



MONASH
University

Big Science Workflows and Orchestration

Ulrik Egede
School of Physics and Astronomy
Monash University

28 August 2020

 @UlrikEgede



Welcome

Very encouraging with above 80 people registered for today

- Science areas with large data processing requirements
 - Radio astronomy, particle physics, gravitational waves, remote sensing, medical physics, free electron lasers, bioinformatics, ...
- Scientific computing providers
 - 9 Universities, NCI, Pawsey, CERN, ...
- Representatives of the technologies to be adapted from
 - AARNet, CERN, Imperial College London, ...
- Other platforms that we can learn from
 - IRIS (UK), ESCAPE (EU), Characteristics of scale, Bioinformatics

The problem to solve

- Many areas in the physical sciences are faced with
 - Very large datasets that are produced at international and central facilities
 - Challenges in managing the data for local processing
 - The requirements for centralised processing of data with output to be distributed elsewhere
 - Individual users having large processing requirements that have vastly outgrown the *point-and-click* level

The ARDC platforms

- The proposal is part of the [ARDC Platforms Open Call 2020](#)
- A proposal should have two aspects:
 - A technical platform that brings together existing technologies and adapt them to solve scientific problems in Australia better and faster
 - A social platform that allows for these technologies to be user by as wide a user community as possible
- The proposals are
 - Not about developing new technologies from scratch
 - Not providing ongoing support for already existing platforms
 - Not proving hardware

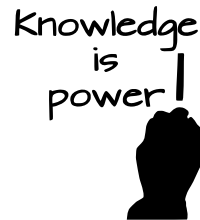
ARDC platform bid

- A platform can seek funding from ARDC of up to A\$1M over a 2.5 year period
- A large number of Expressions of Interest were submitted so clearly it will be very competitive – total envelope is A\$10M
- Deadline for bid is 18 September, so a few busy weeks are ahead of us
 - Use cases are an important part
 - Will be built around the science presentation that we see today

Proposal outline

An **Expression of Interest** has been submitted already

- 1) Data orchestration
- 2) Software access
- 3) Computing access to data
- 4) Large-scale data access
- 5) Training and Community



Workshop – part II

- Next Tuesday from 2pm – 4.30pm
- Focus will be on developing the actual proposal
- Discussion about co-contributions to the bid
- Would like to see representatives there from Universities and Institutions that wish to be part of implementing the platform

Housekeeping

- As a virtual meeting, I would like to pay my respects to the lands of the many traditional owners where we currently are. I would also like to pay my respects to their Elders, past and present
- We are many people in the meeting today
- Please stick to the allocated time for presentations
- For questions, please use the *raise hand* feature in Zoom
 - Questions directly to presentation can go straight after
 - For broader questions and discussion, please leave them to end of each session