Creating Pathways in Disadvantaged Communities Towards STEM and HPC

Elizabeth Bautista
Lawrence Berkeley National Laboratory
Berkeley, CA
ejbautista@lbl.gov

Nitin Sukhija
Slippery Rock University of Pennsylvania
Slippery Rock, PA
nitin.sukhija@sru.edu

ABSTRACT

Today's job market has its challenges in gaining proficient staff but more so in the High Performance Computing area and within a government lab. Competition from industry in terms of the type of perks they provide, being able to negotiate a higher salary and opportunities of remote work all play a part in losing candidates.

At the National Energy Research Scientific Computing Center (NERSC) at Lawrence Berkeley National Laboratory (LBNL), a site reliability engineer manages the data center onsite 24x7. Further, the facility itself is a unique and complex ecosystem that uses evaporative cooling and recycling of hot air to keep the facility cool. This is in addition to the normal areas to be monitored like the computational systems, the three tier storage, as well as infrastructure and cybersecurity.

To explore creating interest into HPC and STEM within the disadvantaged communities near the Laboratory, NERSC partnered with a community college during the pandemic to support high school seniors and freshmen students to provide an educational foundation. In collaboration with the community college, they created a program of specific classes that students needed to take to prepare them for an HPC and/or STEM internships. In certain demographics, students do not believe they can be successful in science or math and require support from the program such as tutors to help them through. With this type of support, students have successfully completed their classes with passing grades.

As part of their recruitment process for site reliability engineers to continue to support diversity initiatives at the Laboratory, NERSC implemented an apprenticeship program. This paper describes the current work that includes partnering with a community college program and then NERSC provides a summer internship for the student so they can gain hands-on experience. The first cohort of students have graduated into their internship programs this summer. This paper demonstrates early results from this partnership and how it has impacted the diverse pool of candidates at NERSC.

KEYWORDS

Site Reliability Engineer, HPC Education, HPC Training, Diversity, Inclusion, STEM

1 INTRODUCTION

According to a 2014 study of the Public Policy Institute of California, the state is likely to face a shortage of staffing, as high as 1.5 million workers who have college degrees, much more than previous projections depending on the industry [5]. In the post pandemic timeframe, most tech industry workers are seeking amenities like a hybrid or remote working environment that provides the flexibility they require incorporated into their normal life. This shortage is more prominent when recruiting for a government laboratory and in high performance computing (HPC). This industry simply does not have the salaries to compete with the neighboring Silicon Valley tech companies even if they are providing a hybrid or remote environment.

To further complicate matters, the site reliability engineer (SRE) at NERSC, requires staff to be onsite at minimum of two to three eight hour shifts to support the 24x7 requirement of the data center control room. Although the position does not necessarily require a college degree or certifications, the job description does require some system administration, networking, storage and facility management understanding at minimum for a candidate to be successful.

Diversity and inclusion has always been a strong component of the Lab's mission. In fact, the Lab has programs that specifically recruit potential candidates and support internship programs where students or candidates are from the underserved communities in the neighboring areas. In spite of these programs, however, according to a 2022 Lab study of our workforce demographics, we continue to see a small percentage of staff who are underrepresented such as Black/African American, American Indian/Alaska Native, Asian, Hispanic or Latino, etc. We see an even smaller percentage of women in these demographics especially within the Lab Senior Leadership roles [4]. The numbers truly tell a compelling story.

Prior to the pandemic, NERSC had an initiative to expand diversity for site reliability engineers through an innovative apprenticeship program [3]. Though this program was successful in filling all the spots in the control room, the pandemic has created a bigger challenge in hiring and recruiting.

How then can we increase our diversity and inclusion percentages in these areas while still being able to recruit much needed skills into our positions? NERSC continues to gather new ideas to recruit staff and also fulfill the much needed diversity requirements of the Lab. Therefore, by partnering with a community college in the neighborhood, we can potentially create a pipeline by influencing students' education with the much needed support and guidance to pass the required classes to qualify for a technical internship at the Lab.

This paper will document the process for creating this new pipeline and tells the story of early success of the first cohort of students from the program. Section 2 will provide the background of how the pandemic created situations where students were

under prepared for their technical classes and the support the program provided. Section 3 will explain the program itself and the creation of an internship program. Section 4 provides the logistics, such as funding and how the summer program worked. Section 5 will provide case studies of positive outcomes and Section 6 will provide final thoughts and future work.

2   BACKGROUND

In the three years where the world experienced the COVID-19 pandemic, there was a disruption in traditional education and learning. Hundreds of students and faculty in California and globally, transitioned to e-learning. What they did not realize is the impact of that disruption to approximately 91% of the population [8]. The assumption at the time was that since most students were used to a digital format, that learning this way, in the comfort of their home, should be easy. However, the challenge comes from not having the flexibility and capabilities to adapt to the new environment where decision making and problem solving would be tested in addition to the learning curve of dealing with technology. Learning formulas, engineering ideas and coding aren't anywhere close to playing with TikTok.

Enforcing e-learning should have lessened the impact of the school closures but no one expected to depend on the use of high-tech tools and platforms to guarantee success in teaching and learning. According to the November 2022 California test scores survey [7], there was an impact in enrollment by at least 6% across the board but a more disturbing fact is that engagement of students became average. This implies that especially in the STEM areas, the students may not have been engaged enough to learn the complex theories and processes to advance appropriately to the next series of classes.

In discussion with Laney College in Alameda, one of the neighboring schools of Berkeley Lab, they thought that perhaps if students had the support to understand STEM concepts that they would get better grades and be more confident in their learning. Further, if we can provide a cohort of students so they can support each other with a program that will prepare them for a hands-on internship, perhaps we can create a pipeline from the school into industry. As a result, Laney College created a first-year engineering program consisting of these classes:

Semester 1

- Introduction to Engineering
- Introduction to Electrical Engineering
- Statistics
- Programming for Engineers using MATLAB

Semester 2

- Engineering Graphics
- Advanced Statistics
- Mechanics of Materials
- Properties of Materials
- Advanced Programming for Engineers using MATLAB

3   ENGINEERING AT LANEY COLLEGE

Once the classes were determined and teachers were assigned to teach the classes, the program needed to create a support program for the students. This included a tutoring lab that had an engineering trained instructor as a main tutor as well as student tutors and an engineering club that engaged external speakers in the industry to engage students to learn about the type of potential jobs they can have after the program as well as engaging university representatives to encourage students to transfer to a four year university.

However, classes and outreach will do very little if there wasn't a real hands-on goal for students to work toward. The solution was to engage in an internship program with the neighboring organizations. The goal was to have the first year cohort participate in a summer internship program.

The next several months were used to engage the companies and Berkeley Lab was one of them. A job description was needed by each organization who would commit to hire interns with a specific component of hands-on work. This means, they needed to be onsite, 40 hours a week, and they would work with various mentors to learn for the next 10 weeks. Further, they were required to present their work using either a poster session or an oral presentation.

4   LOGISTICS

Growth Sector [6] is an organization that creates pathways for young people toward STEM. They collaborate with many educational institutions across the country and provide support for students to get their foot into the STEM workforce. This organization approached Laney College with a grant to help support their new engineering program cohort. They provided resources such as salary support for the tutoring lab, funding support for the engineering club and agreed to help administer the summer program including paying for the students' salaries during their ten week internship. With this type of financial support, the engineering program was geared toward success.

Employers were motivated to hire a Laney student because the salaries were paid for by Growth Sector. Students had to turn in time cards on a biweekly basis and that is approved by the supervisor of record at the organization. In some instances, they even provided the student housing and travel expenses to ensure they are closer to their employers' locations.

By week eight of the program, the students were well on their way to preparing for their presentations. Depending on the organization, a poster session or an oral presentation was going to be part of the last week of the program where students were coached on how to prepare their poster and/or their presentation speeches with power point.

5   CASE STUDIES

This section will discuss the positive outcomes of three students in the program at NERSC.

Student #1 is a female student who had a prior career in HR but she was laid off from her position at the beginning of the pandemic. Rather than getting into more debt for education, she enrolled in Laney College to get training into STEM. Apparently, this was all new therefore, she needed all the support she could get to get through her classes. Being a single parent as well, she had some challenges maintaining her grades, completing her homework and taking care of her small child. However, she persevered.

When the summer arrived, she applied to various internships at the advice of Growth Sector and Laney advisors. However, by the middle of May, she was quite disappointed that she had not

gotten an offer and she was close to giving up. The thought that perhaps, this was not going to happen this summer.

One of the tutors spoke to a NERSC manager, who already recruited two students from the program. He emphasized that this particular student was hardworking and smart. She just needs a chance. After an interview, it was determined that this student had received good grades in the classes and she participated in extracurricular activities that would enhance her studies. Therefore, this manager decided to take a chance and hire her.

Student #1 was very eager to learn anything hands-on from system administration to laying down networking cables to building servers. Half way through the program, she was invited to learn the job as a site reliability engineer. She wasn't sure about her capabilities but she decided to learn it and by the end of the program, the manager extended her time into the fall. A quote from this student at the end of the program, "I've had the simultaneous challenge of learning to leverage my old soft skills in a new context while learning new technical skills."

Student #2 was a recent graduate from San Francisco State University with a business and marketing degree. After many months of job hunting, she became discouraged and decided to pivot into STEM. She also entered Laney College to minimize taking on more debt for her education. She initially was interested in STEM but was very discouraged when the environment in the four year university did not support her dream. She was told that it was a steep learning curve and she was not confident in her skills to attain it therefore, she changed her major into business.

This particular student graduated recently but learning the classes in the cohort was a challenge. She definitely needed the support of the tutors to get through. She was chosen at NERSC to work on updating the 3D model of the HPC floor and she successfully did so by the end of the program including making suggestions to implement additional features. She was also extended after the summer program.

The last student, student #3, is a young freshman student who originally intended to study civil engineering but changed his major to computing science at Laney. Because he took some of the cohort classes, he also participated in some of the engineering club activities, including hearing one of the NERSC managers speak about careers in HPC and about the NERSC data center. After the talk, he approached her and she encouraged him to join the cohort program to take advantage of the tutoring and support, which he did. He was chosen to be one of the interns for the summer. While he consistently volunteered for much of the hands-on work, his primary goal was to research the replacement of ovirt, an opensource virtualization software management program that managed NERSC's Operations Monitoring and Notification Infrastructure (OMNI) [2]. He too successfully completed his research and summer internship and was extended through the fall to continue testing and implementation of the new virtualization software.

The three students not only presented their work at the Lab but also had the opportunity to attend the Practice and Experience in Advanced Research Computing (PEARC23) Conference in July 2023 in Portland, OR, as part of the Sixth Workshop on Strategies for Enhancing HPC Education and Training (SEHET23) where they presented their experiences as part of the program. Further, the Lab wrote a story about their success, which they proudly took to their school in the fall [1].

As a point of reference, though Growth Sector paid for the students' salaries during the summer, NERSC is currently paying for their salaries for the internship extension.

6 CONCLUSION and FUTURE WORK

We've seen the success of just three students at NERSC, however, there were approximately 45 students in the initial cohort who also had very good summer internships. Laney College will continue this engineering program pending any future funding issues. Early feedback from students at the end of the Laney program shows the following outcomes:

- All 44 out of 45 students completed the summer programs successfully. The one student who did not have family sickness and was required to leave the country.

- 10 out of 45 students were extended by their employers into the fall.

- All students had an opportunity to showcase their work through a presentation or poster session. Three students attended a conference to present their experiences.

- All students reported that they are much more confident coming back to school as a result of the summer program. They are much more motivated with school.

- All students agreed that they would not have been successful without the tutoring or support program provided by Laney.

As an extension of the program, NERSC continues to collaborate with Laney in creating the next step, which is a Data Analytics program. For NERSC, OMNI is the heart of monitoring an HPC data center that collects streaming data from everything in the facility into two streams, one real time to monitor the health of the facility and another for archiving. Visualizing this data is key for the site reliability engineers to quickly "see" what is going on and to quickly diagnose the issues to allow the facility to serve its 7000 global scientists and engineers who use the facility on a 24x7 basis.

Understanding data and visualizing data is important to NERSC now and for future work. Therefore, the next step in their staffing is to create a pipeline for students who can do this work while also being able to fulfill the diversity mission of the Lab.

ACKNOWLEDGMENTS

REFERENCES

[1] Elizabeth Ball. *NERSC summer interns feel the thrill of HPC.* Retrieved from https://cs.lbl.gov/news-media/news/2023/nersc-summer-interns-feel-the-thrill-of-hpc/

[2] Elizabeth Bautista, Melissa Romanus, Thomas Davis, Cary Whitney, and Theodore Kubaska. 2019. Collecting, monitoring, and analyzing facility and systems data at the National Energy Research Scientific Computing Center. In *48th International Conference on Parallel Processing:*

Workshops (ICPP 2019), Kyoto, Japan. https://doi.org/10.1145/3339186.3339213

[3] Elizabeth Bautista and Nitin Sukhija. 2021. Employing directed internship and apprenticeship for fostering HPC training and education. *JOCSE 12*, 2. https://doi.org/10.22369/issn.2153-4136/12/2/8

[4] Berkeley Lab. 2022. *Berkely lab workforce demographics.* Retrieved from https://diversity.lbl.gov/berkeley-lab-workforce-demographics-fy2022/

[5] Sarah Bohn. 2014. *California's need for skilled workers.* Retrieved from https://www.ppic.org/publication/californias-need-for-skilled-workers/

[6] Growth Sector.org. n.d. *Growth Sector reimagines the pathway to careers in STEM.* https://www.growthsector.org/

[7] Heather J. Hough and Belen Chavez. 2022. *California test scores show the devastating impact of the pandemic on student learning.* Retrieved from https://edpolicyinca.org/newsroom/california-test-scores-show-devastating-impact-pandemic-student-learning#:~:text=Both%20the%20COVID%2D19%20pandemic,loss%20of%20270%2C000%20students%20statewide

[8] Emad Mushtaha, Saleh Abu Dabous, Imad Alsyouf, Amr Ahmed, and Naglaa Raafat Abdraboh. 2022. The challenges and opportunities of online learning and teaching at engineering and theoretical colleges during the pandemic. *Ain Shams Engineering Journal 13.* Retrieved from https://www.sciencedirect.com/science/article/pii/S2090447922000818

[9] Northwestern Medicine. 2023. *COVID-19 pandemic timeline* Retrieved from https://www.nm.org/healthbeat/medical-advances/new-therapies-and-drug-trials/covid-19-pandemic-timeline#:~:text=By%20March%202020%2C%20the%20World,COVID%2D19%20outbreak%20a%20pandemic