6.896 Lecture 5: Congestion Control Principles

Hari Balakrishnan

September 23, 1998

1 Outline

- **Readings: CJ89**
- Finish Paxson's routing behavior paper (from last class)
- Congestion management: problems and approaches.
- Notions: power, control, avoidance, rate-based, window-based, efficiency, fairness, distributedness, convergence, stability.
- Analysis of linear controls.

2 Congestion management: the problem

- Resource contention: what resource(s)? Caused by (i) overload (ii) heterogeneity. How's this different from flow control?
- Why is this different from resource management in a standard OS?
- Throughput vs. load curve. Draw pic of topology and draw pic of curve. Show knee, plateau and congestion collapse cliff. Ask questions to explain regions of curve.
- Why is this important? Else low (zero) throughput for everyone. Congestion collapse (increasing load, diminishing throughputs). (congestion collapse in mid-1980s in ARPANET).
- Why is this hard? Distributed (incomplete knowledge), changing conditions and traffic, want stability/convergence, want fairness. Want high *power.*

3 Notions

- **Bottleneck:** Draw picture showing a bottleneck link and delays.
- **Power** = Throughput/Delay. Should this be large or small? Ideally, if you had perfect knowledge of the network, what would you make this be for a connection?
- **Efficiency:** want to be as close to available bandwidth as possible.
- **Fairness:** *max-min* definition [Gafni & Bertsekas 1984]: an allocation of a resource is fair if:

  1. no user receives more than its request, ρ_i.

2. no other allocations satisfying (1) has higher minimum allocation
3. condition (2) recursively holds as we remove the minimal user and reduce the total resource accordingly.

- This condition reduces to $\mu_i = MIN(\mu_{fair}, \rho_i)$, where μ_i is the allocation for user i and ρ_i is user i's request. Assumes all users have equal rights to resource.
- **Jain's index:** $(\sum x_i(t))^2/(n \sum x_i(t)^2)$.
- **Distributedness:** no centralized agent. Scales a lot better.
- **Convergence:** minimize oscillations and move towards fairness.
- **Congestion control:** react to congestion (reactive control). Key question: how to decrease rate of transmission? How to carefully probe for bandwidth?
- **Congestion avoidance:** proactive control. Anticipate it and slow down before it happens (or at least before losses happen). E.g., delay-based (CARD, Vegas), marking-based (DECbit, ECN). Pros and cons? Sometimes need router support. Delay-based schemes often don't work well because of variable delays.
- **Window-based:** Use a (sliding) window; corresponding rate is roughly window/rtt.
- **Rate-based:** No window to modulate, at best an upper bound flow control window. Clock packets out using timers. Key point: need router support in some form to preserve rate. (Packet-pair + FQ; explicit rate enforcement at router,...). Also, rate is really 2 parameters, not 1—burst size and burst interval!
- **Router support** for congestion control: FQ, reservations, RED/ECN, backpressure-based hop-by-hop schemes, incentives/disincentives for sources needs to be looked at closely.
- Will hardware advances (faster links, routers,...) solve congestion in the long run, so this is a non-problem?
- NO! Actually could make the problem worse! (diverse apps, more flows, increased heterogeneity). Lots of small flows make the problem a lot harder.
- **Multicast congestion control:** very hard in general, area of active research.
- **Pricing** to discourage bad behavior. Open problems here.
- **Multiple schemes:** a key problem to be understood and solved. Why is this harder? Today, only TCP does any adaptive congestion control.
- Congestion management still has many open and hard problems!

4 Announcements

- Classroom change: 66-168. Yes/No? Where is it?!
- Working on project suggestions. Should have ready by end of week, have about 8-9 days to pick partners and project. Proposal due 10/2 (1 page).

- HW 1 will be assigned end of next week, will be due in about 3 weeks.
- Any questions?

5 Linear Controls

- What form of control? Two aspects: probe for more bandwidth, slow down if you think the network's overloaded.
- In theory, simple enough strategy. But there are some key questions.
  - Congestion signal?
  - How to probe?
  - How to slow down?
  - And remember, want to be efficient, fair, distributed and not oscillate all over the place.
- Signal either from router notification or (in TCP today) from loss of packets.
- Literally an infinite solution space! For starters, look at *linear controls* (from control theory).
- $x_i(t+1) = a_I + b_I x_i(t)$ if no congestion signal, and $x_i(t+1) = a_D + b_D x_i(t)$ if congested. Obviously this is linear because the a's and b's are constants.
- Phase plot pictures. In general, n dimensions. **Ask questions for each!** $X1 + X2 = 1$. Equifairness lines through origin. What is underutilization region? Overload? Optimal point? Digress about proxy hosts, differently priced flows etc.
- Explain in picture what the different controls do and what the good ones are. Additive/multiplicative increase/decrease on picture. **Again ask questions.**
- Efficiency: Means that $\sum x_i(t)$ increases with t as long as there's no congestion. This means that $b_I > 1 - na_I / \sum x_i(t), \forall n, \forall \sum x_i(t)$. Similarly, $b_D < 1 - na_D / \sum x_i(t), \forall n, \forall \sum x_i(t)$. On picture, this goes on 45-degree line through current allocation (assuming distributedness).
- Fairness: Must never decrease. Since $f(x(t)) \leq 1, f(x(t+1)) > f(x(t))$ implies convergence. This yields $a/b \geq 0$. I.e., they must be of the same sign. Show calculation on board. Leads to $a_I \geq 0, b_I \geq 0, a_D \geq 0, 0 \leq b_D < 1$. If pure multiplicative decrease, no change in fairness. Additive increase improves fairness. Additive decrease generally reduces it.
- Distributedness: Means $x_{i+1}(t) > x_i(t)$ when not loaded. Along with previous conditions, this leads to $a_I > 0, b_I \geq 1; a_D = 0, 0 \leq b_D < 1$. Remember, these are *sufficiency* conditions (not necessity). (Assumes no truncation of bandwidth).
- As it turns out, $a_I \geq 1$ and $b_I \geq 1$. What's the best value of b_I? Why? (quickest convergence to fairness).
- Pictorially, ask students to explain why only one of the four schemes works for all values of $x_i(t)$ and n. Reemphasize linear control space.
- Upon congestion, reduce load but don't compromise fairness (pure multiplicative decrease). Try to improve fairness while probing. Neat!

5.1 Are loss rates a sufficient (or necessary) statistic to study?

- Derive simplistic TCP congestion avoidance equation relating throughput to loss probability. Sawtooth pattern of TCP congestion avoidance.
- Is the $1/\sqrt(p)$ relationship sacred?! Can argue both ways... (open question).
- Notion of "TCP-friendliness." How fundamental is this?
- Timeouts and other realistic effects worked out in painstaking(!) detail in Padhye et al. (Sigcomm 98); verifies that the asymptotic relationship holds.
- Possible project: work out details for web workloads and mixes of web & longer connections. Work to date analyzes long connections.
- Possible project: work out set of sufficient statistics that completely characterize the behavior of a congestion management scheme.

6 Summary

- **Theory:** With non-synchronized sources, different round-trip times etc., can't achieve perfect fairness without router support (Jaffe, Shenker).
- Congestion arises from overload and is exacerbated by heterogeneity.
- Control schemes are typically source-based, sometimes use router support.
- Multiplicative decrease and additive increase is a stable, efficient and fair (under many assumptions) scheme for linear controls. Key insight is to use phase pictures for understanding the math.
- Next class: TCP variants, DECbit. On 9/30, queue management schemes that handle congestion better.