Title:
Stability Criteria and Applications for Randomized Load Balancing Schemes

Author(s):
Maury Bramson, University of Minnesota

Abstract:

In this talk, we consider randomized load balancing schemes where an arriving job joins the shortest of d randomly chosen queues from among a pool of n queues. Vvekenskaya, Dobrushin and Karpelevich (1996) considered the case with Poisson input and exponentially distributed service times, and derived an explicit formula for the equilibrium distribution for fixed d as n goes to infinity. Since its tail decays doubly exponentially fast, this distribution is useful in various applications.

Relatively little work has been done for general service times or input. For general service times, the behavior of the service rule at each queue will now play a role in the behavior of the system. In particular, the question of under which conditions the system is stable (i.e., its underlying Markov process is positive recurrent) for fixed n is no longer obvious. Ideally, one would like to understand the limiting behavior of such equilibria (provided they exist) as n goes to infinity, as in the first paragraph.

Here, we discuss results which show that for fixed n such systems are always stable for the appropriate notion of traffic intensity. These results also show that the associated equilibria are tight when restricted to a finite number of queues and hence subsequential limits exist as n goes to infinity. It is anticipated that this behavior will provide a general framework for examining the behavior of such limits under different service rules. In this context, we discuss joint work with Y. Lu and B. Prabhakar on the limiting behavior of the equilibria when service at each queue is given by the standard first-in, first-out rule.