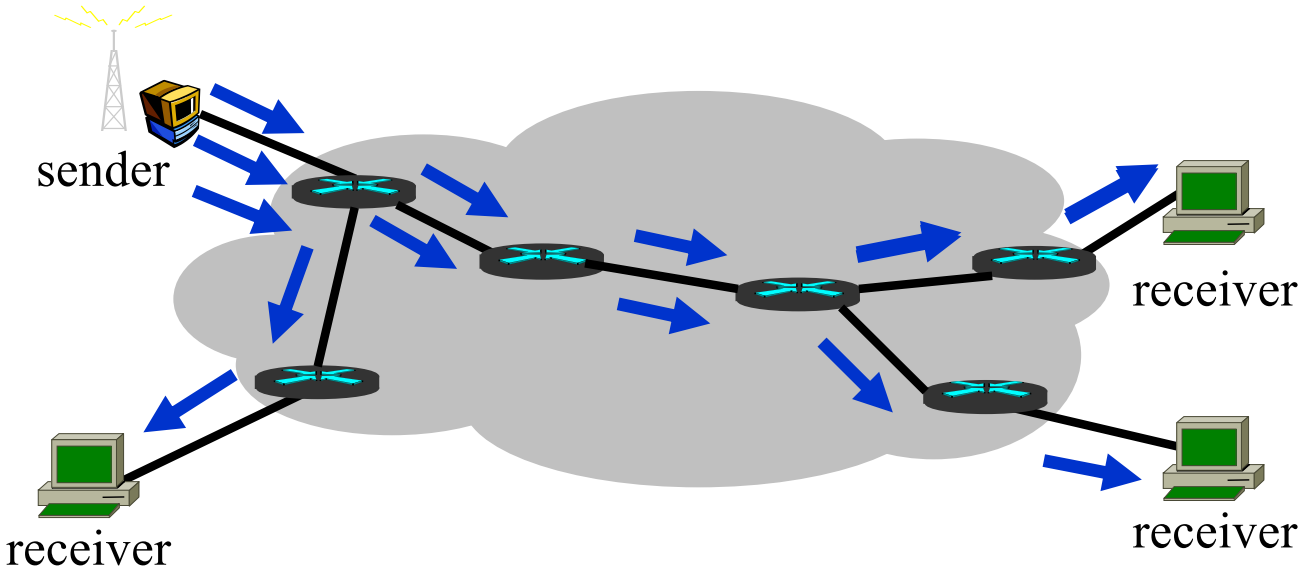
# CSC 525:
# Principles of Computer Networks

# Multicast

- Unicast vs. Multicast
- Objective:
  - Efficient *one-to-many* and *many-to-many* delivery of same data at roughly the same time
  - The concept of "group", i.e., intended receivers of the data.
- Challenges: Multi-party communication with wide range of application characteristics
  - Newss, teleconference, distance learning, resource discovery, etc.
  - Number of group members (large or small groups)
  - Distribution of group members (dense or sparse groups)
  - Dynamic group membership
  - Hard to optimize for all application scenarios in one design.
  - routing, end-to-end reliability, security, etc.

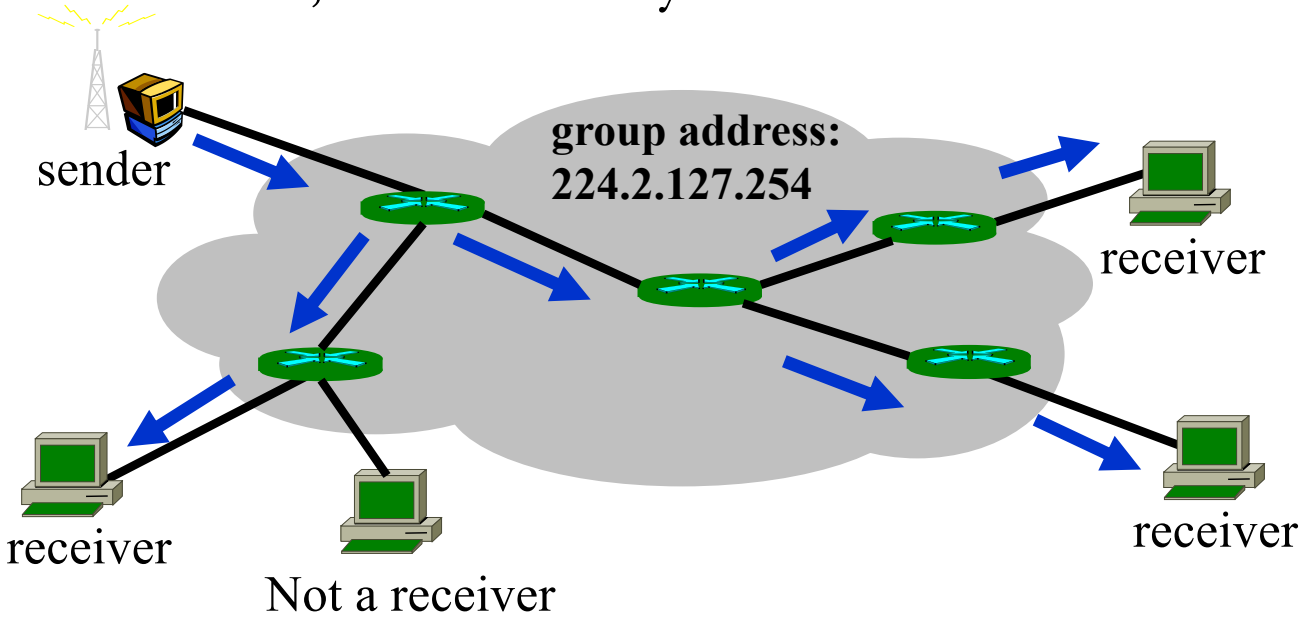
# Repeated Unicast

- Easy to implement and deploy
- Doesn't scale to the number of group members
  - Bandwidth bottleneck at the sender
  - Sender has to keep track of every receiver's join and leave.



# IP Multicast

- Use in-network replication to achieve efficient transmission
  - Only make data copy at branching routers.
  - At most one copy of data on any link, regardless of group size.
- Require router support
  - Proposed in late 80's, implemented in most OS and routers, deployed in some networks, but not widely in the Internet across ISPs.



# IP Multicast Service Model

- Each group is identified by an IP address (class D)
- Groups may be of *any* size
- Members of groups may be located *anywhere* on the Internet
- Members can join and leave *at will*
- Senders need *not* be members
- Group membership is not explicitly known.

analogy: each multicast address is like a radio frequency, on which anyone can transmit, and to which anyone can tune in.

# IP Multicast Addresses

- Class D IP addresses:


224.0.0.0—239.255.255.255

- Two administrative categories:
  - "well-known" multicast addresses, assigned by IANA, e.g.
    - 224.0.0.1 all systems on a LAN
    - 224.0.0.2 all routers on a LAN
  - Ephemeral multicast addresses, assigned and reclaimed dynamically by applications

# Unicast vs. Multicast (address)

- Unicast IP addresses are allocated to networks in blocks, i.e., prefixes.
  - A host obtains an address from the network it is connected to, and keeps using the address as long as it is connected to the same network.
- Multicast IP addresses are needed not for connectivity, but for applications.
  - A host needs to join a group (i.e., use a multicast address) when it runs a particular application.
  - A host may use and stop using different multicast addresses even when it is connecting to the same network.

# Multicast Address

- An IP Multicast address is a *logical* group address
  - Identify the shared interest in particular traffic.
  - Not tied to any particular hosts
  - No structure in the address
- Good for resource discovery
  - How to discover an available application server without knowing any individual server?
  - Servers join a multicast group
  - Clients send queries to the group address
  - The first response comes from the closest server.
- But it also causes significant challenge to routing: figuring out where those member hosts are in order to deliver packets to them.

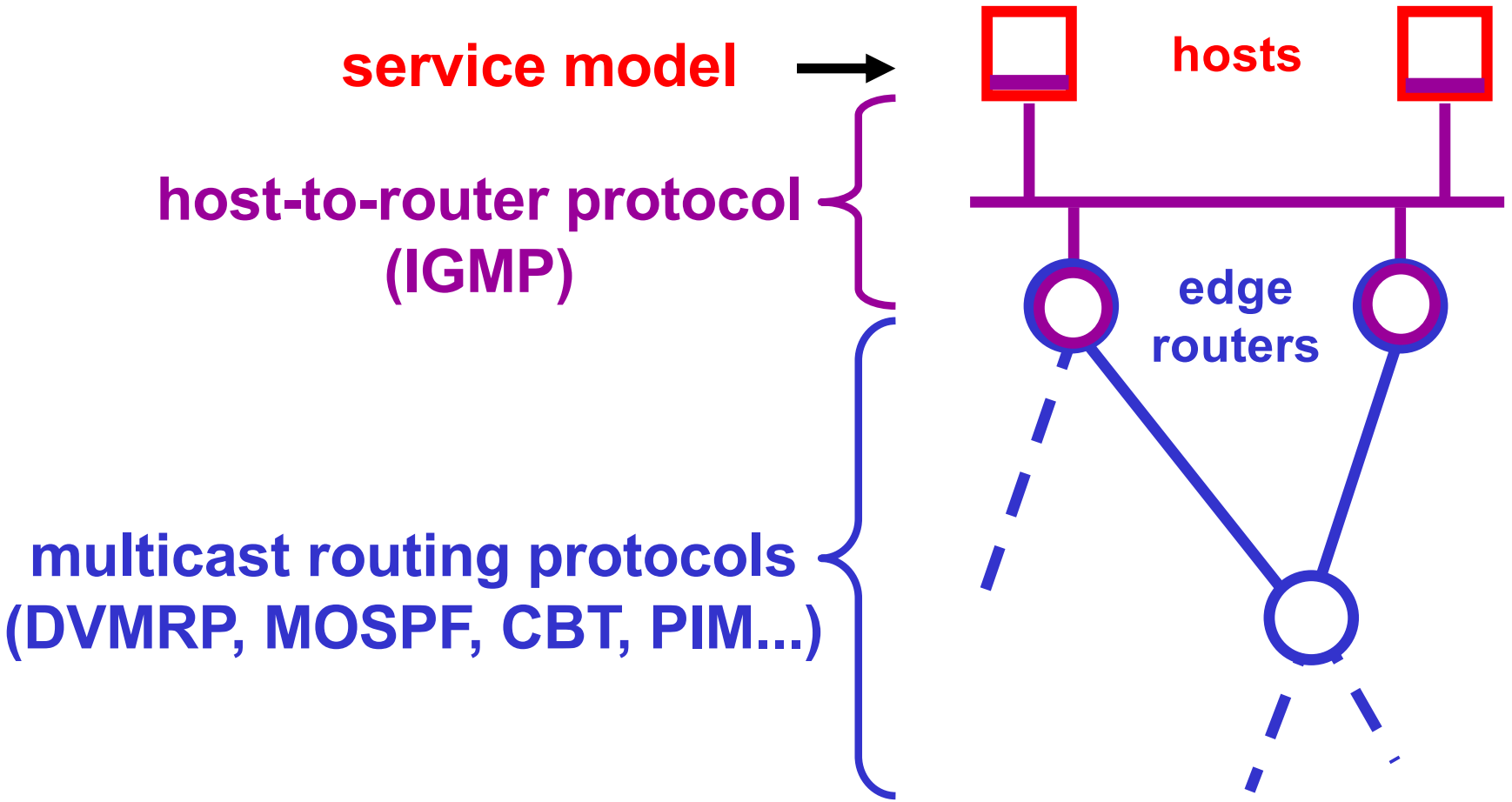
# IP Multicast in LAN

- Receive
  - Two socket options, IP_ADD_MEMBERSHIP and IP_DROP_MEMBERSHIP, to signal the OS and network interface about group membership.
  - The group address will be added to the network card.
  - Network card will pick up packets whose destination is the group address.
- Send via UDP socket; cannot have TCP connections.
- Replacing ARP with static address mapping
  - 224.0.0.0—239.255.255.255 ⟷ 01:00:5E:00:00:00—01:00:5E:7F:FF:FF
  - share the same low-order 23 bits. One Ethernet address will map to 32 IP mcast addresses.
  - The network card will receive pkts destined to all 32 Ethernet addresses, and the OS will do further filtering.

# Unicast vs. Multicast (routing)

- Unicast routing:
  - A network holds a prefix, announces it in routing.
  - Other routers select best path to reach this prefix.
  - Hosts have addresses sharing the same network prefix.
- Multicast routing:
  - Cannot have a multicast prefix associated with a network, because hosts may use and stop using any multicast prefix at any time, depending on the applications.
  - A network needs to keep track of the groups (addresses) that its hosts are interested in.
  - The routing system needs to build a data dissemination tree from the sender to all the receivers.

# IP Multicast Architecture

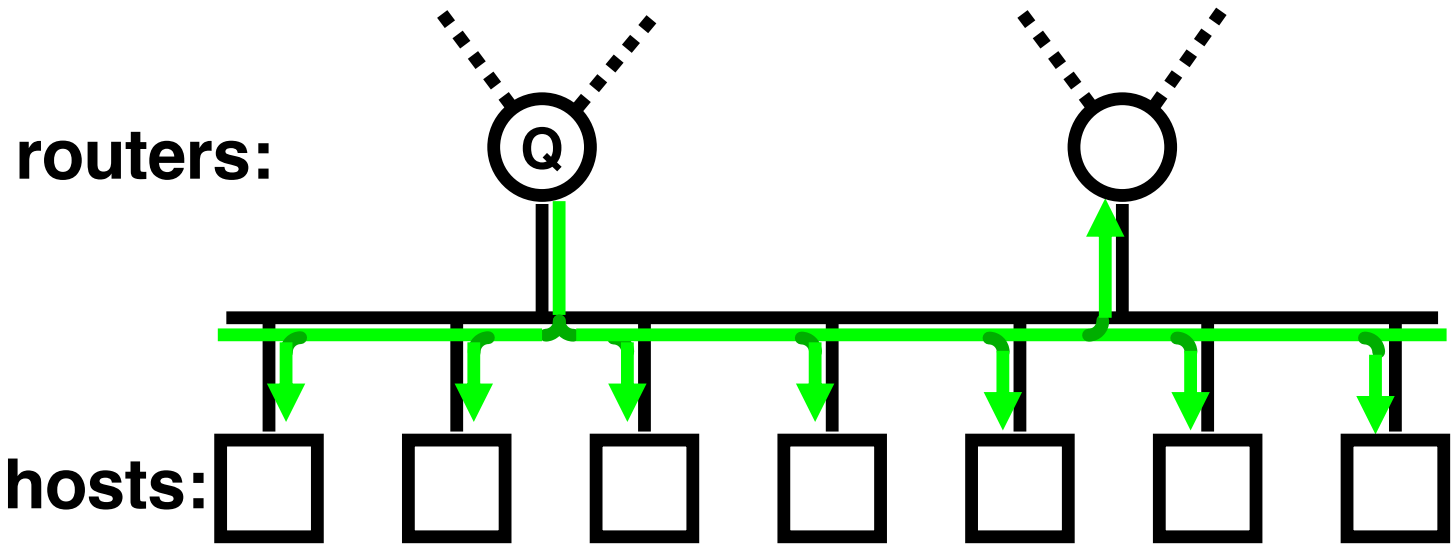


# Internet Group Management Protocol (IGMP)

- IGMP: "*signaling*" protocol to establish, maintain, remove groups on a subnet.
- Objective: keep router up-to-date with group membership of entire LAN
  - Routers need not know who the members are, *only that members exist*
- Each host keeps track of which multicast groups are subscribed to by its own applications
  - via multicast socket options.
- Hosts report membership to the router.
- Router then run routing protocol to get group traffic from outside.

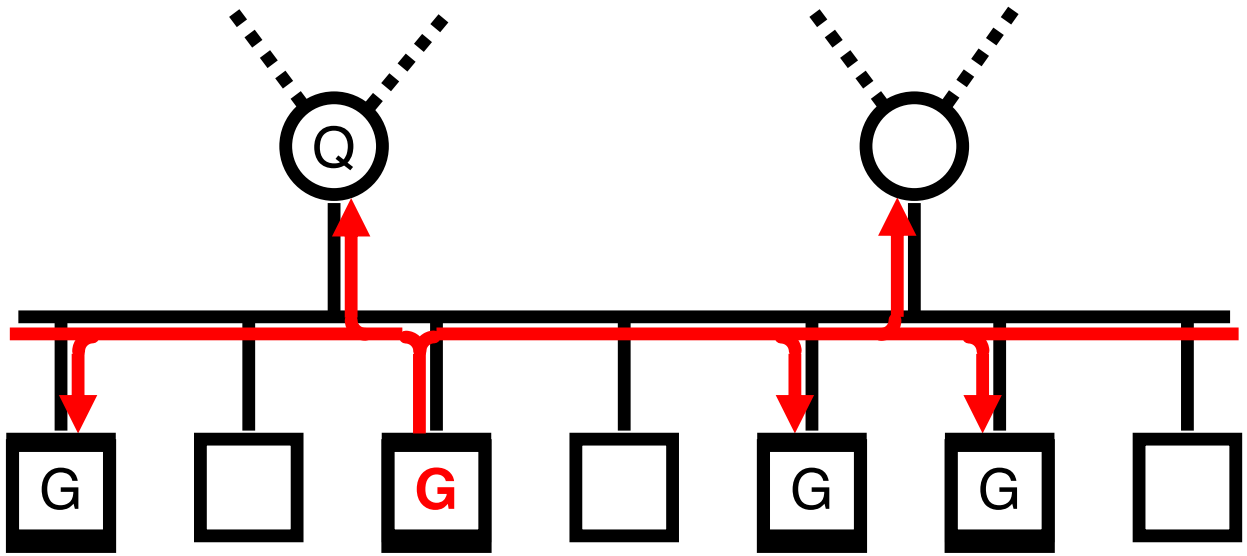
# How IGMP Works

- one router is elected the "querier" on each LAN
- querier periodically sends Membership Query message to the all-systems group (224.0.0.1), with TTL = 1
- on receipt, hosts start a random timer [0, 10 sec] for each multicast group to which they belong



# How IGMP Works (cont.)

- when a host's timer for group G expires, it sends a Membership Report to group G, with TTL = 1
- other members of G hear the report, stop their timers
- routers hear all reports, and time out non-responding groups



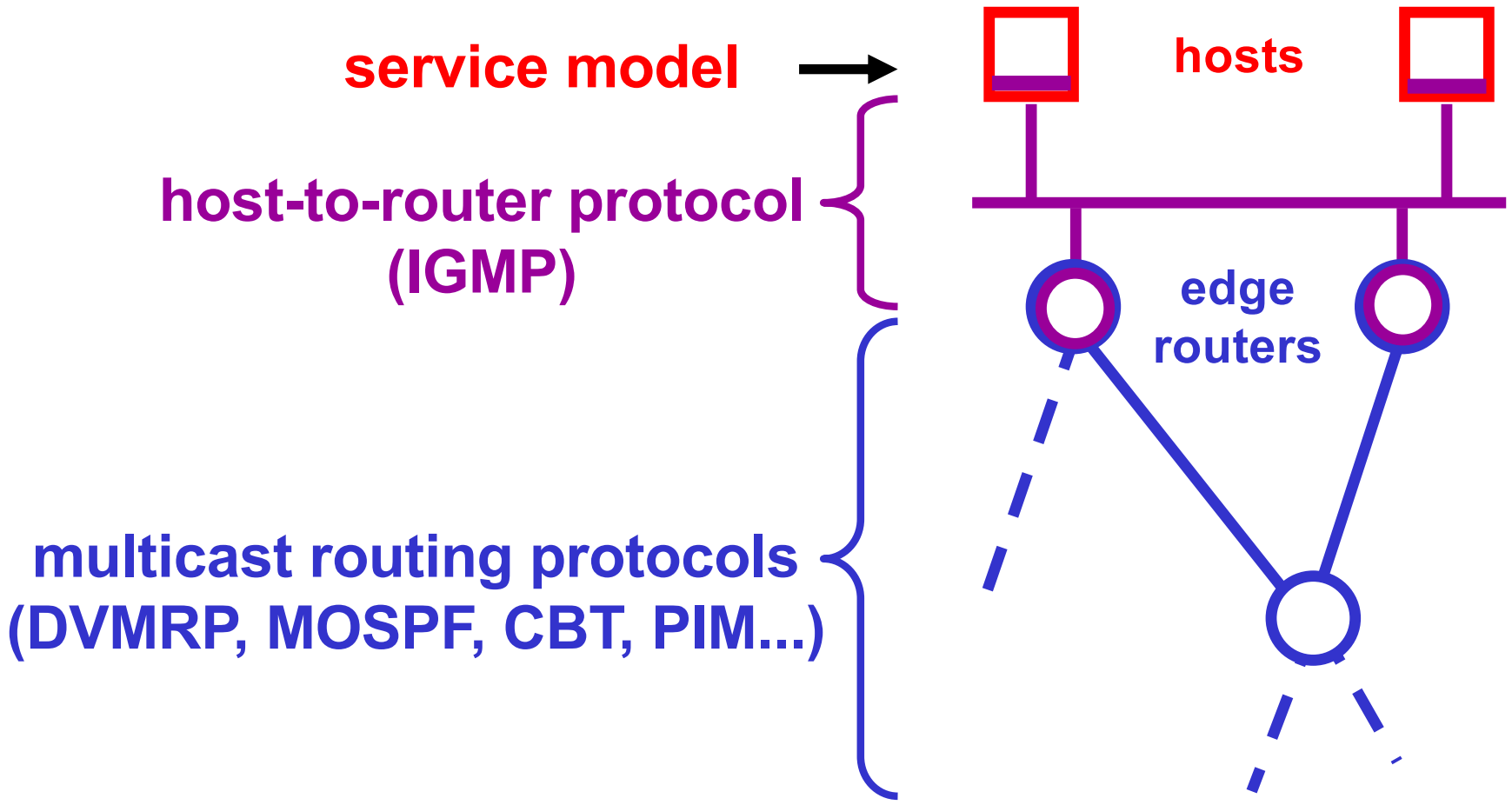
# How IGMP Works (cont.)

- In normal case, only one report message per group present is sent in response to a query
  - Query interval is typically 60-90 seconds
- When a host first joins a group, it sends one or two immediate reports instead of waiting for a query

# Leaving a multicast group

- when a host leaves a group: sends a Leave Group message if it was the most recent host to report membership in that group
  - to all-routers (224.0.0.2) address
- upon receiving Leave Group message, query router sends a couple of group-specific queries, specifying a small max-response-time
- if no report heard, assume the group is no longer present in the subnet.

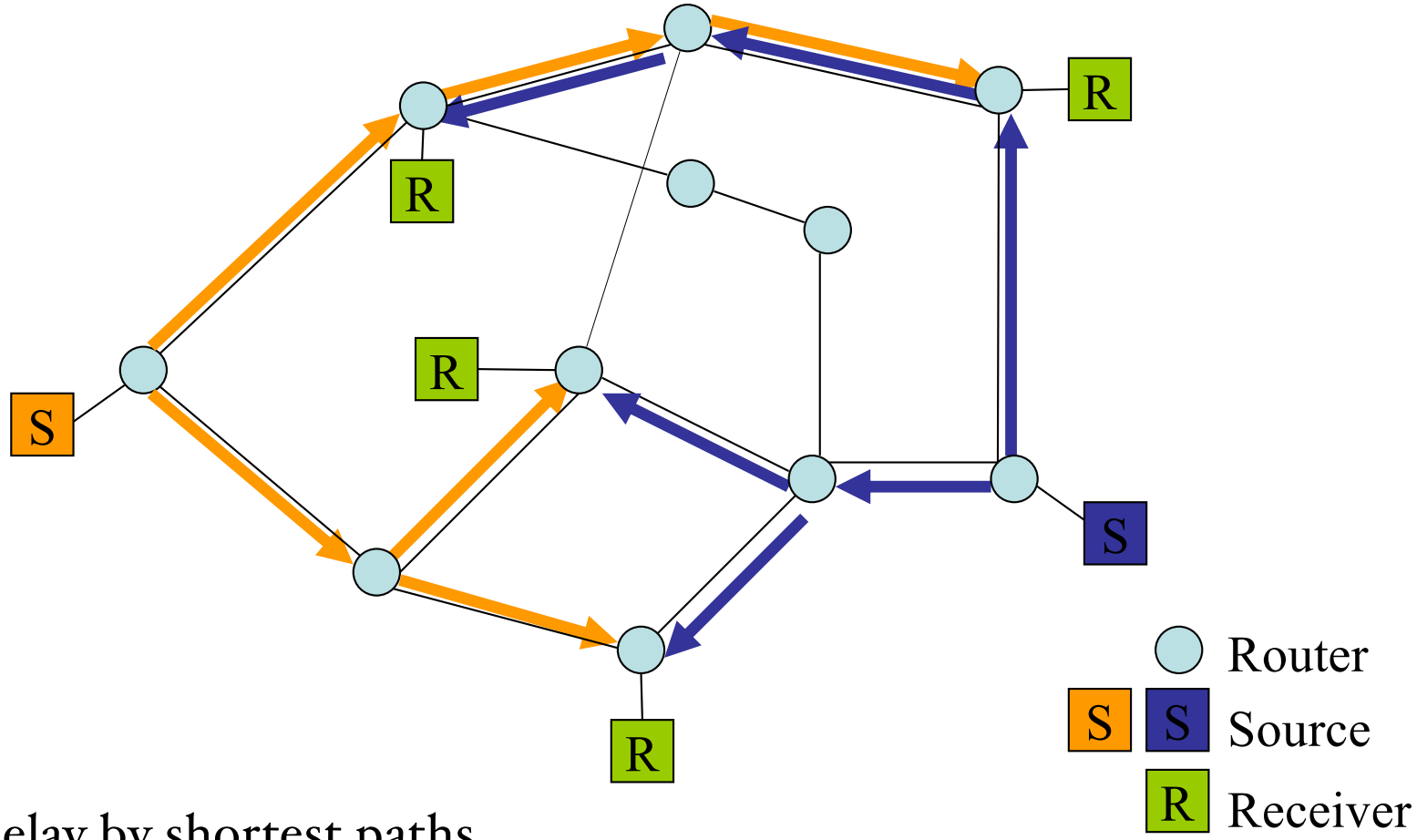
# IP Multicast Architecture



# Multicast Routing

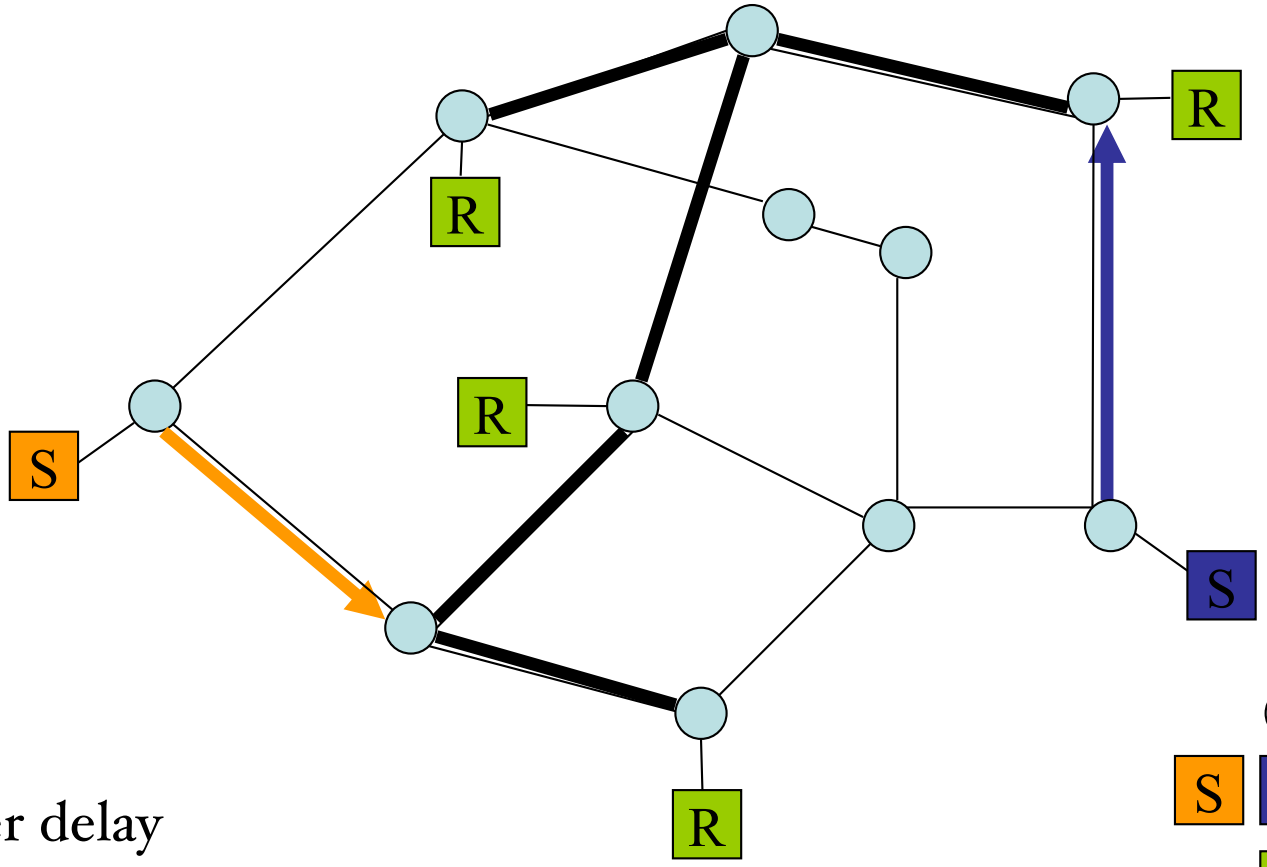
- Objective – build a distribution tree to reach all subnets that have group members.
  - The "leaves" of the distribution tree are the subnets containing at least one group member (reported by IGMP)
  - The "root" of the tree is the sender.
  - As a comparison, unicast routing is to build a path to reach the destination subnet.
- Two issues:
  - One tree per data source, or one tree per group?
  - How do senders and receivers find each other?

# Source Trees




- Low delay by shortest paths
- distribute traffic load on more links
- More routing states #sources * #groups

# Shared Tree




- Longer delay
- Using less number of links
- But much less routing states #groups

Router

S Source

R Receiver

# Distance Vector Multicast Routing Protocol

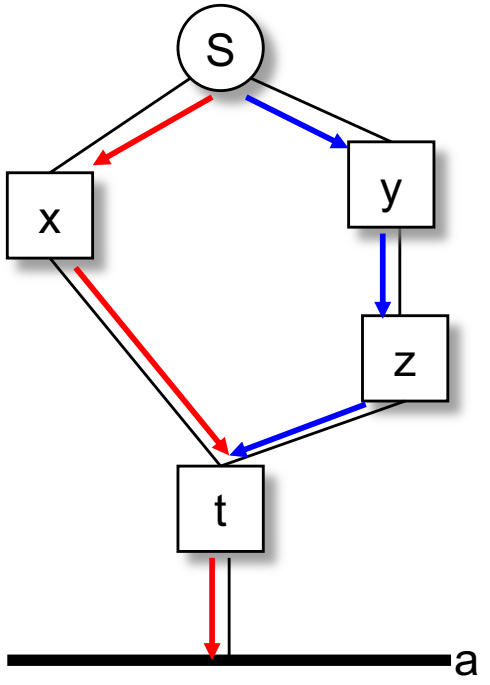
- Two major components:
  - a conventional distance-vector routing protocol (like RIP) which builds a unicast routing table
  - a protocol for determining how to forward multicast packets, based on the distance vector routing table.
- Basic idea: flood and prune
  - Initially flood the network with the multicast packets, then prune branches that have no group members.
  - How to flood without looping?

# Reverse Path Checking

- A router forwards a packet from source (S) **iff** it arrives via the shortest path from the router back to S; otherwise drop the packet

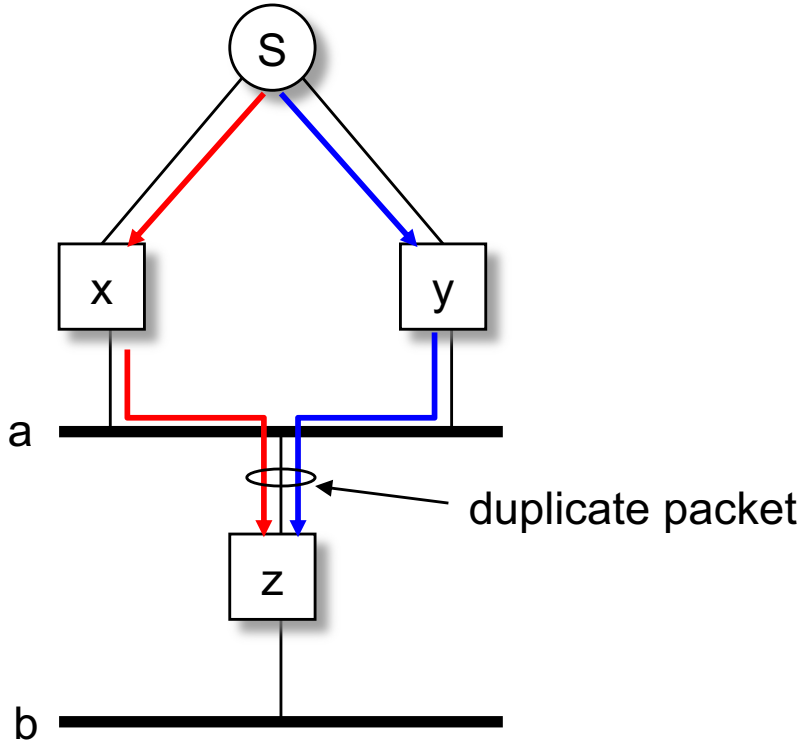
- Reverse shortest paths are obtained from unicast routing tables

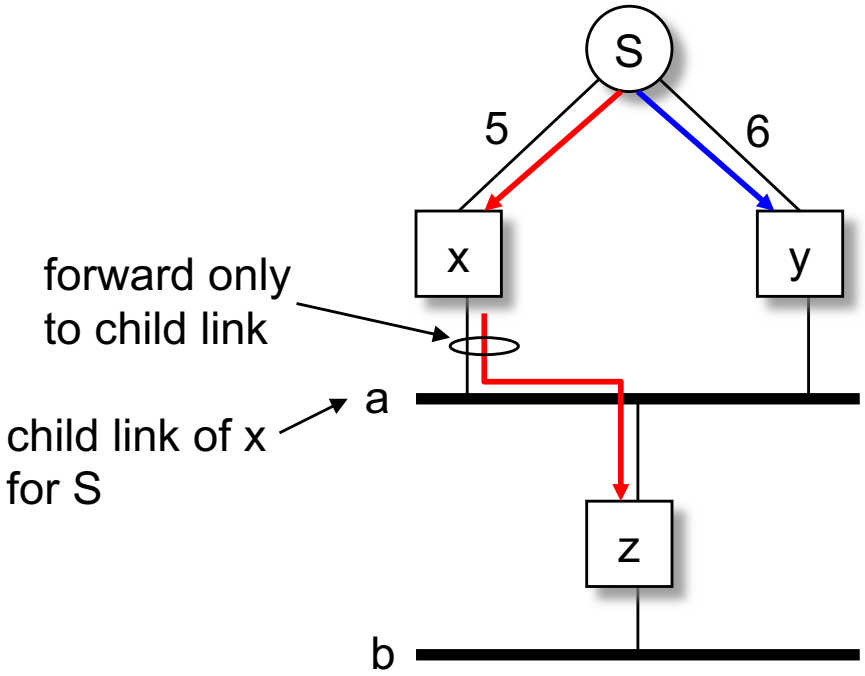
Why cannot use sequence number like in OSPF?



# Reverse Path Broadcasting

- Forward a packet from S only on child links
- Child link of router x for source S: iff z uses x as the next hop in its shortest path to reach S.

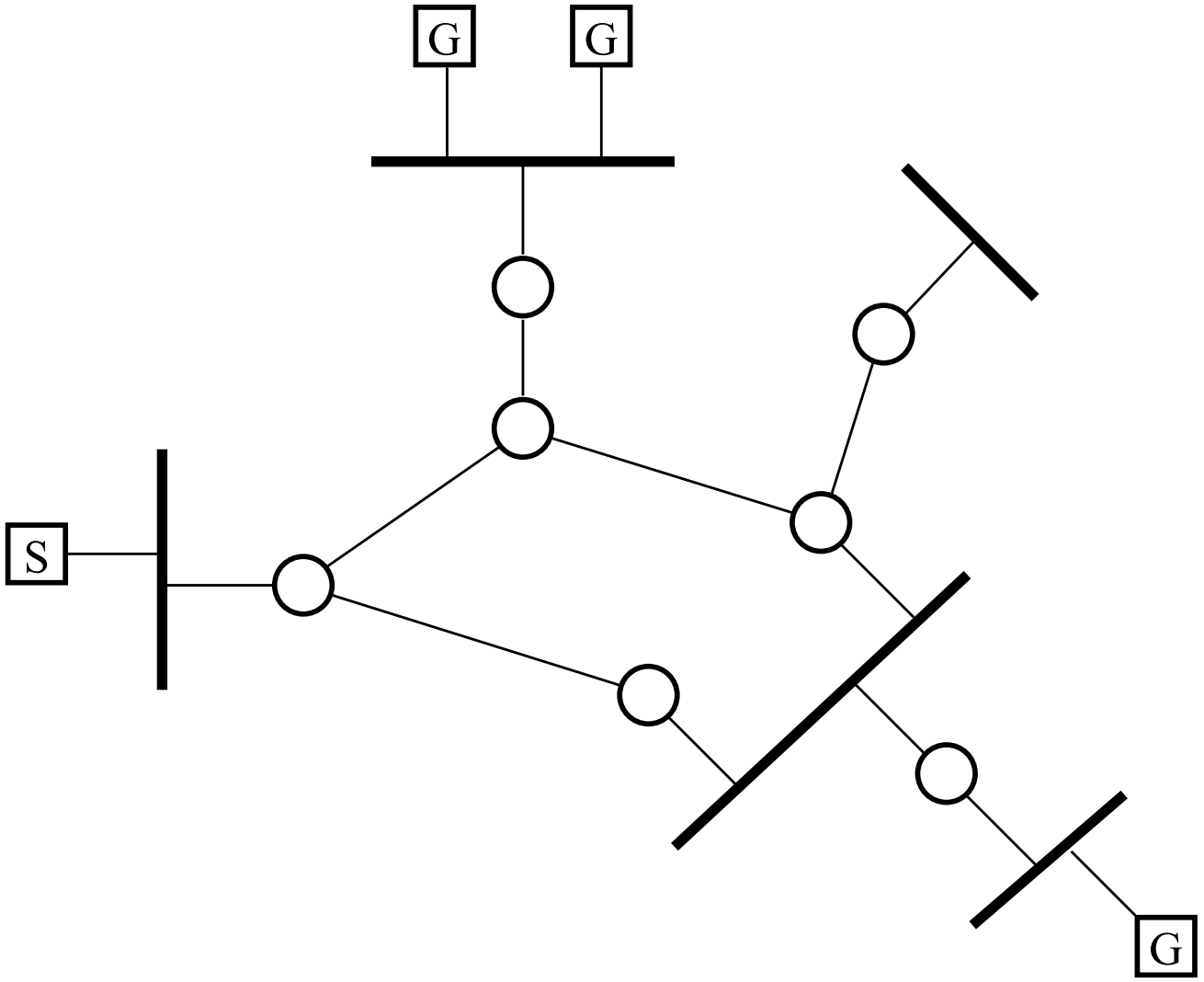


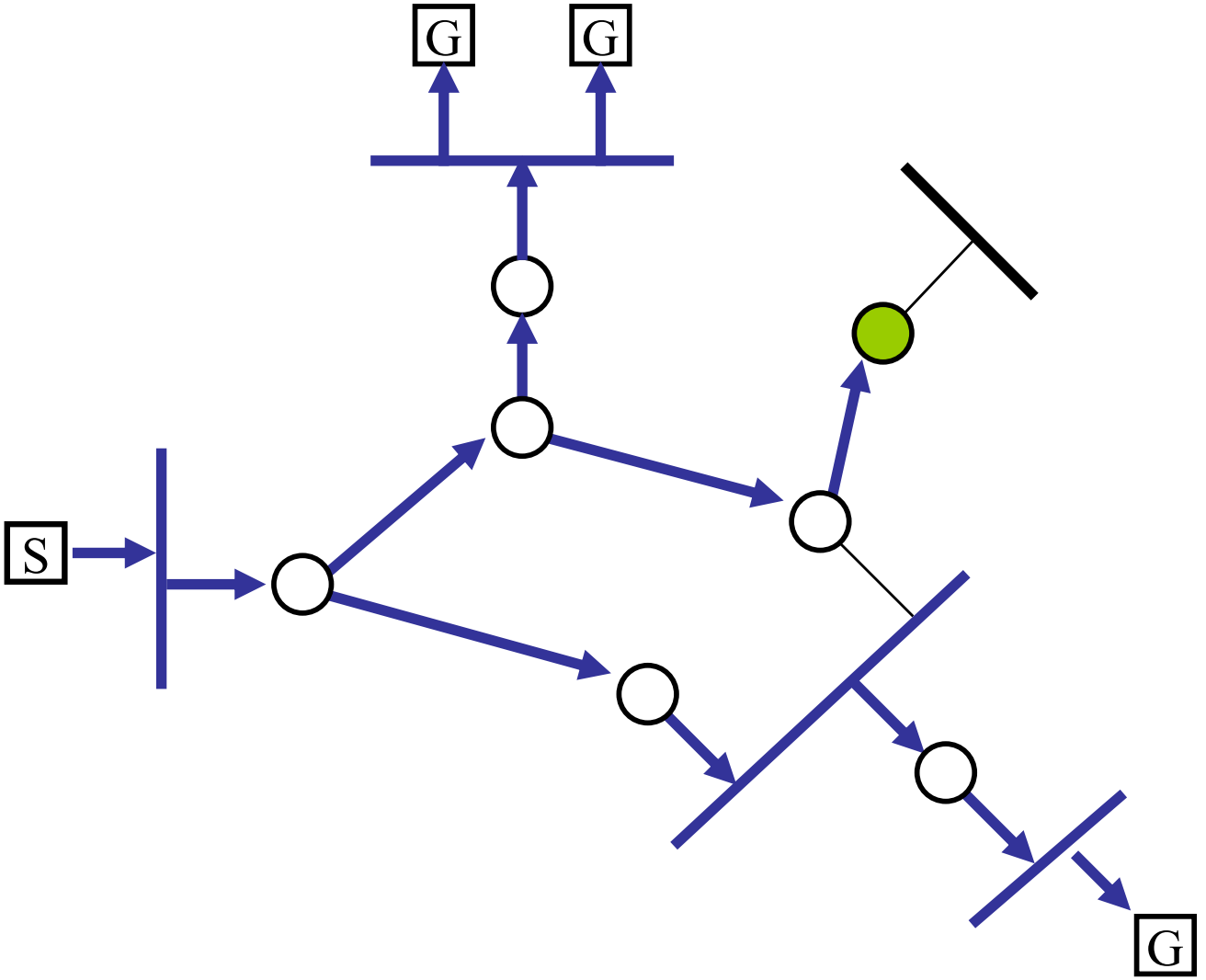
# Reverse Path Multicast

- Truncated RFB: Don't forward traffic onto networks with no receivers
- Prune (Source,Group) at leaf if no members in LAN
  - Send Non-Membership Report (NMR) up the tree
- If all children of router R prune (S,G)
  - Propagate prune message for (S,G) to the parent of R
- On timeout:
  - Prune state dropped
  - Flow is reinstated to allow flooding packets
  - Downstream routers re-prune
  - Note: this is a soft-state approach
- Grafting: Explicitly reinstate sub-tree when
  - IGMP detects new members at leaf, or when a child asks for a graft.

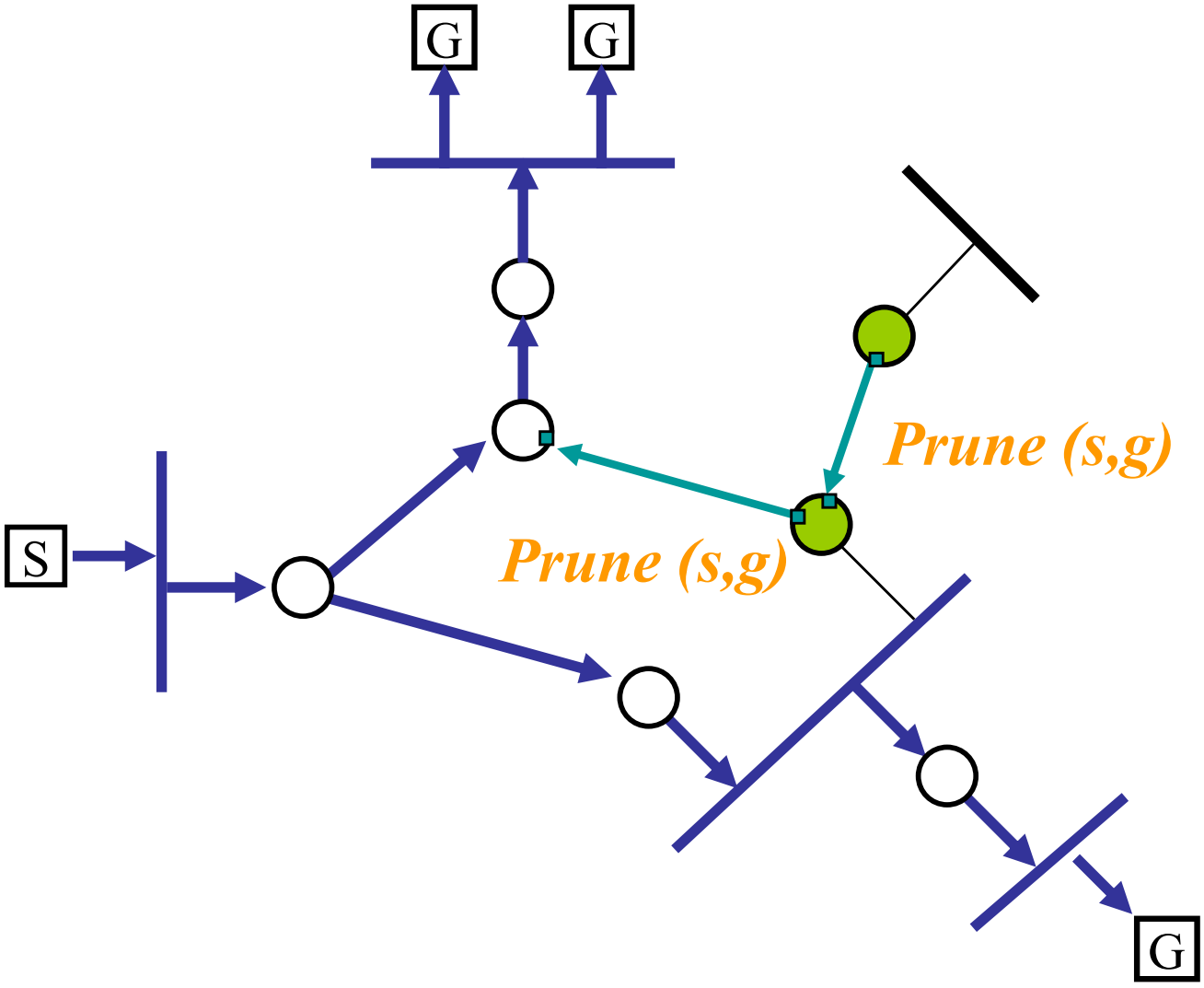
# Putting it together: Topology



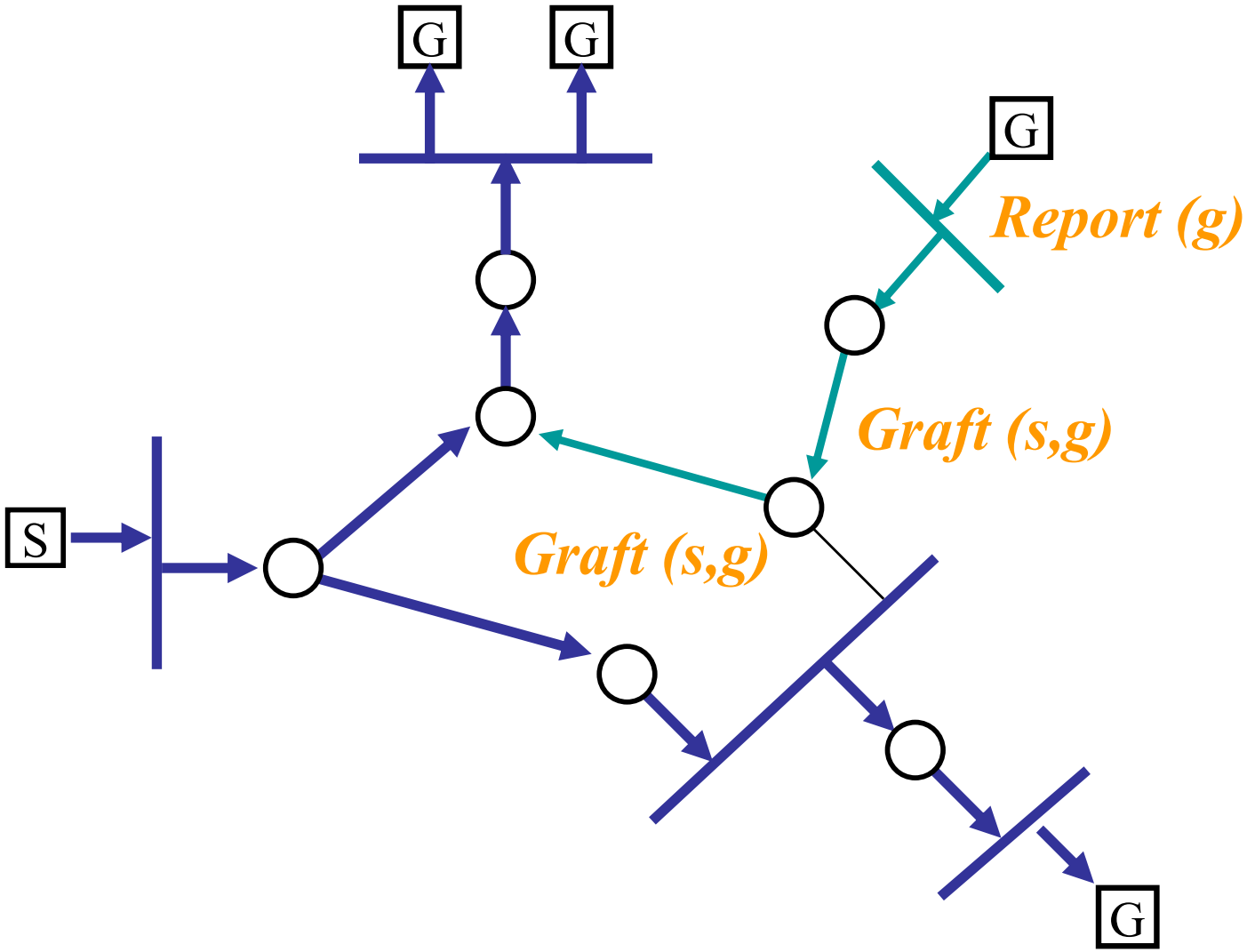
# Flooding with Truncated RFB



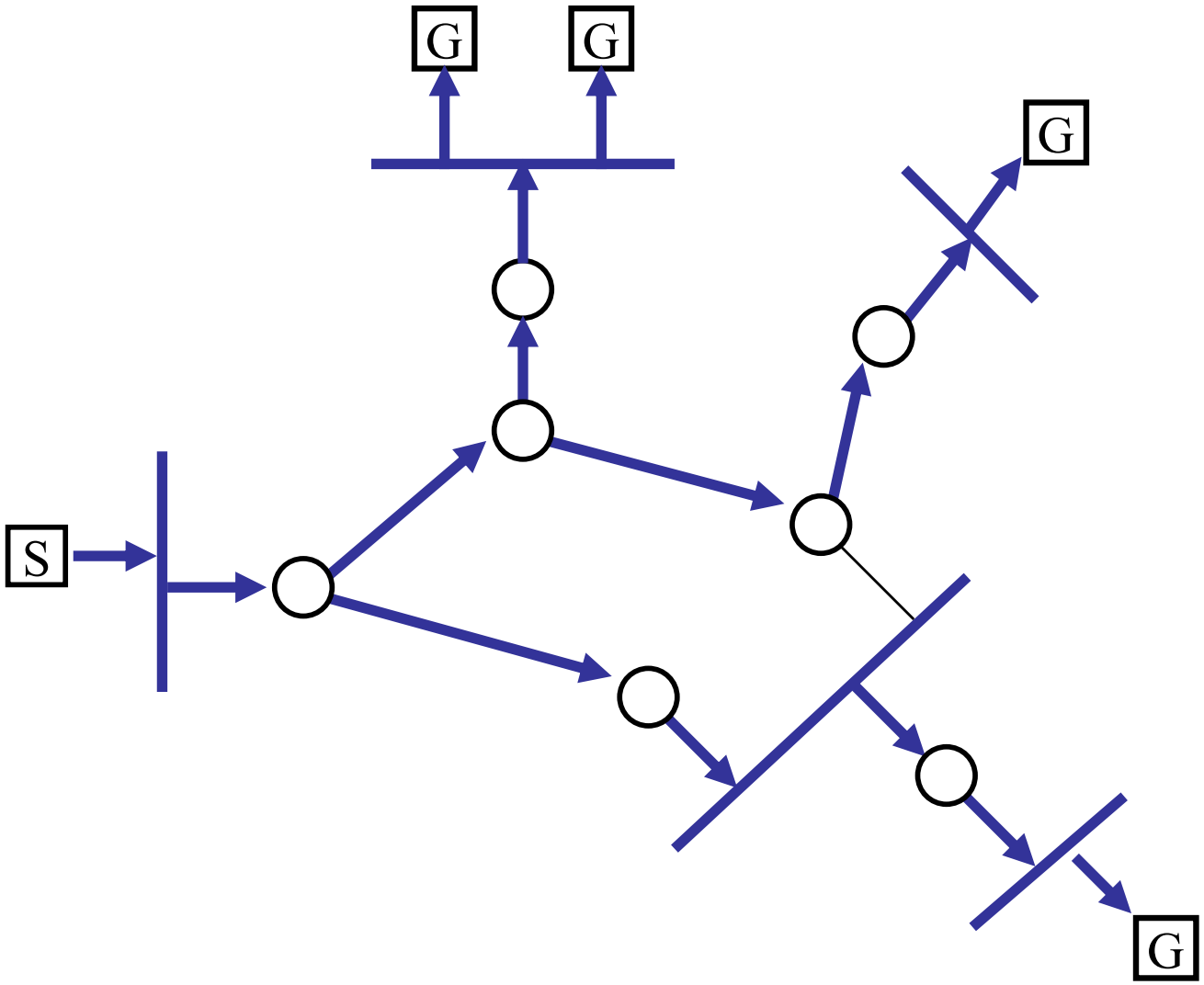
# Pruning



# Grafting



# After Grafting Complete



# Multicast OSPF (MOSPF)

- Extend OSPF to support multicast
- Link-state packets include multicast group addresses to which local members have joined
  - Thus a complete network topology annotated with group membership
- Routing algorithm computes shortest-path distribution tree from a source to all subnets that have members.
  - Every time a new source shows up, it triggers a new round of computation.

# DVMRP Evaluation

- Data-driven, simple and robust
- Failure mode
  - fully distributed design, no single point of failure
  - may deliver packets to unwanted places, but never fails to deliver to all interested receivers

- Transmission overhead of (periodic) initial broadcast
- State overhead
  - Every router maintains (S,G) states
  - Even in pruned parts

# MOSPF Evaluation

- Membership driven
  - get ready for data; idle trees?
- Failure mode
  - Fully distributed; Data delivery reliability relies on membership info distribution reliability

- Computation overhead, cannot pre-compute multicast trees
- (S,G) states