Research Experiences for Undergraduates: Autonomic Computing Research at FIU

Masoud Milani, S. Masoud Sadjadi, Raju Rangaswami, Peter J. Clarke, and Tao Li
School of Computing and Information Sciences
Florida International University
11200 SW 8th St., Miami, FL 33199
305-348-2744
{milani,sadjadi,raju,clarkep,taoli}@cis.fiu.edu

ABSTRACT
According to Computing Research Association, during each year between 2003 and 2007, fewer than 3% of the US's Ph.D.s graduates in computer science and computer engineering were Hispanic or African American and fewer than 20% were women. Such an under-representation precludes the benefits of diversity in computer sciences research and industry and consequently compromises the competitiveness of the US economy. It is therefore imperative that undergraduate institutions introduce students from these groups to research at an early stage of their academic careers and to provide them with the tools necessary for success in graduate school. The School of Computing and Information Sciences (SCIS) at Florida International University (FIU) has been working to strengthen the pipeline of underrepresented students to graduate work in computer science by hosting an NSF sponsored Research Experiences for Undergraduates (REU) site for the past three years. Our REU site has hosted 30 undergraduate students, 23 of them were underrepresented including 8 females, 16 Hispanics, and 4 African Americans, who published 13 technical papers. Six of the ten students who have already graduated, have started their graduate studies.

Categories and Subject Descriptors
C.4 [**Performance of Systems**]: *Fault tolerance, Measurement techniques, Modeling techniques, and Performance attributes*; D.2 [**Software Engineering**]: Testing and Debugging— *Testing tools*; Software Architectures—*Information hiding, Languages.*

General Terms
Management, Measurement, Performance, Design, Security, Experimentation, Languages, and Verification.

Keywords
Research experiences for undergraduates, REU site, autonomic computing, underrepresented students.

Tapia'09, April 1–4, 2009, Portland, Oregon, USA.

1. INTRODUCTION
The lack of early exposure to research (starting at the undergraduate level) contributes, in a significant way, to the lack of understanding among undergraduate students about graduate school and of research careers in academia and industry. This problem is particularly acute among minority students since many of them are first generation college students, few faculty members in Computer Science are from minority groups [1-4], and very few Minority-Serving Institutions (MSIs) are strong research institutions. Minority groups pursue graduate degrees in Computer Science at rates far below their proportionate representation in the United States' population. Hispanics and African Americans respectively represent 12.1% and 12.5% of the US population [5]. The Computing Research Association has reported (via its recent annual Taulbee Survey Report [6], and prior reports) that between 2003 and 2007 an average of 15.75 Doctoral degrees in Computer Science per year were awarded to Hispanics in North America. Fewer than 3% of the U.S.'s approximately 1,500 Ph.D. graduates in computer science and computer engineering each year are Hispanic or African American; fewer than 20% are women. Such an under-representation not only compromises the competitiveness of the US economy, but also deepens the divide and imbalances in our society.

During the past three years, with support from the National Science Foundation's Research Experiences for Undergraduates (REU) program, we established a successful REU site at the School of Computing and Information Sciences (SICS) at Florida International University (FIU) and built upon our past experiences and infrastructure at FIU to provide world-class research opportunities to undergraduate students primarily recruited from Minority Serving Institutions.

The overriding objective of the site has been to positively impact the participating students' beliefs and perceptions of research and academic careers and to strengthen the pipeline of underrepresented students to graduate programs in computer science. Our program exposes the participants to research, life in graduate school, and work in large-scale research projects early in their careers [7] and thus facilitates their transition into graduate school and advanced research careers in computer science. Our REU site has successfully achieved this goal. Six of ten students who have already graduated have started their graduate studies

(60%) in Computer Science: 2 at the Georgia Institute of Technology, 1 at Purdue, 1 at the University of Minnesota, and 2 at Florida International University. From the four students who have joined the workforce, one has accepted a position with Fluke Networks and will begin his graduate work in University College of Colorado Springs while two others have accepted positions with IBM and Citrix Systems and are planning to apply for admission to graduate studies within the next year, upon which our success rate will be 90%.

During each summer, 10 participating students engaged in a 12-week research-intensive program designed to help the students evolve from being highly motivated but dependent learners to independent researchers experienced in designing, implementing, evaluating, and reporting solutions to research problems. The students met with their faculty mentors on a regular basis to report their progress, discuss problems they encounter, and set future goals. In addition, our graduate students acted as peer mentors, discussing the students' research and their progress on a daily basis and providing informal guidance and mentoring while serving as role models [8].

The students participating in this REU program joined our existing research groups in the area of Autonomic Computing [9] and participated in the group's research activities. As computing systems become increasingly complex with a growing number of heterogeneous hardware and software components, they become more and more difficult to monitor, manage, and maintain. Autonomic computing promises to solve the management problem of today's complex computing systems by embedding the management of such systems inside the systems themselves, freeing users from potentially overwhelming details. The REU research program was centered on several key issues in autonomic computing such as self-management in grid computing, autonomic anomaly and misuse detection and response, adaptive network traffic monitoring and management, autonomic storage systems, and testing of autonomic computing systems, areas in which the REU site faculty members had substantial expertise and ongoing research projects.

2. STUDENT RECRUITMENT ACTIVITIES

The recruitment efforts were proactive in nature. The site faculty members gave recruitment seminars at targeted Minority Serving Institutions. These seminars described the overall research theme of the site and the benefits that the research experience will bring to the participating students. To ensure the success of our recruitment effort, the University of Puerto Rico, Mayagüez (UPRM) and Florida A & M University (all Minority Serving Institutions) each supported our activities by identifying a faculty member who coordinated student recruitment at their institution. Additionally, the program was also promoted through our web site (www.cis.fiu.edu/reu) and via direct mail to faculty advisors and student clubs at institutions throughout the US.

3. OVERVIEW OF STUDENT ACTIVITIES

The Site's faculty mentors introduced REU participants to the fundamentals of research including literature search, technical writing, presentation skills, time-management, and various tools and techniques to support research activities. The faculty mentors conducted seminars describing the significance of the overall research work in autonomic computing and the relevance of each research project. As the students understood more about the research projects and became familiar with the process of conducting research, they were encouraged to work independently. This process was closely monitored by the mentors to make sure that the participants were making progress towards the completion of their projects.

The faculty mentors assigned their graduate students to serve as peer role models and mentors to the REU participants. They provided daily contact, helped with the problems that the REU students encountered, and offered information and advice on graduate school. As all student research projects were within the scope of autonomic computing, we facilitated the exchange of ideas and knowledge among the participants, promoted team work, and allowed the students to effectively learn from each other and to gain the confidence necessary for the successful completion of their projects.

4. STUDENT RESEARCH ACTIVITIES

After a detailed introduction to the faculty, graduate student mentors and to the proposed research projects, the participants were introduced to the fundamentals of research including: literature searches, technical writing, presentation skills, research methodologies, and various tools and techniques to support research activities, the expectations in terms of work ethic, team-work, and the requirement for the eventual independence in conducting research activities. Additionally, the faculty mentors motivated the need for each project, explained the relations among the projects, and presented the big picture of the research theme. Once project groups were formed, the assigned faculty and graduate student members provided more details about each individual project separately for each group of students. Although the scope and goals of each project were defined by the faculty mentors, students were encouraged to brainstorm and discuss how each research project should be conducted. This way, students participated in constructing a concrete plan customized for each specific research project. After this introductory period, the students were immersed in research-focused project activities and conducted research in a group setting and incrementally learned how to assume more independent roles. The team members interacted with graduate assistant(s) assigned to their group on a daily basis and formally met with their faculty mentor on a semi-weekly basis to discuss progress and set milestones. A full day was devoted to mid-project review where the participants each gave a 15-minute presentation on their progress and obtained valuable feedback from the faculty mentors and their peers. The mid-project review proved to be essential in bringing all of the students up to date on the progress of individual research projects as well as the entire research program. In particular, the mid-project review helped to 1) re-paint the big picture for everyone, 2) re-enforce project roles and how each individual project contributes to the overall theme of Autonomic Computing, 3) foster collaboration and learning between projects and project teams, 4) re-evaluate the project goals and potentially re-think priorities, and 5) investigate potential student publications and publication venues. During the final project review at the end of the summer session, each REU participant presented his/her work

in detail. The presentations outlined the initial goals, project schedule, challenges encountered, lessons learned, and the current project status. Students proposed follow-through activities for their research projects that were conducted upon return to their home campuses and plans for these activities were laid out. Starting from the follow-through activities proposed by the students, faculty mentors and students designed concrete plans to continue working on publications of the results of the research projects.

In summer of 2006, two REU students worked on designing and implementing Rootsense, a low-overhead Operating System-based intrusion detection system. While one student conceived, designed and implemented a high-performance kernel module for intercepting and correlating Linux kernel system calls, the other student designed and implemented a state-machine-based analysis engine for identifying both malicious and anomalous sequence of system calls [10]. Two other REU students worked on intelligent data layout on the disk drive designed as a core component of the BORG, an autonomic, block-reorganizing storage system. The contribution by one of them included the modeling of I/O access patterns by independent processes as directed weighted graphs. This model was used by the other student to design an algorithm to merge individual process graphs into a cumulative graph and intelligently laying out a directed and weighted graph onto the disk drive. Results from a prototype system indicate the potential for large improvements in disk I/O performance with BORG [11]. One REU student improved the self-managing mechanism of a Network Communication Broker. He developed a software module to enable dynamic configurations of XML-based policies that guide the self-managing behaviors according to application/user preferences. Two REU students performed a literature review on testing autonomic computing systems by surveying several research projects including Impala (Princeton University), OceanStore (UC Berkeley), Model Driven Autonomic Manager (Indiana University), and The Bison Project (University of Bologna). This review led to the development of a new testing framework for autonomic computing systems. The framework was partially validated by testing an autonomic container at runtime [12, 13]. Two other REU students improved the overall performance of the TRAP/J toolkit significantly, overcoming a major inhibitor to using this toolkit for Grid enablement. The improvement was focused on TRAP/J's critical execution path, namely, the section of the code that repeatedly gets executed as part of the program's functionality. They found and eliminated a redundant search loop for a matching delegate method, which was causing unacceptable performance overhead [14, 15]. Finally, the last REU student developed a Remote Composer for TRAP.NET that gives a user the ability to dynamically change the behavior of a .NET application at run time [16].

In summer of 2007, two REU students investigated a component-based mechanism for safe dynamic adaptation to support the validation of autonomic computing systems. A survey of existing static analysis tools was completed and an object-oriented (OO) design for the data structure to store the component dependencies, and build a safe adaptation graph to compute the sequence of adaptation steps was designed [17, 18, 19]. Two other REU students investigated methods to extend the functionality of a Network Communication Broker (NCB), a layer in the Communication Virtual Machine [20] that interfaces with the underlying networks, to utilize multiple communication frameworks by dynamically configuring these frameworks in NCB, and developed static and dynamic models for the self configuring capability of the NCB [21]. One student worked on UPGRAID, an operating system optimization that dynamically replicates popular data to improve both performance and availability in RAID-based storage systems. He participated in the conceptualization, architecture, and design of UPGRAID and developed the initial prototype of the system as a Linux kernel module. Another student worked on storage block-layer 'integrity checker' and 'workload generator' test tools that can be used to test new block-layer extensions in the operating system storage stack. The remaining four REU students joined the Transparent Grid Enablement of Weather Research and Forecasting (WRF) project aiming at adding autonomic computing capabilities to a weather forecasting scientific application, called WRF (over 160,000 lines of code developed in FORTRAN and C), so that it can take advantage of pool of heterogeneous and dynamic resources provided by different interested organizations. The students installed WRF in our cluster environment and produced a step-by-step guide of a model run in WRF, conducted a series of benchmarking experiments to evaluate the behavior of WRF in a cluster computing environment and developed a mathematical model to predict the execution time of WRF on cluster settings [22].

5. RESULTS

During the past three years, we received 25, 51 and 54 applications for years 1, 2 and 3 of this project respectively. Each year, 10 students (six of whom were from outside FIU) attended the summer program. The demographic make-up of the participants is shown in Table 1.

Table 1. 2006-2008 student body profile

	2006	2007	2008	2006-2008	
Applications	21	51	54	126	
Participants	10	10	10	30	
Acceptance ratio	47.62%	19.61%	18.52%	23.81%	
Underrepresented participants	8	7	8	23	76.67%
Females	2	2	4	8	26.67%
Hispanics*	5	6	5	16	53.33%
African Americans**	2	0	2	4	13.33%

* includes female hispanics
** includes 1 female African American

Table 2 shows the current status of our 2006 and 2007 participants. Fifty percent of the students who participated in our summer 2006 and 2007 programs are currently in school completing their undergraduate work. The remaining ten have graduated. Most students in both groups are continuing their mentoring relationship with their faculty mentors. In addition, over 50% of the 2006 and 2007 participants have co-authored at least one article in a variety of journals, conferences and workshops (10 publications) and 40% of the students have either presented at a conference. Additional publications are under preparation as part of the follow through activities. The data clearly shows that our REU site has achieved its goals. 100% of our students have experienced the daily life in graduate school.

Six of ten students who have already graduated, have started their graduate studies (60%): 2 at Georgia Institute of Technology, 1 at Purdue University, 1 at the University of Minnesota, and 2 at Florida International University. From the four students who have joined the workforce, one has joined Fluke Networks and will begin his graduate work in University College of Colorado Springs starting Spring 2009, and two have accepted positions with IBM and Citrix systems and are planning to apply for admission to graduate studies within the next year.

Table 2. 2006-2007 REU results

	2006	2007
Participants	10	10
Completing BS degree*	5	5
Graduated	5	5
In graduate school	2	4
Joined the work force	3	1
Students partcipated at conferences	5	3

6. LESSONS LEARNED

Recruitment: Our recruitment effort has been extremely successful. The recruitment presentations, sending posters to academic institutions, and working with faculty in our target institutions to get direct referrals of qualified students have all proved to be practical and successful. We also learned during our recruitment process that it is important to contact selected applicants on an ongoing basis and be persistent in our recruitment until each applicant makes a decision. In some cases we called applicants 3 or 4 times in order to convince them that the REU program would benefit them in their future career plans. One of our success stories included the recruitment of a student while he was serving in the armed forces in Iraq. While serving his country oversees he sent in his application and based on his qualifications, and our initial email exchanges with him, we decided to accept him into the program for summer 2008. In this case we were unable to contact him by telephone for the initial interview.

Student Research activities: Students generally reported that they were satisfied with their planned activities and that the progress they made in the completion of their research projects and that the organization of our activities was an important factor comtributing to their success.

Mentoring: The students reported that they were satisfied with their research environment but wanted to have a closer mentoring relationships with graduate students. We addressed this issue by assigning students to projects that were led by graduate students while maintaining their close mentoring relationships with the faculty. Preliminary discussions with students during Summer 2008 indicate that this model is working very well and students feel that the graduate students are available to answer their questions throughout the day and that this additional help is making them more productive.

Program length: The twelve-week duration of our current summer program requires the program session to be from mid May until early August (May 19 – Aug 8, for Summer 2008). Such a long period results in scheduling conflicts for students whose Spring semester finishes late and to those whose Fall semester begins early (particularly the students from Puerto Rico). Additionally, many students felt that being involved in research for the entire summer does not give them sufficient time to take summer vacation. While the twelve week period is optimum in that the students will have ample time to start and complete their projects, we learned that a ten-week long REU program is a compromise that can resolve the other issues while maintaining a reasonable research period during which students can complete their work.

Follow-through activities: Our experience from the past two years indicates that most of the publications co-authored by our REU students were submitted for review sometime after the REU program ended. We learned that if we do not lay out a concrete plan for finishing the work for publication, it is very difficult to accomplish this task.

7. CONCLUSION

To positively impact the participating students' perceptions of research and academic careers and to consequently increase the number of students entering computer science graduate programs and research and academic careers, we introduced the students to the challenges and joys of research in computing via an NSF-sponsored REU program for the past three years. The 12 week summer research program transformed the participants from dependent learners, who learn mainly through the direct controlling influence of their instructors, into more mature and semi-independent researchers, who are not only able to learn and create through self-exploration, but also effectively function within a large and internationally spanned research group with a well-defined objective. To achieve this goal, the students were mentored and engaged in carefully designed activities that gave them a wide range of necessary skills including research methods, appropriate domain knowledge, communication skills, team-work skills, social skills, and experience working in large-scale international collaborative research projects.

Thirty REU students have completed their 12-week summer REU program at FIU. The REU students have each learned from their team-members and extended their domain knowledge rapidly while developing their communication and social skills. Our graduate students served as peer mentors and role models, and guided the participants throughout their summer research experience. This non-hierarchical structure provided the undergraduate students with technical expertise, as well as the team work, communication, and social skills. These skills helped lay the foundation for the successful completion of their summer projects and success in their future graduate studies and careers in the competitive market.

The Site's recruiting and outreach efforts focused on students from minority serving universities, leading to broader diversity among computer science researchers. By involving a group of traditionally underrepresented students in REU activities at FIU, the site has increased the diversity of students prepared to pursue advanced degrees and academic careers in computer science. 60% of the participants who have completed their undergraduate studies (6 out of 10) have entered Computer Science graduate programs. These and our future graduate students will later become role models for those who follow and therefore play a major role in broadening the participation of minorities in Computer Science.

8. ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation (NSF) grants IIS-0552555, IIS-0534530, CNS-0747038, and Department of Energy grant DE-FG02-06ER25739. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect those of the above agencies. The authors thank Dr. Chi Zhang for his contributions to the first year of this project as a faculty mentor.

9. REFERENCES

[1] Broadening Participation in Science and Engineering Research and Education: Workshop Proceedings. National Science Board, National Science Foundation. August 2003. 145 pp. http://www.nsf.gov/pubs/2004/nsb0441/nsb0441.pdf

[2] National Science Board. Broadening Participation in Computing Research and Education. October 20-21, 2004. http://www.cra.org/Activities/workshops/broadening.participation/nsf/NSB/NSBproceedings.doc

[3] Broadening Participation Through a Comprehensive, Integrated System – Final Workshop Report. National Science Foundation. January, 2005.

[4] P. Gándara and J. Maxwell-Jolly. Priming the Pump – Strategies for Increasing the Achievement of Underrepresented Minority Undergraduates. The College Board. December 1999.

[5] US Census Bureau. US Census 2000. http://www.census.gov/main/www/cen2000.html

[6] Computing Research Association. "Taulbee Survey Report, 2006-2007." March 4, 2008.

[7] R.S. Hathaway, B.A. Nagda, and S.R. Gregerman. The Relationship of Undergraduate Research Participation to Graduate and Professional Education Pursuit: An Empirical Study. Journal of College Student Development, 43, 614-631, 2002.

[8] Cesar D. Guerrero, Miguel A. Labrador, and Rafael A. Pérez. Graduate Students Mentoring in REU Sites. ASEE SE Section Annual Conference. April 1-3, 2007. Louisville, KY

[9] J. O. Kephart and D. M. Chess. The Vision of Autonomic Computing, IEEE Computer, 36(1), pp. 41–50, 2003.

[10] R. Koller, R. Rangaswami, J. Marrero, I. Hernandez, G. Smith, M. Barsilai, S. Necula, S. M. Sadjadi, T. Li, and K. Merrill. Anatomy of a real-time intrusion prevention system. In *Proceedings of the IEEE International Conference on Autonomic Computing*, 2008.

[11] M. Bhadkamkar, J. Guerra, L. Useche, S. Burnett, J. Liptak, R. Rangaswami, and V. Hristidis. BORG: Block-reORGanization for self-optimizing storage systems. To appear in the *Proceedings of the USENIX Conference on File and Storage Technologies, February, 2009.*

[12] Ronald Stevens, Brittany Parsons, and Tariq M. King, "A Self-Testing Autonomic Container", Proceedings of the 45th ACM Southeast Conference, ACMSE '07, p. 1-6, 2007.

[13] Tariq King, Djuradj Babich, Jonatan Alava, Ronald Stevens and Peter J. Clarke, "Towards Self-Testing in Autonomic Computing Systems", International Symposium on Autonomic and Decentralized Systems 2007, p. 51-58 2007.

[14] S. M. Sadjadi, J. Martinez, T. Soldo, L. Atencio, R. M. Badia, and J. Ejarque. Improving separation of concerns in the development of scientific applications. In *Proceedings of the 19th International Conference on Software Engineering and Knowledge Engineering*, pages 456-461, 2007.

[15] S. M. Sadjadi, L. Atencio, and T. Soldo. TRAP/J v2.1: An improvement for transparent adaptation. Technical Report FIU-SCIS-2007-09-01, School of Computing and Information Sciences, Florida International University, 11200 SW 8th St., Miami, FL 33199, May 2007.

[16] J. Ocasio Pérez, P. I. Rivera-Vega, S. M. Sadjadi, and F. Trigoso. Dynamic adaptation of a math service using trap.net. Technical Report FIU-SCIS-2008-07-01, School of Computing and Information Sciences, Florida International University, 11200 SW 8th St., Miami, FL 33199, July 2008.

[17] Tariq M. King, Alain E. Ramirez, Rodolfo Cruz, Peter J. Clarke. An Integrated Self-Testing Framework for Autonomic Computing Systems. Journal of Computers, Academy Publisher. November 2007, 2(9) pages 37-49.

[18] Tariq M. King, Alain E. Ramírez, Barbara Morales and Peter J. Clarke. A Reusable Object-Oriented Design to Support Self-Testable Autonomic Software. The 23rd ACM Symposium on Applied Computing, Track on Autonomic Computing, ACM. March 16-20, 2008, Pages 1664-1669.

[19] Alain E. Ramirez, Barbara Morales and Tariq M. King. A Self-Testing Autonomic Job Scheduler, Proceedings of the 46th ACM Southeast Conference, ACMSE '08.

[20] Y. Deng, S. M. Sadjadi, P. J. Clarke, V. Hristidis, R. Rangaswami, and Y. Wang. CVM - a communication virtual machine. *The Elsevier Special Issue of the Journal of Systems and Software (JSS)*, 81:1640-1662, 2008.

[21] Andrew Allen, Sean Leslie, Ricardo Tirado, Yali Wu and Peter J. Clarke. Self-Configuring User-Centric Communication Services. The 3rd International Conference on Systems. pages 253-259, 2008.

[22] S. M. Sadjadi, S. Shimizu, J. Figueroa, R. Rangaswami, J. Delgado, H. Duran, and X. Collazo. A modeling approach for estimating execution time of long-running scientific applications. In *Proceedings of the Fifth High-Performance Grid Computing Workshop,* 2008.