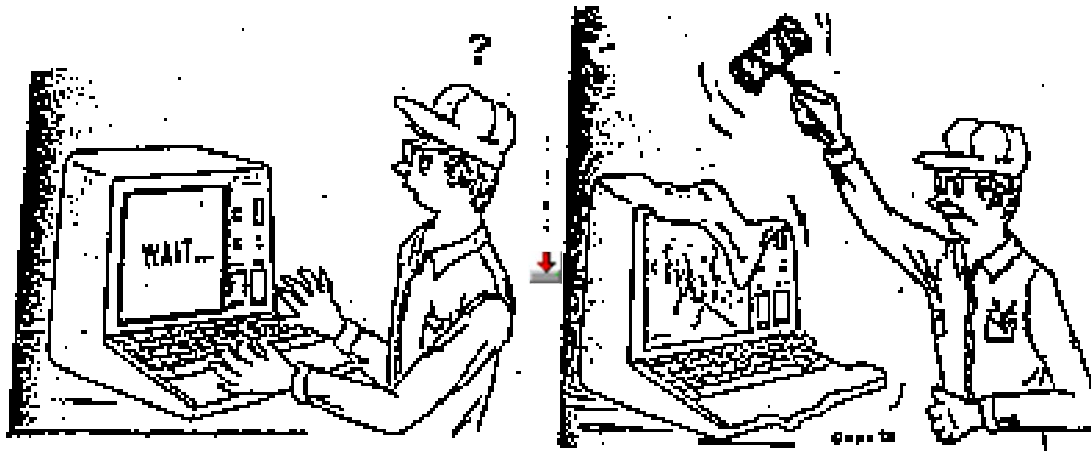


CS5239 Computer System Performance Evaluation

2018/19 – Semester I

www.comp.nus.edu.sg/~teoym/cs5239-18



Teo Yong Meng

Room: COM2, #04-39

Department of Computer Science

National University of Singapore

E-mail: teoym@comp.nus.edu.sg

Tel: 6516 2830

Faster is better

time, cost, energy

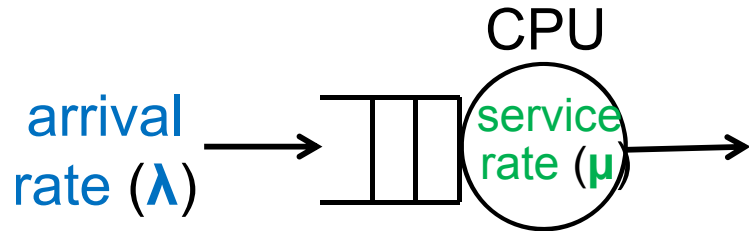
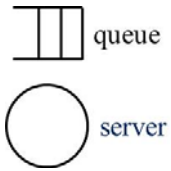
Outline

- What, why and how
- Learning objective
- What we cover
- Module Assessment
- Course Schedule & Webpage
- Resources

What is Performance

- Concept of work
 - **latency** (time)
 - **bandwidth** (rate)
- How well a computer system performs a given job or activity?
- Why do we care about computer performance?
 - **time deadline**, cost (and efficiency) of computing resources
 - energy cost, cloud computing – pay per use
- What is hard?
 - Performance of a computer system is multidimensional
 - Complex component interactions
 - Hard to predict how it will scale
 - ...

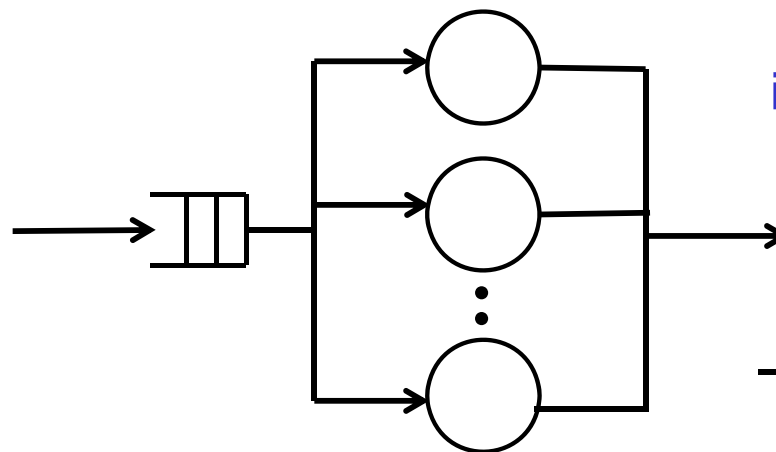
Performance Questions



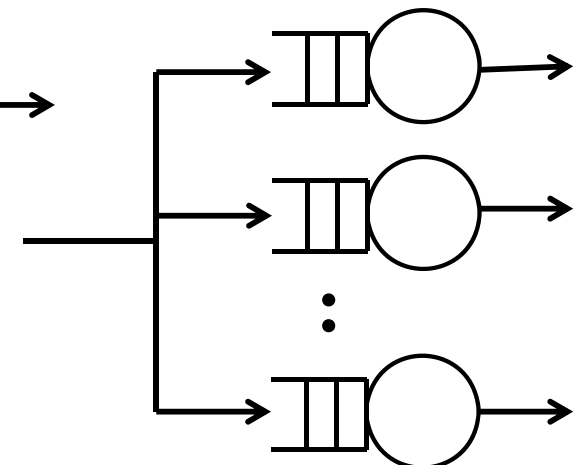
1. What is the average **time** it takes a job to complete service?
2. What is the **throughput** of the system (number of jobs completed per unit time)?
3. If arrival rate is doubled ($\lambda \rightarrow 2\lambda$), how much should μ increase? Do we do nothing or do we need another CPU?

4. If we need more server capacity, what are our options?

- a. buy **a new server** with the needed capacity
- b. buy **a few smaller servers** that adds up to the required capacity
 - i. one queue for **all** servers
 - ii. one queue for **each** server
 - iii. does it matters?

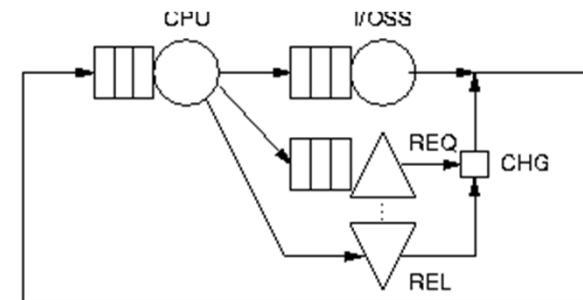


ii. one queue per server

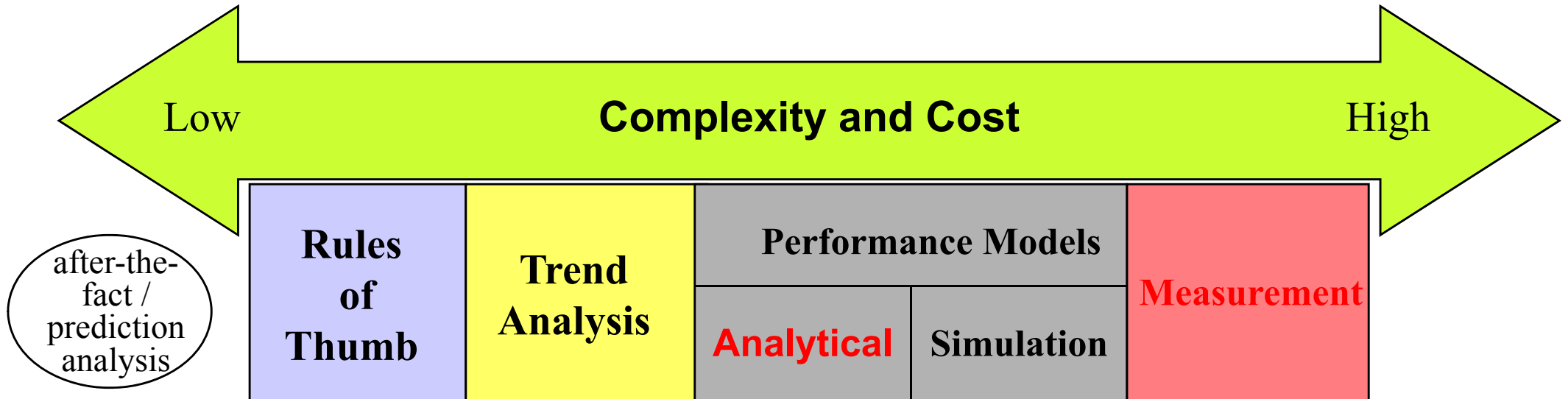


Performance Evaluation: How

- **Measurements** of actual systems
- **Modeling**
 - **Simulation** using software models
 - **Analytical modeling** using techniques such as queuing analysis



Performance Evaluation: How



CS5239 Computer System Performance Evaluation

CS5271 Performance Analysis of Embedded Systems

CS6211 Analytical Performance Modelling for Computer Systems

CS5233 Simulation and Modelling Techniques

CS6205 Advanced Modelling and Simulation

Course Catalogue

CS5239 Computer System Performance Analysis

Modular Credits: 4

Workload: 2-1-0-3-4

Prerequisite(s): (CS1020 or CS1020E or CS2020 or CS2030 or CS2113/T) and (ST1232 or ST2131 or ST2334)

The objective of this module is to provide students a working knowledge of computer performance evaluation and capacity planning. Students will be able to identify performance bottlenecks, to predict when performance limits of a system will be exceeded, and to characterize present and future workload to perform capacity planning activities. Topics include: performance analysis overview; measurement techniques and tools including workload characterization, instrumentation, benchmarking, analytical modelling techniques including operational analysis, stochastic queuing network analysis; performance of client-server architectures; capacity planning; case studies.

Prerequisites

ugrad: (CS1020 or CS1020E or CS2020 or CS2030 or CS2113/T) and (ST1232 or ST2131 or ST2334)

grad: knowledge of computer organization/architecture and statistics/probabilities

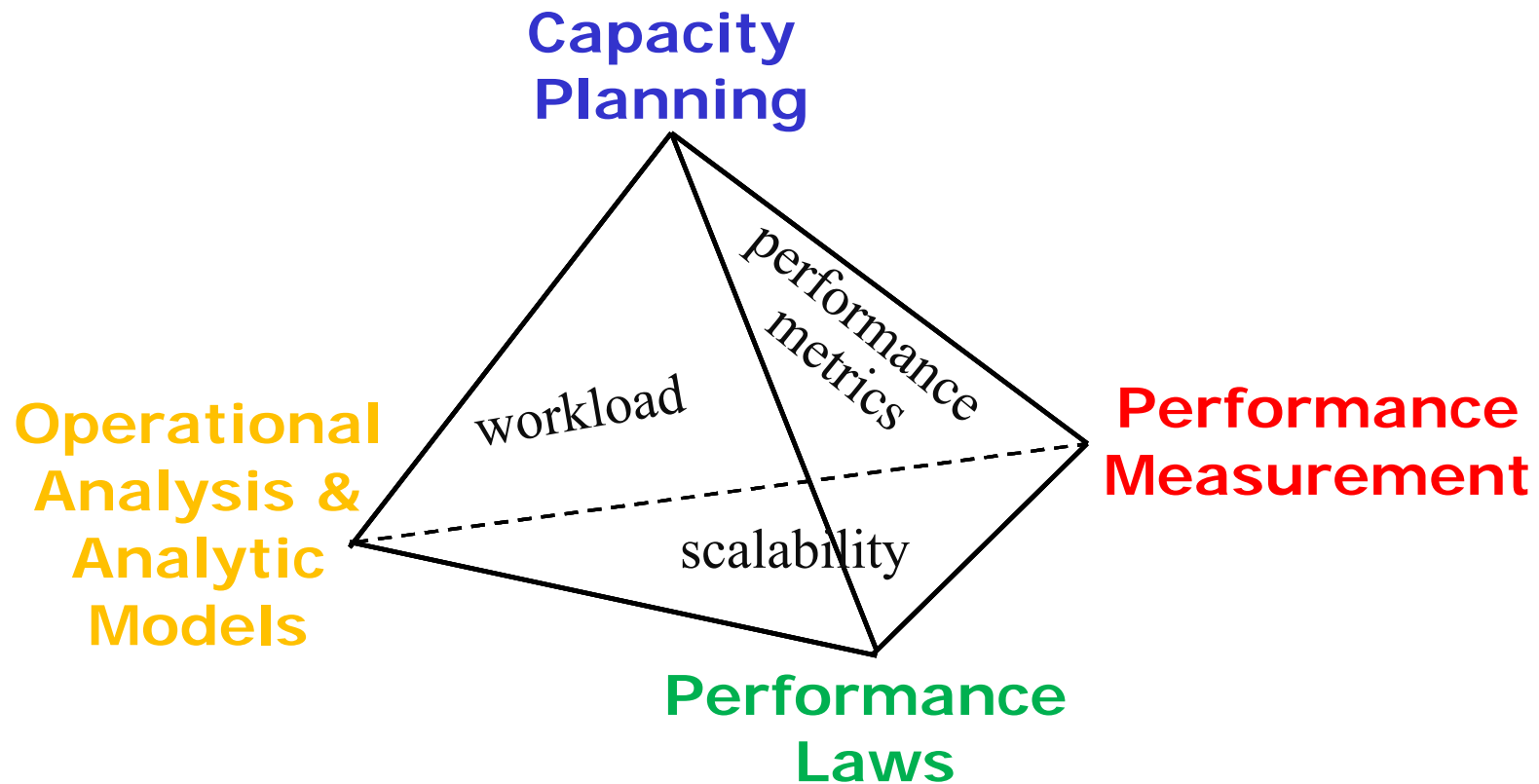
if you are not comfortable with Maths, this module may not be suitable for you

Learning Objectives

performance analysis of computer systems

1. capacity planning
2. bottleneck and modification analyses
3. measurement and analytic model analyses
4. scalability analysis

What we cover



OPERATIONAL ANALYSIS & ANALYTIC MODELS

L#03:
Probabilities
& Statistics

L#04: Queuing
Introduction &
Notation

L#05-08:
Techniques

L#12:
Performance Laws
& Scalability

L#05:
Operational
Analysis

L#06:
Analysis of
Single Queue

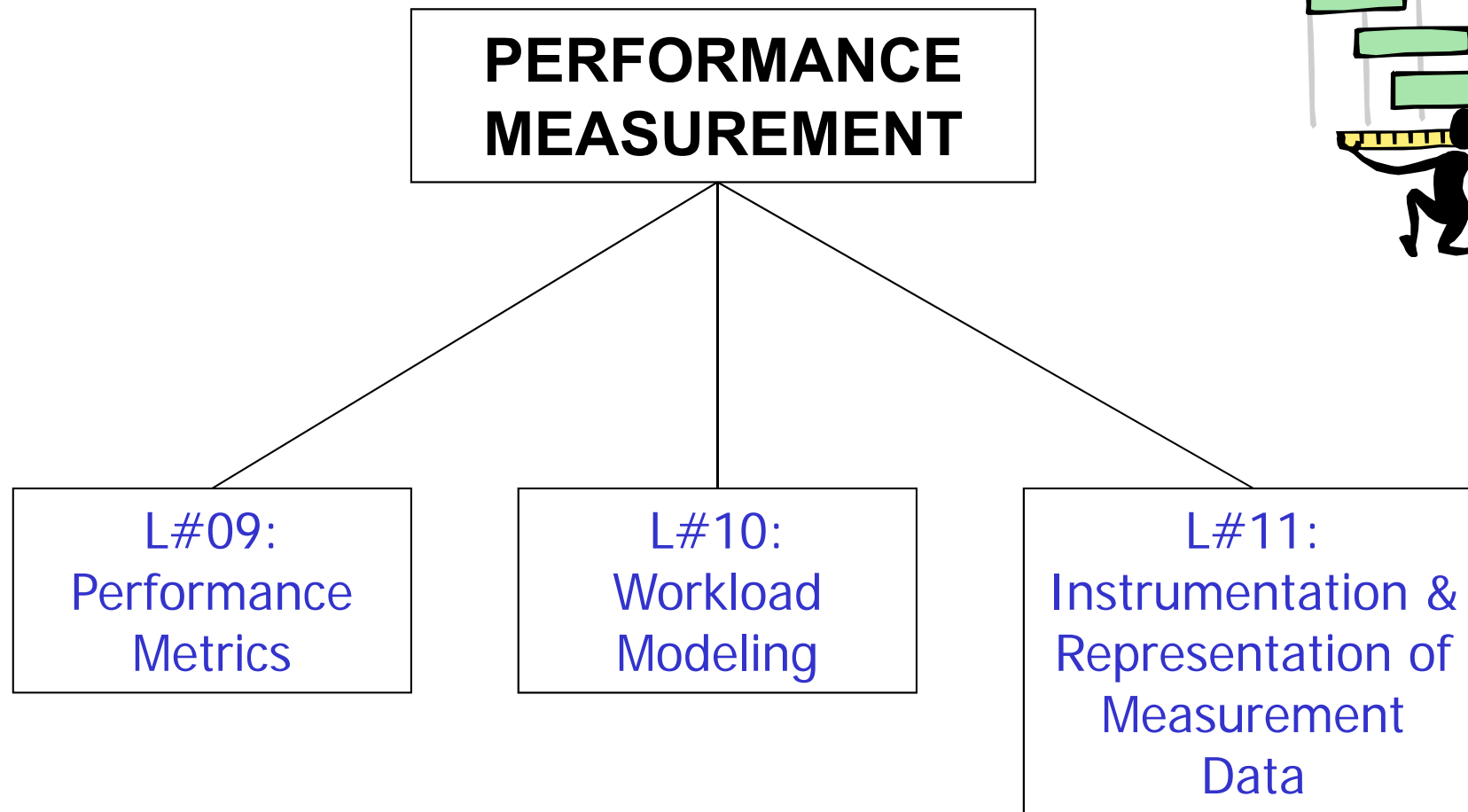
L#07: Analysis
of Queuing
Networks

L#08
Examples

- bottleneck analysis
- performance bounds

System
Component
Workload

- open, closed, hybrid
- fixed capacity, delay, load-dependent
- single, multiple classes



“Measurements are not to provide numbers but insights.”
Ingrid Bucher

Module Assessment

Continuous Assessment (100%)

- Quiz (30%)
- Assignment 1 (20%)
- Assignment 2 (20%)
- Open Book Test (30%)

Course Schedule & Webpage

- Lecture: Mon, 6.30-8.30pm, SR2@Com1, 02-04
- Tutor: Sunimal Rathnayake (Com2, #B1-01)
- Consultation: Wed, 10-12am
- Webpage:
 - IVLE for course announcement
 - www.comp.nus.edu.sg/~teoym/cs5239-18 for lecture slides, assignments, etc.



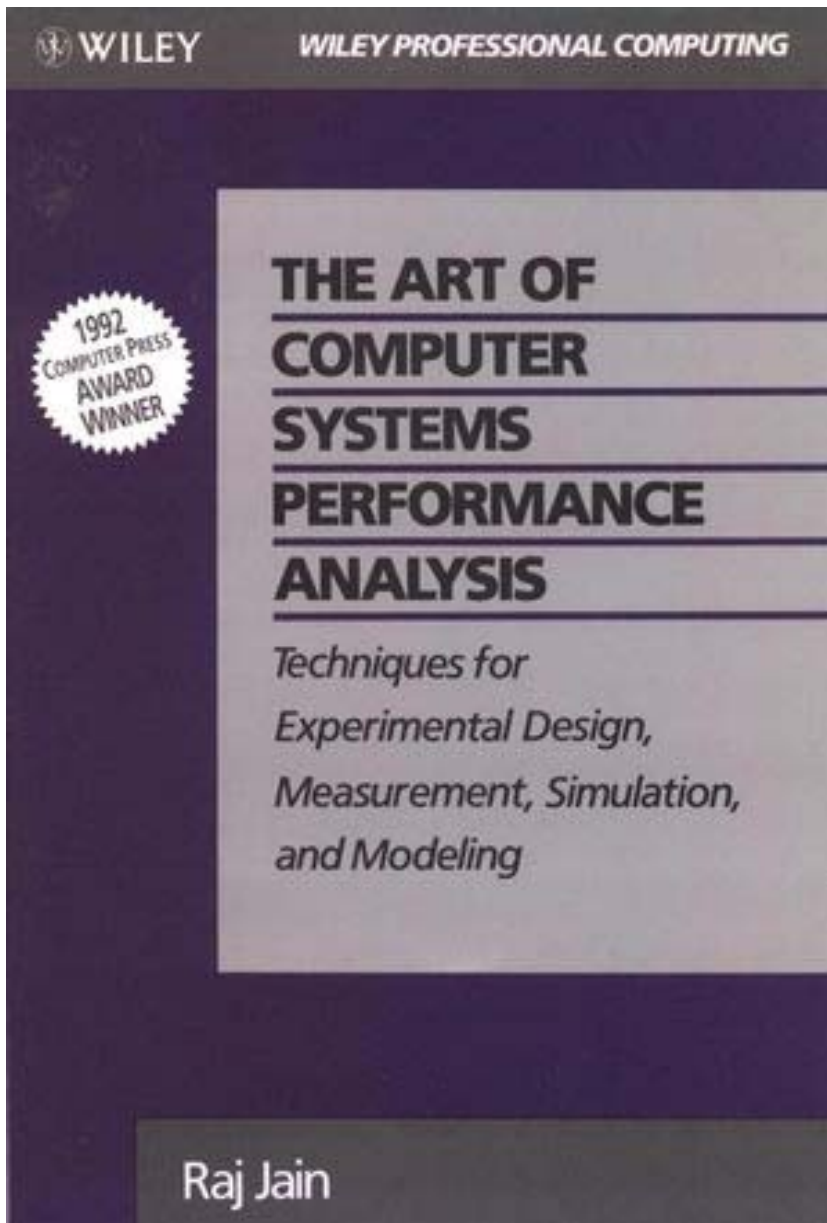
Resources

Main Textbooks

- ◆ **The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation and Modeling**, R. Jain, John-Wiley, 1991.
- ◆ Quantitative System Performance, E.D. Lazowska et al., Prentice-Hall, 1984, <http://www.cs.washington.edu/homes/lazowska/qsp/>.
- ◆ Measuring Computer Performance - A Practitioner's Guide, D.J. Lilja, Cambridge University Press, 2000.

Reference Books

- ◆ Capacity Planning and Performance Modeling - From Mainframes to Client-Server Systems, Daniel A. Menasce, et al., Prentice-Hall, 1994.
- ◆ Capacity Planning for Web Performance – Metrics, Models and Methods, D.A. Menasce, et al., Prentice-Hall, 1998.
- ◆ Simulation Modeling and Analysis, A.M. Law and W.D. Kelton, McGraw Hill, 3rd edition, 2000.
- ◆ Introduction to Parallel Computing, A. Grama, et al., Addison-Wesley, 2nd Edition, 2003.



**The Art of Computer Systems
Performance Analysis: Techniques for
Experimental Design, Measurement,
Simulation, and Modeling**

Raj Jain

ISBN: 978-0-471-50336-1

720 pages

April 1991