to use PIM-DM.

PIM-DM: Evaluation

Cisco.com

- **Primary use:**
  - **Test Labs and router performance testing**
- **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**

RST-1701
9783_05_2004_X2

41

- Evaluation: PIM Dense-mode
  - Appropriate for large number of densely distributed receivers located in close proximity to source
  - Advantages
    - Minimal number of commands required for configuration (two)
    - Simple mechanism for reaching all possible receivers and eliminating distribution to uninterested receivers
    - Simple behavior is easier to understand and therefore easier to debug
    - Interoperates with DVMRP
  - Potential issues
    - Necessity to flood frequently because prunes expire after 3 minutes.

PIM-SM (RFC 2362)

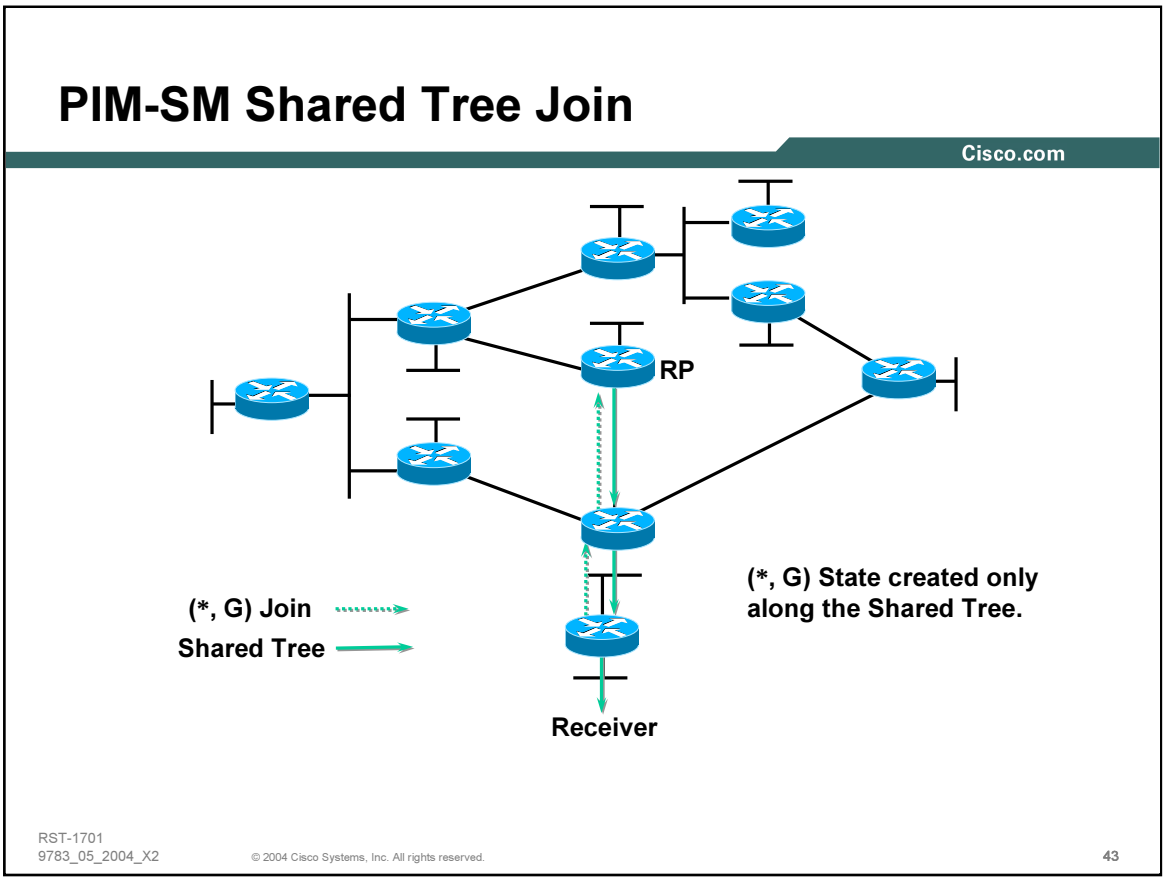
Cisco.com

- **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.**

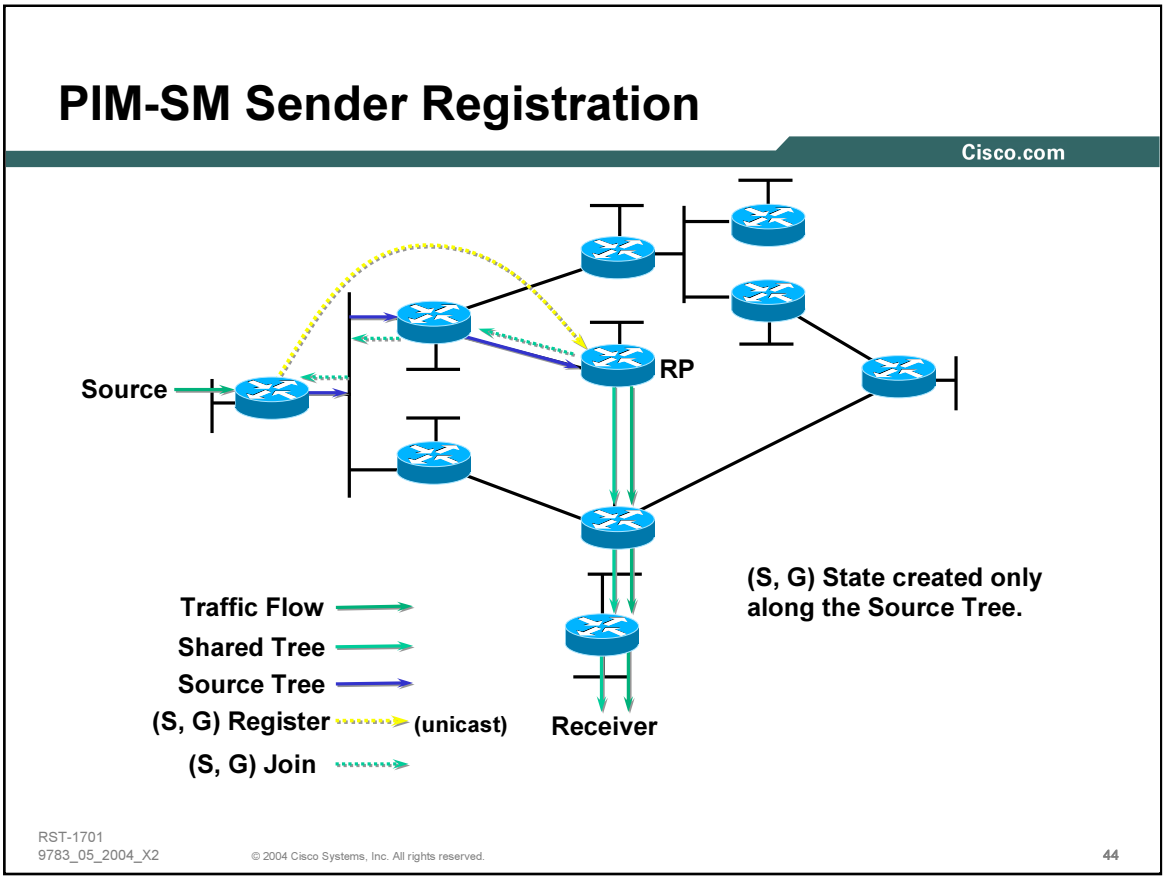
RST-1701
9783_05_2004_X2

42

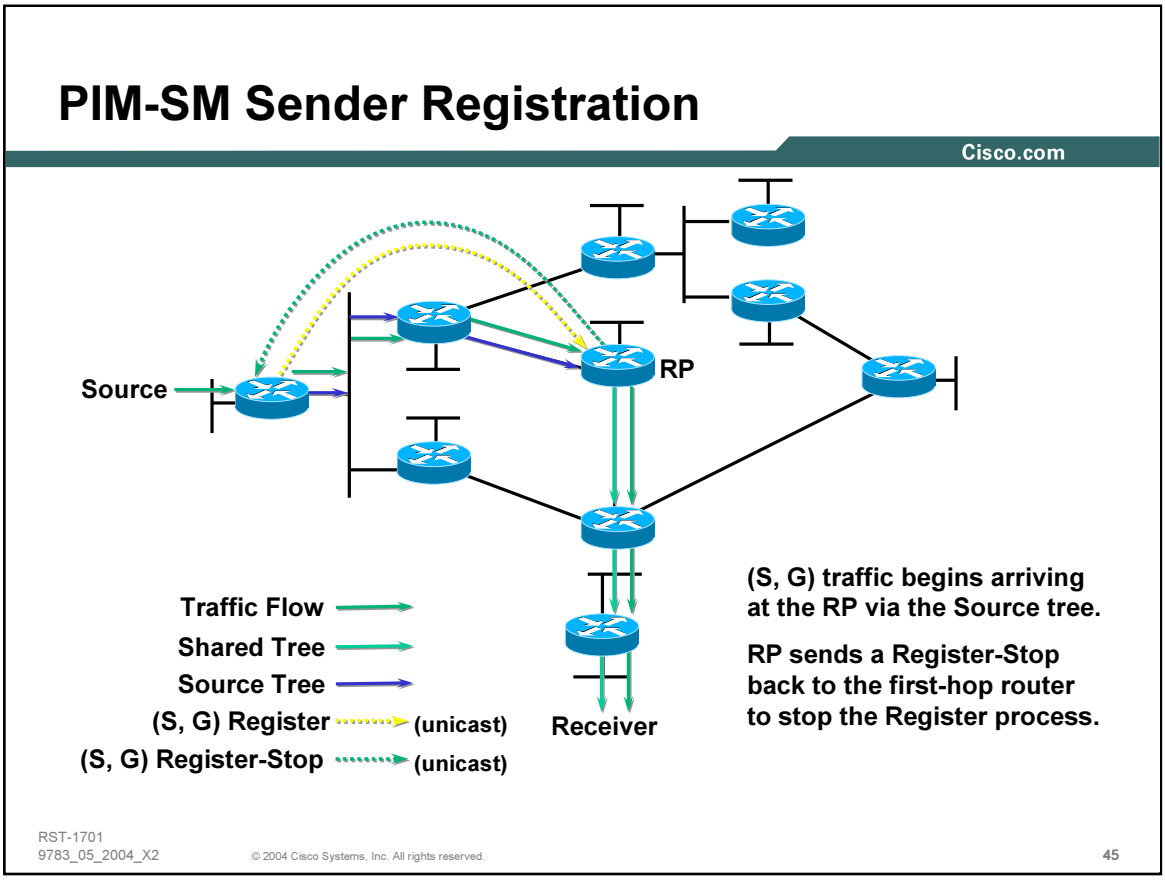
- Protocol Independent Multicast (PIM) Sparse-mode (RFC 2117)
  - Utilizes a rendezvous point (RP) to coordinate forwarding from source to receivers
    - Regardless of location/number of receivers, senders register with RP and send a single copy of multicast data through it to registered receivers
    - Regardless of location/number of sources, group members register to receive data and always receive it through the RP
  - Appropriate for
    - Multipoint datastreams going to a relatively small number of LANs
    - Few interested receivers per multicast group
    - Senders/receivers sparsely distributed or separated by WAN links
    - Intermittent traffic (no necessity to flood each new session)



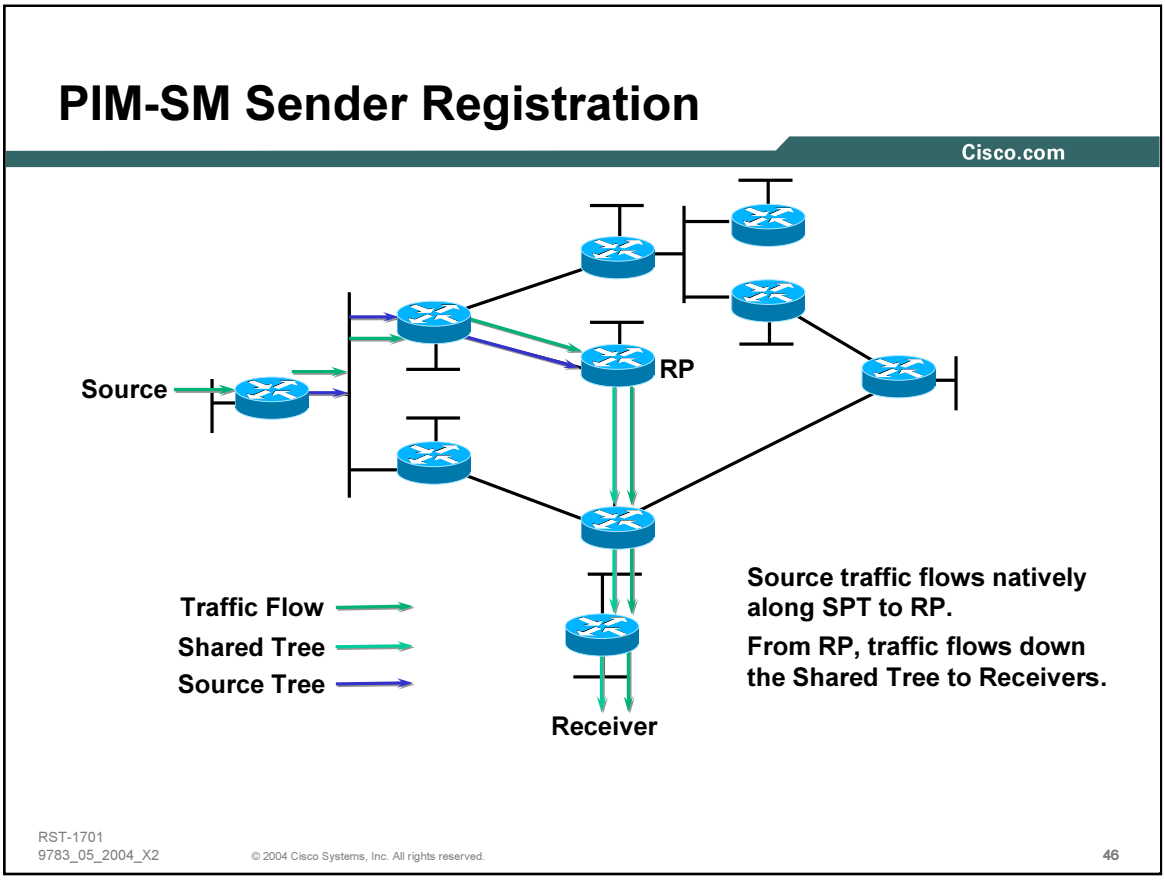
- PIM-SM Shared Tree Joins
  - In this example, there is an active receiver (attached to leaf router at the bottom of the drawing) has joined multicast group “G”.
  - The leaf router knows the IP address of the Rendezvous Point (RP) for group G and when it sends a (*,G) Join for this group towards the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



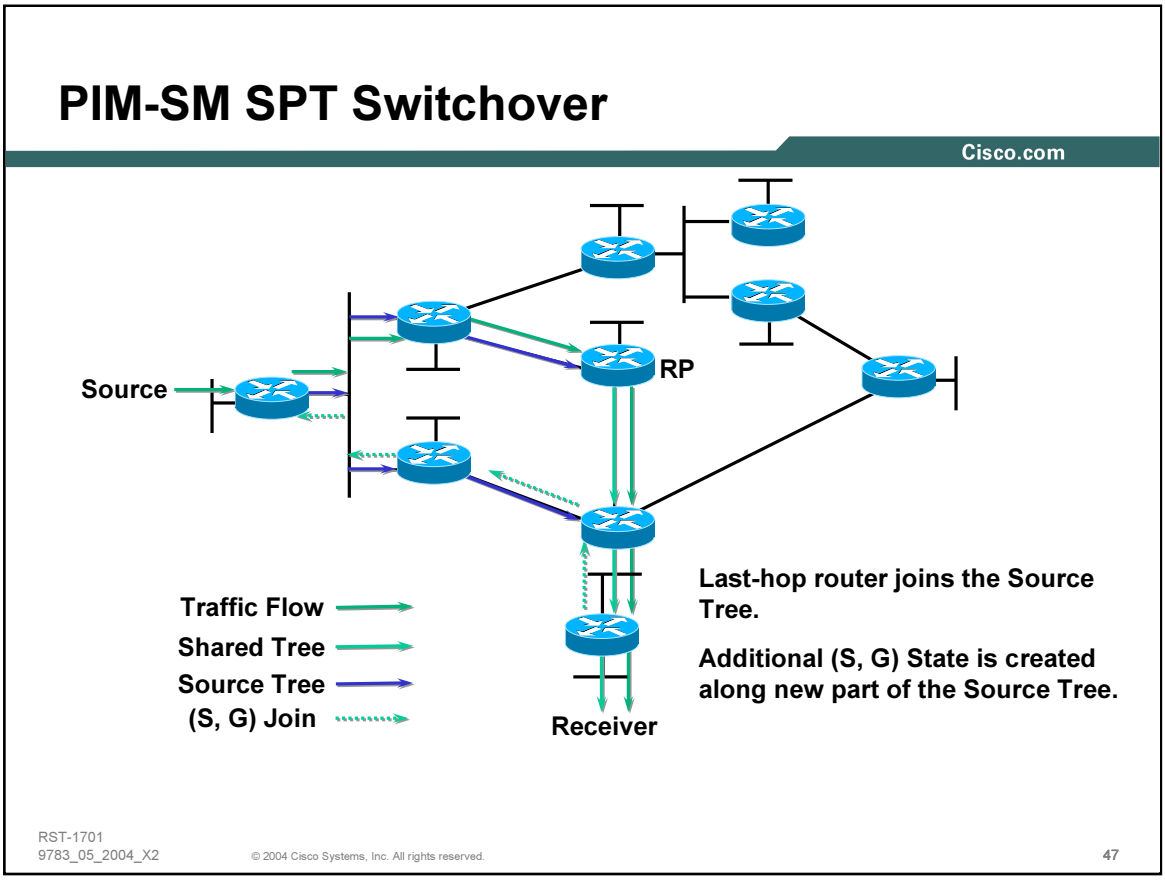
- PIM-SM Sender Registration
  - As soon as an active source for group G sends a packet the leaf router that is attached to this source is responsible for “Registering” this source with the RP and requesting the RP to build a tree back to that router.
  - The source router encapsulates the multicast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



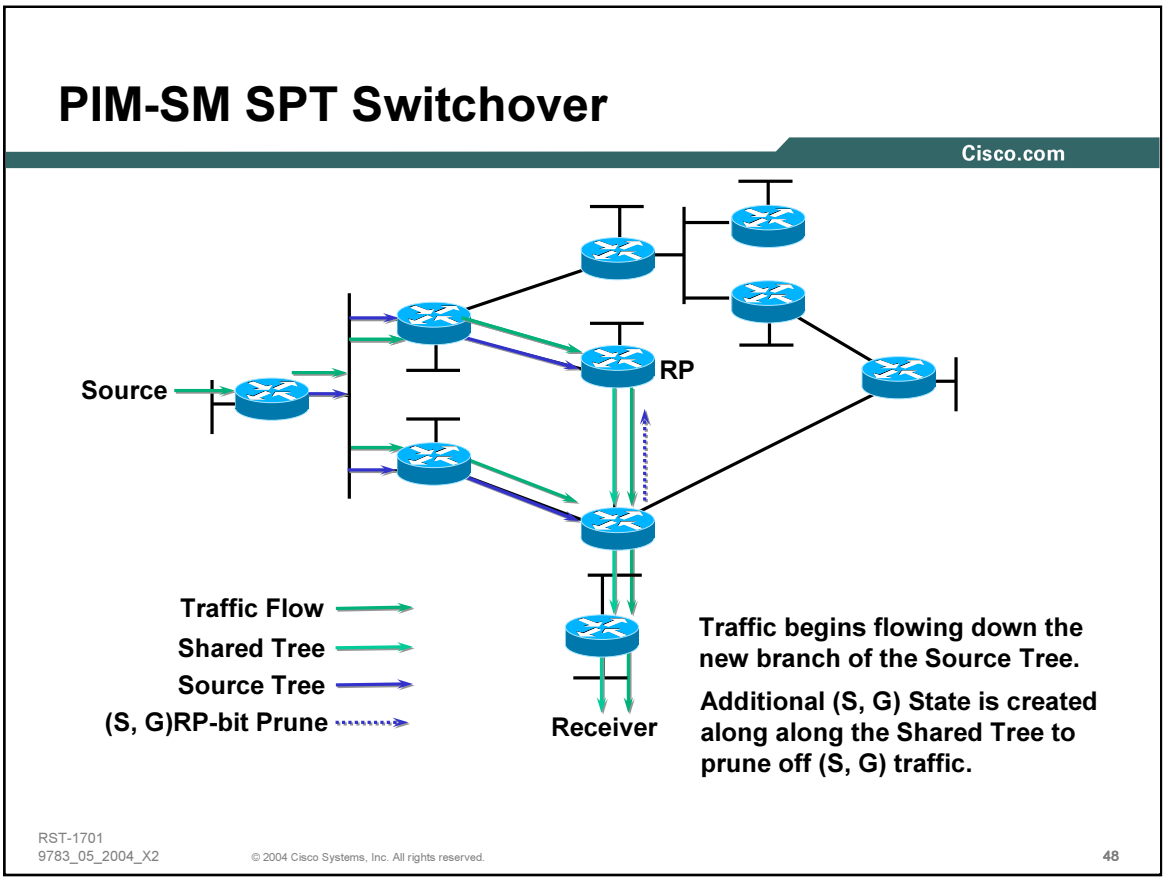
- PIM-SM Sender Registration (cont.)
  - As soon as the SPT is built from the Source router to the RP, multicast traffic begins to flow natively from source S to the RP.
  - Once the RP begins receiving data natively (i.e. down the SPT) from source S it sends a ‘Register Stop’ to the source’s first hop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



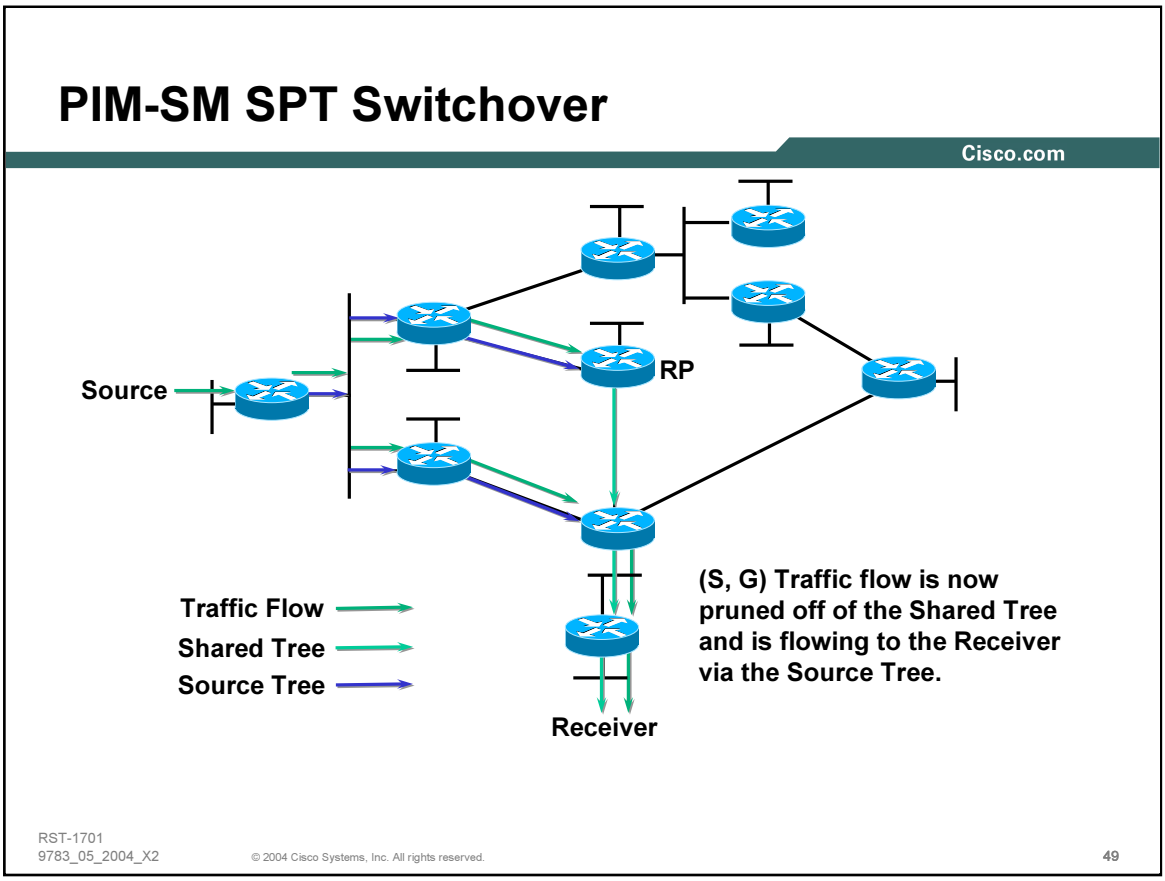
- PIM-SM Sender Registration (cont.)
  - At this point, multicast traffic from the source is flowing down the SPT to the RP and from there, down the Shared Tree to the receiver.



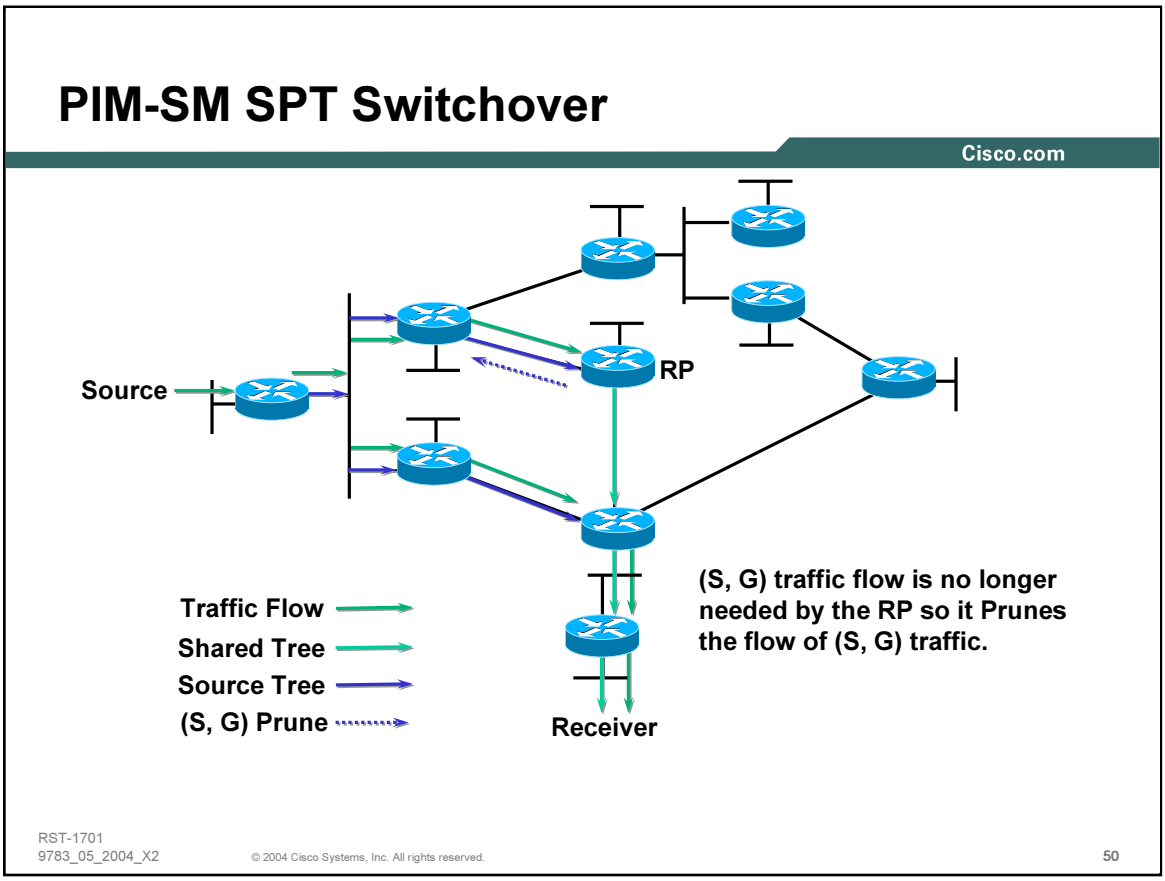
- PIM-SM Shortest-Path Tree Switchover
  - PIM-SM has the capability for last-hop routers (i.e. routers with directly connected members) to switch to the Shortest-Path Tree and bypass the RP if the traffic rate is above a set threshold called the “SPT-Threshold”.
    - The default value of the SPT-Threshold” in Cisco routers is zero. This means that the default behaviour for PIM-SM leaf routers attached to active receivers is to immediately join the SPT to the source as soon as the first packet arrives via the (*,G) shared tree.
  - In the above example, the last-hop router (at the bottom of the drawing) sends an (S, G) Join message toward the source to join the SPT and bypass the RP.
  - This (S, G) Join messages travels hop-by-hop to the first-hop router (i.e. the router connected directly to the source) thereby creating another branch of the SPT. This also creates (S, G) state in all the routers along this branch of the SPT.



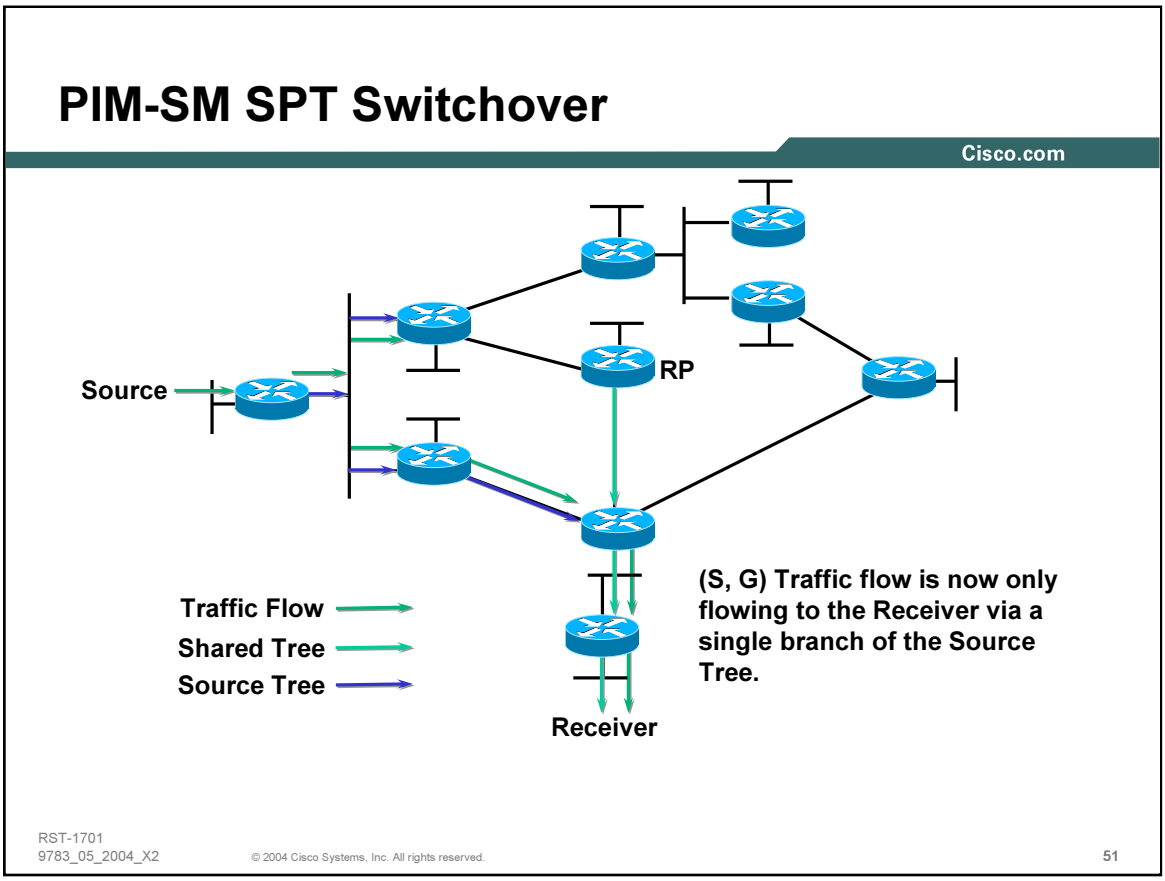
- PIM-SM Shortest-Path Tree Switchover
  - Once the branch of the Shortest-Path Tree has been built, (S, G) traffic begins flowing to the receiver via this new branch.
  - Next, special (S, G)RP-bit Prune messages are sent up the Shared Tree to prune off the redundant (S,G) traffic that is still flowing down the Shared Tree.
    - If this were not done, (S, G) traffic would continue flowing down the Shared Tree resulting in duplicate (S, G) packets arriving at the receiver.



- PIM-SM Shortest-Path Tree Switchover
  - **As the (S, G)RP-bit Prune message travels up the Shared Tree, special (S, G)RP-bit Prune state is created along the Shared Tree that selectively prevents this traffic from flowing down the Shared Tree.**
  - **At this point, (S, G) traffic is now flowing directly from the first-hop router to the last-hop router and from there to the receiver.**



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



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

"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."

PIM-SM Frequently Forgotten Fact

PIM-SM: Evaluation

Cisco.com

- **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, MSDP and/or SSM**

RST-1701
9783_05_2004_X2

53

- Evaluation: PIM Sparse-mode
  - Can be used for sparse of dense distribution of multicast receivers (no necessity to flood)
  - Advantages
    - Traffic sent only to registered receivers that have explicitly joined the multicast group
    - RP can be switched to optimal shortest-path-tree when high-traffic sources are forwarding to a sparsely distributed receiver group
    - Interoperates with DVMRP
  - Potential issues
    - Requires RP during initial setup of distribution tree (can switch to shortest-path-tree once RP is established and determined suboptimal)
    - RPs can become bottlenecks if not selected with great care
    - Complex behavior is difficult to understand and therefore difficult to debug

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)
- RFC 3569 An Overview of Source-Specific Multicast (SSM)

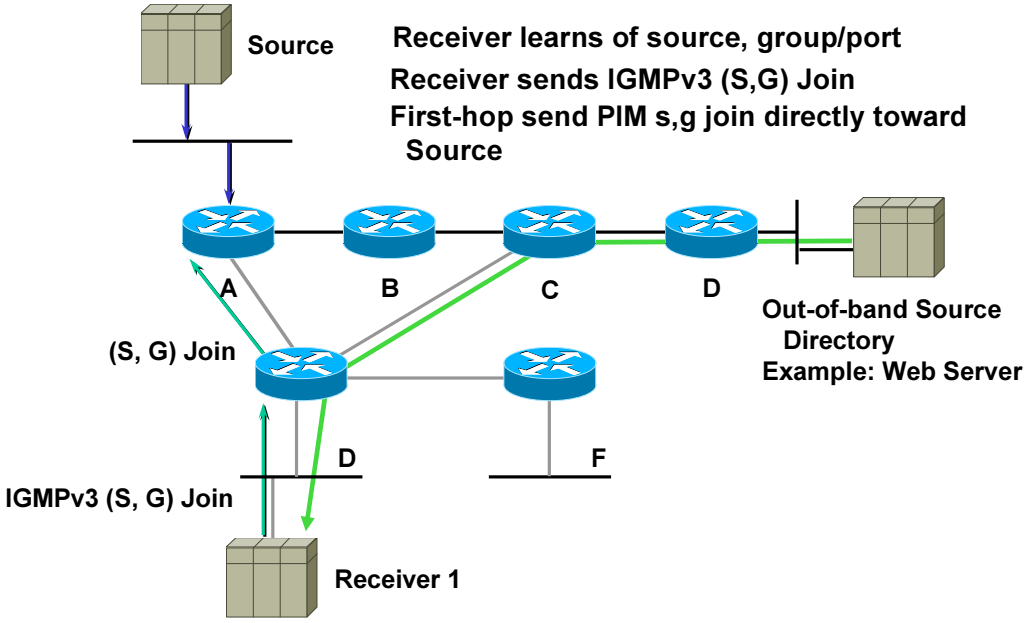
Source Specific Multicast Advantages

Cisco.com

- **Simplifies** Multicast deployment and eliminates the concept of RP and dependence on MSDP for finding sources.
- Optimized and **Reduce latency** for multicast forwarding in case of one to many applications.
- Simplifies **Address allocation** problem for global single source groups.
- Allows immediate use of shortest forwarding path to a specific source, without need to create shared tree.

PIM Source Specific Mode

Cisco.com

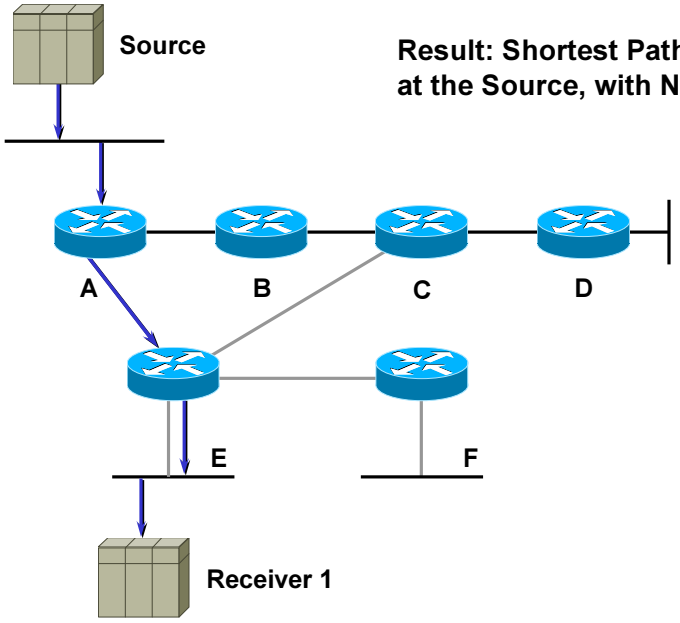


RST-1701
9783_05_2004_X2

56

PIM Source Specific Mode

Cisco.com



RST-1701
9783_05_2004_X2

PIM-SSM: Evaluation

Cisco.com

- Ideal for applications with one source sending to many receivers
- Solves multicast address allocation problems.
  - Flows differentiated by both source and group.
    - Not just by 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.

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

PIM-SM: One Size Fits All?

Cisco.com

- **Effective for sparse or dense distribution of multicast receivers**
- **Advantages:**
  - **Widely deployed**
  - **Source-trees or Shared tree possible**
  - **Basis for current inter-domain multicast with MSDP**

RST-1701
9783_05_2004_X2

60

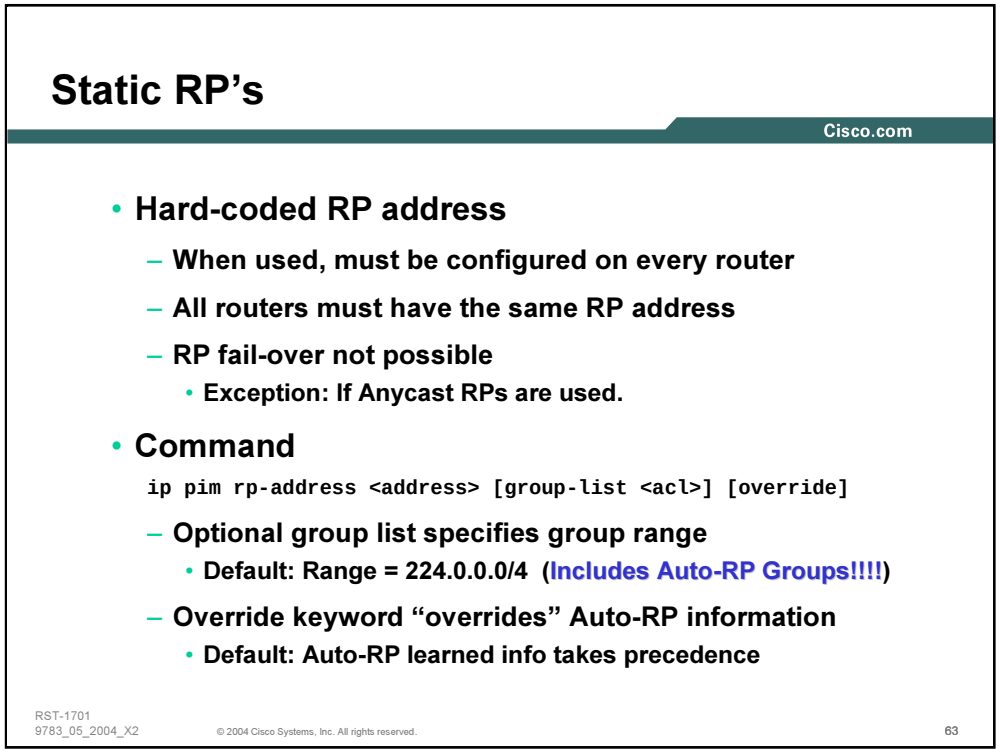
- Evaluation: PIM Sparse-mode
  - Can be used for sparse of dense distribution of multicast receivers (no necessity to flood)
  - Advantages
    - Traffic sent only to registered receivers that have explicitly joined the multicast group
    - RP can be switched to optimal shortest-path-tree when high-traffic sources are forwarding to a sparsely distributed receiver group
    - Interoperates with DVMRP
  - Potential issues
    - Requires RP during initial setup of distribution tree (can switch to shortest-path-tree once RP is established and determined suboptimal)
    - RPs can become bottlenecks if not selected with great care
    - Complex behavior is difficult to understand and therefore difficult to debug

RP CHOICES

How Does the Network Know about the RP ?

Cisco.com

- Static configuration
- AutoRP
- BSR



- Hard-code RP Addresses
  - Requires every router in the network to be manually configured with the IP address of a *single* RP.
  - If this RP fails, there is no way for routers to fail-over to a standby RP.
    - The exception to this rule is if "Anycast-RP's" are in use. This requires MSDP to be running between each RP in the network.
- Command
  - `ip pim rp-address <address> [group-list <acl>] [override]`
  - The 'group-list' allows a group range to be specified.
    - The default is ALL multicast groups or 224.0.0.0/4
    - DANGER, WILL ROBINSON!!!

      The default range includes the Auto-RP groups (224.0.1.39 and 224.0.1.40) which will cause this router to attempt to operate these groups in Sparse mode. This is normally not desirable and can often lead to problems where some routers in the network are trying to run these groups in Dense mode (which is the normal method) while others are trying to use Sparse mode. This will result in some routers in the

Auto-RP Overview

Cisco.com

- **All routers automatically learn RP address**
  - **No configuration necessary except on:**
    - **Candidate RPs**
    - **Mapping Agents**
- **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**
  - ***Warning: Can fall back into Dense mode if misconfigured.***
- **Can be used with Admin-Scoping**

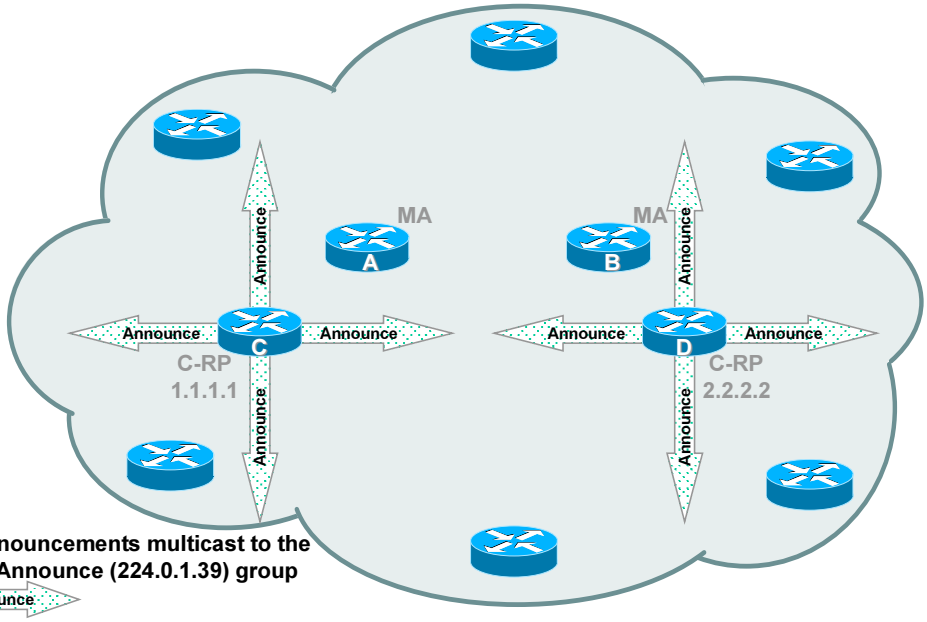
RST-1701
9783_05_2004_X2

64

- Auto-RP Overview
  - Auto-RP allows all routers in the network to automatically "learn" Group-to-RP mappings.
  - There are no special configuration steps that must be taken except on the router(s) that are to function as:
    - Candidate RP's
    - Mapping Agents
  - Multicast is used to distribute Group-to-RP mapping information via two special, IANA assigned multicast groups.
    - Cisco-Announce Group - 224.0.1.39
    - Cisco-Discovery Group - 224.0.1.40
  - Because multicast is used to distribute this information, a "Chicken and Egg" situation can occur if the above groups operate in Sparse mode. (Routers would have to know a priori what the address of the RP is before they can learn the address of the RP(s) via Auto-RP messages.) Therefore, it is recommend that these groups always run in Dense mode so that this information is flooded throughout the network.
  - Multiple Candidate RP's may be defined so that in the case of an RP failure, the other Candidate RP can assume the responsibility of RP.

Auto-RP: From 10,000 Feet

Cisco.com



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

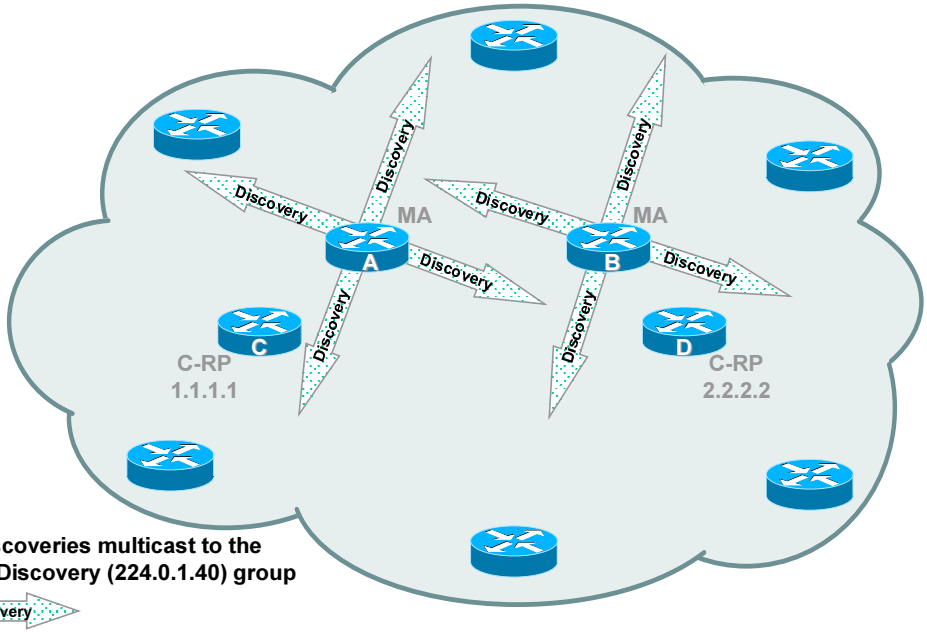
Announce

RST-1701
9783_05_2004_X2

65

Auto-RP: From 10,000 Feet

Cisco.com

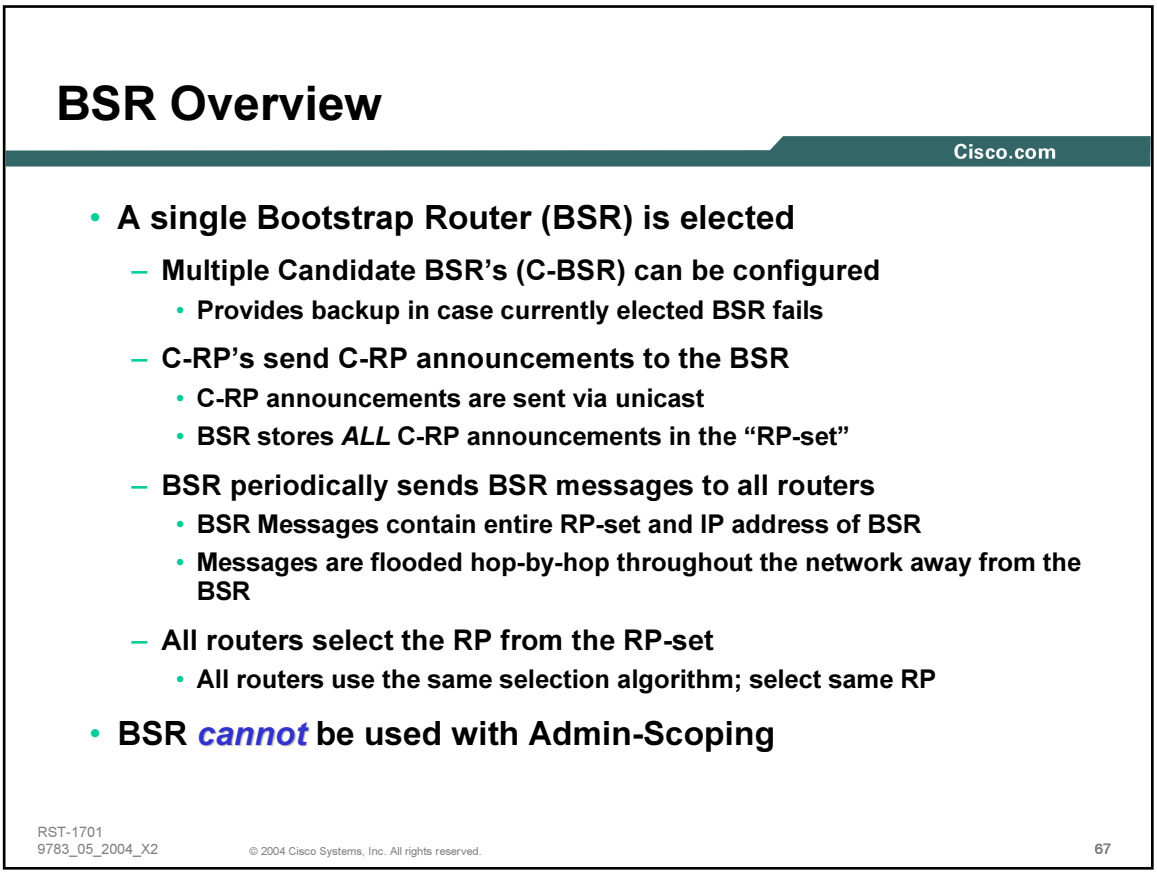


RP-Discoveries multicast to the Cisco Discovery (224.0.1.40) group

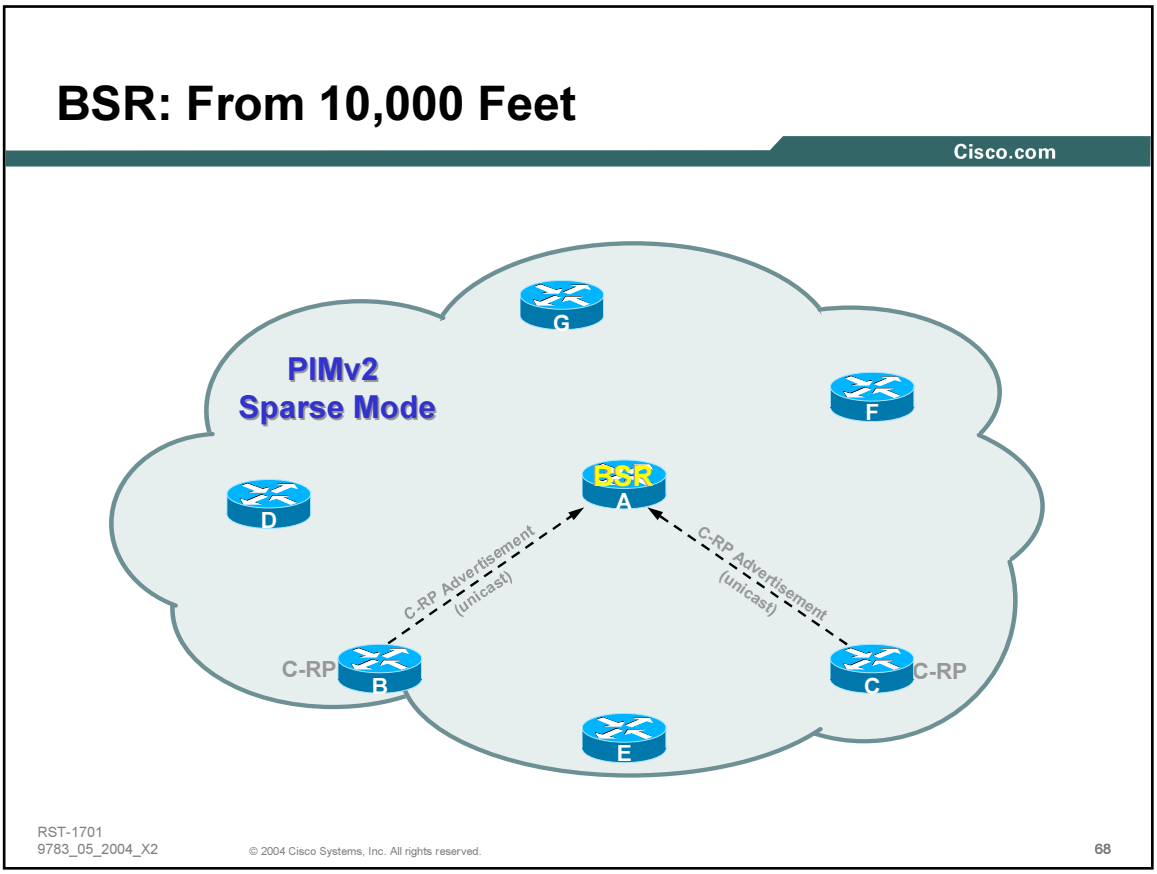
Discovery

RST-1701
9783_05_2004_X2

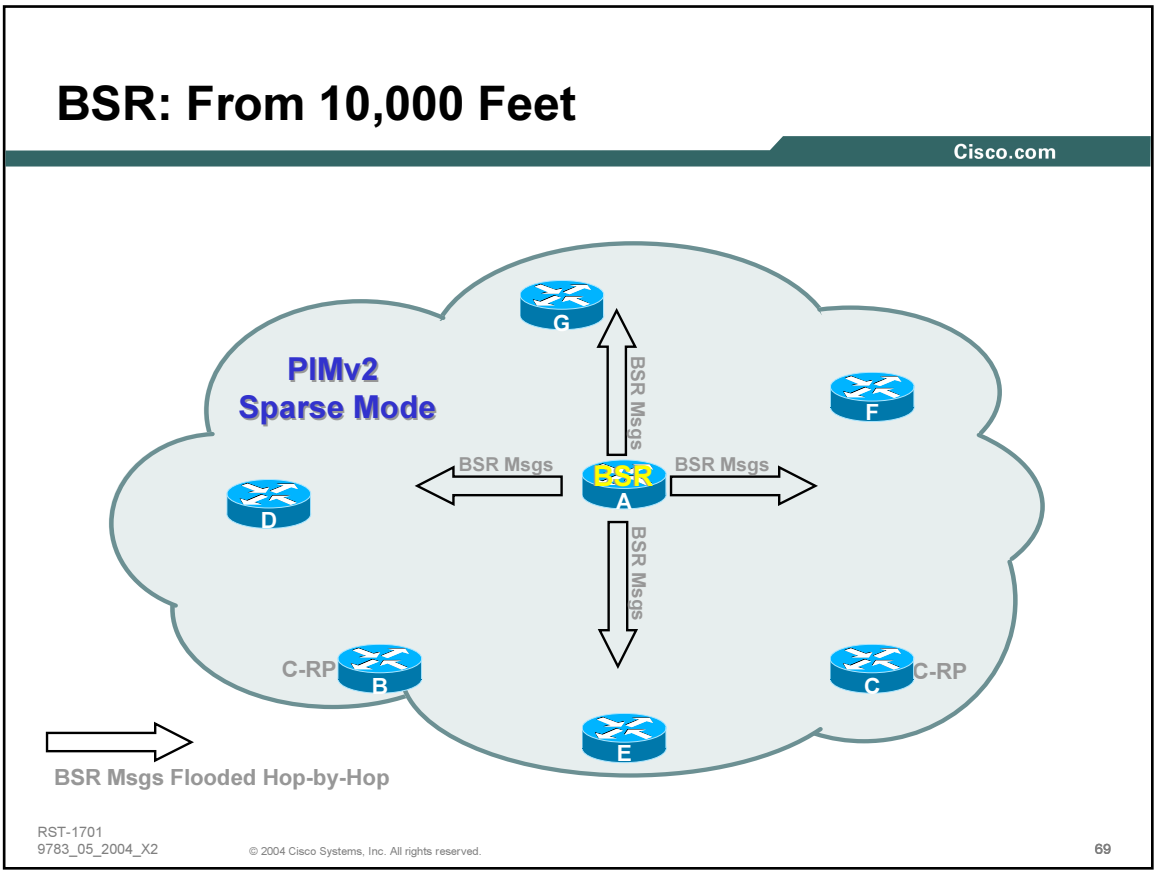
66



- BSR Overview
  - Bootstrap Router (BSR)
    - A single router is elected as the BSR from a collection of Candidate BSR's.
    - If the current BSR fails, a new election is triggered.
    - The election mechanism is pre-emptive based on C-BSR priority.
  - Candidate RP's (C-RP's)
    - Send C-RP announcements directly to the BSR via unicast. (Note: C-RP's learn the IP address of the BSR via periodic BSR messages.)
    - The BSR stores the complete collection of all received C-RP announcements in a database called the "RP-set".
  - The BSR periodically sends out BSR messages to all routers in the network to let them know the BSR is still alive.
  - BSR messages are flooded hop-by-hop throughout the network.
    - Multicast to the "All-PIM Routers" group (224.0.0.13) with a TTL of 1.
  - BSR messages also contain:
    - The complete "RP-set" consisting of all C-RP announcements.
    - The IP Address of the BSR so that C-RP's know where to send their announcements.
  - All routers receive the BSR messages being flooded throughout the network.
    - Select the active RP for each group range using a common hash algorithm that is run against the RP-set. This results in all routers in the network selecting the same RP for a given group-range.



- Router A is configured as the RP mapping agent - Router A listens to group 224.0.1.39 for announcements from C-RPs. It then sends the group-RP mappings to the group 224.0.1.40 RP discovery group for routers D,E,F,G to discover the group-RP mappings automatically
- Router A could also be the administrative boundary to external routing domains
- Routers A and B are both configured as C-RPs for all groups (group range 224.0.0.0/4) and send candidate announcements to the RP announce group 224.0.1.39
- Auto-RP allows for multiple hot backup C-RPs for a given group range. (Highest IP address C-RP is selected.)



- Router A is configured as the RP mapping agent - Router A listens to group 224.0.1.39 for announcements from C-RPs. It then sends the group-RP mappings to the group 224.0.1.40 RP discovery group for routers D,E,F,G to discover the group-RP mappings automatically
- Router A could also be the administrative boundary to external routing domains
- Routers A and B are both configured as C-RPs for all groups (group range 224.0.0.0/4) and send candidate announcements to the RP announce group 224.0.1.39
- Auto-RP allows for multiple hot backup C-RPs for a given group range. (Highest IP address C-RP is selected.)

MULTICAST AT LAYER 2

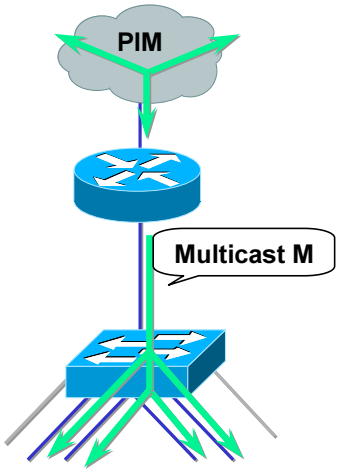
RST-1701
9783_05_2004_X2

L2 Multicast Frame Switching

Cisco.com

Problem: Layer 2 Flooding of Multicast Frames

- **Typical L2 switches treat multicast traffic as unknown or broadcast and must "flood" the frame to every port**
- **Static entries can sometimes be set to specify which ports should receive which group(s) of multicast traffic**
- **Dynamic configuration of these entries would cut down on user administration**



RST-1701
9783_05_2004_X2

71

- L2 Multicast Switching
  - For most L2 Switches, Multicast traffic is normally treated like an unknown MAC address or Broadcast frame which causes the frame to be flooded out every port within a VLAN at rates of over 1 Mpps. This is fine for unknowns and broadcasts but as we have seen earlier, IP Multicast hosts may join and be interested in only specific multicast groups. Again, on most L2 Switches, all this traffic is forwarded out all ports resulting in wasted bandwidth on both the segments and on the end stations.

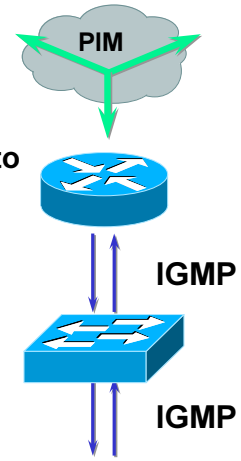
    One way around this on Catalyst Switches is using the Command Line Interface to program the switch manually to associate a multicast MAC address with say ports 5,6,7 so only ports 5,6,and 7 receive the multicast traffic destined for the multicast group. This works fine but again we know IP Multicast hosts dynamically join and leave groups using IGMP to signal to the Multicast Router. This static way of entering the multicast information is not very scaleable. Dynamic configuration of the Switches' forwarding tables would be a better idea, and cut down on user administration.

L2 Multicast Frame Switching

Cisco.com

Solution 1: IGMPv1-v2 Snooping

- Switches become “IGMP” aware
- IGMP packets intercepted by the NMP or by special hardware ASICs
  - Requires special hardware to maintain throughput
- Switch must examine contents of IGMP messages to determine which ports want what traffic
  - IGMP membership reports
  - IGMP leave messages
- Impact on low-end Layer-2 switches:
  - Must process ALL Layer 2 multicast packets
  - Admin. load increases with multicast traffic load
  - Generally results in switch *Meltdown* !!!



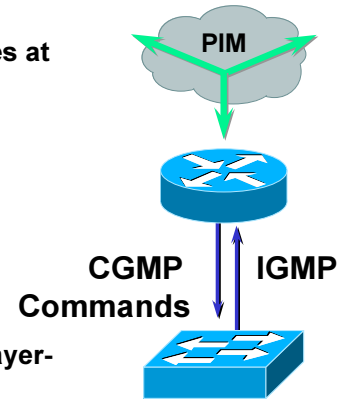
RST-1701
9783_05_2004_X2

L2 Multicast Frame Switching

Cisco.com

Solution 2: CGMP—Cisco Group Management Protocol

- Runs on both the switches and the router
- Router sends CGMP multicast packets to the switches at a well known multicast MAC address:
  - 0100.0cdd.dddd
- CGMP packet contains :
  - Type field—Join or Leave
  - MAC address of the IGMP client
  - Multicast address of the group
- Switch uses CGMP packet info to add or remove a Layer-2 entry for a particular multicast MAC address



RST-1701
9783_05_2004_X2

73

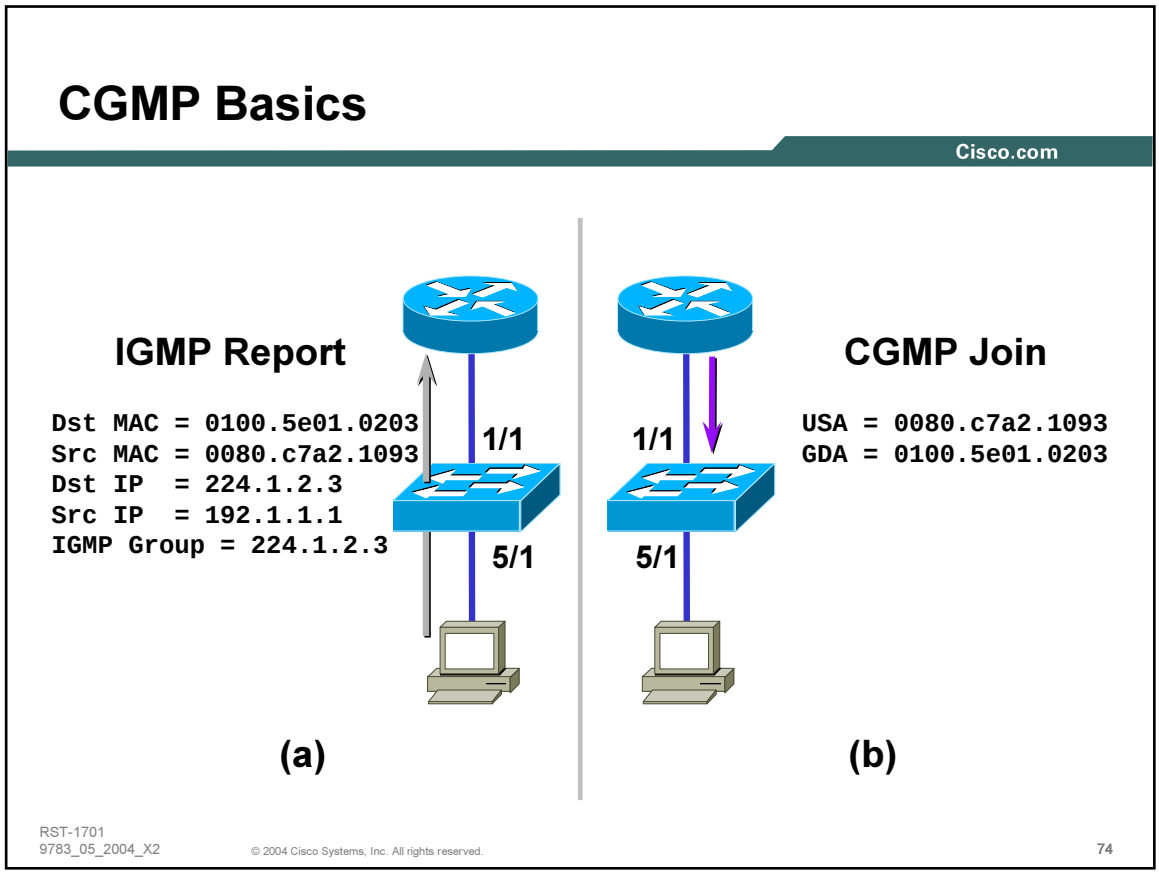
- CGMP
  - CGMP is based on a client server model where the router can be considered a CGMP server and the switch taking on the client role. There are software components running on both devices, with the router translating IGMP messages into CGMP commands which are then executed on the Catalyst 5000 NMP and used to program the EARL's forwarding tables with the correct Multicast entries.

    Since the hosts and routers use well-known IP Multicast Addresses, the EARL can be preprogrammed to direct IGMP Control packets both to the router and the NMP. We will see the NMPs use of these IGMP control packets in a later slide.

    The basis of CGMP is that the IP Multicast router sees all IGMP packets and therefore can inform the switch when specific hosts join or leave Multicast groups. The switch then uses this information to program it's forwarding table.

    When the router sees an IGMP control packet it creates a CGMP packet that contains the request itype which can be a join or a leave, the Multicast Group Address, and the actual MAC address of the client.

    This packet is sent to a well known address which all switches listen on. It is then interpreted and the proper entries created in the



- In this example - the client will asynchronously sent a report when it wants to join the group. This is a formal function of IGMP so the client doesn't have to wait for a general query.
- The debug shows the arrival of the IGMP join
- The creation of group state based on the IGMP join will trigger the router to obtain and begin sending that group onto the client's VLAN

L2 Multicast Frame Switching

Cisco.com

- Impact of IGMPv3 on IGMP Snooping
  - IGMPv3 Reports sent to separate group (224.0.0.22)
    - Switches listen to just this group.
    - Only IGMP traffic – no data traffic.
    - *Substantially* reduces load on switch CPU.
    - Permits low-end switches to implement IGMPv3 Snooping
  - No Report Suppression in IGMPv3
    - Enables individual member tracking
  - IGMPv3 supports Source-specific Includes/Excludes
    - Permits (S,G) state to be maintained by switch
      - Currently not implemented by any switches.
    - May be necessary for full IGMPv3 functionality

Summary: Frame Switches

Cisco.com

- IGMP snooping
  - Switches with Layer 3 aware Hardware/ASICs
    - High-throughput performance maintained
    - Increases cost of switches
  - Switches without Layer 3 aware Hardware/ASICs
    - Suffer serious performance degradation or even *Meltdown*!
      - Shouldn't be a problem when IGMPv3 is implemented
- CGMP
  - Requires Cisco routers and switches
  - Can be implemented in low-cost switches

INTERDOMAIN IP MULTICAST

MBGP Overview

Cisco.com

- **MBGP: Multiprotocol BGP**
  - **Defined in RFC 2858 (extensions to BGP)**
  - **Can carry different types of routes**
    - **Unicast**
    - **Multicast**
  - **Both routes carried in same BGP session**
  - **Does not propagate multicast state info**
  - **Same path selection and validation rules**
    - **AS-Path, LocalPref, MED, …**

MBGP Overview

Cisco.com

- Separate BGP tables maintained
  - Unicast Routing Information Base (URIB)
  - Multicast Routing Information Base (MRIB)
- URIB
  - Contains unicast prefixes for unicast forwarding
  - Populated with BGP unicast NLRI
    - AFI = 1, Sub-AFI = 1
- MRIB
  - Contains *unicast* prefixes for RPF checking
  - Populated with BGP multicast NLRI
    - AFI = 1, Sub-AFI = 2

MBGP Overview

Cisco.com

- MBGP allows divergent paths and policies
  - Same IP address holds dual significance
    - Unicast routing information
    - Multicast RPF information
  - For same IPv4 address two different NLRI with different next-hops
  - Can therefore support both congruent and incongruent topologies

MSDP Concept

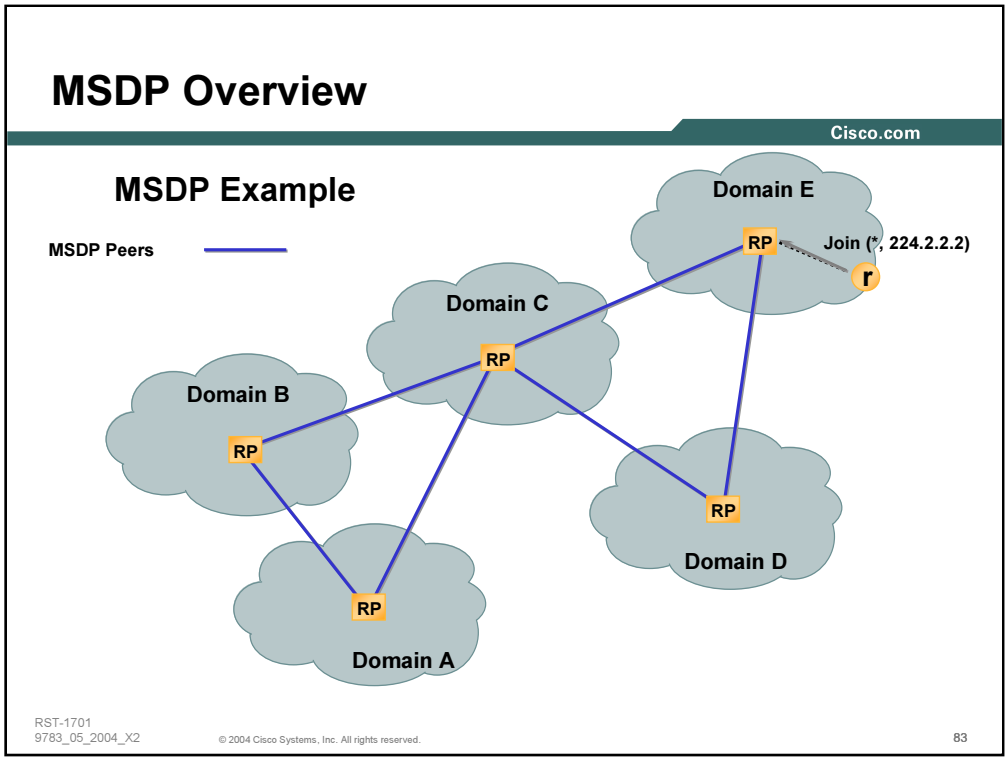
Cisco.com

- **Simple but elegant**
  - **Utilize inter-domain source trees**
  - **Reduces problem to locating active sources**
  - **RP or receiver last-hop can join inter-domain source tree**

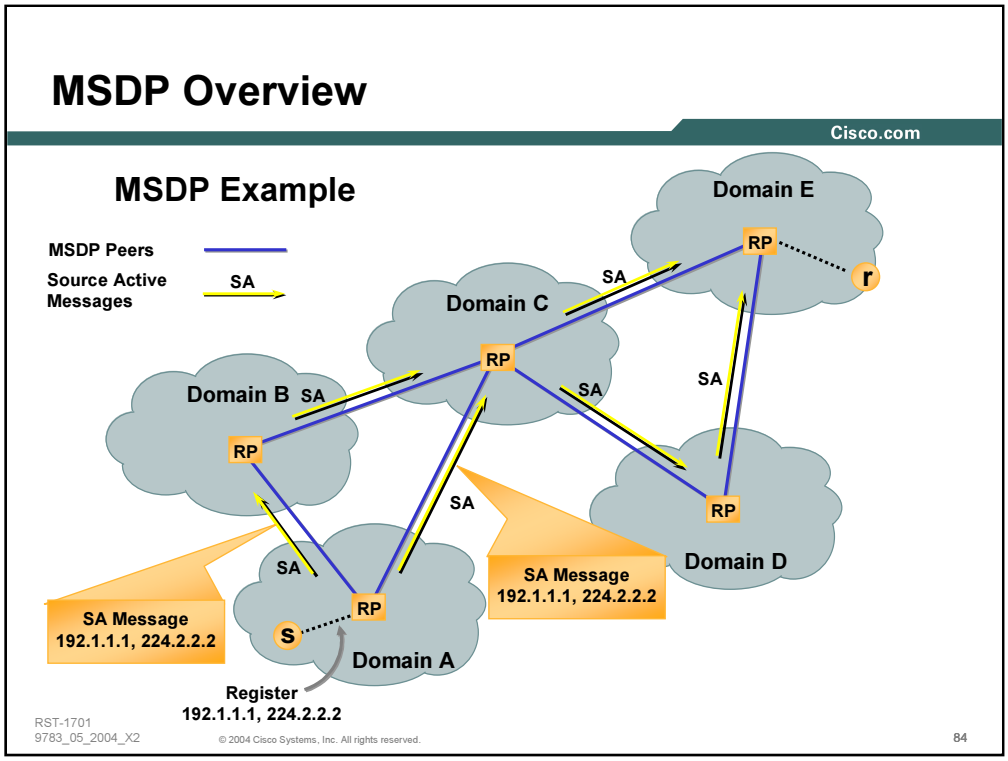
MSDP RFC 3618 Concepts

Cisco.com

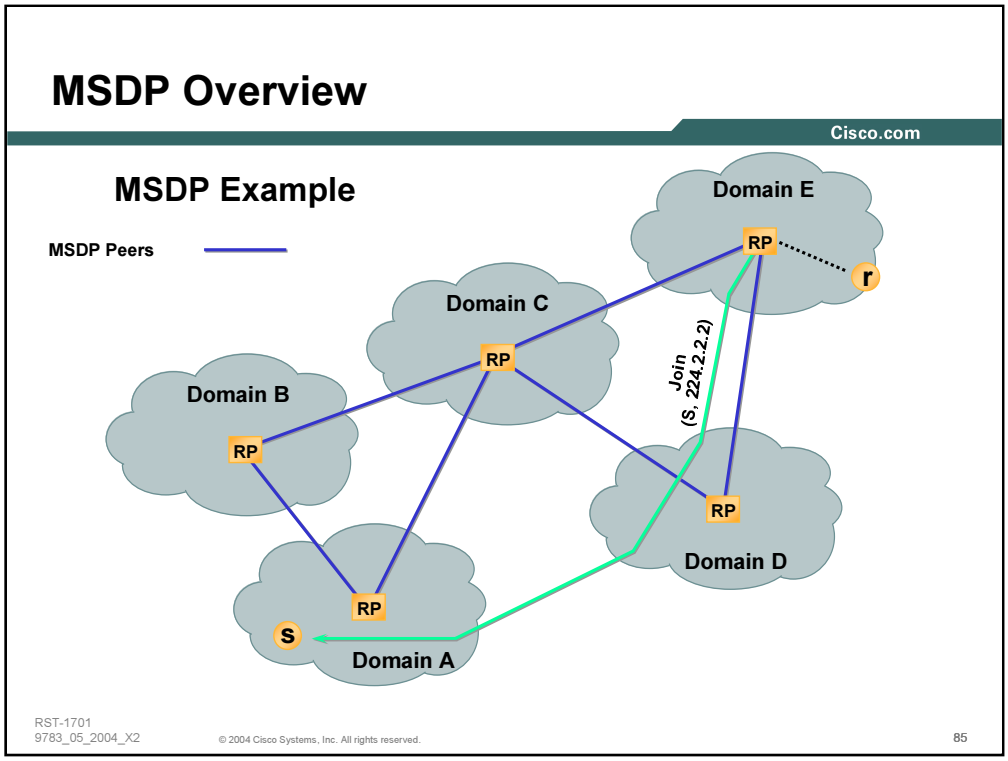
- Works with PIM-SM only
  - RP's knows about all sources in a domain
    - Sources cause a "PIM Register" to the RP
    - Can tell RP's in other domains of its sources
      - Via MSDP SA (Source Active) messages
  - RP's know about receivers in a domain
    - Receivers cause a "(*, G) Join" to the RP
    - RP can join the source tree in the peer domain
      - Via normal PIM (S, G) joins



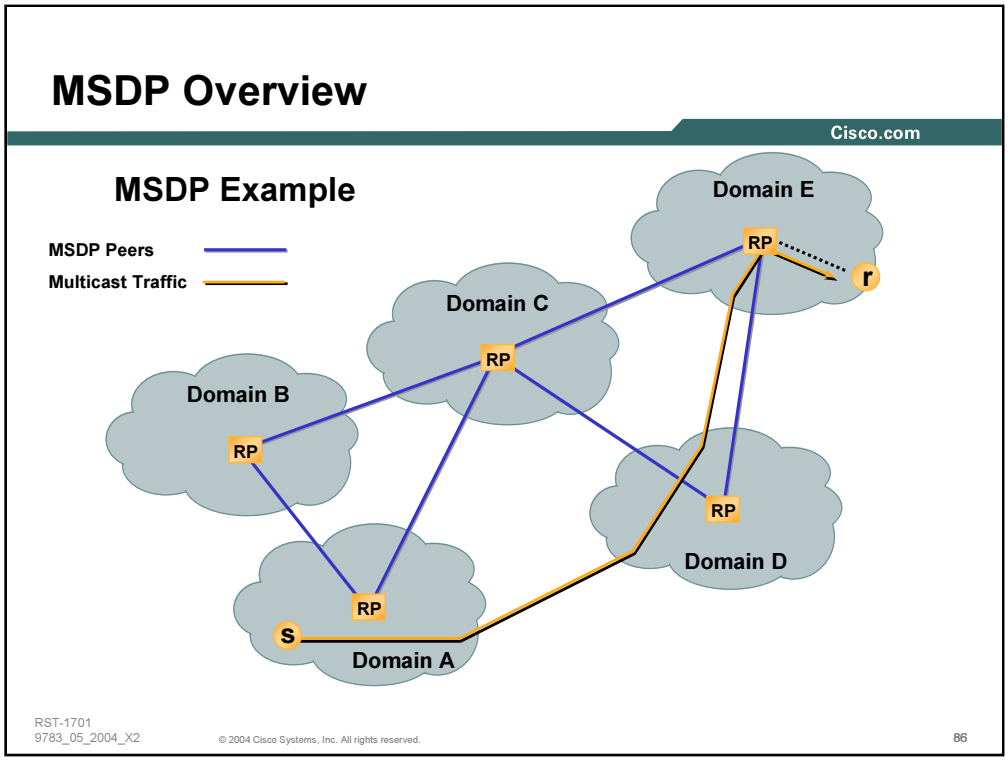
- MSDP Example
  - In the example above, PIM-SM domains A through E each have an RP which is an MSDP speaker. The solid lines between these RP's represents the MSDP peer sessions via TCP and not actual physical connectivity between the domains.
    - *Note: The physical connectivity between the domains is not shown in the drawing above.*
  - Assume that a receiver in Domain E joins multicast group 224.2.2.2 which in turn, causes its DR to send (*, G) Join for this group to the RP.
  - This builds a branch of the Shared-Tree from the RP in Domain E to the DR as shown.
  - When a source goes active in Domain A, the first-hop router (S) sends a PIM Register message to the RP. This informs the RP in Domain A that a source is active in the local domain. The RP responds by originating an (S, G) SA message for this source and send them to its MSDP peers in domains B and C. (The RP will continue to send these SA messages periodically as long as the source remains active.)
  - When the RP's in domains B and C receive the SA messages, they are RPF checked and forwarded downstream to their MSDP peers. These SA



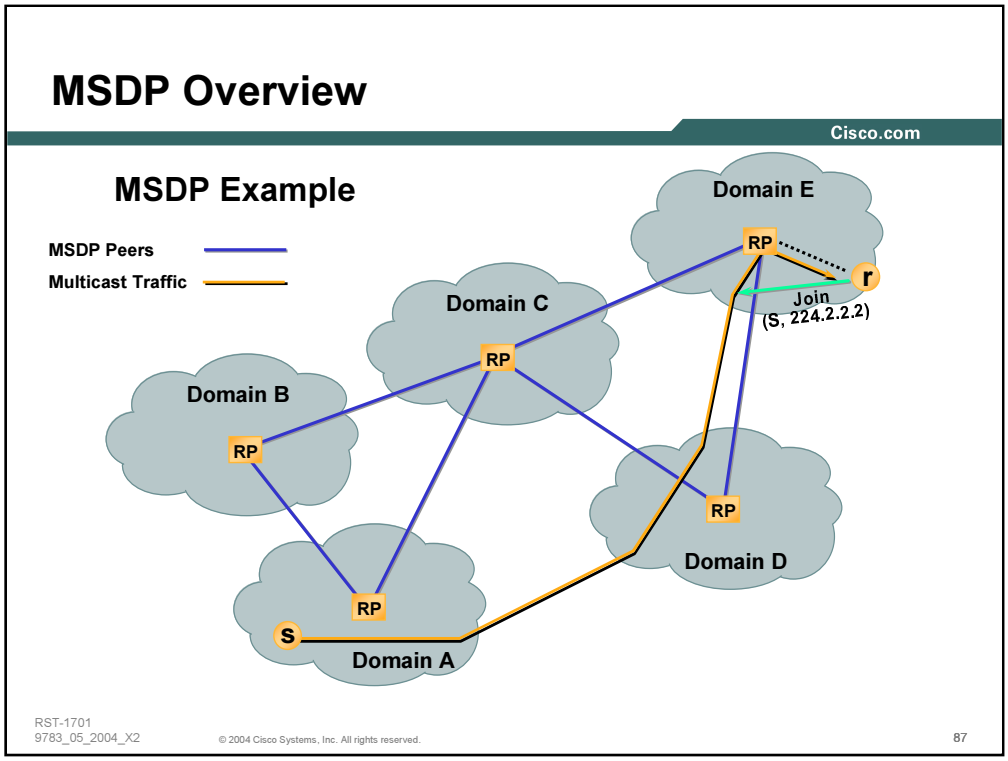
- MSDP Example
  - In the example above, PIM-SM domains A through E each have an RP which is an MSDP speaker. The solid lines between these RP's represents the MSDP peer sessions via TCP and not actual physical connectivity between the domains.
    - *Note: The physical connectivity between the domains is not shown in the drawing above.*
  - Assume that a receiver in Domain E joins multicast group 224.2.2.2 which in turn, causes its DR to send (*, G) Join for this group to the RP.
  - This builds a branch of the Shared-Tree from the RP in Domain E to the DR as shown.
  - When a source goes active in Domain A, the first-hop router (S) sends a PIM Register message to the RP. This informs the RP in Domain A that a source is active in the local domain. The RP responds by originating an (S, G) SA message for this source and send them to its MSDP peers in domains B and C. (The RP will continue to send these SA messages periodically as long as the source remains active.)
  - When the RP's in domains B and C receive the SA messages, they are RPF checked and forwarded downstream to their MSDP peers. These SA



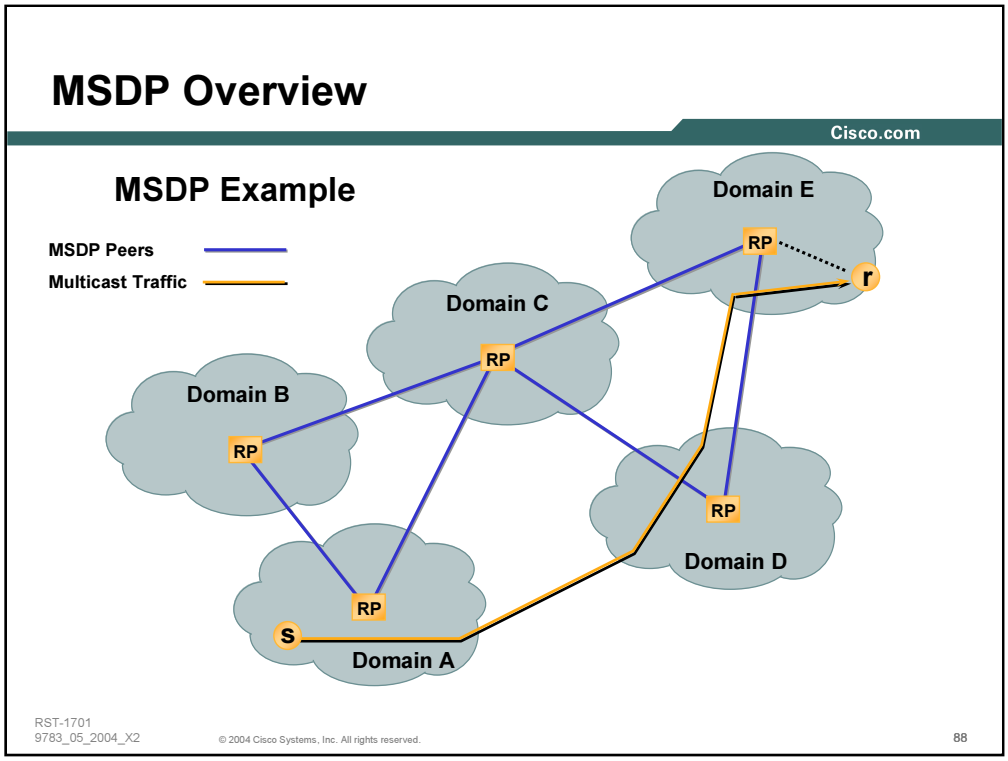
- MSDP Example
  - Once the SA message arrives at the RP (MSDP speaker) in domain E, it sees that it has an active branch of the Shared-Tree for group 224.2.2.2. It responds to the SA message by sending an (S, G) Join toward the source.
    - IMPORTANT: The (S, G) Join will follow the normal inter-domain routing path from the RP to the source. This inter-domain routing path is not necessarily the same path as that used by the MSDP connections. *In order to emphasis this point, the (S, G) Join is shown following a different path between domains.*



- MSDP Example
  - Once the (S, G) Join message reaches the first-hop router (S) in domain A, (S, G) traffic begins to flow to the RP in domain E via the Source Tree shown.
    - IMPORTANT: The (S, G) traffic will not flow over the TCP MSDP sessions. It will instead follow the path of the Source Tree that was built in the preceding step.



- MSDP Example
  - Once the (S, G) traffic reaches the last-hop router (R) in domain E, the last-hop router may optionally send an (S, G) Join toward the source in order to bypass the RP in domain E. This is shown in the above example.



- MSDP Example
  - At this point in the example, the (S, G) traffic is flowing to the last-hop router (R) in domain E via the Source-Tree as shown in the above example.

Anycast RP: Overview

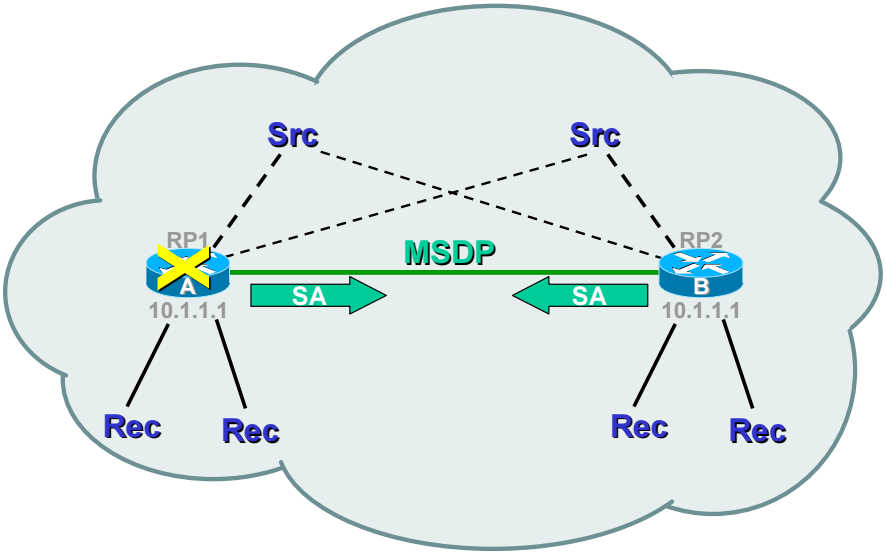
Cisco.com

- Uses single statically defined RP address
  - Two or more routers have same RP address
    - RP address defined as a Loopback Interface.
    - Loopback address advertised as a Host route.
  - Senders & Receivers Join/Register with closest RP
    - Closest RP determined from the unicast routing table.
  - Can *never* fall back to Dense mode.
    - Because RP is statically defined.
- MSDP session(s) run between all RPs
  - Informs RPs of sources in other parts of network
  - RPs join SPT to active sources as necessary

RST-1701
9783_05_2004_X2

Anycast RP: Overview

Cisco.com

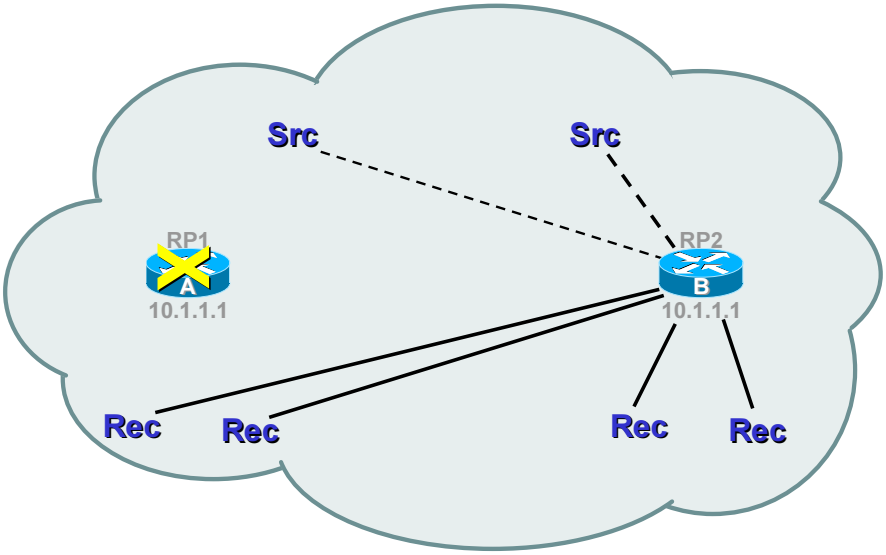


RST-1701
9783_05_2004_X2

90

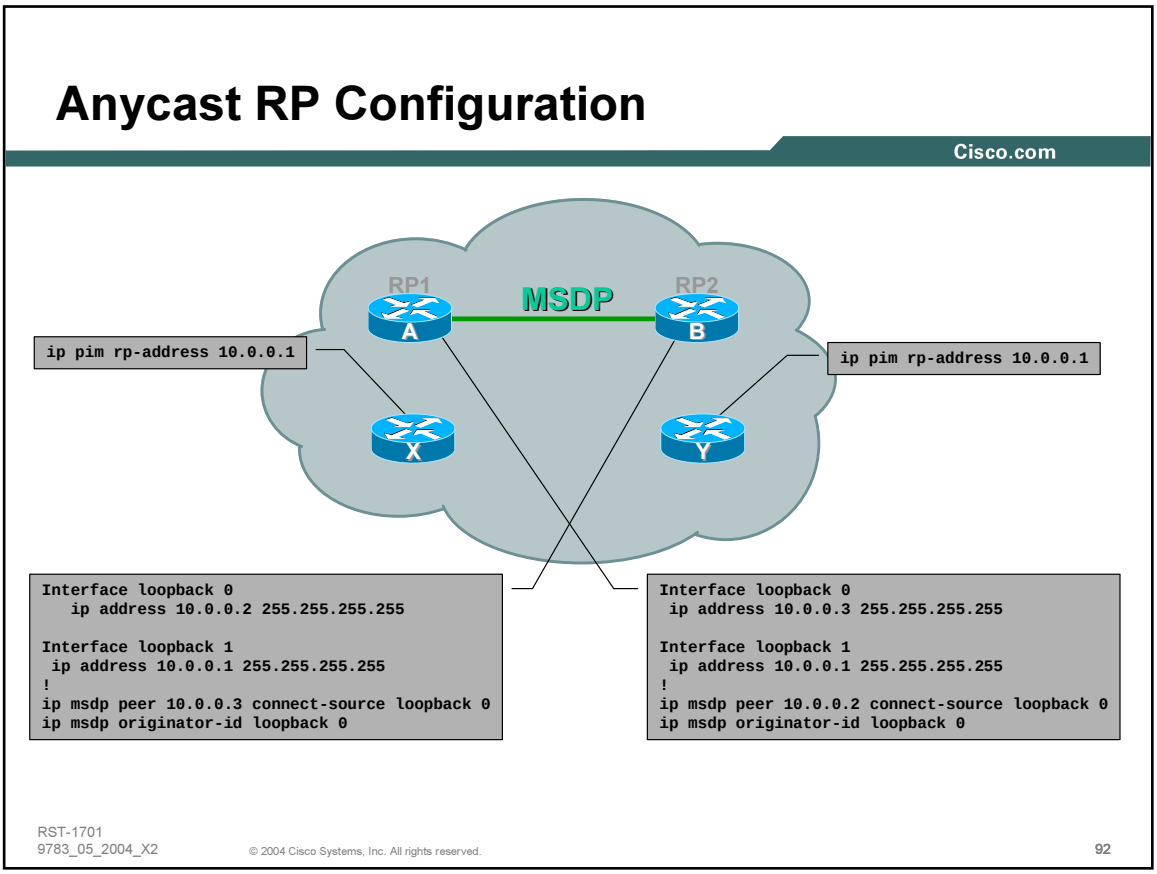
Anycast RP: Overview

Cisco.com



RST-1701
9783_05_2004_X2

91



- Anycast RP Example
  - In this example, two Anycast RP's are configured with the same IP address, 10.1.1.1, using Loopback 0.
  - Each are connected via MSDP using their Loopback 1 addresses, 10.0.0.1 and 10.0.0.2.
    - (Yes, you must use some other address in the 'ip msdp peer' commands than 10.0.0.1.)

You Now Know…

Cisco.com

- Why Multicast?
- Multicast Fundamentals
- PIM Protocols
- RP choices
- Multicast at Layer 2
- Interdomain IP Multicast

RST-1701
9783_05_2004_X2

94