# HECToR Annual Report 2010

## 01 January – 31 December 2010

## Issue: 1.0

# Contents

# 1 Introduction

This report covers the period from 1 Jan 2010 at 08:00 to 1 Jan 2011 at 08:00.

The next section of this report contains an executive summary for the year.

Section 3 summarises service availability, performance and utilisation statistics for the year. Section 3 also covers the Helpdesk statistics. Systems support is covered in Section 4, with the work of the Cray Centre of Excellence described in Section 5 and the Computational Science and Engineering (CSE) Support provided by NAG covered in Section 6.

The Appendices define some of the terminology and incident severity levels and list the current HECToR projects together with their overall utilisation profile to date.

This report and the additional SAFE report are available to view online at http://www.hector.ac.uk/about-us/reports/annual/2010.php

# 2 Executive Summary

2010 was a very challenging period for the HECToR Service with a number of major system changes and periods of disruption for users. The highlights of the service over the year included:

- Average utilisation in 2010 on the Cray XT4 was 73%, compared to 59% in 2009. The X2 was very reliable in 2010. Charging remained suspended and the overall utilisation for the year was 44%.
- In 2010 there were 43 technology-attributed service failures compared to 37 in 2009. In addition to the technology failures, there was one site failure and one external failure. The overall MTBF was 195 in 2010, exactly the same as it was in 2009. The key difference in 2010 was the nature of the service failures and their duration.
- The Phase 2b upgrade ran very smoothly. The hardware was delivered and installed as planned, and acceptance tests were completed in early June. A 20-cabinet Cray XT6 system was installed; the system came online in June. As part of the upgrade, the Phase2a system was downsized from 60 to 33 cabinets in May.
- The transition to the external shared lustre filesystem went ahead in October. This process had previously been aborted in April and again in August due to major difficulties. The Phase2a system was taken out of service for a 10 day period in October to allow for dedicated testing prior to finally going ahead with the migration.
- The upgrade of the Phase 2b system to Gemini was originally planned for January 2011. In response to user concerns, the work was brought forward to December. The Phase 2b system was unavailable to users from the 6$^{th}$ to 17$^{th}$ December to accommodate the upgrade to the new interconnect. The Phase 2b system formally becomes the contractual service as of 1 Feb 2011.
- Utilisation on the Phase 2b system averaged 35% for the 6 month period July to December. Utilisation was initially very low but it increased dramatically in December with the introduction of the Gemini interconnect, and the associated suspension of charging to accommodate the acceptance testing and availability trials.

- A total of 3166 queries were handled in 2010 and the associated Helpdesk statistics for the year were excellent.
- The introduction of the Cray XT6 and its subsequent upgrade to a Cray XE6 in December required substantial effort from all the service support teams, both behind the scenes and in the user facing roles.
- HECToR training continued to be popular. At the end of December the CSE team were just short of having 1000 attendees on courses since the start of the service. The CoE was also engaged in user training and workshops, promoting the HECToR service throughout the UK, Europe and the rest of the world.
- To date over 47 person-years of effort have been allocated through the dCSE mechanism and the benefits of this work are being felt throughout the user communities of the codes that have been worked on, as well as by users of HECToR. The dCSE panel is having to be increasingly selective in where they target resources.
- 2010 saw the establishment of the Cray Exascale Research Initiative Europe which is a research collaboration between Cray, EPCC and CSCS. The focus of the Initiative will be research into technologies, both software and hardware, which are required to get to Exaflop scale systems.
- A review meeting was held with key HECToR user representatives on December 12th. A number of options for HECToR Phase 3 were presented and discussed. The options for Phase 3 will be further reviewed on 1st February by Cray, EPCC and EPSRC.

# 3 Quantitative Metrics

## 3.1 Reliability

The quarterly numbers of incidents and failures (SEV 1 incidents) are shown in the table below:

| | *1Q10* | *2Q10* | *3Q10* | *4Q10* | *2010* |
|---|---|---|---|---|---|
| Incidents | 84 | 92 | 65 | 40 | 340 |
| Failures | 14 | 16 | 7 | 8 | 45 |

The incidents above are primarily related to single node failures. Details on both the service failures and single node fails in 2010 can be found in Section 5.

### 3.1.1 Performance Statistics

- MTBF = (732)/(number of failures in a month)

  Quarterly MTBF = (3x732)/( number of failures in a quarter)

  Annual MTBF = (12*732)/(number of failures in a year)

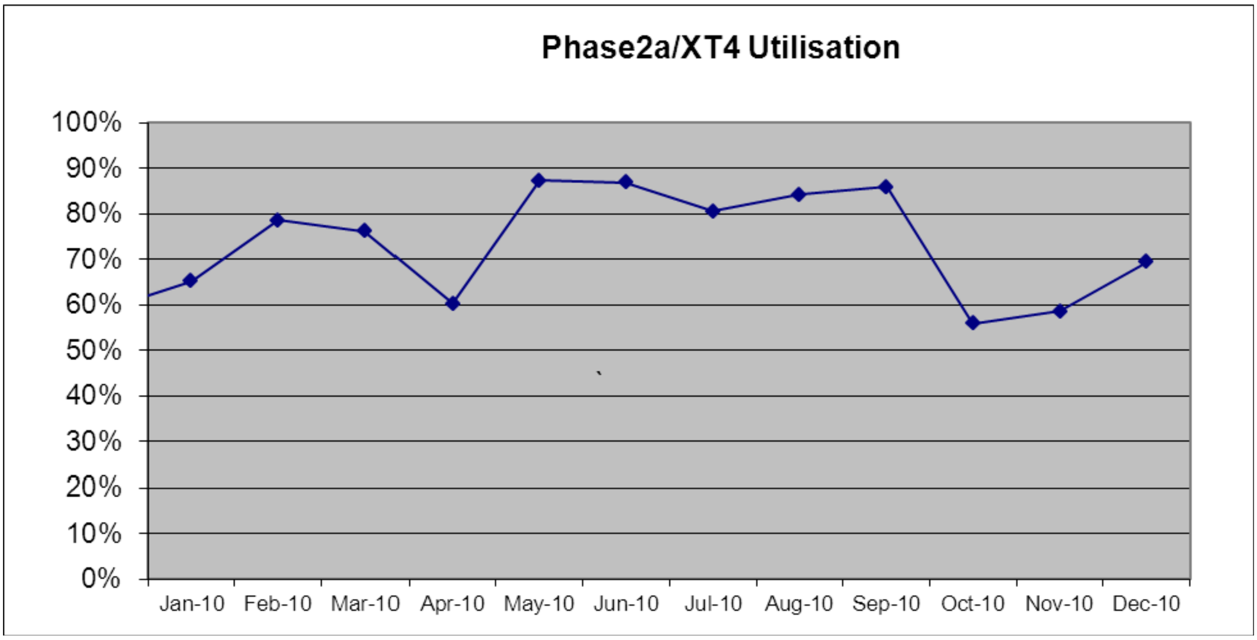
| *Attribution* | *Metric* | *1Q10* | *2Q10* | *3Q10* | *4Q10* | *2010* |
|---|---|---|---|---|---|---|
| Technology | Failures | 12 | 16 | 7 | 8 | 43 |
| | MTBF | 183 | 137 | 314 | 275 | 204 |
| Service Provision | Failures | 1 | 0 | 0 | 0 | 1 |
| | MTBF | 2196 | ∞ | ∞ | ∞ | 8784 |
| External | Failures | 1 | 0 | 0 | 0 | 1 |
| | MTBF | 2196 | ∞ | ∞ | ∞ | 8784 |
| **Overall** | **Failures** | **14** | **16** | **7** | **8** | **45** |
| | **MTBF** | **157** | **137** | **314** | **275** | **195** |

Overall system failures and node failures decreased by a factor of around two in the second half of 2010. This coincides with the downsizing of the Cray XT4 from 60 cabinets to 33 in June. Failure rates on a per cabinet basis were therefor relatively constant.

## 3.2 HECToR Utilisation

### 3.2.1 Phase 2a Utilisation

The Phase2a (Cray XT4) system was reduced in capacity from 60 to 33 cabinets in June 2010.




Overall utilisation in 2010 was 73%, compared to 59% in 2009.

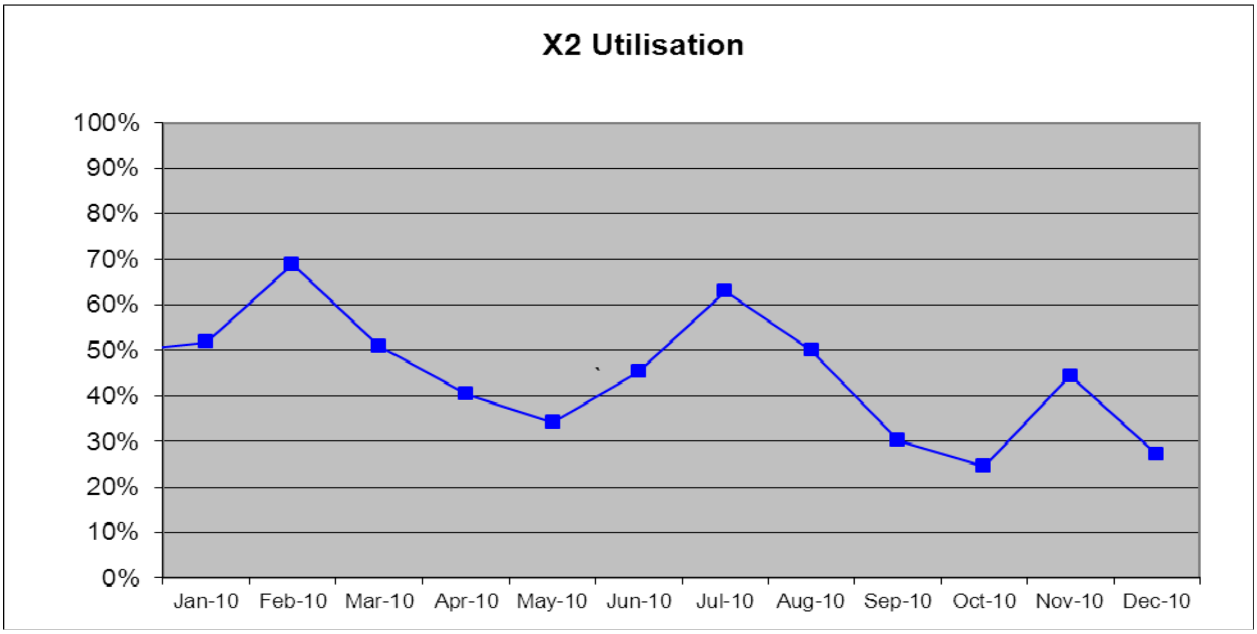
In 1Q10, the results of the first EPSRC 'Direct Access' Call were announced and a new class of project 'Class 1b' was introduced on HECToR. Several new projects have since started in 2010 as a result of the initiative. Further details are available in Section 3.2.6 of this report.

Utilisation was impacted in April due to stability problems on the service at that time. There were a number of service failures relating to the installation of the esFS file system.

Utilisation on Phase 2a recovered in May and remained high through to September despite the introduction of the new Cray XT6 system in June. There was a 14 day period in May when charging was suspended in an attempt to compensate for the service disruption in April.

In October, the Phase 2a system was unavailable for a period of 10-days to enable Cray to perform dedicated testing on esFS. The migration of user data to esFS ran during October and November. During this period many users elected not to use the Phase 2a system whilst their data was being migrated as recommended by the service. This reduced the risk of any data corruption whilst the migrations were in process. There were no such issues and the service are very grateful for the co-operation and understanding of the users.

### 3.2.2 X2 Utilisation




Only a very small subset of HECToR users use the Cray X2 system as can be seen in Section 3.2.5. These groups were all contacted in 3Q10 regarding their future use of the system.

- **UK Turbulence Consortium**

The main use of the Cray X2 is from the UK Turbulence Consortium. There were two main areas where they found the system was useful for their research:

Available Memory Per Core

The group can run a simulation on a single computational core with up to 32Gb of memory. This was important for the group in order to generate files for 3D visualisations and/or in order to generate statistic in time and in space from simulations

Direct Numerical Simulations of Turbulent Flow

The group used the vector component of the HECToR service for two sets of direct numerical simulations of turbulent flow.

The first of these were of turbulent Couette-Poiseuille flows. The simulation code employed for these simulations (SWT) uses an FFT library designed for vector processors, and a one-dimensional domain decomposition that limits its scalability. This means that it performs relatively well on the Cray X2, and also that, barring significant development work, it is not possible to run challenging simulations on a multicore machine with limited memory per core.

The second set of simulations performed was of arrays of actuator disks immersed in a turbulent Ekman layer. The code used for these simulations was a version of the EBL code that has recently been the subject of a dCSE project to improve its scalability. Prior to the completion of this project (and the additional development work required to re-implement the actuator disk code in the new version) it employed a one-dimensional domain decomposition and thus benefited strongly from fast processors with access to large amounts of memory. It will now be possible to run future simulations efficiently on conventional multicore HPC architectures.

- **Other Groups**

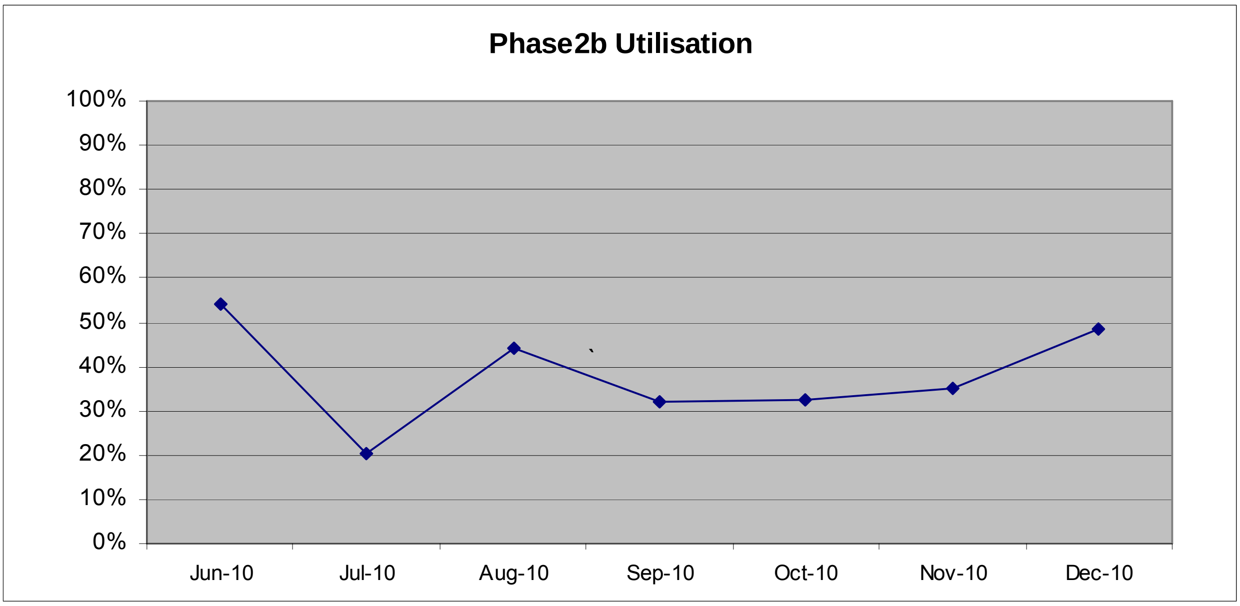
Other groups commented that the low usage on the Cray X2 and the suspended accounting were appealing. They were able to achieve very fast turnaround times. Some users had seen better performance on the Cray X2 than on the Cray XT4 for their code, however this had not been compared against the Cray XE6 as it was not available at the time of asking.

- **Access to Vector Resources**

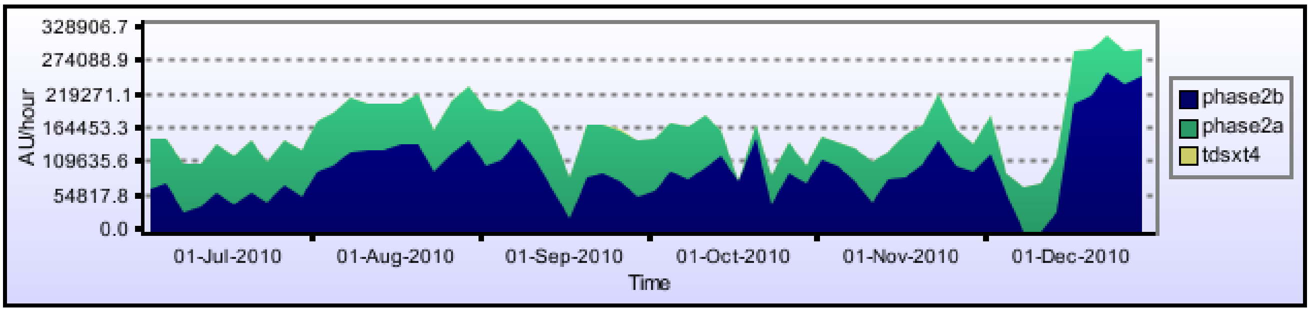
A vector resource is available within the DEISA infrastructure (an NEC SX-9 at HLRS) should UK researchers have a requirement for such a compute resource. Details are available at www.deisa.eu.

### 3.2.3 XT6 Utilisation

The Phase 2b system came online in June. Utilisation averaged 35% for the 6 month period July to December.




Despite the percentage utilisation being lower on the Phase 2b system than on Phase 2a, far more AUs were used.




Utilisation on Phase 2b was lower than we would like, despite the introduction of a highly discounted AU rate in September . There may be a number of reasons for this:

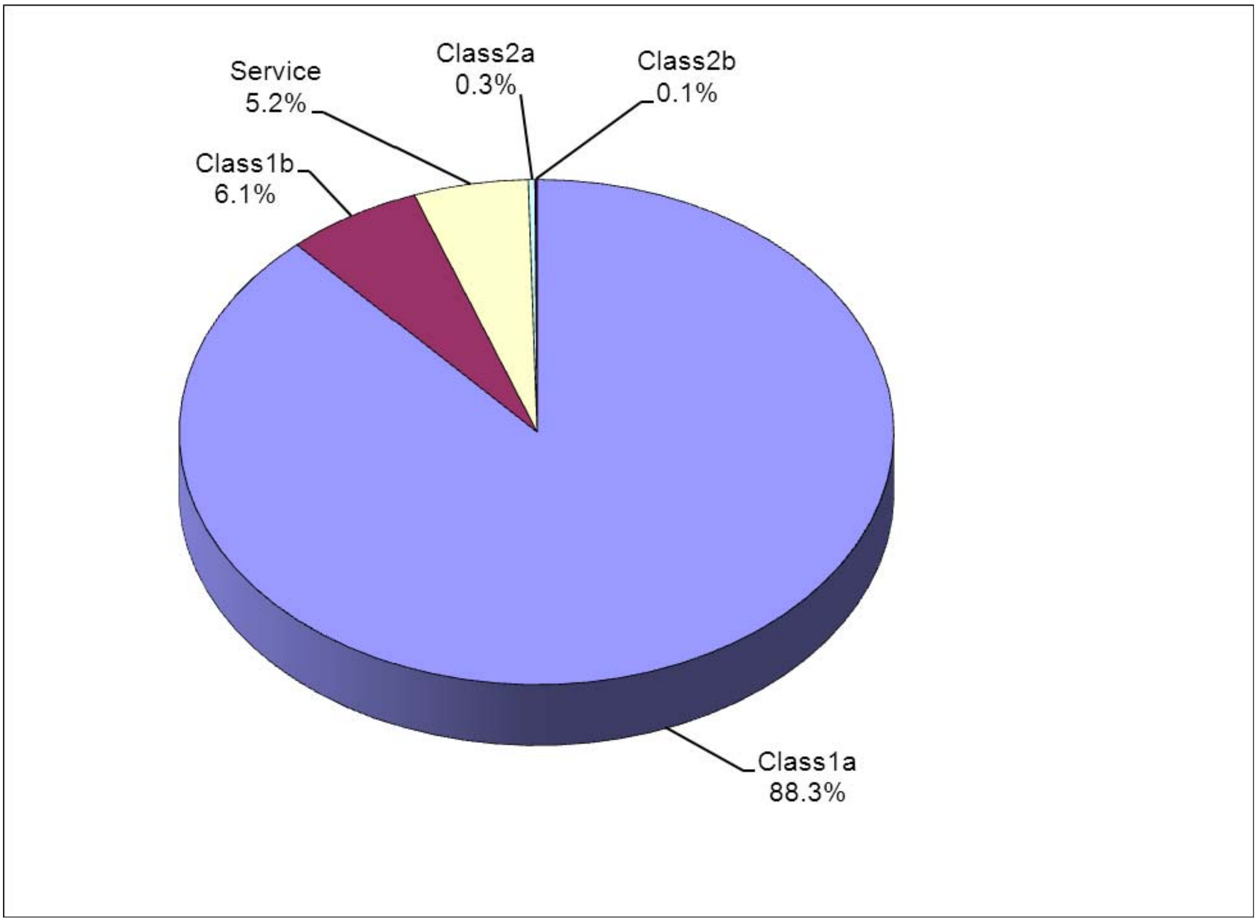
- A number of codes did not perform well on the original Seastar interconnect
- The above performance issue made the Cray XT6 a costly solution for many groups, despite the 40% reduction in the AU rate
- Users waited for the arrival of the shared lustre filesystem in October before using the Cray XT6 as the initial embedded storage capacity was minimal
- The allocations of compute time no longer match the increased capacity of the service. A projection of overall system usage is included in Section 3.2.9.

Utilisation on the Phase 2b system increased dramatically in December with the introduction of the Gemini interconnect, and the associated suspension of charging to accommodate the acceptance testing and availability trials. Charging was re-enabled in January on completion of the 10-day availability trials.

### 3.2.4 HECToR Utilisation by Project Class

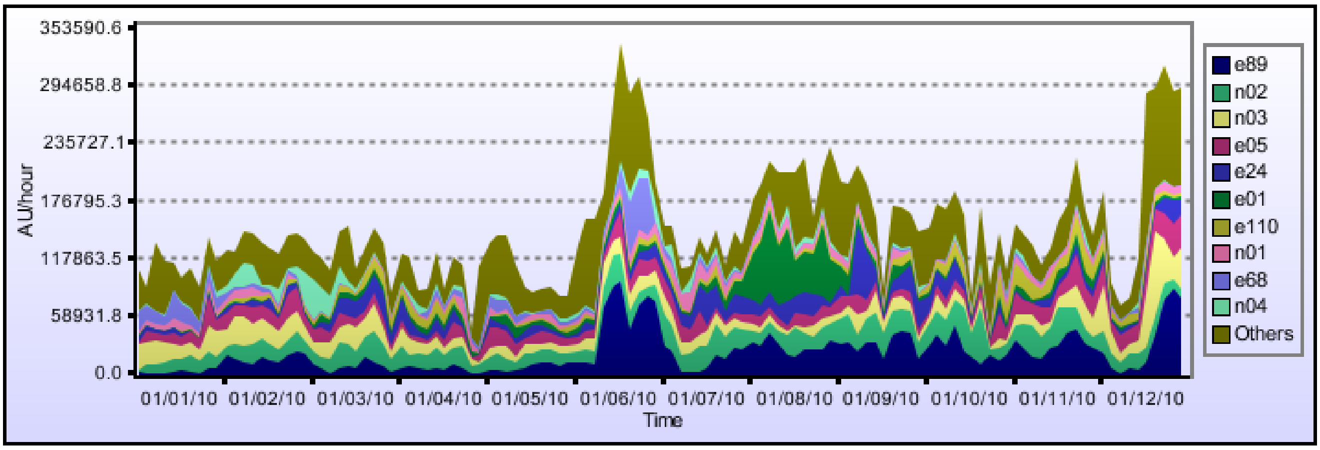
In March 2010 the EPSRC 'Direct Access' mechanism was introduced. A new class of project 'Class 1b' was defined. Direct access provides a modest amount of compute resource, which must be used within a defined period (usually 6 months) or the resource is lost.  A total of 22 direct access projects started in 2010.

As below, direct access accounted for just over 6% of the utilisation on the service during 2010.



### 3.2.5 HECToR Utilisation by Consortium

As below, the main utilisation in 2010 came from the major EPSRC and NERC consortia as one would expect.



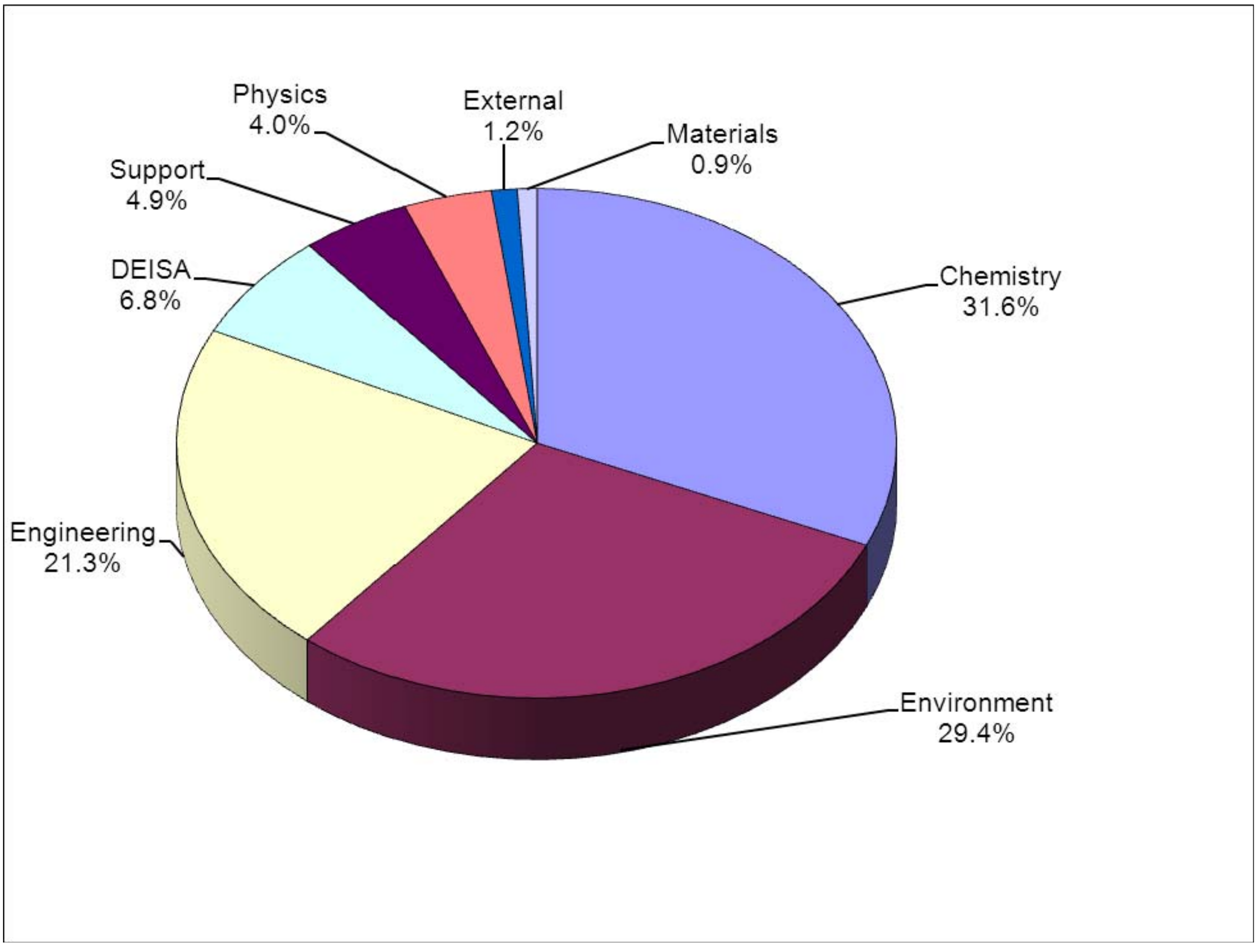

A total of 2,486,315,781 AUs were available during this period. The breakdown across the various systems was a follows:

| Project | phase2a (XT4) | phase2b (XT6/XE6) | phase2a (X2) |
|---|---|---|---|
| y01 | 10.0% | 0.0% | 0.0% |
| y02 | 1.0% | 1.6% | 0.0% |
| y03 | 0.0% | 0.0% | 0.0% |
| y04 | 0.0% | 0.0% | 0.0% |
| y05 | 0.0% | 0.0% | 0.0% |
| y06 | 0.0% | 0.0% | 0.0% |
| y07 | 0.0% | 0.0% | 0.0% |
| y11 | 0.0% | 0.0% | 0.0% |
| z01 | 0.1% | 0.1% | 0.0% |
| z02 | 0.0% | 0.0% | 0.0% |
| z03 | 0.4% | 0.3% | 0.0% |
| z06 | 0.0% | 0.0% | 0.0% |
| Internal Total | 1.6% | 1.9% | 0.0% |
| c01 | 1.3% | 0.3% | 0.0% |
| e01 | 2.2% | 3.6% | 35.0% |
| e05 | 6.8% | 2.6% | 7.0% |
| e10 | 0.1% | 0.0% | 0.0% |
| e19 | 0.0% | 0.0% | 0.0% |
| e24 | 6.4% | 1.6% | 1.4% |
| e35 | 0.3% | 0.0% | 0.0% |
| e42 | 2.1% | 0.0% | 0.0% |
| e59 | 0.0% | 0.0% | 0.0% |
| e63 | 1.3% | 0.2% | 0.0% |
| e68 | 2.6% | 1.1% | 0.0% |
| e70 | 0.1% | 0.1% | 0.0% |
| e71 | 0.2% | 0.0% | 0.0% |
| e72 | 0.0% | 0.0% | 0.0% |
| e75 | 0.0% | 0.0% | 0.6% |
| e76 | 2.0% | 0.0% | 0.0% |

| Project | phase2a (XT4) | phase2b (XT6/XE6) | phase2a (X2) |
|---|---|---|---|
| e81 | 0.0% | 0.0% | 0.0% |
| e82 | 0.1% | 0.4% | 0.0% |
| e84 | 0.3% | 0.0% | 0.0% |
| e85 | 0.2% | 0.8% | 0.0% |
| e89 | 6.5% | 9.4% | 0.1% |
| e90 | 0.0% | 0.0% | 0.0% |
| e92 | 0.0% | 0.3% | 0.0% |
| e93 | 0.0% | 0.0% | 0.0% |
| e101 | 0.0% | 0.0% | 0.0% |
| e102 | 0.3% | 0.0% | 0.0% |
| e104 | 0.2% | 0.0% | 0.0% |
| e105 | 0.0% | 0.0% | 0.0% |
| e107 | 0.1% | 0.0% | 0.0% |
| e108 | 0.0% | 0.0% | 0.0% |
| e110 | 2.7% | 1.6% | 0.0% |
| e113 | 0.0% | 0.0% | 0.0% |
| e117 | 1.4% | 0.0% | 0.0% |
| e120 | 0.0% | 0.0% | 0.0% |
| e121 | 0.9% | 0.0% | 0.0% |
| e122 | 1.3% | 0.4% | 0.0% |
| e124 | 0.5% | 0.4% | 0.0% |
| e125 | 0.5% | 0.6% | 0.0% |
| e126 | 0.4% | 0.0% | 0.0% |
| e127 | 0.0% | 0.0% | 0.0% |
| e129 | 0.1% | 0.0% | 0.0% |
| e131 | 0.0% | 0.0% | 0.0% |
| e132 | 0.0% | 0.0% | 0.0% |
| e134 | 0.0% | 0.0% | 0.0% |
| e135 | 0.0% | 0.0% | 0.0% |
| e136 | 0.4% | 0.1% | 0.0% |
| e137 | 0.0% | 0.0% | 0.0% |
| e138 | 0.0% | 0.0% | 0.0% |
| e139 | 0.0% | 0.0% | 0.0% |
| e141 | 0.0% | 0.4% | 0.0% |
| e142 | 0.0% | 0.0% | 0.0% |
| e144 | 0.0% | 0.0% | 0.0% |
| e145 | 0.0% | 0.0% | 0.0% |
| e146 | 0.0% | 0.0% | 0.0% |
| e147 | 0.0% | 0.0% | 0.0% |
| e148 | 0.0% | 0.1% | 0.0% |
| e149 | 0.3% | 0.1% | 0.0% |
| e150 | 0.1% | 0.0% | 0.0% |
| e151 | 0.1% | 0.1% | 0.0% |
| e152 | 0.0% | 1.2% | 0.0% |
| e153 | 0.2% | 0.0% | 0.0% |
| e154 | 0.2% | 0.0% | 0.0% |
| e156 | 0.0% | 0.0% | 0.0% |
| e157 | 0.0% | 0.0% | 0.0% |
| e158 | 0.0% | 0.0% | 0.0% |
| e162 | 0.0% | 0.0% | 0.0% |
| e163 | 0.0% | 0.0% | 0.0% |
| e166 | 0.0% | 0.0% | 0.0% |
| e167 | 0.0% | 1.4% | 0.0% |
| e169 | 0.0% | 0.0% | 0.0% |

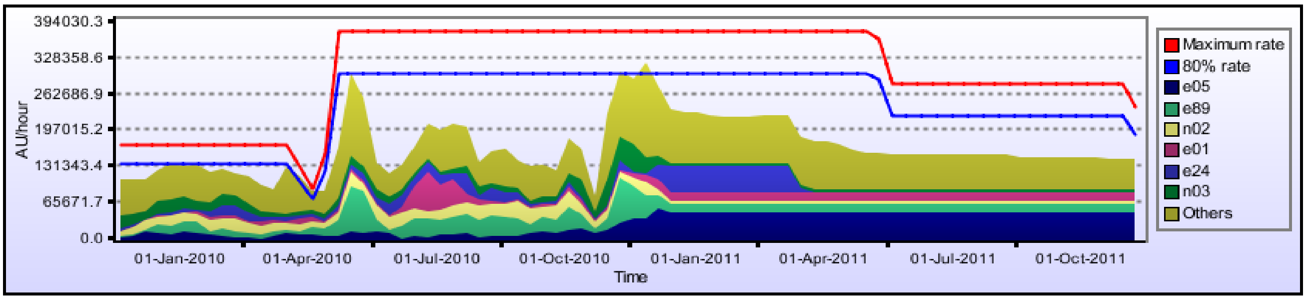
| Project | phase2a (XT4) | phase2b (XT6/XE6) | phase2a (X2) |
|---|---|---|---|
| e170 | 0.0% | 0.0% | 0.0% |
| e171 | 0.1% | 0.0% | 0.0% |
| e172 | 0.0% | 0.0% | 0.0% |
| e173 | 0.0% | 0.0% | 0.0% |
| e174 | 0.0% | 0.0% | 0.0% |
| e177 | 0.0% | 0.0% | 0.0% |
| e178 | 0.0% | 0.0% | 0.0% |
| e186 | 0.0% | 0.0% | 0.0% |
| e188 | 0.0% | 0.3% | 0.0% |
| e189 | 0.0% | 0.0% | 0.0% |
| e190 | 0.1% | 0.0% | 0.0% |
| e191 | 0.0% | 0.0% | 0.0% |
| e192 | 0.0% | 0.1% | 0.0% |
| e193 | 0.0% | 0.0% | 0.0% |
| e195 | 0.0% | 0.0% | 0.0% |
| e199 | 0.0% | 0.0% | 0.0% |
| u03 | 0.0% | 0.0% | 0.0% |
| EPSRC Total | 42.8% | 27.5% | 44.0% |
| n01 | 2.7% | 1.3% | 0.0% |
| n02 | 12.1% | 2.3% | 0.0% |
| n03 | 9.1% | 2.3% | 0.0% |
| n04 | 2.7% | 0.9% | 0.0% |
| NERC Total | 26.6% | 6.9% | 0.0% |
| b09 | 0.1% | 0.1% | 0.0% |
| b10 | 0.0% | 0.0% | 0.0% |
| b100 | 0.1% | 0.0% | 0.0% |
| BBSRC Total | 0.1% | 0.1% | 0.0% |
| p01 | 0.1% | 0.0% | 0.0% |
| STFC Total | 0.1% | 0.0% | 0.0% |
| e168 | 0.0% | 0.0% | 0.0% |
| t01 | 0.3% | 0.0% | 0.0% |
| x01 | 0.9% | 0.3% | 0.0% |
| x03 | 0.0% | 0.0% | 0.0% |
| x04 | 0.0% | 0.0% | 0.0% |
| x05 | 0.1% | 0.0% | 0.0% |
| x06 | 0.0% | 0.0% | 0.0% |
| External Total | 1.4% | 0.3% | 0.0% |
| d03 | 0.1% | 0.0% | 0.0% |
| d04 | 0.0% | 0.0% | 0.0% |
| d11 | 0.0% | 0.0% | 0.0% |
| d14 | 0.1% | 1.2% | 0.0% |
| d15 | 0.0% | 0.0% | 0.0% |
| d16 | 0.0% | 0.0% | 0.0% |
| d17 | 0.0% | 0.0% | 0.1% |
| d18 | 0.0% | 0.0% | 0.0% |
| d19 | 0.0% | 0.0% | 0.0% |
| d21 | 0.0% | 0.0% | 0.0% |
| d22 | 0.0% | 0.0% | 0.0% |
| d23 | 0.0% | 0.0% | 0.0% |
| d24 | 0.0% | 0.1% | 0.0% |
| d25 | 0.1% | 0.0% | 0.0% |
| Director's Time Total | 0.4% | 1.3% | 0.1% |
| **Overall Total** | **73.0%** | **38.0%** | **44.1%** |

### 3.2.6 HECToR Utilisation by Application Area




### 3.2.7 HECToR Projected Utilisation

The chart below shows a projection of system use. This projection assumes that the Cray XT4 has been shut down by the end of June 2011, and that the Cray XE6 system remains with its current configuration for the duration of 2011. The chart is a best guess of the projected future use of the system based on the per project profiles. If the project has over or under used their time allocation relative to the projected profile then the profile is re-scaled to take account of this. For dates in the past the actual use is displayed. It is clear from this that the allocated resources are no longer in line with the scale of the service since the arrival of the Cray XT6.



## 3.3 Helpdesk

A total of 3166 queries with a specified service metric were completed in this period.

### 3.3.1 Helpdesk Targets

| Metric | Pass | Total | Fraction | Target |
|---|---|---|---|---|
| All queries finished in 1 day | 2598 | 2627 | 98.9% | 97.0% |
| Admin queries finished in 1 day | 2370 | 2389 | 99.2% | 97.0% |
| Queries assigned in 30 min | 3034 | 3037 | 99.9% | 97.0% |
| Technical assessments in 10 days | 86 | 88 | 97.7% | 97.0% |

### 3.3.2 Queries by Service Metric

| Service Metric | Queries | Percentage |
|---|---|---|
| Automatic | 1724 | 54.5% |
| Admin | 665 | 21.0% |
| In-depth | 359 | 11.3% |
| Technical | 238 | 7.5% |
| No Metric | 92 | 2.9% |
| Technical assessment class-1a | 40 | 1.3% |
| Technical assessment class-1b | 34 | 1.1% |
| Technical assessment class-2a | 11 | 0.3% |
| Technical assessment class-2b | 3 | 0.1% |

### 3.3.3 Queries by Category

| Query Category | Queries | Percentage |
|---|---|---|
| Set group quotas | 493 | 15.6% |
| New User | 371 | 11.7% |
| Set user quotas | 275 | 8.7% |
| New Password | 269 | 8.5% |
| Access to HECToR | 206 | 6.5% |
| None | 192 | 6.1% |
| 3rd Party Software | 161 | 5.1% |
| Disk, tapes, resources | 143 | 4.5% |
| User behaviour | 135 | 4.3% |
| Add to group | 93 | 2.9% |
| Batch system and queues | 86 | 2.7% |
| Node Failure | 85 | 2.7% |
| Compilers and system software | 85 | 2.7% |
| New Group | 77 | 2.4% |
| Join Project | 72 | 2.3% |
| Other | 66 | 2.1% |
| Login, passwords and ssh | 62 | 2.0% |
| User programs | 58 | 1.8% |
| SAFE | 51 | 1.6% |
| Archive | 39 | 1.2% |
| Update account | 31 | 1.0% |
| Remove account | 21 | 0.7% |

| Query Category | Queries | Percentage |
|---|---|---|
| Performance and scaling | 19 | 0.6% |
| Courses | 16 | 0.5% |
| Delete from group | 15 | 0.5% |
| Create certificate | 15 | 0.5% |
| Static website | 14 | 0.4% |
| Grid | 8 | 0.3% |
| Delete from project | 3 | 0.1% |
| Porting | 2 | 0.1% |
| Network | 2 | 0.1% |
| Delete Certificate | 1 | 0.0% |

## 3.3.4 Queries by Handler Category

| Handlers | Total | Automatic | Technical | In-depth | Admin | No Metric | Technical assessment | %-age |
|---|---|---|---|---|---|---|---|---|
| OSG | 1884 | 1723 | 82 | 23 | 39 | 17 | | 59.51% |
| Cray | 110 | | 37 | 71 | 1 | 1 | | 3.47% |
| USL | 924 | 1 | 117 | 108 | 624 | 74 | | 29.19% |
| CSE | 248 | | 2 | 157 | 1 | | 88 | 7.83% |

## 3.3.5 Helpdesk General

### 3.3.5.1 esFS

The helpdesk worked extensively with user groups in April in order to reduce the volume of data held on the embedded Lustre filesystem. The team worked directly with the Principle Investigators to increase archive usage and reduce work storage quotas.

During October and November, the helpdesk played a major communication role during the migration of data to esFS. The phased transition of data for projects with large volumes of data required a significant co-ordination and communication with individual user groups.

### 3.3.5.2 Phase 2b

The arrival of the Phase 2b system required a number of changes to the Helpdesk software. These changes were mostly transparent to users. A number of technical and process changes were required in order to differentiate between the two systems. Monthly reporting mechanisms were updated to enable split reporting on each system.

The HECToR website was also updated in June for the arrival of the Phase 2b system. Additional status reporting was configured, and Phase 2b specific documentation was created.

### 3.3.5.3 Mailing List Mechanism

A key change in 2010 was the creation of a new mailing list mechanism. At the requests of users, a new system was configured in November, allowing users to opt in/out of different levels of email.

There are currently three mailing list options available:

- Major Announcements
- Service News
- System Status Updates

The 'Major Announcements' mailings contain information on major service upgrades and future plans. This option is enabled for all users by default.

The 'Service News' mailings contain information on upcoming training courses, CSE newsletters, events, and other general announcements. This option is enabled for all users by default.

The 'System Status Notifications' informs users when the service goes up or down, including the weekly reminders of the next planned maintenance shutdowns. This option is not enabled by default. Those wishing to receive this information need to explicitly subscribe to it. System status and planned maintenance information is available via the HECToR web page.

#### 3.3.5.4 Code Porting Activities

The contractual applications as managed by EPCC were ported to the Phase 2b system in June. These were ported again in mid-December to the Gemini platform.

### 3.3.6 User Quality Tokens

A number of quality tokens were received at the Helpdesk in 2010. The majority of negative tokens related to service outages. Positive feedback was received with regard to user support and the improvement in performance with the arrival of Gemini.

A summary per project is below:

| ***Project*** | ***Negative Tokens*** | ***Positive Tokens*** |
|---|---|---|
| e05 | 8 | 5 |
| e24 | 0 | 3 |
| e82 | 1 | 4 |
| e89 | 0 | 5 |
| e102 | 0 | 5 |
| e125 | 0 | 5 |
| e136 | 0 | 4 |
| e142 | 0 | 5 |
| n02 | 39 | 0 |
| x01 | 0 | 9 |
| **Total** | **48** | **45** |

## 3.4 Performance Metrics

| Metric | TSL | FSL | January | February | March | April | May | June | July | August | September | October | November | December | Annual Average |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Technology reliability (%) | 85.0% | 98.5% | 98.4% | 96.9% | 97.9% | 80.9% | 98.6% | 99.3% | 99.3% | 98.6% | 99.9% | 91.6% | 99.4% | 93.8% | 96.2% |
| Technology MTBF (hours) | 100 | 126.4 | 244.0 | 183.0 | 104.6 | 73.2 | 146.4 | 732.0 | 244.0 | 366.0 | 366.0 | 244.0 | 732.0 | 146.4 | 204.3 |
| Technology Throughput, hours/year | 7000 | 8367 | 8504 | 8379 | 8047 | 6984 | 8393 | 8636 | 8601 | 8348 | 8307 | 7934 | 8665 | 8088 | 8068 |
| Capability jobs completion rate | 70% | 90% | 96.9% | 100.0 % | 97.7% | 97.4% | 94.3% | 100.0 % | 97.7% | 98.2% | 96.7% | 94.3% | 92.9% | 100.0 % | 97.2% |
| Non in-depth queries resolved within 1 day (%) | 85% | 97% | 100.0 % | 100.0 % | 98.9% | 98.7% | 98.5% | 98.7% | 100.0 % | 99.0% | 97.7% | 99.3% | 99.5% | 98.2% | 99.0% |
| Number of SP FTEs | 7.3 | 8.0 | 8.1 | 8.7 | 9.4 | 8.6 | 10.3 | 10.7 | 9.0 | 9.0 | 9.1 | 8.1 | 9.0 | 7.6 | 9.0 |
| SP Serviceability (%) | 80.0% | 99.0% | 100.0 % | 100.0 % | 97.2% | 100.0 % | 100.0 % | 100.0 % | 100.0 % | 100.0 % | 100.0 % | 100.0 % | 100.0 % | 100.0 % | 99.8% |

*Colour coding:*

| | |
|---|---|
| Exceeds FSL | (green) |
| Between TSL and FSL | (amber) |
| Below TSL | (red) |

# 4 Systems Hardware

## 4.1 HECToR Technology Changes

2010 was a very busy period for the service, with a number of key technology changes taking place.

### 4.1.1 CLE 2.2 Upgrade

The HECToR operating system was upgraded to CLE2.2 in April. There were no major issues encountered. The lessons learned during the previous upgrade from CLE2.0 to CLE2.1 were clearly beneficial in ensuring that the process went smoothly.

### 4.1.2 Phase2b Installation

As part of the Phase 2b upgrade process, the Phase 2a system was reduced in capacity from 60 to 33 cabinets. This took place on 19th May as planned. No issues were encountered with the change.

The Phase 2b upgrade ran very smoothly. The hardware was delivered and installed as planned, and acceptance tests were completed in early June. A 20-cabinet Cray XT6 system was installed. The system comprises 44,544 cores (464 blades x 8 chips x 12 cores), delivering a peak performance of 338 T/Flops. The system uses the Seastar interconnect, which was scheduled to be replaced by Gemini later in 2010.

Early access users were granted access to the Phase 2b system on 10th June, and the remaining users were given access on 14th June. The Phase 2a system formally provided the main service until such time as the Gemini interconnect was installed and completed all acceptance and availability trials.

### 4.1.3 CLE3.1 Upgrade

The operating system on the Phase 2b service was upgraded on 29th September. All core third party applications, including the Unified Model were tested prior to the upgrade, and no major issues were encountered with the exception of a maintenance overrun. This upgrade was a necessary pre-cursor to connecting the Phase 2b system to esFS.

### 4.1.4 esFS

The migration to external lustre (esFS) was originally planned for 2Q10. Due to the major issues encountered in April, this was then delayed until 3Q10. Details on the issues encountered are in Section 4.2. Further issues with the upgrade were then encountered in July and August and the planned migration of user data was postponed for a second time. The decision was taken in late September to remove the Phase 2a system from user service in October for a period of 10-days to enable dedicated testing and debugging to take place. The testing period was completed and the decision was taken to progress with the migration of user data.

During October and November all user data from both the Phase 2a and Phase 2b systems was successfully transitioned to esFS. The majority of projects had their data migrated during an extended maintenance slot. Groups with large volumes of data were catered for on a case by case basis. Temporary disk space was made available for any users who wished to continue working during the migration process. No issues were reported during the migration process, and there have been no related faults reported since.

### 4.1.5 Gemini Upgrade

The upgrade to Gemini was originally planned for January 2011. In response to user concerns regarding the timing of this, the work was brought forward to December. The Phase 2b system was unavailable to users from the 6th to 17th December to accommodate the upgrade to the new interconnect. During this period, the Seastar interconnects were all replaced, the operating system was upgraded and all applications were ported to the new platform. The system was made available for users on 17th December, at which point charging was disabled. Acceptance tests were completed in early January 2011, and the 10-day availability trial completed on January 27th 2011. Charging has since been re-enabled.

Having been upgraded to use the Gemini interconnect, the Cray XT6 is now known as a Cray XE6. The Phase 2b system formally became the contractual service as of 1 Feb 2011.

## 4.2 Severity-1 Incidents

### 4.2.1 Technology Failures

Cray supplied technology has been responsible for 43 severity-1 incidents in 2010.

This is a breakdown of the failure categories:

- 11 High Speed Network (HSN) hardware failures (including 7 single link failures)
- 7 cabinet Power Distribution Unit (PDU) failures
- 6 module power/control failures
- 5 esFS related failures
- 2 Blue-Arc storage related failures
- 5 system software interruptions (including 2 security vulnerabilities)
- 3 instances of a late return to service following scheduled planned maintenance sessions
- 4 procedural interrupts (safe/web server, job closedown script, rack power supply, PBS restart error)

The largest number of hardware failures occurred on the High Speed Network, including seven single link failures. These failures are fatal to the Seastar based High Speed Network. Six module power/control failures were responsible for system wide interruptions in 2010. These failures impact the system due to the loss of power to the Seastar interconnect processors on the compute modules.

The Gemini interconnect fabric recently installed into the Cray XE6 system has resiliency features that enable the system to continue to operate through most HSN failure scenarios including single link failures and module power/control failures. As the production service moves onto the Cray XE6 system a significant improvement in system resiliency is expected due to these Gemini features.

Failures in the cabinet power distribution units were diagnosed as being related to the variable frequency drive (VFD) component. A Field Change Order (FCO) to improve the electrical grounding of this unit was installed in all Cray XT4 cabinets during August 2010 and has improved the reliability of these devices.

The shared lustre (esFS) filesystem was responsible for five system interruptions during 2010. Several faulty InfiniBand cables which caused some early instability in this system were discovered and replaced. Software patches and configuration improvements were applied to esFS servers as well as the Cray XT systems to provide a more robust and stable software environment.

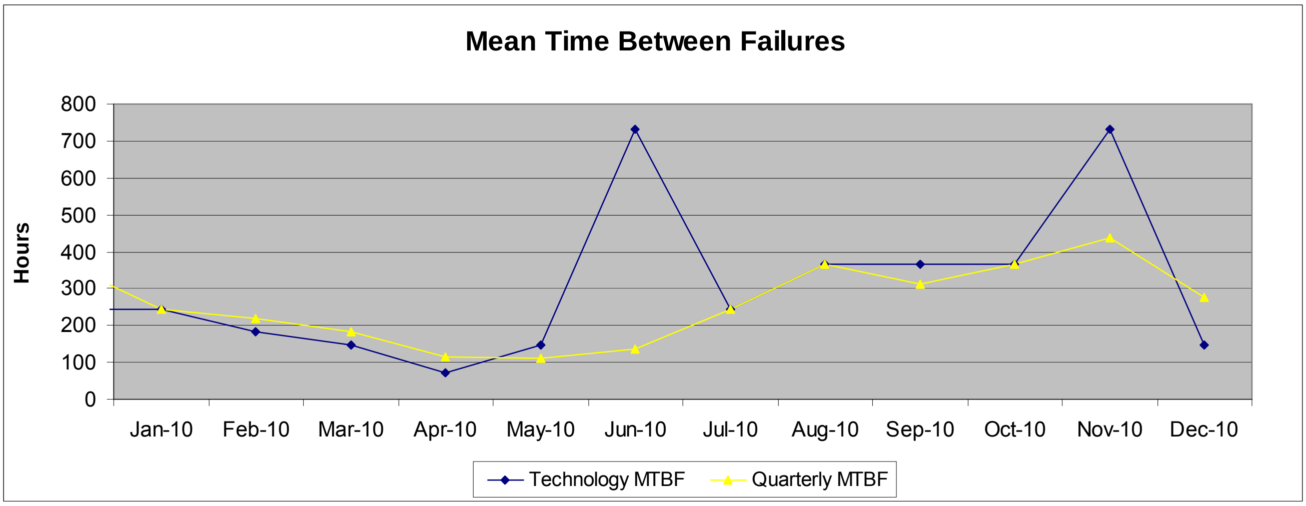
Two BlueArc storage system related incidents in 2010 have resulted in a plan to upgrade firmware on the BlueArc controllers during March 2011.

The system software was upgraded to CLE 2.2 in April 2010 and has proven to be a stable platform. Four software related incidents during the year included one PBS failure, an internal lustre panic, a HSN software problem and a system shutdown to resolve a security vulnerability in the base SuSE Linux software.

On three occasions, the system was not returned promptly at the end of a scheduled maintenance session. Improved maintenance session planning and earlier system reboots have been implemented to avoid these unfortunate incidents whenever possible.

Four procedural interruptions occurred in 2010. Lessons learned from these incidents and improvements to site procedures will continue to be implemented as the service evolves.

2Q10 was particularly challenging due to the high number of service failures, however reliability improved throughout the last two quarters of 2010.




### 4.2.2 Service Provision Failures

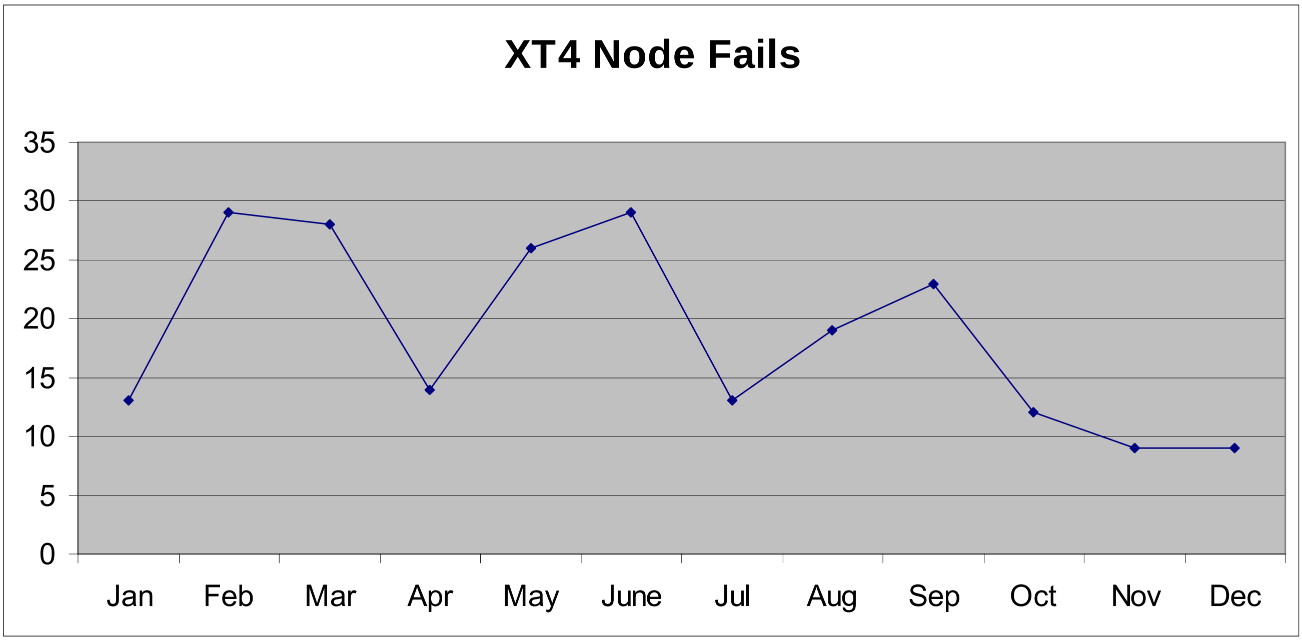
There was one service provision failure during 2010. During March there was a plant failure due to the loss of chilled water as a result of a valve failure in the free-cooling infrastructure. A process has since been adopted to test the valve on a regular basis.

### 4.2.3 External Failures

There was one failure in 2010 as a result of external causes. In March there was a fault on the external 11KV power distribution network.

## 4.3 Single Node Failures

There were a total of 224 node fails in 2010, compared to 250 in 2009.




Node failures as a result of hardware failures were relatively low at 16% of all node failures; with communications errors on the Seastar interconnect being the main fault. We would hope to see a reduction in such failures on the Cray XE6 platform in 2011 with the improved Gemini interconnect.

.

# 5 Cray Centre of Excellence

## 5.1 Executive Summary

The year 2010 saw the Cray Centre of Excellence for HECToR continue to deliver targeted high-end application support to a wide range of user groups. The Centre was heavily involved, along with the wider Cray applications team, in preparing users for the arrival of the Cray XT6 and Cray XE6 systems. This involved both improving scaling of applications to take advantage of the greater number of cores, and assistance in the OpenMP hybridisation of certain applications to take advantage of the increase in cores on a node. The Centre has also participated extensively in the Cray Centre of Excellence Steering Committee which aims to target the Centre's activities to important UK HPC applications and users.

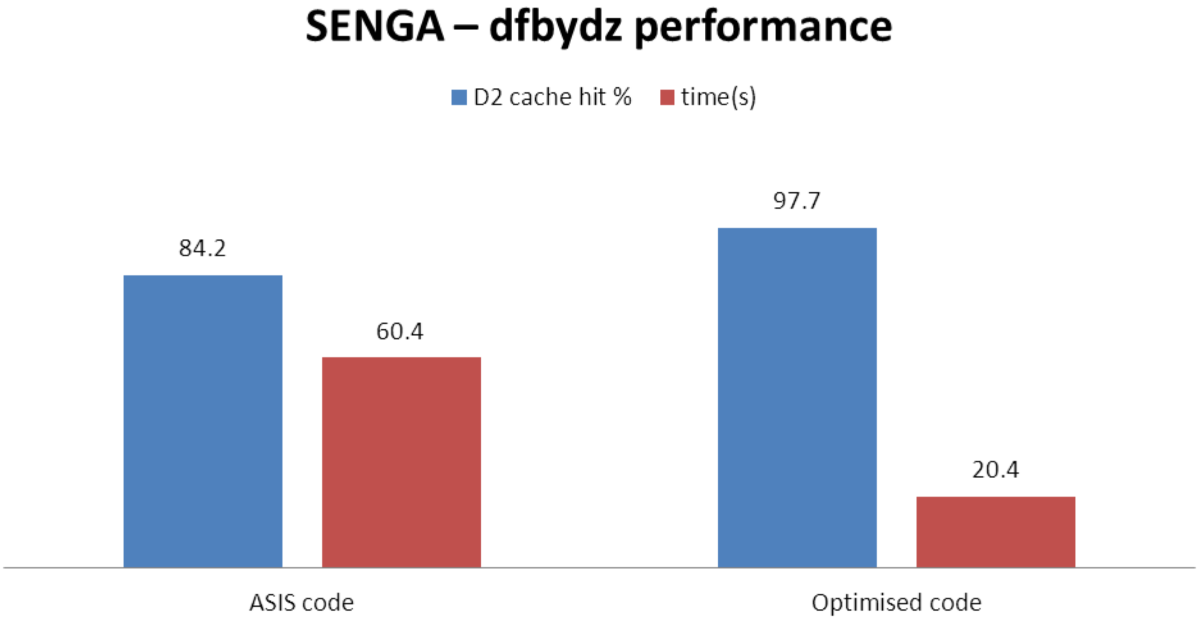
The Centre has engaged in user training and workshops, promoting the HECToR service throughout the UK, Europe and the rest of the world. Members have contributed to the distributed CSE process and Performance Working Groups; provided guidance and support for postgraduate students participating in High Performance Computing training; and contributed to the preparation of the UK's HPC user community and application base for the arrival of HECToR Phase 2b and the Cray XE6 system. This was in addition to the continued involvement of the Centre's core staff and members of Cray's Application and Development teams with HECToR activities where specialist input was essential.

## 5.2 Applications Support

### 5.2.1 CoE Steering Committee

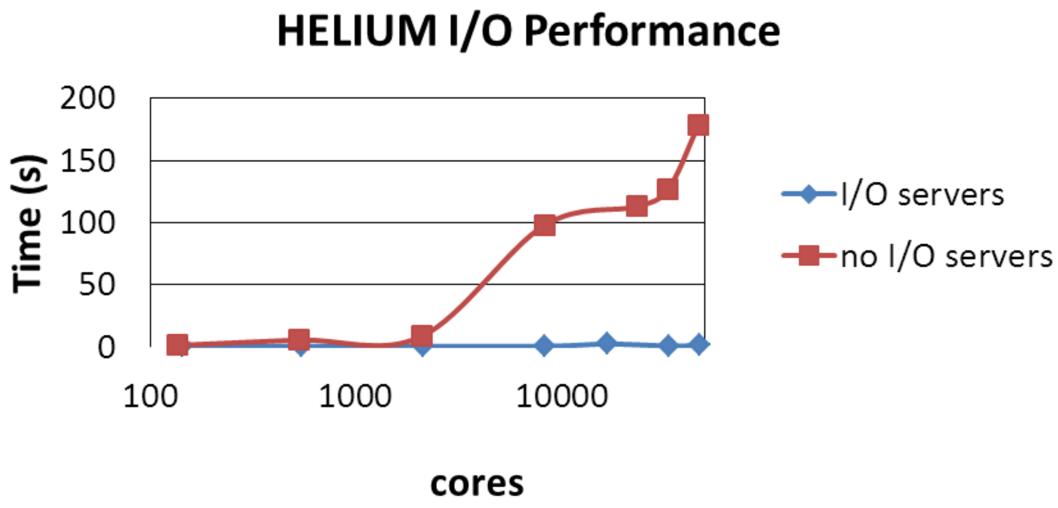
Last year saw the creation of the CoE Steering Committee, a joint venture between Cray, HPCX, and the UK academic community. The Committee, chaired by Professor Richard Kenway, aims to provide additional focus to the Centre's activities by targeting specific user groups which represent important HPC applications in the UK. 2010 saw the continuation of the CoE Steering Committee with several new projects two meetings, in February and July, to review progress. Attendees from a number of different user groups participated and after some evaluation, a number of projects have been identified and carried forward: SENGA, HELIUM, HiPSTAR, HiGEM, GADGET and CARP. The CoE has been actively working on these applications over the last year with the aim of providing the users with an Cray XT optimised version of their code, and making the necessary changes to the applications so that the users could return the science which was discussed at the Committee meeting.

- The DNS combustion application SENGA was optimised for Professor Stewart Cant of the University of Cambridge. SENGA is an important application in the field of combustion research, both here in the UK, and internationally. The Cray CoE first produced some very large scaling runs of SENGA on the Petaflop Cray XT5 system (jaguarpf) at ORNL to find out the scaling potential of the application. As can be seen in the figure below, the application was scaled to 64,000 cores of jaguarpf, which far exceeded the known scaling behaviour of the code. In addition a number of routines in this application have been optimised for cache reuse leading to a factor of three speedup in the targeted routines.




- The HiGEM project aims to test the skill of climate models in predicting decadal scale environmental phenomena like El Nino and help improve confidence in longer term predictions. The project will use a specific version of the Unified Model which will incorporate an I/O server that moves a significant bottleneck away from the compute nodes. The Cray CoE is assisting users at the National Centre for Atmospheric Sciences (part of the University of Reading) to tune the IO server and the UM in order to gain improved performance that will allow more ensemble members to run in the allocated time. The results of these simulations will form evidence that will be presented to the next International Panel on Climate Change.
- HELIUM integrates the time-dependent Schrödinger equation for few-electron laser-atom interactions. It is developed by Professor Ken Taylor at Queen's University, Belfast and scales very well to over 16,000 cores on HECToR. Through collaboration with the Centre of Excellence work has focused on separating the I/O from the computation to allow Helium to scale to the limits of the next phase of HECToR and beyond. By dedicating a small number of processors to handle writing checkpoint information to disk the highly scalable computation sections can continue uninterrupted. Checkpoints which previously took approximately 100 seconds on 32,000 cores on jaguarpf are reduced to less than 1 second by dedicating 125 processors as I/O

servers



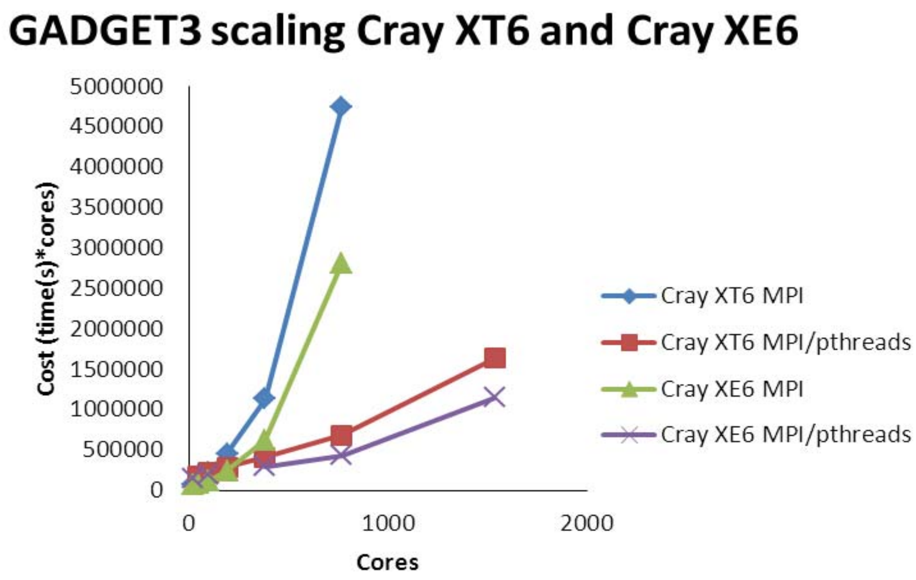

- HiPSTAR is a CFD simulation of subsonic jet engines developed by Dr Richard Sandberg at the University of Southampton. The project will perform a direct numerical simulation of fluid flow around a jet engine to further understand the noise generating mechanisms. A study of this nature has only become possible with the increased computing power available on HECToR. In the first instance, the Cray Centre of Excellence has tuned the serial and parallel sections of the code, creating an 8% overall improvement on smaller numbers of processors and increasing HiPSTAR's ability to scale up to beyond 2048 cores.

  HiPSTAR's original parallel decomposition was restricted to two dimensions using MPI which limited the total number of processor cores any problem could use. The increased number of shared memory cores in each node of the Cray XE6, from 4 to 24 allowed further decomposition along the third dimension using OpenMP.

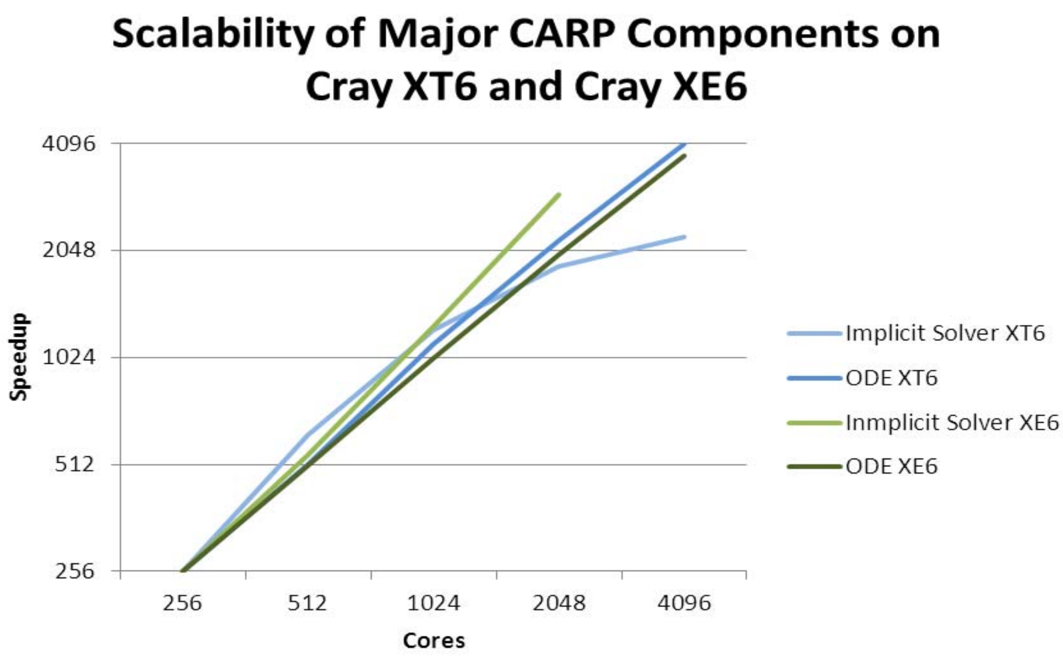
  The steering committee commissioned the Cray Centre of Excellence to adapt HiPSTAR using OpenMP and the resulting code shows much improved performance. The code demonstrated a 42% overall speed up on the Cray XE6 using OpenMP over the Cray XT6 using the same number of cores and the application may now scale up to 6 times further than was previously possible, taking it to within the entire capability of service. Further runs to assess the enhanced scalability are ongoing.

| | Wallclock time per model Timestep (s) | % improvement |
|---|---|---|
| **Cray XT6 without OpenMP** | 1.290 | |
| **Cray XE6 without OpenMP** | 1.087 | +19% |
| **Cray XE6 with OpenMP** | 0.908 | +42% |

- GADGET is a computational cosmology code used by researchers at the Institute for Computational Cosmology at the University of Durham. With the current datasets of interest, GADGET produces significant load imbalance at very small core counts. The CoE have been involved in obtaining and tuning of a new hybrid pthreads/MPI version of GADGET which significantly lessens the load balance issues with the application, and allows for a factor of 10 increase in scaling on the Cray XT6 and Cray XE6.



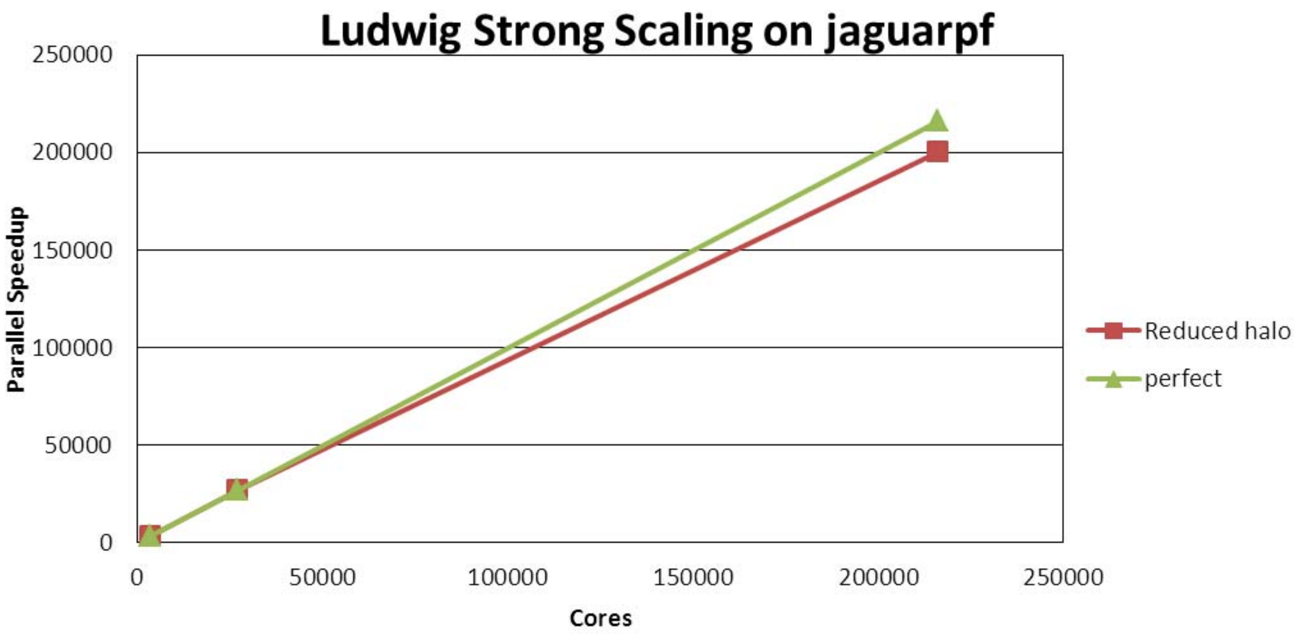

- CARP is used to study Cardiac Arrhythmia and treatments for cardiovascular disease. This project seeks to extend the ability of CARP to study the effect of these treatments not just on the heart but the surrounding areas; this is part of ongoing efforts to allow simulation to play a much greater part, not just in pre-operation scenarios but ultimately during operations themselves.



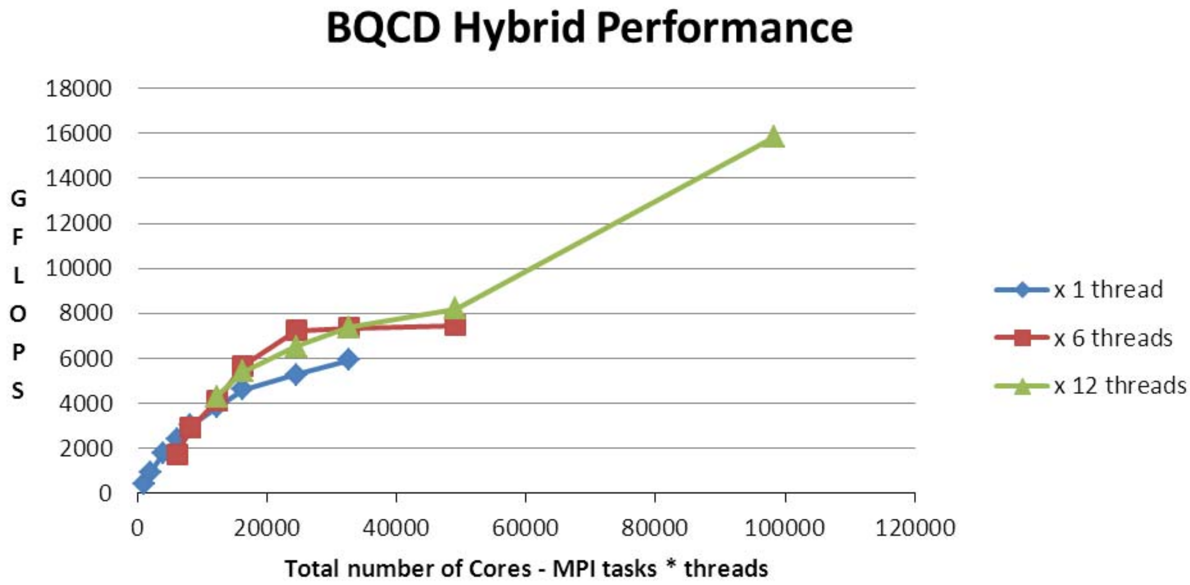
### 5.2.2 Other Applications

The Cray CoE has engaged with a number of other projects within the HECToR project, including:

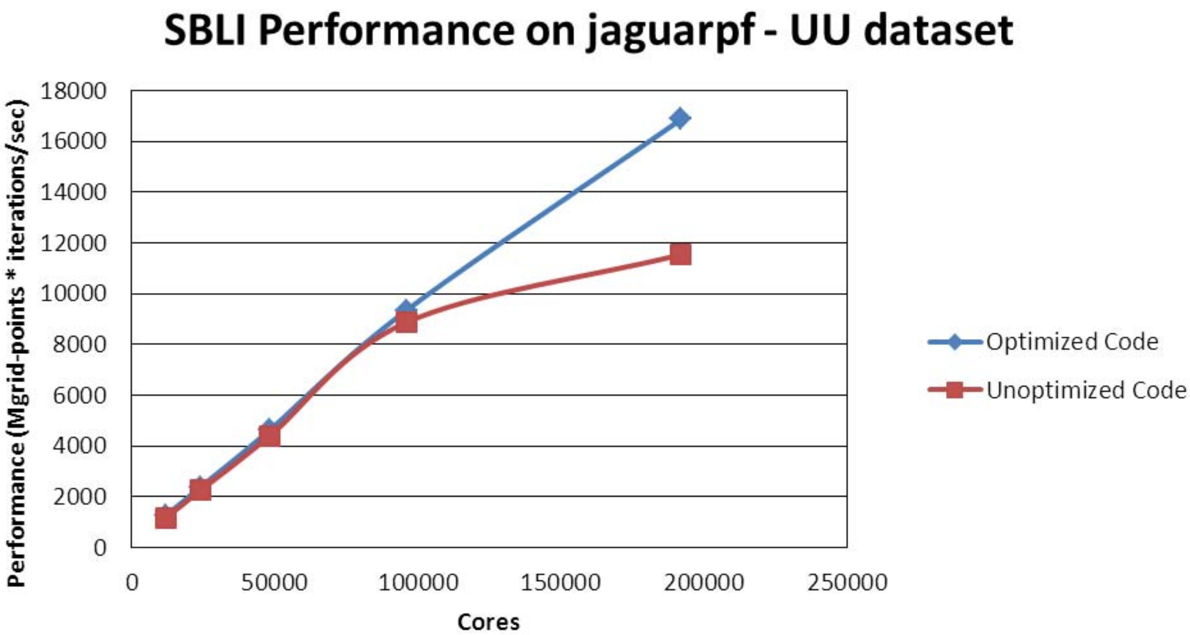
- **Ludwig**: Ludwig is a lattice-boltzmann code used to simulate fluid and multiphase flows. It is useful for both fundamental research into the nature of the multiphase flows, and in practical applications such as industrial and manufacturing flows. The Cray CoE were able to test the Ludwig application at scale on the jaguarpf system with the aim of testing a new halo swap optimisation put in place by the developers which was expected to show benefit at scale. To test this, runs on jaguarpf were performed up to 216,000 cores showing a significant benefit to the new halo swap routines.




- **BQCD**: BQCD is a QCD application which is part of the DEISA benchmarking suite. The Cray CoE assisted a DEISA user in running BQCD using the full HECToR system, and has since followed up this work on jaguarpf where we have run both the pure MPI version of the code along with the hybrid OMP/MPI version of the code. As can be seen in the figure below, the hybrid version of the code does very well at scale, and is good validation of the hybrid OMP/MPI approach on the Cray XT.

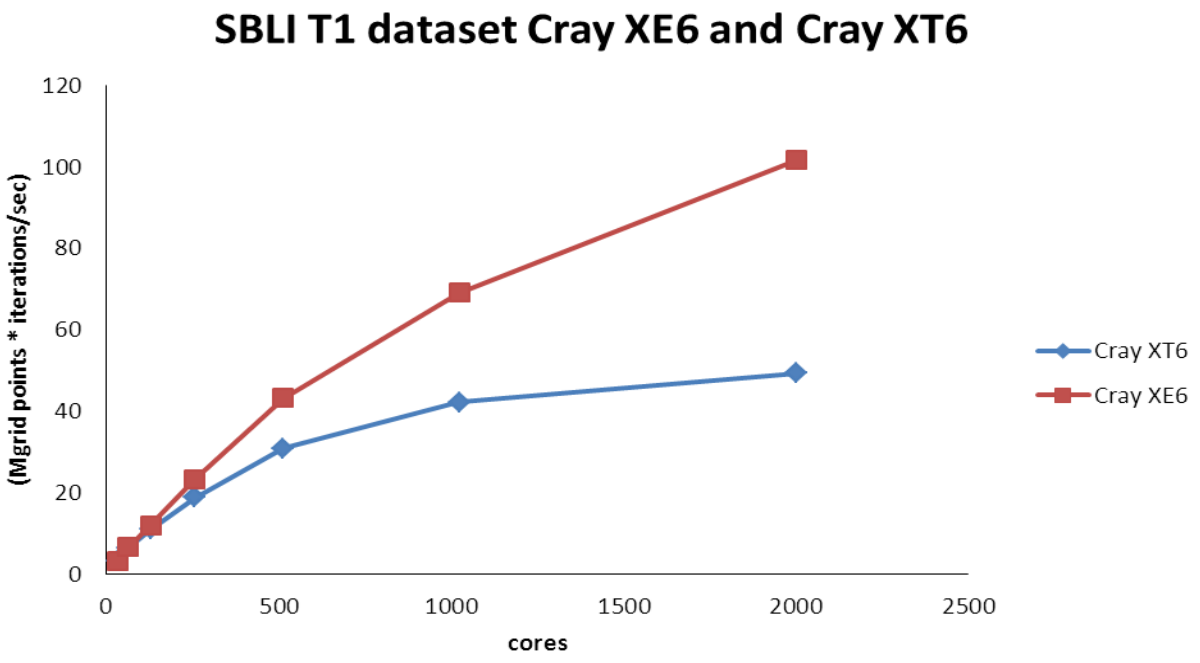



- **SBLI**: SBLI is a CFD code developed by Professor Neil Sandham at the University of Southampton and used extensively within the UK Turbulence Consortium. SBLI is used for direct numerical simulation (DNS) of turbulence and is under active development and use throughout the UK. The Cray CoE performed a number of scaling runs of SBLI on the HECToR and jaguarpf systems, running the code to 200,000 cores on jaguarpf, which is in excess of the scaling that DLR thought was possible with the provided dataset. In addition a number of improvements were made to the communication routines of SBLI to improve the scaling at very high core counts which led to a substantial increase in performance at scale.



### 5.2.3 HECToR GEMINI Preparation

The CoE contributed to the successful update of the Cray XT6 to a GEMINI based Cray XE6 system by engaging with key user groups and assisting them in optimising their applications for the new interconnect. A number of user groups were engaged including the UK Turbulence Consortium, the Materials Chemistry Consortium, and the NCAS users at Reading. The users have been very happy with the performance of the new interconnect, which is exhibiting much improved performance and scaling. For example, users of SBLI were able to increase the scaling of their application by a factor of two.




## 5.3 Cray Exascale Research Initiative Europe

2010 saw the establishment of the Cray Exascale Research Initiative Europe which is a research collaboration between Cray, EPCC and CSCS. The focus of the Initiative will be research into technologies, both software and hardware, which are required to get to Exaflop scale systems. As part of this Initiative, Cray increased the Edinburgh based team with the addition of Dr Harvey Richardson and Dr Alistair Hart. They will be working closely with Cray R&D and EPCC on topics such as programming models for PGAS languages, programming GPU accelerators, improved algorithms for FFTs and network and I/O profiling research. The team has been very active the first year producing a one-day tutorial on coarray FORTRAN, which was given at SC10 and at DEISA and HECToR workshops in Edinburgh. In addition the Exascale Initiative members have been very active in the GPU accelerator area, giving several user workshops on proposed directive based approaches to programming accelerated nodes. Their activities are further detailed in the following section.

## 5.4 Workshops and Training Events

The Cray CoE conducted a number of training events in 2010:

- On the 11-15th January, the Cray CoE gave a presentation titled “Challenges in HPC Application Development” at the EPSRC Extreme Computing Sandpit. The focus of this presentation was the obstacles, and possible solutions, in scaling applications beyond a petaflop.

- On the 2nd February and 1st July, the HECToR CoE gave updates on project status and future plans to the HECToR CoE Steering Committee in Edinburgh.
- At the CUG2010 Conference in Edinburgh, 24-27th May, the CoE gave a presentation titled "Using I/O Servers to Improve Performance on Cray XT Technology". This presentation detailed the work done to the HELIUM code as one of the CoE Steering Committee projects.
- On the 14-16th September, the CoE gave DEISA workshop in Edinburgh titled "Performance Optimisation on Cray XT Systems". This workshop introduced the Cray XT hardware and software and covered serial and parallel optimisation techniques. In addition, the CoE and members of the Cray Exascale Research Initiative in Europe gave a one day coarray FORTRAN tutorial.
- On the 12th October the CoE gave a presentation to at the HECToR User Group Meeting in Manchester titled "Cray Centre of Excellence for HECToR: Activities and Future Projects" outlining progress to date, current project status and upcoming activities.
- On the 13th October the CoE gave, along with NAG, a one-day Cray XT6 Workshop in Manchester to HECToR users. The aim of this workshop was to introduce the Cray XT6 hardware and software and Cray XT6-specific optimisation techniques.
- On the 22nd October there was a one day workshop on the Challenges in Programming Multicore and Accelerated Nodes. This workshop focused on shared memory parallelism for multicore nodes, and covered the challenges and opportunities presented by GPU accelerated compute nodes.
- There were a number of CoE and Exascale activities at SC10 including a one day coarray Fortran tutorial and two posters discussing programming of accelerated nodes, and optimization of scientific applications on HPC systems.
- Visited several user groups throughout the year to meet with a larger group of users, give updates on the HECToR service, and give targeted application support where necessary. This included visits to the UK Turbulence Consortium and Materials Chemistry Consortium meetings, along with several visits to the NCAS researchers in Reading.

## 5.5 Seminars

The Cray CoE has given a number of external seminars of interest to the UK HPC community:

- HPCN Workshop:  The Cray CoE gave a presentation titled “CFD Applications at Scale on Cray Architectures” at the HPCN Workshop in Braunschweig, 4-5th May. This presentation covered CFD application performance on Cray systems, and introduced technologies, such as PGAS, of interest to the CFD application developers.

- NAIS HPC Workshop:  The Cray CoE gave a series of presentations titled “PGAS Programming with UPC and Fortran coarrays” at the NAIS HPC Workshop in Edinburgh on 23-24th June.  These presentations introduced PGAS concepts and gave a general overview of both UPC and Fortran coarrays.

- HLRS Parallel Tools Workshop:  The Cray CoE gave a presentation at the HLRS Parallel Tools Workshop, 7-8th September in Stuttgart, titled “Performance Tools on the Cray XT Systems”.  The focus of this presentation was to introduce the Performance Tool suite on the Cray XT systems, and a look ahead at future tools development at Cray.

- GPUs and Accelerators in HPC:  This workshop was held 28-29 September at STFC Daresbury.  The Cray CoE and Exascale team members gave a presentation entitled “Accelerator Directives – A User’s Perspective” covering a directive based approach to programming accelerated nodes.

- UK GPU Computing Conference:  At this meeting held 13-14 Dec, the Cray CoE and the Exascale team members gave a presentation titled “Directive-based programming for GPUs, accelerators and HPC”.

- KCSE Annual Meeting:  At the KCSE Annual Meeting at KTH, Stockholm, on 16-17th Dec, the Cray CoE gave a presentation titled “Cray and HPC Applications”.  This presentation focused on how Cray deals with applications from presales up to postsales support, and also covered the activities of the Cray CoE.

# 6 The HECToR Computational Science and Engineering (CSE) Support Service

## 6.1 Overview of the CSE Service

The Computational Science and Engineering (CSE) service exists to help the user community to make the best use of the HECToR hardware by providing training, web-based resources, and assistance with porting, optimisation and tuning of software. The service is provided by the Numerical Algorithms Group Ltd (NAG), a not-for-profit Company with offices in Oxford and Manchester, and 40 years' experience developing mathematical and statistical software. The *Core Team*, made up entirely of NAG staff, responds to in-depth software problems reported by users via the HECToR helpdesk, processes Technical Assessments related to applications for access to HECToR, runs a range of training courses and maintains a range of good practice guides and reference material as part of the service website, and undertakes various outreach activities. The *Distributed Team*, made up of a mixture of NAG staff and staff employed via a contract with a third party, provides dedicated resources for projects (DCSE projects) to enhance specific applications or support groups of users for periods of between six months and two years.

## 6.2 Highlights of 2010

2010 was another busy year for the CSE service with the introduction of the XT6 and its subsequent upgrade to an XE6 in December. We have helped many users make the transition to the new architecture and a number of DCSE projects have been concerned with adapting codes to make better use of a large number of many-core nodes (e.g. through better use of shared memory).

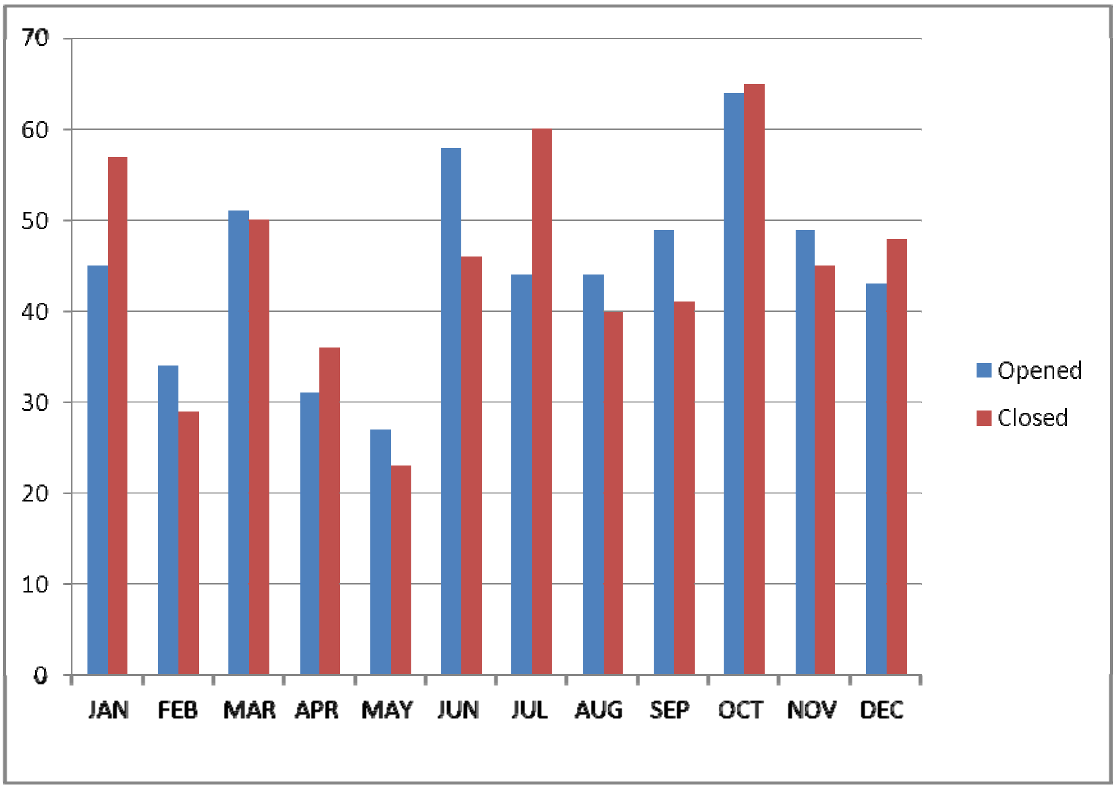
Training has continued to be popular and at the end of December we were just short of having 1000 attendees on our courses since the start of the HECToR service. We continue to be flexible in offering courses at users' institutions.

The DCSE programme is now flourishing and we are having to be increasingly selective in where we target resources. To date over 47 person-years of effort have been allocated through this mechanism and the benefits of this work are being felt throughout the user communities of the codes that have been worked on, as well as by users of HECToR.

## 6.3 The CSE Helpdesk

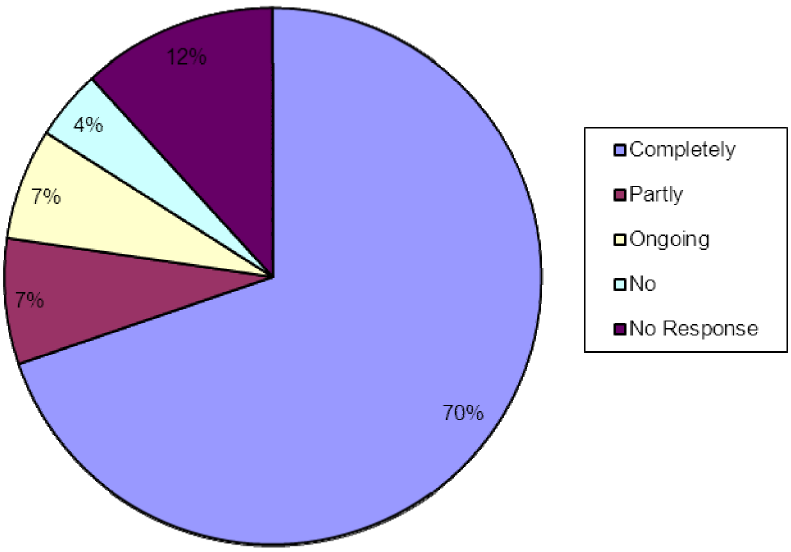
The Core CSE team handles queries from users forwarded by the service helpdesk, carries out technical assessments of applications for HECToR time, undertakes various outreach activities and runs the training courses.

The queries received by the CSE team vary from straightforward requests for advice to requests for assistance in porting, tuning etc. Some queries are resolved straight away while others develop into small projects lasting weeks or even months. The team resolves most queries but if, after investigation, they are found to be connected to system issues, then they will be re-assigned to the Service Provider (UoE HPCx) or to Cray. The following table shows how many queries were opened and closed each month.

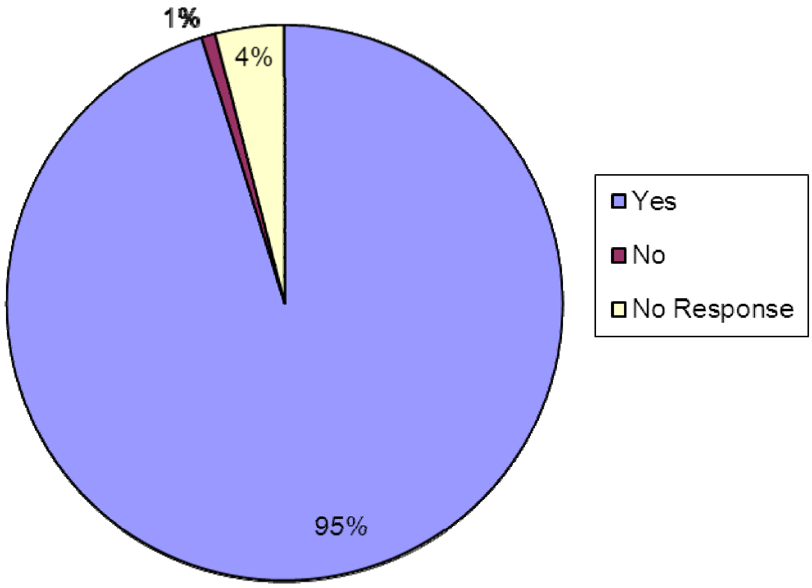



In cases where a technical query (as opposed to a request for a Technical Assessment) is resolved by the team (strictly speaking where the query is closed within the SAFE system by the CSE team), the user is invited to fill in a questionnaire giving feedback about his or her experience and satisfaction with the outcome. This year, 125 have been returned. The responses to the questions are summarised as follows:

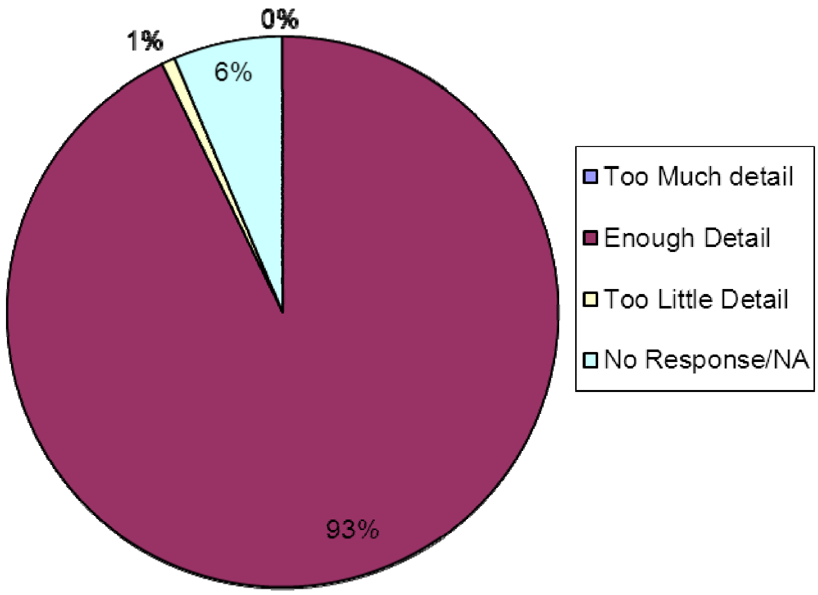
**Has the problem raised in your query been resolved by the information provided by the helpdesk?**

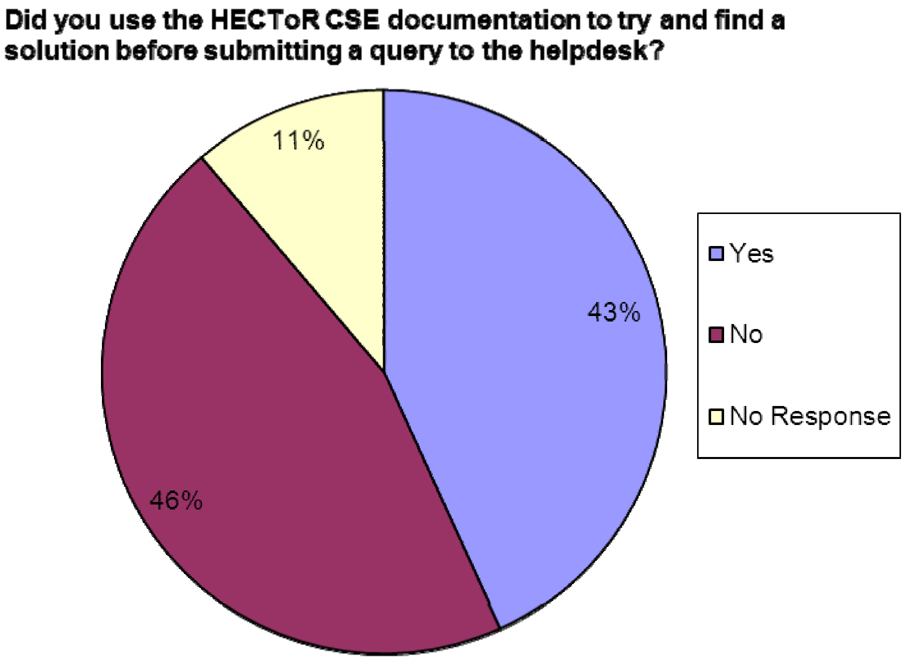


**Were all communications replied to promptly and the text clear and understandable?**




**Were technical issues explained in sufficient detail?**






On the relatively rare occasions that a negative response is received in feedback the CSE team will attempt to understand the reasons behind the response and, if necessary, change or improve their procedures. Most feedback is, however, extremely positive.

## 6.4 Training

During this period NAG has offered a range of training course, varying from HECToR-specific courses (e.g. *Programming the X2 Vector System),* through more general HPC courses (e.g. *Parallel Programming with MPI*), to courses teaching good practice in software engineering (e.g. *Best Practice in HPC Software Development*). Most courses have been run by NAG staff, although we have courses in collaboration with the University of Warwick, STFC and Allinea.

Take-up for the courses this year has been extremely good, with many of the courses being delivered at users' institutions. We have provided training at Imperial College, Bath, Southampton, Nottingham, Leeds, Warwick, Reading, Edinburgh and UCL, as well as the NAG offices in Oxford and Manchester. A total of 77 days of training have been offered and delivered, above our contractual level, and there have been 354 registrations.
The complete list of courses offered in 2010 is as follows. The courses range in duration from one day to a week, and complementary courses are often scheduled together for convenience. Many of these courses were run several times in different locations. In addition we delivered a module on Advanced Computational Methods within the Doctoral Training Centre for Complex Systems at the University of Southampton.

- Parallel Programming with MPI
- OpenMP
- Best Practice in HPC Software Development
- Introduction to HECToR
- Quad Core Training
- Fortran 95
- Exploiting Parallel CASTEP on Large-scale HPC
- Parallel IO

- Debugging, Profiling and Optimising
- Multicore
- Programming the X2 Vector System
- Core Algorithms for High Performance Scientific Computing
- Debugging with DDT
- DL_POLY
- Scientific Visualisation

## 6.5 Other Core CSE Activities

The CSE team publishes a newsletter approximately every six weeks which contains information about upcoming training, deadlines for Distributed CSE applications, tips and hints for programming HECToR, and news of changes to the service, particularly the development environment. We make regular meetings to users, both individually (for example to provide assistance preparing an application for HECToR time or discuss training needs), and collectively (for example attending meetings of the UK Turbulence Consortium). We attend appropriate conferences and workshops, and maintain links with other major international HPC facilities.

## 6.6 Distributed CSE

The aim of the Distributed CSE (DCSE) programme is to provide researchers with resources to enable them to:

- port their codes onto HECToR, in particular to work with new codes or to enable previously unsupported features in existing codes;
- improve the performance of their codes on HECToR;
- re-factor their codes to improve long-term maintainability;
- take advantage of algorithmic improvements in the field of high-performance computing.

DCSE is designed to complement the usual instruments for funding research which often cannot be used to support software development and maintenance of this kind. The necessary staff can be provided by NAG, an appropriate third party, recruited specially, or seconded from an existing project at the host institution. Where necessary the CSE service will provide special training to equip the DCSE team member with the skills necessary to perform his or her project. A DCSE project can be allocated up to four years of effort, and can last for up to two years. From 2011 these limits have been halved, to reflect the fact that we are now in the latter half of the service’s life.

Applications for DCSE support are considered by an independent review panel, chaired by Professor Chris Bischof of the University of Aachen. The panel recommends which applications should be funded wholly or partially, and NAG then works to agree a contract with the applicant who covers the mechanism for resourcing the project, timescales and a programme of work. The panel met twice during 2010, and considered a total of fifteen applications requesting 177 months of effort. Of those, eleven have been recommended to receive a total of 109 months of effort.

Nineteen projects ended during 2010. At the end of December, twenty projects were running with four more scheduled to start in 2011. Since the start of the service the DCSE panel has allocated just over 47 person-years of effort. Wherever possible, code changes are fed back to the code developers, to enable users of machines other than HECToR to benefit.

## 6.7 Completed DCSE Projects (during 2010)

### 6.7.1 HECToR Distributed CSE Support for OpenFOAM

OpenFOAM is a widely used and versatile open-source CFD toolbox. This DCSE project was allocated 12 months of support at the December 2008 panel. Paul Graham of EPCC supervised the work along with Professor David Emerson of STFC. Gavin Pringle from EPCC performed the development between April 2009 and March 2010.

The objective was to install and test key OpenFOAM solvers on HECToR, benchmark a variety of the modules available, and publicise the results to potential users. OpenFOAM was ported to the XT4 successfully. Scaling results for six sets of benchmarks were produced. Documentation about how to use OpenFOAM on HECToR was placed on the third party software wiki page (https://wiki.hector.ac.uk/userwiki/OpenFOAM). There is demand for OpenFOAM, and it is currently being ported to Phase 2b by Cray CoE and EPCC as a core project.

### 6.7.2 Multigrid improvements to the Citcom code

The Citcom code is part of one of the most widely used suite of computational codes in solid-earth geodynamics. This project was allocated 12 months effort at the September 2008 call. Dr Jeroen van Hunen of the Department of Earth Sciences at the University of Durham supervised the work, performed by Sarfraz Nadeem of NAG between January 2009 and April 2010.

The objective was to improve the parallel performance of Citcom by implementing multigrid methods to improve the rate of convergence. During the project, the code was benchmarked for various multigrid schemes which were implemented to enable Citcom to achieve faster convergence. For the best cases Citcom now performs over 31% faster for the V-cycle multigrid scheme and over 38% faster for the W-cycle multigrid scheme in comparison to the corresponding FMG(V) and FMG(W) schemes respectively for the simple 2D test problem on Phase 2a. The benchmarking provided guidance on which multigrid scheme to use for best performance in various situations, and identified key I/O related optimisations for future work. Citcom has used around 1.8 million AUs (allocation units) so far on HECToR. The key result for researchers is the scalability and faster convergence, making it possible to do new science. In particular, the significant development of local mesh refinement in combination with multigrid, and future work based on this important pioneering study, will enable Citcom users to address much larger numerical problems.

### 6.7.3 Scaling Turbulence Applications to thousands of cores

This work was to improve the parallel performance of the Ekman Boundary Layer CFD code, EBL. The project was proposed by Dr. Gary Coleman of the School of Engineering Sciences at the University of Southampton. It was allocated 12 months support at the December 2008 round. David Scott of EPCC worked on the project between April 2009 and April 2010. The EBL code models pressure driven flow over a single surface using a spectral decomposition. The objective was to re-factor EBL to replace the one-dimensional data decomposition with a two-dimensional one in order to improve scalability from the hundreds to thousands of cores. This project has enabled EBL to be re-engineered such that , for a problem size that previously would have used a maximum of 365 cores, good scaling can now be achieved for up to 14,000 cores on Phase 2b. EBL can now run with much larger problem sizes and with higher Reynolds numbers due to the two-dimensional domain decomposition.
The UK Turbulence Consortium has an allocation of 484M AUs of which they have so far used 23M. EBL is only one of several codes that they use. Dr. Coleman's group are

particularly interested in performing numerical experiments on wind turbines and canonical turbulent flows at much higher Reynolds numbers than previously possible.

### 6.7.4 Better Parallelism for the electron-atom scattering codes PRMAT

This project was allocated 12 months effort at the January 2008 call, to improve the parallel performance of the electron collision R-matrix code PFARM. Dr. Martin Plummer of Daresbury supervised the work with Andy Sunderland and Cliff Noble, also from Daresbury, each working a third of their time during July 2008-March 2010. The aims were to:

- Improve performance and scalability of PFARM on HECToR. This would involve a major upgrade to the current parallelization, capable of increased numbers of channels and scattering energies: the scattering resonances which need to be taken into account in future calculations will require grids of energy points increasing from around 10000 to 20000 or more.
- Introduce an alternative propagation method, the Airy Logarithmic Derivative (ALD) method.
- Improve methods of load-balancing, task-harnessing symmetry and energy parallelization, and optimal multicore memory sharing (in readiness for HECToR Phase 2 and beyond) based on passive communications and features of Fortran 2003.
- Use the experience of future-proofing the PFARM code to parallelize HAM optimally. (RMATRX2/95 consists of four codes, RAD, ANG, HAM and PDG95.)

The outcome of this work was that the parallel performance of PFARM was improved and a new Airy Logarithmic Derivative (ALD) propagator code was developed, FARM2, which is much more memory efficient while maintaining accuracy and performance. The external region Hamiltonian matrix parallel diagonalization EXDIG is now nearly four times as fast on Phase 2a with 8192-16384 cores. Improved load balancing and further parallelization for the R-matrix generation code EXAS in the fine and coarse regions gives around a 50% reduction in run time for 1024 cores. For the entire PRMAT simulation an overall performance speedup of 2.89 is now achieved when using 16384 cores.
The external region codes are used by groups at Daresbury Laboratory and Queen's University Belfast. When combined with the inner region PFARM codes this enables:

- Electron scattering and accurate determination of electron impact excitation rates for low ionization stages in the astrophysical elements Fe, Ni and Co and in the study of laser-produced Sn plasmas.
- A broader range of experimentally observed temperatures to be studied in addition to providing a greater range of individual atomic data.
- Electron scattering by atoms with highly term-dependent wavefunctions such as those for C, N and O which are important in atmospheric physics and solar physics.

### 6.7.5 WRF code optimisation for Meso-scale Process Studies (WOMPS)

The Weather Research and Forecast (WRF) Model is a medium to global-scale model for forecasting the weather and for process studies in the atmosphere. This project was allocated 10 months effort at the April 2008 call and was supervised by Dr. Alan Gadian of the School of Earth and Environmental at the University of Leeds. Andrew Porter of Daresbury performed the work between January 2009 and June 2010. The aims were to:

- Improve the efficiency and scalability of WRF on HECToR.

- Provide guidelines to enable users to choose optimal configurations of domain structures.
- Provide optimal compiler options for optimisation of key (numerically/IO intensive) routines.
- Enable around twenty UK scientists to efficiently use the WRF model to examine the scientific processes of cloud physics, pollution dispersion, orographic flow, storm development, idealised and forecast flow.
- Produce a document explaining the advantages and disadvantages of each optimisation.

The outcome of this DCSE project was an optimised WRF model. WRF in hybrid (MPI / OpenMP) mode was found to give the best parallel-scaling performance. Attention was paid to cache usage and compiler options, and recommendations for the best choice of domain decomposition and the optimum I/O configuration were made in a document which was distributed to HECToR WRF users.

The *Modelling the UK Wind Power Resource* project was awarded 3.6M AUs by the June 2010 Resource Allocation Panel specifically to run WRF.

### 6.7.6 Efficient Parallel Algorithms for ChemShell

ChemShell provides a means of integrating quantum mechanical (QM) and molecular mechanical (MM) software packages to perform combined QM/MM calculations. The external programs are used for energy and gradient calculations while ChemShell performs higher level tasks such as geometry optimisation. ChemShell can exploit the parallelism of the external programs but does not run in parallel itself.
This project was allocated 9 months effort at the January 2008 call. Dr. Paul Sherwood at Daresbury and Professor Richard Catlow from the Department of Chemistry at UCL proposed the work with Thomas Keal, also from Daresbury, performing the implementation between January 2009 and June 2010. The aims were to:

- Modify ChemShell to change the parallel strategy from a one-level scheme (all parallel worker tasks dealing with the same chemical system) to a two-level strategy in which a number of systems are dealt with concurrently, each parallelised over a subset of the available nodes. This will allow the rapid computation of finite difference Hessians on the HECToR system, greatly assisting a number of current projects for which Hessian information is available by slower methods.
- Deploy this technology in the nudged elastic band (NEB) TS location algorithm, running in a two-level parallel fashion. This will involve parallelisation of the algorithm as implemented in the (serial) DL-FIND code and application of the task farming strategy to QM/MM Calculations.
- Incorporate parallel search algorithms to provide an automated search for low energy regions of the potential surface, and fast convergence of guessed geometries.

The main outcome of the project was that an extra layer of parallelism was added to the ChemShell environment to enable calculations to scale to over 1000 cores for more efficient studies of heterogeneous catalysis on HECToR. The extra layer of parallelism enables better scaling properties than relying on the external programs alone.
The finite-difference Hessian code within ChemShell was re-written to allow task-farmed execution of the gradient evaluations. In a 1024 core test case, speed-up factors of up to 7 were observed compared to the original (non-task farmed) implementation. Parallel DL-

FIND has been interfaced to ChemShell to make parallel geometry optimisations available. The NEB method for transition-state determination in DL-FIND was parallelised and a speed-up factor of over 8 was found for a QM/MM surface chemistry test case.
ChemShell is used particularly within the themes of chemical reactivity and nanomaterials in the Materials Chemistry Consortium at UCL. On Phase 2a around 14M AUs (allocation units) of HECToR supercomputer resources were used by ChemShell between 13 October 2009 and 15 November 2010, and an additional 3.4M AUs have been used on Phase 2b. The software development from the DCSE project allows more efficient runs across in excess of 1,000 cores and a reduction in runtime of 50-90% compared to the older implementations.

### 6.7.7 HPC Driven Next-Generation Modelling of the World's Oceans (ICOM)

The three-dimensional non-hydrostatic parallel ocean model Fluidity-ICOM uses a control volume finite element discretisation method on unstructured meshes with dynamic adaptivity. This project was allocated 12 months of support at the April 2008 round. Dr. Mike Ashworth of DL supervised the work, performed by Xiaohu Guo (also of Daresbury) during March 2009 to May 2010. The aims were:

- An algorithm for optimal renumbering of the mesh nodes.
- Improvements to mesh adaptivity performance.
- Better I/O by implementing a subset of writers approach.
- General code re-factoring and compiler optimisation.

The outcome of this work:

- Fluidity-ICOM now supports block-CSR for the matrix assembly.
- Interleaved I/O has been implemented to the vtu output. The parallel I/O strategy has not yet been applied to the mesh file output as the end format is still to be decided.
- An optimal renumbering method for PETSc linear solver performance was implemented.
- Fluidity-ICOM has complex dependencies on third party software. Together with core CSE support, several modules were installed for users to build Fluidity-ICOM more easily.
- A scalability analysis was performed for the differentially heated rotating annulus benchmark. This allowed the performance of the parallel mesh optimisation method to be evaluated in the context of a “real” application.

This project has enabled Fluidity-ICOM to be transformed from a code that was primarily used on institution level clusters with typically 64 tasks per simulation into a high-performing scalable code which can be run efficiently on 4096 cores on Phase 2a. The central version of Fluidity-ICOM used 2.1M AUs on HECToR Phase2a between October 2009 and November 2010.

### 6.7.8 Improvements in I/O for DL_POLY_3

DL_POLY_3 is a general purpose package for classical molecular dynamics (MD) simulations developed by I.T. Todorov and W. Smith at Daresbury. The main purpose of this code is to enable the exploitation of large scale MD simulations on multi-processor platforms. This project was allocated 6 months of support at the December 2008 round. Dr. Ilian

Todorov supervised the work which was performed between March 2009 and April 2010 by Ian Bush of NAG.
The aim of the project was to improve the I/O performance so that real science could be performed on many hundreds of cores when simulating large systems. This optimisation would involve large scale data rearrangements to enable both good use of the Lustre file system and also an efficient parallelisation of the I/O. The modular design would also allow it to be adapted very easily to use either MPI-I/O or NetCDF.
The following points summarise the outcome:

- For sufficiently large systems the code does now scale to many thousands of Phase 2b cores, even when performing realistic amounts of I/O.
- All large I/O transactions are now performed in parallel.
- The writing of files is now up to 3 orders of magnitude faster than the original.
- The reading is up to two orders of magnitude faster.
- A good set of default parameters for the method have been found, and so most users need make no change to their input files to get good I/O performance.
- I/O based on both MPI-I/O (the default) and netCDF have been implemented.
- If using MPI-I/O the files are in exactly the same format as those used by the older method.

Between October 2009 and November 2010, the centrally installed version of DL_POLY alone consumed 7M AUs on Phase2a and 3.2M AUs on Phase2b. This is not taking into account usage from those who compile their own version. This six months of work resulted in the I/O for DL_POLY_3 being, on average, 50 times faster. The new code has reduced the time taken to write a single snapshot from 3 minutes to less than half a second, resulting in an overall factor of 20 improvement in net performance. This work enabled Dr. David Quigley from the Department of Physics and Centre for Scientific Computing, University of Warwick to use a specific model to investigate eggshell formation.

### 6.7.9 Hybrid time-dependent density functional theory in CASTEP Part 1

This project was allocated 10.5 months effort at the September 2008 call. Dr. Keith Refson of the STFC Rutherford Appleton Laboratory (RAL) supervised the work which was performed by Dominik Jochym, who is also based at RAL, from February 2009 to March 2010. All the original objectives were achieved, namely:

- Implementation of a scheme (based on the existing DFPT module code in CASTEP), to compute the electronic response to an external electric field of a set frequency. This provided a reference calculation against which further calculations could be benchmarked.
- Production of a releasable implementation of TDDFT using Hutter's published method with a comparison of the performance of two eigensolvers, namely ARPACK and EA19.
- Porting the implementation to HECToR with more extensive testing, debugging and benchmarking against previous calculations. On HECToR Phase 2a, 80% parallel efficiency was achievable on up to 256 processing cores. On Phase 2b, comparable calculation times could only be achieved by using 4 cores per node, i.e. 1 core per hex core die.

A scalable implementation of TDDFT in CASTEP will now give the UK electronic structure community an opportunity to address cutting-edge scientific problems in areas such as inorganic and organic photovoltaic materials, catalytic reactions at surfaces, light-emitting polymer materials for optical displays, and femtosecond laser chemistry. The inclusion of the use of hybrid functionals promises to address some of the limitations that have previously hindered the application of TDDFT to extended systems. Between October 2009 and November 2010, the centrally installed version of CASTEP alone consumed 50.7M AUs on Phase2a and 40.6M AUs on Phase2b. This is not taking into account usage by many who compile their own version of CASTEP.

### 6.7.10 Massive Remote Batch Visualizer – MRBV

The AVS/Express Distributed Data Renderer (AVS/Express DDR) allows the visualization of datasets that are too large to visualize using GPU hardware. The aim of the project was to port the AVS/Express DDR visualization application to HECToR Phase 2a.
This project was allocated 10 months of support at the December 2008 round. The development was mainly concerned with the optimisation of the parallel I/O module.
Dr. Martin Turner of the Department of Research Computing at the University of Manchester led the work. George Leaver who is based in the same department performed the implementation between April 2009 and April 2010. The outcomes:

- Replacement MPI library for the AVS/Express user interface code. This allows it to forward MPI function calls from a login node (on which the user interface code must execute) through a proxy running on a back end node and enables the existing AVS parallel module architecture to be used for module development without the need to introduce another communication API. Custom AVS applications can then be developed where a user interface is required for interactive use.
- A parallel image compositing method using MPI has been implemented to replace the sockets-based code in AVS/Express.
- Improvements to the AVS rendering code allow interactive use of the application with many more processors than any previous installation of the product. Volumes of sizes up to approximately $(7000)^3$ have been rendered.

AVS/Express DDR provides parallel module processing, where various visualization techniques are applied to domain decomposed data. Parallel rendering and image compositing, utilizing the distributed memory which is available on HECToR, allows significantly larger datasets to be visualized. Initially a batch rendering system was proposed, however, modifications to the AVS code allow interactive use of the application on HECToR. The work has been promoted extensively throughout the University of Manchester.

### 6.7.11 Fast Fourier Transformations for Gyrokinetic Plasma Simulations

The GS2 application uses a gyrokinetic approach to simulate micro-turbulence within magnetised fusion plasmas. This code spends significant amounts of time performing fast Fourier transformations and the overall aim of this project was to upgrade the code's usage of the legacy library FFTW2 to the newer FFTW3 library for more efficient usage of SSE3 instructions on multicore processors. For this work an in-depth analysis of how GS2 uses FFTW was required.
This project was supervised by Dr. Colin Roach of the Culham Centre for Fusion Energy. It was allocated 6 months funding at the March 2009 round. Joachim Hein and Xu Guo from EPCC performed the work between August 2009 and May 2010.

- Detailed analysis showed that for the FFT calls relevant for GS2, the benefits from the SSE instructions are at best minimal. This was unexpected, and has shown that there is little benefit from SSE even when using smaller problems to make the problem fit into cache, or when using only a single core of the processor to give the compute task more memory bandwidth and level 3 cache.

- An in-depth analysis of the profile after upgrading the FFTW library has shown the data redistribution routines inside the transformation routines to be very costly. A test case demonstrated that substantial performance gains can be achieved by removing the indirect addressing. For the subroutines c_redist_22_inv and c_redist_22, their time cost was reduced by almost a factor of 2 after removing the indirect addressing.
- Indirect addressing is at the core of the application and a clean, well-engineered solution to this problem would be very worthwhile.

GS2 now has an interface to FFTW3 but there is little performance benefit from using it without the further work highlighted from this project, which is the need for improvement to the data redistribution and method of addressing. Between October 2009 and November 2010, the Phase 2a central versions of GS2 consumed 6M AUs and Trinity (which contains the core of GS2) 7.8M AUs.

### 6.7.12 Scalability Optimization for Large-scale in-silico Simulations of Cardiac Bioelectric Activity

The Cardiac Arrhythmia Research Package (CARP) is a widely-used piece of software designed for large-scale simulations. These are to provide detailed personalised therapies for treatment of medical conditions such as cardiac arrhythmia and involuntary fibrillation. CARP has been developed by Dr. Gernot Plank from the Oxford e-Research Centre at the University of Oxford, who supervised this project. CARP is based on a finite element model and uses the PETSc package.

The project was allocated 8 months support at the September 2009 round. Lawrence Mitchell from EPCC performed the work between December 2009 and July 2010. The aims were to improve the key areas which are commonly the source of scalability issues in irregular parallel data decompositions, namely:

- Performance optimisations on the parallel mesh decomposition scheme. Asynchronous parallel output was found to give improvements of around 250% for small systems, rising to 1800% for large simulations.
- PETSc linear solver performance and preconditioning.
- Optimal data locality optimisation for the ODE solver.
- Non-blocking communications to improve the I/O routines.

As a result of this work, the scientific impact is that a mono-domain simulation of a human heartbeat (1 second activity) can now be performed in under 5 minutes wall-clock time using 16384 cores on Phase 2a. Before this work began, the same simulation would have taken around 75 minutes to complete using 1024 cores. Studies can now be completed within a realistic turnaround time.

### 6.7.13 Improving scalability of CP2K on multicore systems

The aim of this project was to implement mixed-mode OpenMP/MPI parallelism in the Density Functional Theory code CP2K. It was allocated 6 months support at the June 2009 round and follows on from the previous successful CP2K project. Dr. Ben Slater of the Department of Chemistry at UCL supervised the work of Iain Bethune from EPCC between October 2009 and August 2010.

The overall outcome of the work is that performance gains show improved scalability of up to a factor of 8 for a small benchmark, and a larger, inhomogeneous benchmark was shown to scale up to over 9000 Phase 2b cores. An increase in peak performance of up to 60% was also realised. The following was noticed:

- The scalability of the code decreases with an increasing number of cores per node, with the maximum performance of the pure MPI code being achieved on 256, 144, and 144 cores.
- The performance of the MPI-only code has improved 40-70% at around 1000 cores. Although this has not allowed the MPI code to scale any further, it will help improve performance at higher core counts for larger problems.
- Using threads does help to improve code scalability. Suitable numbers of threads to use are between 2 and 6, depending on the balance between performance for low core counts, and the desired scalability. The overall peak performance of the code has been increased by about 30% on HECToR Phase 2a and by 60% on HECToR Phase 2b, due to the fact that it reduces the number of messages being sent through the interconnect.

Between October 2009 and November 2010, the centrally installed version of CP2K consumed 43M AUs on Phase2a and 102M AUs on Phase2b. Users of this module are able to benefit from these DCSE improvements.

### 6.7.14 Performance enhancements for the GLOMAP aerosol model Part 2

GLOMAP MODE MPI is an aerosol simulation code. This project was proposed by Dr. Graham Mann of the School of Earth and Environment at the University of Leeds and was allocated 4 months of support at the March 2009 round. Mark Richardson of NAG performed the work between August 2009 and April 2010. The aim of the project was to improve the multicore performance of the code.
Outcomes:

- GLOMAP MODE has been converted to hybrid OpenMP/MPI and tested on Phase 2b.
- The performance of pure MPI production code on Phase 2b has improved 30% over that on Phase 2a.
- The OpenMP allows the code to use more cores per MPI task.
- Reducing the number of MPI processes maintains good scalability.
- OpenMP is only as effective as the loop count.

The centrally tracked version of GLOMAP used 7M AUs on Phase 2a between October 2009 and November 2010. These users will benefit from the DCSE improvements.

### 6.7.15 SPRINTing with HECToR

SPRINT is an add-on to the R statistical package. SPRINT allows users to write C code to manipulate R objects directly, for more computationally intensive tasks. SPRINT facilitates easy access to the application of HPC for statistical analysis. R is used worldwide and has many contributors, and is available on HECToR as a centrally installed module.
This project was proposed by Mr. Terry Sloan of EPCC who works in close collaboration with the Division of Pathway Medicine (DPM) at the University of Edinburgh to develop SPRINT. The project was allocated 6 months support at the June 2009 round and Savvas Petrou also performed the work at EPCC between October 2009 and March 2010. The main objective of the project was to port two widely used parallelised statistical functions of SPRINT to HECToR.
The outcome is that the parallelised correlation function *pcor* and permutation testing function *pmaxT* were both successfully implemented and scaling to 512 cores on Phase 2a was demonstrated.

SPRINT is used extensively by the Division of Pathway Medicine in order to enable genomic analysis on large data sets through the use of parallel processing.

### 6.7.16 Cloud and Aerosol Research on Massively-parallel Architectures

The overall aim of this DCSE was to improve the computational efficiency and operation of two key numerical cloud models that are of significant importance to the UK atmospheric science community. These models work in synergy to improve our understanding of the impacts of micro-scale processes on the evolution and life-cycle of clouds.
The project was allocated 12 months support at the January 2008 round. Dr. Paul Connolly of the School of Earth, Atmospheric and Environmental Sciences at the University of Manchester supervised the work and Jon Gibson from NAG performed the development between July 2008 and June 2010.
Outcomes:

- A major improvement to the pressure solver in the Met Office Large Eddy Model (LEM) has been implemented by replacing the original Fast Fourier Transform method, which required periodic boundary conditions, with the Jacobi-preconditioned bi-conjugate gradient method.
- This newly implemented iterative solver has been fully tested with a one-dimensional data decomposition and validated from known solutions.

### 6.7.17 Porting and Optimisation of Code_Saturne® on HECToR

Code_Saturne® is a versatile, open-source finite volume CFD code which was used widely on HPCx. It can be used to solve the Navier-Stokes equations on arbitrary unstructured grids. This project was allocated 18 months effort at the April 2008 call. Professor David Emerson of Daresbury supervised the work. Charles Moulinec of Daresbury began working on the project in November 2008, and in March 2009 Zhi Shang took over. The project finished at the end of October 2010.
The objectives of this work were:

- Improvement of the parallelisation for the post-processing and restarting process.
- Implementation of enhanced parallel partitioning which will enable a full simulation on HECToR, i.e. pre/post-processing and running the solver all in parallel.
- Optimisation of Code_Saturne® for multicore platforms.
- Specific optimisation for the vector X2 machine.

Outcomes:

- Initially ParMETIS and PT-scotch were used for testing and data partitioning. However the Zoltan library has now been implemented to facilitate easier use of ParMETIS.
- Meshes of around 200 million cells for arbitrary geometries can now be used.
- The I/O for post-processing and restarting has been improved with MPI-IO.

Code_Saturne® has recently been installed as a central module on HECToR Phase 2b and is expected to be used to produce new scientific results as a result of this DCSE.

### 6.7.18 Improving the performance of GWW

The GWW code performs Density Functional Theory (DFT) calculations based on a localised Wannier method. GWW is built around the more widely known Quantum espresso (QE) code. This project was proposed by Professor Merlyne De Souza from the Department of

Electronic and Electrical Engineering at the University of Sheffield, and was allocated 16 months support at the March 2009 round.
The objectives of the work were to improve the performance of the all-to-all communications and the Fast Fourier Transforms (FFTs) within the QE code and also implement improved I/O within GWW.
Damien Casterman of the University of Sheffield worked from October 2009 for 12 months on the implementation of a 2-dimensional decomposition for the FFTs and the improvement of I/O for GWW. Improvements to the all-to-all data communication were carried out between October 2009 and April 2010 by Iain Bethune of EPCC. The 2-dimensional FFT implementation was not completed due to unforeseen code difficulties, however a detailed report was produced which could help for knowledge transfer in the area of multi-dimensional FFTs for DFT on HPC systems.

### 6.7.19 Micromagnetic modelling of naturally occurring magnetic mineral systems

The Micromag code is a finite-element based application which performs time-dependent magnetic hysteresis calculations for arbitrary magnetic materials. This project was proposed by Professor Wyn Williams of the School of Geosciences at the University of Edinburgh, and allocated 6 months support at the September 2009 round.
The objectives of the work were to implement MPI parallelism within the main data structures, a linear equation solver, and the I/O of a serial code. Chris Maynard from EPCC worked on the project from January 2010 to September 2010. The outcome of the project is a PETSc interface for the code, enabling the potential use of HECToR to investigate theoretical systems which cannot be visualised with experimental imaging techniques and microscopy.

## 6.8 Ongoing DCSE Projects

### 6.8.1 Parallel Algorithms for the Materials Modelling code CRYSTAL

The CRYSTAL code is unique in that it can be used to perform Hartree-Fock (HF), Density Functional Theory (DFT) or HF-DFT hybrid calculations with periodic boundary conditions. The code is under continuous development in order to extend its applicability to emerging problems in materials modelling and to ensure its efficient use on modern computer architectures.
This project was allocated 24 months effort at the January 2008 call. Leading the work is Professor Nic Harrison of UCL. Stanko Tomic from Daresbury started working full time on the project in February 2009 and is due to complete it in March 2011. The following has so far been achieved:

- The memory bottleneck in the code which is caused by a replicated data structure of the HF and density matrix has been removed by replacing those data structures with their irreducible representations. Results show speedups in multiples of problem size/(2 x symmetry of the system).
- Implementation of the new MPP_BLOCK command provides better control of the block size of distributed data. Reducing the block size from the default value of 96 to 64 or 32 achieves more than a 10% speed-up in the diagonalisation part of the code and a 20% speed-up in the back and similarity transform can be demonstrated on 3584 cores.

Work is in progress to:

- Implement coarse-grained parallelism.

- Develop memory optimisation and modularisation to include HF-DFT and GW calculations.

The improvements are essential to future applications of the code and will benefit a large user base on HECToR.

### 6.8.2 Parallelisation of CABARET

The CABARET (Compact Accurate Boundary Adjusting high Resolution Technique) code uses a linear advection scheme based on a modified non-dissipative and low-dispersive second-order upwind leapfrog method. The modification introduces separate conservation and flux variables creating an additional independent evolutionary variable, which preserves the small phase and amplitude error.

The aim of this project is to port the serial compressible Navier-Stokes CABARET to HECToR. It was proposed by Dr. Sergey Karabasov of the University of Cambridge Engineering Department and was allocated 12 months effort at the December 2008 round. Phil Ridley of NAG has been working on the project since March 2009 and is scheduled to complete the work by January 2011.

The outcome will be a hybrid OpenMP/MPI CABARET which is capable of dealing with more than 50M grid cells. Progress so far includes:

- Development of an automated domain decomposition method for the unstructured grid.
- Full testing and validation. Simulations can now be performed in a few minutes that would otherwise have taken several days.
- A test case of 51M cells retains 80% parallel efficiency on 1000 cores on Phase 2b.

Dr. Karabasov is now able to investigate problem sizes at least 100 times larger than previously possible. He has recently secured 7M AUs from the November Resource Allocation Panel to investigate jet flap noise, and a further HECToR resource of 21M AUs on Phase 2b and 21M AUs on Phase 3, for further investigation using CABARET as the main code.

### 6.8.3 Hybrid time-dependent density functional theory in CASTEP Part 2

This follow on CASTEP project was allocated 12 months effort at the December 2008 round. Dominik Jochym began working in March 2010 and is scheduled to finish in April 2011. The objectives are to:

- Implement the exchange-correlation response kernel for hybrid functionals. This will involve developing code to compute the exchange potential response to a response orbital. The current response code in CASTEP has the assumption of semi-local functionals built in at a deep level, so completely new subroutines in the *nlxc* module will be required.
- Implement the calculation of forces using the formalism of Hutter.
- Implement Born-Oppenheimer molecular dynamics addressing issues of stability and Hamiltonian conservation.
- Perform parallel testing and validation on the final code which will be capable of performing molecular dynamics of a molecular system in an excited state.

### 6.8.4 Implementation of established algorithms to extend HELIUM

The HELIUM code solves the time-dependent Schrodinger equation for a two-electron atom or ion exposed to intense laser light. The proposed work involves:

- code extensions for crossed laser fields and 2D Bessel transformations;
- hybrid MPI-OpenMP to resolve memory usage issues;
- transformation of HELIUM results from spherical to cylindrical coordinates; and
- an extension of the techniques to address atoms with more than 2 electrons.

This project was allocated 44 months effort at the December 2008 round. The work is supervised by Professor Ken Taylor of the Department of Applied Mathematics & Theoretical Physics, Queen's University Belfast (QUB). Laura Moore (QUB) began working on the project in April 2009, and subsequently Michael Lysaght took over. Their work has focused upon improving the Taylor and Arnoldi propagators. Jonathan Parker (QUB) started working on the other half of the project in October 2009. Dr. Parker's work concentrates on extensions of HELIUM to allow the calculation of crossed laser fields, post-processing improvements in I/O, and the addition of OpenMP.
Dr. Parker and Dr. Lysaght are scheduled to complete this project by the end of March 2011. The overall outcome will enable the HELIUM code to maintain the ability to produce world-class science and bring further possibility of new science to HECToR.

### 6.8.5 Direct Numerical Simulations (DNS) of turbulent fluid flows

The Incompact3d code is based on an implicit finite difference solver which can handle flow which passes through fractal generated geometries. This project was approved at the December 2008 round for 16 months of effort. Professor J. C. Vassilicos of the Department of Aeronautics at Imperial College is supervising the project, on which Ning Li (NAG) has been working since March 2009, and is scheduled to complete the work by January 2011. The aim of the project is to improve the scalability of Incompact3d on HECToR through the use of hybrid OpenMP/MPI, by implementing an improved data decomposition for the FFT solver.
The outcome of this work has improved the performance and scalability of the code so that it can handle computational domains with excellent efficiency on 24000 Phase 2b cores. The 2DECOMP&FFT library[1] for multi-dimensional FFT decompositions has also been developed by Dr. Li. This Fortran library performs two major tasks: a 2D domain decomposition algorithm for applications using 3D Cartesian data structures; and an efficient FFT interface to perform three-dimensional FFTs in parallel. The library is optimised for large-scale computations on supercomputers and scales well to tens of thousands of processors.
This project has enabled HECToR users of Incompact3d to save at least 20% of their AU allocation and given them the possibility of investigating more realistic, higher Reynolds numbers flows within a reasonable turnaround time. Between October 2009 and November 2010, Incompact3d used over 13M AUs on Phase 2a.

### 6.8.6 Dynamic load balancing and rigid-body dynamics for DL_POLY_3

The aim of this work is to add functionality to DL_POLY_3 which is already available in DL_POLY_2, since DL_POLY_3 is more efficient. The outcome of this project will be to achieve a scalable version of DL_POLY capable of rigid-body dynamics, to maintain the user base of the code on HECToR and to attract more users. This project was proposed by Dr. Ilian Todorov of Daresbury and was allocated 18 months of effort at the March 2009 round. Lawrence Ellison of Daresbury started working on the project in October 2009 and is scheduled to complete the work in March 2011.

---

[1] http://www.hector.ac.uk/cse/distributedcse/reports/incompact3d/UserGuide.html

### 6.8.7 Multicore performance and domain choice for DL_POLY_3

There are two objectives to this project:

- Multicore improvements to enable DL_POLY_3 to perform as well on a Phase 2 quad-core processor as it did on a dual-core processor on Phase 1.
- Enable the code to use FFT grids whose prime factorisation is of the form $2^{n}*3^{(0\text{-}2)}*5^{(0\text{-}1)}$ (i.e. it will allow the prime factors of the order to contain 2 to any power, 3 up to order 2, and 5 up to order 1) rather than $2^{n}$.

This was allocated 10 months of support at the September 2009 round of funding. Valène Pellissier (NAG) implemented the multicore improvements between January and September 2010, and this work has enabled improved scalability for DL_POLY_3 on Phase 2a. Ian Bush (NAG) began work on the FFT grid enhancements in May 2010 and is due to complete this work March 2011. This project is supervised by Dr. Ilian Todorov of Daresbury.

### 6.8.8 Adding molecular dynamics functionality to CASINO

The objective of this work is to effectively couple the Quantum Monte Carlo (QMC) CASINO code to the molecular dynamics (MD) code PWSCF. The result will be a scalable model for use on HECToR which will enable new scientific simulations to be carried out within the fields of both MD and QMC.
The previous CASINO DCSE, performed by Lucian Anton (NAG) and supervised by Professor Dario Alfè, demonstrated excellent performance improvements for shared memory parallelism. To build on this success, this follow-on project was allocated 18 months support at the March 2009 round. Norbert Nemec, who was based at UCL, started working on the project at the beginning of January 2010, but left to take up another post overseas in August 2010. Dr. Mike Towler, who is the main developer of CASINO, continued this work with Professor Alfè from October 2010 and is scheduled to complete the project in July 2011.

### 6.8.9 Optimising the performance of the VASP code

The VASP code is a widely used application on HECToR for performing ab-initio quantum-mechanical molecular dynamics (MD) and has similar functionality to CASTEP. This project was allocated 12 months support at the September 2009 round.
There are two objectives to this project:

- Provide an optimal set of collective-communication routines for VASP by accumulating message data within a node by using explicit shared-memory segments to improve the multicore scalability when using up to 256 processing cores.
- Implement parallelisation over k-points (as used in other ab initio electronic structure codes) for VASP 5.2. This adds further parallelisation in addition to the existing capability of parallelising over bands and over plane waves.

Professor Richard Catlow of the Department of Chemistry at UCL is supervising the project which started in January 2010 with Mina Maniopoulou from NAG working on the k-point implementation. Andrew Turner from EPCC performed the single node optimisation for multicore between January and August 2010. Dr. Maniopoulou is scheduled to complete the k-point work in January 2011.
For the k-points parallelism, speed-ups of nearly 6 times are achievable on up to 288 Phase 2b cores for certain optimised test cases, and most users of VASP are readily able to achieve speed-ups of 3 times for their problems. During October 2009 to November 2010 the central module of VASP 5 used 160M AUs on Phase 2a and 24M AUs on Phase 2b.

### 6.8.10 Improving Load Balancing and Parallel Partitioning in ICOM

This project was allocated 12 months of support at the September 2009 round. Dr. Jon Hill of Imperial College London is leading the work with Paul Woodhams from NAG performing the implementation. The project began in November 2009 and is scheduled to complete October 2011. The main objective of this work is to improve the scalability of Fluidity-ICOM through the introduction of optimised load balancing. This will be achieved by incorporating the Zoltan library within the code as an alternative to METIS for the irregular domain decomposition.

### 6.8.11 Metal Conquest: efficient simulation of metals on petaflop computers

CONQUEST is a Density Functional Theory (DFT) code which can demonstrate linear scalability on up to 4096 Phase 2a cores when applied to the study of metallic insulators. However, in order to improve the scalability for other metallic material simulations there are several areas of development required. This project was allocated 12 months of support at the September 2009 round to implement the developments proposed by Dr. Andrew Horsfield of Imperial College.

Lianheng Tong of UCL began the work in February 2010 and is scheduled to complete in February 2011. The aims of the project are 1) to implement charge mixing and K-point sampling and 2) improve the ScaLAPACK optimisation. The outcome so far is that Pulay mixing, Kerker pre-conditioning, wave-dependent metric and Methfessel-Paxton smearing methods have been successfully implemented in CONQUEST. Test results show that there is a four times decrease in the number of self-consistency iterations required for convergence while maintaining good accuracy.

### 6.8.12 Implementation of a Divide and Conquer Strategy for the Materials Modelling code CRYSTAL

This second DCSE for CRYSTAL was recommended for 18 months support at the September 2009 round. Dr. Barbara Montanari of RAL is supervising the work. Daniel Jones (NAG) began the work in August 2010 and is scheduled to complete it in February 2012.

The objective of this work is to improve the efficiency of the point multi-pole calculations for parallel computation and to perform further optimisation of the code for multicore. The outcome will be of immediate benefit to the large community of CRYSTAL users on HECToR.

### 6.8.13 A New CASTEP and ONETEP Geometry Optimiser

This is the fourth CASTEP proposal which has been recommended for DCSE support. The developments from this work will, in addition, be applicable to the ONETEP (Order-N Electronic Total Energy Package) linear-scaling code for quantum-mechanical calculations which is also based on DFT. This project was proposed by the chair of the UK Car-Parrinello Consortium (UKCP), Dr. Matt Probert, who is based in the Department of Physics at the University of York.

The work was recommended for 12 months support at the September 2009 round. The aims of the project are to improve the geometry optimisation calculations in both codes by improved use of multicore architectures for matrix factorisation methods. It is anticipated that knowledge transfer from this project will be useful to other DFT codes. Jolyon Aarons (NAG) began the work in September 2010 and is scheduled to complete it in September 2011.

### 6.8.14 SPRINTing further with HECToR

This is a follow-on project from the previous SPRINT DCSE. The aim of this work is to implement two more widely used statistical routines within the SPRINT package, the Rank Product and Random Forest algorithms. The parallel version of SPRINT is used extensively by the Division of Pathway Medicine (DPM) at the University of Edinburgh, to solve problems in Biosciences.
The project was allocated 6 months support at the January 2010 round. Lawrence Mitchell from EPCC started the work October 2010 under the guidance of Mr. Terry Sloan (EPCC). The project is scheduled to complete in March 2011.

### 6.8.15 Developing the TELEMAC system for HECToR (phase 2b & beyond)

The TELEMAC-2D unstructured coastal modelling code is new to HECToR and this project targets shared memory improvement to the underlying BIEF finite element library by:

- Employing OpenMP within the linear systems.
- Employing OpenMP within the convection schemes.

The project was allocated 12 months support at the January 2010 round. Professor David Emerson of Daresbury is supervising the work which Zhi Shang, also of Daresbury, started in October 2010 and is scheduled to complete in September 2011.

### 6.8.16 HECToR enabled Step Change in Turbulent Multiphase Combustion Simulations

This project will enable the Direct Simulation of Turbulence and Combustion (DSTAR) code to conduct the world's largest direct numerical simulation of turbulent multiphase combustion. It was allocated 6 months support at the June 2010 round and started in September 2010. The project was proposed by Professor Kai Luo of the University of Southampton, with Lucian Anton (NAG) working on the project until March 2011. The following objectives will allow DSTAR to efficiently utilise more than12000 cores on Phase 2b:

- Implementation of a 2D domain decomposition algorithm.
- Improved I/O for capability size data sets.
- Refactoring legacy Fortran code.

### 6.8.17 Microstructurally Faithful Modelling Of Materials

ParaFEM is a portable library of subroutines for parallel finite element analysis. This project was allocated 9 months support at the June 2010 round and started in December 2010. It was proposed by Dr. Lee Margetts of the University of Manchester with Ms. Louise Lever, also from the University of Manchester, working on the project until November 2011. The objectives are:

- Linear elastic, Strain hardening and Large strain User MATerial implementations tested with ParaFEM and Abaqus.
- Dynamic load balancing with User MATerial implementations.
- Interface the existing ParaFEM-Viewer application to the Massive Remote Batch Visualizer.

### 6.8.18 CP2K - Sparse Linear Algebra on 1000s of cores

This project will improve the sparse matrix (DBCSR) library performance in CP2K. It was allocated 6 months support at the June 2010 round and started in October 2010. It is the third CP2K DCSE to be proposed by Dr. Ben Slater of UCL, with Iain Bethune of EPCC working on the project until December 2011. The objectives are:

- Implementation of a reordered 2D process distribution, automatic message type selection, and row/column communicators in DBCSR.
- Improved load balancing and on-node performance for the key multiplication routines.
- Implementation of automatic selection of optimal communication method and higher OpenMP efficiency.

### 6.8.19 Developing hybrid OpenMP/MPI parallelism for Fluidity/ICOM next generation geophysical fluid modelling technology

The aim of this project is to enhance the multicore performance of Fluidity-ICOM, allowing it to utilise 16384 Phase 2b cores. It was allocated 12 months support at the June 2010 round and started in October 2010. It was proposed by Dr. Andrew Sunderland of Daresbury with Xiaohu Guo working on the project until September 2012. The objectives are:
Implementation of MPI/OpenMP mixed-mode parallelisation of the finite element assembly stage in Fluidity-ICOM.
Optimization of the HYPRE Library usage for linear pre-conditioners/solvers for large core counts.
Benchmarking and code re-engineering for hybrid mesh adaptivity.
Optimisation of VASP Density Functional Theory (DFT)/Hartree-Fock (HF) hybrid functional code using mixed-mode parallelism
This project will enhance the hybrid-DFT performance in VASP 5 enabling the application to utilise more than 256 Phase 2b cores in an efficient manner. It was allocated 12 months support at the June 2010 round and started in November 2010. It was proposed by Dr. Scott Woodley of UCL with Andy Turner of EPCC working on the project until October 2012. The objectives are to add OpenMP parallelisation to PAW routines, implement hybrid DFT routines, and address any remaining bottlenecks in VASP.

## 6.9 New DCSE Projects Starting 2011

### 6.9.1 Enhancing Conquest for accurate, scalable simulation of entire biological molecules

The aim of this project is to attract users from the Bioscience community onto HECToR by adding new functionality to CONQUEST: van der Waals Density Function theory and spin polarisation. It was allocated 12 months support at the June 2010 round and was proposed by Dr. David Bowler of UCL. Lianheng Tong (UCL) will be working on the project from February 2011 until January 2012. The objectives are:

- Implementation of van der Waals Density Function theory for O(N) scaling on Phase 2b.
- Working implementation of spin polarisation.

### 6.9.2 Developing NEMO for Large Multi-core Scalar Systems

This project will support the Oceans 2025 UK marine research programme by targeting three aspects of NEMO code development. It was allocated 12 months support at the June 2010 round and was proposed by Dr. Stephen Pickles of Daresbury. Dr. Pickles will work on the project throughout 2011 and Andrew Porter will join the project in January 2012. The objectives are:

- Provision of an ensemble capability in NEMO.
- Multicore-aware partitioning and halo-exchange optimisation.
- Array index re-ordering.

### 6.9.3 Combined-Multicore Parallelism for the UK electron-atom scattering Inner Region R-matrix codes on HECToR

Following on from the success of the previous DCSE project which targeted the Outer Region R-matrix code, PRMAT, this work will enhance the Inner Region codes for use within the UK-RAMP consortium. It was allocated 12 months support at the June 2010 round and was proposed by Dr. Martin Plummer of Daresbury who will work on the project from April 2011 until the end of March 2012. The common objectives are to upgrade the RAD, ANG and HAM(PDG) codes for improved OpenMP and task management. Specific objectives are:

- Enhanced continuum function generation in RAD.
- Addition of the *xstream* parallel I/O library to ANG.
- Improved dipole matrix calculations in HAM.

### 6.9.4 Boosting the scaling performance of CASTEP: enabling next generation HPC for next generation science

The aim of this work is to improve CASTEP for Phase 2b and other multicore architectures. It is the fifth CASTEP related DCSE project and was allocated 10 months support at the June 2010 round. The work will be supervised by Dr. Keith Refson of RAL. Dominik Jochym, also of RAL, will work on the project full time from April 2011 until the end of January 2012. The objectives are;

- MPI Buffer memory and I/O optimisation.
- Produce a report on parallel efficiency.
- Improved band parallelism for multicore.
- Use ScaLAPACK for orthonormalisation and subspace diagonalisation.

# Appendix A: Terminology

**TSL** : Threshold Service Level

**FSL** : Full Service Level

**SDT** : Scheduled Down Time

**UDT** : Unscheduled Down Time

**WCT** : Wall Clock Time

**MTBF** : Mean Time Between Failures = 732/Number of Failures

**SP** : Service Provision

**SP Serviceability%** = 100*(WCT-SDT-UDT(SP))/(WCT-SDT)

**Technology Reliability %** = 100*(1-(UDT(Technology)/(WCT-SDT))

## Incident Severity Levels

**SEV 1** — anything that comprises a FAILURE as defined in the contract with EPSRC.

**SEV 2** — NON-FATAL incidents that typically cause immediate termination of a user application, but not the entire user service.

The service may be so degraded (or liable to collapse completely) that a controlled, but unplanned (and often very short-notice) shutdown is required or unplanned downtime subsequent to the next planned reload is necessary.

This category includes unrecovered disc errors where damage to file systems may occur if the service was allowed to continue in operation; incidents when although the service can continue in operation in a degraded state until the next reload, downtime at less than 24 hours' notice is required to fix or investigate the problem; and incidents whereby the throughput of user work is affected (typically by the unrecovered disabling of a portion of the system) even though no subsequent unplanned downtime results.

**SEV 3** — NON-FATAL incidents that typically cause immediate termination of a user application, but the service is able to continue in operation until the next planned reload or re-configuration.

**SEV 4** — NON-FATAL recoverable incidents that typically include the loss of a storage device, or a peripheral component, but the service is able to continue in operation largely unaffected, and typically the component may be replaced without any future loss of service.

# Appendix B: Projects on HECToR

| Code | Project Title | Funding Body | Class | Principal Investigator | AUs allocated | AUs used | AUs left |
|---|---|---|---|---|---|---|---|
| **EPSRC Projects** | | | | | | | |
| c01 | Support of EPSRC/STFC SLA | EPSRC | Class1a | Dr. Richard Blake | 30,803,723 | 27,861,147 | 2,942,576 |
| e01 | UK Turbulence Consortium | EPSRC | Class1a | Dr. Gary N Coleman | 483,969,876 | 31,356,628 | 452,613,248 |
| e05 | Materials Chemistry HPC Consortium | EPSRC | Class1a | Prof C Richard A Catlow | 1,139,124,000 | 122,237,924 | 1,016,826,076 |
| e10 | GENIUS | EPSRC | Class1a | Prof Peter Coveney | 10,248,188 | 7,104,643 | 3,143,545 |
| e24 | DEISA | EPSRC | Class1a | Mrs. Alison Kennedy | 233,146,943 | 126,052,911 | 107,094,032 |
| e35 | Non-adiabatic processes | EPSRC | Class1a | Dr. Tchavdar Todorov | 12,246,862 | 4,164,267 | 8,082,595 |
| e42 | Computational Combustion for Engineering Applications | EPSRC | Class1a | Prof Kai Luo | 32,000,001 | 30,021,151 | 1,978,850 |
| e59 | Turbulence in Breaking Gravity Waves | EPSRC | Class1a | Prof Ian P Castro | 708,922 | 444,128 | 264,794 |
| e63 | UK Applied Aerodynamics Consortium 2 | EPSRC | Class1a | Dr. Nick Hills | 30,925,323 | 25,925,815 | 4,999,508 |
| e68 | Hydrogenation Reactions at Metal Surfaces | EPSRC | Class1a | Prof. Angelos Michaelides | 50,000,000 | 44,023,414 | 5,976,586 |
| e70 | Computation of Electron Transfer Properties | EPSRC | Class1a | Dr. Jochen Blumberger | 1,160,000 | 1,154,405 | 5,595 |
| e71 | Simulating the control of calcite crystallisation | EPSRC | Class1a | Prof John Harding | 130,403,522 | 42,731,053 | 87,672,469 |
| e76 | HELIUM Developments | EPSRC | Class1a | Prof Ken Taylor | 42,521,798 | 33,457,247 | 9,064,551 |
| e82 | ONETEP: linear-scaling method on High Performance Computers | EPSRC | Class1b | Dr. Peter Haynes | 1,105,352 | 682,588 | 422,764 |

| Code | Project Title | Funding Body | Class | Principal Investigator | AUs allocated | AUs used | AUs left |
|---|---|---|---|---|---|---|---|
| e84 | Vortical Mode Interactions | EPSRC | Class1a | Dr. Tamer Zaki | 9,600,000 | 2,808,979 | 6,791,021 |
| e85 | Study of Interacting Turbulent Flames | EPSRC | Class1a | Dr. N Swaminathan | 8,088,610 | 2,122,745 | 5,965,865 |
| e89 | Support for UK Car-Parrinello Consortium | EPSRC | Class1a | Dr. Matt Probert | 360,100,001 | 161,367,725 | 198,732,276 |
| e92 | Dynamo Action In Compressible Convection | EPSRC | Class1a | Mr. Paul Bushby | 4,075,000 | 2,073,069 | 2,001,931 |
| e102 | Numerical investigation of aerofoil noise | EPSRC | Class1a | Dr. Richard D Sandberg | 6,484,191 | 6,084,516 | 399,675 |
| e104 | Fluid-Mechanical Models applied to Heart Failure | EPSRC | Class1a | Dr. Nicolas Smiths | 30,400,000 | 2,678,248 | 27,721,752 |
| e105 | Joint Euler/Lagrange Method for Multi-Scale Problems | EPSRC | Class1a | Dr. Andreas M Kempf | 1,300,000 | 297,323 | 1,002,677 |
| e106 | Numerical Simulation of Multiphase Flow: From Mesocales to | EPSRC | Class1a | Prof Kai Luo | 3,650,000 | 0 | 3,650,000 |
| e107 | Parallel Brain Surgery Simulation | EPSRC | Class1a | Dr Stephane P. A. Bordas | 6,000,000 | 338,197 | 5,661,803 |
| e108 | Unsteady Propeller Noise | EPSRC | Class1b | Dr. Sergey Karabasov | 7,684,524 | 158,100 | 7,526,424 |
| e110 | Computational Aeroacoustics Consortium | EPSRC | Class1a | Prof Paul Tucker | 39,100,000 | 33,090,759 | 6,009,241 |
| e117 | Biosurfactant via molecular dynamics simulations | EPSRC | Class1b | Dr. Carmen Domene | 18,889,068 | 12,289,508 | 6,599,560 |
| e120 | [dCSE] FF Transformations for plasma simulations | EPSRC | Class2b | Dr. Colin M Roach | 200,000 | 170,213 | 29,787 |
| e121 | [dCSE] Improving Performance using Wannier functions | EPSRC | Class1a | Prof Maria Merlyne DeSouza | 2,680,305 | 2,299,591 | 380,714 |
| e122 | Multiscale Modelling of Magnetised Plasma Turbulence | EPSRC | Class1a | Dr. Colin M Roach | 65,000,000 | 19,901,902 | 45,098,098 |
| e124 | Compressible Axisymmetric Flows | EPSRC | Class1a | Dr. Richard D Sandberg | 22,887,943 | 6,248,060 | 16,639,883 |
| e125 | Full configuration interaction quantum monte carlo | EPSRC | Class1a | Dr. Ali Alavi | 18,324,825 | 3,572,645 | 14,752,180 |

| Code | Project Title | Funding Body | Class | Principal Investigator | AUs allocated | AUs used | AUs left |
|---|---|---|---|---|---|---|---|
| e126 | Clean Coal Combustion: Burning Issues of Syngas Burning | EPSRC | Class1a | Prof Xi Jiang | 9,984,000 | 4,214,261 | 5,769,739 |
| e127 | Alternative drag-reduction strategies | EPSRC | Class1a | Prof Michael Leschziner | 7,000,000 | 17,041 | 6,982,959 |
| e128 | Rate-Controlled Constrained Equilibrium | EPSRC | Class1a | Dr. Stelios Rigopoulos | 6,230,000 | 0 | 6,230,000 |
| e129 | Novel Hybrid LES-RANS schemes [ICL] | EPSRC | Class1a | Prof Michael Leschziner | 7,500,000 | 599,443 | 6,900,557 |
| e130 | Novel hybrid LES-RANS schemes [MAN] | EPSRC | Class1a | Prof Dominique Laurence | 10,500,000 | 0 | 10,500,000 |
| e133 | Implementation of Established Algorithms to Extend HELIUM | EPSRC | Class2b | Prof Ken Taylor | 800,000 | 0 | 800,000 |
| e135 | DNS of unsteady turbulent flow over a smooth or a rough surface | EPSRC | Class2a | Dr. Shuisheng He | 204,000 | 174,763 | 29,237 |
| e136 | Modelling the UK Wind Power Resource | EPSRC | Class1b | Dr. Gareth Harrison | 5,679,268 | 4,959,825 | 719,443 |
| e141 | A numerical study of turbulent manoeuvering-body wakes | EPSRC | Class1a | Dr. Gary N Coleman | 16,350,000 | 166,810 | 16,183,190 |
| e143 | Numerical Investigation of Jet Noise | EPSRC | Class1a | Dr. Anurag Agarwal | 2 | 0 | 2 |
| e144 | Numerical Simulation of Rotating Stall and Surge | EPSRC | Class1a | Dr. Mehdi Vahdati | 1,266,001 | 24 | 1,265,977 |
| e145 | UK-SHEC Consortium | EPSRC | Class1a | Dr. T.J. Mays | 1,191,899 | 270,364 | 921,535 |
| e146 | G protein-coupled receptor dynamics | EPSRC | Class2a | Dr. Irina Tikhonova | 199,680 | 3,276 | 196,404 |
| e147 | Scale adaptive simulations of turbulent flows | EPSRC | Class2a | Prof Oubay Hassan | 243,495 | 243,221 | 274 |
| e148 | Adding the molecular dynamics functionality to the quantum | EPSRC | Class2b | Prof Dario Alfe` | 638,951 | 263,691 | 375,260 |
| e149 | Fractal-generated turbulence and mixing: flow physics and | EPSRC | Class1a | Prof Christos Vassilicos | 51,920,000 | 3,464,960 | 48,455,040 |
| e155 | Modelling Cholesterol Deposits | EPSRC | Class1a | Dr. David Quigley | 10,000,000 | 0 | 10,000,000 |

| Code | Project Title | Funding Body | Class | Principal Investigator | AUs allocated | AUs used | AUs left |
|---|---|---|---|---|---|---|---|
| e156 | Metal Conquest: efficient simulation of metals on petaflop | EPSRC | Class2b | Dr. David Bowler | 1,600,000 | 1,063 | 1,598,937 |
| e157 | Global stability computations of separated flows | EPSRC | Class2a | Prof. Jitesh S B Gajjar | 299,996 | 98 | 299,898 |
| e158 | Novel Asynchronous Algorithms | EPSRC | Class1a | Prof. Nicholas J Higham | 500,000 | 0 | 500,000 |
| e159 | Multi-layered Abstractions for PDEs | EPSRC | Class1a | Prof. Paul Kelly | 3,816,000 | 0 | 3,816,000 |
| e160 | Sustainable Software Generation Tools | EPSRC | Class1a | Prof. Paul Kelly | 20,208,060 | 0 | 20,208,060 |
| e161 | Properties and Dynamics of Atomic Bose-Einstein Condensates | EPSRC | Class1a | Dr. A White | 69,895,466 | 0 | 69,895,466 |
| e165 | Multi-scale simulation of intense laser plasma interactions | EPSRC | Class1a | Dr. Tony Arber | 4,872,000 | 0 | 4,872,000 |
| e166 | Large Eddy Simulation of LNG Pool Fires | EPSRC | Class2a | Dr. Siaka Dembele | 300,000 | 286,252 | 13,748 |
| e167 | LES of supersonic jets | EPSRC | Class1b | Prof. William Dawes | 2,696,000 | 597,293 | 2,098,707 |
| e170 | CFD Simulations of the BLOODHOUND SuperSonic Car | EPSRC | Class1b | Dr. Ken Morgan | 1,935,360 | 256,728 | 1,678,632 |
| e171 | Conformational switching of tetra-(bromophenyl) porphyrins | EPSRC | Class1b | Prof. Mats Persson | 3,289,521 | 1,734,321 | 1,555,200 |
| e173 | Performance of oomph-lib in largescale parallel computations | EPSRC | Class1b | Prof. Matthias Heil | 4,800,000 | 49,092 | 4,750,908 |
| e174 | 3D instabilities in two-layer flows | EPSRC | Class2a | Dr. Prashant Valluri | 701,899 | 441,024 | 260,875 |
| e175 | Fine-Scale Turbulence | EPSRC | Class1a | Dr. Richard D Sandberg | 50,000,000 | 0 | 50,000,000 |
| e176 | Structure refinement of nanomaterials | EPSRC | Class2a | Prof. Peter G Bruce | 300,000 | 0 | 300,000 |
| e177 | Amorphous structures of mirror coatings | EPSRC | Class2a | Dr .Ian Maclaren | 300,000 | 0 | 300,000 |
| e178 | Conformational changes in macromolecules | EPSRC | Class2a | Dr. Philip Biggin | 300,000 | 208,714 | 91,286 |

| Code | Project Title | Funding Body | Class | Principal Investigator | AUs allocated | AUs used | AUs left |
|---|---|---|---|---|---|---|---|
| e179 | Non-conservative dynamics | EPSRC | Class1a | Dr Daniel Dundas | 87,000,000 | 0 | 87,000,000 |
| e182 | Advanced Modelling of Two-Phase Reacting Flow | EPSRC | Class1a | Dr Edward S Richardson | 8,150,164 | 0 | 8,150,164 |
| e183 | Analysis of Processes in Hydrocarbon Fuel Droplets | EPSRC | Class1a | Prof Sergei Sazhin | 8,640,000 | 0 | 8,640,000 |
| e184 | UK-RAMP | EPSRC | Class1a | Prof Ken Taylor | 130,500,000 | 0 | 130,500,000 |
| e185 | Chemistry of ceramic materials | EPSRC | Class1a | Prof John Harding | 340,000,000 | 3,840 | 339,996,160 |
| e186 | Step Change in Combustion Simulation | EPSRC | Class1a | Prof Kai Luo | 40,000,000 | 105,722 | 39,894,278 |
| e187 | IAGP: Integrated Assessment of Geoengineering Proposals | EPSRC | Class1a | Prof Piers Fosters | 16,030,170 | 0 | 16,030,170 |
| e188 | Hydrostatic compression of energetic materials | EPSRC | Class1b | Prof Colin Pulham | 2,000,000 | 246,199 | 1,753,801 |
| e189 | Towards Biomimetic Nanopores | EPSRC | Class1b | Mr Mark M Sansom | 4,566,240 | 3,353 | 4,562,887 |
| e19 | Edinburgh Soft Matter and Statistical Physics Group | EPSRC | Class1a | Prof Michael Cates | 4,663 | 6,171 | -1,508 |
| e190 | Design of helicopter rotors | EPSRC | Class1b | Dr George N Barakos | 4,500,000 | 1,193,384 | 3,306,616 |
| e191 | CFD Analysis of Flight Dynamics | EPSRC | Class1b | Prof Kenneth Badcock | 3,500,000 | 75,378 | 3,424,622 |
| e192 | Physical properties of carbon nanotubes | EPSRC | Class1b | Dr Michael R C Hunt | 2,534,400 | 0 | 2,534,400 |
| e193 | Colloids in Cholesteric Liquid Crystals | EPSRC | Class1b | Dr Davide Marenduzzo | 12,500,000 | 0 | 12,500,000 |
| e194 | Direct Numerical Simulation of Meso-scale Combustor | EPSRC | Class1b | Dr N Swaminathan | 3,701,520 | 0 | 3,701,520 |
| e195 | Sensitivity study of turbulence model | EPSRC | Class1b | Dr Alistair Revell | 5,279,040 | 4,618 | 5,274,422 |
| e196 | Structure of intrinsically disordered proteins | EPSRC | Class1b | Dr Robert Best | 3,600,000 | 3 | 3,599,997 |

| Code | Project Title | Funding Body | Class | Principal Investigator | AUs allocated | AUs used | AUs left |
|---|---|---|---|---|---|---|---|
| e197 | Potassium on graphite | EPSRC | Class2a | Dr Kai Hock | 300,000 | 0 | 300,000 |
| e198 | Numerical Studies of Droplets | EPSRC | Class2a | Dr Kensuke Yokoi | 300,000 | 0 | 300,000 |
| e199 | Microstructurally Faithful Modelling of Materials | EPSRC | Class2b | Dr Lee Margetts | 800,000 | 230 | 799,770 |
| y08 | Testing | EPSRC | Early use | Dr David Jenkins | 1,000 | 0 | 1,000 |
| **STFC Projects** | | | | | | | |
| p01 | Atomic Physics for APARC | STFC | Class1a | Dr Penny Scott | 3,020,000 | 557,149 | 2,462,851 |
| **NERC Projects** | | | | | | | |
| n01 | Global Ocean Modelling Consortium | NERC | Class1a | Dr Thomas Anderson | 89,243,840 | 53,343,347 | 35,900,493 |
| n02 | NCAS (National Centre for Atmospheric Science) | NERC | Class1a | Dr Lois Steenman-Clark | 273,768,327 | 220,858,164 | 52,910,163 |
| n03 | Computational Mineral Physics Consortium | NERC | Class1a | Prof John P Brodholt | 284,142,416 | 224,198,697 | 59,943,719 |
| n04 | Shelf Seas Consortium | NERC | Class1a | Dr Roger Proctor | 88,202,935 | 62,311,628 | 25,891,307 |
| u01 | Melting of MgSiO3 Perovskite | NERC | Early use | Prof John P Brodholt | 11,000,000 | 11,018,423 | -18,423 |
| **BBSRC Projects** | | | | | | | |
| b08 | Int BioSim | BBSRC | Class1a | Mr Mark M Sansom | 866,000 | 909,998 | -43,998 |
| b09 | Circadian Clock | BBSRC | Class1a | Prof Andrew A Millar | 2,000,000 | 1,393,875 | 606,125 |
| b100 | Widening the BBSRC HPC User Base | BBSRC | Class1a | Dr Michael Ball | 10,000,000 | 632,465 | 9,367,535 |
| b12 | Flu Analysis on HECToR | BBSRC | Class1a | Mr Adrian Jackson | 50,000 | 0 | 50,000 |

| Code | Project Title | Funding Body | Class | Principal Investigator | AUs allocated | AUs used | AUs left |
|---|---|---|---|---|---|---|---|
| b13 | HECToR Allocation | BBSRC | Class1a | Dr David Bowler | 5,587,200 | 0 | 5,587,200 |
| b10 | SPRINTing with HECToR [dCSE] | BBSRC | Class2b | Mr Terry Sloan | 795,120 | 61,996 | 733,124 |
| **Director's Time** | | | | | | | |
| d03 | EUFORIA | Directors Time | Service | Mr Adrian Jackson | 2,400,000 | 2,426,575 | -26,575 |
| d04 | MSc in HPC | Directors Time | Service | Dr David Henty | 343,500 | 204,567 | 138,933 |
| d11 | NAIS | Directors Time | Service | Prof Mark Ainsworth | 416,668 | 201,939 | 214,729 |
| d12 | CoE HiGEM | Directors Time | Service | Dr Len L C Shaffrey | 10,000,000 | 0 | 10,000,000 |
| d13 | CoE SENGA | Directors Time | Service | Dr Stewart Cant | 10,000,000 | 0 | 10,000,000 |
| d14 | CoE HiPSTAR | Directors Time | Service | Dr Richard D Sandberg | 2,000,000 | 2,264,689 | -264,689 |
| d15 | HPC-GAP | Directors Time | Service | Dr David Henty | 2,033 | 1,038 | 995 |
| d16 | ETC | Directors Time | Service | Dr Lorna Smith | 501,000 | 154,897 | 346,103 |
| d17 | CAF Optimisation | Directors Time | Service | Dr Jason Beech-Brandt | 2 | 0 | 2 |
| d18 | FireGrid HPC | Directors Time | Service | Prof Arthur S Trew | 600,001 | 251,409 | 348,592 |
| d19 | OpenFOAM Demo | Directors Time | Service | Dr Alan Gray | 950,000 | 396,957 | 553,043 |
| d20 | CSCS | Directors Time | Service | Dr Alan Gray | 50,000 | 0 | 50,000 |
| d21 | GADGET | Directors Time | Service | Dr Adrian Jenkins | 1,000,001 | 18,584 | 981,417 |
| d22 | Summer Science Exhibition | Directors Time | Service | Prof. Angelos Michaelides | 70,000 | 35,572 | 34,428 |

| Code | Project Title | Funding Body | Class | Principal Investigator | AUs allocated | AUs used | AUs left |
|---|---|---|---|---|---|---|---|
| d23 | TEXT FP7 | Directors Time | Service | Dr Mark Bull | 1,500,000 | 15,441 | 1,484,559 |
| d24 | SBSI | Directors Time | Service | Dr Stephen Gilmore | 2,000,000 | 957,651 | 1,042,349 |
| d25 | Code Scaling | Directors Time | Service | Dr Ken Rice | 1,500,000 | 1,098,949 | 401,051 |
| y09 | Director's Time | Directors Time | Service | Prof Arthur S Trew | 29,685,133 | 82,538 | 764,170 |
| **External Projects** | | | | | | | |
| t01 | NIMES: New Improved Muds from Environmental Sources. | External | Class1a | Dr Chris Greenwell | 4,113,669 | 4,245,424 | -131,755 |
| x01 | HPC-Europa | External | Class1a | Dr Judy Hardy | 16,415,790 | 8,925,197 | 7,490,593 |
| x05 | FIOS | External | Class1a | Mr Davy Virdee | 1,130,100 | 1,074,930 | 55,170 |
| e168 | TEXT | External | Service | Dr Mark Bull | 1,500,000 | 0 | 1,500,000 |
| x02 | BlueArc (TDS) | External | Service | Mr M W Brown | 1,000 | 0 | 1,000 |
| x06 | Rhymney | External | Service | Dr Mark Sawyer | 4,500 | 16 | 4,484 |