Editors: Rubin Landau, rubin@physics.oregonstate.edu | Scott Lathrop, lathrop@illinois.edu

Implementing a Collaborative Online Course to Extend Access to HPC Skills

Steven I. Gordon | Ohio Supercomputer Center
James Demmel | University of California, Berkeley
Lizanne Destefano | Georgia Institute of Technology
Lorna Rivera | University of Illinois

The application of high-performance computing (HPC) simulation techniques has become an integral tool for scientific discovery across a wide range of fields. The ability to use large-scale simulations, find patterns in extremely large datasets, and reduce the costs of product design and testing has proven invaluable in such fields as physical sciences, biological sciences, and engineering.

Knowledge of the basic principles of parallel computing is crucial to making efficient and effective use of today's parallel HPC architectures, yet few university students in the relevant fields of study have access to courses on the topic. This could be due to limitations in faculty expertise, faculty commitments to teach core domain-specific courses rather than interdisciplinary offerings, or a lack of access to the appropriate software and hardware environments.

In recognition of this problem, the Extreme Science and Engineering Discovery Environment (XSEDE) education program designed and provided an online version of James Demmel's University of California, Berkeley course, "Application of Parallel Computers," which was first offered in 2013 as a massively open online course (MOOC). This article describes the initial experience with the course and how it has been revised to overcome problems with the original offering (https://cvw.cac.cornell.edu/apc/default).

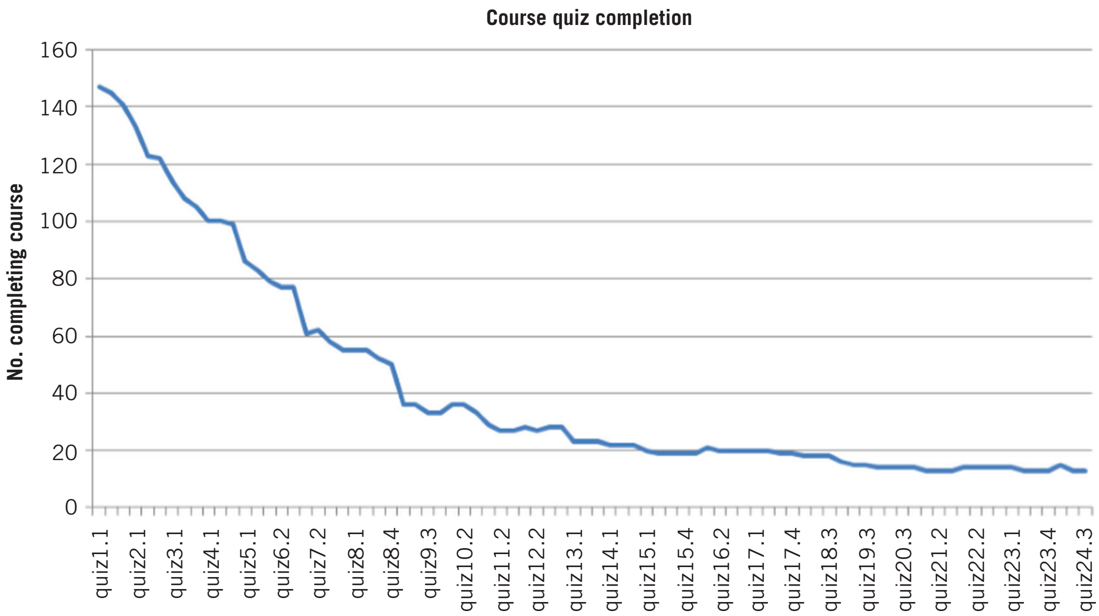

Figure 1. Course quiz completion rates for MOOC.

MOOC Experience

Jon Baggaley provides a good overview of the rapid adoption of MOOCs.[1] It has involved a large number of institutions, many of which are affiliated with one of the major online consortia, such as Coursera and edX. Although originally praised for democratizing access and their focus on competency rather than "seat time,"[2,3] MOOCs have since been criticized for low completion rates and doubts concerning the efficacy of the pedagogy.[4–6]

For the Berkeley course, XSEDE staff at Cornell University edited recorded lectures into shorter segments for online use; they then added quizzes to each of the shorter modules to test student mastery of the lecture materials. A Moodle server worked as a course management system, hosting forums for student questions and discussions and distributing and collecting programming assignments.

When the course was announced, several administrative and technical problems immediately emerged. XSEDE service providers were concerned about potential peak demands on compute resources when assignments were due; the staff time required to enter large numbers of new accounts in a short period of time was an additional challenge. An initial registration limit of 100 was subsequently raised to more than 300 due to demand. Ultimately, 376 people registered for the course.

To deal with concerns about peak loads, we delayed account creation until after introductory materials were introduced—this was under the assumption that some people who initially registered might not remain all the way through. Sure enough, shortly after course startup, the number of active participants dropped to 145. As programming assignments became due, more students dropped, a trend that continued until the end of the course, when only 18 people, or 5 percent of those who originally registered, completed it (see Figure 1).

Participants in the initial course iteration shared details of their experience through evaluation surveys conducted by the Illinois-Science, Technology, Engineering, and Mathematics Education Initiative (I-STEM) as part of I-STEM's ongoing effort to evaluate XSEDE's Training, Education, and Outreach services (TEOS). I-STEM conducted two postcourse surveys for 2013. Those participants who completed the course requirements for certification received a "course evaluation," while the remaining participants received a "general evaluation." Both groups were asked to indicate their motivations for participation, future plans, course strengths, and areas in need of improvement. The largest group of survey respondents identified as graduate students (69 percent), 8 percent as university faculty or equivalent, 7 percent as undergraduate students, Those who didn't complete course certification requirements identified their greatest barriers as being the amount of time required (88 percent) and insufficient background knowledge and skills (30 percent).

and 7 percent as postdoctoral fellows. The two groups (both "course" and "general" respondents) shared similar motivations for participating, including extending their current knowledge of the topic (76 percent) followed by general interest in the area (38 percent) and professional development (26 percent). Thirty-one percent of all respondents planned on applying for an XSEDE user allocation and 21 percent stated they already have an active XSEDE user allocation.

Those who didn't complete course certification requirements identified their greatest barriers as being the amount of time required (88 percent) and insufficient background knowledge and skills (30 percent). Regarding time commitments, these participants cited additional responsibilities such as credit courses at their local institution as taking precedence over their MOOC coursework:

- "The only reason was time. I'm a full time senior at a university who is currently working on his final senior project."
- "The main reason I didn't complete the course was that I got busy with my required first-year graduate courses. I also found that my Matlab addiction has severely hindered my ability to write good C code quickly."
- "The first homework assignment required programming knowledge that I don't have—it was essentially 'beat this BLAS benchmark,' and I would have had enough trouble with an assignment APPLYING the benchmarks. What I feel was unavoidable failure in this assignment removed the pedagogical utility of grading, as a failing course grade seemed guaranteed. I kept taking the quizzes for a while but eventually reprioritized my effort away from the class."
- "Since I don't have background knowledge in parallel computing, I need to spend much more time on this class. I'm also occupied with my own research work, so there isn't enough time for me to complete it."

Most (91 percent) participants failing to complete the course requirements were willing to take another course offered in this format, and 82 percent were interested in learning more about the resources and opportunities available through XSEDE. Respondents earning certificates stated the availability of in-person interaction with other students would have facilitated further comprehension of the more difficult material. More than half (56 percent) didn't feel they had adequate support or assistance with content from other students, and 38 percent felt isolated while taking the course.

A Blended Instructional Approach

Given the poor outcomes from the MOOC version, we converted the second offering of the course into a blended approach that's sometimes called a SPOC (small, private online course). Collaborating faculty from higher education institutions were recruited to participate via local course numbers that their students could register through to receive academic credit. Those faculty members also served as advisors to their local students to answer programming questions and supervise a locally defined set of final projects. Local faculty retained the duty of assigning grades for computing assignments with the aid of Berkeley's autograding programs and of altering assignments to suit the expertise of their local student population as needed. Thus, each participating institution could gauge the progress of their own students and set performance expectations.

The XSEDE instructors still provided lecture materials, quizzes, programming assignments, overall grading, and consultation with faculty and students on both technical and discussion questions. Meetings were held with faculty once every three to four weeks to gauge course progress and address any questions arising from the course. At the suggestion of faculty, partial solutions to programming problems were made available as needed to assist those with more limited programming experience.

A total of 18 institutions with 158 registered students joined the course in 2014. Of those, only 11 or 7 percent dropped, or, put another way, the completion rate was 91 percent.

Table 1. Student presurvey responses to background information (2014 course, $N = 99$).*

Statement	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 computational science or engineering (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 CSE is important to my future career.	4.20	0.91

*1 (strongly disagree) to 5 (strongly agree).

Table 2. Student race/ethnicity (2014 course, $N = 90$).*

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.

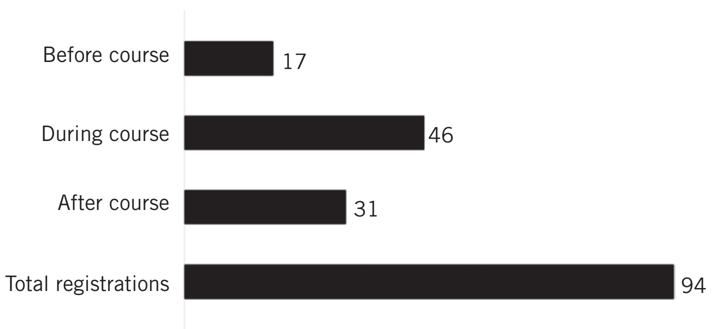

Figure 2. Students participating in other XSEDE training events. The 2014 course encouraged additional student participation.

Evaluating Course Impacts

To more closely examine the dramatic change in completion rates, the I-STEM evaluation team conducted online surveys covering course expectations and background experience for both students and faculty alike during the first two weeks of the course. 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 and aspects about the local implementation. Although many items were intended for immediate interpretation, others were specifically intended for explanatory analysis at course end. Following the course, I-STEM conducted surveys to evaluate outcomes in light of initial expectations.

In addition, the evaluation team tracked the level of participation by faculty and students in other XSEDE services, including the use of computing resources and participation in training activities. Together, these assessments helped shed light on this version of the course.

Student Outcomes

For most of students, this was their first online course—only one-third had taken one before. Table 1 shows additional background information from students.

Although a large number of students saw themselves pursuing careers in CSE, they felt less confident about their prerequisite 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 percent doctoral students, and 27 percent who didn't respond to that question. Students came from a variety of majors, with 43 percent from computer science and the rest from engineering (17 percent), physical sciences (18 percent), biological science (6 percent), or other fields (14 percent). Also notable was the significant number of ethnic and racial minorities in the class (Table 2).

The followup survey for students had only a 15 percent response rate, so it's hard to generalize to the total population of students completing the course. Nevertheless, we can make some interesting observations. Of those who responded, 84 percent indicated that 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 percent 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.

An additional way in which we can gauge course success is student participation in other XSEDE training events or online materials. Figure 2 shows student registrations for such events; we can see that students took up other XSEDE training both during and after course completion.

Finally, we assessed student success by tracking the number and type of XSEDE computational resource allocations to which they had access, excluding the course allocation itself. Figure 3 shows these data relative to the course timeline.

There are three possible options for receiving an XSEDE allocation: campus champions, the local XSEDE representative on campus, can give one to faculty and students starting on a system; education allocations can be given to teach a formal class or workshop; and research allocations can be given for research projects that undergo a peer review. Evident from Figure 3, the 2014 course encouraged students to access champion allocations both during and after the course. Most notably, students accessed educational allocations more frequently after participating in the course.

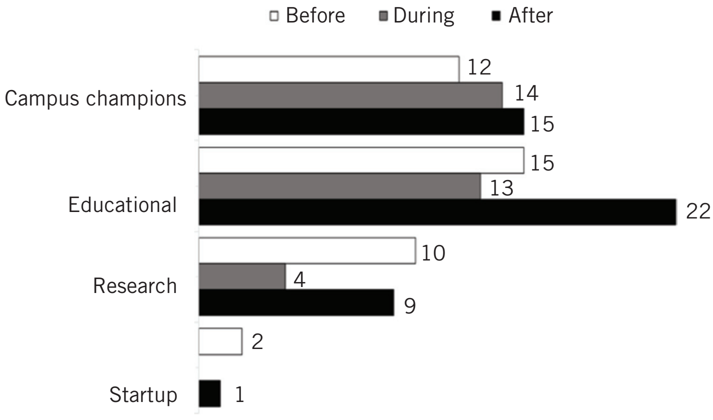

Figure 3. Student access to XSEDE allocations by type. The 2014 course encouraged participants to access champion allocations both during and after the course.

Table 3. Instructor presurvey response to local course implementation items (2014 course, *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/teaching assistants (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 onsite labs at a designated time	5	20

*Respondents could select more than one option.

Faculty Outcomes

In the precourse survey, 96 percent of 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 course content, some instructors looked forward to developing their teaching skills: "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."

Some faculty also felt that the course benefitted their institution. As one person indicated, "It has been eight years since an HPC course has been offered here. 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 his 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, which included joint viewing of lectures, discussion sessions on site, lab time, and office hours for both faculty and teaching assistants.

Fourteen of the instructors participated in the postcourse survey, including only seven of the pri-

Table 4. Instructor postsurvey response (2014 course, *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

*1 (strongly disagree) to 5 (strongly agree).

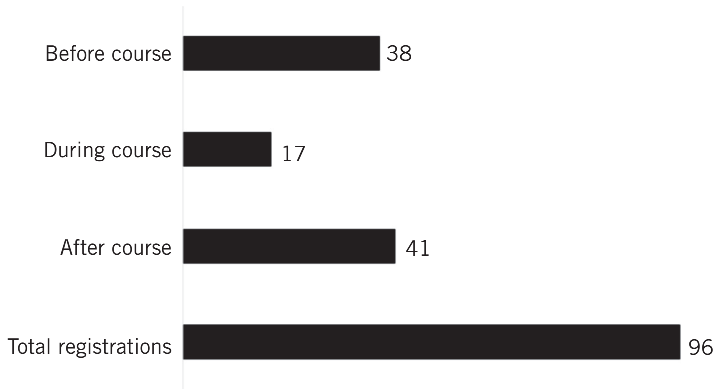

Figure 4. Faculty involvement in XSEDE training. Participants involved as instructors in the 2014 course have become active in other XSEDE training options.

mary instructors. Table 4 summarizes the overall outcomes. The course had problems related to the preparation of participating students—a significant number of faculty felt their students weren't 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 other participating campuses were better prepared, even though they came from a wide variety of backgrounds. The faculty made several suggestions for improving the supporting course materials that have been implemented in the current version, including providing additional partial solutions that faculty can pass along to students who need extra help and guidance on how to apply the autograders.

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

The greatest impact of the course might be the long-term connection between the faculty and the XSEDE project and its services. The vast majority (more than 70 percent) indicated that they would continue to interact with XSEDE through the Web portal, access XSEDE resources, or teach another online course. Over 58 percent indicated they would attend a seminar or workshop or seek an XSEDE allocation.

One measure of these longer-range impacts can be observed by compiling data on faculty registrations for XSEDE workshops and training events. Figure 4 shows that the faculty involved as instructors in the parallel computing course have become active in other XSEDE training options.

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 5 shows these data relative to the course timeline—the course encouraged instructors to access champion allocations after the course but didn't affect their access to other types of allocations

Anecdotally, we also know that several 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 extend XSEDE services to a broader community.

From our experience, it's clear that primarily self-directed, large enrollment approaches to teaching advanced computing techniques aren't an effective way to scale the number of people with the intended skills. Although a few dedicated individuals will have the prerequisite expertise, skill, and time to complete such courses, the vast majority require more guidance. The time required to complete courses such as this one without the in-

Although a few dedicated individuals will have the prerequisite expertise, skill, and time to complete such courses, the vast majority require more guidance.

centive to receive a benefit such as academic credit also contributes to this problem.

Collaborating to teach advanced computational science topics appears to be an excellent way of expanding the availability of such courses at a variety of institutions. Many of the faculty at the collaborating institutions didn't have the expertise to initiate such a course at their institution or sufficient time to prepare one. The collaboration allowed them to provide the course for students with a much lower effort than would be required to prepare and teach it themselves. In at least one instance, a faculty member used the experience to begin teaching the course on his own campus in subsequent years. Given the uneven distribution of expertise in specialized areas and the other teaching and research responsibilities of faculty members, such collaborations could provide a model for specialized graduate education beyond what the XSEDE project can offer. Interuniversity collaborative agreements could be reached to share responsibilities as lead instructors for such courses in return for participation in other courses as local instructors.

Combining online and local instruction could provide a way to scaleup the number of courses and thus students that acquire more specialized computational expertise across a variety of science, mathematics, and engineering disciplines. Our experiment clearly demonstrates the utility of this approach. ■

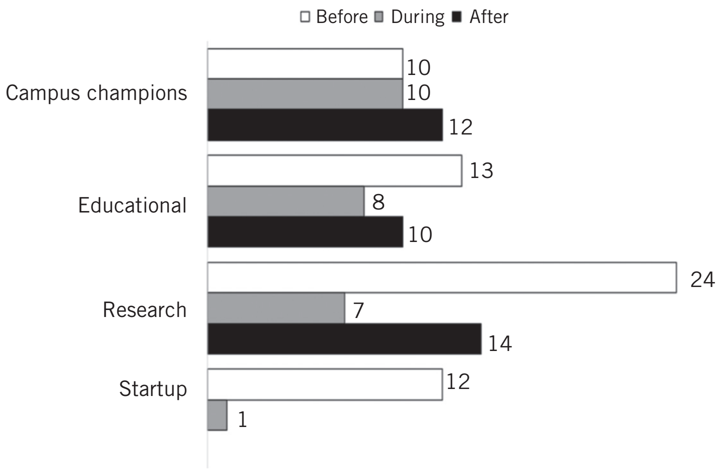

Figure 5. Faculty access to XSEDE allocations by type. The 2014 course encouraged instructors to access champion allocations after the course but didn't affect their access to other types of allocations.

References

1. J. Baggaley, "MOOC Rampant," *Distance Education*, vol. 34, no. 3, 2013, pp. 368–378.
2. J. Daniel, "Making Sense of MOOCs: Musings in a Maze of Myth, Paradox and Possibility," *J. Interactive Media in Education*, 2012, pp. 1–20.
3. H.I. London, "MOOCs on the March," *Academic Questions*, vol. 27, 2014, pp. 313–315.
4. L. Biemiller, "MOOC Busters," *Chronicle of Higher Education*, vol. 60, no. 15, 2013, p. A23.
5. R. DeJong Peterson, "MOOC Fizzles," *Academic Questions*, vol. 27, 2014, pp. 316–319.
6. M. Klawe, "The Evolving MOOC," *Educause Rev.*, vol. 49, no. 2, 2014; http://er.educause.edu/articles/2014/3/the-evolving-mooc.

Steven I. Gordon is a senior education lead at the Ohio Supercomputer Center. Contact him at sgordon@osc.edu.

James Demmel is a professor of EECS and mathematics at University of California, Berkeley. Contact him at demmel@eecs.berkeley.edu.

Lizanne Destefano is professor of psychology at Georgia Institute of Technology. Contact her at lizanne.destefano@ceismc.gatech.edu.

Lorna Rivera is a senior research specialist at the Illinois Science, Technology, Engineering, and Mathematics Education Initiative at the University of Illinois at Urbana-Champaign. Contact her at lirivera@illinois.edu.