

Using OSCAR to Win the Cluster Challenge



University of Alberta

Paul Greidanus, Gordon Klok

The Cluster Challenge

- New challenge event introduced at Super Computing 2007.
- Six team members without a undergraduate degree, a faculty coach from the institution.
- Competition consisted of running the HPCC benchmarks and three applications, GAMESS, POP (Parallel Ocean Program), POV-Ray.
- Power limit of 26 Amps

Team Alberta



From left to right, Gordon Klok, Chris Kuethe, Paul Greidanus, Stephen Portillo, Andrew Nisbet, Paul Lu, Antoine Filion
Not Pictured: Bob Beck, Cameron Macdonel

Our Cluster

- Our vendor partner SGI supplied 5 Altix XE310 servers,
- Altix XE310 1U chassis contains two nodes sharing a single power supply. Each node consisted of:
 - Two quad-core Intel Xeon 5355 CPUs running at 2.67Ghz with 8MB of L2 cache.
 - 250GB SATA drive.
 - 16GB of RAM, we later added 8 GB to two of the nodes to get 24GB of ram.



The competitors

SC07 Cluster Challenge

Team	Sponsor	chip	nodes	sockets	cores	mem/ node	interconnect	operating system
Alberta	SGI	Xeon <i>2.66 Ghz</i>	8	16	64	16 GB	20 Gbit Infiniband	SL
Colorado	Aspen Systems	Xeon <i>2.66 Ghz</i>	6	12	48	8 GB	Dual 10 Gbit Infiniband	CentOS
Indiana	Apple	Xeon <i>3 Ghz</i>	9	18	36	8 GB	Myrinet 10G over 10GE	OS X
NTHU	ASUS	Xeon <i>2.83 Ghz</i>	6	12	48	12 GB	10 Gbit Infiniband	CentOS
Purdue	HP	Opteron <i>2.2 Ghz</i>	14	28	54	4 GB	20 Gbit Infiniband	CentOS
Stonybrook	Dell	Xeon <i>1.86 Ghz</i>	13	26	100	8 GB	5 Gbit Infiniband	Debian

Courtesy: Brent Gorda

Why OSCAR?

- OSCAR allowed us to deploy the cluster quickly and focus on the important thing: Applications.
 - Not everyone used a product like OSCAR.
- Changes can be pushed to nodes quickly.
- Dealt with all the details. SGE, Ganglia, SystemImager
- Used ganglia as part of our visualization strategy.

Cluster preparation

- Installed the head node using Scientific Linux 4.5. OSCAR 5.0 used to build client image and push it out to the nodes.
 - Sun Grid Engine chosen over Maui/Torque.
- Not perfect, needed new kernel, needed SystemImager update.
- No 3rd party compiled MPI libraries, and no Infiniband.

OSCAR Limitations and Future Features

- C3 – timeouts with dead nodes
- SGE Limitations – Does not work after reimage
- MPI Limitations
 - No IB, no OFED, no Alternative compilers
- IPMI support – reboot nodes, predict failures.
- Application Checkpoint/Restart
 - Linux is weak here out of the box
 - This could be a killer-feature.
- No non-headnode /home NFS possible.

Concluding remarks

Clusters have evolved, the tools need to keep evolving.

Students can use tools like Oscar to build clusters, it's not rocket surgery

Using the clusters with high performance interconnect, and with non-standard configurations needs to be addressed.

Acknowledgements

Brent Gorda and the cluster challenge organizers.

SGI Dan St.-Germain,

Jimmy Scott, Martin Pinard, Marcel Bourque, John Baron, Kah-Song Cho, corporate

Computing Science at the University of Alberta

Cam Macdonell, Yang Wang, Neil Burch, Steve Sutphen, Jonathan Schaeffer, Jill Bagwe,
Sheryl Mayko, Carol Smith, Ruth Oh

University of Alberta

Alex Brown, Paul Myers, Mariusz Klobukowski, Ron Senda,

Mark Gordon, Greg Lukeman, Keith Thompson, Yimin Liu, Asia Embroidery



SC08 Cluster Challenge

showcasing the
**Next Generation of
High-Performance
Computing
Professionals**
Austin Convention Center
November 15-21, 2008



Important Dates:
Entries Due: July 31, 2008
Notification: August 15, 2008

Questions: visit:
<http://sc08.supercomputing.org>
email: cluster-challenge@info.supercomputing.org

This year, SC08 invites teams of undergraduate students to rise to a new Cluster Challenge.

The **SC08 Cluster Challenge** is a showcase event in which teams of next-generation high performance computing talent harness the incredible power of current-generation cluster hardware. This challenge brings together an international field of teams that share a "need for speed" and a willingness to earn the top prize. The event promises to be exciting, educational and a truly rewarding experience for all involved.

Taking place **Nov. 15-21, 2008, at the Austin Convention Center** in Austin, TX, six teams of undergraduates working with a faculty adviser and cluster vendors will assemble, test and tune their machines until the green flag drops on Monday night as the Exhibit Opening Gala is winding down. The race now begins as the teams are given data sets for the contest. With CPUs roaring, teams will be off to analyze and optimize the workload to achieve maximum points over the next two days.

In full view of conference attendees, teams will execute the prescribed workload while showing progress and science visualization output on large displays in their areas. As they race to the finish, the team with the most points will earn the checkered flag - presented at the awards ceremony on Thursday.

After the checkered flag drops, teams are invited to partake in the side-show, where they can spin their wheels and show off what they've learned and what they can do with the equipment with demonstrations that defy gravity, simulate blood flow, visualize earthquakes, search the genome, or perhaps even model a cure for AIDS.

Are you up to the challenge?

To be considered for one of the six teams:

- 1) Visit <http://sc08.supercomputing.org> for rules and entry details.
- 2) Form a team of up to six undergraduates, plus a faculty supervisor.
- 3) Contact a cluster vendor for equipment and support.
- 4) Submit proposal.

SC08 Sponsors:
IEEE Computer Society
ACM SIGARCH

