

Interdomain Multicasting in SuperJANET4

Jonathan J. Couzens

Principal Network Engineer

JANET Network Operations & Service Centre

jjc@nosc.ja.net

+44 (0)20-7692-1316



Agenda

- Introduction
- Multicasting within the network
- Multidomain multicasting
- Multicast on JANET

Introduction

- Multicasting within the network:
 - Not end-system
 - Router-to-router
 - Single management domain

Introduction

- Multidomain multicasting:
 - Router-to-router
 - Between different management domains
 - Particular emphasis on JANET

Agenda

- Introduction
- **Multicasting within the network**
- Multidomain multicasting
- Multicast on JANET

Single-domain Multicasting

- Dense/Sparse Protocols
- Reverse Path Forwarding
- Protocol-Independent Multicast
- Rendezvous Points

Dense/Sparse Protocols

- Dense:
 - Typified by ‘flood and prune’ behaviour
 - The last-hop router gets everything unless it explicitly rejects it
- Sparse:
 - Typified by ‘send on request’ behaviour
 - The end-system gets nothing unless it explicitly asks for it

Reverse Path Forwarding

- Crucial to understanding wide-area multicasting
- How a router determines whether to forward a multicast packet and, if so, where to....

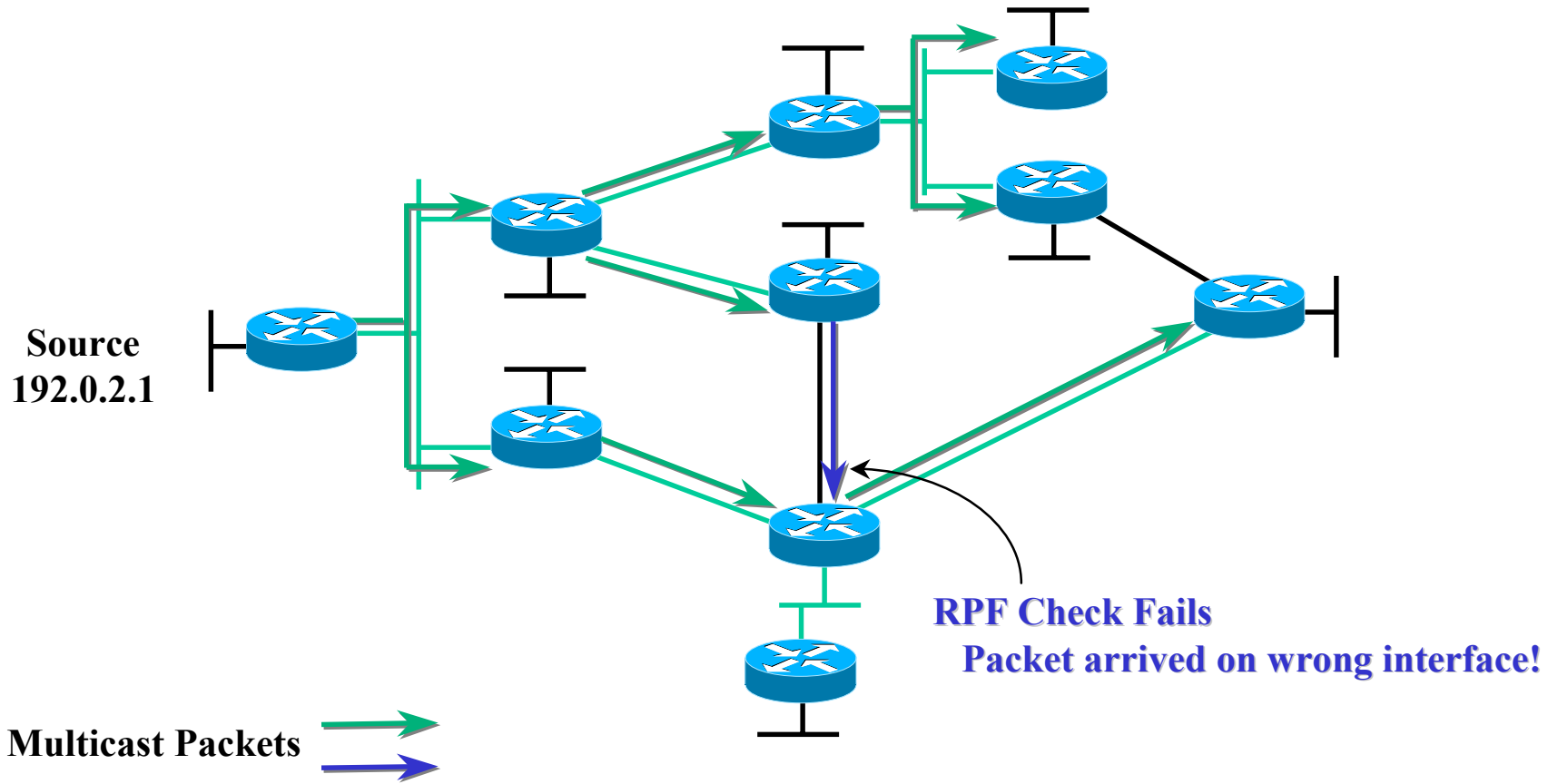
Reverse Path Forwarding

- Unicast forwarding
 - works by looking up the destination address in a routing table and forwarding to the associated interface

Reverse Path Forwarding

- Multicast forwarding
 - look up the source address in the routing table
 - did the packet arrive on the interface to which a packet would have been sent if that address had been the destination address?
 - This is the ‘RPF check’
 - If so, forward it out of all the multicast-enabled interfaces except for the one on which it arrived

Reverse Path Forwarding



Protocol-Independent Multicast

- A bit of terminology:
 - For each particular multicast flow, the router may need to keep the states:
 - $(*,G)$ - all sources for group G
 - (S,G) - source S sending to group G
 - Typically, needs to remember whether to and where to forward packets

Protocol-Independent Multicast

- Router-to-router protocol
- Protocol-independent
 - uses existing routing information
 - hence independent of the underlying routing protocols

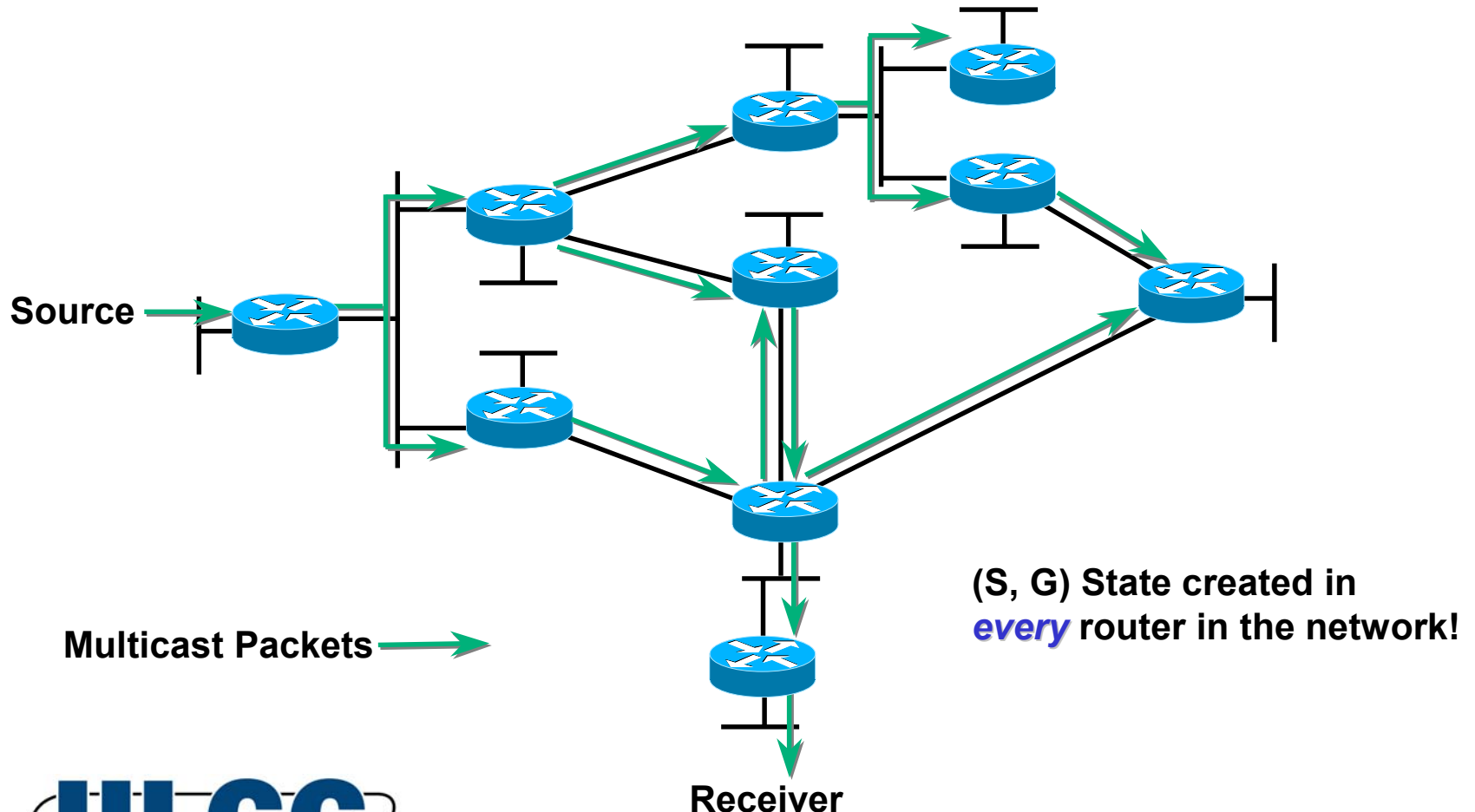
Protocol-Independent Multicast

- Two flavours (at least!)
 - Dense-mode
 - various Internet-Drafts but no RFC yet
 - Sparse-mode
 - RFC2362 (June 1998) - Experimental
 -Single-Source mode
 -Bidirectional-mode

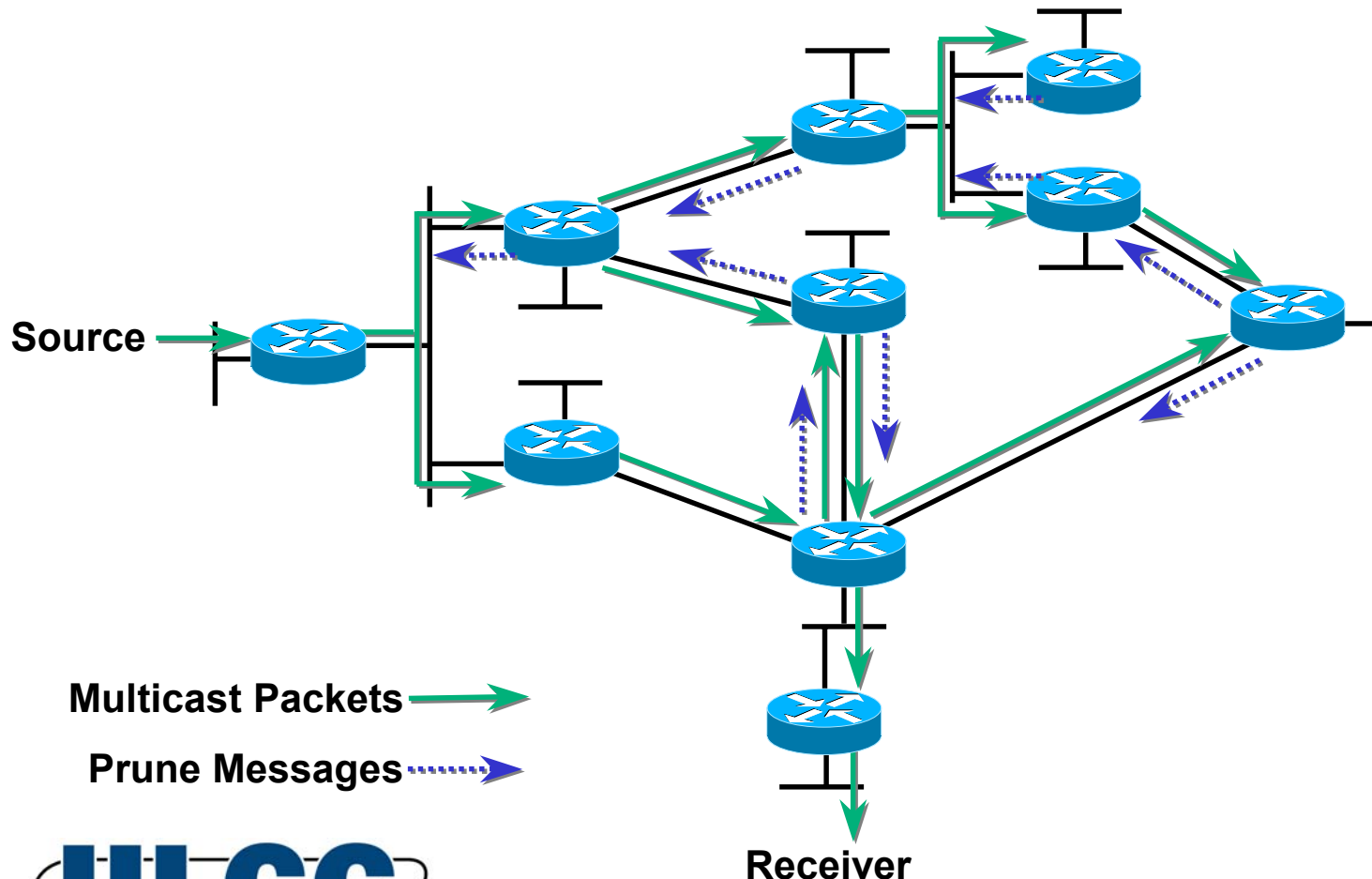
PIM Dense-Mode

- PIM-DM
- A dense-mode protocol
- Flood-and-prune
- Distribution tree for the traffic rooted at the source

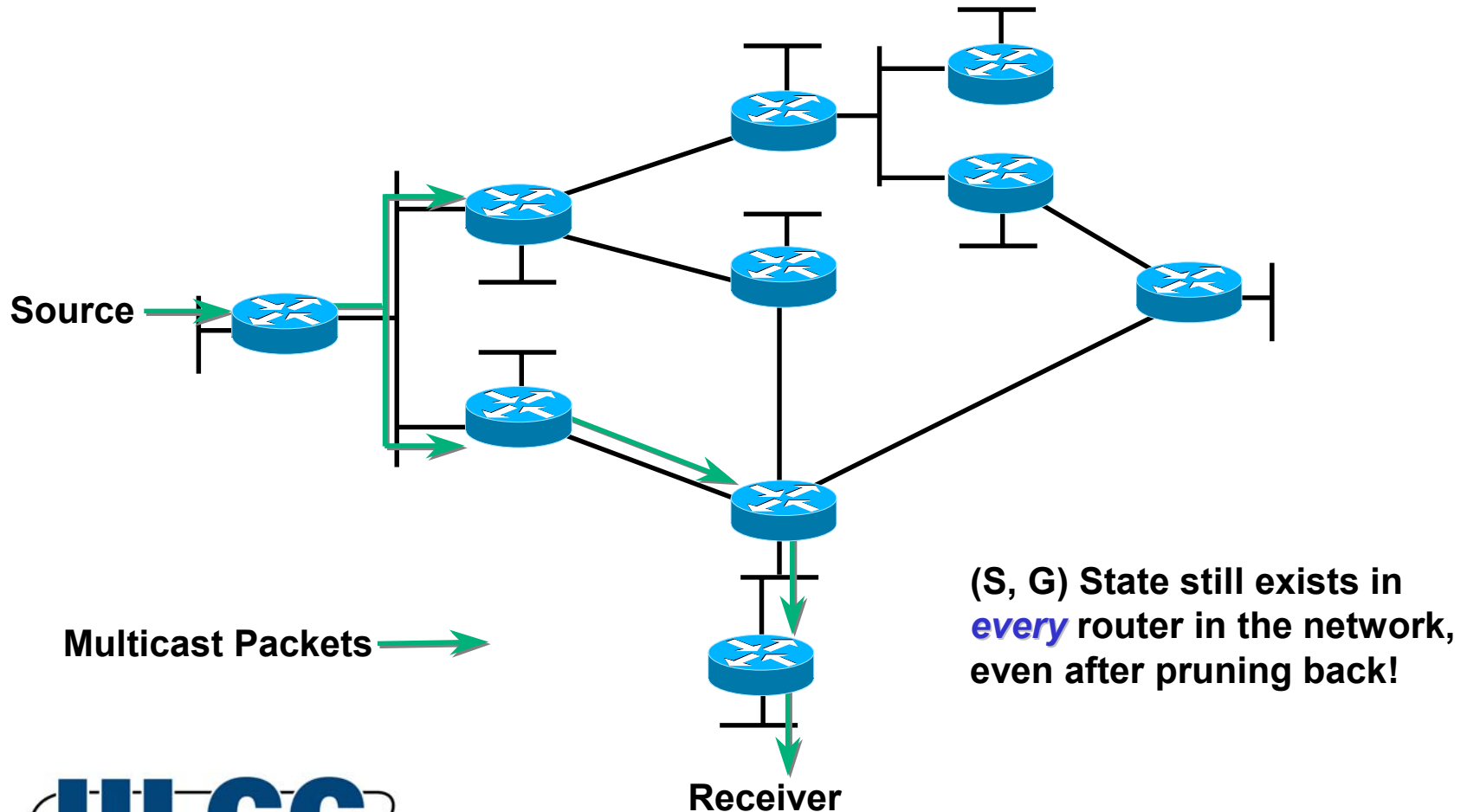
PIM-DM - initial flooding



PIM-DM - pruning back



PIM-DM - after pruning



PIM-DM

- Great if
 - virtually everyone wants to receive the traffic
 - local scope
 - e.g. financial dealing rooms
- Not so good if
 - few want the traffic
 - global transmission

PIM-DM

- Currently, some 12 Mbps of multicast being sent around JANET *continuously*
- Around 2000 sources transmitting globally at any one time
- Do you want that on your network?
- Especially true if your sites has a low-speed access line

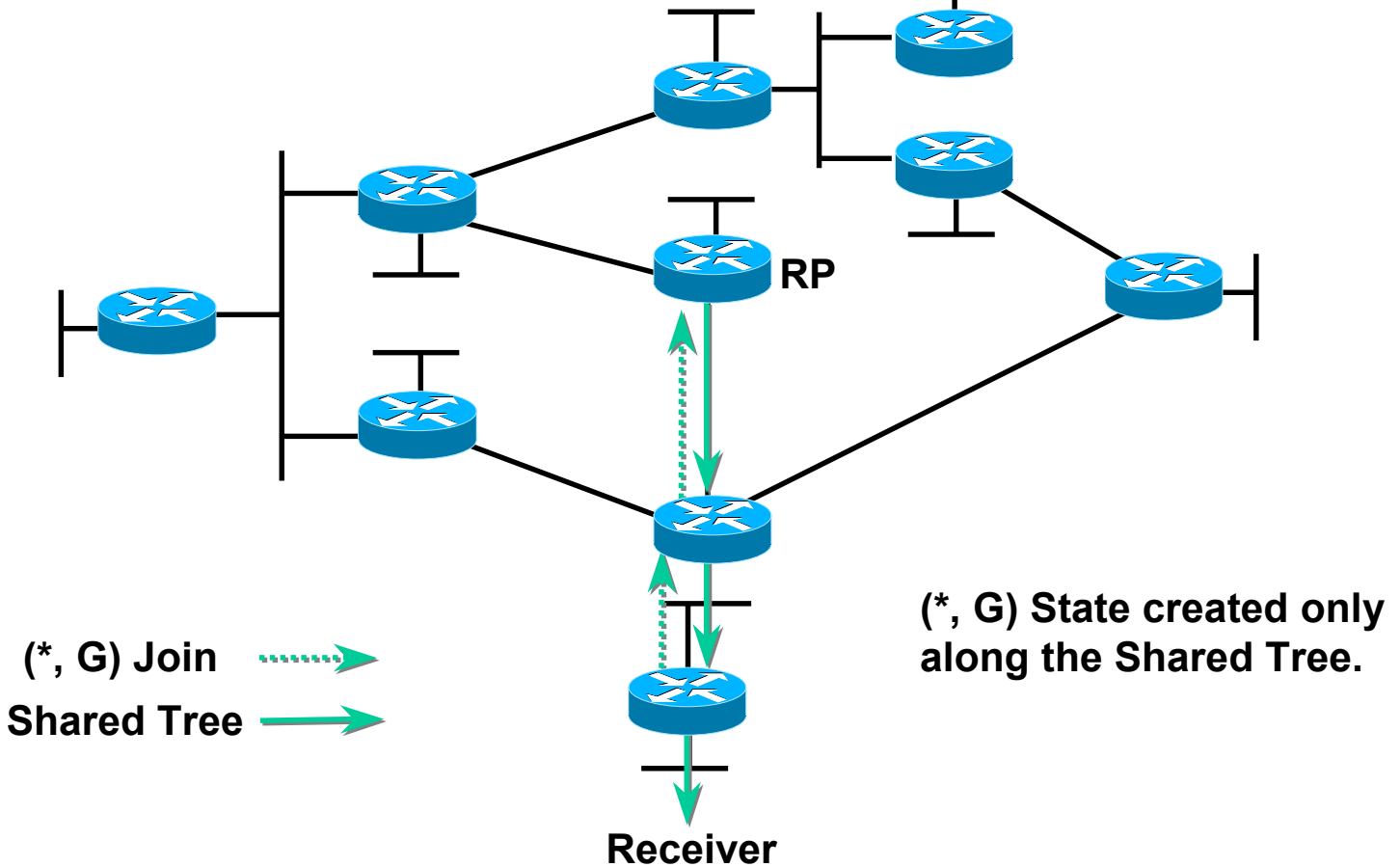
PIM Sparse-Mode

- PIM-SM
- Sparse-mode protocol, so
 - Traffic only received if explicitly requested
 - But you have to know what to ask for....
- To avoid propagation of state throughout all routers, PIM-SM has the concept of a rendezvous point (RP) which knows all!

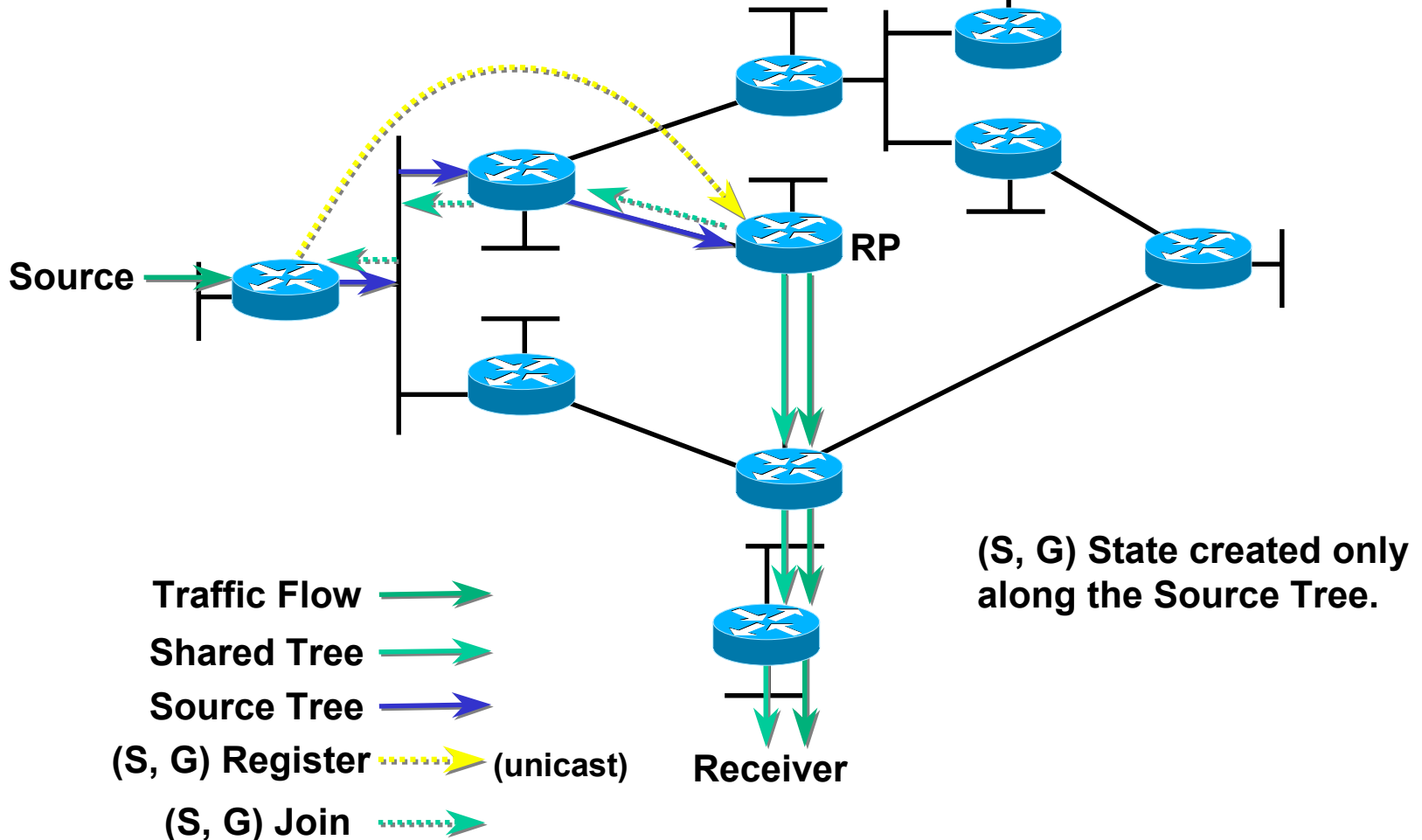
PIM-SM

- Receivers join to the RP
- Senders register with the RP
- Two distribution trees involved:
 - Shared tree, rooted at the RP
 - Source tree, rooted at the source
- All last-hop routers in the domain need to know where the RP is

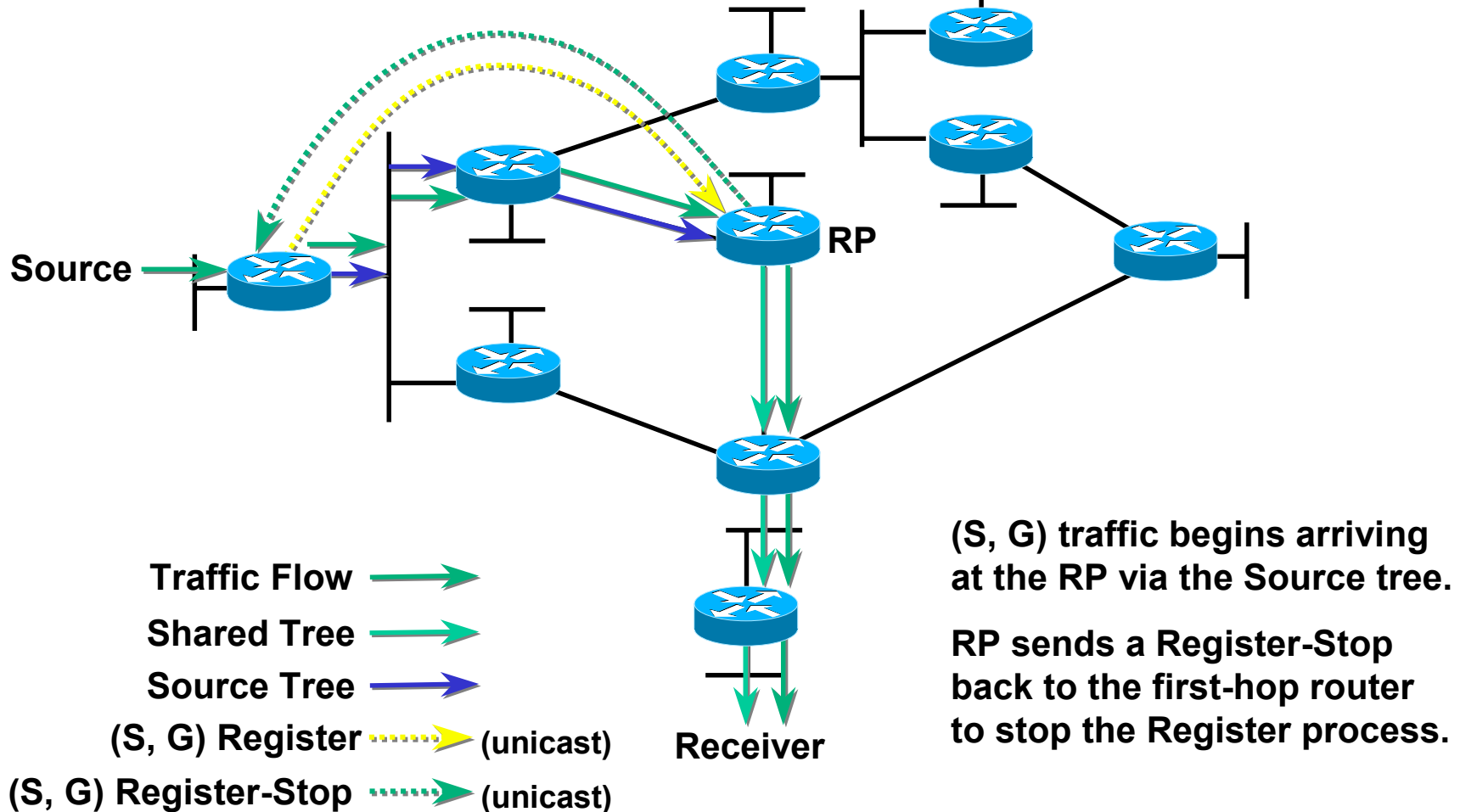
PIM-SM - receiver join



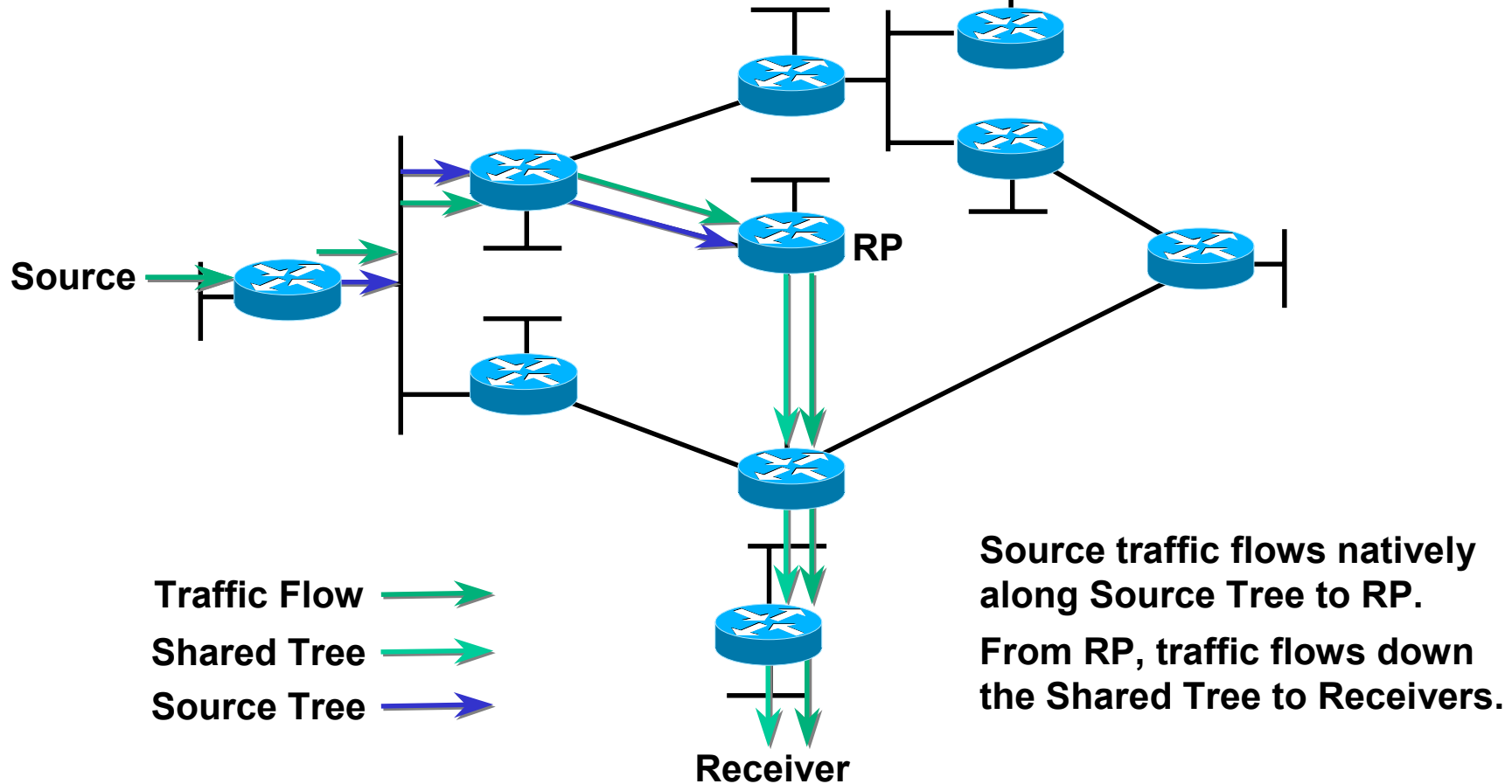
PIM-SM - sender registration



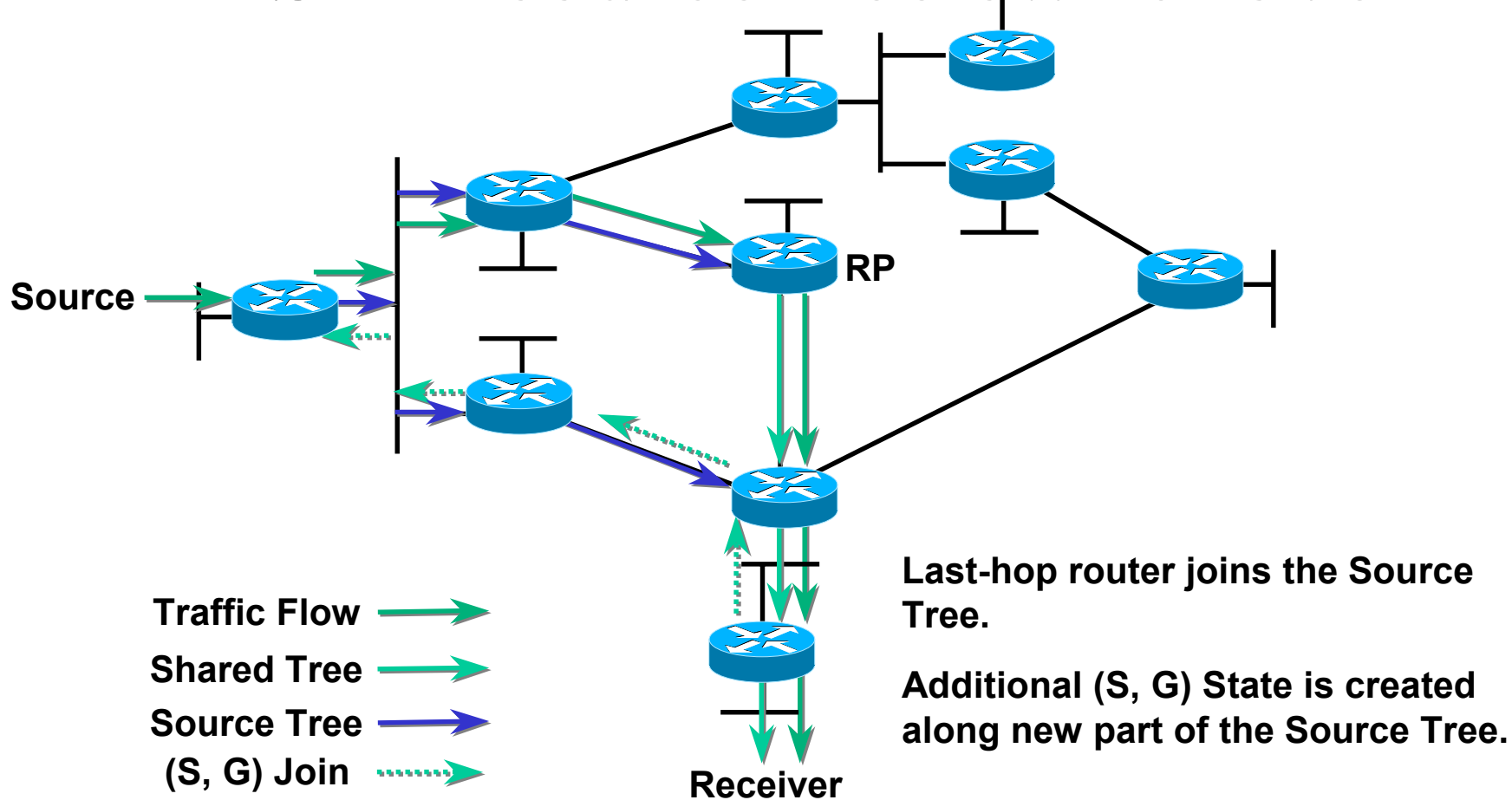
PIM-SM - sender registration



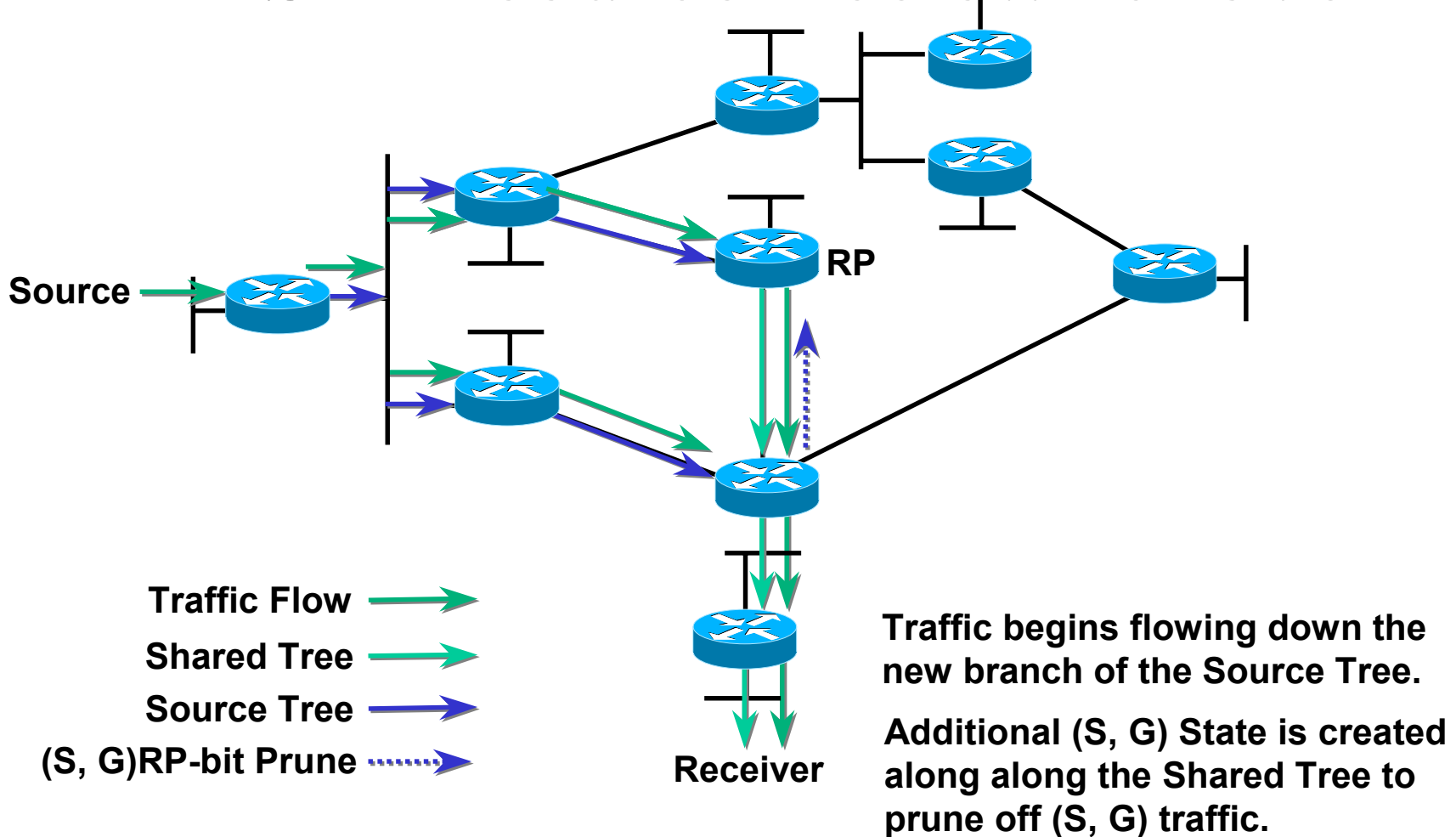
PIM-SM - sender registration



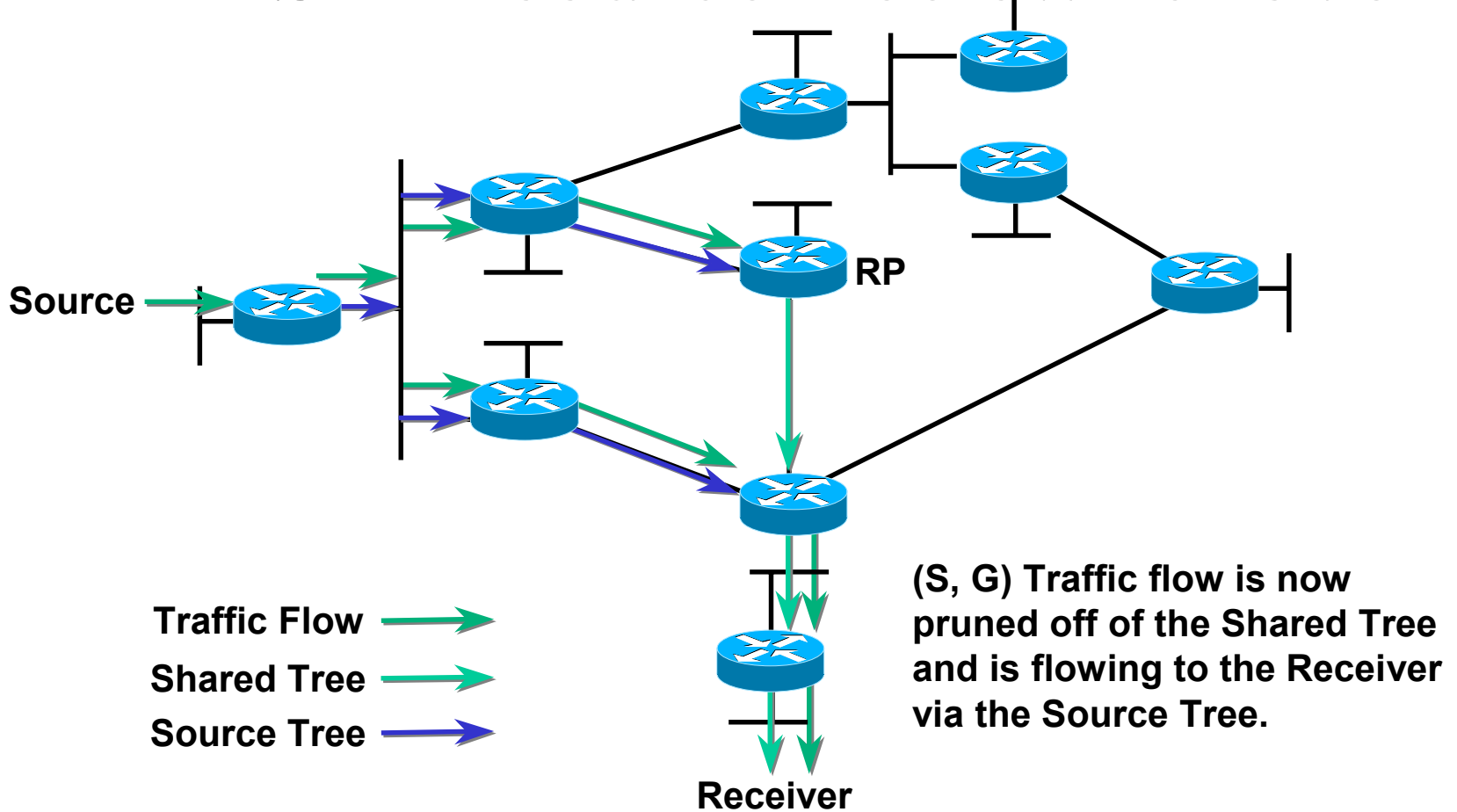
PIM-SM - source tree switchover



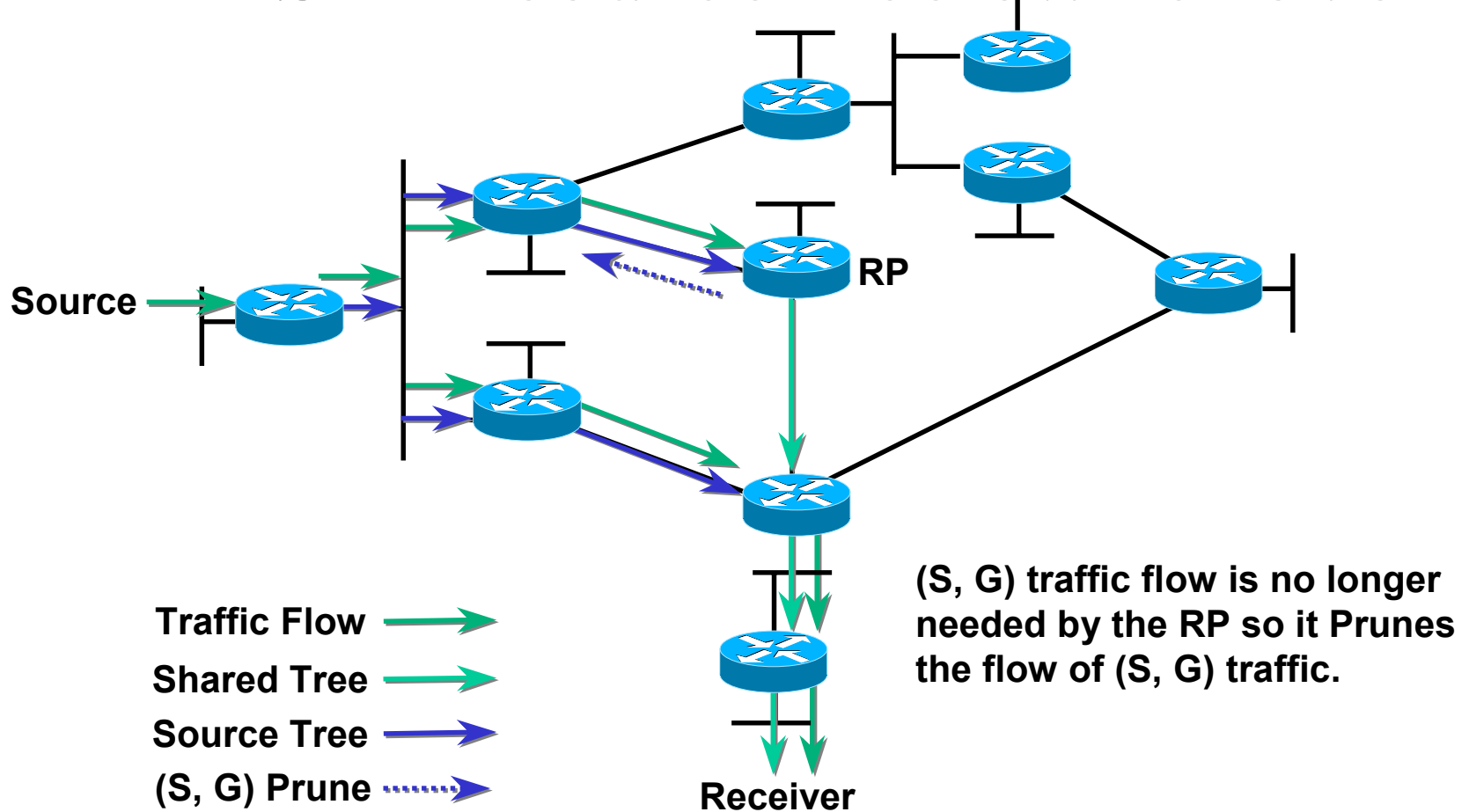
PIM-SM - source tree switchover



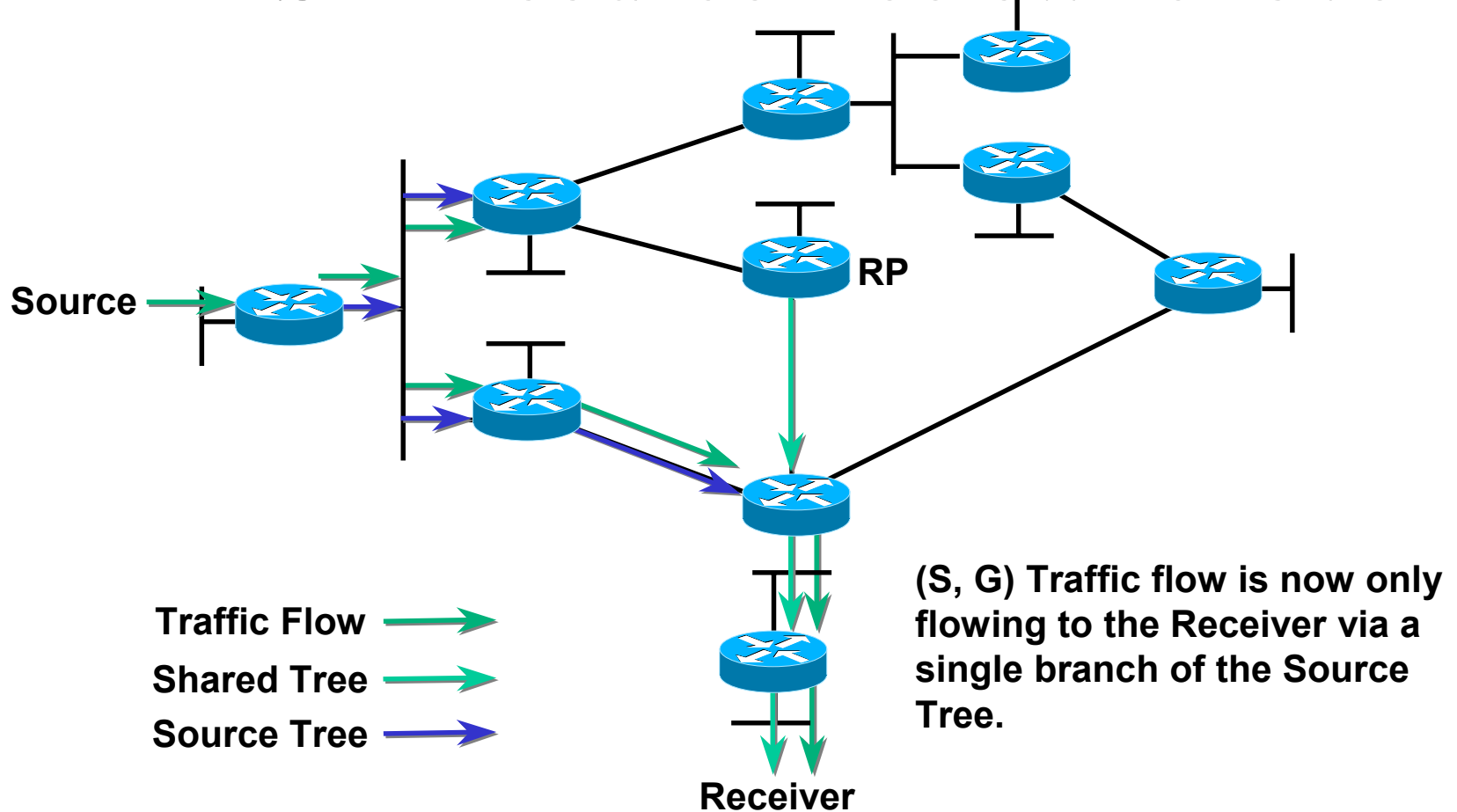
PIM-SM - source tree switchover



PIM-SM - source tree switchover



PIM-SM - source tree switchover



PIM-SM - RP placement

- Which is the best router to be the RP for your network?
 - Needs to be ‘well-connected’
 - In the ‘centre’ of your network?
 - But, if you want global multicast connectivity, your RP will need to talk to your peers’ RPs
 - So, at the edge?

PIM-SM - RP placement

- You can have more than one RP
 - You can dedicate one RP to local groups and place it centrally, and
 - You can have another for global groups and place it more to the edge
- Where ever it is, it needs to be part of your management domain

Agenda

- Introduction
- Multicasting within the network
- **Multidomain multicasting**
- Multicast on JANET

Multidomain Multicasting

- Domains are *management domains*
- Each domain has its own RP which is known to all of the routers in that domain
- The RP is not known to any last-hop routers in any other domain
- So, how do we get traffic flowing between domains?

Multidomain Multicasting

- For PIM-SM to work, the RP needs to know about the sources and receivers for a given group
- It knows about those within its own domain, but how can it know about those in other domains?

Multidomain Multicasting

- Receivers?
- The RP waits for receivers to join
 - then it can establish the shared tree to it
- Receivers in another domain?
 - It doesn't care - it's someone else's problem
 - Specifically, the RP in that domain

Multidomain Multicasting

- Senders?
- The RP sends joins to the senders, once it knows about them
- If the sender is in another domain, the sender will only register with its own RP
- Need a way for RPs to know about senders in other domains

Multidomain Multicasting

- RPs need to exchange information about which senders are active for which groups
- Current solution is to use the Multicast Source Distribution Protocol (MSDP)
- Extend BGP to exchange the routing information for the multicast sources as well as the usual interdomain unicast routes

MSDP

- Still an Internet Draft
- Not the final solution by any means....
But the best that we have at the moment

MBGP

- *Multiprotocol* BGP, not *multicast* BGP
- RFC2858 (June 2000) - Proposed Standard
- Extends BGP4 to carry multiple address families as well as IPv4 unicast:
 - Multicast source routes
 - VPN routes
 - IPv6 routes

Agenda

- Introduction
- Multicasting within the network
- Multidomain multicasting
- **Multicast on JANET**

Implementation in SuperJANET4

- The SuperJANET4 backbone
 - runs PIM-SM throughout
 - uses MBGP/MSDP to exchange source routes with external peers:
 - Internet2
 - KPNQWest
 - GÉANT (planned)

Implementation in SuperJANET4

- Each BAR in SuperJANET4
 - is configured as an RP
 - uses MBGP/MSDP to exchange source routes with the SuperJANET4 core
 - can use MBGP/MSDP to exchange source routes with the associated regional network
 - not currently the default - on request
 - will become the default after Easter

Multicast in JANET

- To join the JANET multicast service, a Regional Network must:
 - run PIM-SM on the connection from its border router to the BAR
 - configure its border router as an RP
 - extend the existing BGP session to use MBGP
 - configure a MSDP peering in parallel to the MBGP peering (use the same addresses)

Multicast in JANET

- To join the JANET multicast service, a Regional Network may
 - use the BAR as its RP
 - this is purely a temporary measure until the proper peerings can be put in place
 - PIM-SM is still necessary between the RN and the BAR
 - It may not work well!

Multicast in JANET

- Individual sites should liaise with their Regional Network
- If you currently have a tunnel into JANET, do this as soon as possible as the tunnel service is scheduled to finish in October 2002

Multicast in JANET

- In theory, it's relatively straightforward to establish the peering with the BAR, but
 - your router may not support MSDP/MBGP
 - lobby your supplier!
 - It is generally non-trivial to rollout native multicast across a network
 - start now!
- All of this is equally applicable to the site

Multicast in JANET

- The SuperJANET4 backbone is already running native, non-tunneled IP multicast
- Several regional networks are peering with the backbone and are providing multicast to their sites
- Not all, however.... is yours?

Acknowledgements

- UKERNA - for letting me play with all of this!
- Cisco - for letting me use some of the flow diagrams
- My wife - for being incredibly patient while I wrote this when I should have been gardening

Interdomain Multicasting in SuperJANET4

Jonathan J. Couzens

Principal Network Engineer

JANET Network Operations & Service

jjc@nosc.ja.net

+44 (0)20-7692-1316

