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)
  ⇒ Lot of independent reading and writing
  ⇒ Project/Survey paper (Research techniques)
Text Book

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:

<table>
<thead>
<tr>
<th>File Size</th>
<th>Link A</th>
<th>Link B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>1200</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>1300</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>50</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>System</th>
<th>Workload 1</th>
<th>Workload 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>B</td>
<td>10</td>
<td>20</td>
</tr>
</tbody>
</table>
Possible Solutions

- Compare the average:

<table>
<thead>
<tr>
<th>System</th>
<th>Workload 1</th>
<th>Workload 2</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>20</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>B</td>
<td>10</td>
<td>20</td>
<td>15</td>
</tr>
</tbody>
</table>

Conclusion: The two systems are equally good.

- Compare the ratio with system B as the base

<table>
<thead>
<tr>
<th>System</th>
<th>Workload 1</th>
<th>Workload 2</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td>0.5</td>
<td>1.25</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Conclusion: System A is better than B.
Solutions (Cont)

- Compare the ratio with system A as the base

<table>
<thead>
<tr>
<th>System</th>
<th>Workload 1</th>
<th>Workload 2</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>0.5</td>
<td>2</td>
<td>1.25</td>
</tr>
</tbody>
</table>

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
- Matrix multiplication and inversion
Prerequisite

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

- **Programming**
# Tentative Schedule

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Chapter</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/14/2013</td>
<td>Course Introduction</td>
<td></td>
</tr>
<tr>
<td>1/16/2013</td>
<td>Common Mistakes</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Selection of Techniques and Metrics</td>
<td>3</td>
</tr>
<tr>
<td>1/21/2013</td>
<td>No Class: Martin Luther King Holiday</td>
<td></td>
</tr>
<tr>
<td>1/23/2013</td>
<td>Selection of Techniques and Metrics (Continued)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Types of Workloads</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Workload Selection</td>
<td>5</td>
</tr>
<tr>
<td>1/28/2013</td>
<td>Workload Selection (Cont)</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Data Presentation</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Ratio Games</td>
<td>11</td>
</tr>
<tr>
<td>1/30/2013</td>
<td>Summarizing Measured Data</td>
<td>12</td>
</tr>
<tr>
<td>2/4/2013</td>
<td>Summarizing Measured Data (Cont)</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Comparing Systems Using Random Data</td>
<td>13</td>
</tr>
<tr>
<td>2/6/2013</td>
<td>Comparing Systems Using Random Data (Cont)</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Simple Linear Regression Models</td>
<td>14</td>
</tr>
<tr>
<td>2/11/2013</td>
<td>Simple Linear Regression Models (Cont)</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Other Regression Models</td>
<td>15</td>
</tr>
<tr>
<td>2/13/2013</td>
<td>Other Regression Models (Cont)</td>
<td>15</td>
</tr>
<tr>
<td>2/18/2013</td>
<td>Mid-Term Exam 1</td>
<td></td>
</tr>
</tbody>
</table>
## Tentative Schedule (Cont)

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Chapter</th>
</tr>
</thead>
<tbody>
<tr>
<td>2/20/2013</td>
<td>Other Regression Models (Cont)</td>
<td>15</td>
</tr>
<tr>
<td>2/25/2013</td>
<td>Experimental Designs</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>2k Experimental Designs</td>
<td>17</td>
</tr>
<tr>
<td>2/27/2013</td>
<td>Factorial Designs with Replication</td>
<td>18</td>
</tr>
<tr>
<td>3/4/2013</td>
<td>Factorial Designs with Replication</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Fractional Factorial Designs</td>
<td>19</td>
</tr>
<tr>
<td>3/6/2013</td>
<td>One Factor Experiments</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Two Factor Full Factorial Design w/o Replications</td>
<td>21</td>
</tr>
<tr>
<td>3/11/2013</td>
<td>WUSTL Spring Break (No Classes)</td>
<td></td>
</tr>
<tr>
<td>3/13/2013</td>
<td>WUSTL Spring Break (No Classes)</td>
<td></td>
</tr>
<tr>
<td>3/18/2013</td>
<td>Two Factor Full Factorial Designs with Replications</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>General Full Factorial Designs</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Introduction to Queueing Theory</td>
<td>30</td>
</tr>
<tr>
<td>3/20/2013</td>
<td>Introduction to Queueing Theory (Cont)</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Analysis of Single Queue</td>
<td>31</td>
</tr>
<tr>
<td>3/25/2013</td>
<td>Mid-Term Exam 2</td>
<td></td>
</tr>
</tbody>
</table>

Washington University in St. Louis

[http://www.cse.wustl.edu/~jain/cse567-13/k_01int.htm](http://www.cse.wustl.edu/~jain/cse567-13/k_01int.htm) ©2013 Raj Jain
### Tentative Schedule (Cont)

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Chapter</th>
</tr>
</thead>
<tbody>
<tr>
<td>3/27/2013</td>
<td>Queueing Networks</td>
<td>32</td>
</tr>
<tr>
<td>4/1/2013</td>
<td>Operational Laws</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Mean-Value Analysis</td>
<td>34</td>
</tr>
<tr>
<td>4/3/2013</td>
<td>Mean-Value Analysis (Cont)</td>
<td>34</td>
</tr>
<tr>
<td>4/8/2013</td>
<td>Introduction to Simulation</td>
<td>24</td>
</tr>
<tr>
<td>4/10/2013</td>
<td>Introduction to Simulation (Cont)</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Analysis of Simulation Results</td>
<td>25</td>
</tr>
<tr>
<td>4/15/2013</td>
<td>Analysis of Simulation Results (Cont)</td>
<td>25</td>
</tr>
<tr>
<td>4/17/2013</td>
<td>Random Number Generation</td>
<td>26</td>
</tr>
<tr>
<td>4/22/2013</td>
<td>Random Number Generation (Cont)</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Random Variate Generation</td>
<td>28</td>
</tr>
<tr>
<td><strong>4/24/2013</strong></td>
<td><strong>Final Exam</strong></td>
<td></td>
</tr>
</tbody>
</table>

- Note final exam is in the last class before the reading period.
Projects

- A survey paper on a performance topic
  - 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 2 to 4 years ⇒ Not in books
- Better ones may be submitted to magazines or journals
Projects (Cont)

- **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

- Performance of Google App Engine and Amazon Web Service
- Availability and Sensitivity of Smart Grid Components
- Modeling and Analysis Issues in x86-based Hypervisors
- Image Sensor Performance
- Performance of Solving Laplace's Equation using Auto-Pipe
- Performance Modeling of Multi-core Processors
- Performance of Named Data Networking
- A Measurement Study of Packet Reception using Linux
- Performance Analysis of Robotics Systems
- Performance and Measurement Issues of Smart Phones Design
- Analysis of Online Social Networks
- Measurement Study on the BitTorrent File Distribution System
- A Survey of Wireless Sensor Network Simulation Tools
Case Studies (Cont)

- 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 2/25**  Topic Selection
- **Mon 3/04**  References Due
- **Mon 3/18**  Outline Due
- **Mon 4/03**  First Draft Due -> Peer reviewed
- **Mon 4/10**  Reviews Returned
- **Mon 4/17**  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.
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 uniform(0,1) variates is 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 _____ = ______