

Enhance Computing Curricula with High Performance Computing Teaching and Research

Dr. Suxia Cui, Prairie View A&M University

Suxia Cui is an associate professor in the Department of Electrical and Computer Engineering at Prairie View A&M University (PVAMU). She joined PVAMU right after she obtained her Ph.D. degree in Computer Engineering from Mississippi State University in 2003. Her research interests include image and video processing, data compression, wavelets, computer vision, remote sensing, and computing education. Her projects are currently funded by NSF, United States Department of Agriculture, and Department of Education.

Dr. Lin Li, Prairie View A&M University

Lin Li is an associate professor of the Computer Science Department at Prairie View A&M University. He received his Ph.D. in computer science from the University of Nebraska at Lincoln in 2004. Before that, he received his B.S. and M.E. from Beijing Institute of Technology and Chinese Academy of Sciences in 1996 and 1999, respectively. His research interests are in Computer Networks, Machine Learning, and Computer Educational Technologies.

Dr. Lei Huang, Prairie View A&M University

Lei Huang is an assistant professor in the Computer Science department at Prairie View A&M University. He received his Ph.D. in computer science from University of Houston in 2006, and M.S. in computer science from Southwest Jiaotong University, China in 1997. His research areas are in High Performance Computing, Cloud Computing, Big Data Analytics, programming models and compiler optimizations.

Dr. Yonghui Wang, Prairie View A&M University

Dr. Yonghui Wang received the B.S. degree in Optoelectronics from Xidian University, Xi'an, China, in 1993, the M.S. degree in electrical engineering from Beijing Polytechnic University, Beijing, China, in 1999, and the Ph.D. degree in computer engineering from Mississippi State University, Starkville, MS, in 2003. From 1993 to 1996, he was a Research Engineer with the 41st Electrical Research Institute, Bengbu, China. From July 1999 to December 1999, he worked as an IT Specialist in IBM China, Beijing, China. From 2000 to 2003, he was a research assistant with the Visualization, Analysis, and Imaging Laboratory (VAIL), the GeoResources Institute (GRI), Mississippi State University. He is currently an Associate Professor with the Department of Engineering Technology, Prairie View A&M University, Prairie View, TX. His research interests include digital signal processing, image and video coding, and wavelets.

Enhance Computing Curricula with High Performance Computing Teaching and Research

Abstract

Today's scientists and engineers depend increasingly on information and tools made available through new and advanced computing technologies, such as networks, large data analysis, and sophisticated simulation tools that assist in the understanding of natural phenomena. High Performance Computing (HPC) now plays a critical role in enabling such scientific and engineering inquiry. However, undergraduate students are still lacking experience in how HPC functions, because our current computing curricula do not adequately cover HPC. To solve this problem, a team of faculty members obtained external funding to improve undergraduate computing education through enhanced courses and research opportunities. The goal is to incorporate HPC concepts and training across the computing curricula in multiple disciplines to motivate students' interests in computing and increase their computing problem-solving skills, thus strengthening and diversifying the future U.S. workforce.

This is a collaborative project with three participating departments: Electrical and Computer Engineering, Computer Science, and Engineering Technology. After the first project year, a diverse environment was established with HPC cluster and embedded HPC platforms. Both platforms supported students' research projects in parallel programming, embedded systems design, and data cloud. In the past year, the project was successfully introduced in undergraduate classes. New course materials integrating parallel and distributed computing concepts were developed and offered to undergraduate students. Class surveys were collected to guide future development. Based on the results, more courses will be revised to accommodate HPC contents in the coming years. A project-based learning scheme will also be introduced to our new course design and implementation including subjects like Computer Vision and Machine Learning. This article presents the current outcomes and findings of the project and a detailed plan of the ongoing education and research activities.

Background

HPC technology has moved beyond bulky multi-rack supercomputers¹, making its way into ever smaller systems, in particular, embedded devices and appliances^{2, 3}. Today, with more compact and powerful embedded processors, embedded systems are becoming HPC-capable⁴. HPC enables U.S. military airborne vehicles to navigate without pilots, home appliances to become "smarter," and security systems to recognize people immediately after they enter the range of vision. In medical applications, large magnetic resonance imaging and CT body imaging scans previously required several days to be processed; now processing can be done within 24 hours, thus shortening the diagnosis cycle. In the case of mobile customers, industries are building HPCs which can fit in the palm of the hand. However, even as HPC is becoming dominant in computing, HPC education is lagging far behind, especially at minority serving institutions.

With the dramatic advancement of computer technology, industry is in great need of graduates with computing-related majors. Nationally, among the fastest-growing occupations, employment

for computer software engineers and engineers for computer information systems and applications is projected to increase 34% from 2008 to 2018. The U.S. Bureau of Labor Statistics forecasts that by 2018, there will be 689,900 jobs created in this category. Due to the aging workforce and emerging techniques^{5, 6}, industry demand for qualified graduates with expertise in computer science, computer engineering and technology is tremendous. In the meantime, studies have shown that computing classes frequently teach students outdated content that is not relevant to current technologies^{7, 8}. As a result, computing majors have been reported by many institutions to have high drop rates and low enrollment especially among females and minorities^{9, 10}. The shortage of qualified computer-related workforce in the U.S. and challenges from rapidly growing countries have resulted in more computing jobs being shifted abroad every year. This tendency will cause U.S. to gradually lose its lead in the computer industry. To help solve the problem, a group of faculty members from Prairie View A&M University acquired funding from the National Science Foundation (NSF) to establish HPC teaching and research platforms and enhance undergraduate computing curricula over three departments.

Prairie View A&M University is the second oldest higher education institution in Texas, founded in 1876. With an established reputation for producing engineers, nurses, and educators, Prairie View A&M University offers baccalaureate degrees in 50 academic majors, 37 master's degrees and four doctoral degree programs through nine colleges and schools. Prairie View A&M University has produced more African American three-star generals than any other Historically Black College and University (HBCU) in the country. Its total fall 2013 enrollment was 8,250 students with 6,731 undergraduates, and its student population is 85.1% African American. The College of Engineering (COE) at Prairie View A&M University consists of six departments— Chemical Engineering, Civil Engineering, Computer Science, Electrical and Computer Engineering, Engineering Technology, and Mechanical Engineering. The overall fall 2014 undergraduate population in the College of Engineering is 1,332.

The Department of Electrical and Computer Engineering (ECE) offers B.S. degrees in the Electrical Engineering (ELEG) program and the Computer Engineering (CPEG) program, respectively. Both programs are accredited nationally by ABET. In addition, ECE department offers M.S. and Ph.D. degrees in Electrical Engineering, with a current graduate enrollment of 55. The Department of Computer Science (CS) offers three degree programs: a B.S. in Computer Science (ABET accredited), an M.S. in Computer Science, and an M.S. in Computer Information Systems. The department offers many areas of concentration, including: Artificial Intelligence, Cloud Computing, Computer Graphics and Animation & Computer Visualization. It also offers computing courses to the entire university to satisfy a three-hour core curriculum requirement. The Department of Engineering Technology (ET) at Prairie View A&M University offers the B.S. degree in the Computer Engineering Technology (CPET) program and in the Electrical Engineering Technology (ELET) program, respectively. Both programs are ABET accredited. The department focuses on offering practical, hands-on experiences in computer science and engineering. This project ties the computing programs (CPEG, CS, and CPET) from the above mentioned three departments. Four faculty members from those three departments are contributing to the project with their respective expertise. Worthwhile to mention is that the three departments also collaborated to obtain a HPC cluster through NSF Major Research Instrumentation (MRI) program.

The three departments' fall 2012 undergraduate enrollment, categorized by gender and racial/ethnic groups, is listed in Table 1. CPEG and CPET program enrollments are enclosed in parentheses beside their department enrollment. Of all the three programs, enrollment of women is 25% and African American students constitute 90.3% of total enrollment. Together, the three departments constitute 44% of all COE undergraduate enrollments.

Category	ECE(CPEG)	CS	ET (CPET)
Female	62 (27)	31	24 (17)
African American	245 (96)	81	96 (70)
Hispanic	14(5)	1	2 (0)
Total	267 (106)	93	107(72)

Table 1: Fall 2012 Undergraduate Enrollment of CPEG, CS, and ET programs

Project Activities

The goal of this project is to increase the number and ability of underrepresented students (African American, Hispanic, and women) in computing disciplines, who learn the HPC hardware, software, and applications skills and demonstrate that they can utilize HPC to solve engineering problems. The following outcomes are anticipated by the end of this project:

- (1) Deploy a novel HPC research and educational platform through engineering computational laboratory;
- (2) Revamp existing courses and develop new computing courses in the context of HPC across three departments;
- (3) Improve student performance and critical thinking in HPC related areas;
- (4) Involve undergraduate students in research, and prepare them with lifelong learning skills;
- (5) Propel faculty professional development and establish a scholar community to promote faculty collaborations through training workshops;
- (6) Increase the awareness of cutting-edge computing techniques to K-12 students as well as general public.

To accomplish the outcomes, HPC teaching and research platforms will be constructed, and a sequence of courses will be redesigned. The two HPC platforms are initialized in both the scales of computing cluster and embedded system. As shown is Figure 1, HPC cluster platform is a 24 nodes IBM iDataPlex, and embedded HPC platform is an unmanned underwater vehicle with ARM technique. These platforms can support HPC parallelism productivity using OpenMP, MPI, and Pthread as well as develop computer vision and machine learning courses and research. Table 2 lists the undergraduate required courses that will be impacted across three departments. Under the "New Content" list, major HPC concepts are listed which are missing in the current curricula and will be incorporated in the revamped courses as sample topics. New class projects centered at HPC techniques will be introduced in lecture classes. Every lab will be enabled to run experiments under HPC embedded and cluster platforms. These listed courses are regularly offered once or twice per year. After three years, a total of around 450 undergraduate students will be equipped with HPC skills from the required courses. This curriculum enhancement is expected to significantly improve the students' skills and confidence, making them better prepared to excel in the computing workforce or continue to graduate school.



(a) HPC cluster platform



(b) Embedded HPC platform

Figure 1 Two HPC platforms

Table 2: Key Computing Course Involved

Dept.	Courses	New Content	Enroll.
ECE	ELEG 3073&3071 Microprocessor	Multi-core microprocessor design;	30
	Systems Design & lab	Task parallelism; Cloud computing;	
	ELEG 4253 Embedded Systems	Embedded HPC; Data parallelism;	20
	Design	Computer Vision	
CS	COMP 1224 Computer Science II	Parallel Computing	60
	COMP 3053 Algorithm Analysis	Parallelism; Cloud computing;	20
	COMP 4063 Artificial Intelligence	Parallel Programming; Machine	20
		learning	
ET	CPET 4053&4051 Computer	Multi-core; Cloud computing;	20
	Systems Design & Lab	Parallelism	

In addition to the above required courses, each department will offer a senior-level special topics class as technical elective. The College of Engineering offers a five-year B.S./M.S. combination program in which undergraduate students can take a specified number of graduate courses as upper-level electives and the credits can be counted for both the undergraduate and graduate degree plans. The goal is to accelerate the students' progress toward achieving advanced degrees. Our project is in synergy with this effort and will positively impact the engineering undergraduate program. Each proposed new course will be an elective offered for both undergraduate students; once enrolled, undergraduate students can learn from and work with graduate students in the same course, allowing them to experience graduate study. These courses will also provide undergraduate students with opportunities to participate in advanced research and create a pathway to graduate study. All of the courses developed will be open to the entire college. The new electives are: Introduction to Parallel Computing, Machine Learning, and Introduction to Computer Vision. All of these courses will not only introduce HPC and engineering applications to undergraduates, but also help to attract more students into graduate programs.

In the first project year, the implementation focuses on integrating HPC components into the following courses: Computer Science II (COMP 1224) and Algorithm Analysis (COMP 3053). The courses were revised with new syllabi by integrating parallel computing concepts; Introduction to Parallel Computing (COMP 4073) was developed as a new course; Embedded Systems Design (ELEG 4253) and Computer Systems Design (CPET 4053) were revamped with pilot modules. The above mentioned classes were offered in spring and/or fall 2014 semesters. Students' feedbacks were collected through class surveys.

Computer Science II (COMP 1224) is a four credit-hour class with four hours for teaching and one hour for computer laboratory. It is continuation of CS1 "Computer Science I" with continued emphasis on program development techniques, array based lists, pointers, basic linked lists, classes, abstraction, data hiding, polymorphism, inheritance, stacks and queues. There were 34 students enrolled in spring 2014 semester when basic parallelism concepts, including threads, data sharing, synchronization, and thinking problem solving in parallel were introduced. In particular, OpenMP was presented to students and it was much easier for them to understand through a hands-on practice. Figure 2 shows an OpenMP example to calculate PI. It stimulated students' interests when they found the performance can be doubled or increased by four times in their laptops after adding a single line of OpenMP directive (line #11).

```
Static long num_steps = 1000000000:
double step;
int main ()
{
       int i:
       double x, pi, sum = 0.0;
       double dif:
       time t start, end;
       time(&start);
       step = 1.0/(double) num_steps;
       #pragma omp parallel for private(x)
       reduction(+:sum)
       for (i=0; i \le num\_steps; i++)
               x = (i+0.5)*step;
               sum = sum + 4.0/(1.0 + x^*x);
        }
       pi = step * sum;
       time(&end):
       dif = (double) difftime (end,start);
       cout << "The value of PI is " << pi << endl;
       printf("It took %.2lf seconds.", dif);
       return 0;
```

Figure 2. OpenMP example to calculate $PI(\pi)$

Feedback: Survey and test results show that it is needy and feasible to introduce the parallel computing concepts to students at the early stage. Students have shown enthusiasm in studying

parallel programming skills after they found the significant difference in performance, compared with sequential code. OpenMP is much easier to understand and can be used shortly after introducing the syntax. However, multithreading is confusing to most of them.

Algorithm Analysis (COMP 3053) is a very popular course in the computer science curriculum. Students in this class developed interest in programming and increased their problem solving skills. There were 20 students in spring 2014 when parallel algorithms including parallel sort and parallel search were added to the class contents. Students were asked to follow an algorithm by themselves (sequential) and then think about how to go parallel if working in a team (parallel). They compared the parallel and sequential algorithms, explored speedup concept, discussed if/how these classical sequential algorithms could go parallel, and debated whether an algorithm is efficient or not in parallel architectures. The goal of these adjustments would let students take the parallelism into consideration when designing an algorithm in solving a problem.

Feedback: Compared with CS II students, the introduction of parallel programming becomes more challenging since more students in this class get used to sequential programming. The comparisons of sequential and parallel algorithms help a lot for them to understand the parallel concepts. Students are more interested and clearly understand parallel algorithms when they work in a team to mimic the parallel process. More practices of using multithreading and MPI are needed.

A new course Introduction to Parallel Computing (COMP 4073) was offered to students the first time at PVAMU computer science department in fall 2014 as an upper-level elective. There were 14 students registered in the course. In this course, the history of HPC, architectures, software packages, parallel programming models and tools, parallel program design, and parallel algorithms were systematically introduced. The students in the course understood the performance of parallel programs, debugging and performance tuning methods. The course topics included parallel computer architectures, parallel programming models in both shared and distributed systems, synchronizations, load balancing, and parallel algorithms. Hands-on parallel programming exercises were designed using OpenMP in the cloud computing platform¹¹. In the class project, a big data problem was assigned to students and let them find the 10 most popular words in a 10 GB text file. They were free to use any programming languages and hardware to achieve the best performance. Some students got good performance and were familiar with parallel programming and performance tuning.

Feedback: Based on the survey data collected and shown in Table 3 and 4, students believe that they have learned some basic HPC concepts and definitely increased their interests in HPC area. Most of them think that they know a great deal of real world HPC applications, and believe that the HPC knowledge will help them better solve problems and be more marketable after graduation. They would also like the HPC knowledge to be included into more courses. For example, some of them think "the class project needs more knowledge such as advanced data structures, and algorithms,""Linux programming environment is also important to implement the project on a cluster," and "the new technologies in HPC such as MapReduce, PGAS, Spark are interesting too."

#	Survey Questions	Conduction
1	Consider your level of awareness about High Performance Computing (HPC) both BEFORE and AFTER this class.	Pre & Post
2	Consider your level of interest in HPC both BEFORE and AFTER this class.	Pre & Post
3	Use the scale to indicate the extent of your gains in understanding of HPC hardware architectures	Post
4	Use the scale to indicate the extent of your gains in understanding of HPC system software	Post
5	Use the scale to indicate the extent of your gains in understanding of HPC real-world applications	Post
6	This class helped me understand the value of the cutting-edge HPC approach to solving problems?	Post
7	Knowing more about using HPC techniques will make me more marketable when I graduate?	Post
8	I would like HPC to be taught in more classes?	Post

Table 3: Pre and Post survey questions of COMP4073

Question #	Know HPC concepts	Know HPC	Know only a few	Only heard	Never heard
	and applications	concepts	about HPC	HPC term	
1 pre		4	3	1	5
1 post	8	5			
Question #	Very Interested in HPC	Interested in HPC	A little interested in HPC	Not at all interested in HPC	Never heard anything for HPC
2 pre		4	3	3	3
2 post	3	6	3	1	
Question #	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
3	2	4	6	1	
4	4	4	4	1	
5	5	2	5	1	
6	5	7	1		
7	5	5	2		1
8	5	6	2		

Table 4: Pre and post survey stats of COMP4073

Embedded Systems Design (ELEG 4253) is a senior level class required for Computer Engineering students and as a technical elective for Electrical Engineering students. The class teaches Microprocessor and Microcomputer structures and applications; programming and design of hardware interfaces. There were eight students enrolled in fall 2014. Two new learning modules related to HPC were pilot tested in both classes: One is HPC benchmark, and the other is Internet parallel searching engine.

Feedback: A survey was conducted at the end of the class with the same questions as shown in Table 3. Survey results are listed in Table 5. From their comments, most of the students said the parallelism modules helped them to understand the value of the cutting-edge HPC approach to solving problems. They also realized HPC applications have entered all categories of engineering, and knowing more about the HPC techniques will make them more marketable when graduated.

Question #	Know HPC concepts	Know HPC	Know only a few	Only heard	Never heard
	and applications	concepts	about HPC	HPC term	
1 pre		1	3	3	1
1 post	4	3	1		
Question #	Very Interested in	Interested	A little interested	Not at all	Never heard
	НРС	in HPC	in HPC	interested in	anything for HPC
				НРС	
2 pre		3	2	2	1
2 post	4	3	1		
Question #	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
3	2	3	1	1	
4	3	3	1		
5	4	2	1	1	
6	5	2	1		
7	3	3	2		
8	4	3	1		

Table 5: Pre and post survey results of ELEG 4253

Computer Systems Design & Lab (CPET 4053 & 4051) are senior level Computer Engineering Technology required courses. The courses cover modern digital design methodologies, microprocessor operation, arithmetic operations, and interfacing. There were nine students enrolled in fall 2014. For this semester, the concepts and applications of HPC cluster and embedded HPC were introduced through class demonstrations.

Feedback: From the students' survey obtained at the end of semester, it is obvious that before this class, the students lacked knowledge in HPC. This was greatly improved at the end of the semester. One student wrote, "I think the way this course is set up is actually a very good way for student who want to learn about HPC to get knowledge they need."

Question #	Know HPC concepts	Know HPC	Know only a few	Only heard	Never heard
	and applications	concepts	about HPC	HPC term	
1 pre			2	5	2
1 post	5	3	1		
Question #	Very Interested in	Interested	A little interested	Not at all	Never heard
	НРС	in HPC	in HPC	interested in	anything for HPC
				HPC	
2 pre		1	4	2	2
2 post	3	4	2		
Question #	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
3		2	7	1	
4		1	6	1	
5	1	1	6	1	
6		7	2		
7	1	8			
8	2	7			

 Table 6: Pre and post survey results of CPET4053 & 4051
 Presented

Based on the experience obtained from 2014, Microprocessor Systems Design & Lab (ELEG 3073 & 3071) were revamped to introduce parallel computing software and HPC cluster hardware. This course and lab are currently offered in spring 2015. Meanwhile, a new course Machine Learning (COMP4073) was also developed and offered in spring 2015 to Computer Science junior and senior students as an upper-level elective. A project-based learning strategy has been adopted for teaching the course to enable students gain more first-hand experience by using Java, R, Hadoop, Spark, etc to implement the learning algorithms and solve practical problems. The goal is to help students "learn from doing" and broaden their knowledge in using various software tools including HPC platforms. Class surveys will be collected by the end of each semester to evaluate the course.

Conclusion and Future Plan

After the first year of implementation, the concept of parallelism was successfully integrated into various undergraduate courses. Students' feedback was positive. In the future, more courses are to be revamped and developed by infusing HPC contents. The results will be disseminated through summer workshops. A two-day summer workshop will be organized each year during the? two project years. Sample topics covered in the workshops include: 1) current HPC development; 2) multi-core and pipelining to improve throughput; 3) parallelism and cloud computing; 4) energy saving simulation and solution through HPC; 5) applications in computer vision and machine learning; 6) K-16 educators' role in computing workforce shortage. Six to 10 participants will be invited to attend each summer. All the teaching and research modules developed in this project will be introduced in the workshops Assistance will also be provided afterward to help the participating schools to adopt the modules in their courses to broaden the impact on more undergraduate and high school students.

Another aspect of the future plan is to continue enhancing undergraduate students' research capability on HPC projects. Aiming at producing qualified graduates, educators must create an environment to involve students. The first year initialized the necessary courses to deliver HPC knowledge. Starting from the second year, students will have some basic ideas about HPC from their classes and will be motivated to do more on research projects. Each undergraduate involved in this project will be assigned an HPC-related research topic with a faculty mentor. They will work closely with faculty and graduate students to build their confidence in HPC research. Research reports and publications will be presented to measure the effectiveness of the research experience. In the meantime, their research will produce preliminary results for new hands-on projects. As of now, one LSAMP scholar has been trained through this project. Four other undergraduate students are supported by the grant to conduct HPC research. They are assisting the faculty to build embedded HPC system, develop parallel software, test learning algorithm, and maintain equipment. All of these research topics give them a valuable experience for their future careers.

Acknowledgment

This project is supported in part by National Science Foundation award # 1229744. The HPC cluster is funded by NSF MRI project with award # 1332566. Opinions, findings, and

conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Bibliography

- 1. Texas Advanced Computing Center (TACC) http://www.tacc.utexas.edu
- 2. B. B. Brey, "Intel Microprocessors Architecture, Programming, and Interfacing", 8th edition, Prentice Hall, 2008 (ISBN: 0-13-502645-8).
- 3. M. Feldman, "HPC Going Embedded", HPC Wire, Dec. 01, 2011; http://www.hpcwire.com/hpcwire/2011-12-01/hpc_going_embedded.html
- 4. P. Decher, "Embedding HPC: A rocket in your pocket", Embedded, Nov. 29, 2011. http://www.embedded.com/design/prototyping-and-development/4230994/A-rocket-in-your-pocket#
- 5. ICAF Industries Studies 2006 Report, National Defense University.
- 6. ICAF Industries Studies 2001 Report, National Defense University.
- S. Cui, Y. Wang, S. Koay, and Y. Yang, "Revamp Computer Education with Multimedia and Game Technologies," *Proceedings of ASEE 117th Annual Conference & Exposition*, Louisville, Kentucky, June 20-23, 2010.
- S. Cui, Y. Wang, F. M. Nave, and K. T. Harris, "Teach Computer Techniques through Multimedia," *Proceedings of the 2010 ASEE Northeast Section Conference*, Boston, MA, May 7-8, 2010. http://acamedics.com/proceedings/aseene2010/papers/faculty_papers/80.pdf
- 9. G. Carmichael, "Girls, Computer Science, and Games", *ASM SIGCSE Bulletin*, vol. 40, No. 4, December 2008. pp. 107-110.
- 10. D. Grant, F.C.Payton, "Career Staging for Girls Moving Toward (Away) from Computing Careers", *Special Interest Group on* Computer Personnel Research Annual Conference, Proceedings of the 2008 ACM SIGMIS
- 11. Lei Huang and Yonggao Yang, "Facilitating education using cloud computing infrastructure." Journal of Computer Sciences in Colleges, published in Vol. 28, Issue 4, April 2013, pp. 19-25