

GRidPP, IRIS, UKRI, BEIS and all that

GridPP42
Coseners House
April 24-25 2019

Pete Clarke
University of Edinburgh



a Community Initiative



- It is an association of peer (computing) interests:
 - Particle Physics: GridPP: LHC + other large HEP
 - Astro: LOFAR, LSST, EUCLID, SKA, GAIA,
 - Astro-particle: LZ, Advanced-LIGO, CTA
 - Nuclear Physics (on-boarding now)
 - STFC Scientific Computing Dept (SCD)
 - DiRAC HPC Facility
 - Diamond Light Source
 - ISIS Neutron Source
 - Central Laser Facility
 - Hartree HPC Centre
 - CCFE (Culham Fusion Centre)
 - Others...



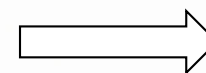
Funding so far



- 2017:
 - £1M Capital
 - 1M capital for hardware
 - 0.5M for digital assets
- 2018:
 - £16M (4M p.a. for 4 years)
 - 11.5M capital for hardware
 - 5.5M for digital assets for National Facilities)

- **Deployed Capacity:**

- Q4 2018 : 4800 cores + 4 PB disk
- Q3 2019: 8000 cores + 10 PB disk + tape
- Q4 2019: 14000 cores + 14 PB disk + tape



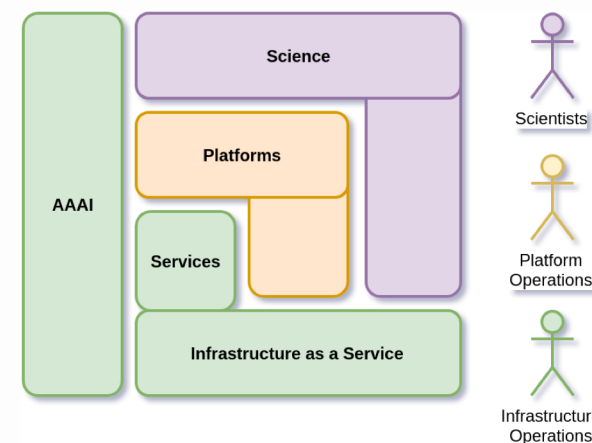
RAL
Manchester
Lancaster
Imperial
QMUL
Glasgow

- **Scientific OpenStack work by StackHPC**

- Development of a Science Optimised OpenStack framework for deployment, federation and support of cloud-native workflows.

- **Other digital assets - ongoing**

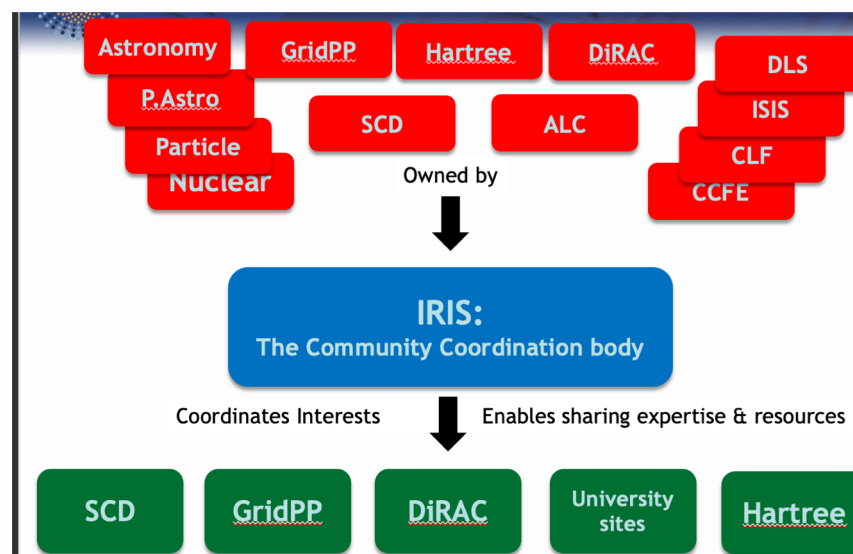
- APEL based accounting for IRIS
- Developing Rucio for multi-community use
- Developing DIRAC for multi community use
- Advancing AAAI across STFC with IAM
- Vcycle development
- Performance monitoring



- **Deployed specialist physical resource:**

- Archive Database test bed
- High performance storage for SCARF
- Fat nodes with > 1 TB memory
- CASU & WFAU support for Astronomy archives

- GridPP is key/founder member of IRIS



- PC is Scientific Director
- Andrew Sansum is Technical Director
- Jon Hays is Chair of Resource Allocation Board (RSAP)
- Daniela Bauer is Technical scrutineer on RSAP



GridPP ↔ IRIS



- **Mutual benefit**
 - Obviously IRIS/STFC benefit from GridPP:
 - Expertise
 - Hosting
 - Direct help to Astronomy and ParticleAstro
 - Particular successes are : LSST, SKA, LZ
 - PPRP sort of relies on GridPP+IRIS
 - GridPP benefits:
 - Cycle scavenging
 - Resource for non-LHC HEP (DUNE, LZ, other new HEP)
 - Does us enormous good seen across STFC
 - GridPP-6 case is greatly strengthened

.....

Put simply, GridPP is doing all that can be reasonably expected of it both for GridPP and for all of PPAN and facilities to promote sharing and avoidance of duplication, and to work in the UKRI group as representatives of STFC to obtain UKRI investment. This all has actual and future monetary value to STFC, but it relies upon STFC policy changes to recognise federal efforts and turn this into a notional benefit to GridPP funding (e.g. to inject some federal funding into GridPP to recognise the wider contribution of its staff).

- We have done a lot of case making
 - Case presented to UKRI for long term investment of ~ £50M p.a. for STFC
 - Lots of “Roadmaps” written
 - People: College of RSEs case



STFC Computing Requirements 2019-2026

Contact editors: Pete Clarke (peter.clarke@ed.ac.uk),
Andrew Sansum (Andrew.sansum@stfc.ac.uk),
Jeremy Yates(j.a.yates@ucl.ac.uk),
Tom Griffin (tom.griffin@stfc.ac.uk)

February 7, 2019

This report shows the aggregated requirements across STFC for CPU, Disk and Tape. This covers all sciences in PPAN and the National Facilities.

Table 2: CPU-cores - net request per activity

CPU-cores	2019	2020	2021	2022	2023	2024	2025	2026
DLS	2999	2999	5299	5299	6399	7399	9799	12799
PANOSC	0	0	0	0	0	0	0	0
ISIS	175	1375	1375	3575	4775	4975	5175	5375
CLF	68	500	700	1100	1100	1100	1100	1100
CCFE	4199	8195	8495	8995	12142	12892	13642	14392
Euclid	3650	3650	6800	6400	6400	7200	7200	7200
Gaia	191	500	690	3400	6800	6800	6800	6800
LSST:UK	1312	1324	1324	189	1876	1699	1699	1699
MERLIN	1039	127	2047	4095	8191	0	0	0
SKA	2559	5119	10239	10239	10239	10239	10239	10239
Lux Zepplin	300	689	1129	1570	2010	2449	2889	3100
CTA	570	570	570	1150	1150	1710	1710	1710
Advanced LIGO	270	270	270	270	270	17698	17698	17698
DUNE	1000	1000	2000	2000	2000	2000	3000	5000
GRIDPP	0	15250	33250	55900	85300	97000	125000	165000
Other at 10.0%	1833	4156	7418	10418	14865	17316	20595	25211
Total IRIS requirement	20165	45724	81606	114600	163517	190477	226546	277323

Series of Tables

- CPU
- GPU
- DISK
- TAPE

All STFC senior management know of this document

Sent to CEO
Sent to Prog.Dir.

Table 3: CPU-cores - Net shortfall summary with estimated cost in kPounds

CPU-cores	2019	2020	2021	2022	2023	2024	2025	2026
Total IRIS Requirement	20165	45724	81606	114600	163517	190477	226546	277323
IRIS existing	7000	7000	7000	7000	5000	3000	0	0
Net IRIS Requirement	13165	38724	74606	107600	158517	187477	226546	277323
Inc.shortfall (10.0% obs.p.a.)	13165	26875	39754	40454	61677	44811	57816	73431
Unit cost (Pounds)	224.4	207.4	192.1	177.65	164.9	153.0	141.1	130.9
Cost (kPounds)	2954	5573	7636	7186	10170	6856	8157	9612

Summary table and plots

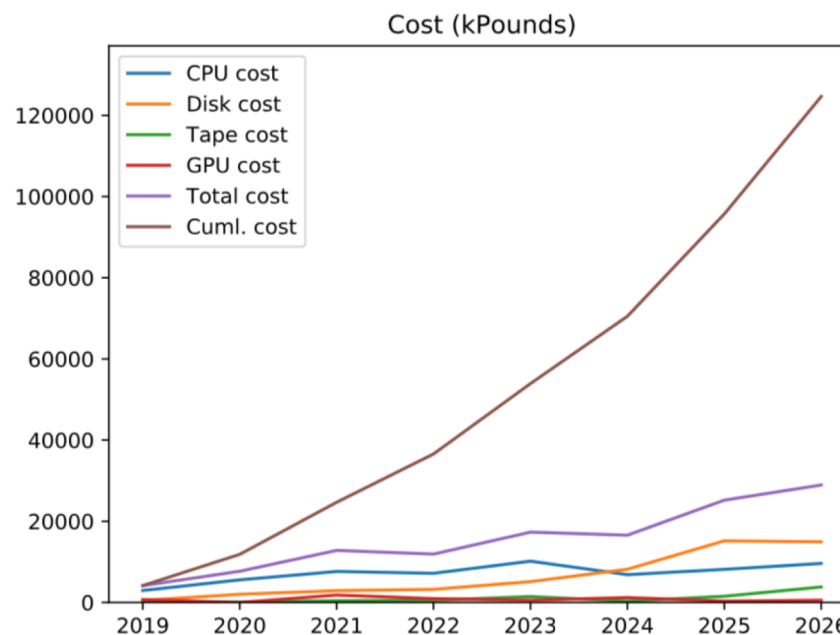
STFC has
HTC+Storage
computing funding
shortfall of A £12M
p.a.

These figures have
been put into UKRI

Will go to BEIS

Table 1: Overall cost summary table in kPounds

Cost (kPounds)	2019	2020	2021	2022	2023	2024	2025	2026
CPU cost	2954	5573	7636	7186	10170	6856	8157	9612
Disk cost	503	2027	2916	3233	5135	8184	15186	14942
Tape cost	0	86	438	552	1457	330	1529	3824
GPU cost	706	4	1836	944	568	1212	332	558
Total cost	4163	7690	12826	11915	17330	16582	25204	28936
Cuml. cost	4163	11853	24679	36594	53924	70506	95710	124646



People for software engineering RSEs

Priority projects – outline case

Project Name

College of Research Software and Infrastructure Engineers

Proponents

Name	Where	Relevant Position
Peter Clarke*	Edinburgh	Director IRIS, Dep PL GridPP
Jeremy Yates	UCL	Dep Dir. IRIS, Innovation Dir DiRAC, ex-Director DiRAC
Dave Britton	Glasgow	Project Leader GridPP
Dave Colling	Imperial	Chair CAP
Andrew Sansum	STFC-RAL	Director IRIS, ex-Head SCD
Mark Wilkinson	Leicester	Director DiRAC
Clare Jenner	UCL	Project Scientist DiRAC
Anna Scaife	Manchester	Lead for SKA RDC
Andrew McNab	Manchester	Dep. Computing Coordinator LHCb
Debora Sijacki	Cambridge	Chair DiRAC Project Board
Mark Heron	Diamond	Head of Computing for Diamond
Robin Pinning	Hartree	Technical Director, Hartree
Stuart Martin-haugh	STFC-RAL	PPD
Catherine Jones	STFC-RAL	Software Engineering Group Leader, ALC Dep. Division Head
Roger Jones	Lancaster	Head of Computing and Head of School Lancaster
Tom Griffin	STFC-RAL	Head of SCD

Scale of investment (Max 500 words)

Please outline the scale, and time profile, of investment required for the construction of the Project. Please provide estimates of future operating costs (if relevant). Identify any expected contributions from sources other than STFC.

Approximate costs over 6 years are shown in the table below, not including inflation.

The baseline is shown in bold. The costs are approximately

£4M per annum

£20M over 5 years

£23M over 6 years

Two optional additions are shown:

A: Adding a further 15 RSEs in year 4

B: Training of PhD students in software engineering in conjunction with CDTs (to be formulated and agreed with CDTs). We estimate this to be approx. 1.25M per annum for 50 students at £25k.

People (RSEs) needed

This is critical to engineering and adapting codebases

Scale £5M p.a. minimum

Optimal ~ £10M p.a

Roadmaps

Networking

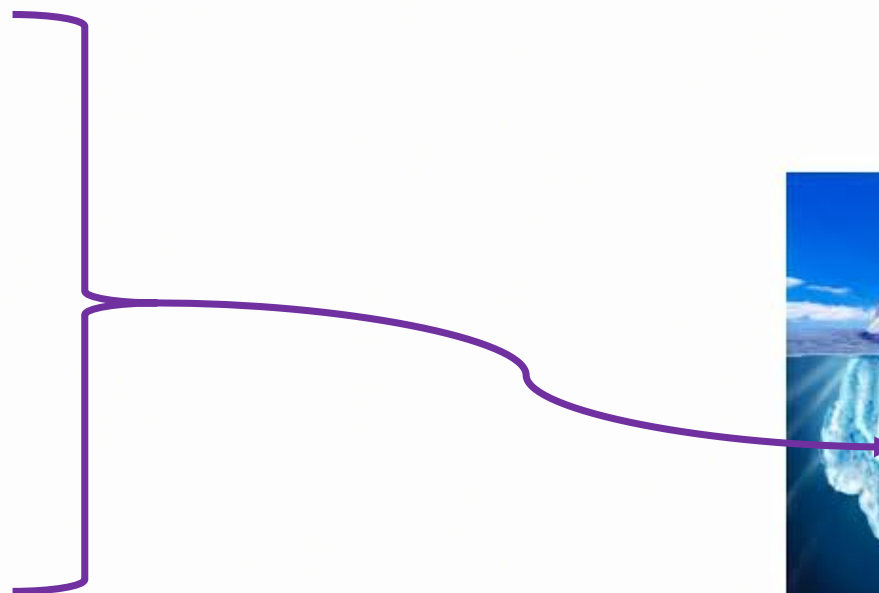
Supercomputing

Data and HTC

Cloud Computing

AAAI

Software and People





Summary



- GridPP “gives-and-gets” from IRIS
- GridPP roles are doing us good in intangible benefits
 - Supporting IRIS
 - Representing STFC at UKRI level
 - This all shows up well in the GridPP6 case
- At high levels STFC has bought into need for computing for Science
 - Across STFC
 - Across UKRI
- GridPP is lobbying everywhere and at every level for investment
 - But STFC funding problems/structure/silos make it difficult to get tangible benefit so far in terms of staff support