Blue Waters Workforce Development
Delivering National Scale HPC Workforce Development

Jennifer Houchins
Shodor Education Foundation, Inc.
Durham, NC
jhouchins@shodor.org

Scott Lathrop
Shodor Education Foundation, Inc.
National Center for Supercomputing Applications (NCSA)
University of Illinois
Urbana-Champaign, IL
lathrop@illinois.edu

Robert Panoff
Shodor Education Foundation, Inc.
Durham, NC
rpanoff@shodor.org

Aaron Weeden
Shodor Education Foundation, Inc.
Durham, NC
aweeden@shodor.org

ABSTRACT

There are numerous reports documenting the critical need for high performance computing infrastructure to advance discovery in all fields of study. The Blue Waters project was funded by the National Science Foundation to address this need and provide leading edge petascale computing resources to advance research and scholarship. There are also numerous reports that identify the lack of an adequate workforce capable of utilizing and advancing petascale class computing infrastructure well into the future. From the outset, the Blue Waters project has responded to this critical need by conducting national scale workforce development activities to prepare a larger and more diverse workforce. This paper describes those activities as exemplars for adoption and replication by the community.

KEYWORDS

HPC, education, training, computational science education, petascale computing, broadening participation

1 INTRODUCTION

The National Science Foundation funds the 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. 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. All of these communities benefit from access to on one of the most advanced computing environments available to the open science research community. Educators, researchers and students are asked to present their research via conference presentations (e.g. the annual Blue Waters Symposium) and publications (e.g. the Journal of Computational Science Education).

https://doi.org/10.22369/issn.2153-4136/11/1/5

2 EDUCATION ALLOCATIONS

The initial proposal for the Blue Waters project requested that 1% of the available computing resources be devoted to educational activities to prepare a larger and more diverse, computationally literate workforce [12]. At that time, 1% of the system was a substantial commitment - providing more computational resources than were available to researchers via all of the other NSF funded HPC systems. This portion of the system is allotted through an education allocation application process that is available to faculty and staff at any US institution. Applications may be made to support undergraduate and graduate courses, training sessions, workshops, webinars, institutes, and Research Experiences for Undergraduates (REUs). We encourage innovative approaches to educating the community. Requests range from one day training events to a full year program of structured learning, such as through internships and fellowships. Allocation requests typically range from 5,000 to 25,000 node hours, although allocations of larger amounts have been granted for programs serving large numbers of participants or for conducting more complex semester course requirements.

3 BROADENING PARTICIPATION ALLOCATIONS

In order to engage a more diverse community of researchers, the Blue Waters project created a new allocations category for Broadening Participation [11]. The purpose was to encourage principal investigators who were women, minorities, individuals at NSF designated EPSCoR institutions [2], andor individuals at Minority Serving Institutions (MSIs) to apply for an allocation of time on the Blue Waters system. These allocations were intended as start-up allocations of up to 200,000 node-hours to allow each research team to scale up their codes to Blue Waters. Twenty-one teams from across the United States were selected in the first year. Included among the Principal Investigators [1] (PIs) were ten females and two underrepresented minorities. In addition, there were four female co-PIs and eight underrepresented minority co-PIs. Among

the lead institutions, five were Minority Serving Institutions and ten were within EPSCoR jurisdictions. After working on the Blue Waters system for nearly a year, one of the teams received a PRAC allocation [3] from NSF, which will allow the team to significantly advance their computational research in the future.

4 STUDENT ENGAGEMENT

The Blue Waters project directly engaged students through the Blue Waters Graduate Fellowship Program and the Blue Waters Student Internship Program. Both programs provide these students with a full year of financial support and access to the Blue Waters system to conduct computational research. Information about each of these student programs follows.

4.1 Graduate Fellowships

The Blue Waters project offers a unique, federally supported program that provides PhD students with a full year of computational science and engineering research support. Each fellow receives an allocation of 50,000 node-hours to pursue their computational and/or data-enabled research on the Blue Waters system [13]. The fellows also receive a $38,000 stipend. Each fellow is able to request up to $12,000 in tuition allowance to help offset their educational expenses. Each fellow is invited to attend the annual Blue Waters Symposium to make a formal presentation and display a poster to share their research progress with the other attendees. They are also encouraged to give a presentation of their Blue Waters supported research at a domain conference of their choosing. There are between four and 10 fellows selected each year through a competitive application process that is open to students in all US academic institutions. After their fellowship year ends, the fellows are encouraged to continue using the Blue Waters system to pursue their research while completing their PhD. They are also encouraged to continue using the system during a subsequent postdoctoral appointment. Many of the fellows have gone on to faculty positions, postdoctoral positions, and professional positions in academia, government agencies and academic institutions. We continue to track the progress of each fellow to facilitate a longitudinal analysis of the impact of the fellowship program.

4.2 Undergraduate Internships

The Blue Waters Student Internship Program is designed to motivate and prepare the next generation of computational researchers by engaging them in year-long research projects. The Internship Program [14] supports about 20 students each year, with a $5,000 stipend spread out over the full year of their appointment. The program welcomes applications from undergraduates at all degree granting US institutions. Each year, the program kicks off with a two-week intensive petascale institute at NCSA that also engages the interns in learning to make effective use of the Blue Waters system. The students are matched with faculty who mentor them through their year-long research projects. Towards the end of their year-long research endeavors, the students apply to present a poster on their research project at the annual Blue Waters Symposium in the May/June timeframe. The faculty report that the combination of a two-week institute and support for a full year have proven to be very effective for the students? learning outcomes and the projects the students pursue. To date, the internship program has benefitted 120 undergraduate students and resulted in numerous papers being published by the students in the Journal of Computational Science Education (JOCSE) [8] as described in the next section.

5 JOURNAL OF COMPUTATIONAL SCIENCE EDUCATION

All of the interns are encouraged to publish their research in the peer-reviewed Journal of Computational Science Education (JOCSE) [8]. In this publication, the students are encouraged to describe their experiences and the impacts of the program on their academic pursuits and career goals. JOCSE accepts articles from the international community that address the teaching and learning of computational science and engineering, the development and applications of instructional materials, projects, and innovative approaches for conducting workforce development. The editors welcome articles that address the assessment of materials or programs, methods for achieving improved learning outcomes, and innovative computational science programs. The journal articles and instructions for submissions are available at http://jocse.org/.

6 EDUCATION AND TRAINING

The Blue Waters EOT team [15] has brought together experts to offer a variety of education and training sessions throughout the year to assist researchers and educators with incorporating state-of-the-art resources, tools, and methods within their research and education endeavors.

The Blue Waters project conducts a variety of training events throughout the year to assist participants in learning computational and data-enabled science and engineering methods, tools, and resources. The training is designed to prepare participants to make effective use of computing resources, with an emphasis on petascale computing. The training events include webinars, workshops, symposia, tutorials, hackathons, and other related activities. These are delivered as in-person events, webcasts, and as self-paced tutorials.

The Virtual School of Computational Science and Engineering (VSCSE) [5] delivered graduate level computational science and HPC workshops and courses to students at colleges and universities across the United States, and to students at international locations. The VSCSE workshops were delivered using high-definition video conferencing to as many as 7 remote sites simultaneously. In total, the 20 VSCSE [5] workshops served over 5,000 people at 54 institutions. The semester courses were led delivered by HPC experts in the field and were conducted in collaboration with faculty at participating institutions in order to provide students with access to course content and mentoring that would otherwise not have been available to them. A total of 7 semester courses were delivered to over 600 graduate students at 29 institutions.

6.1 HPC University Repository

The HPC University portal [7] was established to provide a mechanism for disseminating HPC training and education material. It is built on the foundation and principles established by the Computational Science Education Reference Desk (CSERD) [6], which is among the collections funded by the National Science Digital Library (NSDL) [4] funded by NSF. The Blue Waters EOT team

recruited faculty and staff from across the United States to develop and classroom-test 30 “Undergraduate Petascale modules" [9] appropriate for teaching parallel computational modeling to undergraduate or graduate students in STEM disciplines. Training materials that were developed by Blue Waters were also posted to the repository. Education and training materials developed by the community have also been made accessible via the repository.

In addition to adding substantial education and training content to the HPC University repository, Blue Waters committed staff time to improving the infrastructure and promoting this resource. The repository has proven to be an effective tool for facilitating broad dissemination of the materials. The materials have been downloaded and used to support workforce development at the high school, undergraduate and graduate levels.

6.2 Undergraduate Petascale Curriculum Modules

To facilitate inclusion of petascale computing in undergraduate education, the Blue Waters project funded the development of a set of curricular materials that include classroom-ready, domain-specific examples of HPC applications in science and engineering [9]. These materials were developed by faculty with experience teaching HPC and designed to enable the teaching and use of HPC in undergraduate science and engineering classrooms. The materials are presented in self-contained modules that include instructor materials for domain-specific applications, starter source codes, and sample problem solutions. The modules range in topic (or content area) from parallel simulation of n-body problems to dynamic programming with CUDA. All of the modules have been catalogued in the HPC University repository [7] and are freely available for use in the classroom.

7 THE BLUE WATERS SYMPOSIUM

The Blue Waters Symposium [10] is an annual gathering of Blue Waters staff, researchers, students, and professionals from the computational science and engineering community. The Symposia participants share successes and challenges in utilizing large-scale heterogeneous computing systems. Each of the scientific teams using the Blue Waters system are asked to provide updates on how the petascale system has helped to advance their research. Nationally and internationally recognized leaders are invited as keynote speakers to present innovative, impactful, and at times controversial ideas that advance knowledge and provoke interactions among the attendees. There are numerous opportunities for the participants to discuss challenges, opportunities, and the future of scientific computing. The discussions often times result in new collaborations and cooperative ventures.

8 SUMMARY

The Blue Waters project actively recruits students, faculty, professionals, and mentors in these activities from across the United States, with an emphasis on engaging women, minorities, and people with disabilities. Since going into full-service operations in 2013, over 200 education and training allocations have been utilized for activities ranging from one-day workshops to two-week institutes. The Blue Waters project has engaged more than 3,700 people in learning to make effective use of computational and data-enabled science and engineering tools, resources, and methods. The participants in the activities came from 219 academic institutions, of which 65 are within EPSCoR jurisdictions. The impact and benefits have been widespread, 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 Blue Waters project places a high importance on sharing what we have learned to help others to be even more successful in their own endeavors. We look forward to sharing our experiences with the community and fostering an ongoing exchange of lessons learned and good practices.

REFERENCES

[1] National Center for Supercomputing Applications. 2019. Blue Waters Broadening Participation Awards. Retrieved May 31, 2019 from http://www.ncsa.illinois.edu/news/story/blue_waters_awards_21_broadening_participation_allocations

[2] National Science Foundation. 2019. EPSCoR Program. Retrieved May 31, 2019 from https://www.nsf.gov/od/oia/programs/epscor/

[3] National Science Foundation. 2019. PRAC Allocations. Retrieved May 31, 2019 from https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503224

[4] NSDL. 2019. The National Science Digital Library. Retrieved May 31, 2019 from https://nsdl.oercommons.org/

[5] The University of Michigan. 2019. Virtual School of Computational Science and Engineering. Retrieved May 31, 2019 from http://www.vscse.org

[6] Inc. Shodor Education Foundation. 2019. Computational Science Education Reference Desk. Retrieved May 31, 2019 from http://www.shodor.org/refdesk/

[7] Inc. Shodor Education Foundation. 2019. HPC University. Retrieved May 31, 2019 from http://hpcuniversity.org

[8] Inc. Shodor Education Foundation. 2019. Journal of Computational Science Education. Retrieved May 31, 2019 from http://jocse.org

[9] Inc. Shodor Education Foundation. 2019. Undergraduate Petascale Curriculum Modules. Retrieved May 31, 2019 from http://www.shodor.org/petascale/materials/modules/

[10] Blue Waters. 2019. Blue Waters Symposium. Retrieved May 31, 2019 from https://bluewaters.ncsa.illinois.edu/blue-waters-symposium

[11] Blue Waters. 2019. Broadening Participations Allocations. Retrieved May 31, 2019 from https://bluewaters.ncsa.illinois.edu/broadening-participation-allocations

[12] Blue Waters. 2019. Education Allocations. Retrieved May 31, 2019 from https://bluewaters.ncsa.illinois.edu/education-allocations

[13] Blue Waters. 2019. Graduate Fellowships. Retrieved May 31, 2019 from https://bluewaters.ncsa.illinois.edu/fellowships

[14] Blue Waters. 2019. Student Internship Program. Retrieved May 31, 2019 from https://bluewaters.ncsa.illinois.edu/internships

[15] Blue Waters. 2019. Training Offerings. Retrieved May 31, 2019 from https://bluewaters.ncsa.illinois.edu/training