Guest Editors' Introduction

Transforming Science through Software: Improving while delivering 100X

Richard Gerber, Lawrence Berkeley National Laboratory
Steven Gottlieb, Indiana University
Michael A. Heroux, Sandia National Laboratories and St. John's University
Lois Curfman McInnes, Argonne National Laboratory

The U.S. Department of Energy (DOE) Exascale Computing Project (ECP) [1,2,3] represented one of the largest scientific software projects ever. Charged with establishing a practical high-performance computing (HPC) ecosystem around systems capable of achieving a billion-billion double-precision operations per second, or an exaflop, the ECP funded the development of new (and the transformation of important existing) applications, libraries, and tools that realized much more improvement in performance and capabilities than what the raw hardware performance potential would indicate. Exascale systems are nominally 50 times more capable than the supercomputing systems available at the start of the ECP, but the improvements in observed capabilities were often 100 times or more.

This exceptional gain, achieved in seven years, inspired the title of this special issue: *Transforming Science through Software: Improving while delivering 100X*. Through the six invited papers in this special issue, our goal is to share experiences with the practice and science of scientific software development, with an emphasis on developing a coherent, portable, and sustainable software ecosystem for high-performance computing, with an eye to the needs of next-generation computational science.

The term 100X refers to advancing capabilities in modeling, simulation, and analysis by a factor of 100 or more using some combination of new algorithms, optimization techniques, software libraries, and programming models coupled with the next generation of HPC hardware. The work represented in this issue was performed by teams funded by the ECP who were charged with delivering novel algorithms for emerging exascale computing platforms in robust software libraries, tools, and applications as needed by stakeholders, while simultaneously improving software quality to reduce technical risk and build a firm foundation for scientific advances.

Heterogeneous architectures and exascale computing. For decades, HPC applications steadily increased performance by relying on a relatively constant and straightforward growth in the capability of computing hardware based on CPUs. But, about fifteen years ago, single-core CPU performance growth began to slow due to the end of Dennard scaling and the slowing of Moore's law. New heterogeneous computing architectures (typically incorporating hardware accelerators [4]) arose with the potential to unleash enormous gains in computational capability for science (even at the desktop level), but realizing that potential comes with substantially increased programming complexity. Developing and optimizing applications for heterogeneous systems is challenging due to massive parallel concurrency requirements, reduced bandwidth-to-compute ratios, heterogeneous memory spaces, new programming models and languages, changing storage paradigms, and a proliferation of hardware accelerators.

Beginning in 2016, the DOE Office of Science's Advanced Scientific Computing Research program and the National Nuclear Security Agency's Advanced Simulation and Computing program jointly formed ECP to support research, development, and deployment of science and national security applications, software, and hardware technologies specifically targeted to the DOE mission priorities. Six years after the start of the ECP, the nation saw the delivery of the world's first exascale systems. By the time of ECP's conclusion in December 2023, ECP teams had developed a diverse suite of applications in chemistry, materials, energy, Earth and space science, data analytics, optimization, AI, and national security [5]. In turn, these applications build on a robust software ecosystem [6], including programming models and runtimes, mathematical libraries, data and visualization packages, and development tools that comprise the ECP-developed Extreme-scale Scientific Software Stack (E4S). Achieving the required advances in simulation and scalability in ECP required a high level of collective efficiency, performance, and fidelity in all models, methods, software, tools, and ECP-enabled computing technologies. This, in turn, required a new level of coordination and collaboration among teams of researchers, computer scientists, and software technology specialists with hardware vendors and DOE HPC facilities that build and manage exascale computers.

Transforming science through software: Improving while delivering 100X. To inform future scientific software communities, the papers in this special issue of IEEE CiSE focus on ECP software technologies and their deployment at DOE HPC facilities through the lens of a sustainable scientific software ecosystem, as needed to support applications at exascale and beyond. We also discuss cultural changes that embrace continual improvement in software practices as needed for collaborative and trustworthy computational science. A complementary special issue of IEEE CiSE entitled *The Scientific Impact of the Exascale Computing Project* focuses on advances in ECP applications. Efforts in ECP on scientific libraries and tools represent one of the most significant scientific software projects ever, requiring a high degree of coordination across many teams and institutions to deliver various software capabilities in support of new exascale platforms and their users. In addition, the mission and size of ECP allowed these teams to invest in collaboration to understand and improve how work on scientific software is accomplished. Building on ECP successes, the community is now tackling new opportunities in the computing sciences. For example, ECP outcomes set the stage for future DOE efforts in Integrated Research Infrastructure (IRI) and AI for Science, which will require a rich, reliable, and portable software stack to support uniformity across systems, as well as new libraries and tools seamlessly integrated with those provided by ECP.

Synopsis of papers. The papers in this issue discuss complementary topics related to building this scientific software ecosystem, including cultural changes as we embrace the fundamental role of high-quality software in sustained collaboration and discovery in computational science. Topics include software portfolio leadership and management techniques, community development, increased impact on scientific progress via robust scientific libraries and tools that are portable across advanced computing architectures, and realizing scientific potential through delivery, integration, and performance tuning of application codes built on this software. The following is a brief synopsis of papers in this special issue:

- ***Scalable Delivery of Scalable Libraries and Tools: How ECP Delivered a Software Ecosystem for Exascale and Beyond*** – Explains software portfolio leadership and management techniques in the ECP Software Technology Focus Area, emphasizing applicability to other large-scale research software projects.
- ***Providing a Flexible and Comprehensive Software Stack via Spack, E4S, and SDKs*** – Introduces the multilayer software ecosystem developed for ECP and beyond, which provides a curated distribution of numerical libraries, runtime systems, and tools that lowers the barrier to entry to HPC and AI/ML developer communities, while promoting software interoperability, quality improvement, porting, testing, and deployment across a range of computing environments.
- ***Creating Continuous Integration Infrastructure for Software Development on DOE HPC Systems*** – Presents ECP's approach for creating robust continuous integration (CI) workflows that span multiple protected HPC environments, emphasizing strategies to address challenges in automation, charging models, and security requirements.
- ***Deploying Optimized Scientific and Engineering Applications on Exascale Systems*** – Highlights how ECP's Application Integration (AppInt) area served as an integration point among applications, supporting software, system environments, HPC facilities, and vendors—overcoming challenges in using first-of-a-kind computing systems while promoting portable and sustainable programming models and helping harden systems before general availability.
- ***A Cast of Thousands: How the IDEAS Productivity Project has Advanced Software Productivity and Sustainability*** – Explains how synergistic work—in fostering software communities while incubating methodologies and disseminating knowledge to advance developer productivity and software sustainability—helps to mitigate technical risks by building a firmer foundation for reproducible, sustainable science at all scales of computing, from laptops to clusters to exascale and beyond.
- ***Then and Now: Improving Software Portability, Productivity, and 100X Performance*** – Discusses how transformations in software development methodologies and interdisciplinary teams during ECP led to improvements in software technologies and applications, and why these advances are essential for next-generation computational science.

We recognize that these papers represent only a subset of the broad range of excellent work in scientific software throughout the international community. We hope the papers will benefit research teams and provide a foundation for advancing computational science and our knowledge and understanding of the world. We further hope that these papers will foster expanded community efforts related to the fundamental role of sustainable scientific software ecosystems in advancing the computing sciences.

Acknowledgments. This research was supported by the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the U.S. Department of Energy Office of Science and the National Nuclear Security Administration. We thank all those in the ECP, DOE, and broader computational science community who are advancing software for science. We especially thank people who have contributed to the work discussed in this special issue.

We sincerely thank all reviewers of papers in this special issue for their insights and feedback, which substantially improved the manuscripts: Wolfgang Bangerth, David Bernholdt, Manoj Bhardwaj, John Cazes, Edmond Chow, Ben Cowan, Charles Ferenbaugh, Jennifer Green, Fred Johnson, Alicia Klinvex, Scott L'Ecuyer, Paul Lin, Ray Loy, Mark Miller, David Moulton, Ryan Richard, David Rogers, Joost VandeVondele, and John Quincy Wofford III. This special issue used an open review process, where the co-authors of a paper were provided the names of their paper's three reviewers but not attribution of comments to a particular reviewer.

References

[1] U.S. DOE Exascale Computing Project, https://exascaleproject.org.
[2] P. Messina, The Exascale Computing Project, IEEE CiSE, vol. 19, no. 3, pp. 63-67, 2017, http://doi.org/10.1109/mcse.2017.57.
[3] D. Kothe, S. Lee and I. Qualters, Exascale Computing in the United States, *IEEE CiSE*, vol. 21, no. 1, pp. 17–29, 2019, https://doi.org/10.1109/MCSE.2018.2875366.
[4] M. Rosso and A. Myers, A Gentle Introduction to GPU Programming, 2021, BSSw.io blog article, https://bssw.io/blog_posts/a-gentle-introduction-to-gpu-programming.
[5] F. Alexander et al., Exascale Applications: Skin in the game, *Philosophical Transactions of the Royal Society A*, 2020, https://doi.org/10.1098/rsta.2019.0056.
[6] L. C. McInnes, M. A. Heroux, E. W. Draeger, A. Siegel, S. Coghlan, and K. Antypas, How Community Software Ecosystems can Unlock the Potential of Exascale Computing, *Nature Computational Science*, vol. 1, pp. 92–94, 2021, https://doi.org/10.1038/s43588-021-00033-y.

Author Bios:

Richard Gerber is the Senior Science Advisor at the National Energy Research Scientific Computing Center (NERSC) at Lawrence Berkeley National Laboratory, Berkeley, CA, USA. He also serves as Director of Hardware and Integration for the Department of Energy's Exascale Computing Project. His research focus is on developing and implementing High Performance Computing ecosystems that accelerate scientific discovery. Gerber received his Ph.D. in Physics from the University of Illinois at Urbana-Champaign. Contact him at ragerber@lbl.gov.

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 helped prepare 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 chair of the Executive Board of the Oak Ridge Leadership Computing Facility User Group. Contact him at sg@iu.edu.

Michael A. Heroux is a senior scientist at Sandia National Laboratories, Director of Software Technology for the US Exascale Computing Project, and a Scientist in Residence at St. John's University, MN, USA. His research interests include all human and technical aspects of scalable scientific and engineering software for new and emerging parallel computing architectures. He founded the Trilinos scientific libraries, Kokkos performance portability, Mantevo miniapps, and HPCG Benchmark projects. He is presently leading the Extreme-scale Scientific Software Stack (E4S) project in DOE, a curated collection of HPC software targeting leadership platforms. Mike is a Fellow of the Society for Industrial and Applied Mathematics (SIAM), a Distinguished Member of the Association for Computing Machinery (ACM), and a Senior Member of IEEE. Contact him at maherou@sandia.gov.

Lois Curfman McInnes is a Senior Computational Scientist and Argonne Distinguished Fellow in the Mathematics and Computer Science Division of Argonne National Laboratory, Argonne, IL, USA. Her work focuses on high-performance computational science, with emphasis on scalable numerical libraries (particularly PETSc/TAO) and community collaboration toward productive and sustainable software ecosystems. McInnes serves as Deputy Director of Software Technology for the DOE Exascale Computing Project (ECP) and a leader of the ECP Broadening Participation Initiative. McInnes earned a Ph.D. in applied mathematics from the University of Virginia. She is a SIAM Fellow and member of ACM and IEEE. Contact her at mcinnes@anl.gov.