8th. World Congress on Computational Mechanics (WCCM8)
5th. European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS 2008)
June 30 – July 5, 2008
Venice, Italy

Towards Petascale Computing with Computational Fluid Dynamics codes on IBM's Blue Gene/P

***Ananthanarayanan Sugavanam[1], Pascal Vezolle [2] and Stephen R. Behling [3]**

[1] IBM Deep Computing, 2455 South Rd., MD P925, Poughkeepsie, NY 12601, USA, suga@us.ibm.com

[2] IBM Products & Solutions Support Center, Rue de la Vieille Poste, 34006 Montpellier France , vezolle@fr.ibm.com

[3] IBM Deep Computing, 360 3-rd Ave. South, MD 55402, Minneapolis, MN 5402, USA, sbehling@us.ibm.com

Key Words: *Petascale, Massively Parallel, Mixed Programming (shared memory/MPI)*

ABSTRACT

The IBM Blue Gene is an ultra scalable and affordable system that enables CFD scientists/engineers develop new insights in to physics of petascale CFD applications. The full paper will discuss the details of CFD algorithms that will map better to the Blue Gene, some recommendations in porting, and tuning communication performance for the Blue Gene systems. Several CFD scaling results and experiences on both research, and industrial CFD codes will be presented.

The first generation of the IBM Blue Gene supercomputers, Blue Gene/L (that had 2 CPUs per processor chip with no shared memory support) was developed to address massively parallel applications, while minimizing cooling/power, cost and space requirements through the use of innovative technologies in low-power processing, embedded DRAM memory chips, system-in-a-chip, advanced power, packaging, and cooling. Its successor Blue Gene/P enhanced many of these attributes, and further added computing capability by including four shared memory processors per processor chip, increasing the memory per node, and improving the communication subsystem.

A single rack of Blue Gene/P has 1024 nodes, and each node has 4 PowerPC 450 CPUs, running at 850 MHz. The nodes are interconnected through five networks: a 3-dimensional torus network, a global collective network, a global barrier, an interrupt network, and a gigabit network. Each node can support 2 GB of memory. The system consists of compute nodes, and separate I/O nodes with each I/O node supporting several compute nodes. 72 Blue Gene/P racks will deliver a peak performance of one PetaFLOPS.

The application developers on the Blue Gene/P can use the familiar OpenMP or multi-threads programming to exploit the shared memory in a node, and an MPI interface between the nodes. Based on real experiences and results, the paper will discuss the potential benefits of some models of the mixed programming (shared memory/MPI) to surpass the barriers to scalability.

Massively parallel computing and Computational Fluid Dynamics

Computational Fluid Dynamics applies to a wide range of problems in research, and industry, and these problems are becoming bigger due to requirements of higher grid resolution, turbulence models, and detailed physics. It is not unusual to see a billion cell CFD problems in the aerospace, automotive, and nuclear industry. The research community has been trying to solve many fundamental problems in fluid mechanics (turbulence, combustion, flame structure, mixing) with billions of grid points already. These computations require massively parallel systems such as the Blue Gene/L, and the Blue Gene/P.

The primary challenges with large scale CFD simulations on the Blue Gene systems include domain decomposition of unstructured meshes, communication and computation load balancing, memory considerations, handling massive I/O associated with these simulations, and providing a mechanism for visualization of the results.

With IBM's GPFS file system, and using MPI-IO (or parallel IO libraries like HDF5), the massively parallel I/O can be addressed. Blue Gene with the IBM Deep Computing Visualization system provides support for parallel and remote visualization. We will outline an IBM solution for interactive, parallel and remote visualization using a CFD Open Source code.

Examples

Several research CFD codes such as Livermore's sPPM that simulates shock physics, DNS codes that simulate turbulence, large eddy simulation codes for combustion, and several astrophysics CFD codes scale to thousands processors on the Blue Gene systems. Several industrial CFD codes in aerospace, automotive, and nuclear energy industry have been ported to the Blue Gene system, and they scale to several hundreds of processors. Figure 1 shows the scaling performance of the NASA Overflow code on the Blue Gene systems (Blue Gene/L and Blue Gene/P).

The paper will extend these CFD scaling results to more than ten thousand processors as well as performance comparisons between pure MPI and mixed OpenMP/MPI implementations on Blue Gene/P.

Fig 1. NASA Overflow-2 code scaling on the Blue Gene Systems