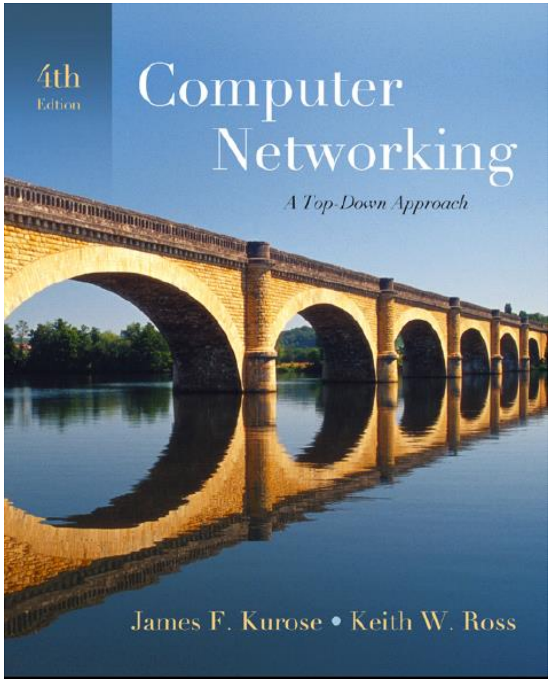
Proportional Fairness



Computer Networking: A Top Down Approach
4^{th} edition.
Jim Kurose, Keith Ross
Addison-Wesley, July 2007.

A note on the use of these ppt slides:

We're making these slides freely available to all (faculty, students, readers). They're in PowerPoint form so you can add, modify, and delete slides (including this one) and slide content to suit your needs. They obviously represent a *lot* of work on our part. In return for use, we only ask the following:

- If you use these slides (e.g., in a class) in substantially unaltered form, that you mention their source (after all, we'd like people to use our book!)
- If you post any slides in substantially unaltered form on a www site, that you note that they are adapted from (or perhaps identical to) our slides, and note our copyright of this material.

Thanks and enjoy! JFK/KWR

2: Optimization-based approach towards congestion control

Resource allocation as optimization problem:

- how to allocate resources (e.g., bandwidth) to optimize some objective function
- maybe not possible that optimality exactly obtained but…
  - optimization framework as means to explicitly steer network towards desirable operating point
  - practical congestion control as distributed asynchronous implementations of optimization algorithm
  - systematic approach towards protocol design

Proportional Fairness

- Vector of rates, $\{X_s\}$, proportionally fair if feasible and for any other feasible vector $\{X_s^*\}$:

$$\sum_{s \in S} \frac{x_s^* - x_s}{x_s} \le 0$$

- Optimal Solution is derived by

$$\max \sum \log(X_s)$$

subject to capacity constraints

- GDS: $$\sum_{s \in S} \frac{x_s^* - x_s}{x_s} = \sum_{s \in S} \frac{\Delta x_s}{x_s} \approx \sum_{s \in S} \Delta(\log x_s) \le 0$$

Implies that $\sum \log(X_s)$ is the maximum possible - Also, this allocation is "fair" in the sense that in order to deviate from it, the total percentage of deviations has to be non-positive, meaning that small flows have to "suffer" relatively more

Optimization Problem

$$\max_{x_s \geq 0} \quad \sum_s U_s(x_s)$$

$$\text{subject to} \quad \sum_{s \in S(l)} x_s \leq c_l, \forall l \in L$$

"system" problem

- maximize system utility (note: all sources equal)
- constraint: bandwidth used less than capacity
- centralized solution to optimization impractical
  - must know all utility functions
  - impractical for large number of sources
  - we'll see: congestion control as distributed asynchronous algorithms to solve this problem

The user view

- user can choose amount to pay per unit time, w_s
- Would like allocated bandwidth, x_s in proportion to w_s

$$x_s = \frac{w_s}{p_s}$$

- p_s could be viewed as charge per unit flow for user s

$$\max \quad U_s\left(\frac{w_s}{p_s}\right) - w_s$$

$$\text{subject to} \quad w_s \geq 0$$

user problem

user's utility

cost

The network view

- suppose network knows vector $\{w_s\}$, chosen by users
- network wants to maximize logarithmic utility function

$$\max_{x_s \geq 0} \quad \sum_s w_s \log x_s$$

network problem

$$\text{subject to} \quad \sum_{s \in S(l)} x_s \leq c_l$$

Solution existence

- There exist prices, p_s, source rates, x_s, and amount-to-pay-per-unit-time, $w_s = p_s x_s$ such that
  - $\{W_s\}$ solves user problem
  - $\{X_s\}$ solves the network problem
  - $\{X_s\}$ is the unique solution to the system problem
  - p_s = Sum of links' shadow prices over the path of s

$$\max \quad U_s\left(\frac{w_s}{p_s}\right) - w_s$$

$$\text{subject to} \quad w_s \geq 0$$

$$\max_{x_s \geq 0} \quad \sum_s w_s \log x_s$$

$$\text{subject to} \quad \sum_{s \in S(l)} x_s \leq c_l$$

$$\max_{x_s \geq 0} \quad \sum_s U_s(x_s)$$

$$\text{subject to} \quad \sum_{s \in S(l)} x_s \leq c_l, \forall l \in L$$

Proportional Fairness Property

- If w_r=1, then $\{X_s\}$ solves the network problem IFF it is proportionally fair
- Related results exist for the case that w_r not equal 1.
- GDS: Solution of the weighted proportional fairness for the derived w_r is an additional property of the socially optimal solution

Solving the network problem

- Results so far: *existence* - solution exists with given properties
- How to *compute* solution?
  - ideally: distributed solution easily embodied in protocol
  - insight into existing protocol

Outline of Approach (GDS):

- Tatonnement process in order to reach equilibrium prices in bandwidth market:
  - Increase (resp. decrease) bandwidth per user if his willingness-to-pay buys him more (resp. less) with the present prices
    - How?? → next slide
  - Increase (resp. decrease) unit-bandwidth price per link if total allocated bandwidth is lower (resp. higher) than link capacity
    - How?? → next slide

Solving the network problem

$$\frac{d}{dt}x_s(t) = k\left(w_s - x_s(t)\sum_{l\in L(s)} p_l(t)\right)$$

change in bandwidth allocation at s — linear increase — multiplicative decrease

where $p_l(t) = g_l\left(\sum_{l\in L(s)} x_s(t)\right)$

congestion "signal": function of aggregate rate at link l, fed back to s.

Solving the network problem

$$\frac{d}{dt}x_s(t) = k\left(w_s - x_s(t)\sum_{l \in L(s)} p_l(t)\right)$$

- Results:
  - converges to solution of "relaxation" of network problem
  - $x_s(t)\Sigma p_l(t)$ converges to w_s
- Interpretation: TCP-like algorithm to iteratively solve optimal rate allocation!
- GDS: In TCP the two phases alternate, while in the above mechanism they are both applied at the same time

Optimization-based congestion control: summary

- bandwidth allocation as optimization problem:
- practical congestion control (TCP) as distributed asynchronous implementations of optimization algorithm
- optimization framework as means to explicitly steer network towards desirable operating point
- systematic approach towards protocol design