CSAR Service

Consolidated Management Report

3rd Quarter 2003

Management Summary

This is the consolidated Management report for the third quarter 2003 of the CSAR HPC facility for UK Academia and Industry, which enables World-Class research and development.

The number of users has grown to a total of 565 to date.

The workload on both the Origin 3000 (Green) and the Cray T3E (Turing) has been fairly evenly spread across the mid- to high-end ranges of PEs during the third quarter of this year, with the T3E running at approximately three-quarters capacity.

CSAR has been granted an 18 month extension of service contract until June 30th 2006. With this extension CfS is implementing a further technology refresh introducing a 256 processor Itanium-2 (Madison) based SGI Altix due to go into production service on 1st October.

CfS remains active in the UK Grid Forum.

Introduction

This Management Report includes a section for each of the main service functions:

1. Service Quality
2. HPC Services
3. Science Applications Support Services
4. Training & Education Services
5. User Registration & New User Services
6. Value-Added Services

Each section includes a status report for the period, including notable achievements and problems, also noteworthy items for the next period.

1 Service Quality

This section covers overall Customer Performance Assessment Ratings (CPARS), HPC System availability and usage, Service Quality Tokens and other information concerning issues, progress and plans for the CSAR Service.

1.1 CPARS

Table 1 gives the measure by which the quality of the CSAR Service is judged. It identifies the metrics and performance targets, with colour coding so that different levels of achievement against targets can be readily identified. Unsatisfactory actual performance will trigger corrective action.

CSAR Service - Service Quality Report - Performance Targets

Service Quality Measure	Performance Targets					
	White	Blue	Green	Yellow	Orange	Red
HPC Services Availability						
Availability in Core Time (% of time)	> 99.9%	> 99.5%	> 99.2%	> 98.5%	> 95%	95% or less
Availability out of Core Time (% of time)	> 99.8%	> 99.5%	> 99.2%	> 98.5%	> 95%	95% or less
Number of Failures in month	0	1	2 to 3	4	5	> 5
Mean Time between failures in 52 week rolling period (hours)	>750	>500	>300	>200	>150	otherwise
Fujitsu Service Availability						
Availability in Core Time (% of time)	> 99.9%	> 99.5%	> 99.2%	> 98.5%	> 95%	95% or less
Availability out of Core Time (% of time)	> 99.8%	> 99.5%	> 99.2%	> 98.5%	> 95%	95% or less
Help Desk						
Non In-depth Queries - Max Time to resolve 50% of all queries	< 1/4	< 1/2	< 1	< 2	< 4	4 or more
Non In-depth Queries - Max Time to resolve 95% of all queries	< 1/2	< 1	< 2	< 3	< 5	5 or more
Administrative Queries - Max Time to resolve 95% of all queries	< 1/2	< 1	< 2	< 3	< 5	5 or more
Help Desk Telephone - % of calls answered within 2 minutes	>98%	> 95%	> 90%	> 85%	> 80%	80% or less
Others						
Normal Media Exchange Requests - average response time	< 1/2	< 1	< 2	< 3	< 5	5 or more
New User Registration Time (working days)	< 1/2	< 1	< 2	< 3	< 4	otherwise
Management Report Delivery Times (working days)	< 1	< 5	< 10	< 12	< 15	otherwise
System Maintenance - no. of sessions taken per system in the month	0	1	2	3	4	otherwise

Table 1

Table 2 gives actual performance information for the period. Overall, the CPARS Performance Achievement for the 3rd quarter 2003 was satisfactory (see Table 3), i.e. Green measured against the CPARS performance targets.

CSAR Service - Service Quality Report - Actual Performance Achievement

Service Quality Measure	2002/3											
	Oct	Nov	Dec	Jan	Feb	March	April	May	June	July	Aug	Sept
HPC Services Availability												
Availability in Core Time (% of time)	99.77%	99.25%	99.21%	99.46%	99.73%	100%	99.74%	97.66%	99.25%	98.83%	98.95%	96.62%
Availability out of Core Time (% of time)	99.52%	99.57%	100%	99.89%	100.00%	99.81%	99.81%	99.33%	99.9%	99.57%	100%	98.48%
Number of Failures in month	1	1	0	3	1	1	1	4	1	2	2	4
Mean Time between failures in 52 week rolling period (hours)	398	417	515	487	487	515	548	461	548	487	461	417
Help Desk												
Non In-depth Queries - Max Time to resolve 50% of all queries	<0.25	<0.25	<0.25	<0.25	<0.25	<0.25	<0.25	<0.25	<0.25	<0.25	<0.25	<0.25
Non In-depth Queries - Max Time to resolve 95% of all queries	<2	<2	<2	<0.5	<1	<2	<3	<1	<2	<1	<0.5	<5
Administrative Queries - Max Time to resolve 95% of all queries	<0.5	<0.5	<0.5	<1	<0.5	<1	<0.5	<0.5	<0.5	<0.5	<1	<1
Help Desk Telephone - % of calls answered within 2 minutes	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
Others												
Normal Media Exchange Requests - average response time	<0.5	<0.5	<0.5	<0.5	<0.5	<0.5	<0.5	<0.5	<0.5	<0.5	<0.5	<0.5
New User Registration Time (working days)	0	0	0	0	0	0	0	0	0	0	0	0
Management Report Delivery Times (working days)	10	10	10	10	10	10	10	10	10	10	10	10
System Maintenance - no. of sessions taken per system in the mont	2	2	2	2	2	2	2	2	2	2	2	2

Table 2

Notes:

1. HPC Services Availability has been calculated using the following formulae, based on the relative NPB performance of Turing, Fermat and Green at installation:

   Turing availability x 143/(143+40+233)] + [Fermat availability x 40/ (143+40+233) + Green availability x 233/(143+40+233)]

2. Mean Time Between Failures for Service Credits is formally calculated from Go-Live Date.

Table 3 gives Service Credit values for each month to date. These are accounted on a quarterly basis, formally from the Go-Live Date. The values are calculated according to agreed Service Credit Ratings and Weightings.

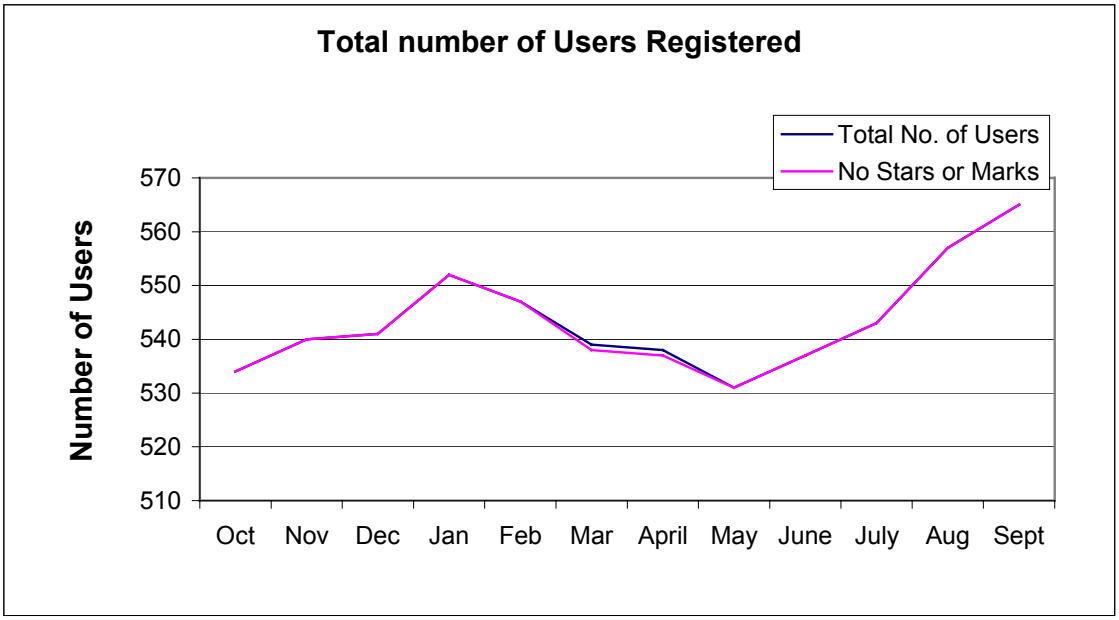
CSAR Service - Service Quality Report - Service Credits

Service Quality Measure	2002/3											
	Oct	Nov	Dec	Jan	Feb	March	April	May	June	July	Aug	Sept
HPC Services Availability												
Availability in Core Time (% of time)	-0.039	0	0	0	-0.039	-0.058	-0.039	0.078	0	0.039	0.039	0.078
Availability out of Core Time (% of time)	-0.039	-0.039	-0.047	-0.047	-0.047	-0.047	-0.047	0	-0.047	-0.039	-0.047	0.039
Number of Failures in month	-0.008	-0.008	-0.009	0	-0.008	-0.008	-0.008	0.008	-0.008	0	0	0.008
Mean Time between failures in 52 week rolling period (hours)	0	0	-0.008	0	0	-0.008	-0.008	0	-0.008	0	0	0
Help Desk												
Non In-depth Queries - Max Time to resolve 50% of all queries	-0.019	-0.019	-0.019	-0.019	-0.019	-0.019	-0.019	-0.019	-0.019	-0.019	-0.019	-0.019
Non In-depth Queries - Max Time to resolve 95% of all queries	0	0	0	-0.019	-0.016	0	0.016	-0.016	0	-0.016	-0.019	0.0312
Administrative Queries - Max Time to resolve 95% of all queries	-0.019	-0.019	-0.019	-0.016	-0.019	-0.016	0	-0.019	-0.019	-0.019	-0.016	-0.01551
Help Desk Telephone - % of calls answered within 2 minutes	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004
Others												
Normal Media Exchange Requests - average response time	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
New User Registration Time (working days)	-0.019	-0.019	-0.019	-0.019	-0.019	-0.019	-0.019	-0.019	-0.019	-0.019	-0.019	-0.019
Management Report Delivery Times (working days)	0	0	0	0	0	0	0	0	0	0	0	0
System Maintenance - no. of sessions taken per system in the mont	0	0	0	0	0	0	0	0	0	0	0	0
Monthly Total & overall Service Quality Rating for each period:	-0.07	-0.05	-0.06	-0.06	-0.09	-0.09	-0.07	0.00	-0.06	-0.04	-0.04	0.05
Quarterly Service Credits:			-0.19			-0.24			-0.13			-0.03

Table 3

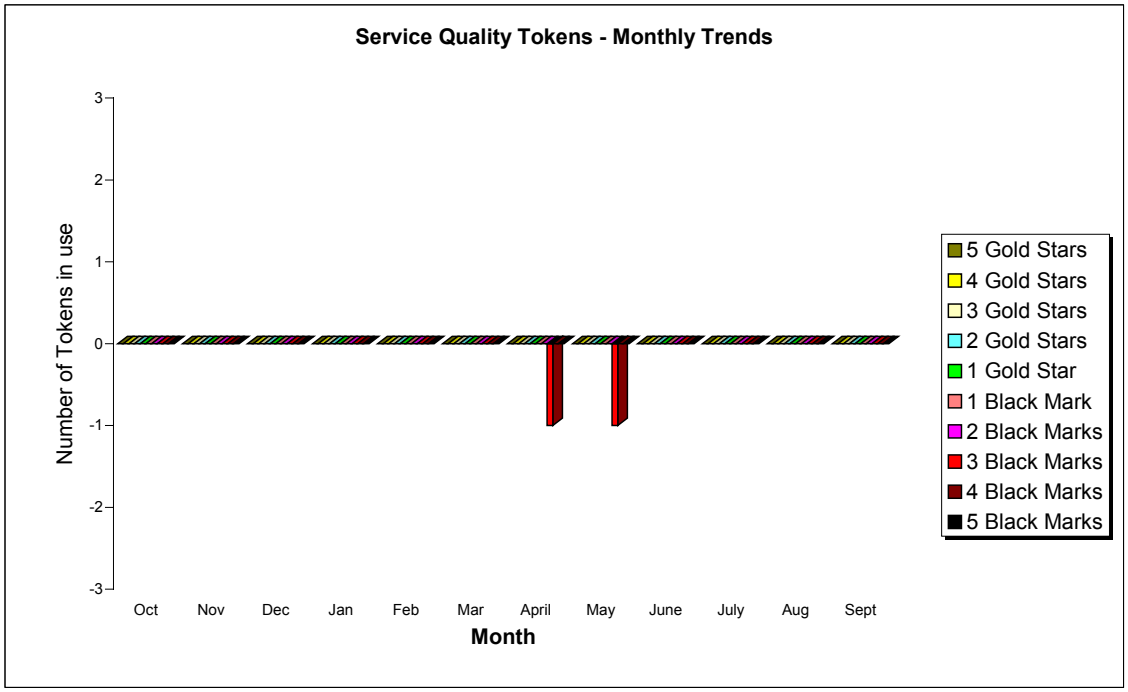
1.2 No. of Registered Users

The current position at the end of the quarter is that there are 565 registered users of the CSAR Service.



1.3 Service Quality Tokens

The graph below illustrates the monthly usage trend of Service Quality Tokens:



Over the course of the quarter the position is that as a management tool the Service Quality Tokens have been available to enable the users to provide qualitative feedback about all aspects of the service. This feedback is used as a mechanism to initiate change in the service where appropriate.

A the end of the quarter there were no black marks or gold stars allocated to the service.

2 HPC Services Usage

Usage information is given in tabular form, and in graphical format. The system usage information covers:

- CPU usage
- User Disk allocation
- HSM/tape usage

This is illustrated in a number of graphs including;

a) MPP (T3E) Usage by month, showing usage each month of CPU (T3E PE Elapsed Hours), split by Research Council and giving the equivalent GFLOP-Years as per NPB. The Baseline Capacity is shown by an overlaid horizontal line.

b) SMP (Origin) Usage by month, showing usage each month in CPU Hours, split by Research Council and giving the equivalent GFLOP-Years as per NPB. The Baseline Capacity is shown by an overlaid horizontal line.

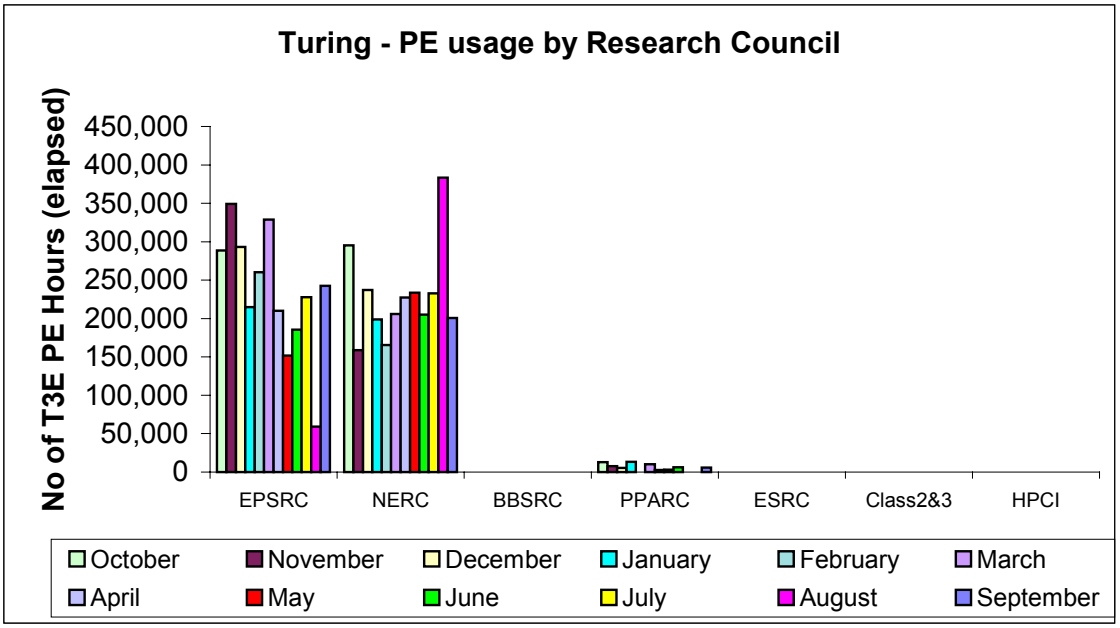
c) High Performance Disk (T3E) allocated for User Data by month, showing the allocated space each month in GBytes, split by Research Council. The Baseline Capacity (1 Terabyte) is shown by an overlaid horizontal line.

d) Medium Performance Disk (Origin) allocated for User Data by month, showing the allocated space each month in GBytes, split by Research Council. The Baseline Capacity (1.5 Terabytes) is shown by an overlaid horizontal line.

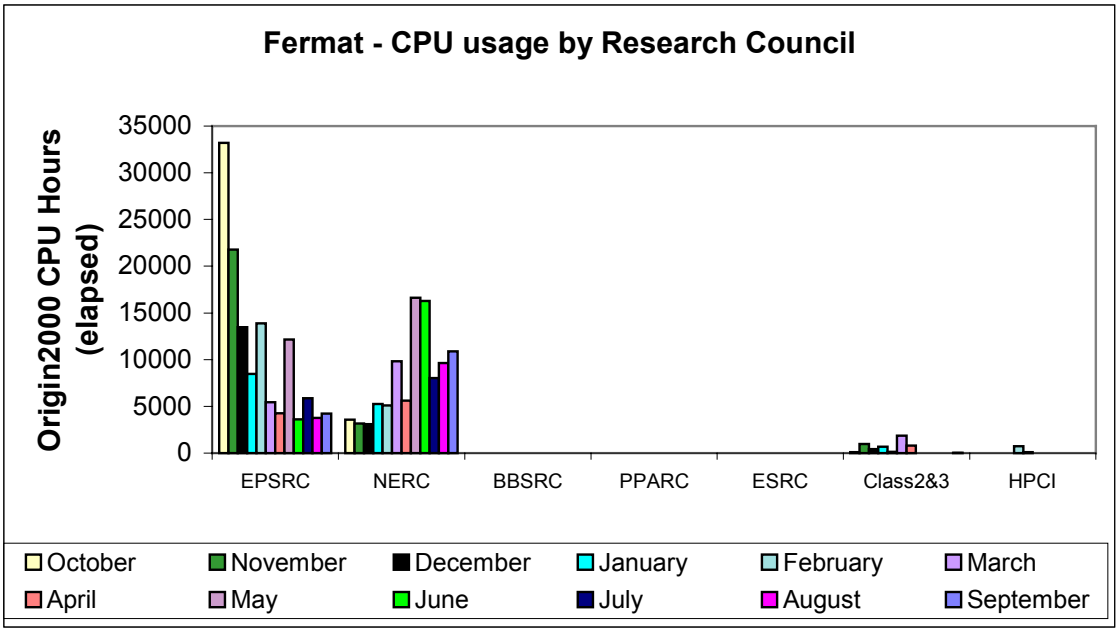
e) HSM/Tape Usage (T3E) by month, showing the volumes held each in GBytes, split by Research Council. The Baseline Capacity (16 Terabytes) available will be shown by an overlaid horizontal line.

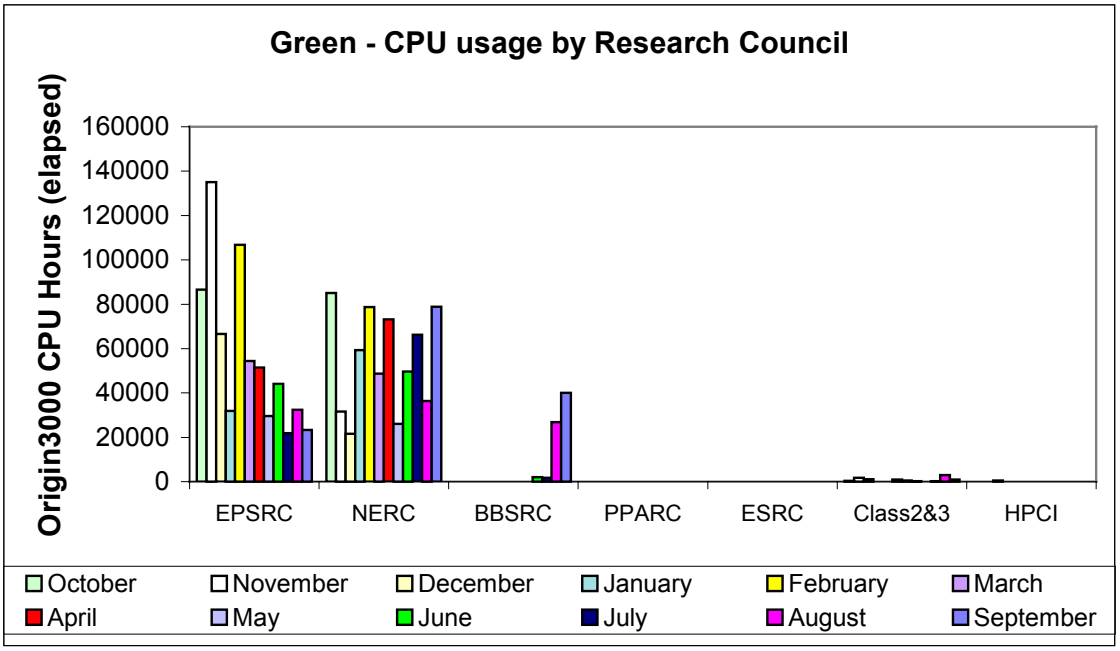
2.1 Service Usage Charts

The graphs below show recent monthly PE, CPU, disk and HSM allocations and usage.

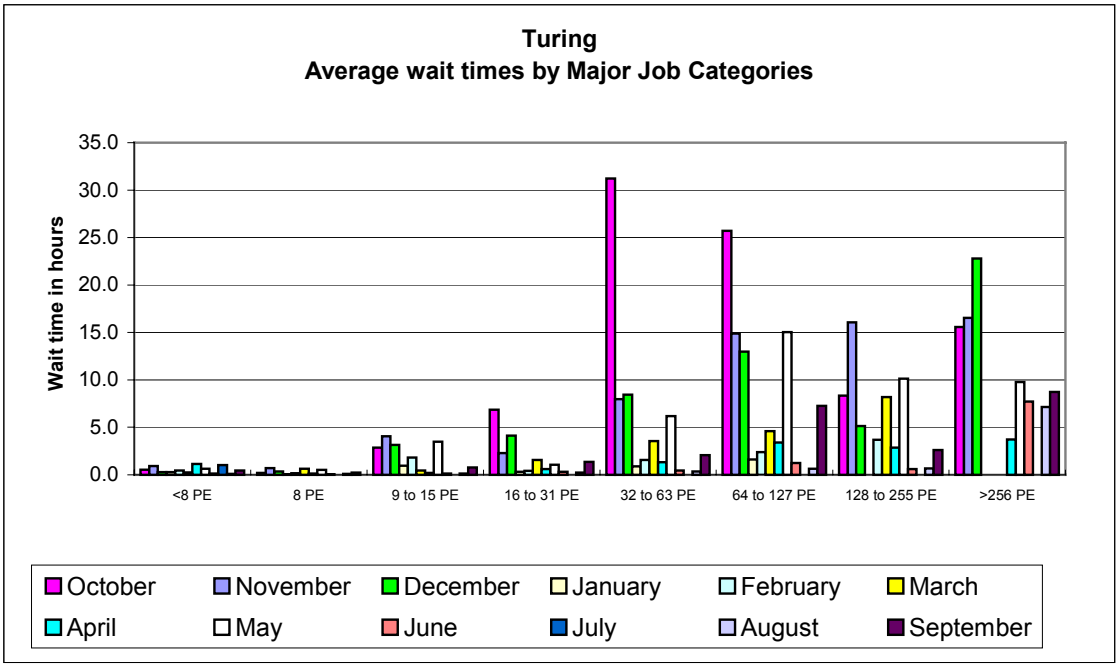


Turing PE usage is shown by Research Council during the last 12 months of service in the above chart.

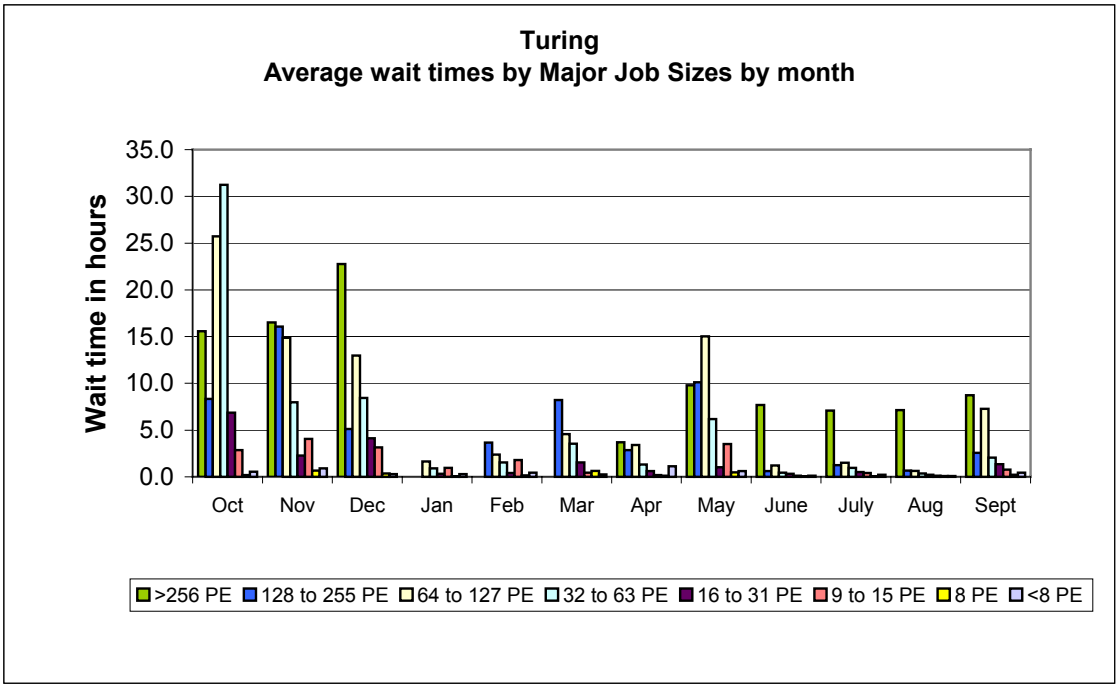




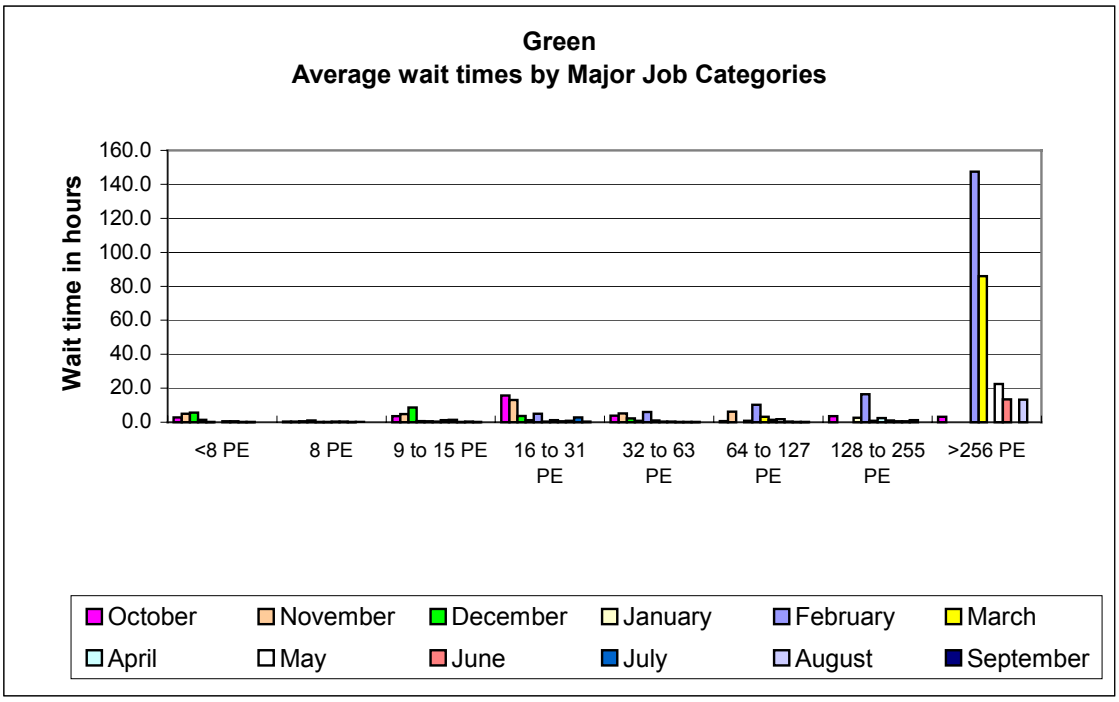
Usage of the two SGI Origin systems, Fermat and Green, is shown by Research Council during the last 12 months of service in the above two charts.

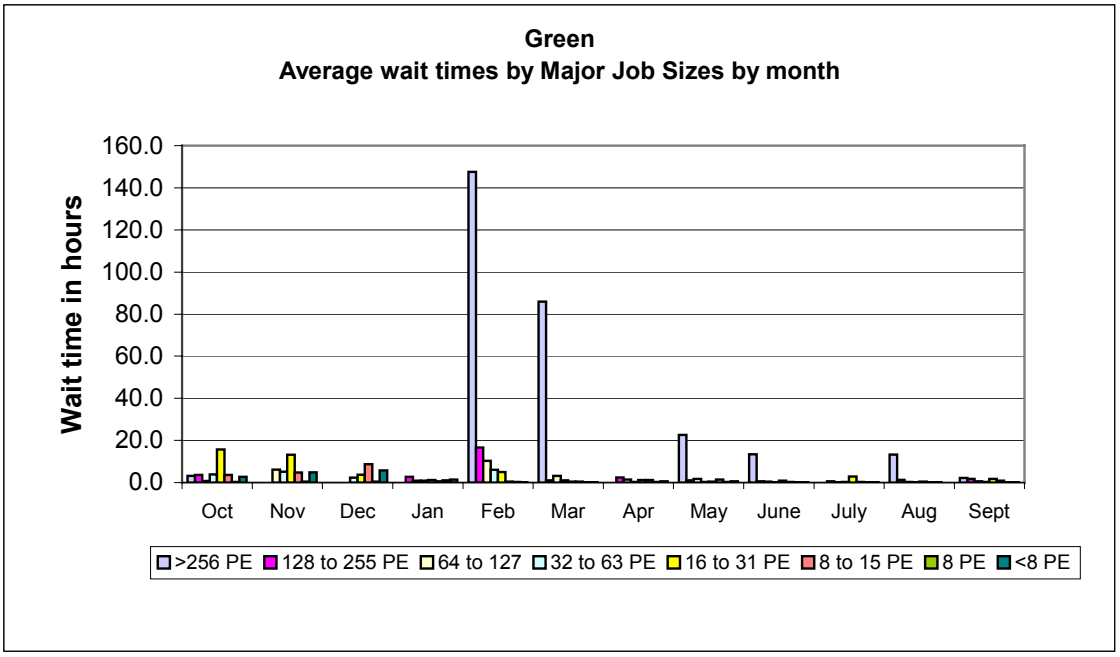


The above chart, and the one below, show the wait time trend in hours on the Turing system.

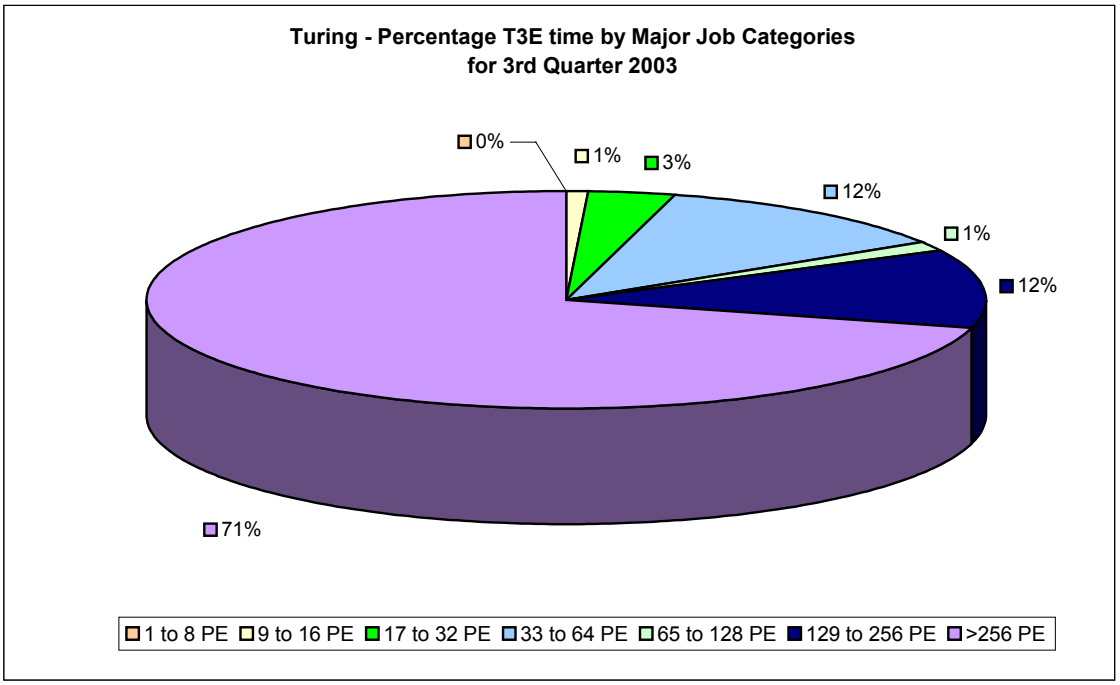


The following two charts show average wait times in hours for the quarter on the Origin 3000 (Green).

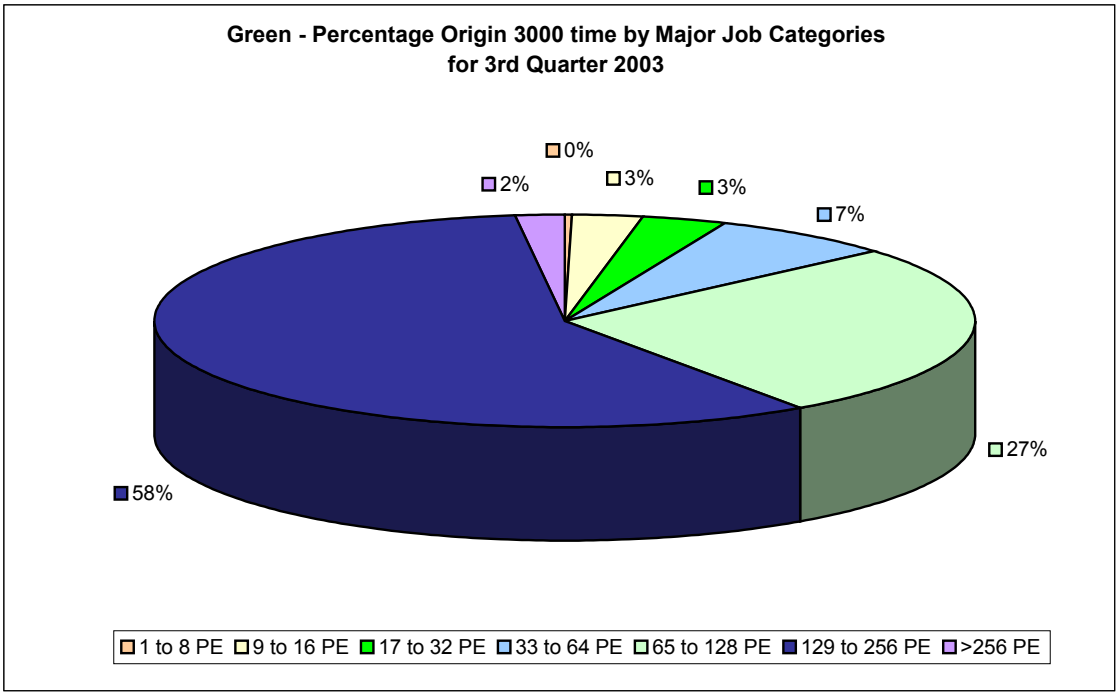




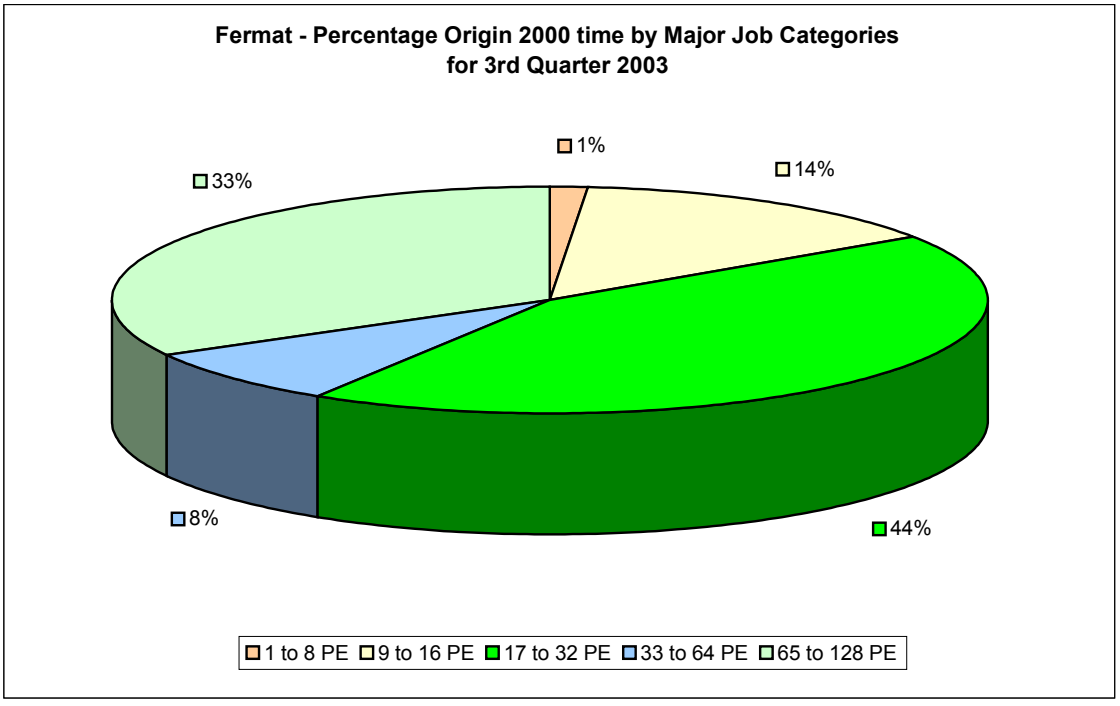
The next three charts show the percentage PE time utilisation by the major job categories on the Turing, Green and Fermat systems for the 3rd quarter 2003.



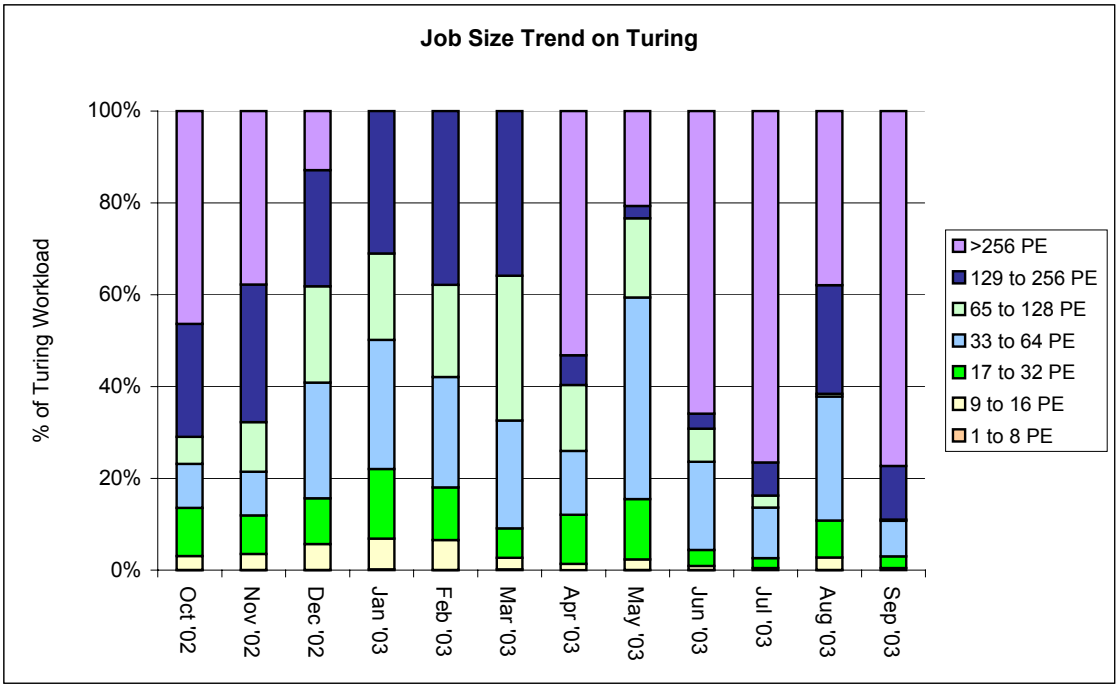
The workload on Turing for the third quarter was predominantly in the greater than 256 PE range.



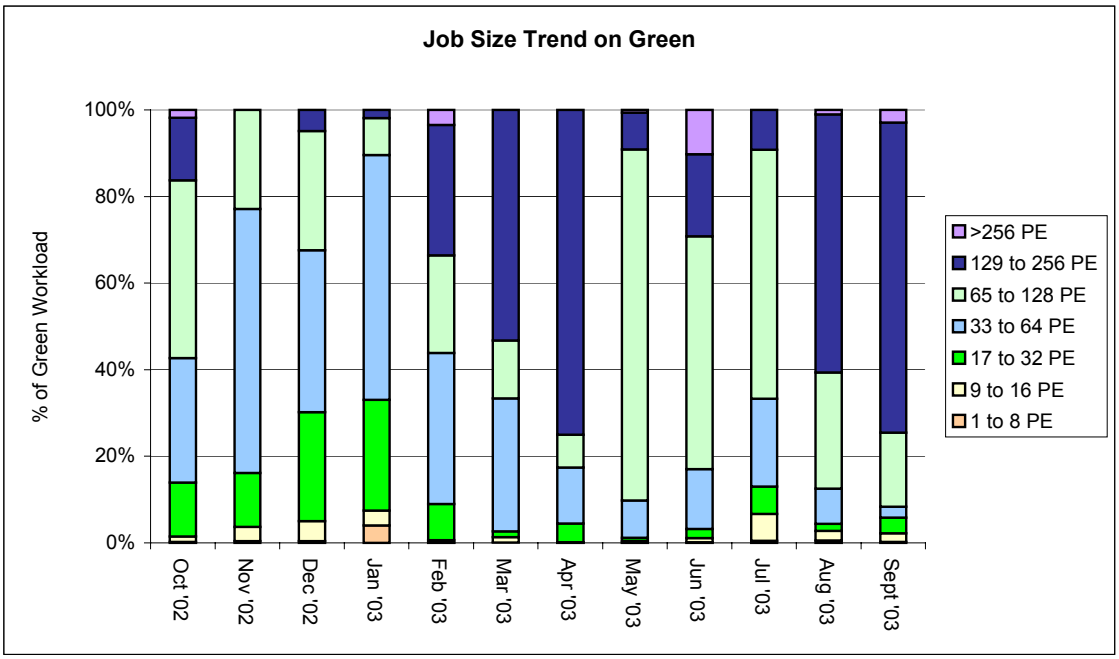
On Green, the 129 to 256 PE range has seen the greatest percentage of workload during this quarter.



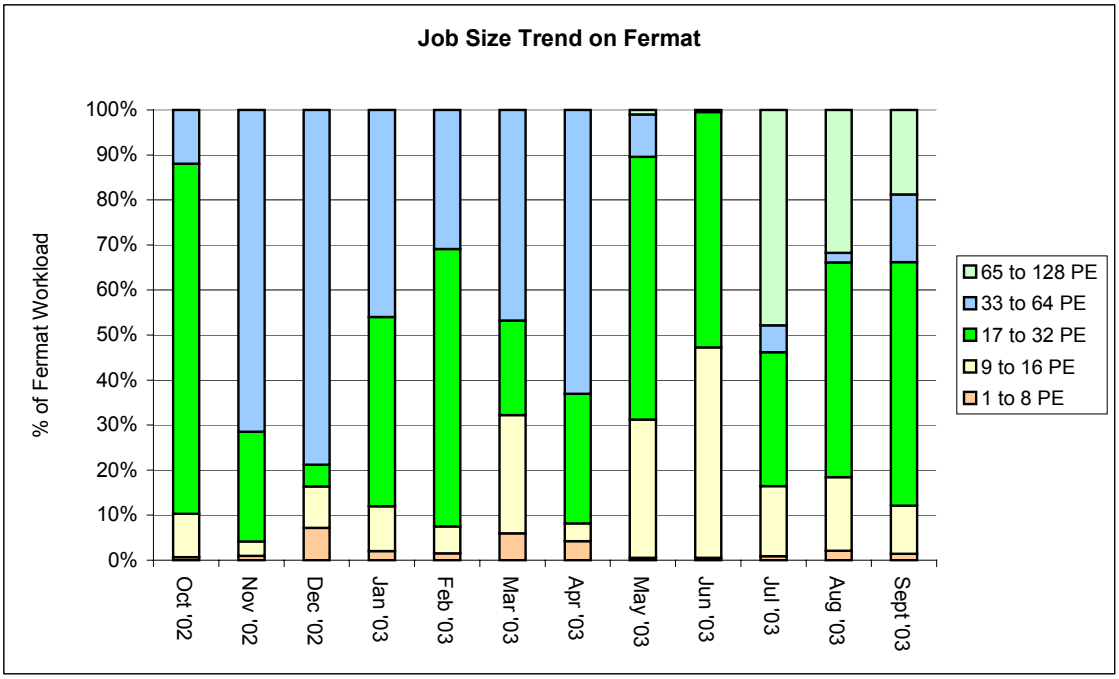
On Fermat, jobs were relatively evenly spread across the various ranges of PEs.



The last few months have seen increasing usage of the T3E for jobs above 256 PEs.



Usage on Green is tending more to the higher-end range of PEs.



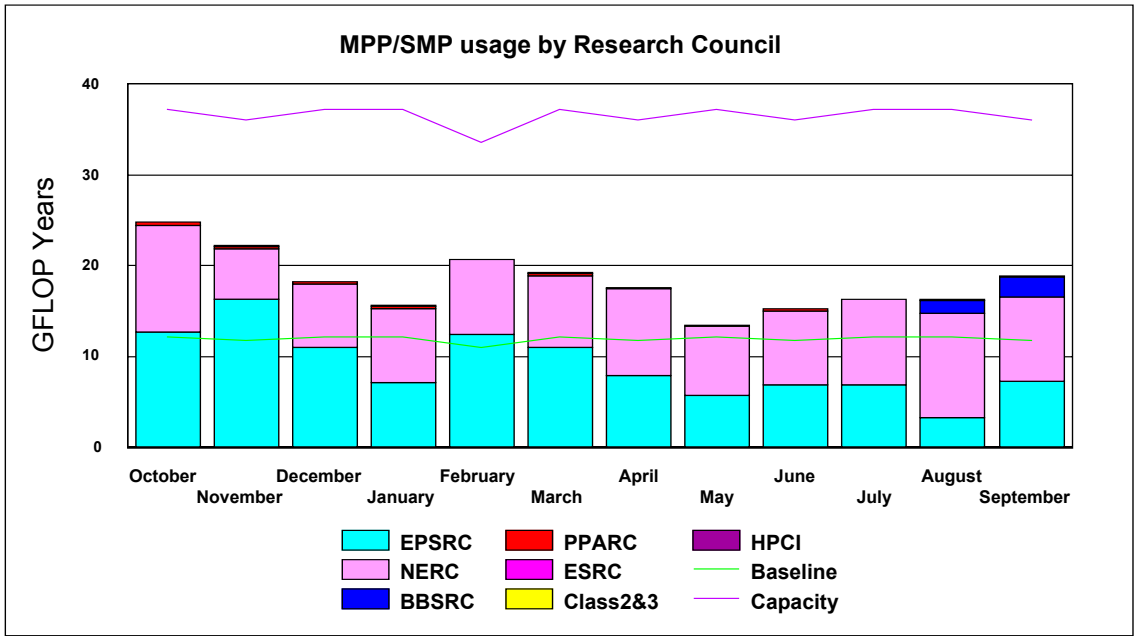
The trend on Fermat currently is that the greatest proportion of the workload is primarily in the mid-range of PEs.

2.2 System Usage Graphs

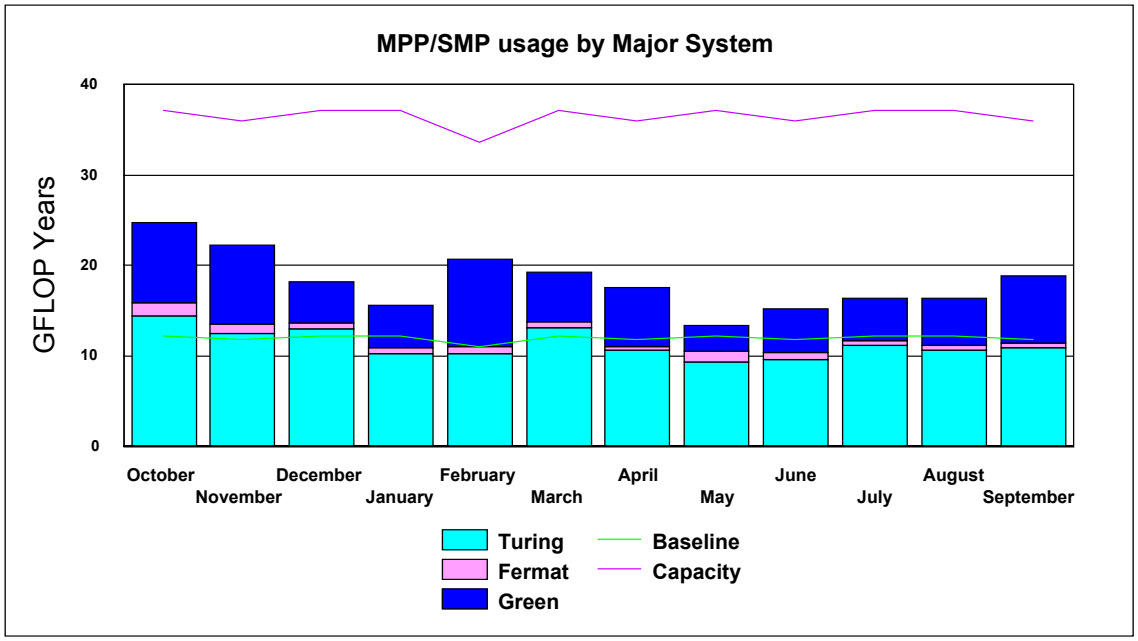
In all the Usage Charts, the baseline varies dependant on the number of days in each month, within a 365-day year.

2.2.1 Baseline System

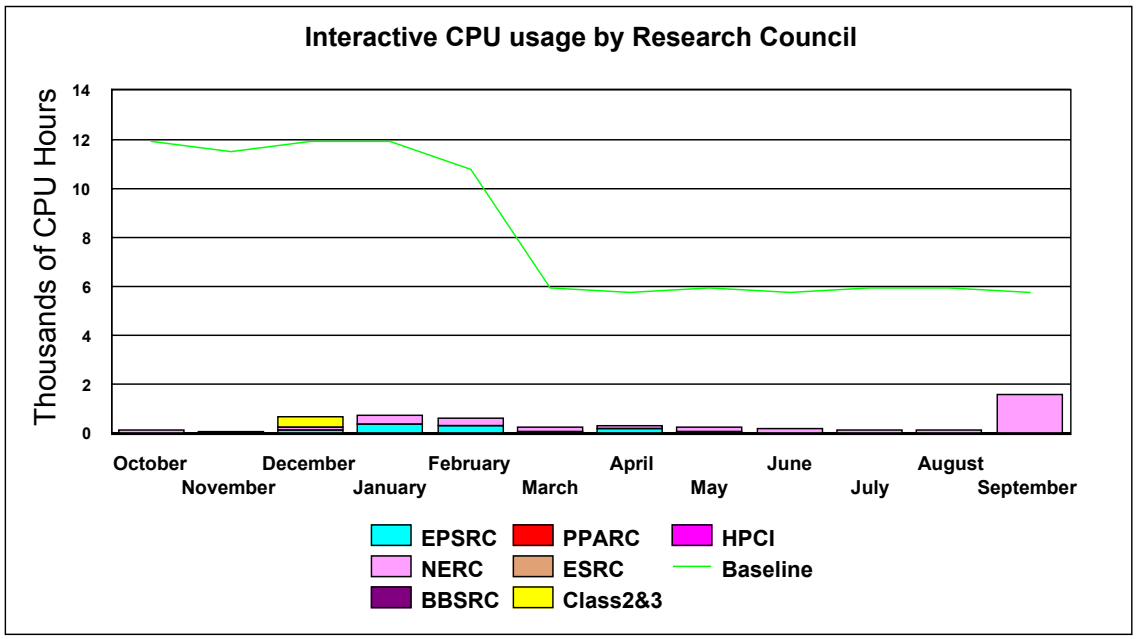
The graph below shows the Gflop Years utilisation on the CSAR systems by Research Council for the last 12 months. BBSRC has become a significant user of Green.



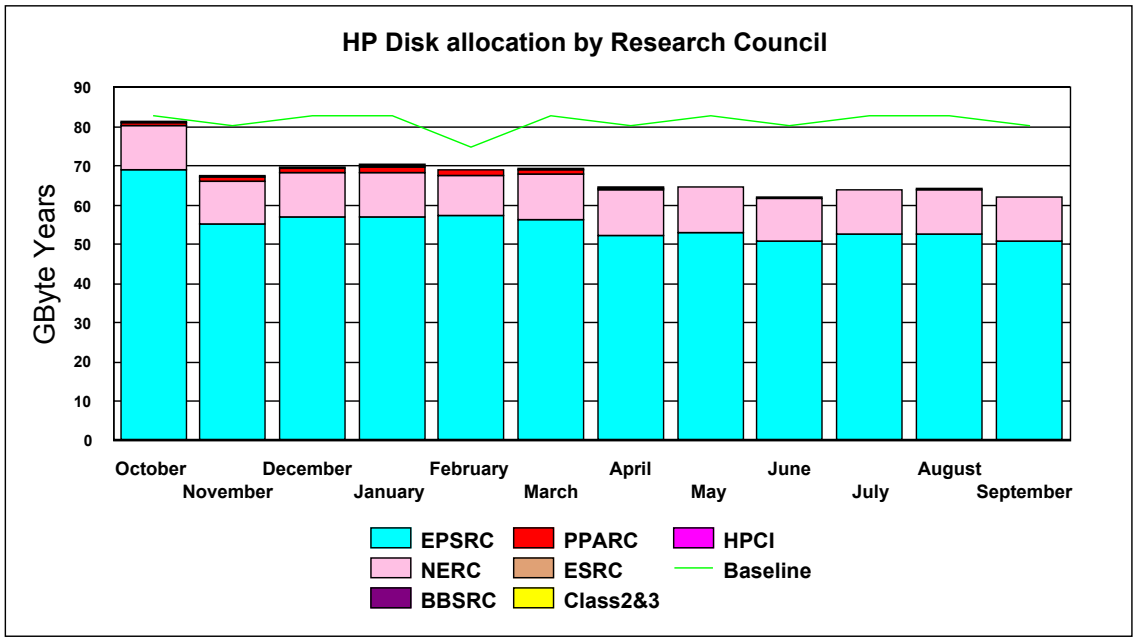
The graph below shows the same service utilisation by major system.



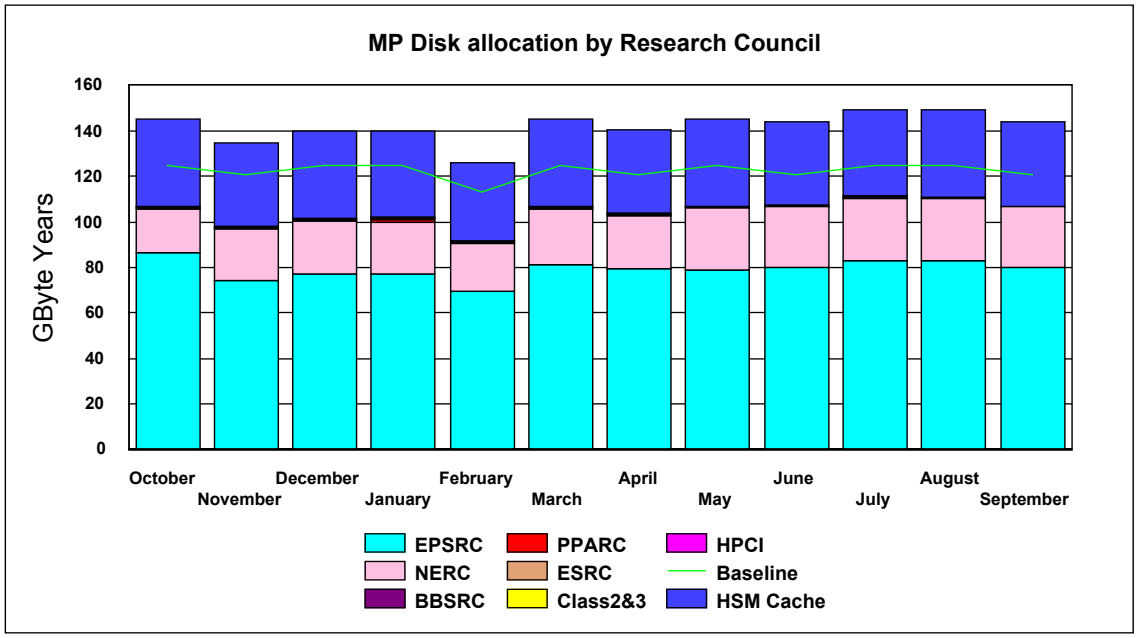
The next chart shows shows the historic interactive usage of the 'baseline' Fermat system (equivalent to 16@250Mhz CPUs) up to the end of February 2003, at which point the interactive usage was transferred to Wren and Fermat became a batch-only system.. Eight of the higher speed 500Mhz CPUs in the Origin 300 system (Wren) deliver the baseline capacity equivalent to that which was previously available on Fermat for interactive usage.



The next series of graphs illustrates the usage of the disk and HSM resources of the system.

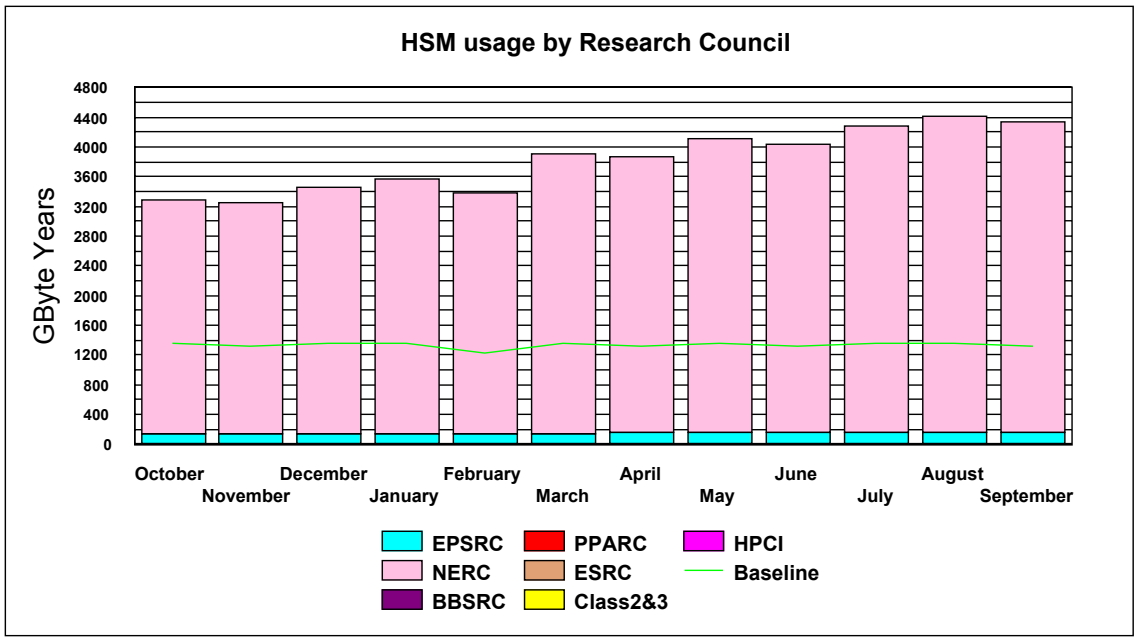


The graph above illustrates the historic allocation of the High Performance Disk on Turing.



This graph illustrates the historic allocation of the Medium Performance Disk on Fermat.

The graph below shows the historic HSM usage by Research Council funded projects, which has exceeded the overall Baseline of 16 Terabytes, and now totals 48 Terabytes.



2.2.2 Guest System Usage

There is currently no Guest System usage.

2.3 Service Status, Issues and Plans

Status

The service has been reasonably utilised throughout the third quarter of 2003, with usage exceeding baseline.

During the quarter there was a relatively balanced spread of work across all major systems.

The four additional fibre-attached tape drives on Fermat have helped to improve the response and reliability of the Data Migration Facility; response times have improved since this addition to the service.

Issues

There are no issues to report for this quarter.

Plans

Newton, the new 256P SGI Altix 3700 Itanium-2 system which will become the flagship of the CSAR service, has now been delivered to the Manchester site. Assembly of the machine together with system configuration and acceptance testing took place throughout September paving the way for Newton to enter full service on 1st October.

3 Project Management, Documentation and User Feedback

This section covers aspects relating to the registration of projects and users, the management of projects and resources, topics associated with documentation and user feedback.

3.1 Project Applications

13 applications for new CSAR projects were received in the third quarter of 2003, requesting a total of 201866 service tokens.

3.2 New Projects

4 new CSAR projects were started with 1242 service tokens being awarded in total.

3.3 Queries

Overall, 173 Class 1, 2 and 3 queries relating to the CSAR service were received by the Helpdesk between 1st July and 30th September 2003.

3.4 Service Quality Tokens

One black mark was recorded for the poor performance of the Origins. The Status Page was updated to include more detailed information regarding the situation and the black mark was removed on the same day that it had been awarded.

3.5 Annual Report

The CfS Annual Report is currently being prepared for publication.

3.6 CSAR Focus

The Summer 2003 edition has been printed and distributed.

3.7 CSAR Website

The current CSAR website has been updated to include pages on the Altix. Fiona Cook is managing the creation of the new CSAR website.

3.8 Launch of Altix

A 32-PE Altix system (Reynolds) was introduced into service in June 2003. This was followed by a 256-PE Altix system (Newton) in the second week of September.

4 Scientific Application Support Services

4.1 Training and Education

The new course timetable for CSAR/HPC courses for Semester 1 of the 2003/2004 academic year is now available at http://www.csar.cfs.ac.uk/using/courses.

Several new courses have been developed including courses on the new Altix system and the Unified Model.

4.2 Consortia Support/Software

Work has concentrated on porting and optimising codes both for the new Altix service and for the removal of Turing. There continues to be great activity in this area. Installation of software on the Altix continues.

Some highlights of the porting and timing process (ongoing):

- Newton is ~3 times faster than Green for most user codes, with some (eg Gaussian) up to 5 times faster
- The Unified Model is 25% faster than on a p690 system
- The molecular dynamics codes NAMD and AMBER are both ~50% faster than on a p690 system

There are a number of consortia that we have provided optimization work for:

CgLES

CgLES is a C/C++ code which required porting to Newton but there was also a request for optimization work to be done for Green. After analysis of the code it was found that the compiler did a poor job of dereferencing arrays. This needed to be done manually on 4 of the key functions in the code. This resulted in an increase in performance of between 200% and 300%. This left a code that was heavily dominated by communication; in fact on 14 processors, 50% of runtime was communication. Further work was done to convert communcations to single-sided MPI communication which reduced the MPI to 20% of the runtime of the code. This work was performed by Neil Stringfellow.

Kai Luo

This began as code porting to Green (bugs in the Intel compiler are preventing this being ported to Newton). Minor serial code optimizations provided no significant gains. Communication also proved to be a problem. The communication involved numerous sends from one processor to its neighbour in order to perform halo exchanges followed by receives. This was replaced by sendrecvs using a derived type; this also opens up the possibility of single copy communication (a significant parallel optimization), and will also allow further optimization to convert the code to single-sided communication. This work was performed by Neil Stringfellow.

H2Mol

H2Mol is an in-house code from Queen's University Belfast. This code is part of the porting effort to Newton. Initial runs of the code showed extremely poor timings, taking longer than Green. The code is dominated by multiplications of the following form:

$$C = A \times B + B \times A$$

The major reason for the poor performance was that this was implemented as a pair of matmul operations and A is not stride one data. Matmul appears to be inlined into the code on the altix which led to really poor cache use and also exhibited a feature unique to the Intel chip, cache bank conflicts. Work was done to replace this function in a convenient manner to further development or optimization exercises, and on this resulted in the code taking half as long to run. Further optimization was performed after discovering that certain matrices were symmetric, and replacing loop structures with library calls to SCSL routines. Work performed by Kevin Roy.

Helium

Helium is another in-house code from Queen's University Belfast. It has also been part of the porting program to the Altix. Work has been performed to increase the performance by nearly 20% on Newton and some minor work on MPI buffering has been completed. The code is currently undergoing a more rigourous examination of its parallel performance. Work performed by Kevin Roy.

Cse082 – EROS-UK

This is work that was begun in the previous quarter, and also reported upon in the previous report on optimization work. Further analysis of this code revealed a dependence between the problem and the number of processors that it ran on optimally. On the problems that we were given this didn't give any concerns of scalability. Serial optimization proved extremely difficult due to the complex datastructures employed. The compiler found it too difficult to provide any optimization beyond –O2. The limited support resources of this project meant that serial optimization could not be performed. Work performed by Kevin Roy.

ftLMPs

FtLMPs is a laser materials processing code which uses a 3D adaptive mesh refinement scheme. Initial work on this code has been done identifying problem areas and will be continued to the next quarter. Work performed by Kevin Roy.

4.3 Netsolve

Work has continued on the Netsolve project.

5 Collaboration and Conferences

5.1 MRCCS Projects

5.2.1 Virtual Prototyping/Finite Elements

Lee Margetts has finished writing the prototype of the new parallel FEA code. Work is ongoing.

5.2 Events

The following events were attended by SVE staff:

- Several staff were involved in the UK e-Science All Hands Meeting in Nottingham, 2nd – 4th September. At this meeting a demonstration was performed that used Manchester's Bezier (SGI Onyx) to visualise data produced in a computational steering experiment run on the Level-2 Grid. The simulation, a parallel Lattice-Boltzmann code which simulated the phase separation of two fluids, was run on eight processors of an SGI Origin at the Welsh e-Science Centre in Cardiff. A grid service providing the steering interface for this simulation was hosted in the OGSI-Lite framework running on a Linux machine in Manchester. The visualization software performed isosurfacing using the vtk library (C++ wrapped by Tk/Tcl) on an SGI Onyx at the University of Manchester and again had a separate grid service providing its steering interface. The application-launching software was built on the Globus 2 Toolkit and the steering client was a Qt-based GUI built on the RealityGrid steering library. Both of these were run on a laptop at the North West Regional e-Science Centre's booth in Nottingham. The steering client connected to the simulation's grid service via SOAP over http. The connection between the laptop and the Onyx in Manchester had a bandwidth of approximately 100 Mbps and used VizServer to compress and transfer the images.

- Zoe Chaplin attended the CAS2K3 (Computing in Atmospheric Sciences) workshop in Annecy, France from 7th to 11th September. The workshop was well attended by representatives from most of the major climate and weather research centres as well as the main computer vendors. Zoe presented a report on the performance of the Unified Model 5.3 for the mes and the global model at this workshop.

- Andrew Jones and Terry Hewitt attended the SGI User Group Committee meetings.

- Andrew Jones and Terry Hewitt attended the HPCy vendor's day in Guildford on 9th September and the NW Grid vendor's day. CSAR will be involved in the NW Grid.

6 Added Value Services

6.1 International Conferences

Preparations are well underway for the exhibition at SC2003 (Phoenix, USA in November) and for SCGlobal 2003.

6.2 Seminars

6.2.1 MRCCS/ESNW Seminar Series

The following seminars have been held during the last three months:

- Delivering e-Science and Visualization to Medical Applications – Nigel John, Joanna Leng, Mark Riding, Manchester Visualization Centre, University of Manchester.
- ESNW Middleware Discussion.
- Pedro: A Tool for Bridging e-Science with Grid Services. A Case Study for Medical Informatics – Chris Garwood, University of Manchester.
- Applications of Ontologies to Grid Middleware Interoperability – Kevin Garwood, University of Manchester.
- A Grid Application Framework based on Web Services Specifications and Practices – Dr. Savas Parastatidis, School of Computing Science, University of Newcastle upon Tyne.

6.3 Summer School 2003

The Summer School 2003 on HPC in finite Element Analysis, jointly organised by MRCCS and the National Science Foundation (NSF) of the USA took place from 1st to 5th September. Speakers attended from France and Japan in addition to the speakers from the USA and the UK. Questionnaires showed that delegates were very pleased with this meeting. We will be in contact with participants to see how these ideas can be taken forward in the future. We are also hoping to obtain funding from EPSRC for further work.

6.4 Visualization

.