

Computational Science and Engineering Department, Daresbury and Rutherford Appleton Laboratories

Contents

Welcome	
Atomic and Molecular Physics Group	
Iron, Nickel and Tin in Astrophysical and Laser-Produced Plasmas:	
Atomic and Molecular Physics at Daresbury in 2006	
Advanced Research Computing Group	
Helicopter Rotor Wake Simulations	
A Parallel Multi-Band K-p Code for the Electronic Structure of Semiconductor Quantum Dots	
Parallel Performance Optimization on HPCx: Materials Chemistry	
Band Theory Group	
Dilute Magnetic Semiconductors from First-Principles Calculations	11
Superconductivity under Extreme Pressure	13
Valence Transitions in Nickelates	15
CCP4 Group	
CCP4 6.0: State-of-the-art Software for the Experimental	
Determination of Protein Structures	17
Computational Chemistry Group	
Extending the Range of Parallel Quantum Chemical Methods	19
Nano-objects in Supercritical Media:	
Solvation Behaviour of Metal Nanoparticles in Supercritical Fluids	21
Molecular Dynamics Simulation Studies of Molecular Complexation Behaviour of Cyclodextrins	23
Molecular Dynamics Study of Radiation Hard Materials: Radiation Damage in Pyrochlores	25
New Tools for the Accurate Study of Biomolecular Reactivity	27
Computational Engineering Group	
Digital Microfluidics: Droplets on Demand	29
Computational Materials Science Group	
Ab Initio Studies of Aluminium Fluoride Surfaces	31
Distributed Computing Group	
Application Performance on Dual-core Processors from Intel and AMD	33
The Challenge of Coordinating a £38M Distributed HPC Procurement	25
In the UK - SKIF3	35
Initial Experiences of FPGA Programming	37
Numerical Analysis Group	
Galahad - A Library of Packages for Nonlinear Optimization	39
Systems and Operations Group	
Less is More – Moving HPCx to Power5	41

© CCLRC, Council for the Central Laboratory of the Research Councils 2006. Designed & produced by CCLRC, Daresbury Laboratory, Media Services. 01925 603831

Welcome to the 2006 issue of CSE Frontiers, an Annual Report of CCLRC's Computational Science and Engineering Department.

Our Department's overall aim is to ensure that UK researchers benefit from the best computational methods, supporting them through research and collaboration, theory and software development, facilities and training. "Frontiers" showcases some of the research and development aspects of this activity over the last year.

Collaboration is central for us. This philosophy, enshrined in the Collaborative Computational Projects (CCPs) and in the UK's consortium approach to supercomputing, was favourably noted in last year's "International Review of Research Using HPC in the UK" (EPSRC, December 2005). Almost all of the articles in this issue describe work done with UK and international collaborators, theoretical and experimental, and this is how it should be. This year we also highlight our work with Astra-Zeneca in the article on cyclodextrins by Chin Yong, Bill Smith and Clive Washington. In this and other calculations we aim to have an impact on problems that are always real, often fundamental and usually big. This year's list is wide-ranging: nanoparticles in supercritical fluids, magnetic semiconductors, aluminium fluoride catalysts, materials undergoing radiation damage, the dynamics of helicopter rotor wakes, superconductivity and valence transitions in correlated solids, microfluidics in "digital" devices, atomic aspects of astrophysical and laser plasmas, and so on. Extending the scope and reality of simulations and the software on which

they depend is a continuing goal, as seen in the articles of Paul Sherwood and of Martyn Winn and their collaborators in computational chemistry and biology. Such calculations are challenging – at the frontiers of the computational sciences in fact. We may be entering a new era for modelling and simulation, one requiring a closer match of applications and algorithms to a diverse range of specialised machine architectures. Our own portfolio of systems, encompassing both robust production platforms like HPCx and experimental testbed machines, is always undergoing upgrade or evaluation as the articles of Steve Andrews and Richard Wain and colleagues illustrate. At the core of both functionality and performance lie the algorithms and the code itself. The articles by Nick Gould on the numerical mathematics of large-scale optimisation and by Ian Bush and colleagues on performance tuning of materials simulation codes for HPCx therefore exemplify crucial activities for us.

I hope you'll find plenty to interest you in this issue of Frontiers. If you do, I hope you'll look at our web site http://www.cse.clrc.ac.uk, and come and see us at the many events, workshops and meetings we hold at Daresbury, RAL and elsewhere around the UK. And if you'd like to collaborate with us, we'd be delighted.

Parl Dol

Professor Paul Durham, Director of CSE

Iron, Nickel and Tin in Astrophysical and Laser-Produced Plasmas: Atomic and Molecular Physics at Daresbury in 2006

Electron-impact scattering cross sections and related effective 'collision strengths' are of crucial importance in the interpretation of spectra of low-ionization stages of elements in the iron-peak region of the periodic table (particularly Fe, Co and Ni) which occur in astrophysical plasmas. For example, electron-impact excitation lines for singly ionized iron (Fe II) have been observed in objects as diverse as the chromospheres of the sun and other cool stars, supernovae and active galactic nuclei. Nickel, though considerably less abundant than iron, is an important contributor to the emission spectra of many stars, with numerous features due to lowionization states of this element. Accurate calculation of the scattering data is made difficult by the open d-shell structure of these ions which gives rise to many lowlying target states that are strongly coupled. Collision strengths must be first calculated over a dense grid of energies sufficient to produce the similarly dense resonance structure, then thermally averaged to produce the effective collision strength for a given temperature.

Figure 1. Hubble Space Telescope images of the very young 'Boomerang Nebula' (1998, 2005): the latter 'Advanced Camera' picture clearly captures the scattered light.

In the laboratory, laser-produced plasmas containing tin are being investigated as possible extreme ultraviolet (EUV) sources for next-generation nanolithography tools. Modelling of plasma behaviour is required to understand and facilitate both debris mitigation (for example, atoms and ions of tin depositing on the collection optics) and, very importantly, the optimization of conversion efficiency (the ratio of EUV energy output to laser energy). Establishing the emission and absorption characteristics of ionized tin is of crucial importance in these investigations.

The ab initio suite of electron-atom codes PRMAT has been developed at Daresbury to handle these scattering calculations. The R-matrix method essentially divides configuration space into two regions: a sphere containing the target atom or ion in which interactions with the scattering electron are treated in full 'manyelectron' detail and the remaining 'external' region in which the problem can be simplified to a (still complicated) series of coupled differential equations for the scattering electron. The parallel features of the PRMAT package have been carefully designed to distribute and minimize memory use as well as to allow fast and efficient computation. Thus, inclusion of many hundreds or in some cases, thousands of real and virtual interacting target states needed for open d-shell calculations is feasible with the package (each target state is itself represented by a 'configuration interaction' or CI expansion over many-electron 'configuration state functions' built from combinations of quantum-mechanical spin-orbitals). The package has been upgraded to allow the inclusion of relativistic effects important for heavy ions: this new package is currently being subjected to rigorous testing.

The atomic, molecular and optical physics group collaborates with universities as part of the computational project CCP2. For example, the PRMAT code is being used in collaborations with The Queen's University of Belfast to study excitation of Fe II (also Fe III, Fe IV and Fe V) and Ni V ions. The ongoing Fe II calculations reported in previous editions of Frontiers have been reasonably converged non-relativistically and now include relativistic effects. The studies of Ni V excitation include 184 target states (thus 550 scattering channels for a given scattering symmetry: the innerregion Hamiltonian matrices to be diagonalized include additional 'bound-bound' components and are much larger). Of particular interest here is the formation of 'two-particle-one-hole' states, intermediate resonant states in which the dominant configuration has two electrons in 'excited' orbitals leaving one electron 'missing' from the Ni V initial-state orbitals (more rigorously, the CI allows for distortion of the initialstate and excited-state orbitals due to electron correlation effects; note also that the electrons are identical fermions and all states are fully antisymmetrized). The work has demonstrated the existence of these states and their important effect on the collision strengths.

The group is also collaborating with University College, Dublin to study excitation of Sn VII. Sn VII out-of-band emission lines should prove valuable as a diagnostic tool for Sn plasma sources. Sn VII valence electrons have the basic ground state configuration 4p⁶4d⁸. Collision strengths have been calculated and the target states that provide the major contributions elucidated (in particular, certain doubly-excited states). Two-particleone-hole states also have an important though more subtle role to play in these collisions.

Along with the PRMAT work, the group has begun a new collaboration to advance the CCP2 flagship project on annihilation rates for positron interactions. Extremely high rates for interactions with various organic molecules are thought to arise from a resonant process involving vibrational motion of the nuclei of the

Figure 2. A tin-based laser-produced plasma.

molecule. The collaboration with Nottingham University is to produce highly efficient code to describe positronelectron interactions directly in scattering systems such as e^+-H_2 or He–*anti*-H. Annihilation rates will provide benchmarks for a generalized positron-molecule *R*-matrix code being developed at University College, London. The attractive positron-electron interactions are included explicitly as functions of the positron-electron distance in addition to the usual CI expansions for the electrons.

Other ongoing calculations relevant to astrophysics include e–O scattering and e–H scattering and ionization. Group members are authors of a major review of the '*R*-matrix theory of atomic, molecular and optical processes': P. G. Burke, C. J. Noble and V. M. Burke, Advances in Atomic, Molecular and Optical Physics 54 (2007), in press.

Figure 3. A typical collision strength spectrum for NiV (various target expansions).

V. M. Burke, C. J. Noble, M. Plummer and A. G. Sunderland, CSED, Daresbury Laboratory; M. P. Scott, C. A. Ramsbottom and P. G. Burke, The Queen's University of Belfast; M. A. Lysaght, University College Dublin.

References:

'On the role of 'two-particle-one-hole' resonances in electron collisions with Ni V', M. P. Scott, C. A. Ramsbottom, C. J. Noble, V. M. Burke and P. G. Burke, Journal of Physics B, 39 (2006) 387-400.

'Electron-impact excitation of Fe II', C A Ramsbottom, C. J. Noble, V. M. Burke, M. P. Scott and P. G. Burke, Journal of Physics Conference Series (2006), 'HCI 2006: 13th international conference on the physics of highly charged ions'. 'Configuration interaction effects in low-energy electron collisions with Sn VII', M. A. Lysaght, V. M. Burke, C. J. Noble, M. P. Scott, A. Hibbert, P. Dunne and G. O'Sullivan, Journal of Physics B, (2006) submitted.

'When antimatter attacks', E. S. Reich, New Scientist, 24 April 2004, 34-37. 'Explicitly Correlated Wavefunctions', eds. E. A. G. Armour, J. Franz and J. Tennyson, CCP2 Workshop Proceedings (2006), CCP6 (on behalf of CCP2).

Helicopter Rotor Wake Simulations

The aerodynamics of helicopter rotor blades is one of the most interesting and challenging problems facing aerodynamicists. Predicting the flow around a blade is a totally different problem to an aircraft wing, which has an (almost) uniform constant speed flow over it, and where what happens downstream of the wing is not that significant. That is not the case for a rotor blade, where the speed of the blade varies from very low at the root, to very high at the tip, and in forward flight the effective velocity the blade sees is different at every point around the azimuth (revolution). Hence, the flow is highly three-dimensional and unsteady. Furthermore, and most significantly, each blade moves into fluid that has already been disturbed by the previous blade(s). The loud buzzing, vibrating, sound caused by helicopters is due to the wake, and particularly the strong vortex from each blade tip, shed from one blade being hit by the following one. The accurate prediction of this blade-vortex interaction (BVI) is essential for both civil helicopters, to attempt to reduce "noise pollution", and military, to avoid detection.

Until recently, numerical simulation of these flows was rare, due to excessive cost, but rapidly increasing computer power and code capability has meant that this is now possible. Computational methods for fluid flow simulation involve filling the physical domain of interest with a computational grid, i.e. filling the domain with a number of cells, each one having the solution stored in it. The solution is the local values of flow variables, i.e. density, pressure, velocity, etc. Using applied techniques from mathematics and physics, methods can be developed which march the solution forward in time, starting from an initial guess and iterating until the solution converges (stops changing), and these methods can be used to simulate steady or unsteady flows. The numerical approximations have a truncation error associated with them, i.e. the difference between the real equations and the numerical approximation, and the error is a function of the grid spacing. Hence, the finer the computational grid used, the more accurate the solution. The error in the numerical approximation is known as dissipation, and so the coarser the grid the more dissipation is added to the flow and, hence, the interesting parts of the flow are "smeared out".

Unlike most fixed-wing cases, rotor simulations require extremely fine meshes *throughout* the grid otherwise the wake from each blade is dissipated by the numerical scheme before it hits the next one. Furthermore, to capture the wake over many turns, particularly for hovering rotors where a helical wake develops beneath the blades, requires a large number of iterations/timesteps. Hence, rotor flow simulation requires many timesteps on a very fine mesh, and this leads to huge computational challenges.

Dr Chris Allen from the University of Bristol has developed both flow-solver and grid generation methods for rotary wing applications. The code is a structured multiblock, upwind solver, using implicit time-stepping for unsteady calculations, with explicittype time-stepping within each real time step. Multigrid acceleration is used to improve convergence. The multiblock approach adopted is ideal for parallelization, and the code has been parallelized using non-blocking MPI communication. Further optimization and performance analysis of the code on several high-end computing platforms has been undertaken by CSED.

Figure 1. Vorticity shading for 1, 8 and 32 million points (left, centre, right) respectively. Vorticity is a measure of local rotation in the fluid.

4

A wake grid dependence study is shown here for a fourbladed rotor in forward flight. The computational domain, and block structure was kept constant and grids of size 1, 2, 4, 8, 16, and 32 million points were generated. Figure 1 shows vorticity shading on selected grid planes, for the 1, 8, and 32 million point grids (the scales are the same in each). A comparison of the results for the different grid resolutions demonstrates the effect of numerical dissipation on the problem.

The parallel performance of the original code was generally quite good. However, profiling the parallel code with performance analysis tools identified certain inefficiencies associated with the data exchange across block boundaries for both planes of solution. Global synchronization points have been removed from the program, thereby facilitating more efficient data transfer mechanisms and also allowing tasks to be scheduled more efficiently by the system. Excellent scaling has been achieved, and the code has been awarded a gold star for its performance on the UK's National Supercomputing Service HPCx (an IBM p5-575-based system).

Figure 2. Performance of the helicopter flow simulation program on high-end computing platforms

The performance of the code on four different high-end computing platforms is shown in Figure 2. The benchmark runs were performed on a 4-bladed rotor case in forward flight. Simulations were run using a mesh density of 20M points. The code scales very well on all the systems, but parallel performance is best on the SGI Altix 3000 system. This is largely due to the excellent memory bandwidth of the Altix combined with good message passing performance from the underlying NUMAflex architecture. A fuller description of program performance analysis and optimisation can be found in [1].

The next stage in this research is to include the helicopter fuselage in the simulation, as the inclusion of the interaction of the fuselage flow with the rotor blades and the downwash from the blades on the fuselage is essential for accurate aerodynamic design. A highly efficient parallel flow-solver is essential as these calculations will require complex gridding structures and extremely fine meshes to capture the wake interactions.

> C. B. Allen, Department of Aerospace Engineering, University of Bristol A. G. Sunderland, CSED, Daresbury Laboratory

Reference:

 [1] 'Application of a Parallel Rotor CFD Code on HPCx', C. B. Allen, A. G. Sunderland,
R. Johnstone, Aero Journal (to appear).

A Parallel Multi-Band k.p Code for the Electronic Structure of Semiconductor Quantum Dots

An efficient parallel implementation of kppw, a multi-band $\mathbf{k} \cdot \mathbf{p}$ code, has recently been developed by members of CSED. This program is used to study the electronic structure and optical properties of zinc blende structure semiconductor quantum dots [1]. In a Quantum Dot (QD) the spatial confinement of electrons and holes along all three dimensions leads to a discrete energy spectrum and a delta-function atomic-like density of states, with sharp optical absorption lines. In that sense, a semiconductor QD can be recognized as an artificial solid-state atom.

Figure 1. Speed-up curves on HPCx for two different ranks of Hamiltonian matrix in k space: the solid symbols correspond to $N_k = 37800$ and open symbols correspond to $N_k = 25688$. The square symbols represent the time taken generating the distributed Hamiltonian matrix. The circle symbols represent the time taken calculating eigenpairs from the whole spectrum by PZHEEVD, and triangles represent the time taken calculating a subset of eigenpairs within the energy band ($E_{CBM} - E_{VBM}$ of a barrier material) by PZHEEVX.

The ability to control systematically the quantum states, so called "band structure engineering", is essential for a variety of applications that range from semiconductor optoelectronics, via spintronic devices, to quantum information processing and quantum computing. Most recently, semiconductor QDs have found applications in nanostructure–bioscience complexes, and as a scanning probe of DNA molecules. The method uses an expansion of the electronic wavefunctions by plane waves, thereby allowing calculations on arbitrary shaped quantum dots. This can be achieved for very complicated shapes by simply taking linear combinations of the Fourier transform of much simpler ones, for instance a box, a cylinder or a cone. A further advantage is that this also allows simple expression of other important quantities, such as the strain elements and the piezoelectric field [2].

Recent improvements to the code have introduced efficient parallelisation of the key stages of the calculation, leading to excellent parallel performance for up to 512 processors, Figure 1, on the HPCx system [1], [4]. These optimisations have enabled both more detailed calculations, for instance including more bands as in the dilute nitrogen quantum dot structures [3], or to perform calculations on bigger quantum dots/supercells structures keeping the same quality of calculation. The performance of the program is demonstrated on the pyramidal shape InAs/GaAs, dilute nitrogen InGaAsN, Figure 2, and the recently found volcano-like InAs/GaAs quantum dot systems, Figure 3.

In order to take advantage of ScaLAPACK parallel numerical library routines the Hamiltonian matrix data is generated using a 2D block-cyclic data distribution [1]. Although this method of data distribution appears relatively complex it ensures good load-balancing across the processors and it allows the dense matrix algorithms from the ScaLAPACK library to take advantage of highly optimised library routines, so maximising performance.

The heart of the calculation is a matrix diagonalization. It is therefore important that the most suitable algorithm is chosen for this stage. Currently the standard way to perform diagonalizations in parallel is to use the ScaLAPACK library, which provides a number of diagonalization routines each of which has different options and different performance characteristics. As the physics of interest only depends upon a few states near the band gap, the ScaLAPACK routine PZHEEVX appears to be the most suitable, as this allows a subset of the eigenpairs to be calculated. However we also tested the ScaLAPACK routine PZHEEVD which is restricted to finding all the eigenpairs of the matrix, but in doing so invokes a highly efficient divide-andconquer algorithm. The parallel performance of the two methods is shown in Figure 1. These results are from HPCx. For the PZHEEVD based calculations, the timings from two problem sizes are given here, involving the generation and diagonalization of matrices of the order 25688 (open symbols) and 37800 (solid symbols). For the PZHEEVX calculations the timings from the order 25688 system are shown. In this latter example we can restrict the required eigenpairs to the energy interval between the conduction band minimum and the valence band maximum of the matrix material. In this range we require 128 eigenpairs for the 25688 order problem. For the smaller problem size, computing 25688 eigenpairs with PZHEEVD takes 357 seconds on 256 HPCx processors, computing 128 eigenpairs with PZHEEVX on 256 processors takes 209 seconds. Figure 1 also demonstrates that the generation of the distributed Hamiltonian matrix and both eigensolver approaches scale very well with number of processors used on HPCx [4].

The modular approach adopted in the design of the kppw code facilitates code expansion, and a number of new features are planned. Studies of wurtzite structure QDs will soon be possible, and we also plan to introduce more bulk band states around the Γ -point, different orientations of the QD about high Miller indices planes, a hybrid atomistic-continuum model description of impurity states in the QD's and further integration with DLV [5], a visualisation software package developed within CSED.

Figure 2. The wave-function squared of the three lowest CB and two highest VB states in (top row) N-free In_{0.7}Ga_{0.3}As/GaAs QD and (bottom row) In_{0.7}Ga_{0.3}As_{0.96}N_{0.04}/GaAs QD. Isosurfaces are plotted at 25% (transparent) and 75% of the maximum charge density. A better electron confinement is visible in the sample with 4% Nitrogen.

Figure 3. The piezoelectric field in the quantum volcano. Blue and red surfaces are at -40 meV and +40 meV respectively. Maximum piezoelectric field value is 90 meV.

References:

[1] S. Tomic, A. G. Sunderland, and I. J. Bush, J. Mat. Chem. 16, 1963 (2006).

[2] S. Tomic, P. Howe, N. M. Harrison, and T. S. Jones, J. Appl. Phys. 99, 093522 (2006).

[3] S. Tomic, Phys. Rev. B 73, 125348 (2006).

[4] www.hpcx.ac.uk

[5] B. Searle, Comp. Phys. Comm. 137, 25 (2001).

S. Tomic, A. G. Sunderland and I. J. Busł CSED, Daresbury Laboratory

Parallel Performance Optimization on HPCx: Materials Chemistry

The effective exploitation of current high performance computing (HPC) platforms relies on the ability of the present generation of application codes to make use of these platforms and their components. HPCx (pictured) is the UK's largest national academic HPC resource, operated by a consortium HPCx UoE Ltd consisting of the University of Edinburgh, Edinburgh Parallel Computing Centre (EPCC), CCLRC Daresbury Laboratory and IBM. HPCx UoE Ltd is funded by the UK Research Councils with the Engineering and Physical Sciences Research Council providing the largest contribution. In 2006 HPCx enabled calculations using up to 1024 processors working in parallel.

Evolving national and international HPC resources allow scientists to plan more ambitious calculations and simulations to achieve more realistic modelling of nature. To exploit these resources to maximum benefit and efficiency scientists should have some understanding of how these systems differ from local PCs and small clusters. As part of the HPCx service the 'terascaling' team, made up of CSE and EPCC staff, exists to further this understanding and works to achieve these aims. Application codes in many disciplines have benefited substantially from optimization performed by team members in collaboration with code authors. The team also provides expertise in new developments in HPC which will allow code improvements to be carried through and extended to make the best use of future resources. On HPCx the team also works to ensure that system changes and system software upgrades are taken fully into account when optimizing code in order to achieve best performance. HPCx is a 'clustered-SMP' system with nodes containing groups of processors which may share memory and with very fast intra-group communication between individual parallel tasks. The nodes are connected to each other by IBM's High Performance Switch. HPCx Phase 2A has been operating throughout 2006 with 16-processor nodes built from IBM Power5 processors, compared to the previous Phase 2 system with 32-processor nodes built from Power4+ processors. Phase 2A nodes have double the memory per processor of Phase 2 nodes and a generally much more efficient internal cache structure.

Calculations on HPCx cover a large range of lengthscales ranging from the structure and dynamics of atoms to simulations of large-scale climate systems.

Scientists studying computational chemistry and theoretical materials science, performing both detailed quantum-mechanical and classical molecular dynamics simulations of solids, liquids, gases and interactions between phases, make up a large user-community on HPCx. This year a special edition of the Journal of Materials Chemistry was published highlighting the work of this community. As well as the impressive array of scientific reports the volume included a review article written by the terascaling team describing the ways in which various codes have been successfully optimized to perform well on the massively-parallel HPCx system. A critical analysis was presented of how codes should be written or modified to approach the ideal of 'capability computing' in which 1000-plus processors on HPCx are used efficiently and necessarily in a single calculation. This is a particularly complex achievement in much of materials chemistry as the performance of the code will often be highly dependent on the physical structure of the real-life system being studied.

Figure 1. HPCx Phase 2A.

Most of the main materials chemistry applications codes tend to be either molecular dynamics codes with classical interaction potentials between atoms or quantum mechanical electronic structure codes which determine the interactions from first principles. The latter codes may expand electronic wavefunctions in a set of (Gaussian) orbitals or as plane waves. The codes analyzed in the review article include, for example, the orbital-based codes SIESTA, CRYSTAL and GAMESS-UK, the plane-wave codes CASTEP and VASP and the classical molecular dynamics code DL_POLY_3. Strategies for parallel optimization depend on the type of code. For the orbital-based applications, efficient algorithms and data-distribution for large matrix operations and diagonalizations are vital. For planewave codes, various strategies for dealing with

collective communications have been developed. These are used for data reorganizations in frequently-occurring three-dimensional Fast Fourier Transforms (3d-FFTs) and make the best use of HPCx's clustered-SMP structure, exploiting the differing intra- and inter-node communication speeds. DL_POLY_3 has benefited from the use of a domain-decomposition distributed data scheme combined with a locally-written 3d-FFT routine that maps directly onto the same data-distribution, thus avoiding large and expensive data redistributions required for standard 3d-FFTs (very important given that simulations with millions of atoms may be run using DL_POLY_3, now scaling to 1024 processors).

The review article also analyzed performance changes brought about by the introduction of HPCx Phase 2A. In general these were very beneficial as expected from the enhanced memory-management capabilities. Some molecular dynamics codes did not initially show an improvement in performance due to certain peculiarities of Power5 cache latency. However this

Figure 2. CASTEP results.

problem has now been overcome with the use of 'simultaneous multi-threading' in which each processor handles two tasks.

Work on optimization in general has continued in readiness for Phase 3 in which the overall size of HPCx will double. As an example we show current timings for a CASTEP electronic energy benchmark test of a slab of Al₂O₃ with 270 atoms in the periodic boundary condition 'supercell'. The overall improvement over the life of HPCx is due to both systems improvement and code improvements obtained by code authors and the terascaling team working together: further improvements in CASTEP are ongoing. We also show how the scaling of CRYSTAL has been improved for medium sized systems. Very recent work has focussed on exploiting parallelism across spin and 'Brillouin zone sampling k points', in particular addressing load balancing. Sample results are shown for pyrope, a mineral which is a member of the garnet family. The 'supercell' consists of 160 atoms. It can be seen the revised code is around 40% faster on 128 processors.

Figure 3. CRYSTAL results.

References:

M. Plummer, I. J. Bush and M. F. Guest, CSED, Daresbury Laboratory; K. Refson, CSED, Rutherford Appleton Laboratory

Full quantitative details of the HPCx system and the codes and optimizations described here may be found at http://www.hpcx.ac.uk and in: 'Terascale materials modelling on highperformance system HPCx', by M. Plummer, J. Hein, M. F. Guest, K. J. D. Mellow, I. J. Bush, K. Refson, G. J. Pringle, L. Smith and A. Trew, J. Mater. Chem. 16, 1885-1893 (2006). 'A DAFT DL_POLY distributed memory adaptation of the Smoothed Particle Mesh Ewald method' by I. J. Bush, I. T. Todorov and W. Smith, Comput. Phys. Commun. 175, 323-329 (2006).

'An Investigation of Simultaneous Multithreading on HPCx', by A. Gray, J. Hein, M. Plummer, A. Sunderland, L. Smith, A. Simpson and A. Trew, http://www.hpcx.ac.uk/research/hpc/technical_ reports/HPCxTR0604.pdf 'A Performance Comparison of HPCx Phase 2a to Phase 2', A. Gray, M. Ashworth, S. Booth,

J. M. Bull, I. J. Bush, M. F. Guest, J. Hein, D. Henty, M. Plummer, F. Reid, A. G. Sunderland and A. Trew,http://www.hpcx.ac.uk/research/hpc/ technical_reports/HPCxTR0602.pdf.

Dilute Magnetic Semiconductors from First-Principles Calculations

Semiconductor Spintronics is the novel electronics which, besides the electron charge, manipulates also the electron spin, with the aim to realize in one electronic device both the information processing and storage (see Figure 1).

Figure 1. The schematic illustrates the importance of both the electron charge and spin for the new spin electronics, and the usefulness of photons.

Recently discovered dilute magnetic semiconductors (DMS) are without a doubt the most promising class of materials for the new spintronics devices. These are based on traditional semiconductors, but are doped with transition metals (TMs) instead of, or in addition to, electronically active elements. In particular, Mn- or Co-doped group III-V and II-VI semiconductors have been extensively studied both experimentally and theoretically. The most important feature of these materials is carrier mediated ferromagnetism which can be easily controlled with voltage. The advantage is that, unlike the conventional ferromagnets, DMS are compatible with semiconductors and can be used as efficient sources for spin injection.

The best understood dilute magnetic semiconductor is Mn-doped GaAs [(Ga,Mn)As] with the highest known Curie temperature, T_C, of about 170 K. The magnetic interaction amongst the spins of Mn²⁺ ions is mediated by holes (p-type) originating from the host material valence band. This is what both the experiments and theory, based on the so-called p-d exchange Zener model (which assumes the S=5/2 Mn spins, interacting via intrinsic holes [1]), agree on when the concentration of Mn impurities is below 10 atomic %. Unfortunately, the situation is not so clear for the other semiconductors from the same group, namely GaP and GaN. Similarly, for Mn- or Co-doped ZnO, there are important issues to be resolved and in particular, whether the p-type ferromagnetism is at all possible.

These important questions are best answered with the first-principles calculations. Here we apply the selfinteraction corrected local spin density (SIC-LSD) method to address the issues of the valence state of the TM elements in these materials, the type of the magnetic interaction and the magnitude of their T_{C} . The advantage of this approach is that it treats on an equal footing both itinerant and localized electrons, and leads also to a definition of nominal valence, $N_{val} = Z -$ N_{core} – N_{SIC}, where Z is the atomic number (25 for Mn), N_{core} is the number of core electrons (18 for Mn), and NSIC stands for the number of localized (self-interaction corrected) electrons (five 3d electrons for a divalent Mn ion or four 3d electrons for a trivalent Mn ion). This valence is equal to the integer number of electrons available for band formation, and differs from the ionic valence in that not only the s- and p-electrons are considered (e.g. for a trivalent Mn ion, four 3d electrons are treated as localized, whilst two 4s electrons and one 3d electron determine the valence).

With the SIC-LSD approach we have studied a variety of valence configurations of Mn impurities (3.125 at. %) in GaN, GaP and GaAs semiconductors, in the zinc blende crystal structure [2]. For GaAs, not only have we been able to predict correctly the divalent state of Mn ion, but also to obtain the fully spin-polarized valence band with holes of predominantly As character, oppositely polarized to the Mn spins. In addition, we have calculated the Mn magnetic moment to be 4.5 Bohr magnetons, in excellent agreement with the XMCD experiment [3]. We have also calculated the so-called *p*-*d* exchange of 67 meV/nm³, in good agreement with the experimental value of 54±9 meV/nm³. We have shown that the Mn impurity in GaAs is very shallow, as seen in Figure 2, where we display the spin moments in all the spheres of a large supercell of 128 sites (including empty spheres, I), with which the compounds are modelled. For GaAs all our findings give full support to the *p*-*d* exchange Zener model, which is however not the case for the other semiconductors. Our SIC-LSD calculations predict correctly a transition from the divalent Mn (d^5) in GaAs to the trivalent Mn (d^4) in GaP and GaN. Regarding the spin magnetic moments for all the spheres in the supercells, we can see in Figure 2 that the Mn impurity is very localized in GaN, and the holes are polarized in parallel to Mn spins. In GaP the situation is somewhat intermediate between GaAs and GaN. Similarly to GaAs, the holes are oppositely polarized to the Mn spins, but the Mn spin magnetic moment of 3.9 Bohr magnetons is smaller than 4 Bohr

magnetons, like in GaN, where it is equal to 3.6 Bohr magnetons. Also, (Ga,Mn)P is on the border line between divalent and trivalent Mn, however with a slight preference for the trivalent state, in agreement with electron spin resonance experiment of Kreissl et al [4].

The ground state of (Ga,Mn)N compound is calculated to be insulating, meaning that the *p*-type ferromagnetism in this system could only be possible with the co-doping of extrinsic holes. By studying both the d⁴ and d⁵ configurations of Mn in these semiconductors, we have been able to access both the weak and strong coupling limits of the exchange interactions, with the d⁴ configuration describing the strong coupling scenario in (Ga,Mn)As, but a weak coupling (the ground state) limit in the (Ga,Mn)N semiconductor, and vice versa for the d⁵ configuration. The corresponding estimates of the Curie temperatures for the doped *p*-type compounds are 190 K for (Ga,Mn)As, 157 K for (Ga,Mn)P and 283 K for (Ga,Mn)N.

For the wurtzite ZnO doped with TM element, where TM=Mn or Co, our calculations [5] have established that the *p*-type ferromagnetism is not possible without co-doping extrinsic holes with concentration exceeding that of TM impurities. The reason being that (Zn,TM)O is insulating, with TM ion in a divalent state. At the same time, the calculated donor level, (+/0), defined by the position of the Fermi energy at which the formation energy of a neutral impurity is equal to the formation energy of a charged impurity (see Figure 3), is situated in the gap, in agreement with experiments. Thus, when co-doping these semiconductors with e.g. Nitrogen, the holes provided by the latter get compensated by delocalisation of one of the TM d electrons, changing the TM valence from divalent to trivalent, and leaving the system insulating, with no free carriers to maintain magnetic interactions. Only by making the Nitrogen concentration larger than that of TM impurities, can one prevent the spin polarized holes, at the top of the valence band, to be compensated, as no further TM d electrons want to delocalise. Considering, however, that co-doping Nitrogen into these semiconductors is notoriously difficult in practice, one may have to concentrate on the *n*-type possibility instead. These issues of co-doping and the *n*-type vs. *p*-type ferromagnetism in the dilute magnetic semiconductors,

T. C. Schulthess, L. Petit, G. M. Stocks, Oak Ridge National Laboratory; Z. Szotek, W. M. Temmerman, CSED, Daresbury Laboratory; W. Butler, University of Alabama, Tuscaloosa; A. Svane, University of Aarhus, Denmark; A. Janotti, University of California, Santa Barbara. and in particular in ZnO- and GaN-based systems, remains an active area of research. The most important outcome of our first-principles study is elucidating the fact that the assumptions of the p-d exchange Zener model apply only strictly to (Ga,Mn)As semiconductor.

Figure 2. Spin magnetic moments in atomic spheres of the supercells for (Ga,Mn)As, (Ga,Mn)P and (Ga,Mn)N. The correct magnitude of the Mn spin moment is not reproduced in the figure as the induced spin moments in other spheres are dramatically smaller. Here V is group V element.

Figure 3. A schematic plot of the formation energy as a function of Fermi energy, between the valence band maximum (VBM) and conduction band minimum (CBM), for (Zn,TM)O. Illustrated is the link between the localization-delocalization picture and the TM donor level, (+/O), in the band gap. Only the donor level between the positively charged TM³⁺ and the neutral TM²⁺ valence configuration is indicated.

References:

[1] T. Dietl et al., Science 287, 1019 (2000).

[2] T. C. Schulthess, W. M. Temmerman, Z. Szotek, W. H. Butler & G. M. Stocks, Nature Materials 4, 838 (2005).

[3] K. W. Edmonds et al., Phys. Rev. B 71, 064418 (2005).

[4] J. Kreissl Phys. Rev. B 54, 10508 (1996).

[5] L. Petit, T. C. Schulthess, A. Svane. Z. Szotek, W. M. Temmerman & A. Janotti, Phys. Rev. B 73, 045107 (2006).

Superconductivity under Extreme Pressure

Some materials, even if looking harmless under normal conditions, can exhibit a surprisingly complex behaviour when exposed to extreme pressures. Putting the squeeze on a material not only gives rise to structural phase transitions, but often also affects other properties, such as their magnetic structure or their tendency to superconduct. Studying the influence of pressure on superconducting properties can give some important insights into the underlying pairing mechanism, but also poses a stringent test on the predictive power of an ab initio method. A prominent example of materials showing strikingly different behaviour under pressure is aluminium and lithium. Both materials are considered to behave like 'simple metals'. Al is superconducting at ambient pressure with a critical temperature $T_c = 1.18$ K, which rapidly decreases with an applied pressure. Li, on the other hand, does not show superconductivity below about 20 GPa and develops superconductivity only at higher pressures. Also structurally, it displays a rather complex behaviour. Below 77K and at zero pressure it shows a series of martensitic transitions between competing closed packaged structures. From 7.5 to 70GPa it undergoes several structural transitions, which indicates strong electron-phonon couplings.

Superconductivity in these materials is caused by a phonon-mediated attraction between the electrons, and is qualitatively well understood in terms of the BCS theory, introduced by Bardeen, Cooper and Schriefer, and its generalization to strong electron-phonon coupling, the Eliashberg theory. The latter approach provides a good description for the electron-phonon interaction, but the treatment of the Coulomb repulsion remains very difficult. In practice, the Coulomb interaction is replaced by a phenomenological pseudo potential μ^* , which is adjusted to reproduce the experimental critical temperature. Therefore, the Eliashberg theory cannot be considered an *ab initio* theory.

The standard tool for material specific first principles calculations of normal-state properties, such as the geometrical or magnetic structure, is density functional theory (DFT). It provides a mapping of the physical systems onto an auxiliary system of non-interacting electrons moving in an effective field - the Kohn-Sham (KS) system - which reproduces the density and the total energy of the fully interacting system. It may be said that DFT is the main working horse for virtually all computational material science, and in particular for the

study of high pressure phases. DFT is a very versatile theory and can easily be generalized to ordered phases, such as spin-polarized or superconducting states. In Frontiers 2005 we reported on a generalization of DFT for superconductors (SCDFT) which treats the electronic correlations and the electron-phonon interaction on the same footing [1]. There, the Kohn-Sham system is defined as a system of non-interacting electrons in an effective field and an effective pairing field, which induces a superconducting state. In addition, the KS system contains the nuclei which interact via an effective many-body interaction, but do not interact with the (KS) electrons. The Kohn-Sham equations take the form of Bogoliubov-de Gennes (BdG) equations for the electrons, and a Born-Oppenheimer like equation for the nuclear motion.

Here we report on an application of SCDFT to study the influence of high pressure on the superconductivity in aluminium and lithium to make contact with experiment. In addition we also study potassium and predict superconductivity for pressures above 20 GPa [2]. The ground-state calculations were performed with

Figure 1. Calculated and experimental critical temperatures for Al (upper panel), K (middle panel) and Li (lower panel). SCDFT refers to the ab initio calculations and the curves marked as McMillan represent results from the McMillan formula with the calculated coupling constants λ and different values for μ^* . the PWSCF pseudo-potential plane-wave code, and the electron-phonon couplings were obtained using densityfunctional perturbation theory. Figure 1 shows the calculated critical temperatures, compared to the experimental values. The calculations, which are completely parameter-free, show excellent agreement with experiments for Al and Li in the pressure range where the fcc structure is stable. No experimental values are available so far for K and our results provide a prediction of superconductivity with a critical temperature of about 2 K within the experimental stability range of the fcc structure.

We relate the appearance of superconductivity in Li and K to an incipient phase transition, which gives rise to phonon softening and very strong electron-phonon coupling that then leads to the unusually high transition temperatures. The different behaviour of Al on one side and Li and K on the other can be understood by analysing the Eliashberg function as a function of pressure, shown in Figure 2. In Al, the phonon frequencies increase as the pressure rises, corresponding to the normal stiffening of phonons with increasing pressure. In addition, the height of the peaks in the Eliashberg spectral function $\alpha^2 F(\Omega)$ decreases with increasing pressure. These factors contribute to a decrease of the overall coupling constant λ (see insets in Figure 1) and, consequently, of the critical temperature T_c. For alkali metals the situation is completely different: due to the incipient phase transitions, a phonon softening at low frequencies increases the value of λ in both materials. This phonon softening is a consequence of Fermi surface nesting, indicated in Figure 3, which in turn originates from a pressure induced change from an s-p character to a d character of the states at the Fermi energy. However, the different topology of the Fermi surfaces and the different range of the phonon frequencies sets the critical temperature much higher in lithium with respect to potassium. Figure 3 displays the Fermi-surface of Li and indicates the nesting vectors, giving rise to the high values of λ .

M. Lüders, CSED, Daresbury Laboratory; A. Sanna, C. Franchini, S. Massidda, University of Cagliari, Italy; G. Profeta, A. Continenza, University of l'Aquila, Italy; A. Floris, M. A. L. Marques, N. N. Lathiotakis, E. K. U. Gross, Freie Universität Berlin, Germany. The results obtained for Al and Li are in very good agreement with experiment, and account for the opposite behaviour of these two metals under pressure. Furthermore, the increase of T_c with pressure in Li is explained in terms of the strong electron-phonon coupling, which is due to changes in the topology of the Fermi surface, and is responsible for the observed structural instability. Finally, our results for fcc-K provide predictions interesting enough to suggest experimental work on this system.

Figure 2. Eliashberg function $\alpha^2 F(\Omega)$ for Al and Li for various pressures. The different behaviour of T_c as function of pressure can be related to the strong increase of $\alpha^2 F(\Omega)$ of Li at low frequencies. The behaviour of K is similar to the one of Li.

Figure 3. Fermi-surface of Li at a pressure of 28.6 GPa. The red colour indicates high values of the energy gap. The cut through the Fermi-surface on the right indicates the nesting vectors, which give rise to the large values of the electron-phonon coupling λ .

References:

[1] See also: M. Lüders et al, Phys. Rev. B 72, 024545 (2005), M. Marques et al, Phys. Rev. B 72, 024546 (2005).

[2] G. Profeta et al, Phys. Rev. Lett. 96, 047003 (2006), A. Sanna et al, Phys. Rev. B 73, 144512 (2006).

Valence Transitions in Nickelates

With the proliferation of portable electronic devices, and the quest for vehicles powered by electricity, the development of high energy density batteries remains a top priority research and development area. The metal oxide compound LiNiO2 is of considerable interest due to its application as a cathode material in lithium-ion rechargeable batteries. The crystal structure consisting of layers of NiO₂ slabs, built from edge sharing NiO₆ octahedra, and separated from each other by a layer of Li cations (see Figure 1), is favourable to fast ion intercalation/removal of the Li ions during the discharge/charge cycle of the battery. Although superstoichiometric Li_{1-x}Ni_{1+x}O₂, $0 < x \le 0.2$, was first synthesized in 1953, attempts to synthesize the stoichiometric LiNiO₂ (x = 0) compound have so far proved unsuccessful.

Figure 1. LiNiO₂ crystal structure, consisting of alternating layers of O (blue), Li (gold), and Ni (green) atoms.

NaNiO₂, although isoelectronic to LiNiO₂, undergoes a crystal structure distortion below T=480K, driven by the cooperative Jahn-Teller (JT) ordering in the Ni-ion layer. In LiNiO₂, such long-range orbital ordering is not occurring, despite the expected Jahn-Teller activity of the Ni ions in an octahedral environment. The difference in magnetic behaviour is equally puzzling, as the observed magnetic order in NaNiO₂, namely ferromagnetic intra-(Ni)layer coupling with antiferromagnetic inter-(Ni)layer modulation, does not seem to occur in LiNiO₂.

The local spin density (LSD) approximation to Density Functional Theory (DFT) is the standard way to perform *ab initio* electronic structure calculations. This approximation inherently assumes delocalised electrons, and the exchange and correlation effects are insufficient to describe narrow d-bands, since they do not account properly for the tendency of strongly correlated *d*-electrons to stay localized at each of their sites. This results in the LSD prediction of metallic behaviour in nickelates, although experiments find them to be insulating. To account for the strong correlations, most model and LSD+U calculations rely on the introduction of the Hubbard U parameter.

For the truly first-principles description of these materials, we have applied the self-interaction corrected (SIC) LSD method to study the ground state spin and valence configuration of nickelates [1]. The SIC-LSD method is an ab initio approach that corrects for an unphysical self-interaction contained in the LSD total energy functional. Whilst for extended band states the resultant error in the LSD energy is generally insignificant, it may be considerable for atomic-like localized states. This correction removes the spurious self-interaction of each occupied orbital. Both localized and delocalised states are expanded in the same basis functions, and are thus treated on an equal footing. Different localized/delocalised configurations are realized by assuming different numbers of localized d-states on the Ni atoms. Since the different localization scenarios constitute distinct local minima of the same energy functional, their total energies may be compared, and the global energy minimum then defines the ground state total energy and valence of the Ni ion. This latter is defined as the integer number of electrons available for band formation, namely N_{val} = Z - N_{core} - N_{SIC}, where Z is the atomic number, N_{core} is the number of core (and semi-core) electrons, and NSIC is the number of localized, i.e. self-interaction corrected, states.

The calculated total energies for LiNiO₂, with Ni ion assuming a number of different valence configurations, Ni²⁺ through Ni⁶⁺, as well as the completely delocalised LSD configuration, are presented in Figure 2. For each Ni valence a number of possible choices of the orbital symmetries are shown, with the corresponding lowest energy scenario indicated by a red line. The global energy minimum is obtained for the insulating state and Ni³⁺, with seven localized *d*-electrons in the $t_{2\sigma}^{6}e_{\sigma}^{1}$ low spin configuration (green line), in agreement with experiment. Thus SIC-LSD finds the low spin ground state for LiNiO₂, based on the firstprinciples total energy considerations, whilst in the model and LDA+U calculations the low spin state is assumed. For NaNiO₂ our calculations deliver the same low spin insulating ground state.

|4

Despite the same low spin ground states, there exist major differences in the electronic properties between LiNiO2 and NaNiO2. Chemical disorder, inherent to super stoichiometric LiNiO₂, but absent in NaNiO₂, has been indicated to contribute to these differences. A possibility has been raised that the complex electronic properties of LiNiO₂ are related to the presence of Ni anti-sites in the Li layers. These we have been able to study straightforwardly using the SIC-LSD method. Replacing one in four Li atoms by Ni in the original LiNiO₂ crystal corresponds to substitution of a monovalent Li by a trivalent Ni, which results in additional eg bands that are partially filled. The gain in hybridization energy associated with this d-band formation is however not large enough to compete with the possible gain in self-interaction energy which results from localizing the corresponding *d*-states. The resulting ground state valence configuration is $(L_{13}^{1+} Ni^{2+})_{Li-layer}(N_{33}^{3+} Ni^{2+})_{Ni-layer}O_2$ and the compound remains insulating even when off-stoichiometry.

Figure 2. Total energy as a function of Ni d-configuration in LiNiO₂. A number of scenarios are shown for each given configuration, $d_{xy}^{\dagger}d_{yz}^{\dagger}d_{xz}^{\dagger}d_{zz^2-r}^{\dagger}d_{x^2-y^2}^{\dagger}$ $d_{xy}d_{yz}d_{xz}d_{yz^2-r}^{\dagger}d_{x^2-y^2}$, where a localized spin up/spin down state is indicated by the \uparrow/\downarrow symbol, and a delocalised state is indicated by the \diamond symbol. The red line marks the lowest energy scenario for each valence. The green line denotes the global energy minimum.

The above interesting behaviour of misplaced Ni-atoms notwithstanding, the question still remains as to the extent that the off-stoichiometry induced valence change can explain the observed differences in electronic and magnetic properties between LiNiO_2 and NaNiO₂. This issue is not of purely academic interest, given that the degree of deviation from stoichiometry strongly influences electrochemical performance of the corresponding batteries. Unlike Ni³⁺ ions, divalent Ni²⁺ ions are not JT-active. Consequently, a random distribution of these ions in the NiO₂ layer could possibly affect the orbital ordering. However, this issue is clouded by the fact that introduction of Ni ions into the Li layer gives rise to additional magnetic interactions between neighbouring Ni layers which then need to be taken into account on top of the preexisting intralayer magnetic interactions. For example, it has been suggested that the substitution of some of the Na atoms by Ni atoms in NaNiO₂ leads to frustration of the anti-ferromagnetic stacking order and the destruction of long range magnetic order.

Given this complexity, it would clearly be advantageous to separate the effect of valence change on the orbital ordering and on the magnetic interactions. Specifically, we have proposed to start from stoichiometric NaNiO₂, which is both magnetically and orbitally ordered, and substitute Mg or Ca ions for some of the Na ions. By replacing monovalent Na with divalent Mg/Ca, we expect no additional interlayer magnetic interactions to occur, but the extra valence electron will, dependent on the relative size of the corresponding self-interaction and hybridization energies, either localize on a Ni-site or occupy an e_{σ} conduction band level. Our calculations have shown that the corresponding self-interaction is indeed the most important energy scale, and that the substitution of Mg atoms in the Na layer, leads to d^8 localization on an equivalent number of Ni-ions, which results in a mixture of JT-active Ni³⁺ and JT-inactive Ni²⁺ ions in the NiO₂ layers. Clearly, investigating experimentally how this would affect the local and medium range orbital ordering is then of considerable interest.

A study of $Na_{1-x}Li_xNiO_2$, where Li is gradually replaced by Na, has shown that both orbital-ordering and antiferromagnetism are conserved for $0 \le x \le 0.2$, and that the collective |T| transition is only suppressed for x > x0.2. This result is in line with our picture, as the substitution of Na by isoelectronic Li does not result in a change in valency of the Ni ions in the NiO-layer. If the substitution of Mg or Ca for Na brought about the suppression of long range orbital ordering for comparatively smaller concentrations, this would be a strong indication that it is indeed the random distribution of Ni²⁺ ions in the NiO₂ layer that is responsible, rather than the introduction of magnetic Ni atoms in the Li layers. By extrapolation, the absence of Jahn-Teller transition in the off-stoichiometric LiNiO₂ could then be partially traced to the very same origin.

References:

[1] L. Petit, G. M. Stocks, T. Egami, Z. Szotek, W. M. Temmerman, Phys. Rev. Lett. 97, 146 405 (2006).

L. Petit, G. M. Stocks, T. Egami, Oak Ridge National Laboratory; Z. Szotek, W. M. Temmerman, CSED, Daresbury Laboratory.

CCP4 6.0: State-of-the-art Software for the Experimental Determination of Protein Structures

Proteins are large biopolymers consisting of a linear chain of amino acids. Genome projects have led to the amino acid sequences of hundreds of thousands of proteins being deduced. However, while the pimary sequence can give an idea about the function of a novel protein through comparison with other proteins of known function, it gives little direct information since there is as yet no reliable way of predicting the three dimensional atomic structure which determines the protein's chemistry.

The experimental determination of protein structures is therefore of great importance for understanding the role(s) of a protein, and the leading technique for doing so is protein crystallography (PX). The SRS at CCLRC Daresbury Laboratory has had several beamlines dedicated to PX, and the importance of this method has been recognised by the inclusion of three PX beamlines out of the seven Phase I beamlines at the new Diamond synchrotron.

A protein crystallography experiment involves collecting a large number of diffraction images from one or more crystals of the protein of interest, each of which may record hundreds of Bragg peaks. To go from here to the desired atomic model remains a challenging problem for several reasons. The most fundamental is the lack of phase information, which is required to generate the corresponding electron density of the protein. Often additional experiments are performed to determine a small substucture of "heavy" atoms, from which initial phase estimates can be obtained. In some cases, the known structure of a related protein can be used to give initial phase estimates. Only rarely can the "direct methods" of small molecule crystallography be used. For all methods of structure solution, progress is

Figure 1. The annual CCP4 Study Weekend regularly attracts over 400 participants to hear about the latest developments in protein crystallography from a strong list of international speakers.

hampered by the weak signal-to-noise ratio of protein diffraction. In addition, the sheer size of proteins introduces a significant level of complexity into all operations.

The Collaborative Computational Project No. 4 (CCP4) was set up in 1979 to coordinate the efforts of UK protein crystallographers who were writing software to help with the structure determination steps. Since then, the project has grown considerably, and now is the leading provider of PX software world-wide. The project is based at CCLRC Daresbury Laboratory, but has strong collaborations with a number of university groups in the UK and abroad. As well as developing and distributing software, CCP4 has a vigorous educational programme, the highlight of which is the annual Study Weekend (Figure 1).

Figure 2. The CCP4 suite includes around 200 programs covering all aspects of structure solution by protein crystallography. The programs are typically accessed via a graphical user interface, shown here, which controls project organisation, job submission and display of results.

Figure 3. Solution of the phase problem leads to an electron density map of the crystal which needs to be interpreted in terms of an atomic model. The graphics program Coot provides a range of tools for building and correcting the model. Equally importantly, it aids validation of the model against the diffraction data and against known chemistry.

CCP4 makes software available through regular releases of the CCP4 suite. Recently, we have released CCP4 6.0 which includes several new programs, and represents a major advance for the suite. The release comprises around 200 applications, based around common library code, and accessed individually from the command line or via a graphical user interface (Figure 2). The release also contains extensive documentation, tutorials, and files of useful reference data.

A common theme of protein crystallography is the correct handling of the errors which arise from the weakness of the data and the complexity of the atomic models. In recent years, PX software has increasingly been based on maximum likelihood techniques, which use Bayesian statistics to incorporate error models in a correct manner. CCP4 6.0 includes three new programs which make use of such techniques. BP3 produces initial estimates of phases for experimental phasing techniques such as "MAD". A correct statistical approach is required when combining diffraction data collected at different wavelengths. PIRATE is a program for improving initial phases, obtained for example from programs such as BP3. It does this by matching the electron density statistics to distributions obtained from a reference structure of known quality. Finally, PHASER applies maximum likelihood techniques to the Molecular Replacement method, which uses a known related protein structure for initial phase estimates.

CCP4 6.0 also sees the introduction of molecular graphics into the suite, via Coot and CCP4mg. Coot (Figure 3) is focused on the problem of building an atomic model into the electron density generated by initial phase estimates. Since the protein model may contain thousands of atoms, this is a non-trivial step and historically was one of the most time-consuming steps. Modern automated model-building algorithms can help here, but are hampered by low resolution data. Coot provides a variety of tools for manual model building, as well as tools for checking the validity of the model against various standard measures.

Because of the complexity of proteins, there are a variety of ways of visualising structures depending on the purpose. Frequently, a cartoon of the overall fold which omits details of individual atoms is sufficient to get a feeling for a protein, and how it relates to other proteins. Some parts of the protein, for instance the active site of an enzyme, may require a more detailed representation. The CCP4 program CCP4mg has been developed to aid the visualisation and structure analysis of proteins (Figure 4). It can also be used to produce high quality figures and movies.

M. Winn, CSED, Daresbury Laboratory

Two on-going themes in CCP4 are useability and automation, both of which follow on from the increasing maturity of the PX field. The new generation of structural biologists view PX as one technique in their toolkit, and software must be easy to apply by scientists who are not expert crystallographers, and may not be strong in computing. The need to automate structure solution software is being driven by faster and better data collection. With data collection times per crystal dropping to a couple of hours or less on modern beamlines, it is desirable to attempt the structure solution in situ, and with minimal user intervention, in order to see whether further data should be collected. CCP4 is currently working on a number of automation schemes to achieve these goals, and one of these, "CRANK", is distributed in CCP4 6.0. In the future, automation schemes will increasingly make use of compute clusters to implement parameter searching algorithms. Of course, not all structure solutions are so straightforward, especially as structural biologists target larger protein assemblies, and CCP4 will continue to support the core programs for manual structure solution.

Figure 4. CCP4mg is a molecular graphics package focused on structure analysis and high quality presentation graphics and movies. Shown here is the surface of dihydrofolate reductase coloured by the estimated electrostatic potential, together with a space-filling representation of the bound ligand methotrexate.

References:

[1] 'An overview of the CCP4 project in Protein Crystallography: an example of a collaborative project' - M. D. Winn, J. Synchrotron Rad. 10, 23-25 (2003).

[2] Proceedings of the CCP4 Study Weekend are published each year in a special edition of Acta Crystallographica section D.

[3] www.ccp4.ac.uk

Extending the Range of Parallel Quantum Chemical Methods

Parallel computers are now well established as the most cost-effective way to provide CPU cycles for computeintensive disciplines such as computational quantum chemistry. In this article, we describe how the Computational Chemistry Group are extending the scope of parallel algorithms from the traditional domain of iterative Self-Consistent Field methods (such as Hartree-Fock and Density Functional Theory (DFT)) to methods designed to treat the excited state.

DFT is arguably the most accurate effective 1-electron model available at a similar cost to Hartree-Fock theory, and to capitalize on the potential of this approach we have been developing a DFT module within GAMESS-UK since 1994. This has led to a code that supports energies, gradients and Hessians as well as to a library of functionals in source code that are available for external users[1]. DFT is essentially a ground state theory but a treatment of excited state systems may be derived by considering how the ground state wave function responds to a time-dependent external electric field. The resulting formulation, Time-Dependent DFT (TD-DFT) essentially amounts to solving a pseudo eigenvalue equation. This can be accomplished effectively by an iterative diagonalization, avoiding the storage of large matrices at any time during the calculation, and distributing those objects that do need to be held (such as trial vectors) using distributed data programming toolkits such as the Global Arrays and DDI [2]. Hence the memory needs scale favourably with respect to system size, and the implicit flexibility in the distribution of work makes them very suitable to parallelization.

Calculating the excitation energies in itself opens the door to a whole field of chemistry but for some areas such as photo-chemistry it is crucial to be able optimise the geometries of molecules in their excited state, (for example to compute Stoke's shifts) and ultimately to perform molecular dynamics of excited state species. To do this efficiently the analytic forces are essential. A recent programme of work, in collaboration with the NWChem group, has led to the implementation of a new, massively parallel code for TD-DFT gradients.

As a first application of this functionality we are considering the study of the effect of defects in ionic materials on their spectroscopy. This is of particular interest for the catalytic processes these materials are used for. For these processes there is spectroscopic data relating to straightforward vertical excitations as well as pump-probe experiments; however, the theoretical analysis is wanting. In Figure 1 we show a cluster model for a defect (Catlow and Sokol) which we are using to test the new functionality.

Figure 1. Cluster model used for TD-DFT studies of a Li defect in Zn0. The Li atom is shown in green, 0 in red and Zn in yellow.

While DFT provides a cost effective solution for many problems, there are limits to its applicability, especially for systems which are poorly described by a single reference configuration (either in the ground or excited states).

To address such systems a highly scalable multiconfigurational self-consistent field (MCSCF) code [3] has been developed within the GAMESS package to exploit the capability of massively parallel resources, such as HPCx. The scalable MCSCF code is based on distributed data methods for the generation of transformed integrals combined with a similar direct diagonalisation to that used for the TD-DFT code.

As an example we consider the prediction of the redox potentials of metalloproteins. Many proteins share common metal cofactors and their redox potentials are thought to arise primarily from changes in the oxidation state of these metal cofactors at the active sites. For instance, mutations of the amino acid sequence lead to subtle distortions in the ligand field surrounding the metal cofactor, changing the difference in energy between their oxidation states, allowing Nature to tune the redox potential for different purposes. While the organic ligands surrounding the metal cofactor may be

Figure 2. The enzyme Ferredoxin showing the Fe2S2 metal cofactor in a ball and stick representation.

described using single-reference methods, the interaction between metal atoms of the cofactor is often highly multiconfigurational. An example of this class of system is the ferredoxin enzyme containing a 2Fe-2S metal core (Figure 2).

Figure 3 shows the scaling on HPCx of a typical MCSCF calculation on the Ferredoxin active site and the plot (Figure 4) of a strongly-occupied natural orbital from the MCSCF calculation is indicative of Fe-Fe bonding through the metal d-orbitals.

The incorporation of advanced configuration interaction (CI) techniques, such as multiple active spaces (ORMAS[4]) and a distributed data sparse CI solver (in development), enables the study of systems that would otherwise be prohibitively large.

The long-term goal of this work is the study of dynamical processes and reaction rates for complex biological and technological systems. This will require high accuracy QM methods combined with detailed environmental models, such as the embedded clusters mentioned above and biochemically oriented QM/MM models (see the article by Kästner and co-workers in this issue). The work illustrates the importance of close collaboration between software development groups (GAMESS-UK, NWChem and GAMESS) to tackle the challenges presented by advanced algorithms on massively parallel computers.

H. J. J. van Dam, G. D. Fletcher, J. M. Rintelman, M. F. Guest, and P. Sherwood CSED, Daresbury Laboratory.

Figure 3. The parallel performance of the parallel MC-SCF algorithm on the HPCx system.

Figure 4. An Fe-Fe bonding contribution to the MCSCF wavefunction for the Fe2S2 metal cofactor of Ferredoxin.

References:

For more information on the software mentioned in this article, visit the web pages www.cfs.dl.ac.uk (GAMESS-UK), http://www.msg.ameslab.gov/GAMESS/ (GAMESS(US)) and http://www.emsl.pnl.gov/docs/nwchem/ nwchem.html (NWChem).

[1] http://www.cse.clrc.ac.uk/qcg/dft/

[2] J. Nieplocha, R. J. Harrison and R. J. Littlefield, "Supercomputing", (1994) 340-349;

G. D. Fletcher , M.W. Schmidt, B. M. Bode et al, Comp. Phys. Comm. 128 (2000) 190-200.

[3] G. D. Fletcher, Mol. Phys., (in press).

[4] J. Ivanic, J. Chem. Phys., 119 (2003) 9364.

Nano-objects in Supercritical Media: Solvation Behaviour of Metal Nanoparticles in Supercritical Fluids

The phenomenal surge of interest in nano-objects (metal-atom clusters, fullerenes, carbon nanotubes, nanowires, nanotapes, nanocomposites, mesoporous nanofilms, protein-nanoparticle conjugates etc) is due to their extraordinary optical, electronic, magnetic and conducting properties that differ significantly, and often qualitatively, from those of the corresponding bulk materials. These properties hold the promise of a technological revolution in areas such as optoelectronics, magnetic storage, sensing technologies, catalysis and fuel storage. In addition to structural and chemical characteristics, factors that feature crucially in the determination of the various properties of nanoobjects are their size and shape. Therefore, a key requirement for pursuing systematic and accurate studies of these objects, with both scientific and technological objectives in mind, is the availability of samples of narrow size distribution. Recently, 'supercritical fluids' have been investigated for their efficacy as fractionating media for nanoparticles and other nano-objects: the tunability of solvent characteristics of supercritical fluids allows the development of promising approaches for the production of monodisperse objects.

Above a certain temperature known as the 'critical temperature', a substance exists in one phase only no matter how great the pressure applied. Supercritical fluids were found to possess extraordinary solvent properties well over a century ago, but their potential as effective media in processing and extraction technologies was first realized in the late nineteenseventies. Today, supercritical fluids feature prominently in this role in technologies as diverse as the manufacture of decaffeinated coffee and nicotine-free tobacco, catalysis, synthesis and purification of polymers, extraction of perfumes and flavours, recovery of low-boiling-point chemicals from crude oil residue, reaction engineering, biotechnology and cleaning of electronic components.

The superiority of supercritical fluids over liquids as solvents and suspending media in the above and other applications is attributed to their high compressibility, low surface tension, low viscosity and, perhaps most significantly, to their special mode of solvating the particles and solute molecules immersed in them. The focus of our studies in the field resides in this last aspect of the behaviour of supercritical fluids. The supercritical regime is characterised by appreciable inhomogeneity and fluctuations in density, particularly close to the critical point, which for simple systems have been shown theoretically to lead to the *clustering* or *depletion* of the solvent around the solute molecule, depending on the strength of the interaction between a pair of the solvent and the solute molecule relative to that between a pair of the solvent molecules. Enrichment or depletion of the solvent in the immediate surroundings of the solute particles would significantly affect the solvent-mediated interactions between them and hence their propensity to aggregate or remain dispersed in the solvent. Thus solvation has an important part to play in controlling the state of material dissolved/dispersed in supercritical media. This makes the development of molecular-based understanding of particle solvation a highly desirable objective.

Figure 1. (a) 8-gold-atom passivated nanoparticle, (b) 38-gold-atom passivated nanoparticle, (c) the 38-atom passivated nanoparticle in ethane. The ligands are alkane-thiolates ($C_{12}H_{25}S$).

Chemical methods developed in recent years yield 'passivated' nanoparticles covered with ligands, in particular alkyl thiolates terminally attached to the particle surface through Au-S bonding. We have been engaged on the study of bare and passivated gold nanoparticles in ethane aimed at the elucidation of the mode and the extent of solvation of the particles in the supercritical regime of the solvent. Our published simulations using the DL_POLY_2 code have elucidated detailed solvation behaviour as a function of (sub- to supercritical) temperature at the critical density and as a function of density over a range of supercritical isotherms. The modes of solvation for bare and passivated particles, deduced from the radial distribution of the solvent about the metal-core centre of mass, are found to be qualitatively different from each other: while the molecules solvating the bare particle form a well-defined, two-region layer around it, those solvating the passivated particle are loosely dispersed in the passivating layer.

We have recently embarked on the computation of partial molar volumes and potentials of the mean force (pmf) for the nanoparticles in the solvent. The aim is to provide a tangible link between our simulations and experimental thermodynamics of nanoparticle solutions

Figure 2. Partial molar volume calculations: T = 310K, p = 0.054 kbar. The partial molar volume $(\partial V/\partial n) = -22 \pm 3 \text{ nm}^3 \text{ per particle.}$

in supercritical solvents. The pmf may be regarded as the effective potential between nanoparticles in the solvent controlling their propensity to disperse or aggregate. High compressibility of supercritical fluids would provide a convenient means to change the pmf by systematically changing the density of the system through application of moderate pressures at constant temperature. Our objective is to explore how the pmf would depend on the solvent density above the critical temperature and in this way to identify those regions of the supercritical regime in which the solvent would serve as a dispersing medium and those in which it would serve as an aggregating medium.

Some preliminary results on the partial molar volume and the relative mean force obtained for 38-atom bare gold particles in supercritical ethane are presented below. The mean force calculation shows a clear distinction between the liquid and supercritical regimes. The computed values of the change in the volume as a function of the number of particles at constant temperature and pressure enable the determination of the partial molar volume with accuracy of \pm 3 nm³ per nanoparticle.

Figure 3. Mean force calculation for two 38-atom nanoparticles and solvent density \sim 8 molecules per nm³ at T = 290K (liquid) and T = 380K (supercritical).

M. Plummer and W. Smith, CSED, Daresbury Laboratory; M. Lal, University of Liverpool, UK.

References:

Full quantitative details of our solvation simulations and general references for nanoparticles and supercritical fluids are given in:'Solvation of metal nanoparticles in a subcritical–supercritical fluid: a computer simulation study', M. Lal, M. Plummer, N. J. Richmond and W. Smith, J. Phys. Chem. B 2004, 108, 6052-6061; 108, 10196. 'Solvent density effects on the solvation behaviour and configurational structure of bare and passivated 38-atom gold nanoparticle in supercritical ethane', M. Lal, M. Plummer and W. Smith, J. Phys. Chem. B 2006, in press. Details of the DL_POLY codes may be found at http://www.ccp5.ac.uk/DL_POLY/.

Molecular Dynamics Simulation Studies of Molecular Complexation Behaviour of Cyclodextrins

Cyclodextrins (CDs) are a group of oligosaccharides which consist of glucose units connected together to form a ring structure. They are readily available from enzymatic hydrolysis of starch. The naturally occurring CDs usually consist of six member (α), seven member (β) and eight member (γ) glucose units and have nearperfect circular shape, adopting approximate truncated cylindrical cone (Figure 1). The arrangements of these units are such that the cavities of the CD molecules are highly hydrophobic whereas the rims are highly hydrophilic. This unique structure allows the CD molecules to dissolve in water while retaining a capability of encapsulating hydrophobic molecules. For this reason, CD-based technologies have been widely exploited in the pharmaceutical, food science and cosmetic industries. In the pharmaceutical application, the encapsulation capabilities of CDs are used to increase the solubility, stability and bioavailability of drugs and to prevent undesirable drug-drug interactions.

Figure 1. Side and top views of the atomic configuration of b-CD, a 7-membered glucose units connected to one another via 1,4 glycosidic linkage to form a ring molecule.

Although a great deal of work has been done to elucidate the physical and chemical characteristics of CD molecules, little is known regarding the underlying atomistic interactions and the detailed complexation mechanisms. Such knowledge is essential to provide a reliable prediction of how a new drug molecule could form a complex with the cyclodextrin molecule. This could lead to direct cost reduction in the oftenexpensive processes of drug production and formulation. We at the Daresbury Loaboratory (DL), in collaboration with the AstraZeneca (AZ) Product Development Team at Macclesfield, have carried out systematic investigations of these complex molecules to gain understanding of the atomistic processes of the molecular encapsulation behaviour. The project is a synergistic combination of computational and experimental approaches which ensures that the challenging research work is carried out efficiently. This combines the molecular simulation expertise at DL, which makes use of the state-of-the-art HPCx supercomputer, and the experimental knowledge of the AZ group.

In the project, we consider the CD derivative 2-hydroxypropyl- β -CD (2-HPBCD). This modified CD is considered to be the most desirable due to its high solubility in water and low toxicity in drug formulations. Unfortunately, as in the case of other chemicallymodified CDs, the precise quantification of 2-HPBCD is very difficult to determine experimentally, if not impossible. This is due to an inherently wide ranging degree of random substitution during the reaction and the practical difficulty of separating individual derivatives from one another.

It is therefore our aim to use molecular dynamics techniques to investigate systematically a selected range of 2-HPBCD derivatives in order to identify the important characteristic features of molecular encapsulation for selected drug molecules and determine the corresponding solubility. We intend that these studies should form the basis for making predictions of complexation and solubilisation of new drug molecules on other forms of 2-hydroxypropyl β -CD substitutions.

The molecular dynamics simulations were performed with the DL_POLY program package developed at Daresbury which exploits the parallel capabilities of the HPCx supercomputer. The example CD derivatives (host molecules) shown in this article are 2HPBCD-DS(O2) and 2HPBCD-DS(O2,O6), of the which the 'DS' notations indicate the presence of 2hydroxypropyl derivatives at only the glucose O2 positions and at both O2 and O6 positions. The example guest molecule (to be complexed with the host molecules) we used is 9-fluorenone, a polycyclic aromatic hydrocarbon that is used in manufacturing antimalaria drugs and other pharmaceuticals. Figure. 2 and Figure. 3 illustrate the complexation of 9-fluorenone within the hydrophobic cavities of the CD derivatives. It can be seen that the guest molecule prefers to locate at the lower parts (with respect to the z-axis as shown in Figure. 2) of the β -CD when only O2 positions are derivatised. However, when the O6 positions are also substituted, the guest molecule moved towards the top part of the β -CD molecule. This may be due to the flexibility of the O6-substituted groups and thus can easily pack and interact with the guest molecule.

Figure 2. Side and top views of the complexation of 9-fluorenone with HPBCD-DS(O2). The blue spheres represent the glycosidic oxygen linkage of the CD molecule and the thick stick models represent the 2-hydroxypropyl group at the glucose O2 positions. The 9-fluorenone molecule is highlighted as the green ball-and-stick model.

The guest molecular orientations are also found to be dependent upon the nature of CD substitutions. For instance, the plane of the 9-fluorenone molecule in the cavity of the (O2,O6)-substituted CD is nearly parallel to the z-axis, whereas the guest molecule has larger angles and over a wider range of value in the cavity of the O2-substituted CD.

These results show that the interactions of the guest molecule within the hosts' cavities are strongly influenced by the locations of the 2-hydroxypropyl derivatives. Such information is valuable in predicting the mechanisms of molecular encapsulation of new guest molecules and their corresponding solubility and complexation behaviour in water.

Figure 3. Side and top views of the complexation of 9-fluorenone with HPBCD-DS(O2,O6). The blue spheres represent the glycosidic oxygen linkage of the CD molecule and the thick stick models represent the 2-hydroxypropyl group at the glucose O2 positions. The 9-fluorenone molecule is highlighted as the green ball-and-stick model.

C. W. Yong, W. Smith, CSED, Daresbury Laboratory. C. Washington, Astra Zeneca.

Molecular Dynamics Study of Radiation Hard Materials: Radiation Damage in Pyrochlores

The urgent need for a new generation of radiation hard materials (RHM) has prompted renewed effort to quantify, understand, and predict the properties of such materials and link these to the underlying solid state chemistry and physics. Applications range from the scientifically dazzling components in fusion reactors to waste-forms for encapsulation of nuclear industry derivatives. The latter remains one of the daunting environmental challenges of this century as there are substantial amounts of plutonium (half-life ²³⁹Pu 24,100 years) and other long-lived contaminants (half-life ²³⁵U 7x10⁸ years, Np, Am, ...) that require immobilisation in waste-forms suitable for geological disposal. The radionuclides need to be bound in an inert matrix capable of retaining its structural integrity during prolonged heavy particle bombardment at temperatures of \approx 300-600 K.

Two types of particle irradiation are responsible for the α -decay damage. The first of these is the α -particle or helium ion itself, with energies of 4.5 to 5.8 MeV and a range of 10-30 µm. The second is the recoil of the heavy actinide nucleus, which has an energy of 70-100 keV and a range of 10-20 nm. An α -particle deposits its energy predominantly by ionisation over its path with a limited number of atomic collisions (\approx 100) during its trajectory; in contrast the α -recoil ion will lose most of its energy by collisions producing up to several thousand atomic displacements. The latter dominates the structural damage, with the initial cascade taking place on a timescale of femtoseconds (10⁻¹⁵ s) followed by a relaxation of the structure over picoseconds (10⁻¹² s). Molecular dynamics (MD) simulations are in principle well suited to studying such atomic processes which take place on the picosecond timescale. Simulations of α -recoil ions cascades at realistic energies, ≈10-50 keV, however, require simulation of the order of millions of particles. This is a far from trivial task as they demand a considerable amount of dedicated CPU resources and also hard-disk storage. Furthermore, the simulation software needs to have powerful tools to quantify the damage in numerical terms as the calculation proceeds, since graphical visualisation of the details of the damage in such large systems is not straightforward and it is easy to be misled especially when similar compounds are compared. Processes such as thermal annealing which take place over longer timescales are currently beyond

the scope of the technique though powerful new methods for this purpose such as hyperdynamics are under active development.

Figure 1. The pyrochlore structure: showing rings of Zr(Ti)O₆ octahedra in green, Gd cations in grey and non-polyhedral oxygens in cyan.

The identification of a prospective waste-form material is contentious, with pyrochlores, Figure 1, such as Gd₂Ti₂O₇ as primary candidates. Nevertheless, studies employing heavy ion bombardment to simulate the effects of α -decay have shown that Gd₂Ti₂O₇ undergoes amorphisation at a dose corresponding to that of α -decay damage, consequently resulting in a ten-fold increase in the leach rate of plutonium. In subsequent experimental studies Wang et al. [1], using heavy ion bombardment, revealed a striking and intriguing variation in radiation tolerance in the series $Gd_2(Zr_xTi_{1-x})_2O_7$ (X=0-1). Increasing radiation tolerance was found with increasing Zr content so that Gd₂Zr₂O₇ was predicted to withstand radiation damage for 30 million years compared with 800 years for Gd₂Ti₂O₇. This is very surprising in view of the chemical similarity of the Group 4 elements Zr and Ti, and the $Gd_2(Zr_xTi_{1x})_2O_7$ system together with other pyrochlores are now receiving considerable experimental attention. The Gd zirconate end member (X=1) is indeed largely insensitive to irradiation damage remaining highly crystalline to high doses even at very low temperatures undergoing a radiation-induced transition to an (oxygen-defective) fluorite structure. In this the Gd³⁺ and Zr⁴⁺ cations are disordered, additional oxygen sites are occupied, and it is itself highly radiation resistant.

Simulations of cascades of 10 keV uranium atom recoils in the pyrochlores Gd₂Ti₂O₇ and Gd₂Zr₂O₇ (193,000 particles) [2] have highlighted the importance of the different rôles played by the (8a) oxygen site in the two systems. The number of defects involving the cation sublattices is much smaller than those on the anion sublattice. When damage takes place in Gd₂Zr₂O₇ the (8a) site is readily occupied by oxygen, consistent with this material also being a good oxide fast-ion conductor. Partial occupation of this site appears to be needed to provide an effective mechanism for efficient impact energy absorption and consequent structure recovery to a crystalline state during and after the cascades. In contrast, the occupancy of the (8a) site is much less in damaged Gd₂Ti₂O₇ than in Gd₂Zr₂O₇ and the activation energy barrier to structure recovery greater. The amorphous state is more stable in the Ticompound. It is worth stressing that the total number of oxygen defects initially generated by the cascade is

actually larger in the Zr compound than in the titanate, consistent with the weaker Zr-O bond. Our simulations thus highlight the importance of the recovery and healing processes and oxygen mobility in these systems rather than the *initial* resistance to damage creation *per se*.

The calculations were performed on the HPCx [3] cluster using the DL_POLY_3 [4] molecular dynamics software.

References:

[1] S. X. Wang, B. D. Begg , L. M. Wang, R. C. Ewing, W. J. Weber and K. V. Govindan Kutty, 1999, J. Mater. Res. 14, 4470.

[2] I. T. Todorov, J. A. Purton, N. L. Allan and M. T. Dove, 2006, J. Phys.: Condens. Matter 18, 2217.

[3] http://hpcx.ac.uk/

[4] I. T. Todorov, W. Smith, K. Trachenko. and M. T. Dove, 2006, J. Mater. Chem. 16, 1911.

I. T. Todorov, W. Smith, J. A. Purton, CSED, Daresbury Laboratory; N. L. Allan, University of Bristol; M. T. Dove, University of Cambridge.

New Tools for the Accurate Study of Biomolecular Reactivity

The simulation of biochemical reactions has become a valuable counterpart to experimental investigations in rational drug design, catalysis, and the unravelling of natural processes. Such reactions pose the methodological challenge of a small reactive centre embedded in a large, inhomogeneous environment, e.g. proteins or DNA strands in aqueous solution.

Three levels of accuracy can be distinguished. (1) classical molecular mechanics force fields used in conjunction with molecular dynamics (MD) or Monte Carlo (MC) simulations. (2) Quantum mechanical methods to describe chemical reactions. (3) The QM/MM approach, which treats the reaction centre by quantum mechanics (QM) but incorporates the environment using classical molecular mechanics (MM). In this article we concentrate on some recent QM/MM enhancements to the programme package ChemShell which is developed by on-going collaboration between Computational Chemistry group at Daresbury, the group of Walter Thiel at MPI Mülheim and the group of Richard Catlow at the Royal Institution. The Mülheim group focuses on the biochemical tool development and it is this aspect of the collaboration, together with some recent applications work, that is summarised here.

Geometry optimisations provide minima and transition states on the potential energy surface. These are sufficient to describe the chemical reactivity of many small systems. The large, inhomogeneous, but comparably well-ordered environments of biochemical reactions, however, require the free energy, which includes the entropy, to be taken into account. We have developed a method to significantly improve the analysis of umbrella sampling simulations, biased MD simulations to calculate the free-energy change of chemical reactions.

Figure 1. In umbrella sampling the system evolves in a biased potential enabling efficient sampling of the transition state. Unbiasing using umbrella integration provides the free energy.

Umbrella Integration

Umbrella integration [1] is a method to analyse umbrella sampling simulations. It calculates the free energy by use of the mean force, which can directly be averaged over the steps, rather than weighting the probability density and finding the factors to combine the windows iteratively, as done in the standard weighted histogram analysis method (WHAM). Umbrella integration greatly reduces the noise inevitably present in MD simulations. The result is independent of other parameters of the analysis, such as the width of the bins for constructing the histograms in WHAM (see Figure 2).

Figure 2. Analysis of umbrella sampling simulations: umbrella integration (UI) converges, while no optimum number of bins can be given for the weighted histogram analysis method (WHAM).

Umbrella integration moreover uses the MD trajectory to provide an estimate of the statistical error in the resulting free-energy difference [2]. This provides welldefined rules for how to optimally choose parameters for the umbrella sampling simulations, such as the strength of the bias, and the number of steps.

The drawback of umbrella sampling simulations, whether analysed with umbrella integration or with WHAM, is the necessity for extensive sampling. While this can easily be done for force fields, the computational costs prohibit sampling for quantum chemical or QM/MM simulations.

QM/MM Free-Energy Perturbation

QM/MM free-energy perturbation allows the free energy change of the most important part of the system to be sampled at high levels of theory. The entropy and enthalpy changes of the QM/MM interactions and the MM part are sampled, while the entropy change of the OM part is neglected or estimated using the harmonic approximation. This allows the sampling to be confined to the MM part and classical MD to be used. The overall computational effort is dominated by geometry optimisations preceding the MD sampling. The QM/MM free-energy perturbation scheme has been implemented into ChemShell and demonstrated in a study of the catalytic bio-degeneration of aromatic compounds by p-hydroxybenzoate hydroxylase (PHBH). The QM/MM setup, containing about 22,700 MM atoms and 49 OM atoms, and the transition state, are depicted in Figure 3.

Figure 3. The protein PHBH was simulated solvated in an 11Å-shell of water. The insert shows the transition state: QM atoms as ball-and-stick model, MM atoms as sticks.

QM/MM Studies on Metallo-Proteins

The ChemShell software has been used in extensive QM/MM studies of cytochrome P450 and related metallo-enzymes. Cytochrome P450 enzymes catalyze a great variety of stereospecific and regioselective oxygen insertion processes, which are of vital importance in biosystems for detoxification and biosynthesis. Given the biological importance of these enzymes and the fact that they are capable of activating inert C-H bonds under ambient conditions, the underlying reaction mechanisms have been intensely studied. The recent QM/MM investigations with ChemShell have addressed all relevant intermediates of the catalytic cycle, in particular the reactive Compound I species [4-6], and most of the key transformations in cytochrome P450cam. The QM/MM study of the mechanism of C-H hydroxylation supports a two-state two-step rebound mechanism in the doublet and quartet state [7], catalyzed by a single water molecule that is hydrogenbonded to the oxo atom [8]. The QM/MM calculations on the formation of Compound I reveal a novel mechanism [9]. This work shows the great potential of QM/MM calculations for the atomistic modelling of enzymes in general, and for mechanistic studies of enzyme reactivity in particular. Such calculations provide unprecedented insight into the way enzymes work.

Outlook

A range of approaches is necessary to tackle biochemical processes at high accuracy with reasonable computational costs. An additional dimension is added by calculating the free energy rather than the potential energy, which is achieved by sampling the phase space. These methods pave the way for studies with high predictive power complementing experimental data. Not only will they lead to a better understanding of biological processes, but they may lend guidance to the development of fine chemicals, biomimetic catalysts, and novel drugs.

References: [1] I. Kästner. W

[1] J. Kästner, W. Thiel, J. Chem. Phys. 123, 144104 (2005).

[2] J. Kästner, W. Thiel, J. Chem. Phys. 124, 234106 (2006)

[3] J. Kästner, H. M. Senn, S. Thiel, N. Otte, W. Thiel, J. Chem. Theory Comput. 2, 452 (2006).

[4] J. C. Schöneboom, H. Lin, N. Reuter, W. Thiel, S. Cohen, F. Ogliaro and S. Shaik, J. Am. Chem. Soc. 124, 8142-8151 (2002).

[5] E. Derat, S. Cohen, S. Shaik, A. Altun and W. Thiel, J. Am. Chem. Soc. 127, 13611-13621 (2005). [6] J. C. Schöneboom, F. Neese and W. Thiel, J. Am. Chem. Soc. 127, 5840-5853 (2005).

[7] J. C. Schöneboom, S. Cohen, H. Lin, S. Shaik and W. Thiel, J. Am. Chem. Soc. 126, 4017-4034 (2004).

[8] A. Altun, V. Guallar, R. A. Friesner, S. Shaik, and W. Thiel, J. Am. Chem. Soc. 128, 3924-3925 (2006).

[9] J. Zheng, D. Wang, W. Thiel, and S. Shaik, J. Am. Chem. Soc., online.

J. Kästner, P. Sherwood, CSED, Daresbury Laboratory; H. M. Senn, S. Thiel, N. Otte, W. Thiel, MPI Mülheim, Germany.

Digital Microfluidics: Droplets on Demand

The ability to handle minute quantities of liquid accurately and reliably, at the nanolitre or picolitre scale, offers many practical advantages. In medical applications, it offers the potential to dispense precisely metered doses while in biological applications, it opens up the possibility of storing and transporting individual biological organisms.

The manipulation of liquids in the form of discrete droplets has become known as *digital microfluidics*. Typically, aqueous droplets are held within a nonaqueous carrier fluid such as silicone oil. The potential advantage of this approach is that each droplet represents a transportable individual reaction volume that does not exchange material with its surroundings. The droplet can be stored, mixed and manipulated in many different ways making it an ideal approach to the design of novel lab-on-a-chip systems.

Figure 1. Digital microfluidics in action. The image shows individual aqueous samples being transported along a microfluidic channel. Each droplet has an approximate volume of 5 nl.

As part of a research collaboration with the Miniaturisation and Nanotechnology Group at the University of Manchester, the factors affecting droplet generation in multiphase microfluidic systems are being investigated. The study is supported through an EPSRC grant (GR/S82978/01) and is exploring the use of both pressure-driven and electrostatic "droplet-on-demand" formation techniques. In addition, the study will address the important factors affecting the transport, merging and splitting of microdroplets. Understanding time-dependent flow of two immiscible phases in microchannels presents a new and significant modelling challenge as there is a strong coupling between the fluid flow and the transported phase. Moreover, this coupling adds significant complexity because it is necessary to account for some or all of the following interfacial effects: surface tension; multi-axis rotation; deformable boundaries; and merging and splitting of the dispersed phase.

Figure 2. Six frames from a sequence illustrating pressure-driven droplet formation from a 100 μ m side channel. The sequence runs from top left to bottom right.

Current numerical methods for simulating two-phase flows with discrete interfaces can generally be classified as either interface-tracking (surface) methods or interface-capturing (volume) methods. In interfacetracking approaches, the free surface is treated as a sharp interface whose motion is followed using an appropriate algorithm. In the second approach, the indicator function is usually a scalar step function (known as a colour function) representing the volume fraction occupied by one of the fluids. In the current project, the latter approach has been adopted using the Volume-of-Fluid (VOF) method because it ensures that conservation can be enforced since a scalar transport equation is solved in an Eulerian manner. However, one of the main difficulties is advecting the step function without diffusing, dispersing, or wrinkling the interface.

As part of EPSRC's Collaborative Computational Project 12, our in-house code, THOR-VOF, has been developed and tested on a range of interfacial flow problems, as illustrated in Figures 3 and 4. The results have shown that THOR-VOF can accurately predict a range of complex interfacial phenomena in micro-scale systems. However, the test cases have also served to illustrate the demanding computational cost of droplet simulations.

A simulation of an air bubble rising in mineral oil is shown in Figure 4. To reduce the computational cost, it was assumed that the problem was axisymmetric. The computations were carried out using 32 processors on an Itanium2 SGI Altix system and required a total cpu time of 40×32 processor hours. Full 3D simulations will therefore be very demanding.

Modelling interfacial phenomena offers significant numerical and algorithmic challenges. However, the benefit of being able to model these flows, accurately and reliably, will enable *next-generation* devices to deliver highly-controlled quantities in a precise and efficient way.

Figure 3. Simulation of Rayleigh-Taylor instability. The sequence illustrates the evolution of the interface between two immiscible fluids of different densities.

D. R. Emerson, R. W. Barber, X. J. Gu, P. W. Hogg and Y. H. Zhang, CSED, Daresbury Laboratory;

P. R. Fielden, N. J. Goddard and S. Mohr, University of Manchester.

(a) $t = 0.5 \times 10^{-3} s$

(b) $t = 42.5 \times 10^{-3} s$

Figure 4. Capturing the interfacial dynamics and velocity field around a rising air bubble.

References:

[1] X. J. Gu, D. R. Emerson, R. W. Barber, and Y. H. Zhang, "Towards numerical modelling of surface tension of microdroplets", in Parallel Computational Fluid Dynamics: Theory and Applications, pp. 365-372, A. Deane et al. (Editors) Elsevier Science, 2006.

[2] P. W. Hogg, X. J. Gu, and D. R. Emerson, "An implicit algorithm for capturing sharp fluid interfaces in the volume of fluid advection method", Daresbury Laboratory Technical Report, DL-TR-2006-001.

Ab Initio Studies of Aluminium Fluoride Surfaces

Aluminium flouride's (AlF₃) have great potential for use as catalysts in many Cl/F exchange reactions. Experimentally, large differences in catalytic reactivity are observed between the different phases of AlF₃. A series of detailed first principles simulations by our group has started to unravel the complex structures of these surfaces which in turn provides clues to their different chemical properties. Little is known experimentally about the detailed atomic scale structure of these surfaces and consequently these calculations provide the only model that explains this catalytic activity.

Figure 1. The two different low energy surface terminations of β -AlF₃.

These results allow us to explain the different reactivity of aluminium fluoride phases in terms of the substantial differences in the accessibility of the reactive, co-ordinatively unsaturated aluminium sites at their respective surfaces. For crystalline α -AlF₃ we have shown that the Al atoms are effectively covered by fluorine atoms [1, 2] leading to a surface that displays no Lewis acidity or catalytic activity. Conversely, we have shown that β -AlF₃, which shows moderate Lewis acidity and catalytic activity, contains co-ordinatively unsaturated five fold aluminium ions at the surface [1, 3]. On the basis of these results we can postulate that the highly reactive, amorphous 'high-surface area' AlF₃ may display acidic sites similar to those found on the β -AlF₃ surface.

It is known that the surfaces of AlF_3 strongly hydrolyse water and consequently it is likely that the surfaces of real operating catalysts will contain significant quantities of hydroxyl groups. To quantify this, a series of calculations have been performed allowing us to predict the surface chemistry and structure of these materials as a function of environmental conditions – namely the partial pressure of water and hydrogen fluoride that the surface is exposed to. These predictions are currently awaiting experimental verification although they appear to explain all known experimental data.

Figure 2. Phase diagram for α -AlF₃(0001). The coloured regions represent the most stable surface terminations that will exist as a function of gaseous HF and H₂O partial pressures.

In order to quantify the strength of the Lewis acid sites present on the surface of the β -AlF₃, studies of the adsorption of NH₃ have been performed. These show the presence of four possible active sites on the surface of amorphous materials (two possible terminations of the crystal each showing two possible adsorption sites). The binding energy of NH₃ to these sites is similar, varying between 1.7 and 2.0eV depending on surface and coverage. These values are comparable to the binding energies at strong Lewis acid sites in zeolite catalysts which are typically around 1.7eV The binding energy data can be used to predict experimental temperature programmed desorption plots and compared to experimental results as shown in figure 3. It can be seen that there is good agreement between experiment and theory. Desorption that occurs at low temperatures experimentally may be due to desorption of NH₃ from the channels that are present in the bulk β -AlF₃.

This ground breaking series of calculations has provided a new, detailed insight into the possible reasons for the catalytic activity of AlF_3 surfaces – materials that could play an important future role in Lewis acid catalysed reactions.

Figure 3. Comparison of the temperature programmed desorption results obtained theoretically and experimentally. The lines marked A and B correspond to desorption from the surfaces in figure 1 labelled A and B. It is expected that both of these surfaces are present on real β -AlF₃

A. Wander, C. L. Bailey, B. G. Searle and N. M. Harrison, CSED, Daresbury Laboratory; N. M. Harrison and S.Mukhopadhyay, Department of Chemistry, Imperial College.

References:

[1] Ab initio studies of aluminium fluoride surfaces, A. Wander, C. L. Bailey, S. Mukhopadhyay, B. G. Searle, N. M. Harrison J. Mat. Chem. 16 (2006) 1906-1910.

[2] The Composition and Structure of the α-AlF₃(0001) Surface, A. Wander, B.G.Searle, C.L. Bailey and N. M. Harrison J. Phys. Chem. B, 109 (2005) 22935-22938.

[3] Identification of Possible Lewis Acid Sites on the β-AlF₃ Surface: An ab initio total energy study, A. Wander, C. L. Bailey, B. G. Searle, S. Mukhopadhyay and N. M. Harrison, Phys. Chem. Chem. Phys., 7 (2005) 3989-3993.

Acknowledgements

We would like to thank the EU for support of part of this work through the 6th Framework Programme (FUNFLUOS, Contract No. NMP3-CT-2004-5005575) and the EPSRC for provision of access to the HPCx facility under the Materials Chemistry Consortium Grant (GR/S13422/01).

Application Performance on Dualcore Processors from Intel and AMD

With dual-core processors dominating server solutions from all leading suppliers, we consider the current status of the leading solutions from AMD and Intel – the dual-core Opteron and Woodcrest processors. Our focus lies in the measured performance of a number of key applications in the areas of Computational Chemistry, Materials and Engineering determined on a variety of commodity-based dual-core clusters.

The performance evaluation work carried out by the Distributed Computing Group continues to assess current and emerging commodity systems in scientific and technical computing through a variety of synthetic and application-based floating point metrics. Numerous systems have been rigorously evaluated using important applications – recent examples include the Cray XD1 and Infinipath Clusters. The primary goals of these evaluations are to a) determine the most effective approaches for using each system, b) evaluate benchmark and application performance, both in absolute terms and in comparison with other systems, and c) predict scalability, both in terms of problem size and in number of processors.

The dominant trend in the semiconductor industry strategy during the past two years has been to increase processor throughput by providing more processing elements (PEs) or "cores", rather than by increasing the operating frequency. Intel, AMD, and others are migrating to multiple (2x, 4x, 8x ...) cores within a single "processor". The next few years of server design will involve more cores and higher integration, with dual-core (DC) server processors now the norm, and quad-cores just months away. We here overview application performance on current dual-core processors, presenting comparisons between two of the more recent offerings, the Opteron250/2.4 GHz from AMD and the Xeon 5160/3.0 GHz "Woodcrest" processor from Intel.

Evaluation Systems

Our on-going assessment of a variety of commoditybased systems (CS) has produced a wealth of data generated through access to some 30 systems since 2002 [1]. Nine such systems have been used in the present study (CS22-CS31, see Table 1), with a particular focus on the emerging dual-core systems from AMD and Intel. Just two of these systems feature single-core processors, CS22 with Xeon EM64T 3.2GHz dual processor nodes and both Myrinet (M2K) and Infiniband interconnects, and CS23, an Infinipathconnected Opteron852/2.6GHz cluster from Streamline. All the remaining systems are based on dual-core processors with a variety of high performance interconnects. Six feature Opteron processors (CS24-CS29) with clock speeds varying from 2.0 to 2.4 GHz, and two - CS30 and CS31- the recently released "Woodcrest" dual-core Xeon processor from Intel. In assessing these systems, we have carried out the same benchmarks on a number of proprietary high-end solutions from IBM, HP, SGI and Cray. Our focus here is on the HP Integrity Superdome, featuring 64 X 1.6 GHz Itanium2 processors with 9 MB L3 cache and a Crossbar memory interconnect.

Application Benchmarks

The present evaluation concentrates on four of the application codes – CPMD [2], PDNS3D [3], ANGUS [4] and DL_POLY [5] – that form part of the DLAB benchmark. For CPMD, the *ab initio* Car-Parrinello plane wave pseudopotential code, three test cases have been considered involving the C₁₂₀ and Si₅₁₂ species. In the former we consider both closed-shell, singlet (S) and open-shell, triplet (T) single point density optimisations using the BLYP functional. Singlet calculations on Si₅₁₂ used the LDA functional.

Commodity System	Compute Nodes / PEs	Number of PEs	Interconnect	Location
CS22	Xeon EM64T/3.2GHz	256	Infiniband, M2K	TACC
CS23	Opteron 852/2.6GHz	32	Infinipath	Streamline
CS24	Opteron 270/2.0GHz DC	256	M2K	Leeds
CS25	Opteron 270/2.0GHz DC	64	Gbit Ethernet	Loughborough
CS26	Opteron 875/2.2GHz DC	128	Infiniband, M2K	HP
CS28	Opteron 275/2.2GHz DC	512	Infinipath	AMD Dev. Centre
CS29	Opteron 280/2.4GHz DC	256	Infiniband	HP
CS30	Intel Xeon 5160/3.0GHz DC	256	Infiniband	Intel
CS31	Intel Xeon 5150/2.66GHz DC	5200	Infiniband	TACC

Table 1. Commodity-based Systems

32

Two engineering direct numerical simulation codes, ANGUS and PDNS3D, have been included given their known dependency on memory bandwidth, the Achilles Heel of multi-core processors. PDNS3D is a simple turbulent channel flow example using the shock / boundary-layer interaction approach; the test case (T1) has a cubic grid size of (120^3). Test cases for the ANGUS combustion code benchmark include two cubic grid sizes – T1 (144^3) and the larger T2 (288^3).

The final application, DL_POLY, is the parallel molecular dynamics simulation package. The three benchmark simulations for the replicated data version of the code, DL_POLY2, feature (i) 27,000 ions of NaCl, (ii) 8,640 ions of NaK disilicate glass with 3 body forces, and (iii) a 12,390 atom macromolecular system, including 4012 TIP3P water molecules solvating the gramicidin A protein. Two simulations have been used for DL_POLY3 These figures suggest that Intel's 3.0 GHz Woodcrest processor has a significant performance advantage over the Opteron280/2.4 GHz processor. On average, runs using 32-cores of the Woodcrest cluster exceed the corresponding performance shown by the HP Superdome (106%), while the 32-core AMD cluster achieves some 80% of the proprietary high-end system. Thus the CS30 Xeon cluster outperforms the CS29 Opteron cluster in ten out of the eleven tests – the only exception is the larger of the two ANGUS benchmarks. Considering each code, we find the largest performance differential is with DLPOLY - on average 158% and 132% for DL_POLY2 and DL_POLY3 respectively. The advantage is less marked in CPMD (122%) and PDNS3D (120%); only in the memory-bandwidth intensive ANGUS code (89%) does the CS29 Opteron system outperform the Xeon cluster.

	Data Set	T _{32-processor} HP Superdome / T _{32-core} CSx		
Application Code		CS30 Intel Xeon 5160/3.0GHz dual-core + IB Cluster (%)	CS29 HP Opteron/280 2.4 GHz dual-core + IB Cluster (%)	
CPMD	C120 (S)	61%	48%	
	C120 (T)	67%	54%	
	Si512	81%	73%	
DL_POLY 2	NaCl	144%	88%	
	NaK disilicate	179%	106%	
	Gramicidin	213%	149%	
DL_POLY 3	NaCl	160%	116%	
	Gramicidin	130%	102%	
ANGUS	T1 (144**3)	27%	26%	
	T2 (288 **3)	64%	84%	
PDNS3D	T1 (120**3)	42%	35%	
	Average Performance Ratio	106%	80%	

Table 2. Application Performance: Percentage of a 32-processor Partition of the HP Itanium2 Integrity Superdome Achieved by 32 cores of the CS30 Intel Xeon 5160/3.0GHz and CS29 HP Opteron280/2.4GHz Linux Clusters.

the distributed data version of the code – a NaCl simulation with 216,000 ions, where the Smooth Particle Mesh Ewald method is used to calculate the Coulombic interactions, and that including a number of Gramicidin A molecules in water, yielding a total 792,960 atoms.

In each case we have generated performance data on the commodity-based systems (CSx), and compared this with results from the HP Integrity Superdome. We limit the discussion here to a consideration of the percentage of a 32-processor partition of the Superdome delivered by the range of commodity based systems (i.e. $T_{32-processor}$ Integrity Superdome / $T_{32-core}$ CSx). Table 2 shows the % of a 32-processor partition of the Superdome delivered by 32 cores of both the Infiniband-connected CS30 Intel Xeon 5160/3.0 GHz and CS29 AMD Opteron 280/2.4 GHz dual core systems. The present results suggest that the balance in the battle for processor supremacy between AMD and Intel perhaps lies in favour of Intel at this point, although the release of Socket F by AMD and the arrival of quad-core solutions strongly suggests that this battle is far from over.

M. Guest, C. Kitchen, M. Deegan, I. Kozin and R. Wain, CSED, Daresbury Laboratory.

References:

[1] http://www.cse.clrc.ac.uk/disco/ Benchmarks/commodity.shtml

[2] http://www.cpmd.org

[3] http://www.cse.clrc.ac.uk/ceg/sbli.shtml

[4] http://www.dl.ac.uk/TCSC/Subjects/ Parallel_Applications/benchmarks/benchmarks/ node31.html

[5] http://www.cse.clrc.ac.uk/msi/software /DL_POLY

The Challenge of Coordinating a £38M Distributed HPC Procurement in the UK - SRIF3

The Distributed Computing Group's role as technical consultants to the UK's SRIF3 coordinated HPC procurement has highlighted a number of key issues which are frequently overlooked or undervalued in cluster procurements. This article considers some of the challenges facing a coordinated procurement involving a diverse user community with differing requirements and expectations, plus some of the lessons learnt along the way.

High-End and Mid-Range Computing in the UK.

The focus of much of the parallel code development and system management & administration effort within CSED has naturally focussed on HPCx, the UK's national supercomputing facility at Daresbury. However, we also have a growing involvement in the procurement and deployment of mid-range commodity-based systems through the activities of the Distributed Computing Group (DisCo).

Figure 1. The spectrum of computer resources.

The Branscomb pyramid in Figure 1 presents a schematic for HPC resources. Leadership computers are currently the preserve of initiatives in the USA and Japan. HPCx is representative of the 2nd tier – National Supercomputers – that enable high-end capability and capacity solutions. Mid-range commodity computing belongs in the 3rd tier of University departmental clusters. While it is tempting to assume that the procurement and support of mid-range clusters is a simpler, scaled down version of that at the high-end, a consideration of current procurement mechanisms reveals important differences.

High-end systems such as HPCx are procured against clearly defined success factors accompanied by well developed scientific and business cases. They are typically funded over a 6-year period through rigorous contractual arrangements with the service provider, have quantified technology refresh points and a consortia-based utilisation model.

Whilst the overall expenditure on mid-range systems is currently compatible with that at the High-end through initiatives such as the Science Research Investment Fund (SRIF), procuring such systems poses a technically challenging problem given the number of sites involved and the variety of commodity-based systems procured (from 128 to 3,000+ "cores"). The modality of use at the high-end and mid-range is also different.

SRIF3: 2006-08

£903 million has been allocated for research capital to English HEIs, of which some £38M has been assigned to HPC procurements. Heriot-Watt University is leading the administrative effort in coordinating these procurements; the idea behind centralising the process is to deliver maximum return on investment while ensuring that the smaller procurements are not overshadowed by their more lucrative £1M+ counterparts. Initial discussions between DisCo and the Heriot-Watt team recognised that while the administrative aspect of the procurements was well understood, a number of technical issues remained largely unresolved. To help address these shortcomings, DisCo has been sub-contracted by Heriot-Watt to provide technical support during the three phases (Tranches) of SRIF3 procurements.

Technical Issues Surrounding Computer Procurements

Our initial requirements analysis highlighted a number of issues where technical guidance and support for the community was crucial:

- Understanding the available technology at the time of the procurements.
- The importance of system and management software in optimizing utilisation of the procured resource.
- The role of Benchmarking in determining the optimum solution given the application portfolio of each University.
- The role of Acceptance Tests in ensuring that the delivered solution is fit for purpose and the options available if this proves not to be the case.
- Understanding the supplier chain i.e. the benefits of procuring solutions from either Integrators, their Tier 1 counterparts or from a partnership involving both.

Recognising the importance of benchmarking, and some of the unrealistic expectations being placed on vendors,

DisCo has designed a core SRIF3 benchmarking suite to test the performance of the key building blocks of cluster-based systems – processors, interconnects and storage solutions – plus, of course, that of the resulting integrated systems. DisCo led in distributing this suite to all vendors involved, in ensuring the benchmarks were understood and successfully run by these vendors, and in analysing the submitted results. This has involved an on-going dialogue with both Universities and vendors to clarify aspects of performance encountered in running the benchmarks. This coordinated approach has enabled key issues to be identified that would otherwise have gone unnoticed.

Evolution of the Benchmarking Suite

A mix of synthetic and end-user codes was chosen to provide a thorough test of the proposed solutions, including CPU speed, memory bandwidth, interconnect performance, compilers, scalability etc. Synthetics featuring in the initial suite associated with the Tranche 1 sites included HPCC, IMB, and IOZONE [1-3], and two end-user applications, CPMD and DLPOLY [4-5]. HPCC and IMB are synthetic benchmarks that test both node attributes and communication fabrics, while IOZONE tests I/O performance of the associated filesystems. Performance characteristics of the end-user applications are well known to DisCo [6]. DLPOLY is the molecular dynamics simulation package; this has little dependency on interconnect, and is fairly insensitive to memory bandwidth and the level of available cache. In contrast CPMD, the ab initio Car-Parrinello package, is critically dependent on interconnect and provides an excellent test of communication capabilities and associated MPI performance.

Our analysis of the Tranche 1 results suggested two significant extensions to the benchmarking suite provided around the Tranche 2 procurements. The first was to add an end-user code that would provide an explicit test of memory bandwidth – a key performance issue given the advent of multi-core solutions. Experience suggested that PDNS3D, the turbulence DNS code [7] and its parallel channel flow benchmark (PCHAN), is well suited given its high dependence on cache and memory issues. The second was to provide revised data sets for the end-user applications that addressed scalability beyond 128 cores. New datasets where incorporated for CPMD and DLPOLY to improve scalability to around 512 processing elements.

Benchmarking Status

Our initial goal when reporting our findings to the Universities was to provide an overall summary that captured the optimum systems as a function of e.g.,

M. Guest, C. Kitchen, M. Deegan, I. Kozin and <u>R. Wain, CSE</u>D, Daresbury Laboratory. interconnect. However, we soon realised it was not this simple, for too many factors are in play to provide such a 'one size fits all', LINPACK-style conclusion [8], even when the Universities base line requirements were firmly understood. Figure 2 provides an example of the performance data for the PCHAN T1 benchmark. Trying to capture this, plus data from the CPMD and DLPOLY benchmarks, into a single number is unrealistic and would mask many important performance issues.

Figure 2. Performance of the PCHAN T1 Benchmark for a range of commodity-based and high-end Computer Systems.

Summary

Our work in technical support of the SRIF3 coordinated procurement has involved the generation of a benchmarking suite that reflects the requirements of the sites involved. This remains an on-going process given the challenge of trying to characterise performance trends and map these back to the Universities user code requirements. Designing a suite of codes and representative datasets that provide a thorough test of the major components of a wide variety of commodity-based systems is, however, a major step in the right direction.

References:

[1] http://icl.cs.utk.edu/hpcc/

[2] http://www.intel.com/software/products/ cluster/mpi/mpi_benchmarks_lic.htm

[3] http://www.iozone.org

[4] http://www.cpmd.org

[5] http://www.cse.clrc.ac.uk/msi/software/ DL_POLY/index.shtml www.cpmd.org

[6] http://www.cse.clrc.ac.uk/disco/ Benchmarks/commodity.shtml

[7] http://www.cse.clrc.ac.uk/ceg/sbli.shtml

[8] http://www.top500.org/about/linpack

Initial Experiences of FPGA Programming

Research performed by the Distributed Computing Group is currently focussed on producing FPGA implementations of kernels that are key to a wide range of scientific applications and using our findings to inform the community about the benefits and drawbacks of FPGA computing. Our primary goal at present, in collaboration with the Advanced Research Computing Group at Daresbury, is to implement a 64-bit floating point matrix multiplication algorithm on Virtex II Pro and Virtex-4 FPGAs in our Cray XD1 system.

What is an FPGA?

A Field-Programmable Gate Array or FPGA is a silicon chip containing an array of configurable logic blocks (CLBs). Unlike an Application Specific Integrated Circuit (ASIC) which can perform a single specific function for the lifetime of the chip, a FPGA can be reprogrammed to perform a different function in a matter of microseconds. Before it is programmed an FPGA knows nothing about how to communicate with the devices surrounding it. This is both a blessing and a curse as it allows a great deal of flexibility in using the FPGA while greatly increasing the complexity of programming it. Over the last two to three years FPGAs have begun to contain enough resources (logic cells/IO) to make them of interest to the High Performance Computing (HPC) community. Recently hardware vendors have begun to offer solutions that incorporate FPGAs into HPC systems where they can act as co-processors, so accelerating key kernels within an application.

This recent surge in interest in High performance reconfigurable computing (HPRC) from the HPC community has come at a time when conventional microprocessors are struggling to keep up with Moore's law. The slowing of performance gains in microprocessors along with the increasing cost of servicing their considerable power requirements has led to renewed curiosity in any technology that might offer a cost-effective alternative. The use of FPGAs in HPC systems can provide three distinct advantages over conventional compute clusters. Firstly, FPGAs consume less power than conventional microprocessors; secondly, using FPGAs as accelerators can significantly increase compute density; and thirdly, FPGAs can provide a significant increase in performance for a certain set of applications.

Potential FPGA Performance

In general terms FPGAs are best at tasks that use short word length integer or fixed point data, and exhibit a high degree of parallelism [1], but they are not so good at high precision floating-point arithmetic [3], (although they can still outperform conventional processors in many cases). The implications of shipping data to the FPGA from the CPU and vice versa [2] must also come under consideration, for if that outweighs any improvement in the kernel then implementing the algorithm in an FPGA may be a pointless exercise. Scientists who are accustomed to continual performance increases provided by each new generation of processor are not easily convinced that the large amount of effort required to port their application will be sufficiently rewarded. If the performance of such codes can be demonstrated by accelerating a number of real-world applications then the wider acceptance of FPGAs will move a step closer. To give an indication of expected performance, data taken from the Xilinx floating-point cores (v2) datasheet [4] is provided. Table 1 illustrates the theoretical peak performance for basic double precision floating point operations in the largest Xilinx Virtex-4 FPGA compared to that of a single core 2.8GHz AMD Opteron.

Device	Multiplier- performance Gflop/s	Add- performance Gflop/s
AMD Opteron 254	5.6	5.6
Xilinx Virtex-4 LX200	23.5	29.6

Table 1. Performance comparison of basic floating point operations running on AMD Opteron and Xilinx Virtex-4.

Summary of programming an FPGA

Our work to date around implementing a 64-bit floating point matrix multiplication algorithm has certainly helped to confirm that programming FPGAs is a non-trivial task. Over the last year, high level programming tools have matured to the point that many now provide support for double precision floating point arithmetic and Virtex-4/Virtex-5 FPGAs. However it is still far from straightforward to get code running on the Cray XD1 platform, in part due to the lack of standardization that has made cross-compatibility between tools, chips and HPRC platforms very difficult to achieve. Until such standards materialise, programmers will still need to make a choice between relying on the assistance of hardware designers and procuring high level tools such as those provided by Mitrion in order to get code running on the Cray XD1.

A number of key points summarise our initial experiences of FPGA programming based on work undertaken using both Hardware Descriptive Languages (HDLs) such as VHDL and Handel-C on development platforms (a Xilinx ML401 development board, see Figure 1) and on the Cray XD1 itself.

Figure 1. Xilinx ML401 Development Board.

1. If you are looking to achieve fast results then it is well worth considering purchasing software such as the Mitrion tools or hardware from vendors like SRC who provide their own high level design tools with their systems. This will save you a great deal of time and effort in getting designs running on your hardware.

2. For a software engineer to achieve ANY useful results it is absolutely necessary that the tool of choice abstracts away the hardware as far as is possible. Thus if the products above are not available, the Cray XD1 platform support library from Celoxica must be fully functional.

3. Even given points 1 and 2, progress is not going to be easy without ready access to an experienced hardware engineer.

R. Wain, M. Deegan, C. Kitchen, I. Kozin, M. Guest, and I. Bush, CSED, Daresbury Laboratory. 4. Only attempt to use HDLs such as VHDL/Verilog if you are in a position to work closely with experienced hardware designers throughout the development cycle.

5. Do not expect to be able to generate easily portable code for FPGAs at this stage. There are currently no recognised standards to support FPGA programming and the current range of tools, languages and hardware are often incompatible.

HPRC is still very much in its infancy, with programming standards and portability between platforms still some way off. This results in a situation where a lot of time and effort can be spent writing software in a soon to be obsolete language for a soon to be forgotten machine. Unfortunately, as far as the machine is concerned, Cray have recently announced that they will no longer be providing software support for the FPGAs on the Cray XD1. They have made it clear that the XD1 does not form part of their future plans. The fact that a system with such potential as the XD1 can fall by the wayside (at least in terms of its FPGA capabilities) so soon after its emergence is an indicator of the unstable nature of HPRC.

Despite the current problems in HPRC there is still a great deal of untapped potential for FPGAs. It is our intention to continue our work in this area in an attempt to quantify that potential. Once we have a clear picture of the real performance gains that are possible we will be able to make an informed judgement as to whether FPGAs are ready for HPC.

A full report can be found on the Distributed Computing group's website [5].

References:

[1] B. C. Schafer, S. F. Quigley, A. H. C. Chan. 11th International Conference, FPL 2001, Belfast, Northern Ireland, UK, August 27-29, 2001.

[2] V. Kindratenko, D. Pointer, D. Raila, and C. Steffen, Comparing CPU and FPGA Application Performance, February 21, 2006.

[3] IEEE Computer Society, New York. IEEE Standard for Binary Floating-Point Arithmetic, 1985. IEEE Standard 754--1985.

[4] Xilinx Logicore Floating-point Operator v2.0 datasheet, Jan 18 2004.

[5] R. Wain, I. Bush, M. Deegan, M. Guest, I. Kozin and C. Kitchen. http://www.cse.clrc.ac.uk/disco/publications/F PGA_overview.pdf

Galahad - A Library of Packages for Nonlinear Optimization

Nature (and mankind) loves to optimize. It is not surprising then, that the optimization (minimization or maximization) of a (linear or nonlinear) objective function of unknown parameters, often with constraints on the permitted values of these parameters, is of crucial concern throughout science, engineering, economics and planning.

Users of scientific software are now accustomed to black-box tools for many commonly-occurring basic problems. For example, few users think twice about solving modest systems of linear equations, Ax = b, their expectation being that software tools will take given A and b and return the required x, perhaps with some assessment of accuracy. In optimization, however, life is not as simple. An optimization problem may have multiple (local or global) solutions. There may be no solution, or the software may not be able to find one. Furthermore, unlike for the case of dense linear equations, it is often not possible to say how much effort it will require to solve a problem of a given size there are fiendishly difficult optimization problems in two or three unknowns. And finally, simply describing a nonlinear optimization problem is frequently far from trivial. Input data is now not simply matrices and vectors, but must somehow include descriptions of objective and constraint functions.

The software library GALAHAD was designed to address many of these concerns. GALAHAD evolved from the highly successful LANCELOT package from the early 1990s, partially as a response to known limitations of LANCELOT, and partially as recognition that there are some specific problem classes – such as those which only involve linear constraints – for which the full generality of LANCELOT would inevitably be inefficient. Thus GALAHAD comprises a number of individual but interdependent packages. GALAHAD is a Fortran 95 library, and makes use of many of the features (data packaging, memory control, modularity) which modern Fortran provides. Version 1 of GALAHAD was released in 2002, and a follow-up is imminent.

An improved version of LANCELOT is one of the cornerstones of GALAHAD. This allows users to solve general optimization problems for which the objective and constraints may be nonlinear. All GALAHAD software is designed with large-scale problems in mind, and in particular LANCELOT uses the concept of group

partial separability (a generalization of sparsity) to store and manipulate data. A crucial design feature of all GALAHAD software is that, as far as is reasonably possible, every underlying algorithm has a corresponding convergence theory under natural assumptions. Alas, it is not possible in general to ensure convergence to a globally optimal point without strong (and often unverifiable) underlying assumptions, and users have to be content with local optima.

Other areas in which GALAHAD provides specialized algorithms include quadratic programming that is minimizing a quadratic function within a feasible region defined by linear constraints, linear programming, for which the objective is linear, and linear and nonlinear feasibility problems. Interfaces to codes for solving linear equations from CSED's mathematical software library HSL are also provided. See Table 1 for details.

GALAHAD is available without charge provided it is not then incorporated within commercial software – a special licence for incorporation is possible. Currently GALAHAD has 560 users worldwide from many application areas. A separate, freely available package, CUTEr, provides access to a large collection of test and application problems, as well as interfaces to external software.

In the near future, GALAHAD will include new methods for solving general optimization problems, and LANCELOT will, at long last, be allowed to retire.

Table 1: primary GALAHAD packages

Package	Purpose
LANCELOT	general nonlinear unconstrained and constrained optimization
FILTRANE	nonlinear least-squares and nonlinear equations
QPA	quadratic programming using an active-set method
QPB	quadratic programming using an interior-point method
QPC	quadratic programming using an interior-point / active-set crossover
LSQP	linear and separable quadratic programming with an interior-point method
PRESOLVE	improve the formulation of a quadratic program before solution
GLTR	quadratic programming within a spherical region
EQP	equality-constrained quadratic programming
WCP	interior-points for linear feasibility problems
LCF	projection method for linear feasibility problem
SILS	symmetric linear systems of equations calling HSL codes
ULS	unsymmetric linear systems of equations calling HSL codes
SBLS	preconditioners for saddle-point problems
SCU	symmetric or unsymmetric block linear systems of equations

Links: http://galahad.rl.ac.uk/galahad-www/ http://cuter.rl.ac.uk/cuter-www/ http://www.cse.clrc.ac.uk/nag/hsl/

N. Gould, CSED, Rutherford Appleton Laboratory; Dominique Orban, École Polytechnique, Montréal, Canada; Philippe Toint, University of Namur, Belgium.

References:

A. R. Conn, N. I. M. Gould and Ph. L. Toint, LANCELOT: a Fortran package for large-scale nonlinear optimization (release A), Springer Verlag, 1992.

N. I. M. Gould, D. Orban and Ph. L. Toint, GALAHAD – a library of thread-safe Fortran 90 packages for large-scale nonlinear optimization, ACM Trans. Math. Software 29(4): 353-372, 2003.

N. I. M. Gould, D. Orban and Ph. L. Toint CUTEr (and SifDec), a Constrained and Unconstrained Testing Environment, revisited, ACM Trans. Math. Software 29(4): 373-394, 2003.

Less is More – Moving HPCx to Power5

The UK's current national high performance computing facility, the HPCx system [1], is managed on behalf of EPSRC by CCLRC, the University of Edinburgh and IBM as consortium partners. Housed at Daresbury Laboratory, the systems administration is provided by the Systems Operations Group of CSED.

The philosophy of HPCx is one of *capability computing*: a robust, general purpose system for calculations and simulations that cannot be done on medium range capacity computing services.

November 2005 saw an interim upgrade of HPCx from IBM's Power4+ (known as Phase2) to Power5 technology to become the Phase2A system. At the operational level the success of this transition was due to the hard work and close collaboration between the HPCx systems group, applications team and IBM.

The move to Power5 keeps us in step with IBM's development of this product line, so that we can get the benefit of their software development. It also means that when we come to upgrade the service to Phase 3 towards the end of 2006, adding more nodes to increase the system's performance to a sustained 12TF will be comparatively simple. The new machine offers several advantages over the old:

- At least the same performance (on Linpack) as the Power4+ system.
- Early sight of what will be the Phase3 machine later this year.
- A doubling of the memory to 2GB per processor.
- A footprint reduction of 75%.

Plans for this transition were first proposed in May 2005 and culminated in the agreed 48-hour interruption of service in November. The hardware was delivered to Daresbury Laboratory and the system built alongside the existing Phase2 machine. This enabled us to perform all the necessary pre-work on the new system without affecting the user service:

- Cabling.
- Power-up.
- System set-up.
- Interim GPFS (general parallel file system) since it was necessary to re-use the Phase2 disk subsystem a temporary independent filesystem was created to allow configuration of the application packages.
- LoadLeveler job scheduler reconfigured and tested for the new 16-cpu units.
- Benchmarking.

40

The Phase 2 service closed on 23rd November 2005 to enable us to:

- Freeze the GPFS filesystems and perform a full backup.
- Perform the GPFS cutover and validation. This involved removing the filesystem whilst leaving the data on the disks, reconnecting these to the new Phase2A system and installing a new version of GPFS on top. All the data successfully reappeared.

The standard measure of establishing performance on such machines is the Linpack benchmark. Phase2A achieved an R_{max} figure of 7.395TF compared with 6.188TF from the Power4+ Phase2 system and well ahead of the contractual requirement of 6TF.

User service resumed on the new machine, as planned, on 25th November. Phase2 had won the reputation of being the most reliable large HPC cluster in the world and we are on track to repeat this achievement with the current system and, we believe, with Phase3 later this year. So what is new? From a systems viewpoint, operation of the new hardware is considerably different from the old p690-based service, requiring considerable changes to be made to our procedures, monitoring and reporting of the system status. In addition to the two login and i/o server frames there are 12 compute frames each containing eight p575 servers. Each of these servers in turn contains eight dual-core Power5 1.5GHz processors giving a total of 1536 cpus. Details of the architecture and initial performance results may be found in an HPCx Technical Report [2].

The combination of AIX5.3 and Power5 offers the potential for running jobs in Simultaneous Multithreaded (SMT) mode where the 16 physical processors on a node can masquerade as 32 virtual ones [3]. We have seen good results for a number of applications, especially those that saw little benefit from Phase2A such as molecular dynamics codes, and this has now been implemented as a user-controlled option.

November 2006 will see the opening of the 12TF Phase3 system and planning for this began in February to ensure that the next transition will be as smooth and unobtrusive as the last.

S. J. Andrews, T. J. Franks, I. A. Georgeson and I. M. Shore, CSED, Daresbury Laboratory; M. W. Brown, EPCC, University of Edinburgh; T. Davis, IBM UK. **References:** [1] http://www.hpcx.ac.uk

[2] http://www.hpcx.ac.uk/research/hpc /technical_reports/HPCxTR0602.pdf

[3] http://www.research.ibm.com/journal/rd /494/mathis.html

Contacts and Further Information

Prof Paul Durham	Director, Computational Science and		
	Engineering Department	p.j.durham@dl.ac.uk	
Prof Martyn Guest	Associate Director, Computational		
	Science and Engineering Department	m.f.guest@dl.ac.uk	
Dr Richard Blake	Associate Director, Computational		
	Science and Engineering Department	r.j.blake@dl.ac.uk	
Mrs Shirley Miller	Personal Assistant	s.miller@dl.ac.uk	
Mr Damian Jones	Administrative Assistant	d.r.jones@dl.ac.uk	
Miss Nina Woodall	Administrative Assistant	n.m.woodall@dl.ac.uk	

Group leaders

Dr Steve Andrews	Systems and Operations	s.j.andrews@dl.ac.uk
Dr Mike Ashworth	Advanced Research Computing	m.ashworth@dl.acuk
Dr Miles Deegan	Distributed Computing	m.j.deegan@dl.ac.uk
Prof lain Duff	Numerical Analysis	i.s.duff@rl.ac.uk
Dr Dave Emerson	Computational Engineering	d.r.emerson@dl.ac.uk
Prof Chris Greenough	Software Engineering	c.greenough@rl.ac.uk
Prof Nic Harrison	Computational Materials Science	n.m.harrison@dl.ac.uk
Dr Bob McMeeking	Chemical Database Service	r.f.mcmeeking@dl.ac.uk
Dr Chris Morris	PIMS	c.morris@dl.ac.uk
Dr Martin Plummer	Atomic and Molecular Physics	m.plummer@dl.ac.uk
Dr Paul Sherwood	Computational Chemistry	p.sherwood@dl.ac.uk
Prof Walter Temmerman	Band Theory	w.m.temmerman@dl.ac.uk
Dr Martyn Winn	CCP4	m.d.winn@dl.ac.uk

Websites

Computational Science and Engineering Department	www.cse.clrc.ac.uk
HPCx Service	www.hpcx.ac.uk
EPSRC HPC Programme	www.epsrc.ac.uk/hpc
CCLRC	www.clrc.ac.uk

Addresses

CCLRC Rutherford Appleton Laboratory

Chilton Didcot OX11 0QX Tel: 01235 445000 Fax: 01235 445808

CCLRC Daresbury Laboratory

Keckwick Lane Daresbury Warrington WA4 4AD Tel: 01925 603000 Fax: 01925 603100