University of BRISTOL

re:search

High Performance Computing issue

University of Bristol • Spring 2008

Drug design

Climate change

Helping child poverty

re: search editorial

High Performance Computing at the University of Bristol

The University has made a substantial commitment to High Performance Computing (HPC) by supporting the development of a new facility that will put the University at the forefront of computational research. BlueCrystal is one of the largest HPC facilities in the UK and will be a major resource both within the University and the wider community.

Over the coming years HPC will contribute significantly to University research income – last year £5.3 million was attributable to HPC – and it will also play an increasingly important role in teaching. Furthermore, we are particularly keen that BlueCrystal should become a University-wide resource used by researchers from all disciplines. Although it is no longer possible to undertake serious research in areas such as climate change, drug design and superconductivity without having access to HPC facilities, they are not the sole preserve of science and engineering and we are already seeing innovative HPC-based research being undertaken in the arts and social sciences.

We have selected the articles in this *re:*search supplement to provide an insight into some of the exciting HPC-based research being undertaken at the University, and to demonstrate the wide range of disciplines where HPC is now essential. I hope you enjoy them.

Professor Malcolm Anderson
Pro Vice-Chancellor for Research
Chair, HPC Board

The challenges ahead

The new HPC facility at the University of Bristol is a very exciting development and we are delighted that the University has chosen to work with IBM for their HPC infrastructure. We are already developing strong relationships with the team running the facility and with University researchers.

IBM recognises the vital importance of HPC to academic research, and there are exciting challenges ahead as we see a rapid increase in the number of processors and the computing power they provide. This transformation in the technology will also require changes to be made to the programming codes, with codes that can operate across many processors being the future. Our Deep Computing Institute is focusing on how to help researchers realise the potential of this increasing computational capability. In addition, efficient storage and retrieval of the data generated will be a key element in the process of solving critical problems for a better future.

The University of Bristol is leading the way by undertaking significant research into these new challenges. I am sure BlueCrystal will lead to extremely productive collaborations across many disciplines, and IBM wishes all users within the University success with their research over the coming years.

Dave Turek
Vice-President Deep Computing
IBM

Contents

Contacts

Caroline Gardiner
Academic Research Facilitator
Advanced Computing Research Centre
Merchant Venturers Building
Woodland Road
Bristol BS8 1UB
email caroline.gardiner@bristol.ac.uk
tel +44(0)117 331 5044
www.bristol.ac.uk/acrc

*re:*search is produced by the Public Relations Office, a department of Communications and Marketing Services.

Editor
Cherry Lewis
Research Communications Manager
Communications and Marketing Services
University of Bristol
Senate House
Tyndall Avenue
Bristol BS8 1TH
email re-search@bristol.ac.uk
tel +44 (0)117 928 8086

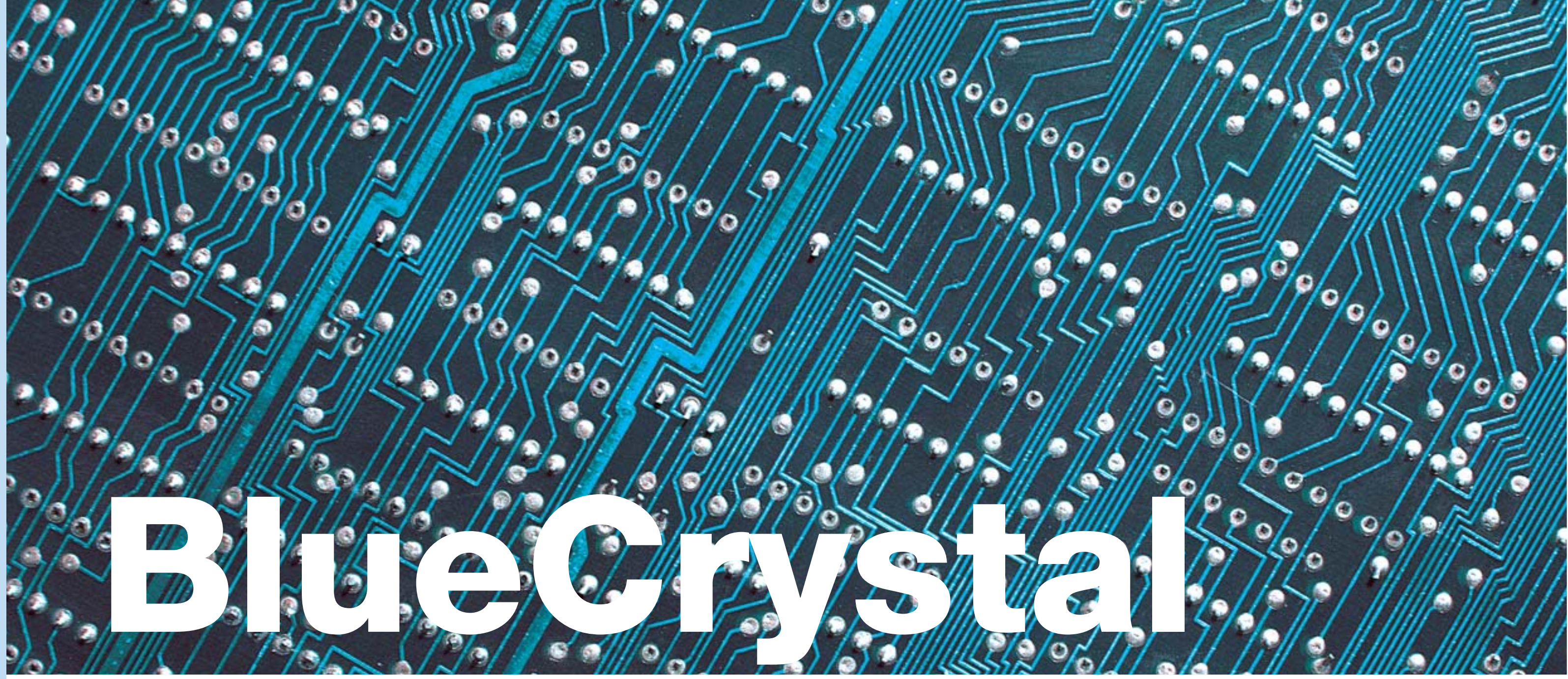
Head of Public Relations Office
Jill Cartwright

Communications and Marketing Director
Barry Taylor

Design, print and production
cw-design.co.uk

If you need part or all of this publication in an alternative format, please contact the editorial office.

*re:*search High Performance Computing issue, Spring 2008



BlueCrystal

turning data into knowledge

Professor Chris Johnson, Director of the Scientific Computing and Imaging Institute at the University of Iowa, recently spoke of a 'golden age' of scientific computing. He pointed out that most of the great discoveries in human history have been preceded by the creation of a new tool. Today, we have another new tool – supercomputers.

Scientific research increasingly focuses on computer simulation; perhaps as much as it does on theory and experimentation. But accurate simulations require huge amounts of computing power. High Performance Computing (HPC) operates on the principle that large problems can almost always be divided into smaller ones, which can then be carried out concurrently, ie 'in parallel'. Thus cluster 'supercomputers' are aggregates of hundreds, or even thousands, of computer processors, much like the one on your desk, that are connected together so that, effectively, they become a single computer with enormous processing and storage capabilities. HPC is therefore used for undertaking research that requires either very large amounts of data to be processed or lengthy computations to be carried out. While this can be expensive, it is clearly cheaper to simulate crashing a plane than to actually crash it.

With HPC underpinning an increasing number of disciplines across the University, it was recognised there was a need for a University-wide HPC facility. Substantial funding was provided via the Science Research Investment Fund and a state-of-the-art system, known as BlueCrystal, recently went live to users. At peak performance, BlueCrystal will be able to carry out more than 30 trillion calculations a second.

The HPC service at Bristol is delivered by the Advanced Computing Research Centre (ACRC). Part of its remit is to ensure that the HPC facility addresses the needs of the whole academic community and not just scientists. For example, ACRC staff are talking to researchers in the Centre for Medieval Studies about the digitisation of medieval manuscripts and to the Department of Drama about digitising the National Review of Live Art (NRLA) Archive, a video archive that holds footage of performances from the prestigious NRLA festival. The Department of Archaeology and Anthropology already runs a mini-cluster to analyse 3D images of archaeological artefacts and in the School for Policy Studies, HPC will be used to help eradicate poverty. But undoubtedly, major users of BlueCrystal will be scientists, mathematicians and engineers.

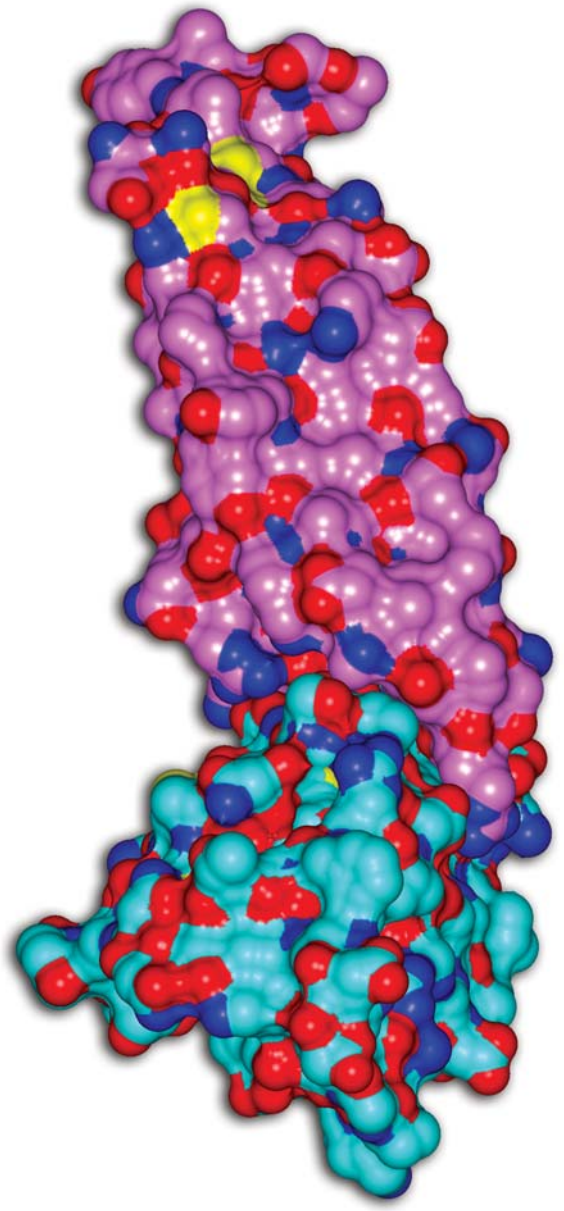
By 2010, virtually everything that has ever been written, composed, filmed, painted, or in any other way 'recorded' by humans, will be on line. All this material will amount to a staggering one 'exabyte' (10^{18} bytes) of data. In response to this ubiquity of data, the scientific method itself will need to change. The conventional model in which data are used to confirm or refute a hypothesis will be replaced by one in which data drive the generation of hypotheses. The mission of the Exabyte Informatics research group in the Department of Computer Science is, therefore, to develop methods and tools for this new data-driven science, and thus make a real difference to the way twenty-first century science moves forward.

It may well be that the 'golden age' predicted by Johnson is about to flourish in Bristol. ■

IAN STEWART

The new state-of-the-art super computer, BlueCrystal.

DESIGNING THE DRUGS OF TOMORROW



Drugs are being sought to prevent metastasis in breast cancer.

RICHARD SESSIONS, HAO WANG AND IAN PATERSON

Developing new breast cancer treatments

A pharmaceutical compound, or drug, usually exerts its effect by specifically binding either to a human protein to correct a biochemical defect, or to a key protein in a pathogen to kill it or reduce its virulence. Until recently, drugs were discovered either through serendipity or by screening hundreds of thousands of chemical compounds. The latter requires looking for appropriate activity using expensive robotic methods, and thus remains the province of large pharmaceutical companies.

"There are a number of fields where modelling studies are advancing biological science more rapidly than experiment alone can achieve." Professor Nigel Brown, Director of Science and Technology, Biotechnology and Biological Sciences Research Council.

With the advent of screening methods performed by computer simulation – called *in silico* screening – it is now possible to use computer algorithms to assess the likelihood of a compound binding appropriately to the target protein. By pre-screening compound libraries in this fashion, researchers are required to test far fewer compounds (hundreds, rather than hundreds of thousands). This greatly increases their chances of finding drug-like compounds. Researchers in the Departments of Biochemistry and Medicine are using this approach to search for anti-cancer drugs that will prevent secondary cancers developing from breast cancer.

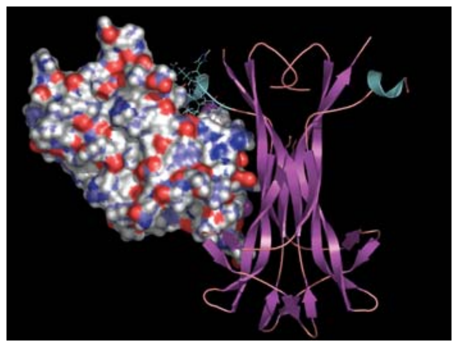
Current *in silico* screening techniques are a compromise between accurately predicting which compounds will be suitable, and high throughput. The team in Biochemistry is thus developing a more accurate method for predicting a compound's ability to bind to the surface of a protein in order to provide a better tool for *in silico* screening. At the same time, BlueCrystal is allowing researchers to use more computationally intensive and accurate commercial software packages to perform this *in silico* screening and hence focus more rapidly on the most promising compounds.

Seeking new treatments for Alzheimer's disease

Using similar techniques to those previously described, scientists in the Henry Wellcome Laboratories for Integrative Neuroscience and Endocrinology are seeking a novel treatment for Alzheimer's disease. Early in the disease, cells responsible for cognition and memory in the basal forebrain are lost. These cells rely on a continuous supply of neurotrophins – a family of proteins that act as chemical messengers between cells. These messengers dock with the appropriate receptors in the brain, thereby initiating a cascade of intracellular signals that ensures the continued survival of the memory cells. The task therefore, is to find a drug that can be taken orally which will activate this process and maintain cell viability.

BlueCrystal is providing the computational resources to exhaustively dock millions of compounds into relevant receptors. These are further screened for drug-like characteristics and the ability to permeate the blood-brain barrier. A combination of characteristics such as the compound's drug-like attributes, chemical diversity and ability to interact, are then used to select compounds for further testing.

DEBORAH SHOEMARK



NGF and its receptor TrkA – important in the prevention of Alzheimer's disease.

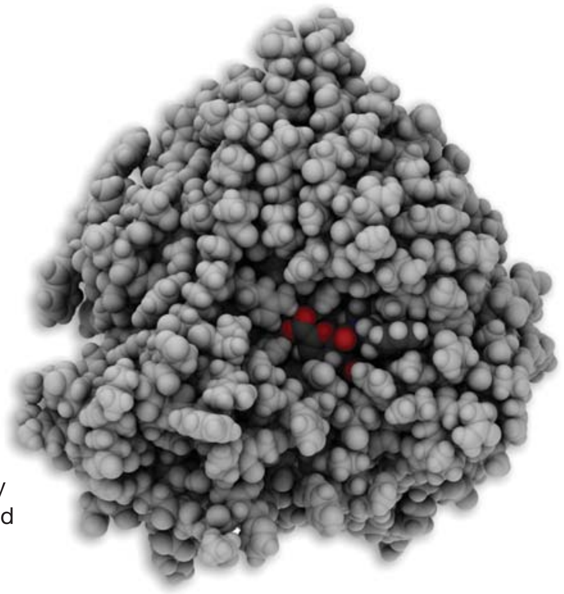
Accelerating drug design

The supply of new drugs seems to be drying up. For example, the US Food and Drug Administration approved 230 new drugs in 1997, but only 28 in 2006. The reasons for this are unclear – it may be that most of the 'easy' drugs have been found – but it is a major worry for both the pharmaceutical industry and for doctors and patients. Serendipity will no longer provide the new drugs we need, but molecular simulations using BlueCrystal may be coming to the rescue. In the School of Chemistry, chemists in the Centre for Computational Chemistry are developing new, more sophisticated computational methods to predict how tightly potential drugs bind to their biological targets.

To model drug molecules and their interactions in proteins, the group is developing methods based on quantum mechanics. These techniques are potentially much more accurate than current simplistic molecular models, because they give an accurate representation of the fundamental physics of molecular interactions. Such complex and computationally demanding calculations were previously out of reach for calculations on biological macromolecules. This situation has changed with the development of new computational techniques which have been adapted to take advantage of the power of new computer processors. ClearSpeed Technology, a Bristol-based semiconductor company, has developed powerful multiprocessor accelerator cards, which only consume a fraction of the electricity used by a supercomputer.

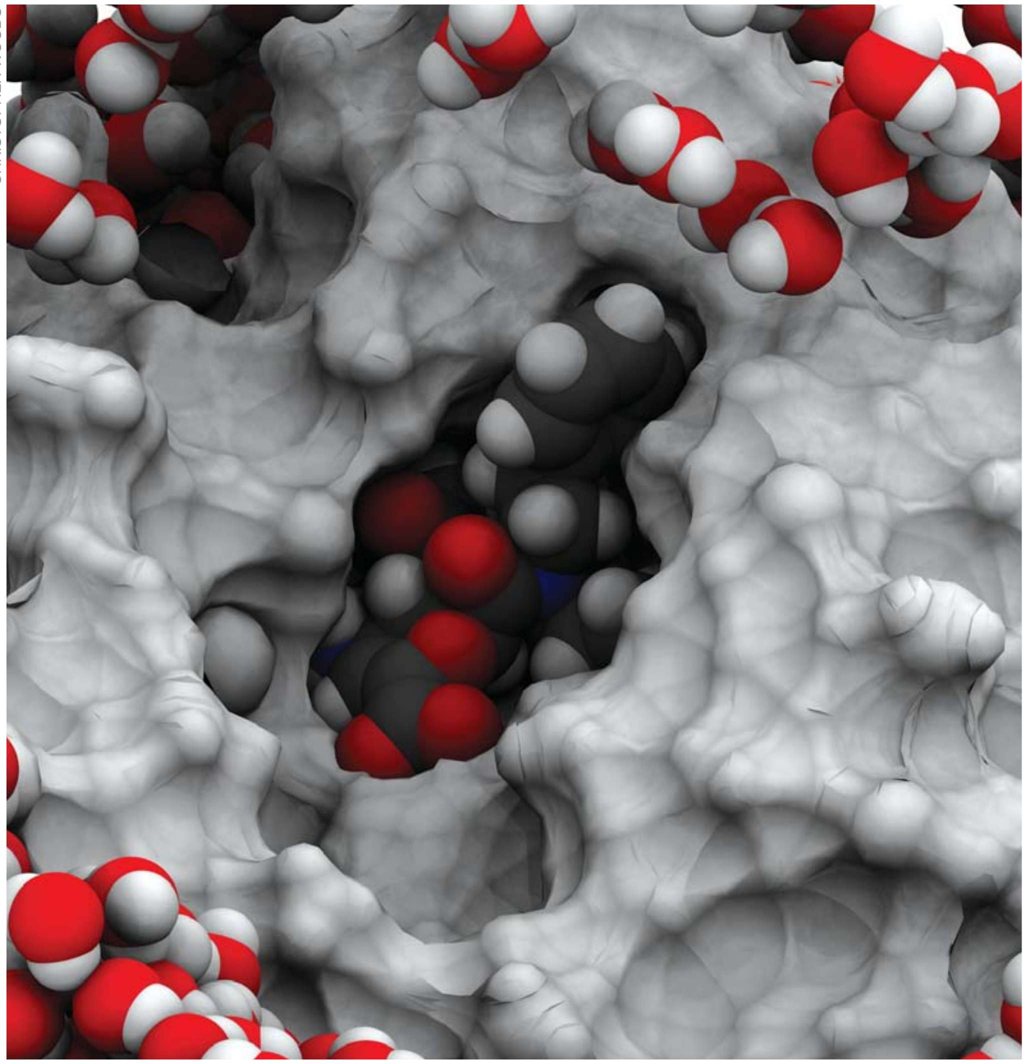
Working closely with ClearSpeed, the group has modelled a target for the development of drugs to combat flu. These calculations use high-level quantum mechanical methods, now made possible because they efficiently use tens of thousands of processors.

The team modelled drugs bound to neuraminidase, an enzyme found on the surface of the influenza virus. Anti-flu drugs work by binding to, and blocking the action of, this enzyme. The calculations predict how strongly potential new drugs would bind to the enzyme, and so identify those most likely to be successful. Results can be obtained in a matter of days, promising to provide reliable input for drug design on a timescale fast enough to be of use to the pharmaceutical industry. This represents a step-change in the speed and accuracy of computational modelling. ■



Above: The neuraminidase enzyme showing a potential drug (red) bound at the active site. Below: Detail.

CHRISTOPHER WOODS



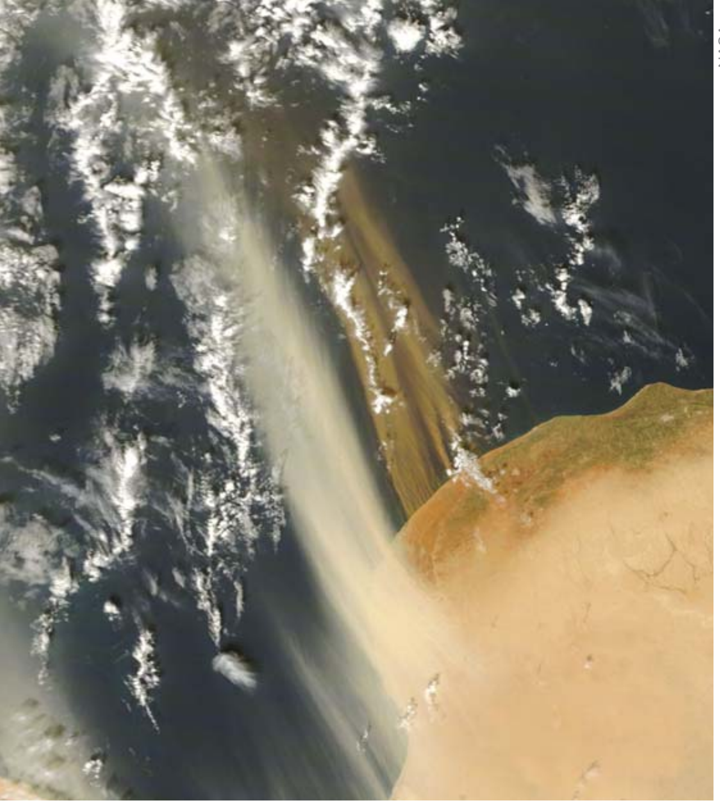
Contributors

Dr Richard Sessions
Department of Biochemistry
www.bristol.ac.uk/biochemistry

Dr Fred Manby, Dr Adrian Mulholland and Dr Christopher Woods
School of Chemistry
www.chm.bristol.ac.uk

Drs David Dawbarn, Shelley Allen and Deborah Shoemark, Laboratories for Integrative Neuroscience and Endocrinology
www.bristol.ac.uk/clinicalsciencesouth/hwline

Climate change



NASA

A dust storm seen from space.

Climate change is rarely out of the headlines. Every day we are bombarded with new results on the latest observations of climate change, and politicians, business leaders and the public are demanding accurate forecasts about what will happen in the future. These predictions of climate change are mainly based on complex computer models and require huge amounts of computer power. The models are based on the best understanding of the physical, chemical and biological controls on climate. They are developed in the context of modern climate, then tested against past changes and applied to predict future climate change. The University is heavily involved in all aspects of this work.

Modelling the past to predict the future

Researchers in the School of Geographical Sciences are developing and improving the latest state-of-the-art computer models used to predict climate change, specialising particularly in the controls and interactions between vegetation, methane, atmospheric chemistry, and climate. They collect and synthesise data on past climate change, in some cases going back millions of years, and then use this information to evaluate whether climate models are able to represent past changes. If the model works well for the past, it gives confidence in the model's ability to predict future change. In addition, modellers are involved in assessing the effectiveness of other 'geo-engineering' solutions for climate change, such as the launch of a giant mirror to reflect sunlight and cool the planet, planned by NASA.

"Prediction is very difficult, especially if it's about the future." Nils Bohr, Nobel laureate in Physics

A further important area is the very nature of the predictions. Climate projections will always have a degree of uncertainty associated with them and researchers are increasingly moving away from a simple 'best estimate' prediction and towards a risk-based assessment where they calculate the probability of climate change. For example, a model simulation may predict that there is an 80 per cent probability that warming will exceed 3°C by the end of the century. Although perhaps more complicated to understand, it is a better statement of the basic science of climate prediction.

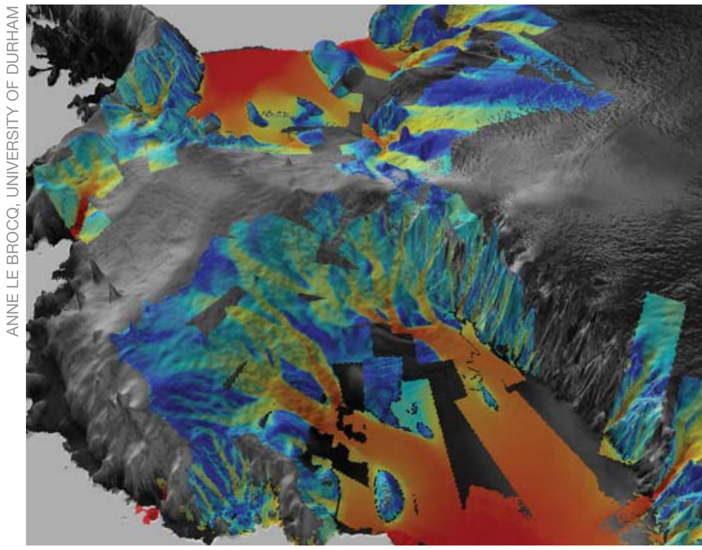
All of this work requires phenomenal amounts of computer power and disk space. A typical simulation may take three months to run and generates an enormous 10 terabytes of model output. BlueCrystal will revolutionise how researchers tackle the challenge of climate change, greatly improving confidence in future predictions on spatial and temporal scales relevant to society.

Challenges in ice sheet modelling

The most recent assessment of the Intergovernmental Panel on Climate Change (IPCC) highlights concerns about the predictive ability of the current generation of ice sheet models. These concerns have been prompted by a revolution over the past decade in our ability to monitor ice sheets, primarily by satellite-borne sensors which have made possible accurate, spatially extensive measurements of both ice thinning rates and horizontal velocity. Models of the type used in the IPCC assessments do not contain the physics necessary to simulate many of the phenomena now being observed. Examples include the local acceleration of virtually all outlet glaciers in Greenland, possibly as a consequence of enhanced lubrication by surface melt waters; the collapse of Antarctic ice shelves by pervasive hydro-fracture; and the rapid response to several large Antarctic ice streams to changes in their fringing ice shelves.

JONATHAN BAMBER

The edge of the floating Filchner-Ronne ice sheet, West Antarctica.



ANNE LE BROCQ, UNIVERSITY OF DURHAM

A perspective view of the Antarctica ice sheets. Colours refer to horizontal ice velocity (blue ~10 m/yr to red ~1 km/yr, with yellow ~100 m/yr). The detailed tributary pattern of the numerous ice streams draining West Antarctica is visible in the foreground.

Model development focuses on three main areas. The first is prompted by the realisation that oceanic change can be at least as important as atmospheric change for the ice sheets, particularly in Antarctica. Traditional coupling between ice sheets and the rest of the climate system is across the wide expanse of the upper surface of an ice mass and involves meteorological variables, such as air temperature and precipitation. New coupling strategies are required that focus on the heat exchanges between ocean and ice in the cavities under floating ice shelves. The second area on which model development concentrates is improving the physical representation of ice flow in these models. This move to improved physics will result in a hundred-fold increase in computer processor requirements. Finally, the gross dynamics of ice sheets appear to be controlled by processes in highly localised areas, such as the grounding line separating floating ice from that resting on bedrock, and fast-flowing ice streams. A strategy involving nested grids that resolve these 'hot spots' may therefore be necessary.

The Bristol Glaciology Group is working with the Met Office's Hadley Centre for Climate Prediction and Research to address many of these issues with the aim of developing a new generation of ice sheet models, using the computing power of BlueCrystal. The results will then be available for inclusion in the next assessment report of the IPCC.

Flooding – where next?

Accurate prediction of flood inundation over complex terrain is a global problem. In the UK alone, two million properties are located in areas at risk from flooding, and the UK spends approximately £1 billion annually on flood defence. Summer 2007 saw extensive flooding over much of the UK with total damage estimated by the Association for British Insurers at £3 billion. The ability to predict areas at risk from flooding is thus critical to the effective management of flood risk, and the last five years have seen significant advances in our ability to model inundation on rural floodplains. This is due to the availability of high resolution data from satellites and aircraft, faster computer processors and the development of efficient, yet accurate, computer models.

The University has been at the heart of these developments, particularly through its involvement in the Flood Risk Management Research Consortium. The Engineering and Physical Sciences Research Council has made available over £14 million for research in this area. As a result, researchers have also begun to develop the capacity to model flows through urban areas, as this is where the majority of at-risk assets are located. This, however, requires very high resolution models because of the complex topology and topography involved. This is problematic as we are currently only able to simulate urban areas in detail of the order of 1 km^2. The grand challenge in flood inundation modelling is therefore to develop efficient systems capable of taking advantage of recent developments in multi-processor computer architectures, as exemplified by machines such as BlueCrystal, to allow detailed, whole-city analysis of flood risk. This research is now underway in the Hydrology group and will provide a major contribution to science and policy in the future. ■

CHERRY LEWIS

Contributors
Professors Paul Bates, Tony Payne and Paul Valdes
School of Geographical Sciences, www.ggy.bristol.ac.uk

© UNICEF/HQ07-0108/THIERRY DELVIGNE JEAN

Helping solve child poverty

The Millennium Development Goals are an ambitious set of eight goals, agreed by world leaders, that include cutting poverty in half by 2015. Researchers in the School for Policy Studies (SPS) plan to contribute to that objective by using BlueCrystal for the analyses of large survey and census data in three major projects.

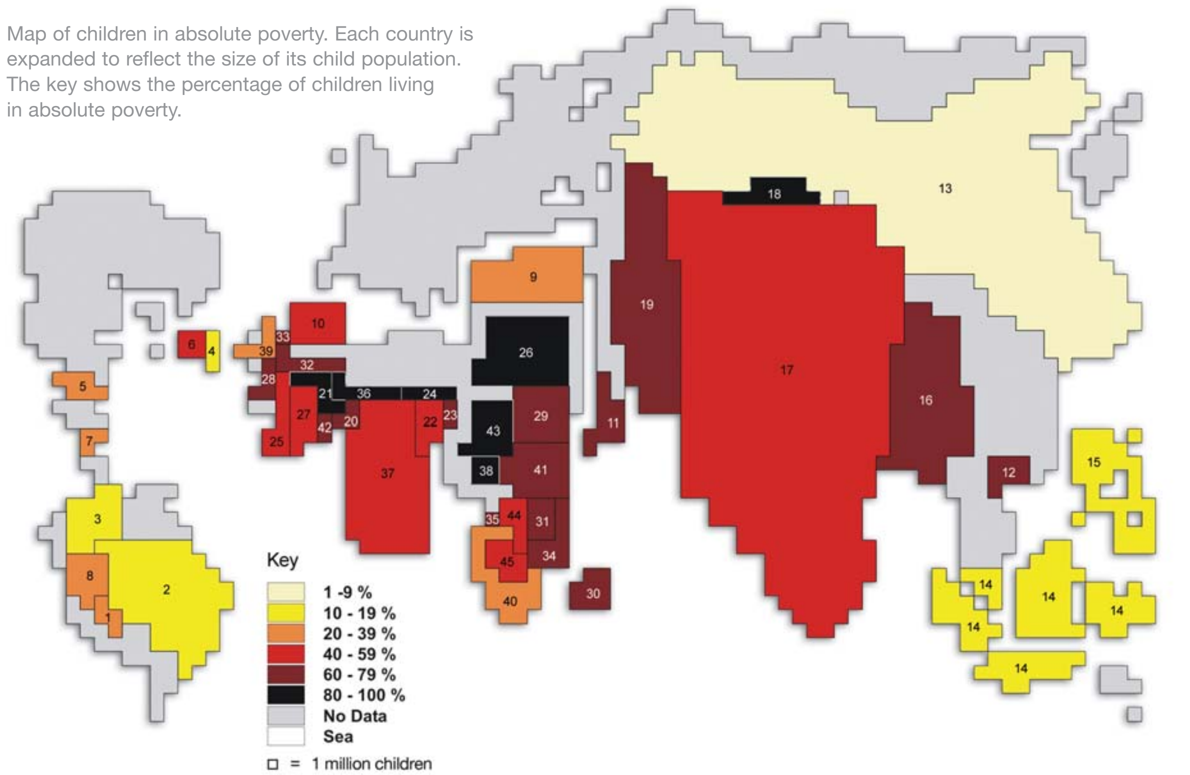
Working with colleagues in UNICEF in 40 countries, researchers in SPS aim to produce the first global study of child poverty and disparities. The primary purpose of this work is to improve national, regional and global awareness of the extent and nature of child poverty in order to produce effective and efficient public policies to rapidly eradicate absolute child poverty during the 21 century.

The lack of safe drinking water is a major problem in developing countries. The World Health Organization has estimated that 1.8 million people around the world die each year from water-borne diseases, of which 1.5 million are children under five. 'Aquatest', an easy-to-use, low-cost device for testing the quality of water in developing countries, has received $13.1 million in funding from the Gates Foundation. BlueCrystal will be used to help identify the optimal sites for the first set of field trials.

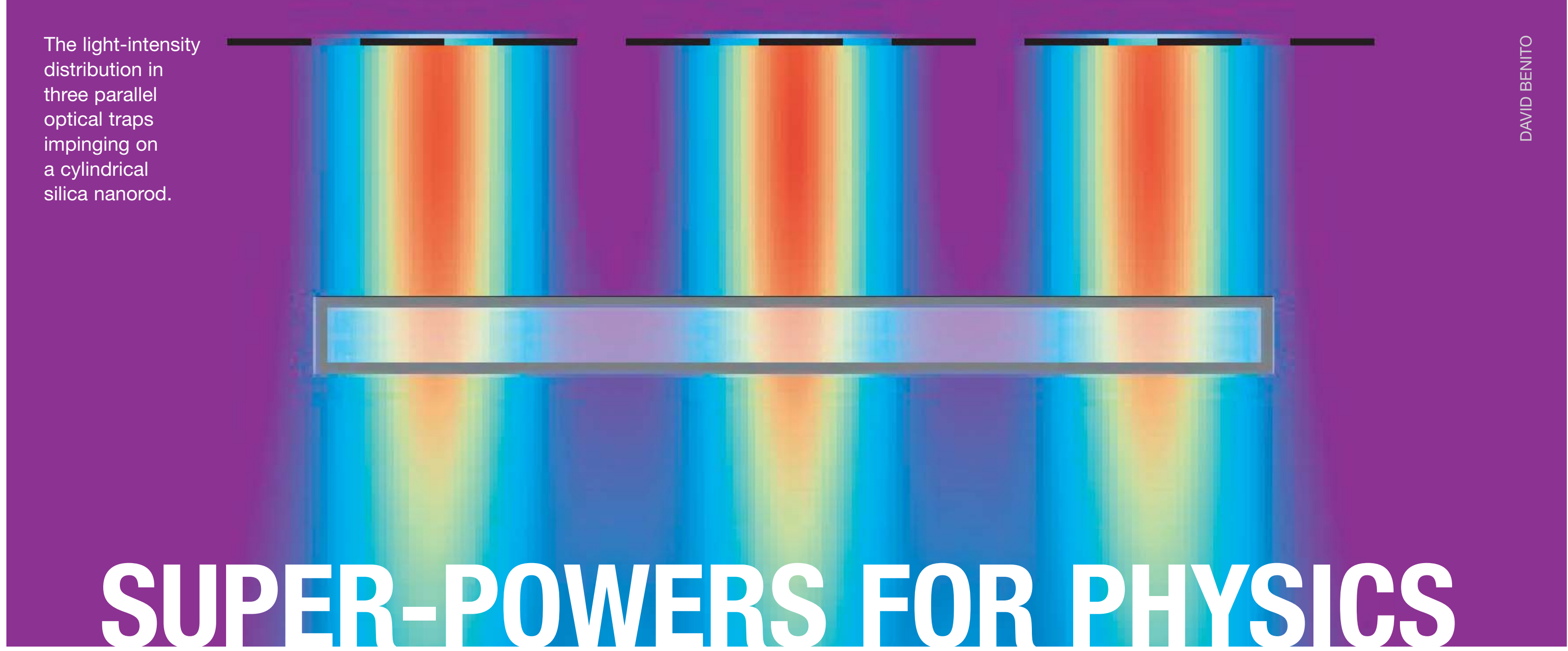
Finally, a collaborative research project is currently being developed with colleagues at the University of Warwick to examine the global situation of disabled children. It aims to undertake the first representative analyses of the living conditions of disabled children using national census data from countries in both the industrialised and developing worlds. Little is currently known about the circumstances of disabled children in many countries and how effective national policies have been at helping disabled children overcome prejudice and disadvantage. This study will help to identify countries which have effective policies and practices to support disabled children and their families. ■

Contributor
Professor David Gordon, School for Policy Studies, www.bristol.ac.uk/sps

Map of children in absolute poverty. Each country is expanded to reflect the size of its child population. The key shows the percentage of children living in absolute poverty.



1. Bolivia
2. Brazil
3. Colombia
4. Dominican Republic
5. Guatemala
6. Haiti
7. Nicaragua
8. Peru
9. Egypt
10. Morocco
11. Yemen
12. Cambodia
13. China
14. Indonesia
15. Philippines
16. Bangladesh
17. India
18. Nepal
19. Pakistan
20. Benin
21. Burkina Faso
22. Cameroon
23. Central African Republic
24. Chad
25. Côte d'Ivoire
26. Ethiopia
27. Ghana
28. Guinea
29. Kenya
30. Madagascar
31. Malawi
32. Mali
33. Mauritania
34. Mozambique
35. Namibia
36. Niger
37. Nigeria
38. Rwanda
39. Senegal
40. South Africa
41. Tanzania
42. Togo
43. Uganda
44. Zambia
45. Zimbabwe



The light-intensity distribution in three parallel optical traps impinging on a cylindrical silica nanorod.

DAVID BENITO

SUPER-POWERS FOR PHYSICS

BlueCrystal is being used in the Department of Physics to assist teams who are at the cutting edge of their fields, such as those working on condensed matter physics (CMP). CMP is by far the largest field of contemporary physics, encompassing all aspects of the study of solid materials and liquids. The applications of CMP research include everything from the invention of silicon transistors and magnetic materials used in computer data storage, to the modern liquid crystal displays in laptop computers and flat screen televisions.

Superconductors are a field of particular interest, since they are compounds that conduct electricity without resistance (and thereby without losing energy) at extremely low temperatures. One area in which superconductors are used is Magnetic Resonance Imaging, which uses the magnetic field derived from superconductors as a non-invasive means of determining what is going on inside the human body.

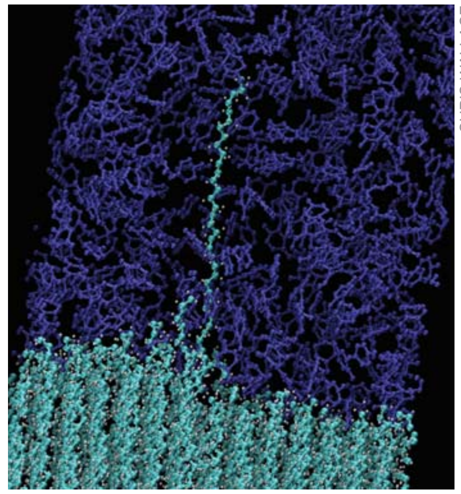
Modern 'high-temperature' superconductors (which are still extremely cold) are based on very complex crystals, and understanding why they become superconducting remains one of the most controversial topics in condensed matter physics today. BlueCrystal is facilitating the study of such materials, as well as the study of the origin and nature of superconductivity. Potentially this could lead to the discovery of new superconducting materials operating at even higher temperatures, possibly even at room temperature.

Of growing interest in CMP is the study of soft matter physics where researchers are using BlueCrystal to model the structure and dynamics of liquid crystals, polymers and other complex fluids. Molecular dynamics simulations are used to understand the relaxation behaviour of long polymer molecules in flow fields, as well as to model atomic force microscope experiments in which single polymer molecules are manipulated. Simulations are used to model the self-assembly of highly complex molecules, which may one day find applications facilitating tissue replacement. They could be used to build biological scaffolds on which new tissue can be grown – replacement cartilage for damaged knees, for example.

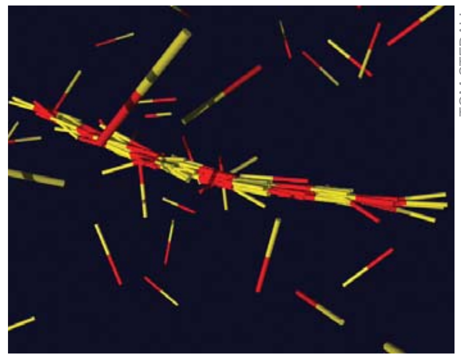
Manipulating light with tweezers

Holographic optical tweezers utilise the power of light beams to manipulate micron-sized particles and biological structures. Multiple tightly focused laser beams allow the assembly of nano-hands and nano-tools for constructing 'photonic' structures that manipulate light, as well as providing a way of probing individual cells. Because the optical fields involved in such systems are so complex, BlueCrystal is used to compute optical forces and torques and to predict novel optical effects. The power of supercomputers is essential in this work as calculations are required at a billion separate points in the sample in order to achieve the require accuracy. Similar techniques are also being used to model the optics of liquid crystalline devices. ■

Contributors
Professor James Annett and Dr Simon Hanna, Department of Physics
www.phy.bristol.ac.uk



CHRIS WALLACE



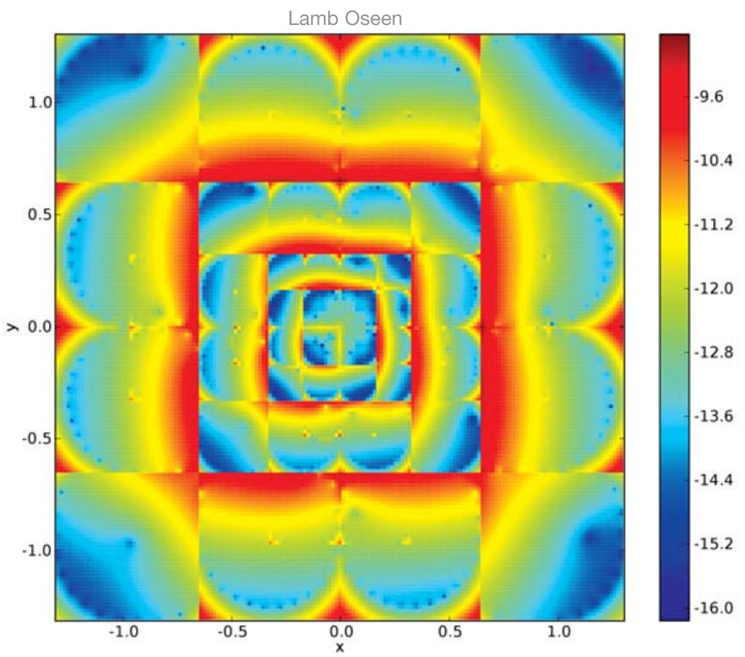
TOM STEDALL

From top:
Molecular dynamics simulation of a single molecule of polyethylene being drawn from a crystal in an atomic force microscope.

Coarse-grained Monte Carlo simulation of the self-assembly of an oligopeptide fibril. Different units are colour coded yellow and red.

Crunching **BIG** numbers

Fluid dynamics is the study of fluids in motion. It includes, for example, examining the way air moves as a plane flies through it, or how cream behaves when poured into a cup of coffee. Many problems involving fluid flows are analysed using numerical algorithms and this branch of the discipline is known as Computational Fluid Dynamics (CFD). Millions of calculations may be required to solve a single problem, which today means running the calculations on a supercomputer such as BlueCrystal.



The coloured dots of the image represent a measure of the error when using an algorithm which is fast, but approximate. Looking at the structure of the errors helps us understand the algorithm and improve its application.

One group currently researching CFD is in the Department of Mathematics where new 'meshless' methods of computational modelling are being investigated. Another is in the Department of Aerospace Engineering where researchers are examining unsteady flows and their interaction with engineering structures, particularly helicopter rotors.

New algorithms, new mathematics

Although many great advances in computer technology were made in the '60s and '70s, it has taken the past 20 years for the tool to develop in such a way that every scientist today can use the computer as her lab. Scientists in every field have been realising that computing is changing the way we do science, but now new technology is changing the way we do computing. New hardware is all well and good, but the real changes will be due to new algorithms – and new mathematics.

For example, at the moment, in order to calculate how the concentration of a chemical changes in a fluid, one writes down the mathematics of what comes in, what goes out, and what was already there, to obtain the difference. The process incurs a small error, which has an 'averaging' effect on the quantity. This is called numerical diffusion. To tackle numerical diffusion errors, the only solution is to construct extremely small computational elements, which results in a very fine 'mesh' and requires large computational time.

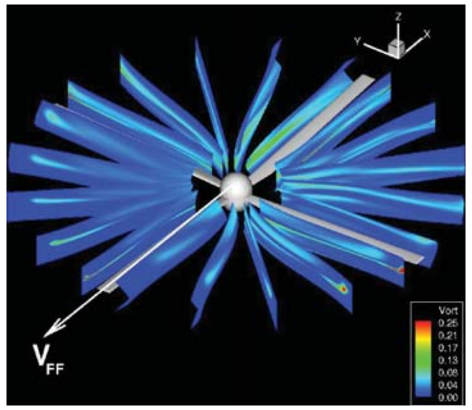
A new wave in computational mechanics is the development of mesh-free methods. They promise to give an answer not only to numerical diffusion, but to a host of other problems, such as dealing with cracks and discontinuities. The challenge lies in developing efficient algorithms to deal with computational 'particles', rather than elements of a mesh. Very complex algorithms are needed, such as the 'fast multipole method', classed as one of the top ten algorithms of the 20th century. And, of course, these algorithms need to be run on supercomputers, such as BlueCrystal. The bonus? These algorithms may be especially well suited to exploit the new hardware technologies which are revolutionising today's science.

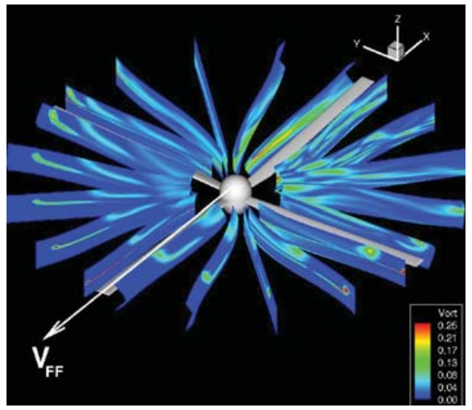
> "What is possible today was only dreamed about just a few years ago."
> Chris Johnson, Director, Scientific Computing and Imaging Institute, University of Iowa

Contributors

Dr Lorena Barba, Department of Mathematics, www.maths.bristol.ac.uk

Professor Christian Allen, Department of Aerospace Engineering www.aer.bristol.ac.uk





Vorticity shading for the four-million-cell grid (top) and the 32-million-cell grid (bottom), which clearly shows greater detail. (Vorticity is a measure of rotation in the flow.)

Unsteady simulations

The aerodynamics and dynamics of helicopter blades are among the most complex in the aerospace field. The aerodynamic loading on a rotor blade is highly dependent on the wake shed by the previous blade. When hovering, a complex helical wake forms below the blades, while in forward flight the wake is swept downstream. Thus each blade experiences large variations in loads as the tip vortices from previous blades vary in both strength and position. The noise and vibration levels and frequencies are also determined by the wake characteristics; for example, the 'chopping' sound made by helicopters is caused by one blade hitting the wake of the previous one.

Predicting time-dependent helicopter loads and vibrations is absolutely critical for rotor blade design and analysis but, as yet, numerical simulation is not used as a primary design tool, due to excessive cost. Forward flight, for example, requires very fine meshes, and many very small time steps, to accurately resolve the wake motion, and so parallelisation is absolutely essential to make the time required to solve these problems realistic. Parallelisation involves splitting the problem, and solving the connected smaller problems simultaneously on multiple processors. A serial code uses a single processor and obviously requires much more time to solve the same problem.

A highly efficient parallel CFD code has been developed in Aerospace Engineering. The code has been tested on 'HPCx', the UK's national HPC facility, and awarded a 'Gold Star' for its performance. It has allowed both hover and forward flight simulations to a resolution not previously possible.

A four-bladed, fully unsteady, forward flight case was run using two grid sizes. A four-million-cell grid was run in serial (this is the limit of the serial code) and a 32-million-cell grid run in parallel, using 256 computer processors. The 32-million-cell task took six days to complete – to run this same simulation using the serial code, if that were possible, would require more than four years. ■

AIMING HIGH – MANAGING THE DATA

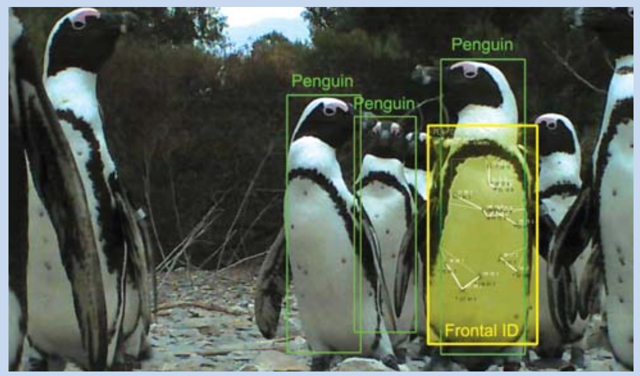
The Exabyte Informatics research group in the Department of Computer Science deals with the scientific challenges and opportunities arising from the enormous amount of data that BlueCrystal will handle. These challenges can be summarised in three letters: AIM – Annotation, Integration, Mining.

Annotation is concerned with capturing the semantics, or meaning, of the raw data in the form of so-called metadata. It is one of the most challenging problems in this area of computer science. For example, in order to illuminate the behaviour of penguins filmed on video, researchers in the department have developed software for tracking individuals from patterns in their feathers. Similar to the way loyalty cards allow supermarkets to track your purchases, thereby discovering patterns in your buying behaviour, the semantically enhanced data produced by the software allows biologists to discover behavioural patterns that were hidden in the raw video data.

Integration involves combining data sources to exploit the implicit relationships between data. For instance, a researcher looking into the functions of genes may want to enhance the data obtained from DNA micro-arrays – a tool for analysing thousands of genes at a time – with a search of the genomics literature, in order to add further knowledge about the genes being profiled, and their relationships. Such analyses often lead to a new representation of the data in the form of 'networks of interaction', to which a range of network analysis techniques can be applied so that the most significant relationships can be exploited.

Data mining seeks to find meaningful patterns in the data obtained after annotation and integration. These patterns can range from finding 'outliers' that behave very differently from the rest of the population, to global models of how the data can be explained. A particular focus of Exabyte Informatics is on highly structured data such as large molecules whose three-dimensional structure is important and needs to be properly represented.

The massive proliferation of data produces yet another challenge: how will all this information be managed? Part of the new HPC facility is a planned £2 million data storage facility that will ensure data can be safely stored, curated, enhanced and retrieved when and in whatever form needed. Eventually, researchers across the University – and potentially anywhere in the world – will be able to retrieve data from BlueCrystal. ■



TILO BURGHARDT

Penguins displaying their biometric patterns.

Contributor

Professor Peter Flach, Department of Computer Science, www.cs.bristol.ac.uk *re:* search

Public Relations Office
Communications and Marketing Services
University of Bristol
Senate House
Tyndall Avenue
Bristol BS8 1TH

www.bristol.ac.uk
Email: re-search@bristol.ac.uk

L E A R N I N G • D I S C O V E R Y • E N T E R P R I S E