

Improving NeSI researchers' productivity with a consultancy service

Chris Scott Alex Pletzer Wolfgang Hayek Georgina Rae

New Zealand eScience Infrastructure (NeSI)

6 September 2019



Outline



① NeSI Consultancy service

② Some recent projects

③ Summary



NeSI Consultancy service

What is NeSI Consultancy?

Introduction



- Goals of NeSI's Consultancy service
 - Enable increased scale and complexity of research
 - Help researchers use NeSI's computing capabilities to increase research outputs and impacts
- How we achieve this
 - Working on projects with researchers; typically 1 day per week for 3-4 months
 - The scope of projects we work on is very broad and covers topics such as:
 - Code optimisation and/or parallelisation
 - Workflows
 - Improving software sustainability
 - Porting code to NeSI's platforms
 - Custom code development
 - Visualisation

Consultancy projects

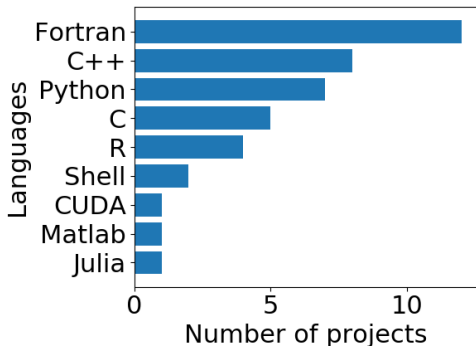
Applying for a project



- Usually researchers who meet the following criteria can access consultancy at no cost to themselves, based on their institution's or MBIE's investment into NeSI:
 - Researchers with Merit or postgrad projects
 - Researchers based at collaborator institutions
 - University of Auckland
 - NIWA
 - University of Otago
 - Manaaki Whenua - Landcare Research
- Others may be able to pay to access the service
- Application form
 - Usually filled out after an initial meeting
 - Defines the current state of the problem
 - Defines the scope of the work

Consultancy projects

Languages



- The programming languages that we have worked with
- Some projects had more than one language
- Fortran, C++/C still dominate
- Python and R are popular too
- We started work on our first project with Julia recently



Some recent projects

Getting closer to more accurate climate predictions for New Zealand

Researchers:

Erik Behrens & Jonny Williams

National Institute of Water & Atmospheric Research (NIWA)

“The NZESM is a community project. With this upgrade we have an exciting tool with which to study the climate.”

NeSI delivered:

- High performance computing resources
- Computational science expertise



Code development

NZESM project

- Multiple models that need to be coupled together (atmosphere, ocean, ...)
- Using a tool called OASIS for coupling these models together
- We added a new feature to OASIS to enable an ocean coupling scheme for coupling a high resolution regional ocean component
- Important step to setting up customised earth, ocean and weather system modelling for New Zealand



Understanding the behaviours of light

Researchers:

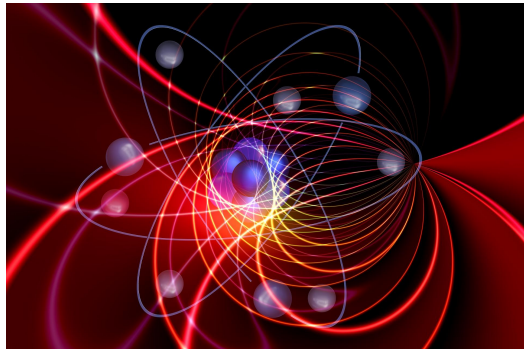
Victor Canela & Jacob Ngaha

University of Auckland

“We are running close to a hundred jobs at the same time and each job has separate parameters...**NeSI was able to supply a script that keeps track of this automatically.**”

NeSI delivered:

- High performance computing resources
- Computational science expertise
- Code optimisation, automation, improved workflow



Cascade trajectories project

- Infrastructure
 - Added a CMake build system
 - Easy to port to different systems
 - Easy to try different compilers – Cray compiler 40% faster
 - Adding tests
- Optimisation
 - Optimised performance of serial Fortran code
 - 22 % boost
- Workflows
 - Instead of running one long calculation they can run many shorter calculations
 - Created Python/Slurm scripts to enable this parallel workflow and give a big boost in performance

Fractal analysis of brain signals for autism spectrum disorder

Researcher:

Stephen Wolfson

Department of Psychology, University of Auckland

“NeSI’s participation in our project is essential...

It’s really only possible to do these calculations on large data sets using a cluster computer.”

NeSI delivered:

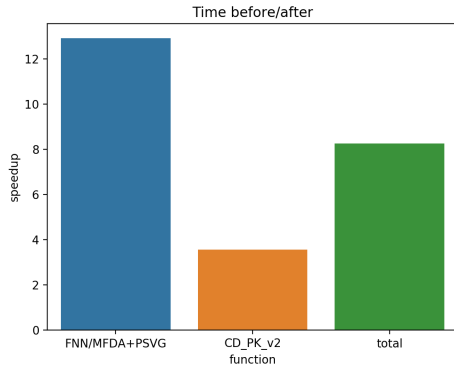
- Computing resources
- Assistance with optimising script to increase computational speed



Fractal analysis project

Code optimisation

- Matlab code
- Identified two functions with big loops that were performance bottlenecks
- Converted these functions to C++ and call from Matlab – mex'ing
- Total speedup of 8x from converting these functions
- Other outcomes
 - Taught the researcher how to use git version control
 - Importance of testing – adding a checksum to easily check result



Trinity project – overview

Workflows and runtime tuning

- Genomics code - de novo assembly
- Complex workflow consisting of two distinct computational phases
 - Initial phase is multithreaded with high memory requirements (100s GB)
 - Second phase has many millions of short, serial embarrassingly parallel jobs with low memory requirements (≈ 5 GB)
- By default Trinity runs everything at once within a single high memory node – not optimal



Trinity project – solution

Workflows and runtime tuning



- HPCGridRunner – launch embarrassingly parallel jobs on a compute grid (cluster)
 - Accumulated I/O bandwidth of many nodes gives much better performance

Type	Num cores	Elapsed time (hh:mm:ss)	Core hours
Single node	16	24:09:36	387
Grid	20	07:59:58	168
Grid	40	04:10:45	171
Grid	60	02:36:58	160

“The NeSI consultancy service has drastically improved my workflow by reducing RNA de novo assembly time from 7 days to 24 hours!” – Alexis Marshall, University of Waikato



Summary

Summary

- NeSI has a consultancy service
- Work on a wide range of projects (code optimisation, runtime tuning, ...)
- We are always looking for new projects – get in touch if you are interested
 - support@nesi.org.nz

<https://www.nesi.org.nz/services/consultancy>