15-744: Computer Networking

L-20 Multicast

Multicast Routing

- Unicast: one source to one destination
- Multicast: one source to many destinations
- Two main functions:
  - Efficient data distribution
  - Logical naming of a group

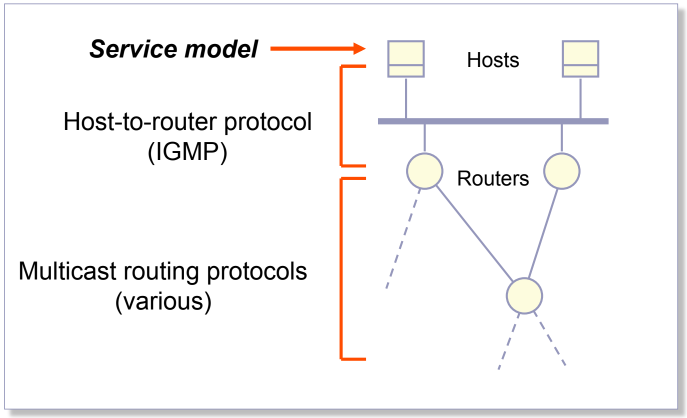
Example Applications

- Broadcast audio/video
- Push-based systems
- Software distribution
- Web-cache updates
- Teleconferencing (audio, video, shared whiteboard, text editor)
- Multi-player games
- Server/service location
- Other distributed applications

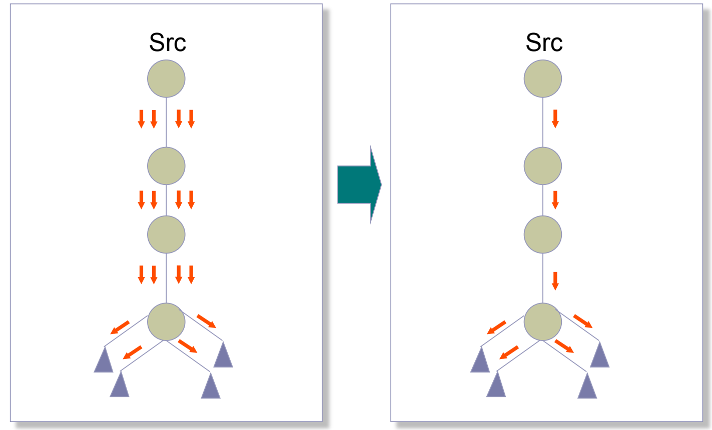
Overview

- IP Multicast Service Basics
- Multicast Routing Basics
- Overlay Multicast
- Reliability
- Congestion Control

IP Multicast Architecture



Multicast – Efficient Data Distribution



Multicast Router Responsibilities

- Learn of the existence of multicast groups (through advertisement)
- Identify links with group members
- Establish state to route packets
  - Replicate packets on appropriate interfaces
  - Routing entry:

Src, incoming interface	List of outgoing interfaces

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 Addresses

- Class D IP addresses
  - 224.0.0.0 – 239.255.255.255

1 1 1 0	Group ID

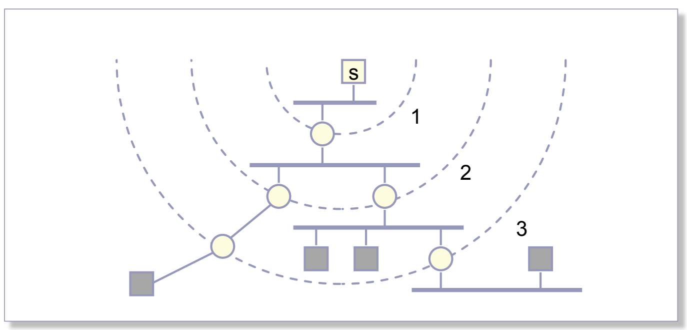
- How to allocated these addresses?
  - Well-known multicast addresses, assigned by IANA
  - Transient multicast addresses, assigned and reclaimed dynamically, e.g., by “sdr” program

Multicast Groups

- Members are the intended receivers
- Senders may or may not be members
- Hosts may belong to many groups
- Hosts may send to many groups
- Support dynamic creation of groups, dynamic membership, dynamic sources

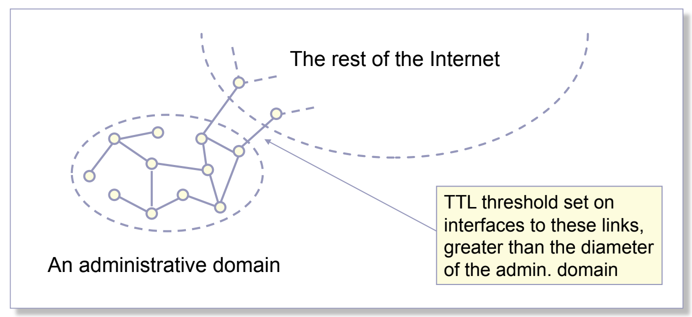
Multicast Scope Control – Small TTLs

- 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

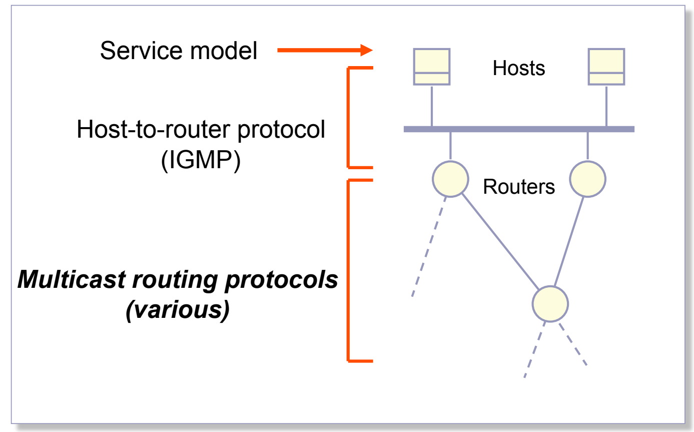


Overview

- IP Multicast Service Basics
- Multicast Routing Basics
- Overlay Multicast
- Reliability
- Congestion Control

13

IP Multicast Architecture



14

Multicast Routing

- Basic objective – build distribution tree for multicast packets
- Multicast service model makes it hard
  - Anonymity
  - Dynamic join/leave

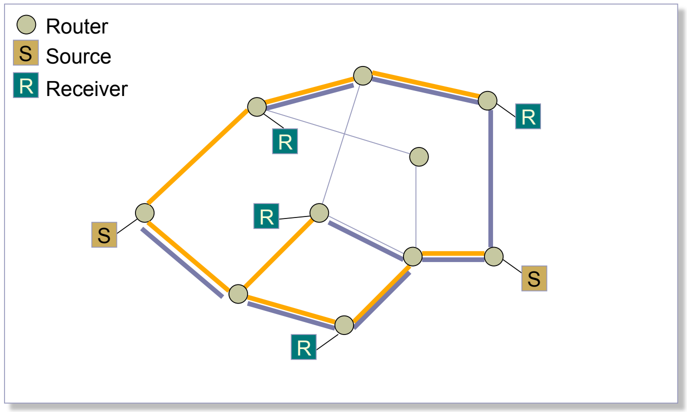
15

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

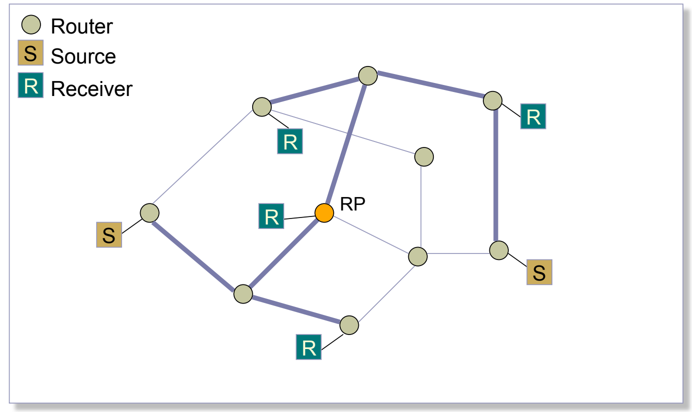
16

Source-based Trees



17

Shared Tree



18

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!

19

Routing Techniques

- 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*

20

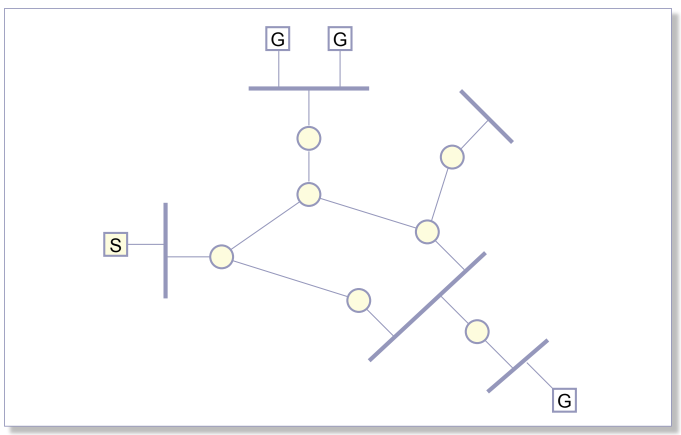
Routing Techniques

- Core based protocols
  - Specify “meeting place” aka core
  - Sources send initial packets to core
  - Receivers join group at core
  - Requires mapping between multicast group address and “meeting place”
  - Examples: CBT, PIM-SM

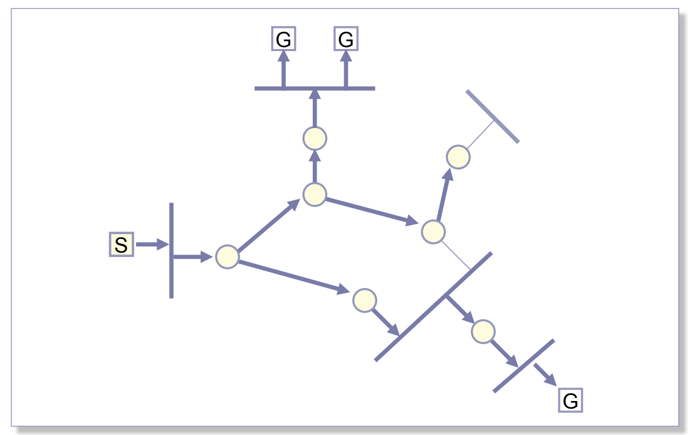
Distance-Vector Multicast Routing

- 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

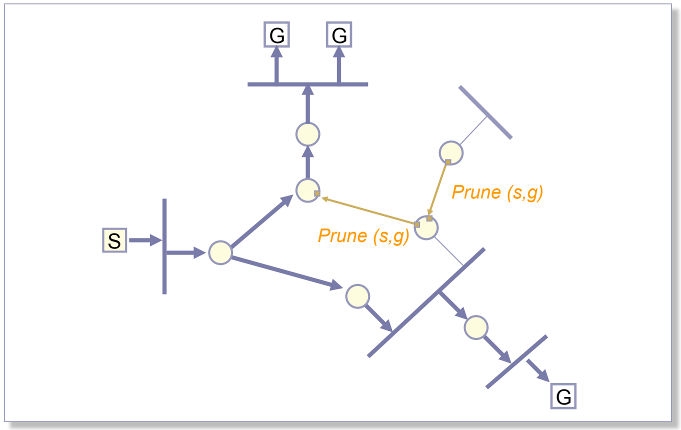
Example Topology



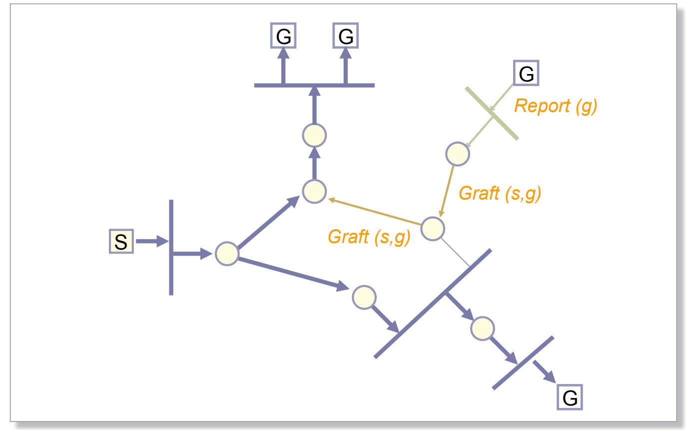
Broadcast with Truncation



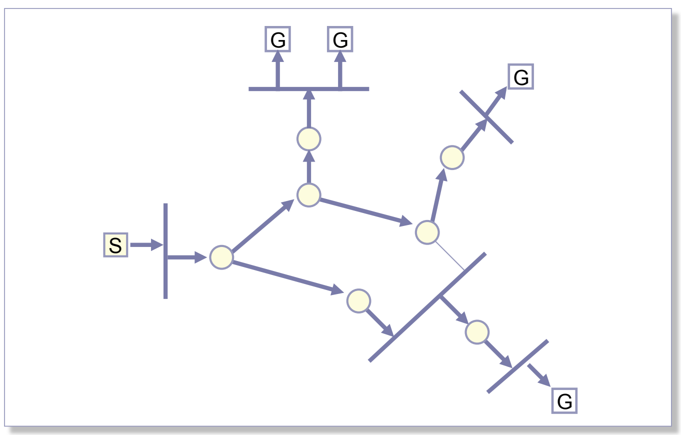
Prune



Graft



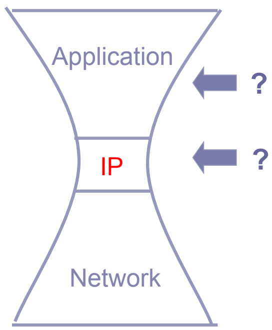
Steady State



Overview

- IP Multicast Service Basics
- Multicast Routing Basics
- Overlay Multicast
- Reliability
- Congestion Control

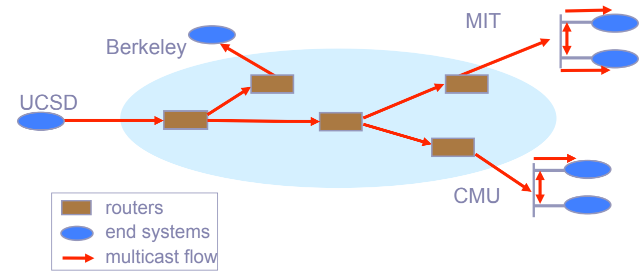
Supporting Multicast on the Internet



Internet architecture

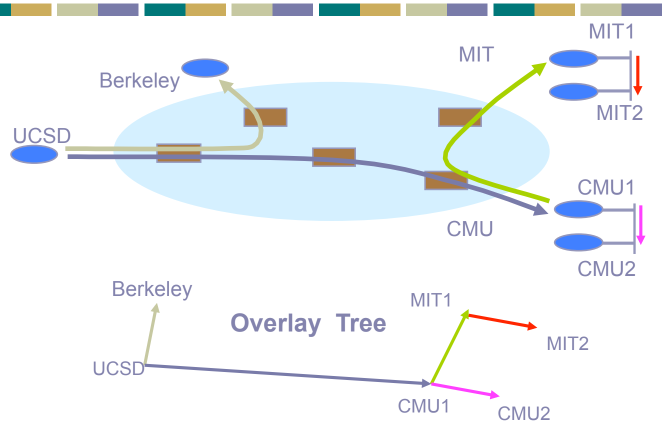
At which layer should multicast be implemented?

IP Multicast



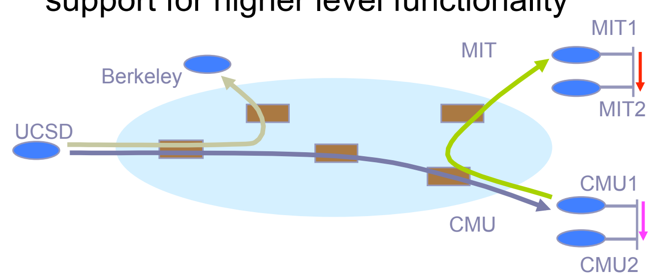
- Highly efficient
- Good delay

End System Multicast



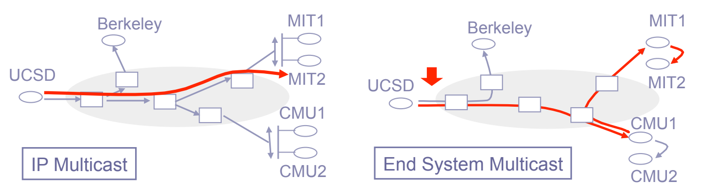
Potential Benefits Over IP Multicast

- Quick deployment
- All multicast state in end systems
- Computation at forwarding points simplifies support for higher level functionality



Concerns with End System Multicast

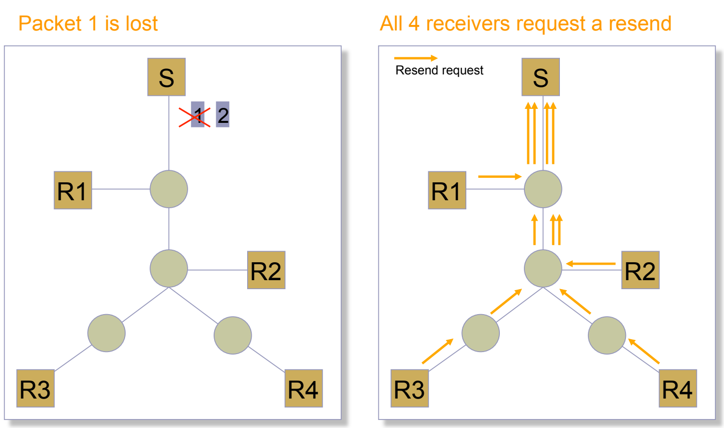
- Self-organize recipients into multicast delivery overlay tree
  - Must be closely matched to real network topology to be efficient
- Performance concerns compared to IP Multicast
  - Increase in delay
  - Bandwidth waste (packet duplication)



Overview

- IP Multicast Service Basics
- Multicast Routing Basics
- Overlay Multicast
- Reliability
- Congestion Control

Implosion



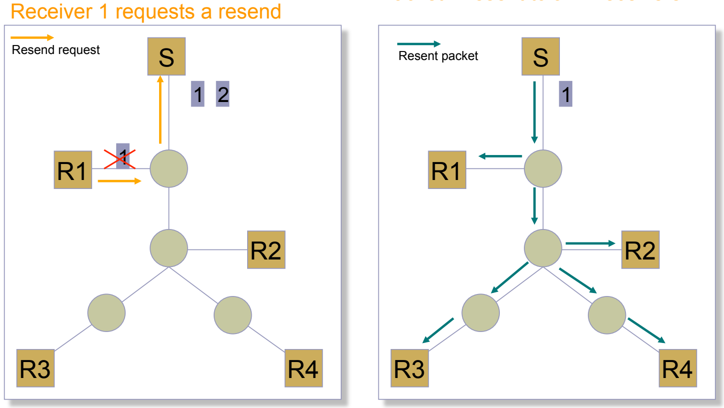
Retransmission

- Re-transmitter
  - Options: sender, other receivers
- How to retransmit
  - Unicast, multicast, scoped multicast, retransmission group, …
- Problem: Exposure

Exposure

Packet 1 does not reach R1; Receiver 1 requests a resend

Packet 1 resent to all 4 receivers

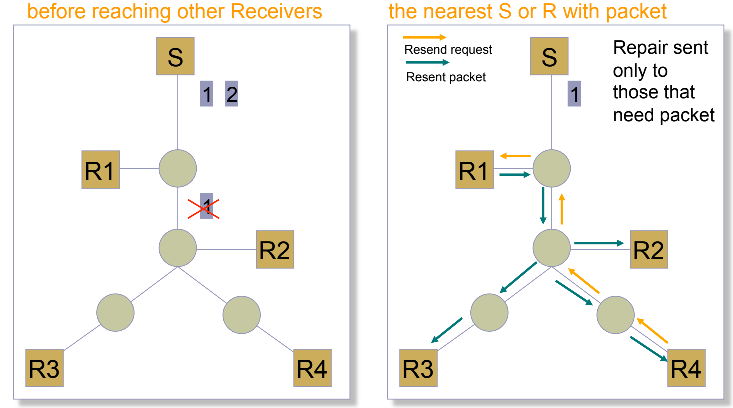


37

Ideal Recovery Model

Packet 1 reaches R1 but is lost before reaching other Receivers

Only one receiver sends NACK to the nearest S or R with packet



38

Scalable Reliable Multicast (SRM)

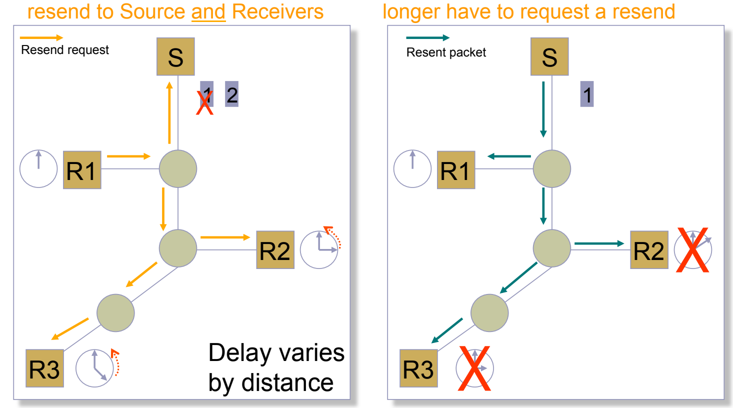
- Originally designed for *wb*
- Receiver-reliable
  - NACK-based
- Every member may multicast NACK or retransmission

39

SRM Request Suppression

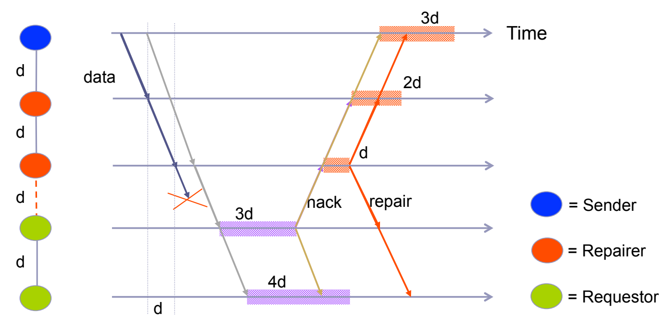
Packet 1 is lost; R1 requests resend to Source <u>and</u> Receivers

Packet 1 is resent; R2 and R3 no longer have to request a resend



40

Deterministic Suppression

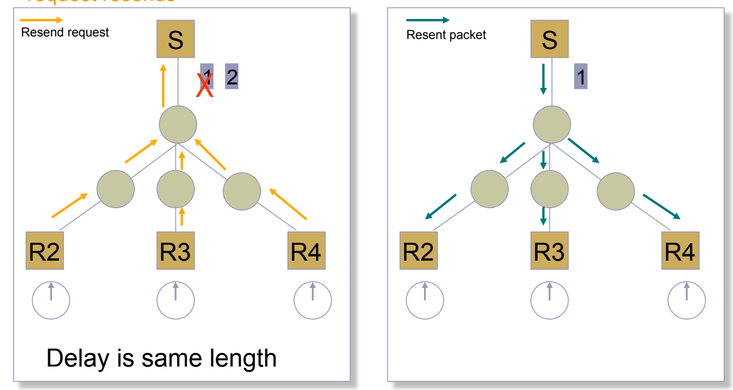


Delay = $C_1 \times d_{S,R}$

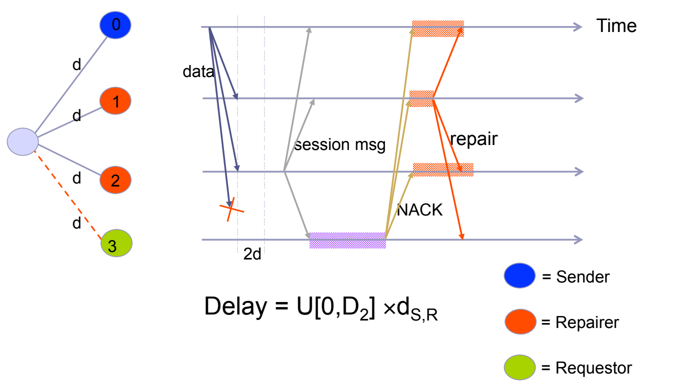
SRM Star Topology

Packet 1 is lost; All Receivers request resends

Packet 1 is resent to all Receivers



SRM: Stochastic Suppression



Delay = $U[0,D_2] \times d_{S,R}$

SRM (Summary)

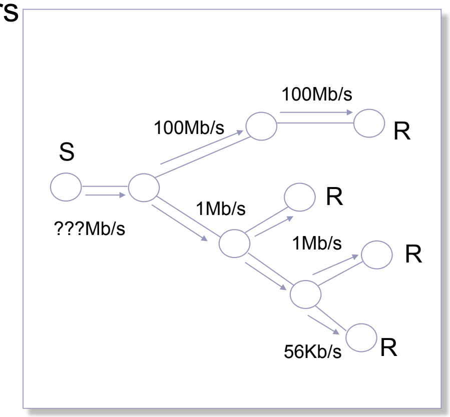
- NACK/Retransmission suppression
  - Delay before sending
  - Delay based on RTT estimation
  - Deterministic + Stochastic components
- Periodic session messages
  - Full reliability
  - Estimation of distance matrix among members

Overview

- IP Multicast Service Basics
- Multicast Routing Basics
- Overlay Multicast
- Reliability
- Congestion Control

Multicast Congestion Control

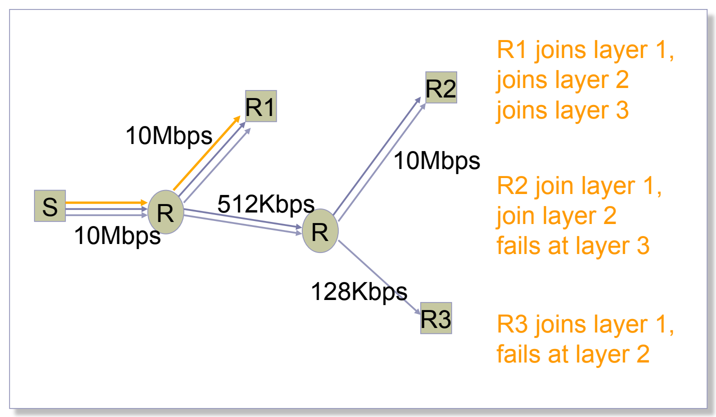
- What if receivers have very different bandwidths?
- Send at max?
- Send at min?
- Send at avg?



Video Adaptation: RLM

- Receiver-driven Layered Multicast
- Layered video encoding
- Each layer uses its own mcast group
- On spare capacity, receivers add a layer
- On congestion, receivers drop a layer
- Join experiments used for shared learning

Layered Media Streams

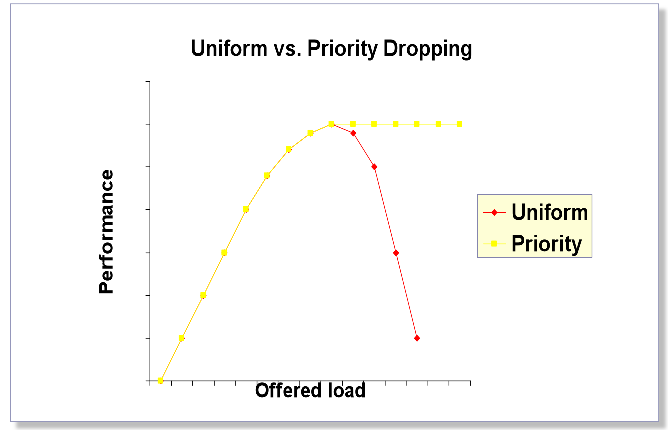


Drop Policies for Layered Multicast

- Priority
  - Packets for low bandwidth layers are kept, drop queued packets for higher layers
  - Requires router support
- Uniform (e.g., drop tail, RED)
  - Packets arriving at congested router are dropped regardless of their layer
- Which is better?

49

RLM Intuition



50

RLM Intuition

- Uniform
  - Better incentives to well-behaved users
  - If oversend, performance rapidly degrades
  - Clearer congestion signal
  - Allows shared learning
- Priority
  - Can waste upstream resources
  - Hard to deploy
- RLM approaches optimal operating point
  - Uniform is already deployed
  - Assume no special router support

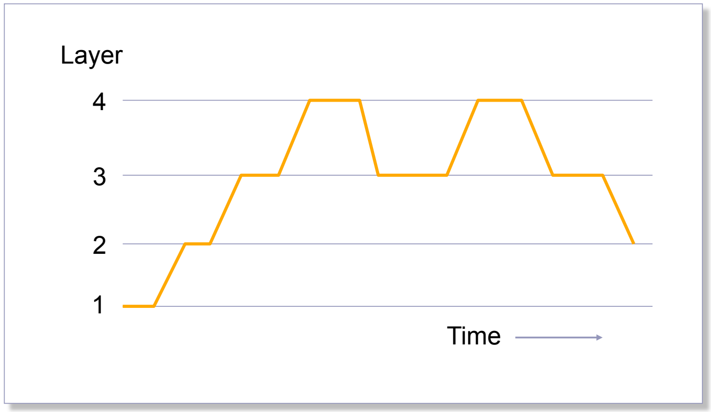
51

RLM Join Experiment

- Receivers periodically try subscribing to higher layer
- If enough capacity, no congestion, no drops → Keep layer (& try next layer)
- If not enough capacity, congestion, drops → Drop layer (& increase time to next retry)
- What about impact on other receivers?

52

Join Experiments



53

RLM Scalability?

- What happens with more receivers?
- Increased frequency of experiments?
  - More likely to conflict (false signals)
  - Network spends more time congested
- Reduce # of experiments per host?
  - Takes longer to converge
- Receivers coordinate to improve behavior

54

Next Lecture

- DDoS and Traceback
- Required reading:
  - A DoS-limiting Network Architecture
- Optional reading:
  - Hash-Based IP Traceback

55