Delay Tolerant Bulk Data Transfers on the Internet
by N. Laoutaris et al., SIGMETRICS’09

Ilias Giechaskiel

Cambridge University, R212
ig305@cam.ac.uk

March 4, 2014
Conclusions

Takeaway Messages

- Need to transfer multiple terabytes daily
  - Postal system for infrequent transfers
  - Direct transfer for small timezone differences
  - Store and forward otherwise
- Take advantage of off-peak bandwidth through “water-filling”
- Mathematical analysis with cost estimates and deadlines
- No concrete implementation!
Context

Motivation
- 1PB of data every day at CERN
  - 10GB of data transfers every second at peak!
  - http://home.web.cern.ch/about/computing
- Data tolerates delays from several hours to a few days
- Postal system and dedicated networks too expensive
95-percentile Pricing

- Allow 5% of traffic to be burst traffic beyond committed rate
- Charges based on peak rate!
- Lots of bandwidth wasted

Figure: http://www.semaphore.com/blog/2011/04/04/95th-percentile-bandwidth-metering-explained-and-analyzed
The problem

Goals

- Transfer data between data centers without dedicated network
- Avoid increasing 95-percentile cost for sender and receiver
- Avoid impact on QoS of interactive traffic

Approach

- Transmit during off-peak hours of both sender and receiver
  - Directly when centers close-by (E2E)
  - With intermediate storage otherwise (SnF)
- Evaluate using bandwidth costs and estimates
Water-Filling

\[ C - x(t) - \Delta \]

\[ q(x) - x(t) - \epsilon \]
Bulk Transfer Policies

End-to-End with Source Scheduling
- Water-filling that respects sender and receiver charge volumes
- If enough to send for free, just use it!

Store-and-Forward
- Two independent water-fillings
- Send minimum of two and store to or transfer from transit
Implementation

Required Predictions
- Next slot load
  - Successive loads highly correlated
- Total charged volume
  - Use current so far or previous month

Meeting Deadlines
- Not all volumes can be sent for free!
- Use existing approach, but modify cost volumes allowed
  - Polynomial exact search or greedy approximation for min cost
- Need prediction for entire month
  - Use same day of previous week
  - 1-2% worse than actually knowing future values
Evaluation

Methodology

- Data given by large Transit Provider for 2008 Q1
- 448 links with 140 ISPs
- Keep 280 that have $>1\text{ Gbps}$ capacity
- Several are unpaid peerings
- Deadline 1 day
- Repeat for all working days of week
Free Volumes

\[ \frac{F(SnF)}{F(E2E-Sched)} \]

vs

time difference (hours)
E2E vs. SnF

CDF of E2E-Sched cost (in USD thousands)
Storage Costs

- Back of the envelope calculation
- $300/\text{TB storage}$
- Server cost $10,000$
- Server lifetime 2 years$
- Double for maintenance$
- $< 1K$ amortized cost
  - $5K$ median for E2E
  - $100,000s$ for constant-rate bulk without scheduling
27TB/day EU to LAT

CDF
SnF
E2E-Sched
FedEx

cost to send 27 Tbytes in USD thousands

Ilias Giechaskiel ig305@cam.ac.uk
Decision Tree

The DTB job

Can FedEx meet the deadline $T$?

Is the network capacity enough for volume $B$?

Does SnF yield non zero cost

$<v,u,T>$

Yes

Yes

FedEx wins...

No

No

SnF wins!

Yes

Case by case

No

SnF wins!
Criticisms

Pricing Model

- Need model based on peak demand
- Network costs defined by peak traffic
- Change for percentile for all traffic
  - Increase, e.g. to 99%, helps SnF
  - Decrease, e.g. to 50%, punishes non-DTB clients
- Transit ISPs claim part of transfer profit?
- Similar idea with electricity [QWB⁺09]
  - Would undermine relationship agreement
Criticisms

Estimates
- Estimates too rough
- Little data
- Irrelevant data
- No transit bottlenecks modeled

Implementation
- Evaluation too theoretical
- Follow-up work NetStitcher [LSYR11]
  - Introduced more intermediate hops
  - Allowed estimation error correction
- GRESE for specific types of bandwidth elasticity [NP12]
- Jetaway for video traffic [FLL12]
Conclusions

Key Contributions
- Model for free transfers (10-30TB for 10-40Gbps links)
- Simple decision tree choice (SnF usually wins)
- SnF useful in different time-zones with similar capacities
- E2E more expensive by $5K in 50% of cases
- Courier better for occasional transfers

Key Questions
- Would more intermediate hops help?
- Is there no transit bottleneck?
- How can you combine jobs and optimize traffic?
- Will the price model change?
- Your questions?

