Optimising RAC and system usage

Alastair Basden

Durham University
Institute for Computational Cosmology

DiRAC
High Performance Computing Facility

The RAC process

- Research Allocation Committee
- Annual calls
  - Typically opening July, closing September
    - Science and technical components
      - Scored separately
  - Deadlines for review early November
  - Panel meetings in December/January
  - Awards announced prior to April start
    - Lack of storage can often cause complications

Science application

- RAC panel reviewed
  - 3 panel members review each proposal
  - A score between 1-10 is given
    - The RAC no longer score individual criteria separately
    - This is combined with the technical score
      - 80%-20% split
    - Targeted cuts may be suggested (subprojects)

Technical application

- Split between sites
  - Reviewed by RSEs and RTPs
    - With moderation
- Points we look for:
  - Does the technical expertise exist?
  - Is the software suitable
  - Is there evidence for scaling up to max job size?
  - Is the workflow clear, and are bottlenecks identified?
  - Is the compute time reasonable and is job breakdown justified?
    - e.g. could less CPU time be used if running on fewer nodes when scaling is poor?
  - Are the storage requests reasonable?
  - Does it require DiRAC, is it suitable for the chosen service?
  - Is it all coherent, do the numbers add up, etc.

Technical review

- A key part of the process (20%)
  - So please don’t rush filling it out!
  - Scoring is fed to the RAC panel
    - Can be used to adjust the award
  - Requests may be moved to more appropriate systems
- Please take note of the box about unusual job requests
  - In particular Slurm reservations

Alastair Basden
DiRAC / Durham University

RSE application

- DiRAC fund 5 RSEs per year
- Applying for RSE time is good
  - Shows a willingness to engage
- RSEs can help with:
  - Code efficiency
  - GPU porting
    - Becoming increasingly important
  - Code hygiene
- DiRAC also offer RSE investigatory scoping studies:
  - Is my code suitable for GPU and how much effort will it take?

Alastair Basden
DiRAC / Durham University

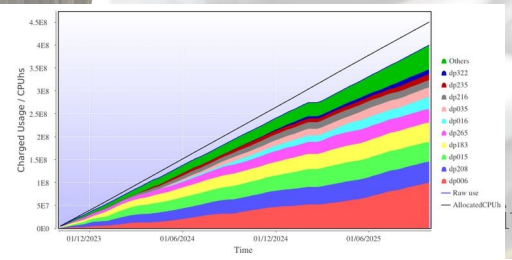
RAC18 notes

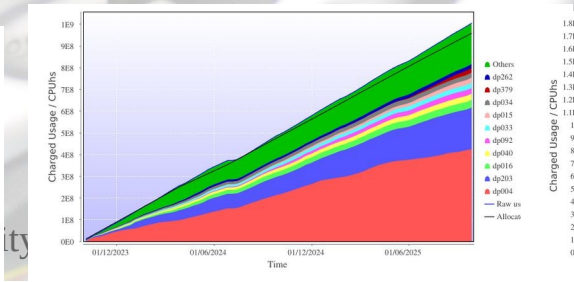
- Oversubscriptions:
  - GPU year 1 request oversubscribed by 1.12x
    - Almost all of GPU requests will be allocated
      - Ask if you need help with code porting!
  - CPU year 1 request oversubscribed by 2.35x
    - Note: for DiRAC-4 we are assuming a large fraction of these projects will move to GPU
      - Based on what you've told us and some assumptions!
  - Year 2 only oversubscribed on COSMA8 (1.45x)
    - Actually more, considering we would want to reserve time for RAC19:
      - ~5x if we want to leave as much for RAC19 as is available for RAC18
  - Year 3 not oversubscribed (2.7x if leaving same amount available for RAC19, RAC20...)
- DIaL2 switching off in 20 days time
  - DiRAC 2.5x system: same vintage as Tesseract, CSD3-Skylake, COSMA7)

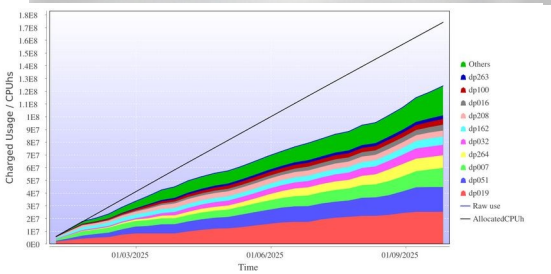
Alastair Basden
DiRAC / Durham University

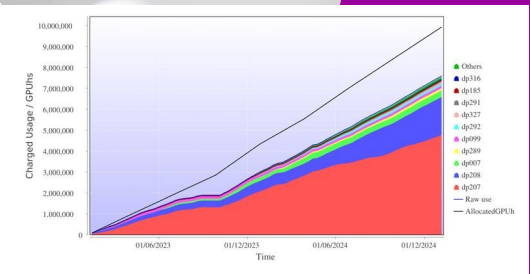
Once awarded

- Please try to make good use of your allocation
  - We like to keep the systems well used
    - Our usage is generally excellent
- Usage is assessed at the end of every quarter
  - If your usage falls below 80% (50% for small projects) of your allocation within a quarter:
    - You will receive an email (from me)
      - It is in your interest to respond – even if there is no valid reason
    - Two quarters in a row without a valid reason are raised with the RAC
      - This can affect future award of allocations









Quota reductions

- Trial for projects on COSMA:
  - Half way through a quarter if you are predicted to underuse by <50%, you will be contacted
    - With the offer of handing time back
      - So that it no longer counts against you
      - And can be reallocated to other projects
    - Most PIs ignore this
      - and then go on to underuse their allocation
    - A small number hand back time
  - This has been successful, may be rolled out to other systems
- Please also be proactive
  - If your past usage has been low, offer to hand time back

Alastair Basden
DiRAC / Durham University

Future trial

- Partial allocation carry-over
  - To avoid a rush at the end of a quarter
    - And avoid needless jobs
- Automatic carry-over of 50% of unused allocation into next quarter
  - Allows for some contingency
  - Details/consequences to be worked through

Alastair Basden
DiRAC / Durham University

Non-DiRAC systems

- We offer some support for DiRAC users using other systems
  - Advice compiling codes
  - Getting codes running
  - Architectural optimisations
  - As far as we are able
  - If you need help, please ask!
    - This also benefits DiRAC
    - And we can talk to the services and help them evolve in the future

Get in touch!

- Do get in touch if you experience problems
  - But don’t leave it too late
- If you want to move time between quarters:
  - Contact us before the quarters start!

Alastair Basden
DiRAC / Durham University

Alastair Basden
DiRAC / Durham University

Alastair Basden
DiRAC / Durham University

Alastair Basden
DiRAC / Durham University

Alastair Basden
DiRAC / Durham University

Alastair Basden
DiRAC / Durham University

Alastair Basden
DiRAC / Durham University

Alastair Basden
DiRAC / Durham University