GUEST EDITORS' INTRODUCTION

Performance Portability for Advanced Architectures

Douglas Doerfler, *Lawrence Berkeley National Laboratory, Berkeley, CA, 94720, USA*

Steven Gottlieb, *Indiana University (AEiC), Bloomington, IN, 47405, USA*

William Gropp, *University of Illinois at Urbana-Champaign, Urbana, IL, 61801, USA*

Barry I. Schneider, *National Institute of Standards and Technology (AEiC), Gaithersburg, MD, 20899, USA*

Alan Sussman, *National Science Foundation Alexandria, VA, 22314, USA*

Performance Portability for Advanced Architectures is a very timely topic as we continue our journey to exascale computers. The leader on the Top 500 list today is the Fujitsu Fugaku computer in Japan, using a modified ARM architecture and 512-bit scalable vector extension (SVE) instructions. In the USA, the National Energy Research Scientific Computing Center (NERSC) has just dedicated Perlmutter, named for the Berkeley Nobel Laureate Saul Perlmutter. Perlmutter, built by Cray/HPE, contains both CPU-only and CPU/GPU nodes. The CPUs are produced by AMD and the GPUs come from NVIDIA. Frontier at Oak Ridge National Laboratory and El Capitan at Lawrence Livermore National Laboratory will contain new AMD GPUs, and Aurora at Argonne National Laboratory will feature new GPUs from Intel. Los Alamos National Laboratory's next-generation machine Crossroads will use Intel CPUs with high-bandwidth memory and AVX-512 vector instructions. Optimizing code for three different GPUs is a daunting challenge for application programmers. Further, there are still many computers that do not feature GPUs, and most computational scientists would not be eager to totally abandon such machines.

Several of your guest editors were involved with the organization of two conferences on the topic of this issue. Douglas Doerfler was on the Steering Committee for the 2020 Performance, Portability, and Productivity in HPC Forum, held in April 2020. Barry Schneider and Alan Sussman were on the Planning Committee of the NITRD-sponsored workshop on Software in the Era of Extreme Heterogeneity, held in September, 2020. This Special Issue of CiSE contains four invited articles that represent some of the important ideas presented in these meetings. Materials from the April meeting[a] and information on the September meeting are available online. [b]

Our first article, "Navigating Performance, Portability and Productivity," by Simon Pennycook, Jason Sewall, Douglas Jacobsen, Tom Deakin, and Simon McIntosh-Smith, provides a rigorous definition of performance portability. The article also examines the important topic of programmer productivity and the need to specialize parts of the code to improve performance on different platforms. You will learn about their tool Code Base Investigator that measures how much code changes as one compiles it for different platforms.

The second article by Michael Wolfe is titled "Performant, Portable, and Productive Parallel Programming with Standard Languages." It provides a valuable historical perspective on how parallel computing has evolved over multiple decades. There were several periods when computers were more amenable to achieving the desired 3Ps. Wolfe introduces a machine model that characterizes the most salient features of today's computers and argues that programming language standards should allow the programmer to express all the necessary parallelism for the compiler to efficiently parallelize the code using all available hardware.

In "The Kokkos EcoSystem: Comprehensive Performance Portability for High Performance Computing," Christian Trott *et al.* focus on a relatively new and popular framework for developing parallel codes called Kokkos. Their article is not a tutorial on writing code using Kokkos, rather, it describes the available tools for development and how they help the programmer to productively target multiple platforms. The article

1521-9615 © 2021 IEEE
Digital Object Identifier 10.1109/MCSE.2021.3104083
Date of current version 23 September 2021.

[a]https://p3hpcforum2020.alcf.anl.gov
[b]https://www.nitrd.gov/software-in-the-era-of-extreme-heterogeneity/

also discusses how the team is working to assure the long term future of Kokkos.

The fourth article, "Performance Portability in the Exascale Computing Project: Exploration Through a Panel Series," is by Anshu Dubey *et al*. This article summarizes findings from a series of panel discussions organized by the Exascale Computing Project and outlines some of the experiences of both software technology and application teams in the project. The software technology teams are developing the Extreme-scale Scientific Software Stack (E4S), which includes tools, frameworks, and libraries that many of the application groups will use.

A fifth article in the issue addresses a topic indirectly related to performance portability. Achieving codes with performance portability requires insights and skills that are not currently widely available. This can be a real challenge for small research groups that do not have a large development team. In "Advancing the Workforce that Supports Computationally and Data Intensive Research," Patrick Schmitz, Scott Yockel, Claire Mizumoto, Thomas Cheatham, and Dana Brunson describe a relatively new computing professional role that they call a Research Computing and Data (RCD) Professional. These professionals have strong skills in such areas as programming, big data management, artificial intelligence, and machine learning, to name a few. They work closely with client research groups, often for months, to provide deep support to solve each group's computing challenges.

We believe you will enjoy reading about the proposed skill set and what type of support is necessary to establish RCD professionals as a standard component of research teams at universities, national laboratories, and other research institutions.

DOUGLAS DOERFLER has recently retired. Prior to retirement, he was a Senior Computer Systems Engineer at Lawrence Berkeley National Laboratory's National Energy Research Scientific Computing Center (NERSC), where he led efforts in benchmarking and architectural evaluation of high performance computing systems and technologies. Prior to NERSC, he was a distinguished member of Technical Staff at Sandia National Laboratories, where he had a decades long career in the field of high performance computing culminating in serving as the Chief Architect for the ACES collaboration with Los Alamos National Laboratory, and the deployment of the Cielo and Trinity Supercomputers for the National Nuclear Security Administration. Contact him at dwdoerf@gmail.com.

STEVEN GOTTLIEB is a Distinguished Professor Emeritus and Provost Professor Emeritus of physics at Indiana University, Bloomington, IN, USA. His research is in lattice QCD. He is a founding member of the MILC Collaboration. The MILC code has been used to benchmark multiple computer acquisitions at NERSC and for acceptance testing of Blue Waters. He is actively preparing the MILC code for Perlmutter through the NESAP Program, and for Frontier and Aurora through the Exascale Computing Project. He is a Fellow of the American Physical Society and has served as Councillor for the Division of Computational Physics. He is an Associate Editor-in-Chief of *Computing in Science and Engineering* and recently joined the Executive Board of the Oak Ridge Leadership Computing Facility User Group. Contact him at sg@iu.edu.

WILLIAM GROPP is the Director of the National Center for Supercomputing Applications and holds the Thomas M. Siebel Chair in the Department of Computer Science, University of Illinois in Urbana-Champaign, Champaign, IL, USA. His research interests include parallel computing, software for scientific computing, and numerical methods for partial differential equations. He received the Ph.D. degree in computer science from Stanford University in 1982, and worked at Yale University and Argonne National Laboratory. He is a Fellow of AAAS, ACM, IEEE, and SIAM, and a member of the National Academy of Engineering, and has received numerous awards for his work in HPC. Contact him at wgropp@illinois.edu.

BARRY I. SCHNEIDER is a Staff Member of the NIST Applied and Computational Mathematics Division. In early 2014, he came to NIST as general editor of the DLMF project after a long career at Los Alamos National Laboratory and the National Science Foundation. His current research interests span a broad number of areas of theoretical chemistry, atomic and molecular physics, numerical methods and high performance computing. His current principal focus is developing novel methods for the solution of the time dependent Schrodinger equation in ultra-short, and intense laser fields. He is the principal architect for the AMPGateway, which is delivering computational capabilities to the entire atomic and molecular physics community. He was a recipient of the prestigious Alexander von Humboldt Prize from the German government in 1986 and a Poste Rouge from France in 1980. He and a small team of other NIST staff were awarded the Department of Commerce Bronze Medal recognizing their work in transforming NIST's infrastructure for computing in the era of data-intensive science. Contact him at bis@nist.go.

ALAN SUSSMAN is a Program Director in the Office of Advanced Cyberinfrastructure, National Science Foundation, in charge of learning and workforce development programs, and is also active in software and data related cyberinfrastructure programs. He is on leave from his permanent position as a Professor of computer science at the University of Maryland. His research interests include systems software support for large-scale applications that require high performance parallel and distributed computing. Working with students and other collaborators at Maryland and other institutions, he has published numerous conference and journal papers, and received several best paper awards in various topics related to software tools for high performance parallel and distributed computing. In addition, since 2010 he has been helping coordinate a curriculum initiative in parallel and distributed computing with the premise that every undergraduate student in computer science or computer engineering must acquire basic parallel computing skills. This curriculum has had wide adoption including its direct impact on ACM's CS2013 Curriculum. Contact him at alasussm@nsf.gov.