"A Performance Comparison of Linux and a Lightweight Kernel on Supercomputing Hardware"

Ron Brightwell, Sandia National Laboratories

Summary

One of the goals of the Lightweight Kernel Research Project at Sandia National Laboratories is to quantify the benefits of lightweight compute node operating systems relative to more heavyweight, full-featured operating systems, such as Linux, on large-scale parallel computing platforms. Lightweight compute node operating systems have been an important factor in applications achieving unprecedented performance and scaling on several large-scale platforms. In 2003, we completed an initial comparison of the Cougar lightweight kernel and Linux on ASCI Red hardware. We are continuing to gather results from this comparison, and we are developing a strategy for implementing future lightweight kernels suitable for running on commodity hardware clusters. We believe such operating systems will significantly increase the performance, scalability, reliability, and usability of large-scale parallel computing platforms.

Linux is currently the preferred operating system of choice for clusters of commodity-based hardware used for scientific parallel computing. In addition to its use on small- and medium-sized clusters, Linux is making in-roads on large-scale machines. In particular, while previous large-scale systems, such as the 9000 processor Intel Teraflops machine installed at Sandia National Laboratories uses a custom lightweight operating system on its compute nodes, many new large-scale machines are using variations of Linux.

Two forthcoming machines, the IBM Blue Gene/L and Sandia/Cray Red Storm machines are notable exceptions. Both of these systems plan to use a specialized lightweight operating system on compute nodes instead of a full-featured, commodity operating system. Red Storm is using a lightweight kernel designed and developed by Sandia, and IBM is using a custom micro-kernel developed by IBM Research.

Previous generation machines that have used lightweight operating systems, such as the Intel Teraflops and Cray T3E machines, have been able to demonstrated unprecedented performance and scaling on real scientific applications. Other large-scale machines with full-featured operating systems have not been able to achieve this same level of effectiveness. However, differences in hardware architecture and other important factors make it difficult to pinpoint the compute node operating system as the main factor in poor performance and scaling.

One of the goals of the Lightweight Kernel Research project at Sandia National Laboratories is to quantify the impact that the compute node operating system has on

*(505) 845-7601,pundit@sandia.gov

parallel scientific applications in terms of performance, scalability, reliability, and determinism. Our first approach to gathering this data is to compare the Cougar lightweight kernel with Linux on Intel Teraflops hardware. Since Cougar is the native compute node operating system on this machine, we needed to port of the Linux operating system to this platform. We have completed a port of the Linux 2.4 kernel to this platform and have also implemented a network driver that allows Linux to communicate using the high-performance network. We gathered initial performance data from several benchmarks and one of Sandia's important application codes.

The following figures show the results of two of the NAS Parallel Benchmarks (version 2.4), CG and IS. For each benchmark, the results are presented in millions of operations per second per processor. For CG, Linux slightly outperforms Cougar on a small number of nodes where communication is a less significant factor. Eventually, scalability issues override the difference in single-node performance and Cougar demonstrates significantly better performance.

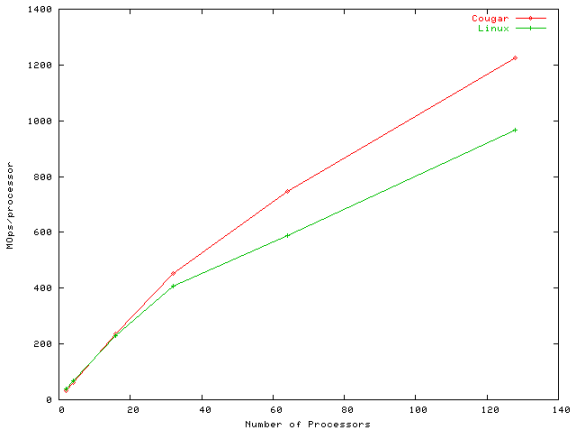

Figure 1: CG Results

The IS results illustrate that Cougar significantly outperforms Linux at small numbers of nodes. This benchmark highlights two of the advantages of Cougar: physically contiguous virtual memory and advanced collective operations.

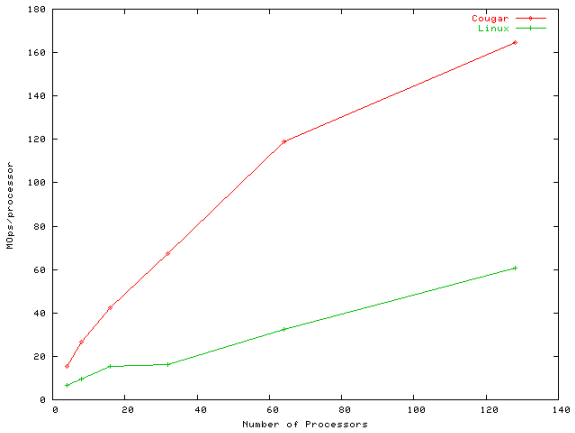

Figure 2: IS Results

While these results are interesting, there are still several issues that need to be addressed in order to perform a true direct comparison. We are continuing to develop our Linux implementation and the network driver to try to better isolate and quantify the impact of the operating system. We also expect to be able to provide more performance results on larger numbers of nodes.

We believe that compute node operating systems are an important factor in the performance and scalability of parallel scientific applications on large-scale machines. The goal of the Lightweight Kernel Project at Sandia National Laboratories is to research and develop scalable operating systems technologies that enable scientific parallel applications to achieve greater levels of performance and scalability.

For further information on this subject contact:
Dr. Fred Johnson, Program Manager
Mathematical, Information, and Computational Sciences Division
Office of Advanced Scientific Computing Research
Phone: 301-903-3601
fjohnson@er.doe.gov