SPEC HPC2002: The Next High-Performance Computer Benchmark

Extended Abstract

Rudolf Eigenmann[1], Greg Gaertner[2], Wesley Jones[3],
Hideki Saito[4], and Brian Whitney[4]

[1]Purdue University
[2]Compaq Computer Corporation
[3]Silicon Graphics Incorporated
[4]Intel Corporation
[5]Sun Microsystems Incorporated

SPEC High-Performance Group

The High-Performance Group of the Standard Performance Evaluation Corporation (SPEC/HPG) [1] is developing a next release of its high-performance computer benchmark suite. The current suite, SPEChpc96 [2] is expected to be replaced by the SPEC HPC2002 suite in the second quarter of the year 2002. HPC2002 represents a minor update to SPEC's current high-performance computing benchmarks. In the last subsection of this document we describe SPEC's longer-term plans.

SPEC/HPG currently develops and maintains two benchmark suites, SPEC OMP2001 [3] and SPEC HPC2002. The OMP2001 suite is targeted at shared-memory platforms that support an OpenMP programming model. By contrast, the HPC2002 makes no restrictions about the target platform. In fact, facilitating cross-platform comparisons was an important motivation that had lead to SPEC's development of the HPC series of benchmark suites.

Compared to the previous suite, SPEC HPC2002 will provide the following changes.

- The current climate modeling benchmark will be replaced by the WRF weather modeling code.
- All HPC2002 benchmarks will include both an MPI and an OpenMP parallel version.
- The suite will make use of the standard *SPEC run tools* for compiling and running the codes, making it much easier to execute the benchmarks.
- New data sets will be defined, representing the small, medium, large, and extra-large benchmark input sets.

The following paragraphs describe these changes in more detail.

Benchmark Applications in SPEC HPC2002

SPEC HPC2002 will include three full-size applications in the areas of seismic processing, computational chemistry, and climate modeling, respectively. The

number of codes is the same as it was in SPEChpc96. However, the climate modeling code MM5 has been replaced by WRF. The other two applications are essentially the same as in SPEChpc96. They include SPECseis and SPECchem. The following paragraphs give a brief overview of the SPEC HPC2002 codes.

SPECclimate, the climate modeling code of the HPC96 suite will be replaced by the WRF weather model. The WRF weather model is part of a next-generation mesoscale forecast and assimilation modeling system. The system has a goal of advancing both the understanding and prediction of important mesoscale precipitation systems, and promote closer ties between the research and operational forecasting communities. The model is being developed as a collaborative effort among several NSF, Department of Commerce, and DoD sponsored institutions, together with the participation of a number of university scientists. Initially, two data sets will be used for the benchmark. Both data sets represent a 24-hour forecast of the weather over the continental United States. One is run at a resolution of 22km, while the larger data set is run at a resolution of 12km. WRF is written in C and Fortran90. It includes 143,000 lines of code.

SPECseis was developed by ARCO beginning in 1995 to gain an accurate measure of performance of computing systems as it relates to the seismic processing industry for procurement of new computing resources. It consists of a modeling phase which generates synthetic seismic traces for any size of data set, with a flexibility in the geometry of shots and receivers, ground structures, varying lateral velocity, and many other options. A subsequent phase stacks the traces into common midpoint stacks. There are two imaging phases which produce the valuable output seismologists use to locate resources of oil. The first of the two imaging phases is a Fourier method which is very efficient but which does not take into account variations in the velocity profile. Yet, it is widely used and remains the basis of many methods for acoustic imaging. The second imaging technique is a much slower finite-difference method, which can handle variations in the lateral velocity. This technique is used in many seismic migration codes today. The current SPECseis is missing Kirkoff and pre-stack migration techniques. SPECSeis contains some 20,000 lines of Fortran and C code.

SPECchem is used to simulate molecules ab initio, at the quantum level, and optimize atomic positions. It is a research interest under the name of GAMESS at the Gordon Research Group of Iowa State University and is of interest to the pharmaceutical industry. Like SPECseis, SPECchem is often used to exhibit performance of high-performance systems among the computer vendors. Portions of SPECchem codes date back to 1984. It comes with many built-in functionalities, such as various field molecular wave-functions, certain energy corrections for some of the wave-functions, and simulation of several different phenomena. Depending on what wave-functions you choose, SPECchem has the option to output energy gradients of these functions, find saddle points of the potential energy, compute the vibrational frequencies and IR intensities, and more. SPECchem contains over 100,000 lines of code written in Fortran and C.

MPI and OpenMP Parallelism

SPEChpc96 did not include shared-memory parallel versions of the benchmarks. Although the benchmark runrules allowed code modifications, so that shared-memory implementations could be created, developing such code versions was non-trivial. By contrast, SPEC HPC2002 includes OpenMP-based implementations, which will facilitate this task directly. The SPECseis benchmark can be made in either an MPI message passing or an OpenMP shared-memory version. Both versions exploit the same parallelism in the application. The other two benchmarks can also be made as combined MPI/OpenMP codes, with the two models exploiting different levels of parallelism.

Benchmark Make and Run Tools

A significant challenge, when obtaining SPEChpc96 benchmark numbers, was the process of making and executing the codes. The different benchmarks included separate make and run environments, which had to be learned individually by the benchmarker. By contrast, the HPC2002 suite will make use of the *SPEC run tools*, which have already proven their value in several other SPEC benchmark suites. These run tools facilitate the process of compiling, executing, validating, and even generating benchmark reports via an easy command line interface. We expect this change will make it significantly easier to generate benchmark results.

Data Sets

Today's high-performance computer systems are significantly more resourceful than at the time of release of SPEChpc96. Therefore, SPEC HPC2002 will include larger data sets. All three benchmark applications include four data sets, ranging from small to extra-large. A small data set will consume approximately an hour of runtime on a single CPU of a state-of-the-art computer. The extra-large data set will run two to three orders of magnitude longer. The memory requirement will follow from the criterion that each data set will represent a realistic workload for the problem being solved.

Future Development of SPEC HPC Benchmarks

While the described updates to the current SPEC high-performance benchmarks warranted the release of SPEC HPC2002 at this time, the SPEC High-Performance Group is actively working towards a next full release of a new HPC suite. The suite will include newer versions of the seismic processing and computational chemistry applications. It will also include applications in a number of additional application areas, such as computational fluid dynamics and computational biology.

References

1. Standard Performance Evaluation Corporation, "SPEC High-Performance Group," http://www.spec.org/hpg/.
2. Rudolf Eigenmann and Siamak Hassanzadeh, "Benchmarking with real industrial applications: The SPEC High-Performance Group," *IEEE Computational Science & Engineering*, vol. 3, no. 1, pp. 18–23, Spring 1996.
3. Vishal Aslot, Max Domeika, Rudolf Eigenmann, Greg Gaertner, Wesley B. Jones, and Bodo Parady, "SPEC OMP: A new benchmark suite for measuring parallel computer performance," in *OpenMP Shared-Memory Parallel Programming*, Springer Verlag, Heidelberg, Germany, July 2001, Lecture Notes in Computer Science #2104, pp. 1–10.