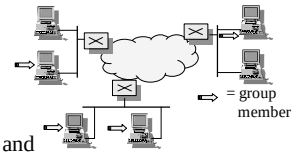
Multi-rate Multicast Congestion Control for the Internet: Challenges, Approaches, and the RLC Algorithm

Sonia Fahmy
Department of Computer Sciences
Purdue University
fahmy@cs.purdue.edu
http://www.cs.purdue.edu/homes/fahmy/

Overview

- IP Multicast
- History and Current Status
- Challenges
- The RLC Algorithm
- Main innovations and open problems

Multicast



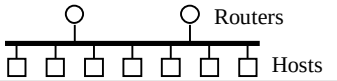
- Multipoint communication = exchange of information among one or more senders and multiple receivers (multicast group)
- Popular applications requiring multipoint support include:
  - conferencing, distance learning, software distribution, searching, server and database synchronization
  - ffnet.com: Architecture for TV-quality streaming
- Complicated by variation in group size and dynamics and bandwidth requirements

IP Multicast: Design Principles

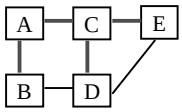
- Single address per group
- Members located anywhere
- Members can join and leave at will
- Senders need not be aware of memberships
  Like a TV channel ⇒ Scalable
- Sender need not be a member
- Soft connections ⇒ periodic renewal

IGMP

- Internet Group Management Protocol
- Used by hosts to report multicast membership
- Join-IP-Multicast Group (address, interface)
- Leave-IP-Multicast Group (address, interface)
- "Leave group" message (version 2) to reduce leave latency
  Sent only if the host that responded to the last query leaves
- Querier then issues a "membership query" with a short response time
- Ref: RFC 1112 (Version 1), RFC 2236 (Version 2), Version 3



Multipoint Routing Algorithms

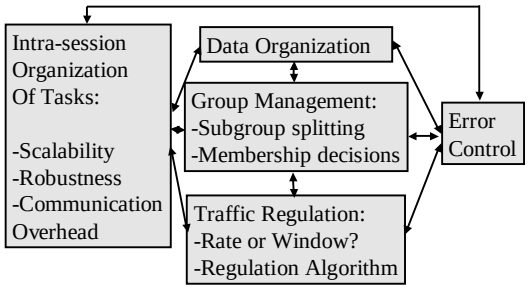


- Flooding
- Spanning Trees
- Reverse Path Forwarding
- Flood and Prune
- Steiner Trees
- Center-Based Trees, e.g., core-based trees

Most routing protocol standards are combination of these algorithms.

Ref: Semeria and Maufer on 3Com page

MCC Components

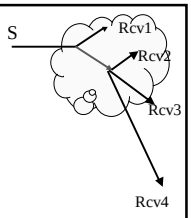


Ref: Golestani, INFOCOM 1999

Multi-rate CC: History and References

- RLM (McCanne): receivers join appropriate layers (groups), shared learning, detection and join timers
- Grouping (GA Tech): e.g., DSG, LVMR
- Sender can use feedback to determine number and rates of layers (UCI/Rutgers: SAMM, AT&T)
- Differential drop: AT&T, Shenker
- Playback and caching: USC/UCLA, Feng at Ohio State
- ---------- RLC ---------- first appeared in INFOCOM 1998
- digitalfountain.com: DLCC, FLID
- See IETF rmt draft [Luby2000], rmt and RMRG minutes, proceedings, and mailing lists, for latest discussions

Goals and Challenges



- TCP friendliness
- Receiver heterogeneity
  - High utilization of resources
  - Low loss and delay
  - Stability and fast transient response
  - Inter-receiver fairness
- Data organization for both bulk data (transfer time) and streaming (quality index)
- Scalability
- Simplicity and generality
- Ease of deployment

TCP Friendliness and Layering

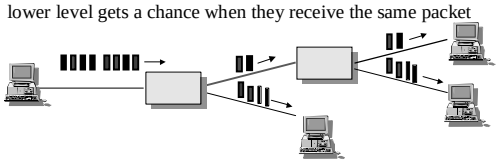
- R = Packet Size × constant / RTT × √loss rate
- Does not consider timeouts, so only holds for small loss rates
- How to mimic TCP? Which RTT?
- Cumulative layering: no duplication, as opposed to simulcast
- Number of layers: small to scale and avoid high join/leave rates

What is RLC?

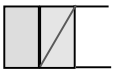
- Based on RLM
- A different layer hopping rule, i.e., when to join a layer, leave a layer, or remain at current level?
  - If loss, decrease
  - If no loss for $t_p(i)$, increase
  - B(i) and $t_p(i)$ double with the increase of i ⇒ TCP friendliness
- Synchronization points to coordinate receivers behind the same bottleneck
- Sender-initiated probes to combat long leave latency after join experiments
- Deaf period to avoid cascading leaves: After leave, do not react to congestion for some time slightly larger than leave delay

Receiver Coordination

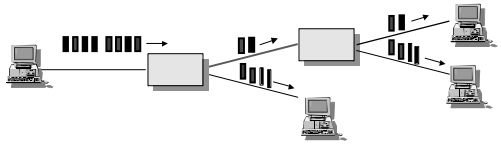
- Avoids mismatch of receivers with the same bottleneck and inter-receiver unfairness
- A receiver can only make a join attempt right after a synchronization point (SP) (special sender packet) based on inter-SP events only
- SPs at each level are always a subset (half) of SPs at previous level. Thus: (1) lower levels get more chances, (2) everyone at same or lower level gets a chance when they receive the same packet



Sender Probes



- Periodically generate short bursts of packets, followed by relaxation period to simulate a join attempt
- Doubles rate for each layer, but is later silent to compensate
- Congestion during burst means do not increase



Performance Evaluation

- Assume one packet lost per cycle (since we use FRR TCP model)
- Average throughput is proportional to 1/√loss rate:
  - No dependence on RTT
  - Rate-based
  - Coarse granularity of rates
  - Minimum rate may be larger than bottleneck bandwidth
- Ns simulations and experiments between Pisa and London with data and streaming applications show that: (1) model for average throughput is valid with RED and slightly high with drop-tail, (2) TCP-friendly (goodputs), (3) fast startup to optimal level with staggered sources
- For more results, see INFOCOM 1998 paper

Key Points

- Main innovations:
  - Synchronization points
  - TCP-friendliness
  - Sender probes
  - Deaf periods
- Questions:
  - RTT effects
  - Rates of layers and number of layers
  - Differential drop
  - Receiver complexity and reliance on well-behaved receivers

Thank You!

Questions?