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

epcc news

Issue 89 SUMMER 2021

Opening up new frontiers in data and HPC

Image: Dieter Meyrl via Getty Images

From our Director

Wherever you work or live, I do not think any of us realised in March 2020 that life would not have returned to normal by summer 2021.

As I write this, the vast majority of EPCC staff are still working from home and have been now for sixteen months. I'd like to take this opportunity to thank everyone in EPCC, not just for coping, but in many cases going well beyond their role to keep our systems and projects running as well as we could throughout the pandemic.

As with all organisations, there have been many challenges along the way but hopefully we are starting to see the beginning of the end. Over the autumn, we will begin returning to our offices in the Bayes Centre and start to bring a full complement of staff back to our data centre (which has remained open throughout).

In parallel, we hope to see the end to the problems we have witnessed with the ARCHER2 system and finally begin to deliver its potential to the EPSRC and NERC computational science communities.

Within this newsletter you will find a taste of the huge number of activities we've managed to undertake despite the pandemic. I hope you enjoy reading it and we look forward to seeing people in person again over the coming months.

Mark Parsons
EPCC Director
m.parsons@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.

Creating virtual humans

Image from Virtual Humans film courtesy of CompBioMed and Barcelona Supercomputing Center.

EPCC is a Core Partner of CompBioMed, a European Centre of Excellence focused on the use and development of computational methods for biomedical applications. We prepare these applications for future exascale systems, to create the first virtual humans: digital twins to enable personalised medicine.

We support medical scientists from both academia and industry, ie hospitals and medical research institutions, who research three main scientific areas: cardiovascular medicine, molecular-based medicine (including COVID-19 research), and neuro-musculoskeletal medicine.

EPCC provides access and support to CompBioMed users on the Cirrus and ARCHER2 systems, but also assists in preparing their applications for future exascale platforms. We are active in CompBioMed's training activities and coordinate its e-Seminar series. EPCC also engages with potential new users, offering free access to the CompBioMed service to port and scale biomedical applications to supercomputers.

We also support those interested in enabling workflows that require safely moving sensitive data to large computers for processing via traditional HPC simulations and/or machine learning. Through a collaboration with the LEXIS consortium, we are investigating a particularly exciting new workflow we have named Resilient HPC, designed for urgent, safety-critical computations.

Coronavirus research

CompBioMed is active in a large international consortium across Europe and USA working on urgent coronavirus research. To date, the consortium has redirected substantial research effort and funding into computational investigations that has improved our understanding of the SARS-CoV-2 virus and the associated COVID-19, and has accelerated the development of treatment options, including antiviral drugs and vaccines.

The work of the consortium includes bioinformatics analysis and simulation, molecular modelling, electronic structure calculations, epitope analysis, machine learning, epidemiological studies, and the creation and hosting of a growing collection of relevant datasets.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823712.

Gavin Pringle, EPCC
g.pringle@epcc.ed.ac.uk

CompBioMed Conference 2021
15-17 Sep, 2021

Our theme for 2021 is "Building the Virtual Human: How to bring experimental research into your digital twin."

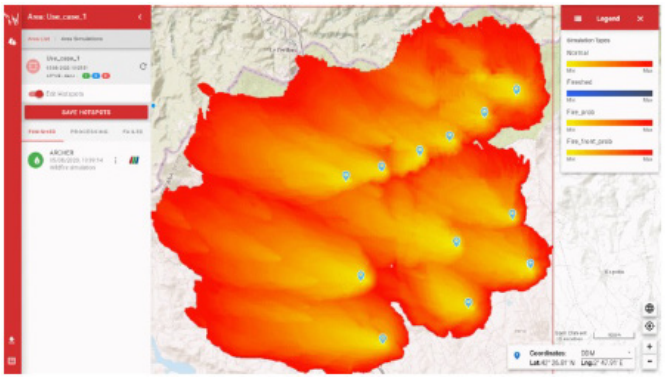
This event will address all aspects of computational biomedicine, from genome to whole human and population levels, embracing data driven, mechanistic modelling and simulation, and machine learning.

For full details and registration please visit:
https://cbmc21.vfairs.com

CompBioMed
www.compbiomed.eu

LEXIS
https://lexis-project.eu

VESTEC: HPC for urgent disaster response



VESTEC is an EU Future and Emerging Technologies (FET) project that is investigating fusing HPC with real-time data and visualisation techniques to provide a step change in the handling of disasters.

With one of the VESTEC project's original use-cases considering the spread of diseases, we never imagined how timely the project would be!

The underlying premise is that technology, and for us in EPCC high performance computing (HPC) especially, has an important but as yet under-exploited role to play in tackling disasters in real time. The problem is so significant that it requires a very diverse skillset, and so VESTEC has involved partners from across numerous areas of expertise, including visualisation, data reduction, mosquito modelling, HPC, weather prediction, satellite data streams, and wildfire simulation.

In EPCC we have been responsible for the core HPC infrastructure, which is effectively middleware connecting the numerous simulations that are required as part of disaster prediction and running on the supercomputers, with the urgent decision makers who are driving these simulations via client GUIs, and real-time data streaming into the system.

There have been numerous interesting discoveries made, for instance we have found that organising the significant number of options and control paths via workflows is very beneficial, as it provides flexibility and gives order to what should run and when.

We have now integrated our three project use-cases (wildfire simulation, space weather, and spread of mosquito-borne disease) with the VESTEC system, running numerous coupled simulations on the HPC machines and enabling emergency responders to interact with these models via custom GUIs or more general-purpose packages including ParaView. The use-cases employ a variety of specialist techniques that have been developed as part of VESTEC, such as the ability to perform very high order data reduction without loss of fidelity, so enabling the streaming of real-time simulation results without impacting the ability of decision makers to take the correct actions.

It is not just the technical highlights that have contributed to the VESTEC project, we have also made significant efforts to further develop the community in this area. This has included two very successful workshops at SC19 and SC20, and in this last phase of the project we have a mini-symposium at The Platform for Advanced Scientific Computing Conference (PASC) about urgent computing for disaster response in July, and additionally a third workshop at SC21 this year in November. VESTEC started in September 2018 and will end in February 2022.

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

If you are interested in this topic, please consider attending our events at PASC or SC21, and also submitting papers to the SC21 UrgentHPC workshop.

VESTEC website
https://vestec-project.eu

PASC21
https://pasc21.pasc-conference.org

UrgentHPC
www.urgenthpc.com

Supporting digital research across the arts, humanities and social sciences

Photo by Jazmin Quaynor on Unsplash.

The Centre for Data, Culture & Society (CDCS) at the University of Edinburgh supports the querying of large collections of digitised texts by researchers in the Humanities and Social Sciences.

defoe is a text and data mining tool based on Apache Spark. EPCC is helping researchers in CDCS use defoe to answer questions over a variety of datasets that CDCS has provided access to, using HPC infrastructure managed by EPCC to process vast amounts of digitised texts efficiently.

defoe has access to and can process digitised collections of various books including the first eight volumes of Encyclopaedia Britannica, spanning 1768-1860; British library books between 1510 and 1899; and newspapers (British Library Newspapers, Times Digital Archive, Paper Past: New Zealand and Pacific Newspapers, and Gazetteers of Scotland).

CDCS works hard to secure additional collections. This year we are ingesting two additional collections: Statistical Accounts of Scotland (1791-1845) and the University of Edinburgh's digitised PhD theses collection, starting from the 17th century. We expect more to come.

We use defoe to answer a wide range of topics, such as the context and tone of debates on economic policy and public finances, or how the reporting of South Asia has changed in British media over time (eg pre- and post-Partition of India).

We have learned from previous collaborations between EPCC and CDCS that it is better to progress from simpler, better defined queries to more complex questions, and so the queries typically start as a search for a set of keywords, n-grams or specific patterns, and gradually become more complex. The preliminary results are in line with expectations based on the archival work, and the outcomes of the queries help advance the research in the respective fields, in some cases forming a base for research publications.

In addition to ingesting more datasets, we are working to extend defoe's core functionality.

At the moment, defoe can retrieve fragments of texts based on keyword or pattern search, and these fragments can then be further processed. We are working to extend the querying ability to also support detecting similar passages between texts, building on external package Passim. This will allow more complex hypotheses to be tested in future, for example when applied to the University of Edinburgh PhD theses collections.

Lisa Otty, Edinburgh Centre for Data, Culture and Society
lisa.otty@ed.ac.uk

Anna Roubickova, EPCC
a. roubickova@epcc.ed.ac.uk

This work is part of a larger project funded by the Data-Driven Innovation network as part of their 'Building Back Better' open funding programme, helping to transform the Edinburgh City Region into the data capital of Europe. It is supported by the Scottish Funding Council Covid-19 Recovery funding to the University of Edinburgh.

Centre for Data, Culture & Society
https://www.cdcs.ed.ac.uk

Data-Driven Innovation network
ddi.ac.uk

defoe
https://github.com/defoe-code

Passim
https://github.com/dasmiq/passim

Advanced AI system comes to EPCC

Mark Parsons, EPCC Director, with the newly-installed CS-1.

The world's fastest artificial intelligence (AI) computer, the Cerebras Systems CS-1, has been installed as part of the International Data Facility at EPCC's Advanced Computing Facility.

Featuring the HPE Superdome Flex Server from Hewlett Packard Enterprise, this leading-edge AI deployment enables the next wave of natural language processing (NLP) and data science research for public, private and academic sectors across the City Region and the UK.

EPCC will employ a unique combination of the Cerebras CS-1 system, powered by the high-performance Wafer Scale Engine (WSE) processor, and an extremely large-memory HPE Superdome Flex Server system for unprecedented AI scalability and massive data handling capability.

This advanced AI supercomputing system will greatly reduce training time, the most time-intensive part of AI, enabling many more ideas to be tested, and will be available in the new Edinburgh International Data Facility (EIDF) for academic researchers and data scientists in the public and private sectors.

This partnership will accelerate AI-powered data science initiatives in the Edinburgh and Southeast Scotland City Region, enabling national-scale genomics research for public health initiatives. It also advances AI research for natural language processing of global academic and industry interest, using deep learning models like BERT in the School of Informatics' Institute for Language, Cognition and Computation.

The CS-1 is built around the world's largest processor, the WSE, which is 56 times larger, has 54 times more cores, 450 times more on-chip memory, 5,788 times more memory bandwidth and 20,833 times more fabric bandwidth than the leading graphics processing unit (GPU) competitor. In AI compute, large chips process information more quickly, producing answers in less time. Depending on workload, from AI to HPC, the CS-1 delivers hundreds or thousands of times more performance than legacy alternatives, and it does so at a fraction of the power draw and space.

EPCC will use the HPE Superdome Flex Server, a powerful and easy-to-use high performance front-end storage and pre-processing solution for the CS-1 AI supercomputer. This will enable users to employ large datasets and application-specific pre- and post-processing of data for AI model training and inference on the CS-1, allowing the CS-1s WSE to operate at full bandwidth.

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

This installation will enable massive breakthroughs in our vision for data science and greatly accelerate our research across genomics and public health, including time-sensitive and pressing issues such as leveraging AI across large models to advance COVID-19 therapeutic research.

Read more about Cerebras Systems' CS-1.
https://cerebras.net/blog/introducing-the-cerebras-cs-1-the-industrys-fastest-artificial-intelligence-computer/

Edinburgh International Data Facility update

Just before Christmas 2020, we took ownership of Computer Room 4 (cr4) at EPCC's Advanced Computing Facility. This will be the new home of the Edinburgh International Data Facility.

Around about the same time we unwrapped the server, storage and network hardware that comprises EIDF Phase 1: 45 petabytes of storage (from fast to archive-slow), 30 high-end compute servers, two 18 terabyte memory HPE SuperDome Flex systems and all the necessaries for an underpinning 20Gb/s core network.

We also took delivery of Europe's first Cerebras CS-1 wafer-scale AI compute engine, now hosted in one of the SuperDome Flexes (see page 6).

Throughout the first quarter of 2021 we've been racking and connecting up this smorgasbord of hardware to create the foundations of the Edinburgh International Data Facility (EIDF). Our job through the second quarter is adding the software layers on top to create the first incarnation of EIDF proper.

Much of EIDF is software defined, a virtualised cloud built around OpenStack and Kubernetes with shared storage and links to EPCC's existing high-performance computing services Cirrus and ARCHER2. If all goes well, a beta version of this "data science cloud" should be available this summer. At time of writing, everything is on course!

EIDF's first customers

The first users of EIDF will be the projects we've been supporting on development systems and older hardware for a year and more now. Over the second half of the year we will move virtual machines, data, and users across to the new EIDF environment, while keeping service disruption to a minimum.

The groups we'll be migrating include the Global Open Finance Centre of Excellence, the ISARIC4C COVID-19 research consortium (see p8), the Scottish Genome Partnership, the iCAIRD digital pathology research service, the Data SlipStream satellite data processing group, and key analytics services for the Scottish Government. Hot on their heels will come other stakeholders from the regional Data-Driven Innovation network including DataLoch, the National Collection of Aerial Photography, the regional IoT Research & Innovation Service network, and the new Research Data Scotland initiative.

2021 is a busy and exciting year for EIDF!

Rob Baxter, EPCC
r.baxter@epcc.ed.ac.uk

The Edinburgh International Data Facility brings together regional, national and international datasets to create new products, services, and research. It is funded by the UK and Scottish Governments under the Data-Driven Innovation network of the Edinburgh and South-East Scotland City Region Deal.

EIDF

www.ed.ac.uk/edinburgh-international-data-facility

Developing an outbreak analysis platform

EPCC is working with ISARIC4C, the Coronavirus Clinical Characterisation Consortium, to develop an integrated analysis platform that will be hosted on the Edinburgh International Data Facility.

The data analysis platform provides a unique combination of linked, curated data from UK sovereign data assets, together with a flexible high performance compute space. Created for COVID-19 research, the ISARIC4C data analysis platform combines the data safeguards of an NHS trusted research environment, with more than £100M of new exabyte-scale computational capacity at the home of the UK national supercomputer. This creates a unique opportunity to combine clinical, biological, genomics and virology research in a secure, openly-accessible framework.

Data held

The outbreak analysis platform was developed by ISARIC4C to encourage and facilitate research by collating, linking and curating clinical and research data, enabling deep integrative analyses of multi-omic disease profiling, stratified by viral variant, clinical phenotype and outcome.

This platform now serves as a hub for a coordinated UK national research response to COVID-19 and hosts data from several sources. The ISARIC research data within the analysis platform is already linked to NHS Scotland secondary care and death records and linkage to NHS England data is currently being incorporated, as well as other research data including COG-UK variant data, GenOMICC genome sequence data and UK-CIC phenotype data. Future plans include primary care, immunisation and ONS data.

Analysis platform structure

There are two routes of access to the analysis platform: 1. NHS Trusted Research Environment (Safe Haven) for access to personal clinical data and data collected without explicit consent. 2. Rapid-access flexible compute for access to non-disclosive research data collected with explicit consent.

Within both these environments there is an additional division in the data: 1. Publishable “open access” data which any user can use and report as they wish, according to data protection and privacy rules; 2. Embargoed active research data, shared by academic investigators and available for linked analysis but not for publication without agreement from all contributors.

This design is intended to build trust in order to encourage immediate contributions of research data from academic collaborators.

Future plans

Rapid addition of viral sequence data from the COG-UK platform will enable real-time detection of the clinical impact of new viral strains, in-depth biological study of reinfection, and host: pathogen interactions at a genetic and mechanistic level.

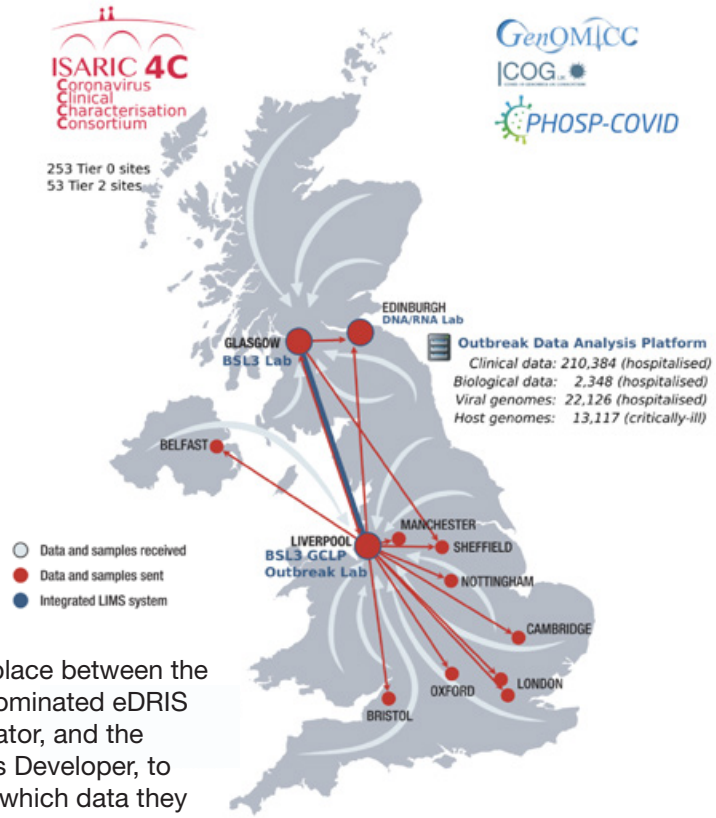
Kenneth Baillie,
Roslin Institute
j.k.baillie@ed.ac.uk

Lucy Norris, EPCC
l.norris@epcc.ed.ac.uk

Research outputs

ISARIC4C is the world’s largest observational study of hospitalised patients with COVID-19.

By generating, integrating and analysing clinical, biological, genetic and virological data on patients with Covid-19 in UK hospitals, ISARIC4C has provided vital information for policy-makers and health providers, including weekly updates to SAGE that guide the public health response. Specific areas informed include vaccine effectiveness and choice of therapeutic agents for clinicial trials.



EPCC's role

EPCC hosts the ISARIC4C and linked data inside the National Safe Haven, alongside the Scottish Covid-19 research database. The latter holds NHS Scotland datasets as well as ISARIC4C data and is used to extract linked data for researchers interested in COVID-related data on Scottish patients.

The main ISARIC4C database includes extracts from the Scottish database and extracts of NHS England datasets for English subjects enrolled in the ISARIC4C studies, received under approvals from NHS Digital. We also plan to include other research study data involving COVID-19 patients, notably the PHOSP and GenoMICC studies. Work is ongoing to ingest all these datasets with suitable anonymised identifiers that allow them to be linked to the ISARIC4C study data.

The ISARIC4C database is large, with over 2 million rows of 902 variables, and on average 11 rows of data for each patient. Cleaning scripts are also being incorporated into the platform, which creates additional outputs of aggregated and summary data as well as updating some outlying values.

Researchers requiring extracts of the data apply via eDRIS (part of Public Health Scotland), which oversees completion of appropriate training to use the Safe Haven and data agreements with the ISARIC4C management committee.

Discussion takes place between the researchers, the nominated eDRIS Research Coordinator, and the EPCC Applications Developer, to determine exactly which data they need and are permitted to have. The EPCC developer creates the extracts, runs checks and transfers the extract data with accompanying documentation to the project area where the Research Coordinator runs further checks before releasing the data to researchers. So far we have provided some ISARIC4C-only data extracts and some linked to Scottish NHS data and variant of concern data. There are ongoing preparations to link to English NHS and other study data soon.

One of our largest challenges over the summer will be the ingest of up to 2PB of genomic data, both viral and host, from the GenOMICC and COG-UK studies. These datasets will be hosted in a secured area of the new Edinburgh International Data Facility (EIDF) and will be supported by significant computational capability from one of EIDF's new HPE SuperDome Flex large-memory systems.

Getting all the data into the Safe Haven and wider analysis platform has been an ongoing challenge, with many stakeholders involved, but this is an exciting project with huge potential for increasing collaboration with researchers. There will be enormous scope for linking different data of interest to a great variety of specialists.

EIDF
Edinburgh International Data Facility

ISARIC 4C

EIDF
www.ed.ac.uk/edinburgh-international-data-facility

ISARIC4C
https://isaric4c.net

On-demand HPC straight to your desktop

We offer direct, secure access to our high-performance computing and data resources.

EPCC provides an on-demand HPC service that brings leading edge supercomputing capability directly to your desktop. Through a simple internet connection you gain cost-effective access to an impressive range of world class high-performance computing (HPC) resources, such as ARCHER 2, Cirrus, and the Edinburgh International Data Facility.

Our on-demand HPC service is targeted at engineers and scientists solving complex simulation and modelling problems in fields such as bioInformatics, computational biology, computational chemistry, computational fluid dynamics, finite element analysis, life sciences and earth sciences.

How it works

Once set up with access to our on-demand HPC service you choose how much resource to use and when. Users have complete control over managing their project resources – jobs, disk space, and archiving. This usage model is based on pay-per-use service and provides an excellent alternative to purchasing expensive HPC hardware. The service can be used to satisfy contingency and peak demand requirements.

Our on-demand service provides direct, secure access to EPCC's HPC platforms, delivering the highest levels of performance. The highest levels of data security are provided, and the service is administered directly by you using a range of administration and reporting functions.

The service is fully supported with an integrated help desk. EPCC support staff are also available to help with usage problems such as compiling codes and running jobs. Access to further technical support is readily available via our large Applications Group with a breadth of software development experience across HPC disciplines including expertise in parallel and distributed computing (using MPI, Open-MP and Web services) that is internationally recognised. Our software developers program on Linux, Windows and in various Unix environments using .NET, Fortran, C, C++, Java, XML, Perl, Python, R, and numerous other languages.

Thomas Blyth, EPCC
t.blyth@epcc.ed.ac.uk

Access to our systems enables a significant competitive advantage for our partners, and our highly experienced staff bring a wide range of expertise to important areas of HPC & data science research and development.

For more information please contact EPCC's Commercial Manager, Thomas Blyth: t.blyth@epcc.ed.ac.uk

Accelerating high performance data pipelines

Illuminate, a commercial partner of EPCC, has a mission to accelerate informed decisions by providing its customers with the means to find specific needle in a haystack of data points from the mass of its network traffic. As networks evolve, it is continuously innovating to ensure its data pipelines operate at maximum efficiency.

Doug Carson, Illuminate
contact@oneilluminate.com

Data analytics pipelines have three main functions: data ingestion, processing, and analysis. The data ingestion stage focuses on obtaining or extracting data from a range of data sources and importing it into systems that can enable processing and analysis. This is followed by data processing to validate, transform and enrich it, and load it into some form of queryable data storage, such as a data lake. Finally, we utilise Artificial Intelligence/Machine Learning processes to build models from the data in the lake and use them to gain insight.

The data ingestion stage was the subject of a recent research project, with EPCC and Intel, where we investigated how network sensor data capture could be accelerated in a cloud environment. While concluding this work, the Cirrus system was upgraded to include 36 GPU nodes, each with 4 Nvidia V100 GPUs. As the data processing and loading stage is a prime candidate for GPU acceleration, we were keen to take the opportunity to run benchmarks on EPCC's Cirrus platform to investigate what opportunities these GPU resources provided for reducing processing and analysis runtimes and thereby generating quicker insight for customers.

Public clouds offer considerable compute resources so why did Illuminate favour a high-performance computing (HPC) facility for our investigation? Our experience is that cloud high performance data pipelines are often I/O bound and it takes considerable system engineering effort to tune the network and storage middleware in cloud deployments to obtain good performance for such pipelines. An HPC facility will offer a high-performance communications fabric coupled to a parallel filesystem that is already tuned to support the needs of the cluster for I/O intensive applications. An HPC service like Cirrus can provide the support required to get applications ported, running, and optimised on the system. Cirrus' excellent documentation meant that only one support request was needed when installing and benchmarking our code.

To evaluate performance, we benchmarked the Cirrus V100 GPUs against some local CPU resources, and on NVidia T4 GPUs (through Google Colab notebook). The V100 GPU provided more than 10x the performance of CPU instances running on a local server and double the performance of a T4 GPU on Colab.

The success of this initial benchmarking has prompted a wider investigation into the application of massively parallel acceleration of data analytics for Illuminate workloads. This work to accelerate high-performance data pipelines once again sees us having the opportunity to use the unique supercomputing facilities at EPCC to ensure that our pipeline remains at the leading edge of data engineering.

Illuminate used Cirrus for this investigation. Cirrus is an EPSRC Tier-2 National HPC Facility which is available for users in both academia and industry.

Housed at EPCC's Advanced Computing Facility, Cirrus is a flexible, state-of-the-art high performance computing system that provides an ideal platform for users to solve their computational, simulation, modelling, and data science challenges.

Cirrus HPC service
https://www.cirrus.ac.uk

Illuminate
https://oneilluminate.com

EuroCC@UK: a new National Centre of Excellence for HPC-based research

The European Commission has launched a network of National Centres of Excellence in high-performance computing (HPC), high-performance data analytics (HPDA) and artificial intelligence (AI). The UK centre will be delivered jointly by EPCC and Hartree Centre.

The objectives of the EuroCC network are to increase the accessibility and availability of HPC, HPDA and AI in each centre's country, and provide a single "front door" into HPC research and development capabilities for industrial customers and academic researchers alike.

The adoption of a strategy based on national centres reflects the differing levels of maturity across Europe of these technologies, allowing countries to focus on their own priorities. In addition to the work done at the national level, the NCCs will also exchange knowledge with each other via the EuroCC network. This will help to establish best practice and elevate the quality of research and innovation in HPC, HPDA and AI to a common high standard across Europe.

Together with the Hartree Centre, EPCC will offer training, access to facilities and an outreach programme to raise awareness of the capabilities of the technologies. We will also promote take-up of HPC, HPDA and AI within industry by providing expertise and small pilot projects.

HPC has a proven track record as a tool for science and industry that delivers impact. Its use has steadily grown over the past three decades into a mature technology that supports many of Europe's most important sectors, including engineering, health, climate and energy production. HPC in combination with data-analytics and artificial intelligence is set to enable further innovation by enriching iterative modelling techniques. Combining simulation, data analysis, visualisation and machine learning will allow large and highly complex problems to be tackled. Realising the impacts of this at the national level is the goal of the UK NCC.

Training

For the training component of the EuroCC project we are investigating the best ways to produce accessible material. Initially we are looking at the suitability of the Software Carpentries style of teaching (note we are not making

Thomas Blyth, Stephen Farr, and Mark Sawyer, EPCC

t.blyth@epcc.ed.ac.uk
s.farr@epcc.ed.ac.uk
m.sawyer@epcc.ed.ac.uk

Image: Nico El Nino via Getty Images

Software Carpentries courses, rather using its website templates and methodology). With this style, instead of the traditional lecture followed by a practical session, the content is served up in one coherent web page which the trainee follows, learning and completing the exercises in a continuous manner. These courses can either be delivered by an instructor, where the trainee and instructor work through the material at the same time, or provided as a self-service course. Due to the web-based nature of content is it easier to keep material up to date and include different types of media such as videos, images, and correctly formatted example code. Furthermore, the accessibility is vastly increased as everything can be accessed via a web browser.

Industry collaboration

As part of its role in the EuroCC project, EPCC is working with several companies to leverage the potential of HPC by developing project proposals for further funding and inward investment. One such collaboration commenced earlier this year with Edinburgh based start-up, Danu Robotics.

Danu Robotics has ambitious plans to disrupt the recycling and waste management industry with a novel blend of AI and robotics. One of the biggest challenges facing recycling and waste management is unsorted contaminant such as nappies, miscellaneous plastic, and other non-recyclable materials which still rely heavily on outdated manual processes that carry various economic and safety risks. Danu Robotics aims to develop an AI driven, vision guided, distributive robotics system to automate this process and revolutionise the recycling and waste management sector.

We are working alongside Danu Robotics to further develop its data architecture and novel AI processing models that will underpin the company's long-term business goals. EPCC is providing technical expertise as part of the engagement alongside prototype testing and these form the basis of a larger project.

"Working alongside EPCC has been fundamentally important for us to design and build a sustainable and scalable IT infrastructure that can support our product development and the growth of the company. The expert knowledge provided by EPCC has been tremendously valuable to us."
Xiaoyan Ma,
Danu Robotics
CEO and founder

EURO

Find out more about the EuroCC network:
www.eurocc-access.eu

New UK National Supercomputing Service

The ARCHER2 Service is a world class advanced computing resource for UK researchers. Hosted and managed by EPCC, it replaces ARCHER, which operated as the UK's National Service for seven years.

The new ARCHER2 service should be capable on average of over eleven times the science throughput of ARCHER, based on benchmarks which use five of the most popular codes on ARCHER.

ARCHER2 is an HPE Cray EX supercomputer and there are several changes from the previous ARCHER system (a Cray XC30 supercomputer) that might have a large impact on application development and compilation, user workflows, and application performance. For example, there is an increase in the number of cores per node from 24 on ARCHER to 128 on ARCHER2, and a change from the Cray Aries interconnect on ARCHER to the new HPE Cray Slingshot interconnect on ARCHER2.

As many readers will know, the ARCHER2 hardware has unfortunately been delayed, with the primary reason relating to scaling of the operating system. While we share the disappointment we know many will feel, there is reason for optimism: users have been successfully operating on a 4-cabinet system and in late June a team of EPCC staff began porting and testing a series of applications and benchmarks on the full system. It is very early days but this is real progress.

EPCC was awarded the other components of the ARCHER2 service, namely the Service Provision (SP) and Computational Science and Engineering (CSE) services following a procurement exercise delivered by UKRI. These contracts involve a range of activities designed to provide users with the best possible support, to ensure researchers are able to maximise the science achieved on ARCHER2.

A selection of activities are described on page 15.

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

Training

There is a significant training programme in place, with a commitment to deliver 60 days of training a year. Courses on offer range from introductions to ARCHER2 and HPC, through to advanced programming techniques, and training for specific application codes.

While the expectation was that courses would be delivered face to face, COVID-19 has led to them being held online. Some positives have been observed here, with online and self-service courses offering greater flexibility and accessibility to users.

Embedded CSE (eCSE)

The eCSE programme is up and running successfully, offering a regular series of funding calls that support Research Software Engineers (RSEs) across the country to enhance and develop application codes for ARCHER2.

ARCHER2 Helpdesk

The Service Desk is available Monday–Friday and provides a helpful, friendly, and responsive face to the service. The in-depth support team assists users with technical problems, working closely with them to ensure they are able to utilise the service.

Publicising research

Finally, the success of ARCHER2 is driven by the success of the science completed on the system. If you have completed a piece of research on ARCHER2 or its predecessor ARCHER and would like to publicise your work, please get in touch. We are keen to showcase a range of science produced on the system.

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

ARCHER2
www.archer2.ac.uk

ARCHER

www.archer.ac.uk

First year of the ARCHER2 Training Service

A fully online ARCHER2 training programme has been delivered this year without major issues, despite the COVID-19 pandemic.

Juan Rodriguez Herrera, EPCC
j.herrera@epcc.ed.ac.uk

During the ARCHER training service, courses were delivered face-to-face, with a few courses and a programme of virtual tutorials delivered online. As the ARCHER2 training service began in the middle of a pandemic, teaching was fully switched to online delivery. The ideal situation would be to offer both approaches, and we hope to offer face-to-face courses again when possible.

Since April 2020, attendees have joined us from all regions of the UK for webinars on a wide range of topics. The average rating given for online courses by attendees is above 4 on a scale of 1-5.

Delivery mechanisms

Content design and delivery has been adapted to online teaching. For instance, The Carpentries workshops are delivered in four mornings rather than two full days. These workshops are usually divided into four topics with each delivered on a different day. It is important to stick to the schedule, have breaks, and not deliver courses on consecutive days if possible.

We have used Blackboard Collaborate software for live broadcast courses, virtual tutorials, and consultancy sessions. Material from all online courses, including videos of the lectures, have been made freely available on the web after the live run. We have also run an Etherpad server to provide collaborative course notebooks for all our courses to allow attendees and trainers to share information, questions, and ideas.

It is not a perfect approach, but one of the advantages of online training is flexibility: attendees and instructors can participate without the need to travel, which may account for the increase in attendance from outside the UK. However aspects like socialising or problem-solving are more challenging online. When you teach, you observe body language reactions and get instant feedback on the teaching level and pace.

Whenever appropriate, ARCHER2 access has been offered before courses. This could be to allow an external training provider to test their materials, or for course

Introductory courses

There will be introductory courses for different user categories:

Package Use on ARCHER2 (1 day)
Efficient use of pre-installed research software packages on ARCHER2.

Development on ARCHER2 (2 days)
The ARCHER2 application development environment, core parallel and scientific software libraries, available debugging and profiling tools.

Data Science on ARCHER2 (1 day)
The essentials of ARCHER2, the basic use of core data science packages on ARCHER2, and data handling best practice on ARCHER2.

An in-person Carpentries course hosted by the University of Edinburgh.

attendees who may wish to port codes in preparation for a course on ARCHER2 tools. We also offered ongoing machine access after the course ended for users to practise their new skills. A follow-up virtual consultancy session around 2-4 weeks after each course allows attendees to ask further questions once they have applied the course material to their research.

At the beginning of the training programme, we increased the number of webinars ("virtual tutorials") to reach a larger audience and mitigate the lack of face-to-face courses. In addition, we introduced self-service courses, where each user follows course materials at their own pace. A self-service course on OpenMP was released in February, and in May an MPI self-service course was released.

What we offer

Recognising that increasing numbers of researchers are using pre-installed research software packages, the training programme has been designed to address the training requirements of users with different needs and levels of experience. It is vital to have a coherent programme of initial courses that work together to meet these needs.

The ARCHER2 training programme directly address the use of centrally-installed research software packages by creating a set of courses to meet the needs of users of HPC research software as well as those who develop it.

Summary

The training plan for the second year of the ARCHER2 service has been tailored considering feedback from course attendees, the ARCHER2 User Training Forum, and the ARCHER2 Training Panel. New courses are added annually to make the course catalogue as much complete as possible. The programme will be delivered online until potential resumption of face-to-face training becomes clearer. Our plan still has the same aims: to address the needs of all users and enable them to make efficient use of the full ARCHER2 system as soon as it is available later this year.

ARCHER2
www.archer2.ac.uk

Supporting ARCHER2

Installing the first phase of ARCHER2 earlier this year.

The HPC Systems Team provides the System Development and System Operations functions for ARCHER2.

One of our main responsibilities is managing the Advanced Computing Facility (ACF) where ARCHER2 is hosted, which we do in coordination with colleagues at the University of Edinburgh Estates department. A system such as ARCHER2 requires a lot of cooling and power, and much effort goes into making sure these are properly provided.

We also play a key role in the development and deployment of ARCHER2, as well as its daily operations.

Infrastructure

We provide a number of services for ARCHER2 including NTP servers, (which ensure the accuracy of the system clocks), authentication servers, (which store and provide details of the users on the system), and various servers to monitor and record the state of the service.

We also provide the networking to integrate the system into our site and provide connectivity to the broader internet. This year we will move to a new network that will offer ARCHER2 100Gbit/s connectivity to the broader internet (about 1500 times the average UK broadband) and 200Gbit/s to other systems and services at the ACF.

Deployment

HPE Cray provide ARCHER2 in a "vanilla" state and we put in place customisations and configurations to best support our users.

This includes deploying appropriate configuration to the system scheduler, implementing a variety of monitoring so that our site systems can keep track of the state of the system, deploying directory structures to various file systems, and implementing the ticketing infrastructure needed during operations.

Operations

This is a wide area of responsibility that is critical to keeping the service running. During working hours a member of the team is responsible for monitoring the system at all times so we can deal with problems as they emerge.

We are responsible for processing tickets on the service. These tickets carry out all creations of user accounts, quotas, and directories as well as any changes to these details. Tickets are created and issued by the SAFE system, which is accessed via a web portal to create and manage a user's account. We also troubleshoot and investigate a lot of user problems reported to the service desk.

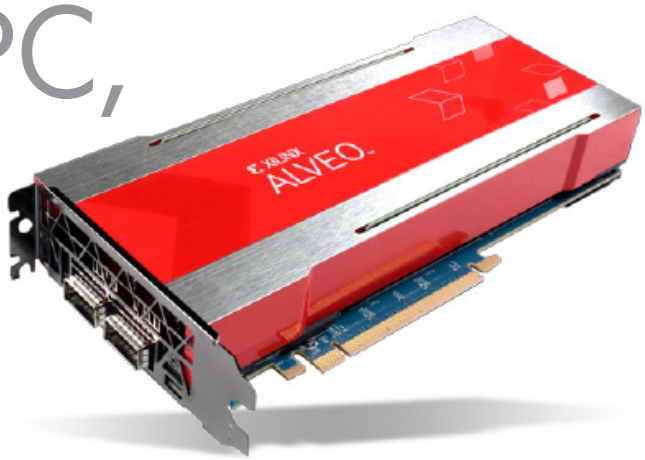
We've been working towards the launch of ARCHER2 and we will be excited for users to get their hands on everything we've been working on!

Kieran Leach, EPCC
k.leach@epcc.ed.ac.uk

We are a team of fifteen System Administrators and Developers who deploy, manage, and maintain the services and systems offered by EPCC, as well as the infrastructure required to host and support all of EPCC's services and systems.

Computing facilities hosted by EPCC
www.epcc.ed.ac.uk/facilities

FPGAs for HPC, but this time it's different



For over a decade our community has enjoyed significant performance benefits by leveraging heterogeneous supercomputers. Whilst GPUs are the most common form of accelerator there are also other hardware technologies which can be complementary.

Field Programmable Gate Arrays (FPGAs) enable developers to directly configure the chip, effectively enabling their application to run at the electronics level. There are potential performance and power benefits to tailoring code execution and avoiding the general purpose architecture imposed by CPUs and GPUs, and as such FPGAs have been popular in embedded computing for many years but have not yet enjoyed any level of uptake in HPC.

There have been several efforts over the years to make FPGAs more mainstream, however factors such as reliance on esoteric programming technologies proved to be barriers to entry. But in the last few years there have been some very interesting developments. Vendors have significantly advanced the hardware offering, providing larger, more capable chips, with some very exciting next-generation FPGAs to be released later in 2021. They have also invested heavily in the software ecosystem, with much improved programming environments and documentation. As shrinkages in CPU process size and the associated performance benefits are slowing significantly, it is worth reviewing how FPGAs can help accelerate HPC computational kernels.

To this end, the ExCALIBUR Hardware and Enabling Software (H&ES) programme has awarded us funding for an FPGA testbed. H&ES addresses the challenges and opportunities offered by computing at the exascale, and is looking to invest in novel hardware testbed systems upon which UK scientific software developers can explore their codes.

EPCC's FPGA testbed will be run in collaboration with University College London and the University of Warwick, and will be physically located in Edinburgh. It will represent a complete ecosystem for FPGA programming: in addition to the hardware itself, all required libraries and licences will be pre-installed, and all the building and emulation nodes, training, and full documentation will be provided. There is also associated effort to develop an enabling software ecosystem, further lowering the barrier to entry for application developers, for instance integrating existing FPGA tools and libraries and developing new ones to support users.

We are currently in the process of installing the hardware, and in addition to the current generation FPGAs we are excited to be hosting Xilinx's next-generation Versal ACAP architecture later in the year.

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

For more information and to sign up, visit our website: https://fpga.epcc.ed.ac.uk

Nick's earlier work exploring the potential role of FPGAs in accelerating the Met Office's HPC codes is described in his blog post "Using FPGAs to model the atmosphere": http://bit.ly/2E6aybl

HPC-Europa3: Transnational access in the time of Covid-19

HPC-Europa3 is an EC-funded Transnational Access programme for collaborative research visits using high performance computing. The last year has been a challenging one for the programme, with pandemic travel restrictions meaning only a few visits were able to take place, and in rather unusual circumstances. Here we hear from a current visitor, and a former visitor who has recently relocated to Edinburgh in the middle of the pandemic.

Adrián Bernal Bermejo arrived in Edinburgh in April 2021 for a 3-month HPC-Europa3 visit.

I am pursuing a PhD in cloud computing modelling and simulation at Castilla-La Mancha University in Spain. I have come to Edinburgh to visit EPCC and to work with my host, Dr Rosa Filgueira (formerly of EPCC, now of Heriot-Watt University), for two reasons. The first is that EPCC manages large amounts of computing resources, which allows me to learn a lot about such systems and to use them for my simulations. The second is Rosa's experience of managing workloads in distributed systems and her experience in collaborating with companies. That makes her the right person to help guide me through the final stage of my PhD.

I had to postpone my visit several times. But noting that the pandemic was giving us a "soft break" and that the government was planning to ease the lockdown, I decided to start my stay and not miss the opportunity. I had to take several Covid-19 tests and have everything prepared and booked before I travelled and then self-isolate for ten days on arrival. Catherine Inglis and Mario Antonioletti from EPCC helped me by giving me all the information I needed and even provided me with an extra monitor so I could work more comfortably from my accommodation!

Once I could go outside, I was able to enjoy the city quietly and without crowds. After a short time, most establishments started to open their doors. So, I am finally enjoying this beautiful city in every way.

Marcos Maroñas was an HPC-Europa3 visitor in 2019, when he was a PhD student at Barcelona Supercomputing Center (BSC). Marcos has now completed his PhD and recently relocated to Edinburgh – in the middle of the pandemic – to take up a new job.

I am a Software Researcher at Edinburgh's Huawei Programming Languages Research Laboratory. I received my bachelor's degree in Computer Science in 2015 and my MSc degree in Innovation and Research in Informatics, specialising in High Performance Computing and Computer Architecture, in 2017, both from Universitat Politècnica de Catalunya (UPC).

Catherine Inglis, EPCC
c.inglis@epcc.ed.ac.uk

Adrián Bernal Bermejo (left) and Marcos Maroñas during their visits to Edinburgh as part of the HPC-Europa3 programme.

I joined the Programming Models Group of Barcelona Supercomputing Center in 2014, where I spent more than six years doing research which was mainly focused on parallel programming models. During my time at BSC, I worked on my PhD thesis “On the Design and Development of Programming Models for Exascale Systems”, and was awarded my PhD in 2021.

In my thesis, I explored different innovative features that state-of-the-art parallel programming models and their runtime systems should include in order to adapt to the future Exascale systems. I also explored new programming models that can ease the use of fault tolerance techniques, which will be crucial in Exascale Systems.

During my PhD, I was awarded an HPC-Europa3 grant to visit Dr Mark Bull at EPCC. Along with my supervisor Dr Vicenç Beltran, we were working on a new feature called worksharing tasks in the OmpSs-2 programming model (similar to OpenMP), and needed to benchmark it to understand its potential benefits. Dr Bull is a leading expert in benchmarking, author of several synthetic benchmark suites for OpenMP, hybrid OpenMP+MPI, and Java. As such, Dr Bull was the perfect person to help me with the benchmarking process, given his knowledge and experience. The project was a complete success, and culminated in a publication for a highly-reputed international conference [1].

However, for me the benefits of the programme were not just the opportunity to work with one of the most experienced experts in benchmarking, nor the boost this collaboration gave to my PhD thesis, but the doors that opened thanks to my stay in Edinburgh.

I was contacted by Huawei to join its Programming Language Research Laboratory, and returned to Edinburgh in February 2021. Currently, my research in the Huawei PL Lab focuses on the design and development of innovative and exciting features for modern programming languages that can help bring Huawei’s products to a whole new level.

[1] M. Maroñas , X. Teruel, M. Bull, E. Ayguadé and V. Beltran, ”Evaluating Worksharing Tasks on Distributed Environments,” 2020 IEEE Cluster Conference (CLUSTER).

The next closing date for applications is 9th September 2021.

More information: www.hpc-europa.org

HPC-Europa3 receives funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no.730897.

More lockdown visit stories from visitors Sudip Kumar Mondal and Alexey Bobrick can be found at : www.epcc.ed.ac.uk/blog/tags/hpc-europa.

ARCHER: the UK's National HPC Service 2013-2021

The ARCHER system was finally switched off in January 2021. Funded by EPSRC and NERC, this remarkable service was the UK's National HPC Service for over seven years.

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

ARCHER began life in late 2013. Over the whole lifetime of the service, just over 5.7 billion core hours were used for computation, with over 5.6 million jobs submitted. That's a lot of resource for UK science.

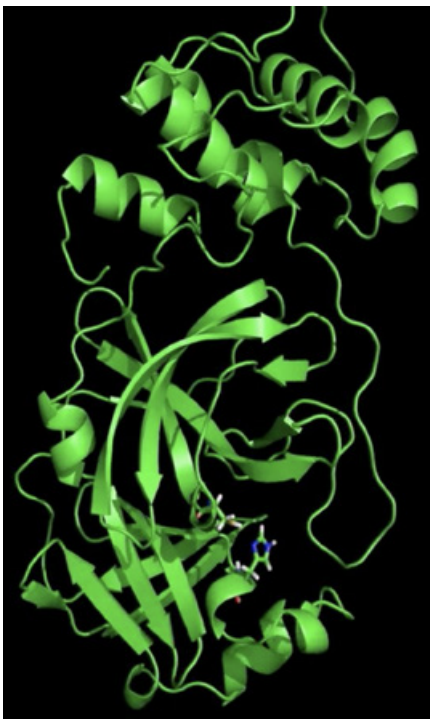
Many people have benefited from ARCHER, with around 3300 users, with PhD (or other post-graduate research degree participants) forming the largest cohort. This is closely followed by post-doctoral researchers.

While ARCHER entered service at number 20 in the TOP500 list, the goal has always been to deliver world class science by supporting an active and diverse community of users from across a wide spectrum of science areas. The range of science carried out on the service has been significant with the software used being quite broad.

To give an idea of the spread, the top 10 applications by use account for around 55% of the total use, the top 20 applications by use account for around 68%, and the top 40 account for around 78% (we identify use from 68 research software packages). Around 20% of use is not associated with known research software packages. The five most heavily used codes (VASP, GROMACS, the Unified Model, CASTEP and CP2K) have all utilised over 100 million core hours each.

EPCC supported the service during the lifetime of ARCHER, which including handling queries, offering training and providing in-depth user support. The User Support and Systems teams (SP) resolved 57,489 contractual queries with 98.6% of them completed within two days. The ARCHER team has also created 6,926 users accounts, which have formed our very broad user base. This ranges from our novice Driving Test users to our very experienced users.

ARCHER served its users well. There was a steady utilisation over the lifetime of the service, with it effectively operating around maximum capacity for the majority of the lifetime.



Metadynamics simulations opening the active site of the SARS-CoV-2 main protease. Image by Mohamed Ali Al-Badri & Khaled Abdel-Maksoud, King's College London, Department of Physics & University of Southampton, School of Chemistry.

Science enabled by ARCHER has been showcased through a series of case studies, highlighting work around many of society's grand challenges. Most recently, ARCHER has been used in the fight against COVID, for example through pandemic modelling. You can explore the case studies on the website at: https://www.archer.ac.uk/casestudies/

The entries to our annual image and video competition is also hosted on the ARCHER website. Last year's winner is shown opposite. https://www.archer2.ac.uk/about/gallery/

Throughout ARCHER's lifetime we aimed to provide an accessible and inclusive service. For example, providing accessible training materials; founding the Women in High Performance Computing initiative; inviting early-career observers to eCSE panel meetings; and delivering a set of outreach activities aimed at under-represented groups.

Back in February 2015, a little over a year after the ARCHER service began, EPCC launched an entirely new and innovative access route, for users to have time on ARCHER via a "Driving Test". The idea behind this was to help those who had never used ARCHER, or possibly any high performance computing (HPC) resource, to gain a small amount of time to familiarise themselves with the HPC environment. The Driving Test has proven to be extremely popular and successful. Since its launch almost five years ago, we have had 587 people pass the test and 455 of them have signed up for an account. 70,460 jobs have been run, using 279,638 kAUs in total.

So as we look back on the ARCHER service we also look forward to ARCHER2 (see p14), which I am confident will see the UK HPC community continue to deliver world-class science.

Many thanks to EPCC's Jo Beech-Brandt, Clair Barrass, and Andy Turner for supplying statistics and other contributions for this article.

archer

ARCHER website
https://www.archer.ac.uk

ARCHER2 website
https://www.archer2.ac.uk

Quantum computing at EPCC

Google's quantum supreme cryostat with Sycamore inside. Photo Credit: Eric Lucero/ Google, Inc.

Computing doesn't come much more novel than quantum computing.

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

Classical computers rely on the manipulation of bits, physical systems (usually transistors) that can be found in one of two states, which we label 0 and 1. Quantum computers on the other hand, use quantum bits (qubits) which can be measured in one of two physical states, still labelled 0 and 1, but can exist in any linear superposition of the two – they can be in 0 and 1.

Such a fundamental change at such a low level means that programming a quantum computer, or even developing quantum algorithms is very different to what we're used to.

Quantum computers aren't likely to replace ARCHER2 any time soon (some things just work better on a classical computer), but quantum computing can achieve impressive speedup of certain tasks. Google has already demonstrated the so-called quantum advantage with a quantum computer with just 53 qubits[1]. Admittedly, it used a contrived application – its quantum computer was tasked with generating a random quantum circuit – but it's clear that if we want to stay at the forefront of computing, now is the time to get involved in quantum.

For this reason, EPCC has joined a collaboration of local experts in quantum computing with members from the Universities of Strathclyde and Glasgow, as well as The University of Edinburgh's School of Informatics. The collaboration's expertise spans both hardware and software, with EPCC contributing our wealth of experience in integrating novel compute devices into traditional HPC, as well as links to potential end-users, and substantial classical compute resource for emulation. We have recently demonstrated emulation of a 41-qubit quantum computer using 512 nodes of ARCHER2[2].

The focus of the collaboration will be on developing applications of quantum computing, with the intent of supporting end-users across industry and academia. We are seeking out practical uses for current generation noisy intermediate-scale quantum (NISQ) computers, as well as looking further ahead at what could be achieved with fault-tolerant quantum computing.

Ultimately, we would like to see quantum computers alongside EPCC's traditional classical supercomputers, ready for exploitation by industry, academia, and commerce. Perhaps one day we will rebrand as EQCC?

Footnotes

1. Arute, F., Arya, K., Babbush, R. et al. Quantum supremacy using a programmable superconducting processor. Nature 574, 505–510 (2019). https://doi.org/10.1038/s41586-019-1666-5

2. Oliver Thomson Brown, Quantum Computing Without A Quantum Computer, ARCHER2 Webinar, (2021). https://www.archer2.ac.uk/training/courses/210331-quantum-webinar/

eqcc

Consortium website
https://sciqcs.org

Video: "Quantum Computing Without A Quantum Computer"
https://bit.ly/epccquantumcomputing

GPU hackathon

EPCC held its first GPU hackathon earlier this year in partnership with NVIDIA, hosting 28 participants and 20 mentors across seven teams.

The event was held virtually due to the ongoing Covid-19 pandemic. Using Zoom and Slack, individual teams were able to work alongside mentors in separate breakout rooms and channels.

Team members originated from universities, companies, and HPC centres across the UK as well as the rest of Europe. The projects they provided involved a range of fields including computational fluid dynamics, machine learning, solid-state physics, molecular dynamics, astrophysics and materials science.

Teams developed and programmed their codes on EPCC's Cirrus platform, which has 144 NVIDIA Volta V100 GPUs available.

Each team had different goals for the hackathon, involving parallelisation from scratch, porting from CPU to GPU, scaling up to larger numbers of GPUs or optimising existing GPU codes. The languages and frameworks used varied from FORTRAN with OpenACC to higher-level frameworks like PyTorch. NVIDIA engineers were on-hand providing Q&A sessions (in particular for tools like the NSIGHT profiler), as well EPCC staff to provide support for any hardware or software issues that arose.

By the end of the hackathon, several teams reported they were able to achieve between 2x to more than 8x speedups over original versions of their codes.

Larisa Stoltzfus, EPCC
l.stoltzfus@epcc.ed.ac.uk

Dates of future training events can be found via the EPCC website.
https://bit.ly/3j07d38

Image: Paul Dodds

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 both Times Higher Education World University Rankings 2021 and QS World University Rankings 2021.

"The MSc in High Performance Computing course had a really good practical focus which allowed me to develop my HPC programming skills and get hands-on experience working on applications using the skills I learned during the MSc on a daily basis."
Holly Judge
MSc student 2018/2019
Now an Applications Developer at EPCC.

"With one of the most experienced teaching and research staff, combined with its cutting-edge supercomputers, EPCC enabled me to establish the fundamentals of HPC and Data Science and to build a solid base for my further academic pursuit."
Liang Liang
MSc student 2019/2020
Now a PhD student at Imperial College London.

www.epcc.ed.ac.uk/msc