

Scientific Applications and Workflows Support

Best Practices, Challenges and Approaches

eResearch Australasia 2025

Scientific Applications and Workflows Support – Best Practices, Challenges and Approaches

Dr Maciej Cytowski¹, Mike Lynch², Glen Charlton³, Dr Rui Yang⁴, Anselm Motha⁵

¹Pawsey Supercomputing Research Centre, Perth, Australia, ²The University of Sydney, Sydney, Australia, ³Intersect Australia, Sydney, Australia, ⁴National Computational Infrastructure, Canberra, Australia, ⁵University of Technology Sydney, Sydney, Australia

Situation

With developments in Artificial Intelligence (AI), scientific research increasingly relies on complex computational workflows that integrate diverse applications, data sources, and infrastructure. Supporting these workflows requires a deep understanding of both domain-specific needs and evolving technologies. Additionally, the diversity of applications and tools used by researchers grows exponentially with the development of new approaches in High Performance Computing (HPC) and AI.

Task

This session invites practitioners, researchers, and support professionals to share and discuss best practices, challenges, and strategies for enabling robust, scalable, and reproducible scientific workflows. We will explore approaches to workflow design, automation, and optimization across a range of disciplines, with a focus on practical and scalable support models, user engagement, and cross-institutional collaboration. Topics may include workflow management systems, reproducibility and containerization, GPU porting and optimisation, data management and storage, application enabling and integration with HPC, cloud, and hybrid environments.

Action

This session aims to foster a collaborative dialogue around what works, what doesn't, and what's next in supporting scientific applications and workflows. The session will begin with short presentations from leading institutions in Australia, focusing on changes and challenges in scientific applications and workflows support over the last couple of years. We will then switch to discussion in subgroups to focus on requirements of specific domains and selected aspects of the support ecosystem.

Result

We hope that participants will leave with actionable insights and new connections to strengthen and improve services (providers) and leverage or integrate with existing services (researchers).

Agenda

- Who are we? Where do we come from? What we would like to talk about? (about 15min)
- Moderated group discussions - challenges, solutions, interesting examples (about 30min)
- Presentation of conclusions (about 10min)
- Wrap-up (5min)

Introductions

Transition to from no Queuing to Queuing

Moving a mainly GUI user base to a queue based system



```
$ qstat
```

Job id	Name	User	Time Use	S	Queue
211.hpcnode@	scaffold.build.	u110234	570:36:5	R	workq
235.hpcnode@	Histomonas.map.	u100123		Q	workq
236.hpcnode@	run_job.sh	u999777		Q	workq

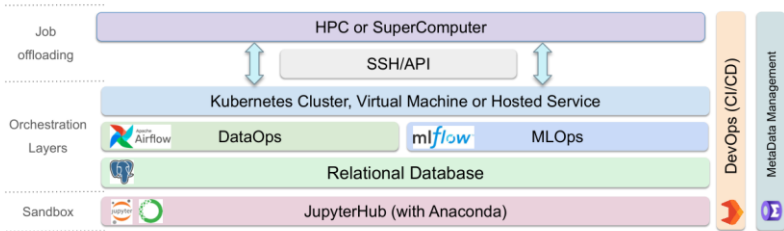
How can we best use containers?

- Traditional HPC: containers? **maybe**
- BioPlatforms and other workflow efforts: containers, **hopefully**
- Modern ML/AI infrastructure: containers are **mandatory**

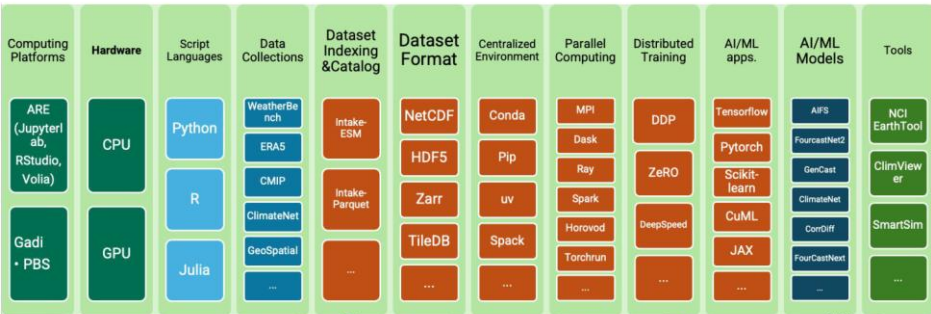
Dockers, Images, and Kubernetes

Researchers are typically proficient in working with Docker. Docker is an isolation level above the operating system which allows creating your own bundle of the operating system + deep learning environment and packaging it within a single file. The file is called a ***docker image***.

:: Operationalising AI on HPC with Hybrid Orchestration Infrastructure



AI/ML Environment @ NCI



Accelerated Computing @ Pawsey

- Pawsey's transition to accelerated computing with Setonix supercomputer
- PaCER and Uptake Projects experiences
- Significant benefits: time to solution, energy efficiency, better software practices
- Main challenges: portability, code maintainability, vendor lock-in, new skills and fast changing technology



Topic groups

How to go from laptop to supercomputer? (Anselm)

How containers can be used to do things better? (Michael)

How to operationalise AI on HPC? (Glen)

How to tackle the complexity of AI/ML software on HPC? (Rui)

How to develop portable scientific codes on accelerators? (Maciej)

Thank You :-)