The newsletter of EPCC, the supercomputing centre at the University of Edinburgh

epcc news

Issue 93 SUMMER 2023

UK Exascale comes a big step closer

The publication of the Future of Compute Review report brings our goal of a UK Exascale service within reach.

Image: Sefa Ozel via Getty Images

From our Deputy Director

Welcome to the summer 2023 issue of EPCC News. This is my first as Deputy Director and I want to express my enthusiasm for the diversity and distinctiveness of the work being done right now at EPCC. I've come here from the Data-Driven Innovation initiative and have worked with EPCC in that context for a number of years.

These are exciting times. In its 2023 Spring Budget the UK Government committed to delivering a key recommendation of the Future of Compute Review, to invest to build an Exascale supercomputer. This is a major milestone, but it is just a milestone, and there is a significant journey ahead to convert it into reality.

This issue of EPCC News illustrates the diversity of EPCC's activities, from the opportunities coming in the (near) future - from the UK Government's Future of Compute Review through the ExCALIBUR projects and quantum computing, to what we are doing right now - with an overview of what it means to host and operate the ARCHER2 service, and updates from the Software Sustainability Institute. There is a lot of ground covered.

Further highlights include our PhD student placements and their experiences at EPCC, celebration of our EPCC MSc students as they prepare to compete in the Student Cluster Competition at ISC High Performance 2023 in May, and a spotlight on two projects: the first in support of the 'Peace and Conflict Resolution Evidence Platform', creating an online platform for global peace process data analytics; the second supporting children with autism, utilising ML-based facial emotion recognition to provide real-time support, working on improving the accuracy of the detection.

I hope you enjoy this issue and, as ever, feedback is always welcome.

Ritchie Somerville
EPCC Deputy Director
r.somerville@epcc.ed.ac.uk

Contents

www.epcc.ed.ac.uk
info@epcc.ed.ac.uk

EPCC is a supercomputing centre based at The University of Edinburgh, which is a charitable body registered in Scotland with registration number SC005336.

UK Exascale comes a big step closer

In June 2022, the world's first officially recognised Exascale supercomputer was announced at the International Supercomputing Conference in Germany. Named "Frontier", the supercomputer is an HPE Cray Ex system, built on the same underlying platform as ARCHER2 but with different processors.

To win the prize, Frontier demonstrated sustained numerical performance of 1.102 Exaflop/s while consuming around 21MW of electrical power. To get this performance within reasonable power and cost budgets, the system is built from a combination of traditional CPUs and GPU accelerators. Each node (or server) in the system has one CPU (an AMD Trento processor) and four GPUs (AMD's MI250X accelerators) with the GPUs providing a huge boost in numerical performance.

As a comparison, ARCHER2 has 5,860 nodes which deliver 19.5 Petaflop/s of performance. Frontier has 9,472 nodes which deliver 1.102 Exaflop/s – making each node 35 times more powerful than ARCHER2 but using 10 times less power per Flop. We can now multiply two 64-bit floating point numbers using 19 picojoules – truly remarkable energy efficiency.

This is a tremendous achievement for the world of supercomputing and hopefully something we'll be able to demonstrate in the UK shortly. For more than a decade EPCC has been working to provide an Exascale supercomputing service for the UK's scientific and industrial research communities. In March 2023, this goal moved much closer thanks to the publication of the Future of Compute Review report and subsequently the Government's Spring Budget. The report builds upon the work by Sir Patrick Vallance's team at the Government Office of Science and recommends that "the Government should commit to the path to Exascale, via the adoption of a phased approach. This would deliver Exascale-ready public capability immediately, adding further hardware that increases the capability to full Exascale by 2026".

The Spring Budget then included the commitment that "in line with two of the key recommendations of the Future of Compute Review, the government will invest, subject to the usual business case processes, in the region of £900 million to build an Exascale supercomputer and to establish a new AI Research Resource, with initial investments starting this year". We are now in that approvals process with the expectation of an initial system in 2024, growing to full Exascale by 2026.

EPCC, with the full support of the University of Edinburgh, has been preparing our ACF data centre for Exascale for the past six years. To date we have invested £20m in Computer Room 4 (currently configured as a 6MW room) and a further £9m in a new 30MW power supply to site dedicated to the new room. We are ready, today, to host the initial system. With the addition of further cooling and power distribution capability we will be ready to install the full Exascale system in 2026.

Mark Parsons
EPCC Director
m.parsons@epcc.ed.ac.uk

The Government's announcements are really important for UK science and industry. They recognise the importance of large-scale computing to drive innovation through advanced modelling and simulation and also the ability to train the largest AI models. The UK will be back where it belongs – at the forefront of computational science and innovation.

HPC services at EPCC
https://edin.ac/hpc-services

Supercomputing: the next wave of opportunity for UK industry

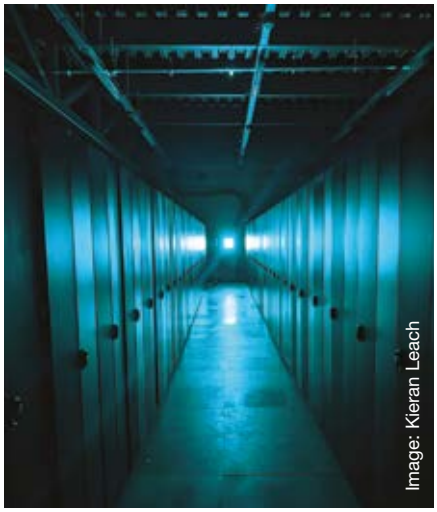


Image: Kieran Leach

In 2020 UKRI published the science case for UK supercomputing. This report set out why investment was required in supercomputing in the UK and sought to demonstrate the breadth of research opportunity that existed to reinforce the UK's scientific research potential/capability and the opportunity to further enhance its productivity.

Seven key research areas were identified in the report:

- Fundamental sciences.
- Engineering and materials.
- Climate, weather, and earth sciences.
- Digital humanities and social sciences.
- Mathematics and science of computation.
- Computational biology.
- Computational biomedicine.

What was less obvious from the 94-page document was the industry case woven within it.

UK industry is already involved in many collaborations with UK academic users of ARCHER2, however it currently accounts for only 5% of the 3,000 users. As the use of supercomputing evolves, we want to expand the amount of industry usage of future systems. There is the potential over the next decade to grow this four-fold; from use by our largest multinationals to our most innovative SMEs (Small and Medium Enterprise), all undertaking research and development to enhance products and services on these systems.

To do this it will be necessary to work across all industry sectors to increase the use of supercomputing whilst also fostering even greater collaboration between industry and academia. This will require dedicated programmes of engagement and development.

The Edinburgh and South East Scotland Science and Innovation Audit published in 2016 foresaw the implications at a regional and national level that accelerating market demand for data-driven innovation would bring, raising challenging new requirements for scientific research and innovation in a broad range of industry sectors, and an ever-growing need for technological skills to support this.

As we work with UKRI and the UK Government on the opportunities for the next generation of supercomputers, we do so with the desire to not only continue to support the UK's position as a global leader in scientific research and innovation, but also to expand the use of these systems to address the social and business challenges of "UK industry Plc".

UK Research & Innovation: Science case for UK Supercomputing: https://edin.ac/3UAeabS

The Edinburgh and South East Scotland Science and Innovation Audit: Enabling a World-Leading Regional Digital Economy through Data Driven Innovation: https://edin.ac/3o3C3fP

Ritchie Somerville, EPCC
r.somerville@epcc.ed.ac.uk

EPCC has worked at the forefront of computing for over 30 years, and we have helped over one thousand industrial organisations. Our key offerings for business include hosting, research, and software and data services.

To discuss any of our services for industry, please contact our Commercial Manager, Julien Sindt: j.sindt@epcc.ed.ac.uk

We can make HPC work for your business

Image: xbrchx/Getty Images

Increasingly easy access to powerful computational resources raises an important question: is it enough to make a code run more quickly by using more resources, or should we consider performance too?

The question is particularly relevant now because optimised code is more energy efficient and less environmentally and financially costly. There is a view that many computational problems can be solved quickly by simply throwing more compute resources at them but, while this approach usually works fine on "personal" systems, it can fail catastrophically on large machines. Quite often running at scale reveals problems that were not visible on smaller systems.

EPCC has been porting codes to high-performance computing (HPC) systems since 1990, and can solve the problem of code performance reducing at scale.

With our code review service, we will work with your development team to understand how your programme performs, pinpoint where improvements can be made to improve high-scale performance, drive efficient code development, and help you understand the limits of your code's performance.

We will work closely with your team in a transparent fashion, and teach the skills required to continue this upkeep in future. This review process ensures you get the most bang for your computational buck, and your code will be ready for the future.

EPCC also offers expertise in domains other than performance. A recent example is our work with eVineyard who contacted us through the EU Hubs4Data programme. eVineyard has developed an app to assist growers in managing their vines. It is looking to include predictive machine learning solutions with its app and has turned to EPCC for expert advice on development and implementation best practices.

In addition to our research and software services, we continue to host a large and varied collection of HPC systems at our Advanced Computing Facility. This includes traditional HPC machines like ARCHER2 (the UK's national supercomputing service, with 750,000 CPU cores), and Cirrus (a hybrid CPU-GPU system with 10,000 cores and 144 Tesla V100 GPUs). We also increasingly offer a variety of cutting-edge systems, such as NextGenIO (a high-memory system optimised for IO throughput, with 34 48-core nodes, each with 3TB of memory) and the CS-2 (a system optimised to improve the speed of deep-learning training).

EPCC provides access to these systems, along with the training required to use them effectively.

Julien Sindt, EPCC
j.sindt@epcc.ed.ac.uk

EPCC has access to some of the world's leading experts in quantum computing (QC) through our Quantum Applications group, and its affiliation with the Edinburgh Quantum Software Lab. We are working hard to understand when QC will be a useful alternative to traditional computing and how it might work efficiently in combination with classical systems. We are always happy to share our understanding and expertise in QC to help you make the best decisions for your business.

To learn more, please contact me at: j.sindt@ed.ac.uk.

Accelerating data-driven discovery


Image: Kieran Leach

Working with our users, we continue to add services to the Edinburgh International Data Facility (EIDF).

Kostas Kavoussanakis, EPCC
k.kavoussanakis@epcc.ed.ac.uk

EIDF Container Service

In November 2022, EIDF opened a Container Service to early adopters in the University of Edinburgh's School of Informatics. Driven by their requirements, the service provides a containerised application platform enhanced with GPU accelerators. In summer 2023, as the service becomes more mature, we will open it to a wider audience across the University and the EIDF user communities.

The service uses NVIDIA A100 40GB GPU accelerators to support Kubernetes users in running GPU-centric workloads. It has exposed GPUs in two forms: one allows multiple isolated workloads to share a GPU by splitting it into smaller distinct units; and the other allows a full GPU to be used by a container. The service currently allows up to 4 full GPUs to be attached to a single Kubernetes pod. This container platform allows users to experiment with GPU technology, and to develop small and large workflows across multiple pods.

Other service developments

The new Container Service will complement the Data Science Cloud, the Ultra2 HPE SuperDome Flex large shared-memory system, and the Cerebras CS-2 AI accelerator.

The Data Science Cloud was recently expanded to over 8,000 virtual CPUs with over 51 TB of RAM, almost doubling in size. From July 2022-March 2023 we initiated 35 projects on the Data Science Cloud, supporting the diverse research of our ever-growing user community.

We are also working to further extend the storage available on the Data Science Cloud and, crucially, to share storage with the existing Ultra2, Cirrus, and ARCHER2 services. We are working with our partner StackHPC to further develop our Ceph-based storage offerings and to expose an S3 service through it.

These two important milestones in our underlying storage capability will underpin the forthcoming Data Archive and Analytics-Ready Data capabilities. We are working on data ingest automation to host large, shared datasets that can be accessed and shared by Data Science Cloud users, also involving work on the corresponding metadata catalogue that will underpin the ability to discover the datasets to be made available through the EIDF. This work is being developed in conjunction with our colleagues at the Schools of Geosciences, Law, and Engineering, whose input is shaping our service.

TRE improvements

We have started the transition to our new Trusted Research Environment (TRE), which will allow us to automate, thus accelerate and improve, many of our standard operating procedures and to then further expand the offering. A GPU cluster will be delivered soon after existing tenants have transitioned.

Cerebras CS-2 cluster

In the coming months we expect to launch an updated Cerebras CS-2 cluster, bringing our two existing standalone CS-2 systems together. This will allow us to work with, and train, models exceeding 1 billion parameters in the most efficient manner possible.

EIDF website
https://edin.ac/edinburgh-international-data-facility

Building a peace process data analytics platform

Image: The Real Estreya via Flickr.com

PeaceRep
Peace and Conflict Resolution Evidence Platform

EPCC has been providing computing and data engineering services for the Peace and Conflict Resolution Evidence Platform (PeaceRep), in support of the project to build an online platform for peace process data analytics. This cooperation is a prime example of the potential for EPCC and the Edinburgh International Data Facility (EIDF) to support data-driven social research.

Based at the University of Edinburgh Law School, the PeaceRep team is researching how peace processes and transitions from armed conflict can be managed and supported, and how more inclusive post-conflict politics and greater opportunities for international development can be fostered.

EPCC is helping the PeaceRep team build a comprehensive system for combining data from a variety of sources, and in a variety of formats, to build interactive web applications that visualise and present peace process data.

PeaceRep's objective is to be the global leader in peace analytics and digital innovation in the field of peace studies. At the core of this endeavour is the data collection taking place for the PA-X Peace Agreement Database and Dataset, which is analysed and presented in parallel with data on other institutional developments (such as constitution-making, or elections), and data on a variety of socio-economic indicators provided by academic and official sources.

The cooperation with EPCC is aimed at building a lasting data infrastructure to store the diverse data collected across the consortium, and in turn streamlining the usage of this data in a range of applications, including interactive peace process trackers that provide an integrated picture of post-conflict transitions in particular countries. The EPCC team is constructing a PeaceRep data warehouse in EIDF and establishing data ingest workflows for harvesting data from a wide variety of sources. They are also developing reproducible methods for enhancing dataflows in a Power BI service to consume – and publish reports based on – the data stored in the warehouse.

These online interactive peace process trackers that PeaceRep and EPCC aim to deliver in 2023 and 2024 are intended to serve multiple purposes: for academic users, they provide a tool for discovery and hypothesis-making; for policy users, they are providing context and history to a problem they may be coming to for the first time as well as an assessment of how transitions from conflict are developing. They are also a means of relying on innovative technology developed by the University of Edinburgh to support policy-making that is adaptive, evidence-based, and in support of global development.

Sanja Badanjak, School of Law, University of Edinburgh
sanja.badanjak@ed.ac.uk

Transitions from conflict often include economic, political, institutional, and military transitions, with the data relevant to PeaceRep coming from various levels – from the global to surveys of citizens. The processing and analysis component of this work produces another challenge, as the textual data in particular is summarised using natural language processing and machine learning methodologies.

PeaceRep is a seven-year research consortium led by the University of Edinburgh Law School.
https://edin.ac/3KJOxBg

EIDF is hosted and managed by EPCC.
https://edin.ac/edinburgh-international-data-facility

Personalised education of children with autism: utilising facial emotion recognition to provide real-time support

Autism Spectrum Disorder (ASD) describes a range of life-long neurodevelopmental conditions which are usually characterised by impairments in three core domains: social interaction, social communication, and social imagination. This is known as the triad of impairments. The prevalence of ASD in the UK is estimated to be more than one in a hundred children.

Research has shown that children with ASD can benefit from educational interventions. In particular, technology-based interventions are very promising as children with ASD often have an affinity with technology. This may be because technology brings a series of advantages for these children, for example it can act an interface between individuals with ASD and other people, and can diminish anxiety by creating emotional and social distancing. Technology also allows such interventions to be customised to the child's particular needs and interests.

The autistic population is extremely heterogenous, hence children with ASD have highly varying learning abilities. While some have normal language and above average intelligence, other autistic children are completely nonverbal and present severe intellectual disability. Consequently, these children require personalised one-to-one guidance to learn both life-long skills and academic skills.

Training adequate number of teachers and assistants to be specialised in providing this type of guidance is challenging to say the least. Therefore we are focusing on building a personalised Intelligent Tutoring System that utilises state-of-the-art machine learning and artificial intelligence techniques and tools.

Currently there is on-going research at EPCC and the School of Informatics to use Children's Facial Emotion Recognition (FER) to identify when and how to, in real time, dynamically adjust the support provided to children with ASD when using a software-based tutoring system.

Our research has so far revealed that:

- The current FER techniques (including ones based on machine learning techniques) are not accurate enough to recognise children's facial emotions.
- It remains uncertain which particular emotions (for example

Charaka Palansuriya, EPCC
c.palansuriya@epcc.ed.ac.uk

Xingran Ruan
xruan@ed.ac.uk

Aurora Constantin
aurora.constantin@ed.ac.uk

Read more in the authors' paper "Real-time Feedback Based On Emotion Recognition for Improving Children's Metacognitive Monitoring Skill".
Available at:
https://edin.ac/3ow6ziL

Image: Paperkites/Getty Images

confusion, frustration, surprise, delight, etc) are crucial indicators for providing effective feedback that can support children's learning outcomes.

We are working on improving the accuracy of detecting children's FER using Vision Transformer (ViT) and Convolutional Neural Network (CNN) based machine learning techniques. The images on the right highlight how the techniques developed so far recognise the emotions of typically developing children.

The training of both ViT and CNN machine learning models are computationally expensive and made feasible by the use of EPCC's GPU-based high-performance computing infrastructure, such as Cirrus, and the Edinburgh International Data Facility's NVIDIA GPU-based Cloud infrastructure.

Our future plans are to perform studies with children in order to identify which of the emotions and their changes are important when a child is undertaking a learning exercise using a software-based educational tool. Here, we will investigate the relationship between emotions and how well a child thinks he or she is learning. We then plan to integrate these techniques and tools to build an Intelligent Tutoring System to help children with autism to learn with appropriate realtime feedback and support to facilitate better educational outcomes.

Special acknowledgement is due to EPCC PhD student, Xingran Ruan, for his extensive and in-depth research on the development and improvement of the accuracy of facial emotion recognition techniques. Additionally, Xingran is exploring ways to incorporate psychological aspects into machine learning models to enable effective real-time feedback.

Technical publications on these topics will be available soon. Findings from Xingran's PhD research will be integrated into the Intelligent Tutoring System (ITS).

Anger

Disgust

Sad

Surprise

Fear

Happy

Education and training at EPCC: https://edin.ac/epcc-training-education

ExCALIBUR: preparing for Exascale

ExCALIBUR is a UK research programme based around five themes that were identified as crucial for the UK's progress towards Exascale computing: RSE Knowledge Integration; High Priority Use Cases; Emerging Requirements for High Performance Algorithms; Cross-cutting Research, and Hardware & Enabling Software (H&ES). EPCC is involved in a number of projects to address these themes.

Nick Brown, Kirsty Pringle, and Anna Roubíčková, EPCC

n.brown@epcc.ed.ac.uk
k.pringle@epcc.ed.ac.uk
a.roubickova@epcc.ed.ac.uk

UNIVERSE-HPC

The UNIVERSE-HPC project will create a comprehensive collection of training materials to train the next generation of research software engineers (RSEs), with the aim of increasing both the skill and diversity of people working in research software engineering.

The project is a collaboration between the Universities of Edinburgh, Oxford, Southampton, and University College London.

The first phase investigated existing training and looked to understand what prerequisite learning is required for each course, this has allowed us to create a map of existing training and develop multiple "learning pathways" to guide potential learners through the material. Secondly, we ran a quick survey to understand the factors that inhibit RSE learning on the job, and the most important was found to be lack of time. To address this, we have developed a new training series called "Byte Sized RSE" which delivers focused training sessions in concise, one-hour long sessions. Learning is then backed up with an accompanying podcast. The series has proved extremely popular; five sessions have been run so far on topics including licensing and code review.

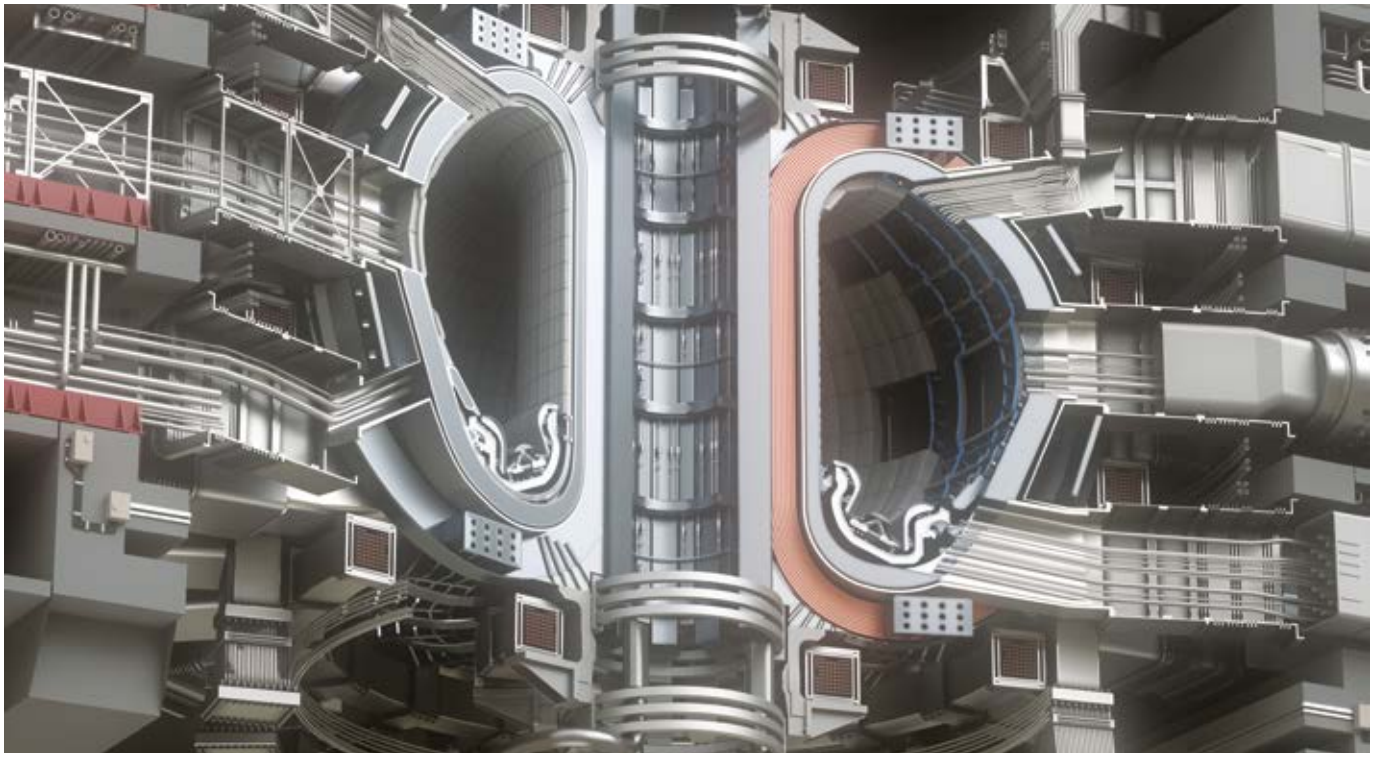
SiMLInt

SiMLInt is a cross-cutting project bringing the speed of machine learning (ML) to large-scale physics simulation, with a particular focus on plasma modelling in tokamak reactors, which is one of the high priority use cases.

The project aims to provide the infrastructure to enable efficient communication between simulation codes and data-driven models. This has the potential to allow scientists to run the simulations in a coarser resolution, using considerably fewer computational resources, and employing pre-trained ML models to supplement the unresolved, sub-grid scale information.

The technical implementation is based on Cray Labs's SmartSim technology and ensures that both the physics and data-driven codes run in a synchronised manner and utilise the HPC resources well. The usefulness of the ML models is determined by the quality of the data they are trained on, and the way the training has been done, both of which require a sound understanding of the underlying ML

The ExCALIBUR programme (Exascale Computing ALgorithms & Infrastructures Benefiting UK Research) is supported by the UKRI Strategic Priorities Fund. The programme is led by the Met Office and the Engineering and Physical Sciences Research Council (EPSRC) along with the Public Sector Research Establishment, the UK Atomic Energy Authority (UKAEA) and UK Research and Innovation (UKRI) research councils.



ITER tokamak reactor.
Image: Filipp Borshch/Getty Images

techniques. Therefore, in the cross-cutting spirit of the project, SiMLInt also provides guidance and support for generating suitable training data, developing and training the ML model, and assessing its properties, while highlighting questions and considerations related to the reliability of the ML model and the whole workflow.

xDSL

xDSL is working to develop a common ecosystem for Domain Specific Languages (DSLs). DSLs significantly raise the abstraction level when programming supercomputers, and this is especially important as we move towards much more complex architectures in the Exascale era. However, they currently share very little underlying infrastructure, resulting in a high barrier to entry for DSLs and long term support challenges. Exposing a Python interface to the ubiquitous MLIR and LLVM, xDSL provides the necessary building blocks for the low-overhead development of HPC DSLs.

ExCALIBUR H&ES RISC-V testbed

In EPCC we host the ExCALIBUR H&ES RISC-V testbed, enabling ExCALIBUR projects and UK HPC developers more widely to experiment with their codes on RISC-V. RISC-V, an open source Instruction Set Architecture (ISA) developed ten years ago and managed by RISC-V International, has the potential to transform computing. With over 10 billion RISC-V CPU cores already produced, the technology is enjoying phenomenal growth. In addition to the hardware itself, this project is also undertaking software development to improve the ecosystem for HPC and benchmarking.

ExCALIBUR H&ES FPGA testbed

EPCC also hosts the ExCALIBUR H&ES FPGA testbed, which aims to provide access to state-of-the-art Field Programmable Gate Arrays for scientists to port their codes to. FPGAs are configurable chips, enabling the electronics of the circuit to directly represent an application. This bespoke tailoring of the electronics to an algorithm can deliver significant performance and energy efficiency gains compared to fixed architectures, such as CPUs or GPUs. However a major challenge is how to best exploit this technology, which the testbed aims to assist with.

https://excalibur.ac.uk/projects/universe-hpc
https://excalibur.ac.uk/projects/simlint
https://xdsl.dev
https://riscv.epcc.ed.ac.uk
https://fpga.epcc.ed.ac.uk

Cray Labs's SmartSim technology:
https://www.craylabs.org/docs/overview.html

Chancellor's Fellowship in Quantum Computing

Quantum computer image by Bartolmiej Wroblewski/Getty Images.

The Chancellor's Fellowships scheme at the University of Edinburgh is a prestigious 5-year programme aimed at guiding outstanding early career researchers towards becoming leaders in their scientific fields and driving innovation.

Although the scheme has been running for several years, the recruitment of the most recent cohort of Chancellor's Fellows was EPCC's first involvement with the programme and we are delighted that two Fellows will be part of EPCC.

Our first Chancellor's Fellow is Dr Oliver Thomson Brown and he describes his Fellowship below.

My fellowship will be focused on answering three fundamental questions about quantum computing and HPC: where, when, and how?

First, I'm interested in seeing where we will be able to apply quantum computing to accelerate or enhance computational methods. This is a major open question within the field, and perhaps the most important currently. Naturally, I'm particularly interested in those areas where HPC is currently applied, but I'm excited to see what new problem domains quantum computing may open up!

Secondly, I'm interested in when we will actually see advantage from using a quantum processor (QPU), be it in terms of performance, computability, or energy usage.

Finally, I'm interested in how we will combine quantum computing and high-performance computing (HPC). Much of my work at EPCC until now has been focused on programming models for heterogeneous systems, and it doesn't get much more heterogeneous than combining quantum and classical compute. There are many exciting open questions about how QPUs may interact with existing HPC programming models like OpenMP and MPI. In the near term, we expect to see one QPU per supercomputer, but will we one day see one per node? Perhaps even one per NUMA region!

I'd also like to investigate what HPC can do for quantum computing. My PhD was in simulating many-body quantum systems, and simulation remains an important tool for quantum algorithm developers, despite the well-known scaling problem. It's also expected that quantum computing may itself require significant classical computing resources, to handle I/O and error mitigation for example.

I am very excited to be starting this new chapter of my life at EPCC, and with the recent launch of the Edinburgh Quantum Software Lab collaboration between Informatics and EPCC, the timing couldn't be better.

Oliver Thomson Brown, EPCC
o.brown@epcc.ed.ac.uk

Edinburgh Quantum Software Lab

The Quantum Software Lab, in collaboration with the National Quantum Computing Centre (NQCC) will identify key challenges in the adoption of quantum computing and investigate new ways in which quantum computers can provide benefits, beyond the reach of traditional computers.

Researchers from the Lab will work closely with industry partners to understand how quantum computers might help address their problems.

Developments at the Advanced Computing Facility

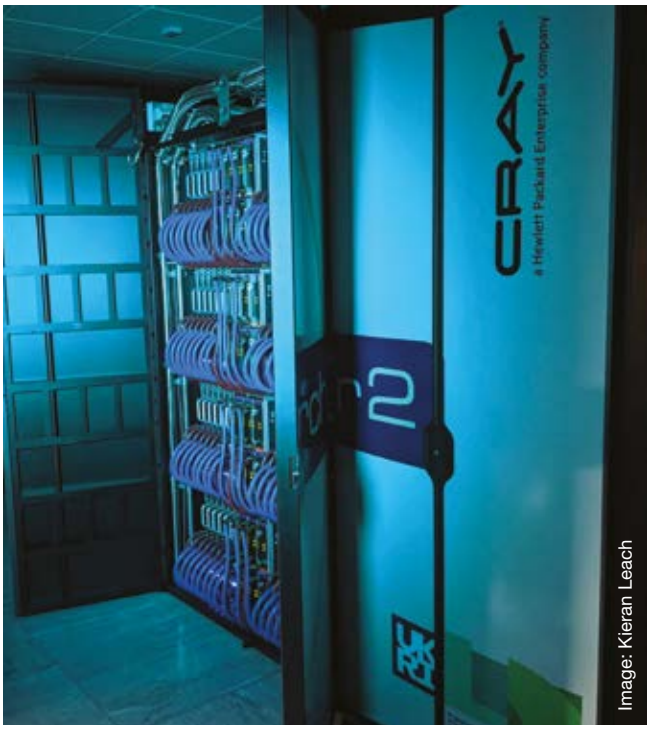


Image: Kieran Leach

Our world-class Advanced Computing Facility (ACF) supports the computing and data services provided by EPCC.

Computer Room 4

We have continued to develop our plans around Computer Room 4 (CR4) at the ACF, currently home to equipment hosting the Edinburgh International Data Facility (EIDF). Over the last 18 months, designs to expand Computer Room 4 and the supporting Plant Room have been prepared in readiness for the UK's first Exascale system, which is expected to need between 20-25MW of power.

In parallel our power supplier, SPEN, has installed power cables from the main electricity distribution site in Edinburgh to a new 30MW substation which sits on the ACF site. Along with this our 200 Gbit/s site Data Centre Network and dual 100 Gbit/s direct connection to the JANET network will allow us to support the significant traffic requirements of such a system.

Computer Room 1

We are also working to upgrade our Computer Room 1. This circa 250m^2 Computer Room currently plays host to the Tursa DiRAC Extreme Scaling system and, to an increasing degree, equipment for the EIDF. We are making changes so that 95% of the equipment in this room will be water cooled (either directly or indirectly). We expect this to offer significant energy savings, decrease our Power Usage Effectiveness number, and greatly reduce our reliance on whole room air cooling CRAC (Computer Room Air Conditioning) units.

As an additional benefit of this work we are able to remove a number of CRAC units, which will add space for an additional row of computer equipment in the room.

Computer Room 3

We are continuing to make energy efficiency improvements for ARCHER2. As part of efforts to increase efficiency on site and to minimise energy use we have been working, alongside HPE, to adjust water temperatures for ARCHER2 to maximise the level of free cooling possible.

Carefully monitoring ARCHER2 alongside our colleagues from HPE we have so far succeeded in raising the temperature of water provided internally to the ARCHER2 compute nodes to 32°C and the temperature of water provided to the ARCHER2 Cooling Distribution Units to 24°C.

Our next phase will be to move the temperature of the chillers in Plant Room C from 14°C to 15.6°C (their design limit). We are progressing design work to use the mixing of water to move our chilled water temperature beyond this and access further free cooling efficiencies.

Paul Clark, EPCC
p.clark@epcc.ed.ac.uk

When providing cooling to our various services on site we endeavour to use as much "free cooling" as possible. Free cooling is when we are able to cool systems without the use of chiller units, directly cooling the warm water with cool Scottish air.

EPCC's Advanced Computing Facility
https://edin.ac/advanced-computing-facility

Hosting and operating the ARCHER2 service

Many of the articles about ARCHER2 contain a statement such as “ARCHER2 is hosted and operated by EPCC at the University of Edinburgh”. But what does this actually mean?

ARCHER2 is a world-class advanced computing resource for UK researchers. It is an HPE Cray EX supercomputing system with an estimated peak performance of 28 Pflop/s. This performance is utilised by academics from across the UK (and beyond) to perform world-class science, with wide ranging benefits to society and the economy. The service has over 2700 active user accounts and regularly sits at around 85-90% utilisation.

Hosting ARCHER2

The system is hosted at the Advanced Computing Facility, our world-class data centre. ARCHER2 is housed securely at the site, with the team ensuring that the appropriate power, cooling, and monitoring are in place for it to run efficiently.

A recent blog post (“Ensuring continuity of service at the ACF”) by my colleague Calum Muir is an interesting read, giving insight into the facilities and activities that are required to host this sort of system. For example, ARCHER2 is hosted in our state-of-the-art Computer Room 3, which has recently undergone an upgrade of the main power distribution units’ (PDUs) supply cabling and sub-floor power supply cables to ensure that the electrical infrastructure has additional resilience when ARCHER2 is at maximum capability.

Service and support provision

In addition to hosting the ARCHER2 service, EPCC is contracted to provide the Service Provision (SP) and Computation Science and Engineering (CSE) support.

The SP team is responsible for maintaining the system, for example ensuring software upgrades are completed, also security management, job management customisation, monitoring compute and storage availability, and job monitoring. This group also runs the Service Desk, offering front-line support to all our users. Developing and maintaining the ARCHER2

ARCHER2 is the UK’s National Supercomputing Service. In addition to providing a world-class resource for UK research, the ARCHER2 service works to demonstrate the value of supercomputing to people across the UK through events, learning materials and work placements.

website and documentation completes the roles of the SP team. This coordinates interactions between different parts of the service, for example working closely with our colleagues at HPE.

The CSE team works more closely with our user community, providing in-depth support. For example we work with users to help port codes, debug problems, and look at optimisations. We run an extensive training programme for users and the wider UK community, running both online and in-person courses. We look to provide the best possible user environment, providing tools and best practice guides. The CSE team also coordinates the "embedded Computational Science and Engineering" (eCSE) programme, which allocates funding to Research Software Engineers across the UK, funding staff embedded in user communities working to enhance the application software on ARCHER2.

Outreach

Outward facing, we engage with different science communities to share knowledge, research, and best practice. We also enjoy delivering activities to school children and the general public, providing hands-on activities at science festivals, and demonstrating the societal benefit of supercomputing and the opportunities available to young people considering a career in computational science.

So all in all, the phrase "hosted and operated" covers a multitude of activities, from electricity supplies, through software management, queries, training, and outreach to schools. It all makes for an interesting and varied job!

ARCHER2 is provided by UKRI, EPCC, HPE Cray and the University of Edinburgh.

Lorna Smith
l.smith@epcc.ed.ac.uk

ARCHER2 Image Comp

All users of the service are annually invited to submit their images of research that is enabled by ARCHER2. You can see all contributions in the gallery on the ARCHER2 website https://edin.ac/43JrYW1

ARCHER2 website
https://edin.ac/archer2

Supporting science with HPC

The DiRAC programme's Tursa system at EPCC's Advanced Computing Facility supports particle physics and astronomy research.

DiRAC (Distributed Research Utilising Advanced Computing) provides distributed high performance computing services to the STFC theory community, providing particle physicists and astrophysicists with both data and compute resources to run their simulations and process the data produced.

EPCC's Advanced Computing Facility is one of four sites hosting DiRAC HPC facilities, and we are involved in DiRAC in several ways.

At Edinburgh we host and manage the GPU-based Tursa system, which comprises 114 GPU nodes each with four NVIDIA A100 Tensor core GPUs connected with fast Infiniband Interconnect, allowing large calculations to run efficiently on many GPUs at once. This year's upgrade will increase the size by around half as much again, allowing even bigger calculations.

Each DiRAC system has particular features which make it appropriate for different kinds of research and this is where the RAC comes in.

The Resource Allocation Committee (RAC), of which I am the Technical Manager, decides how much resource is given to each research group, and on which systems each project is hosted.

DiRAC RAC calls are issued yearly with occasional extra calls if resources become available mid-year. As this is a yearly process, and projects can run for up to three years, this is a large exercise which is usually started in the spring when we assess what resources are available, calls then open in the summer and close in the autumn, with the winter months used for reviewing.

First, technical reviews are carried out (which I organise within the technical manager role). The proposals and completed technical reviews are then passed to the Science and Technology Funding Council who arrange scientific reviews. Once all reviews are complete, resources are allocated and projects begin in April.

Chris Johnson, EPCC
c.johnson@epcc.ed.ac.uk

HPC-based modelling is an essential tool for the exploitation and interpretation of astronomy and particle physics data generated by facilities.

Further information about DiRAC and associated RAC calls can be found on the DiRAC website
https://edin.ac/41oK916

HPC services at EPCC
https://edin.ac/hpc-services

EPCC's Outreach team goes back on the road

Images taken at the Edinburgh Science Festival.
Credit: Ludovic Farine

Last year we were finally able to return to running in-person events after the pandemic, and it was fantastic to see so much interest from people of all ages.

We were particularly excited to see the number of young people who were interested in learning about supercomputing, data science, and how it impacts everyday life.

Following on from a successful New Scientist Live event in October 2022, where we demonstrated our ever-popular Wee Archie system, ball sorting activity and logic puzzle challenges, the new year saw us work on new activities for future events. The first came fast, a four-day exhibition at the Edinburgh Science Festival (ESF) in April as part of the "Discover Our Science" University of Edinburgh program.

At ESF, we ran three activities to describe the inner workings of a computer. We had a binary bracelet-makers' station to help teach how binary numbers represent data inside computers. We also had a deep dive into the type of hardware present in modern computers, with a typical PC motherboard placed next to a compute node from ARCHER2, allowing us to show just what makes a supercomputer so "super".

The last activity is a new addition to our outreach offerings: Spintronics[1]! This is a hands-on activity that allows you to build your own mechanical version of electrical circuits using components linked together by moving chains rather than electricity and copper wires. Attendees used these to create mechanical transistors and logic gates, two important building blocks of computers.

Currently in development is a refresh of Wee Archie, with a view to including NVIDIA Jetson Nano components, which will allow us to demonstrate how supercomputers can be used to accelerate Machine Learning and AI. We are also hoping to make use of the Edinburgh International Data Facility (EIDF) to provide a virtual sandbox in Minecraft, with both a virtual Advanced Computing Facility to tour (which hosts EIDF itself alongside our wide range of systems), as well as Minecraft equivalents of our logic puzzles and spintronics tasks, complementing our in-person events.

We are confident that these new activities will be just as popular as our existing offerings, and we are keen to share the excitement of supercomputing with even more people in the years to come.

Up next for us is a return to the Big Bang Fair in June at the NEC in Birmingham, which is always a favourite for the team. We are looking forward to seeing you there!

[1] https://upperstory.com/spintronics/

Darren White
d.white@epcc.ed.ac.uk

2022 was a great year for EPCC Outreach, and we are looking forward to meeting many more people at events in 2023!

EPCC Outreach
https://www.bit.ly/discover-epcc

Updates from the Software Sustainability Institute

The Software Sustainability Institute (SSI) cultivates better, more sustainable research software to enable world-class research. The Institute helps people build better software and works with researchers, developers, funders and infrastructure providers to identify key issues and best practice in scientific software.

Collaborations Workshop 2023

The Software Sustainability Institute's Collaborations Workshop series brings together researchers, developers, innovators, managers, funders, publishers, policy makers, leaders and educators to explore best practices and the future of research software. Collaborations Workshop 2023 (CW23) took place as a hybrid event in Manchester, UK, from 2–4 May 2023. This year's theme was Sustainable Career Development for those in the research software community: looking after your software, your career, and yourself. The theme encompassed technical, career, and personal development and aimed to explore ways in which RSEs can develop their skills and progress their career while always taking care of their mental health and well-being.

Find out more:
https://software.ac.uk/cw23

Research Software Camp

The Research Software Camps (RSCs) help address the lack of software skills in research through free online introductory workshops and resources. The RSCs allow researchers with limited or no experience in software engineering to learn new coding skills and receive personal guidance in a welcoming and supportive environment. The next Research Software Camp will be held from Monday 19 June to Friday 30 June 2023. It will include a panel discussion, basic skills workshops (R, Python, visualisation, etc) and a mentorship programme. Curated online resources will be published throughout the two weeks of the Camp, addressing varying approaches to learning to code in academia.

Find out more:
https://www.software.ac.uk/research-software-camps

Software Sustainability Institute 2023 Fellows

Earlier this year the Software Sustainability Institute hosted the 2023 Fellows Inaugural Meeting to welcome this year's cohort. The Fellowship Programme provides funding for individuals who want to improve how research software is used in their domains, fields, and

Denis Barclay
d.barclay@epcc.ed.ac.uk

Better software, better research

Since 2010, the SSI has facilitated the advancement of software in research by cultivating better, more sustainable, research software to enable world-class research.

areas of work. This funding can be used for any activities that meet both the Fellow's and the Institute's goals, such as organising or travelling to workshops, nurturing or contributing to communities of practice, running training events, collaborating with other Fellows, or for any other activities that relate to improving computational practice or policy.

Find out more: https://software.ac.uk/programmes-and-events/fellowship-programme

Report on the AHRC Digital/ Software Requirements Survey

The Software Sustainability Institute has recently published the Report on the AHRC Digital/Software Requirements Survey 2021. The original survey run by the SSI on the Arts and Humanities Research Council (UKRI AHRC) community aimed to understand the views on digital/software tools, the experience of developing such tools, the practices, learning intentions and preferences around how to resource projects involving digital/software. The report concluded that there was a clear demand - and need - to provide software training, improve recognition of software outputs, and expand resources for software skills and techniques in the arts and humanities communities.

Find out more: https://www.software.ac.uk/blog/2023-03-01-report-ahrc-digitalsoftware-requirements-survey

Code For Thought

Earlier this year, the Code For Thought podcast hosted by SSI Fellow Peter Schmidt celebrated its second anniversary. EPCC's own Weronika Filinger joined Peter Schmidt on the latest episode to discuss the importance of training. Although there are many opportunities available for RSEs like the Code Refineries and various Carpentries, there is a considerable gap between the demand for experienced engineers and the amount of training provided, which is unsustainable long-term.

Find out more about their discussion and other episodes of Code For Thought: https://www.software.ac.uk/tags/code-thought

The Software Sustainability Institute (SSI) is a leading international authority on research software sustainability, working with researchers, funders, software engineers, managers, and other stakeholders across the research spectrum. EPCC is a founding partner.

Find out more on our website: https://edin.ac/software-sustainability

News from EPCC's Education team

Ben Morse, EPCC
b.morse@epcc.ed.ac.uk

Home is where the coffee is

One of the biggest changes from the previous academic year has been how much more use students have been making of the EPCC facilities within the Bayes Centre. Our first eighteen months in the building, until lockdown in March 2020, saw our students regularly treating EPCC like their second home – with EPCC's coffee machine definitely an important fixture and focal point.

As Jonas Faßbender, a 2020 graduate, put it: "The best thing about the MSc was being able to take a break and grab your lunch or a coffee next to an expert like Mark Bull and just talk about OpenMP, even months after that course."

The return to service, and student access, of the EPCC coffee machine may seem like a small step, but it has been transformational in our efforts to foster a community among our students as it once more demonstrates that EPCC is a welcoming and open environment where staff, PhD students and MSc students can catch up and discuss everything from football and rugby results to the current state of RISC-V CPUs.

That our alumnus, and now colleague, Daniyal describes joining the 'EPCC family' in November, despite his own MSc taking place in the still-COVID-restriction-affected 2021/22 academic year inspires the Postgraduate Programmes team and our teaching colleagues all the more to keep offering that welcome environment.

There's no place like home

Another new colleague who joined us in November is Dr Ludovic Capelli, a 2016 MSc graduate who has since been studying for his PhD under the supervision of EPCC's Dr Nick Brown through the EPSRC Centre for Doctoral Training in Pervasive Parallelism.

In truth, Ludovic has never really left EPCC since beginning the MSc in 2015, beyond taking up internship opportunities during his PhD. His arrival also underlines EPCC's ongoing commitment to helping our graduates develop skills in HPC which are highly in demand – now further supported by the Online Part-time MSc programmes.

Ludovic's arrival officially as a colleague has brought fresh vigour to our educational offerings, such as the launch in 2023 of EPCC's HPC Summer School.

EPCC HPC Summer School

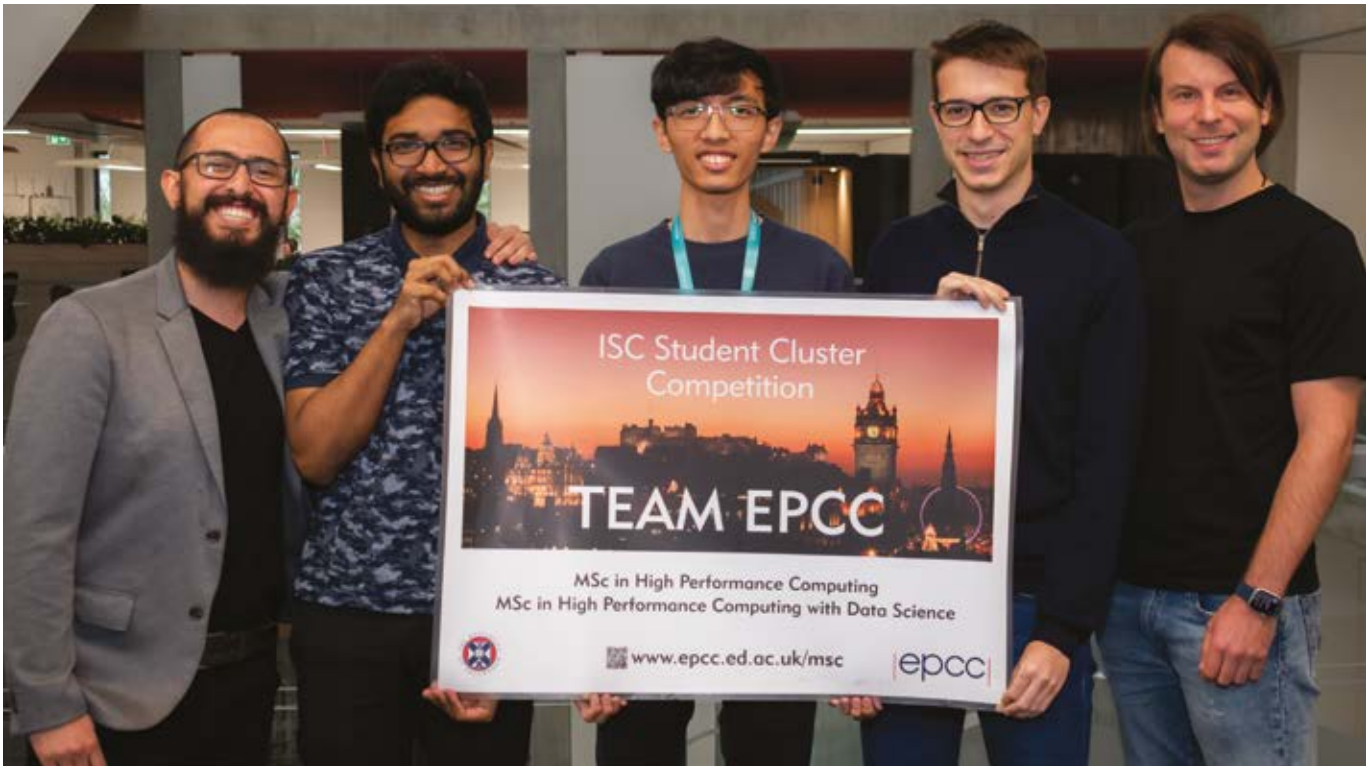
The HPC Summer School will offer opportunities for undergraduates from UK universities to spend five weeks in Edinburgh learning the fundamentals of HPC before working in groups on a variety of HPC and computational science and engineering projects, beginning with a hackathon and rounded off by a more longitudinal project-based approach.

TeamEPCC returns to ISC

Being able to take part in the Student Cluster Competition alongside their studies, through integration with their dissertation projects, opens up many opportunities for TeamEPCC members, with a disproportionately high number of former competitors now working at large companies in the HPC space, including ARM, or continuing with their studies through PhDs either at EPCC or elsewhere.

"The MSc in HPC with Data Science has been a one-of-a-kind experience, giving me exposure to world-class computing facilities, alongside a conducive learning environment. I feel grateful for the constant support provided by Ben and Jemma throughout the programme and Dr Anna Roubíčková for her guidance as my dissertation supervisor. I also had the opportunity to engage with the wider University and Edinburgh community as a Student Ambassador, and since graduating, I am thrilled to have joined the EPCC family as I embark on my professional journey!"

Daniyal Arshad, former MSc student (2021/22) now an EPCC Applications Developer



TeamEPCC, from left to right: Tomas Rubio Cruz, Jaffery Irudayasamy, Ikraduya Edian, Hristo Belchev, and Oleksandr Piekhota.

TeamEPCC, five students from EPCC's MSc programmes, will take part in the on-site ISC23 Student Cluster Competition (SCC) in Hamburg.

ISC SCC is an annual event for all the selected student teams to demonstrate their abilities to design their own clusters and achieve the best performance on a selection of benchmarks and scientific applications within a certain power usage limit. Taking part in the competition is a great opportunity for our students to push their skills and knowledge of high-performance computation in a realistic, high-stakes environment.

With the support of our sponsor, HPE, and the EPCC systems team, TeamEPCC has designed a cluster specifically tailored to meet the computational demands of this year's benchmarks. While multiple designs were considered, the team decided on a GPU-dominated cluster, as most of this year's codes can be GPU accelerated. While this promised the best overall performance, balancing power-constraints is an ongoing challenge. Our SCC cluster for this year will have two nodes, each equipped with eight NVIDIA A100s GPUs, dual-socket AMD 7717 64-core CPUs, and 1TB RAM. To keep the system's temperatures in check on the conference floor, the HPE Apollo d6500 Gen10 chassis employes fifteen 80mm high-RPM fans.

To better prepare for the ISC23 SCC, TeamEPCC also visited EPCC's Advanced Computing Facility where the competition cluster and several UK national HPC systems are located. The visit provided the team with the opportunity to learn how to perform basic hardware maintenance and understand the constituent components and the cooling mechanisms.

TeamEPCC is ready to showcase what they have learned at EPCC and see how they measure up against the world-class teams that will be attending the ISC23 SCC. While aiming to excel in the competition, the team also looks forward to the unique opportunity to network with academic and industrial experts in the HPC world and engage with other teams from around the globe at ISC23.

We would like to thank HPE for being TeamEPCC's sponsor again for the ISC23 SCC.

Xu Guo and Spryo Nita, EPCC
x.guo@epcc.ed.ac.uk
s.nita@epcc.ed.ac.uk

Look for TeamEPCC at booth A116 of the student cluster competition.

EPCC will also take part in ISC'23 with events and a booth in the exhibition hall. Find us at G715.

Education and training at EPCC: https://edin.ac/epcc-training-education

PhD student industrial placements

One of the popular aspects of studying for a PhD at EPCC is being able to leverage our global network of contacts to generate opportunities. Internships are an obvious an example of this, where students pause their study to spend time with a company and apply techniques they have researched and developed to the company's problems. Not only does this give experience of how their research can be translated to the real world, but students are generally paid for their time.

Two of my PhD students, Gabriel Rodríguez Canal and Mark Klaisoongnoen, recently undertook internships with Hewlett Packard Enterprise (HPE). Both moved from Edinburgh to HPE's HPC/AI EMEA Research Lab in Bristol from October last year to the end of March. The Lab's seventeen members drive diverse research in high-performance computing (HPC) and in HPE Centres of Excellence (CoE), working closely with clients. For example the CoE covering the national supercomputing service ARCHER2, which is an HPE Cray EX system hosted by EPCC.

Producing the Cray line of supercomputers as well as many other technologies and software developments, HPE is one of the leading companies in the field of supercomputing. Consequently these internships were a fantastic opportunity for both students to work with a company at the forefront of HPC.

FPGAs for HPC

Both Gabriel and Mark are exploring FPGAs and I think it's fair to say that HPE, as per the HPC community in general, has been more focussed on CPUs and GPUs rather than this technology to date. Gabriel's internship investigated enabling Fortran programming for FPGAs using High Level Synthesis (HLS). To this point when writing code in HLS one must use C or C++, however many HPC codes are written in Fortran which programmers are familiar and comfortable with. Thus, an initial and often time-consuming step when porting to FPGAs is to first translate these into C/C++ which the programmer might be unfamiliar with. The idea of this work is that existing, potentially legacy, Fortran HPC codebases can then be run directly on FPGAs with the programmer then being able to concentrate on tuning their code to the architecture using a language that they are familiar with.

Benchmarking

Throughout Mark's PhD he has been working on benchmarks developed by the Securities Technology Analysis Center (STAC®), specifically exploring how FPGAs can accelerate them in an energy-efficient manner. This is a fantastic opportunity because, with membership comprising over 400 financial institutions and more than 50 technology vendors, STAC provides industry standard financial

Nick Brown, EPCC
n.brown@epcc.ed.ac.uk

We offer opportunities for PhD study in areas related to high performance computing, computational science, data science, software engineering and sustainability, and parallelism.

As one of Europe's leading supercomputing centres, with a wide, varied, and interesting hardware ecosystem available, EPCC offers a unique setting for in-depth research and excellent connections into industry to provide further opportunities for our students and graduates.

Gabriel Rodríguez Canal (left) and Mark Klaisoongnoen.

benchmark suites representing common workloads. While interning at HPE Mark continued to work with these benchmarks, specifically exploring whether they can be accelerated using AMD Xilinx's new AI engine technology which can be found in its latest Versal ACAP.

Not only are the techniques that are being developed of interest here, but furthermore it is also a demonstrator for how beneficial this new technology is to these types of workloads.

Access to novel technologies

Another other important benefit of being an EPCC PhD student is access to the latest novel technologies for research. This is especially important for both Mark and Gabriel as throughout their PhDs and internships they are working with the FPGA testbed hosted by EPCC and funded by the ExCALIBUR H&ES programme (see page 10).

Mark Klaisoongnoen
EPCC PhD student

"With my PhD research I've been focusing on accelerating financial codes using FPGAs and at HPE I have a wider range of opportunities to work on challenges around how to make such devices easily deployable for HPC users. While working with the EMEA Research Lab, I have gained an overview of HPE's global initiatives and involvements with customers, for instance through their various Centres of Excellence. Overall, the research internship is a fantastic chance to connect with researchers and industry-leading clients in HPC and to explore career opportunities."

Gabriel Rodríguez Canal
EPCC PhD student

"My internship at HPE has given me an insight into the industrial approach to research. HPE being one of the leading companies in HPC, this opportunity has benefitted my career with new connections and given me a glimpse of the state of the art industrial side of HPC.

"Our programme included regular seminars on the work undertaken by the lab, which touched on topics that are not directly part of my research and were quite stimulating. Overall, it has been a fruitful experience that has enabled me to gain new skills that I will be applying once I resume my PhD studies."

PhD opportunities at EPCC

We have lots of exciting PhD opportunities available across EPCC. Please contact me if you are interested in novel hardware architectures for HPC (eg RISC-V, FPGAs etc) and/or compilers and runtimes.

Education and training at EPCC: https://edin.ac/epcc-training-education

Study HPC with us

Master's degrees in High Performance Computing (HPC) and in HPC with Data Science

EPCC is the UK's leading supercomputing centre. We are a major provider of HPC training in Europe, and have an international reputation for excellence in HPC education and research.

Our MSc programmes in High Performance Computing (HPC) and HPC with Data Science have a strong practical focus and provide access to leading edge systems such as ARCHER2 (the UK's National HPC Service), and Cirrus (an EPSRC Tier-2 National HPC facility including over 150 GPUs).

MSc students have the opportunity to undertake their dissertations as an industrial project, building on EPCC's strong business links. Recent project partners range from start-ups to multinationals.

Optional course choices include modules from the School of Informatics and the wider College of Science and Engineering.

Our graduates are in high demand in both academia and industry in the UK and abroad.

The University of Edinburgh is ranked in the top 30 universities in the world by Times Higher Education World University Rankings 2023, and 15th by QS World University Rankings 2023.

"The quality of the courses provided is exceptionally high, all lecturers are the field experts in HPC. Additionally the support that EPCC provides to students is second to none, all staff are extremely helpful and supportive."
Xiaoyan Ma
MSc in HPC with Data Science graduate, 2021
Now CEO, Danu Robotics

"The exposure to world-class computing facilities and the National HPC Service coupled with lectures taught by highly experienced tutors has opened a world of new possibilities for me, including skills that I am now further honing at EPCC."
Daniyal Arshad
MSc in HPC with Data Science graduate, 2022
Now Applications Developer, EPCC

www.epcc.ed.ac.uk/msc