

RSAP Update

Daniela Bauer
RSAP Chair

What is the “Resource Scrutiny and Allocation Process” ?

- Gathers the resources needed by IRIS supported activities.
- Allows IRIS to plan, purchase and commission the appropriate hardware.
- The requests are assessed by a panel to insure they are:
 - Supporting the stated science goals of the activities.
 - Credible: Do the requested resources follow from the information provided in the request ?
 - Realistic: Is the activity in a position to use the resources in a timely manner ?
- IRIS aims for continuity and planning security for their activities, and assigned resources **that are being used** will not be withdrawn.
- IRIS does not:
 - Assess the Science Programme of an activity.
 - With a few exceptions: Buy you your own kit.

Input from the activities

The process is outlined in <https://www.iris.ac.uk/rsap/> including detailed guidance.

Science Partner Form contains information to:

- Confirm the science programme lies within the IRIS remit and to determine the peer-review status of the science activity.
- Provide contextual information for any resource requests. The Science Partner forms are made available to the reviewers.
- Allow IRIS to produce reports for STFC/UKRI.

Input from the activities to the panel

Resource requests cover:

- Currently (Oct 2023) available resources:
 - Allows reviewers to get a feeling on how much of a step change the new request is.
 - Serves as a checkpoint for IRIS: The bigger the change, the more advance warning IRIS needs.
- Usage of IRIS (or related if it makes sense in context) resources in the past year (Oct 2022-Oct 2023):
 - It is understood that this will often be different to the resources available in October 2023, so activities need to describe the actual resources on the ground and how they were used. This is used to:
 - Convince the reviewers that the activity is able to make use of the provided resources
 - Inspire confidence in the activities' forecasts
- Forecast for 2024/25 and how the activity arrived at that forecast (aka “show your workings” or “the computing model”):
 - At the end of the document the reviewer should arrive at the same numbers you do.

Pitfalls

- What is a core:
 - IRIS is a hardware provider and CPUs are sold in cores, hence we need projections in something close to “cores”.
 - Unlike e.g. HEP which runs fairly homogeneous workloads and therefore can make use of standards like HEPsScore, there is no meaningful standard which we can use to characterize an “IRIS CPU” for all activities.
 - We encourage activities to make their projections on **based on the hardware they are currently using and use that as their de-facto standard core**. IRIS will work with the providers to translate this into purchase orders.
 - More is better: If there is additional information in a format other than cores (e.g. CPU is allocated in hours etc), it can be helpful to include it in the request.

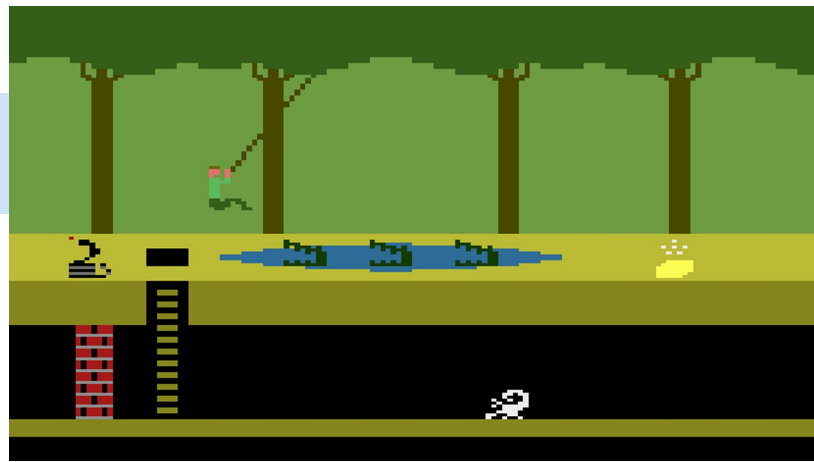
Handle with care

- GPUs:
 - Usage is lower than predicted in the early days of IRIS.
 - Activities are clear on which GPU card they want.
- Accounting:
 - Still no central GPU accounting, though most (all?) cloud providers can supply some usage data. (Deniza as the capacity manager will also be able to help.)
 - CPU accounting is only as good as its inputs, but it's getting closer to realistic usage numbers: **Do not ignore it.**
- Memory:
 - Activities now generally supply memory requirements for CPU and GPU in a consistent manner.
 - Just be wary of any overview plots you see.

More pitfalls

Storage

- Storage requirements vary widely:
 - So do price points :-S
 - It is particularly important that the computing model makes it clear which types of storage are needed for what.
 - For long term planning purposes (hardware provider!) we need to group storage in different categories
 - Bulk storage/longterm/cheap&cheerful (e.g. echo, any grid storage element, typically HDD)
 - High-performance/longterm (not so cheap and cheerful, typically SSD)
 - Temporary storage (“diskspace”) attached/close to CPU/GPU: Once the VM is gone, so is the data.
 - Other?



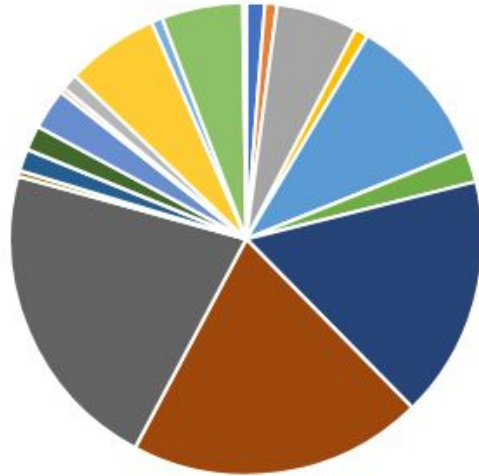
Feedback

- IRIS aims to provide one round of feedback on any submitted request before it reaches the reviewers.
- Feedback is provided by the RSAP team: Matt Doidge, Sophie King, Deniza Chekrygina and me.
- All submissions have been volunteered for and feedback should reach you by tomorrow. Hopefully.
- But: **The final responsibility for the completeness of the request lies with the activity.**

Requests: submission status

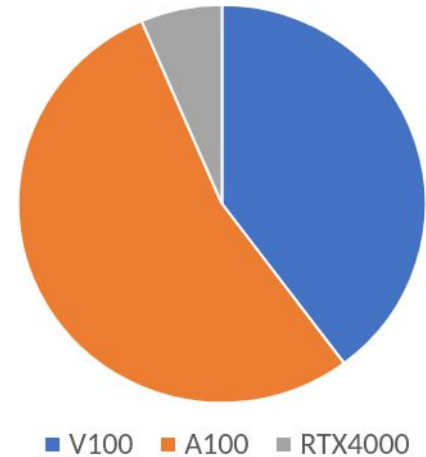
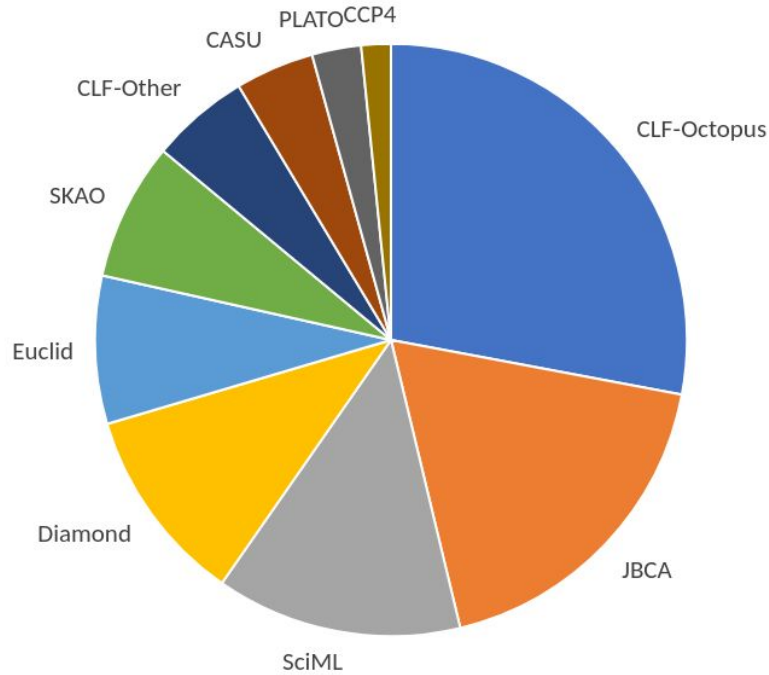
Activity	Request received	Expected	Activity	Request received	Expected
CASU	YES		JBCA	YES	
CCP4	YES		JINTRAC	YES	
CLF-Octopus	YES		JLAB	YES	
CLF-Other	YES		LSST	YES	
CTA	no: change in strategy	N/A	LZ	YES	
DIAMOND	YES		MICROBOONE	YES	
DUNE	YES		PLATO	YES	
EUCLID	YES		SKAO	YES	
GAIA	YES		SCIML	YES	
ISIS	YES		UKSRC (SKA)		YES
Ligo/Virgo	No.	Maybe.	WFAU (new!)	YES	

Very preliminary summary - CPU



~ 40k physical cores
~ 65k requested cores, taking into provider mapping of vCPU to physical cores into account

Very preliminary summary - GPU

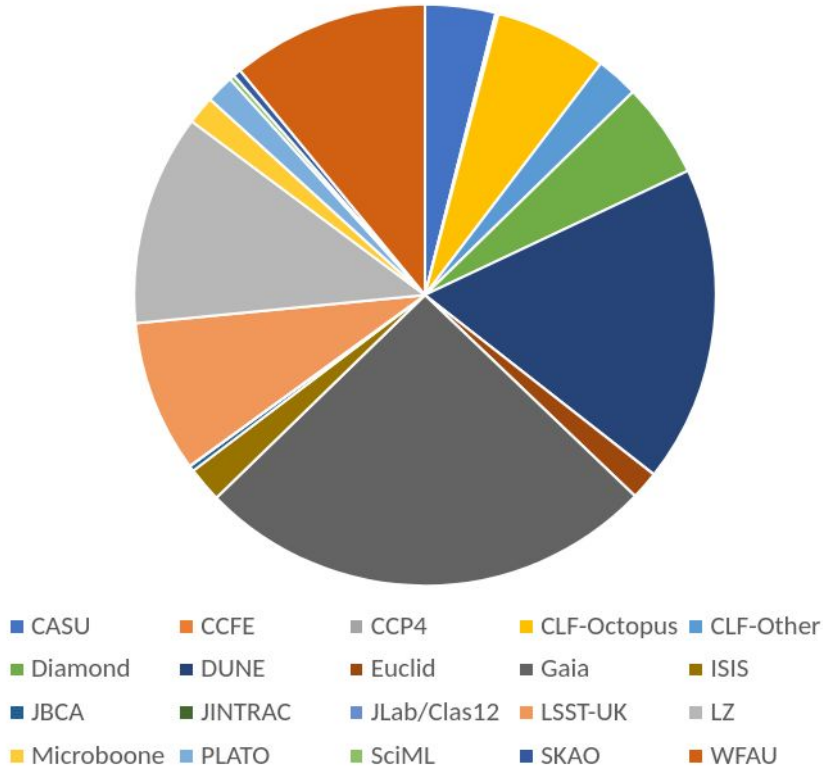


100 GPUs, ~50 % astrophysics

Plot by D. Chrekrygina

Very preliminary summary - Storage

Total: ~ 25 PB



Plot by D. Chrekrygina

Conclusion

- Overall the RSAP achieves its goals:
 - Prediction of future provision needed.
 - Accountability of all parties concerned.
- We are here to help:
 - The goal is to supply the resources needed to the activities, not to withhold them:
 - Please work **with** us.
 - We now have a dedicated capacity manager (Deniza), which allows us to better monitor and allocate the resources we have.
- Questions ? Please come and find me.
 - Alternatively: rsap@iris.ac.uk