EE384Y Project Proposal

Load Balancing and Switch Scheduling

Xiangheng Liu

Department of Electrical Engineering
Stanford University, Stanford CA 94305
Email: liuxh@systems.stanford.edu

Load balancing and switch scheduling are two important algorithms in the effort to maximize the stability region and minimize (average) latency. For a queuing system, load balancing tries to regulate the traffic to conform to the service rate while the switch scheduling tries to allocate the service rates adaptive to the arrival patterns of the queues. Existing algorithms on load balancing often assume simple service disciplines like independent exponential services. Switch scheduling algorithms are often tested with arbitrary arrivals. One would certainly conjecture that the combination of the two would give a better performance than a load balancing or switch scheduling algorithm alone. An example of this is the two-stage router proposed by C.S. Chang [1], where the first stage performs load balancing such that the arrivals destined to the same output are distributed evenly among the inputs at the second stage.

There are many similarities in the algorithms that have already been designed for load balancing and switch scheduling. For example, randomized algorithms combined with the use of memory has been proven stable in both load balancing and scheduling. In this project, we will first explore the similarities and possibly duality in designing load balancing and switch scheduling algorithms. This exercise, hopefully, will benefit the research in both communities. Secondly, we would like to explore a general framework of formulating the two problems into one joint problem with the objective of maximizing the stability region and reducing the latency of the VOQs. We plan to adopt the switch architecture as in the two-stage load balanced switch proposed by C.S. Chang. Different load balancing algorithms give different steady-state distribution of the VOQs at the second stage. We are interested in finding the best distribution for delay minimization. We can also fix a scheduling

algorithm and ask what is the best load balancing algorithm. If the optimal solution is found to be elusive, we will try to propose some good heuristic algorithms that combine existing knowledge on load balancing and switch scheduling. Some simulations are expected to compare the delay performance.

The time-line of this project is roughly as follows:

- Apr 16 to Apr 25 - Literature Survey
- Apr 26 to May 9 - Similarity and Duality in the design of the load balancing and switch scheduling algorithms
- May 14 - Progress Report due
- May 10 to May 30 - Formulating the joint problem of load balancing and switch scheduling; Simulating the two-stage switch with existing load balancing and switch scheduling algorithms.
- May 31 to June 5 - Final Report

References

[1] C.S. Chang, D.S. Lee, Y.S Jou, "Load balanced Birkhoff-von Neumann Switches" *IEEE Workshop on High Performance Switching and Routing,* 2001.