

Lecture Notes in Computer Science

949

Edited by G. Goos, J. Hartmanis and J. van Leeuwen

Advisory Board: W. Brauer D. Gries J. Stoer

Dror G. Feitelson Larry Rudolph (Eds.)

Job Scheduling Strategies for Parallel Processing

IPPS '95 Workshop

Santa Barbara, CA, USA, April 25, 1995

Proceedings



Springer

Series Editors

Gerhard Goos, Universität Karlsruhe, Germany

Juris Hartmanis, Cornell University, NY, USA

Jan van Leeuwen, Utrecht University, The Netherlands

Volume Editors

Dror G. Feitelson

Thomas J. Watson Research Center, IBM

P.O. Box 218, Yorktown Heights, NY 10598, USA

Larry Rudolph

Institute of Computer Science, The Hebrew University

91904 Jerusalem, Israel

Cataloging-in-Publication data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Job scheduling strategies for parallel processing : proceedings

/ IPPS '95 workshop, Santa Barbara, CA, USA, April 25, 1995.

Dror G. Feitelson ; Larry Rudolph (ed.). - Berlin ; Heidelberg ;

New York : Springer, 1995

(Lecture notes in computer science ; Vol. 949)

ISBN 3-540-60153-8

NE: Feitelson, Dror G. [Hrsg.]; IPPS <9, 1995, Santa Barbara, Calif.>;

GT

CR Subject Classification (1991): D.4, D.1.3, F.2.2, C.1.2

ISBN 3-540-60153-8 Springer-Verlag Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

© Springer-Verlag Berlin Heidelberg 1995

Printed in Germany

Typesetting: Camera-ready by author

SPIN 10486444 06/3142 - 5 4 3 2 1 0 Printed on acid-free paper

Preface

This volume contains the papers presented at the workshop on Job Scheduling Strategies for Parallel Processing held in Santa Barbara, California, as a prelude to the IPPS'95 conference. All the papers have gone through the usual refereeing process with the full version being read and evaluated by at least five members of the program committee. We would like to take this opportunity to thank the program committee, Nawaf Bitar, David Black, Jim Cownie, Allan Gottlieb, Mal Kalos, Bradley C. Kuszmaul, Miron Livny, Virginia Lo, Reagan Moore, Charlie Smith, Mark Squillante, Bernard Traversat, and John Zahorjan, for an excellent job. Thanks are also due to the authors for their submissions, presentations, and final revisions for this volume. Finally, we would like to thank Marc Snir and IBM for printing an early copy of these proceedings for distribution at the workshop.

As multi-user parallel supercomputers become more widespread, job scheduling takes on a crucial role. The number of users of parallel supercomputers is growing at an even faster pace and so there is an increasing number of users who must share a parallel computer's resources. Job scheduling strategies must address this need.

There is a spectrum of groups that are interested in job scheduling strategies for parallel processors. At one end are the vendors of parallel supercomputers who supply the scheduling software for managing jobs on their machines. In the middle are researchers in academia, National Labs, and industrial research labs who propose new scheduling strategies and methods for evaluating and comparing them. At the other end of the spectrum are the users and providers of parallel processing resources who have a set of demands and requirements.

The workshop and the papers in this proceedings have been organized along these three groups. After an initial overview come the system implementations — what has been done. The middle section is devoted to evaluation techniques and new strategies — what could be done. The final section is mostly requirements and attempts at solutions for job scheduling, mostly from the point of view of major supercomputing installations — what should be done.

At the workshop there were many interesting discussions between people in the three groups (but we were too busy to be part of many of them). We were encouraged by this since we believe it is important to increase communication so that academics work on the right problems and vendors and computation centers make the best use of the novel solutions. We hope these proceedings help parallel supercomputing to satisfy its fundamental goal of satisfying the needs of the user.

Dror Feitelson
Larry Rudolph
New York, May 1995

Contents

Parallel Job Scheduling: Issues and Approaches	1
Dror G. Feitelson and Larry Rudolph	
Scheduling on the Tera MTA	19
Gail Alverson, Simon Kahan, Richard Korry, Cathy McCann, and Burton Smith	
A Scalable Multi-Discipline, Multiple-Processor Scheduling Framework for IRIX	45
James M. Barton and Nawaf Bitar	
Scheduling to Reduce Memory Coherence Overhead on Coarse-Grain Multiprocessors	70
Christopher Connelly and Carla Schlatter Ellis	
Time Space Sharing Scheduling and Architectural Support	92
Atsushi Hori, Takashi Yokota, Yutaka Ishikawa, Shuichi Sakai, Hiroki Konaka, Munenori Maeda, Takashi Tomokiyo, Jörg Nolte, Hiroshi Matsuoka, Kazuaki Okamoto, and Hideo Hirono	
Demand-Based Coscheduling of Parallel Jobs on Multiprogrammed Multiprocessors	106
Patrick G. Sobalvarro and William E. Wehl	
Multiprocessor Scheduling for High-Variability Service Time Distributions	127
Eric W. Parsons and Kenneth C. Sevcik	
The Interaction between Memory Allocation and Adaptive Partitioning in Message-Passing Multicomputers	146
Sanjeev K. Setia	
Analysis of Non-Work-Conserving Processor Partitioning Policies	165
E. Rosti, E. Smirni, G. Serazzi, and L. W. Dowdy	
Loop-Level Process Control: An Effective Processor Allocation Policy for Multiprogrammed Shared-Memory Multiprocessors	182
Kelvin K. Yue and David J. Lilja	
A Microeconomic Scheduler for Parallel Computers	200
Ion Stoica, Hussein Abdel-Wahab, and Alex Pothen	
On the Benefits and Limitations of Dynamic Partitioning in Parallel Computer Systems	219
Mark S. Squillante	

Intelligent Fuzzy Control to Augment Scheduling Capabilities of Network Queuing Systems	239
Oscar Kipersztok and James C. Patterson	
Parallel Processing on Dynamic Resources with CARMi	259
Jim Pruyne and Miron Livny	
Job Scheduling Under the Portable Batch System	279
Robert L. Henderson	
The ANL/IBM SP Scheduling System	295
David A. Lifka	
Requirements of the Cornell Theory Center for Resource Management and Process Scheduling	304
Marcy E. Rosenkrantz, David J. Schneider, Roslyn Leibensperger, Melinda Shore, and John Zollweg	
Job Management Requirements for NAS Parallel Systems and Clusters ...	319
William Saphir, Leigh Ann Tanner, and Bernard Traversat	
Job Characteristics of a Production Parallel Scientific Workload on the NASA Ames iPSC/860	337
Dror G. Feitelson and Bill Nitzberg	
Author Index	361