Reliable Multicast and Layered Multicast

12/3/2001

Review: Application Layer Multicast

- Narada
  - first builds a mesh
  - then runs DVMRP over the mesh

- Overcast
  - directly builds a tree (the non-leaf nodes are servers)
  - when connecting to the tree, a new node probes from leaf to root
Reliable Multicast: Discussion

- What are the goals of reliable multicast?

- Why is reliable multicast difficult?

Scalable Reliable Multicast (SRM) [Floyd et al. '95]

- Randomize NACKs (requests)
- All traffic including requests and repairs are multicast
- A repair can be sent by any node that heard the request
- A node suppresses its request if another node has just sent a request for the same data item
- A node suppresses a repair if another node has just sent a repair
Estimating Distances

- Every node estimates distance (in time) from every other node
  - why?
  - how?

Request Timer

- Chosen from the uniform distribution on
  \[ [C_1 d_{S,A}, (C_1 + C_2) d_{S,A}] \]
  - \( A \) - node that lost the packet
  - \( S \) - source
  - \( C_1, C_2 \) - algorithm parameters
- If \( A \) receives a request before its timeout triggers, it does an exponentially backoff
  \[ 2^{[C_1 d_{S,A}, (C_1 + C_2) d_{S,A}]} \]
Repair Timer

- Chosen from the uniform distribution on $[D_1d_{A,B}, (D_1 + D_2)d_{A,B}]$
  - $A$ - node that lost the packet
  - $B$ - node that sends the repair
  - $D_1, D_2$ - algorithm parameters
- If $B$ receives a repair for missing data it cancels its timer
- $B$ does not verify whether $A$ has received data (why?)

Chain Topology

- $C_1 = D_1 = 1, C_2 = D_2 = 0$ (why?)
- All link distances are 1

![Chain Topology Diagram]
**Star Topology**

- *Choose* $C_1 = D_1 = 0$ (why?)
- *Given* $C_2$, what is the expected number of requests, expected time to receive a repair?

**Bounded Degree Tree**

- Use both
  - deterministic suppression (chain topology)
  - probabilistic suppression (star topology)
- Large $C_2/C_1 \rightarrow$ fewer duplicate requests, but larger repair time
- Large $C_1 \rightarrow$ contribute to suppressing more duplicate requests
- Small $C_1 \rightarrow$ smaller repair time
- Simulation parameters: $C_1 = C_2 = 2$, $D_1 = D_2 = \log_{10} g$
Discussion: Adaptive Timers

- How would you adapt the timers?

Discussion

- How scalable is SRM?

- Some other ways you can think of to provide reliable multicast?
Remaining Issue: How Fast to Send to the Receivers?

- Why is the problem difficult?

Two Multi-rate Schemes

**Destination Set Grouping (DSG)** [JAZ95]
- sender sends to K groups
- sender sends to group i at $g_i$
- each receiver joins one group

**Layered approach** [Sha92, McC96]
- sender sends to K groups
- assume $g_1 < g_2 < \ldots < g_K$
- sender sends at $g_1, g_2 - g_1, \ldots, g_K - g_{K-1}$
- receivers that can receive $g_i$ join layers 1, ..., i
Layered Multicast

- A receiver:
  - if no loss when joins layers 1 to i, try i+1;
  - if high loss rate when joins layers 1 to i, drops layer i to see the effects