Advanced Research Computing at Cardiff

Annual Report 2007/08

Does your research have the potential to use these techniques?

Curious to find out more?

Then please visit our web site to find out more www.cardiff.ac.uk/arcca or get in touch.

We'd be delighted to discuss what Advanced Research Computing means in practice, and how ARCCA may be able to help you to do your computer-processing based research more effectively.

email: arcca@cardiff.ac.uk

telephone: +44 (0)29 208 79319 (Cardiff University extension 79319)

visit: ARCCA, Redwood Building (Tower), King Edward VII Avenue, Cardiff, Wales, CF10 3NB, UK

www.cardiff.ac.uk/arcca

If you have any queries or comments with regards to the ARCCA Annual Report please contact insrvCommunications@cardiff.ac.uk

The Annual Report is written and produced by ARCCA with Stella Faulkner and Felicity Childs from the insrvCommunications team, and is designed by the Media Resources Centre.

arcca

ARCCA

Contents

What is Advanced Research Computing (ARC)?

Advanced Research Computing is the use of computing and data resources for research, which are beyond the capabilities of the average desktop or laptop computer.

This technique uses leading-edge IT resources and tools to pursue research including computer simulation and modelling, manipulating and storing large amounts of data, and many other methods to solve research problems that would otherwise be impossible.

Front cover shows

Image 1: A simulation showing the formation of a binary star system (Courtesy of Dr Kitsionas & Prof. Whitworth).
Image 4: A simulation of the convection of the Earth's mantle. The yellow plumes are practically steady (Courtesy of Dr Huw Davies).
Image 8: A simulation of stress on a dental implant, the colours showing the areas of high stress in the surrounding bone. (Courtesy of Prof. Middleton).
Image 9: The sensitivity of interferometric gravitational wave detectors depends on the direction of the source (Courtesy of Prof. Sathyaprakash).
Images 2,3,5,6,7: The Redwood Machine Room and ARCCA High-end Cluster.

Preface from Teresa Rees, pro-VC Research

With computational science now widely recognised as the third fundamental pillar of science, alongside theory and experiment, it is timely to consider the potential impact of computing on the research agenda of the University, and the role that we hope ARCCA – Advanced Research Computing at Cardiff – will play in supporting and enabling researchers to use leading-edge IT resources and tools to pursue ground-breaking research.

Computing has already transformed many scientific disciplines, but major advances in computing technology that are already taking shape promise even more profound impacts on research in the next decade. Communication bandwidths of tens of Gigabytes per second will transform the accessibility of information. Masses of multi-dimensional scientific and medical data will be stored and manipulated routinely given affordable Petabyte-scale data storage systems. And, most dramatically, multi-core processors will allow the construction of computers containing hundreds of thousands of processing cores, all working on different parts of the same task.

Scientific progress will be linked so closely to computing that computing will increasingly be used to supplement experiment when dealing with complex phenomena, phenomena where the basic scientific laws are understood and where experiment is difficult or impossible. Already, progress in areas such as climate research, astrophysics, cosmology, protein structure determination, and discovering the chemical origins of life (to name just a few) would be impossible without reliable computational modelling. In the next decade many more areas, such as human biology and the development of new drugs, can be expected to join this list. Just as importantly, computing promises to play an increasingly key role in many new areas of research endeavour – the humanities and social sciences are clear examples – with ARCCA established to provide the support necessary for researchers to exploit this potential.

With the investments made possible through the Science Research Investment Fund (SRIF) leading to the imminent launch of Cardiff's High-end Computing Facility, it is a pleasure for me to introduce this, the first Annual Report of Advanced Research Computing at Cardiff.

Welcome from the Directors

I am pleased to present the first Annual Report of Advanced Research Computing at Cardiff – ARCCA – the new ground-breaking, academically-focused initiative which is dedicated entirely to supporting and enabling research.

ARCCA's prime purpose is to enable researchers to harness the power of leading-edge high performance computing technology to pursue innovative research within and across all academic disciplines. Researchers will now be able to apply these tools to solve problems and undertake original work that was previously impossible.

The genesis of ARCCA marks a significant step in evolving the research infrastructure of the University. By transforming the contribution of advanced computing to the University's research agenda, ARCCA will build and sustain a position which takes the University to the forefront of leading research universities in the UK and internationally in this field.

As you will see from the report, ARCCA has taken a number of years to come to fruition. Despite the existing pockets of excellence across the University in this area, it was confirmed by a consultant's review that it was imperative that the University took a more strategic and coordinated approach to high end computing.

A vital step in realizing the University's ambition was to appoint a Director of ARCCA with the expertise and profile to lead the University's high performance computing agenda. This was more than achieved with the appointment, in April 2007, of Professor Martyn Guest, a world-renowned and acknowledged expert in the field of high end computing and its application to research. In bringing his considerable expertise, experience and global standing to the University, Martyn adds lustre to the University's profile and gives an incomparable edge to the University's research capability.

Finally, I would like personally to acknowledge the vital role played by the personal sponsorship of the Vice-Chancellor, Dr David Grant, who recognised the importance of high performance computing to the University, and the support of Professors Hadyn Ellis, Peter Blood and Richard Whipp, in bringing ARCCA into being.

So, enough from me; let me hand over to the Director of ARCCA, Professor Martyn Guest, to tell you more…

Martyn C. Harrow
Director of Information Services

From a personal perspective, delivering on the vision of those who led the University to invest in Advanced Research Computing is a challenge that I greatly relish, and one that ARCCA is committed to delivering in full. My specific ambition is to lead ARCCA to a position of international standing as an organisation in its own right.

In line with its core mission of research enablement and promoting research outcomes, ARCCA is committed to the provision of enhanced access to the associated facilities, enabling researchers from all academic disciplines to harness their potential.

No small task, but one that we will strive to deliver.

I have great pleasure therefore, in presenting to you this, our first Annual Report. We introduce a selection of the highlights and developments of the past year and outline our strategy for providing support to the University across all levels that characterise the research computing landscape. The report is structured to run through the stages of the development of ARCCA: from its genesis, to how the pioneering and influential work around Condor[1] demonstrated the value of High End Computing (HEC) to the University community, through the SRIF cluster procurement process, to where we are today. Finally we look ahead to the next year of ARCCA and what we believe will be a step change in delivering capability to researchers across the University.

The techniques outlined in this report are already enabling and enhancing research in more than half of the Schools across the University. ARCCA will also work with clients and partners outside the University through a range of training and outreach activities.

If your research has the potential to use these techniques, or even if you are just curious to find out more, please visit our web site (www.cardiff.ac.uk/arcca) or get in touch. We'd be delighted to discuss what Advanced Research Computing means in practice, and how ARCCA may be able to help you to do your computational based research more effectively.

Finally, mention must be given to the crucial role of our colleagues within Information Services in contributing to the work of ARCCA. Indeed, Information Services works in close partnership with ARCCA to provide joined-up support across a huge range of research, teaching and administrative needs.

Professor Martyn F. Guest
Director of Advanced Research Computing at Cardiff

Please note that the terms 'Advanced Research Computing at Cardiff' and 'ARCCA', and 'Information Services' and 'INSRV' are used interchangeably within this report.

[1]The workload management system for compute-intensive jobs

Meet the ARCCA Team

ARCCA has a small but dedicated team in place to help and support the University's research needs. This team will deliver Advanced Research Computing solutions and services to enhance the way research is conducted at Cardiff and beyond.

Below is a brief introduction to the team, which is led by internationally-recognised expert Professor Martyn Guest.

Professor Martyn F. Guest

Professor Martyn F. Guest – previously Associate Director of the Computational Science and Engineering (CSE) Department at the Science and Technology Facilities Council's Daresbury Laboratory – was appointed in April 2007 to lead the ARCCA division. With a strong background in computational chemistry, he is the lead author of the GAMESS-UK electronic structure code, with over 230 articles to his name in the areas of theoretical and computational chemistry, and High Performance Computing.

Internationally recognised as an expert in the field of High Performance Computing (HPC) and application performance, Martyn has been involved in helping to establish a number of HPC centres both nationally (HPCx, HECToR) and internationally (ICHEC, Ireland and CHPC, South Africa).

Professor Guest is an Honorary Professor in Physics at the University of Edinburgh and an Honorary Reader in Chemistry at the University of Manchester.

Dr Christine Kitchen

Dr Christine Kitchen is the ARCCA Division Manager. Prior to joining the group in September 2007, Christine also worked at Daresbury Laboratory in the Distributed Computing group within the Computational Science and Engineering department, where she led the Commercial Outreach and Training activities and undertook a variety of roles, from system administrator through to providing technical expertise to both University and Industrial sites, particularly in the area of mid-range cluster computing. Christine's background is also in Chemistry and she started working life in the Biomolecular Modelling and Drug Design Group at AstraZeneca (Alderley Park, Cheshire, UK), running simulations to predict the electrostatic properties of molecules, as well as gaining her first experiences in the area of system administration.

Mr Huw Lynes

Huw Lynes is the HEC System Administrator for ARCCA. Prior to joining the group, Huw worked in the film industry as System Administrator for the Moving Picture Company in Soho (London, UK), before rejoining Cardiff University in 2006 as a member of INSRV, seconded to the Welsh e-Science Centre (WeSC).

Huw's background is in Microbiology and he spent some time studying Molecular Biology at the MRC Clinical Sciences Centre at the Hammersmith Hospital.

Dr Hugh Beedie

Dr Hugh Beedie is Chief Technology Officer (CTO) for INSRV and ARCCA.

Hugh has been working at Cardiff University for over 25 years, carrying out a range of technology related roles in the computing service. Along with the development of the underlying network, he has also led the teams responsible for developing many IT services for the University.

Hugh's ARCCA responsibilities are to ensure that ARCCA's services will seamlessly integrate with those of INSRV. He has also played a major role in, and was the original INSRV sponsor of the Condor project.

Dr James Osborne

Dr James Osborne is an Application Support Engineer for ARCCA. James joined the group from INSRV, where he was project manager of the Condor Project, a role he is migrating over into ARCCA.

His background is in Computer Science and his research interests include visualization, distributed computing, and computer science to support other scientific disciplines.

Miss Elizabeth Fitzgerald

Miss Liz Fitzgerald is PA to Professor Guest and provides administration support for ARCCA. With nineteen years senior secretarial experience gained in various industrial and Blue Chip companies, Liz provides a friendly, professional liaison to the group.

INSRV Affiliate – Dr Chris Dickson

Dr Chris Dickson co-ordinated the Advanced Research Computing programme at Cardiff University until Autumn 2007. Chris was then appointed to a new role in INSRV, that of Senior Consultant for Research Development. This key role will bring the research community together with IT solutions to enable the University's research agenda. Chris will also continue to work in partnership with ARCCA in developing new and existing client relationships, and in helping to ensure a coherent approach to research support and enablement between ARCCA and INSRV.

The Genesis of ARCCA

The beginning of Cardiff's High Performance Computing Service is set to coincide with the publication of this, the first ARCCA Annual Report. Reaching this point has taken a significant investment in time and effort from what has been a true team effort across the whole University.

In 2004, developments around an offer from IBM of a large supercomputer exposed serious concerns about the University's approach to High End Computing (HEC). This triggered an HEC consultancy study that was commissioned and reported in the same year. The study revealed that in spite of pockets of excellence, there were serious weaknesses that meant the University was not keeping up with other Russell Group Universities or the 'state of the art' generally. This led to a persistent failure to capitalise on undoubted opportunities in research.

The report proposed significant changes in how the University should approach HEC.

In response to this challenge, the University – under the direction of the Vice Chancellor –– established the HEC Ad Hoc Working Group (AHWG) to both mobilise and engage the University's HEC community. Chaired by Martyn Harrow, the Group was charged with proposing the required steps to elevate the University's approach and develop an HEC strategy.

Recognising the need for a catalyst for wider use both within and across disciplines, the group quantified this untapped potential, notably through work such as Condor and the research efforts of the School of Computer Science (COMSC) and the Welsh e-science Centre (WeSC).

The group crafted a compelling vision for HEC (see insert), and under the direction and close involvement of the AHWG – Hugh Beedie, Alex Hardisty, Chris Dickson and Tony Morgan amongst others – led an All-Schools coordinated approach behind a SRIF-3 bid for £2.9M that embraced many new research areas.

The University High End Computing Vision

To deliver to all researchers and collaborators, across all disciplines, the High End Computing services and resources they require to achieve their objectives, with minimum impact on their time and without the need for them to directly understand or 'program' the technology, and in a way which is highly cost effective both for Schools and for the University overall.

We will have the aim before the end of the decade of positioning the broad range of HEC enabled e-Research in the University firmly in the top half of the Russell Group Universities, when measured by any common indicator, and of achieving real leadership in a number of key selected areas.

At the heart of this bid sat the development of an institutional-level Cardiff University HEC organisation and focus, a development that led to the present ARCCA organisation.

ARCCA was to move, step by step, to delivering a new / up-to-date HEC service model, providing an Academic Facility to enable world-leading research across all disciplines. ARCCA itself would ultimately prove of international standing in its own right.

This journey was not straightforward – the first tender attempt had to be abandoned in May 2006 when the Group agreed to withdraw due to an identified lack of a suitable location. Undeterred, the Group persevered, and with significant progress on the space problem during the Summer / Autumn of 2006, the Tender process was restarted in October 2006. The Hardware and room fit-out tender documents were issued in March 2007, coinciding with the arrival of the ARCCA Director. The further steps of this story are developed in other sections of this report.

Many people have been involved in what has been, for the first time, an All-Schools coordinated approach to delivering the resources for enabling research computing across the entire University.

A special vote of thanks must go to the AHWG, the initial membership of which is shown; the Group is still extremely active, having held its 26th meeting in April 2008.

Initial Membership of the AHWG

- Martyn Harrow (Information Services Directorate) (Chair)
- Alex Hardisty (Welsh e-Science Centre) (Secretary)
- Chris Dickson (Information Services Directorate)
- Kathryn Hoose (Planning Division)
- Mike Davies (Finance Division)
- B.S. Sathyaprakash (School of Physics and Astronomy)
- Peter Knowles (School of Chemistry)
- Paul Smith (School of Medicine - Pathology)
- Nick Avis (School of Computer Science)
- Peter Kille (School of Biosciences)
- Huw Davies (School of Earth, Ocean and Planetary Sciences)
- Peter Halligan (School of Psychology)
- Peter Cleall (School of Engineering)
- Ian Matthews (PCAPH)
- David W. Walker (School of Computer Science)
- Plus other delegates as work proceeded

Introducing Advanced Research Computing at Cardiff

ARCCA exists to enable research and is driven by the University's research agenda. It will deliver Advanced Research Computing solutions and services of internationally-recognised quality, innovation and value by harnessing the power of high-end computing to support innovative and interdisciplinary research.

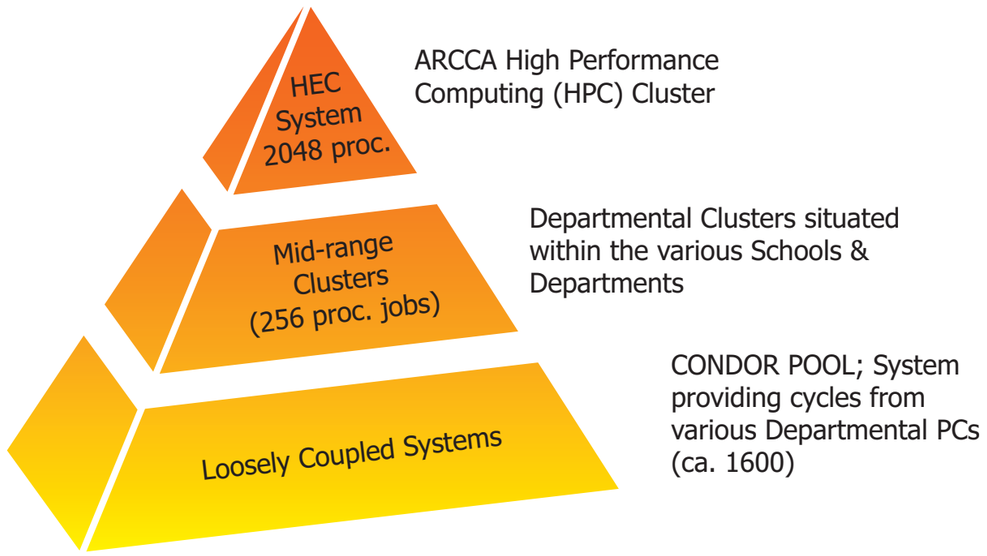
It's about **research pull not technology push.**

To set up the new ARCCA division in 2007, the University provided significant initial funding to staff the division with a small but expert team, led by internationally renowned expert Professor Martyn Guest.

"Cardiff's current research record is outstanding; there is huge potential for new and interdisciplinary opportunities which will be enabled through ARCCA" said Professor Guest

The team's focus will be on assisting new users with the use of this technology whilst collaborating with existing High Performance Computing (HPC) users to further their specific goals. The team is already available to help and support your research needs through a range of services.

The following schematic illustrates the hierarchy of high-end computing resources at Cardiff University.



The Branscomb Pyramid model of High-End Computing

To foster integration of existing systems and to assist users, ARCCA will provide tailored training packages and integrated support across the spectrum of ARC systems.

To enable the University to achieve its vision for ARCCA, it has invested in leading-edge IT equipment. A new supercomputer – funded by a substantial Science Research Investment Funding (SRIF3) grant from the Higher Education Funding Council Wales (HEFCW) – will be available to all academic Schools at Cardiff University.

At the start of 2008 the University selected Bull Information Systems to provide a new High Performance Computing solution to support a variety of advanced research projects.

This partnership will open up a range of new research frontiers. Research projects in the fields of astrophysics, fluid dynamics, geological simulation, materials science, and molecular simulation amongst many others will see major benefits, as well as non-traditional areas such as the humanities and the social sciences.

"We are proud to have made this investment on behalf of the University," said Professor Martyn Guest, Director of ARCCA. "We are working in close partnership with the key academic researchers to ensure the new equipment fulfils their research needs. The new supercomputer will enable us to compete on research terms with the best universities in the world, and provide an unique opportunity to deliver on ARCCA's core mission of Research Enablement – our goal is to establish Advanced Research Computing as an invaluable tool for research across all Schools."

The First Step: Condor

What is Condor?

Condor is a specialized workload management system for compute-intensive jobs.

It is a toolkit that makes use of the spare computing capacity of many PC workstations at Cardiff – turning them into a £500,000 supercomputer!

Users submit their jobs to Condor, Condor places them into a queue, chooses when and where to run the jobs based upon a policy, carefully monitors their progress, and ultimately informs the user upon completion.

Cardiff University's pool began as a pilot service back in April 2004 led by Dr Hugh Beedie in conjunction with staff at the Welsh e-Science Centre. Early adopters of the pilot service included researchers in the School of Biosciences, the School of Business and the School of Optometry and Vision Sciences.

Since January 2006 Condor has helped researchers from across the University speed up the processing of their research results – now ARCCA has adopted Condor as one of its first services available to assist researchers.

Cardiff University's Condor pool is the second largest in the UK with a total of 1600 machines providing a theoretical 1 teraflop of computing power on-demand to the University's researchers. This gives them a significant advantage over their competitors in other universities whilst saving half a million pounds.

James Osborne, Condor Project Manager and Application Support Engineer for ARCCA, said "The Windows Condor Pool is the most widely used computing resource on campus and has delivered over 2 million CPU hours since early 2006. The largest users of Condor are based in the Department of Primary Care and Public Health and are using Condor to help them analyse their data using combinatorial methods."

The Condor pool is also used by a number of other institutions including English Heritage and Velindre Hospital.

In the 4 years since its launch, Condor has provided over 450 years of CPU time!

Condor …

… is one of the largest pools in the UK
…and we plan to expand the pool by a factor of 4
…is probably the most utilised pool in the UK
… has more users than any other pool in the UK
…and we are working hard to keep it that way!

Did you know…?

…Condor is actually named after the bird of the same name! For more information see www.cs.wisc.edu/condor

…Condor can speed up the processing of research data by a huge

6000%!

"Condor has shown the spectacular results that can be achieved by providing research groups with access to high throughput computing services – it has already enabled research which otherwise could not have been done" said Martyn Harrow, Director of Information Services.

User profiles

Tom Beesley, School of Psychology

Tom Beesley, a researcher in the School of Psychology, has used Condor to gain insight into how the human sub-conscious works whilst acquiring new and sophisticated sequential skills - such as learning a language, playing a musical instrument, or driving a car.

Tom uses Condor to run multiple Simple Recurrent Networks (SRNs) - a special type of Artificial Neural Network (ANN) - simultaneously with different probabilities in order to discover the most appropriate SRN for a particular type of sub-conscious behaviour.

Tom used over ¼ million CPU hours in 2 weeks – that's the equivalent of 30 years on 1 PC!

Simon Caton, School of Computer Science

Simon Caton, a researcher working in the field of image processing and analysis, used Condor to classify schizophrenia in diffusion tensor images (DTIs). This provides important insights into the structure of the brain, and helps clinical researchers in the identification of brain regions, enabling evidence-based decisions.

Simon used Condor to run image analysis on a total of 400 machines – whereas previously he ran the processes sequentially on one machine, resulting in long execution times and large running costs.

Using Condor, the computation time for each dataset was reduced from 90 minutes to 10.

Cardiff supplies UK National Grid Service with first Windows Condor Pool

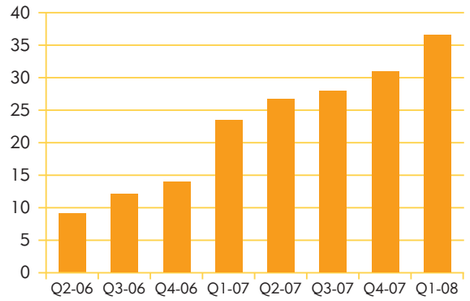
Cardiff University has increased its provision to the UK's National Grid Service (NGS) by making its Condor Pool accessible to NGS users all over the country.

Jonathan Giddy, Grid Technologies Co-ordinator for WeSC, said "The Windows Condor Pool can be used to perform a range of computations, from determining the structure of proteins to calculating radiotherapy dosages. By contributing these resources to the NGS we are enabling researchers nationwide to run a greater number of Windows based programmes, thereby continuing to open up the NGS to new types of user."

Who uses Condor and what are the Key Applications?

The numbers of Condor users have continued to increase since its introduction in 2006.

The University's Condor pool is currently used by researchers from 15 Schools, incorporating a variety of different applications. Some of these are summarised in the table below.



School / Department	Application Description
Architecture	Simulations of heating, cooling, lighting and water flows in buildings.
Biosciences	Comparing nucleotide or protein sequences to a database for statistical significance; Estimating admixture proportion and genetic drift using genetic markers; Generating evolutionary and family trees from genetic sequences using inference; Investigating population structure using multi-locus genetic sequence data.
Business	Simulating econometric models.
Computer Science	Image processing (using MATLAB) simulating radiation transport; simulating WIFI roaming in metropolitan areas.
Engineering	Codes to solve the travelling salesman problem, and to perform geological modelling.
History & Archaeology	Calculating probable age ranges for scientifically dated samples i.e. radiocarbon dating (in collaboration with English Heritage).
Mathematics	Simulation of non-Newtonian fluids.
Medicine	Statistical analysis (using R).
Optometry & Vision Sciences	Performing ab initio shape determination using simulated annealing with dummy atoms and residues.
Physics	Performing Fourier Analysis.
Psychology	Simulating Artificial Neural Networks.
Social Sciences	Performing Social Network simulations.

ARCCA Year One

Following the decision to set up the ARCCA division, a number of elements were put in place to deliver the University's High End Computing vision.

From setting the focus and priorities of the division - underpinned by a new governance structure - to putting in place fora to share good practice and ideas, and running the procurement process, the needs of the University research community have driven every step.

Here are some of the ways in which this has been achieved in our first year.

Focus and Priorities in Research

ARCCA is here to work with the University community to stimulate new research opportunities, and enable existing initiatives to take advantage of the enhanced capabilities for their research. It will build and sustain a position which takes the University to the forefront of leading research universities in the UK and internationally in this field.

To support and oversee the deployment of ARC to the University, a new organisational structure has been put in place including two new groups – the ARCCA Oversight Group and the ARCCA Operational & User Group.

The ARCCA Oversight Group

The University Research Committee, through this Oversight Group, will oversee and own the strategic development, policies and future of ARCCA.

Chaired by the Pro-Vice Chancellor for Research and comprising membership from across the University including the Dean of Interdisciplinary Research, Research Committee nominees, School of Computer Science and Research and Commercial Division representatives and the Director of INSRV to reflect University Board ownership, the group will:

- Advise ARCCA on achieving and promoting excellence in research;
- Monitor ARCCA's funding, sustainability and engagement with the research community;
- Provide guidance to ARCCA on the research directions to be supported, and the levels of support that should be delivered to these activities;
- Guide ARCCA as to the optimum allocation and deployment of University resources, from start-up activities to the on-going deployment of resources in support of mature projects;
- Provide guidance and advice as to the effectiveness of ARCCA's mandate of research enablement, ensuring that this enablement is delivered, and seen to be delivered, to the University community.

The ARCCA Operational and User Group

With representation drawn from all Schools, the ARCCA Operational and User Group will provide oversight across a range of operational activities, focusing on effective delivery to the research community by the ARCCA team.

This will include:

- Processes and procedures to make best use of the University's existing HPC and Grid computing assets;
- Arrangements and schedules for the management, maintenance and upgrade of facilities in line with the needs of researchers.

Members of the Group will use their personal expertise and judgement to contribute to the Group to reflect the needs and interests of Academic Schools and the wider interests of the University.

Clusters on the Cardiff Landscape

A number of Schools are already actively engaged in the support and exploitation of mid-range clusters – examples include Helix, the CUBRIC cluster at the Cardiff University Brain Research Imaging Centre, COMA and Explorer, and a variety of smaller clusters within the Cardiff School of Computer Science.

Integration and support

To foster the integration of these pre-existing clusters into the University's ARC framework and to support users migrating between systems, ARCCA will provide tailored support packages and integrated support.

Current uptake will be expanded through the provision of training courses and workshops to highlight the tools available and how best to deploy them to optimise the performance and usage of existing codes. Information and user guides on the ARCCA website will complement these courses.

Deploying these tools will enable researchers to develop more efficient codes which will, for example, scale to much higher processor counts, enabling larger and more detailed datasets to be analysed.

To promote the integration and support for such systems, ARCCA has initiated a forum for researchers, known as the **Cardiff Cluster Computing Forum (CCCF).**

Helix Cluster

What is a Cluster?

'Cluster' is a widely-used term meaning a number of independent computers combined into a single, unified system through software and networking. At the most fundamental level, when two or more computers are used together to solve a problem, it is considered a cluster. Clusters are typically used for High Performance Computing (HPC) to provide greater computational power than a single computer can provide.

The Cardiff Cluster Computing Forum (CCCF)

The CCCF provides a forum for sharing ideas and best working practice for clusters at the University, while ensuring their optimum utilisation. The forum brings together system administrators and researchers of School facilities.

The agenda and associated training programme is governed by the forum members, ensuring the necessary focus is given to achieve their research requirements.

Initially the CCCF will focus on a number of key operational activities:

1. Enabling the optimum deployment of technology in support of research computing;
2. Disseminating information and discussing all aspects of cluster technology, best-practice cluster management and system utilisation;
3. Providing a common software interface for clusters and their users;
4. Enhancing environments for Code Development and Performance;
5. Providing appropriate training and outreach;
6. Encouraging the usage of Cardiff visualization assets;
7. Providing regular technology updates to ensure Cardiff is at the forefront in understanding technological developments and associated research impact;
8. Mentoring.

An early task to be undertaken by ARCCA and the CCCF is a detailed requirements analysis of the existing users of departmental clusters and the Condor system. This will enable a better understanding of the current levels of awareness and expertise within the community, and the challenges researchers face in pursuing their computer-related research.

The CCCF will help to establish a set of common goals, including:

- Fostering collaboration and communication between ARC System Administrators;
- Agreeing methods to deliver standard reporting and monitoring of information;
- Developing and documenting best practices for the management and administration of ARC facilities.

Would you like to know more?

If you are an administrator of any ARC computer system at the University, a researcher utilising these systems, or a potential user of ARC resources and you are interested in learning more about the systems and associate training available, we invite you to join the group and related email list. For further information, please visit our web site: www.cardiff.ac.uk/arcca

The ARCCA Procurement

The University has invested in leading-edge IT equipment to transform its Advanced Research Computing programme.

2007 saw the official tendering process to procure a new supercomputer for the University. Funded by a £2.9 million SRIF3 award from HEFCW, the process saw ARCCA working with several leading supercomputer vendors, to ensure the University selected the best supplier.

The process culminated in December 2007 in signing a contract with Bull Information Systems Ltd to provide the new supercomputer which will be available to all academic Schools within Cardiff University.

This contract will not only provide Cardiff University with 'state of the art' technology but also a valuable partnership.

Equipment in the new ARCCA machine room

ARCCA is based in the Redwood Building on the Cathays Park Campus

"The University is delighted to be working in partnership with Bull on this project that will open up a range of new research opportunities", said Professor Martyn Guest. "The new supercomputer will enable us to compete on research terms with the best universities in the world, and provide a unique opportunity to deliver on ARCCA's core mission of Research Enablement – our goal is to establish Advanced Research Computing as an invaluable tool for research across all Schools".

ARCCA Implementation Schedule: 2007-2008.

- Technology Evaluation, Tender negotiations with prospective suppliers — Jun-Sep '07
- Bull chosen as partner for ARCCA Cluster System — Oct '07
- Contract Negotiations, Kit Installation, Acceptance Tests, Infrastructure — Nov '07-Apr '08
- Bull "Fast start" programme, underway, Awareness raising - Training Events — Feb '08-Apr '08
- **Start of Service,** "Impact" Programme underway, initial focus on Advanced Users — Apr '08
- Awareness and Outreach Programme in full swing - training Events, Theme Days etc — May '08

Green ARCCA

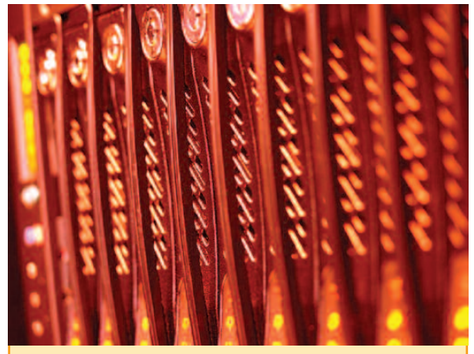
In line with the University's environment and sustainability goals, ARCCA's equipment is one of the greenest Advanced Research Computing services in the UK.

The new ARCCA installation uses water cooled racks which could lead to a £30,000 saving based on the air-conditioning running costs. This will be achieved using new cooling methods which cool the racks of processors directly, rather than cooling the air of the entire room.

The space needed by ARCCA has also been cut by the careful selection of high density equipment.

"Processor chips, for example Pentium processors, are getting smaller but not faster. At first this may not seem beneficial but it means we can put more processors on a single chip" said Dr Hugh Beedie, Chief Technology Officer. "In the space acquired by ARCCA we have not only four times the number of servers but the amount of electricity needed to cool the system is only 20% of the power needed to run it."

The cooling equipment purchased will pay for itself in 3 years and will last for 10 years.



Procurement Challenges

The Technology

During the initial evaluation of the ARCCA tender responses, the processor technology available changed dramatically with the arrival of "Quad Core" technology.

Working collaboratively with all of the short-listed suppliers, together with other SRIF-3 sites (e.g., Bristol and Birmingham), ARCCA staff collated a comprehensive set of benchmark results for the key applications, executing on both single processor units and small cluster systems.

By mapping the performance of real applications to the technology the team were able to make an informed decision on the optimum processor manufacturer - a vital decision in all procurements.

>>>

The Supplier

The ARCCA evaluation teams met with each of the short-listed vendors to discuss the final technology design, the partnership opportunities, plus the benchmarking and acceptance tests.

With the solutions offered by all short-listed suppliers being in principle similar, the differentiators were the quality, robustness and usability of the entire integrated solutions being proposed.

The key driver, and one emphasised throughout the procurement, was to develop a long-term relationship with the supplier – so the differentiator between the bids came down to the commitment and partnership opportunities offered to the University.

The Solution

The solution, endorsed by the ARC Working group, is an Intel-based quad-core solution from Bull which will create what would currently be the 3rd most powerful cluster in a UK University and is at least 10 times more powerful than any other cluster currently available at Cardiff University.

Bull impressed with the quality and expertise of their cluster design and technical knowledge.

What Bull brings to Cardiff University

- Bull develops their software in-house as well as manufacturing their own supermicro-based servers (NovaScale R440)
- Bull provides a "one-stop" shop for all queries downstream.
- Bull demonstrated their commitment to ARCCA and Cardiff University in the partnership opportunities they offered, creating a Centre of Excellence at Cardiff.

Technology-wise, ARCCA and Bull opted to go for:

- A 2048 quad-core Intel Xeon 3.0GHz (Harpertown/Seaburg chipset) cluster, with Infiniband Connect-X interconnect (1.2μs latency, 20Gbps bandwidth) over the entire cluster
- A Lustre parallel file system with 38TB Fibre-channel disk, and an additional 50TB SATA disk storage
- In addition to the cluster itself, the solution features four 16-way SMP systems to support those codes with shared-memory requirements

This solution will provide the highest density (as the R440 are dual motherboard nodes, equivalent to 16 cores per 1U node), with excellent power efficiency. It will give 24Tflops peak performance.

The Cardiff High Performance Computing Centre of Excellence

As part of its procurement process, ARCCA has developed a partnership with Bull Information Systems Limited with a view to establishing an HPC Centre of Excellence.

Bull is one of the leading manufacturers in Europe, large enough to support ARCCA, and yet small enough for Cardiff University to represent a prestigious partnership opportunity for the company. Such a collaborative enterprise is geared towards delivering a step-by-step approach to integrating and exploiting the HEC environment.

Bull can support ARCCA by providing advanced information and access to new technology, ensuring that appropriate and timely investments are made in HEC to assist the University in maintaining its position at the forefront of research enablement.

The formation of the Centre will extend the scope and quality of computer-based research support, enhancing the University's reputation in the UK and beyond.

Standing from left to right: Liz Fitzgerald (ARCCA - PA / Administration Officer), Jenny Newton (Bull), Hugh Beedie (Chief Technology Officer INSRV & ARCCA), Professor Peter Blood (Deputy Vice-Chancellor Cardiff University), Peter Ingram (Bull), Professor Martyn Guest (Director of ARCCA), Steve Moore (INSRV - Business Operations Manager), Christine Kitchen (ARCCA - Manager), Alec McSkimming (Associate Director INSRV - Business Development).
Sitting:Phillippe Miltin (Bull), Martyn Harrow (Director of INSRV)

Focus Areas for the Centre of Excellence

Initially the Centre will focus on five key areas:

1. Fast Track Start-up

To ensure the ARCCA service is fully operational when the cluster is delivered and installed, Bull will provide dedicated expertise in advance to work on porting and optimising the main Cardiff application codes.

2. Programming Support

To attract the highest calibre of programming expertise, Bull will co-fund two HPC specialists for the duration of the contract to address, (1) the technical expertise and experience required to port, optimise and assist in the development of key applications, and (2) as an outreach and training presence to raise the profile of ARCCA and Cardiff University.

3. Training and Development

A variety of training courses will be provided in order to ensure optimal use of the technology. A range of specialist high quality training support has been procured from Intel. Through this, the University will have access to high quality HPC experts from within both Intel and Bull.

4. Technology Evaluation

ARCCA and Bull's skills in performance analysis will be utilised in areas such as: access to, and demonstrations of, new technology; Cardiff acting as a beta-test and evaluation site; the authorship of HPC evaluation reports and white papers; and, the development and exploitation of a standard set of core UK university benchmarks.

5. Workshops and Awareness-Raising

The University will host a range of HPC events, workshops and training courses which will raise ARCCA's profile within the national and international communities and will evolve into information exchange programmes with European collaborators, including other Bull sites in France, Germany and Spain.

Established ARC Disciplines

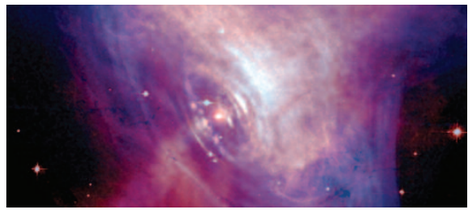
There are a number of well recognised, leading ARC researchers at Cardiff University, all of whom have been progressing their research either remotely, or at the University using existing schools-based clusters or Cardiff's Condor pool. It is these researchers who have collectively made the case for SRIF-3 provision and are expected to dominate early usage of the ARCCA service. A number of examples of their work, grouped according to discipline, are summarised below.

Physics and Astronomy

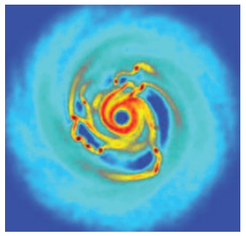
Prof. B.S. Sathyaprakash - "Black Holes, Big Bang and Gravitational Waves"

Prof. Sathyaprakash's research focuses on sources of gravitational waves and their detection. His research group is engaged in the analysis of data from the British-German GEO 600 and American LIGO interferometric gravitational wave detectors, mainly concerned with the detection of coalescences of compact objects (i.e. neutron stars and black holes), transients from supernovae and stochastic background. This research is computationally very intensive, in part due to the huge amount of live data being generated by the network of gravitational wave detectors across the world, and also because of the large parameter space in which the search is carried out.

Facilities used: Large cluster computers, as well as access to the international Einstein@Home collaboration.



Prof. Anthony Whitworth - "Simulating Star Formation"

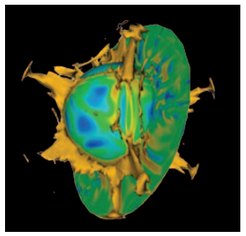


Prof. Whitworth's research focuses on simulations of collapsing molecular clouds, protostellar and protoplanetary discs, brown dwarf and low-mass star formation, binary star formation, computational radiative transfer, transport of ionising radiation in HII regions, massive star formation, and dynamically triggered star formation. Efficient implementation of much of this research has required specialist 'shared memory' cluster computers due to the close physical dependencies between different regions in the star forming region.

Facilities used: Prof. Whitworth currently uses two older shared memory machines which will be superceded by the ARCCA shared memory systems. When appropriate, he also occasionally makes use of national facilities such as the UK Astrophysical Fluids Facility at Leicester University.

Earth, Ocean and Planetary Sciences

Dr J. Huw Davies - "Mantle Convection Modelling"

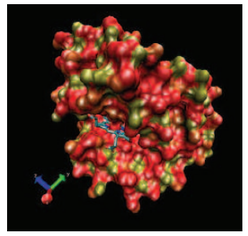


Dr Davies's research interests are in understanding the Earth's mantle, and how it drives plate tectonics. This is carried out by simulating mantle convection numerically using large supercomputing clusters, imaging the deep interior using seismic tomography, and attempting to understand the igneous output of the mantle (especially at subduction zones where one plate moves under another into the planetary mantle). This is computationally very demanding, given the size and complexity of the problem – simulating the fluid dynamics of the whole Earth's mantle using finite element modelling.

Facilities used: Large, high-memory cluster computer, with up to 1000 processors. Also using stereo visualization of 100 million data points in the whole Earth model, for both interpreting science and demonstration.

Chemistry

Profs Allemann, Harris and Knowles, and Drs Willock, Mella and Platts



There is a wide range of computational research underway in Chemistry, focused on applications in chemical biology, in the solid state, and in drug-related areas. Areas of ongoing research include (i) the atomistic simulation of enzyme-catalysed processes and other topics in chemical biology (Allemann), (ii) the crystal structure determination of biologically relevant molecules from powder X-ray data (Harris), (iii) chemical reaction mechanisms on the surfaces of metals, oxides and zeolites (Willock), (iv) the action of anti-cancer drugs on DNA fragments (Platts), and (v) Quantum Monte-Carlo simulations of the dynamics of proton exchange and transport in clusters and solution (Mella).

In addition to applications-based studies, Prof. Knowles, together with Prof. Walker in Computer Science, is working on high performance software engineering methods oriented towards computational chemistry.

Facilities used: Much of the current usage within Chemistry is carried out on the 200 processor Helix cluster, and the computation and 3D visualization facility located in Main Building.

Computer Science

Profs N. Avis, D. Walker, A. Grey, O. Rana and Mr. A. Hardisty

Research interests focus on:

(i) Novel approaches to HPC through the Centre for Reconfigurable Computing and Application Acceleration (a focus on the use of application acceleration techniques based on FPGAs, ASICs or GPUs).

(ii) Experimental Grid systems for developing / piloting Grid middleware, with Cardiff a National Grid Service "Partner".

(iii) Large scale linear algebra, numerical solution of PDEs, grid solvers for Maxwell equations, and numerical methods for inverse and imaging problems.

Wider interests include distributed collaborative computing, parallel algorithms, scientific visualization, programming methodologies and problem solving environments.

Engineering

Dr P.J. Cleall, Prof. B.L. Lin and Dr C.A. Wilson

While HPC has the potential to impact across all of Engineering's six multidisciplinary research institutes, current usage is concentrated within the Institute for Sustainability, Energy and Environmental Management. Research areas include the modelling of geo-environmental problems and the analysis of aerospace transient structural dynamics problems; biomechanics; computational mechanics and structural integrity.

For more examples and case studies see www.cardiff.ac.uk/arcca/enablement

What's to Come ...

Emerging ARC Disciplines

ARCCA is actively engaging with individual researchers and research groups across the University to foster emerging ARC disciplines.

Advanced Research Computing techniques can potentially be used by all Schools and ARCCA aim to facilitate this. To achieve this, an outreach programme is being established to engage with researchers from 'new' disciplines, with metrics identified for research enablement, and the aim of encouraging a broader uptake of ARC techniques though an appreciation of the potential benefits of ARC in many novel areas of research. A programme of School visits has already been initiated to work with researchers to identify how ARC can assist and help in their research.

The University's Modern IT Working Environment (MWE) programme will contribute to this aim by making new tools available to the University's staff and students, with the goal of enhancing their everyday interactions and working practices. In the longer term, the MWE will ensure maximum availability and ease of use of HPC tools for all the researchers who need access to them. The ARCCA team will also provide expertise to widen access to these facilities.

Many of the emerging disciplines are already making use of the ARC capabilities delivered through the Condor pool, incorporating a variety of different applications that have been summarised earlier in this report. The inset serves to reinforce the earlier summary by pointing to current usage.

Emerging Disciplines:

Biosciences: informatics in support of multiple projects, and neuroscience; DNA sequencing; plus support for EMRIC (Experimental MRI Centre)

Mathematics: Non-Newtonian computational fluid dynamics, developing software for predicting complex flows of polymer solutions and melts.

Pharmacy: molecular modelling and medicinal chemistry.

Psychology: Artificial Neural networks, partnership with CUBRIC.

Medicine: Pathology - systems biology/cancer research; Epidemiology, Statistics and Public Health.

Dentistry: biomechanics and biomaterials, using finite element modelling and engineering applications and techniques, to model and predict the movement of teeth, prosthetics and soft tissue in the mouth.

Architecture: heat flow/dissipation, building properties.

Optometry: protein modelling and other applications.

History and Archaeology: More precise dating of archaeological artefacts using Bayesian statistics.

Business: Econometric modelling.

Social Sciences: Social network simulations.

Training and Workshops

ARCCA have organised a programme of events and training courses which coincide with the start of service of the SRIF-3 funded Cluster.

These courses are based around the ARCCA system but will also be applicable to research and code development performed on departmental clusters as well as desktop/laptop systems.

Events range from introductory modules for users new to Advanced Research Computing, through to more advanced topics such as code optimisation for more experienced users. Some of these courses and workshops will be given in partnership with Bull, the ARCCA cluster provider, under the auspices of the Centre of Excellence, utilising the expertise of both Bull and Intel engineers.

The Official Launch of ARCCA

The official launch of ARCCA will take place in June 2008, and will comprise three main events:

The Grand Opening - 2 June

This will take place in the form of a 2-3 hour long event, involving key figures from the world of High Performance Computing, delegates from Bull, and members of staff from all Welsh Universities. It will see ARCCA officially opened by Wales' First Minister, Rhodri Morgan.

Opening of the Cardiff HPC Centre of Excellence - 4 June

The aim of this event will be to showcase the system and the people involved.

Guests and speakers will include key figures from Bull, who will talk about their aims and intentions for the Centre. Delegates from APC and Comtec will talk about the energy efficiency of the machine room infrastructure, while those from Intel will provide information about training packages.

Tours of the machine room will also be provided throughout the event.

Research Seminar Afternoon - 4 June

This seminar event will involve presentations from Cardiff researchers who will speak about their science and research agenda for the ARCCA cluster, highlighting what can be achieved on large systems.

Glossary

ARC (Advanced Research Computing)

Advanced Research Computing is the use of computing and data resources for research, which are beyond the capabilities of the average desktop or laptop computer.

This technique uses leading-edge IT resources and tools to pursue research; including computer simulation and modelling, manipulating and storing large amounts of data, and many other methods to solve research problems that would otherwise be impossible.

Condor

Condor is a specialized workload management system for compute-intensive jobs. It's a toolkit that makes uses of the spare computing capacity of PC workstations.

Cluster

'Cluster' is a widely-used term meaning independent computers combined into a unified system through software and networking.

HEC (High-end Computing)

High-end Computing (HEC) is a hugely powerful technique which uses leading-edge IT resources and tools to pursue research; including computer simulation and modelling, manipulating and storing large amounts of data, and many other methods to solve research problems that would otherwise be impossible. At Cardiff, we use the term Advanced Research Computing to cover the whole range of HEC and e-research techniques that are available to today's researchers.

HPC (High Performance Computing)

Use of (parallel) supercomputers and computer clusters, that is, computing systems made of multiple (usually mass-produced) processors linked together in a single system with commercially available interconnects.

SRIF

Science Research Investment Fund - a major programme of investment in physical infrastructure for research funded jointly by the four UK Higher Education Funding Bodies and the Office of Science and Technology (OST).

Supercomputer

A supercomputer is a computer that is considered, or was considered at the time of its introduction, to be at the forefront in terms of processing capacity, particularly speed of calculation.