



High Performance Computing Report 2014

Introduction

Foreword



he University has long recognised the importance of High Performance Computing (HPC) in supporting our world-class research community and since 2006 has invested £10 million in the BlueCrystal HPC and BluePeta data storage facilities. BlueCrystal is the central University HPC facility, housed

within two state-of-the-art, dedicated machine rooms. BlueCrystal has already provided enhanced computing capability to over 1,000 researchers, leading to cuttingedge research across a wide range of disciplines including climate modelling, drug design and gene sequencing.

The University recently made a further capital investment of £2 million in a new phase of the HPC facility, one which utilises the latest GPU and shared memory technologies to supplement the 5,000-plus cores. This investment provides Bristol with one of the largest and fastest UK university supercomputers, which will be used not only by researchers, but by the growing number of students being trained in the use of HPC, and by a range of industrial collaborators.

This brochure showcases some of the exciting research being undertaken using BlueCrystal. It reflects the range of disciplines now using HPC and provides an insight into both current research and how researchers are planning to exploit the increased capability and the latest technologies that the new machine is providing.

Professor Guy Orpen Pro Vice-Chancellor for Research Chair. HPC Board

Contents

Introduction	
Foreword	
Training tomorrow's programmers	3
Modelling the planet	
Climate, complexity and carbon	
Telling evolutionary time	5
From worlds to molecules	
Computing among the stars	6
Genetic sequencing:	
some reassembly needed	
A broader appeal	
A bigger picture of health	8
Computing beyond science	9
New frontiers	
Graphic exploitation: GPUs in research	10
Making light work of nanophotonics	
Working with industry	
Material gains in composites research	12
Complexity in the air:	
fluid dynamics on BlueCrystal	13
A new era in drug design	14
Facts and figures	
BlueCrystal Phase 3	15

Editorial

High Performance Computing Report 2014

Editor

Managing editors Dr Ian Stewart

Contact Caroline Gardiner Advanced Computing Research Centre Bristol BS8 1HH caroline.gardiner@bristol.ac.uk

Design cw-design.co.uk

Produced by

Printed by

© University of Bristol 2014 Extracts may only be reproduced with the permission of the Public Relations Office.

If you need part or all of this publication in an alternative format, please telephone +44 (0)117 928 8895.

Training tomorrow's programmers

For three years, the Department of Computer Science has been running an undergraduate course in High Performance Computing.

'It's optional but extremely popular,' says Simon McIntosh-Smith, the course's director. 'The course focuses on applied parallel programming, and teaches our students how to use high performance computing systems like BlueCrystal. Almost all eligible students choose to take it, and in the academic year 2012-13 it had the highest student feedback of any of the 75 or so courses run in Computer Science.'

As a result of this, many computer science students have gone on to choose final-year projects in high performance computing, often in collaboration with academics across the University. Projects with researchers in Biochemistry, Aerospace Engineering, Chemistry, Physics and Geographical Sciences have all taken place since the HPC programme began. Having become well-versed in high

Many computer science students have gone on to choose final-year projects in high performance computing, often in collaboration with academics across the University

BlueCrystal has become a core part of teaching as well as research. Whereas most universities have separate machines for teaching parallel computing, students at Bristol are given accounts on BlueCrystal itself. 'That's much more exciting for them and much more realistic,' says McIntosh-Smith; 'they get to see what it's like using real supercomputers in real circumstances.

performance computing, some of these students are then taken on as interns and Research Assistants, helping to run the facility and working closely with users to optimise their software in order to get the most out of BlueCrystal

Bristol becomes an Intel® Parallel Computing Center

Intel announced in January 2014 that it has selected the University of Bristol as an Intel® Parallel Computing Center (IPCC), in recognition of the world-leading research carried out by the University's High Performance Computing (HPC) group into the efficient use of many-core parallel computer architectures, and its leadership in driving open parallel programming standards.

This is one of several IPCCs that Intel is establishing with leading HPC research

groups in order to tackle the challenge of using modern computer systems more efficiently. It is the first IPCC in the UK.

The University will collaborate with Intel to develop parallel algorithms and optimise several of Bristol's HPC applications for Intel's latest parallel computer architectures, such as the Intel® Xeon Phi[™] Coprocessor. Programmes of HPC training will also be developed, drawing on Bristol's undergraduate and Masters courses in high performance computing.



Profile Mike Boulton **Research Assistant** Department of Computer Science

While I was an undergraduate, we did a group project on image processing that involved getting some algorithms to run fast on a tiny little ARM chip. It was challenging, but after taking the **High Performance Computing course** and using BlueCrystal we discovered several ways of speeding it up. That experience paid off when I worked on a fluid dynamics project for the Atomic Weapons Establishment, using what was the fastest computer in the world. Titan at Oak Ridge National Laboratory in the US. Now I work as an RA at Bristol, helping researchers to run their projects on BlueCrystal.

Joe Curley, Director of Marketing for Technical Computing Group at Intel Corporation, said: 'Intel is pleased to expand our Intel® Parallel Computing Center program by collaborating with the University of Bristol, to innovate and optimize applications in the field of computational fluid dynamics and molecular dynamics to benefit research and industry in Europe and worldwide.'

Modelling the planet

Climate, complexity and carbon



Profile: Dr Gethin Williams Senior Support Specialist Advanced Computing Research Centre

Part of my job is teaching 'pragmatic programming' skills to domain scientists. These key 'workbench' skills are targeted at those who use simulation but can safely skip over the more esoteric aspects. Writing code, for instance, is something a lot of researchers pick up on the fly, but that can result in 'spaghetti code' that is incredibly hard to fix or augment. We train them to think more in terms of creating more structured 'Lego code'. We train researchers in programming and scripting languages, but there are other useful practices, like version control, which can save a lot of time in finding bugs in your source code and also make it easier to reproduce your results - a crucial factor in scientific research.

Given the levels of complexity and uncertainty involved in simulating the systems that govern a whole planet's environment, it comes as no surprise that climate modelling accounts for some of the most intensive usage of BlueCrystal.

The Bristol Research Initiative for the Dynamic Global Environment (BRIDGE), established in 2003 by Professor Paul Valdes, aims to improve our understanding of the Earth's climate and environmental variability. Dr Dan Lunt is among several researchers at BRIDGE who are using computational methods to test predictive models of future climate change.

'These models are founded on a set of relatively simple equations of fluid mechanics, but they can produce unexpected, chaotic behaviour – "El Niño" events, for example,' says Lunt. 'Furthermore, making relatively small changes to the model can result in quite different answers for climate prediction, all equally valid. So to get a full appreciation of the uncertainty involved, you have to do a lot of simulations. The great advantage of Phase 3's increased computing power is that we can do a comprehensive series of runs to explore that uncertainty.'

The great advantage of Phase 3's increased computing power is that we can do a comprehensive series of runs



A computer model showing contours of annual average temperature in the North Atlantic during the Cretacious period.

At the core of most climate models is a grid method that divides the planet into boxes, layered up into the atmosphere and down into the oceans, and uses the equations from basic laws of physics, atmospheric chemistry, and biology to assign values (for temperature, wind speed, humidity, chemical composition, vegetation type and so on) inside each box. Where models differ is in the finer details: many processes in the atmosphere and ocean (for example, clouds in the atmosphere or vegetation on the surface) occur on a spatial scale smaller than a single box, so the models must approximate these processes in various ways.

'You put all these things together, run the model, and compare its predictions with meteorological observations from weather stations around the world,' says Lunt. 'Then you can run sensitivity studies by adjusting the model's parameters. If you want to run predictions of future climate, you can increase carbon emissions, for example; or if you want to find out more about how the system works in general, you can do something more drastic like taking away all the ice.'

Lunt and his colleagues at BRIDGE have taken the comparison studies further by examining how their model predicts past climate. For that, they work with colleagues in the School of Earth Sciences to establish a picture of past climate based on the geological record – ice cores from Antarctica, for example, or sediment cores from lake beds. 'Then we run our model as though it was, say, 50 million years ago, by moving the mountains and the continents around, and we see how well it predicts the conditions indicated in the records,' he explains. 'If it does a good job of predicting the climate 50 million years ago, when there was a lot of CO₂ in the atmosphere, we can have more confidence about the model's predictions of the future. There will always be a level of uncertainty, but at some point that level will be low enough to yield clear information that can inform public policy and contribute to decision making.'

Lunt echoes the sentiments of many researchers who have used BlueCrystal when he says: 'One of the great things about supercomputing at Bristol is that it's free at the point of use: if we have an idea that we want to test, we can just do it. That isn't always the case elsewhere. If you're writing a grant proposal for a large project, for instance, you can do a pilot study first and present that in the proposal. Things like that give us a tremendous advantage over our competitors.'

Telling evolutionary time

Reconstructing the family tree of life on Earth is no small task. Add to that the figuring out of the ages of the family members – when one species diverged from another, for example – and the work strains at the limits of what is currently possible.

The field of molecular evolution challenges these limits, attempting to account for biological change at the molecular and cellular level using the principles of evolutionary biology and population genetics, in particular a method known as the molecular clock. Professor Philip Donoghue studies the evolutionary emergence of



What I'm able to achieve now wasn't possible at the beginning of my PhD

animals and plants, using molecular genetics to calibrate the Tree of Life to time using molecular clock theory, and to determine the role of genetic regulators of development in effecting evolutionary change in organisms. His PhD student, Rachel Warnock, is looking at the dating of these branching events on the Tree of Life. When did animals diverge from fungi or plants? When did vertebrates crawl out of the water onto land? When did humans branch from their relatives?

It's a field rife with uncertainty: the fossil record can help to provide estimates that enable researchers to bracket an interval of time within which evolution occurred, but it's imprecise and there are huge gaps in the record. There is a multiplicity of methods for using the molecular clock to construct calibrations of the timespans involved and to date divergences, but, says Warnock, 'we're not satisfied that anyone's ever tested these methods properly.' Warnock uses a Bayesian framework to characterise the uncertainty associated with the calibrations, and is testing these different methods using simulated fossil data – a new approach made possible by BlueCrystal's vast resources. 'The fossil record is incomplete but in a very structured, nonrandom way, which means we can simulate it,' says Warnock. 'We can then build calibrations, run molecular clock analyses, and work out which methods are working properly.'

Producing sophisticated simulations of evolutionary trees, of molecular evolution on those trees, and of the fossil record itself, is, says Donoghue, 'a very complex, multidimensional statistical problem – and that's what we need BlueCrystal for'. It enables Warnock to rerun evolutionary history millions of times and assess whether the molecular clock can correctly predict the ages of those divergence points. 'I'm really excited about using Phase 3,' she says. 'We'll be doing hundreds of thousands of simulations and molecular clock analyses, all with a whole range of variables. Phase 3's GPU nodes will enable the molecular clock analysis to be done much faster, because the specialist programme we use for likelihood calculations is able to exploit this new technology.

'The increase in speed and volume with Phase 3 will be great for me, and will mean simultaneous access to more processors. What I'm able to achieve now, and what I anticipate being able to achieve by the end of the project, wasn't possible at the beginning of my PhD.'

From worlds to molecules

Computing among the stars

How do planets and solar systems form? And why are there so many different types of planet? Dr Zoë Leinhardt is a computational astrophysicist at Bristol whose Planet Formation Group is trying to answer these and other fundamental questions.

Around one thousand planets outside our solar system have been discovered, with at least another two thousand candidates identified. Some are gas giants similar to Jupiter or Saturn; others are smaller, rocky planets. Just about every planet type and planetary system imaginable has been observed.



Profile: Linford Briant PhD student, Department of Engineering Mathematics / School of Physiology and Pharmacology

My PhD project, funded by the EPSRC, is to model the control of our blood pressure by a particular node in the sympathetic nervous system: the sympathetic preganglionic neurone. The system can go into a state of hyperactivity, leading to the development of hypertension (high blood pressure). To investigate the mechanisms underlying the increase in sympathetic preganglionic output that causes hypertension, we developed a mathematical model of this neurone and its control of the blood pressure. The multi-scale nature of this biological model means it requires a lot of computing power to solve. On BlueCrystal, we've been able to run simulations with parameter sweeps identifying key parts of the sympathetic pathway that can account for the increased activity seen in hypertension. The staff at the Advanced Computing Research Centre have been very generous with their time. They helped me install new modules onto BlueCrystal so that I could run my model - a task that would have taken me much longer on my own.

Leinhardt and her group are using numerical models to determine whether it's possible to predict how these diverse planetary systems will evolve from their initial conditions. This work relies heavily on numerical simulations.

Stars evolve from vast, cold molecular clouds. Some event, such as the explosion of a distant supernova, triggers density waves that cause the cloud to start collapsing under its own gravity, forming a star at its centre. The cloud material forms a hot, dusty disk, known as the proplanetary disk, around the star. As this disk cools, some of its gasses condense into solid materials and begin to clump together to form rocky asteroids called planetesimals, which continue to hit each other and build up into planets.

'We have observations of the proplanetary disk,' says Leinhardt, 'and observations of planetary systems, but these are only snapshots. We have to put them in order, but we don't know what the order should be ahead of time so we have to guess a bit. The stages between the proplanetary disk and the final planetary system are pretty much a black box.'



This image shows the intensity of light emitted by dust in a young star system containing a Jupitermass planet, when viewed at a wavelength of 850 microns. The colour scale indicates if a large amount (red regions) or small amount (blue regions) of light is emitted by the dust in the disk. The dust shown here is produced in collisions between planetesimals. A gap can be seen coinciding with the orbit of the planet as it sweeps up and scatters the material it encounters. The dust disk also shows asymmetries resulting from the effect of the planet's gravity upon the planetesimals.

Assistant (Computer Sciences graduate Mike Boulton, see p3), to modify PKDGRAV to work more efficiently. But to get the most out of Phase 3, PKDGRAV will need a node that uses a CPU and GPU together.

Around one thousand planets outside our solar system have been discovered, with at least another two thousand candidates identified

To shine a light into this box, Leinhardt has developed a collision model and integrated it into a large n-body code, which numerically simulates the evolution of a system of bodies that are constantly interacting with each other via gravity. The code, called PKDGRAV, was written at the University of Washington where Leinhardt gained her doctorate, and she modified it for her PhD. 'It's a highly parellised code written in C, and it works on BlueCrystal really well,' she says.

The code wasn't written to run on GPUs, so Leinhardt has had help from a Research

The underlying gravity calculation, which is repeated n log n times, would run on the GPU, and everything else – the prediction of where the particle was going to go and whether it was going to collide with something – would run on the CPU.

With these modifications, Leinhardt's group will be able to start running much bigger models, involving not just millions of particles, but billions. Leinhardt will also be putting BlueCrystal to use on a joint project with colleagues in the School of Earth Sciences to explore the origins of Earth's formation.

Genetic sequencing: some reassembly needed

Sequencing the human genome was a landmark achievement. But as our understanding of the interconnectedness of living systems continues to grow, so does the need to study the genetic makeup of other animal species.

One major reason for this urgency is the appearance in the human population of viruses that have crossed the species barrier: coronaviruses like SARS and MERS-CoV plus others such as Ebola and Henipaviruses. These viruses, all of which originated in bats, have been responsible for fatal outbreaks in the last 50 years.

'We've evolved over millions of years to cope with our viruses,' says Dr David Matthews, a virologist in the Department of Cellular and Molecular Medicine who studies the genetics of cells infected with viruses. 'We've developed genetic responses to defend ourselves against viral infection. But all viruses carry genes of their own that they use to modify the host environment by

Unpicking the genetic content of another species in this way requires access to a huge amount of RAM, and BlueCrystal 3's large memory system provides exactly this: up to several terabytes of RAM that the software can access. 'We use a technique called de novo assembly, says Matthews; 'you fragment an animal's genetic information into millions of bits, each a couple of hundred letters long, and sequence them. So you have millions of short sequence reads that all overlap each other and they need to be reassembled. It's like shredding a couple of thousand copies of an encyclopedia and then putting them back together. You can only do that on a computer that can hold the entire problem in its RAM memory.'

Unpicking the genetic content of another species requires access to a huge amount of RAM, and BlueCrystal 3 provides exactly this

switching the host's genes on or off so that the virus can quickly get on with making new viruses. We're interested in exactly which genes get switched on or off.'

Recording these changes in gene expression, not just in humans but in other species, generates massive datasets that need a computer cluster like BlueCrystal to analyse properly. In a joint project with the Australian Animal Health Laboratories, Matthews has been using Phase 3 to compare bat genes with human genes and ask a series of pressing questions. 'When a virus gets into bat cells, which of the bat's genes are switched on or off? When you put the same virus into human cells, is the response different? And can that help us to understand why these viruses don't seem to make bats ill, but often prove fatal to humans?' The new sequencing technologies can be used for any living thing, and the consequent broadening of BlueCrystal's potential user base has prompted Matthews and his colleagues to plan the installation on Phase 3 of 'software that we hope will make it more intuitive and easy to use for less computationally savvy users in the biology community'.

'High-throughput studies like this are fast becoming routine in modern biological research,' says Matthews, 'and we're able to do world-leading work at Bristol because of BlueCrystal. For researchers, like me, who aren't part of a big conglomerate, it's a phenomenal resource.'



Profile: Dave Platt Heilbronn Research Fellow School of Mathematics

My main interest is in analytic number theory. Recently, I helped Harald Helfgott prove the ternary Goldbach conjecture which states that every odd number greater than five is the sum of three primes. This had been outstanding since 1742 and my contribution was to check a huge number of cases (about 10¹¹ of them) by computer. One key attraction of BlueCrystal is that it uses Error-Correcting Code (ECC) memory which gives us protection against random events like cosmic rays flipping a bit in memory. Without ECC, such events cause programs to crash or, even worse, silently produce incorrect answers. Another advantage is the availability of the experts in the ACRC support team who provide advice and assistance and allow me to focus on the mathematics rather than on the systems.



Black flying foxes (Pteropus alecto) in Kakadu National Park, northern Australia.

A bigger picture of health

When researchers at Bristol recruited 14,000 expectant mothers in 1991 and began to follow the progress of their babies, the project's aims were broad: to collect data on the growth and development of a large cohort of individuals from birth to old age.

Over 20 years later, the Avon Longitudinal Study of Parents and Children (ALSPAC) is an unrivalled resource for researchers across the world, and high performance computers like BlueCrystal offer the tools to mine this resource in ways that could barely have been anticipated when ALSPAC began.

Dr Nic Timpson, Reader in Genetic Epidemiology at the School of Social and Community Medicine, conducts populationbased studies using data collected from ALSPAC participants, ranging from obstetric data and questionnaires to biological samples. 'On any given individual in the ALSPAC cohort we might have over 20 years of observational data,' he says. 'Computationally, that's tractable – most of it consists of single variables – but combining data from many sources, including genetics, is more of a challenge.' This tremendous rise in capacity brings with it an increase in data being generated. Studies can involve hundreds of thousands of individuals, with several million variables being collected simultaneously for every person. And genetics is just one part of a multi-omic data collection regime. 'As well as the genomic data, we often have data on the transcriptome, which is the product of the genome, or on the epigenome – external factors that influence the genome,' says Timpson. 'The really valuable part of bioinformatics lies in integrating these and similar data together, and that demands a massive amount of processing power.'

Phase 3 of BlueCrystal provides this power, and the ACRC's Research Data Storage Facility is also being upgraded to cater for the ever-increasing volume of data generated. 'This will help us keep

Studies can involve hundreds of thousands of individuals, with several million variables being collected simultaneously for every person



An ALSPAC participant generates more data.

The standard scientific hypothesis-based approach proved a bad match for the sheer scale and complexity of the human genome: to test a hypothesis concerning, say, the role of insulin in the presentation of diabetes, previous efforts worked systematically through related genes, screening each one for variations among different individuals to look for an association with diabetes. It was a flawed, restrictive approach that slowed the pace of discovery.

The advent of new technology for the measurement of biological data, coupled with high performance computing, has revolutionised the process: researchers can ask for all variable sites in the human genome at a given frequency and interrogate them all at the same time. The implications for research are huge: 'You don't need a hypothesis anymore,' says Timpson, 'you simultaneously pose all hypotheses and let the data show you where the answer is. We started finding dozens of genes at the same time that contribute to disease, often in regions of the genome that would have been ignored before.'

pace,' says Timpson, 'and our Integrative Epidemiology Unit, funded by the Medical Research Council, is absolutely reliant on BlueCrystal.'

Timpson's mission to translate basic science into clinical application is picking up speed, thanks to BlueCrystal. 'Now that we can study a couple of million measures across, say, 40,000 people, we're getting some traction on problems such as the mechanisms of type two diabetes, or the relationship between adiposity and heart disease,' he says, 'and that should have a direct benefit for public health.'

Computing beyond science

As BlueCrystal's facilities evolve, so does the diversity of its users. What was once the preserve of science and engineering is now helping a small but growing number of academics working in the humanities and social sciences.

Dr Paddy Carter in the Department of Economics has been using BlueCrystal to look at a rule that major development agencies such as the World Bank use to allocate foreign aid. His research focuses on a specific allocation formula based on two characteristics: how poor the recipient country is, and how effectively it is estimated that the country will use the aid. currently exploring the variation in marriage patterns among Austronesian-speaking societies in the Pacific. The question at the heart of her work is why some societies in the region have evolved norms of monogamous marriage while others allow polygyny. She is using language relationships together with phylogenetic comparative methods to work out how marriage practices

Even bearing in mind the effort required of a non-programmer to prepare his model for running on BlueCrystal, he still saved many hours

To examine what determines the optimal weights given to these characteristics, Carter ran a simulation using a simple macroeconomic model of 20 aid recipients. 'The model involves non-linear differential equations,' he says, 'and I needed to solve the model for all 20 countries at once, which increased the running time to half an hour for a single iteration. I'm a programming novice, so I did a very basic grid search to look for the optimal coefficients and then used a finer grid to get a more precise number. That could involve thousands of runs by the time I'd got my optimal rule down to about three decimal places. And that was just for one experiment; I wanted to run lots of different experiments to see what the optimal rule would look like in different setups.'

Carter decided to move the project from his desktop to BlueCrystal. Even bearing in mind the effort required of a non-programmer to prepare his model for running on BlueCrystal, he still saved many hours of watching his computer crunch numbers.

Another academic making good use of BlueCrystal is Dr Fiona Jordan, Lecturer in Anthropology. Jordan studies cultural evolution, especially kinship issues, and is evolved, which is where BlueCrystal comes in. 'If you have more than about 15 different societies this quickly becomes a problem that you can only do with supercomputers,' she explains.

For Jordan the facilities offered by BlueCrysta were a big factor in her decision to bring her research to Bristol from the Max Planck Institute for Psycholinguistics. She is using software developed at the University of Reading by evolutionary biologist Mark Pagel, along with Andrew Meade. 'Pagel created these phylogenetic simulation programmes that don't just work on genetic material, but also have models built into them that are flexible enough to allow for cultural or linguistic traits.

Jordan hopes that more of her colleagues will realise the advantages of using HPC. 'I come from an interdisciplinary background – biology as well as anthropology,' she says. 'Biologists know about parallel computing and know those resources are out there, but a quantitative social scientist probably doesn't realise how handy BlueCrystal could be to them. They often run statistical modelling on their desktop, because the things they're doing don't require huge processing power, but even the ability to do it in minutes rather than hours is really useful.'



New frontiers

Graphic exploitation: GPUs in research

Graphics processing units (GPUs), originally

for the computer games industry, are now

designed to provide smooth, fast-moving graphics

increasingly used by the scientific community to accelerate numerical computations. BlueCrystal

Phase 3 has 78 of them, and that opens up new



A computer model showing Tamiflu® being expelled from the binding site of a drug-resistant mutant of influenza neuraminidase.

Professor Adrian Mulholland and Dr Christopher Woods work in the field of computational chemistry, using computer simulations to examine how drugs bind to proteins and to assess the implications for the treatment of viruses. Woods is developing software that creates a kind of virtual microscope, allowing scientists to see exactly how the proteins move and change. This should in due course help pharmaceutical companies to make their drugs more effective.

A case in point is the H7N9 avian influenza virus, which killed more than 30 people in China in March 2013, having rapidly become resistant to the anti-viral drug Tamiflu®. When H7N9 first emerged, Woods was able to take the DNA sequence of the virus - extracted from patients' blood samples - and use it to show in simulation what the viral protein looked like and which versions of it could be treated by Tamiflu®. This allowed him to predict accurately whether a drug would work or not, and in a fraction of the time it would take doctors on the ground to figure it out.

In 2000, when Woods first started simulating viral proteins on a desktop computer, it took him almost a whole summer to move just 500,000 steps, or frames, of his simulation – the equivalent of one nanosecond (10-9 of a second) of protein movement. The introduction of BlueCrystal enabled him to accelerate his progress dramatically, and the GPU blades in Phase 3 will allow him to run 400 million steps per month. 'Once you hit this point, you can actually see the drug moving and being thrown out,' says Woods. 'We'll be able to explore what's really happening in the binding site and how the protein is evading the drug.

In the School of Physics, several projects are using BlueCrystal's GPUs for computational modelling of soft matter and optics. One such project involves liquid crystal devices (LCDs), the subject of research by Dr Simon Hanna, Professor Robert Richardson, and their PhD student, Tom Fenech.

LCDs, such as displays and optical switches for use in telecommunications, are controlled via the switching of liquid crystal molecules by a small electric voltage. Their evolution from the seven-segment displays of early digital watches to the sophisticated detail of HDTV

reserves of potential for researchers. Fenech has been using BlueCrystal to model what happens to the liquid crystal molecules in the vicinity of an added particle - in this case, a gold nanorod. A small region of liquid crystals, typically smaller than one display pixel, is divided into a cubic grid of cells, each of which is assigned parameters for its cluster of molecules - primarily their orientation and their interaction energy, which indicates how each site will interact with its neighbours. Other energy terms are assigned wherever a cell includes a surface, such as that of a nanoparticle, which has a

particular effect on the liquid crystals.

We could be working with around one million cells and running through millions of iterations. Phase 3's GPUs can run it all in parallel very rapidly

reflects a relentless push to improve their performance and efficiency. A new strategy being pursued is adding nanoparticles carbon nanotubes or gold nanorods, for example - to the liquid crystals.

'Adding nanoparticles causes the liquid crystal molecules to change their orientation,' says Hanna. 'You can switch the nanoparticles by applying a voltage, and they take the liquid crystals with them as they switch around. The idea is that the nanoparticles could pre-orient the liquid crystals, so you get a faster switchover, which requires a lower voltage,' It's an idea that leads to complications, however. Liquid crystals are made of rod-like molecules that like to stay parallel to each other, but adding nanoparticles can sometimes change their orientation too abruptly, leading to small defects in the display.

'We run the simulation to find out which configuration of these nanoparticles involves the lowest energy,' says Hanna. 'We could be working with a grid of around one million cells and running through millions of iterations. Phase 3's GPUs can run it all in parallel very rapidly and spit out the result, then we run the next iteration. At the end we will have a "director map" showing us how the liquid crystal is moving and changing orientation, how defects form, and how the energy of particles varies with their orientation and separation from each other.'

Making light work of nanophotonics

The field of nanophotonics - the study of the behaviour of light at the nanoscale - has begun to yield dramatic improvements in devices that interact with light.

Solar cells, for example, have been around for a while, but using nanotechnology to change the structure and patterning on the surface of the cells can dramatically increase their production of electricity. Martin Cryan, Professor of Applied Electromagnetics and Photonics, is using BlueCrystal to conduct several projects to determine the most effective structures for these cells.

'The starting point is the simplest solar cell, a flat piece of silicon,' says Cryan. 'If you start structuring the surface, maybe adding nanoparticles to it, you can improve its efficiency. To work out how well that modified surface works, you need to know how much light at a particular wavelength gets absorbed by the surface and how much gets reflected. That's where you need to do some serious number-crunching; the equations concerned are relatively simple, but unless you've got the computing power, you can't get enough detail to see whether these tiny materials are helping to absorb the light.

Cryan's methodology involves dividing a simulated solar cell, measuring only 1mm square, into many millions of extremely tiny cells. The equations for the simulations are then solved by running them in parallel on BlueCrystal's processors.

'Making the cells smaller introduces an extra level of detail that can actually change the results quite radically,' says Cryan. 'So we run a "standard" simulation, get an answer, then run a much bigger simulation where the cells are even smaller, and if we still get the same answer, we know that it's definitely right. BlueCrystal gives you that kind of flexibility.'

John Rarity, Professor of Optical Communication Systems, is also looking for new ways to make light interact with matter, in an effort to refine the technology behind nanophotonic devices such as optical switches.

'At present, optical switches are made with very highly nonlinear materials that use 105

photons; we're trying to get five orders of magnitude below that in terms of intensity of switching,' he says. 'So we're working on building structures that can trap a single photon.

Rarity's postgraduate students Mike Taverne and Xu Zheng and Research Assistant Dr Daniel Ho have been using BlueCrystal to build virtual models of three-dimensional structures, made from transparent materials. Using high-refractive-index materials with structure periodic on the wavelength scale to block the propagation of light at certain wavelengths, they model traps (cavities) that hold light for up to a million optical cycles.

'We put in a piece of solid material where light can propagate in the centre, the tiniest amount we can, so that it traps the smallest volume of light possible,' says Rarity. 'Any effects of light interacting with matter are thus enhanced to the ultimate level. Effectively we ring a "bell" in the middle of this structure - an optical bell, at an extremely high frequency - and watch that ring down, as the light slowly leaks away. That tells us what the resonance frequency is for this cavity, and how long the light is trapped.'

This resonance can survive for millions of oscillations, and doing the calculation requires solving a set of equations, called Maxwell's Equations, on a grid that might consist of a million cells. After dividing each oscillation into some 20 time steps, the researchers calculate the resonance field over every cell, then move to the next time step and repeat it, and so on. Millions of time steps, each one involving a million cells; it's easy to see why a supercomputer like BlueCrystal is invaluable for this kind of work.

'Phase 3 enables us to run these simulations faster in parallel to test the effect of changing various parameters or to split up large geometries to run them in parallel over multiple nodes,' says Mike Taverne. 'It also gives us a RAM upgrade from Phase 2's 1GB per core to 4GB per core. That means we can request a single processor per node,

which allows us to run lots of non-parallel simulations at the same time.

The applications of single-photon switches are potentially world-changing - they could be a key component in a quantum computer, for instance - but for now, the work made possible by supercomputers like BlueCrystal is giving Rarity and his colleagues the clearest view yet of the behaviour of light at the quantum level.



Profile Dimitra Simeonidou **Professor of High Performance Networks**

Supercomputers increasingly use optical technologies to reduce latency and energy consumption in server communication and to support very high data volume for processing. We are developing new hardware solutions and software tools to optimise the energy and computational efficiency of high performance computers like BlueCrystal, so that data can be moved around very flexibly while data processing speeds and storage capacity can increase further. These kinds of refinements are part of my group's work which drives the evolution of the next generation of supercomputers and data centres. It's predicted that, by 2022, servers in HPCs like this will have multiple interfaces running at 1Tb per second; the innovations we're working on will prepare for that future while offering huge benefits for current users of high performance computers such as BlueCrystal 3.

Material gains in composites research

The Advanced Composites Centre for Innovation and Science (ACCIS), based in the Faculty of Engineering, brings together several areas of composite research and collaborates with government funding agencies and industry partners to develop and exploit this research.





Rolls-Royce, which increasingly aims to use composite materials in the manufacture of its components, is one of those partners. For the aerospace, automotive and other industries, composite materials have many advantages: they're lightweight, and fatigue and corrosion-resistant. But they also have their downsides, including a susceptibility to impact damage.

Professor Stephen Hallett's research group is examining the mechanisms of failure in these materials. Hallett is Technical Director of the Composites University Technology Centre (UTC), which forms part of ACCIS and is supported by Rolls-Royce. 'It is our expertise in composite materials that has attracted Rolls-Royce to the University of Bristol, but it's also capability,' says Hallett, 'and part of that capability is high

better predict the variations in the material to understand in more depth how they might fail. Green's research concerns the numerical modelling of a repeating volume element a small section of the material, in this case around 2cm x 3cm, that can be used to determine the complex internal architecture of the composite. Meanwhile, FI Said uses Green's model of the repeating volume element to build a complete structure in simulation, in order to test how the material is likely to perform at points where it is most flawed.

Green and El Said are working at different ends of the length scale: Green is using BlueCrystal to run high-resolution models of small sections of material, while El Said is constructing much larger models - with around 10 million elements - but with a

Hallett's group have used BlueCrystal to run increasingly complex models

performance computing.' Since the arrival of BlueCrystal, Hallett's group have used it to run increasingly complex models. Two of his PhD students, Steve Green and Bassam El Said, are looking at modelling strategies for the latest generation of composite material, made from woven carbon fibre, in order to allow increased optimisation and provide benefits over and above the current 3D woven components in service.

Unlike traditional composites, which consist of layers of 2D materials, 3D woven composites have fibres in the out-plane as well the in-plane direction, providing added strength and reducing manufacturing costs, but they can still be prone to failure at points where the weave pattern is disturbed. Research at Bristol is looking at how to

lower resolution. Both require resources on the scale that BlueCrystal offers, but El Said's work is particularly reliant on access to the additional memory afforded by Phase 3's compute nodes - 64GB per node, compared with 8GB per node on Phase 2.

Hallett's students are a good example of the BlueCrystal users who, with Phase 3's increased capacity, will see shorter queueing times than before. 'If the queuing time is longer than the run time, the whole process isn't much faster than running it on your desktop,' says Hallett. 'Phase 3 is a crucial step, because to continue to be successful, BlueCrystal needs to expand in proportion to increased demand.'

Complexity in the air: fluid dynamics on BlueCrystal



A forward flight simulation using 32 million grid points. The shading shows vorticity, which is a measure of rotation in the flow; blue is close to zero.

Simulating the flow over the wing of a plane is relatively straightforward: air moves around it in a fairly well-behaved way. A helicopter rotor blade is a different matter: it has to fly through air made vortical and turbulent by the blade ahead of it. Turbulent air, with its eddies, is much more difficult to simulate, owing to the small-scale effects present.

Professor Christian Allen is Head of the Fluid Dynamics and Aerodynamics research group, which has contracts with the likes of Augusta Westland, Airbus, BAe Systems, Garrad Hassan, and EPSRC. He uses BlueCrystal to tackle the challenge of simulating complex air flows. 'When simulating the flow around a helicopter rotor or a wing, we fill the domain around it with computational cells to make a numerical mesh,' he says, 'In each one of the cells we store the flow variables - density, velocityvector, pressure, and so on - and then we solve a set of conservation equations and update those to get the solution.'

computation increases by around 100 times. Unless, that is, you use a supercomputer with multiple processors or graphical processing units (GPUs).

'Because we use a mesh with a grid structure, we can partition that into smaller zones that can be solved simultaneously on multiple processors,' says Allen. 'These partitions need to communicate at the boundaries, but if you write your code efficiently, you can ensure that portion of the job runs efficiently on each processor, adding minimal overhead."

Computers like BlueCrystal are often hailed for their ability to complete in a few hours or days what used to take years, but Allen points out that 'nobody would actually run a iob that took a year in the first place. You'd run lots of smaller jobs and try to extrapolate them to what a year's job would give you. Now we can actually do those bigger computations and not have to extrapolate.

The difference between real-world flow and simulated flow lies in the number of grid points: the more grid points, the higher the resolution and the more data can be captured

The difference between real-world flow and simulated flow lies in the number of grid points in the simulation: the more grid points, the higher the resolution and the more data can be captured. But increase the grid cells by ten times, and the time needed to do the

Size of single computations is one thing one can generate a billion-cell mesh around a complete aircraft and split it between 10,000 cores - but BlueCrystal Phase 3 also provides the possibility for running a smaller simulation with many different variables simultaneously.



Modelling composite materials (from top): High-resolution model comparison with X-ray CT scans at different levels of compaction: further comparison of different fabric features with CT scans: comparison of feature scale model with CT scan: large-scale model showing local fabric

Computational fluid dynamics is sometimes called a 'numerical wind tunnel', and high performance computing is clothing that metaphor in ever greater layers of detail. Simulating air flow or fluid flow across moving objects of varying sizes - aeroplanes, helicopters, wings, cars, submarines - is a job that varies in complexity depending on the nature of the object.

When trying to optimise, for example, a wing shape to minimise drag, the traditional simulation method would involve making a series of minor changes and measuring each one's effects on the drag, each change requiring the simulation to be run again. But BlueCrystal enables all these changes to be run in parallel. 'If you have 100 different design variables, you can send each one to 100 different processors and get all the answers back at the same time,' says Allen.

Most supercomputers have numerous conventional processors – often thousands of them – but BlueCrystal also features many GPUs. These offer a step-change in computational performance, but they require specialist knowledge to exploit them properly. The Research Assistants attached to BlueCrystal are helping a growing list of users by converting code written for a less powerful computer cluster into something that will run smoothly on Phase 3. The result is data with a finer grain and a larger set of variations, acquired much faster. In Allen's work, this means a better numerical wind tunnel.

'Thirty years ago, an aircraft was designed 95% via simple theory, wind tunnel tests, and flight tests,' he says. 'Now 95% of it is done using simulation, and practical work is mostly kept in reserve in case a result from a simulation doesn't seem trustworthy. It's because of high performance clusters like BlueCrystal that this work can be done with greater and greater accuracy and efficiency.'

Working with industry

A new era in drug design

The field of drug discovery and design was revolutionised with the introduction of highthroughput screening, in which microplates with up to several hundred 'wells', each containing a different chemical compound, are processed using sophisticated robotics.

'Drugs companies use these robotic screenings to test hundreds of thousands of small-molecule drug candidates against a protein and measure what effect each one has on its activity,' says Dr Richard Sessions, Senior Research Fellow in the School of Biochemistry. 'They get a bunch of "hits", which are molecules that seem to have some of the desired effects, and they try to develop those through medicinal chemistry into a drug.' But each compound has to be bought or synthesised before it can be screened, which takes a good deal of time and money. Computational methods can, in principle, evaluate whole libraries of virtual compounds in simulation.

The number of possible molecules that companies have in their virtual compound databases runs to millions. Current computational virtual screening methods have serious limitations: the highthroughput methods lack accuracy, while the more accurate methods require too much computation for high throughput, making them unsuitable for screening large databases. But high performance computers like BlueCrystal, combined with the expertise of researchers like Sessions and his colleagues, now offer an alternative: accurate, high-throughput screening by molecular docking.

'With BlueCrystal we can actually screen several million compounds in simulation in

a couple of days, and identify the top few that are worth a closer look,' says Simon McIntosh-Smith, Senior Lecturer in High Performance Computing, who has been working with Sessions and Dr Amaurys Avila-Ibarra on developing software for this purpose. 'We might also find things that researchers may never have considered, and rule out things that might have looked promising before any physical screening time or money has been wasted on them.

The software in question, called Bristol University Docking Engine (BUDE), is part of Sessions' work to develop better ways of predicting whether a small molecule will be effective at binding to a protein in order to exert its pharmacological effects. BUDE's high computational demands required special techniques to be developed so that the software could make use of the latest

the intrinsic curvature that allows it to assemble into spheres. With Phase 3 we've been able to simulate larger pieces, and we've provoked enough interest in the scientific community to be able to get time on the new UK national supercomputer, Archer, to simulate a whole solvated nanocage, comprised of about 44 million atoms.

80 nm in diameter, comprising a self-assembled set

of three different designed peptides, each 25 amino

acids long (3.720 peptide molecules in total)

With BlueCrystal we can screen several million compounds in simulation in a couple of days

highly-parallel computer architectures. 'BUDE makes very effective use of Phase 3's HPC-oriented GPUs, which allow us to do many more calculations than in the past,' he says.

Sessions has been assisting researchers in other fields who are keen to exploit molecular modelling in their work. Professor Dek Woolfson and his group in the School of Chemistry design and synthesise small polypeptides that can be assembled into structures of varying size, usually via coiled coil interactions. Woolfson and his postdoctoral researcher, Dr Jordan Fletcher, have designed three peptides which, when assembled into a hexagonal network, form 'nanocages' of around 100 nanometres in diameter. These structures, it is thought, could act as delivery vehicles for transporting drugs and biologics into cells, presenting antigens for vaccine development, or compartmentalising enzymes into nanoscale chemical reactors.

'These nanocages are challenging to simulate,' says Sessions. 'We used BlueCrystal Phase 2 to simulate a small patch of this hexagonal network to examine

Meanwhile, in a good example of BlueCrystal's increasing role as a catalyst for partnerships between academia and industry, Sessions' work with McIntosh-Smith on molecular docking attracted the interest of a chemistry software company, Cresset Biomolecular Discovery, who saw the potential of the team's techniques for their own work. This led to the setting-up of a knowledge transfer partnership under the government's Technology Strategy Board. which funded an 18-month secondment for one of McIntosh-Smith's postgraduates, Simon Krige, at Cresset.

'Krige has been applying some of the methods that we developed, and implementing them in the molecular analysis software that Cresset supplies to the pharmaceutical industry,' says McIntosh-Smith. 'This means that parallel software techniques developed in Bristol on BlueCrystal are now helping industrial partners, such as Cresset, develop their own next-generation HPC products.

Facts and figures

BlueCrystal Phase 3

Phase 3 of BlueCrystal is now online and producing new, exciting science. The new Phase 3 system is five times faster in pure peak speed than the Phase 2 system, and as such represents a giant step up in processing power for the University. It is also a more heterogeneous design than its predecessors.

Phase 3 in brief

- Each of the 223 DellTM blades has two eight-core 2.6GHz Intel® Xeon® E5-2670 (Sandy Bridge) CPUs, 64GB of RAM, and a 1TB SATA disk.
- There are a further 18 similar blades configured with 256GB of RAM, giving 16GB RAM per single core. These large memory nodes can be reconfigured to make virtual SMPs (ie virtual systems that can address a single, shared main memory) of up to 4TB RAM using software from ScaleMPTM.
- Some 100 additional nodes also have, or have the capability to include, compute accelerator cards; to this end, Phase 3 currently has 78 NVIDIA® Kepler K20 GPGPUs installed.
- There are several networks installed to link these separate compute systems together. The fastest is the 40Gb/s Intel® True Scale QDR InfiniBandTM network. Providing the data storage element is a 340TB Panasas® 14S storage system, which comprises SSD and SATA devices.

In total, Phase 3 has close to 5,500 cores and a theoretical peak performance of some 200TFlops, which translates to 200 trillion floating point operations per second.





- The Operating System is Scientific Linux.
- Cluster management system is Bright Cluster Manager®.
- The system was supplied and installed by ClusterVision.

Advanced Computing Research Centre University of Bristol 8-10 Berkeley Square Bristol BS8 1HH

Tel: +44 (0)117 331 4375

bristol.ac.uk/acrc