SC10 Student Cluster Competition Proposal

NTHUCS Team

Team Information

Institute:

Department of Computer Science, National Tsing Hua University, Hsinchu, Taiwan 30013, R.O.C.

Supervisor:

Dr. Yeh-Ching Chung <ychung@cs.nthu.edu.tw>

Members:

Yu-Cheng Chen <sachung28@yahoo.com.tw>, Junior
Yu-Chun Hsueh <ox92000@yahoo.com.tw>, Junior
Yi-Man Ma <eye548@hotmail.com>, Junior
Zheng-Wei Huang <fredyj20135@yahoo.com.tw>, Junior
Yi-Chun Chen <rimi7609@gmail.com>, Junior
Meng-Kai Liao <dabeniao@gmail.com>, Sophomore

Sponsors:

HP Development Company

Why are you participating?

We are students major in Computer Science. During the past 2-3 years, we have learned some background knowledge from textbooks. However, instead of sticking our noses on textbooks, we believe that we should have some experiences on real world applications or implementations. At the beginning of this year, we started to discuss with several professors about their research in the hope that we can have some opportunities to use those knowledge we have learned to do researches with other students. After several discussions, our current advisor gave us the information about this competition, which can help us to learn how a cluster built, understand how to develop and analyze parallel programs, and even have chances to exchanges our experiences with other students from other universities. Therefore, we start to find other classmates who are also interested in this competition to form a team. We hope that we can learn more domain knowledge of HPC via this competition and understand the related work in this research area. Also, we hope that we can be the winner team in this competition.

Why do you believe that you’ve put together a winning team?

HPC is totally new to us. Professor Chung has asked two experienced upper-classmates, who have attended the cluster challenge in 2007 and 2008, to train us to master the parallel programming, the cluster system, the Linux operating system, HPL, and some HPC applications. A detailed training schedule is set. Each of us has been assigned two tasks. Currently, we are not a winning team yet. But, we are diligent and talent students in our department. We believe that we can put together a winning team at the end of October if we follow the training schedule and complete all the goals set.

What sorts of diversity in skills does your team possess?

All students have excellent programming skills. In addition, one of the team members is the system administrator of all servers used in the department. He is familiar with Linux operating system.
Since our native language is not English, two of team members can speak English fluently.

**Why will your team work well together?**

Our team formed in January this year. The upper-classmates have assigned some tasks for us to work together. During the team work, we found the strength and weakness of each other. Therefore, tasks can be well distributed to the team members who are the best candidate to do them. Through the team work, we also learned how to help each other when needed. We have gone through this kind of team work many times and will continue to do it until the end of October. In this way, we believe that we can put together as a team and work well together.

**What experience do you and your team members have?**

No team member has attended any competition before or has any experience on HPC or clusters. However, Professor Chung has asked two upper-classmates who have attended the cluster challenge in 2007 and 2008 to train and help us. The upper-classmates have given some tutorials and training to help us to build the background of HPC. They also let us understand what we should do in the future. During the training, we have finished some programs with MPI, pthread, and OpenMP. We also built up our cluster in April. As the training continuing, we will have more experiences on HPC.

**How will your team work together to tune and optimize the application set?**

For the task assignments, one team member will take the role of system administrator and others as application managers. The system administrator is responsible for the kernel building and system parameter setting to make the whole system more stable and faster. The application managers are responsible for studying and tuning application sets. We plan to assign each application to 2 application managers, one plays the primary role and the other plays the supporting role. For each application, the application managers need to analyze the application, understand how to use this application, and figure out the characteristics of the input data and the application. In this way, each application is covered by two team members. With the understanding of applications, we can come up with a best schedule in terms of application data sets, number of processors, and time constraint.

**Desired Cluster Configurations**

**Hardware:**

1. HP ProLiant BL2x220c x 5
2. HP Blade System c3000 Enclosure x 1
3. Intel Xeon @ 3.0 GHz Quadcore 2 x 6 MB L2 (4 Cores x 2 CPUs) x 10 Nodes = 80 Cores x 20
4. PC2-533 DDR2 Registered DIMMs x 40 (4 GB x 4 Slots) x 10 Nodes = 160 GB
5. HP 4X DDR IB Dual-Port Mezzanine HCA for HP BladeSystem c-Class x 10 (20 Gb/s InfiniBand PCI-E card)
6. HP Double Wide Switch x 1 (InfiniBand switch, 960Gbps non-blocking switch capability)
7. Broadcom Corporation NetXtreme II BCM5708S Gigabit Ethernet x 1

Our vendor sponsor HP has provided us with the latest blade system. The blade server could share power and has the cooling system so it does not generate as much heat as traditional servers. Furthermore,
the new double-density HP blade is designed for efficient power utilization and smart control of cooling hardware components. It provides a management system built in the enclosure. The new administration system could automatically monitor all hardware conditions and control different components dynamically in different situations including turn off redundant power supply and turbo the fans under heavier workloads.

With the sponsorship from Intel, we will use special low power CPU to save a lot of power and add more computing nodes to our cluster. This will extend our system to 10 computing nodes which contains eighty 3GHz cores. The Rpeak of Linpak will be 960 GFlops.

We are also working with HP internal research center to improve the performance of our blade system. We are planning to use 20 Gb/s DDR InfiniBand connection to speedup the interconnection between computing nodes, and use Solid-State Disks to reduce the power consumption of each node.

We have found that the power saving of SSDs are beneficial for building a HPC system. In our recent tests, SSDs require only 1/4 of the power consumed by traditional Hard Disks.

With more computing nodes and higher clock speed of each CPU connected under faster network, we believe the result will showcase the incredible computing ability of cluster using limited power.

Software:

1. Operating System
   I. GNU/Linux CentOS 5.4, x86_64
2. Compiler
   I. GNU C/C++ Compiler 4.4.0
3. Parallel Virtual File System
   I. PVFS2 2.7.1
4. Message Passing Interface (MPI) And InfiniBand Support Software
   I. MPICH2 1.2.1p1
   II. MVAPICH2 1.4.1
   III. OFED 1.4
5. Management and Monitoring Tools
   I. Ganglia 3.1.7
6. BLAS Library
   I. GotoBLAS2 1.13

In this competition, we purpose to build up an environment which is matched to our desired hardware platform so that users can use the whole system to solve their problems more quickly with very easy manipulation. In order to achieve this goal, we look into the following issues and uses the corresponding software as the above list:

In order to build up the system, we select CentOS as the operating system on our cluster. CentOS is a GNU/Linux distribution released from RedHat. Based on its stability in the whole system, convenient package management tools, and many supports and information on the web, we think we can build up our purposed system through this distribution.

To make users can concentrate on their work, rather than use much time to get the information of the current status about the whole system, we choose Ganglia to help users to monitor their system. Ganglia is a scalable distributed monitoring system with Web-based UI for HPC systems. It allows us to remotely view live or historical statistics (such as CPU load or memory consumption) for all machines. Those statistics collected by Ganglia can let us compare the performance results with system status.

Besides, in order to reduce the impact of disk I/O, we prefer to use parallel file system to improve
the latency of disk I/O during the computation. As for the parallel file system, we purpose to use Parallel Virtual File System (PVFS). PVFS is a distributed file system which distributes files across multiple servers and allows multiple processes to access the data at the same time. By using one of them, the time of reading or writing a large data can be remarkably reduced.

For tuning HPCC benchmark, to reduce the cost of frequently message passing in local computing node, we've used shared memory to replace message passing in each local node. GotoBLAS2 will be used as our BLAS library because of the optimization in its implementation and multi-threading support during computing. Moreover, GotoBLAS2 supports more recently CPU to do these computations, and we trust this feature will help us to tuning HPCC benchmark very well.

As the above issues and software list, we trust we can reach the goal we purpose, and have more impressive result in the competition.

For the showcase period of the conference, what demonstrations do you anticipate to impress and attract conference attendees?

Based on our experience, visualized result is very important issue for demonstration. By visualizing the system information or the results, we can make attendees understanding the current progress about our cluster and competition results faster and feeling impressive on our team. During the competition, we purpose to design an integral interface like web pages or an application to show the real-time information about our cluster and the result with graphical method to the conference attendees.

Besides, we also hope to find something interested and run them on our cluster when our machine is free. This will help attendees to understand the performance of our cluster, and impress them on our team and cluster.

Explain the commitment of the institution to educating the broader student community about the usefulness and the accessibility of High Performance Computing at your institution; explain how cluster computing is integrated in the educational curriculum of the proposing institution.

Professor Chung has offered Parallel Programming course at NTHU CS for seven years. This course is mainly for graduate students, but in the last two years, there are more and more undergraduate students attending the course. It is a trend that more and more Multi-Core architectures are adopted in mainstream systems and the students are interested in learning programming skills for Multi-core systems. Parallel algorithms and well-known parallel programming models such as MPI, OpenMP and Pthread are lectured in this course. Professor Chung has also asked all students to write parallel programs using aforementioned programming models and run programs on different clusters which have varied hardware and software configurations to compare the results.