Cielo CCC 4 Proposal

Principal Investigator:
Name:Luke Shulenburger
Username:lshulen
Work Phone:505-284-3198
Unclassified e-mail Address:lshulen@sandia.gov
Lab/Department/Division/Group:Sandia National Laboratories/ Org 1641

Please provide a one-line unclassified title for this proposal:
Development of quantum Monte Carlo techniques for high fidelity determination of materials properties at extreme conditions.

Milestones supported (if any): none

Goal/Objective:
This research aims to perfect and validate a high precision alternative to traditional electronic structure techniques for predicting the properties of materials at extreme conditions. Diffusion quantum Monte Carlo (DMC) provides a complementary approach to the more widely used Density Functional Theory (DFT) based techniques used to determine the properties of materials under extreme temperature and pressure conditions. Both of these methods employ approximations to solve the quantum mechanical equations describing the electrons in degenerate materials. DFT uses a mapping of the problem to a noninteracting system with an effective potential. This approach has been used successfully in a number of applications, but suffers in accuracy when considering materials where the physics of the electronic interaction differs markedly from the homogeneous electron gas as is the case for many technologically important materials in the laboratory complex such as actinides or transition metal oxides. QMC takes an alternative approach of constraining the topology of the wavefunction while keeping the Hamiltonian governing the electrons unchanged. This results in a very high accuracy method although at the price of being one to two orders of magnitude more expensive for typical problems.

In this project we will continue earlier work towards validating the utility of DMC for simulating a wide variety of materials under extreme conditions. Previous calculations on cielo have shown a high accuracy for the equation of state of a variety of materials from the first three rows of the periodic table, comparable to that of some of the newest and most computationally demanding extension of DFT.

This project will build on that success in two facets. The first will be in calculating the high pressure phase transitions of lithium and two of its compounds, LiH and LiF. In these cases, the relative errors between calculations of the materials at different densities were well controlled in previous work, but to calculate accurate phase transitions far more demanding absolute energies must be obtained. Also as phase transition pressures are exquisitely sensitive to errors in the total energies in ab initio calculations, these calculations will provide rigorous test of the method. The particular choices of materials are driven by two considerations. Firstly, the expense of DMC

necessitates the use of pseudopotentials to approximate the tightly bound core electrons for heavier elements. When considering compounds of light elements such as lithium, reference calculations with all electrons in the valence can be considered directly so that the accuracy of our pseudopotential construction techniques can be judged. Secondly, the lithium compounds are important in high pressure applications such as secondary design and windows for dynamic materials experiments and therefore high quality data exists with which we may compare our calculations.

The second focus of this project will be to extend our equation of state calculations past the first three rows of the periodic table. In doing so, we will provide valuable benchmark data to understand the size of errors that may be made as the method is applied to more strongly correlated materials such as the actinides. Specifically, we aim to calculate the equilibrium density and bulk modulus and cohesive energies of Mo, Ta, Pd, Ag, Au and Sn.

While no milestones are directly supported by this proposal, this technique will be directly applicable to a large portion of the dynamic materials properties objectives of the C2 campaign as well as providing equation of state information to support the AC and C1 activities.

Readiness Justification - Scaling:
Quantum Monte Carlo code qmcpack

The quantum Monte Carlo algorithm is inherently suited to massively parallel applications. In fact, variational Monte Carlo is an intrinsically parallel algorithm exhibiting linear scaling with the number of CPU cores. Diffusion Monte Carlo (DMC), which will comprise the bulk of the computing resources used in this proposal scales only slightly less well. The only inter-node communication necessary for this algorithm is due to the load balancing between different Monte Carlo trajectories, so called walkers, which happens at set intervals during the calculation. The particular code that will be used for the DMC calculations, qmcpack, has been extensively tuned for massively parallel architectures, taking advantage of shared memory parallelism via OpenMP and distributed memory parallelism via MPI. Recently a 64 atom supercell of Si with an interstitial defect was simulated on the Cray-XT5 system (jaguar-pf) at NCCS. In these calculations it was found that nearly 95 percent parallel efficiency was obtained when scaling from 2400 to 216,000 cpu cores, although checkpointing of the calculations at larger scales had the potential to reduce this significantly as shown in Figure 1.

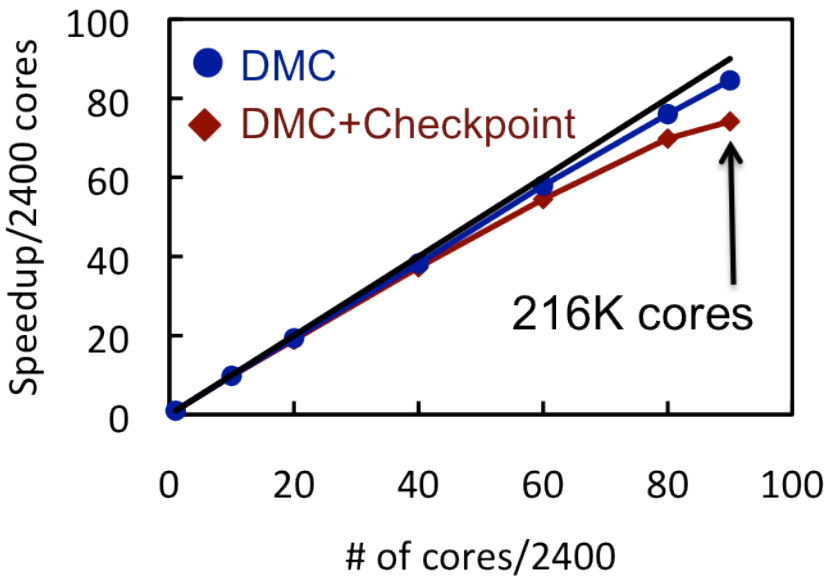

Figure 1: Scaling of a 64 atom calculation of silicon performed on the jaguar-pf. Greater than 95% parallel efficiency was demonstrated when scaling from 2400 to 216,000 CPU cores

Experience:

One of our team members (LS) has previously performed a benchmarking study on the accuracy of quantum Monte Carlo on light elements using DMC on cielo. The resulting data, summarized in figures 2 and 3 allowed for the assessment of the state of the art for the method and showed that without any theoretical advances, its accuracy is better than the most accurate of the new class of DFT functionals.

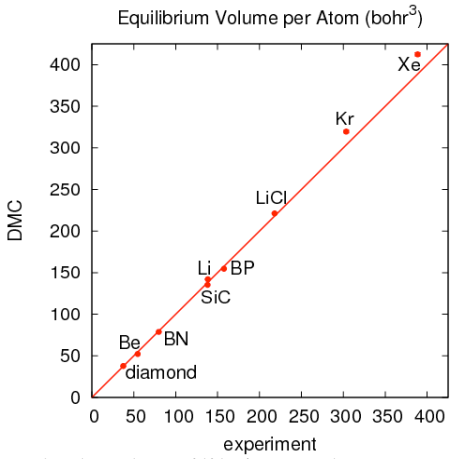

Figure 2: Calculated equilibrium volume per atom of a test set of materials using DMC. The straight line indicates perfect agreement with experiment.

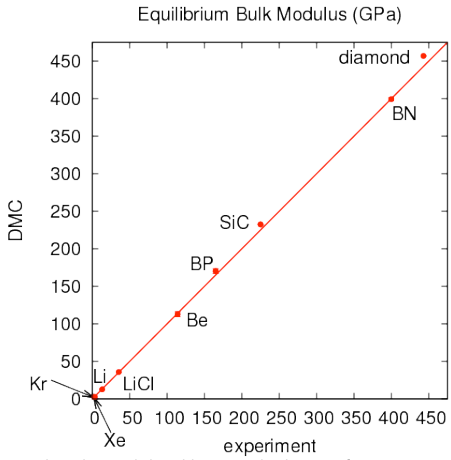

Figure 3: Calculated bulk modulus of a test set of materials using DMC. The straight line indicates agreement with experiment.

C1 Requirements (jobs that use 6375 or more nodes)
Number of jobs in this category that will be run: 0
Size (number of nodes needed): 0
Average run time for jobs in this category: 0
Code used:

C2 Requirements (jobs that use at least 2550 but less than 6374 nodes)
Number of jobs in this category that will be run: 0
Size (number of nodes needed): 0
Average run time for jobs in this category: 0
Code used:

C3 Requirements (jobs that use at least 850 but less than 2549 nodes)
Number of jobs in this category that will be run: 30
Size (number of nodes needed): 1536
Average run time for jobs in this category: 36 hours
Code used: qmcpack

Based on your C-level requirements, please provide the total number of Cielo days' allocation required to fulfill your proposal requests. Base your aggregation calculations on: 8518 compute nodes for 24 hours = 1 Cielo day. Total Cielo days required for this proposal is:
8 days

Data Management:

Based on your C-level requirements, please provide estimates for the following:

How much data do you expect to store in the /scratch file system? (TB) < 5 TB will be necessary as the majority of the necessary data is transient and the results of the calculations are typically on the order of a few megabytes at most.

How much data do you expect to store in the LANL HPSS archive? (TB) 0 TB will be stored on the LANL HPSS. All storage will take place at SNL on the HPSS or at LLNL.

How much data do you expect to transfer across the WAN? (TB)
< 2 TB total data is to be transferred back to SNL and LLNL over the course of the project

For data to be transferred across the WAN: will it be destined for a file system, HPSS, or both?
Data will primarily be destined for HPSS, but some small transfers to file systems for local analysis are expected

Will your application require the use of the read-only “UDSL’ file system? If so, please estimate the size of the data storage required (GB).
UDSL is not necessary

Application Codes and Application Development Teams

Please list the application codes that you will be running on Cielo during the CCC3 campaign.
Codes to be run on cielo are pwscf from the quantum espresso suite and qmcpack

Please list the primary contact for each application development/support team for the applications you plan to run.
Contact for quantum espresso: externally developed, supported locally by Luke Shulenburger
Contact for qmcpack: lead developer is Jeongnim Kim at ORNL, supported locally at SNL by Luke Shulenburger and at LLNL by Miguel Morales

Visualization

What visualization software/tools will you require on Cielo?
None required

List of users authorized to use this CCC bank. Please specify each user's name, login name, e-mail address, and work phone:
Luke Shulenburger, lshulen, lshulen@sandia.gov, 505-284-3198
Miguel Morales, mmorale, moralessilva2@llnl.gov, 925-423-4956
Randy Hood, hood, hood9@llnl.gov, 925-422-5787