HPCx Quarterly Report
January – March 2005

1 Introduction

This report covers the period from 1 January 2005 at 0800 to 1 April 2005 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

- STAC (Scientific and Technical Advisory Committee) were very positive about the progress HPCx made during 2004 and supportive of the objectives in the Annual Plan for 2005, although they asked for more emphasis to be made on Outreach. However, despite a subsequent discussion at the Oversight Committee, the Annual Plan has not yet been signed off by EPSRC.

- This quarter has seen a very successful start to the year with utilisation increasing significantly since December and January. March had the highest utilisation since April of last year and represented the largest number of AUs delivered in any single month of the HPCx service. Capability usage has also returned to more normal levels.

- The system continued to demonstrate its excellent reliability throughout this quarter; there were again only three failures, all of which were due to external network problems.

- The helpdesk again met all the targets for queries during this quarter; indeed every non-indepth query was answered within the 72-hour target.

- EPCC will be hosting the joint ScicomP/SP-XXL meeting from 30 May to 3 June. SP-XXL is the worldwide user group for IBM HPC systems administrators, whereas ScicomP is the major IBM conference for scientific

applications and should therefore be interesting for all users of the HPCx service.

- Two of the Life Sciences molecular dynamics codes (LAMMPS and NAMD) have been benchmarked on HPCx and have now been awarded capability incentives.

- Arthur Trew gave a talk at the Edinburgh International Science Festival entitled *From Blizzards to the Big Bang: what supercomputers can bring to life.* This talk was well attended and was followed by a successful and interesting Question-and-Answer session.

- The Terascaling team have started on a programme of consortia visits to increase their understanding of the scientific drivers and requirements. Four such visits have already taken place.

- We have compiled a set of web pages illustrating the scientific highlights produced using HPCx. This will be combined with similar material from CSAR to demonstrate the scientific benefits of HPC in the UK.

- The Software Engineering team have developed software that can track code usage over time, allowing our optimisation efforts to focus effectively on the most heavily used codes. This work has also enhanced the plotting capabilities from the SAF allowing the production of new graphs showing, for example, the trends in usage by the various Research Councils.

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:

	January	*February*	*March*
Incidents	1	1	1
Failures	18	10	18

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

January

Failure	*Site*	*IBM*	*External*	*Reason*
05.017	0%	0%	100%	External network loss

February

Failure	*Site*	*IBM*	*External*	*Reason*
05.025	0%	0%	100%	External network loss

March

Failure	*Site*	*IBM*	*External*	*Reason*
05.025	0%	0%	100%	External network loss

3.1.2 Performance Statistics

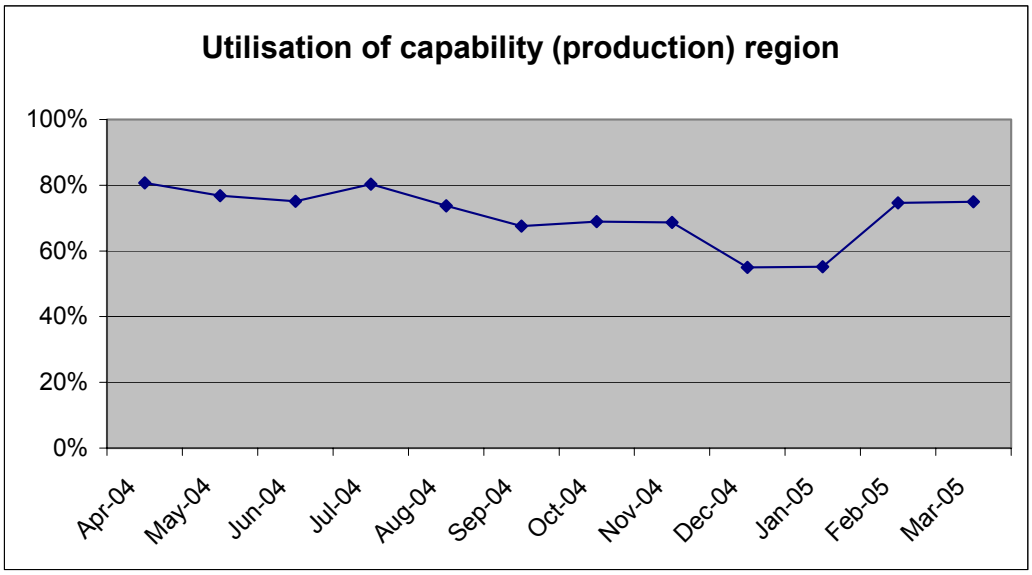
This section uses the definitions agreed in Schedule 7, ie,

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

Attribution	*Metric*	*January*	*February*	*March*	*January*
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	1	1	1
	MTBF	732	732	732	732
	Serviceability	99.9%	99.9%	99.9%	99.9%
Total	Failures	1	1	1	1
	MTBF	732	732	732	732
	Serviceability	99.9%	99.9%	99.9%	99.9%

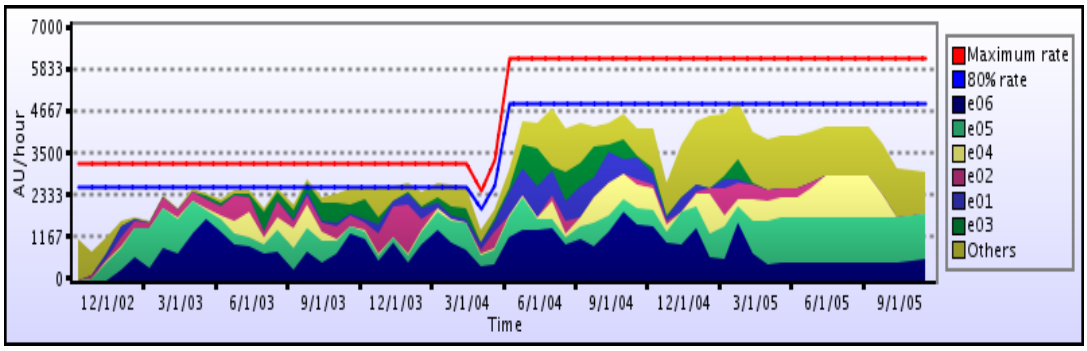
3.2 Capability Utilisation

After a substantial slackening of demand over December and January, utilisation has now recovered strongly. Overall utilisation of the system in March was around 78%, the highest since April last year.



3.3 Capacity Planning

Predicted Utilisation



The graph above shows the utilisation since the start of the project and the projected utilisation until the end of 2005. The scale on the *y*-axis is AUs per hour, where the peak that HPCx Phase 1 could currently deliver is around 3240 AUs per hour, and Phase 2 6188 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 from the SAF, which is often simply that on the original application form.

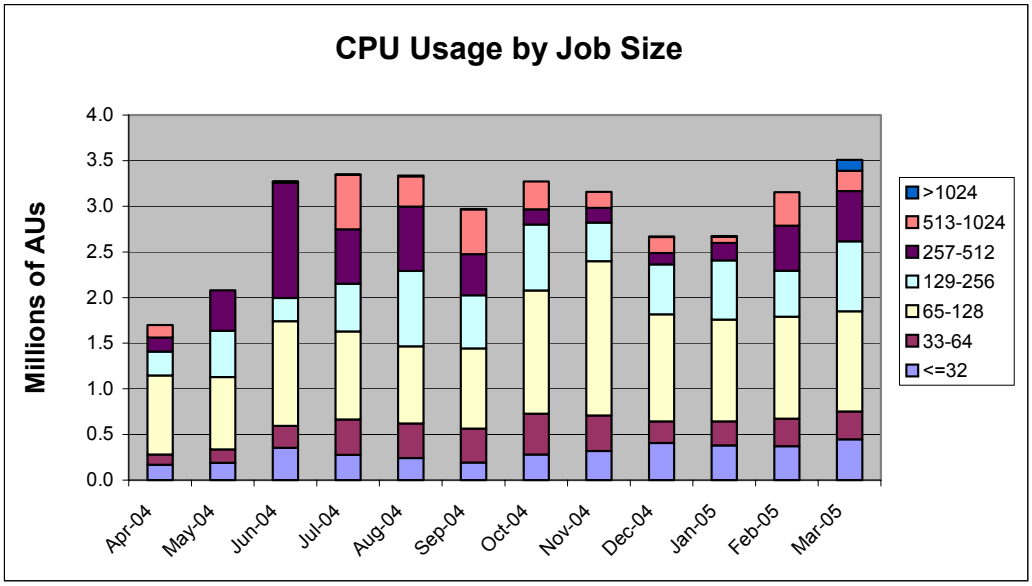
As can be seen from the graph, the projected utilisation of the existing groups falls within the 80% limit. Two other allocations that have already been made, starting in June and October, are not included in the graph. However, this still leaves resources spare for further allocations arising from the current Call for Applications.

Numbers of Research Consortia

There are currently 35 research consortia using the HPCx system. Three other projects have now been closed.

In addition, there are two externally funded projects.

3.4 CPU Usage by Job Size



Deleted: . It nonetheless stood at a respectable 28.7% for the quarter, and reached 39% in June – a record apart from November 2003, when the Tergyroid exercise took place.

3.5 AU Usage by Consortium

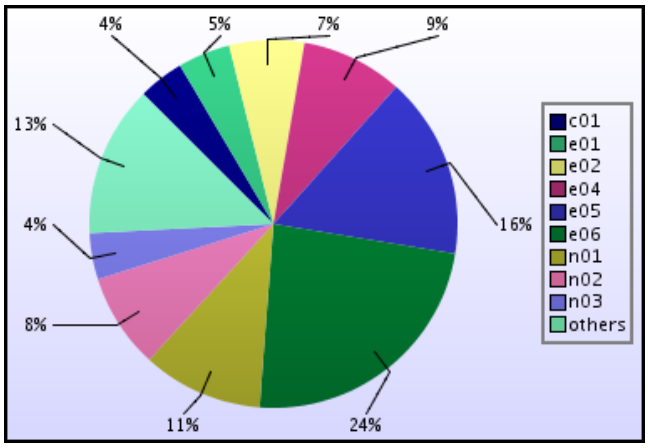
The PIs and titles for the various consortia are listed in Appendix B.

Consortium	*January*	*February*	*March*	*Quarterly*	*%age*
e01	268926	6366	154084	429376	4.6%
e02	41188	78485	494359	614032	6.6%
e03	6226	157	195186	201569	2.2%
e04	31060	596089	203711	830860	8.9%
e05	566461	412275	511595	1490331	16.0%
e06	756873	713075	734040	2203988	23.6%
e07	1	479	1474	1954	0.0%
e08	2348	8102	19549	29999	0.3%
e10		17086	12571	29657	0.3%
e11	12477	27411	8955	48843	0.5%
e14		139	629	768	0.0%
e15	31	2500	3	2534	0.0%
e16			2	2	0.0%
e18	1	21021	6348	27370	0.3%
e19			80240	80240	0.9%
e20	178675	16851	40353	235879	2.5%
e24	10	3956	1044	5010	0.1%
e25		5	0	5	0.0%
e26		743		743	0.0%
z09		8454	8898	17352	0.2%
EPSRC Total	1864278	1913192	2473041	6250511	67.0%
n01	182450	463785	342492	988727	10.6%
n02	333139	298978	148135	780252	8.4%
n03	77419	173827	127878	379124	4.1%
n04	57726	90677	110254	258657	2.8%
NERC Total	650734	1027267	728759	2406760	25.8%
p01	697	5558	3321	9576	0.1%
PPARC Total	697	5558	3321	9576	0.1%
c01	85237	92992	197839	376068	4.0%
CCLRC Total	85237	92992	197839	376068	4.0%

b02	1795	4285		6080	0.1%
b03	6			6	0.0%
b05	32270	64759	11691	108720	1.2%
BBSRC Total	34071	69043	12073	115187	1.2%

x01	211	6108	1502	7821	0.1%
x02			71869	71869	0.8%
External Total	211	6108	73371	79690	0.9%

z001	29079	23751	28460	81290	0.9%
z002		0	0	0	0.0%
z05	1108	8972		10080	0.1%
z06	0	4269	100	4369	0.0%
HPCx Total	30187	36992	28560	95739	1.0%



3.5.1 Discounts

There are now a number of user codes that have qualified for capability discounts. The following table shows the discounts that were awarded during the last quarter.

Consortium	*AUs Used*	*AUs Charged*	*Discount*
b05	131779	108719	23060
e01	435397	429376	6020
x01	7826	7820	5
z001	81313	81290	22

3.6 Helpdesk

3.6.1 Classifications

Category	*Number*	*% of all*
Administrative	112	47.5
Technical	99	41.9
In-depth	24	10.2
PMR	1	0.4
TOTAL	236	100.0

Service Area	*Number*	*% of all*
Phase 2 platform	216	91.5
Website	11	4.7
Other/general	9	3.8
TOTAL	236	100.0

3.6.2 Performance

All non-indepth queries	*Number*	*%*	*Target*
Finished within 24 Hours	167	79.1	75%
Finished within 72 Hours	211	100.0	97%
Finished after 72 Hours	0	0.0	

Administrative queries	*Number*	*%*	*Target*
Finished within 48 Hours	111	99.1	97%
Finished after 48 Hours	1	0.9	

3.6.3 Experts Handling Queries

Expert	*Admin*	*Technical*	*In-Depth*	*PMR*
epcc.ed.ac.uk	97	57	12	0
dl.ac.uk	0	17	9	1
Sysadm	15	25	3	0
Other people	0	0	0	0

3.7 Service Quality Tokens

Date	*Person*	*Value*	*Comment*	*Status*
Mar 20, 2005 1:46:58 PM	*Dr Kostya Trachenko*	* * * * *		

4 Support

4.1 Applications Support *(Dr David Henty)*

4.1.1 Documentation

In response to a number of issues raised via the helpdesk, the guide for the web-based administration system has been completely restructured and rewritten. It can be found at *http://www.hpcx.ac.uk/projects/FAQ*.

Due to the stability of the HPCx hardware there has been no need to change the system documentation. There have, as ever, been ongoing updates to the online FAQ, for example explaining in more detail how best to run the Totalview debugger on the Phase 2 system.

4.1.2 Technical Reports

A total of four reports were due for Q1 in the following areas:

a) Overview of Techniques for Achieving Scalability on HPCx
b) Optimisation of an Environmental Modelling Code: a Case Study
c) Visualisation Tools and Libraries on HPCx
d) Comparison of Parallel Debuggers on HPCx

We have produced the following three reports this quarter:

- **HPCxTR0501**: *Terascaling Techniques on HPCx*, Stephen Booth and Adrian Jackson.
- **HPCxTR0502**: *Performance Optimisation of an Environmental Modelling Code (POLCOMS)*, Michael Holden.
- **HPCxTR0503**: *Using VTK and the OpenGL Graphics Libraries on HPCx*, Jeremy Nowell.

Reports **01**, **02** and **03** correspond directly to the titles **a)**, **b)** and **c)**. We decided to move the production of report **d)** to Q2 as it requires extensive investigation of the DDT debugger. There were some technical problems at the beginning of the year regarding launching DDT from the HPCx batch system. These delayed the start of this work, but have now been resolved.

There are a total of twelve reports due for 2005. The profile given in the Annual Plan was deliberately front-loaded to cope with any unforeseen delays, so with three reports in the first quarter we are still on target for the year.

4.1.3 Training

In Q1 of 2005 we ran the following two courses:

- **Edinburgh, 22– 24 February**: *Message-Passing Programming with MPI.*
- **Edinburgh, 30 March–1 April**: *Shared-Memory Programming with OpenMP.*

Both courses were run over three days to allow more time to discuss issues specific to the attendees; a session dedicated to consultancy with HPCx staff was also included after the end of the taught parts.

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

Metric	*Total*	*Target*
Course days	6	30
Number of courses	2	12
Different locations	1	4
Student-days for HPCx users	108	
Student-days for HPCx staff	0	
Student-days available for HPCx	156	600

In the coming year we plan to run courses at a variety of locations around the UK. Although many of these are in the process of being arranged, the list includes Daresbury, Rutherford, Leeds and London.

4.1.4 Workshops and Conferences

Other than the Third HPCx Annual Seminar, to be held at Daresbury in December alongside the Machine Evaluation Workshop, the major event this year is the joint ScicomP/SP-XXL meeting to be held in Edinburgh from 30 May to 3 June. These are the major conferences for scientific users / systems administrators of IBM supercomputers, and receive significant support from IBM in terms of financial sponsorship and provision of key speakers. As well as contributed talks, there will be updates from IBM on the hardware roadmap, operating system releases, new compiler features and applications performance on the latest machines, including POWER5 and Blue Gene. We also plan to run a tutorial session covering performance optimisation on POWER5 systems.

4.1.5 Newsletter

The fifth issue of *Capability Computing* is in production and will be available in time for distribution at ScicomP11 in late May, in addition to the usual mailshot.

4.1.6 Packages

No major new packages have been installed this quarter. As well as major parallel applications, we also maintain a number of smaller utility programs, written by the HPCx team, within the packages setup. One of these, the `taskfarm` program, allows serial programs to be run in a massively parallel manner using a simple command-line interface. This has proved extremely useful to members of the Chemreact consortium, who are using it for more efficient pre and post-processing of data (an issue that was raised during an HPCx consortium visit).

The helpdesk also raised another issue regarding an external package. In answering a query regarding problems using the MPH package, Joachim Hein uncovered a bug in the library. This bug was fixed, and the solution passed on to the authors of the package who are including it in their next release.

Deleted: delegates.

4.2 Outreach Activities *(Dr Richard Blake)*

4.2.1 Outreach to Lifesciences

Molecular Dynamics Codes

We have focused on two molecular dynamics codes this quarter: NAMD and LAMMPS. Both have received capability incentives (Gold and Bronze respectively). The work has involved benchmarking, performance profiling and bug fixing and is reported in more detail in the TeraScaling report (section 4.3 below).

QDVE Project

The GAMESS-UK task farm has been developed as part of our support for Marcus Durrant's Quantum Directed Virtual Evolution project (b05). The project uses a "Genetic Algorithm" to try and determine an optimal catalyst for the conversion of nitrogen to hydrazine. Several putative catalysts are chosen to create “generations" of catalysts, and, for each catalyst, the energy of various species within the catalytic cycles is calculated using GAMESS-UK and compared with an ideal value. Promising candidates from this generation are selected and a new generation created by shuffling various characteristics of the candidates. This process is iterated until a catalyst of the desired efficacy has been found.

The GAMESS-UK task farm is spawned over a large number of processors on a parallel machine, and handles the running of the numerous small jobs, on a subset of these processors. Considerable effort has been put into making the task farm simple to use for a non-expert and as robust as possible with respect to errors encountered during the individual GAMESS-UK runs. Additional help has involved writing a number of Python scripts to summarise the results from the large number of jobs and to prepare fresh inputs to restart from the last viable point any calculations that may fail.

The parallel efficiency of the task farm has qualified it for an HPCx "Gold Star" capability incentive, which has helped to extend the number of generations that the project will be able to run for. The project is currently nearing completion of the third generation, with the catalysts already settling down to a significantly smaller number of catalytic species of similar charge and geometry.

Based on the results obtained so far, BBSRC have granted the project an extension of 3 months to allow more generations to be processed.

Integrative Biology Project

We have begun to port the Cardiac Arrhythmia Research Package (CARP) to HPCx, for the Integrative Biology project. The code has been developed to carry out large-scale cardiac simulations.

Modelling of the Human Retina

Work has begun on developing a domain decomposition implementation of a C-version of the retina modelling code. The time-dependent model resolves the retina into photoreceptor, horizontal, bipolar, amacrine and ganglion cells with differing densities of cells in each layer, and complex interlayer and intralayer connectivities. A two-dimensional domain decomposition involving columns of different cells is being developed using METIS. Current work is focussed on testing the data swaps between neighbouring domains.

4.2.2 Other Outreach Activities

Operational Research

David Henty has been working with academics from the Operational Research group at Edinburgh University, to develop a proposal for HPCx time. They have been approved access to Outreach time on HPCx to do this. The work will involve a world-beating calculation on HPCx. The actual calculation will be synthetic; however, the application area comes from a similar area to financial risk management and is therefore linked to real world and commercial calculations.

Virtual Screening

Judy Hardy is in early discussions with a group from the Institute of Structural and Molecular Biology at Edinburgh University over the possibility of developing an HPCx proposal to develop a “virtual screening resource for chemical genomics and ligand discovery”.

4.2.3 Improve Public Awareness

Edinburgh International Science Festival

Arthur Trew and Irina Nazarova have put considerable effort into developing a talk for the Edinburgh International Science Festival. The talk, titled: *From Blizzards to the Big Bang: what supercomputers can bring to life* will be presented by Arthur Trew on Sunday, April 10 in the Royal Museum & Museum of Scotland. The HPCx demonstrator has been updated and will be available for attendees to view prior to the talk.

Further details can be found on the Science Festival web site:

http://www.sciencefestival.co.uk/

4.3 Terascaling Applications *(Dr Martyn Guest)*

The main focus of the Terascaling team activities for this quarter has been on facilitating capability usage on the system. This has involved arranging visits to Consortia, to discuss their requirements and scientific output, monitoring the profile of codes on the system, performance evaluation and optimisation of specific codes and promoting capability incentives on the system. The work, future direction and priorities of the team were presented at the HPCx STAC Meeting (1st March 2005).

4.3.1 Capability Science

Consortium Visits

We have begun the process of visiting PIs and leading users to discuss their requirements and scientific drivers. The following consortium visits have taken place:

- Lorna Smith and Mark Bull attended the 5th HiGEM Consortium Meeting at Reading on the 7 February 2005.
- Martyn Guest and Andrew Sunderland attended a ChemReact Consortium Meeting at UCL on the 23 March 2005.
- Joachim Hein, Lorna Smith and Martyn Guest visited UKAEA, Culham, on the 4 April 2005.
- Martin Plummer and Ian Bush attended an e05 Consortium Meeting at the Royal Institution on the 5 April 2005.

Capability Incentives

The Terascaling team have worked with user groups to obtain capability incentives for the following codes in this quarter:

- NAMD (Gold)
- LAMMPS (Bronze)

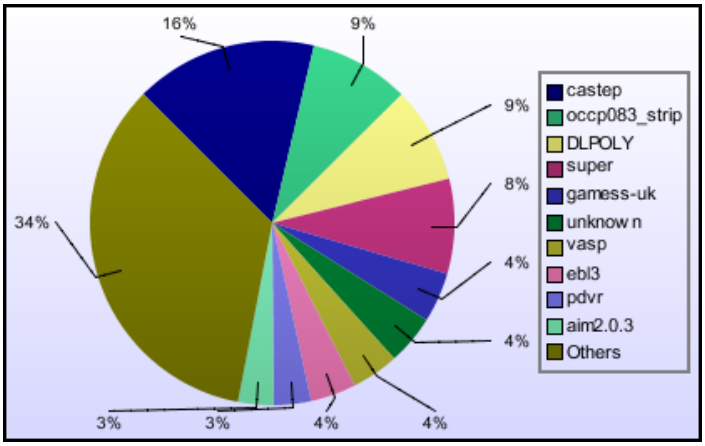
A full list of codes with capability incentives can be found at:

http://www.hpcx.ac.uk/services/policies/starcodes.html

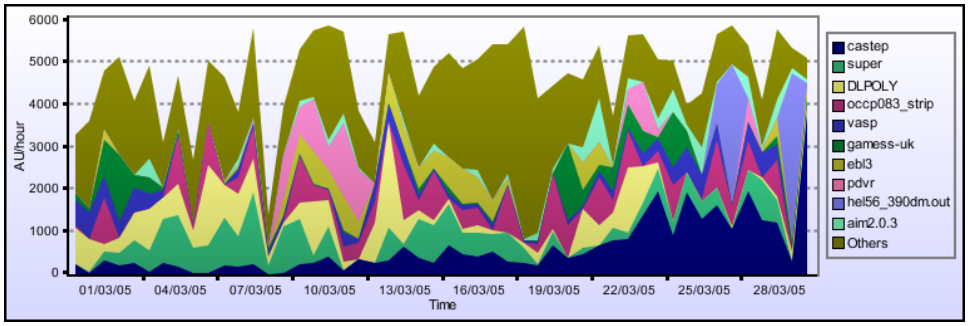
Code Usage

The Software Engineering team have now developed software to allow code usage to be monitored as a function of time. This allows the TeraScaling team to

focus efforts on the codes making most substantial usage of the system. The two figures below shows the percentage code utilisation for March.



Percentage Code Utilisation for March 2005.



Code AU Utilisation for March 2005.

It is clear from these figures that CASTEP used the largest amount of time during March, followed by the OCCAM code. DL_POLY and the `super` code also utilised substantial amounts of time.

Scientific Highlights

Scientific highlights from work done on HPCx have been compiled as a set of Web pages. This comprises published and near-publication results from Materials and Condensed Matter (10 highlights), Atomic and Molecular Physics (2), Life Sciences (1), Engineering (2) and Environmental Science (2). This will be combined with similar material from CSAR to present a UK-wide summary of the scientific benefits of HPC, and publicised on both the HPCx and CSAR websites.

4.3.2 Quantum Chemistry and Ro-vibrational Codes

GAMESS-UK

- Work has been completed on the distributed data MPI / ScaLAPACK analysis module. This now uses very much less memory than original version. The properties and graphical analysis modules were also incorporated into the MPI code, with the memory associated with the latter drastically reduced. Other developments have included:
  - Implemented Fermi-Dirac smearing on the distributed data memory SCF driver. This is a technique that can be useful in accelerating convergence in certain cases by effectively running at non-zero Kelvin and then annealing the system to zero Kelvin. A number of annealing algorithms were implemented
  - Developed a set of wrappers so that the distributed SCF driver can be run in serial, which can be very useful for debugging. A set of routines for printing and writing to file distributed matrices was also developed, again useful for debugging.
  - Fixed a problem with restarts in the distributed data SCF driver.
  - Debugging: especially one particularly problematic bug that only showed up for large systems.

CRYSTAL

- We have developed an initial version of k-point parallelism in the distributed data code. This now allows each k-point to be distributed across a subset of the processors. At present only the diagonalisation is performed in this manner, and this form of data distribution will be expanded to all the other linear algebra. However as the diagonalisation is the dominant linear algebra operation the most important part has been done. However some load balancing issues need to be finalised (some k-points require a diagonalisation of real symmetric matrices, other complex).
- We are also developing an interface between GAMESS-UK and CRYSTAL so that the latter can be restarted from the former. This is useful because:
  - In certain cases it can be easier to converge the isolated molecule and use that as a starting point for the periodic calculation;
  - Related to the above, GAMESS-UK has superior convergence acceleration algorithms, which can make difficult systems converge more quickly, or even converge at all.

MOLPRO

- We have installed and benchmarked the parallel code on HPCx in consultation with the developers (Wilson and Knowles) and are currently analysing the I/O performance. Progress and next steps on identifying the causes of poor performance on HPCx were discussed with Peter Knowles at the ChemReact Consortium meeting on 23rd March.

PDVR3DRZ (Tampa Group - Tennyson and Munro)

- We are examining the parallel performance of this code on HPCx and have demonstrated to the developers the use of multiple BLACS grids which allow the use of multiple ScaLAPACK Diagonalisations. We have also advised switching to the Divide-and-Conquer algorithm in this case. Feedback from the users indicates a several-fold improvement in parallel performance.

4.3.3 Computational Materials Codes

CASTEP

- Castep 3.1 is now publicly available. This combines the LPAR-aware communications techniques developed earlier with the York intelligent task farming. By request of the Castep Developer's Group, the latter capability is not yet being publicised to users who are not already aware of it. Over the year more applications (and by negotiation, more users) should make use of this facility. We are currently testing and optimising an application module for the production of Wannier functions for Prof P Madden and co-workers. General optimisations include a successful re-write of the restart routines to avoid overloading HPCx's memory with read-in data on large jobs (and to speed up the restart), plus ongoing serial-fraction reduction and bug fixing.

4.3.4 Life Science and Molecular Simulation Codes

LAMMPS

- This code forms part of the ETF demonstrator for SC2005, and is therefore an important code to investigate and optimise. The latest version of LAMMPS (LAMMPS 2005) has been ported to HPCx and the performance of this code benchmarked. These benchmark results have been passed to the developer and now appear on the LAMMPS web site. Fiona Reid ported the code, and in the process identified a bug in the MPI implementation. She has fixed the bug and passed the fix to the developers.

  Her efforts have resulted in LAMMPS being awarded a bronze capability incentive. She has also worked with a user from UCL to benchmark a range of real systems, comparing and contrasting the performance with that of the SGI Altix at CSAR. She has looked closely at the IO usage of this code, and worked with a user to ensure an optimal usage of the restart facilities. Fiona is currently focussing on the communication overhead, with an aim of reducing this. A technical report is currently under construction.

- LAMMPS web site: *http://www.cs.sandia.gov/~sjplimp/lammps/bench.html*

NAMD

- This code is also important for the ETF demonstrator for SC2005. Joachim Hein has benchmarked this code on Phase 2 and passed the results to the developers. These results now appear on the NAMD web site. This activity has led to NAMD being awarded a gold capability incentive.

- NAMD web site: *http://www.ks.uiuc.edu/Research/namd/performance.html*

DL_POLY

- We have worked in collaboration with DL_POLY developers in the finalisation of version 3.0.4, a major revision that includes many more Fortran 90 features than before. Initial results from the testing and benchmarking of this version are encouraging. Systems of over 50 million atoms can now be run, and the code is being used to examine radiation damage by the material chemistry consortium. Extensive runs are required because the damage caused by radioactive decay can cover a very large area.

4.3.5 Engineering Codes

ROTORMGMGP (UKAAC)

- We have optimised and benchmarked the ROTORMGMGP code (Chris Allen) on HPCx. The code has been awarded a Gold Star as part of the Capabilities Incentives Scheme. Timings have also been obtained from the Altix system and the results will be presented at Parallel CFD 2005.

Hydra (UKAAC)

- Optimisation work underway on the Hydra code involves an analysis of the code's memory usage.

CFX and Fluent

- The CFX licence is now permanent on HPCx. The Fluent license has been renewed but is still temporary.

4.3.6 Environmental Codes

POLCOMS

- We have investigated the performance of this code on the Phase 2 system. As with Phase 1, this code scales well to thousands of processors. However, the peak processor speed is still low and hence we focused our effort on improving the sequential performance of this code, and in particular, on the use of unnecessary divides and the small inner loop lengths, which prevented the compiler from utilising the vector library effectively. Michael Holden has been able to achieve a performance improvement of 16% on 128 processors.

- M. Holden, *Performance Optimisation of an Environmental Modelling Code (POLCOMS),* HPCx Technical Report HPCxTR0502, See:*http://www.hpcx.ac.uk/research/hpc/technical_reports/HPCxTR0502.pdf*

4.3.7 Tools

Totalview and Paraver

- Both tools have been upgraded. Totalview has been set up to allow users to use the memory debugging facilities of version 6.5. The descriptions of both these tools on the HPCx web site have been updated and expanded.

4.3.8 New Application Development

CENTORI

- Joachim Hein has continued to work closely with this consortium, to optimise their code for HPCx. He has written a report for the consortium on the serial optimisation of the code (*Improving the serial performance of CENTORI*, J. Hein) and has contributed to their internal milestone report (which is copied to the EPSRC). Current work is focused on re-parallelisation of the code, to use a multi-dimensional strategy.

- Joachim Hein, Lorna Smith and Martyn Guest visited Culham to discuss this parallelisation strategy and their scientific drivers.

- J. Hein, *Improving the serial performance of CENTORI*, report to the consortium, January 2005.
  P. J. Knight, A. Thyagaraja, J. Hein, *CENTORI code investigations of transport barrier physics in MAST and JET*, Milestone Report, 2005.

4.4 Software Engineering *(Dr Stephen Booth)*

4.4.1 General Terascaling and Optimisation Techniques

We have completed a technical report on General Terascaling techniques for use on HPCx. This is an introduction to the terascaling process and to the particular issues that are important on HPCx. The report provides references to other HPCx technical reports where the individual techniques have been explored in more detail.

4.4.2 Advanced Data Handling and Grid Computing

During this quarter we have made the following enhancements to the Grid computing infrastructure on HPCx.

We updated the Globus LoadLeveler-jobmanager script. This is the script that Globus uses to translate a job request received by the Globus installation into a batch job request to be submitted to the LoadLeveler batch system, which runs on HPCx. Since the original installation of Globus the LoadLeveler configuration on HPCx has been extended to support new keywords. We therefore extended the jobmanager script to support these new keywords as well, and to provide sensible default values for parameters that are mandatory on HPCx but optional in Globus.

The DEISA project has been investigating installing the UNICORE system on HPCx. Like Globus, UNICORE is a Grid middleware environment; Globus is currently the main environment in use in the US and the UK but UNICORE is used widely in continental Europe, especially in Germany, where it was originally developed. Like Globus, UNICORE uses a series of scripts to map generic Grid job requests to the specific environment installed on the target host. We now have test versions of UNICORE running successfully with the HPCx environment and can move on to starting to set up a production version in the next couple of months.

We have been writing a report on IO performance on the Phase-2 system. A proper analysis of IO performance requires a large number of tests to be run, so that the intrinsic IO performance of the system can be separated from the transient effects of other users. This data has now been collected and the final report is undergoing internal review.

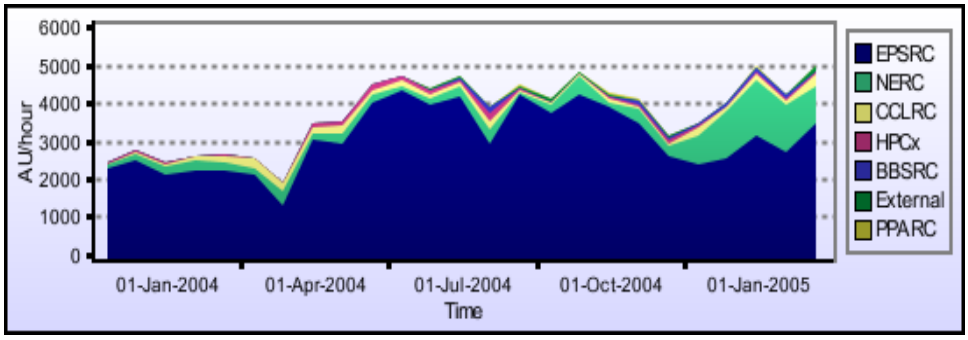
4.4.3 Future Trends

We are intending to investigate the performance and portability issues of important HPCx application codes on alternative computer platforms that may be significant in the future. Initial investigations will use the BlueGene system at Edinburgh as BG systems currently occupy the top positions of the Top500. We have requested access to the OCCAM code for the first of these investigations and hope to start this work at the beginning of May.

4.4.4 Generic Software Support

One of the major activities for this Quarter was to increase the amount of data we gather about the pattern of use of the system. In particular it became clear that the work of the terascaling team would be much easier if they were able to break down the accounting data according to which application codes were being run. This data is not available in the loadleveler accounting records as loadleveler only deals with batch scripts. The actual application codes are started from within these batch scripts (typically using the `poe` job starter program). In order to gather application use statistics we added additional instrumentation to the `poe` program. Once a week the logfiles of `poe` usage are merged with the main accounting database to allow us to identify the applications being run on the system. See section 4.3.1 for an example of the use of this facility.

At the same time we took the opportunity to improve the graph generation code in the SAF to produce additional types of plot. For example the following graph of resource use by project funding body as a function of time, clearly showing the increase in NERC use of the system over the last quarter.



We have also started an activity to develop new tools to aid in the task of optimising code for HPCx. A key part of any optimisation process is to analyse the behaviour of the application in order to understand what parts of the program are expensive and need to be optimised. IBM provides the DPCL library for

writing custom profiling tools. This is a set of C++ classes that allow you to attach to a running program and dynamically alter the running program to add additional instrumentation. The DPCL library is very powerful but is quite complex to learn and use. We therefore plan to use it to build some simple additional profiling tools to measure things that are either impossible or difficult to measure using the standard profiling tools currently available. We are currently developing a profiling tool similar to the standard `prof` tool that additionally uses the hardware counters in the Power-4 chip to provide cache-miss statistics for each routine in the program.

4.5 Operations and Systems *(Mr Mike Brown)*

4.5.1 Staffing

There has been no overall change in staffing levels, although the coverage is still substantially in excess of the "core hours" contractual requirement. Andrew Elwell left the team at the end of 2004 and has been replaced by Ian Georgeson.

4.5.2 Incidents

This quarter there have been no failures (SEV1 incidents) attributable either to the site or to IBM. The three failures recorded were the result of networking problems external to the service.

The pattern of incidents at the lower severity levels reveals no particular areas of weakness or causes for concern.

4.5.3 Outreach

The Operations and Systems Group was represented at the winter meeting of the IBM SP-XXL group of system administrators, and at the UK IBM HPC user group.

4.6 Staffing

AV	*January*	*February*	*March*
DL	5.1	6.1	6.2
EPCC	5.5	7.0	8.5
Total	10.6	13.0	14.7

Systems	5.8	6.1	6.1

5 Summary of Performance Metrics

Metric	*TSL*	*FSL*	*January*	*February*	*March*
Technology serviceability	80%	99.2%	100.0%	100.0%	100.0%
Technology MTBF (hours)	200	300	∞	∞	∞
Number of AV FTEs	7.5	10	10.6	13.0	14.7
Number of training days per month	22.5/12	30/12	0/1	3/2	6/3
Non in-depth queries resolved within 3 days	85%	97%	100.0%	100.0%	100.0%
Number of A&M FTEs	3.75	5.75	5.8	6.1	6.1
A&M serviceability	80%	99.6%	100.0%	100.0%	100.0%

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

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

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

Note that Dr Charlie Laughton's project *Simulation of Radioprobing*, which used to be coded as b07, has been recoded as e26, reflecting its funding by EPSRC.

EPSRC Projects

Code	*Class*	*Title*	*PI*
e01	1	UK Turbulence Consortium	Prof Neil Sandham
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
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
e12	1	Parallel programs for the simulation of complex fluids	Dr Mark R Wilson
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
z09		HECToR Benchmarking	Dr Edward Smyth

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
n05	2	Non-linear Wave-particle Instabilities in Plasmas	Dr Mervyn Freeman

BBSRC Projects

Code	*Class*	*Title*	*PI*
b02	1	Modelling enzyme catalysis	Dr Adrian J Mulholland
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

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 J-C Desplat
x02	OHM Ltd	Mr Mark Westwood

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
e09	2	Molecular Properties and their Geometry	Prof Peter Taylor
e13	1	TeraGyroid project	Dr Richard J Blake