BGW Days – Newsletter

BGW Days is a joint project between the Argonne National Laboratory Blue Gene Consortium and IBM TJ Watson Research Center.

The objective of the program was to provide Consortium members an opportunity to scale applications and Computer Science Projects across the 20 rack, 100Tflop+ system at TJ Watson.

- Demonstrate capabilities of Blue Gene architecture
- Create New Science

The program began in October of 2005 and there have been a total of 4 sessions:

- October 2005
- March 2006
- July 2006
- October 2006

The next session is now being planned for March 2007.

Since its inception, we have had a total of 18 consortium institutions participate in the 17 Application projects and 8 Computer Science projects.

October 2005	
Applications	Computer Science
Oak Ridge National Laboratory (ORNL)	Argonne National Laboratory (2)
Argonne National Laboratory (ANL) & University of Chicago	Japan Science and Technology Agency (JSTA)
Purdue University	
March 2006	
Applications	Computer Science
Allied Engineering	JSTA
National Center for Atmospheric Research (NCAR)	Louisiana State University
University of Chicago	
Edinburgh (EPCC)	
July 2006	
Applications	Computer Science
ANL	Lawrence Berkeley National Laboratory (LLNL) & Princeton Plasma Physics Laboratory (PPPL)
ANL, ORNL, Fermi National Laboratory	

(FNL) & Brookhaven National Laboratory (BNL)	
PPPL	
Georgia Institute of Technology	
Gene Network Science, Cornell University & Purdue	
October 2006	
Applications	Computer Science
KTH (Stockholm Sweden)	University of Illinois at Urbana Champaign & LLNL
UC Berkeley	ANL
ORNL	SDSC
NCAR	
SDSC	

For proposals that were successful at scaling, IBM offered the opportunity to apply for additional
time on the Watson Blue Gene System to produce significant science. The following Institutions
have been awarded time through this mechanism:

Argonne National Laboratory
National Center for Atmospheric Research
Princeton Plasma Physics Laboratory

Program Highlights:

Allied Engineering

Allied Engineering used the BGW Days allocation to create and refine their Gordon Bell Proposal, which
was a Gordon Bell finalist.

Edinburgh – EPCC:

The Soft Matter and Statistical Physics Group at the University of Edinburgh, together with Kevin
Stratford of Edinburgh Parallel Computing Centre (EPCC), have developed a code to study 'complex fluid
problems using the lattice Boltzmann method. The method allows this type of problem, which include
particle suspensions and fluid mixtures, to be handled efficiently and effectively in parallel. The

consortium's BGW Day in March of 2006, allowed us to extend the studies performed on the single-rack BG/L system at EPCC to the larger T.J. Watson system. We aimed to check that studies on very large systems of hydro dynamically interacting particles and mixtures were feasible with the code. Despite a slightly curtailed session owing to time differences between the US and UK, the tests performed showed excellent scaling to 8 racks of the BGW system. Having not run anywhere near 16,384 MPI tasks before, the day proved and enjoyable worthwhile experience!

NCAR - Scaling Community Climate System Model Components on Blue Gene

John Dennis of NCAR has used the access to the Blue Gene Watson (BGW) system to prepare the Community Climate System Model (CCSM) for the upcoming availability of Petascale systems. CCSM is comprised of five component models, of which two have been successfully scaled to 32,000 processors during BGW days. In March, we scaled the ocean component of CCSM, the Parallel Ocean Program (POP) to 32,000 processors, while the Community Ice Model (CICE) was tested on 32,000 processors in October. The results of participation in both BGW days events indicated that POP and CICE at 0.1° resolution can achieve simulation rates necessary for production climate runs on 32,000 processors. A 0.1° resolution simulation represents 100 times the computational cost of our current 1° production simulations. Our longer-term goal is to use very large processor counts to significantly increase the resolution of our climate simulations. Access to Blue Gene Watson is invaluable because it provides the opportunity to test and prepare our applications to use 100-200 times the number of processors they currently utilize.

ANL – PVFS

We installed PVFS on 33 servers at Watson and conducted some scalability benchmarks with up to 4 rows (16k nodes). Running on BGW gave us a chance to verify our design goal that PVFS be able to scale to tens of thousands of clients. We achieved peak read performance of 2.8 GB/sec. Writes topped out at 400 MB/sec. We were happy with the results, especially given how quickly we were able to install PVFS and start running experiments. These results provided a baseline upon which we could compare future tuning and optimization efforts on Blue Gene systems.

KTH

The computational architecture and function of the cerebral cortex is still unknown. However, there is an ongoing frontal assault on discovering the principles of neocortical information processing in neuroscience in general, and in three ambitious projects---Blue Brain, DAISY and

FACETS---in particular. Furthermore, an exciting new development in electron microscopy (Denk 2004) holds the promise of scanning an entire cortical column at the nanoscale level. In this development, large-scale modeling of neural circuits is an increasingly important tool, both for the evaluation of hypotheses of cortical function and as a platform for integrating experimental data.

During the autumn of 2005, we performed a series of optimizations of our parallel neuron simulator SPLIT and a scalability study at the Rochester Deep Computing Capacity on Demand Center, IBM, Rochester, MN, USA (Djurfeldt 2005). During the BGW day's event at Watson, we could conclude that the current version of the SPLIT simulator is capable of utilizing the resources of 8 BG/L racks to simulate a large neural system with reasonable efficiency. Scaling was linear up to 4096 processors but a 23% overhead seemed to be introduced for 8192 processors. We can use data from these runs as a starting point for further optimizations of the SPLIT simulator.

We could also show that it is now possible to simulate models of sizeable fractions of brains of small mammals without large sacrifices of model detail to computation speed. The largest simulation comprised 22 million neurons and 11 billion synapses, corresponding to a cortical surface area of approx. 16 cm^2. This is the largest simulations of this type ever performed. An exciting finding was that the cortical attractor dynamics did not break down due to axonal propagation delays. In fact, the switch to a new memory state was completed in less than 50 ms.

ANL – KEK5000

Nek5000 is a high-order spectral element code used for the simulation of incompressible Navier-Stokes and MHD in general three-dimensional domains that has been specifically developed for deployment on large-scale distributed memory parallel architectures such as the Blue Gene architecture developed by IBM. Early access to the BGW platform at IBM Watson Research center enabled us to test the scalability of every aspect of Nek5000, not just the central kernels. Prior to running on BGW, Nek5000 had been extensively tested on the P=2048 processor BGL at Argonne and on thousands of processors on other machines such as Pittsburgh Supercomputer Center's Alpha-based Lemieux cluster and Sandia's Intel-based ASCI Red. In initial runs on the first BGW Day, we discovered that the computational kernels of Nek5000-dominated by multilevel Schwarz-preconditioned conjugate gradient solvers-performed at scale, as expected. A key kernel in the original computational start-up phase, however, did not scale beyond P=8K and required rewriting. The start-up phase involves a topological discovery process requiring an all-to-all communication and has been rewritten to scale as ~ n/P for an n-point computation.

After the successful demonstration of scalability during BGW Day, our proposal to use BGW to study the magneto-rotational instability (MRI) was awarded several million node hours on BGW by IBM. In astrophysics, the MRI has been proposed as a critical mechanism in the initiation of turbulence in accretion disks. Turbulence provides the means for matter orbiting a compact object to lose angular momentum and fall inward, thereby releasing enormous amounts of

energy in the process. Indeed, the in-fall of matter onto black holes and neutron stars powers some of the most energetic phenomena in the universe. Fortunately, the essential physics of the instability rely only on classic MHD and are thus accessible to laboratory experiments using liquid metals. Our numerical simulations are similar in configuration to liquid metal experiments being conducted by Profs. Goodman and Ji, and co-workers at the Princeton Plasma Physics Laboratory. The experiments and simulations naturally complement one another. High Reynolds numbers can be realized in the experiments, but not high magnetic Reynolds numbers, because of the very small ratio of the viscosity to the magnetic diffusivity characteristic of liquid metals. While it is difficult to simulate at laboratory Reynolds numbers, it is nonetheless possible to achieve significantly higher magnetic Reynolds numbers. Our large scale simulations on BGW have confirmed the presence of an MRI in the laboratory configuration with an externally imposed magnetic field. Moreover, preliminary analysis indicates that the turbulent flows driven by the MRI can actually regenerate the magnetic field thereby forming a self consistent nonlinear system that can maintain the turbulence indefinitely even when the external field is removed. This is the first demonstration of dynamo action driven by magneto-rotational turbulence in full cylindrical geometry.
In summary, access to a machine of the scale of BGW allowed us, for the first time, to realize a full scenario for the MRI process.

To see a full list of results, go to: http://www-fp.mcs.anl.gov/bgconsortium/pastresults.htm