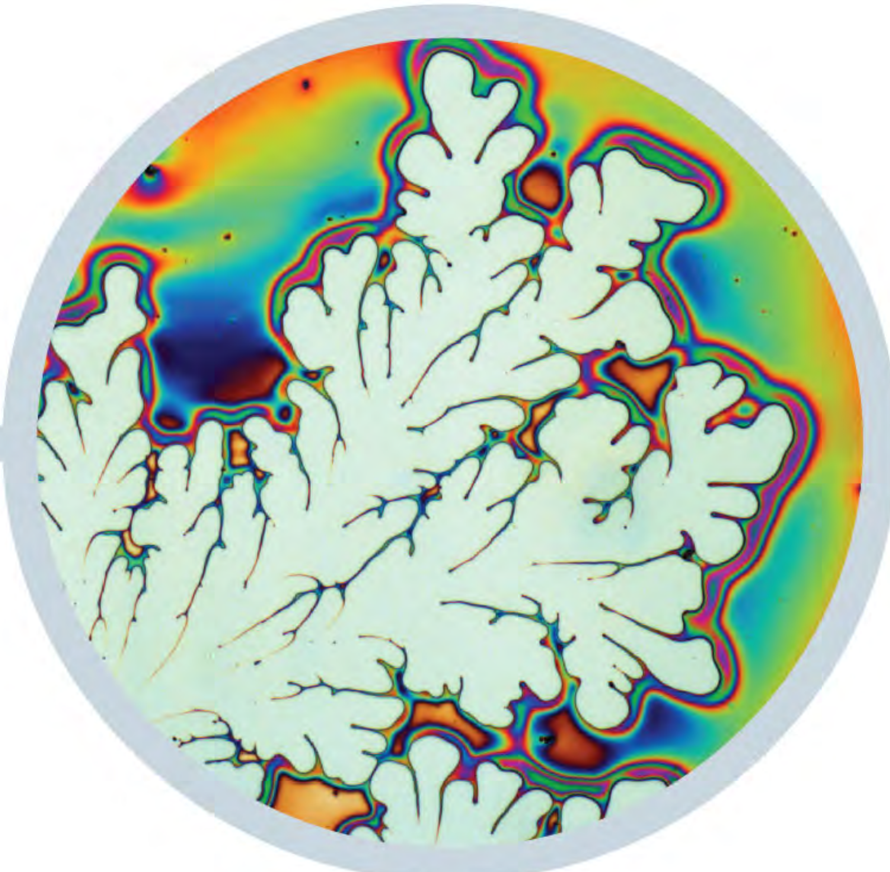


NeSI
New Zealand eScience Infrastructure

Computing Facilities for NZ Researchers

National HPC Roadshow 2012

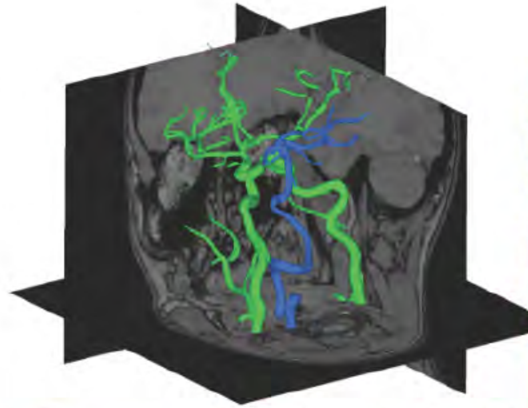
Nick Jones, Director

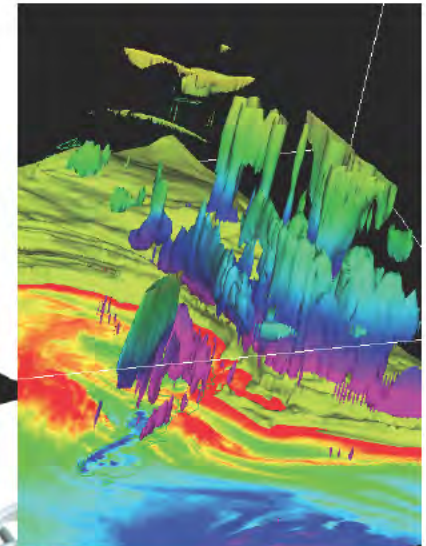
A partnership between NZ's research institutions delivering advanced computational services for leading edge discovery

ministry of science + innovation · The University of Auckland New Zealand · University of Canterbury · Landcare Research Manaaki Whenua · NIWA Taihoro Nukurangi · University of Otago

NeSI, a national research infrastructure

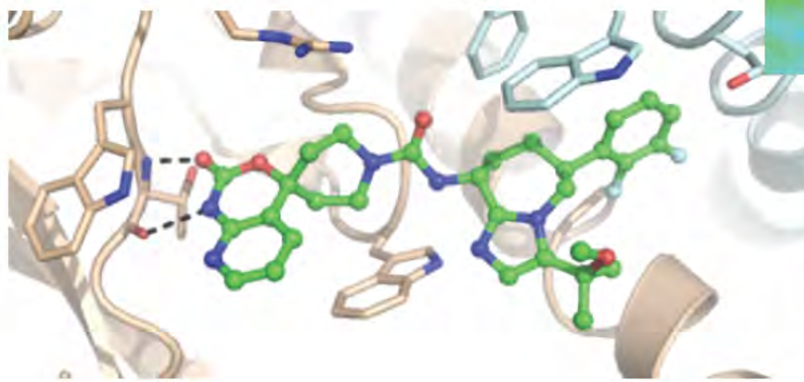
- ✓ High Performance Computing (HPC)
- ✓ & eScience services
- ✓ … for all NZ researchers



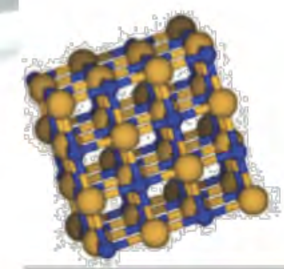


NeSI HPC Facilities

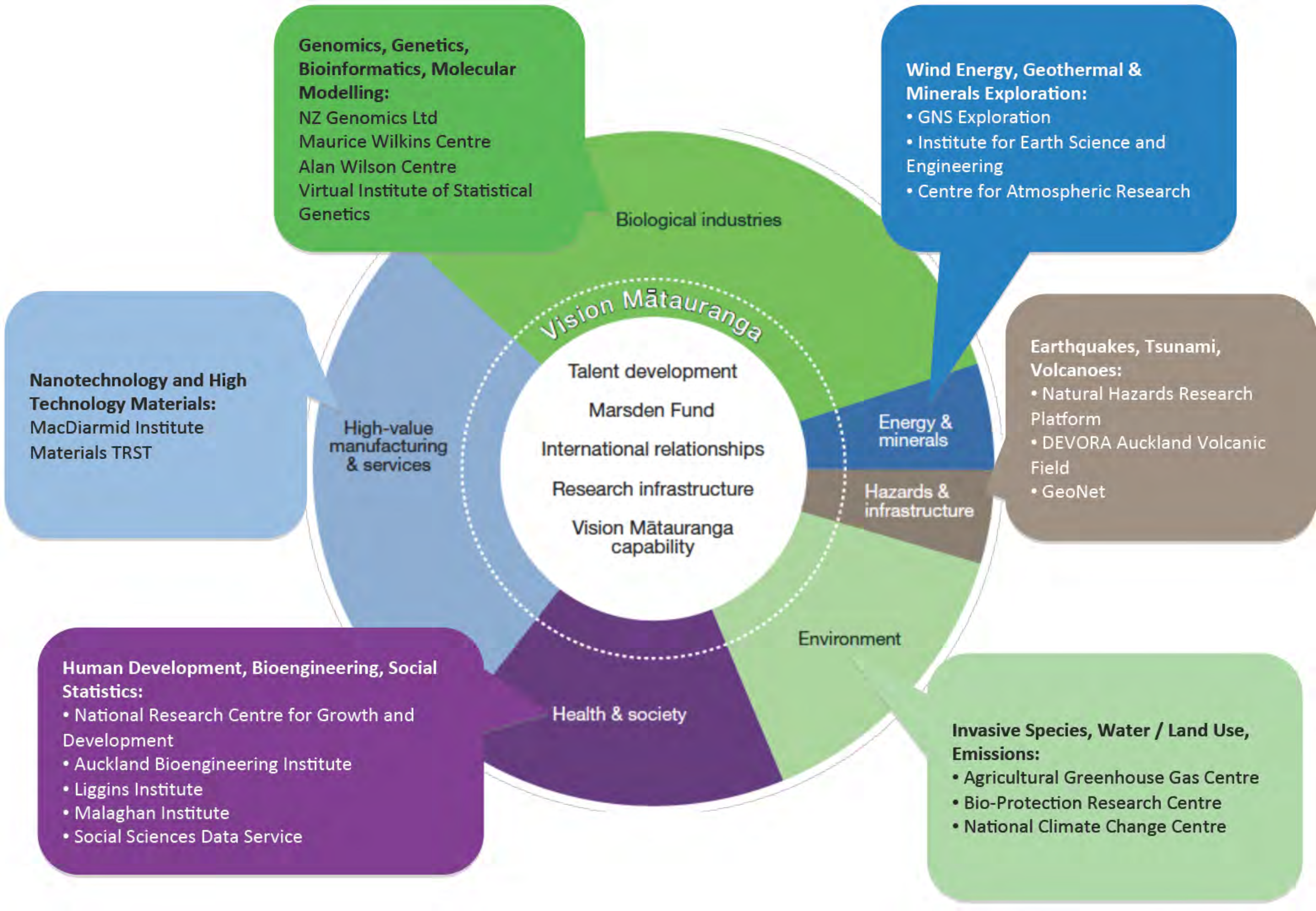
NeSI's Roadmap



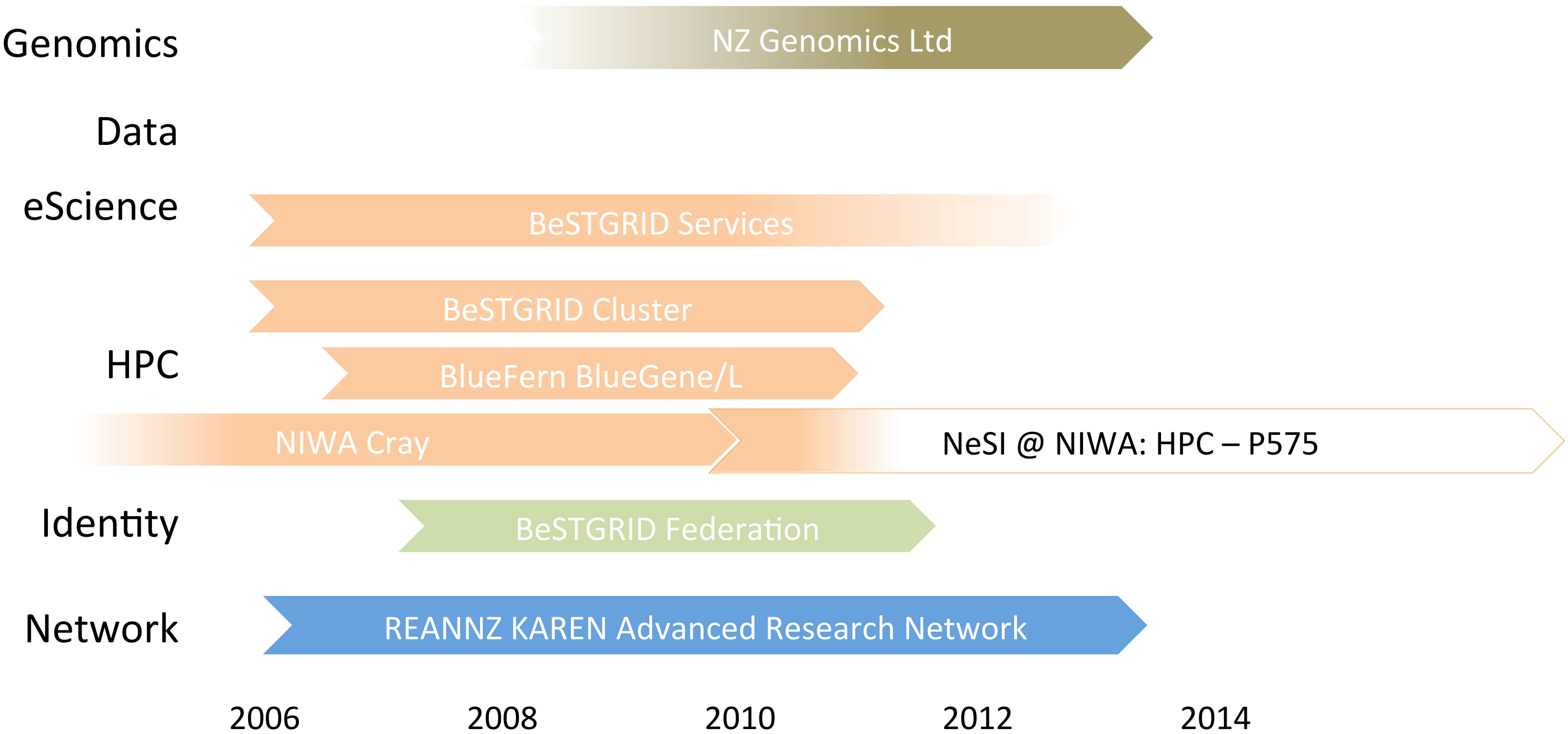
Progress Report



NeSI
New Zealand eScience Infrastructure


Genomics, Genetics, Bioinformatics, Molecular Modelling:
NZ Genomics Ltd
Maurice Wilkins Centre
Alan Wilson Centre
Virtual Institute of Statistical Genetics
Wind Energy, Geothermal & Minerals Exploration:
• GNS Exploration
• Institute for Earth Science and Engineering
• Centre for Atmospheric Research
Biological industries
Vision Mātauranga
Nanotechnology and High Technology Materials:
MacDiarmid Institute
Materials TRST
High-value manufacturing & services
Talent development
Marsden Fund
International relationships
Research infrastructure
Vision Mātauranga capability
Energy & minerals
Hazards & infrastructure
Earthquakes, Tsunami, Volcanoes:
• Natural Hazards Research Platform
• DEVORA Auckland Volcanic Field
• GeoNet
Environment
Health & society
Human Development, Bioengineering, Social Statistics:
• National Research Centre for Growth and Development
• Auckland Bioengineering Institute
• Liggins Institute
• Malaghan Institute
• Social Sciences Data Service
Invasive Species, Water / Land Use, Emissions:
• Agricultural Greenhouse Gas Centre
• Bio-Protection Research Centre
• National Climate Change Centre

Research e-Infrastructure Roadmap - 2010



NeSI
New Zealand eScience Infrastructure

How did NeSI arise?

October 2010 an Investment Case entitled:

> "National eScience Infrastructure (NeSI) High Performance Computational Platforms and Services for NZ's Research Communities"

- Submitted to the Minister of Research, Science & Technology
- Prepared by a Working Group of representatives from: UoA, UC, UoO, Landcare, AgResearch & NIWA (under an indept. Chair)

Investment Case asserted that:

- HPC and related eScience infrastructure are indispensable components of modern science, and are having a major impact on almost every branch of research;
- By taking a sector approach, more efficient coordination and cooperation would be achieved, leading to strategically targeted investment in HPC;
- Thereby providing international-scale HPC to a wide range of communities and disciplines.

Formulated following a "Needs Analysis" during 2010

NeSI
New Zealand eScience Infrastructure

What “we” said about our HPC needs…

In 2010, NZ Researchers were surveyed to determine their existing and anticipated HPC requirements. We (~194 of us) said (e.g.):

Processors to run a code	< 10	10 - 100	100 - 1000	>10,000
In 2010	40%	39%	15%	5%
In 2015	12%	32%	38%	17%

File Space per experiment	< 100GB	1 - 10 TB	100 TB	>1PB
In 2010	64%	35%	0%	0.6%
In 2015	27%	58%	13%	3%

Off-Site Data Transfers/day	< 100MB	1GB	1TB	>10TB
In 2010	33%	42%	21%	4%
In 2015	17%	28%	33%	23%

Why hadn't this already happened?

Coordination failure..

a collective failure to coordinate investment decisions for the greater benefit (value, efficiency)

We see a coordination failure (coordination of capital and resources) leading to HPC facilities that don't achieve the scale required for demanding research applications

Funding strategy aligns Opex and Capex – Crown and collaborators fund both, clarifying intent

Value to Crown

Why do we need Government investment?

- **overcome coordination failure**
- **reach scale and maturity**
- **efficient investment in critical infrastructure**

Value to Institutions

Why co-invest into national infrastructure?

- **new investment into institutions and sector**
- **create scalable mature facilities**
- **trust and control ensure good fit to needs**

What did we decide on?

Shared facilities

shared facilities accessible nationally

National Access scheme

single Access Policy for all HPC facility access

National Team

single Team nationally sharing practices and expertise

NeSI

New Zealand eScience Infrastructure

Investment partners

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

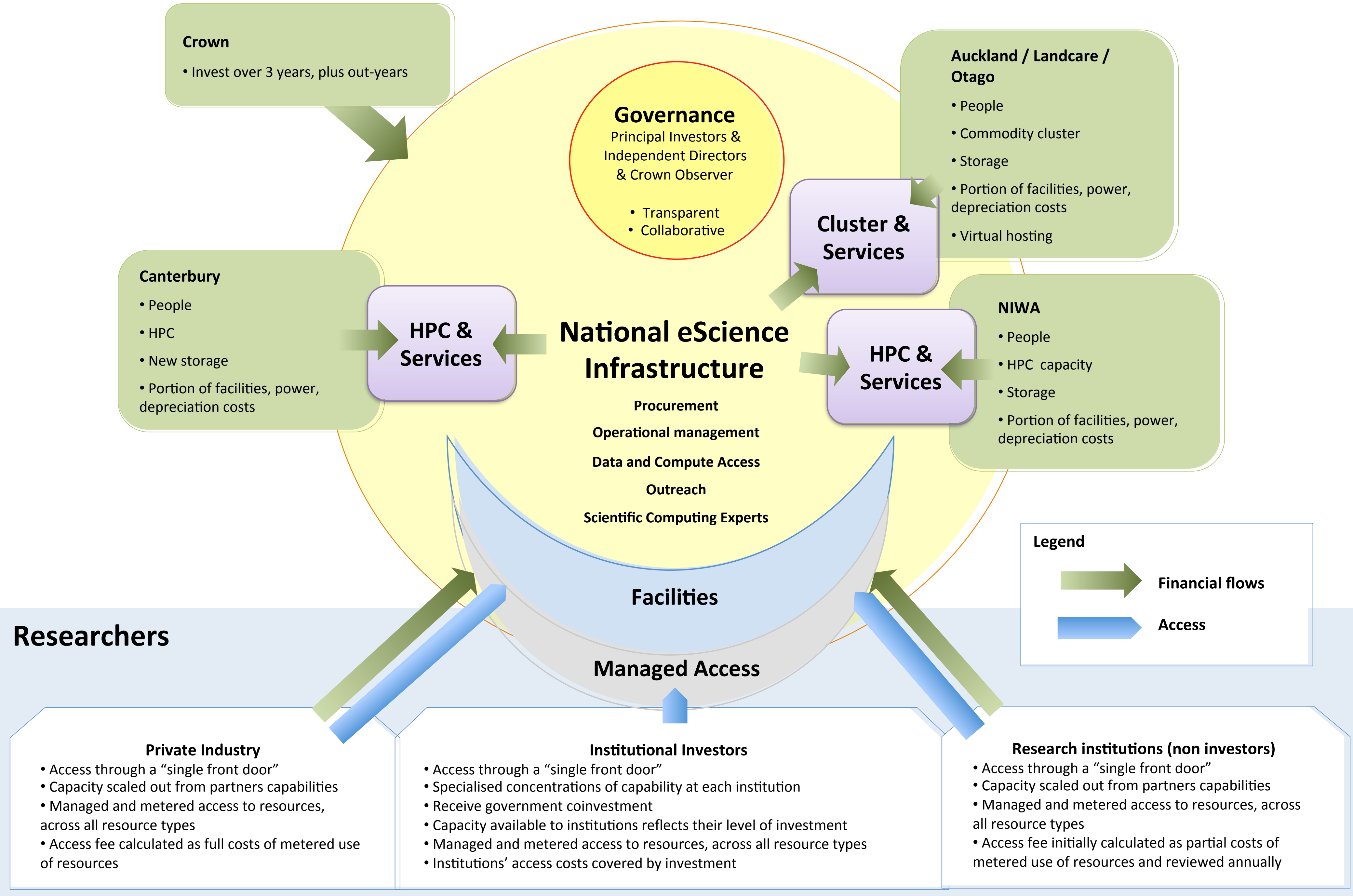
UC UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

NIWA
Taihoro Nukurangi

UNIVERSITY of OTAGO
Te Whare Wānanga o Otāgo
NEW ZEALAND

Landcare Research
Manaaki Whenua

Ministry of Business, Innovation & Employment



Crown
• Invest over 3 years, plus out-years
Governance
Principal Investors & Independent Directors & Crown Observer
• Transparent
• Collaborative
Auckland / Landcare / Otago
• People
• Commodity cluster
• Storage
• Portion of facilities, power, depreciation costs
• Virtual hosting
Cluster & Services
Canterbury
• People
• HPC
• New storage
• Portion of facilities, power, depreciation costs
HPC & Services
National eScience Infrastructure
Procurement
Operational management
Data and Compute Access
Outreach
Scientific Computing Experts
HPC & Services
NIWA
• People
• HPC capacity
• Storage
• Portion of facilities, power, depreciation costs
Facilities
Managed Access
Legend
Financial flows
Access
Researchers
Private Industry
• Access through a "single front door"
• Capacity scaled out from partners capabilities
• Managed and metered access to resources, across all resource types
• Access fee calculated as full costs of metered use of resources
Institutional Investors
• Access through a "single front door"
• Specialised concentrations of capability at each institution
• Receive government coinvestment
• Capacity available to institutions reflects their level of investment
• Managed and metered access to resources, across all resource types
• Institutions' access costs covered by investment
Research institutions (non investors)
• Access through a "single front door"
• Capacity scaled out from partners capabilities
• Managed and metered access to resources, across all resource types
• Access fee initially calculated as partial costs of metered use of resources and reviewed annually

How is NeSI Funded?

There are three anchor partners:

- The University of Auckland (receives Crown funding of ~$2.2M pa);
  - including two associates (Associate Investors): Landcare and University of Otago.
- University of Canterbury (receives Crown funding of ~$2.6M pa);
- NIWA (receives Crown funding of ~$1.0M pa);

Crown investment is $27M over 3 years
Institutional investment is $21M over 3 years

- $15.4M in Capital Expenditure & $5.6M in Operational Expenditure

Which provides:

- 4 × HPC systems at Universities of Auckland & Canterbury, and at NIWA
- NeSI Directorate at Auckland (5 staff positions + admin) 5.5 FTEs
- Systems Engineers ~ 5.5 FTEs, Services & Application Engineers ~ 5.5 FTEs, Site Management 1.8 FTEs
- HPC specialist Scientific Programmers ~5.7 FTEs

And access is allocated in the following way:

- anchor partners (called Collaborators) reserve 60% of the HPC capacity for their purpose
- and 40% is available, by any researcher at a public research institution in NZ with an active peer-reviewed grant

NeSI
New Zealand eScience Infrastructure

What are NeSI's objectives?

1. Creating an advanced, scalable computing infrastructure to support New Zealand's research communities;
   - i.e. International scale as opposed to institutional/department scale.

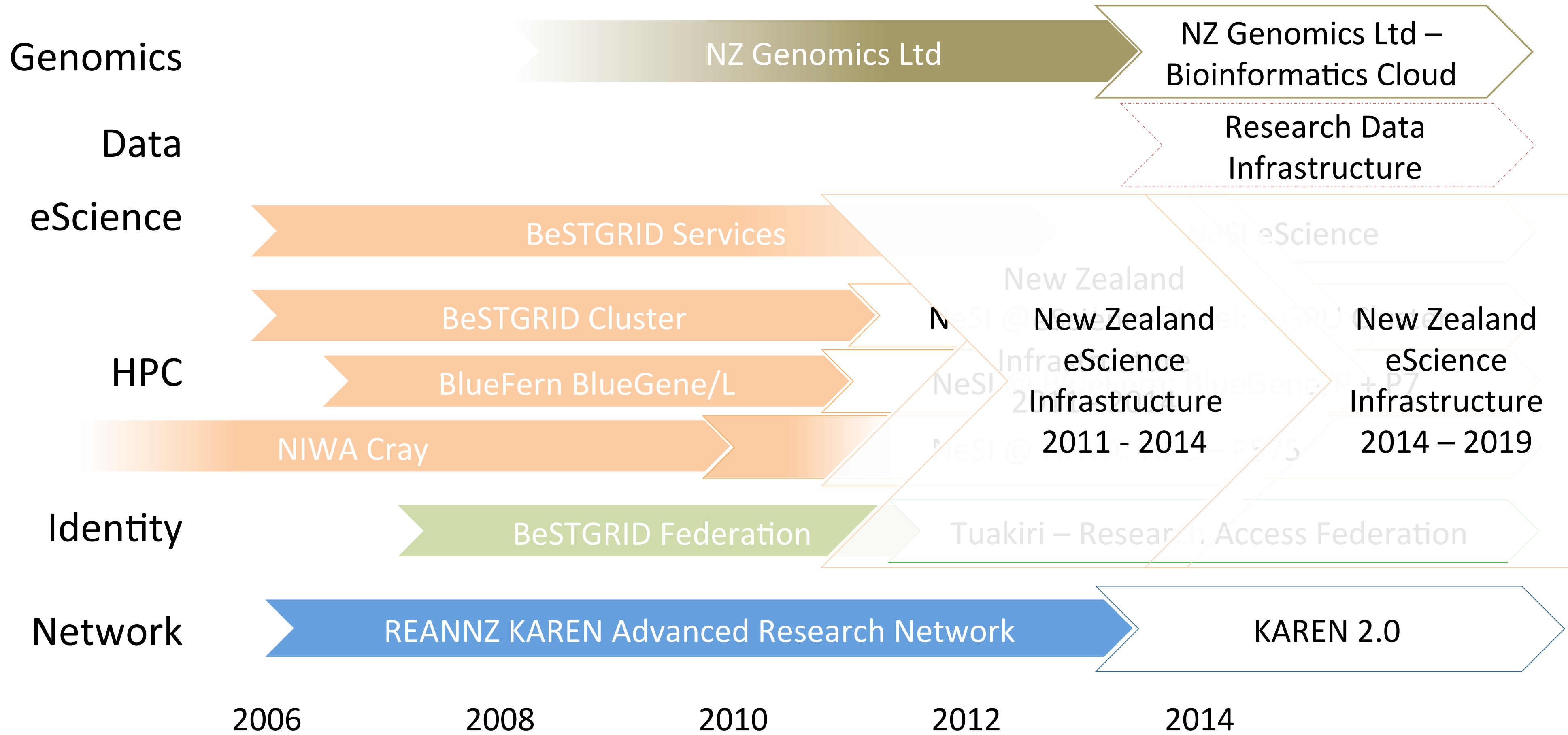
2. Providing the grid middleware, research tools and applications, data management, user-support, and community engagement needed for the best possible uptake (of HPC);
   - i.e. enable efficient use of and easy access to these systems.

3. Encouraging a high level of coordination and cooperation within the research sector;
   - i.e. fit the science to the HPC as opposed to "my institution's resources" (shared services and resources).

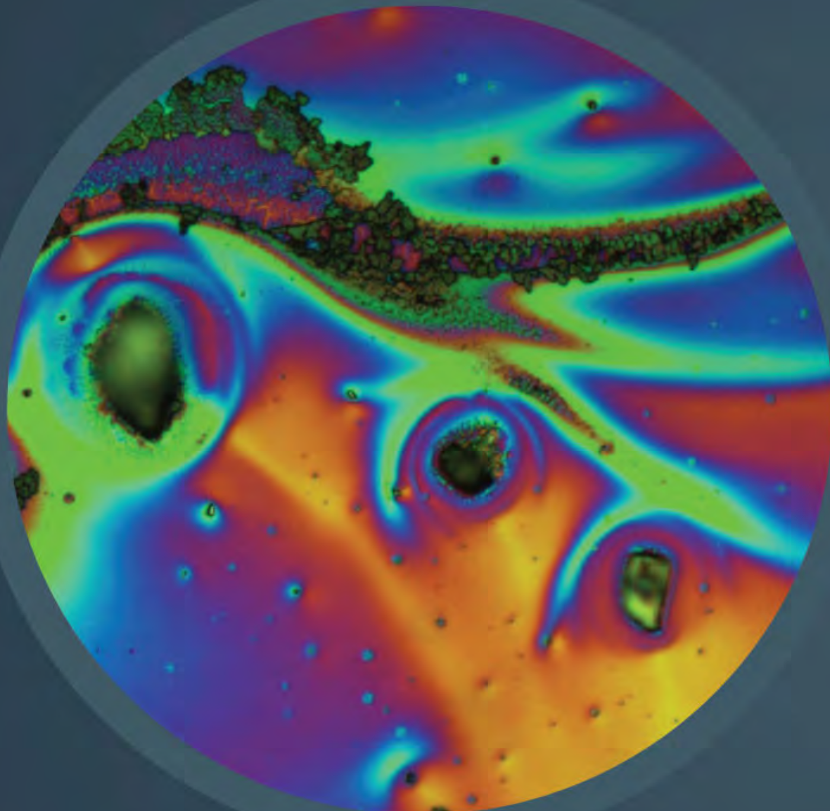
4. Contributing to high quality research outputs from the application of advanced computing and data management techniques and associated services, which support the Government's published priorities for science.
   - i.e. Its all about better science to underpin national outcomes.


LAX
SYD
WRK
AKL
NSH
MTA
HLT
TAU
ROT
NPL
GIS
NPE
WAG
MUP
PNR
NLS
WLG
AVL
POR
CHC
LCN
TPO
DUD
IVM
INV
Single PoP in region
Multiple PoPs in region
10Gb/s
1Gb/s

Research e-Infrastructure Roadmap - 2012



NeSI
New Zealand eScience Infrastructure



NeSI
New Zealand eScience Infrastructure

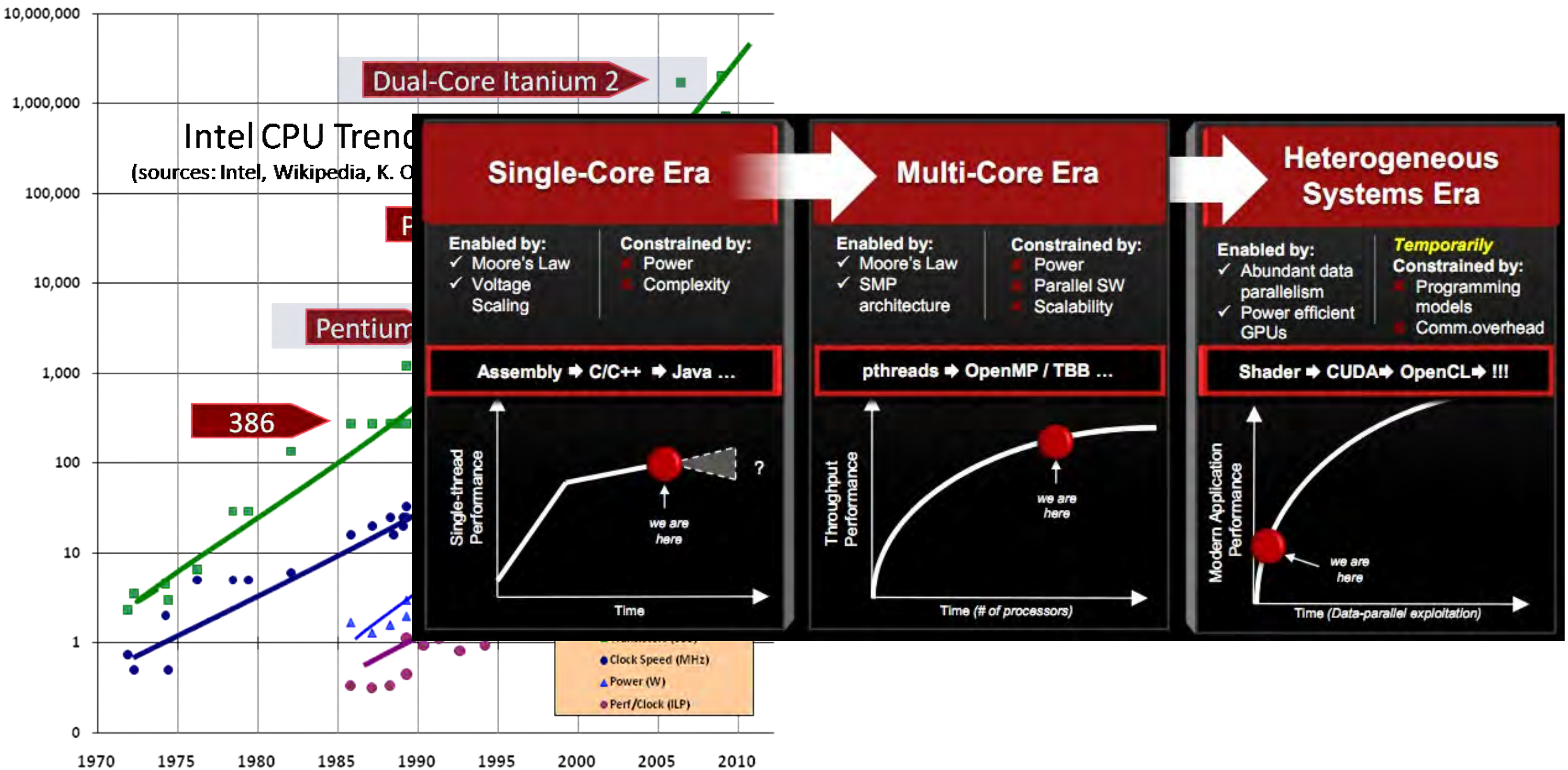
An Introduction to Supercomputing

Dr. Michael J. Uddstrom

Director, NIWA High Performance Computing Facility

michael.uddstrom@niwa.co.nz

Processor Trends and Limits



What is a Supercomputer (or HPC)?

- There are two "basic" types:
  - Capability (aka Supercomputers): provide the maximum computing power available to solve large problems: the emphasis is on problem size (large memory, lots of cores) . e.g:
    - IBM p775/p7 & p575/p6, Cray XK6, IBM BG/Q & BG/P
  - Capacity: typically use efficient cost-effective computing components: the emphasis is on throughput (dealing with loads larger than a single PC/small cluster), e.g:
    - IBM iDataPlex, HP Cluster Platform *n*000

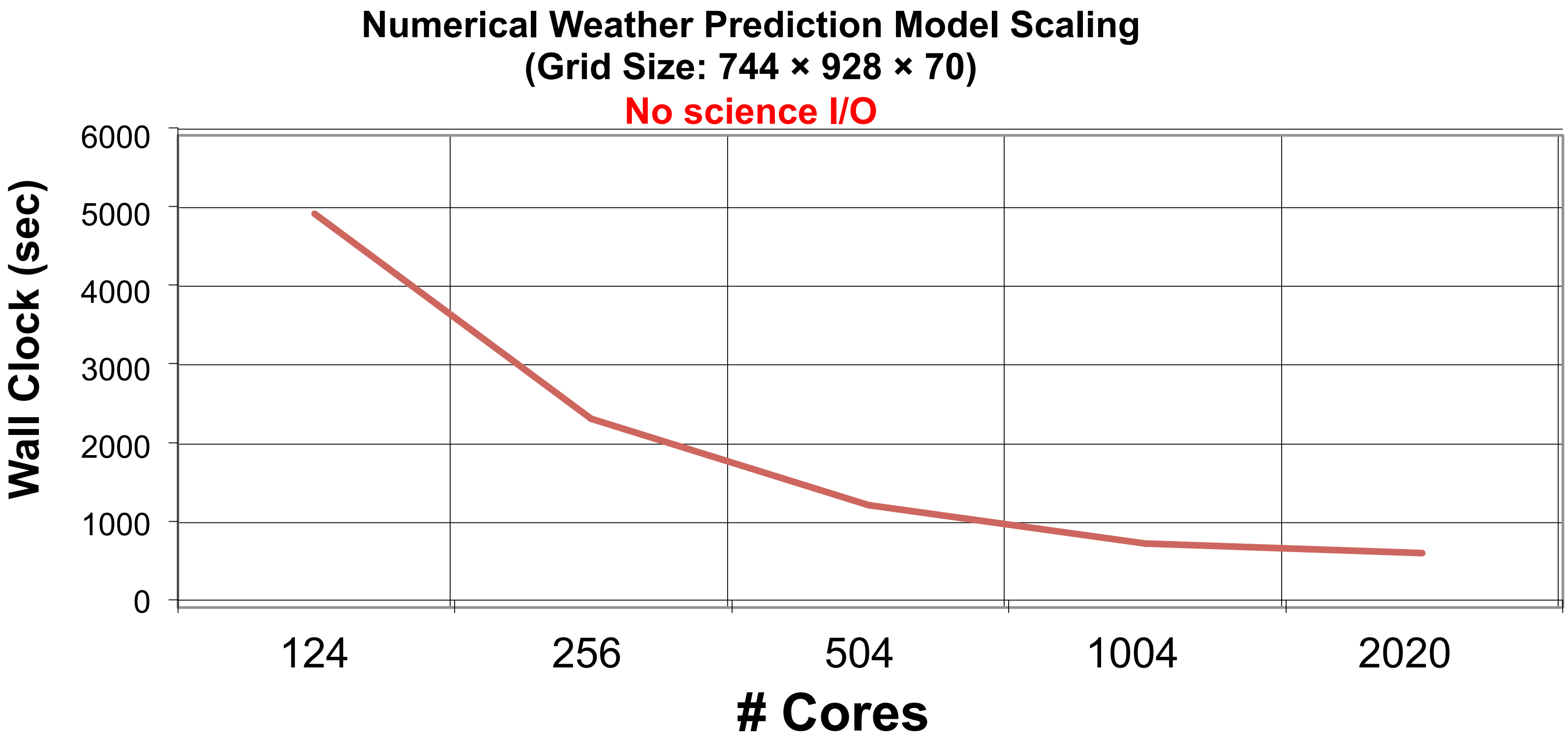
- The essential differences between Capability & Capacity systems are:
  - the interconnect fabric performance;
  - the processor performance, and
  - reliability (i.e. resiliency to component failure).
- Supercomputers have high efficiency:
  - Efficiency = sustained-performance / peak-performance (%).

NeSI
New Zealand eScience Infrastructure

What types of HPC do we need?

- It depends on the problem (and the data locality)
- Is it "Embarrassingly Parallel" (EP)?
  - This means the problem can be split into independent tasks, with each sent to a different processor:
    - Eg: Image rendering, classification, Monte Carlo calculations, BLAST, etc…
- If not EP, then is it highly-scalable?
  - This means the problem does not place high demands on processor performance – because the coupling between processors is relatively "loose" and you can use very many:
    - Eg: materials codes, Schrodinger's equation, DCA++, LSMS, NWChem…
- If not EP and not highly scalable then
  - This means that the problem will place high demands on processor performance and on the interconnect between processors:
    - Examples: numerical weather and climate prediction, Variational data assimilation, combustion, seismology, POP,

NWP: The Computational Challenge



- Speedup on 2020 4.7GHz cores: 7.4
- Relative number of cores : 16.3 $\Rightarrow$ 45% of possible scaling

Scalability: Amdahl's Law

- Amdahl's law states that if *P* is the proportion of a program that can be made parallel, and (1 − *P*) is the proportion that cannot be parallelized, then the maximum speedup that can be achieved by using *N* processors is:

$$\frac{1}{(1-P)+\frac{P}{N}}$$ In the limit as *N* becomes large → $$\frac{1}{(1-P)}$$

- In practice, performance to price ratio falls rapidly as *N* is increased once there is even a small component of (1 − *P*).

e.g. if P=0.99, & N=64, then speedup = 39.3 (perfect = 64) √
if P=0.99, & N=1024, then speedup = 91.2 (perfect = 1024) ×
if P=0.90, & N=1024, then speedup = 9.9 (perfect = 1024) ××!

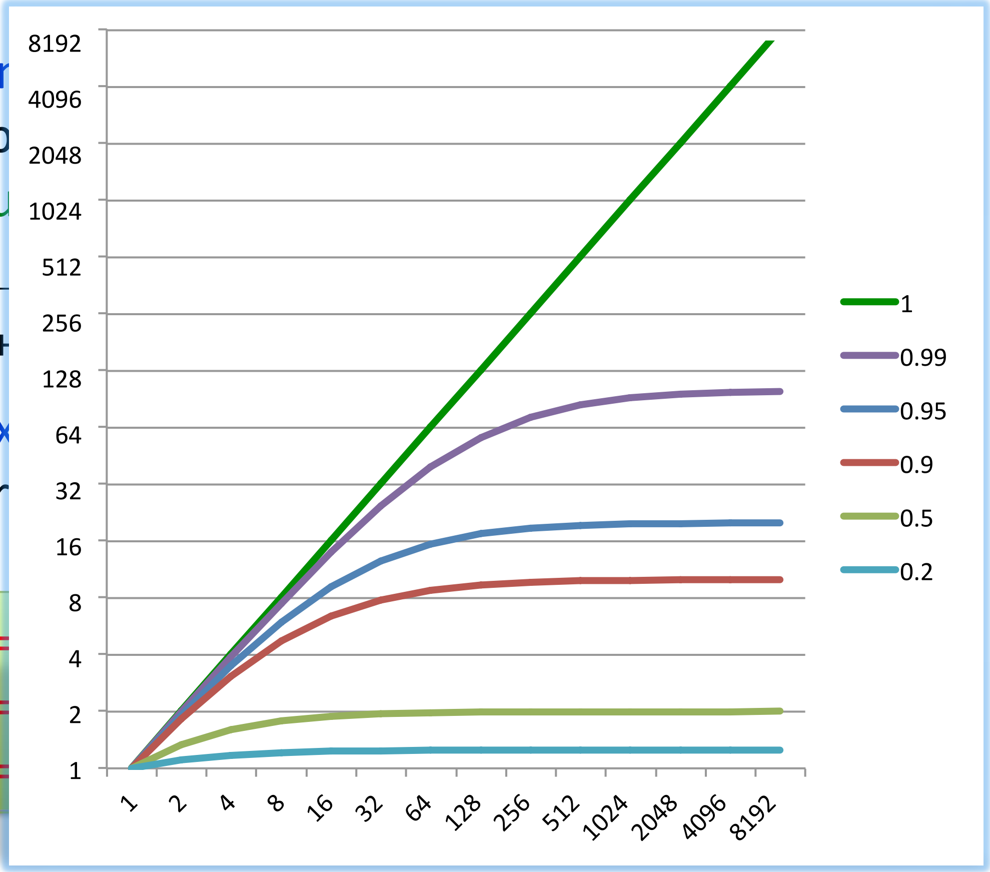
The Limits to Scalability: Amdahl's Law

Amdahl's law states that if *P* is the proportion
and (1 – *P*) is the proportion that cannot be p
maximum speedup that can be achieved by u

$$\frac{1}{(1-P)+}$$

- In the limit, as *N* tends to infinity, the max practice, performance to price ratio falls r even a small component of (1 – *P*).

e.g. if P=0.99, & N=64, then speedup =
if P=0.99, & N=1024, then speedup =
if P=0.90, & N=1024, then speedup =

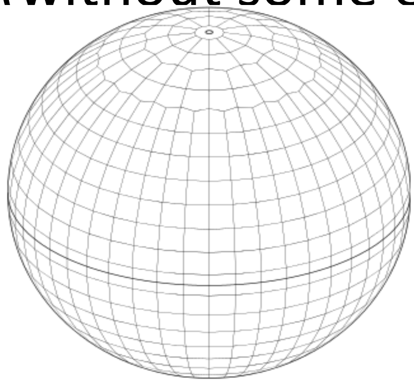


It's a Challenge… but this is "our" future!!

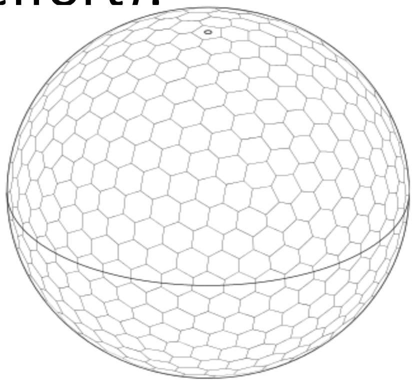
NeSI
New Zealand eScience Infrastructure

Load Balancing

- Consider a global atmospheric model:
  - Decompose the grid across the processors (chess-board like)
  - ~50% of the planet is in darkness all of the time ⇒ only longwave radiation calculations are needed;
  - ~50% of the planet is in sun light all of the time ⇒ both longwave and shortwave radiation calculations are needed;
  - P is going to be less than optimal (without some effort)!
- Then there is the matter of the poles.
- But… HPC means we can at last solve really large problems!



skipped lat-lon



hexagonal-icosahedral

Applications running on Jaguar at ORNL (2011)

Domain area	Code name	Institution	# of cores	Performance	Notes
Materials	DCA++	ORNL	**213,120**	**1.9 PF**	2008 Gordon Bell Prize Winner
Materials	WL-LSMS	ORNL/ETH	223,232	1.8 PF	2009 Gordon Bell Prize
Chemistry	NWChem	PNNL/ORNL	224,196	1.4 PF	
Nanoscience	OMEN	Duke	222,720	> 1 PF	
Biomedical	MoBo	GaTech	196,608	780 TF	2010 Gordon Bell Prize Winner
Chemistry	MADNESS	UT/ORNL	140,000	550 TF	
Materials	LS3DF	LBL	147,456	442 TF	
Seismology	SPECFEM3D	USA (multiple)	149,784	165 TF	
Combustion	S3D	SNL	147,456	83 TF	
Weather	WRF	USA (multiple)	**150,000**	**50 TF**	

Highly scalable ⇒ high fraction of Peak Performance

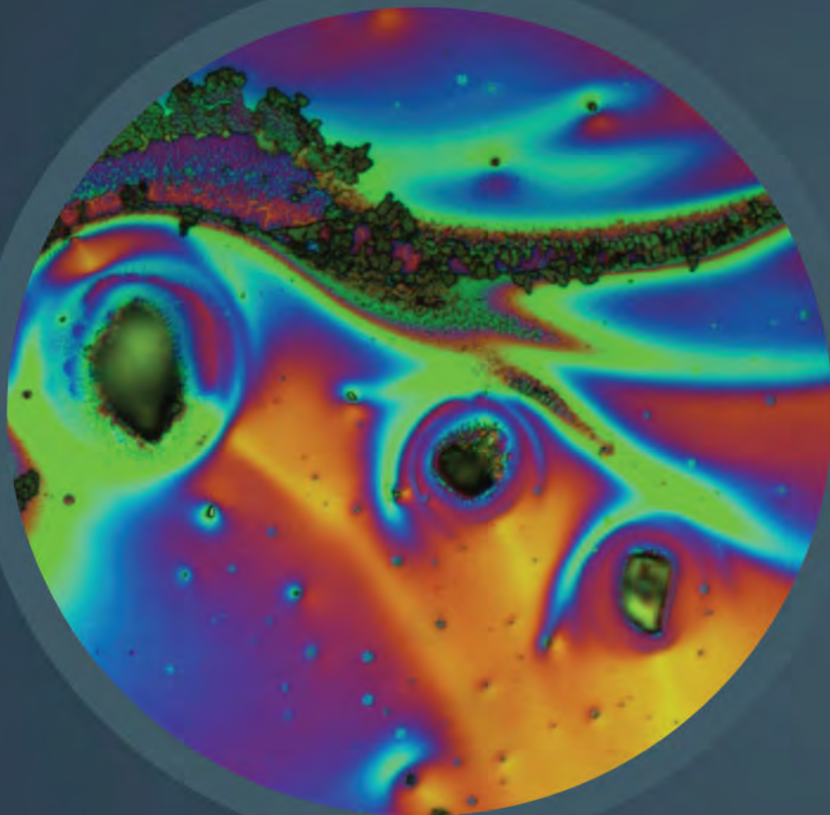
Tightly Coupled Problem ⇒ small fraction of Peak Performance
Just 3.6% of DCA++ Performance

NeSI HPCs in Summary

- University of Auckland:
  - IBM iDataPlex Intel processor Cluster, large node memory + some exotic hardware (i.e. GPGPUs) (Pan)
    - General purpose HPC cluster
    - Optimised for EP and Highly Scalable problems;
- University of Canterbury:
  - IBM BlueGene/P Supercomputer
    - Optimised for EP and Highly Scalable problems;
  - IBM p755/POWER7 cluster
    - General purpose / capability HPC cluster
  - IBM IDataPlex Visualisation Cluster
- NIWA High Performance Computing Facility:
  - IBM p575/POWER6 Supercomputer (FitzRoy)
    - Optimised for tightly coupled (large) problems
    - Operational/production-ready (i.e. IBM Support arrangements)

Other Points about HPCs

- Provide ultimate compute performance;
- Provide ultimate I/O performance (think scalability):
  - Parallel file systems (read / write simultaneously from/to multiple cores);
  - Very high I/O throughput (e.g. multi- GB/s).
- Data management matters – generate vast amounts of data:
  - Hierarchical Storage Management systems.
- Jobs run in a batch-processing environment:
  - There will be time and resource limits on queues;
  - Need to be able to "self checkpoint" – i.e. write restart files;
  - It's a shared resource.
- Often characterised by small number of large users.



NeSI
New Zealand eScience Infrastructure

NeSI HPC Facilities

A mix of specialised HPC facilities to meet varied research requirements

NeSI
New Zealand eScience Infrastructure

If its about HPC….what is an HPC?

There are two "basic" types:

- Capability (aka Supercomputer): provide the maximum computing power available to solve large problems: emphasis is on problem size (large memory, lots of CPUs) . e.g:
  - IBM p775/p7 & p575/p6, Cray XK6, IBM BG/Q & BG/P

- Capacity: typically use efficient cost-effective computing power: the emphasis is on throughput (dealing with loads larger than a single PC/small cluster), e.g:
  - IBM iDataPlex, HP Cluster Platform *n*000

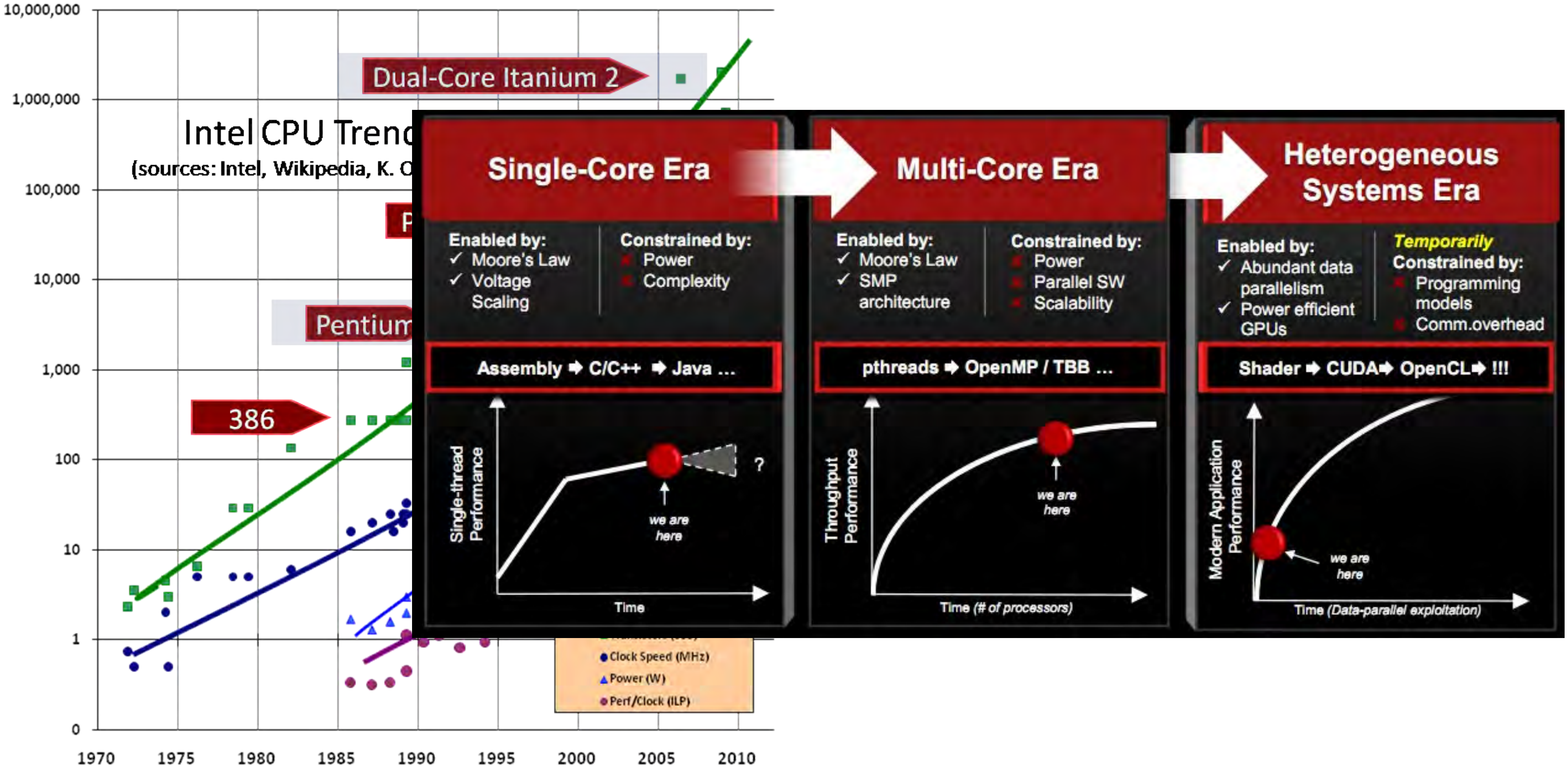
The essential differences are:

- the interconnect fabric performance
- the processor performance
- reliability (i.e. resiliency to component failure)

Supercomputers have high efficiency:

- Efficiency = sustained-performance/peak-performance (%)

Processor Trends and Limits



NeSI
New Zealand eScience Infrastructure

The Limits to Scalability: Amdahl's Law

Amdahl's law states that if *P* is the proportion

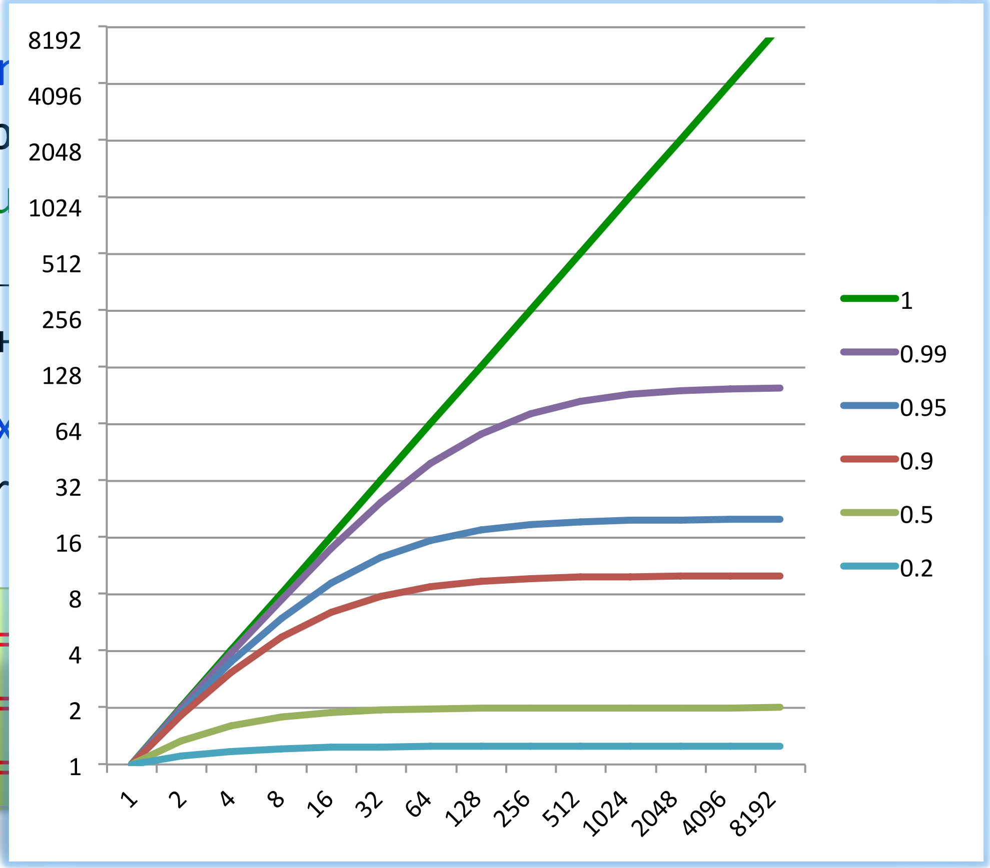
and (1 − *P*) is the proportion that cannot be p

maximum speedup that can be achieved by u

$$\frac{1}{(1-P)+}$$

- In the limit, as *N* tends to infinity, the max practice, performance to price ratio falls r even a small component of (1 − *P*).

e.g. if P=0.99, & N=64, then speedup =
if P=0.99, & N=1024, then speedup =
if P=0.90, & N=1024, then speedup =



It's a Challenge… but this is "our" future!!

What types of HPC do we need?

It depends on the problem (and the data locality)!

Is it "Embarrassingly Parallel" (EP)?

- This means the problem can be split into independent tasks (P≈1), with each sent to a different processor:
  - Eg: Image rendering, classification, Monte Carlo calculations, BLAST, genetic algorithms, etc…

If not EP, then is it highly-scalable?

- This means the problem does not place high demands on processor performance – because the coupling between processors is relatively "loose" (P>0.999) and you can use very many:
  - Eg: materials codes, Schrodinger's equation, DCA++, LSMS, NWChem…

If not EP and not highly scalable – is it tightly coupled?

- This means that the problem will place high demands on processor performance and on the interconnect between processors: (P > ~0.9)
  - Examples: numerical weather and climate prediction, Variational data assimilation, combustion, seismology, POP

NeSI HPCs in Summary

The University of Auckland, Centre for eResearch:

- IBM iDataPlex Intel processor Cluster (incl. GPGPUs)
  - General purpose HPC cluster
  - Optimised for EP and Highly Scalable problems;

University of Canterbury, BlueFern:

- IBM BlueGene/P Supercomputer
  - Optimised for EP and Highly Scalable problems;
- IBM p755/POWER7 cluster
  - General purpose / capability HPC cluster
- IBM IDataPlex Visualisation Cluster

NIWA, FitzRoy High Performance Computing Facility:

- IBM p575/POWER6 Supercomputer (FitzRoy)
  - Optimised for tightly coupled (large) problems
  - Operational/production-ready (i.e. IBM Support arrangements)

NeSI Auckland / Centre for eResearch

- Pan: IBM iDataPlex
  - 912 (→ 1904) Intel cores
    - 76 nodes (Westmere: 12 cores / SMP node, 96 GB/node) (64 bit)
      - 62 nodes (Sandy Bridge: 16 cores / SMP node, 128 GB/node)
    - 12 GPGPU nodes (24 Tesla M2090: 2 per node)
    - InfiniBand Interconnect fabric
    - Red Hat Linux Operating System

- Storage: General Parallel File System (GPFS)
  - 200 TB SAN user disk

NeSI Auckland / Centre for eResearch

- Applications:
  - Preinstalled (some licensing may be required):
    - Math: Gap, Magma, Matlab, Mathematica, R
    - BioInformatics: BLAST, BEAST, beagle, PhyML, MrBayes, BEDtools, Bamtools, Bowtie, Clustal Omega, Cufflinks, FastQC, FASTX Toolkit
    - Computational Chemistry: Gaussian, Gromacs, AMBER, Orca, VASP
    - Engineering: Ansys, Abaqus, OpenFOAM
    - Meteorology: WRF, WPS

NeSI Canterbury / BlueFern SuperComputer

- Foster: IBM BlueGene/P
  - 2048 nodes (4 cores / SMP node)
    - 8192 × 0.85 GHz PowerPC 450 cores (32 bit)
    - 4 GB memory / node
    - 13.6 GFLOPS / node, 23 TFLOPS (peak)
  - 3 Dimensional Torus Interconnect
- Kerr: IBM p755/POWER7 Cluster
  - 13 nodes (32 cores / SMP node)
    - 416 × 3.3 GHz POWER7 cores (64 bit)
    - 1.7 TB Memory (128 GB / node)
    - 769 GFLOPS/node, 10 TFLOPS (peak)
  - Infiniband Interconnect
  - AIX & Linux Operating Systems
- Popper: IBM iDataPlex Visualisation Cluster
  - 5 GPGPU nodes (8 cores / SMP node, 10 Tesla M2070Q: 2 per node)

- Storage: General Parallel Filesystem (GPFS)
  - 180 TB SAN user disk
  - 1 PB Automatic Tape Library storage with Hierarchical Storage Management

NeSI New Zealand eScience Infrastructure

NeSI Canterbury / BlueFern SuperComputer

- Applications;
  - BG/P
    - Molecular Dynamics (NAMD, LAMMPS, VASP etc), Weather Forecasting (WRF), Protein Folding/Docking (AMBER, GROMACS, etc), Monte Carlo & researcher codes
  - P755/POWER7
    - Fluid Dynamics (Fluent/CFX), Genomics (MrBayes etc), Numerical (Octave, R), interpreted languages (Java, Python+SciPy/NumPy) & researcher codes
  - Visualization
    - Visualization tools (VTK, ParaView, VisIt etc.) and high-speed remote graphical sessions (vizstack, turboVNC, etc.)

NeSI NIWA / FitzRoy HPCF

- FitzRoy: IBM p575/p6 Supercomputer

  Being upgraded in Q4 2012

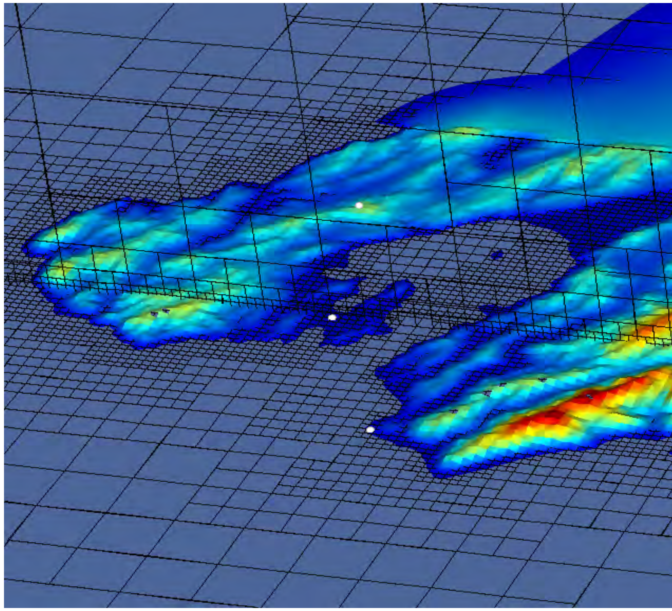
  - 58 (→ 108) POWER6 nodes
    - 32 cores / SMP node (64 bit)
  - 1,856 × 4.7GHz (→ 3,456) cores
  - 34 (→ 66) TFLOPS peak
    - 602 GFLOPS / node
  - 5.3 (→ 8.5) TB Memory
    - 64 and 128 GB memory nodes
  - InfiniBand interconnect fabric
  - AIX Operating System

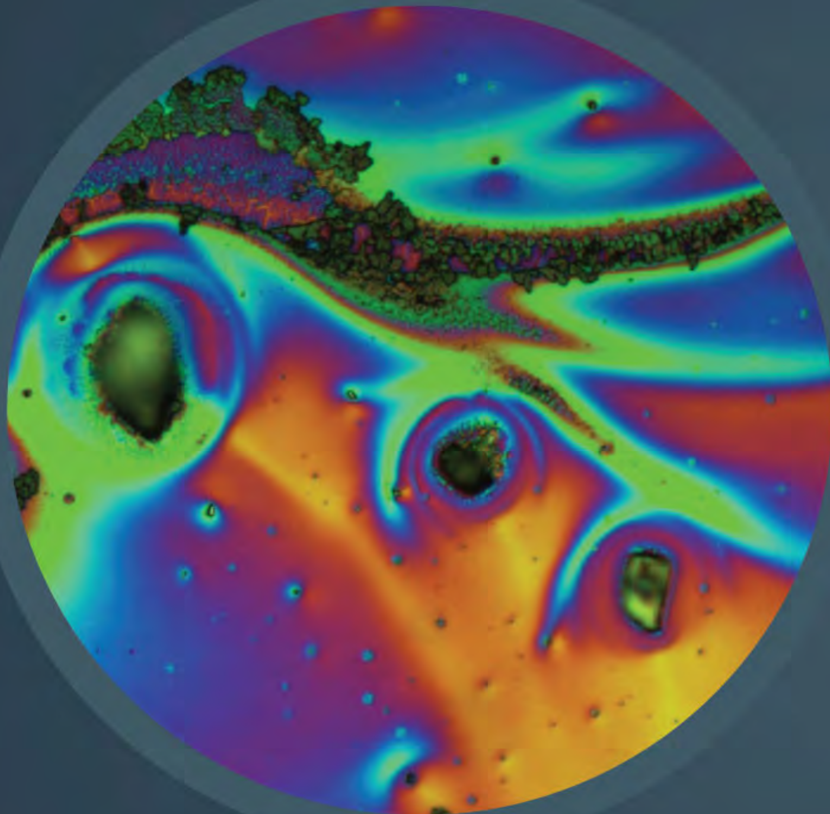
- Storage: Global Parallel File System (GPFS)
  - 790 TB SAN user disk
  - 5 PB Automatic Tape Library storage with Hierarchical Storage Management

NeSI New Zealand eScience Infrastructure

NeSI NIWA / FitzRoy Applications

- Unified Model (UK Met Office / Hadley Centre)
  - Weather forecasting: (global & regional – to 100 m resolution)
  - 3DVAR & 4DVAR data assimilation
  - Regional Climate modelling HadGEM3-RA
  - Coupled (atmospheric, ocean, land, sea ice) earth simulation HadGEM3
  - Chemistry Climate Modelling – UKCA
- Ocean Modelling:
  - ROMS (Regional Ocean Model)
  - NEMO. (Global Ocean Model)
- Wave Modelling:
  - WaveWatch 3, SWAN
- CFD: Gerris (self refining grid)
- Typical job sizes:
  - 64 – 1024 cores & O(10) GB output per job





NeSI
New Zealand eScience Infrastructure

NeSI Roadmap

Access Scheme

Allocation Classes

• Proposal Development	Free
• Research (peer reviewed)	80% subsidy by NeSI
• Educational	Free
• Collaborator	Free
• Private Industry	Full cost recovery

Access Policy Advisory Group

- Development of an Access Policy, in partnership with funding agencies
- Members are senior representatives of science and funding providers

… more from Technical Qualification Panel Chair, Mark Cheeseman, shortly..

Complexity & Computation

Challenges for researchers

- Multiple competencies required for modern computational approaches – "Computational Thinking"
- Computational facilities access and efficient use
- Maintaining links with leaders in relevant fields
- Communicating with institutions and facilities and colleagues in other disciplines

Challenges for institutions

- Comprehending opportunities that support and grow research
- Education and skills training
- Supporting distributed teamwork
- Sustaining and growing research computing infrastructure
- Coordination and collaboration across institutional boundaries

Scaling scientific computing

NeSI plays a variety of roles in addressing these needs…

Research computation – scaling up/out

- Code porting and tuning and scaling
- Bridging from Campus to National
- Bridging from National to International
- Computation as a Service

Experts

scientific applications Experts, available through our Research allocation class

Infrastructure Partnerships – scaling platforms

- Research communities scaling up, collaboration nationally & internationally
- Sector wide research infrastructures e.g. HPC, Research Data
- Institutional/campus bridging
- Innovation, new facilities/services to meet evolving needs

- REANNZ
- Tuakiri
- Institutional IT Services
- AeRO, Caudit

NZers are Computational Research leaders…

Medical & Health Sciences
Biological Sciences
Physical Sciences
Engineering
Math & Statistics
Humanities
Social Sciences

Outreach & Education

Communities of Practice:

3rd year for…

- NZ eResearch symposium – annual meeting that highlights work in the community to address researcher needs
- 2013, Christchurch

1st year for…

- NZ HPC Applications Workshop – focused on building a NZ community of researchers developing scientific applications
- 2013, Christchurch

Planned…

- eResearch leaders forum

Education:

3rd year for…

- Summer of eResearch – annual summer scholarship programme bringing together software engineering students with researchers to tackle IT problems at a national scale
- Aim is to develop a balanced programme, where institutions and Crown jointly invest into building capability thorugh education

Planned…

- Recruiting a national NeSI Education Lead, who will act as a colleague of those teaching HPC and eScience across the sector, collaborate on curriculum development, design, and delivery, and engage in training

eScience Services

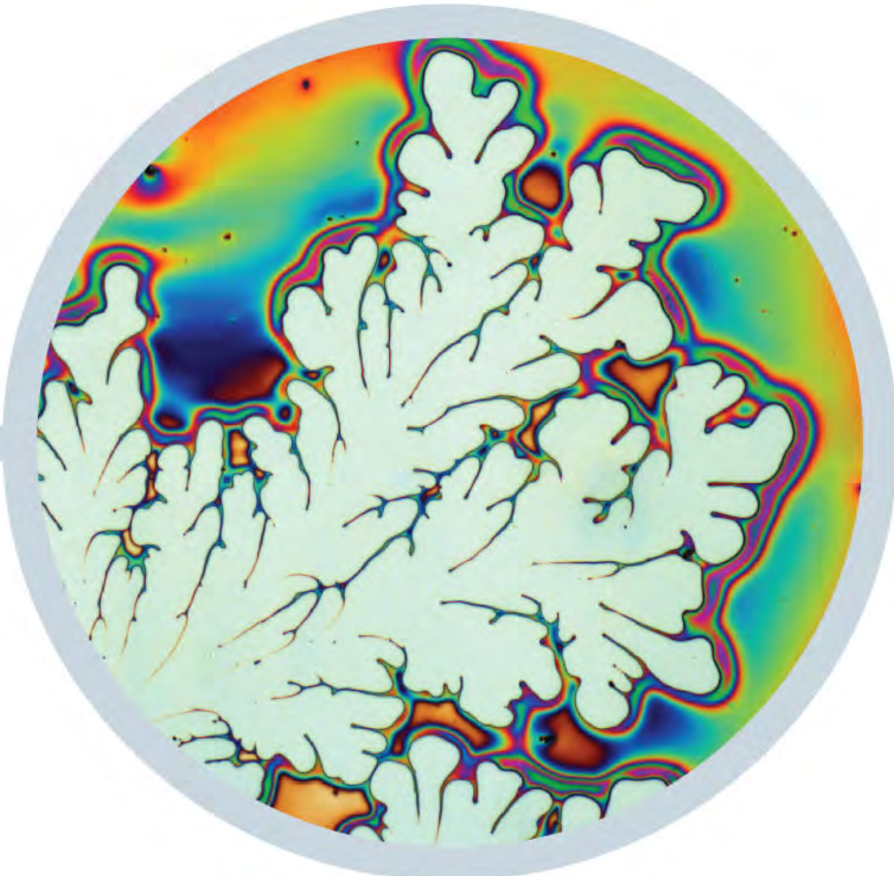
- DataFabric (Federated Data Sharing Service)
  - Hosted at multiple redundant sites
  - Using best of breed platform: Integrated Rule Oriented Data Service (iRODS)
  - Tuakiri authentication
  - WWW / WebDAV / FUSE / GridFTP
  - Status: Prototype (BeSTGRID migration to NeSI)

- Data Transfer
  - High throughput transfer of data across REANNZ network nationally, and internationally
  - GridFTP and GlobusOnline
  - Status: Protoype (BeSTGRID migration to NeSI)

- Hosting (virtual) for national services
  - Hosting of databases and core NeSI services
  - Status: Prototype (New platform)

NeSI

New Zealand eScience Infrastructure



NeSI
New Zealand eScience Infrastructure

Progress Report

ministry of science + innovation · The University of Auckland · University of Canterbury · Landcare Research Manaaki Whenua · NIWA Taihoro Nukurangi · University of Otago

Collaboration

"Big scientific problems require big teams these days and our current institutional arrangements, with their high transaction costs and researcher-scale accountabilities, are ill-suited to meet such challenges. Putting together large, multi-institutional teams to tackle complex problems remains depressingly difficult in the New Zealand environment."

03/2012, Shaun Hendy, President, NZAS

NeSI Status

- Creating an advanced, scalable computing infrastructure to support New Zealand's research communities
  - 3 international scale HPCs operating
- Providing the grid middleware, research tools and applications, data management, user-support, and community engagement needed for the best possible uptake (of HPC)
  - NeSI is staffed (few open positions), grid middleware being developed, community engagement underway (HPC Workshop, Presentations beginning)
- Encouraging a high level of coordination and cooperation within the research sector
  - Has been a challenge – but Auckland, Canterbury & NIWA working together for the good of NZ Science
- Contributing to high quality research outputs… which support the Government's published priorities for science
  - Too early to tell – but we need to do so within the next 12 months

NeSI Statistics

Operational HPC facilities since January 2012

First Call for proposals for national allocations: May 2012

Current Call – September: Third Call

28 Projects across development and research classes

Allan Wilson Centre, ANU, Auckland, Auckland DHB, Canterbury, CSIRO, ESR, GNS, Lincoln, Livestock Improvement Corporation, IRL, MacDiarmid Institute, Massey, Met Service, NIWA, Ohio State, Otago, Scion, VUW, Waikato

http://www.nesi.org.nz/projects

NeSI New Zealand eScience Infrastructure

Strategic opportunities

Research community and sector wide partnerships

e.g.

- National Science Challenges
- Research Data Management
- Callaghan Institute (Advanced Technology Institute)

Leveraging further institutional capabilities?

- Significant institutional investments into eScience and HPC resources that might be better coordinated nationally

Strategy & Plans

2012

- Strategic Plan under development
- Seeking sector feedback late 2012/early 2013

2013

- September: Contract review, to inform renegotiation and continuation from July 2014

NeSI
New Zealand eScience Infrastructure

Summary

NeSI is a big new investment (by the Crown and Sector, into HPC for NZ Science)

- It is making a world-class HPC ecosystem available to NZ Science
- It is a collaboration between Universities & CRIs
- Initial contract funded till June 2014 – but will need to prove its success by September 2013 – is in long term budget, if we perform

To be successful… (i.e. attract ongoing investment to support national HPC access)

- NZ Scientists will need to demonstrate their need for HPC (see User Needs Survey)
- This means transitioning from the problems that I can solve on "my" PC, and/or departmental cluster – to large scale HPC provided by NeSI
  - This is a function of the funding-round cycle too…
  - It will take time to learn new programming methods & tools: MPI, OpenMP, and new Operating Environments

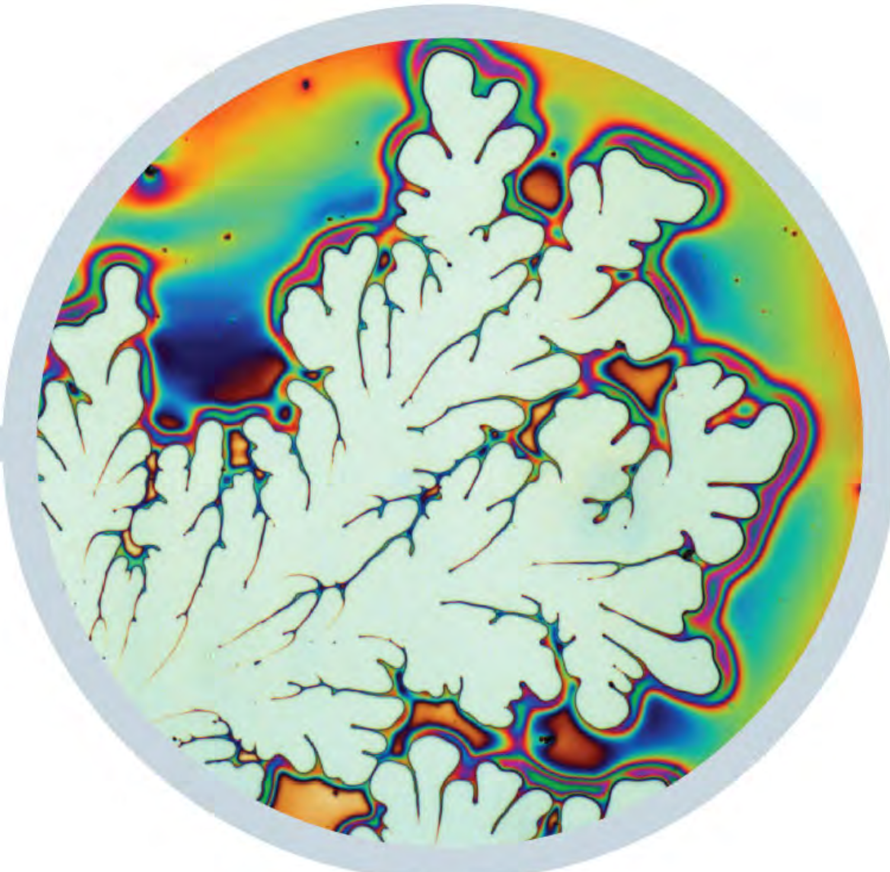
Summary

and, importantly:

- PIs will have to demonstrate the value of NeSI by funding 20% of their access costs
  - This has implications for the way our Institutions provide access to Operational Expenditure (most prefer to provide Capital Expenditure)
- Research Projects using NeSI will need to generate excellent science (that could not be done without these HPC facilities)
- Research Projects and NeSI need to contribute to Government Outcomes

.. in which case we can expect a long period of HPC funding in the years ahead

.. and can aspire to a broader scope of eScience infrastructure, too…



NeSI

New Zealand eScience Infrastructure

NeSI http://www.nesi.org.nz

Access Policy	http://www.nesi.org.nz/access-policy
Eligibility	http://www.nesi.org.nz/eligibility
Allocation Classes	http://www.nesi.org.nz/allocations
Application Forms	http://www.nesi.org.nz/apply
Calls Timetable	http://www.nesi.org.nz/timetable
Case Studies	http://www.nesi.org.nz/case-studies
Projects	http://www.nesi.org.nz/projects

Facilities

- Center for eResearch, Auckland
  - http://www.eresearch.auckland.ac.nz/uoa
- BlueFern SuperComputer, Canterbury
  - http://www.bluefern.canterbury.ac.nz
- FitzRoy HPCF, NIWA
  - http://www.niwa.co.nz/our-services/hpcf

ministry of science + innovation · The University of Auckland New Zealand · UC University of Canterbury · Landcare Research Manaaki Whenua · NIWA Taihoro Nukurangi · University of Otago New Zealand