Computer Systems Performance Analysis

Raj Jain
Washington University in Saint Louis
Jain@acm.org
Indo-US Collaboration on Engineering Education
Leadership Institute 2010

Overview

- Goal of this Course
- Contents of the course
- Tentative Schedule
- Projects

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
- Can be taught as:
  - Senior Undergraduate Courses on
    - Performance measurement
    - Statistical Performance Analysis
    - Design of Experiments
    - Simulation
    - Queueing Theory
  - Graduate course:
    $\Rightarrow$ Add a 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 Students 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
- $2^k r$ 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.

Prerequisite

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

Schedule for a 14-Week Semester

Class	Topic	Chapters
1	Introduction	
2	Common Mistakes	2
3	Introduction to Simulation	24
4	Single Queues	30
5	Selection of Techniques and Metrics	2,3
6	Workload Selection and Characterization	5, 6
7	Data Presentation, Ratio Games	10, 11
8	Summarizing Measured Data	12
9	**Exam 1**	
10	Comparing Systems Using Random Data	13
11	Simple Linear Regression Models	14
12	Other Regression Models	15

Tentative Schedule (Cont)

Date	Topic	Chapters
13	Analysis of Simulation Results	25
14	Random Number Generation	26
15	Test Random Numbers	27
16	Experimental Designs	16, 17
17	Factorial Designs with Replication	18
18	Fractional Factorial Designs	19
19	**Exam 2**	
20	One Factor Experiments	20
21	Two Factor Full Factorial Designs	21, 22
22	General Full Factorial Designs	23
23	Analysis of Single Queue	31

Tentative Schedule (Cont)

Date	Topic	Chapters
24	Queueing Networks	32
25	Operational Laws	33
26	Mean-Value Analysis	34
27	Convolution Algorithm	35
28	**Final Exam**	

Graduate Student 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 student is already working on
- Recent Developments: Last 5 to 10 years ⇒ Not in books
- Better ones may be submitted to magazines or journals

Example of Previous Case Studies

- Measure the performance of a remote procedure call mechanism used in a distributed system.
- Measure and compare the performance of window systems of two artificial intelligence systems.
- Simulate and compare the performance of two processor interconnection networks.
- Measure and analyze the performance of two microprocessors.
- Characterize the workload of a campus timesharing system.
- Compute the effects of various factors and their interactions on the performance of two text-formatting programs.
- Measure and analyze the performance of a distributed information system.

Case Studies (Cont)

- Simulate the communications controllers for an intelligent terminal system.
- Measure and analyze the performance of a computer-aided design tool.
- Measure and identify the factors that affect the performance of an experimental garbage collection algorithm.
- Measure and compare the performance of remote procedure calls and remote pipe calls.
- Analyze the effect of factors that impact the performance of two RISC processor architectures.
- Analyze the performance of a parallel compiler running on a multiprocessor system.

Projects (Cont)

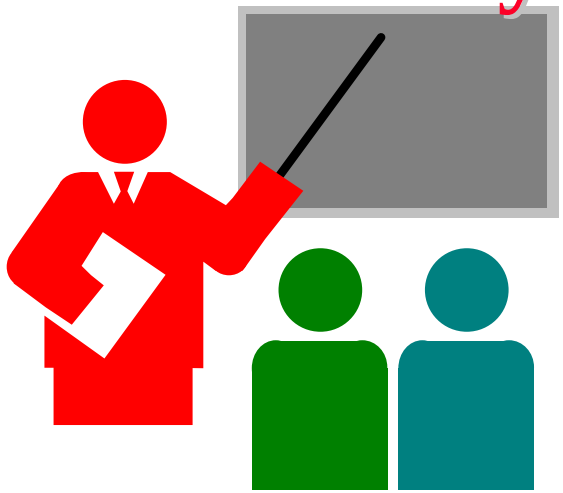
- Develop a software monitor to observe the performance of a large multiprocessor system.
- Analyze the performance of a distributed game program running on a network of artificial intelligence systems.
- Compare the performance of several robot control algorithms.

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

Project Schedule

Class 11	Topic Selection
Class 15	References Due
Class 17	Outline Due
Class 21	First Draft Due
Class 23	Reviews Returned
Class 25	Final Report Due

Summary



- Goal: To prepare students for correct analysis and modeling of any system
- Use measurement, modeling, simulation techniques
- Design experiments for the maximum information with the least effort

Quiz 0: Prerequisites

True or False?

T F

❑ ❑ The sum of two normal variates is normal.

❑ ❑ The sum of two normal variates with means 4 and 3 has a mean of 12.

❑ ❑ 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 ∞.

❑ ❑ 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.

Marks = Correct Answers _____ - Incorrect Answers _____ = _____