CSAR Service

Consolidated Management Report

2nd Quarter 2002

Management Summary

The system continues to be the flagship HPC facility for UK Academia and Industry, enabling World-Class research and development.

The number of users has grown to a total of 542 to date.

The Cray T3E (Turing) system continued to run almost to full capacity again this quarter.

The Origin 3000 (Green) now has 512 CPUs, and usage is growing steadily. 62% of the usage has been greater than 64 CPUs in size.

The T3E has seen 78% of the Quarter's workload at greater than 64 PEs in size.

CfS remains active in the UK Grid Forum.

CSAR is currently installing new facilities to enhance many aspects of the service. These include a new interactive machine 'Wren' and new high performance disks via a storage area network. In addition, NQE (Network Queuing Environment – the batch job queuing system) is in the process of being replaced by LSF (Load Sharing Facility – produced by Platform Computing) on the SGI Origin Systems.

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 was the year 2000 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	2001/2 July	Aug	Sept	Oct	Nov	Dec	Jan	Feb	March	April	May	June
HPC Services Availability												
Availability in Core Time (% of time)	98.49%	98.49%	98.49%	98.60%	98.60%	100.00%	99.86%	99.73%	99.70%	96.17%	96.08%	97.66%
Availability out of Core Time (% of time)	98.49%	100%	99.40	99.50%	99.50%	98.49%	99.89%	99.85%	99.97%	97.75%	99.90%	99%
Number of Failures in month	4	2	2	2	2	4	2	1	2	2	1	4
Mean Time between failures in 52 week rolling period (hours)	438	398	365	365	365	337	350	324	313	302	324	313
Fujitsu Service Availability												
Availability in Core Time (% of time)	100%	100%	100%	100%	100%	100%	100%	100%	100%	96.89%	100%	100%
Availability out of Core Time (% of time)	100%	100%	100%	100%	100%	100%	100%	100%	100%	98.92%	100%	100%
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	<1	<1	<1	<1	<1	<1	<1	<2	<1	<1	<2	<5
Administrative Queries - Max Time to resolve 95% of all queries	<1	<2	<1	<1	<0.5	<2	<0.5	<1	<2	<2	<3	<5
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	<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 month	1	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.

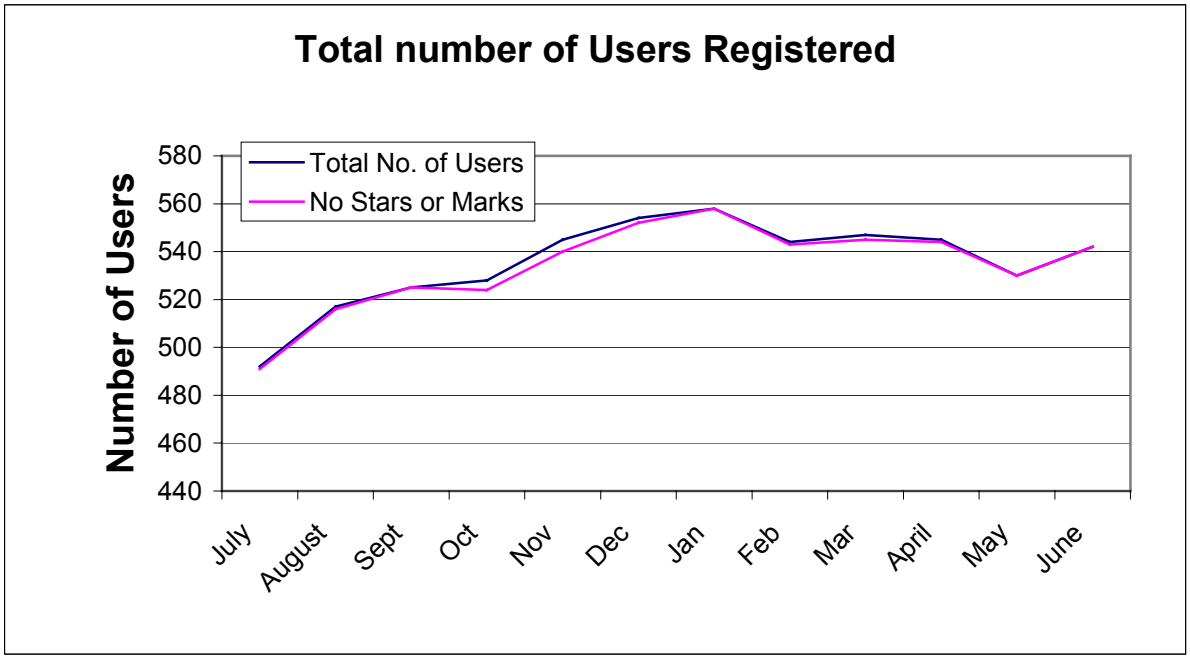
CSAR Service - Service Quality Report - Service Credits

Service Quality Measure	2001/2 July	Aug	Sept	Oct	Nov	Dec	Jan	Feb	March	April	May	June
HPC Services Availability												
Availability in Core Time (% of time)	0.039	0.039	0.039	0.039	0.039	-0.058	-0.039	-0.039	-0.039	0.078	0.078	0.078
Availability out of Core Time (% of time)	0.000	-0.047	0	-0.039	-0.039	0.000	-0.047	-0.047	-0.047	0.039	-0.047	0.000
Number of Failures in month	0.008	0	0	0	0	0.008	0	-0.008	0	0	-0.008	0.000
Mean Time between failures in 52 week rolling period (hours)	0	0	0	0	0	0	0	0	0	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	-0.016	-0.016	-0.016	-0.016	-0.016	-0.016	-0.016	0	-0.016	-0.016	0	0.031
Administrative Queries - Max Time to resolve 95% of all queries	-0.016	0	-0.016	-0.016	-0.019	0	-0.019	-0.016	0	0	0.016	0.031
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	-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	-0.003	0	0	0	0	0	0	0	0	0	0	0
Monthly Total & overall Service Quality Rating for each period:	-0.02	-0.03	-0.02	-0.04	-0.04	-0.05	-0.08	-0.08	-0.07	0.03	0.00	0.05
Quarterly Service Credits:			-0.07			-0.13			-0.23			0.08

Table 3

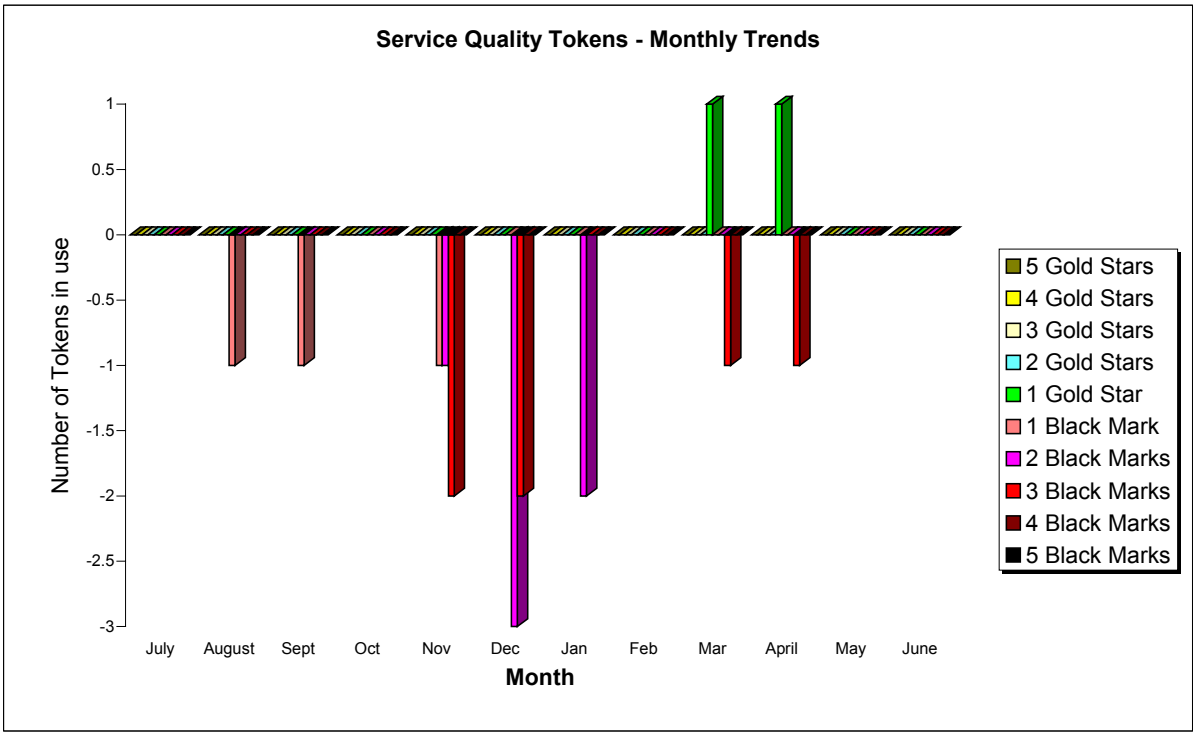
1.2 No. of Registered Users

The current position at the end of the quarter is that there are 542 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.

At the end of the quarter there are 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.

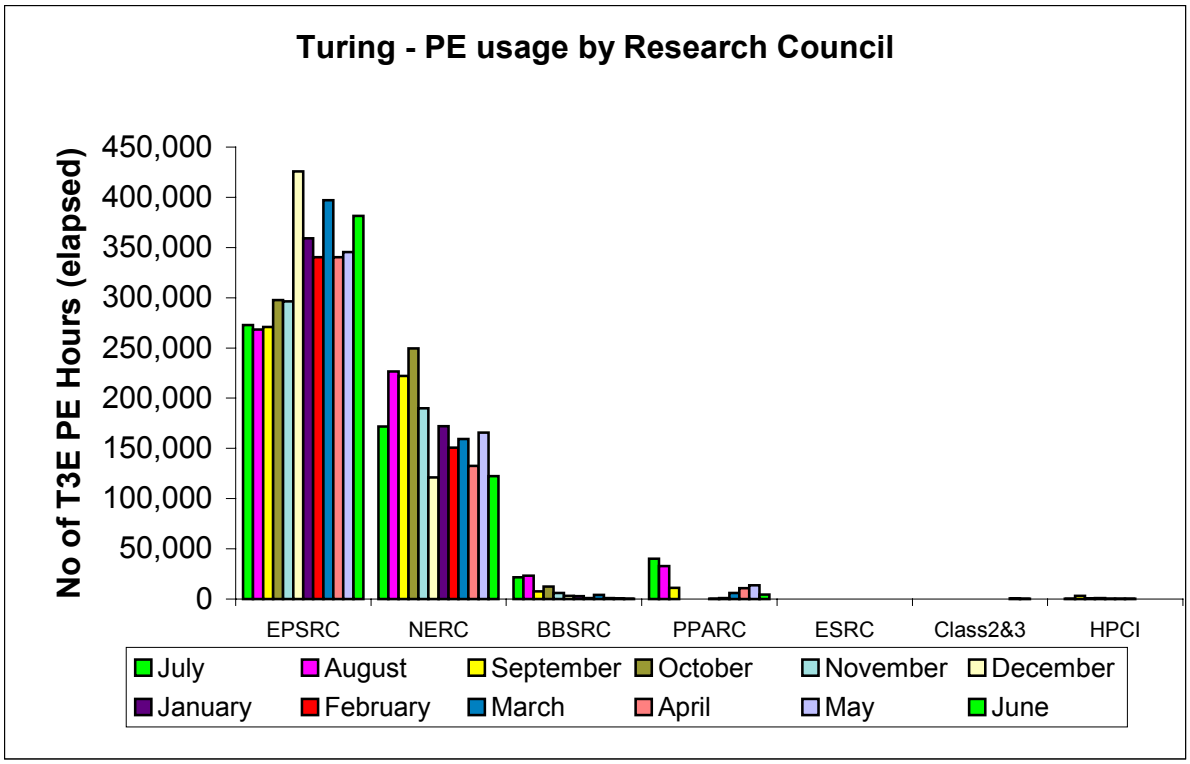
c) High Performance Disk (T3E) allocated for User Data by month, showing the allocated space each month in GBytes, split by Research Council. The Baseline Capacity (1 Terabyte) is shown by an overlaid horizontal line.

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.

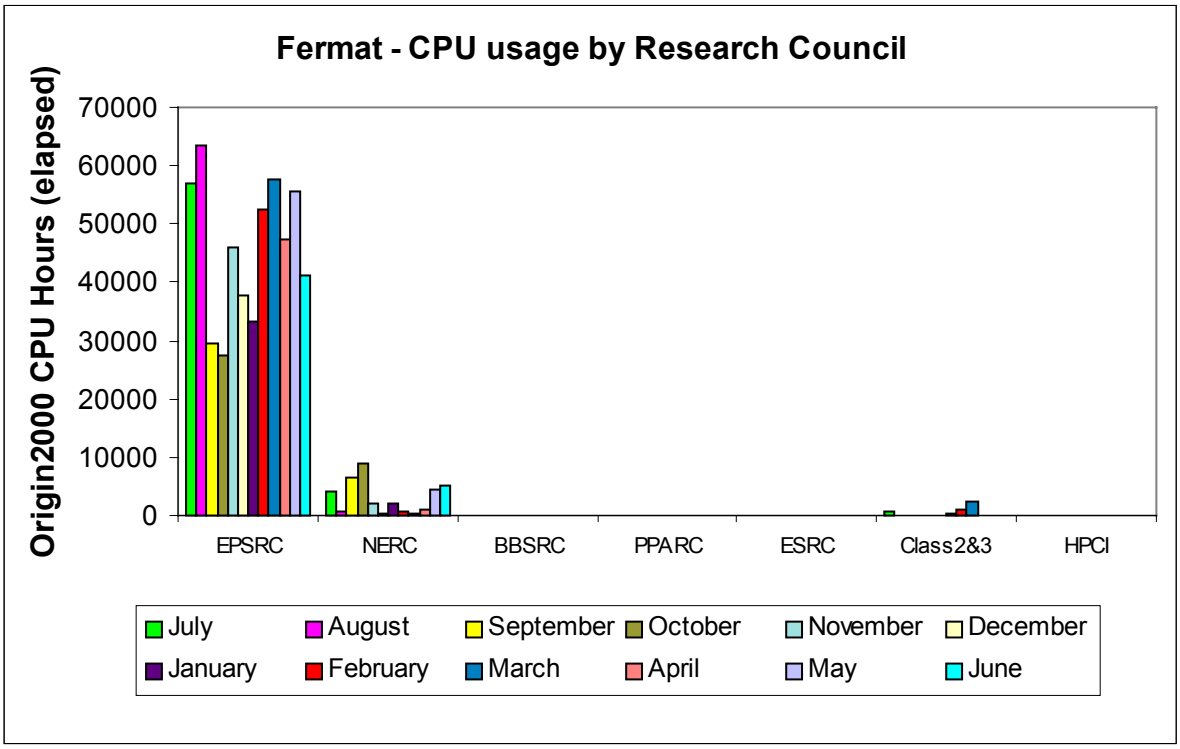
e) HSM/Tape Usage (T3E) by month, showing the volumes held each in GBytes, split by Research Council. The Baseline Capacity (16 Terabytes) available will be shown by an overlaid horizontal line.

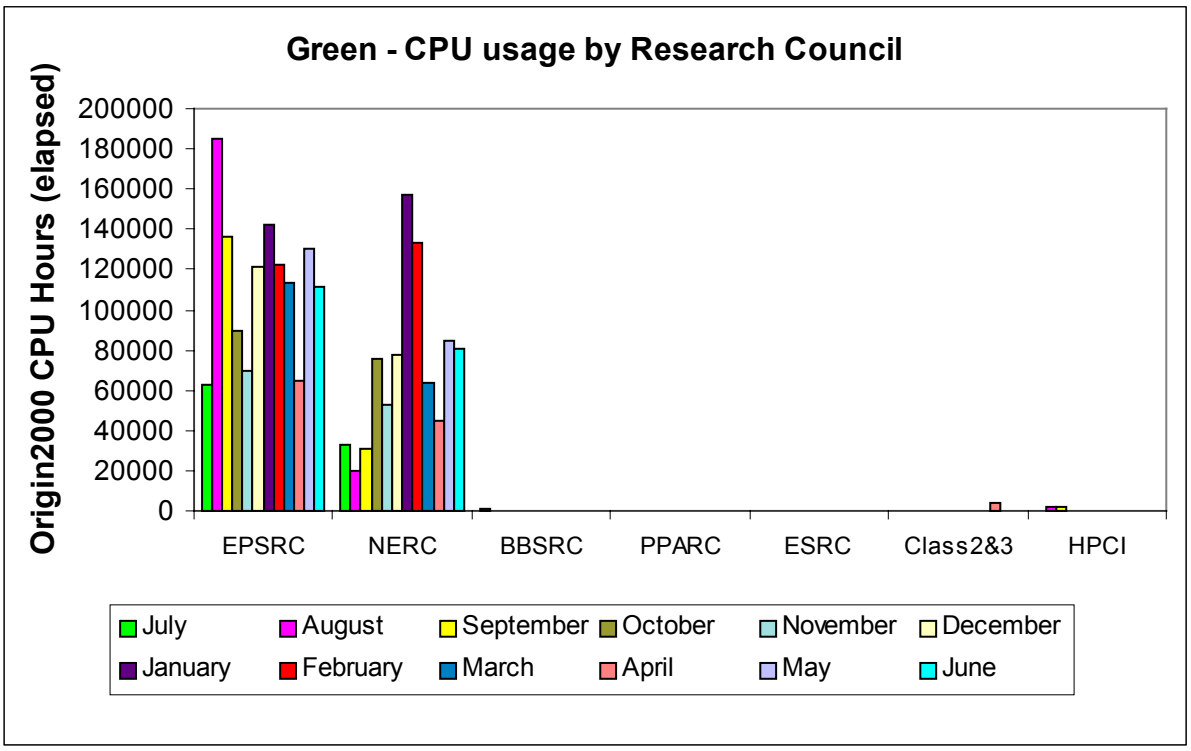
2.1 Service Usage Charts

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

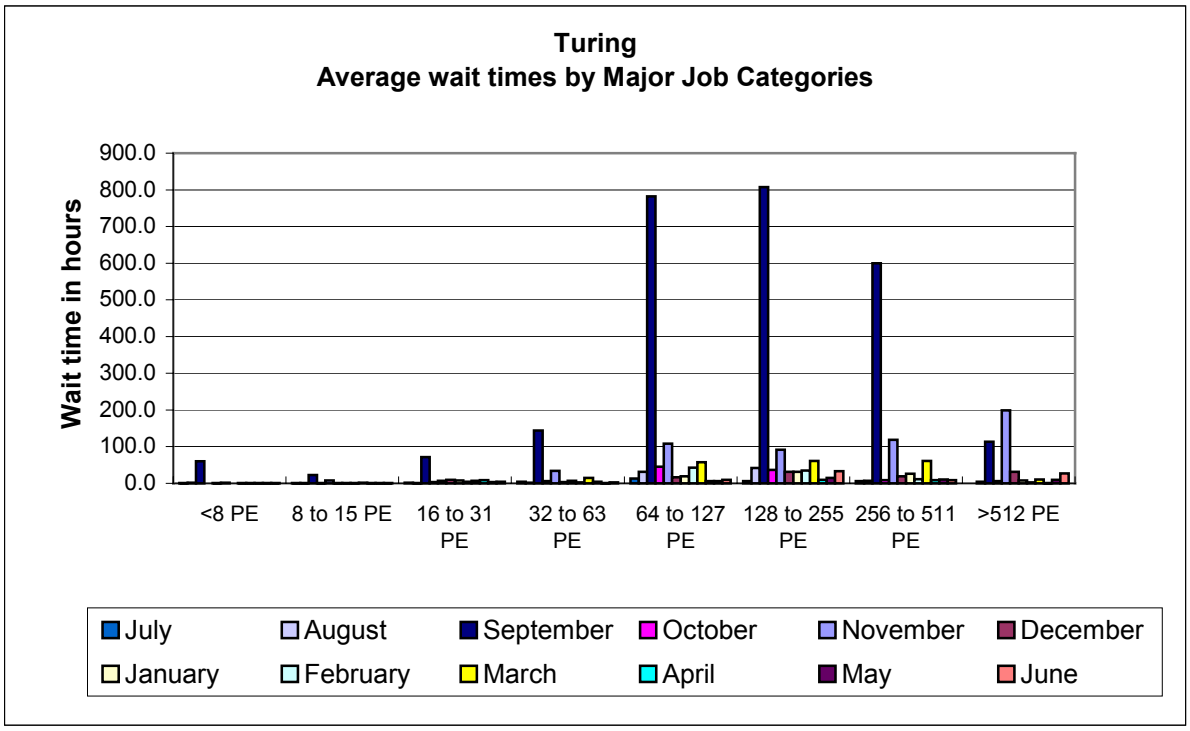


Turing PE usage is shown by Research Council during the last 12 months of service in the above chart.

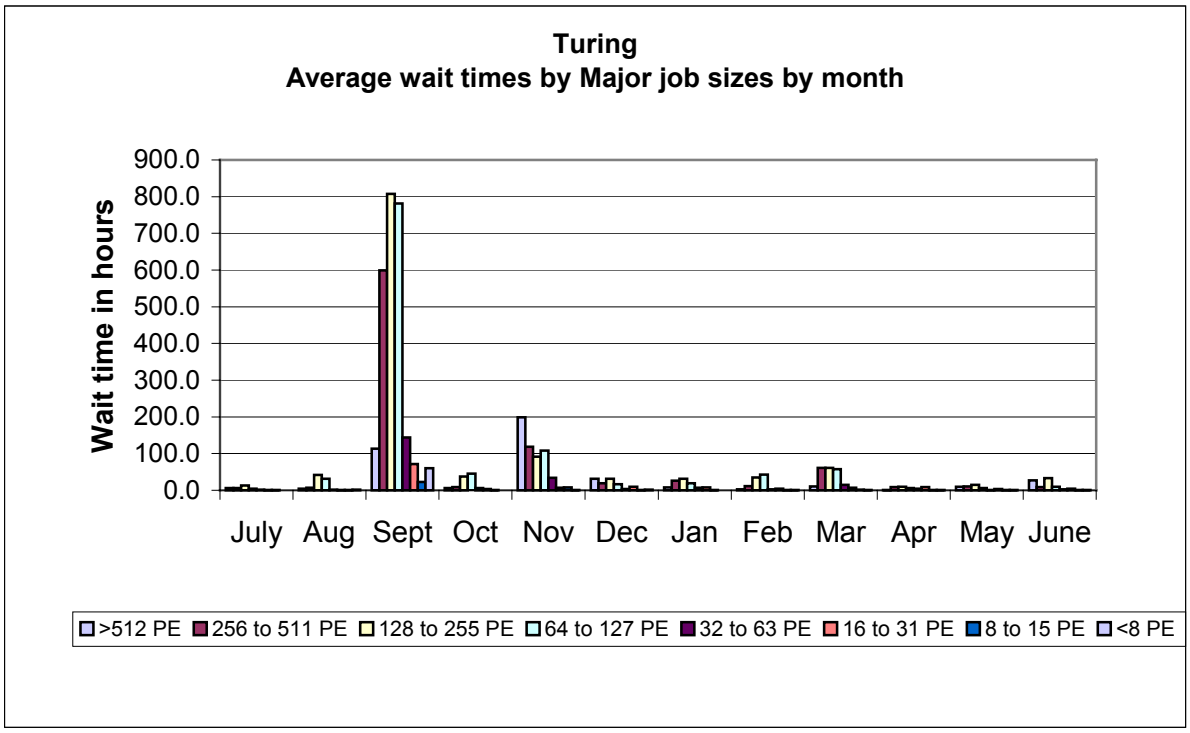




Usage of the two SGI Origin systems, Fermat and Green, is shown by Research Council during the last 12 months of service in the above charts.

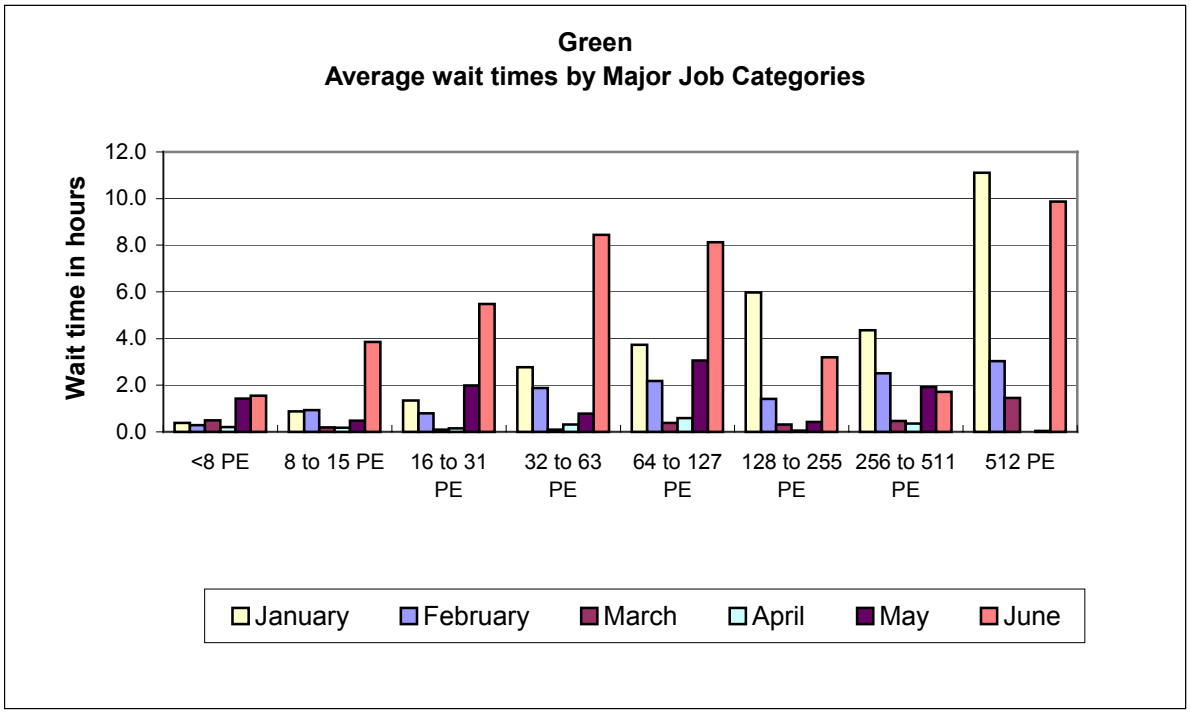


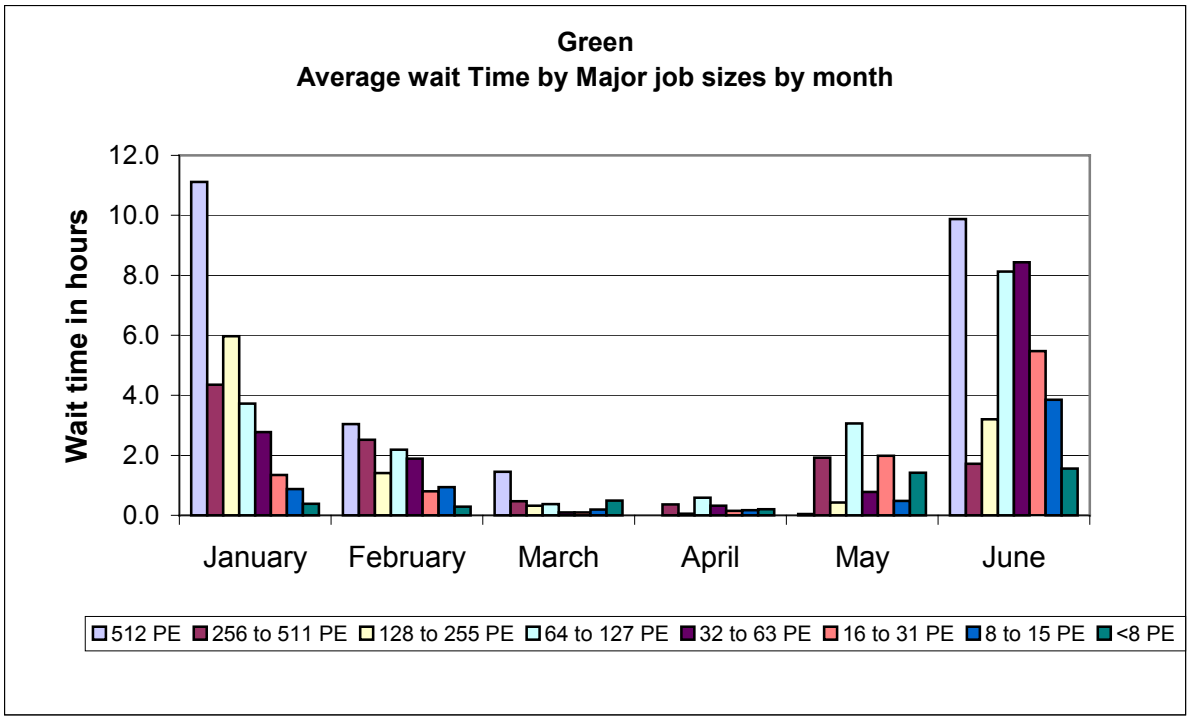
The above chart shows the wait time trend in hours on the Turing system. The quarter has seen job wait times gradually decrease.



It can be seen from the above graph that the additional capacity added to the service is now managing the queue wait times on the T3E (Turing) to an acceptable level.

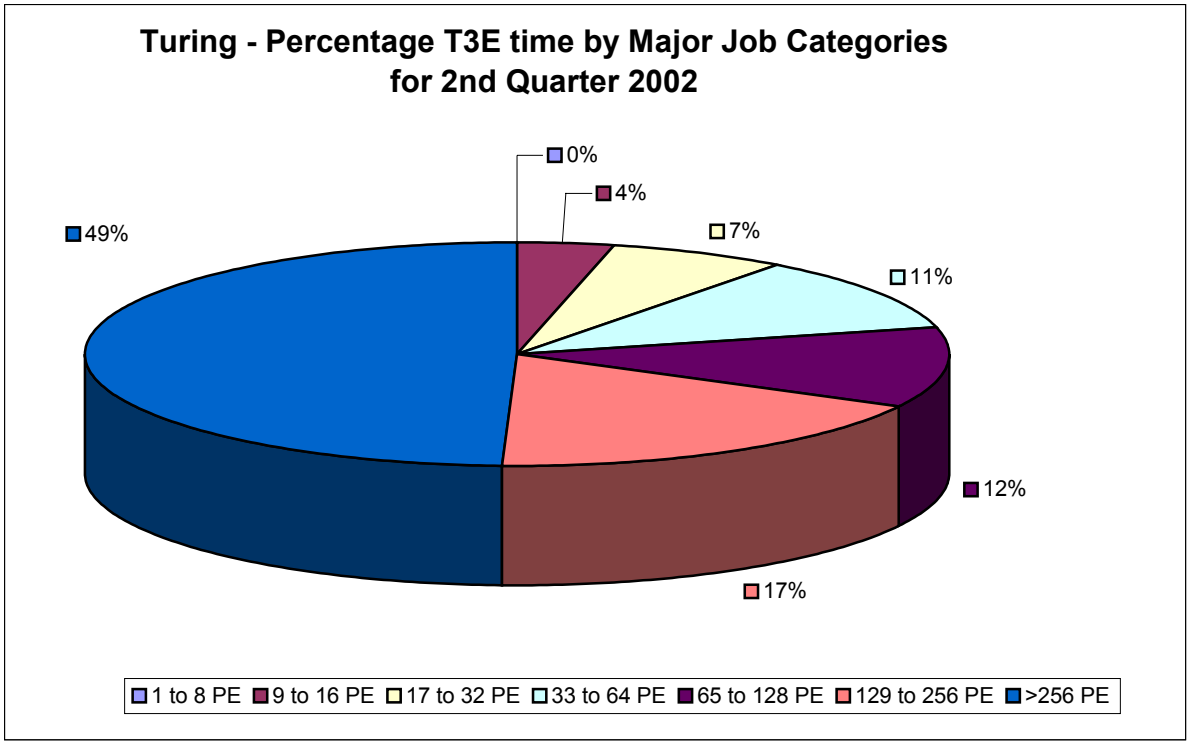
The following two charts show average wait times in hours for the quarter on the Origin 3000 (Green).



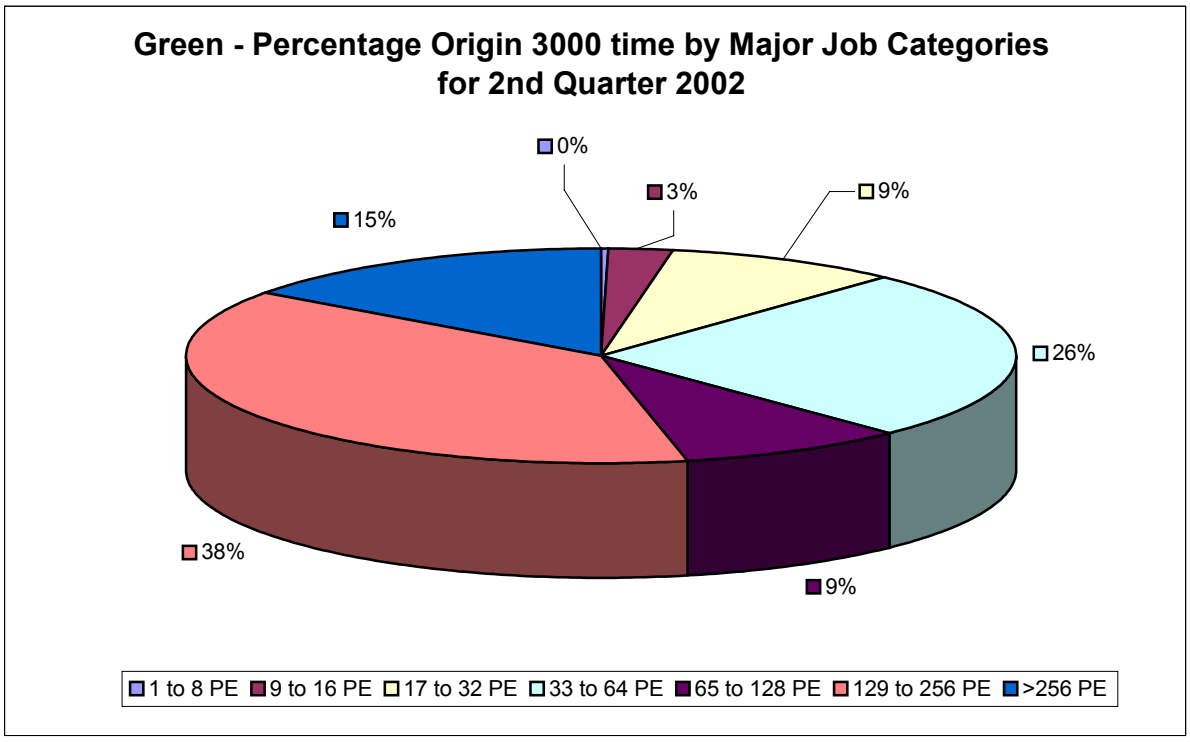


Green wait times are higher due to fuller loading of the system typically requiring 24 hours of processing.

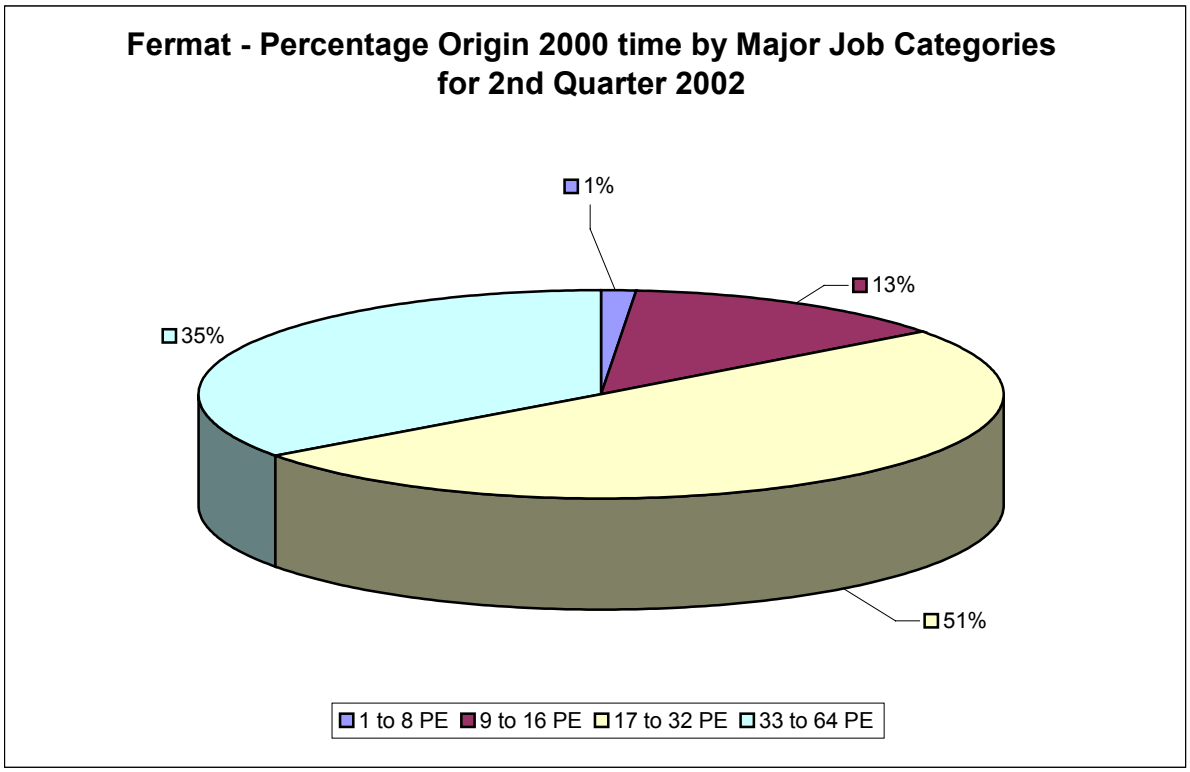
The next three charts show the percentage PE time utilisation by the major job categories on the Turing, Green and Fermat systems for the 2nd quarter 2002.



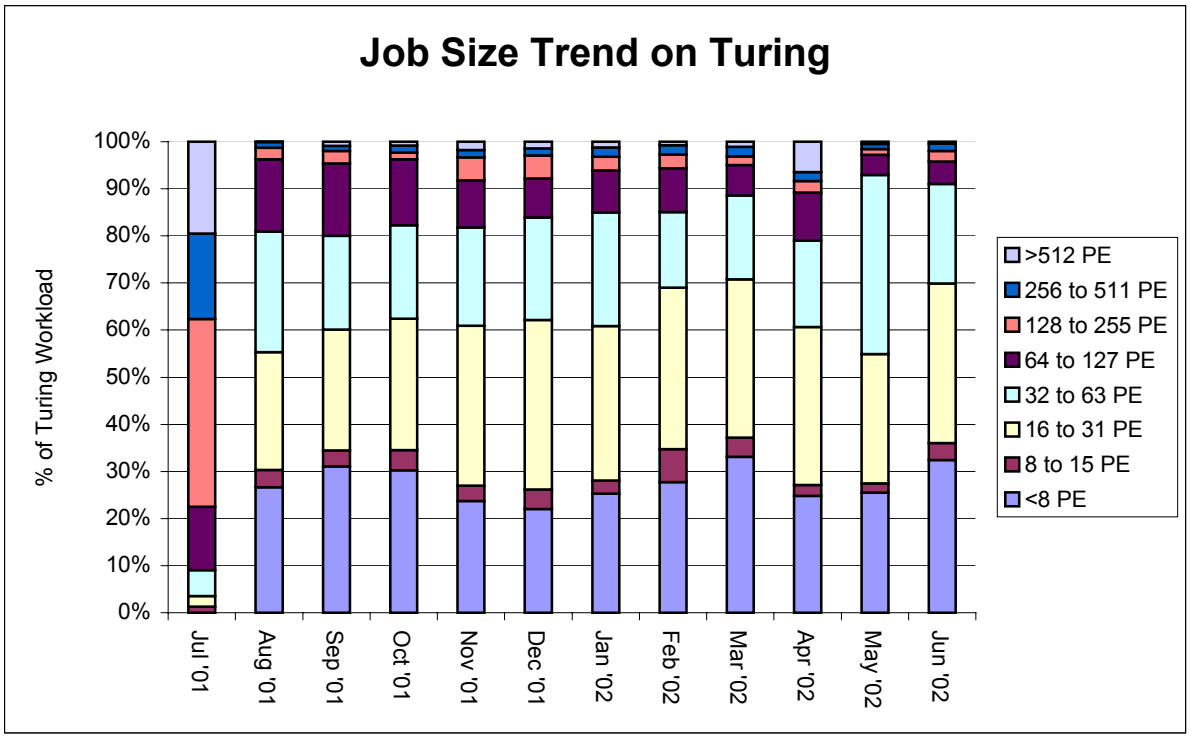
The predominant job size on Turing is large, with the percentage of jobs greater than 64 PEs for the quarter being 78%.



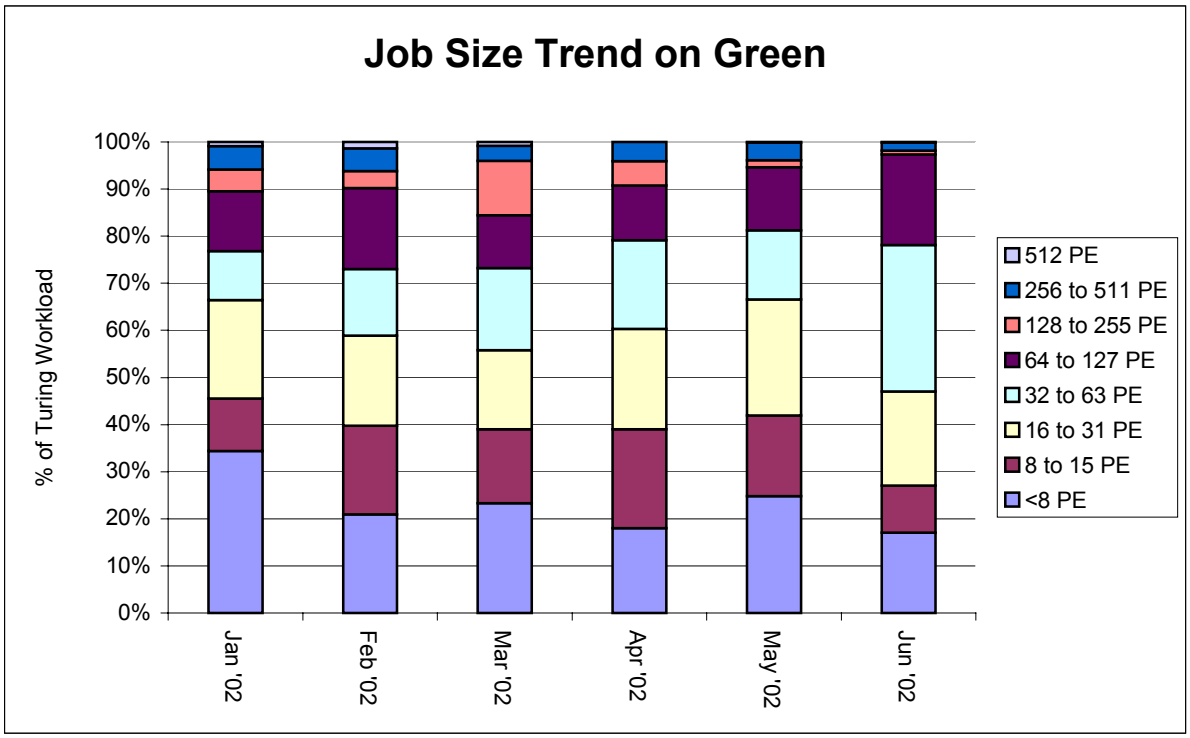
On Green, the percentage of CPU utilisation used by jobs requiring greater than 64 CPUs for the quarter was 62%.



On Fermat the greatest proportion of jobs during the quarter were in the 17 to 64 PE range.at 86%.



From the above chart it can be seen that the workload on the T3E has now stabilised.

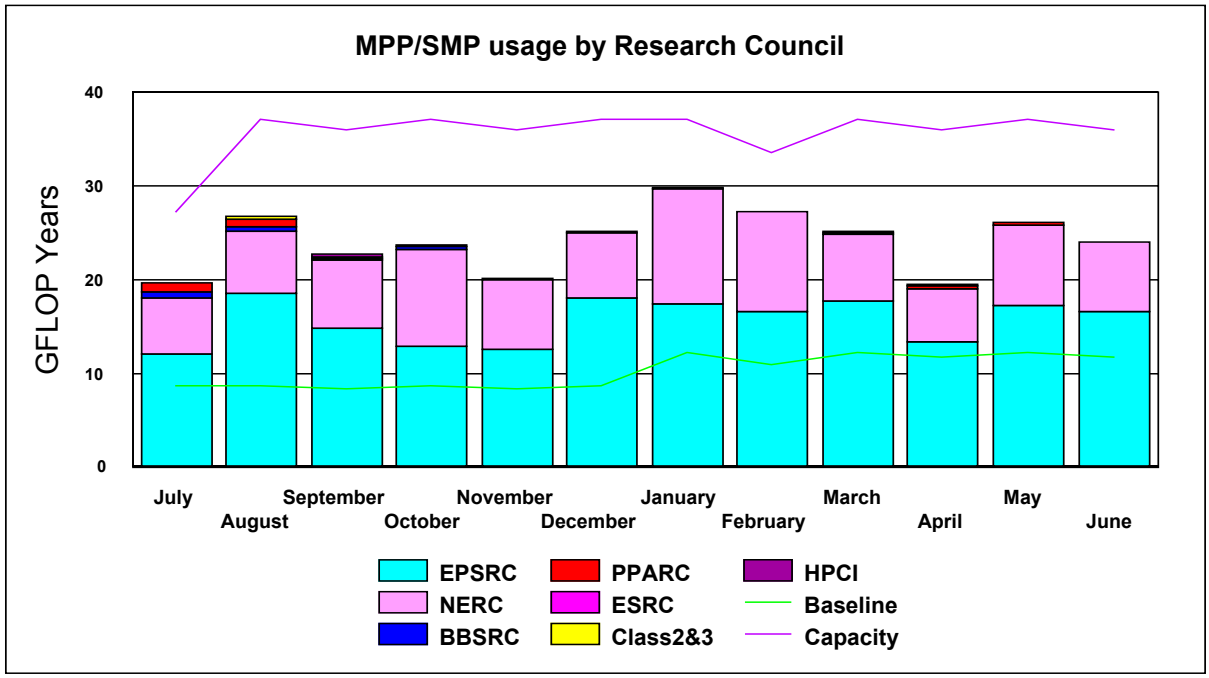

Job Size Trend on Green
% of Turing Workload
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Jan '02
Feb '02
Mar '02
Apr '02
May '02
Jun '02
512 PE
256 to 511 PE
128 to 255 PE
64 to 127 PE
32 to 63 PE
16 to 31 PE
8 to 15 PE
<8 PE

2.2 System Usage Graphs

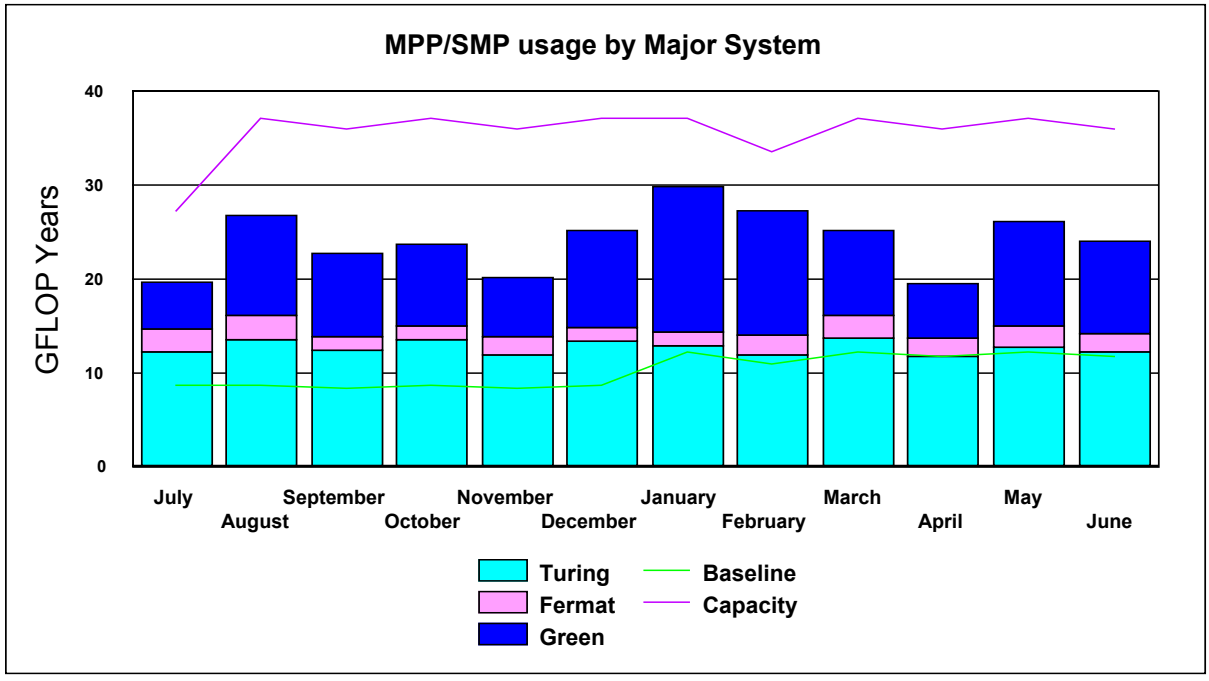
In all the Usage Charts, the baseline varies dependant on the number of days in each month, within a 365-day year.

2.2.1 Baseline System

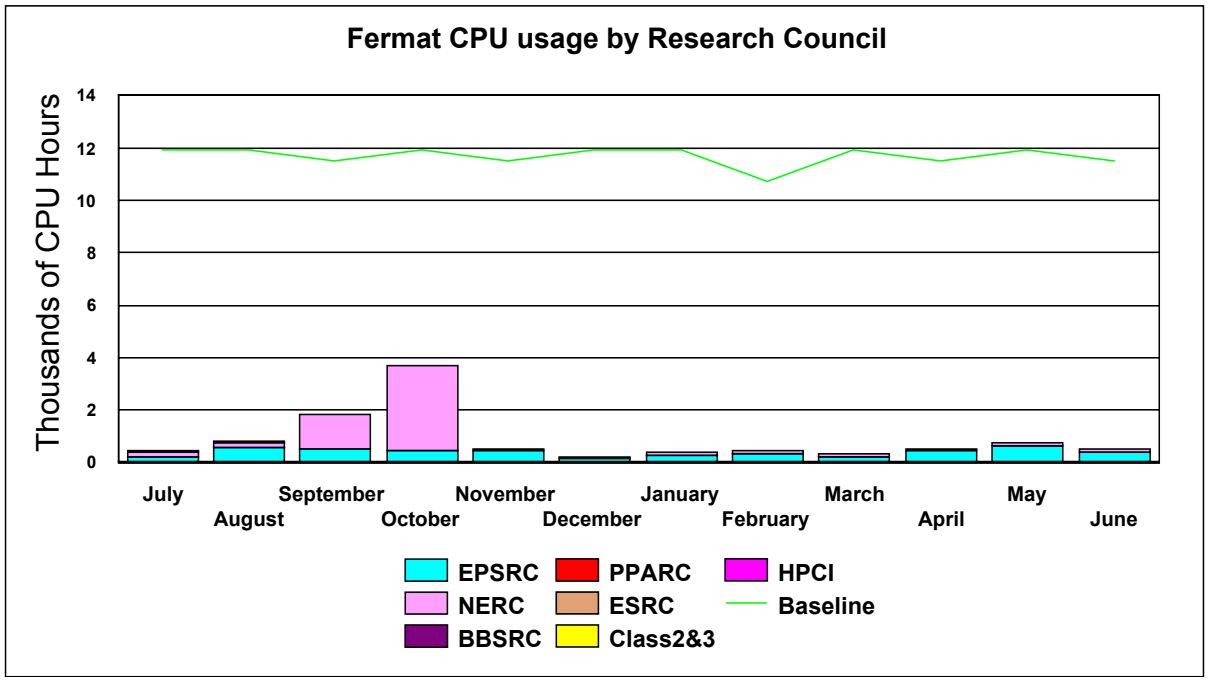
The graph below shows the Gflop year's utilisation on CSAR's systems by Research Council for the last 12 months.



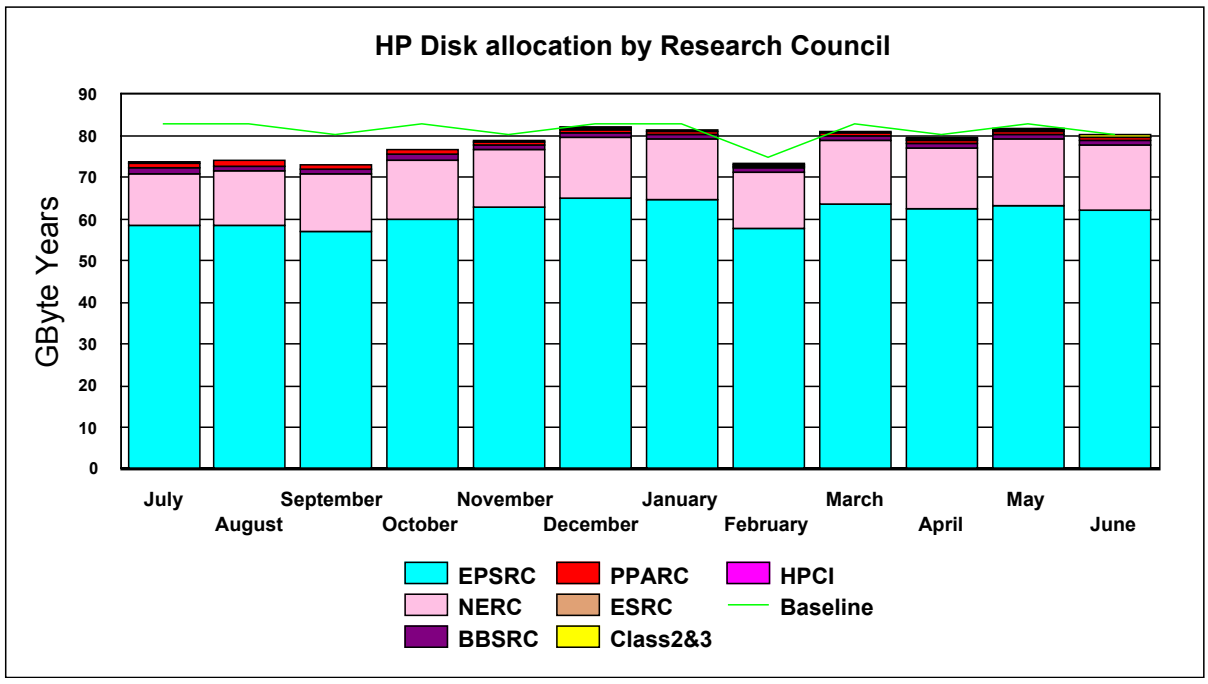
The graph below shows the same service utilisation by major system. .



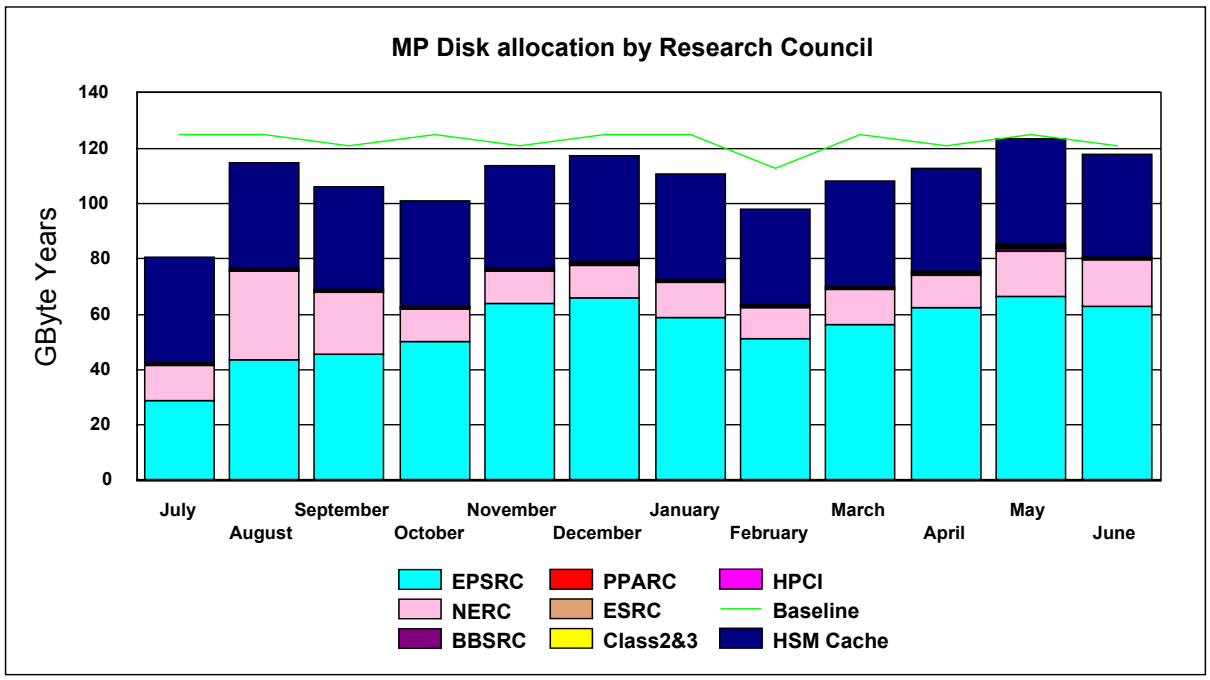
The next chart shows the historic interactive usage of the 'Baseline' Fermat system (equivalent to 16 @250Mhz CPU's).



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

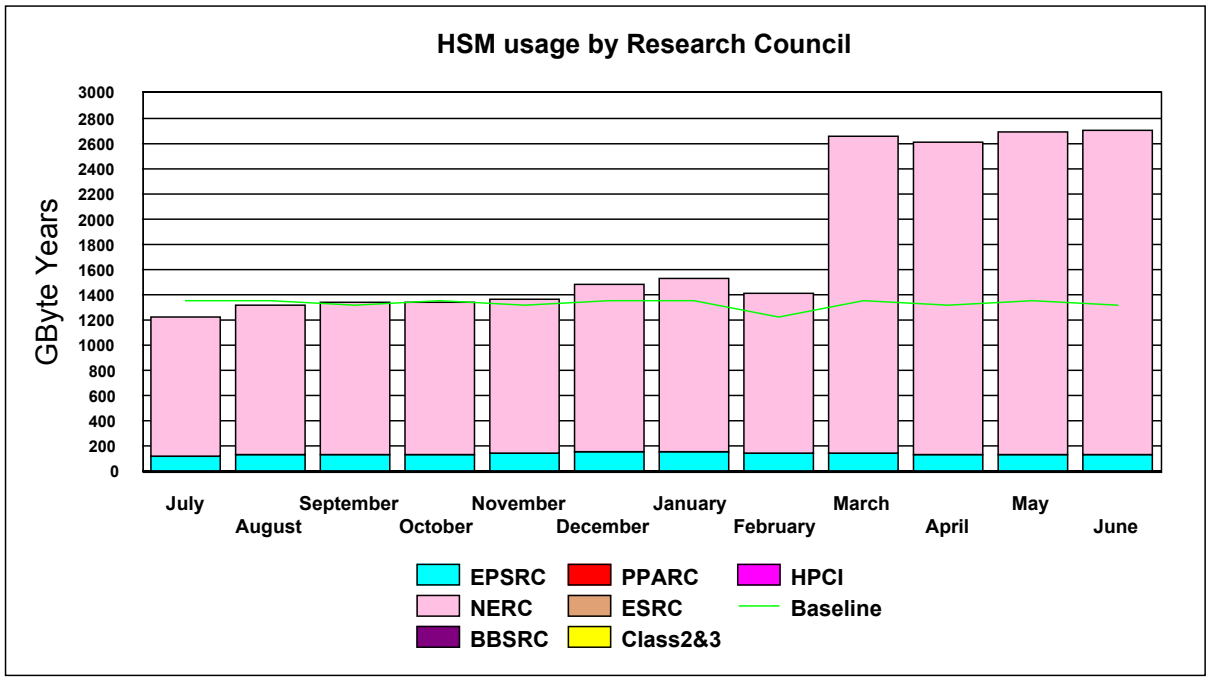


The preceding graph illustrates the historic allocation of the High Performance Disk on Turing, showing baseline disk fully utilised.

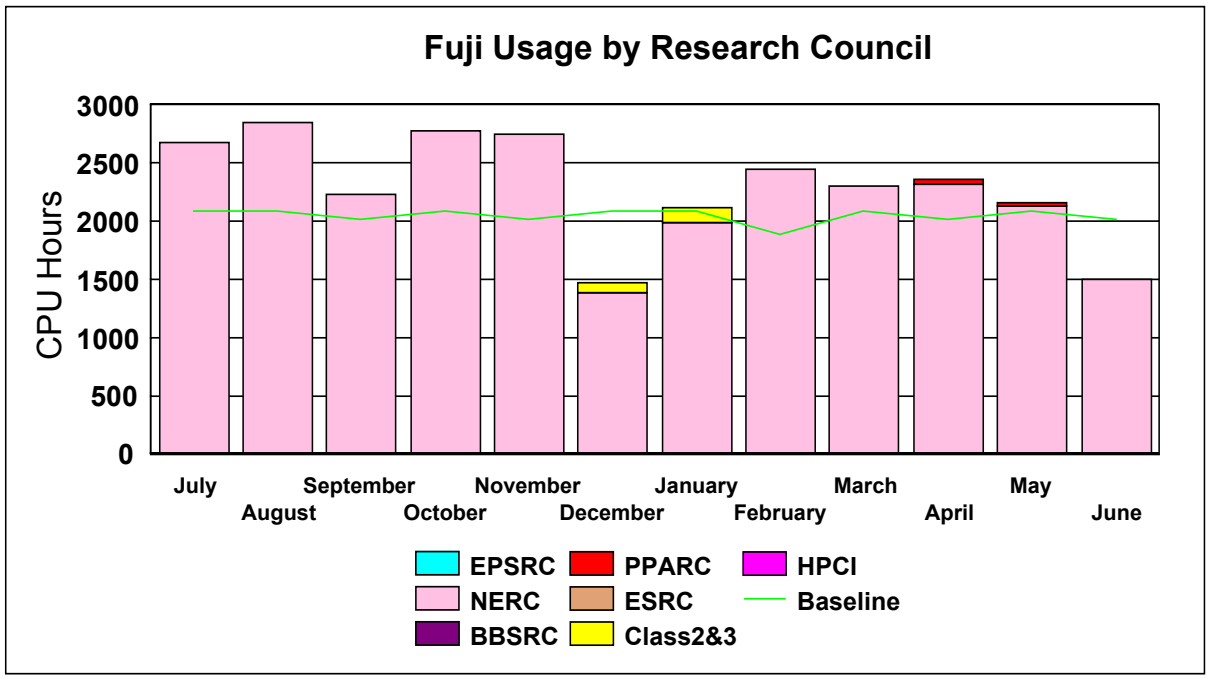


The graph above illustrates the historic allocation of the Medium Performance Disk on Fermat.

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 32 Terabytes. The primary usage is for NERC and, in total, this amounts to about 31 Terabytes of data storage at present.



2.2.2 Fujitsu System Usage Graph



The above graph shows the current CPU usage on the Fujitsu VPP 300 NERC system based at the University of Manchester. The Fujitsu usage exceeded the baseline in two of the three months of the quarter.

2.2.3 Guest System Usage

The Compaq ES40 cluster still shows no guest usage, and the NEC SX4 has now been withdrawn from service due to continued zero usage demand.

2.3 Service Status, Issues and Plans

Status

The proposed SAN implementation is now in its final stages, and is currently undergoing a stabilisation period before CfS transfers the user file systems from current Fermat/Green disks.

Issues

Wait times can still be significant at times due to the time demands on all the machines, however the additional Fermat and Green resources have substantially reduced the queuing times on Turing.

CfS still awaits a decision from NERC regarding the 12 Terabytes of additional data (the ECMWF dataset) that may require storage on the system.

Plans

Wren

A new Origin300, Wren, which has 16 500 MHz processors, will replace Fermat as the interactive system. It will be purely for interactive (and administrative) work, thus providing more processing power for interactive use. Interactive access to Fermat will be withdrawn.

Storage Area Network (SAN)

The SAN will provide a major enhancement to the disk provision and performance for Fermat, Green and Wren. The primary benefits will be improved access to home and /hold filestore from Green, thus removing problems experienced by some users with large I/O requirements. There will also be a globally accessible 'temp' area, files on which can be accessed from all Origin systems without requiring interactive access to the batch systems.

At present there is only one type of disk available for use on Fermat and Green - known as Medium Performance (MP) disk. With the SAN there will also be High Performance (HP) disk (as is currently available for Turing), Ultra High Performance (UHP) disk, which will be faster still - targeting intensive I/O applications, and High Volume (HV) disk, which will provide a cheaper storage facility. Neither UHP disk nor HV disk will be backed up. Existing data will remain on Medium Performance disk but you will subsequently be able to trade resources and use any of the 4 types of disk. In addition to the globally accessible 'temp', there will also be local temporary areas on both Fermat and Green.

Installation is progressing well and it is hoped to have the new facilities available during mid-August.

Load Sharing Facility (LSF)

LSF is currently being installed on Fermat, Green and the new Origin 300, Wren. This will provide better control for running parallel jobs than NQE and in particular should minimise the extent to which resources such as memory are shared between jobs, thus improving job performance particularly on Green.

LSF does use different commands, options and job submission scripts from NQE but the changes are generally straightforward and documentation will be provided to minimise any disruption during the transition period. However we urge anybody who may use NQE commands or scripts in an unusual way (such as embedding NQE script details within programs or other scripts), to contact us as soon as possible so that we can ensure that such cases are tested before LSF is made live.

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

11 applications for new CSAR projects were received in the second quarter of 2002.

3.2 New Projects

7 new CSAR projects were started.

3.3 Queries

Overall, 195 queries relating to the CSAR service were received by the Helpdesk between 1st April and 30th June 2002.

3.4 Generic Token Usage

Implementation of daily checking of project token usage with warning messages being sent to PIs and the Research Councils has taken place with the ability to stop project usage at 100% of token allocation utilisation.

3.5 Documentation

The CSAR web-site has been updated to reflect the status of the Technology Refresh.

3.6 Annual Report

The CfS Annual Report is currently being prepared for publication.

3.7 CSAR Focus

The Spring edition of CSAR Focus has been produced and distributed.

4 Scientific Application Support Services

4.1 Training and Education

The early 2002 course schedule has now finished. The type and content of courses are being reviewed and a new course schedule is being drawn up. The emphasis will be on making best use of the CSAR systems.

Developments are being made to the Advanced MPI course, to include one-sided communications..

In February courses in MPI and OpenMP were given at the National Supercomputing Centre in Linköping, Sweden. The two day course was well attended and subsequent feedback has been very positive

In addition, an "Introduction to HPC and the CSAR Service" course was given to British Antarctic Survey staff (in Cambridge) in March.

4.2 Consortia Support

Further support work is likely in the area of parallel linear algebra optimisation. Consortium cse086 have begun discussions to arrange appropriate support to replace a ScaLAPACK re-distribution subroutine. Earlier work for cse076 has been successfully completed and results were presented at CUG2002. This work has received attention from certain vendors and other research groups and the availability of funding to establish a project is being assessed.

Optimisation work has produced a 20% improvement for another consortium. Work is now concentrating on parallelisation. An optimisation assessment has also been carried out with a separate group of the same consortium.

Work on another code has resulted in it being run regularly on 256 PEs. Further optimisation is being performed.

4.3 Applications Software

4.3.1 Unified Model

We are installing the Unified Model version 5.3 on both Turing and Green. We are also looking at the ways in which this code can be run, in order to improve the flexibility available to users.

4.3.2 GAUSSIAN98

A new version of this software, Rev. 11.3, is on order and will be installed when it arrives.

4.3.3 LAMMPS

We are pursuing a licence for LAMMPS with a view to providing this on the CSAR machines.

4.3.4 ScaLAPACK

Version 1.7 of this is now available on Turing and Fermat.

5 Collaboration and Conferences

5.1 MRCCS Projects

5.1.1 Reality Grid

Recent work includes:

- the parallelisation of visualization software used to visualize the output of RealityGrid simulations and experiments.
- the development of a design for steering software within the RealityGrid framework.

5.1.2 Euro Grid

The EuroGrid project has continued with a project meeting in Bergen in June. The first version of the Eurogrid Resource Broker, produced by Manchester, will shortly be delivered to the project partners. This will allow users to describe the resources required for their job, and then get the broker to provide a list of sites where this job could run. The second version of the broker, due next year, will be able to estimate turnaround times, and broker jobs dynamically based on user-specified criteria, but without user interaction.

Manchester has continued to operate the Certification Authority (CA) for this project. The CA has been so successful that its scope of operation has now been extended to cover the aligned GRIP project.

5.1.3 Access Grid

Manchester's Access Grid node continues to grow in usage - the node now averages about 3 meetings a week (as opposed to 2 in the first quarter of this year). Major usage of the Access Grid is made by e-Science projects such as Geodise, myGrid, RealityGrid and GRIP.

We are expanding our involvement in the global Access Grid community and consolidating our UK leadership in the Access Grid project. One development in this area is that Michael Daw has been co-opted onto the SC Global 2003 committee in the role of Technical Director (SC Global is a distributed conference, held in conjunction with the SC conference, but over the Access Grid).

Many other developments to Access Grid work at Manchester are pending but not yet confirmed. The access grid has proven to be a valuable resource outside the normal seminar series. Several non-scheduled broadcasts have been attended on an adhoc basis as requested by members of the University.

The CSAR User Steering Group Meeting will be held via the Access Grid at both Manchester and Imperial College.

5.1.4 GRIP

GRIP, the Grid Interoperability Project is proceeding well. Manchester is coordinating publicity for this project, which includes running the website. As well as being involved in the recent project meeting at Pallas software's headquarters in Germany, Manchester has also helped organise regular technical meetings which are held on the AccessGrid, between FZ-Juelich (Germany), Manchester, Southampton, and Argonne National Laboratories (USA).

5.1.5 SAMD Project

Work is continuing on SAMD, the grid application to manipulate econometric datasets on HPC resources. (This is an ESRC funded e-science demonstrator project).

5.1.6 North West Centre for Advanced Virtual Prototyping

An existing member of staff, Lee Margetts, has been reappointed to fulfil the MRCCS contribution to this project. Several meetings have been attended in which the objectives of the role have been identified. Lee will be responsible for further developing the parallel finite element software that will be used for various

engineering case studies.

5.1.7 Finite Elements

A large finite element simulation of a 3D CFD problem was carried out using 256 processors on Green. The analysis was run to demonstrate the performance and scalability of a parallel Navier-Stokes solver developed as part of an MRCCS project. An article describing the work is to be presented in the next edition of the UKHEC newsletter. A UKHEC case study has been started on the visualisation of finite element analyses. A report will be released to the website in the future.

The parallelisation of a suite of serial finite element programs developed at the University of Manchester School of Engineering is progressing well. The performance and scalability of certain case studies has impressed engineers from other institutions. It is anticipated that these programs will encourage new users from engineering who currently rely on poorly scaling industrial packages in their research. The programs will also have direct application in the Virtual Prototyping project.

5.2 Events

Lee Margetts attended the 8th International Symposium for Numerical Methods in Geomechanics, Rome, 10-12th April, and the 10th meeting of the Association for Computational Mechanics in Engineering (ACME) in Swansea, 14-17th April. Conference papers describing the parallel finite element work were presented at each.

Adrian Tate attended VECPAR 2002 in Porto from June 24th-27th.

Lee Margetts, Jon Gibson, Adrian Tate (with Patrick Briddon from the University of Newcastle Upon Tyne) and Neil Stringfellow (with Peter Coveney from Queen Mary, University of London) presented papers at CUG.

A number of staff attended a course in Edinburgh on WebServices, in order to learn about new technologies such as XML, SOAP and WSDL, which are becoming increasingly important in the world of Grid Computing. Some of the staff also attended the "Getting OGSA Going" workshop. Open Grid Services Architecture - the extensions to Web Services on which the Globus team and IBM are basing the next version of the Globus Toolkit.

John Brooke and Jon MacLaren attended a one-day meeting at the Sheraton Skyline Hotel at Heathrow Airport called "Grids: Hype meets reality". The meeting was being organised by HP and Platform Computing. Among the speakers was Songnian Zhou, the CEO of Platform.

6 Added Value Services

6.1 International Conferences

CUG (Cray User Group) SUMMIT 2002 was held in Manchester on May 20th-24th, hosted by the University of Manchester. Kevin Roy was the Local Arrangements Organiser. CUG Summit is a major event providing a forum for both Cray and SGI. Presentations included a keynote speech by Dr John Taylor, Director-General of the UK research councils, and covered new developments from the vendors and experiences from the users. This international event was a great success.

The HPC/CSAR Team exhibited at the ISC (International Supercomputing Conference) 2002 in June at Heidelberg. We have suggested to the organisers that next year a link is provided via the Access Grid to Manchester, with the objective of encouraging much greater participation from the UK (there was very little representation from the UK).

We are also preparing for SC2002 in November at Baltimore. There will be an exhibition stand and we have submitted proposals for two tutorials - one in optimisation of linear algebra and the other in the use of the Cray MTA.

The CSAR Team & machines are likely to play an active role in the forthcoming 5th Global grid Forum.

6.2 Cray MTA Optimisation

The work on optimising genomics codes on the Cray MTA-2 was presented at the CUG Summit. The project finished at the end of June.

6.3 Seminars

The Summer programme of seminars has had subjects relating to HPC application areas, Grid computing and clusters. The programme continues to bring in speakers from around the UK and worldwide. Most of the seminars are offered to a worldwide audience via the Access Grid and the speakers often utilise this technology to save time, energy and expense by presenting remotely.

6.4 The VIP Laboratory

Recent visits to the VIP Laboratory have been made by CRAY, IEE, and the University of Leeds. The VIP Laboratory has also been visited as part of numerous events including the ERASMUS Intensive Programme, e-Science North West Opening, the Science and Engineering Awareness Day and the ISCO Conference.