Project Status

Document identifier: GridPP-PMB-147-ProjectStatus.doc

Date: 11/06/2010
Version: 1.0
Document status: Final
Author: PMB
Introduction

Since the last Oversight Committee meeting the LHC achieved proton-proton collisions at a centre of mass energy of 7 TeV on 30 March 2010. This signalled the start of the physics programme and the experiments have now entered a long-anticipated period of sustained data taking in 2010 and 2011. The Figure-1 below shows the rapid rise in luminosity delivered to the experiments over the initial period of the physics programme.

Figure-1: Integrated luminosity delivered to the experiments in the early part of the 2010 run.

Against this backdrop, GridPP has successfully operated the UK Grid with average Tier-1 availability of over 99% since data-taking started. New Tier-1 hardware has been deployed to fully meet the 2010 MOU commitments, and the next procurement cycle has been initiated. The Tier-2s continue to deliver to the experiments over a period where usage patterns have changed considerably due to the advent of data and user analysis. The GridPP4 proposal was completed, submitted, reviewed, and taken through the PPRP and Site-Visit processes.

Project Status

Since the last OC meeting, quarterly reports have been produced covering Q4 2009 and Q1 2010. The latest ProjectMap thus covers the period to the end of the first quarter of 2010. The ProjectMap is online at: www.gridpp.ac.uk/pmb/ProjectManagement/GridPP3_ProjectMap_8.xls, while the reports for each quarter are at: www.gridpp.ac.uk/pmb/ProjectManagement/QuarterlyReports/reports.html.

The overview of the ProjectMap for Q110 is shown in Figure-2 below.
Table 1 below shows the evolution of the ProjectMap milestones and metrics for each quarter since the start of GridPP3. The number of metrics and milestones fluctuates slightly each quarter, as some areas of the project add new milestones, and metrics are reviewed to ensure they still meet requirements. For example, one of the tests used to measure Tier-2 performance has been ‘suspended’, as no longer reflecting site performance accurately. Overall, the number of ‘OK’ and ‘close to target’ metrics remains similar to previous quarters. Of note is the introduction of a new category for milestones that are ‘close to completion’: this mirrors the similar category for metrics that are close to target. It includes milestones that are late but where no remedial action is necessary: for example, reports that have been written and are in the process of review, but have not yet been signed off.
In Q110, as in Q309, there were 22 metrics that were ‘red’ - not met, and not close to the target. These reflect a range of independent issues, most of which were short term. For example:

- An issue with the CASTOR Storage Area Network (SAN) took a week to resolve, resulting in poor availability in February before data-taking recommenced. This led to red metrics for the CASTOR SAM tests and for the number of incidents reaching level 4 in the disaster management system.
- NorthGrid are below their expected level of disk, but procurements are underway that will rectify this.
- APEL suffered a major database corruption that exposed a problem with backups. This was fixed by an upgrade to servers and SSD disk, and RAL do not believe that any data was lost. However, availability was low as a result.

The LHC restarted at the very end of this quarter, and the experiments were happy with the service during early data taking. The next set of quarterly reports (Q210) will explore data taking in more detail.

Overall, only 3 milestones were ‘red’:
- Discussions with NGS are ongoing over the shape of the NGI, and the split of services between GridPP and NGS. The milestone relating to SLA/SLDs between GridPP and NGS is pending the outcome of these discussions.
- A few services remained in old Tier-1 ATLAS centre. The Tier-1 did not wish to migrate the remaining critical servers to the UPS room until the UPS issue was resolved.
- Major users of the ADS service have been migrated, but the service is not yet closed. It is not obvious there is significant benefit from pursuing the minor users, and the Tier-1 is deciding how to progress this.

Of milestones overdue at the time of the last OC, the Tier-1 Disaster and Business Continuity Plan is now complete, and an extremely useful security review of the Tier-2 sites was presented at the GridPP Collaboration Meeting in April.
**Risk register**

The GridPP3 risk register has recently been updated, with the new version available at: [http://www.gridpp.ac.uk/pmb/ProjectManagement/GridPP3_RiskRegister_3.xls](http://www.gridpp.ac.uk/pmb/ProjectManagement/GridPP3_RiskRegister_3.xls)

There are now three high level risks:

- **R1** - Recruitment and retention. This is now becoming a substantial potential problem, as staff start to fear that agreement on GridPP4 funding is likely to be delayed. Previous experience has shown that staff are likely to look for new work if they feel their positions are insecure, and we know of a number of individuals who are considering leaving. In addition, the City is now recruiting again after the financial crisis: this has recently been the destination for several excellent GridPP members. The likelihood of R1 has been raised to 4, with an impact of 3, giving an overall risk of 12.

- **R2** - Sudden loss of key staff. The likelihood of this has been increased to 3, for the same reasons as above. Staff are trained to overlap in roles where possible, but this is difficult in some areas with small numbers of staff, for example at some of the Tier-2s. The new Tier-2 staffing allocation in GridPP4 attempts to mitigate this risk. The impact of this is 3, giving an overall risk of 9.

- **R27** - Uncertain long term funding. Future funding for GridPP is uncertain, and the impact of this increases as we approach the start of GridPP4. The likelihood remains at 4, but impact has been raised to 3, giving an overall risk of 12.

In the previous version of the risk register, R12 (Machine room problems compromise Tier-1) was seen as high risk. There are still occasional issues with the machine room, but it has run well for a long period. R12 has therefore been reduced to likelihood of 2 and impact of 3.

**LCG Status**

**LHC**

On 30 March the LHC started its physics programme with collisions at 3.5 TeV/beam. The next major milestone came on 19 April with a ten-fold increase in luminosity. This was thanks to two simultaneous developments: firstly the number of particles in each bunch was doubled, and secondly the beam size at the interaction point was squeezed down. Four weeks of running under these conditions led to significant quantities of data being accumulated by the experiments. The next milestone, over the weekend of 22 May, was that running started with 13 bunches in each beam. This resulted in a luminosity of $2 \times 10^{29}$, to be compared with $10^{27}$ on 30 March, the design figure for the LHC of $10^{34}$ and the objective for 2010 to reach $10^{32}$. The 13 bunches should also be compared with the design target of 2808 bunches.

All this was achieved during physics running, leading to incredible progress being made by the experiments. They have been running with 90% efficiency; billions of collisions have been recorded and successfully dispatched for analysis via the LHC Computing Grid. The rediscovery of the Standard Model, which is necessary before one can confidently say the detectors are ready for new physics, is well underway. There are even some intriguing observations about the properties of collisions at this new energy. As a measure of their success to date, the experiments have already published or submitted over a dozen papers based on LHC collision data to peer reviewed journals and conferences.
**WLCG**

The grid infrastructure is now working for physics. Real data is being reprocessed, simulated events are being produced, and all data is being delivered to users for analysis. The integrated volume of data and live-time of the accelerator are still lower than the final planning and not all resources are equally utilized, but the activity level is high.

In the first months of data taking, while the data volumes are still low, the experiments have been changing their data handling tactics frequently. For the very first data runs, raw data was distributed from the Tier-0 via Tier-1s to the Tier-2s to give everyone an early look. Then, as datasets started to accumulate, data distribution was reduced as planned, but reprocessing continued more frequently than expected in the steady-state, as a better understanding of the detectors was achieved. In the UK, data has been flowing from CERN through RAL to the Tier-2s more quickly than anticipated, since the data quality checking at CERN has been light. No major problems have been uncovered although data rates to some Tier-2s have been higher than anticipated (for the reasons given above). As the experiments start to settle into the planned production mode, this is expected to abate, but GridPP will keep an eye on this and consider whether UK network plans need revising.

Once the programme for LHC was set early this year, the experiments revised their requirements for 2010-12 in the light of the revised running time. The LHC Resource Scrutiny Group met in April and approved the level of the experiment requirements for 2011-12 on the grounds that they were within 10% of their own estimates. GridPP subsequently requested the UK experiment representatives provide new extrapolated requests up to 2015, using these new numbers as the starting point. These revised requirements have been built in to the latest GridPP4 plan as presented to the Visiting Panel in April and dovetailed (as far as possible) with the reduced spend now foreseen in GridPP3. The UK Tier-1 CPU and disk pledges were in place by the target of 1st June and tape will be provisioned just in time to meet the rising data profiles.

The LHC Experiments Committee (LHCC) also met in April and expressed concerns in the following areas: Alice resources; delays in the progress of the Tier-0; the transition from EGEE to EGI; and the long-term sustainability of the middleware. None of these relate specifically to GridPP, although we are in a position to influence the EGEE-EGI transition with key roles in EGI.

Also at the April LHCC Review of WLCG Sergio Bertolucci, CERN Director for Research and Scientific Computing said, “The committee is impressed and pleased to see that the WLCG and the experiments’ computing models are successfully processing and analysing the first LHC physics data without any major problems. This is the result of many years of careful planning and investment, and the hard work of a great many people. We send our congratulations to all the staff involved, at CERN and the many computer centres around the world.” Figure-3 shows that over the last year, global LHC computing has tripled, with a substantial part of the increase associated with the LHC turn-on in December 2009.

![Figure-3: Growth in Global LHC Computing over the last year.](image-url)
**EGEE/EGI Status**

Following on from the series of EGEE projects (EGEE-III ended in April 2010), responsibility for operation of the European production grid has passed to the European Grid Initiative (EGI). EGI is made up of a consortium of National Grid Initiatives (NGI) and a central coordinating office, EGI.eu, located in Amsterdam. The EGI collaboration is governed by a Council of NGI representatives, with operation of EGI.eu supported by fees paid by the EGI members. For the UK, JISC (Joint Information Systems Committee of HEFCE) has joined EGI and paid the EGI.eu membership fees for 2009 and 2010\(^1\).

EGI.eu was formally created in Amsterdam on February 8\(^{th}\) 2010; the first meeting of the EGI Council took place on March 3\(^{rd}\) 2010. The beginning of May saw the start of full-time operations at EGI.eu with the appointment of a number of key posts in Amsterdam. An Executive Board that reports to the full EGI Council oversees the operation of EGI.eu. Neil Geddes is the UK Council member and a member of the Executive Board.

The structure and operation of the EGI basically follows that proposed by the EGI Design Study (EGI-DS) and described in the EGI Blueprint: [http://web.eu-egi.eu/fileadmin/public/EGI_DS_D5_3_V300b.pdf](http://web.eu-egi.eu/fileadmin/public/EGI_DS_D5_3_V300b.pdf).

In November 2009, EGI submitted a project proposal to the European Commission to co-fund the creation of EGI.eu and the operation of the EGI grid for the first four years of EGI. This project, EGI.Inspire, is currently in the final stages of negotiation and will be coordinated by EGI.eu. From the UK perspective, the EGI funding assumed in the GridPP4 proposal is now assured, providing that the required UK-matching effort is realised; this depends on NGS3 and GridPP3 funding for the first year and then on the future projects beyond that.

---

**Tier-1 Status**

**Overview**

The Tier-1 operated well during the first three months of physics data-taking and there have been very few surprises. During this period, the priority areas of work have been:

- Last minute preparations for first LHC physics data taking.
- Service operation during LHC data taking.
- Delivery and commissioning of 2009 disk and CPU hardware.
- Commencing 2010 disk and CPU capacity procurements.
- Planning and testing of upgrade path from CASTOR version 2.1.7 to version 2.1.9.

**Fabric and Infrastructure**

The new R89 machine room has generally provided a stable environment during this period.

- The cooling system was upgraded to reach its full capacity of 2.2MW and the cooling pumps, which had caused problems last August, were replaced with more powerful models. The upgrade was carried out transparently with no impact on service availability.

- Work continued to identify a solution to address the stability problems in the UPS supply. Two possible solutions have been identified and a final decision will be made shortly.

---

\(^1\) In 2009, STFC, via GridPP, contributed £9k to the EGI subscription.
Large deliveries of disk and CPU hardware were received. The CPU hardware was delivered and commissioned in time to meet WLCG MoU commitments. Sufficient disk capacity was deployed to meet WLCG MoU commitments for disk; however acceptance of one delivery (50% of the FY09 purchase) was delayed while the supplier resolved technical problems. This demonstrates the need for the recently adopted policy of maintaining a buffer of disk capacity equal to 50% of the next annual purchase, though fortunately in this case the problem was resolved. This delivery is now proceeding through the Tier-1’s own acceptance tests and is expected to be deployable in late June. The 2010 disk and CPU procurements have commenced, with target delivery dates of December 2010.

Part (50%) of the 2006 disk server generation (0.5PB) was phased out of operation after 3 years instead of the planned 4 years after drive failure rates made their continued operation untenable owing to the increased risk of data loss. These were the last of the old RAID 5 disk servers making them particularly vulnerable to drive failure. Work commenced to upgrade the main RAL site link to SuperJanet 5 from 10Gb to 20Gb and to add a second 10Gb resilient OPN link to CERN. Work continued to roll out the Quattor Fabric management system. This is not expected to be completed until the end of 2010 however both the disk and CPU farms have now been upgraded.

**Production, Operations and Service**

During February the last few planned changes were implemented ahead of the start of physics data-taking. Unfortunately the migration (to unwind the changes made when the UPS developed problems) of the CASTOR databases back onto their original Storage Area Network (SAN) ran into problems when the SAN proved to be unstable. It took nearly a week to resolve these issues leading to poor availability for February (93%). However, once data-taking started, the service has been reliable with very few operational problems. Availability (ops VO) for March was 100%, April 98%, May (100%). Generally operations were smooth and trouble free; callout rates fell by about 50% during the period as stable operation was emphasised over service development.

Plans to upgrade to CASTOR 2.1.9 (from 2.1.7) and a test strategy were developed. As some previous major upgrades of CASTOR have been problematic, considerable effort has been made to ensure that sufficient and suitable testing is carried out ahead of this upgrade to minimise the risk. The GridPP PMB reviewed these plans in April: the need to upgrade was accepted by the experiments as necessary and the test strategy was considered to be sufficient. The PMB intends to review the outcome of testing in August and if satisfied will approve the upgrade plans, which we expect to implement in the autumn.

Although experiment workload has been high at times, the service infrastructure has generally been relatively unstressed. Where problems have occurred (rarely) they have usually been addressed very quickly. We expect experiment workload to continue to increase but are not aware of any fundamental constraints within the service that will prevent us meeting our commitments in 2010.

**Management, Business Processes and Communications**

- The Tier-1 remains at full staff complement.
- The Tier-1 met its spend and procurement plan for FY09.
- Staff and finance information is now managed through the Research Council’s new Shared Service Centre. This will require changes in the internal systems used by the management team to manage staff and finance information.
- The new change control system is working well and appears to be improving the planning and preparation of changes.
Deployment Status

Overview
During this reporting period there have not been any significant security incidents and operationally the grid has run smoothly. Advances have been made in resource deployment at many sites. In April, EGEE ended and EGI began. Further steps towards regional independence have been taken with the recent commissioning of UK-run Nagios services. The period has seen increased liaison with NGS counterparts as our experts have taken an active role in sharing their knowledge to allow NGS sites to operate gLite services. With low luminosity the grid mechanisms have not been stretched by the LHC VOs, but as expected, a variety of minor problems have arisen and been addressed.

Deployment & Operations
Since the last report, there has been further deployment of CREAM CEs across GridPP sites and new instances of SCAS/ARGUS deployed; these provide a site central service for user authentication. Glexec has also been deployed more widely but further rollout is pending experiment testing and plans for uptake. Data movement and processing have remained the core priority for the experiments. Consequently, the WLCG MB agreed in March that the policy suspension allowing Multi-User Pilot Jobs should be extended until the end of the year. While this does theoretically pose a risk to the infrastructure, it was felt to be small.

Smooth running has been the general experience this year, but there has been a need to iterate and improve in some areas. For example the allocation of Tier-1 to Tier-2 FTS channels has had to be revised to give extra channels to larger sites for which transfer bottlenecks were seen. ATLAS has very recently had to deal with Tier-2 datadisk spacetokens becoming close to full at more than 5 sites. This has been the result of increased distribution of datasets over what was expected (due to reprocessing runs). ATLAS UK is working on a strategy to redistribute space held in other spacetokens. In a related area, ATLAS has observed some issues with poorly justified dataset requests from its users, that in some cases have circumvented the Data Transfer Request Interface – with the result that data ends up in the wrong spacetokens at sites and fills disk.

CMS and LHCb have generally run well at UK Tier-2s. However, there is currently an unresolved issue affecting LHCb work at about one quarter of sites, where there are poor transfer efficiencies (between 5% and 50% failures) of job output to remote SEs from sites using Network Address Translation (NAT) to expose their Worker Nodes. Despite extensive testing the underlying cause remains unknown but currently looks like a gridftp incompatibility with (some) NAT configurations. CMS sites have also recently been affected by poor efficiency jobs, but in this case the problem is understood to result from the implementation of lcg-cp in the experiment layer called CRAB, such that when many jobs complete at a site they compete for available bandwidth and eventually large quantities fail to copy out their output.

There have been two significant site changes in recent months. Firstly, the RHUL cluster has been successfully moved from its temporarily placement in the Imperial College machine room to a new RHUL machine room. There were some teething problems (such as with its network configuration) but the site is now fully operational. Secondly, UCL-CENTRAL resources have been merged with those of UCL-HEP to appear as a single site from an external perspective.

In operations, the move to biweekly (from once monthly) experiment plus all sites meetings has worked well and will be kept as a model for the foreseeable future. Monitoring is an area of current common focus and was explored at the June HEPSYSMAN meeting at RAL. The topic was also raised at GridPP24 and received some discussion at the GridPP Storage Workshop. Following the adoption of Gstat2 in EGEE/WLCG, another area that has received attention is that of correcting site published information. The logical and physical CPU numbers are now picked up by monthly WLCG Tier-2 resource and availability reports showing available capacity. In line with EGEE3 to EGI transition planning, a lot of comparative work has taken place in relation to development and results to satisfy the community that migration from SAM centrally run monitoring to Nagios regional monitoring could take place. The UK was one of the first countries to make the transition and is now running its site testing from a server based at Oxford.
Work by the GridPP data group to interface the UKQCD DiGS interface with SRM based storage has been a success. DiGS can now communicate with DPM and Storm-based SEs. Testing against dCache is in progress.

More widely there has been good progress in the area of data integrity checking (cross checking the accuracy of information systems and differences between what SRM and the experiment tools see as available at a site) and identification and cleaning up of “dark data”. File checksum validation work has progressed in parallel and shown no major issues at any site. Finally, storage optimisation testing has continued and resulted in improved site efficiencies – particularly following experiment reordering of their datasets.

In February there was a problem with the gridpp.ac.uk DNS at Manchester (including the secondary DNS) which prompted work to start on hosting a secondary DNS at RAL. The specific cause of the problem was a kernel panic on the DNS host. The impact was larger than it should have been due to the DNS and several other services being in the process of host migration at Manchester. To mitigate future occurrences DNS backups are being sought in the Manchester computer centre and at RAL. February also saw some disruption when the central APEL database suffered a failure and needed to be rebuilt – many sites had to republish their accounting data for certain periods.

Tier-2 MoU Commitments

Delivery of Tier-2 resources against the pledges is shown in Table-2. Overall, the UK has delivered against the 2009 pledges (with just two sites low on disk). With respect to the 2010 pledges due this month, the March-2010 status shown in this table demonstrates that, overall, the CPU pledge will be met and the disk pledge is not far short. Subsequent to the collection of this data, STFC has released the second tranche of Tier-2 grants (delayed at STFC’s request to FY10) and disk procurements are underway at many sites, which will bring the disk level up to the pledged value within a few months.

<table>
<thead>
<tr>
<th></th>
<th>CPU</th>
<th>Disk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HepSpec06</td>
<td>cf 2009</td>
</tr>
<tr>
<td>Brunel</td>
<td>2446</td>
<td>1997%</td>
</tr>
<tr>
<td>Imperial</td>
<td>4374</td>
<td>311%</td>
</tr>
<tr>
<td>QMUL</td>
<td>7920</td>
<td>145%</td>
</tr>
<tr>
<td>RHUL</td>
<td>2800</td>
<td>284%</td>
</tr>
<tr>
<td>UCL</td>
<td>2446</td>
<td>1997%</td>
</tr>
<tr>
<td>Lancaster</td>
<td>5363</td>
<td>367%</td>
</tr>
<tr>
<td>Liverpool</td>
<td>4374</td>
<td>311%</td>
</tr>
<tr>
<td>Manchester</td>
<td>4374</td>
<td>311%</td>
</tr>
<tr>
<td>Sheffield</td>
<td>1516</td>
<td>1915%</td>
</tr>
<tr>
<td>Durham</td>
<td>5699</td>
<td>978%</td>
</tr>
<tr>
<td>Edinburgh</td>
<td>15320</td>
<td>209%</td>
</tr>
<tr>
<td>Glasgow</td>
<td>1700</td>
<td>708%</td>
</tr>
<tr>
<td>Birmingham</td>
<td>3344</td>
<td>344%</td>
</tr>
<tr>
<td>Bristol</td>
<td>1836</td>
<td>790%</td>
</tr>
<tr>
<td>Cambridge</td>
<td>15320</td>
<td>209%</td>
</tr>
<tr>
<td>Oxford</td>
<td>3332</td>
<td>322%</td>
</tr>
<tr>
<td>RAL PPD</td>
<td>15320</td>
<td>209%</td>
</tr>
<tr>
<td>London</td>
<td>19173</td>
<td>228%</td>
</tr>
<tr>
<td>NorthGrid</td>
<td>22719</td>
<td>270%</td>
</tr>
<tr>
<td>ScotGrid</td>
<td>23212</td>
<td>267%</td>
</tr>
<tr>
<td>SouthGrid</td>
<td>19173</td>
<td>228%</td>
</tr>
<tr>
<td>Total Tier-2</td>
<td>93853</td>
<td>254%</td>
</tr>
</tbody>
</table>

Table-2: Disk and CPU resources at GridPP Tier-2 sites as of the end of March 2010 compared to the pledges due in June 2010.
Users’ Reports

ATLAS

After successful commissioning of the detector, ATLAS has recorded a total of 198.5 million events at an energy of 7 TeV. These events have been reconstructed at CERN and the ten external Tier-1 centres and then redistributed according to the computing model (which foresaw more copies of the ‘raw’ formats distributed to users for early data). Datasets have arrived at the UK Tier-1 on average 4 hours after processing and to the Tier-2s after a further 4 hours (a noteworthy achievement!). We have been performing frequent reprocessing campaigns, as detector and software performance understanding improves, and in preparation for conferences. Ganga has been extensively used for the analysis, and the UK has been leading the minimum bias and $J/\psi$ analyses, which are both early physics measurements and essential benchmarks of the detector and software performance.

ATLAS has, in general, been very happy with the performance of the UK cloud in the last 6 months. The Tier-2s have not only been workhorses for the simulation production, but the larger Tier-2 sites have been prominent in the early data analysis (see Figure-4). The slight dip in production efficiency in 09Q4 was almost entirely down to single-site issues. The Tier-1 has generally been very reliable, with the CASTOR issues much improved; the slight delays in disk deployment were manageable.

A longer-term issue that is currently under consideration is the network bandwidth to the Tier-2s. In general, the earlier projections of the required bandwidth have been followed quite well. However, two issues have led to the larger sites running for periods uncomfortably close to the maximum bandwidth. The first is that the reprocessing campaigns have been more frequent and more intense than originally foreseen and involved both detector and simulated data (leading to more files being replaced in a shorter time). The second effect is that there is more concentration of computing activity into a subset of sites, which means that the peak bandwidth to those sites has increased. At present, the networking is sufficient to our needs, but we are looking to upgrades for the future; such a growth was a feature of our original forecasts.

Figure-4: Analysis job share by site in May 2010 for ATLAS. 3 UK sites (RALPP, Lancaster, Glasgow) are in the top 18 of all ATLAS sites (including T1s) for analysis jobs.
The period since the last oversight meeting has been very exciting for CMS with almost 20 nb$^{-1}$ of data collected at a centre of mass energy of 7 TeV. This data has enabled us to test the performance of the detector through the standard calibration process and the reconstruction of well understood physics objects (π$^0$, K$_S$, J/$\psi$, b-hadrons). These have shown that the detector is working very well and the inevitable initial minor problems have been corrected with impressive speed and efficiency. These studies have also shown a better than expected agreement between simulation and real data.

During this period of data taking the computing system has performed well in all activities, however the volumes of data have not yet been particularly challenging. The positive influence of the computing challenges and exercises that have been carried out over recent years is clearly visible. Nevertheless, the computing system continues to evolve, in a very conservative way, as lessons are learnt from processing real data. One area that has seen a considerable increase is analysis at the Tier-2 centres, especially at those Tier-2s hosting analysis groups who might hope for a discovery in early data. Whilst a large fraction of all CMS user analysis has been performed at the Tier-2 centres for some time, there has been an increase in skimming and production of analysis group specific reduced data formats.

The UK has performed well during this period and has provided a Tier-1 centre at RAL and effectively three Tier-2 centres at Imperial, Brunel and RALPPD. The Tier-1 has been very responsive to CMS’s needs and overall has provided a very good service, which has been recognised by CMS centrally. The UK supports 5 analysis groups at the Tier-2s: e-gamma (Imperial), Electroweak (Brunel), exotica (RALPPD), SUSY (Imperial) and trigger studies (Imperial and Brunel). These are amongst the most active analysis groups and the UK has made a significant contribution. For example the SUSY group (one of the most aggressively active groups in CMS as it has some realistic expectation of discovery or exclusion in early data) has made Imperial its main Tier-2 centre with a complete copy of all of their skims and reduced data formats. The other Tier-2 centres worldwide that support SUSY have only a fraction of the data and are, effectively, a back up for Imperial.

In conclusion, the period of data taking so far has been very interesting and informative; however the rest of the year we expect will be even more interesting as the integrated luminosity grows in an exponential manner.

![CMS PhEDEx - Transfer Rate](image)

Figure-5: Real CMS data flowing to the UK Tier 1 for custodial (and non-custodial storage) and to the UK Tier-2 centres for users to perform the first real physics analyses on real LHC data.
**LHCb**

After successful commissioning of the detector, the LHCb experiment has recorded a total of 473 million events at an energy of 7 TeV and with different polarities of the spectrometer magnet. These events have been reconstructed at CERN and the six external Tier-1 centres and then re-distributed according to the LHCb computing model. Using this early data sample, the applications software and grid production system has been regularly updated and optimised - with the result that the data has now been re-processed six times in preparation for the summer conferences. Physics analyses are being performed on the grid using the Ganga interface (a joint LHCb-ATLAS interface, developed by GridPP) to the DIRAC workload management system.

The particle identification system, based on two RICH detectors, has provided the means to isolate clean resonant peaks, such as the φ and J/ψ, essential for checking calibrations and understanding experimental resolutions. A rich programme of charm physics has started with the isolation of a number of different D meson channels and the first B meson decays have been observed in the data.

The UK Tier-1 Centre has deployed the WLCG resources pledged for 2010 data taking and, aside from some fabric failures (disk servers), the performance has been good with no bottlenecks identified. LHCb continues to successfully exploit the UK Tier-2 sites for Monte-Carlo production. The only issue, still under investigation, has been the occasional problem at five of the sites with uploads of data to some remote SEs. Over the period February - June 2010, the UK has processed 17.1% of the total number of batch jobs submitted by LHCb, in line with the authorship fraction of the experiment.

Figure 6: First reconstructed beauty particle candidate from LHCb

This has been an excellent start to UK data processing and analysis for LHCb and we look forward to the rest of 2010 with great anticipation.

**Other Experiments**

The UKQCD collaboration has reviewed and revised how best to utilise SRM-based storage within GridPP using their existing software infrastructure. It was concluded that this was best done through disk-based storage at Tier-2 sites, rather than tape based storage at the Tier-1 centre, which was assumed previously. With help from the GridPP data-management team, good progress has been made using the SRM version of the DiGS interface to test facilities at Glasgow and Edinburgh. In 2010 the collaboration is preparing to exploit 16TB of storage provided through GridPP.
SuperB were set up as a VO and, after installing software at the QMUL and RAL Tier-2 sites, the UK produced 25% of a large simulation run totalling 1.7 billion events. Progress has also been made on exploiting the Tier-1 centre for a large simulation run in June/July to produce the equivalent luminosity of \( \sim 1.5 \text{ ab}^{-1} \) for the Technical Design Report (UK share \( \sim 20\% \)). The UK design for the silicon vertex geometry will be simulated and a Computing Model and Detector White Paper will also be published later this year using information from this Monet-Carlo data.

At the UK Tier-1 Centre, the MINOS collaboration has re-commenced Monte-Carlo production over the last three months. SuperNemo has started to exploit the centre since the beginning of the year (in addition to Tier-2s) and MICE are expected to run from August.

**Impact**

**Events**

Key events attended since the last OC meeting include:

- A GridPP stand at the launch of EPSRC’s e-Science review in February. One of only four projects invited to run a stand, this was an excellent opportunity to present GridPP to science policy-makers.
- At end of March, GridPP held a stand at the annual IoP HEPP Group meeting at UCL.
- The outreach team’s busiest day was on the 30th of March, when three events took place in parallel. In addition to the ongoing IoP meeting, this was the official opening for the Tier-1 facility at RAL. The same date was also selected by CERN to hold the LHC “first physics” event. In the UK, LHC “first physics” was marked with an early morning STFC event for journalists at UCL, at which GridPP members distributed literature.
- The Tier-1 opening was very successful, with talks, tours of R89 and a reception. To introduce the facility to potential new users, the first part of the day covered the various services provided at the Tier-1. The official opening included talks from Keith Mason the STFC Chief Executive, and David Britton of GridPP.

*Keith Mason talks at the Tier-1 opening*

**EGI and NGI**

April saw the final meeting for the EGEE project at the User Forum in Uppsala, where GridPP shared a stand with the other UKI partners. The launch of EGI means that dissemination within a European and UK context has changed somewhat. To this end GridPP has formed a NGI dissemination team along with NGS, the Software Sustainability Institute and the National eScience Centre, to help promote the UK within EGI and beyond. Of the two EC proposals that GridPP’s outreach team helped develop, the Creating Users for E-infrastructures project was unsuccessful. However, GridTalk-II was approved and will commence in September, when it will fund two FTEs in the UK at QMUL and IC.

**News**

Ten news stories have been published on the GridPP website since the last OC meeting. In addition, GridPP was mentioned in five press releases during the period, covering topics including LHC “first physics”, Viglen computing and the launch of EGI.

**Knowledge Exchange**

*Lumerical Solutions Inc.*

GridPP and Vancouver-based Lumerical Solutions Inc. produced a joint press release in February announcing the donation of ten FDTD Solutions Engine licenses to ScotGrid. The licences allow scientists to perform large-
scale design of devices across applications in photonics, and users of FDTD Solutions can run their simulations on the ScotGrid infrastructure at no additional cost.

Glasgow industry day
At the end of February, the University of Glasgow held an industry day to investigate how researchers, the university and entrepreneurs can work together. Due to the success of ScotGrid, one of the themes on the day was grid computing. Members of ScotGrid demonstrated the benefits of the technology to academia and industry, and discussions included the potential for knowledge transfer projects and industrial placements.

Others
- Due to the start of the LHC, it was decided that ‘data’ would be the theme of GridPP’s Collaboration meeting in April at RHUL. The planned session about knowledge exchange was therefore moved to the next meeting in September.
- QMUL has recently appointed a new KE consultant, Alex Efimov, who already has experience with GridPP. We plan to use his expertise in engaging with new user groups.
- PEGASUS, an LSE research project funded by the EPSRC to study how GridPP works, ran for almost four years and has just handed its final reports to EPSRC.