Guest Editors' Introduction

Race to Exascale

We are pleased to be able to bring you this "Race to Exascale" special issue of *Computing in Science and Engineering* (CISE). Whether called leadership computing, flagship computing, or just plain exascale, over the next few years, governments around the world are planning to spend over 10 billion dollars on a handful of new computer systems that will strive to reach an exascale level of performance. These systems and projects reflect the widespread and expanding recognition that almost all science and engineering endeavors now are intrinsically reliant on computing power not just for modeling and simulation but for data analysis, big data, and machine learning. Scientists and engineers consider computers as "universal instruments" of insight.

The implications of these major investments are enormous for the CISE community, not just because of the cost, but also because of the many issues that will determine how effective exascale systems will be for the wide range of disciplines that currently have important but infeasible problems. The main international projects in exascale computing are in China, Europe, Japan, and the United States. For this special issue, we asked experts from these regions to give you an idea of the complexity and challenges these ambitious and important projects face. Challenges for the authors, and for us, are that some of the information about projects is proprietary and some plans are still not public and/or subject to change.

Our first article, "High Performance Computing Development in China: A Brief Review and Perspectives," describes the history of computing efforts in China and plans to be the first to reach exascale level computing. You may recall that in November 2010, China's Tianhe-1A was the number one computer on the Top 500 list and that from June 2013 to November 2015, China's Tianhe-2 topped the list. The next top computer was China's Sunway TaihuLight, so it should come as no surprise that China may be spending more than any of the others and may have motivated this "race." China is deploying at least three "pre-exascale" systems in 2018/2019 with three different architectures. Based on the results of these systems, China plans to deploy two or more exascale level systems that they hope will be available in 2020 and that would have the ability to execute the high-performance Linpack (HPL) benchmark above one exaflops per second.

"The European approach to the exascale challenge" provides many details about the efforts in Europe to support the future of HPC. On September 28, 2018, the European Union and 25 European countries established a Joint Undertaking, which is a legal framework to help coordinate their efforts to assure a bright future for HPC in Europe. Two exascale systems are planned by 2023, with at least one of them based on European technology. This ambitious project targets application development, co-design, and user training. Funding for the first two years is about 1 billion euros, with more to come as EuroHPC will run to at least 2026. This project will also consider how quantum computing can play a future role in computational science and engineering. Another interesting aspect of the project is the plan to increase the use of HPC in industry. HPC competence centers will provide training and outreach to (especially) small and medium size companies that may face greater challenges in adopting HPC than the largest companies.

The short contribution entitled "Japan's Flagship 2020 Post-K System" focuses on the description of a single system and the project that is the largest leadership/exascale computing project in Japan. It is planned to complete deployment in mid-2020. This effort takes a giant step beyond the petascale K computer installed in 2011. Unlike the U.S. and European efforts, the post-K system

Digital Object Identifier 10.1109/MCSE.2018.2882574
Date of current version 6 March 2019.

1521-9615 © 2019 IEEE
Published by the IEEE Computer Society

relies entirely on general purpose CPUs with vectorization to achieve its broad goals implemented by Fujitsu in a new 48-core ARM processor. A key to this effort is the attention to balance among floating-point speed, memory bandwidth, and interconnect bandwidth. The computer will have over 100 000 nodes. This Post-K project is complemented by a broader balanced investment strategy that will implement systems in at least 12 of the major Japanese research universities, which, in aggregate, will also total an exascale level of computing in the same time period.

The article "Exascale Computing in the United States," about the US exascale computing program, explains a complex project intended to deploy at least three exascale capable systems by 2023 at three different DOE laboratories. The ECP is part of an even larger effort within US federal government called the National Strategic Computing Initiative (NSCI) announced in the summer of 2015. ECP is currently the largest and most obvious part of the NSCI efforts. This article outlines the overall goals of the ECP project, which is focused on "codesign" methods to bring application development teams together with system designers and system software environment developers to not only make a system that is capable of running a simple benchmark such as HPL above an exaflop, but is more significantly able to solve important problems in a wide range of areas. Hence, ECP's definition of capable exascale is developing baseline performance measures on today's systems called figures of merit and then improving the ability to solve problems by at least 50 fold using figures of merit.

We wrap up our special issue with "Reflecting on the Goal and Baseline for Exascale Computing: A Roadmap Based on Weather and Climate" that is related to the European exascale efforts and takes a different tack by exploring the challenges of creating an exascale solution in the important area of weather and climate prediction. The point of this article is not to provide a comprehensive overview of the European exascale related efforts, but rather to provide an in-depth exploration of the complexities facing both application developers and system designers to create exascale solutions and explore what "codesign" actually takes. The article is important because it clearly explains how floating-point operations, the common marketing metric for the performance of computers in the past, are not relevant in understanding performance and time to solution, in particular, for climate and weather. The authors present a new memory use efficiency measure that is much more important to the true, sustained performance of applications in the exascale era.

We hope you find this special issue informative, and we thank the authors for their time and insights sharing these important efforts with you.

JACK DONGARRA
University of Tennessee
STEVEN GOTTLIEB
Indiana University
WILLIAM T. C. KRAMER
University of Illinois at Urbana-Champaign

Jack Dongarra holds an appointment as a University Distinguished Professor of computer science with the Electrical Engineering and Computer Science Department, University of Tennessee, and holds the title of a Distinguished Research Staff with the Computer Science and Mathematics Division, Oak Ridge National Laboratory; a Turing Fellow with Manchester University; and an Adjunct Professor with the Computer Science Department, Rice University. He specializes in numerical algorithms in linear algebra, parallel computing, the use of advanced-computer architectures, programming methodology, and tools for parallel computers.

Steven Gottlieb is a Distinguished Professor Emeritus and a Provost Professor Emeritus of physics with Indiana University. His research is in lattice quantum chromodynamics, and he has been using parallel computers for about 30 years, with a collection of coffee mugs to match. He is a Founding Member of the MILC collaboration and one of the key developers of the MILC code.

William T. C. Kramer leads the National Center for Supercomputing Applications' Blue Waters project, which deployed the highest sustained-performance computational and data analysis system available to the nation's open research community. He also is a Research Professor with the Computer Science Department, University of Illinois at Urbana-Champaign. Previously, he was the General Manager of the National Energy Research Scientific Computing Center, the flagship computing facility of the Department of Energy's Office of Science, and the Head of the High Performance Computing Department, Lawrence Berkeley National Laboratory.