The 2010 Open Science Grid Summer School – Final Report

About the School

The 2010 Open Science Grid (OSG) Summer School was held from July 19-22, 2010 (https://twiki.grid.iu.edu/bin/viewauth/Education/OSGSummerSchool2010). Seventeen participants arrived in Madison, Wisconsin, on the preceding Sunday, and attended the four-day school, which was held in the Computer Sciences Department at the University of Wisconsin–Madison. After settling into their hotel, the participants joined some of the instructors for the introductory group dinner. The next four days consisted of all-day hands-on sessions in which the participants heard from various experts in the field of distributed computing. The participants gained valuable experience with high throughput computing, the Condor Project, Globus, security in distributed computing, and TeraGrid in addition to OSG. The final day of the school consisted of discussions of collaboration between computer science and other science fields that can use distributed computing in their research efforts.

"The OSG Summer School was a great experience," said Vishagan Ratnaswamy, who will begin work on his doctorate in aeronautics this fall at the California Institute of Technology. "I was able to learn more about the script files I was using as well as the systems I was running my simulations on." Before attending school, Ratnaswamy was already an OSG user. But by attending, he learned more about the available computational resources and the many ways in which grid computing can be applied to various types of research.

Making the 2010 OSG Summer School a Success

The seventeen participants were chosen from a pool of forty-five applicants and received funding to attend the school and the TeraGrid '10 conference. Candidates were chosen based on their backgrounds and the ways in which OSG would be useful in their daily research activities. At the school, participants had the opportunity to learn the basics of grid computing – knowledge they could apply to working within a variety of distributed computing environments. "Distributed computing is fundamentally complex because you've got resources you don't control; it's a new paradigm of working and so people have to be got over the threshold to use it. Having an intense collaborative session of a few days where you're resident really helps to get over that threshold," commented Ruth Pordes on the advantages of holding a school in which the participants traveled from a variety of locations. Organizers decided to make residency a

requirement in order for the participants to fully immerse themselves in the distributed-computing experience.

The participants were able to meet OSG staff members and ask questions and experiment with their own accounts at the school, giving the participants true hands-on experience while listening to lectures and engaging in practice sessions. According to anonymous surveys answered by the participants, each of the sections were useful for learning more about OSG. The anonymous surveys were overwhelmingly positive with regard to the practice sessions, calling them "the most beneficial" part of the school.

The organizers' efforts focused primarily on the development of the school's curriculum and content. "We concentrated on providing lots of high-quality, hands-on learning activities. The ultimate goal was to give students the ability to apply the lessons directly to their own research projects. And already, less than a month since the school, we are hearing back from students who are doing just that," said Tim Cartwright, who led the team that organized the school, in August. The school provided computational and storage resources in two locations within OSG (University of Wisconsin–Madison and University of Nebraska–Lincoln) and this not only provided significantly more resources for students to run large jobs, but it gave them a real feeling for what it is like to run in a real grid.

The model for the school was aimed toward finding an audience not necessarily solely within computer sciences. "In the new summer school, we focused on the use of distributed resources and then bolstered that by foundational methodologies, whereas in grid schools we talked about the 'nitty gritty' of the underlying software," Ruth Pordes claimed. In the past, school "faculty" and about ten students would receive sponsorship to attend an international summer school in Europe. In 2010, however, the International Summer School in Grid Computing did not take place, making the school even more important for continuing users' education. The team at OSG had been working on the idea of a summer school in the United States for several years, but this year, the faculty were able to channel efforts normally reserved for the international summer school toward making the OSG Summer School a reality.

Bringing the Participants into the Computing Community

The school was co-located with the annual OSG staff retreat, making it possible for the OSG staff to become more involved and interact with the students in both formal and informal

settings. "Having the OSG staff teach the school gave it a credibility and vitality that could not be matched otherwise, and it seems clear to me that the students felt this, too," Tim Cartwright said.

The school has been enormously successful in fostering interest in distributed computing, as well as teaching users how to be more efficient in their use of distributed computing. While attempts to reach out to the scientific community can be difficult, the school brought in different members of the scientific community lecture about their use of OSG. Faculty of the University of Wisconsin–Madison from the Botany Department, the Laboratory for Molecular and Computational Genomics and the IceCube Project showed the participants how OSG has been beneficial to their research. OSG hopes to continue to foster this interdisciplinary fellowship at succeeding schools.

The participants were each paired with a mentor who has served as a resource over the next year, in order to maintain the knowledge obtained at the school. The mentorship portion of the school was modeled after the Department of Energy's computational science graduate fellowship program. As part of the Department of Energy's fellowship, each fellow and his/her supervisor attend an intensive workshop, creating a more cohesive learning experience. A follow-up conference call consisting of the participants and their mentors was held in November to get final feedback and help the ongoing OSG Summer School program. Forty-two percent of all participants and mentors were able to attend, and they offered helpful tips in planning next year's school.

Measuring the Success of the Summer School

In order to adequately evaluate the success of the Summer School, the organizers asked the participants to fill out anonymous evaluation forms for each of the sessions and then a final, all-encompassing evaluation form at the end of the week. The organizers requested opinions on the quality of the session, whether the content was understandable, if the presentation was clear and useful, the quality of the hands-on exercises, time management, and the instructors' ability to teach and lead the session. The participants were asked to fill out the forms as soon as the session was over so that they would not forget any key information they wanted to include. The final evaluation form was handed out the last day of the school, and the questions focused more on the quality of the logistics than the individual sessions. The organizers requested feedback on

logistics like the food selection, the quality of the accommodations, the quality of the conference room, the quality of network access and the ease of planning. See the Appendix for a sample of the two evaluation forms which were used during the school.

The results of the evaluation forms were overwhelmingly positive. The participants judged the school to be an overall success in both the sessions and the logistics. A professional analysis of the evaluations ruled that, "overall, the majority of the participants were very happy with how the conference went, and raved about the accommodations and organization of it. The majority were also very happy with the instructors and the presentations and activities within the sessions."

Each evaluation was examined for insight into what the participants thought of the organization and the sessions. The following are some of the quotes taken from the evaluation forms:

- The example tutorials are excellent. I'm learning the most by far by these, which just indicates how much time went into their prep.
- The presentations did a good job summarizing information and integrating the exercise.
- The instructors provide useful hints and techniques for solving 'real' problems. It's also interesting to see the advantage of doing exercises on live during lectures – common problems can be identified quickly and solved together.
- The examples were easy to follow, the information was given in digestible chunks.
- I like the way the instructor giving sample examples to explain a bit complicated concepts.

Final Reports

Each participant was asked to complete a final assignment, comprised of a 500-1000 word paper, after the completion of the school. This final assignment was given to the participants to gain insight into what they learned while attending the school and learning to what sort of projects the participants felt they could apply their new knowledge. The participants were asked to pick a project, ideally one they would be working on in the near future. The project needed to have large-scale computing requirements and should have come from their domain area. Each paper included a problem to address, an estimate of necessary resources and the basic approach the participant would take in addressing these issues.

In writing the paper, the participants were asked to consider several questions: Would they use local resources at their research institution, or would they need to gain access to remote resources as well? How would they turn their projects into actual jobs? What are the resource needs of the jobs themselves? What sort of workflow would be utilized? How much data would they need to move around and how would they deal with it? Is their project better suited for HTC or HPC, and why? What security concerns need to be addressed and how would they do so? The titles of the participants' completed papers are listed below:

Beissinger, Timothy	*Running Haplotype Identification Software in a Grid Environment*
Cao, Shanshan	*OSG Final Assignment: HTC for Relativistic Heavy-Ion Collisions*
Deaton, Brett	*A High-Throughput Computing Solution to Calculating Uncertainties Associated with Compact Binary Coalescence Searches in Gravitational Wave Observations*
Fisher, Ryan	*OSG Final Assignment: Ground-Based Gravitational Wave Detection on the Grid*
He, Chen	*A Condor-Based Open Environment Hadoop MapReduce Platform*
Jiang, Wenying	*Machine Design Optimization Based on Finite Element Analysis in a High Throughput Computing Environment*
Joyce, James	*Cluster Computing – a Heuristic Approach*
Lee, Kyungyong	*Creating Condor Pool-On-Demand by Using a Loosely-Coupled Peer-To-Peer Network*
O'Donnell, Michael	*Understanding the Implications of Changing Imagery Map Projections: A High Throughput and High Performance Computing Application*
Ratnaswamy, Vishagan	*Simulating Granular Dynamics Using High Throughput Computing*
Snihur, Rob	*Feasibility Study of PROOF Cluster at a CMS Tier 3 Site*
Somavarapu, Dilip	*Work Flow Management for UCoMS*
Wang, Zhiyong	*A New Probabilistic Model for RNA 3D Structure Prediction Using High-Throughput Computing*
Younts, Alex	*Supporting Bioinformatics Applications Using High Throughput Computing*
Zhou, Yuhong	*Optimizing the Placement of Defibrillators in 3D Indoor Environment on Grid*
Zuo, Luo	*Force Chain Capturing and Analyzing in Cylindrical System Using High-Thoughput Computing*

The organizers of the Summer School read and reviewed the compiled reports, looking to see if the reports met the final assignment guidelines. All the submitted reports met the organizers' standards, and they detailed projects which would use the skills learned at the OSG Summer School. The participants were then invited to submit posters with the status of the projects, which came out of their final reports, to the OSG All-Hands Meeting in March, 2011.

Final Thoughts

The organizers would like to thank NSF for their support in this venture. The 2010 OSG Summer School was very successful, and OSG would like to build upon its success when creating a 2011 OSG Summer School.

Appendix

Evaluation Form Template

Time: Date: Instructor:
M T W Th Topics:

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
1. I would recommend this section to others.	1	2	3	4	5
2. The content was organized and understandable.	1	2	3	4	5
3. The presentations and materials were clear and useful.	1	2	3	4	5
4. The hands-on exercises were interesting and useful.	1	2	3	4	5
5. I will be able to apply the knowledge after the school	1	2	3	4	5
6. There was adequate time for all the activities	1	2	3	4	5
7. The instructor(s) knew the content well.	1	2	3	4	5
8. The instructor(s) did a good job teaching.	1	2	3	4	5
9. The instructor(s) promoted interaction among all students.	1	2	3	4	5
10. The instructor(s) encouraged questions and gave helpful answers	1	2	3	4	5

Additional Questions:

What were some of the best parts of this section?

How would you suggest that we improve this section?

Other comments:

Final Evaluation Form Template

Logistics and Final Thoughts

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
1. The food was excellent.	1	2	3	4	5
2. The conference room was comfortable and worked well for the conference.	1	2	3	4	5
3. Logging into the computers and the wireless was easy.	1	2	3	4	5
4. There were no network problems.	1	2	3	4	5
5. The accommodations were pleasant.	1	2	3	4	5
6. Planning (with us and SDSC) was easy and painless.	1	2	3	4	5
7. The length of time for the entire conference (4 days) was good not too short or too long.	1	2	3	4	5
8. The group dinners were enjoyable.	1	2	3	4	5
9. Overall, the conference went well.	1	2	3	4	5

Additional Questions:

Logistically, what did you enjoy the most this week?

How would you suggest we improve the food, conference rooms, planning, etc?

Overall did you feel there was enough time for all the activities and presentations? Do you recommend any changes?

What stands out the most to you this week (for everything).

Other Comments: