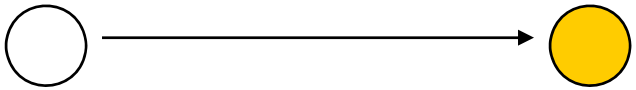
Data Networks
UdS and IMPRS-CS

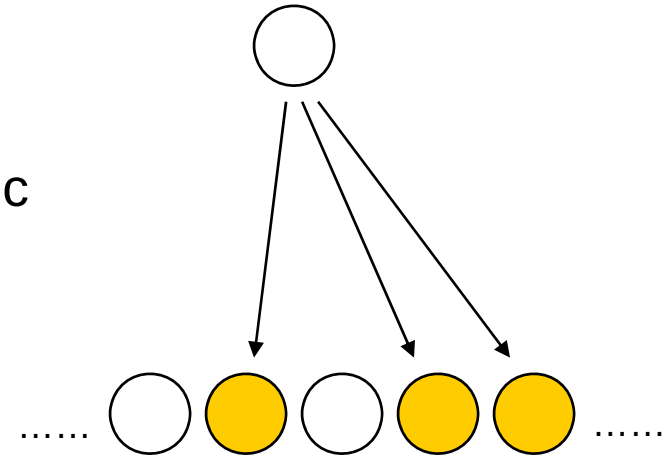
Lecture 18: Multicast Routing

Terminologies

- What is “unicast”?
  - Cast: to send, to throw
    - Broadcast: to send everywhere (recall broadcast in local area Ethernet, ARP, DHCP)
  - Unicast: to send to a single receiver
    - Point-to-point communication
  - Nearly all wide-area Internet traffic is unicast
    - Web traffic, SSH traffic, FTP traffic
    - Two unicast streams, one in each direction



- What is “multicast”?
  - In between unicast and broadcast
  - Each packet is sent to multiple specific receivers
    - Point-to-Multipoint communication
  - What is multicast useful for???



Example Uses

- Internet TV radio
- Stock price update
- Video conference
- Spam?!

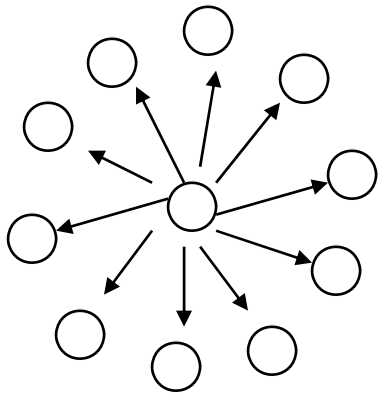
How to Send to Multiple Receivers?

- What are the simplest ways?
- Ex1: Send a copy of the packet to one receiver at a time until all receivers have it
  - i.e. use unicast to implement multicast
- Ex2: Flood a packet throughout the network and have non-receivers discard the packet
  - i.e. use broadcast to implement multicast
- Advantages? Disadvantages?

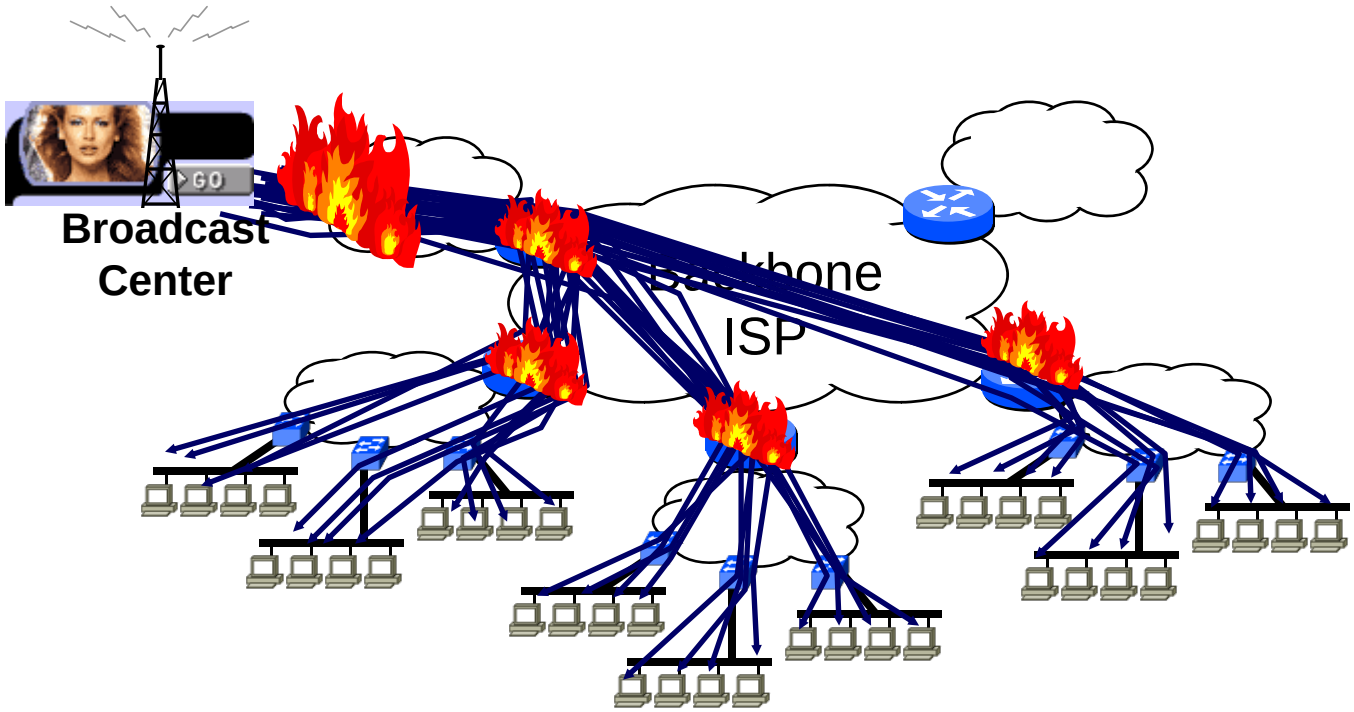
- In general: We want a distribution tree
  - Many ways to do it
  - Big research topic for a decade

Example: Internet Radio

- www.digitallyimported.com
  - Sends out 128Kb/s MP3 music streams
  - Peak usage ~9000 simultaneous streams
  - Consumes ~1.1Gb/s
    - bandwidth costs are large fraction of their expenditures
  - A fat and shallow tree
  - Does not scale!

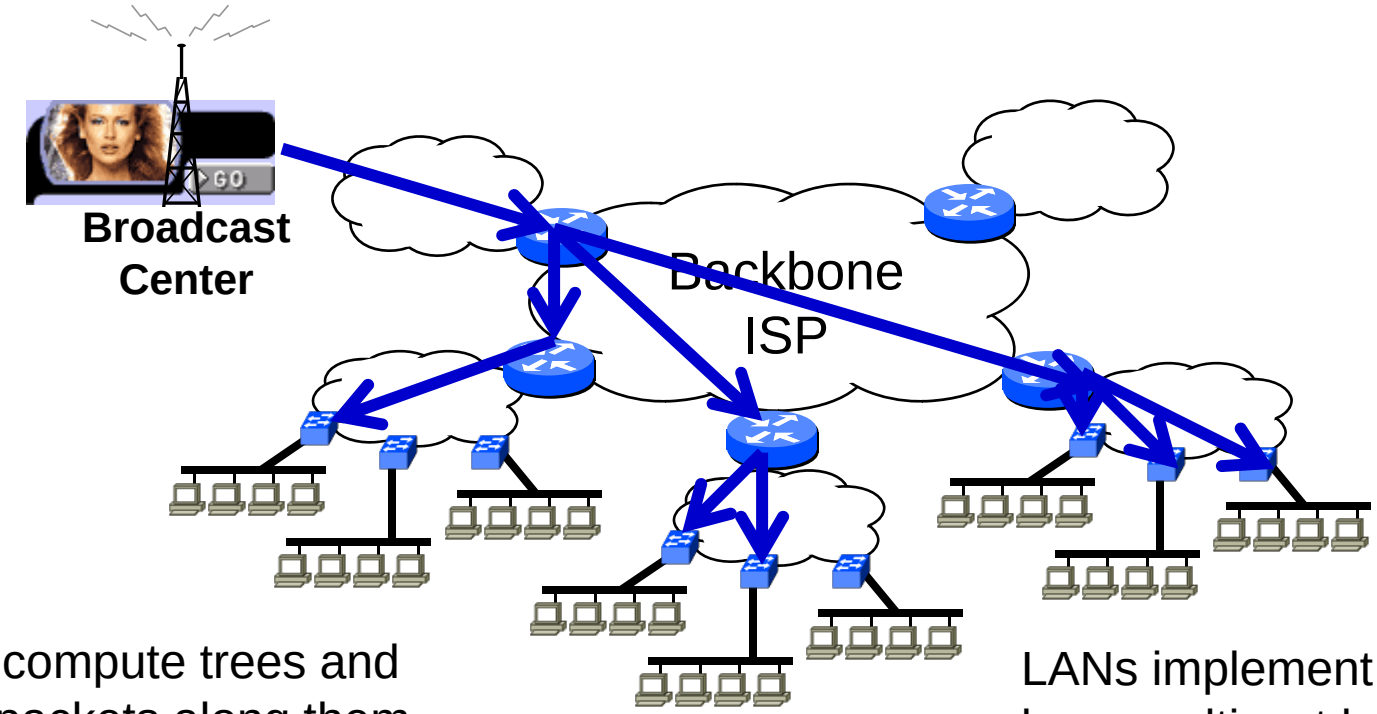


This approach does not scale…



Use routers in distribution tree

Copy data at routers
At most one copy of a data packet per link



Routers compute trees and forward packets along them

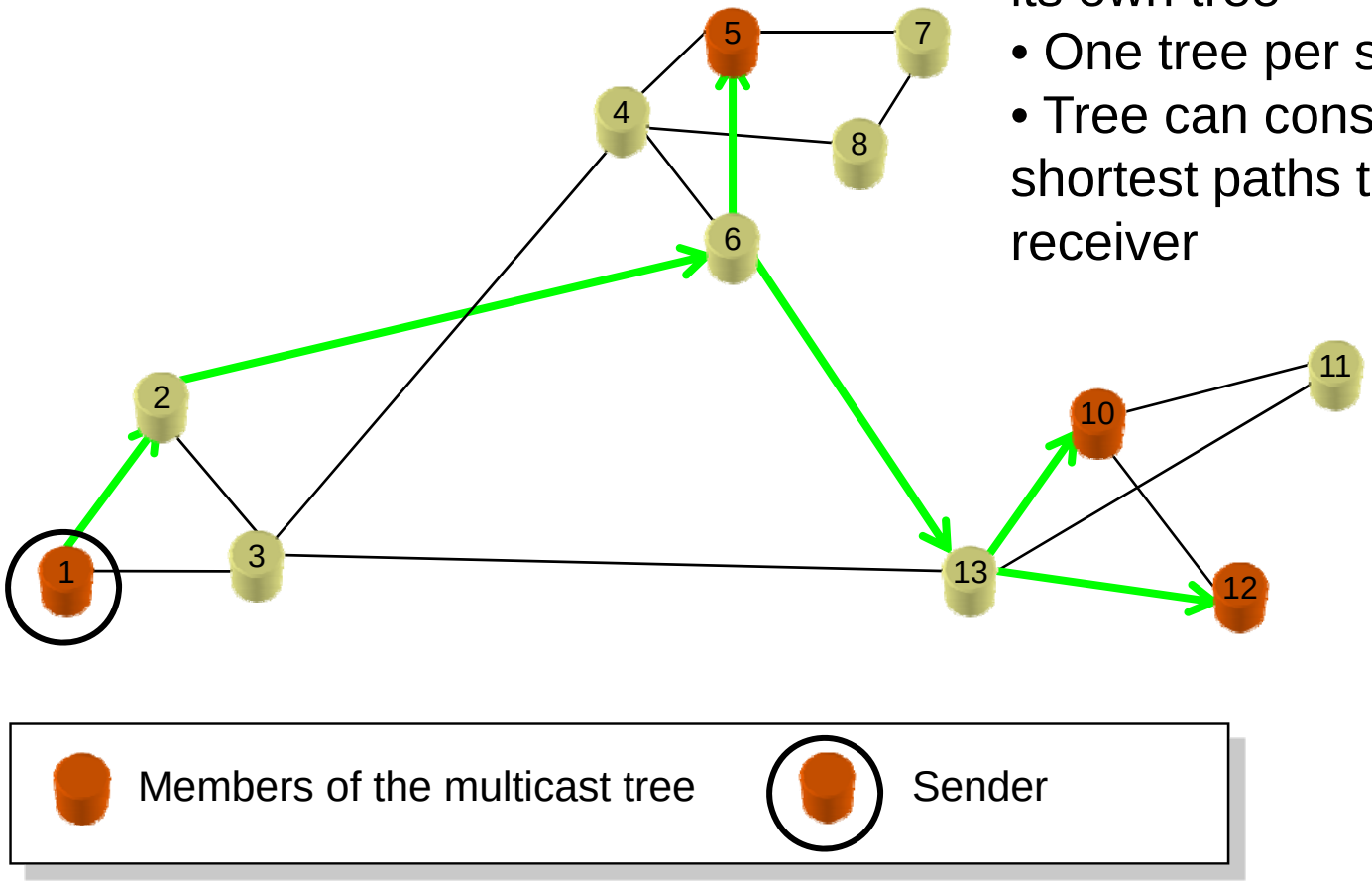
LANs implement link layer multicast by broadcasting

Multicast Routing Approaches

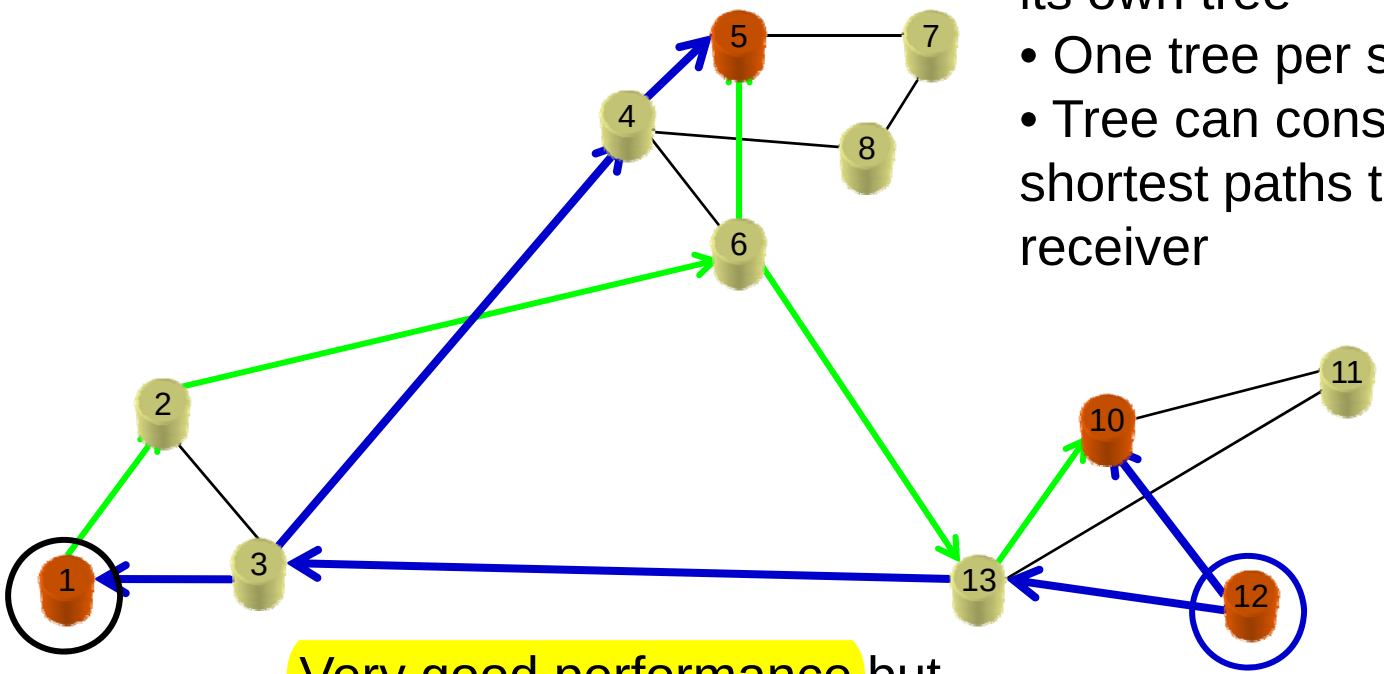
- Kinds of Trees
  - Source Specific Trees
    - Most suitable for single sender
    - E.g. internet radio
  - Shared Tree
    - Multiple senders in a group
    - E.g. Teleconference
- Tree Computation Methods
  - Link state
  - Distance vector

Source Specific Trees

- Each source is the root of its own tree
- One tree per source
- Tree can consists of shortest paths to each receiver



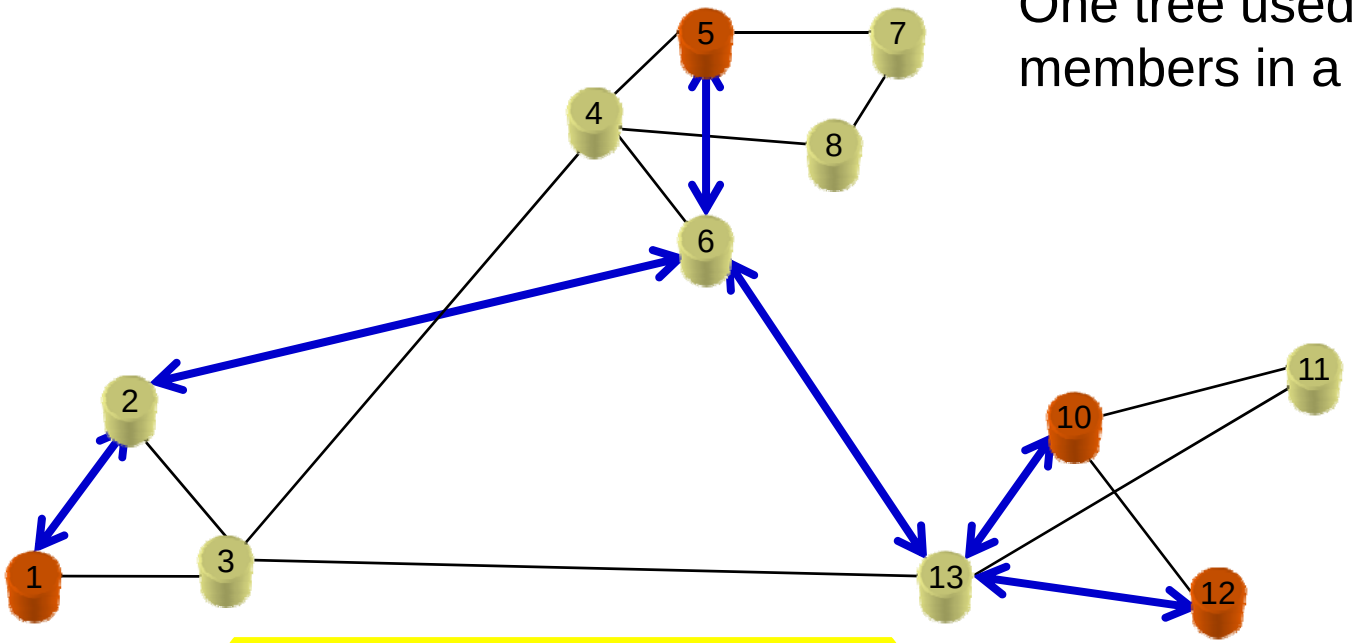
Source Specific Trees



- Each source is the root of its own tree
- One tree per source
- Tree can consists of shortest paths to each receiver

Very good performance but expensive to construct/maintain; routers need to manage a tree per source

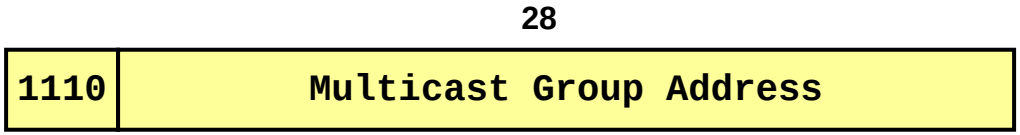
Shared Tree



One tree used by all members in a group

Easier to construct/maintain but hard to pick “good” trees for everyone!

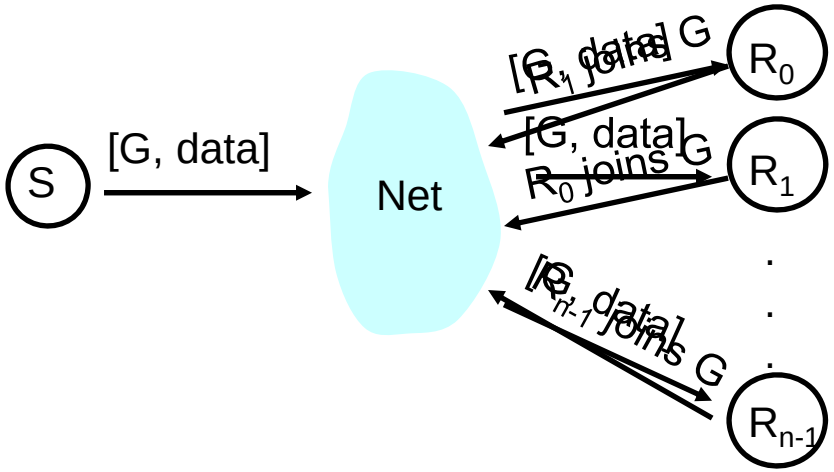
IPv4 Multicast



First octet: 224 - 239

- Class D addresses
  - These are group identifiers
  - Not specific to an end host
  - Flat address space
  - In practice, pick a group address at random, hope no collision
  - No security in the network layer
- Will use “G” to designate an IP multicast group address

IP Multicast Service Model



- Receivers join a multicast group which is identified by a multicast address (e.g. G)
- Sender(s) send data to address G
- Network routes data to each of the receivers

Multicast Implementation Issues

- How is join implemented?
- How is send implemented?
- How much information about trees is kept and who keeps it?

IP Multicast Routing

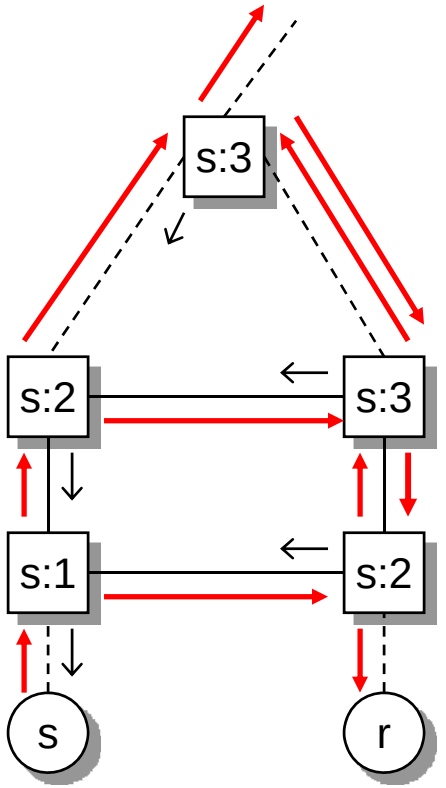
- Intra-domain
  - Distance-vector multicast
  - Link-state multicast
- Inter-domain
  - Protocol Independent Multicast, Sparse Mode
    - Key idea: Core-Based Tree

Distance Vector Multicast Routing Protocol (DVRMP)

- An elegant extension to DV routing
- Use shortest path DV routes to determine if link is on the source-rooted spanning tree
- Three steps in developing DVRMP
  - Reverse Path Flooding
  - Reverse Path Broadcasting
  - Truncated Reverse Path Broadcasting

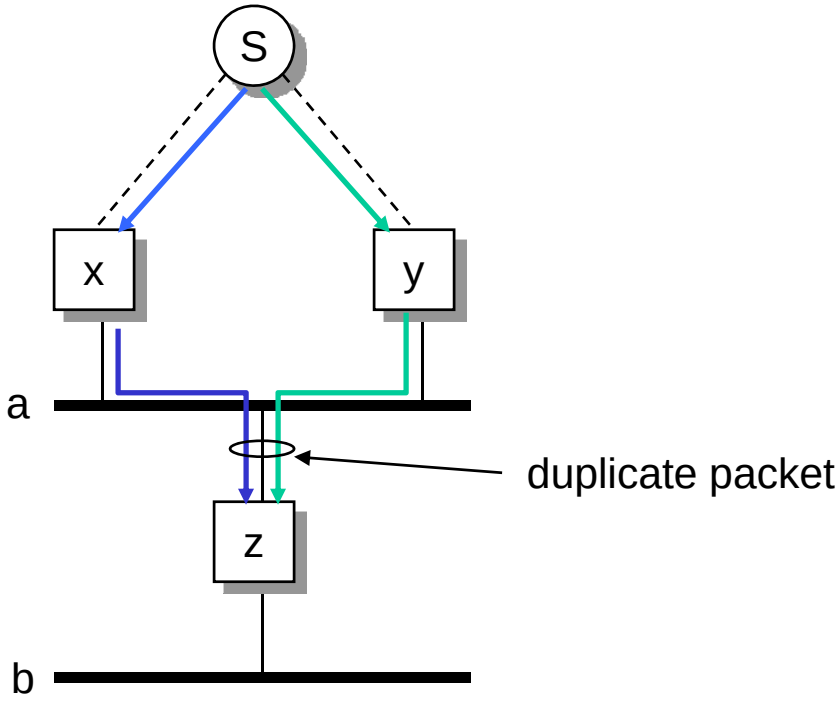
Reverse Path Flooding (RPF)

- Extension to DV unicast routing
- Packet forwarding
  - If incoming link is shortest path to source
  - Send on all links except incoming
  - Packets always take shortest path
    - assuming delay is symmetric
- Issues
  - Some links (LANs) may receive multiple copies
  - Every link receives each multicast packet, even if no interested hosts



Example

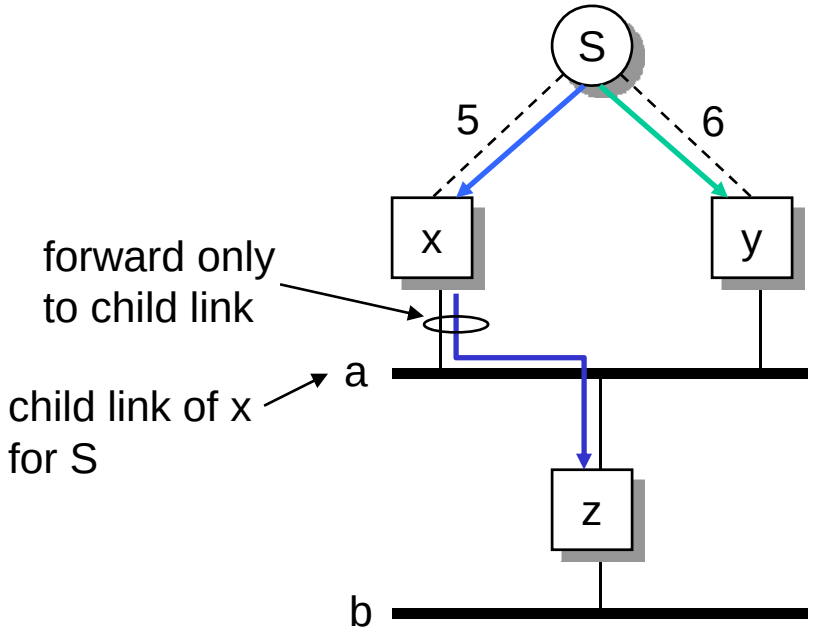
- Flooding can cause a given packet to be sent multiple times over the same link



- Solution: Called “Reverse Path Broadcasting”

Reverse Path Broadcasting (RPB)

- Chose parent of each link along reverse shortest path to source
- Only parent forward to a link (child link)
- Use DV routing update to identify parent

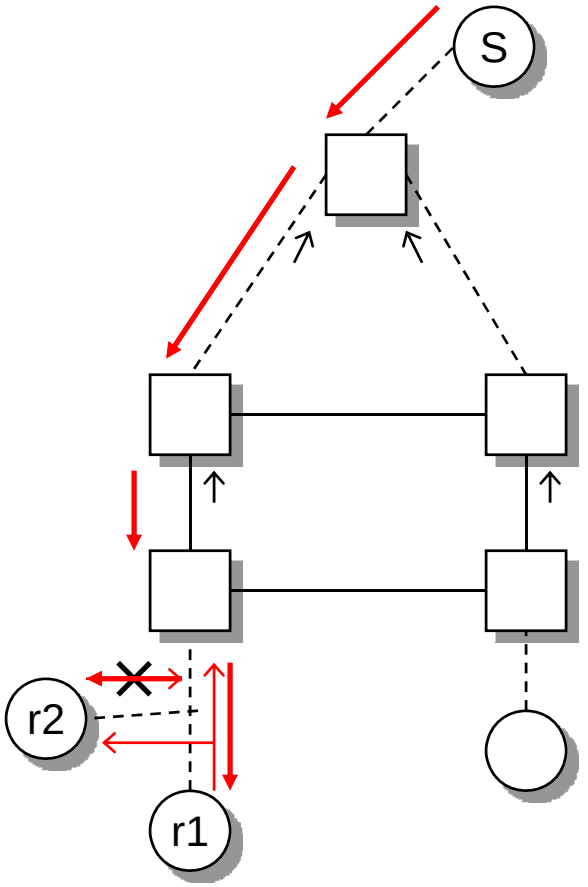


Don’t Really Want to Flood!

- This is still a broadcast algorithm – the traffic goes everywhere
- Need to “Prune” the tree when there are subtrees with no group members
- Solution: Truncated Reverse Path Broadcasting

Truncated Reverse Path Broadcasting (TRPB)

- Extend RPB to eliminate unneeded forwarding
- Explicit group joining
  - Members periodically send “join” requests
  - If another LAN member has joined (overheard join message), other members do not send join message
- Router with no member downstream is removed from tree
  - Router sends “prune” message to upstream router when no member



Distance Vector Multicast Scaling

- State requirements:
  - O(Sources $\times$ Groups) active state

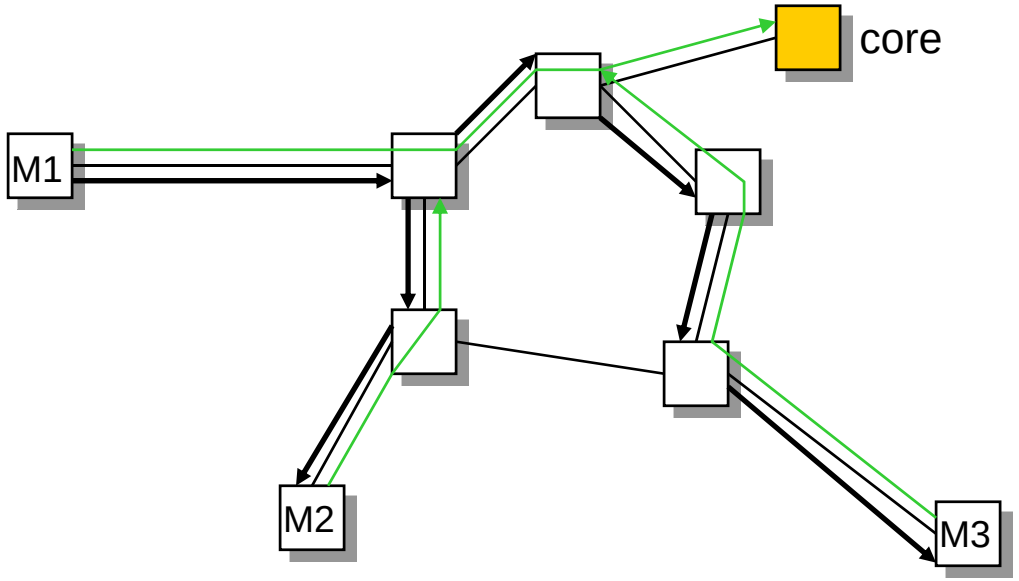
Core Based Trees (CBT)

- The key idea in Inter-domain PIM-SM protocol
- Pick a “rendezvous point” for the group called the core
  - Build a tree towards the core
  - Union of the unicast paths from members to the core
  - Shared tree

- To send, unicast packet to core and bounce it back to multicast group
- Reduce routing table state from O(S x G) to O(G)

Example

- Group members: M1, M2, M3
- M1 sends data



control (join) messages
data

Disadvantages

- Sub-optimal delay
- Single point of failure
  - Core goes out and everything lost until error recovery elects a new core
- Small, local groups with non-local core
  - Need good core selection
  - Optimal choice (computing topological center) is NP hard