

# Galaxy Australia

# What is Galaxy?

- A web interface to a complex mixture of computing resources
- A launching site for analysis tool(s) with pre-determined and scalable resources keyed to a users dataset size
- A repository for reference data required to support analysis
- A workflow creation and run platform
- A data and results sharing platform

# More examples of how Galaxy is used

28 JANUARY 2022

## Virus research tips Galaxy Australia over **3 million** jobs

31 AUGUST 2022

The Galaxy Australia service is being chosen to complete their bioinformatics analyses. submitted across a broad spectrum of critical real world.

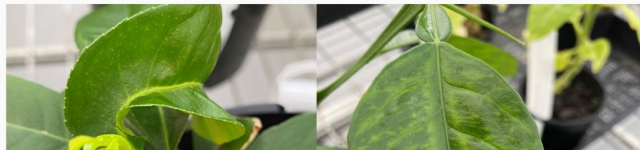
**Uncovering the secrets of parasites through AI and Cloud computing**

**10 million jobs - last week!**

17 MARCH 2021

## Better plant quarantine enabled by Galaxy Australia's **two millionth** job

To celebrate the broad uptake and diversity of their users, Galaxy Australia took the opportunity to spotlight the researcher who





WEEKLY DROP IN

## Top Tips from Galaxy Australia

Drop in for tips, tricks, and nifty ways to use Galaxy!

11 - 11:30 am AEDT, Wednesdays

16, 23 & 30 Oct

[biocommons.org.au/events/galaxy-top-tips](https://biocommons.org.au/events/galaxy-top-tips)

Galaxy Australia is an **open, web-based** platform for accessible, reproducible and transparent computational research. Galaxy supports thousands of documented and maintained tools that are free to use. We facilitate on-demand training capacities and provision **600GB** for Australian institutional (and

History + ≡

search datasets

TIAAS test of 'BPA Fungi  
- Genome Assembly  
Quality Control'

<https://training.galaxyproject.org/training-material/topics/assembly/tutorials/a...>

training

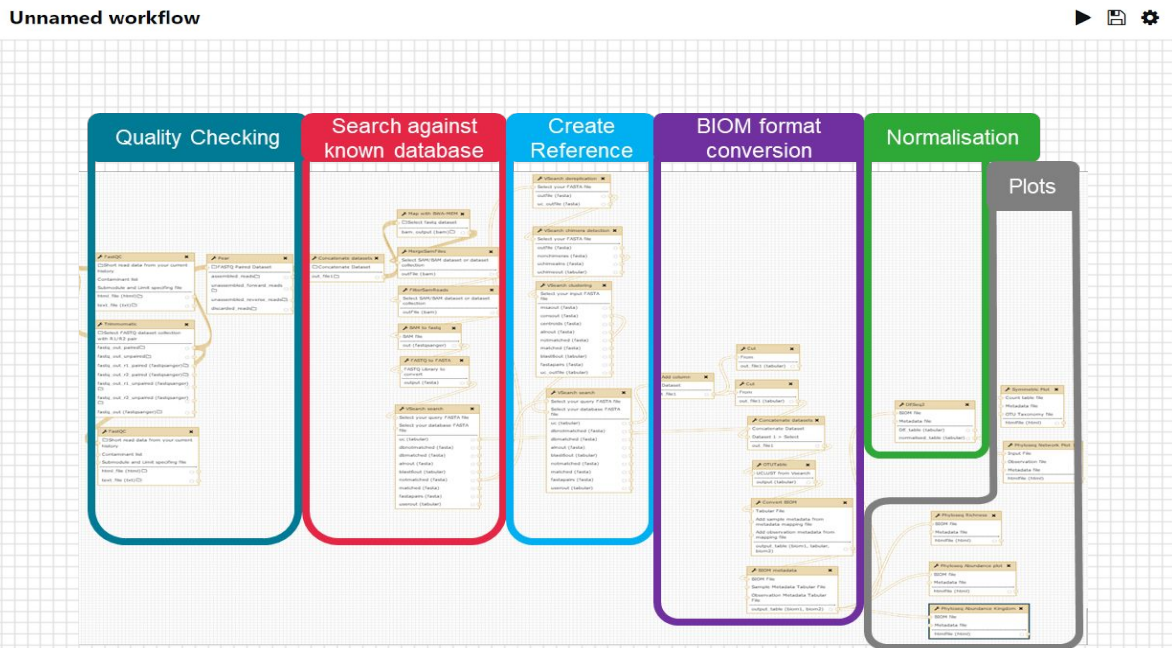
371 MB

61 45

- 92: Chromeister on data 5 and data 6: Comparison score
- 91: Chromeister on data 5 and data 6: Detected events plot
- 90: Chromeister on data 5 and data 6: Detected events
- 89: Chromeister on data 5 and data 6: Comparison metainformation
- 88: Chromeister on data 5 and data 6: Comparison dotplot
- 87: Chromeister on data 5 and data 6: Comparison matrix

# Efficiency through Galaxy controlled scheduling

- Tools
- search tools
- Inputs
- FILE AND META TOOLS
  - Get Data
  - Send Data
  - Collection Operations
  - GENERAL TEXT TOOLS
  - Text Manipulation
  - Filter and Sort
  - Join, Subtract and Group
  - GENOMIC FILE MANIPULATION
  - FASTA/FASTQ
  - FASTQ Quality Control
  - SAM/BAM
  - BED
  - VCF/BCF
  - Convert Formats
  - COMMON GENOMICS TOOLS
  - Operate on Genomic Intervals
  - Extract Features
  - Fetch Sequences/Alignments
  - GENOMICS ANALYSIS



Details

**Edit Workflow Attributes**

**Name:**  
F1000 workflow

**Version:**  
Version 0, 0 steps (active)

**Tags:**  
BIOM X

Apply tags

to make it easy to search for and find items with the same tag.

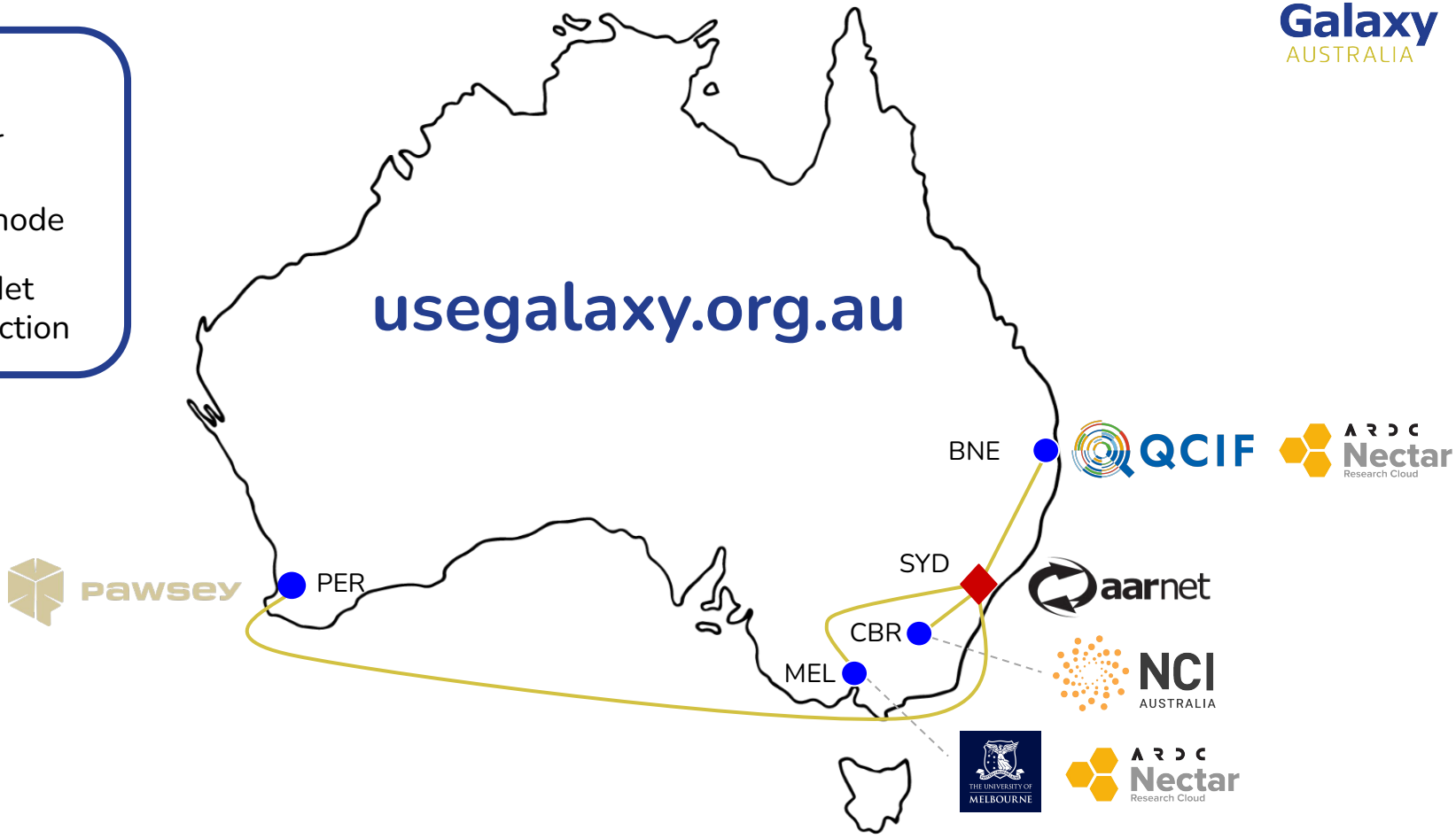
**Annotation / Notes:**  
Describe or add notes to workflow  
Add an annotation or notes to a workflow; annotations are available when a workflow is viewed.

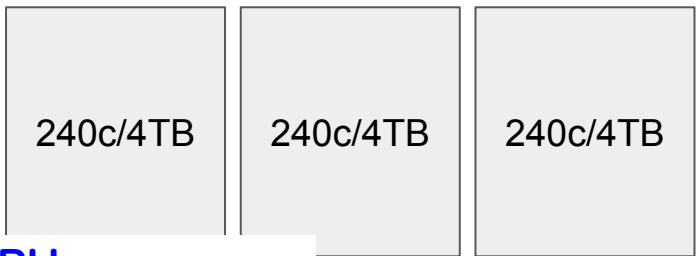
# Galaxy Australia deployment | 2024



## Key

- Pulsar node
- ◆ Main node
- AARNet connection





High CPU

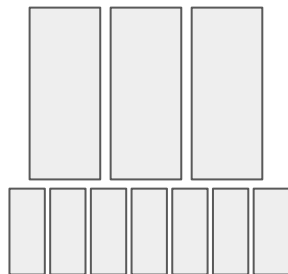
High memory

GPGPU

Rapid I/O

Commercial

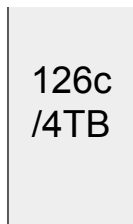
local slurm cluster:  
7  
16c/62GB +  
x 32c/125GB



BNE

SYD

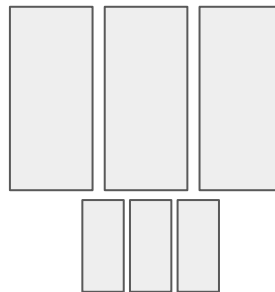
MEL



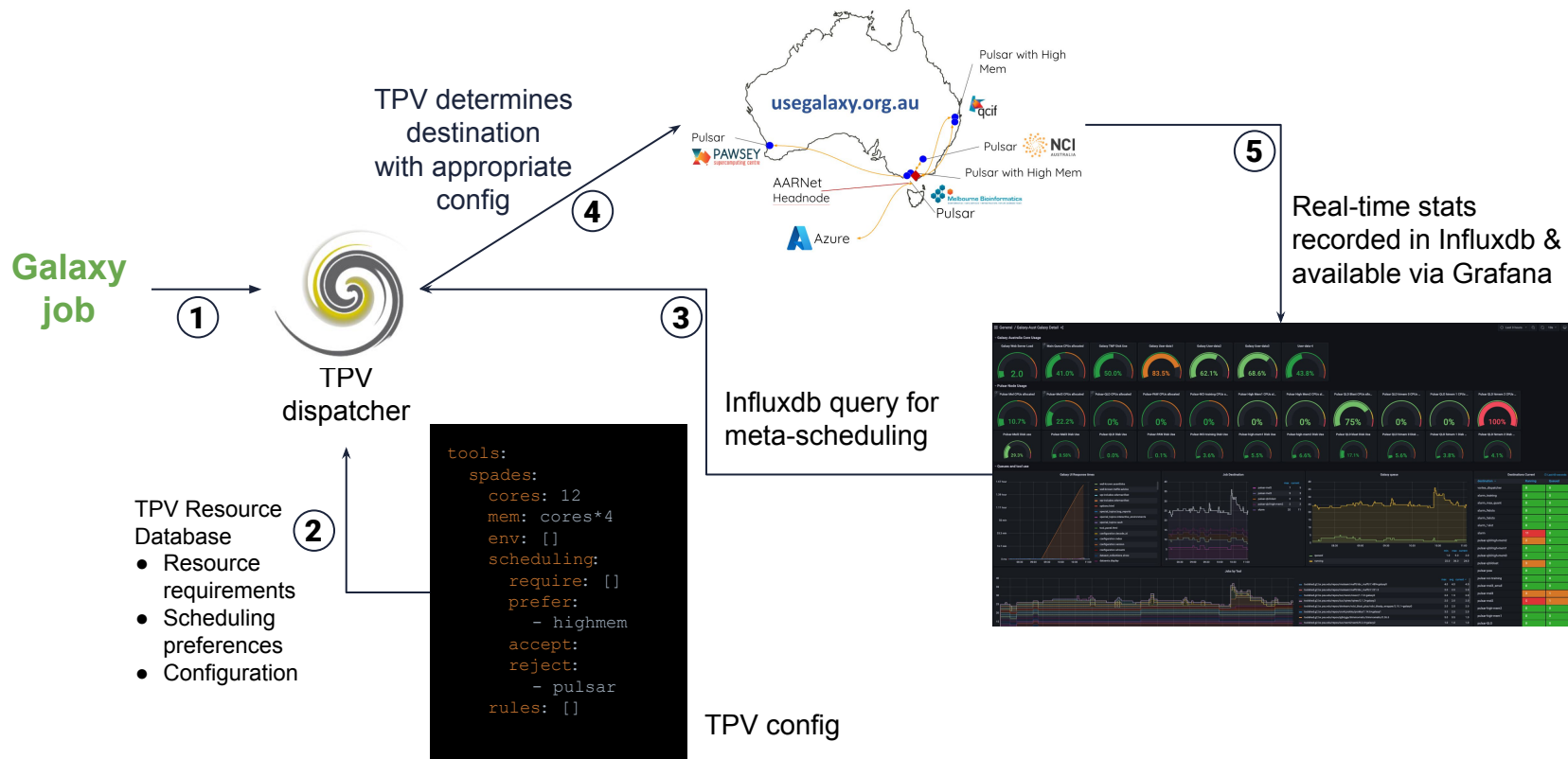
7 x 8c/31GB



3 x 16c/62GB +  
3 x 32c/125GB

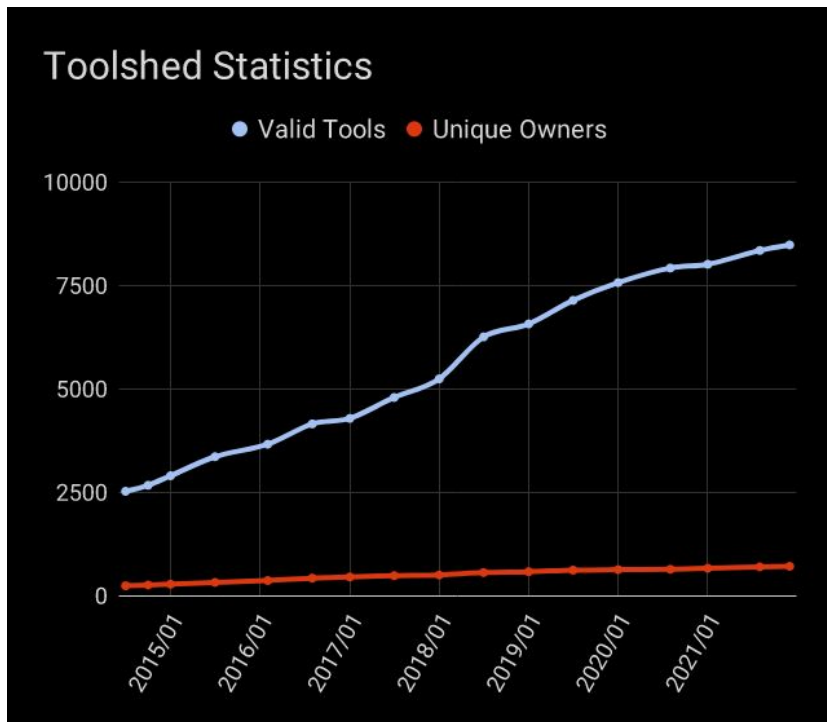


# TPV Job Dispatch Process for Galaxy Australia





# Galaxy Tools and References



## Galaxy Tool Shed

- <https://toolshed.g2.bx.psu.edu/>
- 10,147 tools - October 22<sup>nd</sup> 2024
- 51 different tool categories
- Tool Shed Wiki:  
<https://galaxyproject.org/toolshed/>

## Galaxy References

- <http://datacache.galaxyproject.org/>
- 100's of reference genomes - precomputed for rapid analysis
- Backed by the CernVM File System

# Galaxy Australia Team



Gareth Price



Catherine Bromhead



Justin Lee



Nuwan Goonasekera



Michael D'Silva



Igor Makunin



Michael Thang



Cameron Hyde



Tom Harrop



Anna Syme

