## **Lecture Notes in Computer Science**

13194

#### **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

Gerhard Woeginger

RWTH Aachen, Aachen, Germany

Moti Yung

Columbia University, New York, NY, USA

More information about this subseries at https://link.springer.com/bookseries/7408

Sridutt Bhalachandra · Christopher Daley · Verónica Melesse Vergara (Eds.)

# Accelerator Programming Using Directives

8th International Workshop, WACCPD 2021 Virtual Event, November 14, 2021 Proceedings



Editors
Sridutt Bhalachandra 
Lawrence Berkeley National Laboratory
Berkeley, CA, USA

Verónica Melesse Vergara D
Oak Ridge National Laboratory
Oak Ridge, DE, USA

Christopher Daley D Lawrence Berkeley National Laboratory Berkeley, CA, USA

ISSN 0302-9743 ISSN 1611-3349 (electronic) Lecture Notes in Computer Science ISBN 978-3-030-97758-0 ISBN 978-3-030-97759-7 (eBook) https://doi.org/10.1007/978-3-030-97759-7

LNCS Sublibrary: SL2 - Programming and Software Engineering

#### © Springer Nature Switzerland AG 2022

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**

In today's high-performance computing (HPC) environment, systems with heterogeneous node architectures providing multiple levels of parallelism are omnipresent. Architectures are evolving rapidly as we speak. Nodes in a future exascale system may also consist of GPU-like accelerators combined with other accelerators to provide improved performance for a wider variety of application kernels. The accelerators have become more usable in recent years, often providing high bandwidth memory with sufficient capacity to fit more of a scientific application's working set, hardware-managed caches, and the ability to implicitly access CPU data without the need for explicit data management. As a result, scientific software developers are being offered a rich platform to exploit the multiple levels of parallelism in their applications.

With increasing complexity to exploit the maximum available parallelism, the importance of programming approaches that can provide performance, scalability, and portability is increasing. Historically, the favored portable approaches, and sole focus of our earlier workshops, were OpenMP offloading and OpenACC. Today, we recognize there are other options and have extended the workshop to include use of standard Fortran/C++, SYCL, DPC++, Kokkos, and Raja among several others that can provide scalable as well as portable parallel solutions without compromising on performance. It is highly desirable that programmers are able to keep a single code base to help ease maintenance and avoid the need to debug multiple versions of the same code.

These proceedings contain the papers accepted for presentation at the 8th Workshop on Accelerator Programming using Directives (WACCPD 2021)—https://www.waccpd.org. WACCPD is one of the major forums for bringing together users, developers, and the software and tools community to share knowledge and experiences when programming emerging complex parallel computing systems.

As in previous years, the workshop highlighted improvements to the state of the art through the accepted papers and prompted discussion through keynotes that drew the community's attention to key areas that will facilitate the transition to accelerator-based HPC. The workshop aimed to showcase all aspects of heterogeneous systems, discussing innovative high-level language features, lessons learned while using directives to migrate scientific legacy code to parallel processors, compilation, and runtime scheduling techniques, among others.

The WACCPD 2021 workshop received 11 submissions out of which seven were accepted to be presented at the workshop and published in these proceedings. The Program Committee of the workshop comprised 24 members spanning universities, national laboratories, and industries. Each paper received a minimum of three single-blind reviews. A new role of "Proceedings/Reproducibility Chair" was added to further help with the reproducibility initiative. This role was ably filled by Ronnie Chatterjee from Lawrence Berkeley National Laboratory. Similar to WACCPD 2020, we encouraged all authors to add the Artifact Description (AD) to their submissions and make their code and data publicly available (e.g. on GitHub, Zenodo, Code Ocean) in support of the reproducibility initiative. As a further push, only papers with AD

were considered for the Best Paper Award. Of the seven accepted papers, two had reproducibility information and these manuscripts are highlighted with an 'artifacts available' logo in this book.

The program co-chairs invited Barbara Chapman from Hewlett Packard Enterprise (HPE) to give a keynote address on "New Frontiers for Directives". Barbara Chapman was a Professor of Computer Science for over 20 years, performing research on parallel programming interfaces and their implementation. Currently at HPE, she is defining future directions for the HPE Cray Programming Environment but remains affiliated with the Department of Computer Science and the Institute for Advanced Computational Science at Stony Brook University, where her team is engaged in efforts to develop community standards for parallel programming, including OpenMP, OpenACC, and OpenSHMEM.

Mathew Colgrove from NVIDIA Corporation and Sunita Chandrasekarn from the University of Delaware gave an invited talk titled "Introducing SPEChpc 2021". Mathew Colgrove is an NVIDIA DevTech engineer working with the NVHPC compiler team. Mat is also NVIDIA's representative on SPEC's CPU and HPG benchmarking committees. As well as serving on SPEC's Board of Directors, Mat holds several officer positions including Release Manager for SPEC HPG and SPEC's Vice-President of Operations. Sunita Chandrasekaran is an Associate Professor with the Department of Computer and Information Sciences at the University of Delaware, USA. She is also a computational scientist with Brookhaven National Laboratory. She received her Ph.D. in 2012 on Tools and Algorithms for High-Level Algorithm Mapping to FPGAs from the School of Computer Science and Engineering, Nanyang Technological University, Singapore.

The workshop concluded with a panel on "Publicly-available directive test suites for heterogeneous architectures" moderated by Christopher Daley from Lawrence Berkeley National Laboratory. The panelists included

- Swaroop Pophale, Oak Ridge National Laboratory, USA
- Michael Kruse, Argonne National Laboratory, USA
- Brandon Cook, Lawrence Berkeley National Laboratory, USA
- Mathew Colgrove, NVIDIA Corporation, USA
- Rahulkumar Gayatri, Lawrence Berkeley National Laboratory, USA

Based on rigorous reviews and ranking scores of all papers reviewed, the following paper won the Best Paper Award. The authors of the Best Paper Award also included reproducibility results to their paper, which the WACCPD workshop organizers had indicated as a criteria to be eligible to compete for the Best Paper Award.

 Miko Stulajter, Ronald Caplan, and Jon Linker from Predictive Science Inc.: "Can Fortran's 'do concurrent' Replace Directives for Accelerated Computing?".

An honorable mention for Best Artifact Description/Artifact Evaluation was presented to

 Kohei Fujita, Yuma Kikuchi, Tsuyoshi Ichimura, Muneo Hori, Lalith Maddegedara, and Naonori Ueda from the University of Tokyo, Riken, and the Japan Agency for Marine-Earth Science and Technology: "GPU porting of scalable implicit solver with Green's function-based neural networks by OpenACC".

These winners received prizes sponsored by NVIDIA Corporation.

Like last year, 2021 was (unfortunately) again a challenging year, as the world is trying to recover from the devastating effects of the COVID-19 pandemic. Major events around the world were canceled or scaled-down. Most computer conferences switched to virtual or hybrid formats. To this end, Supercomputing 2021 was held in a hybrid format for the first time. Similar to last year, WACCPD 2021 was again a fully virtual workshop. Thanks to all of you that contributed to its success! Hopefully, we will be able to meet in person again next time. Stay tuned!

December 2021

Sridutt Bhalachandra Christopher Daley Verónica Melesse Vergara

#### **Organization**

#### **Steering Committee**

Rosa Badia Barcelona Supercomputing Center, Spain

Barbara Chapman Stony Brook University, USA
Oscar Hernandez NVIDIA Corporation, USA
Kuan-Ching Li Providence University, Taiwan

Nick Malaya AMD, USA Duncan Poole OpenACC, USA

Jeffrey Vetter Oak Ridge National Laboratory, USA

Michael Wong Codeplay Software, UK

#### **General Chairs**

Sunita Chandrasekaran University of Delaware, USA

Guido Juckeland Helmholtz-Zentrum Dresden-Rossendorf,

Germany

#### **Program Chairs**

Sridutt Bhalachandra Lawrence Berkeley National Laboratory, USA Christopher Daley Lawrence Berkeley National Laboratory, USA

Verónica Melesse Vergara Oak Ridge National Laboratory, USA

#### Proceedings/Reproducibility Chair

Ronnie Chatterjee Lawrence Berkeley National Laboratory, USA

**Publicity Chair** 

Rahulkumar Gayatri Lawrence Berkeley National Laboratory, USA

Web Chair

Weile Wei Louisiana State University, USA

#### **Program Committee**

Daniel Abdi National Oceanic and Atmospheric

Administration, USA

James Beyer NVIDIA Corporation, USA

Maciej Cytowski Pawsey Supercomputing Center, Australia Johannes Doerfert Argonne National Laboratory, USA

Rahulkumar Gayatri Lawrence Berkeley National Laboratory, USA

Antigoni Georgiadou Oak Ridge National Laboratory, USA Millad Ghane Samsung Semiconductor Inc., USA

Priyanka Ghosh Pacific Northwest National Laboratory, USA

Arpith Jacob Google, USA

Haoqiang Jin NASA Ames Research Center, USA

Mozhgan Kabiri Chimeh NVIDIA Corporation, UK

Vivek Kumar Indraprastha Institute of Information Technology

Delhi, India

John Leidel Tactical Computing Laboratories LLC and Texas

Tech University, USA

Ron Lieberman AMD, USA

Chun-Yu Lin National Center for High-Performance

Computing, Taiwan

Meifeng Lin Brookhaven National Laboratory, USA

Simon McIntosh-Smith University of Bristol, UK

Stephen Lecler Olivier Sandia National Laboratory, USA
Tom Papatheodore Oak Ridge National Laboratory, USA

Thomas Schwinge Mentor Graphics, Germany

Gregory Stoner Intel, USA

Volker Weinberg Leibniz Supercomputing Centre, Germany

Rengan Xu Dell, USA

Charlene Yang NVIDIA Corporation, USA











### **Contents**

| Directive Alternatives  |     |
|---|-----|
| Can Fortran's 'do concurrent' Replace Directives for Accelerated Computing?   | 3   |
| Miko M. Stulajter, Ronald M. Caplan, and Jon A. Linker  |     |
| Achieving Near-Native Runtime Performance and Cross-Platform Performance Portability for Random Number Generation Through SYCL  |     |
| Interoperability Vincent R. Pascuzzi and Mehdi Goli   | 22  |
| Directive Extensions  |     |
| Extending OpenMP for Machine Learning-Driven Adaptation  Chunhua Liao, Anjia Wang, Giorgis Georgakoudis,  Bronis R. de Supinski, Yonghong Yan, David Beckingsale, and Todd Gamblin              | 49  |
| Directive Case Studies  |     |
| GPU Porting of Scalable Implicit Solver with Green's Function-Based Neural Networks by OpenACC  Kohei Fujita, Yuma Kikuchi, Tsuyoshi Ichimura, Muneo Hori, Lalith Maddegedara, and Naonori Ueda | 73  |
| Challenges Porting a C++ Template-Metaprogramming Abstraction Layer to Directive-Based Offloading   | 92  |
| Accelerating Quantum Many-Body Configuration Interaction with Directives  | 112 |
| Brandon Cook, Patrick J. Fasano, Pieter Maris, Chao Yang,   |     |

and Dossay Oryspayev

#### xiv Contents

| GPU Offloading of a Large-Scale Gyrokinetic Particle-in-Cell Fortran |     |
|--|-----|
| Code on Summit: From OpenACC to OpenMP                               | 133 |
| Qiheng Cai, Junyi Cheng, Yang Chen, Marcus Wagner,                   |     |
| Christopher Daley, Dossay Oryspayev, Stefan Tirkas, Sophie Redd,     |     |
| and Scott Parker   |     |
| Author Index   | 149 |