

# Introduction to Volume 11 Issue 1: Special Issue on HPC Training and Education

Nitin Sukhija

Slippery Rock University of Pennsylvania  
Slippery Rock, PA

## FORWARD

High performance computing is becoming central for empowering scientific progress in the most fundamental research in various science and engineering, as well as societal domains. It is remarkable to observe that the recent rapid advancement in the today's and future computing and software environments provide both challenges and opportunities for cyberinfrastructure facilitators, trainers, and educators to develop, deliver, support, and prepare a diverse community of students and professionals for careers that utilize high performance computing to advance discovery. This special issue focuses on original research papers submitted to the First Workshop on HPC Education and Training for Emerging Technologies (HETET19), which was held in conjunction with ISC19 conference in Frankfurt, Germany, June 20, 2019 the Second Workshop on Strategies for Enhancing HPC Education and Training (SEHET19), which was held in conjunction with PEARC19 conference in Chicago, Illinois, U.S.A., July 29, 2019, and the Sixth SC Workshop on Best Practices for HPC Training and Education (BPHE19), which was held in conjunction with SC19 conference in Denver, Colorado, U.S.A., November 17, 2019.

This special issue begins with an article by Holcomb et al that presents the lessons learned from the summer school and the internship program operated by University of Virginia in partnership with NASA from the year 2013 to 2018. The greatest challenge they found was to provide a good program for students with widely varying backgrounds, skills, and expectations. They accommodated the diversity by increasing hands-on exercises interspersed with lectures and by expanding their bank of programming exercises to span a range of abilities.

The article by Goodhue describes the Northeast Cyberteam Program, a 3-year National Science Foundation (NSF) funded initiative to increase effective use of cyberinfrastructure by researchers and educators at small and mid-sized institutions in Northern New England (Maine, Massachusetts, New Hampshire, Vermont). The core of their strategy is to build a regional pool of research computing facilitators (RCFs) and a process to share them across institutional boundaries, augmented by knowledge sharing and self-service learning tools that increase the effectiveness of Research Computing Facilitators. They conclude by highlighting the program management portal and the impact of the program on the smaller institutions and its relevance to the potential facilitators.

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Copyright ©JOCSE, a supported publication of the Shodor Education Foundation Inc.

The article by Chakravorty et al presents the Summer Computing Academy (SCA), a weeklong cybertraining program offered to high school students by High Performance Research Computing (HPRC) at Texas A&M University (Texas A&M; TAMU). They provide the best practices that have been adopted in the Summer Computing Academy model and the recruitment strategies along with the selection criteria for the participants. They conclude by discussing the legal and administrative issues encountered while hosting such efforts and the sustainability of the effort.

The article by Cais et al presents a training model of the carpenteries for the HPC space that provides a pathway to collaboratively created training content which can be delivered in a scalable way (serving everything from university or industrial HPC systems to national facilities). They describe some of the training material design principles and how they influence the creation process. They conclude by discussing two distinct evaluation processes, a review process that happens during material creation and learner evaluations that occur during/after training events.

The article by Lathrop et al describes the National science Foundation funded Blue Waters project, which supports an Education, Outreach and Training (EOT) program focused on preparing an HPC-capable workforce with an emphasis on petascale computing competencies. They discuss how the Blue Waters EOT team engages undergraduate students in internships, graduate students in fellowships, researchers as participants in training sessions, trainers and educators as PIs of education allocations, and underrepresented communities as PIs of broadening participation allocations. They conclude by describing the impact and benefits of the project, including directly reaching people located in many foreign countries, as well as freely disseminating materials that have been downloaded and used by thousands of people world-wide.

The article by Kunkel et al presents the current status of the certification program curated by the HPC Certification Forum. They describe the program that consists of three parts: the tree of defined competencies, the examination of practitioners to prove they possess those skills, and finally the certification demonstrating their knowledge. They conclude by discussing the benefits of the program that not only allows the re-use of existing content but also makes it possible to create a new ecosystem in which HPC centers, research labs, academic institutions and commercial companies could offer the best of their teaching material.

The article by Wong et al describes design and plan of a National Science Foundation supported hands-on Research Experiences for Computational Science, Engineering, and Mathematics (RECSEM) program at the University of Tennessee (UTK) in high-performance data sciences, data analytics, and machine learning on emerging

computer architectures. They discuss the experiences and resolutions in managing and coordinating the program, delivering cohesive tutorial materials, directing mentorship of individual projects, lessons learned, and improvement over the course of the program, particularly from the perspectives of the mentors. They conclude by discussing the outcomes and the progress of the students and overall impact of the REU program.

The article by Saravanan et al presents a novel course curriculum to teach high performance, parallel and distributed computing to senior graduate students (PhD) in a hands-on setup through examples drawn from a wealth of areas in computational biology. They provide a sample course outline that details the HPC concepts that can be covered in a one-semester course, along with the suggested bioengineering applications to introduce them. They conclude by discussing the assessment methods and the high evaluation ratings received for this interdisciplinary course.

The article by Wang et al describes an ongoing effort, FreeCompilerCamp.org, a free and open online learning platform aimed to train researchers to quickly develop OpenMP compilers. They explain the challenges faced in giving compilers training and the solutions and present the implementation of the platform. They provide an overview of the design of the web-based tutorials to take advantage of FreeCompilerCamp. They conclude by discussing the feedback received on the design of FreeCC tutorial.

The article by Terboven et al presents a learning status survey, a developer diary to track the student's progress in achieving the learning objectives, and an approach to enable the comparison of different HPC cluster architectures or parallel programming models. They report on the learning objectives of the labs along with their experiences with using various stimuli in their labs to increase the success rate of the learning objectives while fostering creative solutions. For example they use a competition among students to motivate them to optimize their codes for performance and show the opinions that students have towards these concepts. They conclude by discussing the feedback on the concept of self-paced learning and the evolution of the software labs in terms of obtained knowledge, training productivity and programming models, as well as students' feedback based on teaching evaluations.

The article by Colbry et al shares the lessons learned during the Computational Mathematics, Science and Engineering (CMSE) department's development at the Michigan State University (MSU) and the initiatives it has taken on to support computational research and education across the university. They describe how the department has aided in establishing both traditional degree programs and non-traditional options to build computational competency in learners from across STEM. They conclude by discussing the CMSE department as uniquely positioned at the "triple junction" of algorithm development and analysis, high performance computing, and applications to scientific and engineering modeling and data science.

The article by Wofford and Lueninghoener describes methodology followed at the Supercomputing Institute to teach cluster computing to undergraduate and graduate students. They described the redesigned boot camp curriculum and the two-fold approach of the new curriculum, which are to extend the content of the lectures to include more technical depth and more technical areas; and to replace the labs with practica (staged guides that have a mix of free

exploration prescribed steps). They conclude by providing qualitative and quantitative results indicating the positive impact of the new curriculum on the program over recent years.

The article by Betro and Loveless describes the development of a STEM ecosystem where both the science department and math department of Baylor school have implemented an interdisciplinary approach to introduce a spectrum of laboratory and computing research skills. They outline the benefits of this ecosystem that has been an effective tool in allowing several driven and interested students to participate in collegiate-level and joint collegiate projects involving virtual reality, robotics and systems controls, and modeling. They conclude by discussing various critical factors in readying the next generation of computing leaders.

The article by Harrell et al presents the parallel computing portion of the HPC seminar series which are used as a tool to introduce students from many non-traditional disciplines to scientific, parallel and high-performance computing. They describe the two-fold approach to their curriculum: engaging students with hands-on exercises using a real-world scientific application along with regularly lecturing on more general parallel computing topics in the class. They conclude by discussing the student evaluations and the lessons learned to give undergraduate students an opportunity to explore the field of HPC and big data in a non-traditional computer science course setting and build a basic foundation of computational and data skills for their further education and research activities

The article by Chakravorty and Pham presents a review of experiences of using Google's Qwiklabs online platform for remote and in-person training from the perspective of the HPC user. They describe the scaffolded instruction methods supported by the Qwiklabs training platform that support learners with varied skill sets. They conclude by providing recommendations on how the large-scale computing community can leverage these opportunities to work with Cloud service providers to assist researchers nationally and at their home institutions.

The article by Younts and Harrell describes the teaching methods and hardware platforms used by Purdue Research Computing to train undergraduates for HPC system roles. They present the scientific computing track, which provides students with some basic Linux skills but focuses on running and visualizing scientific codes and the HPC Systems Track, which truly focus on important aspects of building systems. They conclude by discussing the system they have designed and the failures and successes they have had teaching HPC system administrators.

The article by Kunkel discusses the current state of the developed Skill Tree in the HPC Certification Program and the process of contributing to the skills to the program. They describe the skills, organized in a tree structure from a coarse-grained to a fine-grained representation, allowing users to browse the skill based on the semantics. They conclude by reporting on the HPC certification forum which plays a virtual central authority to curate and maintain the skill tree and certificates and how the contributions to the skill definitions can be made.