

Course Number: CSC 641/841
Course Title: Computer Performance Evaluation
Number of Credits: 3
Schedule: Three hours of lecture per week.
Prerequisite: CSC 415 or consent of instructor

Catalog Description

Computer performance analysis problems related to system design, selection, and tuning. Modeling using stochastic and operational queuing models. Workload characterization, design, and performance measurement methods. Design of simulation models of computer systems. Paired with CSC 841. Students who have completed CSC 641 may not take CSC 841 for credit.

Expanded Description

Computer performance analysis is required at any stage in the life cycle of a computer system, including *system design*, *system selection/procurement*, and *system use/tuning*. The first problem is to predict the values of performance indicators of a computer system during its design and development. Then, in the area of system comparison and selection the problem is to evaluate and compare the performance of existing competitive systems, which are assumed to be available for performance measurements. In the area of system management the problem is to improve the performance level of existing operational computer systems. The main techniques for computer performance analysis include (1) *analytic modeling* (queuing theory, and operational analysis), (2) *simulation* (the use of specialized simulators, and general purpose simulation systems and languages), and (3) *measurement* (benchmarking based on natural and synthetic workloads). This course includes the presentation of all main techniques for computer performance analysis. Individual topics include the following:

- Performance of basic hardware and software components of computer systems (performance models for processors, memory, disks, tapes, peripherals, compilers, operating system, and application software)
- Using queues and servers for modeling dynamic behavior of computer systems (batch processing, processor scheduling, interactive systems, networked client server systems)
- Simulation of queuing models (machine dependent and machine independent random number generators, writing discrete event simulators of queuing systems)
- Stochastic queuing models. Poisson arrival process, exponential birth-death model, and single server models (M/M/1, M/D/1, MG/1, and GI/G/1). Single queue models M/M/2 and M/M/k.
- Operational analysis. Proof of Little's law. Operational laws and conditions (job flow balance, one-step behavior, homogeneous external arrivals, device homogeneity, homogeneous service times, routing homogeneity). Operational aspects of the birth-death model.

- Modeling open queuing networks. Serial, parallel, and cyclic queues. Open models of computer systems. The problem of load balancing.
- Analytic models of closed queuing networks. Mean value analysis models. Machine repairman model. Analysis of multiprocessor interactive systems. Models of computer networks.
- Workload characterization and benchmarking. Natural and synthetic workloads. Standard benchmark programs. SPEC benchmarks. Software performance monitors.
- System tuning and selection. Finding and eliminating bottlenecks. Load balancing. Computer system evaluation, comparison, and selection.

Course Objectives and Role in Program

The objectives of this course include:

- Develop detailed understanding of dynamic behavior of computer systems.
- Present performance characteristics of all major components of computer systems.
- Provide analytic models for analysis of dynamic phenomena in computer systems and networks.
- Expose students to all major methods for analysis and prediction of computer performance.
- Teach the relationships between theoretical performance models and practical performance problems encountered by system administrators and performance managers.
- Explain and exemplify limits of accuracy of mathematical models that describe the performance of computer systems.
- Present a performance approach to design concepts implemented in computer architecture, and in operating systems.
- Provide background to advanced work in a graduate program.

Learning Outcomes

At the end of this course students will

- Be able to identify performance problems in the operation of computer systems.
- Utilize measurement tools to identify and eliminate bottlenecks, monitor load distribution, perform load balancing, and improve performance
- Be able to participate in computer capacity planning projects
- Be prepared to participate in computer benchmarking and selection projects
- Write discrete event simulators
- Use basic analytic performance analysis models to predict computer performance and participate in system sizing projects

Method of Evaluation

Student learning will be evaluated on the basis of

- Completeness and quality of project assignments (20%)

- Grade on midterm examination (30%)
- Grade on final examination (50%)

The total number of points is used for relative ranking of students. Letter grades are assigned taking into account both the ranking and the total score.

Recommended Literature

Textbooks: R. Jain, *The Art of Computer Systems Performance Analysis*. J. Wiley, 1991. [Recommended]
Menasce-Almeida-Dowdy, *Capacity Planning and Performance Modeling*. Prentice Hall, 1994.
Menasce, D., and V. Almeida, *Capacity Planning for Web Performance*. Prentice Hall, 1998.
Menasce, D. and V. Almeida, *Scaling for E-Business*. Prentice Hall, 2000.
A.O. Allen, *Probability, Statistics, and Queuing Theory with Computer Science Applications*. Academic Press, 1990.

Reader: J.J. Dujmović, *CSc841 Reader: Computer Performance Evaluation*. SFSU, 2003.

Journals: ACM SIGMETRICS Performance Evaluation Review, CMG Transactions, Performance Evaluation, IEEE Transactions on Software Engineering.

Conferences: ACM SIGMETRICS, IEEE MASCOTS, CMG, WOSP (see Proceedings of these conferences)

Prepared by: Jozo Dujmović

Last Revision Date: September 1, 2006