15-744: Computer Networking

L-17 Multicast Reliability and Congestion Control

Multicast Issues

- Reliable transfer
- Congestion control
- Assigned reading
  - [F+97] A Reliable Multicast Framework for Ligh-Weight Sessions and Application Level Framing
  - [MJV96] Receiver-driven Layered Multicast

Overview

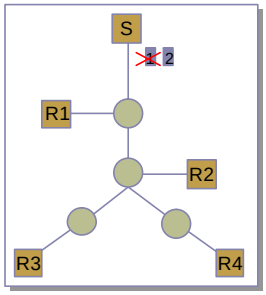
- Reliability
  - Scalable Reliable Multicast
  - Reliable Multicast Transport Protocol
  - Lightweight Multicast Service
- Receiver-driven Layered Multicast
- Other Issues

Loss Recovery

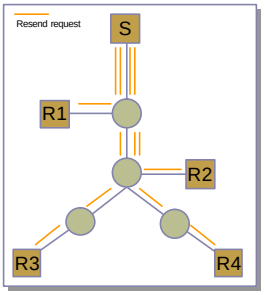
- Sender-reliable
  - Wait for ACKs from all receivers. Re-send on timeout or selective ACK
  - Per receiver state in sender not scalable
  - ACK implosion
- Receiver-reliable
  - Receiver NACKs (resend request) lost packet
  - Does not provide 100% reliability
  - NACK implosion

Implosion

Packet 1 is lost



All 4 receivers request a resend

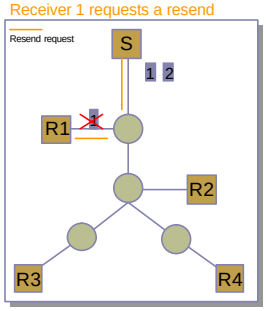


Retransmission

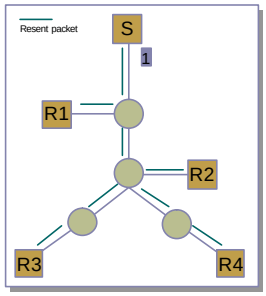
- Re-transmitter
  - Options: sender, other receivers
- How to retransmit
  - Unicast, multicast, scoped multicast, retransmission group, …
- Problem: Exposure

Exposure

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

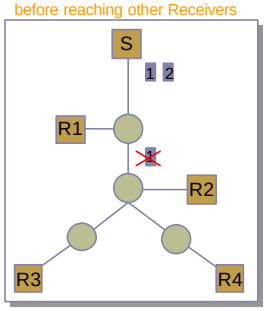


Packet 1 resent to all 4 receivers

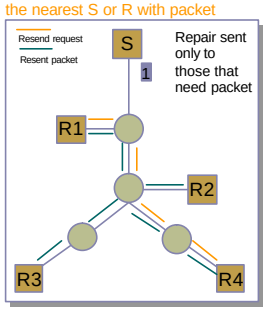


Ideal Recovery Model

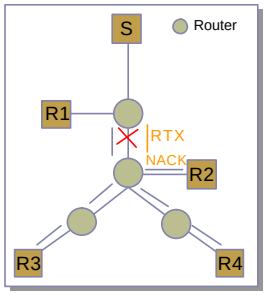
Packet 1 reaches R1 but is lost before reaching other Receivers



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



Aside: Using the Routers



- Routers do transport level processing:
  - Buffer packets
  - Combine ACKs
  - Send retransmissions
- Model solves implosion and exposure, but not scalable
- Violates end-to-end argument

Overview

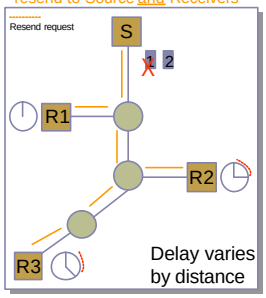
- Reliability
  - Scalable Reliable Multicast
  - Reliable Multicast Transport Protocol
  - Lightweight Multicast Service
- Receiver-driven Layered Multicast
- Other Issues

SRM

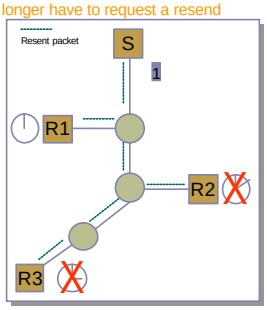
- Originally designed for *wb*
- Receiver-reliable
  - NACK-based
- Every member may multicast NACK or retransmission

SRM Request Suppression

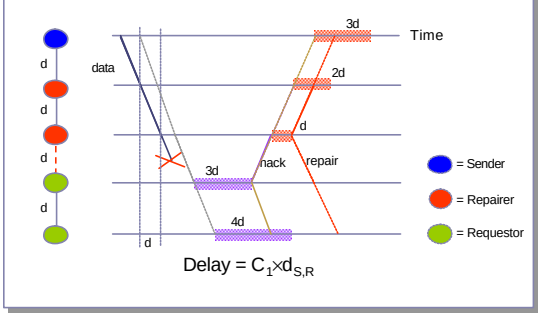
Packet 1 is lost; R1 requests resend to Source and Receivers



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



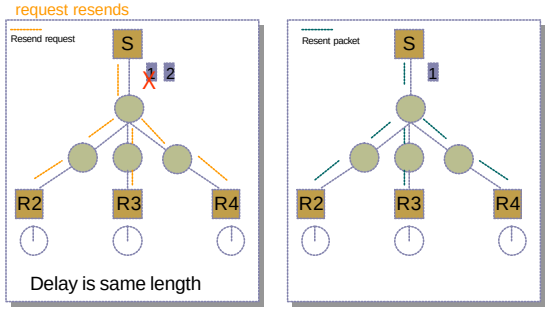
Deterministic Suppression



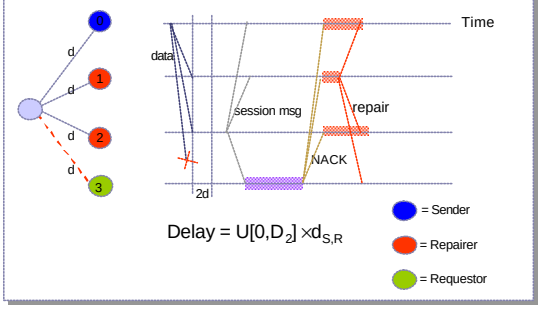
SRM Star Topology

Packet 1 is lost; All Receivers request resends

Packet 1 is resent to all Receivers



SRM: Stochastic Suppression

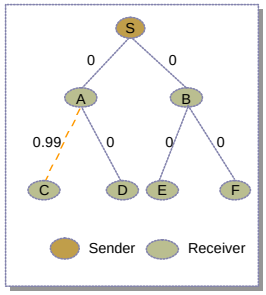


SRM (Summary)

- 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

What's Missing?

- Losses at link (A,C) causes retransmission to the whole group
- Only retransmit to those members who lost the packet
- [Only request from the nearest responder]



Local Recovery

- Application-level hierarchy
  - Fixed v.s. dynamic
- TTL scoped multicast
- Router supported

Overview

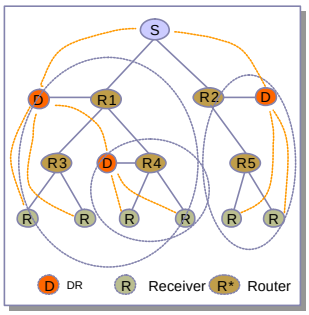
- Reliability
  - Scalable Reliable Multicast
  - Reliable Multicast Transport Protocol
  - Lightweight Multicast Service
- Receiver-driven Layered Multicast
- Other Issues

RMTP

- Reliable Multicast Transport Protocol by Purdue and AT&T Research Labs
- Designed for file dissemination (single-sender)
- Deployed in AT&T's billing network

RMTP: Fixed Hierarchy

- Rcvr unicasts periodic ACK to its Designated Receiver (DR)
- DR unicasts its own ACK to its parent
- Rcvr chooses closest statically configured (DR)
- Mcast or unicast retransmission
  - Based on percentage of requests
  - Scoped mcast for local recovery



RMTP: Comments

- +: Heterogeneity
  - Lossy link or slow receiver will only affect a local region
- –: Position of DR critical
  - Static hierarchy cannot adapt local recovery zone to loss points

Overview

- Pragmatic General Multicast

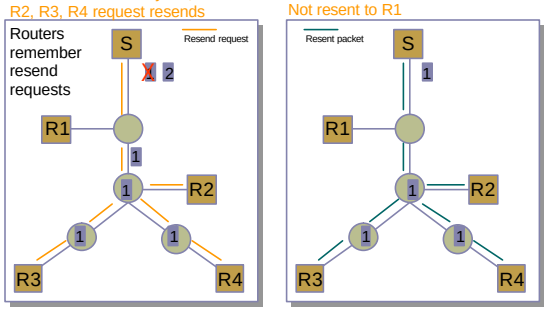
Pragmatic General Multicast

- Cisco's reliable multicast protocol
- NACK-based, with suppression
- Repair only forwarded to the NACKers

Pragmatic General Multicast

Packet 1 reaches only R1; R2, R3, R4 request resends

Packet 1 resent to R2, R3, R4; Not resent to R1



Overview

- Reliability
  - Scalable Reliable Multicast
  - Reliable Multicast Transport Protocol
  - Lightweight Multicast Service
- Receiver-driven Layered Multicast
- Other Issues

Light-weight Multicast Service (LMS)

- Enhance multicast routing with selective forwarding
- LMS extends router forwarding - what routers are meant to do in the first place
- No packet storing or processing at routers
- Strictly IP: no peeking into higher layers

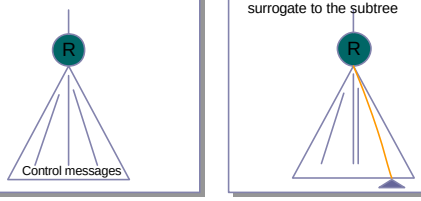
The LMS Concept

Heavy-weight model

- Router stores packets, receives NACKs and sends retransmissions

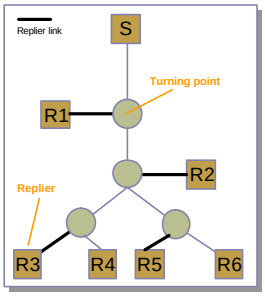
LMS: Receiver acts as surrogate

- Router chooses a receiver as a surrogate
- Router steers all control messages to surrogate
- Router relays messages from surrogate to the subtree



LMS: Definitions

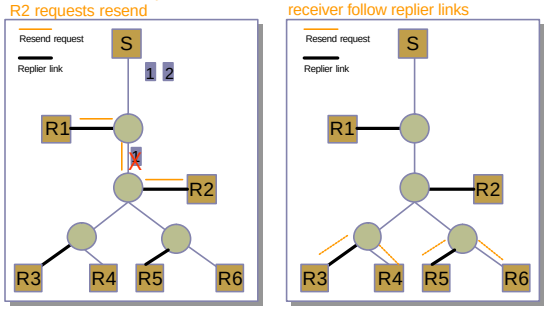
- Replier
  - Receiver volunteered to answer requests
- Turning point
  - Where requests start to move downstream
- Directed mcast
- Mcast to a subtree



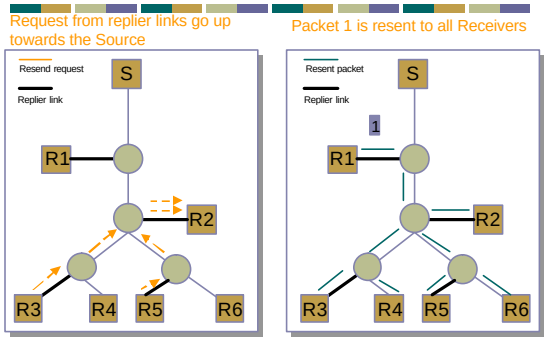
LMS with Replier Links

Packet 1 reaches only R1; R2 requests resend

Resend requests from each receiver follow replier links



LMS with Replier Links



LMS: Comments

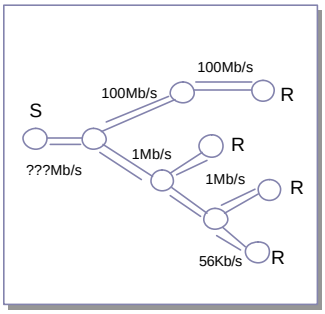
- Replier problems
  - Selection? Fault tolerance?
- Works with unidirectional shared trees (PIM)
  - Needs to relay requests from core/RP to sender
- Difficulties with bi-directional shared trees
  - Needs per-source state

Overview

- Reliability
  - Scalable Reliable Multicast
  - Reliable Multicast Transport Protocol
  - Lightweight Multicast Service
- Receiver-driven Layered Multicast
- Other Issues

Multicast Congestion Control

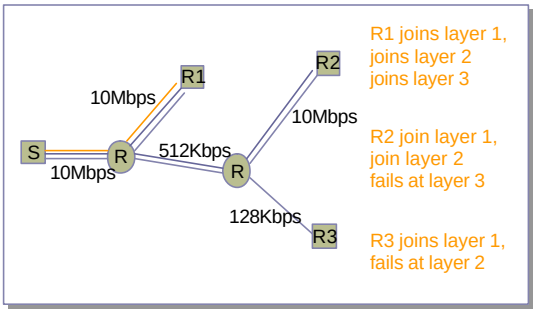
- What if receivers have very different bandwidths?
- Send at max?
- Send at min?
- Send at avg?



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



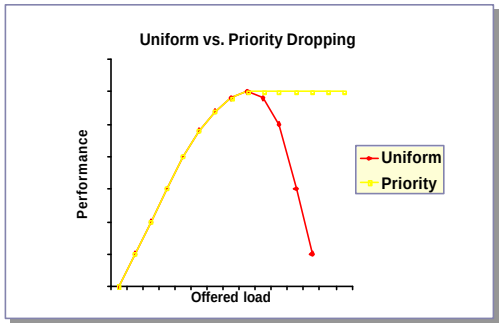
Drop Policies for Layered Multicast

- 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?
  - Intuition vs. reality!

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

RLM Intuition



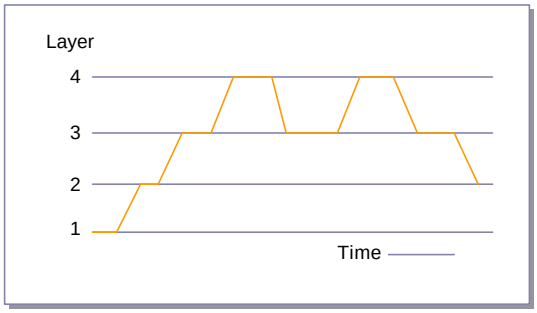
Receiver-Driven Layered Multicast

- Each layer a separate group
  - Receiver subscribes to max group that will get through with minimal drops
- Dynamically adapt to available capacity
  - Use packet losses as congestion signal
- Assume no special router support
  - Packets dropped independently of layer

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?

Join Experiments



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

RLM Receiver Coordination

- Receiver advertises intent to add layer
- Other receivers
  - Avoid conflicting experiments
  - If experiment fails, will see increased drops => don't try adding layer! (shared learning)
  - OK to try adding lower layer during higher layer experiment
    - Won't cause drops at higher layer!

Overview

- Reliability
  - Scalable Reliable Multicast
  - Reliable Multicast Transport Protocol
  - Lightweight Multicast Service
- Receiver-driven Layered Multicast
- Other Issues

Inferring Topology

- What if packet is lost on link?
  - All children of link will not get packet
- Idea: use loss "fingerprints" to identify siblings
  - Siblings will have the most similar fingerprints
- Various techniques to build tree from collection of fingerprints

Session Messages

- SRM
  - Identify what node knows about global state
- Multimedia & other applications
  - Identify list of members
  - Communicate loss rates → possibly for congestion control or other feedback
  - What if it is a large group?
    - Periodic transmissions can flood network!!

Next Lecture: QOS & IntServ

- QOS
- IntServ Architecture
- Assigned reading
  - [She95] Fundamental Design Issues for the Future Internet
  - [CSZ92] Supporting Real-Time Applications in an Integrated Services Packet Network: Architecture and Mechanisms