SSM Saves the World

Mark Williams
miw@juniper.net

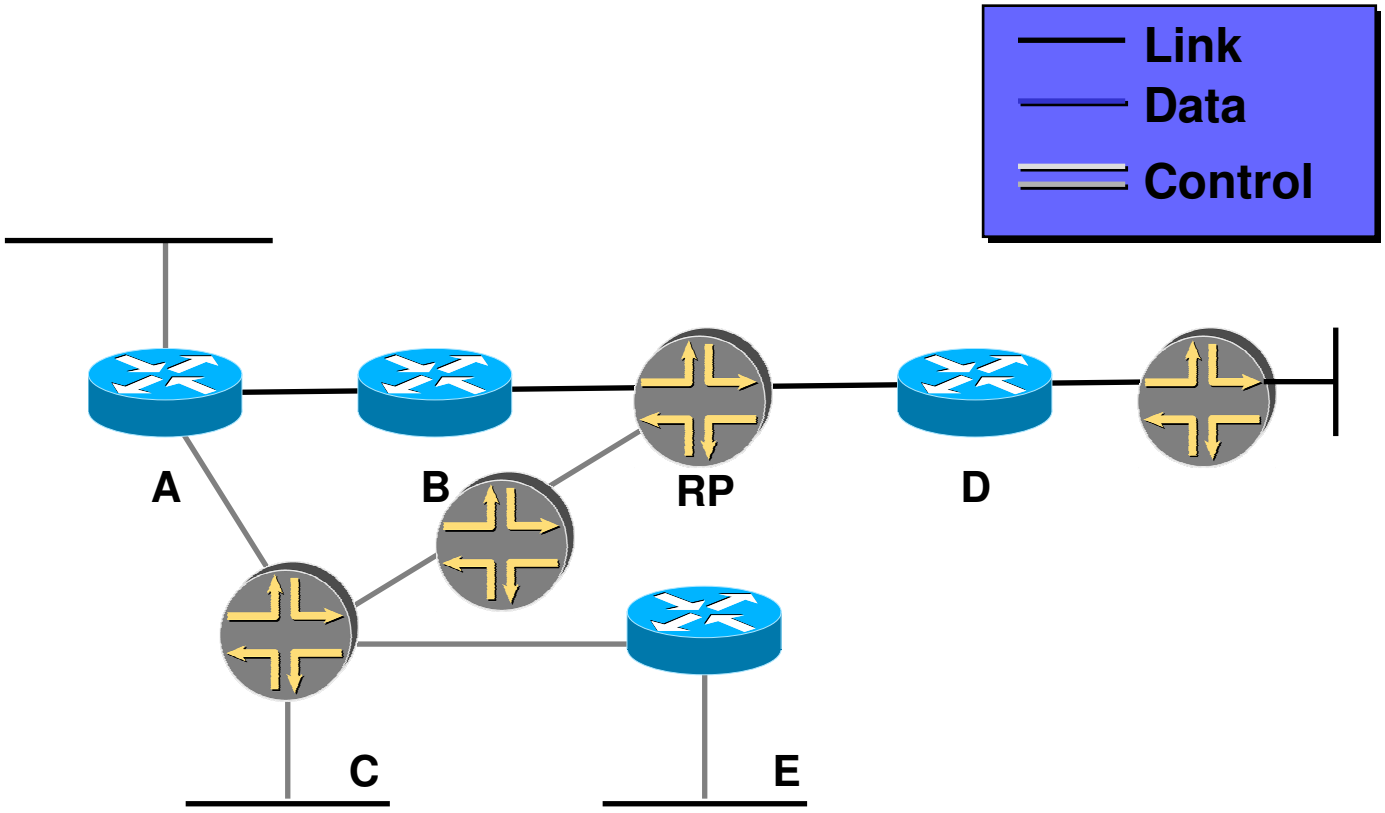
ASM and SSM

- **ASM: Any-Source Multicast. Traditional multicast – data and joins are forwarded to an RP.**
  - **All routers in a PIM domain must have RP mapping.**
  - **When load exceeds threshold, forwarding switches to an SPT. The default threshold is one packet; in this case, the sole purpose of the RPT is to learn which sources are active. (With IGMPv2, the receiver can only specify the group, not specific sources.)**
  - **State increases (not everywhere) as number of sources and number of groups increase.**
  - **SPT state is refreshed when data is forwarded and with Join/Prune control messages.**
- **SSM: Source-Specific Multicast. PIM-SM without RPs – instead, the source is learned out-of-band, and the SPT is built directly to it.**

ASM: the original multicast service model

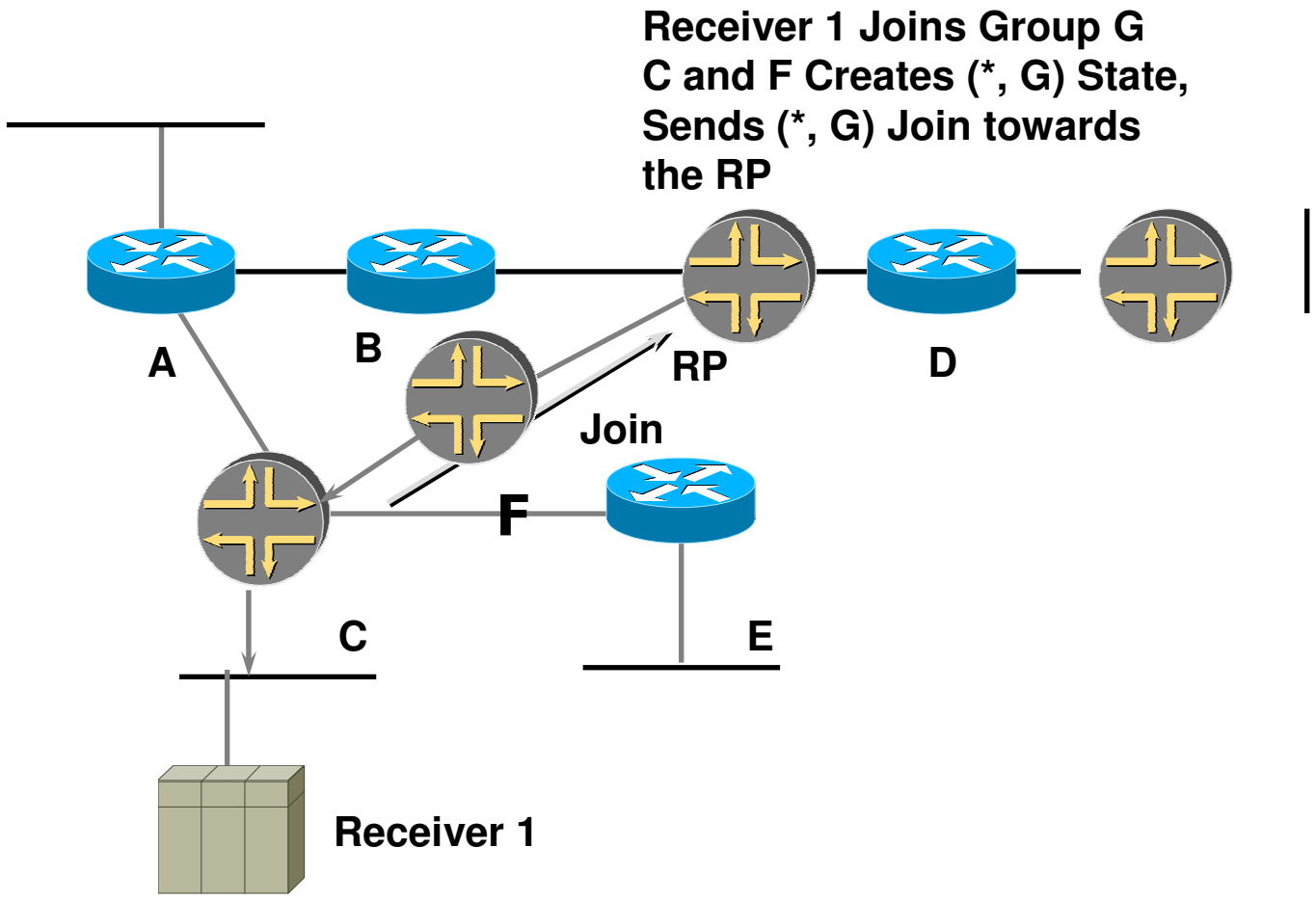
- From RFC 1112 :
  - Packet transmission is based on UDP, so packet delivery is "best-effort", with no loss detection or retransmission
  - A source can send multicast packets at any time, with no need to register or schedule transmissions.
  - Sources do not know the group membership. A group may have many sources and many members.
  - Group members may come and go at will, with no need to coordinate with a central authority.
  - And, critically, group members know only the group. They don't need to know anything about sources — not even whether or not any sources exist.

PIM Sparse Mode Review



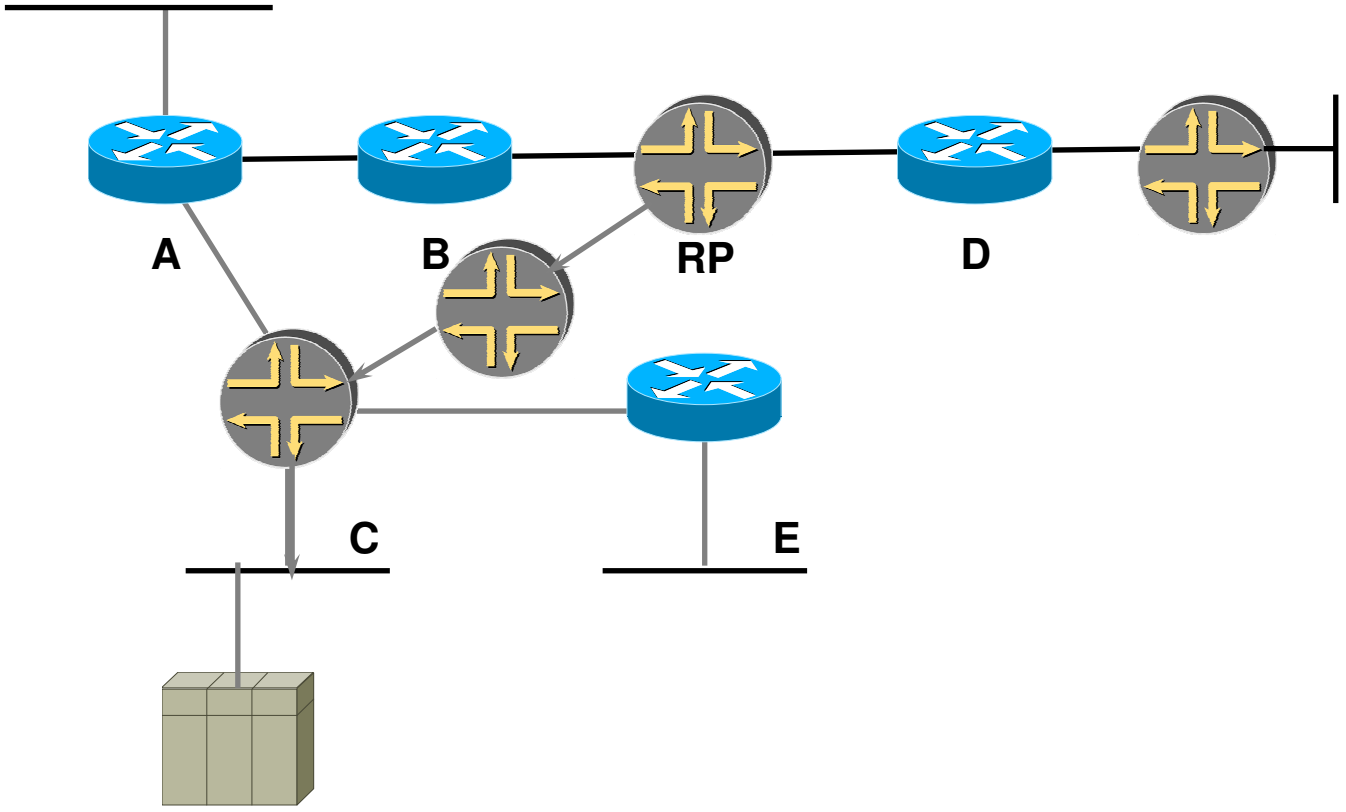
The Rendezvous Point (RP) is the shared root, and is administratively assigned.

PIM Sparse Mode Review

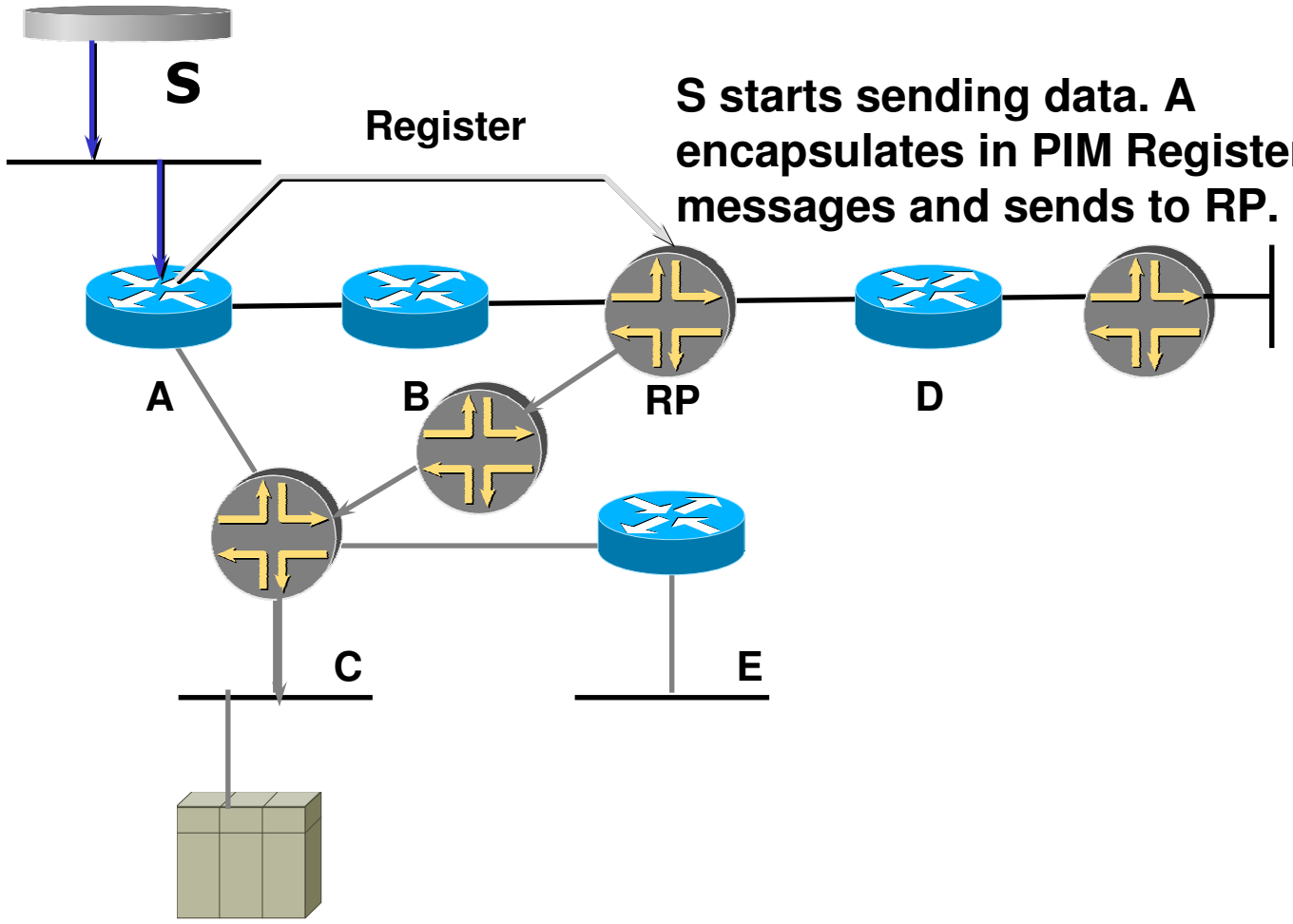


PIM Sparse Mode Review

RP creates (*,G) state.

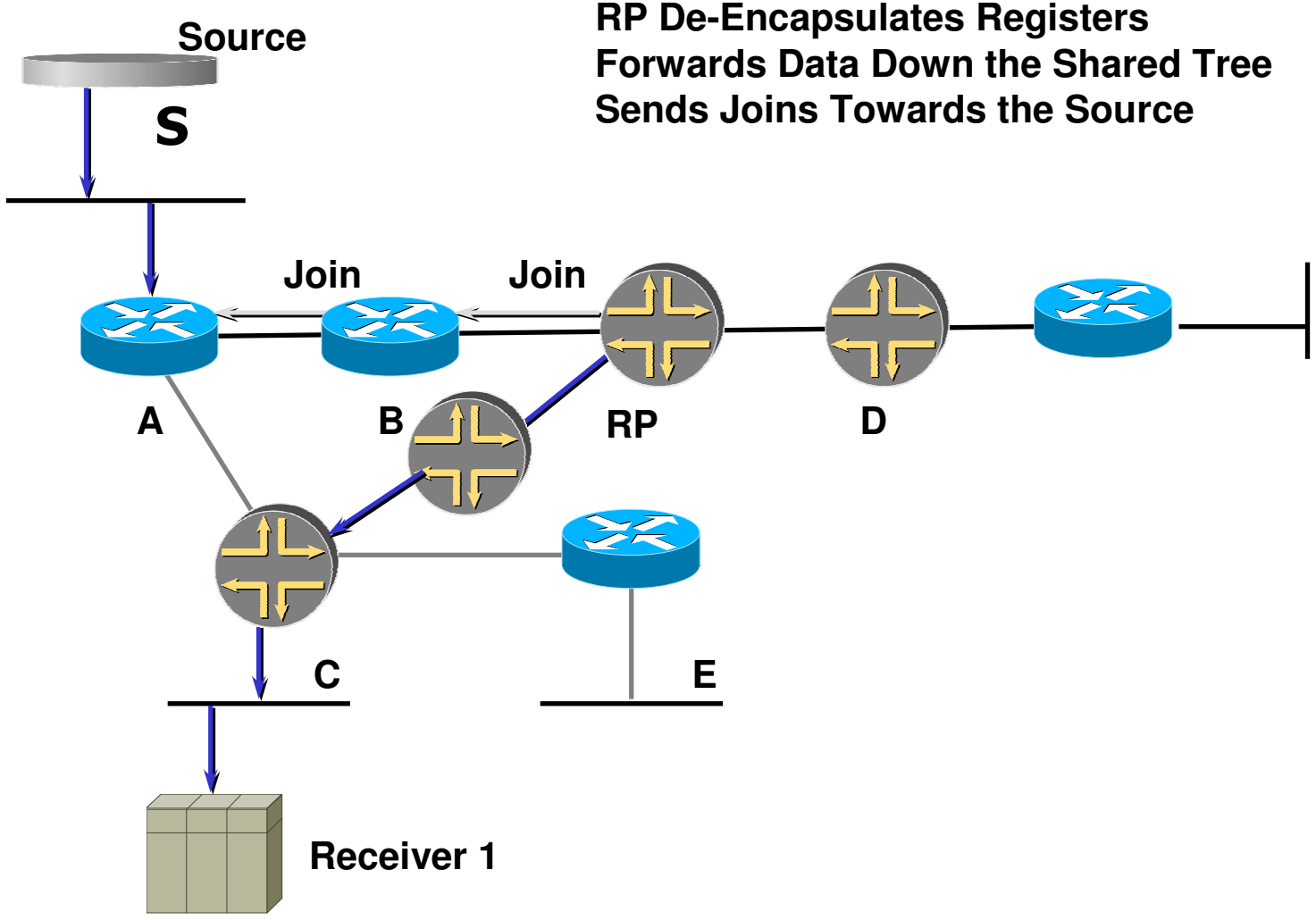


PIM Sparse Mode Review

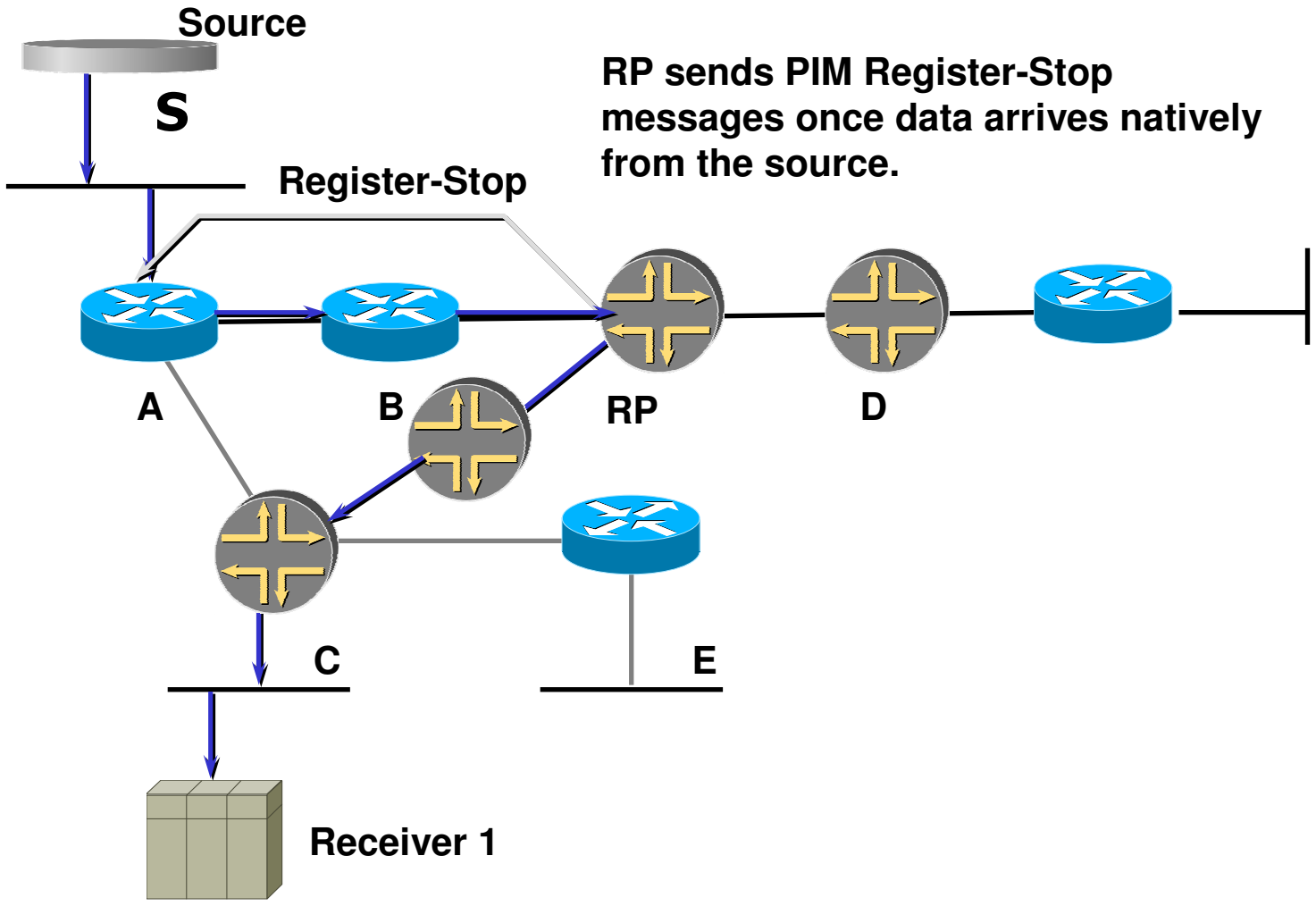


PIM Sparse Mode Review

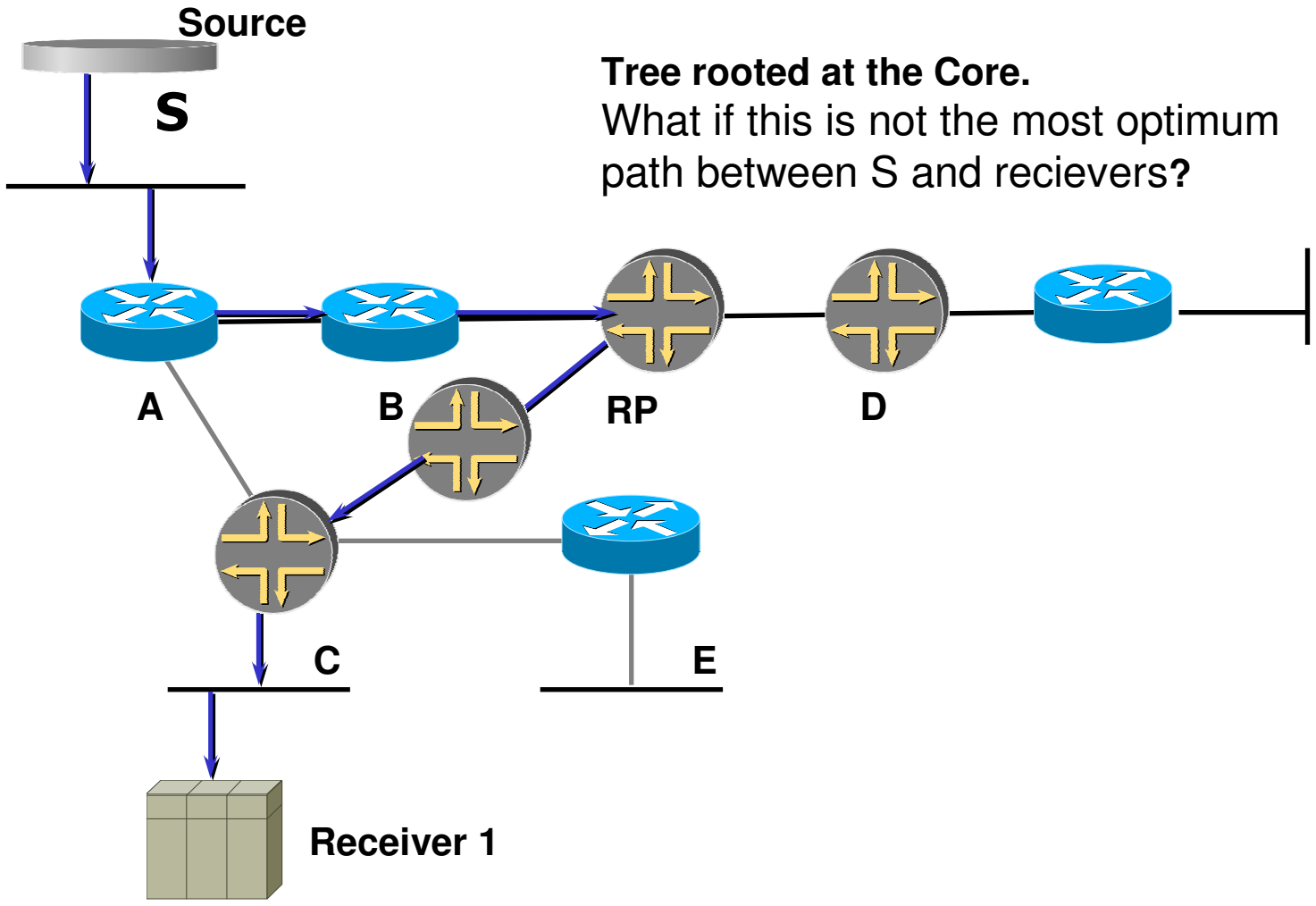
RP De-Encapsulates Registers
Forwards Data Down the Shared Tree
Sends Joins Towards the Source



PIM Sparse Mode Review

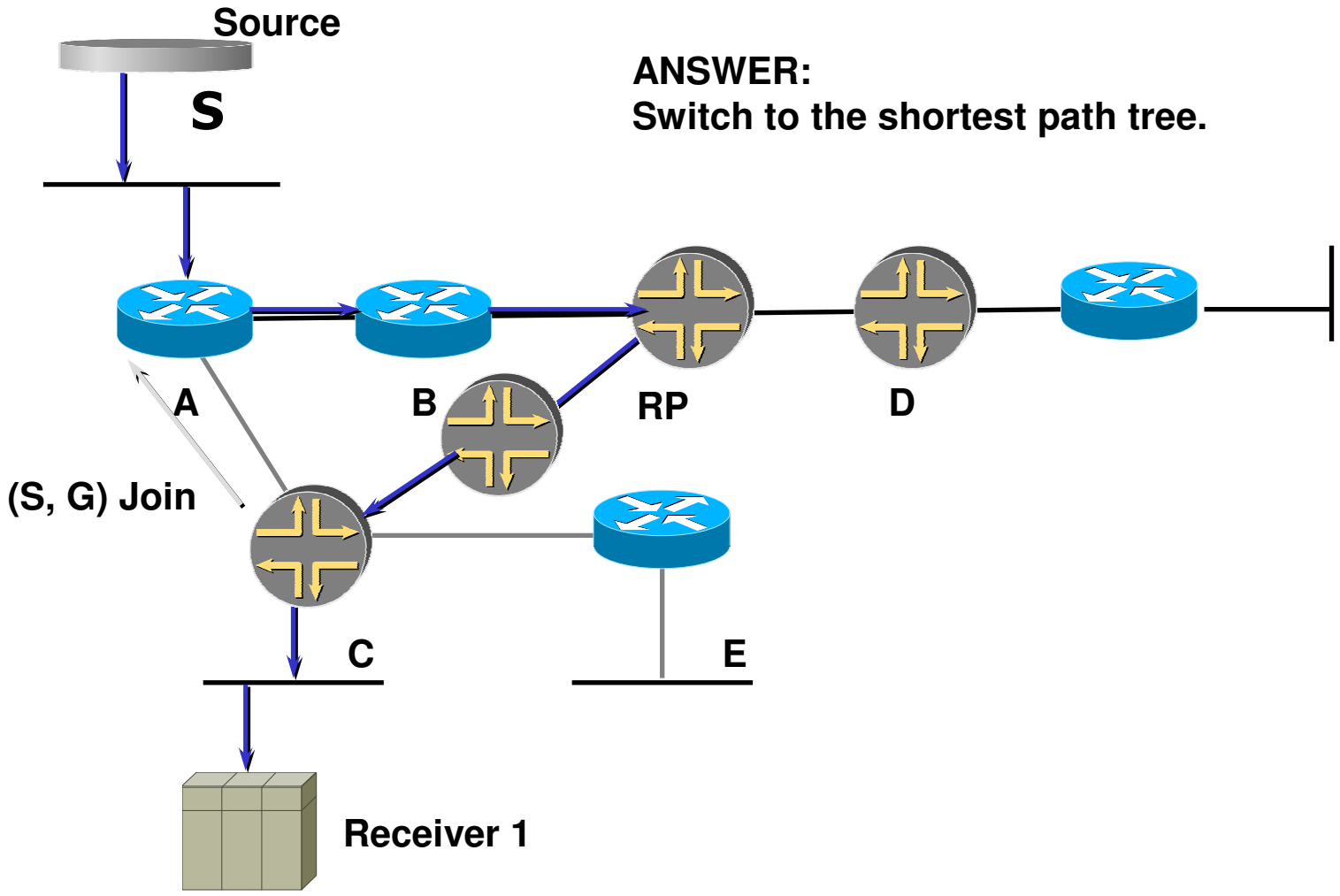


PIM Sparse Mode Review

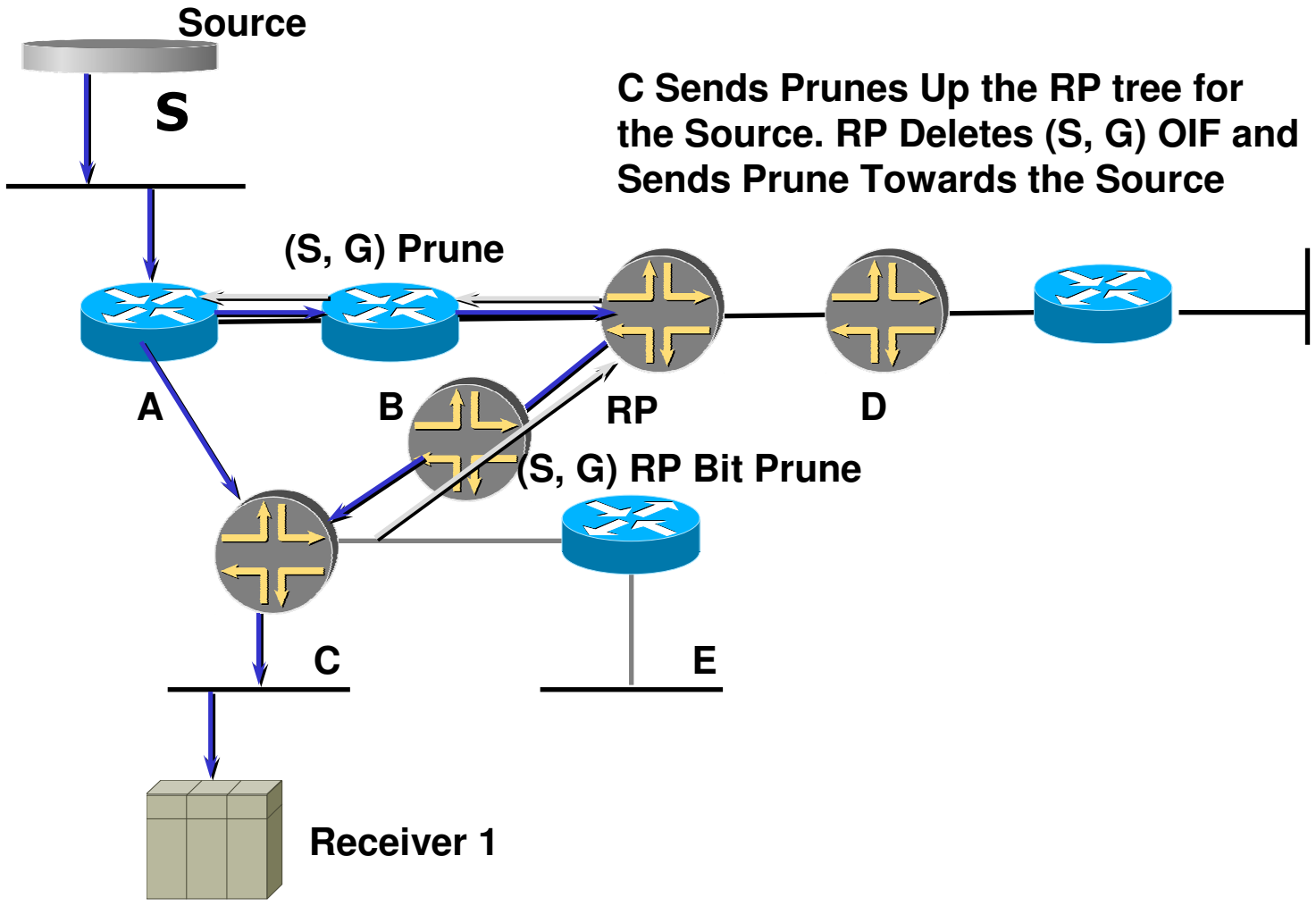


Tree rooted at the Core.
What if this is not the most optimum path between S and recievers**?**

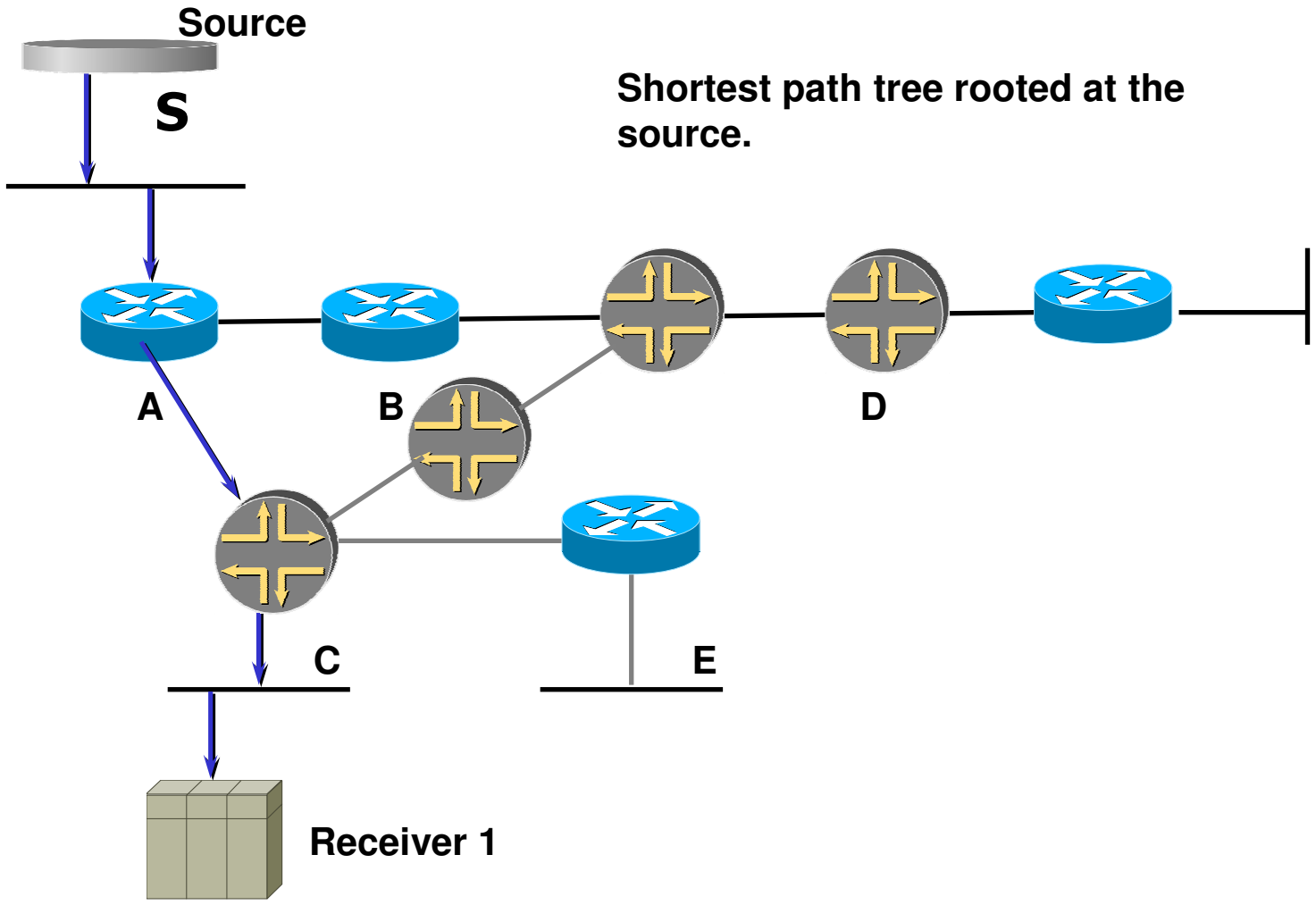
PIM Sparse Mode Review



PIM Sparse Mode Review



PIM Sparse Mode Review



MSDP

- Multicast Source Discovery Protocol
- Used for PIM SM inter-domain support
  - Remember how PIM SM works - Designated router (receive side) initiates the request for the flow to the RP. How will the RP in your domain know about an active source in a neighbor's domain?
    - OPTION 1: RP's in my domain tell your RP's of active source(s) (that's MSDP)
    - OPTION 2: or I simply flood everything to your domain and turn your edge PIM router into the sources DR. This can be done via running DVMRP on the links interconnecting the 2 PIM SM domains.

MSDP Operation

- MSDP peers between RP's (inter or intra domain)
  - (TCP port 639 w/ higher IP addr LISTENS)
- "FLOOD & join"
  - SA (source active) packets periodically sent to MSDP peers indicating:
    - source address of active streams
    - group address of active streams
    - IP address of RP sending the SA
    - only send SA's for your sources w/in your domain (full mesh)
- "flood & JOIN"
  - interested parties can send PIM JOIN's towards source (creates SPT trees between domains)

MSDP

- **Multicast Source Discovery Protocol**
  - **Allows each domain to controls its own RPs**
  - **Interconnect RPs between domains with TCP connections to pass active source info**
  - **Can also be used within a domain to provide RP redundancy**
  - **RPs learn about external source and join source tree externally**
  - **Receivers learn about external source through internal shared tree then may also join external source tree**
  - **MSDP connections typically parallel MBGP connections**

MSDP Source Active Msgs

- Initial SA sent when source first learned
  - May optionally encapsulate first data packet
- Subsequent SA messages periodically refreshed every 60 seconds as long as source still active by originating RP
- Other MSDP peers don't originate this SA but only forward it if received
- SA messages cached on router for new group members that may join

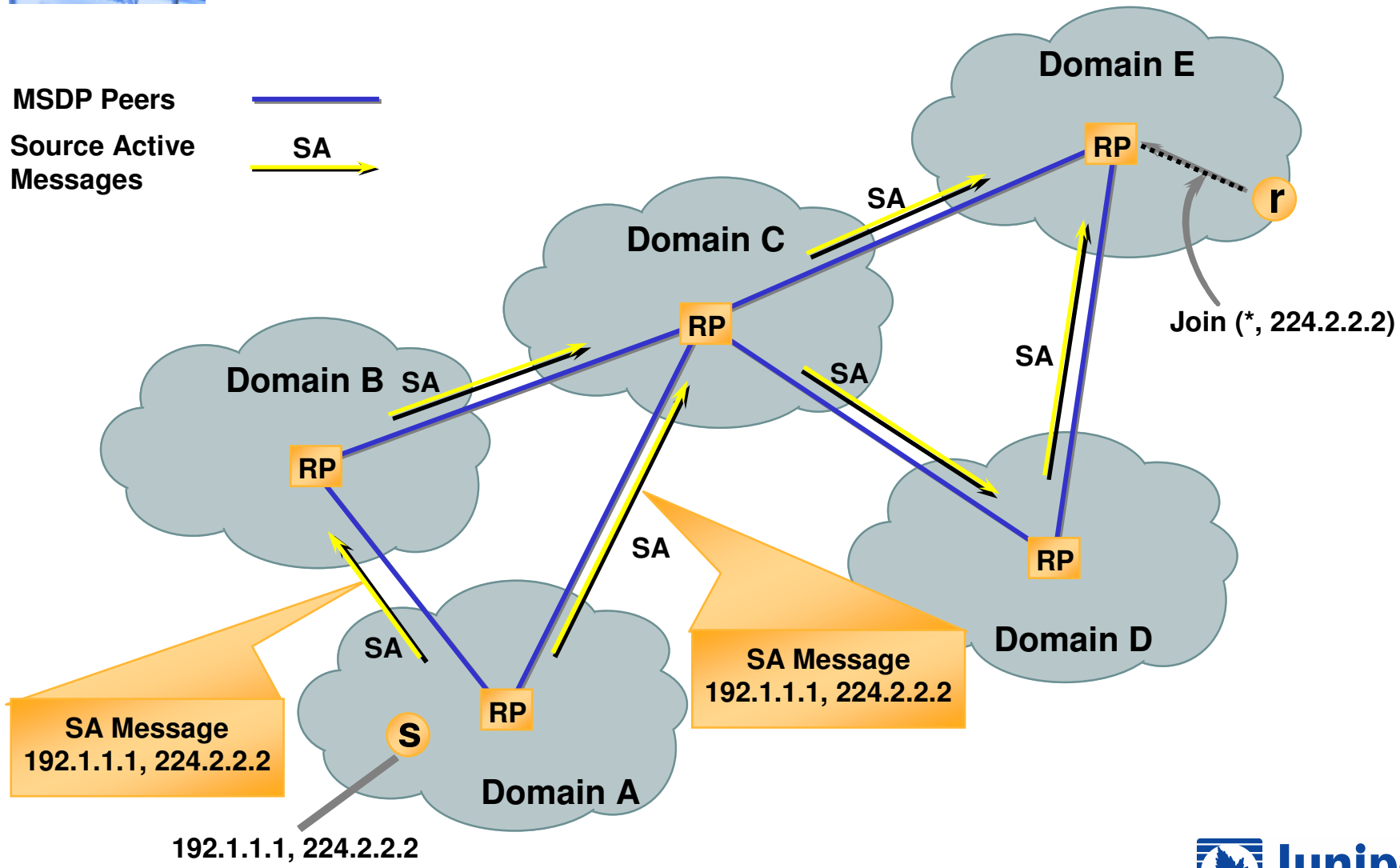
MSDP Peers

- MSDP establishes a neighbor relationship between MSDP peers
  - Peers connect using TCP port 639
  - Peers send keepalives every 60 secs (fixed)
  - Peer connection reset after 75 seconds if no MSDP packets or keepalives are received
- MSDP peers must run BGP!
  - May be an MBGP peer, a BGP peer or both
  - Exception: BGP is unnecessary when peering with only a single MSDP peer.

MSDP Design Points

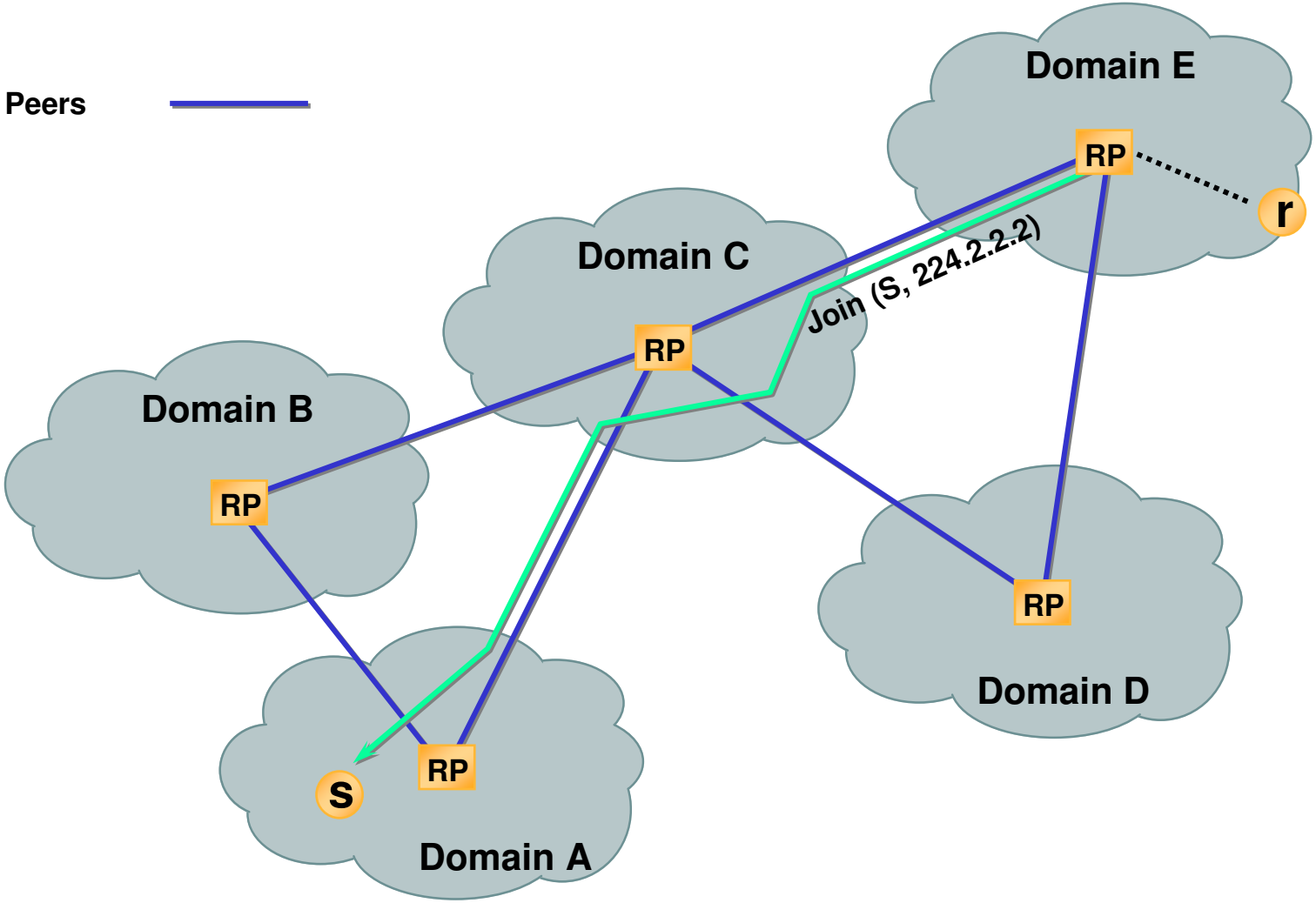
- MSDP has all the scalability problems of a flooding protocol.
- MSDP peers talk via TCP connections
  - UDP encapsulation option
- Source Active (SA) messages
  - Peer-RPF forwarded to prevent loops
    - RPF check on AS-PATH back to the peer RP
    - If successful, flood SA message to other peers
    - Stub sites accept all SA messages
      - Since they have only one exit (e.g., default peer)
  - MSDP speaker may cache SA messages
    - Reduces join latency

MSDP Example

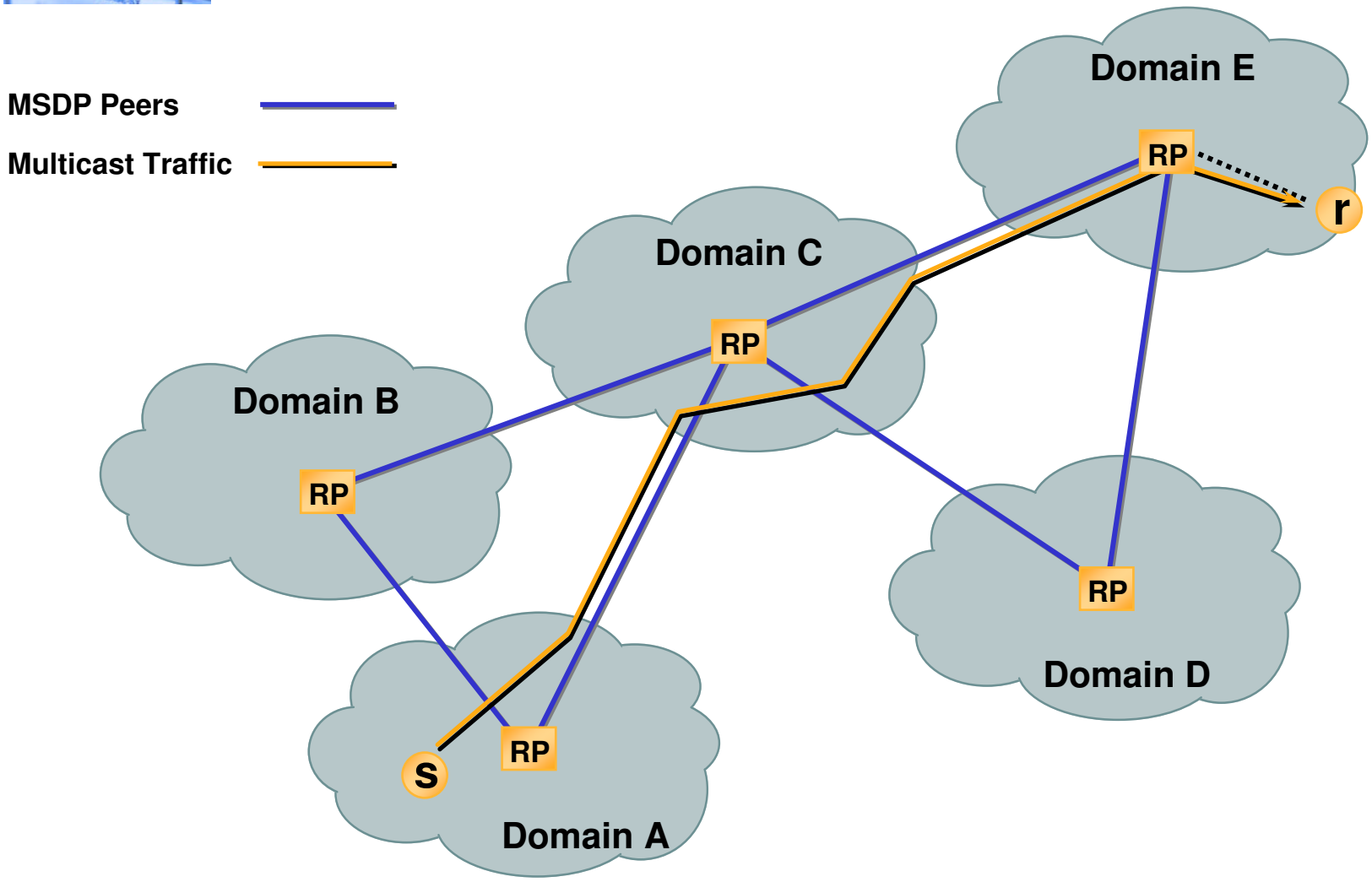


MSDP Example

MSDP Peers



MSDP Example



What's wrong with this picture?

- NOC can't manage this
- Traffic concentration at the RP
- Flooding 150,000 active sources globally over MSDP doesn't scale
- Group address allocation not deployed on large scale
- Lots of Join/Prune state in the network

SSM

- SSM - draft-ietf-ssm-arch-01.txt
  - SSM behavior in 232/8 by default
    - Configurable to expand range
  - No RPTs
  - Guarantees ONE source on any delivery tree
    - Content security – no unwanted sources
  - Reduced protocol dependence – more later…
  - Solves address allocation issues for inter-domain one-to-many
    - tree address is 64 bits – S,G
  - Host must learn source address out-of-band (e.g, from a web page)
  - Host-to-router join request specifies source as well as group
  - IGMPv3 allows for "Include" lists of (S,G) pairs.
  - Source-Specific Multicast (SSM) is a subset of ASM

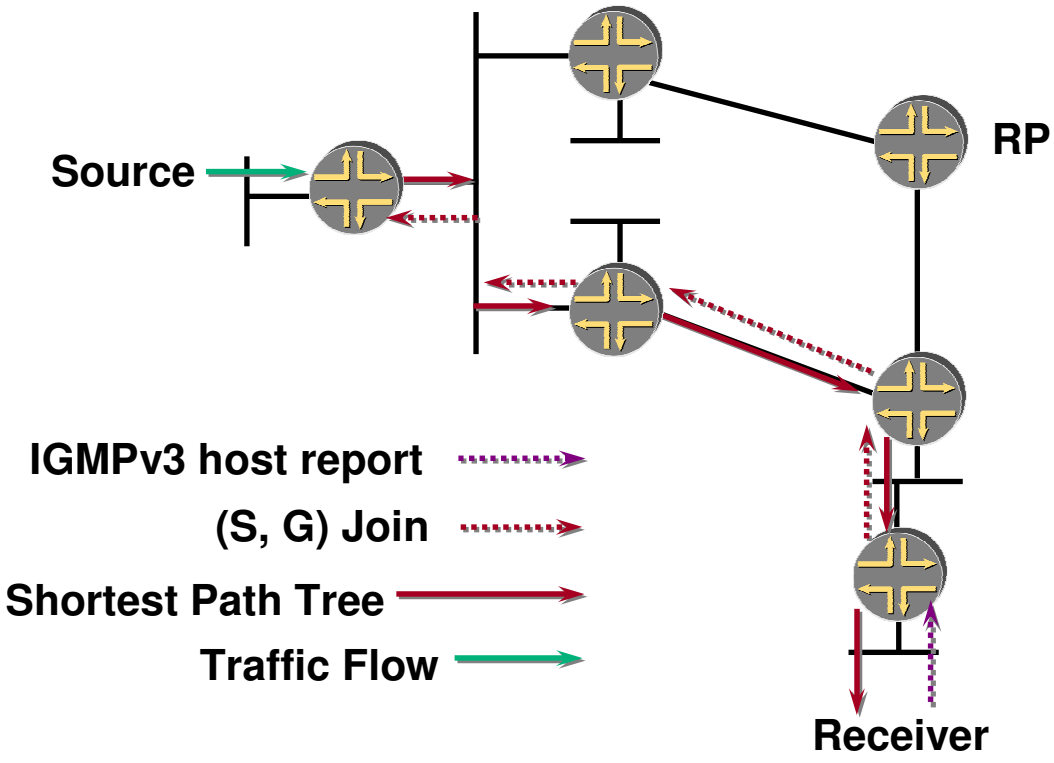
SSM in Action

- Each (S,G) pair listed in the IGMPv3 include list generates a (S,G) Join directly towards the source.
- That's it. It's very simple. All you need to implement is :
  - Edge routers need IGMPv3
  - Interior routers need filters to prevent RP (*,G) Joins & other RP state for the SSM address block

SSM Group Addresses

- 232 / 8 is assigned to SSM as an address space.
  - You don't have to ask, you can just pick one and use it.
- How can this be ?
  - Note that all joins are unique as long as the combination of S and G are unique. Not only can one source support multiple groups, but if there are two sources using the same group address, everything works just fine.

SSM

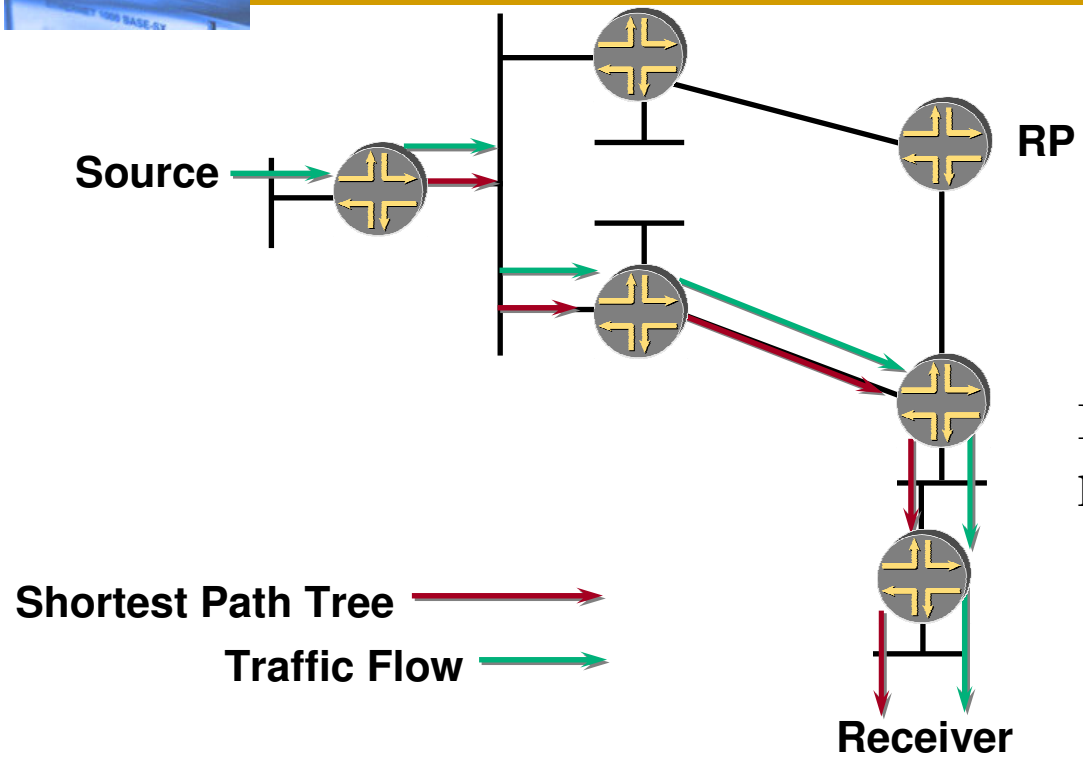


Receiver announces desire to join group G AND source S with an IGMPv3 include-list.

Last-hop router joins the SPT.

(S,G) state is built between the source and the receiver.

SSM

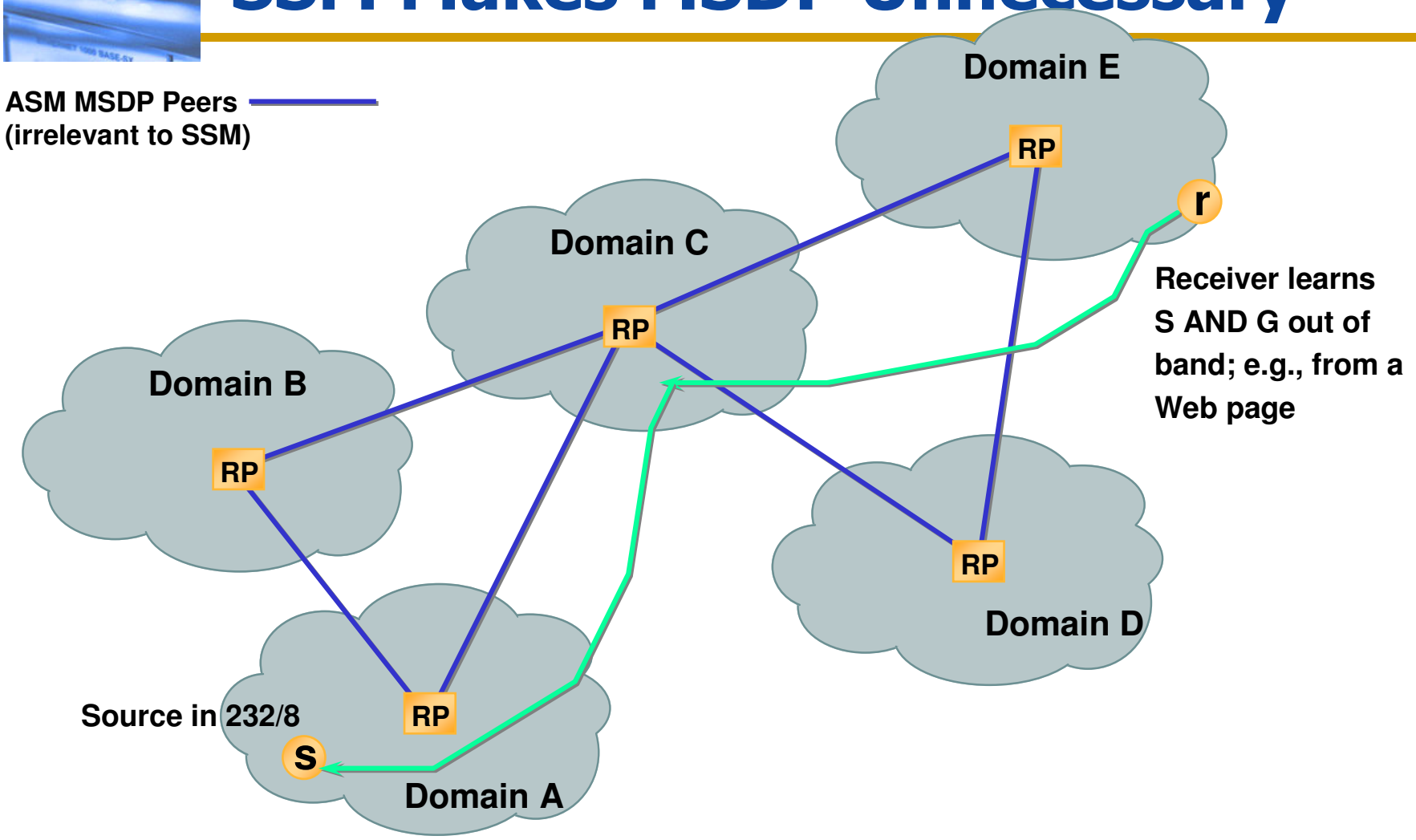


Data flows down the shortest path tree to the receiver.

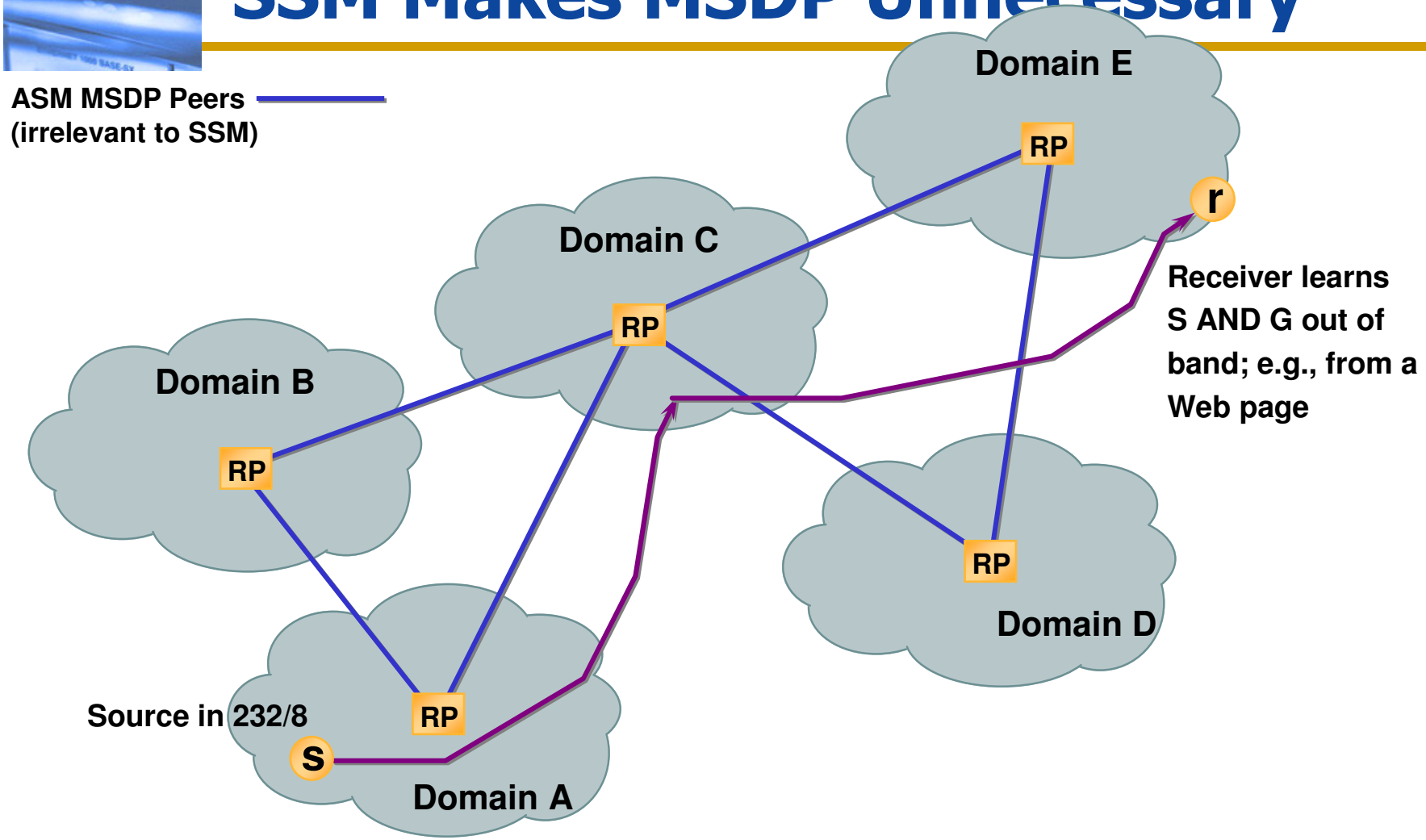
Rationale for SSM

- **Why go to all the trouble involved in using RPs (tree-switching, MSDP) when the RPT is dropped for the SPT as soon as the first packet flows down the RPT?**
  - **The RP is not really forwarding data, just doing source discovery.**
  - **Isn't there an easier way ?**
- **This is the rationale for Source-Specific Multicast (SSM).**

SSM Makes MSDP Unnecessary



SSM Makes MSDP Unnecessary



Summary: Advantages of SSM

- No RPTs
- No register packets
- No RP mapping required (no RP required!)
- No RP-to-RP source discovery (no MSDP required!)
- No RP means no concentration of traffic towards the RP, and no single point to attack
- Rogue sources cannot easily spoof traffic
  - Solves the DoS problem
- SSM can use entire multicast address space, but 232/8 is reserved for SSM exclusively
- Backward-compatible with ASM

What's the catch?

- No Catch!
- Routers use (S,G) forwarding state regardless so this reduces state
- In lab testing, Juniper routers have demonstrated over 150,000 (S,G) pairs
- No need for MSDP or RPs
- Router changes only required at the last-hop for a receiver.
- …One catch: IGMPv3 needed on all clients
  - Windows XP

Who needs to add IGMPv3?

- Routers with directly attached receivers
- Host OS need to support IGMPv3
- Switches doing IGMPv2 snooping
  - Necessary when host OS switches from IGMP v2 to v3
- DSL aggregation or other IGMP proxies currently passing on IGMP v1/v2 membership reports

Resource Location

- IGMP
  - IGMPv1- RFC 1112
  - IGMPv2- RFC 2236
  - IGMPv3- RFC 3376
- DVMRP
  - draft-ietf-idmr-dvmrp-v3-11.txt
- PIM-DM
  - draft-ietf-pim-dm-new-v2-02.txt
- PIM-SM
  - draft-ietf-pim-sm-v2-new-08.txt
- MSDP
  - draft-ietf-msdp-spec-14.txt

Resource Location

- MSDP
  - draft-ietf-msdp-spec-14.txt
- MBGP
  - RFC2858
- Anycast RP
  - draft-ietf-mboned-anycast-rp-08.txt
- SSM
  - draft-ietf-ssm-arch-01.txt
- Multicast On Book
  - *Interdomain Multicast Routing: Practical Juniper Networks and Cisco Systems Solutions* (Addison-Wesley)

Juniper® NETWORKS

Thank you!

http://www.juniper.net