

HPCx Quarterly Report

January – March 2007

1 Introduction

This report covers the period from 1 January 2007 at 0800 to 1 March 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

- The first full quarter since the Phase 3 upgrade was an exceptional one for reliability and for utilisation.
- There was only a single failure in this quarter which is the best since the start of service. Moreover, this single failure was due to an external network problem.
- Utilisation on the main service continues to increase and this quarter included a couple of record months. Overall, for the quarter, the utilisation was almost 82%. As a consequence, capability utilisation was somewhat squeezed.
- The utilisation of the Development system was very high at well over 90%. This high utilisation is possible due to the relatively small job sizes on this system.
- The helpdesk met all its targets throughout this quarter and the various monthly service metrics were generally good.
- As well as making progress against the targets for courses and technical reports, we ran a workshop on the new Cell processor, a novel system possessing internal 8-way parallelism on a single chip. IBM also ran a course on their HPC toolkit which was beneficial both for the users and the HPCx staff who attended.
- As part of our Public Understanding of Science activity, EPCC will be giving a talk in the Cutting Edge theme of the Big Ideas programme at the 2007 Edinburgh International Science Festival and we will be promoting the national HPC services at this.

- The Software Engineering team have produced a Technical Report on their investigation into network performance over SuperJANET5, This work has resulted in a useful approach to maximise the available bandwidth.
- There has also been significant work on the interactive web software called SAFE (Service Administration from EPCC). We have developed new helpdesk software and this is now integrated into SAFE. This provides significant advantages to both users and the support teams and will be rolled out on 1 May.
- The Terascaling team have had productive interactions with major consortia, including NCAS (n02) and UKCP (e06). Martin Plummer has now installed and tested a new version of CASTEP with enhanced functionality, including a new pre-conditioner.
- In collaboration with users at Southampton, Paul Sherwood has been investigating task-farm approaches for CHARMM. This resulted in an approach using MPICH as a “secondary MPI” which is now ready for testing.
- Joachim Hein has improved the serial performance of CENTORI, a code developed by UKAEA, Culham to simulate the turbulent flow of the plasma inside a fusion reactor. Early results show a 33% improvement in execution time.
- Fiona Reid assisted Professor Millar with his clock modelling code as he prepared to apply for HPCx time, providing help with optimisation and parallelisation strategies. Professor Millar has now been successful in obtaining time on HPCx.

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:

	<i>January</i>	<i>February</i>	<i>March</i>
Incidents	8	1	4
Failures	1	0	0

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

January

<i>Failure</i>	<i>Site</i>	<i>IBM</i>	<i>External</i>	<i>Reason</i>
07.005	0%	0%	100%	Network failure at Manchester

February

None

March

None

3.1.2 Performance Statistics

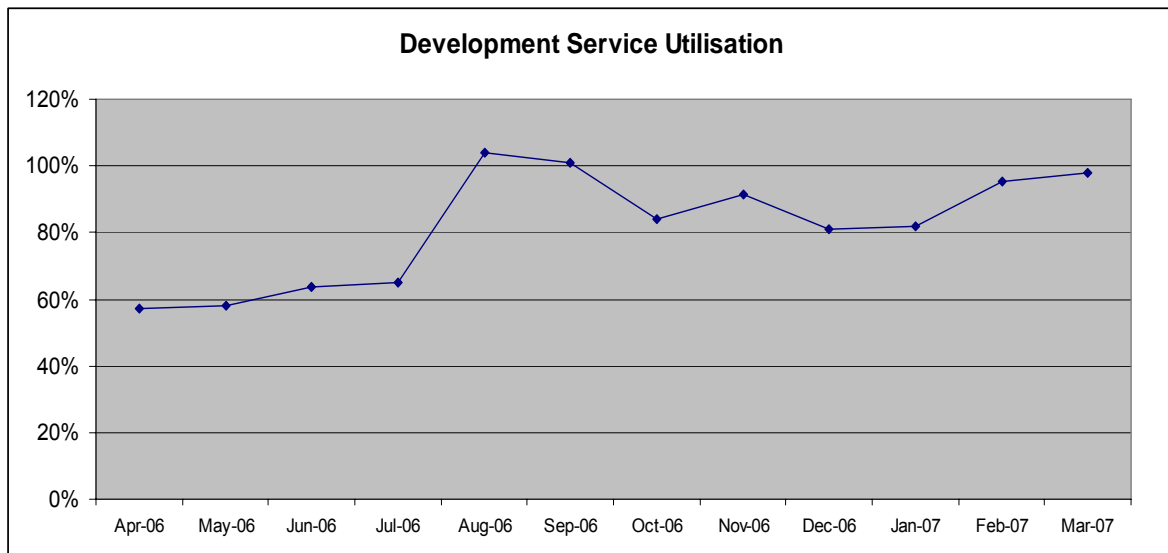
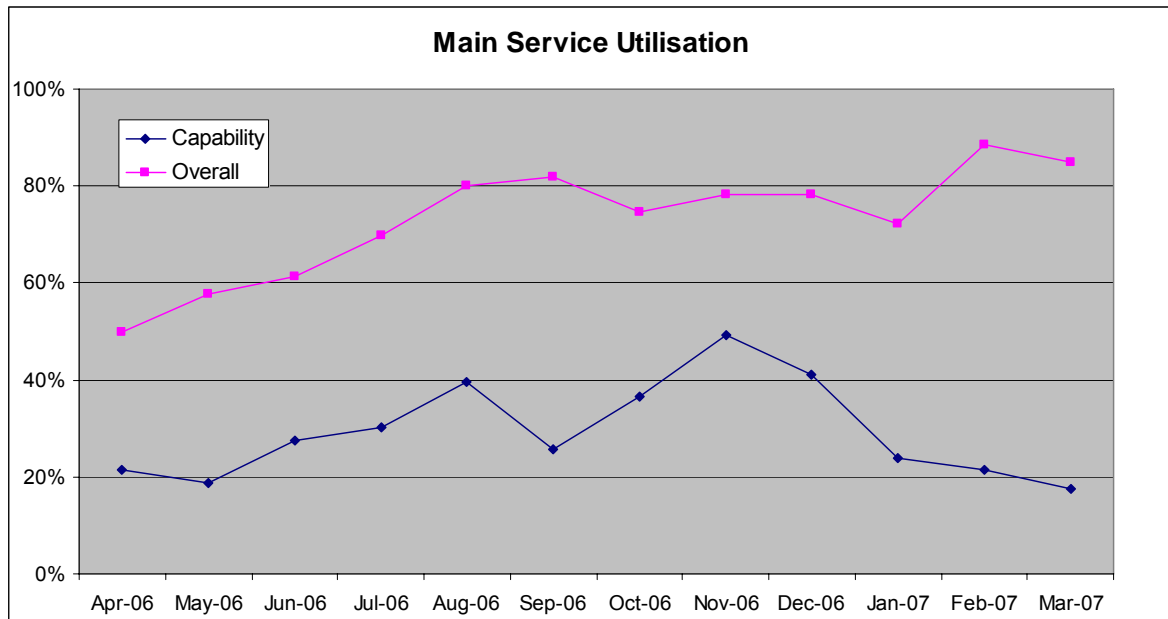
This section uses the definitions agreed in Schedule 7, ie,

- $MTBF = (24 \times 30.5) / (\text{number of failures in month})$
- $\text{Serviceability (\%)} = 100 \times (WCT - SDT - UDT) / (WCT - SDT)$

<i>Attribution</i>	<i>Metric</i>	<i>January</i>	<i>February</i>	<i>March</i>	<i>Quarterly</i>
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	1	0	0	1
	MTBF	720	∞	∞	2196.0
	Serviceability	99.9%	100.0%	100.0%	100.0%
Total	Failures	1	0	0	1
	MTBF	720	∞	∞	2196.0
	Serviceability	99.9%	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.

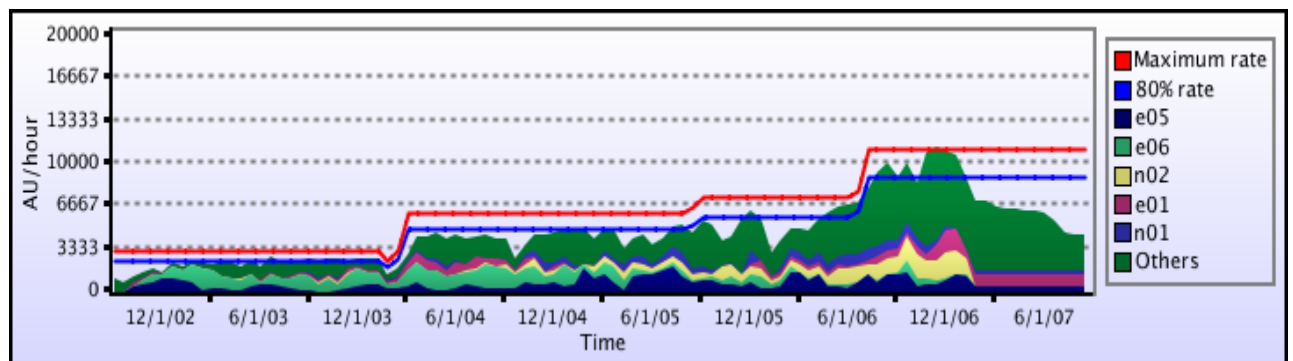


Utilisation figures greater than 100% for the development service correspond to a period in August and September 2006 when the number of processors in the service was temporarily increased.

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.

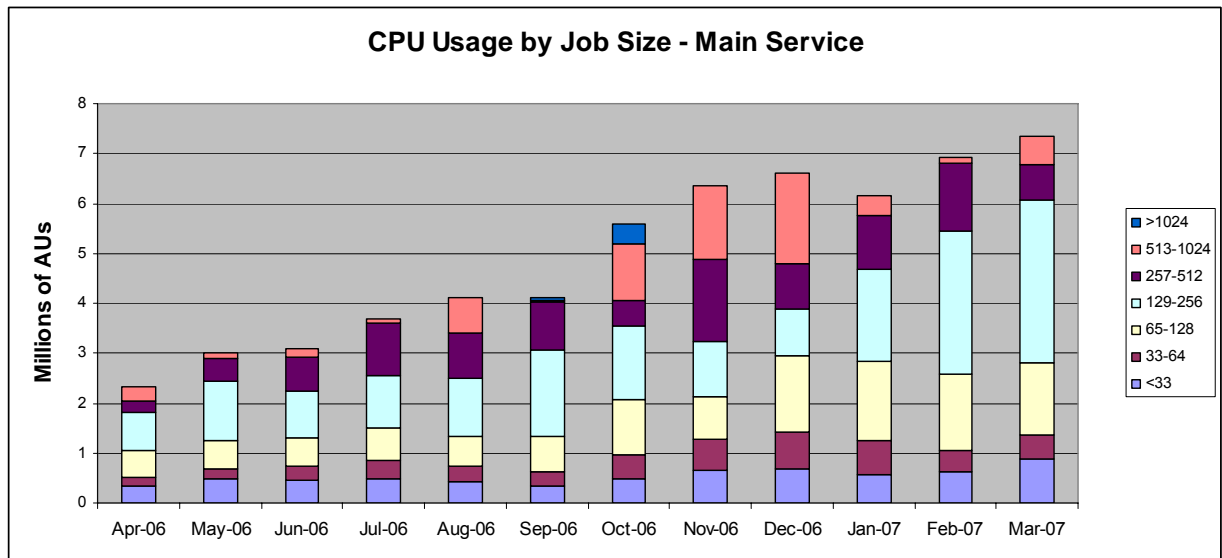
Numbers of Research Consortia

At the end of this quarter there were 49 research consortia on HPCx; another is awaiting final confirmation from BBSRC.

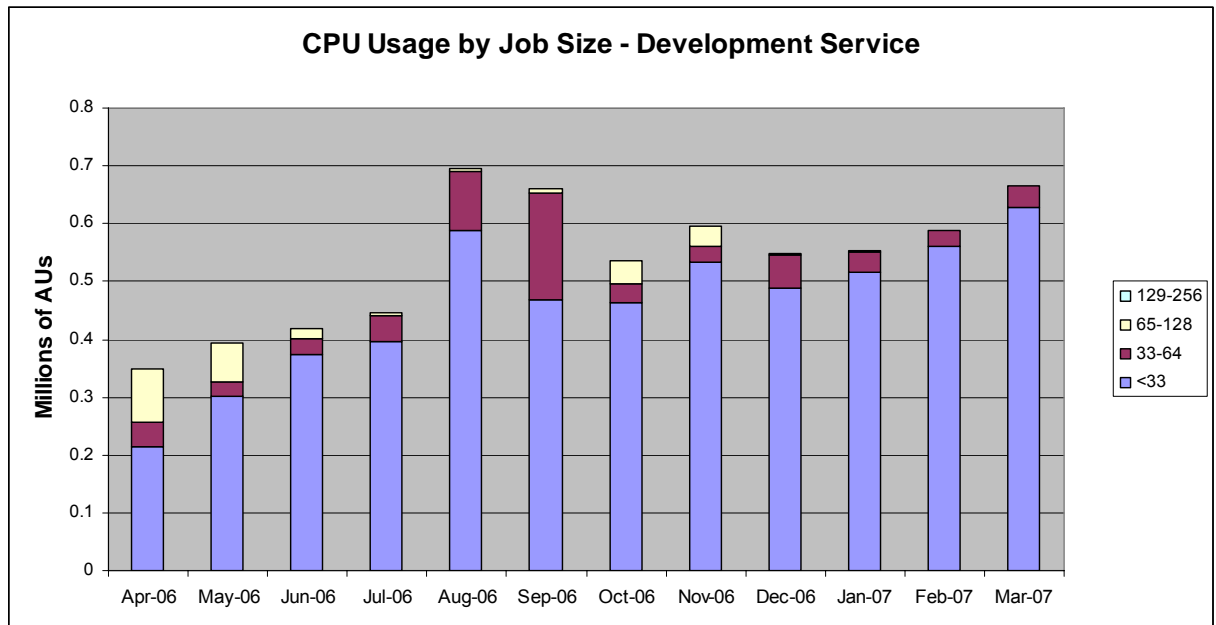
In addition, there is one active externally funded project.

3.4 CPU Usage by Job Size

Main service



Development Service



3.5 AU Usage by Consortium

Main Service

<i>Consortium</i>	<i>January</i>	<i>February</i>	<i>March</i>	<i>Quarterly</i>	<i>%age of charged Aus</i>
e01	374716	590122	1087080	2051918	10.1%
e03	182	413	11	606	0.0%
e05	607554	406834	824648	1839036	9.0%
e06	462025	169383	85169	716577	3.5%
e08	361033	34508	368493	764035	3.7%
e10	12		5650	5662	0.0%
e11	25644	16088	54840	96572	0.5%
e14	170821	31564		202385	1.0%
e17	28851	111795	115869	256516	1.3%
e18	1772			1772	0.0%
e19	278	323		601	0.0%
e20	5			5	0.0%
e21	18994	54161	25411	98566	0.5%
e23	24523	286069	90347	400938	2.0%
e24	11	106	356	474	0.0%
e25	7	6	3	16	0.0%
e26	499	9388	0	9888	0.0%
e27	25606	63234	111870	200709	1.0%
e31	2094	1242		3336	0.0%
e32	14250	75039	118290	207579	1.0%
e33	28703	77301	61237	167241	0.8%
e35	105784	370943	736865	1213592	5.9%
e36	4426	853	666	5945	0.0%
e37	1259868	1487743	959607	3707219	18.2%
e38		2516	75	2592	0.0%
e39	31897	50378	30775	113051	0.6%
e40		3913	2584	6497	0.0%
e41			1568	1568	0.0%
e42	386471	681635	53365	1121471	5.5%
e45	28645		32224	60869	0.3%
e46	2707		55	2763	0.0%
e48	1	10745	3332	14078	0.1%
e49	7308	2039	783	10130	0.0%
e50	44116	143181	20458	207755	1.0%

e51	0	939	3099	4037	0.0%
e53	155	1693	568	2415	0.0%
e54	7396	72103	244777	324276	1.6%
<i>EPSRC Total</i>	4026355	5120875	5179530	14326760	70.2%

n01	395520	293875	143247	832641	4.1%
n02	875024	579378	802806	2257208	11.1%
n03	348041	546350	819038	1713428	8.4%
n04	48340	198771	167962	415073	2.0%
<i>NERC Total</i>	1666924	1618374	1933052	5218350	25.6%

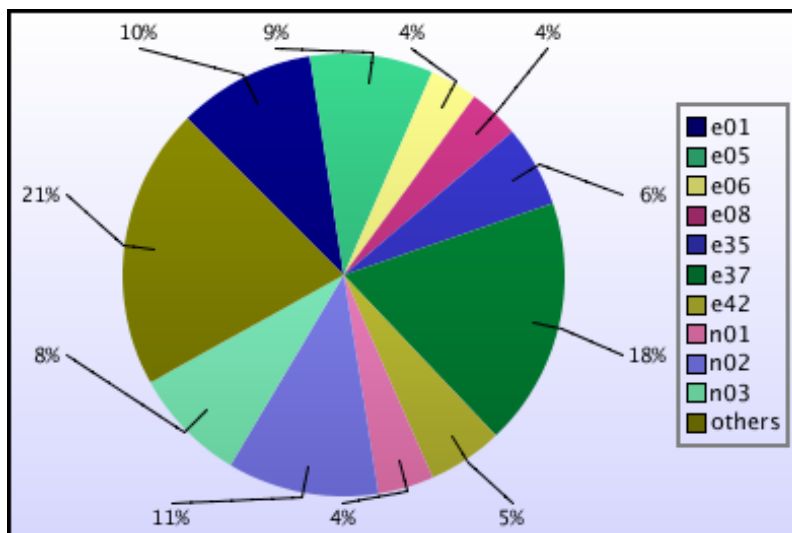
p01	33	3419	2433	5885	0.0%
<i>PPARC Total</i>	33	3419	2433	5885	0.0%

c01	313683	36755	86344	436782	2.1%
<i>CCLRC Total</i>	313683	36755	86344	436782	2.1%

b08	16244	23409	100669	140322	0.7%
<i>BBSRC Total</i>	16244	23409	100669	140322	0.7%

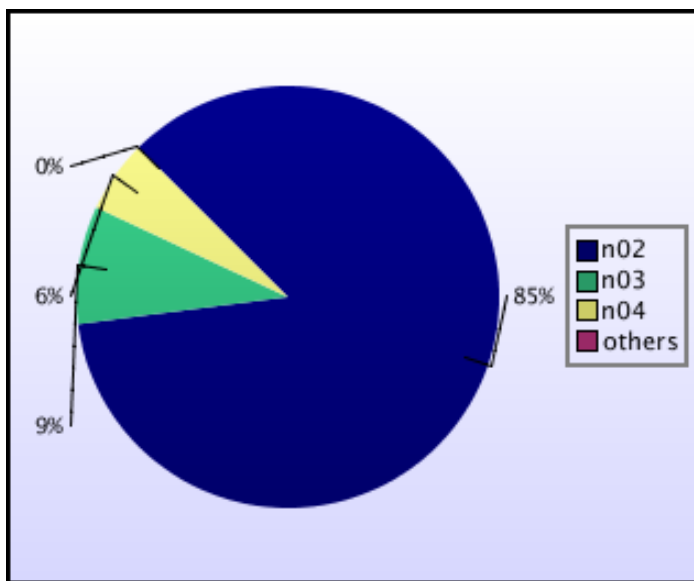
x01	145751	86476	26196	258423	1.3%
<i>External Total</i>	145751	86476	26196	258423	1.3%

z001	6126	10281	12153	28561	0.1%
z002	0			0	0.0%
z004	57	654	199	910	0.0%
z06	163	1	0	164	0.0%
<i>HPCx Total</i>	6346	10935	12356	29637	0.1%



Development service

Consortium	January	February	March	Quarterly	%age of charged Aus
n01			3	3	0.0%
n02	436242	486602	624990	1547834	85.4%
n03	79255	70453	14300	164008	9.0%
n04	39383	33473	27642	100498	5.5%
<i>NERC Total</i>	554880	590528	666935	1812343	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
c01	439202	436781	2421
e01	2052947	2051917	1029
e27	211440	200709	10730
e32	208320	207578	741
e36	5964	5944	19
e58	525278	504064	21214
n03	1732069	1713428	18641

3.6 Helpdesk

3.6.1 Classifications

<i>Category</i>	<i>Number</i>	<i>% of all</i>
Administrative	176	52.7
Technical	132	39.5
In-depth	25	7.5
PMR	1	0.3
TOTAL	334	100.0

<i>Service Area</i>	<i>Number</i>	<i>% of all</i>
Phase 1/2 platforms	250	74.9
Website	2	0.6
Other/general	82	24.6
TOTAL	334	100.0

3.6.2 Performance

<i>All non-indepth queries</i>	<i>Number</i>	<i>%</i>	<i>Target</i>
Finished within 24 Hours	240	77.9	75%
Finished within 72 Hours	306	99.4	97%
Finished after 72 Hours	2	0.6	

<i>Administrative queries</i>	<i>Number</i>	<i>%</i>	<i>Target</i>
Finished within 48 Hours	171	97.2	97%
Finished after 48 Hours	5	2.8	

3.6.3 Experts Handling Queries

<i>Expert</i>	<i>Admin</i>	<i>Technical</i>	<i>In-Depth</i>	<i>PMR</i>
epcc.ed.ac.uk	140	55	18	0
dl.ac.uk	8	35	1	1
Sysadm	27	42	6	0
Other people	1	0	0	0

3.7 Service Quality Tokens

<i>Date</i>	<i>Person</i>	<i>Value</i>	<i>Comment</i>	<i>Status</i>
Mar 12, 2007 1:58:39 PM	Dr Mohamed Maidi	* * * * *		
Feb 5, 2007 3:34:38 PM	Mr Jeremy Rabone	* * * *		
Feb 3, 2007 1:23:58 PM	Mr Oliver JH Browne	* * *		
Jan 16, 2007 10:40:35 PM	Mr Martin Vejmelka	* * * *	speed & service excellent, job waiting time not too long	

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 the tools section following new software releases. We have kept users up-to-date with service developments via four user mailings.

4.1.2 Technical Reports

Two reports were planned for Q1 in the following areas:

- New Fortran Language Features: Interfacing to Unix Shared Memory
- Data Transfer over SuperJANET5

We have produced the following two reports this quarter:

- **HPCxTR0701**: *New Fortran Features: Portable Use of Shared Memory Segments*, I.J. Bush.
- **HPCxTR0702**: *Transferring Large Files to Remote Sites: An Update*, C. Maynard.

which directly address the two planned areas.

The first report shows how the newly defined interface between Fortran and C allows Fortran programmers easily to access shared-memory segments on HPCx. This enables users to take advantage of the shared-memory nature of the computational nodes from within message-passing codes. This has the potential for significant benefits in terms of both memory efficiency and performance.

The second report revisits HPCxTR0603, written last year, which looked at data transfer speeds between HPCx and CSAR. Since then, the SuperJANET5 network has been installed and the location of the new HECToR service has been decided. It is therefore timely to look at the same general issues, but now focusing on data transfer between HPCx and Edinburgh. Although a full end-to-end test is not currently possible as HECToR is not yet installed, the report investigates the achievable bandwidths on the critical sections of the network.

4.1.3 Training

In Q1 of 2007 we ran the following three courses:

- **17-18 January**: Using the IBM HPC Toolkit (Edinburgh)

- **24-25 January:** Shared-Memory Programming with OpenMP (Belfast)
- **6-8 March:** HPC Techniques in Computational Chemistry (Edinburgh)

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

<i>Metric</i>	<i>Total</i>	<i>Target</i>
Course days	7	20
Different course titles	3	6
Different locations	2	4
Student-days for HPCx users	68	
Student-days for HPCx staff	24	
Student-days available for HPCx	167	

We already have six more days of training planned and advertised for Q2.

4.1.4 Workshops and Conferences

In the Annual Plan we committed to running two workshops this year, the first covering “novel systems or programming techniques”. This quarter we chose to run a workshop focusing on the new Cell processor, a novel system possessing internal 8-way parallelism on a single chip which also requires new programming techniques due to the novel nature of the memory hierarchy. This workshop was particularly timely as the Playstation 3, which is based on Cell technology, was launched in Europe at the end of March; blade systems containing Cell processors are also starting to become available.

The two-day workshop *Programming the Cell Processor* was run in Edinburgh on 27-28 March, with material presented by a Cell expert from IBM Germany. Although it is not possible easily to access Cell hardware, attendees were able to develop and run practical examples on their own laptops using the Cell simulator, which was made available in advance. There were 25 registrations for the workshop: 9 registered HPCx users, 2 HPCx staff and 14 HPC academics.

The main conference for the year, the Fifth Annual HPCx Seminar, will be held in Daresbury at the end of 2007 alongside the 17th Machine Evaluation Workshop.

4.1.5 User Group

The first user group of the year will be held over Access Grid some time in Q2. To increase attendance, we normally aim to present at least one talk of general interest to all HPCx users. The topic has not yet been decided, although given the timing of the User Group it might be appropriate to give information on the progress of the HECToR service.

4.1.6 Newsletter

Production of the first edition of Capability Computing for 2007 is well underway. It will be printed and mailed out to users during May, and will be ready in time for distribution at ISC'07, the International Supercomputing Conference in Dresden in June.

4.2 Outreach Activities (*Dr Richard Blake*)

Progress against the 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.*

Datasets have been generated and sequential bottlenecks are being parallelised.

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 is giving a public talk on 15th April as part of the Cutting Edge theme of the Big Ideas programme at the 2007 Edinburgh International Science Festival. The subject is a general overview of the security risks associated with the modern Internet. The UK National Supercomputer services will be promoted at the event via posters, flyers and newsletters.

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

We have submitted a note of interest to a funding line from the Scottish Executive for around £30K to support the promotion of HPC and leading-edge computational science local schools and at other relevant events.

- *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 discussed with the Proudman Oceanographic Laboratory and the plasma fusion simulation group at Culham.

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

This is planned for September 2007.

4.2.4 The Grid

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

- *evaluation of Advanced Reservation;*

Implementation of HARC is being explored on the NW-GRID systems.

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

A 10 Gb/s link has been procured; this is due to go live in May.

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

We are participating in the GENIUS grant proposal to simulate cranial blood flow in a Grid environment.

4.3 Terascaling Applications (*Dr Martyn Guest*)

4.3.1 Key Objectives

Progress against key objectives:

Publications

A technical report has been issued on the use of new features from the Fortran 2003 standard which allow access from Fortran to shared memory segments and semaphores. These can be used both to enable HPCx applications to solve bigger problems and to solve those problems more quickly. Techniques are illustrated with an example using the GAMESS-UK electronic structure code.

"New Fortran Features: The Portable Use of Shared Memory Segments", Ian J. Bush, HPCx Technical Report, HPCxTR0701

Interactions with Consortia

M. Plummer represented HPCx at the UKCP (e06) working group meeting on 30th March at Imperial College. The agenda included an HPCx Q&A session.

M. Ashworth attended a meeting of the Institute of Atmospheric Sciences (part of n02: NCAS) on 28th February at the University of Leeds. There were discussions about the viability of developing a NERC consortia application to develop a community atmospheric model and about Terascaling assistance for installing two new atmospheric codes on HPCx.

K. D'Mellow and L. Smith visited Professor Rex Godby from e40 on the 25th January 2007. There was discussion on the GW Space-Time code, which implements Hedin's GW method and how best to parallelise this.

4.3.2 Computational Materials Codes

CASTEP: M. Plummer has installed and tested the latest version, CASTEP 4.1, on HPCx. Modifications by Dr P Hasnip (following feedback from HPCx staff) to the new, scientifically preferred, 'density-mixing' minimiser and pre-conditioner combination for electronic SCF energies mean that, for small to medium jobs, this now performs more efficiently than the older, simpler but less rigorous, pre-conditioner and minimiser. In the limit of large jobs certain aspects of the new pre-conditioner still affect performance adversely and the old version is still recommended in these cases.

The 'band-blocking' software, in which multiple 3D-FFTs corresponding to several electronic bands take place 'simultaneously', has been written and tested for CASTEP 4.1. Basic band-blocking can improve performance for large test-cases with the 3D-FFT performed across HPCx 16-processor nodes (LPARs). There is a price to be paid in terms of copying data into large buffers which means that the procedure reduces performance if the 3D-FFT is performed within 1 or 2 nodes. As inter-node communication increases, the procedure becomes

more advantageous. However the basic procedure, while impressive for large cases, is not as efficient on HPCx as the current internally developed 'LPAR-customisation' technique which collects information within nodes and reduces the number of processors involved in the MPI-AlltoAllV, while keeping message sizes within the 'eager' buffered-memory limits. At the end of the quarter, a more sophisticated band-blocking technique based on 'LPAR-customisation' but which pushes messages to the RDMA limit is being tested. The main challenge from a user point of view will be to develop a guaranteed automated algorithm to choose which combination of LPAR-combination and band-blocking will give optimum performance. This will be in place in the near future.

The above work retains the current CASTEP model of parallelisation over k-points (near embarrassingly parallel) and 3d-FFTs. As has been noted before, many larger physical problems tend to have few k-points. A proposal was submitted to the EPSRC Software Development Call (PI Keith Refson, M Plummer as a co-investigator) for parallelisation over bands for CASTEP (in conjunction with similar optimisation of the CRYSTAL code). This would improve performance and deal with current concerns over large memory requirements for the current CASTEP code.

CRYSTAL: The latest version of this package, CRYSTAL06, has been beta tested and released to users on HPCx. This involved I.J. Bush, in conjunction with a number of experienced users, performing a wide range of tests on the package, and I.J. Bush fixing the bugs that were found. A number of users have already asked for access to this version.

I.J. Bush has also improved the parallel characteristics of CRYSTAL06. Improvements include

- Full and more convenient integration of the work on distributed DFT grids with the user interface;
- Allowing the user to specify in the input options the use of faster (if potentially less robust) diagonalisers;
- Improved log file I/O options to avoid the user being "swamped" with information.

In collaboration with B.G. Searle, I.J. Bush has produced an optimised version of *properties*, the CRYSTAL analysis program, and made that available to users.

A comparison of CRYSTAL06 on HPCx with BG/L, an XT3 and Mare Nostrum has been started.

VASP: K. D'Mellow has worked on some memory issues in VASP 4.6.23, manifesting in unpredictable runtime behaviour when performing certain I/O operations. These are currently affecting VASP's ability to create xml output. This is a recurrent set of issues across previous versions of VASP, which are in the main due to ageing legacy code.

4.3.3 Computational Chemistry Codes

DL_POLY: This quarter DL_POLY work has focused primarily on additional functionality and safety compliant with the requirements for domain distributed parallelism, corresponding memory distribution and software modularisation. These will become available in the next release scheduled in Q2 of 2007 after the testing phase is completed. The most outstanding new features include:

- Conjugate gradient minimiser (CGM) and re-gaussing routines to help users with the simulation and equilibration of systems away from equilibrium or under extreme conditions.
- Fennel's version of damped, force shifted coulombic potential to substitute or be used as complementary technique to the Ewald summation for electrostatic interactions when the system is large enough to allow cutoffs of $\sim 15 \text{ \AA}$ and larger. Fennel's technique scales as $O(N)$ rather than $O(N \log N)$ of the SPM Ewald, where N signifies the system size. Additional improvements of this technique also take care of residual dipole moment outside the cutoff sphere by using a standard reaction field and implement self-adjusting damping factor according to the choice of the cutoff length.
- Infrequent calculation of the k-space SPM Ewald contributions to system properties. This can be used to speed up system equilibration as well as decrease the price of biochemical simulations which need long-timescales to capture conformations and are moderately insensitive to the k-space SPM Ewald contributions which do not change much during equilibrium simulations in such systems.

Various safety features, hidden from the user, improve input processing with an increased degree of checks for self-consistency and error handling. Various pre-staging algorithms are improved to handle initial conditions of type "idealised systems" such as single strings or single planes along Cartesian axes which were prone to failure for DL_POLY_3 previously. Additionally, minor speed-ups (0.5-2%) are achieved by substituting division with multiplication in nested do loops of the most intensive parts of the k-space Ewald evaluation where possible and without sacrificing or loss of the scientific reading of the algorithms.

The current release has proved itself to be of high quality as only three bugs have been reported since December 2006. This is impressive bearing in mind the greatly increased functionality and complexity of the package with respect to its previous version. Routine work on HPCx detected an xlf9-10 memory overwrite bug which was submitted as a PMR to IBM. Following two months of investigation IBM have supplied a fix, which will be included in the next compiler upgrade.

I.J. Bush has also run DL_POLY_3 benchmarks on two IBM Bluegene BG/L platforms: at Jülich in Germany and at Daresbury Laboratory. HPCx is typically 2.5 times faster than BG/L on a per-processor basis. A comparison with the XT3 at Lausanne has also been started. The BG/L data reveals a superb strong-

scaling performance of the package on a 15 million atom pyrochlore system up to 16,384 processors. This opens the way to access energies of 100 keV in radiation damage simulations. However, the simulation of such large systems on such processor counts indicated problems with data dumping in ASCII and strategies how to alleviate this are now being considered.

Scientific work on radiation damage simulations has come to an end in January and a paper in Nature is targeted for June this year.

DL_POLY_2: there has been a minor adaptation of the statistical output routines to give detailed step-by-step information on 'mean force' calculations between fixed centres without overloading the more general statistical properties output file.

GAMESS-UK: A.G. Sunderland has written an interface routine for the GA version of GAMESS-UK that enables use of efficient Scalapack eigensolver routines rather than PEIGS. This has increased the GAMESS-UK eigensolver performance by up to 50% for some of the large datasets, resulting in an overall increase in code performance of around 20%.

CHARMM: P. Sherwood has provided support for Jonathan Essex and colleagues (University of Southampton) in the installation of a modified version of CHARMM onto HPCx. The scalability of CHARMM for the problem type and problem size in question is quite poor, being limited to around 16 processors, especially considering the group was working with an old version of the code. The group asked if a task farm approach could be used, with a large number of 16-processor jobs.

Existing HPCx task-farming tools were limited to

- i) using the provided task-farming binary to run serial codes or codes using OpenMP;
- ii) adapting codes at the MPI level by introducing additional communicators.

Following option ii) did not seem cost-effective in this case because of the customised version of CHARMM in use. Instead it was realised that Option i) could perhaps be extended to task farm parallel codes if a "secondary MPI" communications library could be used at the same time as IBM's PSSP and not interact in any way with it. The MPICH implementation was chosen as it has a port to AIX and also supports a shared memory communications layer which might be expected to be well suited to running up to 16-way parallel on HPCx. In the event, it was found to be reasonably easy to build MPICH and to modify the build procedures for the Southampton version of CHARMM so that all the MPI calls within the application were resolved against MPICH rather than PSSP (i.e. using xlf95_r instead of mpixlf95_r as a loader).

Other small changes required were an extension to the command length of the task-farming binary and establishing the correct MPICH process group specification (using 'localhost' as hostname for all slave processes) so that once

a single task is started on a given node by poe/taskfarm it can start 15 more processes on the same node using the MPICH machinery.

The relevant Southampton PDRA has been out of the country since the work was completed, but the implementation is now available for the group to start testing.

SIESTA: The latest SIESTA 2.0.1 version has been built, tested, tuned and installed on the service machine by J. Hein. The HPCx web page has been modified to detail its availability.

4.3.4 Physics Codes

LUDWIG: K. Stratford has worked on the performance of the LUDWIG code on HPCx, and has attended (on behalf of projects e10 and e19) the HPCx training course on IBM's HPC toolkit (see below).

GWST code: K. D'Mellow and L. Smith visited Professor Rex Godby from e40 on the 25th January 2007. K. D'Mellow has started working on the GW Space-Time code, which implements Hedin's GW method in which the representation of the Greens function and Coulomb interaction are given in the real-space and imaginary time domain. This is a serial code that requires parallelisation for both speed of computation and to allow access to large amounts of memory. Since an early parallelisation of the code, the serial version has been significantly refactored and expanded. Based on the current serial version, a distributed data strategy is being taken in order to optimally utilise available memory. He has analysed and profiled previous parallelisations of this code, and has worked on a port to HPCx of the current serial code, involving system specific modifications to the build process. Currently, he is isolating and fixing platform specific bugs in the serial version, and drafting a parallelisation and MPI integration strategy. It is intended to parallelise the code whilst preserving its ability to run in serial, to avoid future version discrepancies.

CENTORI: J. Hein has worked on the serial performance of CENTORI. CENTORI, developed by UKAEA Fusion in Culham, uses a fluid model to simulate the turbulent flow of the plasma inside a fusion reactor. Early results show a 33% improvement in execution time. A simple performance assessment of GS2, which is a gyrokinetic code used by UKAEA Fusion to investigate flux-tubes inside a fusion reactor's plasma, has been performed. The Culham group was visited to discuss possible future code developments for CENTORI and GS2 in the light of their forthcoming application for computational resources (HPCx or HECToR).

4.3.5 Engineering Codes

Saturne: Collaboration and a visit on 24th January 2007 from the Saturne_code developers from EDF led to this code being installed on HPCx. A.G. Sunderland has been working further with the developers over testing and benchmarking and it is planned that this freely-available CFD code will be made available for general use in the near future.

Imperial College DNS code: The parallelisation of a direct numerical simulation code for two-dimensional inverse-cascading homogeneous turbulence has been carried out on behalf of Erwan Hascoet and Christos Vassilicos at Imperial College. The code employs the pseudo-spectral method of Goto and Vassilicos (New J. Physics 6, p65, 2004), and is being used to investigate the clustering of inertial particles in homogeneous isotropic turbulence under the influence of gravity. Preliminary results show elongation of particle clusters in the direction of gravity for particle Froude numbers much greater than one, but no influence for Froude numbers below 1.

Cranfield CFD code: A fluid dynamics code is currently being parallelised for Evgeniy Shapiro of Cranfield University. This code is designed to solve arbitrary hyperbolic or hyperbolised (e.g. compressible or artificial-compressibility incompressible Navier-Stokes) conservation laws on multiblock structured grids, with multigrid acceleration. The intention is to use this code to perform large eddy simulations on HPCx.

AMAZON: Some consultation has been provided to David Ingram (e54: free surface simulation of waves overtopping during storms) concerning capability incentives for the Amazon code.

4.3.6 Life Sciences

NAMD: We are experiencing problems building NAMD with the latest IBM C++ compiler xlc8. The problems have been investigated and J. Hein is presently liaising with the developers of NAMD to resolve the problem. We have re-installed catdcd (file processing tool used with NAMD) with version 8.4 of TCL. This install was required for allowing large (>2GB) files to be handled with the new version on NAMD which was installed by J. Hein in January.

Integrative Biology: K. Stratford has been supporting Gernot Planck of e17 in trying to debug a problem in one his codes using Petsc. This is ongoing.

Systems Biology: Previously F. Reid and L. Smith worked with a group led by Andrew Millar, a Professor in Systems Biology to develop his code for a submission to BBSRC for time on HPCx. This submission has now been successful and Professor Millar has been awarded time on HPCx. F. Reid has been investigating the performance of Andrew Millar's Clock Model code. The code runs 4 times slower than the previous version (apparently due to be a change in the operation where a computation is now carried out 4 times). The code and relevant documentation has been handed over to K. Stratford who will work on the parallelisation process.

4.3.7 New Applications

OpenBEAGLE: A fix has been provided by IBM for the Internal Compiler Error in the C++ compiler which was affecting this code. Tests using the patched compiler are currently being carried out.

4.3.8 Libraries

Linear Solvers: A.G. Sunderland has installed the MUMPS package on HPCx and has analysed the parallel performance of the sparse matrix solver. I will feedback performance to Iain Duff.

4.3.9 Tools

Vampir: A.G. Sunderland has installed the latest version of Vampir and VampirTrace (v 5.0) on HPCx. Following the commercialisation of the product by Pallas GmbH and their subsequent take-over by Intel, VampirTrace has now returned to being an academic project by the University of Dresden. It is free, unsupported and requires source-code compilation. Both VAMPIR and VanpirTrace have been tested on HPCx and found to work correctly. One outstanding issue is the incorporation of hardware event counters based on PAPI. This is a new feature that should allow 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. At present this is unavailable on HPCx though it is hoped that this feature will be available for general usage once the latest version of PAPI is installed and tested on HPCx.

IBM HPC Toolkit: An HPCx training course took place on 17th-18th January on IBM's latest HPC toolkit with IBM tutor I-Hsin Chung. The HPC toolkit contains a suite of tools to investigate performance aspects of applications, when using IBM hardware such as HPCx. The course was attended both by users and by HPCx staff.

MATLAB: On 27th March F. Reid met with representatives from MATLAB regarding the possibility of obtaining the distributed version of MATLAB for HPCx (MATLAB contacted us to request the meeting). No figures were available for the performance (scalability) or latency of the communications. The licensing costs would likely prevent it from being available on more than a modest number of processors (~64). Licenses are required for both the basic toolboxes and also for the distributed computing engine. Therefore, MATLAB is unlikely to be suitable for systems such as HPCx/HECToR. It is much better suited to smaller department/University clusters (e.g. the Edinburgh ECDF system).

4.3.10 Applications on New Technologies

As part of the investigation of new and alternative technologies J. Hein has been working on optimisations for parallel FFTs when using a toroidal network. The IBM BlueGene and the Cray XT4 (HECToR hardware) employ toroidal networks. In this context J. Hein and F. Reid also attended a workshop on application porting and performance optimisation on the Cray XT4 at CSC Finland from 5th-8th March. During the workshop LAMMPS 2007 was ported to the CSC Cray XT4. Several benchmarks (both I/O intensive and compute intensive) were run in order to investigate performance of the system. The tools installed on the system

were also investigated, e.g. profiling tools for measuring time spent in user code (non-MPI) and time spent in communications (MPI).

4.4 Software Engineering (*Dr Stephen Booth*)

4.4.1 Migration to HECToR

We have been investigating network performance over the New SuperJANET5 network with particular emphasis on transfers from HPCx to the future location of HECToR. As a result of this work we have produced a technical report:

- **HPCxTR0702:** *Transferring Large Files to Remote Sites: An Update*, C. Maynard.

Unsurprisingly the SuperJANET5 network seems to be problem free with the performance bottlenecks usually located in the more local network. It is clear that very careful tuning of the networking parameters at both ends of a connection are required to get good throughput from a single network stream but that the use of multiple parallel network streams (as used by the grid-ftp protocol) is an easy way to work round any problems of this kind.

4.4.2 In-depth software support

Grid middle-ware support

A new version of Globus, Globus-4.0.4, is now available. A 64-bit version of this release is available on HPCx and a full 64/32 install will become the default version of Globus in the near future.

The Globus job-manager script has been extended to generate a default charging budget when Globus jobs are submitted without an explicit budget being specified. This allows greater compatibility with other Grid software such as the new INCA-2 automated tests being deployed by the NGS. HPCx is now passing the INCA-2 tests and is integrated with the automated testing environment.

We have been working on the Dante port-forwarding software in order to configure it to support network access to the back-end nodes from the dedicated UKLight/TeraGrid network. We are also continuing to support our grid-proxy port forwarder as this is more convenient for some applications.

SAFE

We have been continuing the development of SAFE, the user administration software suite for the HPCx Service. As IBM now support version 5 of the Java language, the entire code base has been migrated to Java-5. Java-5 provides a number of language syntax extensions over previous versions of the Java language. These include support for "Generic classes". We are now using these extensively in the user administration software in order to implement stronger type checking in the code.

We have also been developing new helpdesk software which will be tightly integrated with the rest of SAFE. There are a number of significant advantages to this, including:

- The web interface that allows users to submit new queries and to view the status of their queries will be greatly improved.
- It will be significantly easier to identify which project a particular query has originated from.
- Generating helpdesk statistics and reports will become easier.
- Quota and CPU Budget information will be directly available in the helpdesk making it easier to identify resource allocation problems.
- The helpdesk will have direct access to the user database making it easier to find alternate contact details such as phone numbers.
- Users will only have to register with the system once to apply for accounts and to gain access to the helpdesk.

This change will also make it easier to make future extensions to the system such as gathering additional user feedback when a query is closed. An initial version of the new helpdesk is now complete and we plan to deploy this for the HPCx service on 1 May.

4.5 Operations and Systems (*Mr Mike Brown*)

There was 1 SEV1 incident (ie, contractual failure) this quarter, the lowest of any quarter in the service; it resulted from a network failure at Manchester. The total number of incidents at all levels was 13 – this was also a record.

There were 4 maintenance sessions this quarter. Details of the work carried out can be found on this webpage:

<http://services.hpcx.ac.uk/maintenance/>

Click on a session date to view the actions performed.

Mike Brown and a number of the systems team will be attending the IBM HPC User Group meeting which has been arranged for April.

4.6 Staffing

<i>AV</i>	<i>January</i>	<i>February</i>	<i>March</i>
DL	4.5	4.5	4.8
EPCC	8.5	6.7	9.3
Total	13.0	11.2	14.1

<i>Systems</i>	5.8	5.8	5.8
----------------	-----	-----	-----

5 Summary of Performance Metrics

<i>Metric</i>	<i>TSL</i>	<i>FSL</i>	<i>January</i>	<i>February</i>	<i>March</i>
Technology serviceability	80%	99.2%	100.0%	100.0%	100.0%
Technology MTBF (hours)	200	300	∞	∞	∞
Number of AV FTEs	7.5	10	13.0	11.2	14.1
Number of training days per month	22.5/12	30/12	4/1	4/2	7/3
Non in-depth queries resolved within 3 days	85%	97%	97.9%	100.0%	100.0%
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%

<i>Colour</i>	<i>Meaning</i>
	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

<i>Code</i>	<i>Class</i>	<i>Title</i>	<i>PI</i>
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
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 Decanometer 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
e57	1	Triplex DNA Structures	Dr Hooshang Nikjoo

e58	1	Quantum Simulations for Chemical Biology	Dr Carmen Domene
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

<i>Code</i>	<i>Class</i>	<i>Title</i>	<i>PI</i>
p01	1	Atomic Physics and Astrophysics	Prof Alan Hibbert

NERC Projects

<i>Code</i>	<i>Class</i>	<i>Title</i>	<i>PI</i>
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

<i>Code</i>	<i>Class</i>	<i>Title</i>	<i>PI</i>
b02	1	Modelling enzyme catalysis	Dr Adrian J Mulholland
b08	1	IntBioSim	Prof M S Sansom

CCLRC Projects

<i>Code</i>	<i>Class</i>	<i>Title</i>	<i>PI</i>
c01	1	Daresbury Laboratory Facilities Agreement Consortium	Dr Richard J Blake

Externally-funded Projects

<i>Code</i>	<i>Title</i>	<i>PI</i>
x01	HPC-Europa	Dr Judy Hardy
x03	IBM	Mr Derrick J Byford

HPCx Projects

<i>Code</i>	<i>Title</i>	<i>PI</i>
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

<i>Code</i>	<i>Class</i>	<i>Title</i>	<i>PI</i>
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