IBM Cluster at NERSC Shows High Productivity

William T. C. Kramer,* NERSC, Lawrence Berkeley National Laboratory

Summary

Early users of NERSC's new IBM cluster computer, Bassi, reported that codes ran from 3 to 10 times faster on the new system than on NERSC's other IBM supercomputer, Seaborg, leading one tester to call the system the "best machine I have seen."

On Monday, Jan. 9, 2006, the National Energy Research Scientific Computing Center (NERSC) opened production use on an IBM cluster with 888 processors for parallel computing. During the acceptance testing, users reported that codes ran from 3 to 10 times faster on the new cluster, "Bassi," than on NERSC's other IBM supercomputer, Seaborg, leading one tester to call the system the "best machine I have seen."

Bassi, named for eighteenth-century Italian physicist Laura Bassi, is an IBM p575 POWER5 system, and each processor has a theoretical peak performance of 7.6 Gflop/s. The processors are distributed among 111 compute nodes with eight processors per node. Processors on each node have a shared memory pool of 32 Gbytes. A Bassi node is an example of a shared memory processor, or SMP.

The compute nodes are connected to each other with a high-bandwidth, low-latency switching network. Bassi's network switch is the IBM "Federation" HPS switch, which is connected to a two-link network adapter on each node. The disk storage system is a distributed, parallel I/O system called GPFS. Additional nodes serve exclusively as GPFS servers.

Figure 1. Bassi is a 976-processor IBM p575 POWER5 system with 111 compute nodes (8 processors per node) available for scientific computing.

Robert Duke, of the University of North Carolina, Chapel Hill, is the author of PMEMD, which is the parallel workhorse in modern versions of the popular chemistry code AMBER. Not only is it widely used for molecular dynamics simulations, PMEMD is also part of NERSC's benchmark suite and used for SC05 research activities. Duke has worked with NERSC's David Skinner to port and improve the performance of PMEMD on NERSC systems, including the two newest clusters, Bassi and Jacquard.

"I have to say that both of these machines are really nothing short of fabulous," Duke wrote to Skinner. "While Jacquard is

* 510-486-7577, WTKramer@lbl.gov

perhaps the best performing commodity cluster I have seen, Bassi is the best machine I have seen, period."

One of the early users during the acceptance testing was the combustion simulation team led by Jackie Chen of Sandia National Laboratories. Chen's group had been awarded 2.5 million hours on Seaborg under DOE's INCITE program. Using NERSC's computational resources, Chen's group was able to run the first three-dimensional direct numerical simulation (DNS) of a turbulent nonpremixed $H_2/CO/N_2$-air flame with detailed chemistry. Accessing Bassi during acceptance testing, the group ran their code for an additional 1.5 million equivalent hours on the new cluster.

"We had a major success enabled by Bassi — the successful completion of our INCITE project to perform direct numerical simulation of a turbulent nonpremixed CO/H_2 jet flame with detailed chemistry," recounted Evatt Hawkes, a member of Chen's group. "Our project required a very long stretch of using a large fraction of Bassi processors — 512 processors for essentially an entire month. During this period we experienced only a few minor problems, which is exceptional for a pre-production machine, and enabled us to complete our project against a tight deadline. We were very impressed with the reliability of the machine."

Hawkes noted that their code also ported quickly to Bassi, starting with a code already ported to Seaborg's architecture.

"Bassi performs very well for our code — with Bassi's faster processors we were able to run on far fewer processors (512 on Bassi as opposed to 4096 on Seaborg) and still complete the simulations more rapidly," Hawkes wrote. "Based on scalar tests, it is approximately seven times faster than Seaborg, and one-and-a-half times faster than a 2.0 GHz Opteron processor. Also, the parallel efficiency is very good. In a weak scaling test, we obtain approximately 78 percent parallel efficiency using 768 processors, compared with about 70 percent on Seaborg."

The machine is named in honor of Laura Bassi, a noted Newtonian physicist of the eighteenth century. Bassi was appointed professor of anatomy at the University of Bologna in 1731, and has been cited as the first woman to officially teach at a European university. She was elected to the Academy of the Institute for Sciences in 1732, and was given the Chair of Philosophy at the University of Bologna in 1733. In 1738, she married her colleague Dr. Giuseppe Veratti. They had 12 children.

While raising her family, she successfully petitioned for wider responsibilities and a higher salary to cover the cost of equipment for physical and electrical experiments. She continued her life-long interest in physics, lecturing from her home while her children were small, then returning to the university at age 65 as a professor of experimental physics in 1776. She died on Feb. 20, 1778, at age 66.

For further information on this subject contact:
Barbara Helland, Program Manager
Mathematical, Information, and Computational Sciences Division
Office of Advanced Scientific Computing Research
Phone: 301-903-3127
Barbara.Helland@science.doe.gov