Lecture Notes in Computer Science 13289

Founding Editors

Gerhard Goos

Karlsruhe Institute of Technology, Karlsruhe, Germany

Juris Hartmanis

Cornell University, Ithaca, NY, USA

Editorial Board Members

Elisa Bertino

Purdue University, West Lafayette, IN, USA

Wen Gao

Peking University, Beijing, China

Bernhard Steffen

TU Dortmund University, Dortmund, Germany

Moti Yung

Columbia University, New York, NY, USA

More information about this series at https://link.springer.com/bookseries/558

Ana-Lucia Varbanescu · Abhinav Bhatele · Piotr Luszczek · Baboulin Marc (Eds.)

High Performance Computing

37th International Conference, ISC High Performance 2022 Hamburg, Germany, May 29 – June 2, 2022 Proceedings



Editors
Ana-Lucia Varbanescu
University of Twente
Enschede, The Netherlands

Piotr Luszczek D University of Tennessee Knoxville, TN, USA Abhinav Bhatele University of Maryland College Park, MD, USA

Baboulin Marc Université Paris-Saclay Orsay, France

ISSN 0302-9743 ISSN 1611-3349 (electronic) Lecture Notes in Computer Science ISBN 978-3-031-07311-3 ISBN 978-3-031-07312-0 (eBook) https://doi.org/10.1007/978-3-031-07312-0

© Springer Nature Switzerland AG 2022, corrected publication 2023

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

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

Preface

ISC High Performance Computing—formerly known as the International Supercomputing Conference—was founded in 1986 as the Supercomputer Seminar. Originally organized by Hans Meuer, Professor of Computer Science at the University of Mannheim, and former director of its computer center, the 1986 edition of the seminar brought together a group of 81 scientists and industrial partners who shared an interest in high-performance computing (HPC). Since then, the annual conference has become a major international event within the HPC community, growing beyond its humble beginnings, and moving out of Mannheim into other cities throughout the years: Frankfurt, Heidelberg, Dresden, Hamburg, Leipzig, and this year back to Hamburg. Prior to the coronavirus pandemic, the conference had seen a steady increase in the number of submissions of high-quality research papers and corresponding growth in the number of conference attendees. Benefiting from the improving health indicators, ISC-HPC 2022 was again held in person in Hamburg.

The call for papers to ISC-HPC 2022 was issued in Fall 2021, inviting the researchers, developers, and practitioners alike to submit their latest results for consideration for one of the five conference tracks: (1) Architecture, Networks, and Storage; (2) HPC Algorithms and Applications; (3) Machine Learning, AI, and Emerging Technologies; (4) Performance Modeling, Evaluation, and Analysis; and (5) Programming Environments and Systems Software. In all, 53 full submissions were received from authors all over the world. The Research Papers Program Committee consisted of 71 members from 18 countries. After initial reviews were completed, a rebuttal process offered authors an opportunity to respond to reviewers' questions and help clarify issues the reviewers might have had. A virtual Program Committee meeting was held to discuss all the papers and to finalize consensus on the papers. Finally, the committee selected 18 papers for publication.

For the past several years, the ISC-HPC conference has presented an ISC-sponsored award to encourage outstanding research in HPC and to honor the overall best research paper submitted to the conference. Four years ago, this annual award was renamed in memory of the late Dr. Hans Meuer, who was general chair of the ISC-HPC conference from 1986 through 2014, and a co-founder of the TOP500 project. This year, from all research papers submitted, the Best Paper Committee selected the best paper based on its technical merit, its novelty, and impact on the HPC community. During a live ceremony, the following paper was awarded the Hans Meuer Award: *Remote OpenMP Offloading* by Atmn Patel from University of Waterloo, Canada, and Johannes Doerfert from Argonne National Laboratory, USA. The paper extended the canonical scope of OpenMP, which is traditionally confined to a shared memory domain, by utilizing the standard's modern features to offload the computational workload to GPU accelerators housed in remote cluster nodes. The Best Paper Committee appreciated the paper's unique combination of OpenMP familiarity and ease of use with accelerated and distributed computing that were accompanied by analysis of scaling capabilities. The winning paper also showed the

Preface

vi

solution's versatility and how it could use multiple transport layers, each of which offers a different set of trade-offs between performance, portability, and scalability potential.

As the chairs of the Research Papers Committee, we would like to express our gratitude to our colleagues for submitting high-quality papers to all five ISC-HPC scientific tracks. Also, we wish to extend our thanks to the track, area, and conflict chairs, as well as the members of the Best Paper Committee, and finally to the Research Papers Committee that provided the reviews and manuscript evaluation throughout the submission stages. We hope to express our thanks in person during this year's meeting and upcoming ISC-HPC 2023.

May 2022

Ana Lucia Varbanescu Abhinay Bhatele

Organization

Program Chair

Keren Bergman Columbia University, USA

Program Deputy Chair

John Shalf Lawrence Berkeley National Laboratory, USA

Research Papers Program Committee

Research Papers Chairs

Ana Lucia Varbanescu (Chair) University of Amsterdam, Netherlands

Abhinav Bhatele (Deputy Chair) University of Maryland, USA

Architecture, Networks, and Storage

Jay Lofstead (Chair) Sandia National Laboratories, USA Edson Borin University of Campinas, Brazil

Elsa Gonsiorowski Lawrence Livermore National Laboratory, USA

Mozhgan Kabiri Chimeh NVIDIA, UK

Nectarios Koziris National Technical University of Athens, Greece

Michael Kuhn Otto von Guericke University Magdeburg,

Germany

Jay Lofstead Sandia National Laboratories, USA

Preeti Malakar Indian Institute of Technology Kanpur, India

Dhabaleswar Panda Ohio State University, USA

Guangming Tan Institute of Computing Technology (ICT), China

Osamu Tatebe University of Tsukuba, Japan

Carsten Trinitis Technical University of Munich, Germany

Venkatram Vishwanath Argonne National Laboratory, USA

HPC Algorithms and Applications

Didem Unat (Chair) Koç University, Turkey Sameh Abdulah KAUST, Saudi Arabia

Mehmet Belviranli Colorado School of Mines, USA

Xing Cai Simula Research Laboratory and University of

Oslo, Norway

Organization

viii

Lin Gan Tsinghua University and National

Supercomputing Center in Wuxi, China

Clemens Grelck University of Amsterdam, Netherlands

Fuerlinger Karl Ludwig Maximilian University Munich (LMU),

Germany

Kamer Kaya Sabancı University, Turkey Simon McIntosh-Smith University of Bristol, UK Gabriel Noaje NVIDIA, Singapore

Lena Oden Fernuniversität in Hagen and Forschungszentrum

Jülich GMBH, Germany

Johann Rudi Argonne National Laboratory, USA

Tuğba Torun Koç University, Turkey

Miwako Tsuji RIKEN, Japan

Machine Learning, AI, and Emerging Technologies

Theodore L. Willke (Chair)

Nikoli Dryden

Gurbinder Gill

Intel Corporation, USA

ETH Zurich, Switzerland

Katana Graph Inc., USA

Jiajia LiWilliam and Mary College, USAMaryam Mehri DehnaviUniversity of Toronto, CanadaBogdan NicolaeArgonne National Laboratory, USA

Mostofa Patwary NVIDIA, USA Shaden Smith Microsoft, USA

Edgar Solomonik University of Illinois at Urbana-Champaign, USA

Sofia Vallecorsa CERN, Switzerland

Abhinay Vishnu AMD, USA

Yang You National University of Singapore, Singapore

Performance Modeling, Evaluation, and Analysis

Nathan Tallent (Chair) Pacific Northwest National Laboratory, USA

Ivy B. Peng Lawrence Livermore National Laboratory, USA

Alexandru Calotoiu ETH Zürich, Germany

Marc Casas Barcelona Supercomputing Center, Spain

Tom Deakin University of Bristol, UK

Seyong Lee ORNL, USA

Simon McIntosh-Smith University of Bristol, UK
Xiaozhu Meng Rice University, USA

Bernd MohrJuelich Supercomputing Centre, GermanyScott PakinLos Alamos National Laboratory, USAXian-He SunIllinois Institute of Technology, USA

Jidong Zhai Tsinghua University, China

Tianwei Zhang Nanyang Technological University, Singapore

Programming Environments and Systems Software

Michele Weiland (Chair) EPCC, University of Edinburgh, UK Bilel Hadri KAUST Supercomputing Laboratory,

Saudi Arabia

Guido Juckeland HZDR, Germany

Michael Klemm AMD and OpenMP ARB, Germany

Pouya Kousha Ohio State University, USA

John L. inford Arm, USA

István Z. Reguly Pázmány Péter Catholic University, Hungary

Harvey Richardson Hewlett Packard Enterprise, UK

Martin Ruefenacht Leibniz Supercomputing Centre, Germany

Roxana Rusitoru Arm, UK

Thomas R. W. Scogland Lawrence Livermore National Laboratory, USA

Simon Smart ECMWF, UK

Hiroyuki Takizawa Tohoku University, Japan

Christian Terboven RWTH Aachen University, Germany Justs Zarins EPCC, University of Edinburgh, UK

Birds of a Feather Committee

Roman Wyrzykowski (Chair) Czestochowa University of Technology, Poland

Iosif Meyerov (Deputy Chair) Lobachevsky State University of Nizhni

Novogorod, Russia

Michael Bader Technical University of Munich, Germany Claudia Blaas-Schenner TU Wien, VSC Research Center, Austria

Dominik Göddeke University of Stuttgart, Germany

Aleksandar Ilic INESC-ID and Universidade de Lisboa, Portugal Jacek Kitowski AGH University of Science and Technology,

Poland

Dieter Kranzlmueller Ludwig Maximilian University Munich (LMU)

and Leibniz Rechenzentrum, Germany

Carola Kruse Centre Européen de Recherche et de Formation

Avancée en Calcul Scientifique (CERFACS),

France

Krzysztof Kurowski Poznań Supercomputing and Networking Center,

Poland

Marco Lapegna University of Naples Federico II, Italy

Simon McIntosh-Smith University of Bristol, UK

Iosif Meyerov Lobachevsky State University of Nizhni

Novogorod, Russia

Koji Nakano Hiroshima University, Japan

Gabriel Oksa Slovak Academy of Sciences, Slovakia Dana Petcu West University of Timisoara, Romania Antonio I Peña Barcelona Supercomputing Center, Spain

Thomas Rauber University of Bayreuth, Germany

Lubomir Riha IT4Innovations National Supercomputing Center

and Technical University of Ostrava,

Czech Republic

Masha Sosonkina Old Dominion University, USA

Vladimir Stegailov Higher School of Economics and JIHT RAS,

Russia

Dave Turner Kansas State University, USA Bora Ucar CNRS and ENS-Lyon, France

Project Posters Committee

Christian Perez (Chair) Inria, France

Are Magnus Bruaset Simula Research Laboratory, Norway

(Deputy Chair)

Marco Aldinucci University of Torino, Italy

Bartosz Bosak Poznań Supercomputing and Networking Center,

Poland

Nick Brown EPCC, University of Edinburgh, UK

Theodoros Christoudias The Cyprus Institute, Cyprus

Andrew Ensor Auckland University of Technology, New Zealand

Ana Gainaru Oak Ridge National Laboratory, USA

Andra Hugo Apple, France

Kamer Kaya Sabancı University, Turkey

Francesc Lordan Gomis Barcelona Supercomputing Center, Spain Sano Centre for Computational Medicine, Maciej Malawski Institute of Computer Science AGH, Poland

Kengo Nakajima University of Tokyo and RIKEN, Japan Bogdan Nicolae Argonne National Laboratory, USA

Eric Petit Intel. France Phil Ridley Arm, UK

Jonathan Rouzaud-Cornabas Inria and INSA de Lyon, France

Kentaro Sano RIKEN, Japan Francieli Zanon Boito Inria, France

Ameli Chi Zhou Shenzhen University, China

Research Posters Committee

Aparna Chandramowlishwaran UCL USA

(Chair)

Hartwig Anzt Karlsruhe Institute of Technology, Germany and

University of Tennessee, USA

Maryam Mehri Dehnavi University of Toronto, Canada Jee Choi University of Oregon, USA

Ana Gainaru Oak Ridge National Laboratory, USA Lin Gan Tsinghua University and National

Supercomputing Center in Wuxi, China

Amal Khabou Université Paris-Saclay, France

Mariam Kiran Lawrence Berkeley National Laboratory, USA

Penporn Koanantakool Google LLC, USA

Ronald Kriemann MPI for Math. i.t.S., Germany

Ang Li Pacific Northwest National Laboratory, USA

Jiajia Li William and Mary College, USA
Piyush Sao Oak Ridge National Laboratory, USA
Christian Terboven RWTH Aachen University, Germany
Bo Wu Colorado School of Mines, USA
Rio Yokota Tokyo Institute of Technology, Japan

Rohit Zambre AMD Research, USA

Tutorials Committee

Kathryn Mohror (Chair)

Suren Byna (Deputy Chair)

Ritu Arora

Rosa M. Badia

Wahid Bhimji

Lawrence Livermore National Laboratory, USA

Lawrence Berkeley National Laboratory, USA

University of Texas at San Antonio, USA

Barcelona Supercomputing Center, Spain

Lawrence Berkeley National Laboratory, USA

Philip Carns Argonne National Laboratory, USA

James Dinan NVIDIA, USA

Ann Gentile Sandia National Laboratories, USA
Tanzima Islam Texas State University, USA
Simon McIntosh-Smith University of Bristol, UK
Diana Moise Cray, and HPE, Switzerland

Sarah Neuwirth Goethe-University Frankfurt, Germany

C. J. Newburn NVIDIA, USA

Dhabaleswar Panda Ohio State University, USA

Raghunath Raja Chandrasekar Frau, USA

Michela Taufer The University of Tennessee, USA
Michele Weiland EPCC – University of Edinburgh, UK

Workshops Committee

Hartwig Anzt (Chair) Karlsruhe Institute of Technology, Germany and

University of Tennessee, USA

Amanda Bienz (Deputy Chair) University of New Mexico, USA

Cody Balos Lawrence Livermore National Laboratory, USA

Harun Bayraktar NVIDIA, USA

Natalie Beams University of Tennessee, USA
Luc Berger-Vergiat Sandia National Laboratories, USA
George Bosilca University of Tennessee, USA

Lisa Claus LBNL, USA

Terry Cojean Karlsruhe Institute of Technology, Germany
Anthony Danalis University of Tennessee Knoxville, USA
Edoardo Di Napoli Juelich Supercomputing Centre, Germany
Markus Goetz Karlsruhe Institute of Technology, Germany
Aditya Kashi Karlsruhe Institute of Technology, Germany

Sarah Knepper Intel, USA

Andreas Knuepfer Technische Universität Dresden, Germany Martin Kronbichler Technical University of Munich, Germany Weifeng Liu China University of Petroleum, China

Simone Pezzuto
Università della Svizzera italiana, Switzerland
Enrique S. Quintana-Orti
Universitat Politècnica de València, Spain
Estela Suarez
Jülich Supercomputing Centre, Germany

Nico Trost AMD, Germany

Markus Wittmann Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

HPC in Asia Committee

Kento Sato (Chair) RIKEN, Japan

James Lin (Deputy Chair) Shanghai Jiao Tong University, China

Inclusivity Committee

Laura Schulz (Chair) Frau, Germany

Publicity Committee

Carsten Trinitis (Chair) Technical University of Munich, Germany

Proceedings Chairs

Piotr Luszczek (Chair) University of Tennessee, USA Marc Baboulin (Deputy Chair) Université Paris-Saclay, France

Contents

| Architecture, Networks, and Storage | |
|---|-----|
| Accelerating MPI All-to-All Communication with Online Compression on Modern GPU Clusters Qinghua Zhou, Pouya Kousha, Quentin Anthony, Kawthar Shafie Khorassani, Aamir Shafi, Hari Subramoni, and Dhabaleswar K. Panda | 3 |
| NVIDIA's Quantum InfiniBand Network Congestion Control Technology and Its Impact on Application Performance | 26 |
| LLM: Realizing Low-Latency Memory by Exploiting Embedded Silicon Photonics for Irregular Workloads Marjan Fariborz, Mahyar Samani, Pouya Fotouhi, Roberto Proietti, Il-Min Yi, Venkatesh Akella, Jason Lowe-Power, Samuel Palermo, and S. J. Ben Yoo | 44 |
| SU3_Bench on a Programmable Integrated Unified Memory Architecture (PIUMA) and How that Differs from Standard NUMA CPUs Jesmin Jahan Tithi, Fabio Checconi, Douglas Doerfler, and Fabrizio Petrini | 65 |
| Machine Learning, AI, and Emerging Technologies | |
| "Hey CAI" - Conversational AI Enabled User Interface for HPC Tools Pouya Kousha, Arpan Jain, Ayyappa Kolli, Saisree Miriyala, Prasanna Sainath, Hari Subramoni, Aamir Shafi, and Dhableswar K. Panda | 87 |
| Hy-Fi: Hybrid Five-Dimensional Parallel DNN Training on High-Performance GPU Clusters Arpan Jain, Aamir Shafi, Quentin Anthony, Pouya Kousha, Hari Subramoni, and Dhableswar K. Panda | 109 |

HPC Algorithms and Applications

| Efficient Application of Hanging-Node Constraints for Matrix-Free High-Order FEM Computations on CPU and GPU Peter Munch, Karl Ljungkvist, and Martin Kronbichler | 133 |
|--|-----|
| Dynamic Task Fusion for a Block-Structured Finite Volume Solver over a Dynamically Adaptive Mesh with Local Time Stepping | 153 |
| Accelerating Simulated Quantum Annealing with GPU and Tensor Cores Yi-Hua Chung, Cheng-Jhih Shih, and Shih-Hao Hung | 174 |
| m-Cubes: An Efficient and Portable Implementation of Multi-dimensional Integration for GPUs Ioannis Sakiotis, Kamesh Arumugam, Marc Paterno, Desh Ranjan, Balša Terzić, and Mohammad Zubair | 192 |
| Performance Modeling, Evaluation, and Analysis | |
| Comparative Evaluation of Call Graph Generation by Profiling Tools Onur Cankur and Abhinav Bhatele | 213 |
| MAPredict: Static Analysis Driven Memory Access Prediction Framework for Modern CPUs Mohammad Alaul Haque Monil, Seyong Lee, Jeffrey S. Vetter, and Allen D. Malony | 233 |
| Rapid Execution Time Estimation for Heterogeneous Memory Systems Through Differential Tracing | 256 |
| Understanding Distributed Deep Learning Performance by Correlating HPC and Machine Learning Measurements Ana Luisa Veroneze Solórzano and Lucas Mello Schnorr | 275 |
| A Motivating Case Study on Code Variant Selection by Reinforcement Learning Oliver Hacker, Matthias Korch, and Johannes Seiferth | 293 |
| Programming Environments and System Software | |
| Remote OpenMP Offloading | 315 |

| Hybrid Parallel ILU Preconditioner in Linear Solver Library GaspiLS | 334 |
|--|-----|
| A Subset of the CERN Virtual Machine File System: Fast Delivering of Complex Software Stacks for Supercomputing Resources Alexandre F. Boyer, Christophe Haen, Federico Stagni, and David R. C. Hill | 354 |
| Correction to: "Hey CAI" - Conversational AI Enabled User Interface for HPC Tools Pouya Kousha, Arpan Jain, Ayyappa Kolli, Saisree Miriyala, Prasanna Sainath, Hari Subramoni, Aamir Shafi, and Dhableswar K. Panda | C1 |
| Andhan Indan | 272 |

Contents xv