

CSAR Service - Management Report

November 1998

This is the first of the monthly reports prepared by the Computation for Science consortium (CfS) since the CSAR Service commenced production service on 16th November 1998.

A more comprehensive report will be provided quarterly, which will additionally cover wider aspects of the Service such as information on Training, Application Support and Value-Added services. The first quarterly report will be published in February 1999.

These and other such reports will be made available through the Web to authorised staff within EPSRC and the other Research Councils, to CfS staff and CSAR Service users. The reports will be indexed in a similar way to which other useful information and news is listed for selection.

1. Introduction

The CSAR Service was accepted on 13th of November and users were for a few days restricted to interactive use in order for CfS to make the necessary final adjustments for the Service to be ready for production usage.

In recognition of the importance of the work, CfS organised for residual batch work queued by the Early Users to still be processed, so the results could be copied into the new User areas when allocated.

The production service commenced on Monday 16th November. During the first week, in particular, there were many set-up tasks undertaken:

- First the initial Projects were set up, and the Principal Investigators (PIs) were identified to the CSAR Service
- PIs were given allocations of Generic Service Tokens by their respective Research Councils, and traded their Tokens for resource allocations on the Service (e.g. T3E processing hours, Disk GByte years, etc.)
- Users in each Project registered
- PIs authorised each User and gave them Disk and HSM/tape quotas
- Users set up their user areas, transferred files (programs and data) and commenced usage, both interactively and with batch work.

This will be an ongoing process for new Projects and Users.

This document gives information on Service Quality and on actual usage of the CSAR Service during the reporting period of November 1998. The information, in particular, covers the availability and usage of the following two main CSAR Service High Performance Computing (HPC) systems:

- Cray T3E-1200E/576 (Turing)
- SGI Origin2000/16 (Fermat).

The information is provided in both textual and graphical form, so that it is easier to see trends and variances.

2. 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.

2.1 CPARS

Table 1 gives the measures 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
Help Desk						
Non In-depth Queries - Maximum Time to resolve 50% of all queries (working days)	< 1/4	< 1/2	< 1	< 2	< 4	4 or more
Non In-depth Queries - Maximum Time to resolve 95% of all queries (working days)	< 1/2	< 1	< 2	< 3	< 5	5 or more
Administrative Queries - Maximum Time to resolve 95% of all queries (working days)	< 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 in month (working days)	< 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 scheduled maintenance sessions taken per system in the month	0	1	2	3	4	otherwise

Table 1

Table 2 gives actual performance information for the period of November 16th to November 30th inclusive. Overall, the CPARS Performance Achievement was satisfactory (Green).

CSAR Service - Service Quality Report - Actual Performance Achievement

Service Quality Measure	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
	HPC Services Availability							
Availability in Core Time (% of time)							99.99%	
Availability out of Core Time (% of time)							98.53%	
Number of Failures in month							2	
Mean Time between failures in 52 week rolling period (hours)							400	
Help Desk								
Non In-depth Queries - Maximum Time to resolve 50% of all queries (working days)							1	
Non In-depth Queries - Maximum Time to resolve 95% of all queries (working days)							4	
Administrative Queries - Maximum Time to resolve 95% of all queries (working days)							2	
Help Desk Telephone - % of calls answered within 2 minutes							100%	
Others								
Normal Media Exchange Requests - average response time in month (working days)							0.5	
New User Registration Time (working days)							2	
Management Report Delivery Times (working days)							10	
System Maintenance - no. of scheduled maintenance sessions taken per system in the month							4	

Table 2

Notes:

1. Since the CSAR Service was operational for only half a month from the commencement date of November 16th, the actual number of failures and number of maintenance session values have been doubled.
2. HPC Service Availability has been calculated using the following formulae, based on the relative NPB performance of Turing and Fermat:

$$[\text{Turing availability \%} \times 122 / (122 + 3.5)] + [\text{Fermat availability \%} \times 3.5 / (122 + 3.5)]$$

Table 3 gives Service Credit values for the November period, though they will be accounted on a quarterly basis. These values are calculated according to agreed Service Credit Ratings and Weightings.

CSAR Service - Service Quality Report - Service Credits

Service Quality Measure								
	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
HPC Services Availability								
Availability in Core Time (% of time)							-0.015	
Availability out of Core Time (% of time)							0.000	
Number of Failures in month							0.000	
Mean Time between failures in 52 week rolling period (hours)							0.000	
Help Desk								
Non In-depth Queries - Maximum Time to resolve 50% of all queries (working days)							0.000	
Non In-depth Queries - Maximum Time to resolve 95% of all queries (working days)							0.008	
Administrative Queries - Maximum Time to resolve 95% of all queries (working days)							0.000	
Help Desk Telephone - % of calls answered within 2 minutes							-0.001	
Others								
Normal Media Exchange Requests - average response time in month (working days)							0.000	
New User Registration Time (working days)							0.000	
Management Report Delivery Times (working days)							0.000	
System Maintenance - no. of scheduled maintenance sessions taken per system in the month							0.002	
Monthly Total & overall Service Quality Rating for each period:	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

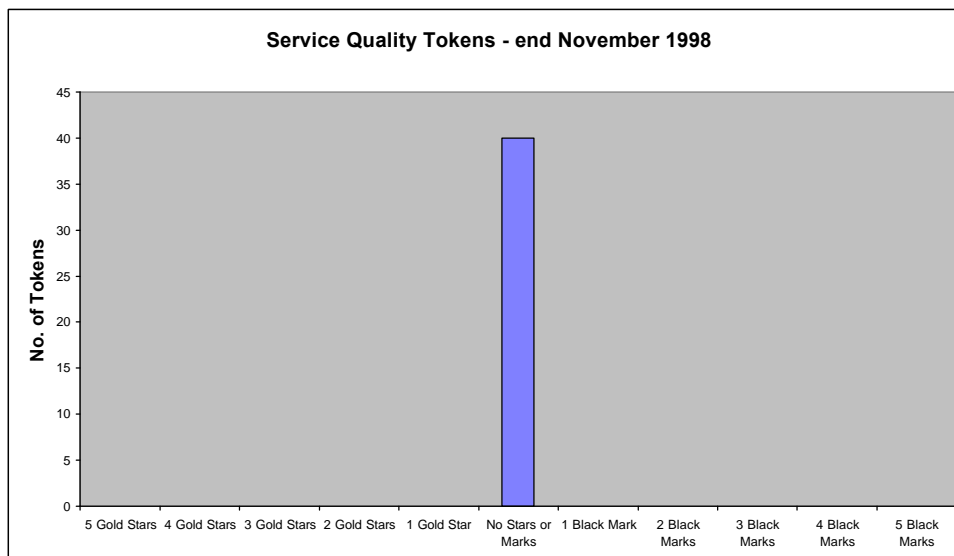
Table 3

2.2 Service Quality Tokens

The current position at the end of November 1998 was that none of the 40 registered users of the CSAR Service had used any Service Quality Tokens. See below:

	Current Position - end November '98											
	Jan.	Feb.	March	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
5 Gold Stars												0
4 Gold Stars												0
3 Gold Stars												0
2 Gold Stars												0
1 Gold Star												0
No Stars or Marks												40
1 Black Mark												0
2 Black Marks												0
3 Black Marks												0
4 Black Marks												0
5 Black Marks												0
Total No. of Users												40

In the form of a bar chart, the current statistics are:



Users are kindly requested to use the Service Quality Tokens facility more regularly, so that the CfS Management Board can better ascertain how the CSAR Service is operating and how well the CfS staff are working in support from an end user perspective.

[Note: Whilst strictly speaking an input to the report for December, CfS was very pleased to see that the first entry of Service Quality Tokens was **** (four stars) from Patrick Briddon, who commented: "The consortium cse006 is extremely pleased with the throughput of work on the new service and we are pleased with the extremely rapid response to straightforward service queries." This response encourages CfS to keep up its good work and standards of practice.]

2.3 Throughput Target Against Baseline

This was an incomplete month and the Baseline Throughput targets were met when there was sufficient workload queued. It should be noted that the capacity usage figures below exclude usage within October and in the period of November 1st to 15th 1998; which amounted to 466,554 T3E PE Hours.

Job Throughput Against Baseline CSAR Service Provision

Period: 16th - 30th November 1998 (15 days)

	Baseline Capacity for Period (T3E PE Hours)	Actual Usage in Period (T3E PE Hours)	Actual % Utilisation c/w Baseline during Period
1. Has CfS failed to deliver Baseline MPP Computing Capacity for EPSRC?	175,066	99,684	56.94%
2. Have Users submitted work demanding > 110% of the Baseline during period?	175,066	101,025	No
3. Are there User Jobs outstanding at the end of the period over 4 days old?		Number of Jobs at least 4 days old at end Period 0	Number of Jobs at least 4 days old at end Period is not zero (Yes/No)? No
4. Have Users submitted work demanding > 110% of the Baseline during period?		Minimum Job Time Demands as % of Baseline during Period 0%	Minimum Job Time Demand above 90% of Baseline during Period (Yes/No)? No
5. Majority of Job Queues contained runnable jobs from Users for more than 97% during period?	4	Average % of time each queue contained jobs in the Period 57.7%	Average % of time each queue contained jobs in the Period is > 97%? No

Baseline Fee reduction due (%) if the answer to item 1 is < 100% and the answers to items 2 - 5 inclusive are "Yes":

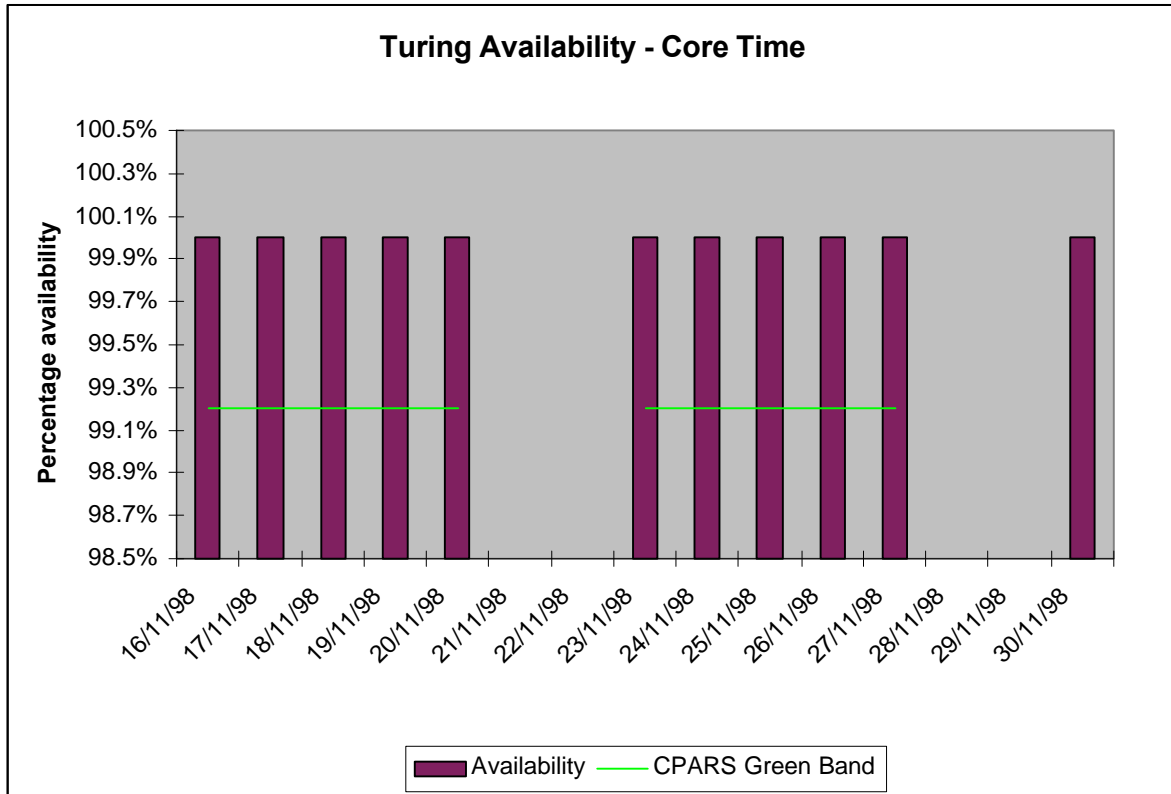
0.00%

3. System Availability

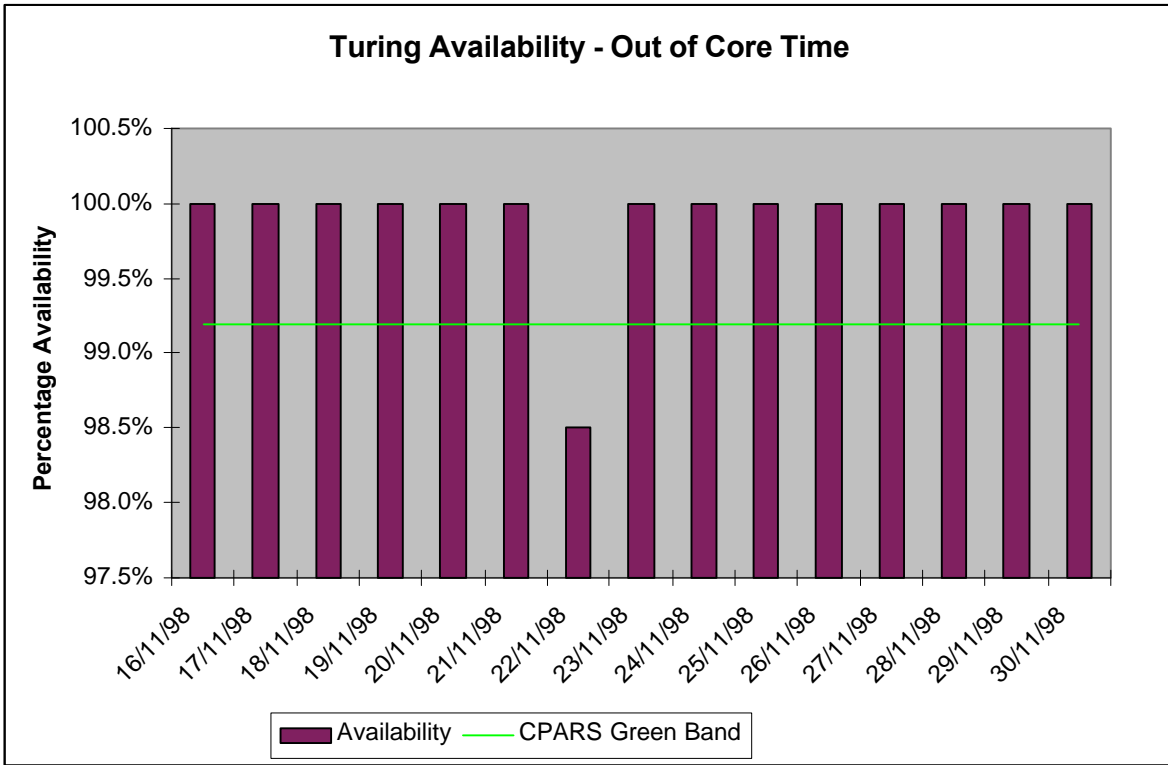
Service availability each reporting period is calculated as a percentage of actual availability time over theoretical maximum time, after accounting for planned breaks in service for preventative maintenance.

3.1 Cray T3E-1200E System (Turing)

The following graphs show the availability of Turing both in core time and out of core time respectively during the period of 16th to 30th November. There was unplanned unavailability out of core time on November 22nd due to the failure of a system PE.



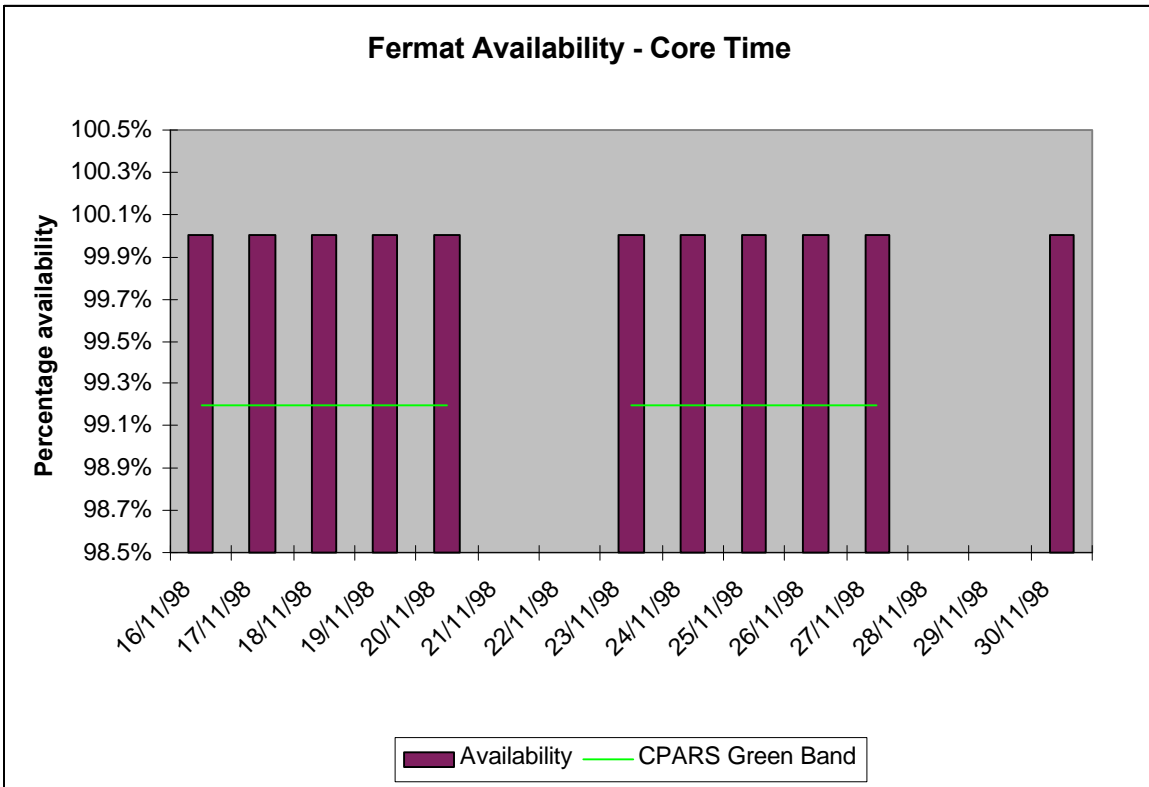
Availability of Turing in core time during November from the Acceptance Date was very good.



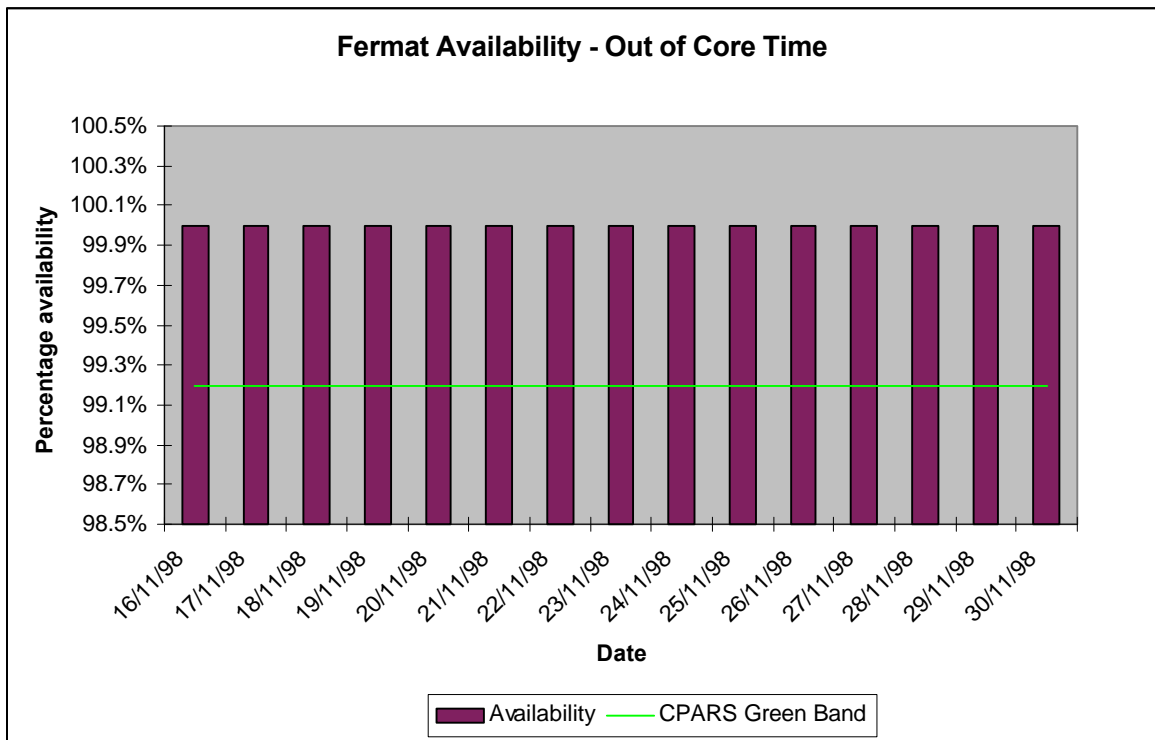
Availability of Turing out of core time during November from the Acceptance Date was satisfactory.

3.2 SGi Origin2000 System (Fermat)

The following graphs show the availability of Fermat both in core time and out of core time respectively. Fermat was continuously available as planned during November.



Availability of Fermat in core time during November from the Acceptance Date was excellent.



Availability of Fermat out of core time during November from the Acceptance Date was excellent.

4. HPC Services Usage

Usage information is given in tabular form, in Appendices, and in graphical format. The system usage information for the period of November 16th to 30th is provided by Project/User Group, totalled by Research Council and overall. This covers:

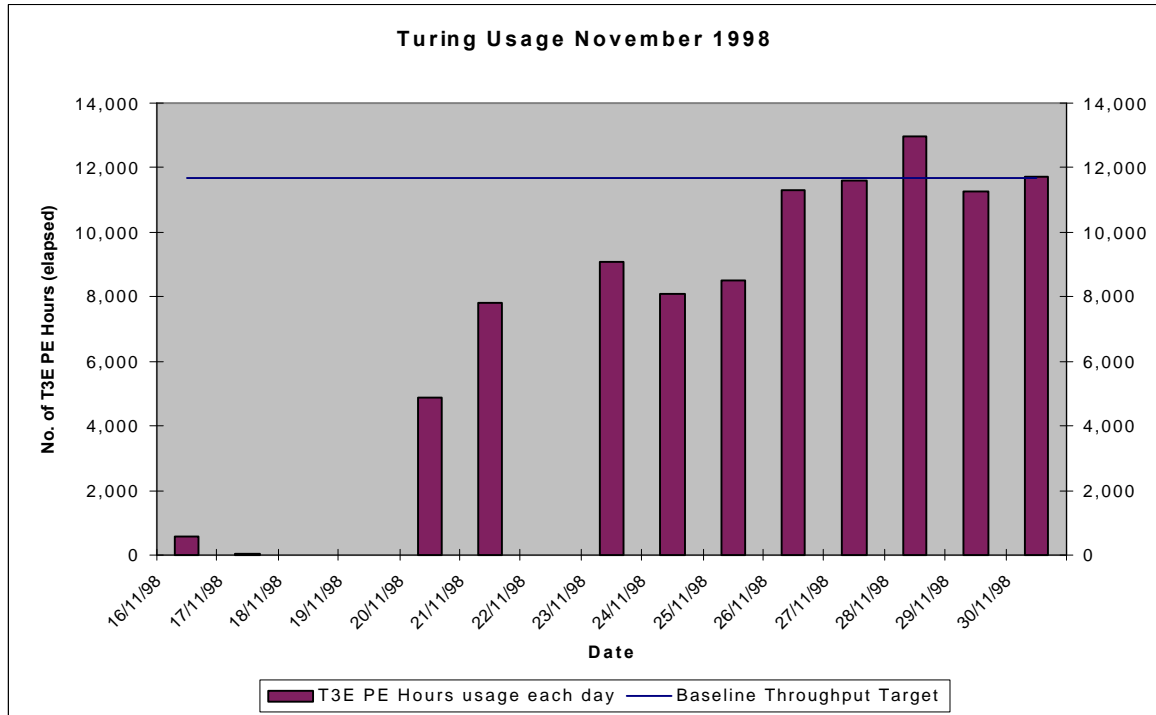
- CPU usage Turing: 99,684 PE Hours Fermat: 853 CPU Hours
 - User Disk allocation Turing: 15.23 GB Years Fermat: 8.56 GB Years
 - HSM/tape usage 0.05 GB Years
- (for the UGAMP project, currently held in UoM & CSC support areas).

Since this is the first month of Service, usage is graphically shown for each day in November from the Acceptance Date. In subsequent reports, the following graphs will be provided to illustrate usage per month:

- 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 (103 GFLOP-Years) will be 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 (3.5 GFLOP-Years) will be 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) will be 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) will be 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.

4.1 Cray T3E-1200E System (Turing)

The following graph shows the usage of Turing during each day from the Acceptance Date. Note that there will be some variance on a day-to-day basis as the accounts record job times, and thus CPU usage figures, at the time of job completion – which could be the second actual day for large jobs. At present, there is a 12 hour limit on jobs, so that they are check-pointed and computational time lost due to any failure is well managed.

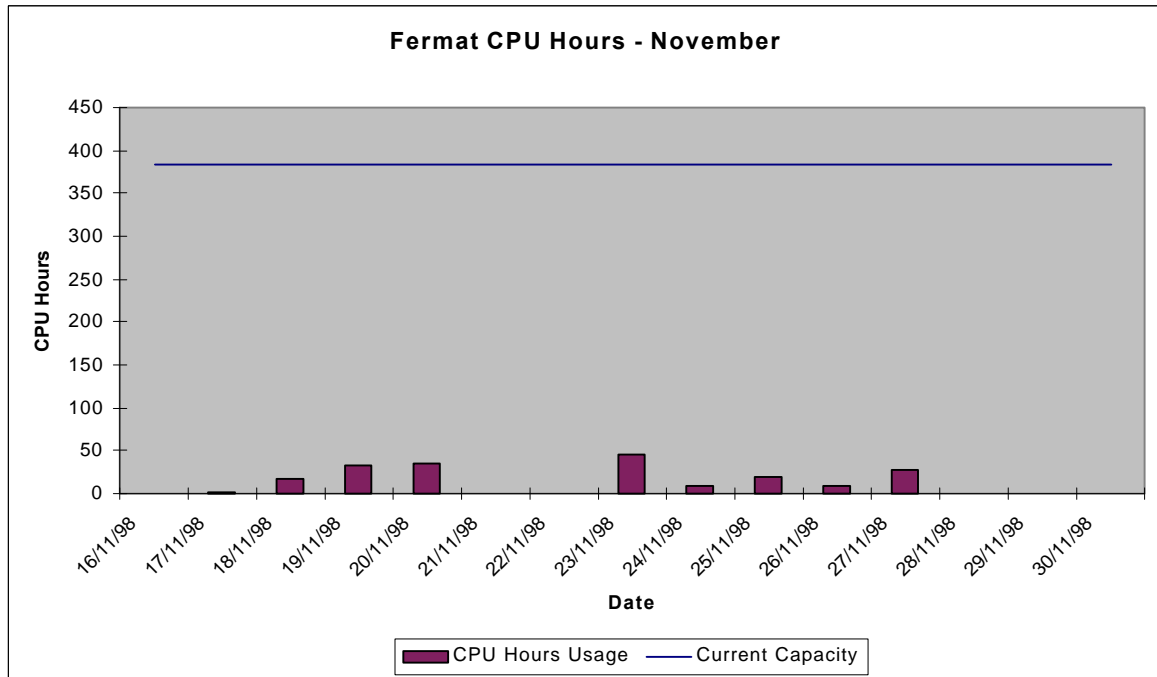


It can be seen that the workload on Turing has steadily increased throughout the period as more users were authorised through the registration system by their PI's and began to load batch work. When fully loaded the Baseline Throughput requirement is being achieved on average ($24 \times 103 \times 576 / 122$ PE Hours each day).

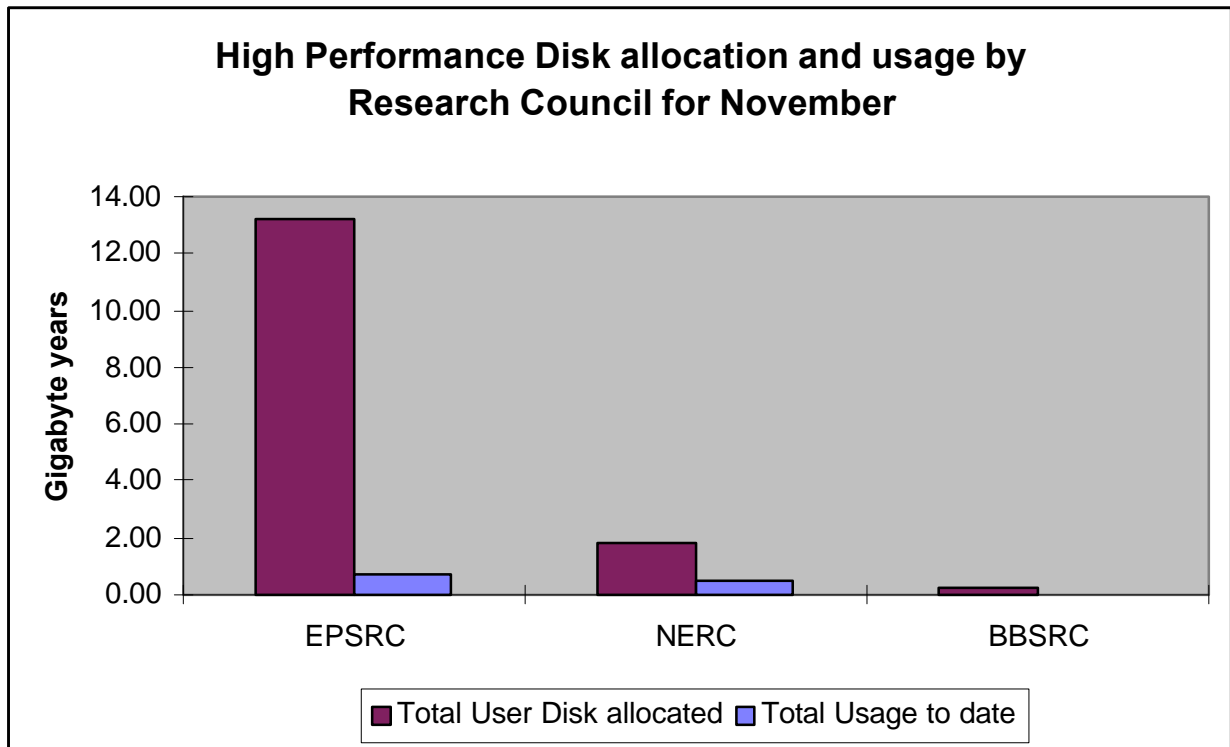
Fine tuning of the CfS scheduling system will continue to ensure minimal wasting of PE resource, to best fit in numbers of different sized jobs (e.g. 32, 64, 128, 256) to maximise throughput.

4.2 SGI Origin2000 System (Fermat)

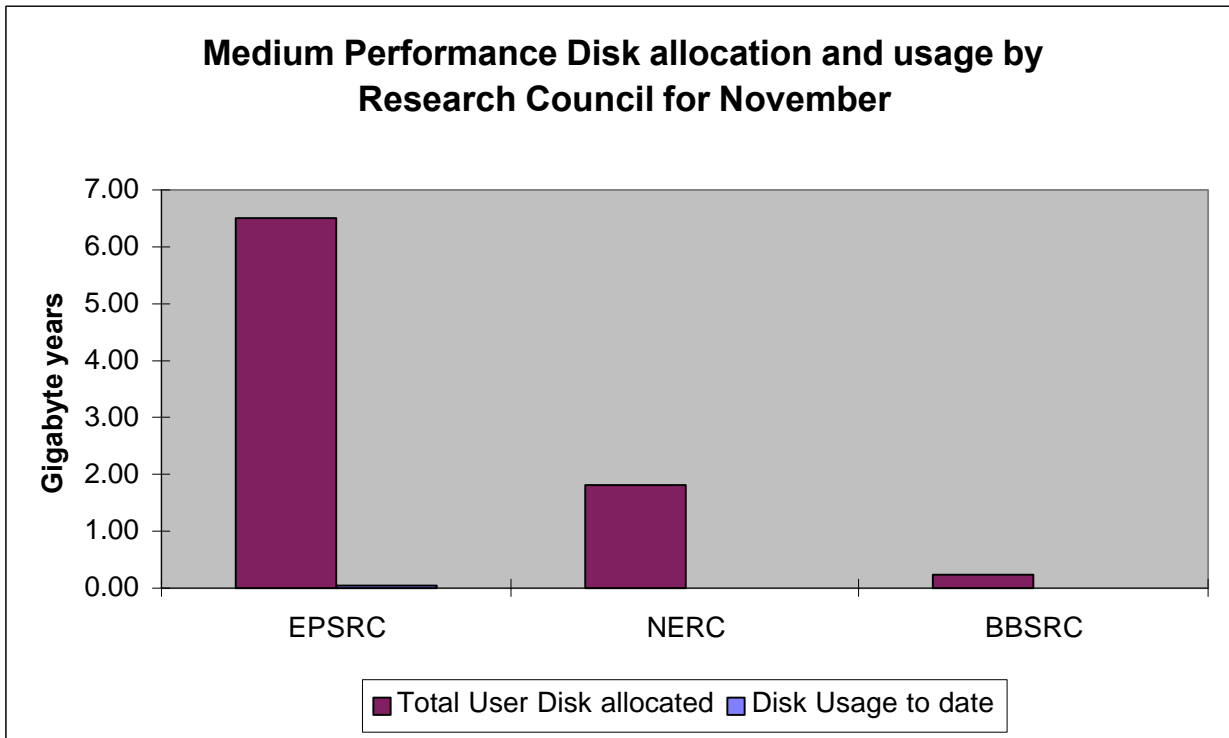
The graph below shows that Origin 2000 CPU utilisation was relatively low during November. This is essentially because Fermat was been primarily used as a file server and pre/post processor during this period, rather than as a computational resource.



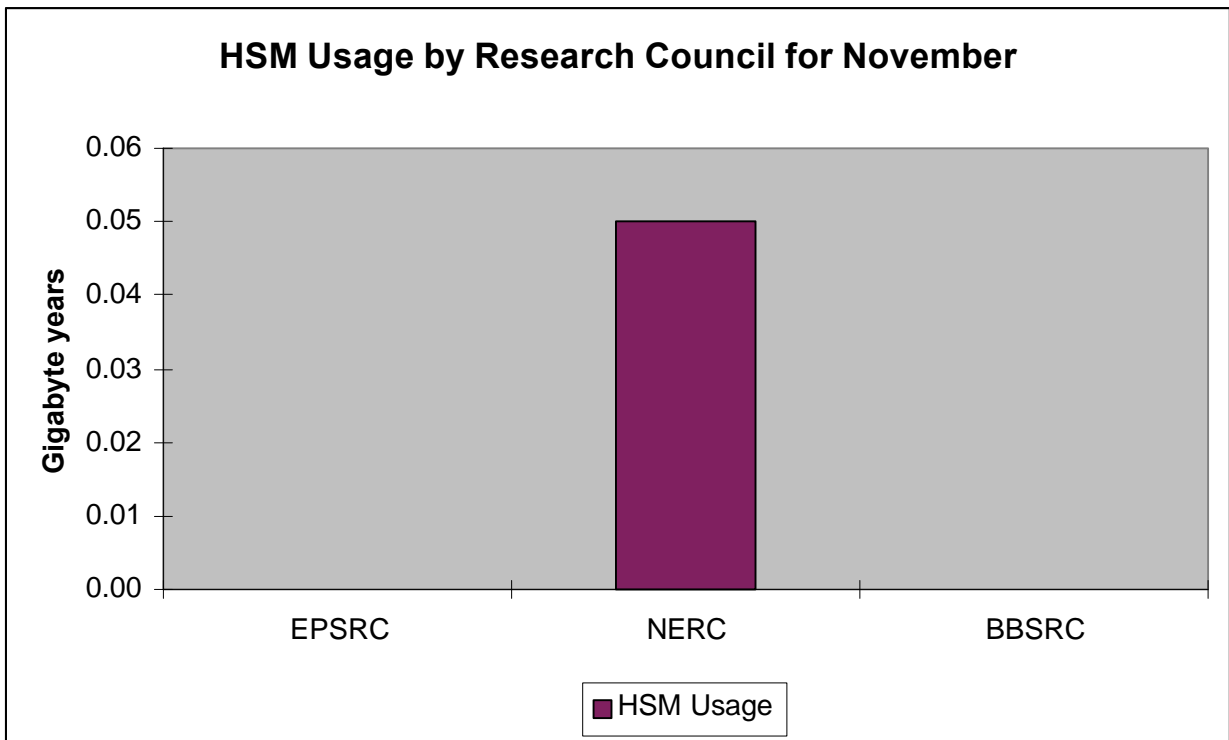
The graph below illustrates the current allocations and usage of the High Performance Disk on Turing. Usage against the present allocations is low at present, as expected after two weeks of production service only.



The graph below illustrates the current allocations and usage of the Medium Performance Disk on Fermat. Usage against the present allocations is low at present, as expected after two weeks of production service only.



The graph below shows the current HSM usage by Research Council funded projects. The primary usage is for NERC, however the volumes are at present recorded under UoM and CSC system support accounts whilst the data is transitioned to new User accounts.



5. Service Status, Issues and Plans

5.1 Status

During the first week of the Service, a number of difficulties were encountered in registering to use the system. Each user had to go up his/her "learning curve" in the process and use of the various facilities; and some problems were caused due to unexpected sequences of events, e.g. Users registering before their PI. A few faults/problems, which had not been previously identified, were found:

- a) A system parameter on Turing had to be increased in order to permit many more Projects to have Disk Quotas. This required the Turing system to be re-booted for the change to take effect. A Preventative Maintenance Session was taken on the evening of 17/11/98 to achieve this, as no batch work then queued. Once this change had been made, the users/Pis were then able to continue the registration and resource allocation process, and could use the system.
- b) Some job limit parameters were not being set on user registration, which initially prevented jobs running on the queues. This was corrected with immediate effect.
- c) The advisory facility used to assist the PI in determining the number of Generic Service Tokens for his/her Project did not correctly calculate the number of Tokens for some of the resource types. This was corrected as soon as the fault was spotted.
- d) Disk quotas were not initially being set correctly through the registration system. The system was fixed as soon as the fault was reported, and previously entered data was corrected.

In addition, there was initially some confusion over the difference between Disk Allocation in GBytes at the Project level, Disk Quotas in GBytes per User and Disk Resource Trading by the PI for his/her Project to obtain Disk capacity in GB Years. The size of the Disk Allocation dictates how rapidly the Service Tokens traded for Disk capacity will be used up over time. Disk Quotas limit the actual volumes of disk space that can be used by any individual user at any time.

Over and above the faults, there were additional requirements/system changes raised as a result of user perceived problems. CfS implemented the additional facilities within a short time of each request, when agreed of general benefit. This included the following:

1. A facility was added so that a PI could allocate Disk capacity/quotas independently of the capacity obtained through trading. This was in recognition that an allocation could be for a Project where the Tokens for Disk space is to run out in, for example, 3 months or last for 3 years.
2. Provision was added for UGAMP Users to be able to change the Owner of individual files, to avoid copying/moving the data.

CfS staff gave close support to each PI/User and, in particular, helped during the Trading Process and in re-setting Disk/Tape Allocations where a PI required more capacity for a shorter time, or less actual capacity than originally set up for his/her Project.

Finally, another Preventative Maintenance session was taken on 24/11/98 in order to implement a number of performance improvements for NFS. This gave about a ten-fold improvement in the speed of NFS mounted file transfers between Turing and Fermat through the HIPPI channels. A new, faster and more comprehensive version of XCP was implemented, which now typically delivers 25-30 Mbytes per second transfer between machines. In addition, a \$TMPTOHOLD facility was added so that large results files can be held in temporary file space for 24 hours after PEs are released from each job; for subsequent usage and copying to more permanent storage areas.

5.2 Issues

There are some applications, such as Gaussian, which require T3E PEs to be allocated to each job, but do not use varying numbers throughout each job. This is wasteful of resource; but there is no easy answer excepting that if the application was run on fermat, the Origin2000 CPUs could be readily shared with much less wastage of processing cycles.

The Turing system is beginning to be overloaded with queued work and some scheduling improvements are in hand to better balance the throughput of large jobs (256 PEs and above) against smaller ones.

5.3 Plans

A Security Audit is planned for 8th December.

A Preventative Maintenance Session is planned for Tuesday 15th December, to replace the present failed PE which is currently swapped out through re-mapping of the Turing configuration.

Initial facilities are planned to be released to enable PIs to undertake basic Capacity Planning, so that their 3 year views on future resource usage can be consolidated.

6. Conclusion

Good progress has been made overall.

Continued management attention will be given to maximise the throughput of the Service, whilst balancing as fairly as practicable the shares between Projects and jobs of the varying sizes.

Appendix 1

Usage report for Research Council Projects

From Monday 16-Nov-98 to Monday 30-Nov-98

Account		----- CPU Usage (Hours) -----					--- Storage (GB-Years) ---		
		Inter	Priority	Normal	Low	Total	D-Usage	D-Allocn	HSM
CSE001 Admin users	turing	-	-	-	-	-	0.00	-	-
	fermat	-	-	-	-	-	0.00	-	-
Total for Subject									
EPSRC Administration	turing	-	-	-	-	-	0.00	-	-
	fermat	-	-	-	-	-	0.00	-	-
CSE002 gr/m01753 Gillan	turing	19.30	-	3288.14	2466.34	5773.78	0.08	3.14	-
	fermat	640.97	-	-	-	640.97	0.01	3.14	0.00
CSE003 gr/m01784 Taylor	turing	13.66	-	46449.10	-	46462.77	0.11	1.25	-
	fermat	39.13	-	-	-	39.13	0.00	1.25	-
CSE006 gr/m05201 Briddon	turing	343.45	-	17441.46	-	17784.92	0.01	1.25	-
	fermat	0.47	-	-	-	0.47	0.00	0.12	-
CSE007 gr m05348 Foulkes	turing	1.08	-	5338.95	-	5340.03	0.01	0.28	-
	fermat	-	-	-	-	-	0.00	0.28	-
Total for Subject									
Physics	turing	377.50	-	72517.65	2466.34	75361.50	0.21	5.92	-
	fermat	680.57	-	-	-	680.57	0.01	4.79	0.00
CSE004 gr/m08424 Sandham	turing	13.48	-	9876.55	-	9890.03	0.37	1.25	-
	fermat	0.37	-	-	-	0.37	0.03	1.25	0.00
CSE010 gr/104108 Williams	turing	0.00	-	-	-	0.00	0.00	0.04	-
	fermat	-	-	-	-	-	0.00	0.04	-
CSE011 gr/k52317 Williams	turing	0.07	-	310.58	-	310.65	0.01	2.68	-
	fermat	-	-	-	-	-	0.00	0.03	-
CSE016 GR/K96519 Cant	turing	-	-	-	-	-	0.00	0.01	-
	fermat	-	-	-	-	-	0.00	0.01	-
cse021 GR/L95427 Staunton	turing	-	-	-	-	-	0.00	0.03	-
	fermat	-	-	-	-	-	0.00	0.03	-
Total for Subject									
Engineering	turing	13.56	-	10187.13	-	10200.68	0.38	4.00	-
	fermat	0.37	-	-	-	0.37	0.03	1.35	0.00
CSE008 GR/M07624 Hillier	turing	2.21	-	2809.68	-	2811.90	0.00	0.96	-
	fermat	-	-	-	-	-	0.00	0.00	-
CSE009 gr/m07441 Catlow	turing	81.87	-	10292.85	-	10374.72	0.16	2.27	-
	fermat	1.28	-	-	-	1.28	0.00	0.36	-
cse017 GR/L58699 Luo	turing	-	-	-	-	-	0.00	0.01	-
	fermat	-	-	-	-	-	0.00	0.01	-
cse018 GR/L68353 Cant	turing	-	-	-	-	-	0.00	0.01	-
	fermat	-	-	-	-	-	0.00	0.01	-
Total for Subject									
Chemistry	turing	84.08	-	13102.54	-	13186.62	0.16	3.25	-
	fermat	1.28	-	-	-	1.28	0.00	0.38	-
CSE019 cr/173104 Berzins	turing	0.00	-	-	-	0.00	0.00	-	-
	fermat	-	-	-	-	-	0.00	-	-
Total for Subject									
Information Technology	turing	0.00	-	-	-	0.00	0.00	-	-
	fermat	-	-	-	-	-	0.00	-	-
Total for Council									
EPSRC	turing	475.14	-	95807.32	2466.34	98748.80	0.75	13.17	-
	fermat	682.22	-	-	-	682.22	0.03	6.51	0.00

Account		----- CPU Usage (Hours) -----				--- Storage (GB-Years) ---			HSM
		Inter	Priority	Normal	Low	Total	D-Usage	D-Allocn	
SOC Core Strategic Webb	turing	0.33	-	52.22	-	52.55	0.47	1.33	-
	fermat	-	-	-	-	-	0.00	1.33	-
csn003 UGAMP O'Neill	turing	0.02	-	-	-	0.02	0.00	0.19	-
	fermat	0.03	-	-	-	0.03	0.00	0.19	-
csn006 GR9/3550 Price	turing	-	-	-	-	-	0.00	0.30	-
	fermat	-	-	-	-	-	-	0.30	-
Total for Council									
NERC	turing	0.35	-	52.22	-	52.57	0.48	1.82	-
	fermat	0.03	-	-	-	0.03	0.00	1.82	-
CSB001 27/B07117 Goodfello	turing	0.00	-	-	-	0.00	0.00	0.23	-
	fermat	-	-	-	-	-	0.00	0.23	-
Total for Council									
BBSRC	turing	0.00	-	-	-	0.00	0.00	0.23	-
	fermat	-	-	-	-	-	0.00	0.23	-
euukcp	turing	173.89	-	-	-	173.89	0.87	0.00	-
	fermat	-	-	-	-	-	-	-	-
eugamp	turing	0.36	-	0.41	-	0.76	0.10	0.00	-
	fermat	-	-	-	-	-	-	-	-
euqub	turing	0.06	-	-	-	0.06	0.29	0.00	-
	fermat	171.07	-	-	-	171.07	-	-	-
euocam	turing	0.08	-	-	-	0.08	0.58	0.00	-
	fermat	-	-	-	-	-	-	-	-
euqmw	turing	37.38	29.68	-	-	67.06	3.79	0.00	-
	fermat	-	-	-	-	-	-	-	-
euhpci	turing	0.43	-	-	-	0.43	0.07	0.00	-
	fermat	-	-	-	-	-	-	-	-
euston	turing	0.00	-	-	-	0.00	0.02	0.00	-
	fermat	-	-	-	-	-	-	-	-
eural	turing	640.81	-	-	-	640.81	1.35	0.00	-
	fermat	-	-	-	-	-	-	-	-
eubbk	turing	-	-	-	-	-	0.02	0.00	-
	fermat	-	-	-	-	-	-	-	-
earlyu	turing	-	-	-	-	-	-	-	-
	fermat	0.02	-	-	-	0.02	0.08	-	0.00
dummy	turing	-	-	-	-	-	0.00	-	-
	fermat	-	-	-	-	-	0.00	-	-
Total for Subject									
eu accounts	turing	853.00	29.68	0.41	-	883.09	7.10	0.01	-
	fermat	171.08	-	-	-	171.08	0.08	-	0.00
Total for Council									
Research	turing	853.00	29.68	0.41	-	883.09	7.10	0.01	-
	fermat	171.08	-	-	-	171.08	0.08	-	0.00
Research Councils	turing	1328.50	29.68	95859.95	2466.34	99684.47	8.32	15.23	-
	fermat	853.33	-	-	-	853.33	0.12	8.56	0.00

 ----- End of Report Thursday 10-Dec-98 15:37:02 -----
 ----- 16-nov 30 /out=nov /user=all/mach -----