ARCHER2
SP Quarterly Report

October - December 2024

EPCC
The University of Edinburgh

Document Information and Version History

Version:	1.0
Status	Release
Author(s):	Clair Barrass, Jo Beech-Brandt, Alan Simpson, Anne Whiting
Reviewer(s)	Alan Simpson

Version	Date	Comments, Changes, Status	Authors, contributors, reviewers
0.1	24/12/2024	Template created	Jo Beech-Brandt
0.2	06/01/2025	Added query feedback data	Anne Whiting
0.3	06/01/2025	Added narrative, graphs, service statistics	Jo Beech-Brandt
0.4	08/01/2025	Added narrative and statistics	Andy Turner
0.5	10/01/2025	Minor additions	Paul Clark
0.6	14/01/2025	Added critical success metrics	Lorna Smith
0.7	15/01/2025	Reviewed	Alan Simpson
1.0	15/01/2025	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: 1 October – 31 December 2024.

1.1 Service Highlights

- We continue to work to improve the user experience with the Slurm scheduler on ARCHER2. An update to the Slurm fairshare priority scheme was implemented based on links between shares within Slurm and the size of allocations that projects have been granted. This change was designed to make wait times for jobs starting to be as equitable as possible between different projects (irrespective of their allocation size). Analysis of the wait times under the new scheme has shown that the change had the desired impact and so this is now the default setup on ARCHER2. We are planning to use this scheme and the experience gained in implementing it to improve the experience for users on other HPC services that we operate.

- ARCHER2 service staff attended SuperComputing in Atlanta in November and the CIUK Conference in Manchester in December. ARCHER2 staff also attended the HPC-AI Advisory Council meeting and the ExCALIBUR workshop in Leicester in October.

- A new training registration process has been implemented which allows users to register for training courses directly within the SAFE. This creates a more streamlined process for users and ARCHER2 staff. It also enables easier management of courses.

- Low priority QoS limits were changed over the festive period to ensure that utilisation remained high and users were notified. They will be returned to the usual limits in January 2025.

- Migration of core services supporting the ARCHER2 service to new Operating Systems to improve performance and reliability.

- Completed management cabinet power swaps to help improve resilience to the service within Computer Room 3 (cr3).

- Due to lower levels of staff availability over the festive period, a successful change freeze was implemented from 16th December to 6th January. This did not impact any business as usual activities for the running of the service.

- An initial analysis of Globus Online data transfers has been conducted which showed that ~900 transfers have been completed to data at rates of up to 17.8 Gbit/s.

1.2 Forward Look

- EPCC are engaged in high-level discussions with HPE about the future of the system software. The aim is to ensure good information security throughout the service, while minimising disruption to users. As part of the visit to SC24, Luke Davies from UKRI and EPCC met with a number of the HPE software team to discuss options further.

- The ARCHER2 Systems team have been working closely with the EPCC training team and a new MSc module in System administration has been developed. This starts in January 2025 and will help to build core skills in HPC system support.

- EPCC are working with HPE towards the deployment of a server running "View for Clusterstor" software. This should enable better understanding of work file system performance and a more immediate ability to identify the origin of problems on work file systems.

- Further analysis of data transfer, including via Globus Online, is planned for the coming quarter.

2 ARCHER2 Performance Report

This is the contractual performance report for the ARCHER2 SP Service for the Reporting Periods from 1 July 2024 until 30 September 2024.

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	Oct 2024		Nov 2024		Dec 2024		Q4 2024	
	Perf	Points	Perf	Points	Perf	Points	Perf	Points
Availability	99.7%	-2	100%	-3	100%	-3	**99.9%**	**-8**
SP_Level1 (MTR)	0.00	-2	0.00	-2	0.00	-2	**0.00**	**-6**
SP_Level2 (MTR)	0.06	-2	0.06	-2	0.06	-2	**0.06**	**-6**
SP_Level3 (MTR)	5.40	-2	0.00	-2	0.00	-2	**5.40**	**-6**
Initial Response (%)	100%	-1	100%	-1	100%	-1	**100%**	**-3**
Query Satisfaction (%)	100%	-2	100%	-2	100%	-2	**100%**	**-6**
Total		**-11**		**-12**		**-12**		**-35**

2.1.2 Service Credits

As the Total Service Points are negative (-35), no Service Credits apply in 24Q4.

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.

October 2024					
Service level	**Assigned**	**Resolved**	**Backlog**	**Correspondence**	**First Response**
SP_Level1	1890	1890	0	0.104	0:00:49
SP_Level2	108	101	21	6.743	0:14:35
SP_Level3	1	2	0	16	0:10:14
November 2024					
Service level	**Assigned**	**Resolved**	**Backlog**	**Correspondence**	**First Response**
SP_Level1	1371	1371	0	0.129	0:00:56
SP_Level2	92	89	24	7.663	0:13:01
SP_Level3	0	0	0	0	0:00:00
December 2024					
Service level	**Assigned**	**Resolved**	**Backlog**	**Correspondence**	**First Response**
SP_Level1	1067	1067	0	0.100	0:00:55
SP_Level2	53	68	9	7.074	0:13:37
SP_Level3	0	0	0	0	0:00:00
Q4 2024					
Service level	**Assigned**	**Resolved**	**Backlog**	**Correspondence**	**First Response**
SP_Level1	4328	4328	0	0.111	0:00:53
SP_Level2	253	258	9	7.147	0:13:47
SP_Level3	1	2	0	16	0:10:14

2.3 Query Resolution

Metric	**Oct 2024**		**Nov 2024**		**Dec 2024**		**Q4 2024**	
Service Level	**MTR**	**Resolved**	**MTR**	**Resolved**	**MTR**	**Resolved**	**MTR**	**Resolved**
SP_Level1	0:00:37	1890	0:00:13	1371	0:00:12	1067	0:00:20	4328
SP_Level2	0:38:01	101	0:34:57	89	0:34:56	68	0:35:13	258
SP_Level3	53:59:09	2	0:00:00	0	0:00:00	0	53:59:09	2
Total		**1993**		**1460**		**1135**		**4588**

A total of 4588 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 October, there were 38 feedback scores received during this period. 100% were Good, Very Good or Excellent with 84% given the highest score of Excellent.

During November, there were 32 feedback scores received during this period. 100% were Good, Very Good or Excellent with 94% given the highest score of Excellent.

During December, there were 13 feedback scores received during this period. 100% were Good, Very Good or Excellent with 77% given the highest score of Excellent.

£83 donation was made to our chosen charity Save the Children with £1 donated per query feedback item received.

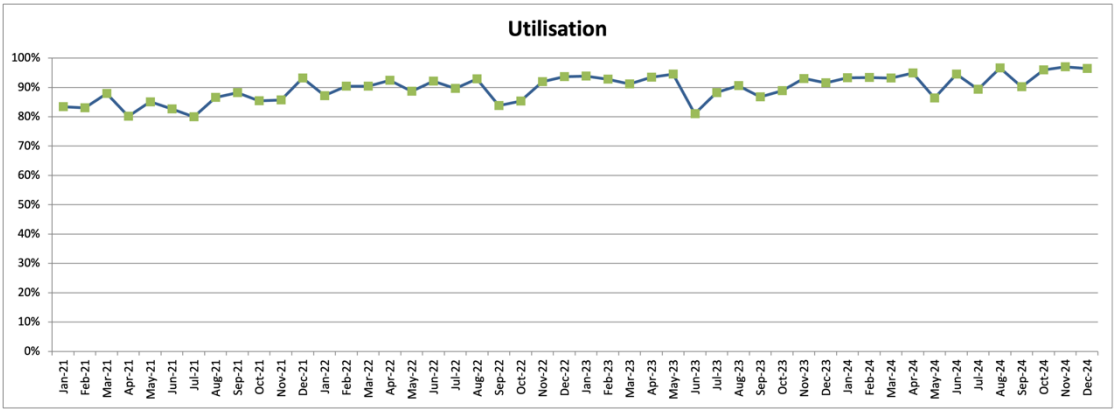
2.5 Maintenance and Outages

Type	Start	End	Duration	User Impact	Reason	Attributable
Partial	2024-12-01 1500	2024-12-02 1330	22 hrs 30 mins	Around 1000 compute nodes are unavailable. Jobs running on affected nodes at the time of failure will have crashed. The rest of the system remains available and jobs on unaffected nodes continue as usual.	Loss of power	Accomm-odation
Partial	2024-11-21 1140	2024-11-21 1200	20 mins	No users can access ARCHER2 from outside of the University of Edinburgh network. Running and queued jobs are unaffected.	Loss of power to part of the communication network	Accomm-odation
Partial	2024-10-24 0930	2024-10-24 1140	2 hrs 10 mins	Running jobs will continue to run, but new jobs will not start and short QoS is unavailable.	Updating slurm batch system	SP

3 ARCHER2 Service Statistics

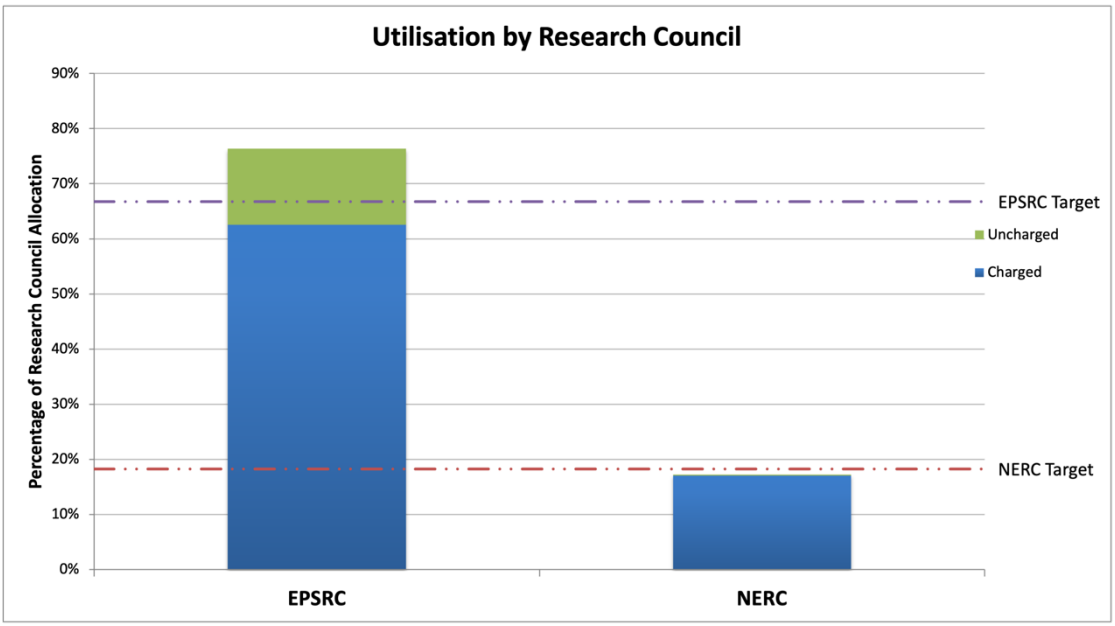
3.1 Utilisation

Utilisation from 1 October – 31 December is 96% which is even higher than the 94% from the previous quarter. Utilisation for October was 96%, for November 97% and for December 96%.

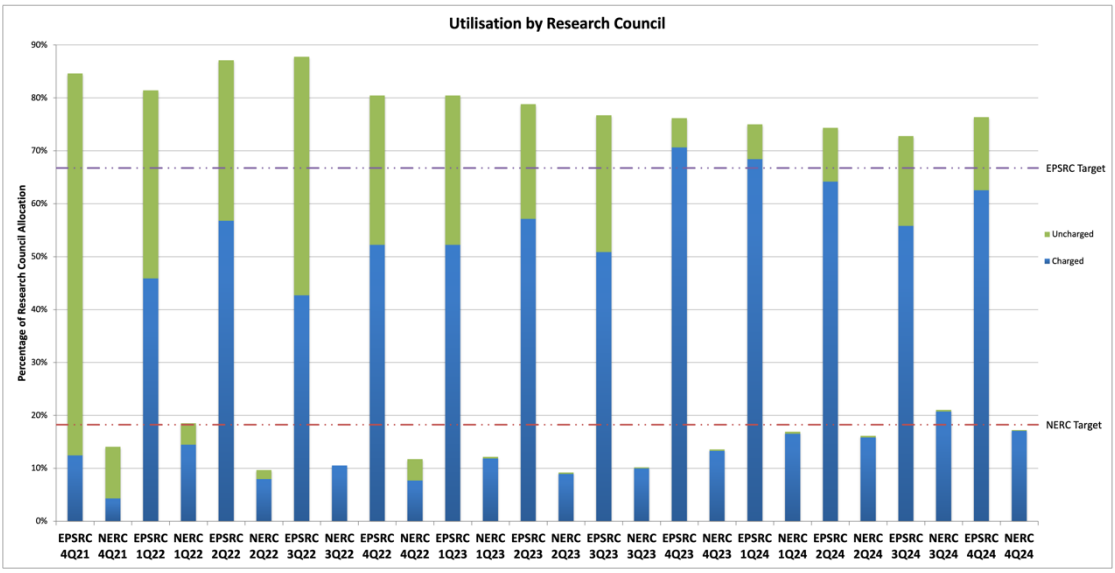


The utilisation by the Research Councils, relative to their respective allocations, is presented below. This bar chart shows the usage of ARCHER2 by the two Research Councils presented as a percentage of the total Research Council allocation on ARCHER2. It can be seen that EPSRC exceeded their target this quarter with their usage being at 76% (against their target of 66.8%). It should also be noted that the proportion of EPSRC's uncharged utilisation decreased this quarter from 17% in 3Q24 to 14% in this quarter.

NERC almost met their target with utilisation being 17% (against their target of 18.2%) which was a decrease from 21% in the previous quarter.

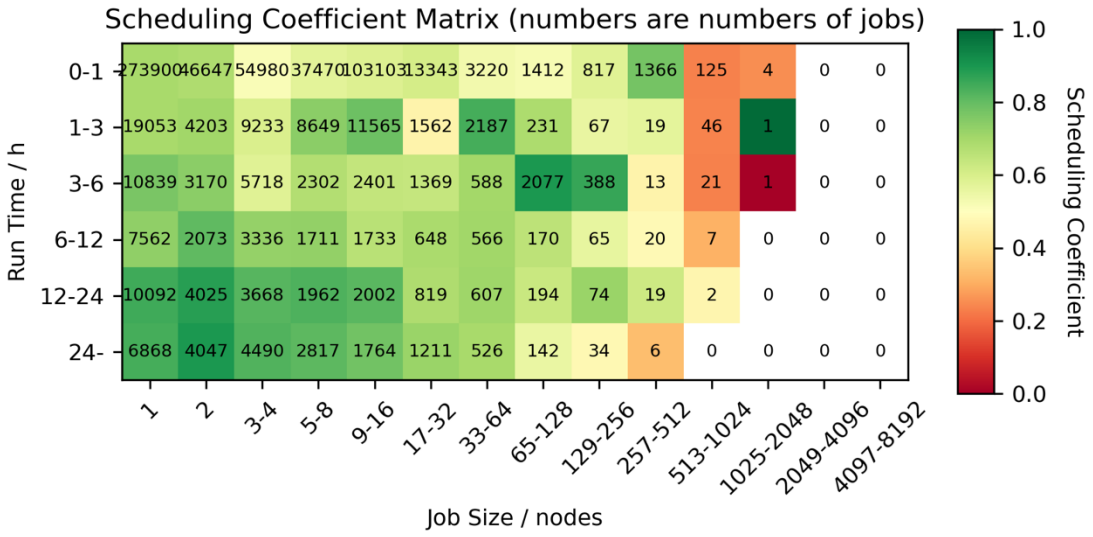


The stacked graph below shows the trend of charge and uncharged utilisation since the start of the service.



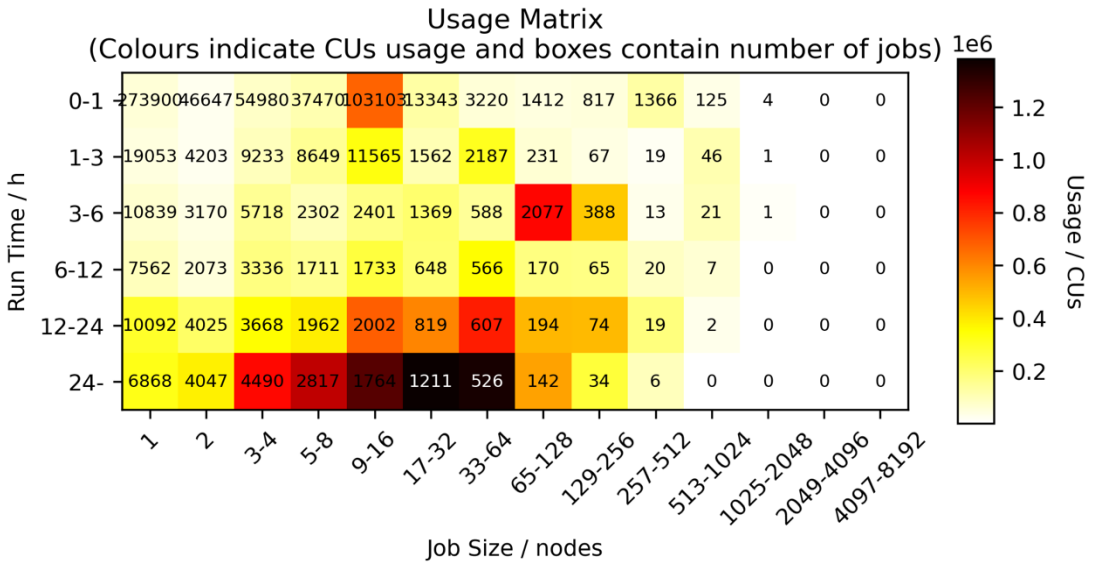
3.2 Scheduling Coefficient Matrix

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.



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)
1e6
Run Time / h
0-1
1-3
3-6
6-12
12-24
24-
273900 46647 54980 37470 103103 13343 3220 1412 817 1366 125 4 0 0
19053 4203 9233 8649 11565 1562 2187 231 67 19 46 1 0 0
10839 3170 5718 2302 2401 1369 588 2077 388 13 21 1 0 0
7562 2073 3336 1711 1733 648 566 170 65 20 7 0 0 0
10092 4025 3668 1962 2002 819 607 194 74 19 2 0 0 0
6868 4047 4490 2817 1764 1211 526 142 34 6 0 0 0 0
1
2
3-4
5-8
9-16
17-32
33-64
65-128
129-256
257-512
513-1024
1025-2048
2049-4096
4097-8192
Job Size / nodes
1.2
1.0
0.8
0.6
0.4
0.2
Usage / CUs

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: 4

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]
- Developed a set of new tools to help users estimate the environmental impact of their computing simulations and workloads [November 2024]

Power Usage

	4Q 21*	1Q 22	2Q 22	3Q 22	4Q 22	1Q 23	2Q 23	3Q 23	4Q 23	1Q 24	2Q 24	3Q 24	4Q 24
Average Power	3.31	3.16	3.15	2.86	2.90	2.51	2.56	2.46	2.53	2.58	2.54	2.64	2.57

* 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

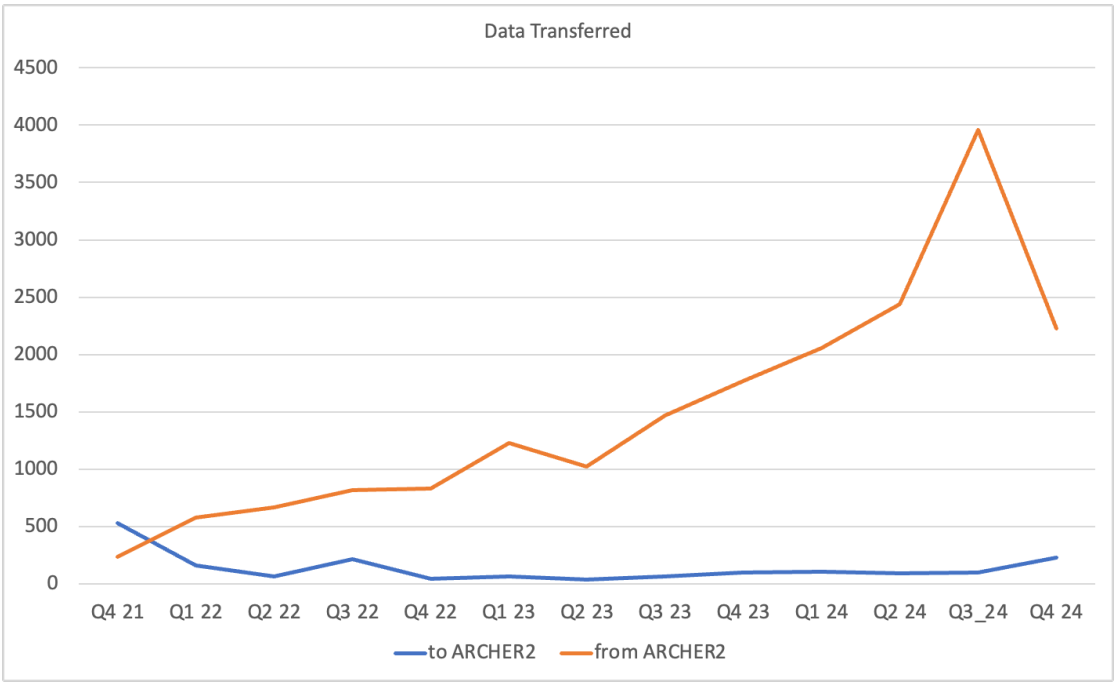
The compute resource will deliver and maintain an efficient, effective and reliable data I/O function which meets the requirements of users and their software. It will evolve and expand to accommodate new software or hardware architectures as required by the Service or its user base.

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	4Q 23	1Q 24	2Q 24	3Q 24	4Q 24
…**to** ARCHER2 (TB)	534	163	68	220	44	67	42	65	99	108	93	98	228
…**from** ARCHER2 (TB)	236	582	667	822	834	1231	1022	1472	1771	2056	2443	3956	2227

* Partial



The amount of data being moved off this quarter has reduced from the previous quarter but is still comparable to Q2 of 2024.

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 2048^3: 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

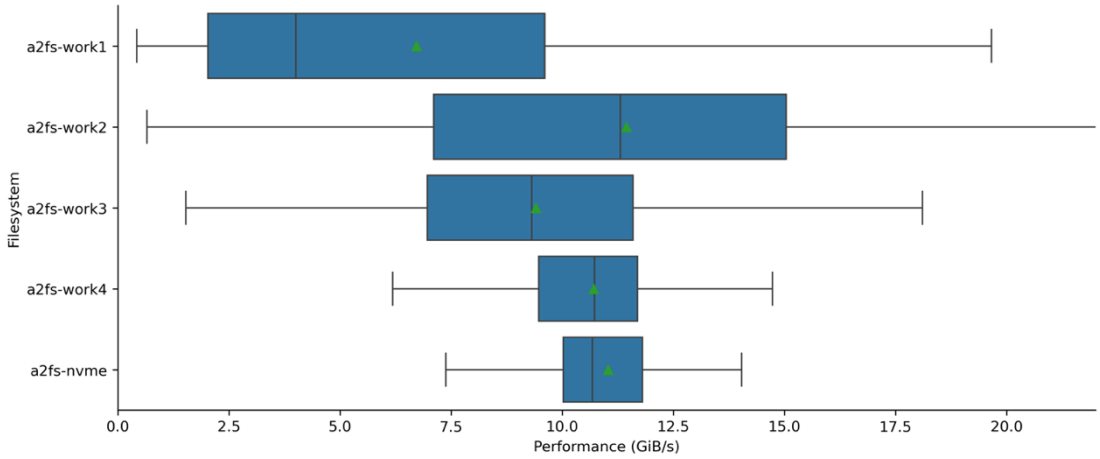
These settings have been found to maximise the IO performance for parallel writes using MPI-IO on the ARCHER2 file systems. Writes using the default settings on ARCHER2 typically have median write values 2-3 GiB/s lower than the optimised values.

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 and provide boxplots illustrating the performance variation. (On the boxplots, the green triangles mark the mean value and the whiskers extend to the last datapoint within the range 1.5 x IQR.)

There is a clear performance drop for the a2fs-work1 file system compared to previous quarters. We have been monitoring and tracking this during ARCHER2 operational meetings. It is thought that this performance issue is due to a combination of the file system use being close to capacity and the use pattern of the file system by researchers. We are taking actions to try and improve this situation and reduce the impact of poor a2fs-work1 performance on the wider service:

- Delete data from old projects from the file system more urgently than would usually be the case
- Move central software installations from a2fs-work1 to a2fs-work4
- Move software modulefiles from a2fs-work1 to a2fs-work4

Benchio MPI-IO medium (GiB/s)	1Q23	2Q23		3Q23	4Q23	1Q24	2Q24	3Q24	4Q24
a2fs-work1	8.2	7.6±0.5		10.5 (8.8:11.8)	10.9 (8.3:12.5)		10.1 (7.0:11.8)	9.7 (6.7:11.9)	4.0 (2.0:9.6)
a2fs-work2	8.5	7.3±0.6		10.4 (7.2:12.4)	10.4 (7.7:13.0)		11.1 (8.0:12.5)	11.1 (8.1:13.1)	11.3 (7.1:15.0)
a2fs-work3	8.3	9.6±0.7		10.0 (8.2:11.6)	10.7 (8.1:11.9)		9.6 (8.4:11.8)	9.6 (7.5:11.8)	9.3 (7.0:11.6)
a2fs-work4					9.7 (9.1:10.2)		10.0 (9.2:10.8)	10.6 (9.4:11.6)	10.7 (9.5:11.7)
a2fs-nvme				10.1 (9.6:11.5)	10.1 (9.5:12.4)		11.1 (10.5:12.4)	11.6 (11.1:12.7)	10.7 (10.0:11.8)



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	01/2023 – 12/2023	1Q 24	2Q 24	3Q 24	4Q 24
Relative Research Output per kWh	100	109	115	115	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	4Q23	1Q24	2Q24	3Q24
Energy per CU (kWh)	0.719	0.713	0.728	0.715	0.650	0.545	0.669	0.590	0.568	0.582	0.585	0.595

Energy Cost per CU Delivered

	4Q21*	1Q22	2Q22	3Q22	4Q22	1Q23	2Q23	3Q23	4Q23	1Q24	2Q24	3Q24
Cost per CU (£)	£0.089	£0.090	£0.098	£0.096	£0.088	£0.074	£0.162	£0.143	£0.136	£0.140	£0.160	£0.164

* 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. There was also an additional increase in the unit cost of electricity from April 2024.