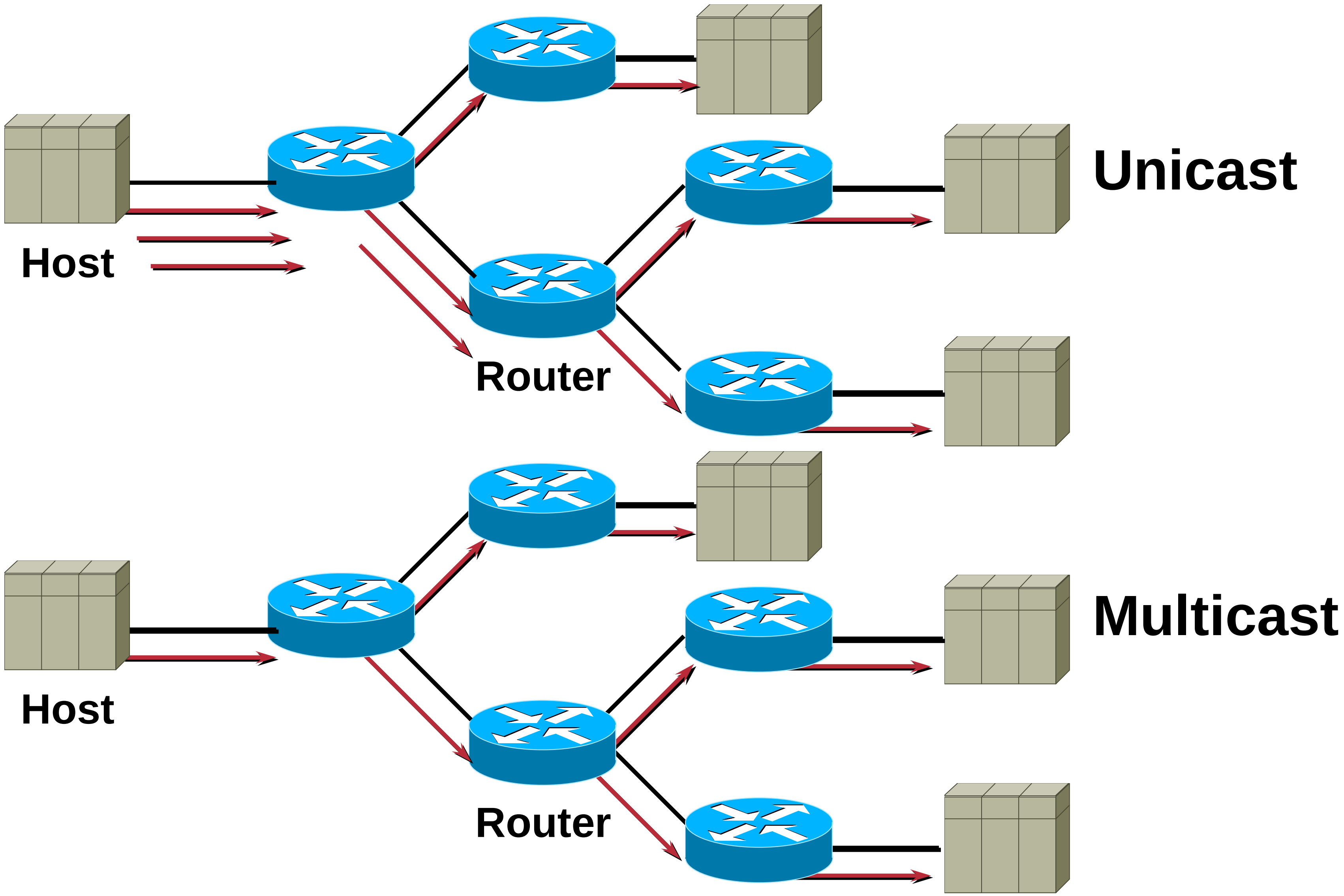
Agenda

- Why Multicast
- Multicast Applications
- Multicast Service Model
- Multicast Distribution Trees
- Multicast Forwarding
- Multicast Protocol Basics

Why Multicast

Multicast Advantages



Multicast Disadvantages

Multicast is UDP Based!!!

- ***Best Effort Delivery*: Drops are to be expected. Multicast applications should not expect reliable delivery of data and should be designed accordingly. Reliable Multicast is still an area for much research. Expect to see more developments in this area.**
- ***No Congestion Avoidance*: Lack of TCP windowing and "slow-start" mechanisms can result in network congestion. If possible, Multicast applications should attempt to detect and avoid congestion conditions.**
- ***Duplicates*: Some multicast protocol mechanisms (e.g. Asserts, Registers and Shortest-Path Tree Transitions) result in the occasional generation of duplicate packets. Multicast applications should be designed to expect occasional duplicate packets.**
- ***Out-of-Sequence Packets:* Various network events can result in packets arriving out of sequence. Multicast applications should be designed to handle packets that arrive in some other sequence than they were sent by the source.**

Multicast Applications

IP Multicast Applications

Live TV and Radio Broadcast
to the Desktop

Corporate Broadcasts

Distance Learning

Multicast File Transfer

Data and File Replication

Training

Video Conferencing

Whiteboard/Collaboration

Video-On-Demand

Real-Time Data Delivery—Financial

Example Multicast Applications

Old Mbone Multicast Applications

- sdr—session directory
  - Lists advertised sessions
  - Launches multicast application(s)
- vat—audio conferencing
  - PCM, DVI, GSM, and LPC4 compression
- vic—video conferencing
  - H.261 video compression
- wb—white board
  - Shared drawing tool
  - Can import PostScript images
  - Uses Reliable Multicast

Multicast Service Model

IP Multicast Service Model

- RFC 1112 (Host Ext. for Multicast Support)
- Each multicast group identified by a class-D IP address
- Members of the group could be present anywhere in the Internet
- Members join and leave the group and indicate this to the routers
- Senders and receivers are distinct:
  i.e., a sender need not be a member
- Routers listen to all multicast addresses and use multicast routing protocols to manage groups

IP Multicast Addressing

- **Multicast Group Addresses (224.0.0.0/4)**
  - **Range: 224.0.0.0 – 239.255.255.255**
  - **Old Class D address range.**
  - **High-order 4 bits are 1110**

Multicast Address Ranges

- **Link-Local Address Range**
  - **224.0.0.0 – 224.0.0.255**
- **Global Address Range**
  - **224.0.1.0 – 238.255.255.255**
- **Administratively Scoped Address Range**
  - **239.0.0.0 – 239.255.255.255**
- **Scope Relative Address Range**
  - **Top 256 addresses of a Scoped Address Range.**

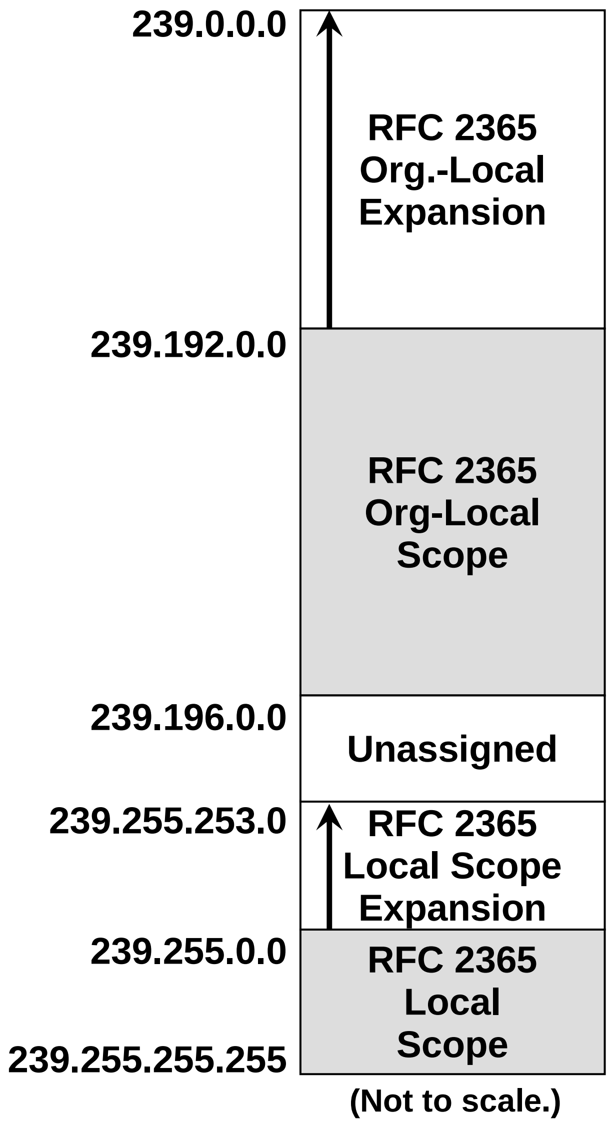
Link-Local Address Range

- **Assigned by IANA**
  - **224.0.0.0 – 224.0.0.255**
  - **Local wire multicast**
    - **TTL = 1**
  - **Examples:**
    - **224.0.0.5 = OSPF_DR's**
    - **224.0.0.10 = EIGRP Hello's**
    - **224.0.0.13 = All_PIM_Routers**
    - **224.0.0.22 = All_IGMPv3_Routers**

Global Address Range

- Assigned by IANA
  - Address Range: 224.0.1.0 – 238.255.255.255
  - Generally intended for “global” Internet scope multicast.
  - Sometimes assigned to specific protocols.
    - Example: Auto-RP (224.0.1.39 & 224.0.1.40)
  - Problem:
    - IANA is coming under increasing pressure from companies to assign them blocks of addresses for their applications or content services.
    - *This was never the intent of this block!*
      - GLOP Addressing or SSM should be used instead!

Administratively Scoped Address Range



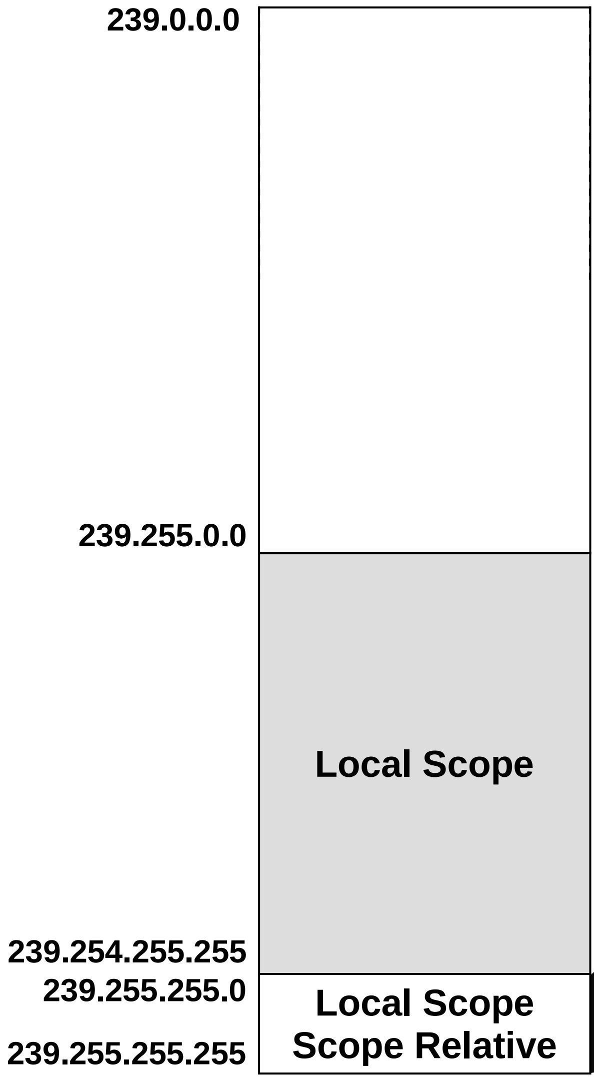
- **Address Range: 239.0.0.0/8**
  - **Private multicast address space.**
  - **Similar to RFC1918 private unicast address space.**
- **RFC 2365 Administratively Scoped Zones.**
  - **Organization-Local Scope (239.192/14)**
    - **Largest scope within the Enterprise network (i.e. Enterprise-wide).**
    - **Expands downward in address range.**
  - **Local Scope (239.255/16)**
    - **Expands downward in address range.**
    - **Smallest possible scope within the Enterprise network.**
    - **Other scopes may be equal but not smaller.**

Scope Relative Addresses – RFC 2365

Top 256 Addresses of every Admin. Scope Range.

Last Octet	Offset	Description
.255	0	SAP Session Announcement Protocol (SDR)
.254	1	MADCAP Protocol
.253	2	SLPv2 Protocol
.252	3	MZAP Protocol
.251	4	Multicast Discovery of DNS Services
.250	5	SSDP
.249	6	DHCPv4
.248	7	AAP
.247	8	MBUS
	9 - 255	Unassigned

Scope Relative Example – Local Scope



(Not to scale.)

Address	Description
239.255.255.255	SAP Session Announcement Protocol (SDR)
239.255.255.254	MADCAP Protocol
239.255.255.253	SLPv2 Protocol
239.255.255.252	MZAP Protocol
239.255.255.251	Multicast Discovery of DNS Services
239.255.255.250	SSDP
239.255.255.249	DHCPv4
239.255.255.248	AAP
239.255.255.247	MBUS

Global Multicast Address Assignment

- Dynamic Group Address Assignment
  - Historically accomplished using SDR application
    - Sessions announced over well-known group(s).
      - Address collisions detected and resolved at session creation time.
      - Has problems scaling.
  - Other techniques considered.
    - Multicast Address Set-Claim (MASC)
      - Hierarchical, dynamic address allocation scheme.
      - Unlikely to be deployed.
  - No really good dynamic assignment method available for Global multicast.

Global Multicast Address Assignment

- Static Group Address Assignment
  - RFC 3180 – GLOP Addressing in 233/8
    - Group range: 233.0.0.0 - 233.255.255.255
      - Your AS number is inserted in middle two octets
      - Remaining low-order octet used for group assignment.
    - EGLOP Addresses
      - Make use of private AS numbers.
      - Assigned by a Registration Authority.

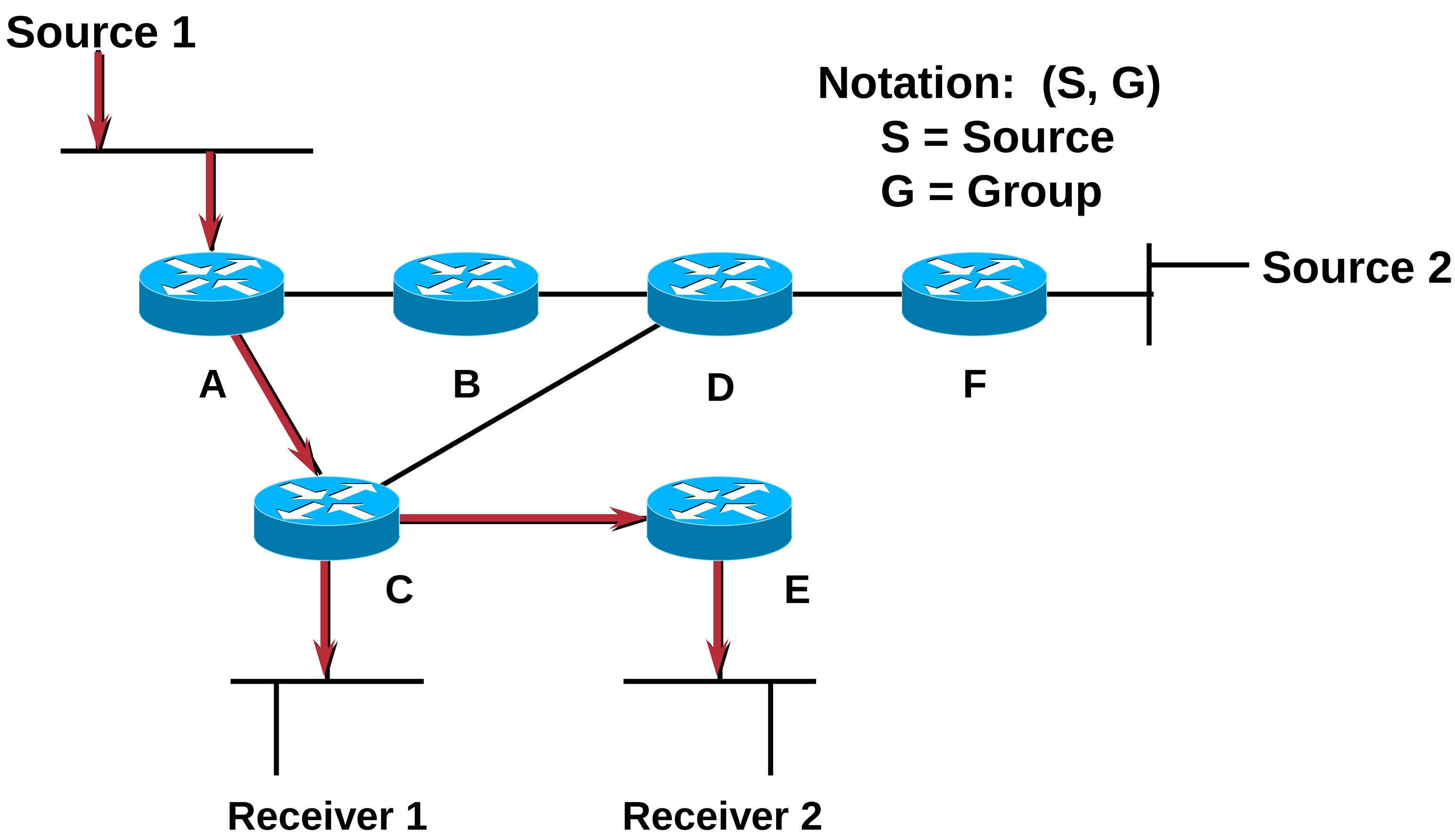
Global Multicast Address Assignment

- **Static Group Address Assignment**
  - **Source Specific Multicast**
    - **Address range: 232.0.0.0/8**
    - **Flows based on both Group *and* Source address.**
      - **Two different content flows can share the same Group address without interfering with each other.**
    - ***Provides virtually unlimited address space!***
      - **Preferred method for Global One-to-Many multicast.**

Multicast Distribution Trees

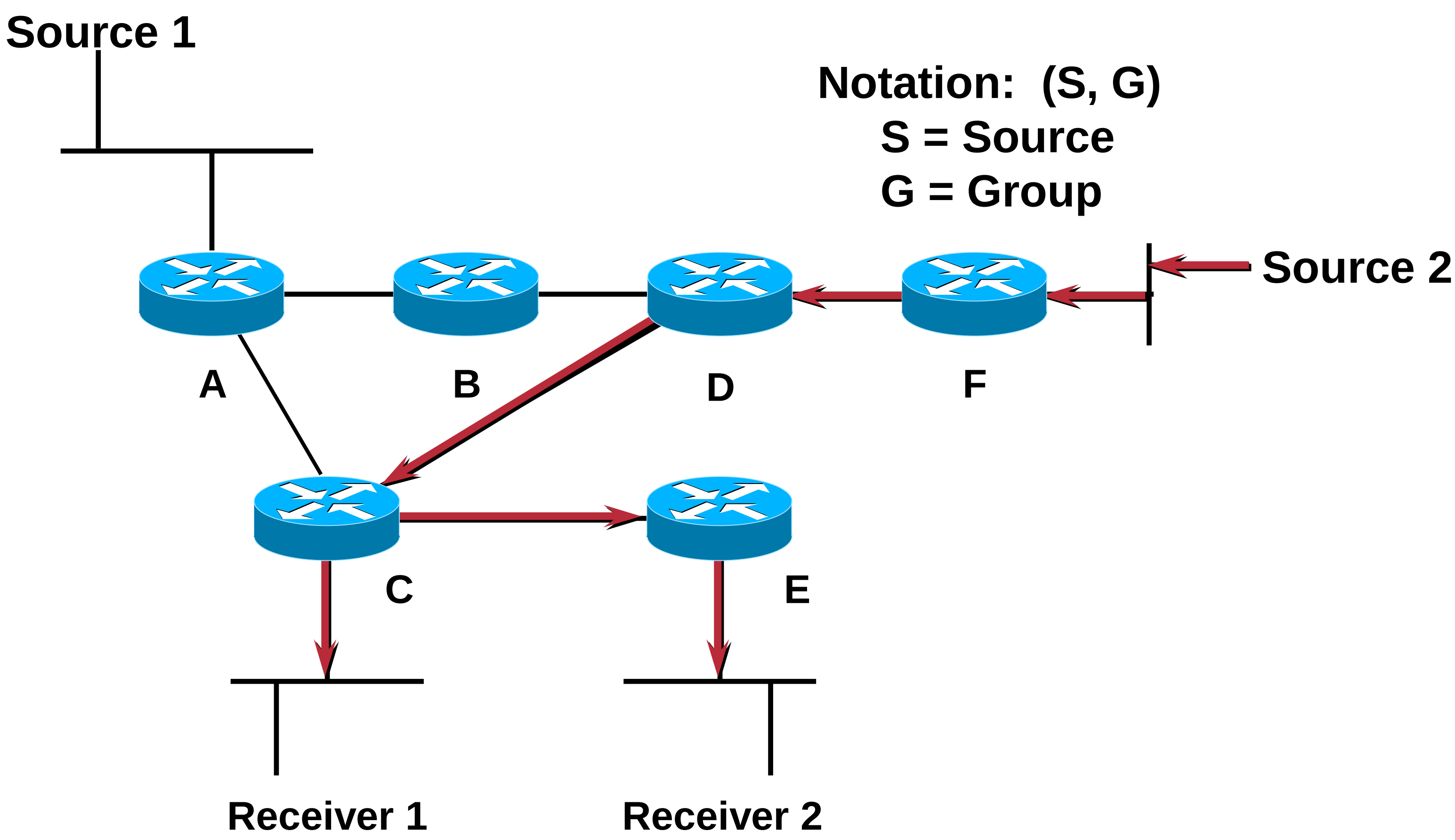
Multicast Distribution Trees

Shortest Path or Source Tree



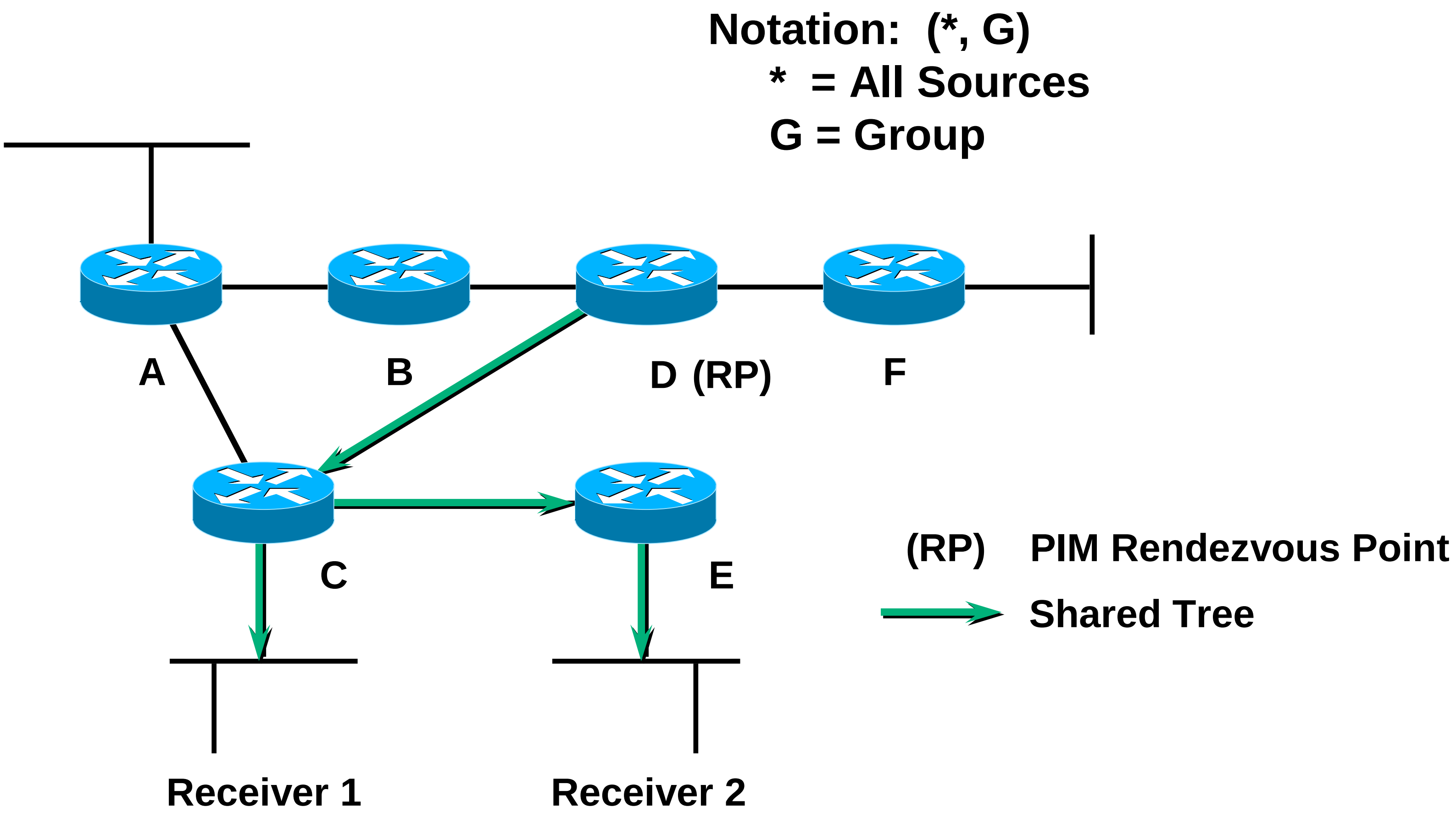
Multicast Distribution Trees

Shortest Path or Source Tree



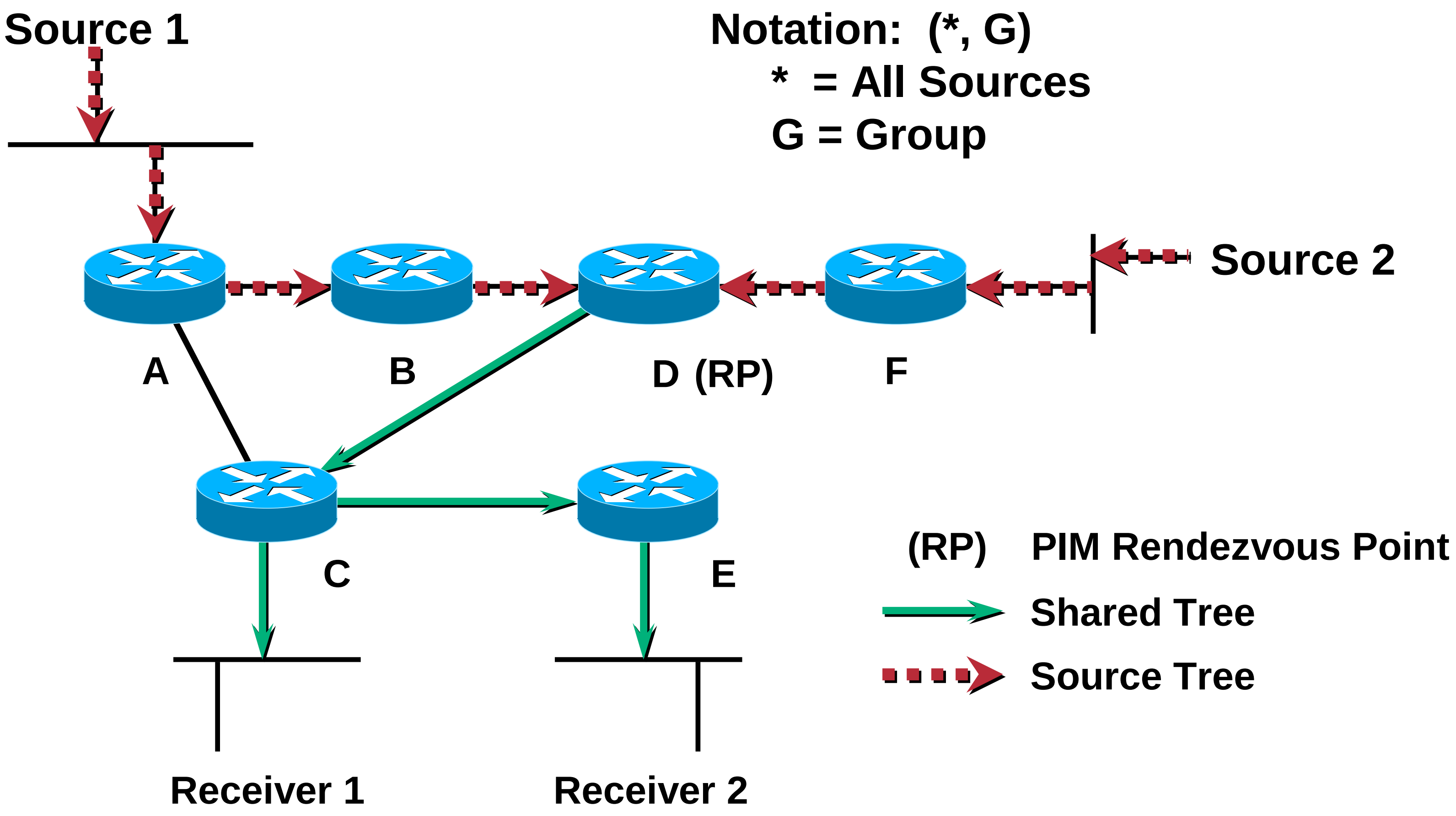
Multicast Distribution Trees

Shared Tree



Multicast Distribution Trees

Shared Tree



Multicast Distribution Trees

Characteristics of Distribution Trees

- **Shortest Path trees**
  - **Uses more memory O(S x G) but you get optimal paths from source to all receivers; minimizes delay**
- **Shared trees**
  - **Uses less memory O(G) but you may get sub-optimal paths from source to all receivers; may introduce extra delay**

Multicast Forwarding

Unicast vs. Multicast Forwarding

- **Unicast Forwarding**
  - **Destination IP address directly determines where to forward the packet.**
    - **Decision based on route table.**
    - **Hop-by-hop forwarding continues even during routing topology changes.**

Unicast vs. Multicast Forwarding

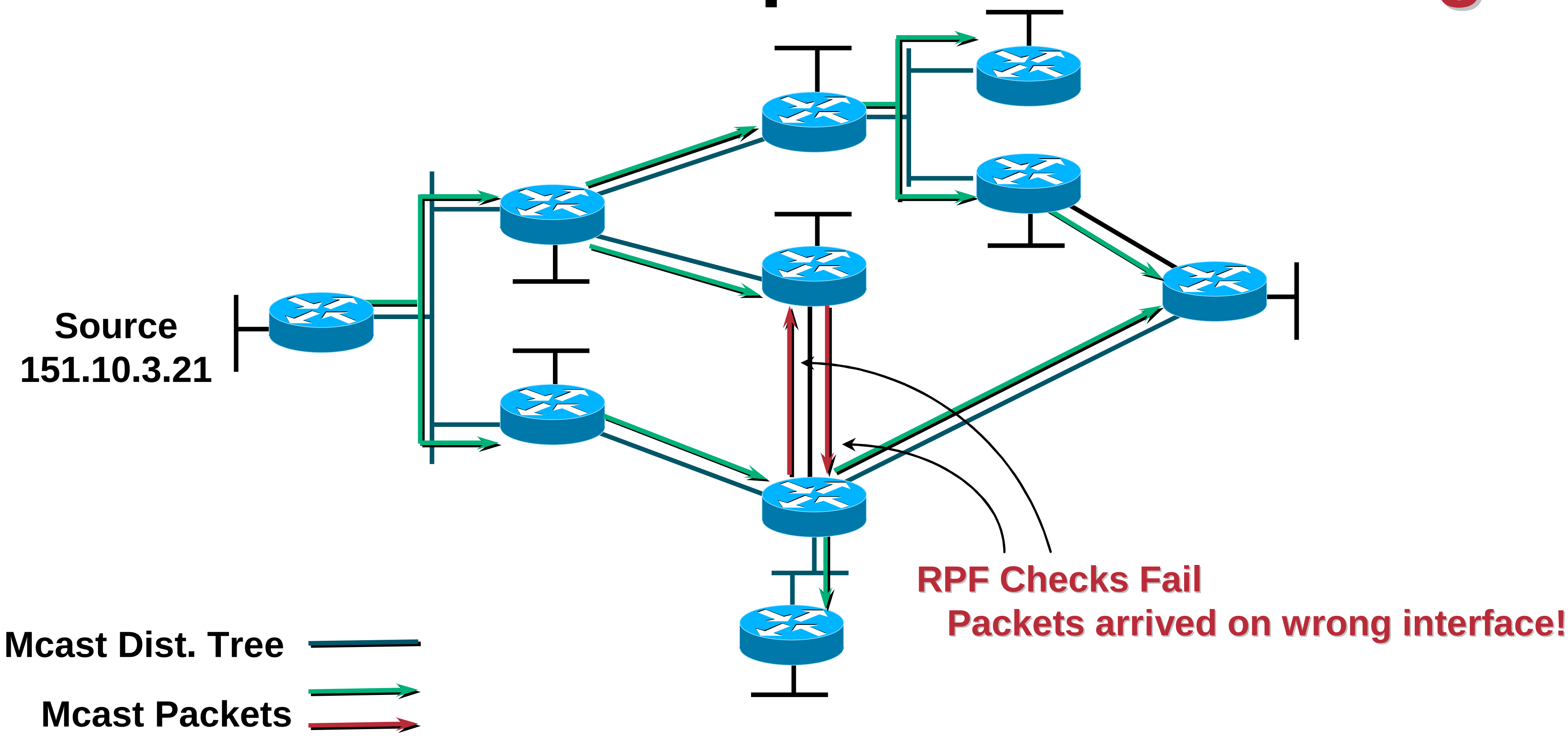
- **Multicast Forwarding**
  - Destination IP address doesn't directly indicate where to forward packet.
  - Forwarding is connection-oriented.
    - Receivers must first "connect" to the source before traffic begins to flow.
      - Connection messages follow unicast routing table toward multicast source.
      - Build Multicast Distribution Trees that determine where to forward packets.
      - Distribution Trees rebuilt dynamically in case of network topology changes.

Reverse Path Forwarding (RPF)

- **The RPF Calculation**
  - **The multicast source address is checked against the unicast routing table.**
  - **This determines the interface and upstream router in the direction of the source to which PIM Joins are sent.**
  - **This interface becomes the “Incoming” or RPF interface.**
    - **A router forwards a multicast datagram only if received on the RPF interface.**

Reverse Path Forwarding (RPF)

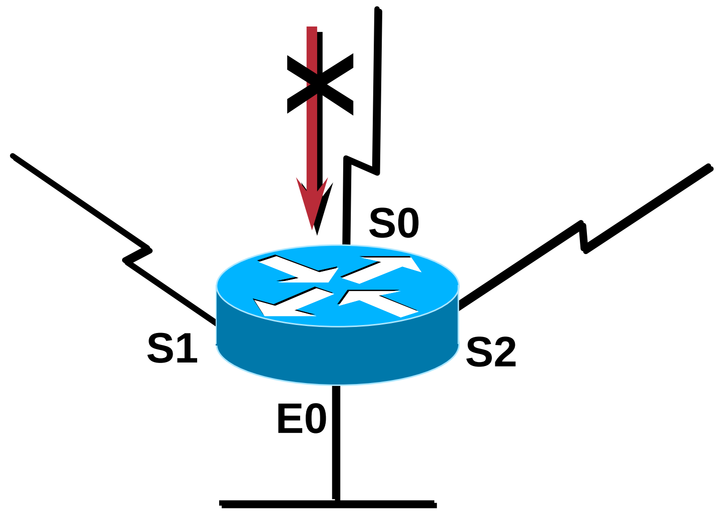
Dense Mode Example: RPF Checking



Reverse Path Forwarding (RPF)

A closer look: RPF Check Fails

Multicast Packet from
Source 151.10.3.21



RPF Check Fails!

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

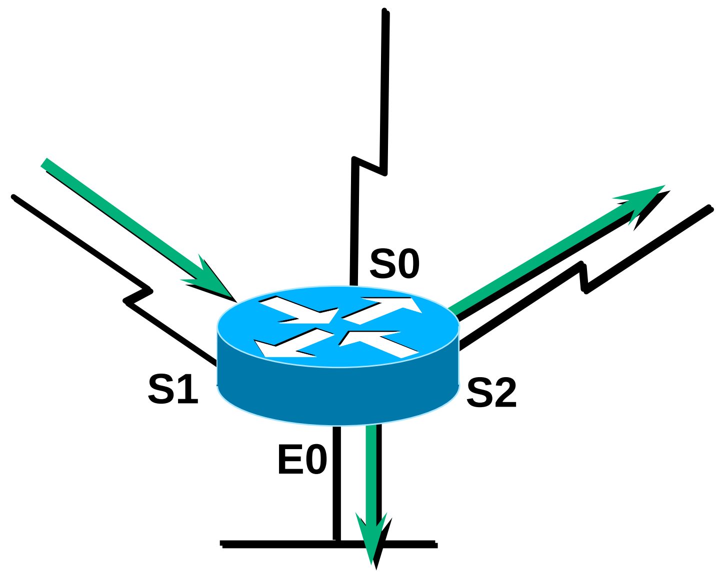
Packet Arrived on Wrong Interface!

Discard Packet!

Reverse Path Forwarding (RPF)

A closer look: RPF Check Succeeds

Multicast Packet from
Source 151.10.3.21



RPF Check Succeeds!

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

Packet Arrived on Correct Interface!

Forward out all outgoing interfaces.
(i. e. down the distribution tree)

TTL Thresholds

- **What is a TTL Threshold?**

  A "TTL Threshold" may be set on a multicast router interface to limit the forwarding of multicast traffic to outgoing packets with TTLs greater than the Threshold.

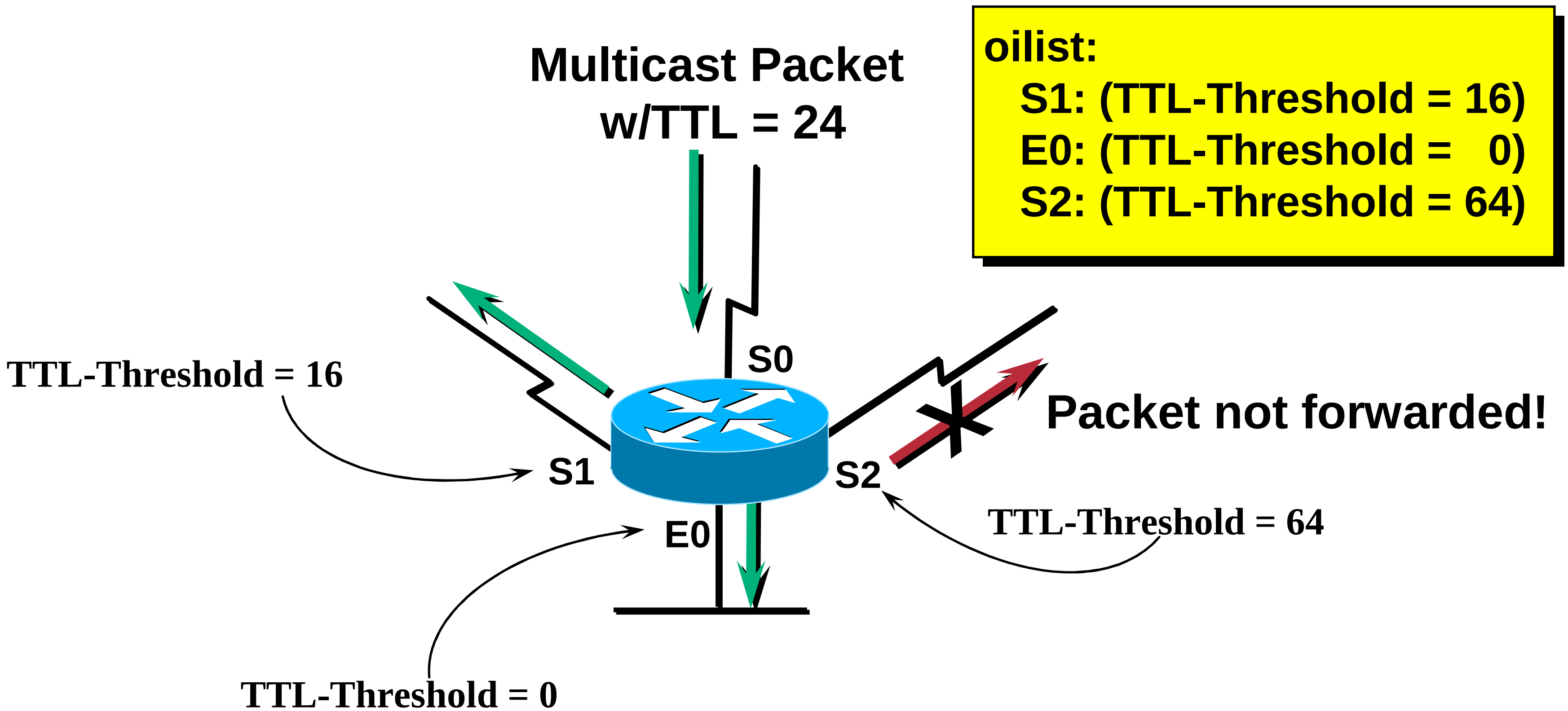
- **The TTL Threshold Check**

  1) All incoming IP packets first have their TTL decremented byone. If <= Zero, they are dropped.

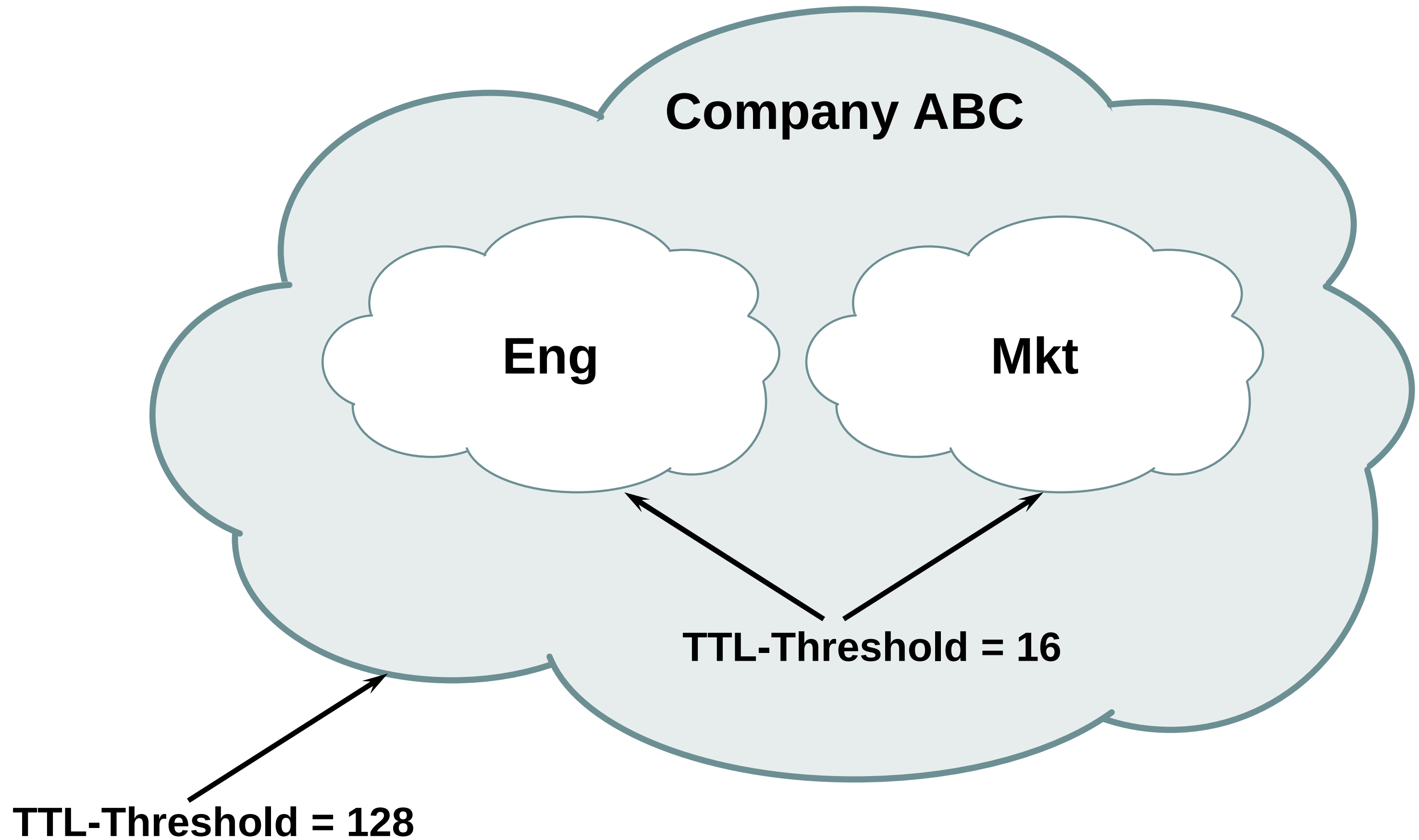
  2) If a multicast packet is to be forwarded out an interface with a non-zero TTL Threshold; then it's TTL is checked against the TTL Threshold. If the packet's TTL is < the specified threshold, it is not forwarded out the interface.

TTL Thresholds

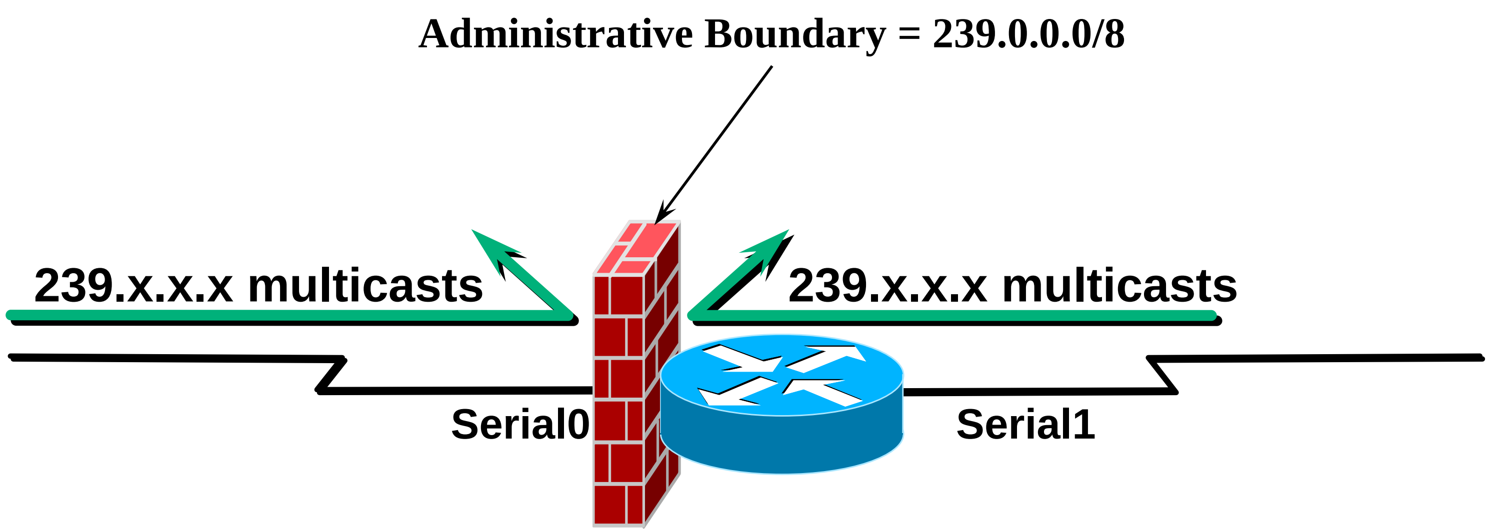
A closer look: TTL-Thresholds



TTL Threshold Boundaries

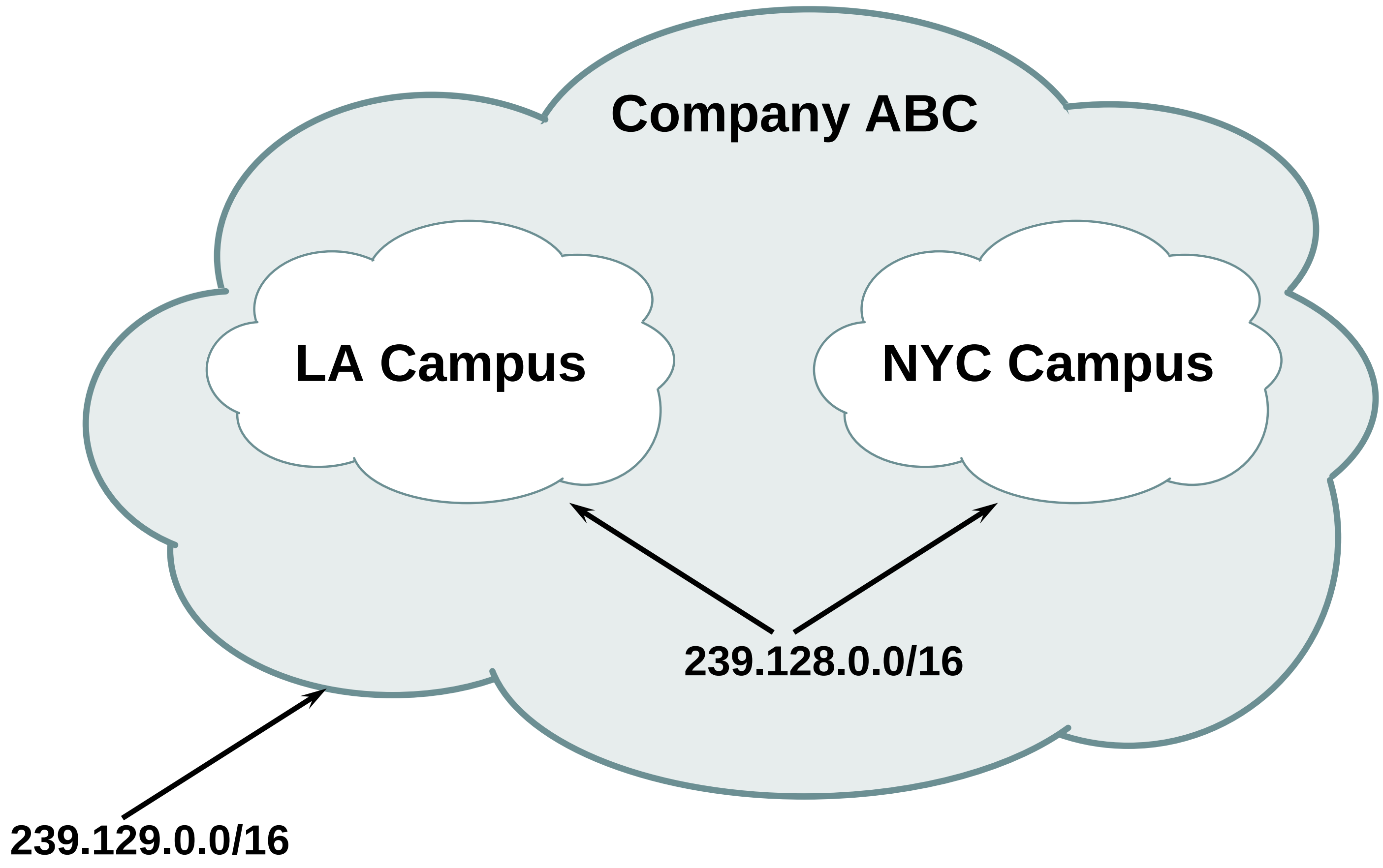


Administrative Boundaries



- Configured using the '**ip multicast boundary <acl>**' interface command

Administrative Boundaries



Multicast Protocol Basics

Types of Multicast Protocols

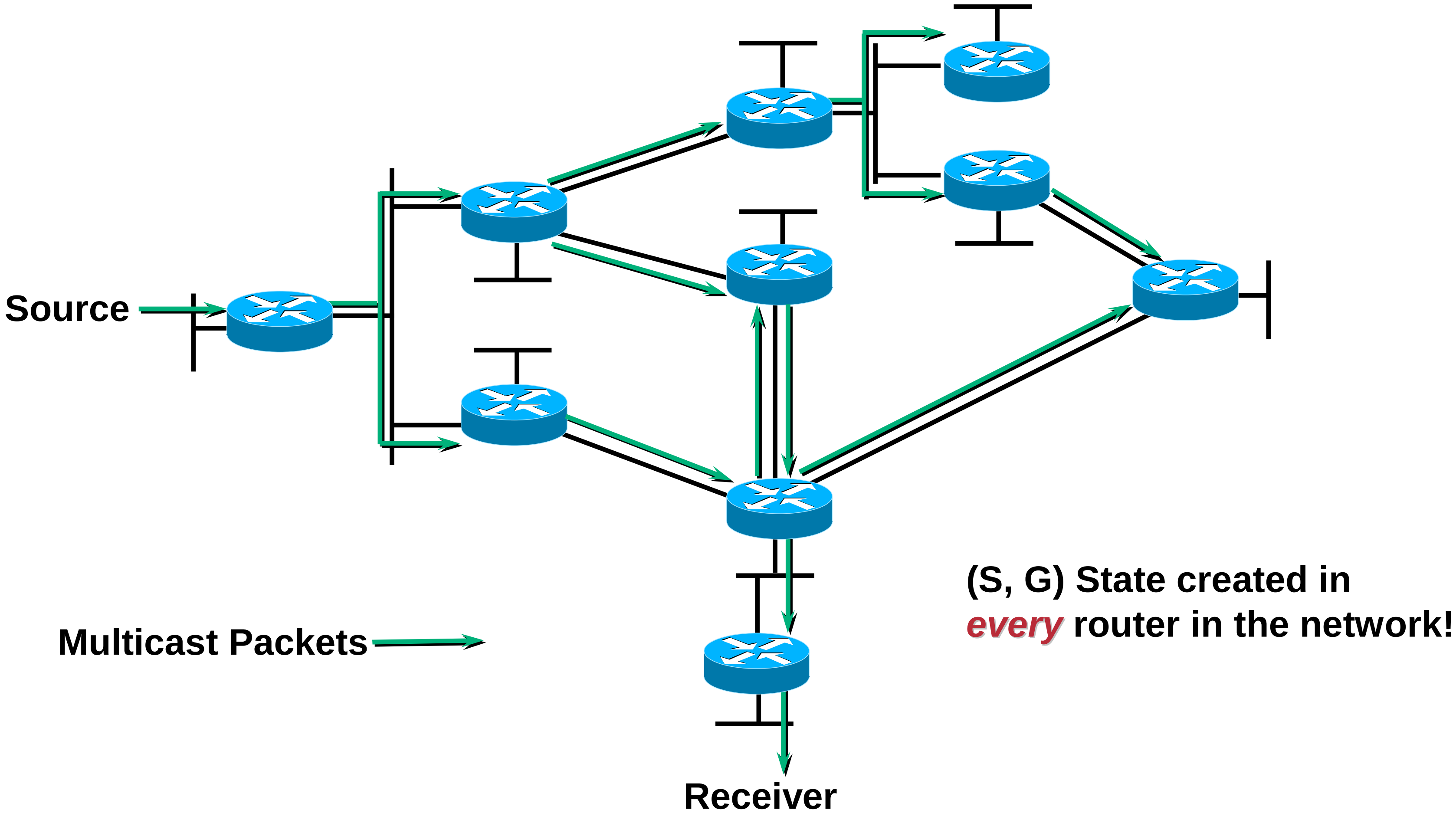
- **Dense-mode**
  - Uses “Push” Model
  - Traffic Flooded throughout network
  - Pruned back where it is unwanted
  - Flood & Prune behavior (typically every 3 minutes)
- **Sparse-mode**
  - Uses “Pull” Model
  - Traffic sent only to where it is requested
  - Explicit Join behavior

PIM-DM

- Uses Flood and Prune model
  - Floods network and prunes back based on multicast group membership
- Data Driven Events
  - Forwarding state created by data arrival.
  - Assert mechanism used to prune off redundant flows.
  - Non-Deterministic Behavior
    - Can lead to black-holes and route loops

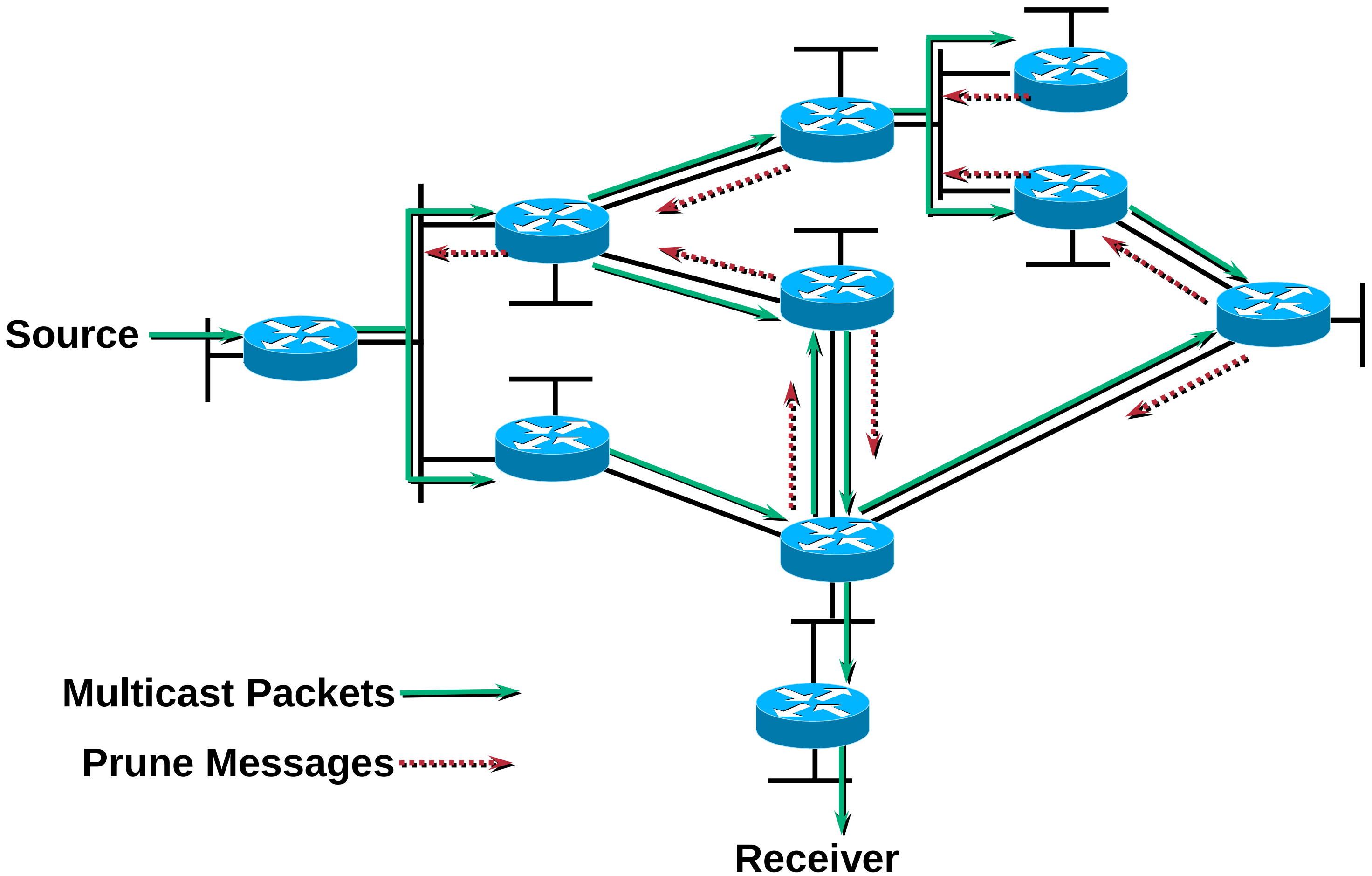
PIM-DM Flood & Prune

Initial Flooding



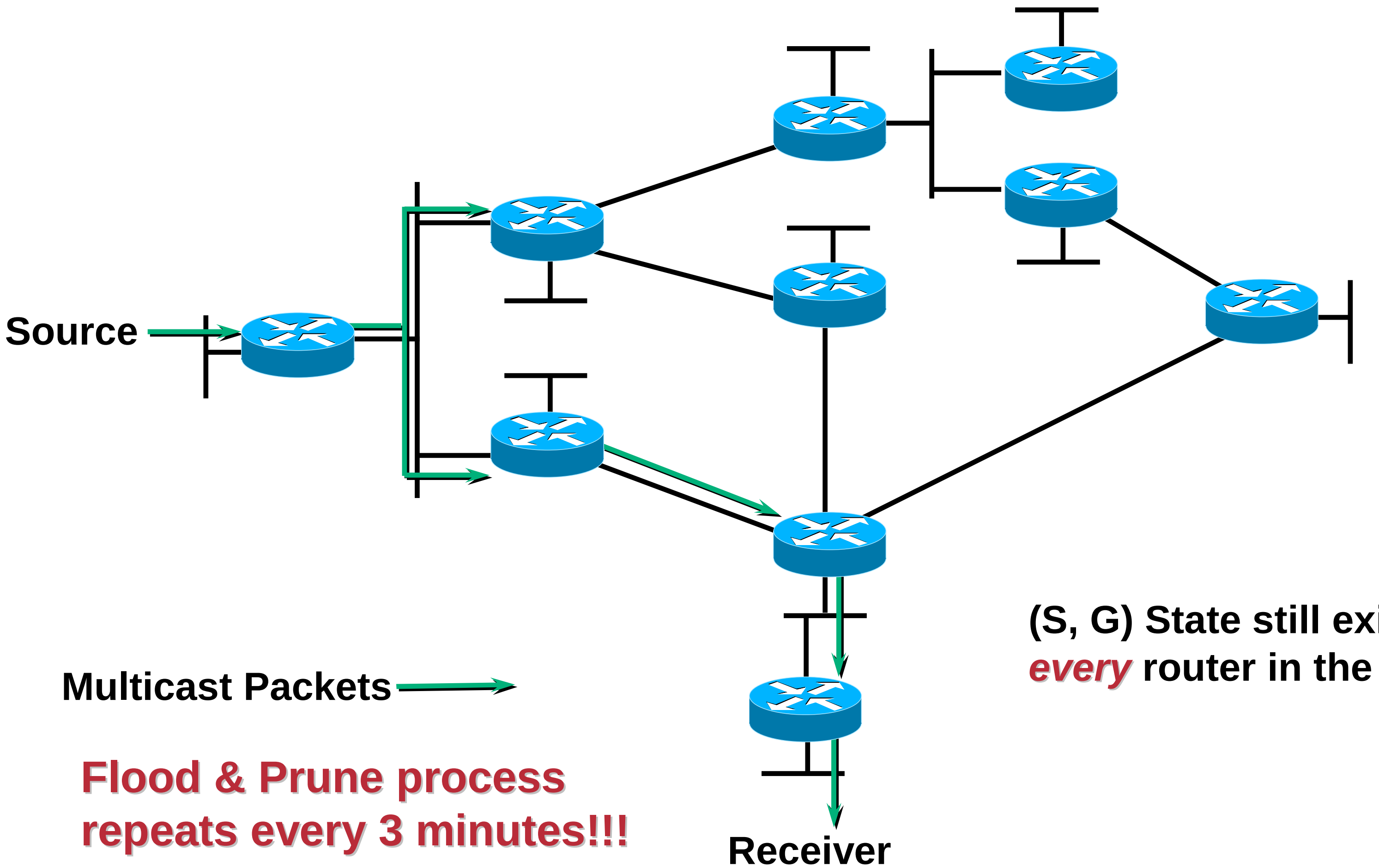
PIM-DM Flood & Prune

Pruning unwanted traffic



PIM-DM Flood & Prune

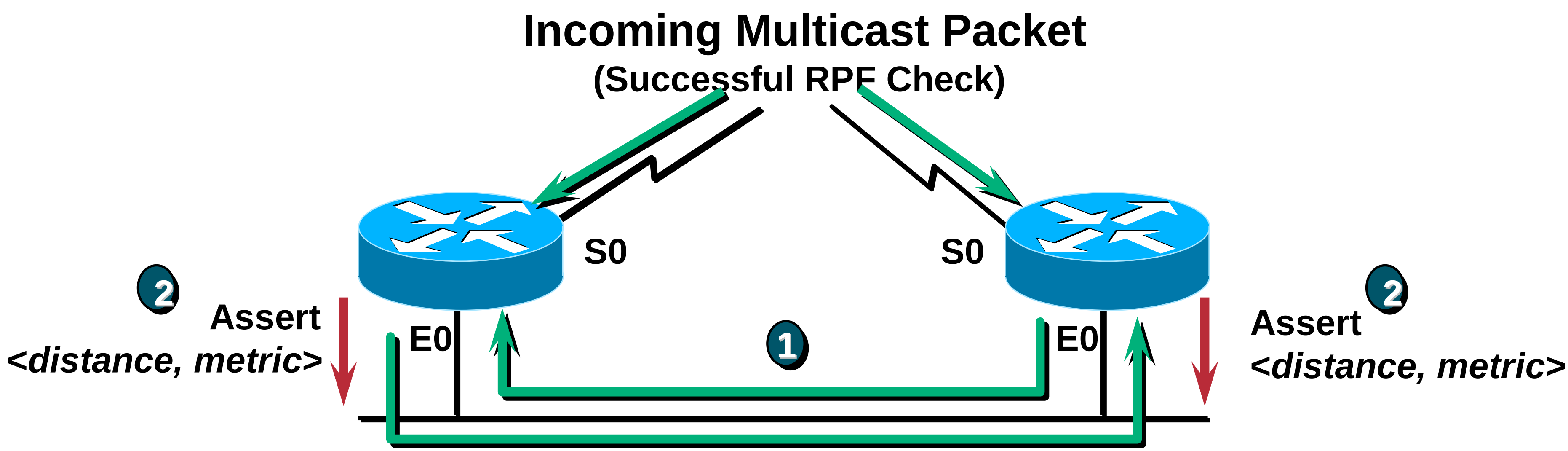
Results after Pruning



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

Flood & Prune process repeats every 3 minutes!!!

PIM-DM Assert Mechanism

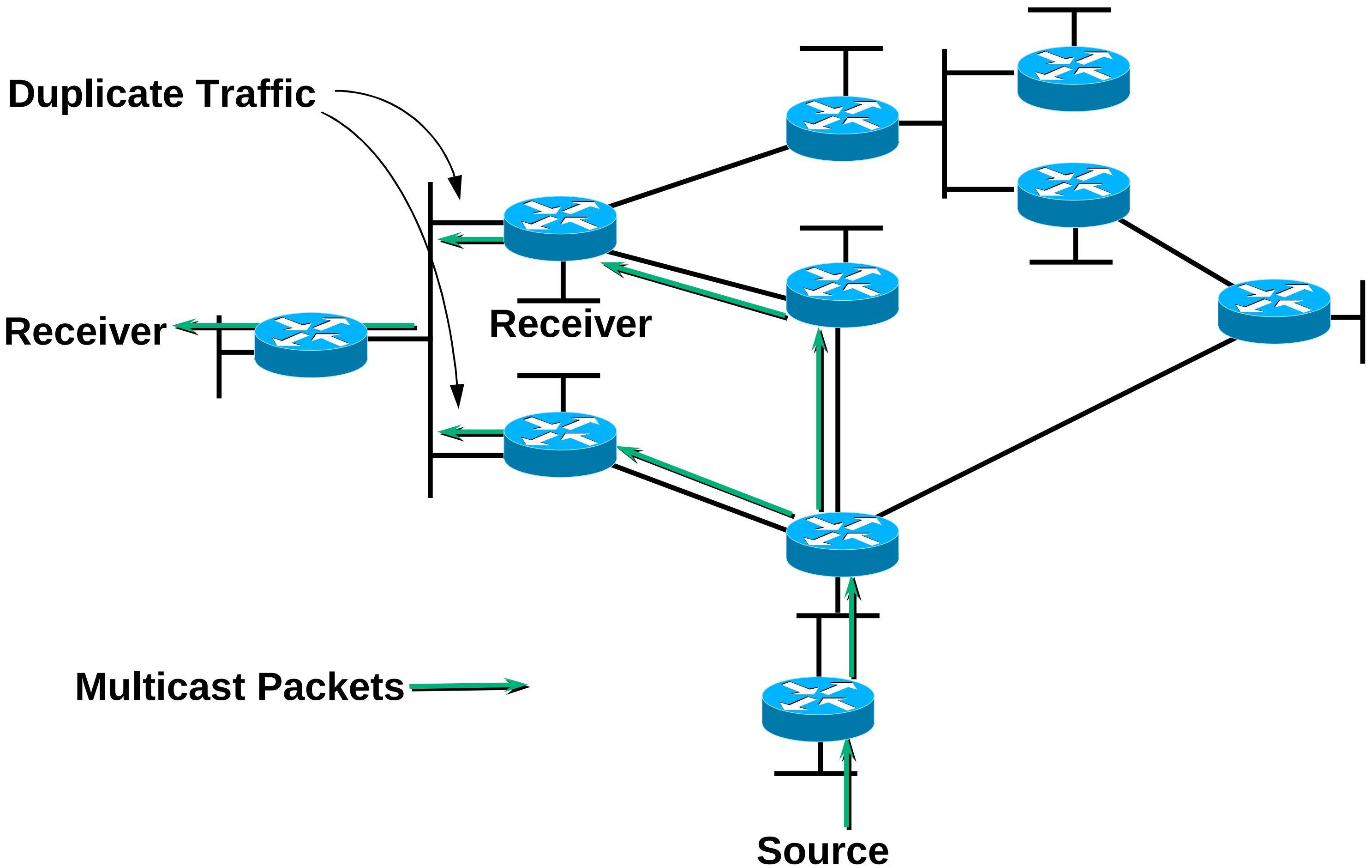


1. Routers *receive* packet on an interface in their "*oilist*"!!
   – Only one router should continue sending to avoid duplicate packets.

2. Routers send "PIM Assert" messages
   – Compare *distance* and *metric* values
   – Router with best route to source wins
   – If *metric* & *distance* equal, highest IP adr wins
   – Losing router stops sending (prunes interface)

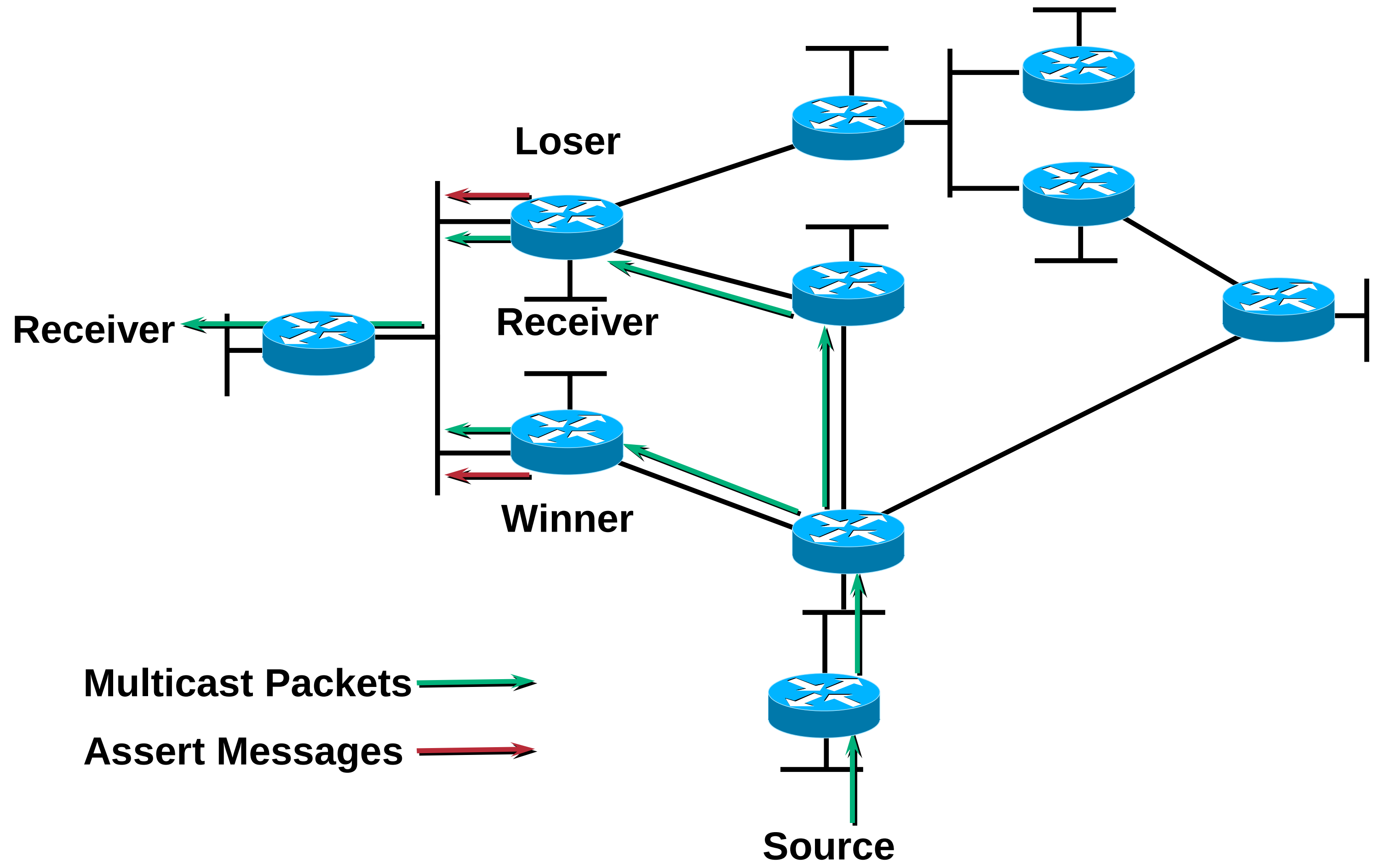
PIM-DM Assert Problem

Initial Flow



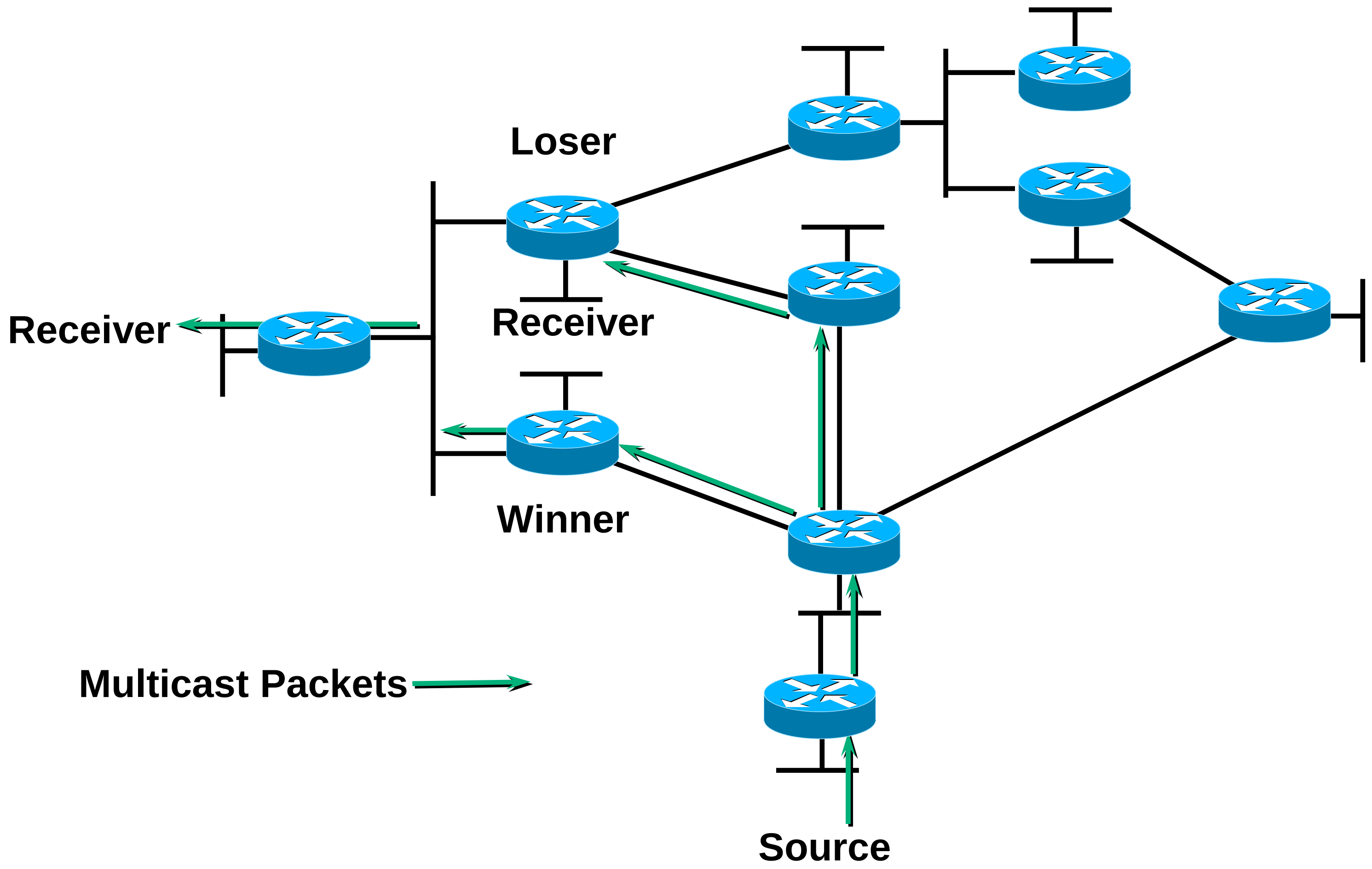
PIM-DM Assert Problem

Sending Asserts



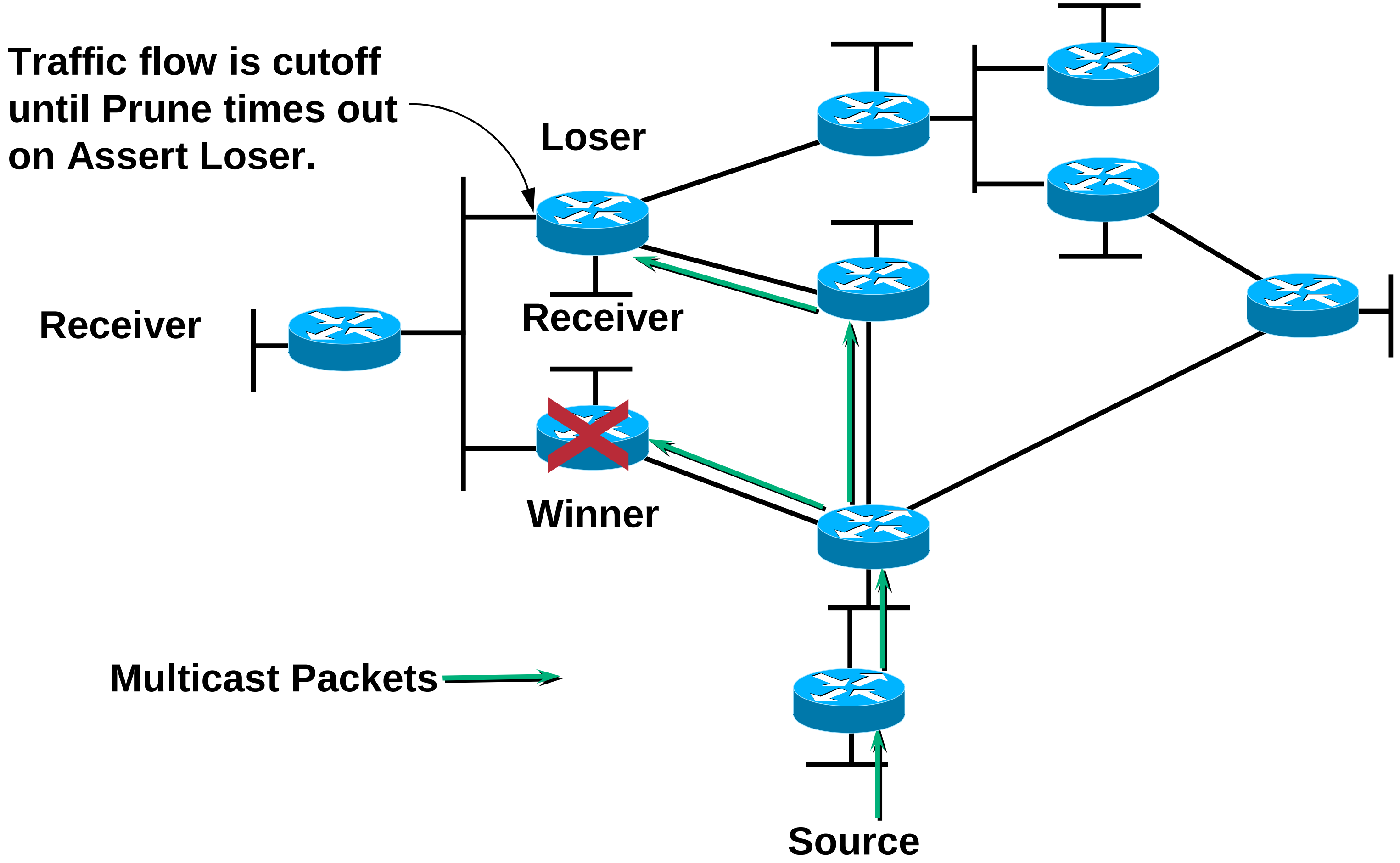
PIM-DM Assert Problem

Assert Loser Prunes Interface



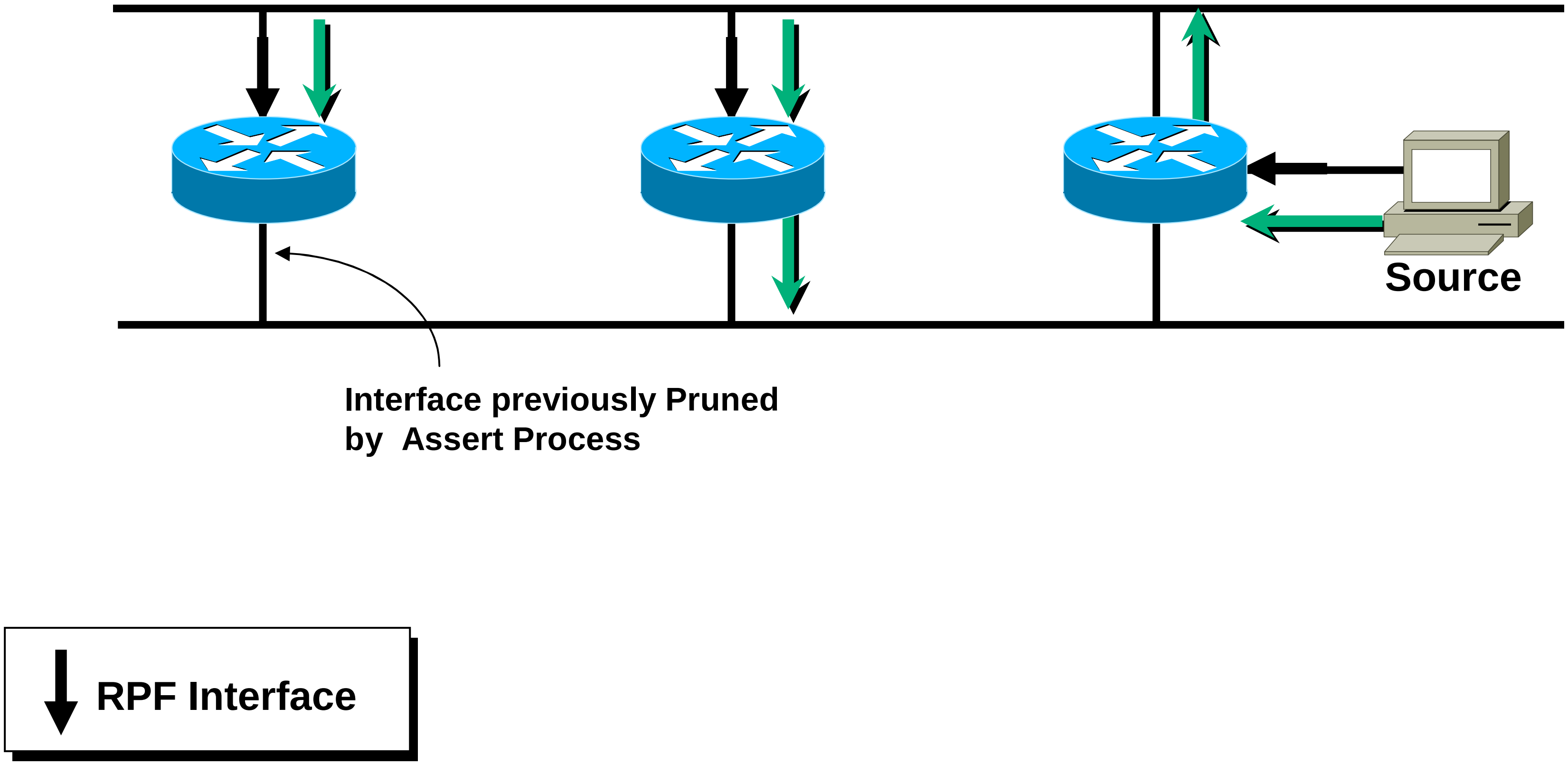
PIM-DM Assert Problem

Assert Winner Fails



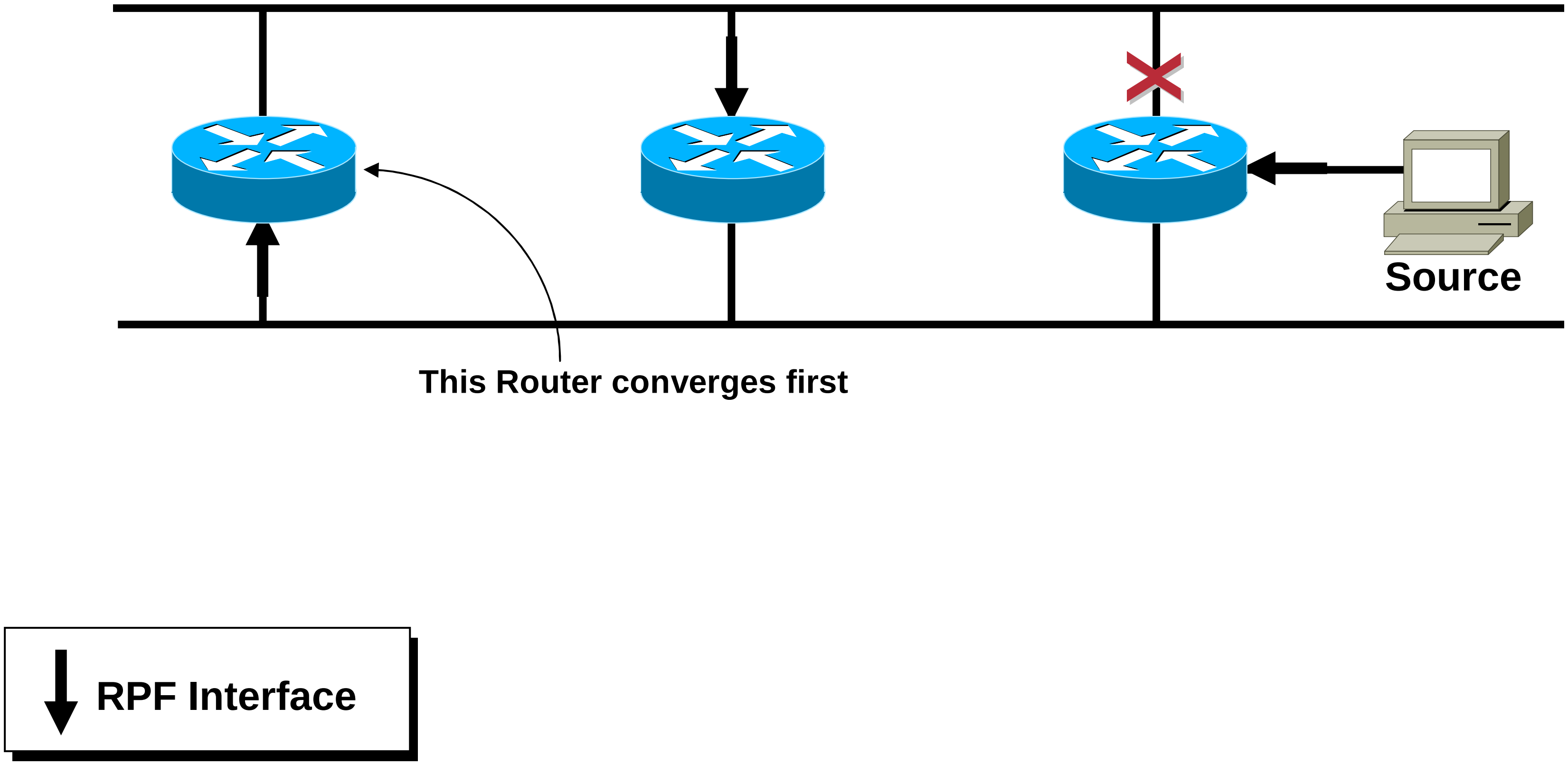
Potential PIM-DM Route Loop

Normal Steady-State Traffic Flow

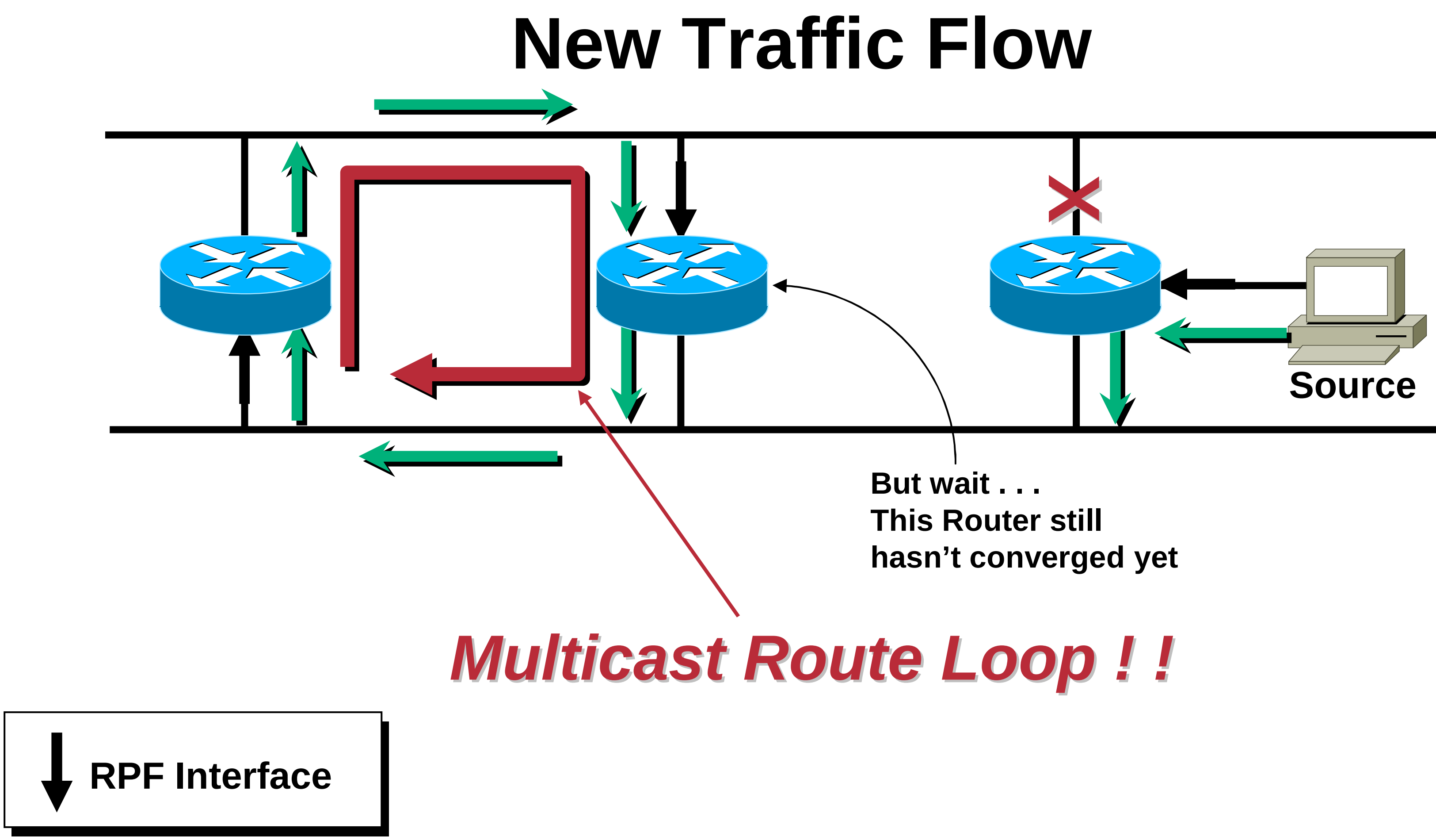


Potential PIM-DM Route Loop

Interface Fails



Potential PIM-DM Route Loop



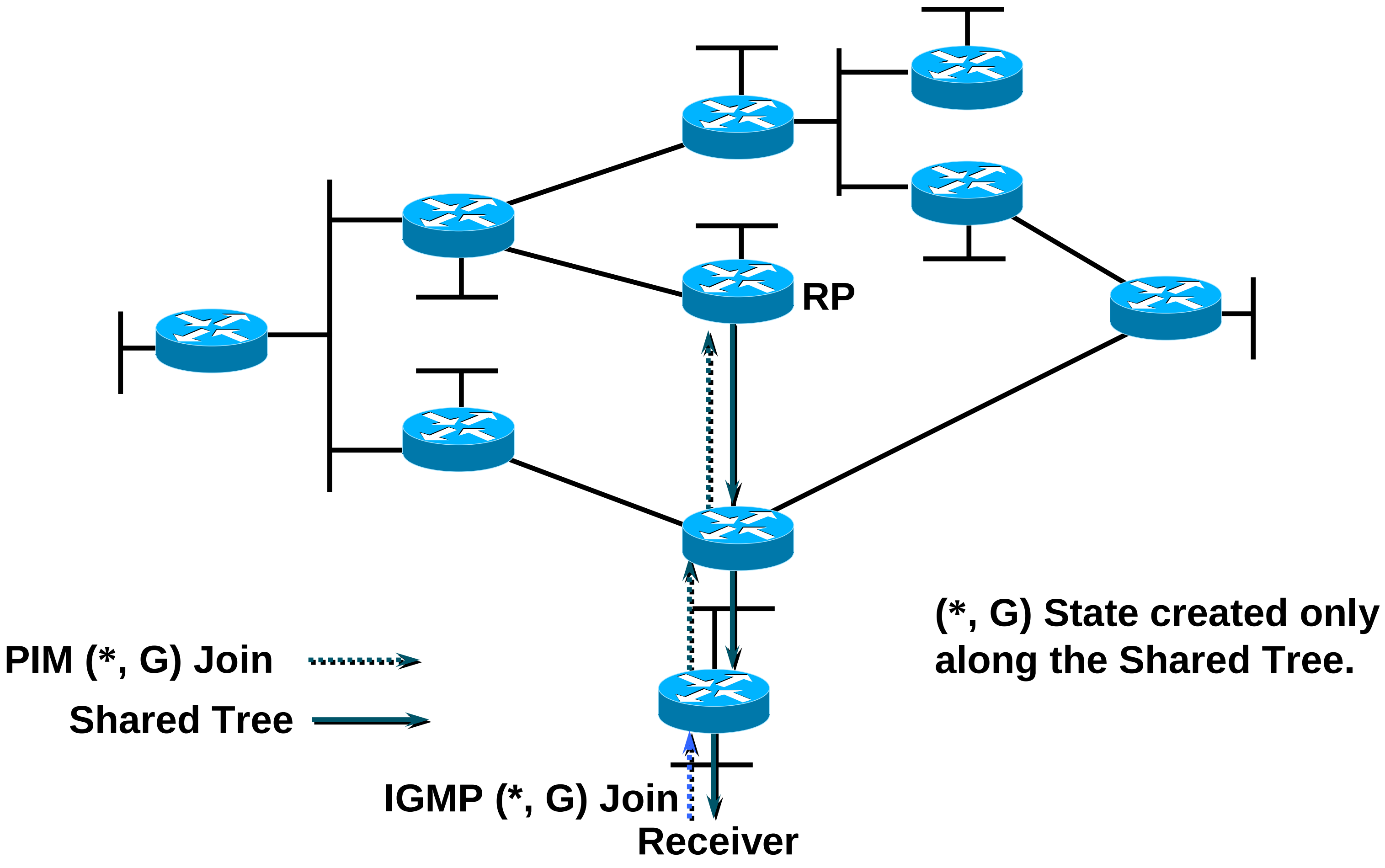
PIM-DM — Evaluation

- **Advantages**
  - Easy to configure.
- **Disadvantages**
  - Inefficient flood and prune behavior
  - Complex Assert mechanism
  - Mixed control and data planes
    - Results in (S, G) state in every router in the network
    - Can result in non-deterministic topological behaviors
- **Primary Application:**
  - Testing Router performance in labs

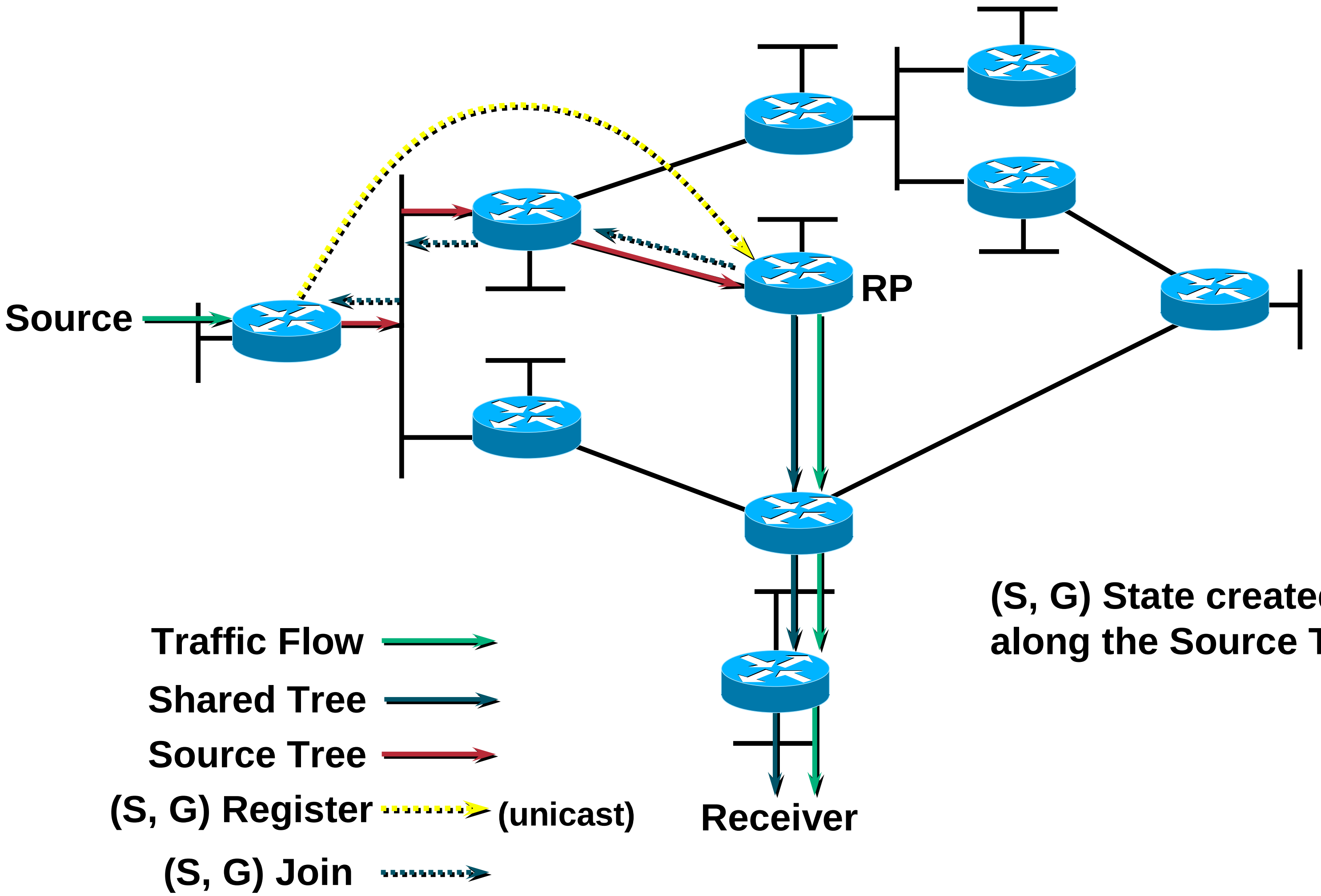
PIM-SM (RFC 2362)

- **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.**
    - **Senders are "registered" with RP by their first-hop router.**
    - **Receivers are "joined" to the Shared Tree (rooted at the RP) by their local Designated Router (DR).**
- **Appropriate for…**
  - **Wide scale deployment for *both* densely and sparsely populated groups in the enterprise**
  - **Optimal choice for all production networks regardless of size and membership density.**

PIM-SM Shared Tree Join

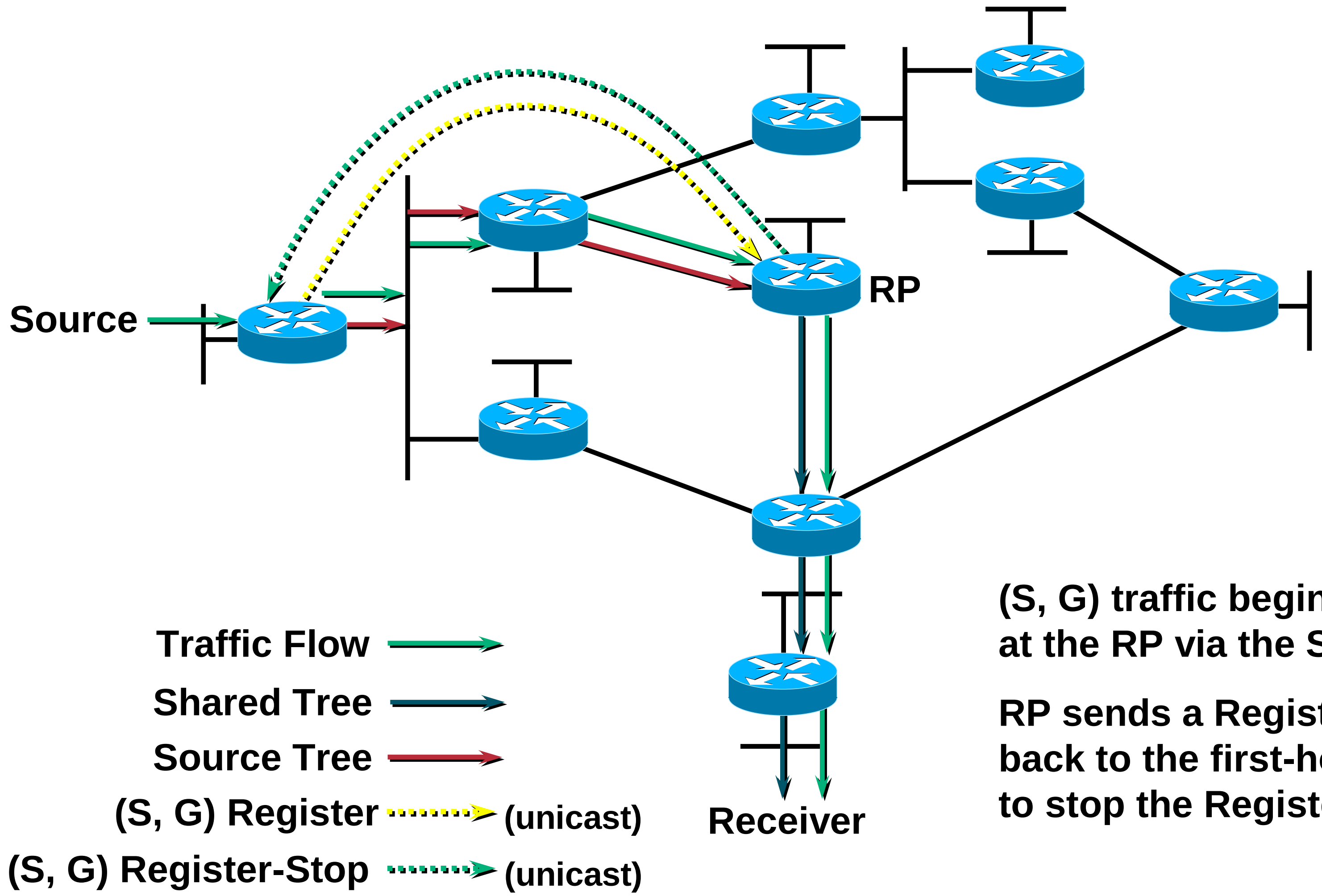


PIM-SM Sender Registration



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

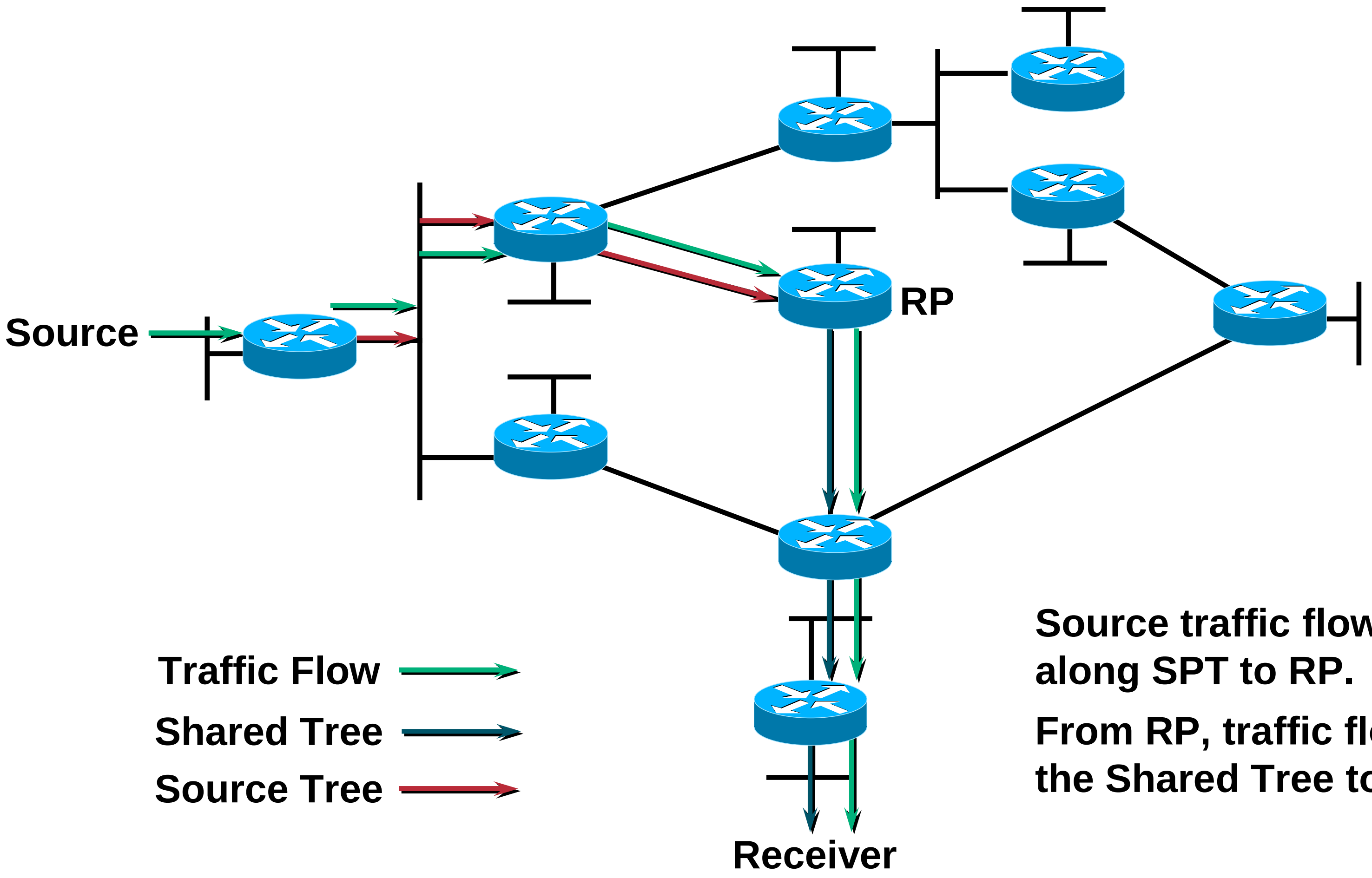
PIM-SM Sender Registration



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

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

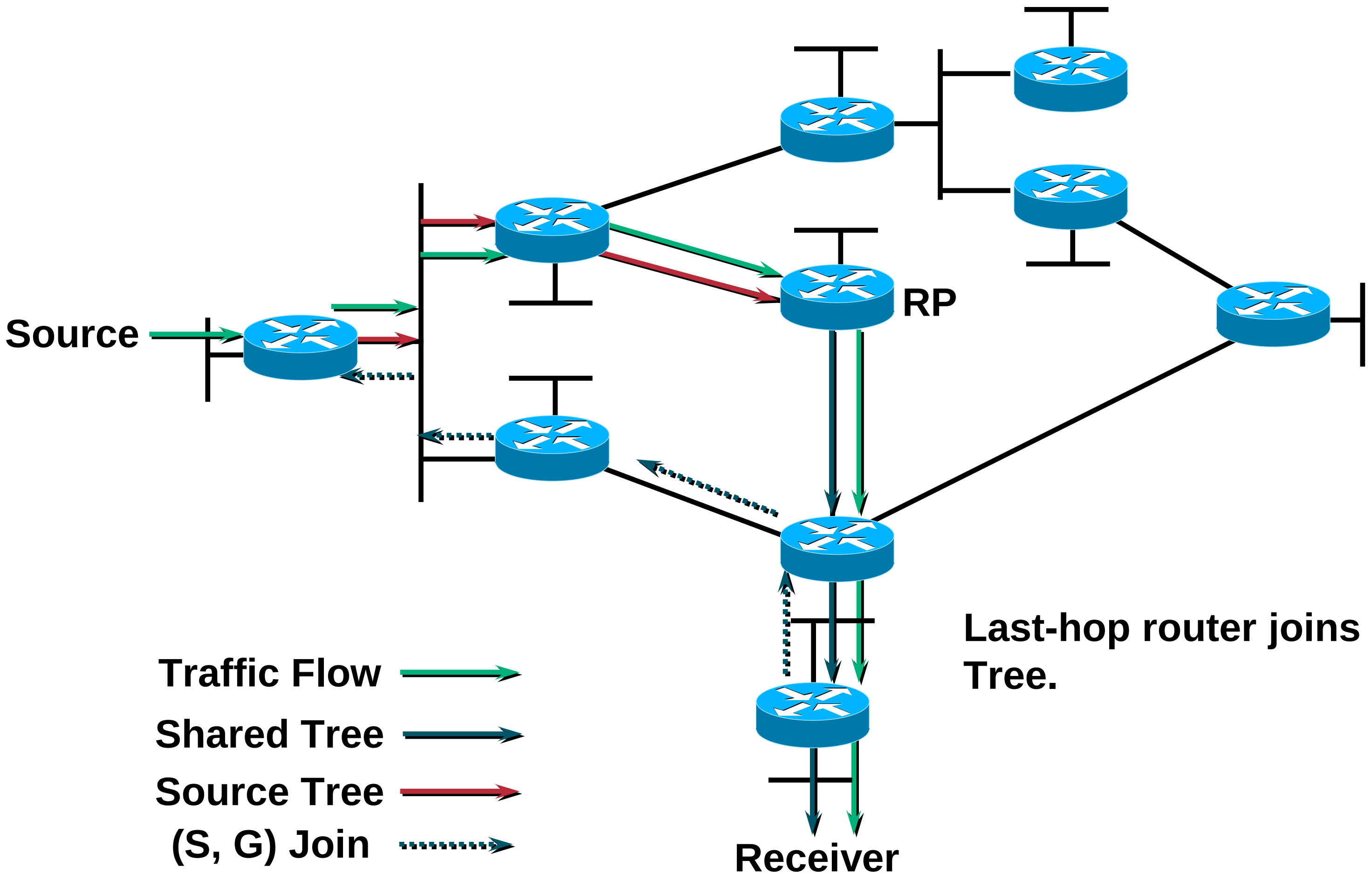
PIM-SM Sender Registration



Source traffic flows natively along SPT to RP.

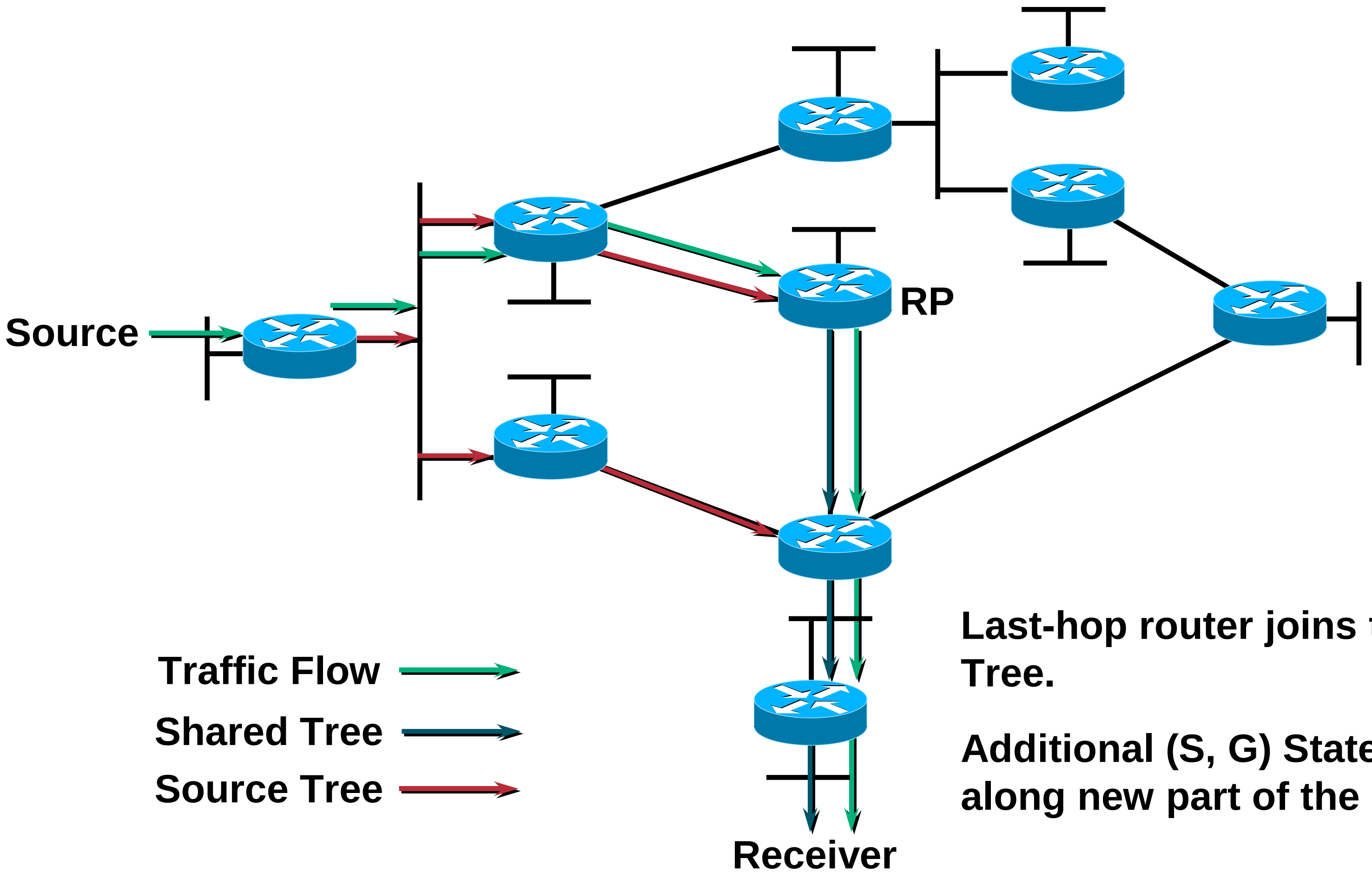
From RP, traffic flows down the Shared Tree to Receivers.

PIM-SM SPT Switchover



Last-hop router joins the Source Tree.

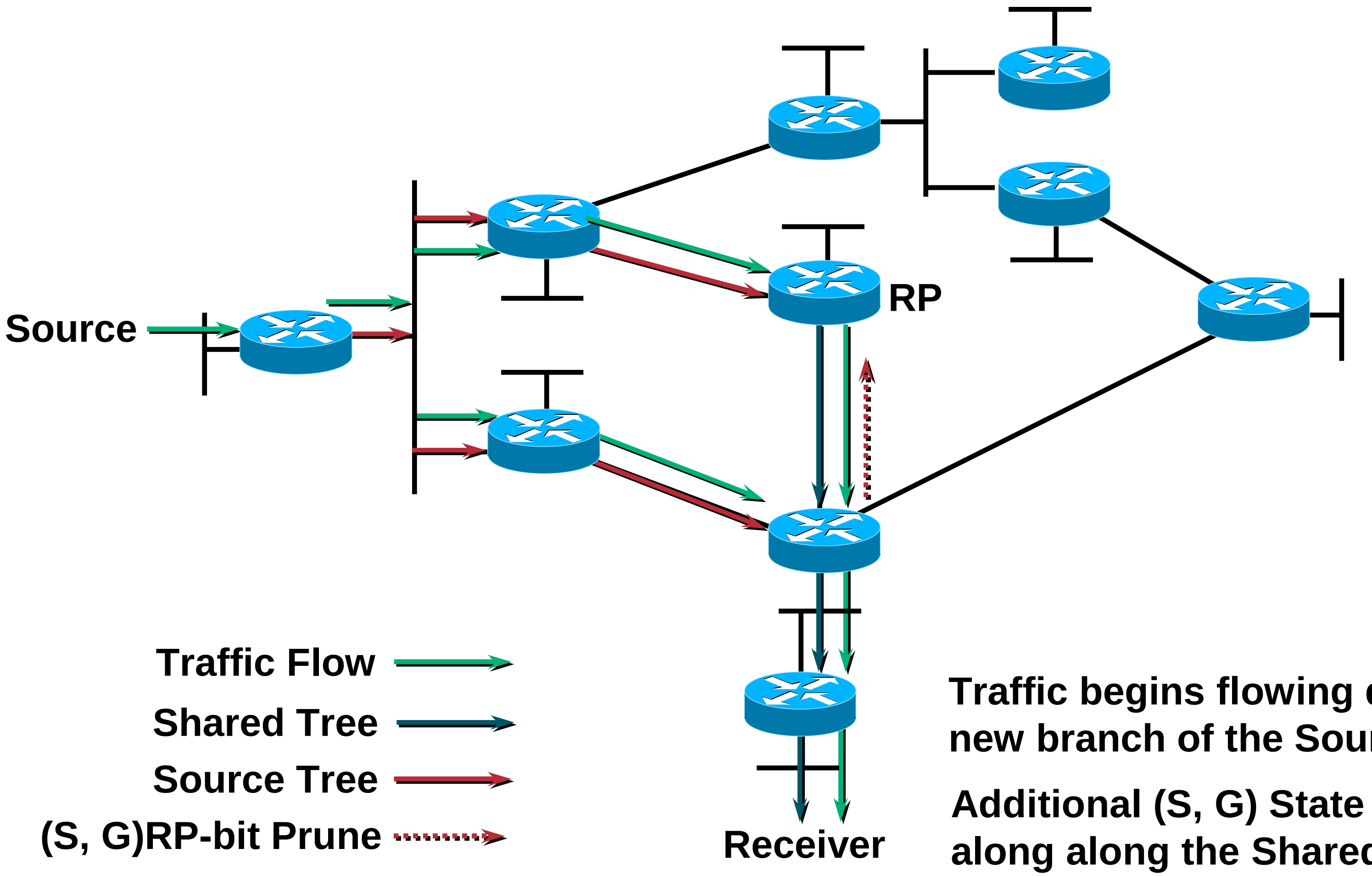
PIM-SM SPT Switchover



Last-hop router joins the Source Tree.

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

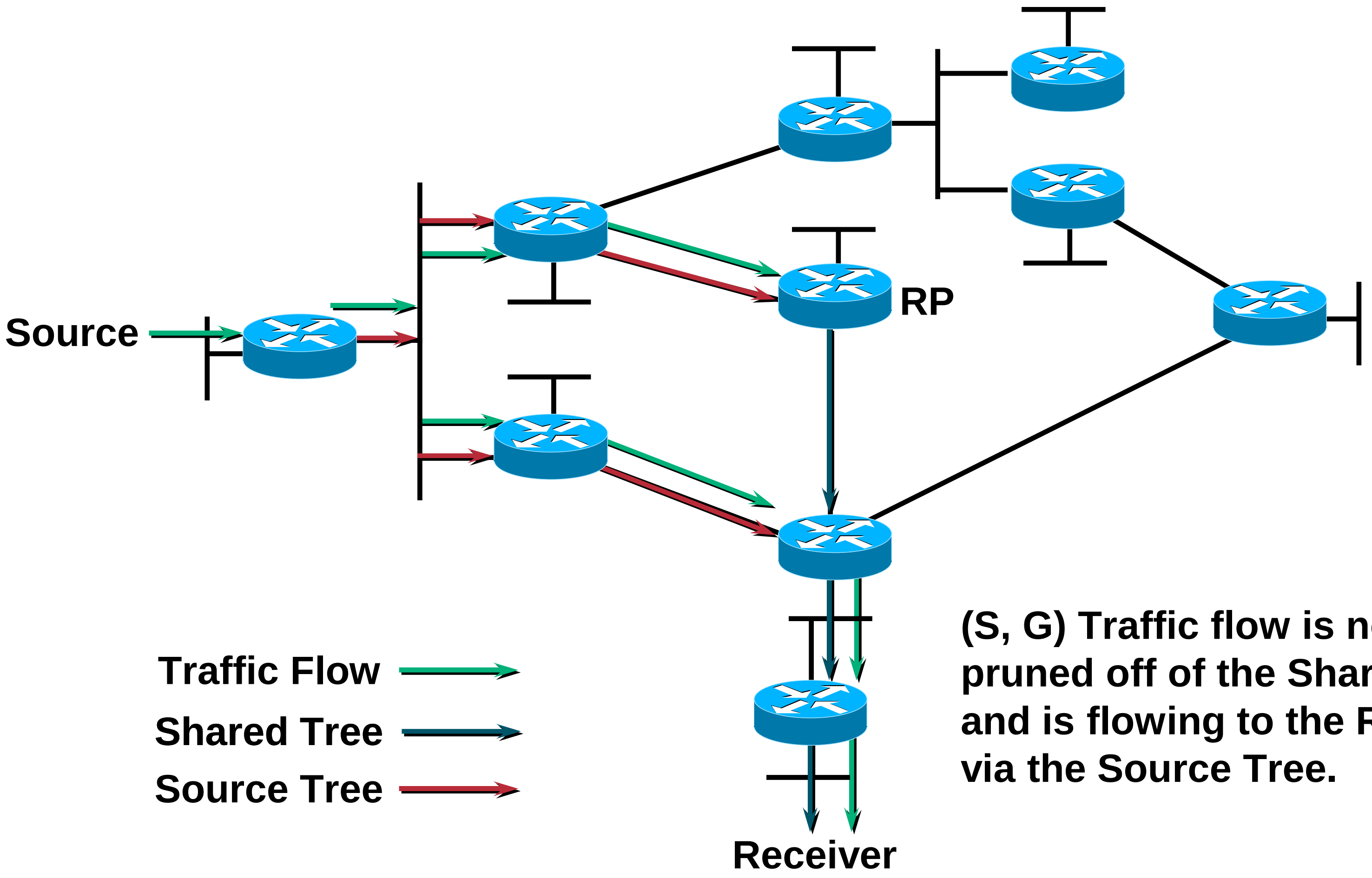
PIM-SM SPT Switchover



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

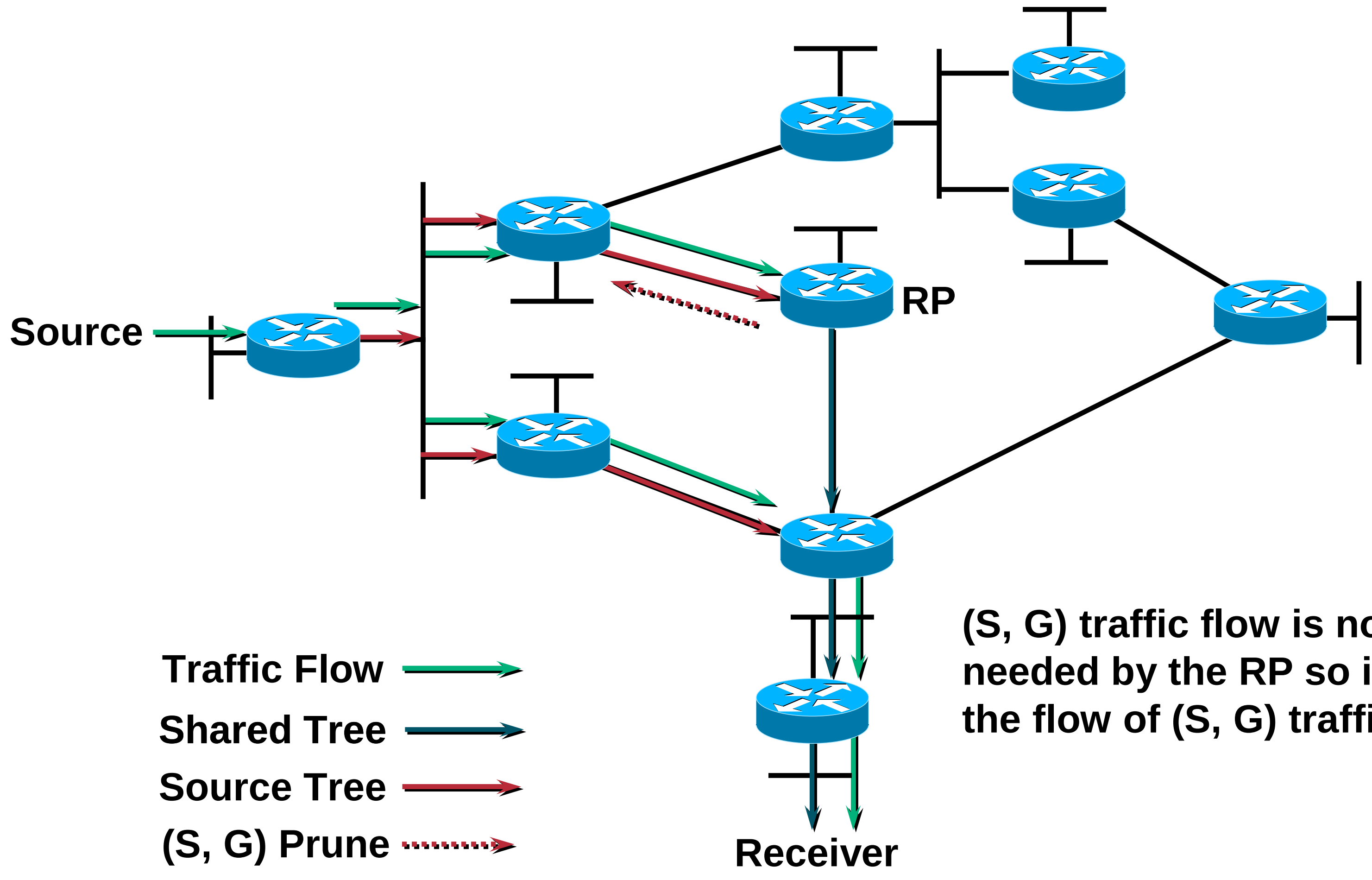
Additional (S, G) State is created along along the Shared Tree to prune off (S, G) traffic.

PIM-SM SPT Switchover



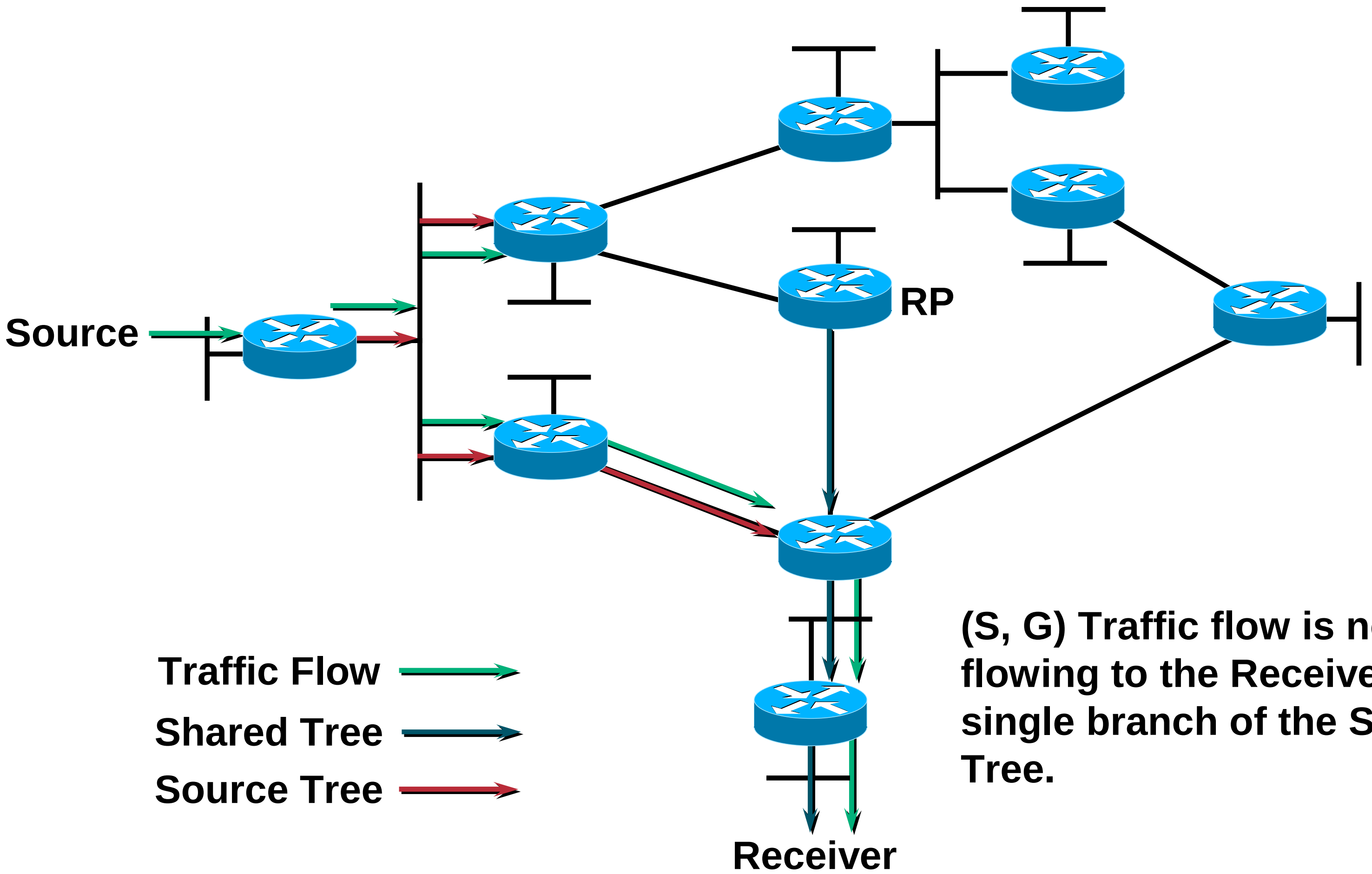
(S, G) Traffic flow is now pruned off of the Shared Tree and is flowing to the Receiver via the Source Tree.

PIM-SM SPT Switchover



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

PIM-SM SPT Switchover



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

PIM-SM FFF

PIM-SM Frequently Forgotten Fact

"The default behavior of PIM-SM is that routers with directly connected members will join the Shortest Path Tree as soon as they detect a new multicast source."

PIM-SM – Evaluation

- **Advantages:**
  - Traffic only sent down “joined” branches
  - Can switch to optimal source-trees for high traffic sources dynamically
  - Unicast routing protocol-independent
  - Basis for inter-domain multicast routing
    - When used with MBGP and MSDP
- **Disadvantages**
  - Few if any
- **Primary Application**
  - All Production Multicast Networks with sparse or dense distribution of receivers