HPCx Quarterly Report
July – September 2006

1 Introduction

This report covers the period from 1 July 2006 at 0800 to 1 October 2006 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 has been another busy but successful quarter for the HPCx service and reliability has remained very high.
- Usage has improved from last quarter and has returned to the normal levels of more than 75%. Capability usage averaged more than 30%.
- The major focus from a Service perspective has been preparations for the Phase 3 upgrade. The additional nodes and disk storage were all in place for an 'early-access' service from early October. The implementation tests are planned for October and should allow a full, production service from the beginning of November.
- The 4th HPCx Annual Seminar, *Moving Science Forward*, went ahead very successfully on the 4th October in Edinburgh with more than 70 attendees. The talks covered a wide range of applications areas and included talks from both IBM and Cray; we received very positive responses from the participants.
- A well-attended User Group was held as the final session of the Annual Seminar. We solicited input from the users on: the batch queue setup for Phase 3; and the Annual Plan for 2007, including topics for technical reports and possible course locations.

- A workshop on Materials Modelling was held on the preceding day in the same location. This was interesting and interactive.

- There was a STAC meeting in London at which the committee members were positive about our progress against the Annual Plan. The various activities are on target and we have clear plans to complete them successfully.

- A final report on the HPCx-IBM Lifesciences project has been prepared and highlights of the various projects are being developed into web pages.

- The use of Simultaneous Multithreading (SMT) on HPCx has been investigated across a wide range of codes. This activity has contributed to a number of technical reports and a talk at ScicomP.

- The Terascaling team have worked on a wide variety of codes during this quarter. Significant progress has been in improving the parallel performance of MPPCRYSTAL for systems of current scientific interest.

- HPCx supported Peter Coveney and collaborators for various demonstrations at the All Hands Meeting in September 2006 and will repeat this for demonstrations at Supercomputing 2006.

- A number of codes (including the Unified Model) experienced problems after the major software upgrade in May that was required for the Phase 3 upgrade. These problems were tracked down to memory leaks in MPI and have now been successfully resolved. A report, including key recommendations, was approved by the Oversight Committee.

- The Software Engineering team have been investigating the performance of Power 5 and OO programming. This work has been published on the web in technical reports.
- There are now 53 projects on HPCx, including the former CSAR projects, with another approved by EPSRC for access. This leaves one spare place within the new maximum of 55.

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:

	July	*August*	*September*
Incidents	2	1	1
Failures	7	13	5

The following tables give more details on the attribution of the failures:

July

Failure	*Site*	*IBM*	*External*	*Reason*
06.070	100%	0%	0%	Nodes taken down due to heatload
06.072	100%	0%	0%	Nodes taken down due to heatload

August

Failure	*Site*	*IBM*	*External*	*Reason*
06.078	0%	100%	0%	GPFS failure

September

Failure	*Site*	*IBM*	*External*	*Reason*
06.089	100%	0%	0%	Emergency power off after aircon failure

None

3.1.2 Performance Statistics

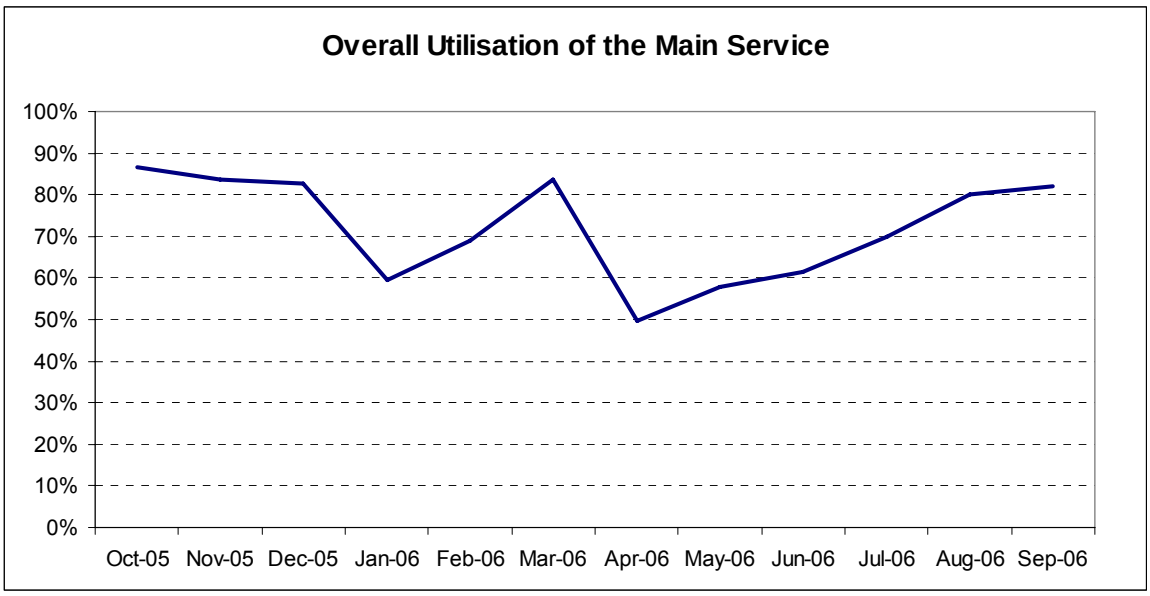
This section uses the definitions agreed in Schedule 7, ie,

- 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	1	0	1
	MTBF	∞	732	∞	2196.0
	Serviceability	100.0%	99.9%	100.0%	100.0%
Site	Failures	2	0	1	3
	MTBF	366	∞	720	732.0
	Serviceability	97.0%	100.0%	99.4%	98.8%
External	Failures	0	0	0	0
	MTBF	∞	∞	∞	∞
	Serviceability	100.0%	100.0%	100.0%	100.0%
Total	Failures	2	1	1	4
	MTBF	366	732	720	549.0
	Serviceability	97.0%	99.9%	99.9%	98.9%

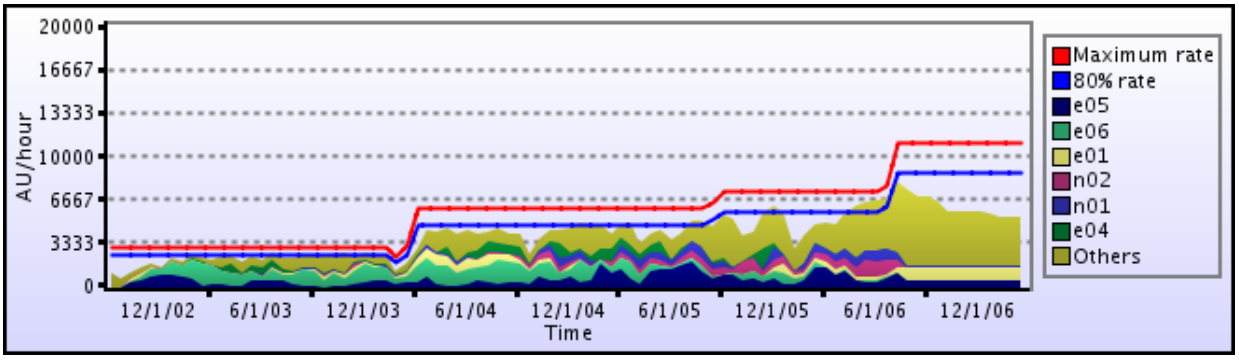
3.2 Utilisation

Normally we report here on the utilisation of the capability region. At present, when there are many 128-processor jobs waiting to use the capacity region, the systems team responds by moving such jobs into the capability region, if there is room there. As a result, the utilisation figures for the capability region do not provide significantly useful information. Instead, therefore, we report here the overall utilisation for the entire main service.



3.3 Capacity Planning

Predicted Utilisation



The graph above shows the utilisation since the start of the project and the projected utilisation (on the main service) until December 2006. 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 we can anticipate some spare capacity in the next few months, now that Phase 3 has come into service.

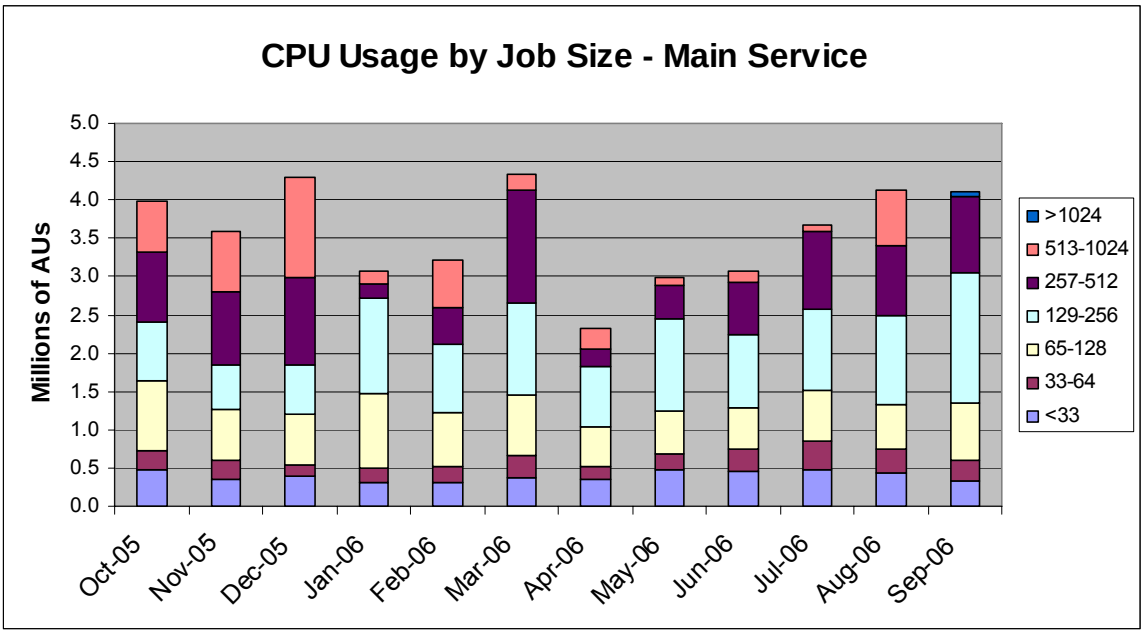
Numbers of Research Consortia

At the end of this quarter there were 53 research consortia on HPCx.

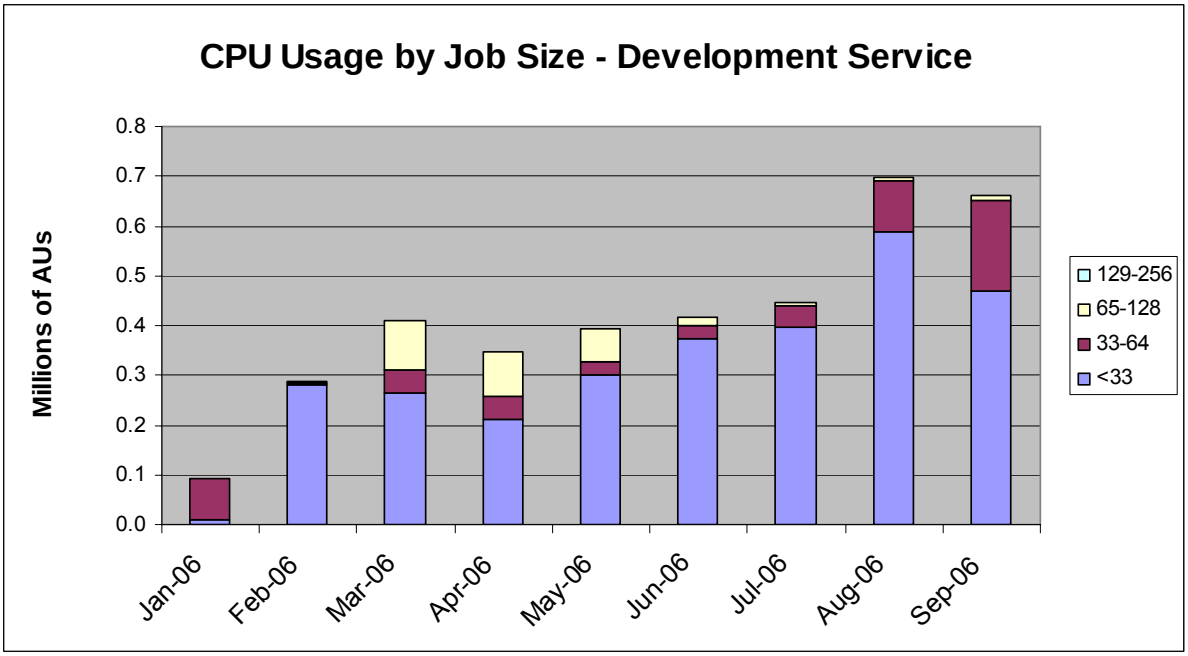
In addition, there is one active externally funded project.

3.4 CPU Usage by Job Size

Main service



Development Service



3.5 AU Usage by Consortium

Main Service

Consortium	*July*	*August*	*September*	*Quarterly*	*%age of charged AUs*
e01	262754	194295	152976	610025	5.1%
e03	2	3983	613	4598	0.0%
e05	905389	388152	410995	1704536	14.2%
e06	44491	81136	112218	237845	2.0%
e08	148994	184462	181259	514715	4.3%
e11	66838	49714	67577	184129	1.5%
e14	151699	353174	64997	569870	4.7%
e15	3	2226	1731	3960	0.0%
e17	30002	11005	17219	58226	0.5%
e18	66973	2242	0	69215	0.6%
e20	270963	153102	152136	576201	4.8%
e21	22781	58	0	22839	0.2%
e23	28		1681	1709	0.0%
e24	2577	120837	215940	339354	2.8%
e25	7227	839	14	8080	0.1%
e26	1041		0	1041	0.0%
e27	301	193081	124014	317396	2.6%
e29	9600		0	9600	0.1%
e31	0	705	2	707	0.0%
e32	61521		0	61521	0.1%
e33	98105	168867	238773	505745	4.2%
e35	1470	548	9730	11748	0.1%
e36	4373	735	3661	8769	0.1%
e37	34825	57371	187956	280152	2.3%
e38	996	917	946	2859	0.0%
e39	61857	35712	212602	310171	2.6%
e40	18626	13971	0	32597	0.3%
e41	9997	2881	0	12878	0.1%
e42	84208	702453	22361	809022	8.3%
e45	113524	57823	131413	302760	2.5%
e49	2294	1024	41848	45166	0.4%
e50	2559	52653	15987	71199	0.6%
e51	67	466	125	658	0.0%
e52	0		1	1	0.0%

e53	1393	620	1764	3777	0.0%
e54	4	868	32	904	0.0%
e56	0	7214	53834	61048	0.5%
EPSRC Total	2487485	2843134	2424403	7755022	65.5%

n01	518463	434914	630522	1583899	13.2%
n02	198312	349059	441871	989243	8.2%
n03	63723	65622	297226	426571	3.5%
n04	102618	240425	41699	384742	3.2%
NERC Total	883116	1090020	1411318	3384455	28.1%

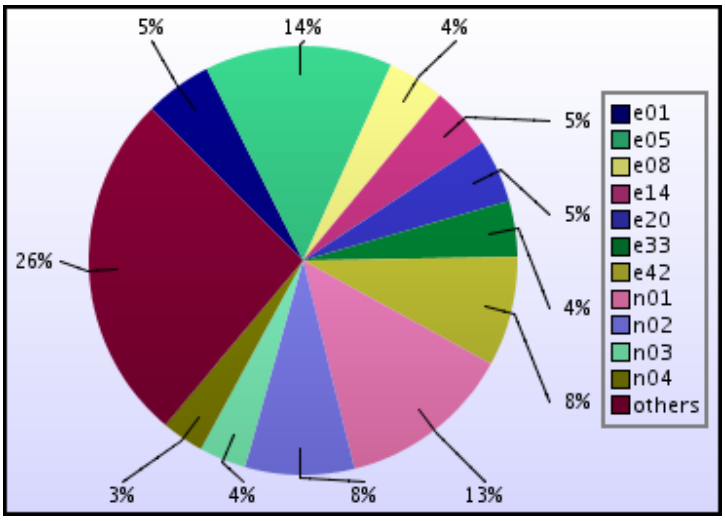
p01	62063	12071	84310	158443	1.3%
PPARC Total	62063	12071	84310	158443	1.3%

c01	128773	104260	76600	309634	2.6%
CCLRC Total	128773	104260	76600	309634	2.6%

b08	170	4810	7109	12089	0.1%
BBSRC Total	170	4810	7109	12089	0.1%

x01	74669	12906	7320	94896	0.8%
External Total	74669	12906	7320	94896	0.8%

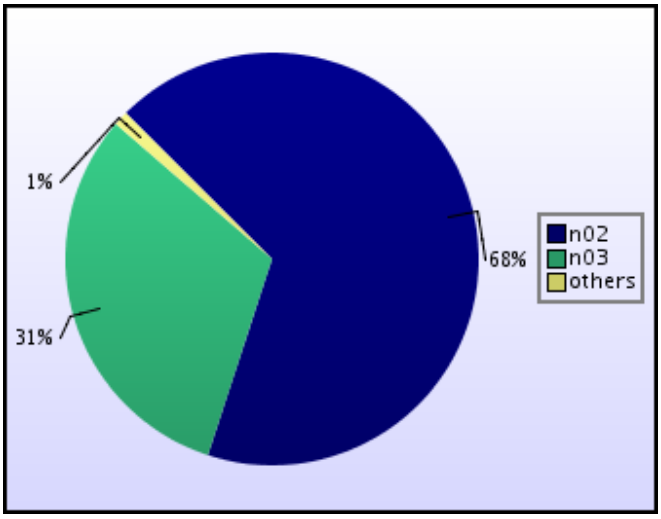
z001	14593	61013	45165	120771	1.0%
z002	5	1	1	6	0.0%
z004	7056	2701	694	10450	0.1%
z06	0	0	58285	58285	0.5%
HPCx Total	21654	63714	104144	189512	1.6%



Development service

Consortium	*July*	*August*	*September*	*Quarterly*	*%age of charged AUs*
n01	5289	958	935	7182	0.4%
n02	273019	473041	475474	1221534	67.5%
n03	170337	218019	178516	566872	31.3%
n04	0	5633	7137	12770	0.7%
NERC Total	448645	697650	662062	1808358	100.0%

z001	0	843	0	843	0.0%
HPCx Total	0	843	0	843	0.0%



3.5.1 Discounts

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

Consortium	*AUs Used*	*AUs Charged*	*Discount*
e05	1739001	1704536	34464

3.6 Helpdesk

3.6.1 Classifications

Category	*Number*	*% of all*
Administrative	125	43.9
Technical	134	47.0
In-depth	25	8.8
PMR	1	0.4
TOTAL	285	100.0

Service Area	*Number*	*% of all*
Phase 1/2 platforms	272	95.4
Website	3	1.1
Other/general	10	3.5
TOTAL	285	100.0

3.6.2 Performance

All non-indepth queries	*Number*	*%*	*Target*
Finished within 24 Hours	197	76.1	75%
Finished within 72 Hours	254	98.1	97%
Finished after 72 Hours	5	1.9	

Administrative queries	*Number*	*%*	*Target*
Finished within 48 Hours	120	96.0	97%
Finished after 48 Hours	5	4.0	

3.6.3 Experts Handling Queries

Expert	*Admin*	*Technical*	*In-Depth*	*PMR*
epcc.ed.ac.uk	88	45	12	0
dl.ac.uk	6	40	11	0
Sysadm	30	48	2	1
Other people	1	1	0	0

3.7 Service Quality Tokens

Date	*Person*	*Value*	*Comment*	*Status*
Aug 4, 2006 12:02:48 PM	Dr Jeffrey M Chagnon	* * *		

4 Support

4.1 Applications Support *(Dr David Henty)*

4.1.1 Documentation

There have been no significant updates to online documentation this quarter. We have kept users up-to-date with service developments via six user mailings.

4.1.2 Technical Reports

Three reports were planned for Q3 in the following areas:

a) Symmetric Multi-Threading
b) Serial Performance of Applications on Power5
c) Eigenvalue Solver Performance on HPCx

We have produced the following three reports this quarter:

- **HPCxTR0607**: *Optimising applications performance with compiler options for the Power5 on HPCx*, C. Maynard.
- **HPCxTR0608**: *Performance of ScaLAPACK's new PDSYEVR eigensolver*, A. Sunderland.
- **HPCxTR0609**: *OO performance optimisations of HPC applications*, S.Booth.

Report (a) was already delivered in Q2 as HPCxTR0604; reports **07** and **08** correspond directly to the titles (b) and (c).

In the Annual Plan we proposed developing material covering advanced techniques in OpenMP for performance optimisation. After some discussion we decided a more immediately useful topic for users was to look at performance optimisation techniques for OO languages. This is an area of wide applicability given the expanding use of C++ in HPC codes, and a growing interest in the potential use of Java. Although OO techniques can greatly enhance a code's robustness and maintainability, some features can compromise performance. Report **09** takes a code kernel as a case study of the pros and cons of using different OO approaches in scientific programming, and comes up with general recommendations for the design and implementation of OO HPC applications.

We have now produced nine reports this year, exceeding the target in the plan.

There are three report titles remaining in the Annual Plan for 2006:

d) HPC Software Survey

e) Performance Portability of Capability Codes

f) Achieving Capability Incentives for HPCx Applications

Report (d) is being implemented as an online user survey, available to users via the HPCx administrative web pages using new technology recently developed by the Software Engineering team. The design of the survey is almost finalised, and we plan to roll it out in October. Report (e) is nearing completion, and plans are in place for production of the final report (f) which will be based on case studies from the work the Terascaling team has done on user codes throughout the year.

By the end of the year we plan to deliver at least twelve technical reports, exceeding the total in the plan, and we will also cover all proposed subject areas.

4.1.3 Training

In Q3 of 2006 we ran the following course at Daresbury:

- **28 September:** *An Introduction to HPCx*

The course was attended by eight users (five registered and three future). Statistics are summarised below alongside annual targets (where appropriate):

Metric	Total	Target
Course days	21	25
Different course titles	7	8
Different locations	2	4
Student-days for HPCx users	215	
Student-days for HPCx staff	41	
Student-days available for HPCx	434	

There was a small issue with how some courses and attendees were classified in the May monthly and second quarterly reports – this has been addressed in the table above.

For the final quarter of the year we have plans for six more days of training, including courses at two new locations (Belfast and Rutherford Laboratory). We therefore expect to meet all the training targets for 2006.

4.1.4 Workshops and Conferences

The second and final workshop for 2006, the *HPCx User Workshop on ab initio plane-wave materials science codes in high performance computing*, took place at eSI (Edinburgh) on 3 October and attracted twelve delegates (four HPCx staff). Following an overview by M Plummer, the workshop had three major talks, given by expert users/developers, on the current state and use of the codes CASTEP (M Probert, York), CPMD (A Curioni, IBM) and VASP (D Alfe, UCL). There were also more HPCx-specific talks on the optimisation of CASTEP and VASP (M Plummer and K D'Mellow). The final session was a wide-ranging

discussion over various topics such as massive parallelism, new technology and the implications of e-science, the relative merits of mixed MPI/OpenMP coding and (MP_SHARED_MEMORY=yes) pure MPI and/or MPI-2 coding. The discussion ended with a diplomatic assessment of whether comparative benchmarks between the codes are possible.

Our major conference for 2006 was the *Fourth HPCx Annual Seminar: Moving Science Forward*, held on 4 October at eSI in Edinburgh. The event was a great success with over 70 delegates from a wide range of institutions and backgrounds. The programme demonstrated the breadth of research done on HPCx with user talks on applications areas including molecular modelling, engineering, biology, materials and fusion. All the speakers were using substantial amounts of HPCx time, with some runs being performed on over 1000 CPUs. The quality of the presentations was extremely high and clearly demonstrated the impact of HPC systems in leading-edge scientific research.

There were also talks from vendors: IBM spoke on the advantages of the new SMT features of the Power5 chip and Cray outlined the science being done on their current XT3 machine and the future roadmap towards petascale computing. HPCx staff gave presentations on the performance of the DL_POLY package and on the usage trends of HPCx over its four-year history.

See http://www.hpcx.ac.uk/about/events/annual2006/ for further details including online copies of the presentations.

4.1.5 User Group

The second and final HPCx User Group for the year was held immediately after the Annual Seminar. A major aim was to inform users of the plans for HPCx Phase 3, and the upgrade timeline and proposed configuration were presented in some detail. Users were asked for their comments on the setup of the batch queues, particularly whether it alleviates the issues regarding turnaround of capacity jobs. Input was also sought for the 2007 annual plan including suggestions for topics of technical reports and locations of training courses.

4.1.6 Newsletter

Production of the second and final edition of Capability Computing for 2006 is well underway. It will be ready in time for Supercomputing in Tampa, where it will be distributed at the HPCx booth in order to promote the HPCx service to the international HPC community.

4.1.7 Packages

We now support some 60 pieces of software under the package mechanism. Recent additions include a suite of performance analysis tools from IBM's Advanced Computing Technology Center and a new release of the HDF5 IO library.

4.2 Outreach Activities *(Dr Richard Blake)*

4.2.1 Outreach to Life Sciences

This area of outreach is now reaching a conclusion with remaining work focussing on tidying up various projects.

- The retina code has been parallelised with reasonable performance on a relatively modest dataset size. The model deals with the possibility of different spatial distributions of cells within each layer and arbitrary couplings of cells between and within layers. The cells are decomposed into 2-dimensional columns using the grid decomposition code METIS. The main post-doc working on the project has moved from ICSTM to Sheffield but is continuing to develop the model to include the peripheral retina, where anatomical structure is more complex leading to much larger simulations.
- A Final Report on the HPCx-IBM Lifesciences project has been prepared and highlights of the various projects are being developed into web-pages.

4.2.2 Other Outreach Activities

Public Awareness

Following on from the success of the talk at the Edinburgh International Science Festival (reported last quarter) we have been working on a proposal for a more formal and unified outreach programme for HPC, aimed in particular at school and general public audiences. The principal aim is to generate an increased interest in the natural sciences by highlighting the link between traditional computing subjects and computational science methods that are becoming an increasingly important tool across the all the sciences. Funding is being sought for the construction of an interactive demonstrator that will illustrate fundamental HPC concepts and applications in a visually exciting manner.

Industrial Outreach

The industrial marketing of HPCx continues with recent presentations to Nexia Solutions, Solvay, Pilkington and AWE. The major interest continues to be in exploiting Grid based cluster systems to simulate many small to medium-scale applications.

Explicit demand for HPCx time to fund contract research appears to be limited. Dr C Pickles, the main industrial marketing resource, left CCLRC in August. A recruitment exercise to replace him is underway and a report on activities to date will be completed in October.

Grid/Visualisation

The CCP1 and DLV GUIs are being developed to incorporate remote job submission via the Grid, graphical display of job status and visualisation of output in support of range of materials and life-science simulation codes.

We are working with the UKAAC Consortium to visualise dynamic simulations of flow around helicopter rotors and flow through engines to be completed in Q4.

Support has been provided to Peter Coveney for various demonstrations at the All Hands Meeting in September 2006 and for demonstrations at Supercomputing 2006. HPCx will be involved in the ETF07 experiments with the Steering Committee due to meet in October 2007.

Experiments with advanced reservation on HPCx will be deferred into next year following the establishment of the Phase 3 service and development of a project plan with interested Consortia – in particular the Reality Grid Project.

4.3 Terascaling Applications *(Dr Martyn Guest)*

4.3.1 Key Objectives

Progress against key objectives:

SMT testing

The use of Simultaneous Multithreading (SMT) on HPCx has been investigated across a wide range of codes. A. Gray presented a talk at ScicomP in Boulder, Colorado, on the Benefits of Upgrading from Power4 to Power5 Technology, which included the results of our investigation into SMT.

Publications

Mark Bull is currently working on a report on the serial efficiency of a range of codes on the Phase 2A system. Results have already been collated by Terascaling team members for most of the major codes on the system.

Capability Incentives

No additional codes achieved incentives this quarter.

4.3.2 Computational Materials Codes

CASTEP: M. Plummer has carried out further bug-fixes to the CASTEP executables. Experiments have been performed (with K. Refson) on the use of FFTW version 3. Whereas the in-house FFT routine (CASTEP-customised GPFA) outperformed FFTW version 2 (and 'all' other FFT library routines including ESSL), tests are indicating that FFTW version 3 now outperforms the 'in-house' routine. Certain irregularities with the FFTW3-CASTEP combination cause jobs to crash on occasions and these problems are being eliminated so that FFTW3 can feature in the next HPCx release of CASTEP.

An ongoing CASTEP project in which electron bands are grouped together in blocks for MPI_Alltoallv data redistribution should be completed in the next quarter. The agreed aim is to incorporate this modification in the top-level CVS version of CASTEP: significant changes have very recently been made to this version by the CASTEP Developers' Group (particularly to the 'wave' module which is directly related to the band-blocking work) which require further study. The project also requires careful testing and comparison with the current treatment as it shifts the MPI_Alltoallv optimisation away from the (well-understood) latency-bound regime towards the bandwidth-bound regime and optimum blocking-levels need to be investigated thoroughly.

CRYSTAL: Over the last quarter I.J. Bush has achieved a marked improvement in the parallel performance of MPPCRYSTAL for medium-sized systems. The scaling of such systems is very important for many users as geometry optimisations are possible for such sized systems. This work will enable these users better to exploit HPCx.

This has involved completing work to exploit the coarse-grained parallelism which is available across spin and k points, and then optimising it for parallel performance on HPCx. The main issue for optimisation is one of load balancing. The problem is that while the majority of k points require complex arithmetic, at some, as a result of symmetry, they can be solved using purely real algebra. Obviously such k points are somewhat cheaper than the complex ones, and on HPCx it was found that real k points take a little under half the time of complex k points. The code can now exploit parallelism at three levels (image, k point/spin and basis functions), a hierarchy that much improves the scaling performance on HPCx (shown below).

The new code will be part of the CRYSTAL2006 release. As an example the scaling for pyrope, a mineral which is a member of the garnet family, is shown below. The calculation used the B3LYP density functional. The unit cell contains 160 atoms, and 3040 basis functions are used to expand the wave-function. 4 k points are used. All times are for a complete SCF calculation. It can be seen that the scaling is very much better, and on 128 processors the new code is around 40% quicker. The times for the calculation are shown in Table 1 and speed-up is presented in Figure 1.

Number of processors	*Original code*	*Updated code*
8	1733	1733
16	964	936
32	601	510
64	441	312
128	380	224

Table 1: Execution times for a complete SCF calculation for pyrope using the original and updated versions of MPPCRYSTAL

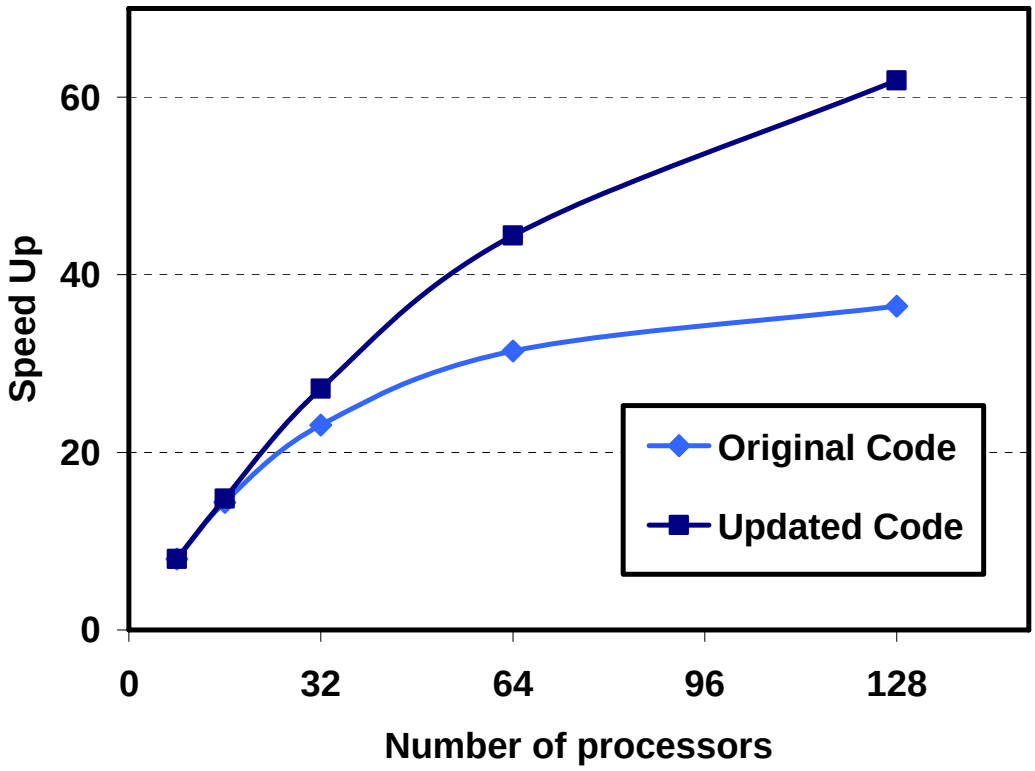


Figure 1: The improvement of scaling of the pyrope calculation using the original and updated versions of MPPCRYSTAL on HPCx, as shown by speed-up relative to 8 processors.

KPPW: A.G. Sunderland has updated the multi-band Quantum dot code, KPPW, to call PZHEEVR, which is the ScaLAPACK implementation of the new MRRR algorithm for double complex Hermitian matrices, and its 'expert' driver routine, PDSYEVR. PDSYEVR is particularly suited to this application as it can be used to calculate the required small subset of eigenvalues at high speed. Timing analysis of PDSYEVR on application matrices indicates that performance improvements of up to 100% can be obtained in comparison to the existing ScaLAPACK 'expert' driver PDSYEVX (see the Eigensolvers section below). We are currently evaluating its effect on the performance of the KPPW code.

VASP: We have continued to investigate and improve the performance of VASP on HPCx Phase 2a, in an attempt to increase both efficiency and scalability. In particular, K. D'Mellow has investigated the use of SMT, and the application of the new MRRR algorithm for matrix diagonalisation which is available in a beta-release of the ScaLAPACK library. This work was presented at the HPCx Workshop on Ab-initio Plane-wave Materials Science Codes in HPC, 3rd October, e-Science Institute, Edinburgh, 2006.

4.3.3 Computational Chemistry Codes

DFT studies: Density Functional Theory (DFT) has emerged as a leading method in the study of chemical systems at the electronic structure level. It is therefore of considerable interest to users interested in studying large biological systems with hundreds of atoms to gain an assessment of the DFT capabilities on HPCx, with its large processor counts and large node memory (28 Gigabytes). In Q3 G. Fletcher has been examining the performance of several quantum chemistry packages offering DFT modules. Of particular interest are those combining fitted-DFT (for non-hydrid functionals) with *in-core* storage of integrals in a scalable way, since these are likely to yield the fastest solution times as problem sizes grow.

Among these, GAMESS-UK (developed by the Computational Chemistry Group at Daresbury Laboratory) and NWChem (maintained at Pacific Northwest Laboratory) have efforts focused in the area of parallel scalability, while Turbomole, a leading commercial package, is focused more on optimising sequential performance but is implemented in parallel for certain platforms. All three packages implement different strategies regarding the screening, computation, and storage of integrals, and the solving and convergence of the self-consistent field (SCF) equations. With different levels of parallel scalability, memory requirements, and numbers of iterations, the overall *time-to-solution* provides a convenient way to compare performance.

To benchmark the three codes, a series of chemical systems corresponding to zeolite clusters (silicon oxide structures terminated with hydrogen) were chosen, starting at around thirty atoms and increasing in size. Results so far indicate that, on a strict *wall-clock* comparison, GAMESS-UK provides the fastest time-to-solution, followed by NWChem, though significant multiprocessing is required (up to 128 processors).

In purely *direct* calculations, those without in-core storage of integrals, calculation costs (in terms of aggregate clock cycles as a function of system size) scale as $n^{2.38}$ (where n is the number of atoms) for GAMESS-UK, compared to $n^{2.75}$ for NWChem, reflecting the impact of faster integral screening and computation methods in the former code. The corresponding in-core results yield $n^{2.92}$ for GAMESS-UK, compared to $n^{2.13}$ for Turbomole.

As noted above, fully in-core calculations require high processor counts to provide sufficient memory and in this processor range the scalability of our current eigensolver technology is known to be fairly flat. While scalability clearly is a key strength of codes such as GAMESS-UK these results highlight areas of improvement and efforts are underway in the Computational Chemistry Group to incorporate the ScaLAPACK eigensolver which exhibits superior performance at high processor counts. These studies also form part of broader survey of the performance of quantum chemistry codes on HPCx which is currently underway, the results of which will be made available in Q4.

GAMESS-UK: H.J.J. Van Dam has carried out further optimisation of the HPCx implementation of GAMESS-UK. This has included a new build configuration in which both the Global Array tools (GAs) and ScaLAPACK linear algebra routines are available. Previously, the GA version was built using TCGMSG as the message passing library and made use of the PeIGS eigensolver. The work involved programming alternatives to some TCGMSG convenience functions and adapting the whole program to work with 32-bit integers (for compatibility with ScaLAPACK and MPI). This change allows a single version of GAMESS-UK to make use both of GA functionality and the more efficient modern parallel linear algebra available in the ScaLAPACK library. PeIGS (using MPI message passing) is also available within the build and we are now doing some performance comparisons. The development also allows a single binary to support all parallel functionality whereas before two binaries were needed.

SIESTA: We have been in contact with the developers over benchmarking the latest version of SIESTA.

4.3.4 Physics Codes

CENTORI: J. Hein has continued to work with UKAEA to develop the 3D parallel version of the CENTORI code. He has developed, tested and documented the communication routines (which allow distribution of data from master to workers and collection back from the workers to the master) for a three dimensional data distribution.

Ludwig: K. Stratford presented his work on the Ludwig code at the International Conference for Mesoscopic Methods in Engineering and Science (ICMMES) in Virginia, 2006

4.3.5 Engineering Codes

Fluent: A.G. Sunderland has compiled a report summarising the Fluent benchmarking exercise he has undertaken on HPCx over recent months. This has been sent to the HPCx Fluent user base and to the Fluent developers team.

ISOJET: K. D'Mellow and K. Stratford have worked on identifying and resolving memory leaks and runtime issues with the ISOJET code - one of the many codes new to the system from ex-CSAR users - resulting in a list of MPI and POE related upgrades and fixes to the system (in conjunction with related investigations on other codes within the Terascaling team). We also supplied the users with a fix to the code.

4.3.6 Life Sciences

LAMMPS: Early in this quarter, a user reported a significant performance degradation of the LAMMPS molecular dynamics code. LAMMPS was found to perform 10-12 times slower than expected when using 16 or more processors. F. Reid determined a workaround which enabled users to continue using the code (either under-populating nodes or turning on SMT resolved the problem) whilst

the reason for the performance degradation was fully investigated. A large number of different tests were performed including examining different environment variables, possible zombie processes, different versions of libraries/LAMMPS, changing interrupt frequencies etc. Profiling of the code suggested that the additional time was being spent within the MPI sections of the code pointing toward a possible problem with the MPI libraries. The MPI libraries were also identified to cause problems in a number of other codes and were subsequently updated around 4th August. This update resolved the LAMMPS performance problem. LAMMPS users were kept informed of developments via email throughout.

GROMACS: A number of reports on the GROMACS mailing list suggest problems have been experienced when attempting to compile GROMACS on the POWER5 system. F. Reid solved these problems by invoking the cross-compiling options under configure. The latest version of GROMACS (version 3.3.1) has been installed and tested on HPCx.

4.3.7 New Applications

OpenBEAGLE: Installation of the OpenBEAGLE package, which is an Advanced Genetic Learning Environment, has been requested by F. Tekiner of the Supercomputing Data Mining project (e51). M. Ashworth has found that compilation of the package fails due to an Internal Compiler Error in IBM's C++ compiler. A PMR has been opened with IBM for resolution of this issue.

K. Stratford and L. Smith visited Dr David Ingram, from one of the new consortia on HPCx (e54: Free surface simulation of waves overtopping during storms) to discuss further parallelisation strategies and check-pointing issues.

4.3.8 Environmental Codes

Unified Model: Problems with the MPI libraries were also identified as the cause of serious problems with the UKMO Unified Model (see the discussion of LAMMPS above). This required special work by IBM. The problems were finally resolved by an additional patch, which was installed on 14-16 August.

4.3.9 Libraries

Eigensolvers: A.G. Sunderland has received several new versions of the pre-release PDSYEVR routine from the ScaLAPACK developers for testing. The performance of the new routines on HPCx has generally shown improvements for a test suite of matrices with a range of eigenvalue spectra. This work has been written up as the HPCx Technical Report *Performance of a New MRRR-based Parallel Eigensolver on HPCx*.

PESSL: A.G. Sunderland has analysed the performance of a new PESSL library on HPCx. This new version includes improved BLACS routines for message passing and added functionality, such as a PESSL version of the divide-and-conquer eigensolver PDSYEVD. Timing results show that the performance of

numerical routines within PESSL is now consistent with their ScaLAPACK counterparts.

PETSc: In response to requests from users in the Integrative Biology project, K. Stratford has built and installed a new version of the PETSc library (version 2.3.1), which is now available in the packages directory.

4.3.10 Tools

HPCT: An updated version of IBM's HPC Toolkit has been obtained from IBM's Advanced Computing Technology Center (ACTC) and installed in the packages directory. C. Maynard and J. Hein are currently evaluating the tools, a process hindered by the lack of documentation from IBM, with a view to updating user documentation.

4.3.11 Talks and Publications

1. *"Electron collisions with Fe-peak elements: Forbidden transitions between the low lying valence states $3d^6$, $3d^5 4s$, and $3d^5 4p$ of Fe III",* B.M. McLaughlin, M.P. Scott, A.G. Sunderland, C.J. Noble, V.M. Burke, C.A. Ramsbottom, R.H.G. Reid, A. Hibbert, K.L. Bell, P.G. Burke, Atomic Data and Nuclear Data Tables, 2006 , in press

2. *"Performance of a New MRRR-based Parallel Eigensolver on HPCx",* A.G. Sunderland, HPCx Technical Report

3. "*Performance Benefits from Upgrading from Power4 to Power5 Technology*", A. Gray, ScicomP 12, Bolder, Colorado, July 2006

4. *"Application Performance on High-end and Commodity-class Computers*", M. F. Guest and C.A. Kitchen, Tyndall National Institute, Cork, 31st August 2006.

5. "*Application Trends on HPCx*", L. Smith, the Fourth HPCx Annual Seminar, e-Science Institute, Edinburgh, 4th October 2006

6. "*The DL_POLY Package: Scalability and Performance*", I. Todorov, the Fourth HPCx Annual Seminar, e-Science Institute, Edinburgh, 4th October 2006

7. "Simulation of Steady Shear Flows in Binary Fluids", K Stratford, the International Conference for Mesoscopic Methods in Engineering and Science (ICMMES), Virginia, 2006

8. Three presentations were made at the HPCx Workshop on Ab-initio Plane-wave Materials Science Codes in HPC, e-Science Institute, Edinburgh, 3rd October 2006:
   - "Common features of plane-wave codes", M. Plummer

- “HPCx-specific optimisations: VASP”, K. D’Mellow
- “HPCx-specific optimisations: CASTEP”, M. Plummer

4.4 Software Engineering *(Dr Stephen Booth)*

4.4.1 Future Look and Strategic HPC Support

OO techniques

Historically HPC applications have been relatively simple software systems. However, the trend over recent years seems to be towards more and more complex programs. This introduces significant problems for the development of HPC software. Object Oriented programming is a style of programming that is intended to help manage the complexity of software systems and a number of HPC applications are now being written in Object Oriented languages such as C++ rather than traditional procedural languages such as C or Fortran. Unfortunately object Oriented languages have a reputation for introducing additional code overheads that result in slower code than procedural languages.

We have therefore been investigating the impact of Object-Oriented techniques on code performance and have produced a technical report outlining the performance issues associated with Object Oriented programming and outlining how applications can be designed to gain the advantages of the OO programming style yet still obtain high levels of performance.

HPCxTR0609: *OO performance optimisations of HPC applications*, Stephen Booth

Low-level Investigations and Analysis

We have continued to investigate and analyse the behaviour of the hardware and software used to provide the HPCx service.

Recent investigations have included a study of the performance of the new Java5 software running on HPCx. This is a new release of Java for AIX implementing the 1.5 version of the Java language.

We have also been investigating the the performance of mixed-mode programming on HPCx. Mixed mode programming is where the a single application is parallelised using a combination of OpenMP and MPI. This work included an analysis on the impact of SMT on the effectiveness of mixed-mode programs.

This work is still ongoing and we will produce reports on the results of our investigations in the next few months.

POWER5: We have completed a number of investigations into the effectiveness of the Power-5 processor used in the Phase 2A and Phase 3 systems. This information has been collated in two technical reports in order to provide

guidance for users of HPCx on how to make the best use of this hardware. The first of these, on the use of the SMT feature of the ot POWER5 processor, was delivered in Q2 (HPCxTR0604). We have followed this with a report on the affect of compiler flags and versions on the performance of the Power-5 processor.

HPCxTR0607: *Optimising applications performance with compiler options for the Power5 on HPCx*, Chris Maynard

Software: A significant amount of effort was expended this quarter investigating a number of apparent memory leaks associated with a recent upgrade to the IBM MPI libraries and drivers (see the discussions of LAMMPS and the Unified Model in sections 4.3.6 and 4.3.8 above).

In addition to producing a number of test cases based on user codes to be passed on to the IBM MPI developers we also produced a tool that uses the MPI profiling interface to help track the allocation of memory within MPI calls.

This work resulted in at least two independent memory leaks in the MPI library being reported to IBM. Fixes for these problems are now available.

4.4.2 In-depth Software Support

User administration software

The user administration software (SAF) has been extended to allow greater control over the allocation of user/group names and UID/GID values. This change is required for full integration of HPCx into the wider DEISA project.

In order to support multiple sites with shared GPFS access DEISA requires user/group names to conform to a particular pattern (depending on the creating site) and UID/GID values to be allocated from particular ranges (depending on the creating site). By default the SAF allocated GIDs and UIDs from a single range, allows any legal non-allocated username and generated group names based on the enclosing project. This remains the default but can now be overridden on a project-by-project basis so that new groups and users created in the DEISA project conform to the wider DEISA conventions.

4.5 Operations and Systems *(Mr Mike Brown)*

4.5.1 Phase 3 upgrade

The principal task undertaken was preparation for the Phase 3 upgrade, with p5-575 nodes being brought on line for integration and testing as they became available. A proto LoadLeveler configuration was devised that it is hoped will enable a smooth take-up of the additional capacity.

4.5.2 Incidents

During the period there were 26 recorded incidents on the service, with 4 SEV1 (major incidents) recorded.

4.5.3 Maintenance sessions

5 maintenance sessions were undertaken, principally to enable IBM to perform preparation work for the Phase 3 upgrade.

4.5.4 Outreach

M W Brown and T J Franks attended the meeting of the IBMUKHPCUG at ECMWF on 25 September.

4.5.5 Staffing

There were no staff changes in the quarter, although Ian Shore resigned with his departure taking place later in the year.

4.6 Staffing

AV	*July*	*August*	*September*
DL	5.4	5.4	5.4
EPCC	9.1	9.5	8.4
Total	14.5	14.9	13.8

Systems	5.5	5.8	6.0

5 Summary of Performance Metrics

Metric	*TSL*	*FSL*	*July*	*August*	*September*
Technology serviceability	80%	99.2%	100.0%	99.9%	100.0%
Technology MTBF (hours)	200	300	∞	732	∞
Number of AV FTEs	7.5	10	14.5	14.9	13.8
Number of training days per month	22.5/12	30/12	21/7	21/8	22/9
Non in-depth queries resolved within 3 days	85%	97%	98.9%	98.8%	96.3%
Number of A&M FTEs	3.75	5.75	5.5	5.8	6.0
A&M serviceability	80%	99.6%	97.0%	100.0%	99.4%

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

Code	*Class*	*Title*	*PI*
e01	1	UK Trubulence Consortium	Dr Gary Coleman
e05	1	Materials Chemistry using Terascaling Computing	Prof Richard Catlow
e06	1	UK Car-Parrinello Consortium	Prof Paul Madden
e07	2	Turbulent Plasma Transport in Tokamaks	Dr Colin M Roach
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
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

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

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