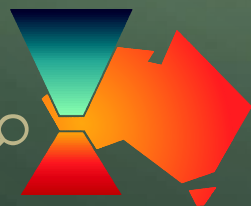


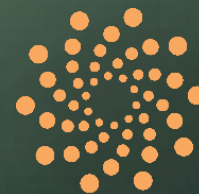


Building a National High-Performance e-Infrastructure for Australian Research and Innovation

Allan Williams
Associate Director, NCI
30th May 2018

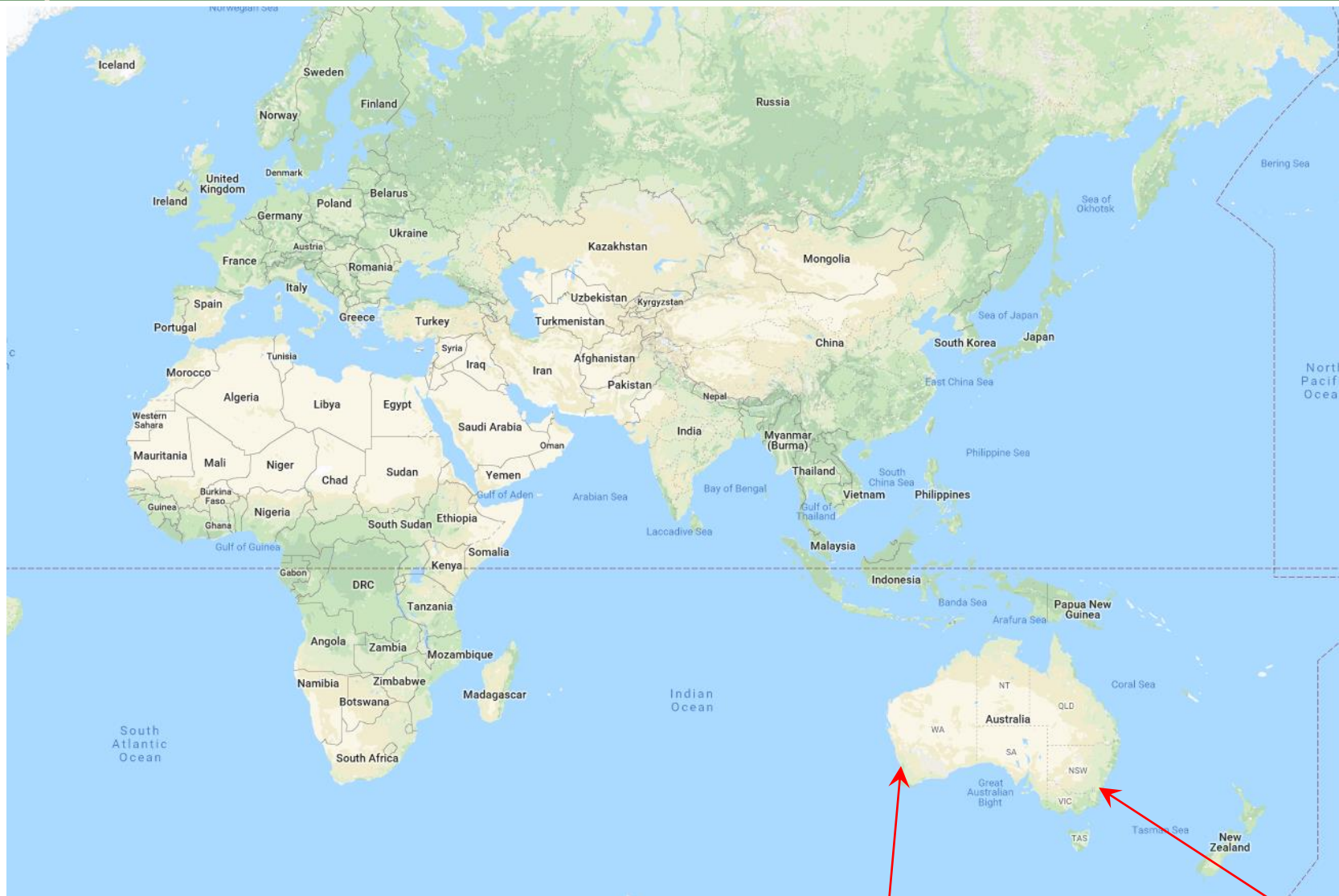


PAWSEY
supercomputing centre



NCI
AUSTRALIA

WHERE ARE WE...





Building a National High-Performance e-Infrastructure for Australian Research and Innovation:

Outline

1. History of HPC in Australia
2. National Roadmaps/NCRIS
3. Drivers shaping HPC services informing a national strategy
4. National HPC at Pawsey and NCI
5. Research outcomes/ impacts
6. 2016 Roadmap and future directions
7. Building blocks for national strategy
8. Collaboration Opportunities

A decorative graphic on the left side of the slide, consisting of a network of thin, light-colored lines and small circles, resembling a circuit board or a stylized tree structure, set against a dark green background.

National High-Performance Computing and Data Services for
Australian Research and Innovation:

1. History of HPC in Australia

HISTORY OF NATIONAL HPC

- 1987 – the first university supercomputer is installed at the Australian National University
- 1989 – the first National Merit Allocation Scheme started with 10% of the machine
(700 hours at a nominal value of \$1M AUD \cong \$700,000 USD)
- 1993 – Australia has 9 machines in the newly created Top 500 list

Rank	Site	Manufacturer	Computer	Processors	Rpeak GFlops
127	Australian National University	Thinking Machines	CM-5/32	32	4.096
140	University of Adelaide	Thinking Machines	CM-5/32	32	4.096
143	University of New South Wales	Thinking Machines	CM-5/32	32	4.096
185	University of Melbourne	Intel	XP/A4	56	2.8
214	Australian National University	Thinking Machines	CM-2/16k	512	3.5
290	ANSTO	Fujitsu	VP2200/10	1	1
291	Australian National University	Fujitsu	VP2200/10	1	1
363	Bureau of Meteorology	Cray Inc.	Y-MP2E/232	2	0.666
365	CSIRO – National Science Agency	Cray Inc.	Y-MP4E/264	2	0.666

HISTORY OF NATIONAL HPC

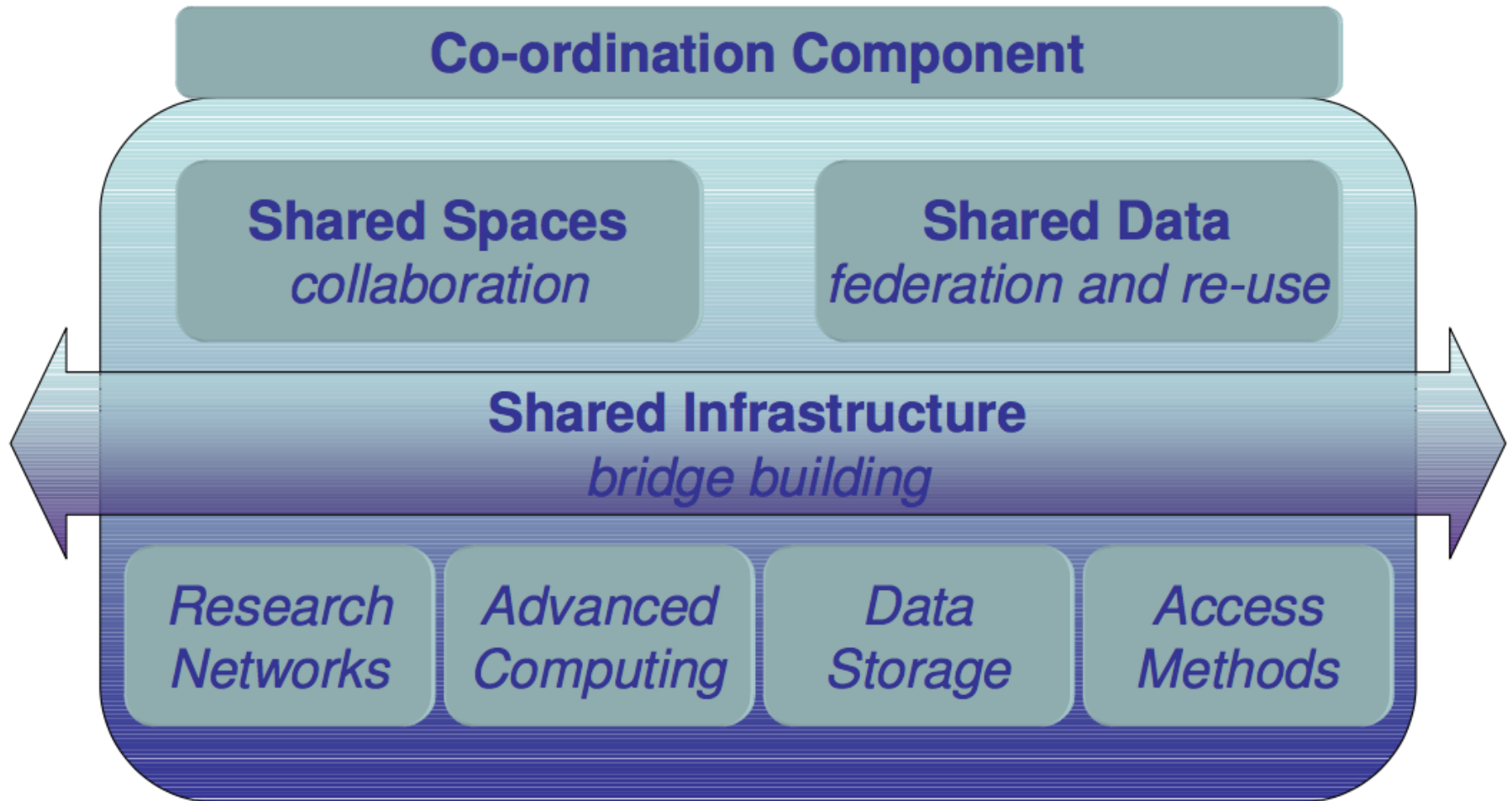
- 1998 – Funding was announced for first national HPC service \$19.5M (€10M)
- 2000 the Australian Partnership for Advanced Computing (APAC) National Facility started with the first compute services commencing in April 2001, debut #31 in Nov'01 Top 500.
(State based partners supporting ~50% was dedicated to merit based access)
- 2000 iVEC (later to be Pawsey supercomputing centre) is the West Australian State member of APAC
- 2005 further injection of funds and purchase of next HPC, debut #26 in Top 500
- 2006 eResearch Coordinating Committee Report - a strategic roadmap to encourage
“a collaborative approach, and provides Australia with the national research facilities and linkages needed to address the economic, social and environmental challenges of the 21st century.”

=> National Collaborative Research Infrastructure Strategy (NCRIS)

A decorative graphic on the left side of the slide, consisting of a network of thin, light-colored lines and small circles, resembling a circuit board or a stylized tree structure, set against a dark green background.

National High-Performance Computing and Data Services for
Australian Research and Innovation:

2. National Collaborative Research Infrastructure Strategy - NCRIS



Rhys Francis : eResearch Infrastructure an NCRIS investment 2006

ROADMAPS

NATIONAL COLLABORATIVE RESEARCH INFRASTRUCTURE
STRATEGY

STRATEGIC ROADMAP

FEBRUARY 2006



Australian Government
Department of Innovation, Industry, Science and Research



2006 – National Computational Infrastructure (NCI) \$26M

- Australian National Data Service (ANDS) \$24M
- Australian Research Collaboration (ARCS) \$20M

2009 – National Computational Infrastructure \$50M

- Pawsey Supercomputer Centre \$80M
- Data Storage (RDSI) \$50M
- Cloud & Collaboration (NeCTAR) \$47M



2012 – No capital funding !
- Agreement to some operational funding

A decorative graphic on the left side of the slide, consisting of a network of thin, light-colored lines and small circles, resembling a circuit board or a neural network, extending vertically from the top to the bottom of the frame.

National High-Performance Computing and Data Services for
Australian Research and Innovation:

3. Drivers shaping HPC services

DRIVERS TO OBTAIN FUNDING

**Comprehensive &
Integrated
Support Framework**

Model/
Simulate/
Generate

High-Performance
Computation
HPC Systems Team

NCRIS
EIF
NCI Collaboration



NCI

Supporting
Researchers

Ingest/
Organise/
Curate

Process/
Analyse/
Visualise

RDSI
ANDS



Data Storage
Data Management

Private/public cloud
OpenStack

NeCTAR
Cloud & Virtual
Laboratories



PAWSEY
supercomputing centre



NCI
AUSTRALIA

A decorative graphic on the left side of the slide, consisting of a network of thin, light-colored lines and small circles, resembling a circuit board or a stylized tree structure, set against a dark green background.

National High-Performance Computing and Data Services for
Australian Research and Innovation:

4. National HPC - Pawsey and NCI

HIGH PERFORMANCE COMPUTING

Ground-breaking Australian computational research is expedited on NCI's high performance cluster, generating experimental results in just days or weeks that would otherwise have taken decades or centuries on a desktop computer.

CLOUD COMPUTING

Enabling collaborative access to specialised computing environments and to more than 10 petabytes of nationally and internationally significant research data.

DATA STORAGE

NCI operates the fastest filesystems in the Southern Hemisphere, linking our high performance computing and high performance data resources, providing an unparalleled service integration for our users.

HIGH-PERFORMANCE COMPUTING INNOVATION

Through maximising research outcomes by developing and optimising codes to run most effectively on a supercomputer, NCI can ensure the most efficient running of high-level models and simulations.

DATA INNOVATION

Meeting the ever-increasing requirements for e-infrastructure investments with 'end-to-end' management of data, including ingestion, curation and delivery.

COLLECTION MANAGEMENT

Simplifying access to petabytes of research data using persistent identifiers and other types of metadata for curation, maintaining a navigable catalogue for users.

VIRTUAL LABORATORIES

Online platforms that make datasets, state-of-the-art analysis and visualisation tools easily accessible in a 'one-stop shop' environment, providing the researcher with everything they need.

VIZLAB

Using cutting-edge visualisations such as image stills, video animations and interactive interfaces to better understand and interpret data across a wide range of scientific disciplines.



New Users



ARC Centres of Excellence



National Agencies



Universities



Nobel Prize Winner



Projects



New Projects



Active Users



Industry Partners

One of Australia's publicly funded tier-1 HPC facility



1500
researchers



170
projects



45
staff

MAKING TOMORROW HAPPEN, TODAY

through



supercomputing



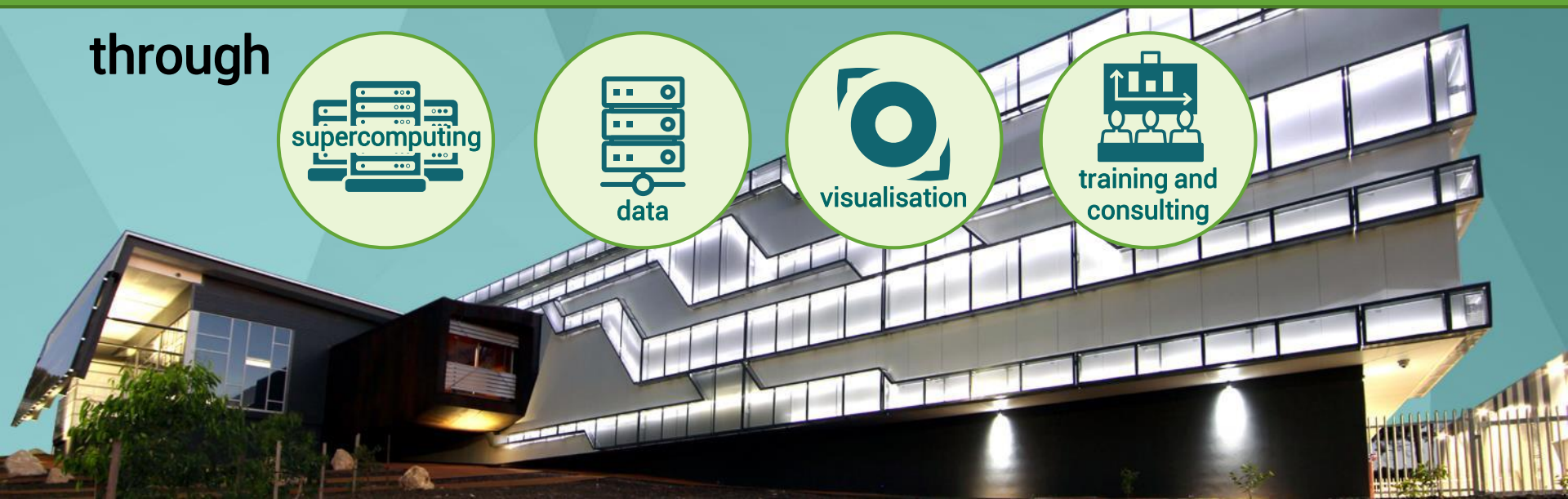
data



visualisation



training and
consulting



A decorative graphic on the left side of the slide, consisting of a network of thin, light-colored lines and small circles, resembling a circuit board or a stylized tree structure.

National High-Performance Computing and Data Services for
Australian Research and Innovation:

5. Research Outcomes /Impacts

Accelerating scientific outcomes



health

Combating Alzheimer's and dementia



170
projects



resources

Protecting Perth's aquifers



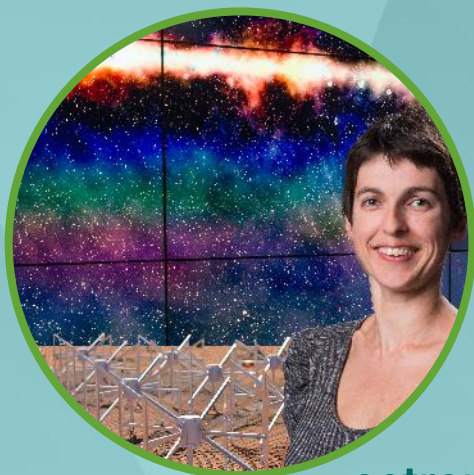
energy

Understanding the ocean to
improve offshore designs



agriculture

AI for targeted weed control



astronomy

A panoramic view of the universe in colour

BIG COMPUTE/BIG DATA — SQUARE KILOMETRE ARRAY



- The Square Kilometre Array (Australia, South Africa) — the largest, most sensitive, radio telescope ever to be built — to answer profound, fundamental questions about the origin of the universe, dark matter and dark energy, and the formation of galaxies.
- Pawsey Centre: a key Australian contribution to international SKA science, now supporting two pathfinder instruments—Australian SKA Pathfinder (ASKAP, left panel—32 mid-freq. dishes, CSIRO) and Murchison Wide-area Array (MWA, right panel—low-freq. antennae, Curtin Uni.)
- Pawsey's Galaxy supercomputer (Cray XC40) provides critical real-time analysis/processing capability for massive data volumes from these telescopes— 2.5 GByte/sec, or 75 PByte/year

RADIO ASTRONOMY



Credit: Photography by Paul Bourke and Jonathan Knispel. Supported by WASP (UWA), iVEC, ICRAR, and CSIRO.

Artemis: A Neonatal Internet of Things



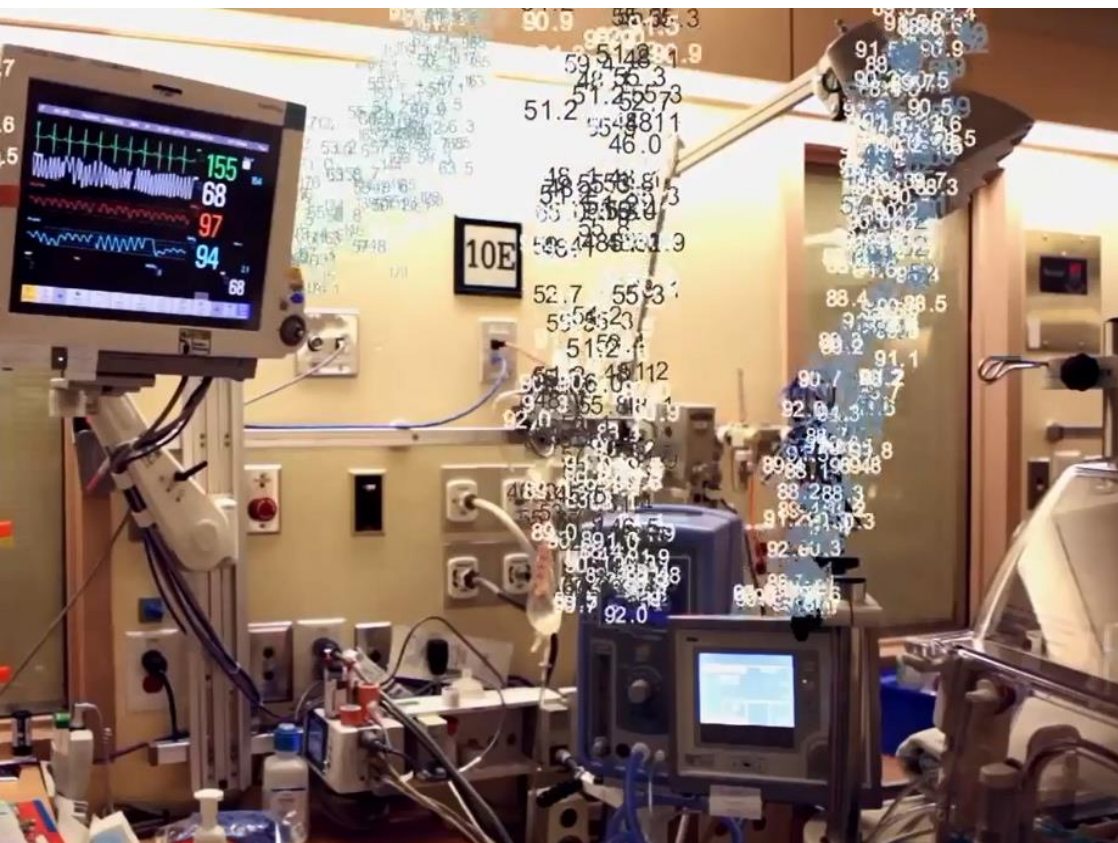
3,300,000 worldwide
neonatal deaths yearly

Prematurity accounts for
41% of deaths in children
under the age of 5

$\frac{3}{4}$ of all newborn deaths
occur within first 4 weeks
of life

Up to $\frac{2}{3}$ of newborn
deaths can be prevented if
known

Artemis: A Neonatal Internet of Things



Collaboration between Pawsey,
WA Dept of Health, UWS, UOIT.

Realtime, remote access to NICU
bed monitoring for clinical
observation

Early identification of health
changes for very vulnerable
patients

Significant possibilities for
expansion in the future

Our Mission:

To provide world-class, high-end computing services for Australian research and innovation

Through:

Enabling a comprehensive, integrated research environment

Backed by

Nationally and Internationally renowned, expert support team

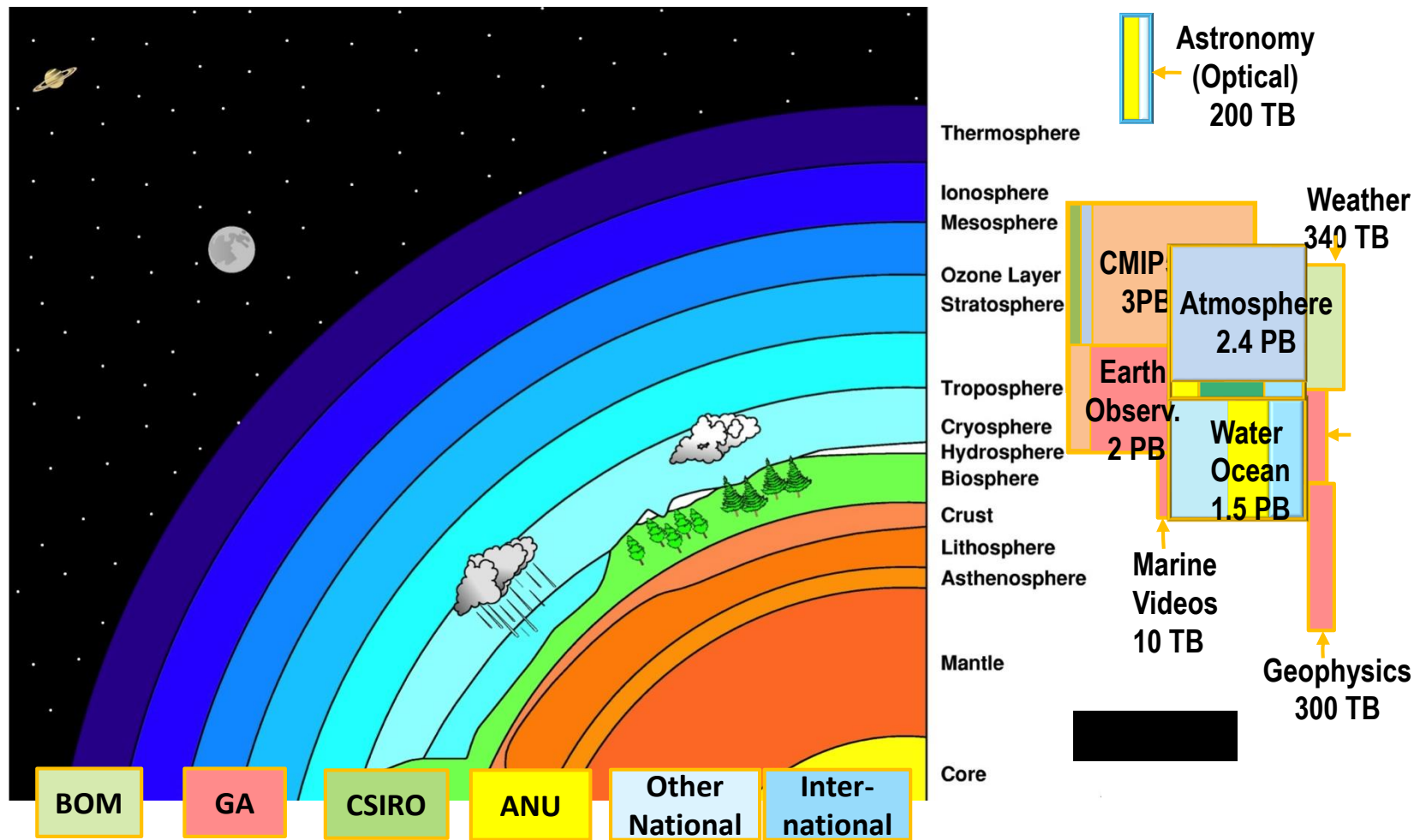


We have established a petascale national data resource that is co-located with high-performance computing.

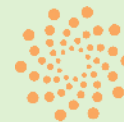
NCI manages 38+ data collections (10+ PB) in multiple categories, including:

- *earth sciences,*
- *climate and weather model,*
- *earth and marine observations and products,*
- *terrestrial ecosystem,*
- *water management and hydrology*
- *biosciences*
- *astronomy*
- *social science and biosciences.*

NCI DATA COLLECTIONS

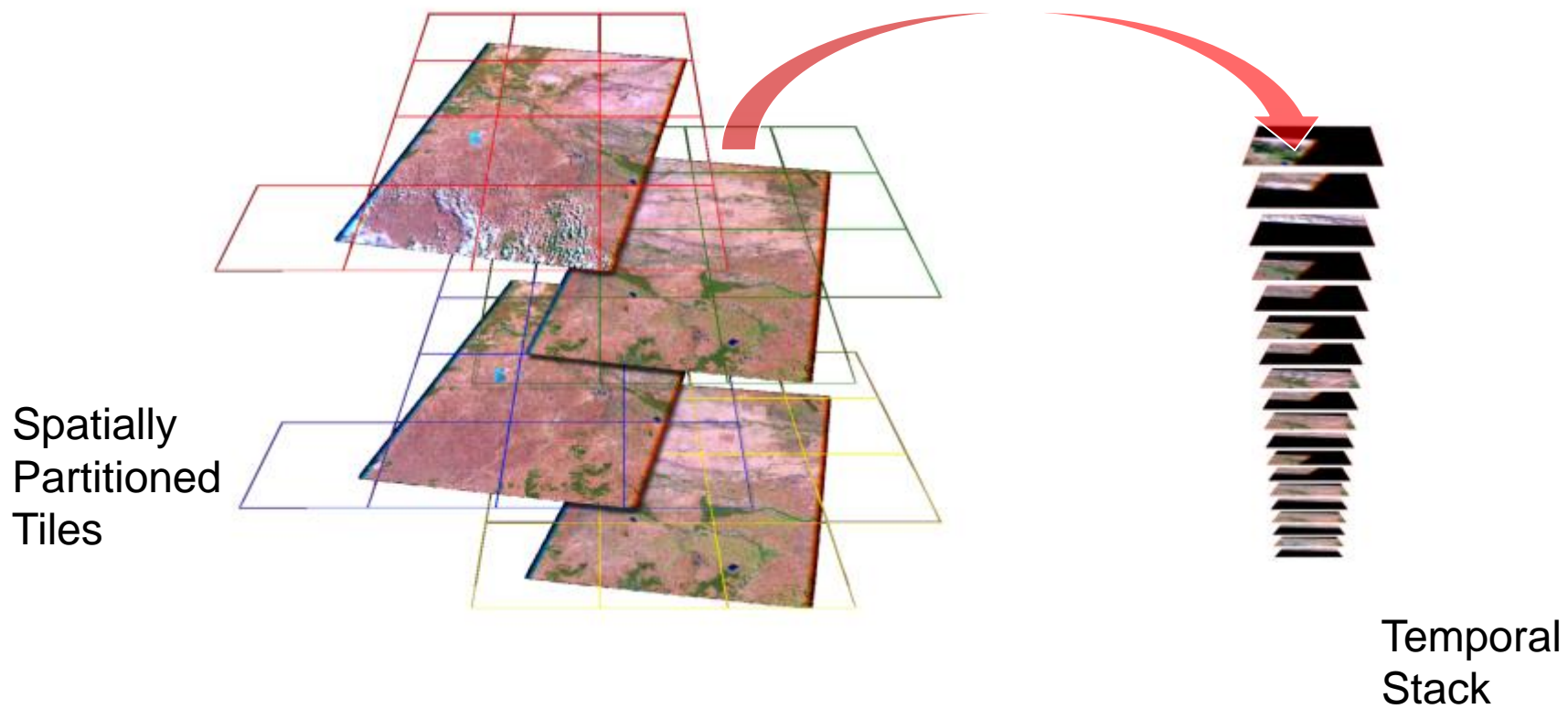


PAWSEY
supercomputing centre



NCI
AUSTRALIA

THE LANDSAT DATA CUBE



Continental Scale Water Observations from Space

27 
YEARS
1987-2014 DATA

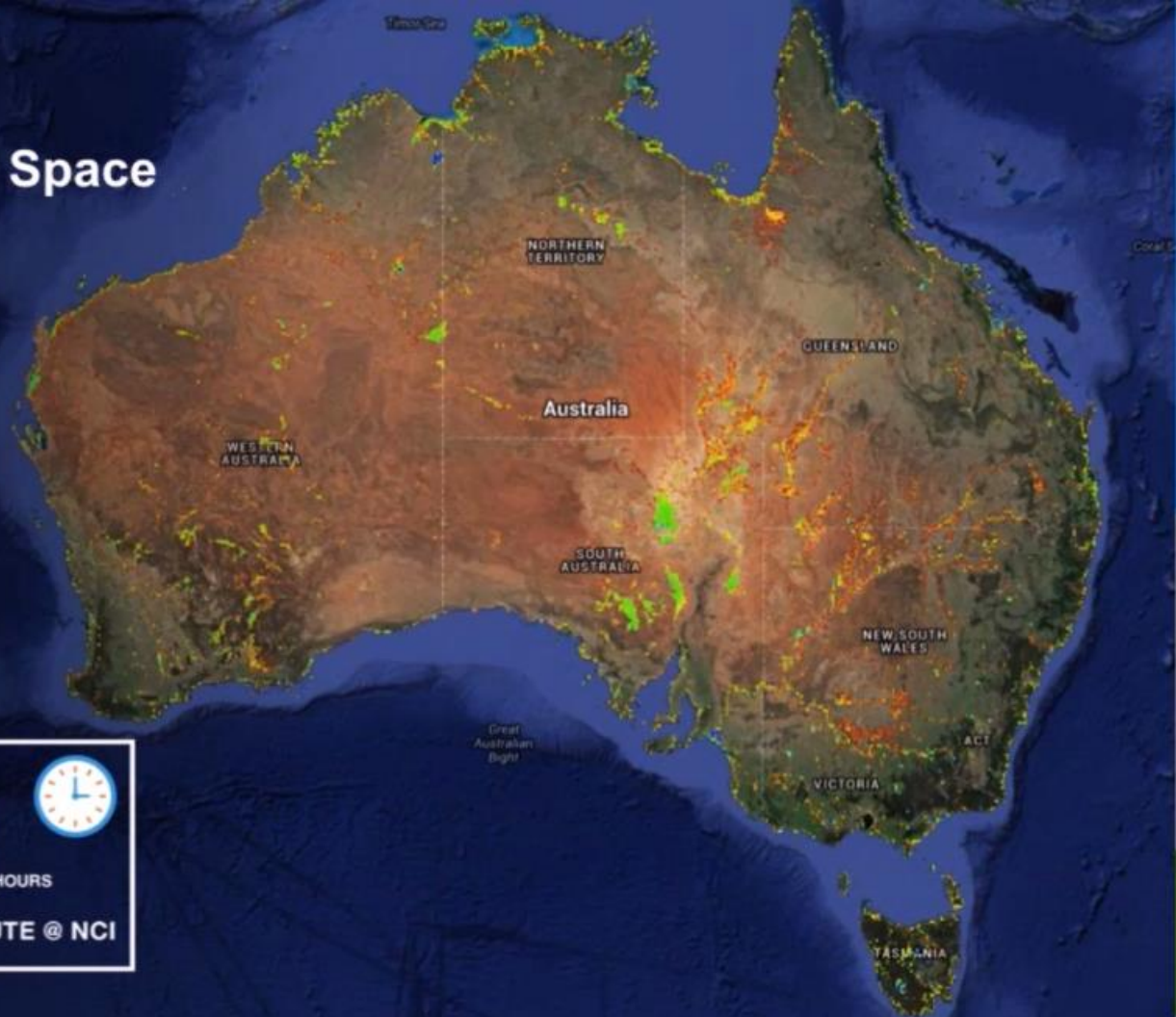
25 
METRE
PIXEL
RESOLUTION

300 000 SCENES
20 000 PASSES

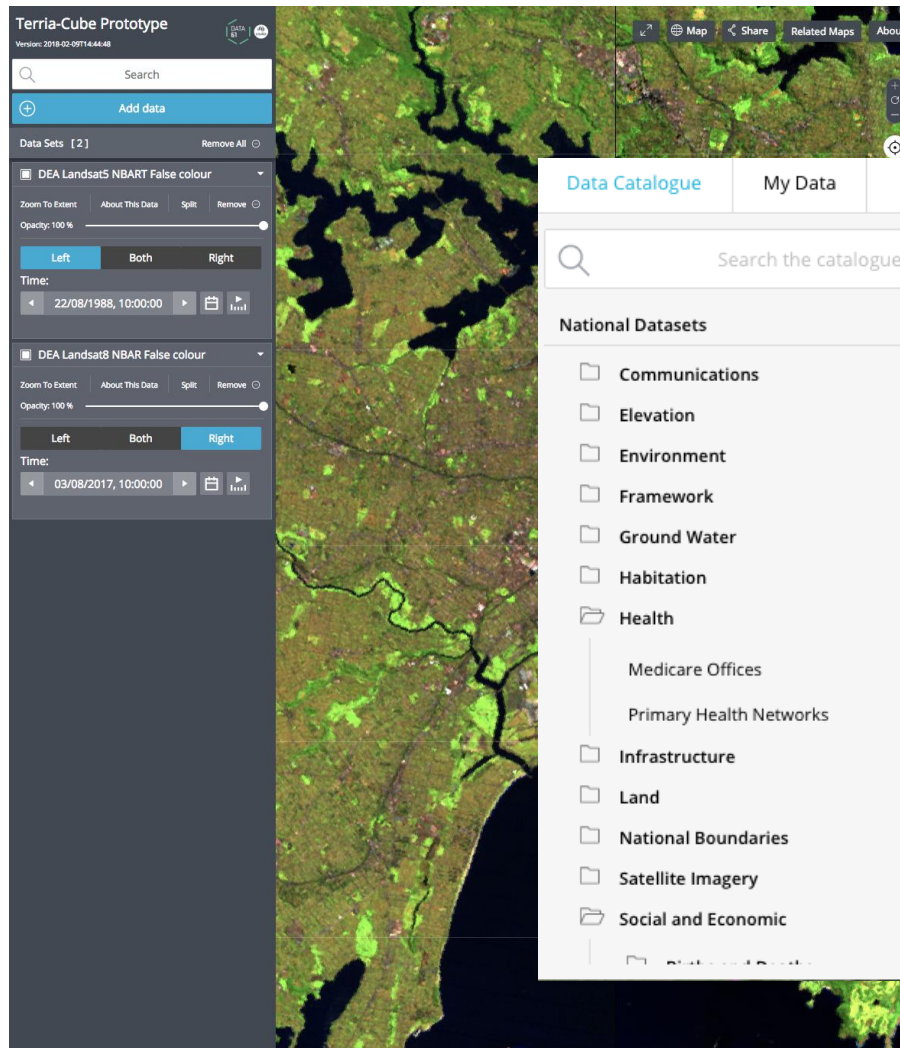
93×10^{12} 
PIXELS

0.75 
PETABYTES

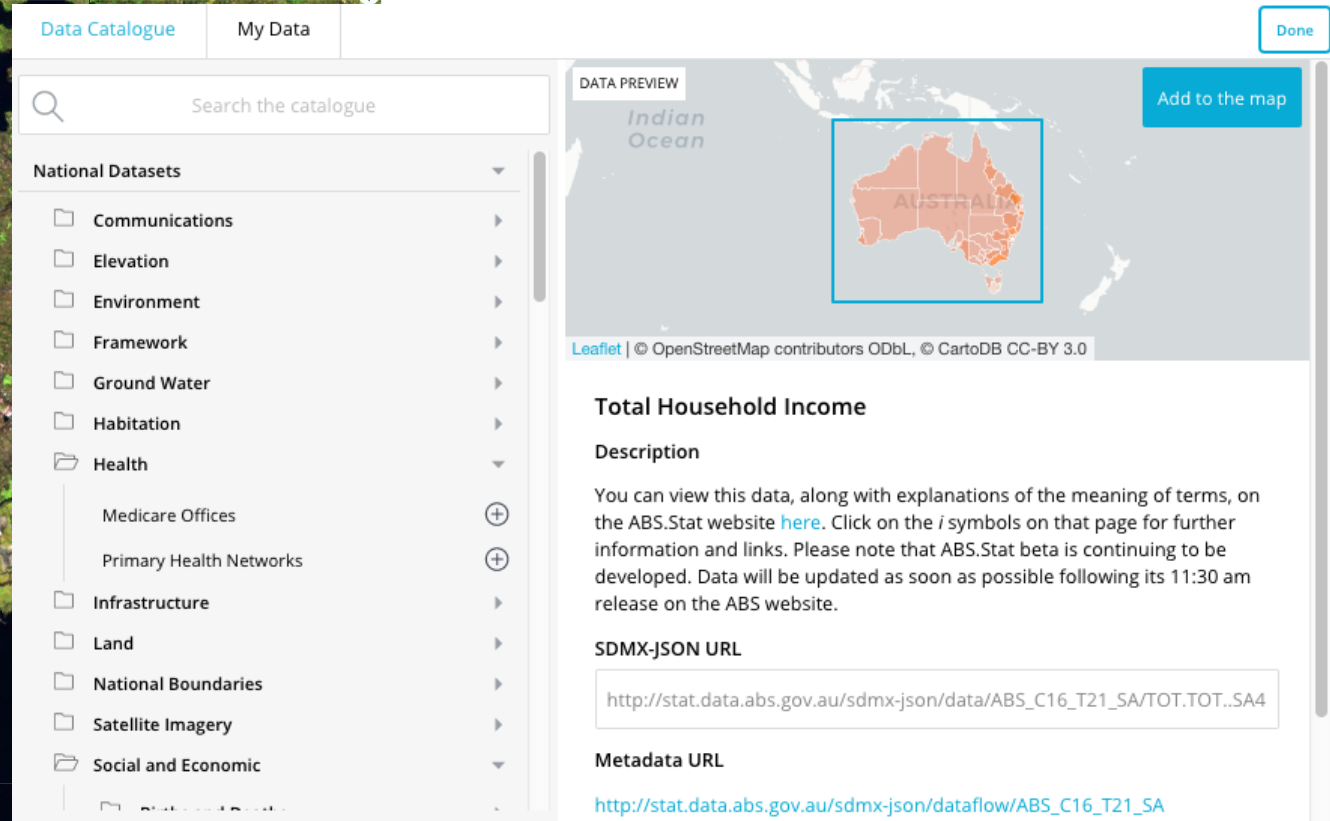
3 
HOURS
COMPUTE @ NCI



GSKY DELIVERING DIGITAL EARTH AUSTRALIA'S 40 YEAR LANDSAT ARCHIVE FOR GEOSCIENCE AUSTRALIA – [HTTP://NATIONALMAP.GOV.AU](http://nationalmap.gov.au)



Dynamic Mapping for Policy



Natural disaster predictions

An international, CSIRO-led collaboration has predicted a tripling in the frequency of natural disasters over the next century—based on greater understanding of the Indian Ocean Dipole, and its impact on the global climate. Improvements to the ACCESS climate model, via NCI, underpin the ability to make significant findings that help inform policy makers.

CSIRO, ARC Centre of Excellence for Climate System Science
& international collaborators



Predicting sea level rise

NCI is assisting the Antarctic Climate and Ecosystem CRC and partners to develop a model spanning the entire Antarctic continental ice shelf at 1km resolution, to predict sea level rise over the next century. Such models will feed into improved weather and climate prediction in the future.



Dr. Ben Galton-Fenzi, The Antarctic Climate and Ecosystems Cooperative Research Centre, The Australian Antarctic Division, & The Antarctic Gateway Partnership

A world-leading genome database

NCI's integrated platform allows Garvan to store and analyse thousands of human genomes, and serve it to clinical and research practitioners. Our co-location of compute and data for genomic medicine is amongst the world's best practice and world-leading.

Garvan Institute's Medical Genome Reference Bank
& the Australian Genomic Health Alliance (NHMRC)

New tools to diagnose auto-immune diseases

NCI provides the computational power to store, analyse and compare the genomes of patients with autoimmune diseases. Diseases like lupus vary significantly at the genetic level, with greater understanding of these variants opening the door to future, targeted treatments.

Professor Carola Vinuesa, ANU
NH&MRC Centre for Research Excellence in personalised immunology

Computer-aided drug design

Using molecular modelling programs on Pawsey Magnus and NCI Raijin, researchers are seeking to understand the origins of Alzheimer's disease and design new drugs that block the functioning of the signalling pathways for such diseases.

Dr. David Wilson, LaTrobe University
Dr Neha Gandhi, QUT



PAWSEY
supercomputing centre



NCI
AUSTRALIA

TODAY'S HPC: A FUSION OF BIG COMPUTE/BIG DATA/EXPERTISE

“Today’s HPC environment has evolved to encompass the needs of big data in addition to its traditional role of computational modelling and simulation.

The contemporary environment comprises tightly-integrated, high-performance infrastructure able to handle the computational and data-intensive workflows of today’s research, together with expertise in computational science, data science and data management.”

2016 National Research Infrastructure Roadmap (Draft) and

NCI submission to Roadmap Capability Issues Paper (September 2016)

- Investment case based on: research excellence/impact/competitiveness and social/environmental/economic benefits to Government and taxpayers
 - **Better Life, Better Environment, More Prosperous Society, Safer Society**



A decorative graphic on the left side of the slide, consisting of a network of thin, light-colored lines and small circles, resembling a circuit board or a stylized tree structure.

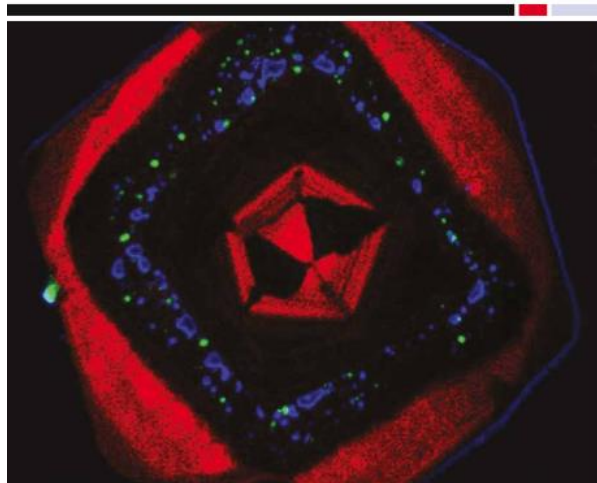
National High-Performance Computing and Data Services for
Australian Research and Innovation:

6. 2016 Roadmap and Future Directions

JUST ANNOUNCED...



2016 NATIONAL RESEARCH INFRASTRUCTURE ROADMAP



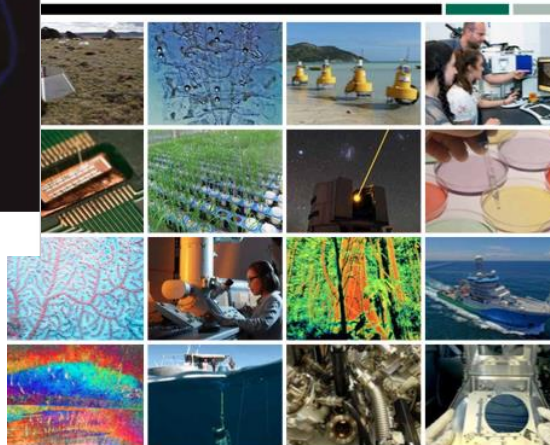
2016 – Urgent case for refunding HPC

- Recommends long term funding for infrastructure
- Need for skill development
- Greater integration of service
- National HPC strategy



FACILITIES FOR THE FUTURE UNDERPINNING AUSTRALIA'S RESEARCH AND INNOVATION

Government Response to the 2016 National Research Infrastructure Roadmap
Research Infrastructure Investment Plan



2018 Capital Investment

- Long term investment \$1.9b /12yr
- 5 year contracts
- Review roadmap every 2 years

In the next 5 years...

- \$70M NCI renewal of HPC
- \$70M Pawsey renewal HPC
- \$72M research platforms
(cloud & storage)



A decorative graphic on the left side of the slide, consisting of a network of thin, light-colored lines and small circles, resembling a circuit board or a neural network, set against a dark green background.

National High-Performance Computing and Data Services for
Australian Research and Innovation:

7. Building Blocks for National Strategy

FUTURE OUTLOOK – BUILDING A NATIONAL SERVICE



- Two facilities:
 - Mitigates a risk of catastrophic failure (e.g., a fire) of either
 - Provide remote backup sites for one another, and dual sites for high-availability services
- Long-term commitment
- Opportunity for Interleaving of replacement (70%) / mid-term upgrade (30%) cycles

The Need

- Insufficient high-level HPC/HPD skills in universities, agencies, industry
 - Dwindling competitiveness rel. advanced economies—NISA context
 - Insufficient expertise to engineer/re-engineer HPC/HPD codes to extract latent performance from new technologies/methodologies for application performance gains—in many cases arising from the failure of Moore's Law
 - Insufficient focus on industry engagement/advancement (cf. UK Hartree)
- Insufficient education/training opportunities for the computational/data scientists needed in future industry, commerce, and R&D environments

The Solution

- NCI and Pawsey Boards adopt a goal of establishing (and co-funding) a **National Hub of Computational/Data Science Expertise**
 - Focus on nationally-significant, program-scale activities (see Roadmap)
 - Comparable in emphasis, but lesser scale, to national labs (USA, UK, ...)
- A shared vision for a critical mass of expertise/skills for:
 - Solutions for R&D challenges—for research/economic competitiveness
 - Enhanced opportunities for participation in international collaborations
 - Vibrant training/education programs in partnership with universities
 - Skilling workforce for an increasingly information/data-rich world in industry, research agencies, medical research, universities, etc.



A decorative graphic on the left side of the slide, consisting of a network of thin, light-colored lines and small circles, resembling a circuit board or a stylized tree structure, set against a dark green background.

National High-Performance Computing and Data Services for
Australian Research and Innovation:

8. Collaboration Opportunities

- Visitor programs and staff exchanges
- Shared resources / access on Australian (NCMAS extension, or Director's shares) and PRACE Tier 0 systems
 - Formalised access for prototyping
- Sharing knowledge and experience:
 - data-intensive computation (integrated environments), etc.
 - Industry linkages
 - Allocation strategies
 - Tools, performance monitoring, porting
 - Governance, planning and future investments (e.g., NCSI), and understanding policy
 - w/shops on operations management; participation in procurements, other exchanges of a technical nature ...

- Opportunities for participation in EU programs, EU Centre of Excellence (BYO funds)
- Contributing to facility/discipline reviews, benchmarking, KPIs, ...
- Sharing expertise in merit-based access programs; external refereeing, external membership of review panels, ...
- Training: shared opportunities for development and participation
Common on-lines courses?

A decorative graphic on the left side of the slide, consisting of a network of thin, light-colored lines and small circles, resembling a circuit board or a stylized tree structure. The lines are vertical and horizontal, with some diagonal branches, and the circles are small and white, scattered along the lines.

Questions ?