IP Multicasting

SWITCH Module 9

Agenda

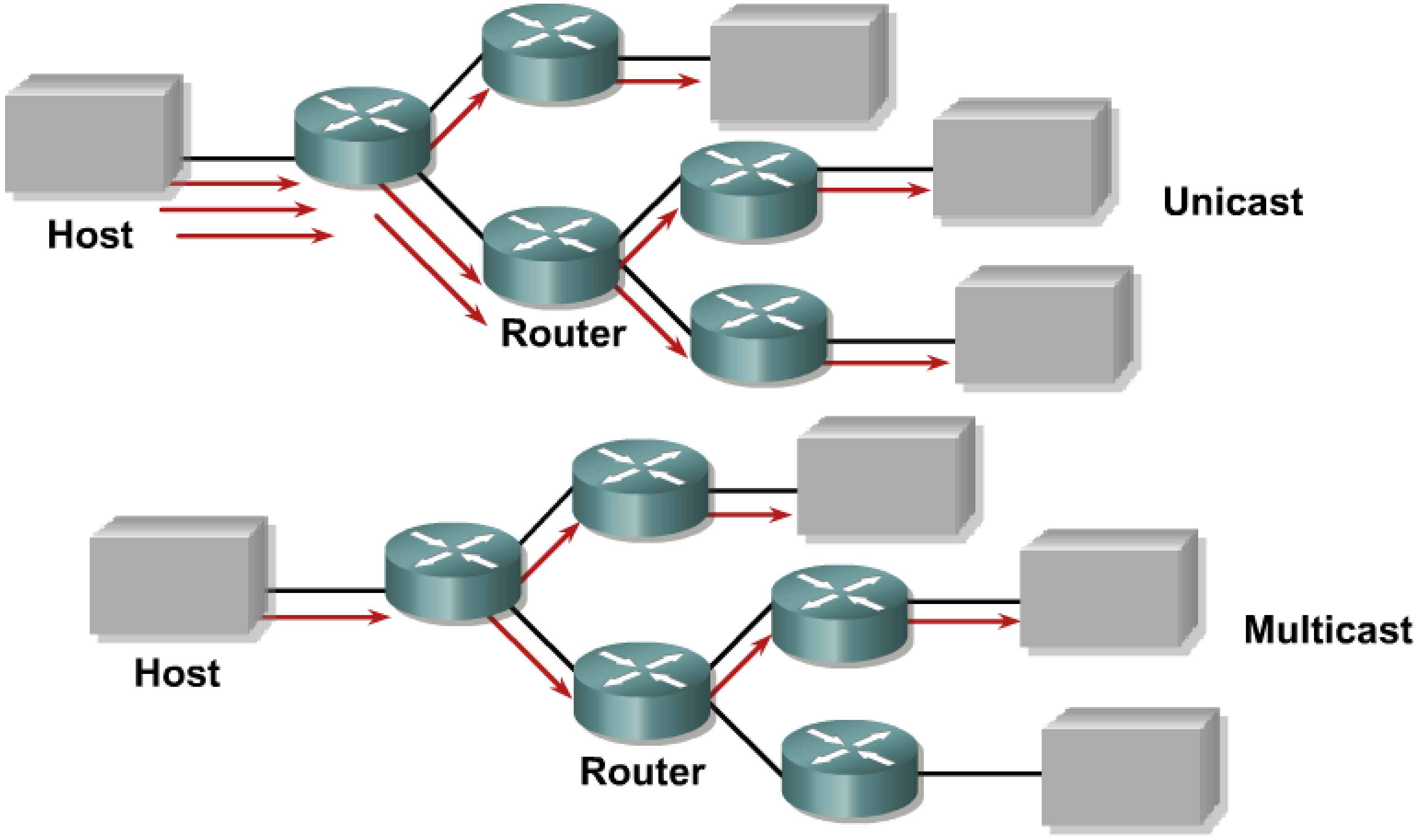
- **Motivation**
- **Addressing**
- **IGMP/MLD**
- **L2 multicast efficient delivery**
- **Protocol Independent Multicast**
  - PIM-DM
  - PIM-SM
- **Mcast Routing Configuration**
- **Verifying and Troubleshooting**
- **Other PIM Modes**
- **VLC Player**

Why Multicast Does Exist?

- Many network applications require to receive multiple data flows simultaneously
  - Internet version of radio/television broadcasts, conference calls
  - Music-on-hold in IP telephony
  - Distribution of information to many (potentially unknown) receivers

- **Multicasting** = sending of one frame/packet that will be delivered to multiple receivers simultaneously
- *Advantages*
  - Better utilization of network resources (effective use of bandwidth)
  - Sender does not have to know identity (address) of each receiver

Unicast vs. Multicast



Disadvantages of Multicast

- Transport protocol for multicast is usually UDP, hence…
  - …it’s best-effort delivery (no packet lost correction)
  - …no congestion control
  - …duplicate packets could be received
  - …no packet ordering
  - …filtering and securing of multicast transfers is more complicated
  - Some of above issues are solved with proposed protocols like PGM
- Routers SHOULD support multicast routing in order to deliver multicast to receivers in different IP networks
- Switches SHOULD support multicast switching in order to effectively deliver multicast only to relevant receivers

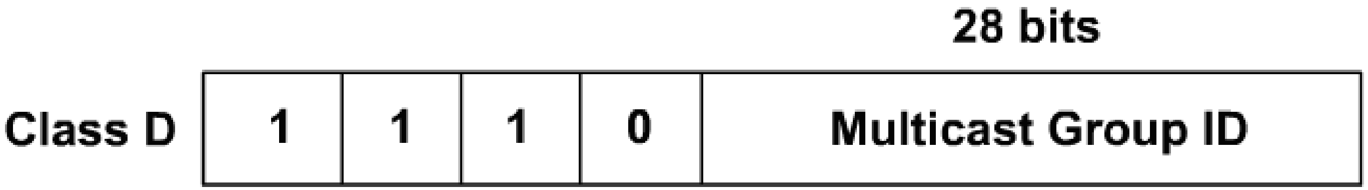
Multicast Application Types

- **One-to-many**
  - One sender, two or more receivers
- **Many-to-many**
  - Any number of senders and/or receivers transmitting data between each other – all of them are members of same multicast group
- **Many-to-one**
  - Any number of senders sending traffic to one receiver
  - Could be also unicast

Addressing

Structure of IPv4 Address

- Class D is reserved for purposes of IPv4 multicast
  - Highest 4 bits are set to 1110 (class D prefix)
  - Rest 28 bits identifies multicast group



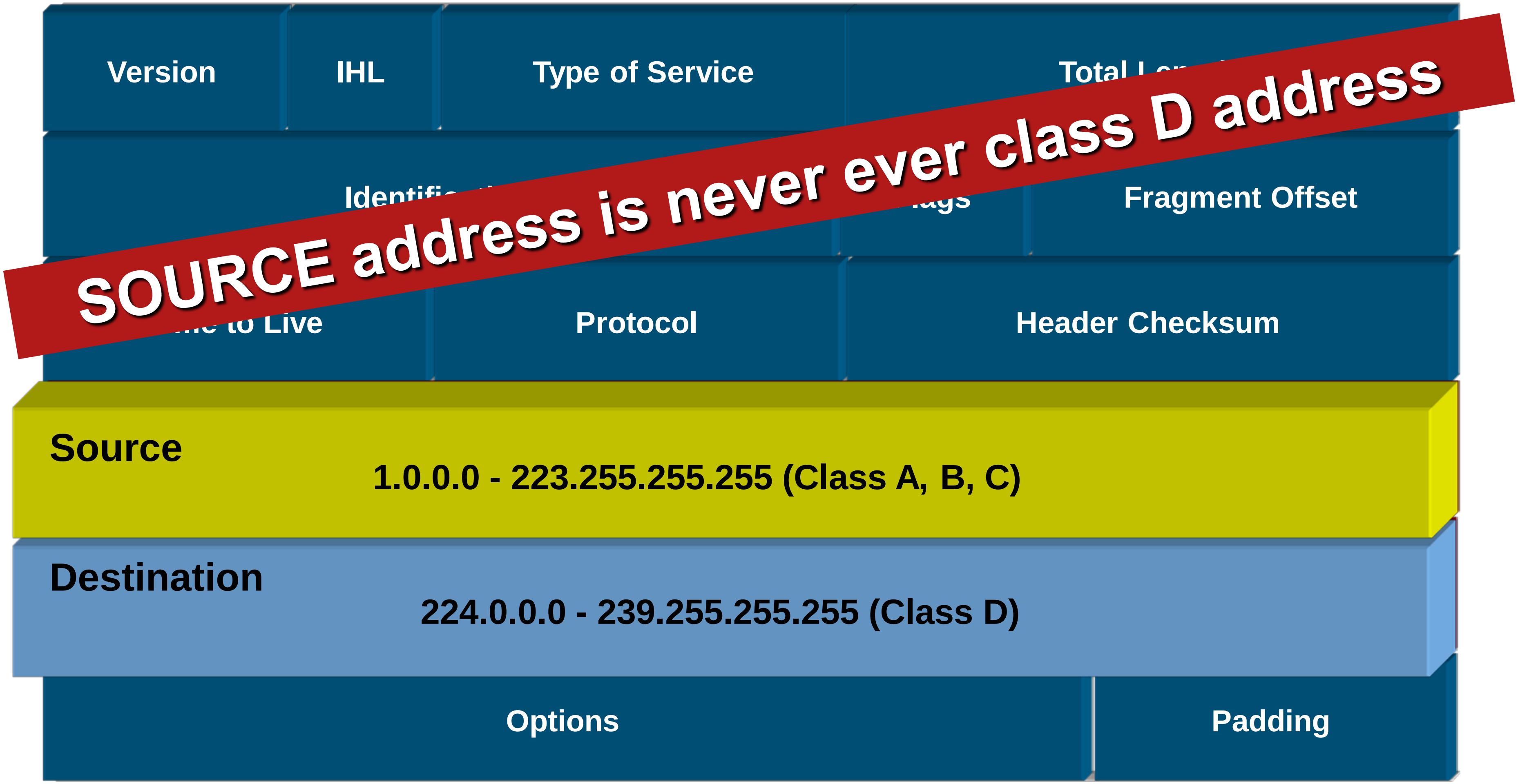
28 bits

Class D | 1 | 1 | 1 | 0 | Multicast Group ID

  - (Multicast) Group consists of members (hosts that declare to be member of this multicast group)
- Class D range is from 224.0.0.0 to 239.255.255.255

Addressing of IPv4 Multicast

IPv4 Header



SOURCE address is never ever class D address

Version

IHL

Type of Service

Total L

Identifi

lags

Fragment Offset

me to Live

Protocol

Header Checksum

Source

1.0.0.0 - 223.255.255.255 (Class A, B, C)

Destination

224.0.0.0 - 239.255.255.255 (Class D)

Options

Padding

Reach of Addresses ①

- **Local scope**
  - 224.0.0.0 - 224.0.0.255
  - Multicast packets with this address don’t expect to leave one broadcast domain – they are link-local
  - Many of addresses from this range are reserved for core network services
- **Global scope**
  - 224.0.1.0 - 238.255.255.255
  - World wide scope
  - *In reality ISPs usually don’t route multicast if they are not forced to* ☹
- **Administratively scoped**
  - 239.0.0.0 - 239.255.255.255
  - Reserved for use in private domains

Reach of Addresses ②

- Some important **Local Scoped** addresses are:
  - 224.0.0.1 (all multicast-capable systems on segment – nowadays de facto broadcast)
  - 224.0.0.2 (all routers on segment)
  - 224.0.0.4 (all DVMRP routers)
  - 224.0.0.5, 224.0.0.6 (all OSPF routers, all DR/BDR routers)
  - 224.0.0.9 (all RIPv2 routers)
  - 224.0.0.10 (all EIGRP routers)
  - 224.0.0.13 (all PIMv2 routers)
  - 224.0.0.18 (all VRRP gateways)
  - 224.0.0.22 (all IGMPv3 clients)
  - 224.0.0.2,102 (all HSRP gateways)

Reach of Addresses ③

- **Global Scope** addresses are not strictly reserved, they are allocated dynamically
  - 224.2.X.X was used in MBONE applications
  - Some parts of global scope are used for new protocols
    - 224.1.0.0-224.1.255.255 ST Multicast Groups
    - 224.2.0.0-224.2.127.253 Multimedia Conference Calls
    - 224.2.127.254 SAPv1 Announcements
    - 224.2.127.255 SAPv0 Announcements (deprecated)
    - 224.2.128.0-224.2.255.255 SAP Dynamic Assignments
    - 224.252.0.0-224.255.255.255 DIS transient groups
    - 232.0.0.0-232.255.255.255 VMTP transient groups
- **Administratively Scoped** has same analogy as private IPv4 addresses
  - Organization-local: 239.0.0.0/8 (RFC 2365)

L2 Addressing ①

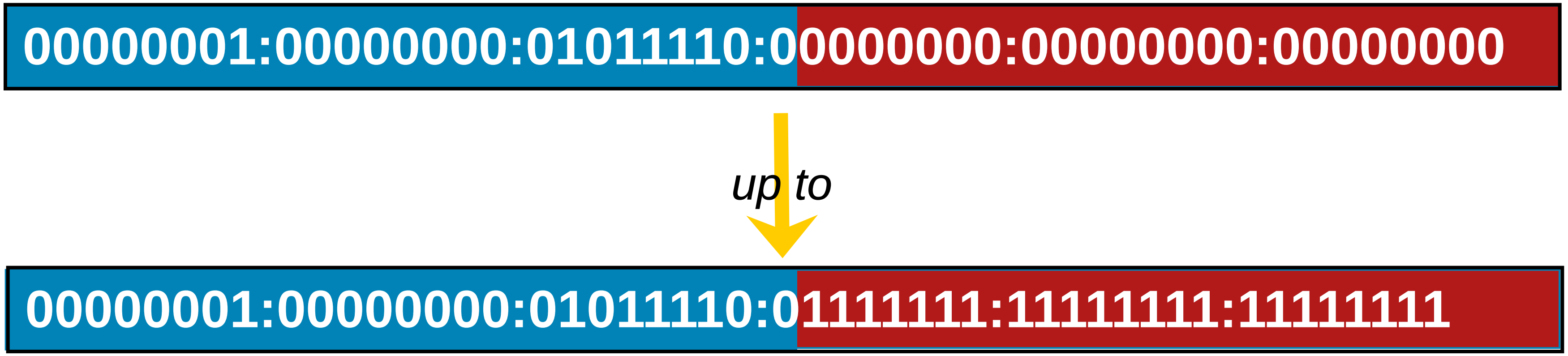
- Until now MAC address points to one target interface
- In reality there are MAC addresses that could point multiple interfaces in one broadcast domain
- MAC address (6B) = OUI (3B) + S/N (3B)
- Format of 1st byte of MAC address:

Bit	7	6	5	4	3	2	1	0
Meaning	n	n	n	n	n	n	U/L	I/G

  - **U/L** = Universal (0) / Local (1)
  - **I/G** = Individual (0) / Global (1)

L2 Addressing ②

- IANA reserved for IPv4 multicast range of MAC addresses from **01:00:5e:00:00:00** to **01:00:5e:7f:ff:ff**



00000001:00000000:01011110:00000000:00000000:00000000

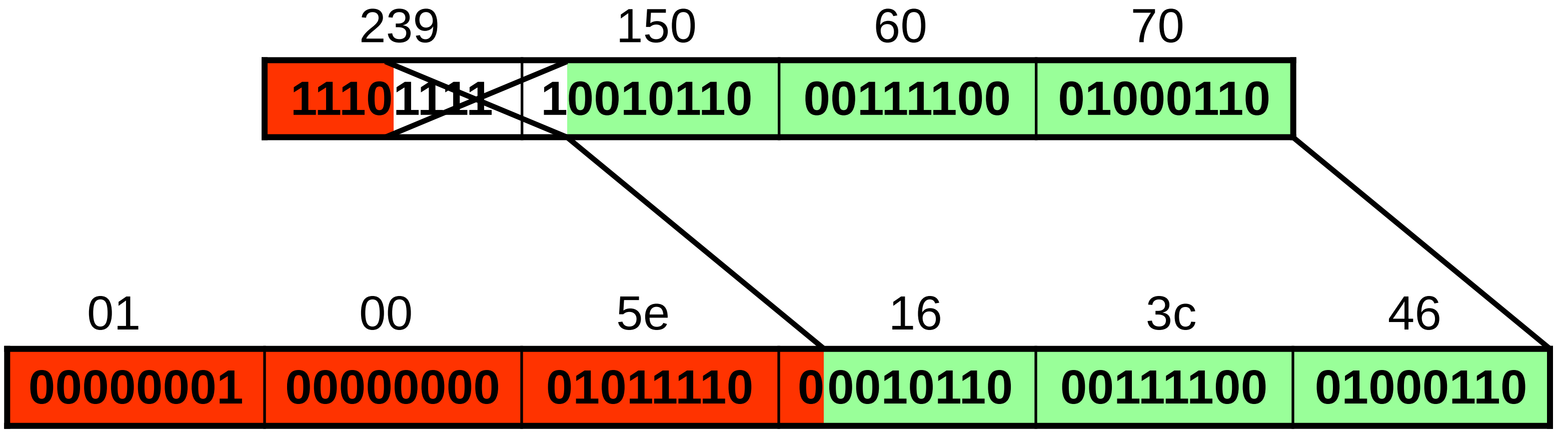
up to

00000001:00000000:01011110:01111111:11111111:11111111

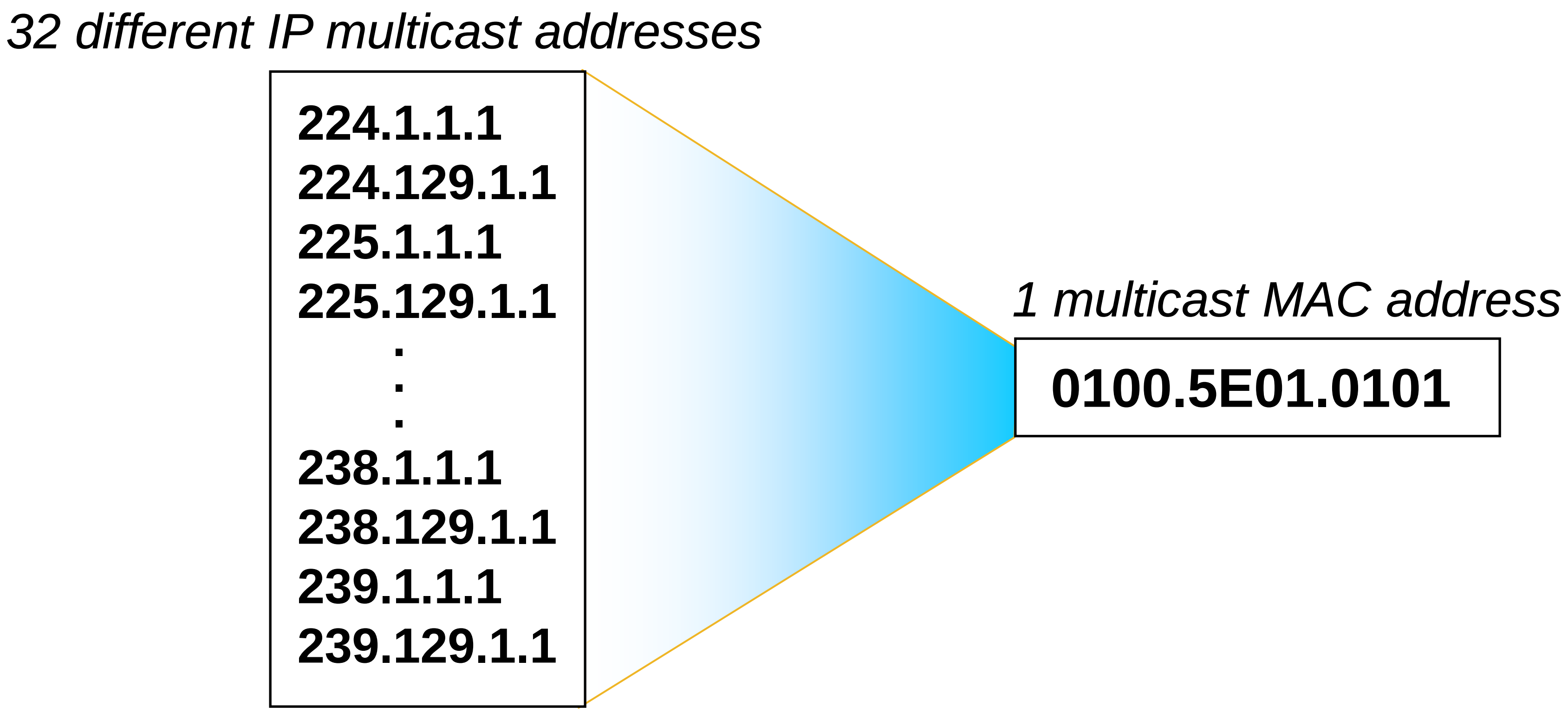
- 1st 25 bits have fixed value
- The remaining 23 bits identifies multicast group

IP to MAC Mapping ①

- *Ordinary IP addresses are mapped to MAC addresses with help of ARP- but this doesn’t apply to class D addresses!*
- Instead of this is used other class D to MAC transformation which is unfortunately not bijective operation



IP to MAC Mapping ②



32 different 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

0100.5E01.0101

- So called **32-to-1 overlapping problem**

Allocating/Finding out Multicast Address

- **Finding out**
  - **Session announcement protocol (SAP)** in RFC 2974, Cisco calls it sometimes SDR
- **Static allocation and finding out**
  - *There are so soooo many rules, exceptions and recommendations which multicast addresses DO NOT use!*
  - e.g. „Guidelines for Enterprise IP Multicast Address Allocation“
- **Allocating according to AS**
  - 233.0.0.0 – 233.255.255.255
  - a.k.a **GLOP addressing** (RFC 3180)
  - 1st byte of IP must be set to value 233
  - 2nd and 3rd byte of IP is ASN
  - 4th byte is multicast group identifier

IGMP

Internet Group Management Protocol (IGMP)

- Whenever host wants to be member of group it announce its request to gateway router
  - `IF` router receives traffic intended to target multicast group `THEN` router forwards it to the host
  - Host is NOT assigned with IP address, instead it initializes support to receive frames with target multicast MAC address
- Protocol to support signing on and off for IPv4 multicast traffic is called **Internet Group Management Protocl (IGMP)**
  - IGMP communication happens between the host and its gateway
- Currently there exist three versions
  - IGMPv1 in RFC 1112
  - IGMPv2 in RFC 2236
  - IGMPv3 in RFC 3376

IGMPv1

- IGMPv1 has two basic messages
  - **IGMPv1 Membership Query**
    - Periodically generated by routers (a.k.a. queriers or query routers) and send to address 224.0.0.1 (all-hosts)
    - Default send interval is 1 minute
  - **IGMPv1 Membership Report**
    - Host sends this messages to destination address as target IPv4 multicast group address to sign on
    - 1 Membership Report is generated for every host member group
    - Report is sent either as solicited (as a reply to Membership Query) or as unsolicited message (whenever host is newly trying to sign on to target multicast group)
    - Every host is waiting for a random period of time (max. 10 seconds) to hear a report from the other host to abstain from generating own report

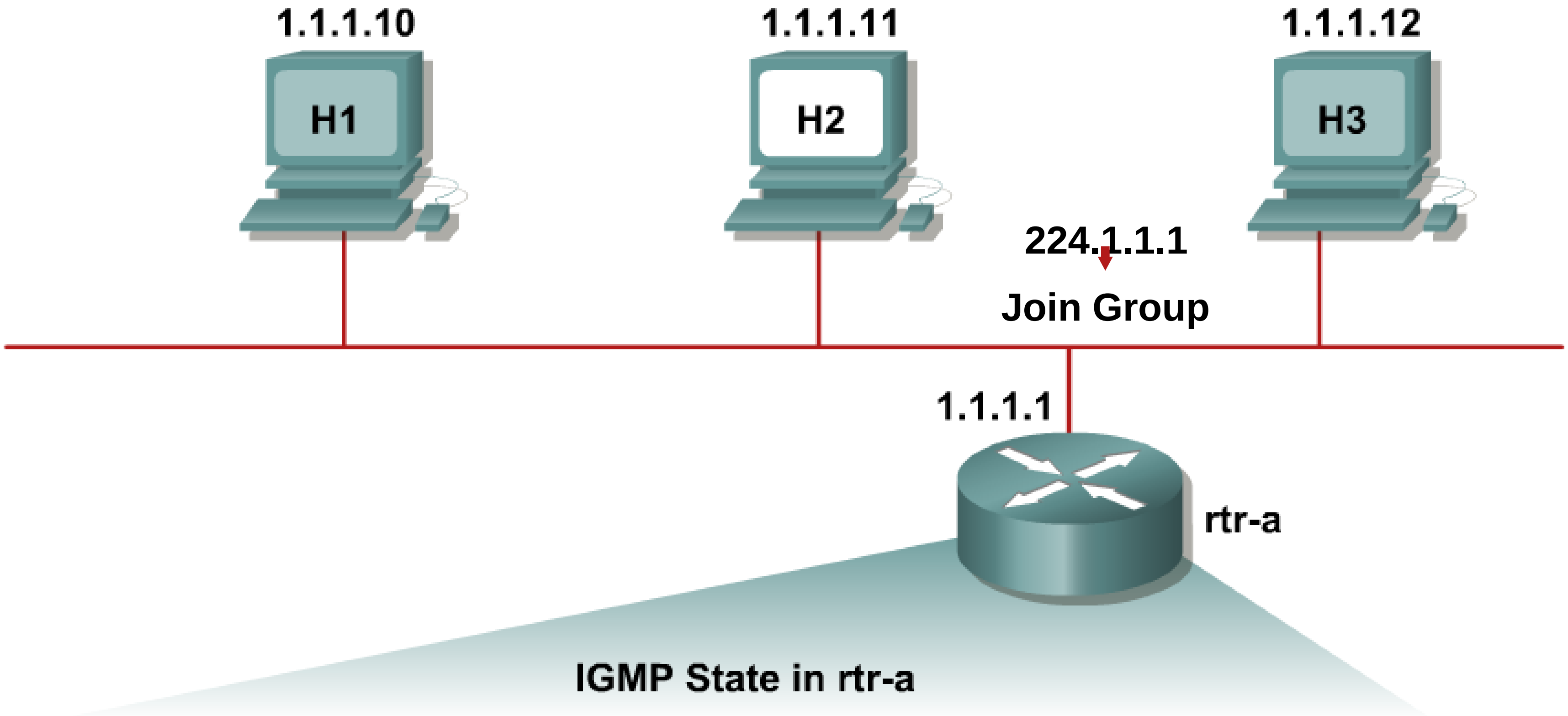
IGMPv2

- IGMPv2 has three basic message types
  - **IGMPv2 Membership Query**
    - Periodically generated by queriers
    - They could be General (sent to 224.0.0.1 every 125 seconds) or Group-specific (sent to target multicast group address)
  - **IGMPv2 Membership Report**
    - Sent to target multicast group address to which host is trying sign on – similar to IGMPv1
  - **IGMPv2 Leave Group**
    - Host announces that it wants to leave target multicast group
    - Message is sent to all routers
    - `IF` response to the last Group-Specific Query was sent by someone else than host which is signing off `THEN` there is no need to send this message
- Query router could enforce maximum time period for waiting to reply on Membership Query (so called **Max Query-Response time**)
- IGMPv2 standardizes who will be elected as **IGMP Querier** on segment – the router with lowest IP address

IGMPv3

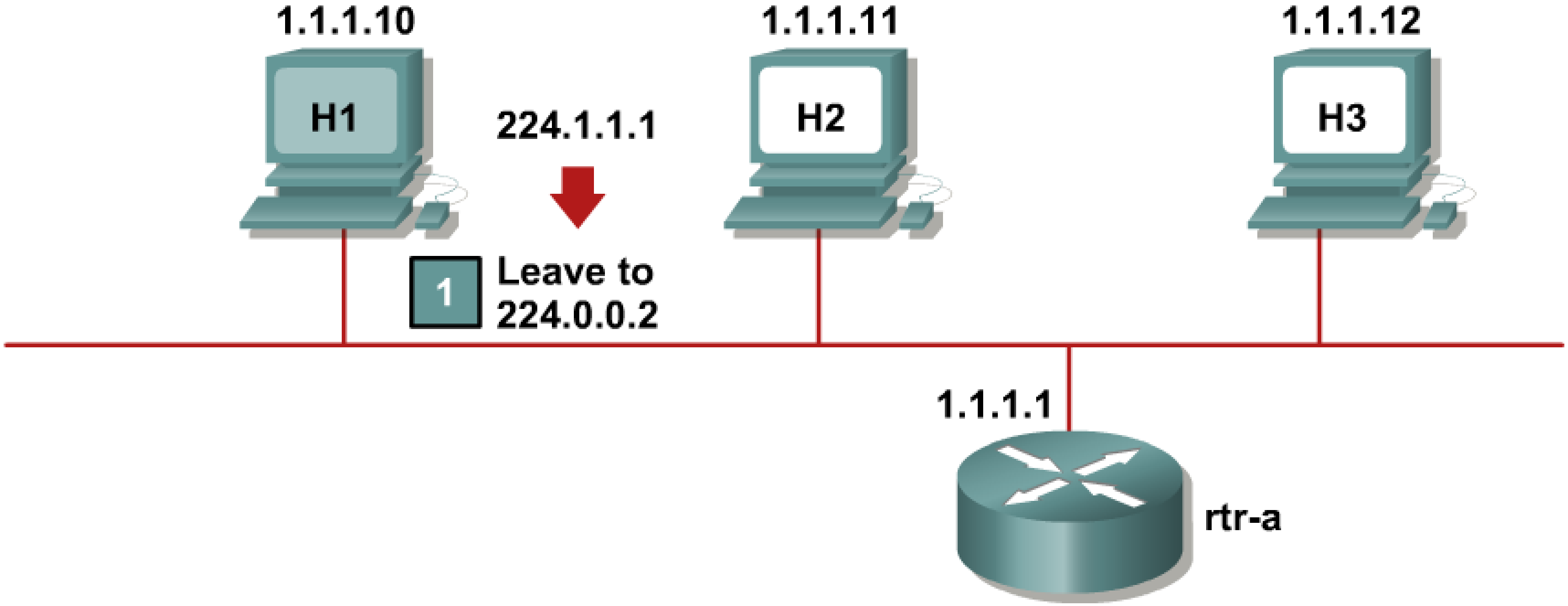
- IGMPv3 has once again two basic messages:
  - **IGMPv3 Membership Query**
    - General query sent to 224.0.0.1
    - Group-specific query (*,G) sent to target multicast group
    - Group-and-source specific (S,G) sent to target multicast group
  - **IGMPv3 Membership Report**
    - Sent to 224.0.0.22 (all IGMPv3 enable devices)
- IGMPv3 MR has much more complicated syntax
  - Include filter = *I want to receive multicast from all sources in the list*
  - Exclude filter = *I want to receive multicast from all sources expect the ones that are present in the list*
  - *How to use filters to simulate Leave Group message and Group-Specific Query?*

IGMPv2: Signing on to Multicast Group

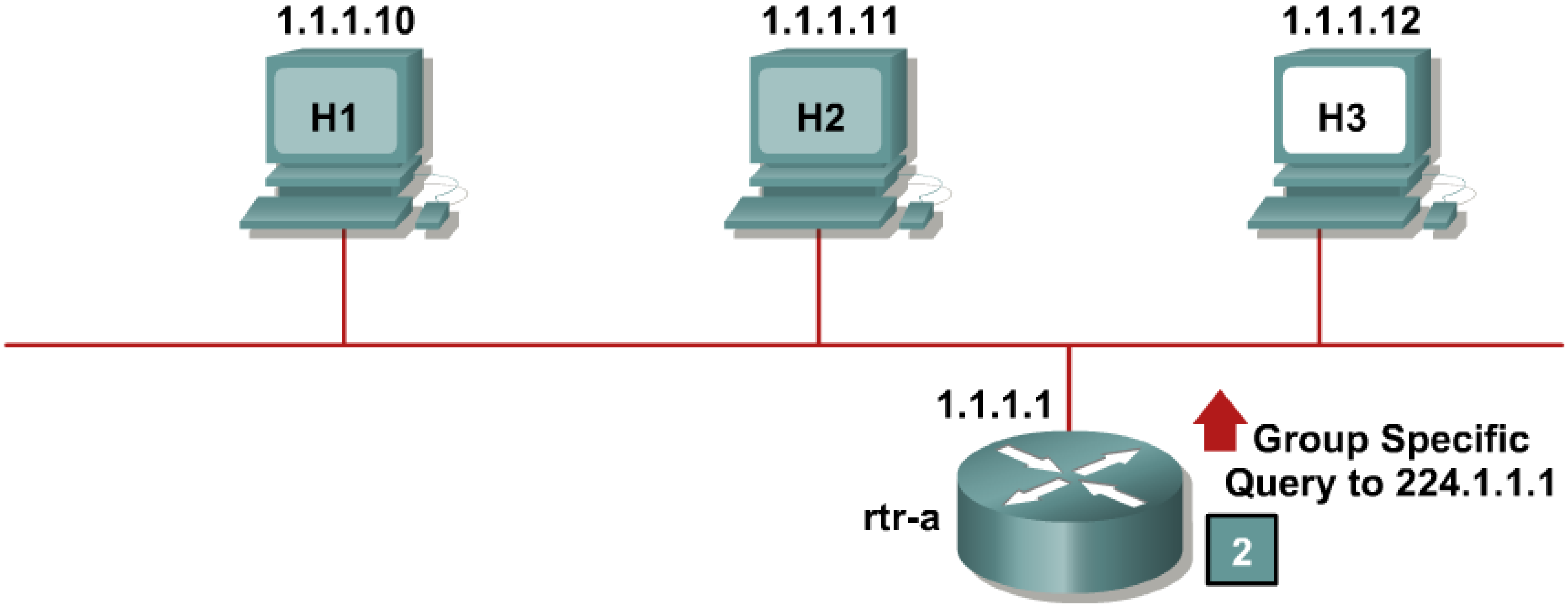


```
rtr-a>show ip igmp group
IGMP Connected Group Membership
Group Address    Interface   Uptime   Expires    Last Reporter
224.1.1.1        Ethernet0   0d1h3m   00:02:31   1.1.1.11
```

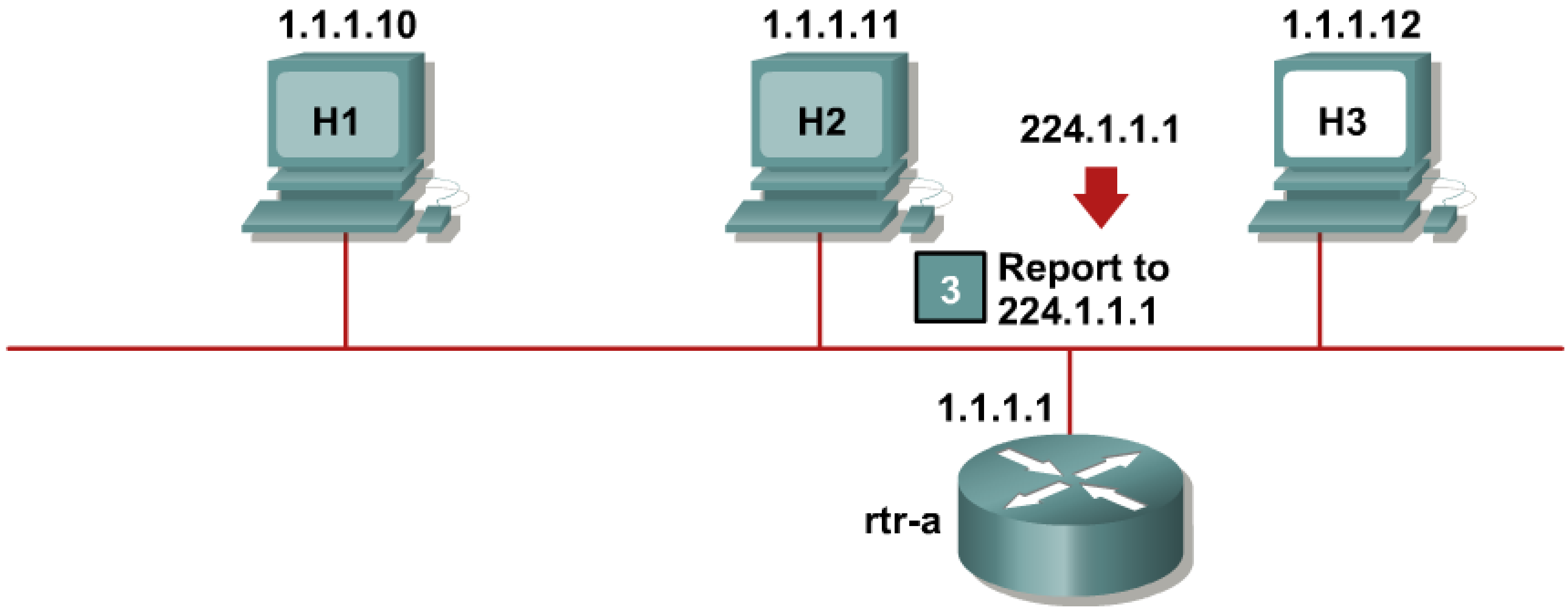
IGMPv2: Leaving the Multicast Group ①



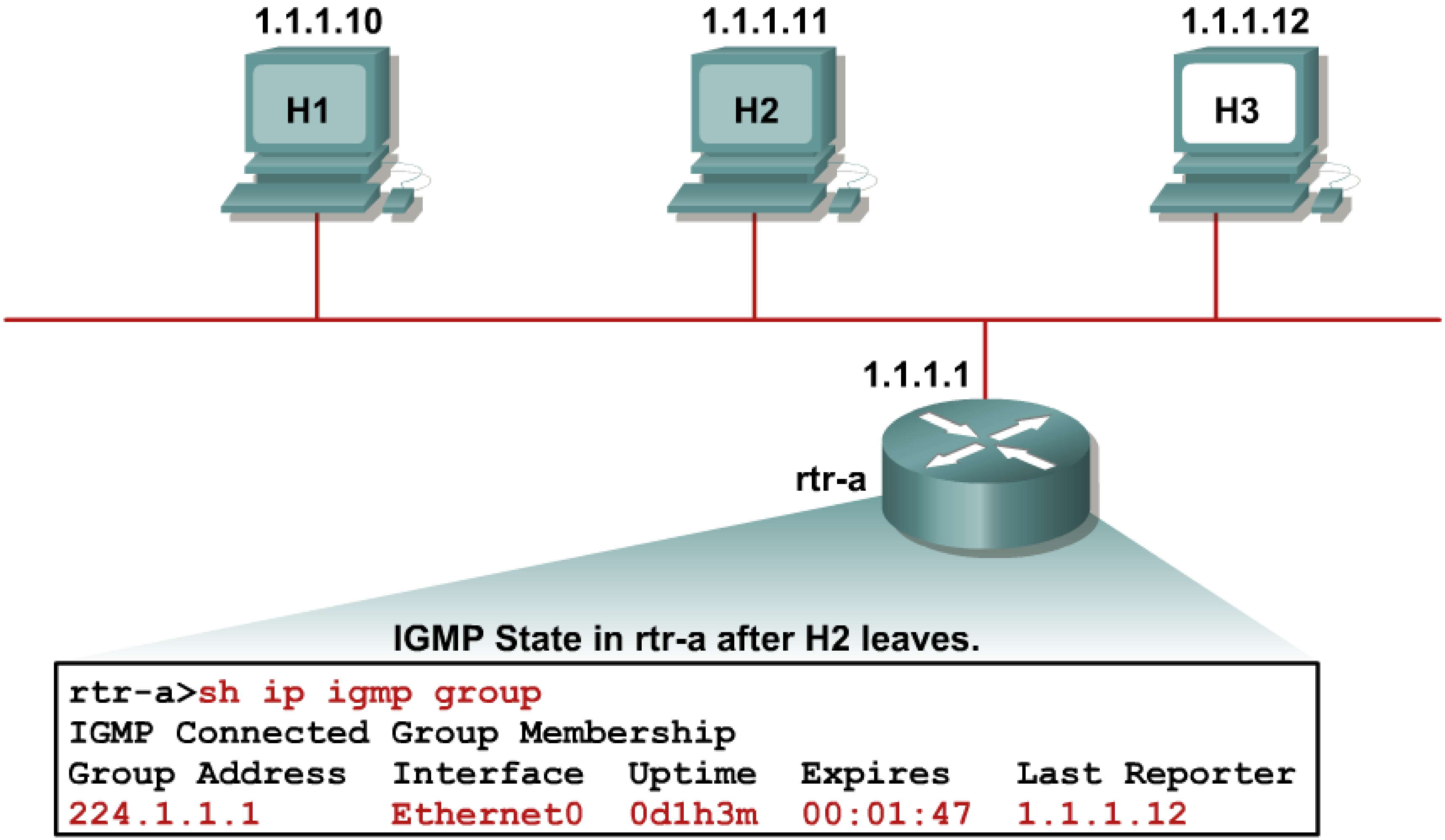
IGMPv2: Leaving the Multicast Group ②



IGMPv2: Leaving the Multicast Group ③



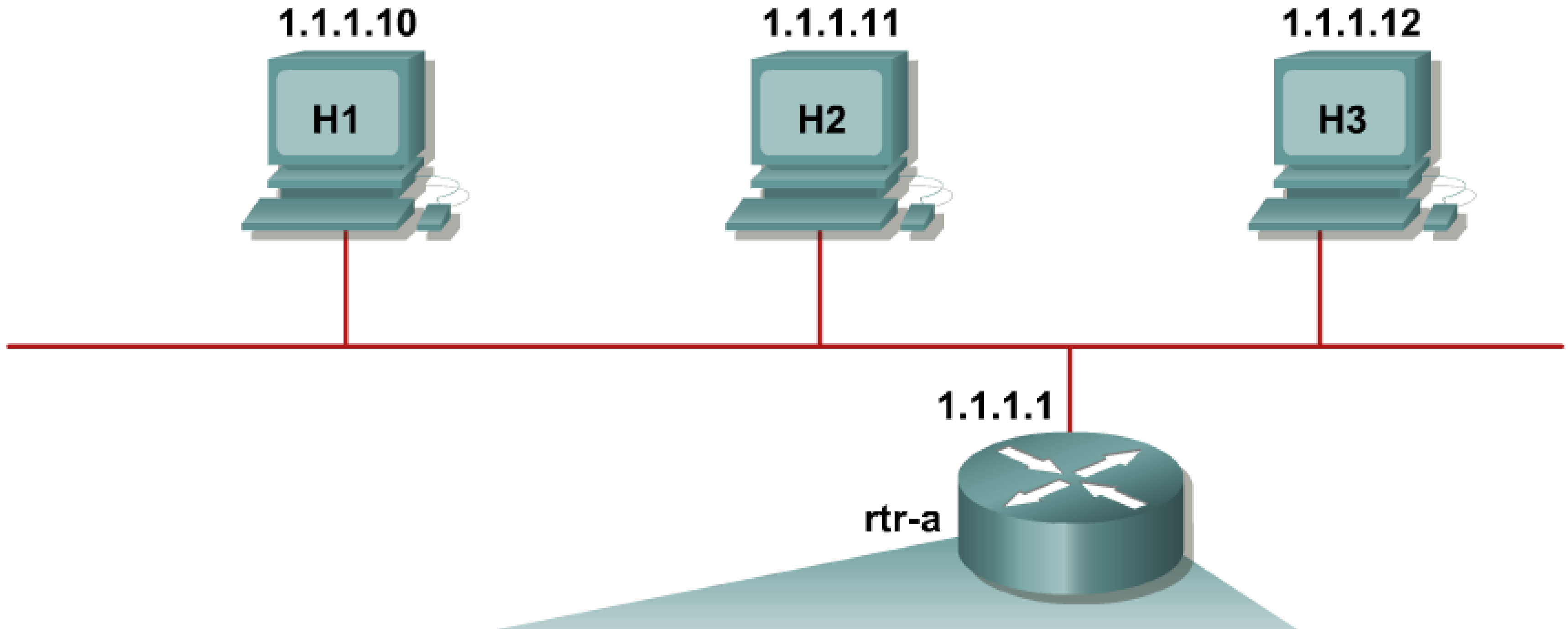
IGMPv2: Leaving the Multicast Group ④



IGMP State in rtr-a after H2 leaves.

```
rtr-a>sh ip igmp group
IGMP Connected Group Membership
Group Address    Interface   Uptime   Expires    Last Reporter
224.1.1.1        Ethernet0   0d1h3m   00:01:47   1.1.1.12
```

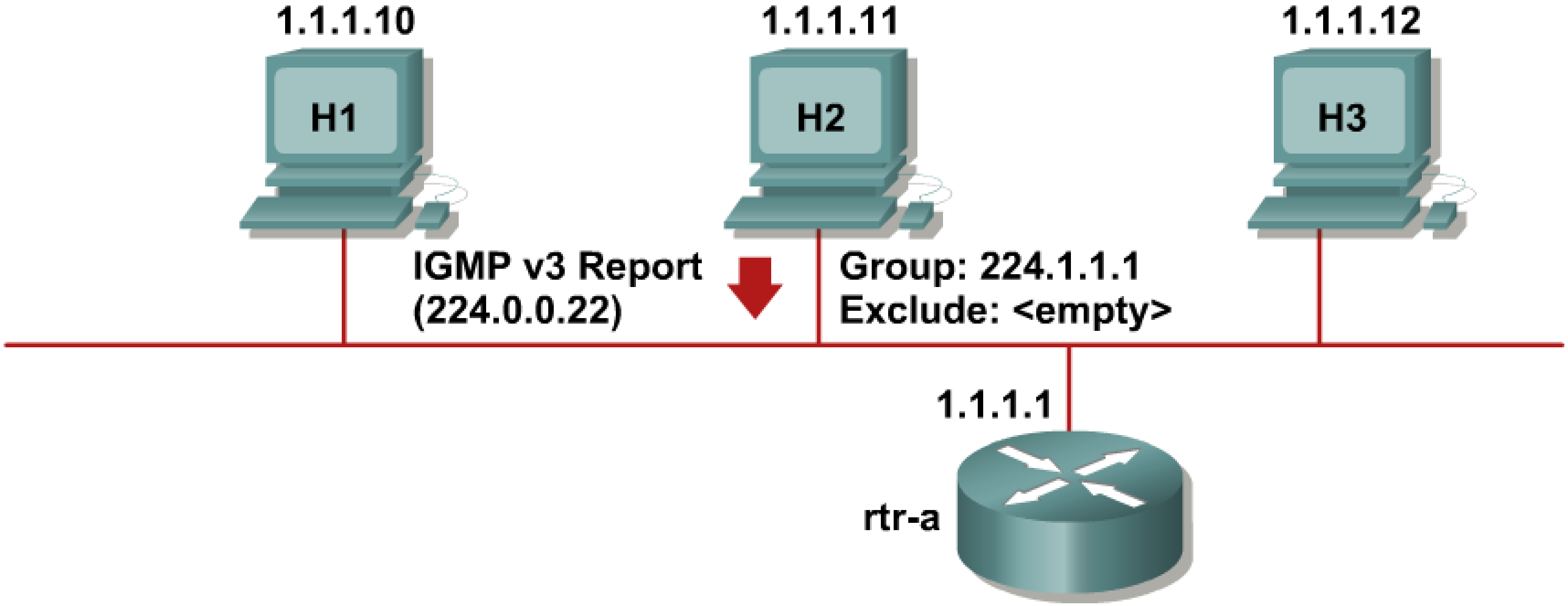
IGMPv2: Leaving the Multicast Group ⑤



IGMP State in rtr-a after H3 leaves.

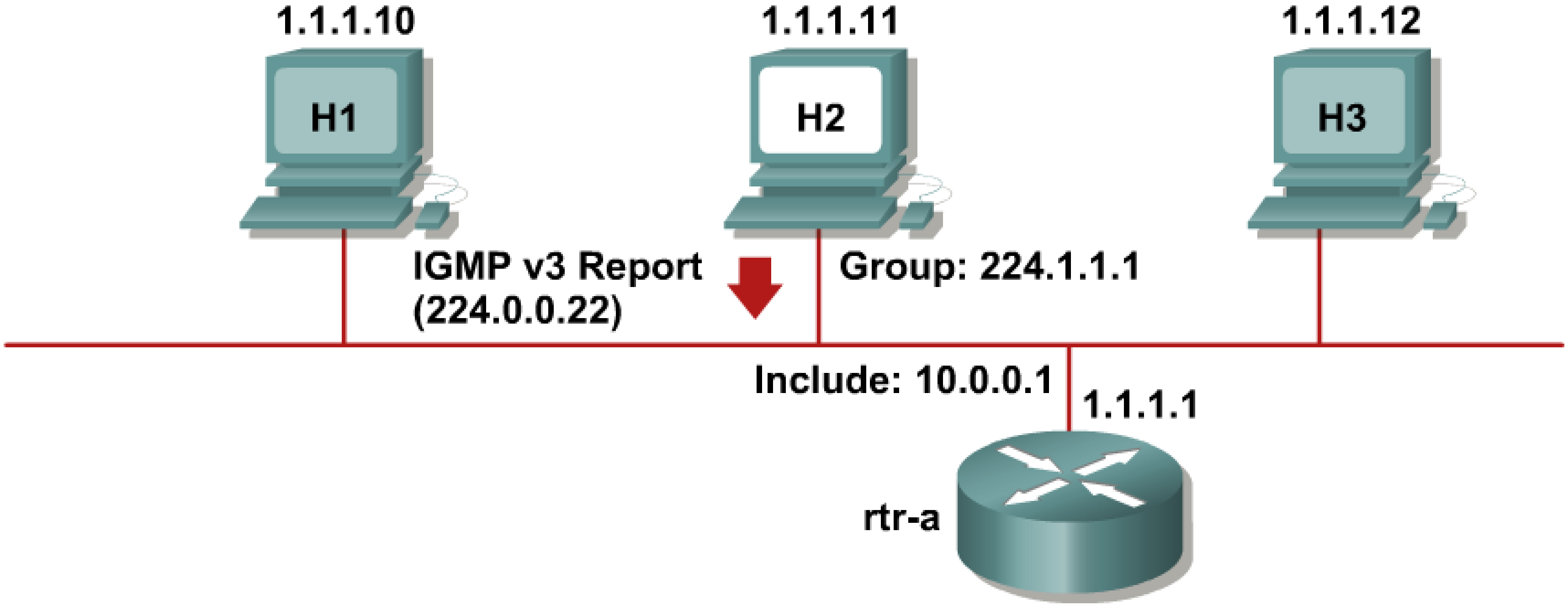
```
rtr-a>sh ip igmp group
IGMP Connected Group Membership
Group Address   Interface   Uptime   Expires    Last Reporter
```

IGMPv3: Signing on to Multicast Group



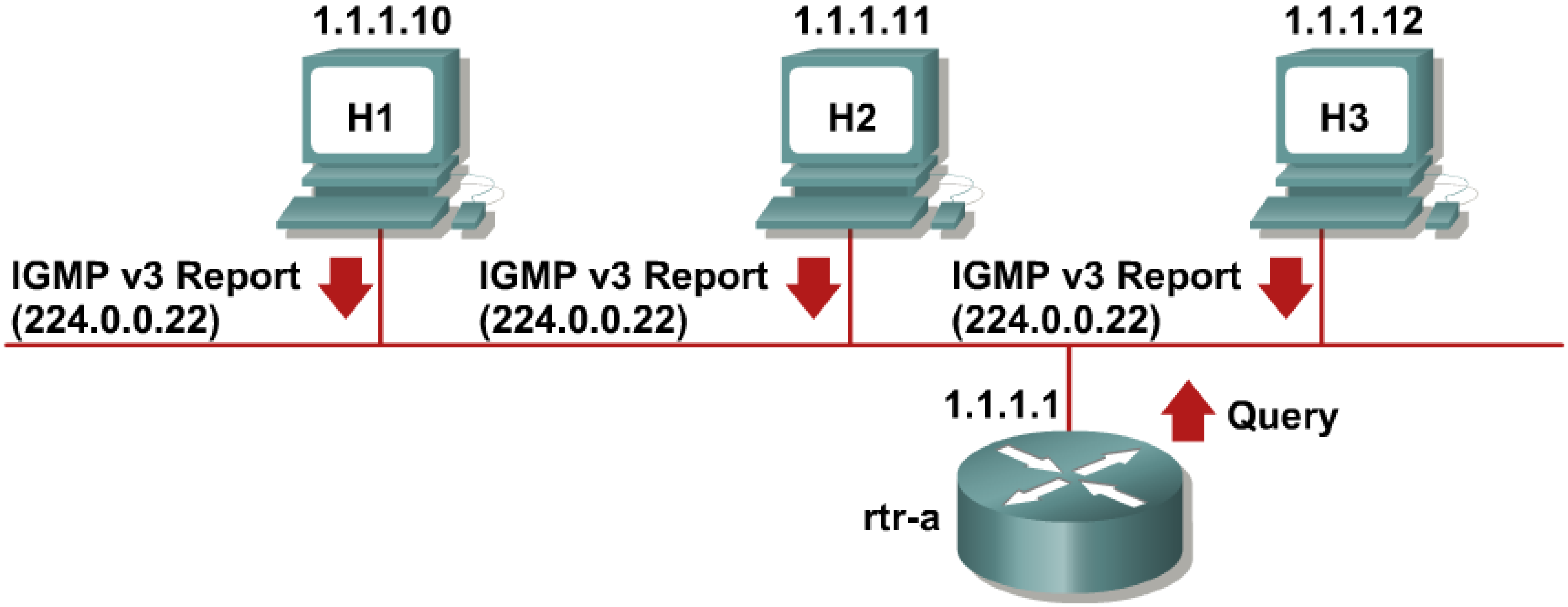
- New host sends IGMPv3 Membership Report to address 224.0.0.22

IGMPv3: Signing on to Specific Source in Multicast Group



- IGMPv3 Reports use INCLUDE/EXCLUDE set operations to direct signing on/off to multicast group and its sources

IGMPv3: Membership State Discovery



- Router sends periodic queries and all IGMPv3 members replies with complete multicast membership information

Verifying Membership State

- Displays relevant information about multicast on the target interface

```
Router# show ip igmp interface [IFACE]
```

- Displays information about multicast groups which members are on router’s local segments

```
Router# show ip igmp groups [group-address | IFACE]
```

Configuring Router as Group Member

- Configuring router to become member of target multicast group itself
  - Router sends IGMP Membership Report on local segment and becomes member of the multicast group
  - Consequence is that IP driver inside router would process all relevant multicast data as they should be delivered also to router

```
Router(config-if)# ip igmp join-group group-address
```

- Interface is included into outgoing interface list for target multicast group

```
Router(config-if)# ip igmp static-group group-address
```

IGMP Timers

- **Query Interval**
  - Period between two consecutive Group-Specific Membership Queries
  - By default 60 seconds

```
Router(config-if)# ip igmp query-interval seconds
```

- **Query Max-Response Time**
  - Each host randomly initializes response timer in range <0, MRT> each time it receives *Group-Specific Membership Query*
  - The host its timer expires first respond to this query
  - By default 10 seconds

```
Router(config-if)# ip igmp query-max-response-time seconds
```

- **Querier Timeout**
  - After expiration of this interval new IGMP Querier is elected on target segment
  - By default 120 seconds

```
Router(config-if)# ip igmp querier-timeout seconds
```

The show ip igmp interface Command

```
rtr-a> show ip igmp interface e0
Ethernet0 is up, line protocol is up
  Internet address is 1.1.1.1, subnet mask is 255.255.255.0
  IGMP is enabled on interface
  Current IGMP version is 2
  CGMP is disabled on interface
  IGMP query interval is 60 seconds
  IGMP querier timeout is 120 seconds
  IGMP max query response time is 10 seconds
  Inbound IGMP access group is not set
  Multicast routing is enabled on interface
  Multicast TTL threshold is 0
  Multicast designated router (DR) is 1.1.1.1 (this system)
  IGMP querying router is 1.1.1.1 (this system)
  Multicast groups joined: 224.0.1.40 224.2.127.254
```

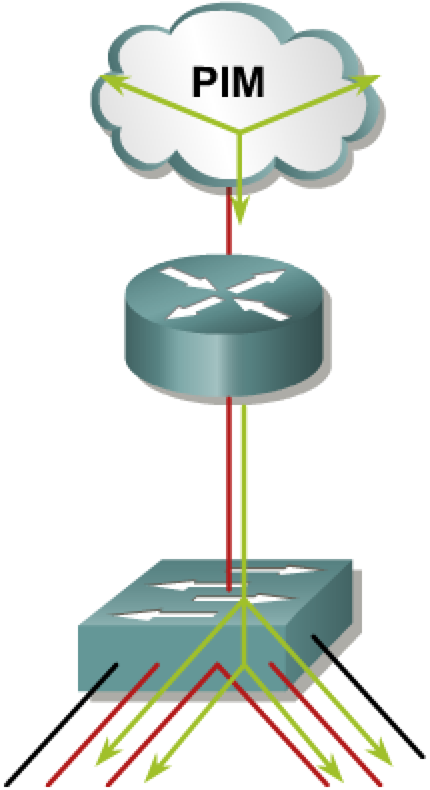
The show ip igmp groups Command

```
rtr-a> show ip igmp groups
IGMP Connected Group Membership
Group Address    Interface      Uptime    Expires   Last Reporter
224.1.1.1        Ethernet0      6d17h     00:01:47  1.1.1.12
224.0.1.40       Ethernet0      6d17h     never     1.1.1.17
```

Effective Delivery of L2 Multicast

Effective Delivery of L2 Multicast ①

- *Problem*
  - Usual L2 switches process multicast frames as frames intended to unknown destination – *they broadcast it!*
  - Efficient delivery of multicast means that multicast data will be delivered only to subscribed members of multicast group

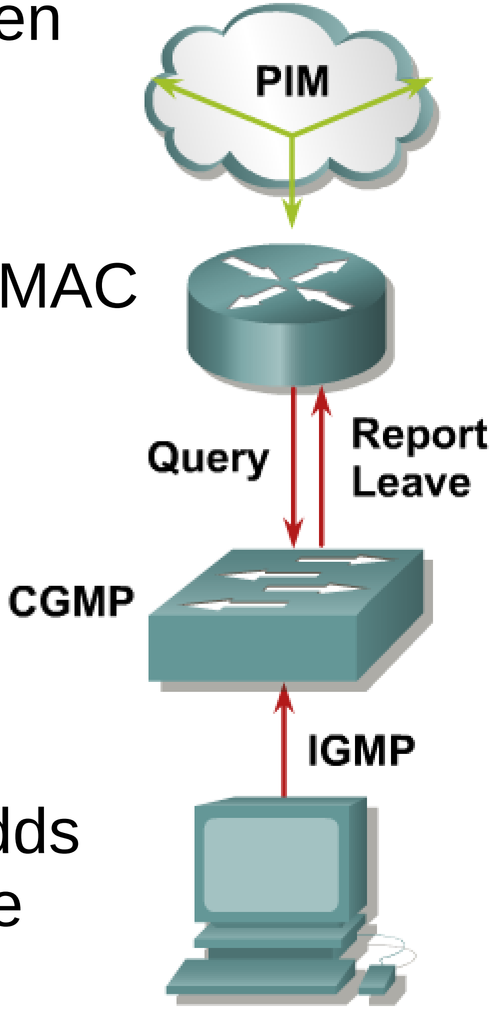


Effective Delivery of L2 Multicast ②

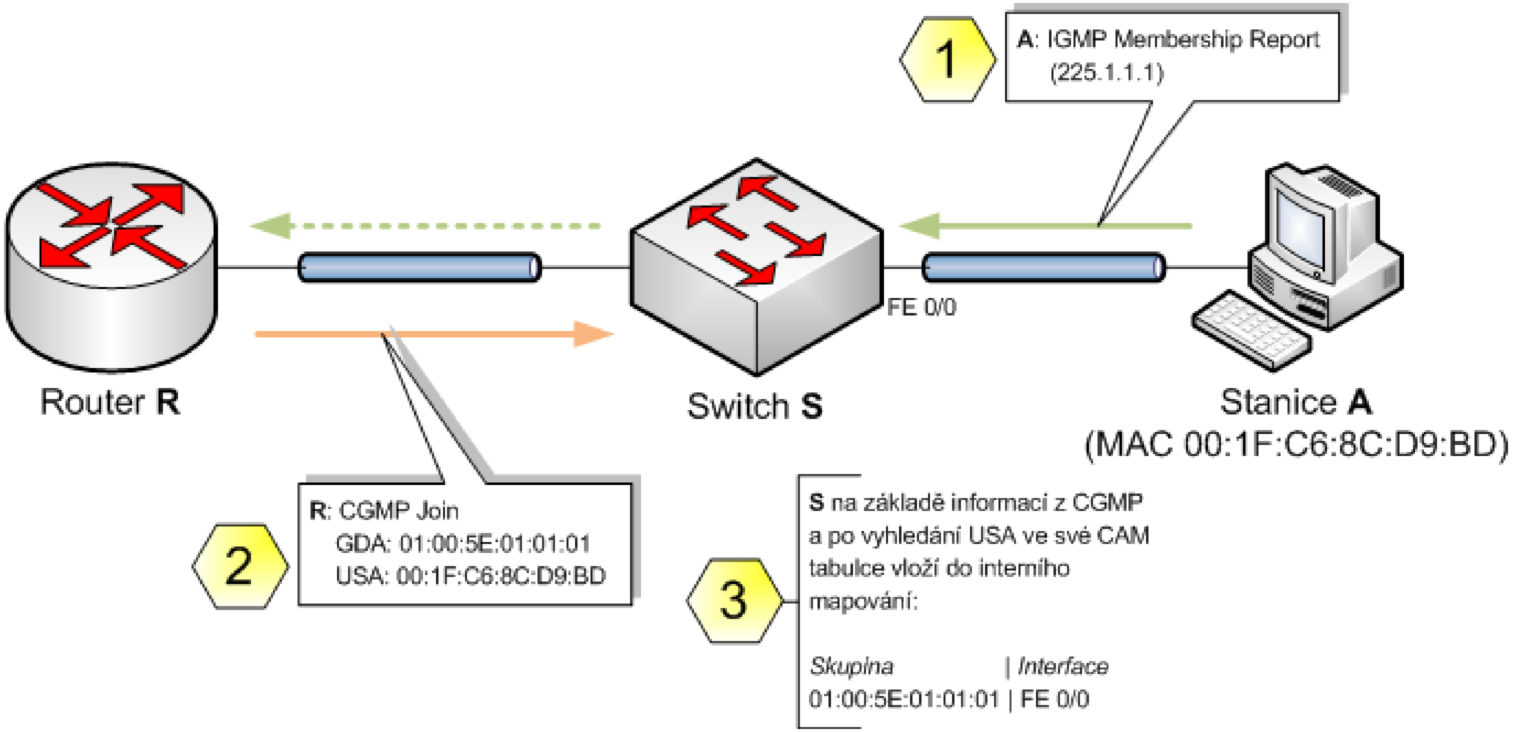
- Idea for dynamic multicast delivery mechanism
  - Switch keeps a track to which multicast groups host belongs to
  - Processed multicast frame is passed only to those interfaces where relevant multicast members (hosts) reside
- Two mechanisms:
  - **Active**: **Cisco Group Management Protocol (CGMP)** simple yet proprietary auxiliary protocol based on client-server model (switch-router) for a multicast membership management
  - **Passive**: **IGMP snooping** is complex but standardized and well accepted solution based on sniffing of relevan multicast information implemented into switches

CGMP

- CGMP is auxiliary signaling protocol between router and switch
  - *It is not a replacement or analogy for IGMP!*
- CGMP frame is sent by router to switch on MAC reserved address **0100.0cdd.dddd**
- CGMP frame contains
  - **Type** = Join or Leave
  - **USA** = MAC address of IGMP client
  - **GDA** = MAC address of multicast group
- According to received information switch adds or deletes multicast MAC on target interface



CGMP: Sign on to Multicast Group



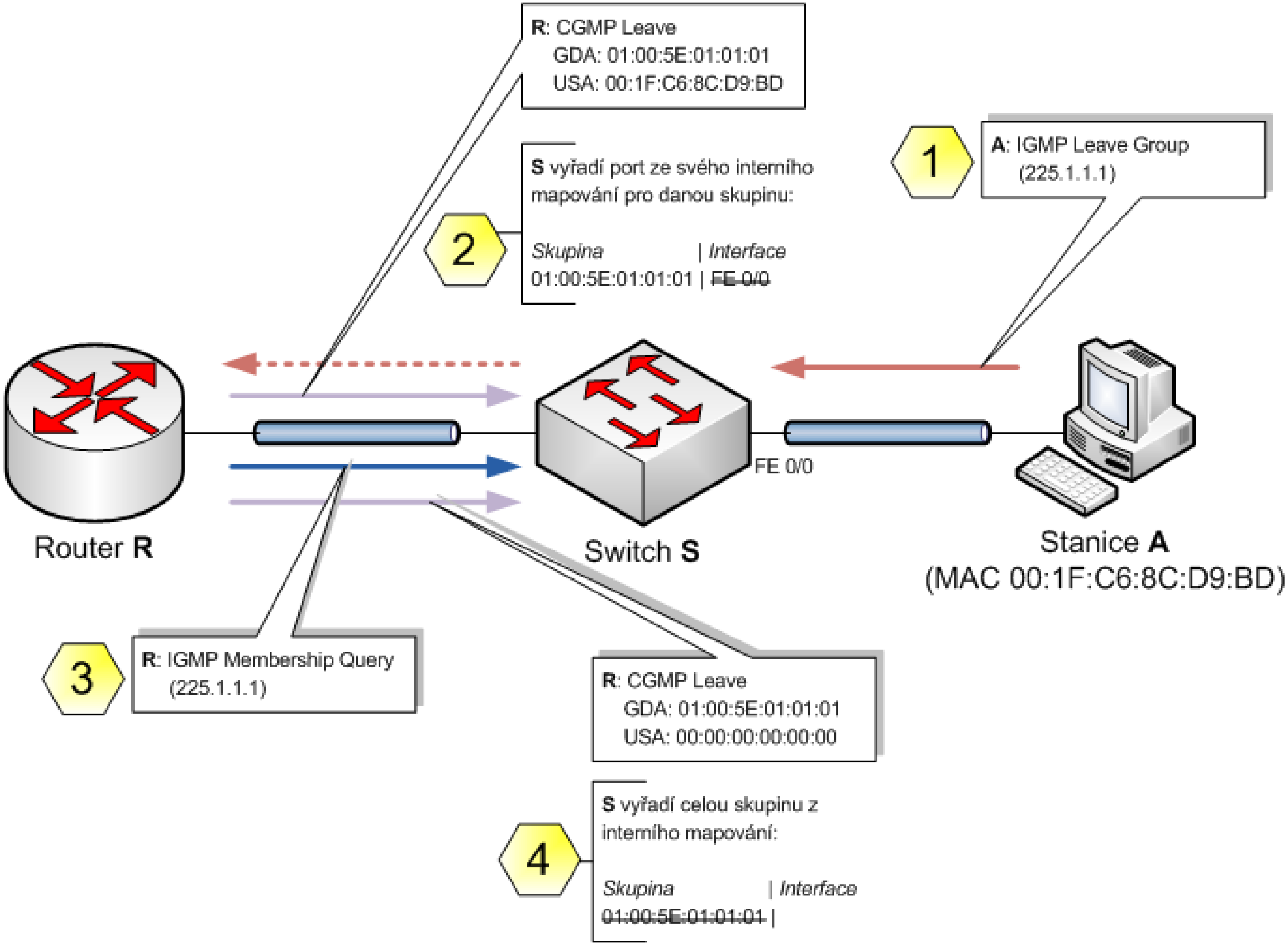
1 **A**: IGMP Membership Report (225.1.1.1)

2 **R**: CGMP Join
GDA: 01:00:5E:01:01:01
USA: 00:1F:C6:8C:D9:BD

3 **S** na základě informací z CGMP a po vyhledání USA ve své CAM tabulce vloží do interního mapování:

Skupina | *Interface*
01:00:5E:01:01:01 | FE 0/0

CGMP: Sign off to Multicast Group



R: CGMP Leave
GDA: 01:00:5E:01:01:01
USA: 00:1F:C6:8C:D9:BD

A: IGMP Leave Group
(225.1.1.1)

S vyřadí port ze svého interního mapování pro danou skupinu:

Skupina	*Interface*
01:00:5E:01:01:01	~~FE 0/0~~

R: IGMP Membership Query
(225.1.1.1)

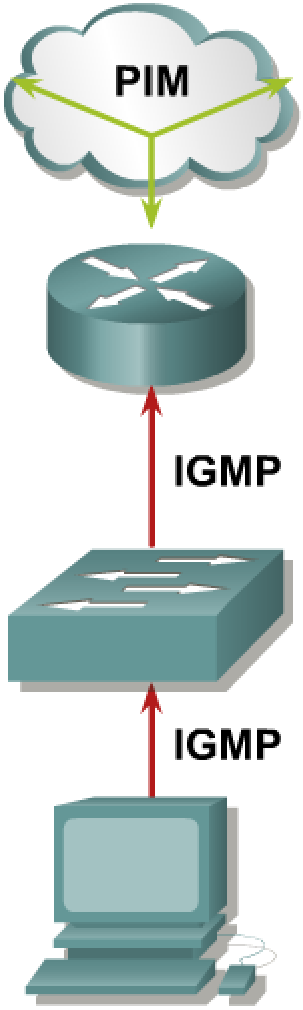
R: CGMP Leave
GDA: 01:00:5E:01:01:01
USA: 00:00:00:00:00:00

S vyřadí celou skupinu z interního mapování:

Skupina	*Interface*
~~01:00:5E:01:01:01~~	

IGMP Snooping ①

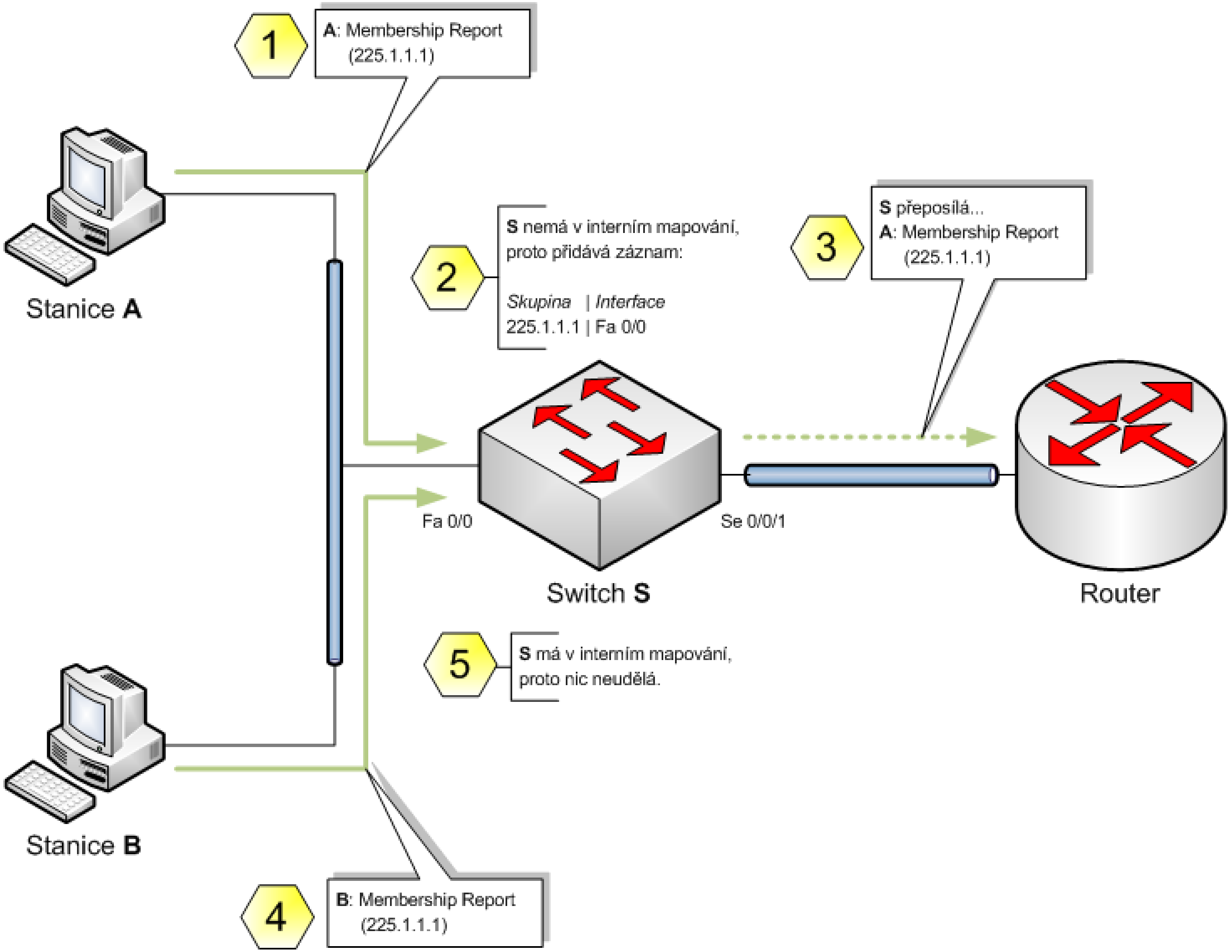
- *Switch sniffs through IGMP in IP packets*
- IGMP packets are processed in the CPU or in the specialized ASIC (Application-Specific Integrated Circuit)
- Switch analyzes content of IGMP messages in order to discover on which ports are present members of multicast groups
- Consequence to switches WITHOUT L3-aware HW/ASIC
    - CPU must processed all L2 frames in order to discover relevant IGMP packets
    - Reduce performance and throughput
- Consequence to switches WITH L3-aware HW/ASIC
    - Performance is preserved, but price is higher



IGMP Snooping ②

- IGMPv3 Membership Reports are sent to reserved multicast address 224.0.0.22
  - It simplifies analysis, because there is no need to snoop on other data traffic intended for 224.0.0.2
  - For SW IGMP Snooping swithces IGMPv3 solution does not pose such a burden
- On current Catalyst switches is IGMP Snooping active by default
  - Only exception are multicast frames with destination address 224.0.0.* (01:00:5e:00:00:**), which are always flooded – *do not forget about 32-to-1 overlapping problem!*

IGMP Snooping: Sing on to Multicast Group



1 **A**: Membership Report (225.1.1.1)

2 **S** nemá v interním mapování, proto přidává záznam:

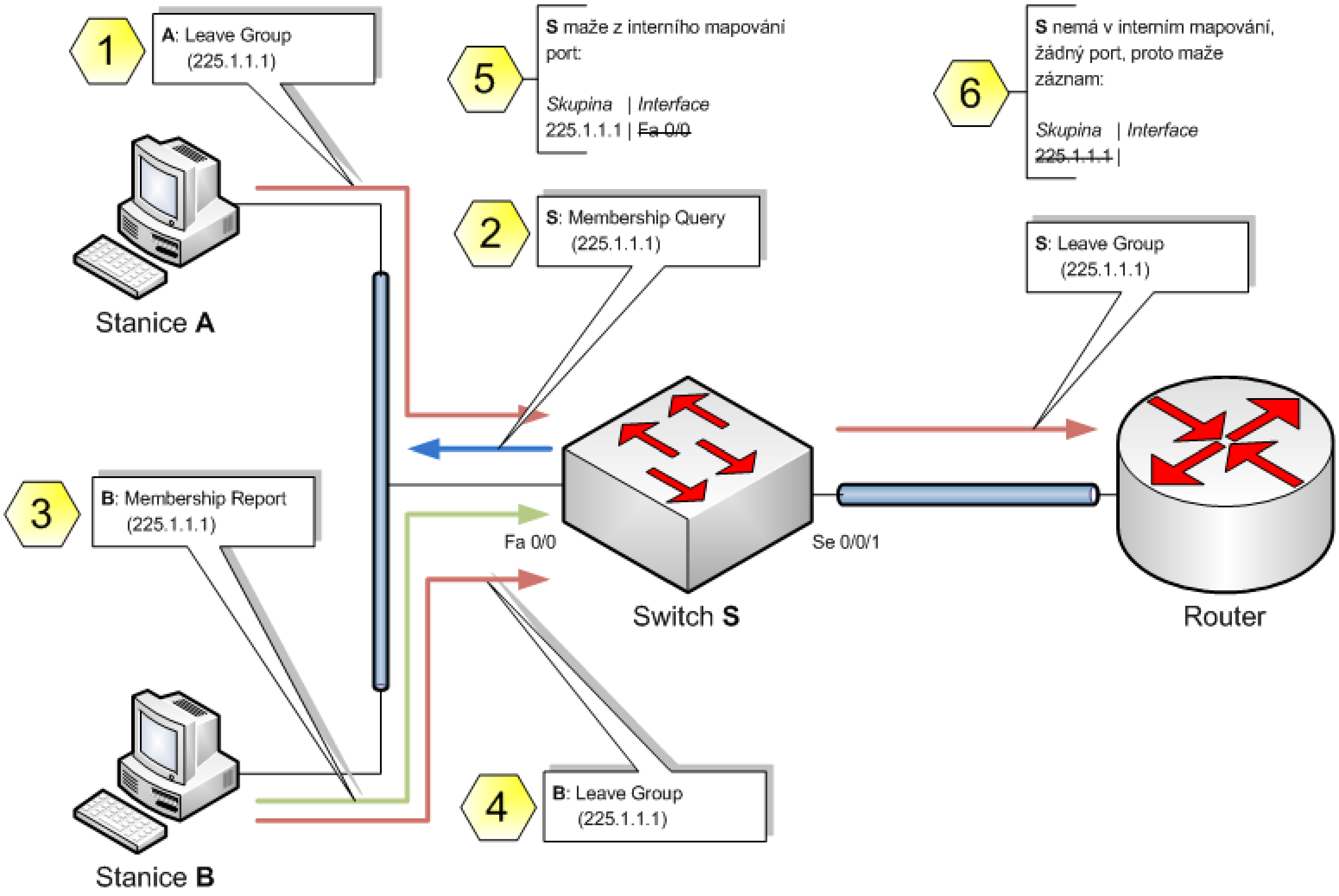
Skupina | *Interface*
225.1.1.1 | Fa 0/0

3 **S** přeposílá... **A**: Membership Report (225.1.1.1)

4 **B**: Membership Report (225.1.1.1)

5 **S** má v interním mapování, proto nic neudělá.

IGMP Snooping: Sing off to Multicast Group



1 **A**: Leave Group (225.1.1.1)

2 **S**: Membership Query (225.1.1.1)

3 **B**: Membership Report (225.1.1.1)

4 **B**: Leave Group (225.1.1.1)

5 **S** maže z interního mapování port:

Skupina | *Interface*
225.1.1.1 | ~~Fa 0/0~~

6 **S** nemá v interním mapování, žádný port, proto maže záznam:

Skupina | *Interface*
~~225.1.1.1~~ |

S: Leave Group (225.1.1.1)

Configuring IGMP Snooping ①

- Enable IGMP snooping globally. (By default, it is enabled globally.)

```
Switch(conf)# ip igmp snooping
```

- Switches add multicast router ports to the forwarding table for every Layer 2 multicast entry by observing PIM or CGMP.
- By default they learn router port using PIM, but it can be changed:

```
Switch(conf)# ip igmp snooping vlan vlan-id mrouter learn
                [cgmp | pim-dvmrp]
```

- Whenever needed, configure the router port statically (by default IGMP snooping detects it automatically):

```
Switch(conf)#
   ip igmp snooping vlan vlan-id mrouter interface IFACE
```

Configuring IGMP Snooping ②

- Enable IGMP snooping globally. (By default, it is enabled globally.)

```
Switch(config)# ip igmp snooping vlan VID fast-leave
Switch(config)# ip igmp snooping vlan VID immediate-leave
```

- By default, all hosts register and add the MAC address and port to the forwarding table automatically

- Static configuration of multicast on port:

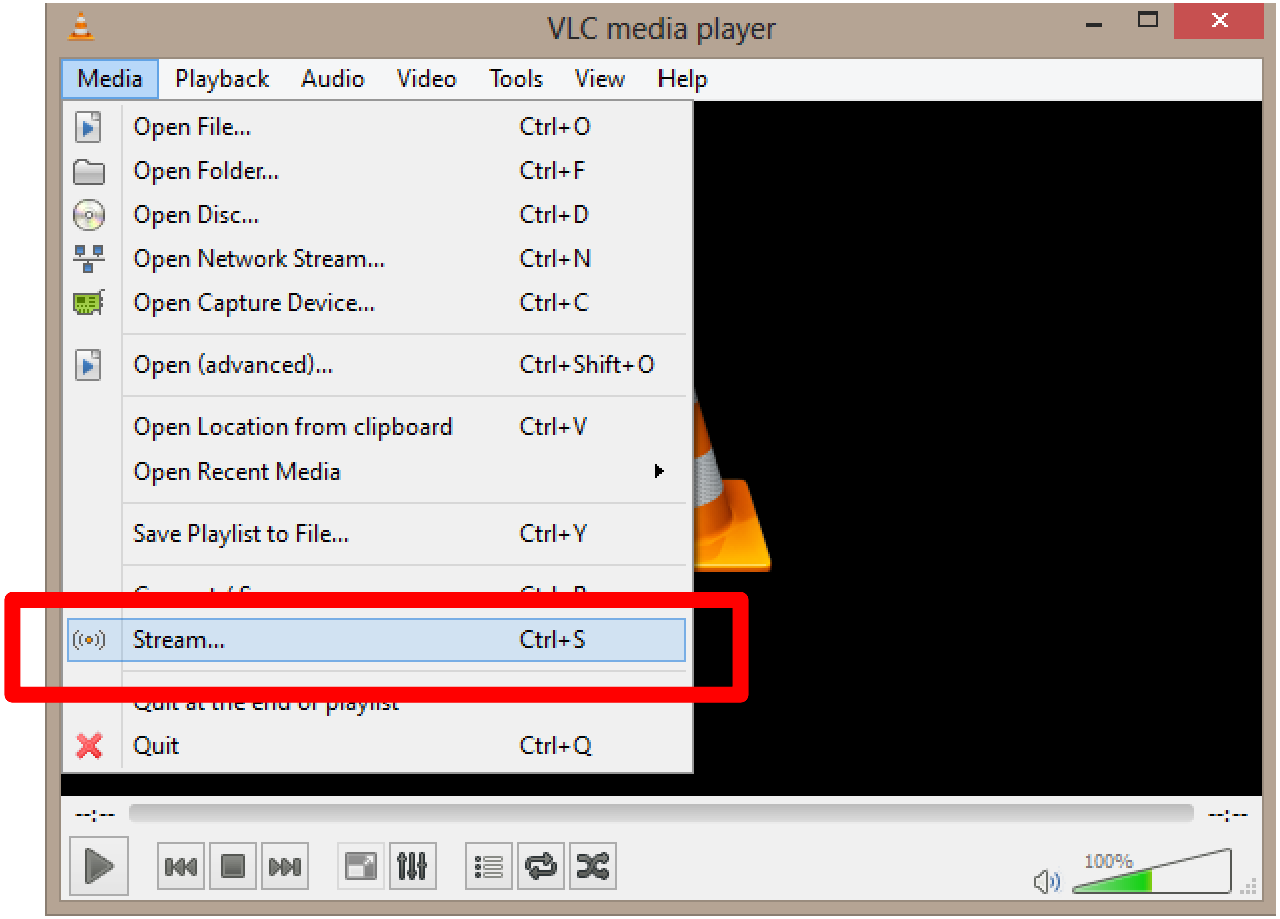
```
Switch(conf)#
  ip igmp snooping vlan VID static MAC interface IFACE
```

Verifying IGMP Snooping

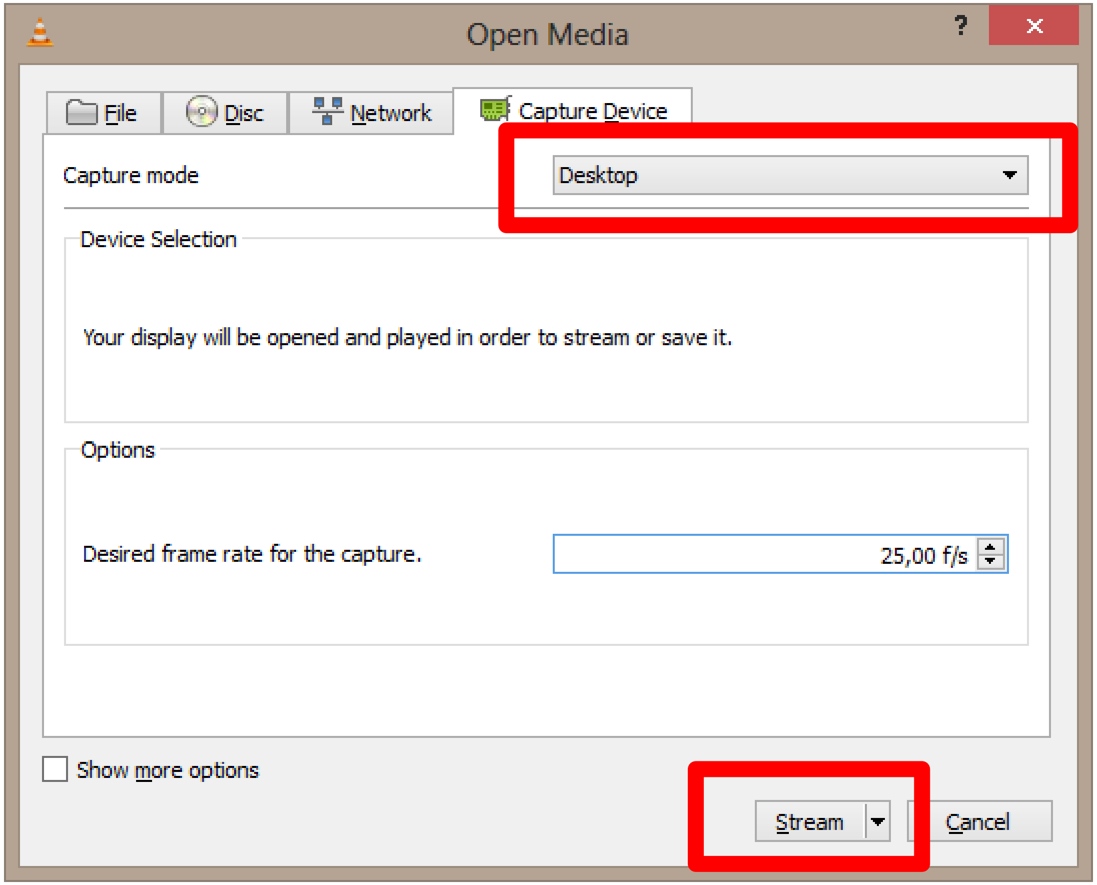
```
show ip igmp snooping

show ip igmp snooping multicast

show ip igmp snooping multicast vlan vlan-id

show ip igmp snooping mrouter
```

VLC Player

Stream Desktop ①



Stream Desktop ②



Open Media

File | Disc | Network | Capture Device

Capture mode: Desktop

Device Selection

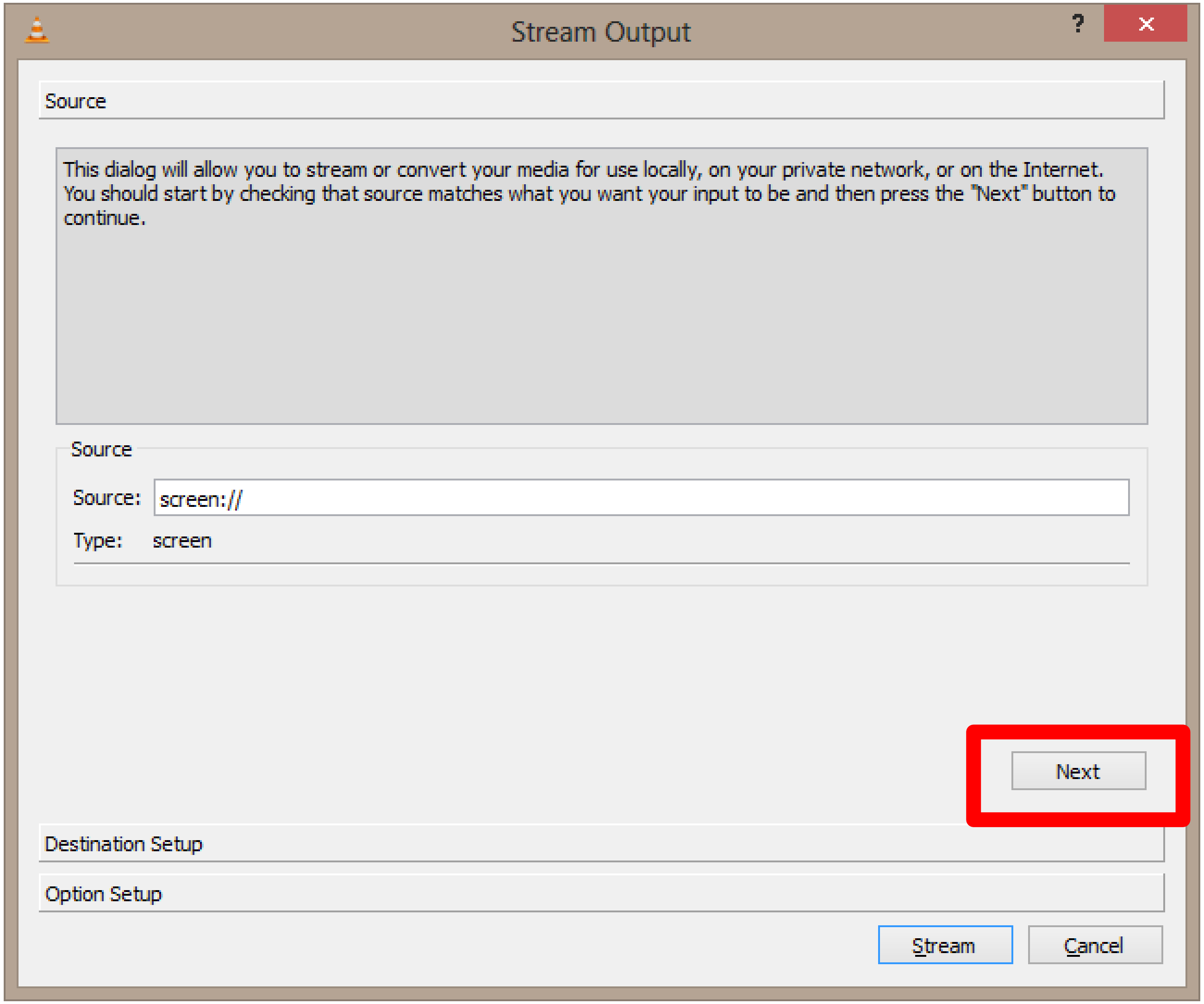
Your display will be opened and played in order to stream or save it.

Options

Desired frame rate for the capture. 25,00 f/s

□ Show more options

Stream Desktop ③



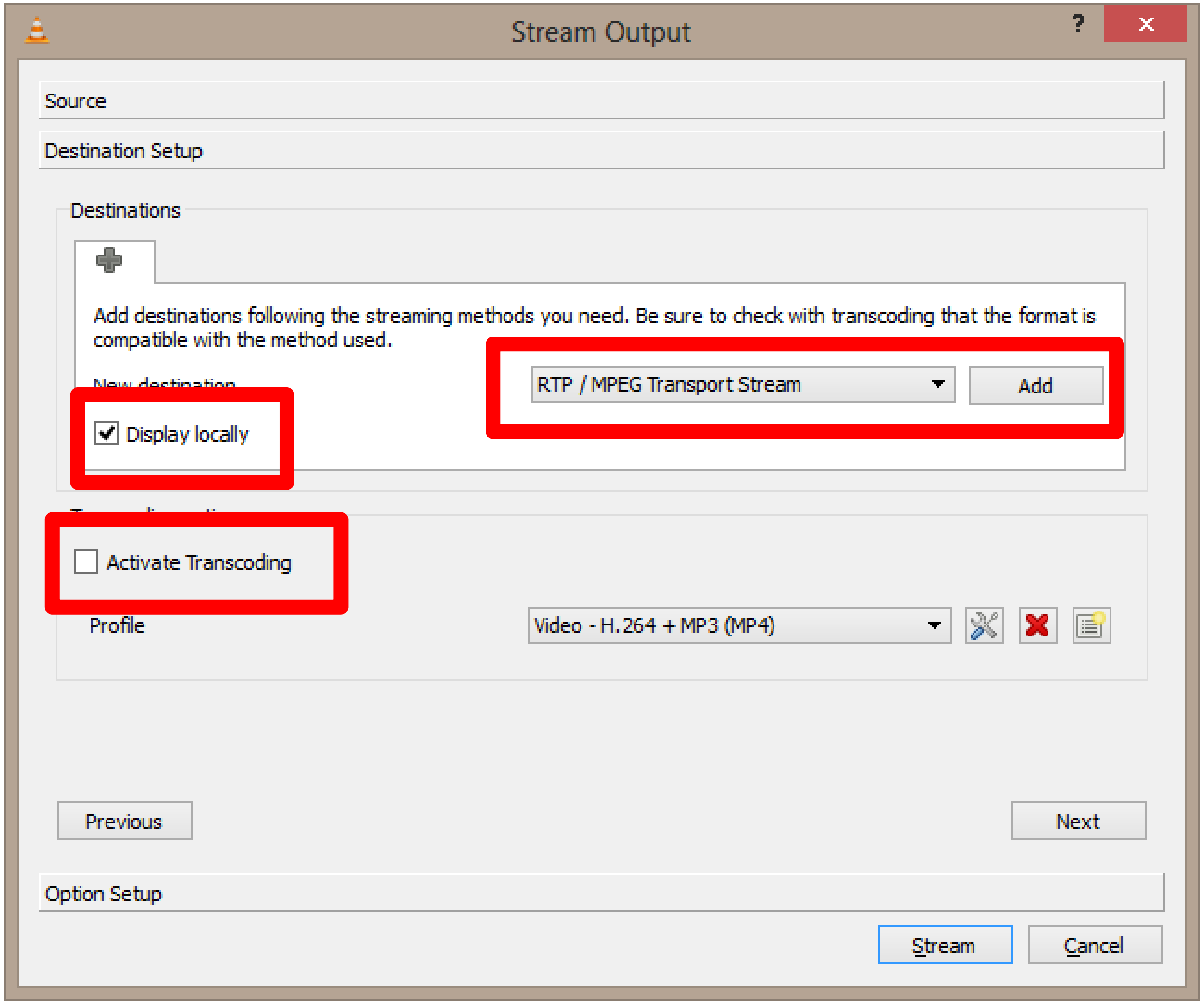
This dialog will allow you to stream or convert your media for use locally, on your private network, or on the Internet. You should start by checking that source matches what you want your input to be and then press the "Next" button to continue.

Source

Source: screen://

Type: screen

Stream Desktop ④



Stream Output

Source

Destination Setup

Destinations

Add destinations following the streaming methods you need. Be sure to check with transcoding that the format is compatible with the method used.

New destination: RTP / MPEG Transport Stream

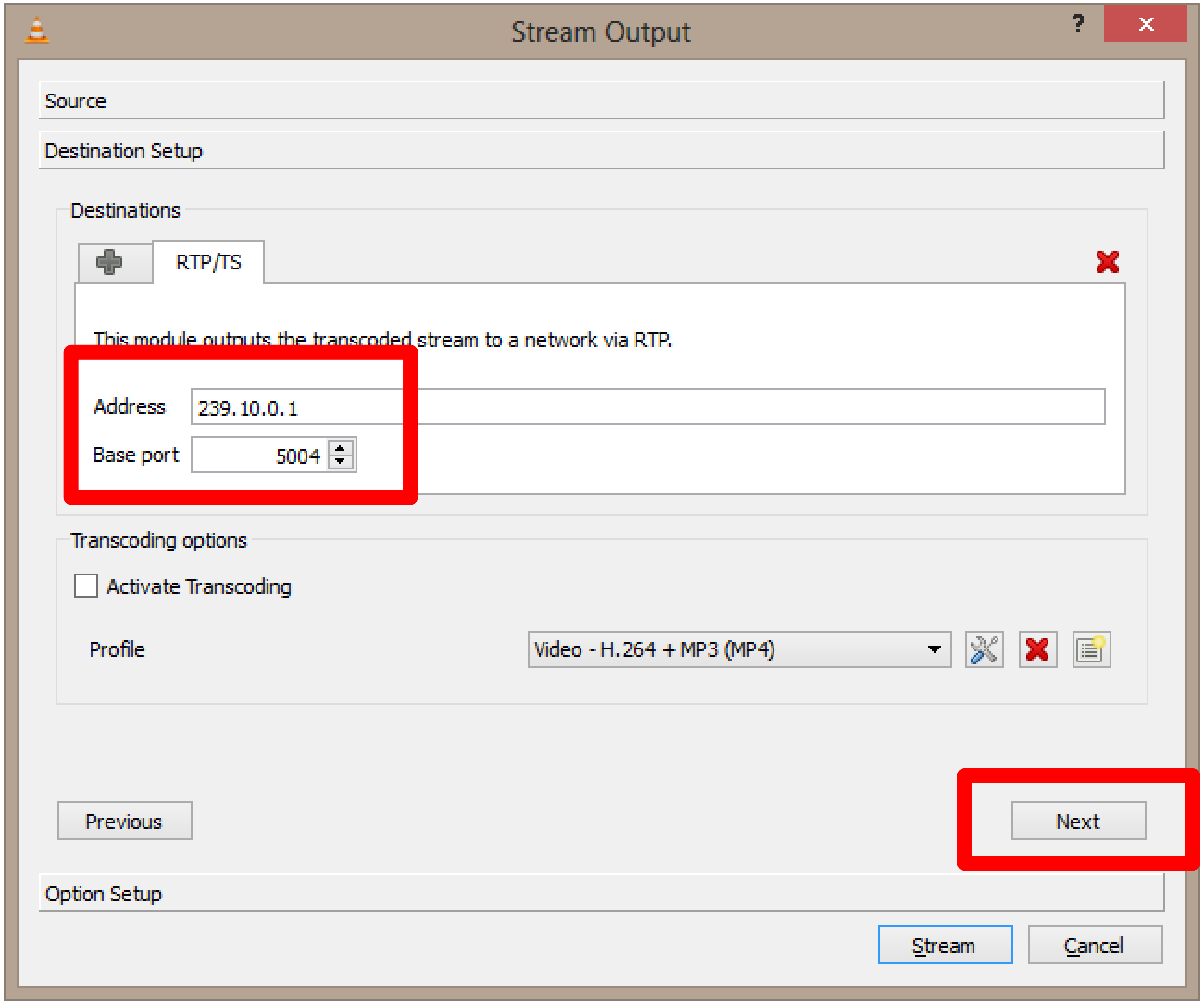
☑ Display locally

□ Activate Transcoding

Profile: Video - H.264 + MP3 (MP4)

Option Setup

Stream Desktop ⑤



Stream Output

Source

Destination Setup

Destinations

RTP/TS

This module outputs the transcoded stream to a network via RTP.

Address 239.10.0.1

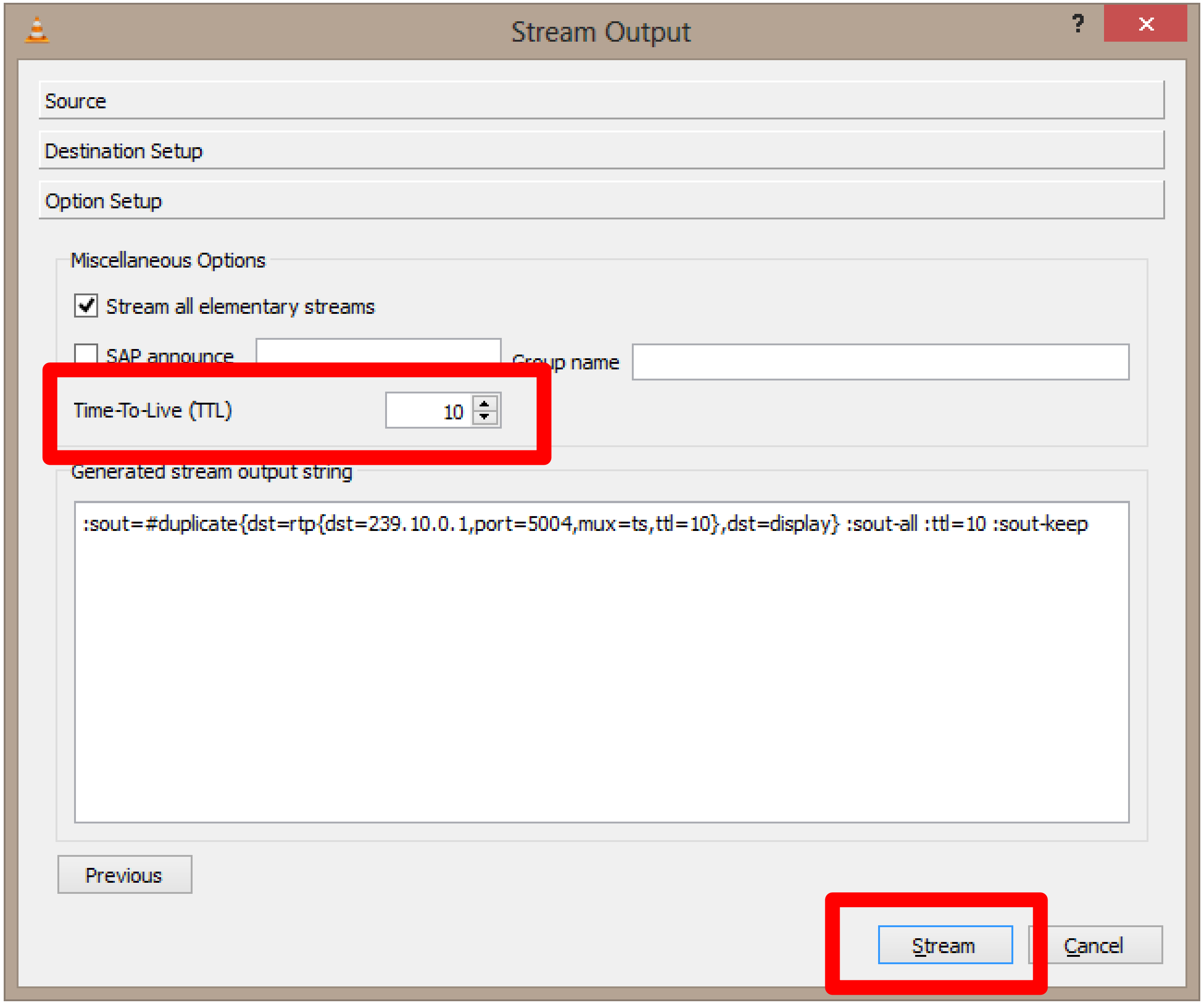
Base port 5004

Transcoding options

□ Activate Transcoding

Profile Video - H.264 + MP3 (MP4)

Stream Desktop ⑥



☑ Stream all elementary streams

□ SAP announce

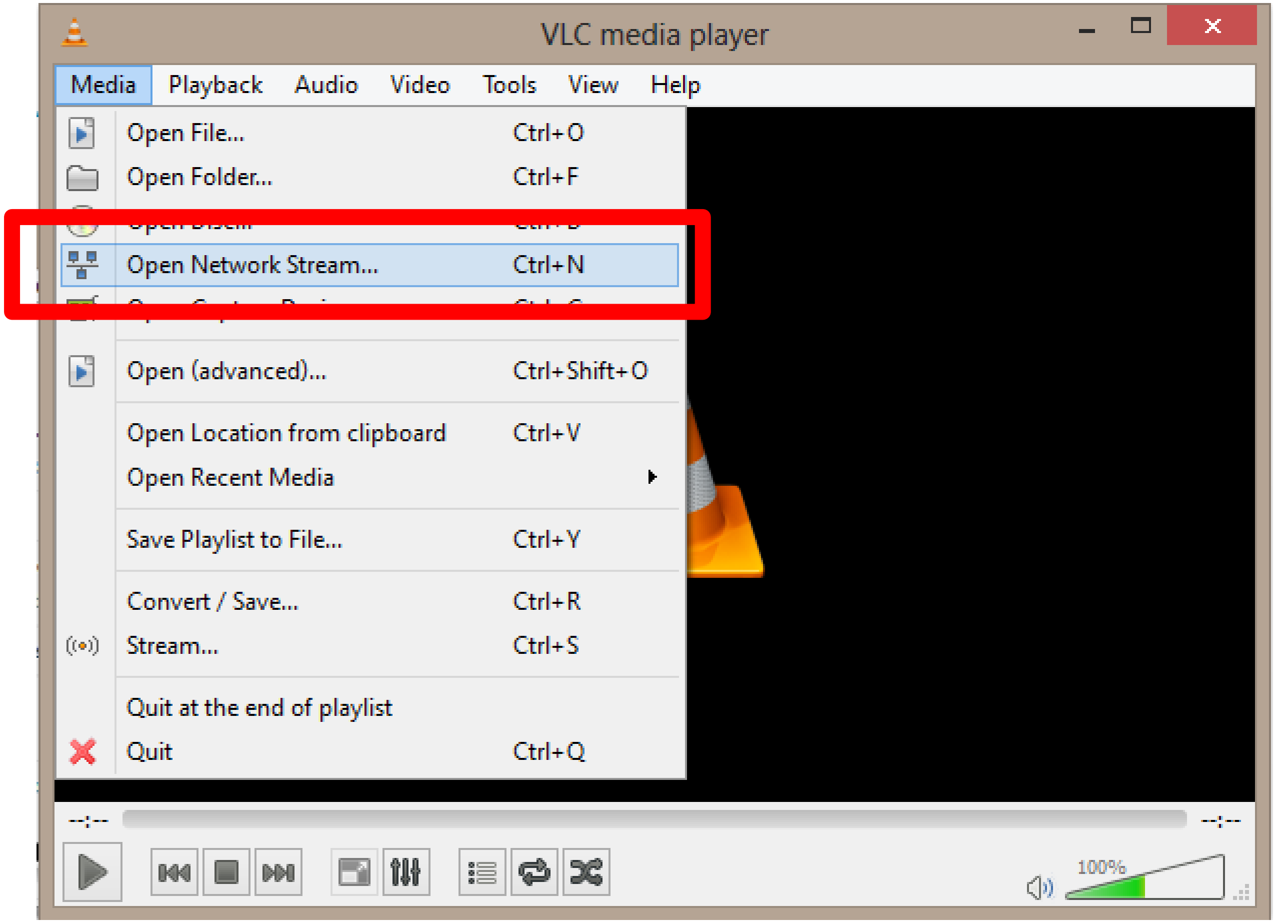
Group name

Time-To-Live (TTL) 10

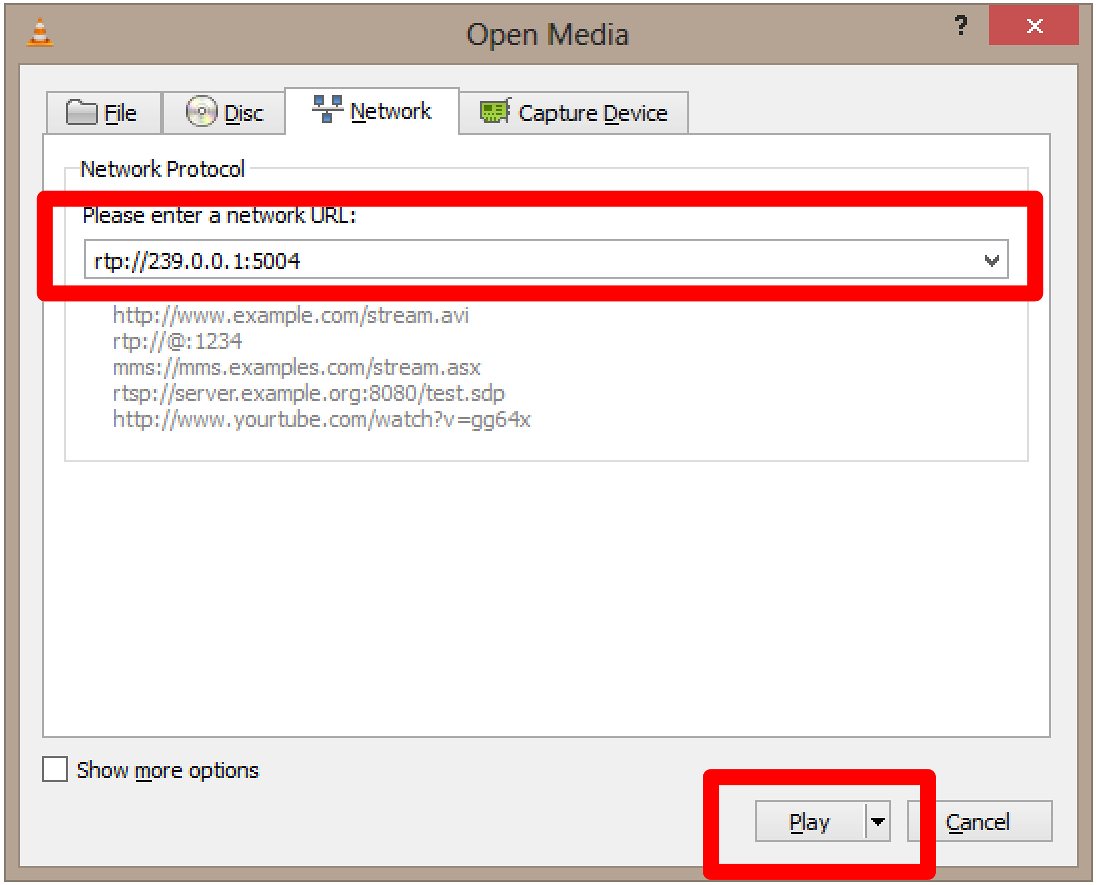
Generated stream output string

```
:sout=#duplicate{dst=rtp{dst=239.10.0.1,port=5004,mux=ts,ttl=10},dst=display} :sout-all :ttl=10 :sout-keep
```

Receive Streaming ①



Receive Streaming ②: ASM



Open Media

Network Protocol

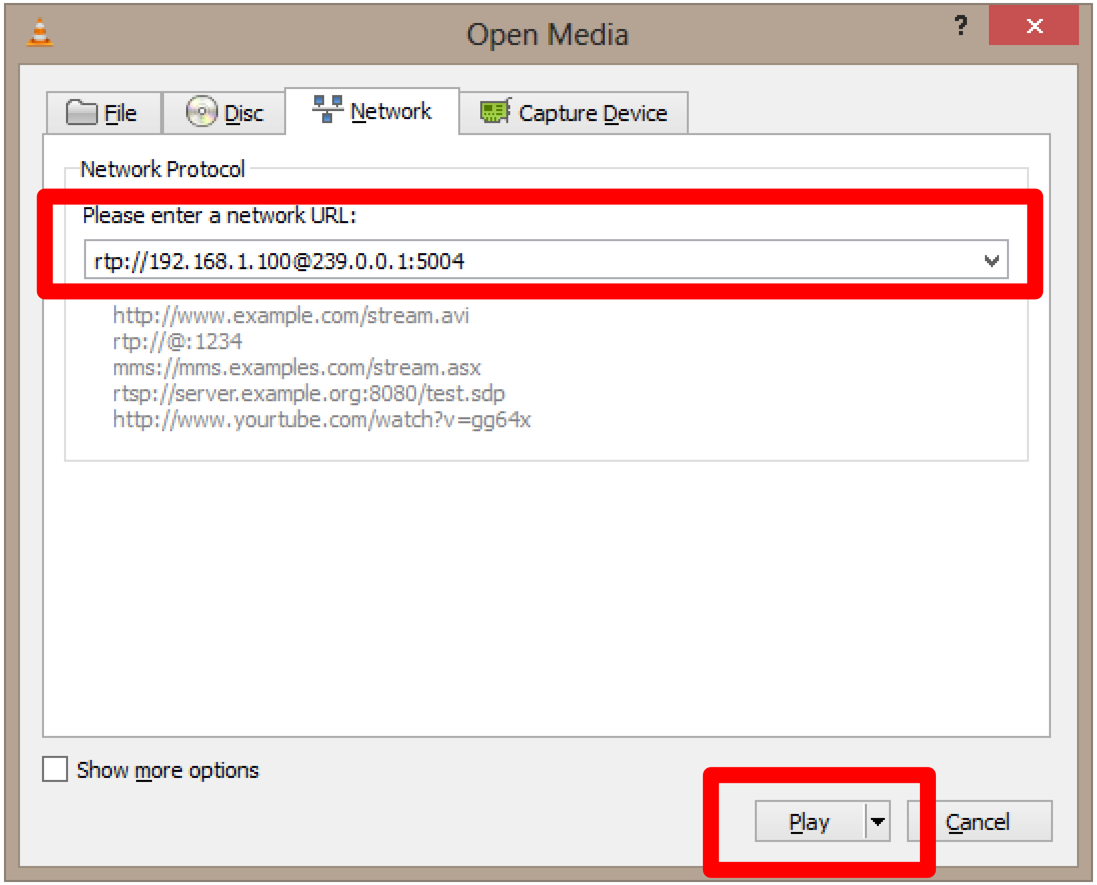
Please enter a network URL:

rtp://239.0.0.1:5004

http://www.example.com/stream.avi
rtp://@:1234
mms://mms.examples.com/stream.asx
rtsp://server.example.org:8080/test.sdp
http://www.yourtube.com/watch?v=gg64x

□ Show more options

Receive Streaming ②: SSM



Open Media

Network Protocol

Please enter a network URL:

rtp://192.168.1.100@239.0.0.1:5004

http://www.example.com/stream.avi
rtp://@:1234
mms://mms.examples.com/stream.asx
rtsp://server.example.org:8080/test.sdp
http://www.yourtube.com/watch?v=gg64x

□ Show more options

Multicast Distribution Trees

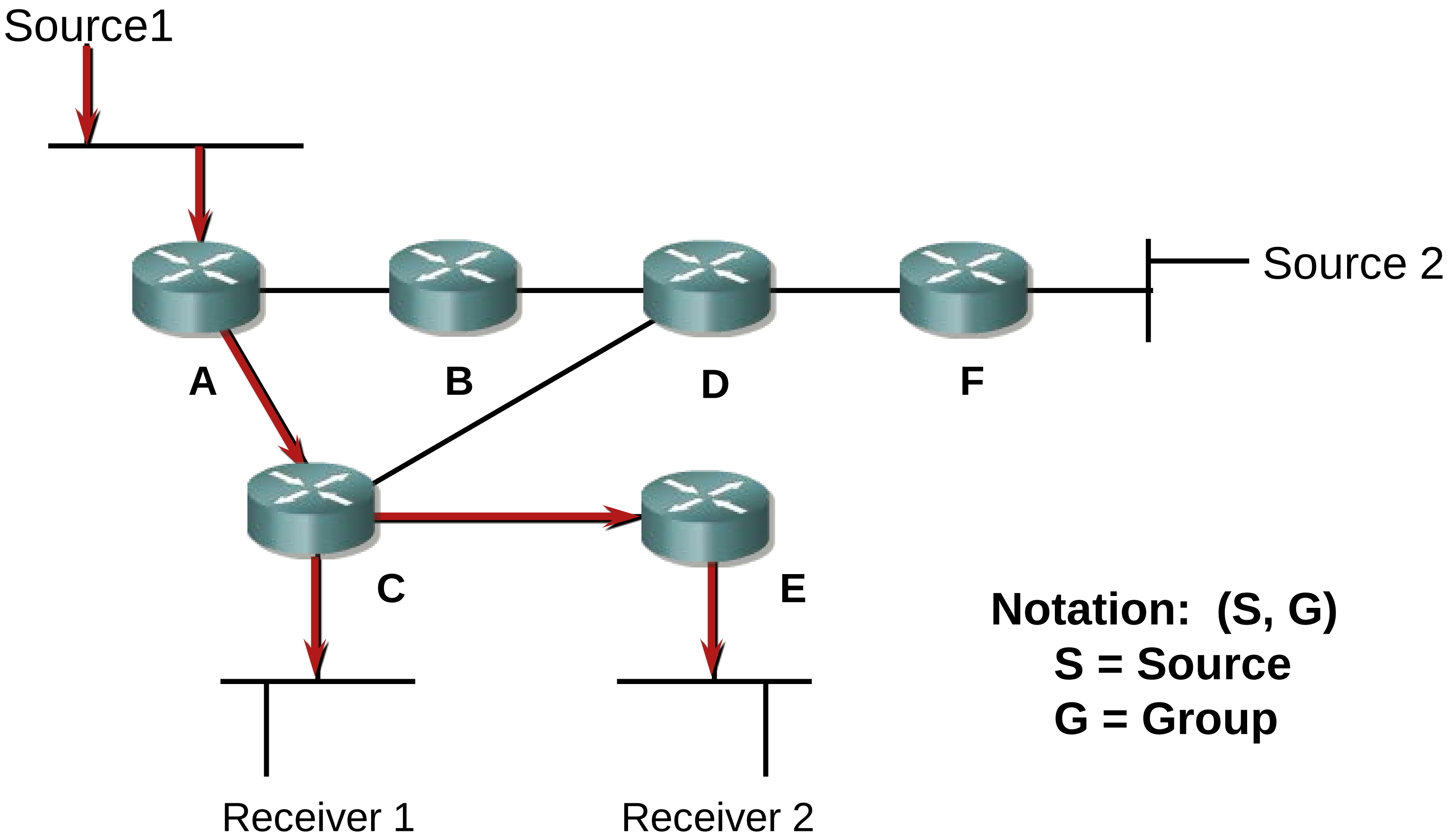
Multicast Distribution Trees

- **Multicast Distribution Tree** = multicast data flows along this path, which creates tree structure
- Two tree types
  - **Source** (distribution) **trees** are shortest path trees (a.k.a. SPT)
    - Source (sender) of multicast data is a root of this tree
  - **Shared** (distribution) **trees**
    - One tree is shared among multiple sources in the target multicast group
    - Rendezvous point (designated router) is a root of this tree

Distribution Trees Characteristics

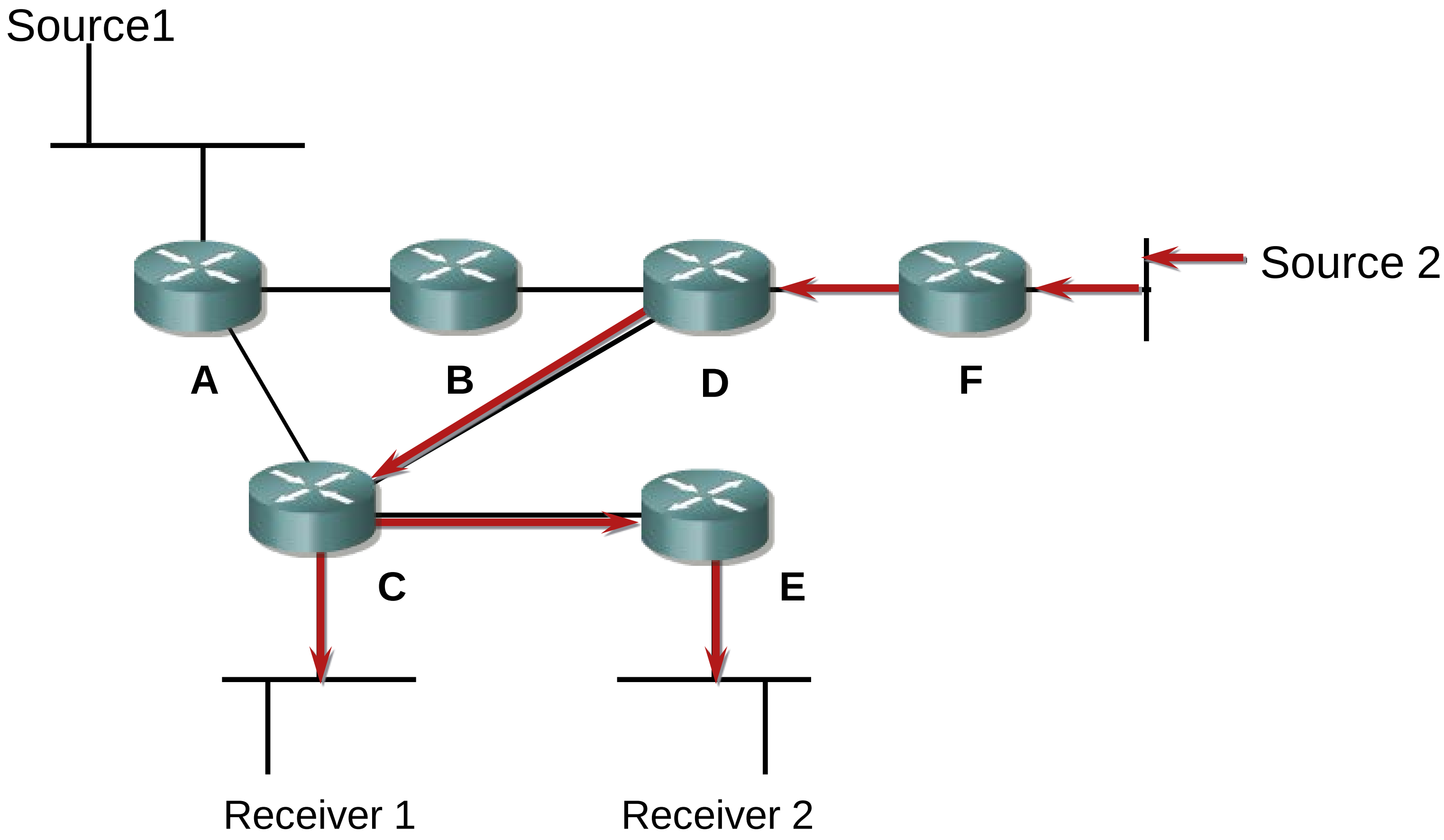
- SPT trees are more memory intensive but guarantee shortest path from sender to all receivers, thus minimalizing delay
- Shared trees consume less router resources but suboptimal routing could occur (introducing extra delay)

Source Trees ①

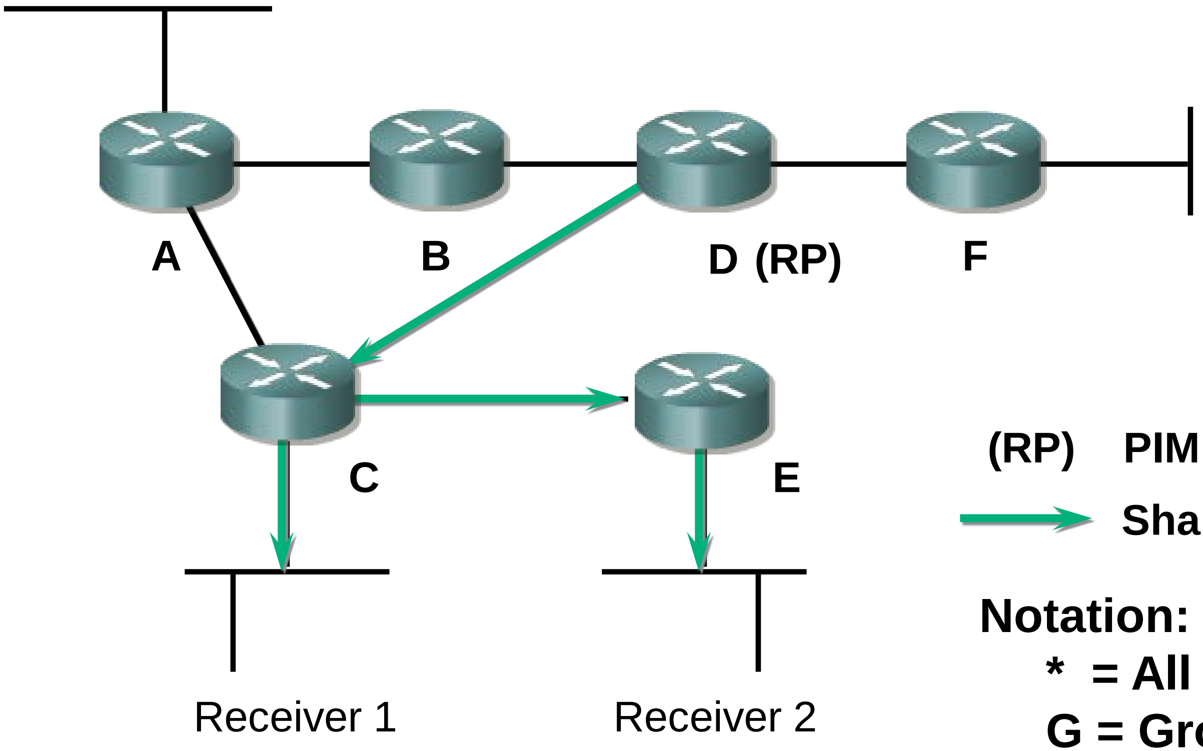


Notation: (S, G)
S = Source
G = Group

Source Trees ②



Shared Trees ①



(RP) PIM Rendezvous Point

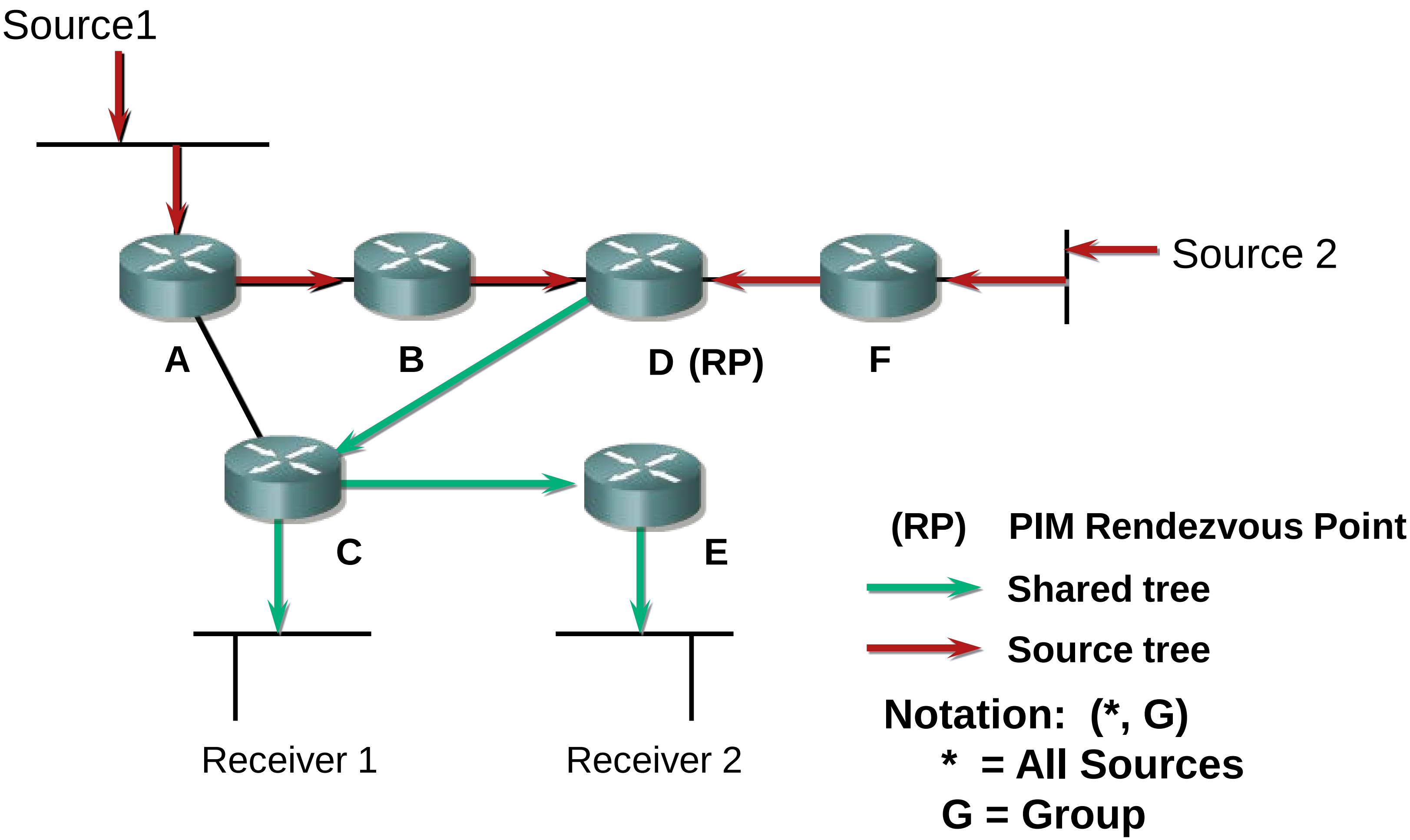
→ Shared tree

Notation: (*, G)

*** = All Sources**

G = Group

Shared Trees ②



(RP) PIM Rendezvous Point

🟩 Shared tree

🟥 Source tree

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

Multicast Routing Table Items

- **Item (S,G)**
  - For each source (S) sending data to multicast group (G)
  - Multicast flows through shortest path between sender and each receiver
- **Item (*,G)**
  - One item for any source (*) sending data to multicast group (G)
  - Multicast flows from source through RP to receivers – not necessarily using shortest path

Verifying Multicast Routing Table

```
Router#
show ip mroute [group-address] [summary] [count] [active kbps]
```

- Display content of multicast routing table
  - **`summary`**: A brief variant of report
  - **`count`**: Displays statistics about groups and sources including number of packets, packets per second throughput or average packet size
  - **`active`**: Shows statistics about every active multicast source (as active is considered every source with throughput higher or equal to *`kbps`* (default is 4 kbps)

The show ip mroute Command

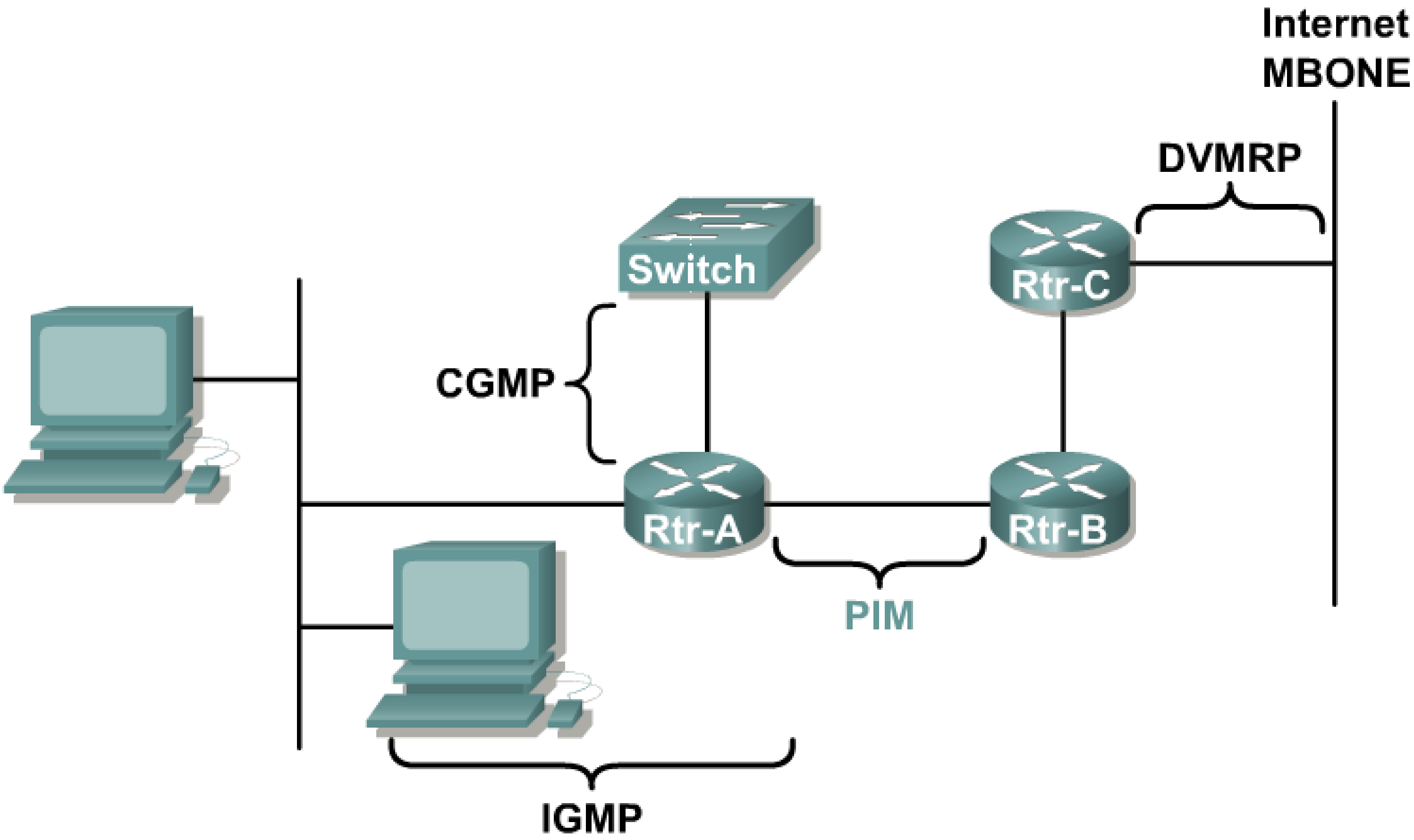
```
NA-1#sh ip mroute
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected
       L - Local, P - Pruned, R - RP-bit set, F - Register flag,
       T - SPT-bit set, J - Join SPT, M - MSDP created entry,
       X - Proxy Join Timer Running, A - Advertised via MSDP, U - URD,
       I - Received Source Specific Host Report
Outgoing interface flags: H - Hardware switched
Timers: Uptime/Expires
Interface state: Interface, Next-Hop or VCD, State/Mode

(*, 224.1.1.1), 00:07:54/00:02:59, RP 10.127.0.7, flags: S
  Incoming interface: Null, RPF nbr 0.0.0.0
  Outgoing interface list:
    Serial1/3, Forward/Sparse, 00:07:54/00:02:32

(172.16.8.1, 224.1.1.1), 00:01:29/00:02:08, flags: TA
  Incoming interface: Serial1/4, RPF nbr 10.139.16.130
  Outgoing interface list:
    Serial1/3, Forward/Sparse, 00:00:57/00:02:02
```

Multicast Routing

Multicast Forwarding

- *Multicast routing has completely different principles than unicast routing!*
  - Unicast routing concerns about where the packet goes to
  - Multicast routing concerns about where the packet goes from
    - Backward path to sender/RP helps in creation of distribution trees
- Multicast routing uses **Reverse Path Forwarding (RPF)** for forwarding loops elimination
  - `IFF` it is received on interface which is in direction to sender/RP in unicast routing table `THEN` multicast packet is processed and forwarded
  - *Receiving interface is along the shortest path to sender/RP*

Protocols Helping Multicast Routing



- Between routers are used protocol **PIM** to manage multicast routing table

Protocol-Independent Multicast (PIM)

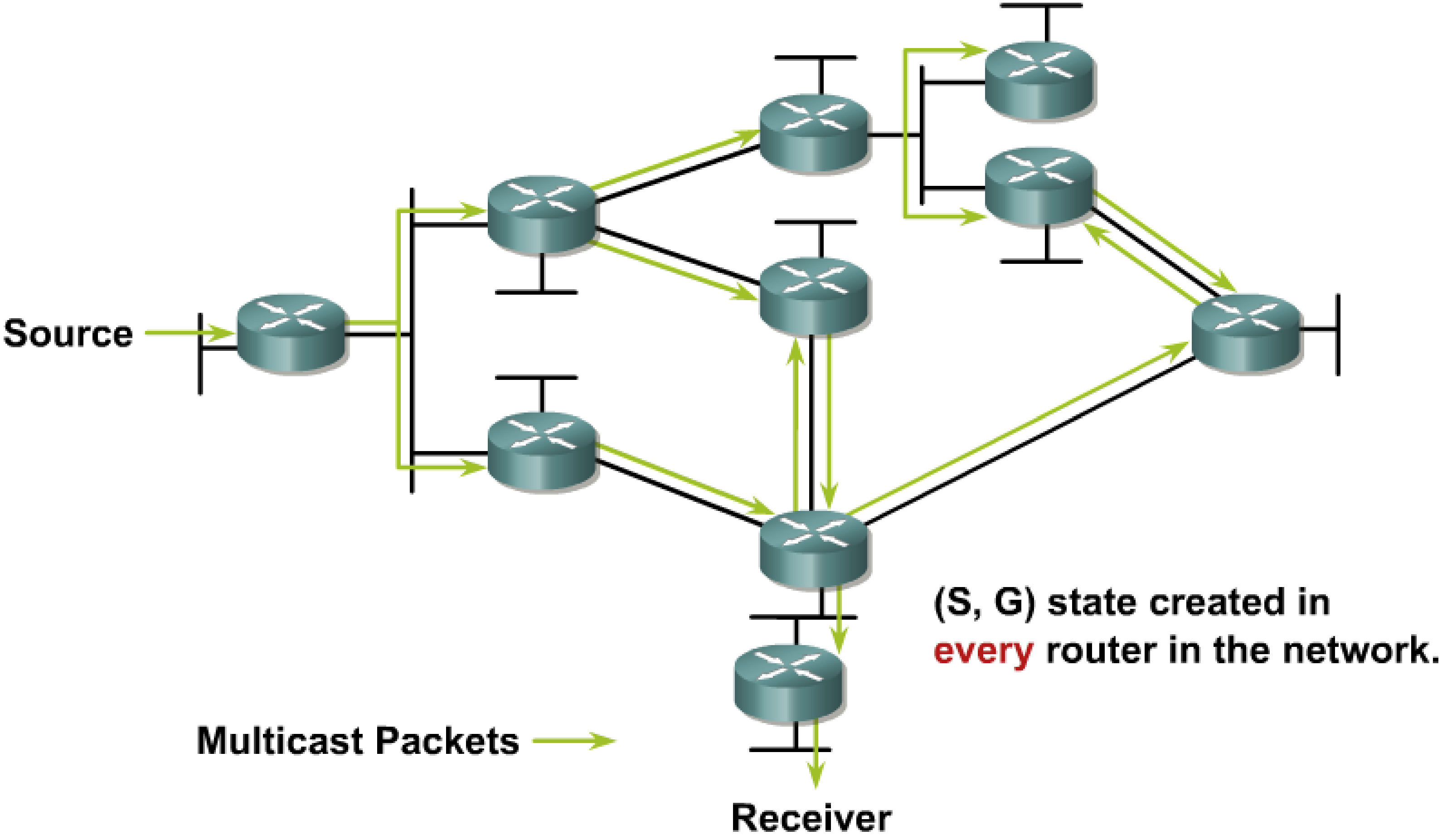
- PIM is actually not a routing protocol which would carried IP prefixes and their metrics – *more less it is signaling protocol*
  - L3 protocol with IP protocol number 103
- PIM needs another unicast routing protocol to be active but it is independent on it – *it doesn't matter whether it is RIP, OSPF or EIGRP*
- PIM routers create multicast routing table to forward multicast datagrams based on unicast RIB
- PIM works in two different regimes by initial design
  - **Dense mode**: Multicast traffic is spread across whole topology. `IF` a router has no multicast members on the some of its segments THEN the router prune itself from multicast distribution tree for target multicast group – a.k.a. periodic flood-and-prune
  - **Sparse mode**: Multicast traffic is sent via distribution trees which are created based on receiving clients requests

PIM and Multicast Distribution Trees

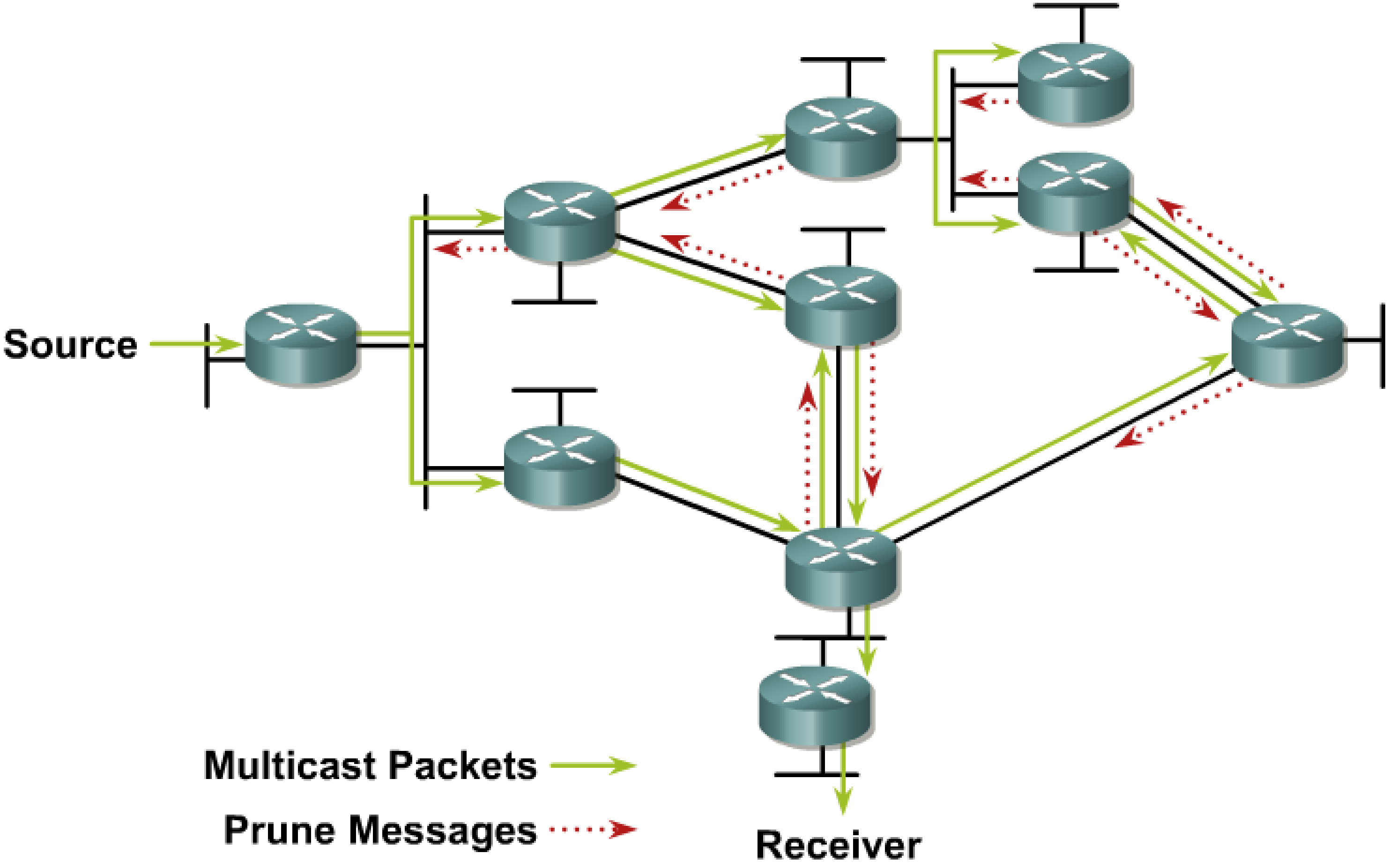
- IF there are more than one PIM router on the local segment THEN as a designated router (DR) is elected one with the
  - highest priority
  - “highest” IP address
- Distribution trees are managed by **PIM Join/Prune** messages
- In case of source trees
  - PIM messages are sent in the direction to sender’s DR
- In case of shared trees
  - PIM messages are sent in the direction to RP

PIM
Dense Mode

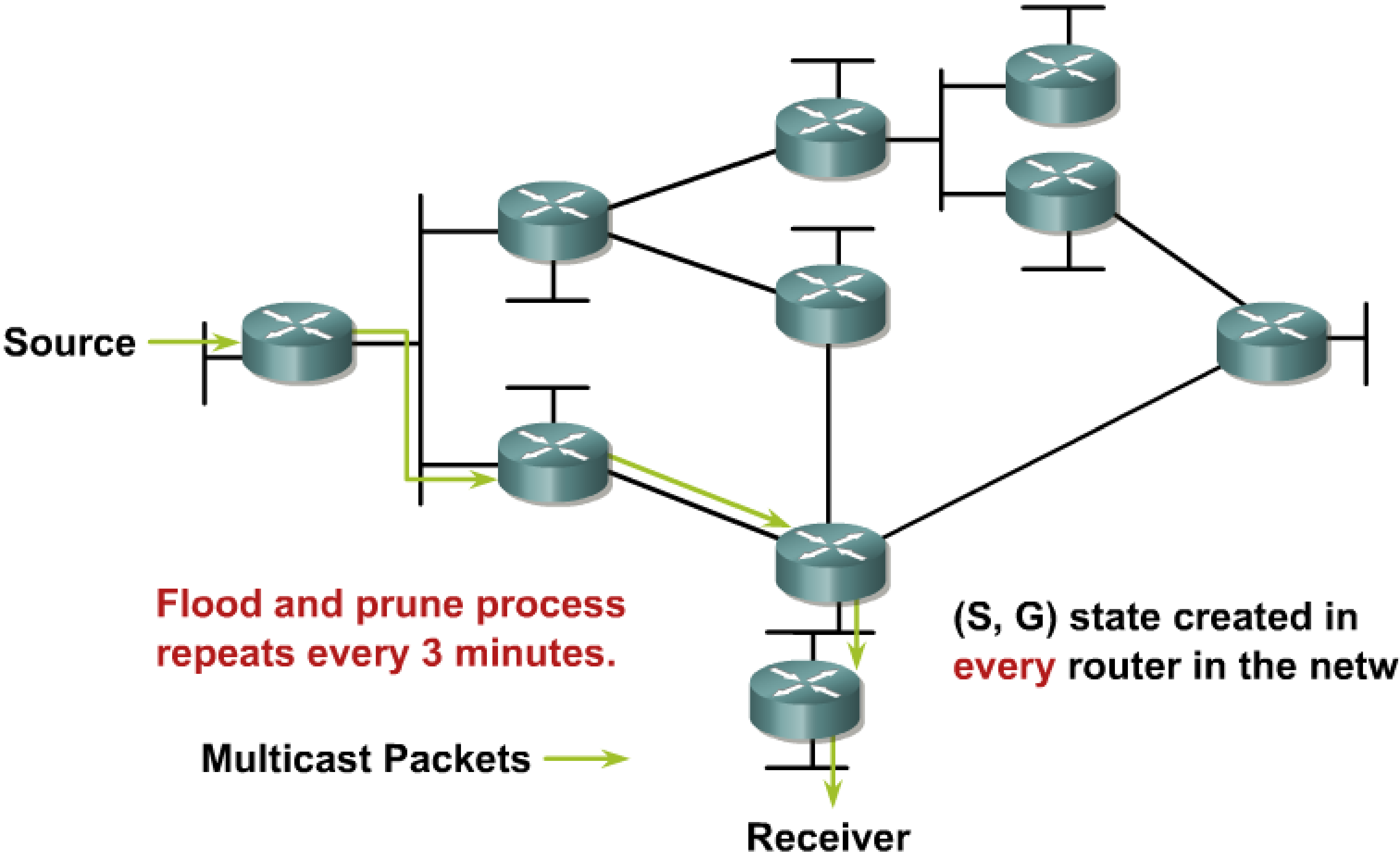
PIM-DM: Initial Flooding



PIM-DM: Pruning



PIM-DM: Converged State



Flood and prune process repeats every 3 minutes.

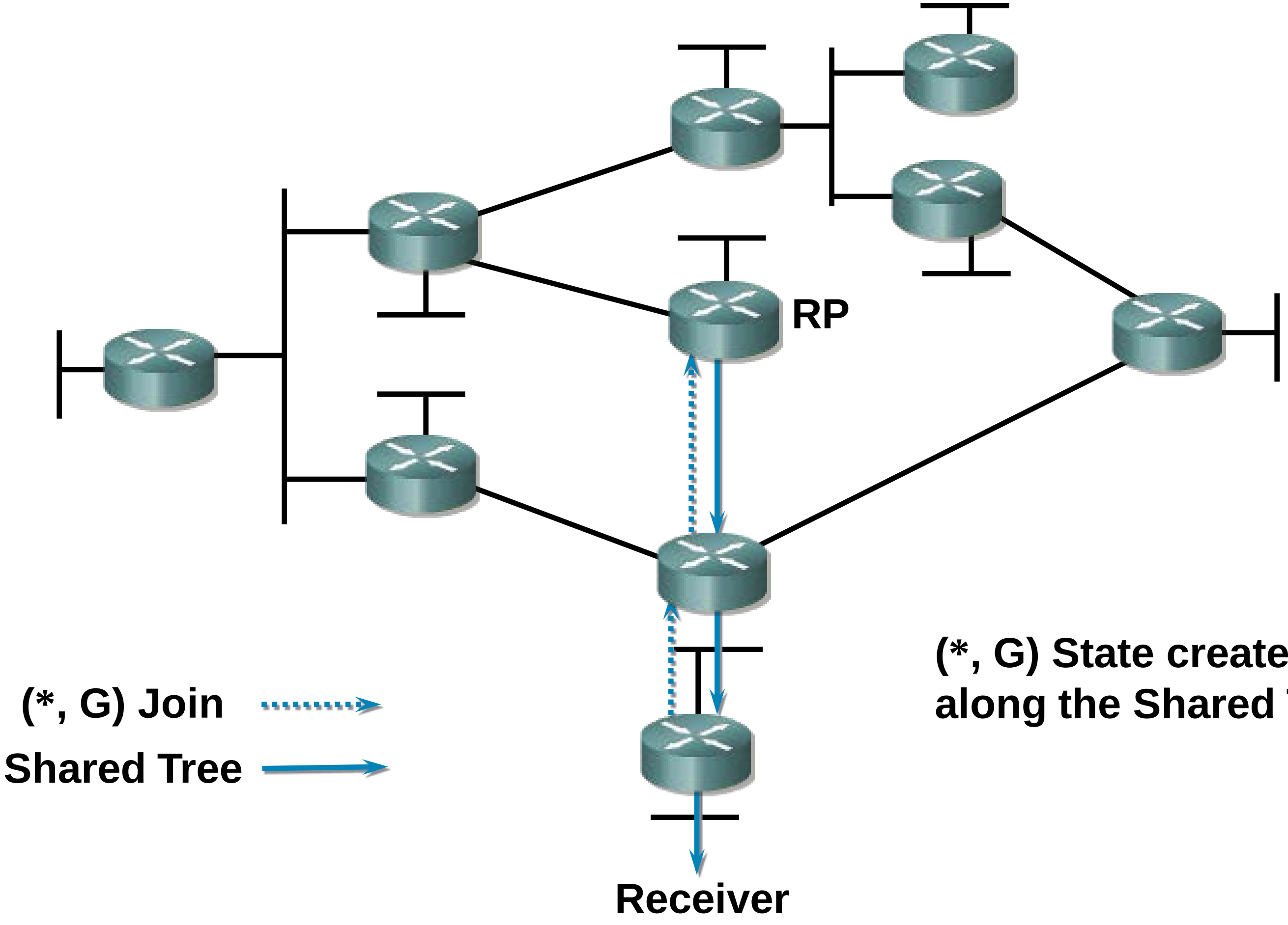
(S, G) state created in every router in the network.

PIM
Sparse Mode

PIM Sparse Mode

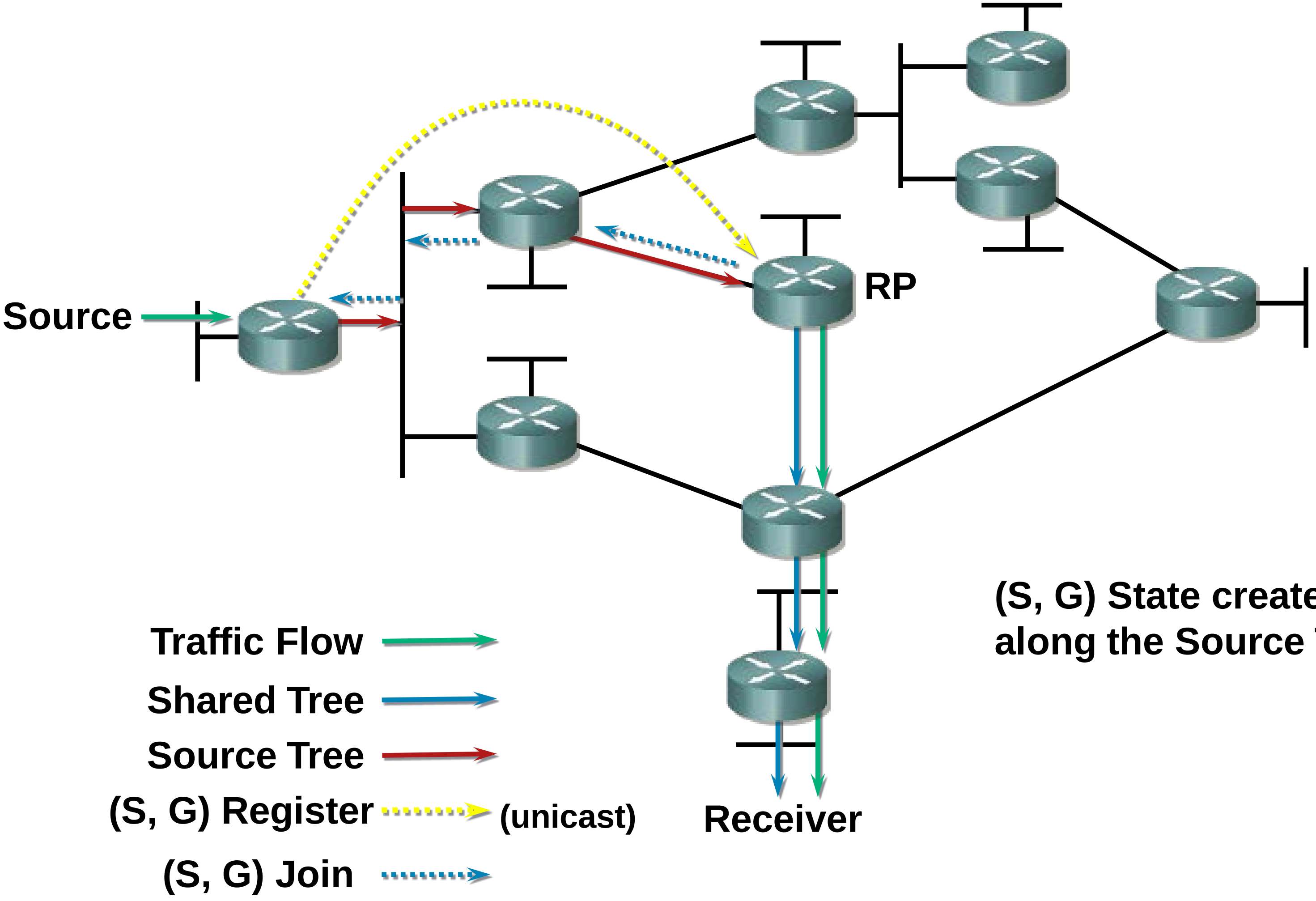
- PIM-SM works with source and also shared trees
  - PIM-DM creates only shortest path trees
- PIM-SM uses **rendezvous point (RP)**
  - *Senders and receivers "meet each other" on agreed point in network*
  - RP is used to coordinate forwarding of multicast traffic from a source to receivers
  - Senders use source tree with RP as leaf via their first-hop routers
  - Receivers use shared trees via theirs DRs

PIM-SM: Shared Tree Join



(*, G) State created only along the Shared Tree.

PIM-SM: Sender Registration ①



Source

RP

Receiver

Traffic Flow 🟩

Shared Tree 🟦

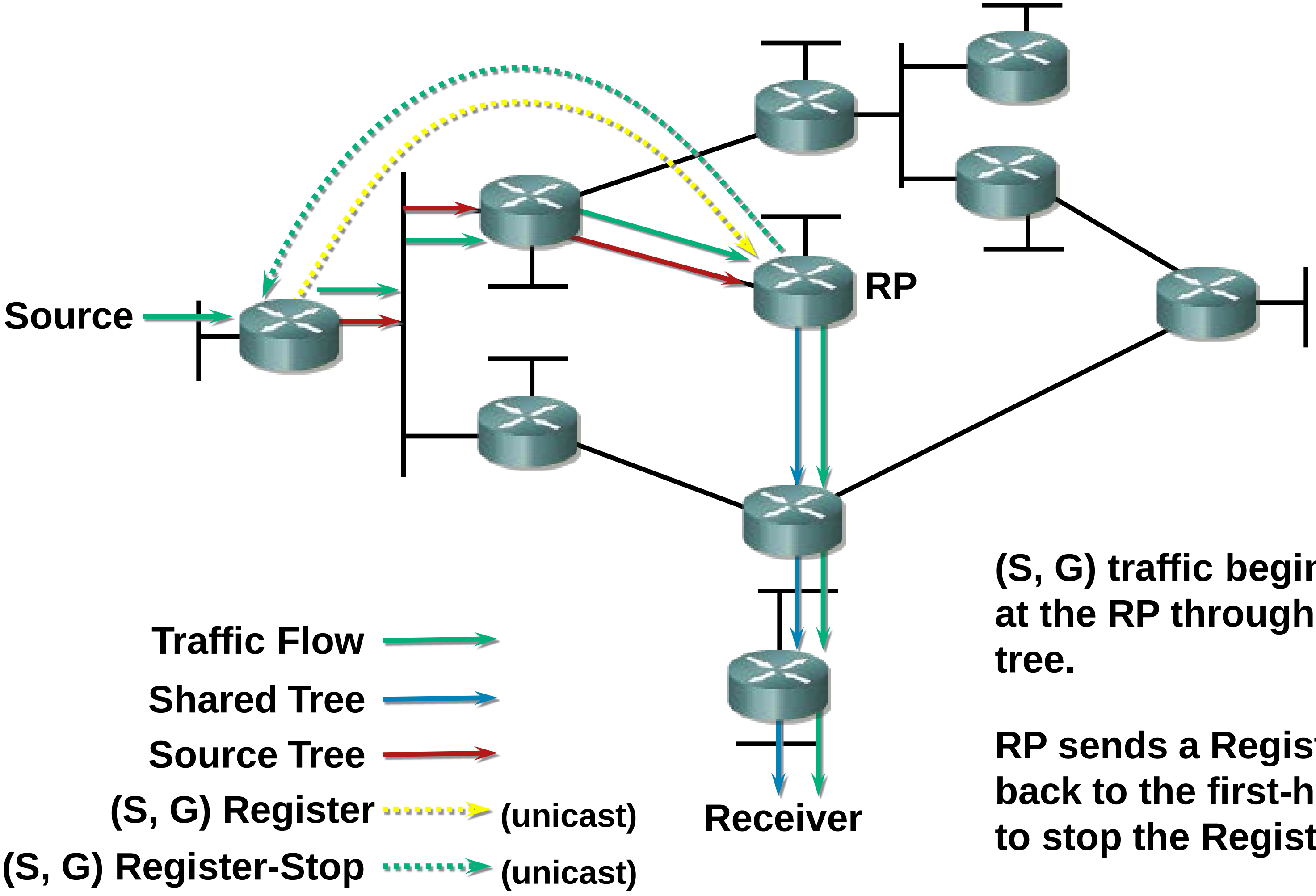
Source Tree 🟥

(S, G) Register ⇢ (unicast)

(S, G) Join ⇢

(S, G) State created only along the Source Tree.

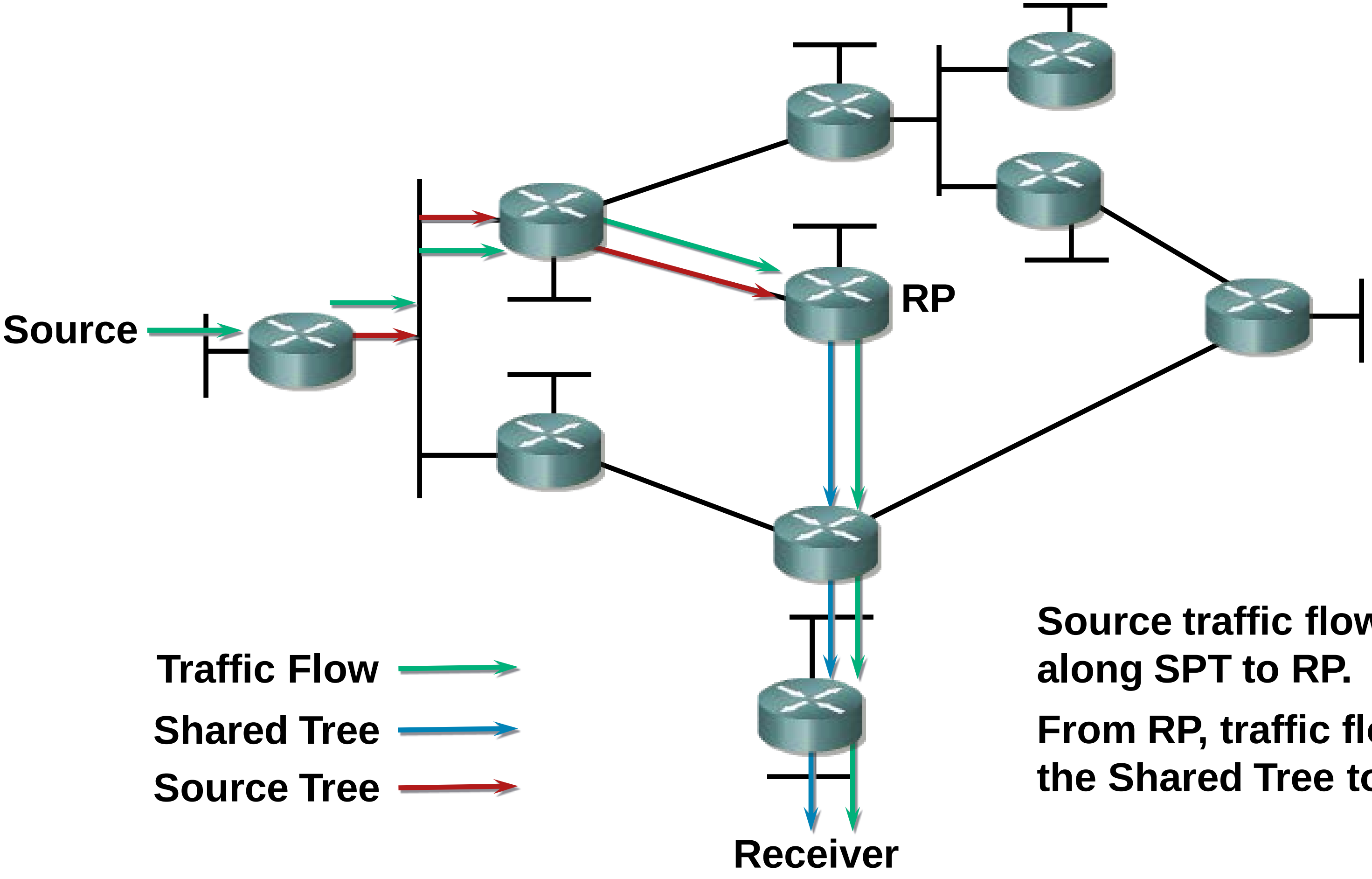
PIM-SM: Sender Registration ②



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

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

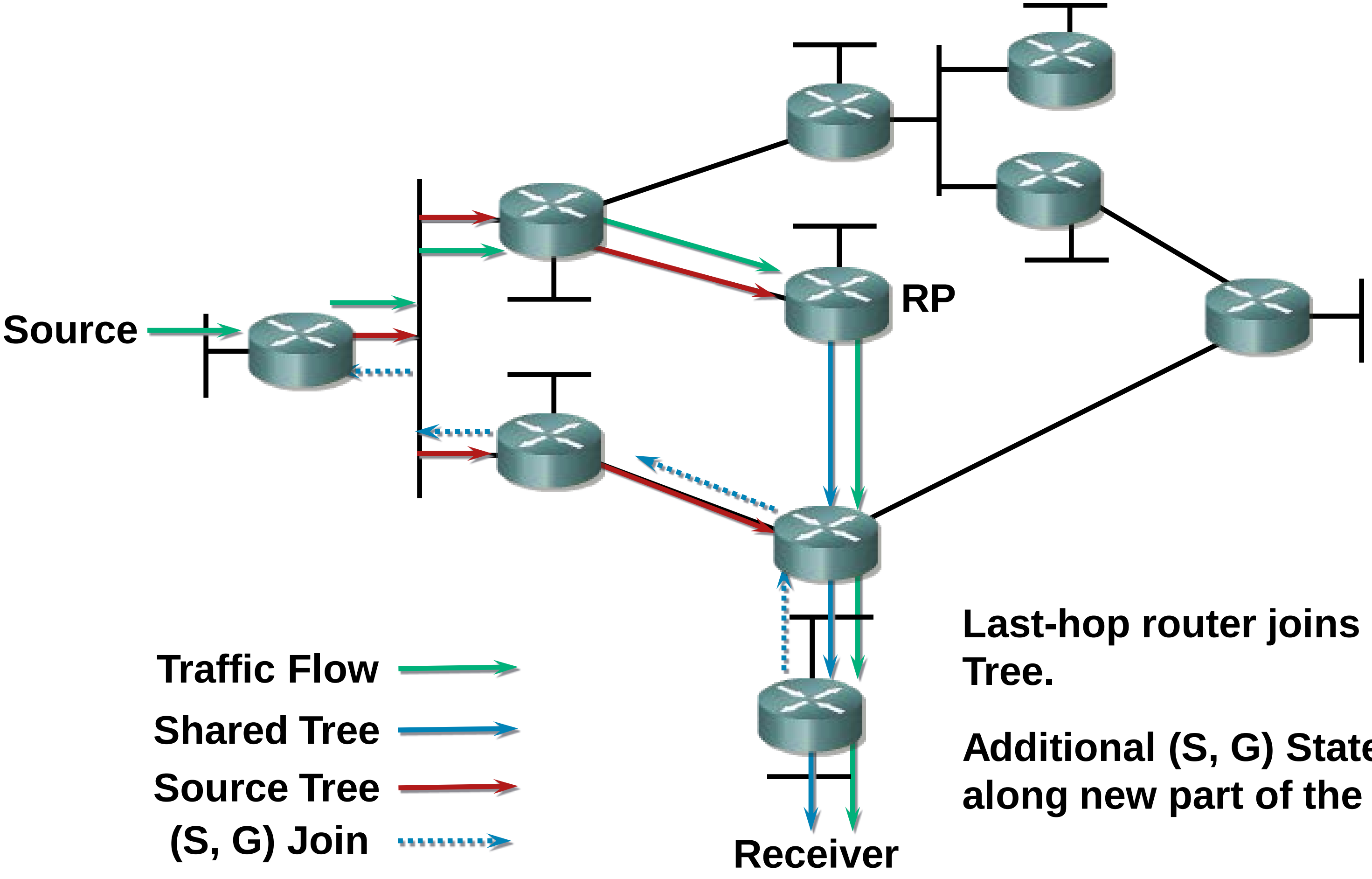
PIM-SM: Sender Registration ③



Source traffic flows natively along SPT to RP.

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

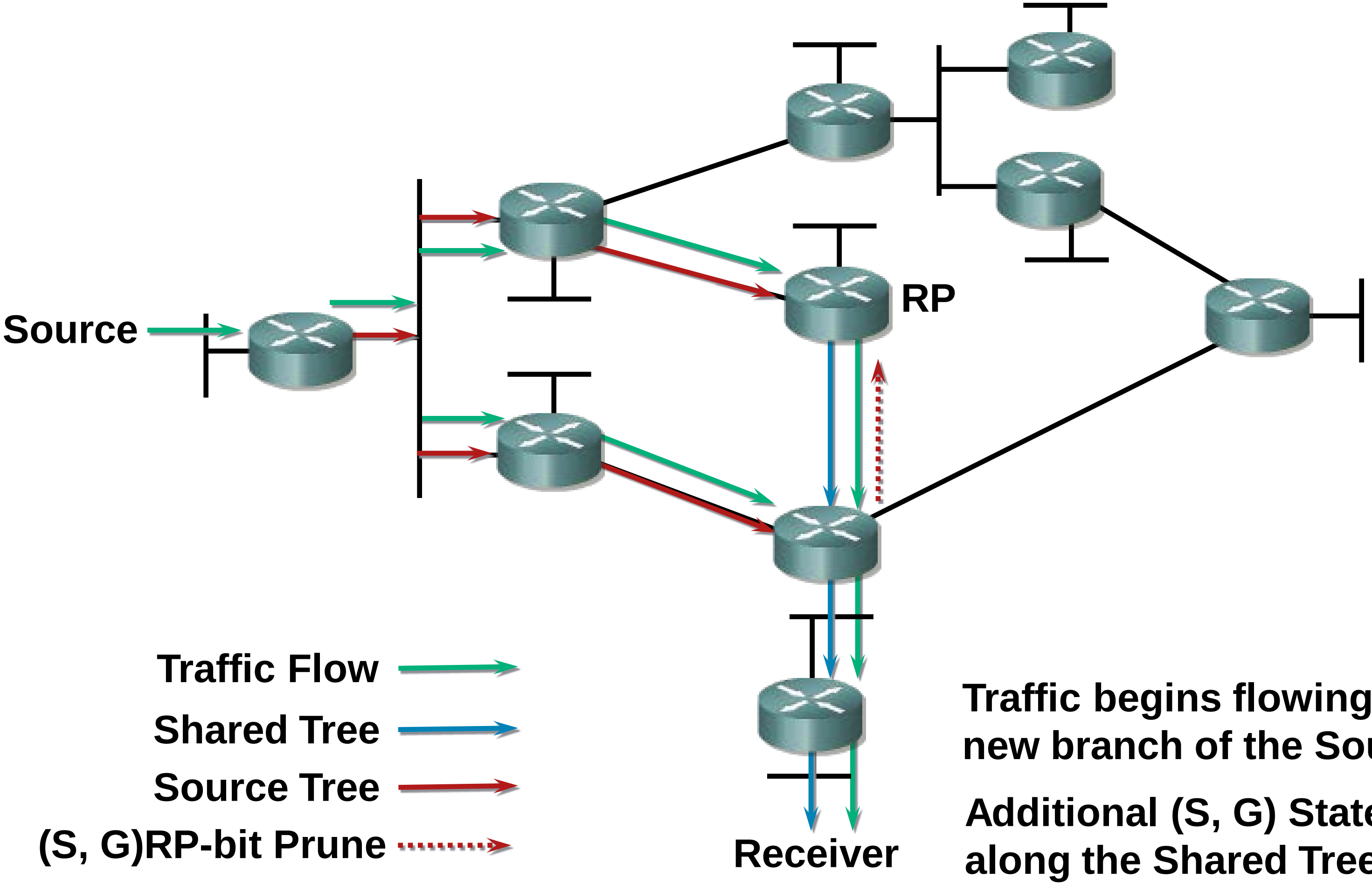
PIM-SM: SPT Switchover ①



Last-hop router joins the Source Tree.

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

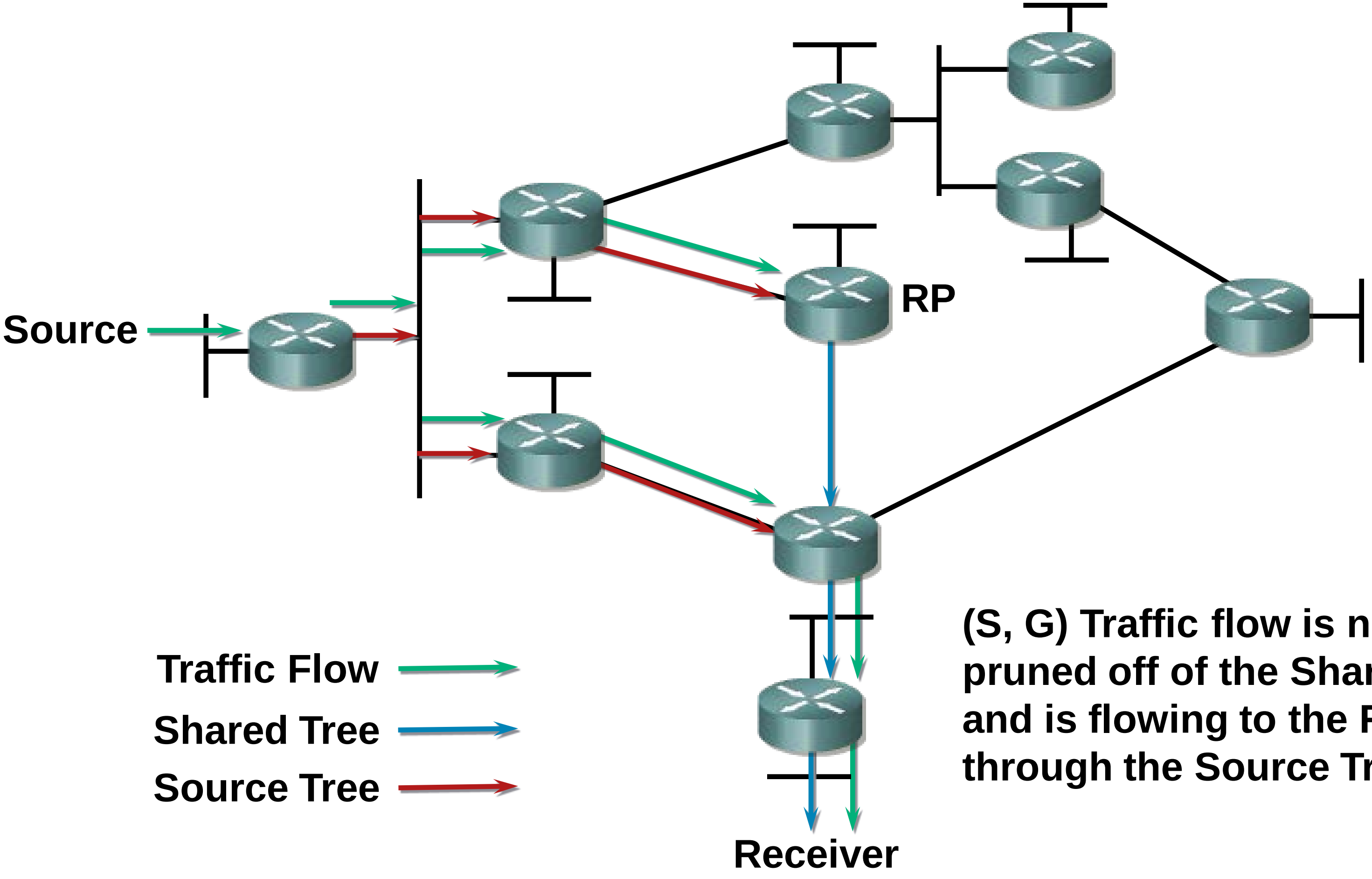
PIM-SM: SPT Switchover ②



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

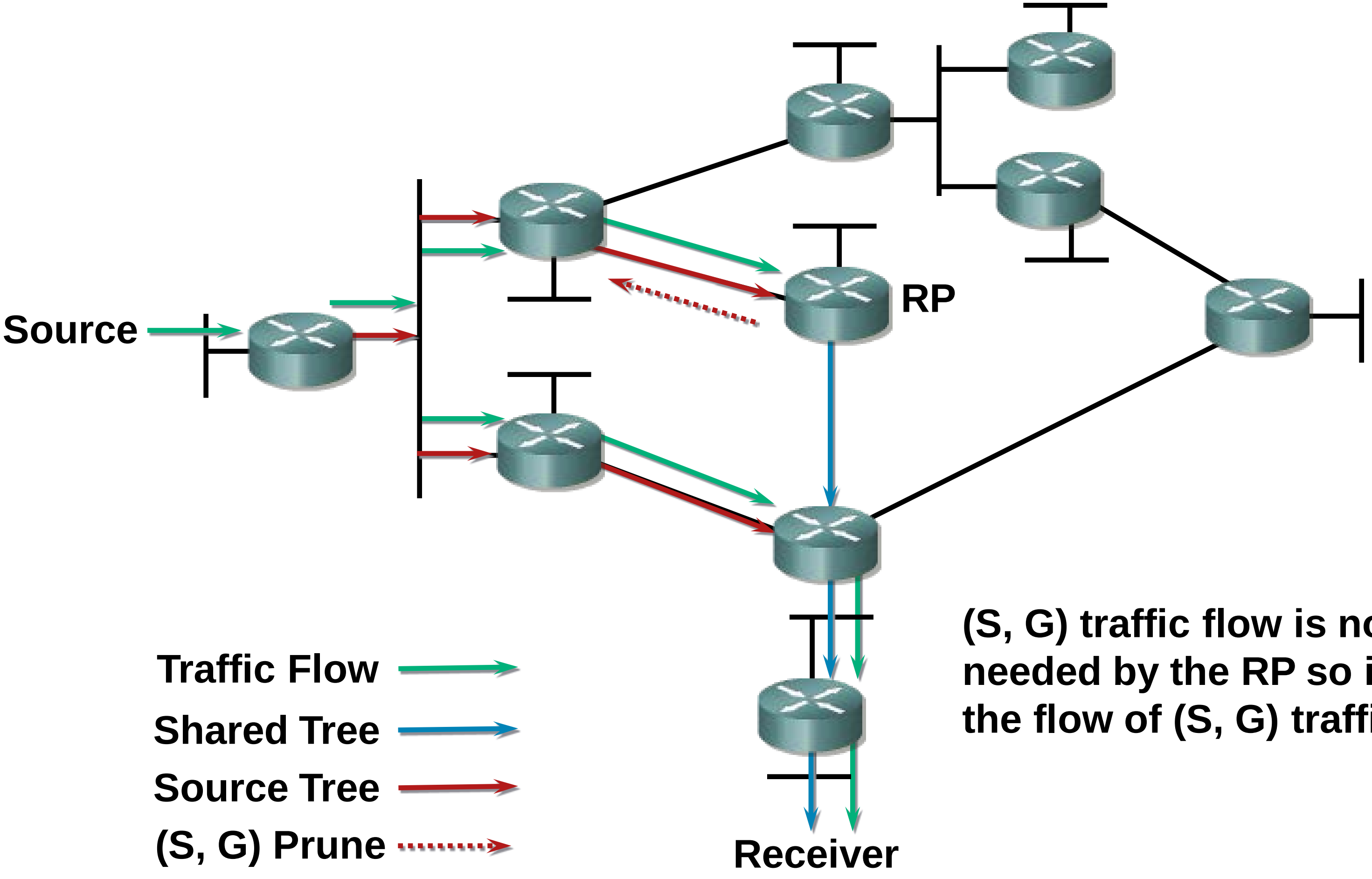
Additional (S, G) State is created along the Shared Tree to prune off (S, G) traffic.

PIM-SM: SPT Switchover ③



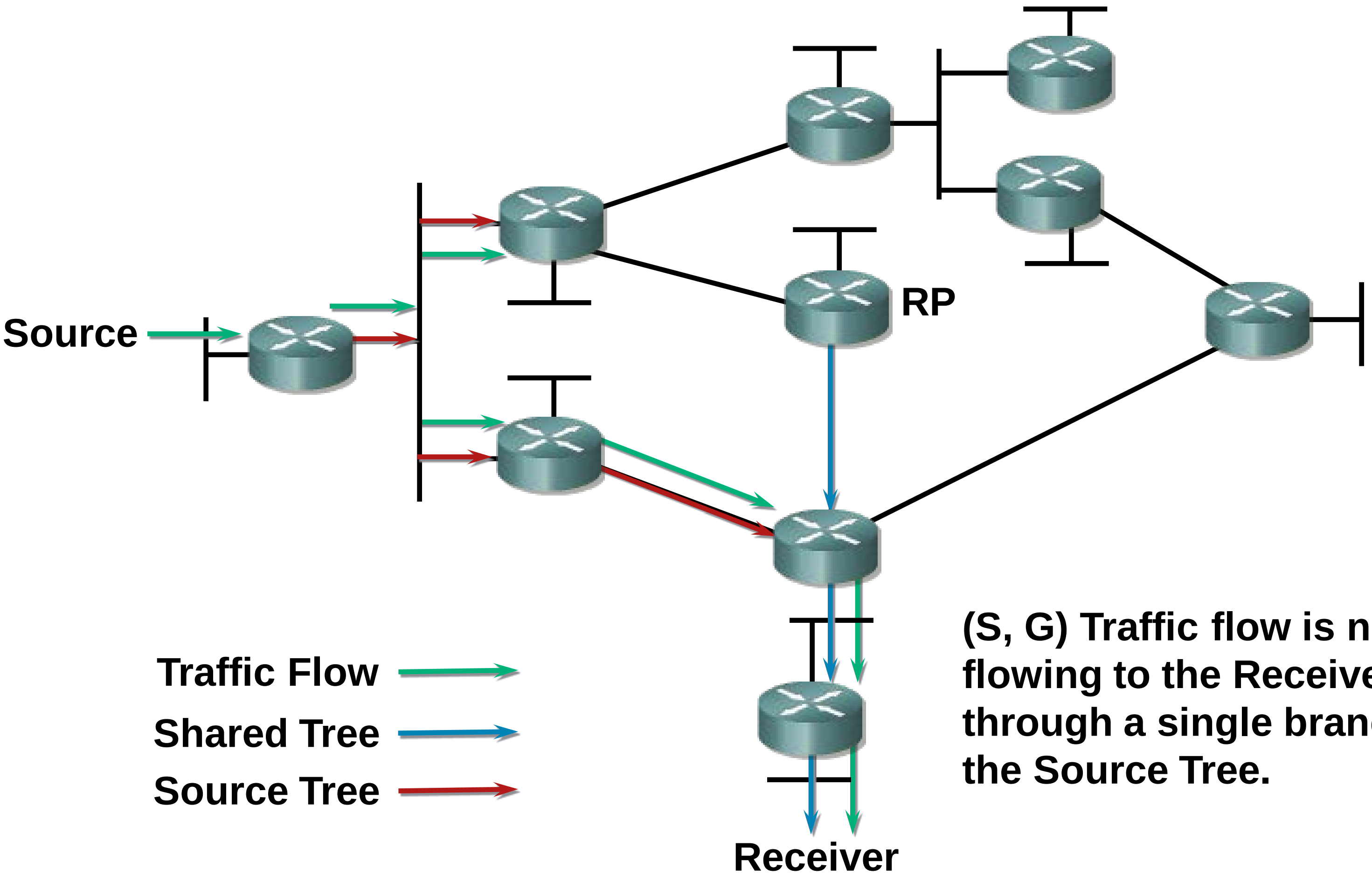
(S, G) Traffic flow is now pruned off of the Shared Tree and is flowing to the Receiver through the Source Tree.

PIM-SM: SPT Switchover ④



(S, G) traffic flow is no longer needed by the RP so it Prunes the flow of (S, G) traffic.

PIM-SM: SPT Switchover ⑤



(S, G) Traffic flow is now only flowing to the Receiver through a single branch of the Source Tree.

Multicast Routing Configuration

Multicast Routing Activation

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

- Activates IPv4 mutlicast routing
- By default this command is disabled; hence, it is necessary to enable it on each router concerning multicast routing

Configuring PIM on Interface

```
Router(config-if)#
  ip pim {sparse-mode | dense-mode | sparse-dense-mode}
```

- Activates PIM on target interface an chooses operating mode
  - Recommended is Cisco sparse-dense-mode where router uses sperse mode if the router knows RP otherwise the router uses dense mode
- Activating PIM automatically activates IGMP on targer interface

Static RP Configuration

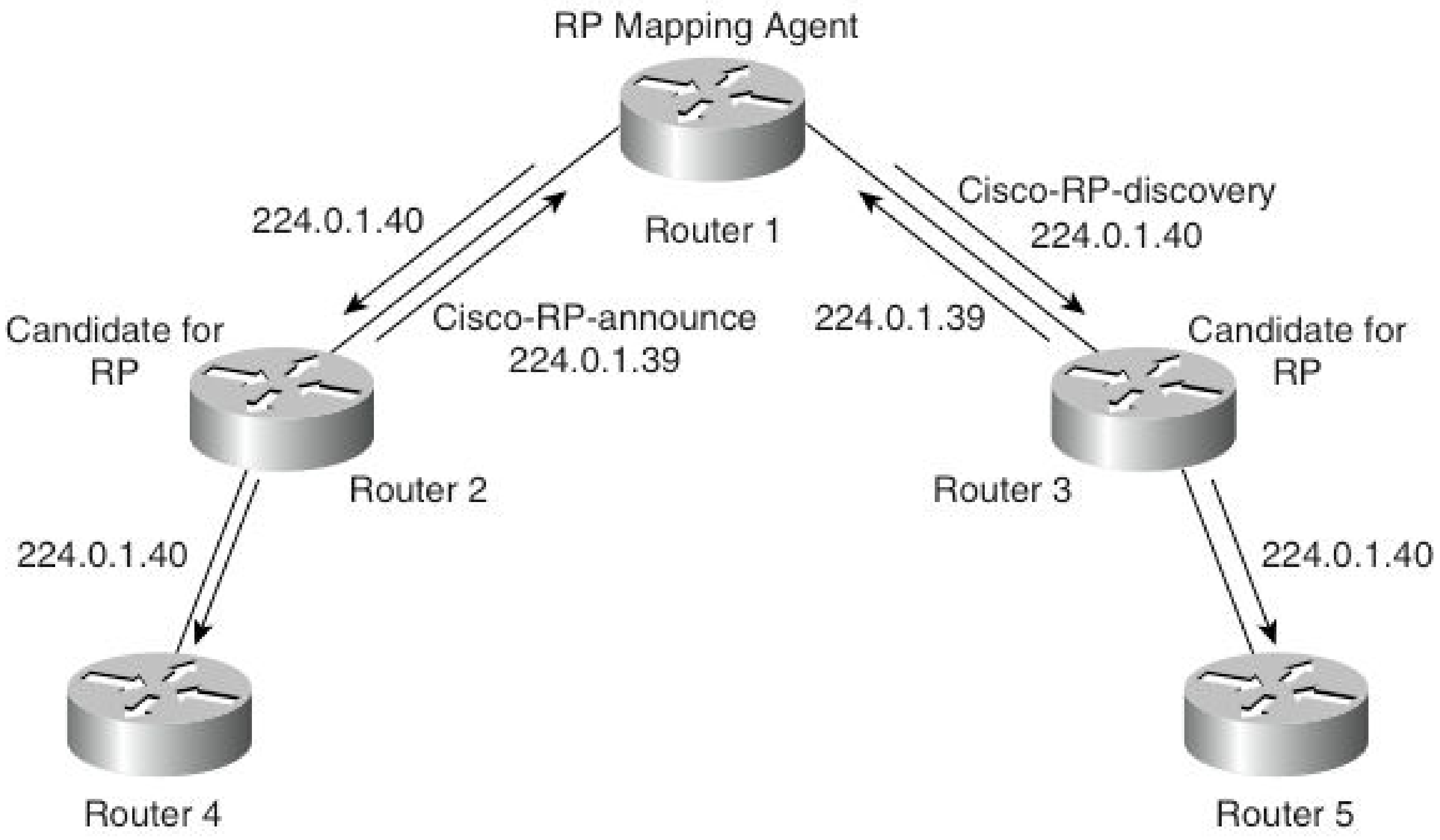
```
Router(config)# ip pim rp-address address [access-list]
```

- When using static RP configuration, it is necessary to configure same command on each and every router in multicast topology
  - In this manner the RP router just appoints oneself
  - Other routers point to RP
  - It is useful to use Loopback address to specify RP
  - By using optional ACL, it could be manipulated to which multicast groups would desired router acts as RP
- *Static configuration is a burden in the large multicast topologies and definitely not an scalable configuration method!*

Auto-RP

- **Auto-RP** is Cisco's own method for automatic discovery and the election of RPs
- Auto-RP has two components
  - **RP Candidate**
    - Routers configured with **`ip pim send-rp-announce`**
    - They announce their willingness to become RP for target multicast group
    - Candidates sends their announcements to 224.0.1.39 to…
  - **Mapping Agents**
    - Routers configured with **`ip pim send-rp-discovery`**
    - They are choosing RP for target multicast group from RP Candidates and sends those mapping to all Auto-RP routers
    - Address 224.0.1.40 is used for communication by all Auto-RP routers

Auto-RP Topology



Configuring RP Candidacy

```
Router(config)#
  ip pim send-rp-announce {interface} scope {ttl} [group-list acl]
```

- Router tries to be RP for allowed multicast group specified in ACL
  - RP Candidacy is advertised to depth of *ttl* number of hops
  - Auto-RP announcements are sent to IP 224.0.1.39 (group name CISCO-RP-ANNOUNCE) on which RP Mapping Agents listen

- *E.g.* Announce this RP Candidate to advertise itself as an RP Candidate for administrative-scope address range using Loopback interface:

```
Router(config)#
  ip pim send-rp-announce Loopback0 scope 16 group-list 1
    access-list 1 permit 239.0.0.0 0.255.255.255
```

RP Mapping Agent

```
Router(config)#
  ip pim send-rp-discovery {interface type} scope {ttl}
```

- RP Mapping Agent is router which collects all announcements from possible RPs and sends the list of this RP-to-group mappings to all Auto-RP routers
  - Auto-RP discovery messages are sent to address 224.0.1.40 (called CISCO-RP-DISCOVERY) which all PIM routers listen to

Bootstrap Router Mechanism

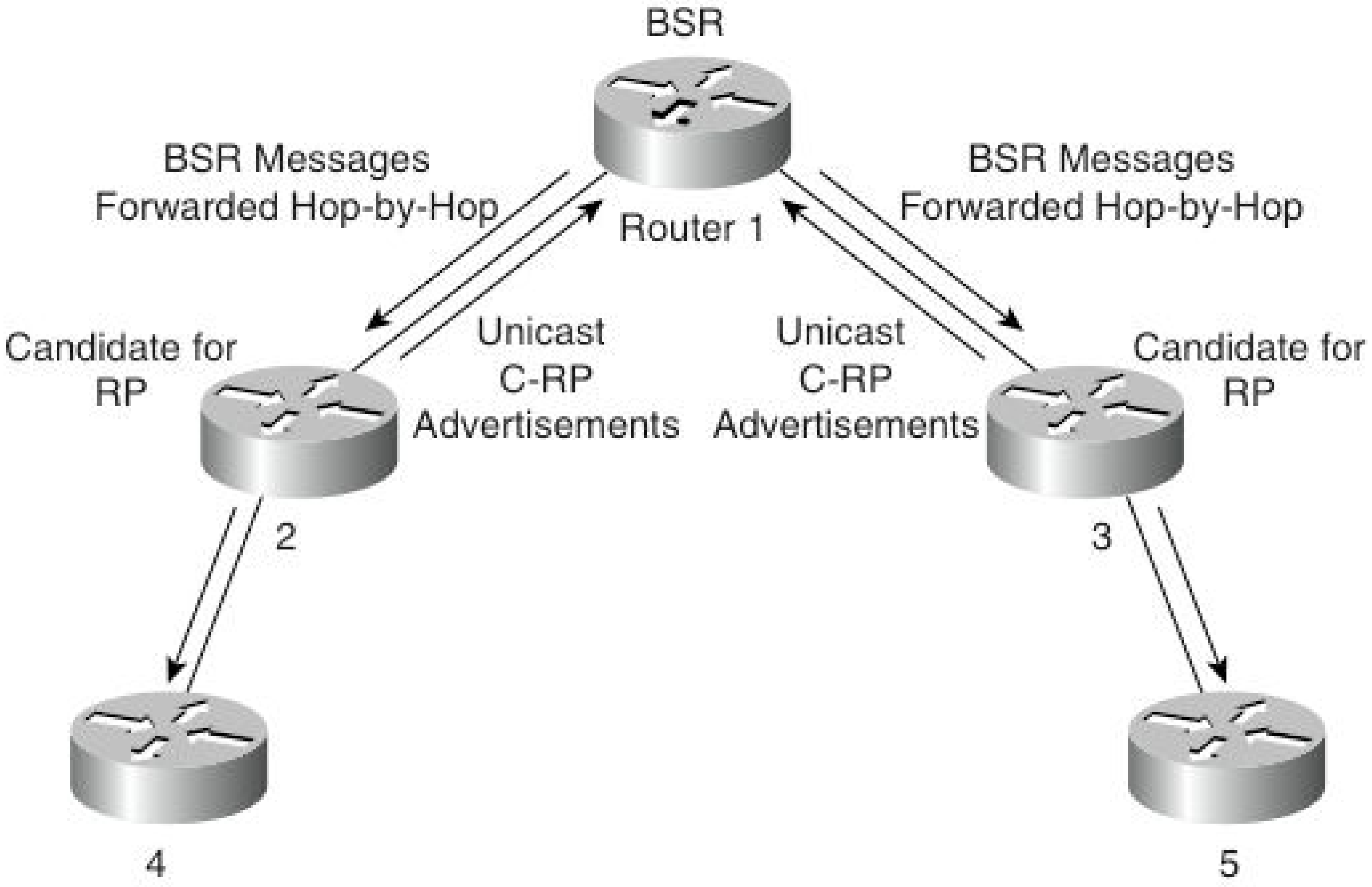
- Auto-RP is Cisco proprietary and does not work in mixed-vendor environment
- Since PIMv2 there is open standard variant of Auto-RP called **BootStrap Router Mechanism (BSRM)**
- RFC 5059
- BSRM is configured analogously
  - Configure **BSR candidate**…

```
(conf-t)# ip pim bsr-candidate interface hash priority
```

  - and **RP candidates**:

```
(conf-t)#
   ip pim rp-candidate IFACE [bidir] [group-list ACL]
      [interval seconds] [priority value]
```

BSRM Topology


BSR
BSR Messages Forwarded Hop-by-Hop
Router 1
BSR Messages Forwarded Hop-by-Hop
Candidate for RP
Unicast C-RP Advertisements
Unicast C-RP Advertisements
Candidate for RP
2
3
4
5

Examples

- PIM-SM in Cisco IOS with RP at 10.20.1.254:

```
Router# conf t
Router(config)# ip multicast-routing
Router(config)# interface vlan 1
Router(config-if)# ip pim sparse-mode
Router(config-if)# interface vlan 2
Router(config-if)# ip pim sparse-mode
Router(config-if)# exit
Router(config)# ip pim rp-address 10.20.1.254
```

- PIM-SM and use router as BSRrouter and also candidate for all private multicast addresses:

```
Router(config)# ip multicast-routing
Router(config)# interface Loopback1
Router(config-if)# ip pim sparse-mode
Router(config)# interface FastEthernet0/0
Router(config-if)# ip pim sparse-mode
Router(config-if)# exit
Router(config)# ip pim bsr-candidate Loopback1 30 200
Router(config)# access-list 1 permit 239.0.0.0 0.255.255.255
Router(config)# ip pim rp-candidate Loopback0 group-list 1
```

Verifying and Troubleshooting

Verifying PIM Neighbors

- Display PIM information relevant to target interface:

```
Router# show ip pim interface [iface-type number] [count]
```

- Shows list of all PIM neighbors:

```
Router# show ip pim neighbor [iface-type number]
```

- Displays information on multicast routers that are peering with the local router:

```
Router# mrinfo [hostname | address]
```

The show ip pim interface Command

```
NA-2# show ip pim interface
Address          Interface          Ver/   Nbr    Query  DR     DR
                                    Mode   Count  Intvl  Prior
10.139.16.133    Serial0/0          v2/S   1      30     1      0.0.0.0
10.127.0.170     Serial1/2          v2/S   1      30     1      0.0.0.0
10.127.0.242     Serial1/3          v2/S   1      30     1      0.0.0.0
```

The show ip pim neighbor Command

```
NA-2# show ip pim neighbor
PIM Neighbor Table
Neighbor          Interface          Uptime/Expires    Ver   DR
Address                                                      Priority
10.139.16.134     Serial0/0          00:01:46/00:01:28 v2    None
10.127.0.169      Serial1/2          00:01:05/00:01:40 v2    1    (BD)
10.127.0.241      Serial1/3          00:01:56/00:01:18 v2    1    (BD)
```

The `mrinfo` Command

- This command shows multicast neighbor router information, router capabilities and code version, multicast interface information, TTL thresholds, metrics, protocol, and status:

```
Router# mrinfo
192.1.7.37 (b.cisco.com) [version cisco 11.1] [flags: PMSA]:
192.1.7.37 -> 192.1.7.34 (s.cisco.com) [1/0/pim]
192.1.7.37 -> 192.1.7.47 (d.cisco.com) [1/0/pim]
192.1.7.37 -> 192.1.7.44 (d2.cisco.com) [1/0/pim]
131.9.26.10 -> 131.9.26.9 (su.bbnplanet.net) [1/32/pim]
```

- The flags in the output indicate:
  - P = prune-capable
  - M = mtrace-capable
  - S = SNMP-capable
  - A = Auto-RP-capable

Verifying RP Configuration

```
Router(config)#
show ip pim rp [group-name | group-address | mapping]
```

- Displays known and active RPs
- **mapping**: With this option all group-to-RP mappings are showed

- To displays RPF information for target source address or RP issue following command:

```
Router(config)# show ip rpf {source-address | name }
```

The show ip pim rp Command

```
P4-2# show ip pim rp
Group: 224.1.2.3, RP: 10.127.0.7, uptime 00:00:20, expires never

P4-2# show ip pim rp mapping
PIM Group-to-RP Mappings

Group(s) 224.0.1.39/32
  RP 10.127.0.7 (NA-1), v1
    Info source: local, via Auto-RP
         Uptime: 00:00:21, expires: never
Group(s) 224.0.1.40/32
  RP 10.127.0.7 (NA-1), v1
    Info source: local, via Auto-RP
         Uptime: 00:00:21, expires: never
Group(s): 224.0.0.0/4, Static
    RP: 10.127.0.7 (NA-1)
```

The `show ip rpf` Command

```
(path in direction to RP)
NA-2# show ip rpf 10.127.0.7
RPF information for NA-1 (10.127.0.7)
  RPF interface: Serial1/3
  RPF neighbor: ? (10.127.0.241)
  RPF route/mask: 10.127.0.7/32
  RPF type: unicast (ospf 1)
  RPF recursion count: 0
  Doing distance-preferred lookups across tables

(path in direction to source)
NA-2# show ip rpf 10.139.17.126
RPF information for ? (10.139.17.126)
  RPF interface: Serial0/0
  RPF neighbor: ? (10.139.16.134)
  RPF route/mask: 10.139.17.0/25
  RPF type: unicast (ospf 1)
  RPF recursion count: 0
  Doing distance-preferred lookups across tables
```

Other PIM Modes

Other PIM Modes

- **PIM Sparse-Dense Mode**
  - Not a different mode, but sort of a behavior
  - IF RP is known to the router THEN use PIM-SM ELSE fallback to PIM-DM
- **PIM Source Specific Multicast (PIM-SSM)**
  - Benefiting from IGMPv3 capable of specifying particular multicast source to receive data from
- **BiDirectional PIM (BiDir PIM)**
  - Senders and receivers communicate

PIM-SSM

- Standardized in RFC3569, reserved addresses 232.0.0.0/8
- Could be deployed in PIM-SM domains
- No need for RPs, hence, it uses only source trees
- (S, G) tuple is called **chanel**
- On routers configure for which multicast groups router should use PIM-SSM with command:

```
Router(config)# ip pim ssm range { default | ACL }
```

- Also on interfaces towards end-station it is needed to enable IGMPv3 support:

```
Router(config-if)# ip igmp version 3
```

- PIM-SSM and SSM mapping

BiDir PIM

- RFC5015 designed for many-to-many applications
- Uses only shared trees and needs RP
- Traffic is sent from source via shared tree through RP to all branches towards receivers
- No source registration (*PIM Register/-Stop messages*) to RP
- Instead of RPF is election of **distinguish forwarder (DF)** where on each segment this router can only forward traffic towards RP to avoid routing loops and multiple data copies
- Globally enable BiDir PIM

```
Router(config)# ip pim bidir-enable
```

- Mark RP to be create bidir shared trees

```
Router(config)#ip pim rp-address address bidir
Router(config)#ip pim rp-candidate IFACE bidir
Router(config)#ip pim send-rp-announce IFACE scope TTL bidir
```

BiDir PIM: Example

```
ip multicast-routing !Enable IP multicast routing
ip pim bidir-enable  !Enable bidir-PIM
!
interface loopback 0
  description One Loopback adddress for this routers Bidir Mode RP function
  ip address 10.0.1.1 255.255.255.0
  ip pim sparse-dense-mode
!
interface loopback 1
  description One Loopback adddress for this routers Sparse Mode RP function
  ip address 10.0.2.1 255.255.255.0
  ip pim sparse-dense-mode
!
ip pim send-rp-announce Loopback0 scope 10 group-list 45 bidir
ip pim send-rp-announce Loopback1 scope 10 group-list 46
ip pim send-rp-discovery scope 10
!
access-list 45 permit 224.0.0.0 0.255.255.255
access-list 45 permit 227.0.0.0 0.255.255.255
access-list 45 deny   225.0.0.0 0.255.255.255
access-list 46 permit 226.0.0.0 0.255.255.255
```

- Cisco IOS Software Releases: “Bidirectional PIM”

IPv6 Multicast

Slides adapted by Vladimír Veselý partially from official course materials but the most of credit goes to CCIE#23527 Ing. Peter Palúch, Ph.D.

The last update: 2013-12-09