Lecture Notes in Computer Science 1162
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 '96 Workshop
Honolulu, Hawaii, April 16, 1996
Proceedings

Springer

Series Editors

Gerhard Goos, Karlsruhe University, Germany

Juris Hartmanis, Cornell University, NY, USA

Jan van Leeuwen, Utrecht University, The Netherlands

Volume Editors

Dror G. Feitelson
The Hebrew University, Institute of Computer Science
Givat Ram Campus, Ross Building, 91904 Jerusalem, Israel

Larry Rudolph
MIT, Lab. for Computer Science, NE43-228
545 Technology Square, Cambridge, MA02139, USA
E-mail: {feit,rudolph}@cs.huji.ac.il

Cataloging-in-Publication data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Job scheduling strategies for parallel processing : proceedings / IPPS '96 workshop, Honolulu, Hawaii, April 1996. Dror G. Feitelson ; Larry Rudolph (ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Budapest ; Hong Kong ; London ; Milan ; Paris ; Santa Clara ; Singapore ; Tokyo : Springer, 1996
(Lecture notes in computer science ; Vol. 1162)
ISBN 3-540-61864-3
NE: Feitelson, Dror G. [Hrsg.]; International Parallel Processing Symposium <10, 1996, Honolulu, Hawaii>; GT

CR Subject Classification (1991): D.4, D.1.3,F.2.2, C.1.2, B.2.1, B.6.1, F1.2, C.2

ISSN 0302-9743
ISBN 3-540-61864-3 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 1996
Printed in Germany

Typesetting: Camera-ready by author
SPIN 10549739 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 Honolulu, Hawaii, as a prelude to the IPPS'96 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, Scott Hahn, Mal Kalos, Phil Krueger, Richard Lagerstrom, Miron Livny, Virginia Lo, Reagan Moore, Ken Sevcik, 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. We would like to thank the MIT Laboratory for Computer Science and the Computer Science Institute at Hebrew University for the use of their facilities in preparation of these proceedings.

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.

This is the second occurrence of the workshop. The previous workshop was held a year earlier, as part of the IPPS'95 conference in Santa Barbara. The proceedings of that workshop have been published as Springer-Verlag Lecture Notes in Computer Science Vol. 949.

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 achieve its fundamental goal of satisfying the needs of the user.

Jerusalem, August 1996

Dror Feitelson
Larry Rudolph

Contents

Towards Convergence in Job Schedulers for Parallel Supercomputers 1
Dror G. Feitelson and Larry Rudolph

Workload Evolution on the Cornell Theory Center IBM SP2 27
Steven Hotovy

The EASY - LoadLeveler API Project 41
Joseph Skovira, Waiman Chan, Honbo Zhou, and David Lifka

A Batch Scheduler for the Intel Paragon MPP System with a Non-contiguous Node Allocation Algorithm 48
Michael Wan, Reagan Moore, George Kremenek, and Ken Steube

Architecture-Independent Request-Scheduling with Tight Waiting-Time Estimations 65
Jörn Gehring and Friedhelm Ramme

Packing Schemes for Gang Scheduling 89
Dror G. Feitelson

A Gang Scheduling Design for Multiprogrammed Parallel Computing Environments 111
Fang Wang, Hubertus Franke, Marios Papaefthymiou, Pratap Pattnaik, Larry Rudolph, and Mark S. Squillante

Implementation of Gang-Scheduling on Workstation Cluster 126
Atsushi Hori, Hiroshi Tezuka, Yutaka Ishikawa, Noriyuki Soda, Hiroki Konaka, and Munenori Maeda

Managing Checkpoints for Parallel Programs 140
Jim Pruyne and Miron Livny

Using Runtime Measured Workload Characteristics in Parallel Processor Scheduling 155
Thu D. Nguyen, Raj Vaswani, and John Zahorjan

Parallel Application Characteristics for Multiprocessor Scheduling Policy Design 175
Thu D. Nguyen, Raj Vaswani, and John Zahorjan

Dynamic vs. Static Quantum-Based Parallel Processor Allocation 200
Su-Hui Chiang and Mary Vernon

Dynamic versus Adaptive Processor Allocation Policies for Message Passing Parallel Computers: An Empirical Comparison 224
Jitendra D. Padhye and Lawrence W. Dowdy

Dynamic Partitioning in Different Distributed-Memory Environments 244
Nayeem Islam, Andreas Prodromidis, and Mark S. Squillante

Locality Information Based Scheduling in Shared Memory Multiprocessors 271
Frank Bellosa

Author Index 291