

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Friedemann Mattern

ETH Zurich, Switzerland

John C. Mitchell

Stanford University, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz

University of Bern, Switzerland

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

University of Dortmund, Germany

Madhu Sudan

Massachusetts Institute of Technology, MA, USA

Demetri Terzopoulos

New York University, NY, USA

Doug Tygar

University of California, Berkeley, CA, USA

Moshe Y. Vardi

Rice University, Houston, TX, USA

Gerhard Weikum

Max-Planck Institute of Computer Science, Saarbruecken, Germany

Dror Feitelson Larry Rudolph
Uwe Schwiegelshohn (Eds.)

Job Scheduling Strategies for Parallel Processing

10th International Workshop, JSSPP 2004
New York, NY, USA, June 13, 2004
Revised Selected Papers

Volume Editors

Dror Feitelson
The Hebrew University
School of Computer Science and Engineering
91904 Jerusalem, Israel
E-mail: feit@cs.huji.ac.il

Larry Rudolph
CSAIL – Massachusetts Institute of Technology
32 Vassar Street, Cambridge, MA 02139, USA
E-mail: rudolph@csail.mit.edu

Uwe Schwiegelshohn
University of Dortmund
Computer Engineering Institute
44221 Dortmund, Germany
E-mail: uwe.schwiegelshohn@udo.edu

Library of Congress Control Number: 2005925176

CR Subject Classification (1998): D.4, D.1.3, F.2.2, C.1.2, B.2.1, B.6, F.1.2

ISSN	0302-9743
ISBN-10	3-540-25330-0 Springer Berlin Heidelberg New York
ISBN-13	978-3-540-25330-3 Springer 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. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media
springeronline.com

© Springer-Verlag Berlin Heidelberg 2005
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Boller Mediendesign
Printed on acid-free paper SPIN: 11407522 06/3142 5 4 3 2 1 0

Preface

This volume contains the papers presented at the 10th Anniversary Workshop on Job Scheduling Strategies for Parallel Processing. The workshop was held in New York City, on June 13, 2004, at Columbia University, in conjunction with the SIGMETRICS 2004 conference.

Although it is a workshop, the papers were conference-reviewed, with the full versions being read and evaluated by at least five and usually seven members of the Program Committee. We refer to it as a workshop because of the very fast turnaround time, the intimate nature of the actual presentations, and the ability of the authors to revise their papers after getting feedback from workshop attendees. On the other hand, it was actually a conference in that the papers were accepted solely on their merits as decided upon by the Program Committee.

We would like to thank the Program Committee members, Su-Hui Chiang, Walfredo Cirne, Allen Downey, Eitan Frachtenberg, Wolfgang Gentzsch, Allan Gottlieb, Moe Jette, Richard Lagerstrom, Virginia Lo, Reagan Moore, Bill Nitzberg, Mark Squillante, and John Towns, 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 the MIT Computer Science and Artificial Intelligence Laboratory (CSAIL), The Hebrew University, and Columbia University for the use of their facilities in the preparation of the workshop and these proceedings.

This year saw a continued interest in scheduling in grid and cluster environments, with a growing representation of real-system issues such as workload studies, network topology issues, and the effect of failures. At the same time, there was also a strong representation of research relating to classical multiprocessor systems, and lively discussions contrasting the academic point of view with that of administrators of ‘real’ systems. We hope that the papers in this volume capture this range of interests and approaches, and that you, the reader, find them interesting and useful.

This was the tenth annual workshop in this series, which reflects a consistent and ongoing interest; the organizers believe that the workshop satisfies a real need. The proceedings of previous workshops are available from Springer as LNCS volumes 949, 1162, 1291, 1459, 1659, 1911, 2221, 2537, and 2862 (and since 1998 they have also been available online). We look forward to the next workshop in 2005, and perhaps even to the next decade of workshops!

September 2004

Dror Feitelson
Larry Rudolph
Uwe Schwiegelshohn

Table of Contents

Parallel Job Scheduling — A Status Report	1
<i>Dror G. Feitelson, Larry Rudolph, and Uwe Schwiegelshohn</i>	
Scheduling on the Top 50 Machines	17
<i>Carsten Ernemann, Martin Krogmann, Joachim Lepping, and Ramin Yahyapour</i>	
Parallel Computer Workload Modeling with Markov Chains	47
<i>Baiyi Song, Carsten Ernemann, and Ramin Yahyapour</i>	
Enhancements to the Decision Process of the Self-Tuning dynP Scheduler	63
<i>Achim Streit</i>	
Reconfigurable Gang Scheduling Algorithm	81
<i>Luís Fabrício Wanderley Góes and Carlos Augusto Paiva da Silva Martins</i>	
Time-Critical Scheduling on a Well Utilised HPC System at ECMWF Using Loadleveler with Resource Reservation	102
<i>Graham Holt</i>	
Inferring the Topology and Traffic Load of Parallel Programs Running in a Virtual Machine Environment	125
<i>Ashish Gupta and Peter A. Dinda</i>	
Multi-toroidal Interconnects: Using Additional Communication Links to Improve Utilization of Parallel Computers	144
<i>Yariv Aridor, Tamar Domany, Oleg Goldshmidt, Edi Shmueli, Jose E. Moreira, and Larry Stockmeier</i>	
Costs and Benefits of Load Sharing in the Computational Grid	160
<i>Darin England and Jon B. Weissman</i>	
Workload Characteristics of a Multi-cluster Supercomputer	176
<i>Hui Li, David Groep, and Lex Walters</i>	
A Dynamic Co-allocation Service in Multicluster Systems	194
<i>Jove M.P. Sinaga, Hashim H. Mohamed, and Dick H.J. Epema</i>	
Exploiting Replication and Data Reuse to Efficiently Schedule Data-Intensive Applications on Grids	210
<i>Elizeu Santos-Neto, Walfredo Cirne, Francisco Brasileiro, and Aliandro Lima</i>	

Performance Implications of Failures in Large-Scale Cluster Scheduling . .	233
<i>Yanyong Zhang, Mark S. Squillante, Anand Sivasubramaniam, and Ramendra K. Sahoo</i>	
Are User Runtime Estimates Inherently Inaccurate?	253
<i>Cynthia Bailey Lee, Yael Schwartzman, Jeniffer Hardy, and Allen Snively</i>	
Improving Speedup and Response Times by Replicating Parallel Programs on a SNOW	264
<i>Gaurav D. Ghare and Scott T. Leutenegger</i>	
LOMARC — Lookahead Matchmaking for Multi-resource Coscheduling . .	288
<i>Angela C. Sodan and Lei Lan</i>	
Author Index	317