

GridPP
UK Computing for Particle Physics

GridPP Project Management Board

Project Map

Document identifier :	GridPP-PMB-139- ProjectMap.doc
Date:	6/09/2009
Version:	1.0
Document status:	FINAL
Author	Sarah Pearce

Introduction

Progress on the GridPP project is monitored by means quarterly reports, which feed into of the 'ProjectMap', a spreadsheet showing the status of each area of the project. Workpackages in the project, such as the services delivered through the Tier-2s, are monitored through a mixture of:

- metrics – figures reported quarterly that show whether a given area is operating as expected, for example, the efficiency of MC production for LHCb, or the number of security incidents reported in the quarter. Thresholds for metrics are defined at the start of the project and revised regularly.
- milestones – tasks expected to be complete by a target date, such as the move to a new building at the Tier-1. These are either 'not yet due', 'complete' or 'overdue'

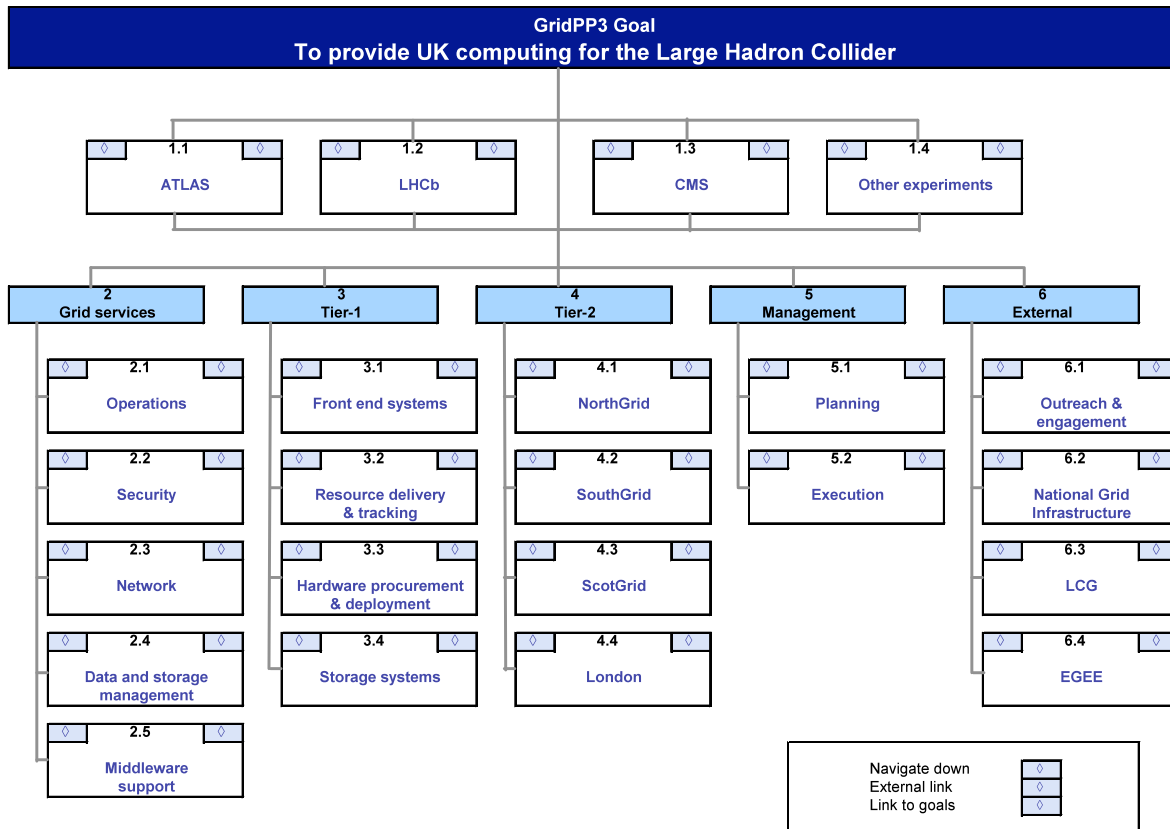
The ProjectMap is updated quarterly using the inputs from quarterly reports in each area. At the end of each quarter, workpackage leaders are prompted by the project manager to submit quarterly reports. First versions of these should be produced one month following the end of the quarter, and are then iterated by the Project Manager before incorporation in the ProjectMap. At this stage metrics and milestones can also be amended, for example to suspend metrics that are temporarily irrelevant to the project, or alter milestones to reflect changes in circumstances.

GridPP3 Project Map

For GridPP3, the ProjectMap reflects the emphasis on running, rather than building, a production grid. This is primarily shown in two features. Firstly, project elements are mostly tracked through metrics rather than milestones, as most elements of the production grid are now in place. Secondly, the structure of the map shows how the elements of the grid provide a service to the experiments. The overall structure is given below – key features are:

- Metrics for the experiments are in boxes along the top of the map. These metrics are defined and measured by the experiments, and show whether GridPP is delivering an acceptable level of service. The 'Other Experiments' box is reported by the Chair of the User Board and summarises the status for the smaller experiments.
- The left hand column is entitled 'Grid services', and, in addition to overall operations metrics, includes services such as security, networking and storage that support the grid as a whole.
- Columns in the centre show the status of operations at the Tier-1 and Tier-2s. Tier-2 figures are reported for each Tier-2 (i.e. not disaggregated per site), as it is the Tier-2s rather than individual sites that are responsible for providing services to GridPP
- The columns on the right show progress of project management (for example, numbers of meetings held), interaction with external projects such as EGEE and LCG, and movement towards an NGI, as well as dissemination metrics.

370 milestones and metrics have been defined in total, each of which is reviewed every quarter.



The version of the ProjectMap discussed here is based on quarterly reports from Q209, from April – June 2009. This is the latest quarter that has been reported to date. The ProjectMap for Q209 is online at http://www.gridpp.ac.uk/pmb/ProjectManagement/GridPP3_ProjectMap_5.xls

Metrics or milestones in the new ProjectMap can have the following status:

- Metric OK (green)
- Metric close to target (orange). This is a new status introduced as a trial in GridPP3. It highlights metrics where the value of the metric is within a small range of the target, so is not critical but might need further monitoring.
- Metric not OK (red)
- Metric not able to be measured (lilac). These are metrics that have been identified as important to the project, but where the infrastructure is not yet in place to enable them to be sensibly measured, for example the ticket response time at the Tier-2s.
- Milestone achieved (dark green)
- Milestones overdue (dark red)
- Milestone not yet due/ metric not applicable this quarter (grey)
- Suspended (black). This is for metrics that are no longer suitable and may be dropped or replaced.
- Awaiting further input. Here the targets may not yet be well defined. Metrics are refined as the project progresses, in particular to reflect external changes.

As of the end of August 2009, the overall status of the ProjectMap is as follows:

Metric OK	Metric close to target	Metric not OK	Not able to be measured	Milestone achieved	Milestone overdue	Milestone not due/ metric n/a	Suspended	Awaiting input	Total
184	22	27	3	57	4	58	12	3	370

Metrics

Metrics in the GridPP3 ProjectMap have been set to be challenging, reflecting the expectations of a high service level when running a production Grid. At any point it is expected that some of these will not be met for individual quarters, but overall the number of 'red' metrics should be decreasing as the project progresses. The table below shows the metrics for each version of the ProjectMap, illustrating a gradual increase in the number of 'OK' metrics.

	Q208	Q308	Q408	Q109	Q209
Metric OK	99	142	155	172	184
Metric close to target	24	47	39	32	22
Metric not OK	41	32	32	21	27
Not able to be measured	27	22	11	10	3
Milestone achieved	11	22	32	42	57
Milestone overdue	2	7	13	17	4
Milestone not due / metric n/a	101	80	69	60	58
Suspended	0	6	6	9	12
Awaiting input	34	5	12	10	3
Total	339	363	369	373	370

At present there are 27 metrics that are 'red' - not met, and not close to the target. The key issues resulting in red metrics are as follows:

Experiments

LHCb saw a variety of issues in Q209 with their application software that resulted in three red metrics (1.2.2: MC production (generation) efficiency in the UK, 1.2.3: T1 MC production (reconstruction, stripping) efficiency and 1.2.4: T1 MC/Event user analysis - UK share/ efficiency). The LHCb application problems were due mainly to memory leaks not picked up in tests at CERN. This led to software hanging without consuming CPU before it was killed by the watchdog, after about 1 hour. Other problems included jobs stalling when events of unusually large size were seen by the application, and issues with the new version of the POOL software causing the LHCb LFC instance at CERN to be overloaded, leading to jobs stalling or failing. These problems led to the reconstruction and stripping being stopped without running all the needed jobs - this additionally leads to lower efficiency as a natural result of low statistics. These problems are all being worked on by LHCb, who note that the support at the UK Tier-1 and Tier-2 sites for Monte-Carlo production has been excellent and the arrival of new staff has improved communications between the sites and the experiment (see GridPP-PMB-138-Status).

The ATLAS quarterly report had no red metrics. This is encouraging, although it is worth noting that of the 8 quantitative metrics for ATLAS, 6 are amber. So, while no individual metric is currently considered critical, the overall situation is being monitored closely.

Tier-1

At the Tier-1, a number of items of scheduled work led to downtime, resulting in red metrics for availability of the CE service (3.1.8) and CASTOR SAM tests for the LHC VOs (3.4.8) and LHCb SAM tests uptime T1 (1.2.11). These were due primarily to three factors: the move to a new building for the Tier-1 machine room (R89), scheduled migration of core RAID hardware on CASTOR, and some network issues following a scheduled

network development. Reliability over the same period, which takes into account scheduled outages, was 99% and met the target.

Tier-1 deployment of disk and CPU in Q209 did not meet the allocations set by the User Board (3.4.4, % met of UB Allocation for Disk and 3.2.9 % met of normalised UB allocation for CPU). The experiment requests are usually compiled every quarter and then matched to the Tier-1 capacity (anticipated over next 2 quarters) and WLCG pledges. There were changes during this period, in particular the move to R89, that resulted in a decreased disk and CPU availability in June compared with what had been planned.

Tier-2s

Site performance is monitored through SAM tests and GridPP has been actively recording these in its own database. GridPP has also produced its own tests to monitor the performance of storage elements at each site, and to run ATLAS jobs at the sites. In Q209 availability and reliability metrics were above the challenging target of 95% for all the Tier-2s apart from LondonGrid, where there have been problems at UCL with their CE and Lustre file system. In addition, LeSC has had few staff and has now been closed as a Tier-2 site. Because the metrics are a straight average over all the sites in a cluster, the UCL results bring down the London results, but the large London clusters are performing well, with Brunel, QMUL, RHUL and UCL-HEP all meeting availability and reliability targets. (4.1.3: Average SAM availability performance over the last quarter, 4.1.4: Average SAM reliability performance over the last quarter, 4.1.5: Average SLL untargeted ATLAS test performance, 4.1.6: Average SLL SE test performance).

In SouthGrid, the % of disk used was lower than the target 20% (4.3.9) due to an additional 380TB added at RALPPD that had not yet been allocated. SouthGrid also failed the metric looking at SLL ATLAS test performance, due to low results from Oxford and Cambridge (4.3.5).

ScotGrid had two red metrics in Q209. The first of these (4.2.5, average SE test performance) was due to jobs no longer running at ECDF. Capacity at ECDF is purchased by the Tier-2, and funds for this period have now expired. The management team are working with ECDF to find a workaround solution acceptable to all parties involved. No decisions have been made yet, but it may be that ECDF will only run GridPP jobs again when more funding for its participation is found. Nevertheless, ScotGrid continues to contribute well above its pledged share of CPU and storage. The second ScotGrid red metric (4.2.9, percentage of disk used) was due to a very substantial disk space upgrade at Glasgow, with little ATLAS data to fill it. Usage is increasing steadily now.

GridPP-wide issues

Demand for CPU resource has been generally low across WLCG. This reflects the change in the LHC schedule and it is expected that this will correct itself once data is flowing; in the meantime we have tried to encourage more support for existing and active VOs such as Camont and Biomed. Metrics reflecting low farm occupancy are 2.1.3 (GridPP Fraction of KSI2k used), 3.2.11 (Tier-1 Farm Occupancy), 4.1.7 and 4.3.7 (Approx. CPU utilisation (wall clock time) at LondonGrid and SouthGrid), 4.1.8 and 4.3.8 (Approx. CPU utilisation (CPU time) at LondonGrid and SouthGrid). These figures are similar to those reported at the last GridPP Oversight Committee meeting, although NorthGrid has seen an improvement in utilisation.

Job success rates across GridPP (2.1.6, Job success rates) are consistently lower than the target of 95% (85.5% in Q209), although this figure increased slightly in the last quarter. The GridPP test jobs that determine this figure have less priority than ATLAS production and so when sites are full these jobs are more likely to queue beyond the pass cut-off time. Therefore, busier sites means more failures; as we do more real work and the experiments get busier we expect this test to show declining success rates. For a general user this may simply mean that the job takes longer to execute.

There were also a number of less significant metrics that were not met, including the number of management meetings held at NorthGrid, the number of Collaboration Board meetings, status of the deployment web pages

and updates to the LHCb GANGA user training material. NorthGrid is being encouraged to meet more regularly, and it is likely that Collaboration Board meetings will be required as discussions on GridPP4 commence. A review and update of the static deployment web pages has started but was not complete by the end of the quarter. For LHCb GANGA, the material will be reviewed in September '09 in preparation for the next round of training.

The last GridPP OC reviewed several red metrics that are now improved and no longer red. These include recruitment and % of staff in post both at the Tier-1 and Tier-2. Document GridPP-PMB-140-Resources sets out more details, but both RAL and Tier-2s are now at nearly their full complement of staff. CASTOR reliability is also improved, with red metrics in this area caused by scheduled downtime rather than crisis management.

Milestones

As expected, the number of completed milestones continues to increase as the project progresses. Of those that have passed their target date, 57 are achieved and only 4 overdue. At the time of the last OC, there were a number of milestones overdue relating to the move to R89 and delivery of new hardware. These are now all complete. In addition, some other milestones have been rescheduled to reflect the new timetable agreed by WLCG, as a result of delays to the LHC. The overdue milestones are:

3.2.16 Tier-1 Disaster and Business Continuity Plan Available and 3.2.18 Disaster Plan fully implemented. As described in GridPP-PMB-138-Status report, a new disaster management system is operational at the Tier-1. This is a four stage system with well defined actions initiated at each stage. The system manages potential disasters in a pro-active rather than reactive manner and has recently been used to handle a number of matters; most recently the Tier-1's response to the current Swine Flu pandemic. The system is working very well, but some contingency plans remain to be completed so this milestones is outstanding.

3.1.22 LHC Monitoring infrastructure operational at RAL. This is being installed by Dante: installation by their engineers is complete but they have yet to commission the service.

3.4.21 General ADS Service Ends. This has not been a priority for the Tier-1 during a period of major activity and the move to R89. However, a plan has been provided to the UB and the closure process has started. The service is now read only and is expected to be terminated before the end of the year.

Risk Register

For GridPP3, the risk register was redesigned to focus on the main risks and make it more streamlined to update. As before, each individual risk is attributed two numbers, a likelihood and an impact, each out of 4. The product of these numbers defines the risk level. Risks between 0 and 4 are regarded as low; between 5 and 8 as medium; and between 9 and 16 as high. These are colour coded green, orange, and red respectively. The current risk register is online at http://www.gridpp.ac.uk/pmb/ProjectManagement/GridPP3_RiskRegister_2.xls

A summary of the risk register is shown below:

ID	Name	L	I	Risk	Owner
R1	Recruitment/retention difficulties	3	3	9	SP
R2	Sudden loss of key staff	2	3	6	SP
R3	Staff time not devoted to GridPP	2	2	4	SP
R4	Service insufficiently resilient wrt data flow	2	4	8	JC
R5	Service insufficiently resilient wrt storage	2	4	8	JC
R6	Service insufficiently resilient wrt CPU	1	3	3	JC
R7	Failure to meet Hardware Pledge	4	2	8	DB
R8	Failure to match the expected Service Levels	2	3	6	DB
R9	Hardware procurement problems	2	3	6	AS
R10	Hardware resources inadequate/insufficient	2	4	8	GP
R11	Additional delays to the LHC	3	2	6	SP
R12	Machine room problems compromise Tier-1	4	3	12	AS
R13	Physical integrity of the Tier-1	1	4	4	AS
R14	Network/OPN breakage	2	4	8	PC

ID	Name	L	I	Risk	Owner
R15	Lack of funding for Network Provision Costs	2	4	8	PC
R16	Sub-components not delivered to project	2	2	4	TD
R17	Software maintainability problems	2	3	6	TD
R18	Technology shifts	2	3	6	TD
R19	Security inadequate to operate Grid	2	3	6	DK
R20	Changes in experiment analysis model	2	4	8	GP
R21	Insufficient effort to support VOs	3	2	6	GP
R22	Users stop production tasks from running	2	3	6	GP
R23	Insufficient/ inadequate user support	2	3	6	SP
R24	Not enough work to fill resources	4	1	4	JC
R25	Poor communications within and outside GridPP	2	3	6	DB
R26	Poor relationship with university computing support	2	2	4	SL
R27	Uncertain long-term funding	4	1	4	DB
R28	Future European Grid initiative fails to meet GridPP needs	2	2	4	RM

Of the 28 risks in the register, there are currently two that score 9 or higher:

R1: Recruitment and retention difficulties. This is currently rated 9, with likelihood and impact both 3. Recruitment since the last OC has led to this being reduced from 12, but it is still seen as a significant risk. The SLA posts are now only 1FTE short of their full complement, with all posts currently filled at the Tier-2s except for LondonGrid. However, several key staff have left this year, and although they have been replaced staff turnover is an issue. GridPP will try to establish future funding early to aid with retention and allow earlier recruitment of new staff, but this depends on funding bodies. Crisis strategies include offering higher salaries and using contract staff. The PMB are keeping close oversight of this risk, with regular reports from the Tier-1 and Tier-2s.

R12: Machine room problems compromise Tier-1. This is rated 12 (likelihood 4, impact 3). The handover of the new R89 building and machine room to STFC (and acceptance by STFC) was considerably delayed; the building was finally accepted in May 2009. Migration of (most of) the Tier-1 hardware to the new building commenced on 22nd June and was completed on schedule on 6th July. However, two separate issues appearing in the week of Aug 10th (an air conditioning failure and a small water leak due to condensation) triggered the Tier-1 Disaster Management protocol.

The GridPP PMB has reassessed all risks on the risk register since the last OC, leading to the following changes:

R5: *Service insufficiently resilient wrt storage*. This has been reduced from (3,3) to (2,4). Issues of service resilience have been explicitly addressed since 2008 and CASTOR is now substantially more stable, with a

consequent reduction in the likelihood. However, as data-taking approaches the impact of storage failure increases.

R7: *Failure to meet Hardware Pledge*, has been increased to (4,2). A problem in the acceptance tests of half of the 2009 disk purchase is being tracked by the Tier-1 Disaster Management protocol. This issue means that the Tier-1 is unlikely to meet the hardware pledge. However, the impact is currently low because the resources are not required on the timescale in which we expect the problem to be resolved.

R11: *Additional delays to the LHC*, has been raised to (3,2). Likelihood of this remains high, but the impact has increased slightly. Ongoing delays to the LHC are having an effect on staff morale and on the funds that will be available for staff and hardware once the LHC starts.

R14: *Network/OPN breakage* has been raised to (2,4), and is felt to be of more concern than when previously assessed. The long LHC data run planned reduces the chance of recovering, or catching-up, from a significant outage; the volumes of data to be transferred (particularly the T1-T1 traffic that also uses the OPN) have been demonstrated by STEP09 to be larger than previous anticipated; the incidence of outages has proved to be larger than expected from earlier experience and the threat to the UK reputation, being the only major Tier-1 without a backup link, is now increased. GridPP proposes to spend £52k of our existing hardware budget to install a backup link, supported by a recurrent cost of between £40k and £60k per annum, depending on negotiations about the end-point costs.

R16: *Sub-components not delivered to project*. The impact of this has been reduced from 3 to 2. A functioning set of software is now available, and any areas that are missing are felt to be minor, for example in monitoring tools.

Conclusions

The GridPP3 ProjectMap tracks the status of project metrics and milestones. Overall the project is considered to be progressing very well, with 206 metrics at our close to the target. Of the 27 metrics that are 'red', current issues include low CPU utilisation, which is expected to be solved once LHC data is flowing, problems with LHCb application software, and low availability at the Tier-1 due to scheduled downtime in advance of LHC start. Several metrics that were red at the time of the last OC are now green, including recruitment and % of staff in post both at the Tier-1 and more widely. The project has now achieved 57 milestones, with only four currently overdue. At the time of the last OC, there were a number of milestones overdue relating to the move to R89 and delivery of new hardware: these are now all complete.

Risks in the GridPP risk register have been reassessed by the PMB since the last OC. Of the 28 risks, two currently score 'high'. The first of these is 'recruitment and retention difficulties', which has reduced in risk since the last OC as more staff have been recruited, but is still considered a significant issue. The second high risk relates to potential problems with the new machine room at the Tier-1. However, the risk that the storage service is insufficiently resilient has been addressed and CASTOR is now substantially more stable, so this risk has been downgraded to 'medium'.