

Department of Industrial and Systems Engineering

“Energy-Aware Scheduling on Heterogeneous Processors”

Joint work with Osman Akgun and Doug Down

Monday, January 28, 2013

1:50 – 2:40

4002 Emerging Technologies Building



Abstract:

ENERGY consumption has been growing exponentially in computers and computer centers, embedded systems, portable devices, etc., and it is also a critical concern in battery-operated devices such as sensors. Indeed, embedded systems are being designed to have heterogeneous processing elements specifically to reduce power consumption. Energy cost has become the critical cost factor for server farms, and there is also increasing concern over the heat and carbon emissions generated. No longer is faster necessarily better; both speed and energy consumption must be considered in the operation of these systems, and often faster servers require more energy. Hence the emphasis is on providing good service while maintaining energy efficiency. We study a multiple-server system where servers are heterogeneous in terms of both their speeds and their (energy) usage costs, and determine the optimal dynamic control policy to minimize total expected waiting and usage costs.

By considering a dual system, in which jobs are given priorities and make individual decisions to minimize their own costs, we show that the optimal control policy is threshold type. That is, servers are ordered in a particular preference order, and there is a threshold for each server such that if the current number of jobs exceeds the threshold the server is assigned a job. We consider both the case with (preemptive) reassignment, or job migration, where jobs can be removed from a server and assigned to a different server or placed in the queue at any time, and the case where reassignment is not possible, so jobs once assigned to a server cannot be preempted. In both cases, determining the server preference order is surprisingly difficult, and depends on the arrival rate.

Bio:

Rhonda Righter is a Professor and past Chair of the Department of Industrial Engineering and Operations Research at the University of California, Berkeley. Before coming to Berkeley she taught at the Leavey School of Business at Santa Clara University. Her PhD is in Industrial Engineering and Operations Research from UC Berkeley, her BS is in applied math and business from Carnegie Mellon. Her primary research and teaching interests are in the general area of stochastic modeling and optimization, especially as applied to service, manufacturing, computer communication, and grid computing systems. She is an associate editor for the Journal of Scheduling. She formerly served on the editorial boards of Management Science, Operations Research, Operations Research Letters, and Queueing Systems. She is the past (founding) Chair of the Applied Probability Society of INFORMS.