

A Digital Fountain Approach to Reliable Distribution of Bulk Data

Scott Weber
CSE 498
06.February.2003
Lehigh University

Problem

- Distribute data to large, changing group
- Must be:
 - Reliable
 - Efficient for both client and server
 - On demand
 - Tolerant of network problems

Ideal Solution

- "Digital Fountain"
 - Like a water fountain, quenches thirst for water (data) no matter which distinct water drops (pieces of data) the user receives
- Uses redundant encoding of k packets
- Client needs only retrieve k distinct packets
- Reliable, efficient, on demand and tolerant

Redundant Encoding

- Reed-Solomon Codes
 - k packets provide all data
 - Slow encode/decode
- Tornado Codes
 - $(k + \epsilon)$ packets provide all data ($\sim 5-7\%$ overhead)
 - Much faster encode/decode (At least 100 times faster)
- Interleaving
 - Overhead can be chosen
 - Better for small total data sizes ($< 200\text{KB}$)

Multicast Layering

- Each layer i sends 2^{i-1} blocks per round
- Clients choose at which layer to participate
 - Server periodically bursts at twice normal rate
 - If client sees no congestion, can move to next layer
- Blocks are scheduled
 - No two layers send same block in same round
 - No one layer repeats a block until all are sent

Implementation

- Uses Tornado coding and multicast layering
- Layering decreases efficiency
 - Hosts switching layers tend to receive more duplicate packets
- Tornado coding nullifies effect of packet loss