Reliable Multicast
CS557
Christos Papadopoulos
Introduction: What is Multicast?

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

- Sender-reliable
- Receiver-reliable
Sender-reliable

- Wait for ACKs from all receivers. Re-send on timeout or selective ACK
  - +: easy resource management
  - -: wait for ACK
  - -: receiver state in sender not scalable
  - -: ACK implosion
Receiver-reliable

- Receiver NACKs lost packet
  +: no state at sender - good for mcast
  -: does not provide 100% reliability
  -: NACK implosion
Implosion
Implosion
Implosion
Implosion
Implosion
Implosion
Implosion
Implosion
Implosion
Implosion
Implosion
Implosion
Implosion
Retransmission

- Re-transmitter
  - options: sender, receiver

- How to retransmit
  - unicast, multicast, scoped multicast, retransmission group, ...

- Problem: Exposure
Exposure
Exposure
Exposure
Exposure
Exposure
Exposure
Exposure
Exposure
Exposure
Exposure
Exposure
Exposure
Exposure
Exposure
Exposure
Exposure
Aside: Using the routers

Routers do transport level processing:
• buffer packets
• fuse ACKs
• send retransmissions
• Model solves implosion and exposure, but not scalable.
• Violates end-to-end argument
• Is this still true with NDN?
Ideal recovery model

- “Ideal router”
- A single request is sent upstream
- A single repair is multicast from the nearest repairer to the subtree down the lossy link
SRM

- Receiver-reliable
  - NACK-based
- Every member may multicast NACK or retransmission
SRM Request Suppression
SRM Request Suppression
SRM Request Suppression
SRM Request Suppression
SRM Request Suppression
SRM Request Suppression
SRM Request Suppression
SRM Request Suppression
SRM Request Suppression
SRM Request Suppression
SRM Request Suppression
SRM Request Suppression
Deterministic Suppression

\[ \text{Delay} = C_1 \times d_{S,R} \]
SRM Star Topology
SRM Star Topology
SRM Star Topology
SRM Star Topology
SRM Star Topology
SRM Star Topology
SRM Star Topology
SRM Star Topology
SRM Star Topology
SRM Star Topology
SRM Star Topology
SRM Star Topology
SRM Star Topology
Delay = \( U[0,C_2] \times d_{S,R} \)
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
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
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

- Rcvrs grouped into local regions
- Rcvr unicasts periodic ACK to its ACK Processor (AP), AP unicasts its own ACK to its parent
- Rcvr dynamically chooses closest statically configured Designated Receiver (DR) as its AP
RMTP: Error control

- DR checks retx “request” periodically
- Mcast or unicast retransmission
  - Based on percentage of requests
  - Scoped mcast for local recovery
- Immediate transmission request
  - Used for late join
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
PGM

- Cisco’s reliable multicast protocol
- NACK-based, with suppression
- Repair only forwarded to the NACKers
PGM: Request forwarding

- NACK + random delay
- Forwarded upstream towards the source
- Only one NACK is forwarded for every packet loss
- NCF: NACK suppression and hop-by-hop NACK reliability
PGM Summary
PGM Summary
PGM Summary
PGM Summary

\[ \text{Src} \downarrow (x+1) \]

[Diagram of a PGM model]
PGM Summary
PGM Summary
PGM Summary

![Diagram]

Src

(x?)
PGM Summary
PGM Summary
PGM Summary
PGM Summary
PGM Summary
PGM Summary
PGM Summary
Problem: Repeated Retransmissions
Repeated Retransmissions
Repeated Retransmissions
Repeated Retransmissions
Repeated Retransmissions
Repeated Retransmissions
Repeated Retransmissions
Repeated Retransmissions
Repeated Retransmissions
Repeated Retransmissions
Repeated Retransmissions
Repeated Retransmissions
Repeated Retransmissions
Repeated Retransmissions
Repeated Retransmissions
Repeated Retransmissions
Repeated Retransmissions
Repeated Retransmissions
Repeated Retransmissions
Repeated Retransmissions
Repeated Retransmissions
Repeated Retransmissions
LMS

- Light-weight Multicast Service
- Enhance multicast routing with selective forwarding, nothing beyond that
LMS

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
Router chooses a receiver as a surrogate
Router steers all control messages to surrogate
Router relays messages from surrogate to the subtree
Receiver acting as a surrogate
Core Ideas

- Each router selects a replier (surrogate)
- Routers steer requests to repliers
- Routers help repliers multicast replies to loss subtree

LMS achieves the efficiency of the heavy-weight model, but without the weight
LMS: Concepts

- **Replier**
  - Receiver volunteered to answer requests

- **Turning point**
  - Where requests start to move downstream

- **Directed mcast**
  - Mcast to a subtree
Request Handling
Request from non-replier link
The Turning Point: stamp incoming iface
Out to replier link
Request handling
Request from replier link
Send upstream unchanged
Request handling
Request from upstream
To replier link unchanged

Src

RQ-XXX
LMS: Request forwarding

- Mcast to the group

- If a request reaches a turning point, it’s forwarded towards the replier

- No request suppression or merging, but scope of requests is limited
LMS: Reply forwarding

- At turning point, `<router_addr:link_id>` is added into request packet
- Replier includes it into its retx packet
- The specified router forwards packet downstream at `link_id`
- No retx suppression necessary
LMS Summary

---

DATA
NAK
RTX
LMS Summary
LMS Summary
LMS Summary
LMS Summary
LMS Summary

Source

- DATA
- NAK
- RTX
LMS Summary

- DATA
- NAK
- RTX
LMS Summary

- DATA
- NAK
- RTX
LMS Summary
LMS Summary
LMS Summary

- DATA
- NAK
- RTX
LMS Summary
LMS Summary

DATA
NAK
RTX
LMS Summary

- **DATA**
- **NAK**
- **RTX**
LMS Summary
LMS Summary

- DATA
- NAK
- RTX
LMS Summary
LMS Summary
LMS: Comments

- Replier problems
  - Selection? Fault tolerance?
  - How well will repliers scale w.r.t. |group|?
- 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