NCI
AUSTRALIA

ANNUAL REPORT
2019-2020

INTRODUCING GADI - THE SOUTHERN HEMISPHERE'S FASTEST SUPERCOMPUTER

This report and additional data are available from the NCI website at nci.org.au/about-us/annual-reports

Requests for authorisation and enquiries concerning the contents of the report should be directed to enquiries@nci.org.au

Produced by
National Computational Infrastructure
143 Ward Road, The Australian National University,
Acton, ACT 2601

Designed & Typeset by Result Design.

Cover Image:
Detail from the Gadi painting
celebrating the Gadi supercomputer,
and its Ngunnawal name and connection. The art symbolises the gathering of knowledge from different places and sources,
and the bringing together of an even greater collective knowledge system.

Artist: Lynnice Letty Church

Tribes: Ngunnawal, Wiradjuri & Kamilaroi (ACT and NSW)

We acknowledge the Traditional Custodians of the ACT, the Ngunnawal people. We acknowledge and respect their continuing culture and the contribution they make to the life of this city and this region.

CONTENTS

INTRODUCTION

About NCI

NCI Australia is Australia's National Computational Infrastructure, a nationally unique facility comprising the country's most powerful supercomputer alongside vast data repositories, data management, data infrastructure and data services. The country's pre-eminent computing facility, NCI delivers on the critical national need for high-performance data, storage and computing services. We provide expert services to benefit all domains of science, government and industry.

The Australian Government and the Australian research sector come together through NCI in a broad collaboration involving the largest national science agencies (CSIRO, the Bureau of Meteorology, Geoscience Australia), research universities, the Australian Research Council and industry partners. This collaboration underpins the services that NCI is able to offer the research community: a shared resource that supports nationally significant research in ways that no one organisation could do on its own. As the home of the high-performance computational science of hundreds of research groups across the country, NCI empowers government agencies, universities and industry across multiple domains of research.

The tight integration of supercomputing hardware, data services and expertise that is made available to researchers at NCI enables advanced scientific outcomes as well as rapid response during times of crisis through a critical modelling and information-gathering capability. Altogether, these provide robust insights that inform and benefit public policy. NCI as a national facility makes possible high-impact research and innovation that could not otherwise be achieved in Australia.

Our Mission

NCI's mission is to radically enhance the high-performance computational methods and capabilities available to Australian researchers. By doing this, we add value to Australian society, fulfill the needs of our collaborators and enable transformative ideas. This allows the research sector to provide direct benefit to Australia's industry, environment, policy and society.

We expand the range, scope and ambition of research that can be undertaken in Australia. Making use of globally competitive systems and services available to thousands of computational researchers, now more than ever Australian scientists are producing breathtaking science that benefits and impacts the country. As Australian research becomes ever more reliant on computational methods, a reliable and innovative high-performance computing platform is required. NCI is pushing the boundaries of what such a facility can offer: our integration of high-performance computing (HPC) and high-performance data (HPD) with innovative and powerful data services expands the reach and range of Australian research.

Our goals are for Australia to have a globally competitive HPC capability and for Australian researchers to experience a significant shift in what they can do computationally. Over time, we are expanding the number of researchers using high-end computing in Australia. Most importantly, we are enabling Australian researchers to produce more high-impact discoveries and innovations.

The NCI building at sunset, on the campus of The Australian National University.

Chair's Report

It is my pleasure to introduce the NCI Annual Report for 2019–20. Even without the pandemic and the other crises that impacted Australia so severely this year, 2019–20 would have been a historic and milestone year for NCI. The crises of 2019–20 have shown the importance of NCI not just as research infrastructure but as vital national infrastructure.

For NCI, and indeed Australian research, the most significant moment of the year was undoubtedly the commissioning of the new Gadi supercomputer and, in particular, Gadi's debut – at number 24 – on the TOP500 list of the world's most powerful research supercomputers. Funded by a $70 million investment by the Commonwealth through the Department of Education, Skills and Employment's National Collaborative Research Infrastructure Strategy, Gadi replaced the aging Raijin supercomputer that served Australian research for over 7 years. Supplied by Fujitsu, Gadi continues their long involvement in Australian high-performance computing, dating back to the installation at The Australian National University (ANU) of a VP-50 in 1987.

Acquiring a new supercomputer is not a trivial task. On behalf of the NCI Advisory Board, which oversaw the procurement and installation of Gadi, I would like to thank all involved for their commitment and professionalism. The sheer physical scale of the installation was captured in an impressive video that I strongly recommend, available at bit.ly/gaditimelapse. Significantly for Australian research, this massive task did not impact the continued operation of Raijin.

Gadi, meaning 'to search for' in the language of the Ngunnawal people, the Traditional Owners of the Canberra region, came online just in time to help respond to the unprecedented national crises of 2020. Firstly, the Bureau of Meteorology used Gadi's enhanced weather modelling capacity to support its emergency response to the bushfires over the 2019–20 summer. Then, as Australian science responded to the COVID-19 pandemic, Gadi enabled important protein-modelling research assisting drug and vaccine design. I have little doubt that as Australia recovers from the pandemic, Gadi will continue to play a significant and critical role. The Commonwealth's investment will be well repaid.

In addition to overseeing the installation of Gadi, the Advisory Board dealt with several other significant issues, including the renewal of the NCI Collaboration Agreement and the development of a formal Strategic Plan for NCI. The Collaboration Agreement underpins NCI's governance and

operational funding. It is pleasing that the renewal affirms the strong commitment of NCI's founding collaborators – CSIRO, ANU, the Bureau of Meteorology and Geoscience Australia. In February 2020, the Board held a strategic workshop day with NCI management to discuss the strategic and policy aspects of NCI's engagement with the national research sector in light of the dramatically enhanced capabilities of Gadi. The outcomes of that day ultimately translated into NCI's new Strategic Plan, approved by the Advisory Board in June and now publicly available at bit.ly/NCIstrategicplan.

As summarised in the Strategic Plan, NCI now moves with confidence – yet not without challenges – into a new and rapidly evolving research environment. How Australia recovers from the pandemic and, indeed, what that 'new world' will look like is, in September 2020, still unclear. What is certain is that NCI and the research that Gadi enables will play a critical role.

On a personal note, I would like to thank my Board colleagues for their commitment to NCI and the vision of NCI as Australia's most advanced and integrated high-performance computing and data facility.

I commend the 2019–20 report to you. It is a record of a historic year for NCI but also testimony to the creativity and innovation of the Australian research community that is enabled by NCI.

Emeritus Professor Michael Barber AO FAA FTSE – Chair, NCI Advisory Board.

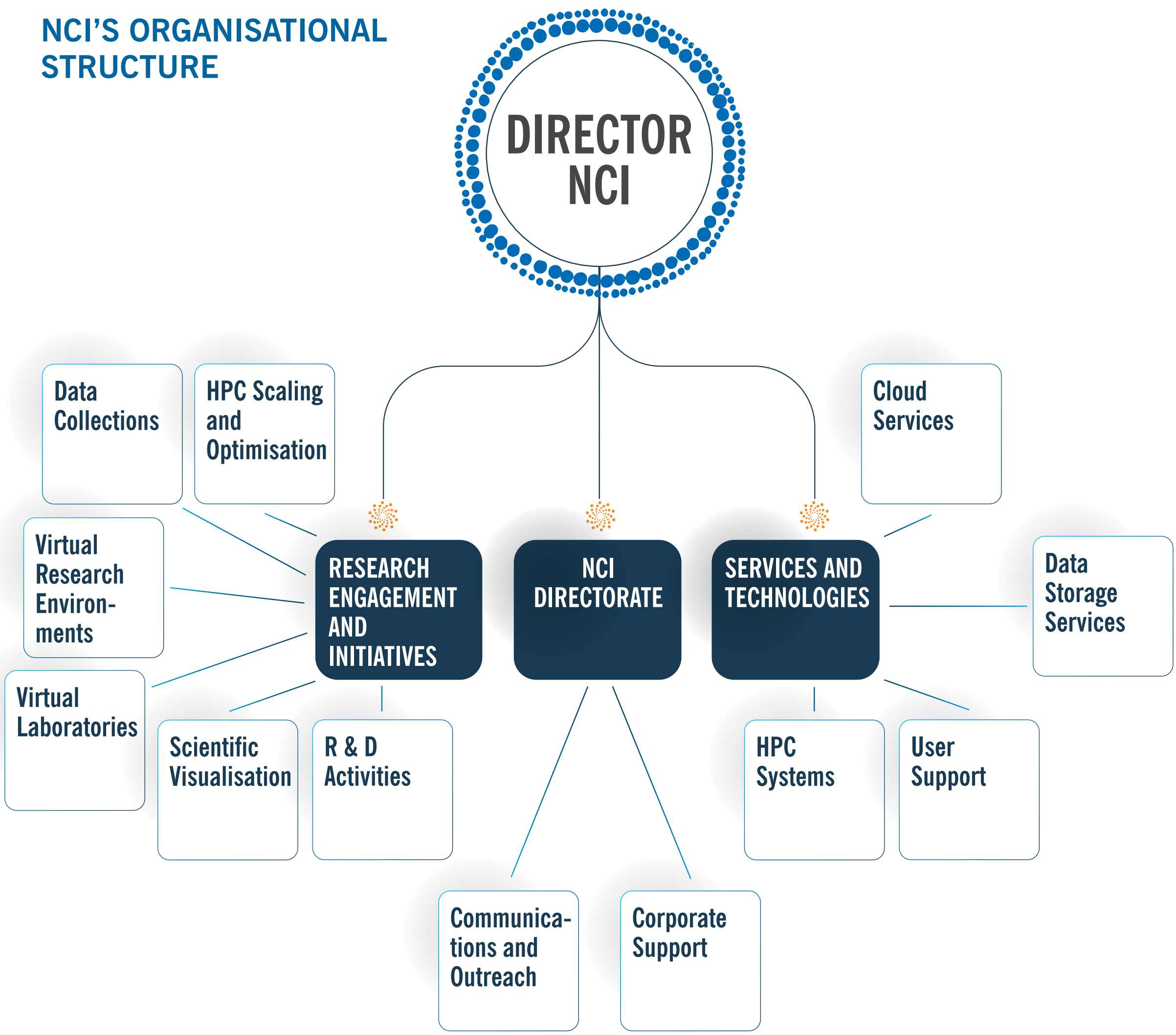
NCI Governance

THE NCI ADVISORY BOARD

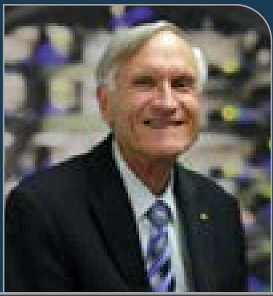
NCI is governed by The Australian National University on the advice of the NCI Advisory Board, which comprises:

> an independent chair appointed by the Advisory Board,
> the Director of NCI,
> one member appointed by each of the Major Collaborators, and
> additional independent board members appointed by the NCI Advisory Board on the basis of their expertise.

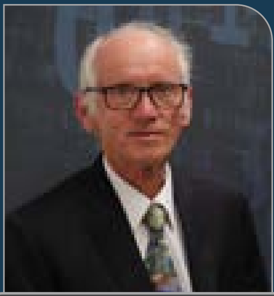
NCI'S ORGANISATIONAL STRUCTURE



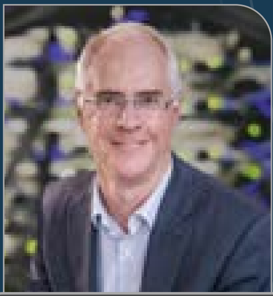
ADVISORY BOARD MEMBERS



Emeritus Professor Michael Barber
AO FAA FTSE
Chair



Emeritus Professor Robin Stanton
Deputy Chair



Professor Keith Nugent
FAA
The Australian National University, Deputy Vice-Chancellor (Research and Innovation)

Professor Sean Smith
FAAAS
NCI, Director

Dr Gilbert Brunet
FCMOS
Bureau of Meteorology, Chief Scientist and Group Executive (Science and Innovation)

Dr James Johnson
Geoscience Australia, Chief Executive Officer.

Dr David Williams
CSIRO, Executive Director, National Facilities and Collections

Dr Simone Richter
ANSTO, Group Executive (Nuclear Science & Technology and Landmark Infrastructure)

Ms Susan Wilson
Independent Member, Founder and Managing Director, Bounce Partners

Director's Report

Colleagues and friends,

What a momentous year 2019–20 has been! Federal funding for a major refresh of NCI's high-performance computing (HPC) engine was announced in 2017. Following an extensive and fiercely competitive procurement process, in July 2019 the NCI Advisory Board approved the preferred vendor bid, led by Fujitsu Australia with a stellar supporting cast of HPC infrastructure providers. This past year, then, the hard work shifted from procurement to logistical planning and physical implementation. Gadi was launched in November 2019 while its predecessor, Raijin, was still running. After the transition of users across to Gadi, the venerable Raijin supercomputer (lauded elsewhere in this report) was stood down and the full Gadi system was fired up, clocking in at 9.3 petaflops which landed it at #24 on the global TOP500 ranking in June 2020. With considerable pride and also a degree of nostalgia, we noted that Gadi secured the very same global ranking Raijin had attained upon its launch, 8 years ago.

We were honoured to be able to christen Australia's most powerful supercomputer with the name Gadi – 'to search for' in the language of the Ngunnawal people on whose lands NCI is located; and equally honoured to highlight its signature artwork featured on the front cover of this report, commissioned from local artist Lynnice Church.

For Gadi, this year has been a baptism of fire in more than one sense. It arrived in the nick of time as NCI's environmental modelling users – including, prominently, the Bureau of Meteorology – rapidly pivoted to address the summer bushfire crisis. The ACCESS climate models developed by CSIRO and the ARC Centre for Climate Extremes, heavily used during the bushfire season, were seamlessly transitioned from Raijin to Gadi with confirmed reproducibility and instantly put to work.

Meanwhile, NCI also worked with a select set of peak HPC projects and communities to stress-test as many performance aspects of the machine as possible with high-performance workloads. Aside from helping the NCI teams iron out performance issues on the new machine, significant scientific results are now emerging from that early program.

Gadi – as a peak national facility – fields a very wide range of national research. Amongst that array of efforts, NCI demonstrates the kinds of achievements that could not possibly be reached in a dispersed, federated compute environment. With this in mind, NCI introduced the Australasian Leadership Computing Grants (ALCG) in a competitive call that was initiated early in 2020. The 4 selected projects, collectively using around 180 million hours worth of computing time between July 2020 and June 2021, are now up and running.

Simultaneously with the ALCG call, it became apparent that a response to an urgent crisis – the COVID-19 pandemic – was needed. Hence, an auxiliary ALCG call was made for research addressing COVID-19 treatment or prevention, with a specialist selection committee allocating a further 40 million hours to 3 outstanding proposals. These projects will continue to run during the next year from July 2020 to June 2021.

With the present and forthcoming upgrades at both NCI and the Pawsey Supercomputing Centre, it was timely to run, in November 2019, the inaugural Australasian Leadership Computing Symposium (ALCS) together with our New Zealand counterpart, the New Zealand eSciences Infrastructure. ALCS was a national high-performance computing and data showcase, research exchange and user community consultation forum with associated training activities that ran over

4 days at the Australian Academy of Science's Shine Dome. The meeting was a resounding success and provided a treasure trove of insights into national and international developments, research highlights and community input.

NCI is internationally recognised for its leadership role in data management and availability, robust FAIR data releases, and data analysis capabilities around ultra-large datasets of critical national significance. Now featuring prominently amongst these are the latest CMIP6 climate datasets as part of our Earth Systems Grid Federation node, and a huge collaborative achievement of local and international teams. We continue to acquire and host further international datasets across earth-observation, geophysics and other disciplines in support of our Australian science research partners.

This truly has been a momentous year for NCI and a significant boost to national research infrastructure with the availability of Gadi. I am immensely proud of the efforts of the teams here at NCI during very trying times to deliver this capability for Australian researchers. The challenges come thick and fast – not least addressing the needs of research communities such as the life sciences new to the HPC sector with contrasting compute and data usage profiles and paradigms. We undertake these challenges as a step up, recognising with gratitude the federal NCRIS funding through the Department of Education, Skills and Employment, and the sustained commitment from our major collaborating organisations that has enabled NCI to reach – and now to build from – its current impressive capability.

Professor Sean Smith, FAAS Director, NCI

Highlights of the Year

International Supercomputing Conference Frankurt

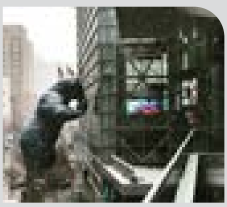
Australasian Leadership Computing Symposium

NCI announces ALCG grants totalling 220 million units of computing time on Gadi

2019: J A S O N D — 2020: J F M A M J

First users on Gadi

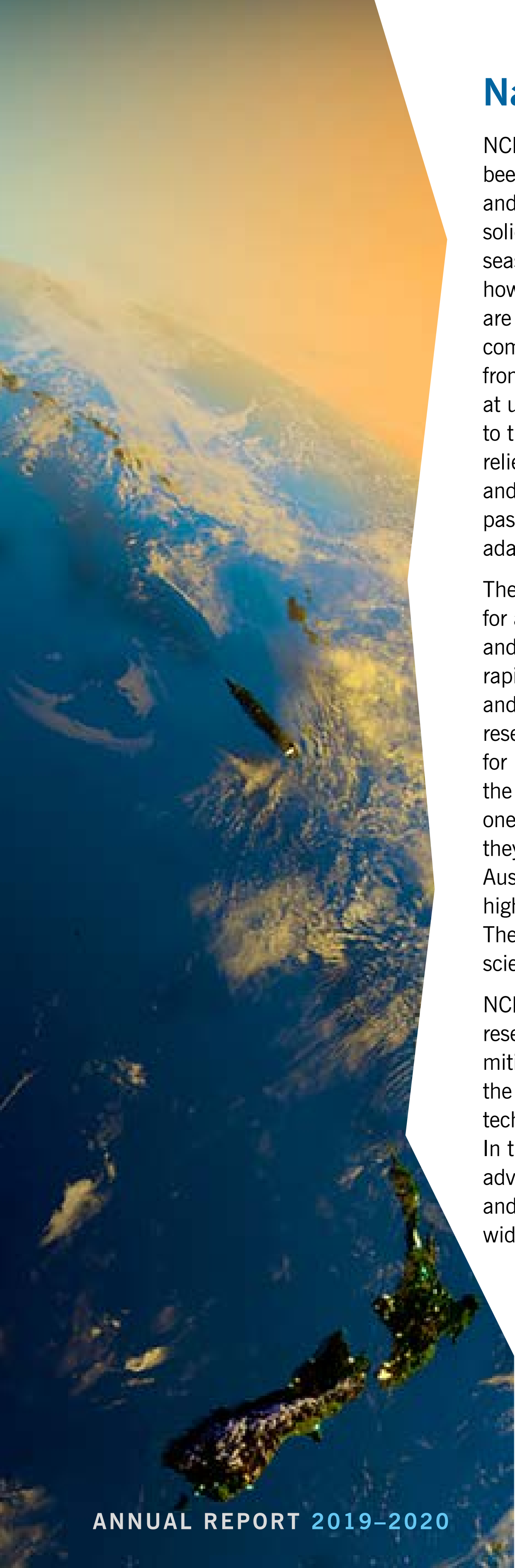
Super Computing Conference Denver



Gadi announced as 24 ranked supercomputer in the world

1

A CENTRAL ENABLER IN A NATIONAL RESEARCH ECOSYSTEM



National Benefits

NCI's contribution to nationally beneficial research has been amplified this year, with multiple environmental and medical crises requiring clear scientific focus and a solid underlying infrastructure. The devastating bushfire season over the 2019–20 summer highlighted just how important short and long-term weather predictions are to the safety and preparedness of Australian communities. Dedicated bushfire modelling activities from the Bureau of Meteorology, CSIRO and researchers at universities all over the country contributed directly to the Australian response over the summer, and these relied in large part on NCI's offering of supercomputing and big data services. Those contributions continue past the summer, with predictions, preparation and adaptation for future fire seasons already taking place.

The COVID-19 crisis overshadowed much of 2020 for all of us, but throughout it all NCI stood strong and opened up new computing resources for specific, rapid-response coronavirus research. It was heartening and powerful to see the variety and breadth of scientific research coming out of NCI with possible solutions for improved diagnosis, treatment and prevention of the disease. Timely support for national priorities is one of NCI's reasons for being. Before crises, after they are long gone, and at every moment in between, Australia's scientific community relies on access to high-performance computing and data infrastructure. The national benefits of cutting-edge, reliable, powerful scientific infrastructure are immense.

NCI provides the foundation for some of the key research helping us predict environmental disasters, mitigate the impact of deadly diseases, respond to the needs of agricultural communities and invent the technologies that will power industries in the future. In the areas of climate and earth-observation, NCI's advances in data collections management, data sharing and computational processing are recognised world-wide. The benefits to the nation are immense.

SUPERCOMPUTER FIRE MODELLING

Australia's need for a powerful and predictive bushfire modelling capability has never been clearer. The 2019–20 bushfire season showed just how widespread, unpredictable and terrifying bushfires can be. The Bureau of Meteorology, with the help of the NCI supercomputer, is developing new modelling tools that are already helping authorities deepen their understanding of fires and their behaviour.

ACCESS-Fire is the model that the Bureau of Meteorology is developing to simulate bushfires and their interaction with the atmosphere. State and Territory fire agencies and the Bushfire and Natural Hazards Centre for Collaborative Research have been key players in this project since its inception.

Fire agencies currently run models in real time to predict how fire may spread. Designed to run on laptops out in the field, those models capture a limited view of the complex structure of the atmosphere. Models run on a supercomputer however, can incorporate a myriad of complex information. ACCESS-Fire includes physics that models the heat, moisture and air flows in and around a fire front at landscape scale.

Taking observations of weather conditions inside a fire is understandably difficult. Wind speed, wind direction, air pressure and air temperature are all key variables that we can only really understand through large, high-resolution computational models. Recent improvements to the ACCESS-Fire model's scientific detail, as well as its computational performance, make it possible to find out new details about bushfires and their propagation. For example, Bureau of Meteorology researchers Drs Mika Peace and Jeff Kepert identified the specific set of weather conditions that led to an unexpected ember attack during the Waroona fire in Western Australia in 2016. With this new knowledge, fire fighters can better plan their tactics and the public safety response during future fires.

The more complex structure of the ACCESS-Fire model makes it an improved tool for dealing with fire extremes and unexpected situations, such as when fires become very large and start creating thunderstorms with the heat and moisture they produce. Supercomputing performance and carefully designed model codes make understanding these phenomena possible.

This model relies on all facets of the Gadi supercomputer: many performant processors, large memory, and fast and powerful storage. Due to the complexity of the model it is not yet suitable for operational use during a bushfire, but the researchers are gradually moving in that direction. Dr Peace says, ‘It’s wonderful to get to work on something that resonates with everyone across our society. Bushfires affect us all, so contributing to that is both exciting and fascinating.’

GLOBAL CLIMATE MODELS AND DATA ANALYSIS

Climate change is a global issue, and understanding its progression and impact requires global science. The amount of scenario-modelling and data analysis required is immense. That modelling produces hundreds of simulations covering all parts of the world to help scientists reach rigorous results about future climate changes. For this reason, climate-modelling datasets are shared worldwide through a global federation of sites, the Earth Systems Grid Federation, of which NCI is one of only a few Tier-1 data nodes (see Case Study *Global Climate Infrastructure* on pages 62–63 for more information about the data management infrastructure supporting this data). Scientists use these datasets to understand in more detail the probabilities associated with various climate outcomes – changes in regional rainfall, hot and cold days, and dangerous weather extremes.

In Australia, hundreds of scientists across a range of research organisations and disciplines are using climate model data to increase our understanding of the regional effects of climate change. A recent focus has been the sixth iteration of the Coupled Model Intercomparison Project (CMIP6), with research in Australia being led by the CSIRO. This generation of data builds on the previous generation CMIP5 from 2013, but at higher resolutions and bigger data volumes. The increases are in part due to improved modelling at regional scales, as well as more detailed integration of various emissions scenarios.

A cluster of national collaborations are underway between researchers from both the university sector and government science agencies, who are currently exploring the information contained within this data. For example, one study by Dr Michael Grose and colleagues shows the potential for greater and earlier temperature increases for Australia, while Dr Anna Ukkola and colleagues' analyses of the CMIP6 data confirm the CMIP5 projections of a significant decrease in rainfall over Southern Australia.

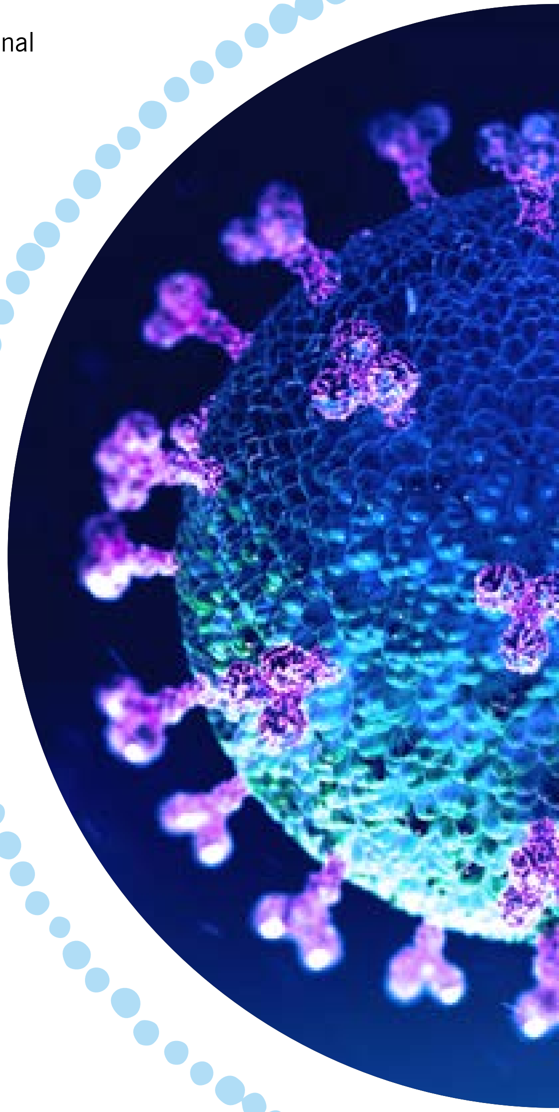
These complex models, refined over decades, produce data that describes these possible future states of the climate. As the whole scientific effort around climate data is international, so are the models. Australia's major contribution comes through the climate variant of our national weather prediction model – ACCESS, the Australian Community Climate and Earth System Simulator. ACCESS is developed collaboratively by the Bureau of Meteorology, CSIRO, the Australian Research Council Centre of Excellence for Climate Extremes, NCI and international collaborators. Scaling from daily weather forecasts to 100-year climate predictions, and from city-scale resolution out to the entire globe, ACCESS and the web of climate research that it supports are some of the most complex tools of Australian science, underpinning both our short- and long-term productivity.

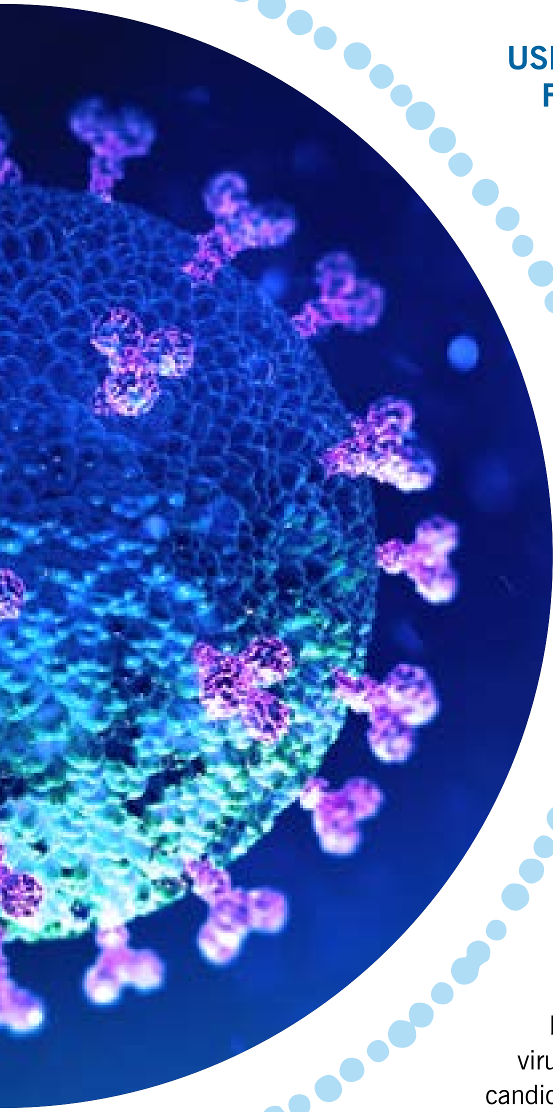
Australian researchers play a large role in global climate-modelling activity, contributing in particular to improvements in the global models and how they represent the local region. Researchers can now run their models on the Gadi supercomputer and use the reference data collections and tools in a high-performance computing environment that supports data-intensive analysis. Together, these enable Australia's steps in understanding future climate impacts, leading the way for social and policy responses based on the latest, most up-to-date information.

COVID-19 Research Highlight

The 2020 coronavirus pandemic highlighted the role of national science infrastructure in responding to global challenges. NCI continued operating throughout the crisis, providing an avenue for researchers to continue conducting their experiments while laboratories and field sites were shut down. We also provided rapid support for those members of our user community able to contribute to the scientific response to COVID-19. Through an expedited merit scheme supported by the Australian Government's NCRIS program, our Australasian Leadership Computing Grants COVID-19 Special Call provided large computing resource allocations to 3 research groups around the country.

Marked by their strong scientific and technical merit, by mid-April 2020 these research groups started simulating protein structures, human cell receptors and more, with an eye to drug and vaccine design. Other research groups, supported by universities and medical research institutes, also took up the challenge, coming at it from various perspectives including genome sequencing and evolution tracking. In total, NCI provided around 40 million units of computing time to 3 different COVID-19 research projects in the first half of the year. Their research is still ongoing, but already some findings have made their way into the hands of medical practitioners and are helping make our populations safer once more.





ALCG Special Call Recipients

USING LARGE-SCALE MOLECULAR DYNAMICS FOR RATIONAL DRUG DESIGN

Associate Professor Megan O'Mara, Dr Katie Wilson, Dr Stephen Fairweather – The Australian National University

This research uses simulations of the approximately 800,000 atoms that make up a key receptor of the human body to understand exactly how the coronavirus uses it to invade human cells. It is only with high-resolution modelling that accurately replicates the true behaviours of these receptors that we can figure out where vulnerabilities in the virus's binding process are. Targeting the interaction between the human receptors and the coronavirus binding protein might well be a useful direction for drug design.

STRUCTURE-BASED DRUG DISCOVERY

Associate Professor Michael Inouye, Dr Sergio Ruiz Carmona – Baker Heart and Diabetes Institute

This project aims to model existing or novel drugs, along with drug binding sites, that could be used in treating COVID-19. As a first step, more than 7 million existing drug compounds from a global repository will be modelled to find ones that may inhibit a key step in the virus replication process. The researchers will focus in on likely candidates through a sequence of molecular dynamics simulations. Expanding further to include more proteins involved in the virus binding process, the researchers will also analyse the complex interactions between the virus and human proteins.

TARGETING STRUCTURAL TRANSITIONS IN THE COVID FUSION PROTEIN

Professor Alan E. Mark, Dr Martin Stroet, Ms Shelley Barfoot – University of Queensland

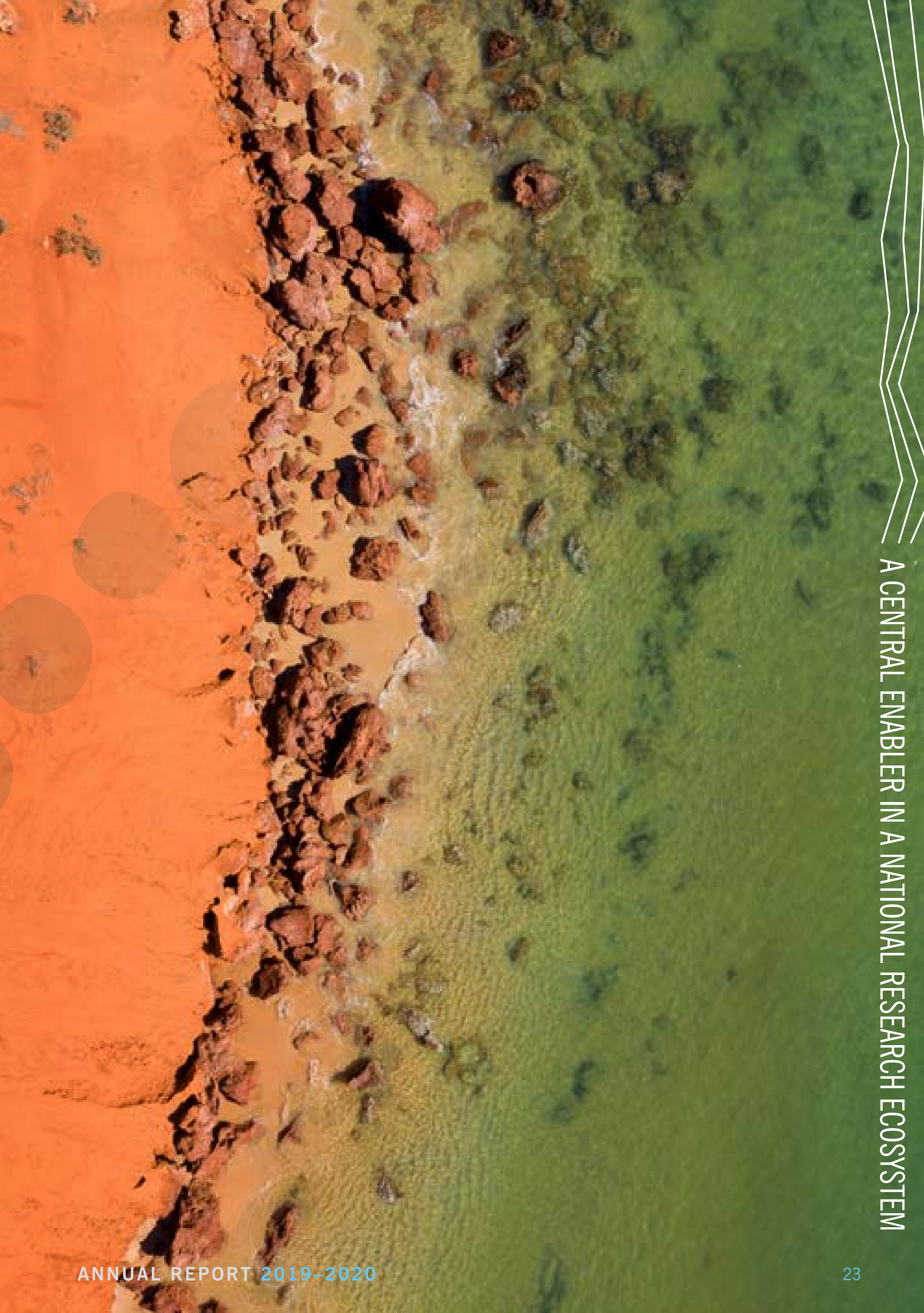
By modelling key structural changes in the protein that enables a coronavirus to enter and infect a human cell, this research project aims to throw light on a critical target for both vaccine development and the discovery of antiviral agents. In particular, working with researchers from one of 3 centres worldwide charged with rapid vaccine development by the World Health Organization, the aim is to understand how potential vaccine constructs that incorporate the coronavirus fusion protein can be optimised. Professor Mark and his team will also use the national supercomputer facilities to assist in the global effort to identify existing drugs that could be repurposed to treat COVID-19.

Impactful Science

NCI's users are some of Australia's leading researchers, conducting research across all domains of science, from the galactic to the microscopic. Enabled by the region's most powerful supercomputer and its large repository of national research data, Australia's scientists are able to produce science that makes a difference and creates an impact on our everyday lives. The Garvan Medical Research Institute's 4,000-strong database of whole human genomes, reprocessed in April 2020 using Gadi's improved parallel performance, marks a major milestone for Australian genomic medicine. Computed over the course of one week, this entire archive of genomic data from healthy seniors is an invaluable baseline for further genomic investigations into cancer, genetic illnesses and autoimmune diseases.

The discoveries and scientific advances made using NCI's tightly integrated compute and data systems yield important results now and into the future. This includes leading science in the realm of fluid flow simulations, the development of new energy storage materials, the modelling of water flows through our river systems and much more. The impacts of the work done at NCI are felt across all disciplines of science and benefit Australia's health, industry, agriculture, environment, tourism and more.

In many cases, the ability to create science impact is extended by the confluence of high-performance computing, large data repositories, data analysis tools and services and expert support, all available in one centre. The tight integration of all these elements within the NCI environment radically simplifies and accelerates the research process. In all, this enables the most powerful and ambitious scientific projects to proceed on an equal footing with international collaborators. From the medical professionals finding new ways of treating rare genetic diseases to the environmental scientists monitoring the health of our waterways and forests, NCI enables science that impacts us all.



A CENTRAL ENABLER IN A NATIONAL RESEARCH ECOSYSTEM

REPROCESSING A HIGHLY VALUABLE GENOMIC DATASET

Genetic medicine in coming years will benefit from the technical innovations that NCI and the Gadi supercomputer are now enabling. The Garvan Institute of Medical Research is refreshing its 4,000-strong database of genomes from healthy seniors – the Medical Genome Reference Bank (MGRB). Free from the markers of heart disease, cancer and diabetes, these healthy genomes form a robust comparison dataset for sequenced patient genomes that doctors might have. Initially processed and stored at NCI in smaller batches, researchers and NCI staff reached a milestone in 2017 by processing 1,000 of the genomes simultaneously over the course of a single night. Now, advances in the complex software controlling the hundreds of steps required have enabled reprocessing of all 4,012 MGRB genomes in one go (see Case Study *Enabling the next generation of genomic medicine* on pages 54–55 for more information about the software advances).

Reprocessing using the latest scientific understanding about human genomes allows researchers to have the most accurate data at hand for diagnosis and treatment. As genetic medicine becomes more central to our medical process, being able to compare patient test results with the rigorous baseline set out by the MGRB will be a key factor. Doctors and clinicians will be able to reach diagnoses much quicker, especially, we hope, for rare genetic diseases with complex and debilitating symptoms.

Reaching this number of processed genomes takes a lot of scientific effort. The sequencing data from thousands of genomes is transferred from Garvan's Sydney laboratory to NCI in Canberra, at which point the reprocessing begins. Thousands of tiny snippets of sequences are compared, lined up and combined into a long string making up the entire human genome. From the snippets to the final sequence takes more than 40 different computational steps, all guided by the expertise of the bioinformaticians and programmers who built the software.

The final product is a treasure trove of valuable genomic data. The entire MGRB is securely shared with approved Australian and international genome researchers. The medical benefits to come from this modern dataset, built using the Gadi supercomputer's performance and filesystem speed, are only beginning to be realised.

A CENTRAL ENABLER IN A NATIONAL RESEARCH ECOSYSTEM

Meritorious Access to NCI

As the nation's largest and most integrated computational science facility, NCI attracts and provides resources for the most ambitious computational researchers in Australia. Through 4 different streams of allocations in 2019–20, NCI allocated more than 510 million units of computing time to deserving researchers. Peer-reviewed and allocated through a stringent, community-led allocation process, the successful proposals are richly deserving for both their scientific excellence and their national significance. NCI's merit allocation schemes are supported by annual investments from the Commonwealth Department of Education, Skills and Employment's National Collaborative Research Infrastructure Strategy (NCRIS) program, the Australian Research Council (ARC) and the support of our collaborators.

The long-standing National Computational Merit Allocation Scheme, together with partner centres the Pawsey Supercomputing Centre, the Monash University eResearch Centre and the University of Queensland Research Computing Centre, allocated this year more than 280 million units of computing time across 154 projects on Gadi. These include vital support for some of the country's biggest users in oceanography, particle physics and materials science, as well as introductory grants for Early Career Researchers looking to gain experience using Australia's biggest supercomputer.

This year saw the introduction of the Australasian Leadership Computing Grants, a high-end scheme catering to the biggest and most experienced computational research projects. Enabled by the new capacity of the Gadi system, the grants provide unprecedented computational resources to some of the country's leading researchers in areas such as fluid dynamics, climate modelling

and astrophysics. They will spend 2020 conducting some of the highest resolution simulations of combustion, and of short- and long-term climate and galaxy formation ever attempted in Australia. The scale of these grants, a total of 180 million units of computing time, and the ambitious science they engender, is only possible due to the installation of Australia's new Gadi supercomputer. Providing next-generation performance and capabilities to all of our science community, it pushes forward the impact of Australian science.

The tragic circumstance of COVID-19 in early 2020 led to the creation of a rapid-response scheme, the ALCG COVID-19 Special Call, which funded meritorious research specifically designed to assist with the global science effort looking for potential vaccines or drugs to treat the disease. NCI and the Pawsey Supercomputing Centre were together perfectly placed to provide computational support for vital and time-critical research in this area. Three groups of researchers with biomolecular modelling expertise and a variety of approaches to finding treatments were awarded between 12 and 15 million units of computing time each.

The fourth merit access scheme at NCI is the Flagship Scheme, supporting the excellent work of ARC Centres of Excellence. The current crop of Flagship Scheme recipients includes Centres of Excellence specialising in astrophysics, climate research, materials science and fundamental physics. These highly meritorious centres receive 110 million units of computing time with the support of both NCRIS and the ARC. Supporting the Centres of Excellence over their entire 7-year operating life, the Flagship Scheme constitutes a key enabling support for the important work of the Centres.

EXTREME SCALE COMBUSTION SIMULATIONS

The ability to create and control fire was one of the earliest human technologies. For more than a million years we have been using flames to cook and power our lives. In the last few hundred years, fire and controlled flame have been used for powering transport and electricity generation. Yet for all that time, the precise mechanisms of how flames form and grow inside engines are still uncertain. This is due to strong interactions between turbulent fluid flow and the combustion chemistry inside the engines, which occur across a huge range of scales – from tens of centimetres down to hundreds of nanometres.

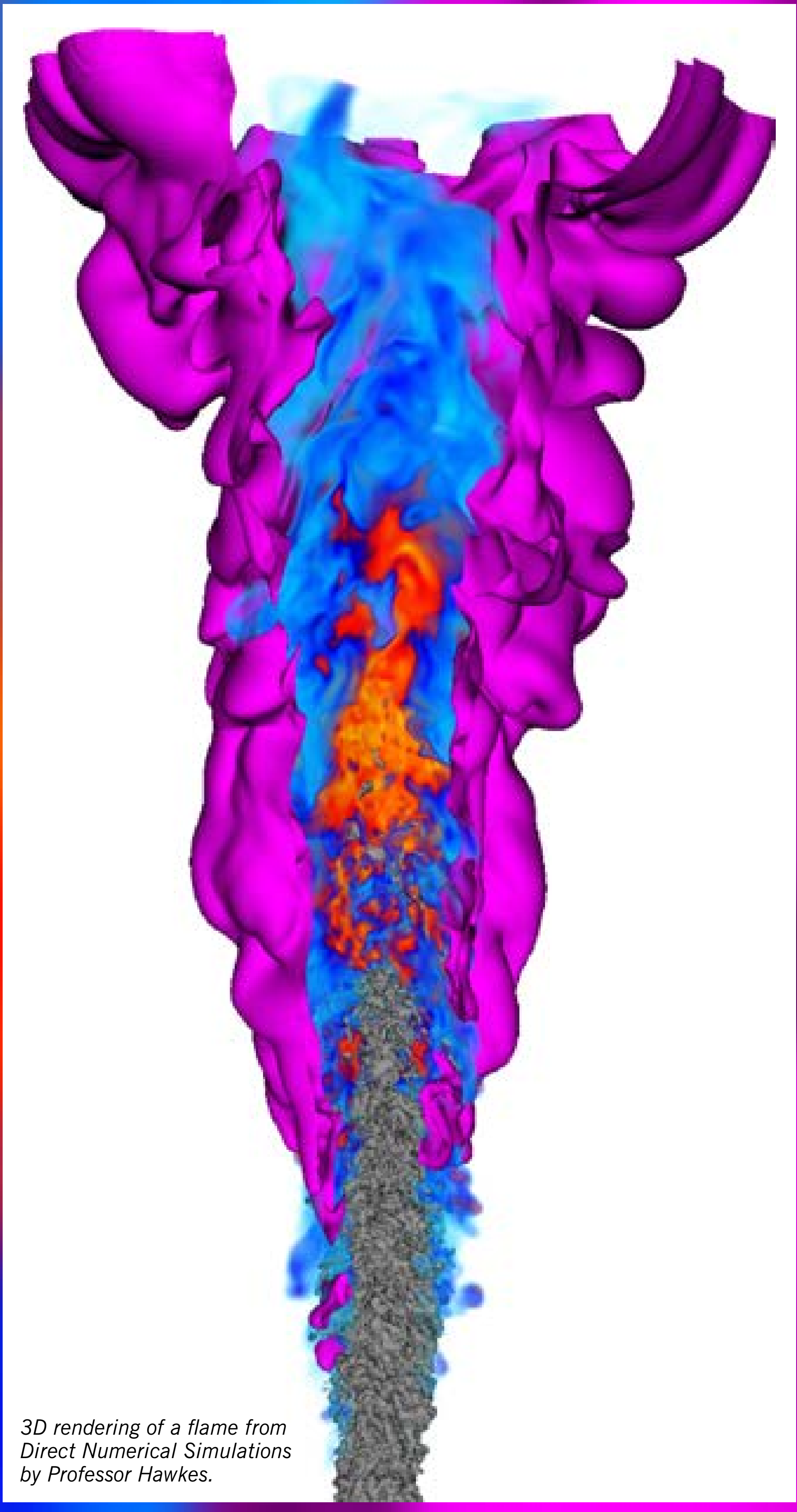
Almost impossible to physically measure with any level of detail, observing flames in full detail in engine conditions requires computational modelling. The most accurate modelling is through Direct Numerical Simulation (DNS). Researchers use all of the physical and chemical equations that govern the reactions and interactions of the molecules involved to simulate each step of flame formation, giving exquisite, physically accurate detail in the results. DNS comes at a cost: the calculations required are extremely intensive, taking millions of hours of computing to produce the result.

This research looking into both diesel and hydrogen combustion is conducted by Professor Evatt Hawkes from the University of New South Wales and colleagues, supported by meritorious grant schemes including the National Computational Merit Allocation Scheme and the Australasian Leadership Computing Grants, as well as early-access allocations on the Gadi supercomputer. The team requires some of the largest computational resource grants in the country to do their simulations. Recent simulations required a grid with more than 1 billion points, evolving through 10,000 time steps for more than 50 variables each.

Depending on the software, each simulation can use up to 20,000 processors at a time, with the code highly tuned for running at the largest scales. Some of the software can adjust the spacing and resolution of the grid it is working with to get the highest performance and best results from the simulation. As the flame being simulated forms, grows and expands into the entire grid, the highest resolution portions of the grid expand to suit it.

Recent research efforts have looked into the exact characteristics of diesel flames as they form inside large engines. More and more, the researchers are turning to hydrogen as a combustible fuel source of the future. While the software is the same, the scientific processes of combustion are quite different for hydrogen compared with fossil fuels. The next generation of engines will use new kinds of fuels and new combustion concepts to power and transport our world.

In coming years, the research team will continue to develop and improve their simulations of hydrogen combustion, learning from the experience about how we should design the future engines of ships and electricity generators.



3D rendering of a flame from Direct Numerical Simulations by Professor Hawkes.

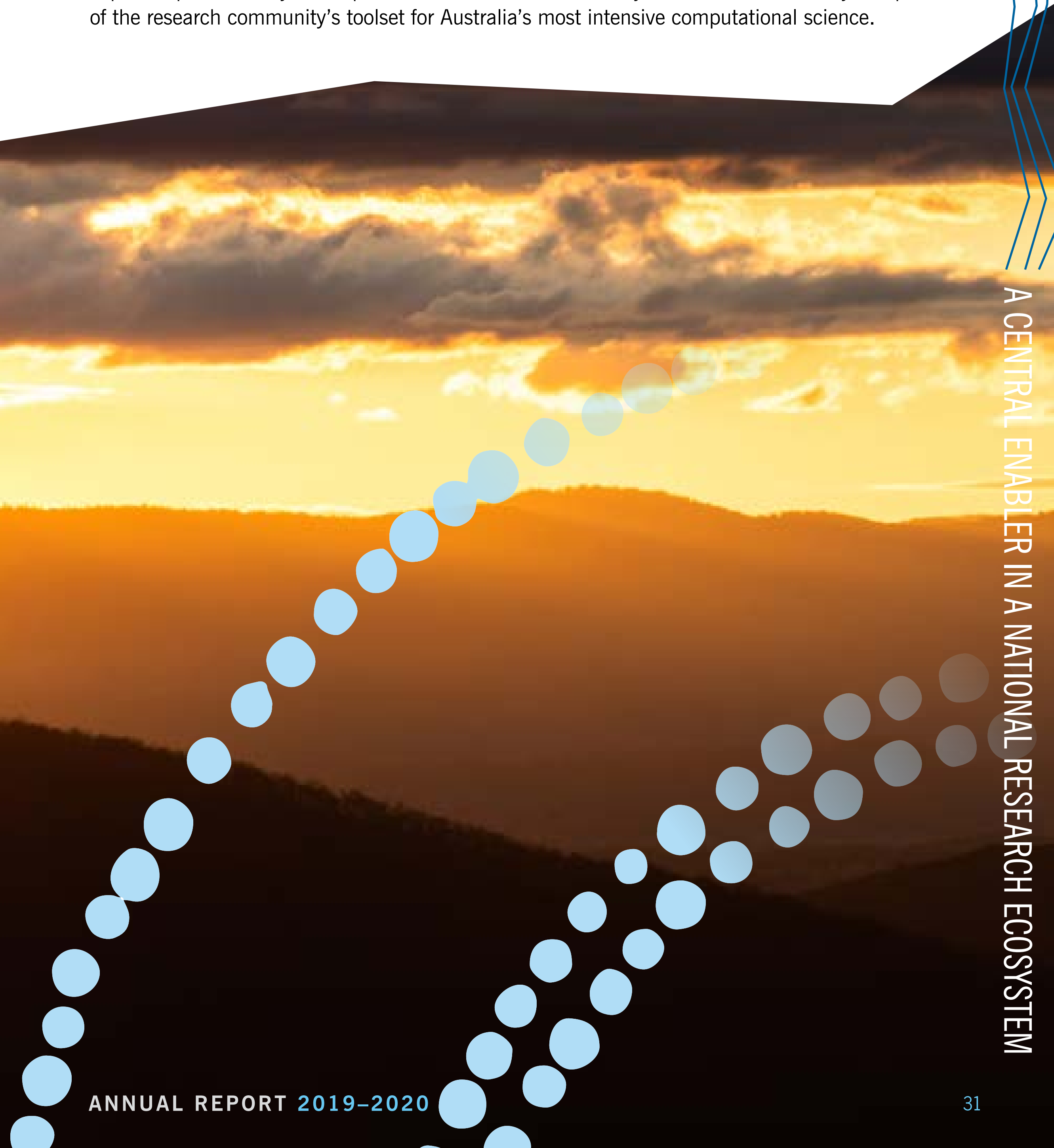
A Hub for the Research Community

NCI is an investment for and by the community. Together, the Australian Government, national science agencies, research universities and medical institutes invest in this key piece of infrastructure and its operations to support their many and varied research goals. Leveraging the combined investment, they support a larger and more capable set of computing and data services than any one of them could build individually. By so doing, they lift up the available national capability and make NCI a veritable hub for thousands of leading researchers across Australia.

NCI's variety of services covers the requirements of different scientific disciplines, workflows, code bases and data sources. We bring together the expert staff, centralised resources and support capacity to enable the biggest and most ambitious scientific research projects. Whether it is big datasets or thousands of processors, NCI's central platform makes it easier for the entire research community to share the resources they need.

One striking example is NCI's central role in the national climate research space. The climate data stored here – including the CMIP6 modelling data used by the Intergovernmental Panel on Climate Change – is produced, accessed and used by researchers from more than a dozen different research organisations around the country. Some of this data belongs to datasets that are shared around the world, thus connecting Australian researchers to the global community. This element of international cooperation and data stewardship is critical for Australian researchers to be involved in internationally collaborative research.

Whether it is aggregated climate data or astronomical data collected by Australian telescopes every night, NCI is a hub for the research of major national research communities. The Gadi supercomputer's finely tuned performance and that of its filesystems makes NCI a key component of the research community's toolset for Australia's most intensive computational science.

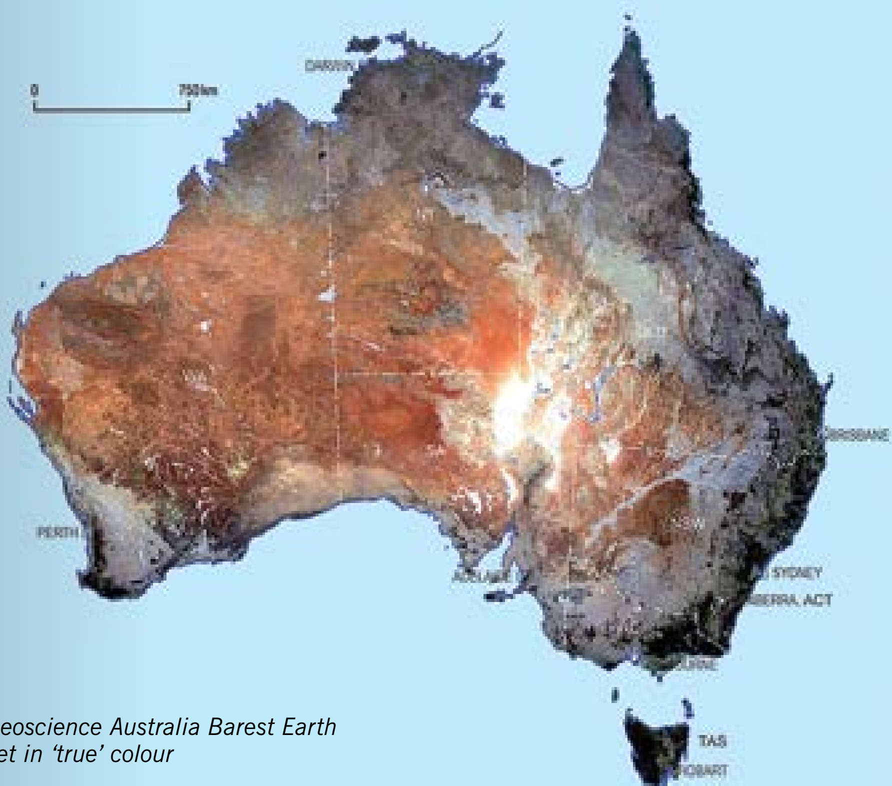


A NEW VIEW OF THE BARE AUSTRALIAN GROUND

Satellites have been orbiting the earth for decades, all the while taking pictures of the ground below. The pictures they take get processed, stored and after a while, archived. With some clever thinking, we can make sure to use the full value of those image archives. The challenge for researchers is using historical archives of satellite data to produce useful science today. Statistician Dr Dale Roberts and geologist Dr John Wilford from The Australian National University and Geoscience Australia respectively have teamed up to reanalyse the entire 30-year archive of Landsat satellite images of Australia and produce an entirely new research data product.

Drs Roberts and Wilford, looking to make new use of existing satellite datasets, have used NCI's supercomputer and data collection management tools to produce a new map of Australia. This map shows the earth in a state with the lowest possible level of vegetation to occur at that location. Built of a systematic analysis of the entire petabyte-scale Landsat archive – itself stored at NCI – the map provides a composite image of the entire country with every single pixel calculated to show the barest possible soil. The map, called the Barest Earth map of Australia, is a stunning view of our country's surface when all vegetation, urbanisation and bushfire scars are removed. The researchers here were motivated to make a new scientific data product that can be used as a foundational data input for other researchers' work.

NCI's unique integration of massive data collections alongside a world-class supercomputer enables data-intensive work that would otherwise be impossible to undertake. The researchers are now remaking the map with newer, higher resolution satellite images that will give even more clarity to the national map. By building even more efficient statistical processing algorithms and leveraging the ever-increasing incoming satellite datasets, national maps revealing the structure and formation of our continent are on their way. Environment management, mineral exploration and land monitoring all stand to benefit from access to these kinds of computer-driven data products.



The Geoscience Australia Barest Earth dataset in 'true' colour

2

A GENERATIONAL SHIFT IN AUSTRALIAN SUPERCOMPUTING

A Look Back at Raijin

In 2013, Raijin became Australia's first petaflop supercomputer, a milestone that heralded new possibilities for Australian computational research. It led not just to an increased technical capacity and the strong focus on data–compute integration that continues today, but also to a new collaborative model of funding and governance for NCI. The Raijin machine ran from April 2013 to December 2019, almost 7 years of continuous operation. Over its lifetime, it performed an incredible 3 billion hours of calculations.

All of the technical innovation was steered towards furthering national scientific ambitions. From an initial focus on climate and weather, and the traditional high-performance computing tasks of molecular modelling and fluid dynamics, NCI's range of scientific disciplines rapidly expanded to include geoscience, materials science, environmental science, astronomy and genomics, among others. Growing in concert with an expanding list of collaborating organisations, over the 7 years of Raijin's life NCI grew to support each of the major computational science disciplines in Australia in bigger and more profound ways.

Within a few years, one of Raijin's strengths quickly became apparent: the collocation of nationally significant research data alongside, and closely tied to, a world-class supercomputer. Still an uncommon capability at even the biggest supercomputer centres worldwide, the availability of collaborative research data at such scales, performance and quality made a whole range of scientific enquiry possible for the first time. Earth-observation was one of the big fields to benefit from this integration, with daily observations from US, European Union and Japanese satellites being made available to Australian researchers for processing and analysis at NCI. The tight integration of compute and storage systems turned into a major differentiating factor for NCI, and a significant enabling technology for some of the biggest and most data-intensive computational projects in the country. It also allowed for the early development of data-intensive science areas such as genomics, a field of science relying very strongly on the performance of data and its deep connections to a robust computing platform for analysis and processing.

In January 2017, NCI procured an additional supercomputing cluster through the support of the National Collaborative Research Infrastructure Strategy and the Australian Government's Agility Scheme. This cluster, several generations ahead of the one it joined, added around 30% to Raijin's computing capability and extended its life for several years. Integrated into Raijin, the newer processors were yet another option for researchers to take advantage of. Given its relative youth compared to the ageing original, the Agility System was retained during the Raijin decommissioning and is now providing its complement of computing power to the Gadi supercomputer.

Raijin ushered in the petascale era of computing in Australia. From its starting position at number 24 in the world, it supported nationally significant developments in weather and climate modelling, materials design and more. Hundreds of millions of hours of computing time were distributed through merit schemes to hundreds of Australian researchers, and hundreds of millions of hours more to the national science agencies CSIRO, the Bureau of Meteorology and Geoscience Australia. Spanning the disciplines and domains of science, Raijin enabled Australia's contribution to the international science community in areas of climate change modelling, geophysics and molecular modelling, among others. Decommissioned but not deconstructed, Raijin is being repurposed in 2020 as part of NCI's cloud computing offering.

TECHNICAL HIGHLIGHTS:

NCI launched with **57,472** **Sandy Bridge** processor cores in **3592** **nodes**

NCI's data storage capacity increased from **10** to **50** **petabytes** over **Raijin's** lifetime

In January 2017, **Raijin** grew by **30%** when around **20,000** **Broadwell processor cores** were added to the system, supported by the **NCRIS Agility Scheme**

RAIJIN'S BIGGEST USER

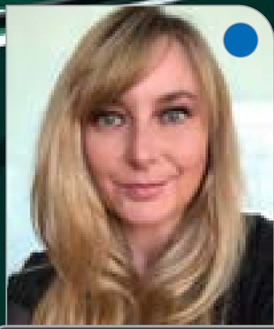
Over the 7 years of Raijin's lifetime, Professor Amanda Barnard was responsible for an incredible 8% of the system's total computing. A multidisciplinary researcher with a strong focus on computational and data science, Professor Barnard is developing new methods for applying machine learning to physical science data.

Using machine learning models appropriately for scientific purposes requires careful consideration. The physical sciences in general produce distinctive kinds of data featuring many dimensions, high variance, outliers and other unexpected variables. While machine learning can potentially be a useful tool for analysing these large datasets, it did not traditionally develop with that kind of data in mind. This is why Professor Barnard is developing new machine learning protocols and models suited exactly for the physical sciences.

Developing these models requires data, and lots of it. She needs to generate large, complex training data to train the models with. This requires huge amounts of computation, up to tens of thousands of jobs all run on the supercomputer. Running on Raijin enabled that scale of computing to take place, leading to improved computational science technologies.

Professor Barnard says, 'I have always been motivated by a desire to develop technologies that enable new and better computational science.' She started by using Raijin with a focus on computational modelling and thermodynamics, which progressed over the past 7 years to statistical learning and machine learning. She says 'As the need for greater complexity in the data grew, so too did the need for more speed and memory.' By the time Raijin was switched off in December 2019, she had used it to do 245 million hours of computing using over 1 million individual compute jobs.

Raijin's tenfold increase in computational performance over its predecessor opened up whole new avenues for researchers to investigate. In Professor Barnard's case, Raijin, Australia's first petaflop supercomputer, allowed her to increase the complexity of her datasets and move into the new and growing area of machine learning. Now a global leader in computational science, she continues to push the field forward.



Professor
Amanda Barnard

RAIJIN

TOTAL NUMBER OF HOURS OF COMPUTING DONE

3,217,489,124 HOURS

TOTAL NUMBER OF COMPUTE JOBS RUN

50,282,262

TOTAL NUMBER OF USERS PER YEAR

5,000+

RAIJIN MILESTONES

2013

First users on board in April **2013**

Official launch in June **2013**. Raijin comes in at 1.2 Petaflops on the November **2012** TOP500 list, Australia's first petaflop supercomputer and ranked number 24 on the global list of top supercomputers

NCI integrates Graphics Processing Units into the Raijin system in May **2016**

A Look Forward to Gadi

Australia's newest supercomputer, Gadi, is a game-changer for the country's scientific research. Gadi strengthens Australia's capacity to perform world-leading research, and directly contributes to the safety and wellbeing of the nation. First made available to users in November 2019, Gadi brings the latest computational technologies and increased performance to the research community.

Coincidentally also ranked 24th in the world as of June 2020, Gadi once again places NCI and Australia among the ranks of the top global supercomputers. The computational performance available on Gadi is unprecedented in Australia, enabling rapid response to national emergencies such as COVID-19 and bushfires that we could not have approached before.

Nationally significant research in areas of cyclone and tsunami monitoring, bushfire and weather prediction, climate change modelling, and more relies on the power and flexibility of the Gadi supercomputer. Further impactful research in materials science, genomics, fluid dynamics, drug design and more, from Collaborators and research groups around the country, is being turbocharged by the new capacity.

NCI's integration of data and compute services gets deeper with Gadi. Larger models are processed, analysed, stored and referenced through NCI services, and earth-observation datasets, among others, bring their impact through our own layer of data analysis and processing tools.

The next generation of Australian high-performance computing is here.

2018

NCI gains early access to Intel's latest generation of processors, Knights Landing many-core CPU chips in September **2016**

The Agility System featuring 800 nodes of Broadwell processors launches in January **2017**, bringing Raijin back up to number 70 in the world

New filesystems beyond the original /short installed in **2015, 2016** and **2018**, bringing NCI's total storage to around 50 Petabytes

AUSTRALIAN WEATHER UNDER THE MICROSCOPE

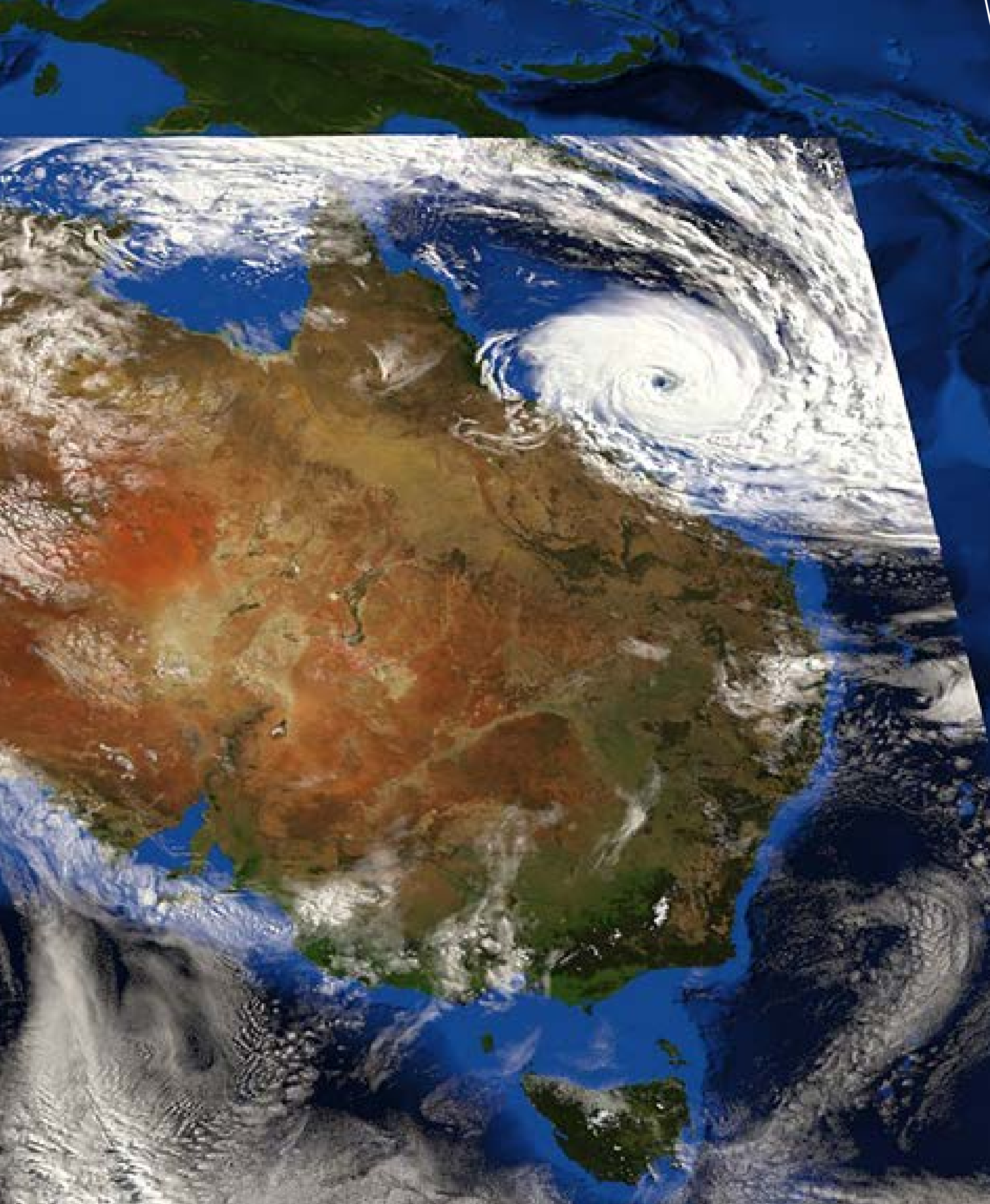
The next big thing in weather forecasting will be national models with much higher resolutions than we currently have. Getting the model resolution down to hundreds of metres across Australia – from the twelve kilometres that is currently state of the art – will make it much easier to simulate localised weather patterns like thunderstorms and cloud fronts, and how they interact from the tropics to the extra-tropics. Modelling the weather at a higher resolution means modelling all of the small interactions between hot, cold, moist and dry air, as well as topography, wind gusts and more.

Professor Christian Jakob, his colleagues Scott Wales and Professor Todd Lane, and collaborators at the Bureau of Meteorology have run an incredible weather model covering the entire Australian continent – a grid 5,600 kilometres from east to west and 4,000 kilometres from north to south – at 400-metre resolution. The grid contained an unbelievable 12.6 billion points. To do this, they used the Unified Model from the UK Meteorological Office, the same model that is used as a key component of Australia's own ACCESS weather and climate model. The simulation contained by far the most grid points that the Unified Model has ever been used on. It took around 3 weeks and 12,000 processors running on NCI's Gadi supercomputer, producing almost 100 terabytes of data and pushing all aspects of the software and hardware to their limits.

As the researchers say, they were unsure that a simulation of this size was even feasible. 'The aim was to do a demonstration simulation to show what might be possible. We wanted to see what simulations at this scale reveal in terms of scientific and technical unknowns,' said Professor Jakob.

The exciting thing now is the number of new scientific questions that will be raised, questions that we can't even begin to imagine. The data we've produced is going to be a veritable playground for not just atmospheric scientists, but potentially users of high-resolution weather information as well.'

The researchers picked a very specific time period over which to run their simulation: around two and a half days in March 2017 containing the landfall of Tropical Cyclone Debbie, one of Australia's most devastating recent cyclones. While the entire cyclone was over 100 kilometres across, lots of critical interactions took place on the edge and in the eye of the cyclone, which is where the increased resolution makes a major difference. Already the researchers have noticed a stunning and welcome similarity in the eye of the cyclone between the model's outputs and satellite images.



Pushing a supercomputer and an international weather model to their limits is exciting and helps researchers learn about the kinds of challenges that will be involved in moving operational weather forecasts to those kinds of unprecedented resolutions. While 400-metre resolutions for areas as large as Australia are still a decade or more away from operational forecasts, this work and the ambitious, much longer simulations that it has inspired, are critical steps that will bring us there that little bit faster.

CELEBRATING THE NCI SUPERCOMPUTER WITH A NGUNNAWAL NAME AND ARTWORK

The Ngunnawal people of the Canberra region have for thousands of years been learning about the world and searching for knowledge in the sky, water, stars and earth all around them.

In celebration of National Reconciliation Week, Australia's largest supercomputing facility, NCI, housed on campus at The Australian National University, has unveiled the name and feature artwork on its newest research supercomputer, built on Ngunnawal country.

The United Nations General Assembly declared 2019 the International Year of Indigenous Languages to raise awareness of the crucial role indigenous languages play in people's daily lives and culture. It was during this year that NCI's new supercomputer was built.

In consultation with the United Ngunnawal Elders Council, the new machine was christened 'Gadi' [Gar-dee], a word of the Ngunnawal people meaning 'to search for'. It perfectly encapsulates NCI's mission of scientific research and high-performance computing: to search for knowledge that can make the world a better place through enabling innovative world-class research.

In addition to the name that will carry the national supercomputer forward over the coming years, Gadi will also be adorned with its own artwork, hand-painted by renowned Ngunnawal artist Lynnice Church. Gadi's artwork visually represents the coming together of western and Indigenous knowledge systems, guided by the Ngunnawal Elders over many generations. Mrs Church says, 'The art symbolises the gathering of knowledge from different places and sources, and the bringing together of an even greater collective knowledge system.'

The bold shapes of the Elders stand in contrast to the fine detail of the knowledge rings. The many sources of knowledge flow and come together in the centre to represent the combination of knowledge systems. Both important and complementary, the two knowledge systems inherent in the world of Australian science are proudly displayed on the Gadi supercomputer.

NCI Director Professor Sean Smith says, 'We are proud to be able to make this connection between the traditional owners of the Canberra region and its newest tool of scientific discovery. It is a great honour for the computational research community to have such a perfectly suited name and artwork as the face of Australia's new supercomputer.'

NCI is proud to acknowledge and celebrate the ongoing contribution of Aboriginal and Torres Strait Islander people to Australian science and discovery. This new supercomputer sits on land cared for by generations of scientists and discoverers. We hope to see it continue that tradition in the years to come.

AUSTRALIAN RESEARCH IS READY FOR ITS NEXT SUPERCOMPUTER

CSIRO says:

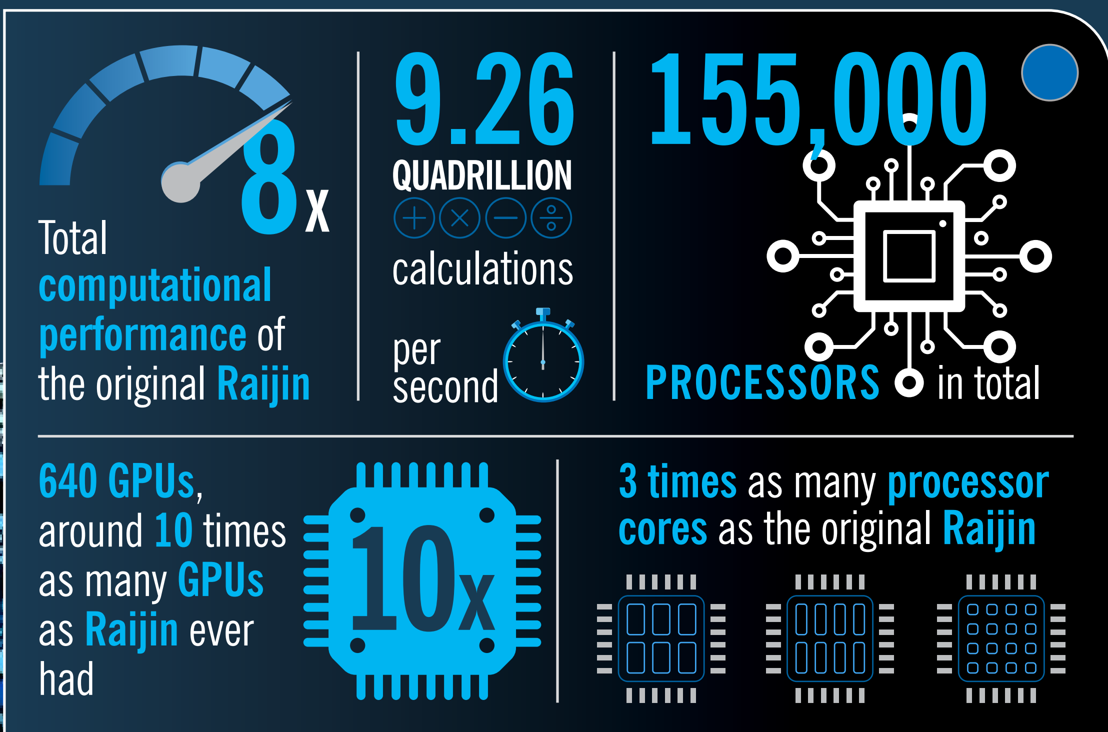
‘National high-performance computing facilities are critical for Australia’s future,’ says Dr Dave Williams, Executive Director of CSIRO - Digital, National Facilities and Collections. ‘At CSIRO, we solve the greatest challenges using innovative science and technology, and Gadi will enable our work in Earth sciences including climate, land, water and fire modelling. The data that Gadi will process and analyse is collected here on Earth as well as by an increasing number of satellites orbiting our planet.’

The Bureau of Meteorology says:

‘The next generation of seamless prediction systems will make weather forecasts more locally relevant, more accurate and more useful for longer periods of time,’ says Dr Gilbert Brunet, Chief Scientist and Group Executive, Science and Innovation at the Bureau of Meteorology. ‘The Gadi supercomputer will be used to support transitioning the experimental models to routine operation. Improvements such as those from the ACCESS-Fire model and specific event case studies are used to better plan fire suppression operations, reduce risk to firefighters and assist the community in affected areas.’

Geoscience Australia says:

‘Geoscience Australia, aided by the National Computational Infrastructure, is improving the safety and resilience of communities around Australia,’ explains Dr James Johnson, CEO of Geoscience Australia. ‘Without NCI, we would not be able to compute the huge data needed to assess the risks associated with hazards such as earthquakes, tsunamis and tropical cyclones. I’m looking forward to seeing how this new supercomputer will enable further work at Geoscience Australia to ensure a better informed nation.’



3

BUILDING NEW TECHNOLOGIES FOR NEW SCIENCE

The Supercomputer and its Filesystems

GADI – THE SUPERCOMPUTER

3,074 nodes each with **2 24-core Intel Cascade Lake Central Processing Units** (CPUs)

155,000 processor cores

9.26 Petaflops

160 nodes each with 4 **Nvidia V100 Graphics Processing Units** (GPUs)

50 CPU nodes also offer **1.5 TERABYTES RAM**

Data transfers inside the **supercomputer** at up to **200 GIGABITS PER SECOND**

THE FILESYSTEMS

Total active storage capacity of more than 60 PETABYTES

Gadi filesystem for **storage** of **active project data**
12 PETABYTES capacity and **performance** up to **500** Gigabytes per second

Four global filesystems for storage of persistent research data with **54 Petabyte capacity** and **performance** up to **140** Gigabytes per second

Replicated archival data storage with **50 PETABYTES capacity**

NCI's new supercomputer marks a major milestone in the life of the organisation and in the capabilities of Australian research. The new Gadi supercomputer is the most powerful in the Southern Hemisphere, supports the work of more than 6,000 researchers and underpins the vast majority of NCI's diverse range of data- and compute-intensive services. A major step forward in technology for Australia, this supercomputer, much like the science it enables, is world-class.

Gadi, meaning 'to search' in the language of the Ngunnawal people, the traditional owners of the Canberra region, will provide the next generation of computational performance to Australia's researchers. Gadi is powered by Intel's latest generation Cascade Lake processors, providing an 8 times speed-up on the original Raijin supercomputer it replaced. Additionally, Gadi features 640 of Nvidia's V100 GPUs (Graphics Processing Units), a type of computing processor specialised for highly parallel workflows. As artificial intelligence and machine learning become a bigger part of high-performance computing, the power and flexibility of GPUs are taking an ever larger role.

Gadi is custom designed to suit the workflows of our wide variety of research users. In particular, the processors, architecture, network performance and components were all chosen with this goal in mind: to accelerate the workflows we already have and to enable new workflows coming in the future. Thus, Gadi is well suited to running some of the biggest climate and weather modelling codes in the world at unprecedented resolutions, while also catering to research users in human genomics, molecular modelling, materials science, fluid dynamics, and more. Gadi is a worthy foundation for the growth of Australian computational science in coming years. Already we are seeing world-leading simulations emerge, all made possible by this new infrastructure platform.

The Gadi filesystems, home to more than 60,000 terabytes of research data, continue to grow to support the influx of national data collections, and the additional performance needed to bring those collections to life in researchers' hands. With performance up to 400 gigabytes per second on the newest filesystems, the data-intensive sciences are well served by NCI's data repositories.

In 2019–20

Total number of **projects**

1,125

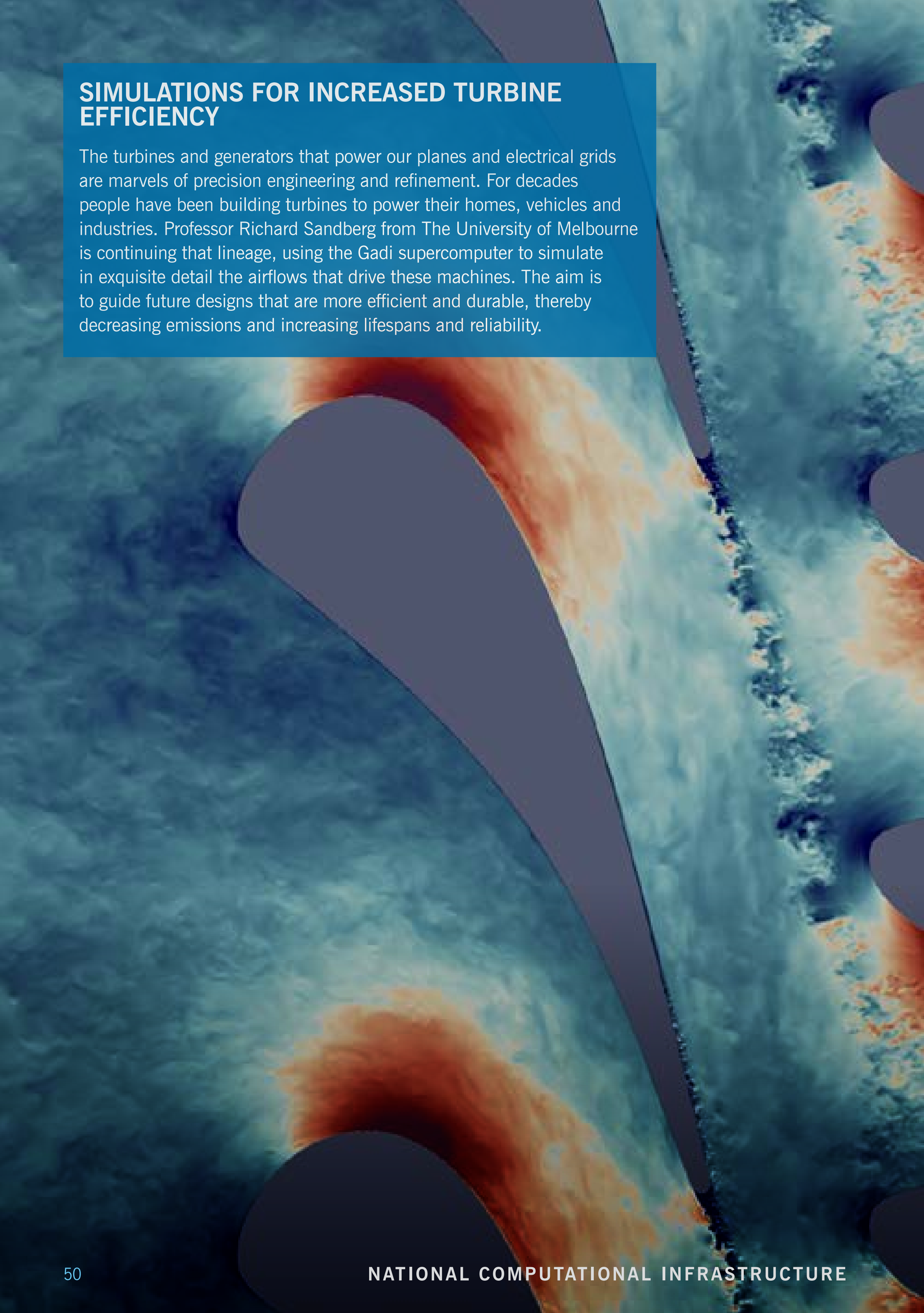
Total number of **users**

6,000+

Total **computing time**

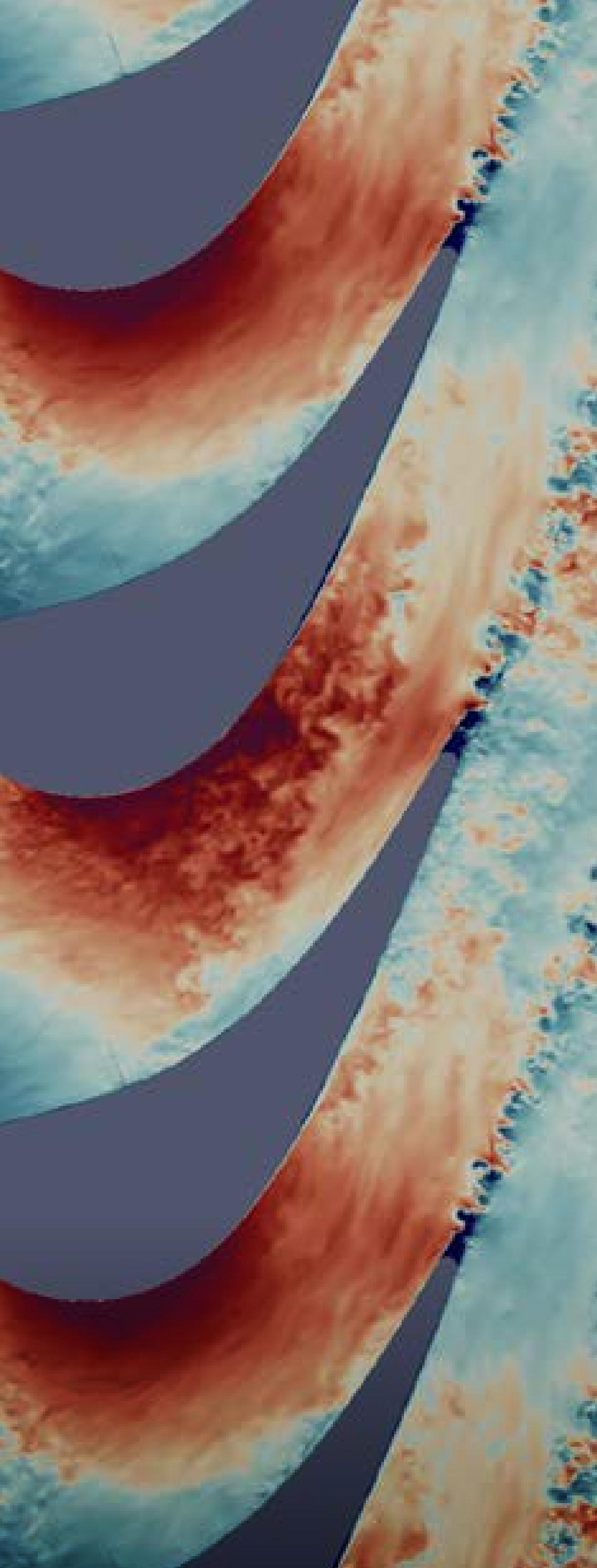
756 MILLION hours

For a full list of all compute allocations on NCI in 2019–20, visit https://bit.ly/NCI19-20CompAllocations.



SIMULATIONS FOR INCREASED TURBINE EFFICIENCY

The turbines and generators that power our planes and electrical grids are marvels of precision engineering and refinement. For decades people have been building turbines to power their homes, vehicles and industries. Professor Richard Sandberg from The University of Melbourne is continuing that lineage, using the Gadi supercomputer to simulate in exquisite detail the airflows that drive these machines. The aim is to guide future designs that are more efficient and durable, thereby decreasing emissions and increasing lifespans and reliability.



A view of turbulent air flows inside a gas turbine using Direct Numerical Simulation produced by Professor Sandberg.

The flows of air in extreme environments, such as in the 1,500 degrees, 50 bars of pressure, and thousands of rotor revolutions that take place every minute inside a gas turbine, are extremely complex. Measurements are practically impossible: the only way to see what is happening inside the machine is with Direct Numerical Simulations (DNS). These first-principle based simulations can capture the physics of the airflows model-free, and recreate the chaos, turbulence and eddies present inside a rapidly spinning turbine. With a resolution down to micrometres, DNS uses billions of grid points and millions of time steps to calculate the fluid flows as they evolve and interact.

Running on Gadi means that simulations can cover the full scale of flows inside the turbine at real-engine conditions, from the tiniest micrometre-sized interactions to the bigger centimetre-sized effects. Being able to simulate them all at the same time means that the interaction between small-scale turbulence and larger eddies becomes clear. The computing limitations that restrict the size of DNS are being pushed back by new technologies, opening up research possibilities that were previously inaccessible.

One of those new technologies is Graphics Processing Units (GPUs), a growing technology in supercomputing that is perfectly suited to running simulations such as these requiring vast numbers of grid points. Gadi's new V100 GPUs offer extreme performance, allowing Professor Sandberg to split his grid into dozens of manageable 20-million-point chunks. This leverages the rapid communication that NCI has built between the GPUs, as well as their own performance characteristics.

Professor Sandberg says, 'It's exciting to be able to see the impact of what we do. If we can find something that makes planes and power generation more efficient, it's an important advancement that benefits everybody. And we shouldn't underestimate the skills development and knowledge that my colleagues and students have built up. Supporting and watching them grow is also very rewarding.'

Computational Science Enhancements

Software enhancements to scientific codes increase efficiency, performance and reliability, and enhance the possible science outcomes. Traditionally focused on the optimisation of large climate and ocean models, this year NCI has extended our expertise in HPC code optimisation to assist more disciplines and their codes. The growing list of domains now includes climate and weather, astrophysics, geophysics, genomics, fluid dynamics, and quantum computer emulation.

Leveraging the technologies offered by Gadi's state-of-the-art Cascade Lake processors can offer researchers improved performance and new capabilities, but achieving that requires expertise. NCI's HPC Simulation Scaling and Data Analysis Optimisation Team helps research groups understand the bottlenecks and inefficiencies in their codes. Then, NCI can provide solutions to implement that save the researchers' time and computational resources, and lead to better scientific outcomes.

A major highlight of 2019–20 is our advance into genomics software, optimising it to run as well as possible on our supercomputer platform. Focused in particular on sequence alignment and processing, genomics software is relatively new to high-performance computing. It also has very specific patterns of computing and data use that distinguish it from the climate and weather codes

we are used to. Working with medical and biological research organisations, we have significantly sped-up an important genomics workflow tool and optimised it to work at scale (see Case Study *Enabling the next-generation of genomic medicine* on pages 54–55 for more information about the software advances). This should allow research teams from the Garvan Institute of Medical Research, The University of Sydney Informatics Hub and others to process and analyse more genomes more efficiently.

Another major achievement of the year was significant speed-ups to the communication portion of a commonly used ocean modelling code, the Modular Ocean Model (MOM5). By leveraging the Gadi supercomputer's highly parallel nature and rapid data network in the core MOM5 code, the NCI team reduced the duration of the communication process by around a factor of 4.

NCI's HPC code optimisation efforts this year extend the fields of research we help support. We have increased our range to include new areas, such as geophysics and genomics, with tremendous HPC needs. As always, getting the most performance from massive supercomputer calculations requires carefully optimised codes. Just as each scientific code is different, so are the required improvements. The benefits are clear, and NCI provides the expertise necessary to achieve them all.

ENABLING THE NEXT GENERATION OF GENOMIC MEDICINE

Bulk processing of genomic data requires high performance from all aspects of a supercomputer. It demands large amounts of computational power, networking speed and storage capacity, called for by dozens of sequential steps for each genome processed. The glue holding all of these together is the software controlling the process, from reading the input data to synchronising the processor cores and collating the final processed genomic data into a useable form. NCI's HPC Scaling and Data Optimisation Team worked closely with the Garvan Institute of Medical Research to improve the existing genomics software workflows and make them as effective as possible on the new Gadi supercomputer (see Case Study *Reprocessing a highly valuable genomic dataset* on pages 24–25 for more information on the science behind the genomics processing).

Genomics processing takes in snippets of genome sequence data, stitches it all together and identifies genetic variations in more than 6 billion letters of DNA. This process takes many steps, each of which has its own technical requirements and bottlenecks. Normally, this would mean that compute nodes can sit idle or under-utilised for a time – for example during a data writing and reading step. To improve the efficiency of processing and increase the utilisation of the compute nodes, NCI helped optimise the workflow to connect the steps more efficiently and process multiple samples at the same time.

Before this optimisation, it took around 23 hours to process a single sample. With the optimised software running, a single sample on Gadi took 7 hours to process. What's more, four samples processed at the same time took just 19 hours all together, or less than 5 hours each. The gains from running multiple samples in parallel on the same compute nodes are significant, and speed up genomic processing in a big way. All 4,012 samples of Garvan's Medical Genome Reference Bank were processed in just over a week on Gadi, with hundreds of genomes split across many dozens of processors running simultaneously.

At the start, it was clear that Gadi had untapped capacity that the existing workflow was not taking advantage of. Understanding how best to pack the different parts of the workflow to maintain high utilisation was an interesting and important challenge. Moving from processing single samples to processing samples at scale required a deep understanding of all elements of the code. We preserved the complex workflow that the genomic processing needed while tuning it to run in the new Gadi environment.

Genomic medicine has the potential to improve medical processes for millions of people, including the diagnosis, treatment and discovery of both rare and common diseases such as cancers and heart disease. Underpinning those benefits is the infrastructure, both the computational hardware and the custom-built software. More than ever, getting the most out of our supercomputers and our data requires carefully optimised codes that leverage new and existing technologies for scientific benefit.

Data Science Technologies

Science in all disciplines is relying more than ever on large distributed data collections and access to high-quality data. On the data science front, NCI this year embarked on a path of modernisation: adding to and improving our data collections' functionality to help users access, analyse and share the data they need every day. We are constantly trying to make sure that the data stored at NCI is safe, secure and accessible. As a trusted repository for some of Australia's biggest datasets, we play a key role in the national data science community.

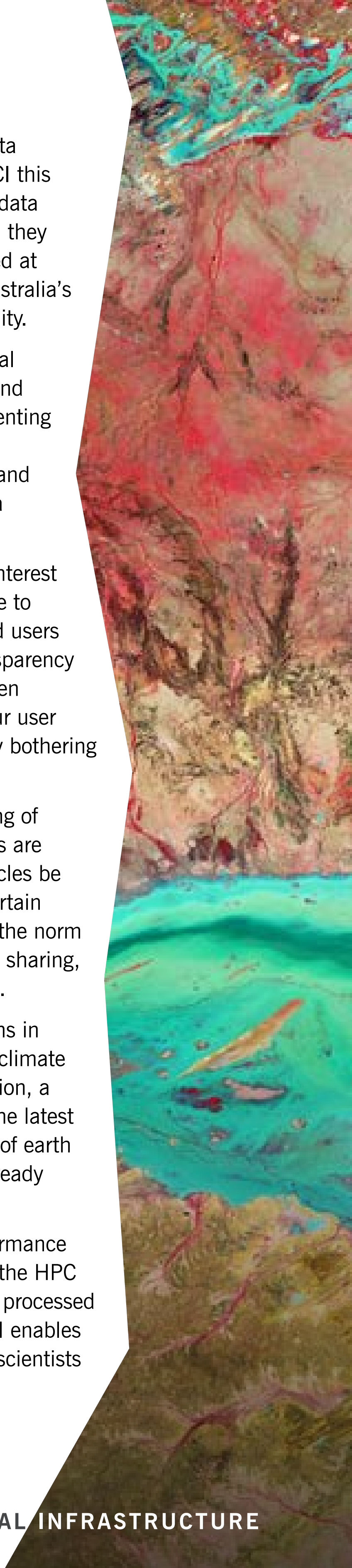
We are continuing to ensure that all data collections meet the international FAIR data standard: data should be Findable, Accessible, Interoperable and Reusable. This includes a large number of improvements we are implementing that are transparent to users but strengthen the underlying foundations. These are steps such as overhauling some of the data structures we use and continuing to work with our data providers on making sure that their data collections are up to date and accurate.

A change this year requiring users to register to access their datasets of interest has simplified communication and access mechanisms. Now, we are able to track usage and access patterns and correspond directly with the affected users when particular data collections receive updates. This increases the transparency for data users who are immediately notified with relevant information when there are changes to a dataset they use. Targeted communication with our user community in this way helps notify relevant parties without unnecessarily bothering our entire user list.

Another data collections modernisation step this year has been our hosting of published reference data. Leading scientific journals and funding agencies are starting to require that copies of the scientific data used in published articles be made publicly available. We are providing this service as data host for certain users with particularly large published datasets. As big datasets become the norm for much of scientific research, new and improved systems for accessing, sharing, referencing and publishing scientific data are becoming ever more crucial.

Some of the biggest and most significant datasets added to NCI collections in the past year include: Australia's contributions to the international CMIP climate modelling project, the latest version of the Landsat satellite image collection, a range of ocean and weather data including models and radar data, and the latest release of astronomy data from the SkyMapper project. Covering a range of earth and environmental science disciplines in particular, these datasets are already getting significant traffic from the scientific community.

NCI is uniquely placed to provide large-scale data storage in a high-performance data and high-performance computing environment. The combination of the HPC and HPD systems enables modelling data to be computed, analysed and processed all in the same place. The wide knowledge and expertise from across NCI enables us to constantly improve the way we store data and make it available to scientists from around the country.

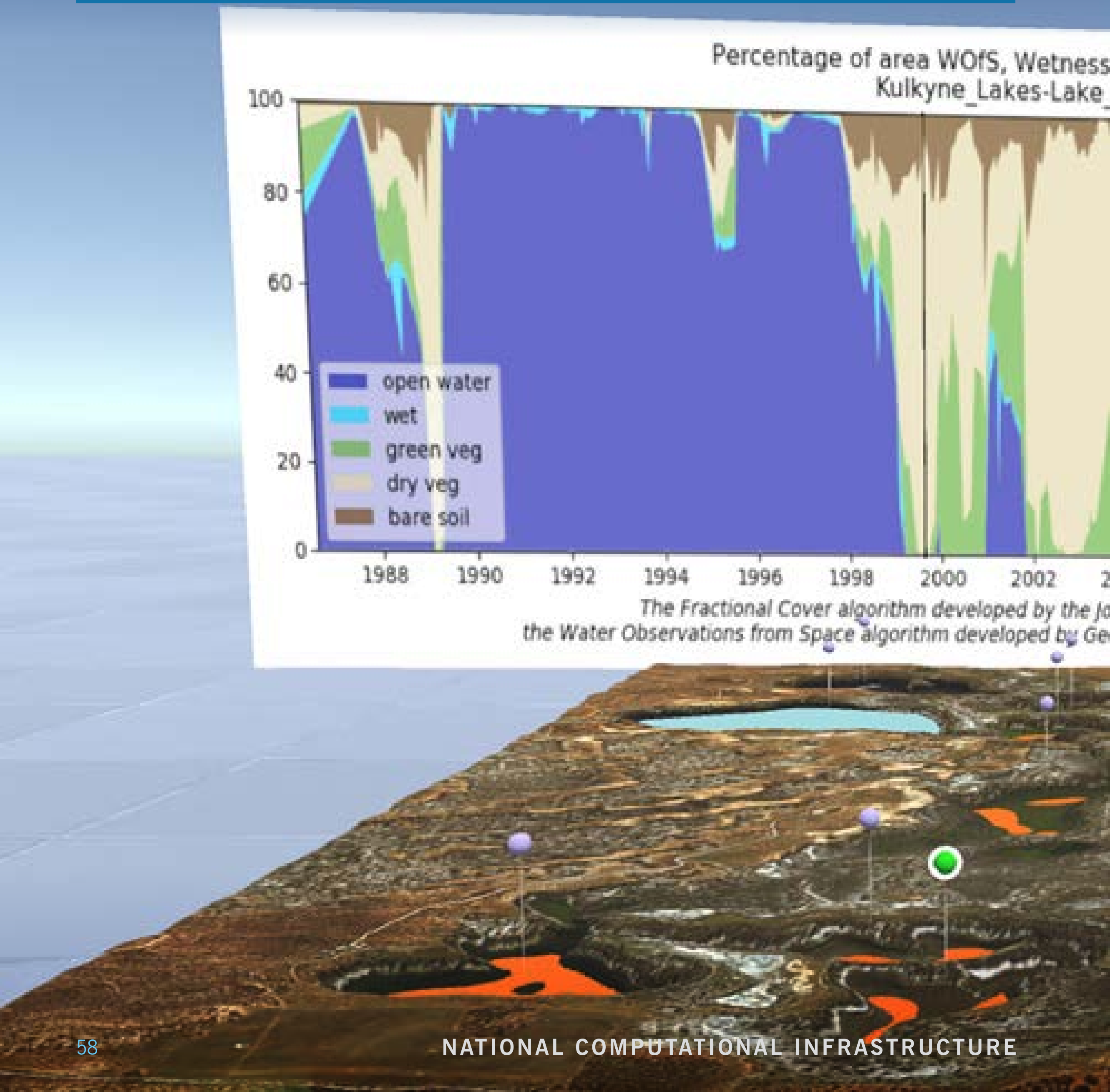


An image of the Channel Country in southwest Queensland, Australia, captured by the European Space Agency's Copernicus Sentinel-2 mission on 5 April 2019. Vegetation in the image is coloured to appear bright red.

TRACKING ENVIRONMENTAL WATER WITH VIRTUAL REALITY

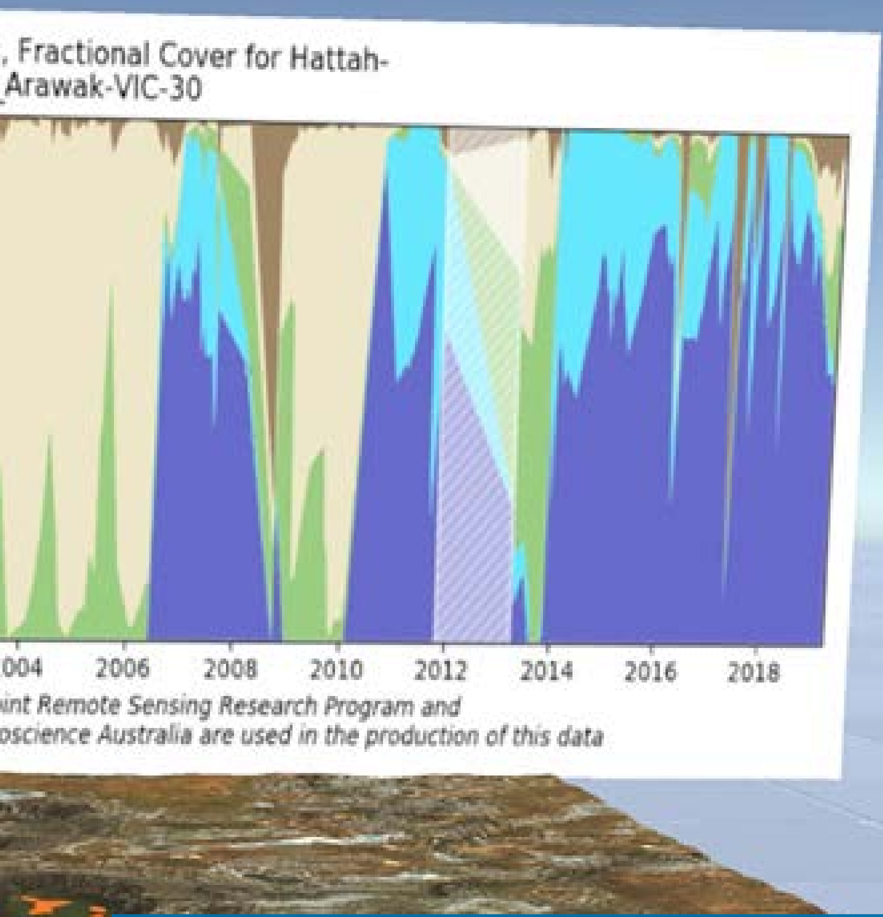
NCI's data science technologies take research data to the next level by bringing to it technologies and expertise that make it more useful and powerful as an input to scientific inquiry. For more than 20 years, the NCI VizLab has been pursuing that mission in the field of scientific visualisation. Evolving over time with shifts in technologies, in 2020 the VizLab is focusing on virtual reality (VR).

In November 2019, the VizLab produced a VR experience showcasing environmental water flows in Africa. This experience was presented at an international summit bringing together thousands of ministers, delegates and scientists from all over the world. The Group on Earth Observations 2019 Ministerial Summit (GEO Week 2019), one of the largest intergovernmental geoscience gatherings in the world, took place in Canberra, providing a great opportunity for NCI to highlight exactly what it can do with research data in the geoscience space.



A VR experience in which you are controlling water flows in lakes and rivers, this immersive demonstration allowed its users to intuitively learn about the role of water in the environment through the sights and sounds of a virtual world. They could also see for themselves an example of the benefits of capturing satellite data of the environment, and the innovative technologies and sciences that this enables.

Actual satellite data from Geoscience Australia provided the structure around which the whole experience was built. Attendees at GEO Week were greatly impressed by the experience, including ministers from South Africa, Senegal and Ethiopia, and delegates from other places around the world including several Pacific Islands nations and the United States. Overall, the attendees who tried out the VR experience found it an exciting and novel way to demonstrate the power of satellite data and the real-world context that it represents.



This use case is one of many possible once the wealth of environmental data stored at NCI is leveraged creatively through virtual reality technologies. Other demonstrations of VR have looked at ocean currents and temperatures around the globe, rock formation, the structure of bones and wood, and more. Virtual reality is opening up new and exciting possibilities for scientific exploration and knowledge sharing. For scientists and industry just as much as for ministers, students and educators, VR presentations of scientific data serve to illustrate and delight.

Data-Intensive Services

Data-intensive services open up opportunities for researchers to take existing datasets and approach them in new, compelling ways. We leverage computational technologies in innovative ways, allowing us to offer powerful analysis and processing services for our users conducting data-intensive science. Services offering tools for data processing and analysis allow scientists to focus on their research without needing to worry about the computational details and data management. NCI's combination of computational performance, data storage capacity and data management expertise offers dynamic ways of interrogating data that extend the capacities and usefulness of research data.

NCI's GSKY data service (pronounced ji-skee) is a central part of data-intensive research now taking place at NCI. The service processes earth-observation data on-demand, delivering to the user all the relevant data for a location of interest, regardless of the data source, date or format. A high-performance service that makes accessible otherwise obscure and complex datasets, GSKY is a well-established research tool for the environmental science community. This year, almost 3 years on from its original introduction, GSKY is continuing to receive major updates that extend its functionality. The continued requests for new features from major user groups demonstrate GSKY's importance to this subset of researchers, including Geoscience Australia who use it as a core part of the Digital Earth Australia project (see Case Study *A new view of the bare Australian ground* on pages 32–33 for more details).

New in 2020, NCI is providing an interactive research environment called Pangeo that gives users interactive access to Gadi's parallel computing nodes, allowing data analysis, image processing and live code testing. A major addition to NCI's existing Virtual Desktop Infrastructure, Pangeo now allows interactive computational tasks to run on multiple Gadi nodes at once, a revolutionary improvement to the current workflow for big data science. Pangeo gathers parallel computing tools in one convenient location for new groups of researchers. Incorporating data analysis and code parallelisation software, the Pangeo improvements provide new ways of processing big blocks of research data with improved flexibility for researchers.

Science in all disciplines is now producing and storing large amounts of research data. This increased volume of data collections provides an opportunity to develop new ways of analysing data going beyond the methods of the past. Computational methods integrated with large data collections extend the science that the data enables in the first place. The combination of supercomputing performance and innovative data services connects and enriches the valuable data stored on site at NCI.

Australian Hits
287,638,305

Australian downloads
4,901,466,195
MEGABYTES

International Hits
106,592,820

International downloads
399,154,185
MEGABYTES

For a full list of data downloads from Australia and around the world in 2019–20, visit https://bit.ly/NCI19-20DataDownloads

GLOBAL CLIMATE INFRASTRUCTURE

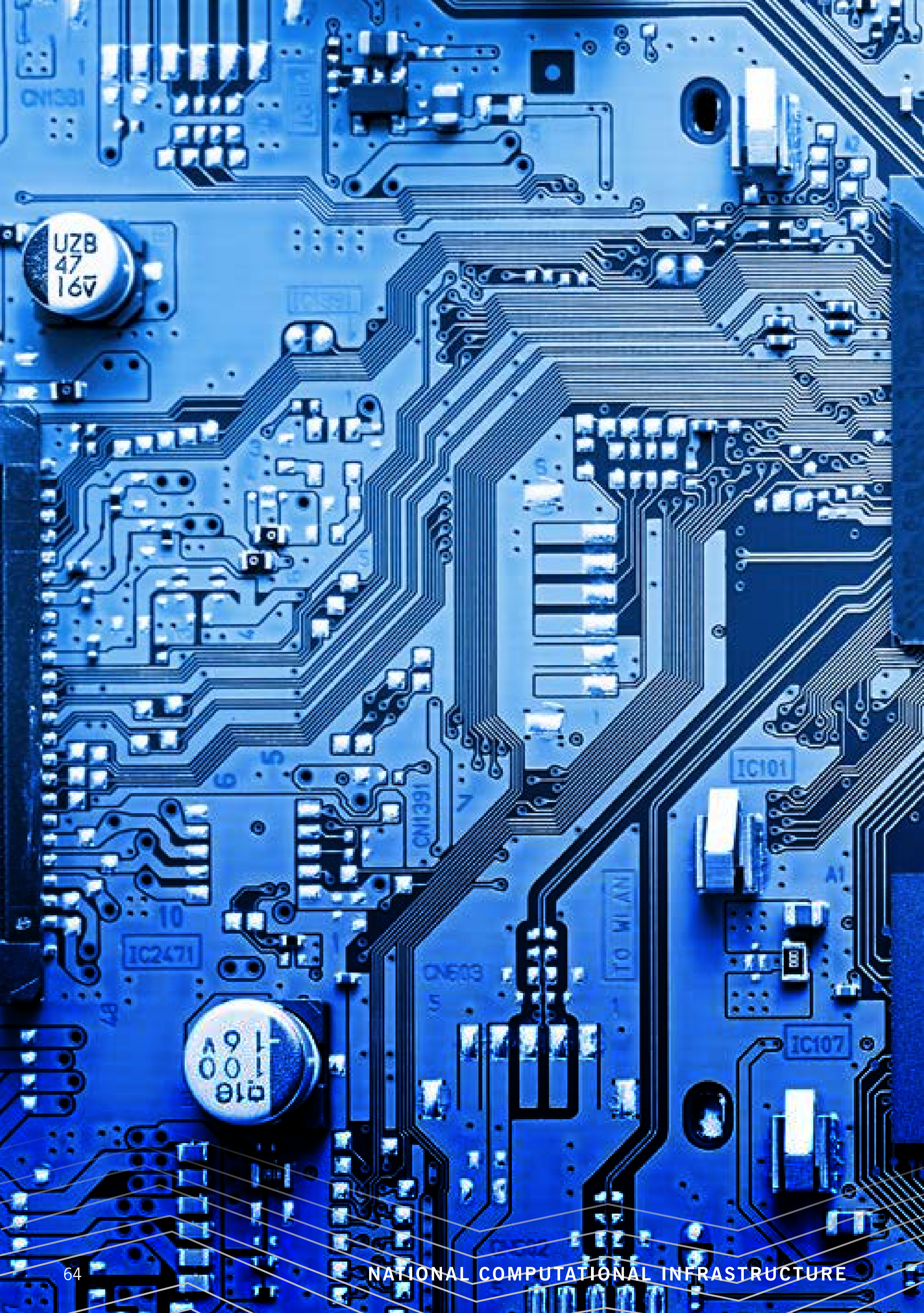
NCI is the only place in Australia where it is possible to run century-scale global climate models and at the same time store, process, analyse and share that data with others in Australia and around the world. As one of the continental-scale Tier-1 climate data centres, NCI joins peer centres from the US, Europe and the UK in providing the international Earth Systems Grid Federation (ESGF). Through the ESGF, NCI is the Australian hub for the global climate data from CMIP6, the Climate Model Intercomparison Project Phase 6 (see Case Study *Global climate models and data analysis* on pages 18–19 for a summary of the science for which CMIP6 data is being used). Supporting the research reliant on CMIP6 means providing robust and scalable solutions for hosting and organising large and growing datasets.

NCI's systems have been designed and tuned to enable Australia's strong participation in CMIP6. Supporting large-scale data activities of this kind requires dedicated focus and an understanding of the needs of research users. Importantly, with incoming CMIP6 data measuring in petabytes, NCI's data repositories need to be much more than just a place for storage and high-performance data analysis. They also need to provide advanced search, citation and sharing functionality for petabytes of data. Much of the work NCI puts into enabling the use of CMIP6 data revolves around the extra requirements that the unprecedented data quantity needs and the opportunities it allows.

The amount of data requires our data download, quality control and publication procedures to scale in response. Due to the scale of CMIP6, NCI has developed automated software- and data-driven processes, which have replaced manual data management systems used in previous international activities. This has meant that users have even more timely access to more quality datasets than in the past. Efficiency gains due to software and process improvements make research easier and quicker, and remove data management barriers to access that have previously hampered research efforts. More timely access to CMIP6 data for Australian researchers allows them to meaningfully contribute to international climate modelling activities.

NCI's data analysis environments support the growing use of climate tools, workflows and open source software. Shareable Jupyter notebooks that use new HPC methods that scale across distributed compute nodes leverage the available supercomputer capacity needed for large-scale climate data analysis. NCI's data-enhanced systems include massively improved search and collaboration capability. This allows researchers to programmatically search for specific criteria when looking through the CMIP models.

NCI is Australia's central hub for access to and use of the internationally significant datasets that comprise CMIP6. Through this, NCI has created a comprehensive platform for access, data-intensive analysis and research collaboration.



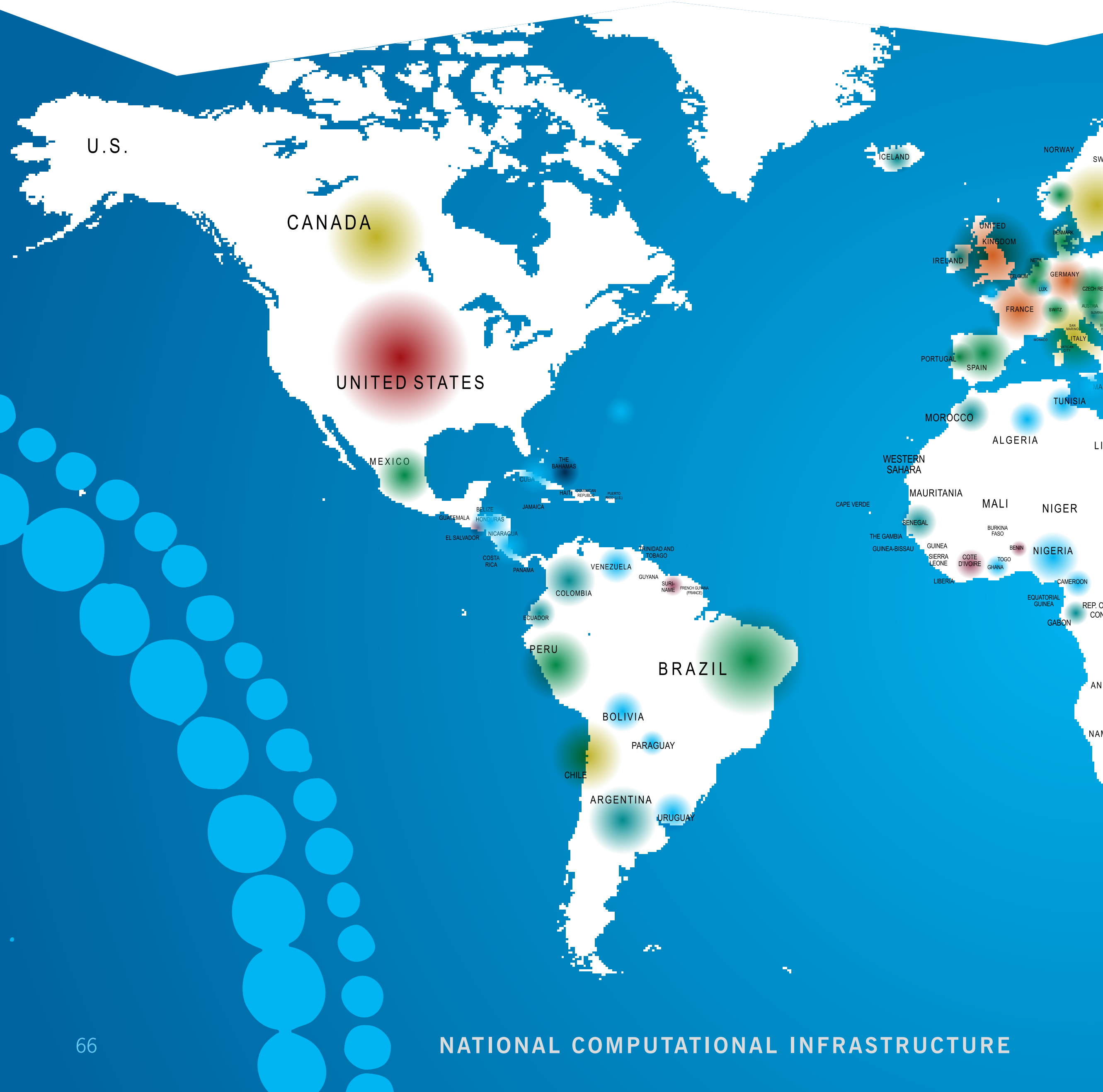
4

A BIG DATA AND BIG COMPUTING LEADER

International Engagements

As a global leader in high-performance data and computing, NCI works closely with numerous international research organisations. NCI plays a central role in the climate research community, as one of the 6 nodes worldwide collaborating to share and make available climate data through the Earth Systems Grid Federation. Alongside partners including NASA, the German Climate Research Centre and others, NCI provides access to petabytes of climate model data for the Australian and global research community. This activity, critical for thousands of researchers worldwide, is a striking example of NCI's deep and continuing engagement with the most central scientific research activities.

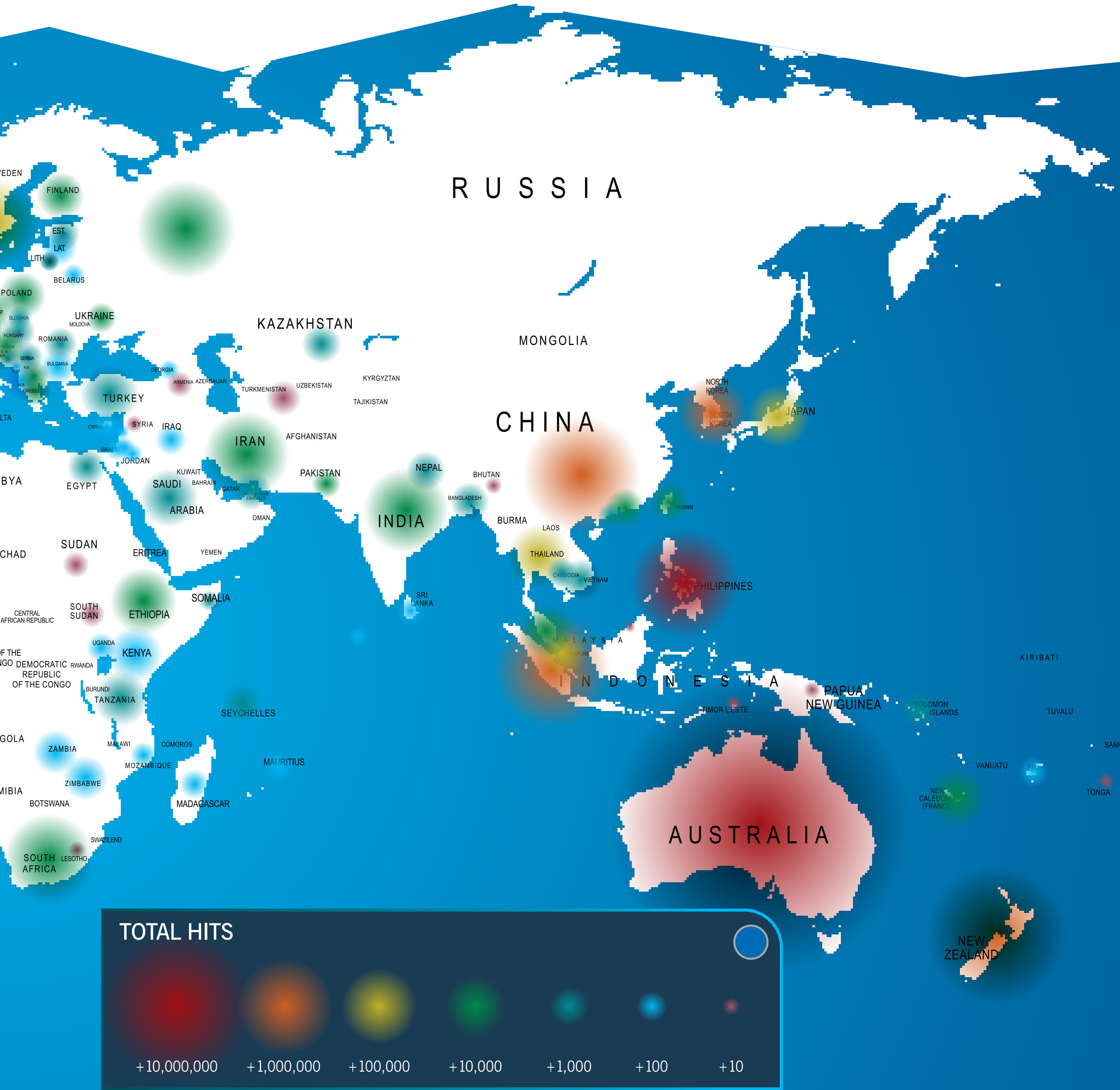
NCI also engages closely with the American Geophysical Union and European Geosciences Union throughout the year, helping guide their annual meetings and contributing to their busy programs and skills sharing. This fits into a broad range of data-related international engagement activities



featuring, for example, the European Space Agency, the United States Geological Survey, the United Kingdom Meteorological Office and others who all work with NCI to share data further and improve scientific outcomes.

On the high-performance computing front, NCI is a member of ADAC, the Accelerated Data Analytics and Computing Institute. Other members include the Swiss Supercomputing Centre, Oak Ridge and Argonne National Laboratories in the US, and RIKEN in Japan, all organisations at the pinnacle of extreme-scale supercomputing innovation around the world. Aiming to develop technologies and skills between collaborating organisations, ADAC, as well as the newly formed COVID-19 HPC Consortium of which NCI is a part, highlight that the future of supercomputing and big data technologies will rely on global cooperation.

One step closer to home, one of NCI's nearest regional analogues, the New Zealand eScience Infrastructure continues to be a valuable collaborator, partnering with NCI to deliver the Australasian Leadership Computing Symposium and other technical workshops and training sessions.



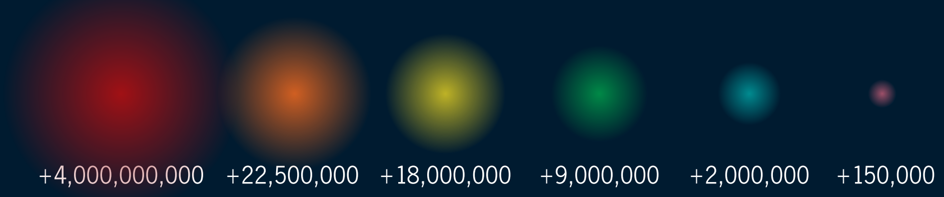
National Engagements

NCI is a central pillar of the Australian computational and data science space. We are at the forefront of national collaborations, working with key partners to further our shared ambitions and anticipate the research needs of our communities. Our primary technical partner for HPC and HPD is the Pawsey Supercomputing Centre, the Western Australian supercomputing facility providing computational and data support for the Square Kilometre Array astronomy project. We work closely with Pawsey on user training, technical development and events.

The other major national organisations we engage closely with are the national science agencies CSIRO, the Bureau of Meteorology and Geoscience Australia. Together they are responsible for the majority of Australia's peak scientific research. As Foundation Collaborators within the NCI

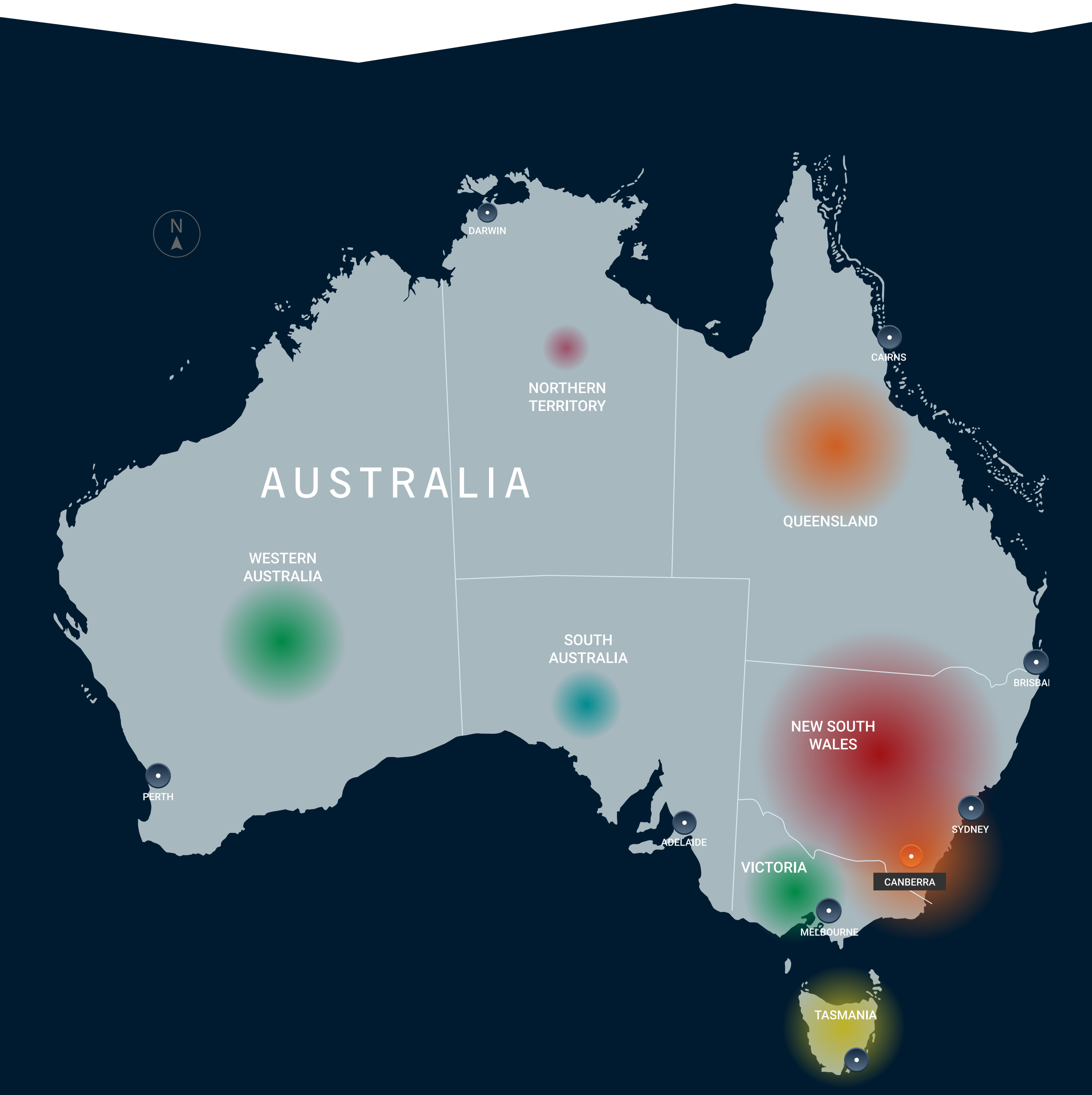
Region	Total Users	Total Hits	Download (MB)
New South Wales	15,270	194,056,455	**4,807,984,933**
Australian Capital Territory	7,213	73,038,770	**24,861,390**
Queensland	6,440	8,230,235	**23,295,884**
Tasmania	1,493	1,363,318	**18,924,021**
Western Australia	3,288	4,187,235	**14,321,426**
Victoria	8,200	5,115,045	**9,437,884**
South Australia	2,233	1,412,725	**2,294,437**
Northern Territory	293	162,010	**178,393**
(unknown)	28	67,970	**164,675**
Grand Total	**44,458**	**287,633,763**	**4,901,463,043**

TOTAL DOWNLOADS



Collaboration supporting and in turn relying on NCI's vast array of services, the ongoing close relationship we have built with them sets the stage for advances in their scientific programs. The other members of the NCI Collaboration include The Australian National University as the fourth Foundation Collaborator, and a host of other universities and research organisations as Major and Minor Collaborators. The NCI Collaboration brings together the country's leading research bodies, all contributing to support and access the computational infrastructure underpinning their research.

Funded under the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS), NCI also works closely with other NCRIS facilities including Pawsey, the Australian Research Data Commons, Bioplatforms Australia, AuScope, the Terrestrial Ecosystem Research Network and more. These national science infrastructure facilities support each other, building on their strengths to leverage even more value and impactful research capabilities for the nation.



A BIG DATA AND BIG COMPUTING LEADER

Financial report

PREAMBLE

NCI is an organisational unit of The Australian National University. The ANU, as represented by NCI, administers numerous funding contracts that support the operations of NCI. In the interests of providing a comprehensive picture of the NCI operation, a financial report consolidating these funding contracts is presented.

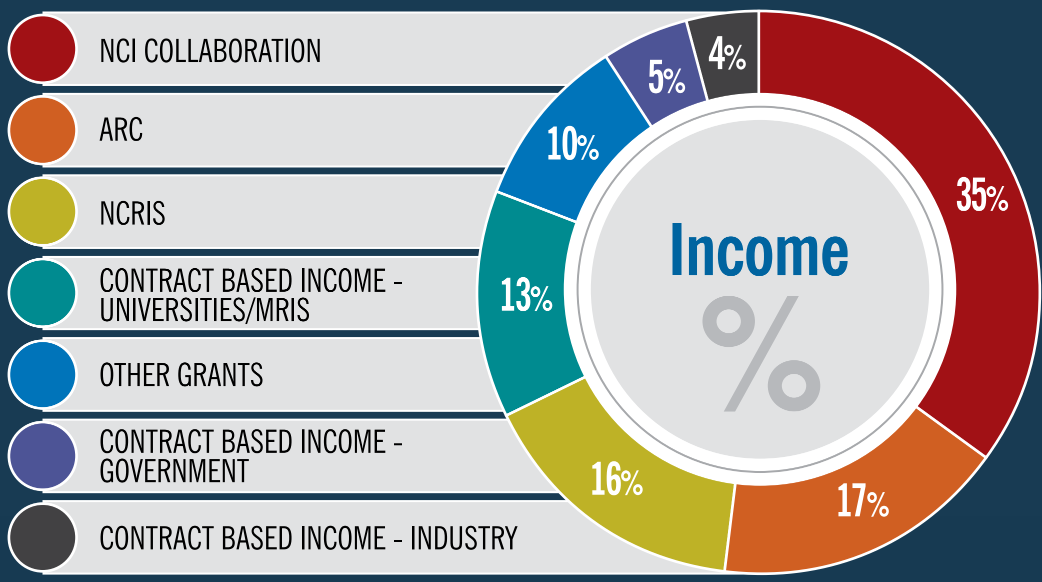
Capital funding provided for NCI's new supercomputer, Gadi, through the National Collaborative Research Infrastructure Strategy (NCRIS), has been noted separately.

STATEMENT OF INCOME AND EXPENDITURE

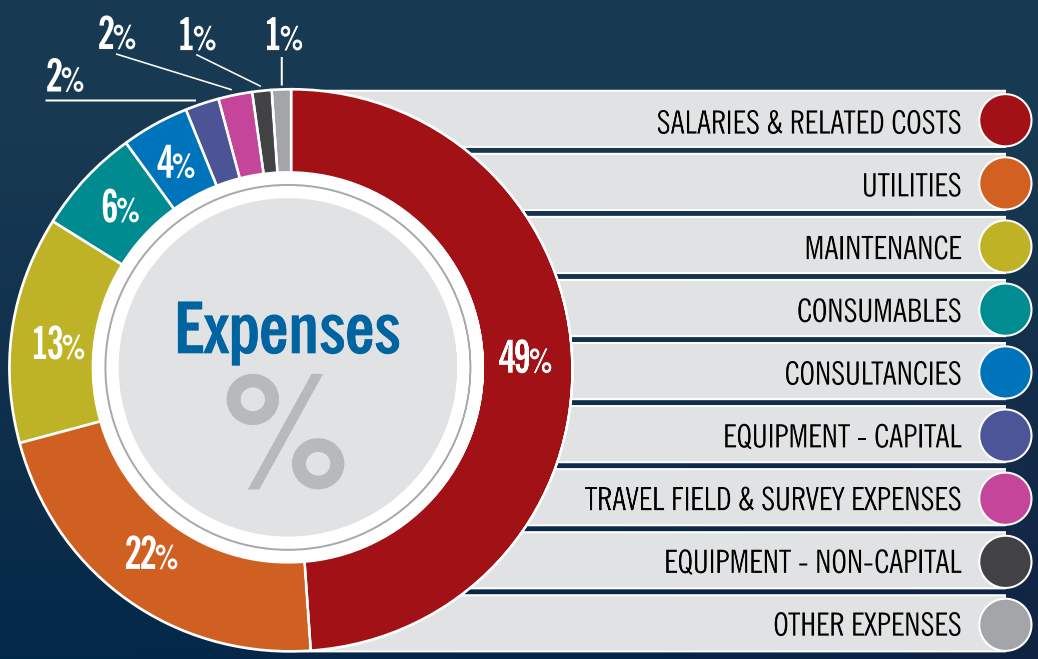
For the period 01 July 2019 to 30 June 2020

For the NCI collaboration and associated project accounts

	NCI $
Income	
NCI collaboration income	9,416,129
Other income	17,863,838
Total income	27,279,967
Expenditure	
Salaries and related costs	8,161,282
Equipment - capital	274,748
Equipment - non-capital	199,574
Utilities and maintenance	5,835,028
Travel, field and survey expenses	320,800
Expendable research materials	111
Consultancies	636,770
Consumables	915,413
Internal purchases	12,162
Other expenses	231,590
Total expenditure	16,587,479
Surplus / (Deficit)	10,692,488



NCI COLLABORATION
ARC
NCRIS
CONTRACT BASED INCOME - UNIVERSITIES/MRIS
OTHER GRANTS
CONTRACT BASED INCOME - GOVERNMENT
CONTRACT BASED INCOME - INDUSTRY
Income
%
35%
17%
16%
13%
10%
5%
4%



SALARIES & RELATED COSTS
UTILITIES
MAINTENANCE
CONSUMABLES
CONSULTANCIES
EQUIPMENT - CAPITAL
TRAVEL FIELD & SURVEY EXPENSES
EQUIPMENT - NON-CAPITAL
OTHER EXPENSES
Expenses
%
49%
22%
13%
6%
4%
2%
2%
1%
1%

Each funding contract is accounted for in a distinct account within the University ledger, and the University facilitates, and where appropriate, acts on, the NCI Advisory Board's directions and resolutions on NCI matters insofar as they are consistent with the relevant funding contract and not contrary to University Statutes and policies.

NCI COLLABORATION INCOME

The NCI Collaboration Agreement enables many of Australia's leading research intensive universities and science agencies to collectively fund a capability beyond the capacity of any single institution. Together, these institutions (including ANU, CSIRO, BoM, Geoscience Australia, the ARC, and a range of other research intensive universities and consortia) fund a significant proportion of NCI's operating costs.

A small but growing proportion of NCI Collaboration income comes from the commercial sector. Some income has not been expended in the current year and will be applied to capital expenditure in subsequent years to meet contractual obligations and emerging infrastructure requirements.

NCI operations are also supported by income received under the National Collaborative Research Infrastructure Strategy (NCRIS).

NCI administers a number of grants and contracts outside of the NCI Collaboration and NCRIS accounts. These special purpose arrangements fund clearly defined projects, infrastructure and services that provide synergistic benefits to the NCI Collaboration.

EXPENSES

NCI, as Australia's national research computing service, provides world-class, high-end services to Australia's researchers. In order to do this, NCI invests significant amounts of money in its expert team of staff and high-performance computing infrastructure – with these two factors driving NCI's expenditure profile.

Two factors combined to see a total expenditure level significantly less than prior years. The first one being, business-as-usual capital expenditure was minimal - the focus NCI's capital expenditure program in 2019/20 was the Gadi project, reported separately. The second, being the COVID-19 pandemic, resulted in the deferral of some expenditure.

REVIEW/AUDIT

Each funding contract held by the ANU as represented by NCI has specific financial reporting and auditing requirements, and NCI in conjunction with the University's Finance and Business Services Division and Corporate Governance and Risk Office acquit individual project funds in accordance with these requirements.

This consolidated statement has been reviewed by ANU's Finance and Business Services Division, which certifies that:

The statement accurately summarises the financial records of these grants and that these records have been properly maintained so as to accurately record the Income and Expenditure of these grants.

NCRIS SUPERCOMPUTER REPLACEMENT PROJECT

	$
Balance as at 1 July 2019	69,152,100
Income	
Other income	2,927,648
Total income	**2,927,648**
Expenditure	
Capital Expenditure	63,043,557
Non-capital Expenditure	726,769
Total expenditure	**63,770,326**
Unspent balance as at 30 June 2020	8,309,421

A BIG DATA AND BIG COMPUTING LEADER

Training

The strength of NCI's research offerings lies not only in its innovative systems and services but also in the creativity and ingenuity of its users. We endeavour each year to build the skills of our users so that they might use the resources available to them to their full potential. For users new to high-performance computing and data services, we guide them through introductory training modules to set them up with the skills they need to make a start with the new technologies. More experienced users have access to regular webinars about how to use NCI's data collections and data services in ways that can extend their research.

Importantly, wide ranging documentation about how to use NCI's wide range of services helps researchers get started and develop their HPC and HPD skills. This documentation features code snippets, interactive examples, descriptions of available resources, and guides for accomplishing standard tasks.

Effective training takes into account the requirements of the learners, which is why much of our training effort is conducted in concert with various research communities and their own needs for data or computing technologies. Targeted training in geophysical data, climate and weather models and broader data science use cases suit those scientists with discipline-specific requirements.

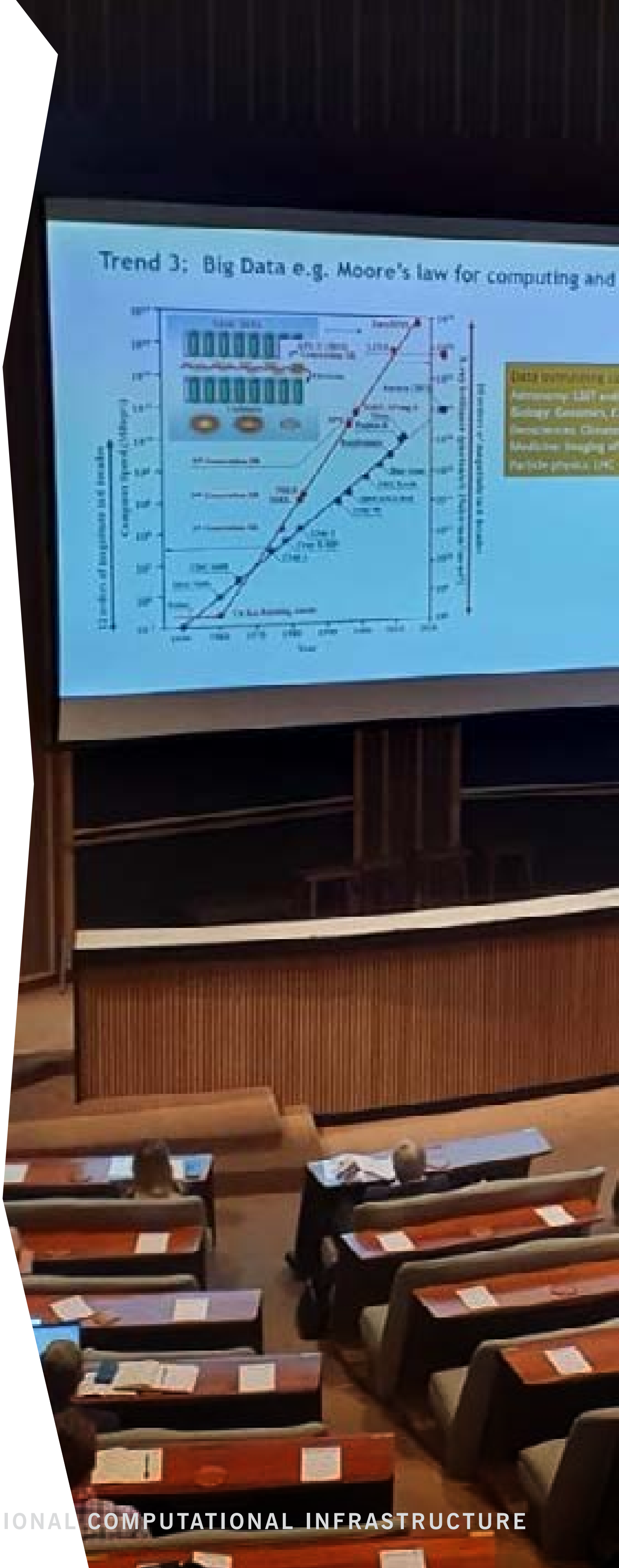
Outreach

NCI's highlight of outreach in 2019–20 was undoubtedly the inaugural Australasian Leadership Computing Symposium (ALCS), held in Canberra and hosted by NCI with support from the Pawsey Supercomputing Centre and the New Zealand eScience Infrastructure. This new and exciting event brought together 150 of Australasia's leading researchers for a mix of scientific and cross-disciplinary talks at the cutting edge of Australasian computational science.

This remarkable event was the first of its kind, and featured leading international and national speakers including Professor Peter Littlewood from the Faraday Institute, Professor Cathy Foley, Chief Scientist at CSIRO, and Anna-Maria Arabia, CEO of the Australian Academy of Science. ALCS received extremely positive reviews from the mix of researchers, research administrators and industry representatives who attended. Such a cross disciplinary event truly allowed for novel conversations to take place and new connections between and across scientific disciplines to emerge. It also provided an ideal forum for direct feedback from the scientists to the facilities about desires and wishes for the development of the HPC and HPD technology going forward.

In 2019–20, NCI also exhibited at several international and national supercomputing events including the International Supercomputing Conference in Frankfurt, Germany in July 2019, Super Computing in Denver, USA in November 2019 and eResearch Australasia in Brisbane in October 2019. These remain an important conduit for growing NCI's visibility among relevant scientists and professionals in the fields of big data and big computing. The international conferences, in particular, allow for NCI and the Pawsey Supercomputing Centre to exhibit together as HPC Australia, thus working together to highlight the integration of Australia's HPC centres. National conferences are a useful platform for engagement with other data-intensive NCRIS facilities. At eResearch Australasia, an inaugural NCRIS Avenue was conceived, connecting all exhibiting NCRIS facilities under one banner and showcasing the diversity and breadth of that program.

We are lucky every year to be able to host student groups from the National Youth Science Forum and other science summer schools. Canberra's fiery summer, and then the COVID-19 pandemic, unfortunately led to the cancellation of many of this year's planned visits and events. Nevertheless, NCI continues to produce and share educational resources, easily accessible research stories and scientific visualisations to interest and delight students and other visitors to NCI.



Professor Peter Littlewood presenting at the Australasian Leadership Computing Symposium at the Australian Academy of Science's Shine Dome in Canberra.

The NCI and Pawsey Supercomputing Centre booth at the International Supercomputing Conference in Frankfurt, Germany.

The NCI Collaboration

SUPPORTED BY

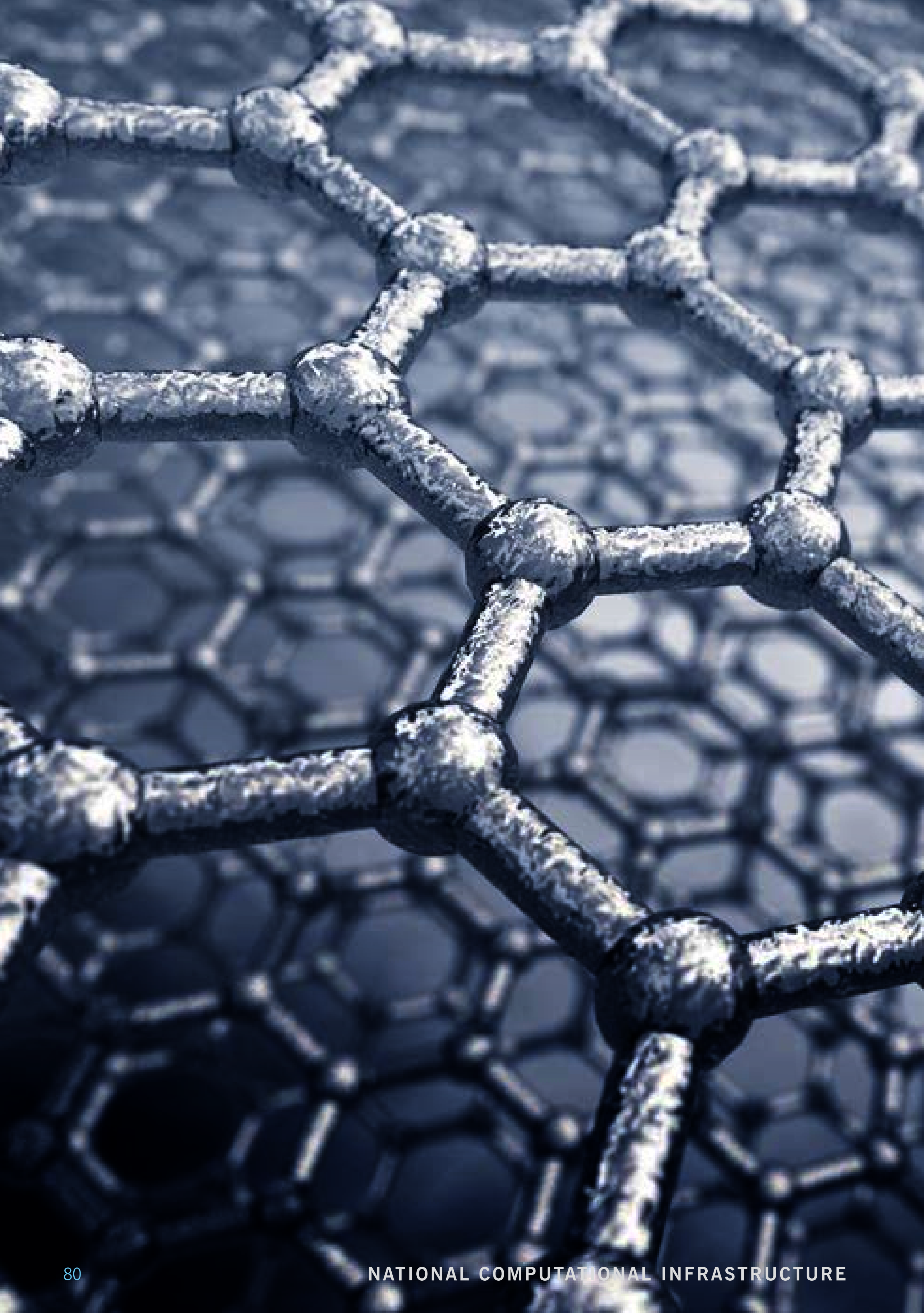
MAJOR COLLABORATORS

COLLABORATORS

OTHER CONTRACTS

MERIT FLAGSHIPS

Our Vendors



4 APPENDIX

NCI Links to Government Portfolios

Government Department Impacted	Programme/Agency Impacted	Activities/Projects supported by NCI
Education, Skills and Employment	Australian Research Council (ARC)	Dependencies from more than 530 projects funded by ARC's National Competitive Grant Programs (NCGP) underpinning in excess of $606m in research investment
	National Collaborative Research Infrastructure Strategy	Support for services provided by numerous other NCRIS capabilities (AuScope, TERN, ALA, ARDC, BioPlatforms Australia)
		Close collaboration with Australian Research Data Commons and Pawsey Supercomputing Centre
Industry, Science Energy and Resources	CSIRO	Australian Community Climate and Earth System Simulator (ACCESS)
		Earth Systems and Climate Science (ESCC) Hub of the National Environmental Science Program
		Climate and Weather Science Data-Enhanced Virtual Laboratory
		Marine Virtual Laboratory
		AuScope Virtual Research Environments Geoscience Data Enhanced Virtual Laboratory
		CMIP6 Climate Dataset
	CSIRO and the Australian Institute of Marine Science (AIMS)	eReefs
	Geoscience Australia (GA)	Digital Earth Australia
		Copernicus Data Hub
		National Reference Grid
		Exploring For The Future Initiative
		Water Observations from Space (WOfS)
		Australian Natural Hazards Data Archive
		AuScope Virtual Research Environments Geoscience Data Enhanced Virtual Laboratory

Government Department Impacted	Programme/Agency Impacted	Activities/Projects supported by NCI
Agriculture, Water and the Environment	National Environmental Science Programme (NESP)	Earth Systems and Climate Science Hub
	Environmental policy development	eReefs
		Coupled Model Intercomparison Project (CMIP)
	Bureau of Meteorology	Australian Community Climate and Earth System Simulator (ACCESS)
		ESCC Hub of the NESP
		Climate and Weather Science Data Enhanced Virtual Laboratory
		Marine Virtual Laboratory
		BARRA Reanalysis
		Bushfire Model Development
	Australian Antarctic Division	Antarctic Climate and Environment CRC (ACE-CRC) ACCESS Southern Ocean and cryosphere models
	Policy development for the agricultural industry and water resources	ACCESS development with BoM/CSIRO
		Development and hosting of Digital Earth Australia Collaboration with GA and CSIRO
		Hosting and curation of Water Observations from Space data with GA
	Murray-Darling Basin Authority	Exploitation of Digital Earth Australia data
		Development of water-related Digital Earth Australia capabilities in collaboration with GA
Health	National Health and Medical Research Council	Dependencies from more than 64 NHMRC funded projects and fellowships
		Australian Genomics Health Alliance (AGHA)
		Support for COVID-19 research projects
Defence	Australian Geospatial Intelligence Organisation (AGIO)	Onshore topographic data and products provided by GA
	Australian Hydrographic Service (AHS)	Raw and processed bathymetric data collections provided by GA
Foreign Affairs and Trade	Policy development for, and by, the tourism sector	eReefs (through CSIRO)
		International Collaboration development (ESGF, ADAC, NeSI)

Government Department Impacted	Programme/Agency Impacted	Activities/Projects supported by NCI
Infrastructure, Transport, Regional Development and Communications	Australian Marine Safety Authority (AMSA)	Managing risks to marine vessels in Australian waters (undertaken with consultant DHI)
	Aviation Programs	Development of specialist weather reporting products for the aviation industry (with BoM)
	Transport Infrastructure Programs	National Reference Grid

Outreach

Tours and events NCI participated in and hosted in 2019–20.

Group	Date
Science Technology Australia Communication Workshop	25 July 2019
Australian Synchrotron Director Professor Andrew Peele Visit	3 September 2019
Geoscience Australia Staff Visit	11 September 2019
Group of 8 Deputy Vice Chancellor (Research) Meeting and NCI Visit	2 October 2019
UNSW Deputy Vice-Chancellor Professor Nicholas Fisk Visit	2 October 2019
De Beers Delegation Visit	10 October 2019
Australasian Leadership Computing Symposium – Hosted by NCI	5-8 November 2019
National Centre for Indigenous Genomics Visit	11 November 2019
SuperComputing Conference, Dallas, USA	18-21 November 2019
Research School of Earth Sciences Japanese Researchers Visit	2 December 2019
Michael Pettersson ACT MLA Visit	17 December 2019
Australian Phenomics Network Board Meeting at NCI	10 February 2020
Intersect Interim CEO Mr Satish Nair Visit	19 February 2020
Rowan Ramsey Federal MP Visit	21 February 2020
Helmholtz International Lab Visit	27 February 2020
International Women's Day Science in the Pub	6 March 2020