

**GridPP**  
UK Computing for Particle Physics

**GridPP Project Management Board**

---

# Review of GridPP3

---

Document identifier :	<b>GridPP-PMB-155-Review.doc</b>
Date:	<b>17/05/2011</b>
Version:	<b>1.1</b>
Document status:	<b>Final</b>
Author	<b>PMB</b>

---

## Introduction

GridPP3 ran from April 2008 until the end of March 2011. The three years were punctuated by six Collaboration Meetings where the majority of the collaboration got together, allowing extensive formal and informal communication. The meetings were loosely structured around themes that provided a focus for discussion about issues of current importance. The themes chosen provide some kind of roadmap to progress over the project. The first collaboration meeting (GridPP21, Swansea, Sep-08) was based on "Information" and was intended to explore what information about the status of sites, the Grid, and the experiments, was available and how it could be accessed, interpreted and used. This was followed by GridPP22, which took place at UCL in Apr-09 with a theme of "Resilience". Here the goal was to explore how to "harden" the Grid services to make a more reliable infrastructure. Then, in Sep-09, GridPP23 took place at Cambridge, by which point it was apparent that a sustained period of real data taking was imminent. The theme of that meeting was "Final Steps to Data". GridPP24 at RHUL in Apr-10, the first meeting after a sustained run of the LHC, was the complement to this and simply had the theme of "Data!" Despite the success of the worldwide Grid in handling the LHC data, GridPP25 in Ambleside in Aug-10, had a theme of "Scaling the Peaks Ahead", to reinforce the point that the LHC had only just started and it was no time for complacency. Finally, in Mar-11, GridPP26 at Sussex provided an opportunity to focus on "Efficiency", an essential tool to help address the increasing contention for hardware resources against a backdrop of financial austerity.

The remainder of this document reviews GridPP3 from a number of angles. The first of these is the overall performance of the UK Grid. This is followed by a specific focus on the Tier-1 as a critical component and then a brief look at some of the operational challenges that were overcome. The remaining sections look at the "Milestones and Metrics", "Finance", "Effort", and "Dissemination". However, the whole is much greater than the sum of these parts, which make no mention of the enthusiasm, trust, collaboration and hard work of a large and diverse community. In addition to the collaboration meetings above, over a hundred Project Management Board meetings took place and an even larger number of national and local technical meetings. All these things have been essential to the success of GridPP3 and allowed the delivery of computing to the LHC experiments.

## Performance Review

The success of the GridPP3 project can be judged from various performance indicators but is ultimately demonstrated by the performance of the UK Grid infrastructure in handling the flow of real data from the LHC and the provision of timely analysis capacity to UK physicists. This can be summed up by the statement at the April 2010 LHCC Review of WLCG where 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.*"

The report from RCUK review of e-science in published in 2010 had a broader perspective, noting that: "*...the investment in GridPP led to the establishment of the start-up companies imense and iLexIR at Cambridge (Camtology). GridPP/ATLAS/LHCb developed the Ganga interface, as well as Econophysica (mathematical models for commodity trading) at QMUL, with access to resources at Total Oil (geoscience research) from Aberdeen.*"

And:

*“There is likely to be a substantial economic impact [in saving] lost work days as a result of the cure for bird-flu, in which GridPP resources played a major role; as is true for their contribution in 2008 to the discovery of three families of molecules that could be effective against the malaria parasite. “*

In advance of the LHC start-up and the sustained flow of real data in 2010, the Grid was tested by various readiness challenges that provided a reasonable indication of the UK status. The Combined Computing Readiness Challenges (CCRC08) took place in two phases in February and May 2008 and enabled most aspects of the Computing Models of each of the experiments to be tested at the nominal transfer rates. As the name suggests, the goal was to do this simultaneously, but due to various problems, the overlap was only for limited periods. Nevertheless, the exercises were, on the whole, very successful and the UK performance was very good. During this period the Tier-1 brought on line significant new hardware to meet the WLCG commitments, and introduced an on-call system to provide 24-hour support. This contributed to an increase in reliability and availability from about the 90% level up to ~98%.

The Tier-2 sites installed new hardware funded by GridPP during the latter part of 2008, which meant they were able to meet the 2009 MoU commitments for CPU and storage. Although individual sites experienced problems at various stages the distributed and hierarchical structure of the UK Tier-2s enabled them to provide a good service to the experiments and an increasing reliability and availability to exceed the 90% level. The manpower provided by GridPP was essential to this success.

The complexities of the repairing the LHC and ensuring that future problems would not occur took most of 2009. There was no real data during this period but this allowed the sites to concentrate on Service Resilience (the theme of the 22<sup>nd</sup> collaboration meeting at UCL). Service monitoring and disaster recovery procedures were set up. The computing infrastructure was re-tested by the “Scale Test of Experimental Programme for 2009” (STEP09) global readiness exercise. The STEP09 exercise was again very successful: All aspects of the experiments computing models, simulation, data processing and analysis, were run simultaneously at the required data rates and the UK performed well in absolute and relative terms. Indeed, the exercise was handled as “routine operations” at many sites with little of the extreme effort and round the clock monitoring by key staff that had been a feature of the CCRC08 tests.

During this period the RAL Tier-1 moved into its new machine room in R89. Handover was delayed around 6 months, which had a knock on effect that the 2009 CPU and disk purchases were delayed in commissioning. Ultimately, this was not a serious problem due to the revised LHC schedule, which led to a relaxation of the 2009 MoU target dates from April to September 2009.

### **CPU Delivered by GridPP3**

The reduced computing required by the LHC in 2009 allowed other communities to make use of the GridPP computing resources. Overall, in the first half of GridPP3, 20% of the resource use was by non LHC VOs in an opportunistic manner, particularly PhenoGrid, BioMed, ComputationalChemistry, NanoCMOS and ILC as shown in the upper plot in Figure-1 below. In the second half of GridPP3, the LHC usage increased, squeezing the non-LHC use down to 10% (lower plot).

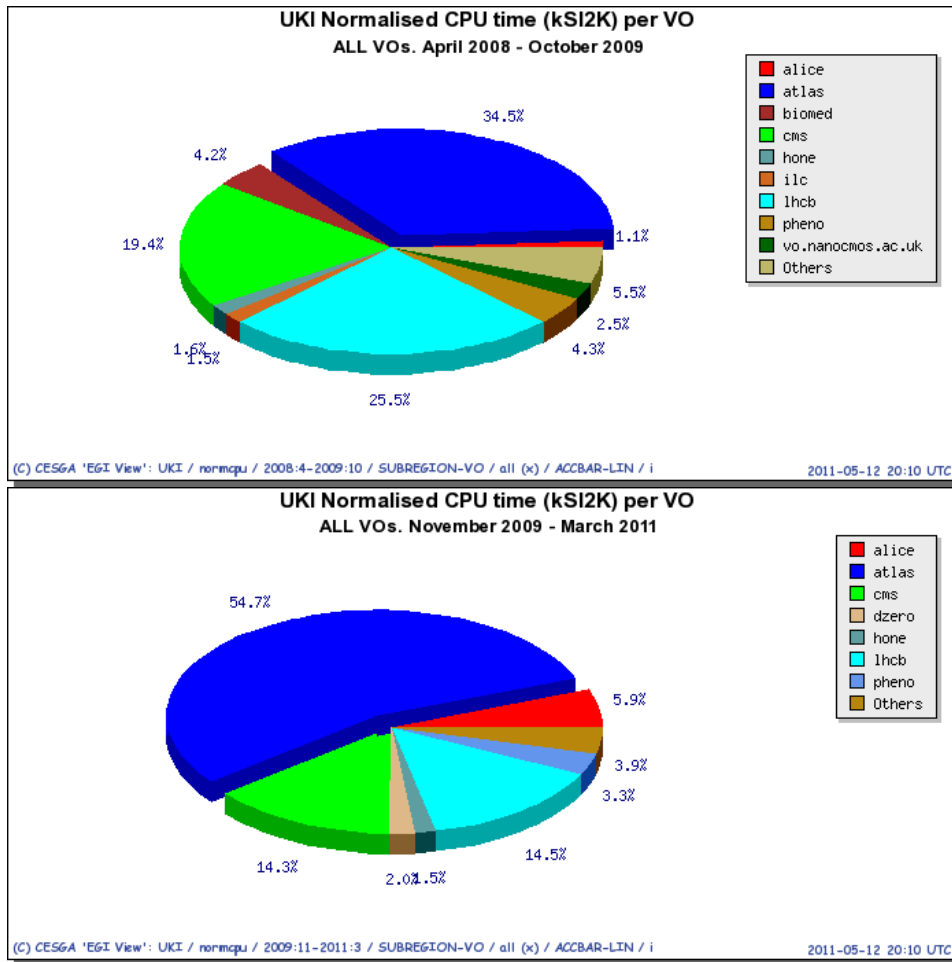


Figure- 1: CPU delivered by VO during first half (upper) and second half (lower) of GridPP3.

The quantity of CPU delivered over GridPP3 is shown in Figure-2, which increases over the GridPP3 project as the deployed resources were ramped up in anticipation of data from the LHC. The plot also demonstrates the relative importance of the Tier-2 infrastructure and the balance between the four Tier-2 centres. The UK contribution is compared with other countries in Figure-3. During GridPP3, the UK was the largest European contributor or resources to the LHC, slightly above, but comparable with, Germany and France. This demonstrates the high availability of the UK resources and the trust of the experiments in UK sites. These plots also cover the first year of LHC running in 2010 and demonstrate the successful ramp up delivered resources in response to the flow of real data and the subsequent reconstruction and analysis.

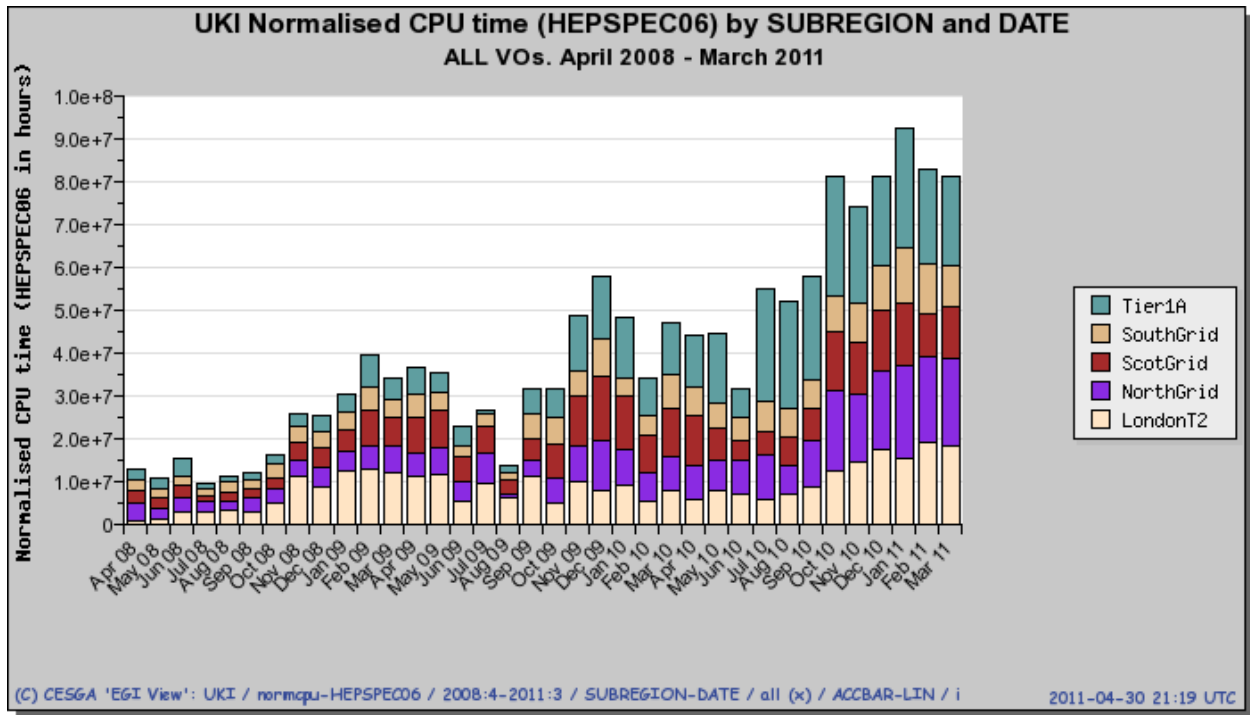


Figure-2: CPU delivered by the UK Tier-1 and 2s over the period of the GridPP3 project

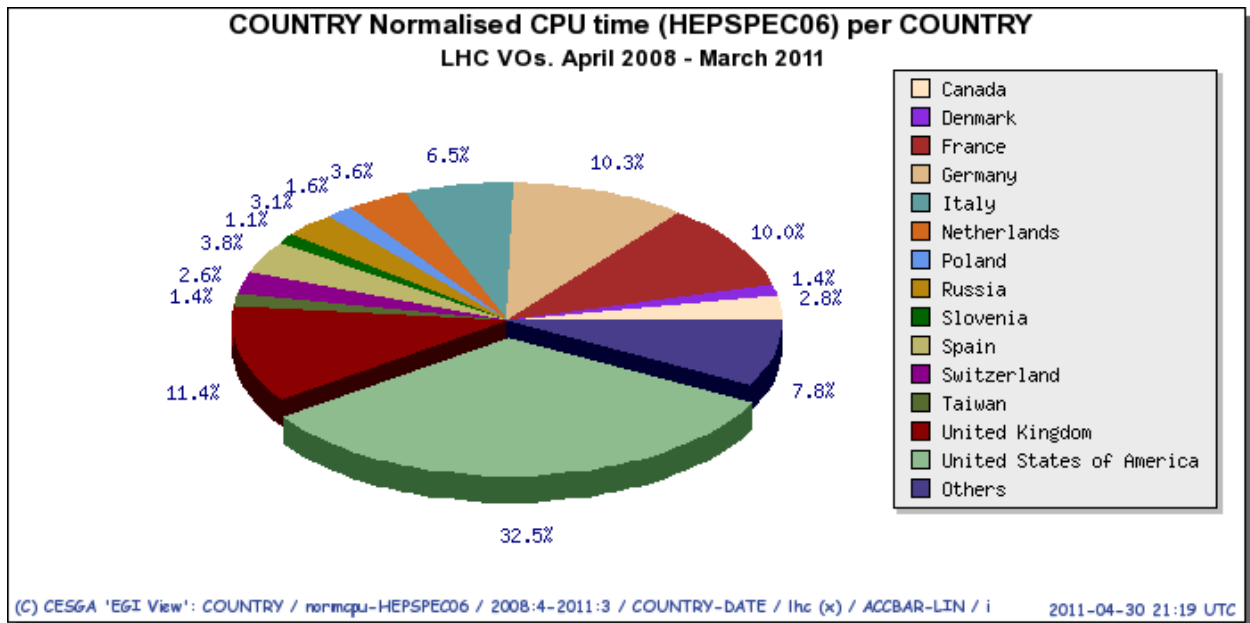


Figure-3 Shows the CPU delivered to the LCG VO's over the GridPP3 Project by all countries

## Security

During GridPP3, security-performance was tested and improved and international leadership roles were reinforced. GridPP3 provided the Deputy Security Officer for EGEE and the leader of the EGI-CSIRT (Computer Security Incident Response Team). During GridPP3 operational security areas included Security Incident Handling, Security Monitoring and Dissemination/Training.

Inevitably, security incidents occurred and GridPP typically experienced one or two per quarter during GridPP3. In such cases it was essential that all system administrators knew the procedure to follow for handling such incidents: who to inform, what to do to contain the incident, how to analyse the incident and prevent it happening again. During GridPP3, the Grid Security Incident Handling procedures have been developed in collaboration with our international partners and have been tested and adopted. To date we have had no experience of security incidents resulting from attacks on the Grid middleware as such, so we have found it very useful to stage a number of Security Service Challenges. One such challenge (SSC3) was designed to mimic a Grid security incident, where a malicious job was run by a Grid user (the security officer!) with a legitimate Grid credential. An alert was then raised some time after this to the site. The challenge was intended to examine three areas in the security incident handling procedure, namely communication, containment and forensic analysis. This challenge was organised by EGEE and started as a test of the WLCG Tier1 sites. The RAL Tier1 came top in 2008 and 4<sup>th</sup> out of 11 in the 2009 challenge.

Next, EGEE then asked all of the national Grids to run SSC3 on its sites. All GridPP sites were therefore challenged and the results were fully analysed. A number of sites did not perform so well in the communication aspects particularly with their local university security team. The containment aspect, on the whole, was well executed, but the forensic analysis also turned out to be rather difficult. The average GridPP site scores for each of the components were as follows: Communication 54/100, Containment 70/100 and Forensics 34/100. We don't have access to the full set of results from the various regions but one federation (DE-CH) presented its averages as: Communication 47/100, Containment 51/100 and Forensics 30/100. Therefore, we conclude that the UK performance was not atypical. All of this feedback was very useful for the planning of a UK Security Training workshop to address these issues which was then given during a subsequent UK HEP Sysman meeting. The next full test of all GridPP sites will happen early in GridPP4.

The Joint (WLCG/EGEE) Security Policy Group, under GridPP leadership, received a revised mandate in 2008 and during GridPP3 revised most of the security policy documents to make them simpler and more general to enable use across many different infrastructures. These policies include documents aimed at users, at VOs, and at Sites. New policies were developed during GridPP3 addressing multi-user pilot jobs, VO Portals, Virtualisation and the data privacy aspects of Job Accounting. JSPG came to an end in April 2010 and was replaced by the EGI Security Policy Group, still under GridPP leadership, which is now addressing the changes required for the EGI era.

The Grid Security Vulnerability Group was initiated by GridPP during GridPP2 to handle vulnerabilities found in the Grid Middleware. The aim is to fix security holes before they are exploited and turn into real security incidents. During GridPP3, the procedures used by this group were developed and approved to ensure that first a risk analysis is performed to define the importance and the urgency and that then patches are produced and made available by the software developers before the deadline is reached for public announcement. During its lifetime, many tens of vulnerabilities have been fixed and patches deployed. With the end of EGEE this was then transferred to EGI, still under GridPP leadership. The procedures have now been broadened to include all middleware stacks in use in EGI and not just gLite. The fact that we have to date not seen any evidence of successful exploits of Grid middleware vulnerabilities is in part due to the success of this group.

## Review of the Tier-1

The main focus of the work at the Tier-1 during GridPP3 was to:

- Manage the considerable increase in hardware volume
- Deploy complete functionality required by the experiments
- Scale up to full experiment workload
- Ensure reliability of both the service and processes.

### Service Operation

As the size and complexity of the Tier-1 operation grew, increasing emphasis was placed on developing processes and procedures to manage service operation. Continuous improvement processes proved to be essential and a number of processes such as Change Control and Incident Management have grown to look quite similar to the ITIL model and there may have been benefit in investing early in the project in ITIL training for more team members.

The introduction of a “production team” to liaise with the experiments, monitor the service and handle day time operation exceptions proved to be a huge success. Responding to the constantly changing operational situation would have been a challenge without a team focussed on ensuring overall service availability. The role of “Tier-1 Production Manager” has become a key player, responsible for all key operational decisions affecting the service.

The out of hours on-call service was a much bigger challenge to introduce than expected, however on-call also brought many unexpected benefits. Not only did it ensure that faults were successfully handled out of hours but also that:

- A well defined process was built to handle exceptions
- A continuous improvement process was put in place to reduce fault rates
- There was a concerted drive to improve resilience
- Service documentation improved dramatically
- Remote operation capability was enhanced giving benefits when access to site is prevented.

Visualisation tools have been barely adequate. It is necessary to have a very clear picture, of everything that is happening on the Tier-1: current exception states, workload on individual servers, performance of larger units such as CASTOR, network rates and flows, experiment job mix and experiment failure rates. Only some of the information rests entirely within the Tier-1's scope, the remainder mainly within the experiments' domain. Monitoring has become a mix of products (eg. Open source, project, experiment or locally written), stitched together by a variety of dashboards. Even so, understanding causes of effects we see on the infrastructure was not always easy. Work in this area continues but the situation is still far from perfect.

### Hardware

As predicted, running a cluster of 500 disk servers has been an “immense challenge”. Hardware failure rates for disk (and CPU nodes) have generally been close to the predicted rate and the basic process of

repair and re-commissioning of hardware was well planned and sized. Typically we expected a disk drive to fail every day during operation (our current rate is indeed about 30-40 per month), but in general we expected to be protected by the redundancy built into the hardware, initially RAID 5 (one parity drive) was found to be inadequate but RAID 6 (2 parity drives) now usually provides good protection and most drive failures can be handled transparently. For a variety of reasons however (middleware, operating system or hardware exceptions) servers need to be temporarily taken out of production operation (typically about once per server lifetime for 1-3 days) and we have found that short term unavailability of servers has had a significant adverse impact on experiments. This indicates that either we need many less fatter servers (potentially reducing the fault rate but increasing the fault impact) or greater hardware layer resilience (probably at considerably greater cost).

Systematic batch related problems with disk servers have been a considerable problem. Aggressive acceptance tests have weeded out most of these problems during the delivery stage, but downstream batch related problems have made it necessary to withdraw entire generations of server from operation while issues are resolved. Detailed monitoring of failure rates and causes has been necessary to identify systematic faults early. Considerable effort (probably 0.25 FTE) has had to be committed to resolving these technically challenging problems with suppliers and manufacturers. Experience shows (and discussion with suppliers confirms) that problems of this kind with storage systems are inevitable. There is no indication that spending more (within a factor 2) on hardware will substantially reduce the risk of problems. Our conclusion has therefore been that reserve disk capacity is necessary (roughly half the annual procurement) in order to ensure that MoJ commitments are always met. Hardware procurements have also been split into multiple purchases from different suppliers (with different hardware configurations) in order to reduce the impact of problems.

Estimations of wide area network traffic rates proved to be accurate and the CERN OPN and SuperJanet 5 links have been able to meet the demands of LHC data taking. Predictions that we would need to be able to deliver average internal rates to disk of 5000MB/s have proven to be reasonably accurate, at least over periods of several days. The local area network solution chosen was based on a tree like structure of commodity network stacks, interconnected by a “low end” enterprise class switch. This has proven to be both reliable and flexible and has easily met the considerable I/O demands of both the batch farm and file moving service. We expect to continue with a similar architecture in GridPP4. The move of disk servers from 1 Gb/s to 10Gb/s connectivity will place increasing load on the 10Gb network backbone and we are investigating recently developed “full mesh” interconnect functionality to provide the means of eliminate outstanding potential hotspots.

The tape system was never well tested during the service challenges leading up to the start of data taking. Nevertheless the robotics system and available tape drive bandwidth met the project requirements, during the first full year, however load on the tape system was lower than expected. Experiments found they were able to hold more data online on disk than planned reducing demand for tape access. This is not expected to be the case in future years and capacity may not meet requirements in future. Experiment data models are evolving in these areas but it is not yet clear what impact this will have I/O capability. Planning flexibility will be required during GridPP4 as the data processing models evolve.

## **Middleware**

At the start of the GridPP3 project we had many concerns about the reliability of the Grid middleware layer (such as CEs, LFC, FTS etc) which still had some truly awful failure rates. By the end of GridPP3 the Grid middleware has ceased to be a major concern, partly this has been down to greater software reliability, partly increased service resilience and partly by additional management layers built by the experiments to handle lower level middleware problems.

Making the storage system work at the scale required has been a huge challenge, not only at RAL (which runs CERN's CASTOR storage system) but at many other Tier-1s around the world. Much of the problem was caused by the relative immaturity of the middleware and the very small deployment base (only 4 sites in the case of CASTOR). In order to reach an acceptable quality of service it proved necessary to introduce a number of continuous improvement processes such as: fault logging and incident reviews, change control and a considerable testing program.

ORACLE requirements grew from almost nothing (just 0.5 FTE of the 3D service) at the end of GridPP2 to a considerable infrastructure in GridPP3 (2.5 FTE), vastly exceeding the planned team capacity (0.8 FTE) and original hardware and licensing projections (essentially nil). This increase in the ORACLE infrastructure was driven by new WLCG components such as the LFC and FTS that used ORACLE as a state-full back end and the CERN decision to base the new CASTOR2 release deployed in 2006 on ORACLE (again as the state-full back end). All these services are vital to Tier-1 operation and the integrity of the LFC and CASTOR databases in particular are absolutely crucial for data retention. A further effect was that as the complexity of the Oracle service grew, the need for increased testing grew too, demanding yet more test instances. At RAL, the e-Science department provided additional unfunded staff effort to support ORACLE and funding support for additional ORACLE license purchases. In 2008 STFC funded an additional GridPP post to support ORACLE. Many Tier-1 centres under-resourced their ORACLE admin effort, those running CASTOR being particularly badly hit.

## **Capacity Planning and Procurement**

Capacity planning and technology foresight proved remarkably adaptable considering how much the LHC timetable evolved over time and how technology roadmaps changed. Projections made in 2006 for tape, disk and CPU technology were reasonably accurate out to the end of the project and prices only departed considerably from plan in 2010 (exchange rates and economic climate being likely causes). The planning and procurement processes have been able to handle unexpected variations in pricing and funding.

Procurement has been a substantial annual undertaking. Significant benefits have been gained from placing our smaller purchases through the Shared Service Centre (SSC) Framework Purchasing Agreement, this has speeded up purchasing and delivery, standardised the equipment we deploy and reduced team effort. Large purchases of disk and CPU have required EU procurements, with consequent lengthy lead times. The process is also inherently risky, with potential legal pitfalls associated with the procurement process itself and the annual challenge of ensuring the equipment works as specified. This has been a substantial undertaking each involving senior management and technical staff. In order to reduce the risk, shorten the process and reduce the effort required future procurements will be made through Framework agreements which are in the process of being set up.

## **Review of Operational Challenges**

The operation of a distributed computing and data Grid throughout the UK is a constant challenge at many levels. However, the nature of the UK Tier-2 infrastructure, wherein sites at 18 institutes are grouped into four regional Tier-2 centres, provides sufficient resilience and redundancy to maintain an excellent service despite individual hardware problems that are tackled at the local level. Thus, GridPP3 at the project-level was not preoccupied with hardware issues at individual Tier-2 sites. A bigger challenge at the Tier-2s, is to ensure that software/middleware releases are up-to-date and configured in the correct manner and the dTeam managed these issues through weekly meetings. The intersection of requirements from

wLCG/GridPP with those from the local institute proved problematic at times, particularly when related to security.

In contrast, the Tier-1 represents a crucial part of the hierarchical structure and, as such, is a single-point-of-failure that is required to function with a high availability. Over GridPP3 the Tier-1 implemented the operational and management structure necessary to achieve this goal. An out-of-hours callout system has been implemented to ensure 24x7x52 operation; disaster-management<sup>1</sup> processes have been developed and exercised; and significant work to ensure the resilience of hardware and services has been undertaken. Nevertheless, due to the critical nature of the Tier-1, challenges at the Tier-1 must escalate to the Project Level when appropriate and this is designed into the disaster-management process. The first issue that came through this process was an assessment of how the Tier-1 service could be maintained in the event of a Swine-Flu epidemic. This turned out to be a useful exercise because the plan to largely operate the Tier-1 remotely was put into action during extreme winter-weather when many staff could not make it to site. This was followed by a series of more tangible problems associated with the Tier-1 move into the new R89 building at RAL, including air-conditioning problems, water-leaks, and problems in the UPS supply causing instability on the core database hardware. There was also a power failure that triggered the disaster management process and an issue where dust from cooling pipe cladding caused health and safety concerns.

## Review of Metrics and Milestones

The GridPP3 project was monitored and managed with the help of the ProjectMap tool developed, with the encouragement of the Oversight Committee, for GridPP1 and refined during GridPP2. The ProjectMap provides a visual overview of the entire project and an ability to drill-down to increasing levels of detail. It was developed because more traditional Gantt charts were found to be unsuitable for a project as broad and adaptive as GridPP. In particular, Gantt charts require time-ordered dependencies and a pre-determined set of steps from beginning to end. The interdependencies within GridPP are often much weaker because the work is embedded in a much bigger global context (thus many dependencies link to things outside, rather than within, the project) and the goal-posts continue to move and evolve over time (for example, the changing experiment computing models and the LHC schedule).

The final status of the GridPP3 ProjectMap is presented and discussed in the current status report document prepared for this Oversight Committee meeting (GridPP-PMB-154-ProjectStatus.pdf) and will not be reproduced here. In summary, only eight out of 236 active metrics were not above or close to their target levels at the end of the project. Fourteen out of the original 250 metrics had been suspended or were currently un-measurable. In terms of milestones, 118 of the 122 milestones had been achieved with a further two close. The two missed milestones related to formalizing an agreement with the NGI on the provision of services critical to GridPP (delayed due to the delay in funding the next stage of NGS) and on the distribution of the APEL middleware (now scheduled for June).

Overall, the metrics and milestones defined for the GridPP3 proved a useful measure of progress and the ProjectMap continued to be a valuable tool. GridPP met the vast majority of the objectives they defined.

---

<sup>1</sup> Previously presented to the Oversight Committee: <http://www.gridpp.ac.uk/docs/oversight/GridPP-PMB-149-DisasterPlan.pdf>

## Financial Review

The financial management of GridPP3 required considerable agility as circumstances evolved within, and beyond, the STFC environment. The first hurdle was the Programmatic Review of 2007/8, which resulted in a reduction in the funding for GridPP3 of £1.24m, in addition to the £1.27m of GridPP2 funding that was previously withdrawn. The review started in the autumn of 2007 and concluded in July 2008, which did allow GridPP to anticipate the outcome to some extent and make savings, primarily by delaying the start of a number of new posts in April 2008. As a consequence of this, and because of subsequent delays in getting approval for new hires at RAL, some WLCG milestones were delivered late. The GridPP3 financial plan was reformulated within the new limits and provided to STFC in July 2008.

By the end of 2008 it was apparent that the LHC was effectively delayed a year and GridPP responded by postponing various hardware purchases to optimise spending-power, which transferred £335k from FY08 to FY09. However, this was complicated by a request from STFC to delay £1m of capital spend at the Tier-1 from FY09 to FY10 to assist with funding profiles. The Tier-2 hardware spend was also delayed at STFC's request. GridPP was also hit by the unforeseen network costs for the OPN network (now £130k/yr) due to a change in the way such costs were handled by RCUK. A site-wide moratorium on hiring at RAL made 2009 an extremely challenging period.

By the end of 2009, the LHC had started in earnest and the UK and worldwide Grid was performing exceptionally well. In December 2009, STFC reversed the previous request and asked for as much capital spend as possible to be brought forward from FY10 to FY09. GridPP successfully spent an additional £661k by purchasing tape-media in advance of the need (there is little or no cost-penalty in this area) and by bringing forward a planned disk purchase. This latter action allowed GridPP to protect itself against future problems with disk-purchases by establishing a buffer of disk equal to half the annual procurement. Although GridPP already mitigated this risk by dividing the disk procurement into two halves with different combination of disk drives and raid controllers, on two previous occasions the failure of procured disk to perform at the required level had led to significant problems reaching the MOU commitments.

During 2010 the external requirements and financial constraints on GridPP continued to fluctuate: The success of the LHC at 7-TeV at increasing luminosities placed significant demands on the Grid infrastructure, requiring a number of areas to run above the design capacity, and the increasing likelihood of running in 2012 foretold a significant increase in the experiment requirements. The Comprehensive Spending Review introduced more financial uncertainty. By June, GridPP were requested to reduce FY10 capital spend by £855k from the figure approved six-months earlier; but by the end of the financial year the pressure was to spend capital before FY11. The backdrop to all this was the GridPP4 proposal that was submitted in early 2010 containing assumptions on requirements, costs, and funding, which had to be frozen at the end of 2009. In an attempt to respond to all this, GridPP brought forward a transition to a new generation of tape technology that was to become available in the first quarter of 2011. This would have had the advantage of bringing useful capital spend into FY10 and reducing the GridPP4 hardware costs. Unfortunately, the drives ultimately arrived in the UK one day after the March 31<sup>st</sup> deadline and the cost of £207k could neither be attributed to FY10 nor to GridPP3.

In summary, it is apparent that the financial management of GridPP3 was made extremely challenging by the changing external factors and the successful delivery of the project objectives under such conditions, was a significant success. However, had there been no delay in the LHC, the various financial constraints would have compromised the project. The figures for GridPP3 will not be final until the RAL outturn is confirmed. However we expect the situation to be close to that presented in the document GridPP-PMB-156-Resources prepared for this oversight committee. The provisional outturn shows an under-spend of £304,341 compared to the funds actually made available. This primarily arose due to the late delivery of Tape Drives (£207,000) and represents a 1.4% under-spend on the total budget of £22,086,295.

## Effort Review

Table-1 shows the effort delivered to GridPP3 over the whole project period. These figures are produced as part of the quarterly reports and it is recognized that there may be some inaccuracies due to personnel changing posts, but the table will give a general overview of the position. The project benefited considerably from additional effort coming from a mix of EGEE/EGI, RAL and university internal funding. Overall, 98% of the funded effort was delivered with the shortfall arising primarily in the “Operations + Experiment” area due to a moratorium on hiring at STFC and difficulties recruiting some posts at STFC. However, in this area GridPP has particularly benefitted from additional effort funded from other sources.

Work Area	GridPP funded effort (FTE)	Fraction of GridPP funded effort delivered	Additional effort delivered (FTE)	Total effort delivered
Management	10.4	100%	0.0	<b>100%</b>
Miscellaneous	3.4	100%	0.0	<b>100%</b>
Operations + experiment	17.5	93%	13.4	<b>164%</b>
Support	21.6	99%	3.8	<b>117%</b>
Tier-1	49.9	98%	13.3	<b>124%</b>
Tier-2	52.4	100%	28.5	<b>154%</b>
<b>Total</b>	<b>155.2</b>	<b>98%</b>	<b>59.1</b>	<b>136%</b>

Table-1: Cumulative effort delivered to GridPP3

## Outreach Review

GridPP has always viewed outreach as an important task, both to potential users and the general public. As a consequence, GridPP has been one of the most visible national projects within both wLCG and EGI. During GridPP3 the project has maintained this position working with numerous organisations and initiatives.

The main focus for the project’s outreach is its users, the LHC and particle physics projects using the Grid within the UK. Interacting with this community at all levels is important and to this end the collaboration regularly exhibited at high energy physics meetings. These included; the annual IoP HEPP group meetings, the CHEP events in Prague and Taipei and the wLCG collaborations workshop in London. GridPP is also embeded in the UK e-Science community as one of the larger projects. As such it is instrumental in encouraging wider use of Grid by UK-based researchers. Since the start of the project the collaboration has had a constant presence at the UK All Hands e-Science meetings. Generally this is in conjunction with the STFC e-Science unit but at the events in Oxford and Cardiff the stand was run alongside the NGS.

As the UK moves closer to the creation of an NGI the outreach activities of GridPP within the EGI project have begun to work closely with the NGS. This has meant that during the GridPP3 project the two have had a joint stands at the major EGEE/EGI events, beginning with the EGEE User Forum (during GridPP2). The two projects continue to work together on events and activities that benefit both.

Due to GridPP's important role within the success of the LHC project it has worked together with STFC and others to ensure Grid input into various activities. One of these activities is the LHC Promotion Advisory Group (LPAG), this has representatives from STFC, the 4 experiments as well as the GridPP dissemination officer. Through this group and the fact the LHC was switched on twice during GridPP3, the project was involved with many related events or exhibitions, including a 2 month long exhibition at Science Oxford and various switch on events in 2008 and 2010. The dissemination team however also organised and attended events separate from those related to the LHC. These included 3 events at the British Science Festival (2008 and 2009), a "Meet The Scientist" event at the Museum of Science and Industry in Manchester and many invited talks around the country.

Over the 3 years the project has also created much dissemination material. This has ranged from news items on the project website, stories for EGEE and iSGTW; the Magic Cubes, brochures, presentations and posters. The dissemination team has also maintained the design and content of the project website including a major redesign at the beginning of GridPP3. All of these were designed with specific audiences in mind, in the case of this website this included a major overhaul of the "Help" section midway through the tenure of the project to make finding information easier for new or existing users.

Throughout the last 3 years GridPP has been a very active participant in all aspects of the international Grid community. This has included dissemination and outreach. Working alongside partners both in the UK and abroad, the project is one of the most recognisable brands within Grid. This is a testament to both the ability of the dissemination activities but also the technical work done by the project. Throughout GridPP4 the project will continue to provide support to users and partners in disseminating success stories while also encouraging greater use of the infrastructure by all users.

## **Conclusion**

GridPP3 successfully developed a large-scale computing Grid in the UK that was able to handle the demands of the LHC experiments, both in terms of handling the real-time flow of data off the LHC detectors during data-taking, and in providing a reliable and performant simulation and analysis infrastructure for UK physicists and their international collaborators. Many hurdles were overcome and the lessons learnt are described in an accompanying document.