STFC Computing Programme Evaluation: Proforma for UKHEP-Compute/GridPP

Project/Programme Name: UKHEP-Compute : Experimental Particle Physics Programme Computing – response through GridPP on behalf of experiments.

Principal Investigator/Lead Contact: Professor David Britton
Organisation: University of Glasgow

Dear Malcolm

As requested here is our response to the questionnaire.

As communicated to the Chair of the Review Committee, the questionnaire is not very well matched to the context of Experimental Particle Physics Computing. We are concerned that answering the questions as formulated might unintentionally provide a misleading picture. At best such input would not help us or the committee; at worst it could lead to damage to the sector that we represent.

The mismatch arises because, whilst we recognise that it is difficult to create questionnaire that can be used by widely different infrastructures, the current formulation is more suited for a leading edge HPC resource that creates new (simulated) data from which science is then extracted. The questions address issues related to the value of that science; whether the data would be better created, or the science done better, by competing facilities; and how the facility producing the data is pushing the innovation boundaries of computing.

None of these issues have much, if any, relevance in our sector. High Energy Physics computing infrastructure is a deterministic commodity resource that is required to carry out the science programmes of the experimental particle physics collaborations. The science has been peer reviewed and approved elsewhere through the PPGP and PPRP, which has led to the construction of detectors and the formation of collaborations. The compute should be regarded as the final part of the detectors themselves and not as the primary source of the science. We will describe in our submission how the required volume of computing resources then follows, is estimated and scrutinised. Without this understanding, the connection of the computing to the large STFC investments already made, might be missed.

In addition, the computing facilities do not "compete", but rather, they "collaborate", with international partners to provide the UK share of the global computing facilities required. Hence the questions on “leadership in compute”, as if it were a self-standing research project, are not applicable.

Furthermore, the GridPP infrastructure is based upon commodity compute clusters and storage that have been deliberately selected at the lowest price point per unit resource to provide the best value for money. So by construction there is no research “adventure” element in the computing itself. Some of the questions, however, presuppose the compute is a self-standing competitive research project using leading edge computing technology, which is not the case.

Given this mismatch, we therefore provide the information in a more natural language, and then as an appendix try to answer the original questions as best we can, but with the caveat that some are difficult to interpret or simply don't apply, and so could be misleading if taken at face value or compared directly to other inputs.

Since no other particle physics experiment was sent a questionnaire (apart from DUNE), our response covers the broad aspects of the compute requirements of almost all of the experimental particle physics programme, including:

- The LHC Experiments ATLAS, CMS, LHCb, and ALICE.
- In operation experiments such as T2K, NA62, SNO+ and MICE.
- In build experiments (simulation) DUNE, HyperK, Calice, Comet, ILC/CLIC.

Dave Britton (GridPP Project Leader) & Pete Clarke (GridPP Deputy Project leader)
Tim Gershon (LHCb UK Spokesperson) & Chris Parkes (LHCb UK Upgrade Project Leader)
Gavin Davies (CMS UK Spokesperson)
Nikos Konstantinidis (ATLAS UK Spokesperson) & Craig Buttar (ATLAS UK Upgrade Project Leader)
Alfons Weber (DUNE UK Spokesperson)

….on behalf of the UK Experimental particle physics community.

UK Computing for Experimental High Energy Physics

Scientific Drivers for Experimental HEP

The high level drivers of GridPP are the scientific drivers of the Experimental HEP programme, which can be found in:

- The PPAP Roadmap.
- The experimental programmes of the experiments themselves, which are peer reviewed and approved by the PPRP for construction and by the PPGP for exploitation.

We do not attempt to summarise the entire programme here and we assume the Committee will obtain these documents directly from STFC if required. The LHC experimental programme currently dominates the computing requirements but computing is no less critical in all the other science areas.

Context of computing for Experimental HEP

When the discovery of the Higgs Boson was announced in 2012, the then Director General of CERN, Rolf Heuer, noted that it would not have been possible without three critical contributions: those of the LHC accelerator, the LHC experiments, and the WLCG computing infrastructure. This demonstrates the intimate and essential function of scientific computing in High Energy Physics (HEP), which together with the accelerators and experiments, leads to the science output.

Experiments are approved, built, and then exploited over decadal timescales and embody long-term strategic decisions. The data that they produce are acquired, calibrated, reconstructed, simulated and stored, by a set of processes called the "production pipeline" that must be executed before any data are available to the collaboration scientists to exploit for specific physics analysis. The computing needed to do this is, in effect, the last stage of the detectors themselves, in the sense that if it were not present then there would be no physics data products to exploit. Production computing is, therefore, a distinct set of computational processes that are different from end-user physics analysis.

Production computing has the following characteristics:

- The volume of resource required is deterministic, depending upon the accelerator running schedule, beam luminosity, approved science programme, event rates and individual event sizes.
- For the LHC experiments, the required resources are requested and robustly scrutinised on an annual basis, as described later.
- Production is carried out by centralised international computing teams that are small compared to the number of scientist collaborators.
- Production is a constant 24x7 process that drives the required service level, though the balance between different workflows changes according to priorities and experiment running schedules.
- The workflows include
    - Triggering.
    - Calibration of the detector.
    - Monte Carlo simulation of events using sophisticated software models of the physical detectors and state-of-the-art physics generators.
    - Reconstruction of raw data (lists of individual hits in detector elements) into the physics objects used for the analysis (tracks, clusters, jets, etc.).
    - Selection/stripping/skimming/slimming of reconstructed event samples via broad criteria into streams suitable for analysis by physicists.

- Re-reconstruction from time to time as detector parameters become better understood or techniques are improved.
- Data distribution according to resilient preservation policies.

Analysis of the data by the physicists within the collaborations then follows, using the same compute infrastructure.

For the LHC experiments, the largest call on CPU is for Monte Carlo simulation production, and this is true for many other experiments because, fundamentally, discovery is made by understanding backgrounds to signal, and understanding and quantifying systematic errors. It is perhaps worth noting that in the last year the LHC has been performing very well, leading to more data than expected. This has led to a shortfall globally in computing resources by some 20% compared to that which has been funded under flat cash. As a result, at least one experiment (LHCb) has seen an impact on its publications because of limited production of Monte Carlo simulated events.

All of this computing requires only high-throughput-computing (HTC) resources. There is, to first order, no true parallel computing requirement (i.e. true HPC where several processors communicate or share memory or pass messages).

The physical computing infrastructure – the WLCG

The computing infrastructure for particle physics revolves around the WLCG (Worldwide LHC Computing Grid). The WLCG infrastructure was conceived as a hierarchical structure of HTC computing centres based on a single Tier-0 at CERN, a small number (currently 13) of national Tier-1 centres, and a large set of Tier-2s (currently about 160) spanning 42 countries. The tier-level is defined by role/responsibility and service level, allowing the use of a wide range of resources from University-based clusters to major data-centres. Over time, and reflecting the much better than anticipated growth in network provision and advances in middleware, the hierarchical structure has evolved into a more matrix-like arrangement and the experiments can increasingly use resources in different tiers in a flexible and transparent manner.

Production computing workflows are increasingly run across the entire distributed computing infrastructure. Historically, the emphasis for production computing (except for Monte Carlo simulation) has been on using the Tier-0 centre at CERN and the national Tier-1 centres, but the Tier-2s are now also a vital component for processing real data. The Tier-2s have always provided the main resource for Monte Carlo production. In contrast to organised production computing, physics analysis is performed by a wide range of individual scientists and teams, with the Tier-2s being the most important facility.

GridPP runs the UK part of WLCG, representing about 10% of the global resources. In addition to the LHC experiments (ALICE, ATLAS, CMS and LHCb) GridPP also provides resources for a range of non-LHC collaborations (SNO+, Solid, T2K, MoEDAL, Icecube, ILC, Pheno, Comet, DUNE, Biomed, Calice, CEPC, Cern@school, ENMR, Geant4, LSST, LZ, MICE, NA62, OSG, SKAtelescope). The UK Tier-1 centre at RAL is one of only three that serves all four LHC experiments. The UK hosts four distributed Tier-2s (ScotGrid, NorthGrid, SouthGrid and London) which group together 17 physical sites across the country into logical interfaces. The UK sites are interconnected via the JANET academic network and the Tier-1 is connected to CERN via the private LHCOPN optical network. Over the years, GridPP has developed and maintained and excellent working relationship with JISC/JANET, regularly providing a network forward look document to flag any issues and future needs.

The Tier-2 sites host compute and disk-storage resources; additionally, the Tier-1 hosts a tape-based archival storage system. The level of resources funded at individual Tier-2 sites through GridPP reflects both the requirements of the experiment(s) they support and a performance element based on past delivery. The latter aspect has encouraged and rewarded well-run sites and institutional investment in both resources and facilities. In the GridPP5 proposal, we estimated the total "leverage" of resources from the UK Tier-2s (including, hardware, electricity, manpower and machine-rooms) at £2.4m/year over a period of eight years between 2005 and 2013.

At both the Tier-1 and Tier-2 sites, the disk and CPU resources are based on commodity hardware that has reached the optimal value-point on the price-performance curve. That is, the hardware is not cutting edge but tends to be mass-produced, tried-and-tested technology, which delivers the largest resource for the money. The distributed infrastructure is of enormous scale, consisting of around 700,000 cores world-wide running up to 2 million jobs per day, supported by over 400PB of disk storage and another 400 PB of tape-based storage. Up to 100 PB of data a month is moved around this infrastructure. For some very specific applications, the HEP community has an interest in other technology, such as GPUs, and GridPP hosts a few such resources that are currently used for testing software development.

There are well established support structures both within the experiments and WLCG for the infrastructure operation and for users. The UK is totally integrated with these systems and contributes to them at an appropriate level. For example, the UK runs a "Regional Operator on Duty" (ROD) team that triages UK-related tickets submitted via the German-hosted GGUS system. The UK hosts the APEL accounting database for the whole of WLCG and the GOCDB Grid Operations Centre database. These two database operations also receive some funding from H2020 projects. The UK infrastructure is supported by a security team, that provides operational advice, security challenges, and front line support in the event of an incident. The UK also leads the international security policy development for WLCG.

HEP is a collaborative rather than competitive, international endeavour, reflecting the large capital investments and long timescales involved. The computing infrastructure is bound together by a common goal and agreed levels of service and resource provision. The UK meets these targets and relies on our international partners to do the same. Formally, for the LHC there is an international MOU signed by STFC with CERN that recognises our share of, and responsibilities within, this global collaboration.

Hardware resource requirements

The UK commitments to the WLCG collaboration are determined by "authorship shares" of the LHC experiments. The UK represents about 2% of the ALICE collaboration; 12.5% of ATLAS, 8(5)% of CMS; and 30(21)% of the LHCb at Tier1(Tier2) respectively. Hence overall the UK represents about 10% of the WLCG. These numbers are used to scale the total computing resources request to obtain the required UK share. The total resource requirement is estimated and scrutinised by a well-established international process involving the LHC experiments, the CERN Computing Resource Scrutiny Group (C-RSG), and the CERN Resource Review Board (RRB). The process starts with agreement on a common set of LHC running parameters as input to experiment computing models. This includes things such as expected energy and luminosity; bunch-structure (which effects pile-up and thus the computational complexity of each event); and expected physics running time in various proton and heavy-ion configurations. The experiments use these numbers to determine how much computing resource is needed to deliver the approved physics output using the experiment-specific computing model. Next, the C-RSG then examines and compares the estimated computing resources levels produced by the experiments and tries to ensure that the requirements are both realistic and minimal; on occasion it will iterate with individual experiments to adjust the levels. Once these are agreed, the C-RSG recommends the levels to the RRB that brings together representatives from the funding agencies across the world, including STFC. Finally, once the RRB formally approves the requested levels, the UK contribution is determined by the UK authorship share.

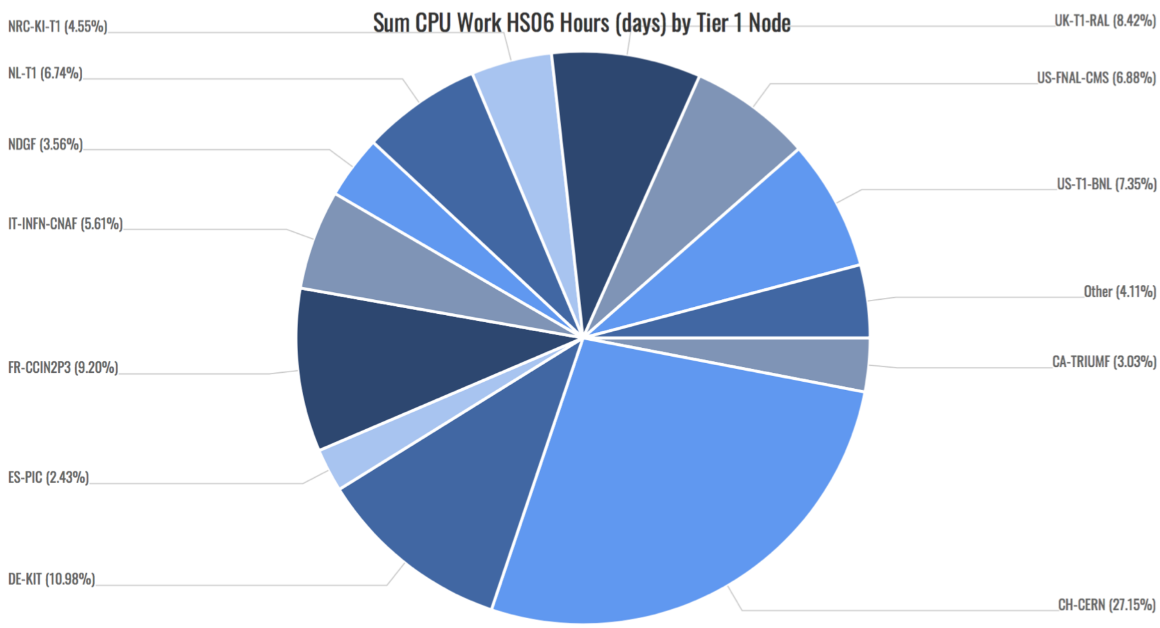
The established resource levels are then delivered according to the agreed (MOU) service levels. For the Tier-1 centre at RAL, this includes a 2-hour response time 24 x 7 x 365 for some critical elements of the service. There is also a de-facto commitment by the UK, as a partner in WLCG, to contribute to operational and working-group that ensure the appropriate function and evolution of the infrastructure.

The table below shows the UK requirement for CPU and storage for 2017-2020 at the Tier 1 and Tier 2s. The figures cover all of the LHC and most other running particle physics experiments. The only major experiments not explicitly included yet are DUNE and HyperK, which add approximately 10% in the next 4 years. The units of storage are PetaBytes (PB). The units of CPU are kilo-HEP-Spec06 (KHS06), which is a measure of power more appropriate to HEP than TFlops. The conversion for a dual Xeon 2.6 MHz processor is approximately 22 HS06/core (0.025 TFlops/core).

As an example of the scale, for 2018/19 the sum of the Tier-1 and Tier-2 requirements from the table below is:

- 535 KHS06.
- 52 PB of disk storage.
- 71 PB of tape storage.

This figure illustrates the international shares provided by the Tier-1 centres in the last 6 months.



		GridPP5 Tier-1 Resources				GridPP5 Tier-2 Resources			
		2017	2018	2019	2020	2017	2018	2019	2020
ALICE	CPU [KHS06]	256	307	375	431	366	438	475	546
(Global)	Disk [PB]	25	31	36	41	31	41	40	46
	Tape [PB]	31	41	41	47				
ATLAS	CPU [KHS06]	921	949	1057	1216	1125	1160	1292	1486
(Global)	Disk [PB]	68	72	88	101	83	88	108	124
	Tape [PB]	188	195	221	254				
CMS	CPU [KHS06]	600	600	650	748	850	900	1000	1150
(Global)	Disk [PB]	57	60	68	78	68	70	78	90
	Tape [PB]	175	188	230	265				
LHCb	CPU [KHS06]	207	253	271	312	116	141	152	175
(Global)	Disk [PB]	22	25	28	32	5	6	7	8
	Tape [PB]	43	46	51	59				
ALICE	CPU [KHS06]	5.1	6.1	7.5	8.6	7.3	8.8	9.5	10.9
(UK)	Disk [PB]	0.5	0.6	0.7	0.8	0.6	0.8	0.8	0.9
	Tape [PB]	0.6	0.8	0.8	0.9				
ATLAS	CPU [KHS06]	115.1	118.6	132.1	151.9	140.6	145.0	161.5	185.7
(UK)	Disk [PB]	8.5	9.0	11.0	12.7	10.4	11.0	13.5	15.5
	Tape [PB]	23.5	24.4	27.6	31.8				
CMS	CPU [KHS06]	48.0	48.0	52.0	59.8	42.5	45.0	50.0	57.5
(UK)	Disk [PB]	4.6	4.8	5.4	6.3	3.4	3.5	3.9	4.5
	Tape [PB]	14.0	15.0	18.4	21.2				
LHCb	CPU [KHS06]	62.1	75.9	81.3	93.5	24.4	29.6	31.9	36.7
(UK)	Disk [PB]	6.6	7.5	8.4	9.7	1.1	1.3	1.5	1.7
	Tape [PB]	12.9	13.8	15.3	17.6				
UK LHC Total	CPU [KHS06]	230.3	248.7	272.9	313.9	214.8	228.4	252.9	290.9
	Disk [PB]	20.2	21.9	25.6	29.4	15.4	16.6	19.7	22.6
	Tape [PB]	51.0	54.0	62.1	71.5				
T2K	CPU [KHS06]	1.5	2.5	3.9	6.2	1.5	0.8	1.3	2
	Disk [PB]	0	0	0	0	0.2	0.3	0.5	0.8
	Tape [PB]	2.6	3.4	4.2	5				
ILC	CPU [KHS06]	0.7	0.9	1	1.2	1	2	3	4
	Disk [PB]	0.3	0.4	0.4	0.5	0.3	0.4	0.5	0.6
	Tape [PB]	0.4	0.5	0.6	0.7				
NA62	CPU [KHS06]	2	2	2	1	3	3	2	1
	Disk [PB]	0.4	0.8	1.2	1.5	0.1	0.2	0.5	1
	Tape [PB]	1	2	3.5	5				
MICE	CPU [KHS06]	0	0	0	0	5	5	5	5
	Disk [PB]	0	0	0	0	1.5	1.5	1.5	1.5
	Tape [PB]	4	4	4	4				
Others	CPU [KHS06]	13.8	14.9	16.4	18.8	12.9	13.7	15.2	17.5
	Disk [PB]	1.2	1.3	1.5	1.8	0.9	1.0	1.2	1.4
	Tape [PB]	3.1	3.2	3.7	4.3				
Operation	CPU [KHS06]	5.8	6.5	7	7.7	6.1	6.4	7	7.7
	Disk [PB]	5	6	6.6	7.5	2	2.1	2.7	3.2
	Tape [PB]	2.7	3.3	3.9	4.5				
Grand Total	CPU [KHS06]	254.2	275.5	303.2	348.8	244.3	259.3	286.4	328.0
	Disk [PB]	27.1	30.4	35.3	40.7	20.5	22.1	26.6	31.1
	Tape [PB]	64.8	70.5	82.1	95.0				

Network Requirement

The wide area network is an integral part of the WLCG and we request that this is taken into account in the computing review. A detailed description of UK networking and the site requirements can be found at:

https://www.gridpp.ac.uk/wp-content/uploads/2017/10/179_20171004_NetworkForwardLook.pdf

In summary:

The UK Tier 1 at RAL currently has 3 x 10 Gbit/s dedicated links to CERN as part of the so called LHCOPN (Optical Private Network). The usage is closely monitored by both the Tier 1 and the STFC networks group, and it was recently increased to this level due to approaching saturation of the 2 x 10 Gbit/s/ links.

The Tier 2s typically have 10 Gbit/s connections to the JANET backbone, though at larger institutions this is now increasing.

This network capacity is provided by JISC-JANET with whom we liaise very closely to ensure that any looming problem would be headed off (although there rarely is, since JANET plans well). We request that the JANET criticality to STFC is noted in the review. We can provide more information on networking if required.

Distributed computing software infrastructure for experimental HEP

Distributed computing software infrastructure is just as essential to HEP computing as the physical hardware, and indeed is required even if all hardware resource were provided by a third party.

Some of this manpower is required within the experiments (layers A and B in the table below) and some within the distributed computing infrastructure of the WLCG/GridPP (layer C). Historically, the experiments have independently developed their own software but there are now significant community efforts to develop common software for the future, where possible, which offers the possibility of reducing the overall effort required.

Layer	Experiment – 1	Experiment – 2	Experiment - 3	Community
A	Experiment specific computing support to maintain and develop reconstruction and analysis software	Experiment specific computing support to maintain and develop reconstruction and analysis software	Experiment specific computing support to maintain and develop reconstruction and analysis software	Community common software development projects
B	Experiment specific effort for production and data management	Experiment specific effort for production and data management	Experiment specific effort for production and data management	
C	Common distributed computing software infrastructure deployment, operation, support and development. Global services such as security response, monitoring and accounting			

Much of the software in layer-A above is specific to each experiment and will always need to be supported within the experiments. The production frameworks in layer-B could in principle be more generic, but they are also the hardest part to change within an experiment's software stack because they link into everything. Production operation will always remain within the experiment's remit. The distributed computing software infrastructure in layer-C is common to all experiments and the manpower to deploy, operate, and develop this is provided by GridPP.

Support is required for the development, deployment and operation of the software infrastructure. As described above, these responsibilities fall within both the experiment and Grid layers. Continued development is essential in order to constrain the growth of the hardware resources required and to reduce the overall amount of manpower needed to operate the combined infrastructure.

The difference between layers A, B and C is often apparently not appreciated by review bodies, and a false assumption seems to be made that GridPP provides computing support to the experiments for layer-A and B. This is not the case (it only provides layer C) and has led to a severe under support for layer A and B in recent experiment awards.

Future software roadmap.

The community development projects referred to in the final column of the table above are now starting to be identified and promoted internationally by the HEP Software Foundation (HSF). In 2017 it produced a roadmap white paper on the software and computing challenges that will be faced during the next decade. GridPP has been significantly involved in this work. The HSF white paper identifies an enormous spectrum of software development work that is needed, which falls in several different domains. Examples are:

- Optimising detector simulations by improved common tools; improved use of vectorisation to make better use of modern CPU architectures; flexible integration of "fast" and "full" simulation methods; and improved modelling of complex detectors. These tools need to be developed by dedicated projects (such as Geant); by cross-experiment working groups; and within individual experiment collaborations.
- Optimising physics generators that demand enormous and increasing amounts of CPU power. These tools are developed mainly by the theory community.
- Data management and organisation must evolve to handle the vast amounts of data expected from the High Luminosity LHC (HL-LHC) and future non-LHC experiments (e.g. the future Neutrino experiments). Here there is a spectrum of issues associated with speed and granularity of access; throughput; consolidation of interfaces; and an evolution towards industrial standards.
- The Infrastructure must evolve to maintain cost effectiveness. This means understanding the costs of public and private "clouds"; the development, as appropriate of hybrid public-private facilities; linkage to other scientific endeavours so that common solutions can be used where possible.
- Security must evolve and develop, both in an operational sense to address the escalating threats, and in terms of trust and policy agreement so that we can integrate with other resources. These issues are driven by the providers such as WLCG.

The HSF is driving a modernisation of HEP software to better exploit new hardware; it is looking for commonalities across experiments to reduce duplication of effort; and it is trying to enable the step-change in performance that will be required to handle the vast increase in data expected from the High Luminosity LHC and future non LHC experiments. It is, therefore, essential that the UK is able to contribute effort to this programme of work across the domains involved.

UKT0 Consortium

UK-HEP/GridPP is integrated with the developing UKT0 consortium, which is the STFC wide science driven initiative to draw all computing interests within STFC together in order to share infrastructure and expertise, and avoid duplication, where possible. UKT0 involves all large PPAN activities, the STFC Scientific Computing Department, as well as the compute needs of the National Facilities (Diamond, ISIS, CLF). This has already led to tangible benefits of sharing of infrastructure with Astronomy and to joint GridPP posts with LSST, SKA and LZ. We also note and support the CERN-SKA formal agreement signed in 2017 and are engaged with CERN discussions at a strategic level through the CERN Scientific Computing Forum.

We understand that UKT0 will be described to the committee separately.

Other related projects

EGI / EOSC-Hub: The EGI Engage project that part-funded the APEL accounting, the Grid Operations Centre Database, and leadership of international security operations and policy development, ended in 2017. This has been succeeded by EOSC-Hub, which is now the EU H2020 flagship-project for e-Infrastructure integration and operation. Running from Jan 2018 until Dec 2020, with €30M of EC funding, EOSC-Hub brings together the EGI Federation, EUDAT CDI and INDIGO-DataCloud to deliver a common catalogue of research data, services and software for research. STFC leads the task on Operational Security, coordinating incident response and harmonising policies/procedures across the EOSC community, and STFC also continues to operate and develop GOCDB and APEL services. Matching funding (15% of the costs) for these activities is provided by GridPP and the EGI Foundation also continues to fund, again at the 15% level (with the balance of 70% funded by EOSC). These services are all critical to WLCG and the continuation of external funding is a notable success. In addition, EGI also agreed to fund CVMFS at 0.1FTE/year. Finally, STFC was also funded 12pm to support a Fusion “Competence Centre” led by CCFE which may link well with its UKT0 activities.

H2020 AENEAS is a project that has started to design a European Science Data Centre for SKA. GridPP is involved directly through co-funded SCD staff and indirectly through informal links between various members of GridPP and the UK SKA team. CERN have now signed a Cooperation Agreement with SKA and attended the autumn AENEAS Collaboration Meeting. In the next year, AENEAS testbed activities may link with STFC’s UKT0 activities.

H2020 HNSciCloud is a project for the joint pre-commercial procurement of cloud services in Europe for use by a range of science projects, including particle physics. Although not a full partner, the UK was able to engage at a reduced level with GridPP funds with this 3 year, €5.3m project. Now starting its final year, the project is taking forward two prototype systems and GridPP hopes to regain the investment in the form of compute cycles.

A note on commercial cloud use

The use of public commercial cloud offering has been explored in some detail by STFC and the science communities involved. The current status is that:

- Commercial cloud is proven to work technically for some types of workflow (e.g. simple simulation) but is not proven to work efficiently for other types of workflow yet (e.g. complex parallel programming and, particularly relevant to the LHC, data I/O intensive analysis).
- The commercial vendor business models are as yet not appropriate to large scale data intensive research, and still yield costs which are much higher than the cost of on-site facilities.

- Work is currently underway in RCUK (and now UKRI) eInfrastructure Group and through BEIS, to develop suitable engagement models and determine where commercial cloud provision can be cost effective for a research NeI. Meetings have recently been held with most of the main Cloud vendors at BEIS HQ to flesh out a way forward, and a clear direction of proof-of-principle exercises with them is envisaged.
- The current working assumption of the RCUK eInfrastructure working group and the vendors themselves, is that the most likely outcome is a hybrid cloud with some NeI resources on premises where that is the most science and cost effective thing to do, and some in the commercial cloud where that is the most cost effective thing to do.

Summary

In summary, the hardware requirements for HEP are a deterministic consequence of the strategic decisions made to fund HEP detectors and collaborations. End-user analysis is not the dominant workflow; rather, the infrastructure is designed to host 24x7 production computing that prepares the data for analysis and performs detailed simulation of the detector responses. One of the strengths of the Grid computing paradigm is that it employs well-understood and widely available hardware that has reached the optimal economic point on the price-performance curve. The benefit is realised by running a software/middleware infrastructure and associated support systems that lie both within the experiment and Grid domains. In addition to continuing investment in the hardware and in this ongoing operation, significant investment is required to refactor/develop software and middleware to fully exploit modern multi-core hardware. This also provides an opportunity to identify and implement common projects that in the longer term will reduce duplication of effort across the experiments.

Appendix: Original questions answered as well as can be with reference to the *main detailed submission* above.

Background Information

1. Please briefly describe the key (scientific) drivers and areas of excellence of UKHEP-Compute/GridPP. *Have there been any significant developments since your PPRP submission or last review or anything you wish to highlight.* **[Maximum 500 words.]**

   The key drivers of GridPP are the scientific drivers of the HEP experiment collaborations who use GridPP resources.

   The "area of excellence of experimental particle physics computing" is in delivering a vast infrastructure in a seamless way across 42 nations and over 170 sites. The UK part of the infrastructure provided by GridPP is excellent in the sense that it is acknowledged as being a very well organised part of the WLCG, responsive, and delivers the pledged resources in a timely way. However, the question is not really applicable if it refers to the leading-edge nature of the computing itself (as It is simply a commodity resource as explained earlier).

2. What is the international relevance of the UKHEP-Compute/GridPP in both the European and global arenas? *You should include the (scientific) relevance of* UKHEP-Compute/GridPP*, and (of) to the science to be carried out, in an international context as well as any international activities in this area and how competitive* GRIDPP UK Compute *is in comparison with others in the field. Please outline how UK investment impacts the international community.*

   The international relevance of UKHEP-Compute/GridPP as a required UK contribution (often as per international MOU's) to international experiments in order to carry out their scientific programmes is described fully in the *main detailed submission* above.

   For the question on *"the science to be carried out, in an international context as well as any international activities in this area"* see the *main detailed submission* and response to Q1.

   The query on the competitiveness of GridPP allows us to address an important misconception. The national compute resources required to deal with the data from an experiment do not compete with others, the UKHEP-Compute/GridPP collaborates as an equal with all international partners. Please see the full description of this issue in the *main detailed submission* above.

3. Please provide comment on the type of facilities that you use:

   The experiments are users of compute resources (of the WLCG). Almost all experimental HEP requirements are satisfied by:

   - Very large HTC commodity compute clusters, almost entirely based upon x86 architecture. Currently WLCG hosts approximately 700,000 cores. These clusters are, by construction, purchased at the lowest cost point available and have no demanding requirements (hence no computing research adventure).
   - Very large distributed resilient storage systems. The WLCG currently hosts ~ 400 PB of disk and 400 PB of tape storage.

   Mostly these resources are made available via a Grid interface but with some resources available using

Cloud technology. In addition, GridPP hosts a relatively small number of GPU and other specialised systems.

a. The type of compute required over the next 5 years (e.g. Extreme Scaling, Data analytic, WLCG infrastructure etc)

WLCG - See response immediately above.

b. The compute facilities currently used and the availability presently accessible (Institutional, DiRAC, Hartree, PRACE, cloud computing providers, etc.).

WLCG - See response immediately above.

c. The current utilisation of the compute facilities currently used.

There is no simple answer to this. The Tier-2 institutions provide additional resources to those funded by STFC and, thus, deliver much more than what is paid for, so the efficiency at the Tier-2s can be said to be much greater than 100%.

The efficiency of an individual CPU, on average, is currently about 90% at the Tier-2s and 95% at the Tier-1 This can be due to I/O wait (jobs waiting for data to be staged in) and/or due to the challenge of using legacy single-thread code on multi-core processors. Investment is needed in software development to address the latter. There is little or no inefficiency due to “insufficient work” because the nature of HEP computing is that resources are used 24x7 by scheduled production jobs, much of which is Monte Carlo simulation production that can use any spare slots at any time of day.

Disk is provided by an annual procurement cycle. When first deployed, there will be significant empty disk and this is then filled up during the year. Thus, it is rather meaningless to talk about average utilisation because this just reflects the frequency of procurement. Aside from the procurement effect, the disk resources are designed to be approximately 85% full on average. This is the “sweet spot” that is neither underused, nor so full that new data can’t be stored.

In a similar way, network bandwidth is monitored and periodically incremented when needed. Network saturation can occur very quickly once spikes in demand start to hit bandwidth limits, but this depends on whether the transfers are automatically re-tried (which leads to rapid saturation). This is extremely complex and means that “average” bandwidth utilisation in not a meaningful measure.

d. The opportunity for provision of compute to be used on different architectures.

In some other countries, there is political imperative to make use of HPC-type resources for HEP computing. It is not an economic decision to use such resources because HEP workloads simply don’t need the sophisticated architecture that drives the costs of such machines. Nevertheless, the LHC experiments have managed to adapt work-loads to make good use of spare-cores on some HPC machines by breaking the workloads into small chunks. Although this adds overhead and introduces inefficiencies for the users it has been used effectively to increase the overall throughput of several HPC machines, particularly those which have a lot of idle cycles. Unfortunately, experience has also shown that bespoke work is required to enable each HPC machine due to their individual architectures and configurations. Simply put, using an HPC machine as a commodity HTC cluster is both difficult and not cost effective, though with the investment of effort it can be used to mop-up space cycles.

4. Please provide comment on the level of support which is currently provided from the compute facilities used? (resource application, job submission, technological and software engineering support) How does this support differ from that provided by your project? Please comment on how this corresponds to the current and future needs.

This is an area where there has been continued confusion and we welcome the opportunity to clarify. Please see the *main detailed submission* under the heading on software support. The summary is:

- UKHEP-Compute/GridPP has staff who run the underlying software infrastructure of the Grid which is used by all experiments.

- UKHEP-Compute/GridPP does NOT support experiment specific software, nor does it provide effort for running experimental production.

The experiments themselves require computing staff support for the experimental specific production software which runs on the e-Infrastructure. This includes data management, storage and cataloguing, calibration, reconstruction, streaming/stripping and Monte Carlo Simulation, and the data analysis software framework.

The misunderstanding between these has led to significant under funding of compute staff within the experiments in recent years. This non-optimal allocation of resources has meant that it has not been possible to properly develop codes to be more efficient on new multi core hardware, or to prepare for the upgrade running of the LHC.

Scale of Current investment

5. Please comment on the number of staff, level of STFC funded computing resources (raw compute, storage, networking, specialised hardware) and the level of externally funded compute resources obtained or used by UKHEP-Compute/GridPP

The GridPP staffing levels are below. The profile falls over the project (2016 – 2020) at the Tier-1 and Tier-2s:

- Tier-1 operations, deployment, support and evolution-development: 17.5 → 14.5 FTE.
- Tier-2 operations, deployment, support and evolution-development: 19.5 → 15.75 FTE.
- Software/Middleware operation, deployment, support and evolution-development: 10.5 FTE.
- Management/Admin/Impact: 2.1 FTE.

Additional non-STFC funded effort is provided by the institutes (~4.5 FTE); by the EOSChub (~3.5 FTE) project; and a small amount by the AENEAS H2020 project (~0.1 FTE).

GridPP provides the following hardware funded by STFC (+ external sources):

- CPU: 440 (+360) KHS06 (provided by 73,000 logical CPUs and equivalent to 36,000 modern CPUs).
- Disk: 34 (+18) PB.
- Tape: 55 (+0) PB.

6. What level of resources of funding for computing do you see as necessary to maintain your current international / national standing over the next 5 years.

The estimated cost of providing the UKHEP-Compute/GridPP compute resources for just the LHC and existing experiments listed in the table in the main submission, is approximately £7M per annum. If we include an uplift for the LHC shortfall, plus rough estimates for DUNE and HyperK, and for computing staff shortfall this rises over time. The approximate profile is, therefore, 7, 8, 8, 9, 9 £M (hardware and staff) over the next five years. Of this approximately £6M per annum has historically come from the GridPP funding line, although we do not know what the future situation will be. Overall, there is a significant shortfall in identified funding for the future.

The experiments themselves also require computing support (research software engineering) staff for the experiment specific computing described earlier (production management, code engineering…). It would take a comprehensive survey of all experiments to identify all of the roles, but we can say that these were very substantially cut in the 2015 Consolidated Grants round. A minimal estimate is about 20 such posts across the UK, of which currently only a small number are funded. The lack of these is having a serious effect upon the ability of the LHC experiments to prepare for upgraded running and means they do not have the manpower to (i) make code efficient and (ii) fully exploit the trend towards many-cores (ii) develop the code needed for the future (e.g. LHC upgrade running). This is a major problem for experiments.

7. What would be the impact if the required resources noted above (Question 5) not available?

The compute resource required is deterministic once the experimental running conditions and science programme are approved. If those resources were not to be available the UK would fail to contribute its fair/expected share of the experiment compute resources. The UK would fail to meet international programme MOU commitments.

If storage were limited, then either data would not be recorded or Monte Carlo simulations would not be performed which would effectively prohibit proper analysis of some or all of the data. If CPU were limited then again data would not be processed and Monte Carlo simulations would not be performed, which would effectively prohibit analysis of some or all of the data.

In summary, the experimental output would drop unless the appropriate balance of computing resources is maintained. This would mean that the STFC investment in the running experiments would not be fully exploited.

8. What is the impact on compute resources in the current flat cash environment?

The last decade of flat cash resources has left the total compute resource across PPAN at about 65% of that required to carry out all of the PPAN programmes at this point. This deficit is increasing by approximately 15% per annum (a detailed assessment of this was recently done by UKT0 and the documentation and spreadsheets are available on request).

9. How much of the required resources noted above (Question 5) were costed into previously submitted SoIs?

Exploitation compute resources have almost never been incorporated into SoIs for experiment construction at PPRP stage.

10. Are there opportunities to get access to required resources from non-STFC sources.

No

Collaborations

11. Have any collaborations or partnerships resulted from the work of UKHEP-Compute/GridPP. These can be multi-discipline and include funding from areas such as Global Challenge Research Fund.
a. Industrial
b. Academic
c. International

Naturally UKHEP-Compute/GridPP is part of very large global collaborations, and details are given earlier.

However, we do not believe that this is what the question is after. We believe this is more aimed at the context where the research of a specific scientist leads to one of the listed collaborations, and/or to further grant applications. E.g. development of an advanced computing architecture in the HPC sector. This is not applicable to UKHEP-Compute/GridPP as the computing resource has no intrinsic research "adventure", as by construction it is composed of the cheapest price point commodity equipment, providing the best value for money to STFC and the research communities.

There is as yet no known way to access GCRF to provide computing resources for particle physics in general.

The experiments themselves may of course become involved in GCRF in other contexts, and in industrial collaboration in respect of detector design and construction.

12. Can you outline the scale of any industrial collaborations or significant contracts with which you are involved which leverage funding from industry for computing?

N/A: Industry does not fund production computing resources for particle physics experiment. There is no R&D element of computing with which industry could be engaged.

At the next level down, small-scale engagement takes place at the local level when industry provides evaluation hardware units in the hopes of selling their more cutting-edge offerings.

Further Funding and Resources

13. Has STFC support for GRIDPP UK Compute led directly to leverage of further resources and/or peer-reviewed, competitively-awarded funding? This can include further funding from STFC or non STFC sources such as non-academic partners, other research councils, or areas of government. Please give an overview of this resources/funding giving a value where possible.

Non-STFC supported resources obtained from the Tier-2 institutes who host GridPP funded hardware averaged £2.4 per annum between 2005 and 2013 (presented in the GridPP5 proposal). This support has continued, though we have not updated our earlier study to quantify the level.

14. Which funding steams are used to enable the use of the required computing? (Consolidated Grant, project funding, industrial, institutional, other Research Councils, etc.)

    The PPGP funds the UK component of exploitation of all HEP experiments through JES FEC awards.

    The University Institutes fund the component not covered by FEC (20% of most costs, and the majority of Academic research time)

15. Has GRIDPP UK Compute undertaken activities to interest, inform, or engage a non-academic audience with the project? Please provide an overview of these activities, clearly indicating the purpose of the activity, the audience(s) initially targeted and eventually reached, a summary of any resources developed, and evidence of key outcomes of the dissemination.

    Over the past 18 years, GridPP has engaged in a wide range of activities to engage non-academic audiences. These are documented in a number of places such as the five main GridPP proposals; the 23 reports to our oversight committee; and many ResearchFish submissions.

    For example, GridPP has been a partner in the CERN@School initiative for several years. This is now part of the Institute for Research in Schools, a charitable trust supporting school students and teachers to develop authentic research in schools. GridPP worked with school students to analyse data from classroom based Timepix silicon pixel detectors, the space-based Langton Ultimate Cosmic ray Intensity Detector (LUCID), the Monopole and Exotics Detector at the LHC (MoEDAL), and the TimPix project that uses Timepix data taken during Tim Peake's Principia mission on the International Space Station.

    [Naturally the Experiments themselves undertake a very extensive range of outreach activities, but these are not centred on compute. We assume these will be assessed by the Particle Physics Programme Review Committee and not this committee]

    Leadership

16. Describe the UK leadership and track record of UKHEP-Compute/GridPP?

    Interpreting the question instead as "how well does UKHEP-Compute/GridPP do its job as part of the global infrastructure", then we are comfortable to assert that we are well respected as a first-rate reliable partner in WLCG.

    GridPP has been in operation since 2001. It has successfully delivered 3 Billion CPU-hours of compute and hosts >100 PetaBytes of storage for the Large Hadron Collider and other particle physics experiments. It supports approximately 10,000 scientists worldwide. As a project, GridPP has delivered more than was planned, on time, with no cost over-runs and indeed, at times, with less money than was initially awarded.

    - The GridPP management structures are detailed at:
        - https://www.gridpp.ac.uk/about/management
        - https://www.gridpp.ac.uk/collaboration/docs/pmbdocs

    - In summary the structures are:
        - A Collaboration Board with overall ownership of the project (one member per participating institute).

- An (executive) Project management Board
- A Technical Team (deals with software deployment, problem resolution, monitoring, service delivery)
- An experiment-liaison group, which acts as an interface with the scientific experiments and projects.
- A resource allocation and monitoring process.
- A close linkage with the CERN Resource Review Board structure to decide resource levels needed for the LHC experiments.

17. To what extent does UKHEP-Compute/GridPP provide STFC with an opportunity to influence policy and funding in the future funding ?

We assume this refers to providing STFC with help in influencing government policy and possible funding streams for computing.

- UK-HEP-Compute/GridPP was a core initiator of the UKT0 association (described elsewhere) which brings together computing interests across STFC, and hence provides a coherent voice in respect of STFC compute requirements that has influenced discussions outside of STFC and has established significant funding opportunities.
- UK-HEP-Compute/GridPP members participate in the RC-UK eInfrastructure working group. This group has for several years been working to make the case to BEIS (BIS as was) for investment in eInfrastructure for UK Science. This resulted in a submission in 2016 and ongoing work for a submission now.
- UK-HEP-Compute/GridPP members have been material to preparing BEIS business cases for STFC specific additional funding starting in 2018/19 <submitted Feb 2018>.
- UK-HEP-Compute/GridPP members have been invited to participate in the UKRI eInfrastructure expert advisory group being set up.
- UK-HEP-Compute/GridPP liaises very closely with JISC-JANET in respect of network requirements for HEP.
- UK-HEP-Compute/GridPP members participate in the JISC Cloud Working Group
- UK-HEP-Compute/GridPP participated in the join meetings between BEIS/RCUK and Cloud Vendors in 2017/18
- UK-HEP-Compute/GridPP members were part of the JISC UK e-Infrastructure Security & Access Management working group that produced five policy documents.

UK-HEP-Compute also has a significant influence on policy in the international scientific computing sector:

- UK-HEP-Compute is a member of the Worldwide LHC Computing Grid (WLCG) and as such participates in its Management Board, Overview Board, Collaboration and Grid Deployment Boards.
- WKCG establishes and sets policies globally in all areas including security and legal compliance with data laws.
- UK-HEP-Compute members represent STFC on the CERN Scientific Computing Forum, Chaired by the CERN Research Director.
- UK-HEP-Compute members are part of the international HEP Software Foundation and were authors of the recent roadmap which sets the direction of computing for HEP into the future with a view to minimizing costs.
- Participation in the EGI (European Grid Infrastructure) and other H2020 projects (HNSciCloud; EOSCHub; AENAES).

Synergies

18. Please outline any synergies UKHEP-Compute/GridPP has with any other computing areas?

UK-HEP-Compute synergises with

- UKT0 (UKHEP-Compute/GridPP is a founding partner).
- STFC Scientific Computing Department (who run the Tier-1).
- DiRAC with whom we have joint projects, and close management liaison.
- The RCUK eInfrastructure group.
- The AENAES SKA H2020 Project.
- EGI (European Data Grid).

Technology

19. Please outline any key technologies and technology development associated with UK-HEP-Compute/GRIDPP and their scientific and technological importance.

There has been very significant and important technological development over the last 15 years in **software infrastructure**. Experimental Particle Physics has developed a very successful and now very efficient distributed computing software infrastructure (historically called the Grid middleware) able, **on a globally distributed physical infrastructure** of the WLCG, to:

- Replicate, transfer, catalogue, manage, discover and access very large (100 PB scale) data sets seamlessly across resources.
- Manage hundreds of thousands of concurrent production compute jobs using comprehensive automatic workload management systems seamlessly across resources.
- Build and distribute large scale reconstruction and analysis software stacks on a nightly basis.
- Run global Virtual Organisation AAAI systems for 3000+ collaborators.
- Run global accounting systems.
- Run a global ticketing system.
- Manage all matters of security across institutions, countries and continents.
- Manage large scale data flows on the wide area network through NRENS and the LHC Optical Private Network.

This infrastructure, which is now stable and in continuous production operation is therefore scientifically important, and provides valuable experience which will benefit other data intensive science activities.

As detailed earlier there is no research “adventure” or technology development associated with UK-HEP-Compute/GRIDPP **physical infrastructure** used for data processing. By construction UK-HEP-Compute/GridPP delivers its obligations by exploiting commodity hardware which has reached the optimum price/performance point.

20. Please describe any economic or societal impacts resulting from such developments. Please provide any information that we cannot currently obtain from Research Fish.

N/A

21. Please outline the progress that has been made with regard to your Pathways to Impact document in relation to computing.

The objectives of GridPP's Pathways to Impact Strategy were:

- Make up to 10% of the Collaboration’s computing resources available to non-LHC users: In 2017 the non-LHC usage of GridPP's computing resources was 7.25%.
- Facilitate access to the tools required to make use of these resources for non-LHC users: The GridUser Toolkit was released and is in use by several non-LHC users, particularly the job submission framework (DIRAC) and the software distribution system (CVMFS). A more secure version of CVMFS is being developed for use by LIGO and medical applications.
- Develop strategic partnerships with non-LHC users to maximise impact: GridPP is developing strategic partnerships with Sno+, Solid, T2K, MoEDAL, Icecube, ILC, Pheno, Comet, DUNE, Biomed, Calice, CEPC, Cern@school, ENMR, Geant4, LSST, LZ, MICE, NA62, OSG, SKAtelescope.

Case Study: As an example, GridPP has a jointly funded post with the Astro-Particle LZ dark-matter collaboration. One of their two worldwide data centres is being run by GridPP staff entirely on GridPP resources (the other centre being the NERSC centre in the US). Their recent international MDC1 data-challenge only succeeded when the US centre ran into problems, because the UK data centre delivered.

Horizon Scanning

22. Please list any computing activities/projects (including those that are interdisciplinary and overlap with other PPAN areas or research councils) that offer potential involvement in new or emerging opportunities.

N/A

Data Management Plan

23. Please detail the long-term data storage, data curation and data sharing facilities available to you and explain how these match the requirements laid out in your data management plan.

The experiments which UK-HEP-Compute/GridPP supports each have DMPs. This is because they create and own valuable scientific data.

UK-HEP-Compute/GridPP does not own or produce any data, but provides the means for the experiments to do so.

The DMPs of each of the experiments are summarised in the DMP submitted to the recent Consolidated Grants Round, and which can be found at:

https://www.gridpp.ac.uk/wp-content/uploads/2018/03/DMP-CG2018.pdf