

Methods of Creating Student Cluster Competition Teams

Stephen L. Harrell
Rosen Center for Advanced
Computing
Purdue University
155 South Grant St.
West Lafayette, Indiana 47906
sharrell@purdue.edu

Preston M. Smith
Rosen Center for Advanced
Computing
Purdue University
155 South Grant St.
West Lafayette, Indiana 47906
psmith@purdue.edu

Doug Smith
Laboratory for Atmospheric
and Space Physics
University of Colorado at
Boulder
1234 Innovation Dr.
Boulder, Colorado 80303
doug.smith@lasp.colorado.edu

Torsten Hoefler
National Center for
Supercomputing Applications
University of Illinois at
Urbana-Champaign
Urbana, IL 61801, USA
thor@illinois.edu

Anna A. Labutina
Nizhni Novgorod State
University
603950, prospekt Gagarina 23
Nizhni Novgorod, Russia
anna.labutina@cs.vmk.unn.ru

Trinity Overmyer
Purdue University
585 Stadium Mall Dr.
West Lafayette, Indiana 47906
tcovermy@purdue.edu

ABSTRACT

This paper aims to describe methods that can be used to create new Student Cluster Competition teams from the standpoint of the team advisors. The purpose is to share these methods in order to create an easier path for organizing a successful team. These methods were gleaned from a survey of advisors that have formed teams in the last four years. Four advisors responded to the survey and those responses fit into five categories: (1) early preparation, (2) coursework specific to the competition, (3) close relationships with the hardware vendors, (4) concentration on the applications over the hardware, and (5) the need to encourage the team members to write papers about their experiences. In addition to these commonalities which may be best practices there are a few divergent but intriguing techniques that may also prove useful for potential advisors. Both will be discussed here and these methods can serve as a primer for anyone looking to start a new Student Cluster Competition team.

Categories and Subject Descriptors

A.1 [General Literature]: Introductory and Survey

General Terms

Undergraduate Education, Supercomputing

Keywords

Student Cluster Competition, HPC Education

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

TeraGrid '11 July 18-21, 2011, Salt Lake City, Utah, USA
Copyright 2011 ACM 978-1-4503-0888-5/11/07 ...\$10.00.

1. INTRODUCTION

The Student Cluster Competition is an event at the Supercomputing Conference that pits teams of six undergraduates against each other in order to complete the most computational science in a given power envelope with whatever cluster they can construct. The rules are fairly simple: Each team gets a power budget of 26 amps at 110 volts of power on two 13 amp circuits. All active hardware for the competition must fit into this envelope. The hardware must also be commercially available by the first day of the competition. The first day of the competition is spent setting up and tuning the cluster and getting base benchmarks using HPCC[16]. Once the final benchmark results are submitted to the judges, the competition begins and the students are running on their own technical knowledge without outside assistance. Historically there have been four applications per competition and have included such scientific codes as POP[3], GAMESS[22], WPP[18], POY[28], WRF[23], NWChem[27], CHOMBO[4], and NAMD[21]. Although there is only one overall winner, the contest is beneficial for everyone.



Figure 1: Undergraduates hard at work during the Student Cluster Competition at the 2010 Supercomputing Conference

The competition provides an excellent opportunity to teach hands-on High Performance Computing (HPC) in a competitive learning environment. During the training for competition, students learn valuable skills that will help them in their technical careers—not only technical skills, but also team skills and for some, technical writing skills. Many of the students in the competition go on to write technical papers about their experiences with the Student Cluster Competition. In addition, many consumers of HPC have looked to the Student Cluster Challenge in order to fill their technical job openings and have donated to the Student Cluster Challenge for this purpose[20]. For educational institutions with HPC foci, the benefits are self-evident: better educated students and name recognition in the HPC space.

Although entering a team into the competition can be daunting, it is a worthwhile endeavor for all participants. Because the competition is young, there is little information shared about starting Cluster Competition teams. Most published information about the competition is from a student perspective[15][31] or is concentrated on technical aspects of the competition[1][10][17][29]. Because of this, past advisors have created their own methods, without the benefit of any best practices or knowledge of pitfalls. This lack of information begs the question, “What are the common practices among the teams and are these the best practices?” This paper aims to be a starting point for answering that question, and a primer for creating a successful Student Cluster Competition team.

Table 1: Institutional Participation in the Student Cluster Competition by Year

Institution	2007	2008	2009	2010
Arizona State University		X	X	
Florida Agricultural and Mechanical University				X
Indiana University (IU)	X			
Technische University Dresden and IU		X		
Louisiana State University				X
National Tsing Hua University	X	X		X
Nizhni Novgorod State University				X
Purdue University	X	X	X	X
Stony Brook University	X		X	X
University of Alberta	X	X		
University of Colorado	X	X	X	X
University of Texas at Austin				X

[13] [14] [25] [26]

2. METHODS

In order to collect the methods that have been used to create a successful Student Cluster Competition team, current and past team organizers were surveyed. The original survey consisted of seven open-ended questions about the inception, execution and success of methods used in funding, institutional support, student education and working with hardware sponsors. The survey was distributed via e-mail.

The participants of the survey were all university staff, post doctorate researchers or graduate students. All potential participants were chosen based on their current and past participation in the Student Cluster Competition. A total of four advisors responded to the survey.

The survey elicited responses in three fairly broad topic areas:

- Resources—This includes hardware sponsorship, university assistance, and other financial support.
- Student education and preparation—Methods of instruction or independent studies
- Other advice—What else is important to know

The questions were designed to extract information about pertinent topics but also to explore those topics that perhaps were overlooked in the survey. The survey questions were:

- How did you get involved with cluster challenge?
- What kind of support has your institution given you?
- How did you get that support?
- How did you train the undergraduates?
- How did you choose your hardware sponsor?
- Did you get any financial or other support from other organizations besides your school/hardware sponsor?
- What haven’t I asked that you wish you would have known when you first started a team?

These questions led to a second survey which targeted time constraints and student training. In addition, follow-up interviews were done to go more into depth about topics that the respondents thought were important.

The information was coded based on common themes throughout the responses. The survey was created to find common methods of team creation and management in hopes of uncovering a list of best practices. This is not meant to be a comprehensive study, but a discussion among current advisors that aspirant teams can look to for direction and guidance.

3. INSTITUTIONAL SUPPORT

All institutions provided a baseline of support:

- Funding
- Computers/training clusters
- Access to HPC experts

In addition to this baseline, half of the institutions provided domain scientists to assist with the understanding of the applications. One advisor encouraged the students to seek out experts themselves:

“Once the applications are announced the students are encouraged to seek out scientists in those application disciplines and learn as much as possible from them via any means possible.”

Another advisor had a different approach:

“We have always had a faculty advisor committed to helping us and he was able to get domain science graduate students to lecture about the specific applications regularly.”

Additionally, extra programming classes were provided to students by two of the institutions.

Institutional support requests seem to come in two varieties: the respondents that were in staff positions generally wrote proposals for support from the institution, while the graduate students and post doctorate researcher respondents were able to get support with more informal inquiries.

Often, the university will provide travel grants to students participating in conferences like the SuperComputing Conference. Some advisors have had success soliciting funds from the deans of the students' home departments. One advisor asked team members to write a letter to their dean, explaining the importance of the competition and asking for travel expenses. The advisor reported that this was typically successful.

4. FINANCIAL SUPPORT

In the last few years, the Student Cluster Competition has provided hotels near the conference center, along with a small stipend for the week and entrance to the SuperComputing Conference. All other funding, including hardware, is the responsibility of individual teams.

As mentioned in section 3, all respondents receive funding from their institutions. Most of the money to support the teams comes from outside sources. Professional organizations such as ACM have donated funds, and sometimes even local alumni that work in the industry have offered mentoring or given testing equipment. Many teams solicit funding from larger companies that do not produce hardware or software resources directly related to the Student Cluster Competition. These companies are looking for qualified job applicants in the HPC field and are willing to provide funding that grants them access to this pool of potential graduates with HPC training. Job fairs on campus are an excellent way to make contact with these recruiters. One respondent has used this method with great success, not only securing funding but also internship and job opportunities for his students.

Another novel approach to the funding issue has been to partner with a company that is small or is new to the HPC space. Two respondents found companies that provided resources beyond what is normally expected from a larger hardware vendor. The vendors' goal of establishing themselves can elicit higher amounts of financial support, PR and technical expertise in an effort to expose their products to a wider market. This translated into financial support as well as help finding ancillary vendor support. For both teams, programming support was provided by the vendor in order to compensate for an atypical architecture.

It is important to include these sponsors during the early stages of student preparation and communicate to them how they are contributing to the educational experience of the students. One advisor said,

“During the preparation for the competition, we sent e-mails and photos of the work we were doing as well as creating a team website and Twitter and Facebook accounts. Appropriate signage of our sponsors at the competition, as well as publishing papers after the competition is important...In every paper and presentation we make sure to attribute our success to our supporters (and we made sure to communicate that to the sponsors themselves).”

5. HARDWARE SPONSORSHIP

In general, all the respondents had a previous relationship with their potential hardware vendors before obtaining competition sponsorship. However, if needed, the Student Cluster Competition organizers also typically have had a list

of vendors and contacts who have been interested in working with institutions that want to field a team[20].

In picking a hardware sponsor it is important to choose one that will be engaged with the team regularly. One respondent even said:

“It’s vital from the very start to have face to face contact with the hardware sponsor.”

In addition, a team must get a hardware platform that can support the Student Cluster Competition applications with only the effort of six undergrads. As an example of what can go wrong:

“We chose [the company] due to our good connections to the company. However, it turned out that the existing software ecosystem was not ideally suited for running the required applications.”

Sometimes when dealing with hardware vendors, time has been a common problem. Hardware sponsors should be courted early and be able to get the hardware into students' hands as soon as possible. Some hardware vendors with the best of intentions have been unable to deliver hardware on time and teams suffer for it:

“The plan was to get the cluster assembled and installed...we would use remotely. But [hardware vendor] was late with the delivery. We had one ‘competition-like’ node for a couple of months, then two nodes for a couple of weeks. We only got the whole system in [town] two days before it needed to be shipped to [the Student Cluster Competition].”

One respondent suggests the best way around this pitfall is to start early:

“The vendor should get that hardware in the hands of the students as fast as possible and be prepared to make the technical contacts from the manufacturer ready to work with the students to get the most out of the hardware as possible. Again communication early and often is very important.”

Table 2: Student Cluster Competition Awards

Year	Winner Overall	Highest Linpack
2007	University of Alberta	National Tsing Hua University 420GFlops
2008	Technische University Dresden and IU	National Tsing Hua University 703GFlops
2009	Stony Brook University	University of Colorado 692GFlops
2010	National Tsing Hua University	University of Texas at Austin 1.07TFlops

[2] [7] [8] [12] [19] [24]

It is also imperative that the hardware vendor know what the parameters of the competition are. When speaking about the power requirements one respondent noted:

“This is not simply a situation where you can throw the fastest hardware at the problem and expect to win. In fact, the team posting the highest Linpack has never won the competition [as shown in Table 2]. Intimate knowledge with the applications is often more valuable than the latest and greatest hardware. It’s important that the hardware vendor provides hardware that is designed with the specific requirements of the competition in mind.”

Getting the hardware vendor’s technical staff engaged and available early has been a challenge for at least two of the respondents.

Hardware vendors often put a tremendous amount of money and time into this and it is important to set their expectations. One participant said:

“I would caution not to set winning expectations and to be very realistic with the vendor about the fact that this is an undergraduate competition.”

Some also found that billing this event as a learning opportunity for students instead of a competition between vendors has been an important step in setting expectations.

6. STUDENT TRAINING

Three of the four participants reported that their universities offered courses either in preparation for the competition or in order for the students to receive course credit for participating. Courses that are tailored for the Cluster Competition are often supplemented by standard HPC curriculum at the institutions.[11][5][9] The only team that does not have dedicated coursework is currently working toward designing a class specifically for the Student Cluster Competition. One team even took a class on presentational speaking before attending the conference. A respondent said this about the coursework:

“...creating a ‘for credit’ class ensures that the students are well motivated to succeed and do their best. The undergraduate curriculum for computer science or engineering is very demanding and students will often fizzle out as their course work begins to demand more of their attention. If their efforts for the cluster competition are integrated with course work they are less likely to drop out of the cluster competition efforts. As a side benefit they tend to regard their course work as more meaningful when its attached to something bigger like the competition.”

In addition to HPC coursework, two of the respondents said that knowing and understanding the applications can be even more important than the HPC fundamentals. One advisor went as far to say:

“In the first year, we focused on broad education in supercomputing with a specific curriculum. While we believe that this was very valuable for the students, we realized that we did not have the time to have enough hands-on sessions to prepare the students for the challenge. In the second year, we concentrated on direct preparation of the challenge (we had some students of

the previous year in the team) which eventually lead us to success...”

A second advisor had a more fundamental approach:

“Early on we focus on basics like linux system administration and the basic taxonomy of parallel high performance computers (Flynn’s taxonomy[6]). For most of the students regardless of what major or year this is the first exposure they have had to hands on high performance computing. Once they have a firm grasp of the basics then we begin to concentrate on the applications...By working with the applications the students begin to understand what these questions mean and why they are important to answer. We typically begin with the HPCC benchmarks since they are consistently used in all the competitions and they are good teaching tools since they stress various aspects of the system like memory or interconnect.”

Independent study by the students, especially on basic, core concepts seems to be essential for all the participants, especially for those universities that do not have background requirements for their team members. Besides access to advisors who are often industry experts, students tend to have access to faculty at their home universities who are doing research in the field. One team encourages attending outside lectures and participation in non-Challenge computer science projects both within the university and in cooperation with partner companies.

Three of the participants put emphasis on continuing the students’ education post-Challenge:

“Writing follow-up papers is important to the process. If we are to train/mentor students to go on and further their education in Computer Science, Engineering or the Sciences then publishing is a very important part of that process.”

Documenting and presenting their findings is one way to not only give those students real world experience, but also to contribute to the scientific community. Many students from several institutions have written papers about their experiences [1] [10] [15] [17] [29] [30] [31].

This student training and post-Challenge effort certainly can pay off. Three of the respondents have had students who have gone on to work in the HPC field with many of them working at national labs and other universities as HPC specialists. One advisor said:

“[A national lab] and [a famous search company] both recruited one of the undergrads based on his submitted papers from [the Student Cluster Competition]. In fact we have competition alumni working at [a university HPC center], [a national lab] and [a major HPC hardware vendor].”

7. TIMELINES AND RECRUITING

In the survey, every respondent emphasized that it is imperative to start planning as early as possible. Originally the application date for the competition was in mid-summer and the acceptance notifications were not sent out until August.

At the behest of the participating teams, the due date for the Student Cluster Competition now falls in mid-April and notifications happen in early summer. When asked about the best time to start, one advisor said:

“As early as possible. At the beginning of the year or even earlier (I started November/December).”

Another respondent joked:

“If you have to ask if you have enough time to prepare, you don’t.”

Respondents indicated that no matter how much time was set aside for preparation, they often felt short on time.

Recruiting students proved to have its pitfalls as well. Two of the respondents said that they have trouble keeping students’ commitments from Spring through Fall. As mentioned above, starting preparation early is important and starting students’ training early is equally as important. One advisor said:

“One big challenge we have is the timing of the competition we start recruiting in the spring and we usually have over 10 students interested by the end of the spring semester. But then the summer comes and students tend to fall out of interest and we have to do another round of recruiting the first couple of weeks of the fall semester.”

Two respondents agreed that a good place to recruit students is the local chapter of the Association for Computing Machinery(ACM). Another novel approach was put forth by another respondent:

“We hire undergraduates as Junior System Administrators to help keep our community clusters running and they are a natural fit for the [the Student Cluster Competition]. Luckily for us, the inverse is true as well.”

Getting students excited is also important. Sometimes the idea of HPC can be daunting for students. One respondent mentioned:

“During [the Supercomputing conference] I was asked a few times: “How did you find students for the team? Ours are scared” ”

Finding students that are already familiar with the software stack was one respondent’s tactic:

“When we recruit we try to get students that are familiar with Linux. We have found that knowledge of Linux is a good baseline to start with and less knowledge than that can be troublesome.”

8. CONCLUSIONS AND FUTURE WORK

The respondents to the survey have agreed on some key points:

- **Start preparing early:** At the beginning of the year or earlier. Much of the work as an advisor is done before the application is due in April.
- **Create a course specifically for the competition:** Find faculty who will support the team.

- **Find a hardware vendor early and create a close relationship with them:** Many universities have existing relationships with hardware vendors; leverage those if possible.
- **Knowing the applications is more important than having the fastest hardware:** Make sure that the vendor understands the requirements of the competition. When instructing, focus on the applications. The team with the fastest and best hardware has not yet won the Challenge. Applications take priority over the hardware.
- **Encourage the undergraduates write papers and share their experience:** Once the competition is over, the undergraduates should write papers about their experiences. This is valuable experience for them and can get them noticed in the job market. Student written papers may also lend credibility with future hardware vendors or sponsors going forward.

Though this survey has uncovered some effective methods for fielding a team, it will be necessary to acquire a larger sampling and more information in order to develop a comprehensive list of best practices. As the teams and competition evolve, the practices will as well. Every year, each returning team hones its own methods and tries to correct last year’s mistakes. These iterations will provide the best practices in the future. While this paper may not have uncovered a comprehensive list of best practices, there are some good starting points and information for those who want to create a team.

In addition to a more comprehensive survey, other future work may include the creation a sample curriculum for HPC basics with a Student Cluster Competition focus that can be shared publicly. Overall competition strategies would also be an important topic to explore. Other aspects of the competition such as hardware and software optimization also warrant a deeper look.

9. ACKNOWLEDGEMENTS

Thanks go to Alex Younts, Mike Shuey and Randy Herban at Purdue University for reviewing this paper.

10. REFERENCES

- [1] R. Beck, A. Filion, P. Greidanus, G. Klok, C. Kuethe, P. Lu, C. Macdonnell, A. Nisbet, and S. Portillo. The cluster challenge - 6 students, 26 amps, 44 hours. In *Proceedings of the 2008 Linux Clusters Institute*, 2008.
- [2] R. Brueckner. Tacc’s student cluster challenge team wins highest linpack award at sc10. <http://insidehpc.com/2010/11/24/15646/>.
- [3] K. Bryan. A numerical method for the study of the circulation of the world ocean. *Journal of Computational Physics*, 135:154169, 1997.
- [4] P. Colella, D. T. Graves, N. D. Keen, T. J. Ligocki, D. F. Martin, P. W. McCorquodale, D. Modiano, P. O. Schwartz, T. D. Sternberg, and B. V. Straalen. Chombo software package for amr applications design document. <https://seesar.lbl.gov/anag/chombo/ChomboDesign-3.0.pdf>.
- [5] J. J. Evans. Ecet 325 - computer architecture, modeling & performance analysis. <http://www.tech.purdue.edu/eet/courses/ecet325/>.

- [6] M. J. Flynn. Some computer organizations and their effectiveness. *IEEE Transactions on Computers*, C-21, No 9:948–960, September 1972.
- [7] B. Gorda. 2008 cluster challenge results. <http://insidehpc.com/2008/12/08/2008-cluster-challenge-results/>.
- [8] B. Gorda. Sc07 holds first cluster challenge. <http://www.hpcwire.com/features/17904509.html>.
- [9] T. Gross. Breadth in depth – a 1st year introduction to parallel programming. In *Proceedings of SIGCSE'11*, March 2011.
- [10] T. Hoefler, T. Schneider, and A. Lumsdaine. A Power-Aware, Application-Based, Performance Study Of Modern Commodity Cluster Interconnection Networks. In *Proceedings of the 23rd IEEE International Parallel & Distributed Processing Symposium, CAC'09 Workshop*, May 2009.
- [11] E. R. Jessup and C. J. C. Schauble. Csci 4576 - high performance scientific computing. <https://csel.cs.colorado.edu/~csci4576/>.
- [12] G. Kline. Sc09 cluster challenge: Purdue, itap student supercomputing team wins 'green' computing award. <http://www.purdue.edu/newsroom/students/2009/091123EvansChallenge.html>.
- [13] J. Leidel. Sc08 cluster challenge teams. <http://insidehpc.com/2008/10/21/insidetrack-sc08-cluster-challenge-teams/>.
- [14] J. Leidel. Sc09 student cluster competition: Go green! <http://insidehpc.com/2009/10/29/sc09-student-cluster-competition-go-green/>.
- [15] D. Leverman, D. Smith, B. Cairns, J. French, T. Knappe, and M. Levy. Sc07 cluster challenge: A student perspective. In *Proceedings of the 2008 Linux Clusters Institute*, 2008.
- [16] P. Luszczek, J. J. Dongarra, and J. Kepner. Design and implementation of the hpc challenge benchmark suite. *CT Watch Quarterly*, 2(4A), November 2006.
- [17] J. Mueller, T. Schneider, J. Domke, R. Geyer, M. Haesing, T. Hoefler, S. Hoehlig, G. Juckeland, A. Lumsdaine, M. Mueller, and W. Nagel. Cluster Challenge 2008: Optimizing Cluster Configuration and Applications to Maximize Power Efficiency. In *Proceedings of the 10th LCI International Conference on High-Performance Clustered Computing*, Mar. 2009.
- [18] S. Nilsson, N. A. Petersson, B. Sjögreen, and H.-O. Kreiss. Stable difference approximations for the elastic wave equation in second order formulation. *SIAM Journal on Numerical Analysis*, 45:1902–1936, 2007.
- [19] D. Olds. Student cluster competition: The runners and riders. http://www.theregister.co.uk/2010/11/09/student_cluster_competition_university_profiles/.
- [20] D. Olds. Webcast: Student cluster competition 2011. <http://whitepapers.theregister.co.uk/paper/view/1974/>.
- [21] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten. Scalable molecular dynamics with namd. *Journal of Computational Chemistry*, 26:1781–1802, 2005.
- [22] M. W. Schmidt, K. K. Baldrige, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery. General atomic and molecular electronic structure system. *Computational Chemistry*, 14:1347–1363, 1993.
- [23] W. C. Skamarock, J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers. A description of the advanced research wrf version 2. Technical Report TN-486+STR, National Center for Atmospheric Research, June 2005.
- [24] B. Sparks. The university of colorado at boulder receives sc10 student cluster competition fan favorite award. http://www.hpcadvisorycouncil.com/pdf/Press_releases/pr_12_6_10_SC_10_Cluster_Challenge.pdf.
- [25] T. Suarez-Brown. The teams - sc10 scc. <http://sites.google.com/site/sc10scc/the-teams>.
- [26] Sc07 - cluster challenge. <http://sc07.supercomputing.org/?pg=clusterchlng.html>.
- [27] M. Valiev, E. Bylaska, N. Govind, K. Kowalski, T. Straatsma, H. van Dam, D. Wang, J. Nieplocha, T. W. E. Apra, and W. de Jong. Nwchem: a comprehensive and scalable open-source solution for large scale molecular simulations. *Computer Physics Communications*, 189:1477–1489, 2010.
- [28] A. Varòn, L. S. Vinh, and W. C. Wheeler. Poy version 4: phylogenetic analysis using dynamic homologies. *Cladistics*, 26:72–85, 2010.
- [29] A. Younts, A. Howard, P. Smith, and J. J. Evans. Bringing disruptive technology to competition. In *Proceedings of the 10th LCI International Conference on High-Performance Clustered Computing (CDROM)*, March 2009.
- [30] A. Younts, A. Howard, P. Smith, and J. J. Evans. Undergraduate education in HPC with the SC Cluster Challenge. In *Proceedings of the 4th TeraGrid Conference*, June 2009.
- [31] A. Younts, A. Howard, P. M. Smith, and J. J. Evans. Undergraduate experience in clustering at the SC07 Cluster Challenge. In *Proceedings of the 2008 Linux Clusters Institute*, 2008.