# Objectives

P erformance Analysis = Analysis + Computer Systems
Performance Analyst = Mathematician + Computer Systems
Persons
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: the criteria used to evaluate the performance of the system. components.
- Workloads: The requests made by the users of the system.

# Part I: An Overview of Performance Evaluation

1. Introduction
2. Common Mistakes and How To Avoid Them
3. Selection of Techniques and Metrics

# Example I

W hat performance metrics should be used to compare the performance of the following systems:

1. Two disk drives?
2. Two transaction-processing systems?
3. Two packet-retransmission algorithms?

# Part II: Measurement Techniques and Tools

1. Types of Workloads
2. Popular Benchmarks
3. The Art of Workload Selection
4. Workload Characterization Techniques
5. Monitors
6. Accounting Logs
7. Monitoring Distributed Systems
8. Load Drivers
9. Capacity Planning
10. The Art of Data Presentation
11. Ratio Games

# Example II

W hich type of monitor (software or hardware) would be more suitable for measuring each of the following quantities:

1. Number of Instructions executed by a processor?
2. Degree of multiprogramming on a timesharing system?
3. Response time of packets on a network?

# Part III: Probability Theory and Statistics

1. Probability and Statistics Concepts
2. Four Important Distributions
3. Summarizing Measured Data By a Single Number
4. Summarizing The Variability Of Measured Data
5. Graphical Methods to Determine Distributions of Measured Data
6. Sample Statistics
7. Confidence Interval
8. Comparing Two Alternatives
9. Measures of Relationship
10. Simple Linear Regression Models
11. Multiple Linear Regression Models
12. Other Regression Models

# Example III

T he 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

1. Introduction to Experimental Design
2. $2^k$ Factorial Designs
3. $2^k r$ Factorial Designs with Replications
4. $2^{k-p}$ Fractional Factorial Designs
5. One Factor Experiments
6. Two Factors Full Factorial Design without Replications
7. Two Factors Full Factorial Design with Replications
8. General Full Factorial Designs With $k$ Factors

# Example IV

T he performance of a system depends on the following three factors:

1. Garbage collection technique used: G1, G2, or none.
2. Type of workload: editing, computing, or AI.
3. 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

1. Introduction to Simulation
2. Types of Simulations
3. Model Verification and Validation
4. Analysis of Simulation Results
5. Random-Number Generation
6. Testing Random-Number Generators
7. Random-Variate Generation
8. Commonly Used Distributions

# Example V

In order to compare the performance of two cache replacement algorithms:

1. What type of simulation model should be used?
2. How long should the simulation be run?
3. What can be done to get the same accuracy with a shorter run?
4. How can one decide if the random-number generator in the simulation is a good generator?

# Part VI: Queueing Theory

1. Introduction to Queueing Theory
2. Analysis of A Single Queue
3. Queueing Networks
4. Operational Laws
5. Mean Value Analysis and Related Techniques
6. Convolution Algorithm
7. Advanced Techniques

# Example VI

T he 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:

1. System utilization
2. Average service time per query
3. Number of queries completed during the observation interval
4. Average number of jobs in the system
5. Probability of number of jobs in the system being greater than 10
6. 90-percentile response time
7. 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 |

# Three Solutions

1. Compare the average:

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

Cnclusion: The two systems are equally good.

2. 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.

3. 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.

# Professional Organizations

P erformance Evaluation = Evalution + Systems

- ACM SIGMETRICS
- ACM SIGSIM
- IEEE Computer Society: Technical committee on simulation.
- CMG: The Computer Measurement Group, Inc.
  - Practical uses of computer performance
  - Regional groups in the United States and abroad.
  - International regional groups in Australia, Canada, Japan, United Kingdom, Belgium, West Germany, France, and Italy.

- IFIP Working Group 7.3
  - IFIP=International Federation for Information Processing
  - Multinational federation of technical societies
  - American Federation of Information Processing Societies (AFIPS) represents the United States.
  - ACM, IEEE, etc are members of AFIPS
  - Several technical committees (TCs) and working groups (WGs)
  - WG 7.3 is devoted to computer systems modeling

# Professional Organizations (Continued)

- The Society for Computer Simulation (SCS)
  - Publishes simulation books, journals
  - Regional councils in the United States, Canada and the United Kingdom.
  - Sponsors:
    * Summer Computer Simulation Conference
    * SCS Multiconference
    * Eastern Simulation Conferences
    * Winter Simulation Conference
- The Society for Industrial and Applied Mathematics (SIAM)
  - Development of new mathematical techniques

- Journals:
  - SIAM Journal on Control and Optimization
  - SIAM Journal on Numerical Analysis
  - SIAM Journal on Computing
  - SIAM Journal on Scientific and Statistical Computing
  - Theory of Probability and Its Applications.

# Professional Organizations (Continued)

- The Operations Research Society of America (ORSA)
  - Interests:
    - Operation research
    - Queueing theory
    - Game theory
    - Network analysis
    - Replacement and inventory theories
    - Scheduling
    - Simulation
  - Geographical sections, technical sections, special interest groups, and student sections.
  - Represents the US in the International Federation of Operational Research Societies (INFORS)

- Journals:
  - Operations Research
  - ORSA Journal on Computing
  - Mathematics of Operations
  - Research (jointly with TIMS)
  - Operations Research Letters
  - Stochastic Models
- Address: ORSA Business Office, Mount Royal & Guilford Avenues, Baltimore, MD 21202.

# Conferences

- SIGMETRICS: Published as Performance Evaluation Review
- CMG
- PERFORMANCE (18 months)
  Alternately in Europe and in North America.
- EDP Performance and Capacity Management
  Applied Computer Research, Inc.
- ACM SIGSIM and IEEE Computer Society Technical Committee
- Pittsburgh Conference on Modeling and Simulation:
  University of Pittsburgh's School of Engineering and IEEE

# Journals

1. Performance Evaluation Review: Quarterly by ACM SIGMETRICS.
2. CMG Transactions: Quarterly by CMG.
3. Performance Evaluation: Twice a year by Elsevier Science Publishers B. V. (North-Holland)
4. EDP Performance Review: Monthly by Applied Computer Research, Inc. Survey in the annual reference issue.
5. Simulation: Monthly by SCS.
6. Transactions of the Society for Computer Simulation: Quarterly
7. Simulation Digest: Quarterly by ACM SIGSIM and IEEE Computer Society Technical Committee on Simulation.

8. SIAM Review: Quarterly by SIAM.
9. Operations Research: Bimonthly by ORSA.
10. See also computer systems journals.
    - Communications of the ACM.
    - IEEE Transactions on Software Engineering
    - IEEE Transactions on Computers
    - ACM Transactions on Computers

# Related Courses

Later you may consider taking courses on:

- Statistical inference operations research
- Stochastic processes
- Decision theory
- Time series analysis
- Design of experiments
- System simulation
- Queueing theory

# Performance Projects

- Select a system:
  - Network mail program
  - Operating system
  - Language compiler
  - Text editor
  - Processor
  - Database
- Do:
  - Perform measurements
  - Analyze the collected data
  - simulate
  - analytically model the subsystem

# Example of Projects

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

8. Simulate the communications controllers for an intelligent terminal system.
9. Measure and analyze the performance of a computer-aided design tool.
10. Measure and identify the factors that affect the performance of an experimental garbage collection algorithm.
11. Measure and compare the performance of remote procedure calls and remote pipe calls.
12. Analyze the effect of factors that impact the performance of two RISC processor architectures.
13. Analyze the performance of a parallel compiler running on a multiprocessor system.
14. Develop a software monitor to observe the performance of a large multiprocessor system.
15. Analyze the performance of a distributed game program running on a network of artificial intelligence systems.

16. 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.

# Exercises

**1.1** The measured throughput in queries per second for two database systems on two different workloads is as follows:

| System | Workload 1 | Workload 2 |
|---|---|---|
| A | 30 | 10 |
| B | 10 | 30 |

Compare the performance of the two systems and show that:

a. System A is better

b. System B is better