The NIWA/NeSI HPC Replacement Project:

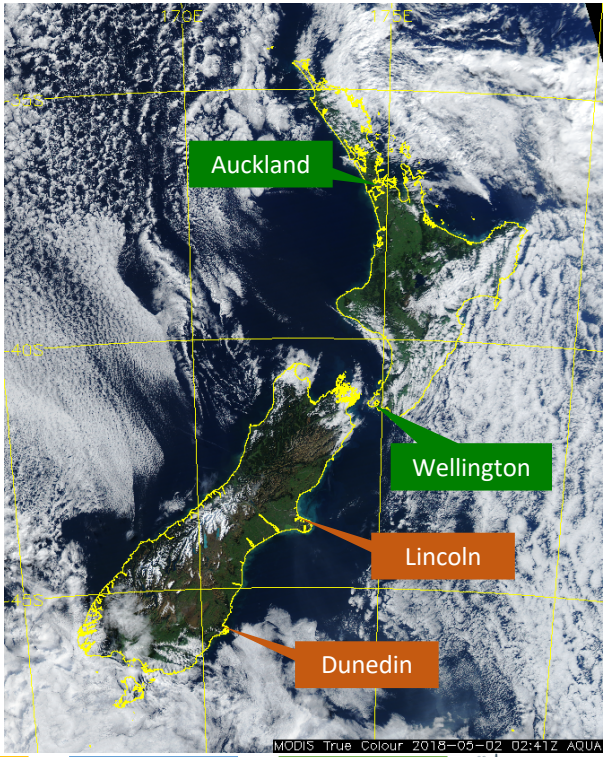
A voyage in complexity integrating (multi-site) XC, CS, ESS and OpenStack systems

- M. Uddstrom[1], B. Corrie[1], N. Jones[1], F. Cantos[1], A. Hicks[1], W. Hayek[1], D. Kelly[2], P. Balle[2], A. Sachitano[2], B. Gilmer[2], J. Altorf[2], D.McCurdy[3], A. Beattie[3].
  [1]NIWA/NeSI, [2]Cray, [3]IBM

New Zealand eScience Infrastructure

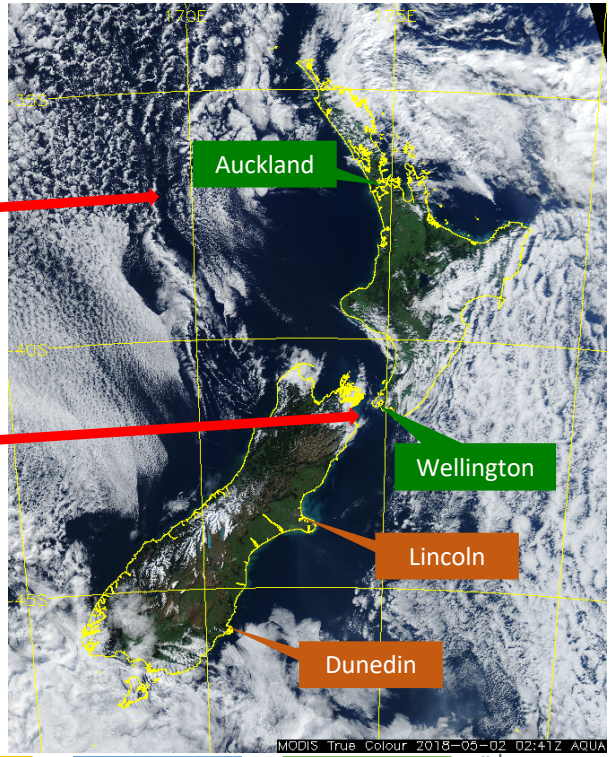
About NIWA and NeSI

- NIWA: National Institute of Water and Atmospheric Research (a Crown Research Institute - CRI):
  - NIWA'S purpose is to enhance the economic value and sustainable management of New Zealand's aquatic resources and environments, *to provide understanding of climate and the atmosphere and increase resilience to weather and climate hazards to improve safety* and wellbeing of New Zealanders.
- NeSI: New Zealand eScience Infrastructure (collaboration: University of Auckland, University of Otago (Dunedin), Manaaki Whenua (CRI, Lincoln) and NIWA (CRI, Wellington))
  - NeSI's purpose is to *grow the computing capability of researchers* to ensure New Zealand's future prosperity
  - ~50% Government funded, provides National HPC services



Background and Context

- NeSI: HPC legacy:
  - Established in 2011;
  - Share of NIWA's IBM P575/P6: 2,208 cores;
  - IBM iDataPlex (2012/13/14): 5,784 cores (WSM, SAB, IVB), GPFS
- NIWA: HPC legacy:
  - Cray T3E 1200e (1999/04): 544 cores
  - IBM P575/P6 (2010/13): 3,392 cores, GPFS
- The Challenge:
  - Four investing institutions (2×CRIs, 2×Uni's);
  - Design - coming to agreement;
  - RFP…

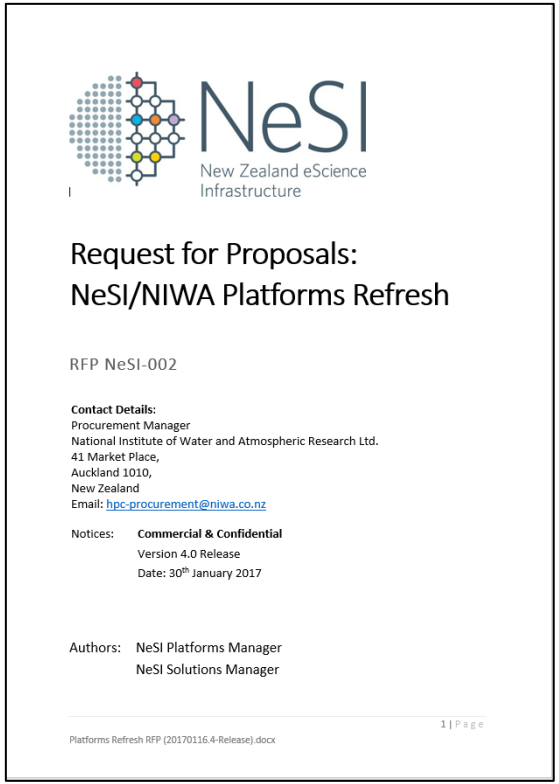


Design Decisions

- National data-centric research and operational computing environment:
  - Single site: NIWA, Wellington;
  - Capacity Cluster: High Throughput and Private Cloud;
  - Capability: Large simulations & NIWA forecasting;
  - Capability (DR): NIWA forecasting (Auckland);
  - High performance filesystems.
- New user services:
  - Virtual Labs;
  - Remote visualisation;
  - OpenStack private cloud;
  - Advanced data analytics;
  - Hierarchical storage management services.
- Back-end services that mitigate the risk of data loss.

Procurement Strategy

- Maximise return on investment:
  - One RFP (3 HPCS);
  - Single Site, shared storage;
  - NIWA ($15.2M), UoA, UoO and MW ($4.8M);
  - Separate NIWA DR site (Auckland).
- Benchmark driven:
  - Capacity: NAMD, ANSYS, GROMACS, GATK, NWCHEM;
  - Capability: Unified Model, NEMO, EMOD3D;
  - I/O: IOR, MDTEST, IOZONE;
  - Workflow:
    - Real use case (cylc NWP cycle including post processing)
  - Workloads.



NeSI
New Zealand eScience Infrastructure

Request for Proposals:
NeSI/NIWA Platforms Refresh

RFP NeSI-002

Contact Details:
Procurement Manager
National Institute of Water and Atmospheric Research Ltd.
41 Market Place,
Auckland 1010,
New Zealand
Email: hpc-procurement@niwa.co.nz

Notices: **Commercial & Confidential**
Version 4.0 Release
Date: 30th January 2017

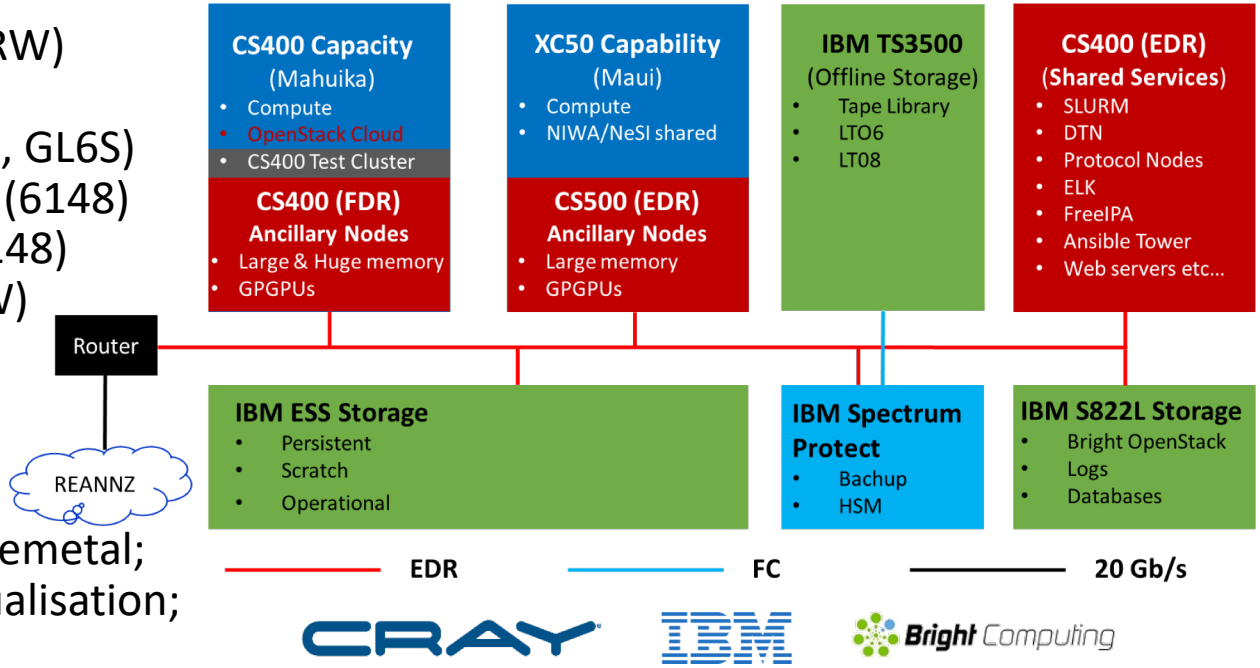
Authors: NeSI Platforms Manager
NeSI Solutions Manager

1 | Page

Platforms Refresh RFP (20170116.4-Release).docx

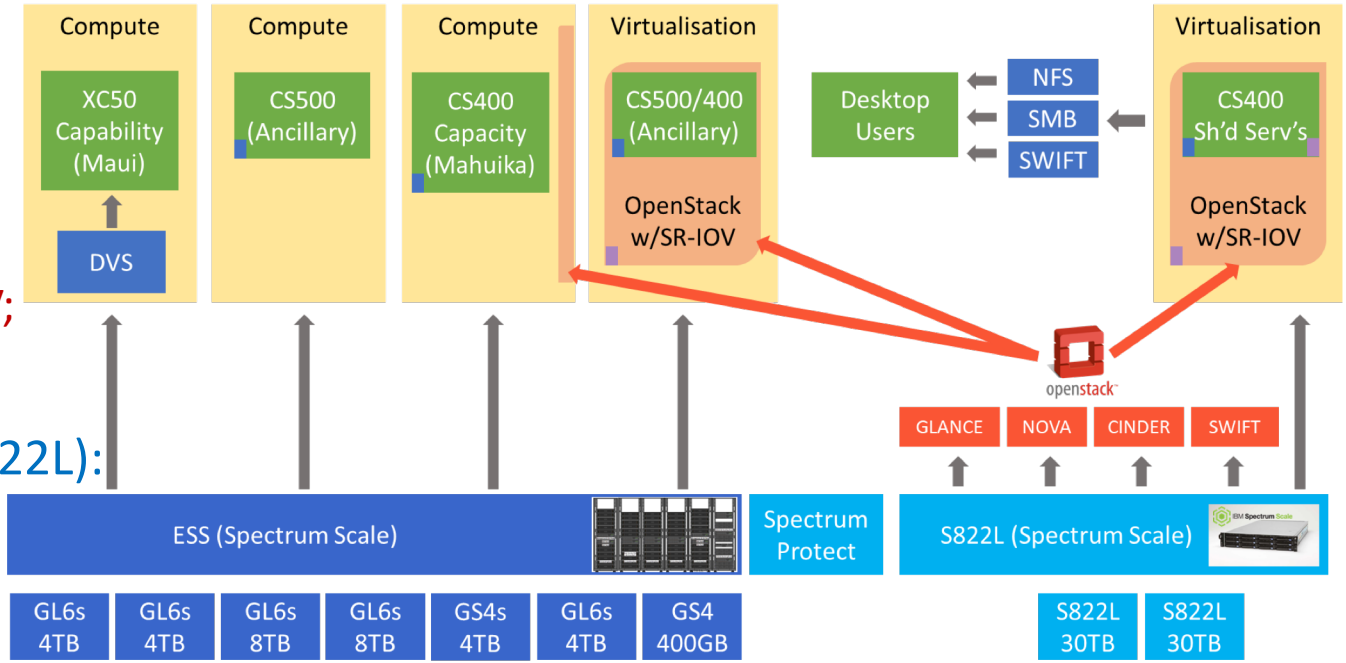
Solution (Wellington)

- NIWA/NeSI:
  - CS400: 9,604 cores (BRW)
  - CS400: Test Cluster
  - ESS:10.1PB (GS4, GS4S, GL6S)
  - XC50-LC: 18,560 cores (6148)
  - CS500: 1,120 cores (6148)
  - CS400: 180 cores (BRW)
  - Nvidia P100: 16
  - S822L: 30TB
- Ancillary Nodes:
  - OpenStack VMs or Baremetal;
  - NICE DCV Remote Visualisation;
  - Virtual Labs;
  - Pre/Post Processing.



Storage Architecture (Wellington)

- User Storage (ESS):
  - GPFS (aka Scale) clients:
    - CS400;
    - CS500.
  - DVS GPFS nodes:
    - XC50.
  - BOS VM via SR-IOV;
  - Other systems:
    - Protocol nodes.
- Flexible Storage (S822L):
  - Provisioning OSs;
  - Databases;
  - ELK, etc.



Some Performance Measures

XC core performance relative to P575/P6 at P6 core counts	Large Simulation Codes: (UM, NEMO, EMOD3D)	Large Simulation Codes + NAMD
BRW E5-2695v4 (2.1 GHz, 18 cores/socket) RFP reference	1.60	1.56
SKL 6148 (2.4GHz, 20 cores/socket) BAFO estimate	2.08	1.97
SKL 6148 (2.4 GHz, 20 cores/socket) Acceptance Tests	2.14 (33.8%)	2.07 (32.7%)

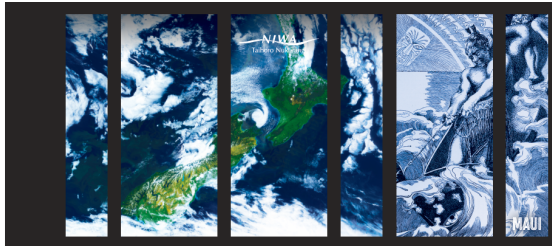
ESS (4×GL6S, 1×GS4S) GPFS v5.0	CS400 (Native GPFS Client)			XC 50 (via DVS – 36 nodes)		
MDTEST (4KB creates, unique dirs.)/s	156,900			35,766 (23%!)		
	Read	**Write**	**Total**	**Read**	**Write**	**Total**
IOR (Single Stream) 4KB (GB/s)	2.1	1.3		3.2 (w/IOBUF!)	3.2 (w/IOBUF!)	
IOR (Single Stream) 8MB (GB/s)	5.1	3.3		2.5	2.3	
IOR (total bandwidth) (GB/s)	59.5	86.7	**146.3**	63.0	64.0	**126.9**

Status

- NIWA DR (XC50, CS500, CS400, OpenStack) installed and operational since 01/2018:
  - P575/P6 (Wellington) users/data transferred (~6PB) 03-Dec-2017;
  - NIWA operational forecasting service 09-Jan-2018.
- NIWA/NeSI (CS400 Cluster, XC50, CS500, CS400, OpenStack) installation started: 26-Feb-2018:
  - Upgrade GPFS 4.3 to GPFS 5.0 (s/w and filesystems);
  - Passed acceptance tests: 30 Apr 2018;
  - Bright OpenStack completed 18-May-2018;
  - Expect to put first users on the system – Mid June.
- Wellington Plant Upgrade:
  - Increased data centre efficiency, power and cooling redudancy.
- Issues: DVS (need a better solution for GPFS sites).

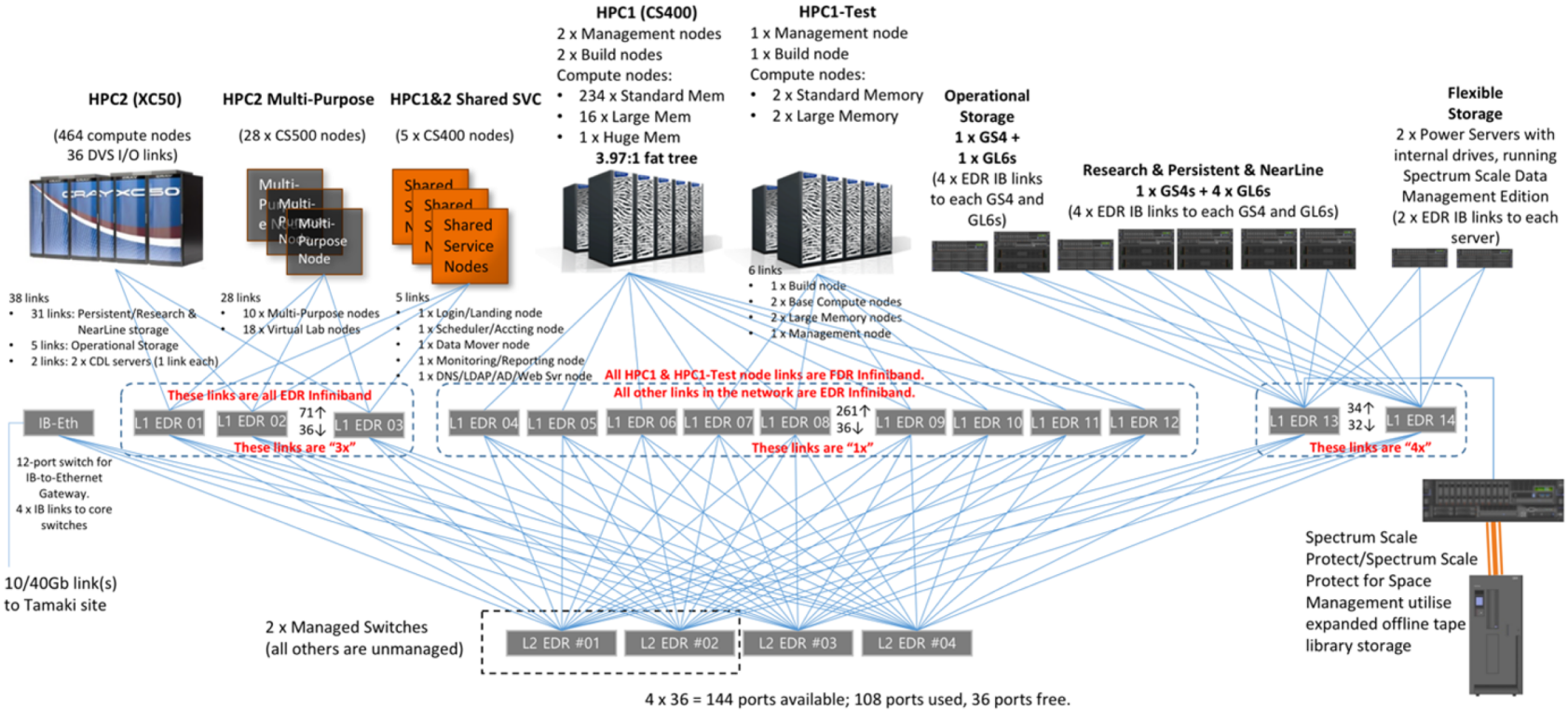

XC50 (Auckland)


CS400, ESS, XC50 (Wellington)



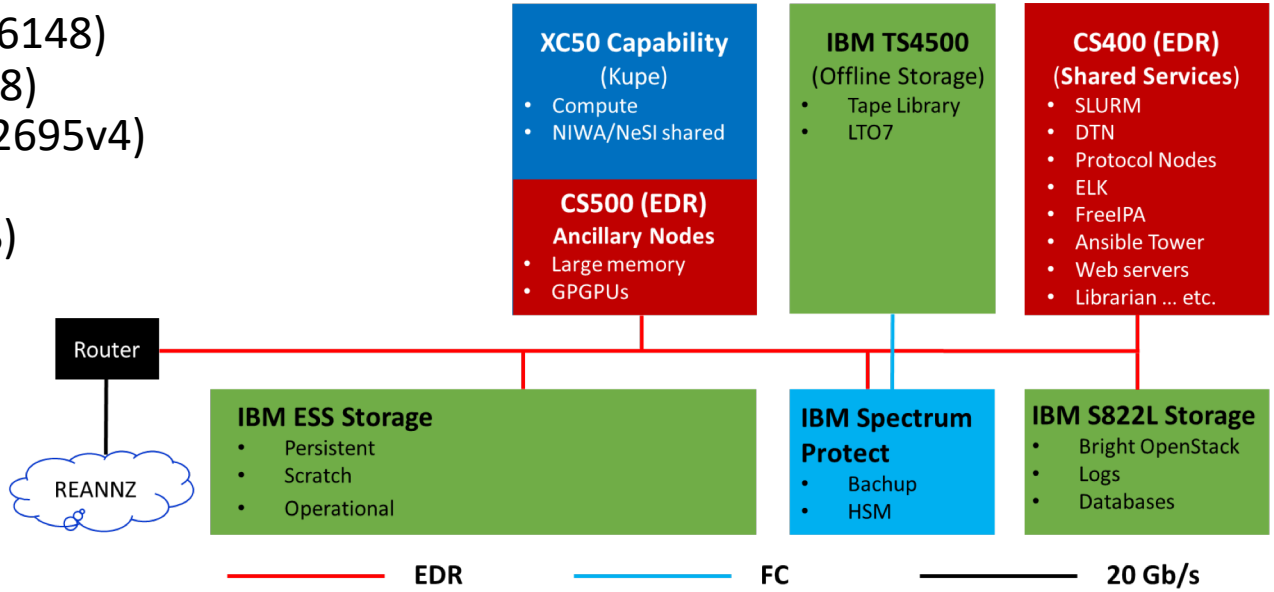
Additional Slides

NIWA/NeSI - Wellington



Solution (Auckland)

- NIWA:
  - XC50-AC: 4160 cores (6148)
  - CS500: 440 cores (6148)
  - CS400: 144 cores (E5-2695v4)
  - Nvidia P100: 2
  - ESS:4.5PB (GS4S, GL6S)
  - S822L: 30TB
- Ancillary Nodes:
  - OpenStack VMs or Baremetal;
  - Virtual Labs;
  - Remote Visualisation;
  - Pre/Post processing…



NIWA DR - Auckland