



ARCHER CSE Service Quarterly Report

Quarter 4 2019



1. Document Information and Version History

Version:	1.0
Status	Release
Author(s):	Lorna Smith, David Henty, Chris Johnson, Xu Guo, George Beckett
Reviewer(s)	Alan Simpson

Version	Date	Comments, Changes, Status	Authors, contributors, reviewers
0.1	07-12-2019	Skeleton document	Lorna Smith
0.2	08-01-2020	Added training section	David Henty
0.3	10-01-2020	Added eCSE section	Chris Johnson
0.4	10-01-2020	Added CSI section and BCDR section	Xu Guo & George Beckett
0.5	14-01-2020	Added Outreach, Central sections	Lorna Smith
0.6	14-01-2020	Draft completed	Lorna Smith
0.7	15-01-2020	Reviewed	Alan Simpson
1.0	15-01-2020	Version for EPSRC	Alan Simpson & Lorna Smith

2. Executive Summary

This report covers the period: 1 October 2019 to 31 December 2019 inclusive.

- Centralised CSE Team:
 - We have focused on supporting users with moving their data from ARCHER to other resources. A virtual tutorial has been carried out and the best practice guide updated. Further virtual tutorials are planned for January.
 - Containers for HPC are currently of interest to the HPC community and may be of benefit to ARCHER 2. We have therefore been investigating their use, with comparative results generated on the Tier-2 system, Cirrus.
 - The system continues to be heavily utilised and the CSE team have responded to a range of in-depth queries relating to compilation and execution issues from a broad range of application codes.
 - The CSE and SP teams ran a joint test of ISO9001 Business Continuity and Disaster Recovery processes on Tuesday 8th October. The test was very successful with no impact on the actual ARCHER service and staff involved are now better equipped to deal with any real major incident
- eCSE:
 - All eCSE calls are now complete with 99 of the 100 awarded projects having finished; the remaining project is in progress and is due to finish by the end of the ARCHER Service. We will continue to support this project and collect the remaining final reports.
- Outreach and Engagement:
 - We attended New Scientist Live in London in October. This 4-day event allowed us to showcase the importance of HPC to society and the economy.
 - We had an exhibit at CIUK, showcasing the ARCHER service to the UK HPC community and beyond. This provided a good opportunity to engage with the Tier 2 community, sharing experiences and best practice.
 - We hosted the RSE Leaders meeting in Edinburgh in December. The focus was on the definition of skills and competencies that define an RSE.
- Training:
 - We delivered 11 days (243 student-days) of face-to-face training in the quarter at 5 different locations (with an average feedback score better than “Very Good”) plus 3 days (80 student-days) of live-broadcast online training. Together with 1 day of online tutorials, this takes the yearly total to 72 days as required.
 - The new 1-day online workshop on the popular LAMMPS package was very well received with 38 attendees and very good feedback.
 - We are continuing our collaboration with the Alan Turing Institute and ran a course on “Introduction to Spark for Data Scientists” at their headquarters in central London. This is an excellent training venue and the course had 35 attendees.

3. Collaborations and Outputs Summary

- Presentations:
 - Analysis of parallel I/O use on the UK national supercomputing service, ARCHER using Cray's LASSi and EPCC SAFE", A. Turner, D. Sloan-Murphy, K. Sivalingam, H. Richardson, J. Kunkel, Speaker J.F.R. Herrera, 2019 CCS-EPCC Workshop, Tsukuba, Japan.
 - "The ARCHER eCSE Programme", Speaker A. Turner, RSE Impact Showcase, CIUK, 5,6 Dec 2019, MCCC, Manchester, UK
- Meetings:
 - A. Turner, HPC User Forum, 10-11 Oct 2019, University of Edinburgh, UK
 - A. Turner, HPC-SIG, 6 Nov 2019, University of Manchester, UK
 - A. Turner, Society of Research Software Engineering Board of Trustees meeting, 25 Nov 2019, University of Southampton, UK
 - L. Smith, A. Simpson, RSE Leaders Meeting, 19 Dec 2019, Bayes, Edinburgh, UK
- Exhibitions and Outreach Events:
 - J. Herrera, J. Kennedy, N. Spyro, C. Wood, A. Grant, O. Brown, E. Belikov, D. White, New Scientist Live, 9-10 October 2019, Location: London, UK
 - A. Grant, Dundee Science Festival, 19th-20th October, Dundee, UK.
 - J. Beech-Brandt and A. Turner, ARCHER Showcase Booth, Research Zone, CIUK, 5,6 Dec 2019, MCCC, Manchester, UK

4. Forward Look

- Centralised CSE Team:
 - Our primary focus is on supporting users as they finalise their work on the system as the end of service approaches. In particular, ensuring users are supported when moving data from the ARCHER system.
 - We are looking to ensure that relevant knowledge and data for application package accounts is maintained and transferred to ARCHER2.
 - We will also focus on supporting the Consortia as they move from ARCHER to the Tier-2 resources for the transition period.
- eCSE:
 - A final review of finances will be conducted to ensure all eCSE projects have claimed their funds.
 - The last few remaining final reports will be reviewed and scientific highlights added to the web page.
- Outreach
 - Preparation is underway to attend the Big Bang Fair in March at the NEC in Birmingham. This is an ideal event to engage and encourage school children interested in a career in computational science.
- Training:
 - To help users in the transition to ARCHER2, we are re-running a virtual tutorial we originally held in September 2019 on “Transferring Data from ARCHER at End of Service”. This is timely as we near ARCHER end of service, and we now have more concrete details on the status of the RDF.
 - The use of containers to package computational applications and aid in portability to new platforms is becoming increasingly popular. This is of relevance to potential users of ARCHER2, so we are running a new 1-day course on this topic in Durham in February.

5. Contractual Performance Report

This is the contractual performance report for the ARCHER CSE Service for the Reporting Periods: October 2019, November 2019 and December 2019.

The metrics were specified by EPSRC in Schedule 2.2 of the CSE Service Contract.

CSE Query Metrics

- **QE1:** The percentage of all queries notified to the Contractor by the Help Desk in a Quarter that the Contractor responds to, and agrees a work plan with, the relevant End User within 3 working hours of receiving the notification from the Help Desk. *Service Threshold: 97%; Operating Service Level: 98%.*
- **QE2:** The percentage of all queries notified by the Help Desk to the Contractor that have been satisfactorily resolved or otherwise completed by the Contractor within a 4-month period from the date it was first notified to the Contractor. *Service Threshold: 80%; Operating Service Level: 90%.*
- **TA1:** The percentage of all technical assessments of software proposals provided to the Contractor by the Help Desk in any Service Period that are successfully completed by the Contractor within 10 days of the technical assessment being provided to the Contractor by the Help Desk. *Service Threshold: 85%; Operating Service Level: 90%.*
- **FB1:** The percentage of End User satisfaction surveys for CSE queries carried out in accordance with the Performance Monitoring System by the Contractor showing the level of End User satisfaction to be “satisfactory”, “good” or “excellent”. *Service Threshold: 30%; Operating Service Level: 50%.*

Period	Oct-19		Nov-19		Dec-19		Q4 2019	
Metric	Perf.	SP	Perf.	SP	Perf.	SP	Perf.	Total
QE1	100%	-2	100%	-2	100%	-2	100%	-6
QE2	100%	-2	100%	-2	100%	-2	100%	-6
TA1	100%	-1	100%	-1	100%	-1	100%	-3
FB1	100%	-2	100%	-2	100%	-2	100%	-6
Total		-7		-7		-7		-21

Training Metrics

- **FB2:** The percentage of all training satisfaction surveys (carried out in accordance with the Performance Monitoring System by the Contractor) in each Quarter that are rated “good”, “very good” or “excellent”. *Service Threshold: 70%; Operating Service Level: 80%.*

Period	Oct-19		Nov-19		Dec-19		Q4 2019	
	Perf.	SP	Perf.	SP	Perf.	SP	Perf.	Total
FB2	97%	-1	100%	-1	100%	-1	98%	-3
Total		-1		-1		-1		-3

*Pink – Below Service Threshold
 Yellow – Below Operating Service Level
 Green – At or above Operating Service Level*

The 97% in October relates to one survey rated as “bad” for a course on “Introduction to Spark for Data Scientists”; this is discussed in more detail in the training section.

Service Credits

Period	Oct-19	Nov-19	Dec-19
Total Service Points	-8	-8	-8

6. CSE Queries

Queries Resolved in Reporting Period

Metric Descriptions

In-Depth	All technical queries passed to ARCHER CSE team
Course Registration	Requests for registration on ARCHER training courses
Course Enquiry	Enquiries about courses
Technical Assessment: <Category>	Request for Technical Assessments of applications for ARCHER time
eCSE Application	Queries relating to eCSE applications

A total of 372 queries were resolved by the CSE service in the reporting period.

Metric	Oct-19	Nov-19	Dec-19	Total
Course Registration	164	53	79	296
In-depth	14	6	6	26
Course Enquiry	15	3	5	23
Technical Assessment: Grant	8	6	3	17
Technical Assessment: Instant	2	1	1	4
Technical Assessment: RAP	1	2	1	4
Technical Assessment: Leader	0	1	0	1
eCSE	1	0	0	1
Total	205	72	95	372

3 query feedback responses were received on In-depth queries in the reporting period. This represents an 8% return rate for feedback forms. All responses registered a score of "Excellent".

Resolved In-Depth queries fell into the following categories:

Category	Number of Queries	% Queries
3rd party software	19	73.1%
Compiler and system software	3	11.5%
Batch system and queues	1	3.8%
Access to HPC	2	7.7%
User programs	1	3.8%

In-Depth Query Highlights

A small number of In-Depth queries have been selected to illustrate the work of the centralised CSE team over the reporting period.

Query 1: "Excellent response, thanks very much for the help"

In this query a user had encountered issues installing a particular code on ARCHER, following the general instructions on the code's web site. In particular the user was struggling to combine the instructions on using python on ARCHER with the general code installation instructions. Kevin installed the software and created a set of tailored installation instructions for the user, which allowed the user to successfully install the code on ARCHER.

This is representative of a number of in-depth queries, where general installation instructions for a code need to be tailored to help a user install the software on ARCHER.

Query 2: "The issue was dealt in timely and possible solution plan was put in place."

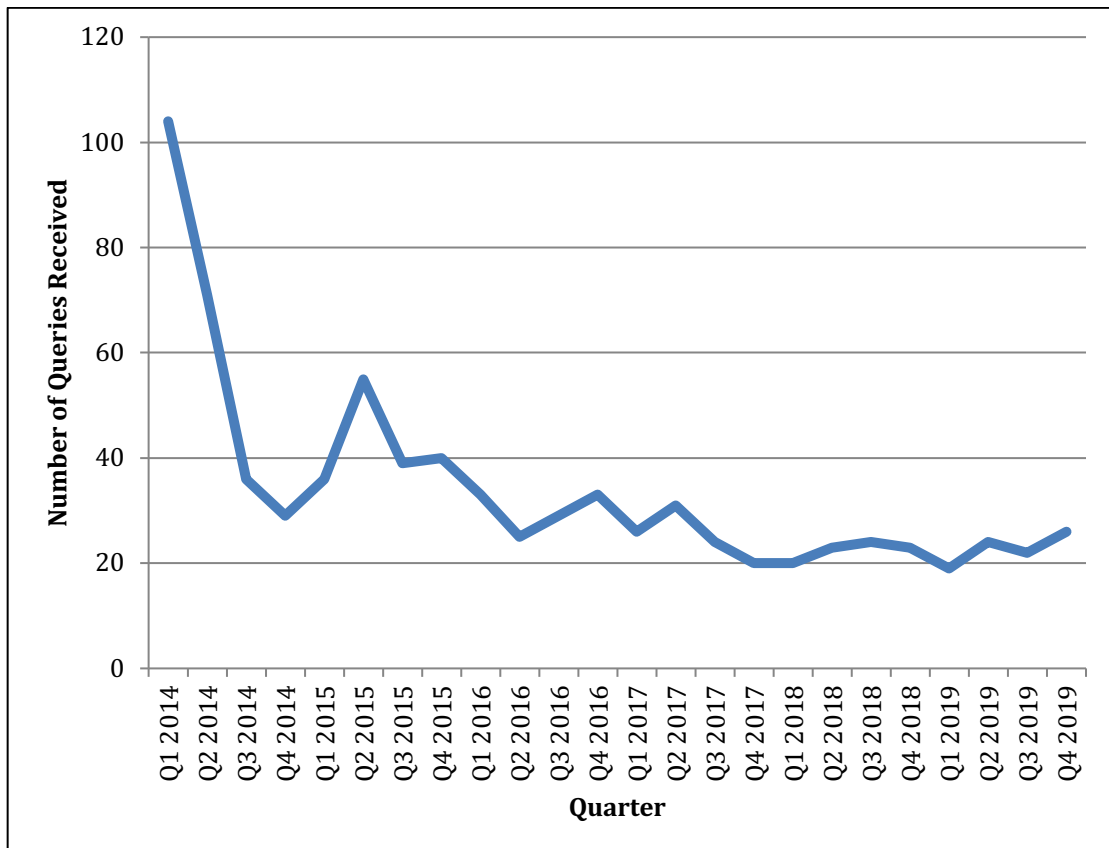
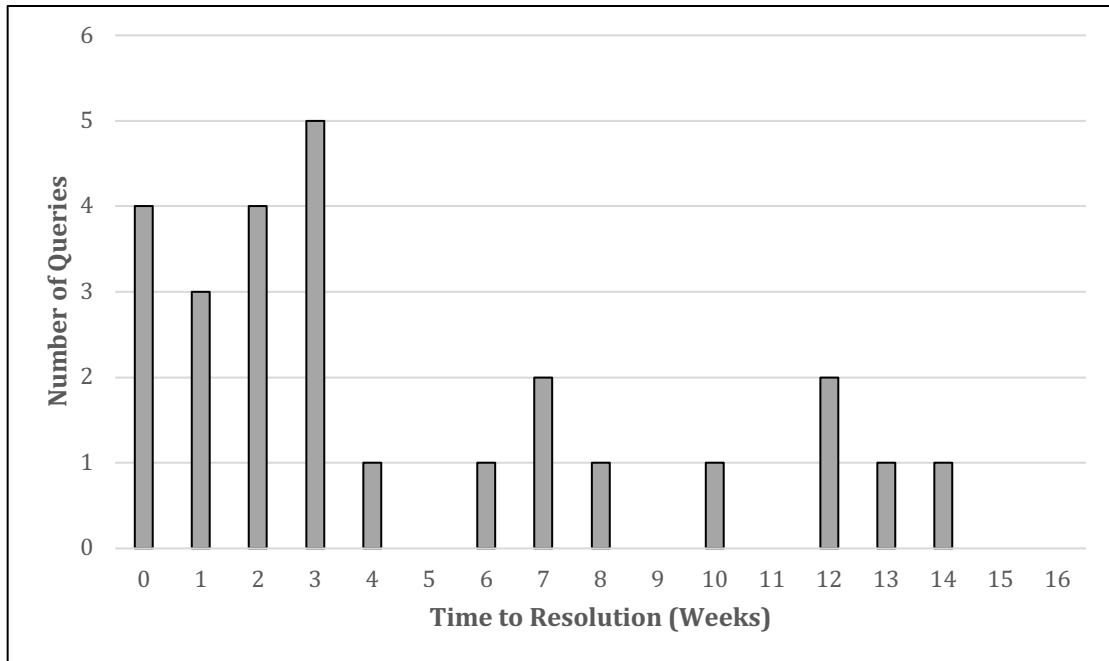
A user experienced a problem when running NAMD on ARCHER. His jobs would run some of the time but would fail most of the time. Mark was able to identify a problem with the SMP version of the code, within CHARMM++. An acceptable workaround was found until the problem is fixed in a later version of CHARMM++.

This is also representative of a number of queries, where a work around is found to enable users to keep running and generating science, while a longer-term fix is implemented by developers in a third party code.

In-Depth Query Analysis

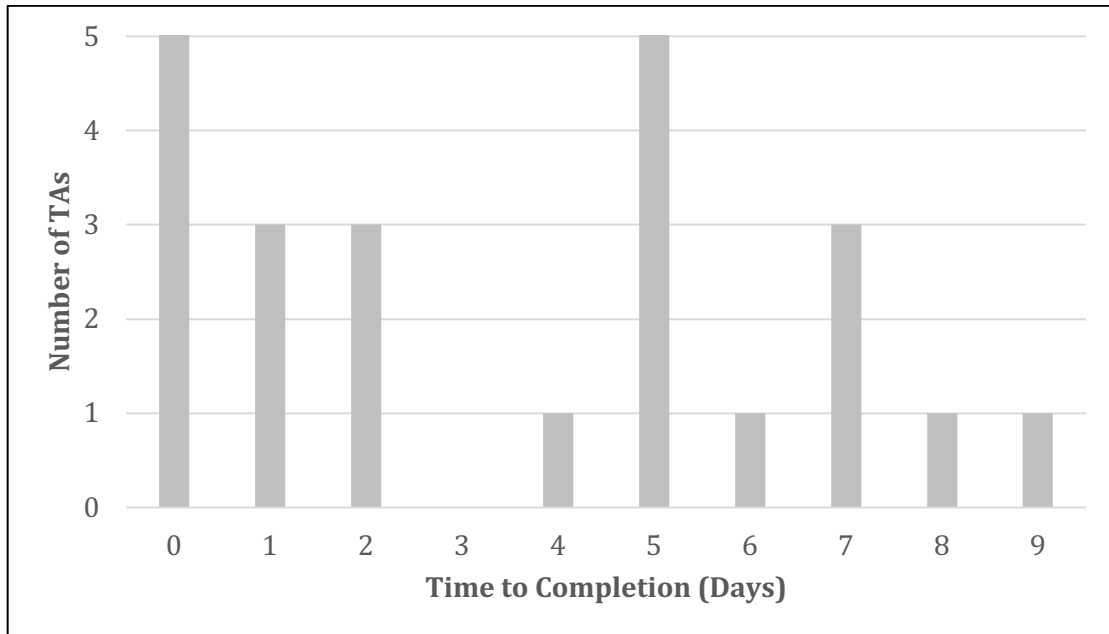
The histogram below shows the time to resolution for In-Depth queries in the current reporting period.

Plot of numbers of In-Depth queries received per quarter:

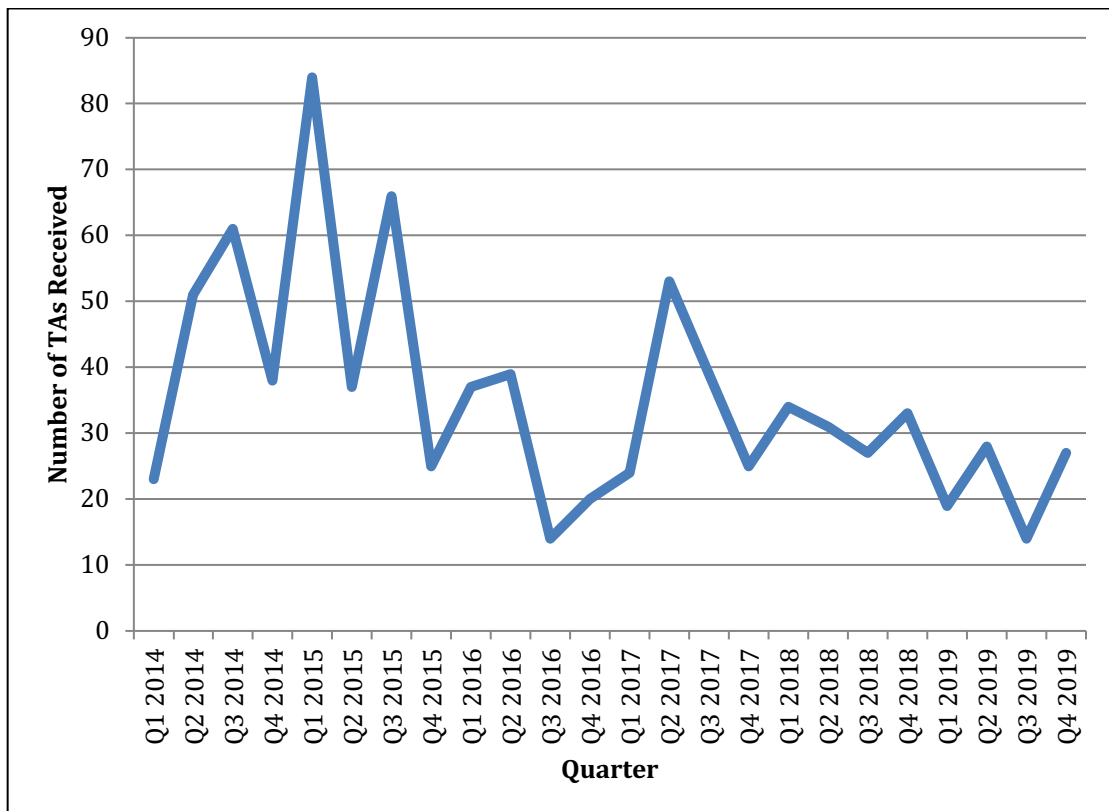


Technical Assessment Analysis

A histogram of the time to completion for Technical Assessments is shown below.



Plot of numbers of Technical Assessments received per quarter:



7. Centralised CSE Team: Continual Service Improvement

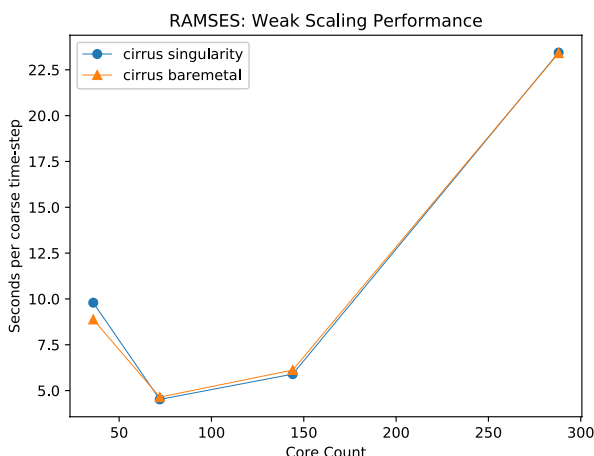
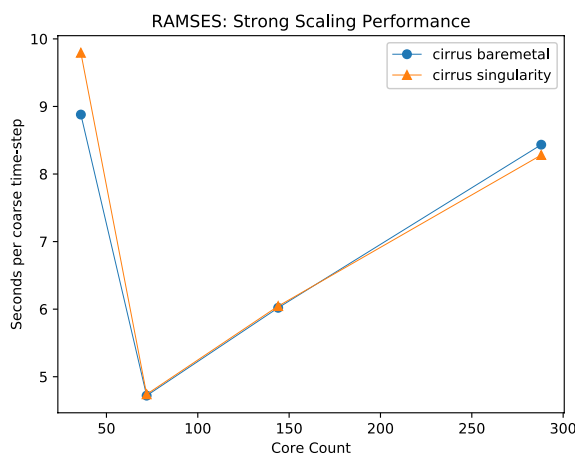
In collaboration with user groups and the other Service partners, the CSE service identified several priority service improvement areas to invest technical effort from the centralised CSE team. This section provides highlights from the reporting period.

Containers for HPC

Containers for HPC is a topic of interest to the community, particularly as we move to ARCHER2. Investigative work has been carried out on the Tier-2 system, Cirrus, with a view to transitioning work to ARCHER2.

Following on from our previous report, the RAMSES code has now been successfully installed within a Singularity v3.2.1 container running Ubuntu 18.04. The containerised RAMSES code was run repeatedly on the Tier-2 Cirrus system using a standard astrophysical benchmark to generate strong scaling and weak scaling results. Such scaling results (the “singularity” results) were then compared with previous scaling results from running RAMSES on Cirrus compute nodes directly (the “baremetal” results) and showed no appreciable difference as demonstrated in the following figures.

The next stage is to setup a Singularity container running ARCHER benchmark codes, namely CASTEP, CP2K and GROMACS. The prerequisites and required work to be done has been analysed and planned.



8. Training

This quarter, the CSE Service has provided a total of 11 days (243 student-days) of face-to-face training across 5 different locations, 3 days (80 student-days) of live-broadcast online training, and 1 day of online tutorials (average attendance 7 per tutorial).

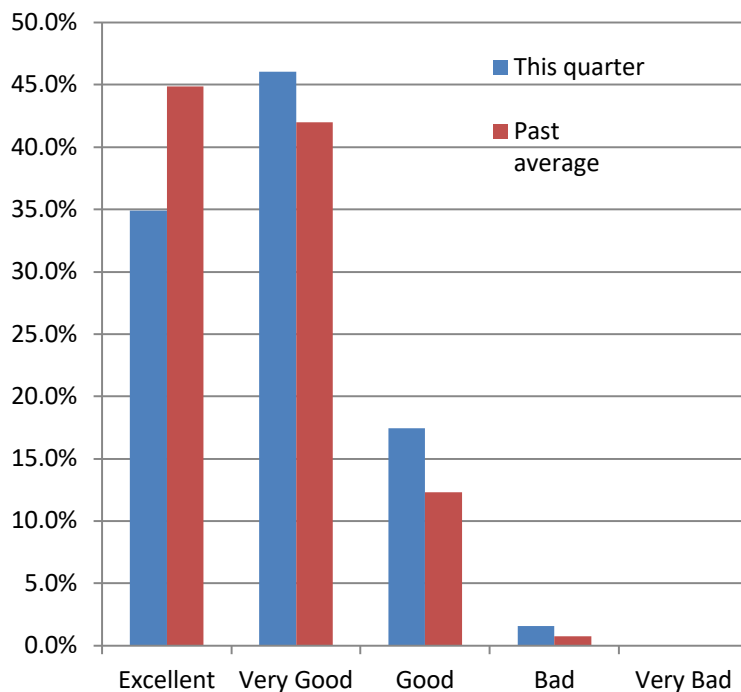
Month	Dates	Course	Location	Days	Attendees
Oct 2019	16 – 23	LAMMPS Workshop	Online	1*	38
	29-30	Data Carpentry	Cardiff	2	21
	30 Oct-1 Nov	Hands-on Introduction to HPC for Life Scientists	Birmingham	3	13
	31 Oct-1 Nov	Introduction to Spark for Data Scientists	London	2	35
Nov 2019	13	Fully Lagrangian Dynamical Core for the UKMO/NERC Cloud Model	Online	0.5	
	13, 20, 27 Nov & 4 Dec	Shared-Memory Programming with OpenMP	Online	2*	21
Dec 2019	4	Parallelisation of LISFLOOD	Online	0.5	
	9-10	HPC Carpentry	Edinburgh	2	16
	16-17	Scientific Programming with Python	Belfast	2	30

* These online interactive courses were run in an extended format over consecutive Wednesday afternoons, giving time for attendees to attempt the practical exercises between sessions and to raise any issues with the trainer at the following webinar.

On the feedback for face-to-face courses, attendees rate the course on a scale of 1-5 (“Very bad”, “Bad”, “Good”, “Very good” and “Excellent”). The average feedback using this metric was 4.1, i.e. better than “Very Good”. Users provided 63 feedback forms, a response rate of 55%.

We received one score of “Bad” for the “Introduction to Spark” course. Looking at the more detailed feedback, this student wanted the course to focus on different subject areas. However, the online course description is explicit about the material to be covered so we have taken no action here.

We are increasing efforts to get user feedback for online training in addition to face-to-face courses. Response rates are typically much lower, but for the online LAMMPS course we did receive 7 feedback forms with an average score of 4.0 (i.e. “Very Good”) so there is some quantitative evidence that the online courses are well received.



We are planning for 9.5 days of training in 2020 – with a service end date of 18th February, this is pro-rata against a quarterly target of 18 days.

Month	Dates	Course	Location	Days
Jan 2020	7 - 8	Hands-on Introduction to HPC	Edinburgh	2
	9 - 10	GPU Programming with CUDA	Edinburgh	2
	15 - 22	Using non-volatile Memory	Online	1*
	22	Transferring Data from ARCHER	Online	0.5
	27 - 28	Advanced MPI	London	2
Feb 2020	29	Cray Compilation Environment	Online	0.5
	3	Reproducible Computational Environments using Containers	Durham	1
	12	The EIS-2 Tool	Online	0.5

* This online interactive course will run over two consecutive Wednesday afternoons, giving time for attendees to attempt the practical exercises between sessions and to raise any issues with the trainer at the second webinar.

9. Embedded CSE (eCSE)

eCSE Scientific Highlights

As we approach the end of the ARCHER service, we continue to create and publish scientific highlights for completed projects. A small number of these have been selected to illustrate the work over the reporting period.

PANDORA Upgrade: Particle Dispersion in Bigger Turbulent Boxes

Prof. John Shrimpton (University of Southampton), Thorsten Wittemeier (University of Southampton), Dr. David Scott (EPCC)

PANDORA is a pseudo-spectral flow solver providing solutions to the Navier-Stokes equations in a cubic domain for isotropic and homogenous turbulent flow. A new version has been implemented within this eCSE project. The main aim of the project was to overcome the limitations of the previous code in terms of memory use and efficiency.

First of all, a significant amount of the data stored in the old code was not necessary. The memory use of the code has been drastically reduced by keeping only those data that are strictly necessary. PANDORA 2.0 uses only about a tenth of memory for fluid simulations compared to the previous version. Another important change concerns the parallelisation of the code. By changing the parallelisation from a one-dimensional to a two-dimensional decomposition, a further significant reduction in memory use could be achieved.

So far the code has been successfully tested on simulation grids up to 8192^3 . The strong scalability of simulations with particles has proven to be nearly ideal. Fluid simulations show reasonably good scaling behaviour within certain limits, which allows for a good performance on all domain sizes. The new code gives access to simulations with large numbers of particles at high Reynolds numbers even where only limited resources are available. Levels of turbulence that have not been studied with particles are now accessible. The new code can be used to perform simulations on bigger computing grids, enabling us to study turbulent flows at higher Reynolds number.

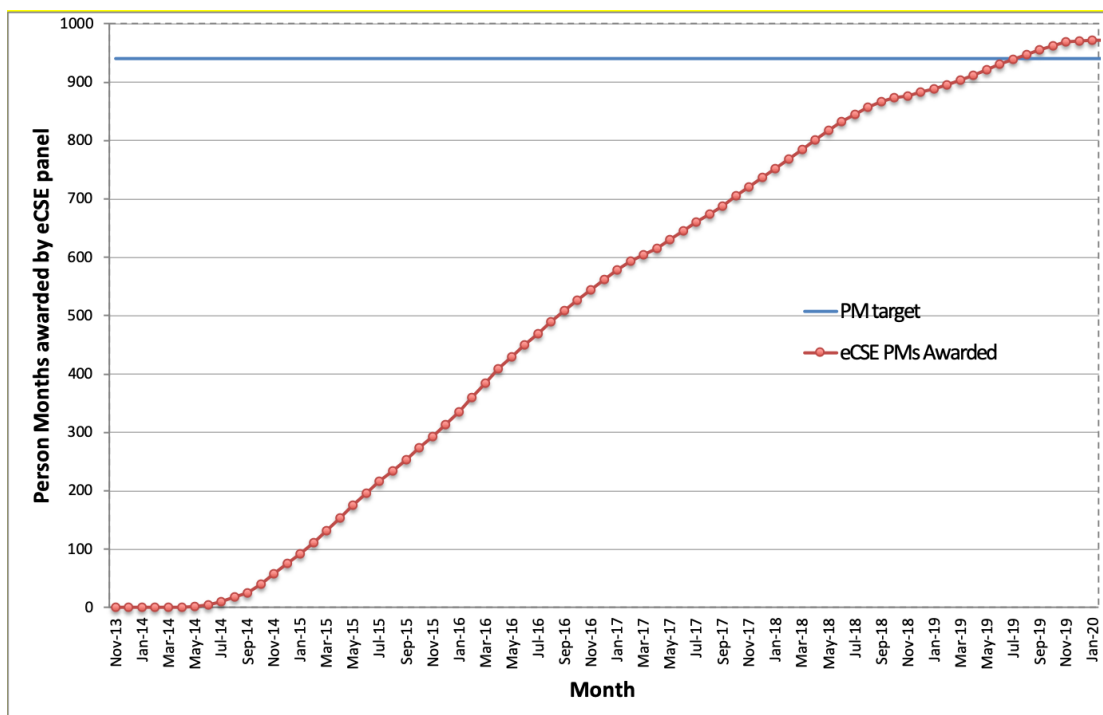
VAMPIRE: Billion-atom Simulations of Magnetic Materials

Dr. Richard Evans (University of York) and Rory Pond (University of York)

Magnetic materials are essential to a wide range of technologies, from data storage to cancer treatment to permanent magnets used in wind generators. New developments in magnetic materials promise huge increases in performance of devices but progress is limited by our understanding of magnetic properties at the atomic scale. Atomistic spin dynamics simulations provide a natural way to study magnetic processes on the nanoscale, treating each atom as possessing a localised spin magnetic moment. Such approaches are computationally expensive, requiring parallel computers to perform simulations of more than a few thousand atoms.

The aims of this eCSE project were to optimise the VAMPIRE code on ARCHER and improve the data input/output routines to enable configuration data to be extracted from the simulation to see the time evolution of the atomic spins. The improved code enables a new class of magnetic materials simulation containing between 10,000,000 and 1,000,000,000 atoms. Such large simulations give unprecedented insight into the behaviour of complex magnetic materials for realistic situations.

Overview of eCSE effort



The eCSE person months awarded up to and including the 13th eCSE call are shown in red.

We committed to awarding at least 941 person months by the end of the project (14 FTEs for 5 years, and 8.4 FTEs for year 6). In practice, 973 person months were awarded across 100 eCSE projects meaning an extra 32 person months were awarded over the duration of the programme.

For each of the eCSE calls, the table below summaries the number of proposals, number of projects awarded, number of person months awarded, number of projects completed and number of final reports received.

All projects have now completed except for one and this is due to complete before the end of the service. Our risk analysis identified this project as of low risk and we anticipate this will complete on time.

eCSE call	No. proposals	No. projects awarded	No. person months awarded	No. projects started	No. projects completed	No. final reports received	Notes
eCSE01	19	14	132	14	14	14	
eCSE02	17	9	82	9	9	9	
eCSE03	16	10	96	10	10	10	
eCSE04	16	8	82	8	8	8	
eCSE05	14	8	94	8	8	8	
eCSE06	9	5	47	5	5	5	
eCSE07	16	5	49	5	5	5	
eCSE08	21	8	88	8	8	8	
eCSE09	19	5	58	5	5	4	Due to extenuating circumstances reported previously, we will not receive a report for project eCSE09-6.
eCSE10	13	6	59	6	6	5	1 final report is late and being pursued.
eCSE11	18	6	49	6	6	4	2 late final reports are being pursued.
eCSE12	23	6	41	6	6	5	1 late final report is being pursued.
eCSE13	21	10	96	10	9	1	The remaining final reports are due by the end of the ARCHER Service.
Total	222	100	973	100	99	86	

10. Business Continuity and Disaster Recovery (BCDR)

The CSE and SP teams ran a joint test of ISO9001 Business Continuity and Disaster Recovery processes on Tuesday 8th October. This is in line with our commitment to run a full and realistic scenario, as part of continual service improvement, every two years.

This particular test was based around a scenario in which a significant portion of the team was affected by food poisoning and unable to work at short notice, following on from an EPCC work function. The test was very successful:

- Staff involved are confirmed to be better equipped to deal with any real major incident that may occur.
- There was no negative impact on the actual ARCHER service.
- Lessons learned from the previous test were confirmed to have been implemented effectively.
- A set of recommended improvements have been agreed as an outcome of the test.

A detailed report on the preparation, execution, and outcomes of the test has been produced and shared with EPSRC, as ARCHER Authority.

This test and the previous one demonstrate the benefits of regular BCDR testing, and even encourage more frequent testing. However, the resource requirements to implement each test makes this challenging. Therefore consideration will be given to running different types of BCDR testing that could be run more frequently---for example, a table-top exercise to review the response to various major-incident scenarios. However, we will also continue to exercise scenarios such as this one on a two-yearly basis.