High Performance TCP/IP Networking

Concepts, Issues, and Solutions

Mahbub Hassan
The University of New South Wales

Raj Jain
The Ohio State University

Raj Jain is now at Washington University in Saint Louis, jain@cse.wustl.edu http://www.cse.wustl.edu/~jain/

Upper Saddle River, New Jersey 07458

Vice President and Editorial Director, ECS: *Marcia J. Horton*
Publisher: *Alan R. Apt*
Associate Editor: *Toni Dianne Holm*
Editorial Assistant: *Patrick Lindner*
Vice President and Director of Production and Manufacturing, ESM: *David W. Riccardi*
Executive Managing Editor: *Vince O'Brien*
Managing Editor: *Camille Trentacoste*
Production Editor: *Joan Wolk*
Director of Creative Services: *Paul Belfanti*
Creative Director: *Carole Anson*
Art Director: *Art Director's Name*
Cover Manager: *Jayne Conte*
Cover Designer: *Suzanne Behnke*
Cover Image: *Credid if Appropriate*
Managing Editor, AV Management and Production: *Patricia Burns*
Art Editor: *Gregory Dulles*
Manufacturing Manager: *Trudy Pisciotti*
Manufacturing Buyer: *Lynda Castillo*
Marketing Manager: *Pamela Shaffer*

Pearson Prentice Hall
Pearson Education, Inc.
Upper Saddle River, NJ 07458

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

ISBN 0-13-064634-2

Pearson Education Ltd., *London*
Pearson Education Australia Pty. Ltd., *Sydney*
Pearson Education Singapore, Pte. Ltd.
Pearson Education North Asia Ltd., *Hong Kong*
Pearson Education Canada, Inc., *Toronto*
Pearson Educación de Mexico, S.A. de C.V.
Pearson Education Japan, *Tokyo*
Pearson Education Malaysia, Pte. Ltd.
Pearson Education, Inc., *Upper Saddle River, New Jersey*

To my parents, my wife, my son Aaron, and all readers of this book

—Mahbub Hassan

To my wife, Neelu, and my sons, Sameer and Amit

—Raj Jain

Contents

Preface

The world is undergoing a revolution in information and communication technology. Not only the lives of citizens but also the networking technology are profoundly affected by this revolution. Traditional wired networks are being replaced or complemented by networks based on wireless, optical, satellite, and other media. TCP/IP has emerged as the global Internet-working solution allowing communication over a wide variety of media and networks. These new networking media and the new ways of communication over these networks have given rise to a host of new performance issues and concepts. To adapt and contribute effectively to such changes, engineers and computer scientists must acquire a solid foundation and understanding of the fundamental concepts that affect performance in TCP/IP networks.

Existing texts on TCP/IP focus on the presentation of the protocol details with little coverage of the performance issues and concepts. These texts are good for a first course on TCP/IP networking but do not provide sufficient material for those advanced readers interested in acquiring in-depth knowledge of the performance aspects of TCP/IP, especially in the emerging networking environment. To address this need, we have written *High Performance TCP/IP Networking: Concepts, Issues, and Solutions*, with a clear focus on the performance fundamentals of TCP/IP.

High Performance TCP/IP Networking: Concepts, Issues, and Solutions is a comprehensive guide to the study of its topic. Our book provides an in-depth coverage of (1) tools and techniques for the performance evaluation of TCP/IP networks, (2) performance concepts and issues for running TCP/IP over wireless, mobile, optical, and satellite networks, (3) congestion-control algorithms in hosts and routers, and (4) high performance implementation of TCP/IP protocol stack. This text has been created with an emphasis on fundamental concepts, such as network measurement and simulation techniques, mathematical modeling of TCP dynamics, and management of implementation overhead, which will continue to guide new developments in TCP/IP. Although many specific networks, tools, and protocols are discussed in the text, a continuous effort has been made to emphasize the underlying performance issues and concepts.

CONTRIBUTING AUTHORS

This book contains contributions from many leading experts actively working on specific performance issues in TCP/IP networks. In addition to the two editors (Hassan and Jain), who themselves wrote parts of the book, there are a total of 24 authors who wrote specific chapters of the book. One of the most challenging tasks was to integrate these individual submissions into a coherent book. As part of the integration effort, the editors have introduced a range of additional materials, including learning objectives, review questions, hands-on projects, and case studies. The editors maintained close liaison with the chapter authors throughout the manuscript preparation process. The manuscript was reviewed and revised twice to address the concerns of the reviewers. Substantial material was added in each revision to further integrate the chapters and improve the quality of the book. While most individual chapter authors were contacted for the revisions, the editors

themselves revised some of the chapters. The list of authors who contributed to this book follows (chapters that were substantially revised by the editors are marked with an asterisk).

Chapter 1 Introduction
Mahbub Hassan, University of New South Wales, Australia
Raj Jain, Ohio State University, USA

Chapter 2 TCP/IP Fundamentals*
Sanjay Jha, University of New South Wales, Australia

Chapter 3 Performance Measurement of TCP/IP Networks*
Yukio Murayama, Kurashilki University of Science and Arts, Japan
Suguru Yamaguchi, Nara Institute of Science and Technology, Japan

Chapter 4 TCP/IP Network Simulation
Mahbub Hassan, University of New South Wales, Australia
Sonia Fahmy, Purdue University, USA
Jim Wu, University of New South Wales, Australia
Abdul Aziz, University of New South Wales, Australia

Chapter 5 TCP Modeling
Sven Östring, University of Cambridge, United Kingdom
Harsha Sirisena, University of Canterbury, New Zealand

Chapter 6 TCP/IP Performance over Wireless Networks
George Xylomenos, Athens University of Economics and Business, Greece
George Polyzos, Athens University of Economics and Business, Greece
Petri Mähönen, Aachen University, Germany
Mika Saaranen, Nokia Mobile Phones, Finland

Chapter 7 TCP/IP Performance over Mobile Networks
Raghupathy Sivakumar, Georgia Institute of Technology, USA

Chapter 8 TCP/IP Performance over Optical Networks
Franco Callegati, Universita' di Bologna, Italy
Maurizio Casoni, Universita' di Modena and Reggio Emilia, Italy
Carla Raffaelli, Universita' di Bologna, Italy

Chapter 9 TCP/IP Performance over Satellite Networks*
Arjan Durresi, Louisiana State University, USA
Sastri Kota, Loral Skynet, USA

Chapter 10 TCP/IP Performance over Asymmetric Networks
Venkat Padmanabhan, Microsoft Research, USA
Hari Balakrishnan, Massachusetts Institute of Technology, USA

Chapter 11 New TCP Standards and Flavors
Sonia Fahmy, Purdue University, USA

Chapter 12 Active Queue Management in TCP/IP Networks
Mohammed Atiquzzaman, University of Oklahoma, USA
Bing Zheng, New Focus, Inc., USA

Chapter 13 Software Implementation of TCP
Jeff Chase, Duke University, USA

Appendix A M/M/1 Queue
Mahbub Hassan, University of New South Wales, Australia
Raj Jain, Ohio State University, USA

Appendix B FreeBSD
Rui Zhao, University of New South Wales, Australia

Appendix C TCP Auto-Tuning
Mahbub Hassan, University of New South Wales, Australia

ORGANIZATION AND OUTLINE

The book is organized into five parts.

- **Part I: Background.** Part I provides an introduction to the book. It contains two chapters. Chapter 1 provides a rationale for the book. Chapter 2 reviews some of the key features of TCP/IP protocols used in later chapters in the book to explain many performance issues. Chapter 2 reviews only the key features of TCP/IP. A comprehensive treatment of TCP/IP protocol stack is beyond the scope of the book.

- **Part II: Performance Evaluation.** Part II consists of Chapters 3, 4, and 5 and provides detailed coverage of the tools and techniques for performance evaluation of TCP/IP networks. Chapter 3 discusses the performance measurement tools available for monitoring, analyzing, and benchmarking the performance of TCP/IP networks. Chapter 4 introduces simulation techniques and discusses two popular simulation tools. Chapter 5 is devoted to the mathematical modeling of TCP congestion control algorithms.

- **Part III: Performance in Emerging Networks.** Chapters 6 through 10 examine the performance concepts and issues for running TCP/IP in the emerging networking environment. Although many of us think modems and Ethernet when we think Internet and TCP/IP, this is no longer the reality. Yes, it is true that nearly every home has a modem for Internet connection, and nearly every organization has some version of the wired Ethernet connectivity to the desktop. Many homes, however, are subscribing to Digital Subscriber Loop (DSL) technologies for high-speed Internet connection, and many organizations are deploying wireless LANs for flexibility. In the wide area, too, we are witnessing new networking technologies such as mobile cellular data networks (e.g., GPRS), high-speed optical backbones, and increasing use of satellite links for long-distance and global coverage. Each of these new technologies has given rise to some new concepts and issues for TCP/IP performance. We have therefore dedicated a separate chapter to deal with each of these technologies: Chapter 6 for wireless, Chapter 7 for mobility, Chapter 8 for optical, Chapter 9 for satellite, and Chapter 10 for asymmetric networks (e.g., ADSL).

- **Part IV: Congestion Control.** With the increase in networking complexities and traffic dynamics, congestion-control algorithms employed at the end hosts and in the network routers continue to evolve. The new congestion-control algorithms in the TCP protocol resulted in many different TCP flavors (e.g.,

Tahoe, Reno, Vegas, and so on). Part IV consists of Chapters 11 and 12 and presents an in-depth coverage of the congestion-control algorithms proposed so far. Chapter 11 discusses various TCP flavors, and Chapter 12 examines the new queue management schemes proposed for the network routers to combat congestion in highly dynamic environment.

- **Part V: Implementation.** For emerging high-speed networks (e.g., 10 Gbps Ethernet), the end-system implementation of TCP can become a performance bottleneck. Part V (Chapter 13) summarizes critical performance issues for TCP implementation in end systems and surveys solutions for improving bulk transfer performance.

HOW TO USE THIS BOOK

The book is designed for use in a second course on networking with a prerequisite course on introductory networking or data communications. Some of the possible courses for which this book can be used include Advanced Computer Networks, Advanced TCP/IP Networks, High Performance Networks, and Internet-working. There is enough material in the book for a one-semester or one-quarter course with 12 or 13 weeks of lecture. Depending on the background of the students, two possible course compositions are given here.

Computer science students with limited background in mathematics and hardware design can exclude Chapter 5 (mathematical modeling) and study Chapter 4 (Simulation) in more detail. Engineering students graduating in computer or electrical engineering can spend fewer weeks on Chapter 4 and one extra week on Chapter 5.

Professionals working as network engineers, R & D managers, research scientists, and network administrators will also find this book valuable as a reference to the most recent advances in TCP performance research.

LEARNING AIDS

There are many learning aids in this book:

- **Learning Objectives.** Each chapter starts with a list of learning objectives. The learning objectives highlight the fundamental concepts (skills) students should understand (master) as a result of reading the chapter and help them organize their study goals. They assist instructors in pointing out lecture objectives.

- **Further Reading Lists.** Annotated reading lists at the end of the chapters provide students with valuable resources for independent exploration on specific topics of interest. These lists are particularly useful for professionals.

- **Chapter Summaries.** Summaries offer students a chance to review their understanding of key concepts in the chapter before moving on.

- **Review Questions.** End-of-chapter review questions evaluate the degree to which the student achieved the learning objectives and force the students to think about the key concepts in the chapter. Answers to most of the review questions can be found directly from the chapter; therefore, students are

forced to reread parts of the chapter to locate the answers. Such rereading is often required to gain a clear understanding of many difficult concepts. The instructor can use some of these questions for classroom discussions or class tests.

- **Hands-On Projects.** For each chapter, a list of performance evaluation experiments are provided for advanced students seeking to gain a deeper understanding of some of the key concepts and solutions described in the chapter. These experiments can be carried out on open platforms using freely available software. The hands-on projects in this book cover a range of difficulty. Some experiments can be completed in a few weeks using *ns*-2 simulation software, without requiring any kernel-level programming. Other experiments require modification of existing TCP/IP stacks in FreeBSD operating system kernel. These experiments are quite challenging and can be given to students as whole semester projects. Students attempting these experiments are expected to have a good background in programming and operating systems. (Appendix B provides a brief tutorial on FreeBSD for students with no prior background in kernel programming.)

- **Case Studies.** A case study is introduced in Chapter 1 based on a fictitious, but realistic organization with TCP/IP networking infrastructure. The same case study is then used in subsequent chapters with some modifications to introduce new performance problems. The running case study holds together different chapters in the text, provides students a realistic context in which to apply the concepts and techniques learned in the relevant chapters, and yields a classroom discussion topic for the instructor.

- **Figures and Illustrations.** Many concepts throughout the book are explained using illustrations. These illustrations help students understand complex performance issues and concepts.

- **Examples.** Examples have been used where applicable to explain the use of techniques learned from the text.

ACKNOWLEDGMENTS

The book would never exist without the contributions from the individual chapter authors. We take this opportunity to thank all chapter authors for their expertise and time and for putting up with our many requests throughout the preparation of the manuscript. We are indebted to the anonymous reviewers for reading the whole manuscript or part of the earlier versions of the manuscript and making useful comments. Their constructive suggestions significantly influenced the revisions of the manuscript. We thank Professor Krzysztof Pawlikowski of the University of Canterbury, New Zealand, for providing early feedback on Chapter 4. The author of Chapter 11 thanks Tapan Karwa, Venkatesh Prabhakar, Farnaz Erfan, and Minseok Kwon for their help with the simulation experiments in that chapter. Jim Wu, a coauthor of Chapter 4, has been instrumental in fixing some of the problems we faced in preparing the manuscript in LaTex. We gratefully acknowledge the support of the entire production team at Prentice Hall. Finally, the first editor (Hassan)

would like to thank Professor Arun Sharma (previous head of school) and Professor Paul Compton (current head of school) at the University of New South Wales for providing a pleasant and stimulating environment in which to work.

Mahbub Hassan
Raj Jain