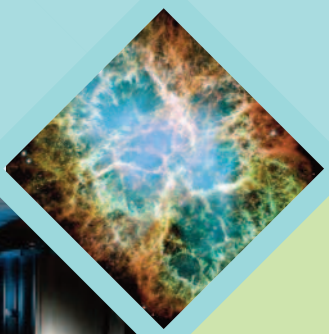
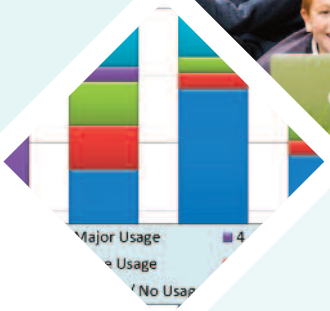
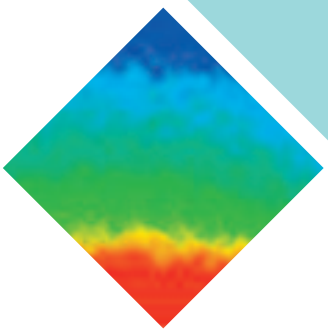
CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

ADVANCED RESEARCH COMPUTING @ CARDIFF







ANNUALREVIEW 2015/16

Advanced Research Computing @ Cardiff

Annual Review 2015/16

ARCCAREVIEW

Foreword 3

Advanced Research Computing @ Cardiff 4
- Cardiff's user community 4
- Project-based overview of the ARCCA facilities 5
- Use of the ARCCA facilities 6
- Research impact 6
- The ARCCA survey 2014 8
- The ARCCA training programme 10
- REF 2014 and ARCCA 12

Projects in review 13
- Enhancing the efficiency of the Redwood datacentre 13
- Upgrading the ARCCA supercomputer software 14
- New Intel processors for Cardiff's Gravitational Physics Group 14
- Hosting equipment for the CLIMB consortium 16
- Lost Visions – Retrieving the visual element of printed books 17
- Pilotomics – A building-block approach to data storage 18

What's next 19
- Promoting utilisation – the Galaxy Gateway 19
- Managing Workflows – Rose and Cylc 20
- HPC for undergraduates – WebMO and the GridChem Gateway 21
- ARCCA and assessing the cloud 22
- Future technology challenges 22

Outreach and collaborative frameworks 24
- The GW4 HPC landscape survey 24
- Usk Code Club visits the Redwood datacentre 25
- Glossary 26

Delivering on Cardiff University's research agenda 27
- Accelerating research through HPC and Big Data 27

Foreword

I am pleased to present the latest review from ARCCA, Advanced Research Computing @ Cardiff. The report introduces a selection of the highlights and developments of the past year. It details some of our steps to provide assistance across all levels of the University that characterise the research computing landscape in support of 'The Way Forward'.

The past eighteen months have been a truly exciting time for ARCCA, a period in which all the hard work invested in the Raven HPC procurement has led to a stable and reliable service. ARCCA has adapted the facility to support 'add-on' services based on research grant income. New partitions have been added to host multi-site research grants such as the STFC Gravity Waves LIGO consortium, as well as new services (e.g. the MRC CLIMB Consortium). This has enabled ARCCA to host technology for research groups that have specialist requirements, reducing their total cost of ownership while offering additional capabilities to the community when not in use by the consortium.

Whilst the availability and support of both the cluster and new partitions are vital components of the service, so too is the increasing range of training and outreach activities – both are described fully in the report.

The report is structured to run through the variety of ARCCA support and enablement activities. Our users have kindly contributed research highlights and associated activities being undertaken on the cluster, illustrating the range of research undertaken using our services.

If your research has the potential to use these techniques, or if you would like to find out more, please visit our website (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 complete your computational based research more effectively.

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

Professor Martyn Guest
Director of Advanced Research Computing @ Cardiff

Advanced Research Computing @ Cardiff

CARDIFF'S USER COMMUNITY

User registrations provide the base level of metric for Raven, capturing where a request has been submitted to use the system and the access subsequently provided. Such registrations show how well ARCCA is promoting usage and interfacing with Cardiff University researchers. Current records show that 741 users have had access to HPC services at Cardiff since January 2009. For comparison, the total number of active accounts/users on Raven in September 2015 is 465 – the difference is due to users leaving the University, at which point their account is deactivated.

Records show that 310 unique users ran jobs using the Portable Batch System (or simply PBS) job scheduler within the same period. This number continues to climb, suggesting that a growing number of users understand the batch scheduler, supported by the ongoing training programme delivered by ARCCA staff. A lack of understanding of the scheduler has often been cited as an obstacle to HPC use.

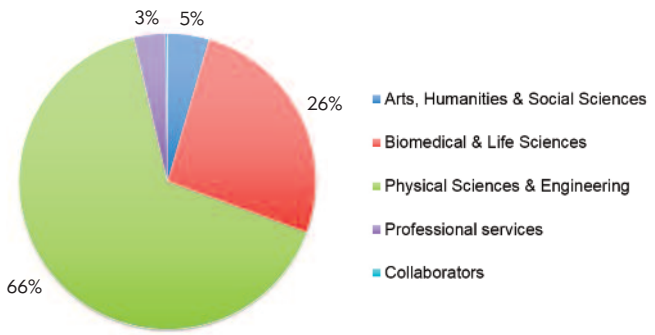


Figure 1 - User registrations by University College, Department, external collaborators

Figure 1 shows the user profile by College and Department in September 2015. It is encouraging to note that the Biomedical and Life Sciences community is becoming a larger section within the ARCCA community, with the present figure at 26%, and is continuing to rise. Our target has been to see an increase in the number of users in this College and there is room to see this share increase still further.

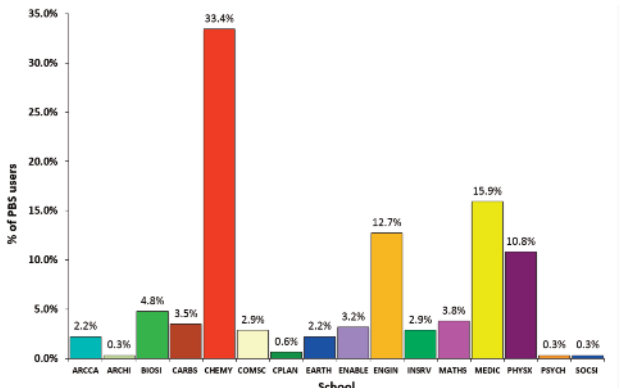


Figure 2 - PBS users by School (or similar)

The distribution by School of users who have run jobs (identified by Unix group) is shown in Figure 2. This again shows the main change has been an increase in the number of users within the Bioscience and Medical Schools, plus the introduction of Social Science compared to previous analyses.

Among the key tasks for ARCCA are to understand why users who log in to the systems do not use the job scheduler, and to increase usage of the system by under-represented Schools.

PROJECT-BASED OVERVIEW OF THE ARCCA FACILITIES

Work was undertaken during 2014 to reinstate project-based accountancy on the Raven system.

Every user of the PBS job scheduler on Raven is required to register to at least one project, but multiple projects are allowed per user. The distribution of the 106 projects by College as of 1st October 2015 is shown in Figure 3. Not surprisingly, this project-based distribution follows that based on the number of users, with Physical Sciences and Engineering providing 80 of the 124 projects, followed by Biomedical and Life Sciences with 27, and the Arts, Humanities and Social Sciences with ten projects.

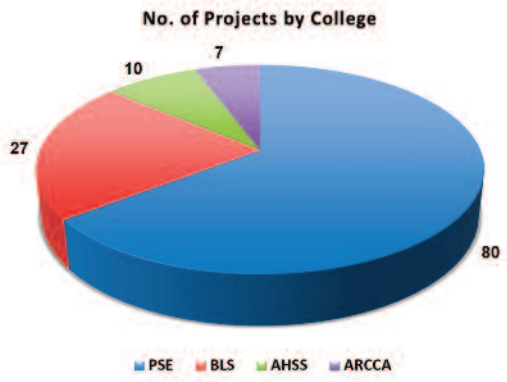
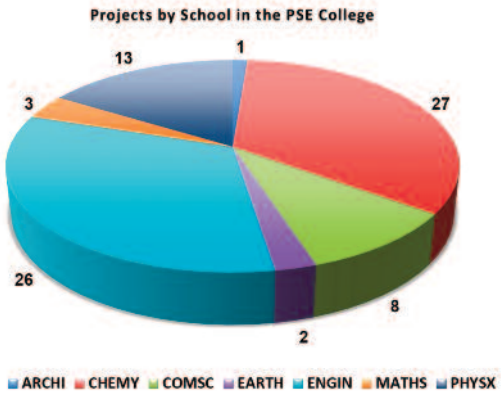


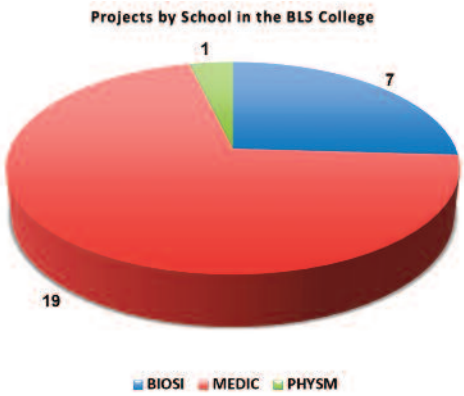
Figure 3 - Number of projects by College.

The distribution of projects by School within each of the Colleges is shown in Figure 4. This suggests that the dominant Schools within the College of Physical Sciences and Engineering are Chemistry (with 27 projects) and Engineering (with 26). The School of Medicine provides the majority of the projects within the College of Biomedical and Life Sciences (with 19 of the 27 projects), while the Cardiff Business School has the dominant share within the College of Arts, Humanities and Social Sciences, with seven of the ten projects.

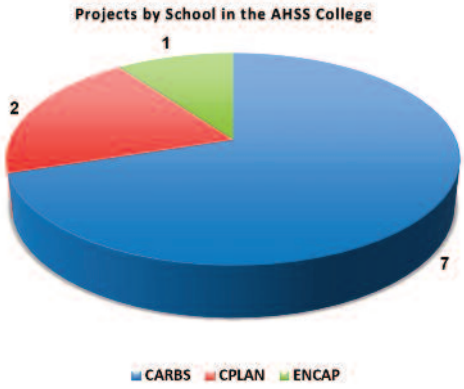
Figure 4 - Number of projects by School in each of the Colleges



(a) Physical Sciences and Engineering



(b) Biomedical and Life Sciences



(c) Arts, Humanities and Social Sciences

USE OF THE ARCCA FACILITIES

Use of the Raven supercomputing service is distributed across the High Throughput Computing (HTC) and High Performance Computing (HPC) partitions through different queues. The High Throughput Computing queue is for predominantly serial jobs that should run within a single node, whilst the High Performance Computing queue requires a larger number of nodes per job and uses a parallel communication library. The HTC partition, originally comprising 864 Intel Westmere cores, has been augmented during 2015 with the arrival of 1,440 Intel Haswell cores.

Figure 5 reflects the use of HPC queues on Raven over the past 12 months between October 2014 and September 2015. Note that processor (CPU) time (in red) should always be lower than the elapsed (wall clock) time (in blue); the ratio between CPU time and wall clock time for a given job is a measure of the efficiency of that job. This suggests that the majority of codes that run in the HPC queues are executing efficiently and are configured to write to the high performance parallel file system (Lustre) rather than to the slower /home (NFS) partition.

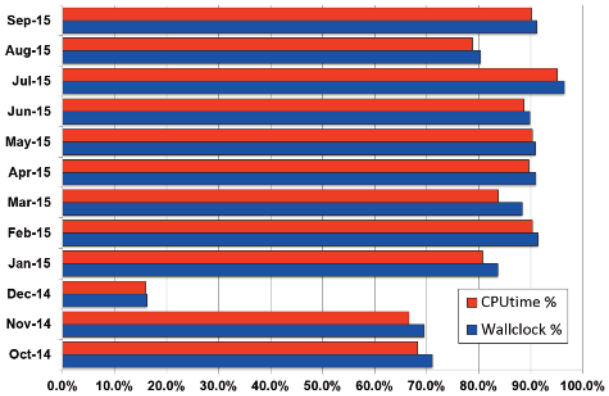


Figure 5 - Raven HPC utilisation between October 2014 and September 2015

THE CURRENT RAVEN SUPERCOMPUTER SERVICE (commissioned in December 2012), has run over 2M jobs, delivering 42.6 Million computing hours while operating at ~

80%

sustained utilisation.

Prior to the turn of the year, the overall utilisation of the HPC queues lay between 75 - 80%, a sustainable level that enabled us to accommodate future research needs. Since that point, however, usage levels have remained at over 90%. This has led to significant wait times for users, and hence a move to accelerate the procurement of the Raven replacement system, which is currently scheduled for mid-2016.

RESEARCH IMPACT

Whilst usage figures demonstrate the demand for a centralised HPC service, effective use of the available resources can be more accurately gauged by the volume and impact of published material generated via access to ARCCA facilities. The published material generated is grouped according to six categories:

1. A list of all journal publications, including those either submitted, in preparation, and accepted/in press, where the associated research has at least in part been facilitated using ARCCA resources.
2. A record of conference proceedings, where again the associated research has at least in part been facilitated using ARCCA resources.

3. A summary of presentations and posters given at national and international meetings and conferences that have included results obtained using ARCCA resources.
4. Other activities that have been facilitated by the availability of ARCCA resources, e.g. research projects, application support, prestige factors, events, reports etc.
5. Specific research projects that have been enabled by access to the facilities; PhD studentships and undergraduate projects.
6. A summary of posters given at national and international meetings and conferences that have included results obtained using ARCCA facilities.

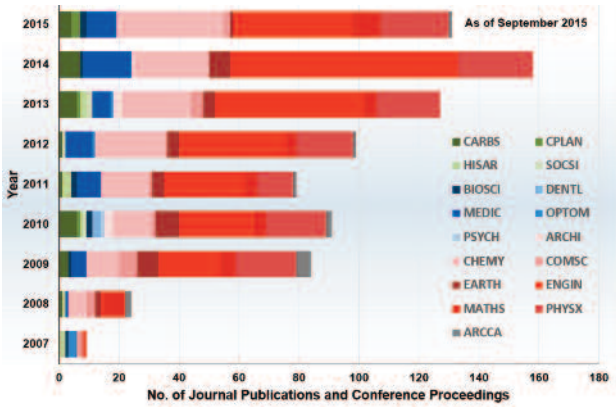


Figure 6 - Number of published journal articles and conference proceedings per year from authors using ARCCA services. Schools are grouped by college: Arts, Humanities and Social Sciences (green), Biomedical and Life Sciences (blue), and Physical Sciences and Engineering (red).

The response is summarised in Figures 6 and 7. Figure 6 shows the number of journal publications and conference proceedings generated using ARCCA facilities. The published papers show a clear increase over the past eight years, which looks set to continue for 2015 when viewing the available nine months of data. Conference proceedings are more variable but also show a clear upwards trend over the period in question. The college from which the publications originate is strongly weighted towards Physical Sciences and Engineering. Biomedical and Life Sciences shows a slow increase, and Arts and Humanities could be argued to show a downward trend.

A pictorial representation of these figures is provided in Figure 7. This shows the number of Research Outcomes as a function of category and School. Not surprisingly, the greatest number of outcomes arises from researchers in those Schools who have dominated usage of Merlin and Raven over the past five – six years (the Schools of Chemistry, Earth Sciences, Engineering, and Physics).

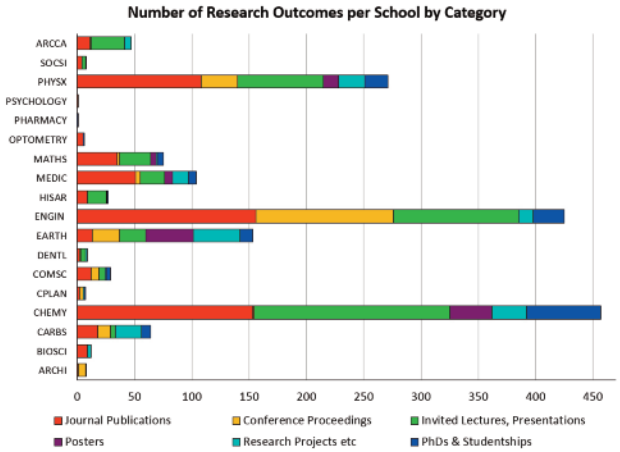


Figure 7 - The number of research outcomes as a function of category and School.

Since its inception in 2007 until September 2015, the use of ARCCA has supported

802 publications, of which

590 are journal outputs.

487 invited lectures and presentations have been made on the work facilitated by ARCCA.

151 PhD students have used the facility.

THE ARCCA SURVEY 2014

The ARCCA survey in early 2014 looked to understand both service development requirements and user satisfaction with the current service and supercomputer facilities. Three surveys were distributed to reflect the diverse community of users, targeting:

- Category 1 researchers (those with potential HPC requirements but who are not currently using computers to support their research);
- Category 2 researchers (those who have started to utilise computational techniques but are still not using clusters); and
- Category 3 researchers (covering experienced computational researchers who are using research computing in progressing their studies).

THE ARCCA SURVEY AND THE 'TOP TEN' ISSUES

1. General awareness of services
2. Accessibility
3. Training and support
4. Optimising job turn-around
5. Optimising application performance
6. Service profile
7. Knowledge awareness of the centralised HPC Services
8. 'Big Data' requirements
9. Accelerator technologies
10. Approach to future surveys

The survey returned 45 complete responses across 15 Schools. Many of these were undertaken at the research group level, and with on average of between three to five members per group, an estimate of the response level would suggest ca. 100 users of ARCCA services.

In many instances, the community's views were consistent across the different survey categories; in particular, the importance of the outreach activities and training/support requirements correlate across the different categories. The major distinguishing feature between the experienced and potential researchers was the requirement to analyse large datasets, requiring a different modality of use compared to traditional HPC facilities.

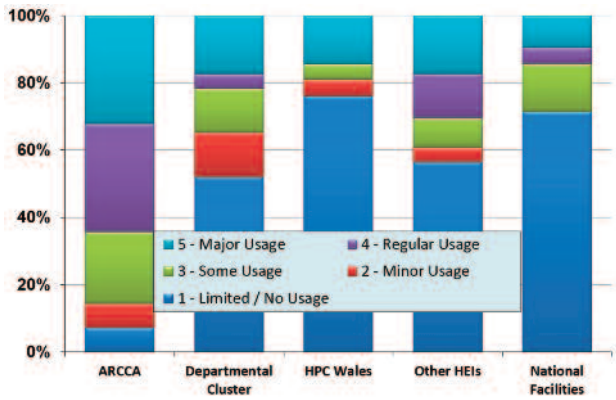


Figure 8 - Breakdown of HPC system usage by category 1 and 2 researchers

Figure 9 - University's environment for supporting HPC (ranked from 1 to 5)

Both the experienced and novice surveys enquired about access to and availability of clusters, from departmental systems through to national facilities. There are a range of different systems across the University, with novice users predominantly using either local or ARCCA systems; only a few of the more experienced researchers have access to national facilities or services at partnering institutions. As can be seen by Figure 8, the dominant computational

facility used was that run by ARCCA – over 90% of the responses were from those who had accessed the Raven supercomputer system, with over 70% either regular or major users of the service.

Figure 9 shows the response to the value of Cardiff University's environment for advanced research computing, using a ranking system of one (insufficient) to five (excellent). Overall, the responses demonstrate that the services are meeting the community's needs, particularly the experienced researchers, with an average score of 4.3. The novice researchers' average was slightly lower with an average rank of 4.0. By way of summary, the key survey findings and ARCCA's response are presented below.

KEY SURVEY FINDINGS

1. **General awareness of services**
   Updating the website to provide a summary of the variety of software currently installed on Raven (brief descriptions of the applications and associated functionality).
   Run an annual event showcasing the research undertaken on Raven.
   Produce a series of case studies highlighting the variety of research supported by ARCCA.
2. **Accessibility**
   Developing web-based interfaces to using Raven (particularly for using the job scheduler).
   Better information on account request processes.
3. **Training and support**
   Developing links to relevant online training facilities and materials (short-term).
   Develop in-house online training modules (longer-term).
   Run training courses on basic scripting techniques and job submission scripts.
4. **Optimising job turn-around**
   Regular reviews of job scheduler policies and associated wait times.
   Diverse demands from 32-128 core jobs, 1024+ core jobs plus rapid turn-around for code development (short-term).
5. **Optimising application performance**
   Profiling the 'primary' codes and optimising to ensure efficient usage of Raven.
6. **Service profile**
   Restoring confidence in the service after the data loss incident (category two users).
7. **Knowledge awareness of the centralised HPC services**
   Organise a series of awareness-raising events in the Colleges to increase the knowledge of ARCCA.
   Greater engagement with the Graduate School – involvement in induction days.
8. **'Big Data' requirements from category two and three researchers**
   Not as great as anticipated, but suspect some of the category two researchers assumed requirements understood as already engaged with ARCCA.
   A requirement to be supporting a variety of data sources for analysis (with appropriate support from ARCCA).
9. **Accelerator technologies**
   Limited demand for novel technologies either now or in the future. This may be as much an awareness issue as a lack of demand.
10. **Approach to future surveys**
    Greater support should have been provided to category two and three researchers in completing the survey – some feedback has suggested that terminology was confusing.

A number of services have been implemented to address the feedback from the research community. The information on the website has been reviewed and refreshed to improve the format of a number of forms (e.g. account requests and project identification). In addition, information has been created describing the variety of software available on Raven.

To simplify access to the Raven supercomputer, a number of web interfaces have been deployed, from the Galaxy solution (Bioinformatics) through to WebMO (Chemistry). Customisable workflows using the Rose / Cylc software are being piloted to assess the impact of web-based interfaces. Any researcher interested in workflows is encouraged to contact ARCCA so we can collaboratively develop these to simplify using the supercomputer.

In terms of raising the profile of ARCCA, a number of drop-in sessions are being trialled to improve the opportunities for researchers to raise speculative enquiries about supercomputing capabilities and ask questions about the service. We are also piloting a live-chat service, to provide instant access to technical staff during specific hours, enabling interactive resolution of problems rather than researchers having to raise service tickets.

Other service improvements are in the pipeline, including new technical forums and a revised training programme to align with the start of the new academic term. Operational activities, such as job scheduler (queue) optimisation and reconfigurations are regularly proposed to the ARCCA Governance groups for endorsement to be deployed on Raven.

Whilst the ARCCA User Survey has highlighted a number of opportunities to refine and improve supercomputing provision to the University, the team is constantly seeking feedback and recommendations on how to continue to develop services to meet evolving computational research requirements. If you do have suggestions on new services that would be beneficial, please contact us to discuss your ideas.

THE ARCCA TRAINING PROGRAMME

ARCCA provides a programme of training courses to help use the Cardiff Supercomputer and associated facilities. These courses are based around the ARCCA Raven system but are also applicable to research and code development performed on departmental clusters as well as desktop/laptop systems. Training ranges from introductory modules for users new to advanced research computing, through to more advanced topics such as code optimisation for more experienced users.

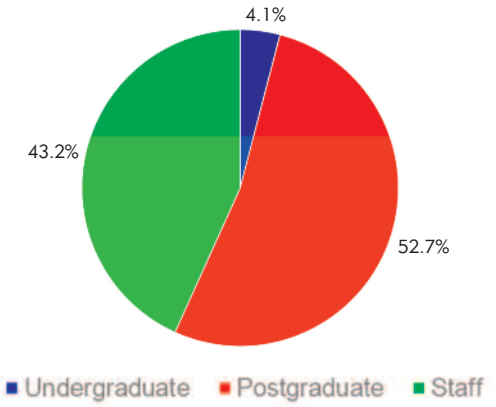


Figure 10 - Registrations by group for ARCCA training in 2014-15.

The majority of courses and workshops are classroom taught with hands-on tutorial examples. Examples of the courses currently available include:

- **An introduction to Advanced Research Computing**
  This module gives a quick insight into what research computing means and how you can use it.
- **An introduction to Linux with command line (and Windows 101)**
  This course concentrates on the use of techniques to improve researchers understanding of the command line, the use of common editing tools, and answers any queries users may have in using the Linux interface.

- **Supercomputing for beginners**
  This half-day course provides an overview of the Raven supercomputer and explains how to use the system.
- **The Raven supercomputer – advanced hints and tips to get the best results**
  The Raven supercomputer is a complex service – this course aims to review some of the techniques that should help users get the best out of the cluster without the need to alter their programme.
- **An introduction to parallel programming for beginners (OpenMP and MPI)**
  This session introduces programmers to the basics of parallel programming using MPI and OpenMP, looking at how the most common methods are used to take advantage of multiple cores.
- **Computational Physics 3rd year undergraduate degree course**
  ARCCA's contribution to the course consists of two lectures on the demands and benefits of programming in a distributed memory high performance environment followed by four lab classes, including course notes and practical examples.

If you think a specific course or workshop would be beneficial for a sub-group of your School or Department (up to ten people), please contact us. We will tailor the course to meet your requirements and run it at a convenient time within the School (subject to available facilities).

Course duration ranges from half-day (three hours) through to a whole day. The full training programme for the sessions taking place in 2015-2016 are available on the ARCCA website (www.cardiff.ac.uk/arcca).

Training sessions take place in the Julian Hodge Study Centre, Cathays campus.

REF 2014 AND ARCCA

The 2014 Research Excellence Framework (REF) was an important milestone for the University. As part of the REF, each assessment area was asked to complete a form, containing a number of templates including research impact and research environment. ARCCA was mentioned across many of the assessment areas, particularly by Schools in the Colleges of Physical Sciences and Engineering and Biomedical and Life Sciences.

At least 12 Schools acknowledged the ARCCA facility in the research environment or research impact returns of REF 2014 as an essential enabling service.

Template	Unit of Assessment	School
REF5	3. Allied Health professions, Dentistry, Nursing and Pharmacy	Mostly Life Sciences
REF5	4. Psychology, Psychiatry and Neuroscience	Medicine and Psychology
REF5	7. Earth Systems and Environmental Sciences	Earth and Ocean Sciences
REF5	8. Chemistry	Chemistry
REF5	9. Physics	Physics and Astronomy
REF5	10. Mathematical Sciences	Mathematics
REF5/3a	11. Computer Science and Informatics	Computer Science and Informatics
REF5	14. Civil and Construction Engineering	Engineering
REF5	15. General Engineering	Engineering
REF5	16a. Planning and Geography	Planning and Geography
REF5	16b. Architecture	Architecture
REF5	17. Geography, Environmental Studies and Archaeology	History, Archaeology and Religion

Projects in review

ENHANCING THE EFFICIENCY OF THE REDWOOD DATACENTRE

ARCCA, in collaboration with colleagues in Portfolio Management and IT Services, Estates, and external contractors, has undertaken a major refurbishment programme of the Redwood datacentre. In addition to improving the Power Usage Effectiveness (PUE) from 1.7 to 1.2 and enhancing the energy efficiency of the facility (by ca. £125K p.a.), the associated carbon emissions have been significantly reduced. Furthermore, the ARCCA team has optimised the footprint of services hosted in the datacentre and enabled the University to support a number of high profile grant awards.

The primary objective of this work was to ensure the datacentre was fit-for-purpose to support existing HPC requirements, at the same time enhancing the infrastructure to enable the full capacity of the room to be realised whilst reducing the operational overheads. This was based on a return on investment over five to seven years to recover the capital infrastructure costs by increasing the efficiency of the datacentre.

With the serious degradation of the original Airedale chiller units (120kW cooling units), the Redwood datacentre was fluctuating between a PUE of 1.7 and 1.8 (originally when the room was first commissioned, it was between 1.5 and 1.6). An ideal PUE would be 1.0 (anything less would mean the room is actually generating power to other locations, costing nothing to cool the equipment in the datacentre). A PUE of 1.7 means that every 1kW of power consumed by the IT equipment requires 700Watts electricity to produce the cooling (and associated datacentre operating overheads such as lighting).

The programme of work required a week's downtime of the services during December 2014, and was completed on time and within budget. Realising the operational efficiencies will be impacted by ambient external temperatures, two tuning exercises were undertaken to determine the optimum set points for the chillers and associated cooling infrastructure. These were performed in January 2015 and repeated during the summer to validate the operational set points to maintain the optimal PUE performance. It is pleasing to note that calculations are demonstrating a PUE between 1.15 – 1.25, so exceeding our estimations.

APC by Schneider Electric

coltherm

comtec power

UPGRADING THE ARCCA SUPERCOMPUTER SOFTWARE

To coincide with the datacentre refurbishment works in December 2014, upgrades to the cluster management software, lustre file system, and job scheduler were also undertaken on Raven, using the planned downtime period to minimise the impact on the user community. This work was important to ensure the cluster was on the most up-to-date version of the Bull cluster management suite, enabling us to benefit from updates and security enhancements.

The upgrade also introduced new functionality to help improve the efficiency and usability of the Raven supercomputer:

1. **Use of Altair's Compute Manager**, providing users with the facility to submit and manage jobs via a web page.
2. **Provision of a purge capability for the /scratch file system**, using the Robinhood (http://arxiv.org/pdf/1505.01448.pdf) utility. This will ensure usability and availability along with file system reliability.
3. **Restricting long batch array jobs to a small subset of nodes**, with the addition of new queues to circumvent issues where long batch arrays dominated the serial queue.
4. **Production of an end of job summary**, with more information presented to the user on job completion to enable fine-tuning of their job scripts.
5. **Enforcing memory limits in the serial queue** to prevent jobs crashing nodes and impacting those users whose jobs had been sharing the stricken nodes.

NEW INTEL PROCESSORS FOR CARDIFF'S GRAVITATIONAL PHYSICS GROUP

The LIGO Scientific Collaboration (LSC) is a group of scientists seeking to make the first direct detection of gravitational waves, to use them to explore the fundamental physics of gravity, and to develop the emerging field of gravitational wave science as a tool of astronomical discovery. The LSC works toward this goal through research on, and development of, techniques for gravitational wave detection; and the development, commissioning and exploitation of gravitational wave detectors.

Gravitational Physics

Using the serial partition of the Raven system, the group search for the signs of gravitational waves by processing large amounts of interferometric data.

Responding to a request for assistance from the group's Dr. Stephen Fairhurst, ARCCA led the negotiations with the University over the expansion of the Raven serial computing partition. The aim was to deliver a four-fold performance improvement over the existing 600 core serial partition procured as part of the original Raven system in 2013, to be available by March 2015.

To identify the optimum solution to meet this criteria, the Gravitional Waves group developed a benchmark that ARCCA and Bull staff ran on a number of different clusters to understand the performance profile of the code. Although there had been significant development by the LSC in the USA to effectively exploit GP-GPU solutions, the majority of the GPU-enabled functionality wasn't currently being used by the UK sites[1]. With multi-core systems now mainstream, performance comparisons became more complex as each processor release has a range of core counts and clock speeds – so a number of different combinations of tests needed to be undertaken to determine the optimum solution

Based on this performance analysis, the final proposed solution comprised 1,440 Intel Haswell cores – the E5-2680v3, 12-way 2.5 GHz core solution with 4GB memory per core. This design was agreed across the grant consortium sites (Cardiff, Birmingham and Glasgow Universities) and an order placed with Bull in December 2014. This was the optimal design, based on the power and floor space available in the Redwood datacentre, to deliver the four-fold performance increase whilst occupying a single cabinet footprint in the datacentre.

The systems were delivered and installed into the Redwood datacentre in the week commencing the 5th January 2015, with a minor outage on 22nd January to integrate this new partition into the Raven cluster management. A few issues were discovered during the testing and these have been documented and will be revisited with Bull – but the basic functionality was operational and would shut down the system in an emergency, the key requirement of these scripts.

The new Haswell partition underwent the formal acceptance tests in February - March 2015, with the expansion now in full production. Although this new queue partition will be prioritised for Gravity Waves consortium member usage, given the burst nature of their computational research, pre-emption will be activated when the system is not in use to enable any parallel jobs with check-pointing capabilities to take advantage of these new resources. This will help free up valuable space on the existing queues, whilst ensuring those jobs can use the most appropriate processor type to maximise efficient usage of the Raven service.

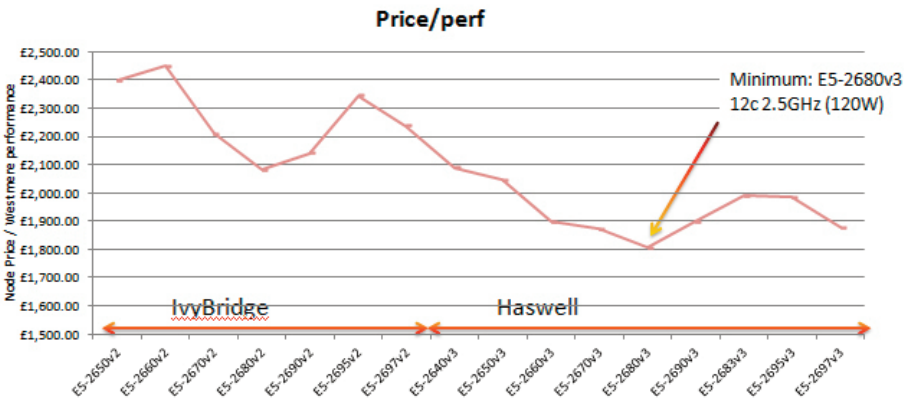


Figure 11 - Gravitational price/performance chart

[1]Note that as part of the consortium, these researchers already have access to these facilities in the USA should this become of interest.

HOSTING EQUIPMENT FOR THE CLIMB CONSORTIUM

A second major service to be hosted in the Redwood Datacentre is the Medical Research Council (MRC) CLIMB consortium system. The Cloud Infrastructure for Microbial Bioinformatics (CLIMB) is an MRC funded project across four sites – Warwick, Swansea, Cardiff and Birmingham Universities. The Cardiff Principal Investigator is Dr Thomas Connor from the School of Biosciences. The project aims to:

- Create a public/private cloud for use by UK academics;
- Develop a set of standardised cloud images that implement key pipelines;
- Establish a storage repository for data that is made available online and within our system, from anywhere ('eduroam for microbial genomics'); and
- Provide access to other databases from within the system.

Developing the requirements for the project, the Principal Investigators scoped out a solution which would be:

- Distributed over the four sites;
- Connected over the Janet academic network;
- Different sizes of virtual machines would be required: personal, standard, large memory, huge memory;
- Must be able to support thousands of virtual machines simultaneously;
- Four petabytes of object storage across four sites;
- 300 terabytes of local high performance storage per site.

A high level schematic diagram of the CLIMB consortium service is shown in Figure 12.

ARCCA and Portfolio Management and IT Services have been involved in the planning and installation stages of the project – with the tender being released (via a framework procurement rather than an open procurement) in Q3/2014. The service has different constraints to those traditionally encountered on an HPC supercomputer – not least the requirements to support a wide variety of operating systems and environments through a virtual infrastructure, with no job schedulers to allocate the computing resources.

Given the distributed nature of the deployment and the cutting-edge design of this solution, there are a number of practical considerations where ARCCA support may be required by this project. The locations of the two Principal Investigators were awarded technical support – with the aim of this providing distributed support to the other sites. This may well work once the system is commissioned and in a stable production state, but during the implementation and design there was a strong recommendation that local technical staff would be required to assist in deploying software and testing networks/performance etc. Furthermore, given the novel approach of this proposal, there is limited expertise across the entire community – so there will be a steep learning curve for the four sites as well as the technology providers.

Pooling all expertise is essential in ensuring this project succeeds, particularly as this requires an understanding of clusters and storage systems not configured for traditional HPC services. The cutting edge approach means the design and performance aspects are not as defined as those for a traditional supercomputer – particularly as some of the configuration considerations need to be flexible to adapt to situations that may arise as the service is tested and developed.

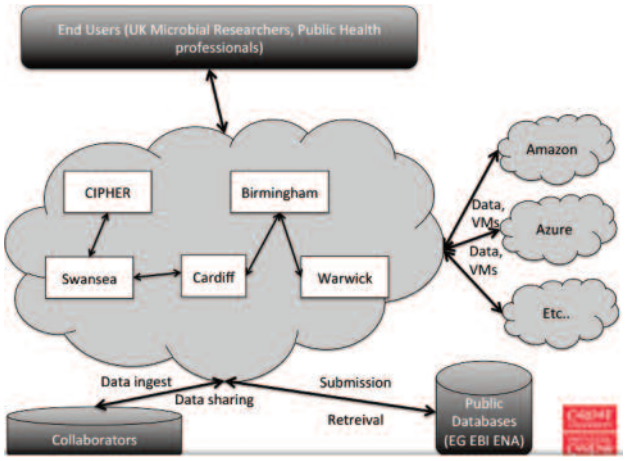


Figure 12 - High-level schematic diagram of the CLIMB consortium service

With the final design in place in late 2014, a preliminary site survey was undertaken in January 2015. To accommodate the increased power, the CLIMB consortium funded the fourth power supply which was installed during the datacentre refurbishment works in December. With the final rack layout agreed, including power and networking designs, the systems were delivered to Cardiff and installed in early March 2015. Cardiff was the first site to take delivery of systems from this procurement (Birmingham went out separately to provide a pilot evaluation set-up to inform the CLIMB tender criteria). With the Cardiff system successfully passing the associated acceptance tests, the service is now operational.

The CLIMB infrastructure will also enable ARCCA to gain greater awareness of new and emerging solutions, helping to define future service directions and aspects of support for diverse community requirements.

LOST VISIONS - RETRIEVING THE VISUAL ELEMENT OF PRINTED BOOKS

'Lost Visions: retrieving the visual element of printed books from the nineteenth century' is one of 21 research projects, funded in 2014 by the Arts and Humanities Research Council, to address the challenges of working with Big Data and making the information more accessible and easier to interpret by a lay audience. Thousands of neglected images dating back to the 18th century will soon be available for the public to search online and find out more about how they have shaped our culture.

Many verbal texts are now stored and delivered by machine, and already search techniques are facilitating research into the meaning, provenance and reception of these texts. However, the illustrations in these texts are frequently omitted and, when included, are often of low quality and without the metadata which are needed for understanding them. There is a risk that tomorrow's readers will be almost unaware of illustration, despite the fact that research over the past century has shown that illustrated texts have qualities, meanings and strategies which are very different from those of un-illustrated texts, and even strikingly different from those of the same verbal texts stripped of their images. Led by Professor Julia Thomas, School of English, Communication and Philosophy, and in collaboration with experts at Cardiff's School of Computer Sciences and Informatics, the project is working on more than a million illustrations from books in the British Library's collection.

The project uses computational methods to identify the visual characteristics of these non-verbal objects, allowing any one object to be assigned multiple characteristics. The available metadata has been incorporated to include full bibliographical details of the book, the exact location and size of the image, and, where possible, a caption or title, and an artist. In addition, tools for identifying the re-use of images will be developed. A crowd-sourcing method was trialled to provide the foundations for future research to describe the content of a small selection of images.

ARCCA supported the technical requirements and provided advice for the successful bid into Arts and Humanities Research Council, and have organised the transfer and copying of the data from the British Library. Providing the computing infrastructure necessary for work on a large dataset, ARCCA's virtual infrastructure has been used to establish a web-based front-end interface to access the databases and undertake the search queries. A searchable database of images with additional metadata is now available to public and academic users.

PILOTOMICS – A BUILDING-BLOCK APPROACH TO DATA STORAGE

The limited provision of large, medium performance storage and the automation of workflows (e.g. genomics pipelines) are inhibiting factors in the use of HPC by communities such as Bioinformatics researchers. In many instances, these researchers simply wish to use HPC as a tool and are not experts in computing techniques. Therefore, the lack of training and associated documentation is also a factor in the limited uptake of advanced research computing by these communities.

Working with Dr. Tom Connor (School of Biosciences), ARCCA and the University's Portfolio Management and IT Services helped to develop a successful bid to the Biosciences Research Infrastructure Fund process to implement a pilot solution, Pilotomics. The project aims to design, develop and implement a pilot, building-block approach to support the data storage, management and compute processes involved in a range of 'Omics' disciplines. The project provides researchers with access to high volumes of data storage on our supercomputer to solve problems from analysing data from gene sequencers, to hosting scanned images for crowd source cataloguing.

This collaborative project involves specifying and purchasing distributed data storage systems, the implementation of control and user interface software, the provision of a long-term archiving solution, and developing training in the use of the pilot system. Although one of the key aims of the project is meeting severe deficiencies that already exist in provision now, we intend to use the formation and operation of the system as a mechanism for defining data use case scenarios within the Health and Life Sciences communities at Cardiff University. This has generated user and system administrator experience that informed the University-wide reviews of data storage requirements, as well as developing and promoting best practice amongst academic staff for data management.

Designing a scalable 'building block' approach to storage supports the University's emerging Big Data requirements. Creating a co-located environment for Big Data and computing resources is essential for a balanced ecosystem to address the large scale data challenges currently facing all major research institutes.

This has been a highly successful pilot service, with solutions being developed for a number of different research groups including:

- Wales Gene Park;
- Illustration Archives (Lost Visions);
- Flexilog – this phrase-searching, ERC-funded project will require processing the entire ClueWeb12 corpus (www.lemurproject.org/clueweb12/specs.php), comprising 733M web pages and taking 27.3TB of space;
- Gravitational Waves;
- Dementia Platform UK;
- Cardiff University Brain Research Imaging Centre (CUBRIC).

In addition, a number of proposals have been produced for a variety of research group grant applications. Due to the success and scale of the service, this solution is being developed in collaboration with colleagues in Portfolio Management and IT Services and will complement that from the forthcoming Research Data and Information Management project.

What's next

PROMOTING UTILISATION - THE GALAXY GATEWAY

Scientific Gateways provide an easier method to access HPC systems. This has been an industry-wide issue across the HPC community, and has resulted in a variety of Gateways which are now becoming popular within their respective communities.

A number of communities within Cardiff University could benefit from Gateways; especially those who might otherwise find the traditional approach to HPC difficult - improving accessibility was also one of the 'top ten' results to come out of the recent ARCCA user survey. The Schools of Biosciences and Medicine in particular are at the early growth stage of using Raven and feedback from that community is that HPC is difficult and requires too much knowledge of unfamiliar technologies such as schedulers and script writing. This has led to deployment of the Galaxy Gateway (http://galaxyproject.org); an open source web-based platform for data intensive biomedical research.

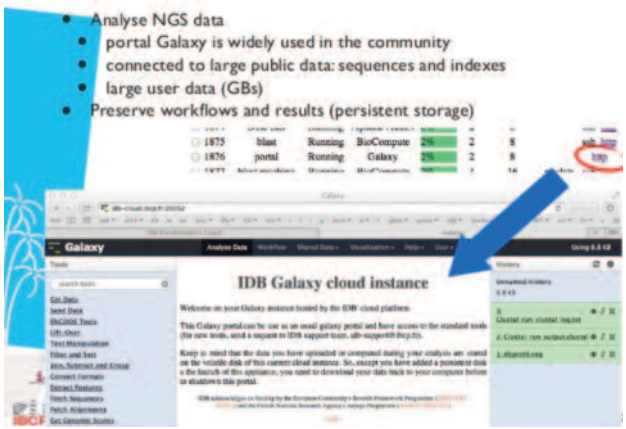
Headlined as 'Online bioinformatics analysis for everyone', Galaxy has emerged as the leading open-source workflow platform for next-generation sequencing data analysis, with many standard processing tools accessible from its web-based user interface. This workflow software is also being applied successfully to proteomics and chemo-informatics.

The installation was driven by the recommendation of the software from within the School of Biosciences; the code is open source and provided for free. Following the initial installation, further tests have been undertaken to make sure it can be interfaced with the large datasets that otherwise cannot be transmitted reliably through a website.

Galaxy is installed on a virtual machine and is based on a mixture of Python, PHP and Bash. The website integrates with University identity management, so that users can protect their job environment with their University username and password. The software also integrates with the PBS pro job scheduler so usage can be monitored through standard processes, but is still to be integrated with our project code system.

The service has been running for a number of months and is awaiting further volunteers from the bioscience and medical community, to further fine-tune this new service.

Example: Galaxy portal for NGS analyses



MANAGING WORKFLOWS – ROSE AND CYLC

Scientific research on an HPC cluster is frequently a multi-stage process, often requiring a number of integrations of a model with differing parameters or initial conditions followed by post-processing, each stage of which must be individually managed.

Cylc (silk), a dependency-based queue scheduler, and Rose, the Met Office-led suite processor, offer a way of simplifying and automating such a series of tasks. This reduces the barrier to new staff and students by allowing the creation of a fully-configurable user interface with built-in help and documentation (see Figure 13) to make task modification as simple as clicking a mouse button.

From the most basic form as a graphical interface and job submission engine, Rose/Cylc can be scaled to manage entire ensembles of tasks, including time and task-based scheduling using Rose's built-in scripting language. An example of the former is a student who needs to quickly become capable of running jobs, producing data for analysis without being obstructed by a complicated and technical configuration system. The latter encompasses the kind of regularly scheduled, massively parallel model runs, followed by automated tape archiving, post-processing and visualisation that the Met Office performs on a daily basis.

Thanks to Rose's graphical suite discovery engine, a job produced by collaborators can be quickly downloaded and modified, with any advantageous changes merged back in under revision control.

ARCCA has been working to implement a set of template Rose/Cylc suites in order to provide practical examples for new users. We currently have suites for the VASP materials modelling package, plus a benchmarking workflow for the Welsh Gene Park available, with further examples pending. With the promise of improved flexibility and simplicity for HPC users, we will provide ongoing development and support for the system.

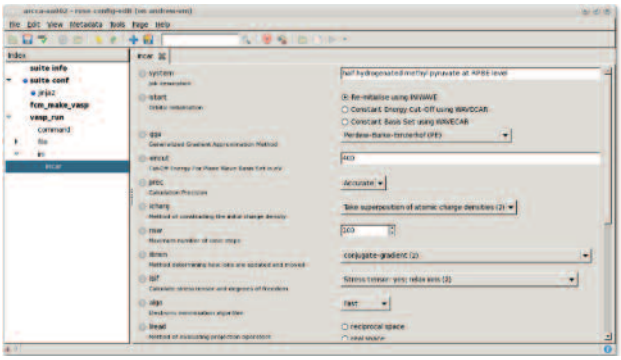


Figure 13 - Rose edit configurable user interface

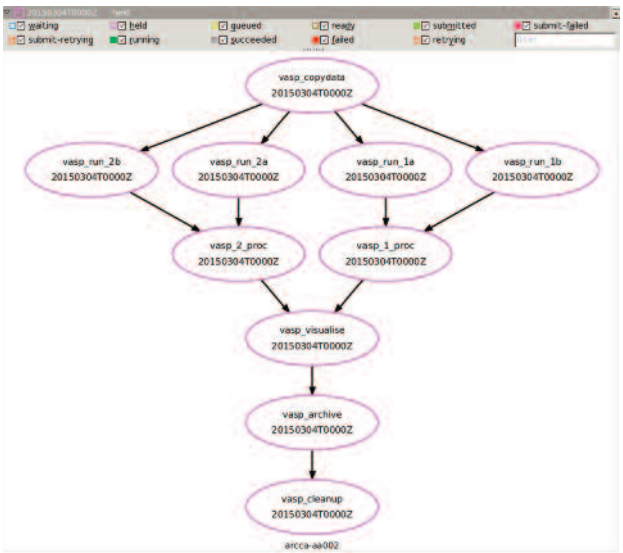
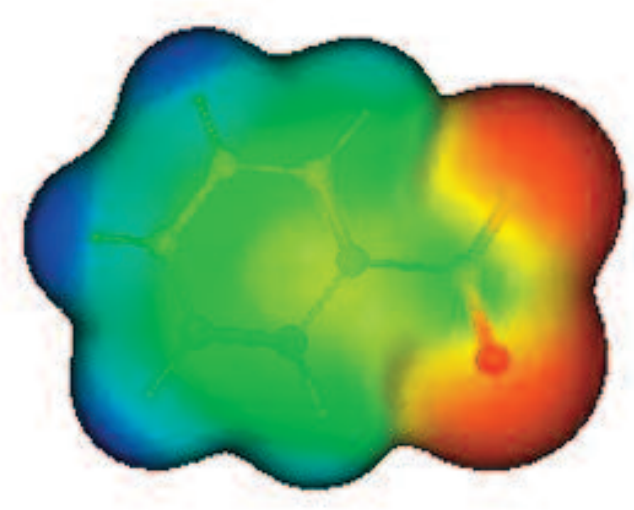


Figure 14 - Dependency graph for a sample workflow

HPC FOR UNDERGRADUATES – WEBMO AND THE GRIDCHEM GATEWAY



Even though the School of Chemistry already utilises a large amount of Raven (around 40 per cent), there is scope to improve the accessibility for non-technical users, such as those within undergraduate learning where the extra cluster utilisation would be minimal but the benefit to teaching would be great.

The School of Chemistry has been using WebMO (www.webmo.org) in their teaching environment for a number of years, and it is seen as a good starting point to learn about computational chemistry. WebMO improves the accessibility and usability of computational chemistry software as users can access and use their software from their web browser. WebMO allows a three dimensional structure to be created using a true 3-D editor, via Java technology. A computational engine is chosen, and job options are selected from dropdown menus. The job is submitted for computation, and the server runs the job. After the job is completed, the output file is parsed and results are presented as 3-D structures, animations, graphs, and tables on a web page.

In addition, WebMO allows for importing existing structures, monitoring the progress of running jobs, job queuing, progress monitoring, and importing previously computed output files for visualization. WebMO also allows many users to take advantage of a single workstation on which the computational chemistry programs are installed. The software is provided by the School and only required an understanding of the installation instructions for subsequent implementation on Raven.

As with most externally facing services within ARCCA, WebMO has been installed on a virtual machine which is connected to the internal Raven network. The Enterprise edition of WebMO interfaces with the underlying PBS pro scheduler and also allows each user of WebMO their own job management environment, protected by their University username and password. The packages that WebMO supports include MolPro (www.molpro.net), Gaussian (www.gaussian.com), GAMESS (www.msg.ameslab.gov/gamess), and VASP (www.vasp.at).

The service is currently undergoing final validation to ensure correct interfacing with our reinstated project code system on Raven.

Currently the plan is to consider another Chemistry package – GridChem (www.gridchem.org) - developed by a number of institutions in the USA, including members of the XSEDE and Teragrid organisations. Due to the long-standing relationship between ARCCA and these organisations, it has been agreed that given the lack of installation instructions for sites outside the USA, XSEDE can provide help. It is envisaged that it might take more time than the currently installed Gateways but that it would be beneficial to learn from a mature community based around GridChem.

ARCCA AND ASSESSING THE CLOUD

As part of an ongoing assessment of the costs of running the existing advanced research computing services directly through ARCCA, against those of outsourcing the comparable services of the Raven cluster through a cloud service, we have recently completed a cost and performance analysis of both options.

Although a contemporary supercomputer centre can be seen to superficially resemble a cloud facility, there are core differences in their intended use and internal architecture. A cloud datacentre is optimized to deliver high quality, low latency access to services used by thousands of concurrent remote users. Supercomputers are optimized to execute small numbers of parallel programs in the shortest possible time. Cloud services are continuously running interactive applications. If the application fails the cloud must restart them automatically so that users do not experience interruptions in service. Supercomputer applications are loaded, execute and terminate and are seldom interactive – such systems are optimised for parallel, closely coupled and/or mixed workload computing (e.g. the simulation of large-scale problems).

Inspection of the resulting costings suggests that replacing Raven with an equivalent cloud computing solution remains considerably more expensive, with some services totalling over three times that of a local datacentre.

As part of this assessment, a project to quantify the relative performance of a variety of widely-used applications from both the physical sciences and life sciences communities is currently in progress with Joel Southgate and Tom Connor from the School of Biosciences. This is looking at the relative application performance found on both "bare-metal" clusters, such as Raven, and virtualised cluster configurations available through cloud-providers such as Microsoft and Google.

Furthermore, for use at an institutional level, guarantees are needed on the performance of such services including those that are not routinely available through external cloud-based provision: the ability for developers to optimise applications for the processor architecture; consistency and response time; the ability to support all HPC problems that ARCCA tackles; system reliability; the ability to access technical support relating to all hardware and middleware used within the cloud; identifying where any data would be physically processed and stored.

ANALYSIS SUMMARY:

- Local HPC costs are currently significantly lower than equivalent hardware in the cloud.
- Cloud HPC solutions are less capable than dedicated hardware for running more complex, highly parallelised codes, but are suitable for less tightly coupled ensemble tasks.
- Remote compute instances distance the data from the point of processing, resulting in significant delays in transferring data.
- Lack of ownership of compute resources potentially introduces limitations as to what can be run and how effectively it can be supported.

We continue to monitor this space with interest as cloud-based HPC services will undoubtedly mature and competition between providers will inevitably drive prices downwards, although it is likely that increased processing capability leads to increased data volume and thus increasing time to transfer data in and out of the cloud.

FUTURE TECHNOLOGY CHALLENGES

As ARCCA looks to replace the Raven system, will this merely be business as usual? We are all used to new machines being relatively simple evolutions of our previous machines. Indeed, the impact of Moore's law, leading to a predictable doubling in performance every 18 months, has introduced a false sense of security. Migrating from Merlin to Raven led to a performance increase by a factor of four or five by merely staying with the latest Intel x86-64 processor and migrating from Intel's Harpertown to Sandy Bridge processor. However, this approach will no longer be true from 2016 onwards.

What's changing? Mainstream multi-core CPUs will continue to evolve, but much more slowly. Intel's x86-64 processor roadmap is well advertised and involves the following evolution:

Intel Sandy Bridge ▶ Intel Ivy Bridge ▶ → Haswell ▶ Broadwell ▶ Skylake CPUs

8 core ▶ 12 core ▶ 18 core ▶ >20 core …

However, the performance per core has effectively ceased. To retain the levels of performance increase we have historically enjoyed, we will have little choice but to adopt radically different architectures.

What are the options? There are a number of viable ways to stay with the previous performance scaling, but all involve major software development efforts. The two principal paths involve many-core CPUs and GP-GPUs; both paths are summarised briefly below:

- **Many-core CPUs**: Most current systems are multi-core. Systems with a large number of processor cores – tens or hundreds – are sometimes referred to as many-core or massively multi-core systems. Intel Many Integrated Core Architecture or Intel MIC is a coprocessor computer architecture developed by Intel incorporating earlier work on the Larrabee many-core architecture. Belonging to the Intel Xeon Phi family, Knights Landing (KNL), the next generation of Intel MIC architecture, is scheduled for Q1 2016. The KNL processor is characterised by O(70) cores - other many-core CPUs are expected to emerge, notably those based on the ARM[2] architecture.

- **GP-GPUs:** GPU-accelerated computing is the use of a graphics processing unit (GPU) together with a CPU to accelerate applications. The market leader here is Nvidia, whose next generation Pascal processor is due for release in 2016. Pascal is to feature in the new systems to come out of the collaboration of Oak Ridge, Argonne, and Livermore (CORAL), featuring GPUs from a close collaboration between IBM, Nvidia and Mellanox. With the addition of key technologies from Mellanox and Nvidia, specifically the latter's NVLINK technology, the new generation – 'Power9' – GPU-accelerated systems will deliver an architecture that can be condensed to a much smaller number of nodes for far higher performance, and a much larger shared memory footprint.

Other technology challenges are also on the horizon, specifically (i) deeper memory hierarchies, with Stacking (HBM), non-volatile memories etc., (ii) integrated interconnects e.g. Intel Omni-Path fabric, and (iii) FPGAs, historically little used in HPC, but now more attractive given the support for OpenCL.

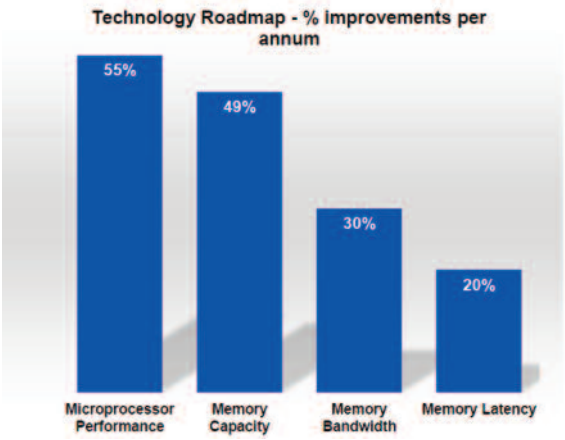


Figure 15 - Technology trends

The graphic above captures the markedly different rates of performance improvements associated with the underlying technology. Our traditional approach to software development has been to ignore these differences; the key challenge ahead is to reflect these differing rates in future software developments.

There is no escaping the fact that a 'business as usual' approach to scientific software development will result in our being left in the slow lane, with developers now faced with the challenging issue of developing performant code on increasingly complex architectures. Cardiff will certainly need a major injection of Computational Science and Engineering/Applications expertise; expertise that already resides in ARCCA and expertise that will prove invaluable in addressing the challenges highlighted above.

[2]ARM is a family of instruction set architectures for computer processors based on a reduced instruction set computing (RISC) architecture developed by British company ARM Holdings.

Outreach and collaborative frameworks

THE GW4 HPC LANDSCAPE SURVEY

As part of the GW4 collaboration, the associated HPC Working Group secured funding to assess the computational research facilities across the four partner institutions. This was led by Dr. David Acreman (Exeter University) who published the resulting GW4 landscape survey report in December 2014. The intention of the survey was to collate information on the current HPC facilities, areas of research supported, software provision and training to help identify possible areas of collaboration (driven by research drivers) and possible complementary support that could be shared amongst the institutions.

Research computing underpins many areas of research throughout GW4, including the most computationally demanding calculations, with around 1,200 researchers across GW4 using campus-level HPC provision. There are different areas of strength within the GW4 partner universities but there are also shared research interests, including computational chemistry and solid state physics, and climate science and meteorology. Throughout GW4 there are increasing requirements from researchers outside the traditional HPC disciplines, particularly life sciences.

All the GW4 partners run campus-level HPC systems similar to those found in other research intensive universities. The systems use similar technology but with some variation in the availability of high memory nodes (frequently required by life sciences) and novel technologies, such as coprocessors (e.g. GPUs). Some researchers make use of larger facilities to run workloads which would not be possible on a campus level system. These larger facilities include EPSRC regional centres, ARCHER (the national HPC facility) and PRACE systems (which are the largest in Europe).

By way of conclusion, the survey provided a series of recommendations captured under the headings of 'Funding opportunities', 'Skills', 'Training and expertise', and 'Resource sharing'.

There was a strong feeling that the four institutions should bid together in response to a number of major funding opportunities in the future, agreeing in principle which large projects would be applied for under the auspices of GW4, and preparing outline bids in anticipation of these opportunities. Examples included a possible bid against EPSRC's next call for regional centres, likely to arise after the forthcoming comprehensive spending review, and to consider a possible bid to host systems for DiRAC-3, the facility that provides HPC services for the UK Theoretical Particle Physics and Theoretical Astrophysics Community. An initial step in bidding for such future funding opportunities would involve maintaining a list of administrative and academic contacts.

Each of the GW4 HPC sites had developed its own programme of skills, training and expertise, and there was agreement that these activities would benefit from a consolidation of these efforts. The report recommended a number of steps to take this forward. First, each GW4 partner should nominate a training contact to progress this consolidation. The group of contacts should first establish a GW4 training web page to collate sharable training resources and publicise their availability within their own institution. Additional recommended steps involved a review of training provision across GW4 with any gaps identified, taking into account the needs of Doctoral Training Centres and Partnerships. Identified gaps in provision should be investigated through external provision where appropriate.

Finally, a number of possible activities and resources across the GW4 HPC centres were identified that might feature in a programme of resource sharing. Considering technology provision, the report identified a clear requirement for establishing a GW4 technology test-bed for evaluating novel hardware and accelerator technologies e.g., GP-GPUs, Intel Phi etc. The possibilities of a targeted approach to VAT expenditure was identified with a recommendation to establish a cost sharing group to enable VAT-free resource sharing. The final area for sharing involved the provision for community-wide access to the HPC centres across GW4, facilitated for example by common mechanisms for cross-site authentication. This could then be used by service providers for benchmarking and testing and by users for access to large memory nodes, and would facilitate general research collaboration.

USK CODE CLUB VISITS THE REDWOOD DATACENTRE

The Redwood datacentre welcomed a group of students from Usk Church in Wales Primary School's Code Club.

Code Club is a nationwide network of free volunteer-led after-school coding clubs for children aged nine - eleven.

20 members of Usk's Code Club and Digital Leaders toured the HPC Wales and ARCCA supercomputers housed in the Redwood datacentre, and spoke with members of the University's technical team. The group also saw for themselves the difference in size between the components that make up a desktop computer and a supercomputer.

The event also offered the children an opportunity to see the variety of projects that can be delivered by supercomputing technology, from research and testing for personalised medicine to aiding the development of the world's first 1,000 mph car.

Dr Christine Kitchen, ARCCA Associate Director, said: "It was great to meet the next generation of computer programmers and scientists from the Code Club and Digital Leaders group, and to talk to them about the coding that lies behind the scenes of the technology they see and use every day."

"I hope we have helped encourage them to continue to experiment and 'play' with technology, and inspired them to consider technology as a future career."

Prof Martyn Guest, ARCCA and HPC Wales Technical Director continued: "We are experiencing a lack of STEM graduates in the workplace, and it is important to encourage children's natural curiosity about technology early on in their education in order to rectify this for future generations."

GLOSSARY

ARC – Advanced Research Computing. The use of computing and data resources for research, which are beyond the capabilities of the average desktop or laptop computer (see HEC).
Archer – UK's National Supercomputer Facility [Cray XC30 at Edinburgh Parallel Computer Centre (EPCC), University of Edinburgh].
CPU – Central Processing Unit or processor.
EPSRC – Engineering and Physical Sciences Research Council.
GPU – Graphical Processing Units (also GP-GPU – General Purpose Graphical Processing Unit).
GW4 – An alliance which combines the intellectual capacity and physical resources of the four leading research-intensive universities in the South West of England and Wales: Bath, Bristol, Cardiff and Exeter.
Haswell – The codename for the fifth generation Intel® Core™ processor microarchitecture developed by Intel as the successor to the Ivy Bridge microarchitecture.
HEC – High-End Computing. Captures the use of 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. The use of (parallel) supercomputers and computer clusters, computing systems made of multiple (usually mass-produced) processors linked together in a single system with commercially available interconnects.
LIGO – The LIGO Scientific Collaboration is a group of researchers who have joined together in the search for gravitational waves: www.ligo.org/about.php.
MATLAB – A numerical computing environment and fourth generation programming language. Maintained by The MathWorks, MATLAB allows easy matrix manipulation, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs in other languages.
Merlin – The first ARCCA supercomputer (a 2048 quad-core Intel Harpertown cluster with 20.12 Tflops peak performance and 164 MFlops/Watt efficiency).
MPI – Message Parsing Interface is a programming technique which uses software libraries to turn serial applications into parallel ones which can run on distributed memory systems. This is one of the standard paradigms for parallel programming in both C, C++ and Fortran languages.
NERC – Natural Environment Research Council.
NGS – Next-generation sequencing, also known as high-throughput sequencing, is the catch-all term used to describe a number of different modern sequencing technologies, including Illumina (Solexa) sequencing, Roche 454 sequencing, Ion torrent: Proton/PGM sequencing and SOLiD sequencing.
Nvidia CUDA – Compute Unified Device Architecture is a compiler and set of development tools that enable programmers to use a variation of C to code for execution on a Graphical Processing Unit (GPU).
OpenMP – Open Multi-Processing is an application programming interface that supports multiplatform shared memory multiprocessing (SMP) programming in C, C++ and Fortran.
PBS – The Portable Batch System job scheduler from Altair Engineering.
PUE - Power Usage Effectiveness, the ratio of total amount of energy used by a computer datacentre facility to the energy delivered to computing equipment. PUE was developed by a consortium called The Green Grid. PUE is the inverse of datacentre infrastructure efficiency (DCIE).
Raven - The current ARCCA supercomputer, a hybrid cluster comprising three separate partitions, with 864 cores of Intel Westmere, 2048 core of Intel Sandy Bridge, and 1440 cores of Intel Haswell processor technology, with a combined peak performance of over 100 Tflop/s.
Robinhood – Robinhood Policy Engine is a versatile tool to manage contents of large file systems.
SMP – Symmetric Multiprocessors: a computer hardware architecture which distributes the computing load over a small number of identical processors, which share memory.
TCO – Total cost of ownership (TCO) is a financial estimate intended to help buyers and owners determine the direct and indirect costs of a product or system.
XSEDE – Extreme Science and Engineering Discovery Environment (www.xsede.org).

Delivering on Cardiff University's research agenda

ACCELERATING RESEARCH THROUGH HPC AND BIG DATA

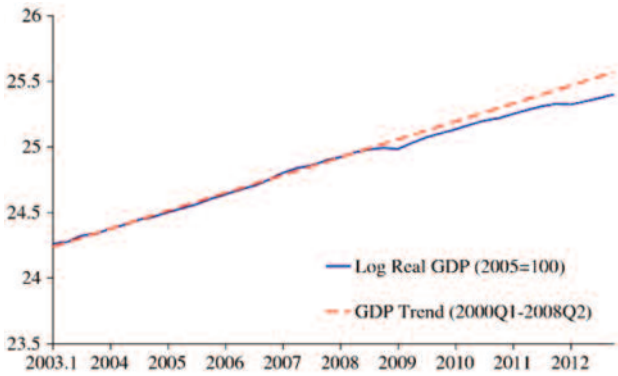
ARCCA provides the computational infrastructure to power current and potential future areas of research that promise to advance the frontiers of science and engineering. The next section details some of the key research which is accelerated by ARCCA; contributions have been provided by many of those who are capitalizing on the availability of Raven to address a variety of scientific or engineering research questions.

Many of the reports are from those areas that have a high potential for a strategically important impact in the short-term - Astronomy and Astrophysics, Chemistry and Materials, Computer Science, Engineering and Mathematics. Consideration is also given to the 'emerging' HPC research disciplines, with contributions from the Schools of Medicine, Biosciences, Planning and Geography, and the Cardiff Business School. A total of 20 contributions from researchers from many of the University's Academic Schools provides compelling evidence of the variety of research now underway using the spectrum of HPC technology and support resources now available for Cardiff researchers.

COLLEGE OF ARTS, HUMANITIES AND SOCIAL SCIENCES

1. **Banking and the Macroeconomy in China: A Banking Crisis Deferred?**
   Le, Vo Phuong Mai; Matthews, Kent Gerard Patrick; Meenagh, David; Minford, Anthony Patrick Leslie; and Xiao, Zhiguo 2014. Open Economies Review 25 (1), pp. 123-161. Doi.10.1007/s11079-013-9301-9

The downturn in the world economy following the global banking crisis has left the Chinese economy relatively unscathed. This paper develops a model of the Chinese economy using a DSGE framework with a banking sector to shed light on this episode. It differs from other applications in the use of indirect inference procedure to test the fitted model. The model finds that the main shocks hitting China in the crisis were international and that domestic banking shocks were unimportant. However, directed bank lending and direct government spending was used to supplement monetary policy to aggressively offset shocks to demand. The model finds that government expenditure feedback reduces the frequency of a business cycle crisis but that any feedback effect on investment creates excess capacity and instability in output.

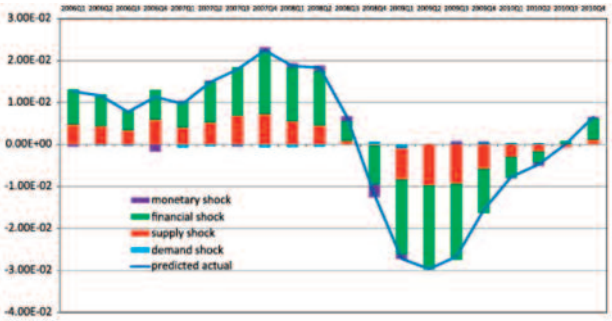


China real GDP per capita and pre-crisis trend

2. **How important is the credit channel? an empirical study of the US banking crisis**
   Liu, Chunping; and Minford, Anthony Patrick Leslie 2014.
   Journal of Banking & Finance 41, pp. 119-134.
   Doi.10.1016/j.jbankfin.2013.12.017

We examine whether by adding a credit channel to the standard New Keynesian model we can account better for the behaviour of US macroeconomic data up to and including the banking crisis. We use the method of indirect inference, which evaluates statistically how far a model's simulated behaviour mimics the behaviour of the data.

We find that the model with credit dominates the standard model by a substantial margin. Credit shocks are the main contributor to the variation in the output gap during the crisis.



Shock decomposition for output during crisis period

3. **Exploring the ripple effect and spatial volatility in house prices in England and Wales: regressing interaction domain cross-correlations against reactive statistics**
   Cooper, Crispin; Orford, Scott; Webster, Christopher John; and Jones, Christopher Bernard 2013.
   Environment and Planning B: Planning and Design 40 (5), pp. 763-782.
   Doi.10.1068/b37062

In this study we conduct an exploratory spatiotemporal analysis of recent UK housing-market data 2000–06, fine grained both in space and in time. We present firstly the exploratory technique itself, and secondly an overview of patterns found in the dataset.

A broadly scoped metamodel is created which points towards fruitful avenues for further investigation and understanding of the driving forces behind price propagation in the housing market. At the core of this model is an 8850 × 8850 cross-correlation matrix representing price linkage between different areas, which is analyzed by a custom regression engine.

Hence this is the first study to unify both local and regional house-price modelling, and the first to look for a quantitative explanation of the structure of any house-price interaction matrix on a large scale.

Findings corroborate existing research on the ripple effect and spatial volatility, and point towards mechanisms through which these might arise, as a combination of local interaction and market segmentation broken by lifecycle migration patterns.

COLLEGE OF BIOMEDICAL AND LIFE SCIENCES

4. Analysis of genome-wide association studies of Alzheimer disease and of Parkinson disease to determine if these two diseases share a common genetic risk

Moskvina, Valentina; Harold, Denise; Russo, Giancarlo; Vedernikov, Alexey; Sharma, Manu; Saad, Mohamad; Holmans, Peter Alan; Bras, Jose M.; Bettella, Francesco; Keller, Margaux F.; Nicolaou, Nayia; Simón-Sánchez, Javier; Gibbs, J. Raphael; Schulte, Claudia; Durr, Alexandra; Guerreiro, Rita; Hernandez, Dena; Brice, Alexis; Stefánsson, Hreinn; Majamaa, Kari; Gasser, Thomas; Heutink, Peter; Wood, Peter; Martinez, Maria; Singleton, Andrew B.; Nalls, Michael A.; Hardy, John; Owen, Michael John; O'Donovan, Michael Conlon; Williams, Julie; Morris, Huw; and Williams, Nigel Melville 2013.
JAMA Neurology 70 (10), pp. 1268-1276. Doi.10.1001/jamaneurol.2013.448

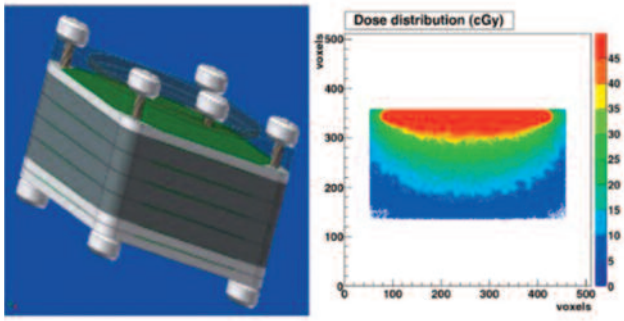
Importance Despite Alzheimer disease (AD) and Parkinson disease (PD) being clinically distinct entities, there is a possibility of a pathological overlap, with some genome-wide association (GWA) studies suggesting that the 2 diseases represent a biological continuum. The application of GWA studies to idiopathic forms of AD and PD have identified a number of loci that contain genetic variants that increase the risk of these disorders.
Objective To assess the genetic overlap between PD and AD by testing for the presence of potentially pleiotropic loci in 2 recent GWA studies of PD and AD.
Design Combined GWA analysis. Setting Data sets from the United Kingdom, Germany, France, and the United States.
Participants Thousands of patients with AD or PD and their controls.
Main Outcomes and Measures Meta-analysis of GWA studies of AD and PD.
Methods To identify evidence for potentially pleiotropic alleles that increased the risk for both PD and AD, we performed a combined PD-AD meta-analysis and compared the results with those obtained in the primary GWA studies. We also tested for a net effect of potentially polygenic alleles that were shared by both disorders by performing a polygenic score analysis. Finally, we also performed a gene-based association analysis that was aimed at detecting genes that harbor multiple disease-causing single-nucleotide polymorphisms, some of which confer a risk of PD and some a risk of AD.
Results Detailed interrogation of the single-nucleotide polymorphism, polygenic, and gene-based analyses resulted in no significant evidence that supported the presence of loci that increase the risk of both PD and AD.
Conclusions and Relevance Our findings therefore imply that loci that increase the risk of both PD and AD are not widespread and that the pathological overlap could instead be "downstream" of the primary susceptibility genes that increase the risk of each disease.

5. Development and validation of RAYDOSE: a Geant4-based application for molecular radiotherapy

Marcatili, S.; Pettinato, C.; Daniels, Stephen; Lewis, G.; Edwards, Peter; Fanti, S.; and Spezi, Emiliano 2013.
Physics in Medicine and Biology 58 (8), pp. 2491-2508. Doi.10.1088/0031-9155/58/8/2491

We developed and validated a Monte-Carlo-based application (RAYDOSE) to generate patient-specific 3D dose maps on the basis of pre-treatment imaging studies. A CT DICOM image is used to model patient geometry, while repeated PET scans are employed to assess radionuclide kinetics and distribution at the voxel level. In this work, we describe the structure of this application and present the tests performed to validate it against reference data and experiments. We used the spheres of a NEMA phantom to calculate S values and total doses.

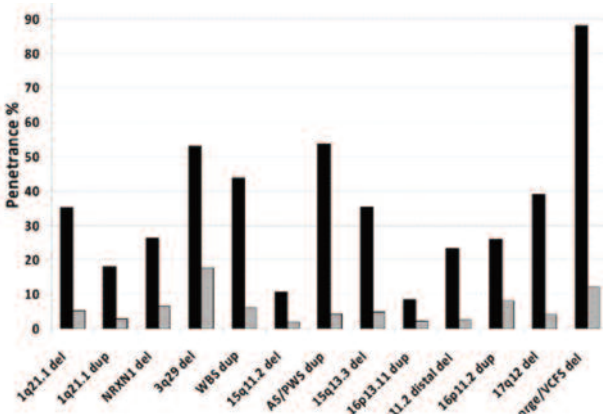
The comparison with reference data from OLINDA/EXM showed an agreement within 2% for a sphere size above 2.8 cm diameter. A custom heterogeneous phantom composed of several layers of Perspex and lung equivalent material was used to compare TLD measurements of gamma radiation from 131I to Monte Carlo simulations. An agreement within 5% was found. RAYDOSE has been validated against reference data and experimental measurements and can be a useful multi-modality platform for treatment planning and research in MRT.



Schematic design of the multi-layer phantom (a) and corresponding simulated dose distribution (b).

6. **The penetrance of copy number variations for schizophrenia and developmental delay**
   Kirov, George; Rees, Elliott; Walters, James Tynan Rhys; Escott-Price, Valentina; Georgieva, Lyudmila; Richards, Alexander; Chambert, Kimberly D.; Davies, Gerwyn; Legge, Sophie E.; Moran, Jennifer L.; McCarroll, Steven A.; O'Donovan, Michael Conlon; and Owen, Michael John 2013.
   Biological Psychiatry 75 (5), pp. 378-385.Doi. 10.1016/j.biopsych.2013.07.022

Background: Several recurrent copy number variants (CNVs) have been shown to increase the risk of developing schizophrenia (SCZ), developmental delay (DD), autism spectrum disorders (ASD), and various congenital malformations (CM). Their penetrance for SCZ has been estimated to be modest. However, comparisons between their penetrance for SCZ or DD/ASD/CM, or estimates of the total penetrance for any of these disorders have not yet been made.
Methods: We use data from the largest available studies on SCZ and DD/ASD/CM, including a new sample of 6882 cases and 6316 controls, to estimate the frequencies of 70 implicated CNVs in carriers with these disorders, healthy control subjects, and the general population. On the basis of these frequencies, we estimate their penetrance. We also estimate the strength of the selection pressure against CNVs and correlate this against their overall penetrance.
Results: The rates of nearly all CNVs are higher in DD/ASD/CM compared with SCZ. The penetrance of CNVs is at least several times higher for the development of a disorder from the group of DD/ASD/CM. The overall penetrance of SCZ-associated CNVs for developing any disorder is high, ranging between 10.6% and 100%.
Conclusions: CNVs associated with SCZ have high pathogenicity. The majority of the increased risk conferred by CNVs is toward the development of an earlier-onset disorder, such as DD/ASD/CM, rather than SCZ. The penetrance of CNVs correlates strongly with their selection coefficients. The improved estimates of penetrance will provide crucial information for genetic counselling.



Penetrance of schizophrenia-associated copy number variants for schizophrenia (grey) and the combined group of developmental delay, autism spectrum disorders, and various congenital malformations (black).

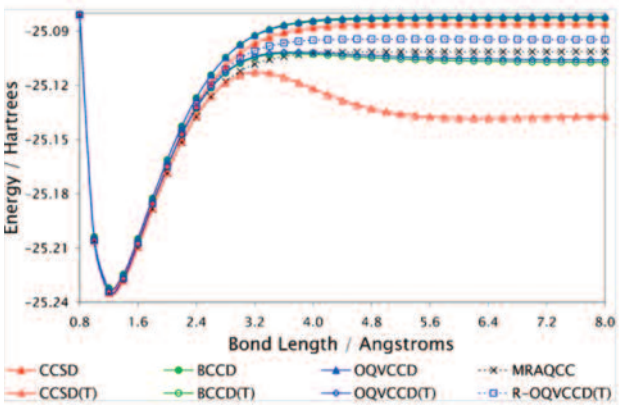
7. **De novo mutations in schizophrenia implicate synaptic networks**
Fromer, Menachem; Pocklington, Andrew; Kavanagh, David; Williams, Hywel John; Dwyer, Sarah; Gormley, Padhraig; Georgieva, Lyudmila; Rees, Elliott; Palta, Priit; Ruderfer, Douglas; Carrera, Noa; Humphreys, Isla; Johnson, Jessica S.; Roussos, Panos; Barker, Douglas D.; Banks, Eric; Milanova, Vihra; Grant, Seth G.; Hannon, Eilis; Rose, Samuel A.; Chambert, Kimberly; Mahajan, Milind; Scolnick, Edward M.; Moran, Jennifer L.; Kirov, George; Palotie, Aarno; McCarroll, Steven A.; Holmans, Peter; Sklar, Pamela; Owen, Michael John; Purcell, Shaun M.; and O'Donovan, Michael Conlon 2014.
Nature 506 (7487), pp. 179-184.
Doi.10.1038/nature12929

Inherited alleles account for most of the genetic risk for schizophrenia. However, new (de novo) mutations, in the form of large chromosomal copy number changes, occur in a small fraction of cases and disproportionally disrupt genes encoding postsynaptic proteins. Here we show that small de novo mutations, affecting one or a few nucleotides, are overrepresented among glutamatergic postsynaptic proteins comprising activity-regulated cytoskeleton-associated protein (ARC) and N-methyl-d-aspartate receptor (NMDAR) complexes. Mutations are additionally enriched in proteins that interact with these complexes to modulate synaptic strength, namely proteins regulating actin filament dynamics and those whose messenger RNAs are targets of fragile X mental retardation protein (FMRP). Genes affected by mutations in schizophrenia overlap those mutated in autism and intellectual disability, as do mutation-enriched synaptic pathways. Aligning our findings with a parallel case–control study, we demonstrate reproducible insights into aetiological mechanisms for schizophrenia and reveal pathophysiology shared with other neurodevelopmental disorders.

COLLEGE OF PHYSICAL SCIENCES AND ENGINEERING

8. **Rigorously extensive orbital-invariant renormalized perturbative triples corrections from quasi-variational coupled cluster theory**
Robinson, James Baden; and Knowles, Peter James 2013.
Journal of Chemical Physics, 138 (7), 074104.
Doi.10.1063/1.4791636

We show that, by making use of the linked tensor objects inherent to the approach, Orbital-optimised Quasi-Variational Coupled Cluster Theory (OQVCCD) leads naturally to a computationally-trivial, rigorously extensive, and orbital-invariant renormalization of the standard (T) correction for the perturbative inclusion of the effects of connected triple excitations. The resulting prototype method, renormalized perturbative triple OQVCCD (R-OQVCCD(T)), is demonstrated to predict potential energy curves for single bond-breaking processes of significantly higher accuracy than OQVCCD with the standard perturbative triple-excitation correction (OQVCCD(T)) itself, and to be in good numerical correspondence with the existing renormalized (R-CCSD(T)) and completely renormalized (CR-CCSD(T)) coupled-cluster singles doubles triples methods, while continuing to provide descriptions of multiple bond-breaking processes of OQVCCD(T) quality.

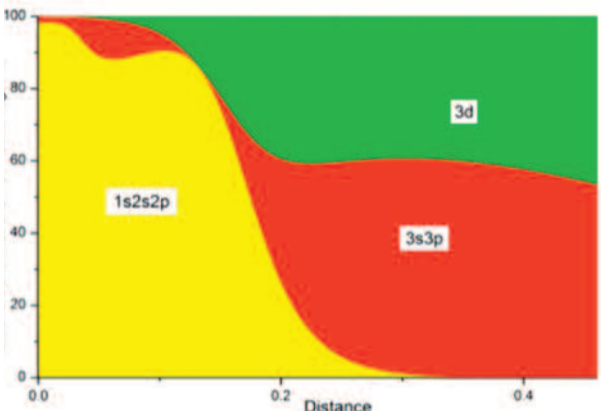


Calculated potential energy curves for the BH molecule with the aug-cc-pVQZ basis set.

9. Electron localisation in Ga-Heterocyclic compounds

Platts, James Alexis; Thomsen, Maja K.; and Overgaard, Jacob 2013.
Zeitschrift für anorganische und allgemeine Chemie 639 (11) , pp. 1979-1984.
Doi.10.1002/zaac.201200498

The present study shows that a successful description of the electron density in a Ga-containing compound using the present multipole model requires separation and independent radial scaling of the three Ga core shells. It is also shown that two similar compounds only differing in the oxidation state of the Ga atom provides similar residual density maps and thus the deficiencies in the standard multipole model are very similar in the two cases. Nevertheless, we find two fundamentally different core modified multipole models. It is found that introduction of the modified core parameters in the modelling of the experimental data has no significant impact. An analysis of the ELI-D is introduced to study the localization of electrons in the Ga-guanidinate moiety. There is a clear analogy between the Laplacian and the ELI-D, however the latter is able to reveal details from the total density which is not available in the Laplacian. Only in an analysis of the valence density alone can the Laplacian reveal the finer details.



Relative orbital contributions to the atomic density of Ga as a function of the distance from the nucleus.

10. Impact of co-adsorbed oxygen on crotonaldehyde adsorption over gold nanoclusters: a computational study

Zeinalipour-Yazdi, Constantinos; Willock, David James; Machado, Andreia; Wilson, Karen; and Lee, Adam Fraser 2014.
Physical Chemistry Chemical Physics 16 (23), pp. 11202-11210. Doi.10.1039/c3cp53691b

Crotonaldehyde (2-butenal) adsorption over gold sub-nanometer particles, and the influence of co-adsorbed oxygen, has been systematically investigated by computational methods. Using density functional theory, the adsorption energetics of crotonaldehyde on bare and oxidised gold clusters (Au_{13}, d = 0.8 nm) were determined as a function of oxygen coverage and coordination geometry. At low oxygen coverage, sites are available for which crotonaldehyde adsorption is enhanced relative to bare Au clusters by 10 kJ mol^{-1}. At higher oxygen coverage, crotonaldehyde is forced to adsorb in close proximity to oxygen weakening adsorption by up to 60 kJ mol^{-1} relative to bare Au. Bonding geometries, density of states plots and Bader analysis, are used to elucidate crotonaldehyde bonding to gold nanoparticles in terms of partial electron transfer from Au to crotonaldehyde, and note that donation to gold from crotonaldehyde also becomes significant following metal oxidation. At high oxygen coverage we find that all molecular adsorption sites have a neighbouring, destabilising, oxygen adatom so that despite enhanced donation, crotonaldehyde adsorption is always weakened by steric interactions. For a larger cluster (Au_{38}, d = 1.1 nm) crotonaldehyde adsorption is destabilized in this way even at a low oxygen coverage. These findings provide a quantitative framework to underpin the experimentally observed influence of oxygen on the selective oxidation of crotyl alcohol to crotonaldehyde over gold and gold–palladium alloys.

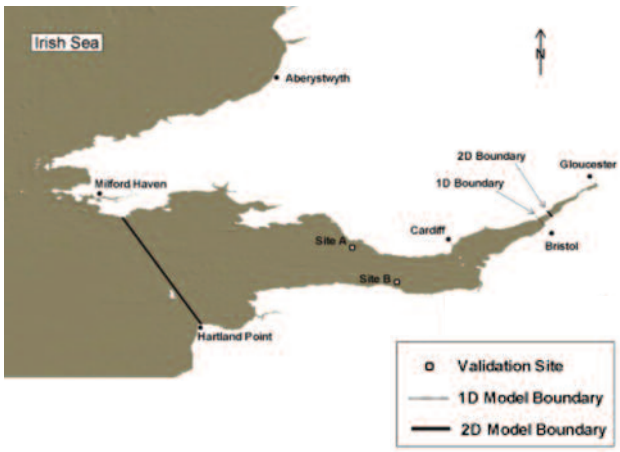
.

11. Cloud computing for the Architecture, Engineering and Construction sector: Requirements, prototype and experience
Beach, Thomas; Rana, Omer Farooq; Rezgui, Yacine; and Parashar, Manish 2013. Journal of Cloud Computing: Advances, Systems and Applications 2 (4), 8. Doi.10.1186/2192-113X-2-8

The Architecture, Engineering and Construction (AEC) sector is a highly fragmented, data intensive, project based industry, involving a number of very different professions and organisations. Projects carried out within this sector involve collaboration between various people, using a variety of different systems. This, along with the industry's strong data sharing and processing requirements, means that the management of building data is complex and challenging. This paper presents a solution to data sharing requirements of the AEC sector by utilising Cloud Computing. Our solution presents two key contributions, first a governance model for building data, based on extensive research and industry consultation. Second, a prototype implementation of this governance model, utilising the CometCloud autonomic cloud computing engine based on the Master/Work paradigm. We have integrated our prototype with the 3D modelling software Google Sketchup. The approach and prototype presented has applicability in a number of other eScience related applications involving multi-disciplinary, collaborative working using Cloud computing infrastructure.

12. Comparison of hydro-environmental impacts for ebb-only and two-way generation for a Severn Barrage
Ahmadian, Reza; Falconer, Roger Alexander; and Bockelmann-Evans, Bettina Nicole 2014. Computers & Geosciences 71, pp. 11-19. Doi.10.1016/j.cageo.2014.05.006

Marine renewable energy is playing an increasing significant role in many parts of the world, mainly due to a rise in the awareness of climate change, and its detrimental effects, and the increasing cost of natural resources. The Severn Estuary, located between South West England and South Wales, has a tidal range of up to 14 m which makes it the second highest tidal range in the world. There are a number of barrage proposals amongst various marine renewable energy schemes proposed to be built in the estuary. The Cardiff-Weston STPG (Severn Tidal Power Group) Barrage, which would be one of the world's largest tidal renewable energy schemes if built, is one of the most publicised schemes to-date. This barrage would generate about 17 TWh/annum of power, which is approximately 5% of the UK's electricity consumption, whilst causing significant hydro-environmental and ecological impact on the estuary. This study mainly focuses on investigating the hydro-environmental impacts of the STPG barrage for the option of two-way generation, and compares this with the commonly investigated option of ebb-only generation. The impacts of the barrage were modelled by implementing a linked 1-D/2-D hydro-environmental model, with the capability of modelling several key environmental processes. The model predictions show that the hydro-environmental impacts of the barrage on the Severn Estuary and Bristol Channel, such as changes in the maximum velocity and reduction in suspended sediment and bacteria levels, were less significant for the two-way generation scheme when compared with the corresponding impacts for ebb-only generation.

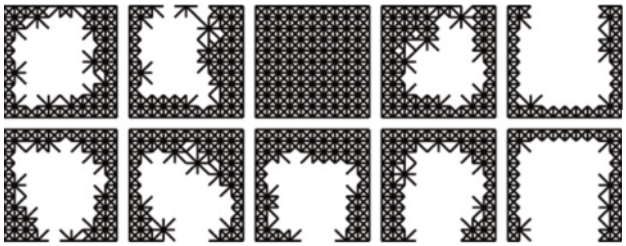


The model domain extent and validation sites.
Site A: Southerndown site, B: Minehead site.

13. A partitioned model order reduction approach to rationalise computational expenses in nonlinear fracture mechanics

Kerfriden, Pierre; Goury, Olivier; Rabczuk, Timon; and Bordas, Stephane Pierre Alain 2013. Computer Methods in Applied Mechanics and Engineering 256, pp. 169-188. Doi.10.1016/j.cma.2012.12.004

We propose in this paper an adaptive reduced order modelling technique based on domain partitioning for parametric problems of fracture. We show that coupling domain decomposition and projection-based model order reduction permits a focus on the numerical effort where it is most needed: around the zones where damage propagates. No a *priori* knowledge of the damage pattern is required, the extraction of the corresponding spatial regions being based solely on algebra. The efficiency of the proposed approach is demonstrated numerically with an example relevant to engineering fracture.

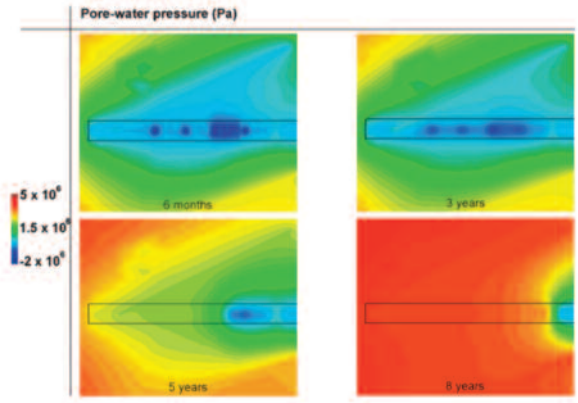


Example of a reduced integration domain. Subdomain 6 is not reduced.

14. Three-dimensional behaviour of a prototype radioactive waste repository in fractured granitic rock

Thomas, Hywel Rhys; Vardon, Philip James; and Cleall, Peter John 2014. Canadian Geotechnical Journal 51 (3), pp. 246-259. Doi.10.1139/cgj-2013-0094

An investigation of the three-dimensional coupled thermohydromechanical behaviour of a prototype repository in fractured granitic rock is presented. The pre-placement behaviour of the repository is first considered, making use of a full three-dimensional simulation of the problem. An effective continuum approach, augmented with discrete features with a high hydraulic conductivity, is employed. The method adopted is found to be able to simulate accurately the highly anisotropic flow regime observed at the pre-placement phase. The major features of the full repository experiment under applied heating conditions were then successfully simulated. The range of buffer hydration rates, the thermal response of the repository, and the associated mechanical response were successfully simulated. Approaches to capture both the transient microstructural behaviour of the compacted bentonite (MX-80 type) and a MX-80 pellet material are incorporated. The repository behaviour was shown to be strongly influenced by complex coupled processes, including interactions between the system components. The adoption of a three-dimensional modelling approach proved to be essential to realistically simulate the behaviour of a repository incorporating anisotropic flow behaviour. Finally, potential impacts of the processes considered on performance of the barrier system and in safety assessment are considered.

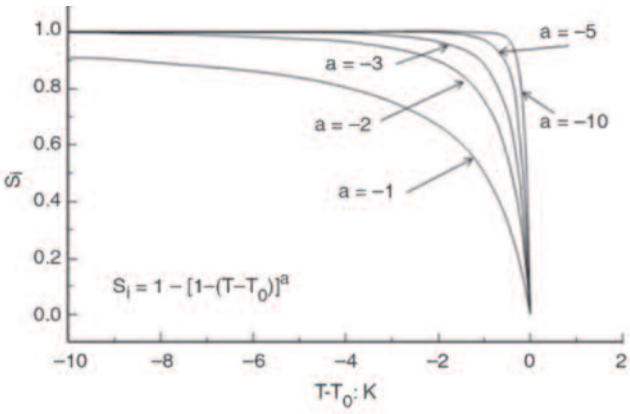


Contour plots of pore-water pressure (Pa) variation in the simulation in a horizontal cross-section intersecting the tunnel.

15. Freezing soil effects on earth-contact heat transfer

Rees, Stephen William; Cleall, Peter John; Li, Y.-C.; and Shao, Kangjie 2013.
Building Services Engineering Research and Technology 34 (3), pp. 259-274.
Doi.10.1177/0143624412441616

The inclusion of soil freezing and snow cover within the context of a building energy simulation is explored. In particular, a method of including soil freezing within the simulation of heat flow from a building to the neighbouring foundation soils is considered. Non-linear thermal conductivity and heat capacity relations are explored that account for the effect of soil freezing. In addition, the work also considers latent heat generated by phase change that occurs as the soil water temperature reduces and ice forms. A simple approach to represent the insulating effect that snow cover may have on the net heat flow at the ground surface is also provided. The approach is illustrated by application to the simulation of a full-scale ground heat transfer experiment performed by others. The results provide a first indication of the potential significance of the inclusion of ground freezing within the context of modelling heat transfer from a full-scale monitored building. Overall transient temperature variations are shown to be dependent on ice content and latent heat effects. Non-linear, ice-content dependent, thermal conductivity and heat capacity are included in the work. Good correlation between measured and simulated temperature variations has been achieved. **Practical application**: The principle application relates to the assessment of heat transfer from buildings through to the underlying foundation soils. In addition, the work is of direct relevance to a wider range of applications, for example, ground source heat pumps, thermo-active piles and foundations. The model lends itself to application with respect to utilisation of ground energy within the context of pavement design, particularly, in relation to the alleviation of winter ice problems. It also applies within the context of assessment of urban heat island phenomenon and heat and moisture transfer beneath freezer foundations.



Relationship between degrees of moisture in the ice phase and temperature.

16. Mixed finite element methods for groundwater flow in heterogeneous aquifers

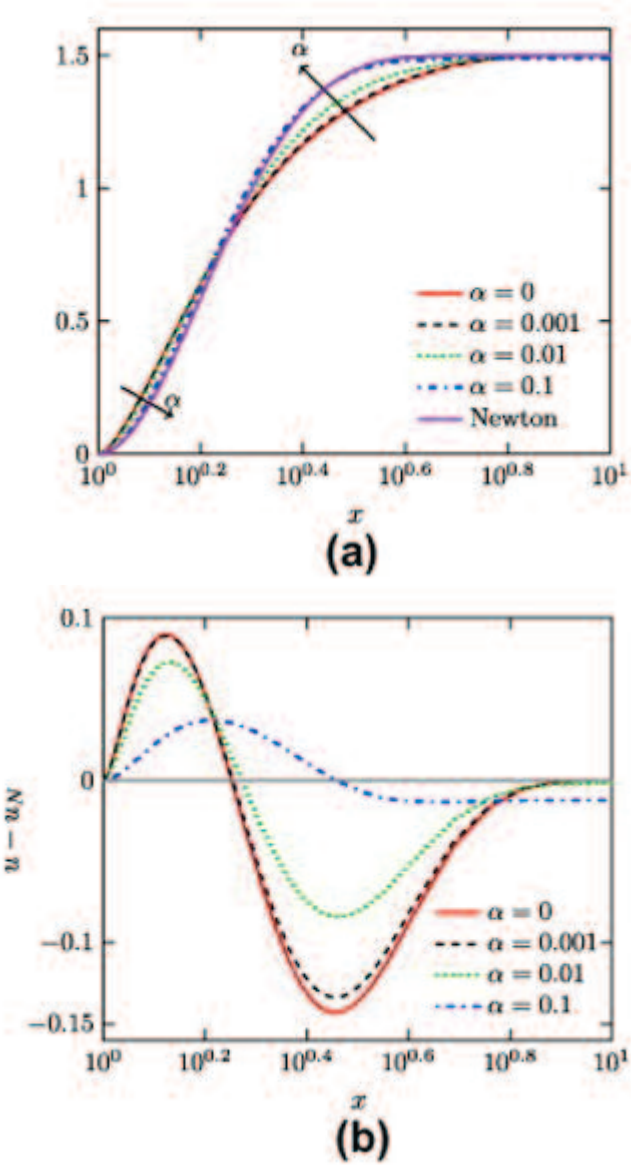
Traverso, L.; Phillips, Timothy Nigel; and Yang, Yuesuo 2013.
Computers and Fluids 88, pp. 60-80.
Doi.10.1016/j.compfluid.2013.08.018

This paper is concerned with a comparison of the performance and efficiency of mixed finite element methods for solving single phase fluid flow in porous media. Particular attention is given to the accurate determination of the groundwater fluxes. The linear systems generated by the mixed finite element method (MFEM) are indefinite. Symmetric positive definite linear systems are generated through the introduction of Lagrange multipliers giving rise to the mixed hybrid finite element method (MHFEM). The convergence behaviour of the numerical approximations is investigated over a range of conductivity coefficients from heterogeneous, isotropic and diagonal to discontinuous, anisotropic and full, on both triangular and quadrilateral, structured and distorted meshes. The robustness and efficiency of various preconditioned solvers is investigated in terms of optimality with respect to both mesh size and conductivity coefficient.

17. Viscoelastic flow around a confined cylinder using spectral/hp element methods

Claus, Susanne; and Phillips, Timothy Nigel
2013.
Journal of Non-Newtonian Fluid Mechanics
200, pp. 131-146.
Doi.10.1016/j.jnnfm.2013.03.004

The benchmark problem of flow of a viscoelastic fluid around a confined cylinder is considered. The governing equations are discretised using spectral/hp element methods. These allow the spatial and temporal variations in the solution that are characteristic of viscoelastic flows, to be resolved accurately and efficiently. A decoupled approach is employed in which the conservation equations are solved for velocity and pressure and the constitutive equation (Oldroyd-B and Giesekus) are solved for the polymeric component of the extra-stress tensor. The computations are stabilized using the DEVSS-G/DG formulation of the problem. Excellent agreement with the literature is achieved for the drag coefficient in the case of an Oldroyd-B fluid. Causes for the breakdown in numerical convergence with mesh refinement beyond some critical value of the Weissenberg number are explored. The high resolution property of spectral/hp approximations has enabled an instability that develops in the shear layer on the cylinder and is convected downstream to be identified. The onset of this instability is shown to occur at the critical value of the Weissenberg number predicted by the theory of Dou and Phan-Thien. The influence of the Reynolds number and, for the Giesekus model, the mobility parameter on the drag coefficient is also investigated and discussed.

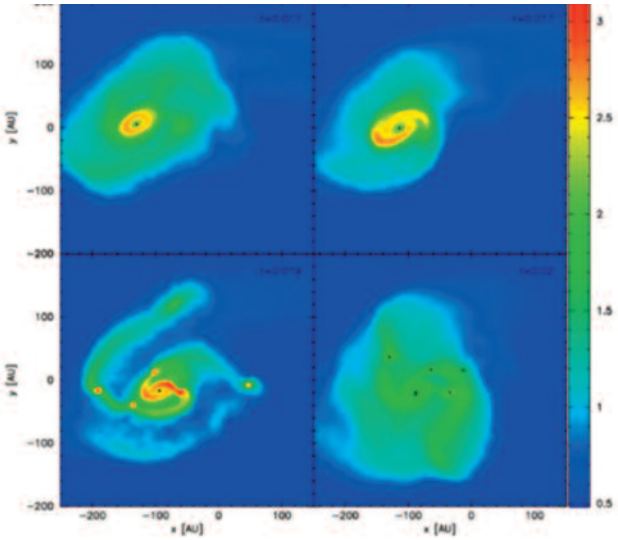


Dependence of (a) U and (b) U-Un, the velocity shift with respect to the Newtonian velocity profile, on Wi along the downstream centreline.

18. Simulating star formation in Ophiuchus

Lomax, Oliver David; Whitworth, Anthony Peter; Hubber, D. A.; Stamatellos, Dimitris; and Walch, Stephanie K. 2014.
Monthly Notices of the Royal Astronomical Society 439 (3), pp. 3039-3050.
Doi.10.1093/mnras/stu177

We have simulated star formation in pre-stellar cores, using smoothed particle hydrodynamics and initial conditions informed by observations of the cores in Ophiuchus. Because the observations are limited to two spatial dimensions plus radial velocity, we cannot infer initial conditions for the collapse of a particular core. However, with a minimum of assumptions (isotropic turbulence with a power-law spectrum, a thermal mix of compressive and solenoidal modes, a critical Bonnor–Ebert density profile) we can generate initial conditions that match, in a statistical sense, the distributions of mass, projected size and aspect ratio, thermal and non-thermal one-dimensional velocity dispersion, observed in Ophiuchus. The time between core–core collisions in Ophiuchus is sufficiently long, that we can simulate single cores evolving in isolation, and therefore we are able to resolve masses well below the opacity limit. We generate an ensemble of 100 cores, and evolve them with no radiative feedback from the stars formed, then with continuous radiative feedback and finally with episodic radiative feedback. With no feedback the simulations produce too many brown dwarfs, and with continuous feedback too few. With episodic radiative feedback, both the peak of the protostellar mass function (at ~0.2 $M_\odot$) and the ratio of H-burning stars to brown dwarfs are consistent with observations. The mass of a star is not strongly related to the mass of the core in which it forms. Low-mass cores ($M_{core} \sim 0.1\ M_\odot$) tend to collapse into single objects, whereas high-mass cores ($M_{core} \gtrsim 1\ M_\odot$) usually fragment into several objects with a broad mass range.



Column density plots from simulation 002_ERF showing disc fragmentation.

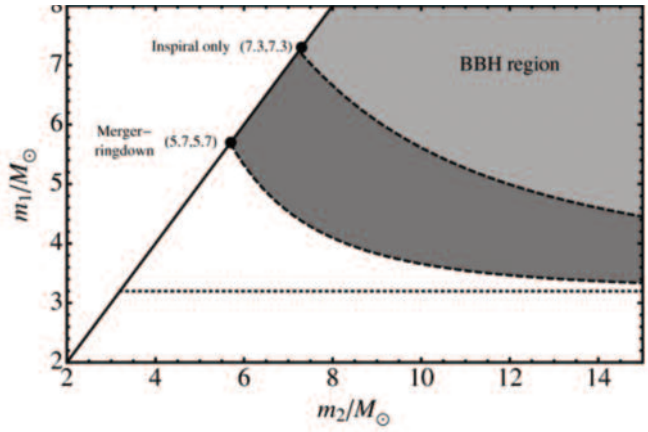
19. When can gravitational-wave observations distinguish between black holes and neutron stars?

Hannam, Mark; Brown, Duncan A.; Fairhurst, Stephen; Fryer, Chris L.; and Harry, Ian W. 2013. Astrophysical Journal Letters 766 (1), L14. Doi.10.1088/2041-8205/766/1/L14

Gravitational-wave observations of compact binaries have the potential to uncover the distribution of masses and spins of black holes and neutron stars in the universe. The binary components' physical parameters can be inferred from their effect on the phasing of the gravitational-wave signal, but a partial degeneracy between the components' mass ratio and their spins limits our ability to measure the individual component masses. At the typical signal amplitudes expected by the Advanced Laser Interferometer Gravitational-wave Observatory (signal-to-noise ratios between 10 and 20), we show that it will in many cases be difficult to distinguish whether the components are neutron stars or black holes.

We identify when the masses of the binary components could be unambiguously measured outside the range of current observations: a system with a chirp mass $\mathcal{M} \leq 0.871\ M_\odot$ would unambiguously contain the smallest-mass neutron star observed, and a system with $\mathcal{M} \geq 2.786\ M_\odot$ must contain a black hole. However, additional information would be needed to distinguish between a binary containing two $1.35 M_\odot$ neutron stars and an exotic neutron-star–black-hole binary.

We also identify those configurations that could be unambiguously identified as black hole binaries, and show how the observation of an electromagnetic counterpart to a neutron-star–black-hole binary could be used to constrain the black hole spin.

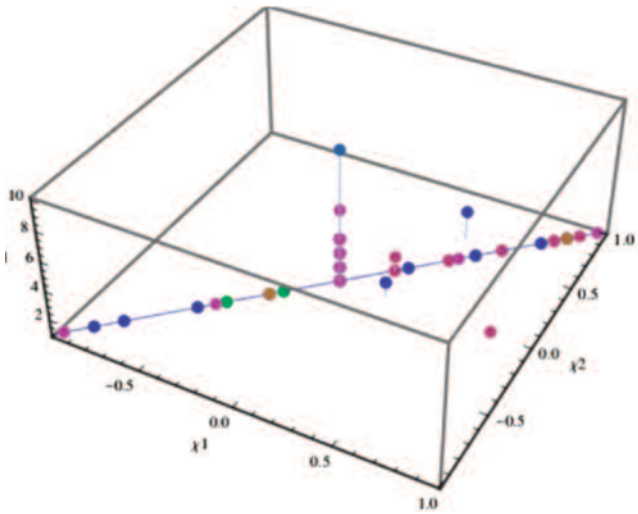


Shaded region indicates binaries would be unambiguously identified as BBH systems, assuming a maximum NS mass of 3.2 solar masses (dotted line).

20. The NINJA-2 project: detecting and characterizing gravitational waveforms modelled using numerical binary black hole simulations

J. Aasi (LIGO Lab., Caltech) et al, 2014. Classical and Quantum Gravity 31 (11), 115004. Doi. 10.1088/0264-9381/31/11/115004

The Numerical INJection Analysis (NINJA) project is a collaborative effort between members of the numerical relativity and gravitational-wave (GW) astrophysics communities. The purpose of NINJA is to study the ability to detect GWs emitted from merging binary black holes (BBH) and recover their parameters with next-generation GW observatories. We report here on the results of the second NINJA project, NINJA-2, which employs 60 complete BBH hybrid waveforms consisting of a numerical portion modelling the late inspiral, merger, and ringdown stitched to a post-Newtonian portion modelling the early inspiral. In a 'blind injection challenge' similar to that conducted in recent Laser Interferometer Gravitational Wave Observatory (LIGO) and Virgo science runs, we added seven hybrid waveforms to two months of data recoloured to predictions of Advanced LIGO (aLIGO) and Advanced Virgo (AdV) sensitivity curves during their first observing runs. The resulting data was analysed by GW detection algorithms and 6 of the waveforms were recovered with false alarm rates smaller than 1 in a thousand years. Parameter-estimation algorithms were run on each of these waveforms to explore the ability to constrain the masses, component angular momenta and sky position of these waveforms. We find that the strong degeneracy between the mass ratio and the BHs' angular momenta will make it difficult to precisely estimate these parameters with aLIGO and AdV. We also perform a large-scale Monte Carlo study to assess the ability to recover each of the 60 hybrid waveforms with early aLIGO and AdV sensitivity curves. Our results predict that early aLIGO and AdV will have a volume-weighted average sensitive distance of 300 Mpc (1 Gpc) for $10M_\odot + 10M_\odot$ ($50M_\odot + 50M_\odot$) BBH coalescences. We demonstrate that neglecting the component angular momenta in the waveform models used in matched-filtering will result in a reduction in sensitivity for systems with large component angular momenta. This reduction is estimated to be up to ~15% for $50M_\odot + 50M_\odot$ BBH coalescences with almost maximal angular momenta aligned with the orbit when using early aLIGO and AdV sensitivity curves.



Mass ratio and dimensionless spins of the NINJA-2 hybrid waveform submissions.

CONTACT US:

Advanced Research Computing @ Cardiff
Redwood Building
King Edward VII Avenue
Cardiff
CF10 3NB

Telephone: +44 (0)29 208 75581

Email: arcca@cardiff.ac.uk