



# Multicast Technology and Troubleshooting introduction

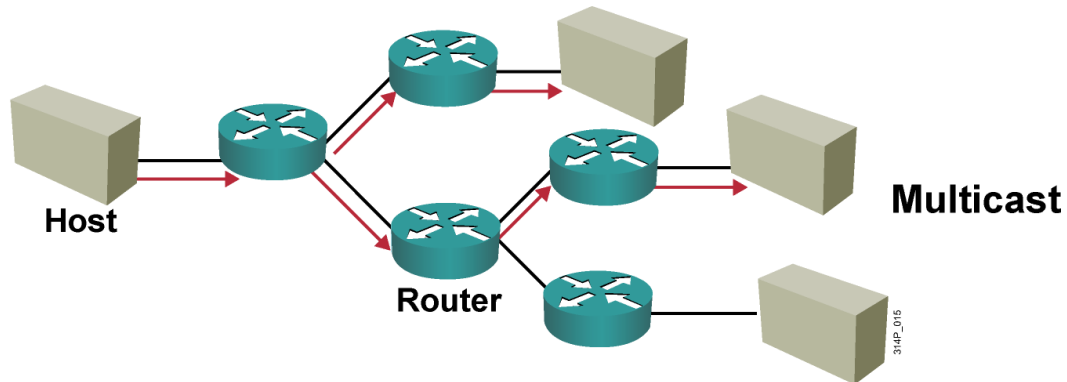
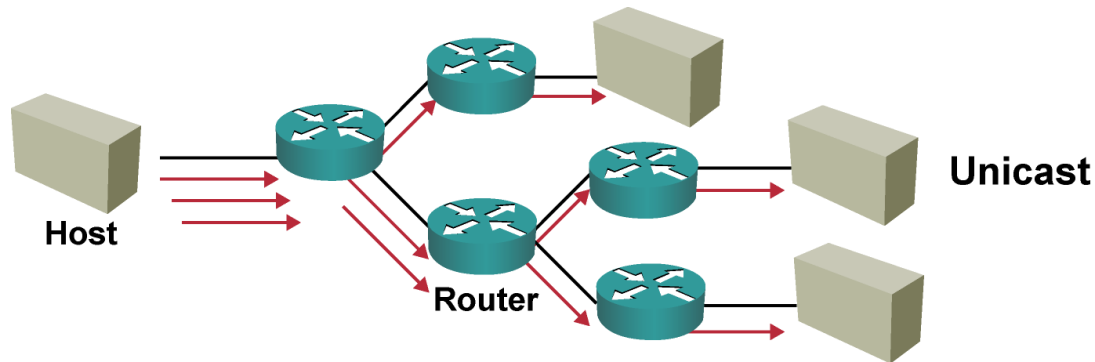
Chen Fei  
High Touch Engineer

2014/7

# Agenda

- Multicast fundamental knowledge
- Multicast technology on Service Provider
- Layer2 Multicast technology
- Layer3 Multicast technology
- Multicast High availability
- Multicast case study

# Why Multicast?



1. Enhanced Efficiency
2. Optimized Performance
3. Distributed Application

# Multicast Application

1. Multi-media and stream media application
2. Training and cooperation communication
3. Data exchanges and replication

# IP Multicast Addressing

- Range from 224.0.0.0 through 239.255.255.255

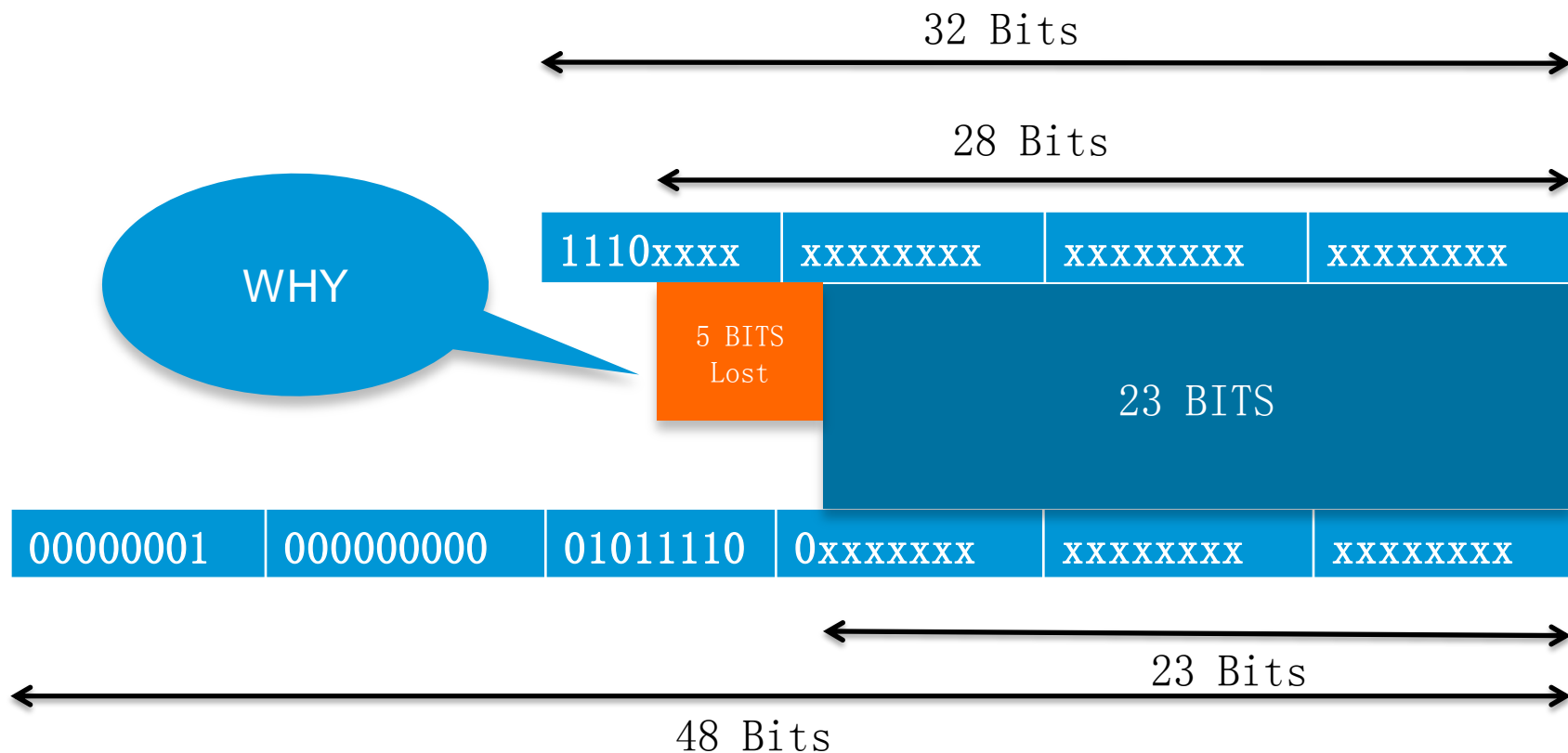
Class D address (high-order three bits are set)



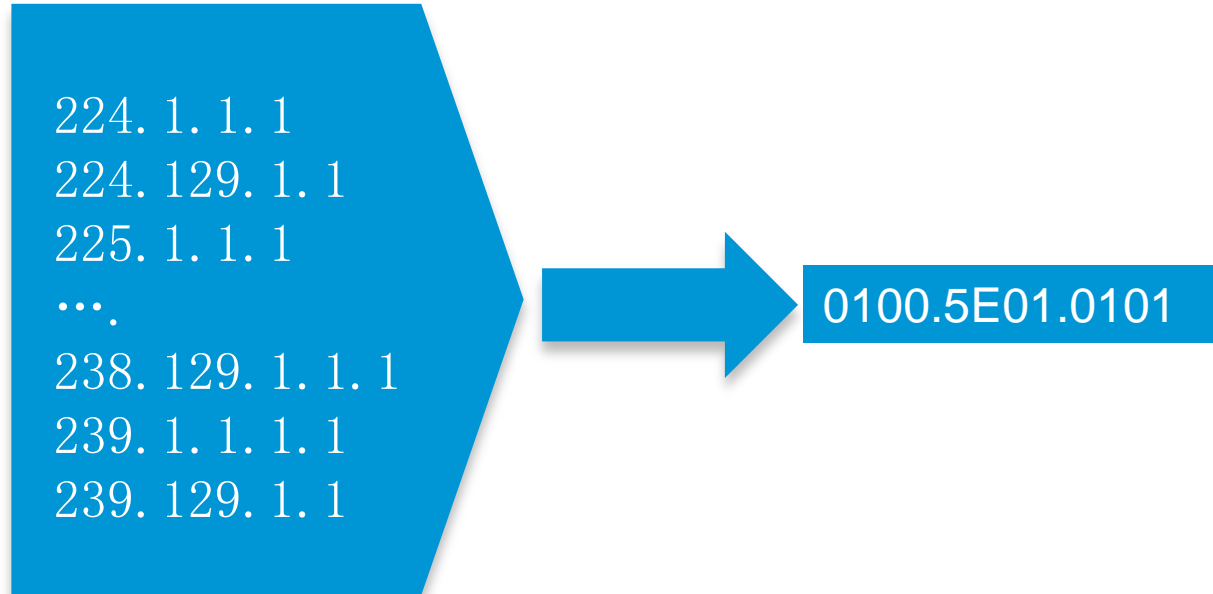
Reserved address:

- 224.0.0.0 - 224.0.0.255  
network protocols on a local network segment  
regardless of their time-to-live [TTL] values
- 239.0.0.0 - 239.255.255.255 administratively  
scoped addresses for use in private multicast  
domains.

# Multicast MAC Address Mapping



# Multicast MAC Address Mapping



Question:

1. Impact host and device CPU utilization
2. Multicast flooding control issue

# Multicast Distribution Tree

## Characteristic of Distribution Trees

- Source or Shortest Path trees

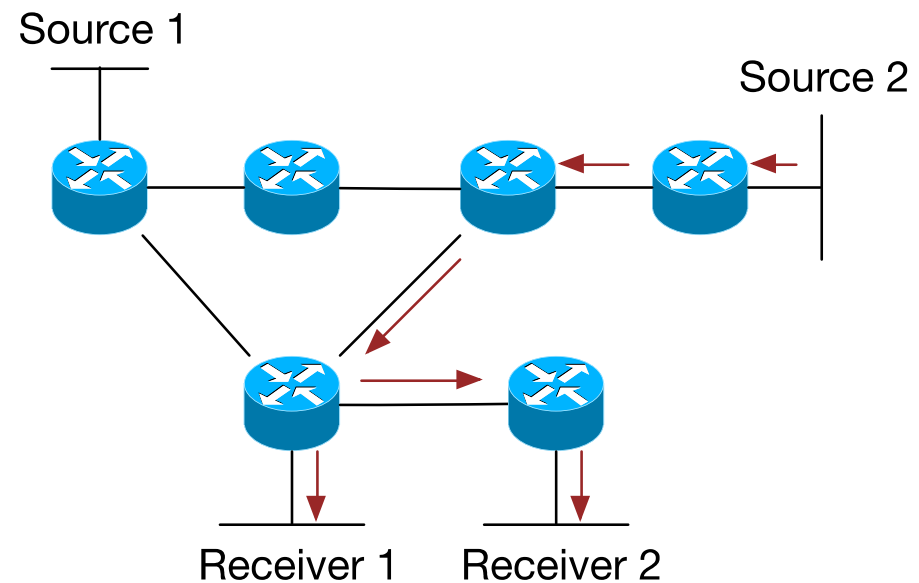
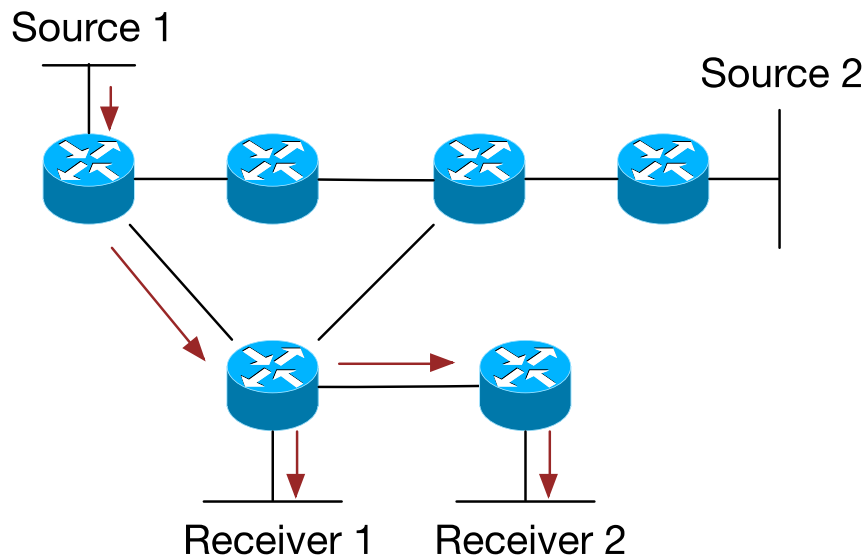
Uses more memory  $O(S \times G)$  but you get optimal paths from source to all receivers; minimizes delay

- Shared trees

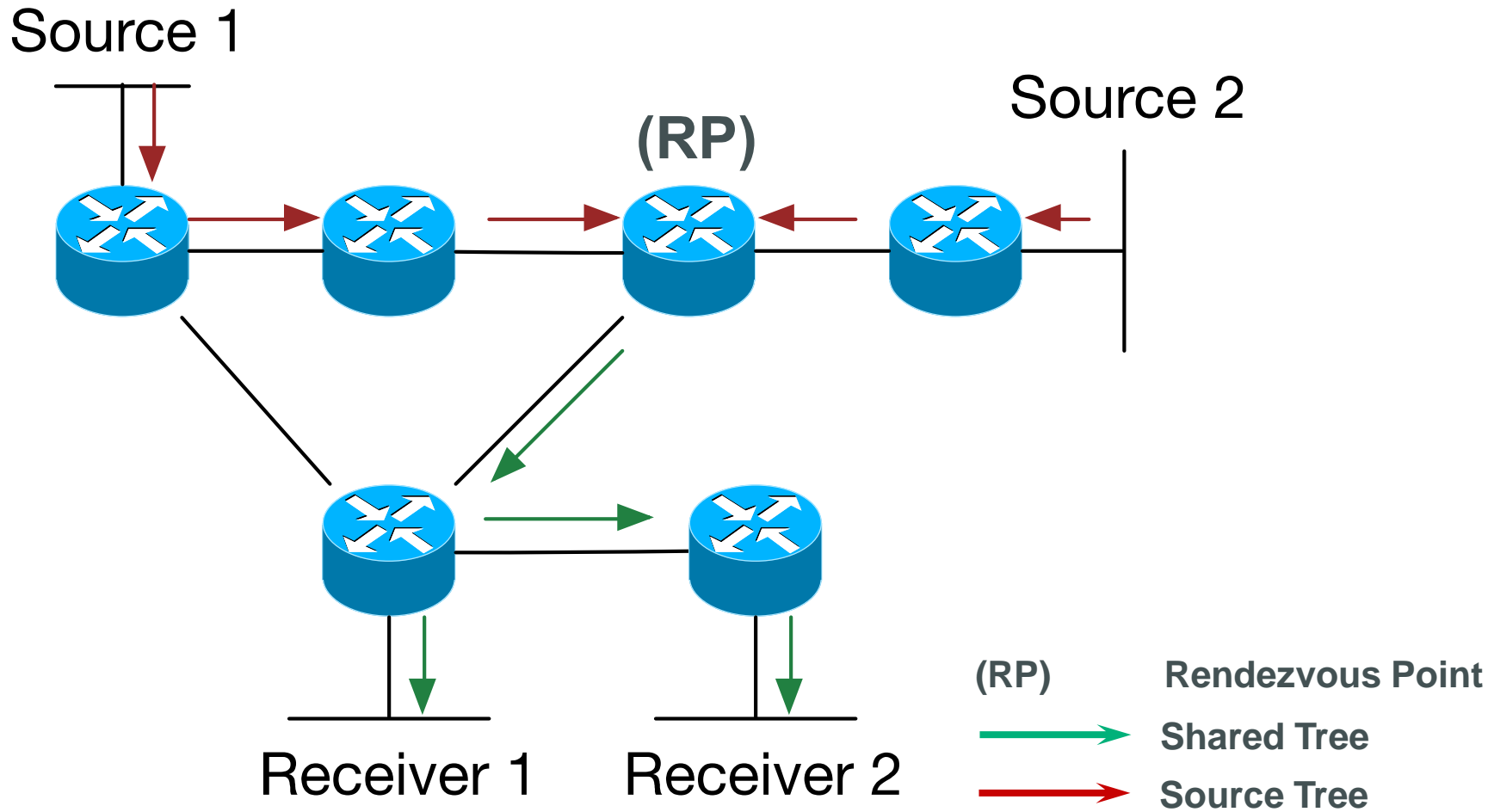
Uses less memory  $O(G)$  but you may get sub-optimal paths from source to all receivers; may introduce extra delay



# Source or Shortest Path trees



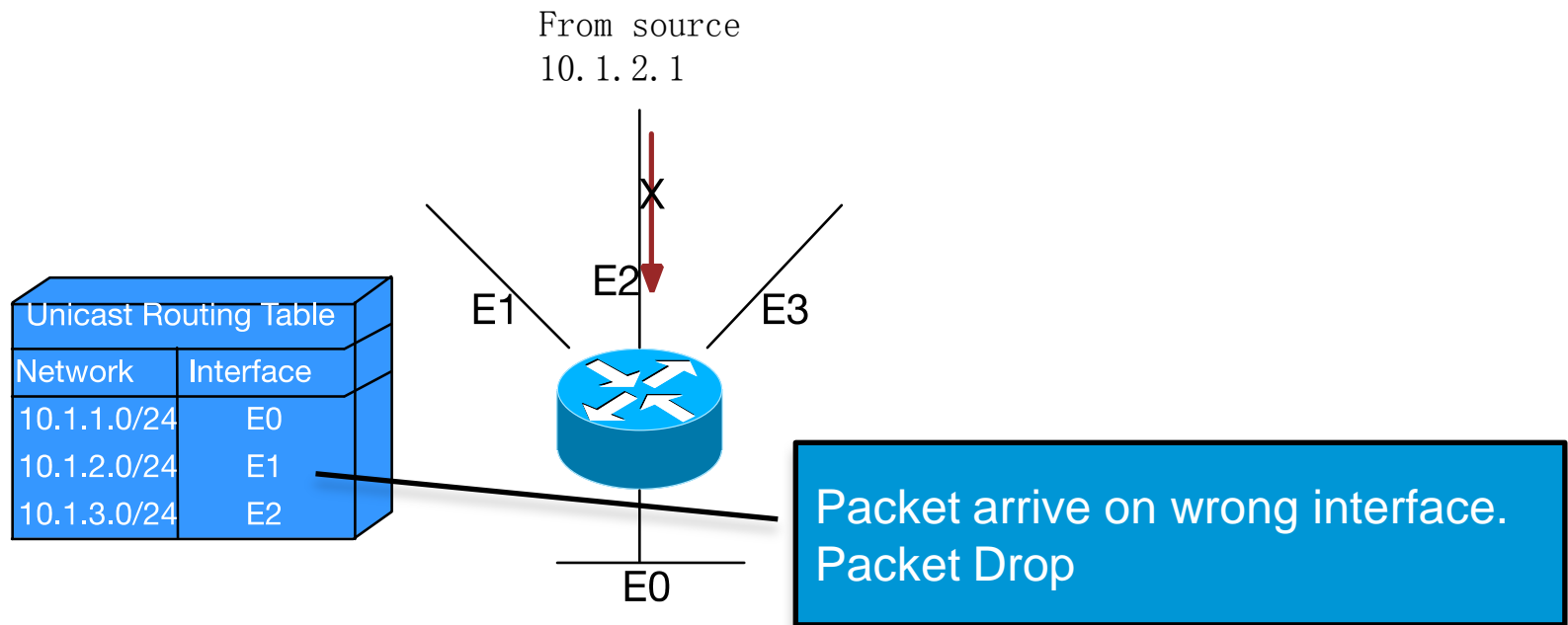
# Shared trees



# Multicast Forwarding

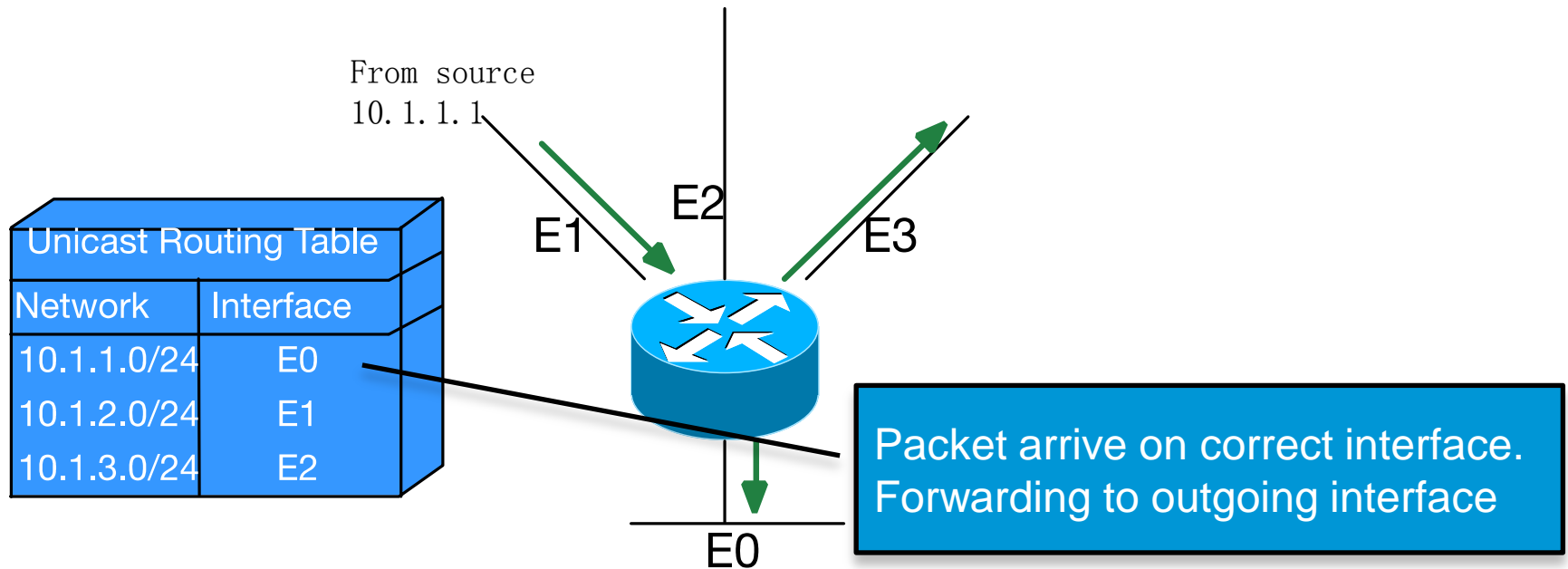
## 1. Reverse Path Forwarding (RPF)

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



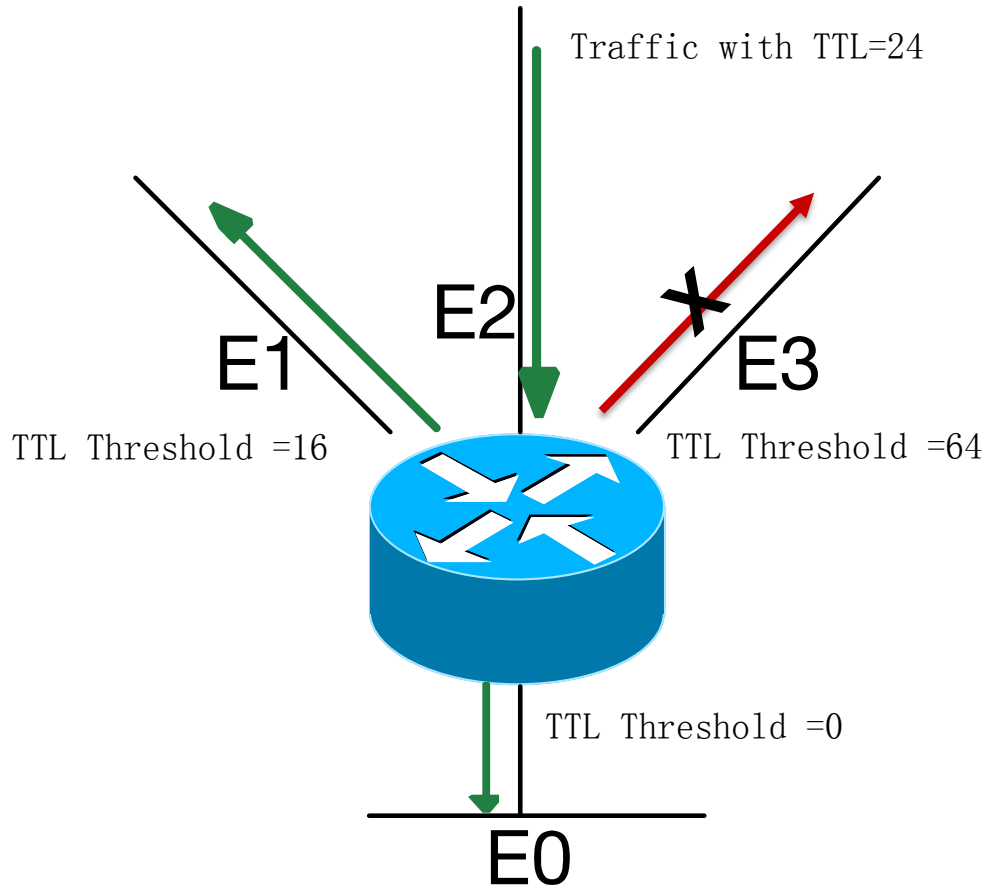
# Multicast Forwarding

## 1. Reverse Path Forwarding (RPF)

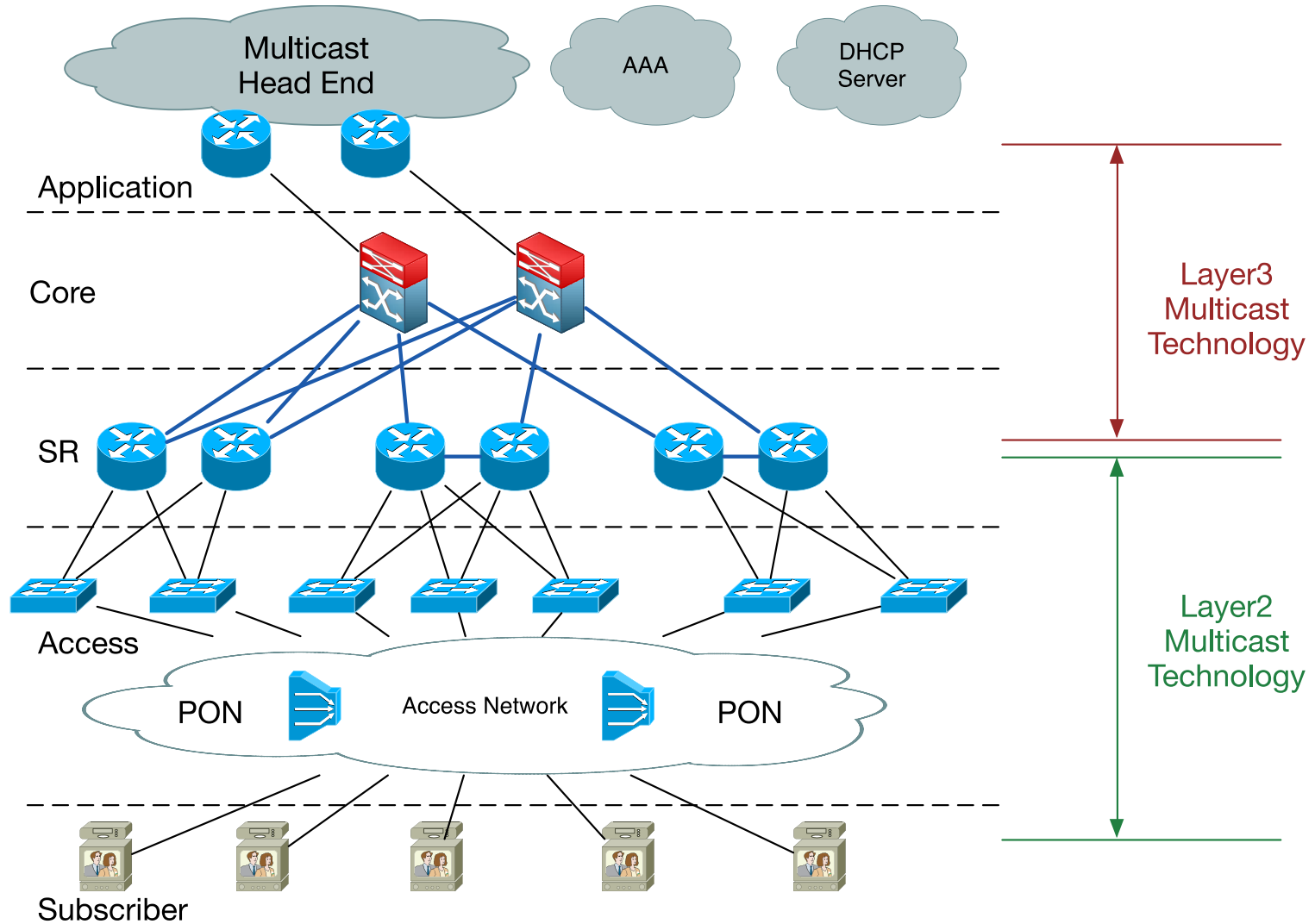


# Multicast Forwarding

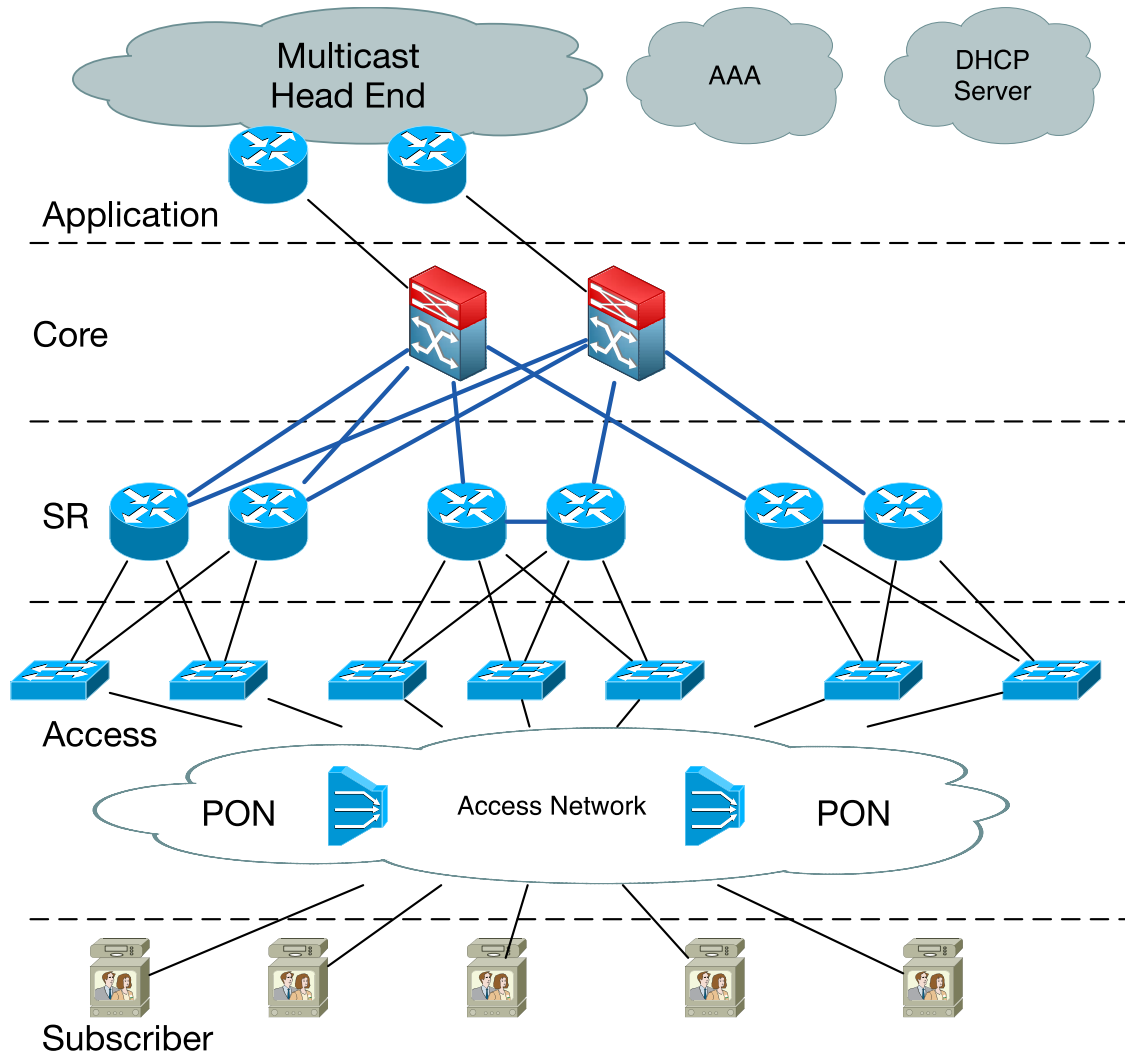
## 2. TTL Threshold



# Multicast Application on SP



# Layer2 Multicast Technology



1. IGMP
2. IGMP Snooping
3. IGMP Proxy
4. Multicast VLAN Registration

# IGMP Introduction

## 1. Purpose

permit hosts to communicate their desire to receive multicast traffic to the IP Multicast router(s) on the local network

## 2. IGMP Version

RFC 1112 specifies IGMP version 1

RFC 2236 specifies IGMP version 2

RFC 3367 specifies IGMP version 3

## 3. IGMP version 1

protocol message :

### **Membership Queries**

sent by the router to the “All-Hosts” (224.0.0.1) multicast address to solicit what multicast groups have active receivers on the local network.

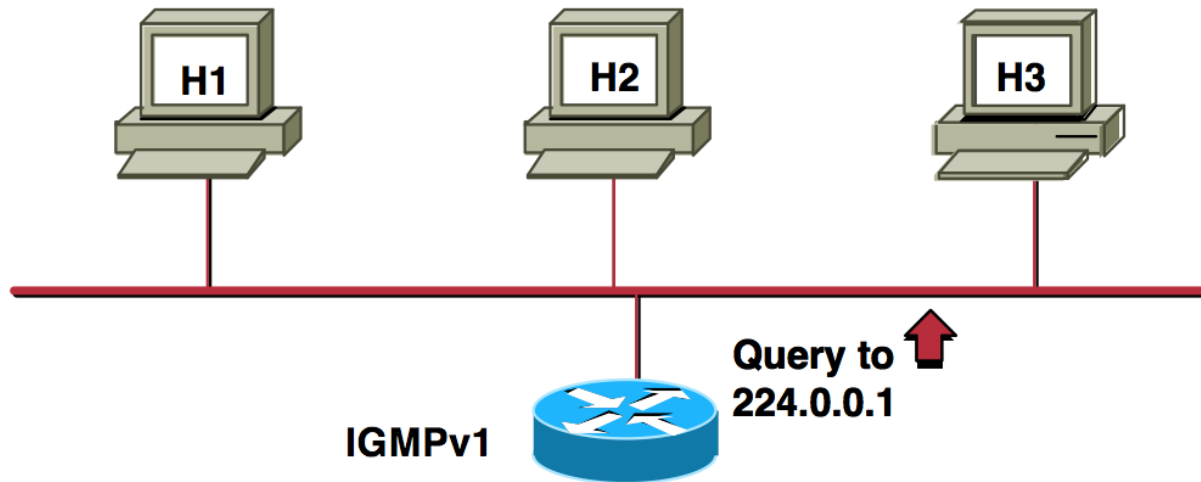
### **Membership Reports**

by hosts wishing to receive traffic for a specific multicast group



# Question of IGMP Version 1

silently leave group , inefficient



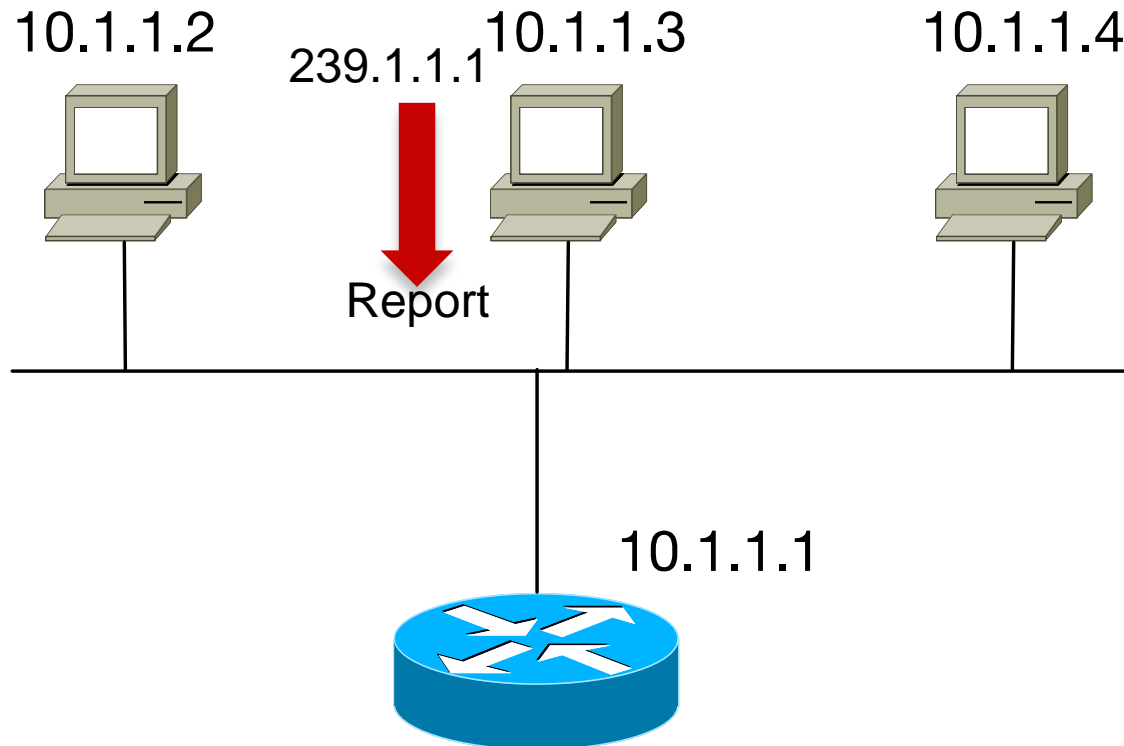
- Router sends periodic queries
- Hosts silently leave group
- Router continues sending periodic queries
- No Reports for group received by router
- Group times out

# IGMP Version 2

To solve inefficient issue, adding two message

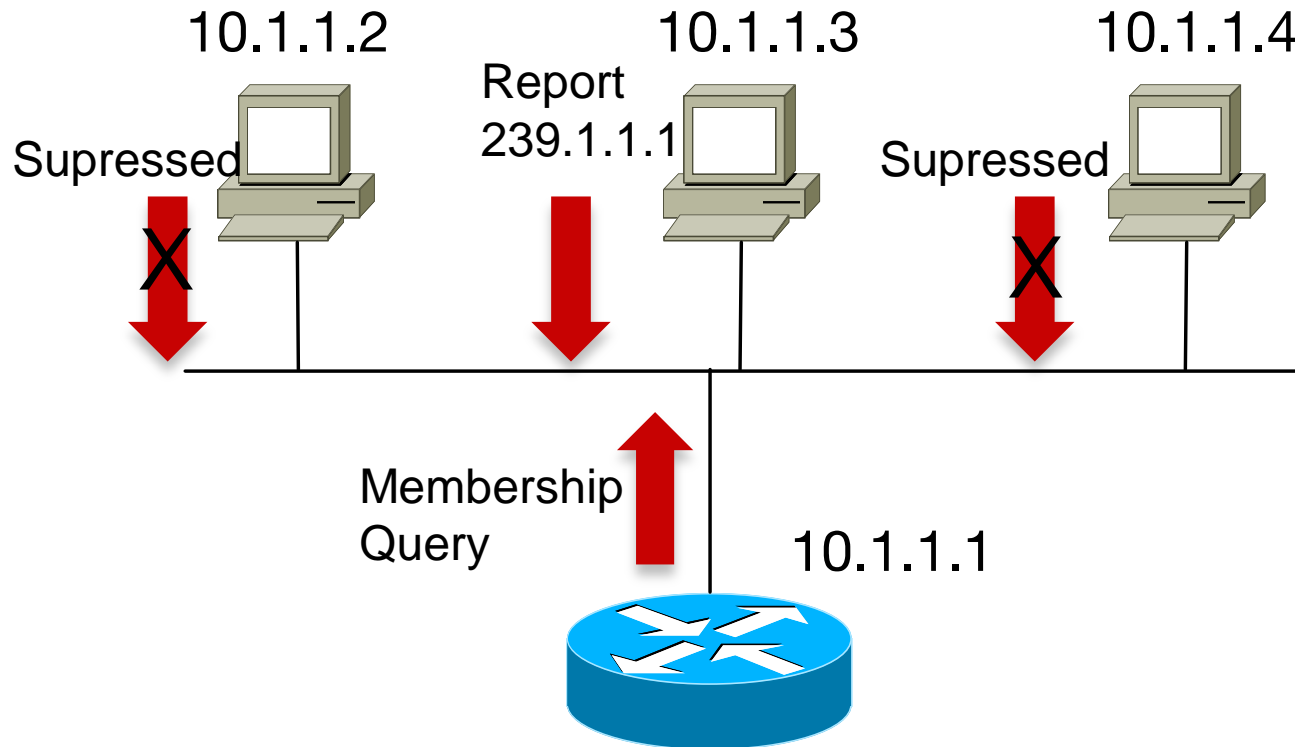
- Group-specific query  
Router sends group-specific query to make sure that there are no members present before ceasing to forward data for the group for that subnet.
- Leave group message  
Host sends leave message if it leaves the group and is the last member (reduces leave latency in comparison to v1).

# IGMP Version 2 Join Group



```
Router>show ip igmp group
IGMP Connected Group Membership
Group Address  Interface  Uptime  Expires  Last Reporter
239.1.1.1     Ethernet0  0d0h3m  00:01:40  10.1.1.3
```

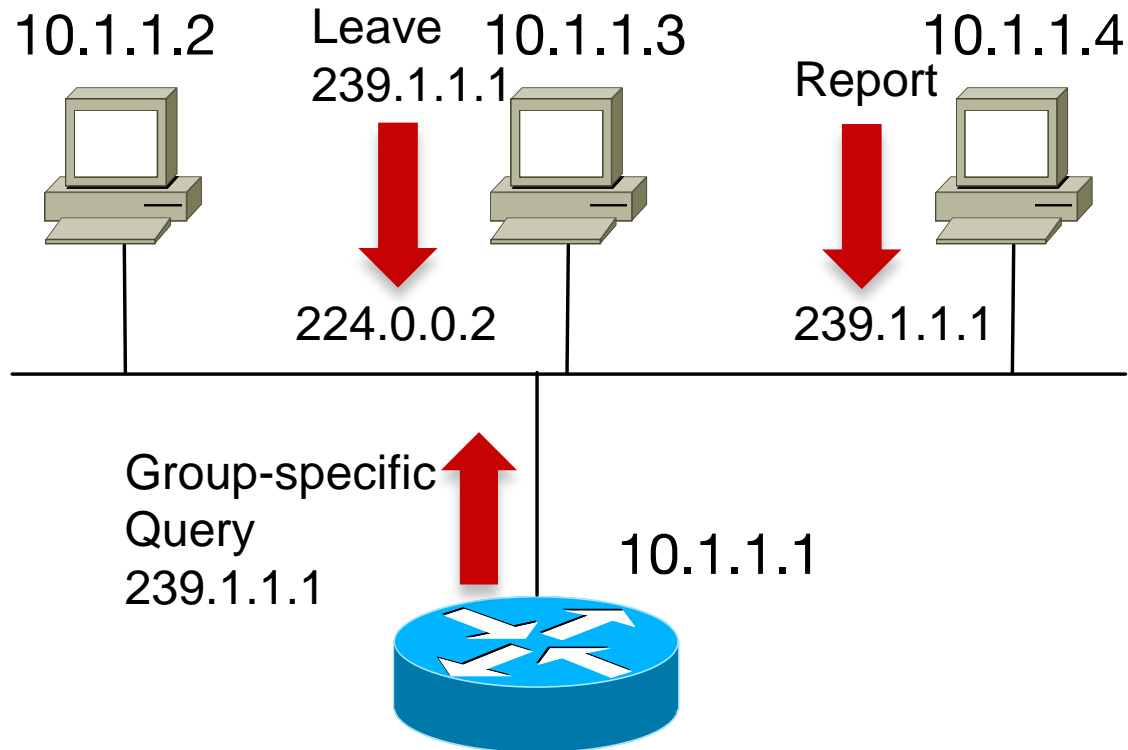
# IGMP Version 2 Maintaining a Group



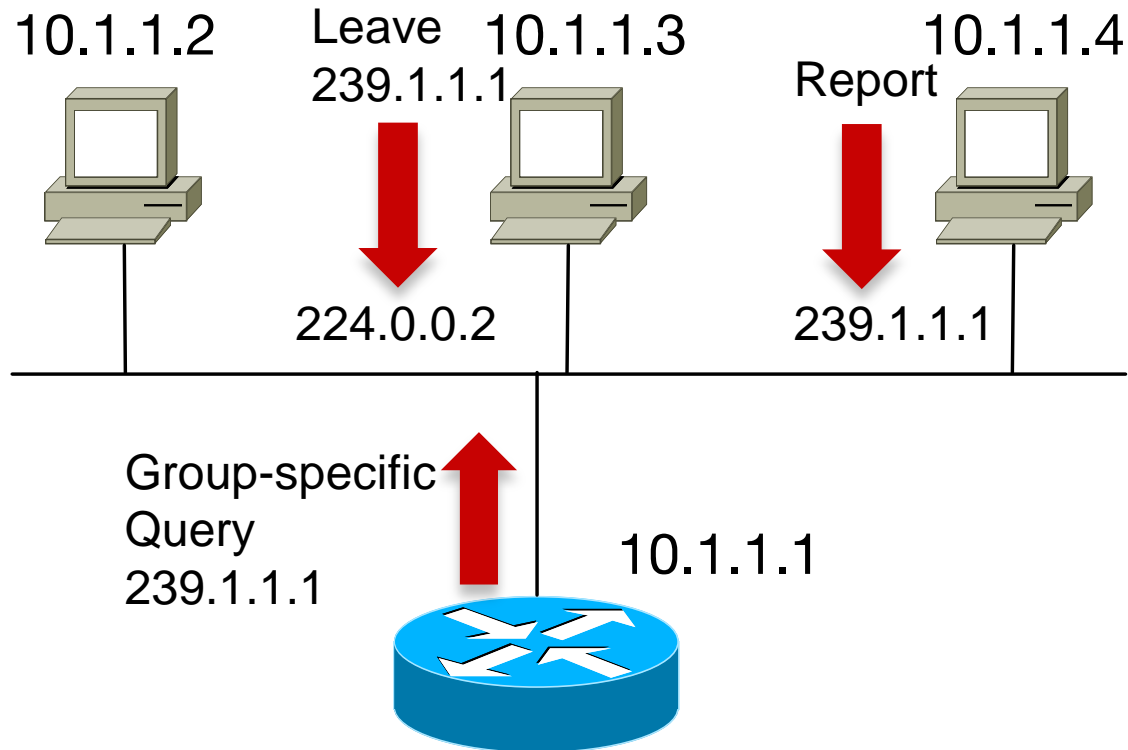
Report Suppression Mechanism

Only one member per group responds with a report to a query

# IGMP Version 2 Leaving Group

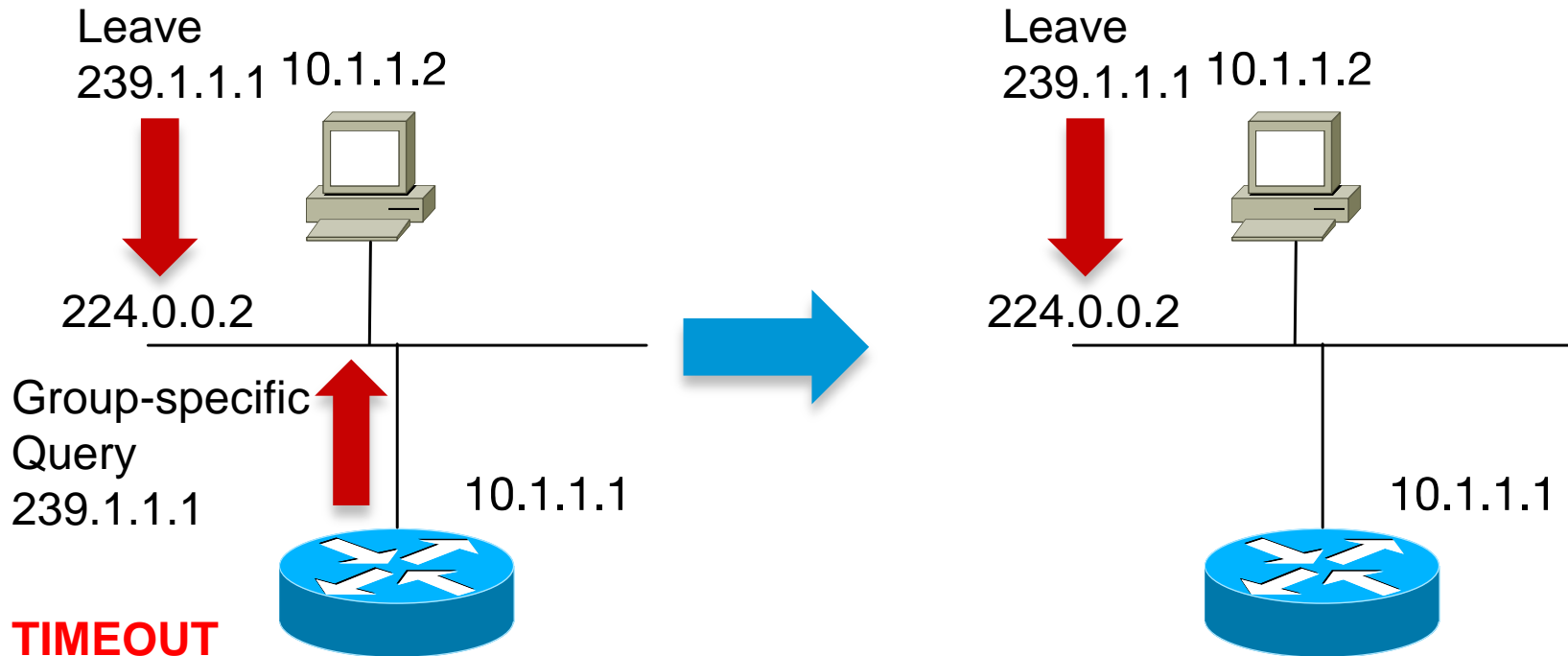


# IGMP Version 2 Leaving Group



# IGMP Version 2 Immediate Leave

LAN one connected one host, still wait group-specific query time out, inefficient.



```
Router>show ip igmp group
IGMP Connected Group Membership
Group Address  Interface  Uptime  Expires  Last Reporter
```

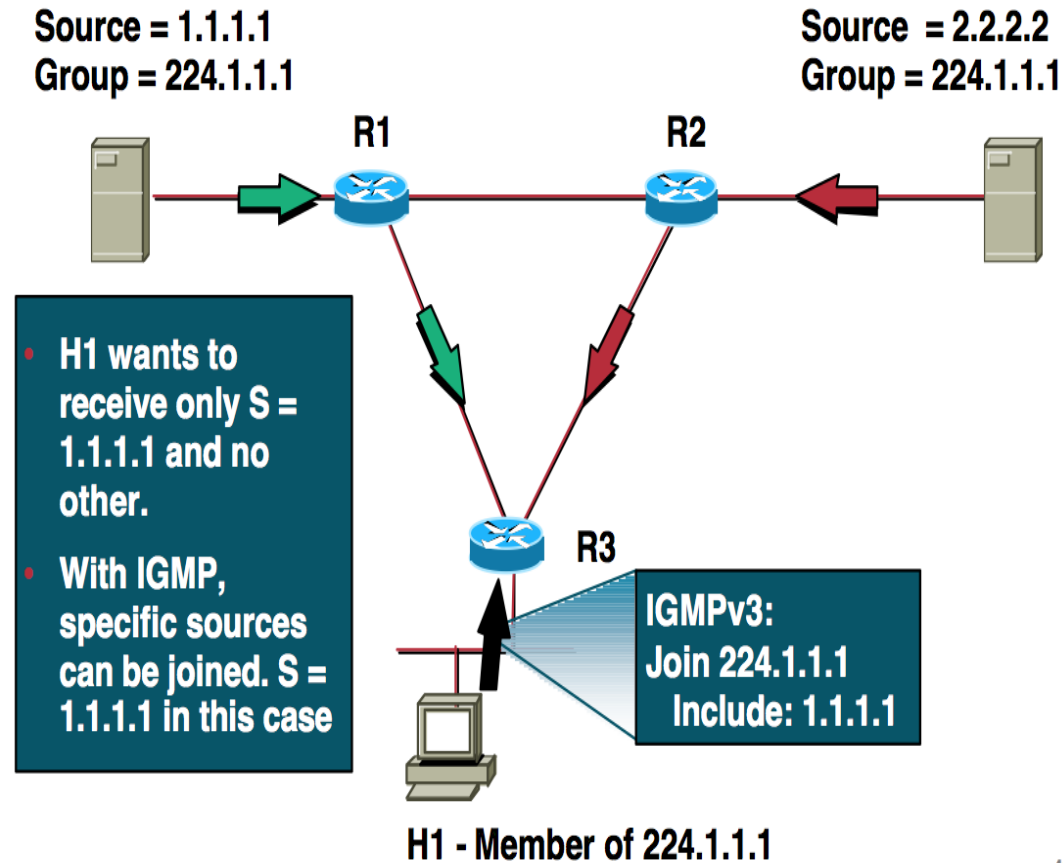
```
Router>show ip igmp group
IGMP Connected Group Membership
Group Address  Interface  Uptime  Expires  Last Reporter
```

# IGMP Version 3

- Adds Include/Exclude Source Lists

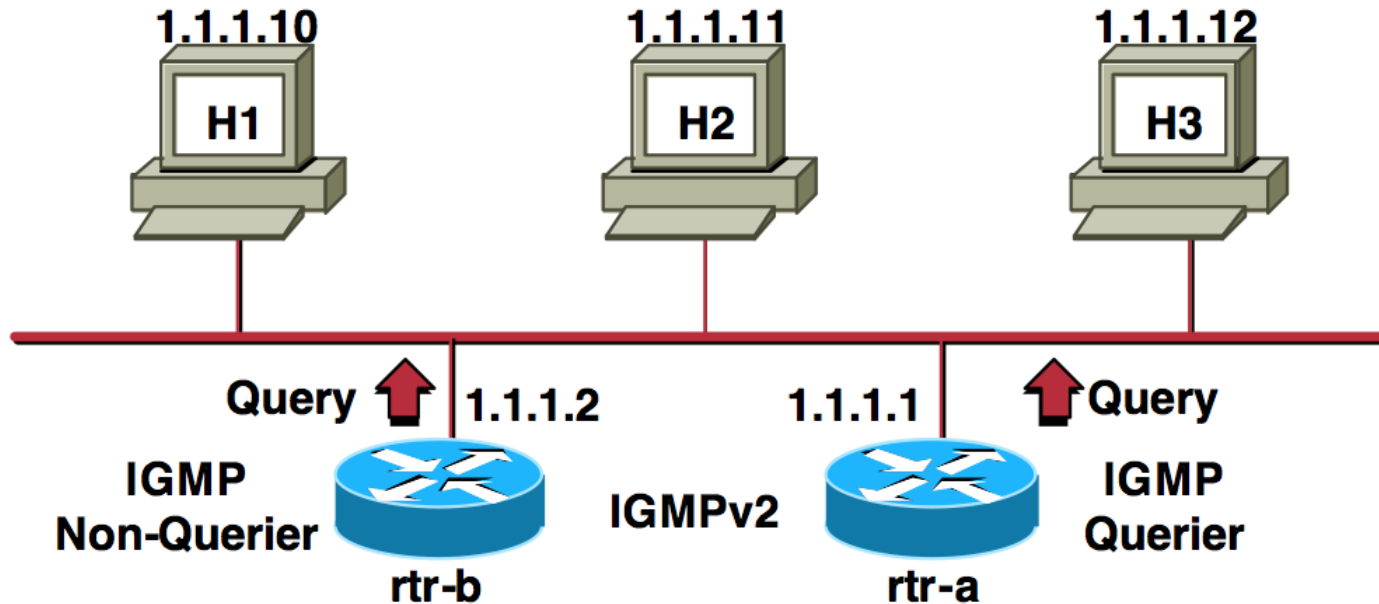
Enables hosts to listen only to a specified subset of the hosts sending to the group

- New Membership Report address  
224.0.0.22 (IGMPv3 Routers)
- No Report Suppression





# IGMP Querier

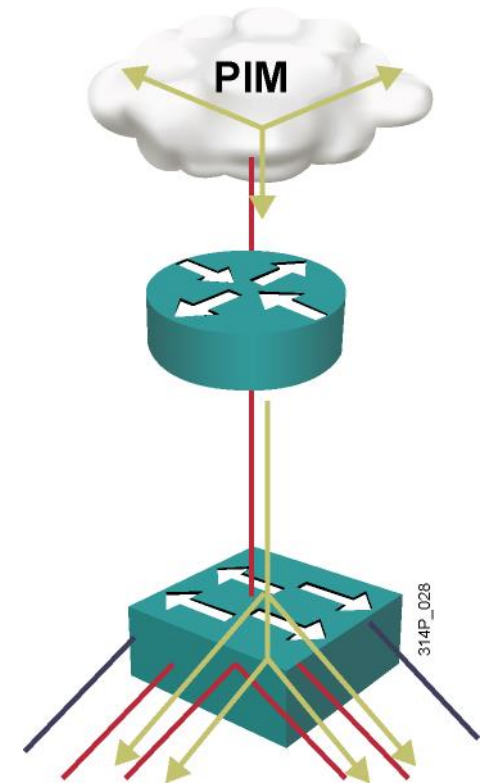


- Initially all routers send out a Query
- Router w/lowest IP address “elected” querier
- Other routers become “Non-Queriers”

# IGMP Snooping Introduction

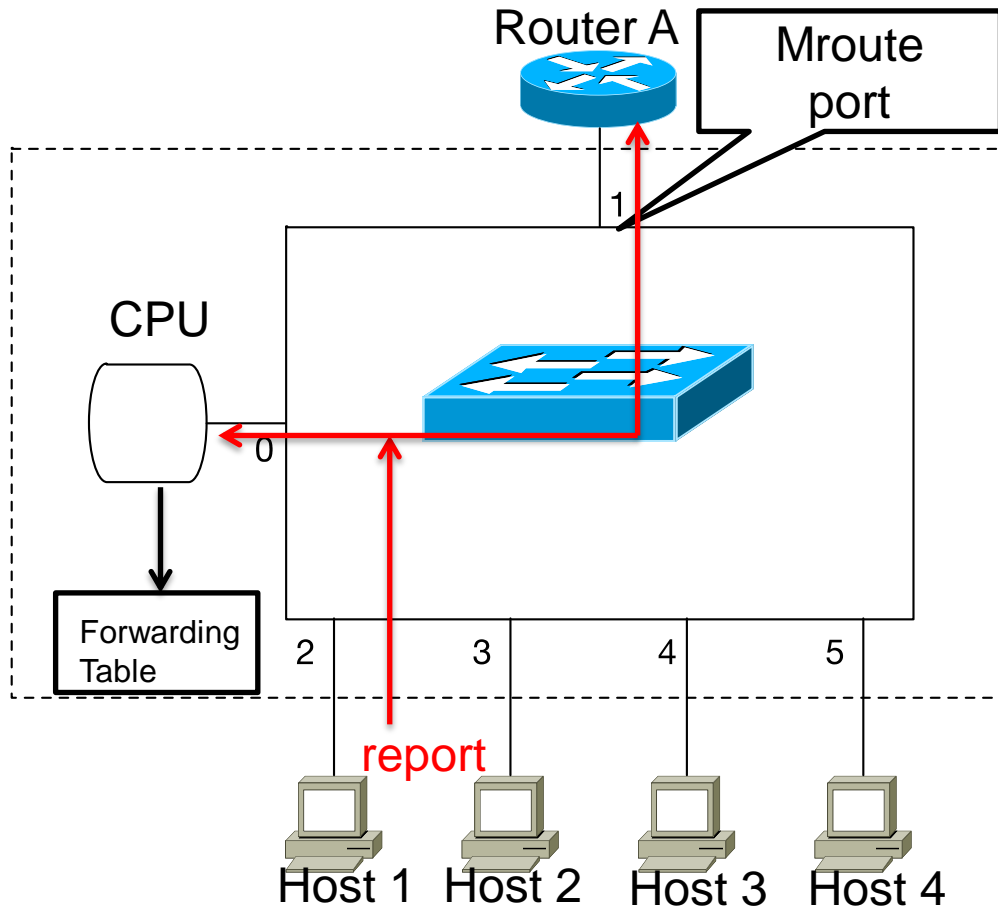
Problem: Layer 2 Flooding of Multicast Frames

- Typical L2 switches treat multicast traffic as unknown or broadcast and must “flood” the frame to every port



# IGMP Snooping Mechanism

1. Reports message only to forward to the port router attached and local CPU



```
switch#show ip igmp snooping
```

```
Vlan 1:
```

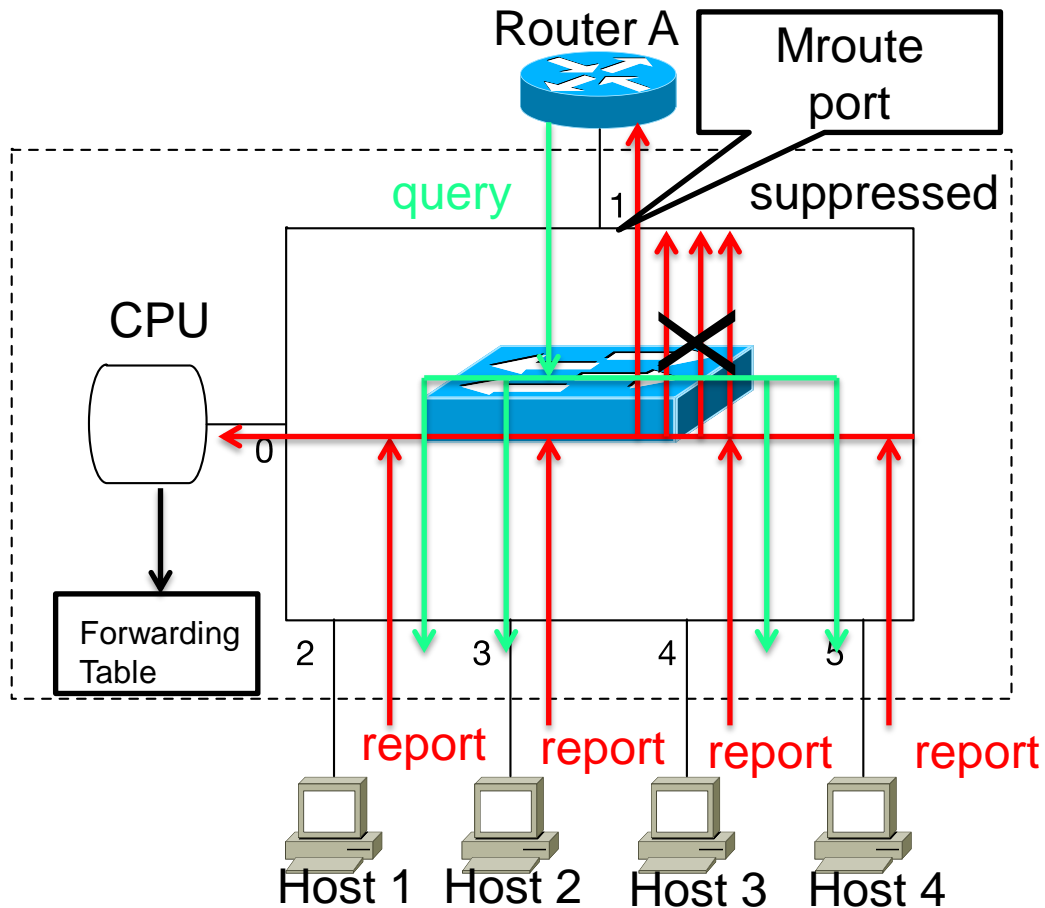
```
-----  
IGMP snooping : Enabled  
IGMPv2 immediate leave : Disabled  
Multicast router learning mode : pim-dvmrp
```

```
switch# show ip igmp snooping mrouter
```

```
Vlan ports  
---- ----  
1 1, Router
```

# IGMP Snooping Mechanism

## 2. IGMP Report suppression



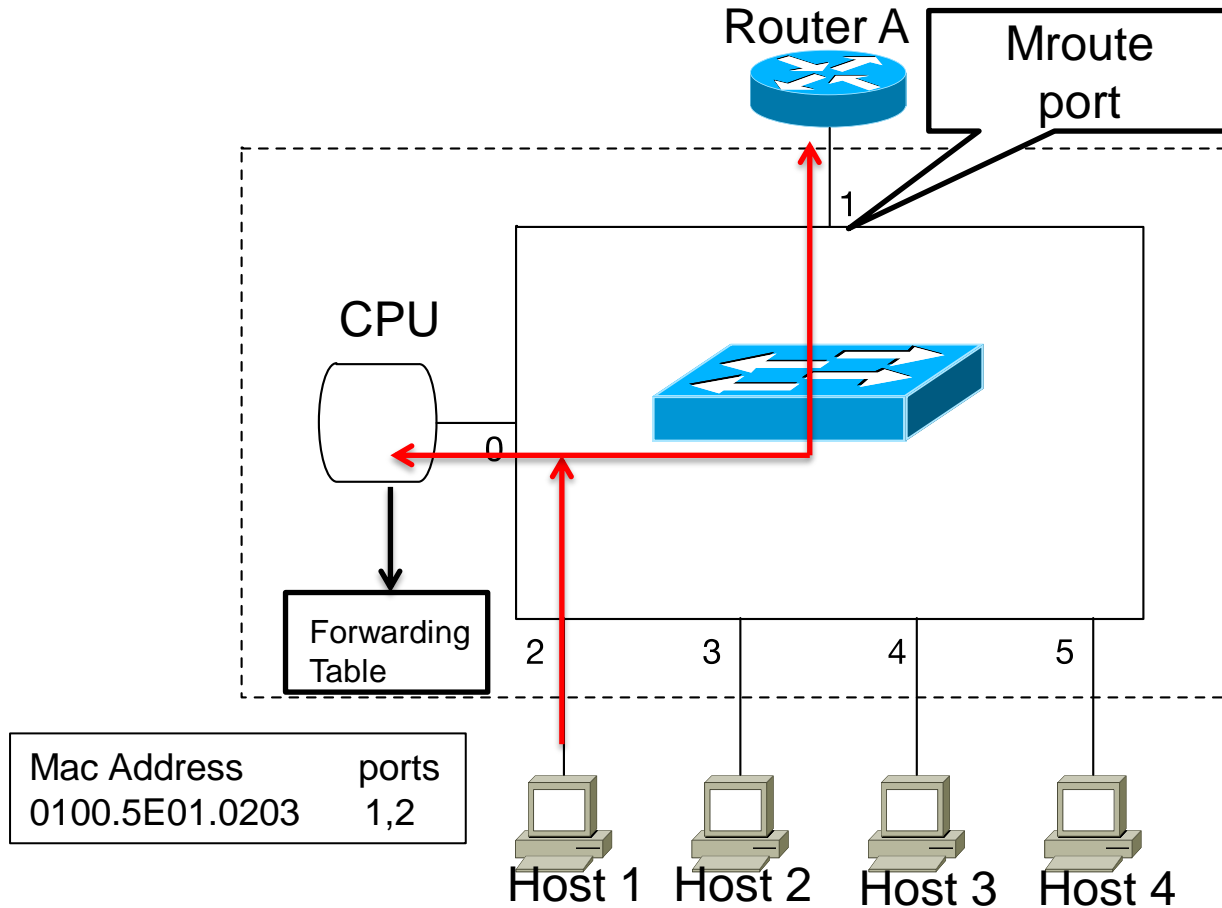
switch sends the first IGMP report from all hosts for a group to multicast routers.

Default enable

Disable command:  
**no ip igmp snooping report-suppression**

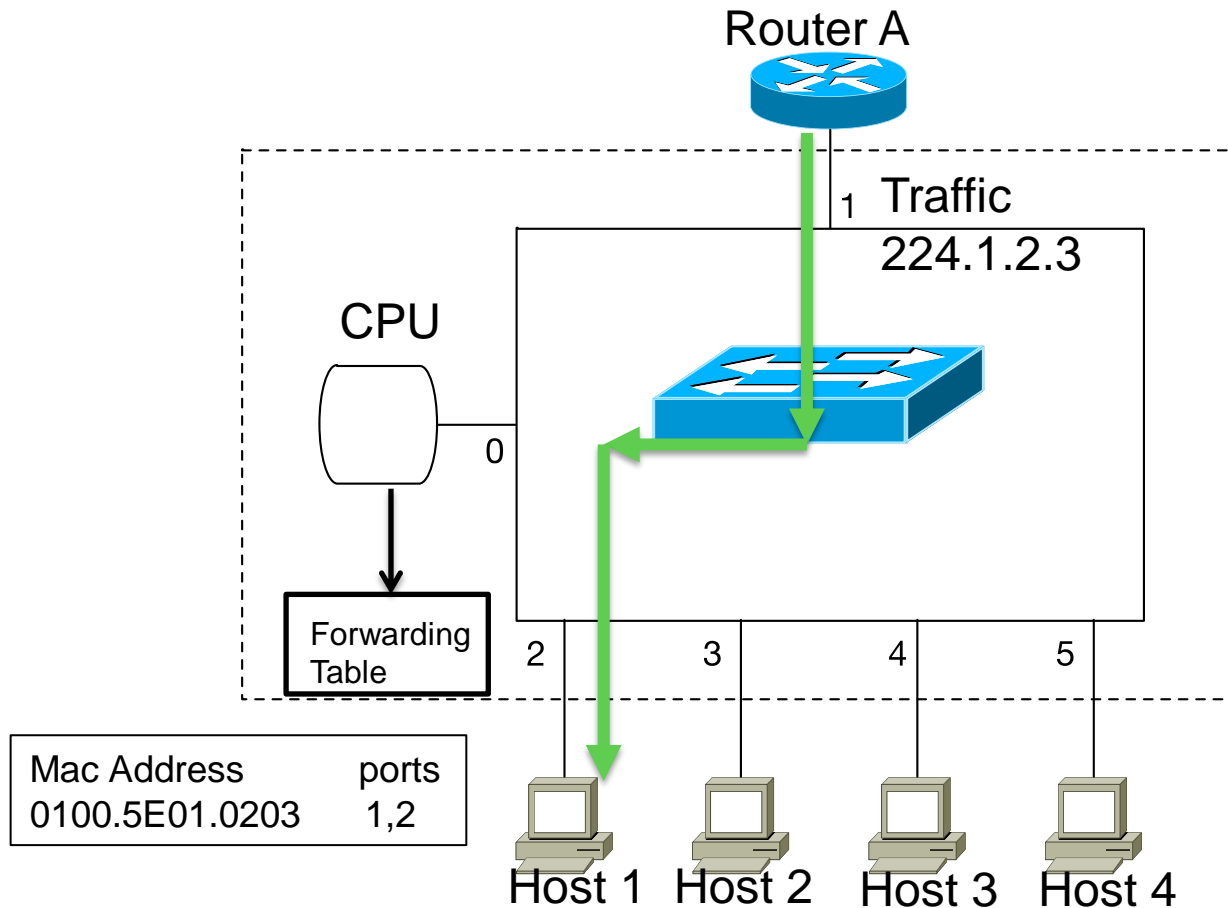
# IGMP Snooping Mechanism

MAC-addressed groups based Bridging



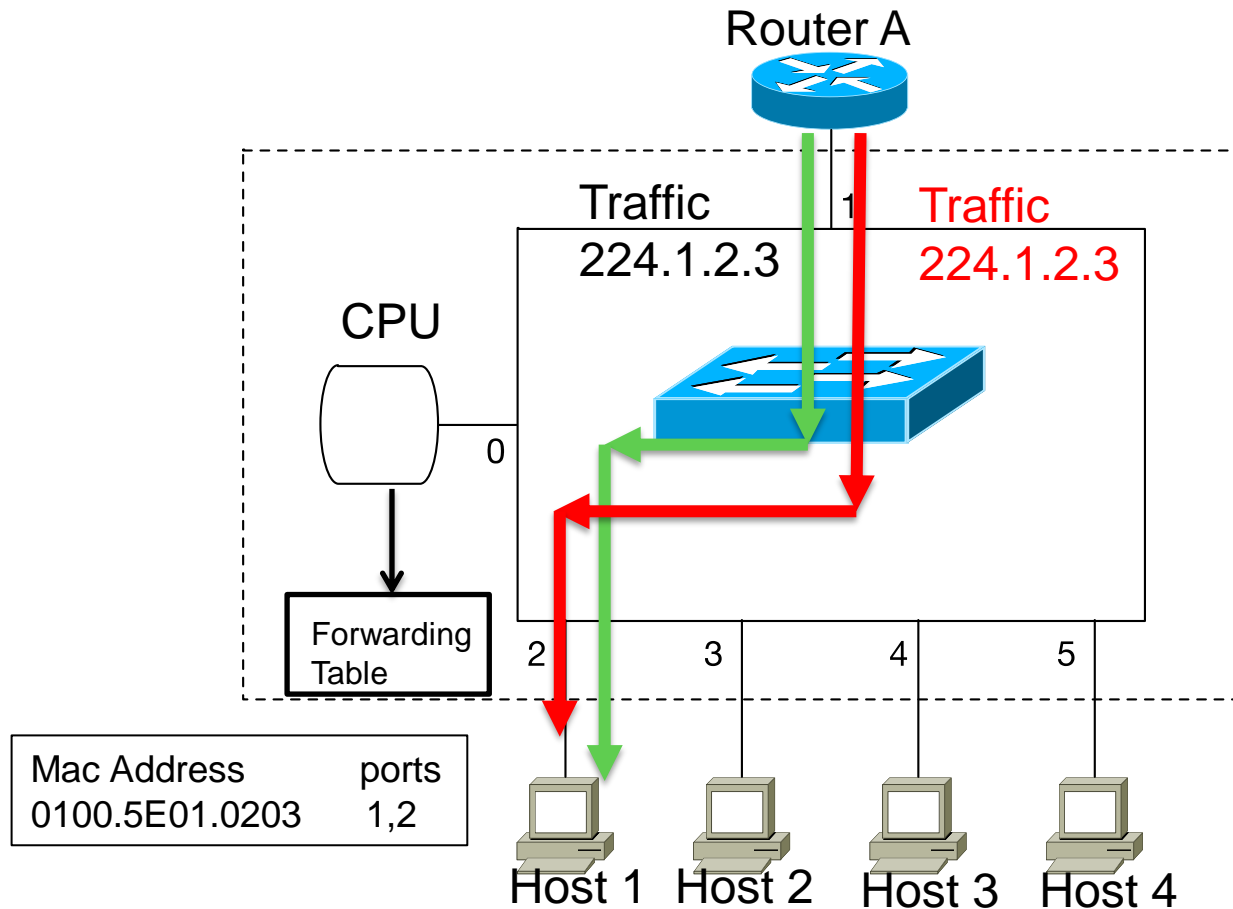
# IGMP Snooping Mechanism

MAC-addressed groups based Bridging



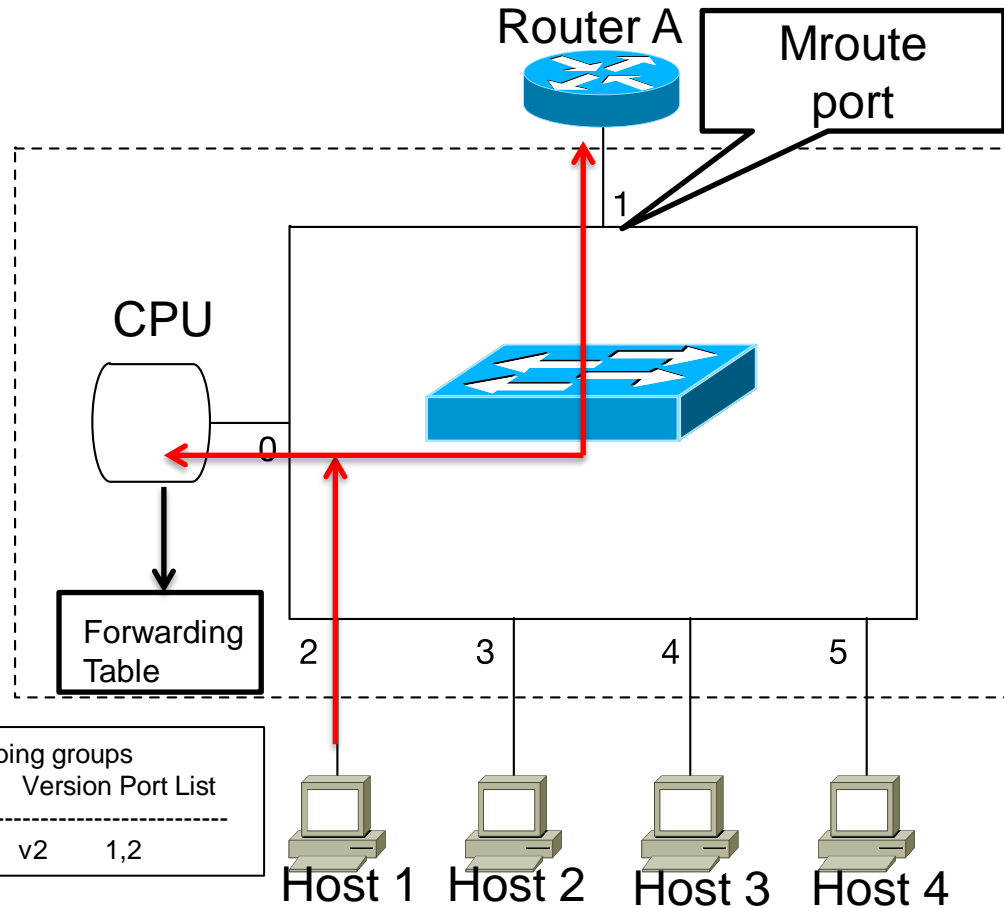
# IGMP Snooping Mechanism

If the incoming traffic is 225.1.2.3, same MAC Address of 224.1.2.3 ?



# IGMP Snooping Mechanism

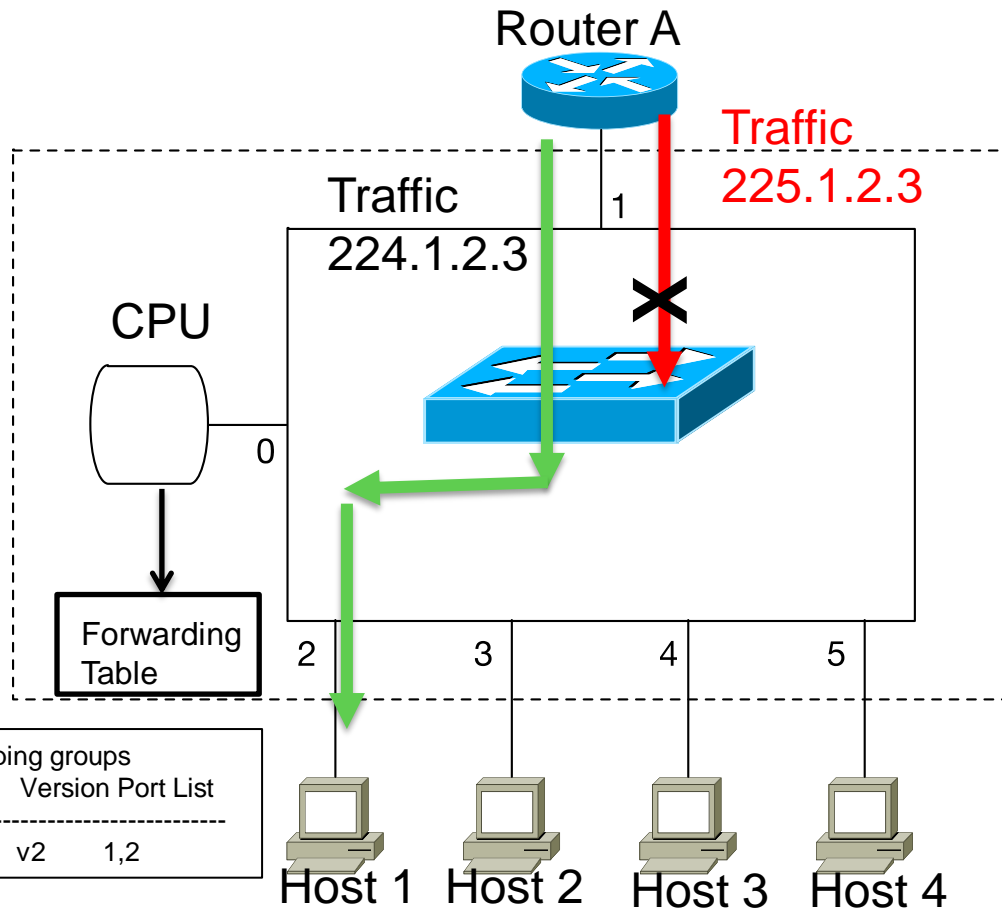
IP multicast group-based  
bridging





# IGMP Snooping Mechanism

Traffic 225.1.2.3 will be drop. No MAC address aliasing issues

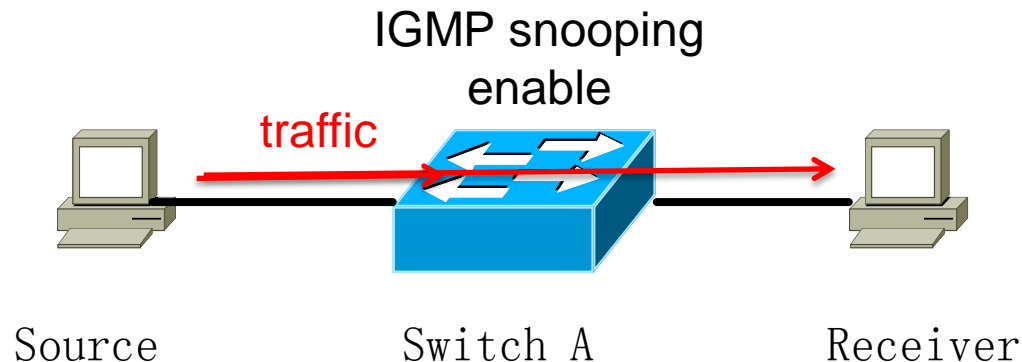


# IGMP snooping case 1

Layer2 switch connected multicast source and receiver.

Default mode , multicast traffic will be flood to all the interface included receiver port.

After Enable IGMP SNOOPING, traffic can not forward to receiver.



# IGMP snooping case 1

Receiver keep sending IGMP report message for this multicast traffic.

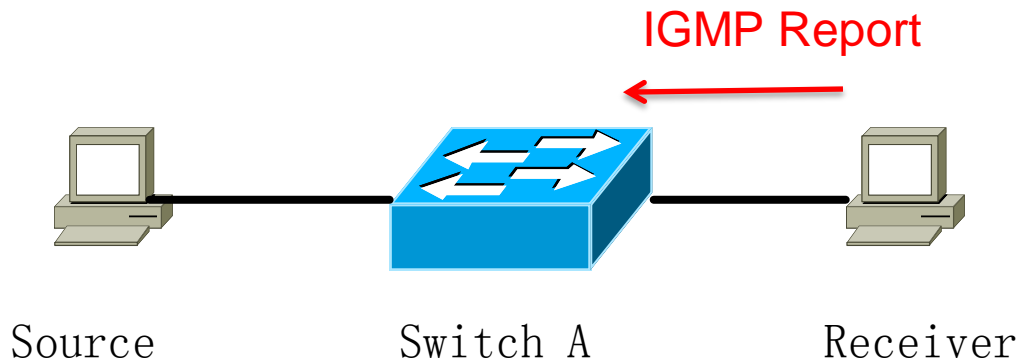
Check the igmp snooping table is NULL and mroute port table is NULL.

```
switchA# show ip igmp snooping group
```

Vlan	Group	Type	Version	Port List
-----				-----

```
switch# show ip igmp snooping mrouter
```

Vlan	ports
----	-----



# IGMP snooping case 1

Reason :

IGMP snooping requires a querier to send periodic queries and receivers can respond to indicate their interest in some groups. without the periodic queries from the querier, the IGMP snooping switch can't keep track of which ports have receivers.

Two solution:

1. Create SVI interface on switch A, enable IGMP protocol
2. Enable feature Internal Querier  
switchA(config)# ip igmp snooping querier

```
switch#show ip igmp snoop querier detail
Vlan IP Address   Version Expires   Port
1    10.1.1.1      v2      00:04:27  Vlan1 (internal)
```

IGMP Report



Source

Switch A

Receiver

# IGMP Snooping case 2

Router A

Enable IGMP  
Become IGMP Querier  
Periodically send out IGMP Query message

Switch A

Switch B

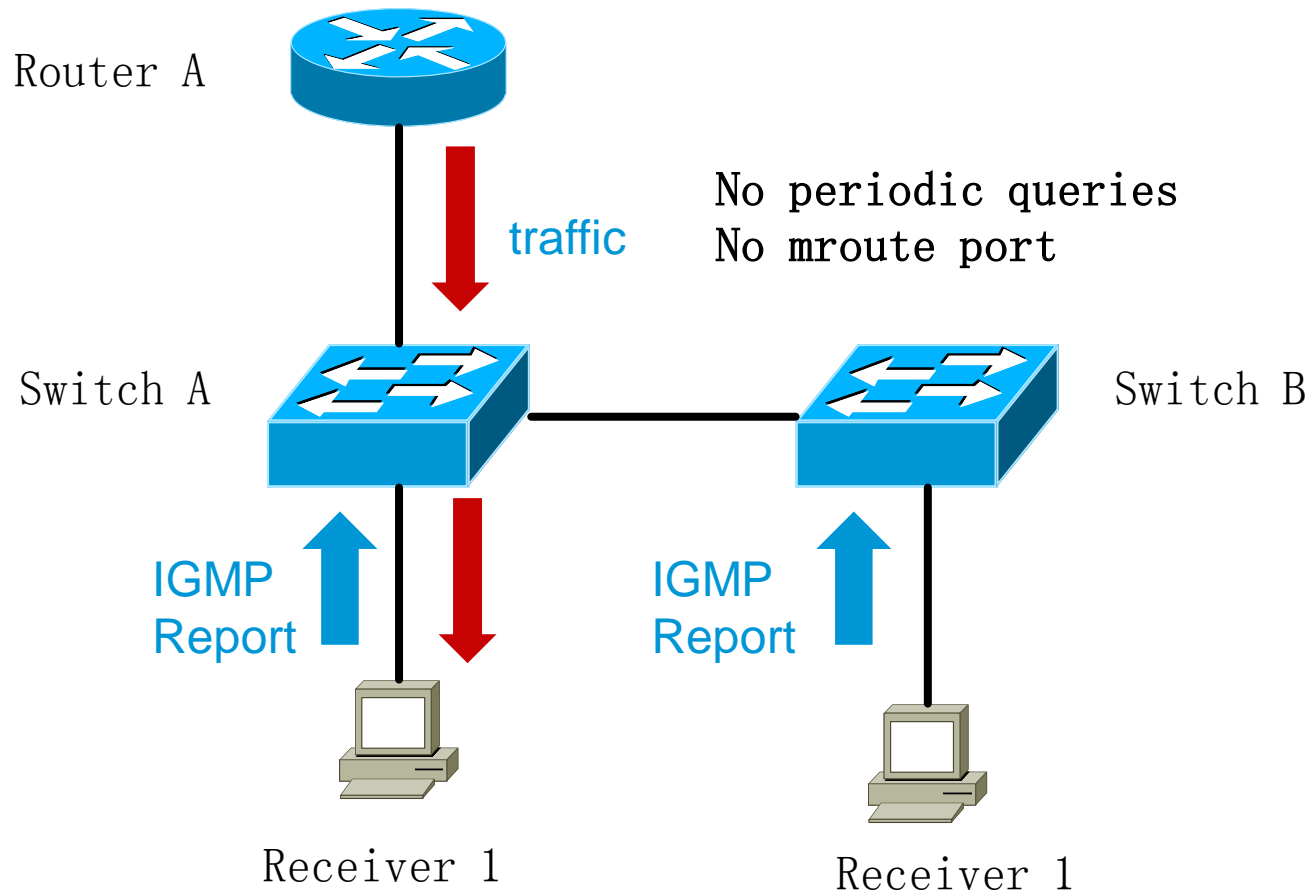
Enable IGMP Snooping

Enable IGMP Snooping

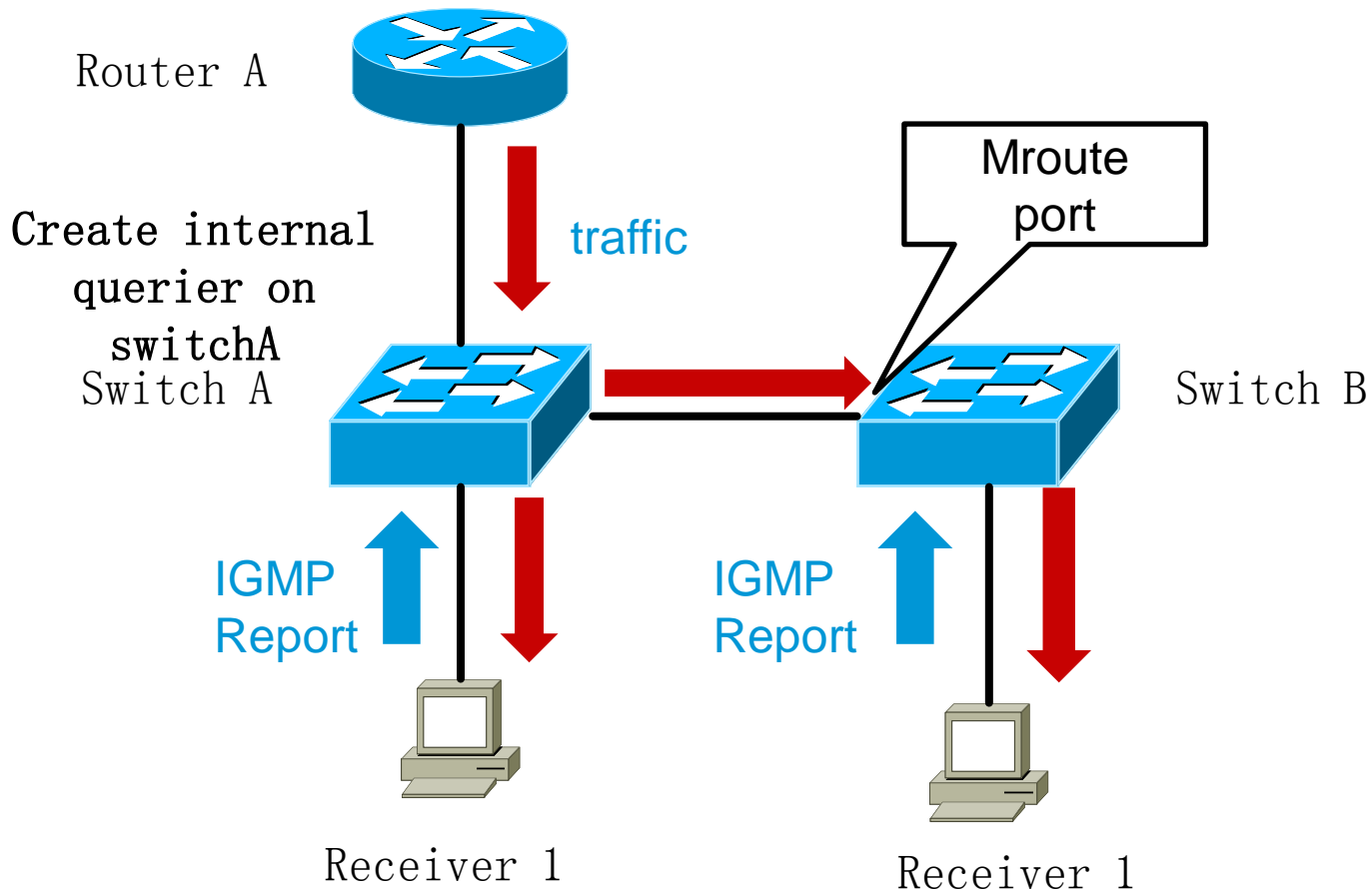
Receiver

Receiver

# IGMP Snooping case2



# IGMP Snooping case2



# IGMP Proxy Routing

## ■ IGMP snooping:

Performed by L2 switch. Intended to be transparent. Many vendor variations.

IETF RFC 4541 – INFORMATIONAL ONLY

**Transparent:** no snooping messages suppressed

**Report-suppression:** guess which IGMP reports are redundant at router (can break explicit tracking, fast leaves).

**Proxy-reporting:** fully emulate host.

IGMPv3: Use source-IP address "0.0.0.0"

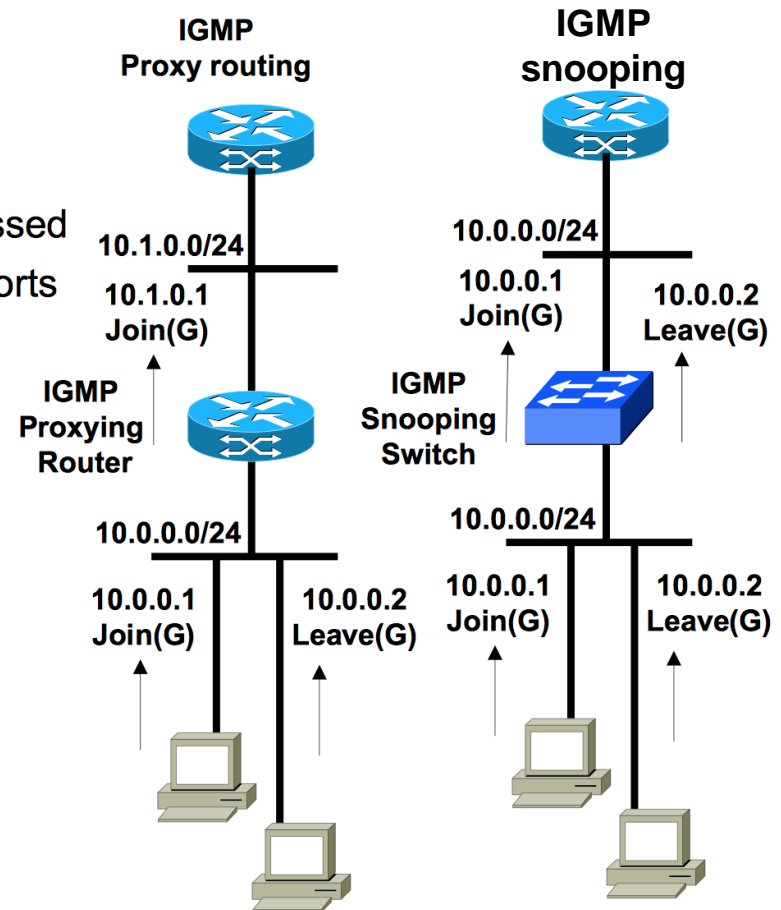
## ■ IGMP proxy-routing:

Performed by router:

IETF RFC4605 – STANDARDS TRACK

IGMP proxy router need to act exactly like a single host on it's upstream interface.

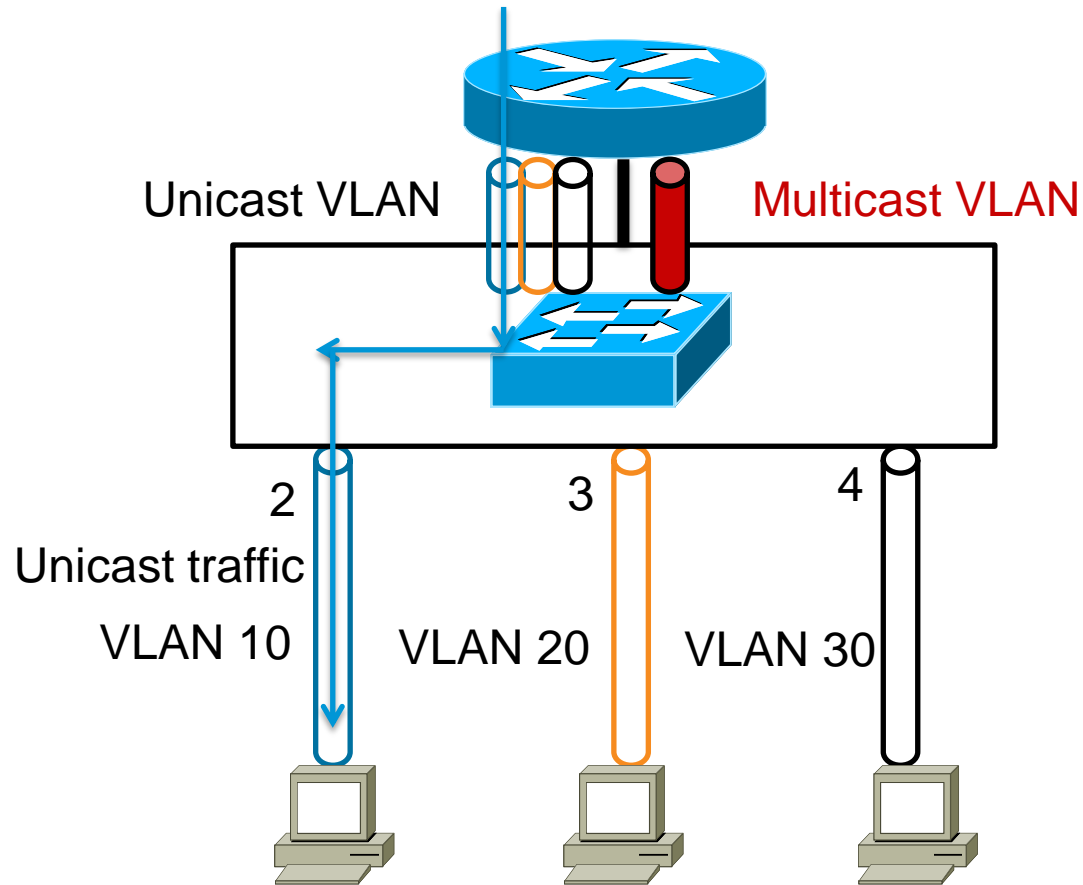
*Router can not transparently pass through IGMP membership packets from downstream hosts: would have incorrect source-IP addresses.*



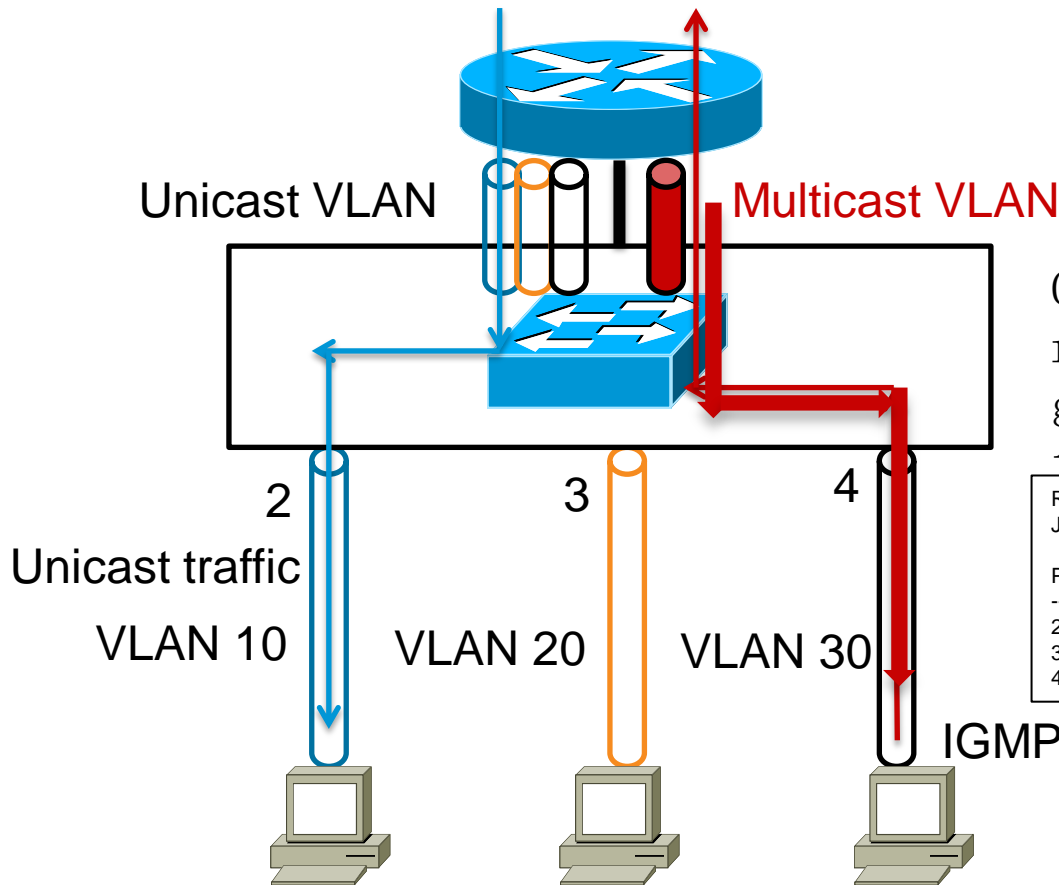


# Multicast VLAN Registration

Subscriber can subscribe and unsubscribe to a multicast stream on the network-wide multicast VLAN.



# Multicast VLAN Registration



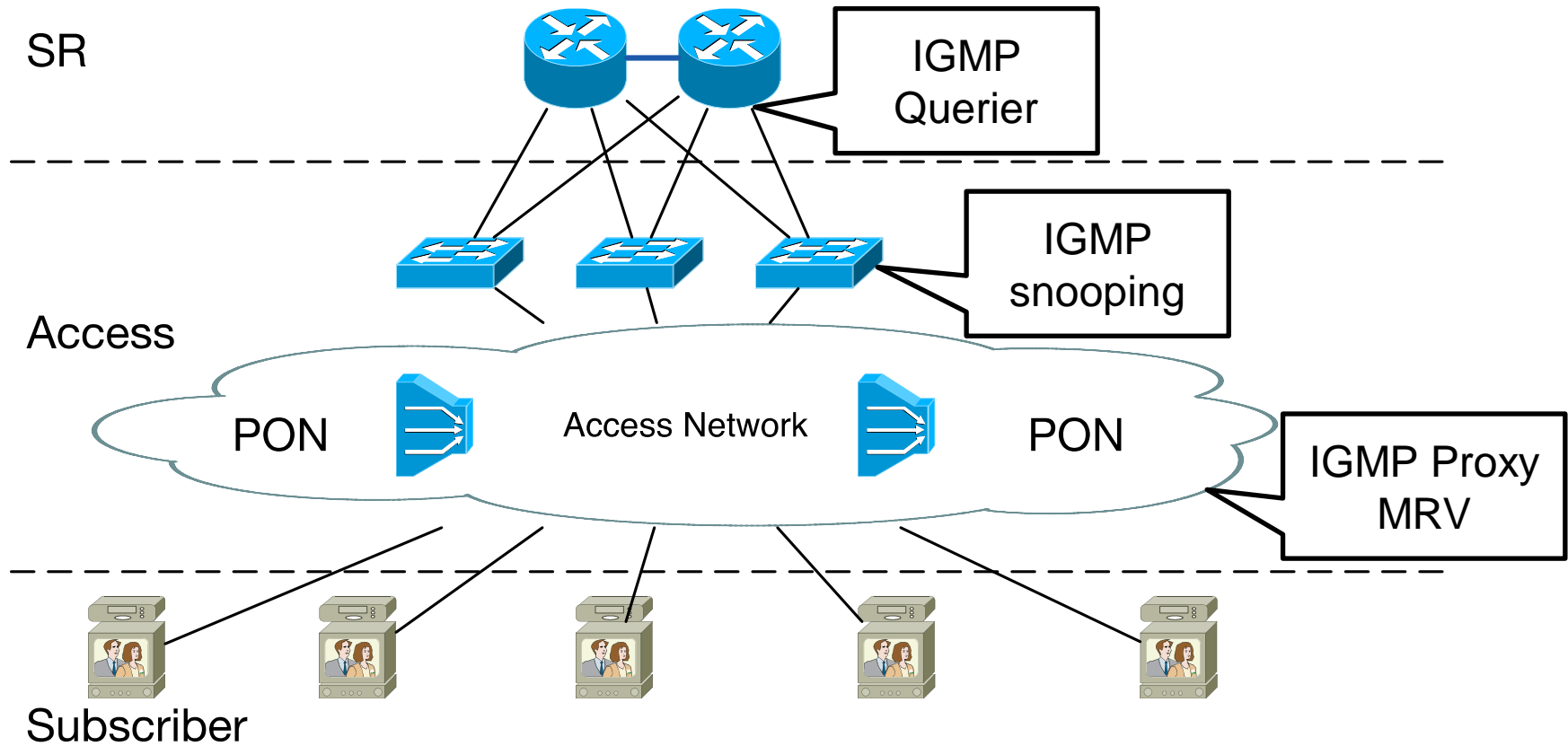
Capture by switch and register multicast group/VLAN/Port info to MVR table

```
Router# show mvr receiver-ports
Joins: v1,v2,v3 counter shows total IGMP joins
       v3 counter shows IGMP joins received with both MVR and non-MVR groups
```

Port	VLAN	Status	Immediate Leave	Joins (v1,v2,v3)	(v3)
2	10	INACTIVE/UP	ENABLED	305336	0
3	20	INACTIVE/UP	DISABLED	4005	0
4	30	ACTIVE/UP	DISABLED	53007	0

IGMP Report

# Layer2 Multicast feature on SP



# Layer2 Multicast Troubleshooting CLI

```
Router#show ip igmp group
IGMP Connected Group Membership
Group Address    Interface    Uptime    Expires    Last Reporter
224.1.1.1        Ethernet1    3d16h     00:01:59   172.16.7.2
224.0.1.40       Ethernet0    4d15h     never      172.16.6.2
```

# Layer2 Multicast Troubleshooting CLI

```
router#show ip igmp interface
FastEthernet1/0 is up, line protocol is up
Internet address is 10.1.4.1/24
IGMP is enabled on interface
Current IGMP host version is 2
Current IGMP router version is 2
IGMP query interval is 60 seconds
IGMP querier timeout is 120 seconds
IGMP max query response time is 10 seconds
Last member query count is 2
Last member query response interval is 1000 ms
Inbound IGMP access group is not set
IGMP activity: 0 joins, 0 leaves
Multicast routing is enabled on interface
Multicast TTL threshold is 0
Multicast designated router (DR) is 10.1.4.2
IGMP querying router is 10.1.4.1 (this system)
No multicast groups joined by this system
```

# Layer2 Multicast Troubleshooting CLI

```
switch# show ip igmp snooping
Global IGMP Snooping configuration:
-----
IGMP snooping           : Enabled
IGMPv3 snooping (minimal) : Enabled
Report suppression      : Enabled
TCN solicit query       : Disabled
TCN flood query count   : 2
Robustness variable     : 2
Last member query count : 2
Last member query interval : 1000

Vlan 1:
-----
IGMP snooping           : Enabled
IGMPv2 immediate leave  : Disabled
Multicast router learning mode : pim-dvmrp
CGMP interoperability mode : IGMP_ONLY
Robustness variable     : 2
Last member query count : 2
Last member query interval : 1000
```

# Layer2 Multicast Troubleshooting CLI

```
switch# show ip igmp snooping group
```

Vlan	Group	Type	Version	Port List
40	224.1.1.1	igmp	v2	Gi1/0/5, Gi2/0/10, Gi2/0/12 40 224.
1.1.2	igmp	v2		Gi1/0/5, Gi2/0/10, Gi2/0/12

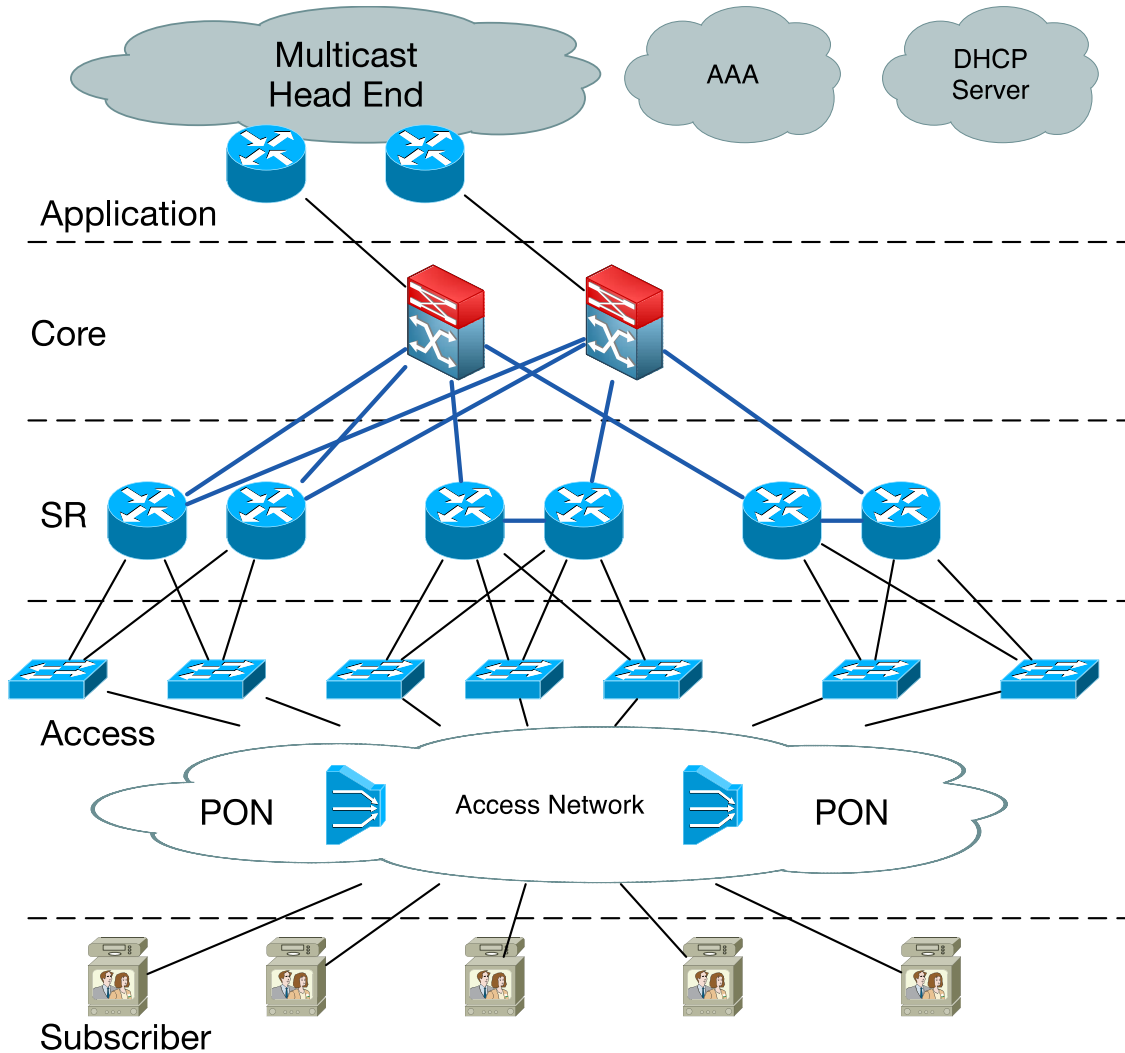
```
switch# show ip igmp snooping mrouter
```

Vlan	ports
16	Gi2/0/12(dynamic), Router
40	Gi2/0/12(dynamic), Router

```
switch#show ip igmp snoop querier detail
```

Vlan	IP Address	Version	Expires	Port
1	10.1.1.1	v2	00:04:27	Vlan1 (internal)

# Layer3 Multicast Technology



PIM SM



# Multicast Protocols

- PIM : Dense-mode

  - Uses “Push” Model

  - Traffic Flooded throughout network

  - Pruned back where it is unwanted

  - Flood & Prune behavior (typically every 3 minutes)

  - Not supported in IOS-XR

- PIM : Sparse-mode

  - Uses “Pull” Model

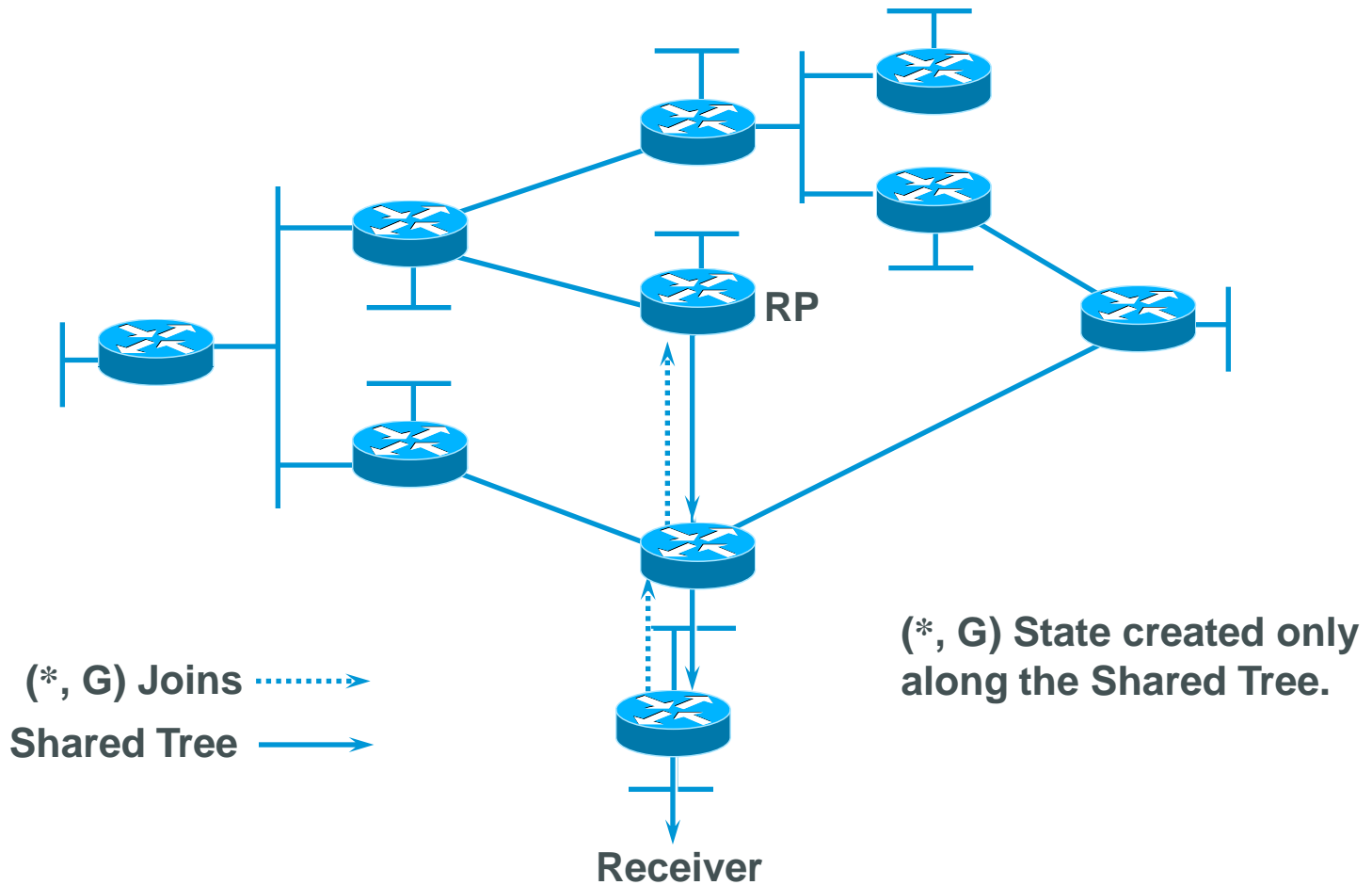
  - Traffic sent only to where it is requested

  - Explicit Join behavior

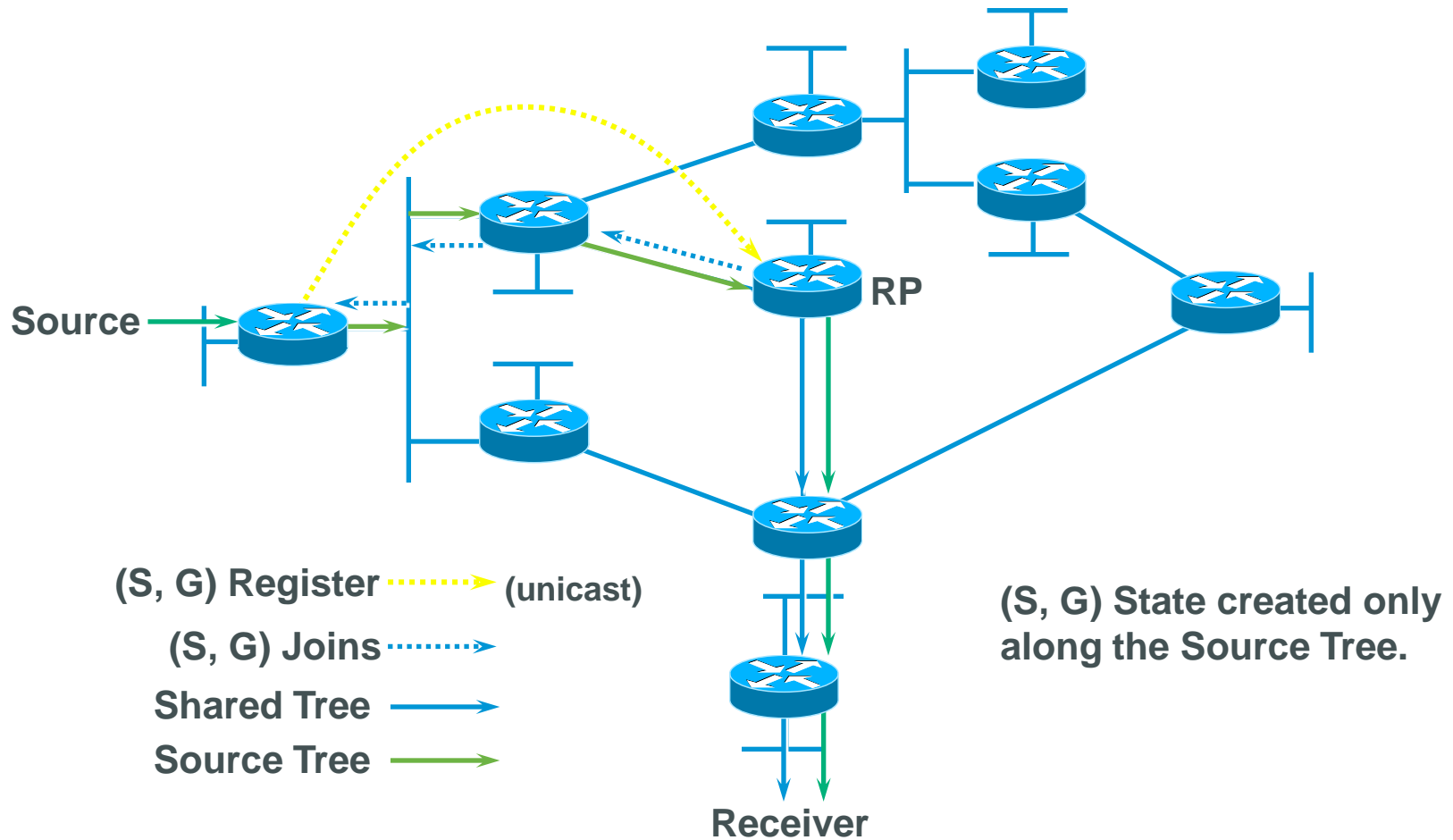
# PIM-SM (RFC 2362)

- Supports both source and shared trees
  - Assumes no hosts want multicast traffic unless they specifically ask for it
- Uses a Rendezvous Point (RP)
- Senders and Receivers “rendezvous” at this point to learn of each others existence.
- Only one RP is chosen for a particular group
- RP statically configured or dynamically learned

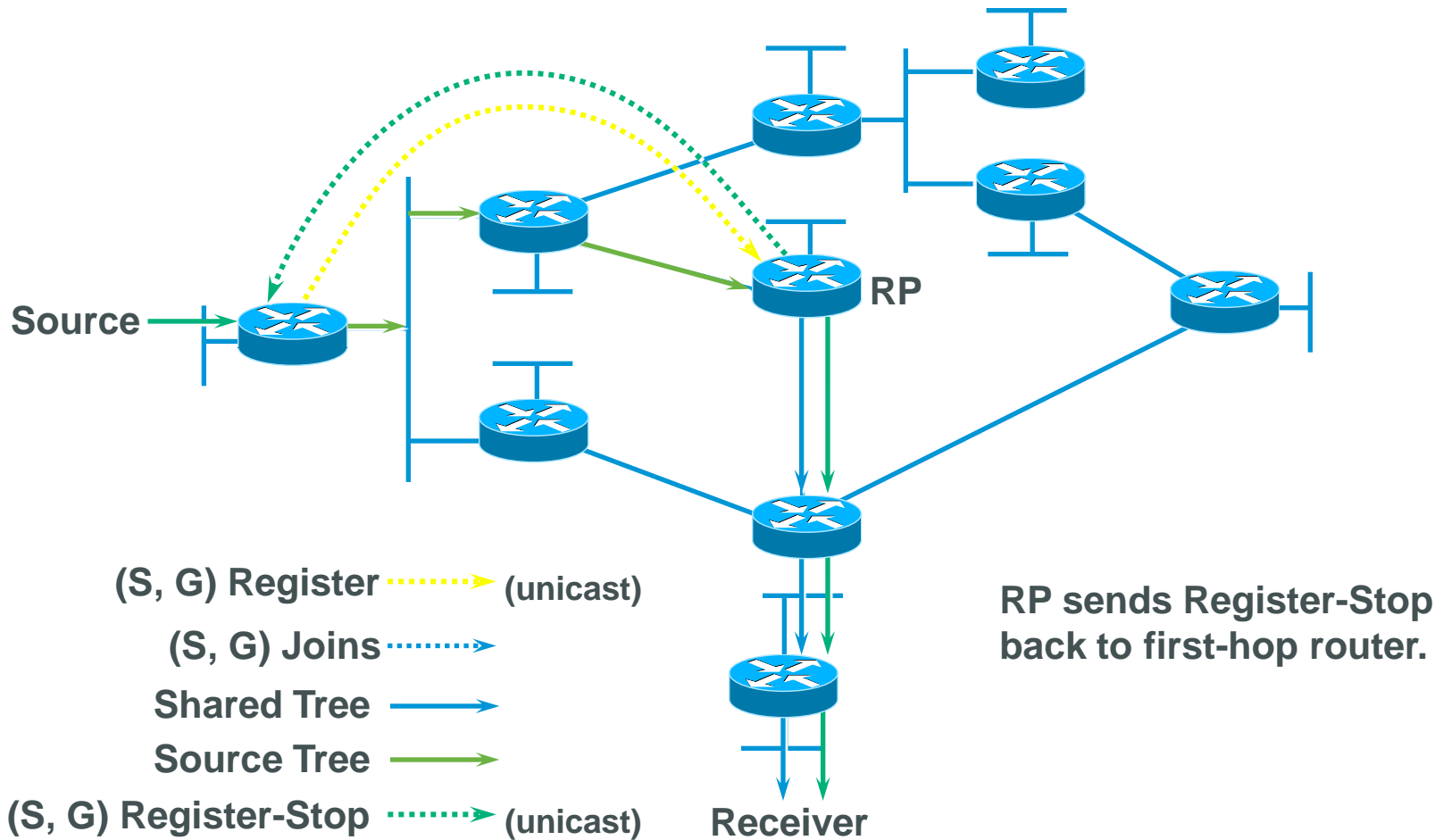
# PIM-SM Shared Tree Joins



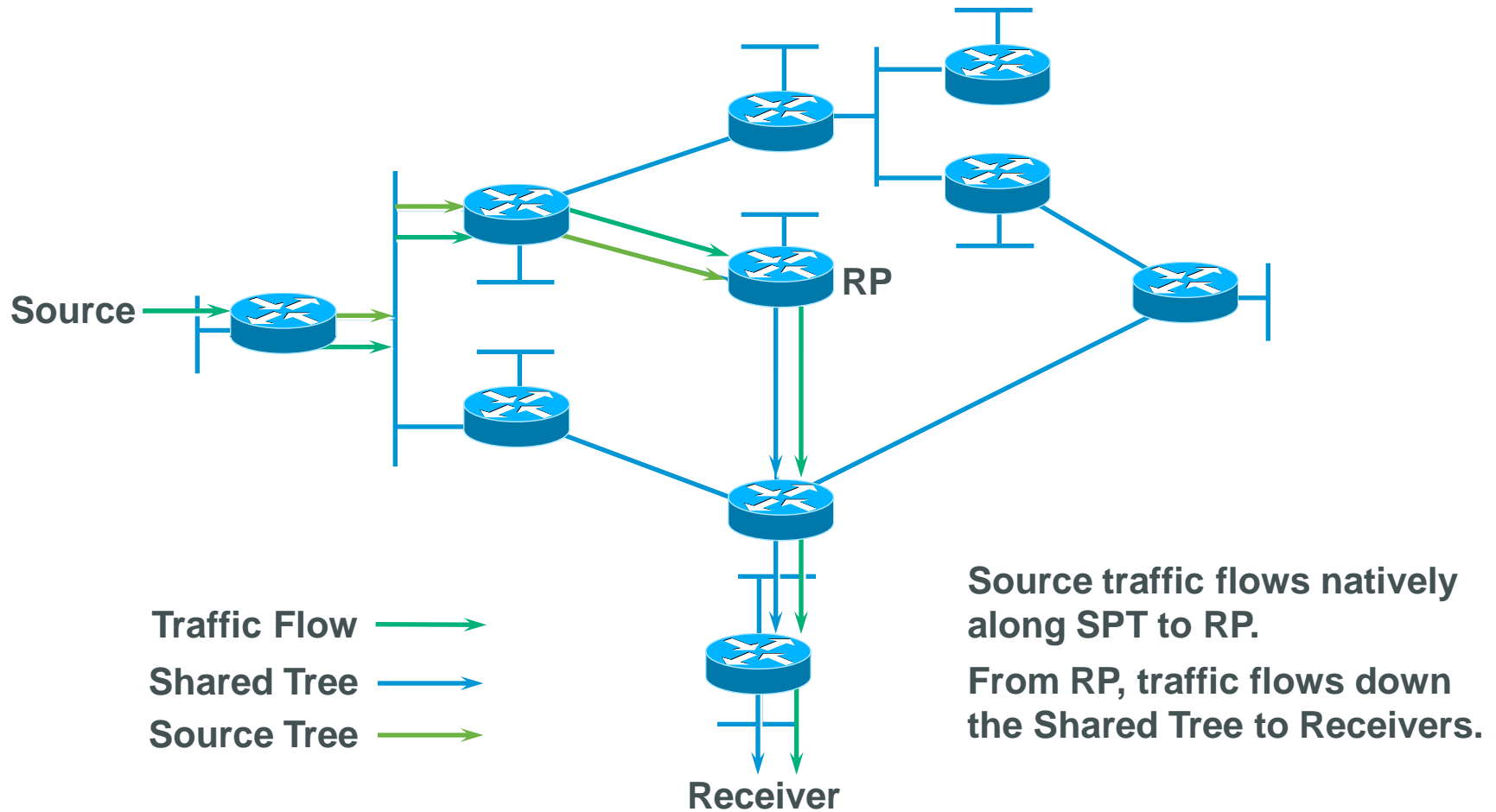
# PIM-SM Sender Registration



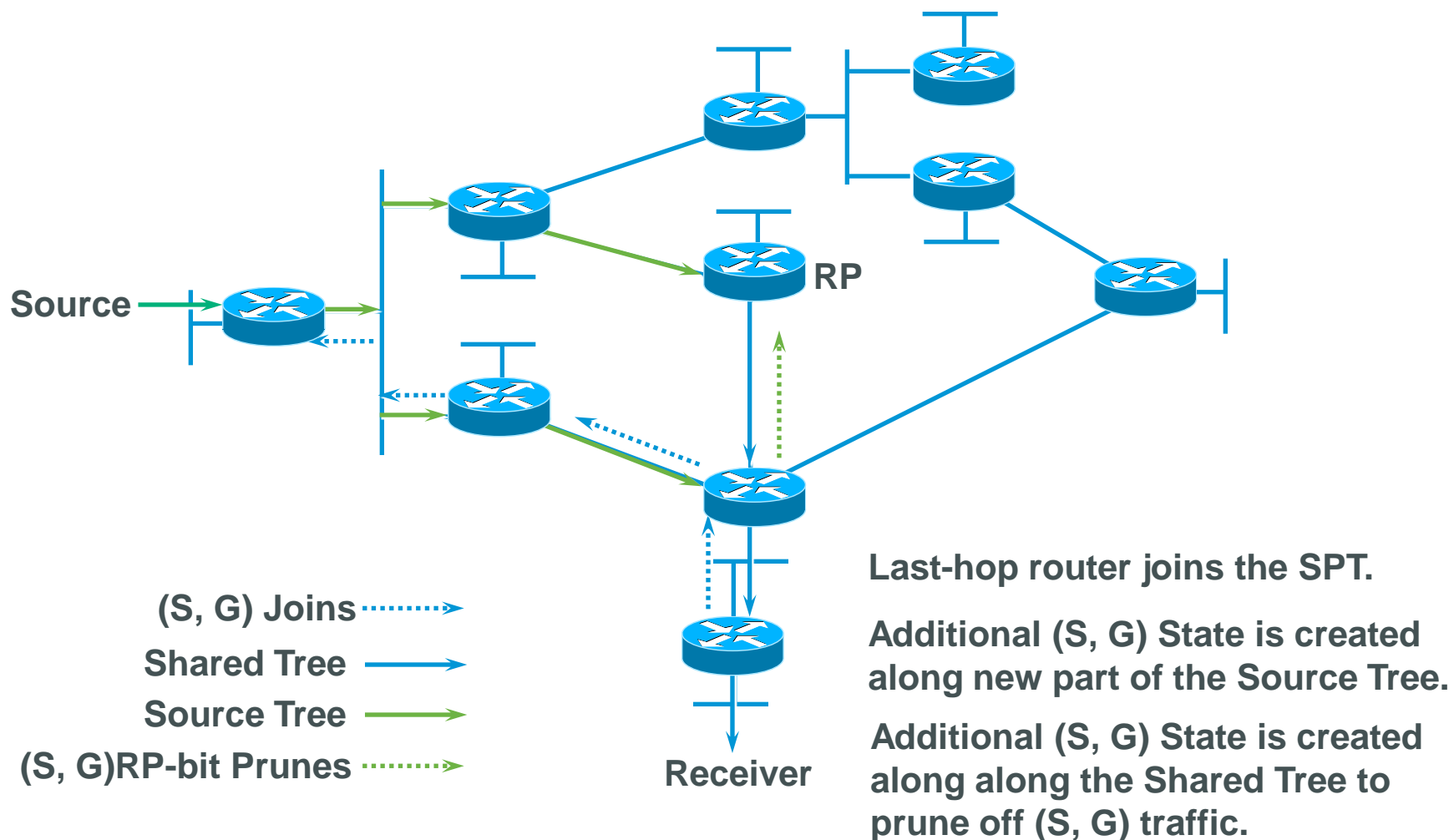
# PIM-SM Reverse SPT



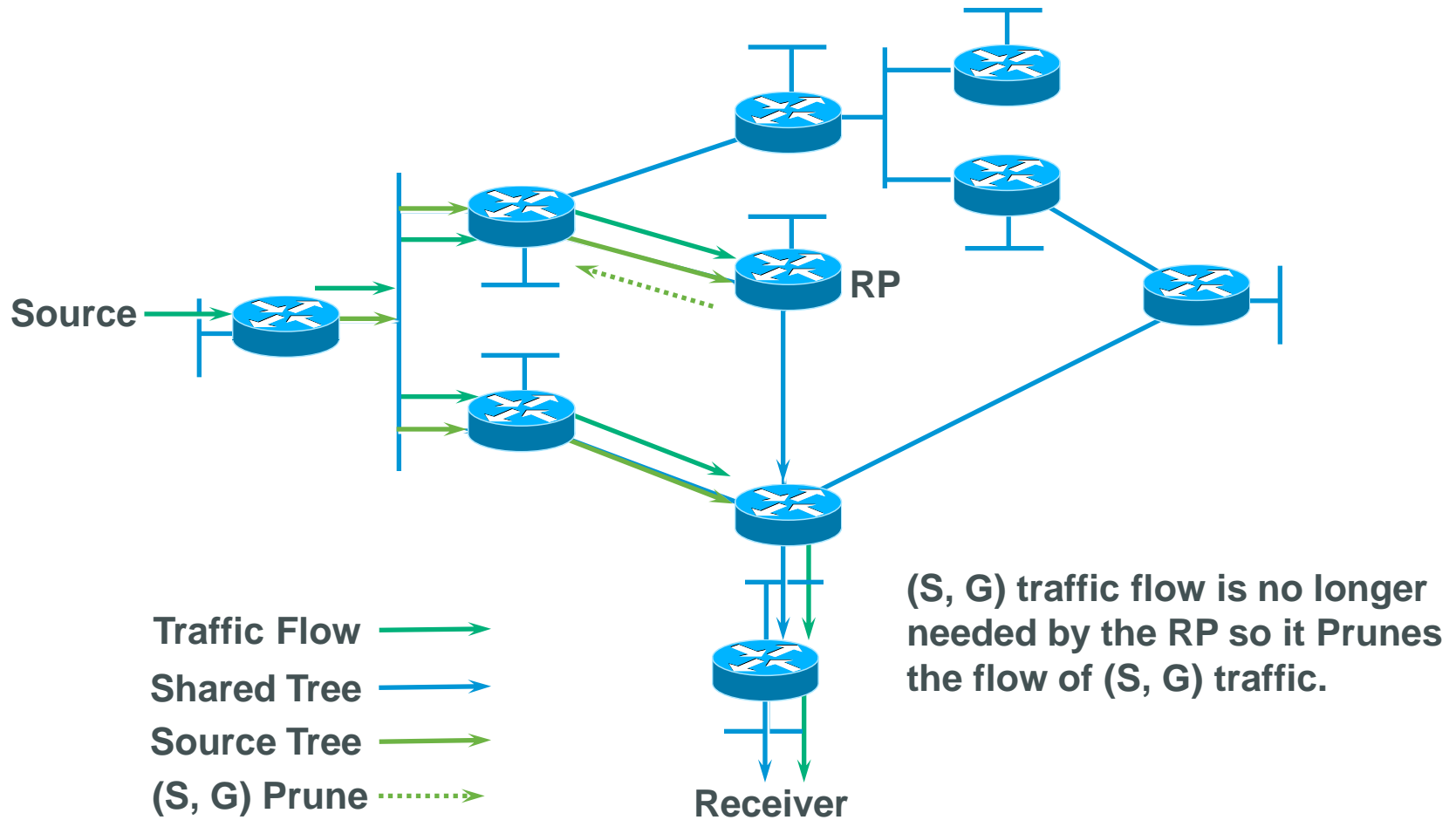
# PIM-SM Traffic Flow



# PIM-SM SPT Switchover

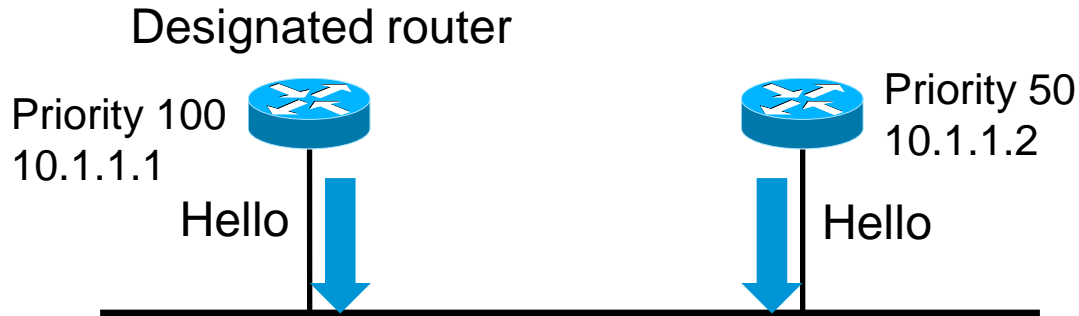


# PIM-SM SPT Switchover





# PIM Neighbor Discovery



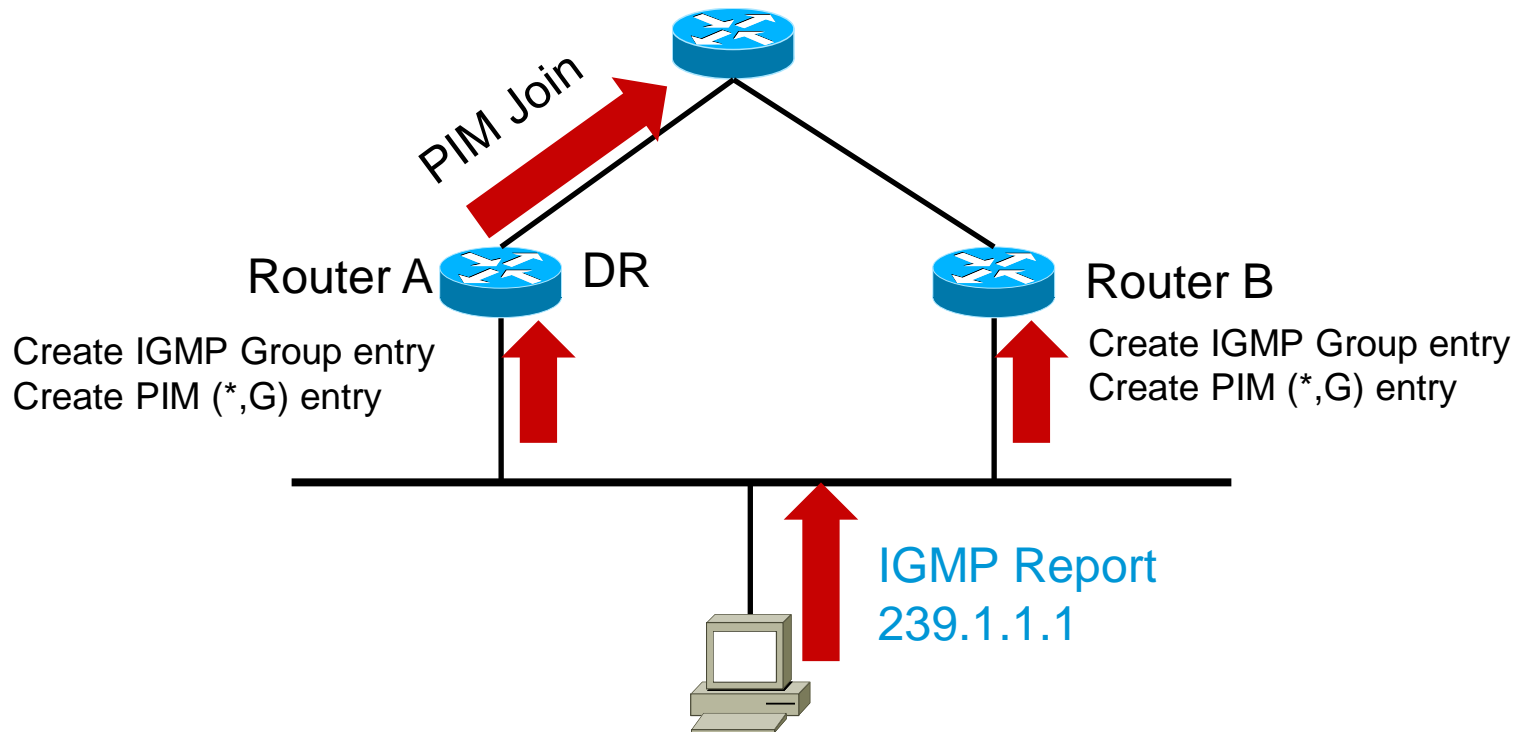
- PIM Hellos are periodically multicast to the “All-PIM-Routers” (224.0.0.13) group address. (Default = 30 seconds)
- Designated Router (DR) Election

Compare DR Priority , larger priority is always preferred

If DR Priority is same, Highest Interface IP address preferred

# Designated Router

- The DR is responsible for sending all Joins and Register messages for any receivers or senders on the network.



# Designated Router

- Show IP PIM

```
router#show ip pim neighbor
```

```
PIM Neighbor Table
```

```
Mode: B - Bidir Capable, DR - Designated Router, N - Default DR Priority,  
      S - State Refresh Capable
```

Neighbor Address	Interface	Uptime/Expires	Ver	DR Prio/Mode
10.1.6.1	GigabitEthernet1/0	01:03:54/00:01:21	v2	1 / S
10.1.4.1	GigabitEthernet2/0	01:04:19/00:01:24	v2	100 / DR S

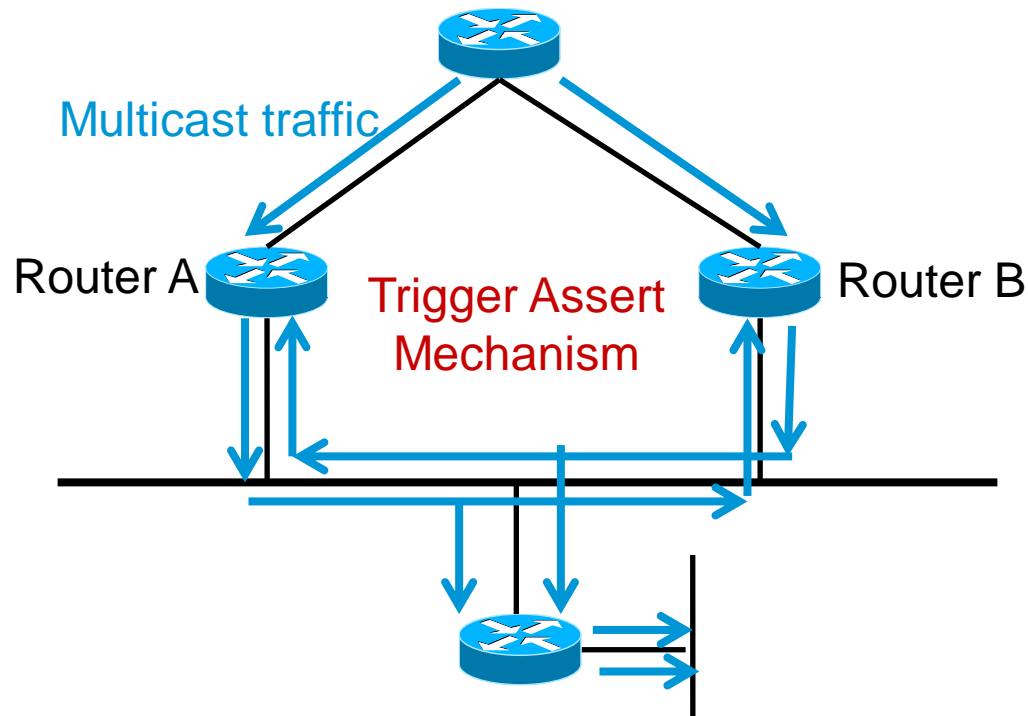
Mode – PIM mode (Sparse, Dense, Sparse/Dense) that the PIM Neighbor is using.

(DR) – Indicates that this PIM Neighbor is the Designated Router for the network.

# PIM Assert Mechanism

Avoid duplicate flows onto the same multi-access network.

Routers detect this condition when they receive an (S, G) packet via a multi-access interface that it is in the (S, G) OIL



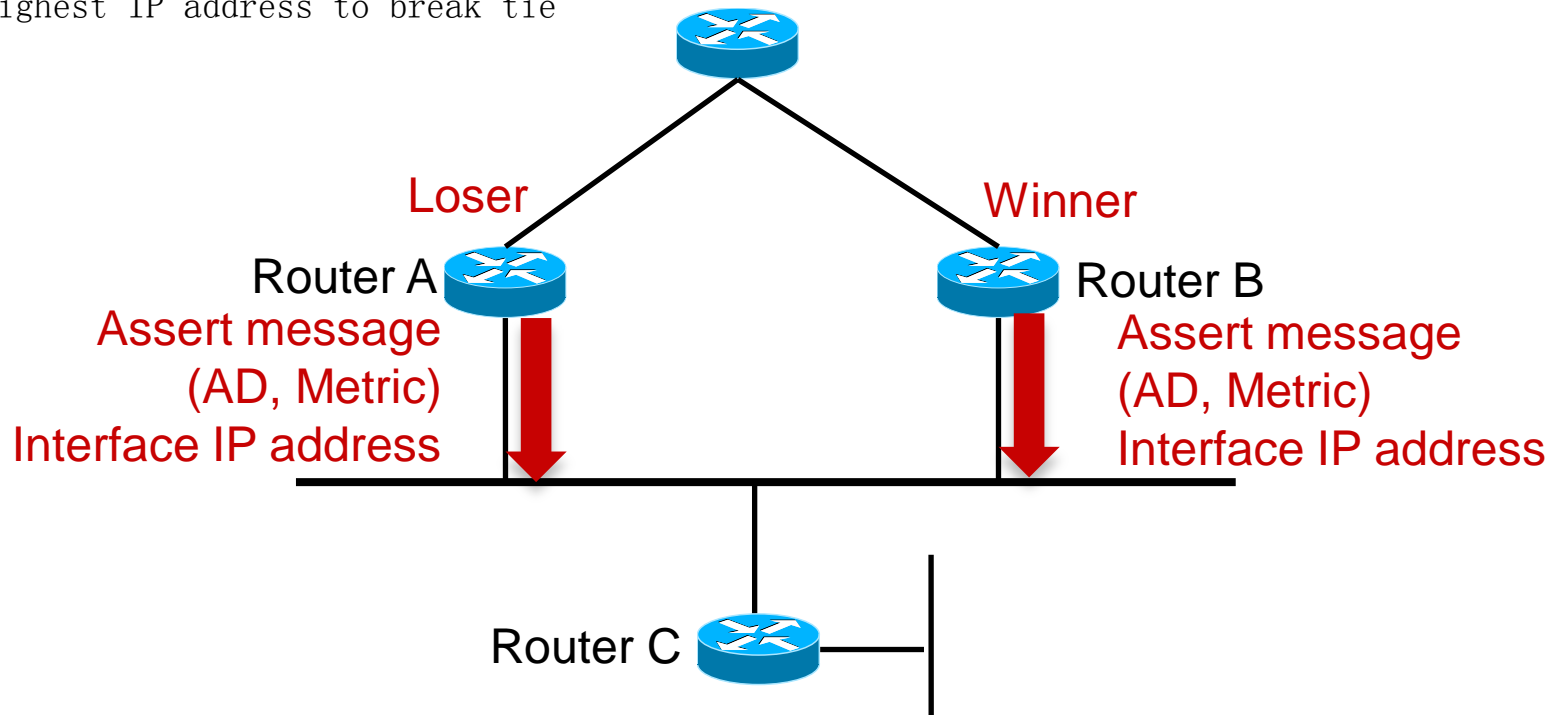
# PIM Assert Mechanism

Trigger assert mechanism

Both send PIM Assert messages that contain their Administrative Distance and route Metric back to the source or RP.

AD/metric the best (lowest) value wins the Assert.

Highest IP address to break tie



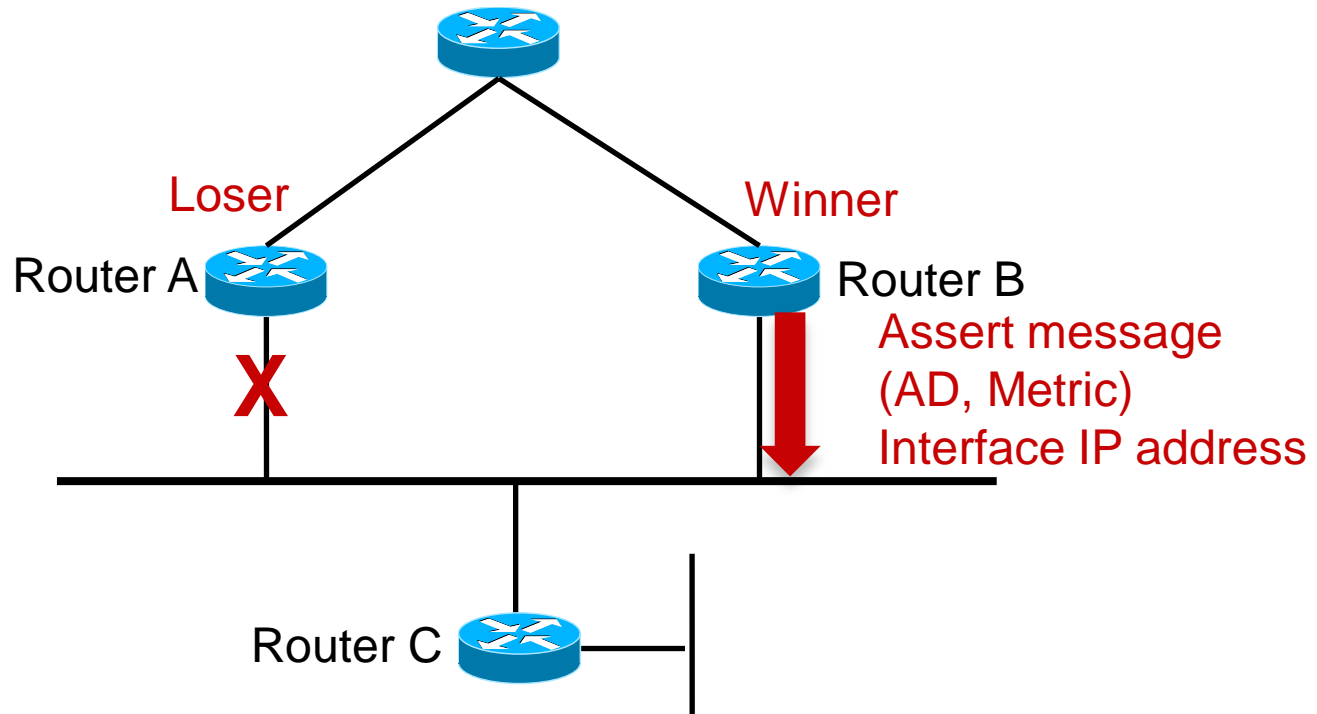
# PIM Assert Mechanism

Trigger assert mechanism

Assert Winner will periodically send out assert message.

Assert Loser will prune its interface from OIL.

Router C also will receive the ASSERT message, update RFP interface towards Assert Winner. Send join or prune message to ASSERT winner.



# PIM SM State Entry

LH#show ip mroute

IP Multicast Routing Table

Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,

L - Local, P - Pruned, R - RP-bit set, F - Register flag,

T - SPT-bit set, J - Join SPT, M - MSDP created entry,

X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement,

U - URD, I - Received Source Specific Host Report,

Z - Multicast Tunnel, z - MDT-data group sender,

Y - Joined MDT-data group, y - Sending to MDT-data group

Outgoing interface flags: H - Hardware switched, A - Assert winner

Timers: Uptime/Expires

Interface state: Interface, Next-Hop or VCD, State/Mode

(\*, 239.1.1.1), 01:17:19/stopped, RP 3.3.3.3, flags: SJC

Incoming interface: GigabitEthernet1/0, RPF nbr 10.1.6.1

(\* ,G) Entry

Outgoing interface list:

E0, Forward/Sparse, 01:17:19/00:00:43

(10.1.3.2, 239.1.1.1), 00:00:04/00:02:55, flags: JT

Incoming interface: GigabitEthernet2/0, RPF nbr 10.1.4.1

(S,G) Entry

Outgoing interface list:

E0, Forward/Sparse, 00:00:04/00:02:55

# PIM SM (\*, G) State Rules

- (\*, G) creation
  - Receipt of a (\*, G) Join or IGMP Report
  - **Automatically if (S, G) must be created**
- (\*, G) reflects default group forwarding
  - IIF(incoming interface) = RPF interface toward RP
  - OIL(outgoing interface list ) = interfaces that
    - received a (\*, G) Join or
    - with directly connected members or
    - manually configured
- (\*, G) deletion
  - When OIL = NULL and
  - no child (S, G) state exists



# PIM SM (S, G) State Rules

- (S,G) creation
  - By receipt of (S,G) Join or Prune or
    - By “Register” process
      - Parent (\*,G) created (if doesn't exist)
- (S,G) reflects forwarding of “S” to “G”
  - IIF = RPF Interface normally toward source
    - RPF toward RP if “RP-bit” set
  - OIL = Initially, copy of (\*,G) OIL minus IIF
- (S,G) deletion
  - By normal (S,G) entry timeout

# Triggering Join/Prune Messages

- $(*,G)$  Joins are triggered when:

The  $(*,G)$  OIL transitions from Null to non-Null

- $(*,G)$  Prunes are triggered when:

The  $(*,G)$  OIL transitions from non-Null to Null

- $(S,G)$  Joins are triggered when:

The  $(S,G)$  OIL transitions from Null to non-Null

- $(S,G)$  Prunes are triggered when:

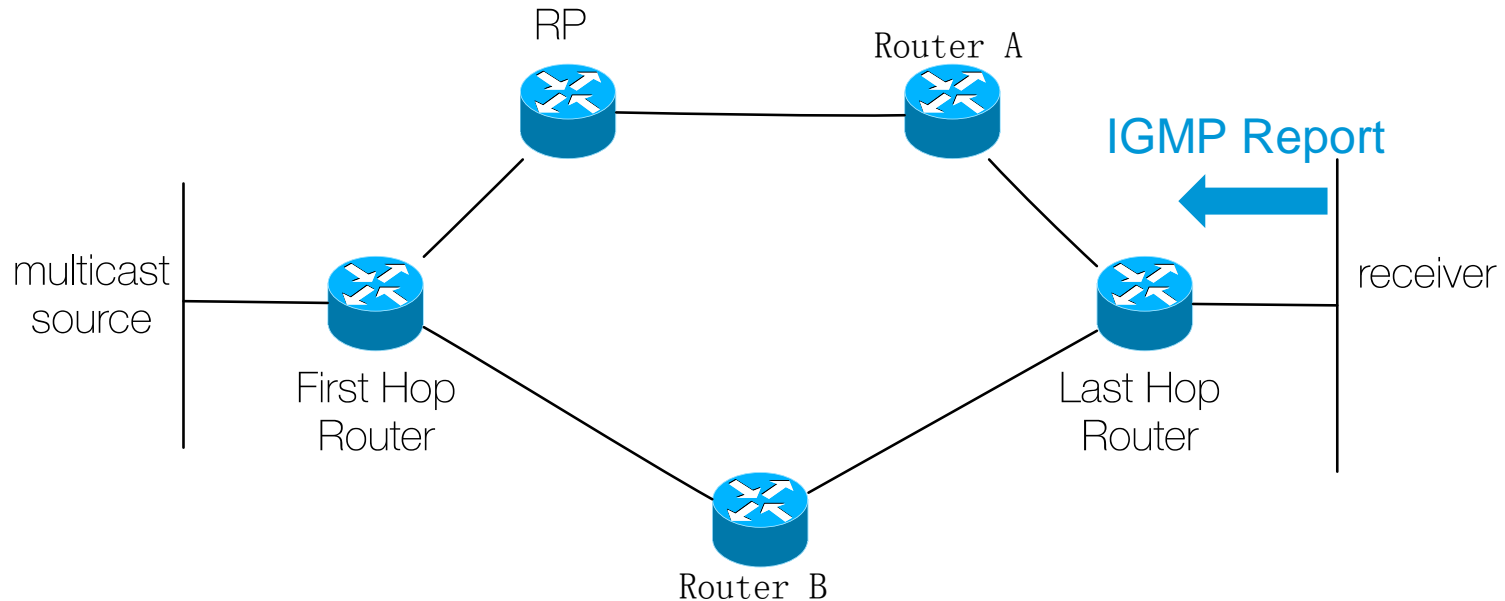
The  $(S,G)$  OIL is Null AND

A packet is received on the incoming interface

- $(S,G)$ RP-bit Prunes are triggered when:

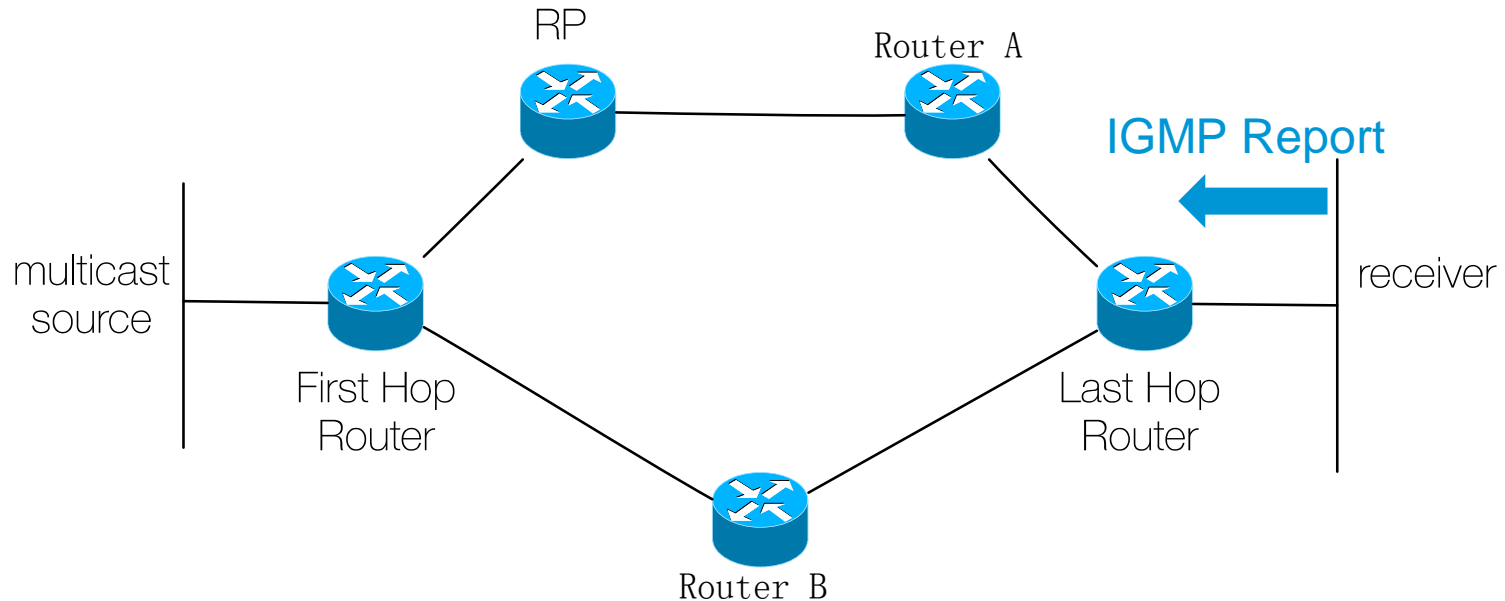
The  $(S,G)$  RPF info  $\neq$  the  $(*,G)$  RPF info

# PIM-SM Review



```
LH#show ip igmp group
IGMP Connected Group Membership
Group Address Interface Uptime Expires Last Reporter
239.1.1.1      E0      00:10:29 00:02:10 10.7.1.1
```

# PIM-SM Review



```
LH#show ip mroute
(*, 239.1.1.1), 00:14:23/stopped, RP 3.3.3.3, flags: SJC
  Incoming interface: GigabitEthernet1/0, RPF nbr 10.1.6.1
  Outgoing interface list:
    E0, Forward/Sparse, 00:14:23/00:00:37
```

S : Sparse mode  
J : Join SPT  
C : Connected

# PIM-SM Review

J Flag : When J flag is set in a (\*, G) entry, it indicates that the rate of traffic flowing down the Shared Tree is above the SPT-Threshold and will cause a switch to the SPT for the next packet received down the shared tree.

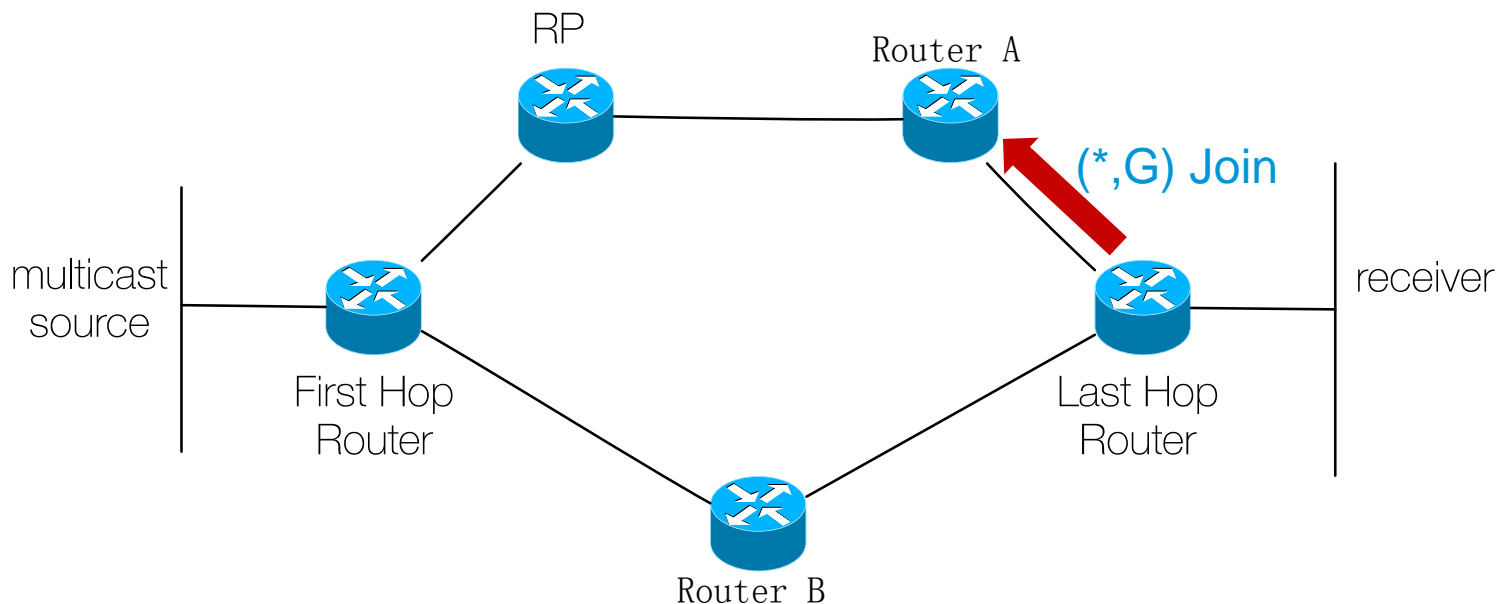
```
LH#show ip mroute
(*, 239.1.1.1), 00:14:23/stopped, RP 3.3.3.3, flags: SJC
Incoming interface: GigabitEthernet1/0, RPF nbr 10.1.6.1
Outgoing interface list:
E1, Forward/Sparse, 00:14:23/00:00:37
```

**LH(config)#ip pim spt-threshold infinity**



```
LH#show ip mroute
(*, 239.1.1.1), 00:21:53/stopped, RP 3.3.3.3, flags: SC
Incoming interface: GigabitEthernet1/0, RPF nbr 10.1.6.1
Outgoing interface list:
Loopback1, Forward/Sparse, 00:21:53/00:02:08
```

# PIM-SM Review

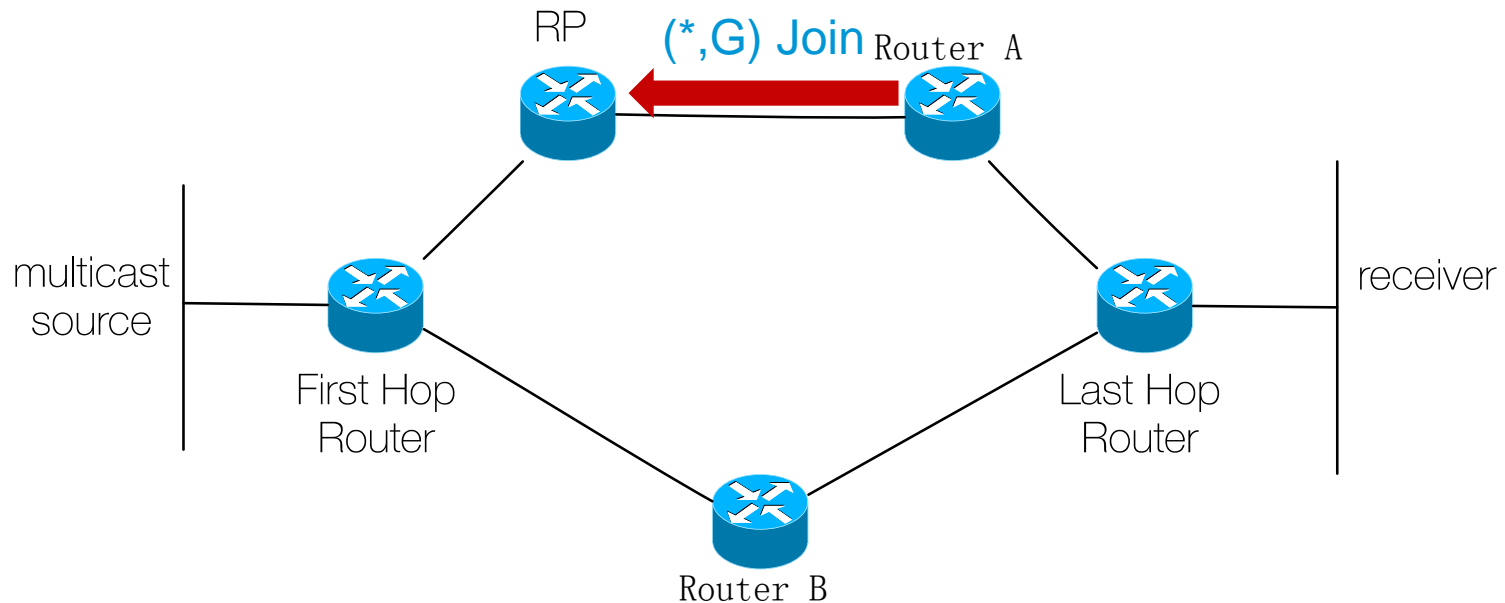


Create (\*,G) entry on Router A

```
Router A#show ip mroute
```

```
(*, 239.1.1.1), 01:58:56/00:02:53, RP 3.3.3.3, flags: S  
Incoming interface: GigabitEthernet1/0, RPF nbr 10.1.5.1  
Outgoing interface list:  
GigabitEthernet2/0, Forward/Sparse, 01:58:56/00:02:53
```

# PIM-SM Review



Propagate Join (\*,G) to RP

```
RP#show ip mroute
```

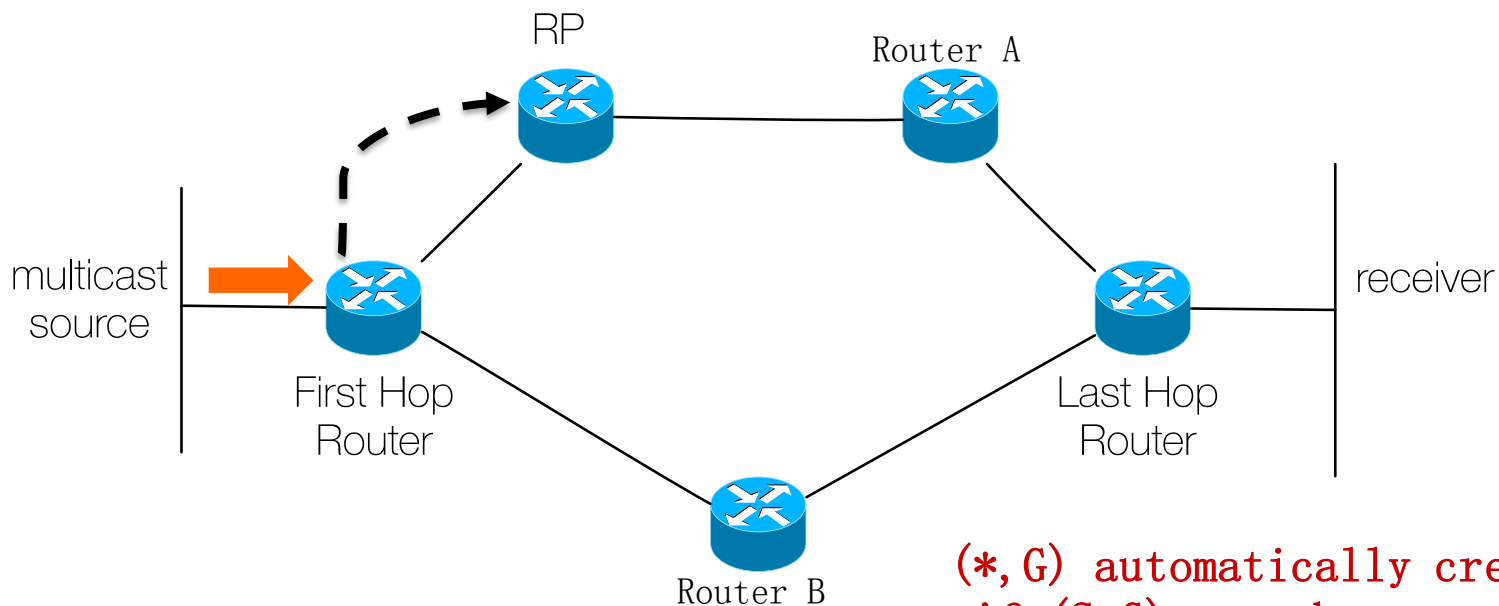
```
(*, 239.1.1.1), 02:02:55/00:02:41, RP 3.3.3.3, flags: S
```

```
Incoming interface: Null, RPF nbr 0.0.0.0
```

```
Outgoing interface list:
```

```
GigabitEthernet2/0, Forward/Sparse, 02:02:55/00:02:41
```

# PIM-SM Review



Source send out traffic to FH router

(\* , G) automatically create  
if (S, G) must be created

```
FH#show ip mroute
(*, 239.1.1.1), 00:00:08/stopped, RP 3.3.3.3, flags: SPF
Incoming interface: GigabitEthernet1/0, RPF nbr 10.1.1.2
Outgoing interface list: Null

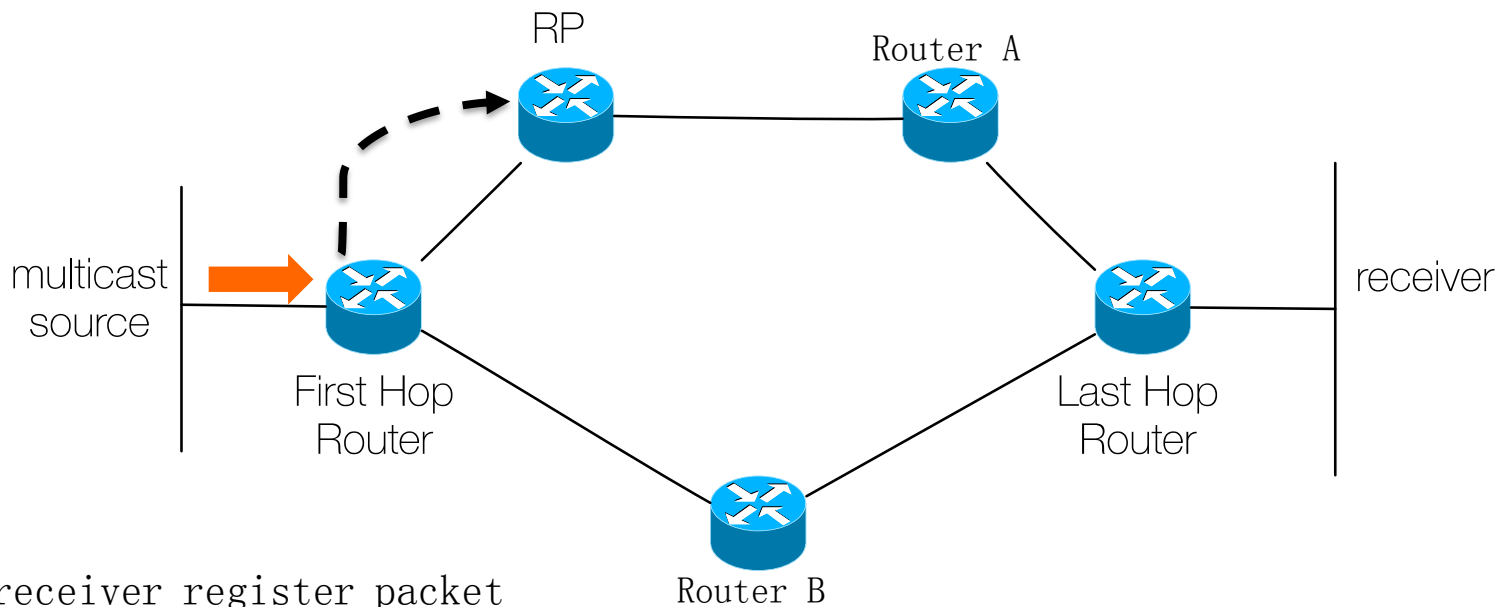
(10.1.3.2, 239.1.1.1), 00:00:08/00:03:23, flags: PFT
Incoming interface: GigabitEthernet3/0, RPF nbr 0.0.0.0
Outgoing interface list:
```

P - Pruned

F - Register flag



# PIM-SM Review

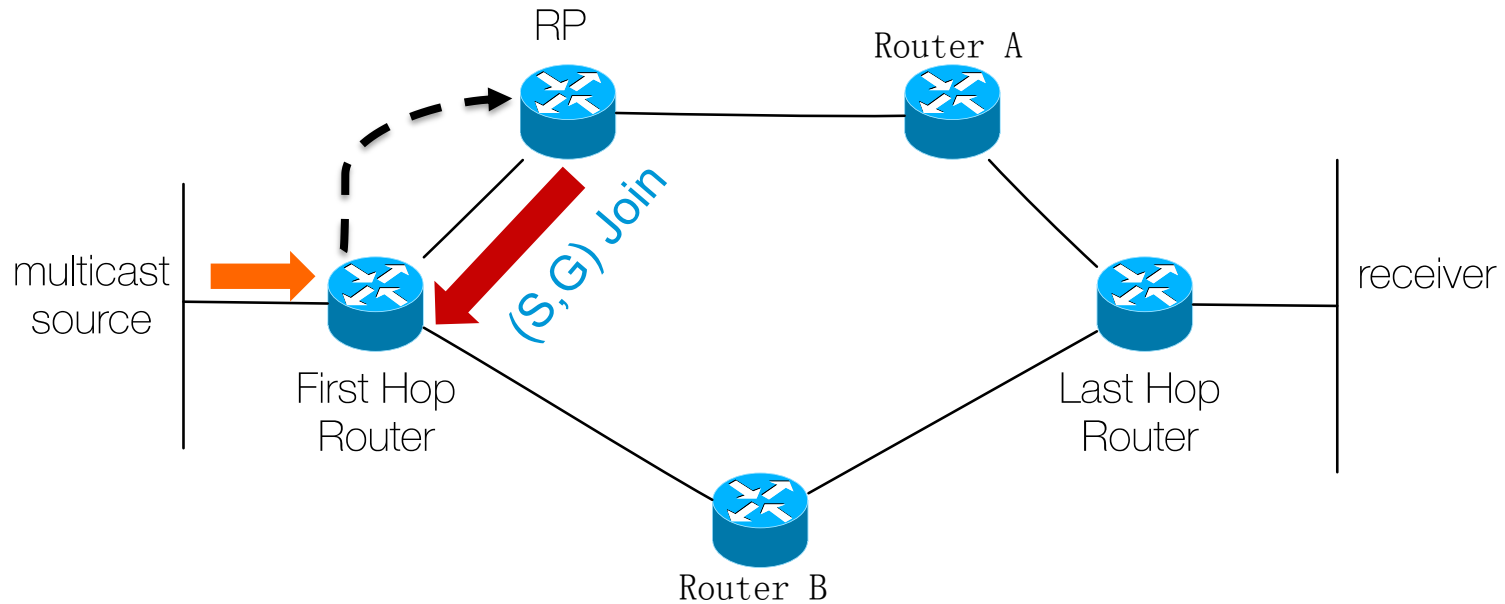


RP receiver register packet

```
RP#show ip mroute
(*, 239.1.1.1), 00:00:06/stopped, RP 3.3.3.3, flags: S
  Incoming interface: Null, RPF nbr 0.0.0.0
  Outgoing interface list:
    GigabitEthernet2/0, Forward/Sparse, 02:14:46/00:02:35

(10.1.3.2, 239.1.1.1), 00:00:06/00:02:53, flags: S
  Incoming interface: GigabitEthernet1/0, RPF nbr 10.1.1.1
  Outgoing interface list: Null
    GigabitEthernet2/0, Forward/Sparse, 02:14:46/00:02:35
```

# PIM-SM Review

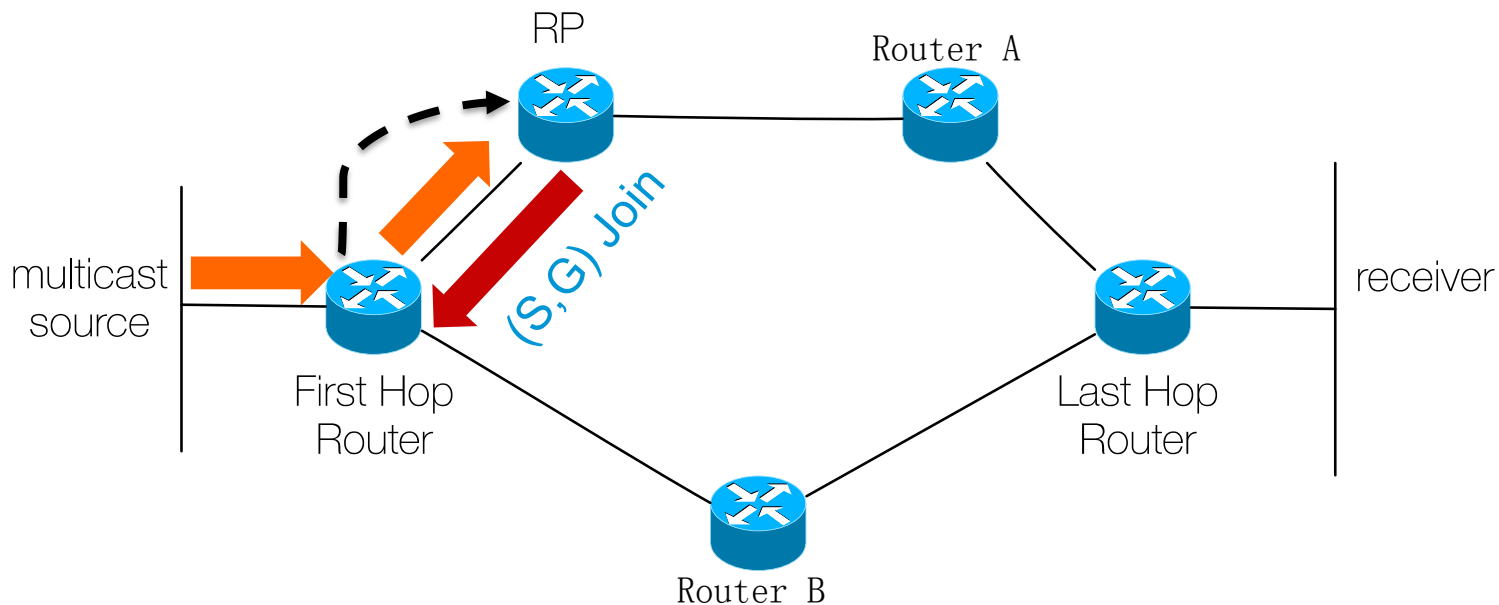


If  $RP (*, G) OIL \neq NULL$ , RP send out (S,G) Join toward source

If not, RP send out register stop to first hop router, to refresh register-stop timer

After register-stop timer expiry, first hop router send out register packet again.

# PIM-SM Review

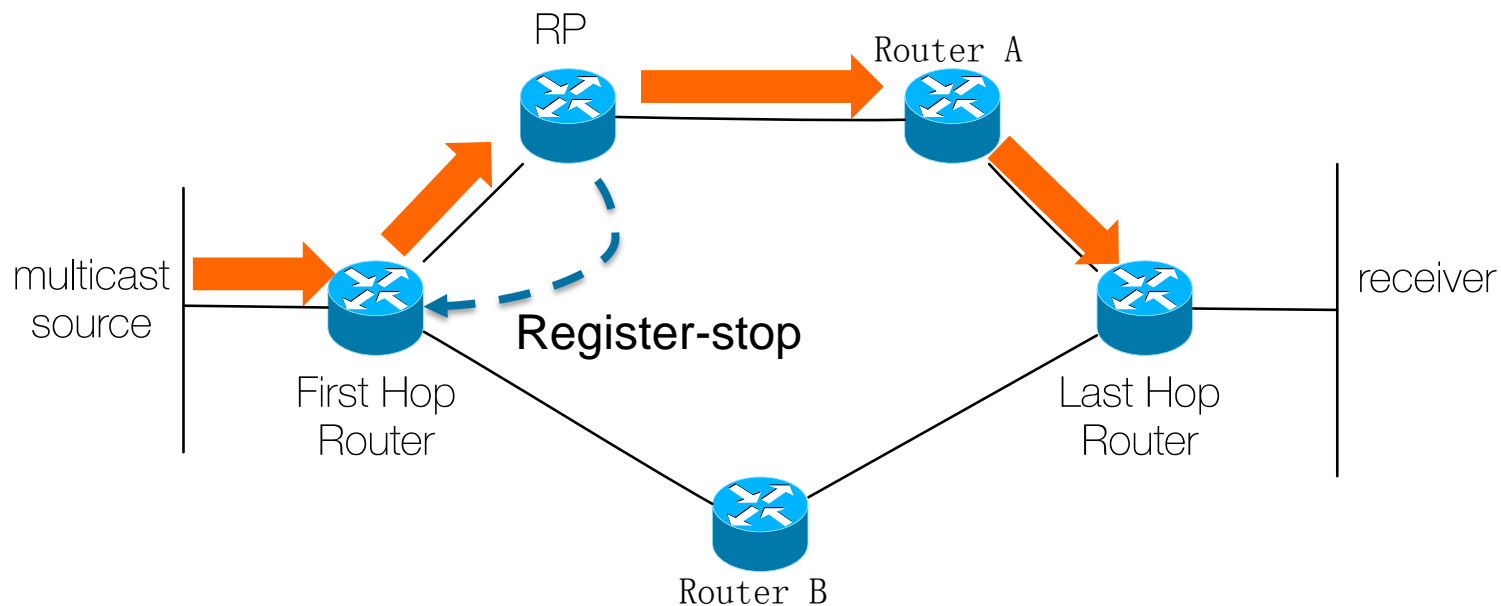


```
FH#show ip mroute
```

```
(*, 239.1.1.1), 00:24:28/stopped, RP 3.3.3.3, flags: SPF  
Incoming interface: GigabitEthernet1/0, RPF nbr 10.1.1.2  
Outgoing interface list: Null
```

```
(10.1.3.2, 239.1.1.1), 00:09:36/00:03:22, flags: FT  
Incoming interface: GigabitEthernet3/0, RPF nbr 0.0.0.0  
Outgoing interface list:  
GigabitEthernet1/0, Forward/Sparse, 00:00:19/00:03:10
```

# PIM-SM Review

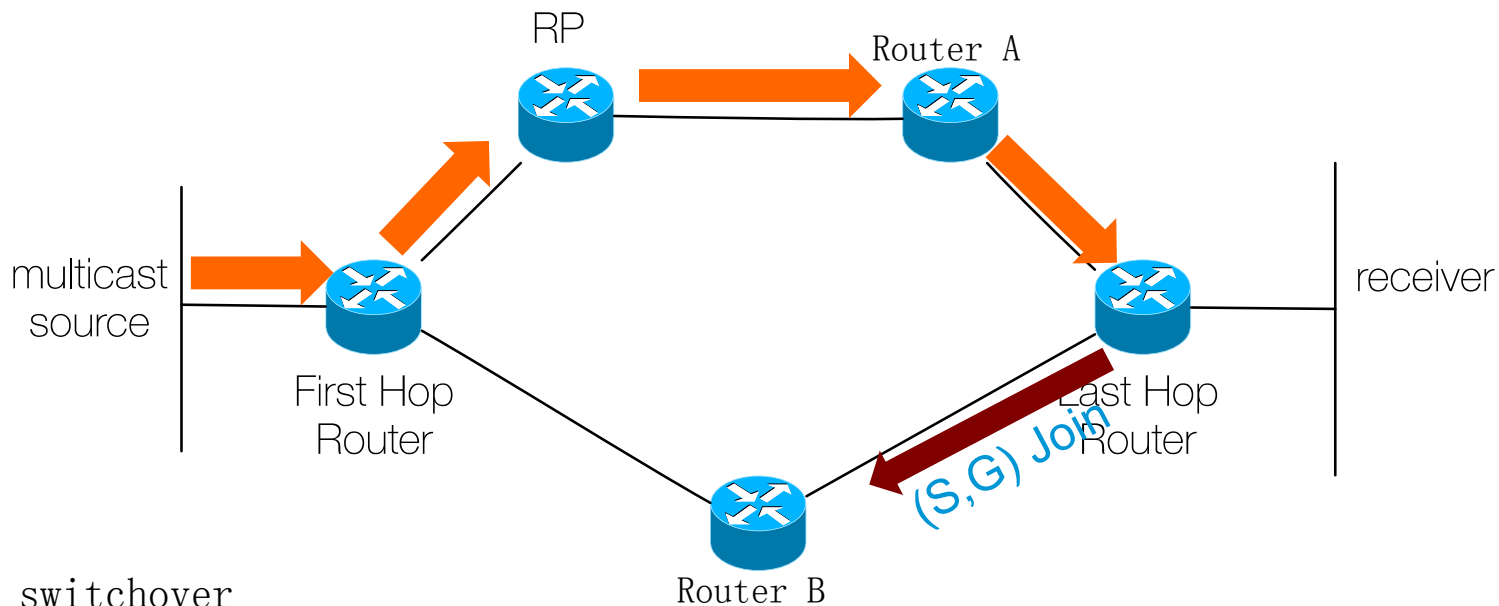


RP send out unicast (S,G) register stop message to first hop router. To stop unicast register Traffic.

On first hop router, when first hop router receive the register stop packet , will stop sending register packet and refresh register stop timer.

If the register timer timeout, FH router will send out data-header register packet again.

# PIM-SM Review

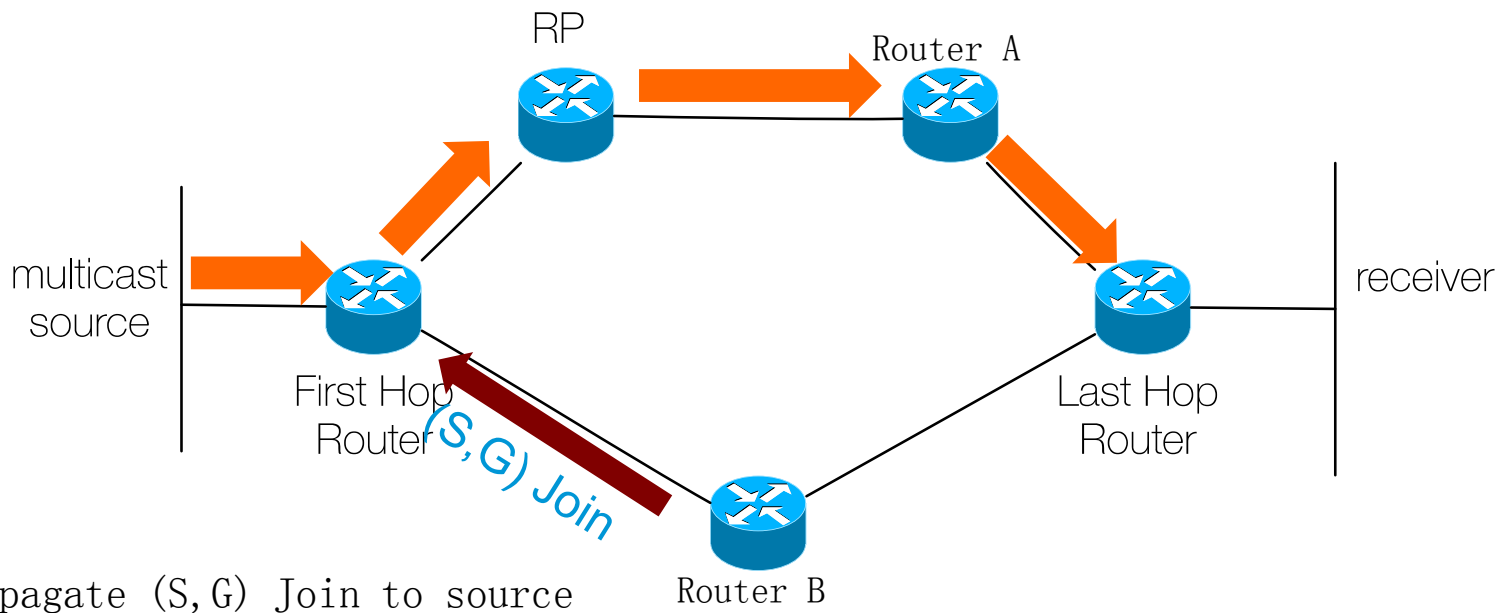


SPT switchover

```
LH#show ip mroute
(*, 239.1.1.1), 00:22:52/stopped, RP 3.3.3.3, flags: SJC
Incoming interface: GigabitEthernet1/0, RPF nbr 10.1.6.1
Outgoing interface list:
E0 , Forward/Sparse, 00:22:51/00:01:09

(10.1.3.2, 239.1.1.1), 00:00:04/00:02:55, flags: J
Incoming interface: GigabitEthernet2/0, RPF nbr 10.1.4.1
Outgoing interface list:
E0, Forward/Sparse, 00:00:04/00:02:55
```

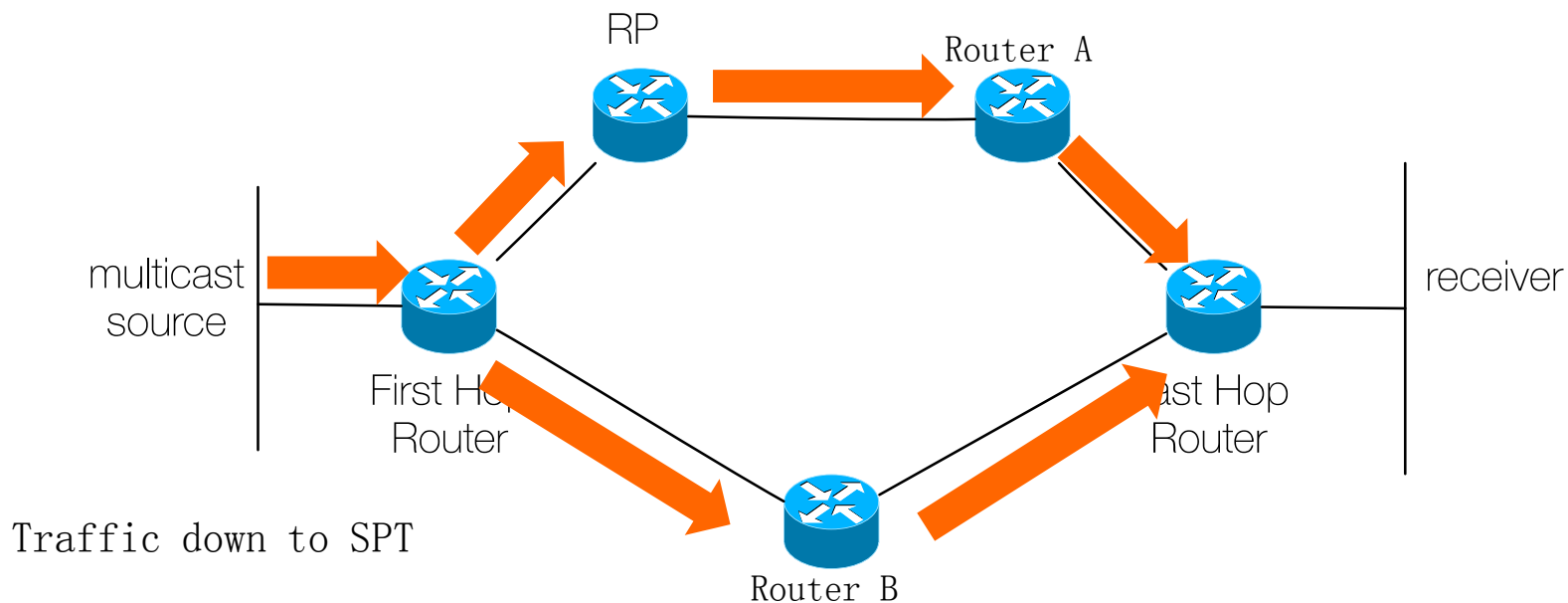
# PIM-SM Review



```
Router B#show ip mroute
(*, 239.1.1.1), 00:03:24/stopped, RP 3.3.3.3, flags: SP
  Incoming interface: GigabitEthernet2/0, RPF nbr 10.1.2.1
  Outgoing interface list: Null

(10.1.3.2, 239.1.1.1), 00:03:24/00:03:26, flags:
  Incoming interface: GigabitEthernet2/0, RPF nbr 10.1.2.1
  Outgoing interface list:
    GigabitEthernet1/0, Forward/Sparse, 00:03:24/00:03:03
```

# PIM-SM Review

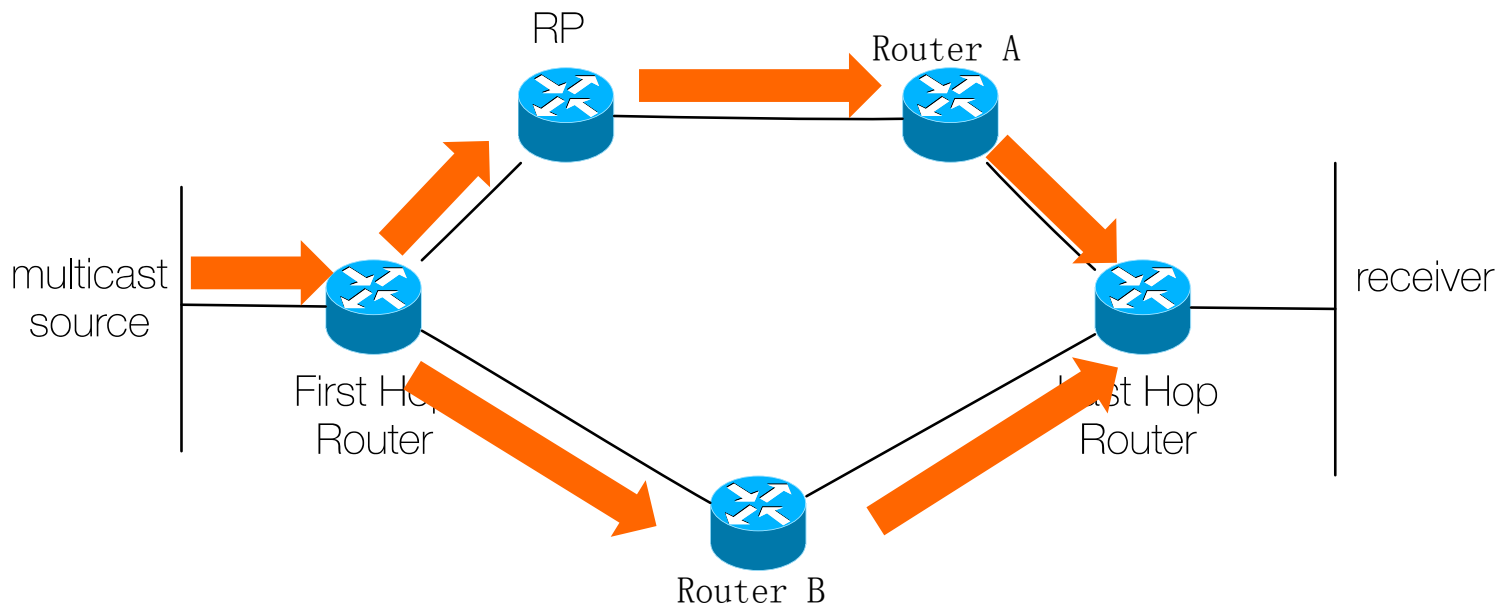


```
FH#show ip mroute
```

```
(*, 239.1.1.1), 00:10:10/stopped, RP 3.3.3.3, flags: SPF  
Incoming interface: GigabitEthernet1/0, RPF nbr 10.1.1.2  
Outgoing interface list: Null
```

```
(10.1.3.2, 239.1.1.1), 00:10:10/00:03:28, flags: FT  
Incoming interface: GigabitEthernet3/0, RPF nbr 0.0.0.0  
Outgoing interface list:  
GigabitEthernet1/0, Forward/Sparse, 00:20:10/00:03:00  
GigabitEthernet2/0, Forward/Sparse, 00:10:10/00:03:10
```

# PIM-SM Review



```
Router B#show ip mroute
```

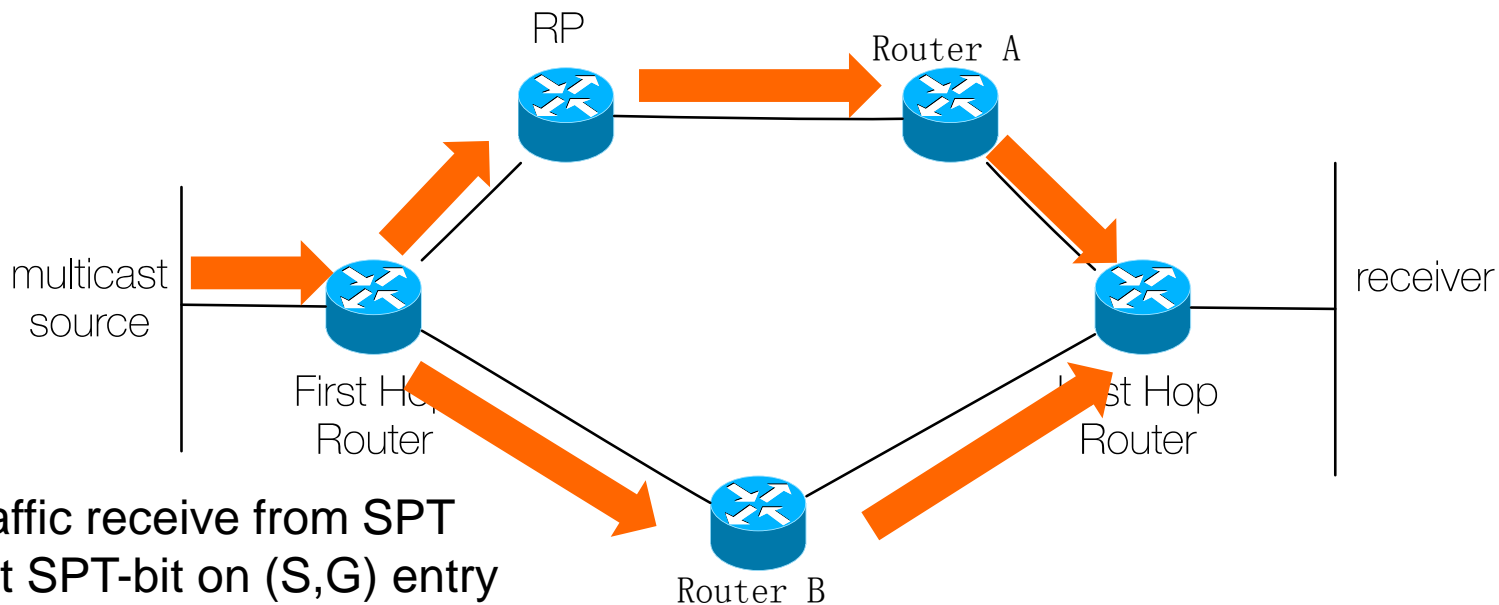
```
(*, 239.1.1.1), 00:15:14/stopped, RP 3.3.3.3, flags: SP  
Incoming interface: GigabitEthernet2/0, RPF nbr 10.1.2.1  
Outgoing interface list: Null
```

```
(10.1.3.2, 239.1.1.1), 00:00:07/00:03:25, flags: T  
Incoming interface: GigabitEthernet2/0, RPF nbr 10.1.2.1  
Outgoing interface list:  
GigabitEthernet1/0, Forward/Sparse, 00:00:07/00:03:22
```

T - SPT-bit set



# PIM-SM Review



Traffic receive from SPT  
Set SPT-bit on (S,G) entry

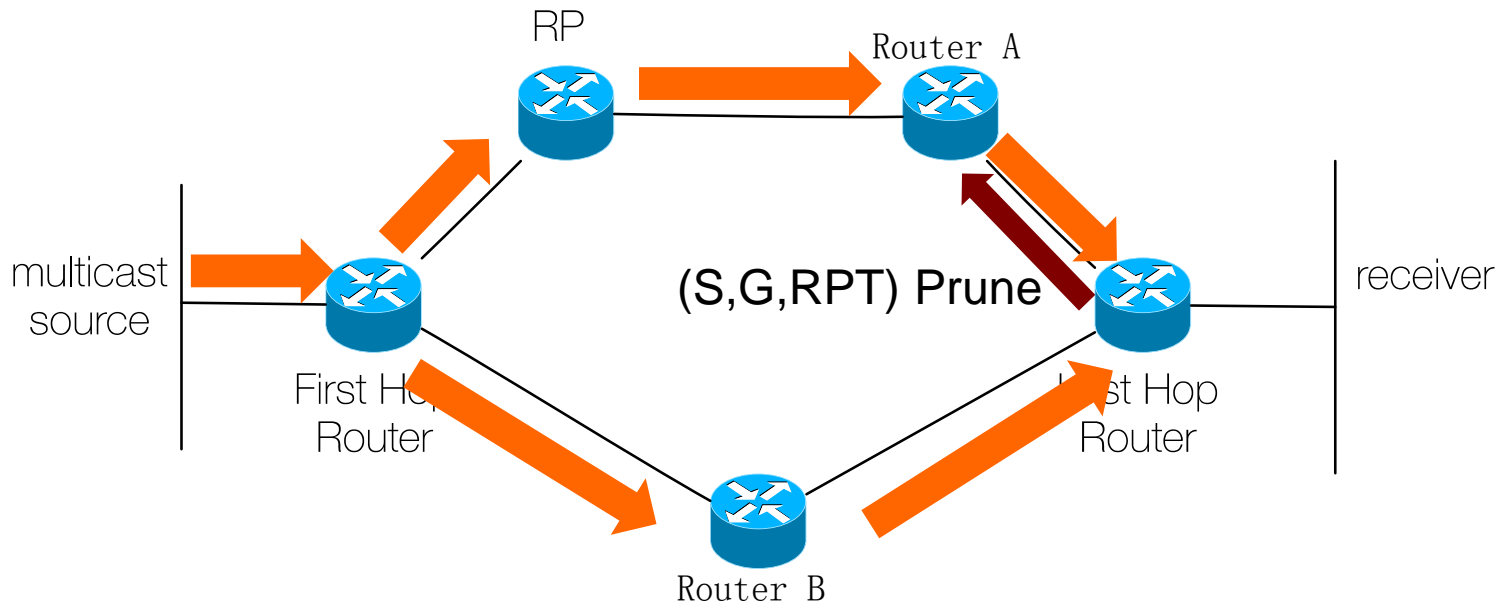
```
LH#show ip mroute
```

```
(*, 239.1.1.1), 00:42:18/stopped, RP 3.3.3.3, flags: SJC  
Incoming interface: GigabitEthernet1/0, RPF nbr 10.1.6.1  
Outgoing interface list:  
E0, Forward/Sparse, 00:42:17/00:02:44
```

```
(10.1.3.2, 239.1.1.1), 00:04:23/00:02:57, flags: JT  
Incoming interface: GigabitEthernet2/0, RPF nbr 10.1.4.1  
Outgoing interface list:  
E0, Forward/Sparse, 00:04:23/00:01:36
```

T - SPT-bit set, the traffic  
Receive from RPT  
will be drop.

# PIM-SM Review

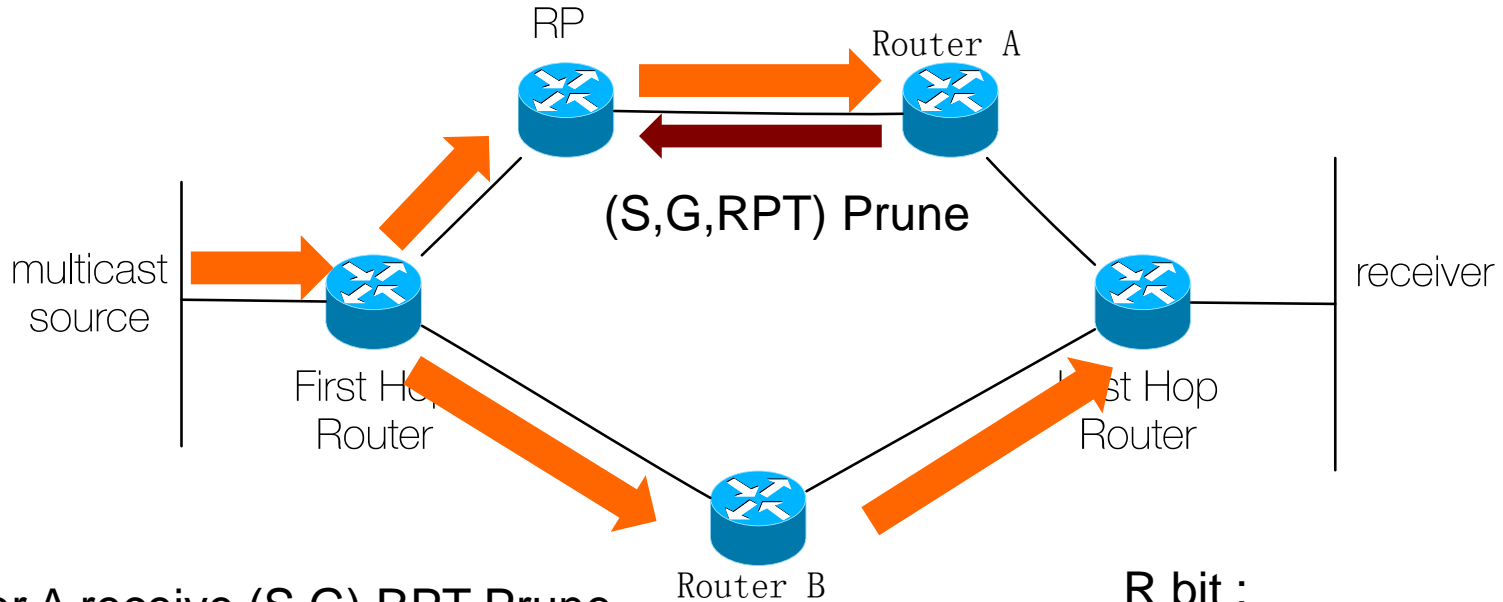


Interface of  $RPF(*,G) \neq$  interface of  $RPF(S,G)$ , and  $(S,G)$  entry with T flag. It will trigger  $(S,G)$  RP-bit prune.

$(S,G)$  RP-bit prune send out towards RP upstream.

$(S,G)$  RP-bit prune entry added to  $(*,G)$  Join packet and send to RP periodically, to modify the results of  $(*,G)$  Joins

# PIM-SM Review



## Router A receive (S,G) RPT Prune

Router A #show ip mroute

(\*, 239.1.1.1), 00:00:13/stopped, RP 3.3.3.3, flags: S  
 Incoming interface: GigabitEthernet1/0, RPF nbr 10.1.5.1  
 Outgoing interface list:

**GigabitEthernet2/0, Forward/Sparse, 00:00:13/00:03:20**

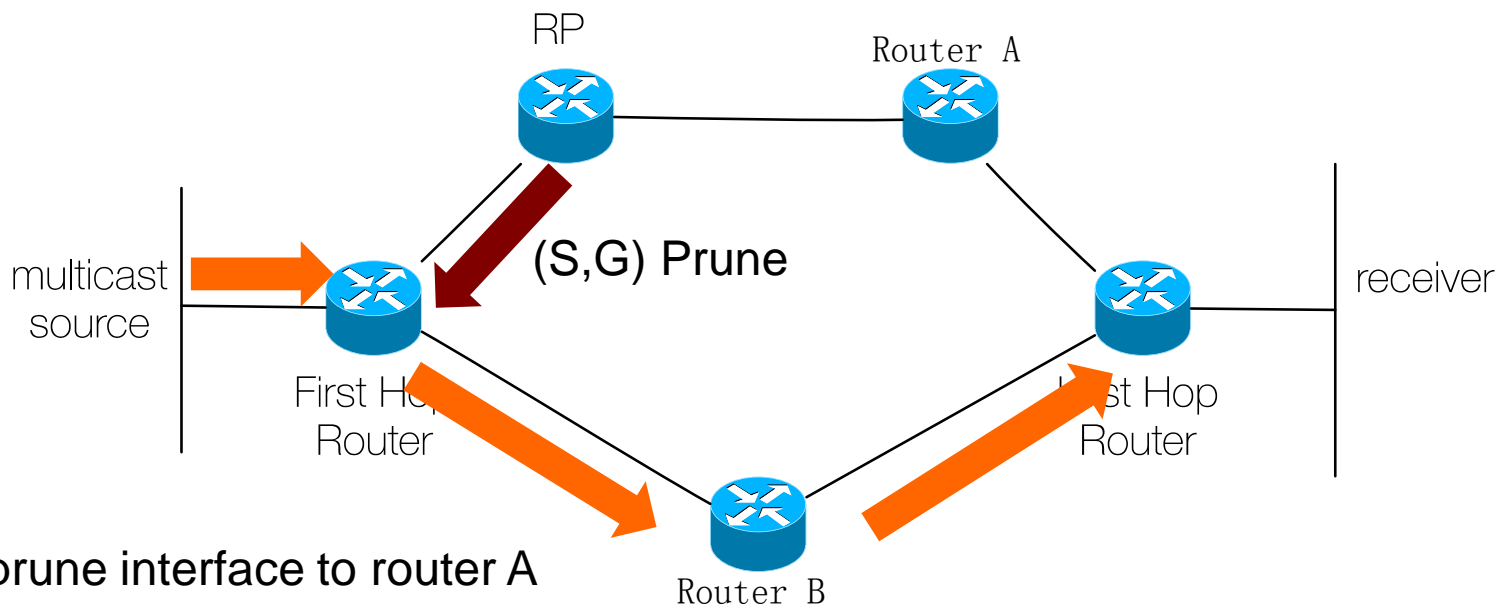
(10.1.3.2, 239.1.1.1), 00:00:06/00:02:53, flags: PR  
 Incoming interface: **GigabitEthernet1/0**, RPF nbr 10.1.5.1  
 Outgoing interface list: Null

R bit :

1. Forward traffic on RPT
2. RPF information must be changed to point toward the RP
3. Receive (S,G) RP-bit prune

P bit : Prune

# PIM-SM Review



RP prune interface to router A  
Null OIL of (S,G) entry, trigger prune

```
RP#show ip mroute
(*, 239.1.1.1), 00:21:45/00:03:19, RP 3.3.3.3, flags: S
  Incoming interface: Null, RPF nbr 0.0.0.0
  Outgoing interface list:
    GigabitEthernet2/0, Forward/Sparse, 00:21:45/00:03:19

(10.1.3.2, 239.1.1.1), 00:09:17/00:01:50, flags: PT
  Incoming interface: GigabitEthernet1/0, RPF nbr 10.1.1.1
  Outgoing interface list: Null
```

# RP Discovery

- Static RP

## Hard-coded RP address

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

# RP Discovery

- PIMv2 Bootstrap router (BSR)

## Candidate RPs

Unicast PIMv2 C-RP messages to BSR

Learns IP address of BSR from BSR messages

Sent every rp-announce-interval (default: 60 sec)

## C-RP messages contain:

Group Range (default = 224.0.0.0/4)

Candidate's RP address

Holdtime = 3 x <rp-announce-interval>

# RP Discovery

- PIMv2 Bootstrap router (BSR)

## All PIMv2 routers

Receive BSR messages

Stored in local Group-to-RP Mapping Cache

Information used to determine active BSR address

Selects RP using Hash algorithm

Selected from local Group-to-RP Mapping Cache

All routers select same RP using same algorithm

Permits RP-load balancing across group range

# RP Discovery

- PIMv2 BSR

## Candidate bootstrap router (C-BSR)

C-BSR with highest priority elected BSR

C-BSR IP address used as tie-breaker

(Highest IP address wins)

The active BSR may be preempted

New router w/higher BSR priority forces new election



# Layer3 Multicast troubleshooting CLI

```
Router#show ip mroute
```

```
(* , 239.1.1.1), 00:28:40/stopped, RP 5.5.5.5, flags: SJCL
```

```
Incoming interface: Null, RPF nbr 0.0.0.0
```

```
Outgoing interface list:
```

```
Loopback2, Forward/Sparse, 00:28:40/00:02:26
```

```
(10.1.1.3, 239.1.1.1), 00:00:10/00:02:52, flags: TA
```

```
Incoming interface: FastEthernet0/0, RPF nbr 10.1.2.1
```

```
Outgoing interface list:
```

```
Loopback2, Forward/Sparse, 00:00:10/00:02:49
```

```
Router#show ip mroute count
```

```
IP Multicast Statistics
```

```
Group: 239.1.1.1, Source count: 1, Packets forwarded: 5, Packets  
received: 5
```

```
RP-tree: Forwarding: 0/0/0/0, Other: 0/0/0
```

```
Source: 10.1.1.3/32, Forwarding: 5/1/100/0, Other: 5/0/0
```

# Layer3 Multicast troubleshooting CLI

```
Router#show ip mroute active
```

```
Active IP Multicast Sources - sending >= 4 kbps
```

```
Group: 233.69.12.16, (?)
```

```
Source: 10.46.1.171 (?)
```

```
Rate: 385 pps/771 kbps(1sec), 771 kbps(last 40 secs), 763  
kbps(life avg)
```

```
Router#show ip pim neighbor
```

```
PIM Neighbor Table
```

```
Mode: B - Bidir Capable, DR - Designated Router, N - Default DR Priority,  
S - State Refresh Capable
```

Neighbor Address	Interface	Uptime/Expires	Ver	DR Prio/Mode
10.1.2.1	FastEthernet0/0	00:03:13/00:01:28	v2	1 / S
10.1.4.2	FastEthernet1/0	00:31:41/00:01:35	v2	1 / DR

# Layer3 Multicast troubleshooting CLI

```
Router#show ip pim interface fastEthernet 0/0 detail
FastEthernet0/0 is up, line protocol is up
Internet address is 10.1.2.2/24
Multicast switching: fast
Multicast packets in/out: 17/0
Multicast TTL threshold: 0
PIM: enabled
  PIM version: 2, mode: sparse
  PIM DR: 10.1.2.2 (this system)
  PIM neighbor count: 1
  PIM Hello/Query interval: 30 seconds
  PIM Hello packets in/out: 10/13
  PIM State-Refresh processing: enabled
  PIM State-Refresh origination: disabled
  PIM NBMA mode: disabled
  PIM ATM multipoint signalling: disabled
  PIM domain border: disabled
Multicast Tagswitching: disabled
```

# Layer3 Multicast troubleshooting CLI

```
Router#show ip pim interface fastEthernet 0/0 detail
FastEthernet0/0 is up, line protocol is up
  Internet address is 10.1.2.2/24
  Multicast switching: fast
  Multicast packets in/out: 17/0
  Multicast TTL threshold: 0
  PIM: enabled
    PIM version: 2, mode: sparse
    PIM DR: 10.1.2.2 (this system)
    PIM neighbor count: 1
    PIM Hello/Query interval: 30 seconds
    PIM Hello packets in/out: 10/13
    PIM State-Refresh processing: enabled
    PIM State-Refresh origination: disabled
    PIM NBMA mode: disabled
    PIM ATM multipoint signalling: disabled
    PIM domain border: disabled
  Multicast Tagswitching: disabled
```

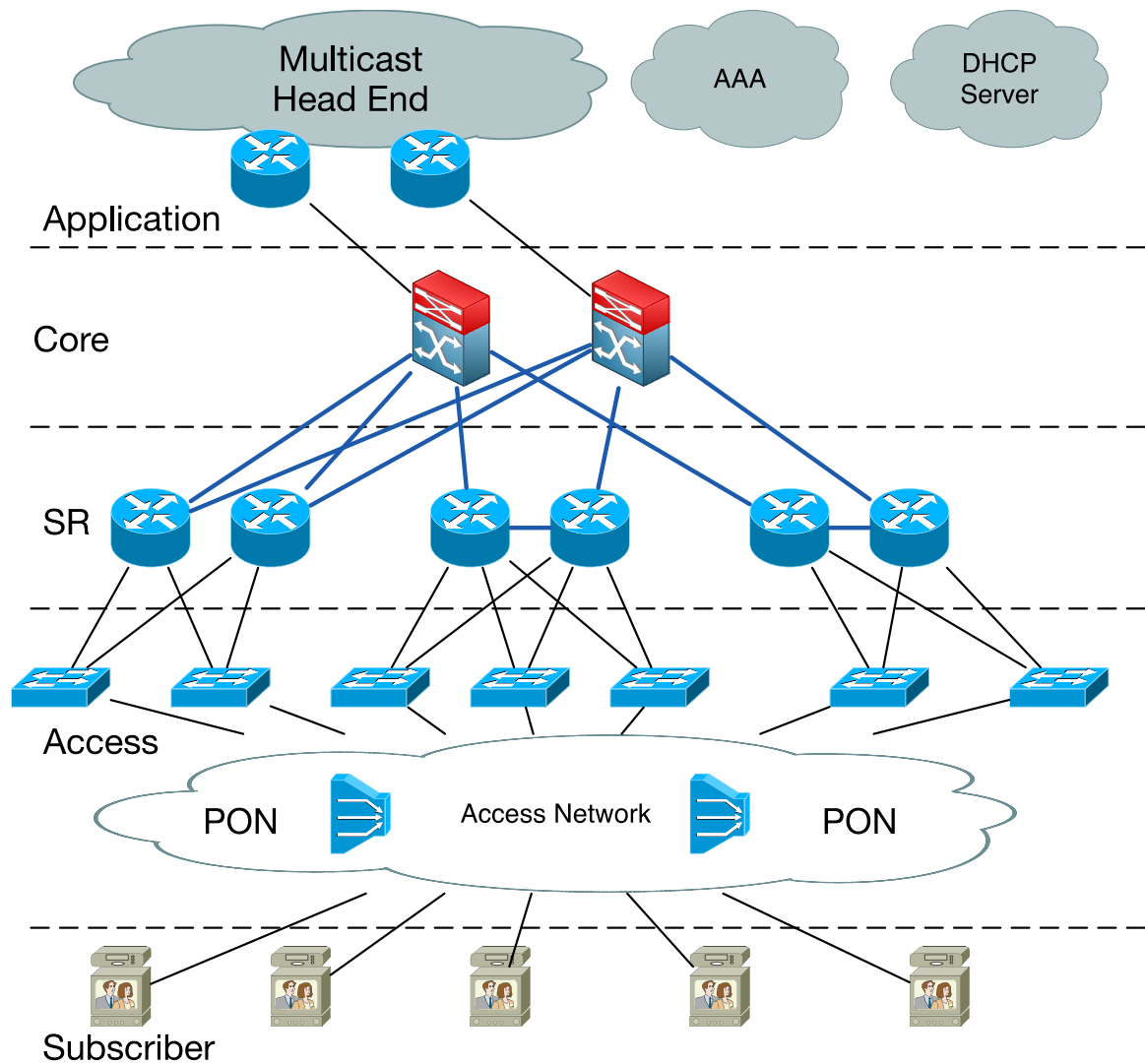
# Layer3 Multicast troubleshooting CLI

```
Router#show ip pim rp
Group: 239.1.1.1, RP: 5.5.5.5, next RP-reachable in 00:00:36
Group: 224.0.1.40, RP: 5.5.5.5, next RP-reachable in 00:00:36
```

```
Router#show ip pim rp mapping
PIM Group-to-RP Mappings
```

```
Group(s): 224.0.0.0/4, Static
RP: 5.5.5.5 (?)
```

# Multicast High Availability



- Last Hop Router protection

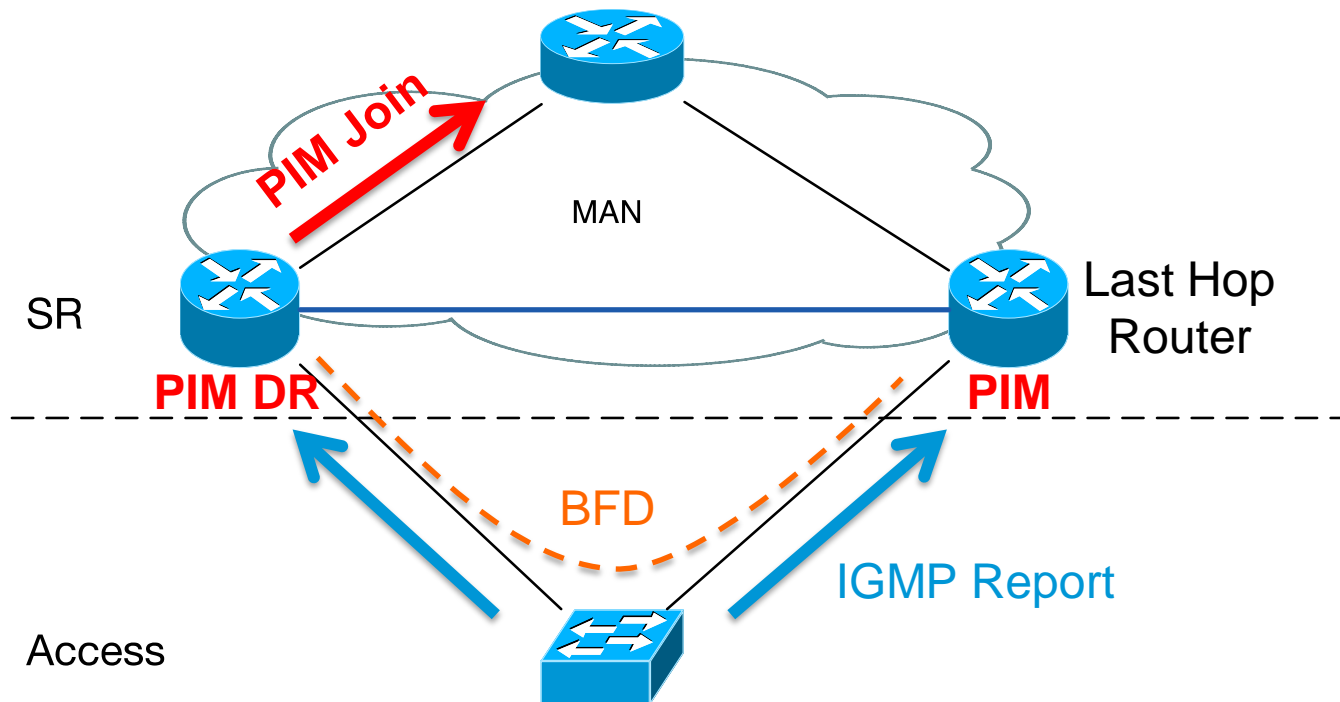
Fast convergency

- RP protection

Anycast RP

# Last Hop Router protection

Interface connected to access switch enable PIM + BFD



# Last Hop Router protection

## IOS Configuration

### 1. Interface enable bfd

```
device(config)# interface fastethernet 1/6
```

```
device(config-if)# bfd interval 500 min_rx 500 multiplier 5
```

### 2. Enable PIM + BFD

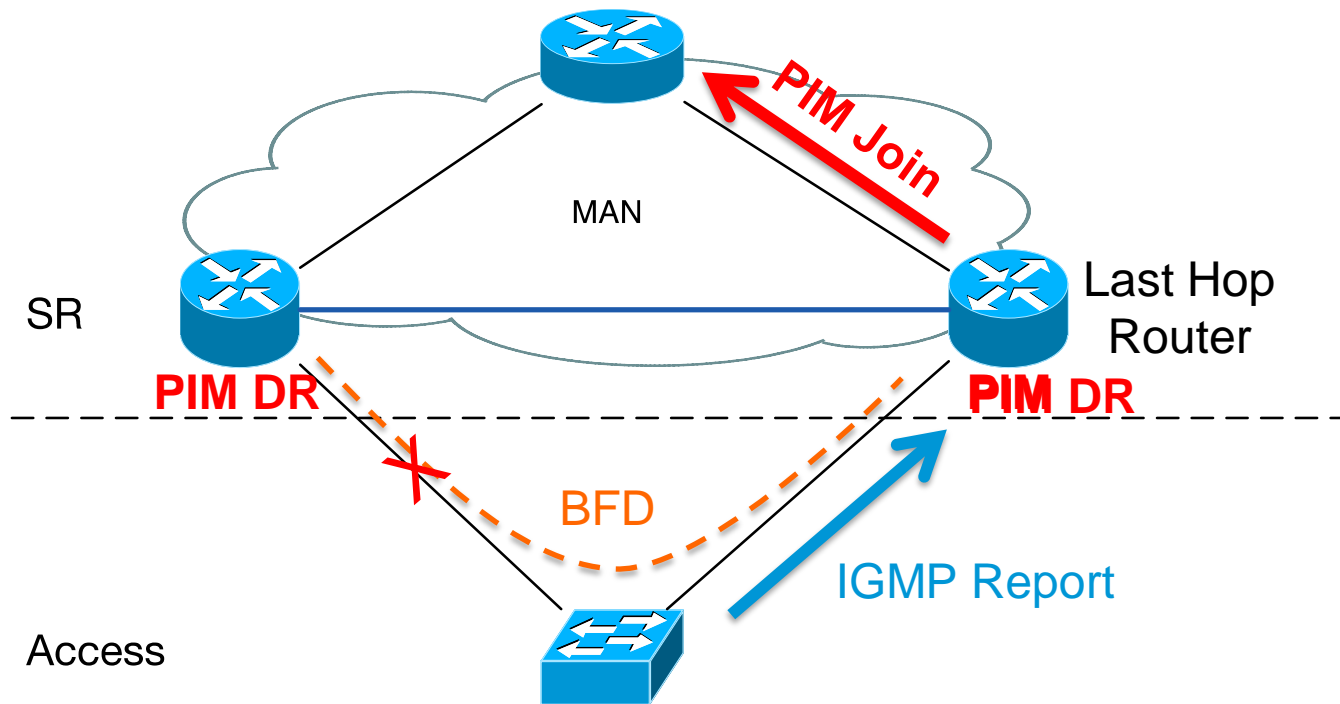
```
device(config)# interface fastethernet 1/6
```

```
device(config-if)# ip pim bfd
```



# Last Hop Router protection

If interface down or SR down, DR role will switchover after BFD session time out.



# Multicast High Availability

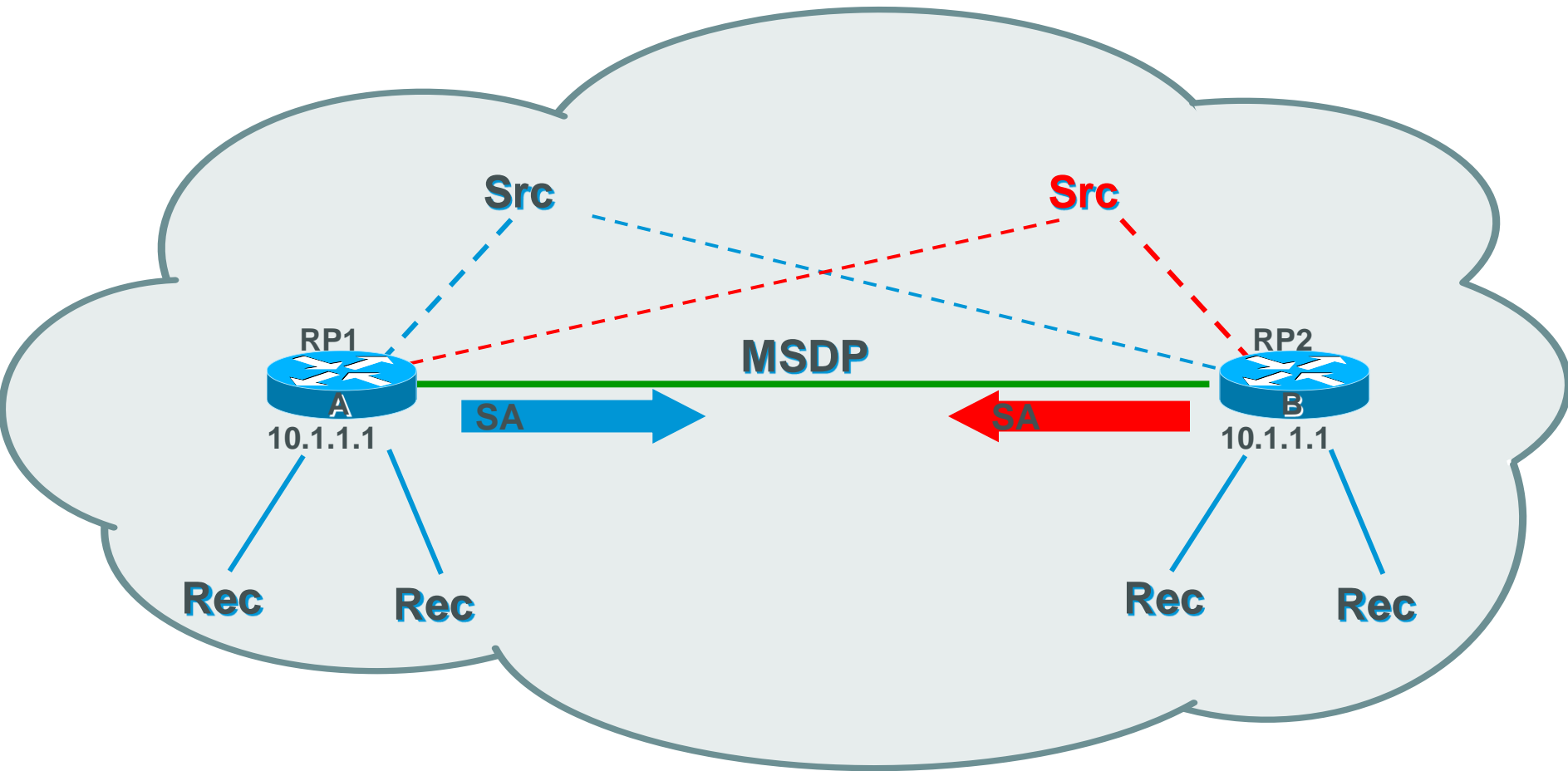
## RP protection - Anycast RP

- Within a domain, deploy more than one RP for the same group range
- Give each RP the same IP address assignment
- Sources and receivers use closest RP
- Use MSDP (Multicast Source Discovery Protocol) to communicate existence of Sources between RP' s.

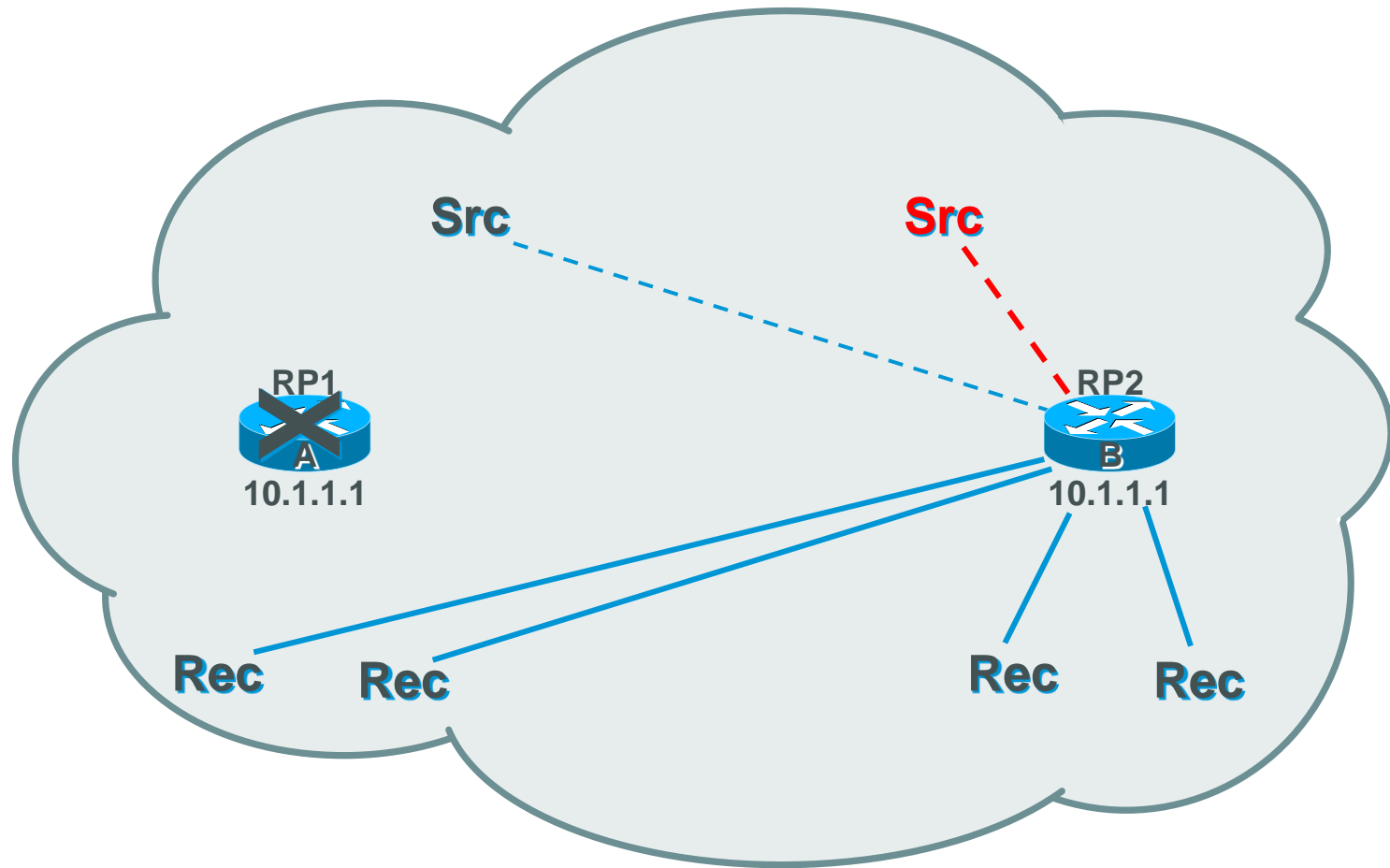
## MSDP :

- MSDP Peers (typically RP' s) are connected via TCP sessions.
- RP' s periodically originate Source Active (SA) messages for sources that are active in their local domain. These SA messages are sent to all active MSDP peers.

# RP protection – Anycast RP



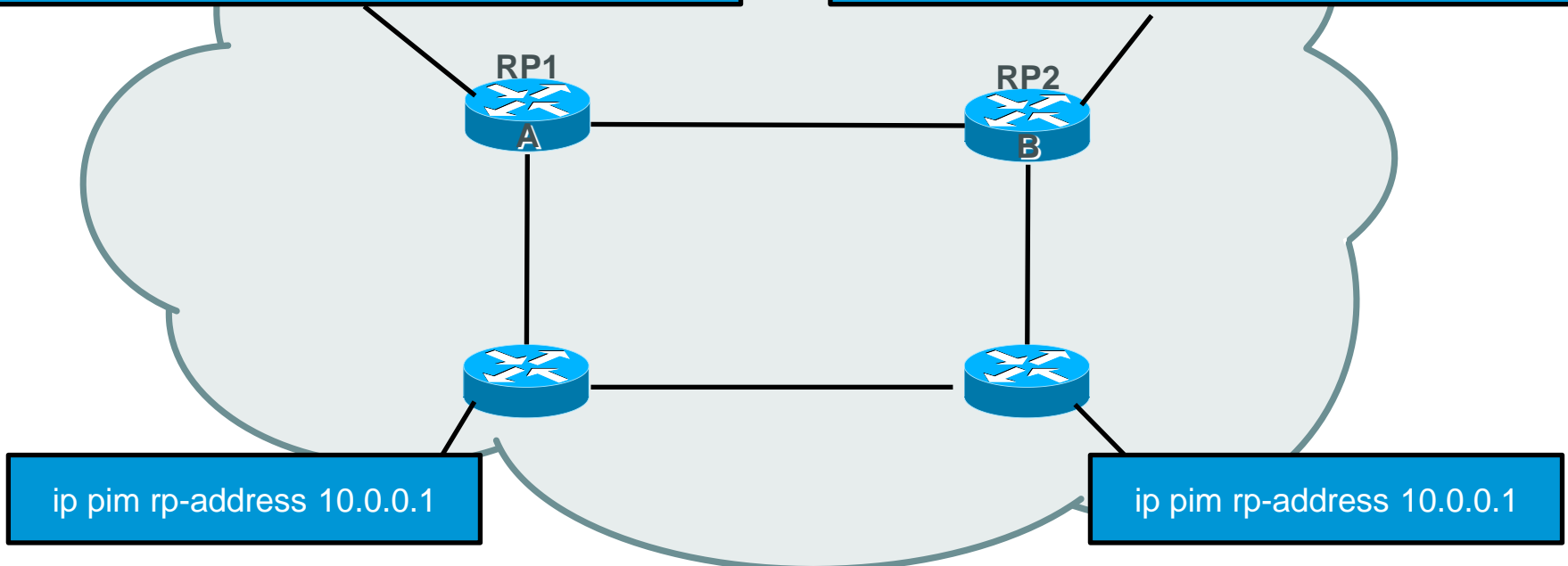
# Anycast RP—Overview



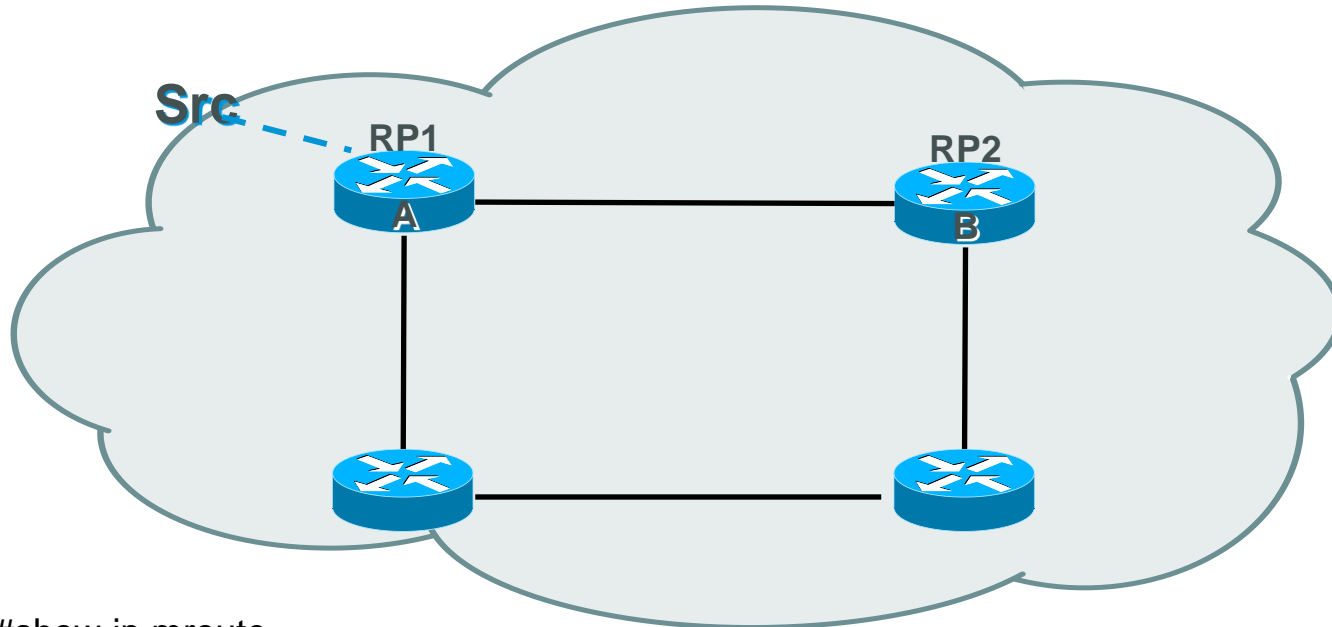
# Anycast RP—Overview

```
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
ip pim rp-address 10.0.0.1
```

```
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 10.0.0.3 connect-source loopback 0
ip msdp originator-id loopback 0
ip pim rp-address 10.0.0.1
```



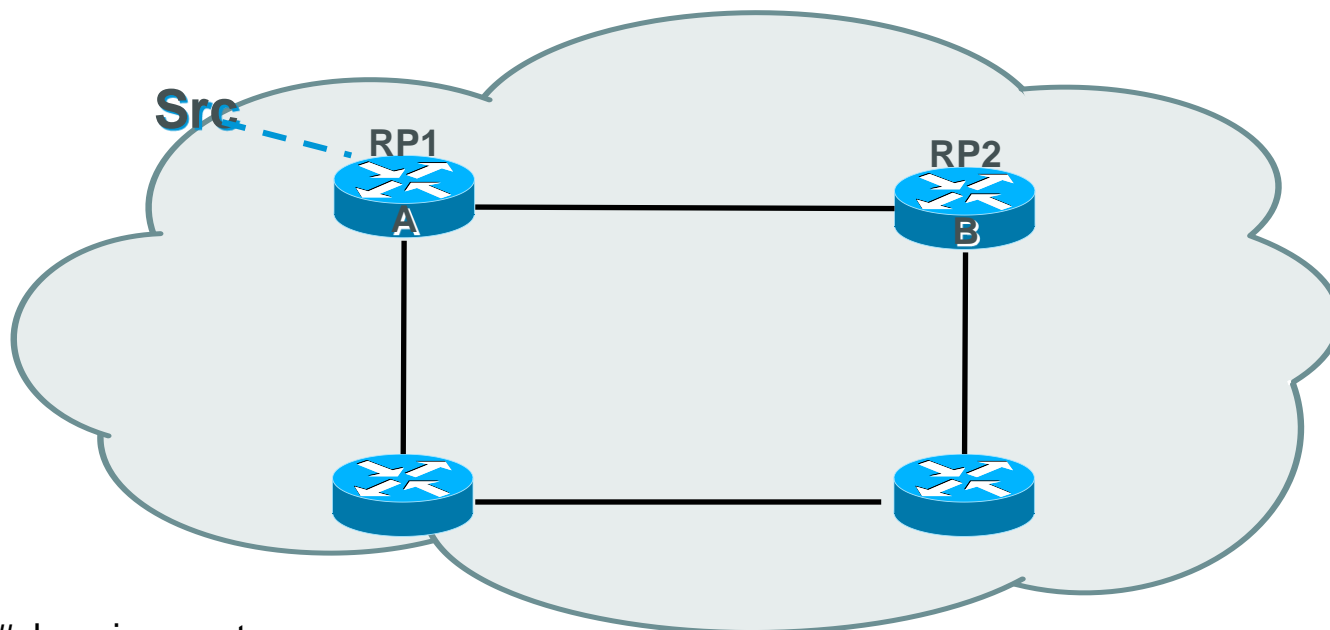
# Anycast RP—Overview



```
RP1#show ip mroute
(*, 239.1.1.1), 00:06:16/stopped, RP 10.0.0.1, flags: SJCL
  Incoming interface: Null, RPF nbr 0.0.0.0
  Outgoing interface list:
    FastEthernet2/0, Forward/Sparse, 00:06:16/00:02:46

(10.1.1.3, 239.1.1.1), 00:02:02/00:02:45, flags: LTA
  Incoming interface: FastEthernet0/0, RPF nbr 10.1.2.1
  Outgoing interface list:
    FastEthernet2/0, Forward/Sparse, 00:02:02/00:02:46
```

# Anycast RP—Overview

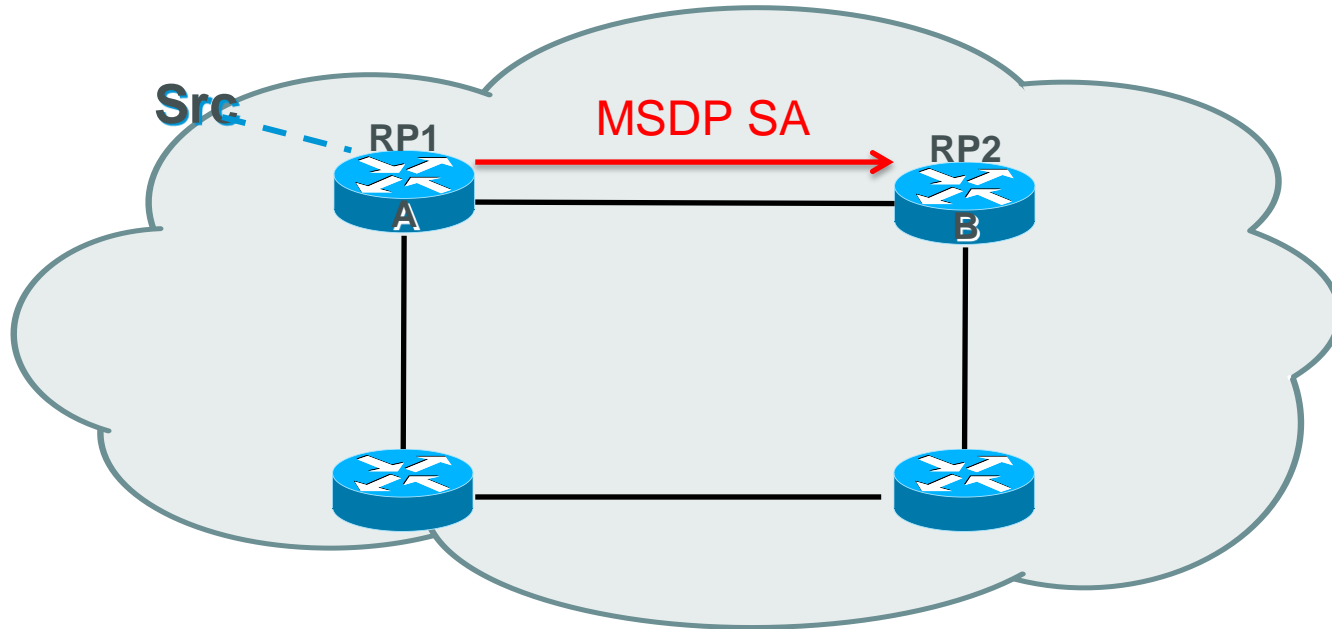


```
RP1#show ip mroute
(*, 239.1.1.1), 00:06:16/stopped, RP 10.0.0.1, flags: SJCL
Incoming interface: Null, RPF nbr 0.0.0.0
Outgoing interface list:
  FastEthernet2/0, Forward/Sparse, 00:06:16/00:02:46

(10.1.1.3, 239.1.1.1), 00:02:02/00:02:45, flags: LTA
Incoming interface: FastEthernet0/0, RPF nbr 10.1.2.1
Outgoing interface list:
  FastEthernet2/0, Forward/Sparse, 00:02:02/00:02:46
```

**FLAG:**  
L : Local  
T : SPT-bit  
A : Candidate for MSDP  
Advertisement

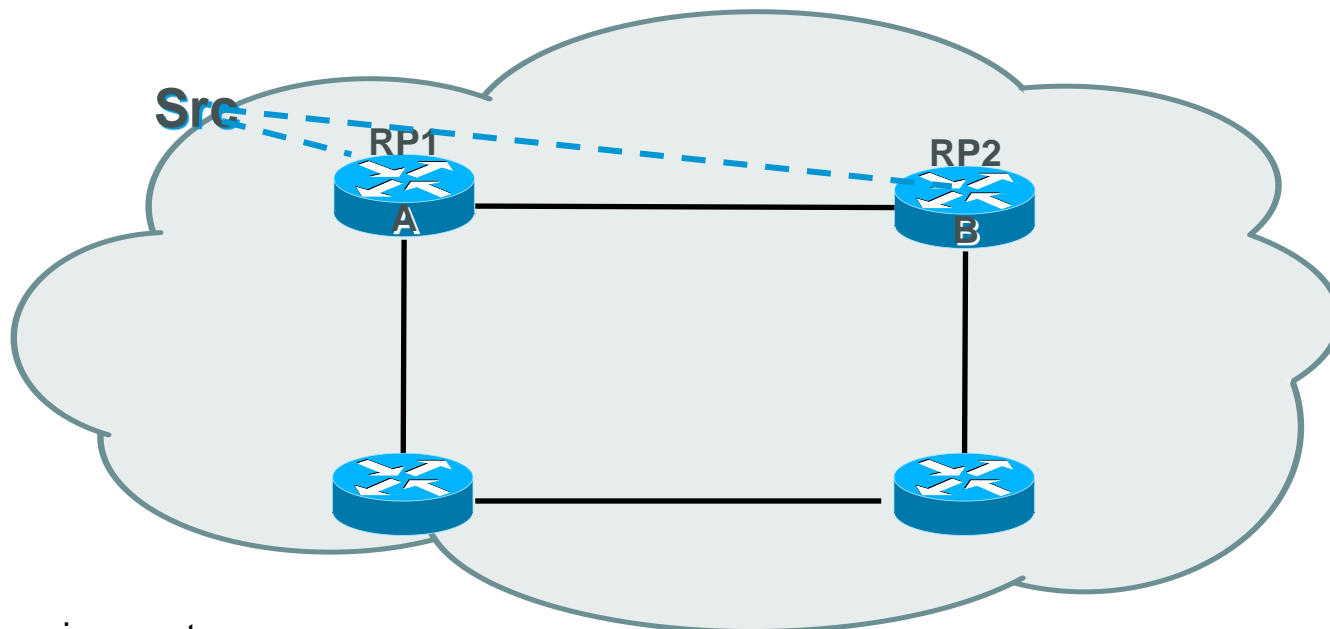
# Anycast RP—Overview



```
RP2#show ip msdp sa-cache  
MSDP Source-Active Cache - 1 entries  
(10.1.1.3, 239.1.1.1), RP 10.0.0.1, AS ?,00:13:07/00:05:04, Peer 10.1.4.1
```



# Anycast RP—Overview



```
RP2#show ip mroute
(*, 239.1.1.1), 00:15:33/stopped, RP 5.5.5.5, flags: SJCL
Incoming interface: Null, RPF nbr 0.0.0.0
Outgoing interface list:
  Loopback1, Forward/Sparse, 00:15:33/00:02:45
```

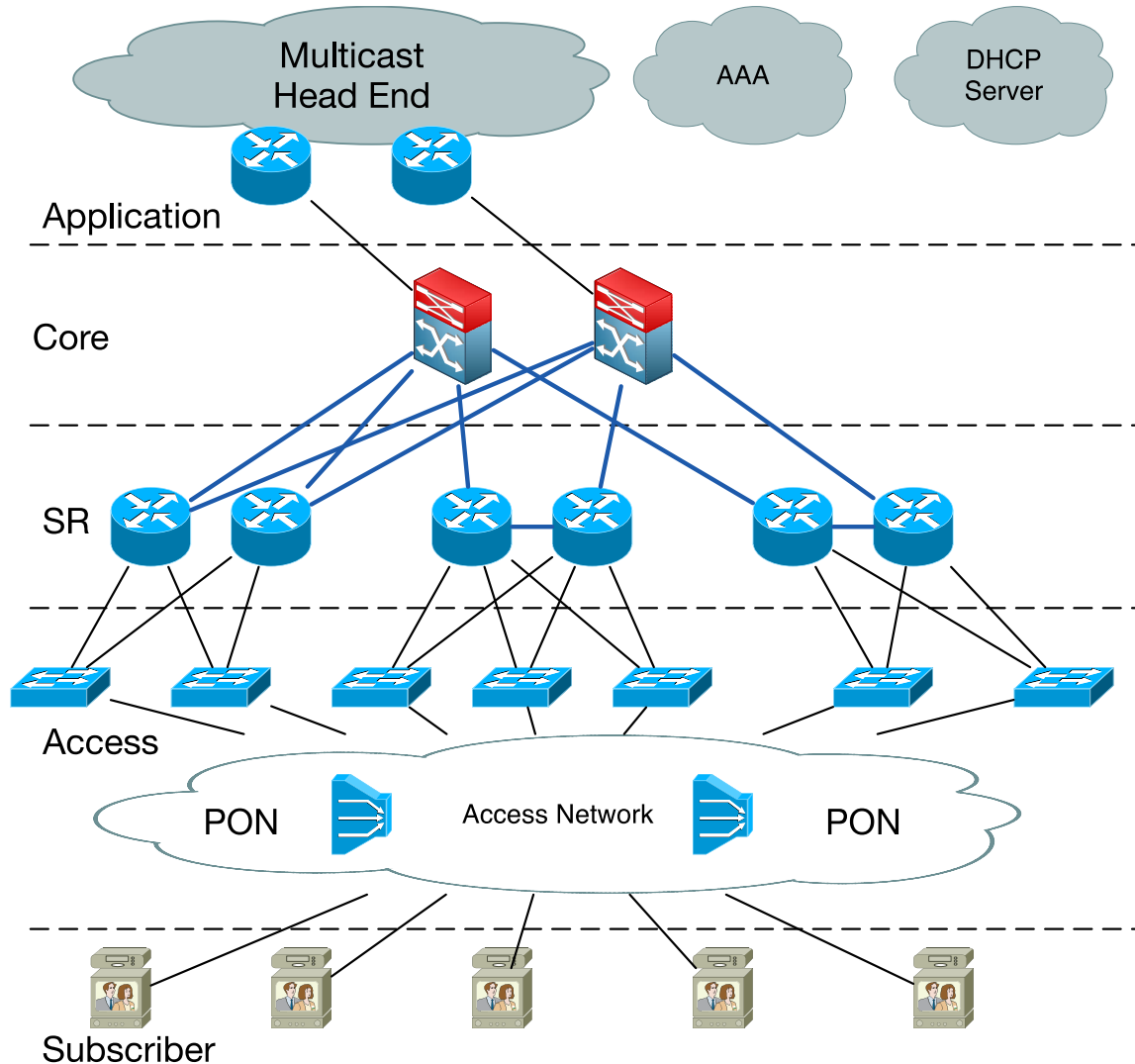
```
(10.1.1.3, 239.1.1.1), 00:03:22/00:02:57, flags: LMT
Incoming interface: FastEthernet0/0, RPF nbr 10.1.3.1
Outgoing interface list:
  Loopback1, Forward/Sparse, 00:03:22/00:02:45
```

FLAG:  
M: MSDP created entry

Thank you.



# Multicast Service on SP



## IPTV Service Flow

1. Set of box(STB) power on. Send out DHCP discovery
2. DHCP relay occur on SR, unicast to DHCP server.
3. After exchange, STB get IP address
4. Establish session to EPG, authentication and download channel list.
5. STB send out IGMP report message.

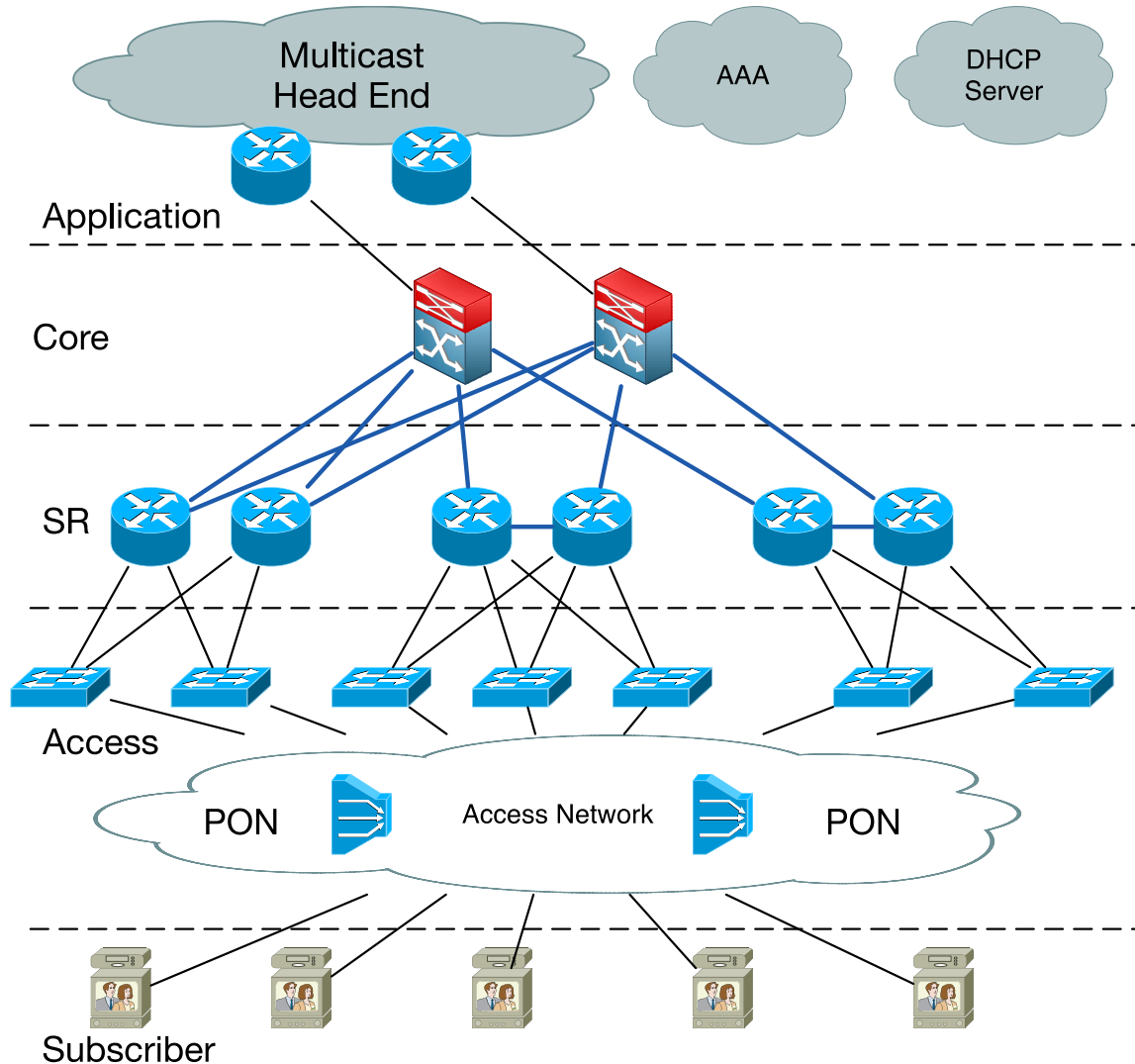
# Multicast Service on SP

## IPTV Service FLOW

6. On PON or DSLAM will implement IGMP Proxy or Multicast vlan replication.  
Propagate IGMP packet to access switch.

7. On access switch IGMP snooping is enable to control multicast flooding.  
after recording the outgoing interface  
Then IGMP packet send out to SR router.

8. ON SR router, after receiver the IGMP report packet. SR will recreate (\*,G) entry , and send out (\*,G) PIM Join packet



# Multicast Service on SP

## IPTV Service FLOW

9. The (\*,G) Join will arrive RP, RP forwarding G group traffic to this outgoing interface.

10. After SR receive the first G group multicast traffic will switchover to SPT. Send out (S,G) Join.

