GridPP Project Management Board

Project Status

Document identifier:	**GridPP-PMB-172-ProjectStatus.doc**
Date:	**08/03/2015**
Version:	**1.0**
Document status:	**Final**
Author	**PMB**

Introduction

The GridPP project was reviewed by the Oversight Committee in November 2013 and by a Strategic Review committee in March 2014. This report provides a further update and summary of the project status.

The LHC ran successfully for 18 months until Long Shutdown-1 started in February 2013. During that time, 25 fb^{-1} of data was processed and analysed on the Worldwide LHC Computing Grid (WLCG) infrastructure, and the Higgs Boson was observed. The Initial announcement on July 4th 2012 was confirmed later in the year as the data set grew, and on 8th October 2013 Peter Higgs and Francois Englert were awarded the Nobel Prize for Physics. The LHC is scheduled to restart in early summer 2015 and good progress is being made with upgrades to the LHC and the experiments. In the meantime, the computing Grid infrastructure has continued to (re-)process data; perform analyses; and produce simulated data. The Grid infrastructure has also continued to grow and evolve in preparation for Run-2 in step with the experiment computing models. ATLAS and CMS are currently performing large-scale tests to validate preparations for Run-2.

GridPP4 is in the final year of a 4-year project. The 32nd and 33rd collaboration meetings were held with themes on “Mini Review: A look at where we are and where we're heading”, and “Preparing for Run 2”. A proposal for a four-year GridPP5 project was submitted and reviewed. However, funding constraints led, instead, to a one-year extension of the current project (to be called GridPP4+) and the GridPP5 proposal is about to be submitted.

International Context

WLCG now provides about 485,000 cores, 270PB of storage and processes in excess of 2 million jobs per day. GridPP currently provides about 11% of the total resources. Figure-1 below shows that the Tier-1 at Rutherford Lab continues to deliver 9% of the global LHC Tier-1 resources, as it has done since the start of data taking. Figure-2 shows 14% of the LHC Tier-2 resources were delivered by the UK over the same period, down slightly from 15% reported for 2011-2013. The difference between the Tier-1 and Tier-2 global fractions reflects the significant leverage of Tier-2 resources from the institutes that GridPP funding has enabled.

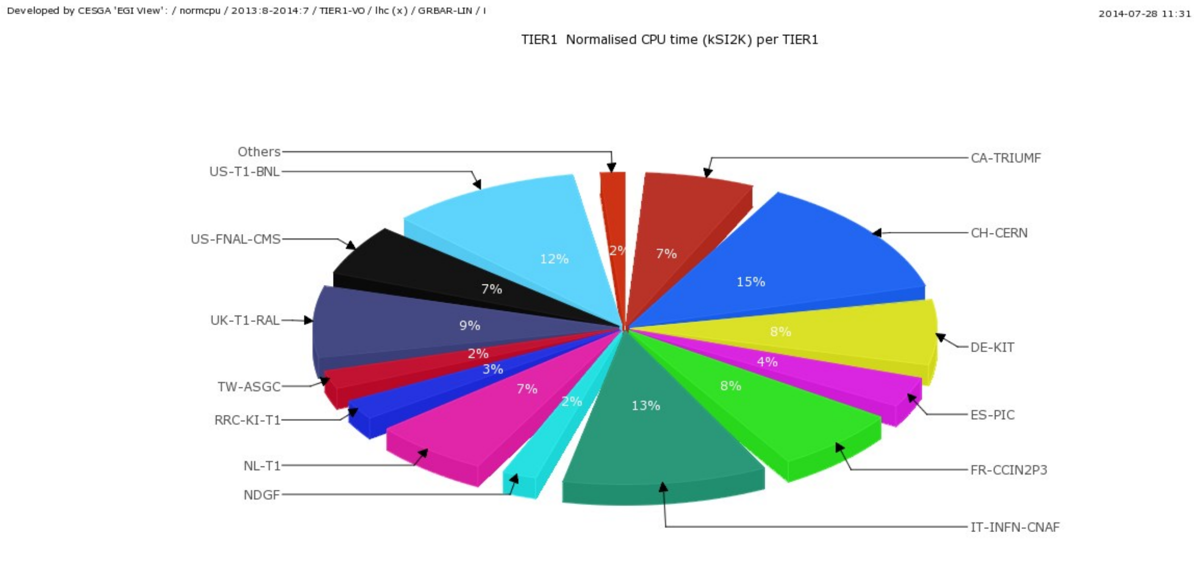


Figure-1: CPU delivered to LHC VOs (Aug 2013 to July 2014) by Tier-1.

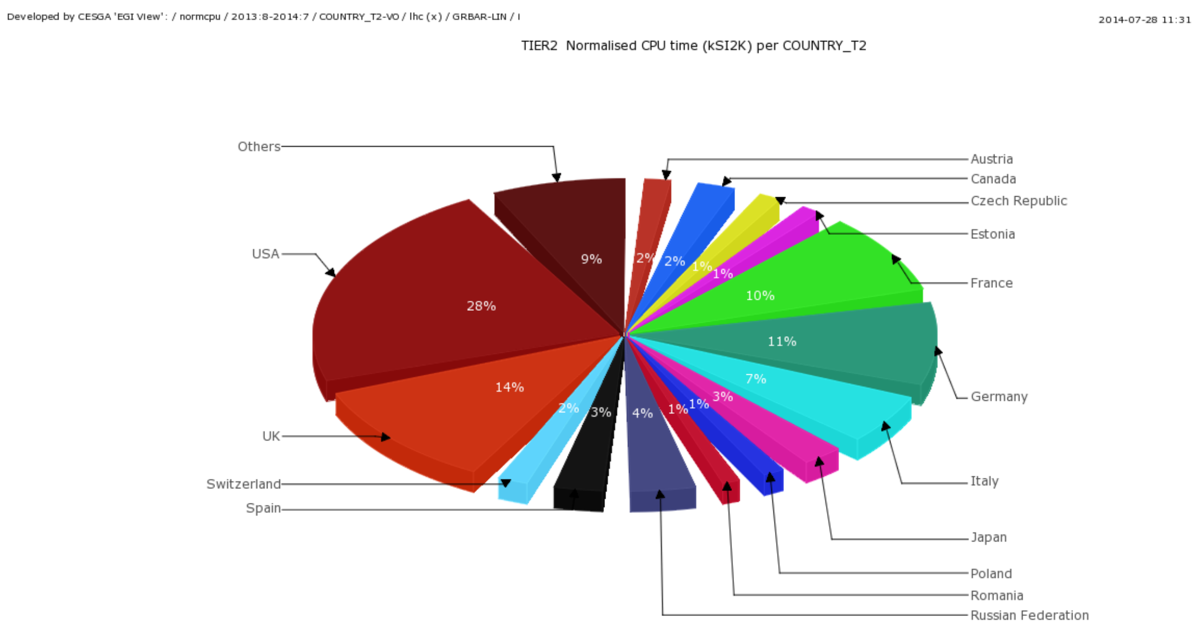


Figure-2: CPU delivered to LHC VOs (Aug 2013 to July 2014) by Tier-2s.

EGI InSPIRE was successfully completed in December 2014 and received the highest accolade, rated Excellent, in its final two reviews by the European Commission (EC). The successor project, EGI ENGAGE, was rated 15/15 and ranked top in its Horizon 2020 call and starts operations on 1 March 2015. The EGI Federation has developed a sustainable strategy to support the operation of core services required to support distributed computing by using its national subscription fee to co-fund these with the NGIs, as well as funding the EGI.eu organisation to provide overall leadership and management.

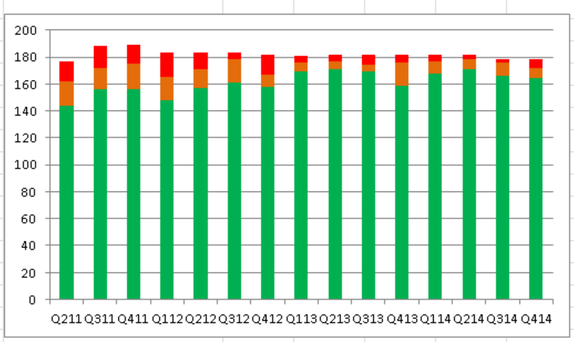
The UK receives support from EGI.eu to provide services that are critical to WLCG (APEL accounting; Grid Operations Centre Database (GOCDB); and leadership of international security operations and policy development). Additionally the EC funded EGI ENGAGE project is fully funding the development of these services to evolve to meet the future needs of its user community. The UK currently chairs the EGI Council and Executive Board to drive on-going discussions about the Governance model and fee-structure of the next phase of EGI.

GridPP is also engaged with and contributing to other H2020 initiatives. This includes the EU-T0 (an association of funding agencies who own substantial computational resources), and as an outcome of this the Indigo Datacloud project has been successful in gaining EC Horizon 2020 funding. GridPP also engages in initiatives to reduce overlaps and in moving towards a common e-infrastructure. This includes GOCDB which has been adopted and part-funded by the EUDAT project and security and access as part of the AARC Horizon 2020 initiative in collaboration with Geant Association. GridPP is also working with the HEP Software Foundation, (a collaboration of HEP software projects designed to promote standards, reduce overlaps and provide common utilities).

GridPP4 Status

Since the last OC meeting the GridPP4 project has been progressing well. The production systems at both the Tier-1 and 2 sites have mainly been running smoothly. Preparation for Run 2 at the LHC has included purchasing and commissioning new hardware to be ready to meet the April 2015 MoU requirements. The progress has continued to be monitored by Quarterly reports from the various Work Packages. The evolution of milestones and metrics over the period is shown in Figure-3 and the current Project Map shown in Figure-4. With the last quarter left before the end of the project we are on track to complete all active milestones by the end of Q1 2015.

	Q211	Q311	Q411	Q112	Q212	Q312	Q412	Q113	Q213	Q313	Q413	Q114	Q214	Q314	Q414
Metric OK	144	156	156	148	157	161	158	169	171	169	159	168	171	166	164
Metric close to target	18	16	19	17	14	17	9	7	6	5	17	9	7	10	8
Metric not OK	15	16	14	18	12	5	15	5	5	8	6	5	4	2	6
Not able to be measured	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Suspended	2	12	15	21	22	18	19	19	18	18	19	19	19	23	22
Awaiting input	13	2	0	0	0	0	0	0	0	0	0	0	0	0	0
Metrics Total	192	202	204	204	205	201	201	200	200	200	201	201	201	201	200
Milestone achieved	7	12	21	23	28	33	34	40	46	54	60	62	62	65	73
Milestone close to completion	0	0	0	1	1	0	0	5	3	0	1	2	4	3	2
Milestone overdue	1	3	3	4	2	2	1	1	0	0	0	0	1	1	0
Milestone not due/ metric n/a	76	65	54	50	46	46	46	36	33	28	20	17	14	12	7
Milestones total	84	80	78	78	77	81	81	82	82	82	81	81	81	81	82
Total	276	282	282	282	282	282	282	282	282	282	282	282	282	282	282



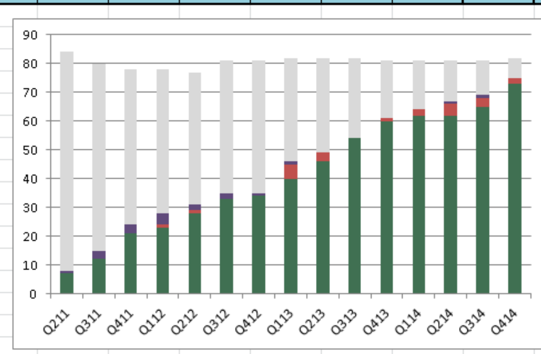


Figure-3: Evolution of Milestones and Metrics for GridPP4 to Q4 2014. The lower left plot shows the number of metrics that are currently met (green) and not met (red). The lower right plot shows the evolution of the number of milestones that have been met (red flags late ones). The data is presented numerically in the table at the top.

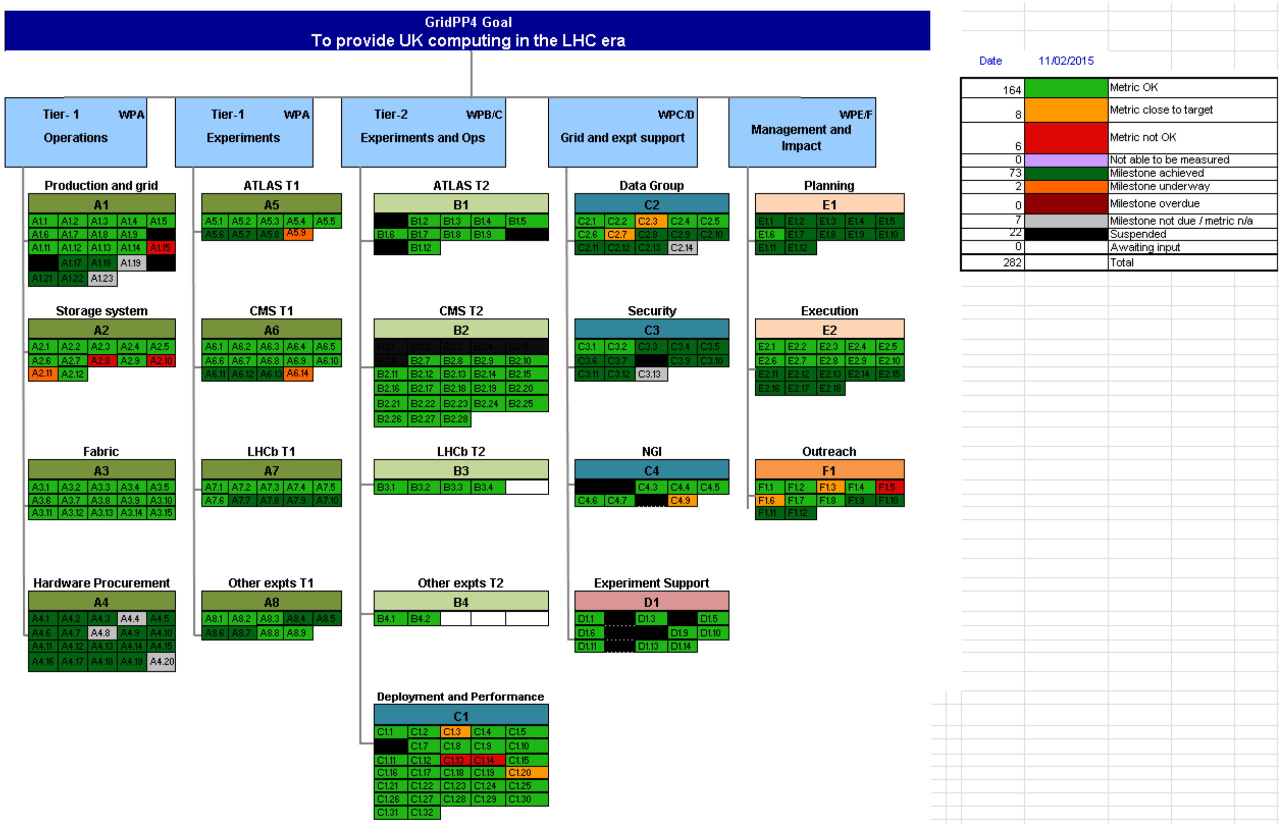


Figure-4: The Project map at the end of Q4 2014

Overall the majority of metrics are being met or are close to target, with just six metrics unsatisfied this quarter.
They are:

- A1.15 ‘Percentage of GridPP4 Staff in post’, Fell to 81% during Q4 2014 after some staff departures and recruitment delays caused by the GridPP5 funding uncertainty. (The target level is 93%) Bookings are projected to increase in Q1 2015 to 89%
- A2.8 A2.10 & A2.11 A variety of problems affected availability during Q4 2014 which caused the CASTOR SAM tests to drop below target for ATLAS, CMS and to a lesser extent LHCb VOs. In December network problems, DNS problems and emergency downtime for security patching were the major causes of this.
- C1.13 & C1.14 ScotGrid CPU utilisation (Wall time and CPU time) both very low this quarter at 39 and 35%. Experiments switching services for Run-2 caused a lack of jobs.
- F1.5 ‘Number of GridPP press releases this year’ is zero. Our target is one. This is partly down to a quieter period of running during the shutdown. However, it also reflects the continued focus on the Impact Agenda, rather than simply dissemination. The GridPP Dissemination Officer appointed in October 2013 has been active in promoting the use of the Grid to small VOs such as ‘CERN@School’ and the University of Central Lancashire GalDyn galaxy simulation group. Greatly improved documentation for the tools to support smaller VOs have been provided on the GridPP Web and wiki pages.

Four milestones have been delayed till Q2 2015 at the earliest. These are:

- A1.19 ‘External review of Operations’ Due to the considerable uncertainty regarding the long term future of the Tier-1 during GridPP5 this will be delayed until at least mid-2015.
- A1.23 ’Strategic and operations plan agreed’ As above this awaits the outcome of the GridPP5 proposal.
- C2.14 'Argus deployment and integration with Storage', delayed due to the slow uptake of Argus by the sites across WLCG not just GridPP.
- C3.13 Security recommendations for the future. With the extension to GridPP4 by one year and EGI-InSPIRE by 8 months it makes sense to delay planning for the future until spring 2015.

The quarterly reporting has continued to be useful to track issues and focus the different groups on their priorities. The overall performance of the different work packages in GridPP has been very successful and this has been reflected in the project map and reports. All detailed quarterly reports are available on the web at:
http://www.gridpp.ac.uk/pmb/ProjectManagement/QuarterlyReports/reports.html

Risk register

The GridPP4 risk register summary is presented below in Figure-5 and the full version is available at:
https://www.gridpp.ac.uk/pmb/ProjectManagement/GridPP4_Risk_Register_v2-Q414.xlsx .

Ref.	Risk Description	Owner	Inherent Risk			Residual Risk		
			Likelihood	Impact	Total	Likelihood	Impact	Total
1	Castor Storage System Problems	AS	0.4	60	24	0.3	60	18
2	Expt. s/w runs poorly on the grid	PG	0.4	60	24	0.5	50	25
3	Middleware at T2s performs badly for user analysis	SL	0.4	60	24	0.1	60	6
4	Outage of UK T1	AS	0.4	60	24	0.1	50	5
5	Loss of experienced personnel at T2s	SL	0.7	50	35	0.6	40	24
6	Failure of T1 to meet SLA or MoU	AS	0.6	50	30	0.2	50	10
7	Security problem affecting reputation	DK	0.7	80	56	0.5	60	30
8	Loss of GridPP service due to security	DK	0.3	50	15	0.1	50	5
9	Insufficient manpower to operate core T2 sites	SL	0.6	40	24	0.4	40	16
10	Recruitment retention problems at RAL	DK	0.9	60	54	0.6	50	30
11	Hardware costs aren't accurately predicted by model	DB	0.25	50	12.5	0.3	20	6
12	Core service funding insuffcient	CD	0.5	40	20	0.1	10	1
13	Insufficient travel funds	DK	0.4	25	10	0.15	25	3.75
14	Expt. Computing models change	PG	0.4	40	16	0.1	40	4
15	Breakdown of NGI/EGI infrastructure	JC	0.3	30	9	0.3	30	9
16	Critical middleware no longer supported	DC	0.4	40	16	0.1	40	4
17	Significant loss of custodial data at the T1	AS	0.1	75	7.5	0.1	50	5
18	Unplanned infrastructure costs	PG	0.25	30	7.5	0.2	10	2
19	Loss or damage to hardware at T1	AS	0.05	100	5	0.05	100	5
20	Disaster at T1 leads to prolonged outage	AS	0.05	90	4.5	0.05	80	4
21	Insufficient funding at T2s for h/w	SL	0.4	40	16	0.3	30	9
22	Loss of EGI.eu	CD	0.5	20	10	0.4	15	6
24	Insufficient Network Bandwidth	PC	0.4	50	20	0.3	50	15
25	Technology Mismatch	DC	0.2	50	10	0.05	50	2.5
26	Insufficient support effort	PG	0.4	40	16	0.2	20	4
27	Financial Uncertainty	DB	0.2	50	10	0.5	60	30
28	Failure to deploy or operate hardware	JC	0.5	70	35	0.3	70	21
29	Conflicting opinions amongst GridPP stakeholders	DB	0.3	50	15	0.3	50	15
30	Over contention for resources	DB	0.3	40	12	0.1	40	4
31	Capital vs Resource	DB	0.8	50	40	0.3	60	18
32	Tier-1 CEPH Project fails	AS	0.8	60	48	0.6	60	36
33	GridPP may fail to broaden its customer base.	JC	0.5	60	30	0.3	60	18

Figure-5: The GridPP4 Risk Register Summary

The risk register is reviewed regularly by the PMB and the key current risks for the next 6 months are:

7 Security problem affecting reputation.

There is a continuous need to be vigilant on all security issues. Security problems may bring down the whole of GridPP or even WLCG, given that all Sites run very similar software. A major outage is likely to attract significant outside interest and GridPP is likely to suffer damage to its reputation. The Tier-1 in particular could be an attractive target for 'Bit Coin Miners'. Attacks of this nature have become common in the last 6 months. We will continue to act on many fronts to reduce the threat by ensuring software is up to date and to contain and handle incidents quickly when they occur. The UK (GridPP) leads the Joint Security Policy Group and the security vulnerability group. The security team operates across GridPP sites to promote best practice and many staff regularly benefit from security training as part of HEPSYSMAN events. We are pleased we have filled the security officer position, to coordinate all security issues and lead the distributed team.

10 Failure to retain or recruit key technical staff at RAL
Losing key staff could impact the projects ability to meet deliverables and key milestones. This is a potential problem that has to be dealt with rapidly by STFC management by speeding up recruitment in the event. It has been noted that recruiting new staff is increasingly difficult due to non-competitive salaries and conditions. Retention has been affected by the uncertainties in GridPP5 funding.

27 Financial Uncertainty
The one year extension of GridPP4+ was welcomed but the uncertainty over the future as we prepare to submit the bid for GridPP5 causes considerable concern. If funding cuts are likely there is a strong possibility that this will provoke staff to leave before the end of GridPP4+.

32 Teir-1 CEPH project fails
If the CEPH project fails during scale testing, we would have to fall back to CASTOR which we expect would not cope with the increased load. If scale testing is successful and CEPH is put into production insurmountable problems may arise leading to an unacceptable degradation of performance or data loss.
Slow and careful testing and early test users will be used to mitigate

Risks that are high but under control are:

5 Loss of experience personnel at Tier-2s
This risk is mitigated by sharing knowledge across sites, but it is important to keep the staff informed of future funding plans to reduce staff loss towards the end of the project. Uncertainty about GridPP5 and possible delays in contract extensions increase the risk at sites, although the risk the overall project is mitigated by the distributed nature of the Tier-2s. If severe cuts are made some key sites will have staff reductions.

6 Failure of Tier-1 to meet SLA or MoU commitments
Failure to meet our WLCG obligations would damage our reputation. Recent availability metrics dipped below the targets set by WLCG. Many systems have been put in place to cover call out, change management, disaster recovery and hardware procurement. These continually focus operations staff on the MoU commitments to ensure excellent performance. However recent staff losses have made this difficult.

28 Failure to deploy or operate hardware
This risk has increased due to the concern the changes in BIS requirements may delay or complicate the next procurement round, which could result in late delivery of equipment. This would make meeting the WLCG pledge levels difficult.

31 Difficulty with budgets due to Capital vs Resource limitations
Classification of different types of equipment as capital or resource can sometimes change causing problems with budgeting. It is important to have good frequent communications with STFC to track such changes and limitations.

33 GridPP may fail to broaden its customer base.
If so GridPP would not benefit from shared infrastructure cost savings. There would be reputational

damage, and lack of endorsement that the GridPP infrastructure service is relevant to a wider community.

Tier-1 Status

Overview

This section of the report covers the period from August 2014 to January 2015.

During this period, the priority areas of work have been:

- On-going routine stable operation of the production service and preparation for LHC Run 2 in 2015
- Procurements and deployment of hardware to meet the 2015 MoU commitments
- Recruitment of staff replacing those lost during the uncertainty in 2014 caused by the GridPP-5 renewal.
- Seek to broaden the funding base by engaging with new user communities and funding opportunities.
- Continuing development of new services (such as cloud and object store) in order to meet potential needs of new user communities

Fabric and Infrastructure

Tenders were carried out for disk and CPU capacity to meet 2015 MoU commitments and orders placed (5PB of disk and 44KHS06 of CPU) in November. Hardware prices were disappointingly high (particularly CPU) and expenditure on disk had to be reduced in order to meet the CPU MoU commitments. Nevertheless disk MoU commitments can still be met, the disk reduction being at the expense of the expanding prototype CEPH storage system. Hardware delivery is scheduled for February for commissioning in April.

In September repacking to T10KD of the T10KA and T10KB generations of tapes commenced and successfully completed in early January. Approximately 8000 tapes containing 5PB of data were repacked, in approximately 4 months, 2 months ahead of schedule. The tape service is now operating on only the T10KC and T10KD generations of tape drive, LHCb and ATLAS on the C media and CMS and other experiments on the D media. It is likely that we will also need to migrate LHCb from C to D media in 2015 in order to meet overall capacity requirements at minimum cost.

In December 2014 a review of hardware reliability and functionality concluded that the 2009 CPU generation (18KHS06) could be run for a further year (6^{th} year) in production. This will provide a short term boost to capacity but will not affect the long range purchase planning which continues to assume a 5 year life cycle.

The R89 machine room has provided a stable operating environment during this period. Electrical safety testing was completed on machine room power distribution infrastructure. This involved shutdown of each machine room PDU, two at a time, over a six day period. This was a major undertaking which required careful planning and was accomplished without operational impact on the Tier-1 service.

The departure of two Fabric team staff has severely impacted the Tier-1's ability to progress necessary developments. Particularly challenging has been the loss of network expertise and effort within the team. Consequently the Tier-1's project to reroute its data traffic flow from its legacy 10Gb/s site link to its new 40Gb/s routing layer has been delayed but is planned to be completed in time for LHC Run 2. Interviews have been scheduled for recruitment of replacement staff. Similarly there have been delays in IPV6 development, partly owing to staff loss and partly owing to limited effort in site networking team to support this development activity when tensioned against other site priorities. It is likely that the Tier-1 will miss a WLCG milestone to have a minimal IPV6 dual stacked monitoring infrastructure in place by April 2015.

Production, Operations and Service

Tier-1 operations were remained reliable over the 3^{rd} Quarter of 2014 but were impacted by several issues in Q4. Nevertheless average RAL (ops test) service availability since August 2014 remained high at 95% compared to a WLCG target of 97%. Average RAL availability since August 2014 for the LHC VOs was 99%, comparable to the previous 12 months covered in the previous report. Reliability broken down by VO was follows:

	3Q14	4Q14
ALICE	100%	99%
ATLAS	98%	97%
CMS	100%	96%
LHCb	100%	98%

Figure-7: Quarterly VO availability at Tier-1

Middleware, Grid and Cloud Computing

CPU utilisation fell slightly this period after record levels in the first half of 2014 as we undergo preparations for the start of Run-2. Farm occupancy averaged 78% (compared to 84%) over the period and job CPU utilisation efficiency fell to 80% (88%). However since August 2014, RAL delivered 10.6% of global LHC Tier-1 CPU computing – a slight increase in our share compared to the previous report. (Figure-8)

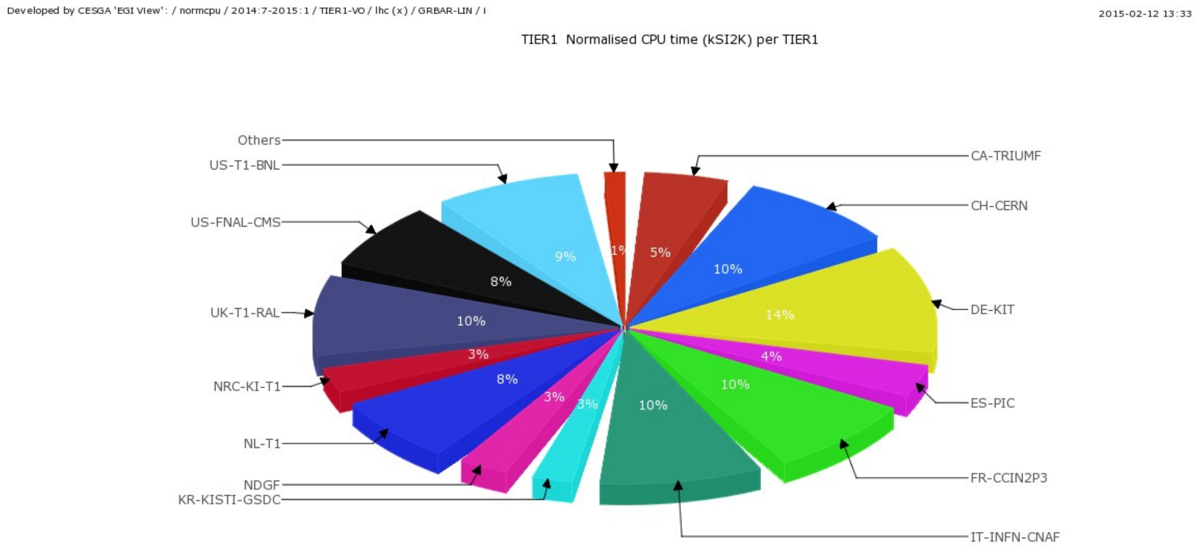


Figure-8: Global Tier-1 usage August 2014 to January 2015

The CEPH storage testing is now fully underway. CEPH is a strategically important project to the Tier-1. If successfully deployed, CEPH will provide an industry standard replacement to the domain specific CASTOR disk system currently deployed. By providing industry standard access protocols, CEPH is more attractive to non-HEP user communities, offering improved opportunity to cost share. Considerable challenges remain to be solved before CEPH is demonstrated to meet the LHC experiments' requirements and operate reliability at large scale. The CEPH project is currently the highest risk item on the GridPP risk register. An additional member of staff has been recruited to work solely on CEPH and priorities of several other members of staff have been raised to work in this area. Work is underway with the experiments to evaluate a test CEPH instance after which a considerably larger deployment will be required to test CEPH at scale. This work is currently entirely funded by GridPP but there is considerable interest in this project from several other communities and we are looking for other funding partners to support this work jointly.

A member of staff is now working full time to deploy a production quality private cloud. Work is also underway with the LHC experiments to test this cloud. This work is part funded by GridPP and part funded by STFC Scientific Computing work. The intention is to be able to enable non-PP projects to exploit the Tier-1 service through industry standard interfaces.

Management, Business Processes and Communications

The Tier-1 team were partners in the successful Horizon 2020 bid: Datacloud (as a member of the EU-T0 project). This work will be closely related to the Tier-1's existing cloud activity. The team were also involved in the ZEPHYR bid which unfortunately just failed to be funded in this round. Work is also underway with LIGO, LOFAR and LSST to explore the possibility of these collaborations exploiting Tier-1 resources.

The Tier-1 manager is currently working reduced hours (50%) owing to other temporary commitments within Scientific Computing Department. This has impacted several activities such as continuous service improvement and several other activities have had to be picked up temporarily by technical staff.

Deployment Status

The section summarises deployment and operations activities for the period October 2013 to March 2015.

Resources deployed have seen significant changes, particularly in Q2 2014 with most Tier-2 sites deploying hardware procured with their GridPP4 2nd tranche hardware awards. There has been a 20% increase in HS06 available, which now stands at about 430,000 HS06, and a 28% increase in disk leading to 31,000 TB nominally available across GridPP sites. There are more deployments to come as acceptance tests are ongoing in some institutes, and several sites had their awards deferred to the 2014/2015 financial year due to constraints on funding from STFC. However, these increases may be partly offset, as there is a non-negligible amount of hardware now out of warranty and becoming too unreliable to be sustained. Even with the disruption of new hardware deployments, GridPP site availability and reliability (which since January 2014 is assessed separately for each of the LHC VOs) has on the whole remained well above the WLCG targets of 95%. In 2014 there was no repeat of the problems that led to some sites having to put worker nodes offline due to high machine room temperatures – over the last year several machine rooms have been renovated while others have had some hardware migrated to central university machine rooms.

The newest hardware is generally well understood, but challenges are being faced as the character of experiment jobs change, and internally machine resources become increasingly stretched with higher core numbers. GridPP has been leading in several areas of deployment testing, particularly in relation to multi-core approaches on different batch systems (A GridPP member is chairing the WLCG task force in this area and also led the migration to SL6 in 2013). Deployment of these (multi-core and whole node) queues is now a priority as the experiments (particularly ATLAS) make more use of multi-threaded jobs. In conjunction with this multi-core work many sites are seeking to move away from Torque/Maui based batch systems and during the current reporting period about 30% of sites have either moved to HT Condor, SGE or SLURM or set up test systems ahead of a move in the near future; most sites have also increased their use of Cobbler and Puppet configuration management (and are actively supporting one another through the sharing of modules). It is expected that the majority of sites will have moved batch system by the end of 2015. To aid sharing of experiences and improve contact with developers, GridPP staff jointly organised an HTCondor workshop at CERN in December 2014.

This period has witnessed a number of changes to the infrastructure with many large-scale changes driven by top-down upgrade campaigns. The first of these was a move of operating system (particularly on Worker Nodes (WNs)) to Scientific Linux 6 with a target of October 2013 that our sites met. The move led to performance improvements that required sites to rerun and publish their HS06 performance[1]. Another successful campaign required sites to move from unsupported EMI-2 to EMI-3 middleware.

[1] HS06 as a benchmark is increasingly challenged as an accurate indicator for WLCG jobs and work will start on a replacement towards then end of 2014.

The middleware products upon which WLCG/GridPP sites depend are generally supported now for the coming year(s), but there are some uncertainties. A notable one is ARGUS, which handles authentication. Following a drive to enable pulling down of banning lists from a WLCG and regional ARGUS server, it has become apparent that the product still has some instabilities and fixes are not being propagated quickly. In December 2014 a grouping of interested parties met at CERN and agreed to community support of various aspects of the component.

Security operations in GridPP continue to be well managed by a distributed team who partake in a duty rota. A new security officer begins work for GridPP in March 2015 which will allow several advancing areas to receive more GridPP input – for example policies related to cloud resources and federated access management. The last year has witnessed a steady flow of security advisories and updates (e.g. the last IGTF release was in December 2014) to which GridPP sites have responded well. This was particularly the case with the openSSL "heartbleed" vulnerability in March 2014. We have seen a couple of site security incidents since October 2013, but none that have caused ongoing concern. Early in 2014 we fully commissioned our resilient network of VOMS servers and, where requested, took over VOMS support for a number of NGS VOs. More widely we have contributed to joint discussions on future security policy and approaches including with EGI, EUDAT and PRACE teams, and to setting timelines for the community to move to SHA-2 – the UK CA switched to issuing SHA-2 certificates in April 2014. The last step for CERN to move to VOMS-admin and remove VOMRS was re-scheduled to mid-February 2015. In June 2014 a security training workshop took place at RAL for our sysadmins and this was very well received; it may now become an annual event.

This period has seen steady progress in GridPP investigations into, and deployment of, cloud and VM technologies. The GridPP cloud at Imperial moved to using OpenStack and the IceHouse release (together with Gluster FS for image storage) and is well advanced in becoming part of the EGI Federated Cloud project. Oxford and the RAL Tier-1 have also deployed Openstack. Lancaster has investigated VMWare cloud, QMUL is deploying a Cloudstack prototype and Manchester is pioneering a new approach based on its own VAC and Vcycle methods. Monitoring of job submission to VMs at sites using VAC/Vcycle and also via a new DiRAC prototype hosted at Imperial (to improve non-LHC VO job submission and data management capabilities) is now in place. In Q3 2014 Manchester expanded VM use onto WNs under its conventional batch system. The GridPP work has been well received by the LHC experiments (who are now using cloud resources for some production work – and of course to get more benefit from HLT resources at CERN) and WLCG management. There have been an increasing number of meetings and workshops aimed at pooling the experiences of the WLCG community into a future strategy in this area. EGI is now making cloud approaches part of its core offering. One area of increasing priority for further investigation is in the area of cloud/VM security – misconfiguration of VNC on some nodes led to a security incident in August 2014. An area that is picking up in WLCG is that of volunteer computing. A workshop towards the end of 2014 demonstrated workable and successful approaches within several of the experiments – it is expected that these will be improved and perhaps consolidated during 2015 before any formal launch.

There have been a number of developments and improvements in the area of operations support. Within the WLCG context GridPP has made significant contributions in areas such as the monitoring consolidation task force (whose mandate was to reduce the effort needed within operations and earlier this year led changes to the SAM infrastructure) and IPv6 studies (Imperial was one of the first sites in WLCG to move to running its services IPv4/IPv6 dual-stack). The UK was one of the first regions to fully adopt perfSONAR at all its sites for network monitoring and also deployed a prototype dashboard. There is still work to do in enabling the full network mesh and resolving some site firewall issues, not least because perfSONAR became susceptible to 'Heartbleed' and the mesh was closed for several months pending a new release. perfSONAR has proved a very useful tool in resolving end-to-end network issues and is now part of the considerations of a new network and transfer metrics group – clear metrics are needed to advance network status aware tools. GridPP is also working with a RIPE initiative using plug-in network probes (confusingly called ATLAS probes) to improve network-monitoring coverage.

The UK is active in the WLCG middleware readiness working group – whose aim is to introduce more stringent testing of products within site infrastructures and experiment workflows before they become part of our recommended baseline. To coordinate in this area CERN has recently appointed a middleware officer. GridPP now tests 3 of the 20+ products in the testing matrix.

Our sites have been responsive to middleware change requests; for example all moved within weeks to CVMFS release 2.1.19 during July 2014 as this was needed for CERN to upgrade its Stratum-0 release. UK sites have also been making good progress in resolving issues affecting the move to Glue2 – a validation exercise started in November. GridPP has also been at the core of developments to get CVMFS more widely used for VO software deployment. RAL T1 now runs a Stratum-0 in support of UK and EGI VOs – progress has been made moving UK regional VOs over to this infrastructure in the last month. This and increasing effort to support and run a stable DiRAC service for 'other VOs' at Imperial shows GridPP's commitment to improving the experience (and ease of) using its infrastructure by non-LHC communities.

Storage and data management has remained a very active area. With a move to Rucio distributed data management, ATLAS required all files at our sites to be renamed – this went smoothly and finished on target in Q4 2013. The main issue was with Rucio creating empty directories when first used in December 2013. Some problems were seen this in June 2014 when ATLAS LOCALGROUPDISK became full at certain sites causing those sites to fail Nagios tests, which impacted their availability. There were also a few niggling issues encountered with federated xroot in December 2013. Although the majority of tickets/issues continue to be resolved quickly, there are a few that have eluded resolution for many months. Top among these is one concerning LHCb pilot jobs losing connectivity at Bristol and RAL which was eventually traced to an HTCondor configuration issue.

We have continued our support of the DPM Collaboration and agreed a revised collaboration agreement and work plan (in areas such as interface testing, documentation and support) in 2014. We organized a DPM workshop in Edinburgh during December 2013. FAX and WebDAV services have been enabled at many sites and GridPP member at Edinburgh led a WLCG wide meeting on data access (and dynamic federations) in May 2014. There have been some major DPM upgrades during the period and these have gone smoothly, largely as the result of a lot of testing beforehand. Another noteworthy change is that LHCb started using storage at Tier-2s (at Manchester and RALPP) and this has gone well.

There have been two WLCG workshops during the period covered in this report: The first took place in November 2013 in Copenhagen and focused on areas such as Run 2 preparations and the evolution of the experiment computing models (including storage), and the outlook for e-Infrastructure. The second was held in Barcelona in July 2014 and prompted thinking beyond Run 2. GridPP presented, particularly on storage aspects, at CHEP in October 2013 and at HEPiX in November 2013.

Core operations work has continued smoothly during the last year. Weekly ticket reviews have been sustained and tools like VOMsnooper updated to aid site alignment with VO portal changes. There has been progress on monitoring visualization based on graphene (an iOS app was developed as proof of concept to show sites how the data might be used), and the deployment of VO Nagios for five of our smaller VOs. The GridPP Nagios service has received regular and on-time updates. Blog posting by team members and site admins, which act as a useful way of disseminating technical findings and solutions, is improving. The ROD activities have strengthened with the Tier-1 now also contributing a team member to the weekly tasks; it should be noted that the work has recently become more difficult due to an untested and somewhat less functional operations dashboard being released into production by EGI. Problems have included non-working email functionality, a loss of the security dashboard and bugs preventing Nagios alarms from being removed. To avoid such poor releases in future an advisory and testing group is being established between NGIs.

Our sites have seen a number of staff changes, for example at Glasgow, Durham, Birmingham, QMUL and Sussex. In December 2014 the WLCG operations coordination area launched a survey for sites to complete in order to better capture and understand ideas for the next phase of improvements in operational stability and efficiency. GridPP sites are already actively exploring new approaches such as using ElasticSearch based analytics on log files to spot problems earlier, and applying cgroups containment to develop understanding of job workflow and dynamics for better utilisation of resources.

In summary, GridPP sites are ready for the start of Run 2 and pro-actively improving the infrastructure and methods to make it more efficient and easily usable by other communities.

Users' Reports

ATLAS

During GridPP4, ATLAS has made extensive use of Tier-2 resources for simulation, group analysis and individual user analysis, and the Tier-1 resources primarily for reprocessing, simulation and data curation. All UK Tier-2s were in the top 20 for delivered HS06-hours in the year from August 2013, and two in the top 5. Inevitably, this pattern has been slightly different in 2013 and 2014 because we are in shut-down mode, so the emphasis has been on simulation, analysis and scaling exercises for the new computing mode; nonetheless, reconstruction (of real data and of simulated data) still represents a significant activity. The overall UK CPU usage by ATLAS in the UK is broken down by activity in Figure-9; this shows a healthy balance between analysis, Monte Carlo production and reconstruction. There is an increasing use of major Tier-2s in dynamic data placement and reprocessing activities; the latter will already be a more crucial role in GridPP4+, with a major revision of the processing and analysis model now rolling out, and will continue to grow in GridPP5. At present, the dynamic data placement in the UK is dominated by analysis formats (ntuples and the AOD), as illustrated in Figure-10. The UK has met its disk pledge to ATLAS within the RRB year, as illustrated in Figure-11.

The UK continues to perform well, delivering roughly 10% of the global Tier-2 capacity (reflecting the UK authorship share) and 13% of the Tier-1 capacity (our authorship share of Tier-1 providers). ATLAS has also made good use of beyond-pledge resources made available by the institutes from their own funds. It has allowed UK groups to exert direct influence on the development of analyses by the provision of over-pledge capacity for work using the institutes' leveraged internal resources.

ATLAS has revised its projected future resources assuming the tightly constraining envelope proposed via the Computing RRB/CRSG. This makes evident the major challenges. The processing time per event has already been considerably reduced (by a factor of two), but this speed-up must continue. More challenging is the growth in data volume; the scope for reduction in disk-resident copies is now exhausted and serious compromises in the physics achievable have to be considered.

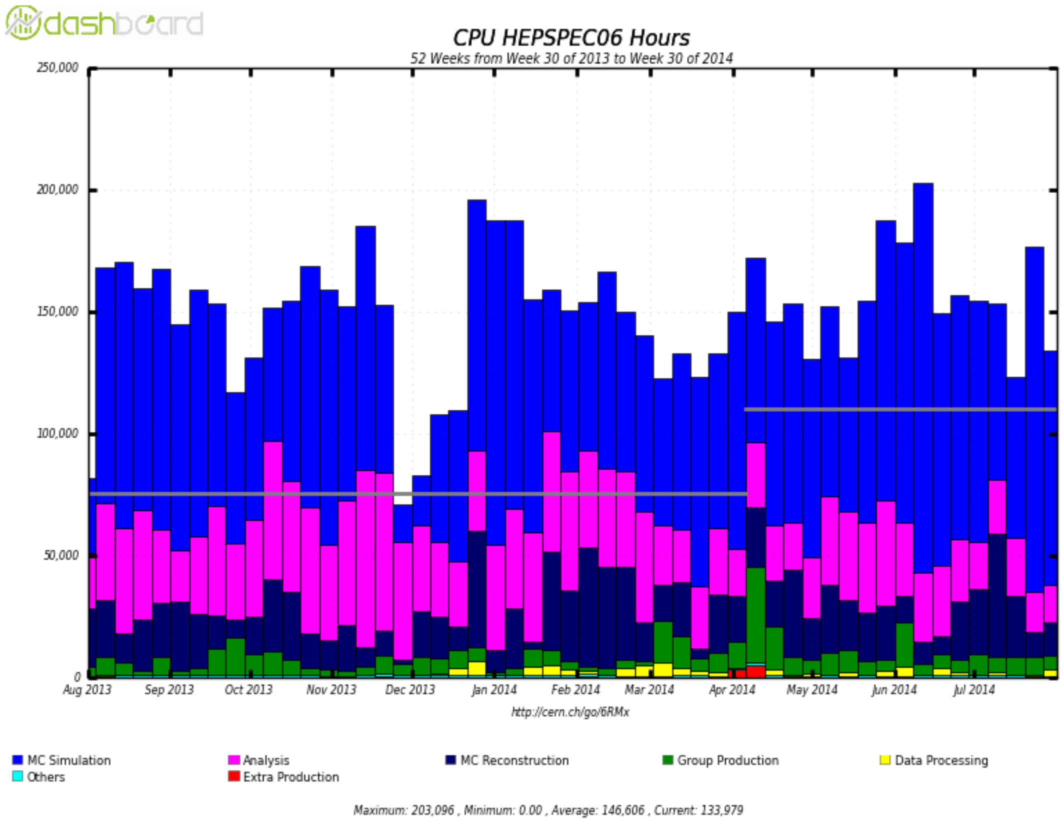


Figure-9 The ATLAS UK CPU usage by activity; the grey line represents the pledged value.

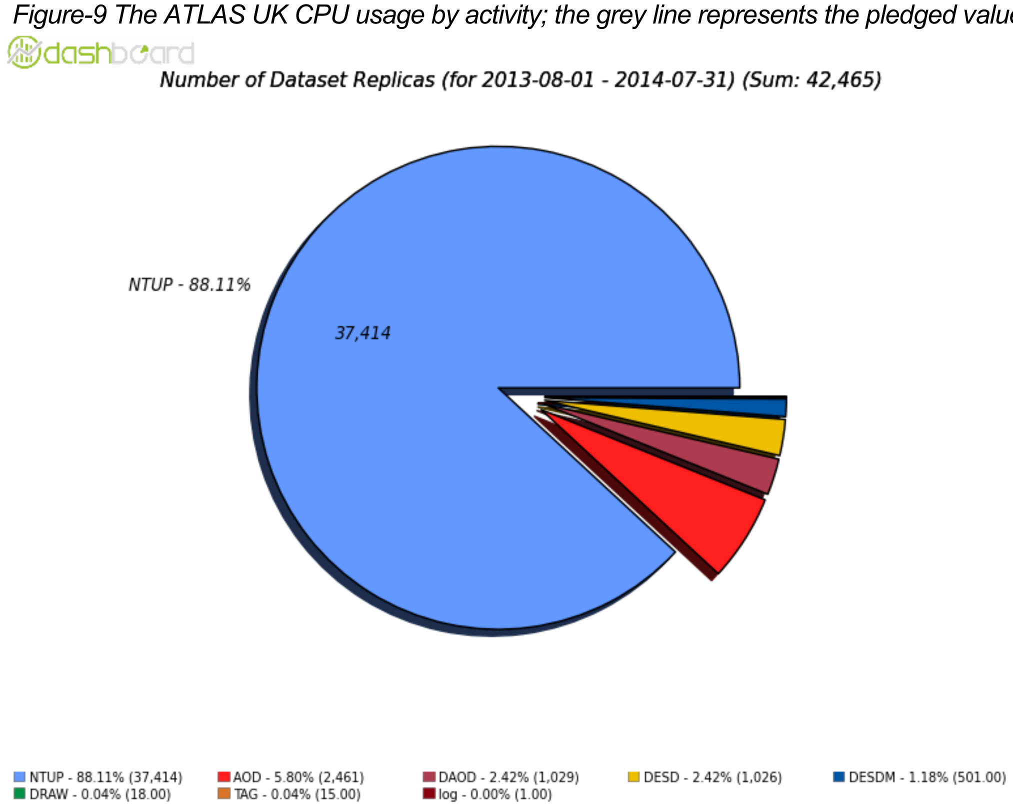


Figure-10 A snapshot of the typical data held in the UK through dynamic data placement by data type; it is currently dominated by analysis formats.

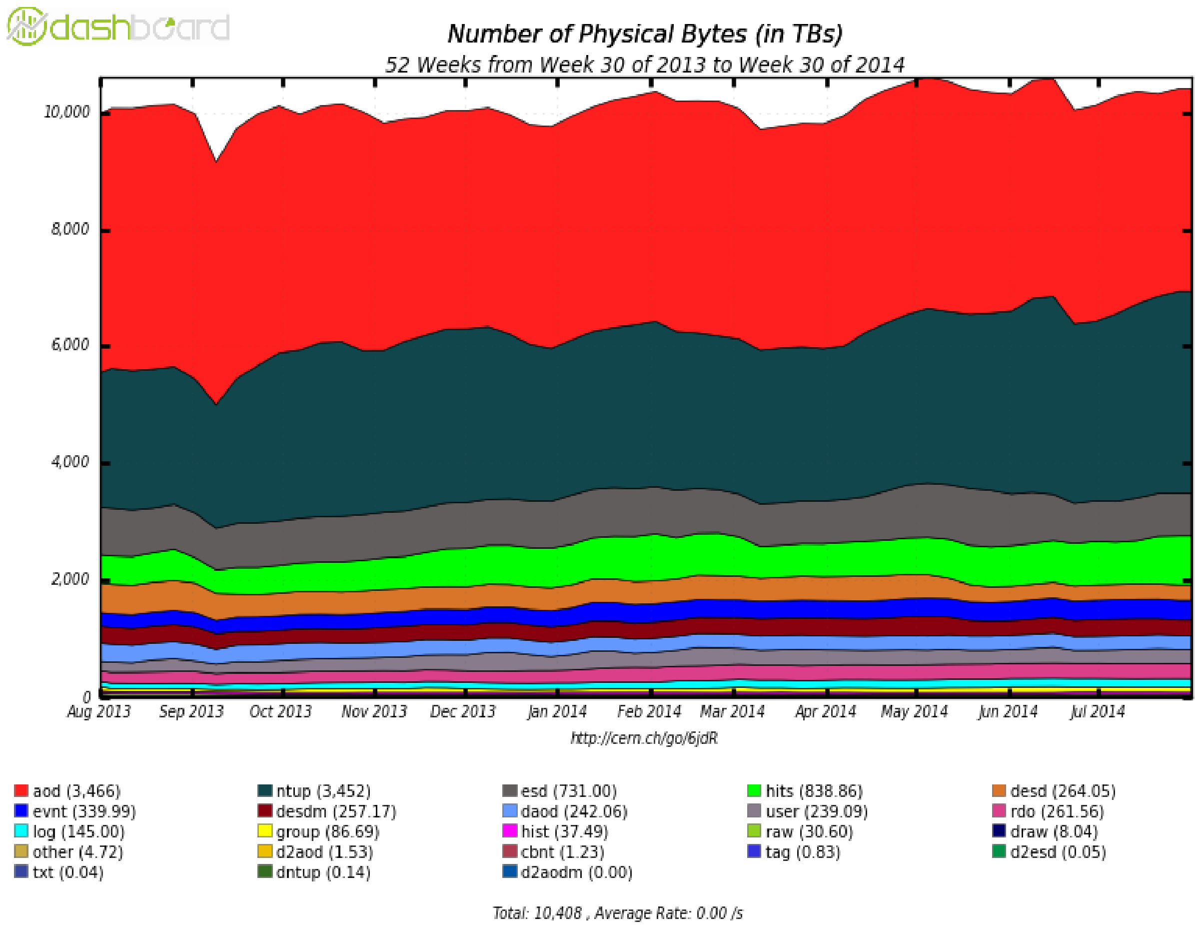


Figure-11 The UK disk capacity delivered to ATLAS through the year from August 2013, broken down by type; the grey line represents the pledge.

The role of the Tier-2s continued to grow in ATLAS in 2013 and 2014, with the major sites beginning to serve data to smaller sites with little storage using data federation. Analysis requires large amounts of disk with excellent tuned bandwidth to the local CPUs and good inbound and outbound network connections to other sites plus many additional services. These sites must be highly responsive to the needs of the supported experiments and their computing teams. If a CPU fails then it can be changed some time later, but if a disk fails it has to be replaced almost immediately and possible data-loss addressed. The transfer to the Tier-2s of workflows that were formerly done at the Tier-1s, such as reprocessing, reinforces the conclusion that the current distributed Tier-2 structure remains a balanced and resilient solution for the ATLAS Tier-2 requirement. ATLAS continues to estimate that a region the size of the UK requires a minimum of 10.5 FTE to run an adequate Tier-2 infrastructure and, in fact, the UK has one of the smallest levels of deployment and operation staff, despite being one of the largest ATLAS regions. In GridPP4, two FTE were allocated to each of five major ATLAS sites in order to provide continuity of operations and support. These have proven invaluable in the last year, with the sites performing with good reliability and availability.

ATLAS globally has investigated distributed computing implementations other than the current Grid model, including the limited use of commercial cloud resources and the opportunistic use of HPC resources. The UK has been particularly active in this work, making contributions to the pilot job system for the clouds, running a pilot factory pointing to UK cloud resources. While most workflows have been tested and made to work at some level, cloud resources are only feasible and efficient for about 10% of the total work, and HPC resources for even less; however, work continues to make best use of all available resources when appropriate. ATLAS-UK has also been adapting and making use of the VAC system for ATLAS.

CMS

The period since the last oversight committee meeting has been one of consolidation within CMS. Activity has been a mixture of producing detailed and refined analyses of the data taken in Run 1 and preparation for Run 2. There has also been evolution of the computing model used by CMS, which aims to make more efficient use of the resources available to CMS.

The UK plays an important role in CMS computing operations by providing a Tier-1 (at RAL) and three Tier-2 sites (at Brunel, Imperial and RAL PPD). The UK only provides a small fraction of CMS' projected hardware needs. However, our very efficient use of these resources means that we are able to deliver a proportionately higher fraction of the useful work done. We are able to do this because of the high quality of the individuals, employed both through GridPP and the Consolidated Grant, who work in CMS computing in the UK. This group of people is smaller than the number of people CMS estimate would be required to deliver the UK contribution and we are regularly awarded more service credit than we have people. The UK's role is very much appreciated within CMS computing. It is pleasing to note that UK sites are at the forefront of the adoption of the changes required to evolve the computing model.

The RAL Tier-1 is pledged to deliver 8% of CMS total Tier-1 hardware requirement and funded by GridPP accordingly, but, as can be seen in Figure-12, the UK has run nearly 11% of the successfully completed jobs. This reflects the fact that CMS has been able to make more efficient use of the resources at RAL than those at some other Tier-1s and is a tangible testament to the efficient operation of the RAL Tier-1 and the CMS specific staff embedded within it. These staff are seen within CMS as being highly responsive. One significant example of this is that RAL was the first CMS Tier-1 to separate the operation disk and tape resources as required by the evolved computing model.

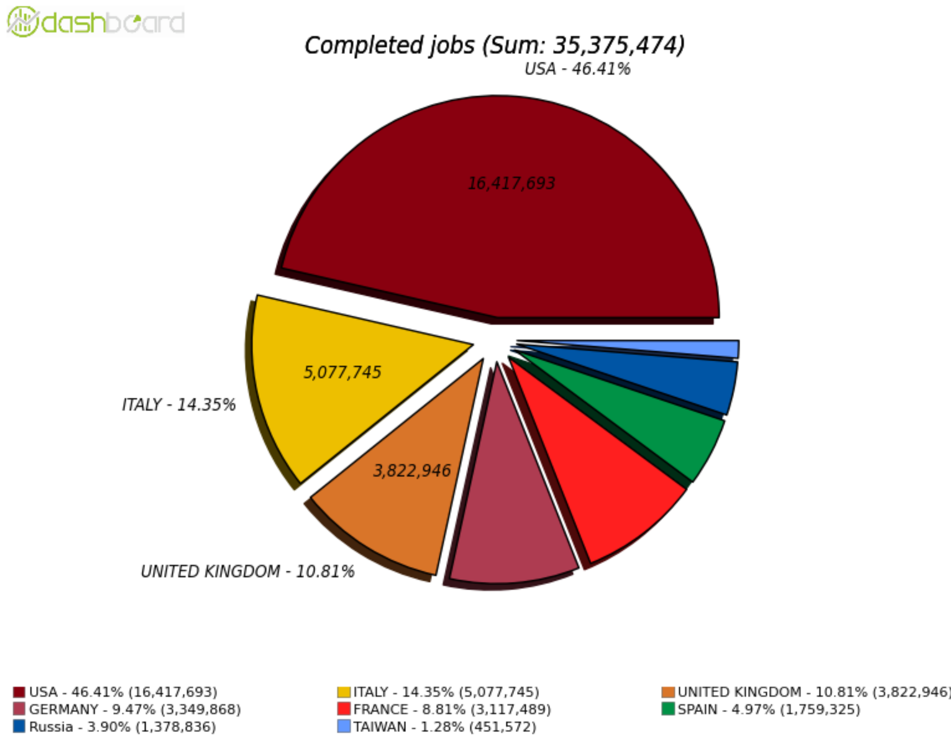


Figure-12 Number of successfully completed jobs at CMS Tier 1s since the last OC meeting (as measured by the number of successfully completed jobs).

GridPP pledges to provide 5% of the Tier-2 computing hardware that CMS projects that it will need. Figures-13 and 14 show the UK Tier-2 actual contributions to CMS (measured by successfully completed jobs). Again it is noteworthy that through the talented people employed at CMS Tier-2 sites, who have enabled particularly efficient use of the resources available and through additional resources provided by the institutes themselves (i.e. not through GridPP) the UK has been able to deliver a far greater fraction of the used resources than the pledged figure.

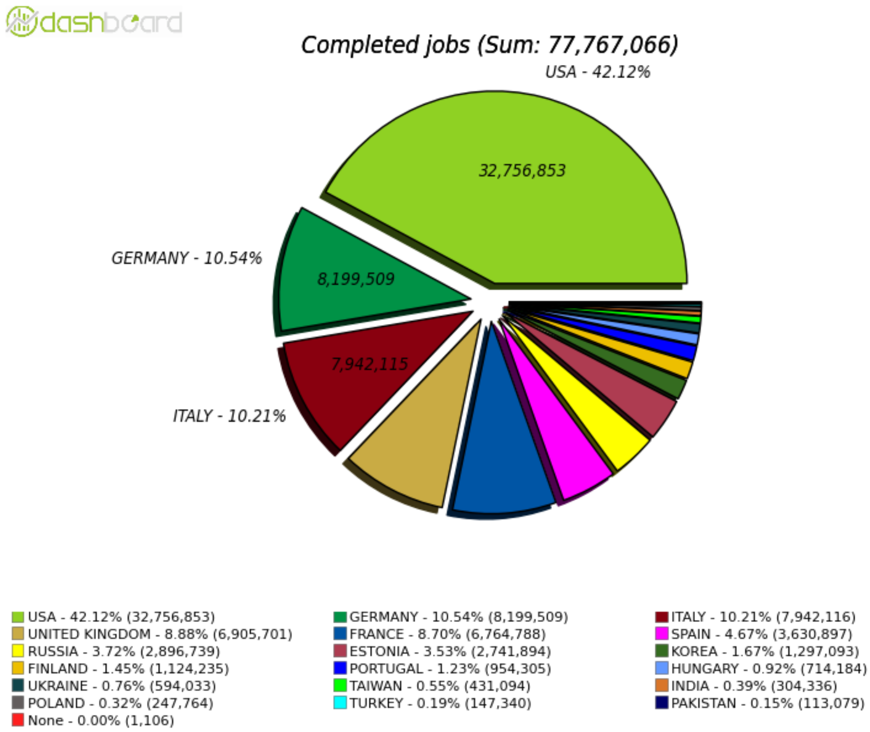


Figure-13 Contribution grouped by country since the last oversight meeting (as measured through the number of successfully completed jobs).

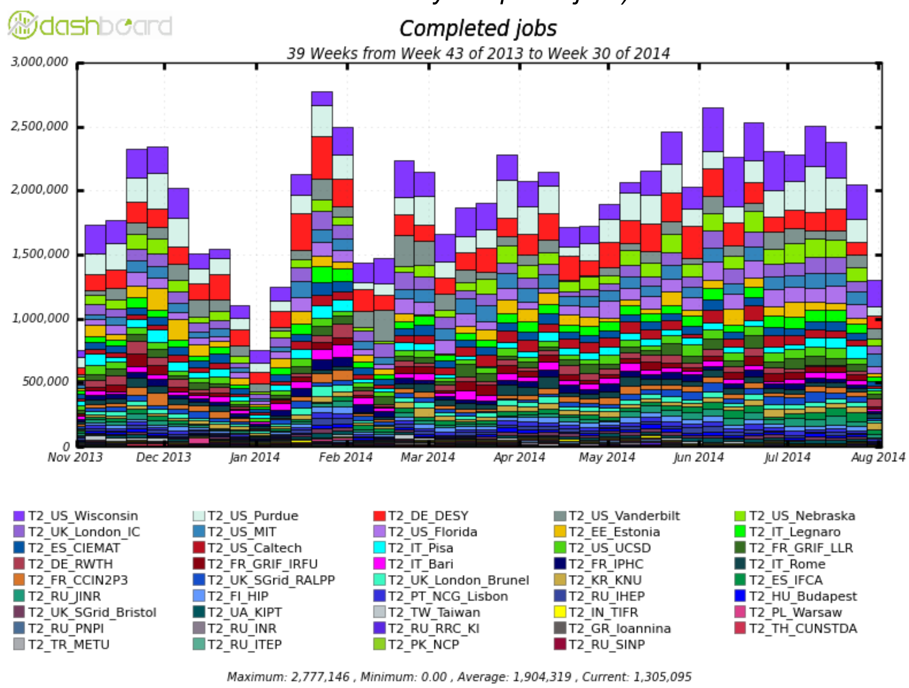


Figure-14 The contribution of individual Tier 2 sites to CMS (as measured through successfully completed jobs).

The same UK individuals who run the UK operations so efficiently for CMS also play significant play significant roles in CMS' central operations and development. One example of this is in the area of cloud computing which we expect will be of increasing importance to CMS computing. Centrally UK efforts have been focused on both using the CMS HLT farm as a cloud resource and on how CMS will interact with the CERN Agile Infrastructure. These efforts are part of a much wider team which is currently led by a UK person.

The UK's contribution to CMS computing, both in terms of operation of resources and the greater role played by GridPP people is highly appreciated by CMS management.

LHCb

The UK continues to be jointly the largest national collaborator in LHCb along with Italy. We are ~20% of the collaboration for Tier-2 purposes and ~30% of the collaboration within Tier-1 countries. During the last period the UK has continued to be a very reliable provider, both at Tier-1 and Tier-2 sites. This reflects both capacity but also as importantly the reliability and availability of resources. Figure-15a below, taken from the most recent figures which will go to the Computing Resource Scrutiny Group (CRSG), shows CPU usage at Tier-1 sites (left) and Tier-2 sites (right) during 2014. Similarly Figure-15b below shows the disk usage at Tier-1 sites (left) and Tier-2 sites (right) during 2014. From the LHCb point of view the UK pulls its weight and responds quickly and positively to any issues that arise. There have been no significant problems due to UK GridPP resources.

UK personnel associated with GridPP hold the positions of LHCb Deputy Computing Coordinator, Chair of the National Computing Board and the LHCb Tier-2-Data coordinator. GridPP staff also provide three Grid operations experts (GEOC roles). These roles ensure very tight coupling of GridPP to the LHCB requirements.

LHCb has now officially changed its computing model to parallel that of the GPDs, with the expectation that more Tier-2 sites will participate in data processing as well as simulation, and that nominated Tier-2D sites will additionally hold DST data for user analysis jobs. The introduction of Tier2-Data centres (Tier-2 centres with data) has proceeded smoothly and successfully. We now have Tier-2 sites in several non-Tier-1 countries contributing and expect ~2 PB of disk to be provided by the end of 2014/15. As usual, the UK T2D sites at Manchester and RAL have proven to be extremely effective and led the way. More UK T2D sites are expected to join (Imperial, Glasgow and Liverpool).

LHCb has evolved its computing model for the approaching Run 2 to include three categories of data (1) The normal prompt data stream to be processed immediately, (2) a parked stream to be kept for later processing and (3) a turbo stream, which is processed within the trigger farm. The prompt stream will be reconstructed after a delay of a few hours, to allow calibration and alignment to be completed. This means that no further reconstruction will be needed as has been the case in the past, when a full reconstruction took place immediately after the end of data taking in any year. The parked stream will not be needed on 2015 and 2016 but may be needed in later years. The turbo stream is mainly for charm physics, and assumes that all necessary information can be created in the trigger farm itself, and therefore no further offline processing is needed. These changes will help to reduce the overall CPU requirement, and LHCb continues to use non pledged resources in Russia, and to make full use of its trigger farm.

LHCb is also developing the means to parallelise its framework, Gaudi, and is very much assuming many core-devices will be the norm in future. LHCb is also adapting its workload management system to support virtual machine architectures associated with the desire to use heterogeneous resources in the future. This activity is being led by the UK in conjunction with the virtualized infrastructure provided by GridPP sites and using the virtual machine lifecycle managers, Vac and Vcycle, developed by GridPP. These developments are already giving LHCb access to resources equivalent to another larger Tier-2.

Following C-RSG recommendations to all experiments, LHCb has put considerable effort into developing better data replication strategy based upon data popularity. In fact LHCb has been very efficient anyway, but still expects to save some data volume.

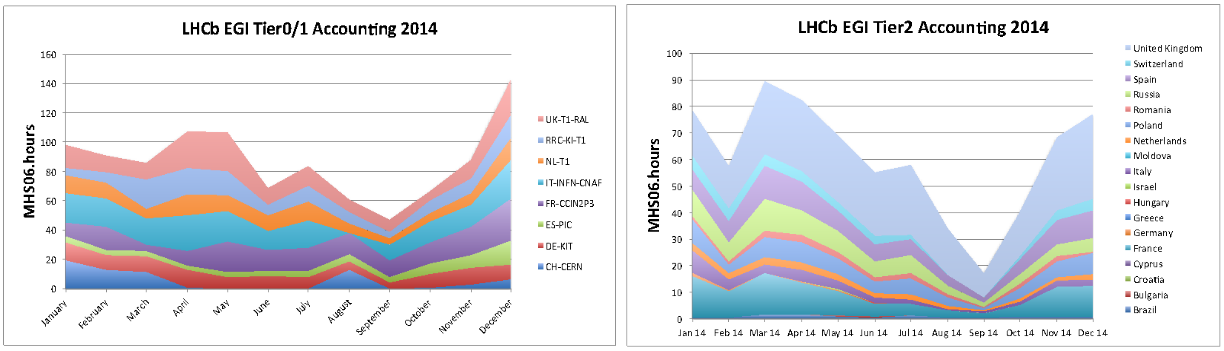


Figure-15a: LHCb CPU usage at Tier-1 (left) and Tier-2 (right) in 2014.

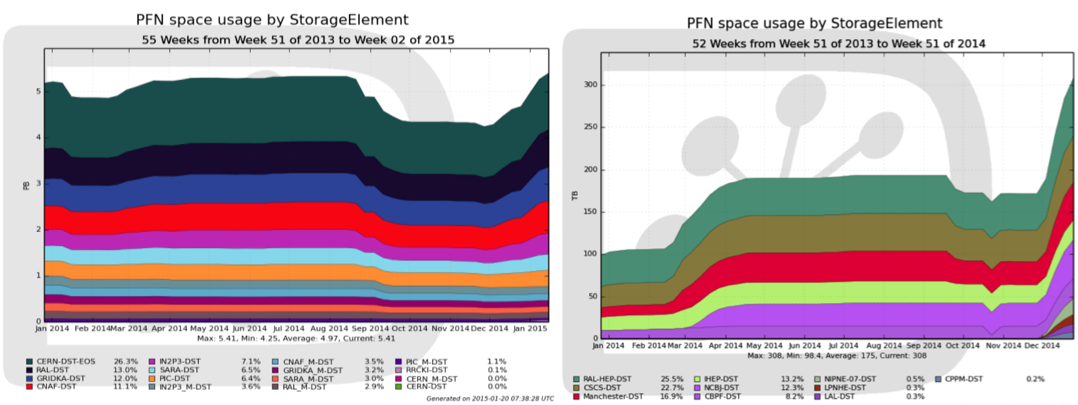


Figure-15b: LHCb disk usage at Tier-1 (left) and Tier-2 (right) in 2014.

Other VOs

We are making steady progress with our other VO communities (i.e. those not one of ATLAS, CMS or LHCb). Excluding ALICE, the highest other user, the fraction of GridPP Tier-1 resources used by "others" increased from 3% in 2013 to 5% in 2014. At Tier-2s the figure rose from 12% to 14% with biomed, followed by phenogrid as the largest users in this category.

Our SL6 migration in 2014 went without major incident for the other VOs, and there was no reoccurrence of the VOMS renewal issues that affected many VOs in early 2013 - when user VO memberships expired with little warning due to a middleware update causing the resetting of all membership accounts to a default date. Problems that did occur during the reporting period include SNO+ writing 1TB files to several sites (e.g. Sheffield) causing disk pools to fill and Nagios tests to fail. Biomed made strong demands on resources such that their job storms at time have saturated a couple of sites (the sites have since implemented maximum job quotas on the VO or in some cases banned individual users – particularly where the users also submitted multi-threaded jobs to single core resources).

The data management strategy for our smaller VOs is still a concern, especially as the LHC VOs move away from tools such as the Local File Catalogue (LFC) and WMS. To mitigate major risks in this area we have been developing a new job management and data management service based on DIRAC (a tool used by LHCb). DIRAC as a service was in test during 2014 and has received positive feedback from several of our early adopter VOs but particularly Landslides, T2K and CERN@School. In December 2014 a proposal to ramp up use of DIRAC such that it becomes a formal production service by mid-2015 was agreed.

GridPP has continued to encourage the smaller VOs to move their software deployment mechanisms to use CVMFS. ILC has been one of the most recent VOs to move to using CVMFS, and the Tier-1 made a final decision to switch off its NFS software server in September 2014. Whilst good progress has been seen with software distribution, more work is needed to address increasing issues associated with use of shared storage resources whereby problems for other users of the same space (or other members of the VO) occur. They have little incentive to use spacetokens so better education remains our best option.

In mid-2014 GridPP operations sent out warnings to our VO communities about a major change that may affect them. In November 2014 there was a move away from lcg-utils and gfal towards the newer gfal2 and gfal2-utils tools; the older versions then became unsupported. This change required the VOs to actively adapt their frameworks which some, but not all, have done (the old rpms remain installed in parallel to the newer ones but are not available when we move to SL7).

There been a small churn in the number of active or supported VOs over the reporting period. Oxford provided support to the European Centre for Medium-Range Weather Forecasts with some network performance tests, but they did not run work on the Grid. QMUL has been working to support HyperK.org and with storage now run on iRods at QMUL there is an open request for more sites to enable the VO. The LUX dark matter collaboration has recently been in contact with our members at Imperial. Our work on DIRAC (which provides the ability to run custom VMs on our testbed of cloud and VAC resources) coupled with CVMFS (which makes pulling in VO software builds much more simple) has led in the last few months to a flurry of new requests/testing including from/with LIGO, LSST, UCLan (galaxy simulations), LOFAR, and a proteomics group at QMUL. Some inactive VOs are returning to activity such as SuperNEMO, whilst others now appear obsolete and are in the process of being decommissioned (for example test VOs associated with the NGS/NES). The Q4 2014 reports indicate that currently GridPP has about 50 VOs enabled[2] and many VOs being developed and supported within regional incubator VOs.

MICE reprocessed their data on the GridPP infrastructure in October 2013 and moved their RAW data to CASTOR. Work on their configuration database has been ongoing for many months. NA62 has made use of resources but their focus has been on software development for their Monte Carlo data distribution (including use of DIRAC as a WMS) and work on asynchronous FTS transfers.

Impact and Dissemination

The GridPP Dissemination Officer, Neasan O'Neill, left the project at the end of July 2013 after several years to become Events Officer at EGI. A new Dissemination Officer, Tom Whyntie, was

[2] http://pprc.qmul.ac.uk/~walker/votable.html

appointed at the end of October 2013. Tom has continued in his role as STFC Researcher in Residence at the Simon Langton Grammar School for Boys in Canterbury for the CERN@school project where he had previously worked actively with GridPP to conduct research based and educational activities. Tom is using the Grid for what it was designed for while promoting its existence and use to a new generation of scientists and engineers, and current researchers in the field. He has worked with various funding and industry bodies in order to gain support and foster strategic partnerships that have enabled the project to flourish.

Tom's appointment is allowing us to leverage this work and fully integrate it with GridPP and will give us access to other school projects such as data from classroom based Timepix silicon pixel detectors and the space-based Langton Ultimate Cosmic ray Intensity Detector (LUCID), recently launched. During this transition in Dissemination Officers, we took the opportunity to update the job description so that although dissemination is still important, the emphasis has changed more towards impact and two-way engagement with industry etc. Tom is very well suited to this changed role.

Prior to Tom's appointment, smaller VOs were presented with a wide array of tools and interfaces for accessing GridPP resources. While this approach offered a great deal of flexibility, sometimes this choice could be overwhelming to new users. Experience had also shown that often new groups would develop (and redevelop) their own ad-hoc scripts and custom tools with the help of local GridPP Tier-2 sys admins, placing additional demands on manpower to essentially reinvent the wheel. It was therefore decided that a more uniform approach, using a standard user interface and toolkit, would offer a more scaleable solution as GridPP sought to widen its impact outside of particle physics. Industry partners - SMEs in particular - simply do not have the time or the staff to develop custom Grid solutions, so this was seen as a temporary priority over the more traditional "news" aspects of the Dissemination Officer role.

With the help of several members of the GridPP Deployment team, and drawing on his experience of working on the CMS experiment (i.e. a larger VO), Tom has assembled and properly documented a suite of tools to make it much easier for new users, such as SMEs and others, to quickly get started using our Grid resources. These tools cover a complete workflow including software deployment, job submission, data processing and metadata handling. Larger VOs (i.e. the LHC experiments) typically have whole teams dedicated to developing proprietary solutions for their data management needs. Adding metadata functionality has proven particularly useful for smaller VOs, as this has not traditionally been covered by LCG toolkits. By offering a GridPP-developed solution for smaller groups, the entry barrier for new users is significantly reduced and the scope for achieving wider impact has been increased.

CERN@school - or rather, cernatschool.org - has proven useful as a "technology demonstrator" VO for this new toolkit, allowing Tom to make use of his STFC Researcher in Residence role. For example, using the newly developed tools, CERN@school has successfully used GridPP resources to store and access data from the school-based Timepix detectors and generate Monte Carlo simulations of the LUCID spacecraft detectors. This has then enabled Tom to help two other small projects get started with the Grid. One consists of two cosmologists at the University of Central Lancashire who are doing massively-parallelisable galaxy simulations. They have successfully joined the NorthGrid regional VO and are submitting jobs. The other is a collaboration between particle physicists and biologists at QMUL to develop an industry-ready suite of proteomics solutions leveraging GridPP's Grid computing capabilities and resources. This one-time investment in a streamlined package of tools and user guides will make it much easier to attract SMEs and others to take advantage of GridPP resources as we move towards GridPP5. GridPP is also developing Cloud interfaces to its resources and these coupled with increased virtualisation of services should also make it even easier to attract new communities to our resources.