Speed and Scalability of Materials Science Simulations (CP2K) Enhanced by HECToR dCSE Team

Ben Slater, University College London (UCL)
Ian Bethune, Edinburgh Parallel Computing Centre (EPCC)
HECToR CSE Team, Numerical Algorithms Group Ltd (NAG)

HPC experts from EPCC, working under NAG's Computational Science and Engineering (CSE) support service for HECToR, the UK's national academic supercomputing facility, have optimized the CP2K materials science code allowing researchers to achieve significantly better scalability and reduce runtimes, thus enabling new science recently published in the Proceedings of the National Academy of Sciences (doi: 10.1073/pnas.1001087107).

Dr Ben Slater of UCL, the Principal Investigator and his colleagues have been able to conduct a detailed study of defect segregation in ice as a results of this HECToR CSE work. Commenting on the success of the project, Dr Slater said "*For the published study, it was imperative to use relatively large models of ice slabs in order to capture the detail of the ice surface structure and probe the surface segregation energies of molecular defects. Using the optimized version of CP2K we were able to routinely model cells containing over 1000 atoms. The CSE investment into CP2K was crucial to making the study tractable.*"

HECToR

A Research Councils UK High End Computing service, HECToR is funded for six-years (2007-2013), providing capability supercomputing resources for researchers. A substantial portion of the funding is devoted to the CSE support service provided by NAG, which ensures users have access to appropriate HPC expertise to effectively exploit advanced supercomputers for their science. A critical feature of the CSE support service is the distributed CSE (dCSE) programme which, through lightweight peer review, delivers dedicated performance and scalability projects on specific codes in response to proposals from users. The dCSE programme now consists of nearly 40 focused projects complementing the traditional HPC user support and training provided by NAG.

The dCSE projects completed so far have delivered outstanding examples of the cost savings and new science that can be enabled through dedicated CSE effort. The CP2K project reported here adds to these success stories with a successful code restructuring and performance improvement.

Project Background

The objectives of this dCSE project were to improve the parallel performance of the Density Functional Theory code CP2K by addressing the MPI communications, load balancing and multi-core performance. Thus, simulations with larger numbers of atoms are made feasible through reduced solution times and more cost-efficient utilization of large supercomputers with multicore nodes (such as HECToR). Dr Slater (UCL) was the Principal Investigator on the project working with Dr Joost VandeVondele (Zurich) and Dr Matthew Watkins(UCL). Iain Bethune of EPCC at the University of Edinburgh undertook the 12 person-month project in collaboration with the NAG CSE team and the CP2K users.

Project Results

The Realspace to Planewave transfer routines were optimized, halving the time taken by halo swaps, resulting in an overall 12% speed up. The performance of the Fast Fourier Transform routines was improved by storing reusable data and eliminating redundant MPI collective operations, giving an additional 12% of speedup. The load balancing for non-homogenous systems was improved resulting in speedups of 25% for 128 cores and 18% for 2048 cores. Additional work, implementing OpenMP in the most computation intensive routines yielded dramatic improvements in scalability, scaling up to 8 times more cores than the original version. Dr Slater, who is both the dCSE PI and a major user of CP2K, estimated that CP2K currently consumes around 4 million AUs per month on HECToR, with usage growing. The potential cumulative performance improvement of nearly 50% represents at this rate of usage a possible saving of around £1.5M in the notional cost of research using CP2K over the life of the HECToR service, in addition to cost savings and new science enabled by use of the dCSE optimized CP2K code on other HPC facilities.

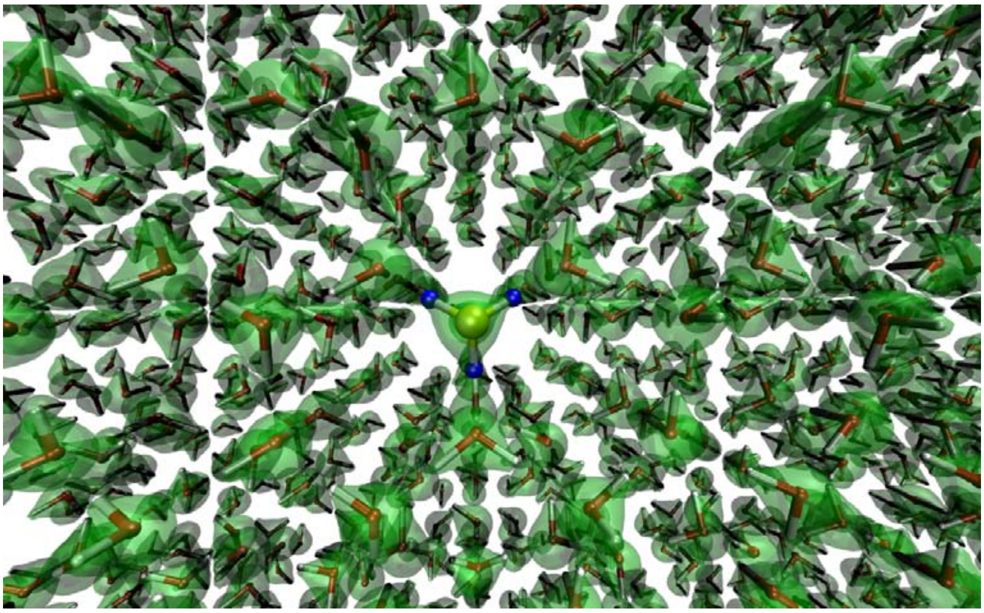
A full technical report can be found at http://www.hector.ac.uk/cse/distributedcse/reports/

About CP2K
CP2K is a freely available density functional theory (DFT) package, licensed under the GPL. CP2K is developed by a number of individuals and groups, and currently numbers over 50,000 lines of Fortran 95 code. The key difference between CP2K and other DFT codes is its implementation of the Quickstep algorithm, which uses a dual basis - atom-centred Gaussian functions to represent the wave-functions, and plane waves/regular grids for the electonic density. Like most other DFT codes, CP2K also makes use of 3D Fast Fourier Transforms (FFTs) to convert from the real space to reciprocal space (plane wave) representations.

About HECToR
HECToR is the UK's national supercomputing service, managed by EPSRC on behalf of the UK Research Councils. Its mission is to support capability science and engineering in UK academia. HECToR's Cray XT supercomputers are located at the University of Edinburgh, managed by EPCC. Computational science and engineering (CSE) applications support, including training and documentation, is provided by NAG Ltd.

HECToR – A Research Councils UK High End Computing Service. **http://www.hector.ac.uk**

For more information contact:
HECToR CSE Team
The Numerical Algorithms Group Ltd, Wilkinson House, Jordan Hill Road, Oxford, OX2 8DR, United Kingdom
Telephone: 01865 511 245 Email: hector-cse@nag.co.uk Web: http://www.hector.ac.uk/cse/



The Hydronium (H_3O^+) defect (highlighted in the centre in yellow) on the ice (0001) surface, obtained with CP2K.