CPE 702 Computer Performance Evaluation

Contents of the course

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)

1-1

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).

1-3

Main Parts of the Course

Basic Terms

1-2

- 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.

1-4

Part I: An Overview of Performance Evaluation

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

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

Example I

• What performance metrics should be used to compare the performance of the following systems:

> Two disk drives?

quantities:

system?

- > 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

1-7

Example II

Use Which type of monitor (software or hardware) would

be more suitable for measuring each of the following

> Number of Instructions executed by a processor?

> Degree of multiprogramming on a timesharing

> Response time of packets on a network?

Part III: Probability Theory and Statistics

1-8

- 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

1-10

Example III

1-9

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

1-12

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?

1-13

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

1-14

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?

1-15

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

1-16

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
 - \succ Probability of number of jobs in the system being greater than 10
 - 90-percentile response time
 - > 90-percentile waiting time

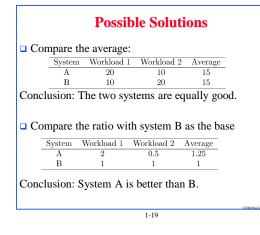
The Art of Performance Evaluation

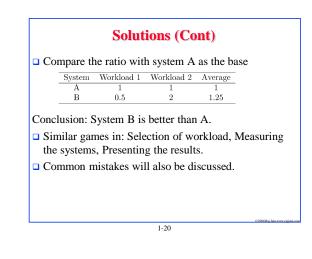
Given the same data, two analysts may interpret them differently.

Example:

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

System	Workload 1	Workload 2
A	20	10
В	10	20





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
- □ Recent Developments: Last 5 to 10 years \Rightarrow Not in books
- Better ones may be submitted to magazines or journals

1-21

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.

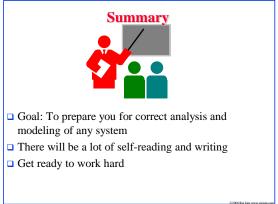
1-22

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.

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.



1-25