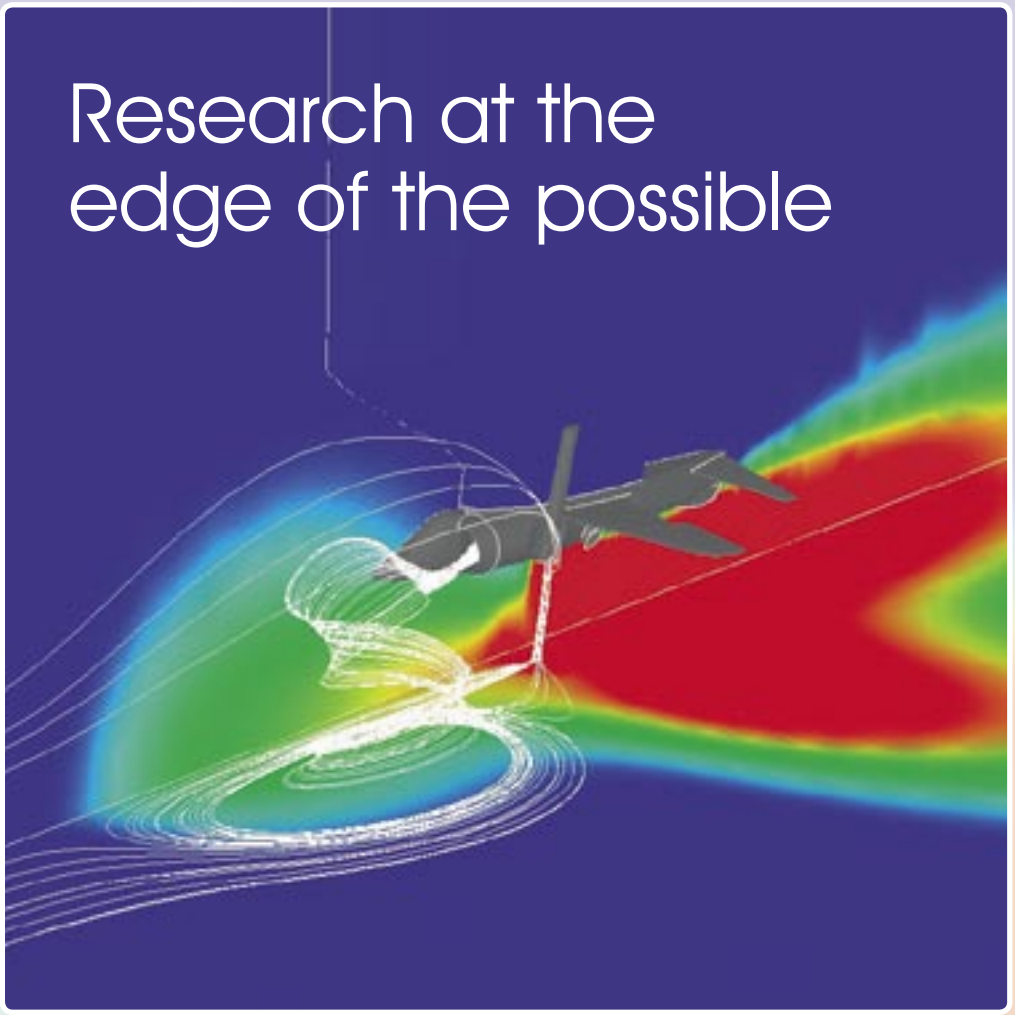
CapabilityComputing

The newsletter of the HPCx community

[ISSUE 4, AUTUMN 2004]



CONTENTS

2 Trends in capability science

3 When simple systems do complex things

5 Lattice Boltzmann algorithms on HPCx

6 Benchmarking molecular dynamics applications

7 UK Applied Aerodynamics Consortium

8 Direct numerical simulation of a neutrally stratified turbulent Ekman layer

9 SIESTA for the HPCx Phase2 system

10 e-Minerals

11 Agulhas rings in a high-resolution ocean model

13 NERC Consortia meeting
Capability usage

14 UK capability science initiatives
Tutorial at SC2004

15 Second HPCx Annual Seminar
ScicomP and SP-XXL in Edinburgh in 2005

16 MSc in High Performance Computing

Editorial

Lorna Smith

It gives me great pleasure to welcome you all to this latest edition of *Capability Computing*. It has been an exciting time for HPCx, with the recent successful upgrade keeping HPCx in the top 20 supercomputers in the world and the most powerful academic supercomputer in Europe.

With this enhanced power, it seemed timely to report on the world-class science being enabled by HPCx. The range of capability science on HPCx is truly diverse and in this issue I have tried to represent this diversity, with articles from users in engineering, physics, chemistry, biology and environmental modelling.

In addition, the HPCx support teams have been working hard to facilitate this science, and you will find a range of highlights from the teams dispersed throughout the newsletter.

With the increased performance of HPCx, allied with the reduced latency and high bandwidth of the new High Performance Switch, I am sure that the UK's computational science community should have an exciting few years to look forward to.

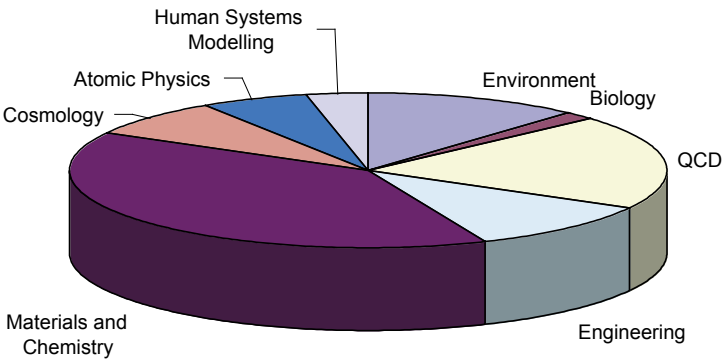
Trends in capability science

Lorna Smith, HPCx Terascaling Team

Capability science – science made possible through HPC resources and expertise – has been the major focus of HPCx from the outset. The UK has historically been strong in this field, with many areas producing world-leading science. The UK's High Performance Computing resources have been a key component of this strength, and will, I believe, continue to play an important role in ensuring this ongoing success.

The breadth and scope of capability science continues to grow, with new fields such as health sciences, bioimaging and disaster simulation joining established fields such as chemistry and cosmology. This in turn is increasing the impact of scientific research obtained from high performance computing resources.

Just after the time of the Scientific Working Group's report *Research Requirements for High Performance Computing* (1992, chaired by Professor Catlow), the UK's national HPC resource, the T3D, had eight 'grand challenge' scientific areas: environmental modelling; biology; quantum chromodynamics (QCD); engineering; materials and chemistry; cosmology; atomic physics and human systems modelling. Taking a snapshot of resource usage on the T3D from April to July 1996 (see below), and comparing this to the current capability usage on HPCx (see Alan Simpson's article on p19) we see an interesting comparison between past and present capability scientific fields.



For example, materials and chemistry continue to be major computational resource users. This reflects their status as world leaders in many disciplines. Atomic and molecular physics also continues to exploit HPC resources within the UK. With particular strengths in multiphoton and electron collision research the community continues to regard computational research as an equal partner to experiment.

The engineering community also utilises a sizeable chunk of the resources. The UK has international standing in computational fluid dynamics, although it is worth noting the growth of interest in HPC from other engineering areas, such as fusion research.

The fields of cosmology and QCD are not major resource users on HPCx; however, both have procured their own HPC resources. In particular, QCD is due to take delivery of a purpose built HPC resource of over 10 teraflops at the end of 2004 (see article on p14).

While not represented on the capability chart, environmental consortia make extensive use of resources below the capability range. High end resources are key to ensuring this community maintains its impressive scientific record.

The absence of the biological community from this chart is also misleading. The trend towards handling 'whole system' problems spans many length scales (e.g. atomic to organism levels) leading to a recognition that HPC is critical to achieving these cutting edge simulations.

To summarise, the impact of HPC on a wide area of science is clear. Together with our traditional HPC users, new and emerging fields will ensure capability science in the UK continues to make an impact on the world stage.

When simple systems do complex things

Simon Bates, The University of Edinburgh

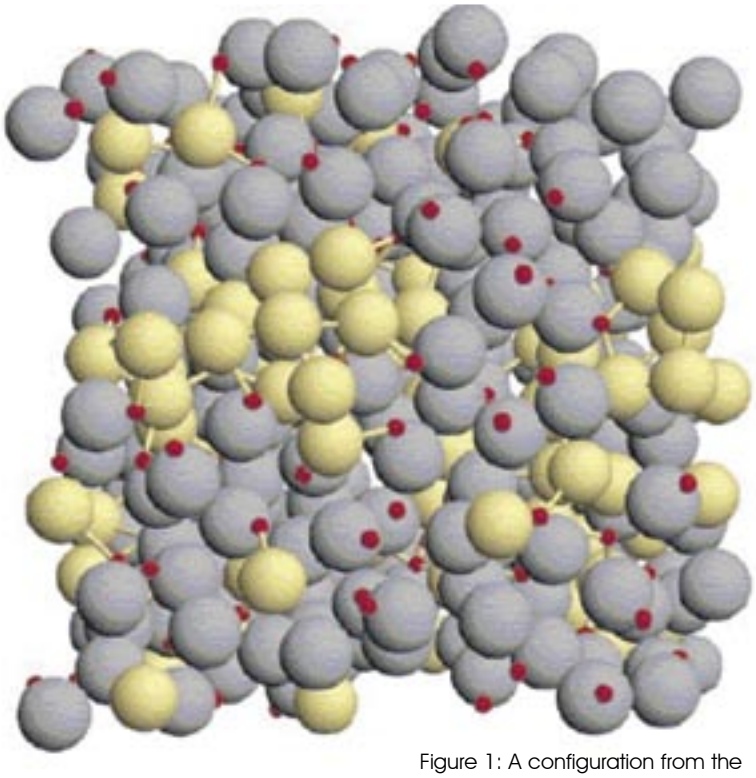


Figure 1: A configuration from the EPSR fit to the ND data illustrating incomplete mixing. The grey spheres denoted hydrophobic methyl groups, the small red spheres the methanol oxygens and the yellow spheres the oxygen atoms of water molecules.

Dr Simon Bates is a member of the School of Physics at the Univeristy of Edinburgh. His research interests focus on investigating the structure and properties of materials on an atomic scale, and area of research at the interface between physics and (physical) chemistry and even molecular biology. He is a member of both the UK Car-Parrinello Consortium (e06) and the Reality Grid Consortium (e10).

Sometimes, even the seemingly simplest of systems can exhibit surprisingly complex and subtle behaviour. This article describes recent computational work within the School of Physics at the University of Edinburgh, looking at the structure and dynamics of simple aqueous alcohols.

We like to think that, as computational scientists, we understand the simple things and can devote our attention, and CPU cycles, to thinking about the more complex problems in science. Sometimes though, even the simplest of systems can exhibit beguiling complexity. Take water as a classic example: the simplest inorganic liquid, the most abundant molecule on Earth and over two-thirds of you and me. It exhibits no less than 41 anomalous properties [1] and some say these anomalies are the basis of its ability to support life. After a long history – over thirty years – of performing computer simulations of water, we are still a long way from being able to say that we understand it. As Guillot has stated in a recent review [2] 'One has a sense of incompletion if one considers that not a water model in the literature is able to reproduce with a great accuracy all the water properties'.

Small surprise then, that simple molecules in solution do strange things as well. Methanol, CH_3OH, is the smallest member of the alcohol family. It is also just about the simplest molecule that displays amphiphilic behaviour; that is, it comprises both hydrophobic (the water-hating CH_3 group) and hydrophilic (the water-loving OH) entities within the same molecule. It exhibits, like all amphiphiles, molecular schizophrenia towards water. Amphiphiles are everywhere in physics, chemistry and biology. Larger amphiphiles self-organise themselves into solution forming spheres (micelles), cylinders (columnar phases) and sheets (lamellae), depending on concentration and temperature. Biologically, amphiphiles are ubiquitous; they comprise the fundamental structural units of cell membranes (phospholipids); DNA forms strong hydrogen-bonds to water and hydrophobic effects are believed to play an important role in protein folding.

It is not immediately obvious that methanol should behave as a 'typical' amphiphile in solution, yet it does. Its solutions exhibit the same thermodynamic non-ideality; it shows a tendency to self-organise in water (as I will discuss) and this organisation is consistent with hydrophobic-like structures where the non-polar hydrophobic groups are in contact. This has led to it being used as a prototypical amphiphilic system, studied using both state-of-the art experimental and computational techniques in a combined research programme at Edinburgh.

In 2002, the experimental group at Edinburgh and RAL published an article suggesting that the anomalous thermodynamics in a methanol-rich solution (methanol mole fraction = 0.7) could possibly be attributed to the micro-heterogeneity that the system exhibited [3]. In other words, despite being fully miscible in all proportions on a macroscopic scale, methanol and water seemed not to want to fully mix on a microscopic scale, and this 'micro-immiscibility' could explain some of the strange thermodynamic properties that had for a long, long time been attributed to the enhanced structure of the water around the hydrophobic group. Their analysis (based on Empirical Potential Structure Refinement (EPSR) of their Neutron Diffraction (ND) results) convincingly showed that the system was far from homogeneously mixed (see Figure 1).

This micro-immiscibility has recently been confirmed by simulation. My group have performed classical molecular dynamics

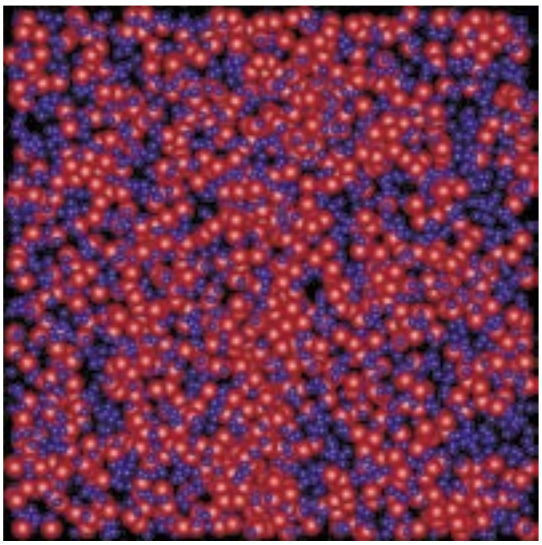


Figure 2: a snapshot of the front face of the MD simulation box for the methanol-rich solution. Water oxygens are red, methyl carbons are blue. The entire box comprises ~20000 atoms. A very large water 'super-cluster', containing over 400 molecules, can be seen running top to bottom of the middle of the box.

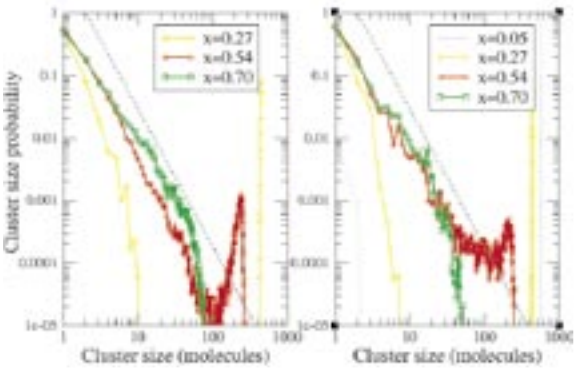


Figure 3: distributions of water cluster sizes, from both experimental data (left) and MD simulations (right) for different mole fractions of methanol (x). The dashed line indicates the theoretical percolation threshold.

simulations using the DL_POLY code that predict the presence of large clusters of water, up to hundreds of atoms. The water molecules break up the structure of pure liquid methanol, which is itself extensively hydrogen-bonded. The methanol molecules re-orient themselves such that their hydrophobic methyl groups come into contact with each other and the hydrophilic hydroxyl groups point at the pockets of water [4]. We were also able to say something about the dynamics of the water clusters, by calculating their average lifetimes. What we found was that large water clusters, such as the one shown in the simulation snapshot in Figure 2, could live for almost the entire duration of the simulation (in some cases several nanoseconds). These 'super-clusters' were far from static though, as they rapidly shed and re-absorbed peripheral water molecules and smaller clusters from their outer surfaces. The average lifetime of the small fraction (about 10%, aggregated over the entire simulation) of water molecules not part of a cluster was only a few picoseconds.

In analysing the MD trajectories, we faced a significant data challenge: thousands of particles running for millions of time steps generate multiple gigabytes of trajectory data. But that is just one composition at one state point. Visualisations such as Figure 2 help a lot, as do movies, but there is still too much information to digest. By coding our own analysis tools, we were able to distil the essential information from gigabytes of data. This allowed us to capture cluster size distributions, calculate their lifetimes and even to partition the members of water clusters to be on either the external or the internal surfaces of the cluster and to investigate the properties of these types of molecules separately.

This micro-segregation was not unique to a single composition. Using interchangeable data formats and the same analysis tools, a combined experimental and computational study [5] showed that the clustering of water persists across a wide composition range (see Figure 3). The similarity of the cluster distributions obtained from two completely independent datasets, one derived from experimental data, the other from MD simulations, is remarkable. Furthermore, both water and methanol (data not shown) are found to cross the theoretical percolation threshold for a range of compositions, making aqueous methanol a bi-percolating liquid mixture. Interestingly, the composition range for which this occurs is in the same region in which the anomalous thermodynamic properties have extreme values.

So if this is how water behaves around just about the simplest of amphiphiles, what does this mean for the structure and dynamics of water around even slightly more complex amphiphiles? We don't yet know the answer, but the current programme of both experimental and computational research in this area certainly aims to try to find out. Work is just getting underway on this and related systems on HPCx. Obvious questions that we would like to answer include: what happens to these water clusters at far from ambient conditions, at higher pressures or lower temperatures? Do these classical simulations predict the same things as quantum mechanical simulations (our own simulations on very small systems suggest that they do)? Another interesting question is just how big do these water clusters get? There is an obvious upper limit to their size since solutions of methanol are clear, but there is a long way between the lengthscale of where we have looked so far and that ceiling. More important than 'how big?' is simply 'why?' What is the energetic basis for the formation of these structures?

While it might be plausible to think that methanol is a good prototypical amphiphile, it may be taking it too far to consider it a realistic model of a biomolecule. So what do we observe for a simple biomolecule? In a collaboration with IBM Research at the T.J. Watson Research Center, classical and quantum mechanical MD simulations of N-methylacetamide (NMA, $CH_3N(H)C{=}OCH_3$) solutions are underway. The hydrophobic methyl groups are there once again, but this time the hydrogen-bond acceptor and donor characteristics are spread across the NHCO linkage, a signature protaceous element. We don't know what the structure and dynamics of the water surrounding this molecule will be, but we're fairly sure it will be anything but simple.

References

[1] www.sbu.ac.uk/water

[2] B. Guillot, J. Mol. Liquids 101 219 (2002)

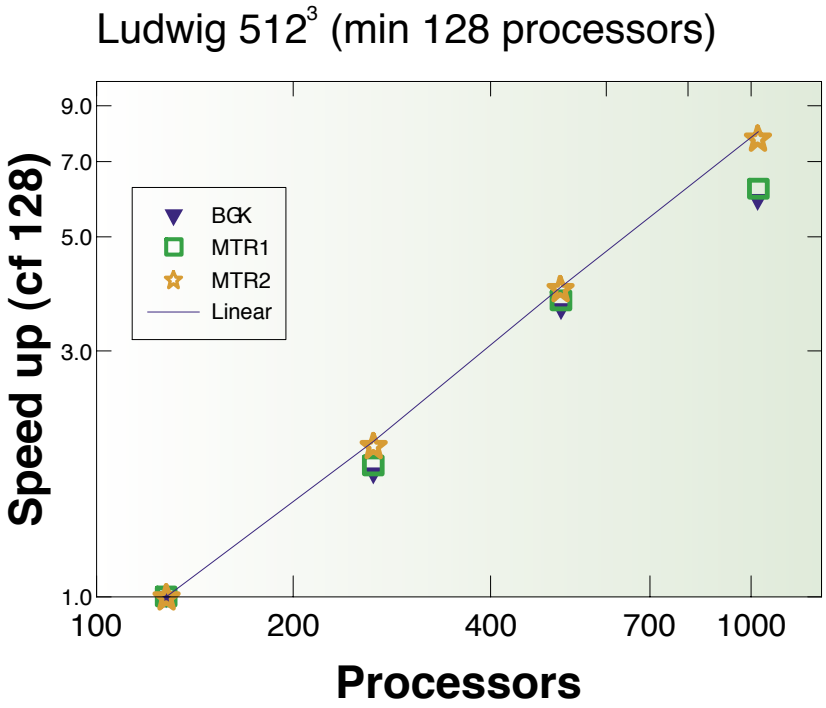
[3] S. Dixit, J. Crain, W.C.K. Poon, A.K. Soper, J.L. Finney, Nature 416 829 (2002)

[4] S.K. Allison, R. Hargreaves, J.P. Fox and S.P. Bates, submitted to Phys. Rev. Lett.: preprint at arXiv:cond-mat/0404380

[5] L. Dougan, S.P.Bates, R. Hargreaves, J.P. Fox, J. Crain, J.L. Finney, V. Reat, A.K. Soper J. Chem. Phys. in press (2004).

Benchmark, then relax... Some different lattice Boltzmann algorithms on HPCx

Kevin Stratford, HPCx Terascaling Team



The lattice Boltzmann (LB) equation provides an important way to solve the Navier-Stokes equations for fluid flow, and one which is particularly suited to parallel computing as all its operations (including the calculation of the pressure) are local.

For some years at Edinburgh, we have developed LB codes that investigate binary fluid mixtures. The state of the system in LB is described by some 30 variables or 'distributions' held at each lattice site, each with an accompanying discrete velocity. The hydrodynamic quantities (density, momentum, etc) and the composition of the mixture are then different combinations ('moments') of the distributions.

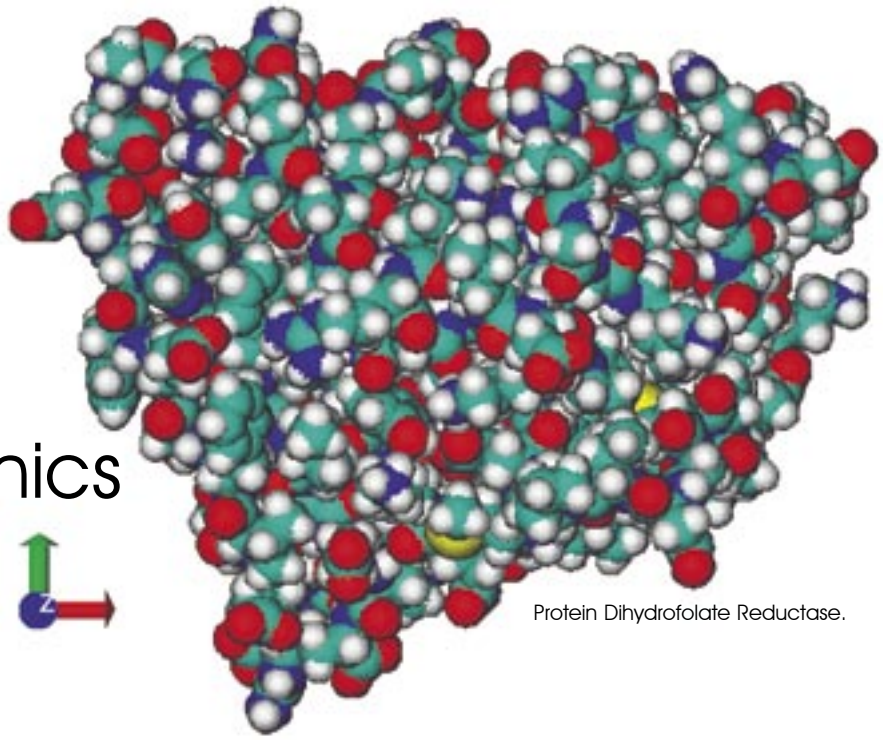
The LB algorithm proceeds in two stages, which are referred to as collision and propagation. In the collision stage, the distributions are relaxed toward some local equilibrium distribution which is determined by the physics in question. In its most simple form, a single relaxation time is used for each moment of the distribution in the collision. A more sophisticated approach, which has benefits for both numerical accuracy and stability, is to relax different moments of the distribution at different rates: a multi-relaxation time approach. Such an approach has been developed for binary fluids at Edinburgh. The propagation stage moves the distributions between adjacent lattice sites according to their discrete velocity.

The LB computation follows these two stages: the collision involves a significant number of floating point operations at each lattice site. However, the propagation involves no floating point operations, merely memory movements. Parallelisation is achieved using domain decomposition plus message passing, for which only halo swaps are required.

The figure shows a number of benchmarks performed on HPCx for the Edinburgh LB code ('Ludwig', after Boltzmann) for a lattice size of 512^3. Memory requirement means that the minimum number of processors needed for this problem is 128, so the speed up figures are relative to 128. The Figure shows a number of different collision algorithms including a single relaxation time (BGK) algorithm, and a multi-relaxation time algorithm (MTR1). Both basic algorithms perform well. Improvements to the propagation scheme (MTR2) ensure that the code scales linearly to 1024 processors for this problem. We have yet to apply for our capability incentive gold star!

Benchmarking molecular dynamics applications on HPCx

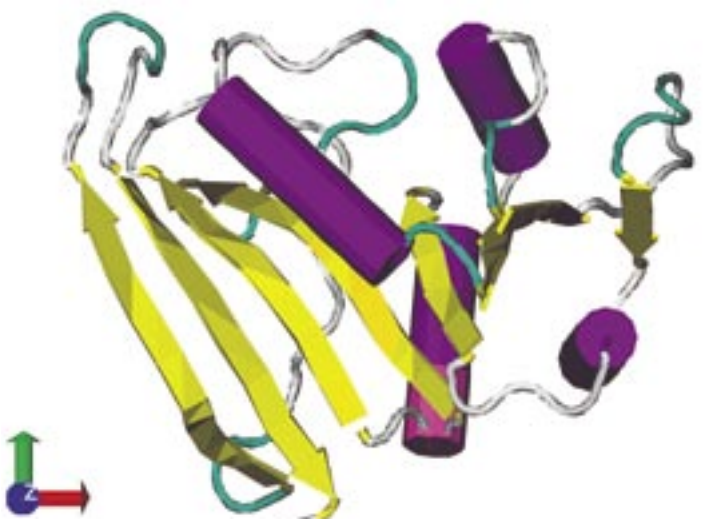
Fiona Reid, HPCx Terascaling Team



Protein Dihydrofolate Reductase.

The HPCx terascaling team are currently undertaking a comparison study of various cutting-edge Molecular Dynamics (MD) applications used on the HPCx system. At present five different MD applications have been identified for this investigation: NAMD, GROMACS, AMBER/PMEMD, CHARMM and DL_POLY.

The aims of the study are to compare the performance, scaling and functionality of the MD applications using a variety of bio-chemical systems for different numbers of processors. The performance and scaling of each application will be measured by running research grade benchmarks consisting of between 50,000 to 1,000,000 atoms. The use of real user benchmarks rather than application specific benchmarks tweaked to perform well on a given application will provide a realistic comparison of the different applications. It is our hope that HPCx users will supply the research grade benchmarks.



In order to make the comparisons between applications unbiased, we plan to run each benchmark system across a number of the application codes (ideally across all five). In addition we will consider the flexibility, ease of use and functionality of the different applications.

We hope that the results of this study will allow HPCx users to choose the most appropriate code for their scientific/research problem, whether specifically for speed of computation or functionality. We also hope that the study will guide optimisation efforts by the HPCx terascaling team. It should be noted that we are not seeking to champion a single application but rather to provide users with information that will allow them to choose the application that best suits their research problem.

A number of research groups currently using molecular dynamics codes on HPCx have been approached to supply benchmark systems. To date the response has been extremely positive. However, more benchmarks are still required, particularly for large (>300,000 atom) systems and also for the AMBER and DL_POLY applications.

If you would like to contribute benchmarks, please contact the helpdesk: helpdesk@hpcx.ac.uk

Dr Ken Badcock is the principal investigator of the UK Applied Aerodynamics Consortium (e20). He is part of the Department of Aerospace Engineering at the University of Glasgow. His research is focused on computational fluid dynamics, and is of direct interest to the UK aerospace industry.

UK Applied Aerodynamics Consortium (UKAAC) established

By K.J.Badcock, University of Glasgow, G.J.Page, Loughborough University, and D.R. Emerson, Daresbury Laboratory

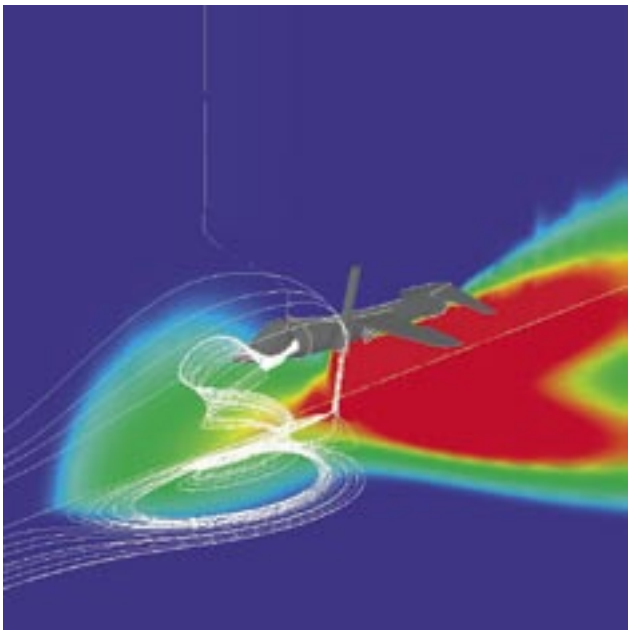


Figure 1: Flow into the intake of a Harrier (provided by Cambridge University).

Computational aerodynamics is a pace setter for the use of major computational facilities in the US and worldwide. It has been estimated that around half of the CPU time used on Department of Defence supercomputers is for aerodynamic studies, contributing to current and future aircraft projects in areas such as flow control, the high angle of attack regime, and design of re-entry vehicles. Research problems in aerodynamics range from fundamental investigations of turbulent flows using Direct Numerical Simulation and Large Eddy Simulation, through to multi-disciplinary analyses of real aircraft configurations at high Reynolds number. The former has been well represented on UK academic supercomputers for a number of years but the latter has been noticeably absent, despite a healthy research community and a clear demand for computational resources.

This situation led to the formation of the UK Applied Aerodynamics Consortium (UKAAC), which started on HPCx in June 2004 (GR/S91130/01). The consortium involves the Collaborative Computational Project 12 (CCP12), with 12 academic institutions and four industrial partners (BAE SYSTEMS, Rolls Royce, Westland Helicopters and Fluent Europe) whose work is split into themes for helicopters, flexible airframes, engine aeroelastics, vertical landing, and engine-air systems. A significant proportion of the work is aligned with research funded by the Defence and Aerospace Research Partnerships (DARPs) for Rotorcraft Aeromechanics and Unsteady Flows (PUMA: the Progress in Unsteady Modelling of Aerodynamics program).

Figure 2: Hawk flutter simulation (provided by Glasgow University)

One target application is the simulation of a vertical landing aircraft, featuring complex geometry, complex, unsteady flow physics due to jet impingement, and a moving mesh as the aircraft descends. A number of research questions are currently being dealt with in the PUMA DARP to provide the methods required. Current simulations on idealised problems, as shown in figure 1, requires several weeks of processing on local HPC facilities. The target of simulating a full aircraft geometry will require an order of magnitude increase in the number of grid points, making this a capability calculation on HPCx.

A further target application is the simulation of fluid-structure interaction, demonstrated on the Hawk aircraft (figure 2). This simulation couples the structural deformations predicted by a finite element model with a CFD simulation of the aerodynamics. At present the aerodynamic simulations have been limited to inviscid flows but there are many potential problems concerned with phenomena requiring turbulent effects to be incorporated. Again, these are capability calculations on HPCx.

The successful development and demonstration of mature methods for these sorts of problems can have a major influence on the way that aircraft are designed in the future and on the benefits of HPC to industrially-relevant problems. The UKAAC community will be making significant advances on current state-of-the-art simulations and providing demonstrators that will re-establish applied aerodynamics as a key UK HPC activity.

Further information: Ken Badcock, gnaa36@aero.gla.ac.uk

Dr Roderick Johnstone is a member of the UK Turbulence Consortium (e01) and is part of the Computational Engineering Group at CCLRC Daresbury Laboratory.

Direct numerical simulation of a neutrally stratified turbulent Ekman layer

Roderick Johnstone, Mike Ashworth, CCLRC Daresbury Laboratory, Gary Coleman, Aeronautics and Astronautics, University of Southampton

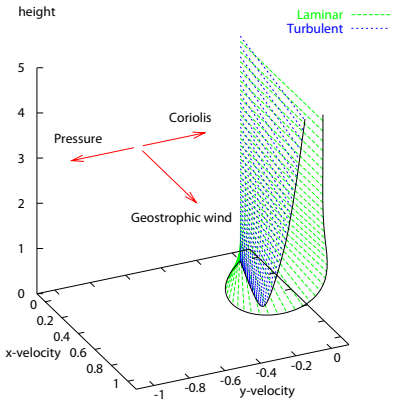


Figure 1: Laminar and turbulent Ekman boundary layer.

The atmospheric planetary boundary layer (PBL) is the region of the atmosphere directly affected by fluxes of momentum, heat, and mass to and from the surface, responding to surface forcings within a short timescale of an hour or less.

Above the boundary layer, the pressure gradient and Coriolis force are in approximate balance, while the velocity is normal to both (this is called geostrophic flow).

As the surface is approached, however, the wind speed decreases to zero, and with it the effects of the Coriolis force; this causes the wind to veer in the direction of low pressure.

In the absence of turbulence, it is possible to obtain an analytical solution to the governing Navier-Stokes equations for this flow; this is the Ekman spiral, shown in Figure 1 as the 'Laminar' solution. However, the PBL is, in reality, turbulent. The 'Turbulent' solution in Figure 1 is a spatially and temporally averaged result from a simulation of the turbulent Ekman layer currently running on HPCx (described below).

The PBL is important for weather prediction – firstly, almost all interaction between people and the atmosphere takes place there. There is therefore a demand for forecasts of near-surface conditions such as wind speed and temperature. Conversely, surface forcings of the atmosphere involve the PBL; for instance, it is from the surface that the atmosphere gains moisture, and most of its solar heat input; the surface also exerts drag on the atmosphere. Turbulence affects all of these fluxes.

Unfortunately, it is not feasible to resolve PBL turbulence in weather prediction models (which instead model the PBL), nor even to simulate it numerically at its real Reynolds number (Re). This is because the computational expense of such simulations scales approximately with the cube of the Reynolds number, and Re is very large indeed in geophysical flows (up to about 10^8 for the PBL).

However, it is possible to use direct numerical simulation (DNS) of the turbulent Ekman layer at a lower Re than that of the PBL, in order to test parametrizations of the PBL, and to obtain values for the empirical constants they require. Idealized simulations are in fact quite suitable as it is possible to exclude from them effects that the theory is not designed to represent (which is difficult to do when using field data for the same purpose).

A simulation of this kind is currently being run on HPCx. The results of this work could be used to improve subgrid-scale models used in Large-Eddy Simulations, and to test similarity theories of the turbulent Ekman layer; the principal aim is in fact to use the results obtained in combination with the theory of Spalart [1].

This theory extends the similarity theory of Csanady [2] to low Reynolds numbers, which makes it possible to extrapolate results from a DNS to the much higher Re of the PBL. The collection of data at a series of increasing Reynolds numbers will permit greater confidence to be placed in such extrapolation.

Coleman et al [5,6], employed a code implementing the method of Spalart et al [3,4], for the simulation of the neutrally-stratified, incompressible, turbulent Ekman boundary layer at Reynolds numbers of 400, 500 and 1000, based on the geostrophic wind speed and viscous Ekman layer depth. This work is now being extended with a further simulation at Re=2000, using HPCx.

This code was originally written in the Vectoral language [7] and run on the Cray C90. Since then it has been translated to Fortran 95 and parallelised using MPI. Because different transforms are applied in the vertical and horizontal directions, it makes sense to decompose the domain alternately into vertical and horizontal planes for the purpose of parallelization. The global character of spectral discretisations makes heavy demands with regard to communication, however. Here, this takes the form of a transpose of large distributed arrays, requiring all-to-all communication.

The code has been optimised to improve the performance. The following were found to be beneficial:

1. Use of dgemm from BLAS for matrix multiplication. Obvious as this is, it is worth a 70% reduction in run time in current production use.

2. Use of the ESSL routine zgetmo to transpose arrays for the FFTs (about 7% reduction in run time).

3. MPI_Alltoallv is used to simultaneously perform the parallel transpose together with a broadcast of some additional data. Separating the broadcast from the all-to-all creates an opportunity to overlap the all-to-all communication with some computation; this was attempted, but found to be slower than simply amalgamating the two.

Continued opposite.

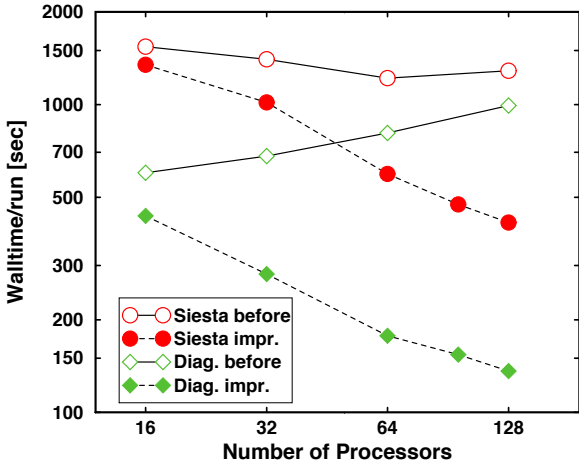
Improved parallel performance of SIESTA for the HPCx Phase2 system

Joachim Hein,
HPCx Terascaling Team

SIESTA is a materials code using self-consistent density functional theory. The code aims at the simulation of materials with a large number of atoms. The name SIESTA is an acronym for 'Spanish Initiative for Electronic Simulations with Thousands of Atoms'. SIESTA uses order-N algorithms allowing linear scaling with the number of atoms [1].

The application is used by several users of the HPCx service. In particular, the e-Minerals consortium is investigating the environment from the molecular level. The science being performed on HPCx is described in the article on p10.

To obtain acceptable runtimes on materials with such a large number of atoms, good parallel scalability is required. Benchmarking by the HPCx terascaling team showed that the performance of the diagonaliser was the key obstacle for the code to scale to a large number of processors. Furthermore, using a larger problem size did not help the parallel performance of the diagonaliser. The performance was clearly inferior when compared to the experience the HPCx terascaling team has gained with eigensolvers [2]. Redistributing the data inside the diagonaliser leads to a dramatically improved performance when using a large number of processors. For our benchmarking configuration SIESTA now runs up to three times faster when used on 128 processors. With the recent hardware upgrade and the change



Performance of Siesta and its Diagonaliser for a benchmark system before and after the improvement

from 8-way to 32-way nodes, this is a very timely improvement. SIESTA is now able to use several 32-way SMP-nodes for a single calculation. These improvements will be incorporated in future SIESTA releases.

SIESTA benefits further from the upgrades to the switch micro-code, introduced to the service in July 2004. On 128 processors we observed an improvement of about 14%.

A full technical report on this work, containing detailed instructions on recommended environment settings is available [3].

References:

[1] J.M. Soler, et al., J. Phys.: Condens. Matter 14, 2745 (2002).

[2] Elena Breitmoser, Andy Sunderland, 'A performance study of PLAPACK and ScaLAPACK Eigensolvers on HPCx for the standard problem', HPCx technical report HPCxTR0406

[3] Joachim Hein, 'Improved parallel performance of SIESTA for the HPCx Phase2 system', HPCx technical report HPCxTR0410

---

4. The recent HPS switch microcode upgrade (SP7) which took place on 28th July 2004 has markedly improved the parallel scaling of this code on HPCx (see figure 2).

This simulation is funded by EPSRC and managed through the UK Turbulence Consortium.

References:

[1] Spalart, P.R., 1989, 'Theoretical and numerical study of a three-dimensional boundary layer', Journal of Fluid Mechanics 205, pp 319-340

[2] Csanady G.T., 1967, 'On the 'resistance law' of a turbulent Ekman layer', Journal of the Atmospheric Sciences 24, pp 467-471

[3] Spalart, P.R., 1986, 'Numerical Simulation of Boundary Layers: Part 1. Weak Formulation and Numerical Method', NASA Technical Memorandum 88222

[4] Spalart, P.R., Moser, R.D., and Rogers, M.M., 1991, 'Spectral Methods for the Navier-Stokes Equations with One Infinite and Two Periodic Directions', Journal of Computational Physics 96, pp297-324

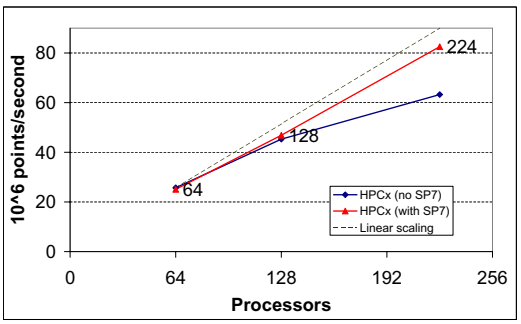


Figure 2: Performance of the Turbulent Ekman Boundary Layer code before and after the HPS switch microcode upgrade (SP7).

[5] Coleman, G.N., Ferziger J.H., and Spalart, P.R., 1990, 'A numerical study of the turbulent Ekman layer', Journal of Fluid Mechanics 213, pp 313-348

[6] Coleman, G.N., 1999 'Similarity Statistics from a Direct Numerical Simulation of the Neutrally Stratified Planetary Boundary Layer', Journal of the Atmospheric Sciences 56, pp 891-900

[7] Wray, A. 2002, http://merrimac.stanford.edu/brook/vectoral.pdf

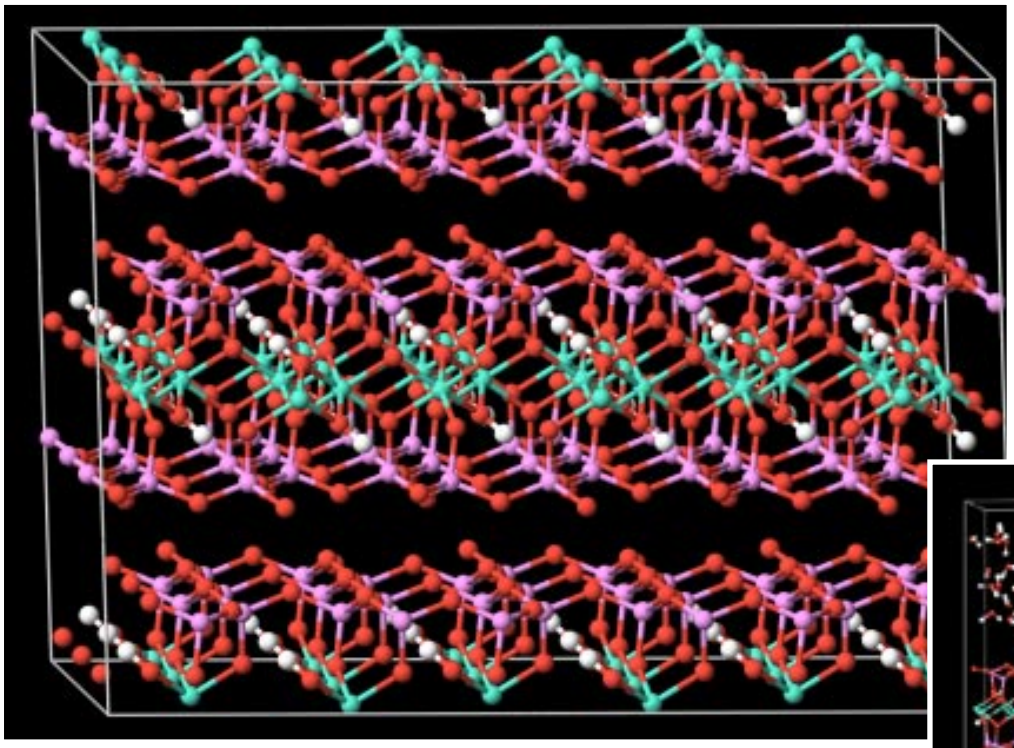
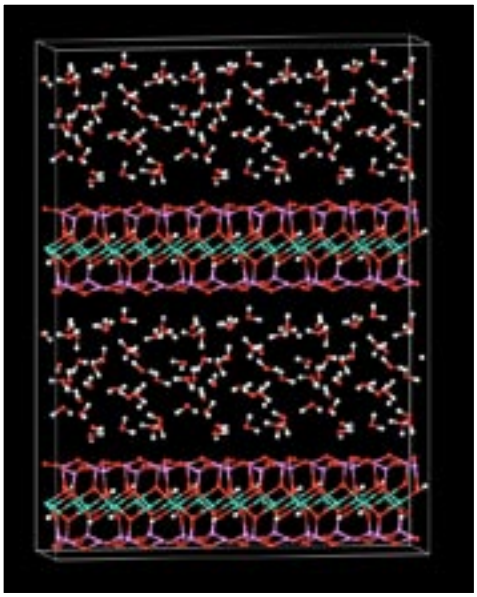


Figure 1: Pyrophyllite

Figure 2: Pyrophyllite at the mineral/water interface.



e-Minerals: environment from the molecular level

John Wakelin, The University of Cambridge

Dr Jon Wakelin is part of the Department of Earth Sciences at the University of Cambridge. He is a member of both the Materials Chemistry Using Terascaling Computing Consortium (e05) and the Computational Mineral Physics Consortium (n03).

We are interested in the processes by which organic pollutant molecules (such as DDT, PCBD and dioxins) bind to soil minerals. This is important because of the wide range of industrial organic pollutants (e.g. DDTs, PCBs, dioxins) found in soils and increasingly being detected in the human food chain, together with the known problems for health and reproduction. An atomistic approach must account for realistic mineral surfaces in contact with fluids containing ionic components with variable pH, and the role of natural organic matter at mineral surfaces.

Initially we are focusing on the interaction 2,3,7,8-tetrachloro dibenzodioxin (TCDD) with the (001) surface of the layered clay mineral Pyrophyllite as a representative system. 2,3,7,8-TCDD makes an excellent test case because it is a known carcinogen and widely regarded as the most toxic of all the dioxin congeners and has been the focus of most experimental research into dioxins.

In total there are 419 dioxin (or dioxin-like) compounds. We intend to perform a systematic study of the interaction of all of these compounds with a range of clay minerals. We estimate that this will take something in the region of 40,000 calculations for a single mineral surface (depending on the methodology). For a given mineral there may be a large number of relevant surfaces, moreover there are potentially thousands of minerals that one might wish to investigate. This is why we are investing so much in setting up these calculations. Certain calculations will need to be performed on extremely large systems and this is why having an optimised SIESTA running on HPCx is so important (and unlike some HPC facilities HPCx is Grid-enabled). Eventually, we want to look at more realistic systems, such as the 'wet' system shown in Figure 2.

This work is being conducted as part of one of the UK e-Science programs (e-Minerals: Environment from the Molecular Level). We are building an infrastructure for performing simulations that exploit the emerging Grid technologies using Globus and Condor. Our HPC focus is on performing simulations on environmentally relevant systems.

Dr Andrew Coward is based at the Southampton Oceanography Centre and is a member of the Large-Scale Long-Term Ocean Circulation Consortium (n01).

A study of Agulhas rings in a high-resolution ocean model

A. C. Coward

James Rennell Division for Ocean Circulation and Climate, Southampton Oceanography Centre, Southampton, United Kingdom

J. Donners and S. S. Drijfhout

Department of Oceanography, Royal Netherlands Meteorological Institute, De Bilt, The Netherlands

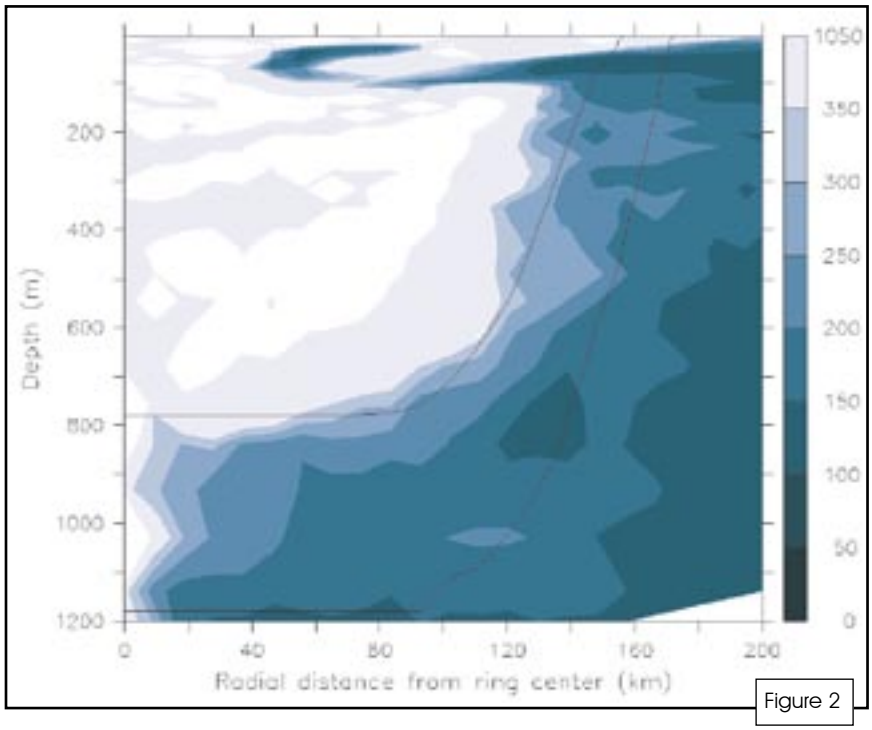


Figure 2

Introduction

The Agulhas Current is the strongest western boundary current of the southern hemisphere. It retroflects to the south of South Africa and flows back into the Indian Ocean as the Agulhas Return Current (ARC). Agulhas rings are shed off the retroflection loop irregularly and move into the South Atlantic Ocean in a north-westerly direction. Present estimates are that roughly six Agulhas rings are being formed each year. These rings form an important link between the subtropical gyres of the South Atlantic Ocean and the Indian Ocean. This link is also thought to play an important role in the upper branch of the global thermohaline circulation. These rings can be followed with satellite measurements of sea surface height (SSH). Such measurements show that the decay of the SSH is strongest during the first five months after the shedding of Agulhas rings. Observational studies of this important section of the thermohaline circulation are hampered by the episodic nature of the shedding process and the reliance on fortunate positioning of research vessels. Complimentary research using realistic ocean models is essential to improve our understanding of the regional processes and hence of any possible changes in a warming environment.

The UK's high resolution Ocean Circulation Model (OCCAM) is now being run at an eddy-resolving resolution of 1/12° globally and has an effective grid size of approximately 7.5km in the Agulhas area. The model employs 66 depth levels, with 20 levels in the top 200 m and an explicit mixed layer. During the integration described here, the ocean surface is forced by a monthly wind stress climatology and a relaxation of temperature and salinity toward monthly climatological values. Integrating this model requires a capability machine such as HPCx and the model runs routinely on 352 processors of the Regatta system. Results from this model provide an ideal framework for investigating regional processes such as those described and this article briefly describes some analysis carried out in collaboration with colleagues from the Royal Netherlands Meteorological Institute (KNMI).For this analysis we used output from the second year of a two model-year integration.. The model provides three-day average values of all its state variables. Each such dataset is 8.5GB large which gives over 1TB of output per model year. These data were used to analyse the leakage of water from Agulhas rings using a novel, off-line, particle tracking approach developed by de Vries and Doos [2001] for the use with time-dependent flows. More than twenty thousand particles were used for the smallest Agulhas ring. Each trajectory was followed for the full second year of the model run. Particles were seeded in the top 2000m of the Agulhas rings wherever the SSH was above 25 cm. The center of the Agulhas ring was defined as the SSH maximum. The Agulhas rings were analyzed also in a two-dimensional, Eulerian framework. To this end, the data was azimuthally averaged around each ring centre and plotted as a function of radial distance and depth.

Figure 1 (on page 12) shows the SSH at the beginning of the year that has been used for the analysis. The Agulhas Current is centered at 24°E. The heavy dashed lines indicate the regions within the three Agulhas rings where Lagrangian particles have been seeded. Also the path of the three rings has been indicated. The youngest (and largest) Agulhas ring is two weeks old. The ring in the middle was shed in mid-November, at the end of spring. The oldest ring was formed in midwinter of the first year.

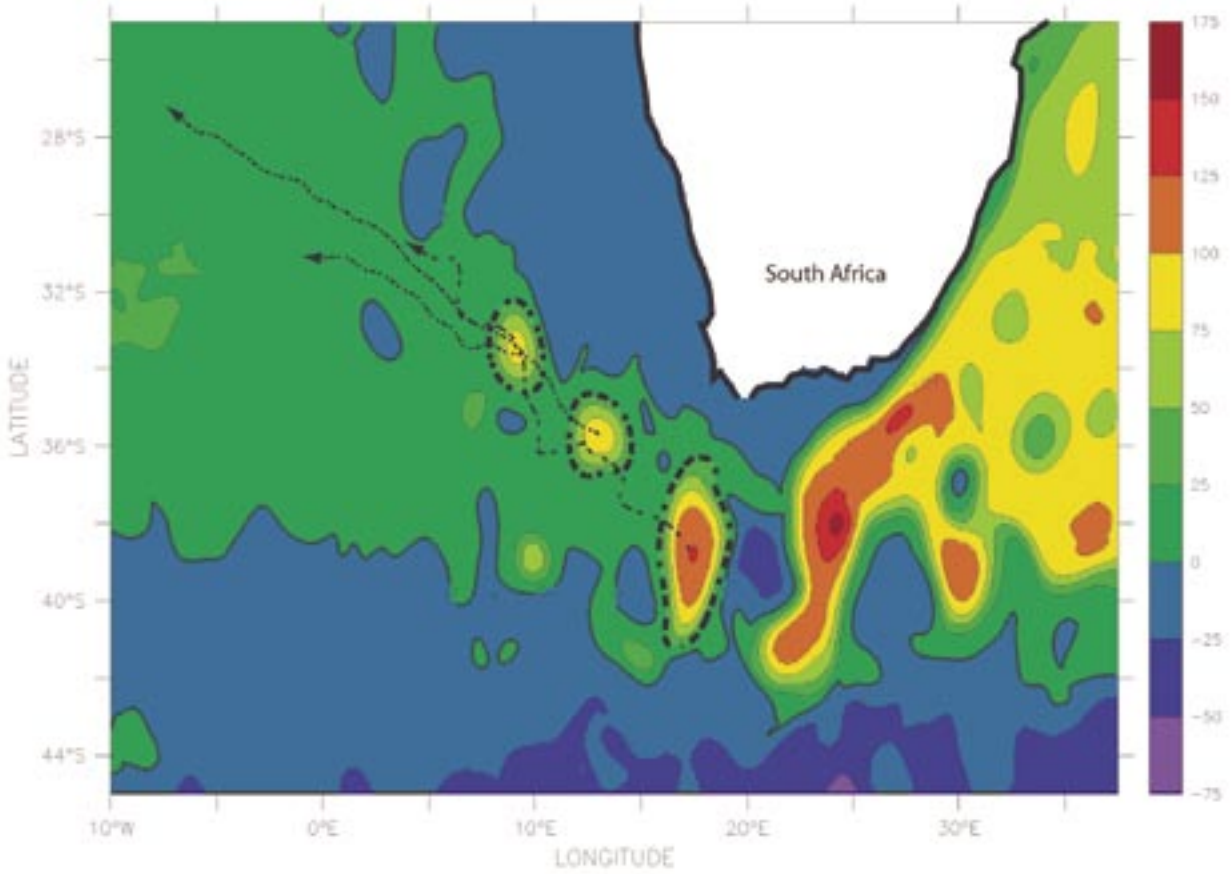


Figure 1

Figure 2 (on page 11) shows the average residence time (or e-folding time scale) of particles within the middle ring as a function of distance from the ring centre and depth. A clear, bowl- shaped division can be seen between particles that reside in the ring (white and light gray) and particles that leak into the environment (dark gray). The bowl-shape reaches to a distance of 140km and a depth of 800 m. The upper solid line in figure 2 shows the contour where the rotational velocity is at least twice the translational speed of the Agulhas ring. Note that the contour shallows at approximately 80km distance from the ring center, where the maximum azimuthal velocity is largest. At radii larger than 80km the criterion is a good indicator of the bowl-shaped ring boundary. The average residence time along the ring boundary is 250 days. In the upper 150m of the Agulhas ring a significantly stronger water mass exchange can be seen, in the form of a curl of water with a lower residence time. This suggests that the surrounding water is drawn into the ring at 100m depth, upwells within the ring and flows radially outward near the surface. This shallow overturning cell is the strongest circulation feature within the bowl shaped ring boundary and is confined to the mixed layer. The cell is present in all Agulhas rings during winter and spring, when heat loss is strong.

We have further analysed the model results to investigate the physical mechanisms behind this previously unobserved phenomenon. This analysis is the subject of an article (Donners et al., [2004]). Briefly, we were able to conclude that:

There is a sharp boundary between particles that stay within the Agulhas ring and particles that mix into the environment. The ring boundary is well predicted by a criterion that the azimuthal velocity is at least twice the translational velocity at any time.

Crucially, mixing with the environment is found to be enhanced when strong surface cooling generates a shallow overturning cell with radially outward flow near the surface and a compensating inward flow at depth. This circulation is limited to the mixed layer. The cell can be explained by Subinertial Mixed Layer theory: cooling creates vertically-sheared pressure gradients that induce vertically-sheared subinertial motions. Vertical mixing is balanced by restratification due to the sheared flow. The overturning cell forms an effective pathway between the edge and the inside of the Agulhas ring and it amplifies the dilution of the anomalous water properties of Agulhas rings near the surface. The surface water is not trapped in the core, but connected with the outside: the overturning cell amplifies this water mass exchange by constantly bringing new water to the edge, where it is mixed with the environment.

The model integrations are carried out as part of Southampton Oceanography Centre's core strategic research program, which is supported by the Natural Environment Research Council (NERC). This work is just one example of the environmental research carried out using the NERC share of HPCx. Realistic simulations of our natural environment will continue to require resources capable of integrating complex numerical models and analysing the large volumes of data generated by such models. The level of realism that can now be achieved is providing an invaluable complement to observational programmes.

References

de Vries, P., and K. Doos (2001), Calculating Lagrangian trajectories using time-dependent velocity fields, J. Atmos. Ocean. Tech., 18, 1092{1101.

Donners, J.; Drijfhout, S. S.; Coward, A. C., Impact of cooling on the water mass exchange of Agulhas rings in a high resolution ocean model, Geophys. Res. Lett., 31, L16312, doi:10.1029/2004GL020644.

Meeting of NERC Consortia

University College London, 26th May 2004.

Kevin Stratford, HPCx Application Support Team

This meeting aimed to allow users to discuss issues arising from the upgrade of the HPCx service with members of the HPCx support team. By request, the morning session of this event was targeted specifically at users from NERC consortia, while the afternoon was available to all HPCx users.

In the morning, representatives from different NERC consortia described the nature of their science, and highlighted the different challenges associated with achieving this science on HPCx.

In particular, Warwick Norton gave an overview of the aims of HIGEM, the coupled climate-modelling effort which aims to feed into the next Intergovernmental Panel on Climate Change (IPCC) assessment in 2007. Paul Burton from the NERC Centres for Atmospheric Science Consortium (NCAS) discussed his work on supporting and developing the Met Office's Unified Model for CGAM and the wider UGAMP community, focusing particularly on performance related issues. Lastly, Andrew Coward from the Large-Scale Long-Term Ocean Circulation Consortium discussed their experiences on HPCx. This highlighted some interesting points:

- There was some discussion of queues, and possible queue changes that would allow a faster turnaround for some consortia jobs on a temporary basis. For example, implementing a much longer queue length eg 144 hours. There was also some discussion of ways of achieving quicker turnaround on small short 32-processor jobs during the daytime.
- There was some discussion of problems with getting the expected amount of memory on the system. IBM agreed to investigate this.
- There was a request for higher priority for capability jobs.

This was a successful event, and we are considering holding similar events for other scientific areas.

Capability usage

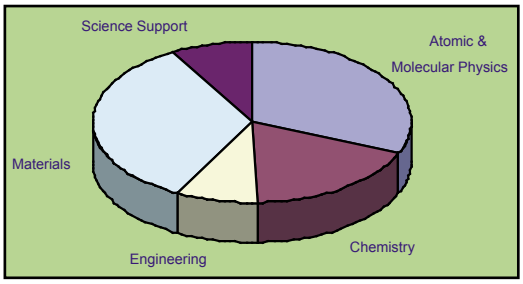
Alan Simpson,
HPCx Project Director

HPCx has been very popular right from the beginning. Within three months of the start of service, the utilisation was already at 80% and this has continued consistently for the last 18 months. The service now supports more than 30 national research consortia from many different applications areas and more than 350 users.

The major mission for the HPCx project is to deliver world-class science from capability computing, where capability computing means jobs that use a significant fraction of the system.

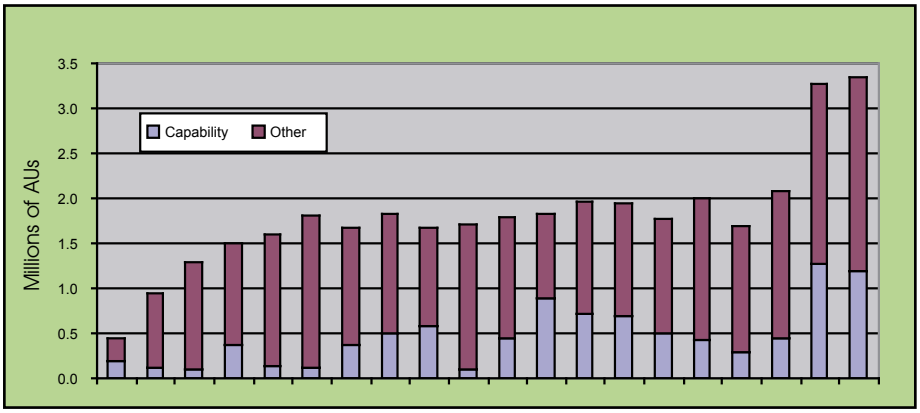
The graph shows the monthly utilisation and capability utilisation of HPCx since the start of service. The upgrade to Phase 2, which was completed by the beginning of June, has allowed the system to deliver significantly more flops and usage has quickly grown to take advantage of this. It is also encouraging to see how the capability usage has increased. At the beginning of 2003, capability usage was around 100,000 AUs (Allocation Units) per month, which represented some 10% of the total utilisation. This has increased over the last 18 months and, since the upgrade to Phase 2, we have been delivering well over 1,000,000 AUs which is over 35% of the total utilisation. This is clearly an encouraging trend.

The pie-chart shows how the capability usage is split up between different applications areas. As Chemistry and Materials Science are the major overall users of the system, it is not surprising that a reasonable fraction of the capability usage also comes from those areas. However, it is interesting to note that Engineering (through the UK Turbulence Consortium) and, especially, Atomic & Molecular Physics also represent significant fractions of the capability usage. The usage from the Science Support effort reflects the work of the Terascaling team to ensure that a wide range of applications codes can effectively exploit large numbers of processors. In fact, more than 20 consortia have already run capability jobs.



It is encouraging to see the continued increase in capability usage of the HPCx system with respect to both the amount of computational power used and the number of consortia. The enhanced High Performance Switch of the Phase 2 should allow many codes to scale even better and the initial indications are very promising.

Monthly values from December 02



Capability science initiatives within the UK

Lorna Smith, HPCx Terascaling Team

It is interesting to note the strong focus on capability science within the UK at the moment. In addition to the activities under HPCx, the Met Office and the European Centre for Medium-Range Weather Forecasts (EWMWF) have both installed state-of-the-art supercomputing systems to enable accurate weather forecasting. Also of note is the UKQCD (UK Quantum Chromodynamics) Consortium, whose computational demands have led them to build a special purpose HPC resource matched to the specific needs of QCD.

The Met Office

The Met Office, a world leader in advising on the weather and the natural environment, was established 150 years ago.

It has recently installed two NEC systems, of 15 SX-6 nodes, each with peak performance of 0.96 teraflops. The sustained performance of the NEC systems is six times more than the two Cray T3E computers they replace. An upgrade in 2005 will double the super-computing computing power available at the Met Office.

These new machines allow the Met Office to continue to enhance the quality of its predictions. Various improvements will be made to the input data and algorithms used in the forecasting systems, but much of the extra power will be used for increases in the resolution of both the weather and climate models run on these machines.

The European Centre for Medium-Range Weather Forecasts (EWMWF)

As the name suggests, ECMWF provides medium-range weather forecasts to meteorological services in a wide variety of European states, including the Met office. Making useful weather forecasts for a week or so ahead requires the most sophisticated numerical models, the most complete database of weather observations, and the most powerful computers. Hence ECMWF require access to state-of-the-art supercomputing systems. Currently, they have two IBM p690+ clusters, each with 66 nodes and a peak performance of 8.9 teraflops.

QCDOC

Fundamental physics research has always relied upon the latest in state-of-the-art computer hardware, with modern lattice QCD simulations now measured in teraflop-years of compute time and the data produced is measured in terabytes. As a result, lattice QCD research demands state of the art supercomputer hardware.

The key performance factors for QCD calculations are memory and communications latencies. These problems are difficult to surmount with commodity processors and interconnects in current computational clusters. The solution is to build a special-purpose computer based on a custom-designed processor with large on-chip memory and communications hardware matched to the specific needs of QCD.

UKQCD, a group of theoretical QCD scientists from around the UK, working jointly with a team at Columbia University, NY, have built a supercomputer, QCDOC, which is due to be installed at the University of Edinburgh at the end of 2004. With over 12,000 processors, this will provide over 10 teraflops of performance and will be the fastest dedicated machine for QCD in the world.

Further information

The Met Office: www.met-office.gov.uk/

The European Centre for Medium-Range Weather Forecasts (EWMWF): www.ecmwf.int/

UKQCD: www.ph.ed.ac.uk/ukqcd/

Tutorial at SC2004

Mark Bull, HPCx Software Engineering Team

EPCC will be presenting a half-day tutorial at SC2004 in Pittsburgh, USA on 8 November 2004. The topic of the tutorial is 'Performance Scaling on Constellation Systems' and it will be delivered by Lorna Smith and Mark Bull, who are both members of the HPCx Terascaling team. The tutorial will draw on material developed for HPCx training courses, as well as including experiences on other constellation (clustered SMP) systems.

The tutorial will provide participants with an overview of the essential features of this type of architecture and of the main techniques for achieving optimal performance and scaling on constellation systems.

Part of the tutorial will focus on the challenges of systems with a non-uniform communication structure; on constellation systems communication within an SMP node is significantly faster than it is between nodes. Topics include cluster-aware collective communications, balancing communication and computation, and optimal process mapping.

The remainder of the tutorial will cover mixed-mode programming, where message passing is used between nodes, and a shared memory model is used within a node, so will consider portability, implementation and performance issues.

The topics covered throughout will be illustrated with a series of performance studies on a range of constellation systems.

See: www.sc-conference.org/sc2004/

The Second HPCx Annual Seminar

Ian Bush,
HPCx Applications Support Team

The second HPCx Annual seminar 'New Science from Capability Computing' was hosted by EPCC in Edinburgh on Friday 9th July. It took place at the National e-Science Centre (NeSC), and followed on from the very successful first seminar, held at Daresbury Laboratory in December 2003. We held a number of associated events on the previous day. These were a new HPCx course, entitled 'Improved Scaling on HPCx', a user group meeting and a buffet/wine reception, the latter being held in the Talbot Rice Gallery. The art on display was from a number of periods, some of the attendees not appreciating the more modern pieces on display!

All events were well attended not only by HPCx users, but also other representatives of European Supercomputing centres and IBM developers from the UK, Europe and the USA.

The eight talks presented at the seminar were from a number of different areas, the majority emphasising the new science that may be studied on top-end computational systems, with the others covering the systems themselves and how to get the best out of them. Wanda Andreoni of IBM Research, Zurich, gave an overview of the impact of HPC on science, concentrating mainly on atomistic modelling. One example was showing how HPC can help to elucidate reaction pathways in biochemical systems. Also using atomistic methods Patrick Briddon presented work on defects in carbon allotropes, which is important both in the electronics industry and in nuclear power generation. Neil Sandham discussed his group's work on turbulent flows and aeroacoustics using a number of different CFD approaches. The final two talks were on environmental modelling. Andrew Coward talked about ocean modelling (and also gave his view on 'The Day after Tomorrow'!), while Paul Selwood presented weather prediction results from the new UK Met Office machine, a NEC SX-6 cluster.

On the systems side Thomas Lippert presented the new large IBM P690+ cluster at Jülich, an interesting comparison with HPCx. Three HPCx support staff gave talks, Joachim Hein comparing the new upgrade with the older system, while Martin Plummer and Andy Sunderland presented some of the work of the terascaling team. PDFs of many of the talks are available at: www.hpcx.ac.uk/about/events/annual2004/

Lastly, the next seminar will be held at Daresbury Laboratory, and we hope to see you there!

ScicomP and SP-XXL in Edinburgh in 2005

David Henty,
HPCx Applications Support Team

EPCC will be hosting a joint meeting of ScicomP 11 and SP-XXL in Edinburgh from 30th May to 3rd June 2005.

ScicomP is an international organisation of scientific/technical users of IBM systems. The purpose of ScicomP is to share information on software tools and techniques for developing scientific applications that achieve maximum performance and scalability on systems, and to gather and provide feedback to IBM to influence the evolution of their systems. Meetings are open to all interested researchers in the area.

SP-XXL focuses on system management for large-scale scientific/technical computing on IBM hardware. It addresses topics and issues important to achieving terascale scientific/technical computing on scalable parallel machines. Meetings are open to members and affiliates of the SP-XXL group.

The meeting represents an excellent opportunity for HPCx users to share experiences with other IBM system users and to meet relevant IBM developers.

We look forward to meeting you in Edinburgh next year.

See:
www.spscicomp.org/
www.epcc.ed.ac.uk/scicomp/
www.spxxl.org/

MSc in High Performance Computing

EPCC at the University of Edinburgh offers a one-year taught MSc in High Performance Computing. This well-established programme provides an excellent grounding in HPC technologies and their practical application.

If you have a keen interest in programming and would like to learn about HPC then this MSc will appeal to you. You will have access to leading-edge HPC platforms and technologies and the course has a strong practical focus. Lectures and tutorials are closely co-ordinated with associated practical sessions. You will also undertake a 16-week independent research project leading to a dissertation.

The MSc provides a doorway to a wide variety of future careers. Previous graduates have gone on to further research for a PhD in HPC and in areas of science that utilise HPC technologies. Other students have found employment directly.

Applications are encouraged from people who are competent programmers in Java, C, C++ or Fortran and who are graduates in science, engineering, computer science or mathematics, or are currently working in a relevant field.

The entrance requirement is a good honours degree or equivalent work experience. No prior HPC knowledge is assumed.

For more information and application details see:
www.epcc.ed.ac.uk/msc/
Email: msc@epcc.ed.ac.uk

The MSc in HPC is supported by EPSRC, who offer a number of studentships.

This MSc is supported by

EPSRC
Engineering and Physical Sciences Research Council

Taught Courses

- Fundamental Concepts of HPC
- Practical Software Development
- Shared Memory Programming
- Message Passing Programming
- Parallel Decomposition
- Core Topics in HPC and e-Science
- Object Oriented Programming for HPC
- Hardware, Compilers and Performance Programming
- Applied Numerical Algorithms
- Exploiting the Computational Grid
- Scientific Visualisation
- Advanced Topics in HPC and e-Science