Computer Systems Performance Evaluation

DOMENICO FERRARI

Computer Systems Performance Evaluation

DOMENICO FERRARI

Computer Science Division
Department of Electrical Engineering
and Computer Sciences
University of California, Berkeley

PRENTICE-HALL, INC., Englewood Cliffs, New Jersey 07632

Library of Congress Cataloging in Publication Data

Ferrari, Domenico (date)
Computer systems performance evaluation.

Bibliography: p.
Includes index.
1. Electronic digital computers—Evaluation.
I. Title.
QA76.9.E94F47 621.3819'58 77-15096
ISBN 0-13-165126-9

To Alessandra and Giuliarachele
for their love, patience and support

Englewood Cliffs, New Jersey 07632

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

PRENTICE-HALL INTERNATIONAL, INC., *London*
PRENTICE-HALL OF AUSTRALIA PTY. LIMITED, *Sydney*
PRENTICE-HALL OF CANADA, LTD., *Toronto*
PRENTICE-HALL OF INDIA PRIVATE LIMITED, *New Delhi*
PRENTICE-HALL OF JAPAN, INC., *Tokyo*
PRENTICE-HALL OF SOUTHEAST ASIA PTE. LTD., *Singapore*
WHITEHALL BOOKS LIMITED, *Wellington, New Zealand*

Preface

Computers are man-made machines. Thus, they belong to the realm of engineering. They are designed by engineers, according to certain cost-performance specifications, to satisfy the information-processing needs of individuals and organizations. The programs they execute are also designed by people who, when engaged in this endeavor, act as engineers.

An essential aspect of any engineering activity is the evaluation of the systems this activity is concerned with. Engineering systems are evaluated by their designers, manufacturers, buyers, managers, and users. One of the dimensions, and a very important one, along which systems are evaluated is the dimension of performance. This book deals with the evaluation of computer systems performance. Its subject is, or should be, of interest to the great majority of computer professionals. System designers, installation directors and staff members, data processing and corporate managers at all levels, systems analysts, program designers, and computer users all have daily to cope with problems whose solution may be made substantially easier and more satisfactory by some knowledge of performance evaluation methodologies, techniques, and tools.

All computer installations have to deal with problems which require a considerable involvement in performance-evaluation activities. Examples of such problems are procurement, configuration design, system tuning, upgrading, accounting and pricing, scheduling and operations management, and short- and long-term planning. Practically all organizations which make use of computer systems must take performance into account in some of their decision-making processes. A system which does not perform as expected, or whose cost/performance ratio can be decreased, causes a waste of resources which must be avoided or eliminated.

The recognition of its practical importance and scope and the fact that it may be seen as the application of the scientific method to the study of computer systems make performance evaluation an attractive subject for an increasing number of researchers and teachers. However, the very fact that performance evaluation is being studied as a discipline distinct from system design is a symptom of an unsatisfactory situation. In other more traditional branches of engineering there are very few cases in which, as quite frequently happens with computer systems, the performance of a system can be improved even by primitive techniques by a very large factor after the system has been designed and implemented. Very few engineers specialize in, or even talk about, the evaluation of the performance of cars, planes, bridges, industrial processes, and electric circuits, since this is an essential aspect of any type of engineering design.

Computer systems engineering is not mature enough yet. It cannot base itself on quantitative laws similar to those which constitute the scientific foundations of other types of engineering: the laws of mechanics, electromagnetism, thermodynamics, and so on. Thus, performance evaluation problems should be studied with the ultimate objective of incorporating the results into design, implementation, and usage methodologies, that is, of making such a separate study meaningless.

In the long run it is reasonable to expect that the quantitative approach will permeate the core courses of computer science and engineering curricula, thereby making the offering of specific courses on performance evaluation unnecessary. However, reaching this point will require extensive and successful research, and even more important and difficult to achieve, a radical change of mentality. The distance between the present state of affairs and the goal mentioned above is perhaps best exemplified by the almost completely qualitative and descriptive approach that is currently taken to the teaching of computer system organization. In this situation it seems indispensable to offer courses entirely dedicated to performance evaluation. These courses should provide students with a new and very important viewpoint on computer systems, and with a picture of this rapidly growing field, its main themes, its problems, and the known approaches and solutions to them.

This book was born from one of such courses, which has been offered several times to Berkeley students. The level of the course, and therefore also the one of the book, has been chosen so as to be appropriate for computer science seniors and first-year graduate students. As for taking the course, the prerequisites for understanding the book are some familiarity with the material usually covered in an undergraduate two-quarter course on computer systems organization and/or operating systems, as well as with the principles and the basic techniques of computer programming and some elementary background in discrete mathematics, statistics, and probability theory. Computer professionals will not have any problems with the first two prerequisites (computer systems and programming) and will find references in the bibliography to a few textbooks that adequately cover the mathematical and statistical material required to make good use of this book.

I have emphasized the conceptual aspects of performance evaluation techniques and problems that I believe should always be given maximum priority. However, I have not neglected those informative aspects that, in my opinion, significantly contribute to providing the reader with a comprehensive view of the state of the art both in research and in the commercial world. The resulting picture is, of course, profoundly subjective and biased. I have stressed those techniques, viewpoints, and approaches that I think will become more and more important in the near future, rather than trying to provide a faithful snapshot of the present situation. I have presented methodologies and tools borrowed from other fields as if they had been invented for evaluating computer systems, and discussed their characteristics only from this standpoint. I have committed a number of other major and minor sins which the hopefully forgiving reader will soon discover.

The unifying theme of the book is a very pragmatic one: the *evaluation study* seen as a set of procedures whose end goal is to gather information on the system being evaluated so as to be able to answer certain performance-related questions. Thus, the book is mostly concerned with the evaluation studies required to solve the main problems which arise in computer systems engineering and with the techniques to be used in these studies.

Chapter 1 introduces the subject of the book by describing the viewpoint from which computer systems will be regarded. It discusses the concept of performance, the most important performance indices, and the most popular classifications of evaluation problems and techniques.

The next three chapters are devoted to the major evaluation techniques and tools. Most of these techniques are well-known in other scientific and technical fields. Except for some minimal amount of background information, these chapters deal only with their applications to computer systems. Chapter 2 describes measurement studies, measurement tools, the design of experiments, and the application of empirical modeling methods to the problem of experimental results interpretation. Chapter 3 deals with the simulation of computer systems. Model formulation, construction, calibration and validation are discussed, as well as the design of simulation experiments and the interpretation of simulation results. Chapter 4 is devoted to analytic modeling. Both deterministic and probabilistic models of systems are treated, with a particular emphasis on the application of queuing network-modeling techniques to performance analysis.

Chapter 5 is concerned with a problem of crucial importance in any evaluation study, the one of work-load characterization. Our lack of knowledge in this area is probably the most serious obstacle to progress in performance evaluation. A separate chapter has been devoted to it since the basic problems in work-load characterization are conceptually the same for all techniques. These problems, together with the known approaches to work-load modeling, are explored in this chapter.

Having been exposed to the techniques which constitute an essential part of any evaluation study, the reader should be ready to apply them to various types of evaluation problems. Three broad classes of problems are examined in the next three chapters. Chapter 6 describes the technical aspects of computer selection and discusses the adequacy of the known techniques and tools with respect to this problem. Chapter 7 deals with performance improvement, the area in which the most striking successes have been obtained, but whose importance is hopefully going to decrease as better system and installation design methodologies become available. Chapter 8 discusses the role of performance evaluation in computer system design. The techniques exposed in the first part of the book are applied there to various design problems.

Finally, Chapter 9 contains a discussion of program performance evaluation. The indices, techniques, and problems related to the performance aspects of programs are briefly described following the organizational scheme of the rest of the book.

Needless to say, this book may be correctly characterized as an introduction to performance evaluation. The reader will easily realize that each chapter might be seen as the introductory chapter, or an extended summary, of an entire book that could be written on the same subject. Besides being usable as a textbook in courses on computer systems performance evaluation or as a reference in courses on computer organization, computer systems design, and operating systems, the book, possibly supplemented with readings from periodical literature and other books, can serve as a textbook in specialized courses or short courses on measurement, modeling of computer systems, computer selection techniques, tuning techniques, and software evaluation.

Many individuals and organizations gave me substantial help in the conception and preparation of this book. The necessary background was provided by all my teachers. Among the most effective of them, I must include my parents, who taught me their philosophy of life, and my students at Berkeley, whose motivation has been a constant source of inspiration and encouragement. A number of my colleagues in the Department of Electrical Engineering and Computer Sciences at Berkeley are also to be thanked for their scientific and moral support and for the continuous challenge of their example. Several people have stimulated my work with their research and their interest in the book. I am really indebted to all the authors listed in the bibliography and to many other workers in the performance evaluation area. My particular gratitude goes to all the individuals who gave me valuable advice on the manuscript or some of its parts, especially to the reviewers, Harold Heath and Daniel Siewiorek, and to Steve Kimbleton, Frank Palermo and Felix Lam. Karl Karlstrom's encouragement and editorial support were extremely helpful throughout my effort. The typing of the manuscript was done with great care, rare patience, and excellent results by Ruth Suzuki, Edith Purser, and Mary

Ann Ratch. Finally, I would like to thank the birds, the squirrels, the trees, and the flowers of my garden in Berkeley, in whose company most of this book has been written.

The Regents of the University of California supported the preparation of this book by granting me a sabbatical quarter and a Regent's Summer Fellowship. The Universities of Pisa and Genoa and the Polytechnic Institute of Milan are also to be credited for their support. Some of the work reported in Chapter 2 was sponsored by the Advanced Research Projects Agency under contract DAHC15-70-C-0274. Some of the material reported in Chapter 9 was supported by the Division of Computing Research of the National Science Foundation under grant DCR74-18375 and the Research Division of the IBM Corporation.

For all the errors and omissions, I am indebted to no one but myself.

DOMENICO FERRARI

Contents

7

8

List of Examples

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Chapter 8