ARCHER CSE Service Quarterly Report

Quarter 4 2018

1. Document Information and Version History

Version:	1.0
Status	Release
Author(s):	Andy Turner, David Henty, Chris Johnson
Reviewer(s)	Lorna Smith, Alan Simpson

Version	Date	Comments, Changes, Status	Authors, contributors, reviewers
0.1	2018-12-20	Skeleton document	Andy Turner
0.2	2019-01-08	Added Centralised sections	Andy Turner
0.3	2019-01-09	Added Outreach items	Lorna Smith
0.4	2019-01-10	Added eCSE sections	Chris Johnson
0.5	2019-01-10	Added Training sections	David Henty
0.6	2019-01-10	Reviewed	Alan Simpson
1.0	2019-01-10	Addressed reviewer comments	Andy Turner

2. Executive Summary

This report covers the period: 1 October 2018 to 31 December 2018 inclusive.

- Centralised CSE Team:
  - In collaboration with Cray, we have developed functionality in the SAFE to import per-job I/O statistics from the Cray LASSi monitoring tool. This data will provide users with insights into the I/O performance of their jobs, and allow service providers to better understand I/O requirements for the research community.
  - We have produced a report comparing the single node performance for the ARCHER benchmarks across the different processor technologies available on UK national HPC resources to help users to choose the best resource for their research. The report is currently with internal reviewers.
  - The ARCHER CSE team were highly visible at the Supercomputing 2018 conference in Dallas, USA. We highlighted the leading role the ARCHER CSE service plays within the world HPC community with activities (listed below in Section 3) on training, outreach, benchmarking, RSE careers and promoting sustainable scientific software development.
- Training:
  - We delivered 15 days (281 student-days) of face-to-face training in the quarter at 5 different locations, with an average feedback score better than "Very Good".
  - To make HPC training as widely accessible as possible, we ran an interactive online OpenMP course for the first time; this was delivered over four consecutive Wednesday afternoons with an average of 21 attendees.
  - In order to prepare UK computational scientists for the availability of new HPC processor hardware, we ran a 2-day hands-on course focusing on the new ARM64 processor as used in the Isambard Tier-2 and EPCC's Catalyst systems.
  - To promote the use of HPC to new user communities, we ran the new HPC Carpentry course at The Alan Turing Institute in London.
  - To assess the long-term impact of the training programme, we issued the sixth Training Impact Survey to all those who attended an ARCHER course in the 12 months prior to September 2018.
- eCSE:
  - All eCSE calls are now complete and 100 projects have been awarded. All projects have now started and are due to complete by the end of the ARCHER service. Effort will focus on supporting these projects to ensure successful delivery of the project work, maximising the benefit to the ARCHER community.
  - A new ARCHER case study has been released, based on an eCSE project to enable DNA simulations on ARCHER. This and other case studies help to showcase the scientific impact of the work of the eCSE programme. See: https://www.archer.ac.uk/casestudies/Archer_DNA_casestudy_LO.pdf

3. Collaborations and Outputs Summary

- Presentations:
  - *Open Source Benchmarking*, **Andy Turner**, Sheffield RSE Seminar, 30 Oct 2018, University of Sheffield
  - *Women in HPC: Diversifying the HPC Community*, **Weronika Filinger**, 11 Nov 2018, SC'18, Dallas, USA. https://sc18.supercomputing.org/session/?sess=sess160
  - *Exascale MPI*, **Daniel Holmes**, 11 Nov 2018, SC'18, Dallas, USA. https://sc18.supercomputing.org/session/?sess=sess174
  - *The Impact of MOOC Methodology on the Scalability, Accessibility and Development of HPC Education and Training*, **Weronika Filinger,** 12 Nov 2018, SC'18, Dallas, USA. https://sc18.supercomputing.org/?post_type=page&p=3479&id=ws_bphpcte110&sess=sess166
  - *Toward an HPC Certification Program*, **Weronika Filinger**, 12 Nov 2018, SC'18, Dallas, USA. https://sc18.supercomputing.org/?post_type=page&p=3479&id=ws_bphpcte106&sess=sess166
  - *Driving Asynchronous Events with Tasks*, **Nick Brown**, 12 Nov 2018, SC'18, Dallas, USA. https://sc18.supercomputing.org/presentation/?id=ws_opensu102&sess=sess175
  - *Heterogeneous Systems and the Road to Exascale for HPC and AI*, **Daniel Holmes**, 12 Nov 2018, SC'18, Dallas, USA. https://sc18.supercomputing.org/presentation/?id=pec330&sess=sess173
  - *Open Source Software*, **Nick Brown**, 13 Nov 2018, SC'18, Dallas, USA. https://sc18.supercomputing.org/?post_type=page&p=3479&id=pec448&sess=sess175
  - *Best Practices for Scaling-Up and Sustaining HPC Education, Outreach and Training*, **Andy Turner**, 13 Nov 2018, SC'18, Dallas, USA. https://sc18.supercomputing.org/presentation/?id=bof189&sess=sess441
  - *Strategies for Inclusive and Scalable HPC Outreach and Education*, **Weronika Filinger and Nick Brown**, 13 Nov 2018, SC'18, Dallas, USA. https://sc18.supercomputing.org/?post_type=page&p=3479&id=bof135&sess=sess391
  - *Innovative Approaches for Developing Accessible, Productive, Scalable HPC Training*, **Weronika Filinger**, 14 Nov 2018, SC'18, Dallas, USA. https://sc18.supercomputing.org/?post_type=page&p=3479&id=pan113&sess=sess297
  - *Software Engineers: Careers in Research*, **Andy Turner**, 14 Nov 2018, SC'18, Dallas, USA. https://sc18.supercomputing.org/?post_type=page&p=3479&id=bof144&sess=sess411
  - *The Message Passing Interface (MPI): Version 4.0 and Beyond*, **Daniel Holmes**, 14 Nov 2018, SC'18, Dallas, USA. https://sc18.supercomputing.org/presentation/?id=bof187&sess=sess437
  - *Software Engineering and Reuse in Computational Science and Engineering,* **Andy Turner**, 14 Nov 2018, SC'18, Dallas, USA.

https://sc18.supercomputing.org/presentation/?id=bof171&sess=sess426

- *Open Source Benchmarking,* **Andy Turner**, 15 Nov 2018, PRACE Booth Presentation, SC'18, Dallas, USA.
- *International HPC Certification Program*, **Weronika Filinger**, 15 Nov 2018, SC'18, Dallas, USA. https://sc18.supercomputing.org/?post_type=page&p=3479&id=bof126&sess=sess385
- *HPC Meets Real-Time Data: Interactive Supercomputing for Urgent Decision Making*, **Nick Brown**, 15 Nov 2018, SC'18, Dallas, USA. https://sc18.supercomputing.org/presentation/?id=bof152&sess=sess415
- *Parallel Programming*, **David Henty**, 20-21 Nov 2018, Data Intensive Science STFC CDT Network, University of Edinburgh

- Meetings:
  - *PRACE WP7 F2F Meeting*, **Alan Simpson, Xu Guo, Chris Johnson, Andy Turner, Dominic Sloan-Murphy**, 23 Oct 2018, University of Edinburgh
  - *UKCOMES Consortium Meeting*, **Kevin Stratford**, 5-6 Nov 2018, Royal Society, London
  - *ARCHER RAP Meeting(remote)*, **Andy Turner**, 28 Nov 2018, EPSRC, Swindon
  - The ARCHER team had a booth at New Scientist Live, in London on the 19th -23rd September. Held across 4 days, the audience had a broad range of backgrounds and ages and our focus was on showing the relevance of Supercomputing to the general public.
  - 6 staff members attended STEM ambassador training on the 6th November, the aim was to increase the pool of staff members involved in Outreach activities, to help increase the diversity and experience of those involved.
  - *HPC – Study on Impact Advisory Board*, **Alan Simpson**, 20 Nov 2018, EPSRC, Swindon

4. Forward Look

- Centralised CSE Team:
  - We will produce publicly-available reports in SAFE on the new I/O statistics imported from LASSi and publicise these to ARCHER users. These reports will help users understand the I/O performance characteristics of their applications and help them use ARCHER more effectively.
  - Based on the LASSi I/O data imported into SAFE, the CSE team will publish a report analysing I/O use on ARCHER by different research areas and applications to provide a better understanding of the I/O requirements of the ARCHER user community. We are also submitting a paper for the Cray User Group 2019 (CUG 2019) conference on this work.
  - We will publish the benchmarking report comparing the performance of the ARCHER benchmark applications across different processor architectures and publicise this to the UK HPC and research community. This will help researchers select the most appropriate HPC resource for their work.
- Training:
  - Exploiting EPCC's new location within the Bayes Centre at the University of Edinburgh, we are running two data-related courses in early January: the popular Data Carpentry course followed by an entirely new course on "Introduction to Spark for Data Scientists".
  - We will continue the interactive online training programme with a repeat run of last year's successful course "Online MPI".
  - To promote the use of modern programming methods in HPC, we will run a new 2-day course on "Modern C++ for Computational Scientists" in the second quarter of this year.
  - We will analyse the results from the sixth Training Impact Survey and produce an updated training impact report that will be circulated to ARCHER management and the Training Panel.
- eCSE:
  - The eCSE team will ensure that the benefits of the eCSE programme are captured and widely disseminated. In particular the focus will be on final reports, case studies and benefits realisation.
- Outreach:
  - We again have a booth at the Big Bang Fair in Birmingham in March 2019. Effort will focus on tailoring demonstrations and practicals for the audience, with an aim of explaining the real-world benefit of Supercomputing, and on encouraging young people to consider careers in computational science.

5. Contractual Performance Report

This is the contractual performance report for the ARCHER CSE Service for the Reporting Periods: October 2018, September 2018 and November 2018.

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-18**		**Nov-18**		**Dec-18**		**Q4 2018**	
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%	-4
Total		-7		-7		-5		-19

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

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-18		Nov-18		Dec-18		Q4 2018	
Metric	Perf.	SP	Perf.	SP	Perf.	SP	Perf.	Total
FB2	100%	-1	100%	-1	100%	-1	100%	-3
Total		-1		-1		-1		-3

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

Service Credits

Period	Oct-18	Nov-18	Dec-18
Total Service Points	-8	-8	-6

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 341 queries were resolved by the CSE service in the reporting period.

Metric	Oct-18	Nov-18	Dec-18	Total
Course Registration	137	87	39	263
eCSE Application	2	2	2	6
In-Depth	11	7	5	23
Course Enquiry	1	11	4	16
Technical Assessment: Grant	6	3	7	16
Technical Assessment: Instant	1	0	1	2
Technical Assessment: RAP	14	0	0	14
Technical Assessment: UoE	0	0	1	1
Total	**172**	**110**	**59**	**341**

5 query feedback responses were received on In-depth queries in the reporting period. This represents a 22% return rate for feedback forms. All responses registered a score of "Excellent". We continue to try to improve the response rate for feedback from queries by offering charity donations for responses and sending additional reminders to users to provide feedback.

Resolved In-Depth queries fell into the following categories:

Category	Number of Queries	% Queries
3rd party software	17	74%
Compilers and system software	3	13%
User programs	2	9%
User behaviour	1	4%

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.

Q1098540 Trouble installing Matlab Runtime

An ARCHER user wanted to run an ice sheet model that was written in Matlab on ARCHER. Although this should be theoretically possible using the Matlab Runtime (MCR) there were a lot of difficulties getting this up and running due to version dependencies between the version of Matlab required by the ice sheet model, the Matlab licence the user had, the MCR versions and the underlying GLIBC runtime on ARCHER. The CSE team explored a number of different MCR versions on ARCHER, identified a compatible version of MCR and produced a set of installation instructions for the user that would allow them to install a version of MCR that they could use to run their model on ARCHER.

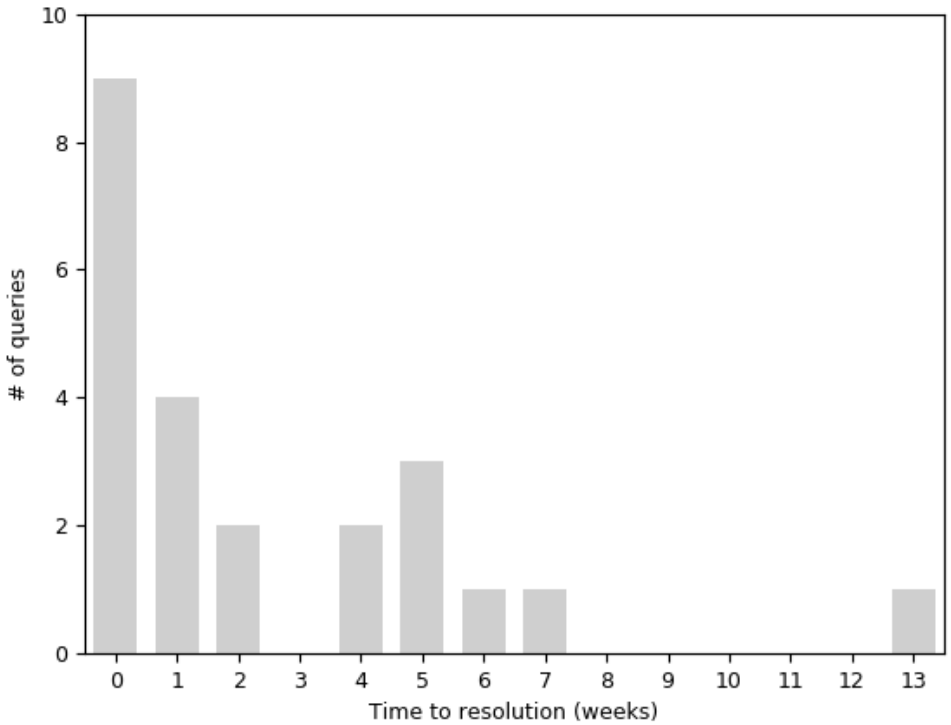
The impact of the CSE support in this query was highlighted by the user in feedback on the handling of the query: *"I think it's safe to close the ticket now - I will let you know if there are any similar problems in the future. But first, thank you so much for all your help!! I am continually impressed with the Archer helpdesk - I've had a couple of "in-depth" queries now and the level of service has just been fantastic. Compared to some of the other HPC facilities I've worked with in different countries, it's allowed me to move along with my science much more quickly, rather than spending ages sorting out these technical issues on my own."*

Q1074844: New OpenFOAM Solver

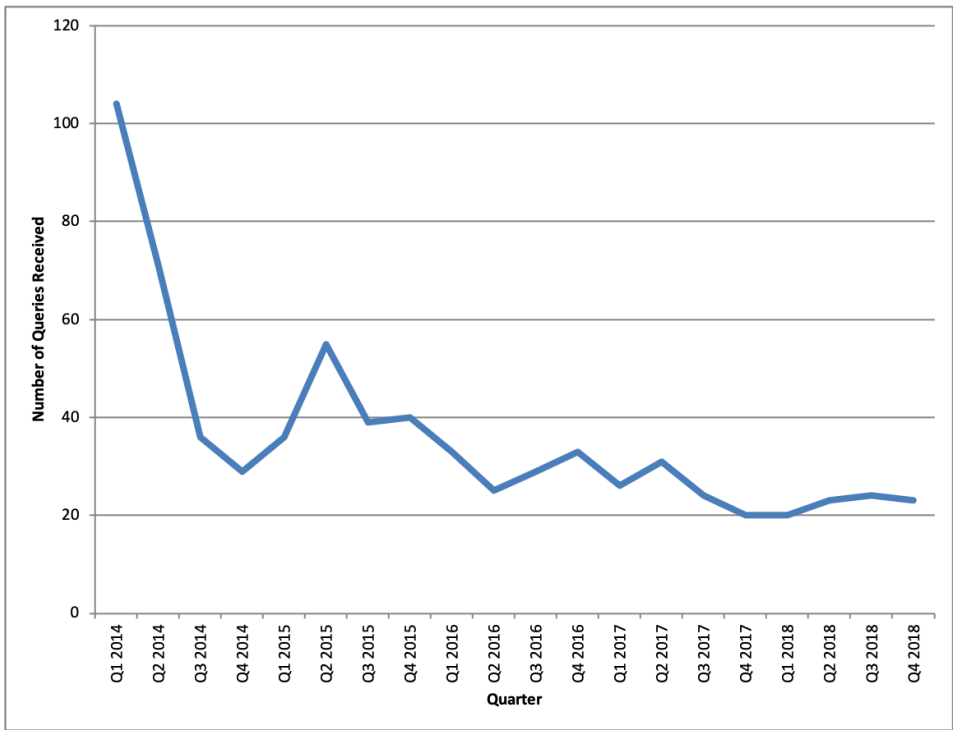
The CSE team received a request for support from a user wishing to use their own custom solver in the OpenFOAM computational fluid dynamics framework. The OpenFOAM framework has been designed to allow researchers to incorporate their own algorithms but it can be quite difficult for researchers to use this functionality within the framework – particularly on supercomputers such as ARCHER where the interconnect and libraries can be quite different from standard HPC clusters. In this case, the researcher required help with two issues. Firstly, ensuring that they were compiling in the correct way to include both the central part of the OpenFOAM tools and their custom routines. Secondly, ensuring that they had setup the runtime environment in their job submission script correctly to make sure the components were available from both their local, custom functionality and the central installation of OpenFOAM. The CSE team were able to provide support that allowed the user to get their complex setup up and running on ARCHER so that they could proceed with their research.

In-Depth Query Analysis

The histogram below shows the time to resolution for In-Depth queries in the current reporting period. The median resolution time during this period is 1-2 weeks (median resolution time since 1 Jan 2014 is 5-6 weeks).

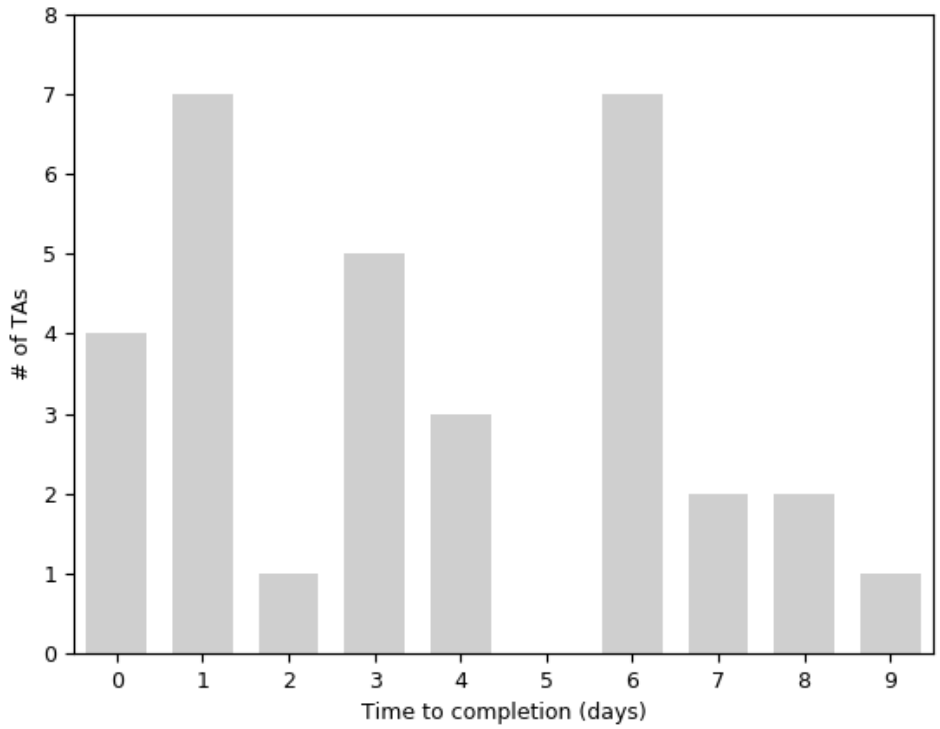


Plot of numbers of In-Depth queries received per quarter:



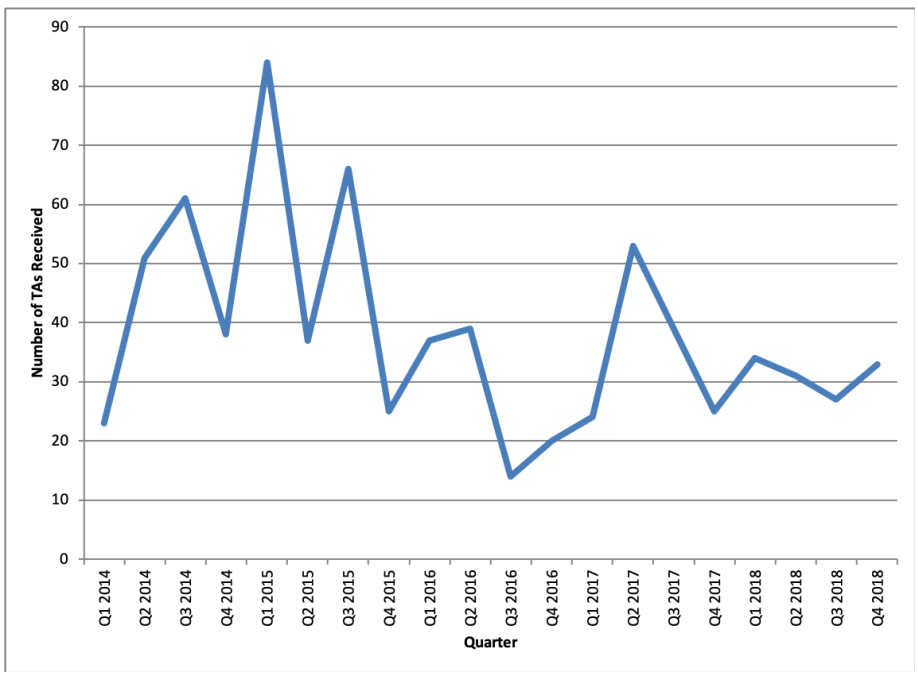
Technical Assessment Analysis

A histogram of the time to completion for Technical Assessments (see below) reveals that the median completion time for this quarter was 3 days (median completion time since 1 Jan 2014 is 3 days).



Plot of numbers of Technical Assessments received per quarter:

1.



2. 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 summarises progress in the reporting period in these areas.

Comparative Benchmarking

We have produced a report comparing the performance of different processor architectures on the ARCHER benchmark set. This report is currently out for internal review within the CSE team, once this review is complete, we will circulate for review by the Tier-2 TWG/RSE groups and, following these reviews, we expect it to be published in early 2019. The information in the report is designed to allow researchers to select the best architecture for their work from the range of different national HPC systems available across the UK.

For this initial architecture comparison, we limited the benchmarks to a single compute node so that we could remove effects from different interconnect types and ease the comparison. We plan to supplement the report with multinode comparisons in the future.

The table below is a sample from the draft report comparing the performance of the 1400k atom GROMACS benchmark across different architectures.

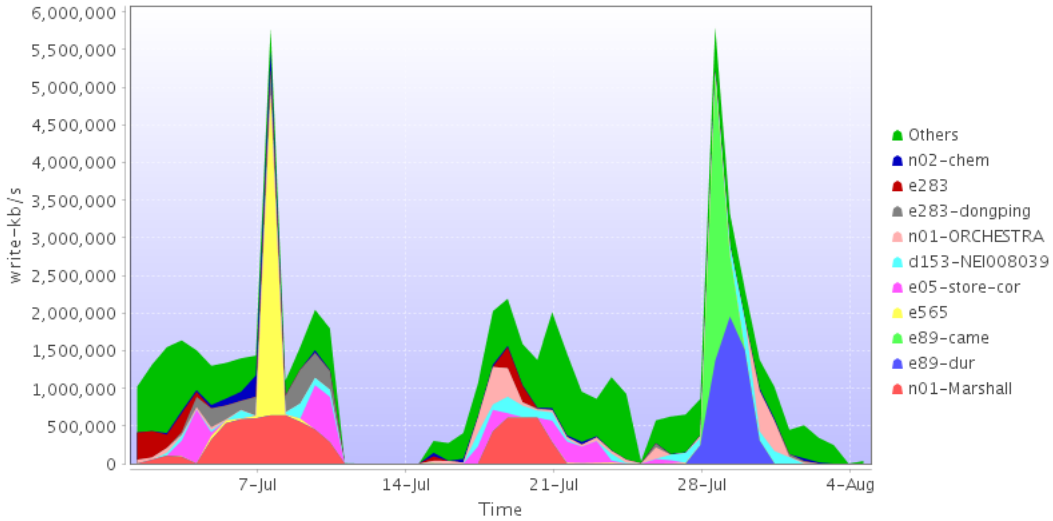
System	Performance (ns/day)	Performance relative to ARCHER node	Notes
Wilkes2-GPU	2.744	2.257	4 MPI tasks, 3 OMP per task, 4 GPU
Peta4-Skylake	2.082	1.712	32 MPI tasks, 1 OpenMP thread per task
Cirrus	1.699	1.397	36 MPI tasks, 2 OpenMP threads per task, 2-way SMT
Isambard	1.471	1.210	128 tasks, 2 OpenMP threads per task, 4-way SMT
Tesseract	1.323	1.080	24 tasks, 2 OpenMP threads per task, 2-way SMT
ARCHER	1.216	1.000	24 tasks, 2 OpenMP threads per task, 2-way SMT

Details of the different architectures can be found online at: https://github.com/hpc-uk/archer-benchmarks/blob/master/SystemDetails.ipynb

LASSi parallel IO data in SAFE

In collaboration with the Cray CoE and ARCHER SP, we have updated SAFE to import per-job I/O performance data from the Cray LASSi tool. This can be used by users and

service providers to gain insight on how I/O is used on the service. An example plot from SAFE on write bandwidth broken down by project group is shown below.



We are in the process of producing SAFE reports that can be used by users to analyse their I/O use; we plan to release and promote this functionality early in 2019.

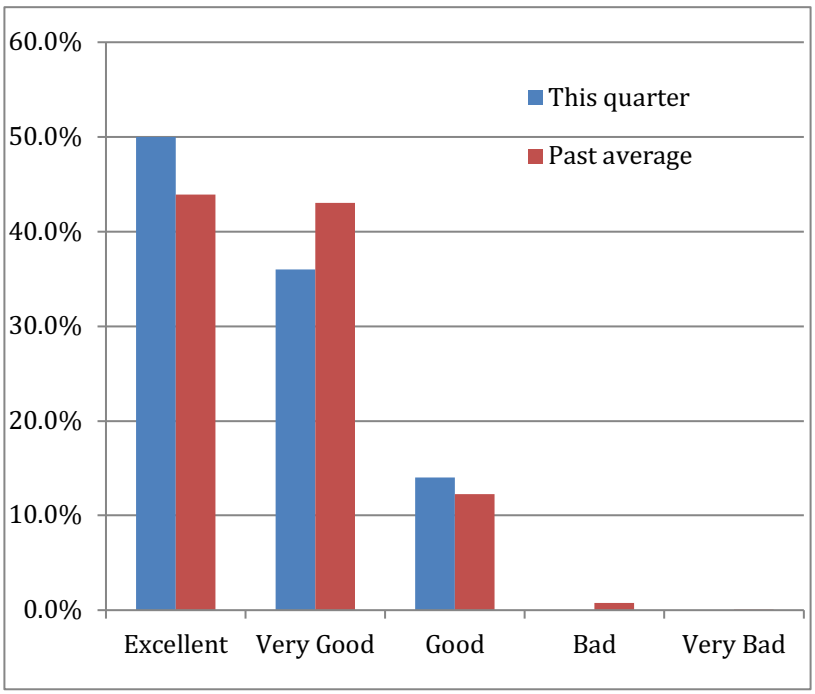
3. Training

This quarter, the CSE Service has provided a total of 15 days (281 student-days) of face-to-face training across 5 different locations, 2 days of interactive online training (average attendance 21 per session) and 1 day of online tutorials (average attendance 14 per tutorial).

Month	Dates	Course	Location	Days	Attend
Oct 2018	9	Hands-on Intro to HPC	Edinburgh	1*	7
	17	The TPLS Solver and PETsC	Online	0.5	
	24	OpenMP online (four consecutive Wednesday afternoons)	Online	2	
Nov 2018	6-8	Parallel Design Patterns	Oxford	3	17
	26-27	GPU Programming with CUDA	London	2	20
	28-30	Message-Passing Programming with MPI	Newcastle	3	23
Dec 2018	3-4	Programming the ARM64 processor	Edinburgh	2	14
	6-7	HPC Carpentry	London	2	21
	6-7	Scientific Python	Newcastle	2	22
	12	Isambard: The UK's Tier-2 ARM-based Supercomputer	Online	0.5	

*This course also formed part of Edinburgh University's celebration of Ada Lovelace Day 2018

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 4.4, i.e. better than "Very Good". Users provided 50 feedback forms, a response rate of 40%. This is a lower return rate than usual but, as several courses were in December, we expect to receive additional forms in the New Year.



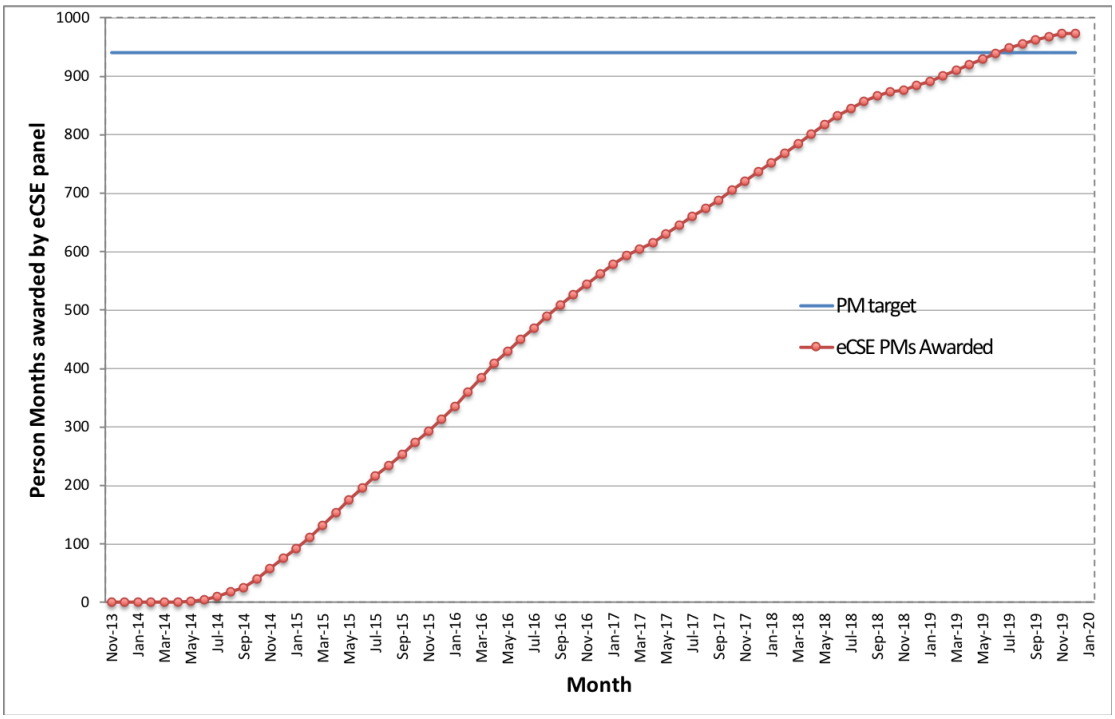
12 days of face-to-face training are already planned for the first quarter of 2019, plus 3 days online.

Month	Dates	Course	Location	Days	Attend
Jan 2019	8-9	Data Carpentry	Edinburgh	2	
	10-11	Introduction to Spark for Data Scientists	Edinburgh	2	
	17-18	Hands-on Introduction to HPC	Aberdeen	2	
	21-24	CCP9-MCC-UKCP-EPCC Workshop on Ab initio Periodic Codes	Daresbury	1*	
Feb 2019	5	Dynamic Load-Balancing in wsiFOAM	Online	0.5	
	20	Online MPI (over 4 consecutive Wednesday afternoons)	Online	2	
	26-28	Hands-on Introduction to HPC for Life Scientists	Birmingham	3	
Mar 2019	TBD	Efficient Parallel IO	London	2	
	20	Multi-level contact detection in Granular LAMMPS	Online	0.5	

*EPCC's role in this collaborative course is to enable attendees to run parallel codes on ARCHER, and we are counting one of the four days as ARCHER training.

4. Embedded CSE (eCSE)

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 FTE for year 6).
- 973 person months have been awarded across 100 awarded eCSE projects, meaning an extra 32 person months have been awarded during the programme.

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	9	1 late final report is being pursued.
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	6	2 late final reports are being pursued.
eCSE09	19	5	58	5	5	3	1 project completed early due to a staff member leaving and a final report is being pursued

							(see risk list below). The other final report is late and promised Q12019.
eCSE10	13	6	59	6	4	2	2 final reports are due during the next quarter.
eCSE11	18	6	49	6	6	1	4 late final reports are being pursued. The other 2 final reports are due during the next quarter.
eCSE12	23	6	41	6	6	4	6 final reports are due in the next quarter.
eCSE13	21	10	96	10	0	0	
Total	**222**	**100**	**973**	**100**	**88**	**74**	

- A risk analysis identified all projects as being of either low or very low risk apart from the following which were identified as being of medium risk:
    - eCSE04-10: the PI indicated that the person named to do the technical work may not be available
        - This project went ahead with the original staffing. There was a short delay to the start of the project which started on 01/01/16. The project is now complete and the final report is presently under review.
    - eCSE08-9: this project had a change of staffing
        - The new staff member was approved by the panel chair and the project has now finished. The project is now complete and the final report is presently under review.
    - eCSE09-6: this project has terminated early after the recent death of Dr Karl Wilkinson who was one of the Co-Is together with the fact that the researcher doing the work resigned from his current post in Cambridge in November 2018
        - The PI confirmed that the first work package is likely be completed and the project used half its allocated effort. Given the circumstances we agreed to this early termination and the unused funds were used to award eCSE12 projects at the final panel meeting. We have requested a final report to describe the work carried out but this has not yet been received.
    - eCSE09-8: this project was awarded 19 person months. This is a higher level of effort than awarded for other eCSE projects where 15 person months is the highest level of effort awarded so far
        - Of the 19 months awarded for this project, 7 were for a member of the ARCHER CSE team and the remaining 12 were for an external member of staff at the PI's institution. . The project is now complete and the final report is presently under review.
    - eCSE10-5: a change of staffing is required

        - We discussed this with the PI and the project was scaled back and re-staffed. The unused funds were used to fund eCSE12 projects. The project is now complete and we await the final report.
    - eCSE12-20: the project runs right up until 31 October 2018 – almost the end of the CSE contract
        - The project is being monitored via regular contact with the PI.
    - eCSE13-11: the project was awarded partial funding with 2 out of the 9 requested person months being awarded. The PI will need to confirm if he agrees to accept the partial funding. The contract for this project is presently awaiting acceptance.