Extending Access to HPC Skills Through a Blended Online Course

Steven I. Gordon
Senior Education Lead
Ohio Supercomputer Center
Columbus, OH 43212
001-614-292-4132
sgordon@osc.edu

James Demmel
Professor
University of California, Berkeley
Berkeley, CA 94720
001- 510-643-5386
demmel@berkeley.edu

Lizanne Destefano
Professor
University of Illinois at Urbana-Champaign
704 South Sixth Street, MC 308
001-217-333-9625
destefan@illinois.edu

Lorna Rivera
Senior Research Specialist
University of Illinois
Urban, IL 61801
001-217-333-9625
lirivera@illnois.edu

ABSTRACT
Extending expertise in parallel computing is critical to all those using high performance computing to gain insights into science and engineering problems. Many campuses do not offer such a course because of course load limits, a lack of faculty expertise, and/or lack of access to appropriate computing resources. MOOCs for this type of course are difficult to scale both because of constraints on computing resources as well as the need for individual attention with programming problems. Using a blended online course with collaborating faculty that offer academic credit for their students, we have been able to facilitate course participation at many institutions that might not otherwise have covered the topic. This has had a significant benefit for both the faculty and students. Our paper summarizes the nature of these impacts and offers some insights on best practices for extending technical courses to multiple institutions.

Categories and Subject Descriptors
D.1.3 Concurrent Programming – parallel programming; D.2.11 Software Architectures; D.3.3 Language Constructs and Features – patterns; D.4.8 Performance; F.1.2 Modes of Computation;

General Terms
Algorithms, Performance, Design

Keywords
Parallel programming; online learning; high performance computing

XSEDE15, July 26-30, 2015, St. Louis, MO, USA.

1. INTRODUCTION
Providing instruction on high performance computing (HPC) topics is a challenge in many institutions because of limitations in faculty expertise, access to HPC hardware and software environment. Faculty across the many disciplines that are using HPC in research also face pressures of course teaching loads in their domains making it difficult to devote efforts to interdisciplinary courses. Massively open online courses (MOOCs) have been accepted by a number of higher education institutions and related consortia [2,4,8]. This approach initially received widespread praise for taking education beyond the traditional classroom, democratizing access, and focusing on competency rather than “seat time” [6,10]. This has been followed by a wide range of criticism associated with the low completion rates for MOOC courses, potential conflicts with academic freedom, and doubts about the effectiveness of the pedagogy [3,7,9]. Reich argues that MOOC completion rates be judged by the intention of those who originally registered rather than compared with the appropriate population associated with regular university courses [11]. Dacey points to several successes with MOOCs and also indicates that there are emerging blended learning models in which a MOOC or other online materials can improve course experiences for students [5]. In this paper we describe the first XSEDE full semester course in 2013 “Applications of Parallel Computers” led by James Demmel at the University of California Berkeley. The course started out as an open, MOOC style course, creating a number of technical, managerial, and pedagogical issues. It then evolved into a blended course model in 2014 that has served students and faculty at a variety of collaborating universities. We describe both experiences and their impacts on both faculty and students.

2. FROM MOOC TO SPOC
2.1 Original MOOC Style Course
In 2013, materials for the course were assembled based on the recorded lectures for the course edited into shorter segments by the XSEDE staff at Cornell University [12]. Quizzes were added to each module to test student’s acquisition of the lecture materials and integrated with the lectures online [1]. Student participation involved using these materials and also participating

in online discussions and assignments facilitated on a Moodle server. A general announcement was made soliciting participants for this open course. Initially, a registration limit of 100 was put in place for this first trial but then raised to 350 because of demand. Eventually, 376 people started the course.

These numbers immediately created administrative and technical concerns. The large numbers raised concerns among the XSEDE service providers associated with the labor costs for creating user accounts on the supercomputers as this requires individual entry and review of accounts and, most importantly, the potential to overwhelm the batch queues at times when programming assignments were due. In response to those concerns, we postponed the creation of accounts until after course startup to see how many of those registered actually participated. Shortly after course startup, the number of active participants dropped to 145. More ominously, dropouts continued especially when the programming assignments were due. The steady decline resulted in only 36 completing homework one, 23 homework 2, and 18 homework 3. Only those 18 completed the course.

The instructional team discussed these results in terms of the overall mission of the XSEDE Education program – to provide the necessary expertise in the use of digital services to a new generation of students. Running a course in this format did not effectively meet those goals, resulting in a trial of a different, blended course approach.

2.2 A Blended Instructional Approach

Given the poor performance of the MOOC, the second offering of the course was converted to a blended approach sometimes called a SPOC (for Small Private On-Line Course). In this case, we invited the participation of collaborating faculty from higher education institutions. The local faculty had the responsibility of creating a local course number for the class so that students could register at their institution and take the course for credit. This was important in creating motivation for students to complete the course as the credits counted toward their degree and distinguished it from a MOOC for which there is no academic credit. The local faculty also served as advisors for their students, assisting them with general questions, programming assignments questions, and devising a localized set of final projects. The faculty also retained the duty of grading with the assistance of program autograders and online quiz results. This allowed them to gauge the progress of their own students and set localized expectations of performance based on their experience and previous programming background.

The XSEDE instructors still provided the lecture materials, quizzes, programming assignments, programming assignment overall grading, and consultation with both faculty and students on both technical and discussion questions. Meetings were held with the faculty once every 3 to 4 weeks to gauge course progress and address any questions arising from the course.

A total of eighteen institutions joined the course in 2014. Those included four minority serving institutions and three non-U.S. institutions. There were 158 registered students. Of those, 11 or 7% dropped out. Another four were auditing the course. Thus, 143 completed the course. In the sections below, we discuss the impacts of the course on both the students and faculty.

3. Evaluating Course Impacts

3.1 Course Surveys

As part of the ongoing evaluation of the XSEDE Training Education and Outreach effort (TEOS), led by the *Illinois*-Science, Technology, Engineering, and Mathematics Education Initiative (I-STEM), surveys were done for both faculty and students at the outset of the course, mid-course, and as follow-up after the course. Online survey invitations were sent to instructors and students during the first two weeks of the course. Invitations were sent both to teaching faculty and to faculty and graduate students in an observing role. The surveys covered course expectations and background experience. For students, the questionnaire included items about learning style, procrastination, and computational science and engineering (CSE) identity. For instructors, the questionnaire included items about how well prepared they felt for the course, as well as aspects of local implementation. While many items were intended for immediate interpretation, other items were specifically intended for explanatory analysis of end-of-course outcomes. Following the course, surveys were conducted to evaluate the outcomes in light of the earlier expectations of the participants.

3.2 Student Outcomes

For most of the students this was their first online course. Only one third of the students had taken on an online course before. Table 1 shows additional background information from the students based on a scale of 1 (strongly disagree) to 5 (strongly agree).

Table 1. 2014 Student Pre-Survey Response to Background Items, N=99

Item	Mean	Standard Deviation
I see myself working in science, technology, engineering, or mathematics.	4.61	0.53
I see myself working with computational research.	3.89	0.91
I am ready to practice CSE	4.10	0.75
I enrolled in this course because I thought it would be interesting.	4.32	0.73
I enrolled in this course because computational science or engineering (CSE) is important to my future career.	4.20	0.91

Although a large number saw themselves pursuing careers in CSE, they felt less confident about their pre-requisite skills. Their self-reported level of experience with parallel computing had a score of 1.42 where 1 was very inexperienced and 5 was expert. Starting expertise was better in linear algebra at 3.10 and in C programming at 2.93. However many comments indicated that students were concerned about their level of computer programming expertise. This was partly due to the mix of students with about half being undergraduates, 22% doctoral students, and 27% who did not respond to that question. Students came from a variety of majors with 43% from computer science and the rest from engineering (17%), physical sciences (18%), biological science (6%), or other fields (14%). Also notable was the significant number of ethnic and racial minorities in the class. This is shown in table 2.

Table 2. 2014 Student Race/Ethnicity, N=90

Race/Ethnicity*	Number	Percent
White	31	34%
Asian	23	26%
Hispanic or Latino	21	23%
Black or African-American	17	19%
American Indian or Alaska Native	4	4%
Other	4	4%
Native Hawaiian or Other Pacific Islander	1	1%

**Respondents could select more than one option.*

The follow-up survey for students had only a 15% response rate so it is hard to generalize to the total population of students completing the course. Nevertheless, there are some interesting observations that can be made. Of those who responded 84% indicated that said the course was moderately or very important to their academic studies. Sixty percent participated in a computational research project or internship following the course and 85% are interested in learning more about parallel computing. Half of those who responded indicated that they would be likely to use the techniques they learned in their future research or studies.

One additional way in which we may gauge the success of the course is the participation of the students registered in the course for other XSEDE training events or online materials. These were obtained by tracking student registrations for those events are shown in Figure 1.

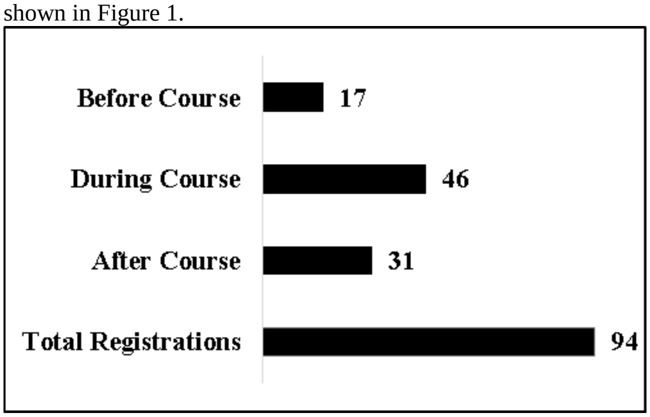

Figure 1: 2014 Registrations of Course Students in Other XSEDE Training

Here one can see that the course encouraged students to undertake other training offered by XSEDE both during and after the course completion.

Finally, student success was assessed by tracking the number and type of XSEDE computational resource allocations to which they had access, excluding the course allocation itself. Figure 2 below shows these data relative to the course timeline.

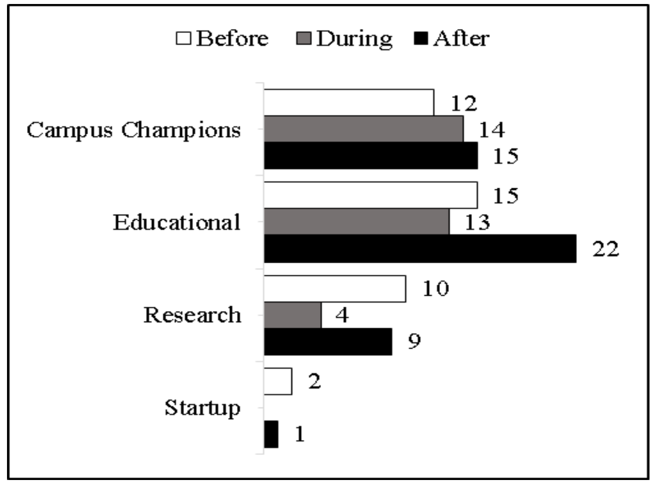

Figure 2: 2014 Course Student Access to XSEDE Allocations by Type

Evident from Figure 2, the course encouraged students to access Champion allocations both during and after the course. Most notably, students accessed educational allocations more frequently after their participation in the course.

3.3 Faculty Outcomes

Faculty were followed in a similar way to students. In the pre-course survey, 96% of the faculty expected the course to benefit them professionally and the same percentage thought it would benefit their students to a moderate or great extent. Only half the faculty had experience with online courses. In addition to becoming more experienced with the course content, some instructors looked forward to developing their teaching skills, for example:

– "This is a great future education model where I can focus more on individual student mentoring and avoid repeating lectures on common parallel computing concepts at each campus, and get to learn the latest with fellow instructors."

Some faculty also felt that the course benefitted their institution. As one person indicated, "It has been eight years since a HPC course has been offered. Within our degree program this course is currently the only feasible way this material and experience can be available to our majors." Another indicated that the course would raise the computational science expertise standard at their institution.

For the most part, faculty ran the course as a hybrid or blended course with both online and face-to-face components. Table 3 summarizes the local course implementation methods. Those included joint viewing of the lectures, discussion sessions on site, lab time, and office hours for both faculty and TAs.

Fourteen of the instructors participated in the post course survey. including only seven of the primary instructors. The overall outcomes are summarized in table 4. Where the course had problems was related to the preparation of the participating students. A significant number of the faculty felt their students were not sufficiently prepared to take the course, resulting in problems associated with the programming exercises and final projects. In large part, this can be attributed to the number of undergraduate students in the course who lacked appropriate backgrounds. Graduate students from Berkeley and the other participating campuses were better prepared even though they came from a wide variety of backgrounds. The faculty made a number of suggestions for improving the supporting materials for the course which have been implemented in the current version of the course. Those suggestions for improvement are discussed in more detail in the next section.

Table 3. 2014 Instructor Pre-Survey Response to Local Course Implementation Items, N=25

Item*	Number	Percent
Students view lectures independently on their own computers	19	76%
Students attend a discussion section on site	15	60%
Instructor/TAs keep office hours on site	11	44%
Local TA is available to assist students	7	28%
Students view lectures in a designated room at a regular time	6	24%
Students participate in on-site labs at a designated time	5	20%

**Respondents could select more than one option.*

Table 4. 2014 Instructor Post-Survey Response, N=17

Question	Mean	Standard Deviation
I would participate in a course with this format (online, multi-institutional) again.	4.19	1.01
Participating in this course was worth my time and energy as a faculty member	4.06	0.90
Overall, I feel that this course was successful.	4.06	0.83
Having multiple institutions participate in this course together was important for its success.	3.88	0.96
I felt that the students at my institution were academically prepared for the course content.	2.69	1.21

An equal percentage of the faculty (38%) felt the course benefitted them to a moderate or great extent. Even given the problem with the preparation of their students, 93% indicated that the course benefitted their students. Almost 77% indicated they are very likely to use parallel computing in their teaching and research.

The greatest impact of the course may be the long-term connection between the faculty and the XSEDE project and its services. The vast majority (over 70%) indicated that they would continue to interact with XSEDE through the web portal, by accessing XSEDE resources, or by teaching another online course. Over 58% indicated they would attend a seminar or workshop or seek an XSEDE allocation.

One measure of these longer range impacts can be observed by the compilation of data on faculty registrations for XSEDE workshops and training events. This is illustrated in Figure 3. That figure shows that the faculty involved as instructors in the parallel computing course have become active in other XSEDE training options.

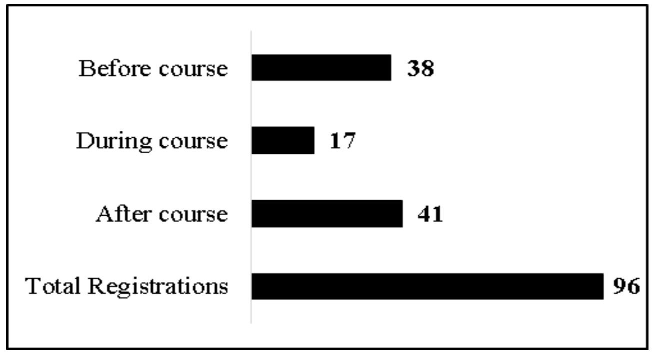

Figure 3: 2014 Faculty Involvement in XSEDE Training

In addition to training registrations, faculty were also assessed by tracking the number and type of XSEDE computational resource allocations to which they had access, excluding the course allocation itself. Figure 4 below shows these data relative to course timeline.

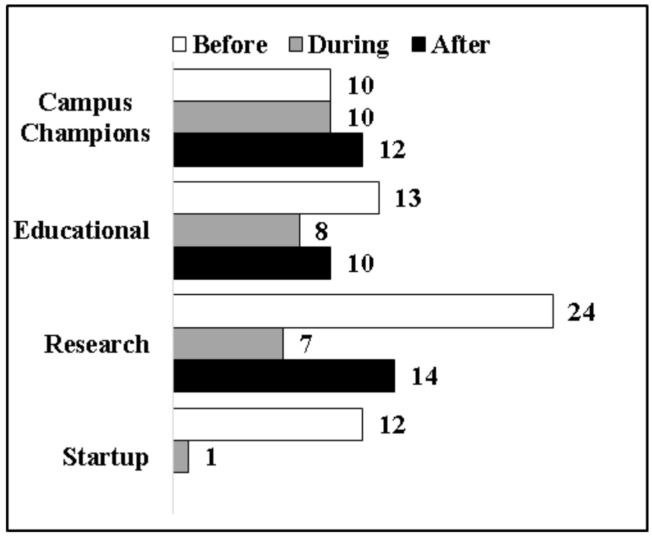

Figure 4: 2014 Faculty Access to XSEDE Allocations by Type

Evident from Figure 4, the course also encouraged instructors to access Champion allocations after the course but did not affect their access to other types of allocations.

Anecdotally, we also know that several of the participating faculty have already participated or agreed to participate in a future online course. We also know that several have participated in the XSEDE conference, in XSEDE summer workshops for faculty, and in other events.

These indicators show that the education efforts have helped us to extend the services of XSEDE into a broader community to the benefit of the faculty, students, and participating institutions.

4. ROOM FOR IMPROVEMENT

As indicated earlier, the first collaborative offering of the course for credit produced a number of suggestions from faculty that will improve the course and reduce the problems that some of the students had in keeping up with the course requirements.

During the first offering of the course, several of the faculty altered the required programming assignment requirements to better fit the experience of their students. In some cases, this meant reaching an intermediate solution that had the appropriate programming approach but did not necessarily meet the program optimization goals of the entire assignment.

Suggestions by faculty on course improvements thus included the potential to add partial solutions or related example programs that could be used to guide less well prepared students through the assignments. These examples could then be used by the local faculty to help guide their students through the assignments while allowing for possible assignment changes better suited to the background of the students.

Such examples have been added to the current version of the course. They are available to faculty on an as needed basis to share with the students that are having trouble making their way through the complete assignments.

Additional suggestions that have not yet been fully implemented are to provide tutorial guidelines for students pointing them to other online modules they can access to improve their underlying programming skills. As the XSEDE training program further develops other online offerings including possible badges for mastering those materials, they can be referred to in courses such as this one so that students who are struggling with basic programming content can enhance their skills. The possibility is also then open to making such badges pre-requisites to the full course to ensure that students have really mastered the underlying programming skills before they undertake the more complex parallel computing techniques.

More generally, we can imagine a version of this (or other courses) in which each topic presented has links to other material that goes into more detail, presents more examples, etc. Ideally this could be customized to the particular background and interests of an individual student, based on what they have looked at before (much like on-line advertising).

In this way, we hope to continue to improve and update the course content so that the experience for both faculty and students is more universally positive.

5. ACKNOWLEDGMENTS

This work was made possible, in part, through a grant from the National Science Foundation NSF OCI-0753287. Any opinions expressed in this paper are those of the authors and not of the National Science Foundation.

6. REFERENCES

[1] Applications of Parallel Computers. https://www.cac.cornell.edu/VW/apc/default.aspx. Viewed on March 17, 2015.

[2] Baggaley, Jon. 2013. MOOC Rampant, Distance Education, Vol. 34, No. 3, 368–378, http://dx.doi.org/10.1080/01587919.2013.835768.

[3] Biemiller, Lawrence. 2013. MOOC Busters. Chronicle of Higher Education. Vol. 60 Issue 15, pA23-A23, 12/13/2013.

[4] Coursera. https://www.coursera.org/. Viewed on March 17, 2015.

[5] Dacey, James. 2015. The MOOC Point. Physics World Archive, March 2014. IOP Publishing Ltd 2015 ISSN: 0953-8585.

[6] Daniel, Sir John. 2012. Making Sense of MOOCs: Musings in a Maze of Myth, Paradox and Possibility. Journal of Interactive Media in Education, JIME http://jime.open.ac.uk/2012/18.

[7] DeJong Peterson, Rachelle. 2014. MOOC Fizzles. Acad. Quest. 27:316–319, DOI 10.1007/s12129-014-9434-3.

[8] edX. https://www.edx.org/. Viewed on March 17, 2015.

[9] Klawe, Maria. 2014. The Evolving MOOC. Educause Review, vol. 49, no. 2 (March/April 2014).

[10] London, Herbert I. MOOCs on the March. 2014. Acad. Quest. 27:313–315 DOI 10.1007/s12129-014-9432..

[11] Reich, Justin. 2014. MOOC Completion and Retention in the Context of Student Intent. Educause Review Online, Published on Monday, December 8, 2014.

[12] These lectures were edited by Susan Mehringer from Cornell University and integrated with the Cornell Virtual Workshop.