NETWORKERS 2004

INTRODUCTION TO IP MULTICAST

SESSION RST-1701

Networkers Multicast Sessions

Cisco.com

- Breakout Sessions
  - RST 1701 – Introduction to IP Multicast
  - RST 2701 – Deploying IP Multicast
  - RST 2702 – Deploying IP Multicast VPN's
  - RST 4701 – Advanced IP Multicast
- Techtorials
  - RST 2T07 – Enterprise IP Multicast
- Multicast BoF

Agenda

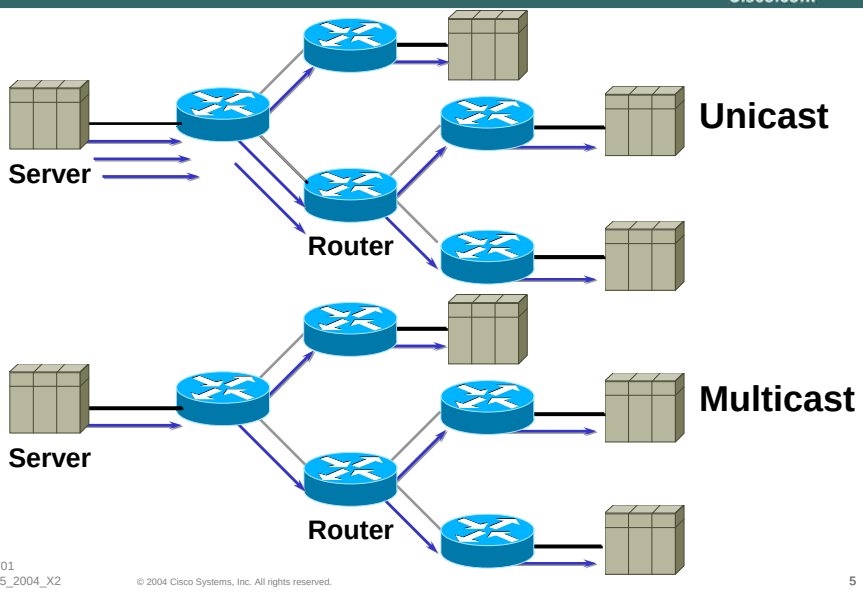
Cisco.com

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

WHY MULTICAST?

Unicast vs. Multicast

Cisco.com



RST-1701
9783_05_2004_X2

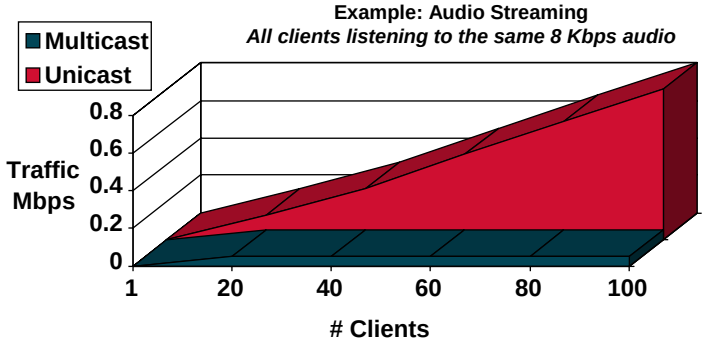
5

Multicast Advantages

Cisco.com

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

Example: Audio Streaming
All clients listening to the same 8 Kbps audio



RST-1701
9783_05_2004_X2

6

Multicast Applications

Cisco.com

Today

- **Finance Applications**
  - –Trading Stocks and Commodities
- **Streaming Multimedia**
  - –E-Learning
  - –Corporate communications
- **Enterprise Resource Applications**
  - –Data warehousing and content synchronization
- **Any one-to-many data push applications**

Tomorrow

- **Broadband access**
- **Video conferencing**
- **Digital TV**
- **Digital audio**
- **Networked gaming**
- **PDAs and Home appliances**

Multicast Disadvantages

Cisco.com

Multicast Is UDP Based!!!

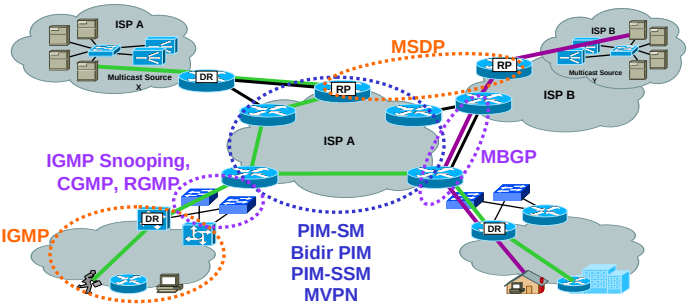
- ***Best Effort Delivery*: Drops are to be expected. Multicast applications should not expect reliable delivery of data and should be designed accordingly. Reliable Multicast is still an area for much research. Expect to see more developments in this area.**
- ***No Congestion Avoidance*: Lack of TCP 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.**
- ***Out of Order Delivery* : Some protocol mechanisms may also result in out of order delivery of packets.**

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

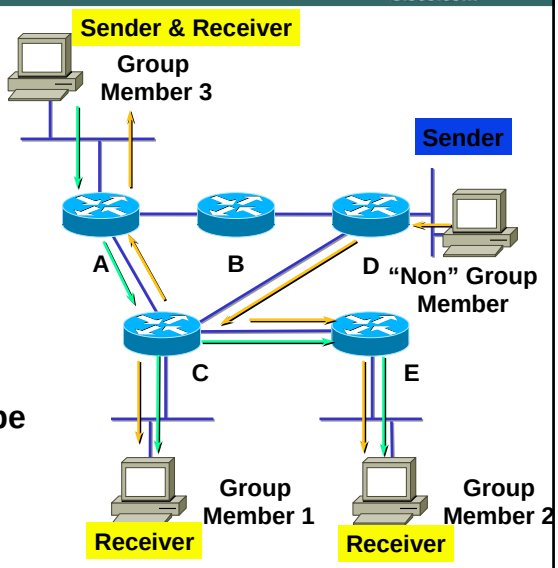
Interdomain Multicast

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

IP Multicast Group Concept

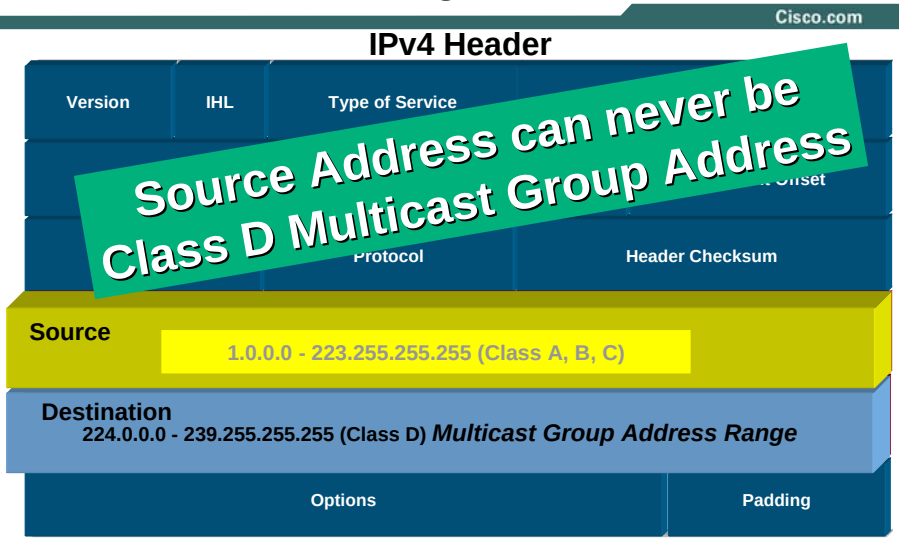
Cisco.com

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



Multicast Addressing

Cisco.com



Multicast Group Address Range

Cisco.com

224.0.0.0 - 239.255.255.255 (Class D)

- 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
    - 224.0.0.5 OSPF routers
    - 224.0.0.13 PIMv2 Routers
    - 224.0.0.22 IGMPv3
- Other Reserved Addresses
  - 224.0.1.0 – 224.0.1.255
  - Not local in scope (Transmitted with TTL > 1)
  - Examples:
    - 224.0.1.1 NTP Network Time Protocol
    - 224.0.1.32 Mtrace routers
    - 224.0.1.78 Tibco Multicast1

Multicast Addressing

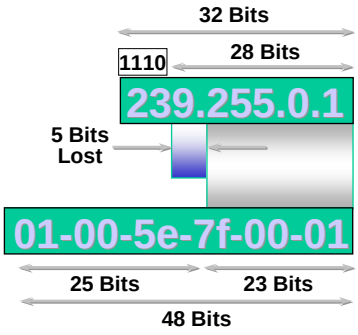
Cisco.com

- 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

Multicast Addressing

Cisco.com

IP Multicast MAC Address Mapping
(FDDI and Ethernet)



Multicast Addressing

Cisco.com

IP Multicast MAC Address Mapping
(FDDI & 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

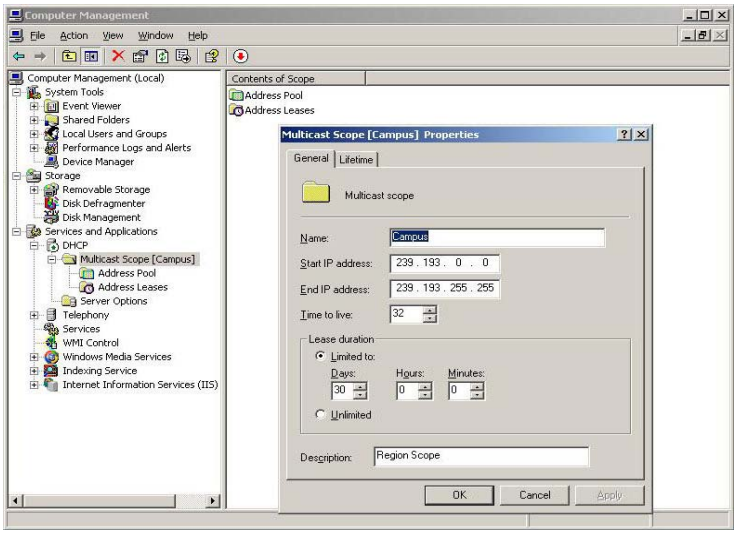
How are Multicast Addresses Assigned?

Cisco.com

- **Dynamic Address Assignment**
  - **Session Directory Tool (SDR)**
    - **Historically used to announce session/group information on a well-known multicast group**
    - **Has problems scaling**
  - **Multicast Address Dynamic Client Allocation Protocol (MADCAP) - RFC 2730**
    - **Similar to DHCP**
    - **Server and Client API shipped in W2K**
    - **Applications need to support to support MADCAP**
  - **Multicast Address Set-Claim (MASC) - RFC 2909**
    - **Hierarchical, dynamic address allocation scheme**
    - **Top of hierarchy is at an Internet exchange**
    - **Children request addresses from parent**
    - **Complex garbage-collection problem**
    - **Not in use**

Madcap in MS Server

Cisco.com



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**
  - **Defined in RFC 2770**
    - **"GLOP Addressing in 233/8"**
- **Manual Address Allocation by the Admin !!**
  - **Is still the most common practice**

Host-Router Signaling: IGMP

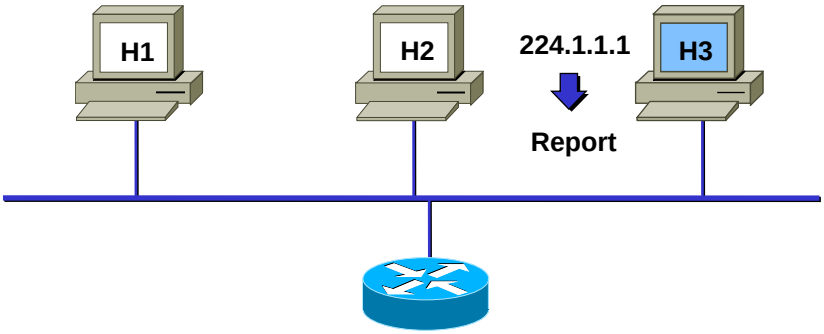
Cisco.com

- **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**
  - **Supported in Window XP and various UNIX systems**

Host-Router Signaling: IGMP

Cisco.com

Joining a Group

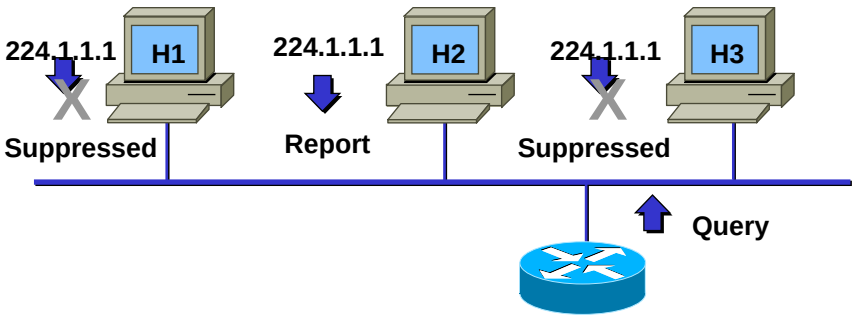


- Host sends IGMP Report to join group

Host-Router Signaling: IGMP

Cisco.com

Maintaining a Group

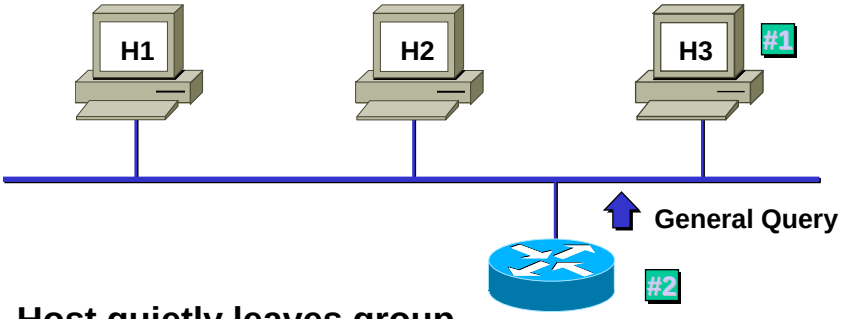


- Router sends periodic Queries to 224.0.0.1
- One member per group per subnet reports
- Other members suppress reports

Host-Router Signaling: IGMP

Cisco.com

Leaving a Group (IGMPv1)

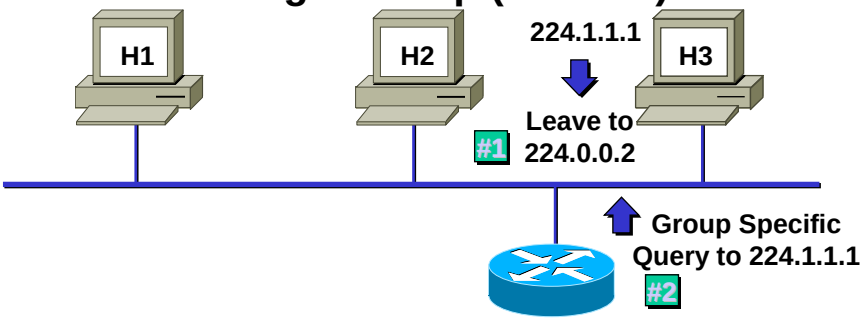


- Host quietly leaves group
- Router sends 3 General Queries (60 secs apart)
- No IGMP Report for the group is received
- Group times out (Worst case delay ~= 3 minutes)

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

- 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

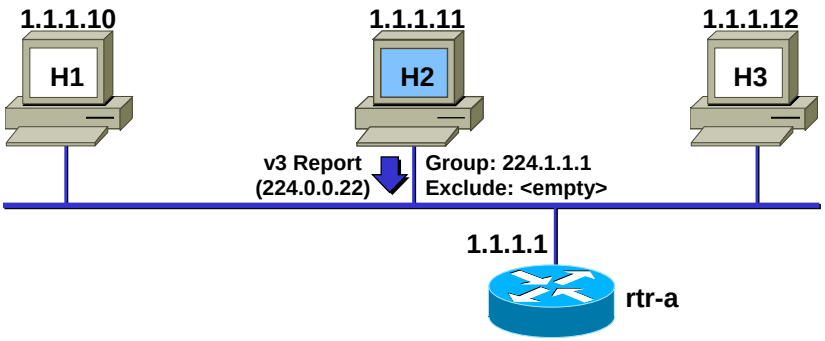
Host-Router Signaling: IGMPv3

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 a Group

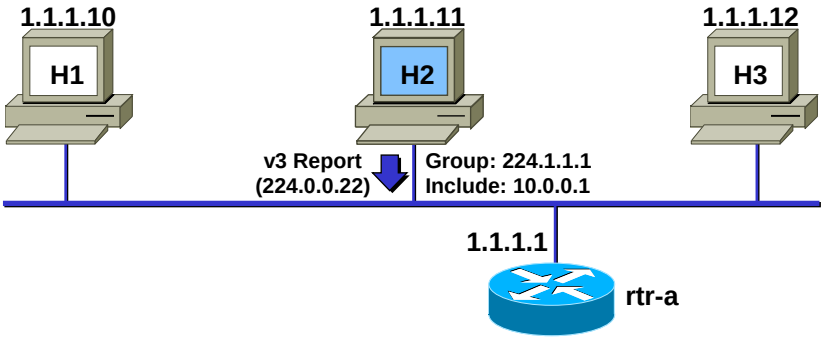
Cisco.com



- Joining member sends IGMPv3 Report to 224.0.0.22 immediately upon joining

IGMPv3—Joining Specific Source(s)

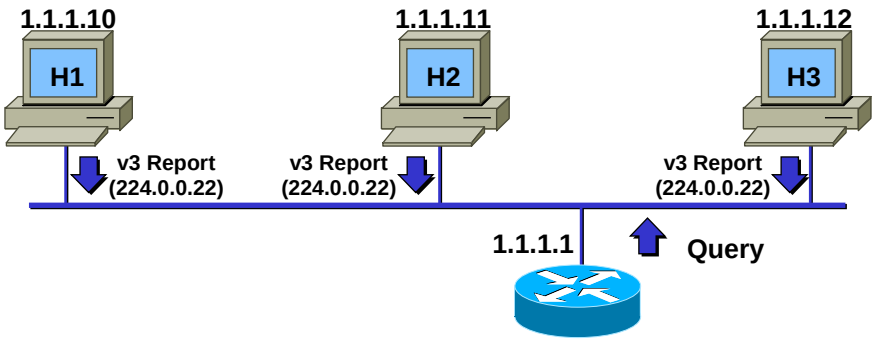
Cisco.com



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

IGMPv3—Maintaining State

Cisco.com

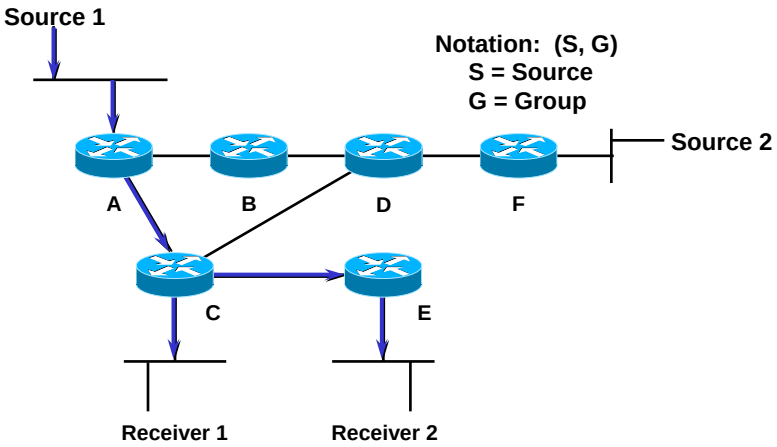


- Router sends periodic queries
- All IGMPv3 members respond
  - Reports contain multiple Group state records

Multicast Distribution Trees

Cisco.com

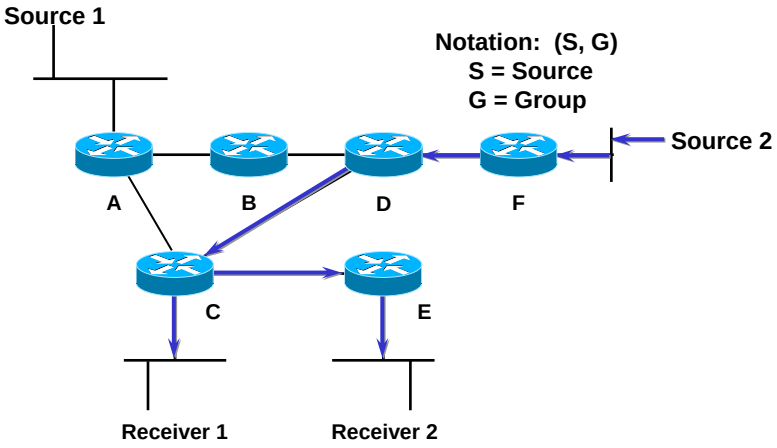
Shortest Path or Source Distribution Tree



Multicast Distribution Trees

Cisco.com

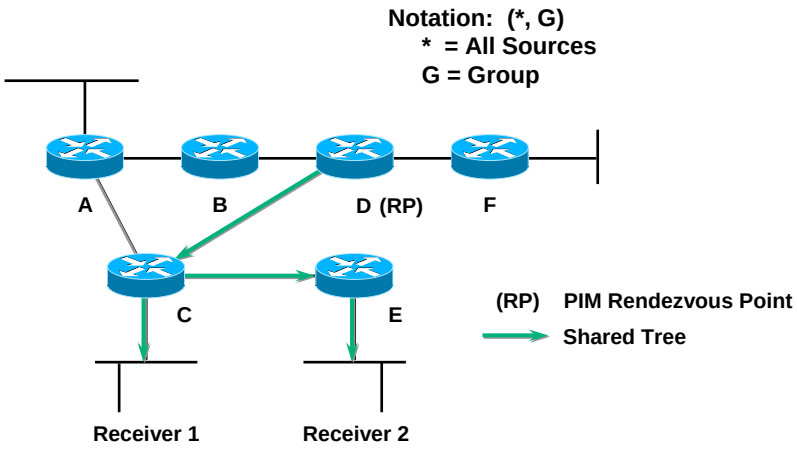
Shortest Path or Source Distribution Tree



Multicast Distribution Trees

Cisco.com

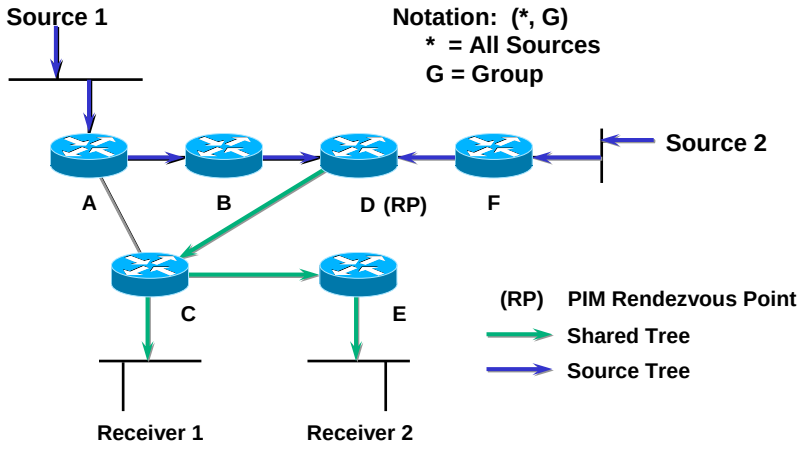
Shared Distribution Tree



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?**

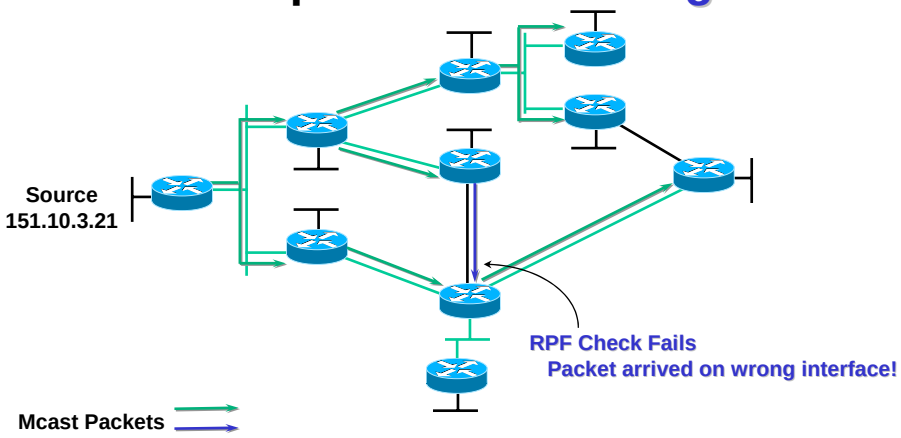
  A router forwards a multicast datagram only if received on the up stream interface to the source (i.e. it follows the distribution tree).

- **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

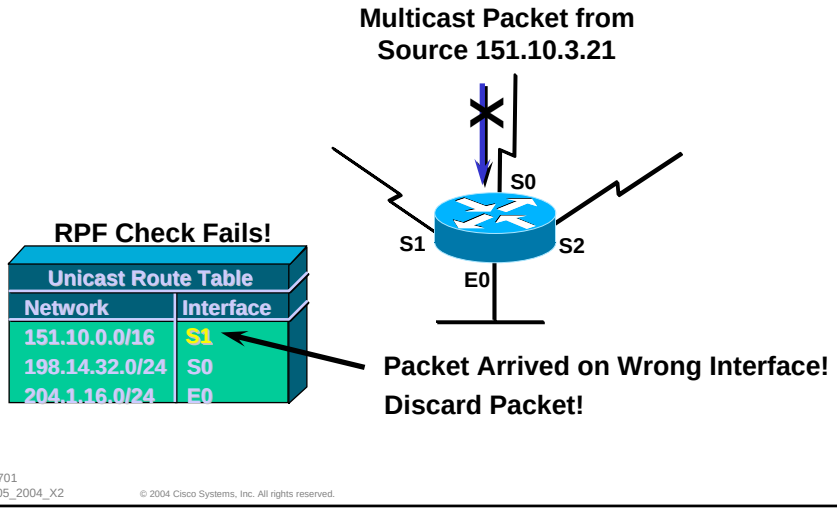
Example: RPF Checking



Multicast Forwarding

Cisco.com

A closer look: RPF Check Fails

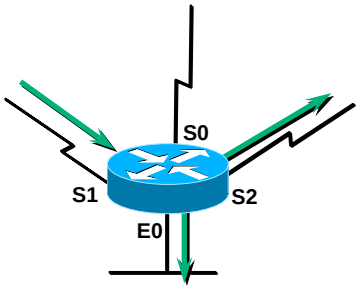


Multicast Forwarding

Cisco.com

A closer look: RPF Check Succeeds

Multicast Packet from Source 151.10.3.21



RPF Check Succeeds!

Unicast Route Table	
Network	Interface
151.10.0.0/16	S1
198.14.32.0/24	S0
204.1.16.0/24	E0

Packet Arrived on Correct Interface!
Forward out all outgoing interfaces.
(i. e. down the distribution tree)

"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

Types of Multicast Protocols

Cisco.com

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

PIM-DM

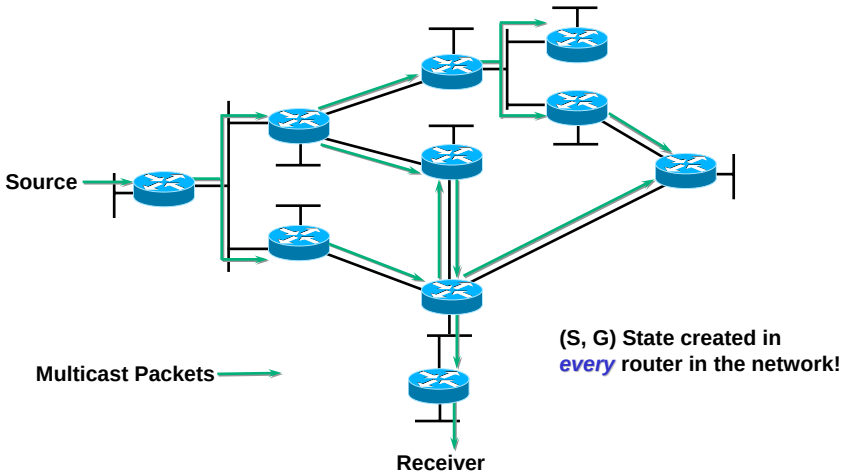
Cisco.com

- 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

PIM-DM Flood and Prune

Cisco.com

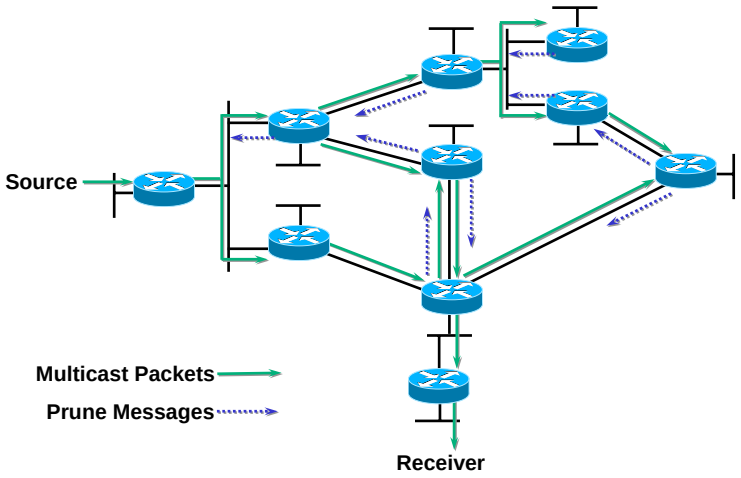
Initial Flooding



PIM-DM Flood and Prune

Cisco.com

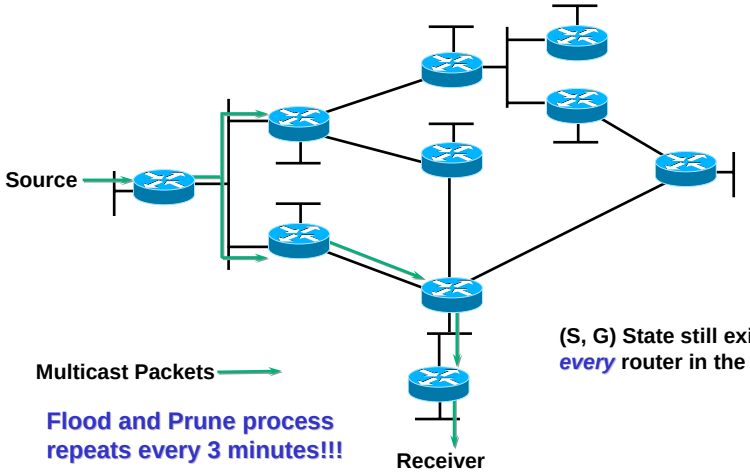
Pruning Unwanted Traffic



PIM-DM Flood and Prune

Cisco.com

Results After Pruning



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

Flood and Prune process repeats every 3 minutes!!!

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

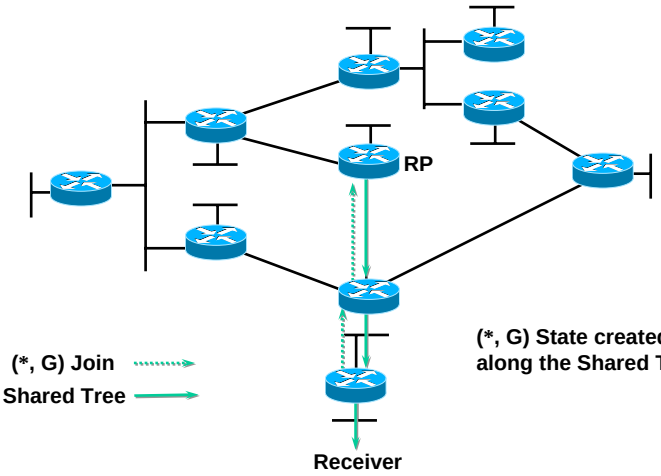
PIM-SM (RFC 2362)

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.

PIM-SM Shared Tree Join

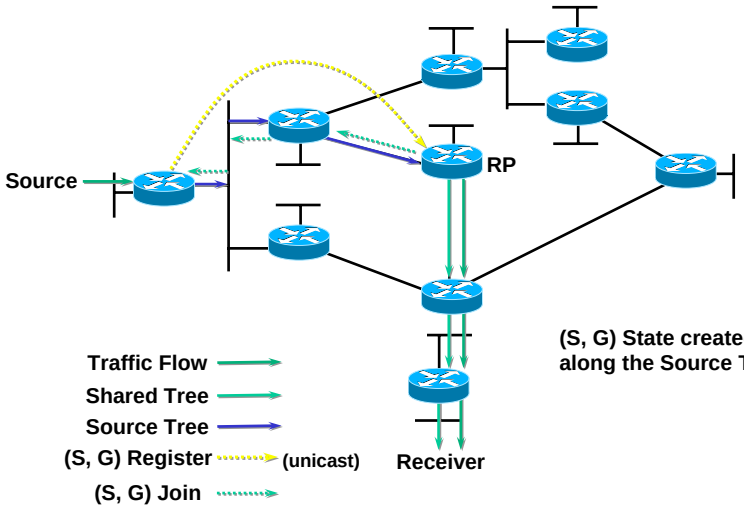
Cisco.com



RST-1701
9783_05_2004_X2

PIM-SM Sender Registration

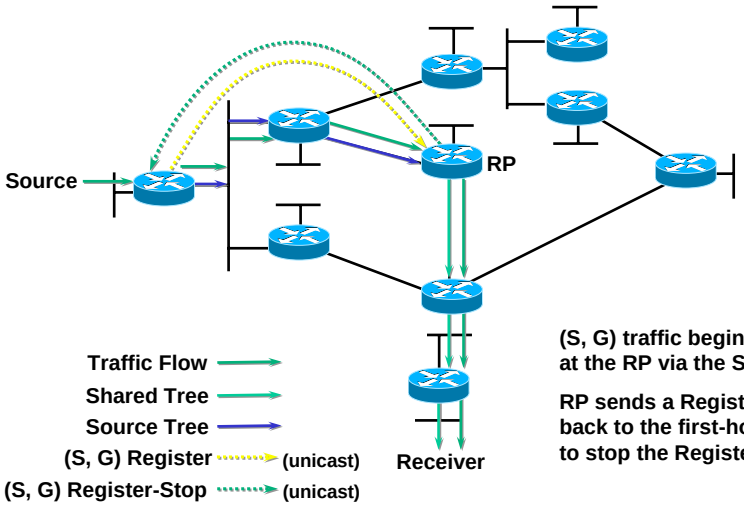
Cisco.com



RST-1701
9783_05_2004_X2

PIM-SM Sender Registration

Cisco.com

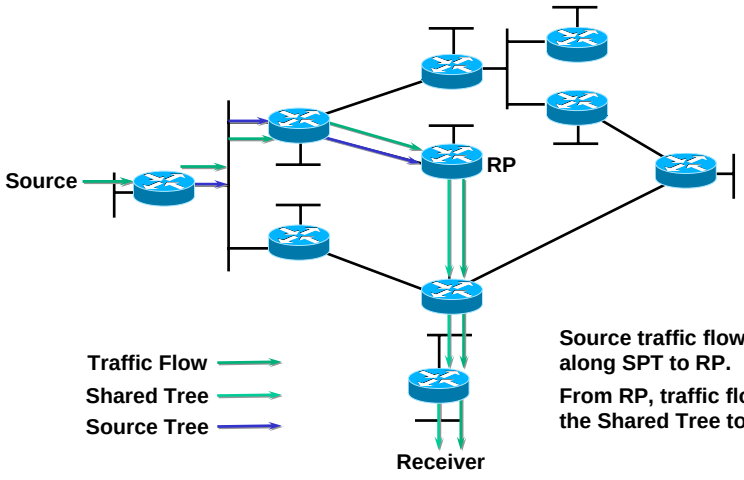


(S, G) traffic begins arriving at the RP via the Source tree.

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

PIM-SM Sender Registration

Cisco.com

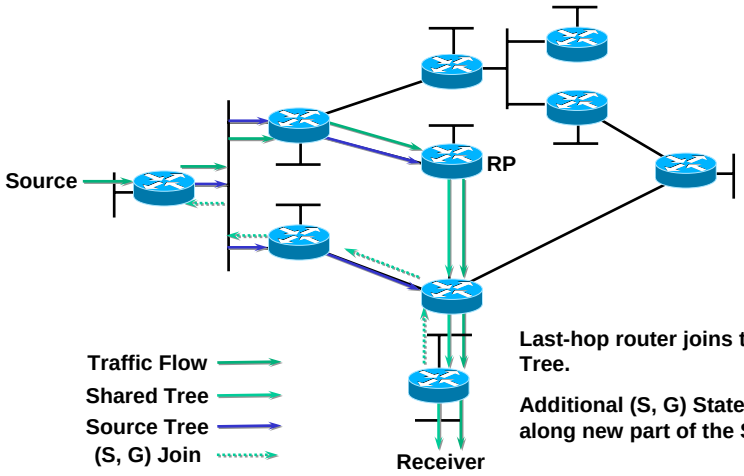


Source traffic flows natively along SPT to RP.

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

PIM-SM SPT Switchover

Cisco.com

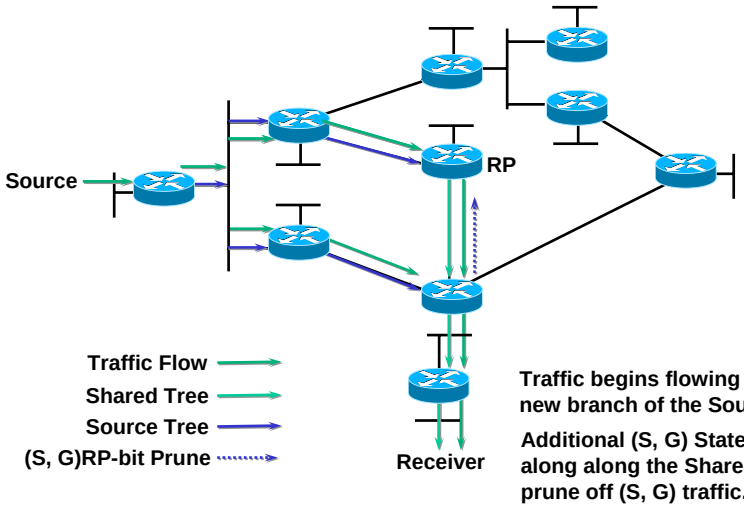


Last-hop router joins the Source Tree.

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

PIM-SM SPT Switchover

Cisco.com

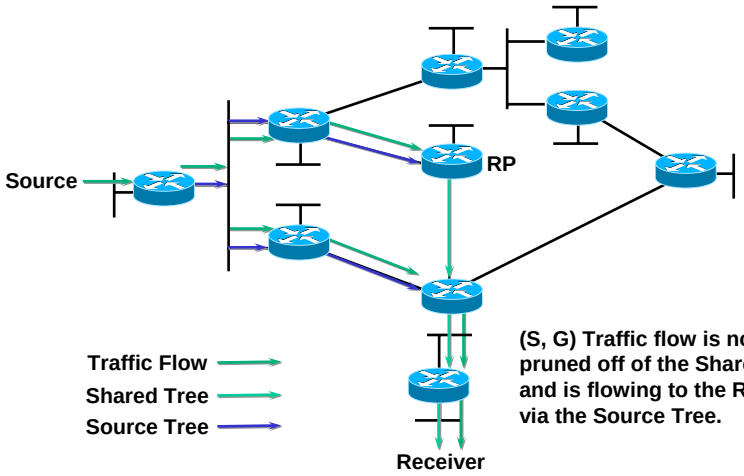


Traffic begins flowing down the new branch of the Source Tree.

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

PIM-SM SPT Switchover

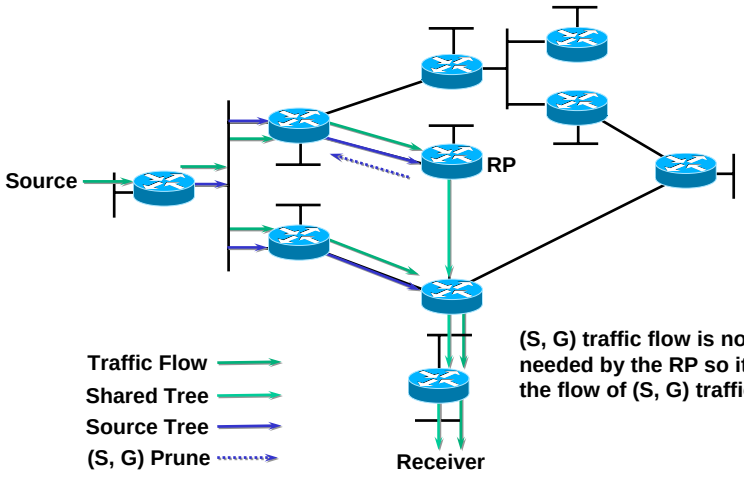
Cisco.com



(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

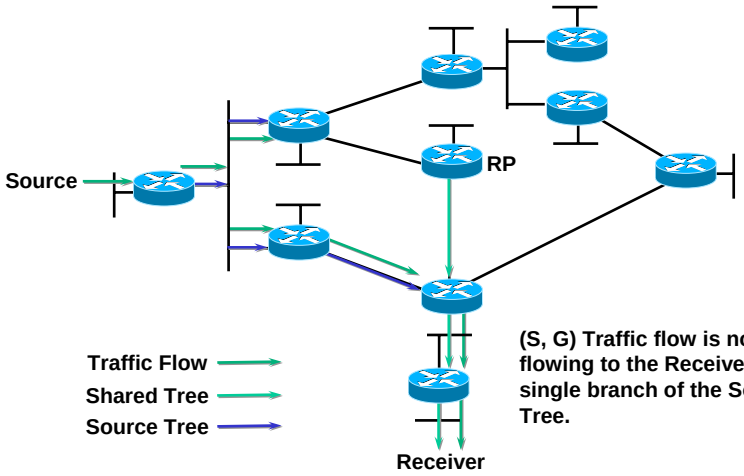
Cisco.com



(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.

"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

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)

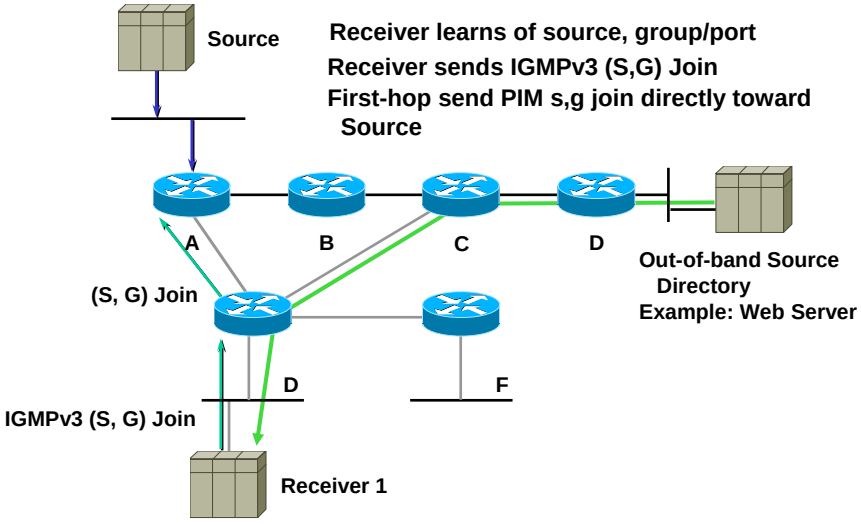
Source Specific Multicast Advantages

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.**

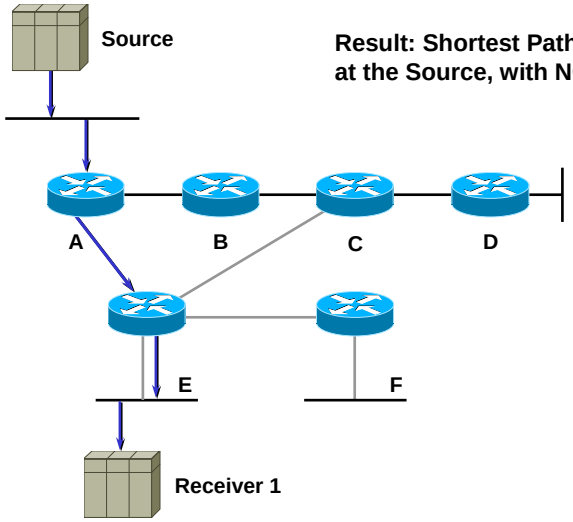
PIM Source Specific Mode

Cisco.com



PIM Source Specific Mode

Cisco.com



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.

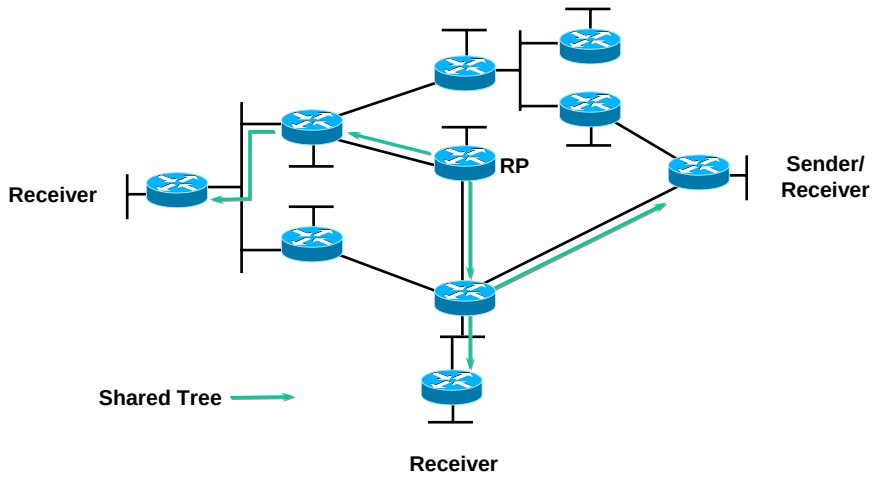
Many-to-Any State Problem

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

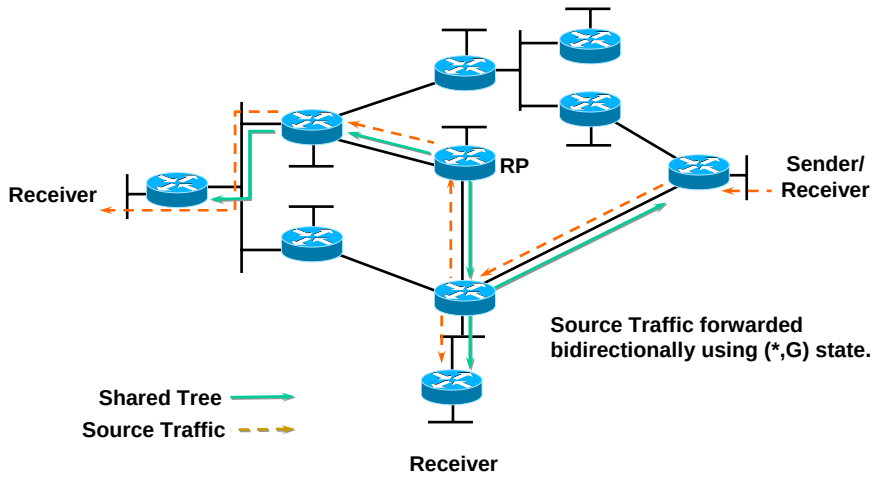
Bidirectional PIM: Overview

Cisco.com



Bidirectional PIM: Overview

Cisco.com



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.

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**

RP CHOICES

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 (Includes Auto-RP Groups!!!!)
  - Override keyword "overrides" Auto-RP information
    - Default: Auto-RP learned info takes precedence

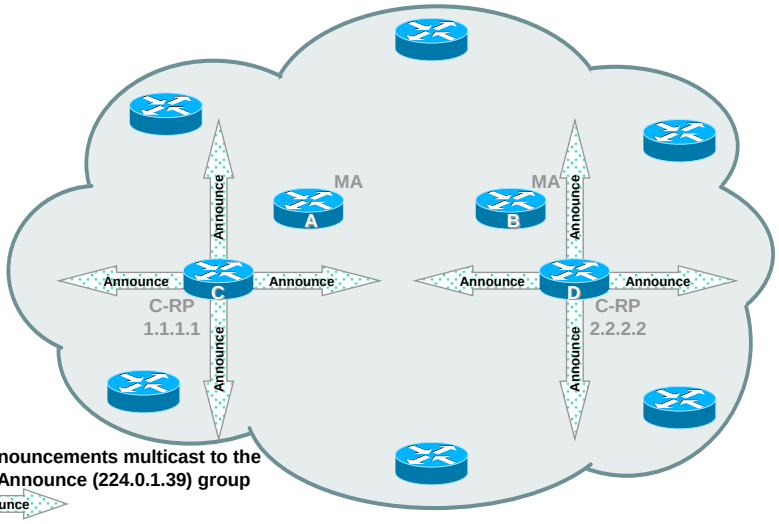
Auto-RP Overview

Cisco.com

- All routers automatically learn RP address
  - No configuration necessary except on:
    - Candidate RPs
    - 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

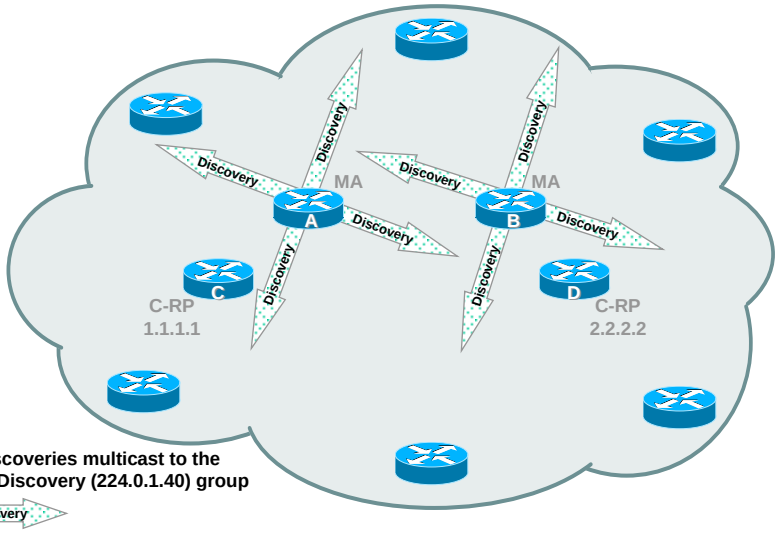
Auto-RP: From 10,000 Feet

Cisco.com



Auto-RP: From 10,000 Feet

Cisco.com



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

Discovery

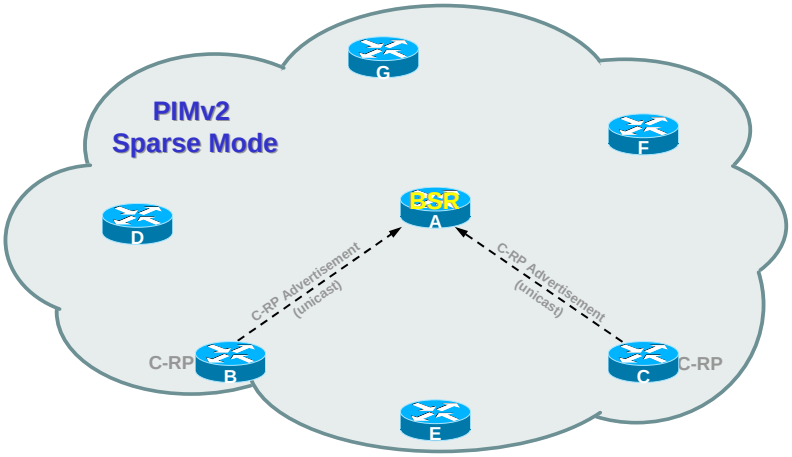
BSR Overview

Cisco.com

- A single Bootstrap Router (BSR) is elected
  - Multiple Candidate BSR's (C-BSR) can be configured
    - Provides backup in case currently elected BSR fails
  - C-RP's send C-RP announcements to the BSR
    - 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

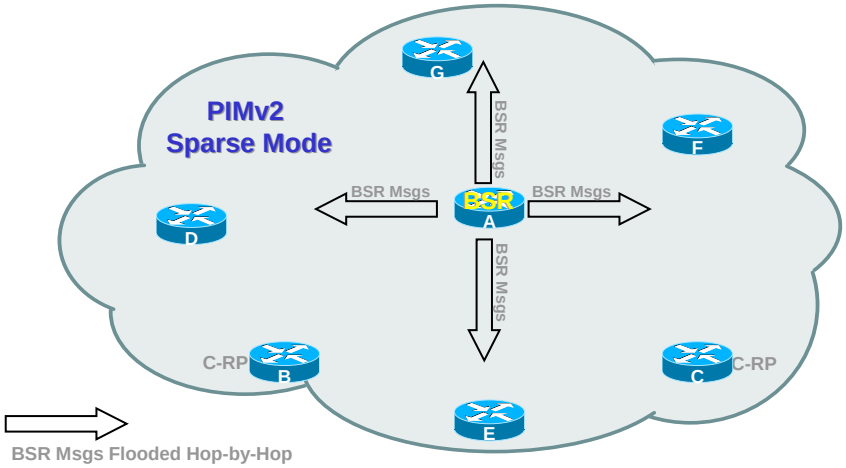


RST-1701
9783_05_2004_X2

77

BSR: From 10,000 Feet

Cisco.com



RST-1701
9783_05_2004_X2

78

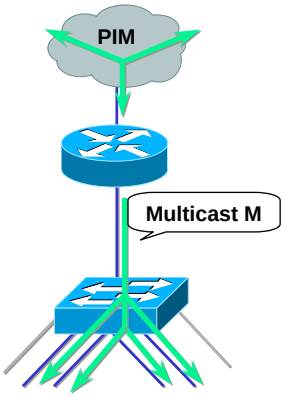
MULTICAST AT LAYER 2

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

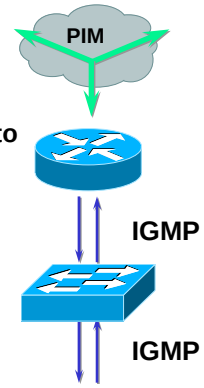


L2 Multicast Frame Switching

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* !!!

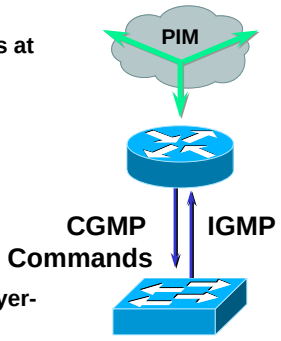


L2 Multicast Frame Switching

Cisco.com

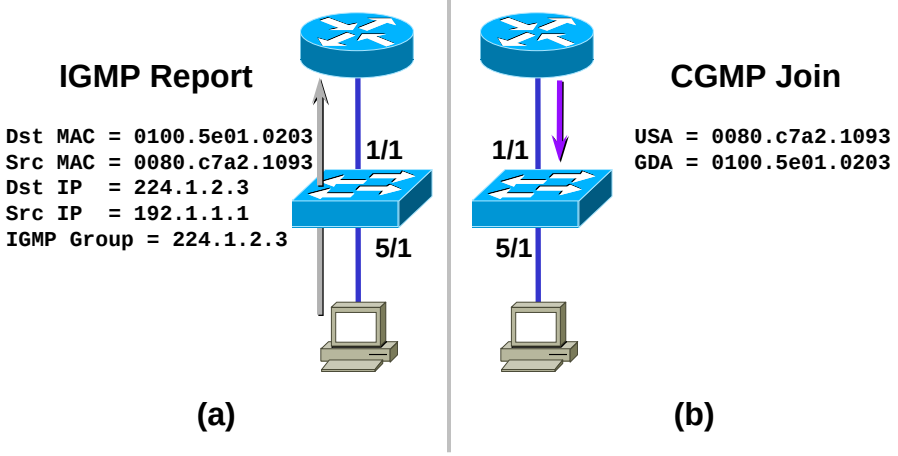
Solution 2: CGMP—Cisco Group Management Protocol

- 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



CGMP Basics

Cisco.com



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

- New Multiprotocol Attributes
  - MP_REACH_NLRI
  - MP_UNREACH_NLRI
- MP_REACH_NLRI and MP_UNREACH_NLRI
  - Address Family Information (AFI) = 1 (IPv4)
    - Sub-AFI = 1 (NLRI is used for unicast)
    - Sub-AFI = 2 (NLRI is used for multicast RPF check)

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

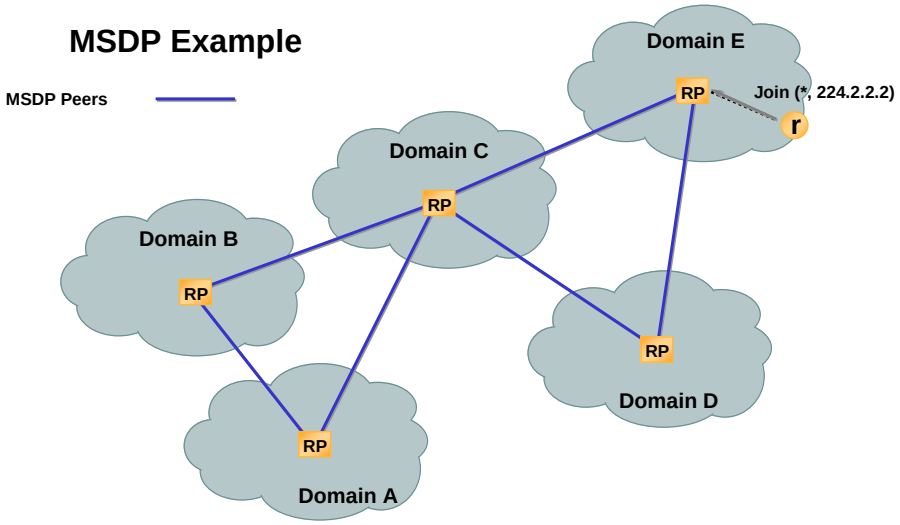
Cisco.com

- 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 Overview

Cisco.com

MSDP Example

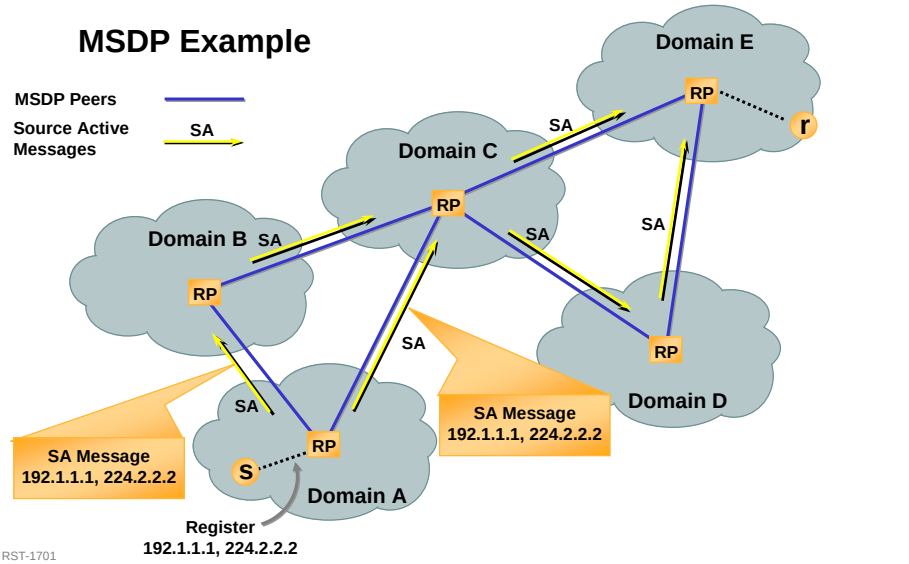


RST-1701
9783_05_2004_X2

MSDP Overview

Cisco.com

MSDP Example

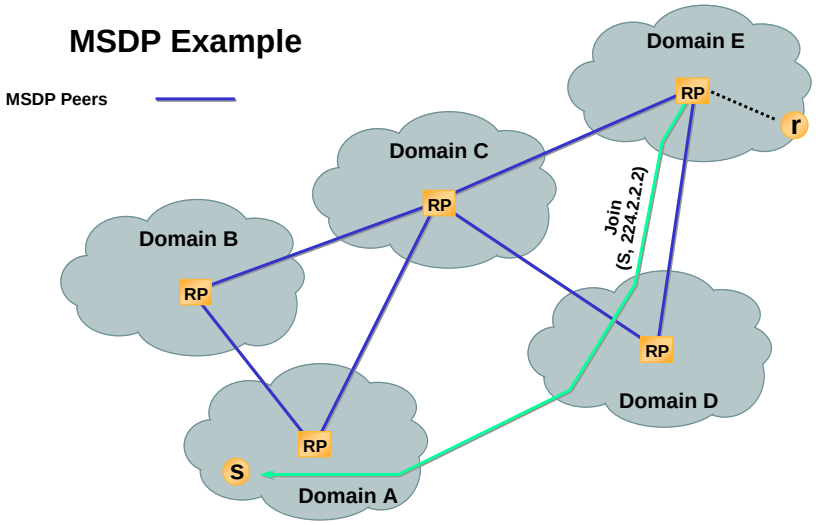


RST-1701
9783_05_2004_X2

MSDP Overview

Cisco.com

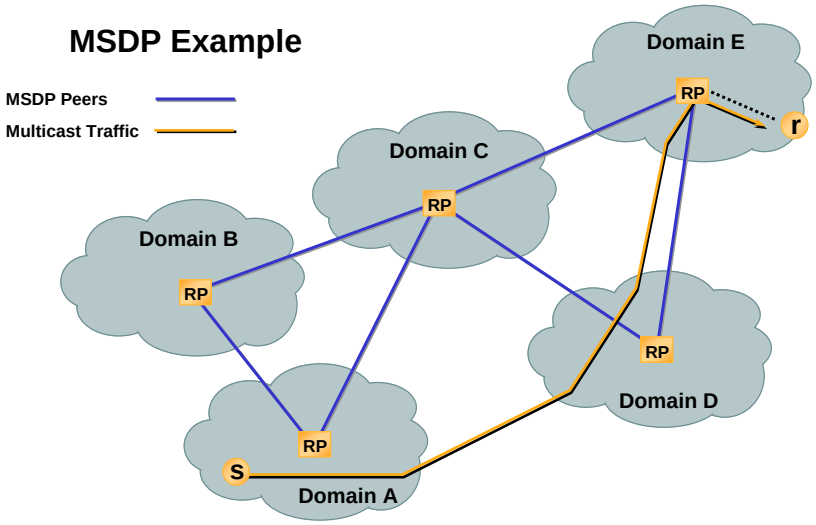
MSDP Example



MSDP Overview

Cisco.com

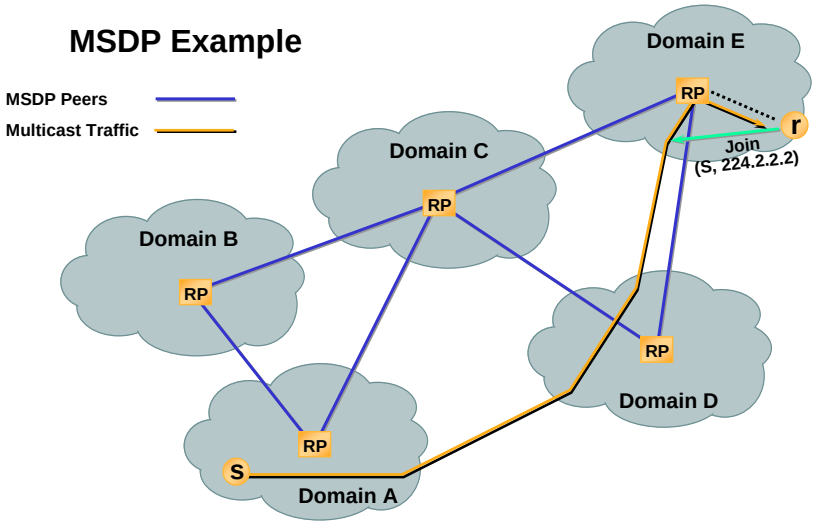
MSDP Example



MSDP Overview

Cisco.com

MSDP Example

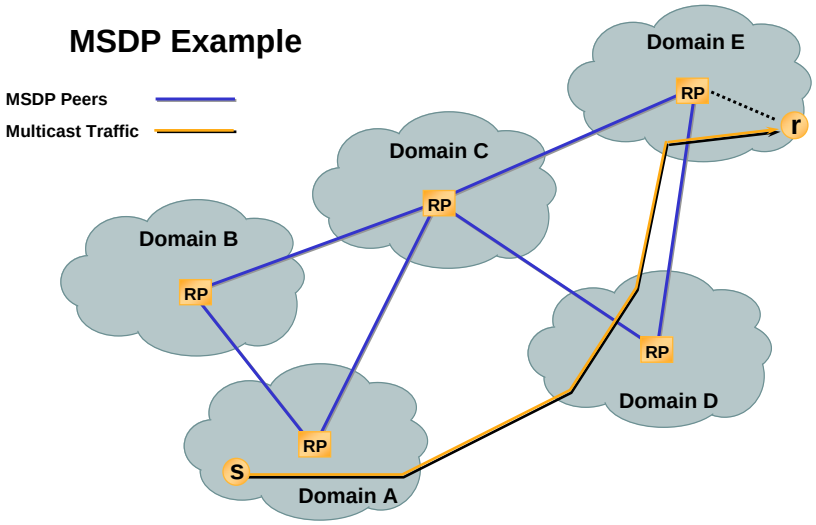


RST-1701
9783_05_2004_X2

MSDP Overview

Cisco.com

MSDP Example



RST-1701
9783_05_2004_X2

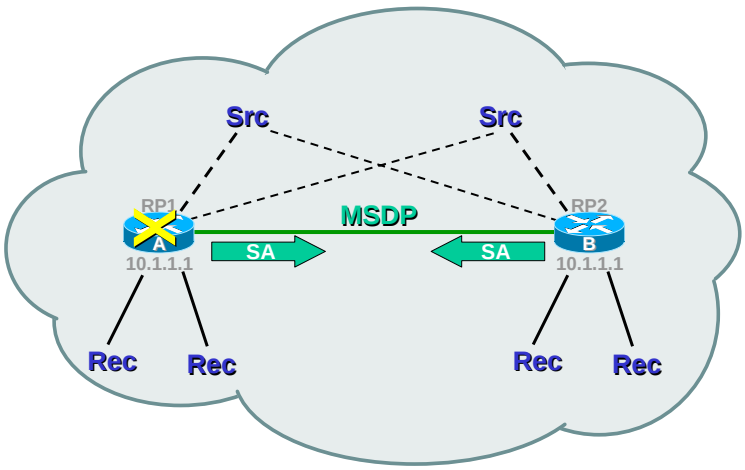
Anycast RP: Overview

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

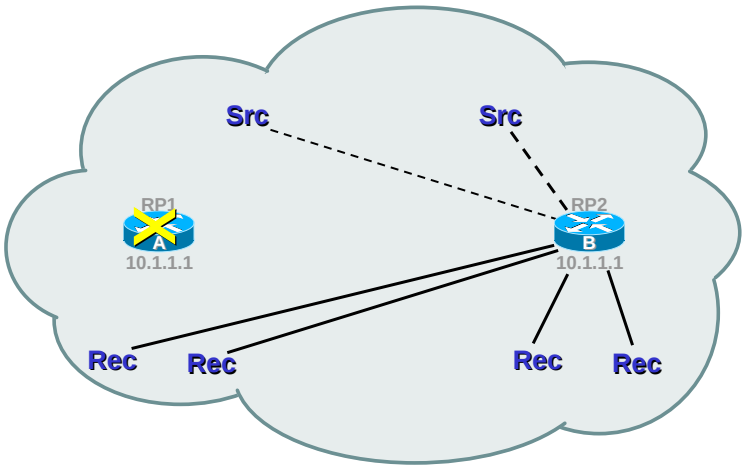
Anycast RP: Overview

Cisco.com



Anycast RP: Overview

Cisco.com

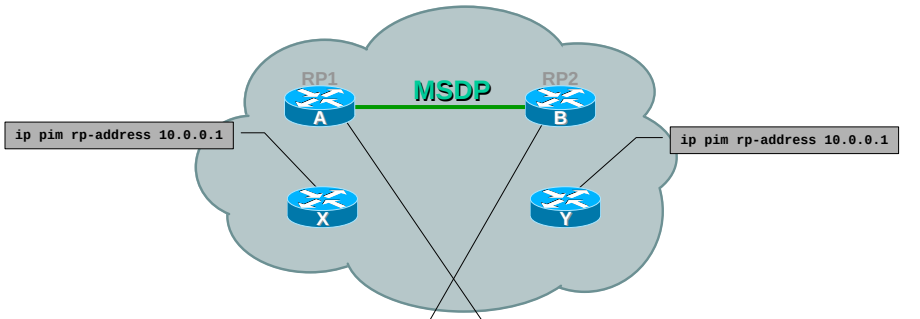


RST-1701
9783_05_2004_X2

101

Anycast RP Configuration

Cisco.com



`ip pim rp-address 10.0.0.1`

`ip pim rp-address 10.0.0.1`

```
Interface loopback 0
   ip address 10.0.0.2 255.255.255.255

Interface loopback 1
 ip address 10.0.0.1 255.255.255.255
!
ip msdp peer 10.0.0.3 connect-source loopback 0
ip msdp originator-id loopback 0
```

```
Interface loopback 0
 ip address 10.0.0.3 255.255.255.255

Interface loopback 1
 ip address 10.0.0.1 255.255.255.255
!
ip msdp peer 10.0.0.2 connect-source loopback 0
ip msdp originator-id loopback 0
```

RST-1701
9783_05_2004_X2

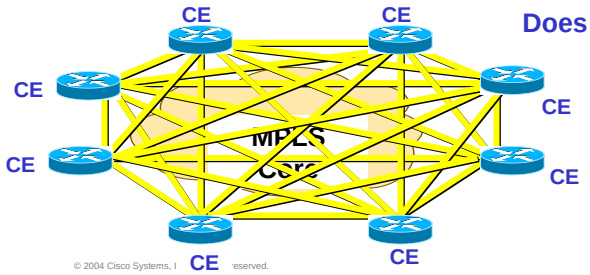
102

LATEST ADDITIONS

Multicast VPN
The Customer Requirement

Cisco.com

- MPLS VPN customers want to run multicast within their VPNs
- Multicast deployment is expanding
- MPLS VPNs do not support multicast today
- Multicast options in MPLS VPNs today
  - GRE tunnels from CE to CE



Does NOT Scale !!!!

Multicast VPN (MVPN)

Cisco.com

- Allows an ISP to provide its MPLS VPN customers the ability to transport their Multicast traffic across MPLS packet-based core networks
- Provides a "Ships in the Night" approach with MPLS
- A scalable architecture solution for MPLS networks based on native multicast deployment in the core
- Uses draft-rosen-vpn-mcast encapsulation and signaling to build MVPN Multicast VPN (MVPN)

Multicast VPN Terminology Used

Cisco.com

- VPN: Virtual Private Network
  - Although different VPN models exist, the discussion here is for MPLS based VPNs
- MVPN: Multicast VPN
  - A VPN that supports multicast natively
- VRF: VPN Routing and Forwarding
  - per-site forwarding tables
- MVRF: Multicast VRF
  - A VRF that supports unicast and multicast tables

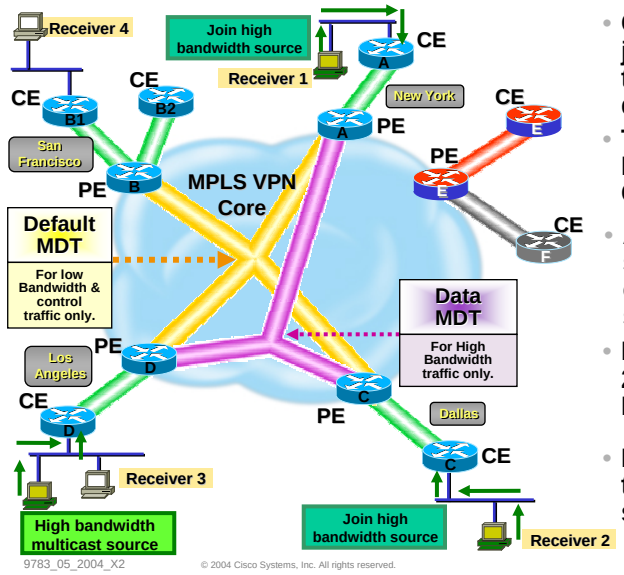
Multicast VPN Terminology Used

Cisco.com

- **MDT: Multicast Distribution Tree**
  - A multicast tree, built in the core network between PE and P routers that distributes multicast traffic between sites
- **Default-MDT:**
  - Default MDT group used for control traffic and flooding channel for dense mode and low bandwidth groups.
- **Data-MDT:**
  - MDT group created on demand for MVPN (S,G) pairs, usually high bandwidth traffic

Multicast VPN (MVPN) Concept and Fundamentals

Cisco.com



- Customer CE devices joins the MPLS Core through provider's PE devices
- The MPLS Core forms a Default MDT for a given Customer
- A High-bandwidth source for that customer starts sending traffic
- Interested receivers 1 & 2 join that High Bandwidth source
- Data-MDT is formed for this High-Bandwidth source

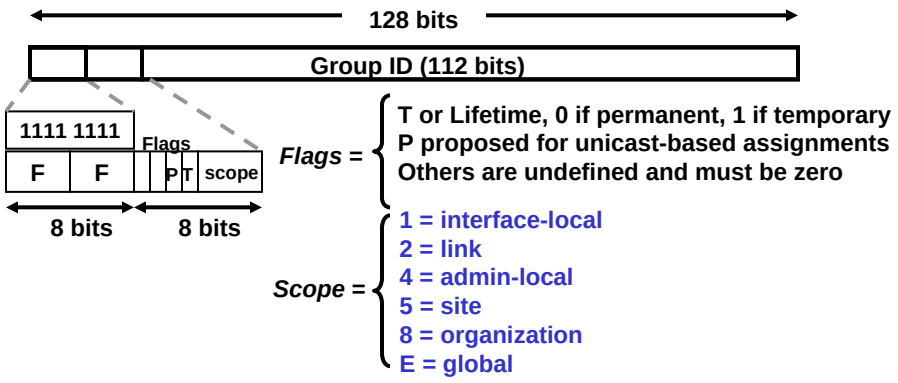
IPv4 versus IPv6 Multicast

Cisco.com

IP Service	IPv4 Solution	IPv6 Solution
Address Range	32-bit, class D	128-bit (112-bit Group)
Routing	Protocol Independent All IGPs,and BGP4+	Protocol Independent All IGPs,and BGP4+ with v6 mcast SAFI
Forwarding	PIM-DM, PIM-SM, PIM-SSM, PIM-bidir	PIM-SM, PIM-SSM, PIM-bidir
Group Management	IGMPv1, v2, v3	MLDv1, v2
Domain Control	Boundary/Border	Scope Identifier
Inter-domain Solutions	MSDP across Independent PIM Domains	Single RP within Globally Shared Domains

IPv6 Multicast Addresses (RFC 3513)

Cisco.com



IP Routing for Multicast

Cisco.com

- RPF based on reachability to v6 source same as with v4 multicast
- RPF still protocol independent:
  - Static routes, mroutes
  - Unicast RIB: BGP, ISIS, OSPF, EIGRP, RIP, etc
  - Multi-protocol BGP (mBGP)
  - - support for v6 mcast sub-address family
  - - provide translate function for non-supporting peers

IPv6 Multicast Forwarding

Cisco.com

- PIM-Sparse Mode (PIM-SM)
  - draft-ietf-pim-sm-v2-new-09.txt,
- PIM-Source Specific Mode (PIM-SSM)
  - RFC3569 SSM overview (v6 SSM needs MLDv2)
  - unicast prefix based multicast addresses ff30::/12
  - -> SSM range is ff3X::/32
    - -> current allocation is from ff3X::/96
- PIM-bidirectional Mode (PIM-bidir)
  - draft-ietf-pim-bidir-06.txt

RP Mapping Mechanisms for IPv6 PIM-SM

Cisco.com

- Static RP assignment
- BSR
- Auto-RP – no current draft

Multicast Listener Discover: MLD

Cisco.com

- MLD is equivalent to IGMP in IPv4
- MLD messages are transported over ICMPv6
- Version number confusion:
  - MLDv1 corresponds to IGMPv2
    - RFC 2710
  - MLDv2 corresponds to IGMPv3, needed for SSM
    - draft-vida-mld-v2-08.txt
- MLD snooping
  - draft-ietf-magma-snoop-11.txt
- CGMP for v6 under consideration

You Now Know…

Cisco.com

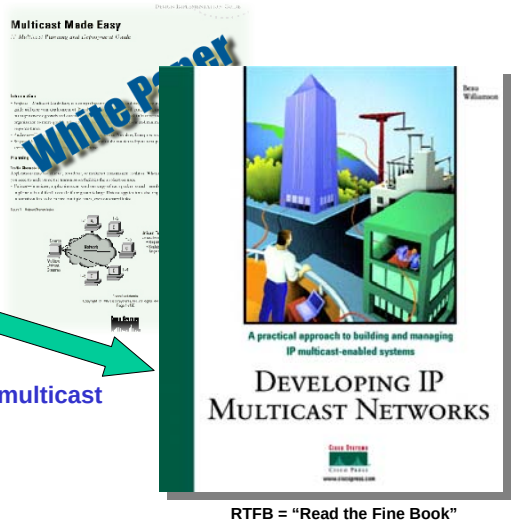
- Why Multicast?
- Multicast Fundamentals
- PIM Protocols
- RP choices
- Multicast at Layer 2
- Interdomain IP Multicast
- Latest Additions IPv6 and MVPN

More Information

Cisco.com

- White Papers
- Web and Mailers
- Cisco Press

RTFB



CCO Multicast page:
http://www.cisco.com/go/ipmulticast
Questions:
cs-ipmulticast@cisco.com
Customer Support Mailing List:
tac@cisco.com

RTFB = "Read the Fine Book"

Complete Your Online Session Evaluation!

Cisco.com

WHAT: Complete an online session evaluation and your name will be entered into a daily drawing

WHY: Win fabulous prizes! Give us your feedback!

WHERE: Go to the Internet stations located throughout the Convention Center

HOW: Winners will be posted on the onsite Networkers Website; four winners per day