

# Understanding IP Multicast Routing

# Keith Bogart

Cisco CCIE #4923

---

# KEY CONCEPTS

- **Multicast communication fundamentals**
- **Multicast addressing & Layer 2 mapping**
- **IGMP & multicast group membership**
- **Protocol Independent Multicast (PIM) Sparse Mode**
- **Rendezvous Point discovery mechanisms**

# MAJOR TOPICS

- Multicast addressing & Layer 2 MAC mapping
- IGMP & multicast membership signaling
- PIM Sparse Mode operation & forwarding behavior
- Designated Routers & multicast forwarding decisions
- Auto-RP & PIM Bootstrap Router RP discovery



# LEARNING OUTCOMES - RECAP

- Understand multicast addressing & traffic flow
- Explain the operation of PIM Sparse Mode
- Configure & verify multicast routing components
- Identify & troubleshoot multicast routing behavior

# PREREQUISITES

- Basic understanding of IP routing concepts
- Familiarity with Cisco enterprise networking environments
- CCNA-level networking knowledge recommended





**LET'S GO!**









# Introduction to Multicast

# Why Multicast?

- » One to Many
- » Many useful applications
- » Locating the Multicast Server
- » Multicast Distribution Methods: Push vs. Pull

# Multicast Addressing (L3)

» IPv4 = Class-D

» Reserved Addresses

» IPv6 = 1111 1111 **xxxx** **yyyy**::  
                  **F**          **F**      Scope(4-bits)  
                                  Flags(4-bits)

- FF02::9(RIPng)
- FF02::5 & FF02::6 (OSPFv3)

# Layer-2 Multicast MAC Addressing

- » Multicast MACs = Partially derived from Multicast IP/IPv6 Addresses.
- » Reserved OUIs for Multicast
  - IPv4 = 0100:5exx:xxxx
  - IPv6 = 3333:xxxx:xxxx
- » What do switches do when receiving multicasts?

# Multicast Terminology

- » Multicast Stream
- » Source
- » Receiver
- » GDA
- » Multicast Distribution Tree
- » PIM



# Introduction to Multicast

# Why Multicast?

- » One to Many
- » Many useful applications
- » Locating the Multicast Server
- » Multicast Distribution Methods: Push vs. Pull



# Multicast Addressing (L3)

» IPv4 = Class-D

» Reserved Addresses

» IPv6 = 1111 1111 **xxxx** **yyyy**::  
                    **F**          **F**      Scope(4-bits)  
                                  Flags(4-bits)

- FF02::9 (RIPng)
- FF02::5 & FF02::6 (OSPFv3)

# Layer-2 Multicast MAC Addressing

- » Multicast MACs = Partially derived from Multicast IP/IPv6 Addresses.
- » Reserved OUIs for Multicast
  - IPv4 = **0100:5e**xx:xxxx
  - IPv6 = **3333**:xxxx:xxxx
- » What do switches do when receiving multicasts?

# Multicast Terminology

- » Multicast Stream
- » Source
- » Receiver
- » GDA
- » Multicast Distribution Tree
- » PIM



# IGMP/MLDP Review

# IGMP Introduction

- » Internet Group Management Protocol
- » Used by Receivers to indicate Mcast interest to directly-connected Routers.
- » Used for IPv4 Multicasts
- » IPv6 Equivalent is MLDP
  - Multicast Listener Discovery Protocol
  - If you learn about one, you've learned about the other.

# IGMP Messages

- » Membership Reports
- » General Membership Queries
- » Leave Group
- » Group-Specific Queries

# IPv6 Multicast with MLDP

- » MLDP = Multicast Listener Discovery Protocol (often called MLD)
- » Structured similar to IGMP
  - MLD version1 based off of IGMPv2 (RFC 2710)
  - MLD version2 based off of IGMPv3 (RFC 3810)
- » Cisco devices support both but default to Version2



# MLDP Message Types

- » MLDP Query
- » MLDP Listener Report
- » MLDP Done
- » MLDP Multicast-Address Specific Query



# Overview of Multicast Routing Protocols

# Why do we need Mcast Routing Protocols?

- » Sources and Receivers...who knows about whom?
- » Multicast Distribution Trees: Concept
- » Creating the tree
  - Dense-Mode protocols
  - Sparse-Mode protocols (triggered joins)

# RPF Checks

- » RPF = Reverse Path Forwarding
- » Why? For Loop Detection
- » How to determine correct Ingress Interface?
  - Based on type of tree used (Source or Core)
  - May be dependent on specific IGP-type



# PIM Sparse-Mode Overview

» Let's look at the topologies we'll be using in our class...

# PIM Overview

- » PIM = Protocol Independent Multicast
- » Two primary varieties:
  - PIM Dense-Mode (RFC 3973)
  - PIM Sparse-Mode (RFC 4601)
- » Objective of both: Build a SPT from Source-to-Receiver(s)



# PIM Sparse-Mode Overview

## » Initially defined in RFC 2117

- Updated several times since
- Currently RFC 4601

## » Relies on the “pull model”

- Builds a Core-Based Tree First (aka Shared-Tree)
- Once Mcast Source is discovered, switches over to SPT.

# PIM Versions

» Original Version of PIM was Version-1

» Version-2 now the de-facto standard

» Similarities:

- Body of PIM packets identical in v1 and v2
- Functionality identical in v1 and v2

» Differences:

- V1 utilized IGMP as transport protocol
- V2 has its own IP Protocol number = 103
- PIM BSR only available in v2

# PIM Comparison to IGP Routing Protocols

## » Similarities

- Dynamic Neighbor Discovery
  - ✓ Hello packets sent to 224.0.0.13
- Information learned via PIM has an expiration timer.

# PIM Comparison to IGP Routing Protocols

## » Differences

- IGPs only care about destinations...PIM cares about Sources and Destinations
- Relies on a host (receiver) or source (server) to trigger any PIM information exchange.
- Populates the Mroute Table
- Performs RPF checks by default
- Cannot work alone: Requires an IGP

# PIM Rendezvous Points

» What is the role of a PIM-RP?

» How does a router know who the RP is?



# Creating the Shared Tree

# Initial PIM Sparse-Mode Configuration

- » Multicast Routing must be globally enabled.

```
Router-1(config)#ip multicast-routing  
Router-1(config)#
```

- » PIM Sparse-Mode must be enabled per-interface.

```
Router-1(config-if)#ip pim sparse-mode
```

- » PIM RP must be defined or learned.

```
Router-1(config)#ip pim rp-address 7.7.7.7
```



# Discovering Neighbors

» PIM will not allow forwarding/receipt of multicast traffic on interfaces unless:

- Another PIM Neighbor has been discovered on that interface.
- A directly-connected Receiver/Source resides on that interface.

» PIM “Hello” packets used to discover PIM neighbors.

# PIM Joins

- » PIM doesn't create a multicast distribution tree until someone wishes to "Join" it.
- » PIM "Join" packets used for this process
  - (\*,G) Joins
  - (S,G) Joins

# Mroute “State”

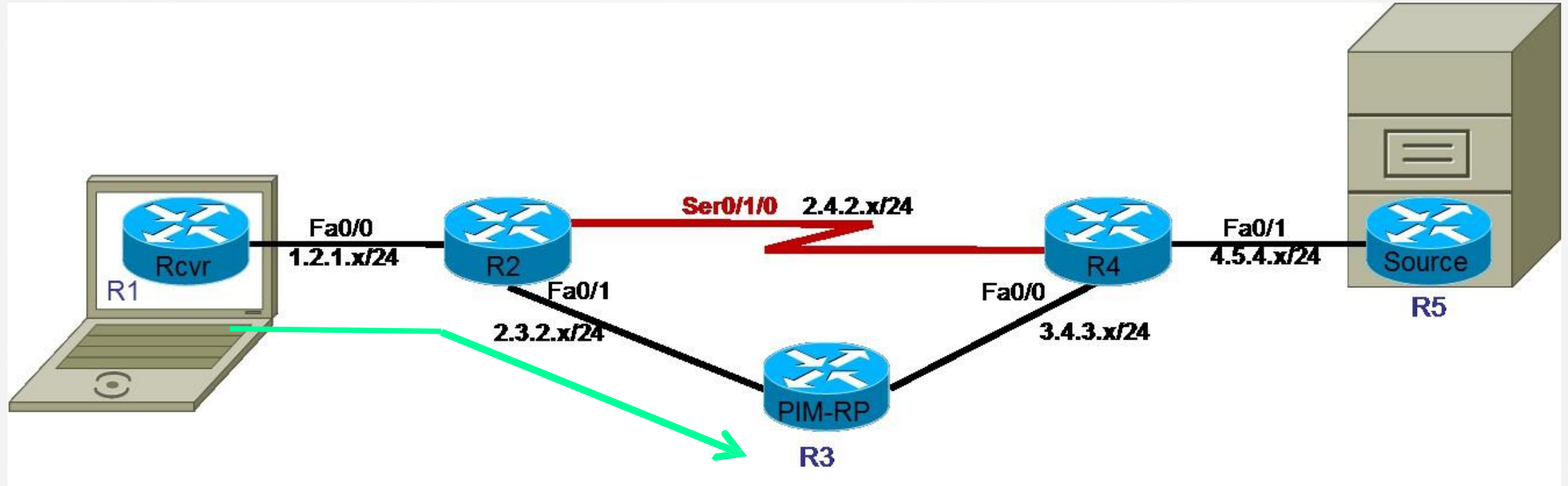
- » Multicast Forwarding determined by Multicast Routing Table
- » Entries in this table called, “Mroutes”
- » These entries are created by:
  - Receiving PIM Joins / IGMP Membership Report
  - Receiving Multicast Traffic
  - Dynamic/Automatic Creation

# Creating the Shared Tree

» Creation of Shared Tree requires three things:

- Knowledge of who the RP is.
- Knowledge of how to reach the RP (unicast routing)
- At least one request from a Multicast Receiver

# The Shared Tree



# PIM (\*,G) Joins

- » All PIM Join packets have the same structure whether they are (\*,G) or (S,G) Joins.
- » What triggers them?
- » What Mroute state do they create?

# PIM (\*,G) Joins

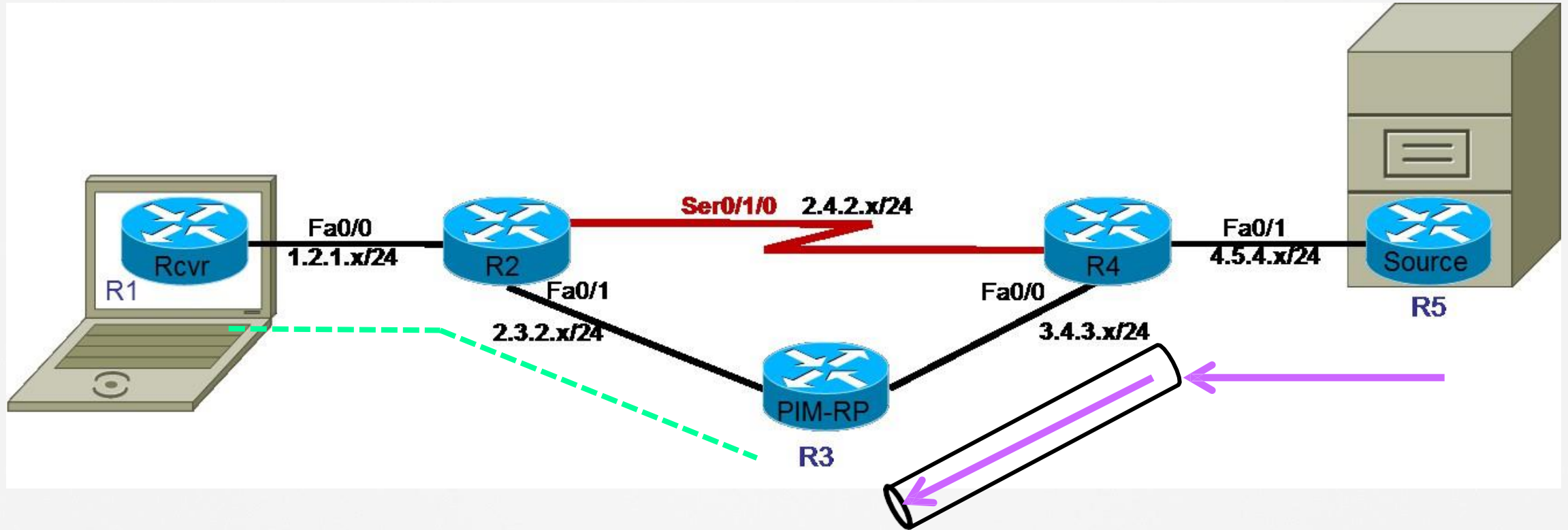
- » Re-Sent every Join/Prune-Period
  - 60-seconds (default)
- » What if RPF info to RP changes?
- » What deletes this Mroute state?



# Registering Sources



# Let the Source Begin!!



# PIM Registration

- » Once Mcast data is received, router connected to Source must “register” those packets with RP.
- » RP uses “Register-Stop” message after its SPT is confirmed and open.
- » Details of PIM Registration Process.

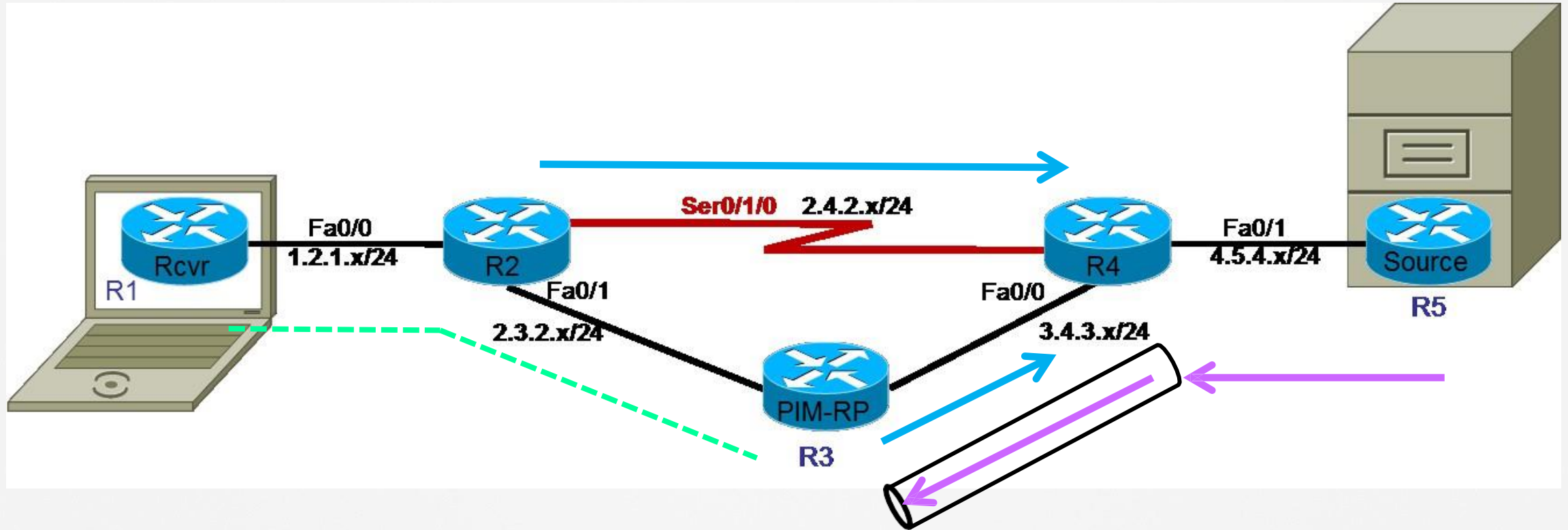
# Null-Register

- » Method used to periodically “remind” the RP that this flow is still going.
- » RP may need to receive a stream again before the “Register-Suppression” timer expires.
- » Router send periodic Null-Registers every 5-seconds (Probe-Time) to RP.



# Joining the Shortest Path Tree

# Joining the Shortest-Path Tree



# PIM (S,G) Joins

- » Once Mcast data is received, we know the unicast “Source” of that data.
- » (S,G) Joins used to open up SPT.
- » Details of (S,G) Joins.
  - What triggers them?
  - What state do they create in Mroute Table?

# Deletion of (S,G) Mroute State

- » (S,G) state created, but must also be maintained.
- » Downstream routers will continue to send periodic (S,G) Joins if interested.
  - Determined by “Join/Prune-Period”
  - 60-seconds (default)
- » When is (S,G) state deleted?
  - Periodic Joins from downstream neighbors stop
  - Mcast Source stops transmitting
  - Receipt of PIM Prune and no outgoing interfaces.

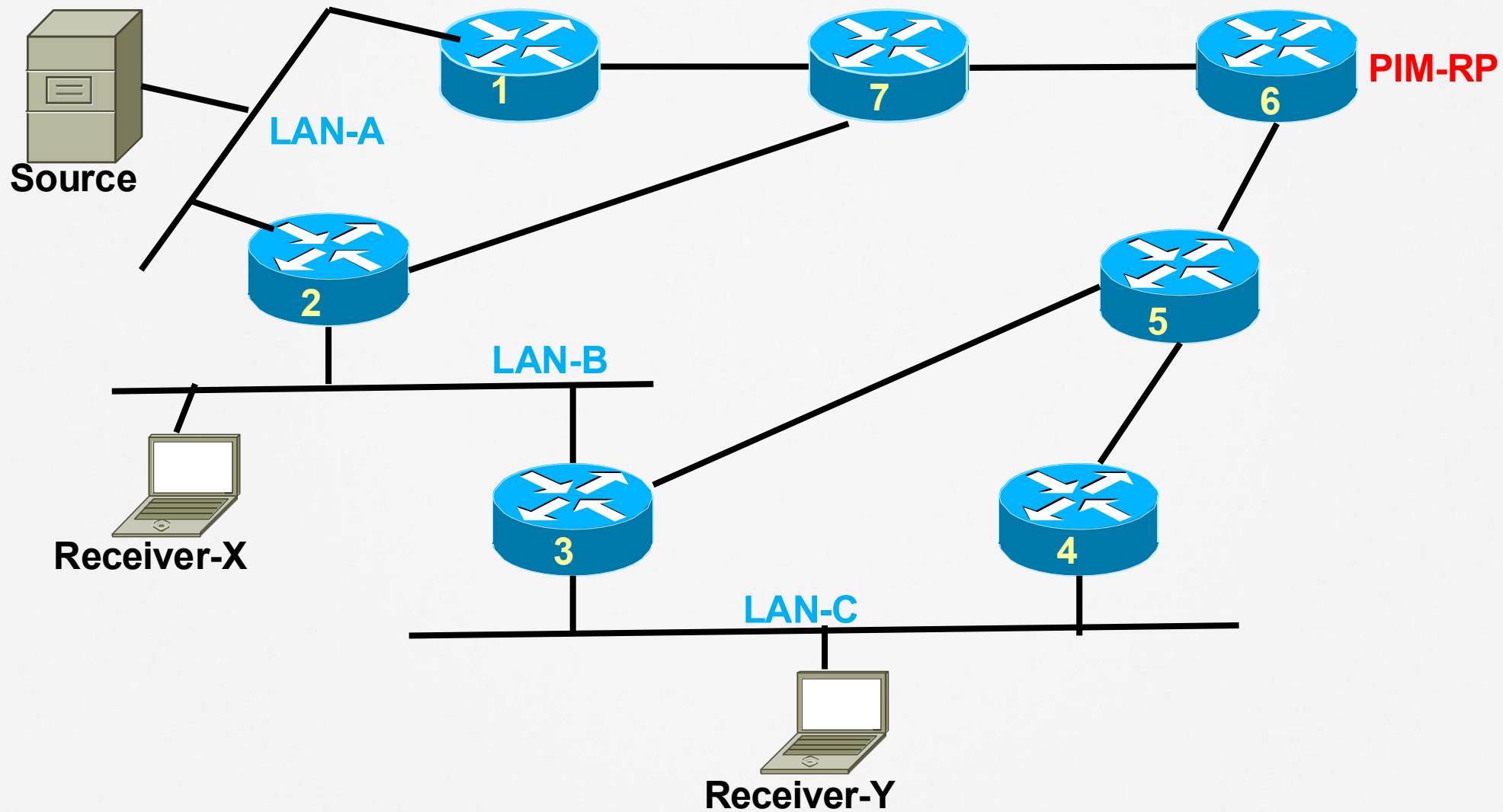
# PIM (S,G) and (\*,G) Prunes

- » PIM Prune uses same packet format as PIM Joins
- » Used to encourage upstream neighbors to stop forwarding Mcast traffic to us.
- » Details of Prunes.
  - What triggers them?
  - What state do they create in Mroute Table?
  - Other details



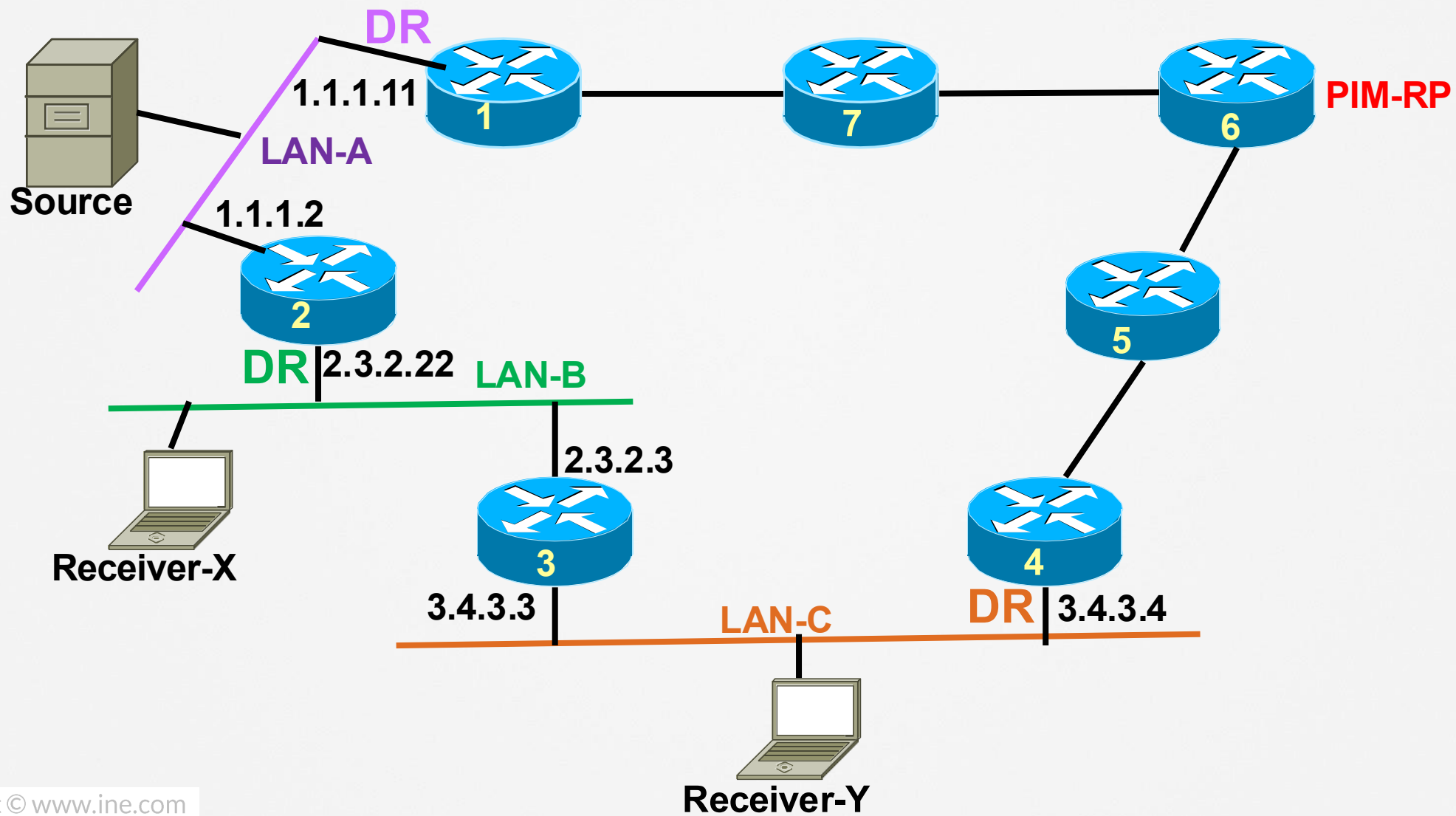


# Designated Routers and Forwarders



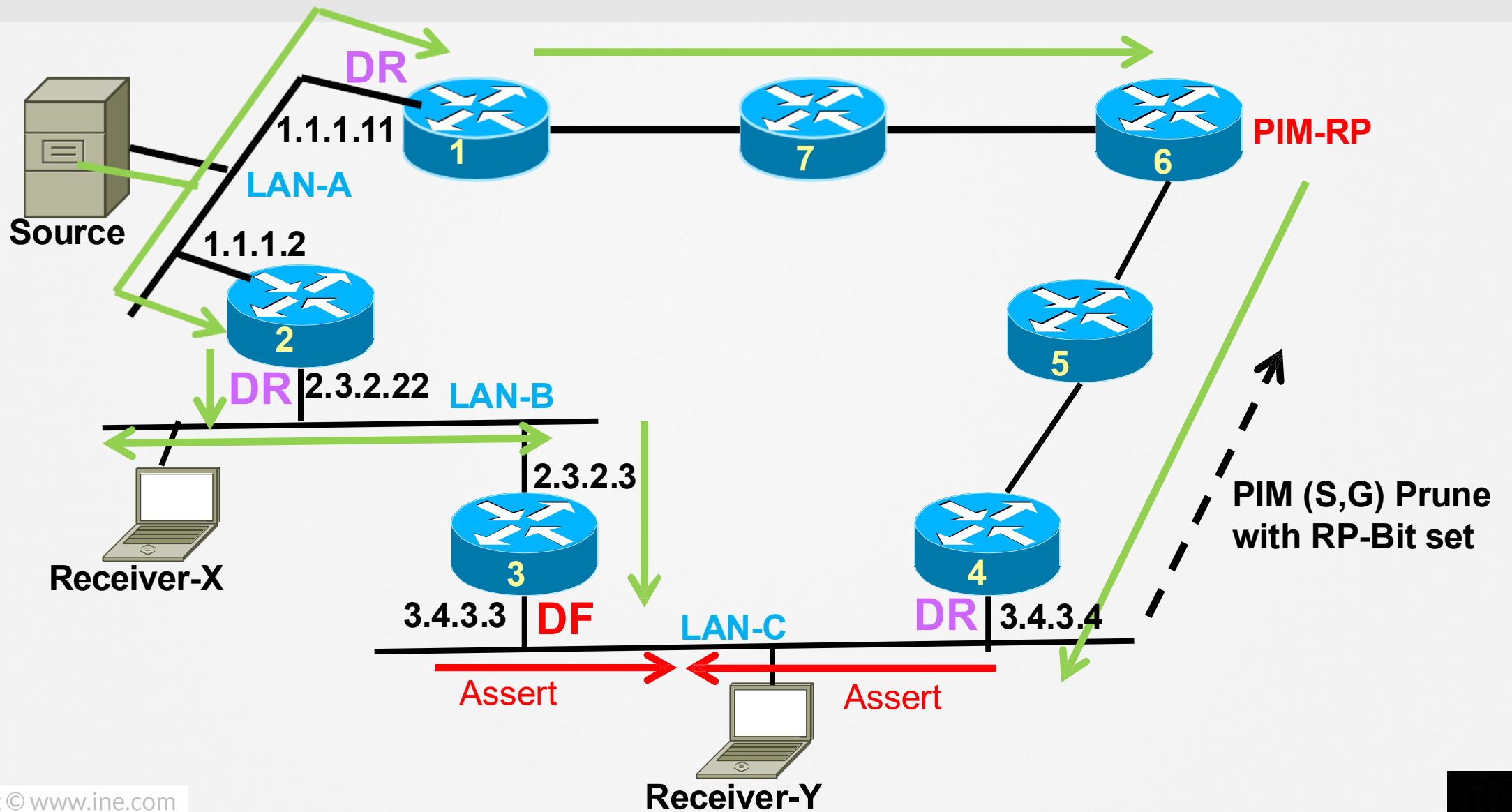
# PIM Designated Router

- » When two routers share the same LAN one of them will be elected as the PIM\_DR.
- » PIM Hellos used for this election
- » DR responsible for sending
  - PIM Joins/Prunes up the Shared and SPT trees.
  - PIM Registers to the RP



# PIM Designated Forwarder

- » If a router is already forwarding Mcast data onto a LAN, and then RECEIVES an incoming mcast packet (from that same source/group) on that LAN...that packet must be a duplicate.
- » A PIM Designated Forwarder will be elected
- » PIM Assert messages used in this process.





# PIM-related and general Multicast Commands

# Verifying PIM Sparse-Mode

- » Show ip pim interface
- » Show ip pim neighbor
- » Show ip pim rp
- » Show ip pim rp mapping



# Modifying PIM Sparse-Mode Behavior

- » (config)#ip pim accept-register list <acl>
- » (config)#ip pim register-rate-limit <bps>
- » (config)#ip pim register-source <interface>
- » (config)#ip pim spt-threshold <0 | infinity>
- » (config)#ip pim sparse sg-expiry-timer <181-57600>
- » (config-if)#ip pim nbma-mode

# “Show ip mroute” flags

- » C = Connected Receiver (or Source)
- » L-Flag: Router itself has joined the group. (igmp join-group)
- » F-Flag: PIM-Register Flag
  - If the flow is currently being registered right now, you'll also see the keyword, “Registering” in the (\*,G) entry.
  - If you only see the F-flag but not the word, “Registering” it means this flow has completed the Registration process and received a “Register-Stop” message from the PIM-RP.

# “Show ip mroute” flags

- » J-Flag: Router has attempted to “Join” the respective tree.
  - If seen in the (\*,G) output, indicates the respective router is trying (or has succeeded) in transmitting a (\*,G) Join up the Shortest-Path-Tree towards the PIM-RP.
  - If seen in the (S,G) output, indicates the respective router is trying (or has succeeded) in sending a (S,G) Join up the SPT.
  - This flag ALONE is not PROOF that any mcast packets have arrived on either tree yet.
- » T-Flag: Data has been received on the SPT
  - Received at least one mcast data packet on the SPT.

# “Show ip mroute” flags (2)

## » R-Flag: RP-Bit Set

- You will see this on an intermediary router (not the PIM-RP) that is only on the Shared path but NOT on the SPT.
- Only seen in (S,G) entries that have no interfaces in OIL.



# Auto-RP (Overview)

# Auto-RP Overview

- » Provides a method for dynamically learning of RPs and RP redundancy.
- » Cisco-Proprietary
- » Routers play one of three (3) roles:
  - RP-Candidates
  - Mapping-Agent
  - Other (neither of the above)
- » Utilizes two, IANA-reserved Multicast Addresses
  - 224.0.1.39 & 224.0.1.40

# RP-Candidates

- » Manually configured
- » Send “**RP-Announce**” messages to Mapping-Agent
- » Can be configured to announce RP-Capability only for certain GDAs.
- » Default = Announce RP-Capability for 224.0.0.0/4

# Mapping Agents

- » Listen for RP-Announcements.
- » If more-than-one RP announces capability for same GDA, Mapping-Agent elects the best one.
- » Transmit “RP-Discovery” messages with list of all, elected RPs for each GDA.
  - Routers listen to these messages and dynamically learn of available RPs for each group.



# Candidates and Agents

- » A single router can be configured to play both roles of RP-Candidate and Mapping-Agent.
- » PIM Dense-Mode used to propagate RP-Announce and RP-Discovery messages.
- » All interfaces must be in Sparse-Dense mode
  - Alternative: `(config)# ip pim autorp listener`



# Auto-RP (RP-Candidates)

# RP-Announcements

- » Encapsulated in UDP
- » UDP Port 496 (source & destination)
- » Sent every 60-secs with an advertised holdtime of 180-secs
- » Sent to 224.0.1.39 (only Mapping Agents pay attention to this).

# RP-Announcement Body

» Contain the following informational fields:

- RP Address
- Group Range
- Holdtime (3x Hello interval)

# RP-Candidate: Configuration

- » (config)# ip pim send-rp-announce <intfc> scope <tth> [group-list acl]
- » Interface: Controls source IP address of RP-Announce messages
- » Scope: Sets TTL in RP-Announce Messages
- » Group-List-ACL: Limit announcements to certain GDAs.



# Auto-RP (Mapping-Agents)

# RP-Discovery

» Mapping-Agents collect all RP-Announce messages and elect an RP for each GDA.

- Highest IP address wins.

» Transmit RP-Discovery messages to 224.0.1.40

- Also encapsulated in UDP
- All PIM routers automatically listen to this group.
- Flooded via Dense-Mode

# RP-Discovery Body

» Contain the following informational fields:

```
▣ Cisco Auto-RP
  ▣ Version: 1 or 1+, Packet type: RP mapping
    0001 .... = Protocol version: 1 or 1+ (1)
    .... 0010 = Packet type: RP mapping (2)
    RP count: 1
    Holdtime: 46 seconds
    Reserved: 0x0
  ▣ RP 2.4.2.2: 1 group
    RP address: 2.4.2.2 (2.4.2.2)
    .... ..11 = Version: Dual version 1 and 2 (3)
    Number of groups this RP maps to: 1
  ▣ Group 224.0.0.0/4 (Positive group prefix)
    .... ...0 = Sign: Positive group prefix (0)
    Mask length: 4
    Prefix: 224.0.0.0 (224.0.0.0)
```



# Configuration & Verification

## » Configuration:

```
Router(config)#ip pim send-rp-discovery scope 10 interval 15
```

## » Verification:

```
Router#sho ip pim rp mapping
PIM Group-to-RP Mappings
This system is an RP-mapping agent

Group(s) 224.0.0.0/4
  RP 2.4.2.2 (?), v2v1
    Info source: 2.4.2.2 (?), elected via Auto-RP
      Uptime: 00:16:18, expires: 00:02:40
Router#
```

# Auto-RP Filters

- » Mapping-Agents collect all RP-Announce messages from any Candidate-RP
  - Rogue RPs could use this to their advantage.
- » Mapping agent can be configured to accept RP\_announce messages:
  - Only from authorized RPs
  - Only for authorized groups.
  - Or both.

# Auto-RP Filters - Configuration

```
(config)#ip pim rp-announce-filter rp-list 1 group-list 2
```

```
access-list 1 permit 10.0.0.1
```

```
access-list 1 permit 10.0.0.2
```

```
access-list 2 permit 224.0.0.0 15.255.255.255
```



# Auto-RP (RP Of Last Resort)

# Why the need for an RP of last resort?

## » Routers running Auto-RP have interfaces configured for Sparse-Dense.

- This means any Mcast traffic received for which there is no, known RP must be flooded via PIM Dense-Mode.

## » RP-of-Last-Resort

- Also called, “Sink RP”
- Forces PIM-Registration of mcast packets that would normally be flooded.
- Can be a “black hole” if desired

# Sink RP - Configuration

```
ip pim rp-address 1.1.1.1 20
```

```
access-list 20 deny 224.0.1.39
```

```
access-list 20 deny 224.0.1.40
```

```
access-list 20 permit 224.0.0.0 15.255.255.255
```

```
ip pim rp-address 2.2.2.2 5 override
```

```
access-list 5 permit 227.7.7.7
```



# PIM-BSR Overview



# PIM-BSR vs Auto-RP (Similarities)

» BSR = Bootstrap Router

» PIM-BSR and Auto-RP Similarities:

Provide dynamic discovery of PIM-RPs

- Candidate-RPs announce their RP Capability
- PIM-BSR transmits messages that allow other routers to learn of potential PIM-RPs (same objective as Auto-RP Agent).



# PIM-BSR vs Auto-RP (Differences)

## » PIM-BSR = IETF Standard

- RFC 5059
- Auto-RP = Cisco Proprietary

## » PIM-BSR = Developed for PIMv2

- Auto-RP: Works in PIMv1 or v2

## » PIM-BSR = No ability to define “Scope”

- Auto-RP demands this as part of configuration

# PIM-BSR vs Auto-RP (Differences) 2

- » PIM-BSR = All interfaces can be in PIM Sparse-Mode
  - Auto-RP requires Sparse-Dense Mode
  - Possible risk of flooding Multicast Traffic
- » PIM-BSR ...there can be only one.
  - Single PIM-BSR elected from multiple Candidate-BSRs
  - Auto-RP: Can have multiple Mapping-Agents
- » PIM-BSR = 224.0.0.13 (or unicast)
- » Auto-RP = Reserved Addresses (224.0.1.39 and .40)



# PIM-BSR Election and Message Formats

# BSR Election Process

» Flooding of PIM “Bootstrap” messages throughout PIM domain.

- Destination = 224.0.0.13

» BSR elected as:

- Router with highest Priority (0-255)
- If tie...highest IP address.

# PIM-BSR Configuration

## » PIM-BSR:

```
(config)#ip pim bsr-candidate fast0/0 0 50
```

Source IP Address of messages.

Hash-Mask  
Length

Priority

## » RP-Candidate:

```
(config)#ip pim rp-candidate FastEthernet0/0 group-list 1 priority 30
```

Source IP Address of messages.

Optional reference  
to ACL

# PIM-BSR Monitoring

» Show ip pim bsr

» Show ip pim rp mapping

# PIM-BSR Message Types

## » PIM Bootstrap Message

- Contains listing of all known RPs, their advertised GDAs, and Priority
- All other routers independently decide which RP to use from this list.

## » PIM Candidate-RP-Advertisements

- Unicast from Candidate-RPs to BSR

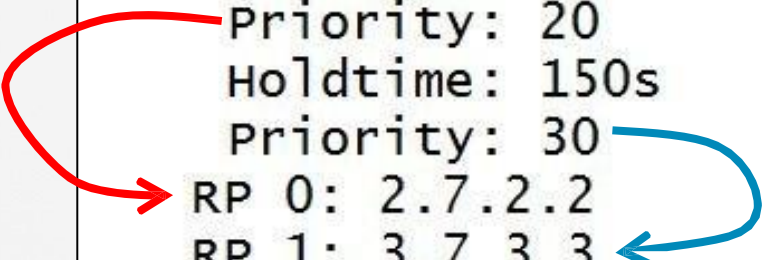
# PIM Bootstrap Message - Details

- » PIM Type = 4
- » Generated only by BSRs
- » Sent to 224.0.0.13
- » Forwarded via PIM Sparse-Mode by all routers.
- » Contains listing of all, known RPs



# PIM Bootstrap Message

```
Protocol Independent Multicast
0010 .... = Version: 2
.... 0100 = Type: Bootstrap (4)
Reserved byte(s): 00
Checksum: 0xa668 [correct]
PIM options
  Fragment tag: 0x0644
  Hash mask len: 0
  BSR priority: 0
  BSR: 6.7.6.6 (6.7.6.6)
  Group 0: 224.0.0.0/4
    RP count: 2
    FRP count: 2
    Holdtime: 150s
    Priority: 20
    Holdtime: 150s
    Priority: 30
    RP 0: 2.7.2.2
    RP 1: 3.7.3.3
```



# PIM Candidate-RP Messages - Details

- » PIM Type = 8
- » Generated only by Candidate-RPs
- » Unicast to elected BSR
- » Includes RP Priority
- » Uses PIM as transport.

# PIM Candidate-RP Message

```
▣ Protocol Independent Multicast
  0010 .... = Version: 2
  .... 1000 = Type: Candidate-RP-Advertisement (8)
  Reserved byte(s): 00
  Checksum: 0xee3c [correct]
▣ PIM options
  Prefix-count: 1
  Priority: 30
  Holdtime: 150s
  RP: 3.7.3.3 (3.7.3.3)
  Group 0: 224.0.0.0/4
```



# Choosing an RP from among the RP-Set

# Which RP to use?

- » Individual routers decide which RP to use for any given GDA from among the total set of RPs listed in the Bootstrap message.
- » If two-or-more RPs contend for the same GDA:
  - RP with lowest priority wins
  - If priority is the same, a hash function is performed.
- » With default hash-mask-length of zero (0) a single RP will always be selected for every GDA.

# Breaking a Tie

- » For every RP IP calculate the hash function value based on the following:
  - Value 1=hash(Group&Mask,R1)
  - Value 2=hash(Group&Mask,R2)
- » Exact algorithm specified in RFC 4601, section 4.7.2
- » Highest hash value wins
- » If a tie, highest IP address wins



# Predicting the RP in Advance

```
Rtr-Sw#sho ip pim rp-hash 228.70.87.65
RP 2.7.2.2 (?), v2
  Info source: 5.7.5.5 (?), via bootstrap, priority 30, holdtime 150
  Uptime: 01:23:56, expires: 00:01:50
PIMv2 Hash Value (mask 0.0.0.0)
  RP 2.7.2.2, via bootstrap, priority 30, hash value 1660980568
  RP 3.7.3.3, via bootstrap, priority 30, hash value 1418243051
```



Thank you!!



# Understanding IP Multicast Routing - Summary

# KEY CONCEPTS - RECAP

- Multicast addressing & traffic distribution
- IGMP group membership signaling
- PIM Sparse Mode operation
- Shared trees & shortest path trees
- Rendezvous Point discovery & control



# LEARNING OUTCOMES - RECAP

- Understand multicast addressing & traffic flow
- Explain the operation of PIM Sparse Mode
- Configure & verify multicast routing components
- Identify & troubleshoot multicast routing behavior

# REAL-WORLD APPLICATIONS

- **Live video & streaming distribution**
- **Financial market data distribution systems**
- **Enterprise content delivery**
- **Large-scale event broadcasting**

# NEXT STEPS

- **Continue with additional ENCOR Learning Path topics**
- **Practice configuration & troubleshooting in labs**
- **Explore advanced multicast design & optimization**

**THANKS FOR WATCHING!**

