HPCx Quarterly Report
April – June 2007

1 Introduction

This report covers the period from 1 April 2007 at 0800 to 1 July 2007 at 0800.

The next section summarises the main points of the service for this quarter. Section 3 gives details of the usage of the service, including failures, serviceability, CPU usage, helpdesk statistics and service quality tokens. A summary table of the key performance metrics is given in the final section. The Appendices define the incident severity levels and list the current HPCx projects.

2 Executive Summary

- This was another exceptional quarter for reliability. Indeed, this was the first quarter without a single failure.
- Utilisation has been generally high, despite a dip in May. Overall utilisation of the main service for the quarter was nearly 74%. It is notable that the top five projects used 65% of the time, the remaining 35% being divided between 39 projects.
- Utilisation for the development service continues to be extraordinarily high, at nearly 93%.
- Dr Rob Baxter of EPCC gave a presentation at the 2007 Edinburgh International Science Festival, at which HPCx and HECToR were promoted. More than 80 people attended this and provided positive feedback.
- We remain on target both for technical reports and training. Plans are in place for the second technical workshop, a visualisation workshop, an Access Grid user group meeting and the Annual seminar.
- The Aeronautical Journal of March 2007 was a special issue on the work of the UK Applied Aerodynamics consortium. It included seven technical papers, all of which cited HPCx.
- Members of the Terascaling team have visited the n01 (Large-scale Long-term Ocean Circulation) and e05 (Materials Chemistry) consortia.
- The Terascaling team has broadened their activities by commencing work on the OCCAM, Quickstep and STAR-CD applications. This work ranges from

installation to major optimisations, and good progress has already been made.

- Exercises to compare the performance of applications codes on HPCx, Cray XT3 and Blue Gene continue.
- As part of the Software Engineering team's activities in HPC tools and languages, Dr Mark Bull represents the UK user community on the OpenMP Architecture Review Board and chairs their Language Committee. He has recently been coordinating the specifications of OpenMP version 3.0 which will soon be available for public comment.
- New helpdesk support software, implemented as part of the administrative software suite SAFE, was rolled out in May. This has been successful and, based on the subsequent experience, we have made a number of enhancements for both helpdesk staff and users.
- The Fortran and C compilers and their run-time environments have been successfully upgraded during this quarter.
- HPCx participated in the recent EPSRC workshop discussing the future of HPCx once HECToR is in service and looking at ways to provide complementary capability services which maximise the benefits to the user community.

3 Usage Statistics

3.1 Availability

3.1.1 Failures

The monthly numbers of incidents and failures (SEV 1 incidents) are shown in the table below:

	April	*May*	*June*
Incidents	3	9	1
Failures	0	0	0

Thus, there were no failures this quarter.

3.1.2 Performance Statistics

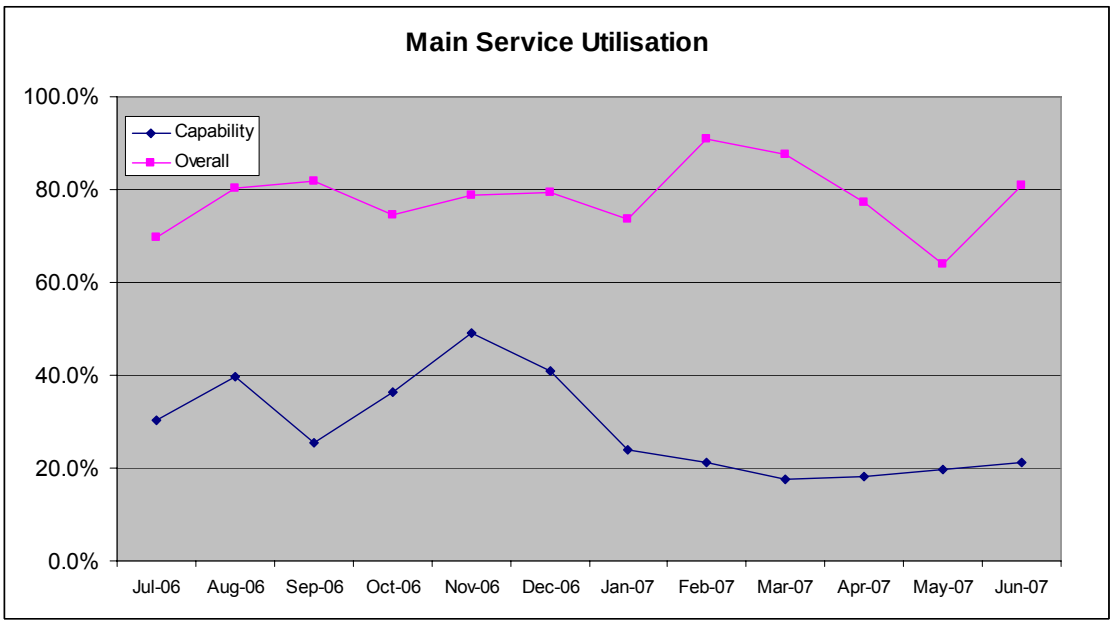
This section uses the definitions agreed in Schedule 7, ie,

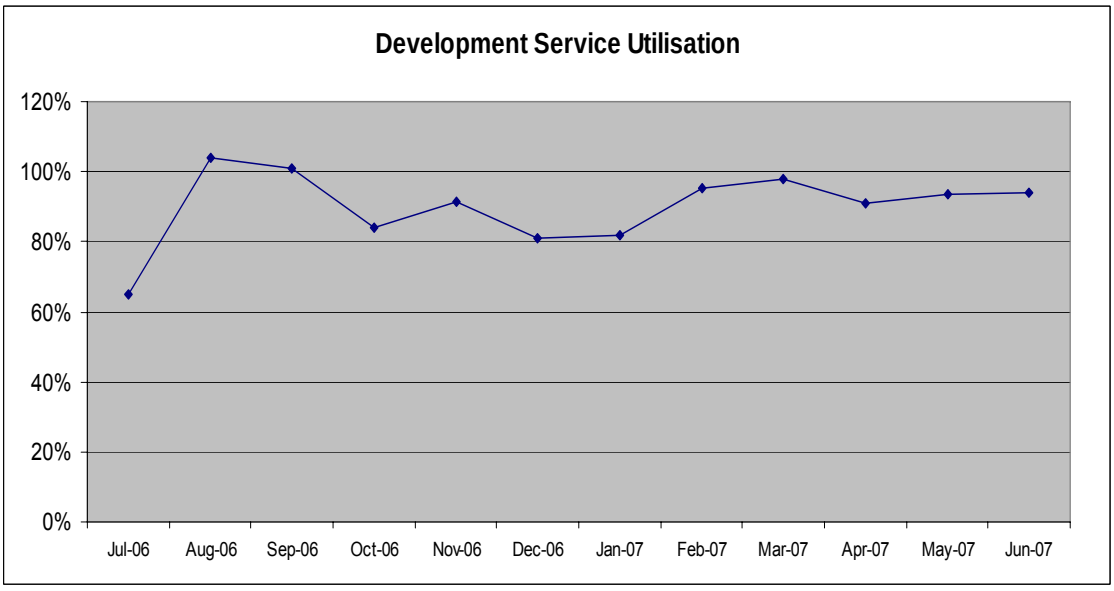
- MTBF = (24 x 30.5)/(number of failures in month)
- Serviceability (%) = 100 x (WCT – SDT – UDT) / (WCT– SDT)

Attribution	*Metric*	*April*	*May*	*June*	*Quarterly*
IBM	Failures	0	0	0	0
	MTBF	∞	∞	∞	∞
	Serviceability	100.0%	100.0%	100.0%	100.0%
Site	Failures	0	0	0	0
	MTBF	∞	∞	∞	∞
	Serviceability	100.0%	100.0%	100.0%	100.0%
External	Failures	0	0	0	0
	MTBF	∞	∞	∞	∞
	Serviceability	100.0%	100.0%	100.0%	100.0%
Total	Failures	0	0	0	0
	MTBF	∞	∞	∞	∞
	Serviceability	100.0%	100.0%	100.0%	100.0%

3.2 Utilisation

The graphs below show the overall utilisation of the two services, and the proportion of the main service utilisation which was classed as capability work – that is, jobs which used more that 256 processors.



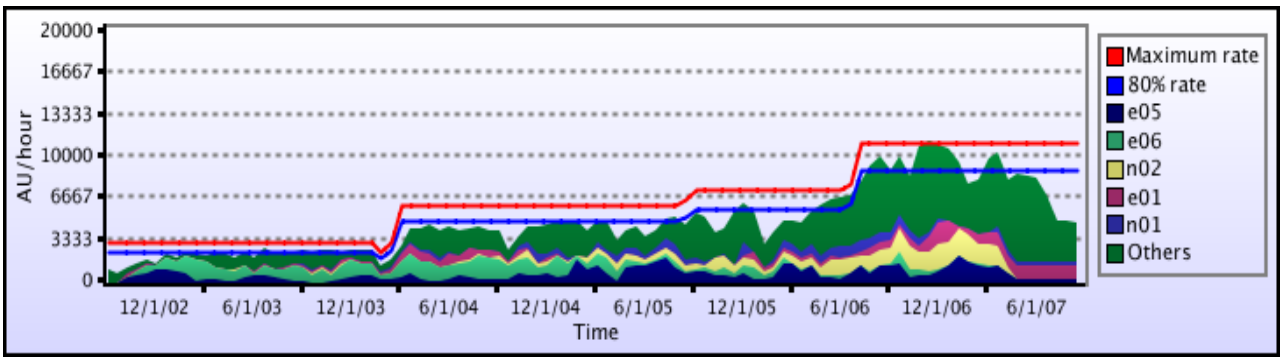


Utilisation figures greater than 100% for the development service correspond to a period in August and September 2006 when the number of processors in the service was temporarily increased.

3.3 Capacity Planning

Predicted Utilisation

The graph below shows the utilisation since the start of the project and the projected utilisation (on the main service) until September 2007. The scale on the *y*-axis is AUs per hour, where at peak Phase 3 can deliver 12034 AUs per hour (the upper red line in the graph). The lower line (in blue) corresponds to the more practicable 80% level.



The graph assumes:

- that each project will use its remaining allocation pro rata with its usage profile as known to the database, which is often simply that on the original application form.;
- that no more projects are given access to the service.

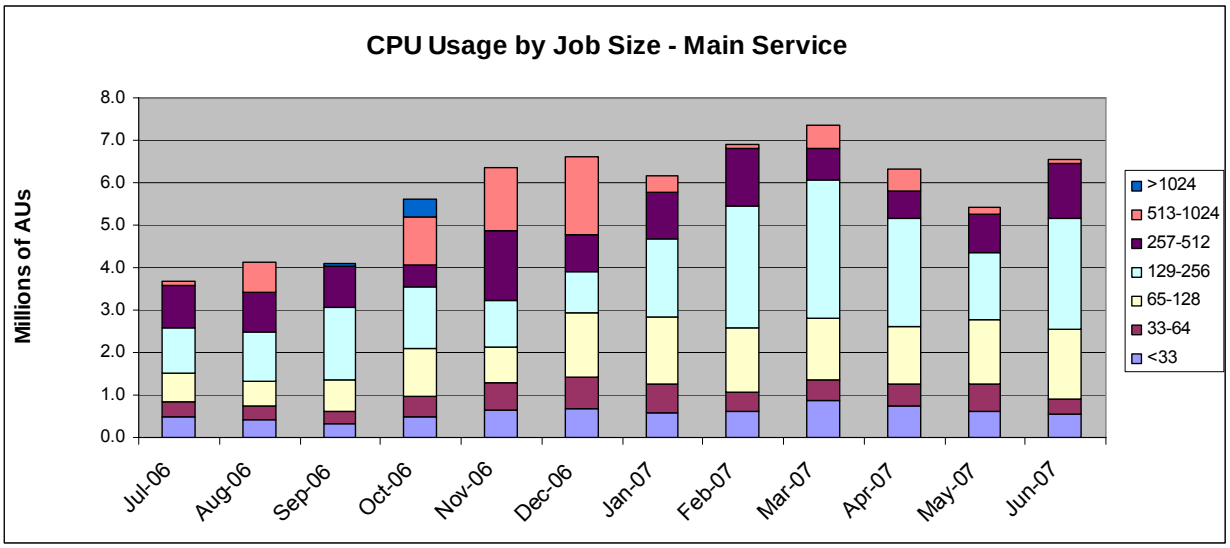
The graph shows that, on the basis of the projects which are currently using the service, we can anticipate some spare capacity later in 2007.

Numbers of Research Consortia

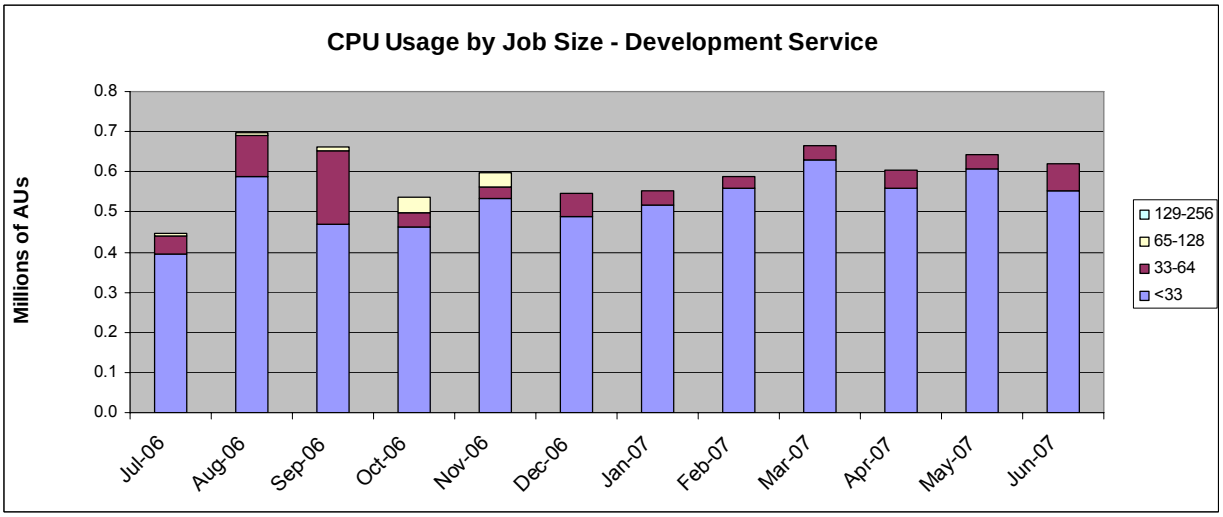
At the end of this quarter there were 45 research consortia on HPCx. In addition, there is one active externally funded project.

3.4 CPU Usage by Job Size

Main service



Development Service



3.5 AU Usage by Consortium

Main Service

Consortium	*April*	*May*	*June*	*AUs charged*	*%age of charged AUs*
e01	555927	507419	768053	1831399	10.1%
e03	1166	299		1465	0.0%
e05	1336103	1265693	1022654	3624449	20.0%
e06	44647	45195	19841	109683	0.6%
e08	204674	84349	140535	429558	2.4%
e10		30473	37538	68012	0.4%
e11	29148	118310	221266	368724	2.0%
e17	93931	28057	368	122356	0.7%
e18	4	1105	4717	5825	0.0%
e24	1	683565	1606188	2289755	12.6%
e26		2168	2130	4298	0.0%
e27	58			58	0.0%
e29		29068		29068	0.2%
e31		14	28069	28083	0.2%
e33	41460	2313	106141	149914	0.8%
e35	1054379	477707	392192	1924278	10.6%
e36	19164	128420	45069	192654	1.1%
e37	673655	331815	364371	1369841	7.6%
e38			85	85	0.0%
e39	35990	85134	89774	210898	1.2%
e41	104934	11	2552	107497	0.6%
e42	46954	18464	376678	442097	2.4%
e45	48377			48377	0.3%
e46	1651		1357	3007	0.0%
e48	29313	13478		42791	0.2%
e49	3625	7673	2292	13590	0.1%
e50	33757	41333	12164	87254	0.5%
e53	91	10	1772	1873	0.0%
e54	8625			8625	0.0%
e58	53534	5548		59081	0.3%
e59			51	51	0.0%
e60	43	23197	112513	135753	0.7%
e61			87694	87694	0.5%
EPSRC Total	4421212	3930820	5446063	13798095	76.1%

n01	4067	1200	109080	114347	0.6%
n02	902070	740441	633063	2275574	12.7%
n03	713049	535050	195939	1444037	8.0%
n04	57439	23217	27123	107779	0.6%
NERC Total	1676625	1299908	965203	3962125	21.9%

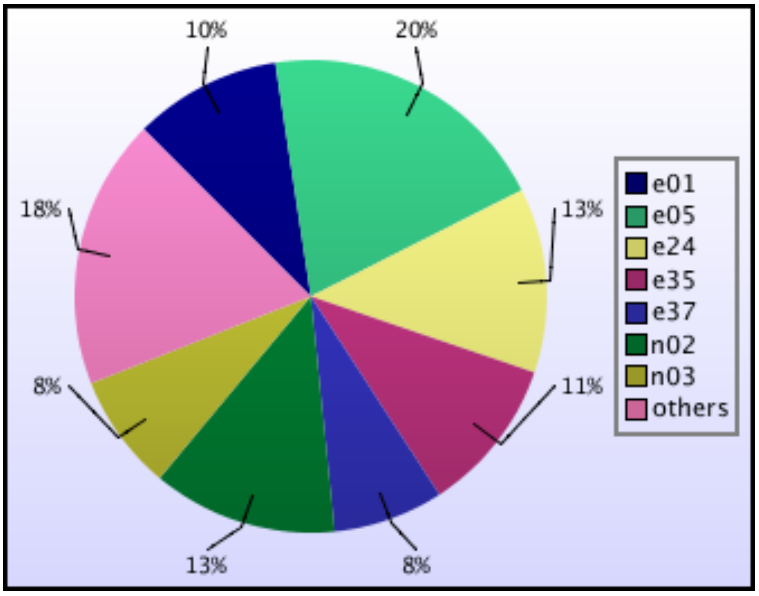
p01	28	17425	13738	31191	0.2%
PPARC Total	28	17425	13738	31191	0.2%

c01	60443	16122	4771	81336	0.4%
CCLRC Total	60443	16122	4771	81336	0.4%

b08	91391	39907		131298	0.7%
BBSRC Total	91391	39907		131298	0.7%

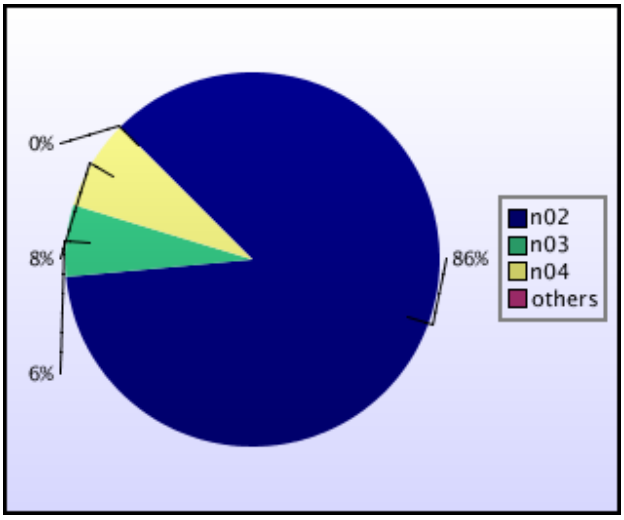
x01	3900	26339	40021	70260	0.4%
External Total	3900	26339	40021	70260	0.4%

z001	14421	31918	3379	49717	0.3%
z002	4			4	0.0%
z004	625	1604	5305	7533	0.0%
z06	207	0	176	383	0.0%
HPCx Total	15257	33522	8862	57640	0.3%



Development service

Consortium	*April*	*May*	*June*	*AUs charged*	*%age of charged AUs*
n01	44		0	44	0.0%
n02	559289	544165	503207	1606661	86.0%
n03	3894	60448	51520	115862	6.2%
n04	42193	38601	65269	146063	7.8%
NERC Total	605420	643214	619996	1874611	100.0%



3.5.2 Discounts

The following table shows the discounts that were awarded during the last quarter.

Consortium	*AU used*	*AU charged*	*Discount*
e01	1881285	1831398	49886
e05	3763935	3624449	139485
e27	60	57	3
e36	213398	192653	20744
e58	59757	59081	676
n03	1460091	1444037	16054

3.6 Helpdesk

3.6.1 Classifications

Category	*Number*	*% of all*
Admin	94	41.2%
Technical	94	41.2%
In-depth	36	15.8%
PMR	2	0.9%
Technical assessment	2	0.9%
Total	228	100.0%

3.6.2 Performance

All non-indepth queries	*Number*	*%*	*Target*
Finished within 24 Hours	137	72.9%	75%
Finished within 72 Hours	178	94.7%	97%
Finished after 72 Hours	10	5.3%	

Administrative queries	*Number*	*%*	*Target*
Finished within 48 Hours	89	94.7%	97%
Finished after 48 Hours	5	5.3%	

3.6.3 Experts Handling Queries

Expert	*Admin*	*Technical*	*In-Depth*	*PMR*	*Technical Assessment*
EPCC	69	39	24	0	0
DL	1	17	8	2	2
sysadm	24	38	4	0	0

3.7 Service Quality Tokens

No quality tokens were set this quarter.

4 Support

4.1 Applications Support *(Dr David Henty)*

4.1.1 Documentation

There have been no significant changes to online documentation this quarter – minor updates were made to give users more detailed information on how to control the production of core files, and to reflect recent updates to some key scientific libraries. We have kept users up-to-date with service developments via four user mailings.

4.1.2 Technical Reports

Two reports were planned for Q2 in the following areas:

a) Serial Efficiency of HPC Applications

b) Use of HPC Tools for Capability Computing

We have produced the following two reports this quarter

- **HPCxTR0703** *Single Node Performance Analysis of Applications on HPCx,* J.M. Bull.
- **HPCxTR0704** *Profiling Parallel Performance using Vampir and Paraver,* A.G. Sunderland.

which directly address these two areas.

Hardware counters can potentially provide useful insights into the behaviour of HPC applications. In HPCxTR0703, we took a set of 19 real user applications from a variety of scientific disciplines running on HPCx. We instrumented these applications using the hpmcount utility. From the raw counter data, we derived metrics which characterise the applications. There were a number of interesting results, for example the percentage of peak floating-point performance had an average of around 16%, with a minimum of 1.5% and a maximum of 50%. Significant correlations were observed between the flop rate and the fraction of instructions which are floating point, and also the rate of Level 1 cache accesses. The factors that affect efficient use of Simultaneous Multi-Threading were also investigated using similar techniques. This work was accepted for presentation at the international Scicomp13 conference to be held in Munich on 16 – 20 July.

Two of the major parallel profiling tools on HPCx are Vampir and Paraver, which are widely available on other platforms. Both tools can simultaneously monitor hardware counters and track message-passing calls, providing valuable information on an application's runtime behaviour which can be used to improve its performance. In HPCxTR0704 we look at using these tools in practice on a number of different codes on HPCx, with the aim of showing users how to utilise such profilers to help them better to understand the behaviour of their own codes.

4.1.3 Training

In Q2 of 2007 we ran the following two courses:

- **14 April - 16 April:** Fundamental Concepts of HPC (Edinburgh)
- **30 April - 2 May:** Message-Passing Programming (Edinburgh)

Statistics are summarised below alongside annual targets (where appropriate):

Metric	*Total*	*Target*
Course days	13	20
Different course titles	5	6
Different locations	2	4
Student-days for HPCx users	167	
Student-days for HPCx staff	27	
Student-days available for HPCx	311	

Our plans for the remainder of the year include running a message-passing programming course at a convenient location in England, perhaps in Cambridge or Leeds, and another course in Daresbury.

4.1.4 Workshops and Conferences

The second of the two workshops for this year will address the practical use of HPC tools for capability computing. The current plan is to hold this at RAL in Q4. The main conference for the year, the Fifth Annual HPCx Seminar, will be held in Daresbury at the end of 2007 alongside the 18th Machine Evaluation Workshop.

4.1.5 User Group

A User Group meeting using the Access Grid is will take place over the summer. Talks will include presentations based on the two technical reports mentioned above, and a discussion of the new helpdesk software from the user's point of view.

4.1.6 Newsletter

Issue seven of Capability Computing, *Simulations at an Atomistic Scale*, was distributed to users in June. We mailed out 3300 copies and made a PDF version available on the HPCx WWW pages. We also distributed the newsletter at our stand at the International Supercomputing Conference in Dresden, Germany, on 26-29 June.

4.2 Outreach Activities *(Dr Richard Blake)*

Progress against key objectives:

4.2.1 Life Sciences

With the completion of the Life Sciences funding, there will be a limited level of resources available for Outreach. Major activities for 2007 will be:

- *demonstration of the retina modelling code on much larger data sets.*

Work is ongoing in terms of analysing sequential bottlenecks in the inclusion of the image field.

Discussions are underway with BBSRC with a view to establishing two focussed workshops on Cell and Physiome Simulation.

4.2.2 Public/Industrial Awareness

We will aim at improving public and industry awareness, in particular through engagement with Science Festivals and marketing activities:

- *continued involvement in the Edinburgh International Science Festival;*

Dr Rob Baxter of EPCC delivered his public talk on 15th April as part of the Cutting Edge theme of the Big Ideas programme at the 2007 Edinburgh International Science Festival. Some 80 tickets were sold for the talk which was very well received. Special flyers were produced for the event to highlight the recent developments in UK Academic Supercomputing, specifically the HECToR project, which reproduced a number of recent press articles on the subject.

- *continued efforts to get funding for a longer term Public Understanding of Science programme around HPC;*

Unfortunately, our application to the Scottish Executive for support for the promotion of HPC and leading-edge computational science was not accepted; the funding line was heavily over-subscribed.

- *ongoing promotion of access to HPCx to industry as part of the general marketing activities at EPCC and CCLRC.*

4.2.3 Exploiting the CCLRC-funded visualisation facility

We plan to enable HPCx users to take advantage of this visualisation resource, which is physically collocated with the HPCx system. To this end we will:

- *develop two new collaborations with high-end computing Consortia;*

Visualisation projects are being explored with the Proudman Oceanographic Laboratory and the plasma fusion simulation group at Culham. These discussions are focussing on the software requirements and potential scenarios for multi-dataset visualisation.

- *hold a Workshop on High-end Visualisation in Engineering in collaboration with Viznet.*

Planned for 24th September 2007 at Daresbury Laboratory. Speakers invited include:

- Ken Badcock (Liverpool) - Overview of visualisation challenges in engineering.
- Roy Kalawsky (Loughborough University) – Human Computer Interfaces in Visualisation
- Roderick Johnstone - Real-time visualisation
- Nick Hill - Using PV3 / Unstructured grid work
- Rob Curr - Working with 2TB Nasa dataset
- Roger Proctor - POL visualisation challenges
- Martin Turner - Overview of new visualisation technologies
- Lakshmi Shastri - Compare and contrast shared memory vs distributed memory visualisation

4.2.4 The Grid

We will continue to support Grid projects which exploit HPCx including:

- *evaluation of Advanced Reservation;*

Implementation of High Availability Resource Coallocation being explored on the NW-GRID systems prior to implementation on HPCx.

- *support for demonstrations of HPCx integration into DEISA;*

10 Gb/s link procured due to go live on July 25 2007 because of delays in UKERNA provision.

- *support for joint EPSRC-NSF collaborations with Teragrid*

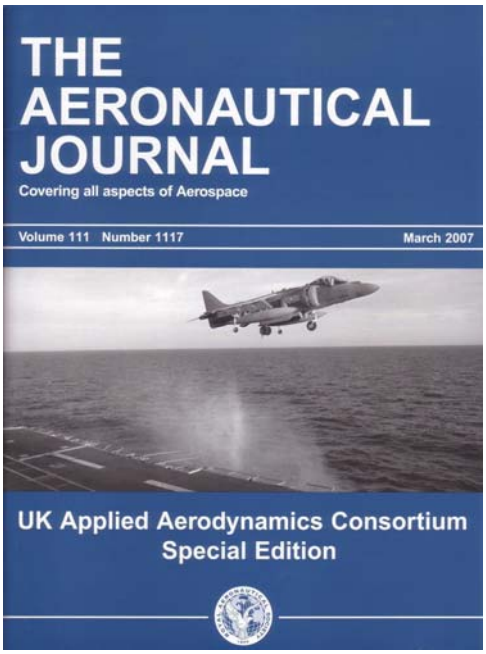
GENIUS grant proposal to simulate cranial blood flow in a Grid environment funded and currently exploring implementation of advanced reservation and MPIg implementation.

4.3 Terascaling Applications *(Dr Mike Ashworth)*

4.3.1 Key Objectives

Progress against key objectives

Publications



The Aeronautical Journal (published by the Royal Aeronautical Society) Vol 111 Number 1117 March 2007 is a UK Applied Aerodynamics Consortium Special Issue. There is a leading article by D.R. Emerson, M. Ashworth, A.G. Sunderland and K.J. Badcock (UKAAC, Liverpool). All seven technical papers reference HPCx. Six papers either are co-authored by or contain acknowledgements to HPCx staff: A.G. Sunderland and R. Johnstone.

Interactions with Consortia: F. Reid visited n01 consortium researchers at the Proudman Oceanographic Laboratory on 7th June 2007. She is working on optimisation of the OCCAM ocean modelling code on HPCx and this visit proved very helpful in furthering this work.

M. Plummer represented HPCx at the e05 Materials Chemistry meeting on 18th May at UCL, which included a detailed HPCx question-and-answer feedback session. There was a specific request for job times longer than the current 12 hour limit. This is one of the issues being addressed as part of the review of the role of HPCx as a complementary service to HECToR.

4.3.2 Computational Materials Codes

CASTEP: M. Plummer has extended CASTEP band-blocking to include LPAR-customisation as well as basic MPI_AllTo_AllV. On HPCx this does not improve performance due to the large-size buffer copying required. Band blocking does improve performance on HPCx compared to the uncustomised code. At the time of writing, ways of speeding up buffer copies and extending the use of band-blocking are being investigated in conjunction with K. Refson and P. Hasnip (certain higher level routines are not yet designed to call the 3d-FFT routines in

blocks of bands). CASTEP 4.2 was delivered at the end of the quarter and is being tested.

A series of comparison benchmarks and investigations have begun comparing HPCx CASTEP performance with BG/L and a Cray XT3 in Switzerland. HPCx performs well compared to the XT3 in dual-core mode (single core XT3 performance was slightly better than HPCx). An initial report has been circulated internally and to the CASTEP Developers Group which will (in Q4) appear as part of a Technical Report. Both BG/L and XT3 have fairly low latency, so straightforward band-blocking does not improve performance. However, single-sided communication (SHMEM) is being developed for the XT3 and the use of ISO_C_BINDING (get hostname) Fortran 2003 features is being investigated on all systems in order to make optimal use of processor layout.

CRYSTAL: I.J. Bush has added calculation of the stress tensor to CRYSTAL. This was at the request of a user (P.Madden) and will allow his group and collaborators to fir potentials for classical MD simulations. Work on reducing the replicated memory usage of CRYSTAL has continued. The introduction of a new optimiser into CRYSTAL06 has resulted in the (re-)introduction of a number of large, replicated arrays that scale with the size of the system squared. In particular, matrices for transformation between a number of different coordinate systems are now the largest objects in the code, and work is in progress to move this to distributed data.

QUICKSTEP: A request for in-depth support has been received from Ben Slater of the e05 consortium for investigation and optimisation of the 'Quickstep' density functional theory code. K. Stratford is currently liaising with the consortium to get this work underway.

VASP: K. D'Mellow has addressed further memory leak issues within VASP's xml output routines that were causing problems for a specific user. VASP's xml module uses irregular character range constructs, and frequently references illegal character elements (e.g. zero, or greater than the length of the character in question). This is a persistent problem seen by several VASP users, across all of the 4.6.x versions so far encountered.

4.3.3 Computational Chemistry Codes

General: The Distributed Data Interface (DDI) has until now only been available as a component of the US-version of the GAMESS quantum chemistry package. G. Fletcher has extracted DDI as a stand-alone library capable of being linked with other programs. DDI offers a virtual shared memory API similar to Global Arrays (GA), with one-sided data copy operations such as GET, PUT and ACCumulate, though, while GA is an on-going research project, DDI maintains a core functionality subset with an emphasis on portability inherited from its usage in GAMESS.

Thus, portability in DDI is based on that of GAMESS, which currently extends to the following types of architecture (most in both 32-bit and 64-bit forms): Linux

platforms - Compaq AXP and SuperCluster, x86, intel, and AMD (with optimisation settings for the ifort, pgf77, g77, and gfortran compilers), Cray platforms - X1, XD1, PVP, T3E, and XT3, Fuiitsu PP, HPUX, IBM platforms - SP (e.g. HPCx) and BG/L, NEC SX, SGI, SUN, Mac 32-bit and G5. At the time of writing the DDI library has been validated on Linux - intel and x86 (using ifort and g77, resp.), HPCx (IBM P5-575), and Cray XT3 (at CSCS in Switzerland), though, ideally, it would be desirable to validate this 'library' form of DDI on all the hardware mentioned above.

Plans are already in place to use the DDI library in the quantum chemistry packages GAMESS-UK and MolPro. At the moment DDI (in the form of a zipped tar file) can be obtained upon request, but in due course it will become available via free download from sites at Daresbury and Iowa State University. The library form of DDI comes with installation instructions and a test program. Other short-term activities will include porting DDI to new hardware such as the BG/P and Cray XT4.

DL_POLY: I.J. Bush has carried out further optimisations to DL_POLY3 and the improved speed has been demonstrated on a number of architectures. For instance on an XT4 the run time for a 3.8 million particle fluortizied pyrochlore has been reduced from 165 seconds to 105, while on a Pentium dual-core system a 12% improvement has been observed for a smaller simulation of NaCl.

Most of the improvements have been in the reciprocal space part of the evaluation of the Ewald terms. The two most important are

a) Optimising the recurrences that calculate the components of the Cardinal B-Splines used in the Ewald summation. This has been a known problem for a while, but profiling on the XT4 showed this to be the dominant routine while on other machines it was less important. Reorganisation of the loops allowed better cache usage and also, on the XT4, the possibility of vectorisation once a number of conditionals had also been eliminated. On the XT4 this reduced the original run time of 165 seconds down to around 125 seconds.
b) Improving the loop structures for evaluating the Cardinal B-splines on the grid. This mostly involved the elimination of a number of conditionals to handle the tails of the B-Splines into/out of the halo data. This reduced the run time from 125 to 110 seconds.

Work is now progressing on profiling more biological systems to see if the work balance is similar there, or further routines need optimisation.

M. Plummer has started work on adapting and generalising a subroutine from DL_POLY_3 to allow an 'intelligent' variation of the simulation timestep during early, pre-equilibrium stages of a simulation. The routine will monitor instantaneous forces, reduce the timestep appropriately to keep nearby 'atoms' from moving into each other and temporarily cap forces if necessary.

GAMESS-UK: The improved ScaLapack version of the GA GAMESS-UK code has been completed and has been tested in collaboration with the code developers. GAMESS-UK users can now access this new version from the central GAMESS-UK repository. Improvements in performance of up to 50% can be achieved for the parallel diagonalisation stages using the new version.

GROMACS: We received a request from Dr G Patargias (working with Professor J Harding, e50, Biological interface with materials), for task-farming assistance with the GROMACS MD package. M. Plummer has built a version of GROMACS which uses MPICH as a secondary message-passing layer to control the task-farming in a similar way to that done by P. Sherwood for CHARMM last quarter. The source code for GROMACS is too complex (and mostly written in C) for simple application of the 'Fortran-wrapper' task-farming approach used for DL_POLY_2 or the full intelligent task-farming approach developed by J. Hein. In order to be compatible with GROMACS, the configuration and make files for both GROMACS and MPICH were adapted to thread-safe 64-bit mode and full compiler optimisation options.

Having confirmed that the MPICH implementation can be useful and robust, it has now been made available in /usr/local/packages/ including the executables and instructions on appropriate LoadLeveler scripts and how to link application codes. This package will be open but not publicised and it will be used and recommended at staff discretion, as a full investigation of how MPICH uses the node has not yet been undertaken. Dr G. Patargias was finding his simulations only scale to around 10 processors, so the task-farming scheme allows more efficient use of HPCx

LAMMPS: F. Reid has installed the latest version (12 Feb 2007) of LAMMPS on HPCx. She has tested and benchmarked this code and updated the HPCx web pages to reflect the changes.

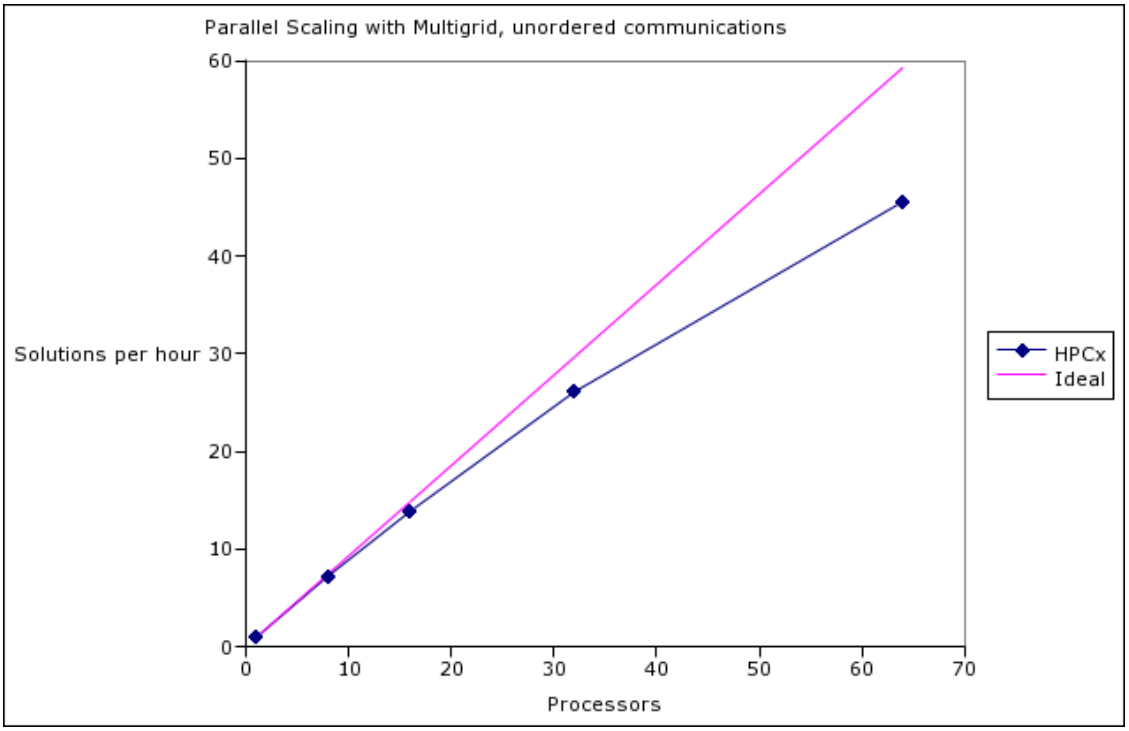
4.3.4 Physics Codes

GWST code: K. D'Mellow has continued to profile and begun to parallelise the GW-Space-Time code, creating communications routines that still allow serial execution if desired. This allows only one code version to be maintained, which is important because the code is being worked on by several of the GWST developers. He has overcome porting issues related to the code itself, and several of the test suite datasets, which were suffering from build and conversion problems related to NetCDF.

4.3.5 Engineering Codes

HyClaws LES code: R. Johnstone has parallelised the HyClaws LES code from Dr. Evgeniy Shapiro, Cranfield (working with D. Drikakis, e37 LESUK_3). HyClaws is designed to solve arbitrary hyperbolic or hyperbolised systems of equations on multiblock structured grids, using multigrid acceleration. Features of the parallelisation include the amalgamation of messages to reduce the effect of

message passing latency, ordered and unordered halo exchanges for debugging and performance respectively, lower precision messages for all or any grids (this is particularly useful for the coarse grids in a Full Approximate Storage (FAS) multigrid method like this, because the converged fine grid solution is not affected). Reasonable scaling is seen for a simple test case of 64 blocks (see figure below). Serial optimisations have shown a speed-up of around 2.5 on the Cray XT4. The improvement on HPCx is only around 20% as the IBM xlFortran compiler does a better job with the original code.



4.3.6 Life Sciences

NAMD: J. Hein has been working to ensure that NAMD was available on HPCx after a compiler upgrade. NAMD 2.6 refused to build with the early versions of the xlc 8.0 compilers. The problem had been narrowed down to a single file, when it was discovered that it was a compiler problem. These issues disappear when using a later version of xlc 8. The executable built using the newly available xlc 8 shows similar performance to the executable built with the now discontinued xlc 7.0.

Integrative Biology: A query from Gernot Planck from the e17 consortium, relating to a code using PETsc, was resolved by locating a bug in the MPI implementation. This has now been fixed by IBM, and the fix has been supplied to the user.

4.3.7 New Applications

STAR-CD: We received a request from the e59 (Turbulence in breaking gravity waves) consortium to install the STAR-CD software as a package on the system. J. Hein is currently performing this task.

OCCAM: We received a request from the n01 (Large-Scale Long-Term Ocean Circulation) consortium to perform scaling work on the OCCAM ocean modelling code. F. Reid has obtained the 2D version of OCCAM from researchers at Proudman Oceanographic Laboratory (POL). She has compiled and tested the code, benchmarked over a number of CPU counts and investigated whether Simultaneous Multithreading would be beneficial. She has investigated using MPICH to task-farm the code. As the code needs to run on more than 16 processors this approach is not possible and a full task-farm harness needs to be written. The code was also profiled with no single routine standing out. Following a consortium visit (see above) priority has been given to the task-farm harness with code optimisation seen as somewhat secondary to their needs. This work is ongoing.

4.3.8 Techniques

RAMDISK: RAMDISK is an IBM feature for allocating memory to be used for reads from and writes to files. This allows small temporary files to be completely memory-resident which may bring significant performance benefits to certain applications. As root permission is currently required to set up RAMDISK, this is being tested on a separate system from the HPCx production platform. The findings here appear promising: RAMDISK I/O from a parallel code is around five times faster then normal disk I/O. However, in order to draw meaningful conclusions, tests need to be undertaken on the HPCx machine (i.e. with GPFS and many more processors available). We shall continue this investigation during an HPCx maintenance session.

4.3.9 Libraries

Linear Solvers: A.G. Sunderland has carried out further testing of the parallel direct linear solvers in the MUMPS_4.6.4 package. Large sparse matrices from the UF sparse matrix collection at Sandia have been downloaded to HPCx in order to get a better analysis of MUMPS performance, but at present the parallel scaling remains very modest. Parallel profiling and analysis is being undertaken to better understand the reasons behind the exhibited performance.

4.3.10 Tools

TotalView: Totalview version 8.1 has been installed on the HPCx. This version has improved memory debugging features.

Vampir and Paraver: Following work with the tool developers at TU Dresden, A.G. Sunderland has upgraded VampirTrace to version 5.3.1 (officially released

on 25th June) on HPCx. This new version works with the incorporation of hardware event counters based on PAPI (only 32-bit at present). This is a new feature that allows users to measure performance parameters such as cache hits/misses and flop rate for user-defined sections of their application code through the VAMPIR GUI. We have produced Vampir traces with hardware event counter information for an eigensolver code and for DL_POLY3 as test cases for an HPCx Technical Report. The report also includes traces from the latest version of the Paraver profiling tool which is also available on HPCx.

"Profiling Parallel Performance using Vampir and Paraver", A.G. Sunderland, HPCx Technical Report, HPCxTR0704

4.3.11 Applications on New Technologies

J. Hein has continued working on optimisations for parallel FFTs when using a toroidal network, which is relevant e.g. for IBM BlueGene and the Cray XT4 (HECToR) hardware.

J. Hein, A. Gray, F. Reid, A. Simpson, M. Bull, M. Weiland and S. Booth have been working on porting a number of codes, many of which currently being used on HPCx, to HECToR. This work will contribute to an HPCx technical report which compares performance of the two systems. In preparation for this report, A. Gray has been running synthetic and application benchmarks on HPCx. In relation to this work, J. Hein attended the CRAY user group (CUG) in Seattle in May in Seattle and reported back from the meeting to the HPCx team at EPCC, M. Ashworth, D.S. Henty and F. Reid attended the Cray Technical Workshop at CSC in Helsinki, Finland, 4-6th June, and a number of HPCx staff attended a Cray workshop organised by and held at EPCC from 11-14 June.

J. Hill has been investigating IBM's Cell technology. This was designed for Sony's Playstation 3 gaming system but is now being offered as part of IBM's blade server range and may provide a good platform for HPC. However, it is still unclear how exploitable the technology is, and whether it is appropriate for the supercomputing field. The aims of this project are to explore the performance of the Cell hardware and to evaluate the usability of the programming paradigms that are currently available. The approach taken is to port simple, but representative, codes to the Cell platform, and this work is ongoing.

4.3.12 Talks and Publications

“High-Performance Computing and Computational Aerodynamics in the UK”, D.R. Emerson, M. Ashworth, A.G. Sunderland and K.J. Badcock, The Aeronautical Journal, Vol 111, No 1117, 125-132

“Application of a Parallel Rotor CFD code on HPCx”, C.B. Allen, A.G. Sunderland and R. Johnstone, The Aeronautical Journal, Vol 111, No 1117, 145-152

"Profiling Parallel Performance using Vampir and Paraver", A.G. Sunderland, HPCx Technical Report, HPCxTR0704

4.4 Software Engineering *(Dr Stephen Booth)*

4.4.1 Novel languages and techniques

Mark Bull acts as chair of the OpenMP Architecture Review Board (ARB) Language Committee, which is the forum that produces the official specifications for OpenMP. The committee meets once a week, by phone conference and is attended by representatives of the 15 commercial and academic institutions that comprise the membership of the ARB. These meetings are backed up by ongoing email discussions and wiki postings. In the last quarter, the committee has been finalising the decision over which revisions and additions to OpenMP will be included in the next major release (OpenMP 3.0). During the next quarter, these proposals will be turned into a draft specification, which will then be available for public comment.

A number of the current EPCC MSc students are investigating novel parallel languages on HPCx. These include UPC and Titanium. UPC is a dialect of C intended to write SPMD programs. UPC supports globally shared arrays where data on remote processors can be accessed using a syntax similar to a normal array access in C. Titanium is a Java like language developed at UC Berkeley and specifically intended for developing HPC applications. Though it is based on Java it is a compiled rather than an interpreted language. It has various additional language features to support parallel processing including a global memory space abstraction (similar to other languages such as UPC or Co-Array Fortran).

A technical report on the use of object-oriented techniques for parallel computing is currently under development. This report investigates object-oriented design patterns that can be used to encapsulate data decomposition in a distributed memory parallel application. Object-oriented techniques are relatively easy to apply when using thread-based parallelism. They are much less easy to apply to MPI-based programs that use distributed memory. The key concept of object-oriented programming is to encapsulate data structures within opaque objects. This encapsulation is much harder to achieve when the data structures are distributed across multiple memory systems.

4.4.2 In-depth software support

User administration software

The new helpdesk system integrated within the SAFE user administration software was deployed on HPCx at the beginning of May. This migration has been relatively trouble free and the new system has been working well. Experience with using this system has resulted in a number of changes and enhancements which have been added since the initial installation. These include:

- improved capabilities to search the query database including text based search;
- improvements to the user interface;
- better handling of badly formatted email messages;
- additional code to generate query reports.

Code has been added to allow project managers to export their projects' accounting records in XML format using the XML schema developed by the Global Grid Forum. This format is used by many grid projects and will allow project managers to upload their projects' usage data into grid accounting systems if they so wish. In particular this will allow HPCx accounting data used by the DEISA project to be uploaded into the DEISA accounting system.

4.5 Operations and Systems *(Mr Mike Brown)*

Reliability: There were no SEV1 incidents (ie, contractual failures) this quarter. The last failure attributable to IBM was in December 2006; the last attributable to the site was in September 2006.

Software upgrades: There were two maintenance sessions this quarter. On 25 April, the Fortran run-time environment was upgraded to version 10.1.0.4. Following extensive testing, on 27 June the Fortran compiler was upgraded to xlf10.1.0.4, and the C compiler to xlc8.0.0.13, to match the runtime environment.

Carbon footprint: In response to a user inquiry, we calculated an approximate value for the power required to provide a CPU hour on HPCx: 0.4kW/hr, or 0.08kW/hr per AU at Phase 3. This takes account of cooling and efficiency losses, and assumes that all ancillaries such as storage etc have their proportions uniformly distributed.

Outreach: Mike Brown, Steve Andrews and Damian Jones attended the UK IBM HPC user group meeting in Mainz, on 17–18 April.

4.6 Staffing

AV	*April*	*May*	*June*
DL	4.4	4.4	4.5
EPCC	7.7	9.5	8.8
Total	12.1	13.9	13.3

Systems	5.8	5.8	5.8

5 Summary of Performance Metrics

Metric	*TSL*	*FSL*	*April*	*May*	*June*
Technology serviceability	80%	99.2%	100.0%	100.0%	100.0%
Technology MTBF (hours)	200	300	∞	∞	∞
Number of AV FTEs	7.5	10	12.1	13.9	13.3
Number of training days per month	22.5/12	30/12	13/4	13/5	13/6
Non in-depth queries resolved within 3 days	85%	97%	97.0%	97.1%	88.5%
Number of A&M FTEs	3.75	5.75	5.8	5.8	5.8
A&M serviceability	80%	99.6%	100.0%	100.0%	100.0%

Colour	*Meaning*
	Exceeds FSL
	Between TSL and FSL
	Below TSL

Note 1: The number of training days is reported as a running total since the start of the year.

Note 2: The above table includes the revised FSL targets for *training days* and *A&M serviceability*, which have been provisionally agreed with EPSRC.

Appendix A: Incident Severity Levels

SEV 1 — anything that comprises a FAILURE as defined in the contract with EPSRC.

SEV 2 — NON-FATAL incidents that typically cause immediate termination of a user application, but not the entire user service.

The service may be so degraded (or liable to collapse completely) that a controlled, but unplanned (and often very short-notice) shutdown is required or unplanned downtime subsequent to the next planned reload is necessary.

This category includes unrecovered disc errors where damage to filesystems may occur if the service was allowed to continue in operation; incidents when although the service can continue in operation in a degraded state until the next reload, downtime at less than 24 hours notice is required to fix or investigate the problem; and incidents whereby the throughput of user work is affected (typically by the unrecovered disabling of a portion of the system) even though no subsequent unplanned downtime results.

SEV 3 — NON-FATAL incidents that typically cause immediate termination of a user application, but the service is able to continue in operation until the next planned reload or re-configuration.

SEV 4 — NON-FATAL recoverable incidents that typically include the loss of a storage device, or a peripheral component, but the service is able to continue in operation largely unaffected, and typically the component may be replaced without any future loss of service.

Appendix B: Projects

B.1 Current Projects

EPSRC Projects

Code	*Class*	*Title*	*PI*
e01	1	UK Turbulence Consortium	Dr Gary Coleman
e05	1	Materials Chemistry using Terascaling Computing	Prof Richard Catlow
e06	1	UK Car-Parrinello Consortium	Prof Paul Madden
e08	2	Organic Solid State	Prof Sarah Price
e10	1	Reality Grid	Prof Peter Coveney
e11	1	Bond making and breaking at surfaces	Prof Sir David A King
e14	1	Blade and Cavity Noise	Prof Neil Sandham
e15	2	CSAR/HPCx Collaboration	Dr Mike Pettipher
e16	1	Cardiac virtual tissues	Prof Arun V Holden
e17	1	Integrative Biology	Dr David Gavaghan
e18	1	DARP: Highly swept leading edge separations	Prof Michael A Leschziner
e19	1	Edinburgh Soft Matter and Statistical Physics Group	Prof Michael E Cates
e20	1	UK Applied Aerodynamics Consortium	Dr Ken Badcock
e21	1	Intrinsic Parameter Fluctuations in Decananometer MOSFETs	Prof Asen M Asenov
e22	1	Preconditioners for finite element problems	Prof David J Silvester
e23	1	Exploitation of Switched Lightpaths for e-Science Applications	Prof Peter Clarke
e24	1	DEISA - Distributed European Infrastructure for Supercomputing Applications	Dr David Henty
e25	1	Turbulent vortex motion in stratified flows	Dr Gary Coleman
e26	1	Simulation of Radioprobing	Dr Charlie Laughton
e27	1	SPICE	Prof Peter V Coveney
e29	1	Free-surface-piercing circular cylinders	Dr Eldad Avital

e30	1	Metal/Oxide Interfaces at the Atomic Level	Dr Nora de Leeuw
e31	1	Lateral Straining of Wall-Bounded Turbulence	Dr Gary N Coleman
e32	1	Rapid Prototyping of Usable Grid Middleware	Prof Peter V Coveney
e33	1	Engineering Functional Coatings	Prof Roger Smith
e34	1	Dissolution of Bioactive Phosphate Glasses	Dr N de Leeuw
e35	1	Non-adiabatic processes	Dr T Todorov
e36	1	Jets in Cross-Flow	Dr Y Yao
e37	1	LESUK_3	Prof J J McGuirk
e38	1	Viscoelastic deformation in 3D non-linear media	Prof Greg A Houseman
e39	1	The Supergen 5 biological fuel cells consortium	Prof FA Armstrong
e40	1	Computational Quantum Many-Body Theory	Prof R Needs
e41	1	Flow in Weapon Bays	Dr George N Barakos
e42	1	Computational Combustion for Engineering Applications	Prof K Luo
e44	1	Extreme Wave Loading on Offshore Wave Energy Devices	Dr Deborah Greaves
e45	1	Metals under extreme conditions	Prof Mike Gillan
e46	1	Advanced materials with complex architectures	Dr Paul Mummery
e47	1	Parallel stochastic analysis for geo-engineering	Dr Michael A. Hicks
e48	1	Organised structure in turbulent flows	Prof Sergei Chernyshenko
e49	1	Integrated Programme of Research in Aeronautical Engineering	Prof Michael Leschziner
e50	1	Biological interface with materials	Prof John Harding
e51	1	Super-computing data mining	Dr Mike Pettipher
e52	1	Spacecraft force modelling	Dr M Ziebart
e53	1	Large-scale communication networks	Prof J M Pitts
e54	1	Free surface simulation of waves overtopping during storms	Dr D M Ingram

e55	1	High-Reynolds-Number Near-Wall Flows	Prof Michael Leschziner
e56	1	Infectious disease threats	Dr Iain Barrass
e57	1	Triplex DNA Structures	Dr Hooshang Nikjoo
e58	1	Quantum Simulations for Chemical Biology	Dr Carmen Domene
e59	1	Turbulence in breaking gravity waves	Prof Ian Castro
e60	1	Device Electronics Based on nanoWires and NanoTubes	Dr Merlyne M de Souza
e61	1	Enhancement of droplet concentrations in clouds	Dr Alan A M Gadian

PPARC Projects

Code	*Class*	*Title*	*PI*
p01	1	Atomic Physics and Astrophysics	Prof Alan Hibbert

NERC Projects

Code	*Class*	*Title*	*PI*
n01	1	Large-Scale Long-Term Ocean Circulation	Dr David Webb
n02	1	NCAS	Prof Alan J Thorpe
n03	1	Computational Mineral Physics Consortium	Dr John Brodholt
n04	1	Shelf Seas Consortium	Dr Roger Proctor

BBSRC Projects

Code	*Class*	*Title*	*PI*
b02	1	Modelling enzyme catalysis	Dr Adrian J Mulholland
b08	1	IntBioSim	Prof M S Sansom

CCLRC Projects

Code	*Class*	*Title*	*PI*
c01	1	Daresbury Laboratory Facilities Agreement Consortium	Dr Richard J Blake

Externally-funded Projects

Code	*Title*	*PI*
x01	HPC-Europa	Dr Judy Hardy
x03	IBM	Mr Derrick J Byford

HPCx Projects

Code	*Title*	*PI*
z001	HPCx Support	Dr Alan Simpson
z002	Systems and Operations	Mr Mike Brown
z003	Test Project	Dr Denis Nicole
z004	HPCx Training	Dr David Henty
z05	Outreach Projects	Dr Richard Blake
z06	Application Porting	Dr David Henty
z07	Package Installation	Dr Mike Ashworth
z10	Globus	Dr Stephen P Booth

B.2 Former Projects

Code	*Class*	*Title*	*PI*
b01	2	Quantum Chemistry Studies of the Rusticyanin Protein Crystal	Prof Samar Hasnain
b03	1	Towards a virtual outer membrane	Prof Mark S Sansom
b04	1	Life sciences software development	Dr Jo L Dicks
b05	1	Virtual forced evolution of catalytic transition metal complexes	Dr Marcus Durrant
b06	2	Biomolecular computational chemistry	Prof Jonathan D Hirst
e02	1	Ab-initio simulation of covalently bonded materials	Dr Patrick Briddon
e03	1	Multi-photon, electron collisions and BEC HPC consortium	Prof Ken Taylor
e04	1	Chemreact Computing Consortium	Prof Jonathon Tennyson
e07	2	Turbulent Plasma Transport in Tokamaks	Dr Colin M Roach
e09	2	Molecular Properties and their Geometry	Dr Mark R Wilson
e12	1	Parallel programs for the simulation of complex fluids	Dr Richard J Blake
e13	1	TeraGyroid project	Mr Mark Westwood
e28	1	Towards the Dynome	Dr Jonathan W Essex
z09		HECToR Benchmarking	Dr Edward Smyth
x02		OHM Ltd	Dr Lucy MacGregor
n05	2	Non-linear Wave-particle Instabilities in Plasmas	Dr Mervyn Freeman
e07	2	Turbulent Plasma Transport in Tokamaks	Dr Colin M Roach