IRIS

eInfrastructure for Research and Innovation for STFC

Pete Clarke
University of Edinburgh
On behalf of IRIS Community

Community Meeting
Virtual
17-18 Nov 2020

Since our last community meeting

- ❑ Computing resource deployed
  - • Capacity CPU
  - • High Memory
  - • GPUs
  - • WFAU & CASU support

- ❑ Digital Assets
  - • ALC Programme
  - • Scientific OpenStack
  - • Infrastructure: Iris-IAM, Accounting, GOCDB
  - • Common user software: Rucio, DIRAC, Dynafed
  - • Jupyter hub

- ❑ Science done !

- ❑ Website

- ❑ Security Policy

- ❑ STFC traction & future funding progress

Capacity ~21,000 cores + 14PB Disk

➔ **See talk by Philip Jackson**

See talks
➔ **Ian Collier**
➔ **"STFC Facilities and ALC "**

➔ **See talks on day-2**
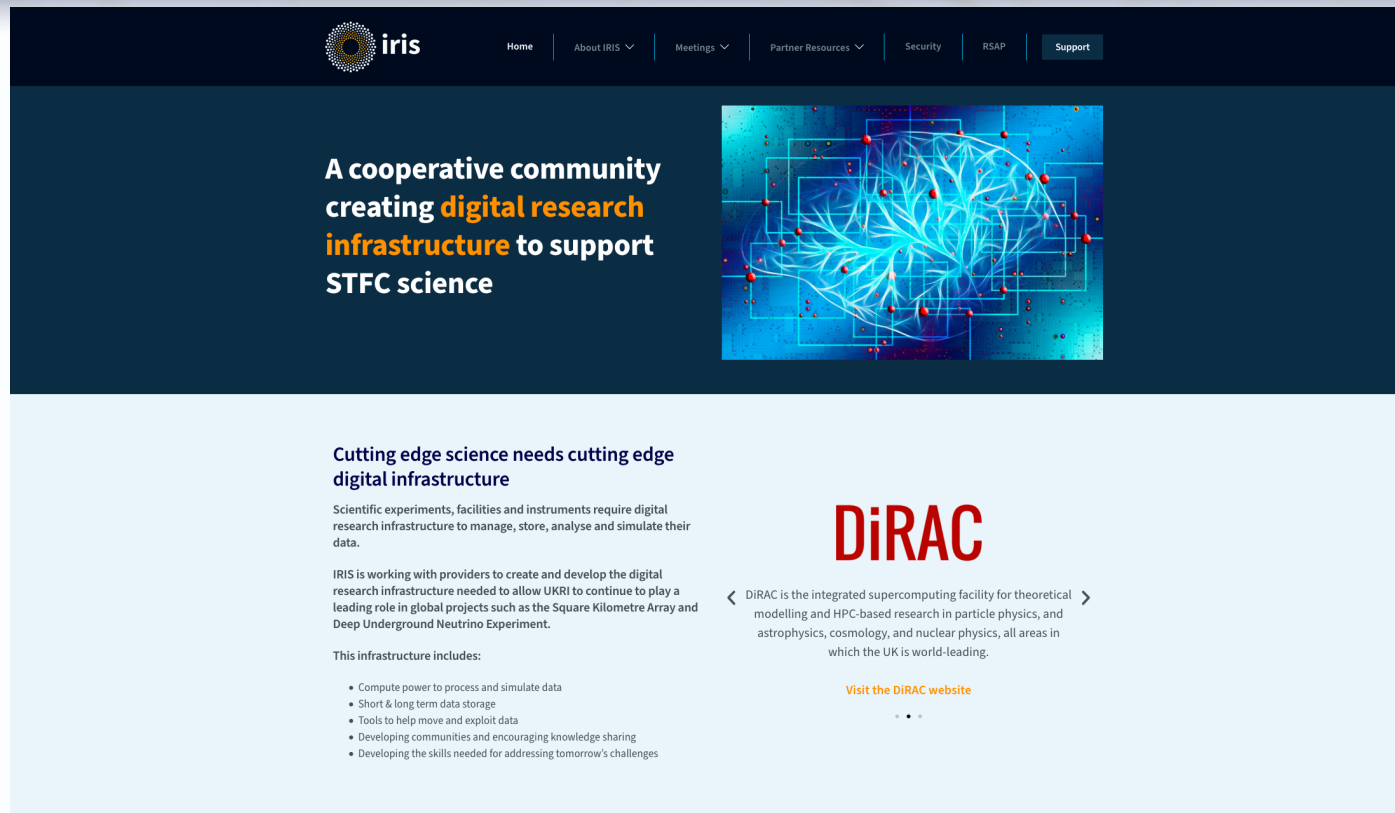
The website is now a useful resource & an excellent front face

Thanks to: Philip Jackson, Andrew McNab, Mathew Banks and others at UKAEA

❑ Iris approved its first high level security policy document

- • It establishes a framework and a principle
- • As I have oft said: as soon as ones computing horizon encompasses truly federated resources across multiple sites across multiple countries this is MANDATORY

This DRAFT IRIS Infrastructure Policy is presented as an interim measure until a full IRIS policy is agreed. It is based on the AARC project Policy Development Kit (https://aarc-project.eu/policies/policy-development-kit/) Top Level Infrastructure Policy, downloaded June 2019)

IRIS Infrastructure Security Policy

This policy, the IRIS Infrastructure Security Policy, is effective from <insert date>.

INTRODUCTION

To fulfil its mission, it is necessary for the IRIS Infrastructure (https://www.iris.ac.uk) to be protected from damage, disruption and unauthorised use. This document presents the policy regulating those activities of IRIS Participants related to the security of the IRIS Infrastructure.

DEFINITIONS

5

STFC Traction & Future Funding Landscape

- **2017:**
  - **£1.5M Capital from Programmes**
    1M capital for hardware
    0.5M for digital assets

- **2018:**
  - **£16M (4M p.a. for 4 years) from BEIS**
    11M capital for hardware
    5M for digital assets for ALC (for Diamond, ISIS, CLF)

- **2017:**

  - **£1.5M Capital from Programmes**

    1M capital for hardware

    0.5M for digital assets

- **2018:**

  - **£16M (4M p.a. for 4 years) from BEIS**

    11M capital for hardware

    5M for digital assets for ALC (for Diamond, ISIS, CLF)

STFC management has believed in Iris

- **2017:**
  - **£1.5M Capital from Programmes**
    - **1M capital for hardware**
    - **0.5M for digital assets**

- **2018:**
  - **£16M (4M p.a. for 4 years) from BEIS**
    - **11M capital for hardware**
    - **5M for digital assets for ALC (for Diamond, ISIS, CLF)**

- **2020:**
  - **£3M extra Capital from World Class Labs Fund**
  - **£2M extra Capital for LSST (Vera Rubin) Data Release Processing**

❑ **Representational work in last year**

    ❑   **Presentation to Executive Board**

    ❑   **Presentation to CAP**

    ❑   **Presentation to Science Board with CAP**

    ❑   **Presentation to STFC council led by CAP**

    ❑   **Interaction with SeIGO**
-   **Submission of "Governance Discussion Document"**

    ❑   **Large amount of work on "pro-formas" for UKRI & CST**
-   **Iris-II**
-   **ALC-II**
-   **DiRAC-3**

Name of Project	**eInfrastructure for Research and Innovation in STFC (IRIS) Phase-2**
Type of infrastructure	Select one from:_[ ] establishing new capability, [ ] significant change to existing capability e.g. upgrade, [ ] combination of new and changes to existing, [ ] decommissioning, [ ] unknown
Lead Council/UKRI team	STFC
UKRI contact	Justin O'Byrne; justin.obyrne@stfc.ukri.org
Stage of development	Select one from [ ] Stage 1, [ ] Stage 2, [ ] Stage 3a, [ ] Stage 3b

Short description of the Project for use in summary briefings to IAC, ExCo

IRIS: A co-operative consortium of STFC science users and compute providers, to deliver High Throughput Computing (CPU and GPU), data management and storage capability to achieve the science goals of the National Facilities (including ISIS, Diamond, CLF), of PPAN projects and instruments (including SKA, LSST, LIGO,LHC, DUNE), and CCFE

[50 words]

Detailed description of the Project for the IAC

This project is to deliver of the next phase of the very successful **IRIS co-operative consortium** for Digital Research Infrastructure (DRI) for STFC Science. IRIS is described as www.iris.ac.uk.

ALC

Name of Project	Ada Lovelace Centre
Type of infrastructure	Select one from: [ ] establishing new capability, [ ] significant change to existing capability e.g. upgrade, [X] combination of new and changes to existing, [ ] decommissioning, [ ] unknown
Lead Council/UKRI team	STFC
UKRI contact	Tom Griffin tom.griffin@stfc.ac.uk 07539 227945
Stage of development	Select one from [ ] Stage 1, [ ] Stage 2, [ ] Stage 3a, [X] Stage 3b

Short description of the Project for use in summary briefings to IAC, ExCo

The Ada Lovelace Centre (ALC) will provide software, data services and skills to exploit data from large scale national facilities including Diamond, ISIS and CLF/EPAC. The ALC will enable researchers to maximize the scientific output of existing investments in the facilities, reduce time to impact, and reduce barriers to industry

[50/50 words]

Detailed description of the Project for the IAC

The UK's national facilities, primarily located at the Rutherford Appleton Laboratory produce data for a wide range of UKRI supported science and industry. There is an increasing gap between our ability to experimentally take measurements (on large-scale infrastructures through to medium scale facilities) and then to be able to

DiRAC

Name of Project	High Performance Computing – Unusual Architectures via DiRAC
Type of infrastructure	Select one from: [ ] establishing new capability, [ ] significant change to existing capability e.g. upgrade, [X] combination of new and changes to existing, [ ] decommissioning, [ ] unknown
Lead Council/UKRI team	STFC
UKRI contact	Justin O'Byrne, STFC; justin.obyrne@stfc.ukri.org
Stage of development	Select one from [ ] Stage 1, [ ] Stage 2, [ ] Stage 3a, [ ] Stage 3b

Short description of the Project for use in summary briefings to IAC, ExCo

Deployment of the DiRAC-3 (Distributed Research utilizing Advanced Computing) HPC facility. DiRAC-3 will benefit the UK by enabling world-leading scientific discoveries, training highly-skilled researchers and driving hardware and software innovation projects. Additional targeted investment would allow DiRAC-3 services to be expanded to support similar workflows across the wider UKRI community.

[50/50 words]

Detailed description of the Project for the IAC

Computational and data intensive approaches are now an essential component of the creation, sharing and

STFC Traction & Future Funding Landscape: Requests made:

❑ **Iris-II Funding Request FY2021 – 2025 (4 years costs) ~ £125M**

- Capacity Hardware ~ £40M

- Digital Assets (RSE staff) for PPAN   ~ £20M

- New Machine Room Capacity across Iris ~ £40M

- Operations costs ~ £25M

❑ **Importance of the RSAP Process**

- Its just good internal practice to keep our house in order

- It results in the annual requirements forecast ➔ taken VERY SERIOUSLY by STFC

  ➔ This in turn justified the additional capital this year

  ➔ Fed directly into UKRI and the CSR request

IRIS Computing Capacity Requirements 2020-2024

Presented on behalf of the IRIS Consortium.
Contact editor: Pete Clarke (peter.clarke@ed.ac.uk)

August 27, 2020

1 Introduction

This report shows the computing requirements for STFC for CPU, Disk, GPU and Tape. These are aggregated across Programmes, National Facilities and SCD.

The figures are compiled from the IRIS RSAP[1] process carried out in Q4-2019 and Q1-2020. Each science Activity submitted a **Resource Request Document**. (RRD) This contained firm requests for 2020 and forward look estimations for 2021-2024. Each RRD was scrutinised by the RSAP to produce final allocations. The 2020 RSAP full report is available in a separate document.

In addition to this

- LSST provided a new estimate for their proposed pipeline processing contribution in lieu of subscription;

- SKA provided a revised estimate to include the SRC;

- GridPP provided an estimate of the shortfall not funded in the GridPP6 award. this is actual during the GridPP6 funding period 2020-2023 and extrapolated for 2024;

Table 1: Overall cost summary table in kPounds

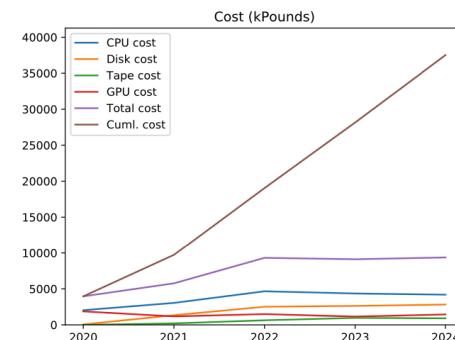| Cost (kPounds)	2020	2021	2022	2023	2024
CPU cost	2036	3048	4664	4358	4192
Disk cost	59	1338	2519	2635	2820
Tape cost	0	206	635	968	906
GPU cost	1867	1176	1495	1152	1447
Total cost	3962	5768	9313	9113	9365
Cuml. cost	3962	9730	19043	28156	37521

Figure 1: All lines show the annual incremental cost summary, apart from the line labelled as Cuml. cost which shows the cumulative total.

❑ **LSST == Vera Rubin Observatory**

❑ **Programmes Directorate ➔ Astronomy ➔ identified explicit VRO funds**

- To implement "Data Release Processing" for VRO in lieu of subscription model

- This is a "big deal" UK commitment (if accepted) ~ same as GridPP to LHC

- Breaks new ground:
  - Funds deployed via Iris to augment common infrastructure
  - But must be hypothecated to VRO as first call when required
  - VRO must have substantial say in where/how funds deployed
  - Requires an MOU
  - Opens up explicit operations costs (power and support staff)

- Similar may happen for SKA-SRC and others

In the next year Iris needs to formalise further how to handle such VO specific funds that STFC wishes to deploy via Iris - to embody:
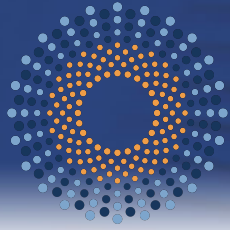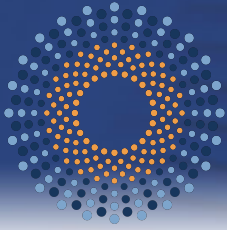➔ Commitment to the VO
➔ Sharing the resources for common good
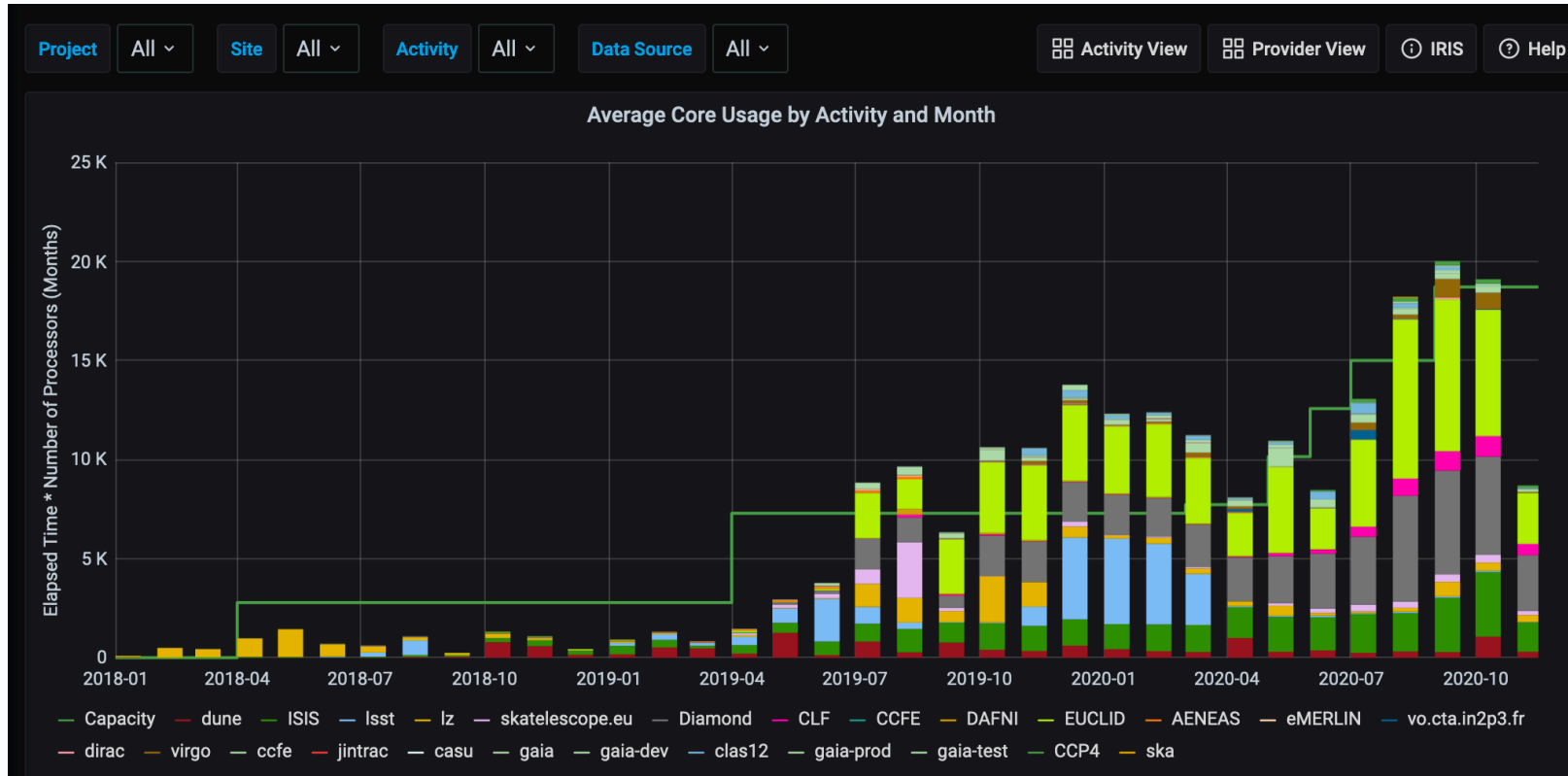
This says it all (well – most of it)
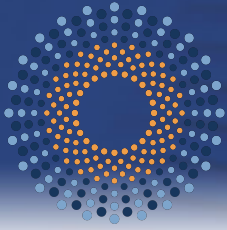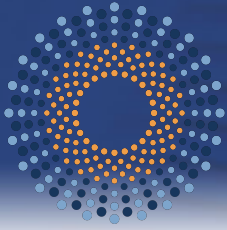
from LastYearsSummary import *

- Exceptional community cooperation across STFC science computing has been built from bottom up.

- Welcome injection of ~ £2.5M p.a. hardware capital for 5 years (FY 2017-2021) for *non-funded* communities
  - Deployed and managed via DiRAC, SCD, GridPP
  - Symbiosis
  - Shared and well used

- eInfrastructure that is *scientist-managed* gets the job done in the most agile way

- ALC has been very successful due to IRIS funds (~£1.3M p.a.)

- Lots of good science has been done that would not have been possible without this IRIS funding from BEIS

- IRIS has robust structures for due-diligence – in particular RSAP. This is far better than any other sector

- eInfrastructure is
  - Hardware infrastructure → capital
  - Software infrastructure → people

- No capital funding past FY 2021 for hardware capacity
  - Large capital shortfall
  - This is all known to UKRI
  - Letter was sent to EB, SB outlining situation
  - Discussed positively at EB

- No people support for large areas of STFC eInfrastructure

  - No staff for software-infrastructure and production-operations for Astro, ParticleAstro, Nuclear

  - No RSE effort for science application development
  - Right hand calls for software advances but left doesn't provide the means

  - No DevOps resource for STFC wide core work
  - Essential pre-requisite to UKRI aspirations.

- No further IRIS support risks losing what has been built. I.e. negating the message that "working together works"

- UKRI
  - New Director of eInfrastructure
  - eInfrastructure progress timescale ?
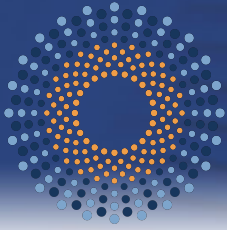 - Needs stop gap money on account

17

- Exceptional community cooperation across STFC science computing has been built from bottom up.

- Welcome injection of ~ £2.5M p.a. hardware capital for 5 years (FY 2017-2021) for *non-funded* communities
  - Deployed and managed via DiRAC, SCD, GridPP
  - Symbiosis
  - Shared and well used

- eInfrastructure that is *scientist-managed* gets the job done in the most agile way

- ALC has been very successful due to IRIS funds (~£1.3M p.a.)

- Lots of good science has been done that would not have been possible without this IRIS funding from BEIS

- IRIS has robust structures for due-diligence – in particular RSAP. This is far better than any other sector

- eInfrastructure is
  - Hardware infrastructure → capital
  - Software infrastructure → people

- No capital funding past FY 2021 for hardware capacity
  - Large capital shortfall
  - This is all known to UKRI
  - Letter was sent to EB, SB outlining situation
  - Discussed positively at EB

- No people support for large areas of STFC eInfrastructure

  - No staff for software-infrastructure and production-operations for Astro, ParticleAstro, Nuclear

  - No RSE effort for science application development
  - Right hand calls for software advances but left doesn't provide the means

  - No DevOps resource for STFC wide core work
  - Essential pre-requisite to UKRI aspirations.

- No further IRIS support risks losing what has been built. I.e. negating the message that "working together works"

- UKRI
  - New Director of eInfrastructure
  - eInfrastructure progress timescale ?
  - Needs stop gap money on account

- Exceptional community cooperation across STFC science computing has been built from bottom up.

- Welcome injection of ~ £2.5M p.a. hardware capital for 5 years (FY 2017-2021) for *non-funded* communities
  - Deployed and managed via DiRAC, SCD, GridPP
  - Symbiosis
  - Shared and well used

- eInfrastructure that is *scientist-managed* gets the job done in the most agile way

- ALC has been very successful due to IRIS funds (~£1.3M p.a.)

- Lots of good science has been done that would not have been possible without this IRIS funding from BEIS

- IRIS has robust structures for due-diligence – in particular RSAP. This is far better than any other sector

- eInfrastructure is
  - Hardware infrastructure → capital
  - Software infrastructure → people

- No capital funding past FY 2021 for hardware capacity
  - Large capital shortfall
  - This is all known to UKRI
  - Letter was sent to EB, SB outlining situation
  - Discussed positively at EB

- No people support for large areas of STFC eInfrastructure

  - No staff for software-infrastructure and production-operations for Astro, ParticleAstro, Nuclear

Still True

But ask put in

- No further IRIS support risks losing what has been built. I.e. negating the message that "working together works"

- UKRI
  - New Director of eInfrastructure
  - eInfrastructure progress timescale ?
  - Needs stop gap money on account

19

- Exceptional community cooperation across STFC science computing has been built from bottom up.

- Welcome injection of ~ £2.5M p.a. hardware capital for 5 years (FY 2017-2021) for *non-funded* communities
  - Deployed and managed via DiRAC, SCD, GridPP
  - Symbiosis
  - Shared and well used

- eInfrastructure that is *scientist-managed* gets the job done in the most agile way

- ALC has been very successful due to IRIS funds (~£1.3M p.a.)

- Lots of good science has been done that would not have been possible without this IRIS funding from BEIS

- IRIS has robust structures for due-diligence – in particular RSAP. This is far better than any other sector

- eInfrastructure is
  - Hardware infrastructure → capital
  - Software infrastructure → people

- No capital funding past FY 2021 for hardware capacity
  - Large capital shortfall
  - This is all known to UKRI
  - Letter was sent to EB, SB outlining situation
  - Discussed positively at EB

- No people support for large areas of STFC eInfrastructure

  - No staff for software-infrastructure and production-operations for Astro, ParticleAstro, Nuclear
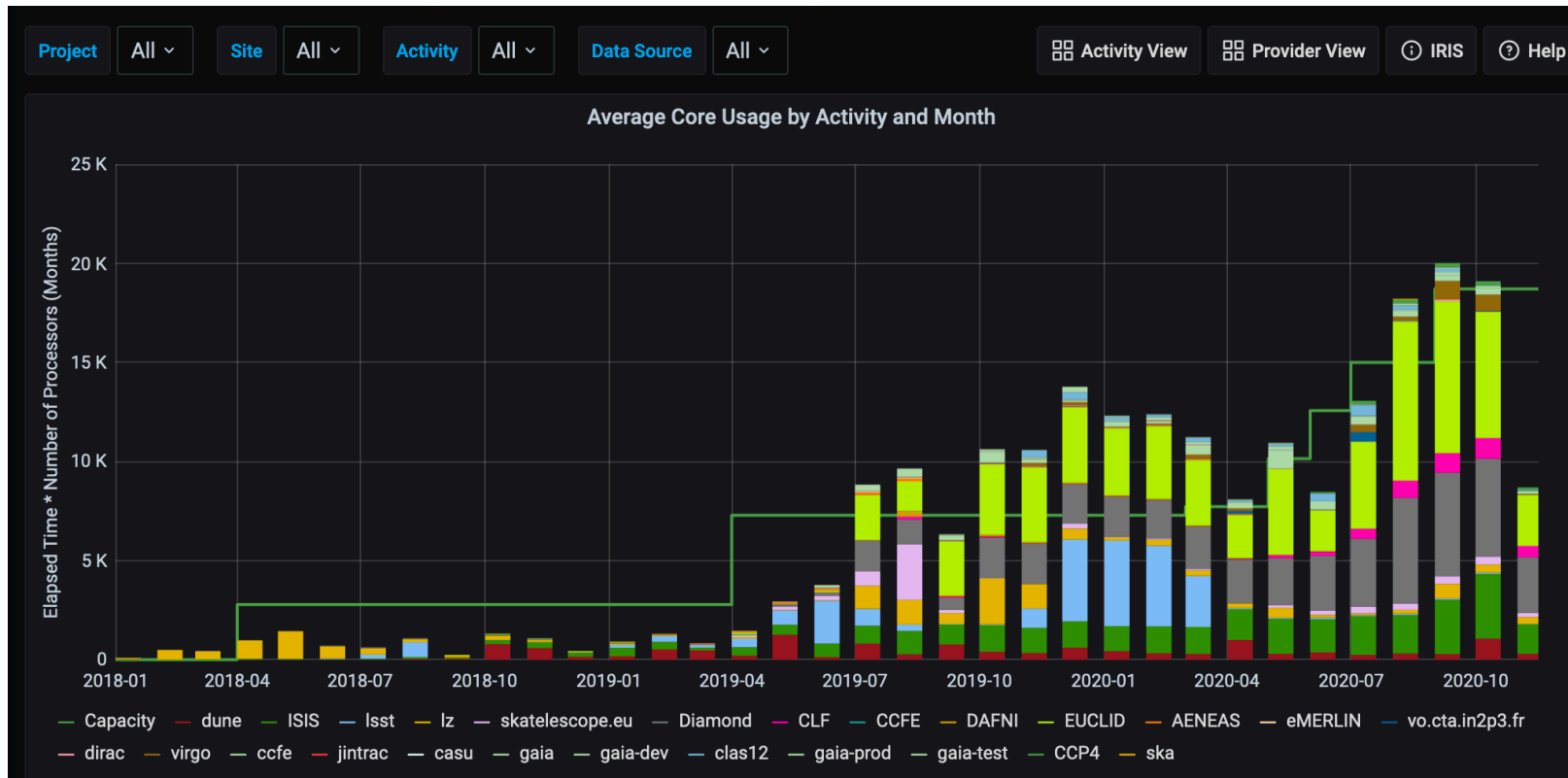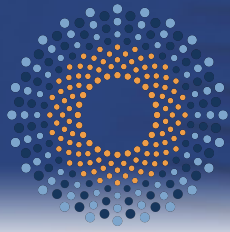
Still True
But ask put in

- No further IRIS support risks losing what has been built. I.e. negating the message that "working together works"

- UKRI
  -
  -
  - Needs stop gap money on account

Underway

This says it all (well – most of it)

BACKUP

❑ Iris is sort of "closed" (to its Members) but "open" (to new Members)

- Iris does not portray itself as a Facility with on open call for anyone to apply.
- I have oft used the phrase "Iris is the wholesaler, SCD, DiRAC, GridPP, ... are the retailers"
- The website says:

Who is eligible to join IRIS?

IRIS is fundamentally a coordination and resource providing body for STFC digital research infrastructure. As such any science area that needs computing, and is supported by STFC, is welcome to participate in IRIS and use its resources. The IRIS philosophy is to be open and inclusive and so seeks to interpret this widely to include closely related activities and activities with which STFC has a de-facto relation.

It is important to understand that IRIS is a self-service infrastructure providing organisation. IRIS can provide resource allocations, tools and building blocks but you must be able to architect and manage your own computing requirements. As a community project we do expect partners to involve themselves in working groups and contribute to the growth of the community.

Potential scientific partners

If you are interested in joining IRIS please reach out to your 'scientifically closest' member of the delivery board(DB) to discuss if IRIS may be able to support you.

Potential eInfrastructure providing partners

Any new potential provider should contact the Technical Director Andrew Sansum andrew.sansum@stfc.ac.uk and Scientific Director Pete Clarke peter.clarke@ed.ac.uk

❑ In practice new(ish) Members/users have come through common people

- Gaia
- JBCA
- SKAO
- CCP4,DAKOTA,RAYSECT
- SOLID
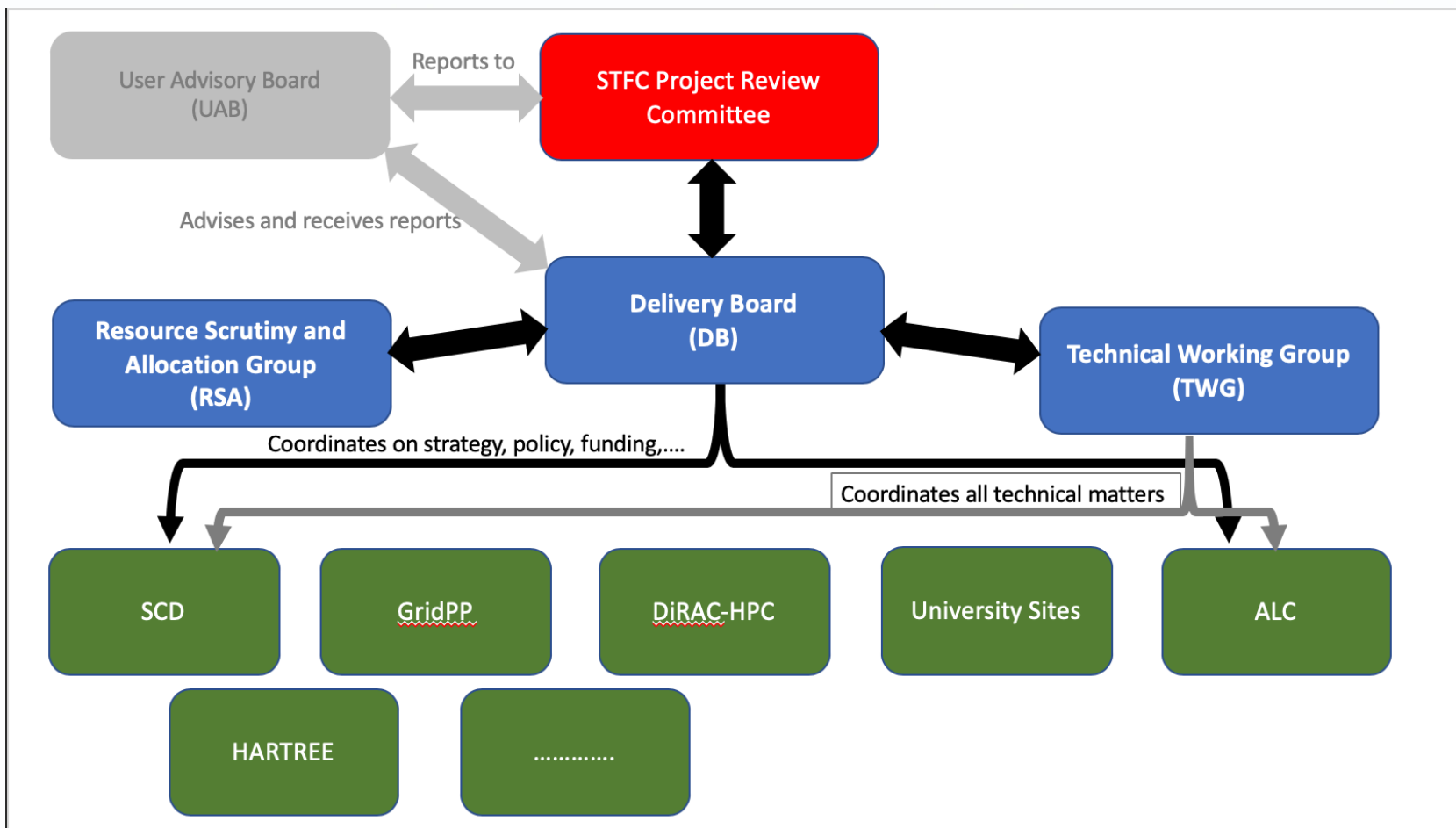
- **Delivery Board is the overall coordination body**

- **1 or 2 members per "thing - but definition is pragmatic and adaptable**

- **Is the final authority for approving all spend, all documents**

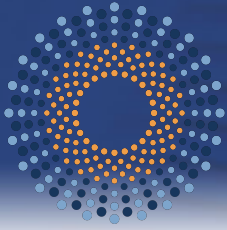- **Meets Thursdays @ 15.00 approx. 3 weekly**
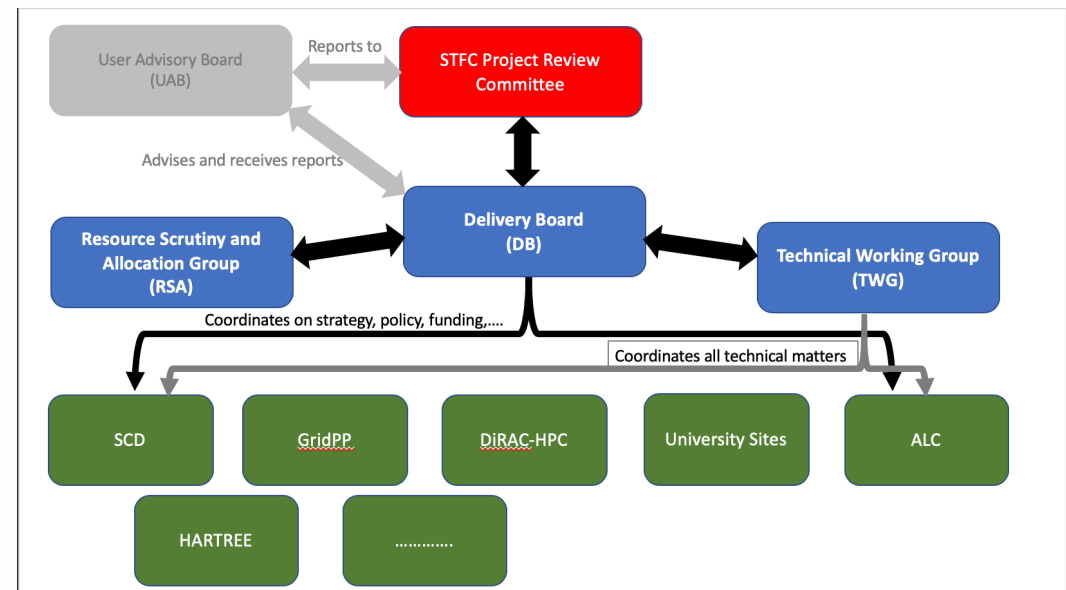
- **Technical forum**

- **Informal mutual help body**

- **Open, anyone can raise anything**

- **Has operations function - regular reporting from user activities in respect of uptake of resources**

- **Meets Tuesdays @ 15.00**
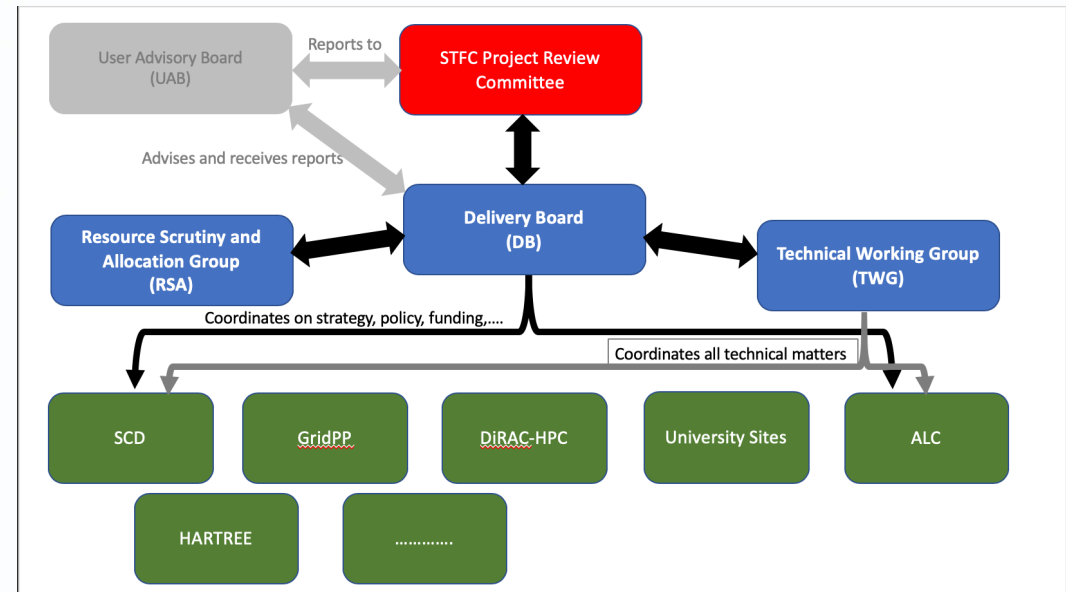
The TWG has been very successful
It has got communities together every 1 or 2 weeks

- **RSAP Modelled on**
  - CERN RRB/CRSG (Computing Scrutiny)
  - DiRAC RAC (Resource Allocation)

- **Partners provide annual resource requirements doc**
  - Justify resources
  - Not science

- **RSAP makes recommendations for allocations**

- **DB process to:**
  - Receive RSAP recommendations
  - Decide "shape" of hardware spend
  - Allocate funds to Providers
  - Make allocations to user Activities
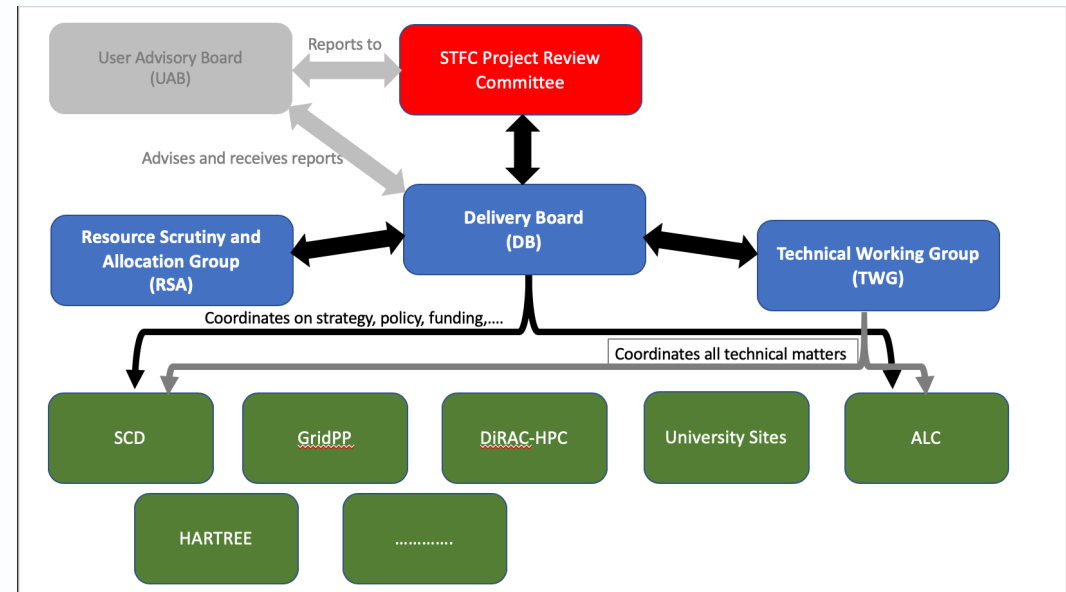
The RSAP has proven crucial.
It is independent
It provides due scrutiny
It means STFC is truly abreast of its needs.
It is far better than any other sector