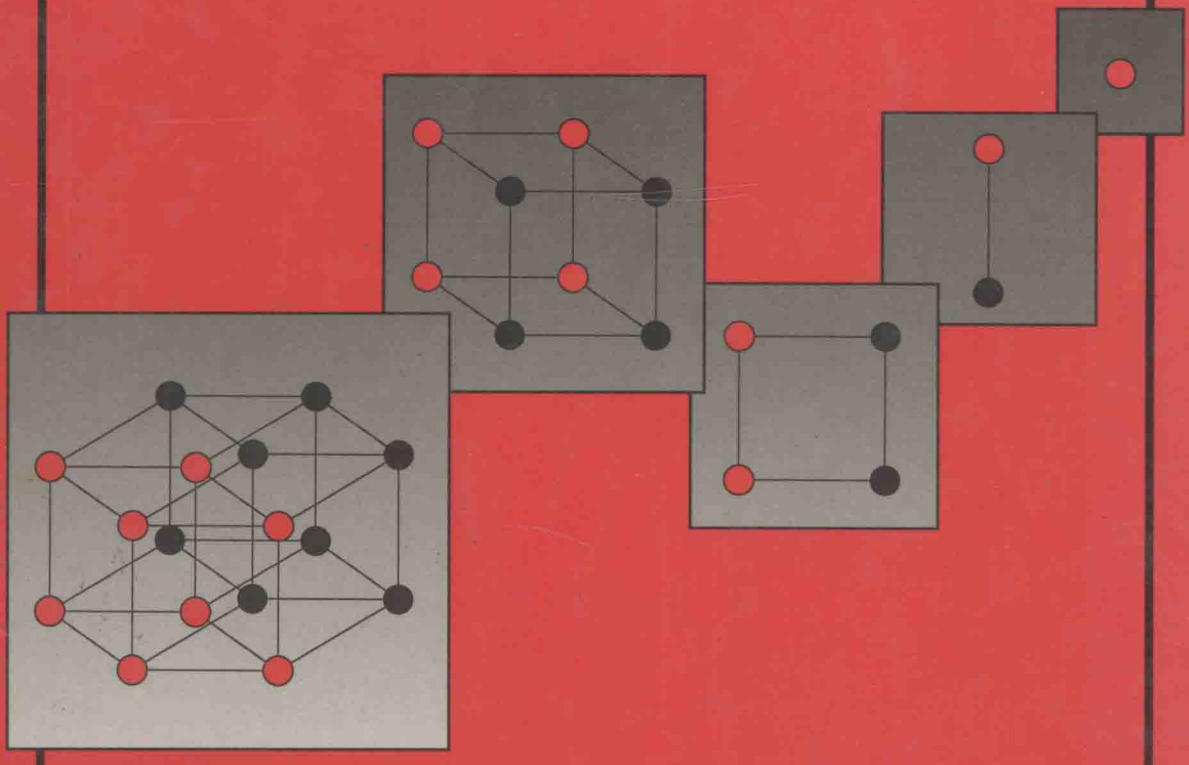
Harold W. Lawson

Parallel Processing in Industrial Real-Time Applications



Parallel Processing in Industrial Real-Time Applications

Harold W. Lawson

Lawson Publishing and Consulting, Inc.
Lidingö, Sweden

with contributions by

Bertil Svensson
Chalmers Technical University, Sweden
Halmstad University, Sweden

Lars Wanhammar
Linköping University, Sweden

江苏工业学院图书馆
藏书章

Prentice Hall, Englewood Cliffs, New Jersey 07632

Library of Congress Cataloging-in-Publication Data

Lawson, Harold W.
Parallel processing in industrial real-time applications / Harold W. Lawson; with contributions by Bertil Svensson and Lars Wanhammar.
p. cm. — (Prentice Hall series in innovative technology)
Includes bibliographical references and index.
ISBN 0-13-654518-1
1. Parallel processing (Electronic computers) 2. Real-time data processing. I. Svensson, Bertil. II. Wanhammar, Lars. III. Title. IV. Series.
QA76.58.L38 1992 91-48015
004".33—dc20 CIP

Editorial/production supervision: *Harriet Tellem*
Cover design: *Karen Marsilio*
Prepress buyer: *Mary E. McCartney*
Manufacturing buyer: *Susan Brunke*
Acquisitions editor: *Karen Gettman*

© 1992 by Prentice-Hall, Inc.
A Simon & Schuster Company
Englewood Cliffs, New Jersey 07632

The publisher offers discounts on this book when ordered in bulk quantities. For more information, write:
Special Sales/Professional Marketing
Prentice-Hall, Inc.
Professional & Technical Reference Division
Englewood Cliffs, New Jersey 07632

This book has been developed as the major component of the Swedish National Information Technology Program (IT4) project concerning parallel processors. The Swedish Defense Materiel Administration (FMV) along with the Swedish National Technical Development Board (Nutek) were the governmental sponsors of the project. NobelTech, Ericsson Radar Electronics AB, and Lawson Förlag och Konsult AB were the industrial partners. The project leader was Gunnar Carlstedt of Carlstedt Elektronik AB.

All rights reserved. No part of this book may be reproduced, in any form or by any means, without permission in writing from the publisher.

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

ISBN 0-13-654518-1

Prentice-Hall International (UK) Limited, *London*
Prentice-Hall of Australia Pty. Limited, *Sydney*
Prentice-Hall Canada Inc., *Toronto*
Prentice-Hall Hispanoamericana, S.A., *Mexico*
Prentice-Hall of India Private Limited, *New Delhi*
Prentice-Hall of Japan, Inc., *Tokyo*
Simon & Schuster Asia Pte. Ltd., *Singapore*
Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*

Parallel Processing in Industrial Real-Time Applications

Prentice Hall Series in Innovative Technology

Dennis R. Allison, David J. Farber, and Bruce D. Shriver *Series Advisors*

Bhasker	*A VHDL Primer*
Blachman	*Mathematica: A Practical Approach*
Johnson	*Superscalar Microprocessor Design*
Kane and Heinrich	*MIPS RISC Architecture, Second Edition*
Lawson	*Parallel Processing in Industrial Real-Time Applications*
Nelson, ed.	*Systems Programming with Modula-3*
Nutt	*Open Systems*
Rose	*The Little Black Book: Mail-Bonding with OSI Directory Services*
Rose	*The Open Book: A Practical Perspective on OSI*
Rose	*The Simple Book: An Introduction to Management of TCP/IP-Based Internets*
Shapiro	*A C + + Toolkit*
Slater	*Microprocessor-Based Design*
SPARC International Inc.	*The SPARC Architecture Manual, Version 8*
Strom, et al.	*Hermes: A Language for Distributed Computing*
Treseler	*Designing State Machine Controllers Using Programmable Logic*
Wirfs-Brock, Wilkerson, and Weiner	*Designing Object-Oriented Software*

此为试读,需要完整PDF请访问: www.ertongbook.com

In memory of

Rear Admiral Grace Murray Hopper

On January 1, 1992, Rear Admiral Hopper died. Her pioneering work in the field of computing will long be remembered. It is with deep respect and appreciation for her guidance in my career that I dedicate this book to her memory.

FORWARD

Real-time systems permeate the industrialized world. They control life-sustaining instrumentation in the intensive care units of hospitals, monitor the operations of petrochemical plants, and control the on-board navigation and propulsion mechanisms of spacecraft, airplanes, and oil tankers. They are an integral component of global communication networks and financial markets and are embedded in many pieces of equipment that we employ on a day-to-day basis, from the appliances we use to the antilock brakes that have increased the safety of our automobiles.

The next decade will witness significant additional growth of such systems. Just the use of real-time data from the international network of geo-positioning satellites will result in a new generation of real-time mapping and navigation systems that will be deployed in the tens of thousands and will affect our daily lives. Not only will commercial aircraft and ships benefit, but fire and police departments will increase their ability to respond to emergency situations in thousands of cities throughout the world. *Important as real-time systems have become, there is very little formal training in most universities today concerning the specification, design, implementation, testing, evaluation, and evolution of the hardware and software components of real-time systems.*

Parallel processing has a similar history. The availability and use of highly parallel systems is on a steep growth curve. No longer is the use of these systems exclusively restricted to the domain of scientific and technical research and development laboratories. The extremely high-levels of circuit integration in today's VLSI chips, coupled with their inherent availability, reliability, and low cost, have allowed the emergence of a new generation of high-performance, cost-effective highly to massively parallel systems. *Furthermore, there are important classes of real-time systems that either can benefit from or require the performance of parallel systems in order meet their design goals.* However, as is the case of real-time systems, there is very little formal training in most universities today concerning the specification, design, implementation, testing, evaluation, and evolution of the hardware and software components of parallel systems, let alone parallel, real-time systems.

The academic and industrial research and development literature in both real-time systems and in parallel systems is growing at a startling rate, but is fragmented and uneven. New journals in both areas are seemingly being announced every few months. Yet, most of this material does not deal with fundamental principles. It rarely gives the practitioner a basis for understanding and evaluating alternative approaches for dealing with the myriad of technical issues that must be addressed when designing, implementing, and testing parallel, real-time systems.

This engaging and important book by Harold Lawson and his colleagues effectively addresses this need. It provides background in both real-time and parallel systems in a broad, yet unifying way. Harold uses a systematic examination of the various "views" of processes, communications, and synchronization inherent in different processing paradigms as a common thread to discuss real-time and parallel systems. Coupled with his *Behavior Mapping Resource* model, used to characterize the development of all parallel processing systems, he provides the reader with a framework for understanding and evaluating these systems. A framework, which as noted previously, has been missing in the journal literature. He focuses on practical implementations through a rich use of case studies and examples. This book should appeal to those having a background in science, engineering, and computer science and engineering disciplines because of its interdisciplinary orientation. It contains extensive references, glossary, and index entries. It is a must read for those interested in either real-time systems, parallel systems, or both.

Bruce Shriver

PREFACE

The utilization of computer technology in a wide variety of industrial real-time applications has been increasing rapidly during the 1970s and 1980s. The utilization has ranged from relatively simple signal processing and real-time control applications up to sophisticated applications that require significant quantities of computer and human resources.

As the degree of sophistication increases in respect to the goals established for industrial real-time systems, it becomes increasingly important to understand the possibilities of applying parallel processing. In some instances, parallel processing may be the "natural" way of viewing a problem or class of problems to be solved. In other cases, due to heavy processing demands, parallel processing can be the best, perhaps only, means of providing sufficient processing power to meet strict real-time deadlines. There are typically further requirements, such as dependability and responsiveness, implying features such as high reliability, availability, testability, and fault tolerance that dictate the use of a parallel processing approach.

This book is written, first, to both inform and stimulate practicing engineers and scientists who deal with concrete real-time problems that must be solved in industrial environments. The **goal** is to provide a broad, yet unifying, view of the seemingly complex heterogeneous subject area of parallel processing. In this regard, the book treats the fundamental issues, as well as considering parallel processing approaches that have been or can conceivably be applied to solving various problems in the industrial real-time problem domain. The book should provide insight into the subject area, as well as serve as a reference book for continual utilization in evaluating alternative solutions to industrial real-time application problems. In particular, the reader should be able to understand and compare the common denominators as well as the differences between various potential approaches.

Parallel processing is the subject of intensive research and development. As a result, several books have appeared that provide an exhaustive cataloging of the field. A listing of several important books is presented in the Literature Review. This book does not

attempt to catalog the general developments in parallel processing by duplicating existing sources. On the other hand, the treatment of those portions of parallel processing research and development as well as products that are or can potentially be of interest for the industrial real-time systems domain are explored in depth.

The book may also be used by faculty and students of engineering and science in an advanced undergraduate or early graduate course. The topic is covered in an inter-disciplinary manner by treating subjects normally addressed in signal processing, automatic control, mechatronics, numerical analysis, computer organization, computer architecture, programming languages, software engineering, operating systems, and real-time systems. Thus, a bridge is formed between science, engineering, computer engineering, and computer science disciplines.

A general knowledge of computer systems technology, its utilization and programming, is required. On the other hand, specific knowledge or experience concerning the subject of parallel processing is not required. A working knowledge of fundamental university level mathematics is assumed.

Organization of the book

The book is organized into the following parts:

Part I: Fundamental Issues
Part II: Selected Paradigms, Models, and Artifacts
Part III: Case Studies

In Part I, we consider fundamental issues in a manner that provides a highly structured presentation of the union of industrial real-time systems and parallel processing. Of particular importance is the introduction of a parallel processing development model called the BMR (**B**ehavior **M**apping **R**esource) model, which is utilized to characterize the development of all parallel processing systems. Further, the "fundamental issues" are identified as being the view of **p**rocesses (P), means of communication (C), and method of synchronization (S). We also find that these issues, in relationship to the BMR model, are universal for all forms of parallel processing systems. The BMR development model and the PCS issues are used as a thread throughout the book; they provide the reader with an ever present mental model of what is important and how the current material is related to the big picture. Further, the model and issues will assist the reader in understanding relevant current and future developments.

In addition to presenting the BMR model and the PCS issues, the introductory chapter provides a general background to the topic and defines several important concepts and terms. In the remaining chapters of Part I, we consider the general properties of real-time systems, philosophies, paradigms, and models for parallel processing, the

applicability of parallel processing to various real-time relevant application domains and the dependability and responsiveness of parallel processing real-time systems. To provide an important perspective, the final chapter provides a brief history of relevant parallel processing developments. In summary, all fundamental information is presented in Part I.

In Part II, we consider several specific hardware and software paradigms, models, and artifacts that are or can be of interest with respect to applying parallel processing in industrial real-time systems. In the prelude to this part, a summary of the various views and mechanisms related to the fundamental PCS (Process, Communication, and Synchronization) issues is provided. While the paradigms, models, and artifacts presented in this part are by no means complete, they are representative of the state of the art in behavioral description, mapping, and hardware resources.

In Part III, we present three case studies that illustrate the utilization of parallel processing approaches that have been applied in concrete industrial and/or military real-time applications. Further, we consider a few relevant ongoing research and development efforts that are aimed at providing new paradigms and models for parallel/distributed processing in industrial real-time applications.

In the last chapter, some concluding remarks are made and future trends are considered. A literature review is provided to serve as a starting point for the readers' access to the important sources of relevant literature. Finally, an extensive glossary and an index are included.

The style of presentation

While the subject area of the book is not sufficiently mature to provide the factual form of a typical handbook, our presentation introduces the subject in a highly structured manner. The structuring has permitted us to achieve the central goals of our Swedish National Information Technology project: to categorize ideas, concepts, and practices, providing a basis for understanding as well as for future reference. Wurman (1989) in his book *Information Anxiety* points to the general problem of assimilating knowledge from the enormous quantities of information around us. Certainly, in the subject area of this book, one can observe problems of this nature. Wurman calls for new inventive approaches to the structuring of information that provide convenient maps between what readers understand and the material to be assimilated. In this respect, we present the material of this book in a style aimed at improving information assimilation.

In Part I, each chapter (in Chapter 3, sections of the chapter) begins with a résumé of the major points made in the text with reference to the page(s) where the point is discussed. It is suggested that the reader review these points first. Perhaps they will provide a guide to skip over (or skim over) information with which the reader is well acquainted or be a signal to concentrate on subjects that are not well known. After

此为试读,需要完整PDF请访问: www.ertongbook.com

reading the chapter or sections, the reader should return to the résumé, which also serves as a summary of the material that has been presented. Thus, the structure utilized in addressing the fundamental issues follows the general structure of a good speech; that is, (1) tell the audience what you are going to tell them, (2) tell them, and (3) tell them what you have told them.

In Part II, we shift emphasis from fundamental information gathering to descriptions of specific software and hardware paradigms, models, and artifacts. Part III provides case studies that illustrate how some designers have viewed the use of parallel processing for real-time applications as well as some relevant ongoing research and development. The style of presentation in these latter parts is rather conventional; however, by the chapter division, information on specific artifacts, and projects is rapidly accessible.

In reading the book or in using it as a basis for course literature, it may be desirable to consider particular paradigms, models, artifacts, or case studies as the process of fundamental information gathering in Part I proceeds. Thus, Parts II and III serve as an easily accessible reference work.

The BMR model and reference to the PCS issues permeate the entire book. The BMR model provides a graphic, appearing throughout the book, which, in the spirit of Wurman's observations, aids the reader in extracting the relevance of the current subject with respect to the scope implied by this universal model.

The organization and style of the book have been designed for accessibility and maintainability. The material in Part I can be updated as the state of the art proceeds. Further, the paradigms, models, artifacts, and case studies in Parts II and III can be continually maintained as important advances are identified.

Acknowledgments

Many people have contributed to this book and a warm expression of gratitude is expressed for their advice, discussions, and contributions: In particular, Gunnar Carlstedt, who proposed the project as an ingredient of the Swedish National Information Technology Program and who, as project leader, provided constant guidance in the development of the book. Ingemar Carlsson, Technical Director of the Swedish Defense Materiel Administration (FMV) recognized the importance of the topic for future real-time application development and was a driving force leading to the initiation of the project. The important contributions of Lars Wanhammar in the digital signal processing area and Bertil Svensson in the SIMD PE array architecture area are hereby recognized. Ragnar Arvidsson, Tor Ehlersson, Kurt Lind, and Robert Forchheimer provided information for the application case studies. The author expresses his deep appreciation to Björn Lisper for the use of material from his Parallel Processing course notes used at the Royal Institute of Technology. Lars Bergman has prepared the photographic illustrations for image processing examples and has granted permission to utilize photographs of his daughter, Lina Bergman.

The project was constantly reviewed by a control board consisting of representatives for the governmental and industrial partners. A sincere expression of appreciation is given to Ragnar Arvidsson, Tor Ehlersson, Göran Tengstrand, Mats Zachrison, Christopher Bengtsson, and Dennis Söderberg for their active participation in this project. The author had discussions with many other colleagues who have provided useful information and constructive criticism. The book has been reviewed in detail, at various stages, by Tor Ehlersson, Sven Holte, and Handong Wu. Borko Furht and Ted Lewis have also provided useful comments and suggestions. Further, appreciation is expressed to Robert Babb, Lars Bengtsson, Miquel Bertran, Wolfgang Halang, Jens Riboe, Björn Sikström, Jan Torin, and Pelle Wiberg for their interest in the project.

Employees of the sponsoring organizations, students at the Royal Institute of Technology, Stockholm, and Halmstad University, as well as participants at intensive four day courses, have endured the early versions of this material and have thereby made a significant contribution to its quality.

A deep expression of gratitude is given to those colleagues whose work has formed the basis for a large part of the presentation and is referenced in this book. The author apologizes for all omissions of important and relevant work and solicits communication from interested parties, thus enabling the material to be updated in future editions.

In addition to my word processor, I express my appreciation to my wife Annika for her assistance in preparing the book. Anders Hedin and Lars-Erik Thorelli kindly arranged for my use of facilities at the Department of Telecommunication and Computer Systems, Royal Institute of Technology. Doris Arenander made significant contributions to the style and form of the final version of the book.

Finally to Annika, Adrian, and Jasmine for their understanding and patience during the many (often late night) hours required for preparing this book.

Harold W. Lawson
Lidingö, Sweden

Reference

Wurman, R.S. (1989). *Information Anxiety*, Doubleday, Garden City, New York.

CONTENTS

Forward xix

Preface xxi

Part I Fundamental Issues 1

Chapter 1 Introduction 1
1.1 Philosophies, Paradigms, Models, and Artifacts 4
1.2 Analogical Reasoning 5
1.3 Resources and their Utilization 6
1.4 Human Parallel Processing Environments 8
1.4.1 Human Matrices 8
1.4.2 Models of Production 10
1.4.3 Pandemonium 11
1.5 Common Denominator Concepts and Terminology 13
1.6 Resource Structure Building Blocks 14
1.7 The Process Abstraction 15
1.8 Parallel Processing System Development 16
1.9 Structure and Behavior of Models 17
1.9.1 Model Restrictions 17
1.10 Application versus System Parallelism 18
1.11 Goals and Reasons for Parallel Processing 19
1.11.1 Natural Implementation 19
1.11.2 Performance 19
1.11.3 The x-abilities 21
1.11.4 Fault Detection and Fault Tolerance 21
1.11.5 Making Trade-offs 21

1.12 Fundamental Mapping Philosophies 22
1.12.1 Direct Mapped versus Dynamically Mapped 22
1.13 Classification of Parallel Computer Systems 24
1.13.1 Flynn's Classification 24
1.13.2 Granularity (fine, medium, coarse, and very coarse) 25
1.13.3 Coupling (tightly, closely, and loosely) 25
1.13.4 General Purpose versus Special Purpose 26
1.14 The Fundamental Issues 26

Chapter 2 Real-time Systems 30
2.1 Real-time Applications 32
2.2 Real-time Processes 34
2.2.1 Periodic and Aperiodic Processes 34
2.2.2 Static and Dynamic Processes 34
2.2.3 Relative Importance of Processes 35
2.2.4 Concurrent and Parallel Execution 35
2.3 Requirements for Real-time Systems 36
2.3.1 Predictability of System Behavior 36
2.3.2 Supervision of Timely Process Execution 36
2.3.3 Synchronization and Communication 37
2.3.4 Avoidance of Deadlock 37
2.3.5 Detection and Treatment of Transient Overloads 37
2.3.6 Safety and Security 37
2.4 Varying Views of Development 38
2.5 Real-time Behavioral Description 39
2.5.1 Abstract Behavioral Descriptions 39
2.5.2 Real-time Programming Languages 40
2.6 Contemporary Real-time Paradigms and Models 42
2.6.1 Direct-Mapped Systems 42
2.6.2 Continuous Operation (Cyclic) Systems 43
2.6.3 Asynchronous Event-Driven Systems 45
2.7 Parallel/Distributed Real-time Systems 47
2.8 Schedules for Process Sets 48
2.8.1 Dynamic Scheduling 50
2.8.2 Some Process Priority Problems 51
2.8.3 Scheduling Strategies 51
2.9 Toward Parallel Processing Environments 54
2.9.1 Elements of Uncertainty 54
2.10 Further Challenges 55

Chapter 3 Parallel Processing 58

Introduction 58
3.1 A Common View of Behavior 59
3.2 Generic Parallel Processing Resource Model 60
3.2.1 Memories as Processing Elements 61
3.2.2 Increasing Bandwidth 61
3.2.3 Elephants, Ants, and Beavers 61
3.3 Parallelism and Real-time Processing 62
3.3.1 Schedules, Resources, and Interprocess Communication 63
3.3.2 Nondeterminism 63
3.4 Chapter Structure 64

Hardware Resource Structures 65
3.5 How are Processing Activities Controlled? 66
3.6 What Controls Processing Execution? 68
3.7 What is the Unit of Parallelism? 68
3.7.1 Processes as a Granularity Common Denominator 69
3.7.2 Varying Process Granularities 70
3.8 How is the Interconnection Network Organized? 70
3.9 How are Memory Processes Organized? 71
3.10 Pipelined Systems 72
3.10.1 Wavefront and Systolic Mechanisms 73
3.10.2 Pipelines for DSP Applications 73
3.11 SIMD Architectures 74
3.11.1 PE Arrays 75
3.11.2 Very Long Instruction Words 77
3.12 MIMD Architectures 78
3.12.1 Multiprocessors versus Multicomputers 78
3.12.2 LANs and WANs 79
3.12.3 Some Important MIMD Issues 79
3.12.4 Mutual Exclusion and Synchronization Mechanisms 80
3.13 Other Architectural Variants 82

Behavioral Description 83
3.14 Abstract Models 85
3.15 Description of Hardware Processes 85
3.15.1 Hardware Description Languages 86
3.15.2 Hardware Abstraction Levels 86
3.16 Direct-mapped Systems 88

3.17 Programmed Processes 90
3.17.1 Parallelism Visibility versus Transparency 90
3.17.2 Programming Paradigms 91
3.17.3 Asynchronous versus Synchronous Parallel Programs 95
3.18 Programming of PE arrays 95
3.18.1 Instruction Level 96
3.18.2 Programming Languages 96
3.19 Concurrent Parallel Programming 97
3.19.1 Mutual Exclusion 98
3.19.2 Process Synchronization and Communication 103
3.20 Object-oriented Programming 109
3.21 Functional Programming 110
3.21.1 Lambda Calculus 111
3.22 Logic Programming 112
3.22.1 Concurrent Logic Programs 114
3.22.2 Additional Logic and Functional Approaches 114

Interconnection Networks 118
3.23 Network Properties 120
3.23.1 What Is Connected to What? 120
3.24 Transport Method 121
3.24.1 Circuit Switched 121
3.24.2 Packet Switched 122
3.24.3 Comparison 122
3.25 Network Connections 122
3.25.1 Network Operations 123
3.25.2 Permutations and Switches (Network Building Blocks) 124
3.25.3 Single-stage Network 126
3.25.4 Multistage Network 127
3.25.5 Examples of Pairwise Couplings 128
3.25.6 Comparison 129
3.26 Control of Networks 129
3.26.1 Global Control 129
3.26.2 Local Control 129
3.26.3 Routing in Packet-switched Networks 130
3.26.4 Network Deadlock 130
3.26.5 Alternative Routing 132
3.27 Performance and Cost Properties of Networks 132
3.28 Review of Selected Network Topologies 133
3.28.1 Clos Networks 134
3.28.2 Rings and Trees 134

此为试读,需要完整PDF请访问: www.ertongbook.com