CSE 567M Computer Systems Analysis

Raj Jain
Washington University in Saint Louis
Saint Louis, MO 63130
Jain@cse.wustl.edu

These slides are available on-line at:

http://www.cse.wustl.edu/~jain/cse567-11/

Overview

- Goal of this Course
- Contents of the course
- Tentative Schedule
- Project
- Grading

Goal of This Course

- Comprehensive course on performance analysis
- Includes measurement, statistical modeling, experimental design, simulation, and queuing theory
- How to avoid common mistakes in performance analysis
- Graduate course: (Advanced Topics)
  $\Rightarrow$ Lot of independent reading and writing
  $\Rightarrow$ Project/Survey paper (Research techniques)

Text Book

- R. Jain, "Art of Computer Systems Performance Analysis," Wiley, 1991, ISBN:0471503363 (Winner of the "1992 Best Computer Systems Book" Award from Computer Press Association")

Objectives: What You Will Learn

- Specifying performance requirements
- Evaluating design alternatives
- Comparing two or more systems
- Determining the optimal value of a parameter (system tuning)
- Finding the performance bottleneck (bottleneck identification)
- Characterizing the load on the system (workload characterization)
- Determining the number and sizes of components (capacity planning)
- Predicting the performance at future loads (forecasting).

Basic Terms

- **System:** Any collection of hardware, software, and firmware
- **Metrics:** Criteria used to evaluate the performance of the system. components.
- **Workloads:** The requests made by the users of the system.

Main Parts of the Course

- Part I: An Overview of Performance Evaluation
- Part II: Measurement Techniques and Tools
- Part III: Probability Theory and Statistics
- Part IV: Experimental Design and Analysis
- Part V: Simulation
- Part VI: Queueing Theory

Part I: An Overview of Performance Evaluation

- Introduction
- Common Mistakes and How To Avoid Them
- Selection of Techniques and Metrics

Example I

- What performance metrics should be used to compare the performance of the following systems:
  - Two disk drives?
  - Two transaction-processing systems?
  - Two packet-retransmission algorithms?

Part II: Measurement Techniques and Tools

- Types of Workloads
- Popular Benchmarks
- The Art of Workload Selection
- Workload Characterization Techniques
- Monitors
- Accounting Logs
- Monitoring Distributed Systems
- Load Drivers
- Capacity Planning
- The Art of Data Presentation
- Ratio Games

Example II

- Which type of monitor (software or hardware) would be more suitable for measuring each of the following quantities:
  - Number of Instructions executed by a processor?
  - Degree of multiprogramming on a timesharing system?
  - Response time of packets on a network?

Part III: Probability Theory and Statistics

- Probability and Statistics Concepts
- Four Important Distributions
- Summarizing Measured Data By a Single Number
- Summarizing The Variability Of Measured Data
- Graphical Methods to Determine Distributions of Measured Data
- Sample Statistics
- Confidence Interval
- Comparing Two Alternatives
- Measures of Relationship
- Simple Linear Regression Models
- Multiple Linear Regression Models
- Other Regression Models

Example III

- The number of packets lost on two links was measured for four file sizes as shown below:

File Size	Link A	Link B
1000	5	10
1200	7	3
1300	3	0
50	0	1

Which link is better?

Part IV: Experimental Design and Analysis

- Introduction to Experimental Design
- 2^k Factorial Designs
- 2^kr Factorial Designs with Replications
- 2^{k-p} Fractional Factorial Designs
- One Factor Experiments
- Two Factors Full Factorial Design without Replications
- Two Factors Full Factorial Design with Replications
- General Full Factorial Designs With *k* Factors

Example IV

- The performance of a system depends on the following three factors:
  - Garbage collection technique used: G1, G2, or none.
  - Type of workload: editing, computing, or AI.
  - Type of CPU: C1, C2, or C3.

How many experiments are needed? How does one estimate the performance impact of each factor?

Part V: Simulation

- Introduction to Simulation
- Types of Simulations
- Model Verification and Validation
- Analysis of Simulation Results
- Random-Number Generation
- Testing Random-Number Generators
- Random-Variate Generation
- Commonly Used Distributions

Example V

- In order to compare the performance of two cache replacement algorithms:
  - What type of simulation model should be used?
  - How long should the simulation be run?
  - What can be done to get the same accuracy with a shorter run?
  - How can one decide if the random-number generator in the simulation is a good generator?

Part VI: Queueing Theory

- Introduction to Queueing Theory
- Analysis of A Single Queue
- Queueing Networks
- Operational Laws
- Mean Value Analysis and Related Techniques
- Convolution Algorithm
- Advanced Techniques

Example VI

- The average response time of a database system is three seconds. During a one-minute observation interval, the idle time on the system was ten seconds.

Using a queueing model for the system, determine the following:

- System utilization
- Average service time per query
- Number of queries completed during the observation interval
- Average number of jobs in the system
- Probability of number of jobs in the system being greater than 10
- 90-percentile response time
- 90-percentile waiting time

The Art of Performance Evaluation

- Given the same data, two analysts may interpret them differently.

Example:

- The throughputs of two systems A and B in transactions per second is as follows:

System	Workload 1	Workload 2
A	20	10
B	10	20

Possible Solutions

- Compare the average:

System	Workload 1	Workload 2	Average
A	20	10	15
B	10	20	15

Conclusion: The two systems are equally good.

- Compare the ratio with system B as the base

System	Workload 1	Workload 2	Average
A	2	0.5	1.25
B	1	1	1

Conclusion: System A is better than B.

Solutions (Cont)

- Compare the ratio with system A as the base

System	Workload 1	Workload 2	Average
A	1	1	1
B	0.5	2	1.25

Conclusion: System B is better than A.

- Similar games in: Selection of workload, Measuring the systems, Presenting the results.
- Common mistakes will also be discussed.

Grading

- Exams (Best of 2 mid terms + Final) 60%
- Class participation 5%
- Homeworks 15%
- Project 20%

Prerequisites

- CSE 131: Computer Science I
- CSE 126: Introduction To Computer Programming
- CSE 260M: ~~Introduction To Digital Logic And Computer Design~~ (Not required)
- Basic Probability and Statistics

Prerequisite

- Statistics:
  - Mean, variance
  - Normal distribution
  - Density function, Distribution function
  - Coefficient of variation
    Correlation coefficient
  - Median, mode, Quantile
- Programming

Tentative Schedule

Date	Topic	Chapter
1/19	Course Introduction	
1/24	Common Mistakes	2
1/26	Selection of Techniques and Metrics, Types of Workloads, Workload Selection	3, 4, 5
1/31	Workload Characterization	6
2/2	Data Presentation, Ratio Games	10,11
2/7	Summarizing Measured Data	12
2/21	Comparing Systems Using Random Data	13
2/23	**Mid-Term Exam 1**	

Tentative Schedule (Cont)

Date	Topic	Chapter
2/28	Simple Linear Regression Models	14
3/2	Other Regression Models	15
3/7	Experimental Designs	16
3/9	Factorial Designs with Replication	18
3/14	Fractional Factorial Designs	19
3/16	One Factor Experiments	20
3/21	Two Factor Full Factorial Designs with Replications	22
3/23	General Full Factorial Designs	23
3/28	**Mid-Term Exam 2**	

Tentative Schedule (Cont)

Date	Topic	Chapter
3/30	Introduction to Queueing Theory	30
4/4	Analysis of Single Queue, Queueing Networks	31, 32
4/6	Operational Laws	33
4/18	Analysis of Simulation Results	25
4/20	Random Number Generation	26
4/27	Random Variate Generation	28
5/2	**Final Exam**	
5/4	Final Grades	

Projects

- A survey paper on a performance topic
  - Workloads/Metrics/Analysis: Databases, Networks, Computer Systems, Web Servers, Graphics, Sensors, Distributed Systems
  - Comparison of Measurement, Modeling, Simulation, Analysis Tools: NS2
  - Comprehensive Survey: Technical Papers, Industry Standards, Products
- A real case study on performance of a system you are already working on
- Average 6 Hrs/week/person on project + 9 Hrs/week/person on class
- Recent Developments: Last 5 to 10 years ⇒ Not in books
- Better ones may be submitted to magazines or journals

Projects (Cont)

- Develop a software monitor to observe the performance of a large multiprocessor system.
- Analyze the performance of a distributed game program running on a network of artificial intelligence systems.
- Compare the performance of several robot control algorithms.

- **Goal:** Provide an insight (or information) not obvious before the project.
- **Real Problems:** Thesis work, or job
- **Homeworks:** Apply techniques learnt to your system.

Example of Previous Case Studies

- Database Systems Performance Evaluation Techniques
- A Survey of Hardware Performance Analysis Tools
- Analytical Modeling of Beyond Visual Range Air Combat
- A Survey of Network Simulation Tools: Current Status and Future Development
- Performance Analysis of Data Encryption Algorithms
- Case Study of the Performance of a Gamma-Ray Event Parameterization Application
- Survey on Performance Analysis of MAC protocols
- A Summary of Network Traffic Monitoring and Analysis Techniques
- A Survey of Network Performance Monitoring Tools

Case Studies (Cont)

- Survey of Network Performance Monitoring Tools
- SNMP and Beyond: A Survey of Network Performance Monitoring Tools
- A Survey of Network Traffic Monitoring and Analysis Tools
- Operating System and Process Monitoring Tools
- A Survey of Performance Analysis Tools
- Processor workloads
- Case Study: Performance Analysis of a Diversified Router
- Performance Analysis of Wireless Sensor Networks
- An Overview of Software Performance Analysis Tools and Techniques: From GProf to DTrace
- Survey of Software Monitoring and Profiling Tools

Case Studies (Cont)

- From Poisson Processes to Self-Similarity: a Survey of Network Traffic Models A Historical View of Network Traffic Models
- A Survey of Network Traffic Models
- Verification and Validation of X-Sim: A Trace-Based Simulator
- A Performance Model for a Thermally Adaptive Application Implemented in Reconfigurable HW

Project Schedule

Mon 3/07	Topic Selection
Mon 3/14	References Due
Mon 3/21	Outline Due
Mon 4/11	First Draft Due -> Peer reviewed
Mon 4/18	Reviews Returned
Mon 4/25	Final Report Due

Office Hours

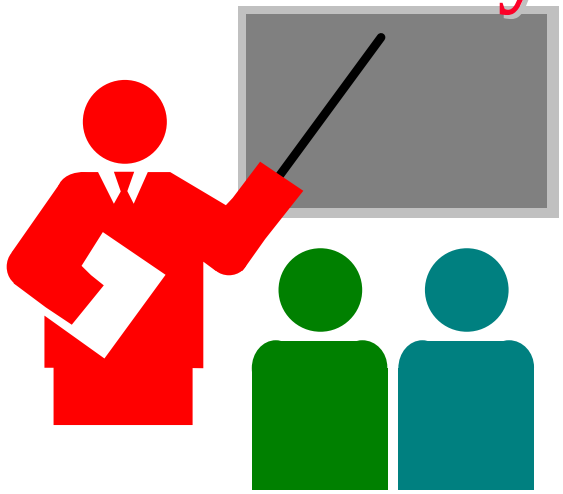
- Monday/Wednesday:11 AM to 12 noon
- Office: Bryan 523

- Teaching Assistant:
  - Michael Hall, Bryan 405G, mhall24@wustl.edu
  - Office Hours: Thursday/Friday 2PM-3PM

Frequently Asked Questions

- Yes, I do use "curve". Your grade depends upon the performance of the rest of the class.
- All homeworks are due on the following Monday unless specified otherwise.
- Any late submissions, if allowed, will *always* have a penalty.
- One 8.4x11 sheet allowed in the exam. Book not allowed. Time limited.
- Exams consist of numerical as well as multiple-choice (true-false) questions.
- There is negative grading on incorrect multiple-choice questions. Grade: +1 for correct. -1/(n-1) for incorrect.
- Everyone including the graduating students are graded the same way.

Summary



- Goal: To prepare you for correct analysis and modeling of any system
- There will be a self-reading and writing
- Get ready to work hard

Quiz 0: Prerequisites

True or False?

T F

❑ ❑ The mean of a uniform(0,1) variate 1.

❑ ❑ The sum of two normal variates with means 4 and 3 has a mean of 7.

❑ ❑ The probability of a fair coin coming up head once and tail once in two throws is 1.

❑ ❑ The density function f(x) approaches 1 as x approaches ∞.

❑ ❑ Given two variables, the variable with higher median also has a higher mean.

❑ ❑ The probability of a fair coin coming up heads twice in a row is 1/4.

❑ ❑ The difference of two normal variates with means 4 and 3 has a mean of 4/3.

❑ ❑ The cumulative distribution function F(x) approaches 1 as x approaches ∞.

❑ ❑ High coefficient of variation implies a low variance and vice versa.

❑ ❑ If x is 0, then after x++, x will be 1.

Marks = Correct Answers _____ - Incorrect Answers _____ = _____