HPCx Service Final Report
Dec 2002 - January 2010

Contents

1 Introduction

1.1 Overview

By the key metrics of utilisation and reliability, HPCx has been a remarkably successful service. Throughout its whole lifetime, utilisation averaged more than 70%, reaching an all-time high of 96% in November 2008.

Since 2007, there has not been a single service failure on the main system; this is a record that it is unsurpassed on any comparable service in the world. The research success of the service is clearly demonstrated from the high utilisation and from the case studies reported in sections 2.3 and 3.5. As included in EPSRC's 2005 International Review of Research using HPC*, "The UK has been very well served by the leading edge HPCx…computational system"*. During the period that HPCx has run alongside the new national service on HECToR, the concept of Complementary Capability Computing has been developed. The service innovations introduced during this period have been very successful ensuring that HPCx leaves a lasting legacy for future national and international HPC services.

1.2 Key Lessons Learned

This section briefly highlights key lessons learned from the HPCx service. More discussion of these issues can be found throughout the rest of the document.

- *Flexibility:* taking a flexible, responsive approach to service provision is vital to ensure that the service better matches researchers' needs.

- *Partnerships:* formed before the start of service result in fewer contractual revisions and a more closely integrated approach, e.g., between hardware providers and service staff, and between science support and the helpdesk.

- *Charging Mechanisms:* are important influences in determining users' behaviour which can result in resource wastage.

- *Low Priority Access:* can help ensure that cycles are used for research that would otherwise be unused during quieter periods.

- *Long Jobs:* these have been consistently popular with a wide variety of user groups, although they are difficult to accommodate on a single national service.

- *Training:* the most consistent demand is for introductory training for new users. Attendance can be significantly enhanced by combining events and by running courses at locations convenient for users.

- *Data:* is a key output of national HPC services and it is important to have a long-term plan for storage and access.

- *Grid:* while this may have an important role in capacity computing, a single job submission interface has only a limited role to play in national services, due to their unique capabilities and to differences in allocation culture.

- *Helpdesk:* building long-term relationships with users through effective communication is essential.

- *SAFE:* The Service Administration from EPCC software significantly reduces admin overhead to provide a better quality, more cost-effective service.

- *Scientific Visualisation:* this can have a key role in aiding public understanding of science and helping to make the case for future services.

- *Commercial Usage:* it is challenging to run industrial and academic applications on a single system as commercial applications often scale poorly and commercial users are more concerned over data security. Collaborative research is often a better option.

- *Terascaling:* investment in this area helps ensure cost-effective use of major facilities through enhanced performance.

- *Upgrades:* a planned series of upgrades can ensure that the service remains competitive throughout its lifetime.

2 HPCx Phase 4 Review

HPCx Phase 4 ran from December 2008 to January 2010. The key aim of Phase 4 was to provide a complementary capability computing service to HECToR trading overall utilisation in favour of a more flexible service. Phase 4 should also provide useful information for running future services. In this section we review the metrics and achievements made during this final phase of the service.

2.1 Phase 4 Quantitative Metrics

2.1.1 Reliability

The following table shows the number of incidents in Phase 4 at the various severity levels.

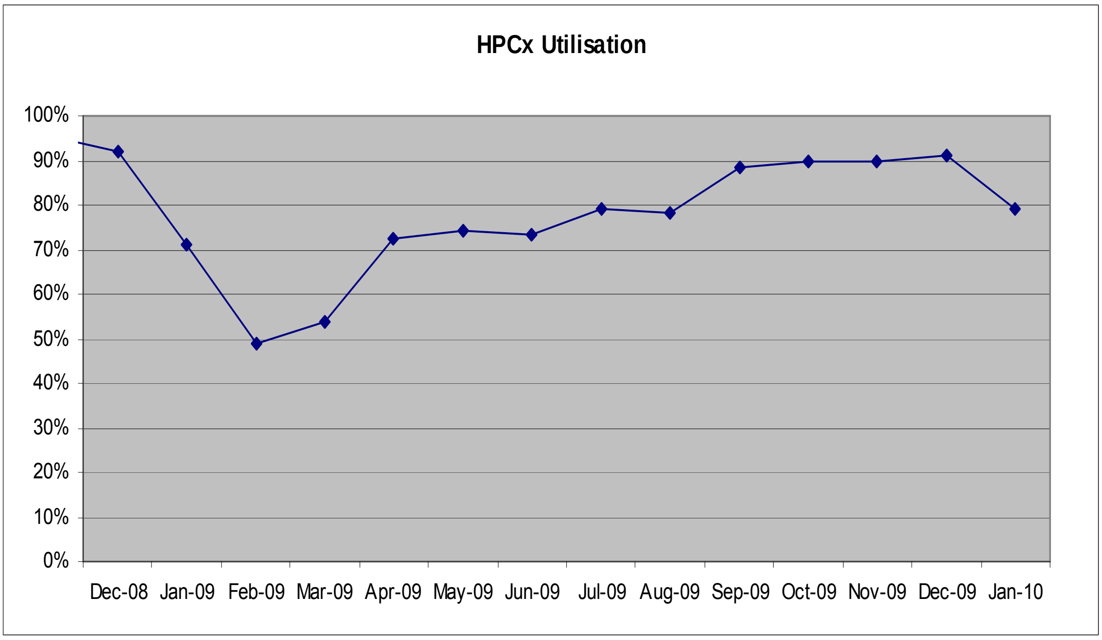
Severity	Number
1	1
2-4	72
Total	73

HPCx reliability was yet again exceptional. There was only one hardware failure during Phase 4. This was a failure of the tape archive and no user jobs were affected.

The following table shows the number of failures, the mean time between failures and the serviceability. These are then attributed to the site, to IBM and to external factors.

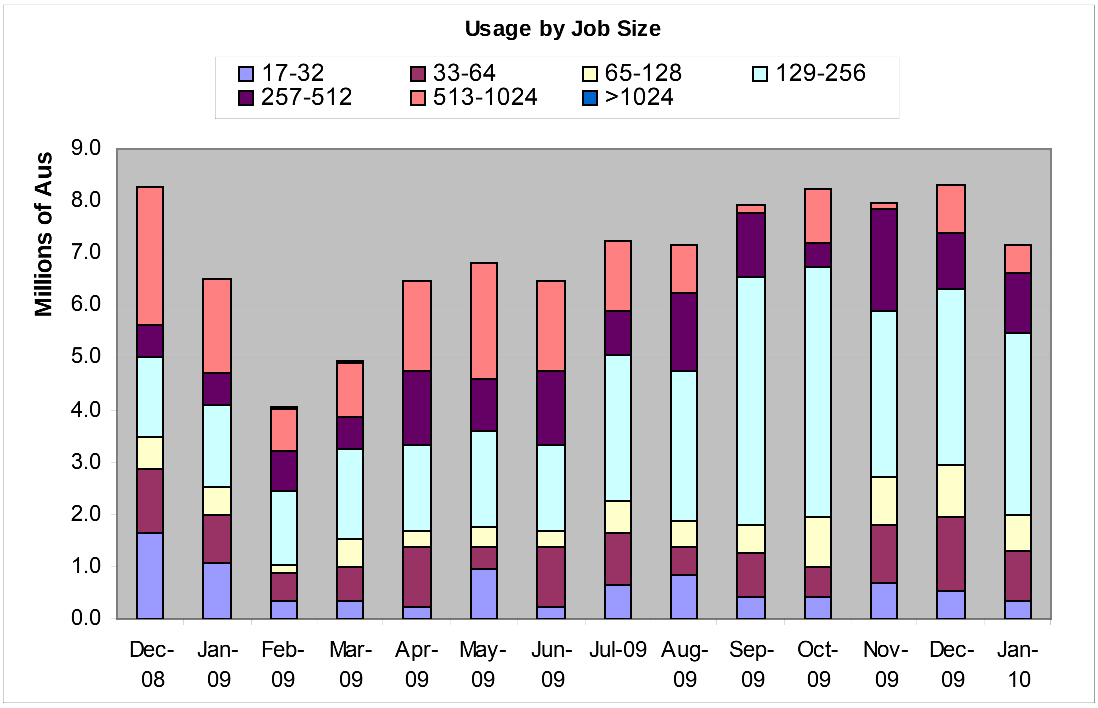
Attribution	N	MTBF	Serviceability
IBM	1	10248	99.96%
Site	0	∞	100.00%
External	0	∞	100.00%
Total	1	10248	99.96%

2.1.2 Utilisation



HPCx remained well utilised through to the very end of service. The average utilisation in Phase 4 was 77%.

2.1.2.1 Usage by Job Size



2.1.2.2 Usage by Consortium

Project	AUs	Raw AUs	Discounted AUs	Number of Jobs	%age of Use	Raw %age	Utilisation
e01	5,082,198	12,692,351	7,610,153	8815	9.65%	13.01%	10.06%
e05	6,556,600	9,917,517	3,360,917	8172	12.45%	10.17%	7.86%
e08	99,585	99,585	0	169	0.19%	0.10%	0.08%
e10	52,888	52,888	0	21	0.10%	0.05%	0.04%
e18	16,761	16,761	0	26	0.03%	0.02%	0.01%
e19	1,212,587	1,362,019	149,432	573	2.30%	1.40%	1.08%
e20	0	48,794	48,794	9	0.00%	0.05%	0.04%
e24	3,937,526	3,937,526	0	25403	7.48%	4.04%	3.12%
e30	210,449	2,830,555	2,620,106	1166	0.40%	2.90%	2.24%
e33	1,105,831	1,341,585	235,753	649	2.10%	1.38%	1.06%
e34	1,221,702	1,227,097	5,395	648	2.32%	1.26%	0.97%
e35	370,033	396,043	26,010	2474	0.70%	0.41%	0.31%
e36	297,573	402,237	104,664	783	0.57%	0.41%	0.32%
e37	999,939	1,257,620	257,681	276	1.90%	1.29%	1.00%
e38	1,304	1,304	0	211	0.00%	0.00%	0.00%
e39	247,144	279,333	32,190	379	0.47%	0.29%	0.22%
e40	44,372	251,445	207,073	23	0.08%	0.26%	0.20%
e41	99,837	491,157	391,320	27	0.19%	0.50%	0.39%
e42	24,774	50,299	25,525	89	0.05%	0.05%	0.04%
e44	296,625	335,606	38,981	161	0.56%	0.34%	0.27%
e46	513,536	735,538	222,002	1131	0.98%	0.75%	0.58%
e50	215,531	231,176	15,645	197	0.41%	0.24%	0.18%
e58	224,479	2,116,599	1,892,120	4073	0.43%	2.17%	1.68%
e59	68,969	958,932	889,963	59	0.13%	0.98%	0.76%
e60	2,145,853	3,681,420	1,535,567	2440	4.07%	3.77%	2.92%
e62	0	18,013	18,013	31	0.00%	0.02%	0.01%
e63	1,740,663	7,018,197	5,277,534	1911	3.31%	7.19%	5.56%
e64	4,221,466	7,737,016	3,515,550	954	8.02%	7.93%	6.13%
e65	1,332,473	1,493,442	160,969	3283	2.53%	1.53%	1.18%
e66	795	795	0	5	0.00%	0.00%	0.00%
e67	928,662	1,295,077	366,415	618	1.76%	1.33%	1.03%
e68	932,305	2,651,002	1,718,697	1878	1.77%	2.72%	2.10%
e69	269,138	269,138	0	7	0.51%	0.28%	0.21%
e94	19,333	19,333	0	13	0.04%	0.02%	0.02%
e96	138,765	138,765	0	195	0.26%	0.14%	0.11%
e97	346,115	346,115	0	139	0.66%	0.35%	0.27%
e98	1,987,229	2,016,263	29,034	595	3.77%	2.07%	1.60%
e99	2,110,615	14,641,030	12,530,415	9383	4.01%	15.01%	11.60%
e100	2,672,675	2,672,675	0	2953	5.07%	2.74%	2.12%
e101	1,656,670	1,656,670	0	545	3.15%	1.70%	1.31%
e102	2,530,815	3,366,879	836,064	393	4.81%	3.45%	2.67%
e103	3,296,660	3,296,660	0	223	6.26%	3.38%	2.61%
e104	101,105	101,108	3	53	0.19%	0.10%	0.08%
e105	22,500	24,067	1,567	764	0.04%	0.02%	0.02%
e106	0	0	0	3	0.00%	0.00%	0.00%
EPSRC Total	**49,354,080**	**93,477,635**	**44,123,555**	**81920**	**93.71%**	**95.82%**	**74.07%**

Project	AUs	Raw AUs	Discounted AUs	Number of Jobs	%age of Use	Raw %age	Utilisation
n01	1	1	0	1778	0.00%	0.00%	0.00%
n02	21,331	21,331	0	10	0.04%	0.02%	0.02%
NERC Total	**21,332**	**21,332**	**0**	**1788**	**0.04%**	**0.02%**	**0.02%**
p01	882,704	901,377	18,673	5169	1.68%	0.92%	0.71%
c01	1,287,262	1,290,864	3,602	9327	2.44%	1.32%	1.02%
STFC Total	**2,169,966**	**2,192,241**	**22,275**	**14,496**	**4.12%**	**2.24%**	**1.73%**
d01	320	320	0	128	0.00%	0.00%	0.00%
d02	119,342	119,342	0	4	0.23%	0.12%	0.09%
x05	365	365	0	50	0.00%	0.00%	0.00%
z001	250,178	296,473	46,296	4139	0.48%	0.30%	0.23%
z002	530	530	0	20	0.00%	0.00%	0.00%
z004	15,572	713,368	697,796	1411	0.03%	0.73%	0.57%
z06	6,539	6,539	0	307	0.01%	0.01%	0.01%
z07	5	5	0	1	0.00%	0.00%	0.00%
z10	7	7	0	8091	0.00%	0.00%	0.00%
HPCx Total	**392,856**	**1,136,948**	**744,092**	**14151**	**0.75%**	**1.17%**	**0.90%**
x01	727,705	727,705	0	4582	1.38%	0.75%	0.58%
External Total	**727,705**	**727,705**	**0**	**4582**	**1.38%**	**0.75%**	**0.58%**
Total	**52,665,939**	**97,555,861**	**44,889,922**	**116937**	**100.00%**	**100.00%**	**77.30%**

A total of 126,207,898 AUs were available during Phase 4.

2.1.3 Performance Metrics

Metric	TSL	FSL	December	January	February	March	April	May	June	July	August	September	October	November	December	January	Value for Phase 4
Technology serviceability (%)	80%	99.2%	100.0%	97.2%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	99.8%
Technology MTBF (hours)	200	300	∞	732	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	10248.0
Total AV FTEs	-	74.0	6.8	6.4	7.0	8.7	7.0	6.6	6.8	5.4	4.7	5.2	4.3	2.7	1.8	2.1	75.5
Non in-depth queries resolved within 3 days (%)	85%	97%	98.0%	97.8%	97.7%	98.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	98.0%	100.0%	100.0%	100.0%	99.3%
Number of A&M FTEs	2.75	4.75	4.8	4.8	4.8	4.8	4.8	4.8	5.1	4.8	4.8	4.8	4.9	5.0	4.8	4.8	4.9
A&M serviceability (%)	80%	99.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

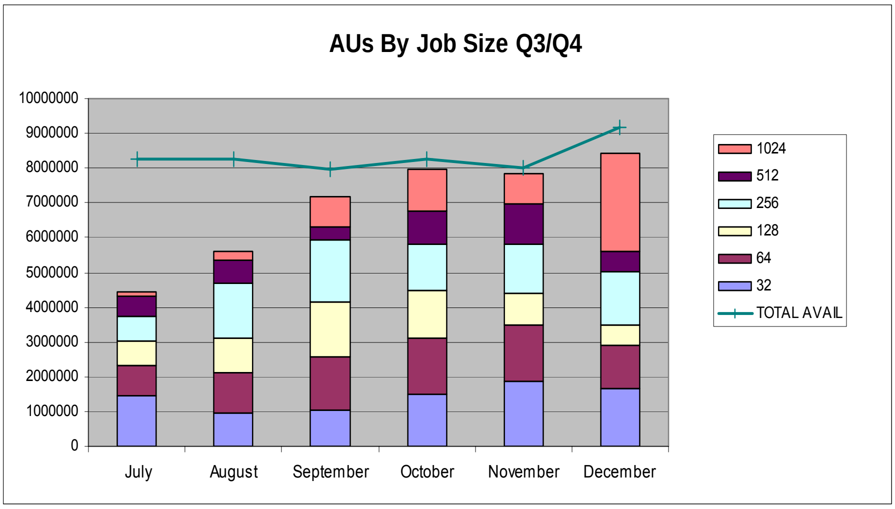
Colour	*Meaning*
(green)	Exceeds FSL
(amber)	Between TSL and FSL
(pink)	Below TSL

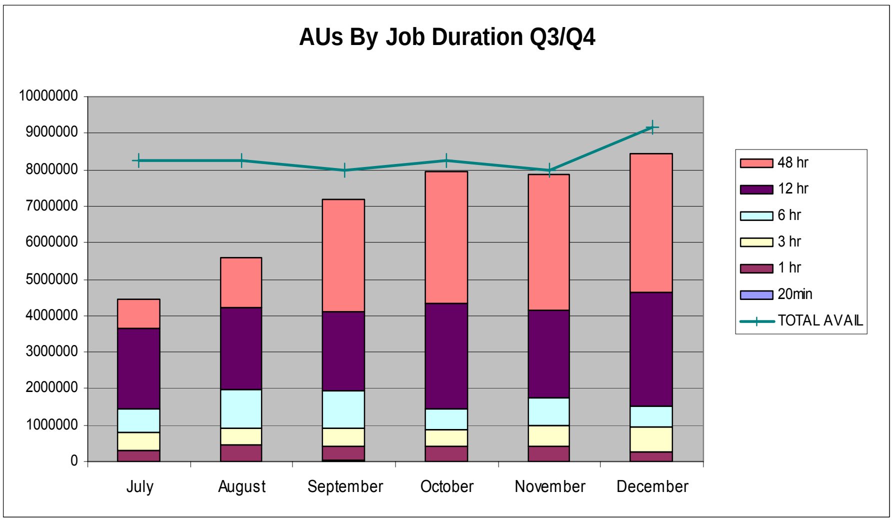
**The staffing profile in Phase 4 varies from month to month*
Target AV FTEs in Phase 4 74.0

2.2 Phase 4 Initiatives

2.2.1 Open Access Initiative

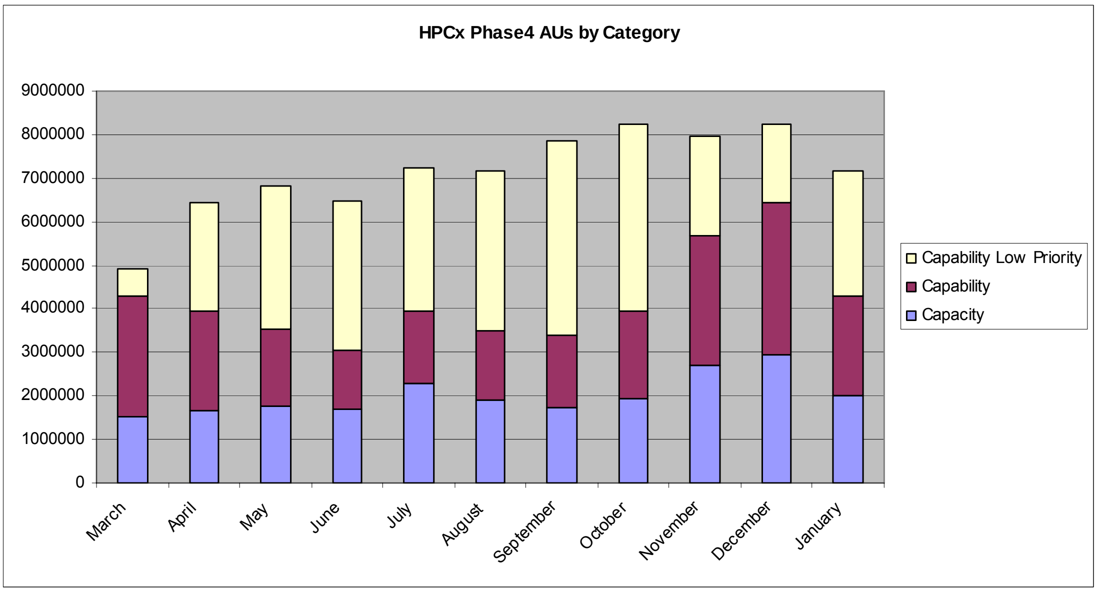
In August 2008, the EPSRC Open Access Initiative was launched with the aim of encouraging users to explore complementarity. During this initiative, charging was suspended. The graphs below illustrate the impact of this initiative on utilisation of HPCx. This initiative ran from August 2008 to January 2009.





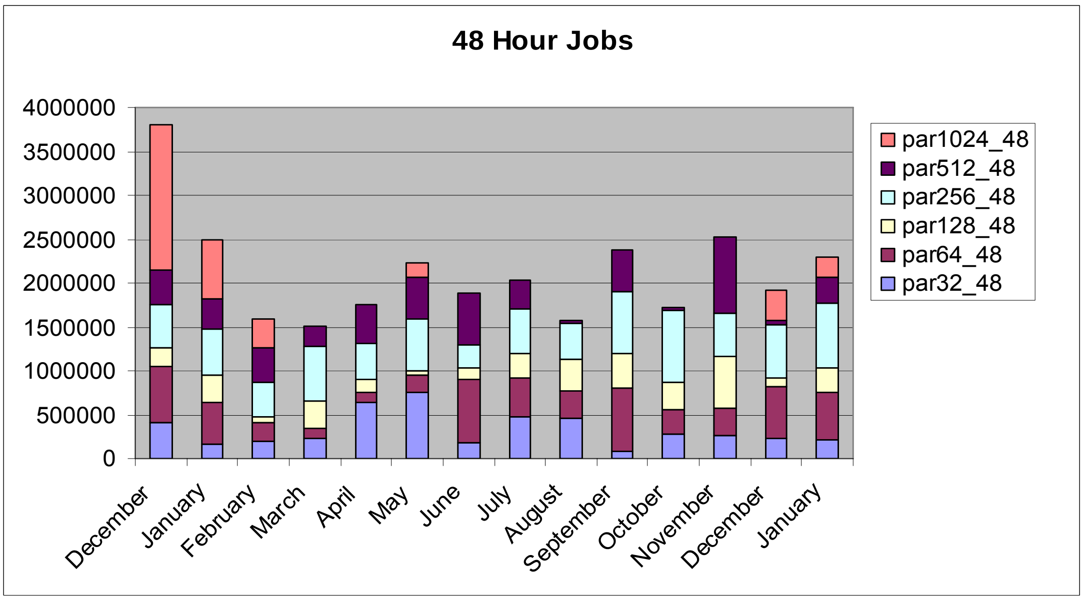
During the Open Access Initiative, 13.4 Million AUs were used over and above the average utilisation of the previous months (60%). This equates to a notional cost of £1.34 Million.

2.2.2 Low Priority Access



Low priority access queues were introduced on 25th March. These allowed short capability jobs to run when the system was quieter. A total of 32.5 million AUs were used by these jobs at a notional cost of ~£3.4 million. In Phase 4, 32% of the overall AUs used were low priority.

2.2.3 Long Jobs



48 hour queues were introduced as part of Phase 4. These were used extensively during the Open Access Initiative when charging was disabled. These long jobs were not part of the Low Priority Access initiative; as such there are were less of the very large 1024 processor, 48 hour jobs in 2009.

2.2.4 Advance Reservations

Advance reservations were used by the FireGrid project to secure resources on HPCx as part of a large-scale fire simulation. This was not one of the CCC funded projects, but was awarded time by the Service Director in support of Complementarity. The project aimed to develop a prototype Integrated Emergency Response System. This is working towards a system in which live emergency situations, such as building fires, can be modelled in super-real-time using HPC resources, such that predictions from the simulations can be used to make informed decisions regarding the response. For example, had such a system been in place in the World Trade Centre before the September 11th 2001 incident, it may have been possible to predict that the buildings would collapse in time to evacuate.

This project utilised HPCx, along with local University of Edinburgh cluster resources, for successful live demonstrations involving real fires. We implemented advanced reservation functionality on HPCx to ensure that the resource would be available at the correct time.

A full article on FireGrid is available in the final edition of the HPCx newsletter.

2.3 Complementary Capability Computing

The main focus of the science support activities during the HPCx extension was the support of those HPCx projects funded under the Complementary Capability Challenge call.

2.3.1 e96 Coupled Real-Time Fluid and Flight Simulation

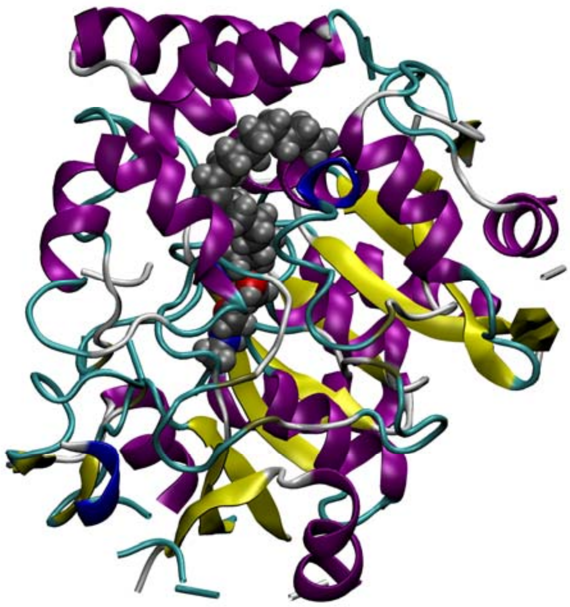
This project aimed to perform flight simulations in real time, where the HPCx calculations are computationally steered by human operation of a flight simulator in Southampton, to allow the evaluation and improvement of air vehicle design by studying operation in flight situations more challenging than otherwise possible.

The consortia did not take up our offers of a visit at the outset or the provision of in-depth support, but we did assist the project with port-forwarding on HPCx. This enables a connection from a machine outside the firewall to a listening compute node on HPCx. Port-forwarding has been used previously on HPCx by users of the NAMD application.

2.3.2 e97 Multiscale Ensemble Computing for Modelling Biological Catalysts

This project involved using HPCx to perform a detailed investigation of the mechanism of chemical reactions catalysed by the enzyme fatty acid amide hydrolase, an important target for drug development, using mixed Quantum Mechanics/Molecular Mechanics (QM/MM) techniques. Multiple individual simulations had to be coupled into a single supra-simulation.

Kenton D'Mellow visited Dr Mulholland and Dr Woods at Bristol University on 13th March. Before the researchers could proceed with production runs on HPCx, they had to complete the porting of two coupled applications, Sire and Molpro: an extremely complicated task. Sire is a C++ MM code, developed by the researchers, with a python interface, and has the ability to spawn and interact with instances of Molpro (QM) via remote procedure calls (RPC), disk access, and socket communications. We confirmed that the proposed range of inter-communication methods between the codes could all work in principle on HPCx, and we specified that disk-based I/O methods are preferable for this work. The code manages its workload rather well, but requires occasional (daily) monitoring to catch and terminate/reschedule unresponsive instances of Molpro. We advised on the best practice for performing this and developed improvements to the current implementation of manual identification of these "bad jobs".



Fatty Acid Amide Hydrolase (FAAH), an important target for the development of drugs to treat pain, anxiety and depression.

Since the simulation was expected to take several weeks we advised on a strategy involving the use of a combination of check-pointing and the long 48 hour queues on HPCx. We also recommended that they make use of the soft limit to direct a clean termination of the code before it is killed. We advised on the best configuration in terms of tasks and threads per node, and also that there might be scope for utilising SMT given that Molpro and Sire will need to be concurrently active.

We discussed code optimisation with the researchers and gave a demonstration of Vampir as a profiling and analysis tool.

The researchers proceeded to complete the complex setup which involved comprehensive porting and code refactorisation. They then performed test runs, finalised the wrappers required for simulation, and set up the production runs. At this point it became clear that the science behind the simulations had advanced, so the users tested some local small-scale simulations, which they then scaled up on HPCx. We advised on the best job profile to meet their needs in doing this, such that they were in a position to move into production mode.

A full article on this project is available in the final edition of the HPCx newsletter.

A continuation of this project has been funded under the EU DEISA Extreme Computing Initiative, which will allow the researchers to further this work after HPCx is decommissioned. They will have access to a range of European machines, and it is expected that they will choose an IBM service similar in nature to HPCx, such that the comprehensive porting performed on HPCx will be transferable, allowing a smooth transfer of work and the ability to move into immediate production on the new system.

2.3.3 e98 Global Turbulence Simulations of Spherical Tokamak Physics

This project involved addressing key issues in the quest towards viable fusion power plants. The work involved extended time simulations on HPCx to study electromagnetic plasma turbulence.

Joachim Hein visited Dr Thyagaraja and his colleagues at UKAEA Culham on 29 January. By the time of the visit the project was already in production on HPCx. As expected, the researchers had not encountered any significant issues in porting: the required libraries were already supported on the system.

The typical run was between 256 and 1024 tasks, and they utilised the long queues on HPCx: the typical job-length was 48 hours. The HPCx queue setup was sufficient for the project.

The preferred mode of working was to transfer the results from HPCx back to local machines for post-processing. We enquired about a previous issue they had regarding poor network bandwidth between HPCx (i.e. the Janet network) and the Culham network, which hampered these transfers and caused delays. They advised that these issues were now resolved, and they now could efficiently make such transfers.

The researchers reported that, on HPCx, the system buffers data in a way which sometimes causes application result data to be lost when a checkpoint/restart technique is used (as opposed to the execution of full single runs). We provided a solution to this problem involving adapting the application to explicitly flush the data at appropriate points.

On the whole the project ran very smoothly. We held regular phone meetings with the researchers, to ensure that they were not being slowed down by any problems. They used up the compute budget at the required rate and reported that the simulations were producing some excellent scientific results.

2.3.4 e99 Computer Simulations for Chemical Biology

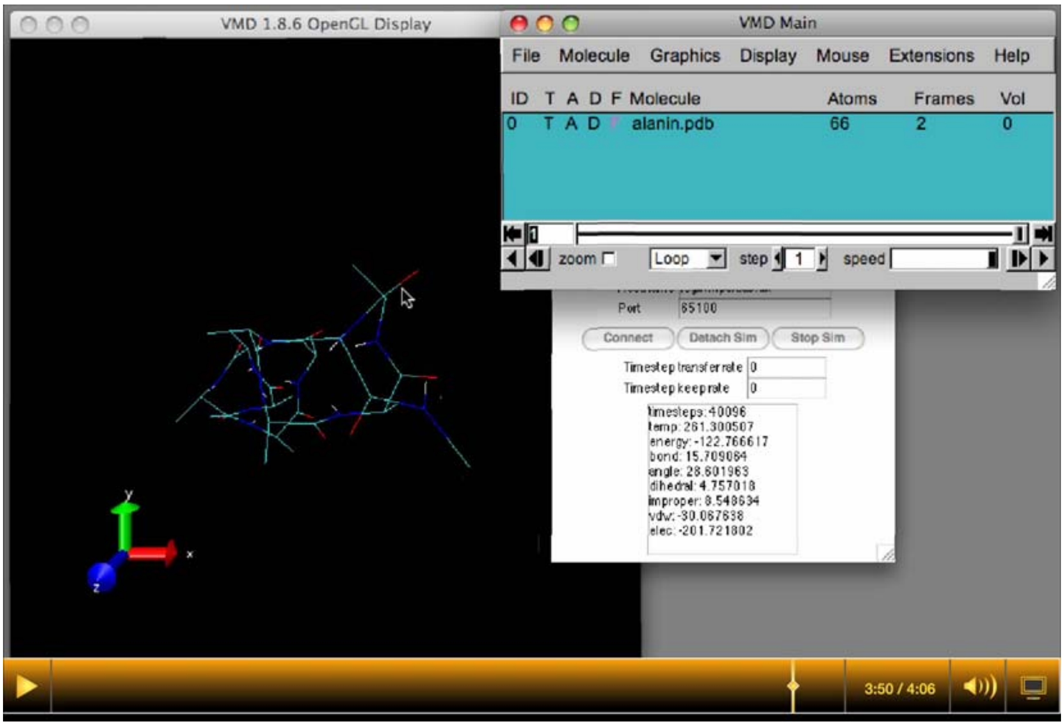
This project involved studying the mechanisms by which proteins control the transport of ions and water molecules across them, a key process for living organisms. Due to the limitations of traditional methods, novel techniques involving real time interaction were utilised involving the computational steering of the simulation.

Iain Bethune visited Dr Carmen Domene at Oxford University on 4th March. This project involved the use of computational steering using the Interactive

Molecular Dynamics (IMD) mechanism. This involves using the Visual Molecular Dynamics graphical package and is used interactively on a local machine to control an instance of the NAMD application running on the back-end of HPCx. Human interaction via this mechanism allows for otherwise impossible manipulation of the molecules.

We enabled IMD on HPCx: we devised a mechanism using the Proxycontrol tool to allow remote connections from a local instance of VMD to NAMD running on the back end of HPCx. We have made this available to all users and documented the procedure in the HPCx User Guide. We also developed a video demonstrating the process (see above figure), which is available on the HPCx web site.

The IMD work involved relatively small resources (where the typical job was 32 or 64 cores for 1-2 hours), to provide starting configurations for larger runs using stand alone back-end NAMD runs that in turn provided results.
NAMD version 2.6, with some metadynamics extensions written directly by the researchers, was initially used: this code had already been compiled on HPCx with help from the support team. The version of NAMD was later upgraded.



Screenshot of IMD Video available on HPCx web site

Dr Domene made extensive use of the HPCx Open Access Initiative, and through this obtained many results. She spent the first part of 2009 analysing and writing up this work, the results of which would influence the specifics of the CCC project. She then went on to perform most of her CCC production runs in the space of a few months in late summer.

Dr Domene expressed concerns that there might be a network performance issue between HPCx and her local institution which could restrict the size of system she could run using IMD. We performed network bandwidth and latency tests between Edinburgh and HPCx, and between Oxford and HPCx. Since the results were generally comparable, we then had confidence that any tests done at Edinburgh would perform similarly to those done at Oxford. We experimented with the computational steering of large systems from Edinburgh, and the results were encouraging: the network coped well with the demand. This gave assurance to the researchers that their intended work was feasible.

An article on this project was included in the HPCx newsletter. This work was also presented at the HPCx Annual Seminar.

2.3.5 e100 Simulations for Synthetic Biology: Mapping Biological Switches

This project involved studying the mechanism by which proteins can switch between states via the kinking of the alpha-helice component, which will in turn allow improved development of synthetic molecules.

Iain Bethune visited Dr Mark Samson and Dr Phil Fowler at Oxford University on 3rd March. The project initially used NAMD 2.6 with TCL scripting, but when it was released we installed NAMD 2.7 in which the TCL interface has been replaced by native routines resulting in improved performance. The project-enabling work was focused into two strands. Firstly, they were simulating a single helix in vacuo and measuring the free energy map. This prototype simulation was run on a local workstation, and the expected behaviour in the results gave confidence in starting simulations of larger systems on HPCx. Secondly, the project was working on integrating with the Application Hosting Environment, to ease the launching and monitoring of jobs.

The project initially required a long simulation to equilibrate the system: the long HPCx queues were utilised for this. The main production runs involved four large batches of jobs, each incorporating a very high number of jobs being submitted concurrently. This job profile did not fit well with our existing queuing setup, which restricted the number of concurrent jobs to three, so we introduced a new aspect of flexibility to the service: we relaxed the concurrency limit for this project while retaining it for others. We also advised on the best job profile (in terms of processor number and job time) to maximise throughput.

The user successfully got through the workload, including the submission of several hundreds of jobs at one time. This was new territory for HPCx, since it resulted in queues containing more jobs than previously experienced on the service, but all went smoothly. The project then scaled up to produce new larger simulation unit cells.

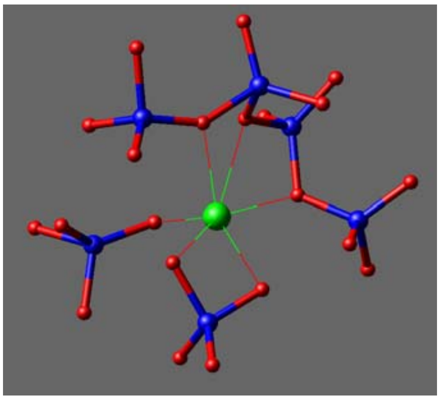
This work was presented at the HPCx Annual Seminar.

2.3.6 e101 Modelling Ion Migration in Bioactive Glasses

This project aimed to improve the understanding of the structure of bioactive glasses, and in turn enable improved application of these materials in biomedicine as restorative and regenerative implants.

We offered to visit this consortia, but they got back to us by stating that they felt remote communication was sufficient for their needs at this stage.

The Quantum Espresso code suite (including the CPMD module) was utilised for this work, which was found to run smoothly on HPCx with good scalability.



Octahedral coordination shell of a sodium ion (green), extracted from a configuration of the CPMD run of 45S5 Bioglass®.

The researchers got underway quickly with production runs, and took advantage of the flexible policies on HPCx which allow them significantly longer parallel jobs than previously possible. In particular they took advantage of the long 48-hour queues on HPCx, which have helped the work proceed quicker than expected (in real time). Midway through the project it transpired that the requested AU allocation was an underestimate of that required for the work, so the researchers requested and were awarded additional compute time.

A full article on this project is available in the final edition of the HPCx newsletter.

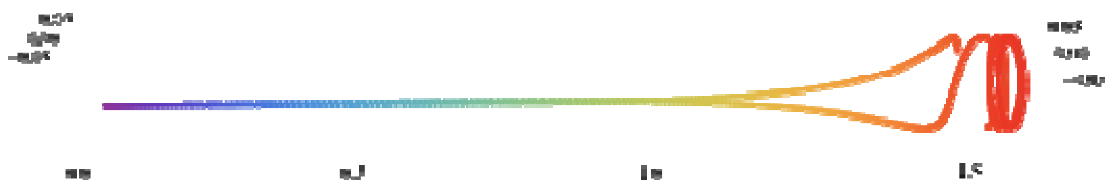
2.3.7 e102 Simulations of antihydrogen formation

This project involved performing studies into the formation of anti-hydrogen, necessary to design ways of cooling and trapping such particles experimentally. Trapped antihydrogen will allow high precision tests of fundamental principles such as the CPT invariance of relativistic quantum theory and the equivalence principle of general relativity. The interaction of antihydrogen with atoms and molecules is a novel area of physics and chemistry, and such interactions may also occur as part of the trapping process, either deliberately or due to impurities in the apparatus.

We offered to visit this consortium at the outset, but they stated that they felt remote communication was sufficient for their needs at that stage. The codes utilised were developed by the researchers themselves (and colleagues) using C++ and MPI, and were already running on HPCx.

We held a meeting with this consortia during the CCP2 workshop 'Atoms and Antimatter' at UCL in June. Dr Svante Jonsell reported that the project was running well. At this stage they had finished performing a phase of intense computations, and were going through the results before re-commencing production runs. We discussed profiling and use of the coretrace tool (very occasionally he suffered from a seemingly random crash which he thought might be due to an accidental divide-by-zero on an unfortunate trajectory; these errors do not easily reproduce, however). The code basically performs many simultaneous Runge-Kutta Newton's-laws-of motion coupled-equations solutions. We also discussed the possibility of adding a dump-and-restart option so he could make use of the low-priority queues in addition to the normal runs.

This project continued to make good progress. We continued to advise on compiler option optimisation, and suggested performing some few-trajectory test runs using gprof, or Totalview, to check for the most time-consuming routines. We obtained a copy of the codes for study purposes. However as noted above the very occasional crash error was not easily reproducible as a few-trajectory test. In the meantime the group generated enough further data to submit articles for publication.



Part of the trajectory of an antiproton, showing helical motion inside the positron plasma (right) and reflection by the electric field at the end of the Penning trap (left).

An article on this project was included in the HPCx newsletter. This work was also presented at the HPCx Annual Seminar.

2.3.8 e103 Towards understanding the mechanism of fast proton transport in biological systems

This project aimed to further the understanding of fast proton transport between hydrogen-bond atoms, which is an important fundamental biological and chemical mechanism. Such an understanding is necessary for technological developments in areas such as hydrogen fuel cells.

Alan Gray visited Dr Carole Morrison at the University of Edinburgh on 21st January. The project wished to use the CP2K ab initio molecular dynamics package for production runs. Immediate support was required to enable this software: we installed CP2K as a centrally available package on the system. This work involved overcoming technical difficulties occurring due to specifics of the HPCx system environment. After this work was complete, the project was able to perform production runs. They took advantage of the long 48-hour queues available on HPCx.

As part of the project progression, the researchers attempted to implement a part of the code which includes dispersion forces, but this failed. We investigated and discovered that the version of CP2K on HPCx did not include this functionality. The enabling of a suitable version of CP2K was not straightforward due to incompatibilities with the current HPCx compilers. We overcame this problem by equipping a test machine with a suitable software environment, building CP2K, and copying the executable across to HPCx. This solved the problem and the consortia were subsequently able to perform the dispersion force calculations.

We consulted with this consortium over the development of a web-portlet interface for submitting CP2K jobs, and provided to them the newly developed web submission functionality, as described later.

This project made very good progress. In particular the simulations have allowed determination of a number of keys factors needed to establish fast proton transport in an alpha-helical domain. This work has been written up for publication.

We requested an article from this consortium for the HPCx newsletter describing their work. They initially agreed, but understandably had to withdraw when it became apparent that the HPCx newsletter publication timescale may cause problems with copyright and as such reduce the chance of the work being accepted for peer-reviewed publication.

2.3.9 HPCx Annual Seminar

The seventh and final Annual Seminar was held at Daresbury Laboratory in December 2009. This included presentations from HPCx users and staff which:

- looked back on research achievements enabled by HPCx;
- described current work exploiting the flexibility offered under the complementarity initiative;
- looked forward to future HPC services.

Details of the event are available at:
http://www.hpcx.ac.uk/about/events/annual2009

2.3.10 Capability Computing Newsletter

The ultimate issue of Capability Computing was published in 2009.

This was a fitting tribute to the HPCx service and included:

- some words presented from those responsible for delivering the HPCx service: Arthur Trew (Service Director) and Caroline Isaac (IBM HPC UK executive);
- articles from two of the most active projects over the years, in the areas of materials research and environmental modeling describing how HPCx has helped further their research;
- a showcase of the fruits of the Complementarity Initiative.

3 HPCx Full Service Review

3.1 The History of HPCx

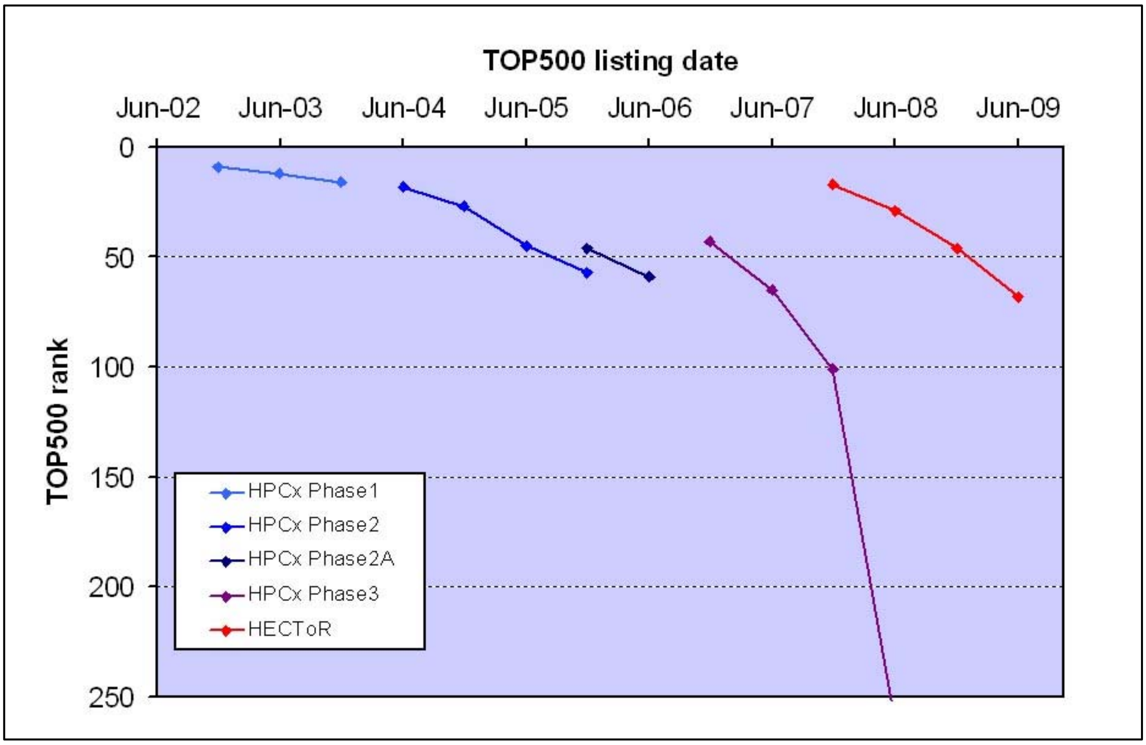
Phase	Phase1 (Dec-02)	Phase2 (Jul-04)	Phase2a (Nov-05)	Phase3/4(Oct-06)
Nodes	40 p690 Frames 1.3 GHz Power4 Processors 1280 Processors	50 p690+ Frames 1.7 GHz Power4+ Processors 1600 Processors	96 eServer 575 nodes 1.5GHz Power5 Processors 1536 Processors	160 eServer 575 nodes 1.5GHz Power5 Processors 2560 Processors
Interconnect	SP switch 2 with Colony PCI Adaptor	HPS Switch	HPS Switch	HPS Switch
Disk	18 TB GPFS	36 TB GPFS	36 TB GPFS	36 TB GPFS
Rmax Linpack	3.4 Tflop/s	6.2 Tflop/s	7.4 Tflop/s	12 Tflop/s

HPCx was the UK's first Terascale resource offering researchers:

- Access to over 1000 processors for single jobs;
- Access to a clustered system of shared-memory nodes (8-way LPARs, 32-way P4 then 16-way P5 nodes).

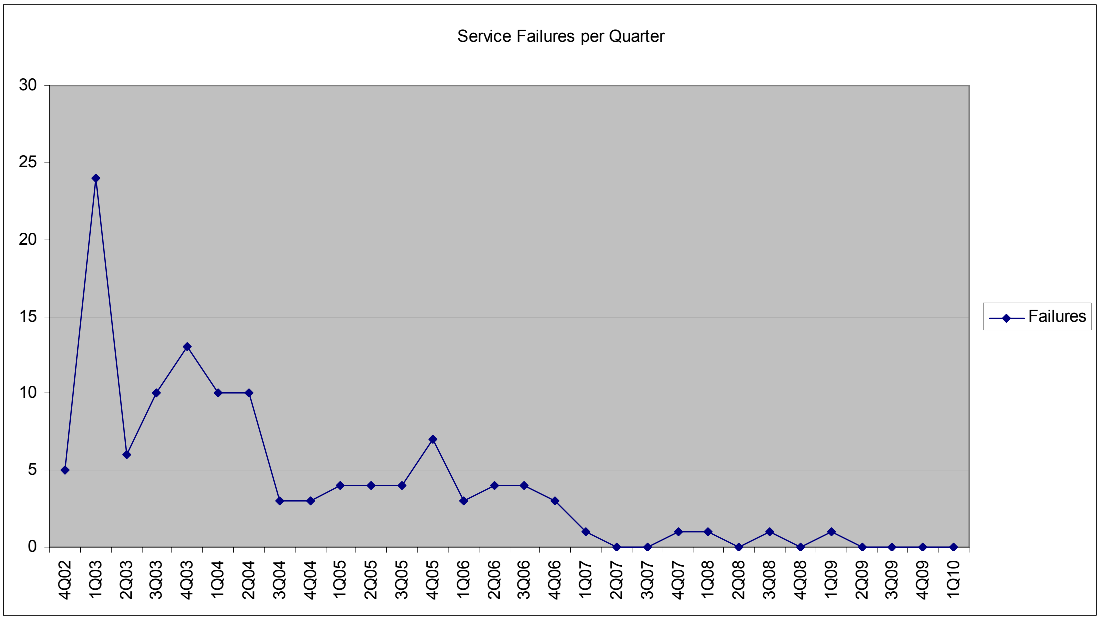
3.1.1 HPCx in the Top500

HPCx Phase1 entered the Top 500 at position 9 in Nov 2002. HPCx remained competitive through performance doubling every two years, and was one of the top 50 systems until 2007.



3.2 Quantitative Metrics

3.2.1 Reliability

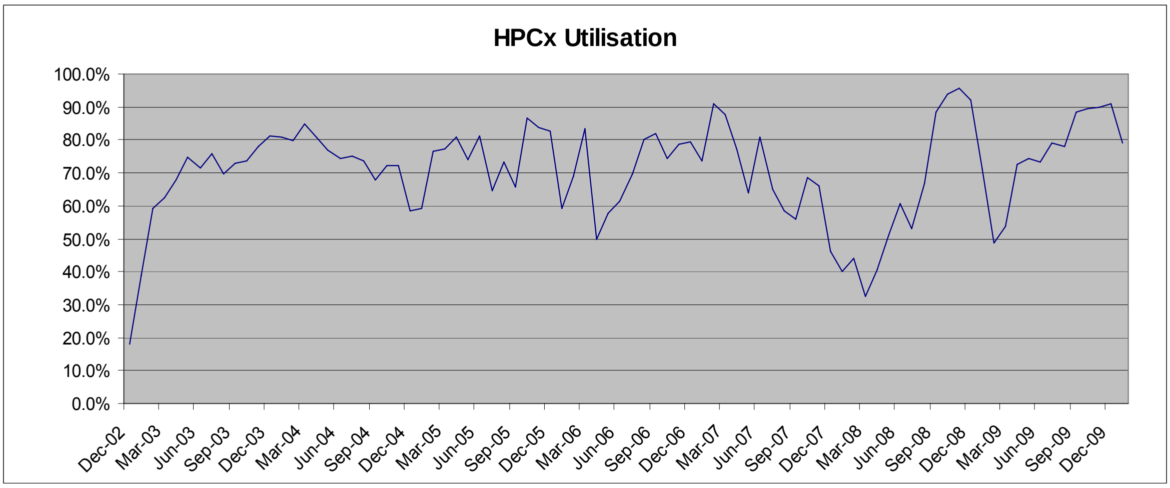


The above graph shows all service failures. Reliability improved considerably from Phase 2 onwards.

Since 2007 there have been no failures of the main service. The contractual failures above in 2007-2010 were due to maintenance overruns, and one failure of the tape archive.

That is a record that is unrivalled at any similarly-sized facility in the world. The results are testament to IBM's build quality and maintenance, and the dedication of the systems support team.

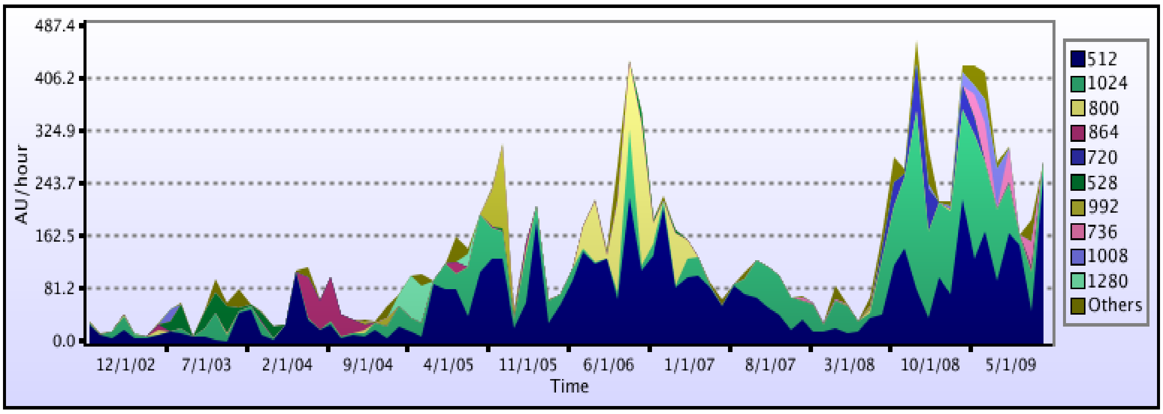
3.2.2 Utilisation



The graph above illustrates utilisation of HPCx through the life of the service. In the period December 2002 through to January 2010, average utilisation on HPCx was 71% (365,050,484 Aus from 845,744 jobs).

HPCx utilisation decreased in line with HECToR coming online in September 2007 but later recovered, reaching an all time high of 96% in November 2008. The Phase 4 initiatives described in Section 2.2 ensured that HPCx continued to be heavily utilised through to the end of service.

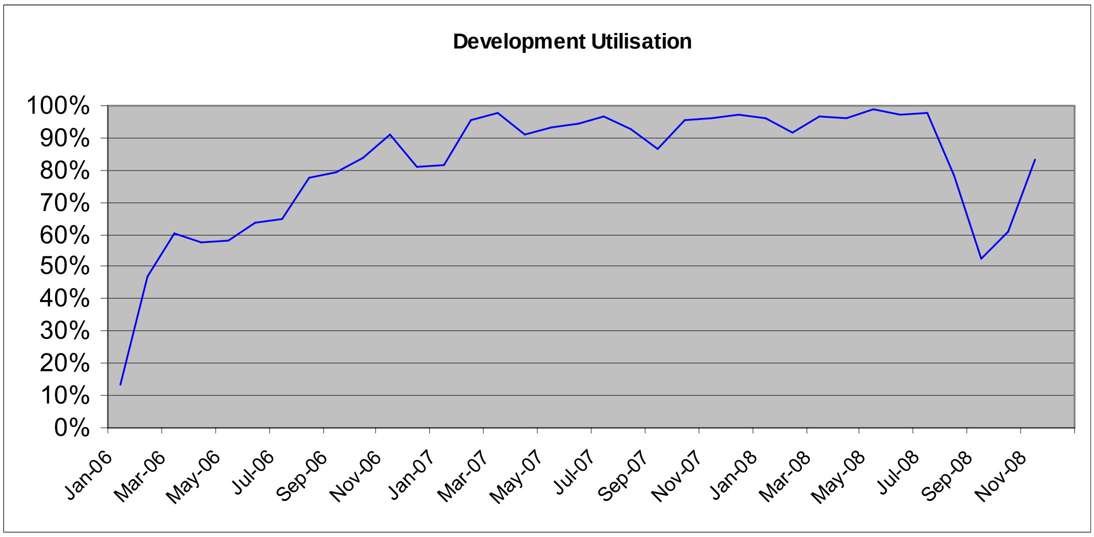
3.2.3 Capability Use



There was a substantial growth in capability use over the lifetime of the HPCx service. The decrease in 2007 coincides with the introduction of the HECToR service. The suspension of charging in August 2008 followed by the introduction of the various Phase 4 initiatives clearly reversed the trend from September 2008.

3.2.4 Development Environment

From January 2006 up until December 2008, a dedicated environment was in place, primarily for the NERC users. This consisted of 12 nodes (192 processors). This enabled NERC to run jobs which did not fit with the overall capability nature of the main machine. This was used for running a high volume of smaller job sizes.



The 12 nodes were added to the main service when NERC left the HPCx service in Phase 4.

3.2.5 Projects on HPCx

In total, 150 projects were setup on HPCx between 2002 and the end of service.

- EPSRC
  - 106 projects (e01 to e106)
- NERC
  - 4 large-scale projects (n01 to n04)
- BBSRC
  - 10 projects (b01 to b10)
- Others
  - 7 Early use projects (y001 to y007)
  - 11 Internal projects (z001 to z11)
  - 1 CCLRC/STFC project (c01)
  - 2 PPARC projects (p01 and p02)
  - 2 Director's projects (d01 and d02)
  - 5 External projects (x01 to x05)

3.2.6 Top Ten Projects Utilisation

Project	Title	PI	Usage Main (M AUs)	Usage Dev (M AUs)	Overall (M AUs)
e05	Materials Chemistry	Richard Catlow	59.3	-	59.3
e01	UK Turbulence Consortium	Gary Coleman	38.5	-	38.5
n02	NCAS	Alan Thorpe	19.4	16.9	36.3
e06	UK Car-Parrinello	Paul Madden	26.6	-	26.6
e99	CCC4	Carmen Domene	14.6	-	14.6
n03	Mineral Physics Consortium	John Brodholt	9.0	2.0	11.0
n01	Ocean Circulation	Thomas Anderson	10.7	0.1	10.8
e24	DEISA	David Henty	10.3	-	10.3
e63	UK AAC	Nick Hills	10.0	-	10.0
e58	Quantum Simulations for Chemical Biology	Carmen Domene	9.0	-	9.0

3.3 Applications Support (Dr David Henty)

The applications support team was responsible for technical reports, training courses, newsletters, workshops and events over the first six years of the HPCx service. From 2009, the focus of the HPCx service switched primarily to targeting complementary capability computing projects so most of these more general activities were not continued. In this section we cover the highlights and lessons learned from these activities over the years.

3.3.1 Training

Training was an important activity – the statistics for the entire service are summarised below.

Metric	Total
Number of courses	83
Number of course-days	178
Student-days training offered	3623
Student-days training delivered	2250
Occupancy	62%

As can be seen, a 'typical' course was around 2 days in length, with 20 places available of which 13 were occupied.

Although the training programme was very popular and successful as a whole, it is difficult to identify particular highlights. However, some important lessons were learned:

- There is an ongoing strong demand for introductory-level courses on fundamentals of high-performance computing and core parallel programming techniques. This audience clearly does not comprise the major national users, who are already experienced in these areas, but it is essential to maintain this training activity to broaden the knowledge base among new researchers. Despite large-scale computational science having been a well established research discipline for some two decades, the required training still appears to be hard to find outwith the national supercomputing centres.

- Advanced training appeals by definition to a much smaller audience. It is difficult to attract sustainable levels of attendance at such specialised courses unless they are combined with other training activities. Wherever possible we have leveraged off EPCC's MSc in HPC, but without this the range of courses we would have been able offer within the effort budget would have been much reduced.

- Travelling to research groups around the country and delivering tailored courses is perhaps the most effective training method, although it requires substantial preparation and effort.

3.3.2 Technical Reports

Some 70 technical reports were produced over the six-year period. A great success of these reports was the part they played in dissemination of the work outside of the UK. A study of two years' web access data from mid-2006 to mid-2008 showed that 1400 reports were downloaded each month, with 50% of these coming from China and the US.

Technical reports were often used as a method of disseminating work done for a particular consortium (e.g. some new performance optimisation or parallel algorithm); this was a useful activity and provided a resource for subsequent reference. However, the alternative approach of using a technical report as a motivation for investigating a completely new area also proved very successful, particularly where this was allied to a planned conference submission. For example, this lead to investigations of new parallel languages and novel object oriented techniques which, although of great interest to the wider HPC community, might not have arisen directly from day-to-day applications support work.

3.3.3 Events

We typically organised five events per year: two user group meetings, two workshops on specific HPC topics and one annual seminar. The annual seminars consistently attracted an average audience of more than 50 users, showcasing the research being done on HPCx and other similar systems. However, the highlight of the events was probably the workshop on *Novel Parallel Programming Languages for HPC*, which took place at the eSI in Edinburgh on June 18th 2008. Initially planned as a small meeting with perhaps 20 attendees, it attracted over 60 registrations and we had to relocate to a larger meeting room. Its success was partly due to its co-location with the HPCx Annual Seminar (held the day before), but mainly due to good timing with the topic of new HPC languages suddenly becoming an active research area after perhaps a decade of relative stagnation.

It is clear that co-location of events is essential: there are so many discipline – specific meetings already crammed into a researcher's year that it is hard to justify travelling to attend a single HPC-focused event. This is particularly true of short user-group meetings, which worked best when held as one of the sessions at the annual seminars. Although the internet should in principle make it possible to hold virtual meetings with a low barrier to participation, technology does not yet appear mature or pervasive enough to make this a practical proposition. Despite concerted attempts, attendance at user group meetings held over Access Grid was always low.

3.3.4 Helpdesk

The query statistics for 2002-2010 are shown below. All query targets were successfully met.

All non-indepth queries	Number	%	Target
Finished within 24 Hours	1501	80.9%	75.0%
Finished within 72 Hours	1827	98.5%	97.0%

Administrative queries	Number	%	Target
Finished within 48 Hours	1108	97.6%	97.0%

Category	Number	% of all
Administrative	1135	49.2%
Technical	720	31.2%
In-depth	355	15.4%
PMR	7	0.3%
Technical Assessment	38	1.6%
Total	2255	100.0%

As above, almost half of the queries received were basic administrative type queries. The functionality provided in the service administration software which enables users to request very basic changes (such as account creation, password resets, quota changes) without having to raise individual queries via the Helpdesk is of great benefit. This element of 'self-service' ensures that there is as little delay in fulfilling such requests. Users can then concentrate on the science as opposed to the administration of their project.

3.3.4.1 Quality Tokens

Users continued to post positive quality tokens through to early 2008. No quality tokens have been received since. The last negative quality token on HPCx was received back in May 2005. Users were clearly very happy with the service.

3.3.4.2 User Communications

In 2008, the Service Administration functionality was updated to enable direct user mailings to all registered users, PIs, or specific project groups. All users have the ability to access these mailings in the SAFE. Previously, mailing lists were manually maintained with mails posted on the HPCx website. The updated software allowed PIs to contact all of their users directly also. This has proven highly useful on both HPCx and now HECToR. The ability to quickly contact all or a subset of users is very beneficial.

3.4 Outreach (Dr Richard Blake)

In the following we report on highlights and lessons learned from the Outreach activities supported by the HPCx service. The effort dedicated to Outreach activities was loaded towards the early years of the service and focussed on a program of support for life-science activities; new applications in particular exploiting international Grid resource; industrial outreach/ public understanding of science; and visualisation. Whilst the HPCx service achieved much with limited resources, future services should develop a more sustained and larger effort in these areas to broaden and embed the engagement beyond the lifetime of individual project grants.

3.4.1 Life Sciences

As part of the added value proposal for the HPCx Service, IBM committed to support a life sciences activity to stimulate the exploitation of high-performance computing in this important and growing area of scientific activity. The advances in the scale and accuracy of simulation capabilities made it timely to explore the role of these methods in addressing questions of interest to biologists. Highlights included:

- An Enzyme Catalysis study which undertook the highest-level QM/MM calculations performed to date for an enzyme, enabling the study of new classes of enzymes of vital biological importance, for example, in drug metabolism.
- A Virtual Outer Membrane study which undertook benchmark simulations of an outer membrane using NAMD, and explored trans-membrane signalling mechanisms for vitamin B_{12} release.
- A Quantum Directed Virtual Evolution study which generated transition metal complexes closely related to the known system and informed future work to search for a new catalyst that will be capable of reducing di-nitrogen to hydrazine.

The projects noted above focussed on large-scale atomic and electronic simulations but is was noteworthy that in the middle to latter stages of the service there was increasing interest from researchers involved in simulations of the physiological and biological response of systems at ‘engineering’ and macroscopic length and time-scales. Highlights of these included:

- The Integrative Biology project - where heart modelling codes were optimised on HPCx to undertake much longer timescale, larger and more realistic simulations of the electro-physiological response of the heart.
- Simulations of the retina based on the electrical properties of a network of cells.
- Cranial blood flow modelling as part of an international Grid project.

The UK projects are now heavily involved in international projects such as the Virtual Physiome funded through the EC. Many of these biological projects require the integration of physical, chemical and biological processes taking place over many length and time-scales. This will require bringing together teams of researchers with very different skills for extended periods of time to tackle the fundamental science and develop the next generation of codes that can run efficiently and effectively on the next generation of hardware. It is noteworthy that the success of the biological modeling activities fed through into a larger commitment of funding to the latter services.

3.4.2 New Applications

The Outreach activity encouraged new science areas to exploit the HPCx service: Highlights included:

- three-dimension seismic wave propagation in collaboration with international researchers in Mexico
- plasma modeling of accelerators and tokomak fusion plasma physics codes
- support for international award winning Grid activities such as the Teragyroid project exploring multi-phase fluids (HPC Challenge award for the most innovative data intensive application at SC 2003), SPICE - Simulated Pore Interactive Computing Environment (HPC Analytics Challenge Award SC 2005) and the GENIUS project to simulate cranial blood flow in a Grid environment.
- Bio-informatics applications including gene-regulatory networks which require computations on large data-sets.

3.4.3 Public/Industrial Awareness

The HPCx Service promoted itself through a number of mechanisms to its various stakeholders. The scientific and technical staff gave presentations at numerous national and international workshops and conferences and worked with the user community to secure the publication of special editions of various learned journals. The Capability Computing Newsletter regularly presented highlights. The service was successful in securing talks at the Edinburgh International Science Festival and STFC runs an active schools program. One notable success was that the BBC featured HPCx and the Computer Room in a series on climate change on Sunday the 21st of September at 9pm. The title was “Earth: The Climate Wars”, and it was the third and final episode in the series. Public Understanding of computational Science and the role of HPC is clearly a big job and the community needs to come together as a whole to raise awareness – not just with the public but also to maintain and enhance support with policy makers.

In terms of industrial outreach the service engaged in workshops and one-on-one visits with industrial end-users and commercial software vendors. We ported and benchmarked codes in the oil exploration, engineering and pharmaceutical areas and worked with a number of commercial software vendors to explore scalability of their codes. Within these projects we were more successful in attracting staff effort to support research and development on the projects than in securing funding for access to the system per se. Many of the codes used commercially do not scale, though the software vendor licenses scale linearly with the number of processors; cluster systems are much more cost-effective than capability systems for low-processor-count applications, JANET does not support commercial access and there are industrial concerns on data security on shared resources. There is work still to do here.

3.4.4 Visualisation

STFC invested in a significant shared memory visualisation system to support its internal program of work and the HPCx service. We explored a number of packages on the visualisation system and running on HPCx itself. The system was used by the engineering, environmental and plasma modelling Consortia to pre- and post-process datasets. As we move to future systems the demands for a visualisation capability to process the ever growing data sets is likely to become ever more important.

3.5 Terascaling Applications (Dr Mike Ashworth)

The HPCx Terascaling Team has throughout the seven-year lifetime of the service provided support for a huge number and range of codes; too many to list here. Full details of support for individual codes and users are given in the Quarterly and Annual Reports. In addition the Team installed a wide range of software (application codes, libraries and debuggers and profiling tools) and investigated their performance and capabilities with feedback in many cases to the developers and to users through training courses, workshops and technical reports. Above all the Team provided a flexible and responsive science support service, with low management and reporting overhead. As well as a program of work which was defined and refined annually in the HPCx Annual Plan, the Team handled a wide range of requests for in-depth support from the Help Desk. No reasonable user request was refused.

The Team provided support for the Capability Incentive Scheme, benchmarking codes and in many cases putting forward codes as a result of optimisation work. The totals for the Scheme are as follows:

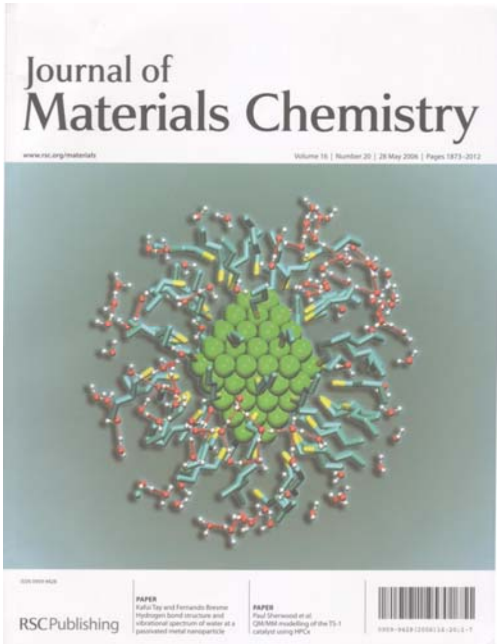
- 18 Gold standard codes (scaling from 512 to 1024 processors)
- 4 Silver (256 to 512)
- 5 Bronze (128 to 256)

The Team also produced many of the 71 Technical Reports available from the HPCx website and were co-authors on many journal papers. This, together with presentations at Consortium meetings, HPCx workshops and Annual Seminars, and International conferences: e.g. Parallel CFD, IBM ScicomP, Cray User Group, resulted in a very high national and international impact for the Team’s work.

The following sections are selected highlights of a vastly greater body of work.

3.5.1 Journal Special Issues

Work of the HPCx Terascaling Team was also featured in special issues of two journals. In May 2006 a special issue of the Journal of Materials Chemistry focused on Materials Modelling using HPC.



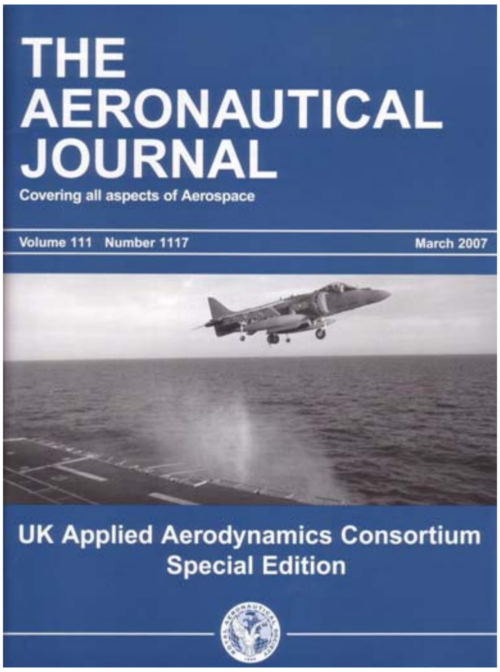


Figure 1: HPCx Terascaling work was featured in two special issues of the Journal of Materials Chemistry and of the Aeronautical Journal

Six out of the fifteen papers were co-authored by HPCx staff. In 2007 a special issue of the Aeronautical Journal featured work of the UK Applied Aerodynamics Consortium. There was a leading article by Dave Emerson, Mike Ashworth, Andy Sunderland (HPCx staff) and Ken Badcock (PI of the Consortium). All except one of the seven technical papers were either co-authored by or contain acknowledgements to HPCx staff.

3.5.2 CASTEP through the Ages of HPCx

The plane-wave materials science code CASTEP (http://www.castep.org/) is developed in the UK and licensed by over 150 UK academic research groups and scientists, with many more worldwide via sales by the commercial partner Accelrys Inc. Plane-wave methods can be used to compute almost any chemical, optical, mechanical or electronic property of the system. Applications include prediction of new crystal structures, interpretation of EELS, Raman and NMR spectra, catalyst discovery, corrosion and crack propagation. CASTEP has been used on HPCx by the UKCP consortium, the

Materials Chemistry Consortium and others. References [1] and [2] list recent scientific publications using CASTEP (on all platforms).

The successive releases of the Fortran 90/95 version of CASTEP [3] have been available on HPCx throughout the lifetime of the system. We have contributed significant effort into parallel performance improvement, particularly from 2003 to 2007. This development was fed back into the main code repository, in collaboration with the CASTEP Developers' Group, for use on other platforms. The transition to HECToR was managed particularly smoothly, with performance development work building on HPCx experience and expertise.

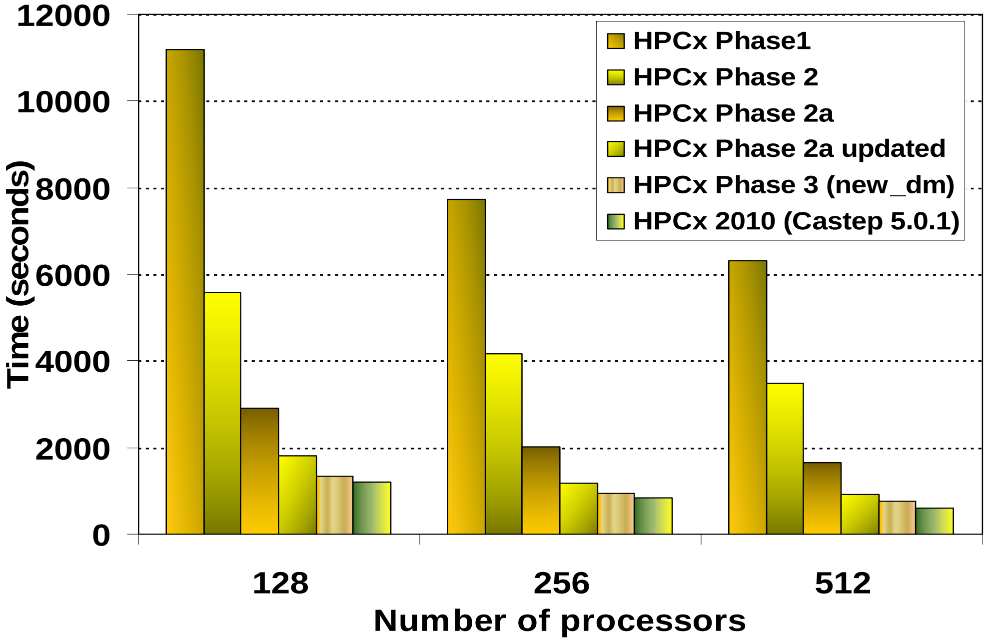


Figure 2: The performance of CASTEP for a 270-atom AL_2O_3 benchmark case for different hardware phases and successive code optimisations

CASTEP's 3D-FFT (Fast Fourier Transform) requires the use of MPI_AllToAllV in the transform. Terascaling effort concentrated on improving the latency-bound MPI_AllToAllV by taking account of the SMP architecture of HPCx and splitting the call into a gather operation, followed by MPI_AllToAllV between sub-group leaders and scatter operations within the sub-groups. These sub-groups utilize fast communication within an SMP and their size is chosen dynamically to maximize performance, taking account of inter-SMP communication speed, which for typical CASTEP grid sizes is highest close to the limit of fast-buffered communication.

Figure 2 shows the performance of CASTEP for a 270-atom AL_2O_3 benchmark case for different hardware phases and successive code optimisations. The performance improves by a factor of ten on 512-processors from Phase 1 to Phase 3. The hardware upgrades contribute only a factor of 2.0 to this (Linpack performance per processor) – the remainder is due mainly to the optimisation efforts of the CASTEP Developers Group and the HPCx Terascaling Team.

This work and other improvements are described in HPCx Technical Reports [4, 5] and a peer reviewed paper [6] with updated performance figures presented at the 2009 HPCx Annual Seminar. The method has transferred to modern multi-core systems such as HECToR, and recent developments on HECToR have incorporated the use of shared-memory segments (developed on HPCx [7]) to increase the performance gain [8]. Work has continued through two HECToR dCSE projects to introduce a new level of 'band-parallelism' [9, 10] and implementation of a TD-DFT method.

[1] http://accelrys.com/products/materials-studio/publication-references/castep-references/
[2] http://portellen.phycmt.dur.ac.uk/sjc/Castep_Publications/publications.html
[3] SJ Clark, MD Segall, CJ Pickard, PJ Hasnip, MJ Probert, K Refson and MC Payne, Zeit Für Kryst 220 (2004) 567
[4] M Plummer and K Refson (2004) http://www.hpcx.ac.uk/research/hpc/technical_reports/HPCxTR0401.pdf
[5] M Plummer and K Refson (2005) http://www.hpcx.ac.uk/research/hpc/technical_reports/HPCxTR0507.pdf
[6] M Plummer, J Hein, MF Guest, KJ d-Mellow, IJ Bush, K Refson, GJ Pringle, L Smith and A Trew, J Mat Chem 16 (2006) 1885
[7] I J Bush (2007) http://www.hpcx.ac.uk/research/hpc/technical_reports/HPCxTR0701.pdf
[8] A Maniopoulou and C Armstrong (2009) http://www.hector.ac.uk/cse/reports/castep_m.pdf
[9] PJ Hasnip (2008) http://www.hector.ac.uk/cse/distributedcse/reports/castep/castep_performance_xt.pdf
[10] PJ Hasnip, MIJ Probert, K Refson, M Plummer and M Ashworth, Cray User Group Proceedings (2009), http://www.cug.org

3.5.3 High-Resolution Coastal Ocean Modelling with POLCOMS

The Proudman Oceanographic Laboratory Coastal Ocean Modelling System (POLCOMS) has been developed to tackle multi-disciplinary studies in coastal/shelf environments. The central core is a sophisticated 3-dimensional hydrodynamic model that provides realistic physical forcing to interact with, and transport, environmental parameters. Integrating from ocean to coast or vice versa, biological production and the fate of contaminants can be determined. The standard workhorse model has been a 12 km resolution grid covering the whole of the north-west European shelf. The capability resources of HPCx allowed POL scientists, in collaboration with the Terascaling Team, to develop a 1.8 km model of the same area.

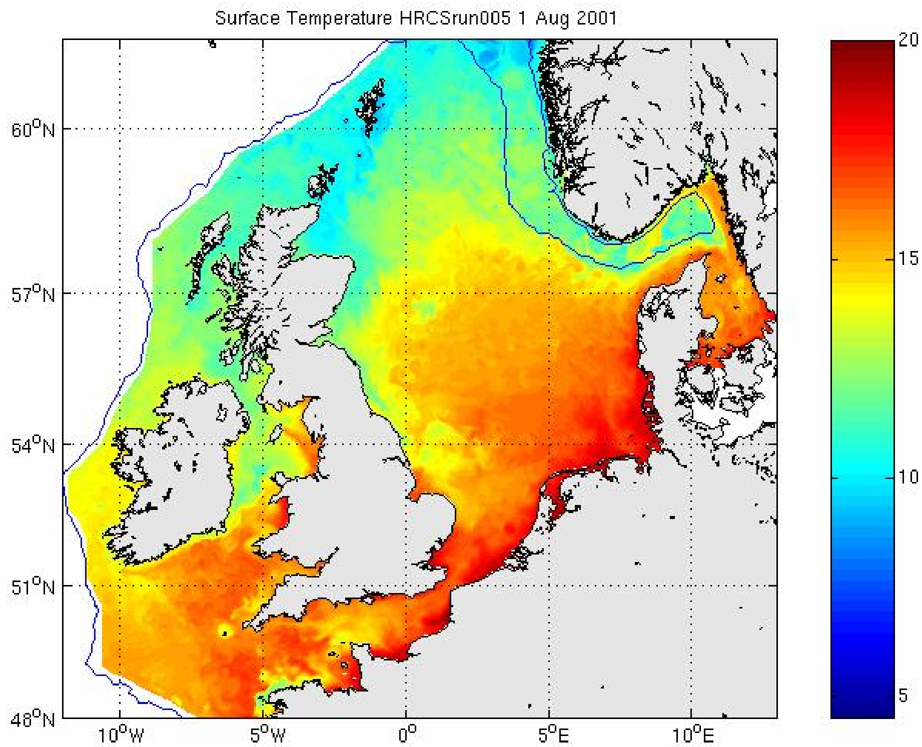


Figure 3: Sea surface temperatures from a 1.8 km simulation of the north-west European shelf using POLCOMS

Results shown in Figure 3 reveal unprecedented detail in the resolution of fronts and of eddies, which are key features to improved understanding of the circulation and its mixing processes.

Work by the Terascaling Team also included coupling of the physical ocean code to the ERSEM marine ecosystem model for studies of marine productivity, coupling with the WAM wave model, inclusion in the code of an ensemble capability and interfacing to data assimilation software using ensemble Kalman filtering. POLCOMS was also coupled with the CICE sea-ice code for use on HPCx in a project funded through the NERC RAPID programme.

Developments in POLCOMS were sustained through a NERC e-Science project on Global Coastal Ocean Modelling (GCOMS) and the techniques learned on HPCx are now being transferred to the NEMO ocean code through the GSUM project (EPSRC Software Development funding) and follow-on work.

3.5.4 Helicopter Rotor Wake Simulation

The helicopter rotor wake simulation code ROTORMGMGP developed by Chris Allen, University of Bristol, under the UKAAC consortium was benchmarked, analysed using the Vampir performance analysis tool and optimised for HPCx. The high level of parallel performance of the code, along with the large memory resources available on HPCx allowed researchers to model 3D unsteady flow blade-vortex interactions in much finer detail than was previously allowed.

Figure 4 shows simulations of vortices using grids with 8 million and 32 million points respectively. The superior resolution of the flow in the 32 million point case is evident. Rotor simulations of this size were only made possible at that time by the availability of HPCx computing resources. In recognition of the code's outstanding scalability performance, ROTORMGMGP was awarded a Gold Award under the HPCx Capability Incentive Scheme. The scientific outcomes and HPC aspects of the project were presented at an invited speaker presentation and a technical presentation and paper at the International Conference Parallel CFD '05.

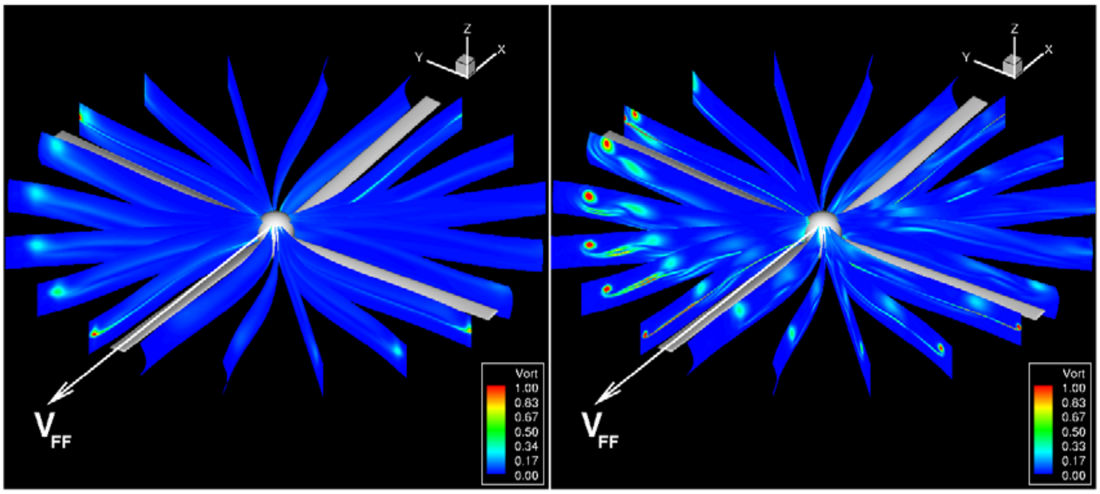


Figure 4: Vorticity shading showing local rotation of fluid for two grid sizes: 8 million points (left) and 32 million points (right)

Parallel performance of a UKAAC helicopter code on HPCx and other large-scale facilities", A.G. Sunderland, D.R. Emerson, and C.B. Allen, Proceedings of Parallel CFD '05, 24th -27th May 2005, Maryland
This paper was also issued as HPCx Technical Report HPCxTR0512, 2005
http://www.hpcx.ac.uk/research/hpc/technical_reports/HPCxTR0512.pdf

3.5.5 Diagonalizing Huge Matrices in PDVR3DRZ

The PDVR3DRZ code is produced by the TAMPA Group led by Jonathan Tennyson and James Munro at UCL and performs high accuracy rotation-vibration nuclear motion calculations on three-atom molecules. The electron motion calculations are performed separately and form an input to the program in the form of a potential energy surface. On HPCx the research involved performing high accuracy, near dissociation calculations on H_3^+ in order to match experimental results. These calculations have helped to improve understanding of the behaviour of simple molecules at dissociative energies; a problem that has been under investigation for over 20 years. In collaboration with the HPCx Terascaling Team, optimisations to the code for HPCx included

i) switching to the ScaLAPACK Divide-and-Conquer eigensolver routines; and
ii) designing a wrapper code to use multiple BLACS grids and multiple MPI communicators within a single parallel job.

These developments increased both the scalability and the performance of the code markedly and PDVR3DRZ was awarded a Gold Award under the HPCx Capability Incentive Scheme. At the time of this research, the optimised PDVR3DRZ code was able to achieve some of the largest diagonalizations of dense real symmetric matrices ever computed For example, a 3D Hamiltonian matrix of order 143,541 was diagonalized in 10,259 seconds on 1280 processors of HPCx.

3.5.6 Molecular Dynamics with DL_POLY

A new version of the DL_POLY package, DL_POLY_3, was developed at Daresbury through CCP5 in collaboration with the eMinerals project. DL_POLY_3 is radically different from DL_POLY_2 in its design and uses a domain decomposition approach to parallelisation and memory distribution. The HPCx Terascaling team provided optimisation and scaling support for DL_POLY_3, which led to a great increase in the number of processors that could be exploited and the time-scales and length-scales of the physical systems that could be simulated. From DL_POLY_2 simulations of around 30,000 atoms on 100 processors, researchers were able to conduct DL_POLY_3 simulations of order 10,000,000 atoms on 10,000 processors, thereby opening a vast new range of feasible studies. This was highlighted by carrying out a successful proof of concept simulation of a NaCl crystal containing 300 million atoms (visible by the naked eye) with full electrostatics evaluation on 1024 processors of HPCx Phase 2.

DL_POLY_3 was first exploited by the eMinerals project at Cambridge in a series of large scale studies of radiation damage in ceramics, an important investigation of these materials as media for immobilising radioactive elements for long term storage [1]. The new code has also found application in the biological area, with novel studies of superoxide dismutases, enzymes that are implicated in motor neurone disease [2]. Meanwhile the capability of DL_POLY_2 was enhanced to include new and valuable methodologies. In collaboration with Sheffield University, hyperdynamics was introduced, which extended the effective time scale of molecular dynamics simulation of rare events by many orders of magnitude. This was proved by applications to studies of surface diffusion in ceramics [3].

The high level of support and continual development of the DL_POLY packages coupled with a large number of projects set to exploit their new accessibility often led to the codes being amongst the five most used codes on HPCx.

[1] Atomistic simulations of resistance to amorphization by radiation damage, K. Trachenko, M.T. Dove, E. Artacho, I.T. Todorov and W. Smith, Phys. Rev. B, 73 (2006) 174207

[2] Structural investigation of human wild-type Cu, Zn superoxide dismutase – a biomolecular case study using DL_POLY_3, C.W. Yong, W. Smith, R.W. Strange and S.S. Hasnain, Molecular Simulation 32 (2006) 963

[3] Surface diffusion and surface growth in nanofilms of mixed rocksalt oxides, D.J. Harris, T.S. Farrow, J.H. Harding, M.Y. Lavrentiev, N.L. Allan, W. Smith and J. Purton, Phys. Chem. Chem. Phys. 7 (2005) 1839

3.5.7 Understanding plasma turbulence in fusion reactors

We worked closely with researchers from the Culham Science Centre, and performed in-depth technical improvements to their nuclear fusion applications to enable more effective utilisation the large-scale resources available at HPCx. ITER[1] is the world's largest machine dedicated to research into the production of energy from controlled thermonuclear fusion. It is designed to demonstrate that safe, clean electricity can be produced economically. In ITER and similar machines plasma must be confined for several seconds in a carefully designed 'magnetic bottle' well away from material walls. The group's effort is devoted to the problem of understanding electromagnetic turbulence in plasmas, and its undesirable effects on enhancing energy losses (called 'anomalous transport') of the plasma far above those due to collisional processes. Global tokamak turbulence calculations present truly 'grand challenges' to the most powerful computers in the world.

We performed in-depth optimisations to the CENTORI and CADENCE codes. We made serial improvements which alone had significant impact. We redeveloped the codes with new parallelisation strategies, introducing multi-dimensional decompositions. This was extremely effective in improving the performance on HPCx, particularly due to the fact that it made the Fourier Transform sections (a key bottleneck) much more efficient. We also worked with the consortia to implement other performance improvements using

[1] International Thermonuclear Experimental Reactor, http://www.iter.org

techniques such as algorithmic substitutions and the introduction of high performing numerical libraries.

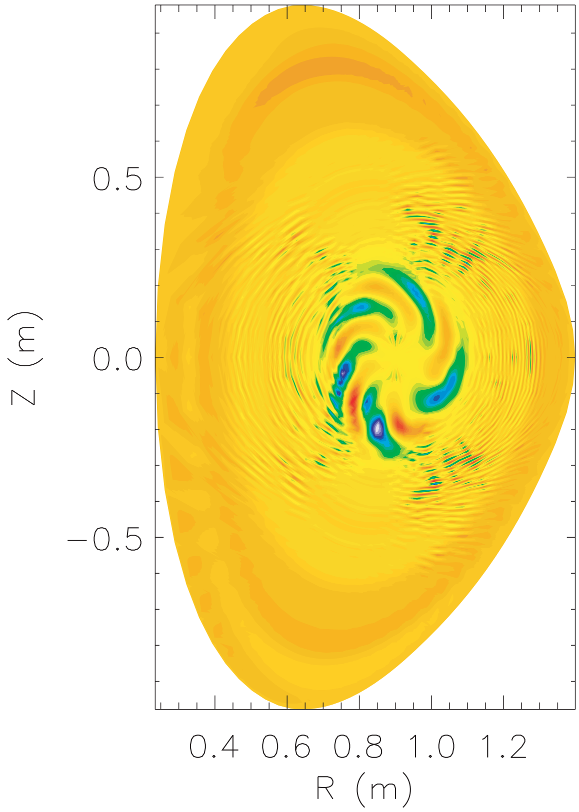


Figure 5. A snapshot of the time-varying fluctuations in the particle density, exhibiting various vortices and other characteristics typical of plasma turbulence.

A range of promising results were been obtained on HPCx. For example, non-linear simulations on HPCx have calculated the saturated state of electron temperature gradient (ETG) driven drift wave turbulence in MAST[2] plasmas. These calculations predict a level of electron heat transport which is, remarkably, comparable to that measured. This result is of considerable interest as ETG turbulence has previously been dismissed as unimportant.

The consortium was successful under the Complementary Capability Challenge call during the HPCx extension, as reported elsewhere in this report. The collaboration between HPCx staff and Culham was expanded, through a jointly-supervised PhD student funded through a CASE studentship. We have also recently been awarded a HECToR dCSE grant to improve GS2, another code used by Culham on HECToR.

Microstability physics as illuminated in the spherical tokamak, C M Roach, D J Applegate, J W Connor, S C Cowley, W D Dorland, R J Hastie, N Joiner, S Saarelma, A A Schekochihin, R J Akers, C Brickley, A R Field, M Valovic and the MAST Team, Plasma Physics and Controlled Fusion 47, B323, 2005

Electron temperature gradient driven transport in a MAST H-mode plasma, N. Joiner, D Applegate, S C Cowley, W Dorland and C M Roach, Plasma Physics and Controlled Fusion 48, 685, 2006

[2] Mega Amp Spherical Tokamak, http://fusion.org.uk/mast/index.html

Arithmetising plasma turbulence: the 'final frontier' of classical physics, P J Knight, C M Roach, A Thyagaraja, D J Applegate and N Joiner, HPCx Capability Computing Issue 6, Autumn 2005 pp 4-6

3.5.8 Modelling heart attacks with supercomputers

One of the goals of the Integrative Biology project is to improve the understanding of the operation of the human heart, and in particular the causes of and treatments for cardiovascular disease, which kills approximately one third of the UK population. Although the incidence of cardiovascular disease continues to fall, it remains an important cause of death in the industrialised world. In many cases, the lethal event is likely to have been a catastrophic arrhythmia called ventricular fibrillation (VF).

Despite its immense social and economic impact, the mechanisms underlying formation of arrhythmias (arrhythmogenesis) are still debated and remedial therapies are, at best, suboptimal. In the case of life-threatening arrhythmias like VF, electrical defibrillation, the application of an electric shock within minutes of VF onset, is currently recognized as the only effective therapeutic option. Despite the critical role that defibrillation therapy plays in saving human life, our understanding of the mechanisms by which electric shocks halt life-threatening arrhythmias remains incomplete. Realistic computational models of the heart are currently the only viable approach to allowing us to understand all relevant parameters at the required spatial and temporal resolutions.

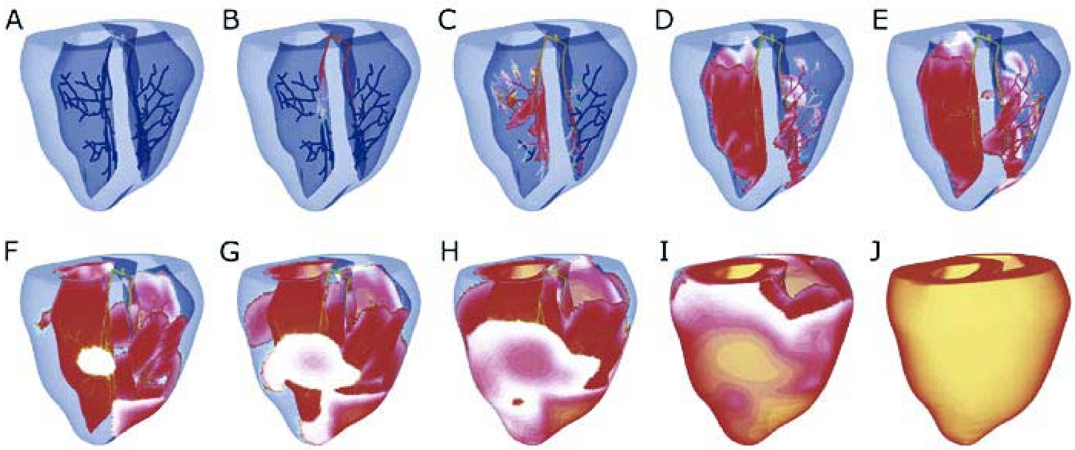


Figure 6. Normal Activation Sequence of the Ventricles.

We successfully ported the Cardiac Arrhythmia Research Package (CARP) and the Sheffield Cardiac Arrhythmia Model (SCAM) to HPCx, and performed a range of optimisations such that the codes could more effectively take advantage of the large-scale resources. This enabled a preliminary study on the feasibility of carrying out a 'virtual experiment' using HPCx. The results suggested that realistic simulations of a human heart including the torso are feasible on HPC platforms. This work is now being taken forward on HECToR, including a collaborative dCSE project between EPCC and the researchers to further improve the scaling of CARP. Such work is required to achieve the aim of using HPC resources for heart modelling to aid real clinical situations.

This project provided an article for the HPCx newsletter, and a presentation at the 2006 HPCx Annual Seminar.

Integrative biology: modelling heart attacks with supercomputers, G Plank, R Clayton, D Boyd and E Vigmond, Issue 7, HPCx Capability Computing, Summer 2006, pp 4-9

Modelling heart attacks and therapeutically treatment regimes with supercomputers, G Plank, HPCx Annual Seminar, 4th October 2006

3.5.9 Summary

The Terascaling Team delivered throughout the course of the seven-year lifespan of HPCx a focused program of work providing support for optimisation and scaling of codes in a wide range of scientific areas in order to improve the efficient and effective use of HPCx and to enable new and world-leading science.

The structure of this support was set out each year in the Annual Plan in consultation with the major user groups. We were very fortunate in having the advice and support of EPSRC and of the HPCx Science and Technology Advisory Committee (STAC) in guiding this process. This planning process allowed us to take a long-term approach in key subject areas and dovetail well with the code development activities taking place in the Collaborative Computational Projects and elsewhere.

In addition to the in-depth plan we retained the flexibility to respond to requests from users for in-depth support through the Helpdesk. Later on in the service we also carried out a series of visits to key groups and Consortia to canvas views on their requirements for HPC in general. This pro-active approach generated a number of additional Terascaling activities.

When scaling applications, any performance improvements are usually used not simply to save resources but to perform larger and more detailed runs resulting in better quality, more realistic simulations and science with a greater impact. This is the true benefit of Terascaling support and the value of this is difficult to quantify. As HPC systems move from multi-core to many-core, it becomes more and more critical that investment in skilled terascaling resources is sustained.

3.6 Software Engineering (Dr Stephen Booth)

The role of the Software Engineering team in the HPCx service was to provide software support for the service and for users of the service with a scope wider than that of a single application.

3.6.1 Machine Investigations

In order to make efficient use of any HPC system it is necessary to have a detailed knowledge of the characteristics and capabilities of the hardware.

As the HPCx service ran on several generations of IBM Power processor and several generations of interconnect we performed a series of investigations into the performance characteristics of each generation of hardware and disseminated the results as technical reports for the HPCx user community

3.6.2 Grid Computing

A significant part of the responsibilities of the software engineering team was to support grid computing on HPCx. The term Grid computing covers a number of novel network-centric technologies. On HPCx support for Grid Computing included:

- Installation and support of the Globus-2 grid middleware.
- Installation and support of the UNICORE grid middleware
- Development and support of the grid-proxy port-forwarding tool to support computation steering of jobs running on the back-end nodes.
- Installation and support of the DANTE network proxy to support tightly integrated meta-computing where a single application is distributed across multiple geographically distributed systems.
- Installation and support of the HARC co-scheduling system to allow meta-computing

3.6.2.1 Lessons Learned

On HPCx users made very little use of the grid job submission interfaces. These aim to provide common interfaces for accessing different compute resources. As the HPCx service provided unique capabilities that other sites on the grid did not, users appeared to find little additional benefit in being able to treat it in an interchangeable fashion with other systems. For future services, it should be noted that only a subset of ‘grid’ criteria may be appropriate as far as a national service is concerned. For example, having a means of common data exchange such as bbFTP is key; however the implementation of a single job submission interface may not be relevant. Such a requirement is clearly not user driven and does not enhance scientific output.

On a European scale, all DEISA users on HPCx used the DEISA interface to submit jobs. This again is a policy decision and again does not appear to be user driven. If a user is granted resources on a facility at a single remote site, having to use such an interface as opposed to a simple login does not appear to provide any scientific benefit.

The ability to make network connections directly to and from running application codes and a greater ability to control when these jobs were scheduled to run was important to a small group of users. Special efforts need to be made if these requirements are to be supported. Such a service may be classed as an 'on-demand' service as opposed to coming under the traditional capacity or capability banner. There will be a small subset of users who require such a service.

3.6.3 Data Management

Data management can be a significant challenge for the use of HPC systems. During the HPCx service we performed a number of investigations into data handling techniques including:

- Efficient strategies for parallel file I/O (including the use of MPI-IO);
- Use of the HPCx archive system;
- Mechanisms for transferring large data-sets across the JANET network.

3.6.3.1 Lessons Learned

Though the 'normal' tools used for data transfer such as *scp* and *sftp* are relatively inefficient mechanisms for the transfer of large amounts of data, alternative tools such as *grid-ftp* and *bbFTP* exist that are capable of utilising a significant fraction of the available network bandwidth. For HPCx the majority of data was stored in a tape archive tightly integrated with the rest of the service. In future it might be possible to move to a less tightly-coupled model where remote data archives are accessed through the network.

3.6.4 Novel Techniques

We explored a number of novel techniques for the development of HPC software. These included:

- Mixed mode OpenMP/MPI programming;
- Novel language approaches (such as PGAS languages);
- Object orientation and performance optimisation;
- Object orientation and parallel design.

3.6.4.1 Lessons Learned

Many of these techniques show great potential for improving the efficiency of HPC software development and the performance of the resultant code.

3.6.5 In-depth Software Support

The software engineering team was responsible for the development and support of the SAFE service administration application. This system handled the majority of administration functions for the HPCx service including: user registration, accounting, resource management, reporting and helpdesk. This has been highly successful throughout the HPCx service not only supporting the service itself but also providing tools to help project mangers manage the resources and people within their own projects.

We also developed a number of alternative implementations of MPI collective operations exploiting the clustered-SMP architecture of HPCx. These could be substituted for the default IBM implementations using the MPI profiling interface. Optimisations included:

- Optimising reductions and barriers using shared memory segments;
- Optimising MPI_alltoallv using message combining.

3.6.5.1 Lessons Learned

The SAFE service administration software has been a key success. From a service perspective the Helpdesk functionality works very well. The SAFE has provided the ability to report quickly and accurately on key metrics such as utilisation and reliability. Without the flexibility in the software, supporting the HPCx initiatives such as Open Access, and Capability Incentives would have been much more complicated.

From a user perspective it has provided Principle Investigators with the means to simply manage their own resources, and to extract key reporting data as projects progress.

The accounting principles behind the SAFE are now being considered as a benefit to HPC users on a European scale, with the DEISA project now investigating the principles of the accounting system.

3.7 Systems Support (Mr Mike Brown)

3.7.1 Multi-Site Support

Systems support for HPCx was provided by operations groups at two geographically-divided sites – Daresbury Laboratory and Edinburgh. In order for this to succeed a high level of co-operation and enthusiasm was required from day one.

This excellent level of co-operation started during the bidding process, and was developed through the service planning stage, the complex refurbishment of the Computer Room, and the delivery and build-up of the system. The two groups were rapidly able to form an integrated team with a wide background skill-set, and were thus able to give the level of professional support planned.

The value of this co-operation goes beyond the HPCx project since the joint Edinburgh and Daresbury group forms the core of the Operations and Systems Group that has gone on to support the HECToR service.

Of similar value was the early close co-operation with IBM that started during the service planning stage. IBM rapidly deployed staff to be based at Daresbury almost from the time that the contracts were signed, and mention must be made of the massive contribution made to the service from Terry Davis who devoted seemingly all of his time to this service until his untimely death in 2009.

3.7.2 Technology

The HPCx service was upgraded every two years, with almost a doubling of capacity at each stage. Not only was this essential to remain competitive, but it was also good for staff motivation and interest.

3.7.3 Relationships with other HPC Sites

Close relationships were established with other large scale IBM HPC sites, initially in the UK, and we were able to enjoy very useful co-operation from sites such as ECMWF which were deploying similar technology.

This model of establishing close contact with like-minded sites continues in the HECToR service and is of immense value to the individuals that take part and the services that they represent. The challenges faced by such services are not unique, and there is much to learn from working with both UK and global HPC sites. Best practices can be shared to the benefit of the user community.

3.7.4 Reliability

The fact that the service continued to provide an effective, reliable and popular service after more than seven years was extremely satisfying. A lot of investment went into developing the service, and the longer lifetime of the service was able to benefit from that.

A very major highlight was the significant development in service reliability after Phase 1. After a challenging first year, the service bedded down to become what must surely have been the most stable and reliable large-scale multi-user HPC platform in the world.

3.7.5 Lessons Learned

- The initial lack of a test and development system was a very significant deficiency in the provided technology. Such a system is necessary to enable software and systems to be tested out in (reasonably realistic) circumstances before being deployed out onto the service platform so as to mitigate risk. Modern HPC systems are highly complex and it became clear that it was essential to be able, as far as is reasonable, to mimic the operational service hardware and software configuration.

- Whilst setting up a new (multi-site) team from scratch was a challenging task, this was made even harder by there being a low level of technology training provided by IBM. This was not part of the technology specification in the procurement, but it meant that the service had to deploy additional resources at the project outset.

- Effective on-site support from the equipment manufacturer is an essential requirement for services of this scale if the service is to be properly supported. IBM were able to base appropriate staff for, in effect, the lifetime of the service, but the overall reliability of the service would have been adversely affected if we had not had access to the regular, skilled support staff from IBM who were implicitly familiar with the service and the physical site.