GridPP Project Management Board

Project Status

Document identifier:	**GridPP-PMB-190-ProjectStatus.doc**
Date:	**25/08/20**
Version:	**0.23**
Document status:	**Final**
Author	**PMB**

Introduction

The Oversight Committee last reviewed the GridPP project in May 2019. Since then, the LHC has been in Long Shutdown-2 (LS2), which commenced at the end of 2018 and was originally scheduled to end with beam-commissioning for Run-3 in May 2021. Unfortunately, the global pandemic has delayed upgrade work on both on the experiments and the accelerator complex and LS2 has now been extended. There is a complex interplay between the various components of the delayed upgrade work, which has proved to be a difficult problem to re-optimise in terms of costs, risks, and physics-output for Run-3. The current plan anticipates a re-start delayed to February 2022 to allow the possibility of the new ATLAS NSW-c detector to be installed during this shutdown. This delay would also allow CMS to undertake some shielding work that will be needed for the higher luminosity part of Run-3. However, if the construction of the ATLAS NSW-c detector gets further delayed, it is possible that the re-start will be brought forward, but a longer mid-run shutdown (5-months) will then be needed to install the new detector. The current intention is that Run-3 will end as originally planned at the end of 2024 to allow LS3 to commence in 2025 to start preparations for Run-4 (HL-LHC).

As can be appreciated, the significant uncertainty in the LHC schedule has a similar effect on the experiment computing resource requirements. There has been no concerted attempt to re-plan global resource requirements at this point, and the Worldwide LHC Computing Grid (WLCG) Management Board has noted that the previously established 2021 resource requirements should stand, but it is understood that there will less pressure to deliver these all by the pledge-date of April 1st 2021. However, it is also clear that the delay does not simply move the resources requirements back a year because the amount of data now anticipated in 2022 is much greater than originally planned for 2021. The delayed restart will allow a much quicker ramp-up of luminosity because various commission works, such as magnet training, can be done in the absence of beam. From the UK point of view, the procurement process for the 2021 pledge was launched at the start of this financial year and is well underway. Therefore, the new schedule does not change our plans but just reduces the pressure, somewhat.

The UK Grid infrastructure has continued to provide an excellent service as part of the WLCG. The main points to note from this period are:

1) Science Board considered the GridPP6 proposal on October 24th, 2019 and awarded £18.3M resource funding and £9.38M capital funding, noting that the latter would only fund ~70% of the GridPP6 MoU commitments to WLCG (the UK fair-share of LHC computing resources).

2) The GridPP infrastructure ran seamlessly during the COVID-19 lockdown, with no noticeable drop in capacity or service level. Before lockdown, the Tier-1 did an extended test of remote-working and was well prepared. Across the sites, the lockdown has delayed a number of routine upgrades, migrations, and development work that required on-site access. There was also, inevitably, a delay installing new hardware that had been procured before the end of the financial year.

3) Significant resources, both computing and manpower, were quickly diverted to support COVID-19 modelling efforts, as follows:

   3.1 GridPP submitted a RAMP proposal[1] (Rapid Assistance in Modelling the Pandemic).

   3.2 GridPP appointed an internal technical lead to manage all technical activities related to COVID-19.

   3.3 The protein-folding simulation workflow 'FoldingAtHome' (FAH) was enabled across the GridPP infrastructure to scavenge free cycles and the 'GridPP Team[2]' currently ranks 333rd out of over a quarter of a million teams.

   3.4 GridPP brokered an international meeting between the USA-based FAH team, the CERN Rucio team, and UK-experts at Daresbury and elsewhere, to offer assistance in data-movement (FAH had been overwhelmed by the 1000x increase in work).

   3.5 As a consequence, Tier-1 Storage is now being used to consolidate output from FAH jobs run on volunteer computers around the world before being sent to HPC resources for further processing.

   3.6 GridPP worked with the Scottish COVID-19 Response Consortium (SCRC) and computing experts from UKAEA to look at containerising and executing COVID-related workloads.

   3.7 GridPP personnel at Brunel are engaged in a short project with Brown University within their modelling group for infection diseases, trying to develop and optimize their agent-based modelling simulation.

   3.8 GridPP personnel are working on modelling the COVID-19 vaccine pipeline.

4) A new £1.3M Tape Library has been installed at RAL, currently providing 63PB of storage. This has entered production and the migration of data from the old library has started, reducing a number of risks.

5) GridPP has worked closely and seamlessly with the other STFC e-infrastructures such as IRIS and the developing ExCALIBUR project:

   5.1 The GridPP infrastructure enabled LSST-UK to fully contribute to the LSST-DESC Data Challenge-2.

   5.2 GridPP successfully bid for IRIS digital-asset funding to develop a UK accounting dashboard, the development of multi-VO Rucio, AAAI development and the integration of the DIRAC workflow management system with Rucio.

   5.3 GridPP were prime instigators in a successful bid to UKRI (via EPSRC) for a part of the ExCALIBUR exa-scale computing initiative and are part of the complementary SWIFT-HEP bid recently submitted to STFC.

[1] https://cernbox.cern.ch/index.php/s/wdQzMdBJcdNTIGR
[2] https://stats.foldingathome.org/team/246309

GridPP5 Project Summary

The GridPP5 project completed at the end of March 2020, corresponding to the end of the first quarter (20Q1) and the final status of the GridPP5 metrics and milestones are presented here. The GridPP project successfully delivered computing resources to the LHC experiments and other users at the appropriate level and the project management systems allowed us to monitor progress in each area. The evolution of the milestones and metrics since the start of GridPP5 is shown in Figure-1 and the current Project Map, which summarises the current/final status, is shown in Figure-2.

	Q216	Q316	Q416	Q117	Q217	Q317	Q417	Q118	Q218	Q318	Q418	Q119	Q219	Q319	Q419	Q120
Metric OK	140	137	132	121	121	122	126	122	124	122	126	126	123	124	122	124
Metric close to target	8	10	10	15	17	17	16	18	17	16	12	13	13	11	11	12
Metric not OK	0	0	1	7	6	5	2	2	1	1	1	3	6	7	9	6
Not able to be measured	0	0	0	0	0	0	0	0	2	3	3	2	2	2	2	2
Suspended	0	0	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Awaiting input	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Metrics Total	148	147	146	146	147	147	147	145	147	145	145	147	147	147	147	147
Milestone achieved	9	9	16	18	23	25	32	34	35	40	44	51	56	58	60	67
Milestone close to completion	0	0	0	1	2	1	0	1	2	0	0	0	0	0	0	0
Milestone overdue	0	1	2	0	0	2	2	2	2	3	2	1	3	3	2	3
Milestone not due/ metric n/a	58	59	52	51	44	41	35	34	30	29	26	18	11	9	8	0
Milestones total	67	69	70	70	69	69	69	71	69	72	72	70	70	70	70	70
Total	215	216	216	216	216	216	216	216	216	217	217	217	217	217	217	217

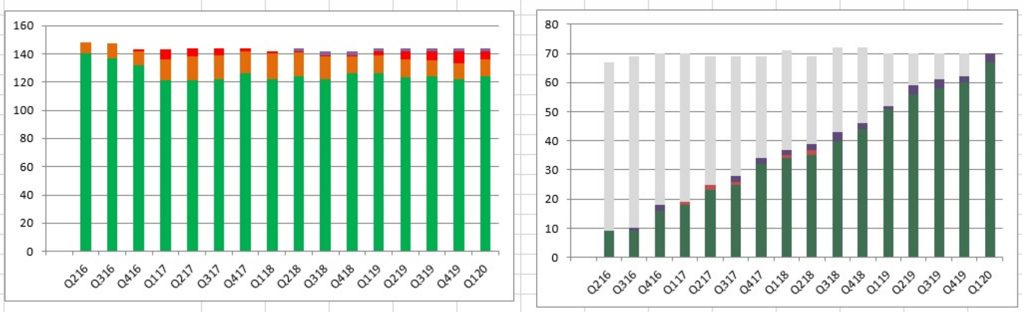


Figure-1: GridPP5 metrics and milestones: Numerically presented at the top, the lower left plot then shows the number of metrics that are currently met (green), close (amber) and not met (red). The lower right plot shows the cumulative total of milestones completed over the project.

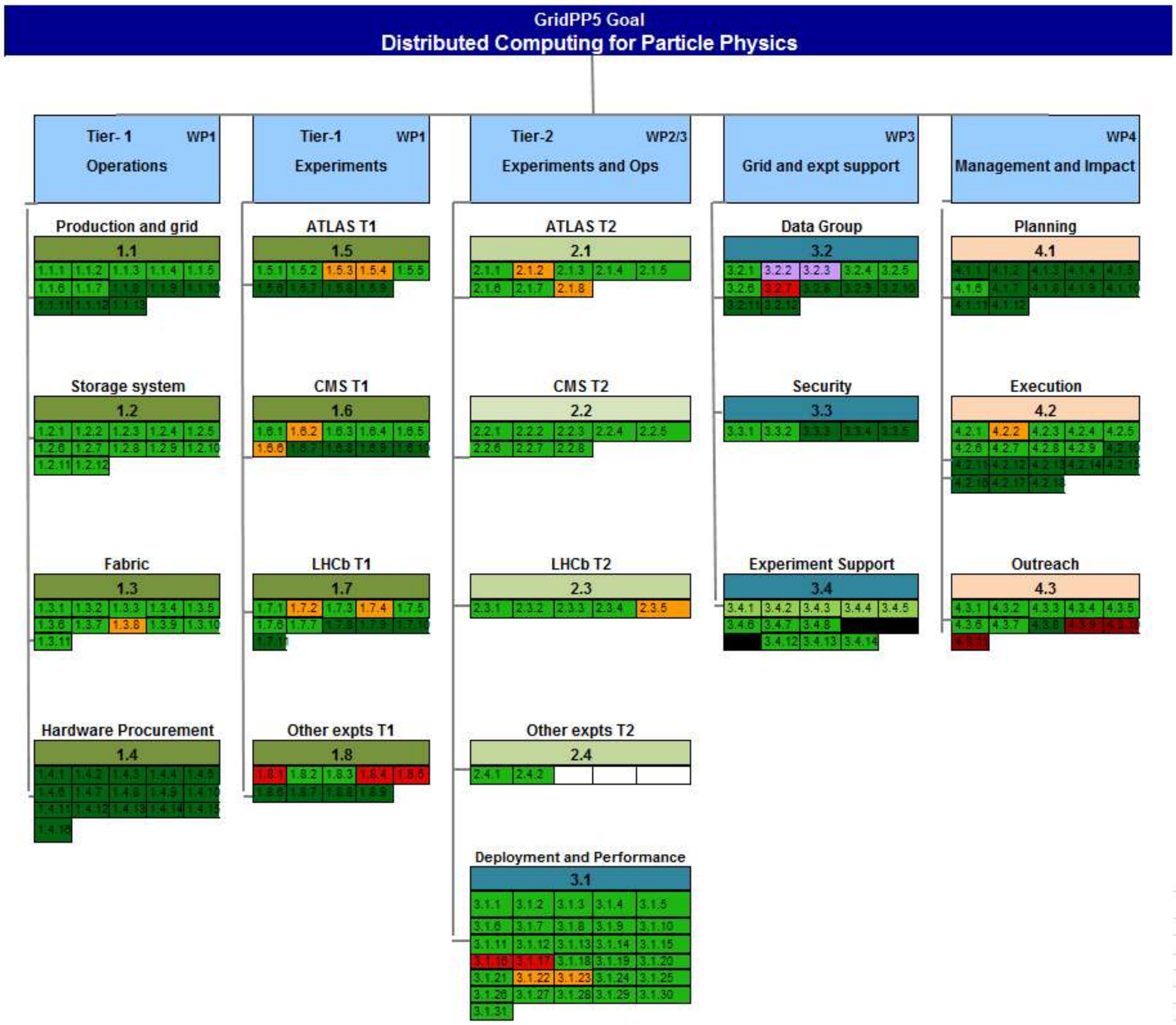


Figure-2: The GridPP Project Map at the end of 20Q1.

The status at the end of 20Q1 shows that the majority of metrics were being met (green) or were close to target (amber). We discuss below the six red metrics and the major amber ones, the majority of which were very close to their target value.

The red metrics this quarter are:

- 1.8.1: This is the "Efficiency of the Tier-1 CPU for non-LHC VOs". Normally around 90%, this dropped to 57% in the final quarter because 86% of the non-LHC jobs run at the Tier-1 were for the international LSST-DESC Data Challenge-2, which had poor efficiency. This was a known issue: to meet the memory requirements; LSST jobs had to reserve more cores than needed. In the longer term LSST will need to moderate their memory requirements or fund higher-memory cores.
- 1.8.4: The Tier-1 CPU efficiency for ALICE jobs at 61.7% was well below our target of 75% and below their usual performance of around 85%. In January ALICE migrated from Castor to Echo;

during this time ALICE production work used a different authentication mechanism that was not configured at RAL. This meant that ALICE jobs were failing to access Echo and falling-back to fetch their data from off-site.

- 1.8.5: The fraction of CPU time used by ALICE was below their historic usage level of 5%. However, this was well above our pledged level because ALICE had historically been very good at scavenging cycles. The metric was set at this level to give an early indication of problems, but in this case, it simply reflected the fact that our CPU capacity had recently been boosted with new procurements and ALICE was still receiving well above their pledge level. We should re-phrase this metric in GridPP6.
- 3.1.16 and 17: These two metrics refer to the availability and reliability of SouthGrid. Sussex has been in extended downtime (the site upgraded to CentOS7 and a new version of the ARC CE were underway when the COVID-19 lockdown halted work). Other major SouthGrid sites were all OK.
- 3.2.7: The number of blog posts by the Data group was zero this quarter. The total this year has been two as the focus of the group has moved away from blogging. This is no longer a useful measure of the work they do supporting the GridPP project and we will remove this metric in GridPP6 as we will no-longer be able to fund all the effort in the storage group.

The amber metrics are:

- 1.5.3 & 1.5.4: Tier-1 data availability on the dashboard and data acceptance for SAM tests for ATLAS were 97.8% slightly below the 98% target.
- 2.1.2 & 2.1.8: The ATLAS production job success rate at Tier-2 sites dropped 1% below our 95% target this quarter. Some storage issues and a configuration problem brought the ATLAS user analysis rate down to 92% (3% below target).
- 1.6.2 & 1.6.6: The availability and use of resources at the Tier-1 for CMS has been slightly low this quarter. Investigations by the Liaison officer and CMS contacts are on-going.
- 1.7.2 & 1.7.4: LHCb Production and User jobs at the Tier-1 were a few percent below target this quarter.
- 2.3.5: The fraction of LHCb work carried out by the UK Tier-2s was 11.2% compared to our target of 15%. However, this reflects the installation of new hardware and the UK more than delivered its CPU pledge to LHCb.
- 3.1.22 & 3.1.23: LondonGrid availability and reliability fell 1% below our target of 95% this quarter. This was caused by upgrade work at QMUL moving to CentOS7 and Arc6 CEs.

All GridPP5 Milestones except three were completed, which were associated with reviewing and updating the Knowledge Exchange part of the GridPP website: The GridPP5 Outreach post was phased-out halfway through the project when the post-holder left. In agreement with the Oversight Committee, we did not attempt to refill this post and re-deployed the effort to help with a more pressing need for technical development of tools for non-LHC communities.

Table-1 summarises the total effort delivered to the GridPP5 project over the four years. The Tier-1 staff profile was planned at 17.5 FTE in the first two years, falling to 14.5 FTE in the last two. In practice, due to retention and recruitment difficulties, the Tier-1 averaged 15.6 FTE during the first two years. Bookings in the third and fourth year averaged 14.6 FTE. The overall project average was 15.08 FTE compared with a target of 16 FTE, giving 94.3% of the target.

GridPP5 Work Area	GridPP funded effort (FTE)	Unfunded effort delivered (FTE)	GridPP funded effort delivered (%)	Unfunded effort delivered (%)	Total effort delivered %
Management	1.1	0.0	100.0	0.0	**100.0**
Miscellaneous	0.6	0.0	100.0	0.0	**100.0**
Operations + experiment	9.9	1.3	90.7	11.9	**102.6**
Tier-1	15.1	0.0	94.3	0.0	**94.3**
Tier-2	17.1	5.6	98.2	31.8	**130.1**
Total	**43.8**	**6.8**	**95**	**15**	**110.0**

Table-1: Effort delivered during GridPP5 (FY16-19)

The operations and experimental support area delivered 90.7%. The lower level was due to several factors: Slight under booking on the GOCDB/APEL support line (the work was done and this really just reflects the internal complexity of booking effort against projects at RAL); a significant gap caused by recruitment problems for the CMS liaison post (after long recruitment delays, the first appointee took a job elsewhere at the last moment and the post had to be re-advertised); a small gap in the ATLAS liaison post when the post holder moved to become the Tier-1 manager; and the loss of the production manager half post in the last year at Cambridge who, for internal reasons, was directed to move to support SKA. However, the shortfall in these areas were compensated by additional unfunded (by GridPP) effort, resulting in total effort delivered to this area being 102.6%

The total effort delivered over the whole of GridPP5 is shown above, including effort funded from external non-GridPP sources (EGI and University funds). Some of the missing funded effort will be delivered during the current six-month no-cost extension of the GridPP5 project.

All detailed quarterly reports are available on the web at:
https://www.gridpp.ac.uk/collaboration/docs/gridpp-quarterly-reports/

GridPP6

Introduction to GridPP6

GridPP were invited to submit a proposal for the continuation of the project for 4-years from April 2020. The brief stipulated "flat cash" with ±10% scenarios and, in addition, a 70% scenario. It quickly became apparent that, for hardware, there was a significant mismatch (more than a factor of 2x) between the "flat-cash" guideline of around £8M and the existing estimates of the experiment's requirements. This quickly led GridPP to initiate a debate within WLCG about the planning assumptions used by the experiments and, eventually, to revised WLCG guidelines related to how quickly computing costs should be assumed to fall. In addition, the LHC experiments have traditionally ignored the cost of replacing end-of-life hardware in their resource planning. This was a relatively small effect when Moore's Law was driving a rapid increase in computational power-per-core but is now a significant additional cost not considered by the experiments when claiming to have stayed within the (over optimistic) WLCG "flat-cash" guidelines.

Fortunately, since the submission of the GridPP proposal, renewed competition between AMD and INTEL resulted in a major price drop, which has alleviated the situation to some extent. Of course, it is impossible to know whether this represents the start of a long-term trend or will just be a brief "blib" in the pricing evolution. There are additional large uncertainties and unknowns that make any estimate of the cost of future hardware, very difficult. These include COVID-19 and the global economic climate, Brexit and the UK exchange rate, delays to Run-3 at CERN and the evolution of experiment requirements, and growing but often ill-defined resource requirements from non-LHC communities. These factors conspire to give an uncertainty that is unprecedented over the 20-years of GridPP, but our current estimate is that there is about a £6M capital shortfall in the GridPP6 era.

The GridPP6 proposal was prepared based on five work packages (WP), consisting of:

- **WP1** that contains the effort needed to keep the WLCG physical and software infrastructures working in the UK as a coherent part of the worldwide federation. This work package is subdivided into:
  - WP1a: Operation of the Tier-1 service
  - WP1b: Operation of the Tier-2 services
  - WP1c: UK Grid Operations
- **WP2** that contains the effort required to interface with the LHC experiment computing teams and support LHC and non-LHC user workloads to ensure efficient use of the infrastructure.
- **WP3** that contains the effort necessary to fulfil federal responsibilities within the worldwide community allowing the UK to both influence and contribute to WLCG at the appropriate level.
- **WP4** that contains the effort needed to evolve the Tier-1, to reduce the overall operational effort; to broaden the user base to other UK communities; and to meet future data rates from the LHC.

- **WP5** that contains the project costs for leadership, management, administration, impact and travel.

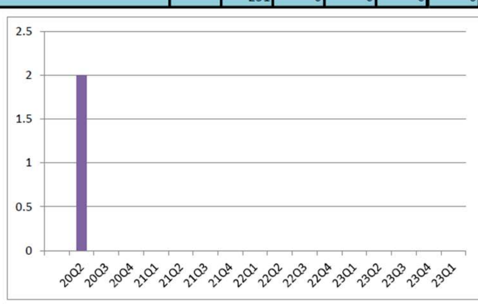
The flat-cash request for manpower (i.e. a 10% reduction from GridPP5) was further descoped by 10% by the removal of half the development effort requested in WP4, corresponding to the longer-term work envisage to prepare the infrastructure for Run-4. It is effectively this reduced scenario that has now been funded. GridPP has engaged with other funding initiatives such as ExCALIBUR, SWIFT-UK, and IRIS (described later) to help address the longer-term development shortfalls.

GridPP6 Status Report

In the first quarter (20Q2) of GridPP6 a new Project Map, consisting of metrics and milestones, was created to track the work taking place in the five GridPP6 Work Packages. As part of this process, the metrics that had been used during GridPP5 were revisited and streamlined to better meet the needs of the new project. New metrics were introduced to measure and monitor commitments to supply computing to IRIS and other VOs along with the four main LHC experiments that are core to GridPP's provision of compute to the WLCG. At time of writing, we are still processing the first quarterly reports of the GridPP6 project (the process was slightly delayed to allow the revision of metrics at the start of the new project), so only the structure of the GridPP6 Project Map can be presented in this report.

A complete set of statistics for all the metrics and milestones is shown in Figure-3 and their relationship to each individual work package is shown in the Project Map, in Figure-4. In GridPP6, 132 metrics will be tracked each quarter and are currently awaiting the input to be extracted from the first quarterly reports received on schedule by the start of August 2020.

		20Q2	20Q3	20Q4	21Q1	21Q2	21Q3	21Q4	22Q1	22Q2	22Q3	22Q4	23Q1	23Q2	23Q3	23Q4	23Q1
Metric OK		0															
Metric close to target		0															
Metric not OK		0															
Not able to be measured		2															
Suspended		0															
Awaiting input		130															
Metrics Total		132	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Milestone achieved		9															
Milestone in Progess		3															
Milestone overdue		0															
Milestone not due/ metric n/a		87															
Milestones total		99	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total		231	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0



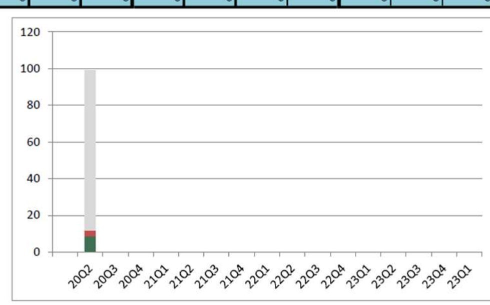


Figure-3: GridPP6 metrics and milestones: Numerically presented at the top, the lower left plot then shows the number of metrics that are currently met (green), close (amber) and not met (red). The lower right plot will show the quarterly evolution of completed milestones (currently 9 out of 99).

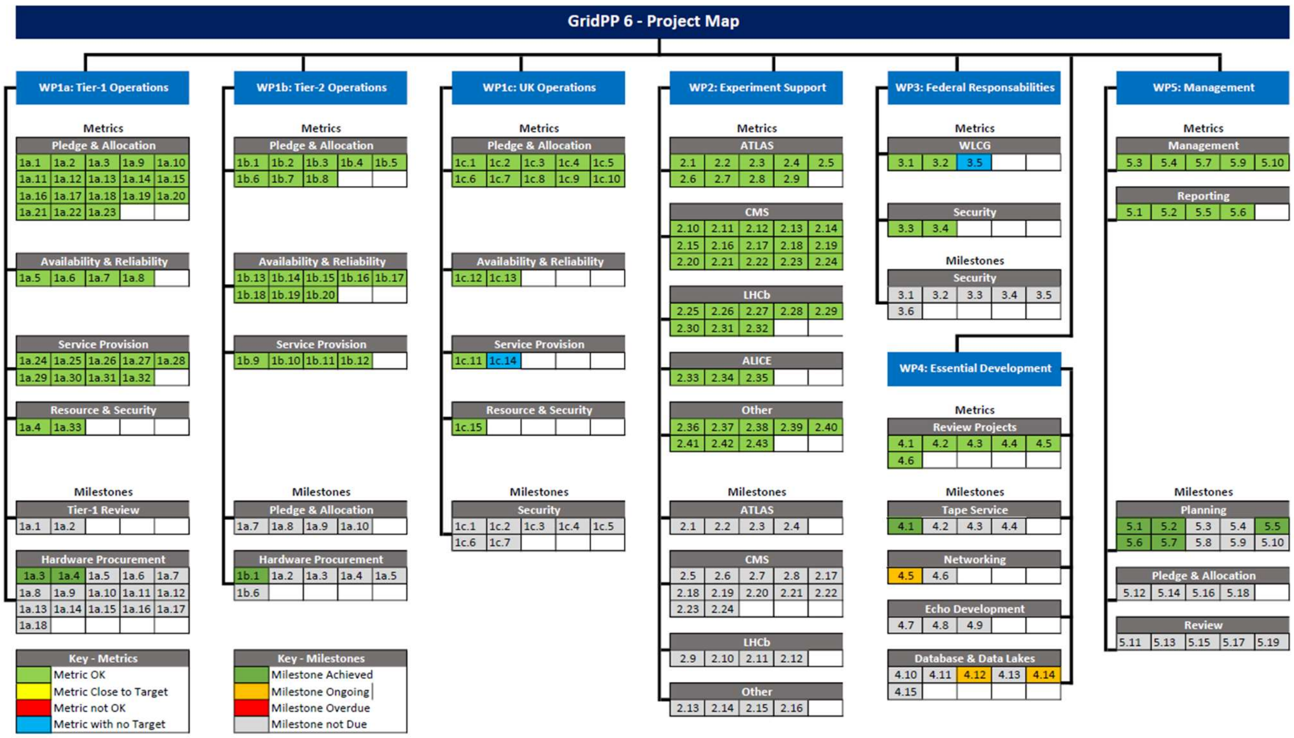


Figure-4: The GridPP Project Map at the beginning of GridPP6 (20Q2).

At the end of 20Q2 nine project milestones had been completed on time, with 3 underway as follows:

- WP4.5: Full Tier-1 LHCONE Integration: This is progressing with a procurement in progress, with ongoing work between the Tier-1 and RAL networking.
- WP4.12: Caching for diskless Tier-2s deployed. Sheffield, Birmingham and Sussex are currently access remote storage through different mechanisms, with performance and reliability and the impact on the remote site, being evaluated.
- WP4.24: Participation in DOMA. This is a long-term milestone to contribute to the DOMA WLCG development. Currently the UK is heavily involved in Third Party Copy evaluation with the Tier-1 and many Tier-2s being involved.

The quarterly reports used to track progress have been collected for 20Q2 and are in the process of being reviewed. A complete set will be made available at:

https://www.gridpp.ac.uk/collaboration/docs/gridpp-quarterly-reports/

A complete and detailed copy of the current GridPP6 Project Map can be found at:

https://www.gridpp.ac.uk/collaboration/docs/gridpp-project-map/

GridPP6 Risk register

The GridPP5 risk register was fully reviewed and used to create a new GridPP6 risk register. Following our normal practice, the risks have all been examined and evaluated with a forward look of 12-months and the summary of the current version is presented below in Figure-5. The full version, including the newer format suggested by STFC, is available at https://www.gridpp.ac.uk/collaboration/docs/gridpp-risk-register/ .

Ref.	Risk Description	Owner	Inherent Risk			Residual Risk		
			Likelihood	Impact	Total	Likelihood	Impact	Total
1	Castor Storage System Problems	AD	4	3	12	1	3	3
2	Tier-1 Tape service replacemet	AD	5	5	25	4	5	20
3	Outage of UK T1	AD	4	6	24	3	5	15
4	Failure of T1 to meet SLA or MoU	AD	7	5	35	4	5	20
5	Significant loss of custodial data at the T1	AD	6	8	48	3	5	15
6	Loss or damage to hardware at T1 >£2M	AD	0.5	10	5	0.5	10	5
7	Disaster at T1 leads to prolonged outage	AD	0.5	8	4	0.5	8	4
8	Recruitment and retention problems at RAL	DK	9	6	54	8	5	40
9	Failure to procur, deploy or operate hardware	AD,GR	6	7	42	4	6.5	26
10	Insufficient Network Bandwidth	PC	4	5	20	3	5	15
11	Over contention for resources	DB	4	5	20	3	5	15
12	Increase non-LHC use	DB	5	5	25	3	5	15
13	Unquantfiable impact of BREXIT	DB	7	8	56	7	8	56
14	Capital vs Resource at the Tier-1	DB	4	6	24	4	5	20
15	Insufficient funding at Tier-1	DB	5	5	25	2	2	4
16	Technology Mismatch	DC	6	7	42	4	5	20
17	Loss of experienced personnel at T2s	RJ	7	5	35	4	4.5	18
18	Insufficient funding at T2s for h/w	RJ	2	4	8	2	2	4
19	Tier-2s are not fit for purpose	RJ	5	6	30	3	6	18
20	Expt. s/w runs poorly	GR	7	6	42	4	4	16
21	Security problem affecting reputation	DK	7	8	56	4	7	28
22	Loss of GridPP service due to security	DK	7	5	35	5	5	25
23	Insufficient VO/user support effort	GR	7	5	35	3	5	15
24	Mismatch between budget and hardware costs	DB	9	7	63	4	6	24
25	Core service funding insuffcient	DB	5	4	20	4	2	8
26	Breakdown of NGI/EGI infrastructure	GR	3	3	9	3	3	9
27	Insufficient travel funds	DK	4	2.5	10	3	2.5	7.5
29	Critical middleware no longer supported	DC	5	6	30	3	5	15
30	Unplanned infrastructure costs	GR/PG	4	4	16	2	2	4
31	Loss of EGI.eu	DB	4	3	12	3	2	6
32	Financial Uncertainty	DB	5	6	30	3	4	12
33	Conflicting opinions amongst GridPP stakeholders	DB	3	5	15	2	4	8

Figure-5: The GridPP6 Risk Register Summary.

The following updates have been made to the Risk Register following its recent review:

- Risks 1, 2, 16, 17 and 24 have had their residual likelihood decreased from the previous version of the register.
- Risks 3, 14, 19 and 21 have had their residual likelihood increased from the previous version of the register.
- Risk 25 has had both its residual likelihood increased and its impact reduced due to a shared risk across the EU and WLCG.
- The text of Risk 19 has had its description updated to include the text "Complex middleware takes significant effort to plan upgrades, deploy and test" to highlight that it is not only broken or non-functional middleware that takes significant staff effort.
- It should also be highlighted that Risks 4, 8, 9 and 10 should have seen an overall reduction in residual likelihood due to improvements in processes, workflows and capacity. However, these have been left at the current level due to increased difficulties around the current COVID epidemic.

After a formal review, there is currently one residual risk marked as red.

Risk-13: Unknown impact of Brexit
The impact of Brexit falls into 3 broad categories:
1) Exchange rate and purchasing power volatility leading to a difficulty in procuring the required capacity to meet MOU commitments.
2) The ending of existing funding streams from EOSC and other EU sources.
3) New, restrictive operational constraints due to diverging legal frameworks (data safeguarding requirements etc.).

It is difficult to fully assess and mitigate these risks with the COVID-19 epidemic masking some of the issues. GridPP will continue to monitor and react appropriately.

The current residual risks that are moderately elevated (amber in last column of Figure-7 above) are discussed below:

Risk-8: Failure to retain or recruit key technical staff at RAL
Losing key staff impacts the project's ability to meet deliverables and key milestones. Recruiting new staff is now generally more difficult due to non-competitive salaries (though, more recently, the STFC situation has eased during lockdown). GridPP will continue to look at methods for improving its ability to retain and recruit staff, such as the use of apprenticeships and internships where appropriate. Recruitment is further hampered by the current COVID-19 situation which makes it difficult to hire and train staff; this difficulty is offset against better retention of key staff due to a reduced likelihood to move.

Risk-9: Failure to procure, deploy or operate hardware
This remains at amber level. Progress has been made at streamlining procurements at the Tier-1 and better understanding of the timescales and processes required at the Tier-2s, which has led to faster

procurements. Delivery and installation times have been made more difficult due to the restricted access associated with COVID-19. After review, GridPP has no plans for Tier-2s to procure in this financial year with Tier-1 procurements well underway which should mitigate the current issues.

Risk-21: Security problem affecting reputation
The was increased to an amber lever at our first review due to significant and sustained attacks against research infrastructures in the past year, particularly at infrastructures that are co-hosted or within the same university domains as GridPP sites and services. To mitigate this staff have recently received and will continue to receive security training from the GridPP security officer. New, closer relations are being developed with IRIS and JISC security teams and a stronger monitoring and reporting structure on patching and site status (focusing on security) is being developed.

Tier-1 Status

During the period from June 2019 to August 2020, the priority areas of work at the Tier-1 have been:

- Completing the migration of VOs to the Echo storage system and the consolidation of the Castor service into a Tape-only service.
- On-going routine operation of the production service supporting LHC work during LS2.
- Deployment of a new Spectra-Logic tape robot to replace the Oracle robot.
- Deployment of hardware to meet the 2020 LHC MoU commitments.
- Modernisation of the Tier-1 network, with goals of having a 100Gb/s connectivity to CERN via the LHCOPN as well as joining the LHCONE.

Performance Metrics

Tier-1 availability metrics were excellent over the reporting period: The average RAL availability for the LHC VOs was 99.7%, well above the 97% target and consistently above the global Tier-1 average. The availability, broken down by VO, is shown in Table-2.

WLCG Target = 97%	ALICE		ATLAS		CMS		LHCb	
	RAL	< T1 >	RAL	< T1 >	RAL	< T1 >	RAL	< T1 >
19Q3	98.3%	97%	100%	96%	98.7%	98.3%	100%	99.3%
19Q4	100%	98.7%	100%	96.7%	99.7%	99%	100%	100%
20Q1	99.3%	99%	100%	99.7%	100%	99.3%	100%	100%
20Q2	99.7%	98.3%	100%	99.1%	99%	98.7%	100%	99.7%

Table-2: Quarterly VO availability at the Tier-1. <T1> is the average over all Tier-1s.

Overall CPU utilisation was high throughout the period, averaging 92.9%, a slight drop compared to 96.5% in the previous reporting period. Average job CPU efficiency (defined as CPU time / Wall Time), however, fell to 69.9% from 80% in the previous reporting period. There were several different contributing factors to this reduction, all of which have been understood and either resolved or with a known solution that we are working towards. These are described in the Operational Issues below.

It should be noted that low CPU efficiency is not necessarily a problem, it is often the result of a necessary but unusual job requirement (e.g. requiring significant amounts of offsite data or needing a large amount of memory). Jobs are packed densely on nodes so that even if a job is not using a CPU other jobs can, this ensures that hardware is not idle. The Tier-1 Liaisons play a vital role in investigating how these jobs run and helping the infrastructure evolve to meeting these changing needs.

Installed and Delivered Capacities

During this period the Tier-1 took delivery of both disk and CPU procurements for FY19/20. At the time of writing the FY20/21 tenders have been launched with procurement decisions expected to be made in the first half of September. Significant work has been put into improving the procurement process, which resulted in better hardware and excellent value for money.

For CPU, the FY19/20 tender was awarded to Dell. The machines have more memory per core than previous years as well as SSD (rather than HDD) storage. The price drop compared to FY18/19 was 53% per HS06! This enabled the Tier-1 to buy significantly more CPU than originally planned. The main reasons for this reduction were.

- A combined procurement with the SCD Cloud increased the total value of the tender to £1.5M.
- The release of the AMD EPYC CPUs meant genuine competition between the CPU vendors.
- Making the tender more vendor friendly e.g. removing unnecessary requirements, giving vendors plenty of time to submit their bids and greatly simplifying the criteria.

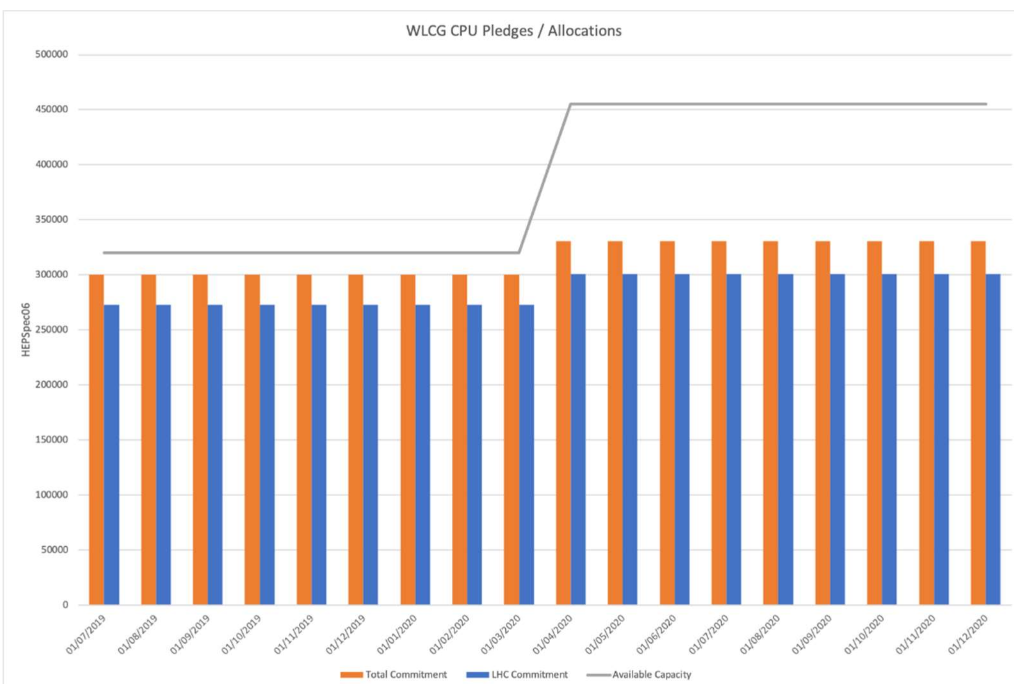


Figure-6: Tier-1 CPU capacity. The grey line shows the resources available for allocation; the blue bars are the resources allocated to the LHC experiments to meet the pledges; and the orange bars are the total allocated to all communities.

Figure-6 above shows that GridPP is meeting its WLCG commitments.

For storage, the tender was also awarded to Dell. The hardware now has a 5-year warranty (previously 3-years) to meet the longer hardware lifetime that was required during GridPP6 planning in order to manage the capital shortfall. The price drop compared to FY18/19 was 29% per TB, again better than the expected fall of 20%.

Figure-7 below shows the monthly disk capacity in the Echo (yellow line) and the phasing-out Castor (black line) systems; these add up to the grey line. The blue bars show the total allocated to the LHC experiments and the orange bars show the total allocated to all experiments, demonstrating that the WLCG pledge was met.

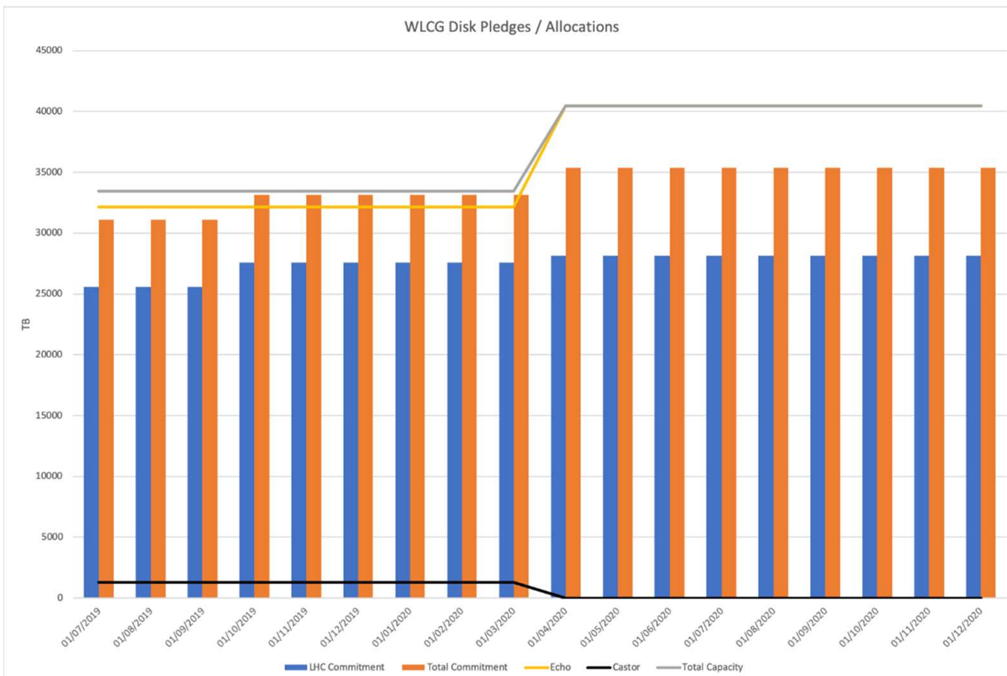


Figure-7: Tier-1 Disk Resources. The yellow and black lines are the resources available for allocation in Echo and Castor, respectively, which sum to the grey line; the blue bars show the resources allocated to the LHC experiments to meet the pledges; and the orange bars show the total allocated to all communities.

Tape Archive Evolution

The purchase of a new tape library was planned for the start of GridPP6. However, in November 2019 additional capital funding was made available to GridPP, £500k of which was used to advance these plans. A tape library was purchased from Spectra Logic and installed in February 2020. Fortunately, all the necessary work that required physical access to the tape library was completed before the COVID-19 lockdown limited access to the data centre. The new tape library entered production in May 2020.

The library is currently equipped exclusively with IBM enterprise drives and media. However, we have the ability to fit the cheaper community LTO drives and media at a later date when they release their next generation of drives and media which should meet our capacity density requirements.

Since the start of GridPP6, additional capital has been spent on the tape library, bringing forward the tape migration plans. To date, £1.3million has been spent on the system, which currently has 63PB of available storage. There is now enough free space so that further expansion to the library will not be required in GridPP6, unless new user-groups are adopted. The migration of data from the Oracle library to the Spectra Library system, started in August 2020 with LHCb data and is expected to progress at 5PB per month for 12 – 15 months.

Work has also been ongoing to replace Castor, the CERN developed storage management software that has been used for many years at RAL to provide both access to both disk and tape storage ("Castor" referred to in Figure-7, is the disk-storage system at RAL managed by the Castor software). The replacement project is managed and run by SCD (because Castor is used to provide tape-access to other communities), although GridPP is contributing 1 FTE of effort to the development work. The hardware has been procured and the majority is now cabled up, although the work to connect it to the RAL core network has been delayed due to the lockdown. Some delays have been introduced, although we still expect the new service to be production ready by the start of Run-3. Figure-8 shows the total tape storage commitment required to meet the UK pledge to the LHC experiments (blue bars) and the commitment to all communities (orange bars). It also shows the capacity of the Oracle SL8500 Robot (black line), the growing capacity of the Spectra Logic TFinity Robot (yellow line) and the total capacity (grey line). The actual LHC Pledge (green line) is also shown as this is slightly higher, reflecting the fact that we install capacity as needed and the LHC experiments did not quite require the full pledge.

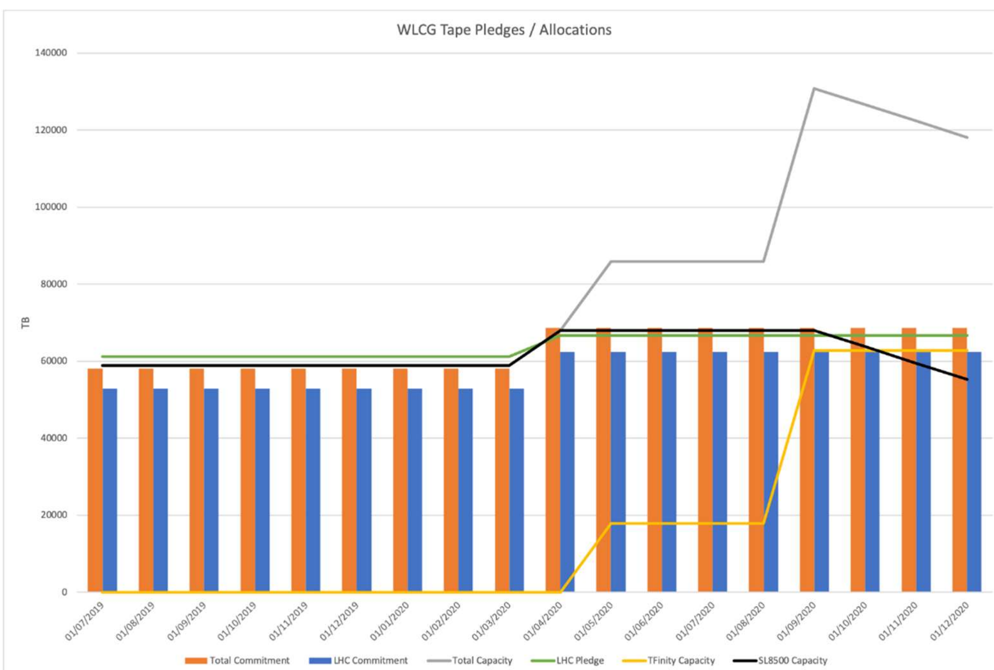


Figure-8: Tier-1 Tape Resources. The blue bars are the resources allocated to the LHC experiments; the orange bars are the total allocated to all communities; the yellow line shows the capacity available in the TFinity Robot; the black line shows the capacity available in the SL8500 Robot; the grey line shows the total available capacity; and the green line shows the LHC Pledge.

Operational Issues

During the last period, the following significant operational issues were observed:

- The migration of the ALICE VO to the Echo storage system did not go as smoothly as hoped. After the switch to Echo was completed it was found that the ALICE production jobs were different from the test jobs sent to the Tier-1 to validate the Echo setup. Two issues were found, both requiring bug fixes in the XRootD software, which took several months to fix and resulted in a period of lower efficiency jobs for ALICE.
- Despite significant improvements over the reporting period, CMS still observe lower than expected CPU efficiencies for some jobs. We believe that the dominant factor is the ability of CMS jobs on our Worker Nodes to access data from external sites. This should be resolved by the planned network improvements.
- In March 2020 the Tier-1, in addition to running Folding@Home jobs submitted by VOs, also enabled Folding@Home jobs to run in the background on worker nodes in case there were any spare CPU cycles. Unfortunately, this caused a severe impact on the CPU efficiency of some job types and was disabled after three weeks but continued to be run on dedicated slots on the farm.
- Since July 2019 LHCb have been observing higher analysis-job failure rates at RAL, as compared to other Tier-1s. Due to the subtle nature of the problem it has taken until August 2020 to really understand due, primary, to additional monitoring tools only being available in later releases of the software. This is now fully understood: The cause was due to an undocumented feature of XRootD and at the time of writing we are working on development of the necessary new features to permanently fix the problem.

Developments

During this period significant evolution took place:

- In December 2019, we decommissioned the ATLAS Frontier service.
- All VOs completed their move to the Echo storage system.
- The four separate Castor instances were consolidated into a single instance, just providing a tape service.
- We completed the migration of services from the High Availability HyperV virtualisation cluster, to one running VMWare. Other groups within SCD use VMWare so the Tier-1 can make use of a shared management infrastructure as well as a depth of expertise in VMWare across the department.
- The Tier-1 has been migrating its services from SL6 to SL7 as support for SL6 ends in November 2020.

Forward look

The Tier-1 plans for the remainder of LS2 are to finish the rationalisation of services as well as upgrading the network to cope with the demands of Run-3:

- Some legacy middleware is proving difficult to decommission. The Tier-1 remains committed to decommissioning the LFC, myproxy and BDii services before Run-3 starts.

- The Tier-1 has ambitious plans to improve its network by the start of Run-3. This includes upgrading the LHCOPN link to 100Gb/s which should be completed by October 2020, joining the LHCONE network (the RAL Tier-2 in PPD has successfully done so) and improving internal connectivity by moving to a spine/lead structure and joining the SCD bridging network.
- The Tier-1 aims to replace the Castor Tape service with the CERN Tape Archive (CTA). This project will also include the migration of all tape backed data from the Oracle Robot to the new TFinity Robot.
- With IRIS funds, the Tier-1 has developed and is in the process of deploying a multi-VO Rucio instance. This will provide a data management services to non-LHC VOs as well as being a key part of the UK Data Lake development.
- The Tier-1 is working with other groups within SCD to improve its monitoring services. For certain services (e.g. on call), the Tier-1 intends to move to a commercial cloud-based solution as this should be significantly cheaper and more reliable.
- In March 2019, the 2019 generation of WN entered production. These have SSD storage and we have noticed a significant improvement in both job success rate and CPU efficiency for many job types run on them. It will take several years to completely phase out HDD Worker Nodes so we will work towards ensuring that I/O intensive jobs are primarily run on the SSD nodes to maximise the benefit they bring.

Tier-2 and Operational Status

This section of the report covers Tier-2 operations and status from Apr 2019 (19Q2) to Apr 2020 (20Q1). During this period, the activities have included:

- Procurement of hardware at the final five sites to meet the FY19/20 capital spend profile requested by STFC.
- Procurement and installation of IRIS funded hardware as well as procurements based on additional capital allocated by STFC.
- Nearly all site upgraded to the latest version of DPM with a new more efficient core called DOME.
- Migration is underway at all sites to the newer ARC6 or HTCondor-CE with the end of EGI support for Cream-CEs.
- Imperial completed its move from an older data centre to the JISC shared data centre run by Virtus at Slough.
- Completion of a £14M data centre at Glasgow; the migration of Tier-2 resources underway but delayed due to the lockdown.

Performance Metrics

Tier-2 availability metrics for the LHC experiments were good over the reporting period; they are broken down below by experiment.

ALICE Availability

Alice	19Q2	19Q3	19Q4	20Q1
UKI-SOUTHGRID-BHAM-HEP	39%			
UKI-SOUTHGRID-OX-HEP	98%	100%	96%	100%

Table-3: Quarterly ALICE availability at supporting Tier-2s

ALICE is only supported at two sites (excluding the Tier-1) within GridPP. During 19Q2, Birmingham decommissioned its last Grid facing CE and supplied its full capacity to ALICE via VAC[3] to reduce demands on manpower. This has meant standard WLCG availability monitoring is no longer available and the low 19Q2 number does not flag any issue. Oxford was above the 90% threshold for all four quarters covered in this report.

ATLAS Availability

ATLAS	19Q2	19Q3	19Q4	20Q1
UKI-LT2-QMUL	87%	97%	88%	94%
UKI-LT2-RHUL	96%	87%	98%	100%
UKI-NORTHGRID-LANCS-HEP	89%	93%	99%	100%
UKI-NORTHGRID-LIV-HEP	98%	98%	97%	100%
UKI-NORTHGRID-MAN-HEP	97%	94%	99%	98%
UKI-NORTHGRID-SHEF-HEP	99%	79%	72%	72%
UKI-SCOTGRID-ECDF	71%	99%	97%	90%
UKI-SCOTGRID-GLASGOW	99%	95%	100%	97%
UKI-SOUTHGRID-BHAM-HEP				
UKI-SOUTHGRID-CAM-HEP	92%	67%		
UKI-SOUTHGRID-OX-HEP	92%	91%	87%	100%
UKI-SOUTHGRID-RALPP	99%	100%	96%	99%
UKI-SOUTHGRID-SUSX				

Table-4: Quarterly ATLAS availability at supporting Tier-2s

As with ALICE above, Birmingham in 19Q2 and Cambridge in 19Q3 moved to provide CPU resources to ATLAS via VAC[4] (reduces required manpower) hence the lack of availability monitoring. Sussex has been classified by as a Tier-3 and is, therefore, not monitored as a Tier-2 but does supply resources to ATLAS and local researchers.

- In 19Q2 QMUL was below the 90% target due to migration of worker nodes from SL6 to Centos6, Lancaster was below 90% because of issues with the site's DPM storage system, and Edinburgh was below 90% due to a failed CE which required replacement.
- In 19Q3 RHUL was below 90% because of decommissioning of its Cream-CE and replacement with HTCondor-CE, Sheffield was below 90% because of a migration to Centos7 worker nodes.

[3] https://www.gridpp.ac.uk/vac/

- In 19Q4 QMUL was below 90% because of transfer problems with the site's StoRM SE, Sheffield was below 90% because of issues with the site's DPM storage system and Oxford was below 90% because of issues with the site's DPM storage system.
- In 20Q1 Sheffield was below 90% because of issues related to upgrading the site's CE to ARC6.

CMS Availability

CMS	19Q2	19Q3	19Q4	20Q1
UKI-LT2-BRUNEL	97%	99%	98%	97%
UKI-LT2-IC-HEP	88%	100%	100%	98%
UKI-SOUTHGRID-BRIS-HEP	88%	96%	85%	96%
UKI-SOUTHGRID-RALPP	98%	99%	97%	95%

Table-5: Quarterly CMS availability at supporting Tier-2s

The overall availability for CMS was well above the 90% target for most of the report period with the following exceptions:

- In 19Q2 Imperial was below 90% availability due to its move to the JISC Slough facility, and Brunel was below 90% due to upgrading network links leading to some disruption accessing CMS data.
- In 19Q4 Brunel was unavailable due to electrical upgrades in its Data Centre.

LHCb Availability

LHCb	19Q2	19Q3	19Q4	20Q1
UKI-LT2-BRUNEL	98%	99%	100%	99%
UKI-LT2-IC-HEP	89%	100%	99%	100%
UKI-LT2-QMUL	89%	93%	100%	46%
UKI-LT2-RHUL	95%	87%	97%	92%
UKI-NORTHGRID-LANCS-HEP	95%	99%	98%	100%
UKI-NORTHGRID-LIV-HEP	100%	100%	100%	100%
UKI-NORTHGRID-MAN-HEP	100%	100%	100%	100%
UKI-NORTHGRID-SHEF-HEP	98%	92%	66%	92%
UKI-SCOTGRID-DURHAM	88%	100%	100%	96%
UKI-SCOTGRID-ECDF	89%	99%	100%	93%
UKI-SCOTGRID-GLASGOW	100%	98%	100%	100%
UKI-SOUTHGRID-BHAM-HEP	100%	100%	100%	100%
UKI-SOUTHGRID-BRIS-HEP	98%	100%	97%	98%
UKI-SOUTHGRID-CAM-HEP	99%	100%	100%	100%
UKI-SOUTHGRID-OX-HEP	97%	100%	96%	100%
UKI-SOUTHGRID-RALPP	100%	100%	97%	99%

Table-6: Quarterly LHCb availability at supporting Tier-2s

The overall availability for LHCb was well above the 90% target for most of the report period with the following exceptions:

- In 19Q2 Imperial was below 90% due to its move to the JISC Slough facility, QMUL was below 90% because due to migration of worker-nodes from SL6 to Centos7, Durham was below 90% because of issues rebuilding an ARC-CE disallowing LHCb payloads and Edinburgh was below 90% because of a failed CE.

- In 19Q3 RHUL was below 90% due to decommissioning of its Cream-CE and replacement with HTCondor-CE.
- In 19Q4 Sheffield was below 90% because of problems with the site's DPM storage system.
- In 20Q1 QMUL was below 90% because of upgrades to StoRM (Storage system) that lead to failures in LHCb data transfers.

Overall job CPU efficiency averaged 81% in 19Q2, 83% in 19Q3, 80% in 19Q4 and 76% in 20Q1 and is well above the 50% target set for Tier-2 resources. Overall, the Tier-2 service is operating within normal parameters.

Installed and Delivered Capacities

During this period five of the Tier-2s carried out procurements, but due to COVID-19 not all sites were able to complete the process before universities were shut for delivery. However, the relevant grants have been given a no-cost extension and it is anticipated that all hardware will be delivered and installed before the associated grant closure deadline. The integrated installed capacities at the Tier-2s provides sufficient headroom to meet the pledges due to (a) hardware funded by the institutes; (b) hardware funded by IRIS and used opportunistically by LHC workloads; (c) old hardware that the institutes have chosen to keep running well beyond end-of-life.

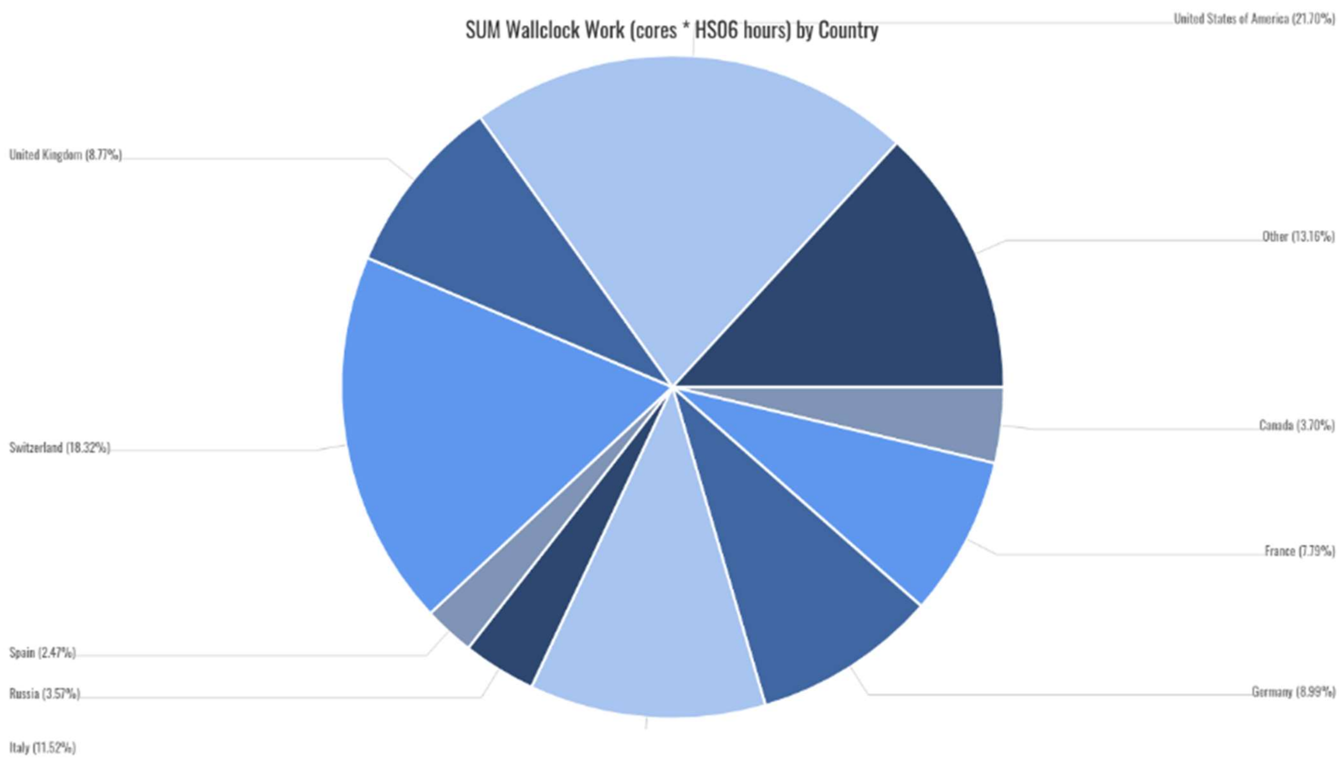


Figure-9: UK Global contribution of computing resources to the LHC experiments.

Operational Issues

There were relatively few operational issues:

- End of support for the Cream-CE has forced sites to move to new software stacks (ARC and/or HTCondor-CE) which has led to some downtimes and associated loss of availability. Experience developed at Liverpool with HTCondor-CEs has been used to successfully set up CEs at Imperial and Bristol to replace Cream-CEs.
- Migration from ARC5 to ARC6 has taken place at many sites preparing for the deprecation of CentOS/SL6 and the potential deployment of service to CentOS8.
- Aging storage systems have been affected by increased data movement from the ATLAS data carousel system and from deletion campaigns to move to the new system. Glasgow, Lancaster and Manchester have reported some (non-critical) lost data due to equipment failures.
- Sites have moved to the latest version of DPM and encountered problems associated with newer version of XRootD that led to reduced availability.
- During this reporting period a major security incident was recorded at UK HPC facilities, often within the same institution as GridPP resources. GridPP took an active role in responding to this situation to ensure that GridPP resources were safe. We engaged with partners in IRIS, JISC and EGI to support investigative procedures.
- GridPP continued to participate in the *Duty Security Officer* rota. GridPP security staff continue leadership and participation in the WISE (Wise Information Security for Collaborating e-Infrastructures) coordination body.

Developments

During this period the following development/testing work took place:

- Sites have begun experimenting with CentOS8 in preparation of new EGI/UMD deployments.
- The DIRAC middleware has been developed further to add functionality for non-LHC VOs, augmenting its ‘transformation’ capability, this includes the inclusion of Glue2 publishing allowing it to work with newer CE technology.
- A CEPH-based storage system is being deployed at Glasgow as a replacement for DPM. Glasgow is working closely with the Tier-1 to leverage their gained operation knowledge and has identified problems with newer version of XRootD plugins.
- Birmingham and Cambridge completed the move to VAC and became storage-less for ATLAS, with storage requirements being met by associated sites (Manchester and QMUL).
- Sheffield also became a storage-less site with the RAL Tier-1 acting as it’s remote SE endpoint.

Forward look

The Tier-2 plans for the remainder of 2020 focus on work related to lightweight sites, and reducing operational overhead:

- Glasgow plans to complete its move of Tier-2 resources to new data facilities to ensure more resilience in site operation.
- The remaining non-IPv6 connected Tier-2s intend to continue working with local network teams to progress dual-stacking of storage services.

- Sites will work with local university networking teams to provide cases for updated network links where needed with increased usage due to remote storage.
- Development of the Data Lake model of distributed storage partially funded by new proposals to enable experiments and local researchers to better access LHC and other data.

Users' Reports

ATLAS

The UK has continued to meet its pledges to ATLAS in terms of capacity and availability. This has been done in the face of growing experiment demand and limited budgets and has benefited from 'extra' capacity donated from universities. The Tier-1 has delivered the expected disk capacity, as shown in Figure-10, with small outages of fractions of the capacity in the Castor-based storage before the migration to the Echo system was complete. The Tier-1 provided an additional 2PB of disk, as requested, in August 19, as an advanced deployment of the 2020 pledge. Over the last few months, the Tier-1 has remained available to receive and serve ATLAS data (see Figure-11) and there is no evidence of COVID-19 affecting performance.

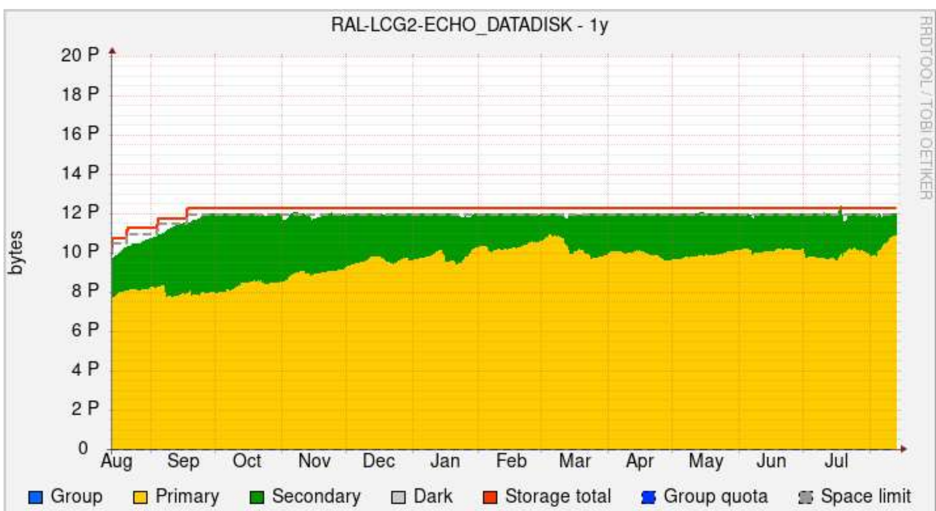


Figure-10: The available and used disk capacity at the UK Tier-1 for ATLAS.

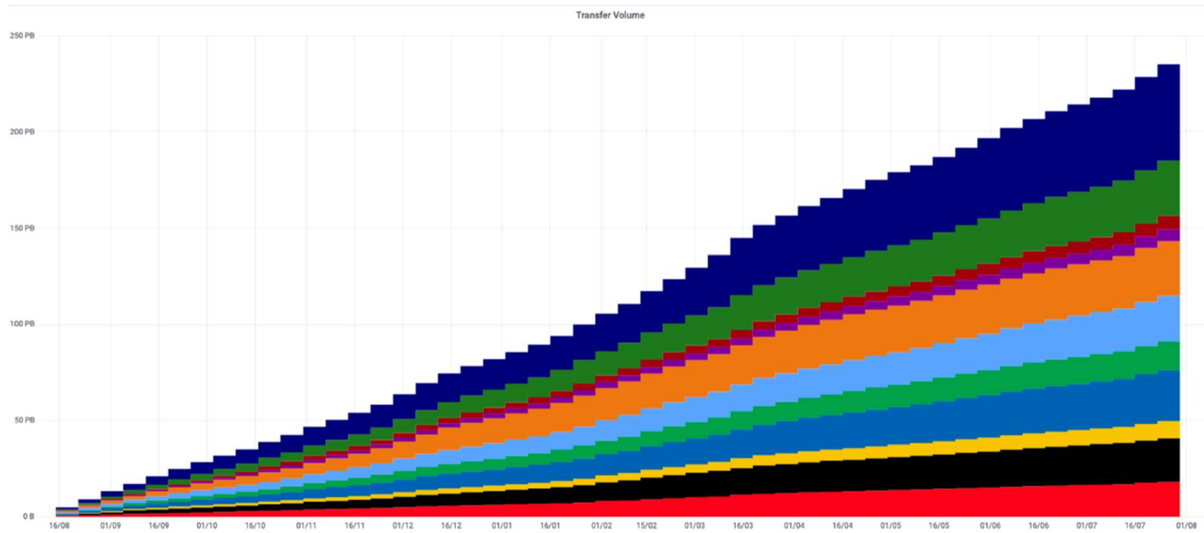


Figure-11: The cumulative ATLAS data transfers into the UK Tier-1 (second to top green contribution) Aug-19 to Aug-20 compared to other countries.

Overall, the service has been excellent. Some data has been lost, inevitably, due to disk errors, but all necessary datasets have been recovered. Some reductions in capacity were seen as the consequence of a necessary and valuable migration to new storage technologies such as CEPH; and from the planned migration of some Tier-2 sites to storage-less mode. This is illustrated in the 'DATADISK' available capacity for at our Tier-2s, shown in Figure-12.

Figure-12: The Tier-2 ATLAS DATADISK available storage capacity for ATLAS in the UK. The temporary drop in available capacity reflects the transition period to CEPH storage at one major site.

The CPU delivery has also been excellent, with the UK meeting its expected share at both the Tier-1 (Figure-13) and Tier-2s (Figure-14).

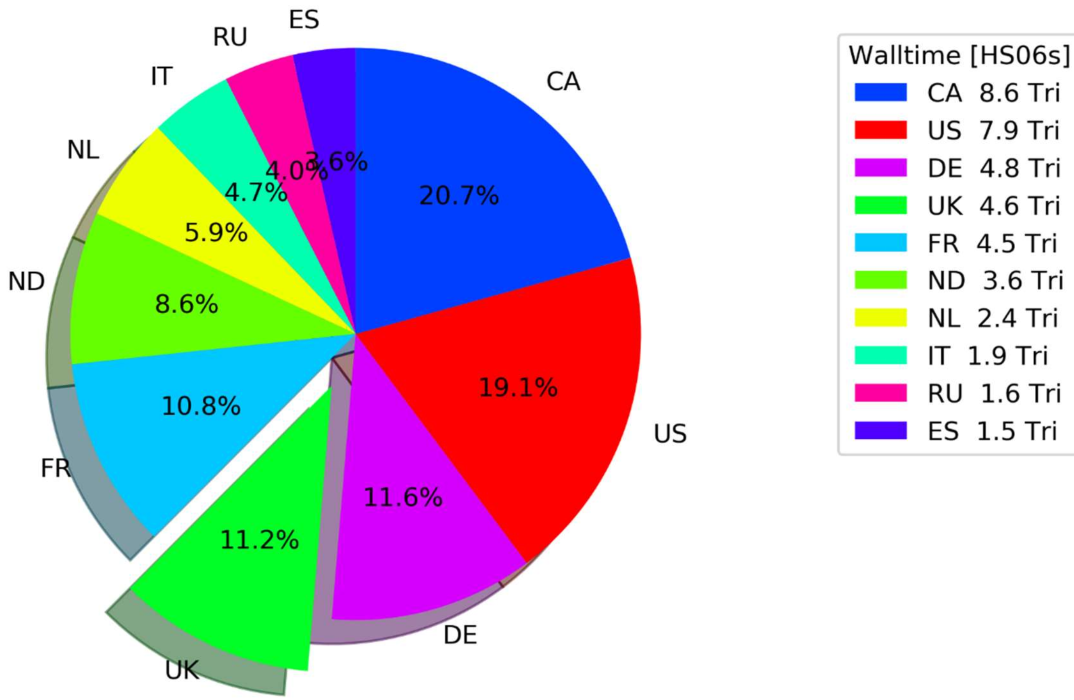


Figure -13: The UK ATLAS Tier-1 wall-time delivered Aug 2019-Aug2020.

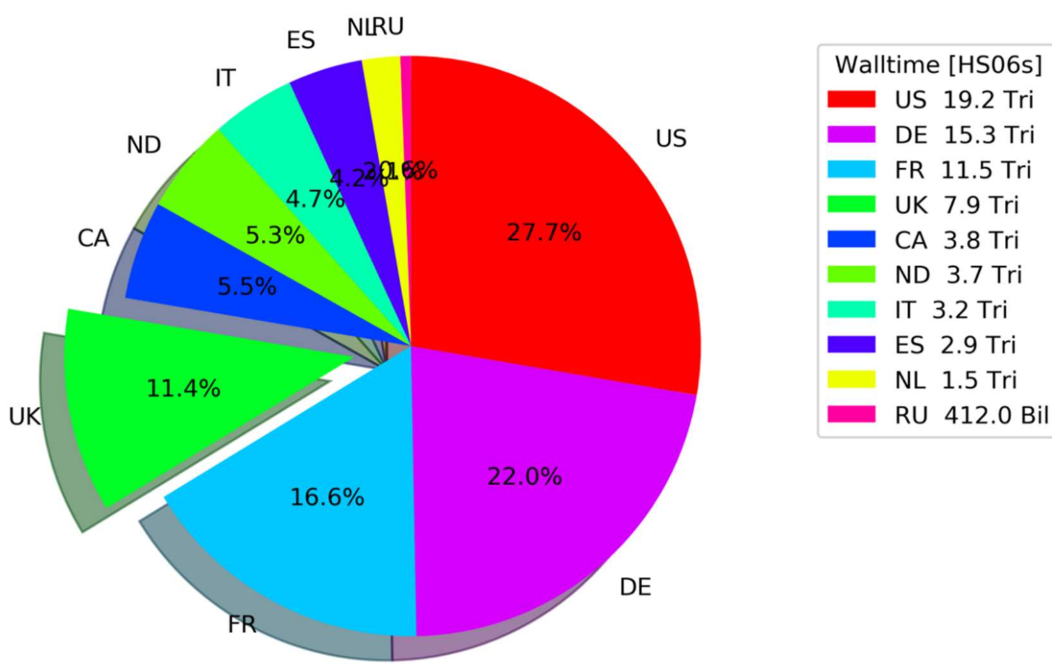


Figure-14: The UK ATLAS Tier-2 wall-time delivered Aug 2019-Aug 2020.

Much of the ATLAS activity in the last year has focused on testing the increased use of data staging to-and-from tape to minimise disk costs, and in particular the data carousel model. In this, data is selectively recalled from tape storage to disk in a coordinated manner (sequential rather than random) and accessed by a set of related computing tasks, minimizing the data traffic. The UK has both engaged on the ATLAS side and delivered high performance in the data staging at the Tier-1, attracting explicit praise. This builds on the generally excellent performance from UK, and the good response when issues arise. This was true both before and during the effects of COVID-19. While machine room access has slowed some processes, the sites have been engaged and pro-active in minimising disruption.

Both ATLAS-facing GridPP staff and the ATLAS UK experiment staff have shown good engagement with WLCG and HSF working groups. One area that might be highlighted is around DOMA, and we note that ATLAS effectively had in-house, many of the elements needed for data lakes/federated data as part of the experiment's existing toolkit.

Both GridPP and ATLAS demonstrated swift and effective redeployment of resources (both hardware and some effort) to support COVID-19 research using the ATLAS jobs framework and UK capacity. This was limited to 20% of the total ATLAS capacity and done in agreement with the experiment management.

The UK ATLAS Tier-1 liaison post holder changed hands in February with no negative effect on the service, and the new hire is integrated and delivering excellent performance both for ATLAS and the Tier-1.

CMS

The period since the last oversight committee has been during LS2. The large datasets that were collected before LS2 have made this a period of unprecedented productivity with many outstanding physics analyses going to publication. CMS submitted its 1000th paper for publication on the 19th June 2020. The ability to produce such a variety of physics results shows that in general, the computing system continues to work well. LS2 is also a period of preparation for Run-3 and even beyond to the high luminosities of Run-4. COVID-19 and other delays have pushed back the start of Run-3 until 2022.

Along with other LHC experiments, CMS is trying to improve efficiency in all areas of computing. This will be required if CMS is to meet the challenges of Run-4 and beyond without a significant increase in spending, which is unlikely to be available. Significant effort is going into maximising the performance of existing CPU and storage resources. There is likely to be an increased dependence on GPUs and Machine Learning techniques, which are already ubiquitous in physics analyses, during some elements of reconstruction, and even at the trigger level.

Opportunistic resources continue to play an increasing role in CMS computing. These include the extensive use of the High-Level Trigger (HLT) farm outside of data taking, and time that is made available on HPC resources, most notably in the USA but now also in Europe. It is worth noting that the HLT farm is roughly the size of all of CMS' Tier-1 resources combined.

CMS plans retire its data management system of the last ~20 years, PhEDEx, by the end of November 2020. Work continues to develop, test and migrate to the Rucio software, which has been used by ATLAS for many years. Rucio is becoming a community standard and is being adopted by several experiments both inside and outside of HEP. Currently all CMS NanoAOD datasets have been moved into Rucio and are no longer managed by PhEDEx. This migration went smoothly without any user disruption. CMS is optimistic that the migration of other data tiers will be as smooth.

In the same time-frame CMS will be migrated from the Castor tape system at CERN, to CTA. Preparations are being made to make this move hand-in-hand with the Rucio transition. RAL Tier-1 tape will also adopt the CTA software during 2021.

A new Tier-1 liaison for CMS started in September 2018 and has become established within the Tier-1 and the broader CMS community.

<u>CMS' Use of the Tier-1, Tier-2 and Tier-3 UK sites</u>

The UK pledged to provide 8% of CMS' global Tier-1 requirement during GridPP5 which was reduced to 6.8% in GridPP6. As can be seen from Figure-15, the UK has significantly exceeded that in CPU delivery during the period of this Oversight Committee with the Tier-1 site at RAL. This is due to additional capacity available at the Tier-1 after the unexpectedly good price of the last procurement.

	total	percentage
US	208 Mil	33.7%
RU	109.4 Mil	17.8%
DE	74.4 Mil	12.1%
UK	73.7 Mil	12.0%
IT	64.2 Mil	10.4%
FR	63.7 Mil	10.3%
ES	23.0 Mil	3.7%

Figure-15: Core Hours used by CMS at Tier-1 sites since the last Oversight Committee.

The Echo disk storage used by CMS at RAL Tier-1 is much more tolerant of hardware failures than the former Castor storage system. Thanks to the massive parallelism offered by the architecture, it is able to sustain much higher levels of total data throughput. This is illustrated in Figure-16, which shows a short period of high rate access (all VOs), for read (green) and write (orange).

Figure-16: The data rate read (green) from and written (orange) to Echo over a period of 24 hours.

CMS jobs running at RAL Tier-- have become increasingly inefficient when compared to other Tier-1 sites in the period since the last Oversight Committee. In the previous period RAL had an efficiency 10-15% lower than other Tier-1s. In this period, the average efficiency of jobs at RAL was 58%, which is ~20% lower than the other Tier-1s, all of which sat between 72% and 80% on average. The efficiency of each Tier-1 across the period is shown in Figure-17, with the UK in blue.

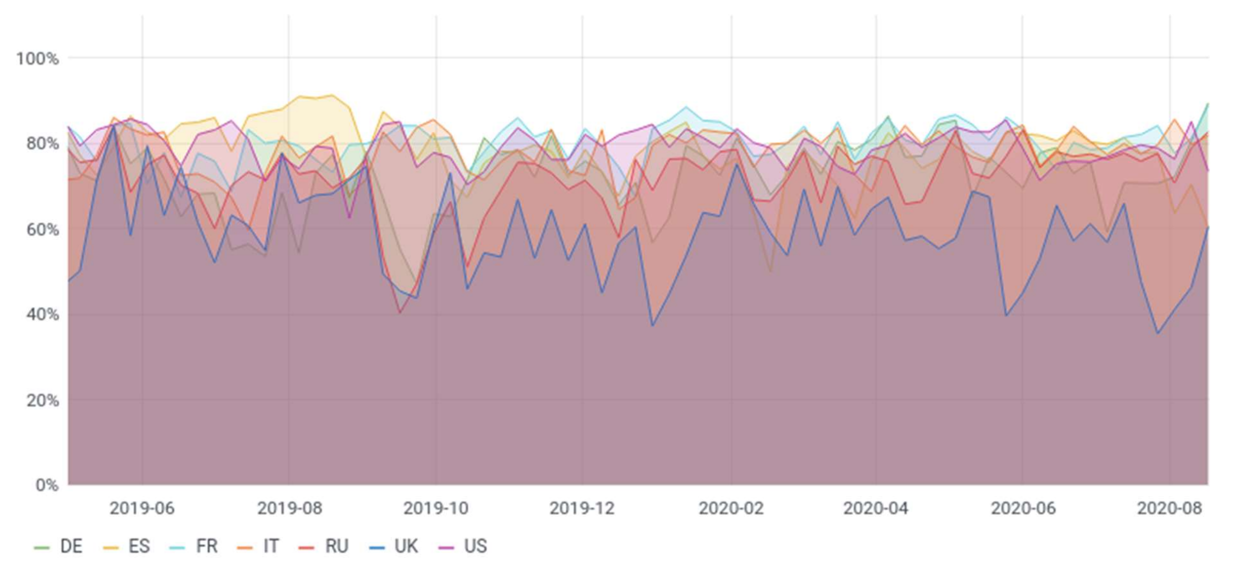


Figure-17: Average CPU efficiencies at Tier-1 sites since the last Oversight Committee.

The cause of low efficiency is believed to be dominated by the reading of input data, and in particular data stored offsite which is accessed via the AAA service. This is currently a subject of vigorous investigation and the CMS liaison is pursuing possible reasons why offsite reads are affecting RAL more than any other Tier-1 site. The amount of data read by RAL jobs tends to be high, but the read time is disproportionately much higher than other sites, as seen clearly in Figure-18. Major Tier-1 changes planned for the near future, such as becoming part of the LHCONE network may go some way to reduce the read time, particularly from offsite storage.

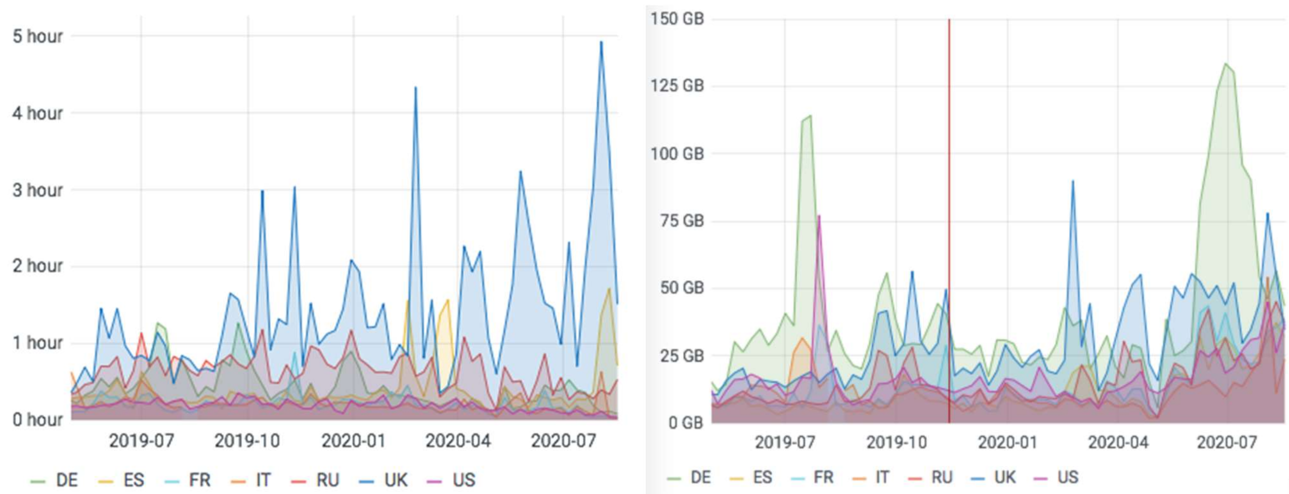


Figure-18: CMS job input data read time (left) at Tier-1 and data volume (right).

The job failure rate at RAL Tier-1 is in line with other Tier-1 sites, and in fact 30-50% of these are frequently of type ‘Log Collect’. This job type has been failing at RAL at a rate of 100% for several years, and although the problem has been identified, the CMS developer has not found the time to implement the change because the issue is not that important.

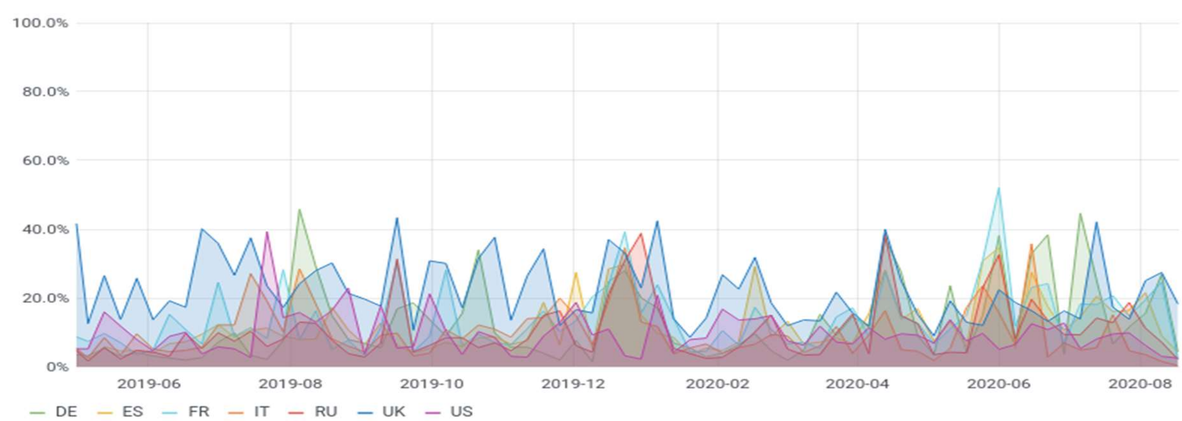


Figure-19: Failure rates as a percentage of completed jobs at Tier-1 centres since the last Oversight Committee.

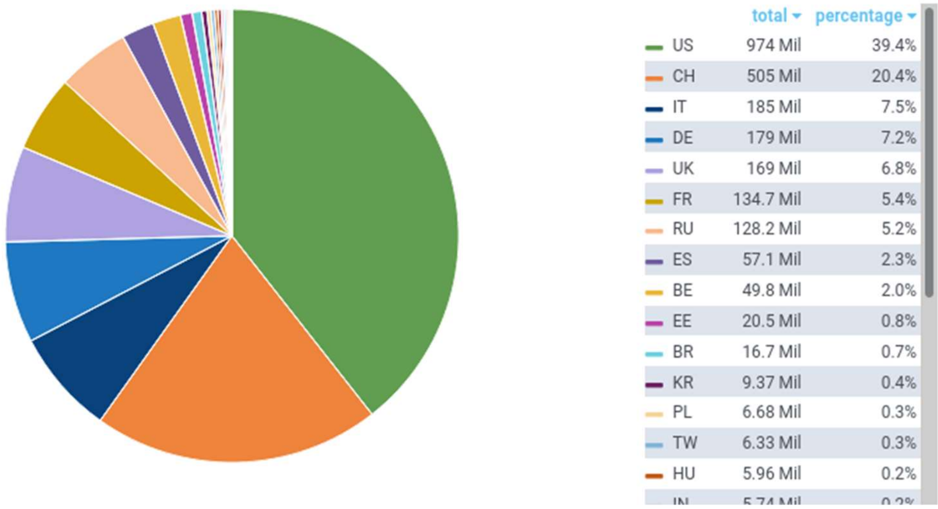


Figure-20: Core Hours used by CMS at Tier-2 sites since the last Oversight Committee.

The UK pledged to provide 5% of CMS' Tier-2 needs during GridPP5 and is almost the same in GridPP6 (4.9%). Figure-20 shows that the UK has provided 6.8% of the Tier-2 CPU used by CMS during the reporting period, primarily due to the sites running older hardware beyond planned lifetimes.

While CMS makes extensive use of its significant Tier-2 sites at Brunel, Imperial and RALPP (Bristol is called a Tier-2 site for historical reasons but no longer fulfils the core role of a Tier-2 site and so should be regarded as a Tier-3 site) it also makes opportunistic use of resources at Tier-3 sites, many of which are sites with no CMS physics activity. Since the last Oversight Committee meeting this has been about 12% of the non-Tier-1 CPU used by CMS in the UK. This can be seen in Figure-21.

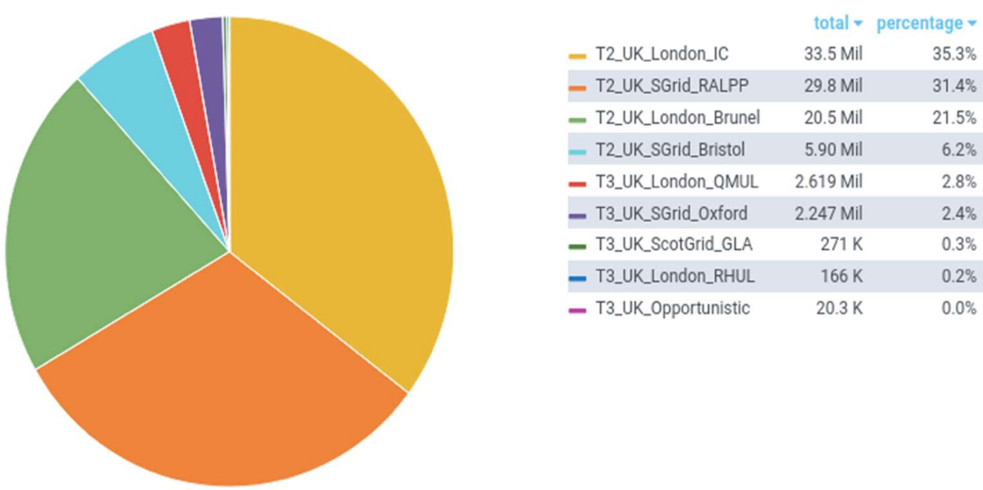


Figure-21: Core hours for CMS, split by UK site (excluding Tier-1), including the very small use of cloud sites (both private and commercial) in the UK.

LHCb

The UK has continued its leading role in LHCb and has been the largest national contributor of CPU power, reflecting the size of the UK collaboration. Between 19Q2 and 20Q3, UK has supplied 25% of the total CPU used by LHCb and provided by national contributions, and 18% of the Tier-0/1 CPU, meeting the MOU pledges. LHCb has appreciated both the capacity and, as importantly, the reliability and availability of the resources. Figure-22 below shows CPU usage at Tier-1 sites (left) and Tier-2 sites (right) during 19Q2 to 20Q1 across all countries, with larger UK sites in leading positions. Similarly, Figure-23 below shows the disk usage at Tier-1 sites (left) and Tier-2 sites (right) during the three quarters. From the LHCb point of view the UK pulls its weight and responds quickly and positively to any issues that arise.

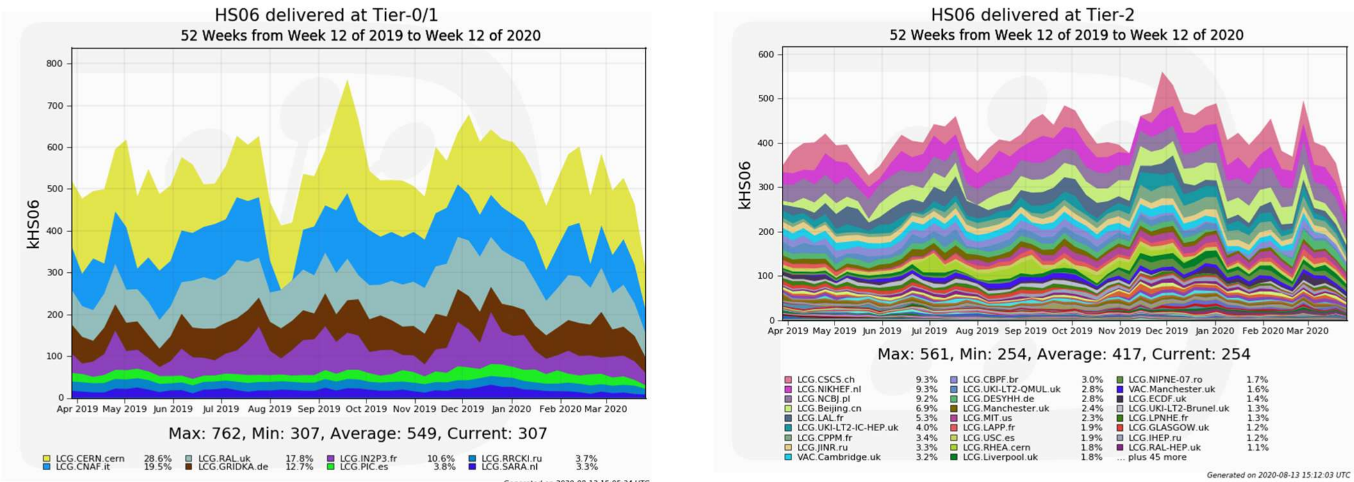


Figure-22: LHCb CPU usage at Tier-0/1 (left) and Tier-2 (right) in 19Q2 – 20Q1

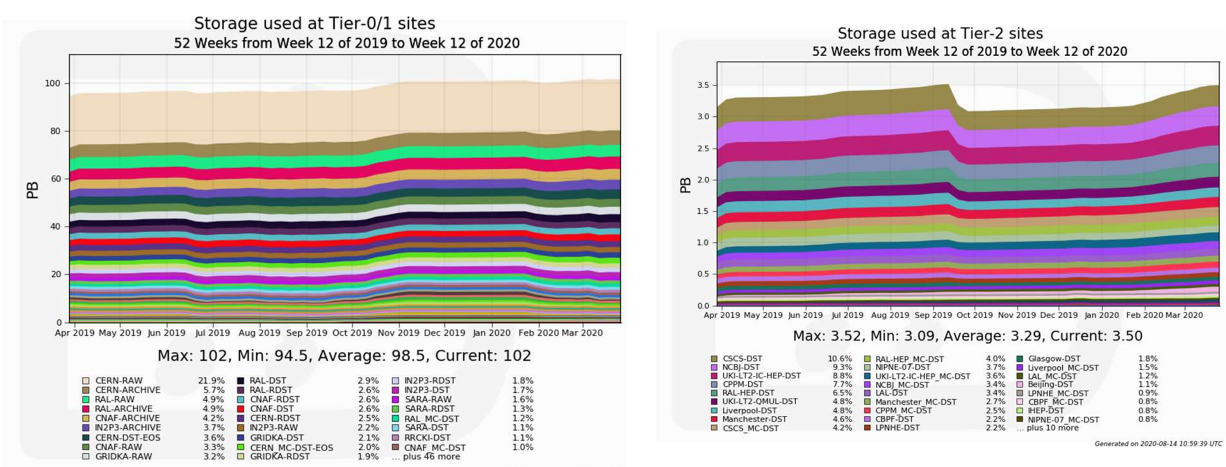


Figure-23: LHCb storage usage at Tier-0/1 (left) and Tier-2 (right) in 19Q2 - 20Q1

During 2019 and into 2020, LHCb used computing resources for the production of simulated events and in running jobs for individual users and for physics working groups. 2019 also saw a campaign to reprocess all the proton-proton collision date from Run-1 and Run-2 of the LHC in a consistent way. Finally, processing of heavy ion and fixed-target collisions data taken in 2018 was done, as a way of smoothing the profile of computing resources needed by pushing work forwards into a shutdown year. UK sites played significant roles in all these activities.

In preparation for the increased data rate during Run-3 with the upgraded LHCb detector, the experiment has looked for non-Grid capacity, from High Performance Computing Sites and Cloud Computing providers. The UK has contributed to this effort, in particular support for the commercial cloud provider Yandex. Although useful, these non-Grid resources are an order of magnitude smaller in total than the Grid.

To mitigate the increasing requirements for computing capacity, which is dominated by simulation work, LHCb designed and implemented various fast simulation options in the past years, offering speed improvements over the full simulation. In 2019, fast simulations accounted for the majority of produced events. As a consequence, the number of simulated events per year correspondingly increased by a factor 4 over the previous years while the total CPU work increased only by 20% with respect to 2018. To save storage resources, only events passing analysis-dependent selection criteria are persisted on storage. The Fast Simulation activity is led by the UK.

LHCb changed its original computing model to make use of storage at larger Tier-2 sites as well as its original focus on Tier-1 sites. The UK now provides the largest number (Manchester, RAL PPD, Imperial, QMUL, Liverpool, Glasgow) of the 14 Tier-2 sites with storage worldwide.

During the period the LHCb data management team, Tier-1 liaison and Tier-1 staff have completed the transition from Castor storage at the RAL Tier-1 to Echo for LHCb. The RAL Castor tape service continues to be used by LHCb.

Other VOs

Under GridPP5, Non-LHC experiments are permitted to make use of up to 10% of GridPP's CPU resources (and 5% of storage). Figure-24 shows, for the period from August 2019 to August 2020, the relative use made of this facility by non-LHC experiment by the amount of compute (HS06) consumed. The figure includes all work done on GridPP nodes, including some contributions from work done for IRIS.

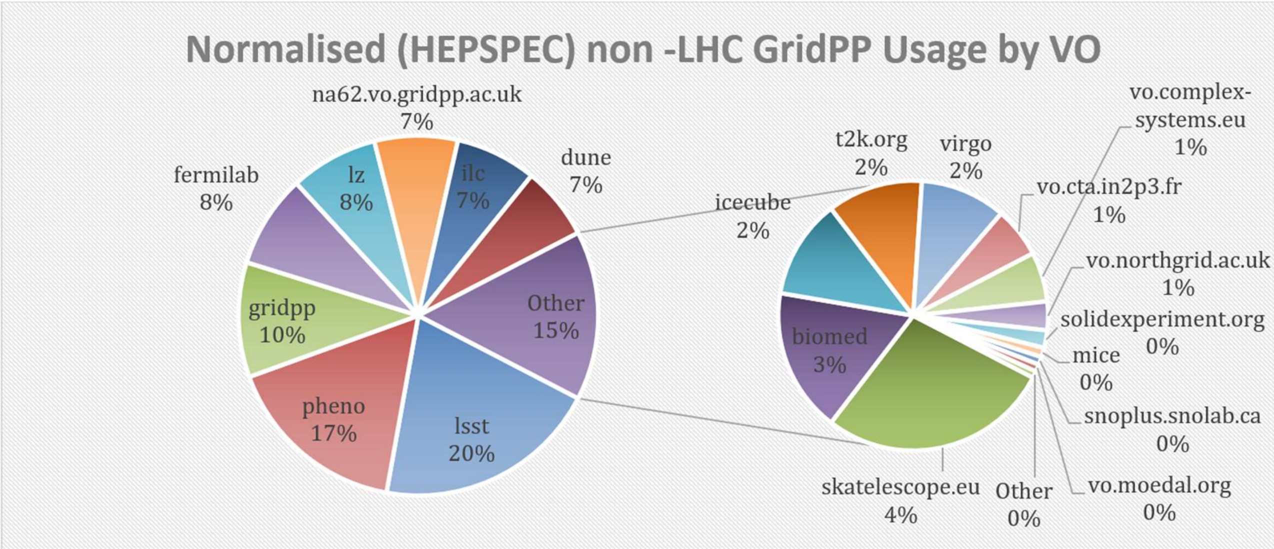


Figure-24: Excluding ALICE, this pie chart shows by normalised CPU time usage (HEPSPEC06) the main non-LHC users of GridPP resources between August 2019 and August 2020. It should be noted that the GridPP VO is typically used to 'incubate" early adopters.

This rest of this section describes work undertaken with, or in support of, the non-LHC VOs:

CVMFS: GridPP continues to provide a CVMFS Stratum-0 service at the RAL Tier-1 for non-LHC VOs.

DUNE: GridPP continues to support DUNE as a key international partner (GridPP members are also the DUNE Technical Coordinator and Chair of its Computing Board). GridPP provided ~ 20% of the processing power and is now storing ~ 2.3PB of data in the UK as per the DUNE policy request for major partners. The CPU and storage hardware were funded by GridPP, local contributions, and IRIS, operated and integrated by GridPP funded staff.

NA62: GridPP continues to support NA62, in particular at Glasgow for production databases and servers, working in conjunction with DIRAC with additional support from Imperial. Services are managed with support from GridPP and local PPE IT at Glasgow. All of the collaboration's Monte Carlo data is generated by the UK production system and GridPP recently agreed to provide tape storage for a backup copy of half of the collaborations data, matching an offer from Italian groups. This is currently 2PB and migration is underway.

LSST: GridPP continues to support LSST through IRIS, and support from GridPP funded staff. LSST-DESC ran a major Data Challenge (DC2) in the last year and demonstrated the ability to get several tens-of-thousand jobs a day through GridPP to contribute to an international activity. LSST will carry out its DC3 on GridPP

SKA: GridPP members at are helping SKA with their SRC (Science Regional Centre) development. They have in particular supported approximately ten high memory nodes integrated into the Manchester GridPP Tier-2. Members at Imperial, RAL and Edinburgh are helping with testing of the use of Rucio data management software for SKA use. The GridPP Project Leader is on the SKA Science Board.

LIGO: GridPP has supported LIGO peripherally in the last year (CVMFS, some use at RAL. However, LIGO are now looking to increase their use of Grid jobs submitted via OSG, and members in RAL, Glasgow and Lancaster are assisting.

EUCLID: Has recently started making use of cloud resources hosted by GridPP at Imperial using Openstack access. This was enabled by GridPP staff and now supplies approximately 30% of EUCLID CPU time.

LZ: In the past year LZ made use of both of the standard GridPP compute resources on a variety of sites and also used high memory nodes provided by IRIS through the GridPP DIRAC instance. The focus was on the last Monte Carlo data challenge before the start of data taking. The high memory nodes were used to generate calibration data, which by construction contains orders of magnitude more particles than the (expected) data and hence has a larger memory footprint than the average 'Grid' compute node can provide. All these activities received strong support from GridPP personnel. Storage (900 TB via GridPP, 300 TB via IRIS, all presented to the experiment as one continuous storage element) was provided by the Tier-2 site at Imperial.

JLAB/CLAS12: The Nuclear Physics experiment at JLAB called CLAS12 has been running ~ 500 job continuously at the Glasgow Tier-2 via the OSG.

T2K: All T2K data is now handled through the DIRAC File Catalogue and job submission continues to be done through the DIRAC instance hosted by Imperial and supported by GridPP and IRIS resources. All raw data for the T2K near detector and the processed data, MC simulation and calibration data are stored on GridPP resources. T2K CPU usage is dominated by Monte Carlo production but also used for calibration jobs. Additionally, T2K has two non-GridPP SEs, at IN2p3 and SFU, which the GridPP DIRAC group have configured to work with the DFC. T2K has recently been working on a containerised model for distribution/development/continuous-integration and now tests are underway to extend this to running Grid jobs using container images unpacked onto CVMFS.

GridPP funded staff also provide modest user support to facilitate joint activities by T2K, HyperK and Belle II, in particular running T2K jobs on the INFN ReCaS Napoli Cloud which uses VCYCLE to manage VMs for DIRAC job submission. This is facilitated by the GridPP/Imperial DIRAC group, who have configured this cloud into the GridPP instance of DIRAC for the purpose of this short-term proof of concept test.

SoLiD. SoLiD has been a GridPP supported VO for a while and we have seen intermittent usage. GridPP is coordinating with IRIS to ensure that the group gets the resources needed, either through direct use of the GridPP non-LHC capacity, or via a more formal request to the IRIS resource review board to use GridPP-hosted IRIS resources. At present, the VO is not consistently using significant resources but there have been some expectations that this will increase in 2020.

SNO+: GridPP provides one of the three main storage sites used for SNO+ data and Monte Carlo production. SNO+ continues to make use of tape resources at RAL (increased from 200 TB to 300 TB over the last year) and modest levels of compute using GridPP. The VO members continue working with the Tier-1 to migrate away from using the LFC.

Hyper-K: At the start of 2019 HyperK began using the Grid for test Monte Carlo productions with a small amount of storage allocated at QMUL (20TB). Job submission was handled by DIRAC and all output was registered in the DIRAC File Catalogue (DFC). This work was completed and at present activity is focussed on a new round of testing prior to a full production run later in the summer.

GridPP: GridPP-VO continues to be used as a catch-all incubation VO.

Others: In addition to the VOs with operational/deployment changes as noted above, GridPP resources have also run (production) work in this reporting period for: Biomed, Cepc, Enmr, Icecube, ILC, Pheno, MoEDAL, CERN@School, and complex-system.eu (which supports work in Complexity Science).

Wider Context

GridPP sits within a wider context of computing across all of STFC and now UKRI. We describe our relationship to some of these external bodies below. This is largely unchanged from the previous Oversight Committee document, but we include it for the benefit of the new members.

IRIS: GridPP is a Partner in IRIS (iris.ac.uk), which is coordinating all computing activities across PPAN and the National Facilities. GridPP is a member of the IRIS 'Delivery Board' (governing board) along with other PPAN interests (DiRAC, the large Astronomy and AstroParticle projects). GridPP supports other IRIS partners on a best-effort basis. GridPP was a part of the BEIS submission which resulted in a £16M (2018-2022) award to cover the greatest computing urgencies across STFC. This provides some of the resource GridPP needs to fully support the new larger HEP experiments (DUNE and HyperK), as well as LZ. The GridPP Deputy Project Leader is also the Scientific Director of IRIS. GridPP members have also been leading an important Security Policy and response initiative.

UKRI: GridPP members participate in many UKRI eInfrastructure activities. Some are continuous such as the JISC led AAAI work. Some are ad hoc such as the recent call to enable COVID-19 computing. We ensure that the facts and practicalities of STFC eInfrastructure requirements are well represented in all discussions and decisions. We are particularly involved in the JISC AAAI group. We also maintain constant communications on networking issues.

CERN SCF: The Project Leader and Deputy attend the CERN SCF (Scientific Computing Forum) Chaired by the CERN Research Director. This body meets twice yearly and considers matters of strategic import to the future of HEP computing.

HSF: GridPP is an active member of the HEP Software Foundation which facilitates coordination and common efforts in high energy physics software and computing internationally and in 2017 produced a roadmap white paper on the software and computing challenges that will be faced during the next decade.

ExCALIBUR: We were prime instigators in a successful bit to UKRI (via EPSRC) for a part of the ExCALIBUR exa-scale computing initiative. This pilot award of approximately £250k is for software work needed for exa-scale LHC experiment software, and includes: Data management, real time tracking on FPGAs, detector simulation and portable parallel strategies. The PI is a member of the GridPP PMB.

SWIFT-HEP: We are a member of the SWIFT-HEP consortium that includes all HEP experiments and GridPP. SWIFT-HEP has recently submitted a proposal to the PPRP for approximately £1.2M over 4 years for the first phase of a UK software-for-HEP project to parallel international initiatives such as IRIS-USA and the HSF. If successful, this will develop the experiment centric advanced software needed for the HL-LHC.

COVID-19 Computing: GridPP participated fully to calls made to make computing resources available for COVID-19 work. Our principle contribution was to 'FoldingATHome' where we have so far run over 200,000 jobs (ranked higher than 400 in the world) and as part of WLCG (ranked 21 in the world). We have been involved in introducing the Rucio data management system to FoldingATHome following a meeting arranged with their management. We also put in significant effort to deploy and make available the Imperial College Group epidemiology simulation code used to provide the intervention advice to the government.

EGI / EOSC-Hub: EOSC-hub is the EU H2020 project for e-Infrastructure/Research Infrastructure integration and operation 2018-2020, with €30M of EC funding, EOSC-hub brings together the EGI Federation, EUDAT CDI and INDIGO-DataCloud into one project, to deliver a common catalogue of research data, services and software for research. STFC leads the task on Operational Security, charged with coordinating incident response and harmonising policies/procedures across the EOSC community, and STFC also continues to operate and develop the Grid Operations Centre Database (GOCDB) and APEL accounting services. Proposals have been submitted in June 2020 for the final round of EU Horizon 2020 funding for EGI/EOSC for the years 2021-2023. STFC is a partner in both projects: "EOSC Future" with €40M and "EGI-ACE" with €8M proposed EC funding. STFC continues to lead Operational Security for EGI-ACE and is also a major partner in the Security team for EOSC Future. STFC has also achieved a proposed increase in funding for GridPP's GOCDB and APEL services and significant new funding for Rucio, FTS and CVMFS. The outcome of these proposals will be known in the autumn.

Impact

The GridPP collaboration has impact beyond the direct technical and scientific impact as the provider of the UK's computing contribution to the WLCG. The indirect impact of the technical work has benefitted organisations within the collaboration, such as:

- Imperial College London and QMUL received prestigious awards from the UK IPv6 Council for IPv6 adoption as a result of supporting WLCG use cases. These awards acknowledge our leading role in IPv6 roll-out.
- STFC was invited to become a founding member of the CEPH Foundation as a result of their experience running the largest 'Erasure Coded' CEPH cluster in the world for the WLCG.

Over recent months, the rapid-response from GridPP, sites, and personnel to requests to support COVID19-related investigations, has allowed significant tangible impact, as described earlier. Not only were relatively large amounts of computing resources brought to bear on problems, very quickly, but technical assistance was also provided that has benefitted FoldingATHome and the Scottish Centre for Research on COVID.

The wider public has also benefitted through the diverse public engagement efforts that GridPP staff undertake, though this kind of impact has not been systematically recorded in recent times because of the lack of funded project effort. However, in GridPP6 with the funding of 0.1FTE, there will be a focus on systematically capturing data relating to the impact generated by the project and its staff. Each quarterly report will ask the reporter to outline any outreach and knowledge exchange activities that may have taken place, such as publications, dissemination events, awards, academic and non-academic collaborations etc. These reports will be followed up and developed to produce cases studies highlighting the impact produced. These case studies will be published on the website and shared through social media channels. They will cover both large-scale impact stories as well as multiple smaller stories that share a theme. Impact will also be highlighted at ongoing GridPP Collaboration Meetings, with the intention of highlighting partnerships and awards similar to those indicated above.