AC 2009-1434: DEVELOPING A RESEARCH AND EDUCATION LABORATORY FOR HIGH-PERFORMANCE COMPUTING AND CYBER INFRASTRUCTURE

Thomas Hacker, Purdue University, West Lafayette

Dr. Thomas Hacker is an Assistant Professor in Computer & Information Technology, and a Research Assistant Professor in the Discovery Park Cyber Center at Purdue University in West Lafayette, IN. Dr. Hacker's research interests include high performance computing, high performance networking, grid computing, and operating systems.

Krishna Madhavan, Purdue University

Dr. Krishna Madhavan is an Assistant Professor in the Department of Science and Engineering Education at Clemson University in Clemson, NC. Dr. Madhavan's area of interests involve Cyber-tools and Cyber-environments for Engineering Education, and Semantic Grids.

Developing a Research and Education Laboratory for High Performance Computing and Cyberinfrastructure

Abstract

High performance computing (HPC) and computational science are critical drivers of economic and research competitiveness in global science and engineering. The growth of open source software and the universal availability of low cost, high performance computer components make it possible to build powerful and inexpensive high performance computing systems. Fully exploiting the power of these systems, however, is a significant challenge. We are developing a high performance computing and cyberinfrastructure signature area, and as part of this effort, working to develop a flexible world-class laboratory that can be used for both research and education. The goals of this effort are to: (a) educate a cohort of students from backgrounds in computing and the domain sciences in the development, deployment, and use of high performance computing and cyberinfrastructure systems; (b) establish a research facility to support discovery in high performance computing, cyberinfrastructure, and bioinformatics among a group of faculty and graduate students; and (c) seek new ways to effectively and efficiently share and leverage limited space, equipment, and high-end computing resources. We are using the laboratory for two hands-on courses and four research projects, and have found that the flexible and reconfigurable concepts we have embedded into the design of the laboratory have already proven their worth. In this paper, we describe our efforts in developing this lab, the challenges we face, specific goals and objectives, and specific outcomes we have already observed in the process of developing and utilizing this lab. The results presented in this paper will be useful and interesting to groups seeking to develop curriculum and research programs in high performance computing and cyberinfrastructure.

1.0 Introduction

The department of Computer and Information Technology in the College of Technology at Purdue University is in the process of developing a new curriculum in the areas of high performance computing and cyberinfrastructure. This effort, which is a collaborative effort that includes faculty from across the College of Technology as well as researchers and faculty from Purdue University Discovery Park, and Rosen Center for Advanced Computing, is working to integrate research and development efforts in the area of cyberinfrastructure and high performance computing that are taking place across campus, which are focused on developing information technology solutions to address significant problems in science and engineering.

As part of this effort, two years ago the College initiated plans to develop a research and education laboratory that would aid these efforts, and become a focal point of activity that would allow students to become involved. Construction on the laboratory has been recently completed (August 2008), and the facility has been in use for one complete semester and two months of the second semester. This paper describes the motivation that led to the creation of the laboratory, the features of the facility that we designed to meet the competing needs of research, education, and collaboration in a shared space, and some of the initial results and outcomes from the use of the lab thus far.

2.0 History of the facility and motivations for its development

The Purdue College of Technology was established in 1964 with the founding vision of providing excellent application-oriented degree programs. One of the critical aspects of the programs across the entire college is an emphasis on the incorporation of hands-on laboratory work in combination with a traditional theory-based lecture. This theory to practice approach is one of the distinguishing features of the college, and has proven to be useful and valuable for students and organizations who seek out graduates of the college. The department of Computer and Information Technology (CIT) within the college focuses on learning through applied, hands-on research and education that allows students to see and experience the interconnections between practice, the emergence of difficult problems arising from experience in the field, and the resulting research into new technologies and approaches based on practice and theory that aims to solve these emerging problems. With this emphasis, to maintain leadership in computing education and research, it is critical that the programs, facilities, faculty, and courses constantly respond and adapt to emerging trends and forces in computing, and the application of computing technology to important problems in the world. Over the course of two decades, CIT has created and built nine new programs, which include: rapid application development, local and wide area networking technology, systems and network administration, wireless technologies, computer forensics, and biomedical informatics and health IT. As part of this trend, faculty in the department identified a new emerging need for practitioners and researchers skilled in the development, deployment, and use of high performance computing systems (supercomputers) and cyberinfrastructure.

High performance computing and computational science are critical drivers of economic and research competitiveness in global science and engineering. The growth of open source software and the universal availability of low cost, high performance computer components make it possible to build powerful and inexpensive high performance computing systems. Supercomputers build from commodity computer components and open source software have become the dominant platform in use today for computing in support of science and engineering research. These systems, known as commodity cluster systems, have undergone tremendous growth in the last decade. Unfortunately, the technical workforce trained to operate these systems has not grown at the same pace and the lack of skilled practitioners limit the adoption of these systems. Moreover, enrollments in Computer Science programs have steadily declined by 50% since the beginning of the century [1]. Employment in computer and mathematical science occupations are projected to increase by 30.7% by 2014, and positions in network, systems, and database administration are also projected to increase by 38% [2]. Individual courses exist today in many institutions in isolation for topics such as parallel computing and parallel architecture. However, these efforts do not address the link between theory and practice "in the field" that serve the needs of the science, engineering, manufacturing, and life science communities – namely the design, development, deployment, and use of HPC resources to create a functional Cyberinfrastructure (CI) that provides a research and analysis "power tool".

These gaps represent a critical problem for several reasons: (a) there is a growing unmet need for skilled HPC practitioners; (b) there is a significant untapped potential to increase the competitiveness and quality of goods and services produced by U.S. industries; (c) the lack of skilled practitioners is slowing progress in science and engineering research; and (d) science and engineering research is becoming globalized through the use of cyberinfrastructure, and students trained to meet the needs of this community must have hands-on experience with this emerging technology.

Figure 1. The CIT High Performance Computing and Cyberinfrastructure Research Laboratory (HPC-CRL).

To address these challenges, CIT initiated the development of a new signature area in high performance computing and cyberinfrastructure. The mission of the HPC and CI program aims to: (1) establish an interdisciplinary curriculum that provides world-class training in HPC systems at both the architectural and utilization level, and (2) utilizing the computational and information capabilities of HPC to support discovery research and development in bioinformatics, computer graphics, information security, nanotechnology, product design and manufacturing, computer science, physics, and Science, Technology, Education, and Mathematics (STEM) educational outcomes.

To support the development of the HPC and CI signature area, a new project was established to develop a world-class facility (Figures 1 and 2) that would be flexible, adaptable, and responsive to the needs of faculty, students, and researchers. The goal of the facility was to provide an environment that would accommodate the needs for a collaborative space where students, faculty, and industry representatives could interact face-to-face and remotely [3]. In partnership with the CIT Industrial Advisory Board and representatives from IBM, the department initiated a development plan to design and build a world-class facility to support research and education in the new high performance computing and cyberinfrastructure signature area. Funding and equipment donations for the facility were provided by IBM Corporation (the most founding and most significant industrial contributor), Hewlett-Packard, Force-10 Networks, the Northwest Indiana Computational Grid project (funded by the Department of Energy), the College of Technology, ITaP (the Purdue central IT organization), and departmental funds. Construction started in March 2008, and was completed on schedule by the end of August 2008.

Figure 2. HPC-CRL Collaboration Area.

3.0 Description of the facility

The High Performance Computing and Cyberinfrastructure Research Laboratory is a comprehensive facility centrally located in the department, and is arranged into three major activity areas: research, collaboration, and sharing. The lab is the first Technology Exploration Center created by the department and college, and is a showcase for the kinds of new spaces and uses envisioned by the College. This lab is comprised of three major areas.

3.1 Data Center Research Laboratory

The first space, the Data Center Research Laboratory shown in Figure 3, is designed to house the most dense high performance computing equipment available today, and can support 25 KW of power and cooling per rack across six racks. The development of a skilled student cohort who can develop and deploy working systems requires hands-on experience with HPC systems, of which there are very few in operation in the field, and even fewer that are dedicated to research and education in HPC/CI as a field.

Figure 3. HPC-CRL Data Center Laboratory.

One of the major goals of the facility was to create a high density machine room with state-of-the art power and cooling systems, in which students would work on research and class projects.

There were many design and construction challenges faced in developing this space. One challenge was the very high power and cooling density required for systems to be housed in this space. Architectural design requirements for data centers has recently been updated, and modern data centers now routinely house systems requiring over 20 KW per rack.

Figure 4. Data Laboratory Power Distribution.

To adequately power and cool the center, a 4" chilled water line from the campus chilled water facility was put in place from the basement of the building to the data center. A portion of the capacity was utilized for a 20 ton Liebert air conditioning unit and for cooling a nearly Network Technology laboratory with a fancoil unit. This cooling infrastructure provided adequate ambient air cooling, but next generation HPC systems are emerging that are evolving to use rack and node level cooling systems. To support the future development of these types of

Figure 5. Data Laboratory Supplemental Cooling Manifold Ports.

cooling systems, a portion of the chilled water capacity was retained for future use. Additionally, an overhead distribution manifold (Figure 5) was installed to provide a future path for rack and node level cooling systems, and ports were installed and capped off for future use. In terms of power, 800 amps of power are installed in the facility for both powering systems installed in the racks and cooling infrastructure. The power distribution is controlled by panels installed in the wall, and is distributed underfloor through a systems of 208V 30 Amp twistlock plugs in a plugmold for each row of racks (Figure 4).

We decided to install 208V power rather than the standard 110V power for several reasons: high density systems such as blade centers require 208V power, and many modern computing systems contain integrated auto-switching power supplies designed to support power sources in countries outside North America. There is sufficient power for 6 30 Amp sockets per rack, or 180 Amps per rack.

One of the design goals for the lab is to support research in "green" computing technologies for data centers, which seek to minimize power and cooling load while maximizing the efficiency of computing technologies. To provide the basic infrastructure to support this work, the A/C unit and all of the power distribution units in the racks have network connections, which can support remote power monitoring and control. Faculty involved in the lab have submitted three grant proposals to fund research in this area to further instrument the facility with sensors and power control system.

In terms of computer hardware presently housed in the data center laboratory, there is a mix of systems consisting of high density blade and storage systems donated by IBM Corporation, and a blade system donated by Hewlett-Packard. There are currently 60 systems installed in the facility, which are a mix of 1U, 2U and blades that are tightly packed into the racks. The systems are from a variety of vendors: IBM, Sun, Dell, and H-P, and support research and education for bioinformatics classes, database and data intensive computing courses, and systems classes in the HPC/CI signature area.

One unique feature of the lab is the wide variety of interconnect technologies available and used in the facility. Force-10 networks donated two gigabit Ethernet switches, which are heavily used for both classes and research. In addition to gigabit Ethernet, other available interconnect technologies include Fibre Channel, InfiniBand (1X and 4X), Myrinet-2000, and Myrinet-10G. This variety of interconnection technologies allows students to gain exposure to a number of networking technologies, and provides an excellent platform with a variety of technologies to support benchmarking, research, and evaluation of many different combinations of architectures and configurations. We have a 10 Gb/sec fiber link to the campus backbone, and are working on establishing links to the TeraGrid, the Northwest Indiana Computational Grid, and BoilerGrid (the campus computing pool). This path also gives us the capability to work with research projects that generate and process enormous amounts of data, such as is the case for bioinformatics and particle physics.

What is especially unique about this laboratory is that it is designed to be OPEN for students, faculty, and industry partners for course work, research, and discovery. It is not intended to be a production facility, so students are welcome to enter the lab and engage in projects, which helps them gain experience in working with large-scale data center environments. One additional advantage of this approach is that when course modules are delivered on data center design, we can literally walk across the hall and provide the students with a hands-on demonstration of the technologies and the problems they solve. We are continuing to work on building up the capabilities in communications, data intensive computing, and benchmarking in this space.

3.2 Collaboration Area

The second area is the collaboration area, which is a flexible reconfigurable space designed to support the needs of students, researchers, vendor partners, and faculty. This area functions as an innovation space, equipped with technologies to support small group interactions and group projects. All of the equipment and furniture in this space are on wheels, to allow quick reconfiguration of the space for many different uses.

The way in which we have the space configured today supports five different activity areas: (a) meeting and collaboration; (b) white boarding; (c) student project space for two classes; (d) graduate student offices; and (e) seminars and lectures. The meeting and collaboration area (shown in Figure 7) is a small conference area with two tables and chairs, a whiteboard, and projection area. This space is useful for small group meetings, and for student project meetings. The white boarding space shown in Figure 6 is an area designed to facilitate the quick sketching up and discussion of ideas in a small group setting, with seating and tables focused on a whiteboard area. This area has proven to be popular with students, who sit in a group with their laptop computers to write lab reports and papers. A portion of the room is utilized as student project space. Students in the courses that require access to hardware are assigned resources in the data center laboratory, as well as a set of commodity desktop computer systems for their use to complete class projects and research projects. Each student station consists of a cart (on wheels), a small gigabit Ethernet switch, monitor, keyboard, mouse, and four Dell Optiplex 260 computers. These project stations are on wheels, and students can wheel them out during work on a project, and move them back into a corner when they are not being used. The beauty of these student stations is that they are very cost effective. The cart and wheels can be purchased at a home center, and the computers and monitors were repurposed desktop computing resources that were slated for retirement. Students working on projects using these carts can easily access the back and front, and combine them together to form larger clusters for high performance computing benchmarking projects.

Figure 6. Collaboration area whiteboarding space.

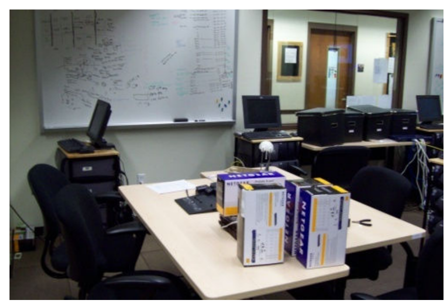

Figure 7. Collaboration area conference space.

We also have been able to work on establishing use of the collaboration area space for graduate student research assistant offices using mobile work carts. Each student is assigned a mobile cart (shown in Figure 8) for their use, and students can select a work area in the lab to set up their workspace. Students working on systems projects have found this arrangement to be useful when the need to link systems to the high density systems in the data center laboratory, since they can wheel their cart to the back of the systems for systems configuration and debugging. We have used the collaboration area occasionally for seminars and lectures by moving the equipment into the corner or a short-term storage space. Thus far, we have found this arrangement to be quite effective – it is meeting its goals of providing an efficient use of space while encouraging collaboration and joint research projects between faculty, researchers, and students.

Figure 8. Graduate student workstation.

3.7 Sharing and Education Area

The third area is the sharing space, which is a small space that can be used for meetings, courses, and seminars (Figure 9). This space can accommodate small classes, and when completed will include videoconferencing equipment as well as large screen displays on the walls, which will be visible from the hallways through a glass "storefront" partition wall. Currently, this space is used for small classes, graduate student meetings, and larger group projects. Future plans for this space include the installation of videoconferencing equipment, and the installation of an AccessGrid node.

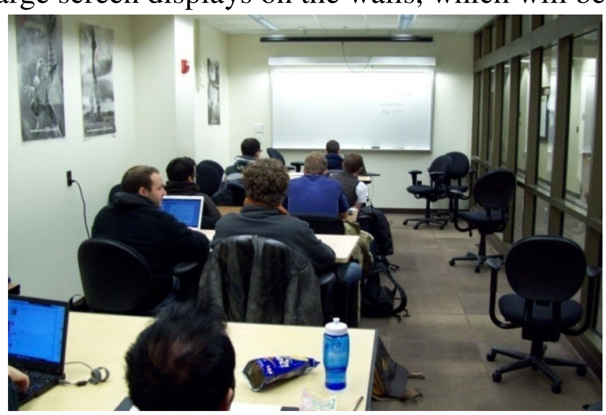

Figure 9. Sharing area.

The overall design of all these spaces is open, to facilitate maximum collaboration and sharing. As we use this space, we will continue to refine how we configure the space, adding additional technologies, such as video conferencing, to further encourage local and remote collaborative activities.

4.0 Use of the Facility and Outcomes

4.1 Education

Since the opening of the laboratory in August 2008, six courses have used the space, with approximately 50 undergraduate and graduate students enrolled in courses or involved with research projects. Currently, there are six faculty involved in the lab from the College of

Technology, with the research and education areas of interest of these faculty ranging from bioinformatics and genomics, to computer graphics, computer networking, and computer science.

The courses that have been developed in the HPC/CI signature area that are using the facility are: CIT 449M High Performance Computing Systems; CIT 581M Advanced High Performance Computing Systems; CIT 581A Advanced Parallel Data Systems; CIT 499A Parallel Data Systems; and CIT 581C Introduction to Computational Life Sciences. The curriculum path we are developing is shown in Figure 10. This is designed to build on the existing strengths and courses in the department. Thus far, we have developed new courses in 4 of the new areas, and are working on creating a new topical course on Grid Computing & Wide Area Networking for the Fall semester.

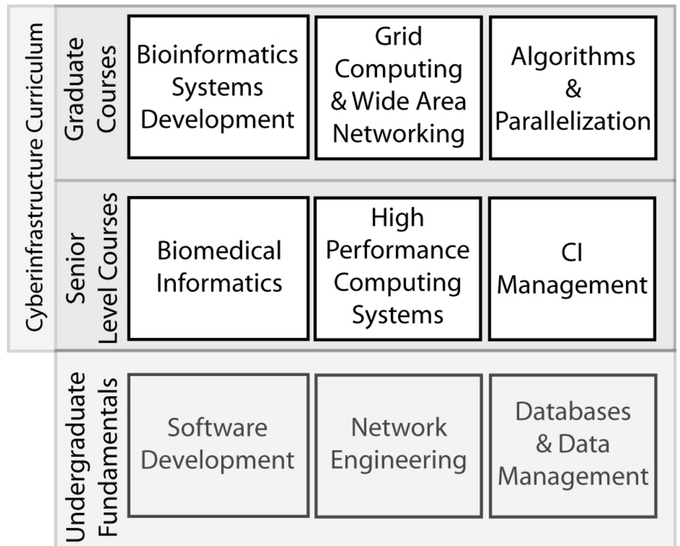

Figure 10. HPC and cyberinfrastructure curriculum.

The systems courses in particular, CIT 499M/CIT 581M High Performance Computing systems, heavily utilize the lab. Students are assigned a semester project to build a small cluster computing system using the four Dell Optiplex 260 systems assigned to them at the beginning of the semester. Through the course of the semester, students first design their cluster as their first project, planning the space, power, and cooling requirements that will be needed for their cluster, and provide a network map of how their systems will be configured. The second project is focused on using CentOS (an open source version of RedHat Linux), and the OSCAR [4] clustering toolkit to install and configure their nodes. The third project is to benchmark their cluster computer using High Performance LINPACK (HPL) [5], and to link their clusters together as a team project to improve performance. New projects planned for this semester include the use of cycle scavenging systems (Condor) on Purdue's campus [6] (BoilerGrid) for job submission and benchmarking.

The HPC-CRL lab has been open for a short time, and over the next year we will monitor student and faculty satisfaction with the facility. Feedback from the first semester of students using the lab was positive. Student response to the new facility has been very enthusiastic, with many positive comments on course evaluations. Students liked using the whiteboarding and collaboration areas, the nearness of the laboratory to the learning environment, and the freedom to utilize the space. Issues encountered during the first semester were: lack of power and networking infrastructure in the collaboration area, and the failure rate of hard drives used in the Dell systems for student projects. To address the power and networking issues, we added additional 10-outlet power bars to the lab and added a central Netgear switch for students. To address the reliability issues, we procured additional used Dell systems that students could use for spare parts.

4.2 Research

The facility has proven to be a useful platform for research, and for the integration of research and teaching. In terms of research, there are now eight sponsored research projects using the lab, with funding from the National Science Foundation, Microsoft Research, the U.S. Department of Labor, the U.S. Department of Justice, and Purdue University. The list of publications published by faculty involved with the lab has totaled 29 to date. Research projects involved with the laboratory are submitting research proposals in the areas of high performance computing, bioinformatics and life sciences, computer science, energy efficient data center design and operation, parallel computing using GPUs (graphics processing units), reliability of large-scale supercomputing systems, and computer forensics.

4.3. Bridging Research and Education

One significant goal of our effort is to act as a model for linking research and education in the training of next generation of scientists for conducting research. To this end, the work presented here attempts to bring in students very early on into the research process. This is a model used at Clemson University also wherein the problem of expanding research capacity not only in cyberinfrastructure, but also other related disciplines is considered to be of primary importance. In essence, our goal is not only to use the laboratory as an opportunity for collaboration among scientists, but also among scientists and students. This leads to the students being included as part of a "community of practice"[10]. As students become part of the community of cyberinfrastructure researchers, they learn the traditions, methodologies, and broad thinking within this field early in their careers. This leads to students who are better prepared to enter the workforce and who are intrinsically well-trained in thinking about problem solving within the field. Also, our goal is to affect the pipeline significantly enough to increase the number of students entering engineering, science, and technology programs at Purdue University and Clemson University significantly.

5.0 Related Work

Purdue is one of the first universities to establish an integrated research and education program in high performance computing and cyberinfrastructure. Several other institutions are now pursuing a similar goal. Clemson University has recently (January 2009) established a new Cyber Innovations Laboratory as part of their efforts in the Cyberinfrastructure Research Group [7], which includes space and facilities for graduate students and computational resources. The facility includes a Cyber Innovations Lab, a cubicle-based office with desks for graduate students, a large conference area and future visualization center, and a small low-density machine room space for systems. The purpose of the Clemson Cyber Innovations lab is to allow groups of students working in "project-based" environment to learn difficult concepts in cyberinfrastructure. Students are assigned a very specific set of research problems. They work with faculty members in a heavily scaffolded environment to understand fundamentals of the problem they are trying to solve. They then apply this fundamental knowledge towards developing innovative solutions to the research problems. Most of the students in the Clemson Cyber Innovations Lab are supported by research funds from the National Science Foundation. Areas of research range from virtual environment design to problems of engineering education epistemology and learning mechanisms[9] involving ultra-scale data. The Purdue HPC/CI lab is similar to the Clemson effort in that it is built up from activity areas, but differs in several ways. First, the Purdue collaboration area is built up from mobile furniture components, and is heavily multiuse. Second, the data center portion of the laboratory at Purdue can support much higher density systems, which is a "slice" of a high density supercomputing center. Third, the physical footprint of the Purdue center is smaller than Clemson.

Another effort is at Rochester Institute of Technology (RIT), located at the Center for Advancing the Study of Cyberinfrastructure (CASCI). CASCI is the focus center for cyberinfrastructure research at RIT. Within CASCI in the Service Oriented Cyberinfrastructure Laboratory (SOCIL), led by Dr. Gregor Laszewski [8]. The objective of the SOCIL lab is to establish an interdisciplinary research effort oriented towards service oriented computing approaches. Current activities in the lab include the development of advanced scientific workflow paradigms, service virtualization, and the development of cyberinfrastructure toolkits that make the user of advanced cyberinfrastructure easier.

6.0 Conclusions

At the outset, the motivation for developing the laboratory was to create a world-class facility to help the high performance computing and cyberinfrastructure area realize its strategic goals of establishing an interdisciplinary curriculum that provides world-class training in HPC systems at both the architectural and utilization level, and building on the computational and information capabilities of HPC to support discovery research and development in bioinformatics, computer graphics, science, physics, and STEM educational outcomes. Since the opening of the laboratory six months ago, there has been significant progress towards meeting these goals, which has been helped greatly by the resources and capabilities provided by the new lab. Altogether, the investment in developing the basic laboratory infrastructure has made a significant impact on attracting new students, research funding, and in the development of a new curriculum. As we continue our work in developing this signature area, we hope to further integrate the lab with complementary efforts and facilities at Purdue, and to work together with peer institutions seeking to develop similar facilities. The next few years are likely to see the emergence of a national curriculum for high performance computing and cyberinfrastructure, and the application of these technologies to solve practical problems. We are grateful to be able to be a part of this effort.

References

1. Zweben, S. (2007), *http://www.cra.org/statistics/survey/0506.pdf* 3ed, pp. 7-22.
2. Hecker, D. Occupational employment projections to 2014. *Monthly Labor Review Online*, **128**.
3. Bentley, L. (2007) The Future of IT Education is Here. *Purdue University Department of Computer and Infomation Technology Report*.
4. Sloan, J. (2004) *High Performance Linux Clusters with OSCAR, Rocks, OpenMosix, and MPI (Nutshell Handbooks)*. O'Reilly Media, Inc.
5. Feitelson, D.G. (2005) The supercomputer industry in light of the Top500 data. *Computing in Science & Engineering [see also IEEE Computational Science and Engineering]*, **7**, 42-47.
6. Smith, P., Hacker, T. and Song, C. (2008), *Second Workshop on Desktop Grids and Volunteer Computing*, Miami, FL.
7. Goasguen, S. (2009) Clemson Cyber Innovations Laboratory. *http://cirg.cs.clemson.edu/facilities*.
8. Laszewski, G.v. (2009) Rochester Institute of Technology Center for Advancing the Study of Cyberinfrastructure. *http://www.lac.rit.edu/*.
9. Journal of Engineering Education. (2006). The research agenda for the new discipline of engineering education. *Journal of Engineering Education, 95*(4) , 259-261.
10. Wenger, E. (1998). *Communities of Practice*. Cambridge University Press.