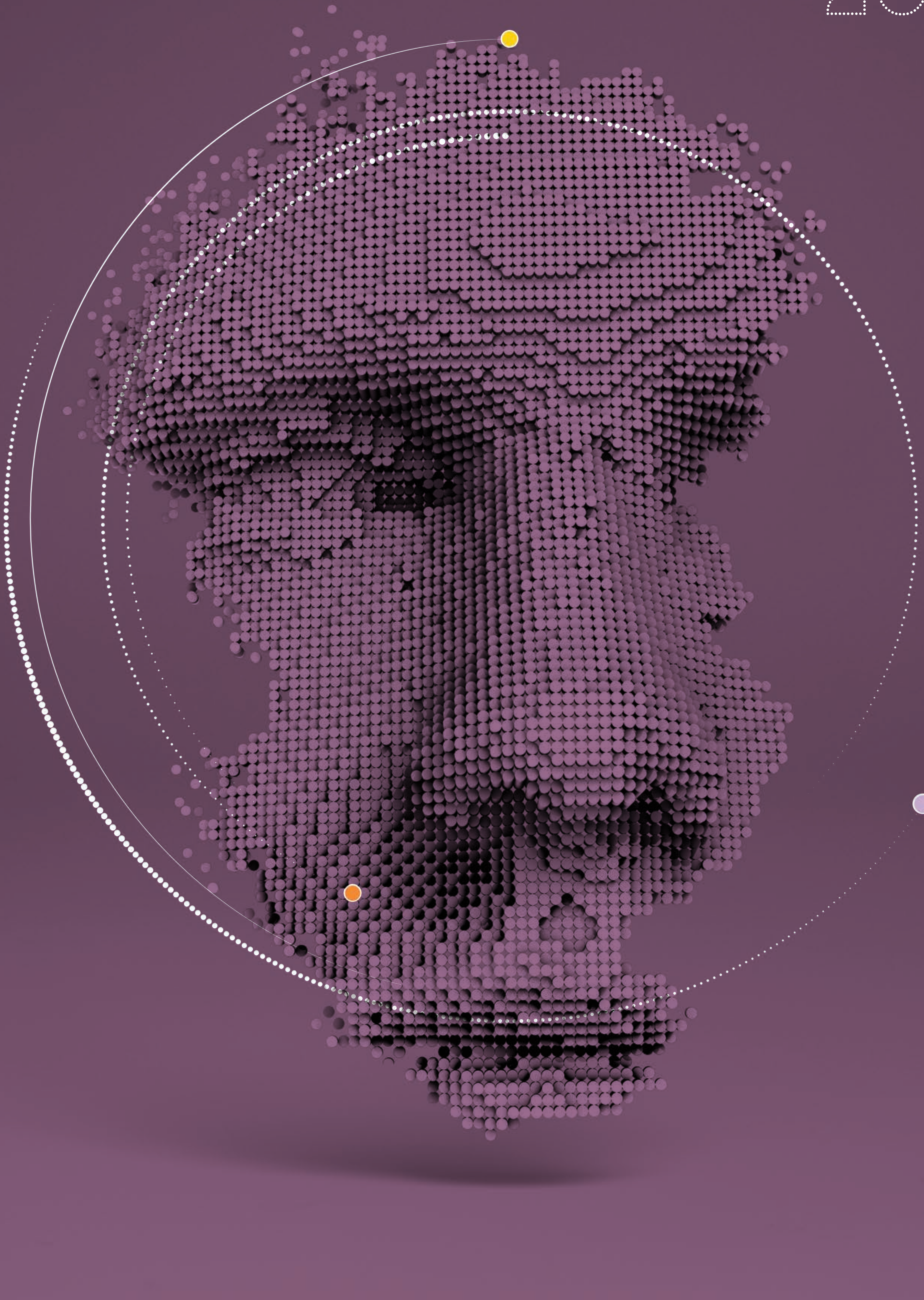
NeSI
New Zealand eScience Infrastructure

2018



NeSI
New Zealand eScience Infrastructure

MINISTRY OF BUSINESS, INNOVATION & EMPLOYMENT HĪKINA WHAKATUTUKI
THE UNIVERSITY OF AUCKLAND NEW ZEALAND
NIWA Taihoro Nukurangi
UNIVERSITY OF OTAGO Te Whare Wānanga o Otāgo NEW ZEALAND
Manaaki Whenua Landcare Research

CHAIR'S REPORT

NeSI is becoming a recognised leader in facilitation and capability sharing, and the team has the values and the heart to engender trust and collaboration sector-wide.

As 2018 closes, we have in place our first nationally co-designed and shared computational research infrastructure, hosted at NIWA's purpose-built Greta Point HPC Facility with a replicated site at the University of Auckland's Tamaki Data Centre. Researchers are now making great use of its new scale, performance, and advanced technologies, to power their research. Early feedback is of a step change in the capabilities on offer, and of the science we can now approach from within New Zealand.

Looking back over 2018, I'm reminded how far we've come over the last 9 years. We started as a network of supercomputers, operating severally, providing national access to separate and in some cases pre-existing assets. We established NeSI on the principle of collaboration, as an embedded team operating from within partner institutions. We built NeSI's foundations from the capabilities of the partners, with investment into national leadership and outreach to broaden our impact.

I credit much of our success to our strong management in NeSI, who have put in place a truly collaborative culture within NeSI, and exhibit the best of these qualities out to the wider sector. As leaders in a space of ever-changing and bleeding edge technologies, NeSI's team are dealing day to day with the most demanding of our science needs for computing, and doing so across institutional and community boundaries. Not to put too fine a point on it, NeSI is an asset to the New Zealand science system.

As we look to the next phase of NeSI there are many opportunities in front of us. We should look to further develop NeSI alongside the ongoing evolution of the science system as a whole. We can see from our international benchmarks that there are significant gaps in our science system which are constraining our ambitions. While our computational mindsets are first rate, we have much more limited skill sets in the advanced computational technologies essential to modern science. If we're to achieve desired levels of productivity and impact in the science system, we should look to NeSI as part of our national investment supporting broader science system transformation.

Rick Christie

Rick Christie
Chair, Board of Directors

> "As we look to the next phase of NeSI there are many opportunities in front of us. We should look to further develop NeSI alongside the ongoing evolution of the science system as a whole."

DIRECTOR'S REPORT

Looking back, 2018 was a year of consolidation for NeSI together with our partners the University of Auckland, NIWA, University of Otago and Manaaki Whenua – Landcare Research.

- We now have the underpinning infrastructure to support a new era of advanced computational research in New Zealand.
- We've revisited our partnerships, seeking stronger alignment with our collaborators, science communities, and our science agency to enable our shared ambitions.
- We've reviewed eResearch ecosystems across comparator science systems internationally, and are arriving at an ambitious vision for our science sector, powered by NeSI.

A new era of advanced computational research
Working in partnership with global HPC and AI leader Cray, we've created the most advanced, creative and high-performance computing and data platform for science ever seen in New Zealand. Together we've connected communities and integrated technologies to create a truly innovative and powerful environment for research at the leading edge of capability internationally.

Across 2018 we maintained our strong position and reputation for focusing on researchers' needs, while identifying key pain points in their computational journeys. Our challenge from here is to take these advanced technologies and deliver them more effectively to New Zealand scientists, to enable a shift in the horizons and frontiers of science, and to lift science excellence and impact.

A strengthened partnership to enable our ambitions
Looking back across the year and the preceding years with our partners, it is clear in retrospect that NeSI has grown beyond its initial ambitions. Working with our science communities, we've sought out and responded to the challenges and needs of researchers, and in doing so adapted NeSI to be better able to support their science. Working together as a partnership we've evolved NeSI, in significant part by enhancing the existing expert capabilities of our partners for the benefit of the science system as a whole.

As with any partnership it is important to stay close, to recognise and respect the heritage and capabilities each offers, and to map out a sustainable pathway to realising our shared ambitions. In short, NeSI and its partners see a renewed and more strategic commitment to the partnership as an essential underpinning to our ongoing successes.

An ambitious vision for our science sector, powered by NeSI
As the year closes, our vision for the future is shaping up within our SSIF investment case, which lays out our strategies for our third round of investment starting from 1 July 2019. We'll continue to work alongside MBIE's SSIF team across 2019 to inform policy directions as they look to provide a roadmap for future eResearch investments.

Meanwhile, we're facilitating a discussion with local and international experts, and shining light into the current health of our eResearch ecosystems. Across this ecosystem we've mapped where NeSI has been, and where we see major risks, challenges, gaps, and opportunities in our current situation. This informs our future strategies, underpinning how we position NeSI as a facilitator of sector transformation in adopting advanced computational technologies.

Nick Jones
Director

Advancing New Zealand's digital capabilities

NeSI empowers New Zealand researchers to tackle large or highly complicated problems and to investigate scientific challenges that were previously impossible.

Enabling New Zealand's science sector to compete and excel globally

Supporting research today and tomorrow requires an inclusive partnership with New Zealand researchers, communities, and Te Ao Māori, underpinned by a specialised and powerful technology ecosystem. NeSI is collaborating with research communities across domains and across the country to strengthen their computational capabilities and enable New Zealand's science sector to compete and excel globally.

As computational capabilities develop, research becomes possible at larger scales, with finer-resolutions, and addressing greater complexity. Months of work can be optimised into weeks, days, or even hours through NeSI's research software & data engineering contributions. NeSI is broadening its sector uptake and enhancing computational capabilities through training programmes, consultancy in research software engineering, and recruitment of NeSI team members with stronger knowledge of and connection into research communities.

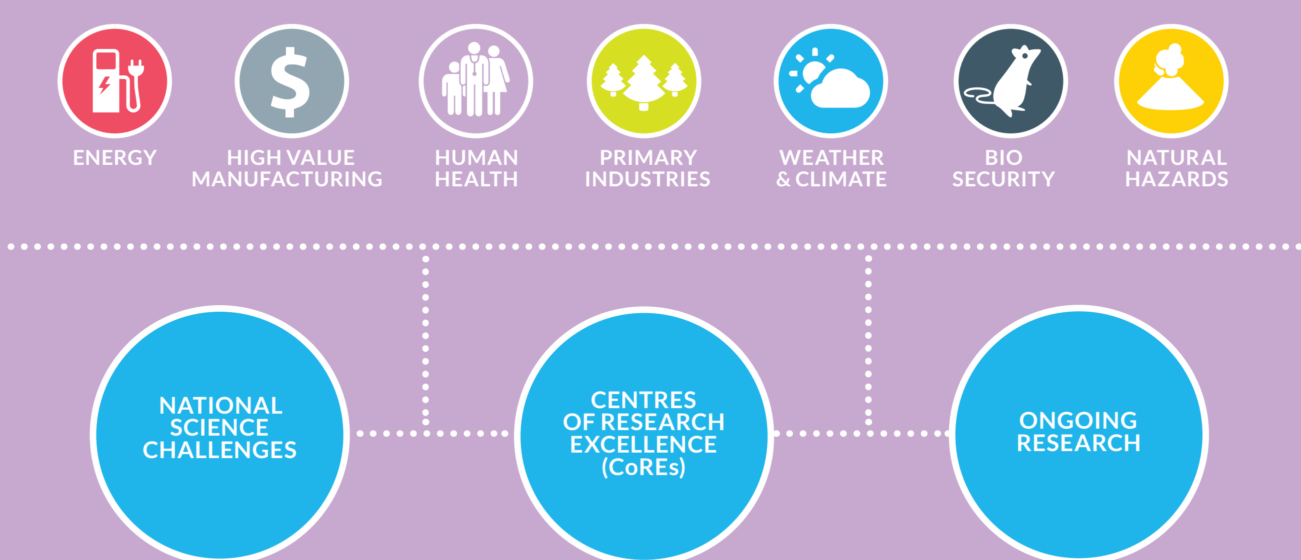
NeSI is contributing national capabilities in:

- Compute Infrastructure
- Storage Infrastructure
- Research Artifact Management
- Analysis and Modelling Tools
- Research Platforms, Virtual Labs, Scientific Gateways

NeSI is enabling others through:

- Partnering with Science
- Service and Resource Governance
- Data, Methods, Models, Code
- Research Software and Data Engineering
- Advice, Training, Outreach
- Cyber security

NeSI grows computational research capabilities and skills

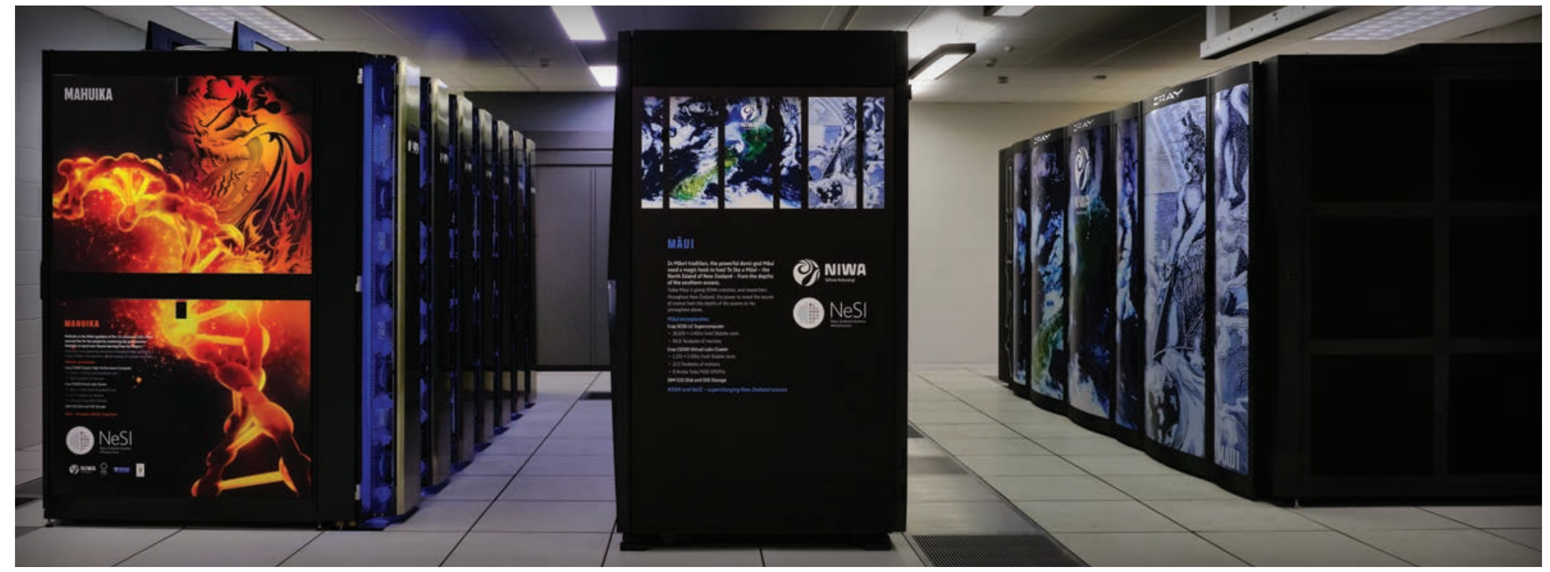


NeSI's enhanced services, enabled by new infrastructure capabilities, are creating new opportunities for researchers.
Growth in data, increasing complexity of models, and an increasing diversity of research drivers are fuelling new growth in demand for digital capabilities within New Zealand's research sector.

Building upon its existing platforms and services, NeSI is developing new and enhanced services that will leverage its raw computing power and integrate tools like virtual lab environments, visualization capabilities, and other essential data management services.

Each of these new service areas enables more varied modes of access and use, and broadens NeSI's impact and performance.

NeSI's investments



NeSI's new infrastructure at NIWA's Greta Point campus in Wellington (photo by Michael Uddstrom)

NeSI's investments in advanced research computing are driven by and co-designed with research communities.
On the 7th of November 2018, Research, Science and Innovation Minister Dr Megan Woods officially opened the new High Performance Computing Facility for New Zealand hosted by NeSI and NIWA. The facility is comprised of three new interconnected supercomputers, two of which are supported by and available through NeSI to meet the needs of New Zealand researchers to investigate scientific issues of national significance.

Consolidated, highly-integrated, and fit-for-purpose, this new infrastructure was co-designed with key stakeholders following an assessment of national research needs. It lifts NeSI's maturity in key ecosystem areas, especially those related to research data, in research platforms and virtual labs, and in security. The new platform represents a culmination of years of collaborative work, led by NeSI with its investors, to coordinate its procurement, design, and launch.

All researchers in New Zealand have access to these specialised high performance computing (HPC) resources, including tailored software environments and data management services, and skills training in computational research tools and approaches.

NeSI core hour availability and usage

In terms of available core hours, NeSI's new platforms provide a significant increase in capacity for the research sector. Use of NeSI has increased over time, and the new platforms provide significant headroom for broadened sector uptake and growth in use in the years to come.

"This marks a step change for science in New Zealand and a further advancement towards an innovative, future-focused society. The supercomputers are a significant upgrade with 10 times the computing capability of their predecessor. This will have a whole range of benefits for scientific research, including better understanding issues around climate change, genomics, the management of New Zealand's freshwater assets and resilience to natural hazards."

Dr Megan Woods, Research, Science and Innovation Minister

Finding genetic help for Māori health

Dr Tony Merriman is a research professor at the University of Otago. There, he is studying the genes of New Zealand and Cook Island Māori to understand why Māori people have one of the highest rates of gout and hyperuricemia in the world.

"Our current project maps 130 genomes of Māori people. Ancestrally, Māori are Eastern Polynesian and we have very little information on the background genetic structure of those populations. We need that information to help us do our research," Tony said.

Hyperuricemia is a condition where too much urate is in the blood. Over time, this urate can crystallise, leading to a number of painful and potentially fatal diseases, including gout.

"With genetics research we may be able to identify genes causing hyperuricemia or gout itself. The pharmaceutical industry can assess whether those genes are valid targets for new drugs or repurposed existing drugs."

A common trigger of gout attacks is diet – particularly purine-rich foods such as alcohol, meat and seafood. However, this doesn't explain the high frequency of sufferers among Māori people – which can be up to double the rate of European-ancestry New Zealanders.

"We know it goes back to pre-colonisation, so it's not just our modern diet," Tony said. "One theory is having elevated urate levels was a defence mechanism against malaria. But we don't really know. You can come up with ideas but they're difficult to test."

The genetics research for these drugs relies on mapping human genomes. A genome is the total set of genetic information within each human – similar to a blueprint for that person. Each genome is unique, but people with a shared ancestry will have more genetic information in common than those more distantly related.

This genetic variation becomes important when designing medicine because differences in genes can change how the medicine reacts to a person. A medicine which helps one person may harm another. Most medicines are designed for people of European ancestry, which can make treatment of people with different ancestries, such as Māori people, difficult.

"There are hundreds of thousands of genomes mapped for Europeans. There are less than two hundred genomes of Māori people that have ever been done in a medical context. What we're aiming for is to have equivalent research done as to other populations," Tony said.

It is difficult to compare exact numbers of genomes mapped for European and Māori ancestries due to a lack of data. However, there is a widely-acknowledged bias towards European genome mapping in medical research. Tony's team are hoping to counteract this bias.

The sequencing of these genomes is done in Australia, at the Garvan Institute of Medical Research. There, DNA is broken up into small snippets and read by a gene sequencer, before being sent back to Tony's team to reassemble using NeSI's Mahuika supercomputer.

Finally, the mapped genomes are compared to what is called a reference genome. Comparison helps Tony's team spot differences between Māori and other ancestries' DNA. Once they've compared the genomes, the next step is to find out what those different segments control in the human body.

"We want to figure out what the genetic variants do in a biological context. What are they doing in the body? We're going to start moving more into functional work: what and when they're expressed; how the variants regulate biological process. Interpreting them is challenging."

For many Māori people, gout is a painful inheritance. It takes close cooperation with Māori communities, passionate researchers like Tony and resources like NeSI's Mahuika supercomputer to provide better treatment to those at risk. Working together may help improve the lives of many people.

"We're using NeSI to allow us to do gene alignments – sequencing millions of small segments of DNA."

Tony Merriman, Professor, University of Otago

A national collaboration

NeSI New Zealand eScience Infrastructure

Collaborating with
MINISTRY OF BUSINESS, INNOVATION & EMPLOYMENT HĪKINA WHAKATUTUKI · THE UNIVERSITY OF AUCKLAND NEW ZEALAND · NIWA Taihoro Nukurangi · UNIVERSITY OF OTAGO · Manaaki Whenua Landcare Research

Foundations in partnerships and collaboration
NeSI is a strategic investment into a national science infrastructure platform with anticipated long-term beneficial impact on New Zealand's health, economy, environment and society. NeSI works nationally and locally with research institutions and research communities as a facilitator and catalyst to grow researcher computational capabilities and skills for future productivity and wellbeing. These activities are underpinned by the nation's most advanced computation and data science facility operating on assets owned by NeSI's Collaborators. Together, we are building high-performing science and innovation systems that transform New Zealand into a more diverse, technologically advanced, and smart nation.

NeSI-at-a-glance

The power behind researchers

*Data replicated in Auckland

Disciplines Supported

Biology · Engineering · Astronomy · Physics · Computer Science · Medical Science · Earth Science · Social Science · Mathematics

Core Services

High Performance Computing & Analytics · Consultancy · Training · Data Transfer & Share

Our Infrastructure

Māui · Mahuika

>136 million	>1.7 petaflops	>130 GB/s
CPU core hours available per year	peak performance	IO bandwidth

Governance

Board of Directors
NeSI's Board of Directors is comprised of an independent Chair, an independent member representing the research sector at large, and three Directors appointed by NeSI's four collaborator institutions – the University of Auckland, NIWA, Manaaki Whenua - Landcare Research, and the University of Otago. The Board approves major initiatives and investments, and provides oversight on strategy, policy, and organisational development.

In 2018 the NeSI Board was comprised of the following members:

- Rick Christie, Chair, Independent
- Dr Fiona Carswell, Chief Scientist, Manaaki Whenua - Landcare Research
- Dr Murray Poulter, former Chief Scientist, NIWA
- Prof. Andrew Rohl, Professor of Computational Science, Curtin University, Independent
- Stephen Whiteside, Chief Digital Officer, University of Auckland
- Justine Daw, General Manager, Manaaki Whenua - Landcare Research


Rick Christie
Chair, Independent Director


Dr Fiona Carswell
Chief Scientist, Manaaki Whenua – Landcare Research


Dr Murray Poulter
Chief Scientist, Atmosphere, Hazards and Energy NIWA (Retired)


Prof. Andrew Rohl
Professor of Computational Science, Curtin University, Independent


Stephen Whiteside
Chief Digital Officer, University of Auckland


Justine Daw
General Manager, Manaaki Whenua – Landcare Research

Research Reference Group
NeSI's Research Reference Group is made up of nine research community leaders with strong technical knowledge on the impact of eScience as an enabler of research. The group has a key role in providing timely advice and input on strategic and policy matters of interest to NeSI, especially those most relevant to researchers.

- Dr Joseph Lane, Chair, Associate Professor (Physical & Theoretical Chemistry), University of Waikato
- Dr Jane Allison, Associate Professor, Biological Sciences, University of Auckland
- Dr Blair Blakie, Professor, Department of Physics, University of Otago
- Murray Cadzow, Teaching Fellow and Postdoctoral Fellow, Department of Biochemistry, University of Otago
- Dr Barbara Chapman, Professor, Institute for Computational Science, Stony Brook University
- Dr Sam Dean, Chief Scientist - Climate, Atmosphere and Hazards, NIWA
- Dr Ian Foster, Professor, Department of Computer Science, University of Chicago
- Dr Nauman Maqbool, Group Leader Knowledge & Analytics, AgResearch
- Dr Cristin Print, Professor, Molecular Medicine & Pathology, University of Auckland


Dr Joseph Lane
Chair, Associate Professor (Physical & Theoretical Chemistry), University of Waikato


Dr Jane Allison
Associate Professor, Biological Sciences, University of Auckland


Dr Blair Blakie
Professor, Department of Physics, University of Otago


Murray Cadzow
Teaching Fellow and Postdoctoral Fellow, Department of Biochemistry, University of Otago


Dr Barbara Chapman
Professor, Institute for Computational Science, Stony Brook University


Dr Sam Dean
Chief Scientist - Climate, Atmosphere and Hazards, NIWA


Dr Ian Foster
Professor, Department of Computer Science, University of Chicago


Dr Nauman Maqbool
Group Leader Knowledge & Analytics, AgResearch


Dr Cristin Print
Professor, Molecular Medicine & Pathology, University of Auckland

Optimising tools used for genetic analysis

> “NeSI helped our team profile our code to detect inefficiencies and helped improve our code, allowing us to run our analyses faster.”
>
> **Ken Dodds, Statistician/Geneticist, AgResearch**

Knowledge of genetic relatedness is important in a number of genetic applications. Genetic relatedness estimates can be used to infer the relationship between individuals (e.g. mother-son, siblings) and can also be used by animal and plant breeders to predict the characteristics of offspring.

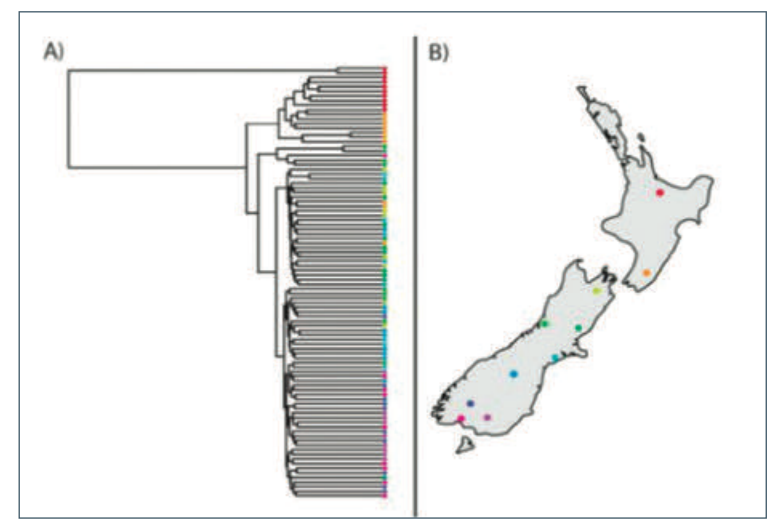
Related to this, genetic linkage maps-one-dimensional representations of genetic inheritance across the genome - are also valuable as they form the starting point for a number of downstream analyses. NeSI's computational science team has been working with Timothy Bilton, a PhD student in the Department of Mathematics and Statistics at the University of Otago and at AgResearch, and Ken Dodds, a Statistician / Geneticist at AgResearch, to incorporate new computational methods and tools into their analysis of genetic relatedness and construction of genetic linkage maps. “Both these applications can have high computing resource requirements with 10,000 to 1 million genomic positions and 100 – 100,000 individuals in a study, depending on the particular case,” says Ken.

“Chris Scott and Wolfgang Hayek from NeSI helped our team profile our code to detect inefficiencies and helped improve our code, allowing us to run our analyses faster.” In one test case, their code ran nearly 20 times faster (24 mins vs 8 hrs) after they implemented the code optimisation and computational tools recommended by NeSI.

Many of Timothy and Ken's projects use genotyping-by-sequencing (GBS), a relatively new sequencing technology that enables many samples to be sequenced at high-throughput. The advantages of GBS over previous technology is that it is significantly cheaper and faster, and is easily adapted to a diverse range of species. “GBS allows researchers to obtain genotype information across the genome of species that previously were not possible due to limited budgets, such as New Zealand native species,” says Timothy.

“However, data generated by GBS contains additional errors that can severely affect the reliability of the results if they are not taken into account.” Consequently, they need to use specialised software packages that can handle the nuances of GBS data in particular contexts. One of those is KGD, a package developed for estimating the genetic relatedness between individuals in a population. KGD can have an important role of performing quality control of GBS data, and a common application of KGD software is to check the accuracy of a recorded pedigree. Another package, GUSMap, uses GBS data to construct genetic linkage maps. Errors in these maps can negatively impact the reliability of downstream analyses, so GUSMap is designed to handle any errors in the data and enable researchers to construct accurate linkage maps from GBS data at a significantly reduced cost compared to other sequencing technologies. To improve Timothy and Ken's productivity with these packages, NeSI introduced them to tools such as Rcpp, used for converting sections of R code into C++ for higher performance, and to OpenMP, an application programming interface for parallelisation (using multiple processors simultaneously to calculate results faster).

The support from NeSI has already benefited Timothy across multiple projects. “Just recently, I managed to implement some OpenMP parallelisation in another R package I'm working on, so that has been a big help,” he says. “It would have taken a lot time to work out how to use Rcpp and OpenMP due to our limited programme knowledge. Chris and Wolfgang have saved us a great deal of time and pain trying to learn how to use these techniques to improve the efficiency of our code.”



Application of KGD to infer population structure in NZ black-billed gulls. (A) Dendrogram of individual gull samples clustered according to relatedness estimates computed from KGD. Colours correspond to the locations of samples. Mischler et al. (2018)

NeSI team

NeSI is powered by a high-performance, nationally-distributed, and multi-disciplinary team empowered to accelerate research outcomes across New Zealand.

Robin Bensley
Operations Manager,
University of Auckland

Blair Bethwaite
Solutions Manager,
University of Auckland

Thomas Berger
Product Management and
Software Engineering Lead,
University of Auckland

Fabrice Cantos
Site Manager,
NIWA

Laura Casimiro
Operations Coordinator,
University of Auckland

Brian Flaherty
Data Services Product Manager,
University of Auckland

Marcus Gustafsson
Site Manager,
University of Auckland

Greg Hall
Systems Engineer,
University of Auckland

Yuriy Halytskyy
Systems Engineer,
University of Auckland

Wolfgang Hayek
Scientific Programmer,
NIWA

Matt Healey
Application Support Specialist,
University of Otago

Aaron Hicks
Systems Engineer,
NIWA

Jose Higino
Systems Engineer,
NIWA

Jun Huh
Business Innovation
and Growth Manager,
University of Auckland

Nick Jones
Director,
University of Auckland

Anita Kean
Analyst Programmer,
University of Auckland

Mike Ladd
Strategic Projects Manager,
University of Auckland

Nancy Lin
Data Analyst,
University of Auckland

Nooriyah Lohani
Research Communities Advisor,
University of Auckland

Jana Makar
Communications Manager,
University of Auckland

Peter Maxwell
Application Support Specialist,
University of Auckland

Alexander Pletzer
Scientific Programmer,
NIWA

Nitharsan Puwanendran
Analyst Programmer,
University of Auckland

Georgina Rae
Engagement Manager,
University of Auckland

Kumaresh Rajalingam
Analyst Programmer,
University of Auckland

Ben Roberts
Application Support Specialist,
Manaaki Whenua –
Landcare Research

Albert Savary
Application Support Specialist,
University of Otago

Mandes Schönherr
Application Support Specialist,
NIWA

Chris Scott
Scientific Programmer,
University of Auckland

Dinindu Senanayake
Application Support Specialist,
University of Auckland

Anthony Shaw
Application Support Analyst,
University of Auckland

Nick Spencer
Site Manager
Manaaki Whenua –
Landcare Research

Michael Uddstrom
Platforms Manager,
NIWA

Callum Walley
Application Support Analyst,
University of Auckland

Shen Wang
Analyst Programmer,
University of Auckland

Damian Wheeler
Site Manager,
University of Otago

Organisational health

Building a multi-institutional, multi-disciplinary team like NeSI takes time. The NeSI team continued to grow in 2018 to support our expanding user base and new infrastructure, as well as bring in new capabilities across multiple teams and service areas.

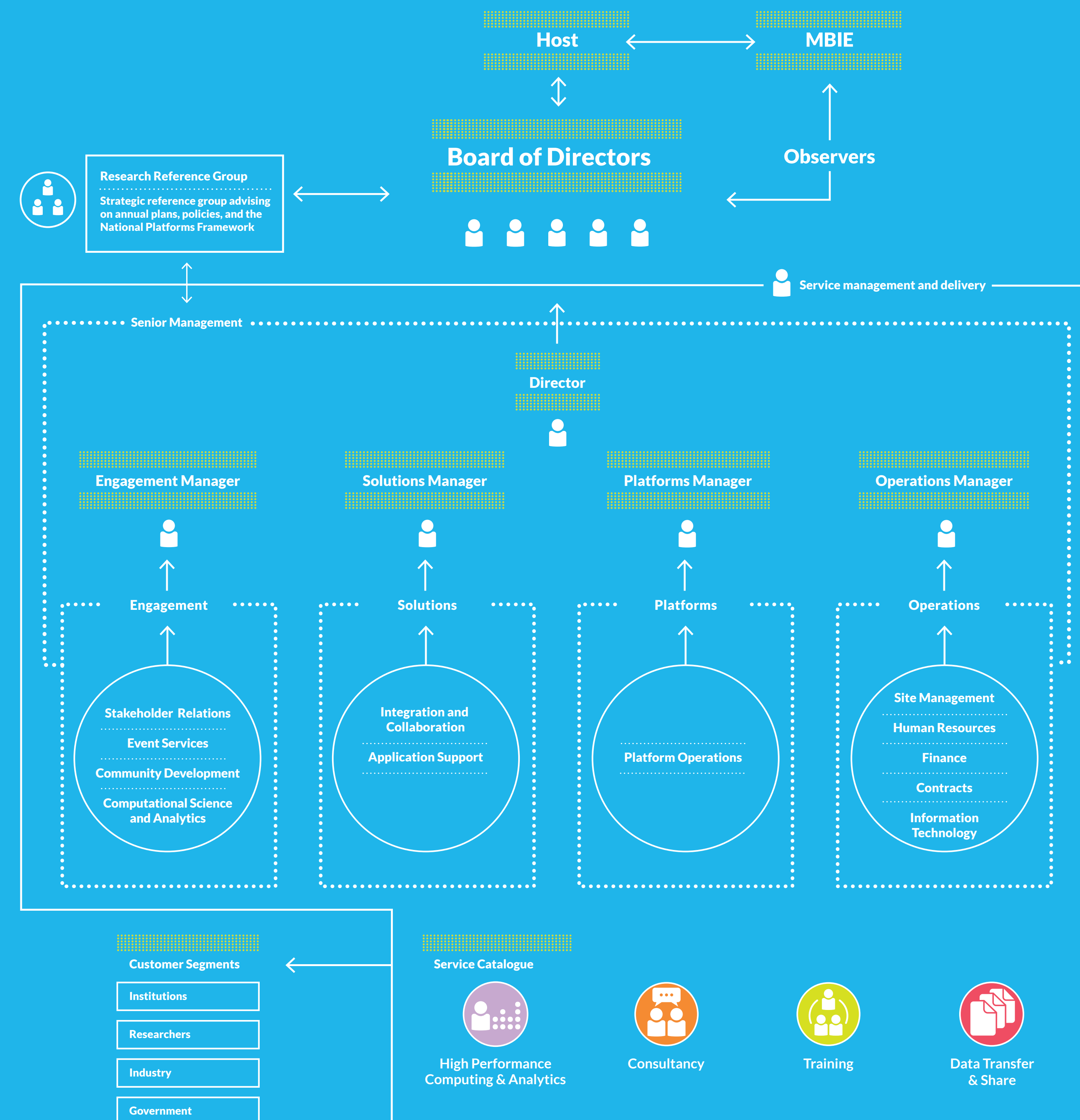
- Blair Bethwaite joined NeSI's senior leadership team as Solutions Manager, bringing significant experience and international connections across HPC and research cloud engineering, platforms, and solutions.
- NeSI's support team expanded to welcome two new members, Dinindu Senanayake and Callum Walley, to expand our domain expertise in genomics and engineering
- Our collaboration and integration team welcomed Nitharsan Puwanendran, a new analyst programmer to help improve our business systems and websites.
- Nooriyah Lohani joined the engagement team as a new Research Communities Advisor to work more closely with and understand our user requirements.
- To support NeSI's growing product management practice, Brian Flaherty was hired as a data services product manager near the end of the year to help us roll out new and improved data services in 2019.

Michael Uddstrom, NeSI's Platforms Manager, will be retiring in 2019. Since NeSI's start, Michael has been working nationally with scientific communities to identify and respond to their HPC and data science needs, and extracting patterns, issues, and requirements that have helped drive positive change to NeSI and NIWA's shared infrastructure and services platform. Planning for Michael's departure began in Q4, with two new related positions being recruited.

In October 2018, the full NeSI team gathered near Christchurch for its annual retreat. The gathering was a great chance for the team to reflect on the success of NeSI in navigating its evolution over the last few years, and to celebrate our recent achievements and team culture.

Organisational design

NeSI operates as a virtual organisation (by contract) with teams formed along functional lines.



NeSI New Zealand eScience Infrastructure

Collaborating with

MINISTRY OF BUSINESS, INNOVATION & EMPLOYMENT · THE UNIVERSITY OF AUCKLAND · NIWA · UNIVERSITY OF OTAGO · Manaaki Whenua Landcare Research

Computer models could help reduce surgery cancellations

Each day, a hospital's Intensive Care Unit (ICU) needs to make important decisions around staffing and resources to prepare for the next day's incoming patients and scheduled surgeries.

Unfortunately, ICUs are highly variable environments, known for unexpected changes in urgency for its services. June Lau, a Doctoral Candidate in the Department of Statistics at the University of Auckland, is using NeSI supercomputers to help ICUs better predict and respond to the day-to-day demands for staffing, beds, and equipment. "It's about ensuring they always have sufficient capacity for every patient that presents themselves to the ICU," says Lau. Her work to date has focused on cardiovascular intensive care. Generally, a cardiac ICU needs to account for three types of patients each day. Nearly 75% of patients are admitted for planned surgeries, such as those who require heart surgery but who have been on a waiting list. Alongside the planned surgeries, between seven to ten elective surgeries are also scheduled each day. The remaining capacity is held for emergency patients, such as those who have been in an accident, suffered a heart attack, or who are unexpectedly transferred in from a different hospital. These patients are the hardest to plan for because it is difficult to predict how many will arrive, when they will arrive, how long they will stay, and what they will require for care.

Each ICU has on-call nurses and other resources it can tap into as spare capacity if needed. However, an unexpected rush of acute trauma patients can use up space capacity without warning. When demand is at its height, one of the ways to free up staff and beds is to postpone the elective surgeries. "That puts the [elective surgery] patient and their family in obviously a pretty anxious state," says Lau. "Also, if the patient and their family have come from rural parts of NZ to support the patient and are told that their surgery is going to be postponed, that is not a great patient outcome." Lau's work aims to change that. She's developing a computer model using discrete event simulation methods to help ICUs better predict and plan for short – and long-term capacity in the ward. Ultimately, she hopes, it will minimise the chances of elective surgeries being cancelled.

"It's a difficult problem because there are really short stays in the ICU unit but also really long stays," she says. "As a result, there is a lot of variability involved in trying to plan for and predict future occupancy of the unit. In addition, the unit itself is quite small, so any small changes or inaccuracies in a model really stand out." Lau has built an anonymised replica of hospital historical data for her computer model. To account for the variability of emergency patients, the model runs thousands of years worth of simulations. It's possible to do this on a desktop computer, Lau says, but the enormity of the task means it would take at least two to three days to complete. "Whereas on NeSI, in the parallel computing performance that we have access to, it means we can run thousands of years of simulations in just over two hours," she says. "It's been amazing to be able to use NeSI to do that because it means we can change different parameters in our simulation model and through that we can do wider scenario analysis."

The better analysis they can do, the better metrics ICUs will have for predicting the number of incoming patients. In particular, knowing the chances of receiving acute trauma patients will hopefully mean fewer 'surprises' and so fewer cancelled elective surgeries. Without access to a supercomputer, her project would have been especially challenging to complete, she says. "I think it would have been very frustrating for me to do my research without NeSI," she said. "I may have had to purchase another computer and just dedicate that one to just run models the whole time." Wrapping up this project in a timely manner was a priority for Lau, as it is her PhD project. She's hoping to complete her thesis within the next 12 months and they are currently in the final stages of a paper on the outputs of simulation for submission to a clinician-based journal. Alongside this project, she is also working with colleagues on another proposal to optimise and model treatment pathways in elective surgeries. It's another simulation that requires them to run a large number of iterations of the same model, but using different data and performing scenario analysis around it. "If we get funding for that research program, then it would definitely be something that is very suited to using NeSI," she said.

"It's been amazing to be able to use NeSI to do that because it means we can change different parameters in our simulation model and through that we can do wider scenario analysis."

June Lau, Doctoral Candidate, University of Auckland

High performance computing and analytics

Driven by the needs of researchers for high-performance productivity, NeSI integrates advanced digital capabilities into a range of eResearch services, and ensures advanced computational research projects are backed by the power necessary to make them a reality.

NeSI lifts New Zealand researchers' ability to tackle large, complex problems, and enables computational and data-intensive research to be carried out much faster and at greater scale, across all domains. NeSI's facilities are the country's most powerful scientific computers.

Transition to new platforms

A major focus for 2018 was seamlessly migrating users from Pan and Kupe across to our new machines Mahuika and Māui. Moving Pan users to Mahuika was the primary focus as FitzRoy users had already migrated across to the NIWA platform Kupe, which operated in similar ways to Māui.

The whole process of transitioning to these new platforms required a massive team effort, from early platform design with key stakeholders through to the tender process and then the final project delivery through NIWA, Cray, IBM and NeSI. By the end of 2018, the new infrastructure was in operation and performing well.

In early 2018, the new infrastructure was installed at the High Performance Computing Facility at NIWA's Greta Point campus in Wellington (photos by Dave Allen)

Making it easier to start

Another focus area in 2018 was to improve our induction processes and support for new users. We worked towards this goal in a number of key areas, such as:

- introducing a new level of basic support within our support team, which included running 'Introduction to HPC' sessions in-person in Auckland and via video conference nationally
- improving our support site with new and updated documentation (in line with the new platforms)
- launching a new and improved project application form through NeSI's new user portal, my.nesi.org.nz

As part of the transition to the new platform, NeSI has been integrating enhanced cybersecurity procedures to protect our users and their data. These new security requirements saw NeSI introduce multi-factor authentication on the new platforms. In some cases, this made the registration and logging in process more difficult for users outside of NIWA's firewall, however the team worked to assist users and make the transition to these new processes as smooth as possible.

In an effort to fully understand the scope of any issues that arose (or still existed) from users' migration onto the new platforms, NeSI surveyed all users in December. This survey identified issues in the logging in process, lack of working data transfer service, and the ability to find the right documentation. This feedback was documented, a follow up plan initiated for any issues still outstanding, and has helped NeSI prioritise its service improvements in early 2019.

Development of new capabilities

NeSI's new platforms enable us to lift our service capabilities to researchers in areas related to data management, research platforms, virtual labs, and security. In terms of data management, we will be able to provide long-term storage for research programmes and data resilience through disaster recovery replication. The new platforms also provide new data analytics tools and the ability to provision visualisation capabilities along with cloud integration.

Virtual labs will provide dedicated and persistent interactive environments for different projects or domains. Genomics Aotearoa have provided some initial requirements for the first iteration of virtual labs due for release in early 2019. In 2018, we also installed the Cray Urika-XC big data analytics tools on Māui which will provide new problem solving methods for researchers. NeSI's recruitment of new staff with expertise in cloud computing, genomics, and engineering domain applications has supported the introduction of these new services.

Transition to new infrastructure

Early 2018

Workshop held with Pan users to understand specific migration related user requirements. From this, NeSI planned the key steps in the migration process which involved sector awareness through communications and webinars and then a detailed training plan to help users make the transition.

January - July 2018

NeSI hosts a series of User Q&A webinars to provide updates on the progress of the new platforms and to answer any questions users have about the new resources.

August 2018

Mahuika goes live on the 1st of August. First test users begin logging in. Full training programme rolled out to get users started on the new systems. Pan is decommissioned at the end of August and Kupe is transitioned back to its primary purpose as a NIWA disaster recovery machine.

September 2018

Māui goes live on the 21st of September. First users begin logging in.

November 2018

Research, Science and Innovation Minister Dr Megan Woods officially opens the new High Performance Computing Facility.

Consultancy

NeSI delivers specialist computational science expertise to the research sector, embedding NeSI's team members within research teams during the course of a project. NeSI experts come from a variety of science domain backgrounds, have extensive programming experience, and can assist with code optimisation, parallelisation, custom code development, and many other tasks. Through these collaborations, NeSI lifts researchers' productivity, efficiency and skills using research computing tools and resources.

Strengthening our specialist science capabilities

During 2018, NeSI focused on broadening uptake of its Consultancy service by a wider range of organisations and improving visibility of the work undertaken by the NeSI Consultancy team. The efforts paid off, with 2018 being the first year that NeSI succeeded in identifying and working alongside research groups from all Collaborators (University of Auckland, NIWA, University of Otago, and Manaaki Whenua – Landcare Research) and the first time NeSI worked with AgResearch on a Consultancy project.

The Consultancy team also looked closely at how the service is promoted, with the goal of ensuring all NeSI researchers are aware of the Consultancy service and how to access it. Existing processes were optimised and new tactics were put in place, including presenting talks at various conferences, advertising the service within the compute allocations process, and adding information about Consultancy to the platform's login interface.

With the growing opportunities for researchers in data science, NeSI wanted to build its internal expertise in this area. Unfortunately, we were unable to identify a suitable candidate for a data science role in our team, but recruitment is likely to continue in 2019.

Consultancy projects in 2018

Project name	Contributions and benefits realised	Science outcomes	Principal investigator
Global Climate Change	This project enabled the use of the second global climate model that will have been run and tested in New Zealand. The results from this project can be compared with those obtained from the HadGEM model currently used by NIWA. Without this consultancy project the researcher would not have been able to carry out this research on NeSI.	Understanding the relationships between atmospheric CO2, water vapour, clouds, and global circulation for better prediction of climate change. Investigating how changes in CO2 level is linked to changes in atmospheric humidity and clouds, which feedback on the radiative energy budget of the Earth's climate system, and in turn affecting climate sensitivity.	Tra Dinh University of Auckland, Department of Physics
Optimisation of tracking algorithm for precipitation systems	Code is now able to run on the new platforms. Without these changes the researcher will not be able to run the full analysis and therefore will not be able to publish results from this work. The researcher was also upskilled in version control and software engineering practices. Performance gains of 3.5 - 5.5x were realised against the initial working version.	Tracking coastal precipitation systems in the tropics. Coastal precipitation plays an important role in the economy of island nations. Impacts from too much or too little precipitation can range from losses in agricultural productivity, to unexpected infrastructure costs, to spikes in sales of particular products or services.	David Coppin University of Auckland, Department of Physics
NZESM Model Development	NIWA NeSI team member helped migrate the climate model to the new platforms. Accelerated the set-up of the New Zealand Earth System Model (NZESM) on the new platforms; improved efficiency; reduced core-hour usage; ensured downstream users get data as quickly as possible. 10.3 million core hours usage on the new platforms in 2018.	Producing useful climate simulations using a hierarchy of models The NZESM is at the core of the Deep South Challenge, a large-scale science challenge project that investigates climate change and its consequences for New Zealand.	Olaf Morgenstern NIWA, Deep South Challenge
Code Optimisation for Genotyping-By-Sequencing	Researchers learnt how to optimise and parallelise R codes and have already applied this to other codes they are working on. Performance improvements of up to 20x speed up.	Incorporating new computational methods and tools into analysis of genetic relatedness and construction of genetic linkage maps.	Jeanne Jacobs AgResearch
Porting Community Earth System Model (CESM) to NeSI	Assisted in code porting, so that the researcher could use NeSI. This involved comparing available compilers and choosing the best performing one. CESM is a community code, so having it available on the NeSI platforms should be useful for other researchers going forward.	Exploring how changes in the sun, particularly an effect known as energetic particle precipitation or "solar wind", influences things like ozone balance in the polar atmosphere and how that impacts other climate elements.	Annika Seppälä University of Otago, Department of Physics
Modelling discretized point processes with extra zeros and long-range dependence	Consultancy enabled better throughout and efficiency on the NeSI platforms and support for more complex models and larger datasets. Performance improvements of ~3x speed up overall. This project enabled the development of more advanced models for earthquake modelling. Code has also been shared with the broader scientific community.	Modelling the first of its kind in the type of seismic events analysis. Will help forecast large destructive earthquakes.	Ting Wang University of Otago, Department of Maths and Statistics
GRATE Sediment Transport Calculator	This project is underway but it is expected that it will increase the usability of the code, and make it suitable for running on NeSI platforms. We are also aiming to improve performance of the code.	GRATE is a research tool that can be applied to numerous practical problems in NZ rivers. It will be an effective educational tool, as well as a valued resource for assessing field data.	Jon Tunnicliffe University of Auckland, School of Environment
Generation of sub-Poissonian radiation fields of large photon number	Introduced research software engineering best practices by implementing a new build system, test cases and improving the accuracy of the calculations. Improved the performance of individual simulations, saving 39% of run time. Implemented scripts for automating parameter sweeps involving running large numbers of simulations in parallel, which will save time compared to manually running those simulations.	This code will be applied to simulations of open quantum systems, in particular simulating the fundamental interaction of light and matter.	Howard Carmichael University of Auckland Department of Physics
Monitoring land cover changes with TMASK_SENTINEL	The code was parallelised, resulting in a 53x speedup over the original, serial version. The master - worker approach implemented here can be used as a framework to parallelise other similar codes (upskilling the researcher).	The code is intended to play an important role in monitoring land use in New Zealand and overseas.	James Shepherd Manaaki Whenua - Landcare Research
Conservative Interpolation for Forecasting Weather Hazards	A method for conservative interpolation was developed during this consultancy project, enabling the remapping (regridding) of vector fields from and to arbitrary unstructured grids made of quadrilateral cells.	Achieve conservation of water and energy when regridding vector fields between different grids, e.g. from the cubed-sphere grid used by the weather/climate LFRic code to a longitude-latitude grid. Can be used to compute the amount of water flux and energy entering a region over time to much higher accuracy than was previously the case.	Michael Uddstrom NIWA

Powering global climate simulations

> "The fact that NeSI can provide support to things like getting the code to work puts New Zealand in an amazing position worldwide."
>
> **Annika Seppälä, Senior Lecturer, University of Otago**

Within their first weeks of operation, NeSI's new platforms are already making a big impact on how New Zealand researchers tackle challenging problems.

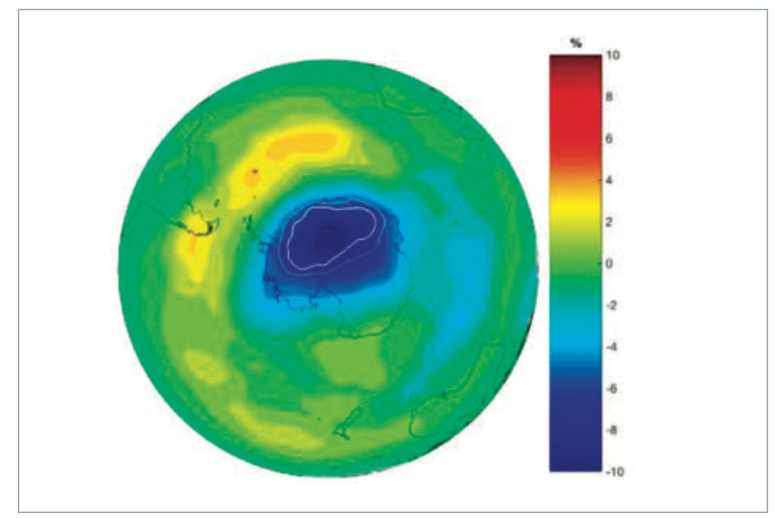
At the University of Otago, Dr. Annika Seppälä is a Senior Lecturer in the Department of Physics and studying the influence of solar activity on climate in the southern polar region, including in New Zealand.
Access to high performance computing (HPC) resources is essential to her work. "Without HPC access we would be unable to do any global climate simulations," she says. Next to HPC, her other most valuable research tool is the Community Earth System Model (CESM), a tool widely used around the world to investigate and simulate how atmospheric, ocean, land surface, carbon cycle, and other components in the climate system are connected. Annika's latest project explores how changes in the Sun, particularly an effect known as energetic particle precipitation or "solar wind", influences things like ozone balance in the polar atmosphere and how that impacts other climate elements such as temperatures and winds elsewhere.

"NZ scientists hear a lot about the New Zealand Earth System Model at the moment, but the Sun-related processes that my group is investigating are not included in the NZESM model," she says. Solar influence produces aurora in Earth's polar regions. Adding energetic particle precipitation from solar activity to the CESM revealed an additional up to 10% wintertime Antarctic ozone loss. (Figure modified from Andersson et al. 2018) So when NeSI's Māui platform launched, a door of opportunity opened for Annika to run the CESM on NZ's newest supercomputer, but she needed help from NeSI's Consultancy Team to make it happen.

"These types of models are rather complex to port, which essentially means getting the model code to work in a new environment," she says. "I'm a physicist, not a computer scientist, so this would have been a massive task for me. So when Wolfgang Hayek from NeSI contacted me and asked if I would benefit from their expert help in porting the code in the new NeSI platforms, I was practically jumping in excitement in my office. I knew this would save months of my research time and things would be done correctly from the start."

Working with Wolfgang, Chris Scott and Alexander Pletzer from NeSI's Consultancy Team meant that Annika could keep her focus centred on her project, while they dealt with the technical details of getting her model properly set up on the cluster. "I know many colleagues working in similar fields around the globe who have to make all of this work on their own and a lot of the time this means that there is a lot of guesswork going on," she said. "We don't know the nitty-gritty of the HPC workings, and, because of the previous point, it can take several months to make complex models work.

The fact that NeSI can provide support to things like getting the code to work puts New Zealand in an amazing position worldwide." NeSI's team was able to get the CESM code properly ported in a matter of weeks, and Annika has been able to get her projects up and running on the new platform fairly quickly. She even picked up some porting tips along the way, and the performance optimisations that were implemented by the team will allow her to get the most out of her HPC core-hour allocation. "I'm especially thankful to Wolfgang for explaining things to me and making sure I understood what certain pieces of HPC terminology meant," she says. "Thanks to this I now have confidence in actually doing code porting and testing in the future. I started referring to the team helping me as the NeSI wizards when I was describing the progress to my colleagues, because they did an amazing job and to me it seemed like magic!" Annika and her students won't be the only ones using the CESM on the new platforms. "I already know there are other scientists looking forward to using the CESM model on Māui," she said. "This project that got the model started will benefit many NZ researchers."



Solar influence produces aurora in Earth's polar regions. Adding energetic particle precipitation from solar activity to the CESM revealed an additional up to 10% wintertime Antarctic ozone loss. (Figure modified from Andersson et al. 2018)

Training

NeSI shares expertise and builds capability in New Zealand's research communities and institutions, growing digital research skills and improving researchers' abilities to use advanced digital tools. NeSI focuses on sustainability by embedding training skills and programmes within institutions and communities, taking a leadership role in growing capacity and changing research culture.

›120 Researchers trained

31 Events

Key topics delivered

- Software Carpentry
- Library Carpentry
- Data Carpentry
- The Carpentries Instructor Training
- NeSI Introductory Workshop
- NeSI New Platforms Training

Special events

- eResearch New Zealand
- ResBaz 2018
- Science Coding Conference
- Machine Learning Day

Increasing opportunities for researchers to gain computational skills

Through late 2017, NeSI reviewed its effectiveness against its Training Strategy. These findings culminated in the Kubke Report (Growing researchers' computational skills to meet future needs, DOI: 10.17608/k6.auckland.7482116.v1) published in 2018. The report reflected on NeSI's achievements to date as well as considering what success looks like for the New Zealand research sector in terms of researcher skills in digital eResearch tools. These findings were presented and discussed at various forums including eResearchNZ 2018, C3DIS 2018, and eResearch Australasia 2018.

Building on our ongoing Carpentries (Software, Data and Library) activities, NeSI had a focus this year on building a routine around delivery of NeSI Introductory Workshops to make it easier to start using NeSI. We also supported existing users during the migration to Māui and Mahuika with nationwide New Platform Training events in August including an online event for those outside the main centres.

In late 2018, we were able to prioritise development of more advanced training material to be delivered at events in 2019. This material will target existing NeSI users who want to become more efficient users of the platform.

NeSI training events for 2018

Date	Event	Location
23 January	Introduction to Unix	University of Otago
24 January	Introduction to HPC Using NeSI	University of Otago
31 January	Collaborating on research projects using version control	Manaaki Whenua – Landcare Research
7-9 February	Software Carpentry	University of Otago
7- 8 February	Introduction Unix and HPC using NeSI	Lincoln University
14-15 February	Software Carpentry Training (NZ-wide)	Massey University
23-24 April	The Carpentries Instructor Training	Victoria University Wellington
27-29 June	Auckland ResBaz - Software Carpentry	University of Auckland
4-6 June	Otago ResBaz - Software Carpentry	University of Otago
4-5 July	Data Carpentry	University of Otago
5 July	Library Carpentry	University of Otago
9 July	Library Carpentry	University of Canterbury
11 July	Library Carpentry	Massey University
12-July	Software Carpentry	University of Auckland
12 July	Library Carpentry	University of Auckland

Date	Event	Location
6 August	NeSI New Platform Training	Auckland
7 August	NeSI New Platform Training	Christchurch
10 August	NeSI New Platform Training	Hamilton
10 August	NeSI New Platform Training	Dunedin
16 August	NeSI New Platform Training	Auckland
16 August	NeSI New Platform Training	Wellington
20 August	NeSI New Platform Training	Online
23 August	The Carpentries Instructor Training	University of Otago
13-14 November	Data Carpentry	University of Otago
13-14 November	Software Carpentry	University of Otago
20-21 November	Software Carpentry	Manaaki Whenua – Landcare Research
28 November	NeSI Introductory Workshop	University of Auckland
29 November	Data Carpentry	ESR
4 December	NeSI Introductory Workshop	University of Otago
5 December	NeSI Introductory Workshop	Massey University
19 December	NeSI Introductory Workshop	University of Auckland

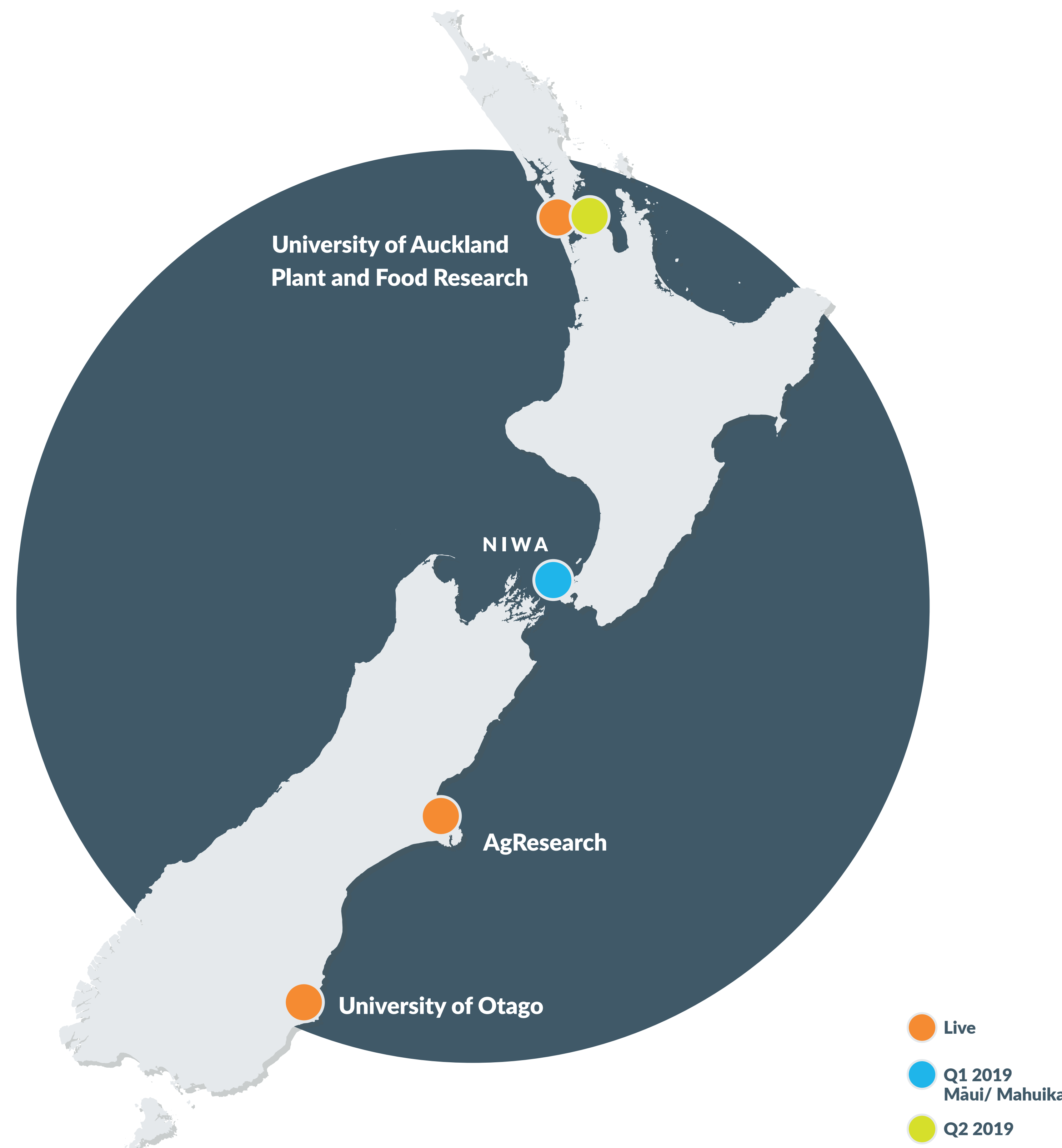
Data transfer and share

NeSI supports researchers to transfer and share their data, working with research communities, institutions, and REANNZ to improve end-to-end performance and reduce time to solution.

Re-establish and refine data transfer service
Through 2018, NeSI continued to look at how it enhances the essential advanced data network connectivity provided by REANNZ. A major component of this work is to establish a data transfer platform, Globus, across sector institutions, providing a common platform for accessing, transferring, and sharing data nationally and internationally. Early in 2018 NeSI reviewed its current implementation of Globus, building a backlog of areas for improvement. The existing data transfer service was taken down as part of Pan's decommissioning, and a new and improved data transfer service was designed to take its place. This new service is due to be implemented in early 2019.

Data share
Work was underway before the end of the year to document and productionise the Globus data sharing function for NeSI users and collaborators. Sharing between managed and personal endpoints via GridFTP is being augmented by native HTTPS transfer with new features in the Globus product. This will provide underpinning functionality for development of secure data sharing and publishing services, the first of which will be explored in 2019, targeting the needs of the Genomics Aotearoa (GA) community. NeSI's partnership with GA is helping us to understand broader issues in data management and custodianship, with discussion underway on access management requirements for Māori and other sensitive genomics data.

Data transfer nodes



Meeting our objectives

NeSI focuses on the following objectives, tracking performance against related Key Performance Indicators (KPIs).

1 | Support New Zealand's research priorities

Annual case studies published aligned with government priorities

Number	Status
≥20 Case studies	20

2 | Grow advanced skills that can apply high-tech capabilities to challenging research questions

Annually, users report significant benefit to any project from NeSI services

Number	Status
>50 Users	208

3 | Increase fit-for-purpose use of national research infrastructure

Annual change to utilisation of national platforms

Number	Status
+20% Utilisation	20.5%

4 | Make fit-for-purpose investments aligned with sector need

Annually, users indicate that NeSI services meet their needs

Number	Status
>80% User satisfaction	91%

5 | Enhance national service delivery consistency and performance to position NeSI for growth

Annual availability of services

Number	Status
>98% Availability	99.6%

6 | Realise financial contributions and revenue targets to enhance NeSI's sustainability

Contract-to-date, ratio of collaborator commitments to Crown contribution

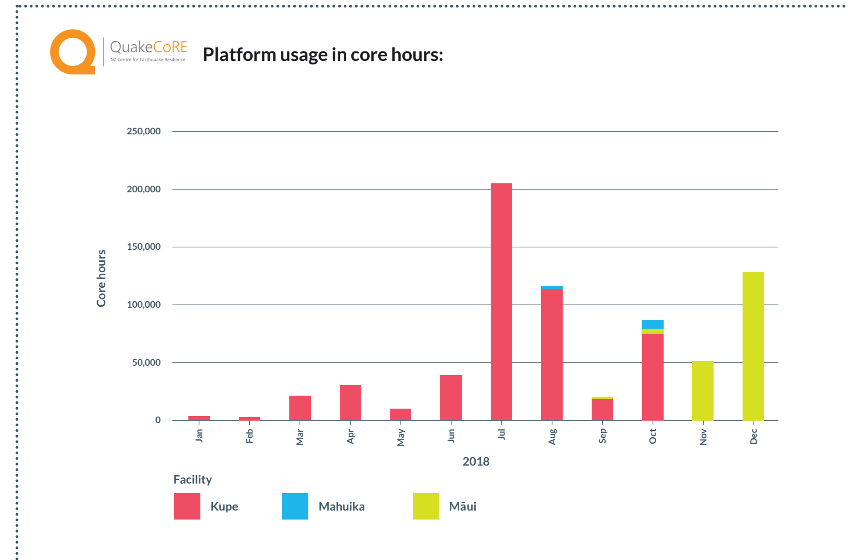
Number	Status
>95%	88.6%

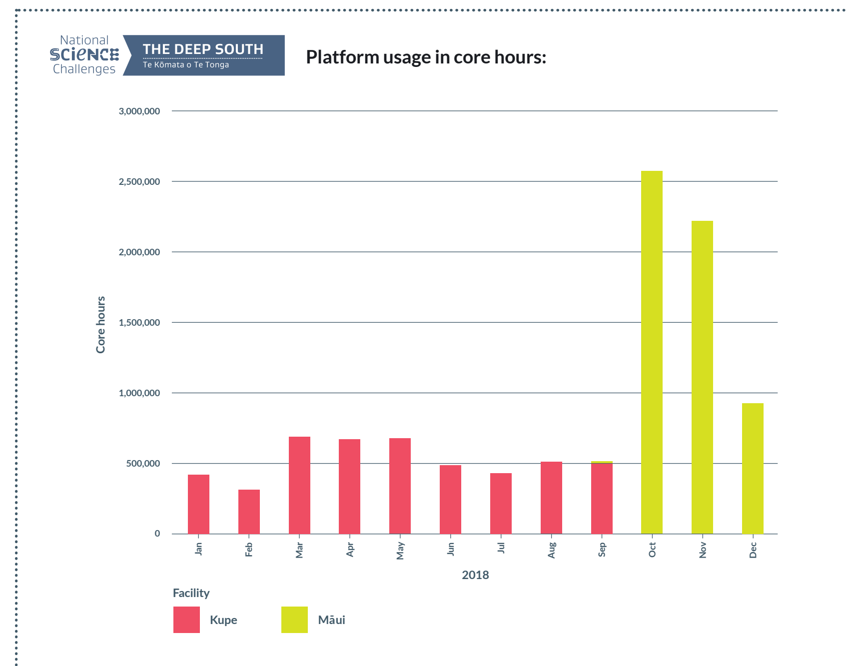
Objective 1

Support New Zealand's research priorities

Building capabilities within national research collaborations

NeSI continues to provide support to some of the largest national research collaborations in New Zealand, including QuakeCoRE and the Deep South National Challenge. Although the transition between platforms in 2018 led to a period of lower usage for some research groups, this quickly picked up once the new platforms were in operation.





Continued on page 36.

Simulating earthquakes for a safer future

"We understand the basic physics, but the complexity is such we can't use intuition to understand these problems. To understand and quantify these phenomena, we need to throw a supercomputer at them."

Brendon Bradley, Deputy-Director, QuakeCoRE

Life in New Zealand is shaped by earthquakes. Just as the land is sculpted by the slow grind of tectonic plates, raising alpines and flooding land, the residents of New Zealand and our cities are similarly carved.

Dr Brendon Bradley is an earthquake engineer and Deputy-Director of QuakeCoRE, a research centre for earthquake resilience. Brendon is using NeSI supercomputers to improve earthquake simulations and build better structures. "We simulate how earthquakes cause waves that propagate through the earth. We model how they shake buildings, bridges and other sorts of infrastructure. We can use that computation to determine the earthquake's physical impact," Brendon said. Brendon's team are working on building more accurate earthquake simulations. At the moment, earthquake damage is predicted by mapping past tremors and building a statistical model. But these models often fail to account for microzones – areas where the earth may contain different types of soil and rock. "We started on this after the 2010 Canterbury earthquake. There was over $40 billion dollars of damage, which was 20 per cent of our yearly GDP," he said. The Canterbury earthquake's aftershock also caused the Christchurch earthquake in 2011.

Combined, these earthquakes killed over 185 people and injured thousands more. It was a devastating moment in New Zealand's history. The Canterbury and Christchurch earthquakes began an effort to better understand why some buildings survived while others didn't, leading to Brendon's research into microzonation – mapping microzones of varied geology. "Looking back on the recordings of shaking from 2011, it became clear there was microzonation. Certain areas had higher amounts of shaking, leading to more damage."

"Now we're modelling the specific geology of the earth at each location we care about. Say it's Wellington, Auckland or Christchurch, we model the geometry of the fault beneath those places – the angle and bend of that fault and how it moves." With this new modelling, Brendon's research can be used to zone different areas by earthquake damage risk. Engineers can design sturdier buildings in those zones to withstand the greater amount of shaking, leading to less damage. This refined form of earthquake simulation also means a high demand for supercomputing power. The geology of the earth is turned into a 3D model and tremors are simulated within it. How those tremors interact with different soil and rock types shows how buildings built atop them will shake. "We use millions of core hours every year from NeSI. Our research team actually includes former NeSI software developers, so we push the bleeding-edge boundaries more. We challenge NeSI with new things and their team gives us the familiarity with hardware and storage systems we don't have." Brendon's team use both the NeSI Mahuika and Māui supercomputers. One recent example of their simulation takes into account Wellington's unique earth properties. Wellington is built upon a rocky slope that flattens out into Wellington harbour. As the land flattens to meet the sea, the rocky earth gives way to deeper soil. The soft, shifting soil causes more shaking during earthquakes and puts the buildings built upon it at greater risk of damage. "The November 2016 Kaikoura earthquake caused a lot of damage to multi-storied buildings in Wellington. Some parts of the CBD were much more damaged than others, where the deep soil had caused more shaking."

The simulations are already leading to surprising insights into how New Zealand's infrastructure reacts to the shaking ground beneath it. Brendon is optimistic about how far this technology can be applied to preventing destruction in the country's future. "The technology is already better in some ways, but the rate of improvement is really the great thing. We're bringing in engineers to talk about how they can use these scientific results in building and construction." The future of Brendon's work will require even more work with NeSI. Hopefully one day, research like Brendon's will help buildings better weather the devastating earthquakes common to New Zealand, saving the lives of many.

Objective 1 (continued)

genomics aotearoa

Advancing genomics capabilities in New Zealand

In December 2018, NeSI and Genomics Aotearoa announced a partnership to deliver advanced capabilities and enhanced tools for genomics research in New Zealand. Working with NeSI's newly launched platform and its specialist team of experts, Genomics Aotearoa research projects will be able to tackle the largest scales and complexities of data available in New Zealand, enabling them to participate in and lead the rapidly developing fields of genomics and bioinformatics.

To underpin this partnership, NeSI recruited new bioinformatics expertise to work alongside Genomics Aotearoa researchers to better understand their current and future needs, while also assisting with code optimisation and deployment, genomics application installation, onboarding researchers new to using NeSI's platform, and other support activities.

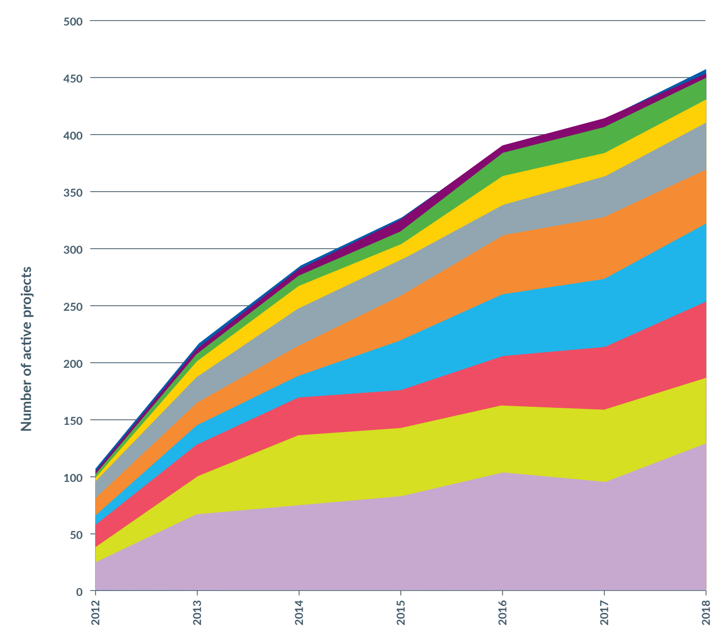
"This capability will allow New Zealand to undertake world-class research in what is a very fast-changing field, and will support researchers to respond to New Zealand's unique needs in health, environment and primary production."

Genomics Aotearoa Director, Professor Peter Dearden

Embracing a diverse research sector

Throughout 2018, NeSI continued to support a broad range of research disciplines from its traditional big user communities of earth science and chemistry, to newer communities such as medical science and biology. There has been good growth, across a range of disciplines, in the number of active projects each year since NeSI started, with plenty of potential to grow outside of the core user disciplines.

Number of active projects per year



During 2018, NeSI's resources, services and support enabled New Zealand researchers across a wide range of research disciplines to greatly increase the size and speed of processing their scientific workloads, and to investigate scientific problems that were previously impossible.

REANNZ

Setting data transfer records on REANNZ

Strategic Science Investment Fund (SSIF) infrastructure plays an essential role in the delivery of services and support to New Zealand's research sector. A prime example of this was demonstrated in May 2018 when a record level of data travelled over the Research and Education Advanced Network New Zealand (REANNZ) as NeSI began moving the first wave of its users' research data to its new platforms.

Tens of millions of files and hundreds of terabytes of data were migrated between NeSI's Pan cluster at the University of Auckland and its new supercomputer, Mahuika, at NIWA's High Performance Computing Facility (HPCF) in Wellington. Over the course of several weeks, NeSI support staff copied and synchronised the research data on Pan to move it over to the new platform's storage facilities at the HPCF. Using REANNZ was essential for supporting this massive transfer, as REANNZ's high capacity, high-speed network was able to handle the loads with little to no disruption to its regular operations.

Overall, NeSI and REANNZ's working relationship remained strong in 2018, with both organisations focused on contributing to priority research areas and delivering long-term beneficial impacts to New Zealand's health, economy, environment and society.

Expanding and targeting support efforts

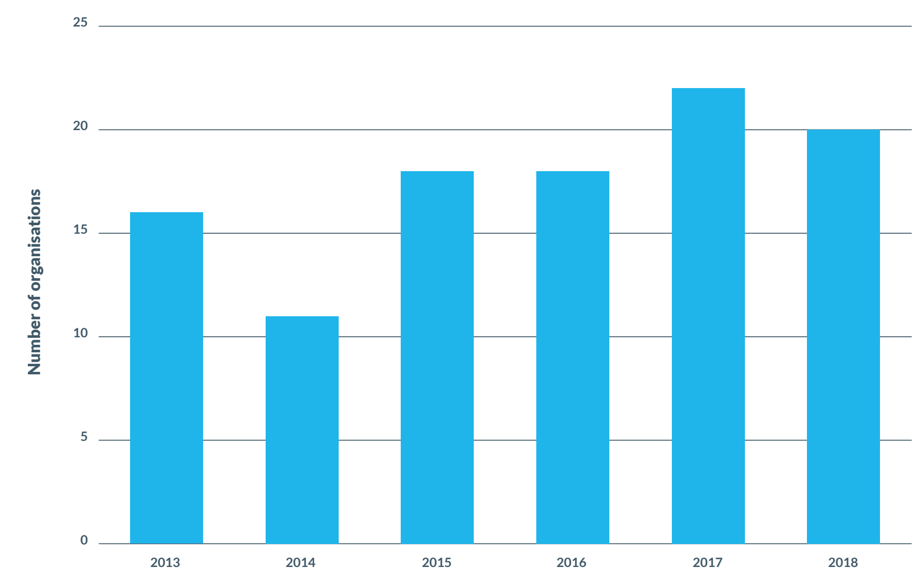
Understanding the research we support is critical to improving our services and identifying which disciplines or research groups should be targeted for further support. In 2018, a new Research Communities Advisor was recruited to take a more active role in understanding NeSI's user base. Further growth in this team is also planned for 2019.

Also, NeSI's project application form was reviewed in 2018, and changes were implemented to improve our information collection methods via this process.

Growing broader sector uptake and impacts

NeSI has seen continued uptake of subscriptions within the research sector, comprised of new organisations as well as maintained relationships with subscribers from previous years. By the end of 2018, NeSI had active subscription contracts with nine organisations outside of NeSI's four collaborators.

Organisations that have actively used NeSI



Showcasing impacts and benefits to research

Twenty case studies were published in 2018, highlighting the ways NeSI resources, support, and expertise delivered impact and benefit to the New Zealand research sector. The case studies profiled researchers based at NeSI's collaborator institutions as well as those working with AgResearch, GNS Science, QuakeCoRE, Victoria University of Wellington and projects supporting the Biological Heritage National Science Challenge, Genomics Aotearoa, the UK Met Office, the New Zealand Software Carpentry community, and the New Zealand and Australian Research Software Engineering communities.

Taking machine learning further

We've all heard the saying 'practice makes perfect'. While this is normally a piece of advice shared with someone learning a new skill, it applies to machine learning as well.

When a robot uses a trial-and-error process to navigate a building layout with zero collisions, that is practice in action. Known as reinforcement learning, this subfield of artificial intelligence (AI) is at the heart of Yiming Peng's work with the Evolutionary Computation Research Group at Victoria University of Wellington. Yiming has been using NeSI computing resources to develop new algorithms that will make it easier for computers to solve complex reinforcement learning problems.

His recent PhD project explored how to improve the efficiency of existing algorithms and how to design better representations of critical reinforcement learning components. "We have developed several new reinforcement learning algorithms that have achieved superior and competitive performance in comparison to several state-of-the-art algorithms reported in the literature," he says. Access to computational resources can be a critical factor for research success in the AI domain, he notes. For example, the problems he was investigating for his PhD project relied on his ability to run large-scale computational experiments. "I couldn't imagine how long my PhD journey would go if I hadn't used the high performance computing resources of NeSI," he says.

The results of his work already show promise for real-world applications. Yiming has been collaborating on a biology project to help several ecologists identify animals' activities from images captured by trap cameras in Wellington. "In the project, I have been mainly adopting image recognition techniques, which had given reasonably good results," says Yiming. "Some recent findings suggest that reinforcement learning techniques can be utilised to facilitate image recognition tasks.

So, it is worthy to provide an attempt of getting my algorithms involved in the project." With his PhD now complete, Yiming is currently looking for an AI-related job in industry, but says he also has other research project ideas he'd like to pursue. "One of the projects I am going to start is to use reinforcement learning to facilitate supervised learning tasks such as classification, which undoubtedly will involve large computations," he says. "Thus, I am pretty sure that I will continue to use NeSI."

> "I couldn't imagine how long my PhD journey would go if I hadn't used the high performance computing resources of NeSI."
>
> **Yiming Peng, Evolutionary Computation Research Group, Victoria University of Wellington**

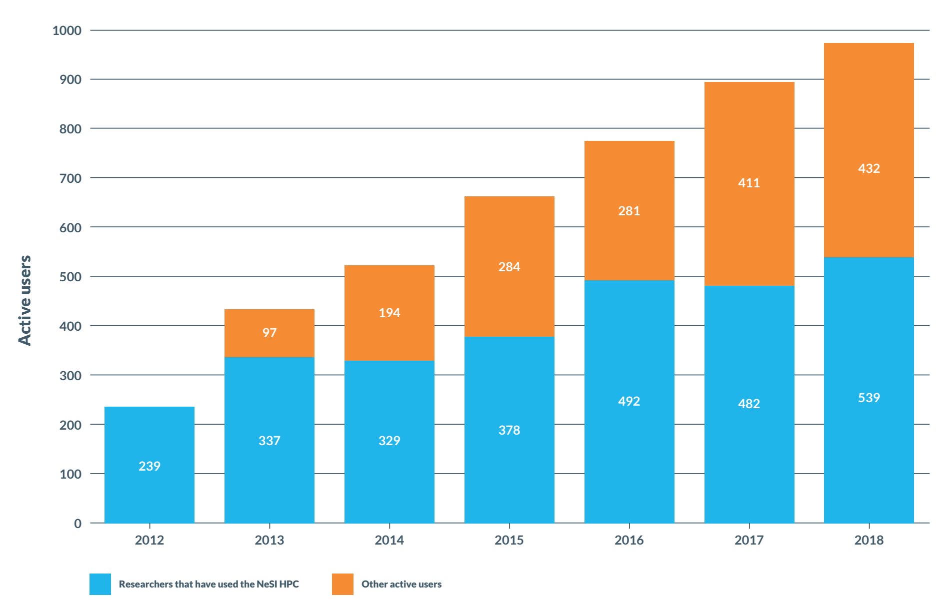
Objective 2

Grow advanced skills that can apply high-tech capabilities to challenging research questions

Sustained growth in users applying HPC to research

NeSI has seen a continued year-on-year increase of users on our platforms. This trend is true for users actively submitting compute jobs, as well as users accessing the system for peripheral uses such as contributing raw data and accessing the output data.

Number of active users



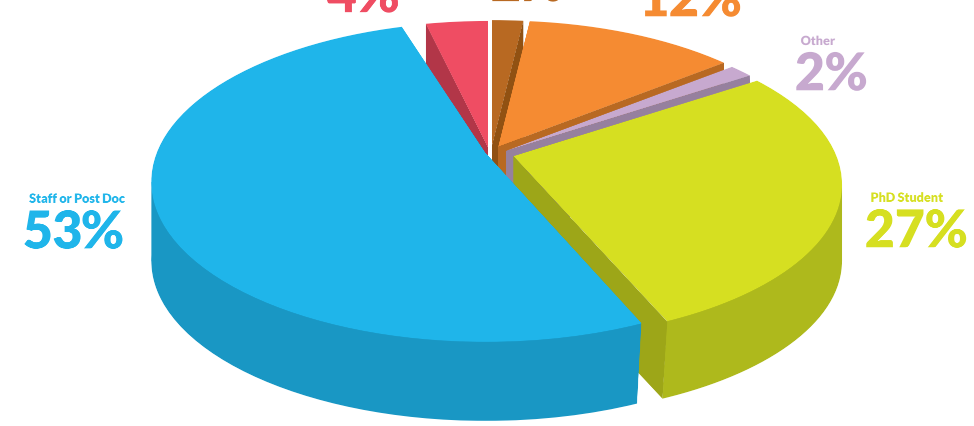
Although there is growth in our user base, particularly with our new platforms, there is still room for greater uptake of our services. Through recruitment of our Research Communities Advisor and specialist support team members in engineering and genomics, we are looking to create greater awareness and support of NeSI's services across New Zealand's research sector.

Fostering partnerships with research collaborations and institutions

NeSI has contributed a broad range of new local and international initiatives to the research sector during 2018, including:

- NeSI hosted the Science Coding Conference, now its third year in collaboration with Scion (host organisation) and Science NZ.
- Continued to support the establishment of the Research Software Engineers community in partnership with Australian colleagues.
- Continued support of Research Bazaar event series across a number of New Zealand institutions (created at the University of Melbourne, Australia).
- Hosted eResearch NZ 2018 alongside co-host REANNZ in Queenstown in February.

Career stage of NeSI active users as of 2018 (HPC user)



> "NeSI Staff provide tremendous support in software maintenance and general 'assurance that things are done well', which is difficult to attain by individual researchers."
>
> **Klaus Lehnert, University of Auckland**

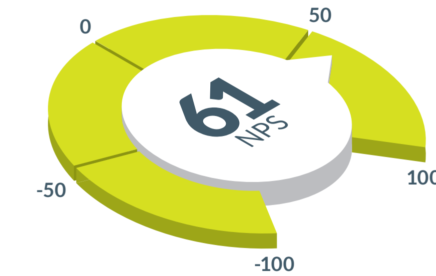
Growing researchers' digital computational literacy

During the current contract NeSI has established a successful training programme, including a key lever to scale its benefits through focusing on train-the-trainer. NeSI has matured its approach to research software engineering, with a well subscribed and recognised consultancy offering which lifts computational capabilities of research teams. Lastly, NeSI's researcher support for access to and use of its infrastructure is highly valued by researchers, consistently receiving high praise for responsiveness and expert support.

NeSI Project Survey

Of all projects surveyed that had allocations ending in 2018, we had 238 responses to the question, "Did your project receive significant benefit from using NeSI?". Of those responses, 87% responded positively that their research project received significant benefit from using NeSI. Of the remaining responses, 20 were neutral, 5 disagreed and 5 strongly disagreed. Reasons given for these responses often refer to NeSI not actually being used as anticipated due to other available resources or timing.

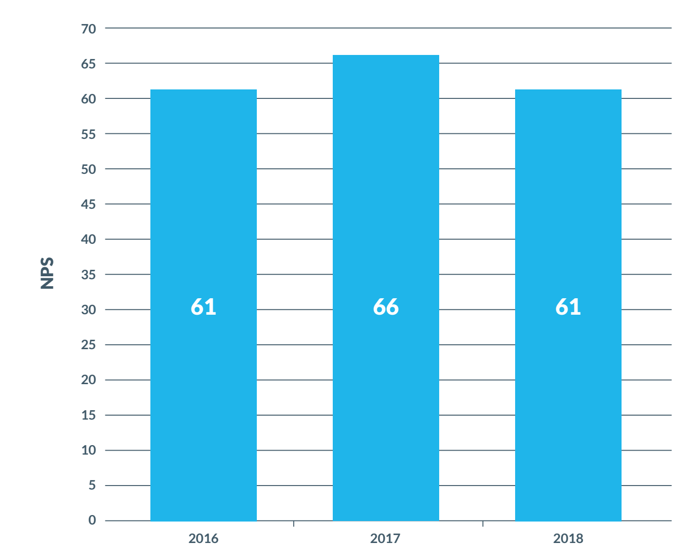
NeSI's routine project survey, shared with researchers at the end of a project with NeSI, asks whether they would recommend NeSI's services. Based on our survey results to date, NeSI has maintained a Net Promoter Score (NPS) of 61. This is a very high score – any positive score is seen as good, with scores between +50 to +80 seen as high performing. This shows that NeSI has loyalty and satisfaction from its user base.

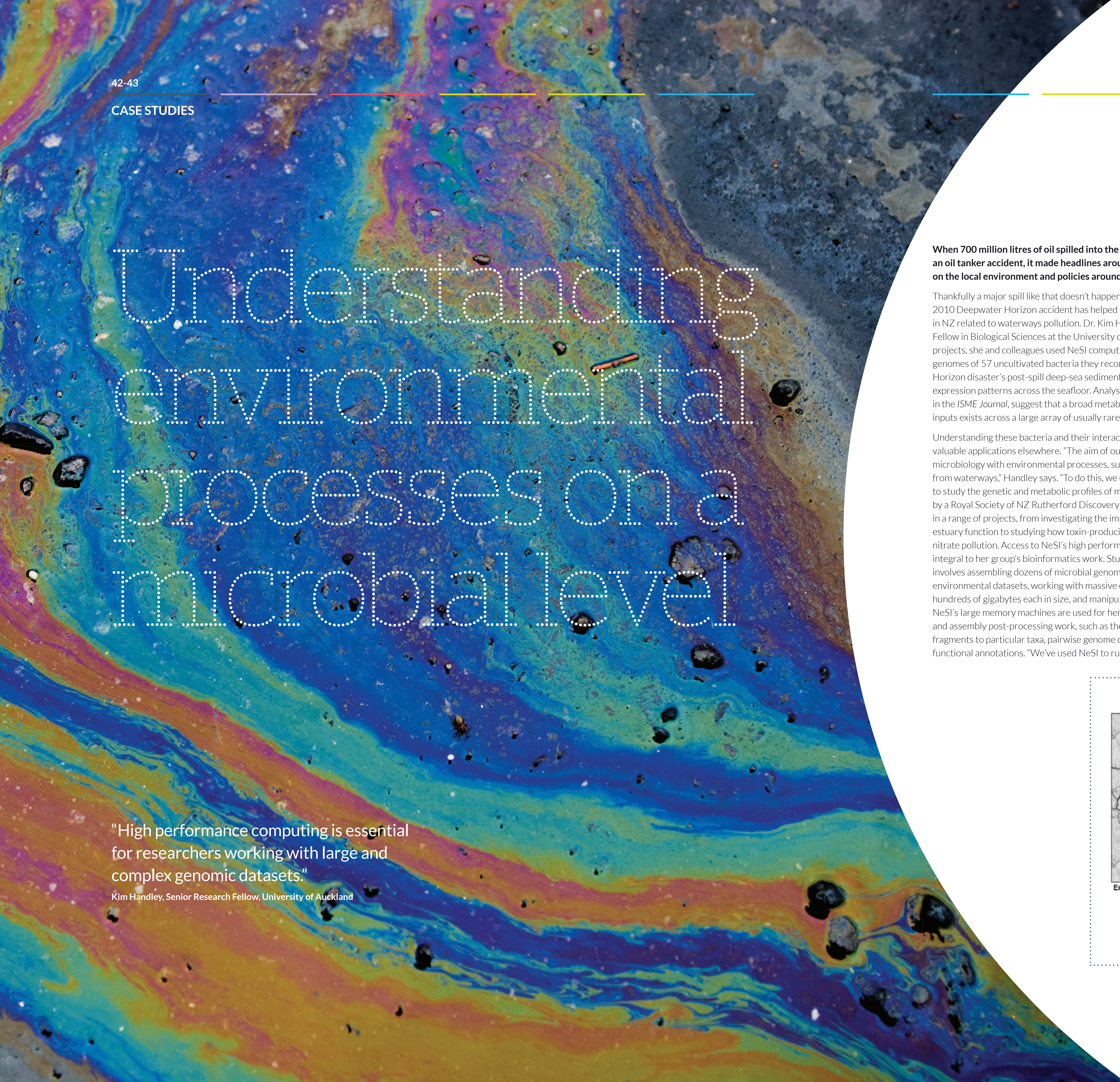


Looking at the change in our NPS score over time, it is consistently high, with perhaps a slight drop off in 2018 due to issues some users experienced during the migration between platforms.

> "It was great working with NeSI… The new code implemented is much faster and runs more efficient on NeSI platforms. Thanks to continuous collaboration along the code building, I am now able to maintain and develop it on my own. I learnt a lot from this project and I am really glad that NeSI offers us this kind of expertise for long-term projects that require a lot of work."
>
> **David Coppin, University of Auckland**

NeSI net promoter score (NPS) by year





Understanding environmental processes on a microbial level

> "High performance computing is essential for researchers working with large and complex genomic datasets."
> **Kim Handley, Senior Research Fellow, University of Auckland**

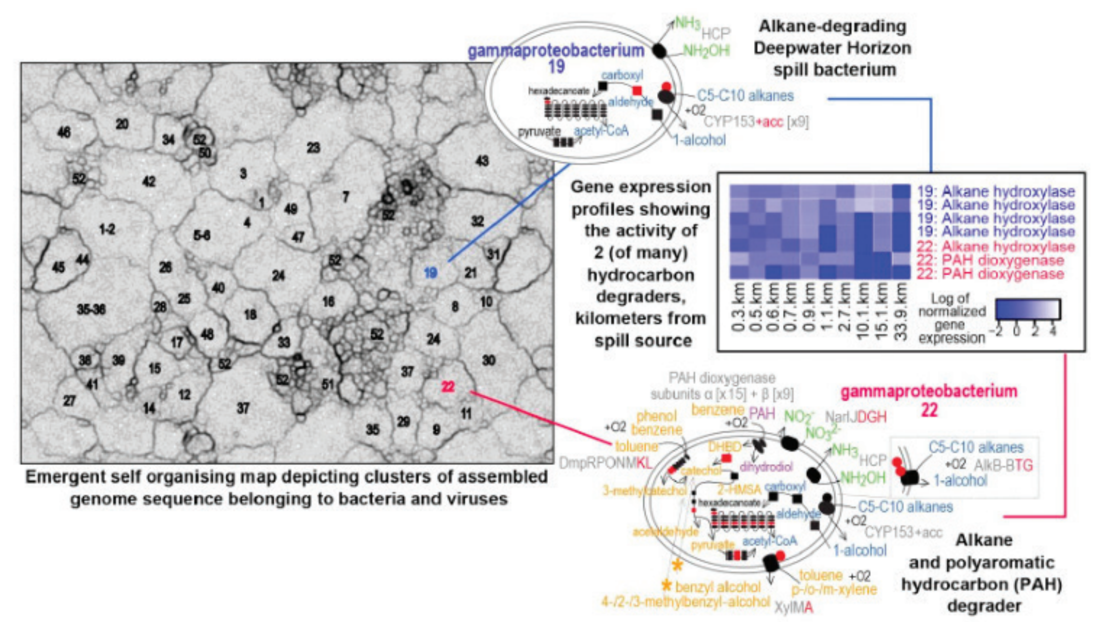
When 700 million litres of oil spilled into the Gulf of Mexico in 2010 from an oil tanker accident, it made headlines around the world for its impacts on the local environment and policies around drilling practices.

Thankfully a major spill like that doesn't happen every day, but studying the 2010 Deepwater Horizon accident has helped inform important research in NZ related to waterways pollution. Dr. Kim Handley is a Senior Research Fellow in Biological Sciences at the University of Auckland. In one of her latest projects, she and colleagues used NeSI computing resources to analyse the genomes of 57 uncultivated bacteria they reconstructed from the Deepwater Horizon disaster's post-spill deep-sea sediments and to study their gene expression patterns across the seafloor. Analyses from the study, published in the *ISME Journal*, suggest that a broad metabolic capacity to respond to oil inputs exists across a large array of usually rare indigenous deep-sea bacteria.

Understanding these bacteria and their interactions with the oil could have valuable applications elsewhere. "The aim of our group's research is to link microbiology with environmental processes, such as the removal of pollutants from waterways," Handley says. "To do this, we use approaches that enable us to study the genetic and metabolic profiles of microbial communities." Funded by a Royal Society of NZ Rutherford Discovery Fellowship, Handley is involved in a range of projects, from investigating the implications of sediment inputs on estuary function to studying how toxin-producing stream bacteria respond to nitrate pollution. Access to NeSI's high performance computing resources is integral to her group's bioinformatics work. Studying microbial communities involves assembling dozens of microbial genomes at a time from large environmental datasets, working with massive output files, some of them hundreds of gigabytes each in size, and manipulating large search databases. NeSI's large memory machines are used for her group's genome assemblies and assembly post-processing work, such as the assignment of assembled fragments to particular taxa, pairwise genome comparisons, and gene functional annotations. "We've used NeSI to run arrays of hundreds of genome-related jobs in order to achieve rapid simultaneous processing," Handley says. "In addition, use of a shared, remote computing resource enables the team to access data, databases, and software without redundancy, and from anywhere at any time."

Technical staff at NeSI and Auckland's Centre for eResearch have helped Handley and her team make the most of the supercomputing resources. In addition to providing standard support like assisting with software installs or troubleshooting occasional job errors, Martin Feller and Markus Binsteiner at the Centre for eResearch and Ben Roberts at NeSI also helped Handley and her students with large data transfers and provisioned custom spaces on NeSI's Pan cluster for databases and software installations that were shareable across all projects.

"NeSI support have also assisted in working with local university IT to provide access to additional resources for data storage and machines for graphical user interface software and larger genome assemblies beyond Pan's capacity," she said. Looking ahead, Handley has other projects on the horizon that will make use of NeSI resources, including an MBIE-funded project studying the genomic content of aquifers to understand groundwater as a living and reactive system, and a study that's part of Genomics Aotearoa, the national advanced genomics platform, that looks at high quality genomes from waterways. For these and future projects, NeSI resources will remain an essential component of their work, she says. "High performance computing is essential for researchers working with large and complex genomic datasets. Researchers attempting to assemble multiple (potentially hundreds of) genomes from diverse sediment microbial communities, for example, need access to large memory machines," she says. "With that in mind, we are looking forward to migrating to NeSI's new and enhanced-capacity high performance computing platform this year."



Objective 3

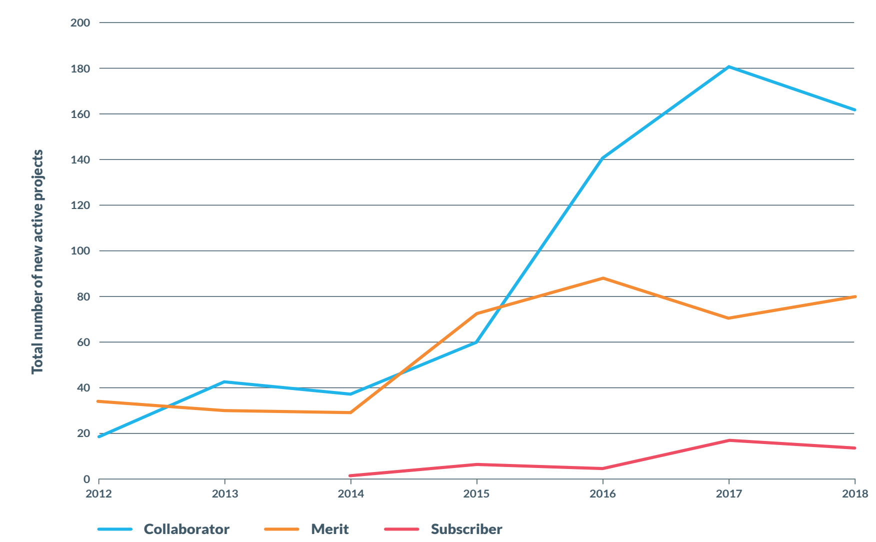
Increase fit-for-purpose use of national research infrastructure

Incentivising efficient use of infrastructure

2018 has seen continued increase in the number of new project allocations across all classes. This is partly driven by the increase in available capacity for new projects but also many projects restarting with a new allocation on Māui or Mahuika.

Core hour usage against available core hours in each allocation class:

Distribution of new active projects by year:



"This project would not have been possible at all without NeSI's support; it has generated substantial benefits for the university and industry partner."

NeSI user

Increased usage - broadening uptake

Usage of NeSI's platforms steadily grew through 2018 and has increased by 20.5% compared to the previous year. This overall growth is very encouraging as usage in the last year was affected by the two major platform migration events, from FitzRoy to Kupe in December 2017 and Kupe and Pan to Māui and Mahuika in September 2018. During these periods there was a noticeable reduction in usage.

Platform	12 months ending on 31.12.17 (CPU core hours used)	12 months ending on 31.12.18 (CPU core hours used)	% Change
Pan	36,833,079	28,358,301*	
FitzRoy	13,938,276	30,135*	
Kupe	19,208*	11,677,260*	
Mahuika		11,450,971*	
Māui		9,674,849*	
Aggregated	50,790,563	61,191,515	20.5%

*Platform was not operational for this whole year

Supporting new innovations in image processing

One of the most essential and important tasks in image processing and computer vision research is background estimation. This is the task of determining what parts of a video are static background, and therefore not of interest, so that those parts can be ignored to save computation time.

Consider a surveillance camera that is monitoring a hallway. "Most of the time, there is nothing happening, and so it would be wasteful to run a computationally expensive person detection algorithm across every frame of the video," says Andrew Chen, a PhD Candidate in the department of Electrical and Computer Engineering at the University of Auckland. Instead, he explains, they can use comparatively lighter background estimation to detect when there has been a change (e.g. a moving object), and then run person detection on that part of the frame to see if it is a person or not. Using NeSI supercomputer resources, he is developing and refining a new superpixel-based algorithm that completes the background estimation process faster and more accurately.

To develop this algorithm, Andrew and his colleagues had to run the algorithm against a large benchmark dataset, which takes about three to four hours on a standard desktop computer. The need for NeSI came when they wanted to take their development and analysis a step further and find parameter values for the algorithm. Even though they did discrete sampling of the values (i.e. not checking every possible value), they still needed to run the algorithm a few hundred times in order to identify the relationship between each parameter and accuracy, as well as compare accuracies across the various parameters.

The figure represents how the SuperBE algorithm isolates the background and foreground in an image.

They wanted to ensure they had the best possible accuracy in order to publish their best results in papers. "This algorithm has five to six parameters, and there are a lot of different combinations of these parameters that need to be checked," says Andrew. "Multiply three to four hours by a couple of hundred and it would have taken a really long time on a single desktop PC." Having access to NeSI meant that they were able to run their jobs on multiple cores at once in parallel.

Due to the size of the job, they figured it would take a few days to complete. "We were surprised when the scheduler allocated more than a hundred cores, and allowed us to complete in one day what otherwise would have taken weeks on a PC in the lab," says Andrew. Along the way, NeSI staff were helping in the background. For example, at one point, when Andrew was having trouble with some of his code, a NeSI staff member even spotted a bug in one line of the code. "All of the NeSI staff members that I interacted with were super helpful," he says. "I attended a tutorial that was expertly delivered and gave me the necessary skills to access and use the supercomputer cluster, and then when we needed specific software (OpenCV) to be installed on the cluster, the NeSI staff did it very quickly." Andrew and his team used NeSI for the later stages of their project as well, to run their benchmarking tests comparing their algorithm with others in the literature.

They've published their findings in the *Journal of Real Time Processing*. With much of the current groundbreaking work in computer vision reliant on access to powerful computing resources, NeSI has become an important tool for researchers such as Andrew to keep pace with the field. "In the computer vision academic community, no one blinks an eye at the thought of needing hundreds or thousands of cores in order to run really complex deep convolutional neural networks, or spending thousands of dollars on GPUs," he says. "Having access to NeSI computing resources means that we can keep up with the state-of-the-art research in this area, otherwise we simply would not be able to run our code in a reasonable time."

"Having access to NeSI computing resources means that we can keep up with the state-of-the-art research in this area, otherwise we simply would not be able to run our code in a reasonable time."

Andrew Chen, PhD Candidate, University of Auckland

Objective 4

Make fit-for-purpose investments aligned with sector needs

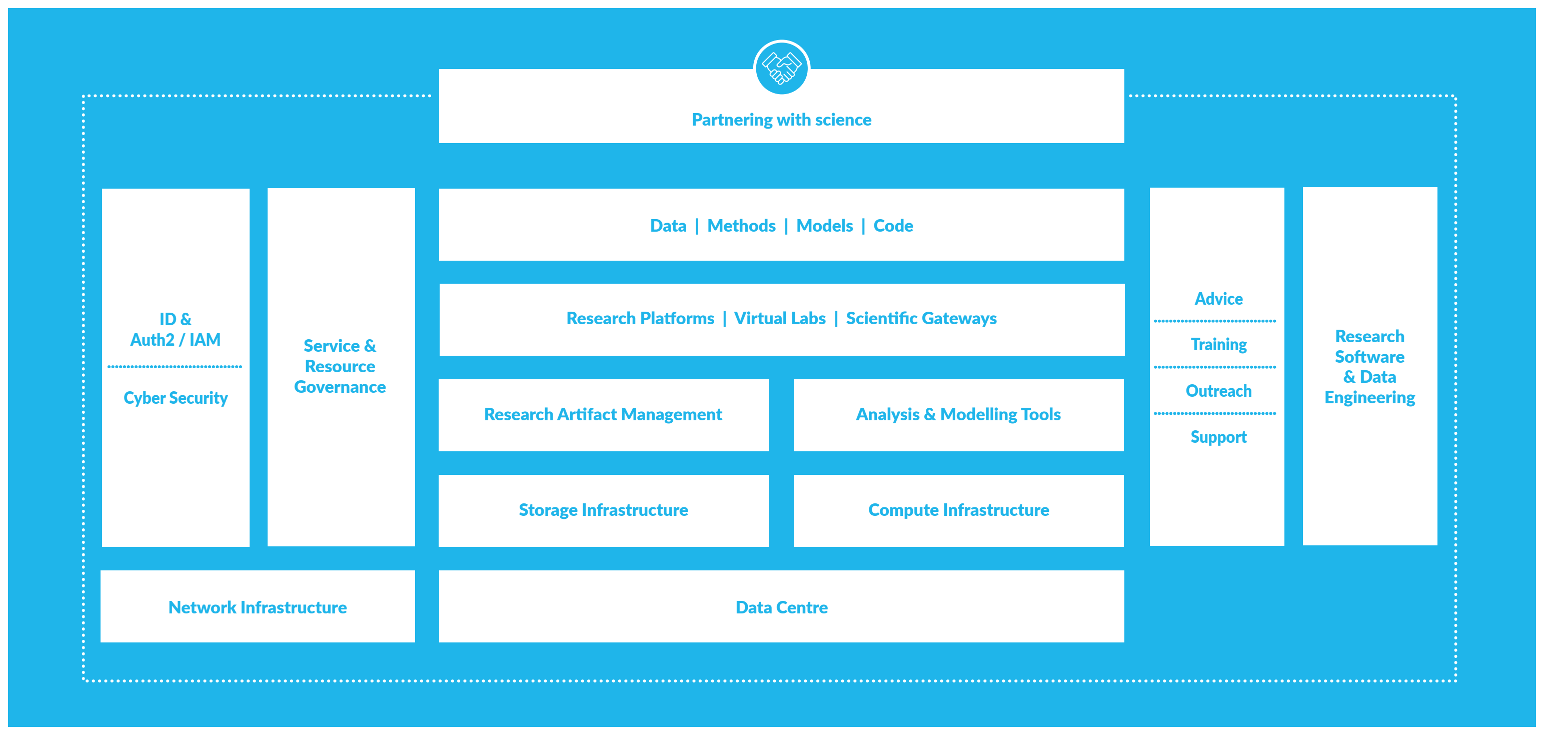
NeSI Futures

Over 2018, NeSI has carried out a range of activities to inform its strategic plans, including a researcher consultation, an international benchmark study, a survey of its migration to its new infrastructure, and a review of its training activities.

The researcher consultation involved interviews with 26 leading researchers from New Zealand from across disciplines and institutions. These discussions presented areas of strategic aspiration and risk for researchers over the next 5+ years, covering the lifetime of the proposed investment.

An international benchmarking study was also undertaken in 2018, exploring both the "what" and "how" of eResearch. This work helped define the components of eResearch ecosystems found around the world, describing how they deliver those services, and identifying best practices relevant to the New Zealand eResearch community.

A national eResearch ecosystem



"The work on this project would not be possible without NeSI. The GPU capabilities at NeSI are especially valuable as they accelerate our calculations by about a factor of 10 and we see this becoming more important in the future."

Blair Blakie, University of Otago

NeSI 3 Investment case 2019-2026

In relationship with the NeSI futures work, in 2018 NeSI began preparing an investment case to seek approval for the Crown to further invest in NeSI from 2019 through to 2026. The case is due for completion in the first half of 2019.

NeSI performed well during its first 7 years as assessed by MBIE through an external evaluation in 2017. Across 2018, NeSI undertook a course of actions aligned with the recommendations from that review, concluding with forming its longer term strategy which will be outlined in the final investment case.

The investment case is for NeSI's third investment period (NeSI.3) in a collaboration between the Crown through MBIE's SSIF (Strategic Science Investment Fund) and NeSI's Collaborators (the University of Auckland, NIWA Manaaki Whenua - Landcare Research, and the University of Otago).

The investment case will aim to:

- clarify the role NeSI plays within the broader research system by articulating its drivers and necessary capabilities and identifying key contributions made by NeSI in lifting computational capabilities and skills
- propose a simpler and more stable longer-term investment to enhance NeSI's sustainability
- outline a revised investment sustaining core infrastructure owned by NeSI's Collaborators and an increased investment into enhancing NeSI's capabilities and services in alignment with MBIE's eResearch Roadmap
- support NeSI evolving to underpin science excellence and impact across a broader range of research communities over the medium to long term
- describe an inclusive science infrastructure platform supporting advanced computational and analytical methods and skills, delivered from within a high-performance facility supported by a critical mass of expertise and sector knowledge

Responding to diversifying needs of research community

In addition to the major platform investments realised in 2018 with the new platforms coming online, NeSI has also implemented an incremental purchase approach to cater for emerging needs in a timely way. The initial example of this is NeSI's additional investment in targeted compute and storage to support our partnership with Genomics Aotearoa. NeSI recognises a growing need for tailoring of specialised virtual data lab solutions and an increase in data-intensity across science investments. NeSI anticipates similar partnerships in support of other SSIF and sector investments such as MBIE's prospective Data Science Platform alongside a renewed investment in REANNZ and scientific databases and collections.

Services are meeting researcher needs

Of the 218 unique responses to a question in our end-of-project allocation survey, 199 researchers (91%) Agree or Strongly Agree that NeSI services are meeting their needs. Of the other responses 15 were neutral, 3 disagreed and 1 strongly disagreed. Some of the comments for these latter responses noted delays and difficulty getting access to the new platforms.

"The benefits beyond the physical computing and storage necessary to make this project happen included the help in setting up the workflow within NeSI and even helping to refine our method because of the specialist knowledge of the NeSI support outside of the HPC. This was fantastic and unexpected too."

Jack Flanagan, University of Auckland

Providing global context to the marine methane paradox

“Parsing through this amount of information requires machines with plenty of memory and computing power.”

Scott Lockwood, PhD Student, University of Otago

Methane is a potent greenhouse gas contributing substantially to global warming. Due to methane’s significant impact on atmospheric warming, understanding the global methane budget is key for formulating realistic plans to mitigate climate change.

Scott Lockwood, a PhD student at the University of Otago, is using NeSI’s newest supercomputer, Mahuika, to investigate some of the causes and contributing factors to methane production. “Currently there is much uncertainty in quantifying the sources and sinks of methane, as atmospheric methane predictions are challenged each year by unexpected fluctuations,” he says. Over-saturation of methane relative to the atmosphere is a universal feature of oceanic surface waters, suggesting surface waters as a source of methane.

However, oceanic surface waters are highly oxygenated, and the classical model of biological methane production can only occur in environments without any oxygen. This inconsistency is dubbed the “marine methane paradox” and over the past decade, researchers have begun unravelling the mystery behind it. Studies have demonstrated the ability for microbes to produce methane through an alternative pathway that involves degrading a specific organic compound, methylphosphonate (MPn).

“Our current research investigates what organisms are involved in MPn cycling and if these key players change in different water masses throughout the year, and the relevance of microbial MPn production and consumption across different oceanic water masses and during different seasons,” Scott says. With sequencing technology improving at an exponential rate and researchers having access to higher and higher sequencing resolution, high performance computing (HPC) resources like NeSI’s have become important tools supporting microbial ecology research like this. “Primary bioinformatics analyses that push the field forward, such as functional screening of metagenomes and assembling genomes from environmental samples, require tremendous amounts of memory and computing power,” Scott says. “Frankly, without NeSI we would only be able to perform a fraction of the metagenome analysis we do now, and an even smaller fraction of the analysis we want to perform in the future.” Much of Scott’s work involves collecting water samples from the ocean, sequencing the DNA found in those samples, and then building enzyme databases and environmental metagenomes from the results.

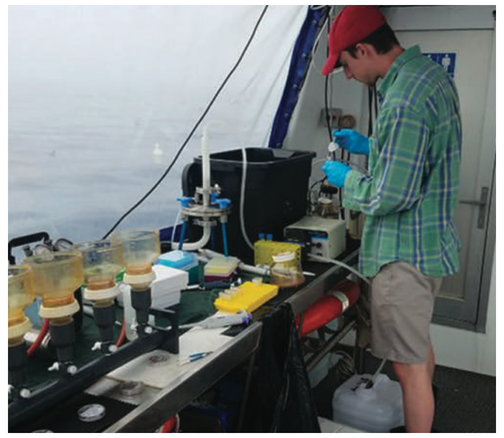
NeSI’s Mahuika cluster has been serving as their bioinformatics homebase. “All of our bioinformatics analysis, from the creation of enzyme databases to screening the metagenomes and assembling genomes, is performed using NeSI’s HPC platform,” he says. Their primary model system for studying the marine methane paradox is the Munida Time-Series Transect, a long-standing transect located up to 65 km off the Otago Coast. This transect spans several types of water, including coastal, Subtropic, and Subantarctic waters that represent all of the Southern Ocean and 40% of global oceans.

Observations from this transect provide global context to the marine methane paradox, giving valuable insight to the players and relevance of an unaccounted methane source in the aerobic ocean.

By looking at metagenomes from different water masses at different times of the year, Scott and his colleagues are able to see whether or not the genes required for MPn production and consumption are present in the microbial community and what organisms are responsible for these functions.

Their current dataset has over 80 million pieces of sequenced DNA representing a vast diversity of microbes. That’s where access to NeSI becomes essential. “Parsing through this amount of information requires machines with plenty of memory and computing power,” says Scott. “If we were to use our in-house servers, analysis that only takes a few days on NeSI’s cluster would take weeks if not months! Not to mention, our in-house server is a shared resource for the entire department and by taking up so much of our limited computing resources, we would ultimately delay everyone else’s work.” When Scott began working on this project at the start of 2018, he had no previous experience in coding or working in the Linux environment. One-on-one support from NeSI’s Support Team – particularly Albert Savary – ensured that wasn’t a barrier. “Albert took the time to meet with me in person, walking me through all of the basics from connecting NeSI to submitting jobs,” Scott says. “I have submitted plenty of tickets and always receive quick, informative responses even if the issue is something silly like changing a password.”

As their next research endeavour, Scott and his colleagues hope to sequence more than four years of samples from the Munida Time-Series Transect Ocean using NovaSeq, the most cutting-edge technology for high-throughput sequencing. “This dataset will provide an incredibly comprehensive view of the microbial community and functional profile across a long-term spatial and temporal gradient,” he says. “It will allow us to perform sample-specific genome assembly to compare how a particular organism adapts to seasonal or spatial changes.” It will also be orders of magnitude larger than the metagenomes they’ve been working with to date. Once again, NeSI resources will play a key role in supporting this work. “Our lab will most definitely continue to use NeSI for the continuation of this project and any other projects that the data may generate,” says Scott.



Scott Lockwood, aboard the Polaris II, working with his sample collections

Objective 5

Enhance national service delivery consistency and performance to position NeSI for growth

Building an adaptive and responsive organisation
The NeSI team moved to a more adaptive way of working through 2018. We work in a complex environment and need to be open to adapting to changes in responsive and proactive ways. NeSI teams have been adopting agile practices and we have introduced a new agile coach role into the organisation to help on this journey. Team culture and purpose was a focus at our 2018 annual retreat. We saw very consistent messages coming from each team around what they enjoyed about working at NeSI.

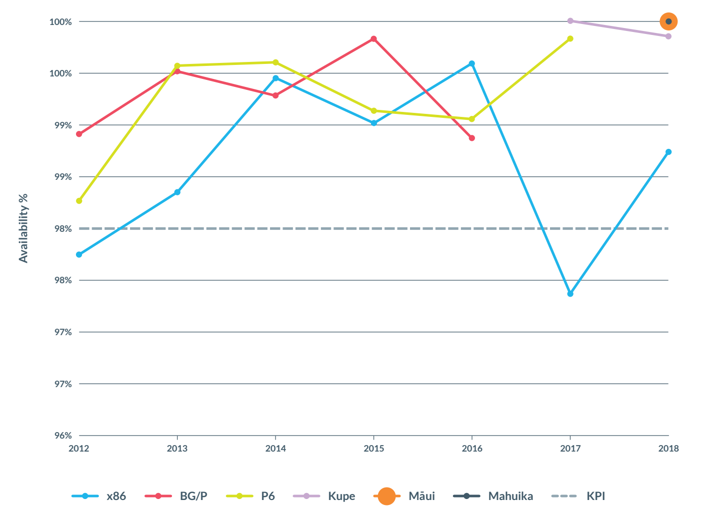
Service delivery consistency
NeSI's new platforms provide a more consolidated and highly-integrated environment for our users. The new infrastructure significantly reduces the barriers to users moving between classes of computing and data systems and they also provide users with a common development and job management and data storage environment. This makes it more efficient to move projects and uses between platforms and also reduces the costs to maintain the platforms.

"It's a fantastic service, that has been, so far, of immense benefit. It's not so much the fast computers, as the coding provided by people, that is the major benefit."

James Sneyd, University of Auckland

Annual availability of services
The annual KPI of availability measures service delivery consistency and performance. Both Māui and Mahuika have been 100% available since they began service for NeSI. This has helped bring our KPI 5 measure up to 99.6% for the year ending 2018.

Availability of NeSI Platforms (percentage of operational time that the platforms are able to be used)



Objective 6

Realise financial contributions and revenue targets to enhance NeSI's sustainability

Our Collaborators' financial requirement to invest at an agreed ratio to Crown contributions, helps track rate of spend against the budget and helps ensure the sector stays engaged with NeSI funding.

The ratio includes all costs across sites, at the directorate for both Opex and Capex spent as well as the accrued Obligation for future spending.

KPI 6

Contract to date, ratio of collaborator commitments to Crown contribution

Ratio of Collaborator commitments to Crown contribution from contract start to date (life-to-date).

TARGET (Annual)	ACTUAL (Last 12 months)
88.6%	95%

As of 31 December 2018, the ratio of Collaborator commitments to Crown contributions is 88.6% LTD (life to date) up from 79.49% in 2017. The target KPI defined in schedule 3 of the Funding Agreement is 95% and accruing reinvestment obligation means we are forecasting hitting target by 30 June 2019.

Controlling invasive predators with local research

Brush-tailed possums are an ecological disaster in New Zealand. An invasive species, they eat native bird eggs, damage trees, and can even carry bovine tuberculosis - a potentially fatal disease for cattle.

Dr Daniel White, a population geneticist and bioinformatician for Manaaki Whenua - Landcare Research, is using NeSI to map the genetic information of brush-tailed possums. His team are working on a project called Achilles Heel. Together, they're designing a way to switch off genes specific to brush-tailed possums, for use as a targeted form of population control.

"The idea is to find variation in genes involved in the vital processes of our target species - the brush-tailed possum - compared to non-target species. Then we design a gene-silencing assay; in this case we're looking at small interfering ribonucleic acid to knock down these genes," Daniel says.

Daniel is searching for specific gene combinations within RNA. RNA acts like a messenger within a cell, taking information from DNA in the cell's nucleus and bringing it to areas such as the ribosomes, which work like factories to produce needed proteins. Silencing or interfering with this RNA can shut off vital functions for the body, causing effects like heart or liver failure. By finding segments of RNA that exist only in brush-tailed possums, Daniel and his team may discover a mechanism for targeted pest control. Unfortunately, the genetic information of brush-tailed possums is not yet well-mapped. "That's where our use of NeSI comes into play.

The challenge is to find sequences empirically because there isn't anything existing to work from," Daniel says. A single sample of tissue can have 180 million reads - snippets of genetic information recorded by a gene sequencer. Daniel and his team are focusing on two sample tissues from the heart and liver of brush-tailed possums. They send samples to the Seoul-based genetics company, Macrogen, before reading the sequences using NeSI's new Mahuika cluster. From each of those 180 million reads the team then have to rebuild how the possum's transcriptome - all the RNA expressed within the sample - was originally put together.

Finally, they compare that genetic information with marsupial relatives of the brush-tailed possum that are already mapped. "At the moment there's information on koalas, Tasmanian devils, the Tammar wallaby and the North American opossum. We compare them to pull out the relevant orthologue gene - meaning genes unique to the brush-tailed possum." Daniel's team were planning to discover target genes by looking at the DNA sequence of these close relatives. Then they would develop RNA-silencing techniques by cloning and sequencing genes in the lab, then testing them in a lengthy trial-and-error approach.

However, winning funding from the Bioheritage National Science Challenge allowed them to use NeSI to streamline this process by directly mapping the brush-tailed possum transcriptome.

"I have to say that to get to the end point of the project we saved a lot of time. The assembly of the transcriptome hasn't taken too long. It took three or four weeks to assemble and pull out the relevant genes. If we'd been in a lab, to get to that stage it might have taken six months, maybe more," Daniel says.

Daniel worked with Peter Maxwell from NeSI's support team to collect and compare this genetic information. Daniel says Peter played an important role in the project, helping the team with software troubleshooting and working with the supercomputer. Another advantage of using NeSI is that the full transcriptome from each sample will be publicly available. Lab experimentation leaves researchers with valuable results but they can be too specific to have wide-ranging applications.

With the transcriptome mapped, future research projects can use that same information, saving time and money. Through their research, the team have found unique target RNA and are now testing silencing techniques in vitro. If effective, the research may be used by other organisations to control the population of brush-tailed possums.

Daniel and project leader Dr. Brian Hopkins (MWLR) hope their research will feature in the Predator Free NZ project. This project aims to rid New Zealand of invasive predators by the year 2050, restoring this country's native wildlife and protecting livestock. Genetic research is a vital part of finding safe, effective population control and NeSI is helping local researchers work towards a predator free 2050.

"The support at NeSI has been great. If I had a problem with software running, or need a new program installed, they were there. It's saved a lot of time."

**Daniel White, Population Geneticist/Bioinformatician
Manaaki Whenua - Landcare Research**

1

Tracking coastal precipitation systems in the tropics

Coastal precipitation plays an important role in the economy of island nations. Impacts from too much or too little precipitation can range from losses in agricultural productivity, to unexpected infrastructure costs, to spikes in sales of particular products or services.

"Land-sea breezes, mountains, and the influence of large-scale circulation all have a role in modulating the daily cycle that brings most of the rain that falls on islands," says University of Auckland researcher David Coppin. "Using a high-resolution model we've been able to simulate the cloud processes over a virtual island of idealized geography, with mountains and vegetation cover. By varying the altitude, size, and aspect ratio of the island, we can investigate the influence of these geographic characteristics, and study the fundamental mechanisms of the daily precipitation. On the other hand, we have a lot of satellite data, but most of the time, they only give you an estimate of precipitation and it is really hard to get a precise idea of what is generating this rain."

To bridge the gap between models and observations, they decided to focus on the cloud systems themselves and to study their temporal and spatial evolution over several years.

"In order to get precipitation right, we first need to get the clouds right and to understand where they come from," says David. "This is a challenge because, during the day, clouds and precipitation will evolve in size, merge with other precipitating systems, or break into smaller systems."

David and his colleagues had a model that labels and tracks precipitation systems, associating a colour to each precipitation system so that researchers can track its origin and temporal and spatial evolution. The model is initialized by satellite data and took about three to four hours to simulate one day globally over the Earth's tropics.

The model was effective on a small scale, but it would have taken two to three months of compute time to label precipitation systems over an entire year. Effectively, the researchers would have had to wait for many months to get their results. So, David sought the help of Alexander Pletzer and Chris Scott from NeSI's Consultancy Team.

First, Alexander and Chris profiled David's code to identify any performance bottlenecks. They found that 34-37% of the execution time was spent in one function. Reimplementing this function in C++ improved the performance by 30%. However, additional performance gains would be difficult to achieve without a complete re-design of some key data structures.

NeSI identified three areas of improvement. First, the code used Python's "numpy" arrays, which are highly optimised to run on high performance computers. These arrays were, however, not a good fit for the problem because they associated precipitation labels to every cell in the grid, even cells with no precipitation or away from any coastline, which was not very computationally efficient. This could be addressed by using a sparse representation of precipitation, which does not consume any memory over "dry" regions.

Secondly, a problem was that precipitation systems arise and disappear over time. In the accounting, labels had to be created and removed from a list dynamically, which was difficult to implement correctly. The solution was to use data structures that can grow and shrink dynamically in time.

Lastly, NeSI's team recommended David replace the least square computation used to approximate systems with ellipses, with one that uses the method of inertia. David's code relied on a function from the skimage library to approximate the contour of a precipitation system with an ellipse. This caused problems when the contour of the system touched the boundary of the domain or when the system had split into more than one part. The method described below overcomes these shortcomings.

Over large regions, NeSI's team expected significant improvements to productivity from the three above recommendations and convinced David to re-design his code around these data structures. NeSI also delivered a first implementation of the code.

With NeSI's help, David re-implemented his code and it now runs nearly 4-6 times faster, uses less memory, and his entire model is more robust and efficient at handling the complexities of his precipitation model. Thanks to NeSI's guidance, David moved the code development under git version control, which allows collaborative code development and ensures that it is always possible to roll back to an earlier version.

"It was great working with NeSI and especially Alex and Chris," said David. "They helped me solve technical problems I would never have been able to sort out on my own. The new code implemented is much faster and runs more efficient on NeSI platforms, which allows me to concentrate more on my research and less on computer science aspects. Thanks to continuous collaboration along the code building, I am now able to maintain and develop it on my own. I learnt a lot from this project and I am really glad that NeSI offers us this kind of expertise for long-term project that requires a lot of work."

> "The new code implemented is much faster and runs more efficient on NeSI platforms, which allows me to concentrate more on my research and less on computer science aspects."
>
> **David Coppin, University of Auckland**

2

Understanding the rupture process of the Kaikōura earthquake

What triggers severe earthquake ground shaking? And when a large earthquake hits, what changes does it cause in surrounding areas?

Using deterministic, physics-based models and a NeSI supercomputer, GNS Seismologist Dr. Yoshi Kaneko and his collaborators are recreating earthquake events to try and answer those and other important questions about earthquake rupture processes and aftermath effects.

In particular, they've been looking at the 7.8 magnitude earthquake that struck Kaikōura in 2016, the country's biggest quake since 2011 when Christchurch was struck by a 6.3 magnitude event.

With help from Alexander Pletzer from NeSI's computational science team, Dr. Kaneko has used an advanced 3D wave propagation code on NeSI supercomputing resources to simulate and understand the quake's rupture process as well as corresponding ground motions and stress changes in the Earth's crust.

"The results of our studies have helped us understand the complex Kaikōura quake and better prepare for the next large earthquake in New Zealand," says Dr. Kaneko. "Without NeSI's supercomputer, we cannot run the software with a sufficient resolution for the research problem."

Part of what made the Kaikōura earthquake so complex is that it not only generated large ground motions and resulted in multiple onshore and offshore fault ruptures, it also caused a profusion of triggered landslides and a regional tsunami.

In his study, Dr Kaneko and his colleagues examined the rupture evolution based on analysis of local seismic and GPS (Global Positioning System) data, combined with numerical simulations of seismic wave propagation.

Their models demonstrate a complex pattern of slowly propagating ruptures from the south to north, with over half of the seismic energy release occurring in the northern source region, mostly on the Kekerengu fault, 60 seconds after the origin time. These results have illuminated the details of the rupture progression and corresponding ground motions throughout New Zealand. They were published in Geophysical Research Letters in October 2017.

In a separate study, Dr Kaneko and his colleagues used the same simulation approach to interpret the cause of widespread slow slip events – gradual fault movement that can last days to months – following the Kaikōura quake.

This time, one of the most intriguing observations was that the Kaikōura quake triggered slow slip events off the Gisborne coast, up to 600 km away from its epicentre in North Canterbury.

Numerical simulations were used to estimate stress changes in the Earth's crust during the passage of the Kaikōura earthquake waves.

"Our results suggest that the slow slip was triggered by stress changes in the Earth's crust caused by passing seismic waves from the Kaikōura quake," says Dr. Kaneko. "We further show that the triggering effect was likely accentuated by an offshore 'sedimentary wedge' – a mass of sedimentary rock piled up at the edge of the subduction zone boundary under the seafloor off the North Island's east coast."

These results also show that layers of more compliant rock is particularly susceptible to trapping seismic energy, which in turn, promotes fault slip at the base of the sedimentary wedge where the slow-slip events occur. These findings were published in Nature Geoscience in September 2017.

As modelling techniques evolve and computing performance improves to make earthquake simulations more realistic and informative, Dr. Kaneko will continue to rely on NeSI's supercomputing resources as an essential tool for his research.

> "Without NeSI's supercomputer, we cannot run the software with a sufficient resolution for the research problem."
>
> **Yoshi Kaneko, GNS**

3

Investigating climate sensitivity

The more we understand about the relationships between atmospheric CO2, water vapour, clouds, and global circulation, the better we will be able to predict climate change.

That's what University of Auckland researcher Tra Dinh is exploring. Recently, she authored an article for the university's "The Big Q", to answer some of the most frequently asked questions about climate change in the IPCC special report.

"I use theory and numerical tools, in combination with observations, to study how the multi-scale interactions of atmospheric processes underline the basic structure of the atmosphere and how these interactions will contribute to future climate change," she says.

Using a circulation model and NeSI's computing facilities, she is investigating how changes in CO2 level is linked to changes in atmospheric humidity and clouds, which feedback on the radiative energy budget of the Earth's climate system, and in turn affecting climate sensitivity.

"The numerical simulations and predictions of the Earth's climate, which are an essential component of my research, are not possible without NeSI's high performance computing facility," she says.

Using NeSI facilities, Tra is able to run and analyse her simulations in a much shorter time, making it easier to identify the results and answers to her research questions.

"Tra is using the High Resolution Atmospheric Model (HiRAM), so access to NeSI is critical for her to run her numerical simulations. The model runs most efficiently on parallel computing platforms," says NeSI's Chris Scott, who worked with Tra on this project.

With multiple interacting variables (e.g. solar insolation, greenhouse gases, humidity, temperature, circulations, clouds) our climate system is incredibly complex and dynamic. For Tra, the multi-variable interactions and their roles in determining Earth's climate are particularly interesting.

"Some of these interactions are nonlinear and they may not be intuitive at first," she says. "For example, recently I found that the seasonal variations of solar insolation have a large influence on the annual-mean spatial distribution of water vapour in the atmosphere. This is work in progress, and this result was found in numerical simulations of Earth's climate using the NeSI computing facility."

"The numerical simulations and predictions of the Earth's climate, which are an essential component of my research, are not possible without NeSI's high performance computing facility."

Tra Dinh, University of Auckland

4

Growing the NZ Carpentries Community

New Zealand eScience Infrastructure (NeSI) has long been a supporter, advocate, and partner of Software Carpentry and now The Carpentries. The nurturing and collaborative values of this global community align well with NeSI's own values and enable us to increase the capability of New Zealand researchers - our organisational purpose.

Herein lies NeSI's challenge - as a relatively small team, how do we have a genuine positive impact on digital literacy in New Zealand? So far, our approach has been to invest in developing our local instructor community, to ensure more local institutions have the skills to deliver their own Carpentry training.

In April this year, we took another step forward in that strategy when NeSI facilitated its fourth Instructor Training event in New Zealand, hosted by Jonny Flutey and his team at Victoria University in Wellington. It attracted a diverse group from across the country, representing a range of domains and organisations. This year's participants included: NeSI, NIWA, Plant & Food Research, Victoria University of Wellington, University of Canterbury, University of Auckland, University of Otago, Massey University, QuakeCoRE, and the Ministry for the Environment.

As Engagement Manager at NeSI, I have been involved in the leadership of Carpentries in New Zealand for a couple of years, but this was my first time diving in as a participant. As someone with little to no background in digital methods, this was a little intimidating. However, the inclusive values of The Carpentries really shone through as I learnt the skills needed to help others develop their own digital confidence.

Our training was led by three of New Zealand's (and indeed the world's) experienced and well-recognised instructors: Aleksandra Pawlik, Jonah Duckles and Murray Cadzow. The contribution of each was well received by those attending. As a newly-minted instructor, I am excited to be part of this growing community and look forward to making my own contributions from now. Thanks again to all of those who made this event a success and I look forward to seeing you all at workshops in the future.

If you would like to learn more about NeSI's Training Strategy check out 'Growing researchers' computational skills to meet future needs', DOI: 10.17608/k6.auckland.7482116.v1; and if you are interested in hosting a Software Carpentry or Data Carpentry workshop at your institution, contact training@nesi.org.nz.

Georgina Rae
Engagement Manager

A Carpentry Workshop: How it Works

A typical workshop runs over two consecutive days and has 40 learners, two instructors, and a few helpers who assist with answering questions during practical sessions.

In a Software Carpentry workshop, instructors teach the core skills needed to be productive in a small research team. The main topics covered include:

- automating tasks using the Unix shell
- structured programming in Python, R, or MATLAB; and version control using Git or Mercurial

In a Data Carpentry workshop, instructors teach core skills required for working with data effectively and reproducibly. To date, Data Carpentry curriculums have been developed in:

- Ecology
- Genomics
- Social Sciences

In all workshops, participants are encouraged to bring their own laptops so that they leave the workshop with a working set of tools installed and operational. Each workshop's instructors will circulate a list of required software (and installation instructions) on the workshop website well in advance of the start.

Short tutorials alternate with hands-on practical exercises, and participants are encouraged both to help one another, and to apply what they are learning to their own research problems during, between, and after sessions.

No pre-requisites are necessary to attend a workshop. The only truly necessary skill is a desire to become more efficient and effective when using a computer in scientific research.

5

Model reveals islands at risk of coastal flooding

Dr. Eddie Beetham, a Research Fellow at the University of Auckland, is studying wave processes on coral reefs to understand the impact of sea level rise on coral reef islands. His work relies on access to NeSI, as his projects involve a combination of field measurements and modelling techniques. All simulations for his research featured below were computed on NeSI's Pan cluster, through the University of Auckland.

The following article, written by Jamie Morton, appeared in *The New Zealand Herald* on 04 November 2018.

Wave-driven flooding is soon expected to make many of the world's coral-fringed coastlines uninhabitable – and now Kiwi scientists have found a way to reveal which are most at risk.

Coral reefs provide a buffer to nearby low-lying communities, through breaking down the energy of waves and effectively regulating how they interact with shorelines.

This natural breakwater effect has also enabled dense urban development around coastal margins on many atoll islands, now extremely vulnerable to sea level rise.

Small and low-lying atoll islands were prone to flooding from spring tides, storm surge and wave "over-topping" - all of it predicted to become increasingly costly in the near future.

At present sea level, destructive flooding events happened every 10 to 30 years, allowing time to recover and rebuild.

But under projected sea level rise - average levels could be around a metre higher by the close of the century - wave-driven flooding events would happen so often they might force exposed areas to be abandoned.

Reef-fringed and atoll islands, alone, were home to nearly 200 million people.

In five nations in particular, more than 500,000 residents had little land to move to and could become climate change refugees.

The scale of the threat made it crucial to find which areas were most endangered.

University of Auckland researchers have devised a simple, accessible method to calculate threshold conditions for wave-driven flooding on reef coastlines, by comparing modelled flooding levels with two measures of island morphology.

Those two measures were relative reef width, which captured how effective the reef was at dissipating waves, and island elevation relative to offshore wave height.

By analysing some 60,000 simulations, they were able to find the point at which waves were likely to breach a coral island reef, which they dubbed the Reef Island Overtopping Threshold, or RIOT.

Their RIOT model was also tested against known flooding events on a number of different atoll islands.

"Our approach, using RIOT, will enable wide-scale and general assessments of island vulnerability, at present and future sea level, based on easily-measured morphological characteristics," study co-author Dr Eddie Beetham said.

"The simplicity of our method makes it broadly applicable to calculate threshold wave heights and sea levels associated with overtopping on different islands."

With this information, island nations would be able to plan for when and where flooding would strike in the future.

The research, co-authored by Professor Paul Kench, also of the university's School of Environment, further showed the impact of sea level rise wouldn't be the same everywhere.

"Sea level rise will have a compounding effect for coastal flooding on reef coastlines," Beetham said.

"The associated increase in reef depth will allow larger waves to reach islands, while the decrease in island height above sea level means that less energy is required for waves to overtop the shoreline."

Islands with maximum shoreline elevation less than 1.5m above spring high tide would be significantly more prone to future overtopping, even if positioned on a shallow and wide reef that effectively dissipates wave energy at present sea level.

On these lower islands, Beetham said, wave overtopping would occur more frequently even with just 10cm to 40cm of sea level rise.

Their study has been published in major scientific journal *Nature Communications*.

> "Sea level rise will have a compounding effect for coastal flooding on reef coastlines."
>
> **Eddie Beetham, University of Auckland**

6

Addressing key questions in evolution

How and when species came to be is the fundamental question in macroevolution. Attempts to answer it use a variety of data sources including genome sequences, morphology, and fossil discoveries. Currently, there is no method for exploiting all of this data simultaneously, while analysing different data sources individually often produces conflicting results.

At the University of Auckland' s Centre for Computational Evolution, a research team led by Professor Alexei Drummond, Dr David Welch and Dr Nick Matzke are using NeSI computing resources to try and address this challenge.

They are developing new mathematical models that will combine genomic, fossil and phenotypic data from multiple sources to give us the best possible understanding of evolutionary history.

"The methods we develop to analyse data are very computationally intensive," says Walter Xie, a research programmer at the Centre. "To analyse a single data set under one set of assumptions may require days for computing time on a regular machine. We typically analyse data under multiple scenarios and need to check and double check that the analysis is valid, so we need a resource like NeSI to be able to use many machines in parallel."

In fact, NeSI resources are an essential tool even before the researchers get to their analysis stage.

"A large part of our work is developing new methods to analyse multiple forms of data in a combined analysis," Walter says. "Since we are developing new methods, we need to validate them on multiple simulated data sets before we can even start using them to analyse real data."

There are two important steps in their work. First, they need to demonstrate that the method works and produces the correct result under the assumptions of the model. This is called model validation. A typical validation involves simulating data sets under a model with known parameter values, then using their methods to see if they can recover the correct parameter values.

"NeSI is used both to simulate the data and run the subsequent analysis," Walter explains. "The analysis of a simulated data set may take hours to run and 100-500 simulated data sets may be analysed. This whole process is repeated many times as the model is refined and debugged."

Once they are satisfied their methods are correct, they can then apply it to real data. They will analyse data sets of interest under various models and assumptions, and use model selection tools to compare them and find the most robust answer. This often comes with its own set of challenges.

"Real data is typically much more difficult to analyse than simulated data," says Walter. "For a start, we don't know the answer before we begin. The methods and models usually require lots of tweaking before reliable results start to appear."

The final output of an analysis is typically a phylogenetic tree that displays the ancestral relationships between the species being analysed and the estimated values of parameters of the model, for example, fossilisation rates over the epochs.

"Without NeSI, these computationally intensive studies would simply not be viable," says Walter. "The batch processing using NeSI saves a lot of our time, and the parallel computing power in NeSI is another big benefit for us."

Looking ahead, Postdoc Fábio Kuriki Mendes is looking to expand their model's capabilities even further, to better accommodate sampling bias in fossil samples. They will continue to use NeSI resources for their testing work.

More information on the project, which is supported by Marsden funding, can be found here. Outside of its New Zealand collaborators, the project also engages overseas experts Associate Professor Tanja Stadler and Dr Tim Vaughan from ETH Zurich, Dr. Mana Dembo and Dr. Mark Collard from Simon Fraser University, and Dr. Graham Slater from the University of Chicago.

> "Without NeSI, these computationally intensive studies would simply not be viable."
>
> **Walter Xie, University of Auckland**

7

Bose Einstein condensate may reveal supersolid's secrets

Dr. Blair Blakie and Dr. Danny Baillie at the University of Otago used NeSI computing resources to support a paper recently published in *Physical Review Letters*. The findings from this research were featured in Ars Technica, an online technology publication. Written by Chris Lee, the below article first appeared in *Ars Technica* on 29 November 2018.

Bose Einstein condensates (BECs) have been around for more than 20 years now. One of the best applications of the BEC, it turns out, is as a tool to explore other quantum things, like solids—yes, the properties of solids are determined by quantum mechanics.

Among those solids is a controversial and possibly nonexistent theoretical one: the supersolid. Now, a pair of theoretical physicists from the University of Otago, Dr. Danny Baillie and Dr. Blair Blakie, have shown that a recently observed BEC "droplet" state may be a way to create a supersolid-like material. That may lead to a way to explore the properties of a supersolid without the difficulties associated with conventional materials.

Bose Einstein condensates

A BEC is a state of matter that requires a specific type of particle. Essentially, the particle world is divided in two: you are either a fermion or a boson. Fermions don't like each other, so they stack themselves in order of energy, from low to high. Any two fermions that are within touching distance of each other must be different. That might mean different energies, or different spins, or some other property, but they must be different. Pretty much everything is made of fermions, and this stacking is what makes the Universe the way it is. Bosons are different. Bosons will happily get together and party in the same state. Not only do they not mind having the same state, they love it. If a couple of bosons find themselves in the same state, they will immediately call all their friends and try to get them to join in. And if the bosons have a way to shed enough energy, they will. The result is that it's relatively easy to create a very cold gas of bosons that are all in exactly the same quantum state, called a BEC.

However, not all bosons and BECs are the same. Although the bosons love to be in the same state, most bosons are composite particles, with some of their components fermions. As a result, there are also forces at work trying to drive atoms out of the BEC. The relative strength between the attraction that pulls the BEC together and the repulsive forces that try to blow it apart is something that is under experimental control.

Exploring weird solids

One of the slightly weirder observations of BECs is that certain types of BEC will form arrays of droplets when the conditions are right. We've seen them form a series of evenly spaced globs of relatively high density, but we don't know why they do this. We also don't know what properties these droplets have. In a reasonably thorough exploration of the phenomena, a pair of theoretical physicists has answered some of these questions. First, droplet formation represents a balance of forces. As we said above, the attractive forces pull the BEC together as the repulsive forces push it apart. For the types of BECs that form droplets, the repulsive forces are not just a function of the density but also the atomic number. A small number of atoms can huddle together nice and close. That number cannot grow very large, though, as that would drive atoms away. That means the BEC forms a number of high-density droplets that all repel each other. Throw in the confines of the trap that controls the BEC, and a nice regular 2D array of droplets emerges.
The researchers wondered if the droplets are really still part of the same BEC. The thing that makes a BEC a single entity is a kind of collective behavior among its atoms. However, that collective behavior requires that the different BEC droplets be connected. That connection takes the form of atoms moving among them.

The researchers showed that if the trap pressed the droplets close together, atoms are regularly transferred among the different droplets, which would allow the BEC to retain its collective behavior.

But critically, they found that under these conditions, the BEC would resemble something called a supersolid. As the droplets move farther apart, atoms can't move between the droplets anymore. At that point, the BEC doesn't resemble a supersolid anymore.

The supersolid transition is important because supersolids are still controversial. Supersolids are, well, solid. That is, they retain their shape like conventional solids. But they also do weird things like slide along surfaces without friction, for instance. There have been several experimentalists that claim to have observed the supersolid state, but none of the experiments have been clear enough to make the claim unequivocal.

Here, we have a tool to explore the supersolid form and its phase transition in detail: a clean environment where the state's properties can be mapped out thoroughly. That should make it easier to refute or support experimental claims in more traditional solid state physics experiments.

Physical Review Letters, 2018, DOI: 10.1103/PhysRevLett.121.195301.

8

Squeezing performance from community code SPECFEM3D

Imagine you buy a new car which you take for a test drive. The speedometer reaches 100 km/h and you discover a switch, hidden somewhere, that suddenly boosts the vehicle's speed to 150 km/h. This scenario often happens in the realm of computing where, thanks to small tweaks, the performance can be greatly improved without adverse side-effects.

Dr. Yoshihiro Kaneko is a seismologist at GNS Science who runs computer simulations on NeSI's platform to compute seismic waves involved in earthquake events, such as the damaging MW 7.8 Kaikoura quake in 2016.

His bread and butter code is the Galerkin spectral finite element code SPECFEM3D, which he ported from the now decommissioned FitzRoy system to NeSI's Cray XC50 supercomputer.

Running SPECFEM3D as fast as possible is important to Dr. Kaneko because it enables him to achieve higher productivity, a larger research output, more papers published, and ultimately a better understanding of how earthquakes are triggered and propagate.

There are many ways an application's performance can be improved. First and foremost, the execution time depends on how a "compiler" translates the program into a set of instructions that can be understood and executed by the computer. Just like each human will express the same concept in slightly different ways, each compiler will map a program more or less efficiently.

There are multiple compilers and each compiler takes myriad different options, only a subset of which impact performance and not necessarily in a positive way. In addition, the NeSI's Cray XC50 supercomputer has the latest of Intel Skylake processors, which support larger vectorization pipelines (AVX-512). This means more operations can be performed simultaneously.

With the help of Alex Pletzer, HPC software engineer at NeSI, Dr. Kaneko was able to identify the compiler and target that produce the best outcome. The graph below shows a 50% speedup on the Cray XC50 supercomputer after switching from the GNU compiler 4.9.3 to the Cray compiler and applying Skylake processor optimizations.

A benchmark simulation ran about 2.5 times faster on the Cray XC50 supercomputer compared to running on FitzRoy. While the Cray compiler was found in this case to produce the fastest code, some software will perform better with either the GNU or Intel compilers. Users on Cray's supercomputers are fortunate to have a choice of compilers, which provides them with more opportunities for higher execution efficiency.

"The computing power available through NeSI is essential for my research," said Dr. Kaneko. "I appreciate the NeSI staff's assistance in helping me move to NeSI's Cray XC50 supercomputer and achieving this significant boost in productivity. I'm looking forward to continuing my projects using this powerful system."

Further details on the speedup work performed for Dr. Kaneko are available on Github: https://github.com/pletzer/perf_kupe/blob/master/specfem3d/specfem3d-tuning.md

> "I appreciate the NeSI staff's assistance in helping me move to NeSI's Cray XC50 supercomputer and achieving this significant boost in productivity."
>
> **Yoshi Kaneko, GNS**

9

How NeSI and NIWA help the UK Met Office develop next generation weather analysis tools

Interpolation is ubiquitous in climate and weather models. Alexander Pletzer and Wolfgang Hayek at NeSI have developed a new interpolation method that will improve the UK Met Office's ability to forecast weather events.

Weather forecast permeates many aspects of daily life, whether it is for planning a weekend with the family, determining whether or not to irrigate crops, identifying the most optimal flight path of an aircraft or evaluating the risk of flash floods.

Behind every numerical weather prediction (NWP) lies a sophisticated computer model, which simulates temperature, pressure, wind, cloud formation and precipitation. Combined with satellite observations, the prediction capability of weather forecast models has improved significantly over the past decades.

The National Institute for Water and Atmospheric research (NIWA), a heavy user of NeSI platforms, runs the Unified Model developed at the UK Met Office daily to predict the weather up to 10 days ahead for New Zealanders.

To improve forecasting capability of high impact weather, resolution must be increased from ~16 km worldwide today to about one kilometre. This is the resolution required to start representing deep air convection, which is responsible for the formation of storm clouds. Unfortunately, an increase of 16x in horizontal resolution translates into a 300+ times higher demand for computational resources, making km-scale models computationally very expensive to run.

To address this challenge, the UK Met Office and partners (NIWA, Bureau of Meteorology Australia, Korea Met Admin, NCMRWF India) are developing a high fidelity weather and climate code named LFRic, after Lewis Fry Richardson, a meteorologist who dreamt about applying parallel computing for weather prediction in the 1920s.

LFRic will rely on a cubed-sphere to alleviate of the problems of numerical instabilities affecting longitude-latitude grids at the poles, apply a "separation of concerns" approach where the mathematical models are decoupled from their implementation to achieve high, exascale parallel performance on a variety of hardware, and leverage a new discretisation scheme based on mixed finite elements where fields are attached to grid nodes, edges, faces and cells.

This new discretisation scheme offers the promises of achieving higher numerical fidelity with fewer degrees of freedom.

Few tools, however, are currently designed to handle vector fields attached to edge and face cells. Simple and widely used post-processing operations, for instance visualisation and interpolation, require fields to be either defined on nodes or cells.

NeSI high performance computing research software engineers Alexander Pletzer and Wolfgang Hayek have developed a new interpolation method [1] that is suited for vector fields on grid cell edges and faces.

Unlike other methods, the new interpolation method conserves line and flux integrals, a property that makes the method "mimetic". This work was recently published in the American Meteorological Society's *Monthly Weather Review Journal*. To read the paper, Mimetic Interpolation of Vector Fields on Arakawa C/D Grids, https://journals.ametsoc.org/doi/10.1175/MWR-D-18-0146.1.

The picture below shows the improvement in accuracy from the dashed, cyan line obtained by applying the most common interpolation method (bilinear), compared to the solid, magenta line obtained using the new, mimetic interpolation method. The error for the mimetic method is more than five orders of magnitude lower.

The mimetic interpolation method has been implemented in a library (MINT), which is openly accessible (https://github.com/pletzer/mint). The library supports arbitrary 2D meshes made of quadrilateral cells and allows users to compute flux integrals.

Many earth science phenomena rely on conservation of water, energy, vorticity and other quantities so it is important to ensure that both dynamic models and post processing tools honour conservation regardless of the grid resolution.

References

[1] https://journals.ametsoc.org/doi/10.1175/MWR-D-18-0146.1

10

Bridging gaps and building collaborations within NZ research communities

NeSI's new Research Communities Advisor is aiming to connect researchers with digital resources and services that can help advance their projects.

Navigating the ins and outs of New Zealand's computational research system can be tricky, particularly if a researcher is new to using digital tools to support their work.

That's where Nooriyah Lohani, NeSI's new Research Communities Advisor, can help. Since joining the NeSI team in June 2018, she has immersed herself in the community, participating in research events, training sessions, and networking activities in order to connect researchers with digital resources and services that can help advance their projects.

With a background in bioinformatics and past roles as a bioinformatician at Pacific Edge in Dunedin and the Bioinformatics Institute at the University of Auckland, she has first-hand experience in both the commercial and academic research spaces.

"I'm very aware and appreciative of the growing need for high performance computing and other digital tools in today's shifting research landscape," she says. "I'm looking forward to connecting with more researchers, particularly those in bioinformatics as that's my specialty, to bridge gaps and build collaborations so that more members of our research communities have the computational tools and skills needed to advance their projects."

Alongside her engagement work with researchers, Nooriyah has also been collaborating with research software engineer (RSE) community leaders in New Zealand and Australia. They've launched a website to connect and build awareness of the NZ & AU RSE communities, and they led NZ & AU's participation in a recent international RSE survey. They hosted a workshop at eResearch Australasia and will be organising other events to allow RSEs to meet, exchange knowledge, and collaborate on methods to create greater recognition and career opportunities.

Locally, Nooriyah helped coordinate NeSI's 2018 Science Coding Conference in Rotorua in August, and is assisting with building the program for eResearch NZ 2019, NeSI's joint conference with REANNZ to bring together researchers across disciplines and communities who are exploring or innovating the ways digital tools are used to advance research outcomes in NZ and internationally.

If you have questions about how computational tools or how NeSI services could support your research, get in touch with Nooriyah by email nooriyah.lohani@nesi.org.nz or visit our Services page on the NeSI website.

NOTES

NeSI
New Zealand eScience Infrastructure

MINISTRY OF BUSINESS, INNOVATION & EMPLOYMENT
HĪKINA WHAKATUTUKI

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND

NIWA
Taihoro Nukurangi

UNIVERSITY OF OTAGO
Te Whare Wānanga o Otāgo
NEW ZEALAND

Manaaki Whenua
Landcare Research