Porting Some Key Caltech & JPL Applications to a PS3 Cluster - A Wild Ride

Ed Upchurch (Caltech/JPL), Paul Springer (JPL), Mark Stalzer (Caltech),
Sean Mauch (Caltech), John McCorquodale (Caltech), Jan Lindheim (Caltech).
etu@cacr.caltech.edu, stalzer@cacr.caltech.edu, pls@jpl.nasa.gov, sean@cacr.caltech.edu, mcq@cacr.caltech.edu, lindheim@cacr.caltech.edu

Introduction

Last October Caltech's Center for Advanced Computing Research (CACR) purchased 13 PS3's (a lucky number) to build a high performance parallel algorithm testbed for under $10K with a peak potential of a little over 2 TFLOPS single precision. Well, as a kid in Texas I remember my uncle always saying "A cheap horse is hard to ride". It has been interesting so far and we will share "the ride" with you.

Figure 1: CACR's PS3 Cluster

Approach

The PS3 cluster offers a rich test environment of heterogeneous multi-core nodes with MPI clustering plus the promise of low cost high performance and low power/weight/volume. Low cost high performance is attractive for exploring ground based applications such as compute intensive SSA and QMC. High performance low power/weight/volume is of interest for space based applications where greater autonomy for deep space missions and downlink data volumes could be significantly reduced by on board processing.

Stochastic Simulation Algorithm (SSA)

The small number of reactant molecules in living cells results in chemical behavior that is discrete and stochastic rather than continuous and deterministic. Discrete stochastic models such as SSA [1] better capture this but their computational complexity becomes overwhelming for practical biological systems. One of the authors (Mauch) has made extensive progress in accelerating these simulations.

ROI_PAC

Repeat Orbit Interferometry PACkage (ROI_PAC) is a library of FORTRAN code to processes SAR images to create InSAR (interferogram) images [2]. Significant reduction in downlink data in possible by onboard processing

SVM

Support Vector Machines (SVM) [3] FFTs and the overlap-and-add technique can be used to efficiently apply SVMs to sensor data in resource-constrained environments (crater recognition, dust devils on Mars). Speedup would allow greater autonomy and opportunistic science.

Very Large Discrete FFT

Combined Ultraviolet, Visible & InfraRed Spectrometer (CUVIRS) Very Large Discrete FFT [4] associated with TES (Tropospheric Emission Spectrometer) is an infrared Fourier Transform Spectrometer designed to make global measurements of tropospheric ozone and its chemical precursors. Order of magnitude reduction in downlink data could be achieved by doing large FFTs in real time on board.

Quantum Monte Carlo (QMC)

Advances have been made in recent years in the development and application of QMC methods to quantum chemistry [5]. QMC is still in its infancy compared to other *ab initio* methods, but the high accuracy and simplicity of the method hold bright promise for the future.

Preliminary Observations

Software releases of YDL & SDK have caused coding pain and legacy codes have been written assuming plenty memory available (not true on the PS3). We have achieved 100x speedups over current spacecraft processing.

References

[1] D. Gillespie, "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions", *J. Comput. Phys.* 22:403-434, 1976.

[2] M. Simons, P. A. Rosen, "Treatise on geophysics, interferometric synthetic aperture radar geodesy", Schubert, G. (ed.), Geodesy, Elsevier Press, Vol. 3, pp. 391-446, 2007.

[3] M.C. Burl, W.J. Merline, W. Colwell, E.B. Bierhaus, C.R. Chapman, "Automated detection of craters and other geological features", *Int. Symposium on Artificial Intelligence, Robotics and Automation for Space*, June 2001.

[4] J. Worden, X. Liu, K. Bowman, K. Chance, R. Beer, A. Eldering, M. Gunson, H. Worden, "Improved tropospheric ozone profile retrievals using OMI and TES radiances", *Geophys. Res. Let.*, Vol. 34, 2007.

[5] M. T. Feldmann, J. C. Cummings, D. R. Kent, R. Muller, W. A. Goddard, "Manager-worker-based model for the parallelization of quantum Monte Carlo on heterogeneous and homogeneous networks", *J. Comp. Chem*. 29 (1): 8-16, 2008

PS3 cluster purchased & maintained by Caltech with development work funded in part by CACR, JPL, US Army (as part of Institute for Collaborative Biotechnologies - Systems biology modeling and simulation languages and software)