Lecture Notes in Computer Science **10773**

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

More information about this series at http://www.springer.com/series/7407

Dalibor Klusáček · Walfredo Cirne
Narayan Desai (Eds.)

Job Scheduling Strategies for Parallel Processing

21st International Workshop, JSSPP 2017
Orlando, FL, USA, June 2, 2017
Revised Selected Papers

Springer

Editors
Dalibor Klusáček
CESNET
Prague
Czech Republic

Narayan Desai
Google
Seattle, WA
USA

Walfredo Cirne
Google
Mountain View, CA
USA

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Computer Science
ISBN 978-3-319-77397-1 ISBN 978-3-319-77398-8 (eBook)
https://doi.org/10.1007/978-3-319-77398-8

Library of Congress Control Number: 2018934363

LNCS Sublibrary: SL1 – Theoretical Computer Science and General Issues

Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG
part of Springer Nature
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This volume contains the papers presented at the 21st Workshop on Job Scheduling Strategies for Parallel Processing that was held in Orlando (FL), USA, on June 2, 2017, in conjunction with the 31st IEEE International Parallel and Distributed Processing Symposium (IPDPS 2017). The proceedings of previous workshops are also available from Springer as LNCS volumes 949, 1162, 1291, 1459, 1659, 1911, 2221, 2537, 2862, 3277, 3834, 4376, 4942, 5798, 6253, 7698, 8429, 8828, and 10353.

This year 20 papers were submitted to the workshop, of which we accepted ten. All submitted papers went through a complete review process, with the full version being read and evaluated by an average of four reviewers. We would like to especially thank our Program Committee members and additional reviewers for their willingness to participate in this effort and their excellent, detailed reviews.

From the very beginning, JSSPP has strived to balance practice and theory in its program. This combination has been repeatedly shown to provide a rich environment for technical debate about scheduling approaches. This year, building on this long tradition, JSSPP also welcomed papers providing descriptions of *open problems in large-scale scheduling*. A lack of real-world data often hampers the ability of the research community to engage with scheduling problems in a way that has real world impact. Our goal in this new venue was to build a bridge between the production and research worlds, in order to facilitate direct discussions, collaborations, and impact.

It was our pleasure that out of the ten accepted papers, two directly address the novel "open problems" track, sharing valuable insights into production systems, their workloads, usage patterns, and corresponding scheduling challenges. In their paper, Allcock et al. present details on job scheduling at the Argonne Leadership Computing Facility (ALCF). The paper described the specific scheduling goals and constraints, analyzed the workload traces from the petascale supercomputer Mira, and discussed the upcoming challenges at ALCF. Klusáček and Parák present a detailed analysis of a shared virtualized computing infrastructure that is used to provide grid and cloud computing services. In their work, they analyzed the differences between cloud and grid workloads and addressed some of the problems the infrastructure is facing, such as (un)fairness or problematic resource reclaiming.

In 1995, JSSPP was the venue where the seminal and widely used backfilling algorithm was presented for the first time. Now, 22 years after its introduction, many researchers are still focusing on improving its performance. This year we had two papers that directly focus on improving the performance of backfilling. Lelong, Reis, and Trystram propose a framework to evaluate the impact of reordering job queues using various policies in order to improve on average/maximum wait time. N'takpé and Suter evaluate a model where a part of the job workload is not sensitive to waiting as long as it is completed before a given deadline. This allowed them to perform some interesting optimizations for regular jobs in order to decrease the average wait time and slowdown.

Wang et al. focus on a somehow similar problem of supporting priority execution for high-priority real-time jobs while minimizing the delays for ordinary workloads in a classic batch scheduling scenario. Their solution investigated several techniques starting from a plain high-priority queue to somewhat more advanced approaches including pre-emption and application checkpointing. Friese et al. present a detailed methodology using a genetic algorithm for cost-efficient resource selection when scheduling complex scientific workflows with uncertainties in forecasted demands on distributed computing platforms such as "pay-per-use" public clouds.

Lohrmann et al. focus on optimizing the execution of complex I/O critical simulations that are performed using iterative workflows. To minimize I/O delays, in situ processing is commonly used to minimize the need for time-consuming disk operations. For this purpose the authors extended the Henson cooperative multi-tasking system that enables multiple distinct codes to run on the same node and share memory to speed up computations. Their major extension is a scheduler for Henson, which is used to schedule iterative trials of a complex simulation. Trials results are used as an input into a relaxed (computationally cheap) surrogate model that generates new, refined parameters for consecutive expensive trials. This iterative approach is used to increase the chance that the expensive simulation converges quickly.

While the majority of batch schedulers are based on job queues, there are few honorable mentions of pure planning systems, where each job is planned ahead upon its arrival, i.e., a complete schedule about the future resource usage is computed and made available to the users. In his paper, Axel Keller presents a detailed description of such a system called OpenCCS, focusing in detail on data structures and a heuristic that are used to plan and map arbitrary resources in complex combinations while applying time-dependent constraints.

Two papers focus on evaluating the system performance using newly proposed simulators and benchmarks, addressing the needs of current HPC systems, where both the workload and the infrastructure become more complex and heterogeneous, thus urgently requiring more advanced scheduling approaches. Rodrigo et al. propose a novel scheduler simulation framework (ScSF) that provides capabilities for workload modeling and generation, system simulation (using embedded Slurm simulator), comparative workload analysis, and experiment orchestration. This simulator is designed to be run over a distributed computing infrastructure facilitating large-scale tests. Lopez et al. present the Dynamic Job Scheduling Benchmark (DJSB), which is a novel tool allowing system administrators to evaluate the impact of dynamic resource (re)allocations between running jobs on the overall system performance. They use a set of experiments from the MareNostrum supercomputer to demonstrate how DJSB can be used to evaluate the impact of different dynamic resource management approaches on each job/application individually, as well as the overall dynamics of the system.

Enjoy the reading!

We hope you can join us at the next JSSPP workshop, this time in Vancouver, Canada, on May 25, 2018.

November 2017

Walfredo Cirne
Narayan Desai
Dalibor Klusáček

Organization

Workshop Organizers

Walfredo Cirne	Google, USA
Narayan Desai	Google, USA
Dalibor Klusáček	CESNET, Czech Republic

Program Committee

Henri Casanova	University of Hawaii at Manoa, USA
Julita Corbalan	Barcelona Supercomputing Center, Spain
Carlo Curino	Microsoft, USA
Hyeonsang Eom	Seoul National University, South Korea
Dick Epema	Delft University of Technology, The Netherlands
Dror Feitelson	Hebrew University, Israel
Liana Fong	IBM T. J. Watson Research Center, USA
Eitan Frachtenberg	Facebook, USA
Alfredo Goldman	University of Sao Paulo, USA
Allan Gottlieb	New York University, USA
Zhiling Lan	Illinois Institute of Technology, USA
Bill Nitzberg	Altair, USA
P-O Östberg	Umeå University, Sweden
Larry Rudolph	Two Sigma, USA
Uwe Schwiegelshohn	TU Dortmund University, Germany
Leonel Sousa	Universidade de Lisboa, Portugal
Mark Squillante	IBM, USA
Wei Tang	Google, USA
Ramin Yahyapour	University of Göttingen, Germany

Additional Reviewers

Helder Duarte
João F. D. Guerreiro
Diogo Marques
Jiaqi Yan
Xu Yang

Contents