Available online at www.prace-ri.eu

**Partnership for Advanced Computing in Europe**

# CASINO for large solid catalyst systems: Configuration numbers and population control

M. Boiarciuc[a]*, P. Hoggan[b]

[a]*IDRIS – CNRS, Orsay, France*
[b]*Pascal Institut, University of Clermont, France*

**Abstract**

The main goal of this project was to port the CASINO code to a Blue Gene/Q system (JUQUEEN) and use it as a test reaction benchmark of hydrogen dissociation on copper surfaces. Some bottlenecks have been identified and investigated in this project.
The main restriction to efficient use of super-computers is efficient use of the RAM, in particular for loading and handling configuration data.

Project ID: 2010PA1186

## 1. Introduction

The test reaction for hydrogen dissociation on copper surfaces is well-documented. Some geometries for the system are available [1].

A very accurate benchmark using Quantum Monte Carlo simulations is required and could be obtained using the CASINO code. This involves handling large input files for the trial wave-function and the initial equilibrium configurations. The code is known to scale linearly up to at least 100,000 cores [2] and it has been shown that the shared memory allows it to run efficiently within the RAM on Blue Gene/P at IDRIS (France).

The unknown aspect is how to best handle such large systems for over 8192 cores and how to be sure the configuration file does not increase in size so much as to saturate the RAM (population explosion) in the early phases of equilibrating the Diffusion Monte Carlo calculation. A fail-safe input flag–trip weight avoids this. The population can be also monitored step-by-step. These precautions are temporary, as the population is more stable after equilibration. This peak in population is largely caused by variance in the trial wave-function.

The aim of this project[**] was to improve the trial wave-function quality and subsequently reach this equilibrium state more efficiently.

* Corresponding author. *E-mail address*: margareta.boiarciuc@idris.fr
** The project was supported by PRACE through the Type C Preparatory Access process.

## 2. Porting the CASINO code to JUQUEEN

CASINO [3] is a code for performing quantum Monte Carlo (QMC) electronic structure calculations for finite and periodic systems. It is a modern, general code capable of treating arbitrary systems (e.g. atoms, molecules, polymers, slabs, crystals and electron phases). The main characteristics of this code are generality, speed, ease-of-use and portability over a wide range of computational hardware.

The 2.8 and 2.10 versions of the CASINO code have been installed on JUQUUEN and an architecture specific file has been generated for this system. This architecture specific file compiles the normal, debug, OpenMP and shared memory (Shm) versions of the code.

Currently users of the CASINO code can easily prepare their own Blue Gene/Q binaries. In order to install the version 2.10 of the CASINO code, an environmental variable CASINO_ARCH has to be declared as following:

**$ export CASINO_ARCH="ibm_bgq"**

The LAPACK library is also needed for the compilation and execution:

**ENVIRONMENT_COMMAND = module load lapack**

In order to increase the default size of the shared memory region, the following environmental variables are used: "BG_SHAREDMEMPOOLSIZE" for a Blue Gene/P system and "BG_SHAREDMEMSIZE" for a Blue Gene/Q system.

Simulations can be started on Blue Gene/Q with the following command:

**RUN_PARALLEL: runjob --np &NPROC& --ranks-per-node 16 --exp -env "SHAREDMEMSIZE=1000" --exe &BINARY&**

It results in runs with optimal scaling in hybrid parallel form on 8192 cores with 4 threads per core.

## 3. Optimization work

The bottleneck of CASINO identified within this project is the trial wave-function which should be compact and have low variance. In practice, for such systems, it is usually a single Slater determinant from a plane-wave DFT code such as ABINIT. An expansion over determinants for several states or mixed plane-wave/atomic orbital input has shown improvements [4] but is laborious.
Tests of a Jastrow optimization have been carried out in order to find a more stable solution satisfying strict optimization criteria. There is a slight improvement regarding the population explosion in the early phases of equilibrating the Diffusion Monte Carlo calculation. Therefore, additional tests have to be done in future work.
Tests on the configuration number have also been carried out. It was determined that, in order to allow for population growth, a maximum of 80% of the available RAM space has to be used.
A single precision numerical wave-function and low-variance expansion of complex plane-waves up to the 1.5 order, which fit RAM and reduce population growth, have been successfully tested.

# 4. Conclusion

The CASINO code (versions 2.8 and 2.10) has been successfully ported to the Blue Gene/Q system (JUQUEEN). Some tests of a Jastrow optimization and configuration number have been carried out in order to investigate how trip-weight choice and trial wave-function variance limitation cure population explosion in DMC equilibration, within the limits of RAM on over 8192 cores. Results are encouraging but additional tests are needed. Especially the RAM saturation by DMC configurations in the case of large systems with variance of the trial wave-function is difficult to reduce.

The code is now able to run efficiently in a production context on Blue Gene/Q systems.

## Acknowledgements

This work was financially supported by the PRACE project funded in part by the EUs 7th Framework Programme (FP7/2007-2013) under grant agreement no. RI-283493.

## References

[1] **C. Diaz, E. Pijper, R.A. Olsen, H.F. Busnengo, D.J. Auerbach and G.J. Kroes, Science, 2009, 326, 832-834.**
[2] **M.J. Gillan, M.D. Towler, D. Alfè, 2011 Psi-k Scientific Highlight of the Month, http://www.tcm.phy.cam.ac.uk/~mdt26/casino2_thingstoread.html**
[3] **R.J. Needs, M.D. Towler, N.D. Drummond and P. López Ríos, J. Phys.: Condensed Matter 22, 023201 (2010).**
[4] **P.E. Hoggan IJQC, 2013, 113, 277-285.**