# NIWA

Taihoro Nukurangi

Climate, Freshwater & Ocean Science

2023 ECMWF Workshop on HPC in Meteorology

# Exploration of public cloud computing by an operational site running the Unified Model –

*Jeff Zais, NIWA*

20th ECMWF Workshop on HPC in Meteorology

October 2023

Climate, Freshwater & Ocean Science

NIWA
Taihoro Nukurangi

# … about New Zealand

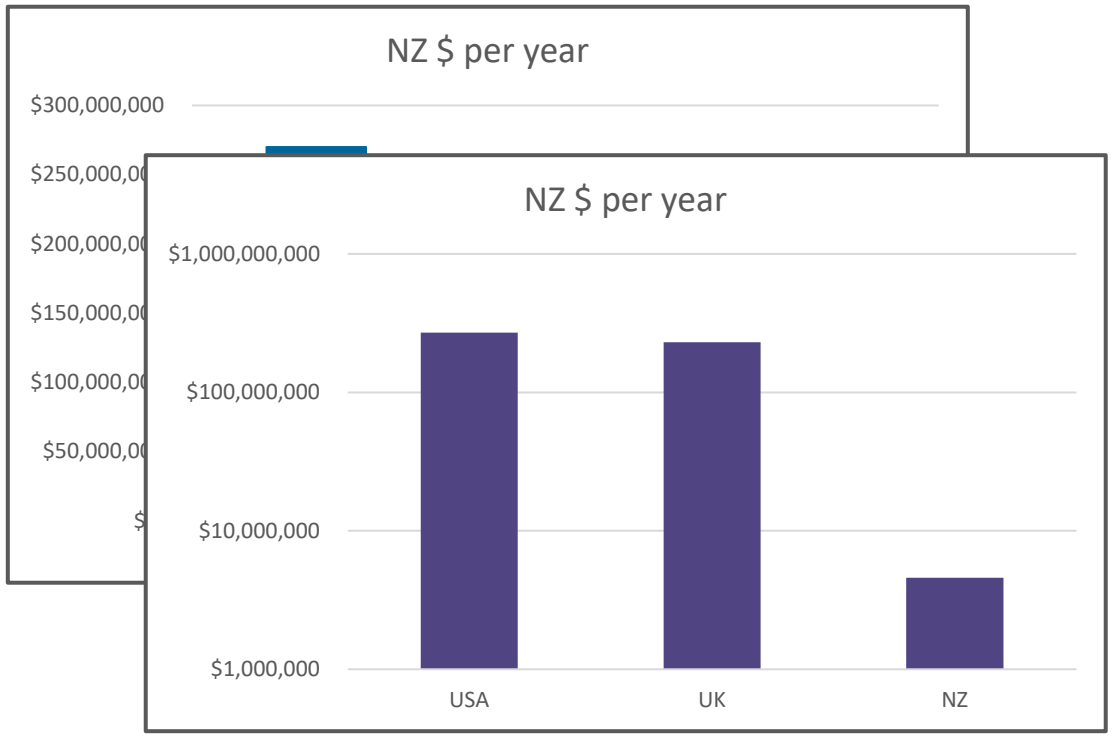
*2 main islands*

*5 million residents*

*11 time zones away*

*diverse landscapes*

# From exa to zetta – and back down to a dose of Kiwi reality




**New Zealand observations**

- Practicality
- Use our natural advantages
- Exports & imports
- Study & survey the rest of the world, then use their success as a guide
- Playful banter with those across the Tasman Sea

# … about NIWA

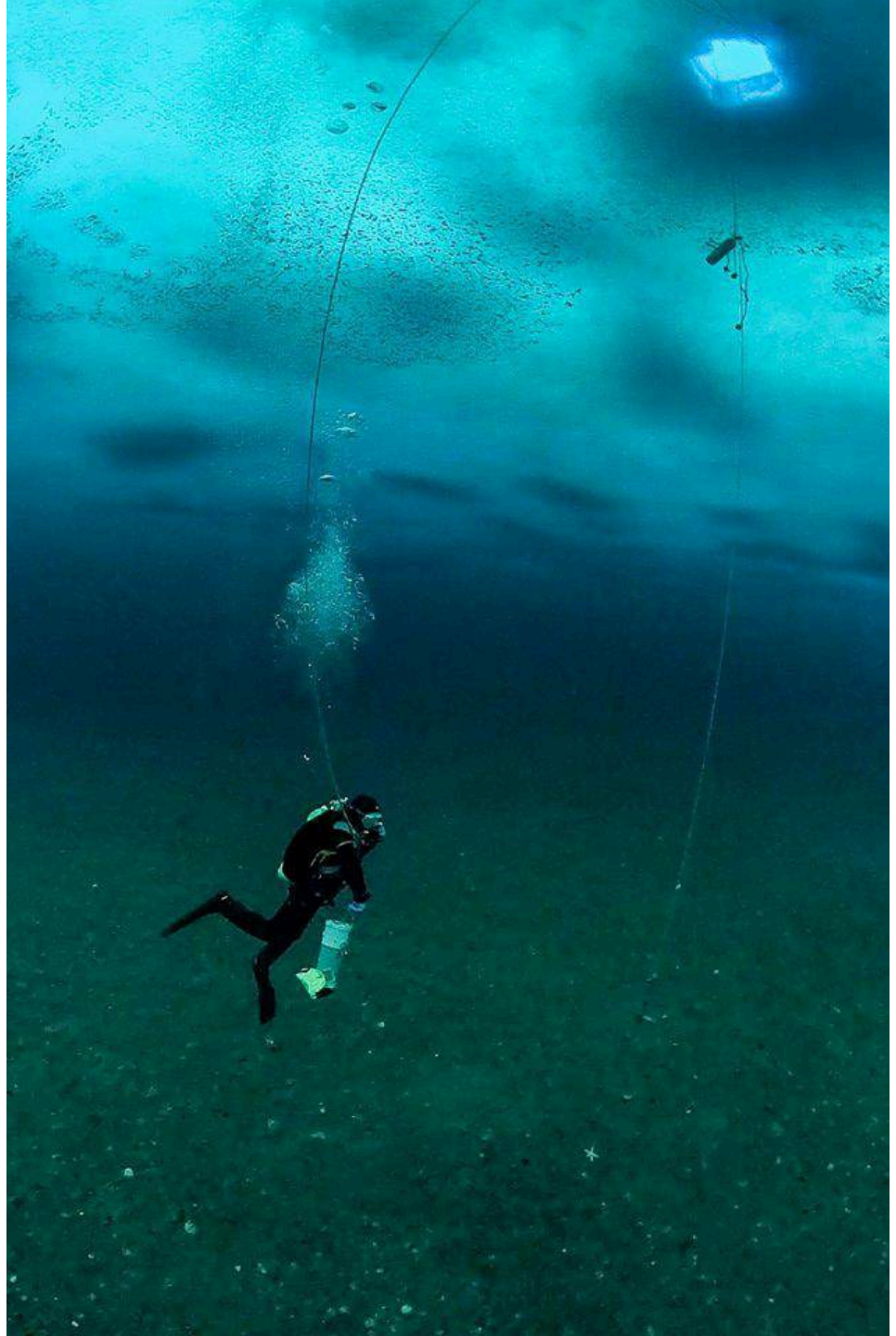
National Institute of Water & Atmospheric Research

- 1992
- 1 of 7
- 13
- 700

NIWA
Taihoro Nukurangi


RV Tangaroa

# Field work *under the ice shelf*



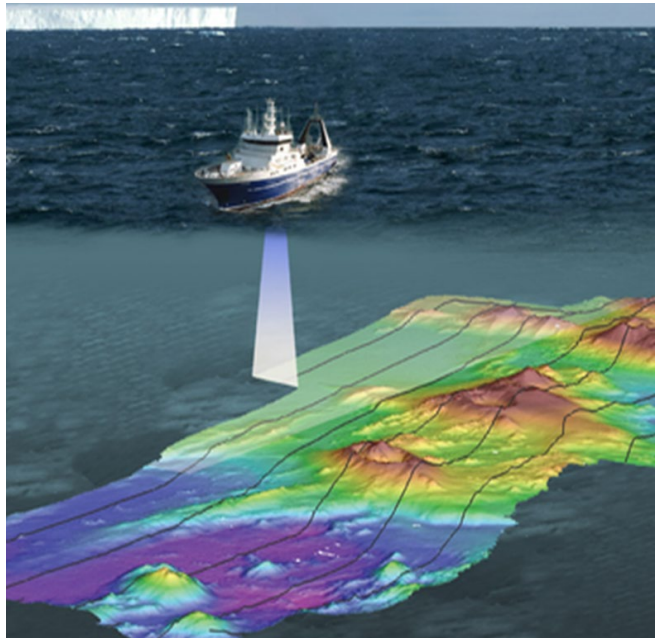
Climate, Freshwater & Ocean Science

NIWA
Taihoro Nukurangi

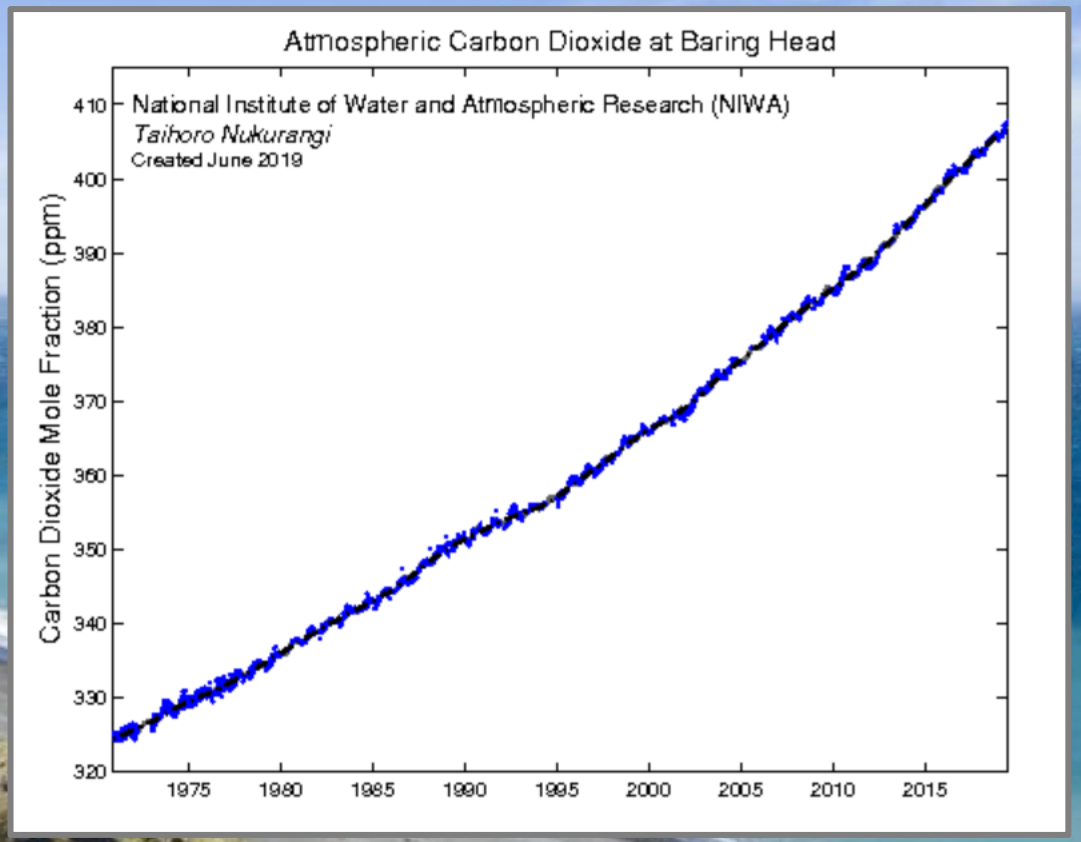
# *update from the 2023 Antarctic voyage*


NIWA
Taihoro Nukurangi

# A growing number of remote sensing data sources



# Baring Head – 50 year record of data



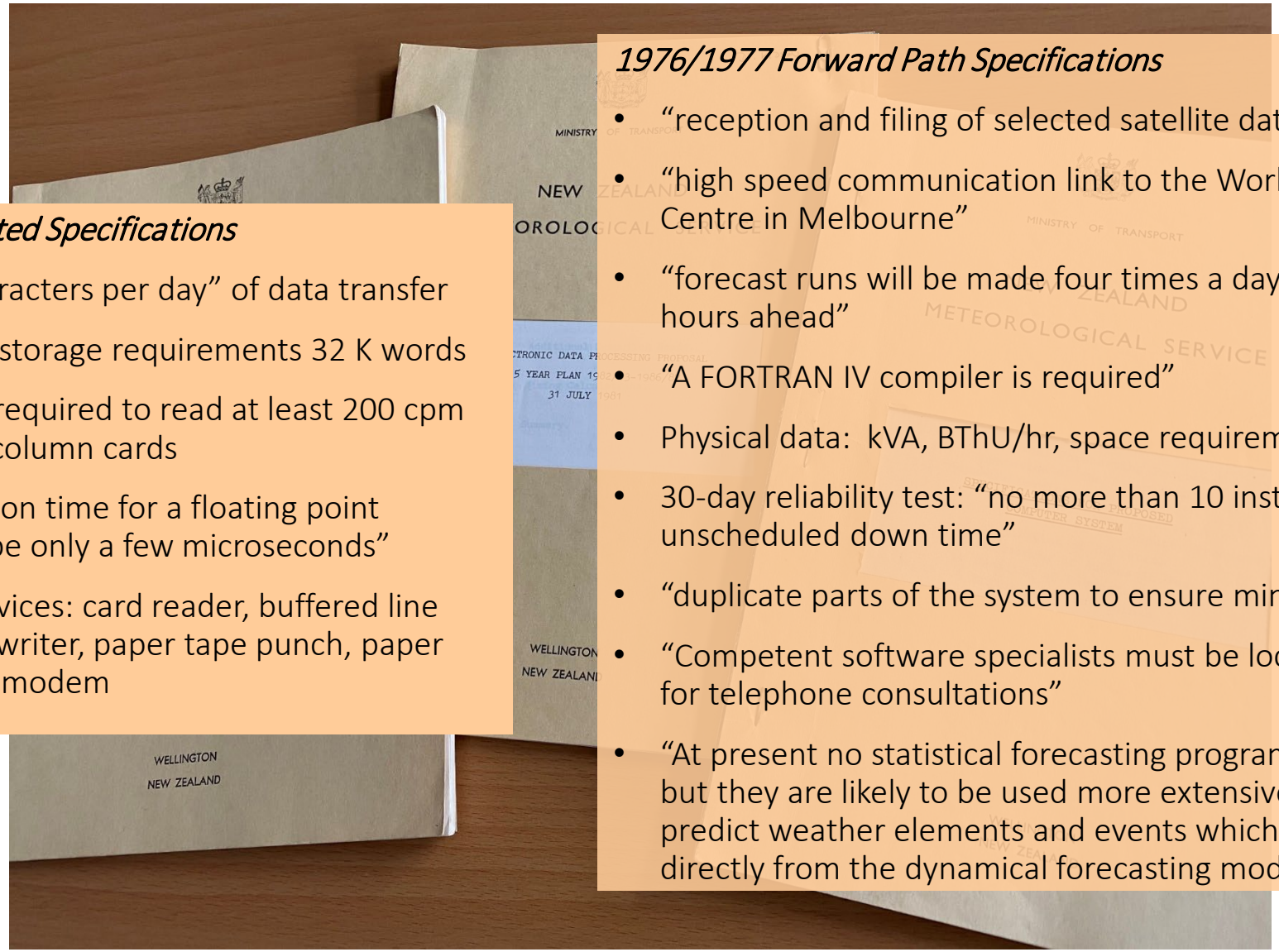
# Data on a journey


*Rain drops that make their way into one of NIWA's smart rain gauges mark the beginning of a long and involved data journey that snakes through cellular networks, desktop applications, data centres and top-of-the-line supercomputers.*

Climate, Freshwater & Ocean Science

NIWA
Taihoro Nukurangi

# … about HPC at NIWA




*1976/1977 Antiquated Specifications*

- "one million characters per day" of data transfer
- "core (memory) storage requirements 32 K words
- a card reader is required to read at least 200 cpm on standard 80 column cards
- "average execution time for a floating point instruction can be only a few microseconds"
- input/output devices: card reader, buffered line printer, teletypewriter, paper tape punch, paper tape reader, ICC modem

*1976/1977 Forward Path Specifications*

- "reception and filing of selected satellite data in real time"
- "high speed communication link to the World Meteorological Centre in Melbourne"
- "forecast runs will be made four times a day for periods of 12-36 hours ahead"
- "A FORTRAN IV compiler is required"
- Physical data: kVA, BThU/hr, space requirements
- 30-day reliability test: "no more than 10 instances of unscheduled down time"
- "duplicate parts of the system to ensure minimal down time"
- "Competent software specialists must be located in New Zealand for telephone consultations"
- "At present no statistical forecasting programs are operational but they are likely to be used more extensively in the future to predict weather elements and events which can not be done directly from the dynamical forecasting models"

# 1999

- CRAY T3E-1200
- #89 on the November 1999 TOP500 list
- *144 processors, 0.17 peak teraflops*


# 2010

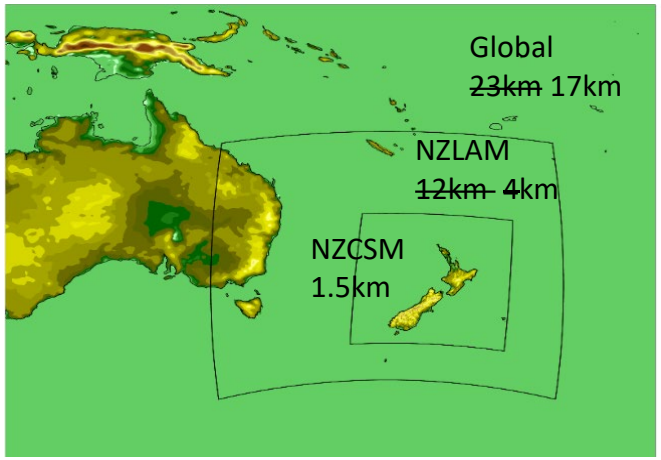
- IBM p575 POWER6 system
- 700 kilowatts for power & cooling

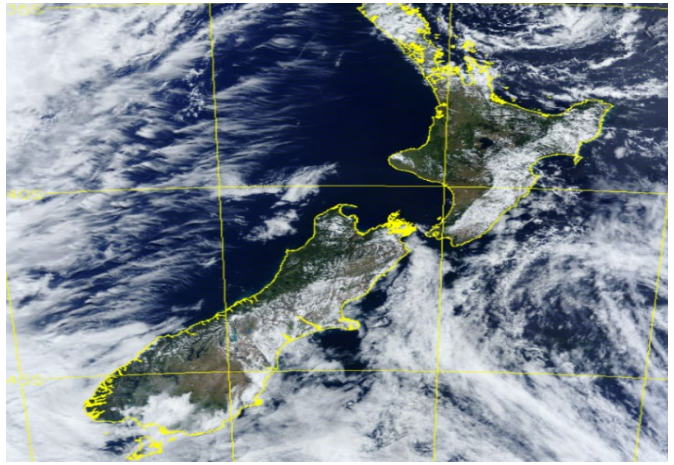
NIWA
Taihoro Nukurangi

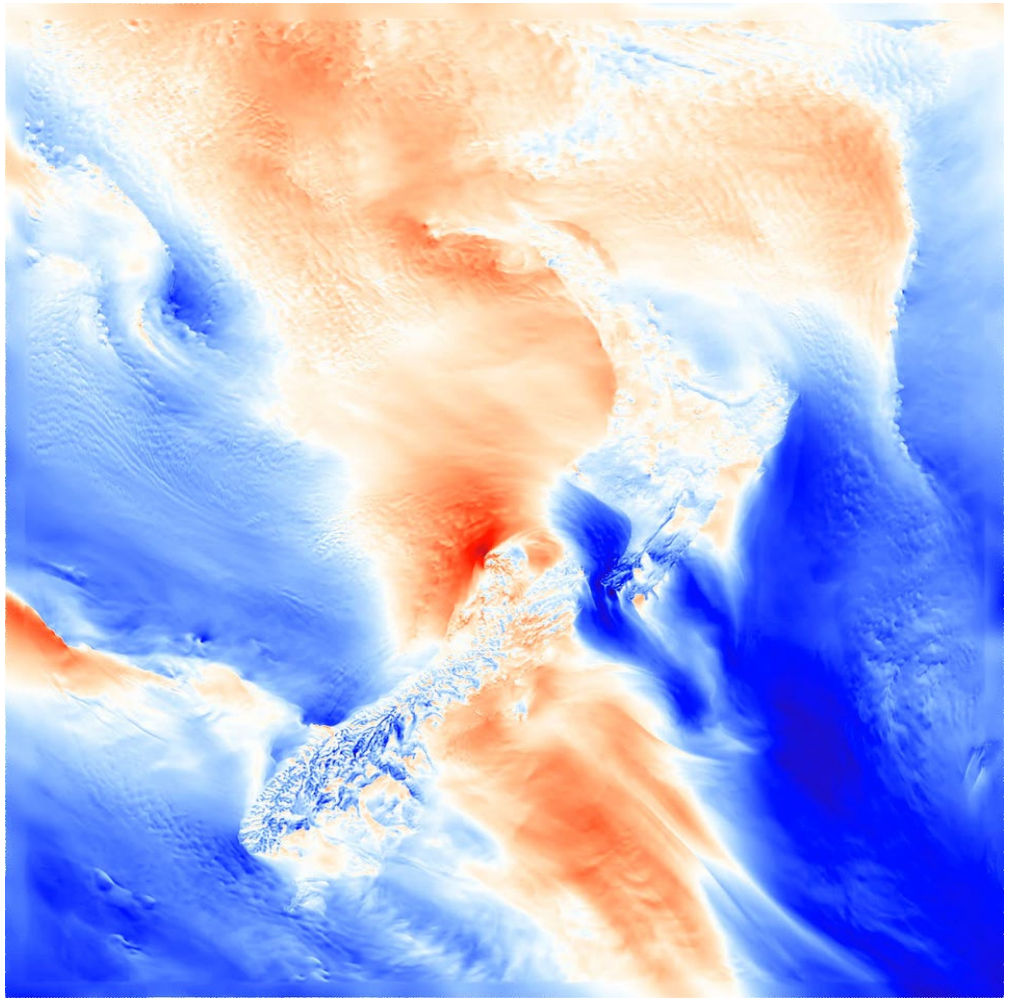
# 2017/2018

- Cray XC50, Cray CS400 & CS500, Spectrum Scale
- 400 kilowatts sum; *'Maui' could be #500 in 2019?*
- smaller disaster recovery system 'Kupe' in Auckland

# … about numerical weather prediction at NIWA – as part of the Unified Model Partnership



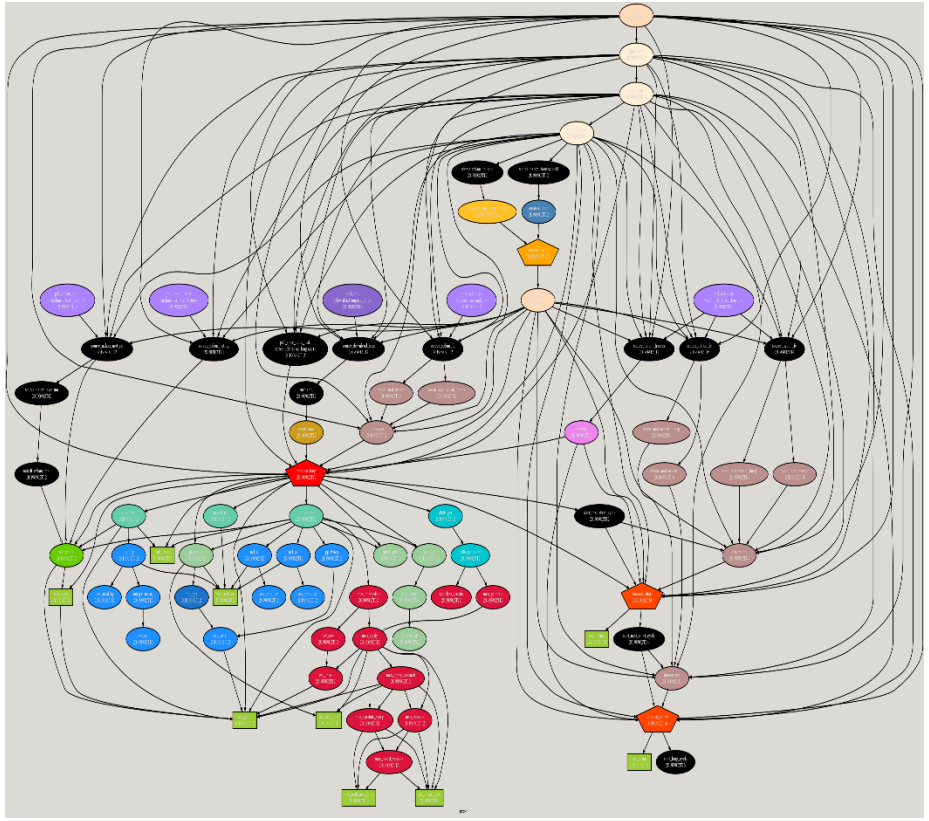





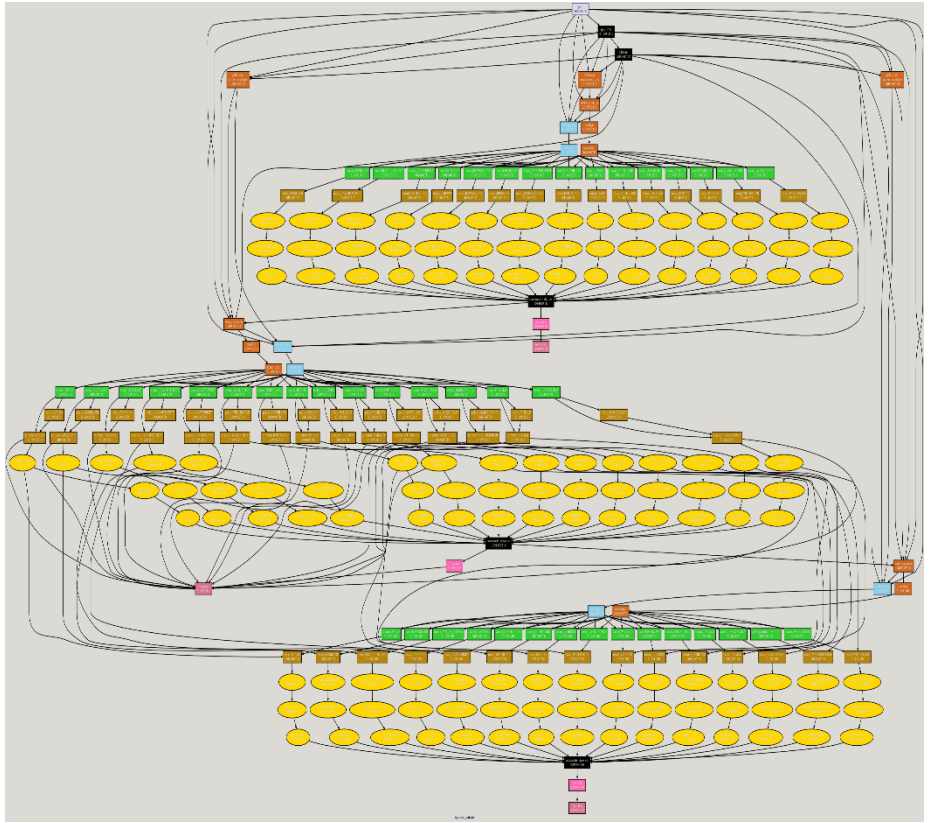


NIWA
Taihoro Nukurangi

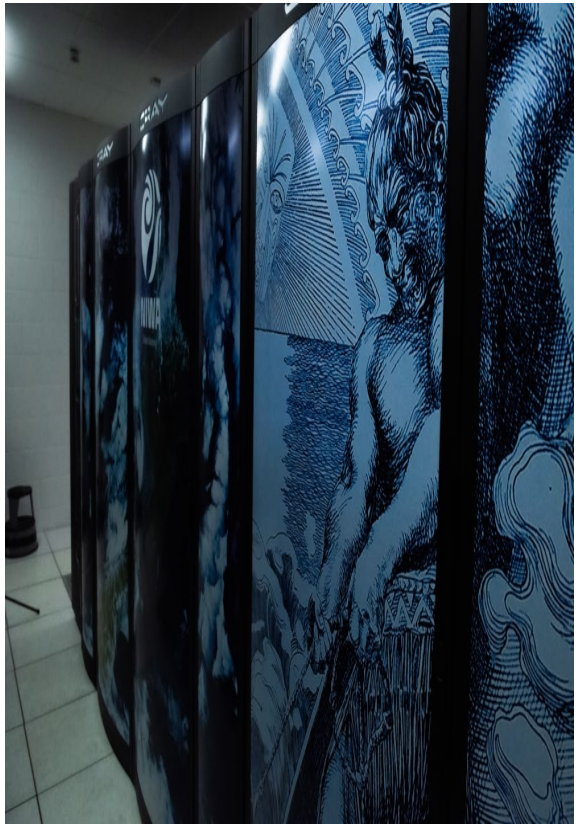
# The scientific workflow for operational forecasting is very complex – and handled by the Cylc (“silk”) Workflow Engine



Weather forecast suite



National River flow forecast suite



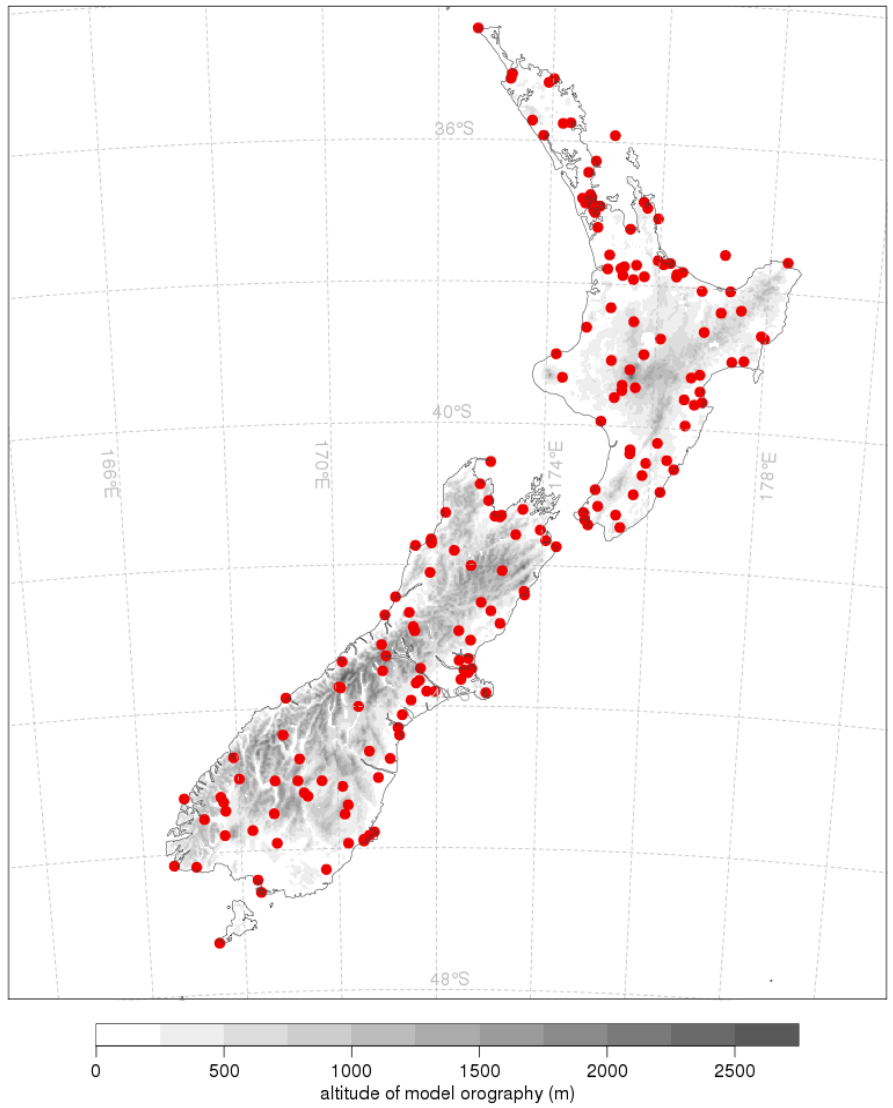
High Performance Computing

H. Oliver, *et al.*,"Workflow Automation for Cycling Systems" in *Computing in Science & Engineering*, vol. 21, no. 04, pp. 7-21, 2019. doi: 10.1109/MCSE.2019.2906593

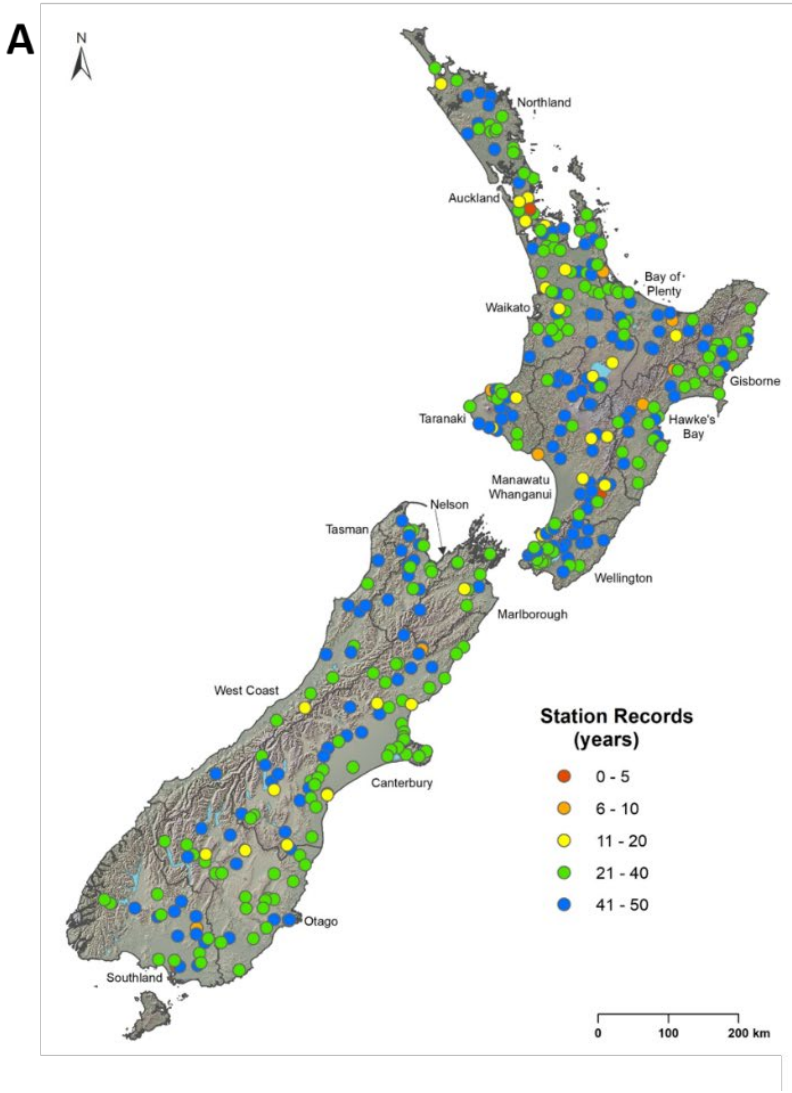
NIWA
Taihoro Nukurangi

# Combine weather with rainfall data, river flow data, and catchment models

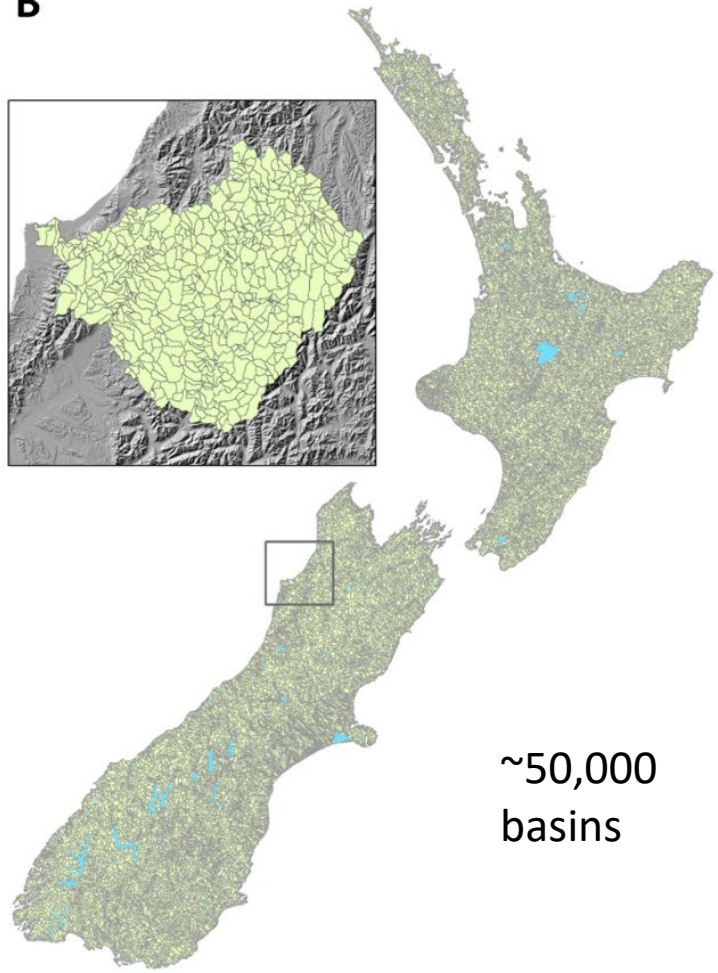
## Rain data at NIWA




## Flow data: *no national database*




## Modelled rivers




~50,000 basins

# Flood event July 2021 – Buller model

## TIMELINE OF EVENTS:

**Calibrated flood model developed for West Coast by NIWA in 2016.**

Various forecasts running for Buller, Grey, Karamea, Hokitika since 2018.

**Wed 14 July 2021**

MetService issues red alert (only third ever for NZ) for heavy rainfall event for region.

**Thu 15 July**

WCRC contacts NIWA to provide river forecast information.

**Fri 16 July 2021**

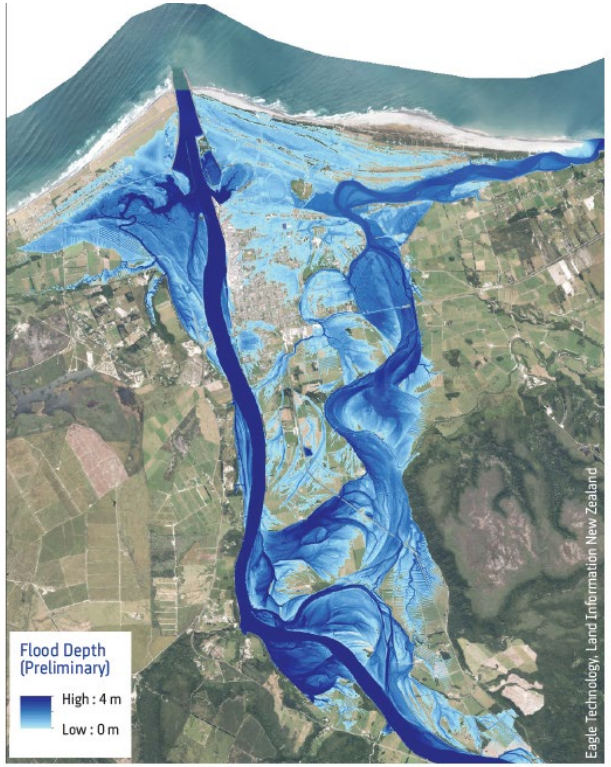
**8am** NIWA provides weather briefing to FENZ and river forecasts to WCRC, CDEM and FENZ.

**1pm** NIWA sends updated river flow and sea level forecasts to CDEM.

**1:15pm** Buller declares local emergency, voluntary evacuation.

**Sat 17 July**

WCRC/CDEM request guidance from NIWA – who advise not to let evacuees return home yet as river still rising. River peaks at 3pm, inundation in Westport continues until 10pm.




*NIWA inundation map of 2021 event in Westport from post-event survey information.*

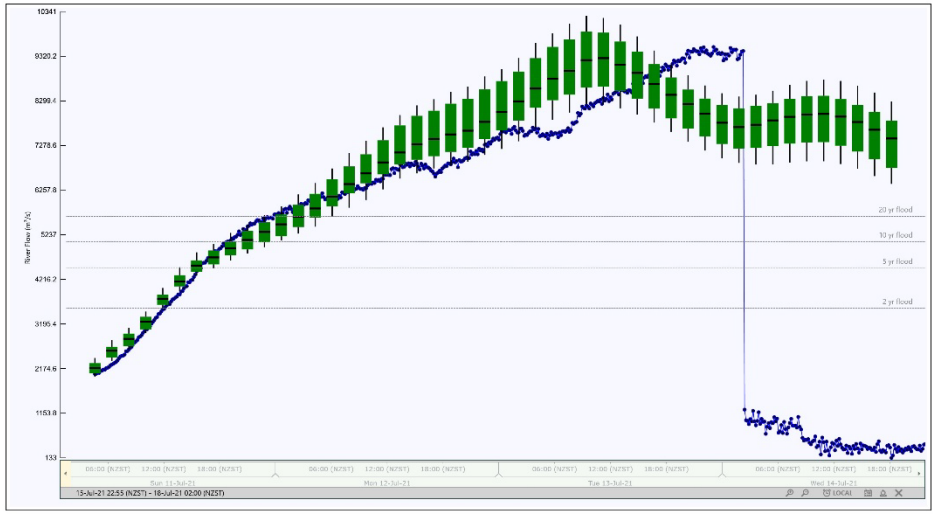



Chart 1
(Printed 10-Aug-2021 14:58)

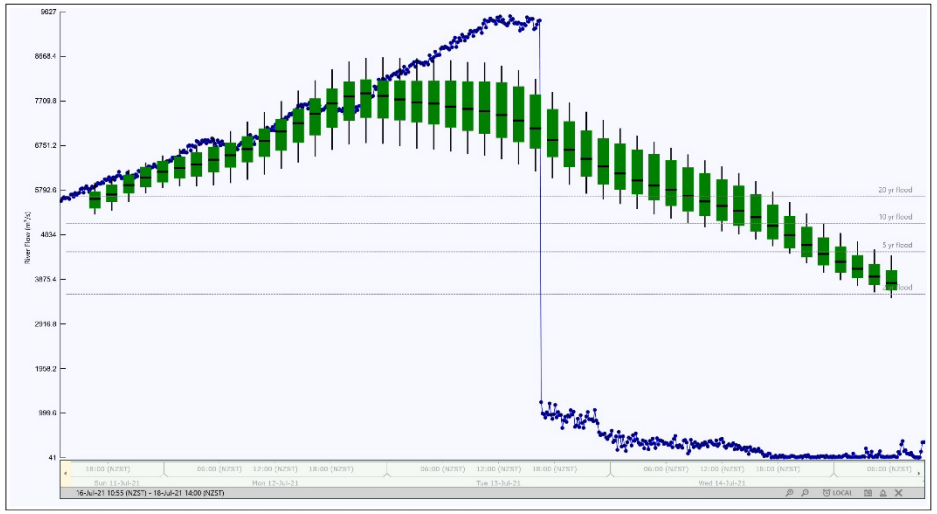


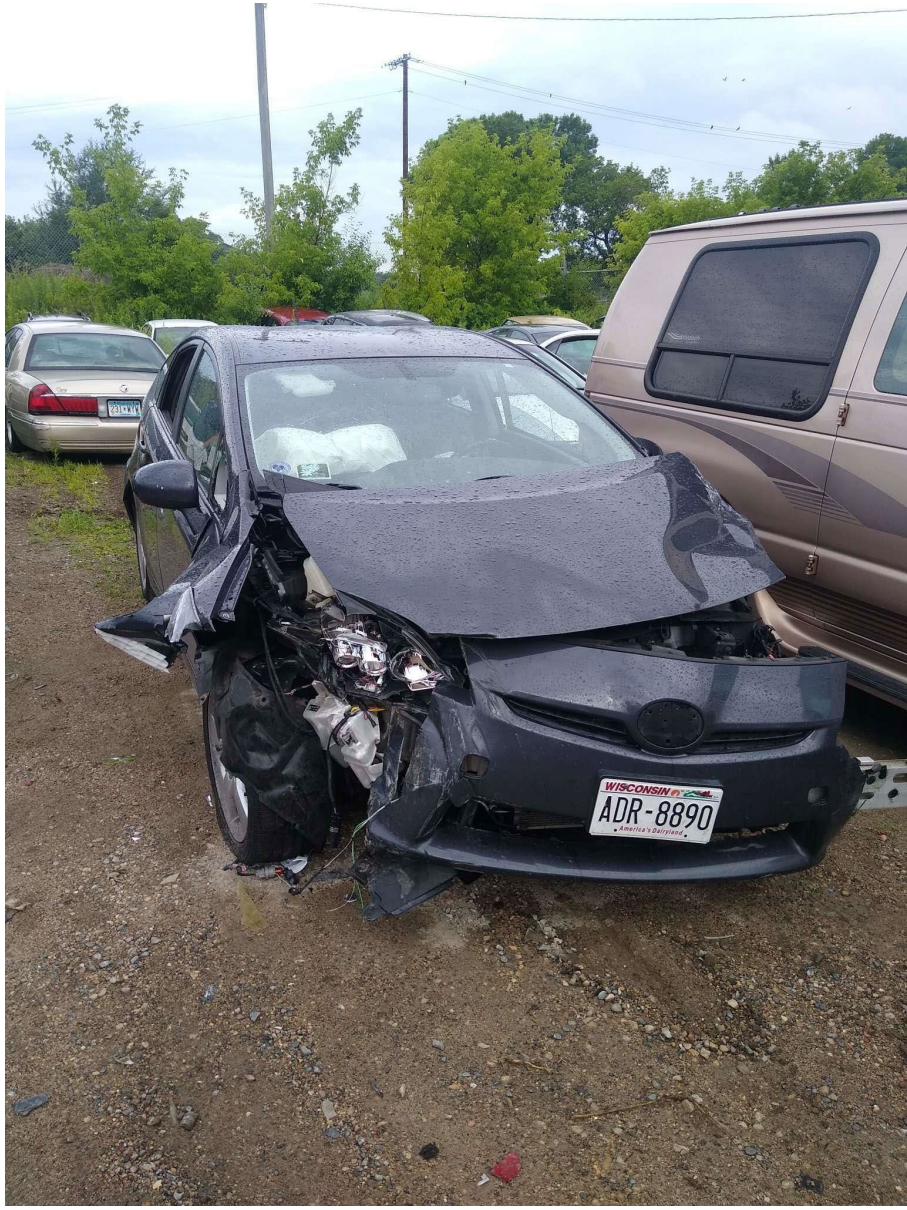

Chart 1
(Printed 10-Aug-2021 15:00)

- Early forecasts tracked well
- 2 out of 3 sites went down

NIWA
Taihoro Nukurangi

# … about my HPC experiences

| | | |
|---|---|---|
| *1977* | *University of Wisconsin* | *Fortran, punch cards* |
| 1985-1991 | Ford Aerospace | CDC7600, Cray-1 user, COS to UniCOS |
| 1991-1999 | Cray Research | first vector to RISC (LS-DYNA) |
| 1999-2014 | IBM | first RISC to X86 (LS-DYNA) |
| 2014-2019 | Lenovo | first Dragonfly Plus (Toronto, Canada) |
| 2019+ | NIWA | *Fortran (!), HPC refresh 2024-2025* |



# … about the HPC 2025 refresh at NIWA

- March of 2025 is the end of year #7 of the systems installed in 2017, accepted in 2018
- RFP active now, goal is selection/order before the end of 2023
- delivery & enablement 1H 2024, workloads moved during 2H 2024
- focus on “data first”
- both ‘hosted’ and ‘public cloud’ approaches welcome

1. new archive (primary & secondary)
2. new compute & high-performance storage (primary & secondary)
3. new hosting location (primary & secondary)

*… and about that abstract/title:*

*Exploration of public cloud computing by an operational site running the Unified Model*

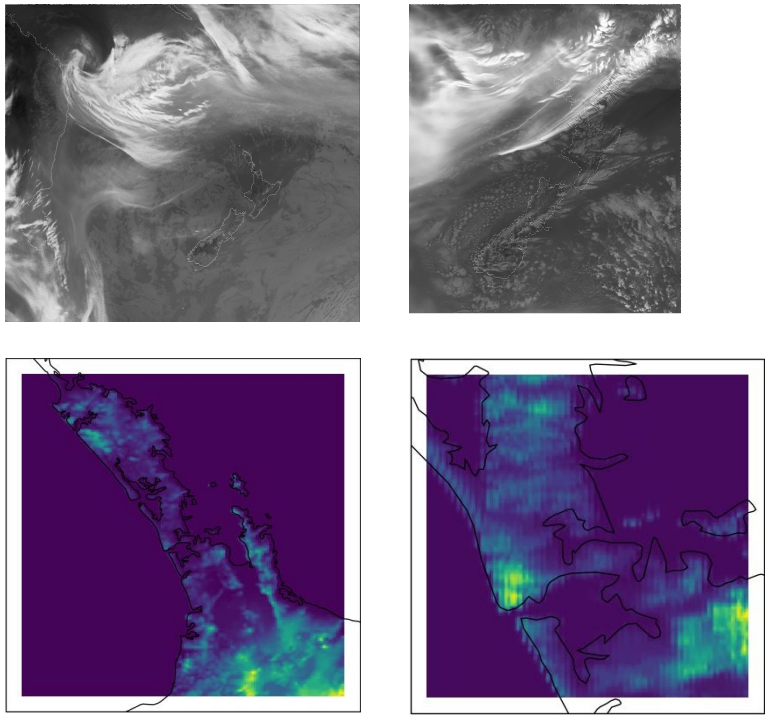
"No person ever steps in the same river twice, for it's not the same river and they aren't the same person"

*Heraclitus (544-483 B.C.)*

NIWA
Taihoro Nukurangi

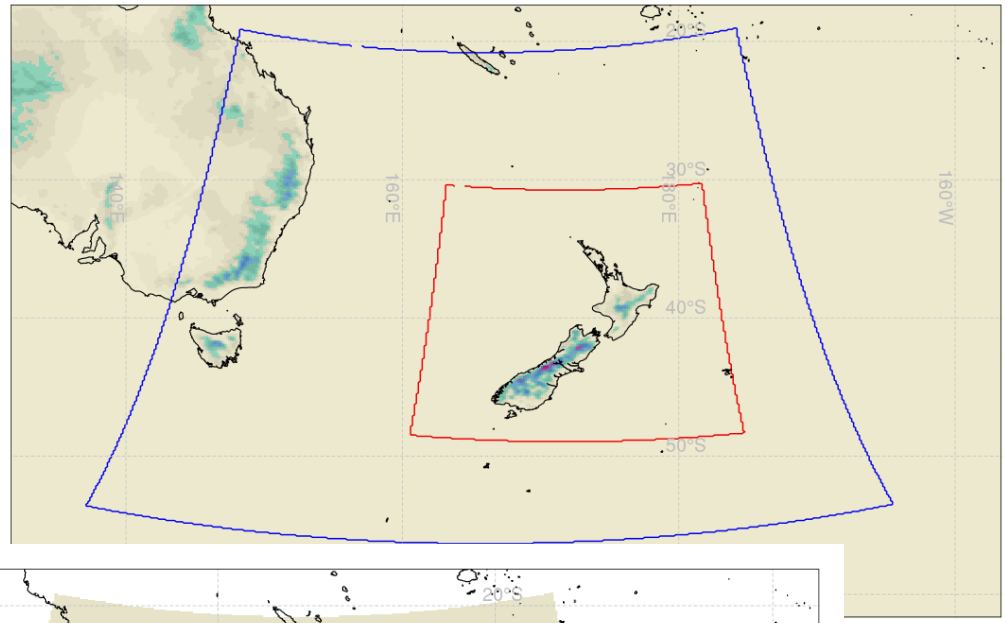
# NIWA Operational Weather Forecasting

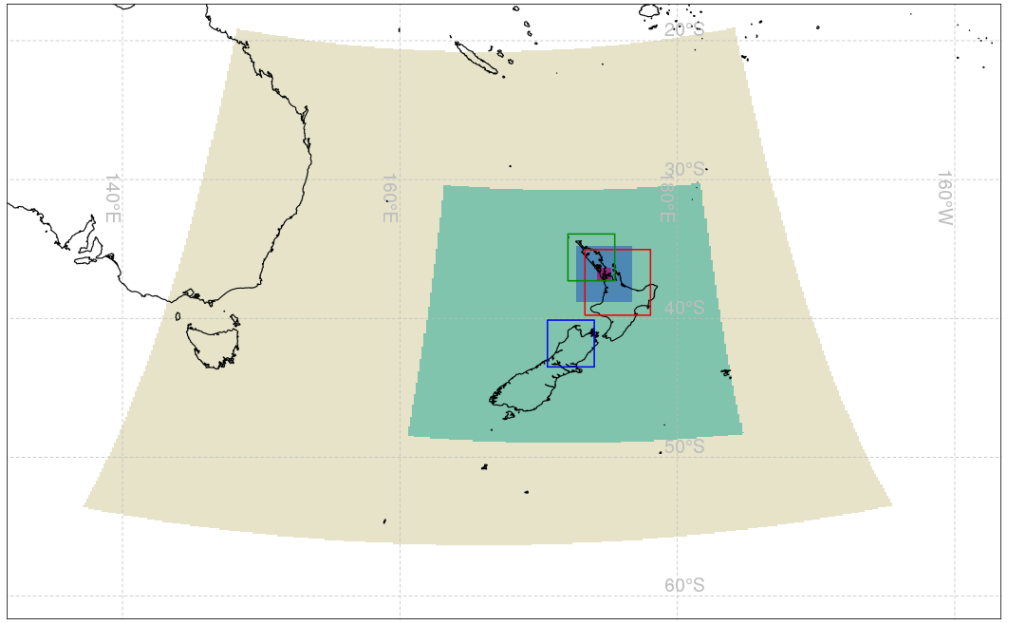
- Based on the Unified Model (UK Met Office)
- Currently 4 operational model configurations
- Ensemble forecasting now in production
- Additional models (e.g., wave and tide forecast)
- Complex workflows with numerous pre- and post-processing tasks



# Recent (and future) growth in compute demand

| | year | forecasting operational workload |
|---|---|---|
| *12 km NZLAM* | **2018** | 4,511,400 |
| *+4 km NZLAM* | **2019** | 6,039,533 |
| *no significant adds* | **2020** | 6,039,533 |
| *+detailed model* | **2021** | 7,061,533 |
| *+ensemble* | **2022** | 15,733,933 |
| *no significant adds* | **2023** | 15,733,933 |
| *next gen* | **2024** | 183,421,260 |
| *next gen* | **2025** | 183,421,260 |






NIWA
Taihoro Nukurangi

# Is public cloud possible? What do vendors offer?

| Service | AWS | Azure | Google | Oracle |
|---|---|---|---|---|
| **Compute** | EC2 c5i, hpc6a, … | HBv2, HC, … | C2, C2D, … | BM.HPC2.36 |
| **Network** | Elastic Fabric Adapter | HDR InfiniBand | Low-latency Ethernet | 100 Gb/s RoCI v2 |
| **Storage (object + high perf)** | S3 and Lustre | Blob, NetApp | Cloud Storage and Lustre | marketplace: BeeGFS, Lustre, GlusterFS, … |
| **Configuration** | ParallelCluster | CycleCloud | Cloud HPC Toolkit | Terraform |

- HPC services also offered by other vendors (HPE, …)
- Fairly similar concepts across vendors
- Config and provisioning can usually be automated

# General impressions – from the proof of concept

- It works ☺
- *Well, for one vendor, it was a bit of a struggle (measured in days) to get a version running.  For another, it was just a few hours.*
- Performance results are encouraging – competitive compute and IO performance, straightforward to use, very similar to on-prem
- Great strong scaling
- Small ssh lag with Australia-East region, workable with US-East region
- Great tools, e.g., AWS cluster CLI or Azure cloud portal
- “DIY HPC” – build custom solution in minutes

# Challenges and open questions

- Data has gravity, need to choose data centre carefully
- Cost - fast storage is expensive
- Move from persistent resource use to ephemeral use – lots of scripting, need to educate users (!)
- (Un)availability of compute resources – reserve, migrate, …?
- Tailor HPC solutions to each project? Let some projects share?

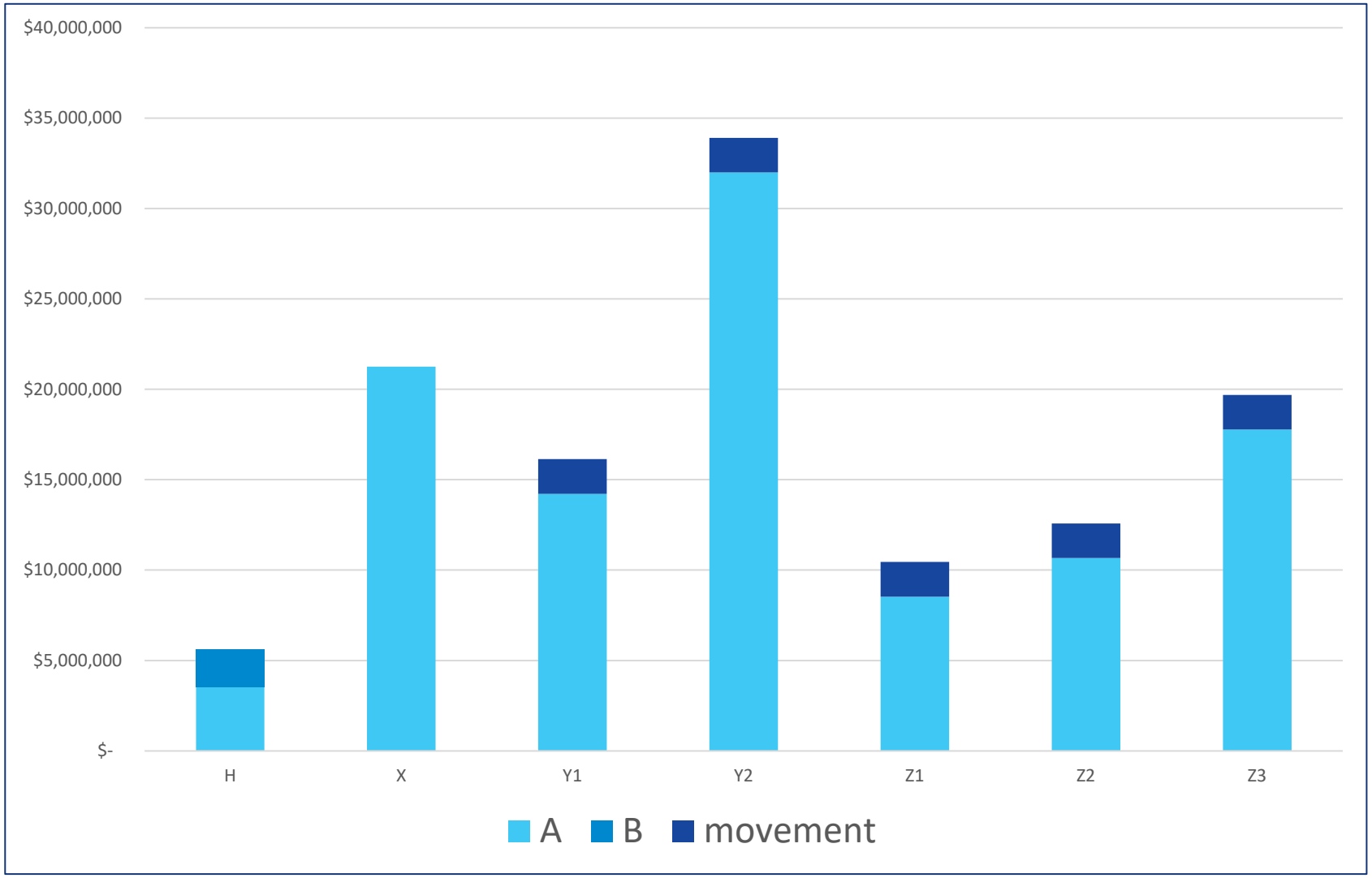
# Public cloud performance summary

- Encouraging results for this Unified Model forecasting pilot
- Relatively straightforward to use, good tools
- Need to embrace automation and careful data staging
- Storage needs a lot of care – what, how, where to transfer and store?

# Cost comparisons

- *old thinking:* cloud is more expensive now, but will eventually be lowest because of the advantages high volumes
- *new thinking:* cloud will always be more expensive, but will have certain advantages which are worth the $$$
- many opinions on how *much* more expensive, factors such as 3x, 4x, 7x, 7x, 10x all mentioned
- so many options – guidance (internal & external) is really required

# Archive solution early cost estimates




- 6 year cost
- 18 petabytes, growing to 150+

*Location matters!*

- 1.2 = USA
- 1.5 = Singapore
- 2.5 = Australia
- ??? = Auckland



NIWA
Taihoro Nukurangi

# Is operational weather possible in the public cloud?

**NIWA**

*Greta Point campus*

**MetService**

*Kelburn campus*

*two approaches have evolved*


Thank you
NIWA
Taihoro Nukurangi