

Building world-class infrastructure for national resilience

In this issue

ARCHER2 Phase1 installation

ExCALIBUR: The Exascale programme of the UK Strategic Priorities Fund

Edinburgh International Data Facility

Global Open Finance Centre of Excellence

Supporting the UK COVID-19 response

From our Director

Welcome to the Summer 2020 edition of EPCC News. Like everyone in Europe, EPCC has had to cope with the effects of the coronavirus pandemic as best we can. I've been incredibly impressed by the way the team has settled into working from home – often juggling domestic and space challenges.

Even without the pandemic, this would have been one of the busiest periods EPCC has ever had. Having won the ARCHER2 SP and CSE contracts, we are in the middle of bringing the new National HPC service to life. At the same time we have installed a major upgrade – 144 top-end GPUs and new fast storage – to the Cirrus Tier-2 HPC service and also concluded the £100m EIDF IT infrastructure procurement which has been won by Hewlett Packard Enterprise.

The City Deal Data Driven Innovation programme is also gathering pace. In the next issue of EPCC News I hope to report on the opening of our new computer room at the ACF which has only been delayed a few weeks due to the pandemic – our contractors, Robertsons, returned to the building site in early May to continue the project having implemented

numerous safe working practices to allow this.

In regard to the pandemic, we have worked very closely with eDRIS (part of NHS Scotland) to create a COVID-19 data repository within the National Safe Haven. This was available less than 4 weeks after lockdown and is currently supporting over 20 COVID-19 research projects. We are also supporting some of the largest UK-wide research projects such as ISARIC. This has been a major activity for EPCC and will continue to be so for the foreseeable future.

We intended to hold a major celebration of EPCC's 30th birthday this year. I hope this issue demonstrates that we are as busy and successful now as we have ever been. Although we haven't been able to hold a face-to-face celebration, we will look back on the past 30 years in the next issue of EPCC News.

I hope you enjoy reading this issue, and look forward to a time when we can once again meet with our many friends and partners in the Bayes Centre and at meetings world-wide.

Mark Parsons
EPCC Director

Continuity of EPCC services

At EPCC's Advanced Computing Facility (ACF) data centre we host and manage a number of National Services, including ARCHER, the DiRAC "Tesseract" system, the Tier 2 Cirrus system, and the National Safe Haven. We know how important keeping these services running is to the research community.

To adhere to government restrictions during the pandemic, we have adopted a slightly reduced, condensed working pattern. However, the ACF is open and

functioning as normal.

We have a number of well-rehearsed business continuity and disaster recovery contingencies in place. Our staff work from home where required, and a rota of on-site resources ensures the maintenance and upkeep of the site and our systems to the best of our abilities. The service helpdesks operated by EPCC continue to function as normal.

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

Contents

- 3 COVID-19**
Supporting the national response
- 4 ARCHER2 Phase1**
Installing the new UK service
- 6 ARCHER2**
EPCC awarded service contracts
- 7 Tier-2 service upgrade**
CIRRUS Phase II will run to 2024
- 8 Exascale**
EPCC joins UK Government's ExCALIBUR programme
- 10 Edinburgh International Data Facility**
Update on the building of this unique data infrastructure
- 12 Global Open Finance Centre of Excellence**
Using financial data for societal and economic benefit
- 14 Working with business**
Recent collaborations
- 15 Exascale for engineering**
EXCELLERAT programme launches Service Portal
- 16 Cloud technology for business**
Collaboration with Intel and Illuminate Technologies
- 18 Data accreditations**
EPCC gains accredited processor status
- 19 Training: ARCHER2**
New programme launched
- 20 HPC research visits**
HPCEuropa3 brings together researchers, HPC experts and world-class facilities
- 21 Software training for researchers**
Carpentries moves online
- 22 EPCC's MSc programmes**
Our MScs prepare students for real-world challenges
- 24 Software Sustainability Institute**
Collaborations workshop report

Supporting the national COVID-19 response

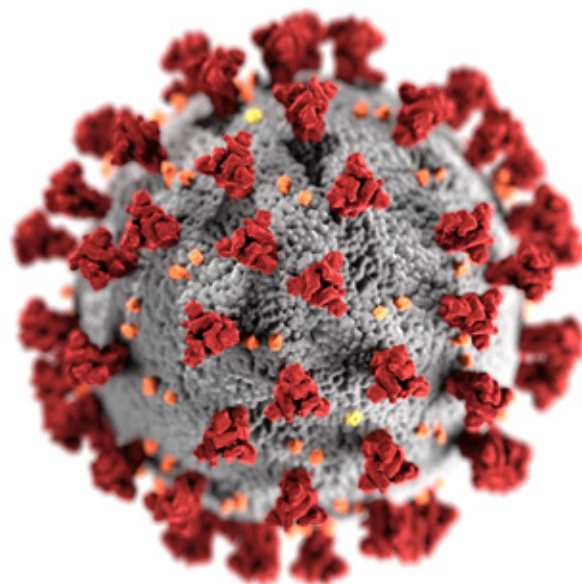


Image: CDC/ Alissa Eckert, MS; Dan Higgins, MAM/Public domain

EPCC has been working closely with public agencies to support responses to COVID-19.

EPCC has been working with the NHS and Public Health Scotland (PHS) to create a secure data and computing environment for urgent research in Scotland into COVID-19.

Collaborating closely with our colleagues in the eDRIS team from PHS, we've assembled a major data resource inside the National Safe Haven (see page 10). This resource, the Scottish COVID-19 Data Repository, brings together COVID-19 testing data with key clinical datasets to provide approved health researchers with the tools they need to understand SARS-CoV-2 and COVID-19.

The National Safe Haven is a logical part of the Edinburgh International Data Facility, but physically separate from it. It provides a secure data management zone with highly restricted access, and a trusted research environment for approved researchers to work with designated linked health datasets.

Scottish COVID-19 Data Repository

The Scottish COVID-19 Data Repository has been assembled within the secure data management zone over the last few months, and is now operational and receiving weekly updates of testing and clinical data from eDRIS. Research projects authorised under a COVID-19 fast-track approvals process by the Scottish national Public Benefit and Privacy Panel are already underway.

We are also very pleased to be able to support the Scottish Government's Test and Protect initiative through additional specialised services within the National Safe Haven, and to be working with colleagues across the UK on the ISARIC4C project.

International Severe Acute Respiratory and Emerging Infection Consortium

ISARIC, the International Severe Acute Respiratory and Emerging Infection Consortium, is a global federation of clinical research networks, providing a proficient, coordinated, and agile research response to outbreak-prone infectious diseases. ISARIC4C is the consortium's response to COVID-19, and EPCC is working with clinical and genomic experts from Oxford, Liverpool, Manchester and Edinburgh universities, including Edinburgh's Roslin Institute, to provide a data and computing environment to help search for possible genetic markers in COVID-19 patients.

COVID-19 has taken a terrible toll in the UK, both directly and indirectly, but a ray of light has been shone by the world-leading research happening at UK universities and life-science firms. Here at EPCC we're glad that we've been able to contribute, even in a small way, to these efforts.

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

The Edinburgh International Data Facility (EIDF) is the underpinning data infrastructure of the activities of the Edinburgh and South-East Scotland Data Driven Innovation (DDI) Programme. Built by EPCC, EIDF provides highly-secure Safe Haven services to health and government users, supporting the linkage of complex personal data for public benefit research and policy-making.

Supported by



THE UNIVERSITY
OF EDINBURGH



Data-Driven
Innovation

Read more about the work of the Edinburgh International Data Facility at:
<http://bit.ly/2POsTQs>

Data Drive Innovation
programme: <https://ddi.ac.uk>



First phase of ARCHER2 arrives in Edinburgh

ARCHER2, the new UK national supercomputing service, is a world-class advanced computing resource for UK researchers. The service is due to commence later in 2020, replacing the current ARCHER service.

The four-cabinet Shasta Mountain system completed its journey from Cray's Chippewa Falls factory in the US to EPCC's Advanced Computing Facility in July. This is the first phase of the 23-cabinet system of ARCHER2, the UK's next national supercomputing service.

Moving these specialist systems and getting the right people here to install them is a logistical challenge at the best of times, but with the COVID-19 restrictions this was considerably more challenging than usual.

We are grateful to our colleagues at Cray/HPE for all their planning and perseverance! It is a huge step forward to see these systems on site.

Related to this, the ARCHER2 Test and Development System (TDS) is operational in Wisconsin, with the ARCHER2 team accessing this remotely. This has allowed testing

and preparation work to get underway with user documentation, training courses and user application code support.

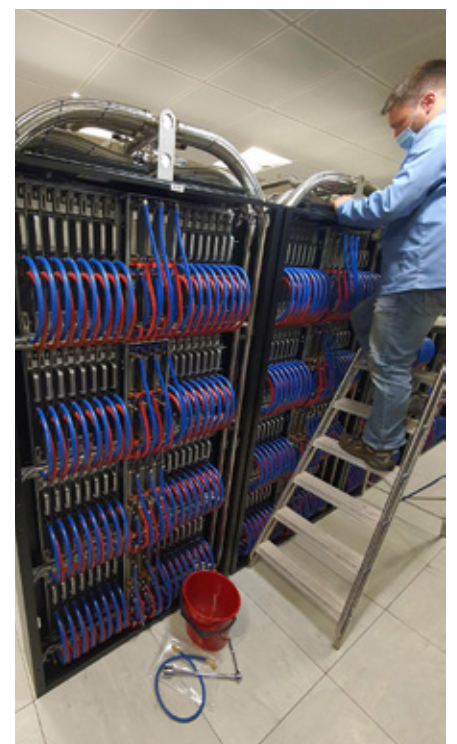
The system

ARCHER2 will be a Cray Shasta system with an estimated peak performance of 28 PFLOP/s. The machine will have 5,848 compute nodes, each with dual AMD EPYC Zen2 (Rome) 64 core CPUs at 2.2GHz, giving 748,544 cores in total and 1.57 PBytes of total system memory.

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

ARCHER2 is provided by UKRI, EPCC, Cray (an HPE company) and the University of Edinburgh. It is hosted and managed by EPCC at the Advanced Computing Facility.

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





Images by Greg Muir.

Gregor Muir, a summer intern at EPCC's Advanced Computing Facility, describes the week-long installation of the first phase of ARCHER2.

Months of planning and work went into preparing for this week, including running large power supplies, installing cooling pipework, strengthening the flooring for the 3500kg cabinet, and vacuuming the computer room.

Monday

We were greeted by four enormous trucks and the Cray team who had helped ship and pack the ARCHER2 system. By 2pm everything was unpacked and sitting in the right place, which was impressive, considering the size of the cabinets.

Tuesday

Power connections were quickly made, but we had to wait for the arrival of the de-ionised water used in the system's internal cooling before it could be switched on. However, the CDU (cooling distribution unit) was powered up.

The CDU can be imagined as two parts. One is attached to the site's water supply, incoming at 16 degrees. This chilled water travels through a heat exchanger, effectively two interlinked radiators with entirely separate water

supplies. The other half of the CDU pumps the system's internal cooled water around, through the heat exchanger, maintaining a specified temperature without connecting its own water circuit to the site. This is for robustness and to separate it from the more expensive 'water' needed to cool the supercomputer.

While powering up the CDU, the site's water was connected and proved to all be working perfectly.

Wednesday

The CDU work continued, with the de-ionised, additive-enhanced water arriving. There was also lots of configuring and some troubleshooting. By the end of the day, management and storage were sorted, and the CDU was operating on the site's water supply and circulating the system's supply.

Thursday

The last day on site for the installation teams, which meant all trouble-shooting had to be finished.

Friday

Load testing has begun. There will now be a period of training before the system is made available to users, and it will operate concurrently with ARCHER for a few months before the full 23 cabinet system is delivered and built.

**Gregor Muir
Summer intern,
Advanced Computing Facility**

I was lucky enough to watch the diagnosing and repairing of blades. The first was done in less than two minutes, just a direct swap of two DIMMs with fresh spares. The workbench trolley that the blade was sitting on was hinged and, with the aid of a pneumatic arm, the blade was lifted from horizontal to vertical, ensuring ease of access for the blade lifter. The lifter itself was pre-programmed to the height of each row of blades, allowing a perfect matching up through the use of magnetic sensors. Everything was beautifully engineered.



For further information and photographs, see:
www.archer2.ac.uk/news



EPCC awarded ARCHER2 services contracts

We have been awarded contracts to run the Service Provision and Computational Science and Engineering services for ARCHER2, the next UK National Supercomputing Service.

The ARCHER2 Service is a world-class advanced computing resource for UK researchers. It represents a significant step forward in capability for the UK's science community, with the £79 million service among the fastest CPU-based systems in the world. It will be hosted by EPCC at our recently expanded Advanced Computing Facility (ACF), a state-of-the-art home for this vital new service.

The Service Provision contract (SP) provides essential system maintenance and upgrades, it is also the access point for users of the ARCHER2 system. The Computational Science and Engineering contract (CSE) supports users with their individual projects as well as providing ongoing help, software support and development. These are both integral parts of the full ARCHER2 service.

"EPCC is exceptionally proud to have been selected as the SP and CSE provider. For the past 30 years we have hosted novel computing services and during that time we have had the privilege of working closely with the incredible

community of people that use them. The UK community has a rich heritage of conducting breakthroughs thanks to the UK National Supercomputing Service. As we enter our fourth decade, I look forward to the many future breakthroughs that ARCHER2 will deliver. Ultimately it's the community of users who make a service successful and the team here at EPCC look forward to supporting all of them on the new ARCHER2 system."

Prof. Mark Parsons, EPCC Director

ARCHER2 eCSE programme

The Embedded CSE (eCSE) programme provides funding to the ARCHER2 user community to develop software for the system. Following on from the very successful ARCHER eCSE programme, we are pleased to announce that a similar programme will run as part of the ARCHER2 service. For information on the programme and associated calls, see www.archer2.ac.uk/ecse

ARCHER2 is provided by UKRI, EPCC, Cray (an HPE company) and the University of Edinburgh. The service is due to commence operation in 2020, replacing the current ARCHER service.

"The award of the contracts to provide these crucial services is a major milestone in commissioning the ARCHER2 service, and providing a step-change in capability for the UK's digital research community. The people and skills that make advanced computational science possible are one of the most important parts of our computational science infrastructure. UKRI is delighted to continue its long-standing relationship with the team in Edinburgh, one of the UK's leading centres of excellence in high performance computing."

Dr James Hetherington, UKRI Director of Digital Research Infrastructure



www.archer2.ac.uk



CIRRUS Tier-2 service upgrade

We are pleased to announce a major upgrade and extension to the life of the successful Cirrus EPSRC Tier-2 HPC service

EPCC has received £3.5m funding over four years from the Engineering and Physical Sciences Research Council (EPSRC) to continue the Cirrus service until early 2024. We will add 144 NVIDIA V100 GPUs to the system and a 256TB high performance storage layer for the most demanding data streaming applications.

Cirrus has already proven to be a valuable Tier-2 service. The Phase II funding will allow us to continue and grow the service we provide. I'm particularly pleased about the addition of the GPU blades – this will provide a key resource for users to explore programming frameworks and accelerated versions of their most commonly used applications.

The new capability has two purposes:

Preparing for heterogeneity at the Exascale.

Two routes to the Exascale exist – a cores-only route or a cores plus accelerator approach. Worldwide,

many supercomputers, including the world's fastest, now use GPUs. The UK has not, to date, chosen this route but a large number of application codes now run well on heterogeneous accelerated platforms. The new GPU resource on Cirrus allows users to experiment with this potential future architecture.

Supporting the growth in AI and ML.

EPCC's limited investment in GPUs on Cirrus to date has clearly demonstrated demand for GPUs for AI and ML applications. We therefore expect a strong portfolio of projects seeking to use Cirrus Phase II for such applications.

Cirrus will be upgraded over the next couple of months. Some downtime will be inevitable, but we will do our best to minimise this. We will keep all users fully apprised of our plans and give early notice of all downtime.

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

Housed at EPCC's Advanced Computing Facility, Cirrus is a state-of-the-art advanced computing facility available to industry and academia. It provides an ideal platform for solving computational, simulation, modelling, and data science challenges.



www.cirrus.ac.uk

EPCC joins the UK's ExCALIBUR programme to address the challenges of Exascale

ExCALIBUR is a £45.7m programme to address the challenges and opportunities offered by computing at the exascale (high performance computing at 10^{18} floating point operations per second). The programme will address problems of strategic importance, and how to approach them in an efficient, effective, and productive fashion on the world's largest computers.

The Strategic Priorities Fund is one of the UK Government's largest programmes to work on multi-disciplinary and inter-disciplinary research and innovation, and has the ultimate aim of boosting the productivity and competitiveness of the economy.

In the context of ExCALIBUR, this includes work on currently intractable problems of strategic importance such as drug and vaccine research, climate and weather prediction, and fusion power and green energy sources.

Specifically, the aim is the redesign and reimplementation of high priority codes and algorithms to maintain a leading position for the UK in high performance computing.

With the first exascale machines perhaps becoming available between 2021–2023, the challenges are certainly formidable. A range of development is required from systems software and libraries, applications themselves, and utilities to prepare, analyse and visualise large amounts of data.

A very high degree of parallelism will be relevant throughout – perhaps $O(10^9)$ – $O(10^{10})$ simultaneous threads of execution will be employed at the exascale.

Applications must maximise efficient use of computational resources, minimise costly data transfers and synchronisations, and at the same time deal with significant algorithmic complexity.

All this must be done with an eye not just to flexible and efficient computation and robust numerical results, but to productivity of those doing the development and to energy efficiency at the time of execution.

In October 2019, as the first stage in ExCALIBUR, UKRI put out a call for high priority use cases around which to form design and development working groups. EPCC is now active, in collaboration with universities and industrial partners around the UK, in three of the resultant eight groups which have been funded: see opposite.

The three groups will concentrate on different areas: mesh generation and manipulation for finite element problems, materials simulations, and lattice field theory.

Future issues of EPCC News will return to cover each of these strands in more detail. It will be the results of such programmes which will enable the benefits of efficient exascale computing to be achieved.

Kevin Stratford, EPCC
k.stratford@epcc.ed.ac.uk

ExCALIBUR is led by the Met Office and the Engineering and Physical Sciences Research Council. It is a UK Research and Innovation (UKRI) Strategic Priorities Fund programme.



EPCC's ExCALIBUR collaborations

The Materials and Molecular Modelling Exascale Design and Development Working Group

This project brings together Research Software Engineers and experts from mathematics and computer science with domain experts in Materials and Molecular Modelling (MMM) to ensure this key research area is ready to exploit future exascale computing resources. This will cover all components required by the community as pre-exascale and exascale computational resources become available to researchers. The project will equip the UK MMM community with the tools and communities required to use exascale computational resources efficiently to address many EPSRC Grand Challenges.

The project is led by UCL with EPCC among the co-leaders.

ELEMENT

ELEMENT addresses the high priority use case of meshing for the Exascale (ensuring that meshes are of sufficient quality to represent Exascale problems and can be partitioned efficiently to minimise load imbalance) as well as meshing at the Exascale (creating highly

scalable solutions able to exploit extreme levels of parallelism).

Meshing and geometry management remain a significant bottleneck for complex applications on HPC platforms, posing a challenging obstacle that must be overcome to enable Exascale simulations. From a technical perspective, these issues include improved geometric handling, mesh adaptation and optimisation, intelligent meshing, automation and robustness, all within a large distributed environment that lies outside of our current capabilities.

The project is led by EPCC.

EXALAT

EXALAT (Lattice Field Theory at the Exascale Frontier) is a collaboration led by the University of Edinburgh, with total funding of £365k to build on the recognised expertise in the Lattice Field Theory (LFT) community to use High-end Computing, to inform and upskill the wider scientific ecosystem for Exascale computing.

The project will develop a roadmap for LFT, as an example for other domains, to prepare for science on the first Exascale computers. It will also develop initial resources (best practice, training, and algorithmic/software templates) to begin the journey.

ELEMENT project
<https://epcced.github.io/ELEMENT>

EXALAT project
<https://bit.ly/2xKoCrE>

The Edinburgh International Data Facility

When it comes into service in 2021, the Edinburgh International Data Facility (EIDF) will be a unique place to store, find and work with data.

Part laboratory and part repository, EIDF is the underpinning data and computing infrastructure of the Data Driven Innovation (DDI) programme.

The DDI programme is one of six within the Edinburgh & South-East Scotland City Region Deal. It has ambitious targets to support talent, research, commercial adoption, and entrepreneurship across the region through better use of data. DDI targets 10 industry sectors, with interactions managed through five DDI Hubs. The activities of these Hubs are underpinned by EIDF.

On the repository side, EIDF provides long-term hosting and curation of datasets for a wide range of stakeholders. On the laboratory side it offers cloud-like and high-performance computing environments for researchers and innovators to work with data.

EIDF will grow and mature with the DDI programme, expanding in capacity and capability, responding to the needs of the innovation Hubs and, through them, to learners, researchers, innovators and entrepreneurs from across the region and beyond.

What will it look like?

Most EIDF users will work in the **Data Service Cloud**, which will offer a rich set of data science and analytics tools. We aim to create ready-to-use environments with

pre-installed, pre-configured toolsets backed by CPU, GPU and storage resources.

The Data Service Cloud will sit on top of an **Analytics-Ready Data Layer (ARD Layer)**, where EIDF data can be shared and re-used. The ARD Layer will grow as we collect more data. Users will search the **Data Catalogue** to find what analytics-ready data EIDF has, and how to access it.

EIDF data managers will work with data depositors at the **Data Ingest Gateway**, ensuring that incoming data are safely stored in the **Data Lake Archive Layer**, and well-described in the Data Catalogue. Data in the Data Lake will be stored for the long term, following best practices in digital preservation.

EIDF data wranglers will work in the **Data Preparation Layer**, often in collaboration with data depositors and others, to turn archived data from the Data Lake into analytics-ready data products in the ARD Layer, so completing what we hope will be a virtuous innovation circle.

Safe Haven services

EIDF will offer **Safe Haven services** to health and government users, following best practice in independent governance and supporting the linkage of complex personal data for public benefit research and policy-making under

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

The Edinburgh International Data Facility, which is being developed by EPCC, will facilitate new products, services, and scientific studies by bringing together regional, national and international datasets.



national and regional safeguards.

Building on EPCC's expertise in operating the National Safe Haven for NHS Scotland, we will offer Safe Haven services for organisations to host and govern access to their data assets in a highly secure environment. Safe Havens will be isolated from the rest of EIDF, and regulated by independent information governance bodies.

How will EIDF evolve?

Apart from getting bigger, the most noticeable change will be in the richness and variety of the datasets available. We aim to collect and curate a large number of datasets and make them "analytics ready". Some datasets might be small, but we hope that many will demand the petabyte scales of the underlying hardware. We'll be working on making them as useful as possible – easily findable, accessible, linkable and interoperable.

Where are we now?

Hardware. A £100m procurement deal has been finalised with Hewlett-Packard Enterprise (HPE). Using a building block approach to meet different needs, we now have the flexibility to build out EIDF incrementally. This will enable us to respond to the rapidly changing nature of data driven innovation. We are working closely with HPE on the first major "Phase 1" equipment purchase, with a number of early

adopters firmly in mind.

Software. EPCC software and systems engineers have been putting in place the last few pieces of a rich cloud-computing environment for data science.

EIDF's first service offerings, expected in April next year, will include virtual desktops pre-configured with data science tools and with access to GPU acceleration and a rich and expanding layer of analytics-ready data. We're also planning to roll out data hosting and archiving services, a new version of the University's Internet of Things research service, and a first version of the central EIDF data catalogue. Other services will follow through the year.

National Safe Haven. A lot of activity has centred recently on the National Safe Haven, the existing secure research environment EPCC run for the NHS, Public Health Scotland and the Scottish Government. In this context, we've been very pleased to be able to assist in the fight against COVID-19 (see page 3).

The first quarter of 2021 will be crunch time for EIDF. While predictions in the current climate inevitably involve guesswork and a degree of luck, we're confident we have all the pieces in place for a rich and powerful new service for data science in Scotland.

Despite the challenges of COVID-19, we have continued our work with EIDF's early adopters, including:

- UK Research and Innovation
- The Scottish Government
- Fife Council
- The National Collection of Aerial Photography
- The Industrial Centre for AI Research in Digital Diagnostics
- The Global Open Finance Centre of Excellence.



EIDF website:
<http://bit.ly/2POsTQs>

The EIDF roadmap is available at:
<https://bit.ly/2WK9NOU>

Data Drive Innovation
programme: <https://ddi.ac.uk>

The Global Open Finance Centre of Excellence: delivering innovation and public benefit

In June, a unique collaboration between the University of Edinburgh, the Financial Data and Technology Association, and Fintech Scotland was awarded £22.5m from UK Government to harness financial data for societal and economic benefit. The Global Open Finance Centre of Excellence (GOFCoE) will be an independent collaboration between governments and regulators, the financial services, industry and academia.

To support this ground-breaking work, EPCC is developing the Financial Data Safe Haven, the underlying data and computing facility that will host a rich variety of financial data, including the transactional data of individuals and organisations. The facility will allow subsets of data to be linked and made available within a strict framework to a variety of actors, including academic researchers, policy makers, regulators, financial institutions, and companies that use technology to improve the delivery of financial services (fintechs).

Data use will be managed by transparent ethical- and privacy-impact assessments, and access will only be granted to trained and vetted individuals from approved organisations for specified purposes. Projects must align with GOFCoE's ethical principles, which aim to safeguard individual privacy and to promote fairness, sustainability, and evidence-based policy development.

EPCC is working closely with the GOFCoE team as an early adopter of the Edinburgh International Data

Facility (EIDF). EPCC has provisioned suitable infrastructure in its Advanced Computing Facility (ACF) for GOFCoE's initial projects and is working with GOFCoE on the specification and design of the EIDF Safe Haven.

Access to insights from granular, linked financial data could help public bodies understand the full impacts of decisions and the true operation of the economy. The Financial Data Safe Haven will be a trusted data resource, allowing analyses in a secure, transparent and privacy-preserving environment, subject to ethical and privacy review and information security controls.

Open Banking for social good

GOFCoE will boost developments in Open Banking and in the broader area of Open Finance that will be grounded in the rights of customers to access, manage and control their data and enable trusted third parties to use that data on their behalf through informed consent.

Kevin Collins
Assistant Principal for Industry Engagement,
University of Edinburgh

"GOFCoE is taking the lead in innovation driven by financial data. We are proud to support its endeavours, particularly its focus on common good, which will be vital in the era of COVID-19 and beyond."

Mark Parsons
EPCC Director

"EPCC was the natural choice to host the GOFCoE data infrastructure, given its long history of processing and managing data. EPCC has acquired unique skills by operating Public Health Scotland's Safe Haven. Its ISO 27001 and UK Digital Economy Act accreditations combined with GOFCoE's Information Governance will reassure the public, regulators, Government and data providers in the GOFCoE ecosystem that data controlled by GOFCoE is safe"

Damien McGarrigle
Programme Lead, GOFCoE



Image: metamorworks/Getty

Financial data has major and largely untapped potential for transformative societal good by enabling better understanding of the economy and more effective policies and detailed assessment of outcomes in areas including financial exclusion and poverty.

The GOFCoE team believe that a balanced approach is needed that respects both the right of individuals to control their data and society's need to gain detailed insights. For example, measuring and assessing policy initiatives to enable informed, evidence-based debate.

The team believe this balance can be achieved through GOFCoE, as a trusted, neutral data facility with a strong and transparent ethical approach that includes citizen consultation and public consent.

Trust and ethics

There have long been ethical and trust questions in the financial sector, but the level of concern is at an all-time high. GOFCoE hopes that the Financial Data Safe Haven, with strong security, privacy and

ethical safeguards, will foster the ethical use of data for public and societal benefit.

The centre also aims to establish a specialist practice in the area of trust and ethics to advise industry on best practice, offer staff training, and create auditing facilities, allowing models and data to be securely tested for bias and potential privacy violations.

Education partnerships

Education and training will be a major focus, developing a practice in data ethics and trust, and in financial data in particular. GOFCoE plans to offer studentships, continuing professional development training and certification, as well as collaborative research. Partnerships with, for example, the University of Edinburgh and as other Scottish and UK universities will be a crucial part of GOFCoE's mission to educate.

Supported by



GOFCoE's key objectives

- Catalyse financial data collaboration to enable research and innovation, so speeding industry adoption of Open Finance at scale
- Decrease time-to-market and development costs
- Train new talent to enable sectoral expansion
- Undertake data-driven research
- Work with regulators to establish ethical standards and best practice
- Build partnerships to achieve financial sustainability

GOFCoE is hosted by the Edinburgh Futures Institute, supported by EPCC, and led by the University of Edinburgh, the Financial Data and Technology Association and FinTech Scotland.

Sharing our expertise and facilities with industry



Image: tdub303/Getty

EPCC supports businesses across a wide range of industries and sectors.

We recently completed a successful collaboration with Oil and Gas integrity management specialists, **PIM Ltd**, to investigate the benefits and opportunities for using novel automation and data analytics technologies to improve the effectiveness and productivity of asset integrity management services. Over the last eight years PIM has delivered risk-based inspection services to the Oil and Gas sector and has built up a considerable amount of data which can be leveraged to optimise future inspection services.

Also in the Oil and Gas sector, we have been working on a long-term project with innovative SME **Exnics** on the management and processing of extremely large and high-velocity datasets to provide a platform for subsequent machine learning modelling.

EPCC is working with **Bo-Create**, a local tech start-up in the property and construction sector. EPCC is providing expertise to boost the development and delivery of a suite of new products, including decision support tools to improve the speed of housing delivery, data analysis tools to enhance information flow across the supply chain, and impact

measurement tools offering pre-investment assessment of the social, economic and environmental impacts of housing and mixed-use developments.

EPCC's world-class HPC infrastructure continues to be extremely well-used by industry, with ongoing collaborations across the energy, manufacturing and environmental sectors, plus a variety of new local and international users such as **Ingrid Cloud** (Wind Simulation), **Risktec** (CFD) and **Jacobs** (Nuclear energy).

We continue to provide software-as-a-service platforms to companies including **Engys** and **Gexcon**, offering access to proprietary advanced CFD and fire simulations running on EPCC's HPC infrastructure.

We are also supporting several commercial projects as part of our role in the Bayes Centre and the wider Edinburgh City Region Deal, including organisations such as the **National Centre for Aerial Photography** (one of the world's largest collections of aerial imagery), **Trace Data** (IT solutions and services), and **GOFECO** (developers of high-quality engine lubricants).

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

“The oil and gas industry has unquantifiable amounts of historical data. Working with EPCC, one of Europe's foremost HPC centres, we plan to harness this legacy data and consider how data captured in the future can be used to best effect. We will explore opportunities for greater service standardisation and automation.”

Tyler Stewart
Project lead, PIM Ltd



Contact Thomas Blyth, our Business Development Manager, to discuss how we can work with you: t.blyth@epcc.ed.ac.uk



Service Portal for engineering excellence launched

EPCC is a partner in the EXCELLERAT programme, which brings together key players from industry, research and HPC to provide services that will enable the European engineering industry to advance towards Exascale technologies.

There are great opportunities for engineering applications in exploiting Exascale computing. EXCELLERAT brings together expertise in HPC and engineering software to support the development of applications that can exploit computing capability at this level.

The EXCELLERAT Service Platform includes training, access to codes and software, and expert consultation. By presenting a wide variety of resources in an easily-accessible format, it will play an essential role in EXCELLERAT's mission to create a Centre of Excellence in Engineering. The Portal offers participants the most suitable type of support, based on their own current needs.

We have adopted a user-centric approach in designing the Portal, and it therefore offers separate dashboards tailored towards engineers, software developers, and the general engineering community. Each dashboard quickly directs users to relevant content, and delivers tailored information accordingly, providing a quick route to any related content and offers an easy way to request services from EXCELLERAT.

Engineers' dashboard

The Engineers' dashboard is designed for EXCELLERAT's end-users. It contains technical content and presents the codes, tools, and use cases under development by project partners. There is also a repository of the datasets that have been produced and publicly released by EXCELLERAT.

Developers' dashboard

The Developers' dashboard presents codes and use cases, with accompanying content tailored to the developers of software codes. And importantly, it presents a detailed list of the consulting topics for which EXCELLERAT can provide support.

Community dashboard

The Community dashboard brings together content of interest to the entire engineering community. It includes the EXCELLERAT training portfolio, a repository of all material published by EXCELLERAT, advice on accessing events organised by project partners, a careers portal, and funding opportunities for the Engineering community.

Claudio Arlandini, Cineca
c.arlandini@cineca.it

The EXCELLERAT Service Portal is a single point of access for expertise on how data management, data analytics, visualisation, simulation-driven design and co-design with high-performance computing can benefit engineering.



EXCELLERAT Service Portal
<https://services.excellerat.eu>
We invite you to send your comments about the portal to support@excellerat.eu

Advanced systems engineering for cybersecurity

EPCC has collaborated with Intel® and Illuminate Technologies to investigate innovative performance-enhancing, packet-processing approaches within the type of Cloud environment expected to be seen in 5G deployments.

Illuminate Technologies focuses on telecommunications networks, providing innovative capabilities for cyber-threat detection, legal compliance for critical communications infrastructure, and public safety.

The company collaborated with EPCC, and Intel® to investigate the suitability of new computer networking technologies for a new cybersecurity solution.

The challenge

Telecoms networks are undergoing a revolution in which network elements are fully virtualised and hosted in a data centre environment. The highest profile example of this is the 5G Service Based Architecture (SBA) that will require Network Function Virtualisation (NFV) to deliver advanced services such as Internet of Things (IoT) connectivity. This NFV architecture means that the traditional approach of monitoring links on network elements is no longer valid as the network elements and links are virtualised and the mapping to physical elements may change dynamically.

Our collaboration's first aim was to create an extremely reliable Cloud environment using standard components. Next, we wanted to establish a benchmark for data-traffic processing against which

further code optimisations could be measured. Finally, EPCC would design and develop performance optimisations utilising emerging programmable networking approaches, optimised for a Cloud environment.

It quickly became apparent that the project was best suited to answering questions about the suitability, reliability and ease of use of the novel technologies involved, notably OpenStack, P4C-XDP, eBPF, and DPDK for our use-cases. This area is important for Illuminate Technologies as it allows network-processing code to be targeted at multiple platforms from a common specification. The work was also expected to yield cybersecurity findings of interest to Intel®.

Results

EPCC developed and tested a new P4 code that aggregates Illuminate Technologies' algorithms handling network traffic. EPCC optimised the code, accelerating it by over 8% in some cases. This acceleration is already significant in the context of cybersecurity and lays a foundation for further improvements to the Illuminate Technologies P4 code.

EPCC also developed and documented a hardware and software toolkit that demonstrates the capabilities of eBPF in the Illuminate Technologies use-case.

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

Jeremy Fuller and Emma Cadzow
Illuminate Technologies
contact_uk@illuminate.solutions

"Working with industry leading partners in a cutting-edge testbed environment allows us to fast-track new concepts into prototype. The excellent facilities and technical expertise provided by this project were perfectly suited for this."

Roy Macnaughton
VP Engineering,
Illuminate Technologies,

"The Illuminate Technologies CodeOptScotland project clearly demonstrates how Scotland couples local industrial and academic excellence with its international links to exploit cutting-edge technology."

David Smith
Director National
Opportunities,
Scottish Enterprise

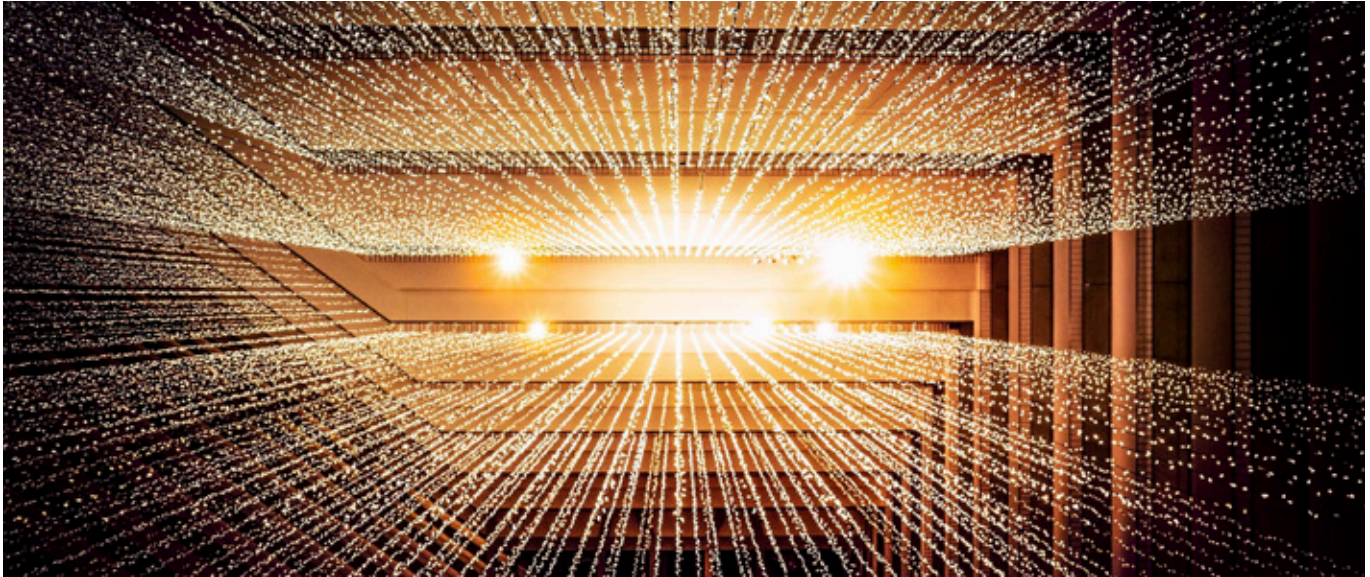


Image: Joshua Sortino on Unsplash.

This includes a cutting-edge unit-test framework that allows the functionality of Illuminate Technologies' code to be validated, and also benchmarked under certain conditions.

Benefits

Society's migration to Cloud technology means the need to identify and mitigate threats to provide a safe operating environment for economies, critical infrastructure and individuals is vital. The collaboration has highlighted many aspects around the maturity of intelligent packet processing in Open Source Cloud technology, highlighting areas of focus for future development.

Telecoms networks infrastructure is undergoing a rapid migration from central-office technologies, based on proprietary hardware, to data centre technologies with commodity, high volume hardware. With these new technologies it is no longer possible to handle higher traffic volumes through hardware optimisation. Instead extensive system engineering is required to optimise all the software components in the packet flow.

This project has demonstrated that software-based packet processing, using data centre software, can be applied to the essential network visibility workload and can scale to

the traffic volumes expected in next-generation networks such as 5G.

Collaborative work is central to the activities of Illuminate Technologies, EPCC and Intel®. The technologies used in the project were innovative and at the outset neither Illuminate Technologies nor EPCC had a mature toolkit to support this development work. We have now consolidated the knowledge and tools that will underpin important future developments in networking and code optimisation for cyber security functions.

This knowledge can support Illuminate Technologies' decisions when selecting target platforms and technologies.

By bringing together the expertise of the three partners, we were able to explore new innovations in packet processing. While it will take time for this technology to mature, the expertise gained by the partners places them in a strong position to bring the next iteration of this technology to market.

This project further enhances Scotland's growing reputation at the forefront of Cloud technology innovation, with growing companies like Illuminate Technologies in Edinburgh becoming a leading R&D centre of excellence in Scotland.

 illuminate technologies
accelerating informed decisions



"The adoption of technologies such as DPDK, XDP/eBPF and P4 helps evolve the capabilities of network infrastructure in cloud environments. Companies like Illuminate Technologies can use these capabilities to advance their development efforts in cloud networking."

Dr Nash Palaniswamy
General Manager AI & HPC
Solutions,
Intel®

CodeOpt Scotland

This project was part-funded by the CodeOpt Scotland programme, which was designed to increase company competitiveness by optimising the performance of business-critical software. CodeOpt Scotland is a partnership of EPCC, Scottish Enterprise and Intel®.



Image: Piranka via Getty Images

EPCC awarded new data accreditations

We have become an accredited processor under the Digital Economy Act, following an audit process conducted by the Office for National Statistics. EPCC is one of only seven UK organisations to gain this accreditation.

The Digital Economy Act (DEA) came into force in 2017 and covers issues related to electronic communications infrastructure and services. Under the Act, an organisation may become accredited for the purposes of processing data, principally with the linkage or de-identification of data, or the storage and provision of secure access to the de-identified data. EPCC is accredited for storage and provision for the National Safe Haven that it operates on behalf of Public Health Scotland.

The accreditation came after an on-site audit carried out by the Office for National Statistics (ONS), the government appointed auditors.

The audit was the culmination of several months of effort to collect and present evidence showing how EPCC defines, implements and records processes for information security.

Overall EPCC was assessed as having 9 out of 15 security controls in a 'mature' status (the highest category) with the remainder being 'good' (the second level). Accreditation is only given when all controls are at least 'good'.

Achieving this accreditation complements EPCC's existing ISO270001 certification and is an important step as it develops the Edinburgh International Data Facility and associated services.

Mark Sawyer, EPCC
m.sawyer@epcc.ed.ac.uk

UK Statistics Authority
<https://edin.ac/2WMqdXI>
Edinburgh International Data Facility
<https://edin.ac/36s3HXb>

ARCHER2 training: meeting our users' needs



Photo by Carl Heyerdahl on Unsplash

Training is one of the functions of the ARCHER2 Computational Science and Engineering (CSE) service, which is run by EPCC. We are fully committed to providing a rich, diverse programme of training that is responsive to our users' evolving needs. We have therefore designed a programme that addresses the training requirements of users with different needs and level of experience.

Our plan aims to enable all users to make efficient use of ARCHER2 as soon as it is available. There will be different introductory courses for different user categories:

Package Use on ARCHER2: efficient use of pre-installed research software packages on ARCHER2.

Development on ARCHER2: the ARCHER2 application development environment, core parallel and scientific software libraries, available debugging and profiling tools.

Data Science on ARCHER2: the essentials of ARCHER2, the basic use of core data science packages, and data-handling best practice.

Further details about the range of courses we offer can be found in the ARCHER2 training pages ([link is external](#)).

Given the current COVID-19 restrictions, we will deliver the courses online for the foreseeable future. We will investigate returning to our original plan of delivering face-to-face courses across the UK

once the lockdown measures ease.

A weekly programme of virtual tutorials (VTs) started in April. The VTs consist of interactive live webinars where experts can share their knowledge on a range of intermediate and advanced topics. Previous VTs can be found in the ARCHER2 training materials repository.

We will use Blackboard Collaborate software for live broadcast courses, virtual tutorials, and consultancy sessions. Material from all online courses, including videos of the lectures, will be made freely available on the web after the live run. In addition to this, we plan to make appropriate ARCHER2 courses available via a self-service model, with the first appearing later this year.

Our programme also features improved accessibility and inclusivity of HPC training through a variety of formats and delivery routes, using best practice to make content accessible and developing policies that promote diversity and inclusion in the community.

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



Further information

To learn more about our training plans, watch this webinar: bit.ly/3hwxgen

Please keep an eye on the ARCHER2 website to find out upcoming training courses. We also have a Twitter account.

Follow us on:
[@ARCHER2_HPC](https://twitter.com/ARCHER2_HPC)

HPC-Europa3

visitor programme

The HPC-Europa3 programme has now been running for three years. Its main focus is its visitor programme, which funds short collaborative research visits in any discipline in which high performance computing can be used.

EPCC is one of nine centres offering HPC-Europa3 visits, along with partners in Finland, Germany, Greece, Ireland, Italy, the Netherlands, Spain, and Sweden.

The visits continue to be highly successful, with over 100 publications having resulted from the work done so far by visitors to all of the participating centres. There is also much evidence of many collaborative links continuing long after visits are complete.

In EPCC News 85 we reported on two visitors (Mats Simmermacher and Javier Gallego Sánchez) who returned to Edinburgh to work in their former host research groups. Here we report on another such success.

In 2019 Giuseppe Negro was in the final year of his PhD at the University of Bari, Italy, when he undertook his first HPC-Europa3 visit, to collaborate with Prof. Davide Marenduzzo of the Institute for Condensed Matter and Complex Systems in the Department of Physics & Astronomy, University of Edinburgh.

After submitting his thesis that December, Giuseppe carried out a second HPC-Europa3 visit to the same group in Jan-Feb 2020. Following on from these successful visits, Giuseppe secured a post-doctoral research contract in the group. He also had his PhD defence exam in March which, due to the COVID-19 pandemic, he did from Edinburgh via Skype.

Giuseppe tells us more...

“My research involves modelling and simulations of soft composite

materials focusing on biological fluids.

“During my first visit in April and May 2019, we worked on simulations of 3D droplets of active cholesteric liquid crystals embedded in an isotropic liquid, understanding the outcome of the interplay between chirality and activity in active fluids. This work resulted in a very important publication co-authored by myself, my PhD supervisor Prof. Giuseppe Gonnella (University of Bari), my HPC-Europa3 host Prof. Davide Marenduzzo, and Livio Carenza, a PhD colleague from Bari who also made two HPC-Europa3 visits to the same host group at around the same time.

“During my second visit at the start of 2020, we performed simulations on active nematic liquid crystals confined in shells. Imposing topological constraints on the orientational order allows global control of active materials. Through high-demanding simulations we were able to discover new hydrodynamics regimes, ranging from regular motility modes to chaotic motion. These studies suggest new strategies to control active materials, and use the self-sustained flows they produce in new bio-inspired materials.

“My visits led me to secure a contract within my host group, where I will continue to study the complex interplay between topological defects and their induced flows in confined active liquid crystals.”

1. www.pnas.org/content/116/44/22065.short

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

Over 600 applications have been approved for visits to collaborate with more than 350 different host researchers in the nine participating countries. There have been 88 HPC-Europa3 visits to EPCC, by researchers working in 19 different countries.

HPC-Europa3 is funded by the European Union's Horizon 2020 research and innovation programme under grant agreement no. 730897. The programme runs until October 2021.



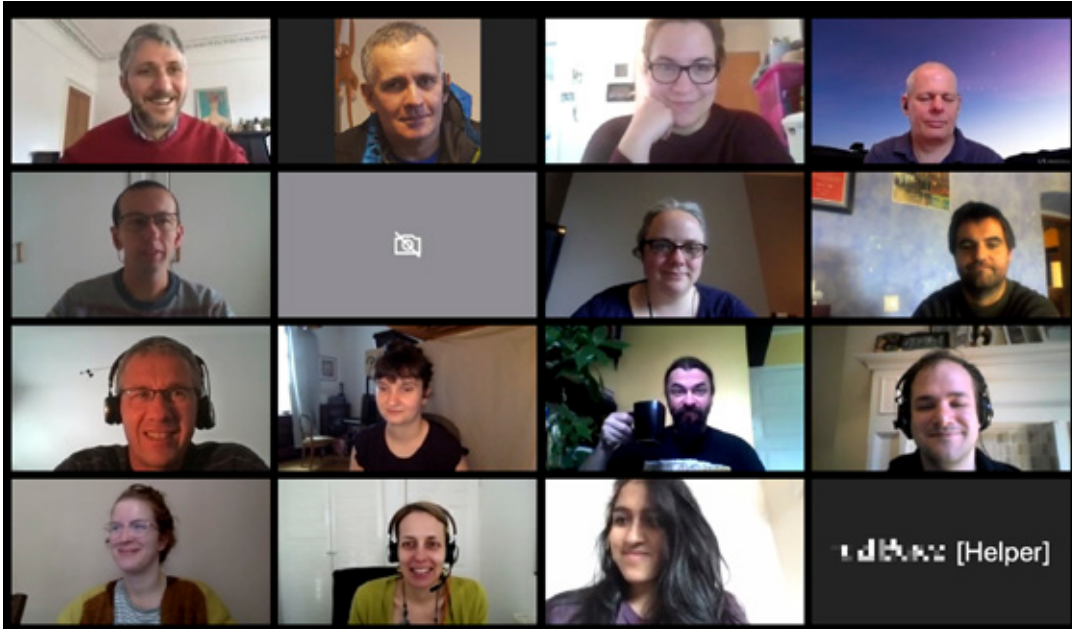
Next HPC-Europa3 closing date: September 17, 2020.

Further information

www.hpc-europa.org

www.twitter.com/HPCEuropa3

staff@hpc-europa.org



Edinburgh Carpentries go virtual

The Edinburgh Carpentries provide researchers with training in basic computing and data skills, even when the country goes into lockdown.

Carpentries workshops aim to improve research methods by demonstrating tooling and processes that will make researchers more productive. They also enable researchers to produce more sustainable and reproducible research outputs in alignment with the Software Sustainability Institute (SSI) motto: “Better software, better research”.

Since 2018 the Edinburgh Carpentries, supported by SSI and instructors from EPCC, has organised 25 free workshops across Edinburgh’s academic institutions, training over 300 staff and students in tools such as R, Python, Unix shell, git, and OpenRefine. Courses are very popular and quickly oversubscribed.

Planning for the 2020 workshops was well underway before the COVID-19 lockdown in March, with support from the Scottish Funding Council’s Upskilling Fund allowing us to expand the scope of the Data Carpentry workshops, opening them up to researchers working outside academia. The COVID-19 restrictions were implemented less than two weeks before the first

workshop of the year, forcing us to quickly adapt our delivery plans.

Several delivery platforms were investigated. The University of Edinburgh prefers Blackboard’s Collaborate but we found it created a number of problems for the agile delivery needed for the Carpentries’ style of workshop, especially the timelag when presenting. We opted to use Zoom, despite the reported privacy problems, partly because the Software Sustainability Institute had successfully run its Collaborations Workshop 2020 on this platform (see p24). So, taking on board a number of precautions, in April we delivered our first virtual Data Carpentry course to 15 participants.

On the whole, the course went very well. We are still receiving comments from the first set of attendees and will use it to improve future runs of the course, but are happy we can continue to deliver training while everybody is working from home. Regardless of the world we will emerge into after the COVID-19 pandemic, Carpentries workshops will never be the same again.

Mario Antonioletti, EPCC
m.antonioletti@epcc.ed.ac.uk

Participants’ screen real-estate is very important in an online workshop, and we highly recommend at least two screens: one for typing along with an instructor and the other to see the videostream of the virtual session.

Further information

More Carpentries workshops are planned, see:
<https://www.ed.ac.uk/bayes/about-us/our-work/education/workforce-development/courses/data-carpentry>.
<https://edcarp.github.io/>
<https://carpentries.org/>
<https://software.ac.uk/>
<https://openrefine.org/>

An update on our MSc programmes

We have seen increasing demand for data analytics skills for large-scale data, and in computation that can scale up. Our MSc programmes in HPC and HPC with Data Science address this demand.

The University's rich ecosystem of streaming and collaborative tools enabled us to switch to online teaching and supervision from the first day of the COVID-19 lockdown. Preparations are well underway for next year, and we will be ready to offer a mix of hybrid, online and blended teaching as required.

This fast transition is the result of our long-term investment in making the most popular courses available online. We started in 2015 by offering "Practical Introduction to Data Science" and "Practical Introduction to High Performance Computing" online. The online versions of the MSc programmes in HPC and HPC with Data Science are accepting applications for their first cohort to commence in September.

Our online MSc programmes are designed as an entirely separate offering to the on-campus versions, providing an alternative route for students to access the same material. Both programmes are designed to take 3-6 years, part-time.

The programmes are built around the same core set of material and compulsory courses to provide a firm grounding in the basics of HPC and Data Science, as well as ensuring our graduates are comfortable using good

programming and development practices.

Taught component

The core courses include absolute necessities in the IT field, such as Software Development and Programming Skills, as well as subjects specialising in high performance computing, like Message-passing Programming and Threaded Programming on the code development side, or Practical Introduction to HPC/HPC Architectures for the hardware angle.

We also offer courses in data analytics and data management, which are mandatory for the HPC with Data Science programmes, but accessible to all our students.

These core subjects form a half of the taught content. Students are then able to take a number of optional courses from a selection to reach the 120-credit requirement for the taught component – usually in Semester 2 for the on-campus version or from year 2 onwards for the online version to engage with further advanced material.

At EPCC, we strive to teach skills that are applicable to a wide range of real-world challenges, and we use this approach both in our taught courses as well as in the dissertation projects.

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

The on-campus MSc programmes in HPC and HPC with Data Science are growing rapidly. With 70 students, the class of 2020 is nearly double the size of 2019's class.

Thanks to the efforts of staff and with the University's support, we have been able to continue teaching activities despite the COVID-19 restrictions. We plan to carry on as normal, even though this might require redefining "normal".

www.epcc.ed.ac.uk/msc



Some of the class of 2019-2020 and EPCC teaching staff in the Bayes Building.

Dissertations

With the taught component successfully completed, students have the opportunity to utilise the skills and techniques learned on the programme in their dissertations.

Dissertation projects are often undertaken on real-world problems in collaboration with partners from other area of the University, wider Higher Education section, or externally.

Our students work on industry-motivated problems using state-of-the-art and/or emerging hardware solutions under the supervision of experienced staff.

As in the previous years, the 2020 dissertation projects include a number of collaborations ranging from data analytics solutions for the hospitality industry through high-performance computation for image analysis in astronomy, to the use of novel hardware, such as FPGAs and spinnaker boards.

With the launch of the online MSc programme and the ability to deliver even on-campus teaching online, we are as ready for 2020/21 as can be.

We will be looking for collaborations suitable for the dissertation projects, which will be offered to our students in early autumn. The dissertation projects typically

concern (but are not limited to) data exploration and analysis, enhancing performance via parallelisation, and emerging applications of AI and ML techniques to new domains.

Further information is available on our website or you can email our dissertations team!

MSc programmes content

Our online and campus programmes are built around the same core set of material and compulsory courses to provide a firm grounding in the basics of HPC and Data Science:

- Message-passing Programming (10 credits)
- Threaded Programming(10 credits)
- Programming Skills (10 credits)
- Software Development (10 credits)
- Project Preparation (10 credits)
- HPC Architectures (10 credits) (on-campus)* / Practical Introduction to HPC (20 credits) (online)
- Fundamentals of Data Management (10 credits) & Data Analytics with High Performance Computing (10 credits) (on-campus)** / Practical Introduction to Data Science (20 credits) (online)**

* Compulsory for HPC only

**Compulsory for HPC with Data Science only

EPCC offers two MSc degree programmes: The MSc in High Performance Computing and the MSc in High Performance Computing with Data Science. Both programmes offer flexibility with a rich choice of units and excellent job prospects on graduation.

EPCC MSc programmes' website:
www.epcc.ed.ac.uk/msc

To contact the EPCC MSc Dissertations team, please email:
industrial.dissertation@epcc.ed.ac.uk

How to move a collaborations workshop online fast!

The Collaborations Workshop (CW) is the annual flagship event run by the Software Sustainability Institute. For the last ten years researchers, developers, innovators, managers, funders, publishers, leaders, and educators have gathered to explore and share best practice.

The Workshop is a highly interactive event with about 100 participants, and designed to facilitate meeting as many new people as possible. With the global pandemic forcing us to stay at home, could we successfully recreate the experience online, and with only three weeks to prepare?

The COVID-19 Crisis

On March 5th, just over three weeks before Collaborations Workshop 2020 (CW20) was due to start in Belfast, everything was in place. The programme had been finalised, keynote speakers announced, catering and accommodation booked.

At this point the number of reported cases in the UK was small, but things were moving rapidly. Given the amount of work that had already been put into organising the event, the decision was made on March 10th to move the event online but stick with the same dates: April 29th–May 1st.

From physical to virtual

For in-person CWs, the days are long to maximise the opportunities for collaboration and discussion. Our agenda was packed with talks, group sessions, social activities and

mini-workshops and we wanted to incorporate as much of the original programme as possible. We did not organise online social activities during the breaks because the days were still quite long and we thought people would need screenbreaks, but this was ultimately something that participants missed the most.

We chose the Zoom video conferencing platform because we had the most experience with it and it had the functionality we needed (namely breakout rooms and easy recording of the event).

Although Zoom has chat functionality, it can be distracting for hosts/speakers and difficult to manage when there are many participants using it. The CW20 Slack workspace enabled participants to create channels for their discussion groups and Hack Day teams, and also allowed for a more informal engagement for sharing photos. Shared Google Sheets and Google Docs were used to link and share resources, agendas and take collaborative notes.

Lessons Learned

- Keep people informed: we increased communication with

Rachael Ainsworth, University of Manchester, Rachael.Ainsworth@manchester.ac.uk

Shoaib Sufi, University of Manchester, Shoaib.Sufi@manchester.ac.uk

Neil Chue Hong, EPCC, N.ChueHong@epcc.ed.ac.uk

Over the course of three days, we had keynotes on Open Research and the Replication Crisis, Data Privacy, GDPR and DataShield.

Participants gave a series of ten lightning talks, sharing projects and initiatives including the FAIRsFAIR project, the eLife Innovation Sprint 2020, and Sarah Gibson's Cross Stitch Carpentry lesson! Almost 40 discussion topics were proposed, ranging from open research, data privacy and software sustainability to various topics spurred by COVID-19. Sixteen mini-workshops were also delivered, teaching people new skills.



Photo by Trà My on Unsplash

participants, to ensure that they always knew what was happening, and were part of the decision-making process. This also meant that they were much more appreciative of the challenges.

- Negotiate with suppliers. We were able to postpone most bookings and renegotiate what sponsors would receive.
- Online is more tiring than in-person: participants are in one place, in one position. Having longer breaks and shorter days helps reduce screen fatigue.
- Test your infrastructure: we evaluated options early on and ran tests to ensure we were confident we knew how to use it to our advantage.
- As organisers it's important to walk through the agenda, modify it and talk through any implications for attendees; so that clear instructions, communication needs and the flow of the events is planned for.
- Clear guidance and practice session for facilitators. We found that stepping through the process and controls with facilitators before the event helped everything run smoothly.

- Breakout Rooms are a great way of encouraging interaction. Breakout rooms at the start of each day with small groups of random participants allowed the personal conversations to flourish.

- Keep it informal and social.
- Review your Code of Conduct: many CoC's are designed for in-person events.
- Have a dedicated helpline for participants
- Screen space is extremely valuable. A second monitor and/or large screen is necessary for keeping Zoom, Slack and notes in view.
- Reach out to experienced members of your community for help if you need it.
- Be honest and authentic throughout the event, it will empower your participants to do the same.

Conclusions

Running an online event is just as challenging as running an in-person event, perhaps harder. Yet many of the rules of running a successful event apply equally well - just give yourself more than three weeks to prepare!

97% of participants who provided feedback said that the workshop was useful and enjoyable. One participant wrote: "I was much more personally engaged and active in CW20 than any of the other online events I attended. It makes me feel like I contributed rather than just passively consuming information."

The Software Sustainability Institute cultivates better, more sustainable, research software to enable world-class research.

Find out more at:
www.software.ac.uk



Image: Paul Dodds

Study HPC in the heart of the city

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 ARCHER (the UK's National HPC Service), and Cirrus (an EPSRC Tier-2 National HPC facility).

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.

"Studying the MSc in HPC at EPCC has given me the benefit of a thorough practical grounding in supercomputing and the once-in-a-lifetime opportunity to participate in the Student Cluster Competition at ISC High Performance 2018, alongside the wider opportunities afforded by the student experience at the University and in the city of Edinburgh."

Wilson Lisan, 2018 MSc in HPC graduate

Optional course choices include modules from the School of Informatics. 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 20 universities in the world by QS World University Rankings 2019.

EPCC is using its expertise in both online and face-to-face training to deliver the full MSc in 2020/21 while adhering to Scottish Government COVID-19 guidance.

"Modules covered the full range of HPC and Data Science skill sets from core 'best-practice' ways of working to the latest technologies. These were well-structured and delivered at a good pace by lecturers who were more than happy to engage in discussion in response to questions."

Dr Andy Law, Roslin Institute,
2017 MSc in HPC with Data Science graduate