



University of
South Australia

TOWN HALL

High-Performance Compute
(HPC) information session

Agenda

- Introduction and Overview of Advanced Computing @ UniSA - **Professor Marnie Hughes-Warrington**
- The progression of HPC @ UniSA - **Mr Ryan Brown – UniSA HPC Specialist**
- What is at your fingertips – **Dr. Katherine Howard – Intersect eResearch Analyst and Consultant**
- Pilot Adopters with the New HPC Service
 - **Associate Professor Lin Lui (STEM)**
 - **Dr. Michael Ricos (Clinical Health Sciences)**
 - **Dr Zen Lu (Business)**
- Q&A session.





University of
South Australia

Advanced Computing @ UniSA

Professor Marnie Hughes-Warrington

Key Research Infrastructure



- National Collaborative Research Infrastructure Strategy (NCRIS)
- National Computational Merit Allocation Scheme



National Research Infrastructure

Emerging Trends

- Artificial intelligence
- Machine learning
- Internet of Things
- Automation

Emerging Research Areas

- Next generation omics
- Quantum technology
- Synthetic biology
- Climate modelling
- Environmental monitoring
- Advanced manufacturing
- Renewable energy



HPC in Key Industries

Key Industries

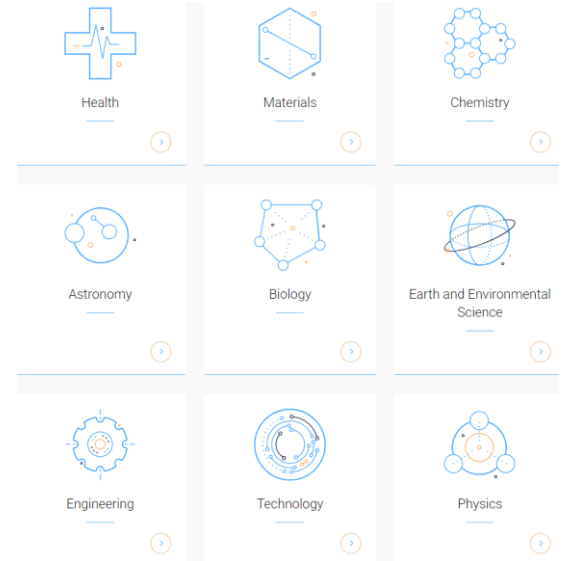
- Financial technology
- Research Labs
- Media and entertainment
- Oil and gas
- Healthcare and Life Sciences
- Government and Defence



National Computing Infrastructure



- Leading edge performance
- High-impact research
- Innovation in research outcomes



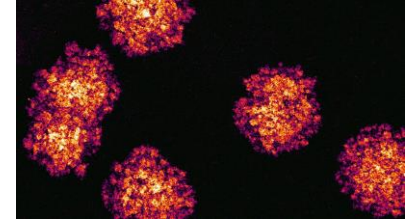
UniSA Research leveraging HPC



Artificial Intelligence
Machine learning
Image processing
Data mining



Civil & structural
engineering



Medical research
Biomolecular
Modelling
Bioinformatics
Human performance
Psychology



Cyber security
Threat detection



Aquaculture



University of
South Australia



University of
South Australia

The progression of HPC @ UniSA

Mr Ryan Brown
UniSA HPC Specialist

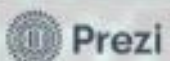
Tango 1.0

Tango 2.0

Engagement

Successes

The Future



eResearch SA



eRESEARCH SA

Tango 1.0 HPC
- 14 node High
Performance Compute
cluster

Tango Cloud
- 8 node cloud
virtualisation platform

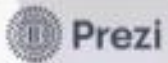
Closure



University of
South Australia

Closure

- Middle of 2019, eResearchSA closed their doors
- Infrastructure was still contracted on lease
- Adelaide and Flinders had their own HPC solutions
- UniSA offered to take on the equipment

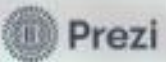


Tango 2.0

As the equipment was still viable

- Re-designed the cluster
- Moved away from virtualisation
- Provided an enhanced 22 node cluster
- 616 CPU Cores
- 9.7TB RAM
- 60TB Direct Attached Storage
- ~25TFlops Peak Performance

expansion



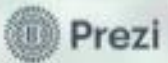
Introduction of GPU

A 23rd node was introduced in late 2020 with GPU compute capability

This allowed for the acceleration of machine learning and AI workloads



Engagement



Successes



Over 27,000 Jobs submitted to cluster

Over 3.5 million CPU Hours Used

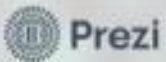
Equivalent to a 4 core desktop running for 100 years

Over 90 unique users

Simon
Doe

Akzam
Saidin

Jisu
Shin



University of
South Australia

Australian National
Fabrication Facility

Worked with Dr Moein
Kashani on computational
fluid dynamic simulations
using Ansys

Processing speed was 2
and a half times faster
than standalone desktops



Simon
Doe



University of
South Australia



Akzam
Saidin

Working with Dr Michael
Ricos in bioinformatics
and genome sequencing

Use of Tango saw massive
speed increases in
processing of data

What used to take 1-2
days was now able to be
performed in 3 hours!



University of
South Australia

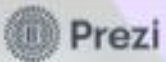
Utilisation of the UK
Biobank Dataset

Working with the
Statistical Genetics group
under supervision of
Hong Lee

Allowed the ability to
offload work from their
National Computation
Merit Allocation Scheme



Jisu
Shin



University of
South Australia

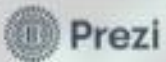
The Future

- Moving to HPC as a Service provided via Intersect
- Reduces in house maintenance
- Availability of eResearch specialists to assist
- Service unit allocations for better tracking of projects
- Utilising NCI Gadi cluster



INTERSECT
RESEARCH FASTER

Gadi



University of
South Australia

NCI Gadi Cluster

- Southern Hemisphere's fastest HPC Cluster
- Ranked 57th in the world on the list of the Top 500 List in June 2022
- 3,200 nodes, 155,000 CPU Cores, 567TB RAM
- 160 nodes with 4 Nvidia V100 GPUs
- Peak performance of over 9 Petaflops
- About 360 times faster than Tango 2.0





University of
South Australia

What is at your fingertips

Dr. Katherine Howard

Intersect eResearch Analyst and Consultant

Intersect

Intersect Membership for the University of South Australia

- HPC as a Service
- Research Support
- Training

University of South Australia Town Hall
1 July 2022

Dr Katherine Howard
eResearch Analyst

Intersect: Overview



Established 2008



48 staff nationwide

Physical presence in NSW, SA, VIC, ACT

Intersect Board

2:1 ratio of member-nominated to independent directors

National eResearch Organisation
Not-for Profit
Membership Based and Founded by Universities



University of South Australia



Members



Nectar Node

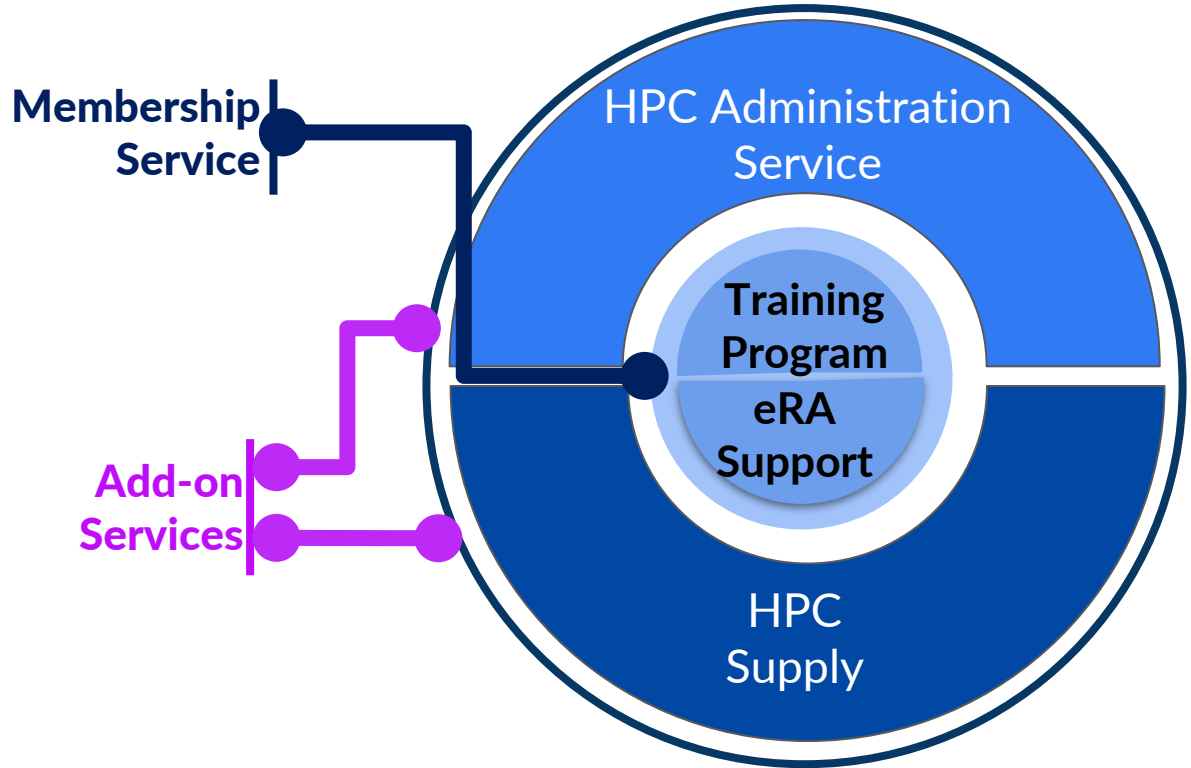


HPC Collaborator



Service Overview

Training
Research Support
HPC as a service



Research Support and the eRA



On-site Expert support and the Services Team



Training

Programming, scripting, and analytics



HPC Access and Support



Researcher Support

Local eResearch Analyst

Primary contact for researcher and organisational support

Researcher Engagement



Research Application Support



Access to Research Technology



Relevant personal expertise...

Opportunity to recruit a local eRA with locally relevant skills and experience ...

...augmented by the Services Team

... while accessing over 50 years of combined skill and experience, in diverse disciplines, from the full eResearch Analyst team



HPC as a Service



Supply and Administration

HPC as a Service



Access to HPC from the National Computational Infrastructure, with allocation support and service from Intersect



HPC supply from NCI



HPC Allocation Service from Intersect

Service Units (SUs) purchased annually upfront (no 'bill shock')



SUs allocated quarterly to ensure maximum usage



HPC Specialist "Expert Hours" (optional/additional)



Access

HPC as a Service



1. Complete an HPC Request form: <https://bpi.unisa.edu.au>
2. Determine how many Service Units you will need. Use this as a guide: <https://intersect.freshservice.com/a/solutions/articles/75000059108>
3. Register with NCI for access to Gadi: <https://my.nci.org.au/mancini/signup/0>

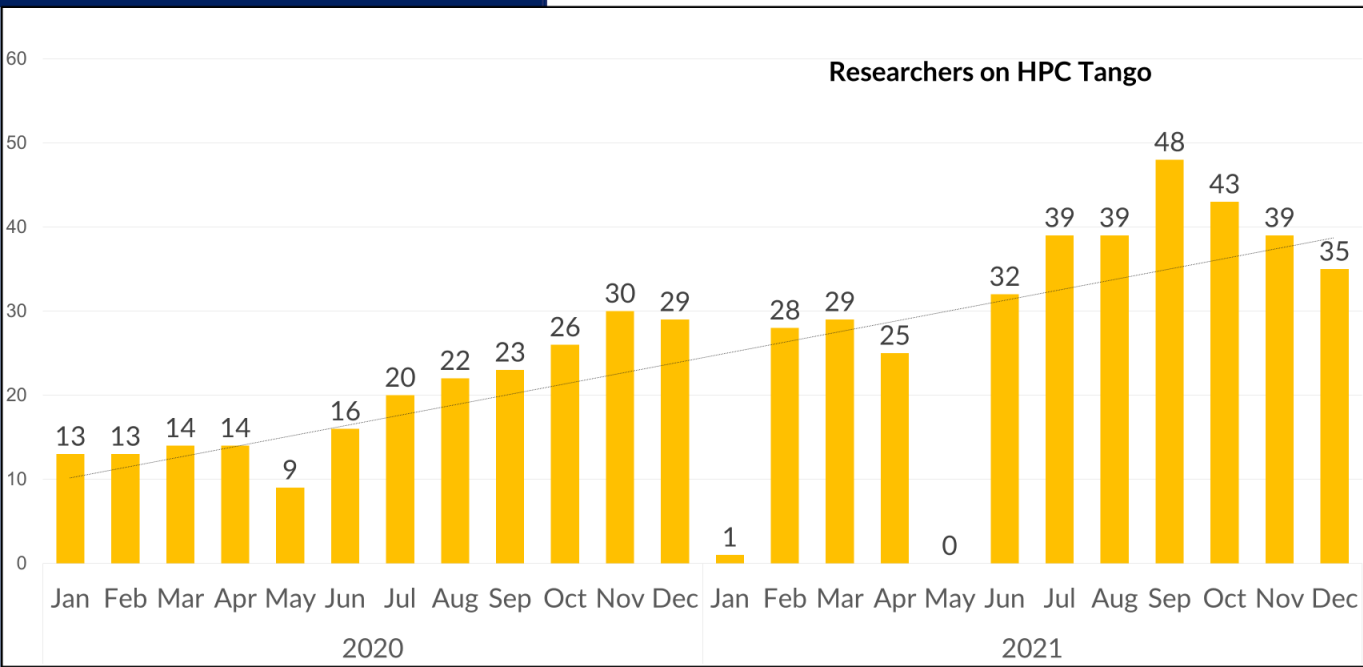
Knowledge base article here:

<https://intersect.freshservice.com/a/solutions/articles/75000058541>



Early adopters

HPC as a Service



NCI Gadi



56

Researchers



Training



Quality Training Requires a Team



Network of 20+ eResearch Analysts & Trainers

New course development

Services Team collectively develops new courses & maintains/updates existing courses

Stakeholder Engagement



Post-training researcher support & engagement



Training Portfolio



Local eResearch Analyst

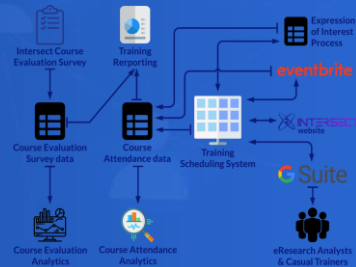
Primary contact for organising & delivering training

Coordinated training strategy and systems



Training Reporting & Analytics

Bespoke Training System



Intersect Course Catalogue

Awareness
Introductory
Intermediate
Advanced

Courses	Number
Programming	25 +
Data Science	20 +
Data Analytics	20 +
Statistics	13
Data Management	8
Data Visualisation	7
Machine Learning and AI	6
Data Collection	6
Research Computing	5



INTERSECT TRAINING 2021

EVALUATION



How likely is it that you would recommend Intersect training courses to colleagues?

Promoters: 9 - 10

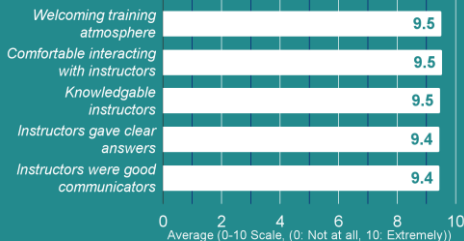
Passives: 7 - 8

Detractors: 0 - 6

Net Promoter Score = % Promoters - % Detractors

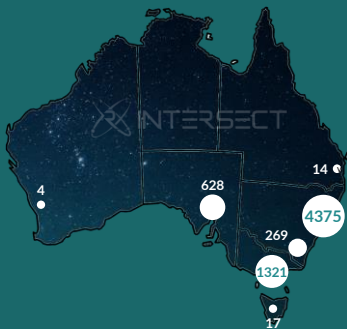
* Based on 2521 survey responses

QUALITY



ATTENDANCE

2021 ATTENDANCE PER STATE



*+100 participants from national organisations.

349 COURSES
Online

0 COURSES
In Person

6728 ATTENDEES
Online

0 ATTENDEES
In Person

5812 ATTENDEES
At their local university

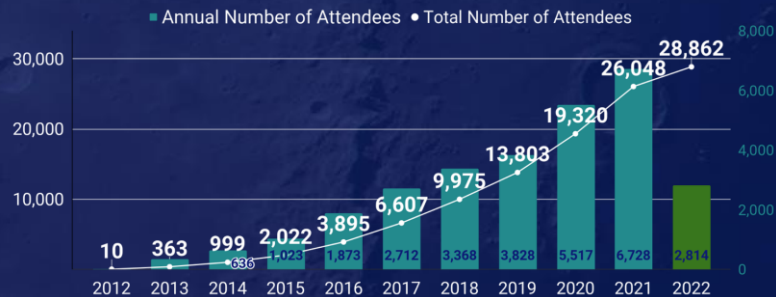
916 ATTENDEES
At other Intersect members

REASON FOR ATTENDING

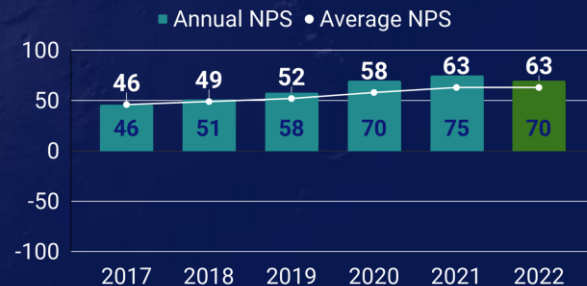
>82% of participants register to Intersect training to **learn skills that they can apply to their work now or in the future**

HISTORICAL TRENDS

ATTENDANCE



EVALUATION



For more information about Intersect Training please contact us at training@intersect.org.au



Who to contact

Training
Research Support
HPC as a service



For more information, see the AskResearch web pages:

<https://i.unisa.edu.au/askresearch/tools-services/computing-resources/>

HPC Request form: <https://bpi.unisa.edu.au/>

Email for assistance: hpc@unisa.edu.au

IT Help Desk: 830 25000 (ext. 25000)

Katherine Howard, eResearch Analyst: katherine.howard@unisa.edu.au





INTERSECT

RESEARCH FASTER



University of
South Australia

Machine Learning Research with UniSA HPC

Lin Liu and Ji-Young Park
UniSA STEM

Research Supported by UniSA HPC

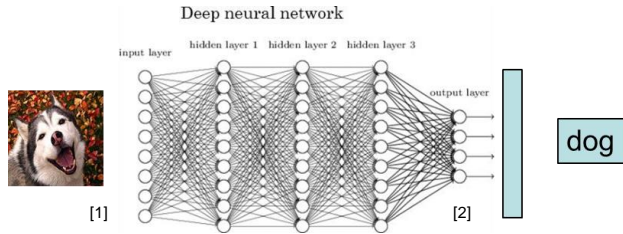
- Research
 - Deep learning & Adversarial machine learning
 - Text mining
 - Bioinformatics
- Outcomes
 - A*/A conference and Q1 journal publications
 - Funding applications
- Continuous demand of UniSA HPC



Thanks to the UniSA HPC Team & The Intersect Team



Machine Learning Research Needs HPC



- Deep neural networks (DNNs)
- Training DNN takes time:
 - PC (> a month)
 - HPC-CPU (12-16 days)
 - HPC-GPU (4 days)

Adversarial Machine Learning

- The goal of my PhD research:
 - build robust defence methods to safeguard DNNs
- Have developed two effective defence methods, one published in a CORE A conference; another submitted to an A* conference
- Support by UniSA HPC:
 - **Tango & NCI Gadi**





University of
South Australia

Pilot

Dr. Michael Ricos

Akzam Saidin

UniSA Clinical Health Sciences

Improving Health Outcomes –Disease Gene Discovery - Genetics, Genomics & Big Data - HPC

- Mutation hunting - Need to find a single letter typo in 6 Billion
- e.g. - single letter typo in 2,255 copies of War and Peace
- Massively Parallel DNA Sequencing
- Fragment DNA to pieces 300 letters long
- Randomly sequence fragments 30 to 100 x
- Assembly fragments (without a reference) or
- Mapping fragments to Human Reference Genome

High Performance Computing : Gamechanger

- Assisted UniSA’s HPC team in Tango 2 Pilot – 2019
- Tango allowed analysis of large cohorts of WES data
- Assist UniSA’s HPC Pilot transition to NCI – GADI – 2022
- NCI - GADI allows analysis of large cohort WGS data
- High Impact Publications - Disease Gene Discovery
- High Impact Publications - Benchmarking Software Tools
- Novel Diagnostic Tools and impacting patient care

Metric	Exome – 60 Million letters	Genome - 6 Billion letters
% Genome analysed	1 %	100 %
Raw Data Size	~ 20 GB	~ 100 GB
Processed Size	~ 60 GB	~ 300 GB
2011 - Desktop	> 1 week	Months
TANGO	~ 1 day	> 1 week
Processed to date	800 WES > 16 TB storage	~ 12 WGS ~ 1.2 TB storage
NCI - GADI	~ 2 hours	~ 1 day
Processing 2022	~ 150 WES ~ 3 TB storage	~ 200 WGS > 20 TB storage
Projection 2025	1000’s WES ~ 20 TB storage / 1000	~1,500 WGS ~ 150 TB storage

- 2022 will have analysed ~ 50 TB of Sequencing Data
- Discovered >50 Genes, Diagnoses for >1000’s
- High Impact Publications – Highly Cited
- 2025 plan-analyse > 2500 Genomes 170 TB data



HDR Experience with HPC facilities

Akzam Saidin | Molecular Neurogenomics and Genetics Group
Supervisors: Prof Leanne Dibbens & Dr Michael Ricos – CHS & ACPReH

2018 – 2019

To identify causative genes & variants for neurological disorders in familial cohort/trio

NGS (WES & WGS) & Bioinformatics

Bioinformatics pipeline testing (Lab PC)

Planned to use HPC@ERSA



500 + 300 WES...

2019 -2022

HPC rebirth - **Tango2**



Deployed analysis pipelines (thanks to Ryan, Scott and HPC team)

Completed Consensus Genome Reads Alignment approach (manuscript in prep.) [1]

Completed 500 WES: Benchmarked joint variant calling (Natasha; manuscript in prep.) [2]

Completed 300 WES: Identified Genes-Variants for 30 Families with Epilepsy & Rare Diseases (manuscript in prep.) [3]

Ongoing WGS: To identify Gene-Variants for 50 unsolved epilepsy cases

2022+

Pilot User for Intersect – NCI



Test with WES & WGS data. Run time

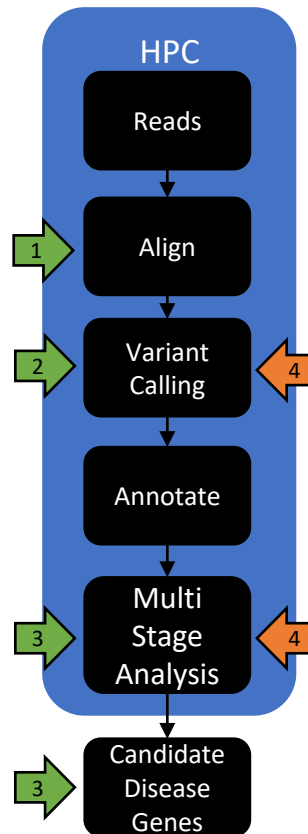
Potential for Deep/Machine Learning on Genomics [4]

Trainings – R, Python, Machine Learning

SIDS Cohort: Pilot WGS Study for SIDS - To identify Gene-Variants for unsolved SIDS disease cases

Genomics Data

Compute Power





University of
South Australia

Pilot

Dr Zen Lu

UniSA Business

Zen's HPC work

- Recently, I was privileged to be part of the UniSA pilot team to use NCI Gadi.
- I use HPC for computations related to a research project on model selection through multiple hypothesis testing which is of computation-intensive.
- For example, for a particular simulation design it could take about a week for a typical desktop machine to complete job. But with NCI Gadi clusters using 28 workers it takes a few hours to complete.
- My experience with Gadi has so far been positive overall.
- Staff at UniSA and Intersect provide excellent assistance to get you start.
- In terms running jobs at Gadi, it appears to take longer queuing time in Gadi than in UniSA's Tango clusters, perhaps due to higher demand for Gadi than for Tango.
- However, once jobs start running, it seems quicker to complete for the same computing resource specification for Gadi than for Tango.





University of South Australia

For enquiries, see AskResearch web pages

<https://i.unisa.edu.au/askresearch/tools-services/computing-resources/#hpc>



University of
South Australia

Q & A

?