

Tony Dale
BlueFern Supercomputing
University of Canterbury
December 15, 2009

Supercomputing 2009 Conference Report

Introduction

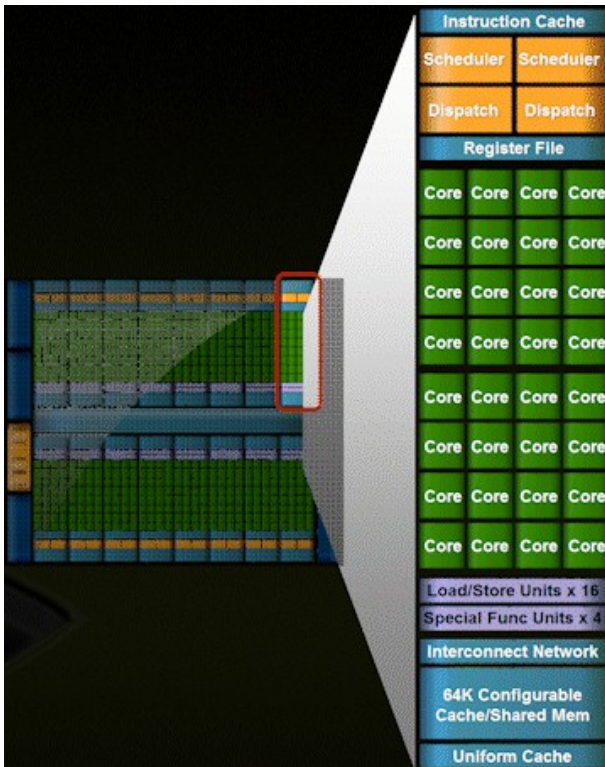
Thanks to support from the KAREN Travel Fund and BlueFern®, I attended the Supercomputing 2009 (SC'09) conference in Portland, Oregon, from November 14-20th. SC'09 (<http://www.sc09.supercomputing.org/>) is the largest supercomputer conference in the world and some brief statistics are:

- Total registrations approaching 12,000, with 1,878 from 70 countries outside the USA
- 261 technical papers submitted with over 1100 reviews, and an acceptance rate of 22.6%
- 28 tutorials accepted from 71 submissions (39% acceptance)
- 29 panel discussions
- Innumerable meetings and Birds of a Feather sessions
- Total of 318 exhibitors and 1.2 hectares of exhibit space
- 123 Research Exhibitors – a record
- 195 Industry Exhibitors

Here I will begin by presenting four of the larger technical themes from SC'09 for which I attended workshops and presentations: GPU Computing, Grid and Cloud Computing, PGAS Languages and Exascale Computing. I will then briefly mention some of the Birds of a Feather sessions and meetings I attended. Further information is available on the web page for the KAREN E-symposium I ran, at <http://www.wiki.karen.net.nz/index.php/BlueFernUserGroup141209>.

GPU computing

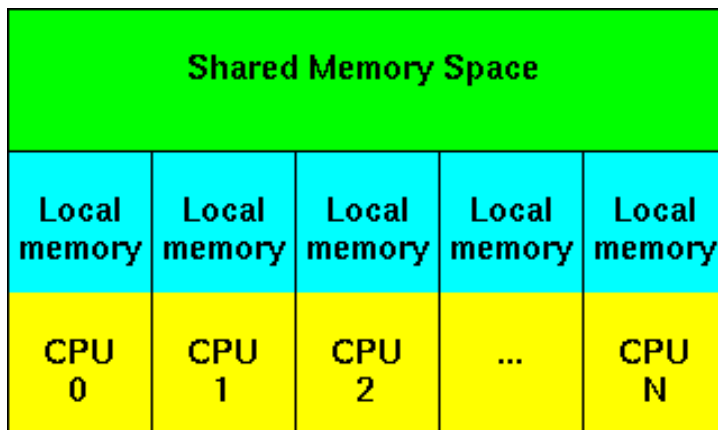
Graphics Processing Unit (GPU) computing was galvanised by the release of Nvidias CUDA software development kit three years ago. It was the subject of a day-long tutorial at SC'09, as well as featuring in papers, posters, and being heavily promoted at the exhibition by Nvidia and other manufacturers. GPU computing makes use of high-end commodity graphics cards manufactured by companies such as Nvidia and ATI, costing, for example: \$NZ3,500 each for the latest Nvidia “Fermi” cards. The Fermi cards contain a GPU made up of 512 powerful vector processors, each of which approaches the compute power of a Cray supercomputer (also a vector processor) from the 1970s. When operating at peak speed, these new GPU cards offer 768 peak GFLOPs of double-precision math performance - such huge performance has resulted from the innovative drive of the computer gaming market. CSIRO regard this new technology as important to their e-Research strategy, and in November they commissioned a GPU cluster with 61,440 compute cores, also training thirty of their scientists to code their experimental software for GPUs.



Architecture of the NVidia "Fermi" GPU. 16 "Streaming Multiprocessors" each control 32 processing units.

The challenge for GPU computing is to utilize the amazing performance of GPUs for real scientific problems: scientists have been grappling with the limitations of GPU architectures, which require data to be kept in a very small, high-speed, cache, and explicitly manipulated with library calls. SC'09 presented technical papers and posters showing that the problems of GPU computing are beginning to be solved. For example: one programming strategy is to sacrifice computation speed by such methods as recalculating values rather than storing and retrieving them, for the sake of keeping data in the high-speed cache. More commercial software tools for GPU computing are now also emerging, with the Portland Group introducing an "Accelerator" GPU programming model for their compilers, and Nvidia making available a plugin for Microsoft Visual Studio called Nexus.

PGAS languages



The basic PGAS programming model, illustrating hybrid shared and local memory

Partitioned Global Address Space (PGAS) languages offer a convenient programming model for distributed-memory architectures such as GPUs and MPP machines like the BlueGene. PGAS makes explicit the discontinuous nature of distributed-memory machines – this is actually an advantage from the point of view of efficiency, because memory local to a CPU can be easily accessed, enhancing locality of execution. Other advantages of PGAS include it having an easy programming idiom for globally sharing data between, potentially, billions of processes, and also asynchronous parallel execution of GPU/CPU and cell processor code. The PGAS tutorial, birds-of-a-feather session and conference exhibits at SC'09 evidenced plenty of activity in the PGAS space, with presentations on languages extending the syntax of Fortran (Co-array Fortran), C (United Parallel C) and Java (Titanium and X10).

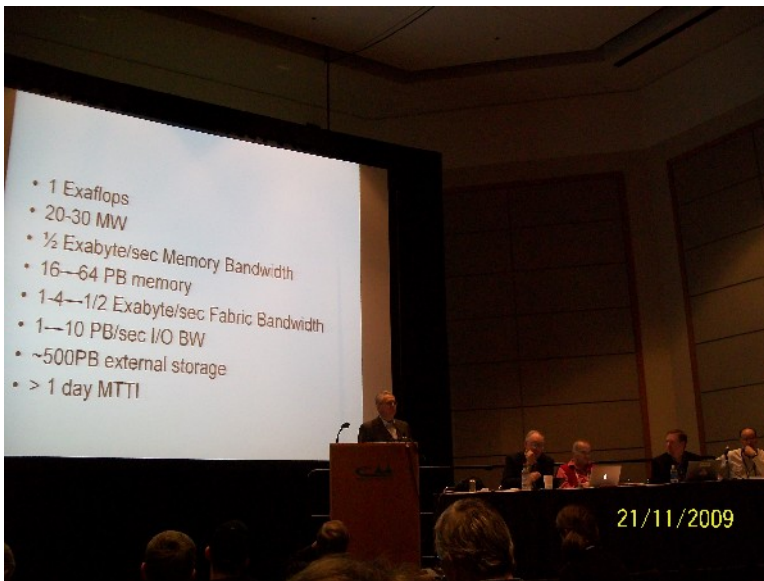
Grid and Cloud Computing

The two modern faces of widely-distributed computing: Cloud Computing and Grid Computing, where both well-represented at the conference, with tutorials, workshops, birds-of-a-feather sessions, papers and posters. It's difficult to see why Cloud Computing was being touted (by some) at SC'09 as a replacement for Grid computing, when computational Grids offer a superior technical infrastructure. Perhaps the approach of Cloud Computing is “just enough parallelism” for some users. I went to a presentation about the Magellan scientific Cloud project (<http://www.lbl.gov/cs/Archive/news101409.html>) and asked why they are implementing a Cloud and not a Grid, but didn't get a clear answer from the presenter - I think the reason is that DARPA wants to maintain diversity in their eResearch ecosystem, as much as anything else.

After the conference, I briefly surveyed two well-known grids: the Teragrid (<http://www.teragrid.org/>) and New Zealand's BeSTGRID (http://www.bestgrid.org/index.php/Main_Page) for their facilities. At the time of writing, I found that the Teragrid offers access (via such diverse interfaces as Apples iPhone, for example) to over 30 petabytes of data storage and over 650 widely-deployed scientific applications and over 1600 local applications from 11 HPC provider sites. BeSTGRID offers access to 8 clusters or HPCs, with a few dozen deployed applications and more than 100 terabytes of distributed storage. This is significant eResearch infrastructure that appears to be increasing in power by the day. Certainly our

experience here at BlueFern is that more and more users are coming in via BeSTGRID.

Exascale Computing

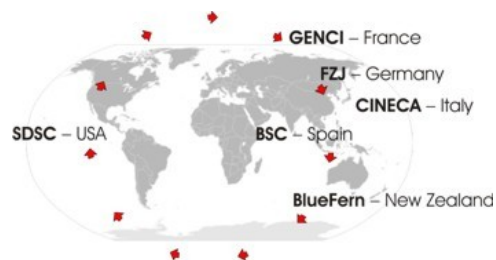


Goals for constructing an exascale computer, a slide from the SC'09 workshop

One of the most interesting features of SC09 for me was a workshop on the race towards exascale computing (10^{15} FLOPS). There are many challenges to be faced for exascale – not the least of which is that no-one is sure exactly how to build a realistic exascale supercomputer, or whether it could be effectively used once built. Nevertheless, exascale computing offers such possibilities as atomic-scale simulations that last for milliseconds rather than nanoseconds, and weather forecasts and global climate modelling of unprecedented accuracy.

The SC'09 exascale workshop presented the current progress towards exascale computing, with speculation that such computers will be built before 2020. Energy consumption is already being addressed by, for example: GPU-type technology, which currently is achieving in the order of 120W per TFLOP (eg: AMD claims this for their GPU cards). However, that would still be a 120 MW energy consumption for an exaflops machine with a quite limited architecture, even if other problems of construction could be solved.

HPCWorld Project



As BlueFern's representative, I participated in the HPCWorld project kick-off meeting run by CINECA, the Italian national HPC centre, at SC'09. HPCWorld (<http://www.hpcworld.eu>) is an international research project for which BlueFern is partnering with HPC centres from the EU (where the project is based) and the USA.

The aim of HPCWorld is to survey and analyze the access models used by HPC centers worldwide, and find how they manage requests for allocation of resources that come from researchers. The consortium will then create a model and procedure for reviewing applications and selecting the

most challenging, suitable for potentially many different kinds of Research Infrastructures, not just HPC. The procedure will then be offered as best practice to HPC centres and the scientific community.



Dr Francesca Garofalo speaking at the HPCWorld kick-off meeting, SC'09

The HPCWorld project will run through 2010-11, beginning with a survey of local & national HPC Centres.

Small HPC Group

The Small HPC group is a mailing list (small-hpc@googlegroups.com) that resulted from a birds-of-a-feather session at SC'09. The participants were small organizations that operated HPC systems, including Sections of IT Depts, College HPC units and independent organizations such as BlueFern. The session discussed issues such as staffing, money, users, etc, from the "small centre" viewpoint of having fewer staff, specialists and less hierarchy. Email addresses were collected to create the mailing list, which subsequently was created on Google Groups and has seen a significant amount of activity.

Conclusion

The SC'09 conference was another highly successful showcase of the newest and best technology in the HPC field that I could survey and bring back for BlueFern and my HPC colleagues. It also provided a nexus for meetings (eg: the HPCWorld meeting) between other HPC institutions and BlueFern, who I was fortunate enough to represent. I would be happy to further discuss anything from this report with the reader; my Email address is Tony.Dale@canterbury.ac.nz.

Acknowledgements

I'd like to thank the KAREN travel fund for enabling me to travel to SC'09, and BlueFern for supporting my travel with conference leave. I'd also like to thank Peter McKay and IBM for organizing two very interesting IBM presentations on Blue Gene and Power7, which sadly I can't describe because they are covered by non-disclosure.