ARCHER SP Service Quarterly Report

Quarter 1 2015

Document Information and Version History

Version:	1.0
Status	Draft
Author(s):	Alan Simpson, Jo Beech-Brandt, Andy Turner, Mike Brown, Stephen Booth
Reviewer(s)	Alan Simpson

Version	Date	Comments, Changes, Status	Authors, contributors, reviewers
0.1	2015-03-16	Initial Draft	Alan Simpson, Jo Beech-Brandt, Stephen Booth, Andy Turner
0.2	2015-04-10	Updates and additions	Jo Beech-Brandt, Andy Turner
0.3	2015-04-16	Additional content	Andy Turner, Jo Beech-Brandt, Mike Brown
0.4	2015-04-17	Review	Alan Simpson
1.0	2015-04-17	Final version for EPSRC	Alan Simpson

1. The Service

1.1 Service Highlights

This is the report for the ARCHER SP Service for the Reporting Periods:

January 2015, February 2015, March 2015

- Utilisation on the system during 15Q1 was 85%, very similar to 14Q4. However, as both EPSRC Consortia and NERC had allocation periods that ended on 31 March 2015, the queues became significantly longer during March.

- The implementation of period allocations for consortia and large RGs went smoothly, with only minor niggles.

- All service levels were met during the period.

- UK-Federation authorisation to the SAFE was enabled. This gives users the option of accessing the SAFE using the same authorisation credentials as they use in their home institution.

- Implemented a monitoring script to highlight if a user is under-utilising requested nodes i.e. if a user requests 120 nodes but only actually uses 24 nodes there are then 96 nodes idle. The user is now alerted to this to ensure it is not user error and assistance is offered if required.

- There has been a significant amount of activity on security matters with additional monitoring, recording and hardening of both external-facing and internal security processes.

- ARCHER Annual User Survey results:
  - A short paper analysing the results has been provided and published on the ARCHER website
  - Comments and suggestions have been reviewed and a prioritised list of possible improvements presented to the ARCHER Operations Meeting for discussion and possible implementation.

- Long term Edinburgh DataShare collection created for ARCHER reports, white papers and eCSE reports:
  - http://datashare.is.ed.ac.uk/handle/10283/689
  - All items submitted to the service will be available beyond the end of the ARCHER service.
  - All items submitted gain a persistent DoI allowing them to be referenced and accessed beyond the lifetime of the ARCHER service.

- Although not directly part of the ARCHER SP contract, there has nevertheless neen a significant uplift to the RDF data store in terms of capacity.

1.2 Forward Look

- Further preparation for the Cray Linux Environment (CLE) upgrade from CLE 5.1 to CLE 5.2.
    - Advice will be provided to users on actions required following upgrade (e.g. recompiling applications, verifying output).

- RSIP node implementation will allow the use of applications requiring licence servers on the compute nodes, for example, compilers and ISV applications.

- SP staff will collaborate with CSE/Outreach on the workshop for ARCHER Champions planned for May 2015.

- We will continue to progress the collection of data describing ARCHER users' career stage
    - Initial summary of data on ARCHER user demographics will be provided to Research Councils

- Finalisation of the end-of-lifecycle ARCHER policy for ending projects and former users.

- Provide information on historical average queuing times for different job classes to users:
    - on the website Service Status page;
    - through a command line utility on ARCHER itself;
    - as part of the output from the checkScript tool that provides feedback on validity of PBS job submission scripts.

- Detailed analysis of queue load and waiting times during the period of exceptionally high demand at the end of initial EPSRC allocation period to:
    - understand how well the scheduler and scheduling policy dealt with the high load;
    - look at correlations between load levels and helpdesk queries;
    - and report on any differences in job size/length balance between "normal" operation and high load.

- There is ongoing work on SAFE through the SAFE Development project that will improve the service for users. Upcoming targets include:
    - Providing greater support for Technical Assessments of new project applications within the SAFE
    - Updates and enhancements to the reports generated by the SAFE
    - Final testing of improved automation for courses
    - New forms for requesting access to packages with access restrictions

- There will be a formal security audit run during 15Q2.

- There will also be improvement in external networking with the configuration of the 40 GN capacity and also the direct connection of the ACF to JANET. This activity is intended to provide enhanced resiliency as well as eventual performance.

2. Contractual Performance Report

This is the contractual performance report for the ARCHER SP Service.

2.1 Service Points and Service Credits

The Service Levels and Service Points for the SP service are defined as below in Schedule 2.2.

- **2.6.2 - Phone Response (PR):** 90% of incoming telephone calls answered personally within 2 minutes for any Service Period. *Service Threshold: 85.0%; Operating Service Level: 90.0%.*
- **2.6.3 - Query Closure (QC):** 97% of all administrative queries, problem reports and non in-depth queries shall be successfully resolved within 2 working days. *Service Threshold: 94.0%; Operating Service Level: 97.0%.*
- **2.6.4 - New User Registration (UR):** Process New User Registrations within 1 working day.

Definitions:

Operating Service Level: *The minimum level of performance for a Service Level which is required by the Authority if the Contractor is to avoid the need to account to the Authority for Service Credits.*

Service Threshold: *This term is not defined in the contract. Our interpretation is that it refers to the minimum allowed service level. Below this threshold, the Contractor is in breach of contract.*

Non In-Depth: *This term is not defined in the contract. Our interpretation is that it refers to Basic queries that are handled by the SP Service. This includes all Admin queries (e.g. requests for Disk Quota, Adjustments to Allocations, Creation of Projects) and Technical Queries (Batch script questions, high level technical 'How do I?' requests). Queries requiring detailed technical and/or scientific analysis (debugging, software package installations, code porting) are referred to the CSE Team as In-Depth queries.*

Change Request*: This term is not defined in the contract. There are times when SP receives requests that may require changes to be deployed on ARCHER. These requests may come from the users, the CSE team or Cray. Examples may include the deployment of new OS patches, the deployment Cray bug fixes, or the addition of new systems software. Such changes are subject to Change Control and may have to wait for a Maintenance Session. The nature of such requests means that they cannot be completed in 2 working days.*

2.1.1 Service Points

In the previous Service Quarter the Service Points can be summarised as follows:

Period	Jan 15		Feb 15		Mar 15		15Q1
Metric	**Service Level**	**Service Points**	**Service Level**	**Service Points**	**Service Level**	**Service Points**	**Service Points**
2.6.2 – PR	100.0%	-5	100.0%	-5	100.0%	-5	-15
2.6.3 – QC	99.2%	-2	99.1%	-2	99.4%	-2	-6
2.6.4 – UR	1 WD	0	1 WD	0	1 WD	0	0
Total		**-7**		**-7**		**-7**	**-21**

The details of the above can be found in Section 2.2 of this report.

2.1.2 Service Failures

There were no Service Failures in this Service Quarter. However, there was an unplanned maintenance session due to a security concern during which the service was closed to user logins although batch system jobs continued to run. See Section 2.3.2 of this report.

2.1.3 Service Credits

The total Service Credit applicable for each Service Quarter is calculated in the following way:

[redacted]

Where:
"Applicable Charge" = the relevant Annual Maintenance Charge divided by four (4) (to form the Maintenance Charge relevant for the Service Periods being assessed)
"SC" = Service Credit
"TSP" = Total Service Points for the Service Quarter

As the Total Service Points are negative (-21), no Service Credits apply in 15Q1.

2.2 Detailed Service Level Breakdown

2.2.1 Phone Response (PR)

	Jan 15	Feb 15	Mar 15	15Q1
Phone Calls Received	35 (6)	54 (13)	46 (22)	**135 (41)**
Answered 2 Minutes	35	54	46	**135**
Service Level	**100.0%**	**100.0%**	**100.0%**	**100.0%**

The volume of telephone calls remained low in 15Q1. Of the 135 calls received in total above, only 41 were genuine ARCHER user calls that resulted in queries or answered user questions directly.

2.2.2 Query Closure (QC)

	Jan 15	Feb 15	Mar 15	15Q1
Self-Service Admin	444	485	793	**1722**
Admin	179	245	230	**654**
Technical	36	43	39	**118**
Total Queries	*659*	*773*	*1062*	***2494***
Total Closed in 2 Days	*654*	*766*	*1056*	***2476***
Service Level	**99.2%**	**99.1%**	**99.4%**	**99.3%**

In addition to the Admin and Technical queries, the following Change Requests were resolved in 15Q1.

	Jan 15	Feb 15	Mar 15	15Q1
Change Requests	3	1	4	**8**

2.2.3 User Registration (UR)

	Jan 15	Feb 15	Mar 15	15Q1
No of Requests	67	119	127	**313**
Closed in One Working Day	66	118	127	**311**
Average Closure Time (Hrs)	1.12	0.83	0.95	**0.95**
Average Closure Time (Working Days)	0.12	0.09	0.10	**0.10**
Service Level	**1 WD**	**1 WD**	**1 WD**	**1 WD**

To avoid double counting, these requests are not included in the above metrics for "Admin and Technical" Query Closure.

2.3 Additional Metrics

2.3.1 Target Response Times

The following metrics are also defined in Schedule 2.2, but have no Service Points associated.

Target Response Times	
1	During core time, an initial response to the user acknowledging receipt of the query
2	A Tracking Identifier within 5 minutes of receiving the query
3	During Core Time, 90% of incoming telephone calls should be answered personally (not by computer) within 2 minutes
4	During UK office hours, all non telephone communications shall be acknowledged within 1 Hour

1 – Initial Response

This is sent automatically when the user raises a query to the address helpdesk@archer.ac.uk. Users may choose not to receive such emails by mailing support@archer.ac.uk.

2 – Tracking Identifier

This is sent automatically when the user raises a query to the address helpdesk@archer.ac.uk. Users may choose not to receive such emails by mailing support@archer.ac.uk. The tracking identifier is set in the SAFE regardless of which option the user selects.

3 – Incoming Calls

These are covered in the previous section of the report. Service Points apply.

4 - Query Acknowledgement

Acknowledgment of the query is defined as when the Helpdesk assigns the new incoming query to the relevant Service Provider. This should happen within 1 working hour of the query arriving at the Helpdesk. The Helpdesk processed the following number of incoming queries during the Service Quarter:

	Jan15	Feb15	Mar15	15Q1
CRAY	10	8	8	**26**
ARCHER_CSE	95	1249	139	**3711**
ARCHER_SP	944	148	1518	**382**
Total Queries Assigned	1049	1405	1665	**4119**
Total Assigned in 1 Hour	1049	1405	1665	**4119**
Service Level	**100%**	**100%**	**100%**	**100%**

2.3.2 Maintenance

SP is allowed to book a maximum of two maintenance occasions in any 28-day period, and these shall last no longer than four hours; these are defined as Permitted Maintenance. Such Maintenance Periods are recorded in the Maintenance Schedule. A 6-month forward plan of maintenance has been agreed with the Authority.

If greater than 4 hours downtime is required for maintenance, 20 days prior approval is required from the Authority. Where possible, SP will perform maintenance on an 'At-risk' basis, thus maximising the Availability of the Service. The following planned maintenance took place in the Service Quarter.

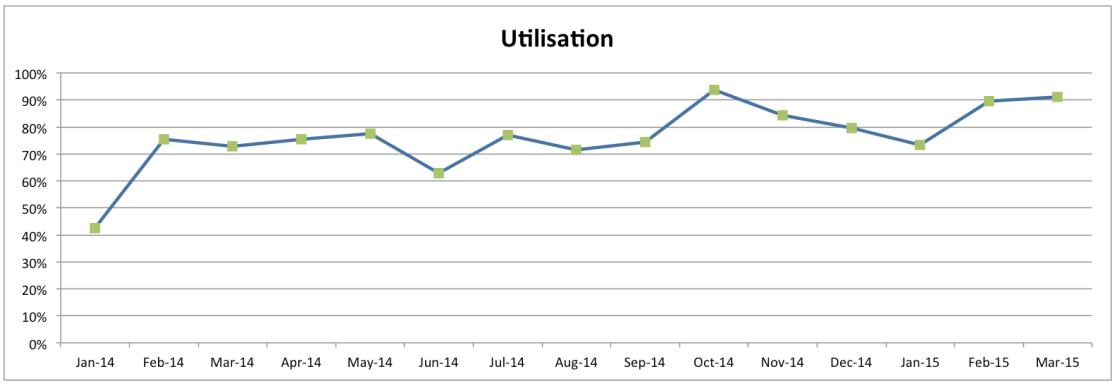
Date	Start	End	Duration	Type	Notes	Reason
23rd Jan 2015	1845	1140 (+1)	16:55	Unplanned	Security Alert	Login access closed for security concerns
11th Feb 2015	0900	1700	8:00	Pre-Approved	EPSRC Approved 0900 - 1700	Planned Maintenance
11th Mar 2015	0900	1633	7:33	Pre-Approved	EPSRC Approved 0900 – 1700	Planned Maintenance

3. Service Statistics

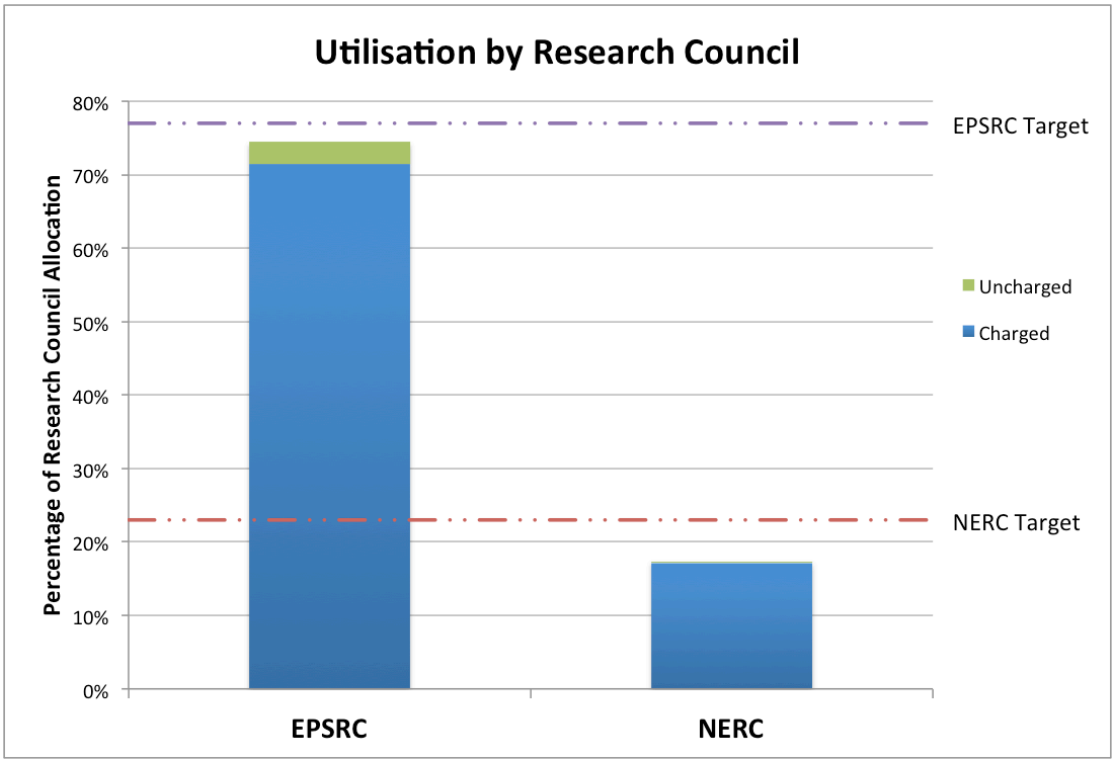
This section contains statistics on the ARCHER service as requested by EPSRC, SAC and SMB.

3.1 Utilisation

Utilisation over the quarter was 85%.

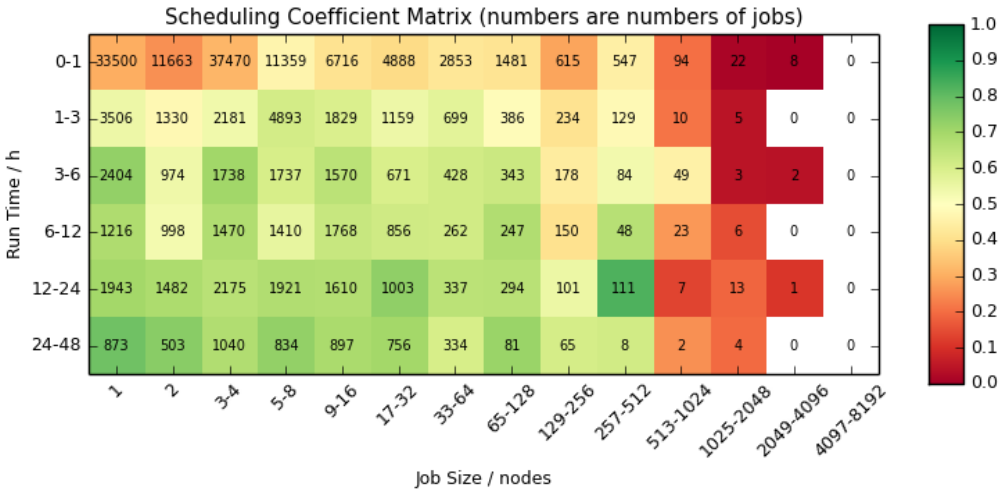


The utilisation by the Research Councils, relative to their respective allocations, is presented below.



This bar chart shows the usage of ARCHER by the two Research Councils presented as a percentage of the total Research Council allocation on ARCHER.

3.2 Scheduling Coefficient Matrix



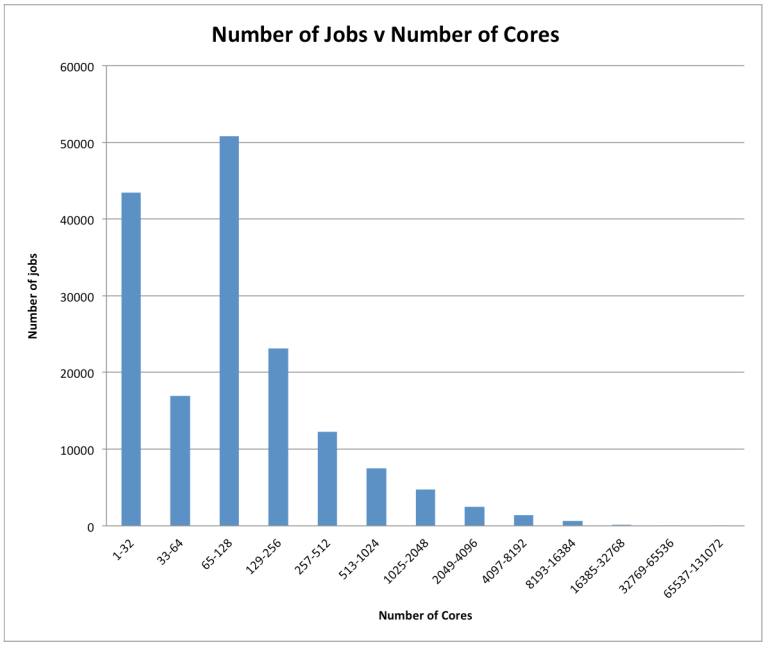
The colour in the matrix indicates the value of the Scheduling Coefficient. This is defined as the ratio of runtime to runtime plus wait time. Hence, a value of 1 (green) indicates that a job ran with no time waiting in the queue, a value of 0.5 (pale yellow) indicates a job queued for the same amount of time that it ran, and anything below 0.5 (orange to red) indicates that a job queued for longer than it ran. We would expect lower Scheduling Coefficients (i.e., longer wait times than runtimes) both for very short jobs (as there is always a scheduling overhead) and for very large jobs (where the system has to drain compute nodes to make space for the jobs).

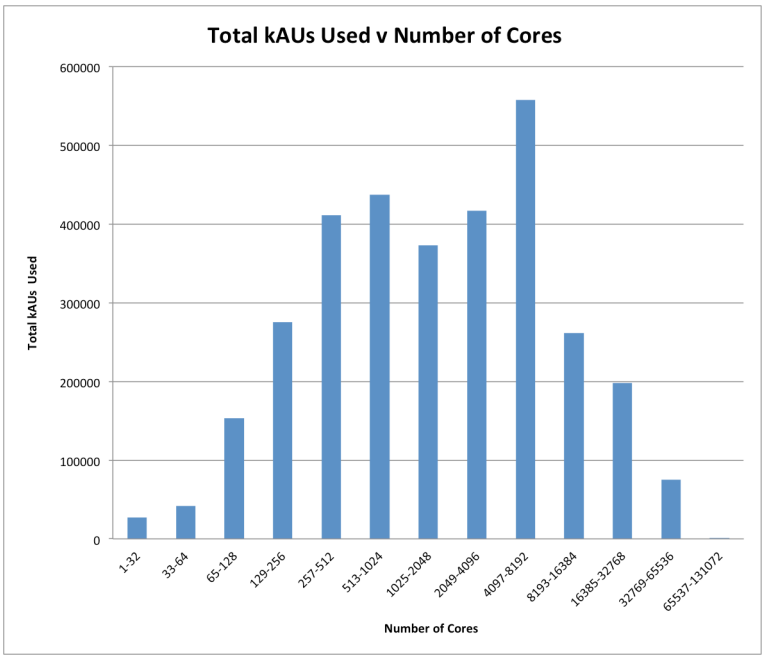
The matrix shows that, on average during this quarter, queuing times appear generally reasonable. Although it is not obvious from the averaged heatmap matrix above, ARCHER was very busy towards the end of the period allocations on 31 March. The queues got significantly longer and so Scheduling Coefficients decreased. We are working on a more detailed analysis of this particularly busy period.

3.3 Additional Usage Graphs

The following charts provide different views of the distribution of job sizes on ARCHER.

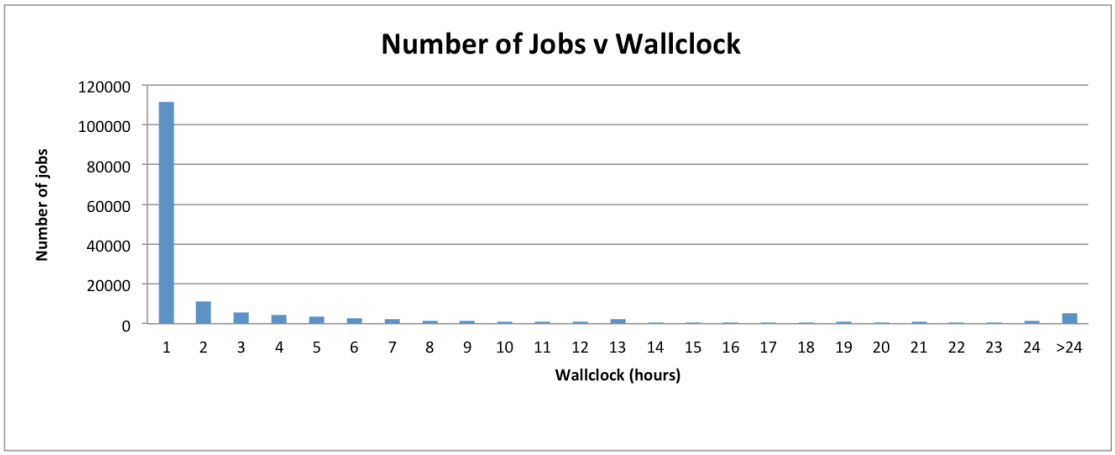
Number of Cores

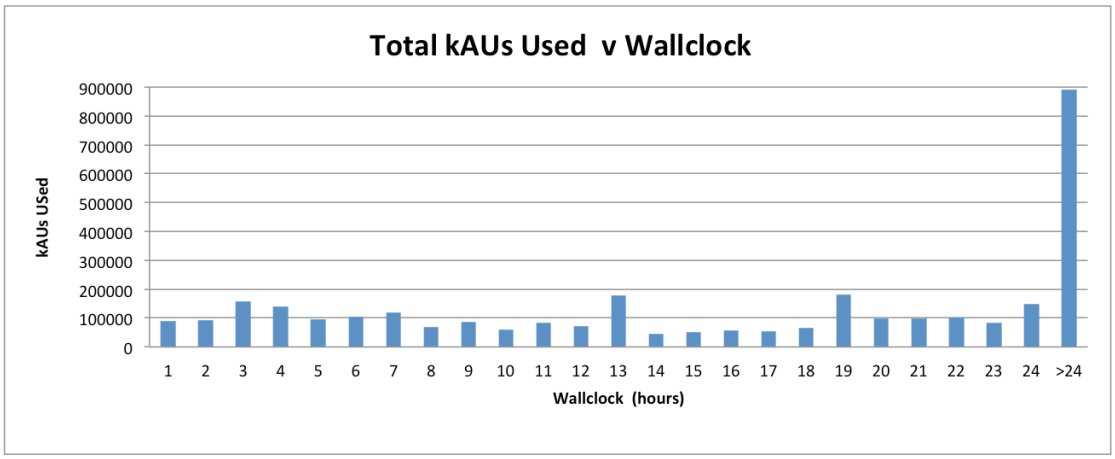




The first graph shows that, in terms of numbers, there is a significant number of jobs using no more than 512 cores. However, the second graph shows that most of the kAUs were spent on jobs between 257 cores and 8192 cores. The number of kAUs used is closely related to money and shows better how the investment in the system is utilised.

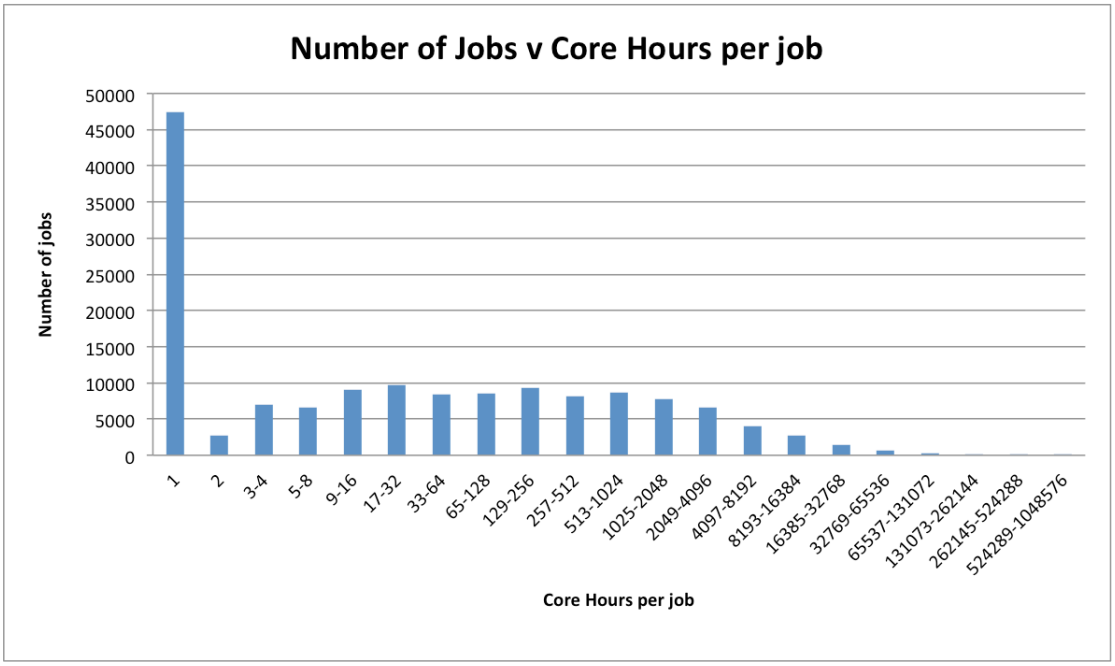
Wallclock

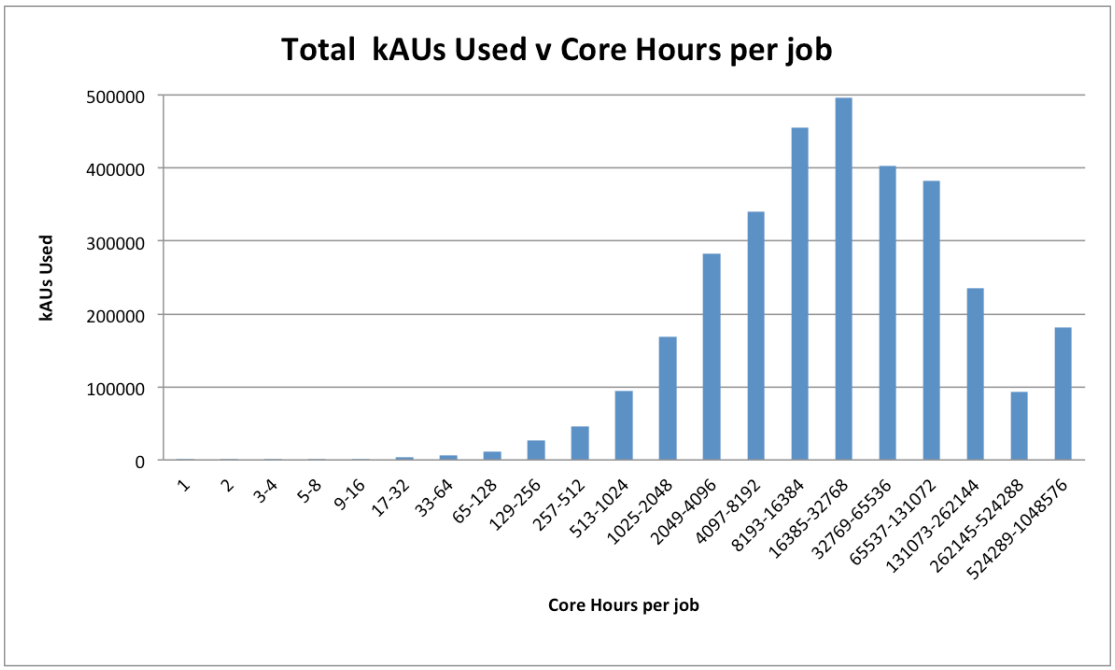




From the first graph, it would appear that the system is dominated by short jobs. However, the second graph shows that the system utilisation is actually dominated by long jobs.

Core Hours





The above graphs show that, while there are quite a few jobs that use only a small number of core hours per job, most of the resource is consumed by jobs that use tens of thousands of core hours per job.