TASK SCHEDULING FOR PARALLEL SYSTEMS

THE WILEY BICENTENNIAL–KNOWLEDGE FOR GENERATIONS

Each generation has its unique needs and aspirations. When Charles Wiley first opened his small printing shop in lower Manhattan in 1807, it was a generation of boundless potential searching for an identity. And we were there, helping to define a new American literary tradition. Over half a century later, in the midst of the Second Industrial Revolution, it was a generation focused on building the future. Once again, we were there, supplying the critical scientific, technical, and engineering knowledge that helped frame the world. Throughout the 20th Century, and into the new millennium, nations began to reach out beyond their own borders and a new international community was born. Wiley was there, expanding its operations around the world to enable a global exchange of ideas, opinions, and know-how.

For 200 years, Wiley has been an integral part of each generation's journey, enabling the flow of information and understanding necessary to meet their needs and fulfill their aspirations. Today, bold new technologies are changing the way we live and learn. Wiley will be there, providing you the must-have knowledge you need to imagine new worlds, new possibilities, and new opportunities.

Generations come and go, but you can always count on Wiley to provide you the knowledge you need, when and where you need it!

William J. Pesce

WILLIAM J. PESCE
PRESIDENT AND CHIEF EXECUTIVE OFFICER

Peter Booth Wiley

PETER BOOTH WILEY
CHAIRMAN OF THE BOARD

TASK SCHEDULING FOR PARALLEL SYSTEMS

Oliver Sinnen
Department of Electrical and Computer Engineering
The University of Aukland
New Zealand

WILEY-INTERSCIENCE
A JOHN WILEY & SONS, INC., PUBLICATION

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data:

Sinnen, Oliver, 1971-
Task scheduling for parallel systems / by Oliver Sinnen.
p. cm.
Includes bibliographical references and index.
978-0-471-73576-2
1. Parallel processing (Electronic computers) 2. Multitasking (Computer science) 3. Computer scheduling. I. Title.
QA76.58.S572 2007
044′.35—dc22 2006052157

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

à Patrícia

CONTENTS

PREFACE

Even though the area of parallel computing has existed for many decades, programming a parallel system is still a challenging problem, much more challenging than programming a single processor system. With the current dual-core and multicore processors from IBM, AMD, Intel, and others, mainstream PCs have entered the realm of parallel systems. The investigation and understanding of the foundations of parallel computing is therefore more important than ever.

One of these foundations is task scheduling. To execute a program consisting of several tasks on a parallel system, the tasks must be arranged in space and time on the multiple processors. In other words, the tasks must be mapped to the processors and ordered for execution. This so-called task scheduling is a very complex problem and crucially determines the efficiency of the parallel system. In fact, task scheduling is an NP-hard problem; that is, an optimal solution generally cannot be found in polynomial time (unless P = NP). This has been motivating the development of many heuristics for its near optimal solution.

This book is devoted to task scheduling for parallel systems. Anyone who gets involved for the first time in task scheduling is overwhelmed by the enormous wealth of heuristics, models, and methods that have been contributed during the last decades. One of my main objectives for this book is to bring order into this jungle of task scheduling. However, the book does not simply categorize and order scheduling heuristics. Instead, it investigates and presents task scheduling by extracting and discussing common models, methods, and techniques, and by setting them into relation. Hence, this book is not a mere survey of scheduling algorithms, but rather an attempt at a consistent and unifying theoretical framework.

Another objective I have with this book is to go beyond the classic approach to task scheduling by studying scheduling under more advanced and accurate system models. These system models consider heterogeneity, contention for communication resources, and involvement of the processor in communication. For efficient and accurate task scheduling, a realistic system model is most crucial. This book is the first publication that discusses advanced system models for task scheduling in a comprehensive form.

Task Scheduling for Parallel Systems is targeted at practicing professionals, researchers, and students. For those who are new to task scheduling, the first chapters carefully introduce parallel systems and their programming, setting task scheduling into the context of the program parallelization process. Practitioners involved in parallel programming will gain an understanding of fundamental aspects of the parallelization process. This knowledge will help them to write more efficient code.

Compiler and parallelization tool developers will benefit from a deeper understanding of the scheduling problem, which is also a generalization of many other problems they face (e.g., loop scheduling). A chapter on graph models promotes the understanding of these relations. For task scheduling researchers, this book serves as a comprehensive reference, based on a unifying framework. The research community will especially value the later chapters on advanced scheduling and sophisticated scheduling models. Graduate students of parallel computing and compiler courses can use this book to thoroughly study task scheduling, which is supported by the exercises at the end of each chapter. The extensive index and the large number of bibliographic references make this book a valuable tool for everybody interested in task scheduling.

For a brief introduction to task scheduling and an overview of this book, including a short summary of each chapter, refer to Chapter 1.

OLIVER SINNEN
The University of Aukland
New Zealand

ACKNOWLEDGMENTS

This book has its roots in my PhD thesis, which I presented at the Instituto Superior Técnico, Technical University of Lisbon in Portugal. I therefore want to sincerely thank Prof. Leonel Sousa again for his academic advice, the fruitful discussions, and his criticism during that time. In this context I also thank Prof. Frode Sandnes, Oslo University College, Norway, who introduced me to genetic algorithms and has been an invaluable research colleague for many years.

I would like to thank The University of Auckland for the opportunity to write this book while being a lecturer in the Department of Electrical and Computer Engineering.

Cover design: Digital image of a painting created by Patrícia Raquel de Vasconcelos da Silva (2006).

OLIVER SINNEN
The University of Aukland
New Zealand