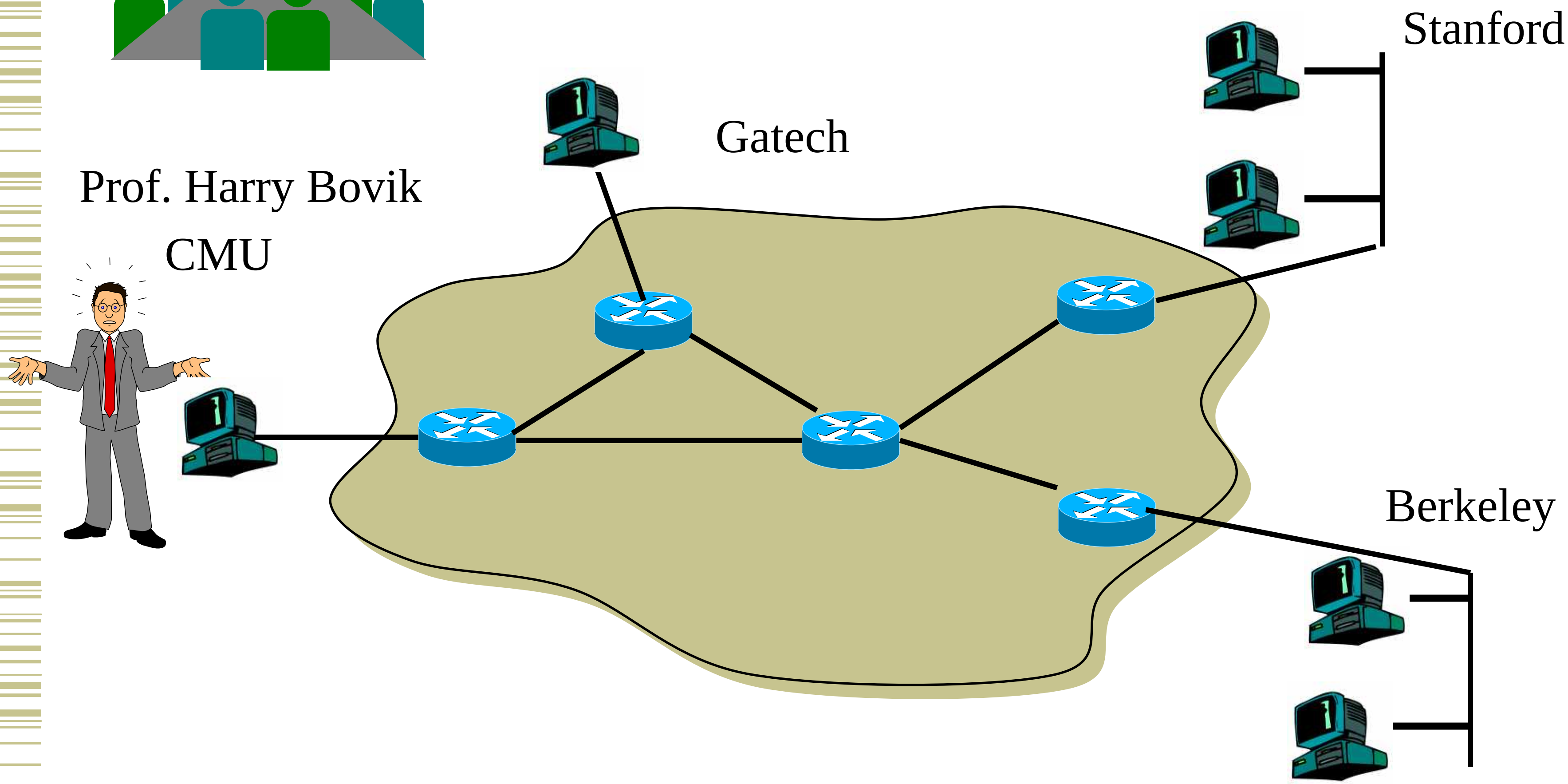
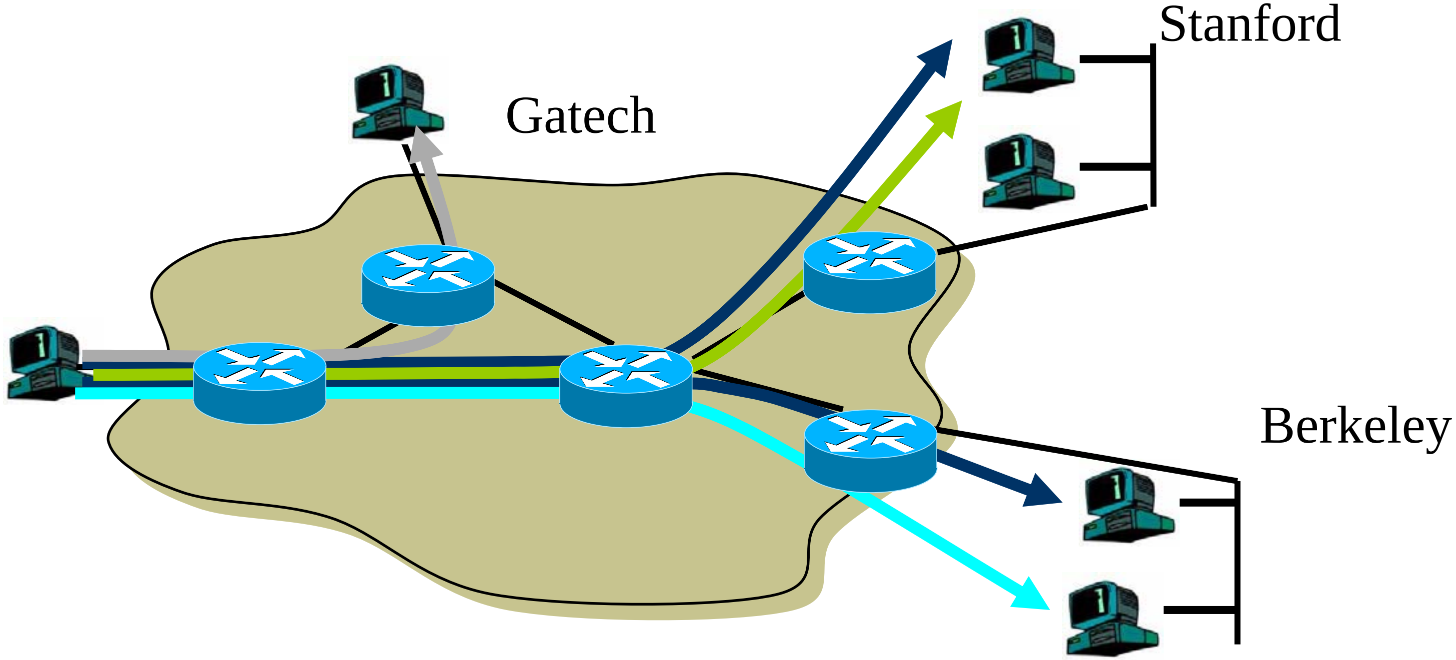
15-441 Computer Networking

Lecture 11 – Multicast

A Virtual Classroom



A Virtual Classroom

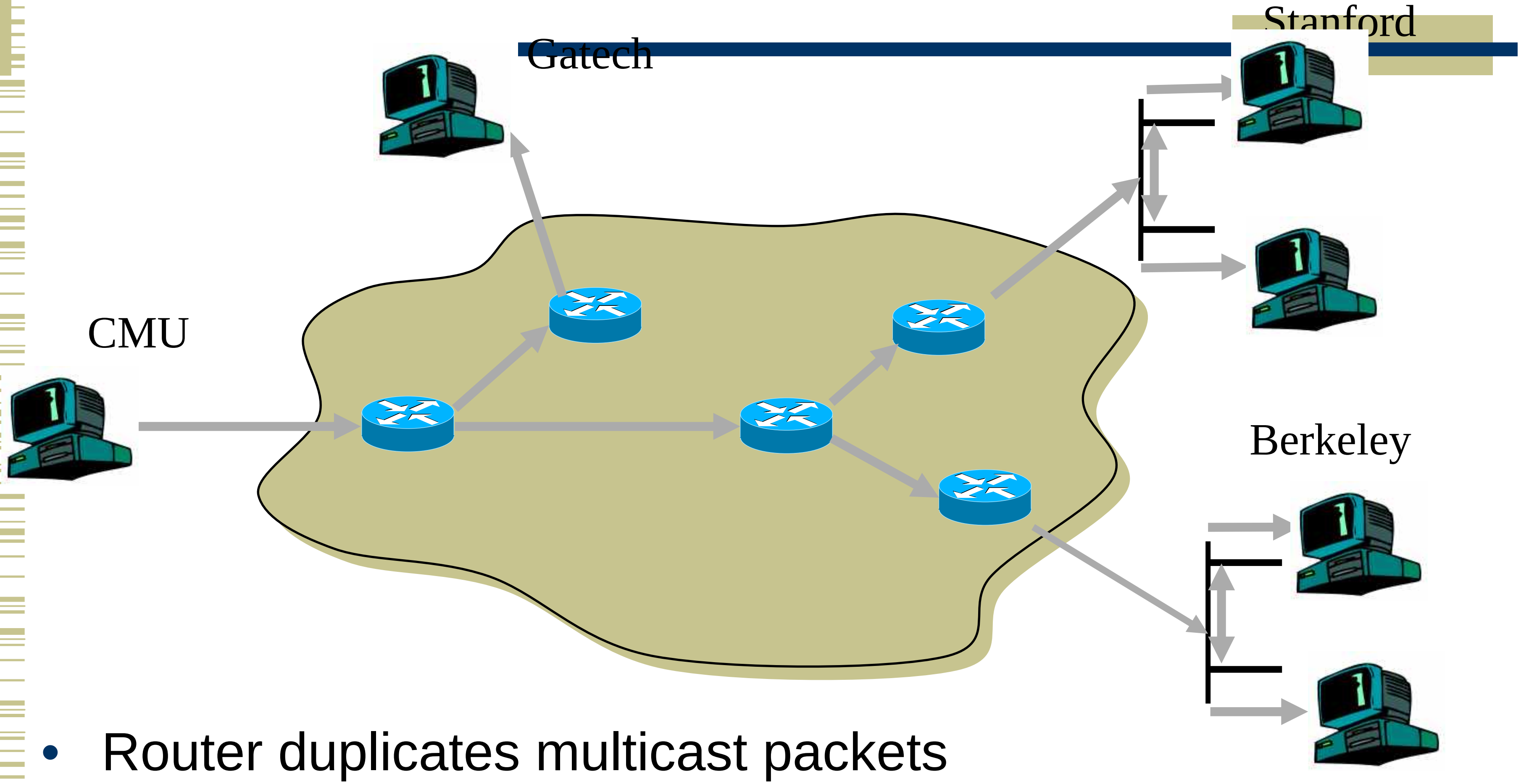


- Poor performance scalability
  - delay, throughput
  - sender, network

The emerging Internet

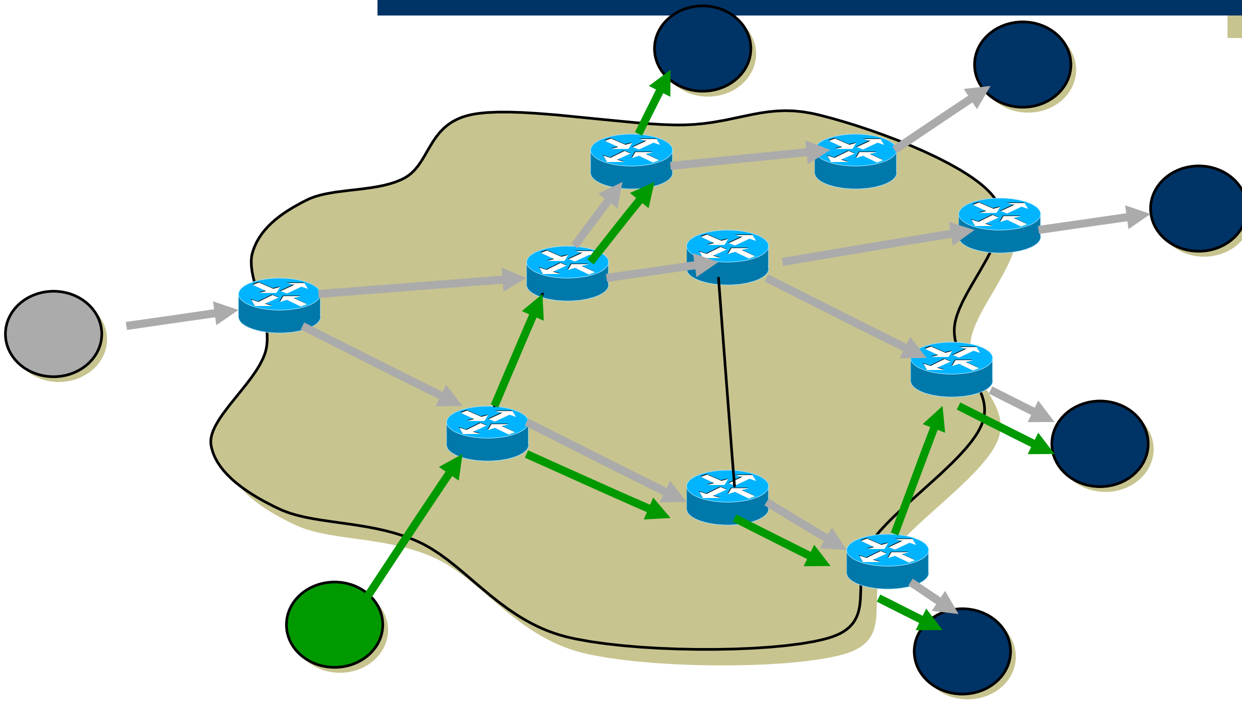
- A plethora of multi-party applications...
  - Audio/video conferencing
  - Multi-party games
  - Software distribution
  - Internet Television
- And now consider a world with ...
  - Millions of groups
  - Each group with tens to several thousand members

IP Multicast



- Router duplicates multicast packets
- One packet on each link
- Good performance scaling property

IP Multicast



- How to tell a packet is multicast?
- How to decide where to branch?

Standard Questions for Any New Network Functionalities

- What does the data plane look like?
  - What is format of the forwarding table entry?
  - What is the key to the lookup table?
- What does the control plane look like?
  - How is the forwarding table constructed?
- What is the service interface?

IP Multicast Addresses

- Class D IP addresses
  - 224.0.0.0 – 239.255.255.255

1 1 1 0	Group ID

- How to allocated these addresses?
  - Well-known multicast addresses, assigned by IANA
  - Transient multicast addresses, assigned and reclaimed dynamically, e.g., by “sdr” program

Multicast Router Data Plane

- Replicate packets on appropriate interfaces

Src IP Address, IP Multicast Address	List of outgoing interfaces

Address or Name?

- Single name/address maps to logically related set of destinations
  - Destination set = multicast group
- Key challenge: scalability
  - Single name/address independent of group growth or changes

IP Multicast Service Model (rfc1112)

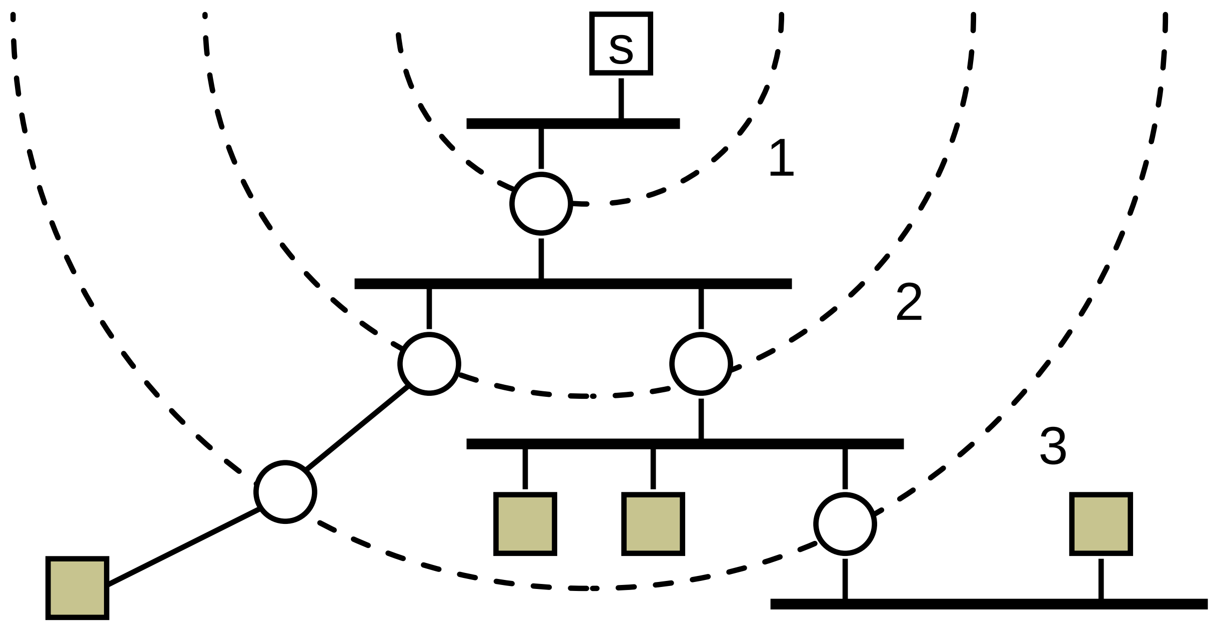
- Each group identified by a single IP address
- Groups may be of any size
- Members of groups may be located anywhere in the Internet
- Members of groups can join and leave at will
- Senders need not be members
- Group membership not known explicitly
- Analogy:
  - Each multicast address is like a radio frequency, on which anyone can transmit, and to which anyone can tune-in.

IP Multicast API

- Sending – same as before
- Receiving – two new operations
  - Join-IP-Multicast-Group(group-address, interface)
  - Leave-IP-Multicast-Group(group-address, interface)
  - Receive multicast packets for joined groups via normal IP-Receive operation
  - Implemented using socket options

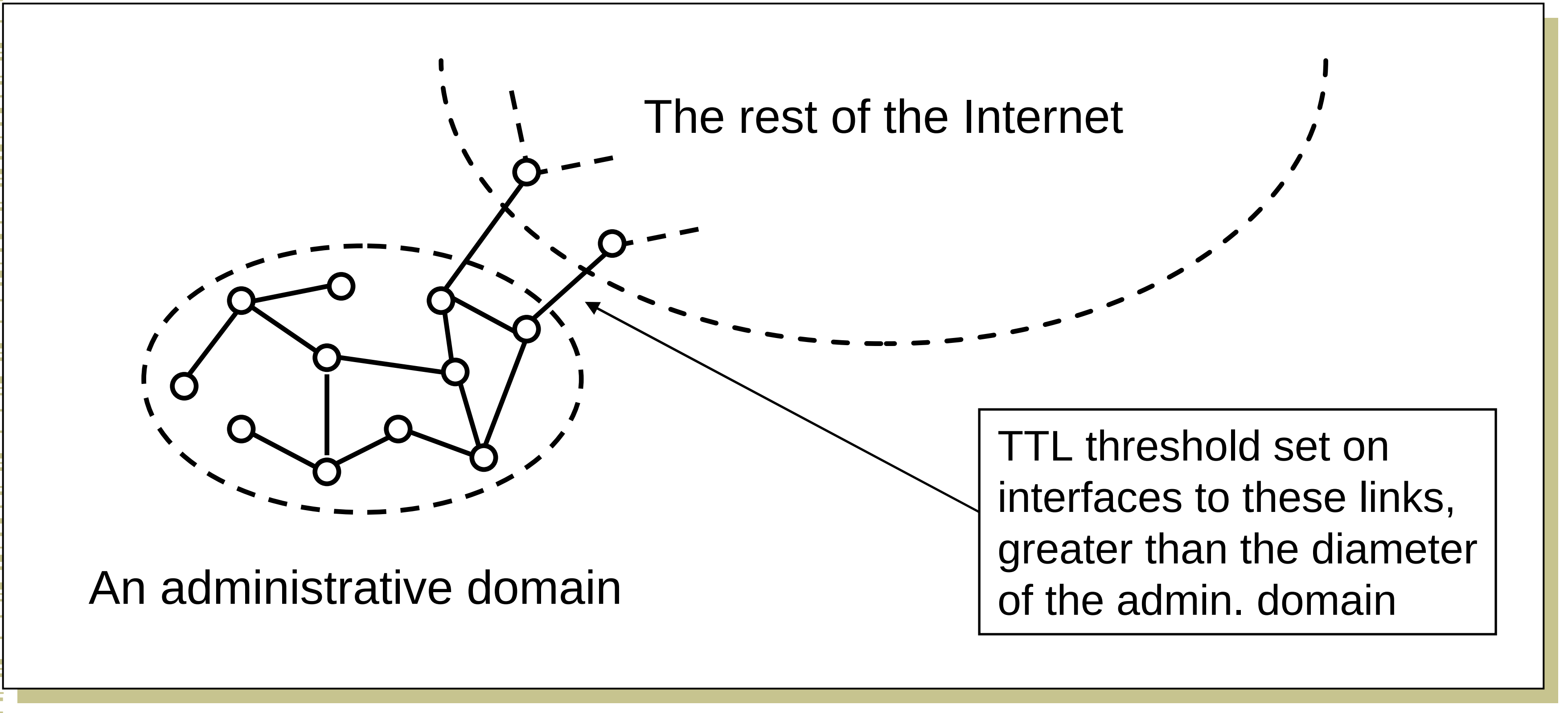
Multicast Scope Control – Small TTLs

- TTL expanding-ring search to reach or find a nearby subset of a group



Multicast Scope Control – Large TTLs

- Administrative TTL Boundaries to keep multicast traffic within an administrative domain, e.g., for privacy or resource reasons

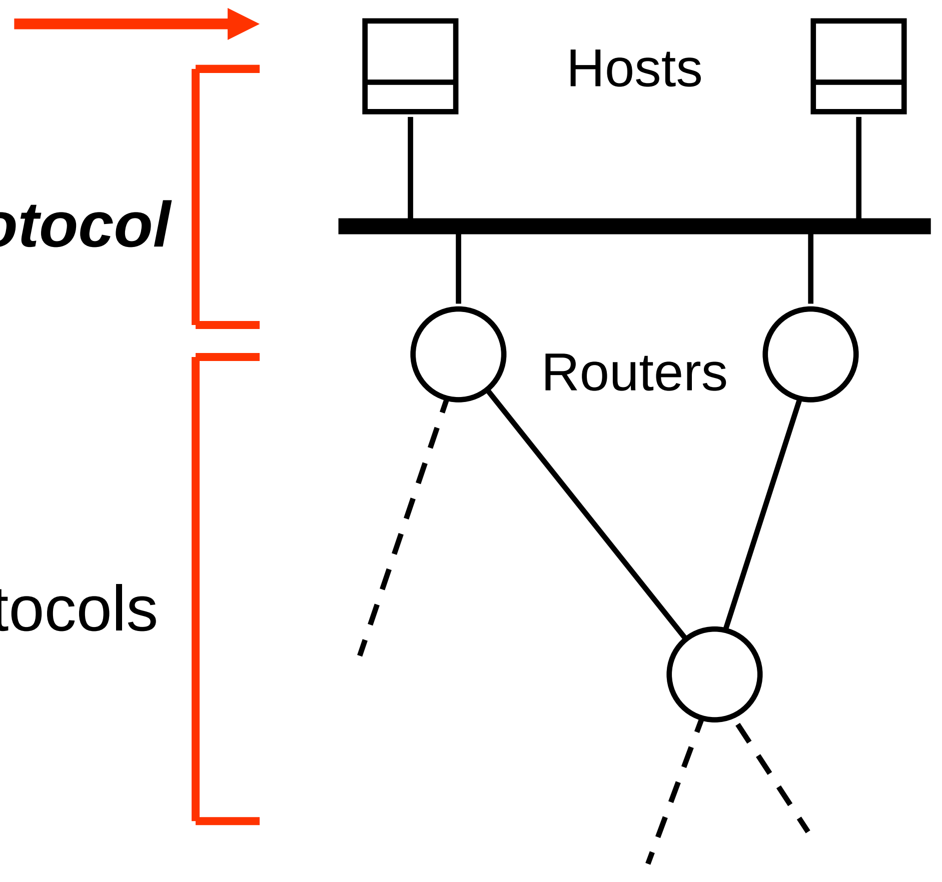


IP Multicast Control Plane

Service model

Host-to-router protocol (IGMP)

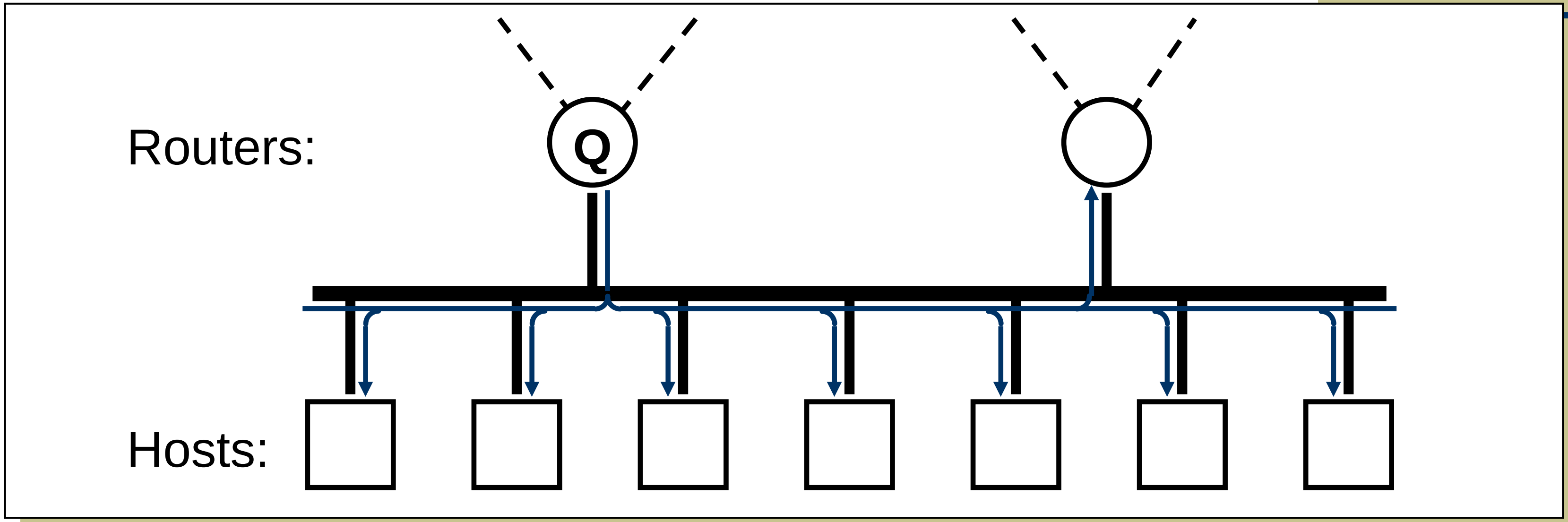
Multicast routing protocols (various)



Internet Group Management Protocol (Part 1 of Control Plane)

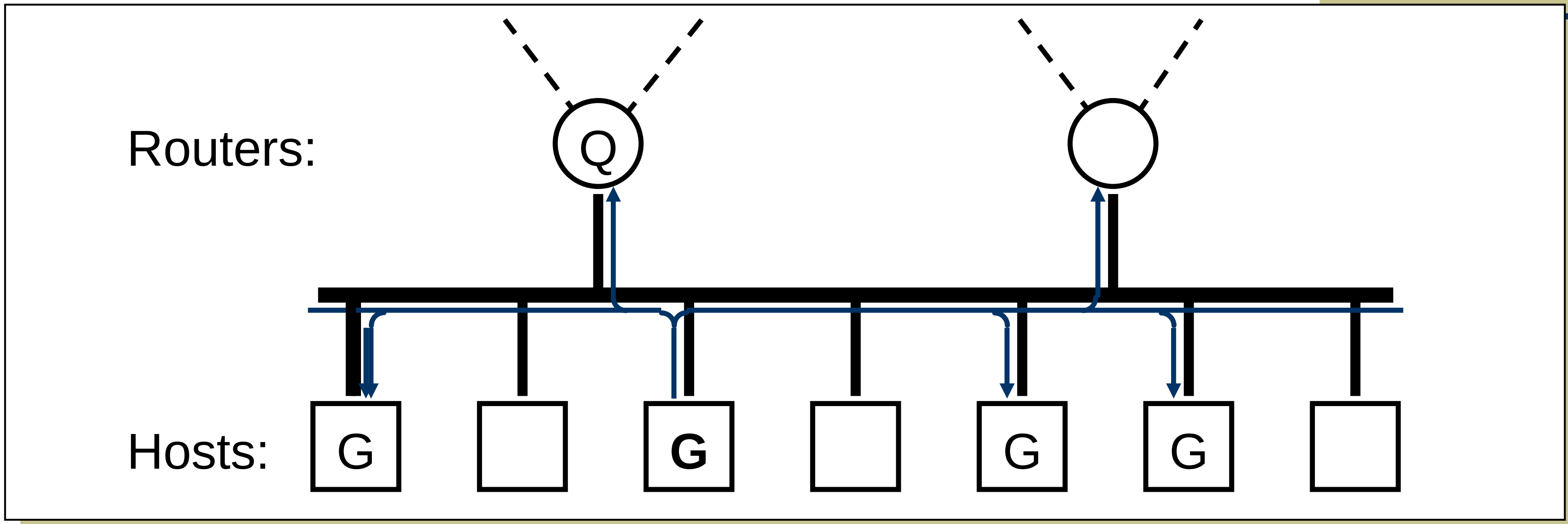
- End system to router protocol is IGMP
- Each host keeps track of which mcast groups are subscribed to
  - Socket API informs IGMP process of all joins
- Objective is to keep router up-to-date with group membership of entire LAN
  - Routers need not know who all the members are, only that members exist

How IGMP Works



- On each link, one router is elected the “querier”
- Querier periodically sends a Membership Query message to the all-systems group (224.0.0.1), with TTL = 1
- On receipt, hosts start random timers (between 0 and 10 seconds) 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 and stop their timers
- Routers hear all reports, and time out non-responding groups

How IGMP Works (cont.)

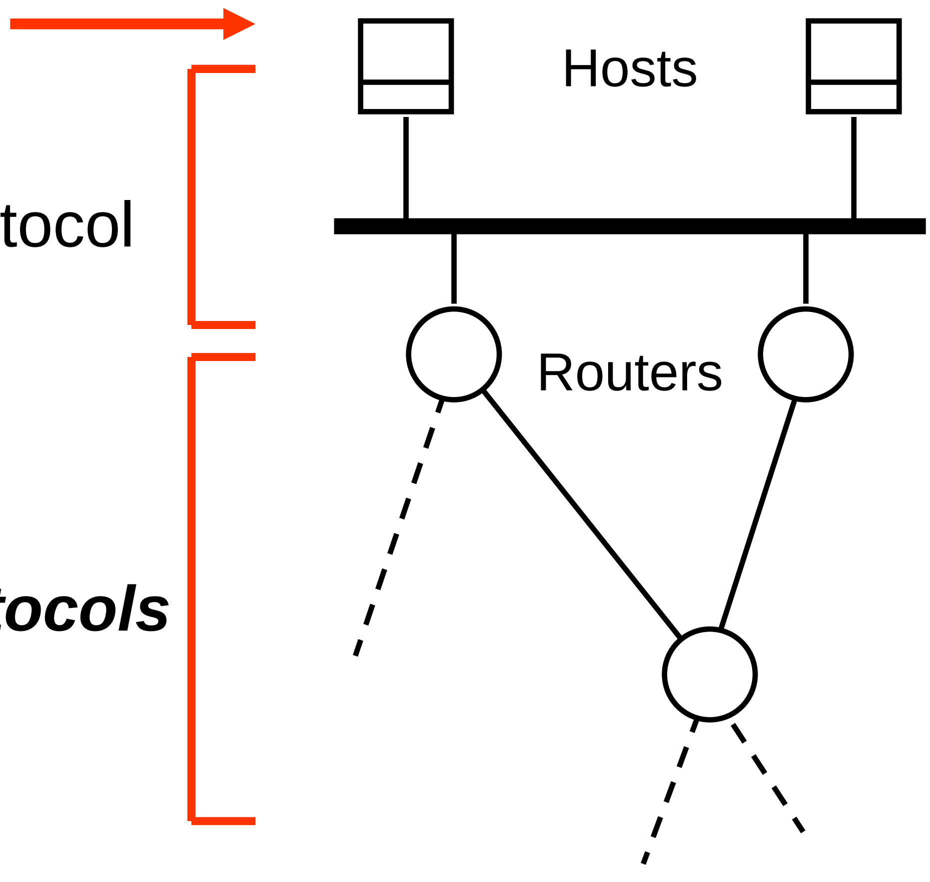
- Note that, in normal case, only one report message per group present is sent in response to a query
  - Power of randomization + suppression
- Query interval is typically 60-90 seconds
- When a host first joins a group, it sends one or two immediate reports, instead of waiting for a query

IP Multicast Control Plane

Service model

Host-to-router protocol (IGMP)

Multicast routing protocols (various)



Multicast Routing Protocols (Part 2 of Control Plane)

- Basic objective – build distribution tree for multicast packets
- Flood and prune
  - Begin by flooding traffic to entire network
  - Prune branches with no receivers
  - Examples: DVMRP, PIM-DM
  - *Unwanted state where there are no receivers*
- Link-state multicast protocols
  - Routers advertise groups for which they have receivers to entire network
  - Compute trees on demand
  - Example: MOSPF
  - *Unwanted state where there are no senders*

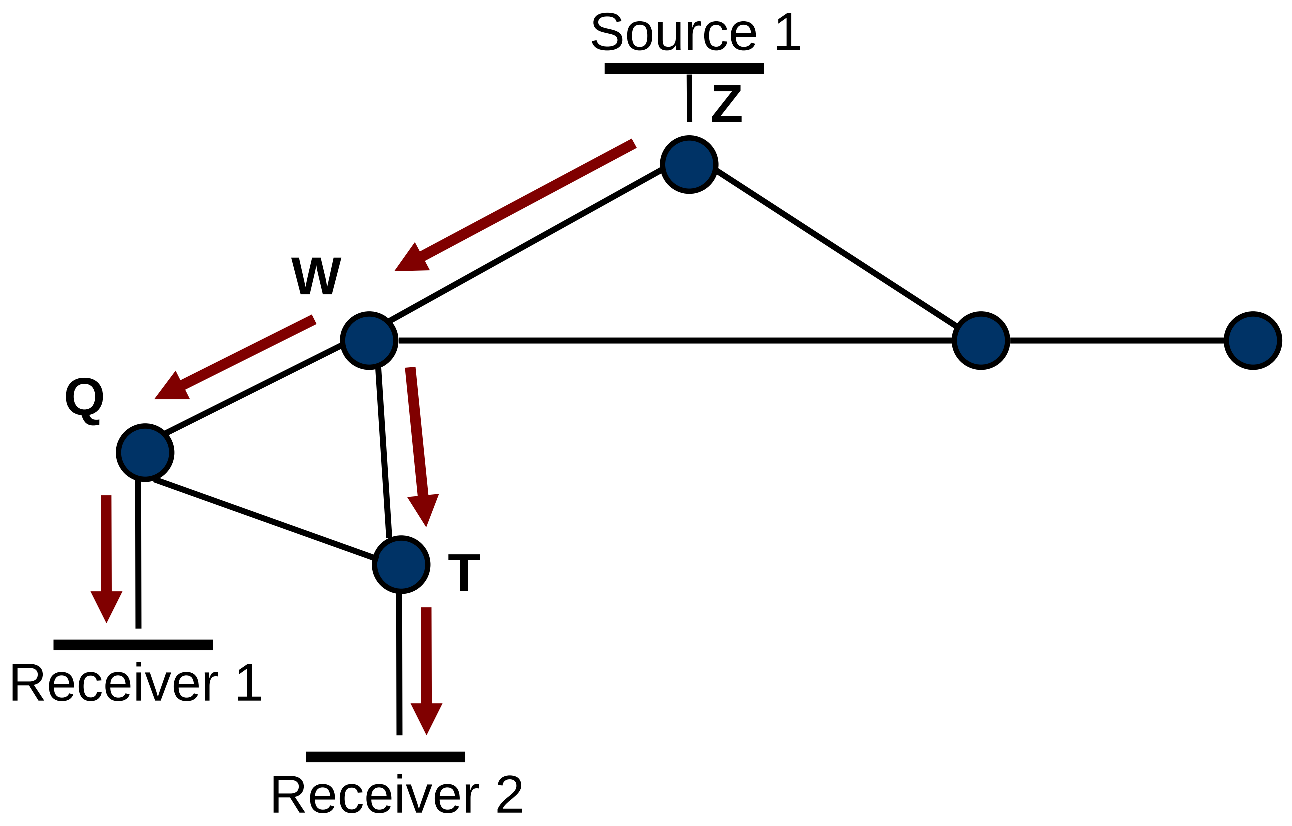
Multicast OSPF (MOSPF)

- Add-on to OSPF (Open Shortest-Path First, a link-state, intra-domain routing protocol)
- Multicast-capable routers flag link state routing advertisements
- Link-state packets include multicast group addresses to which local members have joined
- Routing algorithm augmented to compute shortest-path distribution tree from a source to any set of destinations

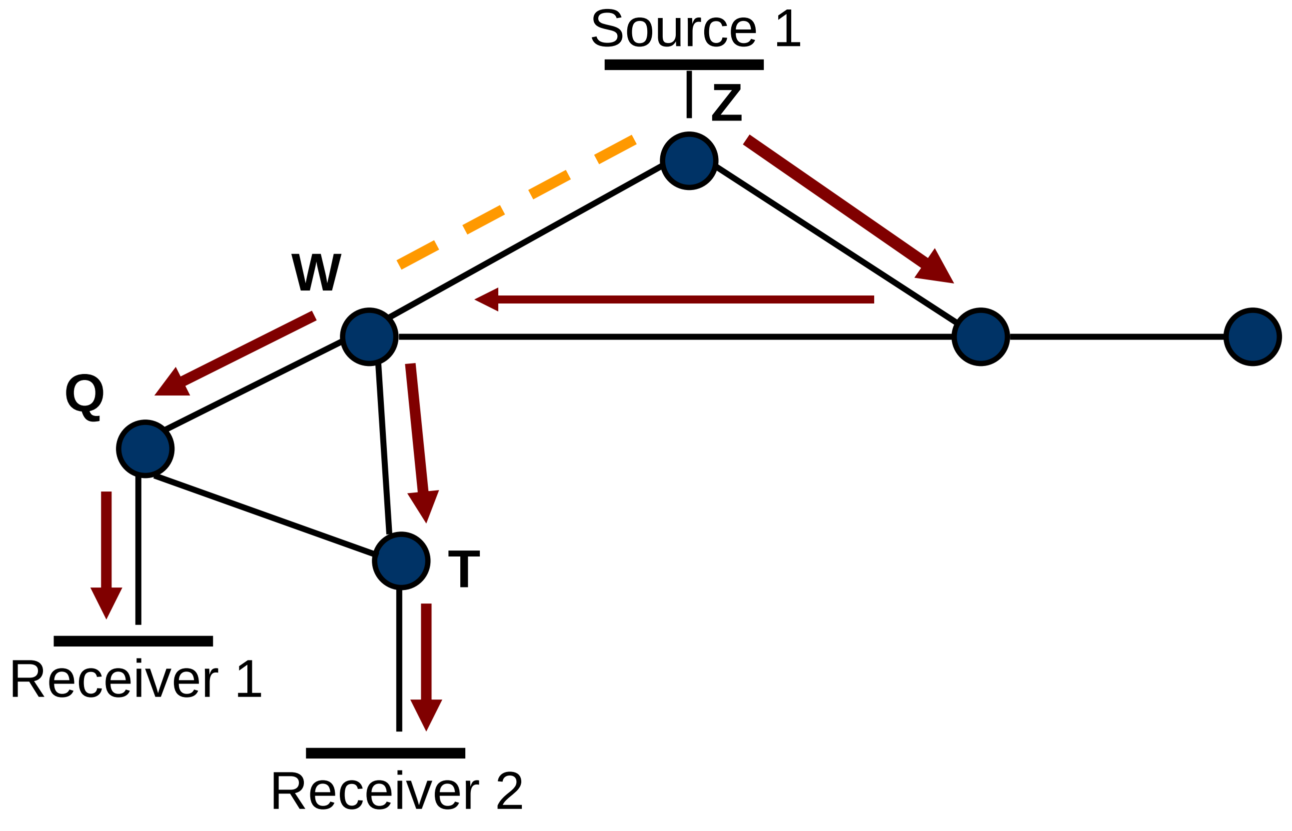
Impact on Route Computation

- Can't pre-compute multicast trees for all possible sources
- Compute on demand when first packet from a source S to a group G arrives
- New link-state advertisement
  - May lead to addition or deletion of outgoing interfaces if it contains different group addresses
  - May lead to re-computation of entire tree if links are changed

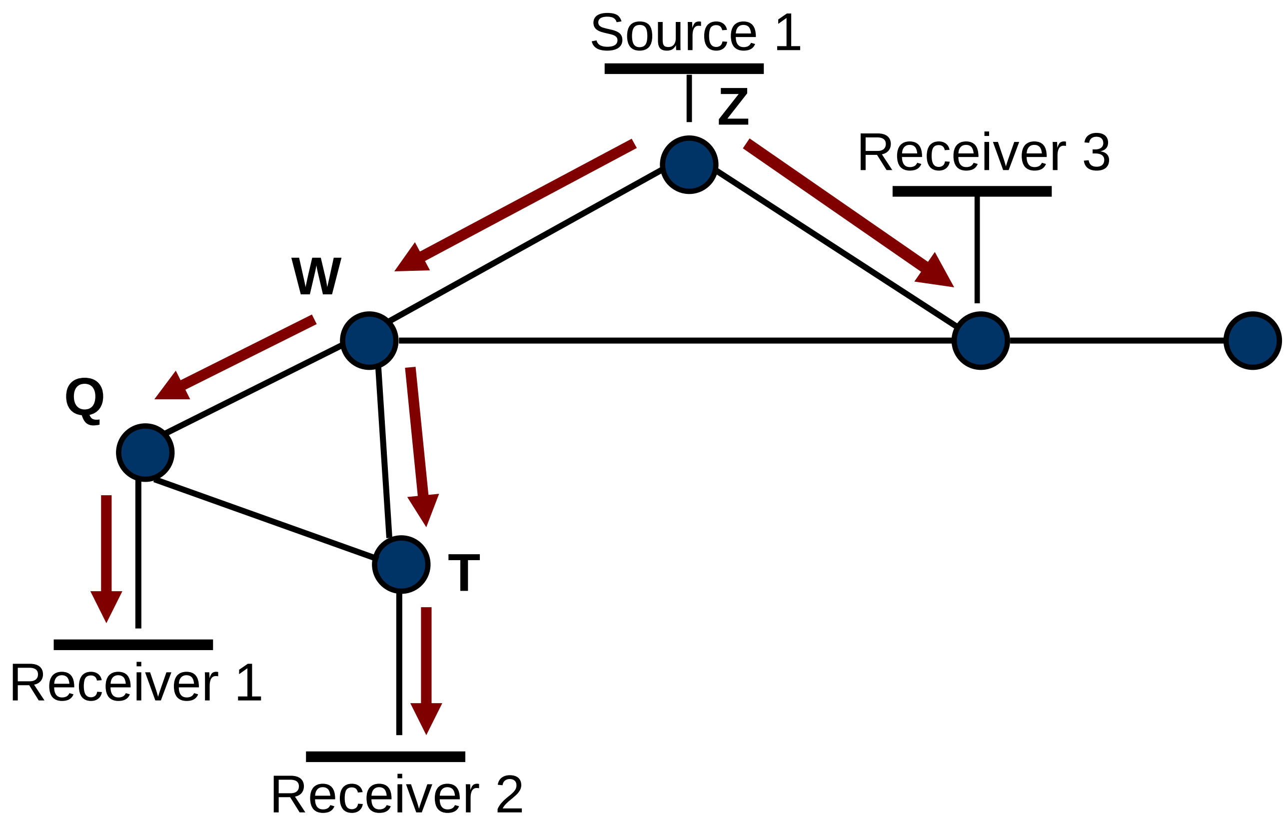
Example



Link Failure/Topology Change



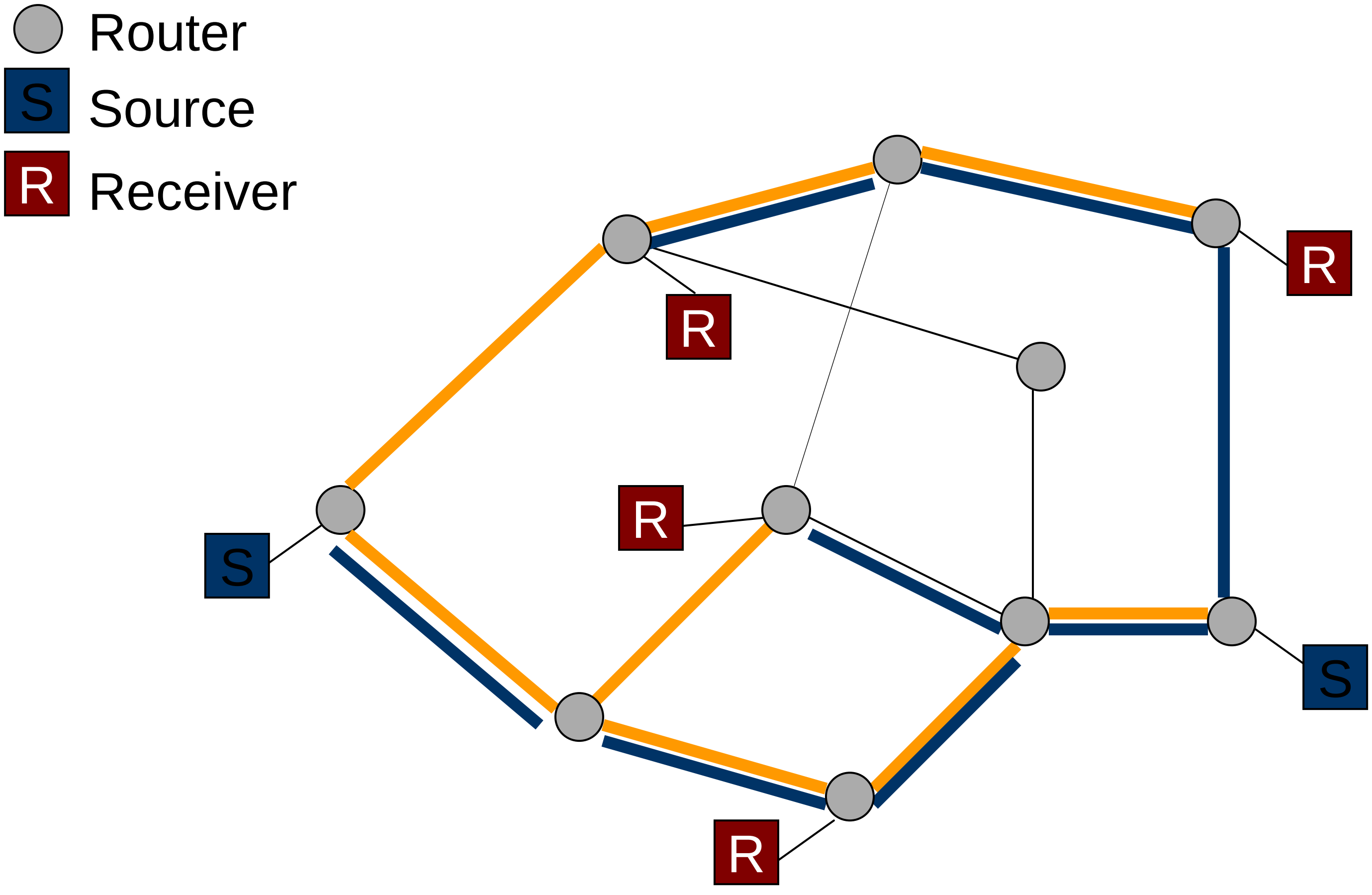
Membership Change



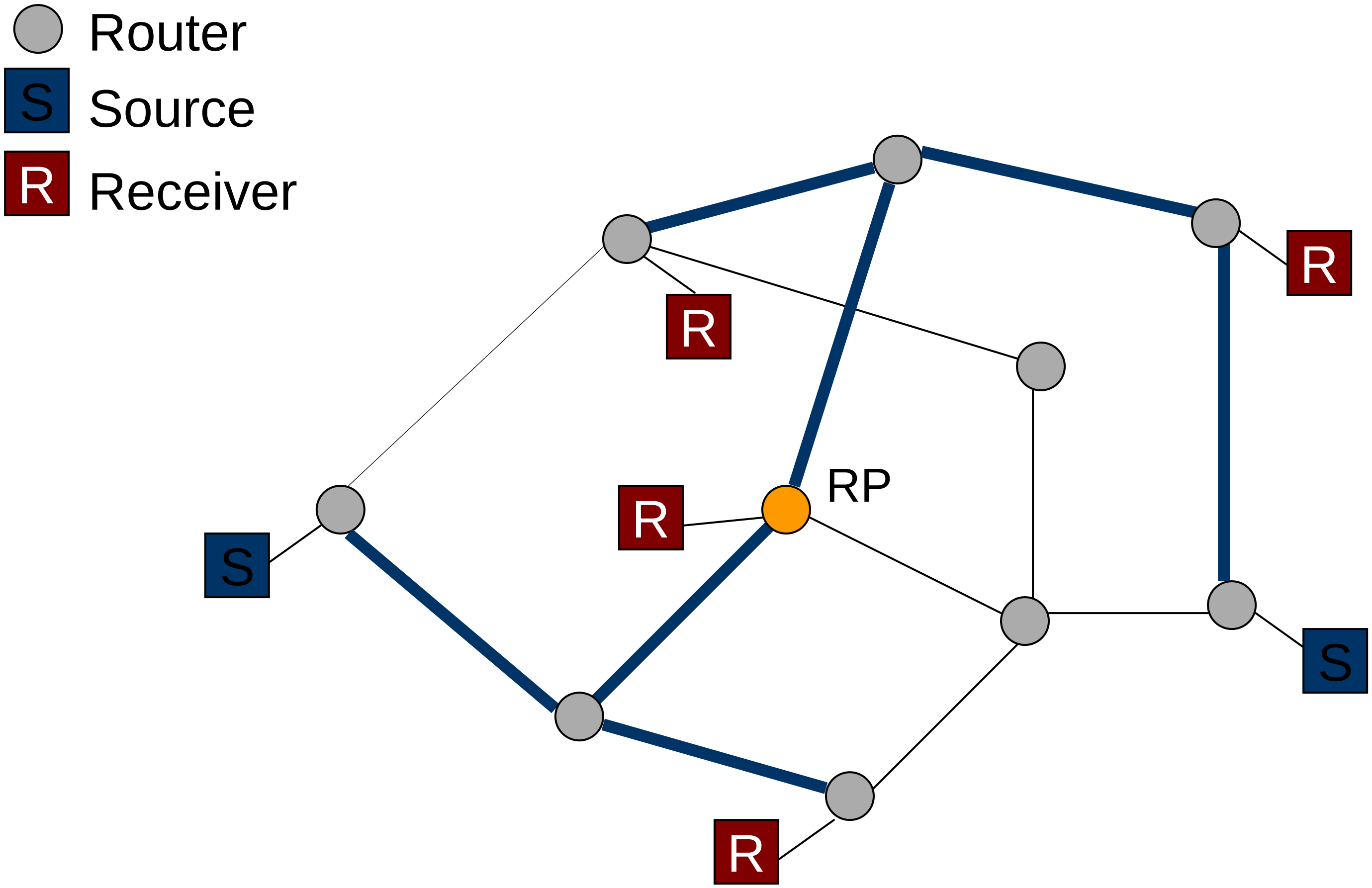
Shared vs. Source-based Trees

- Source-based trees
  - Separate shortest path tree for each sender
  - DVMRP, MOSPF, PIM-DM, PIM-SM
- Shared trees
  - Single tree shared by all members
  - Data flows on same tree regardless of sender
  - CBT, PIM-SM

Source-based Trees



Shared Tree



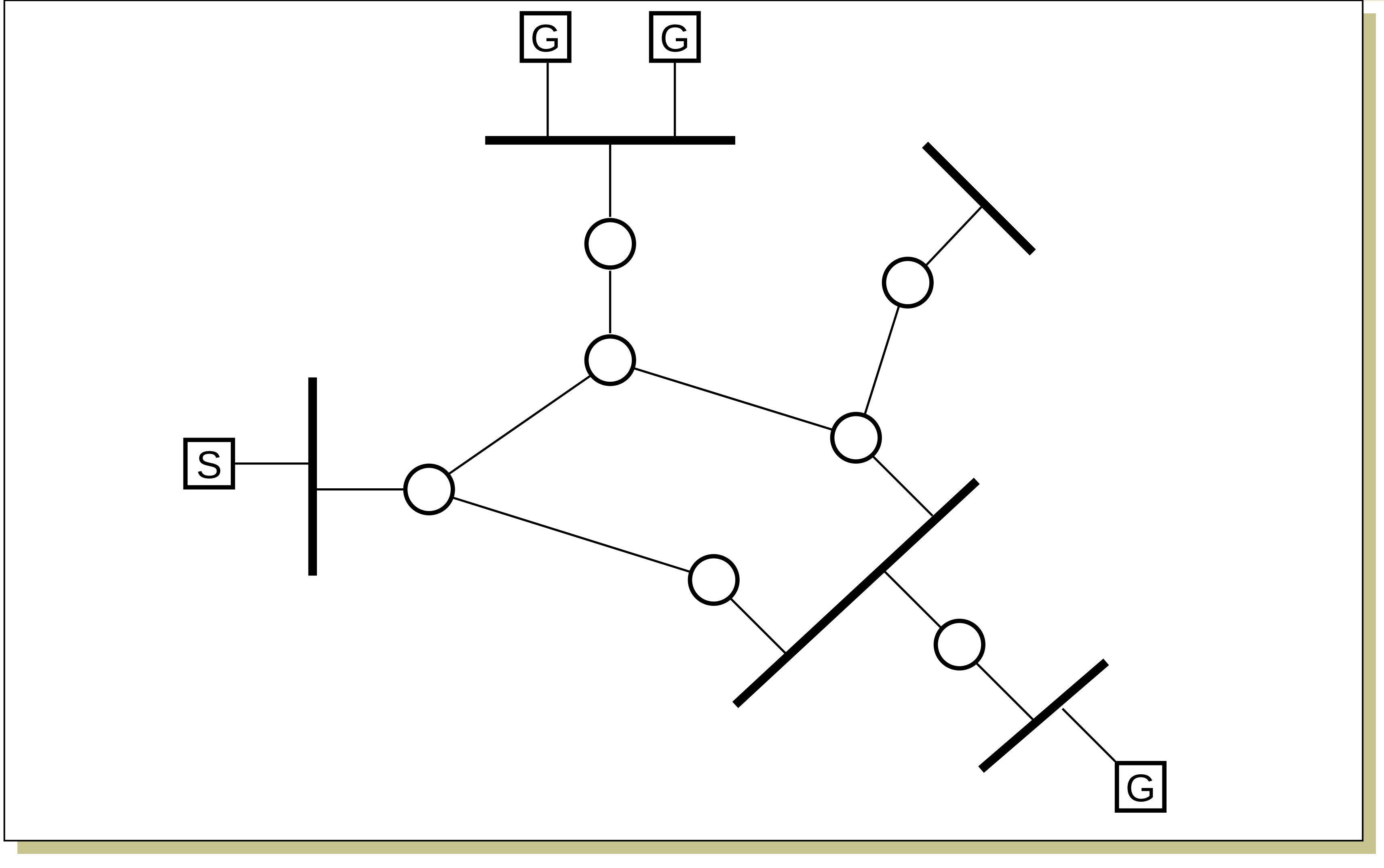
Shared vs. Source-Based Trees

- Source-based trees
  - Shortest path trees – low delay, better load distribution
  - More state at routers (per-source state)
  - Efficient for in dense-area multicast
- Shared trees
  - Higher delay (bounded by factor of 2), traffic concentration
  - Choice of core affects efficiency
  - Per-group state at routers
  - Efficient for sparse-area multicast
- Which is better? → extra state in routers is bad!

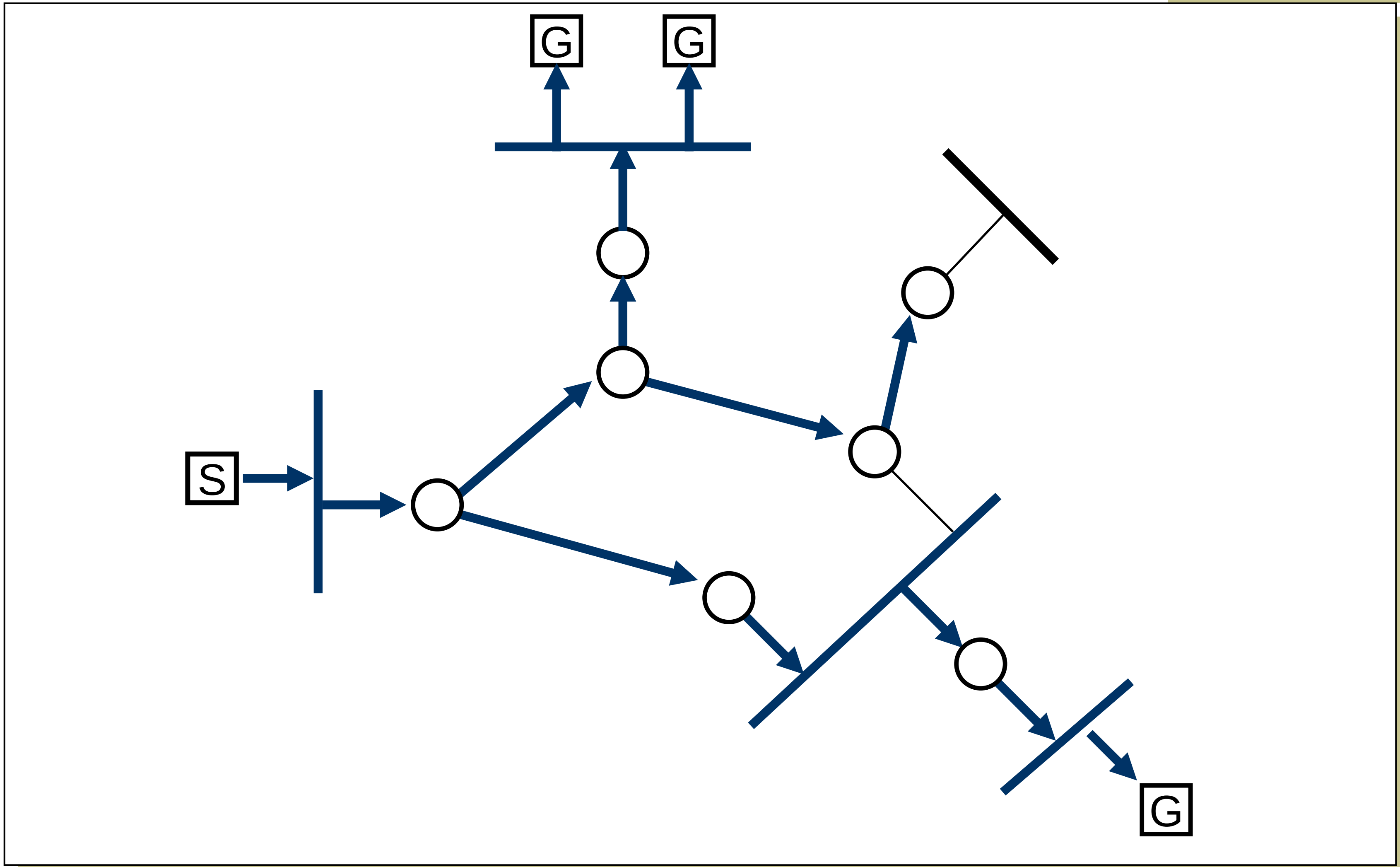
Distance-Vector Multicast Routing

- DVMRP consists of two major components:
  - A conventional distance-vector routing protocol (like RIP)
  - A protocol for determining how to forward multicast packets, based on the routing table
- DVMRP router forwards a packet if
  - The packet arrived from the link used to reach the source of the packet (reverse path forwarding check – RPF)
  - If downstream links have not pruned the tree

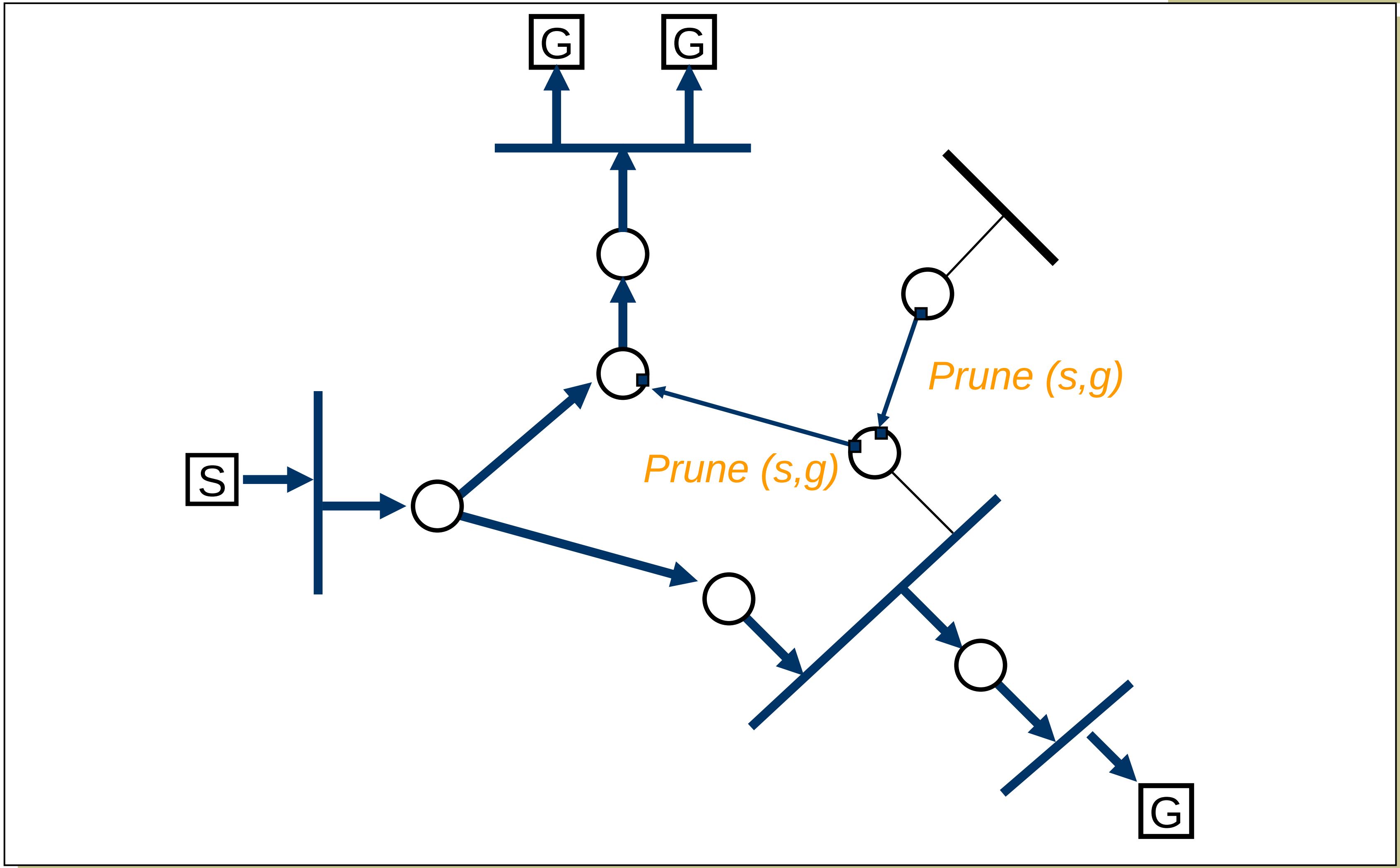
Example Topology



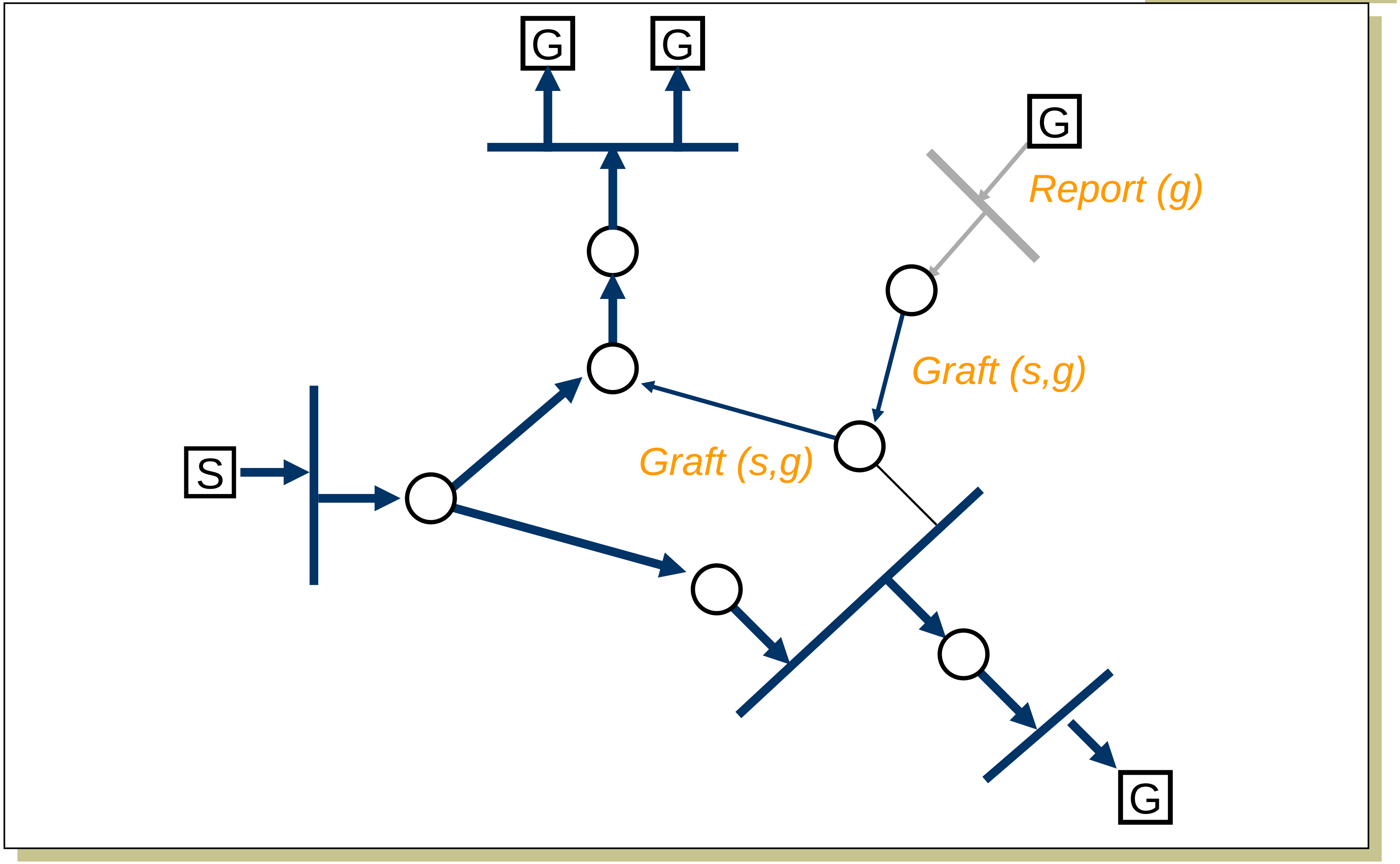
Broadcast with Truncation



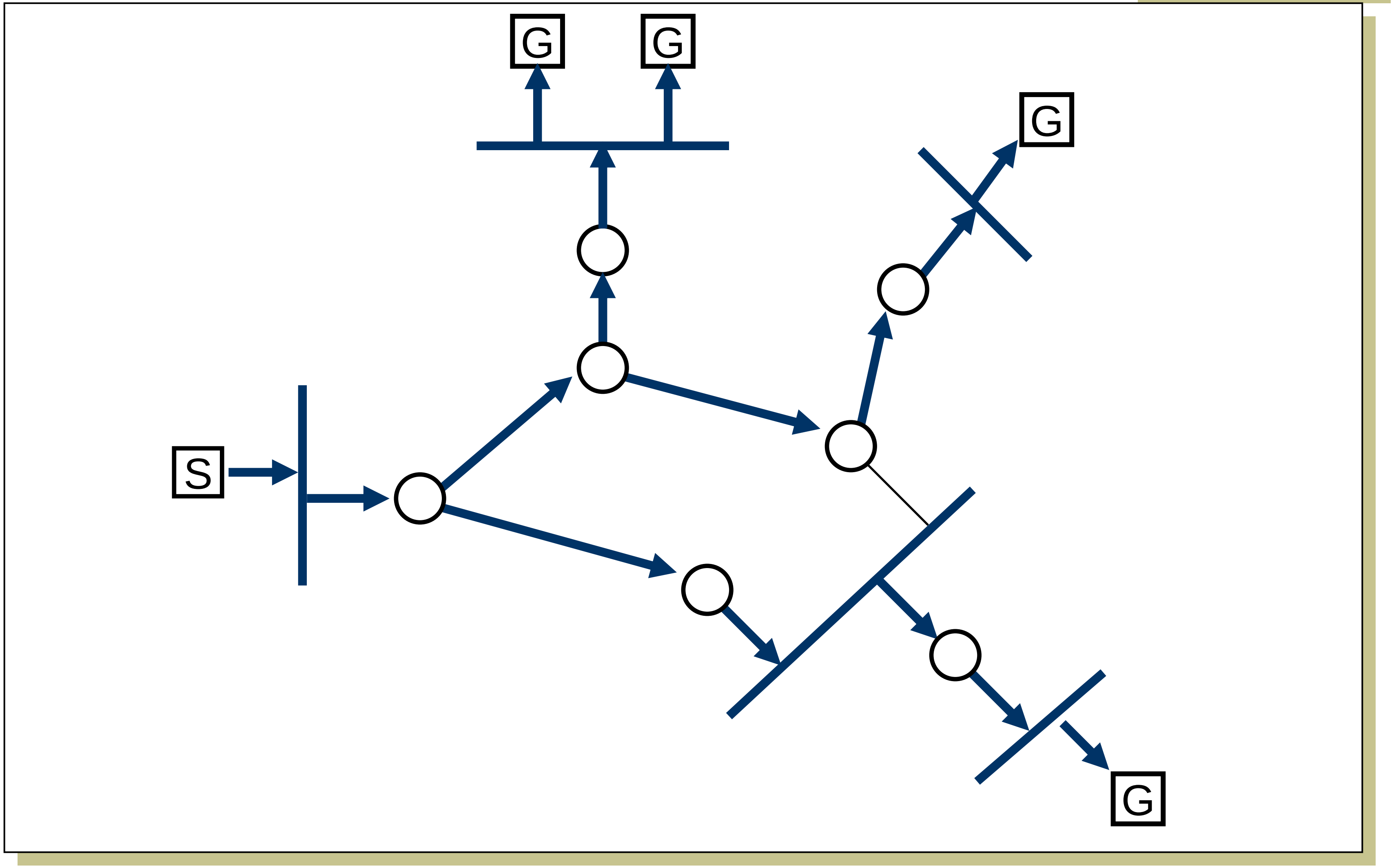
Prune



Graft



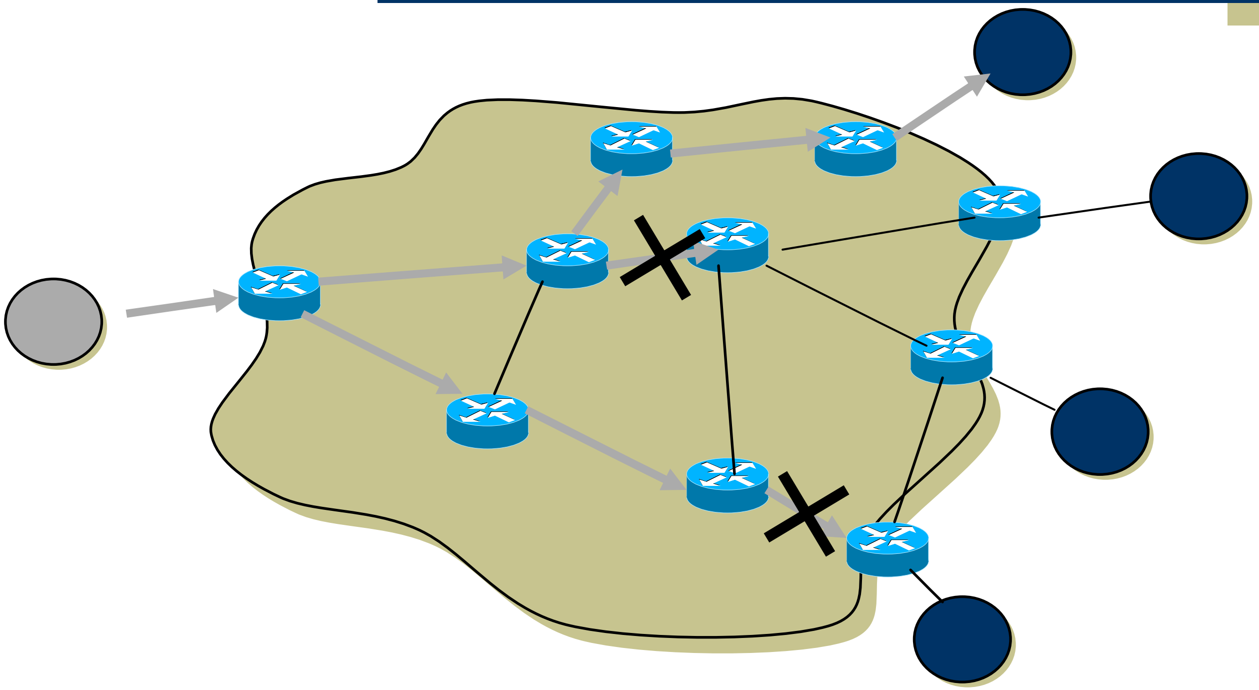
Steady State



Failure of IP Multicast

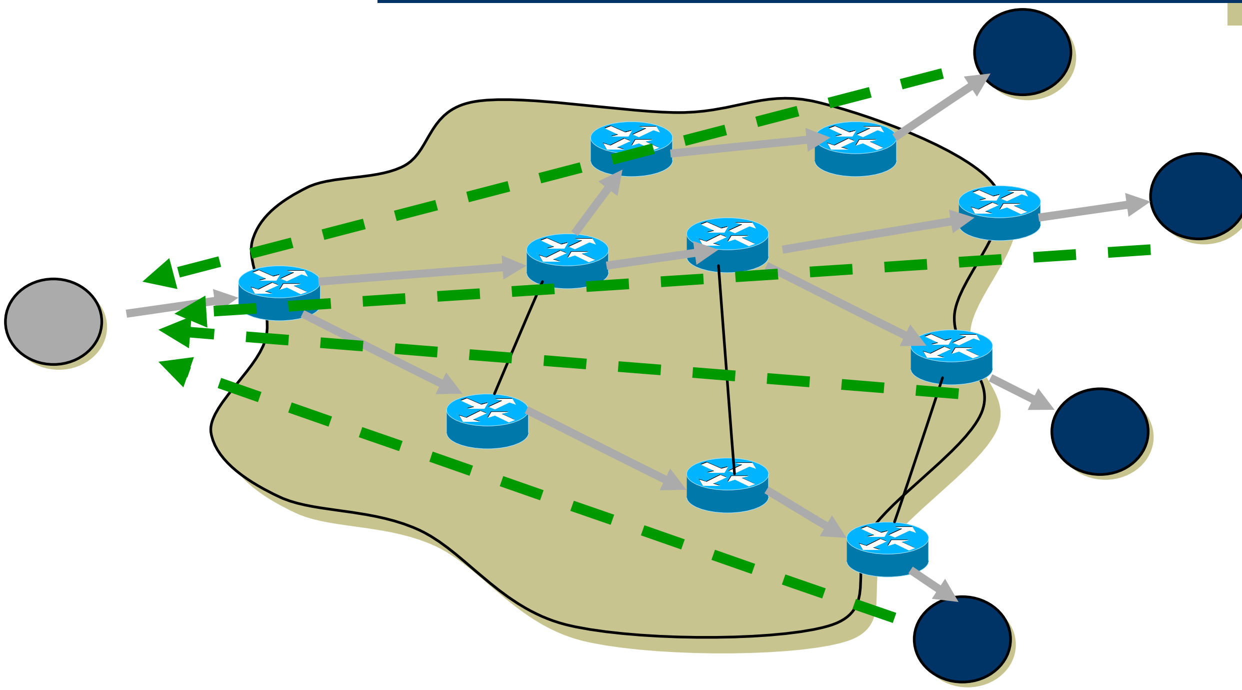
- IP multicast is a powerful service abstraction
  - Too general, too powerful?
- Not widely deployed even after 15 years!
  - Use carefully – e.g., on LAN or campus, rarely over WAN
- Various issues

Error Control: Reliable Multicast



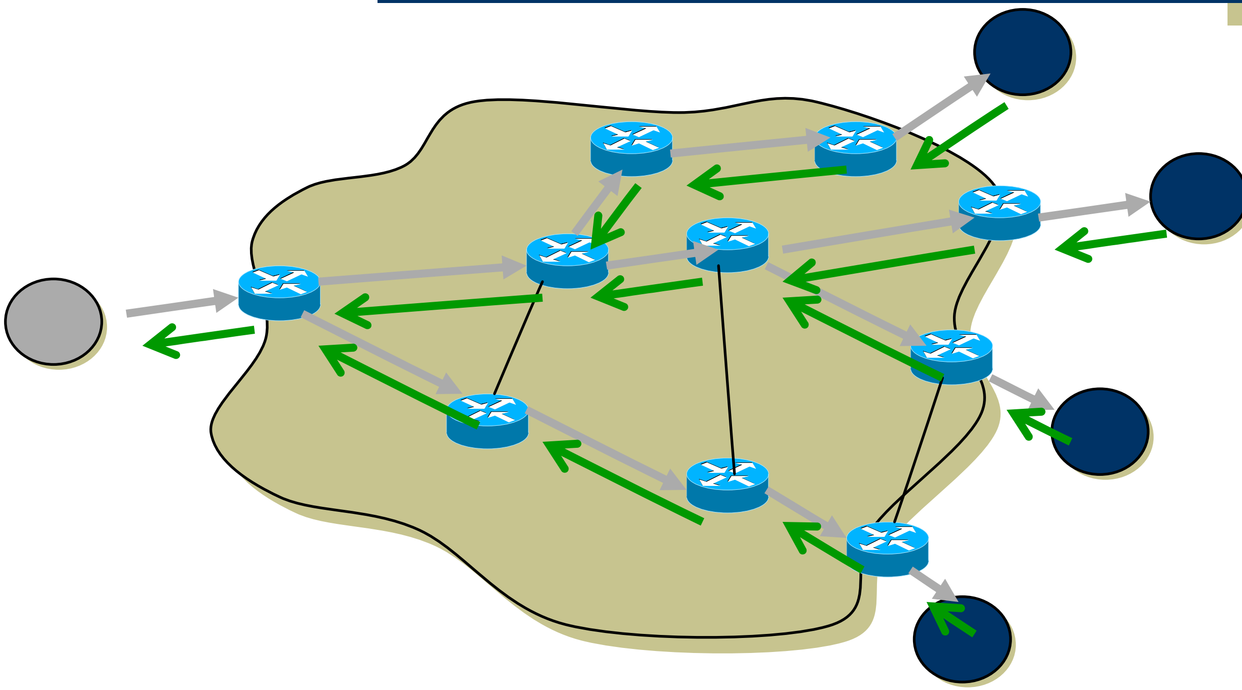
- IP multicast is best-effort
- How to achieve reliable delivery?

Ack Implosion



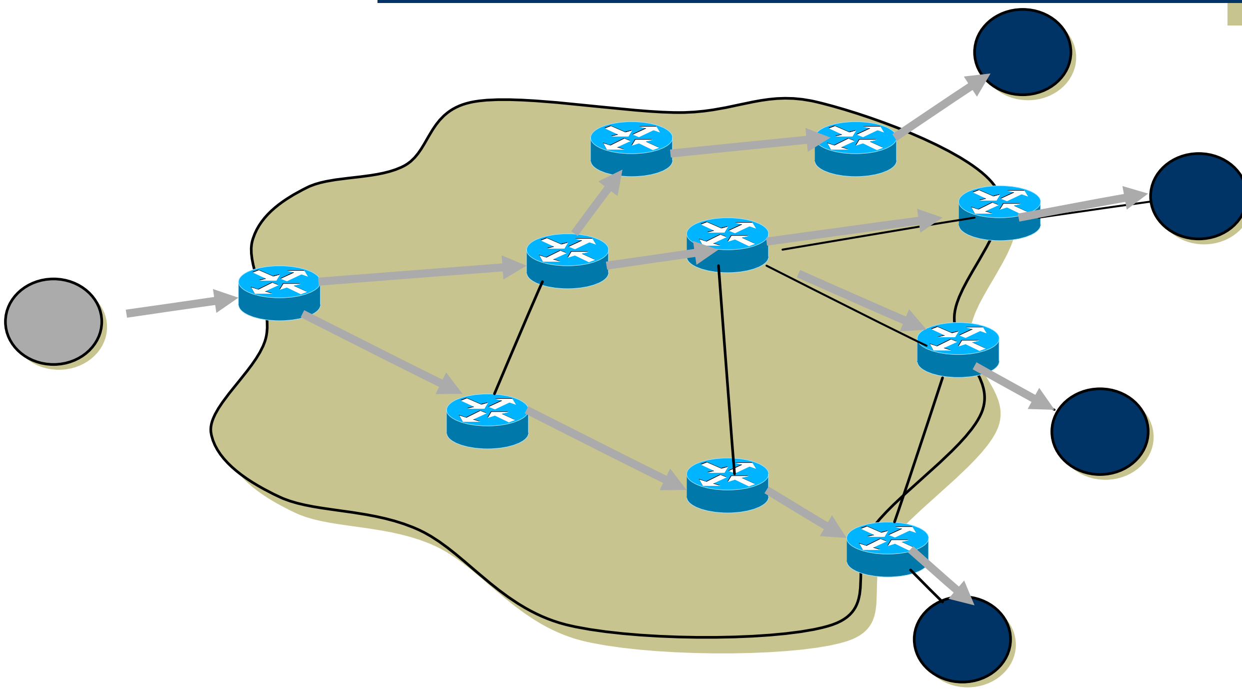
- Scalability: number of acks increase with number of receivers

Routers Collect Acks



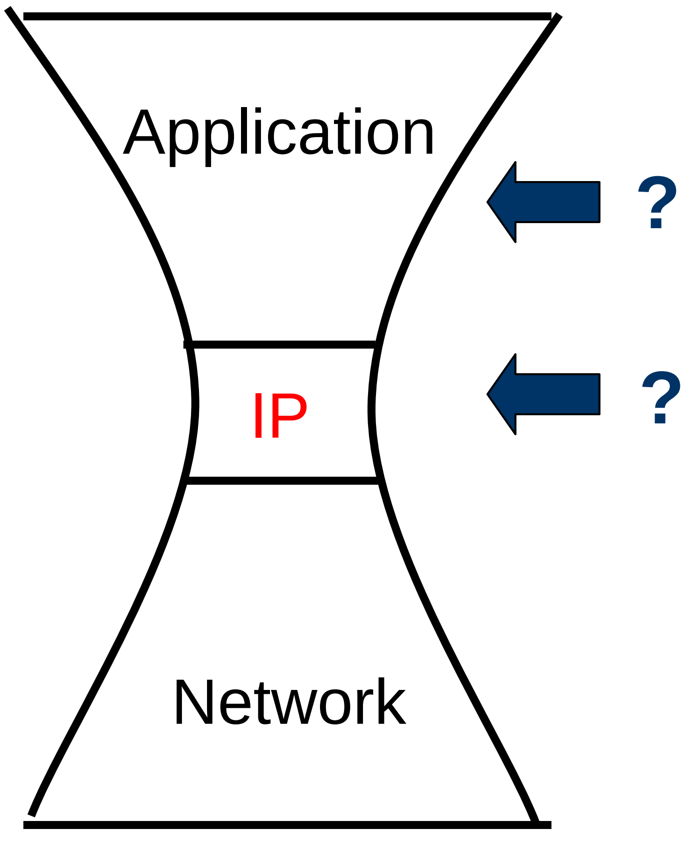
- Overload router functionalities
  - even more per group states

Congestion/Flow Control



- Diverse link technologies: different rates on each link
- Dynamic network condition: available bandwidth changes on each link
- What rate should sender transmit?

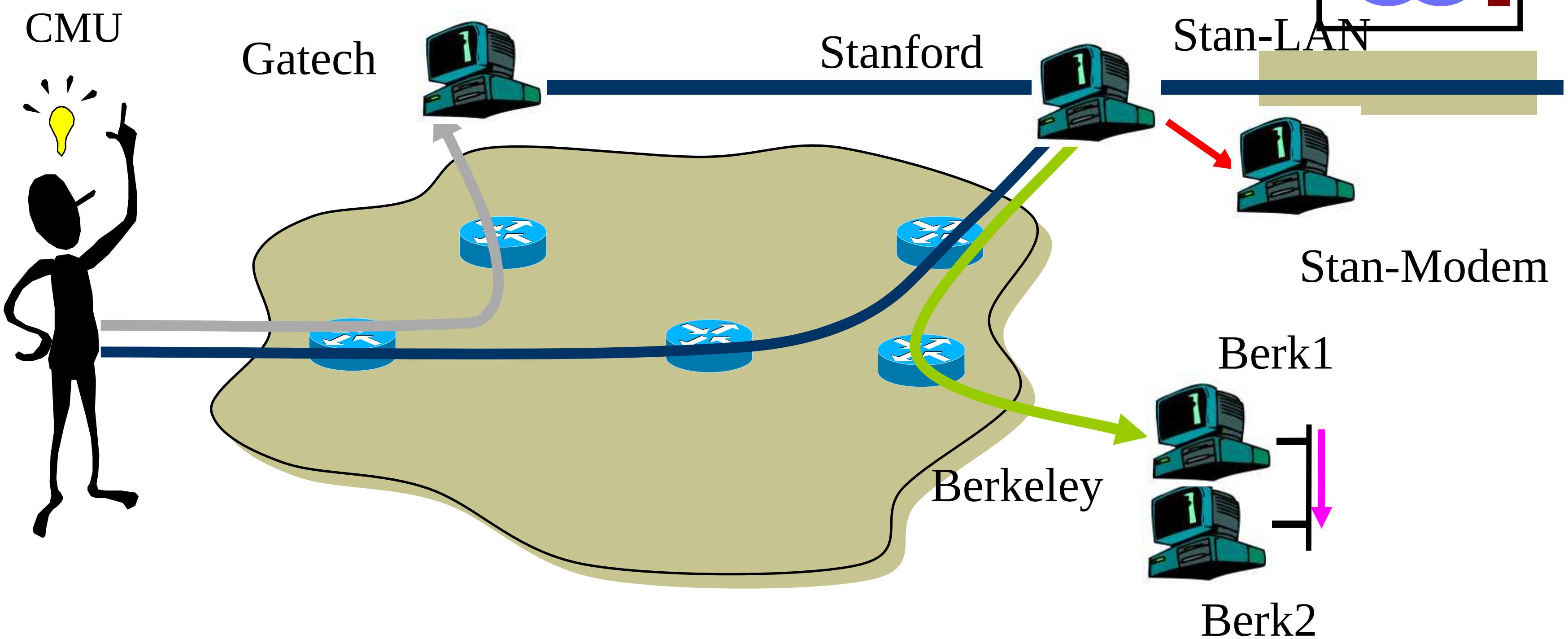
Supporting Multicast on the Internet



Internet architecture

At which layer should multicast be implemented?

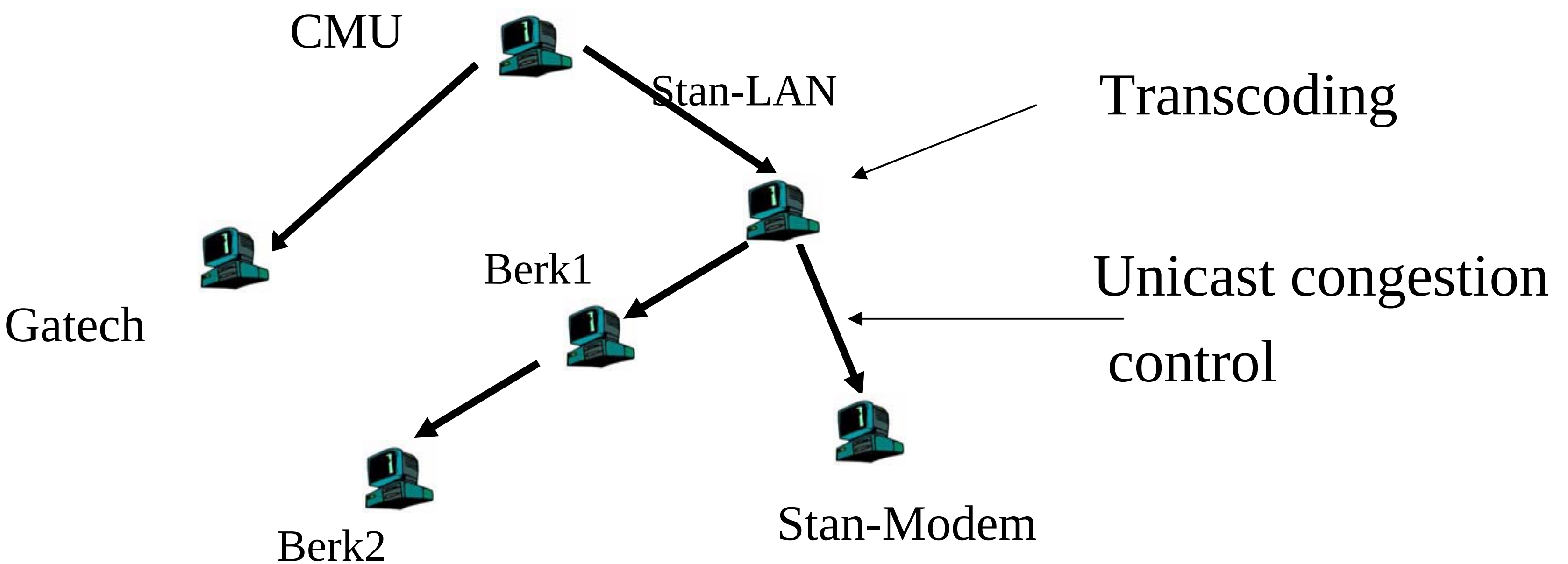
End System Multicast



Overlay Tree

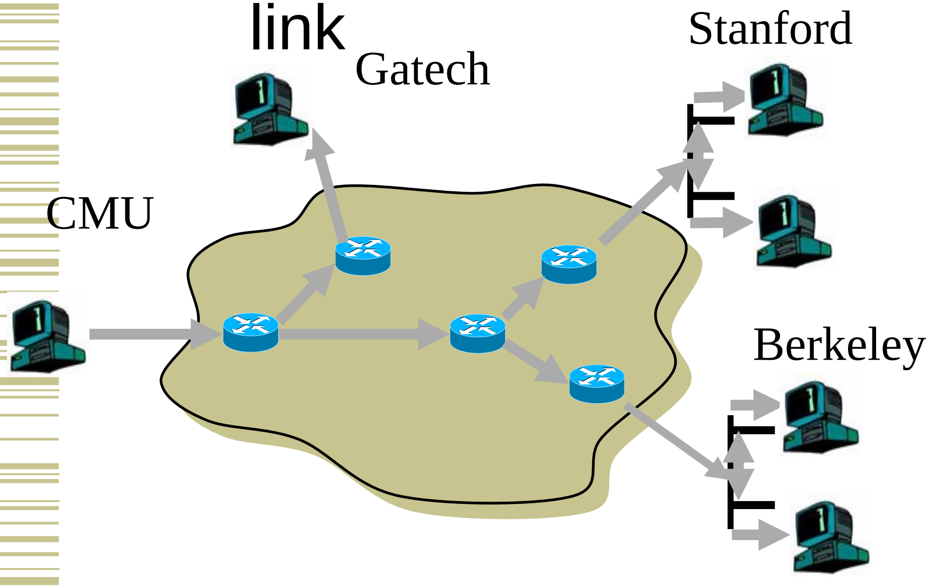
End System Multicast: Benefits

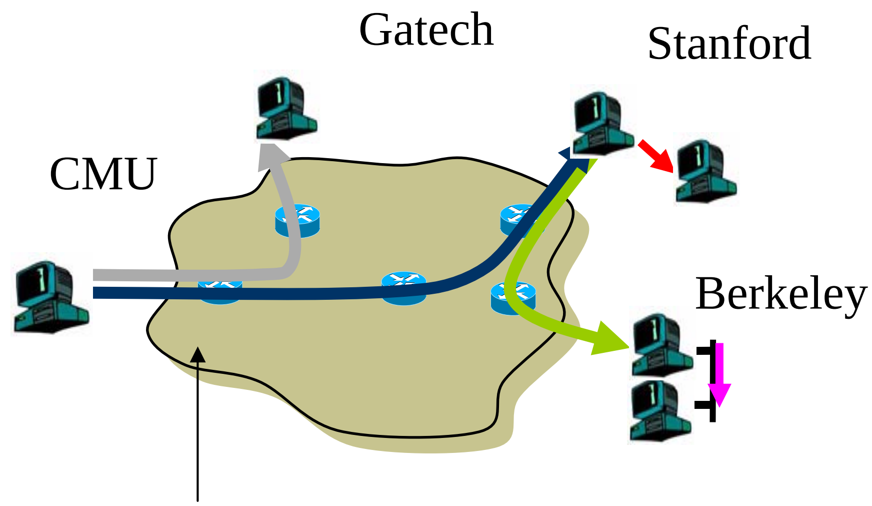
- Scalability
  - Routers do not maintain per-group state
- Easy to deploy
  - Works over the existing IP infrastructure
- Can simplify support for higher level functionality



Performance Challenges

- Degradation in application performance: delay, throughput
- Network overhead: packet duplication over the same link





Two copies of the same packet

Important Concepts

- Multicast provides support for efficient data delivery to multiple recipients
- Requirements for IP Multicast routing
  - Keeping track of interested parties
  - Building distribution tree
  - Broadcast/suppression technique
- Difficult to deploy new IP-layer functionality
- End system multicast