TECHNISCHE UNIVERSITÄT DARMSTADT

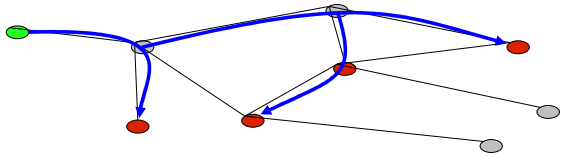
Content Dissemination in Distributed Systems

Wesley Terpstra
Databases and Distributed Systems Group

Motivation

- Providing content costs
  - Per-consumer uploads
  - Bursty demand requires worst-case bandwidth
  - Must remain on-line
- Roaming devices
  - Variable reachability
  - Download sessions are longer than connection
- Security updates
  - Millions of receivers
  - Initial download spike
  - Timely delivery critical
- Popular private publishers
  - Low source bandwidth
  - High consumer demand
  - Often a free service

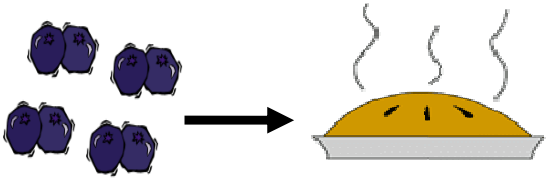
Leverage Multicast Trees



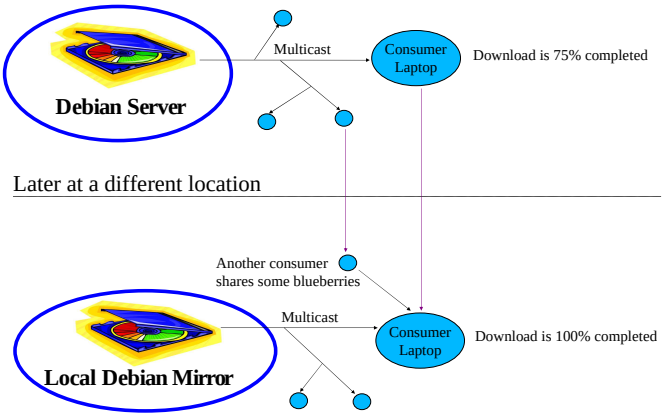
- Content is delivered in stream leading to many destinations
  - The stream is replicated wherever delivery paths branch
  - The source transmits only once even for many consumers
  - Bandwidth is saved inside the network from shared paths
- Multicast provides neither flow control nor reliability

Mobility and Blueberry pies

- In order to bake a blueberry pie
  - collect blueberries whenever a blueberry bush is nearby
  - exchange blueberries with other parties
  - bake the pie when blueberries meet recipe's requirement
- Mobile devices downloading content can
  - locate local (multicast) transmitters
  - exchange packets with other nearby consumers
  - decode any 4GB of packets to recover the 4GB content

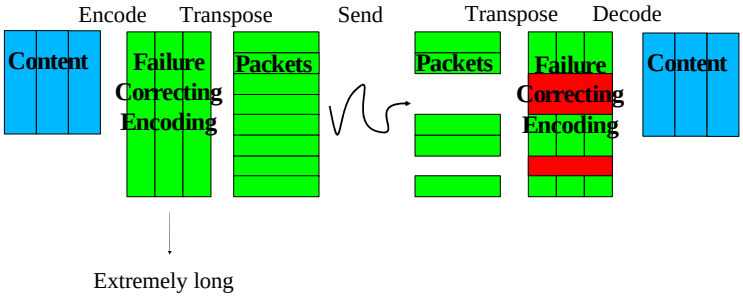


Example Scenario: Debian DVD



Apply Failure Correcting Codes

- Multicast transmission is unreliable (has packet failure)
  - Failure correcting codes can recover lost pieces of content
  - The transmitter need never communicate with individual consumers
  - Enables small publishers to manage popular content
- Mobile clients download from many sources
  - Reassembly of content is the same as decoding a very long code
  - Autonomous transmitters have low probability of repeating data
  - Decoding may be preformed off-line



Use the Number Theoretical Transform

- Interpret columns of the content to be a polynomial
- Encoding consists of evaluating the polynomial at points
- Decoding consists of interpolating to recover the polynomial
- Degree n polynomials are uniquely determined by n points

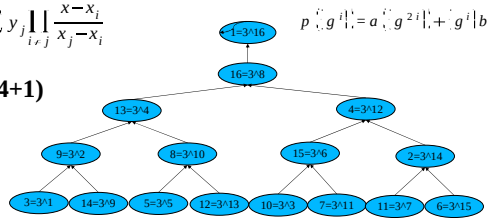
$$y_j = p(x_j) = \sum_{i=0}^{n} a_i x_j^i$$

$$p(x) = \sum_j y_j \prod_{i \neq j} \frac{x - x_i}{x_j - x_i}$$

$$p(x) = \sum_{i=0}^{15} a_i x^i = \sum_{i=0}^{7} a_{2i} x^{2i} + x \sum_{i=0}^{7} a_{2i+1} x^{2i}$$

$$p(x) = g(x^2) + x\,b(x^2)$$

GF(2^4+1)



In practice, GF(3*2^30+1)

Ongoing and Future Research

- Bridging multicast barriers
  - Not all ISPs on backbone
  - Use P2P to merge ISPs
  - Leverage mobility
- Optimize the decoder
  - Use SIMD instructions
  - 64bit prime numbers
  - Use 2D NTT with on disk transpose for large files
- Distributed automatic caches
  - ISP-sniffs traffic and records it
  - Cut off backbone subscription
  - ISP now multicasts locally
- Integrated search
  - Reduce advertising costs
  - Enable free information markets
  - Try to leverage multicast to build efficient search algorithms

Summary

- Enables publishers with poor throughput
- Enables consumers with poor connectivity
- Copes with bursty download behaviour
- Extreme bandwidth savings
- Decouples content from its source
- Assists ISPs with traffic shaping