HPCx Service
Summary of Annual Report for 2004

1 Summary of Progress

1.1 Overview

Clearly the most significant feature of 2004 from the perspective of the HPCx project was the upgrade to Phase 2. This was the dominant focus of our Annual Plan and of the work of both the Science Support teams and Systems team throughout the year.

The upgrade to Phase 2 provided a doubling of the capability of the system. This required:

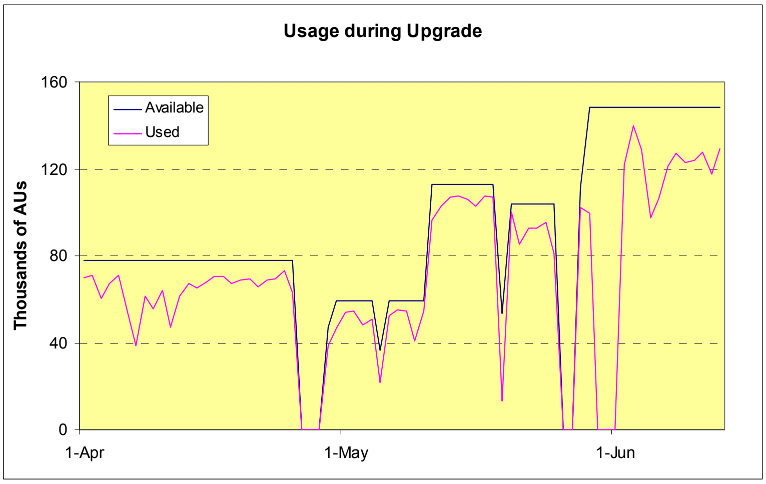
- Complete replacement of all frames to deliver faster processors
    - 40 p690 frames to 50 p690+ frames (for a total of 1600 CPUs)
- New service model to utilise 32-way LPARs
- Complete replacement of the switch with new High Performance Switch
- New parallel support environment and AIX upgrade
- Addition of 18 Tbytes disc and 24 LTO tape drives
- Migration of all existing GPFS data

This was probably the major technical challenge of the whole HPCx project. We agreed a phased transition plan to ensure early access to the technology, to minimise disruption to users and to allow effective risk management. This phased transition plan was very successful and we were able to stick very closely to the timetable. Phase 2 passed all its implementation tests and was accepted into service on schedule on 17 June 2004. HPCx was then ranked 18th in the TOP500 list of the world's fastest supercomputers which made it once again the most powerful academic supercomputer in Europe.

To maximise the scientific benefits from the enhanced system to the UK research community, the Science Support teams ensured that:

- Users had early access to a document containing all the information needed to transfer to Phase 2
- Many key applications were ported and tuned on Phase 2 right from the start of service.

The success of this work is shown in the following graph of daily usage during the transition period. It shows that the total amount of downtime was low and that utilisation remained around 75% throughout. There was minimal lag to the usage as the system increased in performance and the levels of capability usage were also higher than anticipated.



Completing this upgrade on time and with only modest disruption to users is a major source of satisfaction and pride for all involved. The upgraded system has significantly increased capability and the new interconnect has high bandwidth and much-improved latency. We are confident that this should allow an even broader range of applications to effectively use 512 processors or more and that this should ensure that the remaining 4 years of the project see the system deliver even more high quality science. This should be a very exciting and productive period for the UK computational research community.

1.2 Key Objectives

The Annual Plan for 2004 identified the following specific objectives and targets and this section indicates the progress we have made.

1. *Identify key codes within the various application areas and collaborate with the users to ensure that they can perform effectively on large numbers of processors under the new High Performance Switch (HPS). We will characterise the performance and identify any performance bottlenecks for at least ten such codes.*

We have investigated the performance and determined the bottlenecks to scaling for a range of application codes from areas such as environmental modelling, materials chemistry, physics and engineering. These included: CENTORI, SIESTA, LUDWIG, H2MOL, GAMESS-UK, NWChem, NAMD, AMBER, LAMMPS and GROMACS.

2. *During the course of this year, we will work to produce an average performance improvement of at least 25% on 128 CPUs for at least five codes from the following range of areas: Chemistry, Environment, Engineering and Physics.*

We have worked to improve the performance of various applications from the chemistry, environment, engineering and physics environments. In particular, we have focused on the following 5 applications: SIESTA, NWChem, CENTORI, GAMESS-UK, and LUDWIG.

Application	*Original Execution Time (seconds)*	*Optimised Execution Time (seconds)*	*Performance Improvement (%)*
SIESTA	1288	414	211
NWChem	3348	1434	133
CENTORI	233	155	50
GAMESS-UK	5129	3788	35
LUDWIG	0.84	0.72	17
Average Performance Improvement across all 5 codes (%)			*89.2*

Table 1: Performance improvement for five applications on 128 processors, computed as (Original/Optimised – 1) so that 100% improvement corresponds to one half of the execution time. LUDWIG results are per iteration.

3. *Investigate the key elements that characterise the performance of the Phase 2 system to ensure that users can benefit from the enhanced system. This should include:*
   - *Low-level communication protocols;*
   - *Shared memory and mixed mode programming;*
   - *Data handling;*
   - *Provision and support of optimised library routines, eg, MPI collectives and parallel multi-dimensional FFTs.*

During 2004, we performed a series of investigations into the performance of the Phase 2 system.

- The low level communication performance has been studied extensively and the results of this investigation have been reported in a number of technical reports
- We have investigated the shared memory performance characteristics of HPCx using the new Version 2.0 of our OpenMP Micro-benchmarks and delivered a technical report comparing them with those on other shared memory platforms.
- We have delivered a technical report on the use and performance of mixed mode programming on HPCx.
- For data handling, we produced a technical report on the tape archiver, investigated MPI-IO performance and installed the Storage Resource Broker (SRB) client toolkit.
- We have produced a benchmarking report on the performance of parallel eigensolvers and FFTs on Phase 1 and Phase 2.
- We also developed a general technique for optimizing global exchange operations that are commonly used to implement parallel FFTs.

4. *During the transition to Phase 2, it is important to ensure that the key elements of a successful service are in maintained, ie,*
   - *Continue to provide a prompt and accurate response to queries;*
   - *Keep the documentation portfolio up-to-date during and after the transition to Phase 2;*
   - *Publicise the HPCx service to encourage uptake through coordination and participation in newsletters, user groups and events, including an international conference.*

These targets were the major focus for the applications support team: this year's transition between Phase 1 and Phase 2 represented the biggest single challenge for the six-year HPCx service. Providing prompt and useful information to users is essential to enable their science programmes to move smoothly between the systems. This was successfully achieved, with the highlights including:

- answering more than 99.5% of non-indepth queries within 3 days, with 85% answered in less than 24 hours
- providing comprehensive documentation throughout the upgrade process, available separately for early users of Phase 2 and fully integrated immediately it became the production system
- distributing over 7000 copies of two newsletters by post and at key events; running a user group; two workshops and an Annual Seminar with international speakers; contributing talks to numerous HPC events worldwide.

5. *Facilitate at least two high profile, world-leading calculations as demonstrators of capability computing. Likely candidates include: high-resolution coastal ocean simulations using POLCOMS; and further TeraGrid experiments.*

There were a number of high-quality capability computing calculations from a range of applications areas performed using HPCx during the last year, including:

- High Resolution Continental Shelf modelling with POLCOMS
- GAMESS-UK and the Modelling of Isocitrate Lyase

Unfortunately, there were no further Teragrid experiments this year although HPCx are represented on the Steering Committee overviewing a Call for Proposals for future experiments; this should result in further capability computing demonstrators during 2005.

6. *Produce at least twelve technical reports covering the topics discussed in 2.1.3, with the following number of reports in each quarter: 2 (Q1), 4 (Q2), 3 (Q3), 3 (Q4).*

We have produced a total of 17 reports this year (quarterly breakdown: 3, 6, 4, 4), exceeding the target in every quarter. The topics cover the areas suggested in the Annual Plan, chosen to give as much information as possible to users on the new features of the Phase 2 system and how best to exploit its enhanced capabilities. We have produced both detailed low-level studies of system characteristics, as well as case studies of how to apply this knowledge to the benefit of real applications. At the request of STAC, we also wrote a report on the tape archive system.

7. *Ensure that the HPCx service plays a key role in the UK's National Production Grid which is due to come into operation during 2004.*

HPCx has been active in the formation of the UK National Production Grid (now called the National Grid Service, NGS). We are involved in the NGS Technical board meetings via Access Grid that are used to manage the technical aspects of the NGS. We have maintained a version of the Globus Toolkit on HPCx that is compatible with the rest of the NGS and have added local mechanisms to make it easier for our users to register their Grid certificates for use on HPCx. As mentioned previously, we have also installed SRB to allow users to access remote data repositories from HPCx via the Grid.

8. *We will use the lessons from the first year of the training activity to maximise the benefits to users. Key elements will include:*
    - *Offering easier access to training material by running courses at a larger number of sites;*
    - *Promoting the use of profiling, debugging and other useful tools to users via HPCx-specific examples and case studies;*
    - *Developing 2 new tailored courses, including one focused on the new HPS.*

The training programme has matured significantly since the start of service. We now have three bespoke courses focusing on the HPCx architecture, in addition to the many covering parallel programming and applications-specific topics. The HPCx community mostly comprises experienced users, so the demand for training tends to be at quite a low level. Even so, it is essential to provide a wide range of courses at convenient times and locations. Although registered users always get top priority, we make strong efforts to fill free spaces with interested UK academics by advertising nationwide by email. As a result, the training programme also represents an important outreach activity to potential new users.

The highlights this year have included:

- providing 30 course-days at five locations around the UK (London, Loughborough, RAL, Daresbury and Edinburgh), thereby delivering in excess of 600 student-days of training;
- having a strong focus on the practical use of tools in all the HPCx-specific courses;
- the development of two new courses, including an advanced course focusing on exploiting the enhanced capabilities of Phase 2, especially the HPS.

9. *Ensure that the science support effort is dominated by experienced staff and that this experience benefits the other staff and the user community at large.*

Science support effort has continued to be delivered primarily by the same staff who are experienced and well-known to the user community. Some 85% of the science support effort has come from the staff named in the contract; this represents 116% of the agreed Full Service Level. The senior support staff regularly provide advice and guidance for the other staff. The following science support staff spent more than 50% of their time working on HPCx: Mike Ashworth, Stephen Booth, Mark Bull, Ian Bush, John Fisher, Joachim Hein, David Henty, Adrian Jackson, Martin Plummer, Gavin Pringle, Fiona Reid, Alan Simpson, Lorna Smith, Andy Sunderland.

To ensure that the experience of the staff benefits the user community:

- Technical queries are answered by relevant experts from the science support team rather than a separate helpdesk team
- Results from the science support are disseminated to users via technical reports and presentations
- Key experienced staff are active in the user community and are readily available to users via all user groups, seminars and workshops

10. The outreach activity will focus on the following key areas:

- *Establishing the Life Sciences initiative within the BBSRC and MRC areas;*

The programme is now well established with three projects in full scientific production, projects in gene pathway decomposition and radiation damage about to start, at least two life-science projects bidding for the UK-US Computational Experiment to be showcased at SC 2005, two major new projects developing heart simulations and modelling the retina to be supported in 2005 and interest being explored in the BBSRC community to support the formation of a Collaborative Computational Project in Biomolecular Simulation.

- *Initiating discussions with the industrial community;*

The industry day was reasonably well attended and is being followed up with discussions on requirements with various industrial customers and commercial software vendors. A proposition governing Access-Codes-Solutions has been developed and will be marketed to up to 20 parties in 2005.

- *Improving public awareness;*

HPCx forms part of the core presentation to school and HEI visits. There was a press release on the successful Teragyroid experiment and we also disseminated scientific results from HPCx via the Annual Seminar and a special issue of the newsletter.

- *Supporting Grid projects exploiting HPCx and seeking to undertake joint experiments with the TeraGrid sites.*

No joint UK-US experiments were planned for 2004, although we are represented in the Steering Committee for demonstrators for 2005. We also provide ongoing support for leading HPC Grid users including RealityGrid, BioSimGrid and Integrative Biology e-science projects.

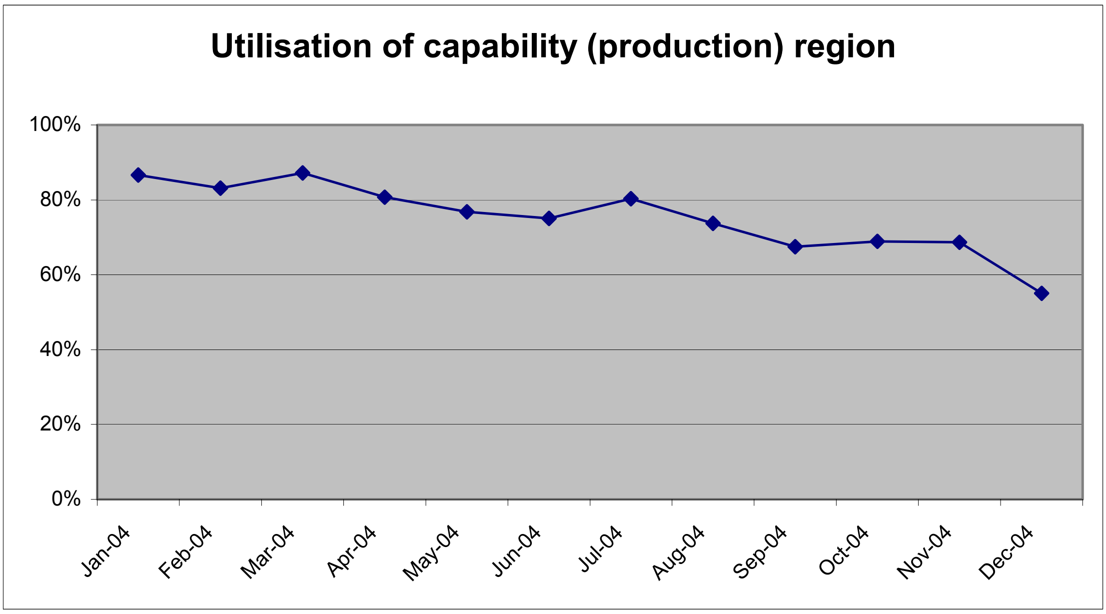
2 Quantitative Metrics

2.1 Utilisation

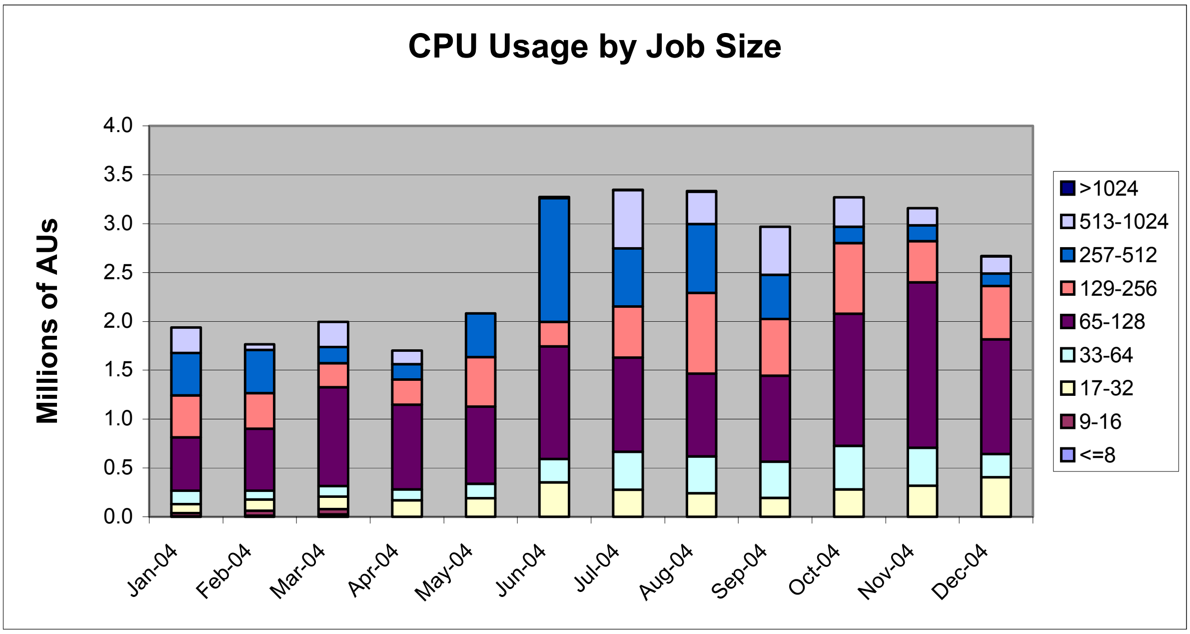
2.1.1 Overall

The monthly utilisation for the production region is shown in the graph below. This has averaged 75.3% for the year, which generally good for a capability computing service, although the recent lower usage is a subject of concern.

Note that on Phase 1, this region had 1024 processors; on Phase 2 it has 1280.



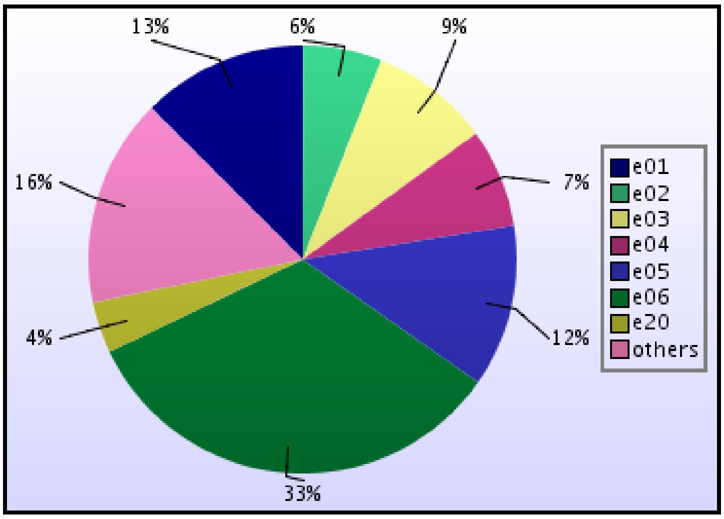
2.1.2 Usage by Job Size



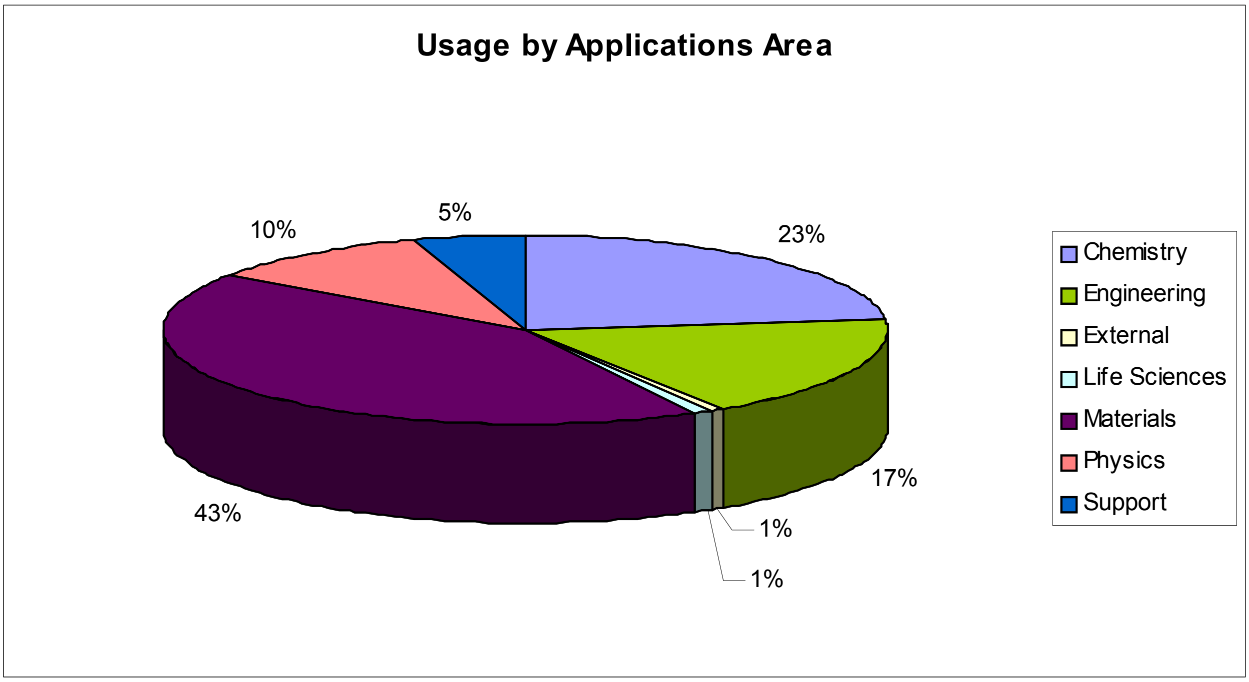
The above graph illustrates the way that job sizes have varied throughout the last year.

2.1.3 Usage by Consortium

A list of consortia and PIs is given in Appendix A.



2.1.4 Usage by Applications Area



2.2 Helpdesk

Category	*Number*	*% of all*
Administrative	467	36.3
Technical	743	57.8
In-depth	65	5.1
PMR	10	0.8
Total	1285	100.0

All non-indepth queries	*Number*	*%*	*Target*
Finished within 24 Hours	1209	85.0	75%
Finished within 72 Hours	1206	99.7	97%
Finished after 72 Hours	4	0.3	

Administrative queries	*Number*	*%*	*Target*
Finished within 48 Hours	465	99.6	97%
Finished after 48 Hours	2	0.4	

2.3 Performance Metrics

Metric	*TSL*	*FSL*	*January*	*February*	*March*	*April*	*May*	*June*	*July*	*August*	*September*	*October*	*November*	*December*	*Annual Average*
Technology serviceability (%)	80	99.2	100.0%	99.8%	99.6%	100.0%	98.4%	98.0%	99.8%	99.7%	100.0%	100.0%	99.9%	100.0%	99.6%
Technology MTBF (hours)	200	300	∞	366	732	∞	244	366	732	732	732	∞	1464	∞	764
Number of AV FTEs	7.5	10	14.2	14.4	16.1	13.9	13.0	15.5	14.7	11.2	13.2	14.0	14.2	10.3	13.7
Number of training days per month	22.5/12	30/12	5/1	8/2	11/3	16/4	21/5	22/6	23/7	23/8	24/9	28/10	30/11	30/12	30/12
Non in-depth queries resolved within 3 days (%)	85	97	100.0%	98.7%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	99.0%	100.0%	98.6%	97.7%	99.5%
Number of A&M FTEs	3.75	5.75	5.9	5.8	6.6	5.8	5.8	6.0	6.2	6.0	6.3	4.9	6.6	5.4	6.0
A&M serviceability (%)	80	99.6	100.0%	97.8%	99.7%	100.0%	100.0%	99.3%	98.8%	100.0%	99.4%	100.0%	99.9%	97.8%	99.4%

Colour	*Meaning*
(green)	Exceeds FSL
(amber)	Between TSL and FSL
(red)	Below TSL

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

Appendix A: Projects

A.1: Current Projects

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
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
b07	1	Simulation of Radioprobing	Dr Charlie Laughton

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

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

A.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