



BeSTGRID: Developing NZ research infrastructure

Nick Jones
Director, BeSTGRID
Co-Director eResearch, Centre for eResearch
University of Auckland
n.jones@auckland.ac.nz

www.bestgrid.org
[@bestgrid](https://twitter.com/bestgrid)

BeSTGRID: since 2006

“**BeSTGRID** aims to enhance e-Research capability in New Zealand by providing the **skill base** to help the various research disciplines engage with new eResearch services”

“**BeSTGRID** aims to kick start **centralised infrastructure** with some capital investment at key institutions”



BeSTGRID: since 2006

“BeSTGRID could well be New Zealand’s earliest dedicated eResearch infrastructure providers – even predating KAREN. Since 2006, BeSTGRID has been delivering services and tools to support research and research collaboration on shared data sets, and in accessing computational resources.”

**KAREN hypen
Issue 09
September 2009**





Massey University



AUT
UNIVERSITY
AUCKLAND, NEW ZEALAND



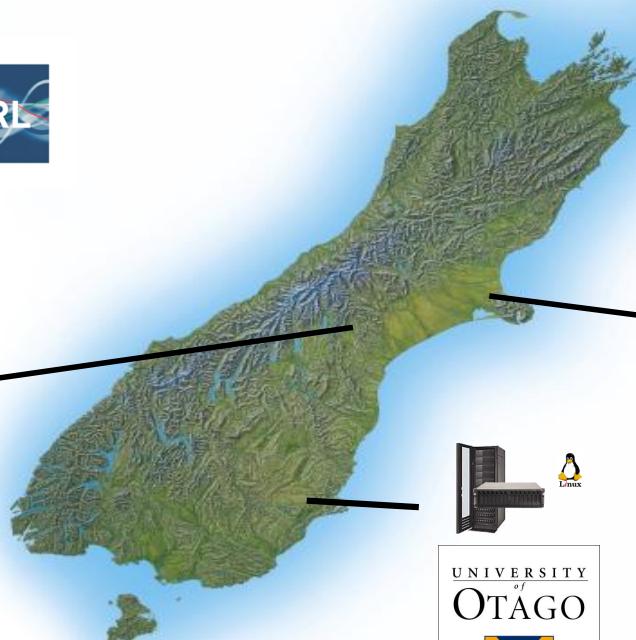
Manaaki Whenua
Landcare Research



INDUSTRIAL RESEARCH LIMITED
Te Tauhi Pūtaiao



Lincoln
University
Te Whare Wānaka o Aoraki



THE UNIVERSITY
OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau



Sakai

EVO



THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI
VICTORIA
UNIVERSITY OF WELLINGTON



UC
UNIVERSITY OF
CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND



UNIVERSITY
OTAGO
Te Whare Wānanga o Otāgo
NEW ZEALAND



Computation
Cluster



Identity provider



Storage



Virtual Applications

— — — Planned
— Current

Successes: within NZ

National grid infrastructure, primarily focused on Bioscience and Geoscience applications, services, data, computation, along with many others

Piloted first **federated identity management** service in NZ, now moving to production

Established the foundations of **first shared-infrastructure service delivery** model across research sector institutions (Universities, CRIs)

Built foundations for sector wide HPC and eResearch programme:

- **National eScience Infrastructure** proposal : 2010 – 2015

What did we learn?

Stay Connected, Collaborate

- With Scientists, Administrators, Technologists, Policy makers, Funders
- Locally AND Internationally
- move much faster !!

Buy / Translate - don't build

- Source from international community: technologies, approaches, strategies, configurations, programmes

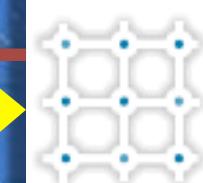
Align

- Essential to work across groups, towards agreed goals
- Fostering loyalty and clear lines of management essential

* UK eScience - JISC
* eFramework – JISC
Shibboleth – JISC, SWITCH
SLCS - Switch
Sakai - JISC

* VDT - OpenScienceGrid
* Globus – Argonne
* Shibboleth – Internet2
* iRODS – DICE / RENCI
* Sakai – Sakai Foundation
* OCI - NSF
* PRAGMA – UCSD
* EVO – CalTech
* AccessGrid - Argonne
* GenePattern – MIT Broad
* caBIG - NIH

* Grisu – ARCS
* Confluence - Atlassian
* NCRIS - DIISR
* QFAB - QCIF



BeSTGRID
Broadband enabled Science and
Technology GRID

What did we learn?

Change difficult within Institutions

- Medium to Long term timeframes required with researchers, ITS, SMT, research support units
- ***“Final Report: A Workshop on Effective Approaches to Campus Research Computing Cyberinfrastructure”***
Klingenstein, Morooney, Olshansky. 2006

Collaboration seen as critical, yet complex and expensive

- **Working with user communities** essential to mutual understand of needs and opportunities, and developing capabilities
- Need significant capability before international collaborations (and related learning) possible
- New forms of organisation and support required
- ***“Beyond Being There: A Blueprint for Advancing the Design, Development, and Evaluation of Virtual Organization”*** Cummings, Finholt, Foster, Kesselman, Lawrence. 2008
- ***“The Importance of Long-Term Science and Engineering Infrastructure for Digital Discovery”*** Wilkins-Diehr, Alameda, Droege, Gannon. 2008

Government and Institutional Investment well below international levels

- Australia NCRIS: \$540M ++ over 5 years alongside State and Institutional investments
- EU FP7 e-Infrastructure call + National Grid Initiatives + Identity + HPC + Networks in member states
- US Office of Cyberinfrastructure, NSF, NIH, DoE all funding major Cyberinfrastructure initiatives

Education programs lacking

- Who is training the next generation?

MoRST eResearch programme

In Vote RS&T 2008/09, funding was appropriated to develop capability within the Kiwi Advanced Research and Education Network user group to make effective use of KAREN:

- **Advanced Video Conferencing Collaboration and Support Centre**
- **Federated Identity and Access Management**
 - Towards a Federated Approach – strategy development
 - Technical Support and Resources
- **Semantic Data and Public Access to Research**
- **BeSTGRID Grid Middleware Initiative**
- Hardship fund **for remote site connection to KAREN**

BeSTGRID 2009 – 2010

- MoRST Dec 2008 funding request for
Middleware development for eResearch in NZ
- Provides ongoing support for BeSTGRID
 - Widened from founders through active participation of new Universities and CRIs
- Aims:
 - Coordinate access to compute and data resources
 - Provide discipline specific services and applications
 - Build a sustainable resource administration, middleware and applications & services development community

People: Technical Working Group

Terms of Reference

- Coordinates nationally distributed compute systems work programme
- Meet monthly: present and discuss proposals, progress reports, and services and applications deployed

Membership

- Institutions contracted to BeSTGRID + resource providers + collaborators

Terms of Reference, Meeting notes, etc on BeSTGRID website

http://www.bestgrid.org/index.php/Grid_Technical_Working_Group

Computation: Grid Resources

Resource name	Location	Architecture	OS	#CPU cores	Freq	Memory	Notes
Auckland Cluster	Auckland	x86 Xeon	Linux	200	2.8 GHz		Operational, since June 2008. Current Status
AUT Cluster	AUT		Linux	32			Connected end September 2009
BeSTGRID Prototype Cluster	Canterbury	x86 Xeon	Linux	4	3.0 GHz		Prototype cluster, serving the community since June 2007.
BlueFern p575	Canterbury	Power5+	AIX Linux	128	1.9 GHz		Midyear 2007. 8 nodes, 16 CPU cores each.
BlueFern BlueGene	Canterbury	PowerPC	Linux	4096	770 MHz		Midyear 2007. Available only together with a local account .
Maggie	Otago	x86 Pentium 4	Linux	10	3.0 GHz		2009 No MPI available
Massey Cluster	Masseu	x86 Xeon	Linux	160			2007 Infiniband interconnect
SCENZ Grid Cluster	Landcare	x86	Linux	104		384 GB	2010 Fibre Channel SAN

Data: Resources and Services

- 90TB storage, 30TB @ foundation BeSTGRID site
 - Auckland, Canterbury, Massey
- Data transfer & access services
 - GridFTP
- Tiered storage performance
 - Secure, Backup, Archive
- Distributed Storage
 - iRODS – under development with ARCS

User-Centered Design

- Aligned with national research & economic growth agenda
- Lead Users from each community
 - Define requirements
 - Evaluate developments
 - Promote services
 - Disseminate learning
- Technology group
 - User needs lead requirements
 - Scans for technology to acquire
 - Incremental developments

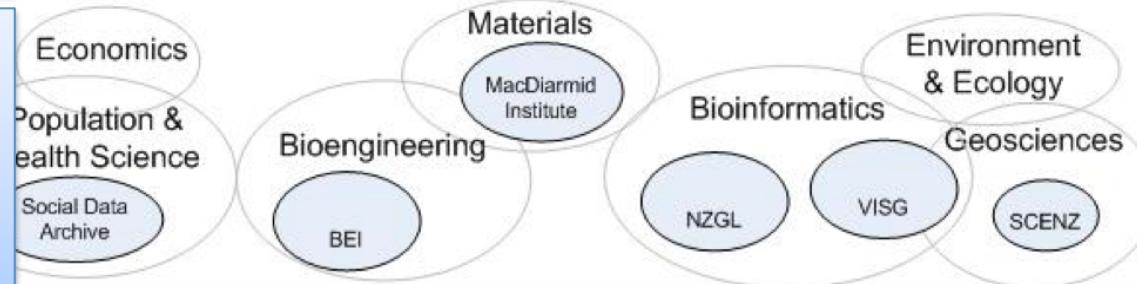
Governance / Priorities

Nimble, responsive & fair
Connects to economic drivers



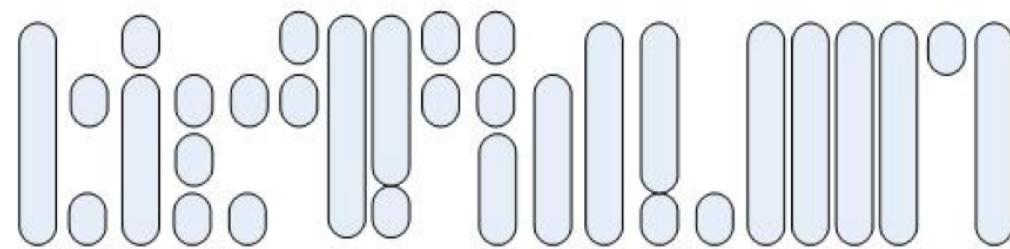
Research communities

Deep engagement to map problems onto HPC and Grid
Increased throughput, efficient



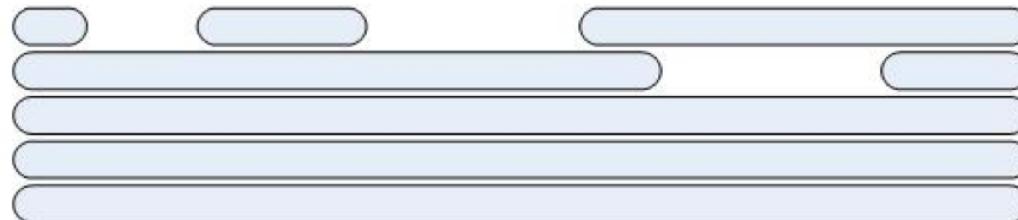
Applications and tools

Increase productivity
Improve collaboration
'Science as a service'



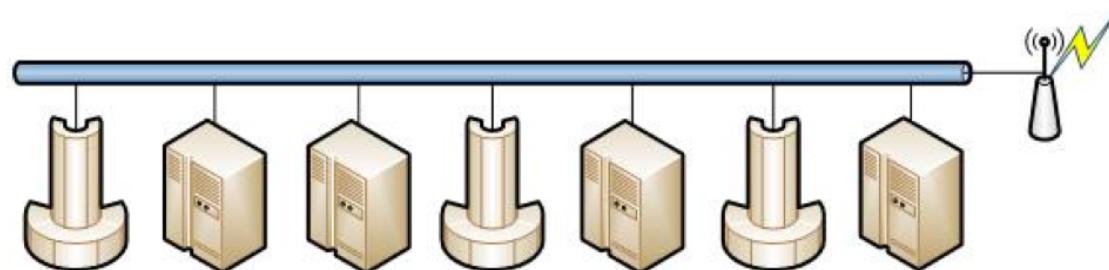
Grid Middleware

Provides simplified access
Identity, security, queues
Allows apps to 'move laterally'
Reduces risk, connects offshore



Compute and storage platforms

Shared investment
Heterogeneous architectures
Scale up science
Create new opportunities



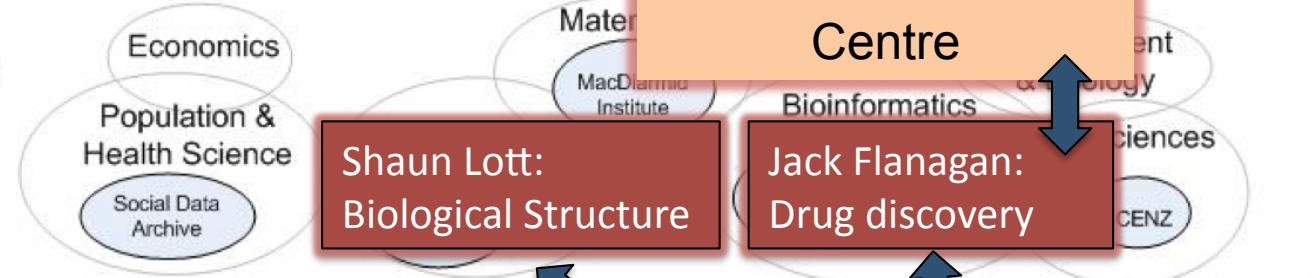
National Priorities

- Transformational Research, Science, Technology
- Science Infrastructure
- Roadmaps for Science



Science & Research communities

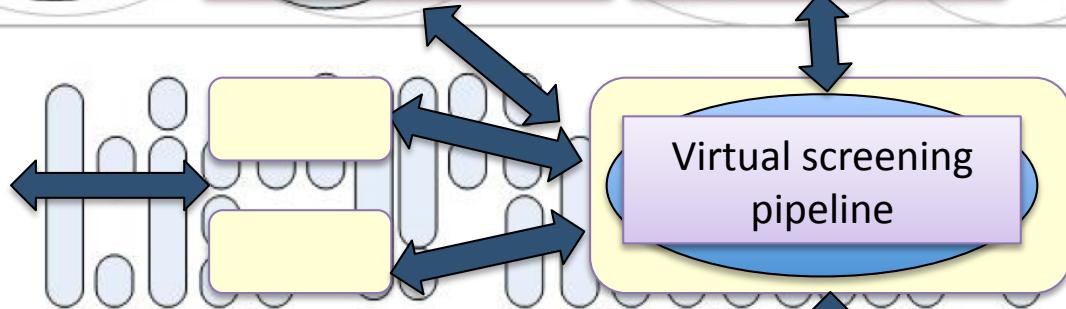
- Research Consortia, Centres, and Institutes
- Bio & Medical sciences
- Health & Population sciences
- Geo, Climate & Environmental sciences
- Maths, Physics, Engineering
- Arts & Humanities



Applications and Tools

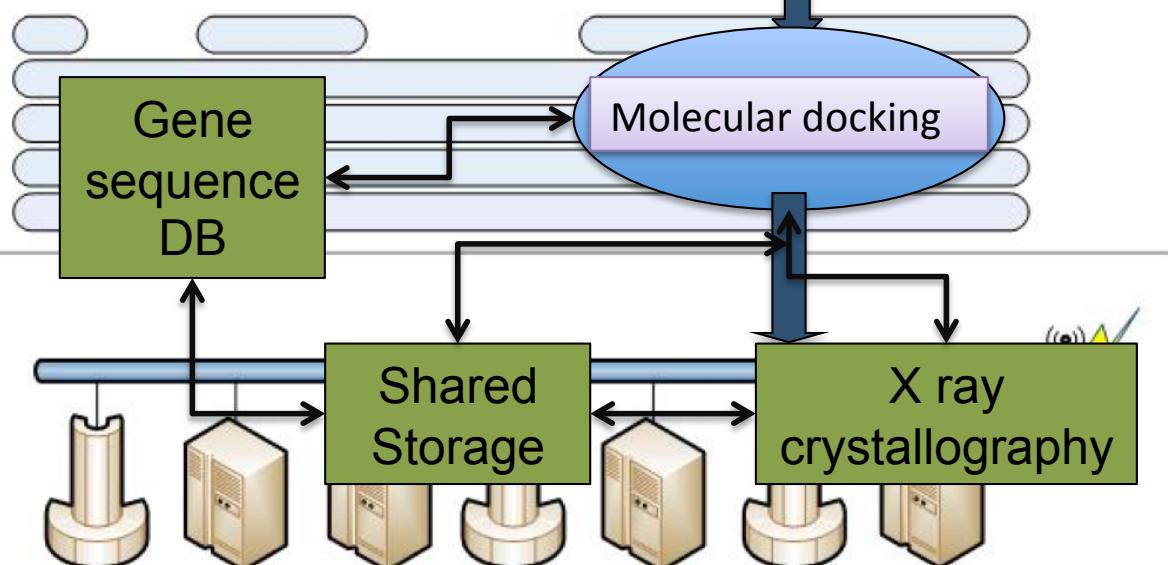
GenePattern – MIT Broad Local Analysis Modules

- Auckland
- Otago



Middleware and Platforms

Virtual servers
Shibboleth Service Provider
Databases
3D Docking libraries



Physical Resources

Data Storage
Computation Grid

- Grid Gateway Servers
- Large Storage Systems
- Backup Systems
- Video Conference Facilities
- Advanced / High Speed Research Network

How to bring NZ into eVLBI?

- Follow Australian model closely
- Demonstrate success at each step
- Leverage and expand existing strengths

National Priorities

- Transformational Research, Science, Technology
- Science Infrastructure
- Roadmaps for Science

Science & Research communities

- Research Consortia, Centres, and Institutes
- Bio & Medical sciences
- Health & Population sciences
- Geo, Climate & Environmental sciences
- Maths, Physics, Engineering
- Arts & Humanities

Applications and Tools

- Science Community Gateways
- Data Management Tools
- Open Data Publish and Consume Tools
- Scientific Applications and Libraries
- Job Submission and Management Tools
- Scientific and Research Workflow Management

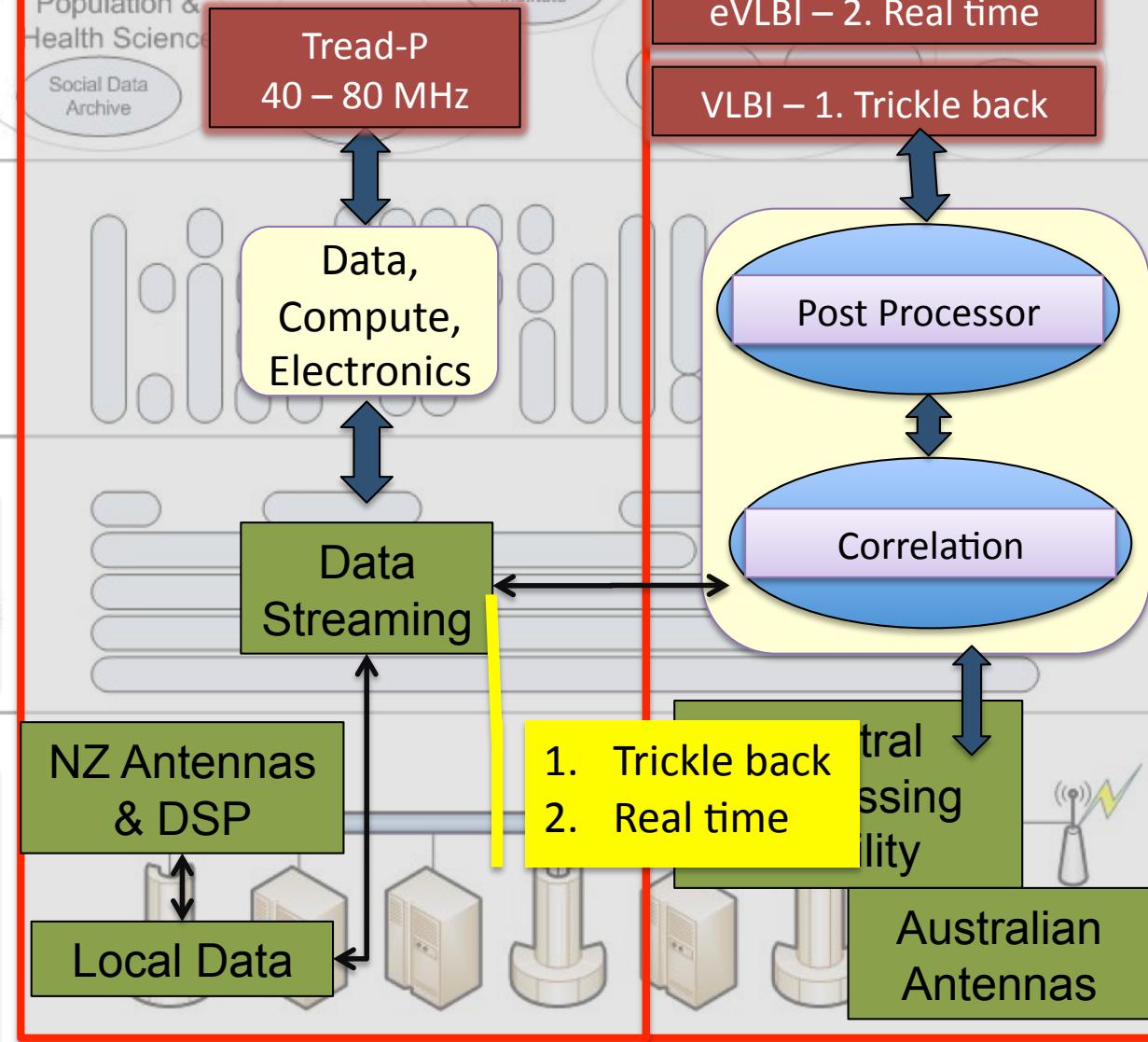
Middleware and Platforms

Data Transfer

1. Trickle back – Grid FTP
2. Real time – UDP

Physical Resources

Antenna site connection < 1 Gb/s
Trans-Tasman bandwidth
1. Trickle back ~ 200 Mb/s
2. Real time ~ 1 Gb/s dedicated
Antenna Storage ~ 10 – 20 TB



Australian NZ links

Australia

- SKA



- ARCS



- AARNET



New Zealand

- AUT IR&SR
- SKA R&D Consortium
- BeSTGRID



- REANNZ





End

Questions / Discussion

Nick Jones
Director, BeSTGRID
Co-Director eResearch, Centre for eResearch
University of Auckland
n.jones@auckland.ac.nz

www.bestgrid.org
[@bestgrid](https://twitter.com/bestgrid)