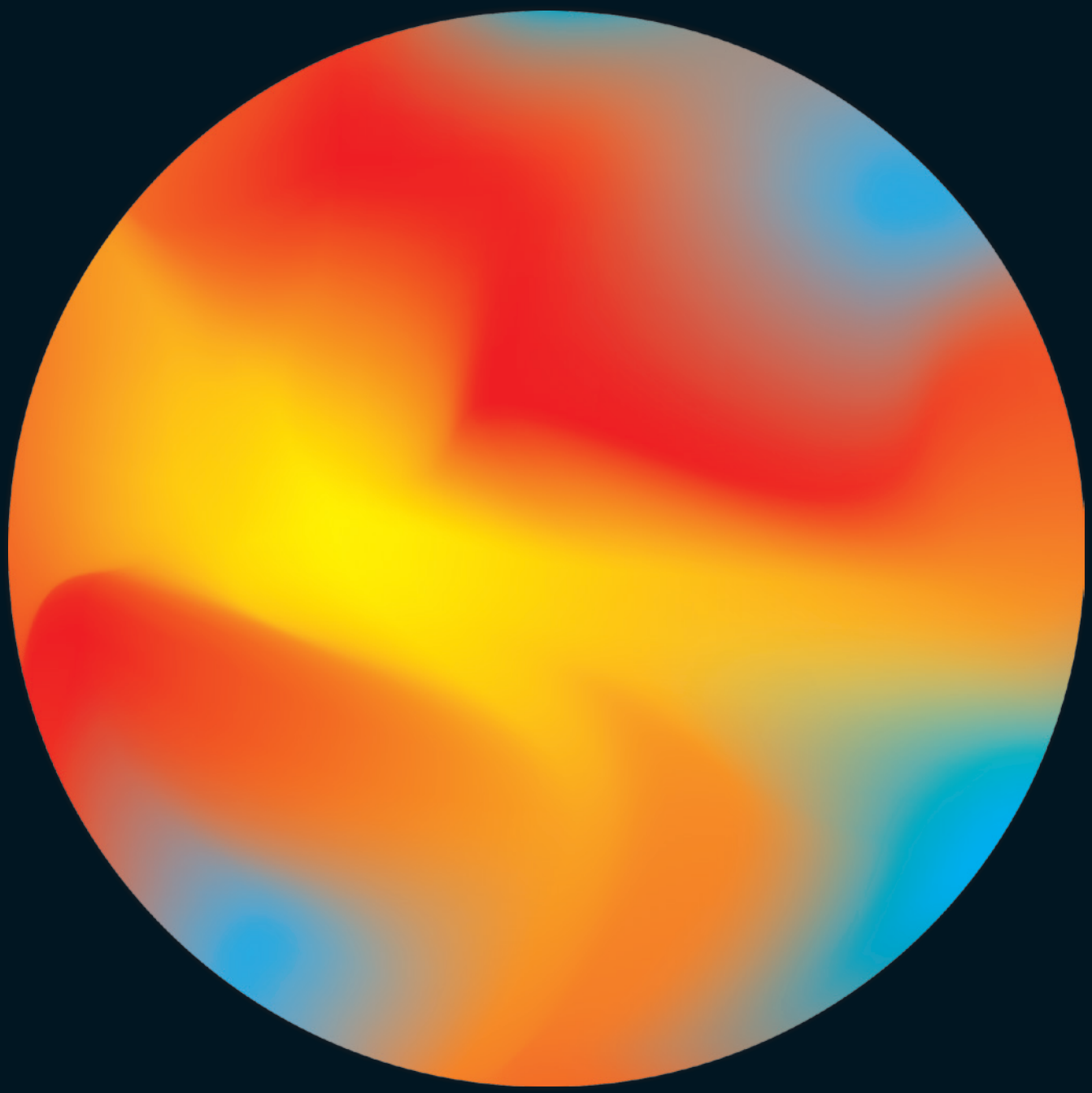
NeSI 2016

New Zealand eScience Infrastructure



NeSI
New Zealand eScience Infrastructure

We live in interesting times, working against the backdrop of a New Zealand research sector undergoing significant development. Sector scope is expanding, via National Science Challenges, Centres of Research Excellence and Strategic Science Investment Fund, to support collaborative research addressing our larger scale challenges as a nation.

Our investments are coming under greater scrutiny and governance oversight; early signs are of our longer term impact starting to show. The year ahead looks to be NeSI's most demanding yet. With the restructured team performing at a high level, our plan for 2017 is ambitious. We will replace our entire infrastructure, refresh and expand our portfolio of services, contribute to the growing efforts around skills and capability development, all while supporting an investment evaluation and subsequently developing our future business case.

Financially, the organisation is on a sound footing. Management has been prudent in allowing time for restructuring and rebuilding of the organisation, while maintaining high standards of service and support within our ongoing business. We completed a review of our long term capital plan early in the year, working with end users, institutional investors, and other stakeholders to arrive at a plan which opens NeSI to broader engagement. This will enable us to adapt to and deal more effectively with the ever increasing volumes of research data and digitally-enabled collaboration, which in turn will drive towards greater excellence and increased impact of research.

One of the most challenging events this year, and a significant distraction for the Board and management, was the withdrawal of the University of Canterbury from the NeSI collaboration, due to ongoing financial pressures which we acknowledge. We have now successfully decommissioned the Canterbury facilities and staff have been redeployed to sustain and grow capabilities on campus.

Over the year the team laid out designs and plans for our future infrastructure. We're already demonstrating effective governance of NeSI's new national shared infrastructure investment approach – our National Platforms Framework. Work on our National Platforms Framework started with research users and stakeholders in late 2015 and was completed early 2016. We took the decision to further consolidate our primary infrastructure down to a single site nationally in NIWA's purpose-built, dedicated facility in Wellington and repurpose the University of Auckland's Tamaki Data Centre as a secondary site for disaster recovery and resilience of our users' data. We've successfully incorporated significant additional aligned investment from NIWA, and along with the parties to NeSI, all are committed to a single integrated investment.

At year's end we released a comprehensive and integrated $19.3M infrastructure RFP jointly between NeSI and NIWA, with NeSI's contribution being $10.8M. We are working with leading HPC and Big Data vendors globally as they respond to the RFP. In late April we anticipate selecting a strategic commercial provider group to join us as the key infrastructure partner over the next six years, and by the end of the year ahead we will have transitioned onto a new infrastructure.

Our eResearch 2020 governance group, comprising Chairs and Directors from NeSI, Research and Education Advanced Network New Zealand (REANNZ) and New Zealand Genomics Ltd (NZGL), struck a few challenges this year. With the release of the Strategic Science Investment Fund (SSIF) we now have a clearer view of the portfolio of research infrastructure investments, along with signals from Government on investment priorities.

The most recent output of this work remains the National Research Data Programme, which is as relevant today as it was when first released early last year. We have supported a small programme of demonstrator projects, which were displayed at the eResearch NZ 2017 conference. Through these we are demonstrating a collaborative model across organisations for lifting our research data capabilities. With the SSIF investment plan signalling a desire for a future research data investment, this work is highly relevant.

NeSI has shown we can share strategic research capabilities effectively while enhancing outcomes for the sector. NeSI is ready to extend its scope of services and capability development to embrace data intensive needs of research communities, offering an obvious base from which to efficiently scale up and embrace an extended mission. With a huge diversity of interesting research challenges ahead, NeSI and the sector can work together to grow data intensive research programmes and attract and retain world-leading talent.

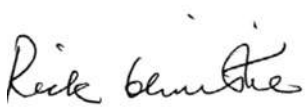
As he launched the SSIF, Minister Joyce's introductory words are more relevant than ever:

"The SSIF is a central part of achieving MBIE's National Statement of Science Investment vision as it will provide:

- *a framework for strategic discussions between the Government and research organisations;*
- *consistent, transparent decision-making and performance monitoring;*
- *stability from longer-term investments;*
- *a means of signalling upcoming investment opportunities."*

As we look to the future, we too hope for the transparency and stability essential to our research infrastructure investments to underpin the excellence and impact of our research.

In conclusion, thanks again to management and my refreshed Board for a very busy and productive 2016, which has ushered in what will be an even busier year in 2017 as we continue to develop and refine New Zealand's research e-infrastructure.



Rick Christie
Chair, Board of Directors

DIRECTOR'S REPORT

2016 was marked for NeSI by a renewed leadership team and refreshed strategic direction, and I've appreciated and enjoyed the commitment and experience of the team as we pursue our goal of growing the advanced digital capabilities of New Zealand's researchers. The team was joined by recognised international experts in advanced digital capabilities and capacity building. We've grown a truly national culture within our collaboration, realising a step change in our ability to support complex challenges of digitally-powered research. It is a privilege to work together on shaping our national advanced computing capabilities for research, and to witness a collective culture that is high-performing, self-assured and welcoming.

Over the arc of a demanding year, a large number of the team has worked closely with key stakeholders and technology vendors in the co-design of our future infrastructure. In a step towards this future, our current infrastructure footprint was consolidated down to two sites, with the mid-year removal from service of HPC platforms at the University of Canterbury. While this decommissioning was part of our long term plans, subsequent to this planned scaling down of investment by Canterbury, they made the hard decision that continued participation in NeSI was no longer financially viable. Canterbury withdrew from being a NeSI collaborator as the year closed, while signalling their intent to re-join the partnership once they reach financial stability.

NeSI's consolidated infrastructure showed increased efficiency and established the form of our future infrastructure footprint. As one of several milestones in moving from multiple local services to a national platform within the current contract, the mid-year decommissioning was aligned with NeSI taking on the national role of fit-for-purpose review and allocation of all workloads across all platforms. Significant gains in efficiency and performance resulted from reallocating and migrating these workloads onto NeSI's two remaining platforms, while doing so provided a smaller-scale test of migration processes essential to smooth transitions onto new infrastructure later in 2017.

Our infrastructure co-design activities spanned most of the year, building on the end-user workshops held late in 2015 as presented in the National Platforms Framework. Starting from a clean slate in Q2, this nine-month co-design process opened space for NeSI and our partners to step back from previous investments to reassess how we might best meet the core needs of New Zealand researchers over the next seven years.

With NeSI's infrastructure scope focused on high performance, the team identified critical gaps of high value in scientific visualisation and in research domain focused virtual laboratories which our new infrastructure will support. To meet demands for better support of research data from our largest and most strategic research communities across National Science Challenges and Centres of Research Excellence, our designs establish big data storage and archive capabilities as core within the national platform, to be harnessed by first class support for high performance data analytics. We anticipate essential discussion on the role of research data in society, and on challenges related to sensitive data and its broader use.

Taking a longer term view, we completed the next phase of our cloud strategy this year, piloting two integrations for cloud-bursting. Both have proven successful and our work provided insights into the relative ease of the integration, and the useful scope of such resources to our traditional markets. We will continue these developments early in 2017, to ensure we are well positioned to exploit such platforms as they evolve the functional support for and performance characteristics required to meet a larger segment of researchers' needs.

Our new National Platforms Framework represents a significant evolution in collaboratively investing into shared infrastructure. This truly collaborative process of co-design and investment is unique in the sector, offering a strategic platform for further advanced computing investments. As future investments in advanced genomics, space science technologies, data science and clinical medicine take shape, NeSI is well positioned to support development in advanced digital capabilities.

Building on our first efforts across 2015, our work on developing researcher digital skills is becoming well recognised. We continue to focus on the sustainability of these initiatives, building up our institutions and research communities to become mostly self-sufficient. We are seeing early signs of a shift in awareness and practice within communities and institutions. Our collective impact in this area is likely to grow in proportion to investment for some time yet, until all research organisations recognise advanced digital skills as essential to improving research excellence and impact.

At mid-year we passed the halfway mark on our current contracts. The renewed team built momentum during the first half of the year, and made strong progress against our milestones across the second. Financially the organisation is performing well, coordinating a longer term capital plan across our partners to achieve the most efficiency out of shared investments. The core of the business is now on a sound footing.

Our next contract evaluation is scheduled early in 2017. Over the past year we've agreed terms of reference and committed to timeframes; this evaluation will be timely to inform future scope and operations of our national investments into research infrastructure platforms. While we continue to expand the range of institutions with whom we work, sustaining and growing this base isn't without challenge. We're seeing early growth in contracting activities with research institutions and high-tech commercial businesses, yet diffusing advanced digital research capabilities into the workforce requires long term partnerships and clarity of common strategic intent; it is still early in our journey.

Looking ahead, through October 24 - 27 2017 we host the international IEEE conference on eScience in Auckland, where we'll be joined by the brightest minds in advanced research computing to share experiences, best practice and explore future directions. Hosting eScience 2017 in New Zealand is a sign of the international regard in which New Zealand is now held, and the trust of our colleagues in NeSI's central role in the region. The year ahead holds much promise as we continue to grow the computing capabilities of New Zealand research.

Nick Jones
Director

Progress against our strategic goals

The NeSI.2 Business Case contained the following action plan identifying key actions to improve performance. NeSI is showing a strong record of progress against this plan, as summarised here.

NeSI.2 Action plan progress to date

Objective	Action Plan	Progress to date
Support New Zealand's research priorities	**1.** Revise Access Policy in Q4 2014 to remove cost recovery for Merit users.	Shifting cost recovery to institutions in 2015 resulted in significant sustained growth in Merit projects and wider sector uptake.
	2. Implement Engagement programme.	Established service governance routines with institutions aligning both parties' goals.
	3. Establish Research Reference Group to advise on strategy and policy.	Group is highly engaged and proactively identifying further opportunities for input.
Grow advanced skills base that can apply high-tech capabilities to challenging research questions in a fit-for-purpose way	**4.** Implement a national training programme early in 2015.	Training delivered to 500+ researchers; 40+ instructors certified. Early stage of impact.
	5. Deliver six training events nationwide each year	Delivered significantly more events, with demand which NeSI is unable to meet at our current scale despite training 40+ additional instructors sector wide.
Increase fit-for-purpose use of national research infrastructure	**6.** Implement a programme of project-led delivery practice across NeSI's internal and customer facing activities, aligned with annual planning cycle.	Programme and project management embedded as team practices, with emerging practice areas of change management and service design.
	7. Scope a service delivery partnership with NZGL in 2014.	After multiple years' development of a NeSI NZGL Alliance, MBIE signalled investment in NZGL to end mid-2017. Work to date has informed NeSI's infrastructure co-design, allowing for expansion to meet genomics needs.
Make fit-for-purpose investments aligned with sector needs	**8.** Define the Platform Acquisition Fund and National Platforms Framework and related processes during Q4 2014, with annual Roadmap reviews thereafter aligned with annual planning cycle.	Fund and Framework are providing effective coordination mechanisms to support co-design of critical national research infrastructure. Linkages to primary research drivers are in place, with a broader set of drivers under development.
Enhance national service delivery consistency and performance to position NeSI for growth	**9.** Reorient NeSI into a Service Line delivery structure in Q4 2014.	Organisation restructured and operating effectively as national functional teams.
	10. Define a programme of service delivery improvement, starting with baselines in Q3/4 2014 and evolve through annual maturity assessments aligned with annual planning cycle.	Annual service planning now driven from maturity assessments and benefits realisation. Improvements delivered through maturing programme and project practices.
Realise financial contributions and revenue targets to enhance NeSI's sustainability	**11.** Achieve budgeted Investor investment levels each year.	Investment levels achieved, with some shifts in timing due to changes to the capital plan.
	12. Agree a consistent pricing and allocation model to provide equity across Investors and Subscribers.	Reporting of clear and consistent pricing and value received has removed ambiguities for all stakeholders.
	13. Implement institutional subscriptions in Q4 2014 and add one additional subscriber annually.	A growing set of subscribing institutions is driving a broadening of NeSI's stakeholders, into industry and public sector agencies.

ACHIEVED SUSPENDED

Our advanced digital capabilities

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

High performance computing - compute and analytics

NeSI makes high performance computing (HPC) available to New Zealand researchers, lifting their ability to tackle large, difficult problems, and enabling research to be carried out much faster. NeSI's facilities are the country's most powerful computers for scientific computing.

During 2016, NeSI fundamentally transformed its service delivery from local to national, driving changes in core infrastructure to underpin these services through the next phase of infrastructure replacement and service evolution. NeSI maintained a high level of service, receiving strong feedback from researchers indicating satisfaction and significant benefits, shown in strong Key Performance Indicator (KPI) and Net

Development of HPC capabilities during 2016

- With HPC infrastructure design driven by research demand, NeSI refreshed a suite of scientific software benchmarks which capture the character of future performance needs.
- NeSI consolidated its infrastructure footprint by decommissioning the University of Canterbury platform, lifting asset utilisation while giving form to future investments.
- NeSI successfully piloted seamless integration with two commercial cloud vendors, demonstrating bursting into cloud providers to meet short-term capacity constraints and tight deadlines.
- NeSI delivered a business case covering a significant evolution of national HPC infrastructure, expanding capabilities into visualisation, virtual labs, and end-to-end integration.
- NeSI's team provides frontline support to get research projects up and running smoothly on our systems and help keep them on track. In 2016 the team responded to over 1,000 support tickets with 99.5% positive feedback.

Fabrice Cantos
Systems Engineer

Gene Soudlenkov
Systems Engineer

Yuriy Halytskyy
Systems Engineer

Kumaresh Rajalingam
Analyst Programmer

Ruan Malan
Solutions Architect

Aaron Hicks
Systems Engineer

Jose Higino
Systems Engineer

Greg Hall
Systems Engineer

Ben Roberts
Application Support Specialist

Peter Maxwell
Application Support Specialist

Jordi Blasco
Systems Integrator

Albert Savary
Application Support

Matt Healey
Application Support Specialist

CASE STUDY

Terrestrial Data Analysis for the Ross Sea region

Terrestrial ice-free areas constitute less than 1% of Antarctica, but represent the most biologically-active sites. The biology within these areas survives in some of the most extreme environments on the planet and is adapted to the unique climate, geology, soils, and geomorphology found in the ice-free areas of Antarctica. Terrestrial areas are thus a focus for conservation efforts because of this uniqueness, but also because they will be the most affected by changing climates.

The Ross Sea region contains more than half of the ice-free ground in Antarctica and so experiences a high impact from scientific and national operations and tourism. This terrestrial environment faces significant challenges, including pressures from increasing human activity, accelerating climate change, and the introduction of non-native species. To understand and mitigate these pressures, the Scientific Committee on Antarctic Research (SCAR) highlights an urgent need to "provide an integrated, comprehensive and dynamic approach to the conservation of Antarctica."

To achieve this we have looked to data science to unlock and explain these pressures. The explosion of data over the last 15 years has had a massive impact on science. Disciplines have moved from data-poor to data-rich and Antarctic terrestrial science is no exception. Utilising the inherent value of Antarctic data is imperative because of the unique nature and importance of the region, and also because collecting data in Antarctica is difficult and expensive. Using this data, we are undertaking a variety of analyses to highlight the range of environmental pressures on the ice-free areas of the Ross Sea region.

Some of these analyses rely entirely on the resources provided by NeSI. In particular, NeSI enables: data transfer and use between members of the project team; undertaking continental scale climate analysis at a 1km spatial resolution on a daily basis over the last 16 years; modelling the regional scale effects of climate change; and modelling thousands of individuals in their likely activity patterns while on the ice-free parts of Antarctica using Agent-based modelling.

Access to NeSI has improved the quality of our analyses and has enabled us to highlight the range of environmental pressures that this part of Antarctica faces. Our MBIE-Funded four-year research programme has key policy outputs and we work alongside the New Zealand Government to ensure these are adopted and used at the Antarctic Treaty meetings to improve environmental management of the continent.

The science goals of the programme are to improve the understanding of how the terrestrial Antarctic environment may respond to climate change and other human impacts; and to facilitate the ongoing improvement of policy development and management of human impacts in Antarctica.

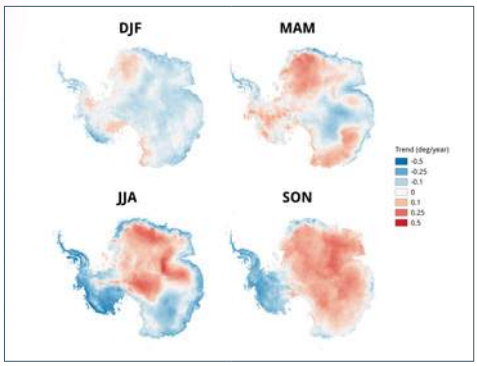


Figure 1. The RegCM recreation of an Antarctic severe wind event (15 May 2004) where extreme winds, sustained 44 metres per second (158 km/h), with gusts greater than 187 km/h. This resulted in torn roofs and blown doors at McMurdo base.

Regional Climate Modelling

Using the Regional Climate Model version 4.4.5 (RegCM), the programme is analysing and predicting the historic, current, and future climate for the Ross Sea region (RSR) of Antarctica. We do this through the temporal and spatial dynamical downscaling of atmospheric state variables (such as wind velocity and temperature) and surface parameters (such as land surface temperature and precipitation) to the Ross Sea region scale. The approach is using existing climate station data and output from coupled global atmospheric-oceanic circulation models (GCM) as input.

This method will resolve regional climate dynamics at a 10 km spatial resolution, a significant improvement over the 100–300 km resolutions from GCMs. To tune the model for the RSR, we will compare surface-based (climate station data) and satellite-derived climate observations (land surface temperature).

The use of the satellite-derived climate observations is a novel approach to tune these models and a necessary one in the RSR because of the paucity of climate stations. The RegCM model for the Ross Sea region can also be initialised by GCM output of future climate change effects under various levels of greenhouse gas emission scenarios such as the Intergovernmental Panel on Climate Change (IPCC) scenarios. This will enable the forecasting of the effects of these scenarios on climate in the RSR. Computationally this is the most intensive utilisation of NeSI, with RegCM being run on FitzRoy.

Continental scale climate analysis

Using data products from the Moderate Resolution Imaging Spectroradiometer (MODIS) on NASA's Terra and Aqua satellites, we have developed a processing pipeline to build a land surface temperature data cube at a 1km spatial resolution over the entire Antarctic continent. The data cube has a daily record for every 1km cell in Antarctica over the last 14 years. Processing, harmonising, storing and analysing this data are only possible through the use of NeSI. The data has also been validated through a comparison with the ground-based stations (Meyer et al 2016). Once complete, we are able to analyse the data cube for a variety of patterns and trends usually hidden when observing on short time periods, such as the way in which the seasons are changing over the last 14 years (see figure below). These types of analyses have assisted in the understanding of environmental changes which contain biological activity through additional analyses focusing on changes in bioregions.

The redder areas in Figure 1 constitute a warming trend while the bluer constitutes a cooling trend. Summer = DJF (December, January, February), Autumn = MAM (March, April, May), Winter = JJA (June, July, August), and Spring = SON (September, October, November).

REFERENCE: Meyer, H.; Katurji, M.; Appelhans, T.; Müller, M.U.; Nauss, T.; Roudier, P.; Zawar-Reza, P. Mapping Daily Air Temperature for Antarctica Based on MODIS LST. Remote Sensing. 2016, 8, 732.

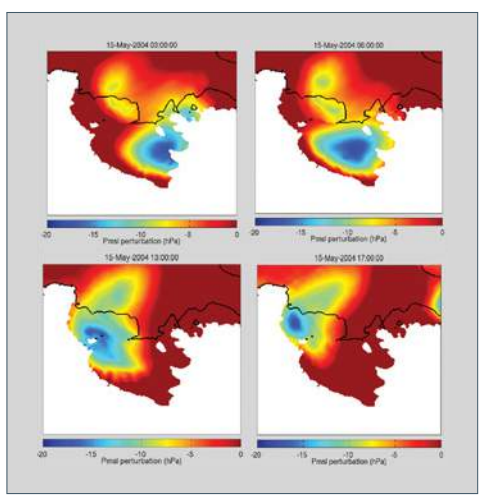


Figure 2. The trends found for each season over the previous 14 years for Antarctica.

"Access to NeSI has improved the quality of our analyses and has enabled us to highlight the range of environmental pressures that this part of Antarctica faces."

Dr Fraser Morgan, a Senior Geographer at Landcare Research, discusses his comprehensive analysis of environmental pressures on the Ross Sea region in Antarctica.

Consultancy

Delivering computational science expertise into projects

NeSI delivers specialist computational science expertise out into the sector, embedding NeSI's team members within research teams for a significant portion of their time. Through this collaboration with researchers, NeSI improves the performance of computational and data analytics software codes and data-driven research workflows, lifting researcher productivity and efficient use of HPC.

Lift in Consultancy capabilities during 2016

- NeSI is improving consultancy practices, with early work completed in 2016 identifying key opportunities to lift capabilities prioritised for delivery during 2017.
- NeSI contributed consultancy services significantly improving speeds, scalability and functionality for researchers involved in key research projects featured in these case studies:

Helping communities anticipate flood events

Improving heart disease prediction and prognosis

Speeding up Basilisk with GPGPUs

Finite element modelling of biological cells

- With the decommissioning of the University of Canterbury (UC) platform, a significant proportion of the team's time focused on supporting the smooth transition of all UC users to NeSI's other platforms. This led to significant performance improvements arising from a better fit-for-purpose allocation of workloads.

Chris Scott
Scientific Programmer

Wolfgang Hayek
Scientific Programmer

Alexander Pletzer
Scientific Programmer

NeSI Consultancy projects for 2016

Project name	Principal Investigator	Affiliation
IO and workflow optimisation for 'Ground motion simulation'	Brendon Bradley	QuakeCoRE, University of Canterbury
Ground motion simulation code migration to Fitzroy	Brendon Bradley	QuakeCoRE, University of Canterbury
Multiscale Modeling of Saliva Secretion	James Sneyd	University of Auckland
Exploring tsunami & flooding code speed-up using GPU processing	Emily Lane	NIWA
TopNet HPC Optimisation	Gabriella Turek	NIWA
Coupled Arterial Cells Simulations of Calcium Dynamics	Constantine Zakkaroff	University of Canterbury

Speeding up Basilisk with GPGPUs

"Helped by the NeSI support crew, we profiled the code and identified some obvious bottlenecks we could remove to start the speed-up process. Pretty soon we had a simple test-case of a trans-oceanic Tsunami making its way across the Pacific being simulated on the Pan cluster".

Dr Emily Lane, a Coastal Modeller and Hydrodynamics Scientist at the National Institute of Water and Atmospheric Research (NIWA), describes her experience as a NeSI user.

I first put in my NeSI application to call someone's bluff. I knew broadly what NeSI was, but had not really looked into it much. My colleagues and I use numerical models to study tsunami and flood inundation and we had heard about codes that had been sped up orders of magnitude by moving the calculations onto GPGPUs (general-purpose graphics processing units). I liked the idea but no one at NIWA had the technical background to translate Basilisk, the code we use for modelling tsunami and flood inundation, to make it run on GPGPUs. We were discussing this at a meeting and someone suggested that I applied for a NeSI project. It sounded appealing but then others told me NeSI didn't 'do that kind of thing' and would offer support but go no further than that. This was just at the start of NeSI 2, though, so I figured what did I have to lose? Worst case scenario they'd just say no.

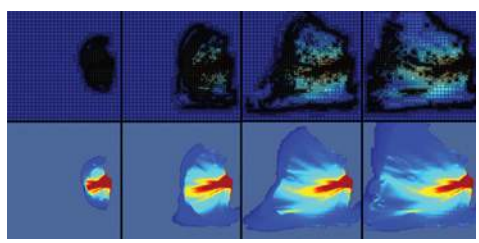
So, I went onto NeSI's website and filled out a form for a proposal. Next thing I know I get an email saying that John Rugis has been assigned as my dedicated contact person – I suddenly realised that it was almost Christmas and I was about to be away for most of the next month. A hastily sent email and we arranged to get back in touch towards the end of January. So in the New Year I arranged a Skype meeting with John. My expectations weren't high but I came out of that first meeting practically bouncing off the walls in excitement! I'd explained a little about the code that I wanted to run and how we were keen to convert it into a GPGPU code. John seemed equally excited about working with us and what we could achieve together. He had ideas about who else he could bring into the project and the whole conversation seemed suffused with a can-do attitude.

Over the next few meetings we worked through the technical details of getting access to the Pan computing cluster and porting the code to it. Helped by the NeSI support crew, we profiled the code and identified some obvious bottlenecks we could remove to start the speed-up process. Pretty soon we had a simple test-case of a trans-oceanic tsunami making its way across the Pacific being simulated on the Pan cluster.

This was the point we'd reached at the conclusion of the preliminary proposal. With assistance from John we filled out an application for a Merit project. We were excited to get approval, but then learned that John was stepping back from the project as he was taking up other opportunities. We were a little worried that we were losing such a valuable member of the team. Our replacement support person was Wolfgang Hayek, who works 50/50 NIWA and NeSI.

John transitioned us over to Wolfgang and pretty soon we realised that, although he didn't start with the experience that John had, he made up for it with enthusiasm in spades. Even better, we still had John's know-how in the background supporting Wolfgang, so it was the best of both worlds. Wolfgang quickly got stuck into the thick of the code and we likewise had our horizons expanded learning the ins and outs of GPGPUs.

One of the defining features of the numerical solver Basilisk, is its adaptive grid. Subdividing squares into four smaller squares in areas where resolution is needed, Basilisk is able to ensure that the grid is resolved enough to calculate accurately where necessary, but is not too fine in other places where nothing is happening. This makes it efficient but also harder to code for GPGPUs. As a starting point, Wolfgang decided to first look at the Cartesian version of Basilisk (i.e. where the equations are solved on a regular square grid). This was an obvious starting step as it more closely aligns with earlier GPGPU algorithms. Another feature of Basilisk is its flexibility. We challenged Wolfgang to design a version that kept a lot of the flexibility while running on GPGPUs.



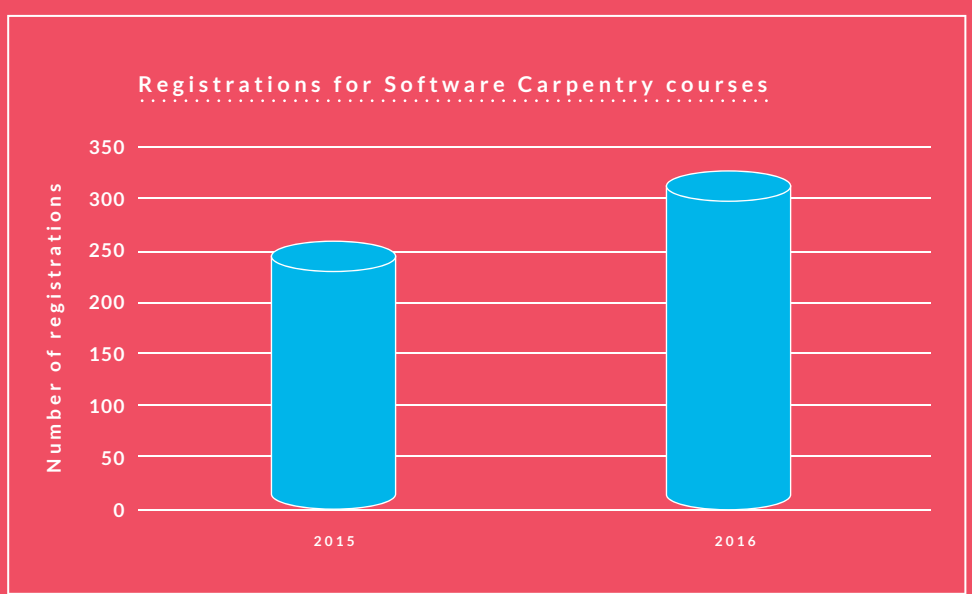
So far we are a year into the project and we've learnt a lot. We've got a Cartesian version of Basilisk that runs on GPGPUs with a good speed-up, while still keeping a lot of the flexibility of the original program. We've also learnt that some of the speed-up claims might be true going from an older, less efficient code to a modern GPGPU code, but aren't realistic for a new, well-written code. With all this under our belt we are excited about the next challenge – seeing whether we can convert the adaptive grid version to GPGPUs. We're pleased that we've got the support of NeSI and Wolfgang behind us because we couldn't do it without them.

Training

NeSI shares expertise and builds capability in research communities and institutions, growing digital research skills and improving researchers' abilities to make use of advanced digital capabilities. NeSI focuses on sustainability by embedding skills to carry out training within institutions and communities, taking a leadership role in growing capacity and changing research culture.

Lift in Training capabilities during 2016:

- A nationally focused Research Communities Manager, Dr Aleksandra Pawlik, joined NeSI from her previous role as Training Lead for the Software Sustainability Institute, UK.
- NeSI supported institutions to run Research Bazaar (ResBaz) events at three locations in 2016: University of Auckland, University of Otago and Victoria University of Wellington.
- NeSI supported two community events – the University of Auckland Digital Skills Winter Bootcamp, and the CRI coding Conference.
- NeSI delivered Software Carpentry events at Victoria University of Wellington, Lincoln University, University of Otago, Plant & Food Research and Scion.
- NeSI has seen a 26% increase in the number of researchers signing up for Software Carpentry training events compared to 2015.
- NeSI trained and supported certification of 16 new Software Carpentry instructors (A further 17 have started the certification process) from across 7 institutions nationally.



NeSI training events for 2016

Date	Event	Location
28-01-2016	Software Carpentry instructor training	University of Auckland
02-02-2016	ResBaz 2016	University of Otago/ University of Auckland
09-02-2016	eResearch NZ 2016 – 10 presentations & workshops	Queenstown
18-04-2016	Plant & Food Software Carpentry workshop	Plant & Food Research
26-04-2016	ResBaz Wellington	Victoria University of Wellington
26-04-2016	VUW Software Carpentry workshop	Victoria University of Wellington
07-06-2016	Palmerston North Software Carpentry (Massey Uni.-Landcare- NeSI)	Massey University (Palmerston North)
07-06-2016	NeSI-Massey Albany Software Carpentry	Massey University (Albany)
09-06-2016	NeSI Seminar	Massey University (Palmerston North)
22-06-2016	Canterbury Software Carpentry (Uni. of Canterbury-NeSI)	Canterbury University
29-06-2016	NeSI University of Otago Software Carpentry Workshop	University of Otago
11-07-2016	Software Carpentry with R (University of Auckland and NeSI - Winter Bootcamp)	University of Auckland
11-08-2016	Software Carpentry with Python (University of Auckland and NeSI - Winter Bootcamp)	University of Auckland
15-07-2016	Introduction to OpenMP - entry level parallelisation	University of Auckland
15-07-2016	Introduction to MPI (Message Passing Interface)	University of Auckland
01-08-2016	CRI Coding Conference	NIWA (Auckland)
08-08-2016	ResBaz Promotion and Organisation	University of Canterbury
22-08-2016	University of Otago Medical School NeSI Seminar	University of Otago (Christchurch campus)
29-08-2016	Queenstown Research Week Conference	Nelson
01-09-2016	NeSI & Victoria University Software Carpentry Workshop	Victoria University of Wellington
15-09-2016	Centre for Advanced Composite Materials NeSI Seminar	University of Auckland
19-09-2016	NeSI presentation at the Waikato University	University of Waikato
22-09-2016	Open Data Charter workshop - LINZ and Auckland Council	Auckland Council
11-10-2016	Te Punaha Matatini eResearch workshop	University of Auckland
14-10-2016	Te Punaha Matatini seminar	University of Auckland
17-10-2016	Lincoln CRI Digital Skills - Software Carpentry Workshop	CRIs at Lincoln and Lincoln University
21-10-2016	Practical Introduction to High Performance Computing with NeSI	Landcare Research
31-10-2016	NeSI at Otago Day	University of Otago
01-11-2016	University of Otago & NeSI Software Carpentry Workshop	University of Otago
14-11-2016	Supercomputing 16*	Salt Lake City, USA
15-11-2016	NIWA Wellington & NeSI Software Carpentry Workshop	NIWA
24-11-2016	NeSI & Scion Software Carpentry Workshop	Scion
29-11-2016	Internal NeSI Software and Data Carpentry Instructor Training	NIWA
14-12-2016	Online Software and Data Carpentry Instructor Training	Online

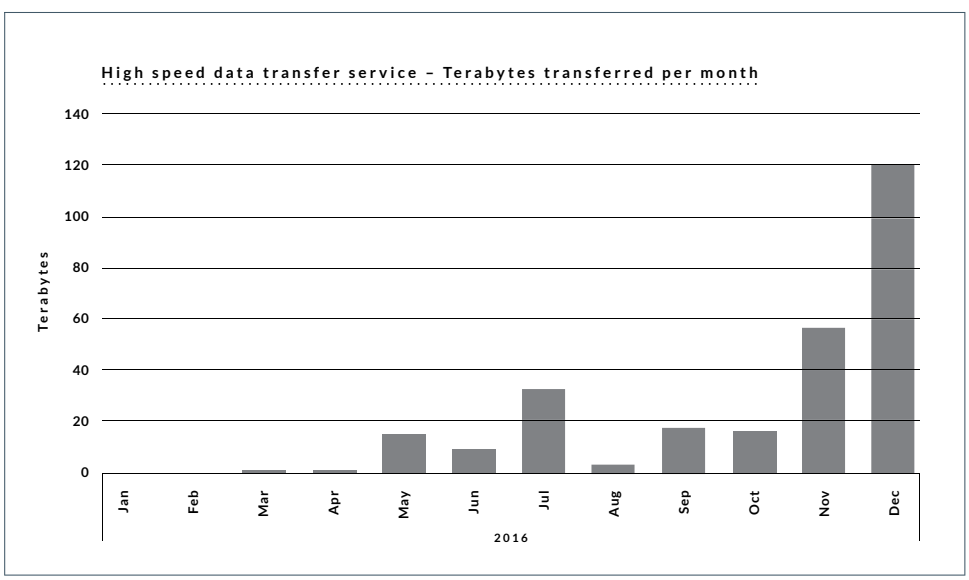
***Data Carpentry in HPC talk; NCSA collaboration workshop NeSI presentation; discussion on training strategies with other national initiatives.**

Data transfer and share

NeSI supports researchers to transfer and share their research data sets, working with research communities, institutions and REANNZ to improve end-to-end performance and reduce time to solution. During 2016, NeSI saw a significant increase in the number and scale of data-intensive research project proposals. The largest projects are driving rapid growth in data transfers into NeSI.

Lift in data transfer and share capabilities during 2016

- Working with the University of Otago and REANNZ, NeSI launched new high-performance data transfer infrastructure, supporting rapid growth in data-intensive research at Otago. This is a best-in-class implementation that has moved almost 300 Terabytes of research data in its first six months of service and shows signs of ongoing growth.
- Improvements were made to service monitoring, robustness and reliability, leading to further improvements in design and implementation of NeSI's core infrastructure.
- NeSI, working within eResearch 2020 alongside NZGL and REANNZ, supported three projects aimed at lifting research data capabilities across the sector, to be completed in early 2017.
- NeSI has worked on Genomics projects in 2016 where the data sets are 40+ Terabytes, with more prospective big data projects planned for 2017. These projects were invaluable to help us identify the requirements and capabilities to manage big data transfer, storage, backup and analysis that are now intrinsic to the design of the services to be run on the new NeSI platforms due to be installed in 2017.



CASE STUDY

University of Otago researchers join the international high-speed data superhighway

Researchers already influencing global leaders have now started another world-leading project thanks to cutting-edge technological advances at the University of Otago. None of the country's commercial services could easily provide the service that the Dunedin Multidisciplinary Health and Development Study Research Unit needed to send MRI scan images securely at high speeds to a partner laboratory at Duke University in the United States, for analysis.

The unit, led by Professor Richie Poulton, had been dreaming about including neuroimaging in its research for more than a decade and had secured the funding. So, early last year, the researchers started planning to scan the brains of about 1000 New Zealanders they have been studying since their births more than 40 years ago.

From megabits to eternity

Even though the University was struggling to send data off campus faster than about three hundred megabits a second, Information Technology Services staff were determined to make the project work. Otago researchers had been resorting to couriering external hard drives because the University's security and bandwidth controls were slowing down big data transfers.

Stupendous study

The research unit, also known as the Dunedin Study, is well respected and worldwide interest in its findings is so acute, it has attracted more than $18 million in offshore funding over the years. The researchers have produced a paper for a scientific journal every 13 days on average for more than 40 years, providing a treasure trove of information about human health and well being.

The researchers' finding that most youths grow out of crime even sparked a law change in the United States, so juveniles can no longer be sentenced to death, saving the lives of more than 70 young people then on death row.

Duke University's Professor Ahmad Hariri says the study's next step – analysing study member's brain scans while also referring to their life histories – will help understand how differences in brain structure and function could reflect people's past experiences and predict their mental health.

Ducks in a row

The University of Otago partnered with Pacific Radiology to obtain a powerful 3 Tesla Siemens MRI scanner and have professional help running it, says David Maclaurin – University of Otago Information Technology Services (ITS) Infrastructure and Applications, Systems Services Manager.

To create the supporting IT infrastructure, the University's Brain Research New Zealand Co-Director Professor Cliff Abraham, formed a project team that met every few weeks for about eight months. It included people from ITS Infrastructure and Applications, Pacific Radiology and subcontractors.

Securely Bridging Networks

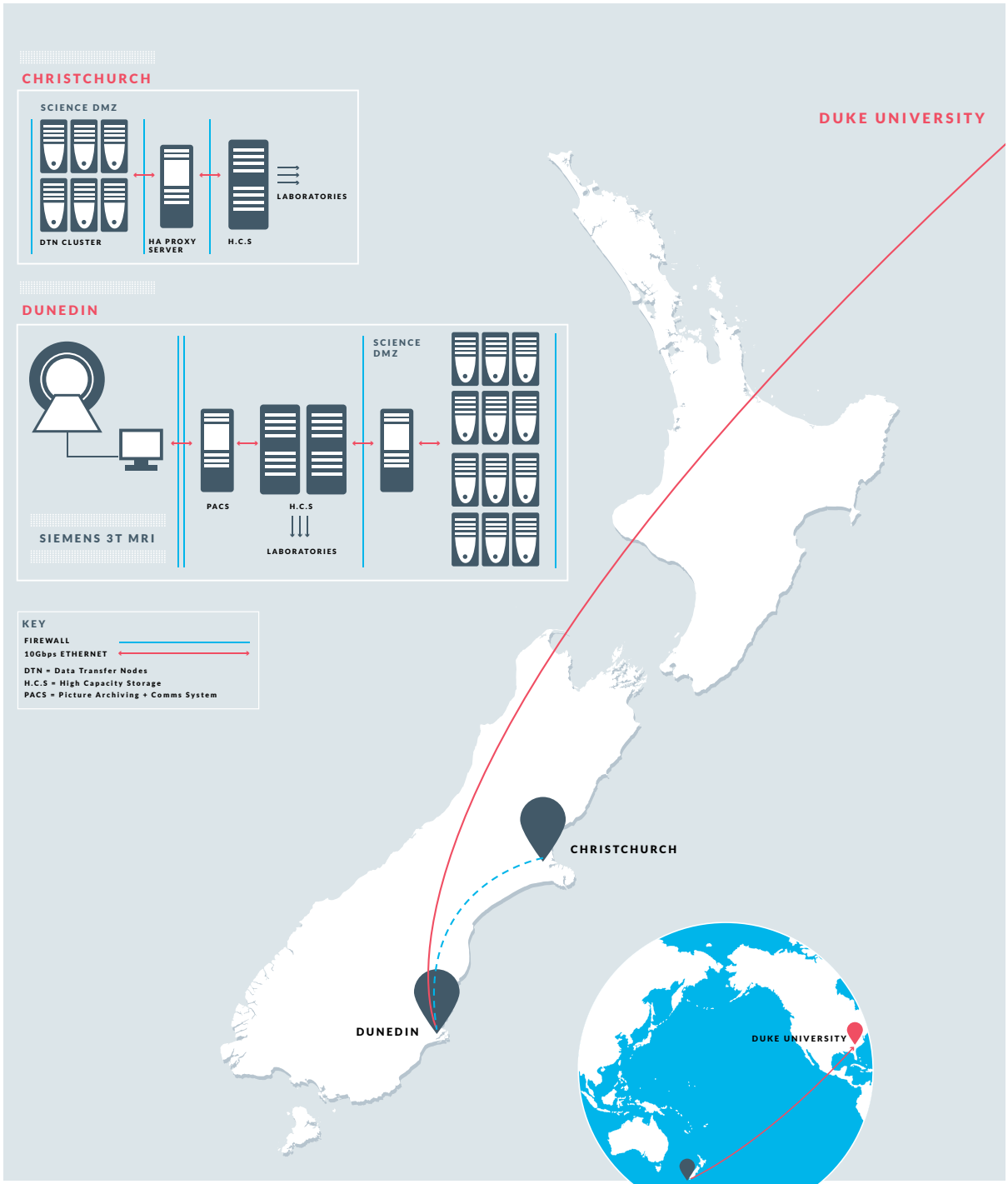
They had to make sure only the appropriate scan images flowed between Pacific Radiology and the University, protecting the privacy of both the University's study members and the company's patients. Designing and planning a network-bridge, servers, storage and firewalls took several months, then several more weeks to get test scans to flow through them at up to 10 gigabits a second.

Speaking different languages

The team also had to get the MRI scanner and the University's servers to "speak the same language." The scanner used a data exchange protocol called Digital Imaging and Communications in Medicine (DICOM) which the ITS team members had never worked with before. They phoned the manufacturer and other research groups numerous times to unearth software to communicate with DICOM, and discovered an open source Picture Archiving and Communication System (PACS) server.

Once it was operating in a University Data Centre, the PACS server became visible to the MRI scanner as a DICOM endpoint. Then, the ITS team taught themselves how to run the server, and discovered the best way to anonymise each scan image and tag it, working with the researchers and Pacific Radiology.

From left: Professor Richie Poulton, Professor Cliff Abraham, and Pacific Radiology Otago-Southland Managing Partner Dr James Fulton with the new MRI scanner before it went live.


CHRISTCHURCH
SCIENCE DMZ
DTN CLUSTER
HA PROXY SERVER
H.C.S
LABORATORIES
DUNEDIN
SCIENCE DMZ
PACS
H.C.S
LABORATORIES
SIEMENS 3T MRI
KEY
FIREWALL
10Gbps ETHERNET
DTN = Data Transfer Nodes
H.C.S = High Capacity Storage
PACS = Picture Archiving + Comms System
DUKE UNIVERSITY
CHRISTCHURCH
DUNEDIN
DUKE UNIVERSITY

Massive storage
The ITS team decided to configure the new PACS server so the data could be kept in the University's High Capacity Storage clusters, which authorised "Dunedin Study" researchers can access easily and securely. Fortunately, about three years previously, the University had created storage clusters that can hold petabytes of data.

An on-ramp to the superhighway
Otago still struggled to send data off campus faster than about three hundred megabits a second and still had to get the MRI scan images to Duke University's Laboratory of NeuroGenetics. To create a new, fast, data pathway, the ITS team members started doing their own research and discovered an architecture known as a 'Science DMZ', from the Energy Sciences Network (ESnet).

Demilitarised zone
The Science DMZ is a demilitarised zone between the edges of two networks, specifically designed to let big data sets flow in and out with the fewest obstructions, at the highest network speeds possible.

To create a Science DMZ from a combination of new hardware called a network managed edge and several types of software, the ITS team sought help from the Research and Education Advanced Network New Zealand (REANNZ). It says the Science DMZ gives Otago security similar to that of concierges in office buildings: "They've seen you before and they know to let you in … we think of them as on-ramps to the superhighway."

Now, everyday University internet traffic can still can arrive and leave with the usual checks, while authorised research data bypasses any checks it does not need.

Collaborating with NeSI
ITS team members still needed an 'engine' to drive the transfer of large volumes of data at continuously high speeds and so sought more help from the New Zealand eScience Infrastructure (NeSI). Mr Maclaurin and the equivalent of 1.5 FTE of his team work part-time for both the University of Otago and NeSI, and the University is a NeSI collaborator.

An engine to drive data transfers
The ITS team worked with NeSI staff at the National Institute of Water and Atmospheric Research to create a secure Data Transfer Service with Data Transfer Node software – provided by Globus – which has not been combined with Science DMZ in New Zealand before.

NeSI helped create the blueprint for the service, then the ITS team worked out how many servers it needed, how the servers had to be organised in relation to each other, and internally – which included gauging how much data they could each transfer and at what speeds.

Mr Maclaurin says the software can "hurl huge volumes of data down the network 'pipeline' in one go by transferring many files at once, and by splitting files into a number of parts that can all be streamed simultaneously, then reassembled at the other end."

NeSI relationship vital
Otago would not have gone down this route without NeSI; their help was critical.

"NeSI provides its partners with access to best High Performance Computing practices from around the world; and with the concepts, ideas and practical help to make these sorts of research IT services a reality," Mr Maclaurin says. "Research is very near to my heart and is what makes a University special. I'm determined that by collaborating with partners like NeSI, we're going to deliver the best IT services we can for our researchers."

The University also used NeSI's Globus encryption option to keep the 'Dunedin Study' data confidential on its way to the United States.

Duke out at Otago
At Duke University, a receiver for the scan images – also based on Globus software – was installed by Research Project Manager for the Dunedin Brain Imaging Study Annchen Knodt, with help from New Zealand. While the programmer and analyst's end point at Duke had less capacity and speed than Otago's, it could do the job.

When Miss Knodt arrived in Dunedin for six months, she was ecstatic the IT infrastructure had already been created by the ITS team, NeSI and REANNZ.

"I was expecting a whole lot of stress, moving across the world and getting things set up, and I didn't have to do that. The pipelines were in place."

"NeSI provides its partners with access to best High Performance Computing practices from around the world; and with the concepts, ideas and practical help to make these sorts of research IT services a reality," Mr Maclaurin says. "Research is very near to my heart and is what makes a University special. I'm determined that by collaborating with partners like NeSI, we're going to deliver the best IT services we can for our researchers."

"This is the kind of direction it's going in. Both these organisations are pushing the boundaries and know about the technology all around the world. They're sowing the seeds and people like us are running with it – it's fun and very rewarding."

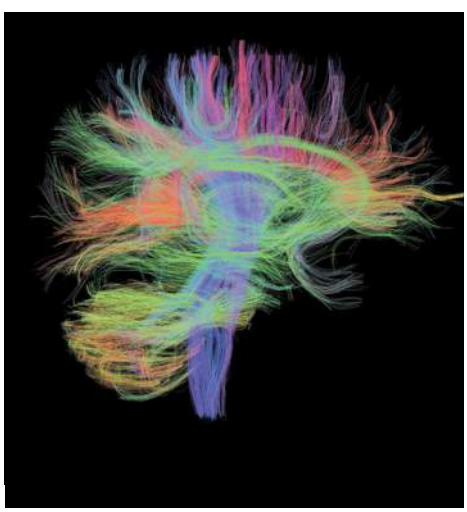
Perfecting scan images

She could concentrate on her main task; training "Dunedin Study" researchers to perform the scans, which includes study members playing games in the scanner that are designed to highlight how specific areas of their brain are functioning – taking about 70 minutes for each person.

One game involves the possibility of winning money if study members respond fast enough, sparking activity in the reward centre of their brain, showing how each person's response varies. That information is combined with each person's life history to investigate how the variations in responses are reflected in other aspects of their personalities.

The results

The study's scan images started reaching the United States about two months ago. So far, about 116 study members have been scanned and the researchers hope to scan all study members every five years, to find out what has changed or stayed the same, and how changes are reflected in people's health and wellbeing.



3T MRI diffusion scan of connecting axon pathways

Health checks

Meanwhile, the quality of the scan data is regularly monitored by medical physicist Dr Tracy Melzer of the New Zealand Brain Research Institute, Christchurch. He helped set up the scanner's parameters, and examines it remotely from Christchurch regularly to ensure its images maintain a consistent quality.

Success at last

Now the data pipeline is in place and working, Mr Maclaurin knows being a research-intensive organisation at the bottom of the world is not a problem anymore: "We can transfer research data up to 15 times faster." Scan images direct from the MRI can even reach Duke University in well under three minutes if necessary.

The flow on

Since the data transfer system was developed, more than 250 terabytes of data – about 15 times the University's total centralised storage in 2010 – have been securely transferred around the country and the world for a range of Otago projects, at speeds of between two and five gigabits a second.

Mr Maclaurin believes that would not have been possible without NeSI and REANNZ.

"This is the kind of direction it's going in. Both these organisations are pushing the boundaries and know about the technology all around the world. They're sowing the seeds and people like us are running with it – it's fun and very rewarding," he says.

The future

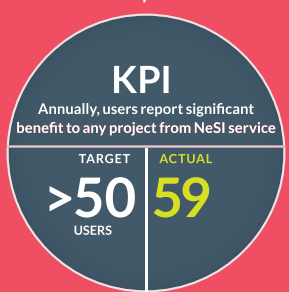
Now, Mr Maclaurin and his team are trying to get even more speed out of the system, and are still working closely with the Dunedin Study, because "it's no good just saying 'here's a thing, see you around'. It has to be a partnership."

Meeting our objectives

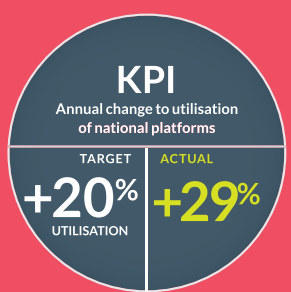
NeSI focuses on the following objectives, tracking performance against related Key Performance Indicators (KPIs).

1 Support New Zealand's research priorities

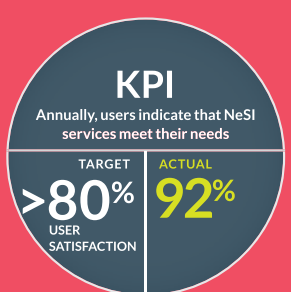
2 Grow advanced skills that can apply high-tech capabilities to challenging research questions



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



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



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

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

Across 2016 NeSI evidenced strong indications that the renewed organisational strategies and structures are bedding down successfully and delivering impact sector wide.

*During 2016 NeSI adjusted its capital investment plan, aligning all major capital investments inside one investment round, due to complete early in 2018, at which point KPI 6 will be back on track.

Objective 1

Support New Zealand's research priorities

Georgina Rae
Engagement Manager

Engaging with and influencing research policy

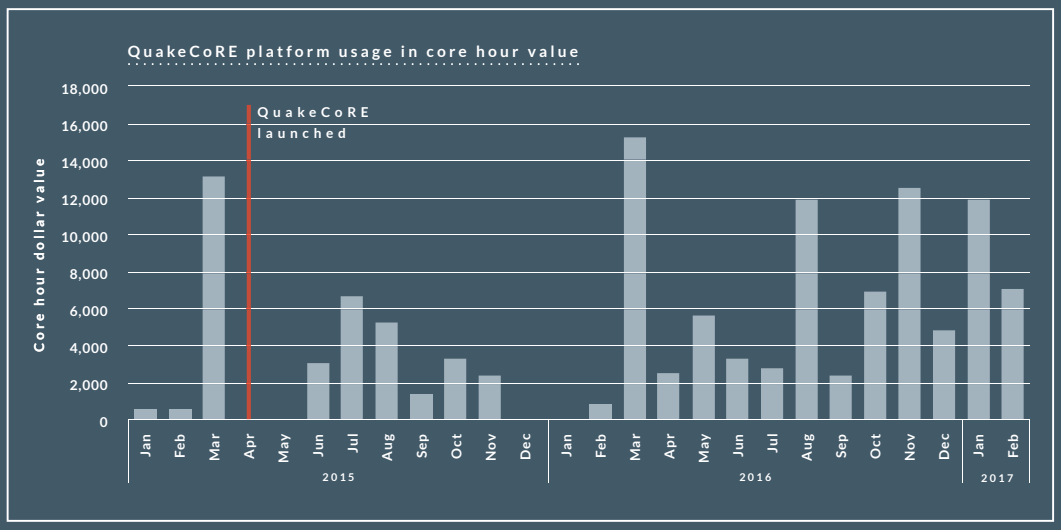
With the release of the National Statement of Science Investment in 2015, an early focus for MBIE in 2016 was on policy and prioritisation of research infrastructure investment. Released alongside the Crown Budget 2016, MBIE announced the formation of the Strategic Science Investment Fund (SSIF) in Q2, folding all research infrastructure appropriations into a broader investment portfolio. Later in 2016 MBIE released the SSIF Investment Plan, outlining investment logic and signaling future investment intentions. Within this framework, MBIE has signaled ongoing support for High Performance Computing, with the investment seen as strategic, addressing path dependency problems and thereby unlocking new opportunities for New Zealand.

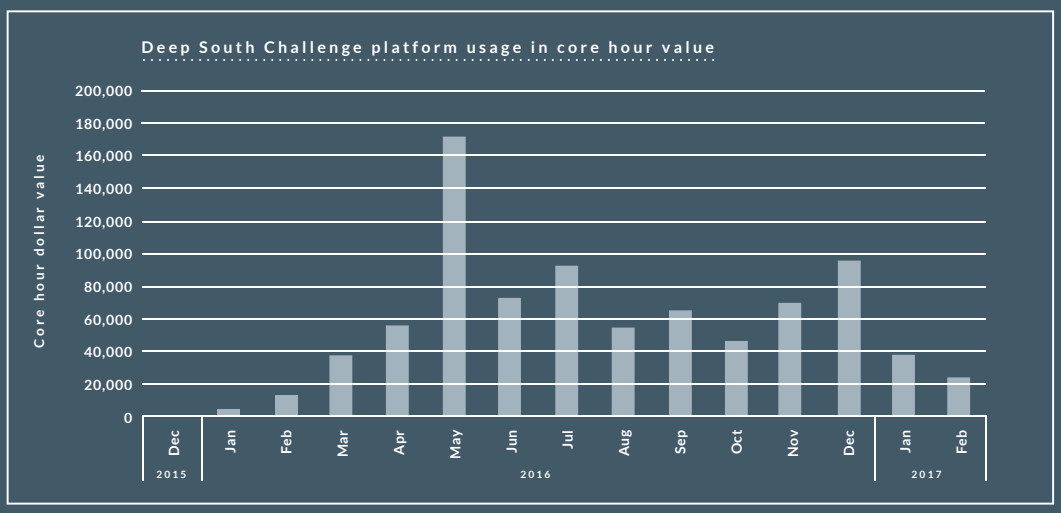
NeSI was consulted on several policy developments across the year, including the Australia – New Zealand Science, Research and Innovation Cooperation Agreement. NeSI has been working alongside Australian investments from it's inception, including participation in a mid-March workshop on the eResearch Framework mapping out Australia's direction through to 2025. Released early in 2017, the trans-Tasman Cooperation Agreement outlines key areas of future collaboration, including prioritising trans-Tasman research and research infrastructure collaborations.

Building capabilities within national research collaborations NeSI focused on building capabilities within key research sector investments during 2016, including supporting proposals for the new Advanced Genomics Platform, a Centre for Space Science Technologies regional research institute, and transferring skills and capabilities into existing investments in the form of National Science Challenges and Centres of Research Excellence.

Two key national collaborations NeSI worked closely with during 2016 are QuakeCoRE and the Deep South Challenge. Both carried out research projects requiring significant sustained HPC support, and in both cases NeSI provided consultancy to enable this work.

> "Its whole implementation was helped by NeSI staff. They did a fantastic job, and covered technical aspects that we hadn't even considered."
>
> **Jack Flanagan**



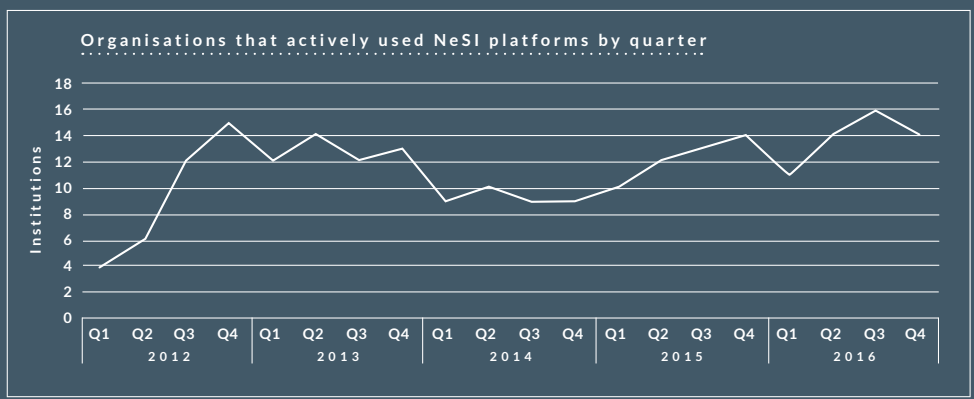


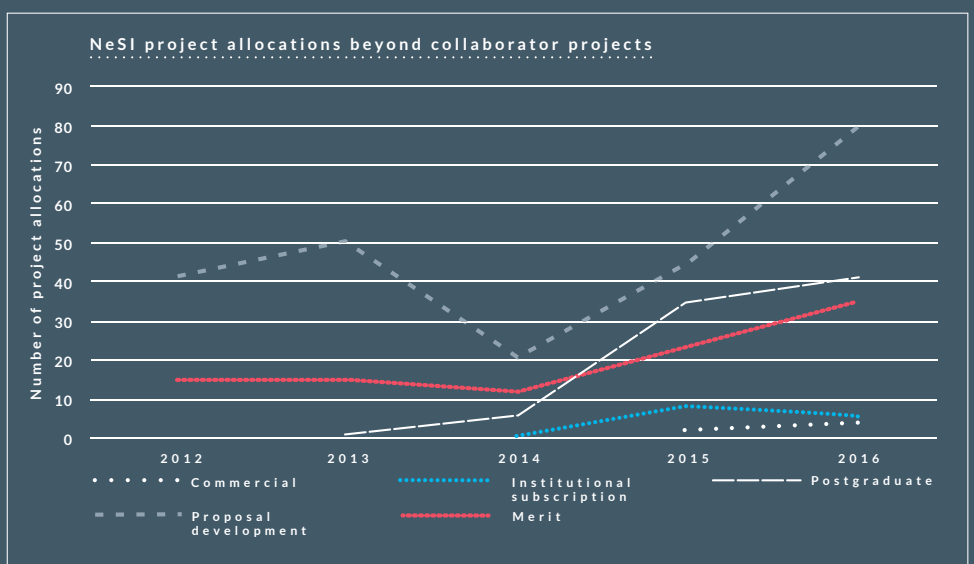
Growing broader sector uptake and impacts

Early in 2013 NeSI experienced a noticeable decline in projects after introducing pricing to the researcher. The impact of this pricing was mostly on non-collaborator institutions, reducing the breadth of institutions receiving value sector-wide. During 2016 NeSI's Access Policy was revised, substituting pricing to the researcher for pricing to institutions via subscriptions. 2015 showed a strong recovery in active institutions which has continued through 2016, highlighting the breadth of impact NeSI is now sustaining.

"I would not have been able to complete my thesis within my deadline without utilising NeSI HPC resources."

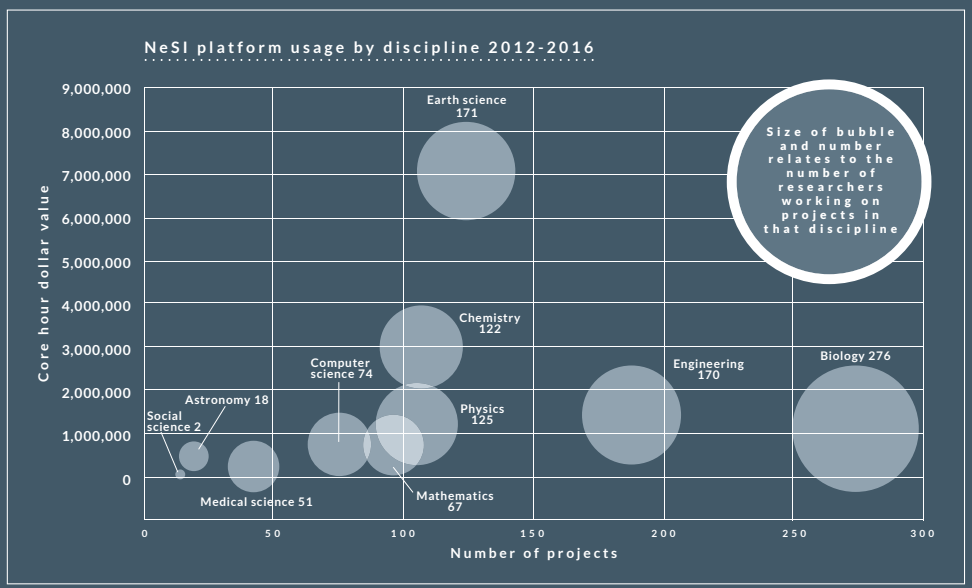
Calum Chamberlain



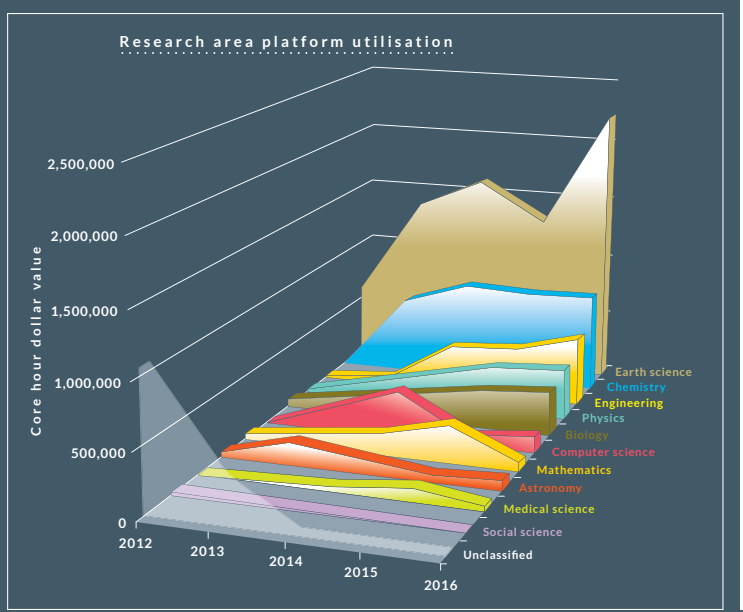


Embracing a diverse research sector

During 2016 NeSI supported a mix of research needs from engineering and the physical, biological, medical and social sciences. The needs of these communities are met through a variety of responses, with some of the largest users in the earth sciences requiring NeSI's highest performance capability HPC platform to meet their simulation and computational challenges. Meanwhile for some disciplines, such as biology, a broader base of smaller projects is typical, requiring NeSI's higher throughput capacity HPC platform.



Across research disciplines, NeSI is witnessing changes in the skills and needs of researchers over time, and responding in several ways. NeSI's training activities focused on lifting confidence to apply NeSI's services to research. NeSI's training strategy is rolling out successfully across the sector, with more and more training activities led by institutions. NeSI will review its training strategy early in 2017, to identify whether shifting focus onto advanced skills might delivery higher value to the sector. NeSI will extend application support to incorporate core skills in data science, to build user skills and extend support across a range of increasingly data-intensive analysis needs.



KPI 1

Annual case studies published aligned with Government priorities

CASE STUDIES

TARGET (Annual)
≥20

ACTUAL (Last 12 Months)
20 POSITIVE

Twenty case studies were published in 2016*, presenting work from all of NeSI's collaborator institutions, and including case studies highlighting research undertaken by researchers at the Dodd-Walls Centre, GNS Science, Massey University, Victoria University of Wellington, Scion, Plant & Food Research, MacDiarmid Institute, Riddet Institute, the Deep South Challenge and QuakeCoRE

1. *When getting back to basics is better than fancy new tools* ***2.*** *Retired but not forgotten: Algorithm research with a Supercomputer* ***3.*** *Molecular Fingerprinting* ***4.*** *Cellulose unchained* ***5.*** *New frontiers for milk proteins: functional foods and drug delivery systems* ***6.*** *Growing computational capacity among wet-lab scientists* ***7.*** *Finite element modelling of biological cells* ***8.*** *Modelling heterogeneous catalysis* ***9.*** *The University of Otago is at the technological frontier* ***10.*** *Helping communities anticipate flood events* ***11.*** *University of Otago researchers join the international high-speed data superhighway* ***12.*** *NeSI's supercomputer helps shed light on future climate* ***13.*** *Speeding up Basilisk with GPGPUs* ***14.*** *NeSI explores innovative uses of cloud services for advanced computing* ***15.*** *Terrestrial Data Analysis for the Ross Sea region* ***16.*** *Improving heart disease prediction and prognosis* ***17.*** *Golden mystery solved* ***18.*** *Otago physicists' prediction of gas "droplets"* ***19.*** *Fresh approaches for modelling geothermal* ***20.*** *UC scientists make biomolecular breakthrough*

**The case studies in full can be found in Appendix 1.*

Objective 2

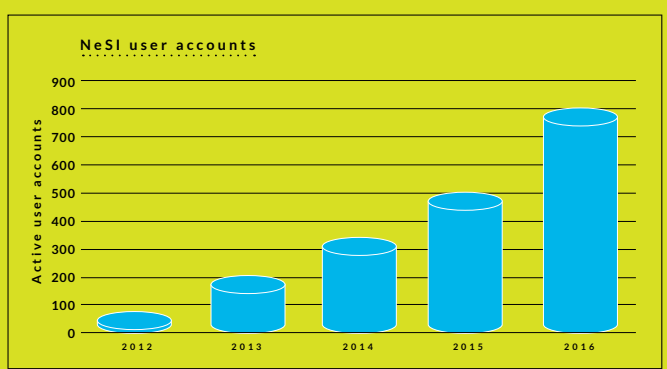
Grow advanced skills that can apply high-tech capabilities to challenging research questions

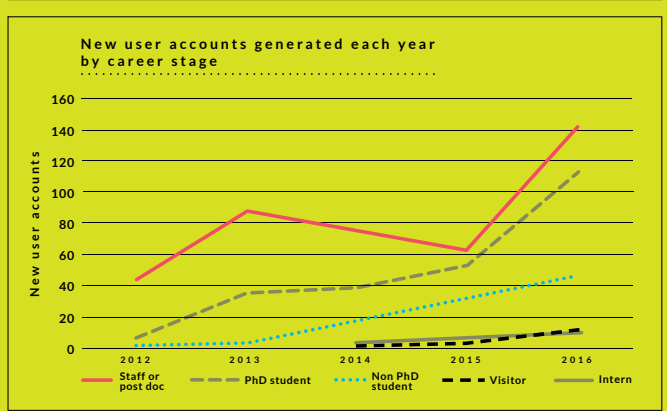
Sustained growth in users applying HPC to research

NeSI saw continued strong growth during 2016, reaching over 800 user accounts. As NeSI provisions accounts when approving a proposal for a research project, this is a direct indicator of the scale of user community NeSI has supported. Growth was sustained during 2016, confirming a continued recovery since the removal of Merit pricing in 2015.

Aleksandra Pawlik

Research Communities Manager





"The help of the NeSI support team is one of the best features of NeSI. Staff respond quickly and are very interested in providing the best possible resources. No improvements necessary."

Anna Garden

Growing researchers' digital computational literacy

Following on from starting a formal training programme in 2015, in 2016 NeSI invested in its leadership to increase impact on the sector. Aleksandra Pawlik took a role as Research Communities Manager in 2016. Aleksandra was previously Training Lead for a UK-wide programme run by the Software Sustainability Institute, and is one of a handful of Software Carpentry instructor trainers worldwide.

NeSI aims to build a sustainable skills base of advanced digital research capabilities, achieving impact across a broad range of disciplines by focusing on actions which enable research communities to be more self-sufficient. Across 2016 NeSI supported eight training events (including Victoria University of Wellington (VUW), Lincoln University, University of Otago, Plant & Food Research and Scion) with a 26% increase in the number of researchers signing up for Software carpentry training over 2015.

NeSI also trained 33 new instructors (16 certified) who will each go on to lead their own workshops. NeSI supported key community building events such as ResBaz (supported in three locations in 2016: University of Auckland, University of Otago and VUW), University of Auckland Digital Skills Winter Bootcamp, and the CRI coding Conference hosted by NIWA.

Partnerships with research collaborations and institutions

NeSI supported the 7th annual eResearch NZ meeting in February 2016, working with REANNZ and NZGL to broaden engagement on opportunities presented by advanced digital capabilities.

Throughout the year, NeSI focused support on specific national research collaborations, leading to future events now planned with the Bio-Protection CoRE at Lincoln ("Is HPC for me?") and with Te Punaha Matatini on PhD student Data Management training. Plant & Food Research proposed a joint workshop including contributions from NeSI; AgResearch invited NeSI to contribute to the planning of the symposium for the Our Land and Water Science Challenge; and the University of Otago ran a 'NeSI @ Otago' day, broadening connections into Otago research communities.

KPI 2

Annually, users report significant benefit to any project from services

59 users reported significant benefit when using the NeSI services and platforms during 2016.

DEFINITION

Change in core-hours consumed over the 12 months to date, compared with the same period 12 months earlier (expressed as a percentage).

TARGET (Annual)	ACTUAL (Last 12 Months)
>50	59 POSITIVE

A routine project closure survey has been up and running from Q3 2016. So far, we have sent out approximately 140 surveys and received 68 responses, a response rate of 49%. Of these responses, 59 researchers have agreed that their project received significant benefit from using NeSI (87%). Of the 9 responses that weren't positive, only one of these was due to a significant concern with service stability, which is being followed up on. The primary reason the other responses were not explicitly positive is because NeSI use was not a significant part of their project or they had not used NeSI in the last 12 months. We are continuously improving our survey to ensure we are asking the right researchers the right questions.

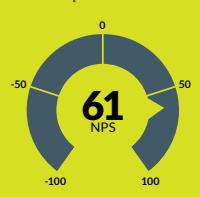
Researcher's experience of NeSI rated high performing

The Net Promoter Score (NPS) is an index (range -100 to 100) which measures customer experience through asking how likely a user would be to recommend NeSI to a friend or colleague.

Based on our NPS survey results to date, NeSI has an NPS of 61. This is a very high NPS 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.

Net Promotor score generated from responses to NeSI project survey in 2016:

How likely is it that you would recommend NeSI to a friend or colleague?



Cellulose unchained

Scion scientists can now generate hundreds of thousands of model cellulose molecules and their x-ray diffraction patterns to help reveal the structure of the most common organic polymer on the planet, thanks to NeSI's support and parallel processing opportunities.

Dr Stefan Hill is Research Leader for advanced chemical characterisation at Scion, the Crown Research Institute for forestry, wood and bioproducts, based in Rotorua. He is intensely interested in the structure of cellulose.

"Cellulose is the most ubiquitous and abundant polymer by weight on the entire planet, but its structure, in particular the number of chains that make up the material, is still unclear. If we can understand how cellulose is put together then we can better understand how to take it apart or modify it to take advantage of its amazing properties."

Direct interpretation of the complex X-ray diffraction patterns generated by cellulose is currently limited and based on a number of assumptions. The Scion team is taking an alternative approach of creating model diffraction patterns and comparing with them with experimental data.

"It's a simple idea," says Stefan, "but first we need vast numbers of input datasets, then we need to generate model X-ray diffraction patterns, and finally compare the model diffraction pattern with actual X-ray diffractograms."

The late Dr Roger Newman started working on this problem in the 1980s. He realised that Bragg's Law, which is normally used to predict diffraction patterns and assumes crystals are infinite in all directions, did not hold for long, nano-fibre-like crystals like cellulose. Dr Newman's solution was to write his own software to process datasets hand-entered into Excel spreadsheets. Generating one diffraction pattern took a day or more.

Stefan updated the program in 2015 and reduced running time to minutes, but was still left with the need to generate millions of input cellulose crystal structures.

"Coincidentally, around this time, I attended a presentation about NeSI's capabilities given by Ben Roberts. The access to supercomputing, the ability to run the model software in parallel on a number of processors seemed to be the perfect way to do one thing many, many times," says Stefan. "Ben was enthusiastic and encouraged us to have a go."

The team at NeSI modified Scion's software slightly and ran it successfully on NeSI's FitzRoy platform at NIWA, showing it was definitely possible to create vast numbers of model cellulose X-ray diffraction patterns in a short time.

The initial proof of concept used a limited number of datasets. A Monte Carlo approach that creates many datasets in a short time has replaced the old Excel method and now the Scion team is planning a run of tens or hundreds of thousands of structures, probably in the second half of 2016.

The larger test is the subject of a proposal for a development project undertaken by Stefan, with Ben taking care of many of administrative actions.

Stefan expects the larger scale testing will allow the team to get a handle on the final part of unravelling cellulose's structure – comparing model diffraction patterns with experimental data.

"Cellulose crystals do not exist alone. They are part of a matrix of hemicellulose, lignin and other components that make up trees and plants. This produces a background that complicates comparing models and actual diffraction patterns."

Removing the background from experimental data, or adding components of it to model diffraction patterns, (the approach that Stefan favours) followed by diffraction pattern comparisons, will also be a task for supercomputing.

NeSI's expertise and computing power have also opened up the possibilities of using new strategies to tackle the problem. One promising approach is the use of machine learning to remove humans (and their bias) from the generation of input datasets. With machine learning, the machine/computer generates datasets, compares model diffraction patterns with actual diffraction pattern and 'learns' which datasets are a good fit and which are not. The 'good' are then 'bred' and mutated to build up model structures that come closer and closer to resembling the actual structure.

"NeSI is an amazing resource for the New Zealand science community," says Stefan, "Not only is our dream to unchain the structure of cellulose within reach, it looks like we will be able to achieve it more smartly and quickly than ever."

"NeSI is an amazing resource for the New Zealand science community. Not only is our dream to unchain the structure of cellulose within reach, it looks like we will be able to achieve it more smartly and quickly than ever."

Objective 3

Increase fit-for-purpose use of national research infrastructure

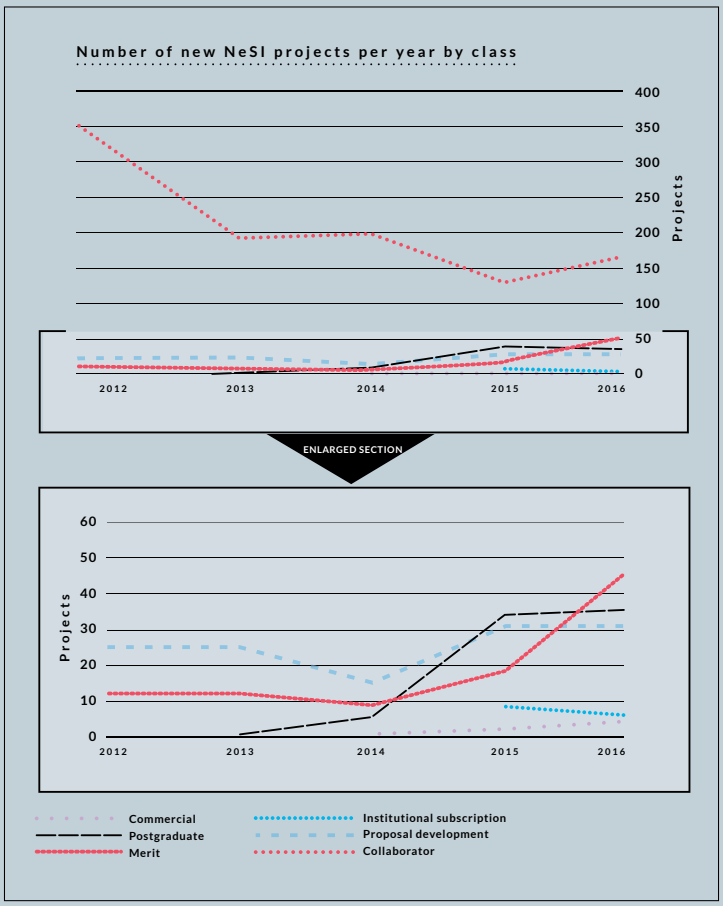
Michael Uddstrom
Platforms Manager

Incentivising efficient use of infrastructure investments
NeSI maintains an Access Policy aimed at improving utilisation of its platforms and increasing fit-for-purpose allocations. This policy was revised in 2015, removing pricing to researchers to ensure a broader uptake by the sector, and creating a subscription contracting model to shift this revenue from researchers to institutions, ensuring consistency of treatment across those researchers at institutions directly investing into NeSI, and those not. Since this change, project allocations and utilisation have recovered to earlier levels, and continue to show significant growth.

With the decommissioning of its HPC platforms, the University of Canterbury confirmed it would withdraw from being a collaborator in NeSI by the end of 2016. This has raised concerns about incentives within the current Access Policy, and whether these promote the best behaviours to support NeSI's sustainability. With these concerns in mind, the Access Policy will be reviewed early in 2017.

Growing partnerships to increase NeSI's impact
Outside of NeSI collaborators, six subscription agreements were signed in 2016 (Massey University, University of Waikato, Livestock Improvement Corporation, Department of Conservation, Caldera Health, Hawkes Bay Regional Council). NeSI is now working with each of these institutions to baseline their current needs, to deliver joint outreach and training, and inform institutional strategy.

Informing institutions of their researcher needs
Working in support of hundreds of research projects each year offers NeSI direct insights and information on researcher needs, which are of high value to research institutions. NeSI shares this information with institutions that contract with NeSI as part of a service governance relationship, informing them of the value of their contracts in terms of the research supported and outcomes achieved. During 2016 NeSI developed and stabilised quarterly service governance routines, ensuring consistency of information and insights across institutions and growing the collective ability to work together to grow use of and value from NeSI.



KPI 3

Annual change to utilisation of national platforms

NeSI realised 29% year-on-year growth in utilisation of its current platforms by the end of 2016. Triggered by the decommissioning of the BG/P platform at the end of June 2016, the scope of this measure was revised, now measuring the growth in utilisation of NeSI's consolidated infrastructure.

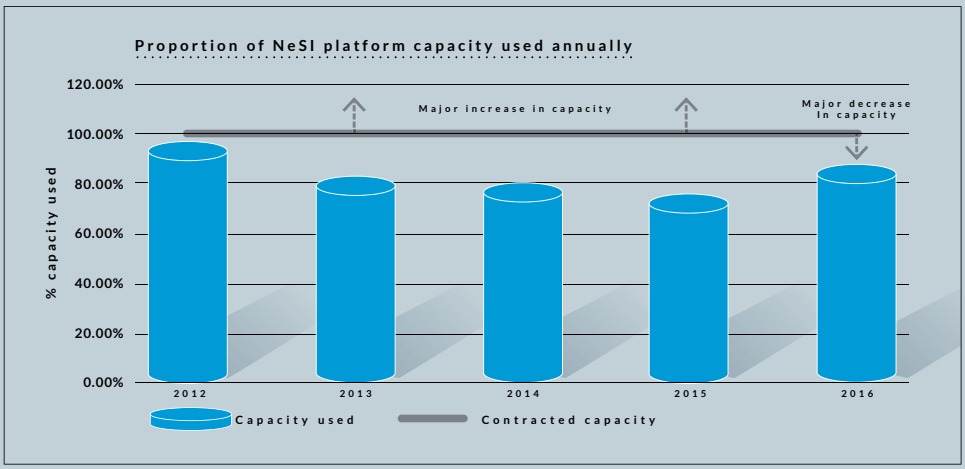
Against this reduced absolute capacity, utilisation of the remaining infrastructure increased significantly, in small part due to the migration of all users onto this smaller infrastructure footprint. However, the increase was primarily driven by a significant increase of usage on P6 (FitzRoy) by the Deep South National Science Challenge, and on x86 (Pan) from the growth by various institutions (University of Otago, Massey University, Victoria University of Wellington).

DEFINITION

Change in core hours consumed over the 12 months to date, compared with the same period 12 months earlier (expressed as a percentage).

TARGET (Annual)	ACTUAL (Last 12 Months)
+20%	+29%

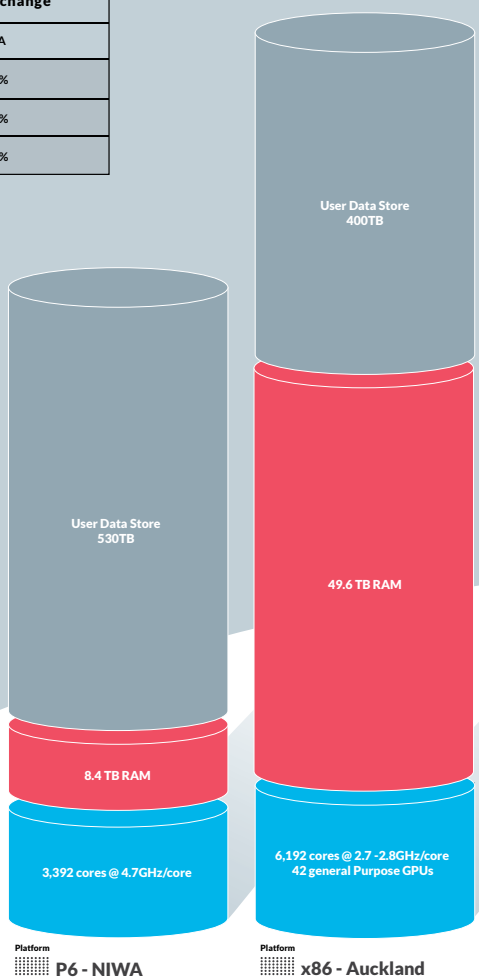
Reviewing both current and recently decommissioned assets shows slightly lower growth in utilisation across all platforms due to a year-on-year drop in utilisation of the decommissioned BG/P platform, which has varied in average utilisation between 80-90% over the last 18 months.
The year-on-year use of NeSI's contracted capacity (below) shows this slightly lower growth trend.



Platform	12 months ending on 31/12/2015	12 months ending on 31/12/2016	% change
BG/P	47,072,364	22,888,235	N/A
x86	32,038,105	36,511,012	14%
P6	8,159,749	15,348,268	88%
Aggregated	40,197,854	51,859,280	29%

Annual Utilisation by platform

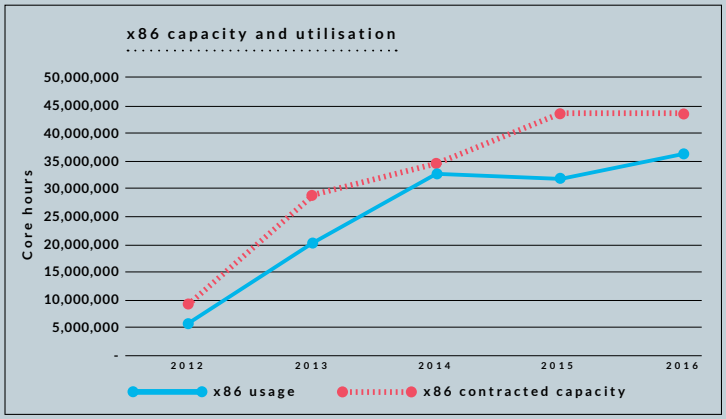
The table above compares utilisation in CPU core hours by platform year on year, year-to-date. The consolidated infrastructure incorporates the P6 and x86 platforms, with the decommissioned BG/P platform utilisation shown for completeness, though its utilisation figure in 2016 represents only six months operations prior to decommissioning.



Annual utilisation (core hours) by platform, 2012-2016

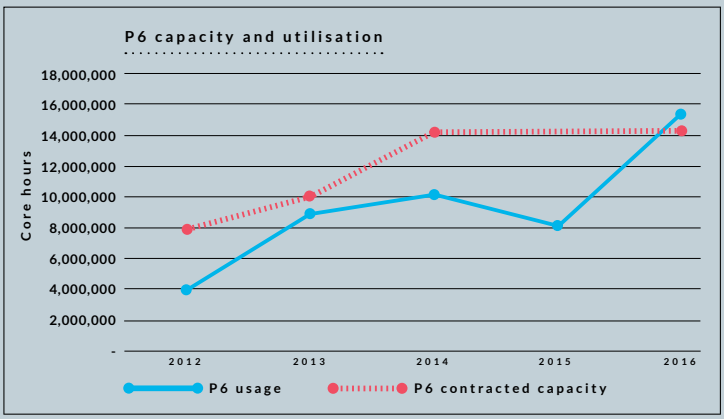
x86

This platform sustains near maximum utilisation over time, with some growth in 2016.



P6

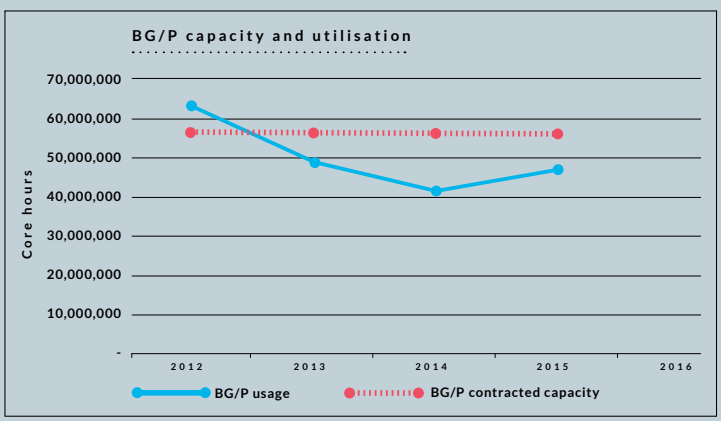
The primary users are running very large research jobs. As anticipated in 2015, 2016 saw a significant lift in utilisation based on new computational capabilities coming online.



Note: It is possible to exceed contracted capacity as NeSI has shared access to a larger machine.

BG/P

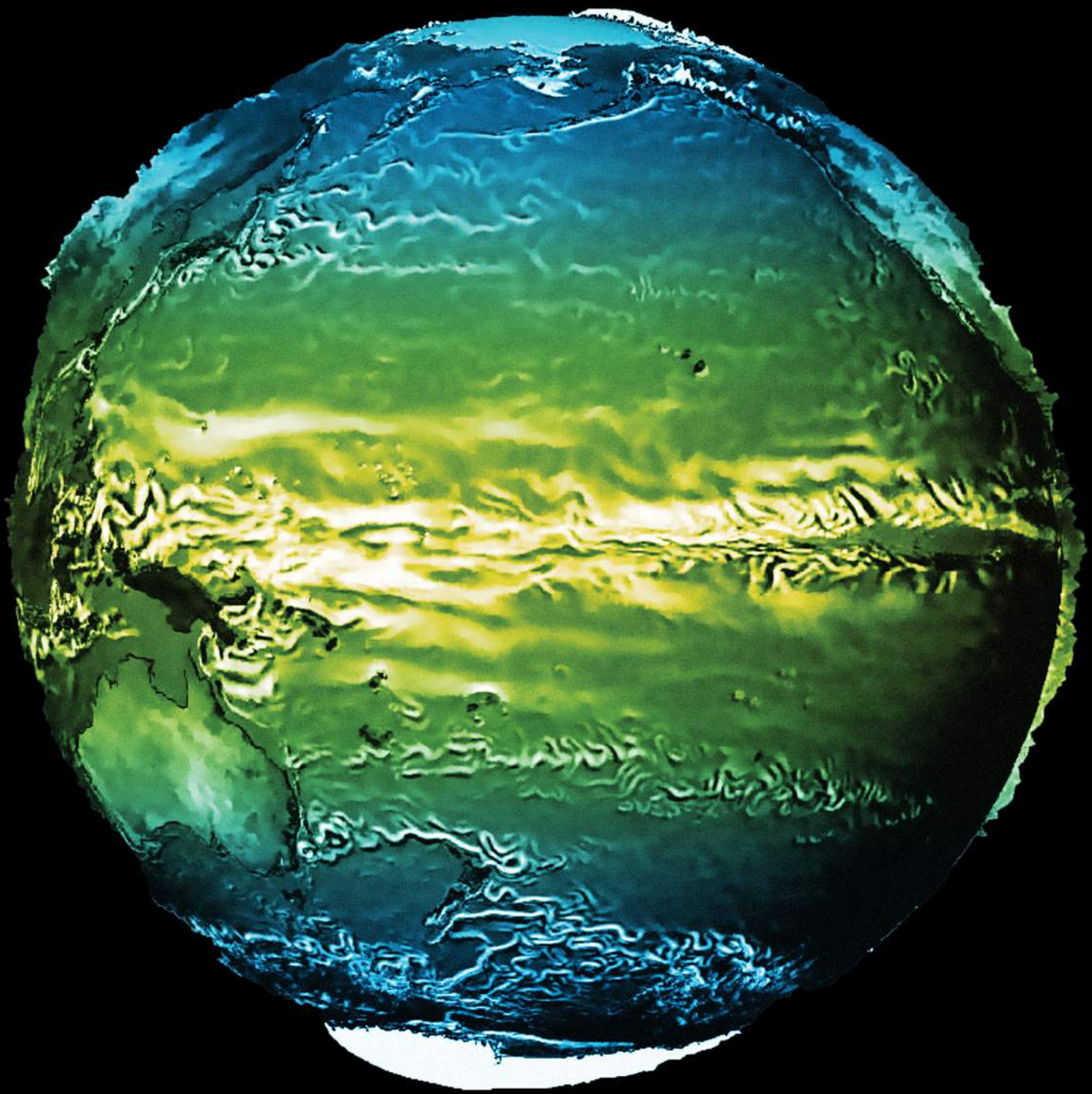
This platform reached end of life and was decommissioned in mid-2016.



Note: It is possible to exceed contracted capacity in cases of exceptionally efficient jobs.

NeSI's supercomputer helps shed light on future climate

Climate change is widely seen as a leading problem of our times. The Intergovernmental Panel on Climate Change (IPCC), in its 2013 assessment report, states that "human influence on the climate system is clear, and recent anthropogenic emissions of greenhouse gases are the highest in history. Recent climate changes have had widespread impacts on human and natural systems."



The impressive ability of climate models to capture many aspects of the climate system has made them the cornerstone of all IPCC assessment reports; they are widely used to quantify the human influence on climate. However, the Earth is a highly complex system, and there remain numerous challenges to improving climate models.

Three competing tensions act to increase the computational cost of running a climate model: Firstly, various processes are simplified or absent in models; addressing this requires us to increase the complexity of models. Secondly, the representation of climate usually improves with improved resolution, so there is a tendency to operate these models at as high a resolution as possible. Thirdly, climate is subject to an element of chaos. Therefore large ensembles of simulations are needed to accurately represent the most likely climate evolution and particularly extreme events such as severe storms, floods, or heat waves. For these three reasons, climate modellers often encounter limitations imposed by the computational resource at their disposal, and climate modelling can be the main motivation for upgrading ageing computing infrastructure.

The aforementioned IPCC report identifies several key southern high-latitude processes that are poorly represented in contemporary climate models, including clouds over the Southern Ocean, sea ice, and Southern-Ocean heat storage and transport. These deficits affect some aspects of climate projections available for the New Zealand region. In response, and in view of the substantial investment decisions that depend on accurate information on climate change, the New Zealand Government has launched the Deep South National Science Challenge, the objective of which is to "understand the role of the Antarctic and the Southern Ocean in determining our climate and future environment". To achieve this, a next-generation Earth System Model will be developed which, in addition to the physics of climate, also represents some chemical and biological feedbacks.

Developing such a model from scratch requires a large team; on the New Zealand scale this would be impossible to achieve. The Deep South therefore partners with a major international modelling centre, the UK Met Office, and focuses its efforts only on improving selected aspects of their model, relevant to the New Zealand region. The model development will benefit from observational campaigns to be launched under the Challenge and will ultimately lead to the development of the New Zealand Earth System Model (NZESM). Simulations produced using this model will then be used to inform high-resolution limited-area downscaling simulations similar to those that have been produced for several years using international climate model projections. Such high-resolution climate projections for the New Zealand region are then

"NeSI is providing the supercomputing infrastructure on which we are producing climate simulations, both globally and using a regional climate model. You cannot operate an Earth System Model without a supercomputer."

taken up by various adaptation projects covering sectors of the economy including agriculture, hydroelectricity, tourism, building and infrastructure, and insurance, to name a few.

The NZESM sits at the heart of the Deep South Challenge. The model couples atmosphere, ocean, and sea ice modules, which have all seen significant progress in recent years (fig. 1). The atmosphere component of the model comprises a whole-atmosphere chemistry package whose main aim is to adequately represent ozone and aerosol – both of which are important climate agents. Compared to the predecessor model used in the aforementioned IPCC report, the model top has been raised from 39 to 85 km. The ocean is represented by physics and marine biogeochemistry modules.

Sea ice only occupies the top metre or so of the ocean but completely changes the interactions between the atmosphere and the ocean if present. The Deep South aims to improve representations of surface wave propagation into the sea ice, which breaks up the ice into floes, and also the injection of freshwater from melting ice shelves which constitutes an important aspect of the hydrological cycle in Antarctica and affects sea ice formation close to the continent.

A single simulation of the NZESM will use roughly 1000 processors when complete, or nearly a third of the present NeSI supercomputer at NIWA, FitzRoy. According to Dr Olaf Morgenstern, part of the Science Leadership Team for the Deep South Challenge, "NeSI is providing the supercomputing infrastructure on which we are producing climate simulations, both globally and using a regional climate model. You cannot operate an Earth System Model without a supercomputer and we are looking forward to FitzRoy being replaced by a more powerful successor in 2017."

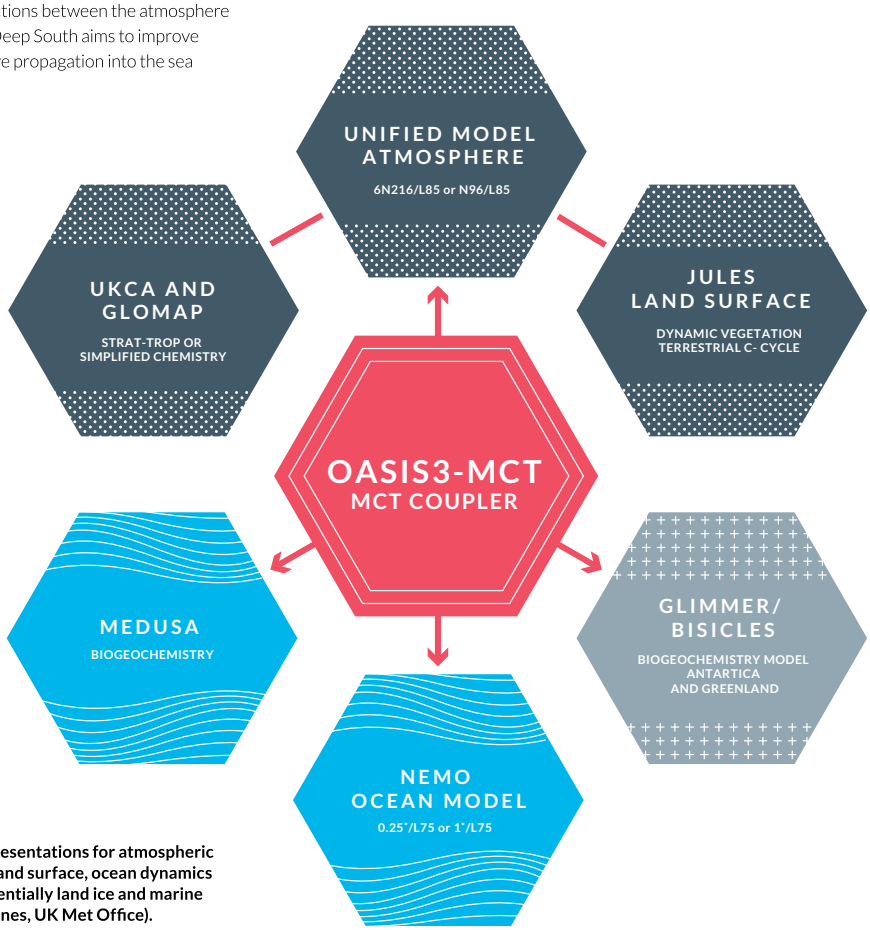


Figure 1 above comprises representations for atmospheric dynamics and chemistry, the land surface, ocean dynamics and biogeochemistry, and potentially land ice and marine ice shelves (Courtesy: Colin Jones, UK Met Office).

Objective 4

Make fit-for-purpose investments aligned with sector needs

Nick Jones
Director

Taking a national perspective - NeSI's National Platforms Framework

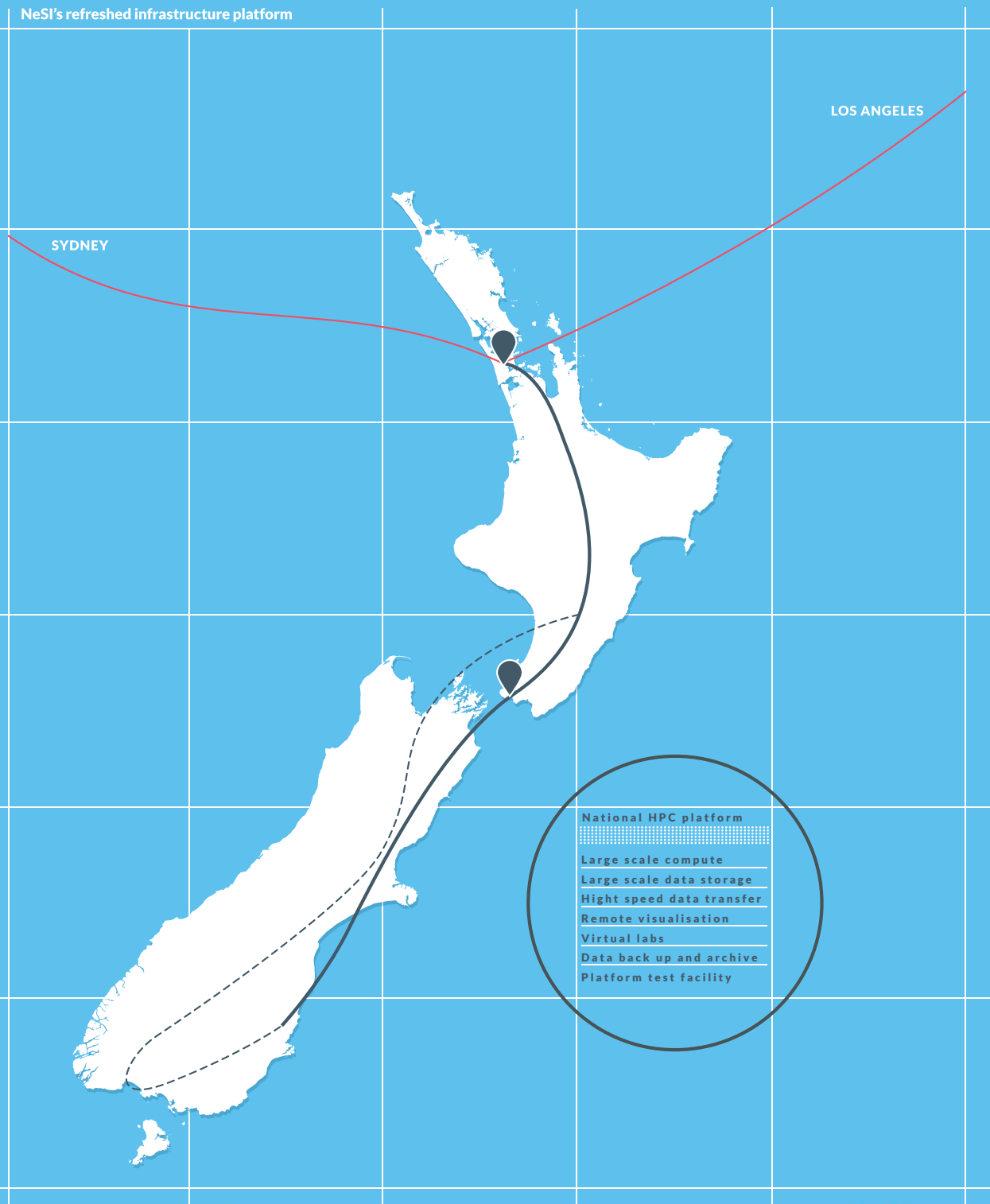
During 2016 NeSI ran the first investment case through its nationally focused and governance-led approach – the National Platforms Framework. Investing in infrastructure this way exploits the deep expertise of stakeholders, builds on the strength of sector institutions and the power of a coordinated approach, in order to realise the most effective investments possible. The process revealed key insights and presented important options for review. As an example, during the process it was identified and agreed that consolidation of the NeSI platforms to one primary site would improve the researcher experience, service capability, and return on investment.

Identifying future digital capabilities to power New Zealand research

Over 2015 and 2016, NeSI captured sector needs for future platforms through the NeSI 2 Business Case, the 2015 review of the National Platforms Framework, stakeholder engagement during design of the future platform, eResearch 2020's demand-driven foresight, and from sector requirements for genomics support through the NeSI - NZGL Alliance and the Advanced Genomics Platform discussions. These needs informed the consolidation of platforms mid-year, the initiation of procurement for new platforms, and investigations into cloud services.

NeSI's HPC platforms in 2016

Platform	Site	Specifications
BlueGene/P	University of Canterbury	8,192 cores @ 0.8 GHz/core with 8 TB RAM User data store: 182 TB
P575/POWER6	NIWA	3,392 cores @ 4.7 GHz/core with 8.4 TB RAM User data store:1856 TB
x86 Intel	University of Auckland	6,192 cores @ 2.7-2.8 GHz/core with 49.6 TB RAM 42 General Purpose GPUs User data store: 400 TB


NeSI's refreshed infrastructure platform
LOS ANGELES
SYDNEY
National HPC platform
Large scale compute
Large scale data storage
Hight speed data transfer
Remote visualisation
Virtual labs
Data back up and archive
Platform test facility

"Without NeSI this research would simply not be possible. For this research I strongly rely on HPC facilities."

Lukas Hammerschmidt

Consolidating infrastructure paves way for future

Decommissioning of UC's HPC platforms was completed mid-year, and the data storage platforms were decommissioned in October 2016. NeSI gained valuable experience through smoothly transitioning a large number of researchers onto the remaining NeSI platforms, building confidence in how to approach user support during the infrastructure refresh later in 2017. The final location of NeSI's remaining two facilities also lays the shape of future investments, establishing a primary and a disaster recovery data centre locations to host NeSI's distributed infrastructure.

Delivering a refreshed infrastructure platform

Early in the year, NeSI started preparations to replace both the NIWA IBM P575/POWER6 (FitzRoy) and Auckland IBM iDataPlex x86 (Pan) platforms, selecting a single procurement process with NIWA as the lead agent working alongside NeSI's other collaborators. The procurement is running through to the end of Q2 2017, while implementation is planned to be completed in early 2018. The project has reached all key milestones during 2016 including running a sector wide solution design process, building a robust business case aligned with key government strategies, confirming Investor commitments to the proposed investment, and delivering an RFP to the market in Q4 2016.

Designing a platform for known unknowns

NeSI's collaborative design of a refreshed infrastructure platform identified key gaps in the current platform offering, in particular when considering future genomics and space science Big Data requirements. Specifically, support for: end-to-end workflows across institutional boundaries; high performance interactive data analyses and visualisation; linking of research projects into a coordinated research programme working on shared data. These capabilities have existed at NeSI's collaborator sites, though haven't been available nationally. NeSI's refreshed infrastructure will deliver these capabilities into the national platform, offering new opportunities to further integrate research workflows, data, and to better support sector-wide research collaboration.

Innovation in infrastructure access

Alongside the consolidation and refresh of NeSI's platforms, the NeSI Business Case prioritised ongoing development of NeSI's cloud strategy, focusing on support for highly scalable research computing and access to emerging technologies. In 2016 NeSI investigated the viability of bursting seamlessly from the NeSI x86 (Pan) platform onto two public clouds, to run a number of less demanding research workloads. NeSI successfully proved the viability of seamlessly integrating an HPC environment to public clouds while delivering scientifically equivalent outputs. NeSI now has capability to work with the cloud and is focusing on implementation of an operational cloud environment to run actual research cloud-based workloads in 2017.

Aligning research and research infrastructure priorities

NeSI and NZGL established an alliance in 2015, focused on identifying common needs of Genomics users across the sector and to explore alignment of capacity building, service delivery, and infrastructure consolidation. With the shift in the sector to a refreshed Advanced Genomics Platform through an MBIE RFP late in 2016, NeSI provided independent support for several proposals, aiming to align infrastructure capabilities and services with the needs identified in any successful proposal. Similarly, as MBIE's Regional Research Institute investment into the Centre for Space Science Technologies (CSST) was confirmed, NeSI established a working relationship with the institute to ensure further alignment and fit-for-purpose design of national research infrastructure.

Meeting pressing needs for better support of research data

Continuing the initial work done in 2015 within eResearch 2020, NeSI supported release of a strategic and economic case for investment into research data – "The Case for Research Data" – the next phase of the National Research Data Programme. This programme identifies opportunities for action at a national level that will yield the greatest impact and benefit for New Zealand researchers with regards to research data.

An exemplar of the potential that coordinated action can have in addressing research data issues is provided by two NeSI case studies were published in 2016. These highlight NeSI's proactive work to better support Big Data workflows – outcomes of joint work between the University of Otago, REANNZ and NeSI. As part of the eResearch 2020 initiative three more exemplar projects were initiated in the second half of 2016. These are proof-of-concept studies to illustrate the opportunities to better align research data strategy and services in New Zealand. The studies are continuing into early 2017 and will be presented at the eResearch NZ Conference in February 2017.

Growing capabilities to meet the data challenge

During review of the National Platforms Framework, NeSI identified a need to build out its data science and analytics capabilities, in order to drive aligned development of services and infrastructure. During 2016 NeSI engaged with collaborators on how best to meet these needs, and agreed to invest in a new position based at the University of Otago, to start mid-2017. Bringing these skills and perspective into NeSI ensures future directions on the challenges of data-intensive research are well informed.

KPI 4

Annual change to utilisation of national platforms

This KPI is measured by a survey which asked users to rate their satisfaction with NeSI services on a scale from 1 to 5, 1 being 'strongly disagree' and 5 being 'strongly agree'.

DEFINITION

Percentage of users in the NeSI annual survey who agree that NeSI services are meeting their needs.

TARGET (Annual)	ACTUAL
>80%	92%
USER SATISFACTION	USER SATISFACTION

From project closure surveys sent out for 2016, we received 72 responses to a question asking if users agree that NeSI services are meeting their needs. Of these responses, 66 researchers have agreed that NeSI services do meet their needs. The other six responses were all neutral rather than negative, with some good constructive feedback on possible areas for improvement, such as more user friendly access and improved cluster stability.

"This project is long running, and without access to NeSI it would not be possible to complete the computer based experiments in a suitable time, or cope with the throughput we wish to achieve."

Jack Flanagan

Objective 5

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

Brian Corrie
Solutions Manager

Taking next steps on national collaboration

A key goal of NeSI's current business case is to shift from a distributed and arms-length collaboration to a unitary virtual organisation. Ensuring the right structure is essential to enabling strategy and NeSI is positioned to make the most of a collaborative and embedded approach. Having team members located within sector institutions builds the direct relationships and provides invaluable insights into their needs. NeSI's structure is built along functional reporting lines nationally, ensuring single clear reporting lines and enabling NeSI to evolve a national culture of service and high performance. While any single institution may achieve great things in some aspects of NeSI's activities, together NeSI brings out the best.

NeSI refers to this shift as being one from "local to national", with focus placed on the areas where collaborative culture can drive the best outcomes, services can be optimised for increased impact, and to build in flexibility to support future growth and development.

Managing change with confidence

Establishing advanced digital capabilities at the leading edge of high performance information technologies requires passion and care in abundance. During 2015 NeSI adopted portfolio management practices across its projects, which were bedded down in 2016. NeSI's tracking is now standard practice and builds on the new functionally organised team structures to provide visibility, and supports national management over resource pools. Achieving this national team-oriented programme practice including direct operational management across institutions is a first for the research sector, and lays foundations which are adaptable to meet future needs.

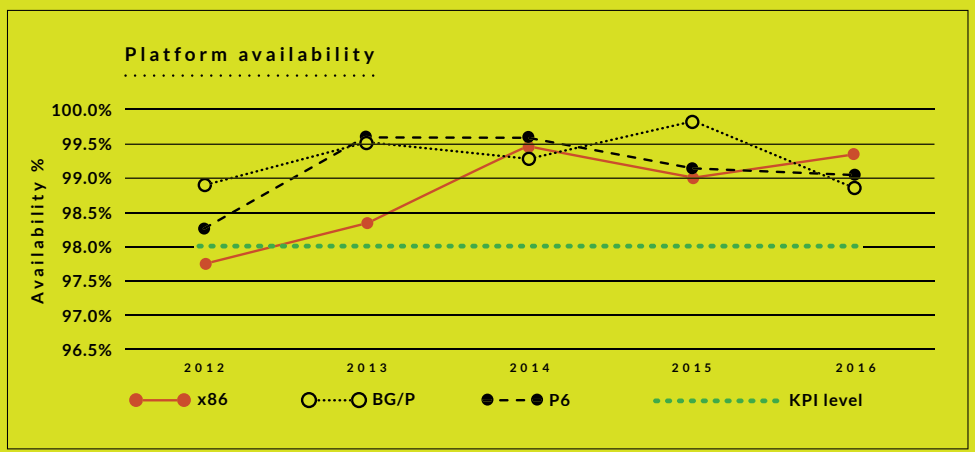
A culture of responsive end-user focus and service

Central to NeSI's high performance service reputation is the expertise of the team and the channelling of this expertise into service delivery. Two key areas underpinning NeSI's services are Applications Support and Computational Science. NeSI initiated reviews of both Applications Support and Computational Science functions during 2016, focusing on how they support NeSI's HPC Compute and Analytics and its Consulting services respectively. These reviews aim to standardise practices in both areas, recognising previous practices focused on distinct value propositions, around two quite specific products – NeSI's highest performance capability HPC platform and its more broadly applied capacity HPC platform. These projects lay the groundwork for NeSI's refreshed infrastructure, where the two platforms will be integrated around a common high performance data infrastructure, driving opportunities for greater efficiency as NeSI scales to support the anticipated growth in users and projects that will follow.

Improving performance of data-intensive workflows

NeSI now has the foundations of a national high-speed data transfer service operating across all NeSI sites, enhancing service delivery consistency, improving time-to-solution and making it easier to support computational workflows across NeSI platforms. These foundations will be built on in 2017, through partnerships with institutional Information Technology Service providers and REANNZ on last mile connectivity to research data stores and data-intensive instruments and equipment.

Availability of NeSI platforms, percentage of operational time that the platforms are able to be used:



Fabrice Cantos
Site Manager - NIWA

Nick Spencer
Site Manager - Landcare Research

Marcus Gustafsson
Site Manager - University of Auckland

David Maclaurin
Site Manager - University of Otago

KPI 5

Annual availability of services

The annual KPI of availability measures service delivery consistency and performance. In 2015 NeSI achieved more than 98% availability.

DEFINITION

Hours annually for which the Platform is able to be used to provide the operational service, divided by total hours in a year, less scheduled downtime, expressed as a percentage (The measurement period for this KPI is 3 months).

TARGET (Annual)
≥98%
AVAILABILITY

ACTUAL (YTD)
99.3%
AVAILABILITY

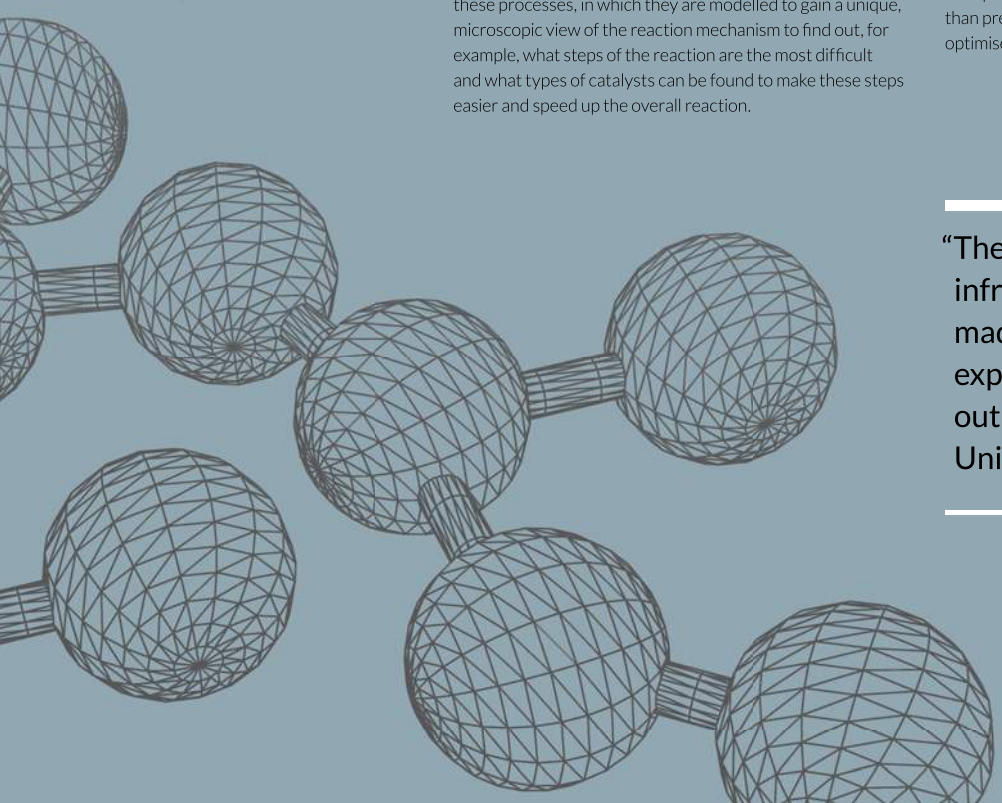
NeSI has managed its platforms to deliver a very high level of availability to researchers in 2016 despite the ageing nature of the current platforms. During 2016, NeSI has increased storage capacity, worked on file system stability and replaced parts of the interconnect hardware to maintain performance levels and will continue a programme of active maintenance through to replacement of platforms late in 2017.

Modelling heterogeneous catalysis

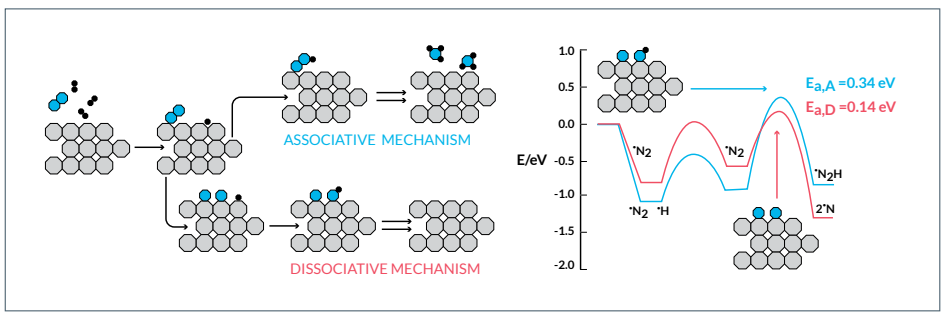
Dr Anna Garden and her research group are based in the Department of Chemistry at the University of Otago. Anna leads a highly successful group of young researchers whose research requires use of large-scale computational resources, making access to NeSI's national services very useful. NeSI's Solutions and Applications team has been supporting Anna and her group throughout the research process.

Anna's research interests are specifically in heterogeneous catalysis, which is a process of catalysis where the catalyst is in one phase (typically a metal surface or nanoparticle) and the reactants and products are in another phase (typically gas/liquid). Anna's group takes a computational approach to these processes, in which they are modelled to gain a unique, microscopic view of the reaction mechanism to find out, for example, what steps of the reaction are the most difficult and what types of catalysts can be found to make these steps easier and speed up the overall reaction.

As Anna explains: "One famous example on which I work is ammonia synthesis, in which ammonia (used for fertiliser) is formed by gaseous nitrogen and hydrogen reaction on a metal catalyst surface. This is what is used in the Haber-Bosch process, which produces 140 million tonnes of ammonia per year. HB process has been studied a lot experimentally but the high temp, high pressure conditions of the HB process mean that information that can be gained experimentally is limited. By simulating the reaction, however, precise information can be gained at any temp/pressure. We have recently used this approach to study different mechanisms of ammonia formation on different structures of catalysts and find that, on nanoparticle catalysts, the mechanism could be quite different than previously thought, throwing the door wide open to optimise the process further."



"The successful utilisation of NeSI infrastructure for her research made Dr Anna Garden a local expert, fully capable of helping out experimental chemists at the University of Otago."



Calculated energy profiles for two possible mechanisms of ammonia formation on a stepped ruthenium catalyst.

The outcomes of this research have been reported in several publications and the internationally recognised work done with a number of collaborators will now flow outwards from the Dunedin campus. Most members of Anna's team hold majors in at least one other discipline apart from Chemistry, which allows them to expand and build their expertise. In order to progress with their research they needed to develop their computational skills and become fluent in High Performance Computing tools and environment. They focused on this task with the dedicated NeSI support.

Knowledge transfer through NeSI support

NeSI's Solutions and Applications team has been supporting Dr Anna Garden and her team in a number of aspects of their computational work. As Anna reports: "Our main software was already installed on NeSI when we began using it, and the support team were able to quickly customise a version for our specific needs, which is built with state-of-the-art optimisers and libraries, useful for surface science."

The close collaboration between NeSI staff and the Computational Chemistry group at Otago brought a lot of benefits in terms of development of transferable computational skills. It was particularly important in the case of incoming postgraduate students.

"Support to the junior students is really useful, both with my group, who have some Unix experience and with my experimental collaborators, whose students are arriving at NeSI having never worked in a Unix environment before. With my group, NeSI's support team has offered more technical support, such as helping to optimise memory usage of large jobs. With the experimental students, they have offered quick and friendly help with getting started in an HPC environment."

Sharing the expertise

The successful utilisation of NeSI infrastructure for her research made Dr Anna Garden a local expert, fully capable of helping out experimental chemists at the University of Otago:

"I also perform support calculations for experimental colleagues on problems ranging from designing optimal prodrugs for chemotherapy to understanding unexpected supramolecular structures."

The outcomes of this collaborative work have recently been published (Y. Abghoui, A. L. Garden, J. G. Howalt, T. Vegge, E. Skúlason, "Electroreduction of N2 to ammonia at ambient conditions on mononitrides of Zr, Nb, Cr and V: A DFT guide for experiments", ACS Catal., 6, 635-646 (2016)).

$300,000 for Fast-Start

The importance and high profile of Anna's research has been widely recognised, and she was recently awarded a Marsden-funded Fast-Start grant of $300,000. This grant will allow Dr Garden to focus on finding solutions to help mitigate the growing pollution of New Zealand's waterways with nitrate, an issue which arises mainly due to agricultural intensification and overuse of fertilisers. The grant will allow Anna to expand her international collaboration, working further with Professor Egill Skúlason from the University of Iceland with whom she has already published a number of high-impact papers.

NeSI is thrilled to support research which not only makes significant contributions to the field of computational chemistry, but also provides solutions for tackling New Zealand's environmental challenges.

Objective 6

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

Robin Bensley
Operations Manager

Our collaborators' financial commitment, as a ratio of Crown contributions, is a financial indicator of the health of the NeSI partnership.

KPI 6

Contract to Date, Ratio of Collaborator Commitments to Crown Contribution

DEFINITION

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

LTD (BUDGET)	LTD (ACTUAL)
95.6%	74.9%

As at December 31, the ratio of collaborator commitments to Crown contributions is 74.9% LTD (life to date). The target financial contribution by contract end is 95.6%.

The significant variance from budget is due to a shift in timing of capital equipment replacement funded from collaborators' reinvestment obligation. The shift in timing aligned all collaborators' capital investments within one procurement, which will be completed in mid-2017. With commissioning of equipment running through to early 2018, this variance will be resolved early in 2018.

Mike Ladd
Strategic Projects Manager

Kirsten Brown
Operations Coordinator

Laura Casimiro
Operations Administrator

A national organisation

NeSI is a collaboration of four institutions coordinating investments in partnership with the Crown. This collaboration is constituted through a legal agreement between the University of Auckland, the National Institute of Water & Atmospheric Research (NIWA), the University of Otago and Landcare Research, with the University of Auckland as the legal contracting entity with the Crown, commonly referred to as the Host. A set of operating principles defined in the Collaboration and Services Agreements forms the basis for governance, management and operations.

Governance
NeSI's Board is responsible for strategy development, policy, approving major initiatives and investments, as well as monitoring the NeSI risk register. The four collaborator institutions appoint three Directors, alongside an independent Chair and another independent member with expertise in the field representing the research sector at large. All Board members are focused on the interests of NeSI, being the effective delivery of national research infrastructure services.

In 2016 the NeSI Board was comprised of the following members:

- Rick Christie, Independent Chair
- Prof. Andrew Rohl, Independent, Professor of Computational Science, Curtin University
- Dr Murray Poulter, former Chief Scientist, NIWA
- Stephen Whiteside, Chief Digital Officer, University of Auckland
- Prof. Rob Ballagh, Department of Physics, University of Otago
- Prof. Steve Weaver, Deputy Vice-Chancellor (Research), University of Canterbury (retired May 31 2016)

The NeSI Research Reference Group, established in late 2015, is an advisory group made up of eight research community leaders with strong technical knowledge on the impact of eScience as an enabler to 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, Senior Lecturer (Physical & Theoretical Chemistry), University of Waikato
- Dr Nauman Maqbool, Group Leader Knowledge & Analytics, AgResearch
- Prof. Cristin Print, Molecular Medicine & Pathology, the University of Auckland
- Prof. Ian Foster, Department of Computer Science, University of Chicago
- Dr Sam Dean, Chief Scientist - Climate, Atmosphere and Hazards, NIWA
- Prof. Blair Blakie, Department of Physics, University of Otago
- Prof. Barbara Chapman, Department of Computer Science, University of Houston and Professor, Institute for Computational Science, Stony Brook University
- Dr Susan Wells, Associate Professor, Epidemiology and Biostatistics, the University of Auckland (retired from the group December 31 2016)

Rick Christie
Chair, Independent Director

Prof. Andrew Rohl
Independent Director and Professor of Computational Science Curtin University

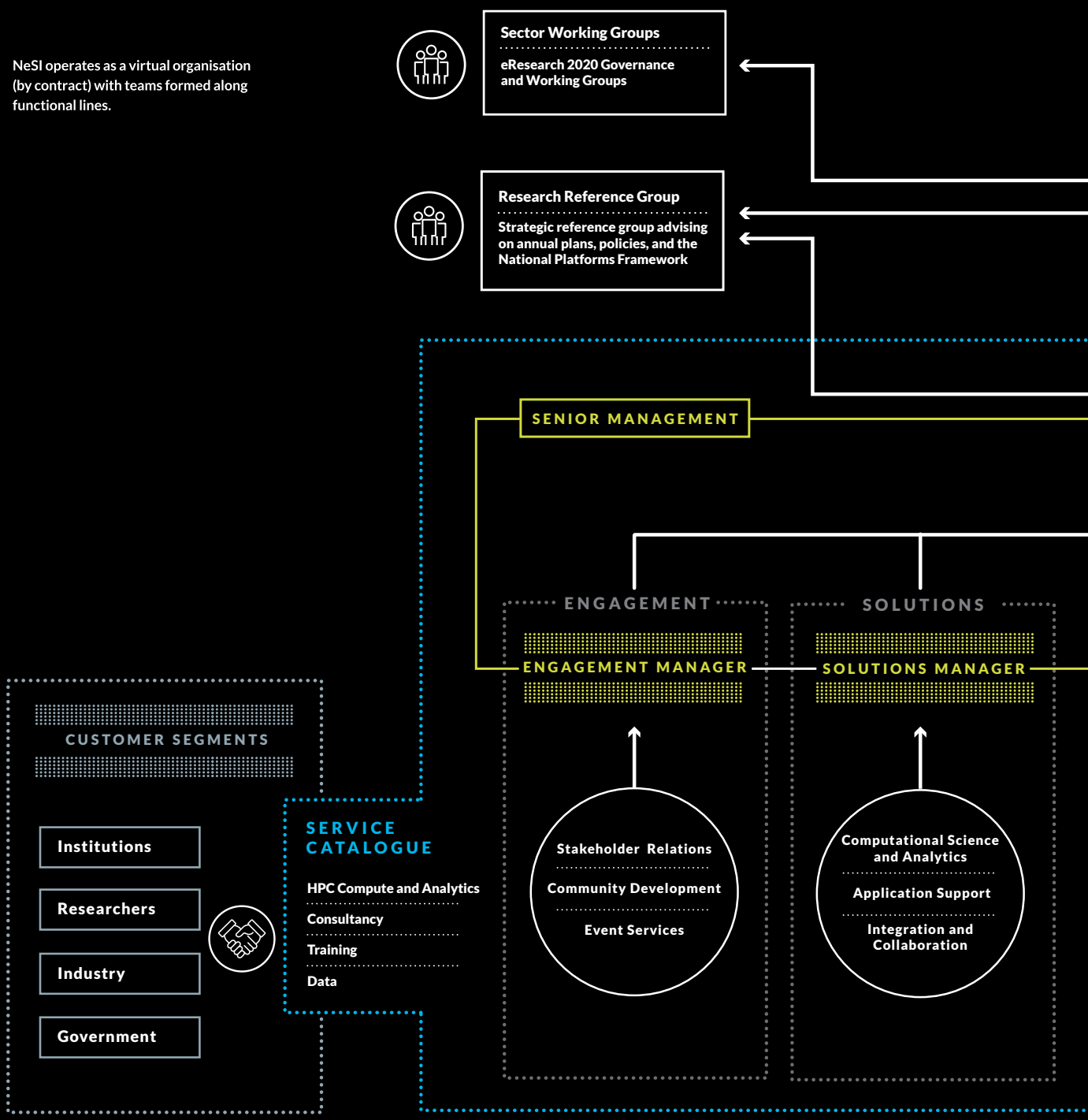
Dr Murray Poulter
Chief Scientist, Atmosphere, Hazards and Energy NIWA (Retired)

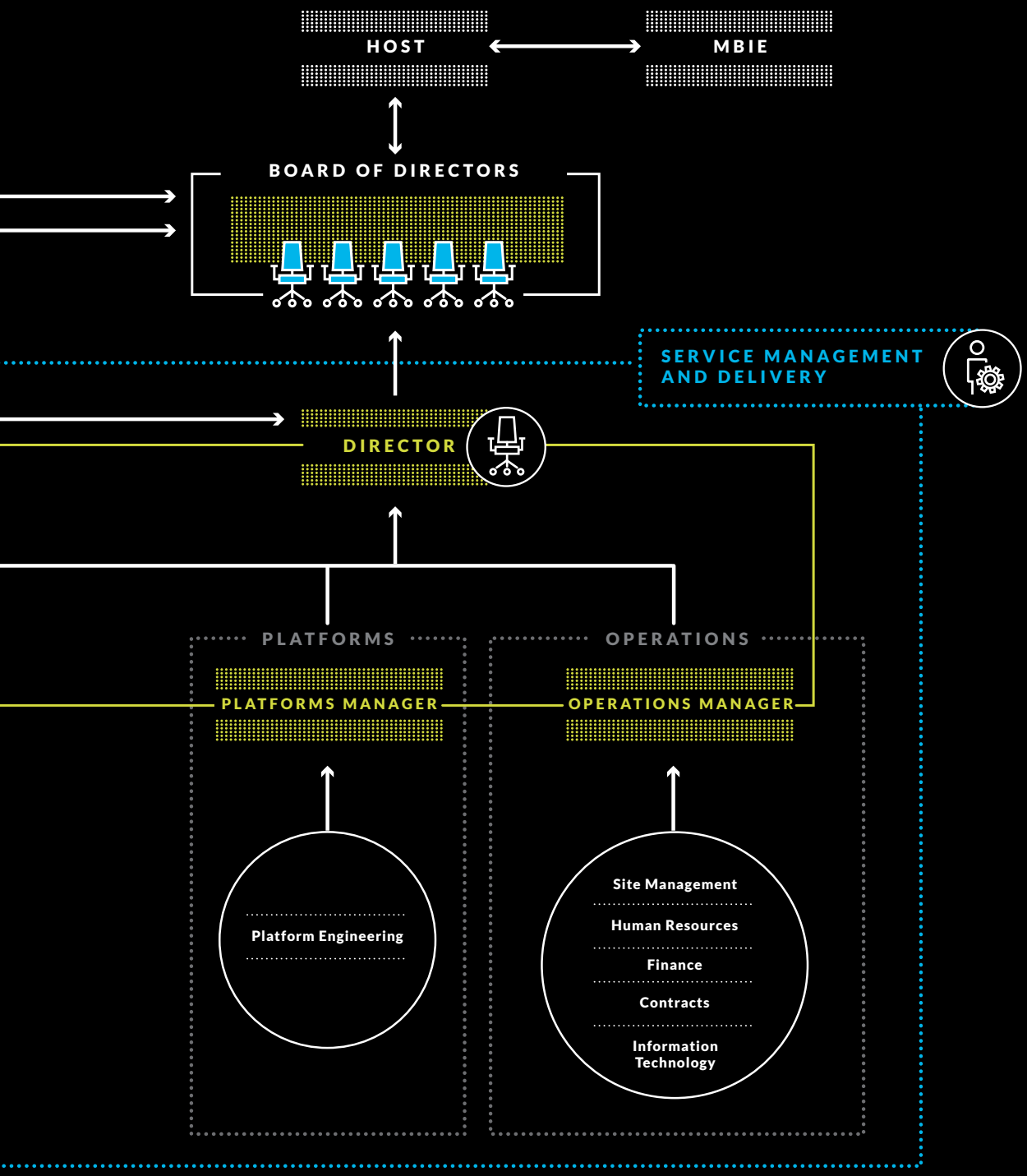
Stephen Whiteside
Chief Digital Officer University of Auckland

Prof. Rob Ballagh
Department of Physics University of Otago

ORGANISATIONAL DESIGN

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




HOST
MBIE
BOARD OF DIRECTORS
SERVICE MANAGEMENT AND DELIVERY
DIRECTOR
PLATFORMS
OPERATIONS
PLATFORMS MANAGER
OPERATIONS MANAGER
Platform Engineering
Site Management
Human Resources
Finance
Contracts
Information Technology

Recommendations and updates to Crown Funding Agreement

Funding Agreement

1. With regard to NeSI's contract budgets, it is recommended these are revised and a new baseline is established to align with NeSI's forecasts, against which future financial and KPI performance would then be measured.

2. With reference to the NeSI Objectives, there are two sets of related objectives – the NeSI Objectives specified within the Crown Funding Agreement, and those which NeSI uses within its KPI framework (also specified within the Crown Funding Agreement). Recommend reviewing whether maintaining both sets of objectives is appropriate.

KPI updates and recommendations

There was a recommended and agreed amendment to KPI 3 in 2016. This KPI measures Objective 3 and is calculated as:

Annual change to utilisation of national platforms +20%

In line with NeSI's National Platforms Framework, any commissioning or decommissioning of platforms should lead to a subsequent revision to KPIs, to reflect any relevant impacts on NeSI's objectives. It is necessary to reflect the change to NeSI's national platforms, specifically addressing the decommissioning of the University of Canterbury BlueGene/P platform during Q2 2016.

Previous calculation scope

The previous calculation of platform utilisation was applied to the scope of operational NeSI platforms, which included all three national platforms, being x86 (Pan), POWER6 (FitzRoy), BlueGene/P (Foster). As the scale of NeSI's national research infrastructure has now reduced, platform capacity has decreased and utilisation has consolidated onto two remaining platforms. We can no longer compare utilisation of the Canterbury platform year-on-year as it was not operational for the full year in 2016.

Proposed calculation scope from Q3 2016

From and including Q3 2016, the scope of platform utilisation to be measured is proposed to reduce down to the two remaining operational platforms, being x86 (Pan) and POWER6 (FitzRoy). The measure will reflect the consolidation of utilisation onto NeSI's remaining national platforms, and demonstrate a resultant lift in utilisation and asset efficiency.

Crown funding

This report has been prepared to meet contractual annual reporting requirements per the Crown Funding Agreement, to show income and expenditure for the NeSI Funding over the calendar year, and the closing cash position. NeSI Funding means all Crown contributions, other than the Host Fee.

	2015	2016
NeSI Funding received[1]		
Collaborator Crown funding	5,434,645	4,435,933
Total NeSI Funding	**5,434,645**	**4,435,933**
Amounts paid/(owing) to Collaborators		
The University of Auckland	2,156,988	1,553,224
The University of Canterbury[3]	1,126,219	457,847
National Institute of Water and Atmospheric Research	1,865,564	1,512,312
Landcare Research	285,195	182,022
The University of Otago	674,322	334,818
Total paid/(owing) to collaborators	**6,108,288**	**4,040,223**
Net year end position	**(673,643)**	**395,710**
Reconciliation to year end cash balance		
Add/(deduct) timing differences		
Add expenses incurred in prior period but paid this period[4]	1,196,641	
Deduct expenses incurred this period but paid in next period[5]	(1,196,641)	(418,924)
Cash flow movement for period after adjustment of timing differences	522,998	(382,007)
Add/(deduct) special items		
Repayment of Crown Funding for Depreciation during NeSI 1, by the University of Canterbury[5]	3,064,801	-
Cash flow movement for period	522,998	2,682,794
Opening Cash balance - Crown Funding	1,359,304	1,882,302
Cash movement	522,998	2,682,794
Closing cash position - Crown Funding[5]	1,882,302	4,565,096

1. Crown Funding received from MBIE per Crown Funding Agreement

2. Includes all collaborator payments incurred for the year.

3. Q3 & Q4 schedule payments ($270,780 in total) are held by Host without paying due to the withdrawal since September 2016. Decision will be made on either utilising the funding within NeSI 2 or returning back to MBIE.

4. 2015 period collaborator payment invoices paid in 2016.

5. The closing cash balance of $4,565,096 is mainly contributed to by the special item "Repayment of Crown Funding for Depreciation during NeSI 1, by the University of Canterbury" together with a timing difference of 2016 payments made in 2017. In January 2017, NeSI agreed with MBIE to return these funds back to the MBIE SSIF appropriation.

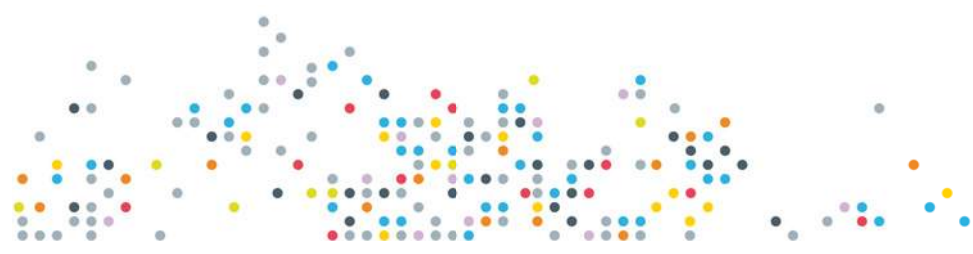
All figures are GST exclusive

NeSI expenditure

This report summarises total NeSI expenditure in the previous and current years, noting the sources of funding and the lines of expenditure, and providing an outlook for the end of NeSI's contracts with the Crown and collaborators through to the end of June 2018.

	2015	2016	*Forecast to June* 2018
Contribution from			
Crown Funding	6,265,261	5,632,525	21,839,998
Collaborator new investment	4,609,047	4,198,501	14,949,040
Collaborator Reinvestment	80,000	-38,405	5,778,999
Platform Access Fund (Co-funded Crown & collaborator)	0	0	10,580,842
Sector Revenue	34,069	15,213	515,884
Total Contribution to NeSI	**10,988,377**	**9,807,834**	**53,664,763**

Expenditure			
People	5,792,388	6,280,876	24,871,376
Operating	771,365	542,649	2,342,594
Platform	2,117,131	1,755,481	5,053,837
Depreciation	2,253,086	1,302,822	4,521,233
Capex	80,000	-38,405	16,875,725
Total Expenditure	**11,013,969**	**9,843,423**	**53,664,765**



Synopses of case studies

Otago physicists' prediction of gas 'droplets' confirmed

Ground-breaking theoretical work by University of Otago physics researchers showing that under certain conditions gases can form into stable droplets – as liquids do – has now been confirmed experimentally by scientists in Germany.

It has been a longstanding question whether it is possible to make a gas self-bind in the same way that water molecules coalesce into stable droplets, says University of Otago researcher Professor Blair Blakie.

Prof. Blakie's team in the Department of Physics determined that an extremely cold gas of highly magnetic atoms will self-bind into gas droplets that stabilise themselves, even in vacuum.

Their predictions were made using quantum calculations on high performance computing facilities provided through New Zealand eScience Infrastructure (NeSI).

The gas needs to be prepared under very specific conditions to observe the formation of droplets, he says.

"We worked out that it had to be at a temperature of a few millionths of a degree above absolute zero, a density more than a billion times lower than liquid water (about 100,000 times lower than air), and in a suitably adjusted magnetic field."

Under these conditions the gas will spontaneously develop into filament-shaped droplets of micrometer dimensions which persist as stable packages even after the gas is released from its container, he says.

This new ability to produce gas droplets is exciting as it opens the door to a range of potential applications.

"These droplets could, for example, form pristine nano-laboratories for performing chemistry reactions or making highly precise measurements of magnetic fields," Professor Blakie says.

Professor Blakie is a member of the Dodd-Walls Centre for Photonic and Quantum Technologies, a national Centre of Research Excellence involving five universities, hosted by the University of Otago.

This work was supported by a grant to Professor Blakie from the Marsden Fund of New Zealand, which supports the lead author Danny Baillie (a research fellow at Otago). The two other co-authors on the paper are Otago PhD graduate Dr Russell Bisset, who is now a postdoctoral researcher in Italy, and Professor Ryan Wilson, a collaborator at the US Naval Academy.

Fresh approaches for modelling geothermal systems

"We worked out that it had to be at a temperature of a few millionths of a degree above absolute zero, a density more than a billion times lower than liquid water (about 100,000 times lower than air), and in a suitably adjusted magnetic field."

The Geothermal Institute at the University of Auckland is one of the premier geothermal research and training centres in the world, and GNS Science is New Zealand's leading provider of Earth, geoscience and isotope research and consultancy services. A major focus for the partnership is working with the geothermal energy industry to improve the efficiency of existing field management operations and developing assessments of new geothermal resources. This work has helped double the contribution that geothermal makes to New Zealand's electricity production in the past 17 years – a strong contribution to the Government's goal of having 90% of New Zealand's electricity supply generated by renewables by 2025.

The Geothermal Institute and GNS Science are developing next-generation approaches to the numerical modelling of geothermal systems as part of an innovative MBIE-funded project led by GNS. Sustainable use of these multiple geothermal systems in the same region over many generations is critical. We need new techniques to reliably predict the impact of using the energy from a geothermal system on both that system and any neighbouring systems including new computer modelling software and the coupling of models from different scientific areas.

The software, currently in its final year of development, will be released as open-source software that can be used by the geothermal sector worldwide. It will take advantage of the increased computing capability that has become available in recent years to provide a robust environment for building large-scale models. A beta version is currently being tested on a range of computing environments, including high-performance computers from New Zealand eScience Infrastructure (NeSI), which provides support systems to enable New Zealand researchers to tackle the world's biggest problems.

The Geothermal Institute and GNS utilise computational resources provided by NeSI to apply software and modelling techniques to a diverse range of real-world problems in the geothermal sector.

Internationally, there is strong interest in the outputs of this programme as they will enable offshore companies to improve the management of geothermal fields. This uptake will strengthen New Zealand's reputation as an international leader in geothermal research and development.

The new modelling suite will add significant capabilities to existing software and be more efficient and reliable. It will integrate seamlessly with existing geothermal research in New Zealand, increasing knowledge of the structure, hydrology and hydrothermal processes that control geothermal systems.

Golden mystery solved

A long-standing discrepancy between experiments and theory concerning the electronic properties of gold has now been resolved. Gold's lustrous colour is due to unusually strong relativistic effects. The same effects also complicate theoretical computations of gold's electronic properties. Indeed, theorists working on this precious metal have struggled for decades to resolve a discrepancy between their predictions and experimental observations. New work has solved this problem by calculating the electron correlation contribution to an unprecedented level of precision that incorporates "pentuple" interactions between five electrons.

Calculating an atom's electronic properties is never easy, especially for heavy atoms whose strong Coulomb potential implies relativistic energies for its electrons. In gold's case, relativistic effects cause a smaller than expected gap between the 6s and 5d orbitals, which is why gold absorbs blue frequencies and reflects a yellowish tint. But other aspects of gold are more difficult to explain. Calculations of the ionization energy (energy to remove an electron) and electron affinity (energy to add an electron) have consistently underestimated the experimental values by tens of milli-electron-volts.

Peter Schwerdtfeger and his colleagues from Massey University, Auckland have performed precise calculations for gold using NeSI's high performance computing resources. Their model accounts for relativistic effects, as well as for the contributions from electron correlations and quantum electrodynamics. Electron correlations embody all the electron-electron interactions that occur in a multielectron atom. Previous studies have dealt with electron correlations between the 79 electrons in gold, but typically they have only gone as far as triple interactions between three electrons. Schwerdtfeger's team extended these calculations to quadruple and pentuple interactions. By doing so, they reduced the discrepancy in the ionization energy and electron affinity to just a few milli-electron-volts – a factor of 10 improvement over past results. The methodology could be applied to even heavier elements.

This research is published in Physical Review Letters and this case study is adapted with permission from an article by Michael Schirber in the American Physical Society publication, Physics.

Improving heart disease prediction and prognosis

"How simulating arterial flow might lead to decreased heart disease mortality."

Coronary heart disease (CHD) resulting from plaque build-up in the arteries (atherosclerosis) remains the most common cause of hospital admissions and mortality in many developed countries. It is thought that arterial geometry, cell function and blood-flow dynamics play an important role in triggering CHD. Wall shear stress and disturbed flow in the vicinity of wide angle bifurcations have also been pinpointed as potential triggers of arterial plaque build-up.

In collaboration with NeSI and the University of Canterbury's High Performance Computing department (UC HPC), PhD student Stewart Dowding, Dr Constantine Zakkaroff and Prof. Tim David have developed the Coupled Arterial Cells project. At its core, this project has a massively parallel, multi-scale software framework for modelling cellular-level signalling with the goal of providing invaluable insight into coronary heart disease. Computer (or in silico) simulations enable the rapid exploration and pruning of the parameter search space for the refinement and integration of micro-level models into biologically realistic macro-scale models. The coupled cells simulations provide the opportunity to perform physiological experiments otherwise impossible to conduct in a clinical environment.

Current simulations include approximately 1.2 million coupled arterial cells whereas a complete coronary tree might contain on the order of 100 million arterial cells. One of the milestones of this project is to increase the simulation size in order to get a full picture of the emergent macro-scale behaviour of plaque deposition at the level of a complete coronary tree.

NeSI was called in to help port the Coupled Arterial Cells project from UC HPC's BlueGene platform to Fitzroy, the supercomputer operated by NIWA for national weather forecasting. In the process, NeSI improved the CMake build system by adding unit tests and the option to instrument the code with TAU (Tuning and Analysis Utilities). The team has now embarked on generalising the domain decomposition in the code to allow for more cells per sub-domain. This will make the Coupled Arterial Cells code ideally suited to run on massively parallel architectures with many cores per node, such as Fitzroy and future NeSI platforms.

Our small-segment arterial simulation results show that wide bifurcation angles can suppress arterial wall calcium signalling.

UC scientists make biomolecular breakthrough

A team of University of Canterbury researchers has made a scientific breakthrough in biomolecular interactions, which will help discover the determinants of gene expression.

UC Bioinformatics Senior Lecturer Dr Paul Gardner says, "We have discovered a previously overlooked mechanism that controls additional variation in expression. Our new approach considers the extent of how random interactions between messenger RNAs (that are translated into proteins) and non-coding RNAs can influence gene expression."

The UC team has proven, using some large and complex datasets, that these interactions play a major role in gene expression. The findings, which used UC High Performance Computing, have been published today by eLife scientific journal in a research paper titled: Avoidance of stochastic RNA interactions can be harnessed to control protein expression levels in bacteria and archaea, proving that stochastic mRNA and non-coding RNA hybridisation has a major impact on protein expression.

Dr Gardner says there are a number of potential biotech applications as a result of the research findings. "Particularly in the space of designing mRNAs for genes so that protein production can be improved."

Computational tools were used to predict the strength of binding between the RNAs found in the dominant forms of microbial life on Earth: the bacteria and the archaea. This approach revealed that the majority of messenger RNAs bind more weakly to the most common RNA molecules found in cells than would be expected by chance. Weakened binding should prevent the RNA molecules from becoming tangled with each other and ensure that protein levels are not perturbed by unintended interactions between highly expressed messages and other RNAs.

To test this hypothesis further, the researchers generated versions of the gene for a green fluorescent protein that differed only in how well their messenger RNAs could avoid interacting with the most abundant RNAs in E. coli cells. Those messengers that were designed to avoid interacting with other RNAs yielded far more protein than those that were not.

The findings show that taking this kind of avoidance into account can improve predictions about how much protein will be produced and should therefore make it easier to control protein production in experimental systems.

NeSI explores innovative uses of cloud services for advanced computing

In 2016 NeSI created a seamless integration from a current platform to two public cloud vendors offering diverse services. This pilot project was successful, resulting in a working capability to burst from the NeSI platform to both cloud services, and building our knowledge and experience in cloud integration.

This kind of integration of public or private cloud computing will allow NeSI to be more responsive to researcher needs. Cloud bursting can provide more capacity at peak times by moving certain workloads onto the cloud, and in some cases mitigates the risk of service unavailability due to infrastructure failure. Access to specialised computing environments can enable researchers to apply technologies and solutions not already deployed on NeSI. Development of virtual labs provides a well-curated community-centric solution for meeting common data workflow, processing, and visualisation needs.

The pilot involved designing two different solutions to evaluate the feasibility of offloading workloads to the cloud. They were based on Infrastructure as a Service (IaaS) and Software as a Service (SaaS) models. It became clear through the project that the design and implementation of the IaaS solution was more straightforward and would lead to a better experience for the researcher. Both solutions could form part of NeSI's future infrastructure.

In the first half of 2017 we will take the learnings from the 2016 pilot and implement and trial cloud services with a set of researchers open to exploring this new way of working. In addition to the seamless cloud burst capability we will be looking at the life-cycle aspects of working with cloud environments, to ensure our future integrated infrastructure supports appropriate policy, reporting, and accounting.

Helping communities anticipate flood events

"The aim of this collaborative consultation project was to enable scientists to obtain results faster, run larger simulations with wider catchment areas, and execute bigger ensemble runs."

The ability to forecast flood events in response to intense rainfall is critical to saving lives across New Zealand and preserving infrastructure.

TopNet is a hydrological code developed at NIWA , which uses concepts of runoff generation controlled by subsurface water storage and topography. It combines a water balance model within each sub-catchment to simulate water flow, soil moisture, lake levels, discharge in rivers and streams over time, taking into account precipitation, snow, evaporation and plant transpiration. It is run four times daily to predict potential flooding in New Zealand. The model is also being applied to assess the impact of climate change on water availability and to estimate changes in the probability of future flood events as global temperatures increase.

NeSI was called in to help NIWA improve TopNet's performance on High Performance Computing (HPC) machines. The aim of this collaborative consultation project was to enable scientists to obtain results faster, run larger simulations with wider catchment areas, and execute bigger ensemble runs for improved uncertainty quantification.

Our first task was to introduce a CMake build system, which supports out of source builds, automates to a large extent the search of library dependencies, and makes it easy to switch between compilers and compiler options.

Much of the data written by TopNet is sparse and thus can easily take advantage of netCDF/HDF5's on-the-fly compression – the amount of disk space used by sparse data can be greatly reduced at very little computational cost. By enabling compression in forecast simulations we achieved an approximately 80% reduction in file sizes. This will reduce load on the file system when many instances of TopNet run concurrently.

The University of Otago is at the technological frontier

University of Otago researchers can now send and receive enormous amounts of data at ultra-fast speeds to stay at the forefront of their field – prompted by Dunedin gout researchers wanting to develop IT capability in New Zealand.

Biochemistry Professor Tony Merriman wanted to confirm results of studies about which genes underlie gout, while also isolating new genes that could raise the risk of developing gout, influence how serious the condition gets, and be a factor in flares.

A human genome contains more than three billion DNA bases, and the Otago researchers need the sequencing to reveal the order those bases appeared in each whole genome. The researchers can then compare individuals and look for differences relevant to gout, on a scale never attempted before.

The data pipeline is so complex, that over the project's lifetime about half a petabyte of data will be transferred and stored.

The ITS team's first challenge was to get the 100 terabytes of raw data from the USA to Dunedin. They created a suite of services known as a Science DMZ, a demilitarised zone between networks that is specifically designed to allow big research data sets through with the fewest possible obstructions, at the highest speeds possible.

The ITS team still had to get the raw sequencing data from Dunedin to the NeSI PAN supercomputer in Auckland. So, Otago postgraduate biochemistry student Murray Cadzow wrote a program which calls the University's new Data Transfer Service to get it to send 80 gigabyte slices of data north on the superhighway provided by REANNZ, then gets the supercomputer to start processing the data.

The supercomputer will now process gout data for the next 12 months – day and night – to return about 250 terabytes of data to the University in total.

"Without the help of Otago ITS and NeSI … this project would not have been possible to do within New Zealand." Murray Cadzow, University of Otago Biochemistry Department postgraduate student.

Growing computational capacity among wet-lab scientists

"Coding is something we shouldn't be afraid of and it is something we should embrace as scientists moving into the future of big data."

Computational lab skills are becoming essential for researchers across disciplines. Use of High Performance Computing infrastructure, such as the one provided by NeSI, requires such skills and one of the key challenges is that learning them is not typically a part of the curriculum. Plant & Food Research (PFR) has successfully leveraged the computational skills training programme supported across New Zealand by NeSI. The basis of this training is the Software Carpentry (SWC) workshop that teaches researchers how to automate their workflows using command line, adapt best practices in programming and use version control.

In the first half of 2016, PFR hosted their first Software Carpentry workshop, hosted in Auckland and supported by NeSI. The workshop received a lot of positive feedback and provided a great experience for the PFR researchers involved in it.

Amali Thrimawithana, Sarah Pilkington and Lara Brian, researchers at PFR, co-ran the workshop, helping to set it up and then becoming a part of the instructor team.

As Amali said: "We had to make sure we had backups especially in terms of IT/network issues, so thinking about those possible issues ahead of time and getting a few of the attendees to test the system and always having a plan B was important. For example, in one session of our workshop we got the attendees to use some test RNASeq data and knew there was a possibility someone may accidentally delete the raw data. That meant some of the trainees had to know how to access the backup data. We had to go to plan B for this one a few times on the actual day! Also having to deal with different operating system issues."

The main audience was wet-lab plant scientists (biochemists, molecular biologists, physiologists, entomologists) with zero coding experience and a fear of the black screen of doom.

"As fellow scientists with no coding background we could relate to their fears of learning something so foreign and make it seem a bit more user friendly."

The talk was well received by both scientists and bioinformaticians. It gave people the confidence to explore coding for the first time.

As a follow up, PFR Auckland has introduced drop-in sessions onsite for those wanting to continue work and had people dropping in to get assistance, self-motivated by their new coding skills. For the teachers and helpers involved in the workshop it was clearly a great experience which contributed to their professional development. This experience has sparked the momentum for increasing digital literacy at PFR.

New frontiers for milk proteins: functional foods and drug delivery systems

"Our ongoing use of NeSI has allowed us to investigate the process of adsorption of b-Lactoglobulin at different oil/water interfaces, and so determine the physico-chemical forces that govern adsorption behaviour."

The ability to control and manipulate protein behaviour at an interface can influence a broad range of technological processes and physicochemical and biological phenomena. In some cases, such as the development of pharmaceutical products and functional emulsion-based food, protein adsorption is a desirable feature that enhances the end product. In other situations adsorption is undesirable.

Although many experimental studies have contributed to rapid progress in the fundamental knowledge of protein behaviour at interfaces, detailed molecular-level understanding of the mechanism of protein adsorption at an interface is still remarkably lacking. The high performance computing resources available through NeSI allowed Davoud Zare of Victoria University of Wellington, supervised by Prof. Kate McGrath (MacDiarmid Institute/Riddet Institute/ Victoria University) and Dr Jane Allison (Massey University), to use atomistic molecular dynamics simulations to characterise the adsorption of the milk protein ß-lactoglobulin (ß-LG) at three different oil/water (O/W) interfaces.

Each simulation comprised a periodic cubic box containing one molecule of ß-LG (1604 atoms), nine sodium ions and large oil and water phase regions, giving 64,115 – 124,367 atoms, or interacting particles, in total for each system. As the pairwise interactions between the particles must be evaluated at each integration step, the computational cost scales up enormously as system size increases.

In addition to these being large simulation systems, it was also necessary to test whether the initial orientation of ß-LG with respect to the oil surface affected its adsorption, and to run each simulation for sufficiently long for stable adsorption to occur. At least two independent simulations of each of three different initial orientations were run for each oil, making 22 simulations in total. Finally, separate simulations of each type of oil interfaced to vacuum and to water were also run to verify that the simulations accurately capture the interfacial tensions of the various phase combinations. Each simulation was run for 200 ns, with an integration time-step of 2 fs, requiring 100,000,000 integration steps. The NeSI Postgraduate allocation class allowed these simulations to be carried out using the GROMACS MD software across 192 cores on the University of Auckland Pan cluster or 64 cores on the University of Canterbury Power755 cluster.

Not only did the nature of the adsorption observed in the simulations largely agree with experimental findings, the simulations revealed the structural and mechanistic detail behind each mechanism of adsorption, paving the way for the use and engineering of protein-stabilised emulsions as functional food products and drug delivery systems.

Molecular fingerprinting

Infrared spectroscopy is often referred to as 'molecular fingerprinting', enabling molecules to be identified by their unique vibrational signature. It is the main technique used for identifying contraband substances such as narcotics, explosives and illegal pharmaceuticals. However, as with standard human fingerprinting, the 'culprit' can only be identified by comparison with a database of known substances. If the substance isn't in the database, then no identification can be made.

While it is impossible, both in theory and in practice, to predict what human fingerprints will look like, even if an entire individual's genome is known, it should be possible to predict a molecule's fingerprint. Developing the theoretical and computational tools required to do this has been the focus of Dr Marat Sibaev's PhD project (Department of Chemistry, the University of Canterbury).

The key insight from Dr Sibaev's work is that it is easiest to mathematically model how the energy changes as the molecule moves in 'chemically recognisable' ways – as bonds stretch, and angles bend, and rotations around bonds occur – but determining the kinetic energy due to molecular vibrations is much easier if the atoms are assumed to move in straight lines.

Dr Sibaev has developed new software to help interconvert between these two different perspectives on atomic and molecular motion and then deployed it on NeSI's supercomputing facilities at the University of Canterbury, enabling him to predict the vibrational fingerprints of molecules from first principles, using a completely automated procedure.

This potent combination of theory and code development, and access to the supercomputing resources available through NeSI, now allows routine prediction of molecular fingerprints for molecules with up to 25 atoms, encompassing most known drug-like or potentially explosive molecules. This is a substantial increase in capacity beyond what is possible using other current methods, which typically have only been applied to quite small molecules, comprising less than 10 atoms.

According to Dr Sibaev's supervisor, Dr Deborah Crittenden, "Access to the increased computing capacity of the NeSI supercomputers facilitated Dr Sibaev's research by speeding up the rate at which he was able to generate the data required for his project. In particular, he made extensive use of the NeSI computers to calculate how a molecule's energy changes as the molecule vibrates and predict molecular fingerprints from this data."

Retired but not forgotten: Algorithm research with a Supercomputer

Professor Emeritus Tadao Takaoka, a recently retired computer scientist from the University of Canterbury, has devoted his research career to designing efficient algorithms for classical computer science problems – shortest path, data structures, combinatorial generations and subarray problems to name a few.

In the area of parallel algorithms, Prof. Takaoka has long been fascinated by 2D mesh architecture, a two-dimensional network of computing nodes, where each node is relatively light with little computing power and is allowed to communicate with its four neighbouring nodes only. This is a strict constraint, but with clever orchestration of data movement, sending the right data to the right neighbour at the right time all happening in parallel, one can still compute the solution very fast.

This architecture is easy to mass-produce in the form of a chip, such as an Application Specific Integrated Circuit (ASIC) or Field Programmable Gate Array (FPGA). These chips are widely used in most modern digital devices to carry out a special function. A parallel algorithm designed for this architecture potentially gives the power of a parallel computer in a tiny form factor at a correspondingly tiny cost.

Takaoka and his students have been designing 2D mesh algorithms for various problems – all-pairs shortest path, maximum subarray, matrix multiplication and maximum convex sum – since the 1980s. One of his PhD graduates, Sung Bae, recalls: "When the first Blue Gene/L was installed at BlueFern [now UC HPC] in 2007, I knew it was the ideal machine to test my algorithm, but sadly I had already finished my PhD." Since those days, Takaoka reinstigated 2D mesh algorithm research and used NeSI's technical support to drive Blue Gene. New solutions for old challenging problems came one after another.

Takaoka says: "Parallel computing is a fascinating area where theory and practice meet in a productive and tangible way. We need a lot of skills for parallel programming through practical experience. It is for this reason that I introduced the use of Blue Gene into my Stage 4 course 'Advanced Algorithms' as well as for my own research. We are very grateful to NeSI for the use of Blue Gene and the supporting team. Thanks to their generosity, we were able to develop a culture of supercomputing and parallel programming at the University of Canterbury."

Prof. Takaoka, even though retired from the University of Canterbury, is still active and now leads a "research club" called Algorithm Research Institute (ARI) with his PhD graduates Tongwook Shinn and Sung Bae. NeSI's Blue Gene/P supercomputer was retired in June 2016. Like Prof. Takaoka himself, it will leave a tremendous legacy and a large number of research papers with its name attached.

When getting back to basics is better than fancy new tools

Prof. Brendon Bradley of QuakeCORE (based at the University of Canterbury) has used a parallelised software, designed for 3D seismic wave propagation simulations on NeSI's BlueGene/P supercomputer, to develop a sophisticated understanding of the effects of earthquake-induced ground shaking. This software, developed by US collaborator Rob Graves (US Geological Survey), enabled the team to simulate all the events in the Canterbury earthquake sequence. NeSI has been supporting this research from the early stages, and has been providing HPC and visualisation services to help develop an intuitive understanding of the model.

Prof. Bradley's team has recently completed the simulation of the Alpine Fault scenario using 8192 cores of BlueGene/P, and discovered that post-processing the output data was frustratingly slow, taking 5.5 days to complete (longer than the main computation itself). NeSI's algorithm specialist, Dr Sung Bae, embarked on an IO optimisation project to overcome this problem.

Bae devised an algorithm that reassembles a long contiguous segment on-the-fly in memory before writing to the disk. This approach eliminated the need for frequent jumps on the disk surface and showed a dramatic improvement. The new parallel solution using four POWER7 cores completes the merge in 12 minutes, which used to take 5.5 days – a 660-fold speedup. While it still took substantial work to materialise this idea, the solution, which yielded an outcome beyond expectations, relied primarily on simple programming concepts.

Prof. Bradley says, "The new approach has been critical for our research workflow. We couldn't have continued with post-processing taking so long for the large computations, and it would have taken several months (if at all) for someone in our research team to learn the necessary MPI IO skills and develop a feasible solution. This is a classic example of the huge advances that are being achieved through our excellent collaboration with NeSI."

Prof. Bradley and QuakeCoRE accessed their compute hours and consulting support through a NeSI Merit Project. Merit Projects are designed to support high-quality, investigator-driven research which clearly aligns with New Zealand's research priorities, providing free allocations of platform time and consulting, and high-priority status in the system.

Finite element modelling of biological cells

Understanding how salivary glands work to secrete saliva has important applications for oral health, since a lack of saliva can cause a range of medical conditions. With the aid of NeSI's high performance computing infrastructure, Professor James Sneyd and his research group (Department of Mathematics, University of Auckland), in collaboration with Professor David Yule (School of Medicine and Dentistry, University of Rochester), have been using mathematical modelling to answer this interesting physiological question.

This collaborative consultation project was aimed at optimising the custom finite element method (FEM) code the research group use to run simulations of the salivary gland acinar cells, enabling the researchers to scale their simulations up and make the best use of NeSI's platforms. The primary goals were to obtain the best possible performance from the code, while ensuring its robustness and ease of use. The main developer of the code within the research group was John Rugis, with NeSI support for this project provided by Chris Scott, through NeSI's Scientific Programming Consultancy Service.

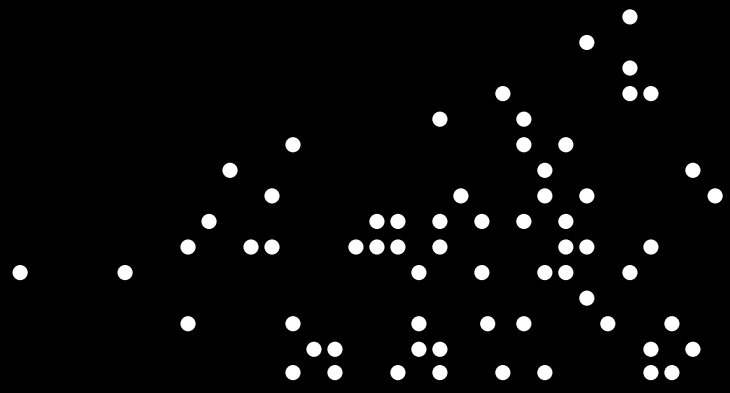
As part of this project a series of benchmarks and comparisons were performed between libraries for solving sparse linear systems. The ViennaCL library was eventually selected, as it performs well and supports a variety of parallelisation options (CUDA, OpenCL and OpenMP), allowing the use of accelerators such as GPUs.

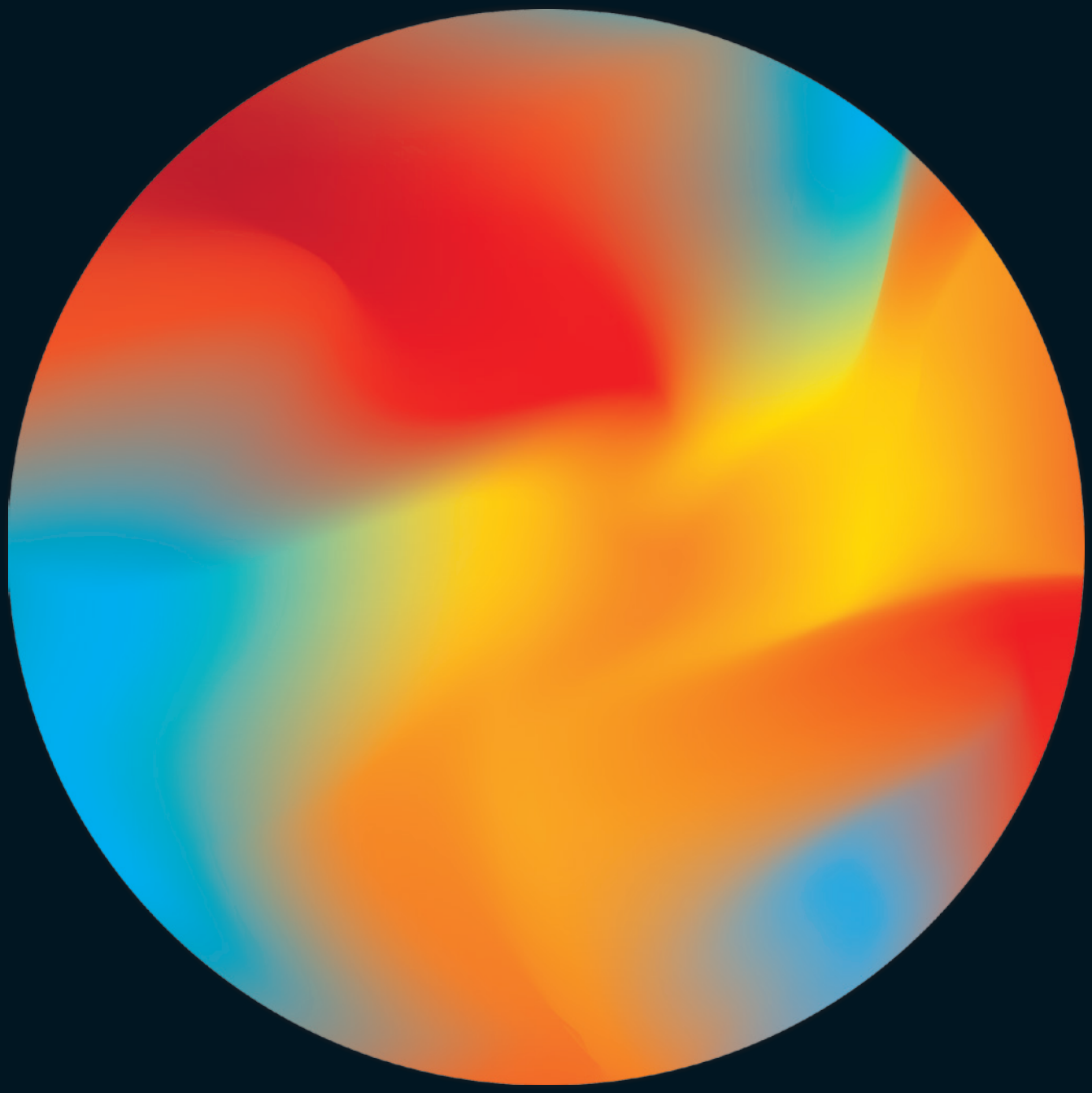
When profiling the serial version of this code it was found that only a small proportion of the run time was spent solving the sparse linear system, with a large amount of time spent on other operations outside of the solver. After further investigation, an 8x increase in performance was gained by changing the matrix storage type and removing this bottleneck. After these improvements we benchmarked the program showing that the GPU version performed twice as fast as the serial version.

Continuous Integration (CI) testing was implemented using the Travis-CI service, which is easy to integrate with GitHub and free for open source projects. A weekly test of the code on the NeSI Pan cluster was also implemented. This testing gives the researchers confidence that their simulations are behaving as expected.

As a result of this project the researchers' simulations are now running faster, it is easier for the researchers to submit and analyse large parameter sweeps and robust software development practices are being utilised, giving the researchers confidence in their results and making future development easier.

NOTES





NeSI
New Zealand eScience Infrastructure

nesi.org.nz