ARCHER2
SP Quarterly Report

July - September 2023
EPCC
The University of Edinburgh

Document Information and Version History

Version:	1.0
Status	Release
Author(s):	Clair Barrass, Jo Beech-Brandt, Stephen Booth, Paul Clark, Daniel Gleed, Kieran Leach, Alan Simpson, Anne Whiting
Reviewer(s)	Alan Simpson

Version	Date	Comments, Changes, Status	Authors, contributors, reviewers
0.1	29/09/2023	Template created	Jo Beech-Brandt
0.2	02/10/2023	Updated	Anne Whiting
0.3	03/10/2023	Added heatmap and usage graphs	Clair Barrass
0.4	04/10/2023	Added narrative, metrics and graphs	Jo Beech-Brandt
0.5	04/10/2023	Added narrative	Paul Clark, Kieran Leach
0.6			
0.7	05/10/2023	Minor additions to forward look	Paul Clark
0.8	09/10/2023	Added Critical Success Factors; Review	Alan Simpson
1.0	09/10/2023	Version for UKRI	Alan Simpson, Jo Beech-Brandt

1 The ARCHER2 Service

This is the report for the ARCHER2 SP Service for the Reporting Period: 1st July – 30th September 2023.

1.1 Service Highlights

- EPCC were delighted to be able to announce that they have passed their annual external ISO audits (ISO 9001 for service delivery, ISO 27001 for information security and ISO 22301 for business continuity and disaster recovery) with flying colours. This reflects the importance we put on delivering a good and secure service for our users.

- As part of continued improvements to the underlying infrastructure at the ACF we upgraded our Data Centre Network software to new "Cumulus" software during September. Changes were made to the QoS layout on ARCHER to try to ensure that system utilisation remained high during the interruption to user access caused by this work. It was noted that a higher than usual number of large-scale jobs ran during this period, with users taking the opportunity to submit these via the low priority QoS.

- During the interruption to user service, scratch directories were deployed to both the fourth /work files system (fs4) and the NVMe file system, along with relevant ticketing. This is currently being integrated with SAFE for deployment to users.

- A mitigation to the 'Zenbleed' vulnerability which impacts the AMD CPUs in ARCHER2 was also deployed during the interruption to user service.

- We are investigating the implementation of Globus online to facilitate better data transfer to and from the Service.

- A detailed review of the Mechanical and Electrical design of site is being undertaken with new, more efficient hardware expected to be installed over the next year. This will include some of the ARCHER2 supporting infrastructure.

1.2 Forward Look

- We are awaiting the arrival of the GPUs for ARCHER2. We will work alongside our HPE colleagues to transition them into service and a User Forum is being planned to publicise their arrival and assist in user transition.

- Auto-deletion of files older than 28 days will be enabled on the /scratch file systems for both the /work file system, fs4 and the scratch NVMe file system. Users will be informed of this policy, and it will then be deployed.

- As a follow-on from the Mechanical and Electrical planning works commented on previously, we will be carrying out some cleaning of the internals of the large chillers which support the cooling processes. These will be communicated via the Change Advisory Board process and is not expected to be service impacting.

- A large number of the ARCHER2 SP team will be undertaking ITIL v4 training in November and Q1 2024.

- ARCHER2 will be represented at SuperComputing '23 in November which takes place in Denver. There will also be ARCHER2 SP representatives at CIUK '23 in December which is taking place in Manchester.

- We are working through the approach for the 2023 Change freeze and festive cover period for the ACF and service desk / SP functions.

2 ARCHER2 Performance Report

This is the contractual performance report for the ARCHER2 SP Service for the Reporting Periods from 1st July 2023 until 30th September 2023.

2.1 Service Points and Service Credits

The Service Levels and Service Points for the SP service are defined by EPSRC in Schedule 2.2 of ARCHER2 SP Service Contract.

The Working Day (WD) for the ARCHER2 Service is 10 Working Hours (WH) as the Service operates from 0800-1800. The Median Time to Resolution is measured in WD.

- **Availability**: *Service Threshold: <=96.5%; Operating Service Level: >98.0%, ≤ 98.5%.*
- **ARCHER2_SP_Level1 (MTR)**: The Median Time to Resolution, of all SP queries falling within Level 1 resolved by the Contractor in the Reporting Period. *MTR Service Threshold: >1 WD; Operating Service Level: >0.3 WD, ≤ 0.45 WD.*
- **ARCHER2_SP_Level2 (MTR)**: The Median Time to Resolution, of all SP queries falling within Level 2 resolved by the Contractor in the Reporting Period. *MTR Service Threshold: >8 WD; Operating Service Level: >2 WD, ≤4 WD.*
- **ARCHER2_SP_Level3 (MTR)**: The Median Time to Resolution, of all SP queries falling within Level 3 resolved by the Contractor in the Reporting Period. *MTR Service Threshold: >25 WD; Operating Service Level: >12 WD, ≤16 WD.*
- **Initial Response to Queries (%):** The percentage of the total number of SP queries assigned to the Contractor in the Reporting Period responded to within 3 Working Hours. *Service Threshold: <96.00%; Operating Service Level: 98.00 – 98.99%.*
- **Query User Satisfaction (%):** The percentage of the total number of query satisfaction surveys completed in each Reporting Period, rating the quality of the resolution of Queries by the Contractor as "Good", "Very Good" or "Excellent". *Operating Service Level: 82.00 – 87.99%*

2.1.1 Service Points

Metric	Jul 2023		Aug 2023		Sep 2023		Q3 2023	
	Perf	Points	Perf	Points	Perf	Points	Perf	Points
Availability	100%	-3	100%	-3	100%	-3	100%	**-9**
SP_Level1 (MTR)	0.00	-2	0.08	-2	0.00	-2	**0.00**	**-6**
SP_Level2 (MTR)	0.04	-2	2.32	0	0.07	-2	**0.07**	**-4**
SP_Level3 (MTR)	2.28	-2	0.00	-2	0.22	-2	**1.41**	**-6**
Initial Response (%)	100%	-1	100%	-1	100%	-1	**100%**	**-3**
Query Satisfaction (%)	100%	-2	100%	-2	100%	-2	**100%**	**-6**
Total		**-12**		**-10**		**-12**		**-34**

2.1.2 Service Credits

As the Total Service Points are negative (-34), no Service Credits apply in 23Q3.

2.2 SP Query Statistics

The metrics were specified by EPSRC in Schedule 2.2 of ARCHER2 SP Service Contract.
- **Assigned:** The number of SP queries assigned to the Contractor within each query resolution category in the Reporting Period.

- **Resolved:** The number of SP queries resolved by the Contractor within each query resolution category in the Reporting Period.
- **Backlog:** The number of SP queries assigned to the Contractor that remained unsolved within each query resolution category in the Reporting Period
- **Correspondence:** The average number of pieces of correspondence generated for SP queries in each query resolution category.
- **First Response:** The average time taken for the Contractor to first respond to the Originator of the SP query.

July 2023					
Service level	Assigned	Resolved	Backlog	Correspondence	First Response
SP_Level1	860	860	0	0.14	0:00:48
SP_Level2	83	99	19	8.73	0:14:09
SP_Level3	0	1	2	7	0:07:20
August 2023					
Service level	Assigned	Resolved	Backlog	Correspondence	First Response
SP_Level1	923	923	0	0.12	0:00:45
SP_Level2	83	92	10	7.91	0:17:39
SP_Level3	1	2	1	8.5	0:06:46
September 2023					
Service level	Assigned	Resolved	Backlog	Correspondence	First Response
SP_Level1	1147	1147	0	0.20	0:01:19
SP_Level2	73	67	16	7.32	0:27:10
SP_Level3	0	1	0	10	0:04:53
Q 2023					
Service level	Assigned	Resolved	Backlog	Correspondence	First Response
SP_Level1	2930	2930	0	0.15	0:01:03
SP_Level2	239	258	16	8.07	0:18:47
SP_Level3	1	4	0	8.5	0:06:26

2.3 Query Resolution

Metric	Jul 2023		Aug 2023		Sep 2023		Q3 2023	
Service Level	MTR	Resolved	MTR	Resolved	MTR	Resolved	MTR	Resolved
SP_Level1	0:00:10	860	0:00:02	923	0:00:34	1147	0:00:16	2930
SP_Level2	0:25:40	99	0:47:18	92	0:44:17	67	0:41:11	258
SP_Level3	22:48:56	1	23:14:55	2	2:14:56	1	14:03:57	4
Total		960		1017		1215		3192

A total of 3192 queries were resolved by the ARCHER2 SP Service in the Reporting Period. The percentage of user queries responded to within 3 hours was 100%.

2.4 Query Feedback

During July, there were 44 feedback scores received during this period. 100% were Good, Very Good or Excellent with 98% given the highest score of Excellent.

During August, there were 32 feedback scores received during this period. 100% were Good, Very Good or Excellent with 63% given the highest score of Excellent.

During September, there were 32 feedback scores received during this period. 100% were Good, Very Good or Excellent with 78% given the highest score of Excellent.
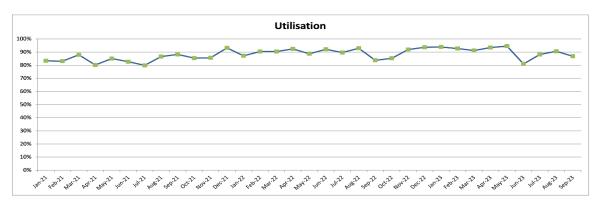
2.5 Maintenance and Outages

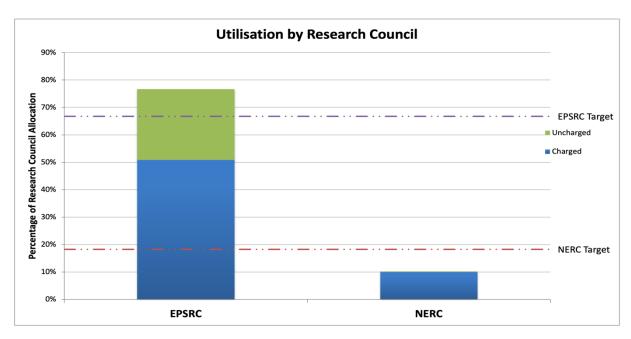| Type	Start	End	Duration	User Impact	Reason	Attributable
Partial Maintenance	18/09/23 0900	22/09/231200	4 days 3 hrs	Users were unable to connect to ARCHER2 or access their data. Jobs continued to run on the compute nodes.	Data Centre Network	Accommodation
Partial Maintenance	23/08/23 1000	22/09/231050	50 mins	Users were unable to submit new jobs	Change to slurm configuration	HPE

3 ARCHER2 Service Statistics

3.1 Utilisation

Utilisation from 1st July – 30th September is 89% which is slightly decreased from 90% the previous quarter. Utilisation for July was 88%, for August 91% and for September 87%.

Please note there was an outage during this period from 18th – 22nd September for the data centre network upgrade. This had little impact on the utilisation as changes were made to the QoS layout on the service before the upgrade to allow users to submit a larger number of jobs to the service. Despite users being unable to connect to ARCHER2, the compute nodes were operational, and jobs were able to run on the service throughout the outage.
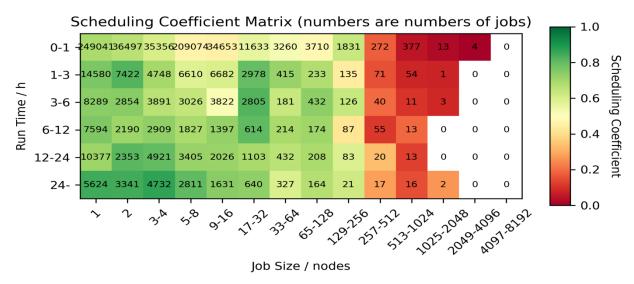

The utilisation by the Research Councils, relative to their respective allocations, is presented below. This bar chart shows the usage of ARCHER2 by the two Research Councils presented as a percentage of the total Research Council allocation on ARCHER2. It can be seen that EPSRC exceeded their target this quarter with their usage being at 76.7% (against their target of 66.8%) but NERC missed their target with utilisation being 10.2% (against their target of 18.2%). The increase in the proportion of uncharged utilisation for EPSRC may be due to the relaxation of the low priority QoS limits during the Data Centre Network maintenance session.
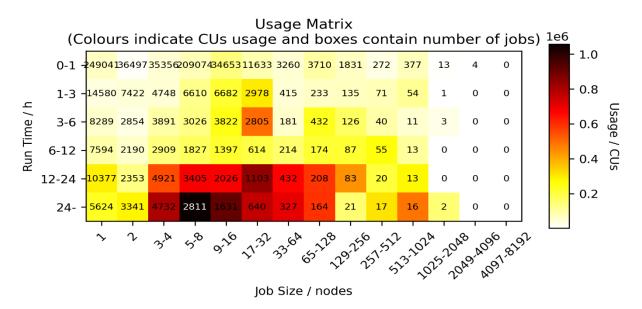
3.2 Scheduling Coefficient Matrix

The colour in the matrix indicates the value of the Scheduling Coefficient. This is defined as the ratio of runtime to runtime plus wait time. Hence, a value of 1 (green) indicates that a job ran with no time waiting in the queue, a value of 0.5 (pale yellow) indicates a job queued for the same amount of time that it ran, and anything below 0.5 (orange to red) indicates that a job queued for longer than it ran.

Scheduling Coefficient Matrix (numbers are numbers of jobs)

The usage heatmap below provides an overview of the usage on ARCHER2 over the quarter for different job sizes/lengths. The colour in the heatmap indicates the number of CUs expended for each class, and the number in the box is the number of jobs of that class.

It should be noted that there was an increase in the number of larger sized jobs during this quarter as users were encouraged to submit larger jobs during the data centre network maintenance session.

Usage Matrix
(Colours indicate CUs usage and boxes contain number of jobs)

Appendix: Critical Success Factors

1. Context

EPCC have been asked by UKRI to provide quarterly data for a number of critical success factors:
- CSF04 Implementation of environmentally considerate energy policies
- CSF07 Deliver and maintain a reliable data I/O function
- CSF08 Be cost-effective, cost-efficient and drive towards lowering of operational costs

In the sections below, please find the relevant metrics and data.

2. CSF04 Implementation of environmentally considerate energy policies

Implementation of environmentally considerate energy policies with a drive to reducing costs and environmental impacts.

All electricity provided to the ACF and ARCHER2 is on a 100% green, renewable energy tariff.

Environmentally considerate policies: 3

Since the start of full Service, EPCC have worked on implementing the following policies:
- Move from High Performance Mode to Low Power Mode: reduced average power draw from 3.2 MW to 2.9 MW (9%) with negligible input on performance [May 2022]
- Reduced default processor frequency: further reduced average power to around 2.5 MW (19%) [December 2022]
- Increase in coolant temperatures: this will result in an increase in passive cooling ("free cooling") [ongoing]

Power Usage

	4Q21*	1Q22	2Q22	3Q22	4Q22	1Q23	2Q23	3Q23
Average Power	3.31	3.16	3.15	2.86	2.90	2.51	2.56	2.46

* Partial

So far, the average power draw has been reduced by around 0.7MW (21%) which will reduce electricity usage by up to 6M kWh per annum, significantly reducing annual running costs.

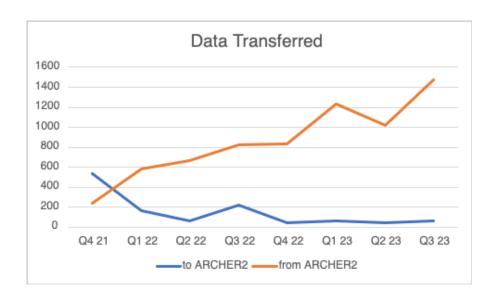# 3. CSF07 Deliver and maintain a reliable data I/O function

The compute resource will deliver and maintain an efficient, effective and reliable data I/O function which meets the requirements of users and their software. It will evolve and expand to accommodate new software or hardware architectures as required by the Service or its user base.

Data Transferred

EPCC monitor the data transfer rates in and out of the ARCHER2 system. Based on this, we now estimate the total amount of data transferred on and off ARCHER2 each Quarter.

Data Transferred…	4Q 21*	1Q 22	2Q 22	3Q 22	4Q 22	1Q 23	2Q 23	3Q 23
… **to** ARCHER2 (TB)	534	163	68	220	44	67	42	65
…**from** ARCHER2 (TB)	236	582	667	822	834	1231	1022	1472

* Partial

Parallel IO Write Performance

We regularly monitor the parallel write performance between the compute nodes and the parallel Lustre (/work) file systems. We use the benchio synthetic IO benchmark application (https://github.com/davidhenty/benchio) and report the MPI-IO write performance with the following settings:
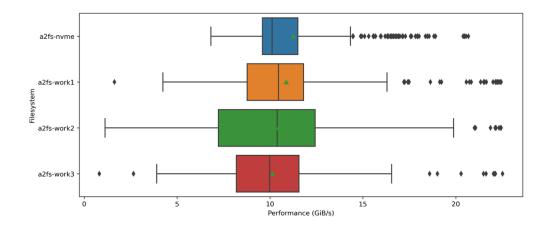
- Global data structure of 20483: writes a single file of 65,536 MiB (64 GiB).
- Uses 16 compute nodes and 128 MPI processes per node.
- Uses UCX as the MPI transport protocol.
- Sets the following environment variables:
    - FI_OFI_RXM_SAR_LIMIT=64K
    - MPICH_MPIIO_HINTS="*:cray_cb_write_lock_mode=2,*:cray_cb_nodes_multiplier=4

These settings have been found to maximise the IO performance for parallel writes using MPI-IO on the ARCHER2 file systems. Writes using the default settings on ARCHER2 typically have median write values 2-3 GiB/s lower than the optimised values.

Original reporting of this data (Q1 and Q2 2023) used the means from a small number of runs on the HDD-based Lustre file systems. From Q3 2023 onwards we have been monitoring performance regularly on both HDD and NVMe-based Lustre file systems throughout the quarter and report median (Q2) and

lower (Q1) and upper quartile (Q3) performance, as well as providing a quarterly boxplot that illustrates the performance variation.

Benchio MPI-IO medium (GiB/s)	1Q23	2Q23		3Q23
a2fs-work1	8.2	7.6±0.5		10.5 (8.8:11.8)
a2fs-work2	8.5	7.3±0.6		10.4 (7.2:12.4)
a2fs-work3	8.3	9.6±0.7		10.0 (8.2:11.6)
a2fs-nvme				10.1 (9.6:11.5)

4. CSF08 Be cost-effective, cost-efficient and drive towards lowering of operational costs

The Service shall be cost-effective and cost-efficient across its elements during its lifetime and drive towards lowering of operational costs by seeking efficiencies in delivery such that TCO presents an acceptable and cost-effective solution for the public. The Service will monitor and report its Power Usage Effectiveness (PUE) and strive to make efficiency savings where possible.

Relative Research Output

Measure	11/2021 – 5/2022	5/2022 – 12/2022	1Q23	2Q23	3Q23
Relative Research Output per kWh	100	109	115	115	115

We define the initial measure of research output per KWh on ARCHER2 to be 100, and then estimate how this has changed with the introduction of the various environmentally considerate policies discussed under CSF04. This is estimated using applications benchmarks similar to those defined by UKRI for the procurement.

Energy Used per CU Delivered

	4Q21*	1Q22	2Q22	3Q22	4Q22	1Q23	2Q23	3Q23*
Energy per CU (kWh)	0.719	0.713	0.728	0.715	0.650	0.545	0.669	0.585

*partial

Energy Cost per CU Delivered

	4Q21*	1Q22	2Q22	3Q22	4Q22	1Q23	2Q23	3Q23*
Cost per CU (£)	£0.089	£0.090	£0.098	£0.096	£0.088	£0.074	£0.162	£0.143

*partial

The two tables above are calculated using the total CUs delivered by ARCHER2, the total kWh of electricity consumed, and the unit cost for kWh. The increase in "Energy Cost per CU Delivered" from 2Q23 is caused by a significant increase in the unit cost of electricity from April 2023. For 2Q23, there is also an impact on the "Energy Used per CU Delivered" from the major software upgrade that took 3 weeks.