Australian Government
Bureau of Meteorology

# Supercomputer Programme

A seven-year programme to enhance the computational and numerical prediction capabilities of the Bureau's forecast and warning services.

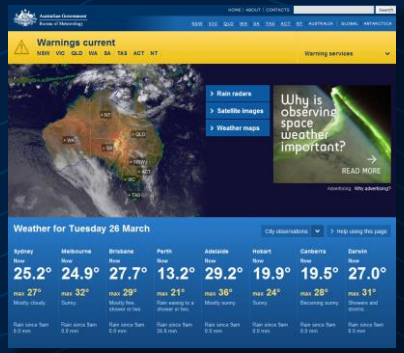
*Tim Pugh, Lesley Seebeck, Tennessee Leeuwenburg, Andrew Khaw, Warwick Howland*

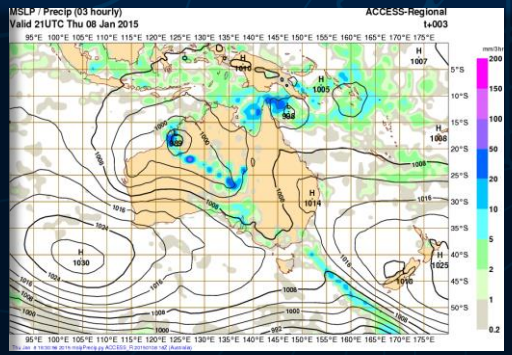
*Bureau of Meteorology, Melbourne, Australia*

*Wednesday 26 October 2016*

# Reliable, resilient, national capability

- National/Global observing system: atmosphere, marine, water, land, space
- 24/7 Operational forecasting systems for weather, climate, oceans and flooding
- Supercomputing and massive data storage
- High uptime internet communications and disaster recovery
- Professional forecasting capability across multiple disciplines
- Experts out posted in the Australian Defence Force, State Emergency Centres and Aviation Operation Centres




# Establishing the Supercomputer Programme

## New funding announced by Australian Government in May 2014

Seven Year Programme from July 2014 to July 2021

- Funding for Supercomputer system, Supporting Data Processing and Storage systems, Data Centre and Networks, and Numerical Prediction Project (Transitions to Operations)

Programme Investment Areas across People, Processes, Science and Technology

» Benefit Planning and Realisation (Supercomputer and Services Board)

- Investments, Priorities, Delivery and Schedules, Social Economic Value, Return on Investment

» Infrastructure (Information Systems and Services)

- Data centre, networks, HPC and Data Intensive Computing, Software Services, Suite and Job Scheduling, UM Modelling Infrastructure, System and Application Monitoring

» Delivery (Science to Services)

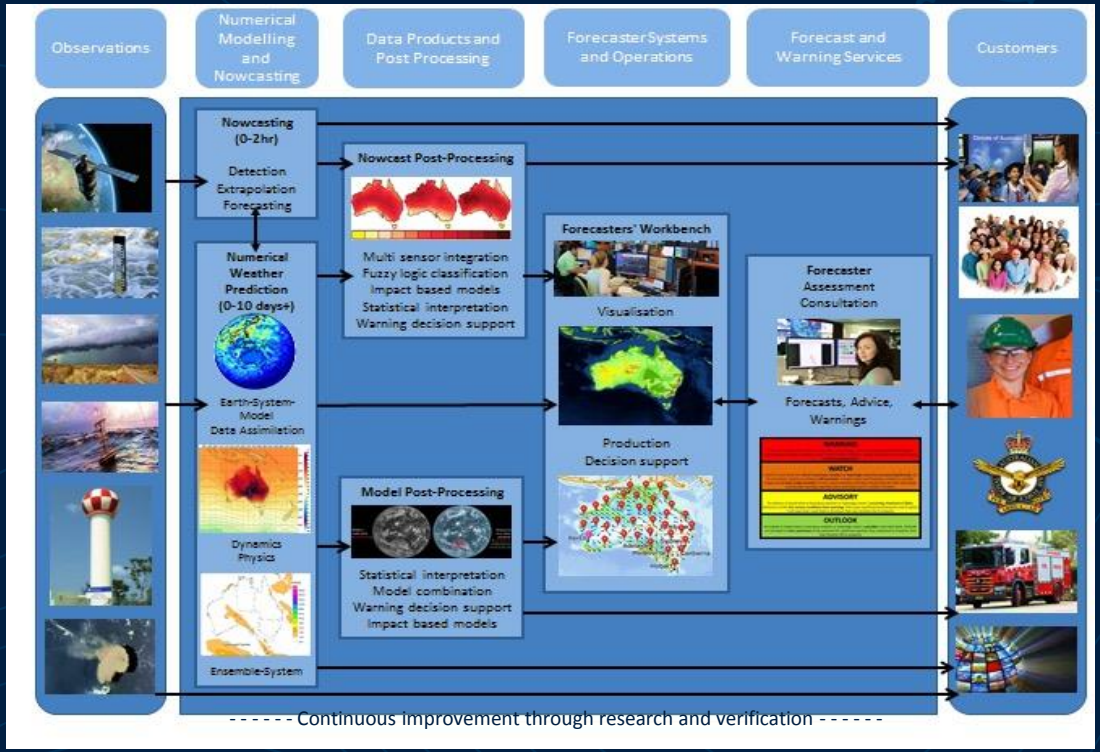
- Scientific Computing Service, Model Build Team, Numerical Prediction, Guidance Post Processing, Model Data Services, Software Lifecycles, Verification Frameworks, Software Services

» Scalability (Research and Development)

- Future architectures, Growth in Compute and Data, Software Engineering, Skills

# Forecast Production Value Chain

Investments and Outcomes




More accurate - particularly for the location, timing and direction of rainfall, storms and wind changes

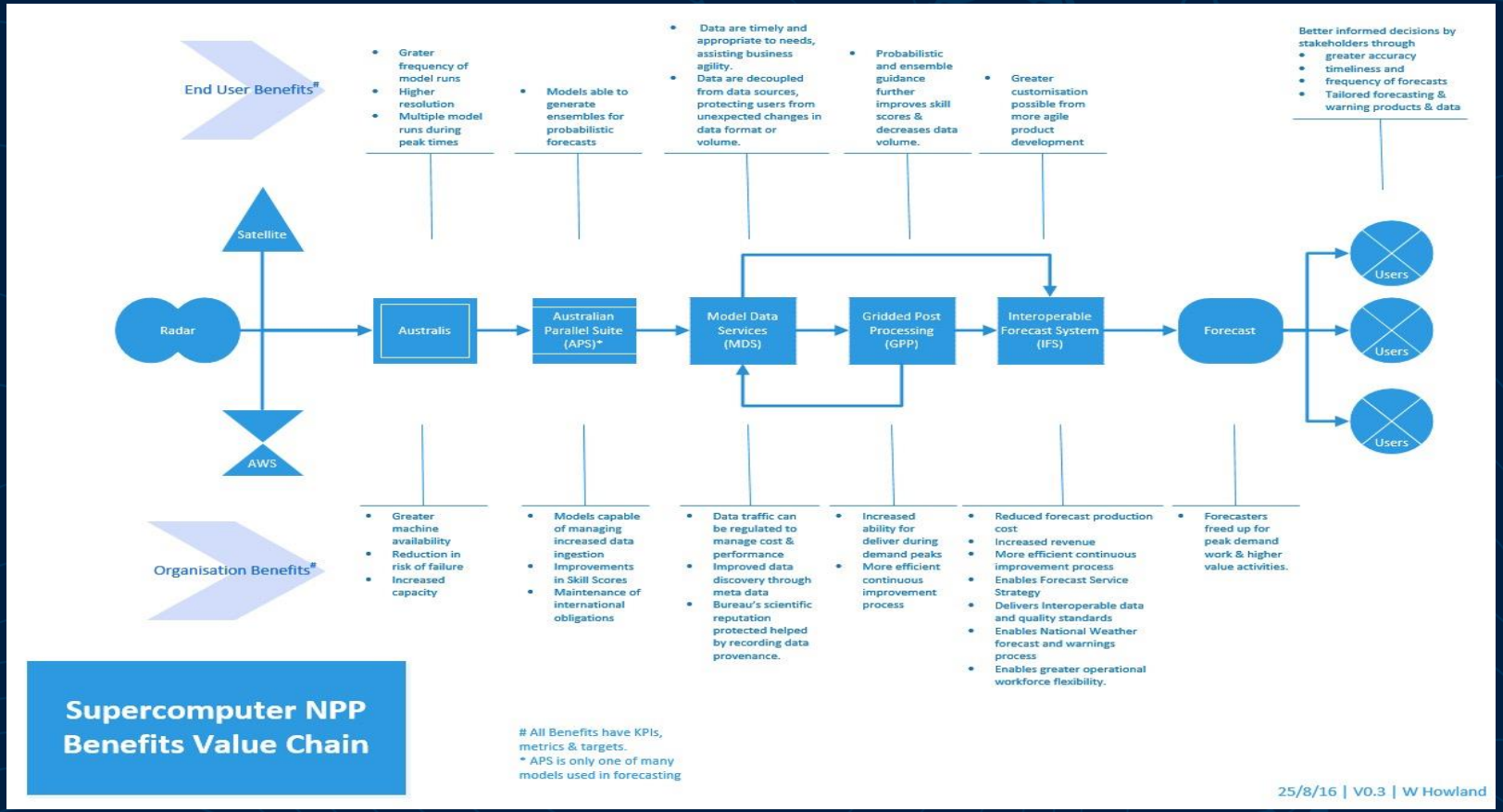
More up-to-date - more frequent forecasts available

More valuable - for decision makers, by quantifying forecast outcome probabilities using ensembles

More responsive - through capability to produce additional, on-demand, detailed forecasts for multiple extreme weather and hazard events across Australia.

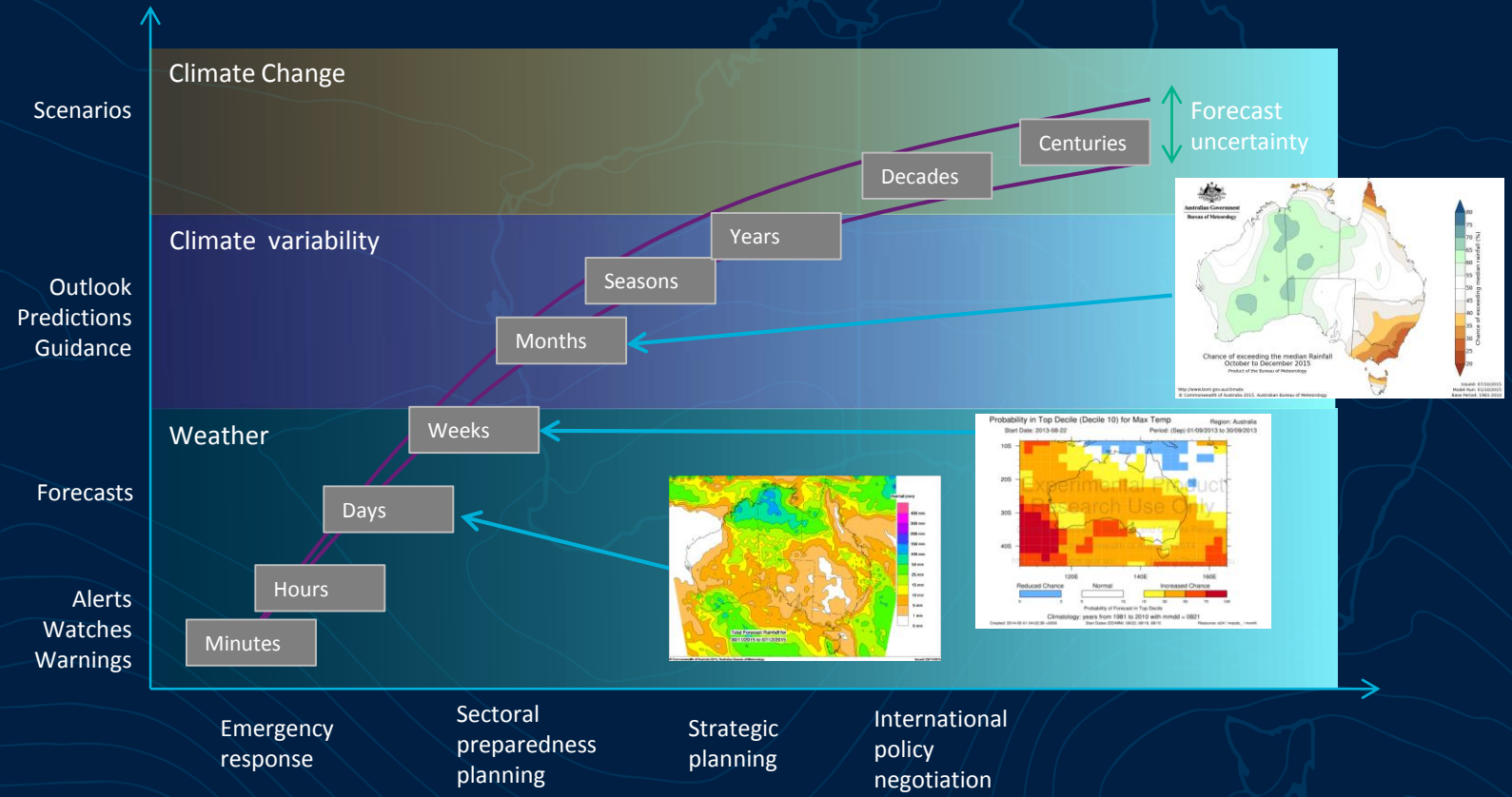
Australian Government
Bureau of Meteorology

# Key Projects in Value Chain - Benefits



Australian Government
Bureau of Meteorology

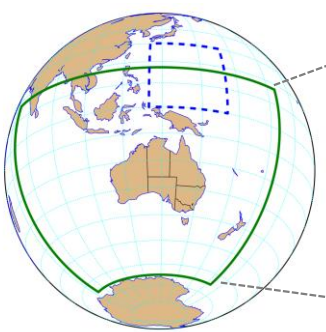
# Environmental Modelling in the Bureau

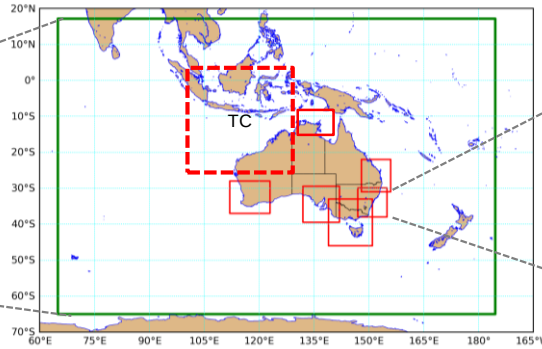


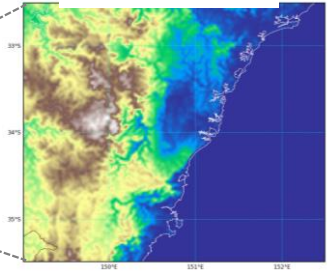
# Modelling to Achieve Outcomes

| **Capability** | **2014 HPC system** | **New HPC systems (2016 to 2021)** |
|---|---|---|
| *Model grid resolution (horizontal only)*<br>ACCESS-G (global)<br>ACCESS-R (regional)<br>ACCESS-C (city) | <br>40 km<br>12 km<br>4 km | <br>25 km > 10 km<br>12 km > 4.5 km<br>1.5 km |
| *Regular forecast updates (times per day)*<br>Global<br>Regional<br>City and on-demand | <br>4<br>4<br>4 | <br>4<br>8<br>Up to 24 |
| Tropical cyclone forecasts<br>*(horizontal grid resolution)*<br>*(forecast length)* | <br>12 km<br>Out to 3 days | Up to 3 concurrent events<br>12km > 4.5km<br>Out to 5 days |
| Ensembles Forecasts<br>*(Certainty for decision makers)* | None | Yes (Global, City, TC, Relocatable) |
| Capability to produce additional, on-demand, high-resolution forecasts for extreme weather | None | 1.5 km<br>Up to 4 concurrent events<br>Up to 24 times per day |

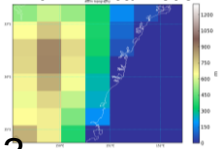
# Numerical Weather Prediction Roadmap



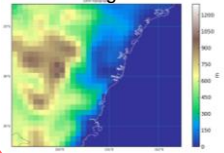


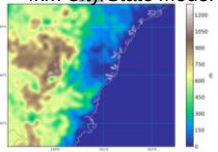

2 x daily 10-day & 3-day forecast
40km Global Model



4 x daily 3-day forecast
12km Regional Model



4 x daily 36-hour forecast
4km City/State Model



Increasing model resolution for improved local information

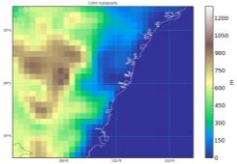
Future model ensembles for likelihood of significant weather

2013

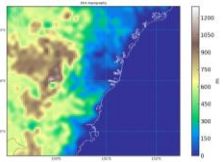
**Model Topography of Sydney, NSW**

2020

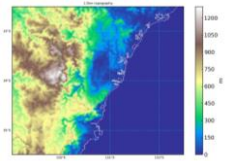
2 x daily 10-day & 3-day forecast
12km Global Model



8 x daily 3-day forecast
4.5km Regional Model



24 x daily 18h or 36h forecast
1.5km City/State Model



Projection of Nominal Modelling Resolutions for Future Computing Systems

## High impact weather modelling of TCs, Aviation, Severe TS, Rainfall, Bushfires, etc.

To improve our understanding of and ability to predict severe weather, including for bush fires, tropical cyclones, severe thunderstorms and heavy rainfall, through the use of high-resolution modelling in conjunction with available observations

A requirement for operational NWP is to provide outputs that meet societal needs, and severe weather prediction is a key deliverable

- High resolution grids to simulate convective-scale high impact weather and hazards
- Rapid update cycles for emergency response
- High resolution ensembles for probabilistic forecasting – Certainty of prediction

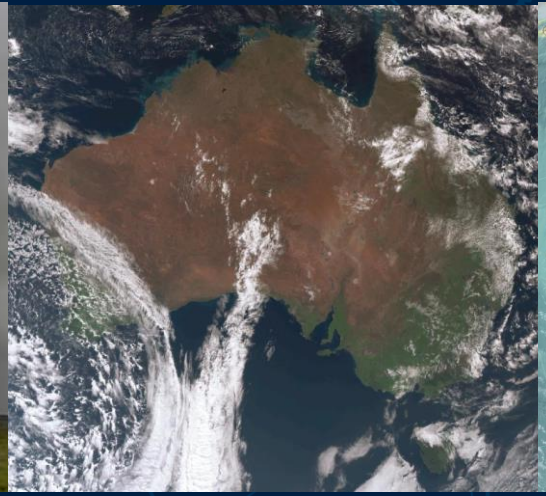
# Australia Extended-Range TC Forecasts

- **Main aims:**
  - Early detection of conditions for cyclone genesis
  - Forecasts to 5 days
  - Better forecasts of cyclone genesis and intensity
  - Wave forecasts under TCs
  - Storm surge prediction and impacts to coastal communities

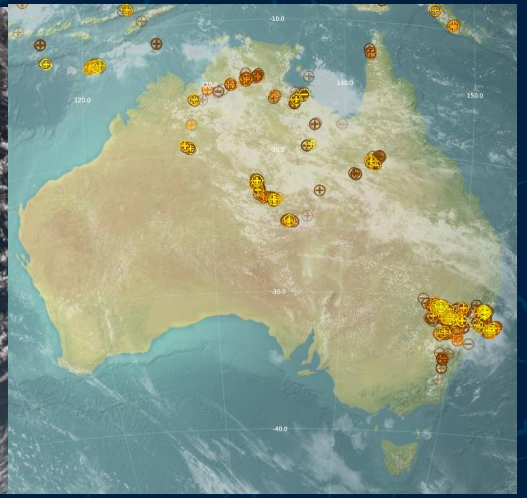
- **In development:**
  - NW Australia region
  - Grid resolution increase from 12 km to 4.5 km
  - Simulate TC events from 3-day to 5-day forecasts
  - 4D-Var with vortex specification (bogussing)
  - Early detection, daily running, not just when a cyclone is named
  - One-way coupling to wave model

# Remotely sensed observations

Advanced radar – more accurate rain and hail



New satellites – greater spatial detail, more frequent updates



Lightning network – weather watch in areas not covered by radar

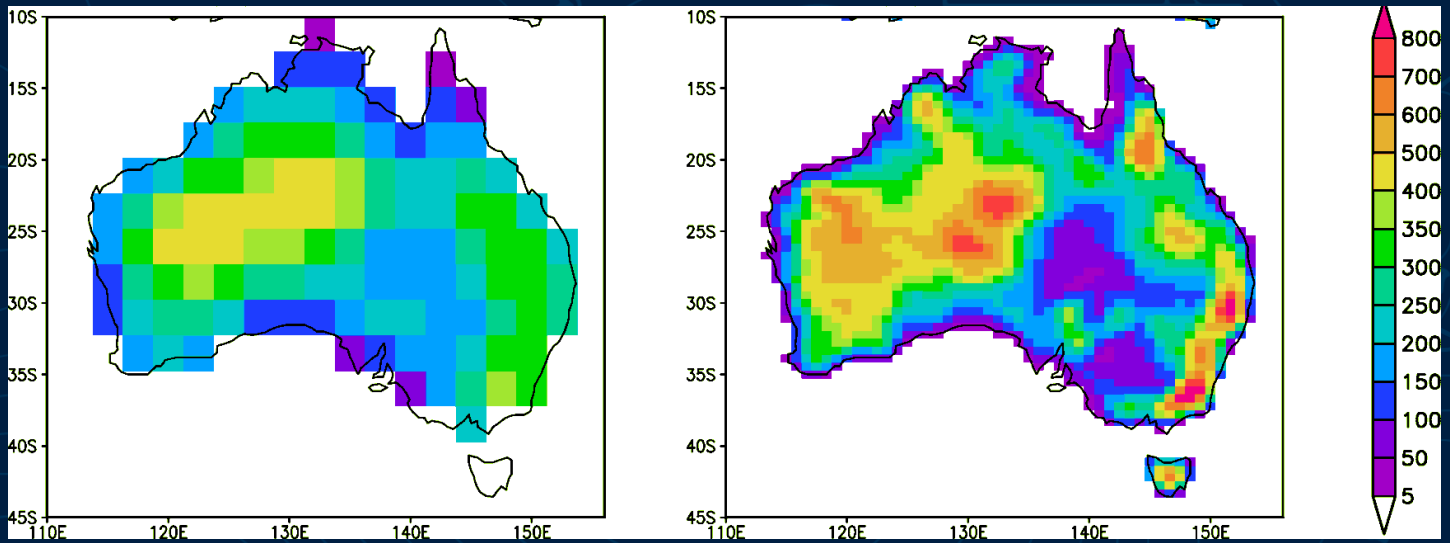
Observations from remote sensing networks let us see what weather is happening everywhere in Australia. They also provide essential input for weather forecast models.

# ACCESS Multi-week to Seasonal prediction

- A new, improved (better physics), high-resolution multi-week to seasonal forecast system
  - 250km **(POAMA2)** → 60km (25 km ocean) **(ACCESS-S)**
  - 17 levels → 85 levels
- Better estimates of uncertainty and risk

Current 250 km system

New 60 km system



Australian Government
Bureau of Meteorology

# BETTER SEASONAL OUTLOOKS

## Finer model detail

Moving from 250 km to 60 km resolution

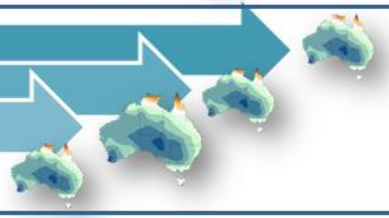
Australia:120 to 2000 grid points

*meaning*
more localised information
*by accounting for local conditions*

## More outlook periods

Seamless:
*filling the gap between 7-day and monthly outlooks*

Season
Month
Fortnight
Week



Outlooks updated weekly

## Higher outlook skill

Likely 10% improvement in outlook accuracy

*meaning*
the best outlooks for Australia
*of all international models*

## World class service

*meaning*
information is clear, concise and available when and where you need it

**More intelligence possible:**

- Evaporation
- Humidity
- Wind
- Drought
- Extremes
- Tropical Cyclones

Not only rainfall and temperature

## Bigger user returns

**Reduce losses:** agricultural production lost from 2010-11 La Niña:

**More than $2 billion**

ABARES

Potential value of improved seasonal forecasts:

**More than $1 billion per year**

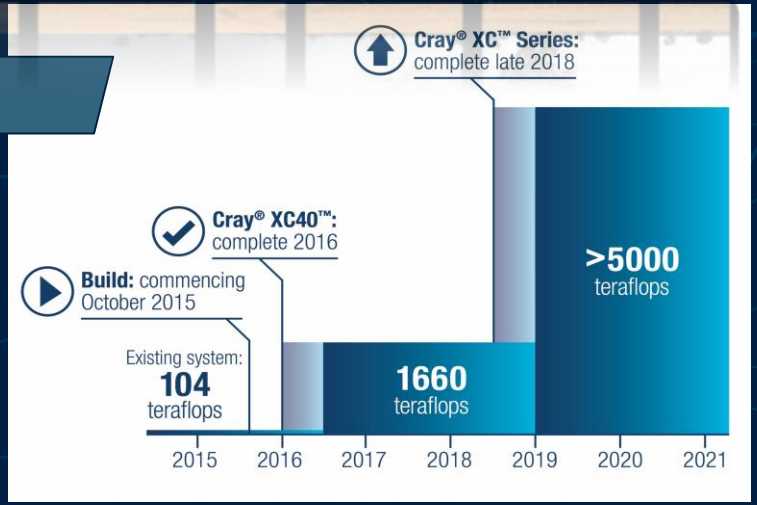
Centre for International Economics 2014

# AUSTRALIS, the Bureau's Production Supercomputer


## Supercomputer details

**$59** MILLION (USD) HAS BEEN ALLOCATED FOR THE PROJECT

**CRAY Inc.** WILL SUPPLY THE NEW SUPERCOMPUTER



# Australis Production System

## 38x performance, 8x electrical power

| | 2015 Australis (delivered) | 2018 Addition (projected) | 2018 Australis (projected) | Ngamai HPC System (Retired Oct'16) | Relative Increase |
|---|---|---|---|---|---|
| Processor | Intel Xeon Haswell 12-core, 2.6 GHz | Intel Xeon Skylake | Intel Xeon Haswell + Skylake | Intel Xeon Sandy Bridge 6-core, 2.5 GHz | Increase relative to Ngamai System |
| Nodes | 2,160 + | 1,952 = | 4,112 | 576 | 2015: 3.8x<br>2018: 7.1x |
| Cores | 51,840 | 78,080 | 129,920 | 6,912 | 2015: 7.5x<br>2018: 18.8x |
| Aggregate Memory | 276 TB | 375 TB | 651 TB | 36.9 TB | 2015: 7.5x<br>2018: 17.7x |
| Usable Storage | 4,320 TB | 4,320 TB | 8,640 TB | 214 TB | 2015: 20.2x<br>2018: 40.4x |
| Storage Bandwidth | 135 GB/s | 171 GB/s | 306 GB/s | 16 GB/s | 2015: 8.4x<br>2018: 19.1x |
| Sustain system performance (SSP) | 253 | 365 | 618 | 16 | 2015: 15.6x<br>2018: 38.1x |
| Typical Power Use | 865 kW | 783 kW | 1,648 kW | 200 kW | 2015: 4.3x<br>2018: 8.2x |

# Infrastructure Investments

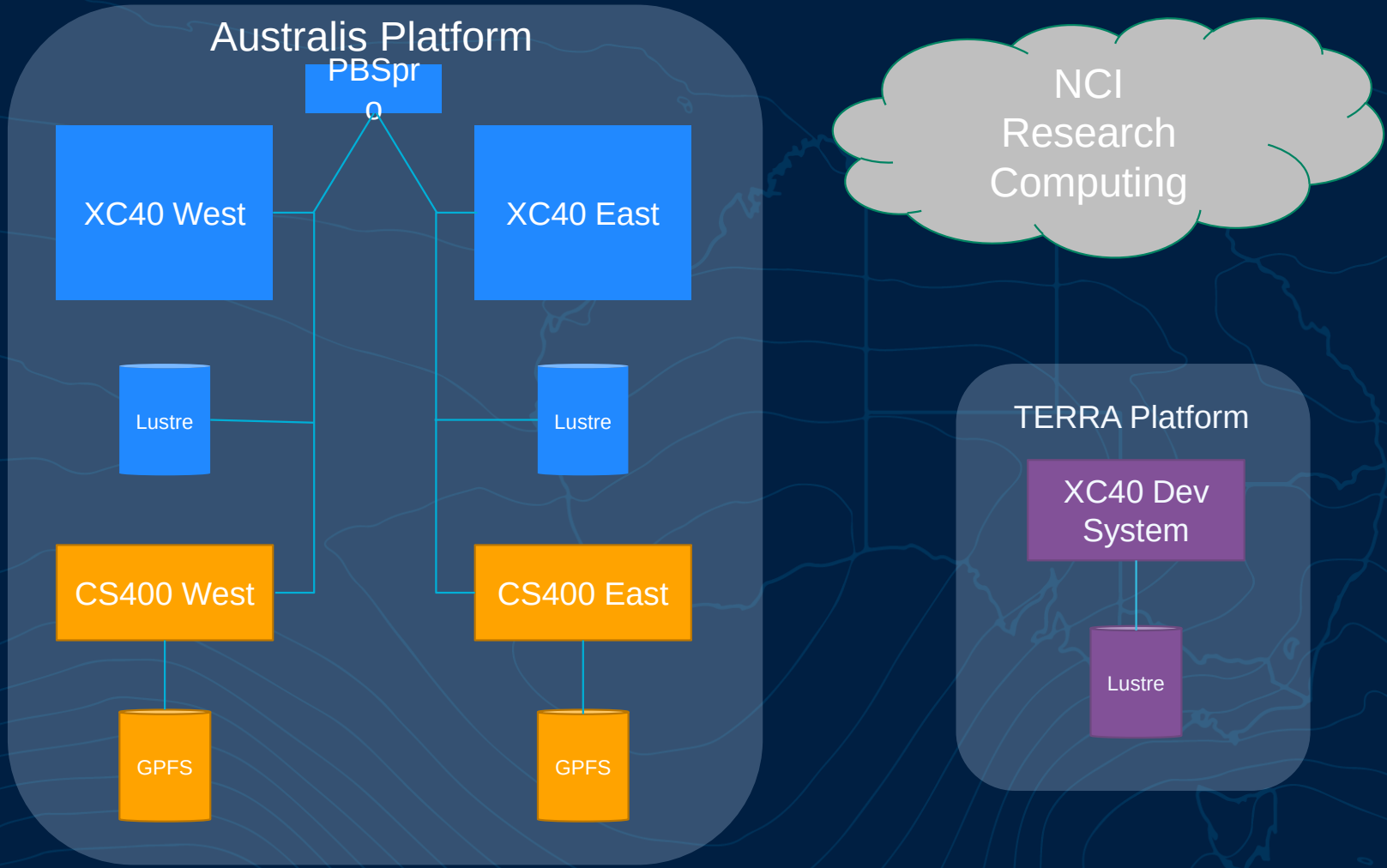
Addition of small Cray XC40 dev system “TERRA” to support delivery pathway from NCI to Production

- 144 Nodes, 3456 Intel Haswell cores, 34TB memory, 1.44 PB Lustre data storage
- Addition of six Cray DataWarp I/O nodes (48TB) for workflow optimisation
- Addition of NVIDIA GPUs and Intel Xeon Phi planned
- 2018 TERRA Cray upgrade (projected): 321 nodes, 10536 cores, 2.88 PB Lustre data storage

Integrated CS-400 data processing system with 40 compute nodes to AUSTRALIS

- Dual socket nodes with Intel 18-core Broadwell, 256GB DDR4, FDR Infiniband interconnects
- Addition of 1.6 TB Intel NVMe flash on all compute nodes and a handful of NVIDIA K80 GPUs
- GPFS data storage system based on DDN GS14K system with 150TB SSD & 2PB HDD storage.

Commercial data centre hosting Cray XC40 (~860 KW) and later, XC+ system (~1650 KW) and network links


Australis Platform
PBSpro
XC40 West
XC40 East
Lustre
Lustre
CS400 West
CS400 East
GPFS
GPFS
NCI
Research
Computing
TERRA Platform
XC40 Dev
System
Lustre

Australian Government
Bureau of Meteorology

# Supercomputer Programme Challenges

- Improving Delivery (time, quality)
- Reducing Complexity (decoupler)
- Improving Software Engineering
- Scaling infrastructure for next generation

# Improving Delivery

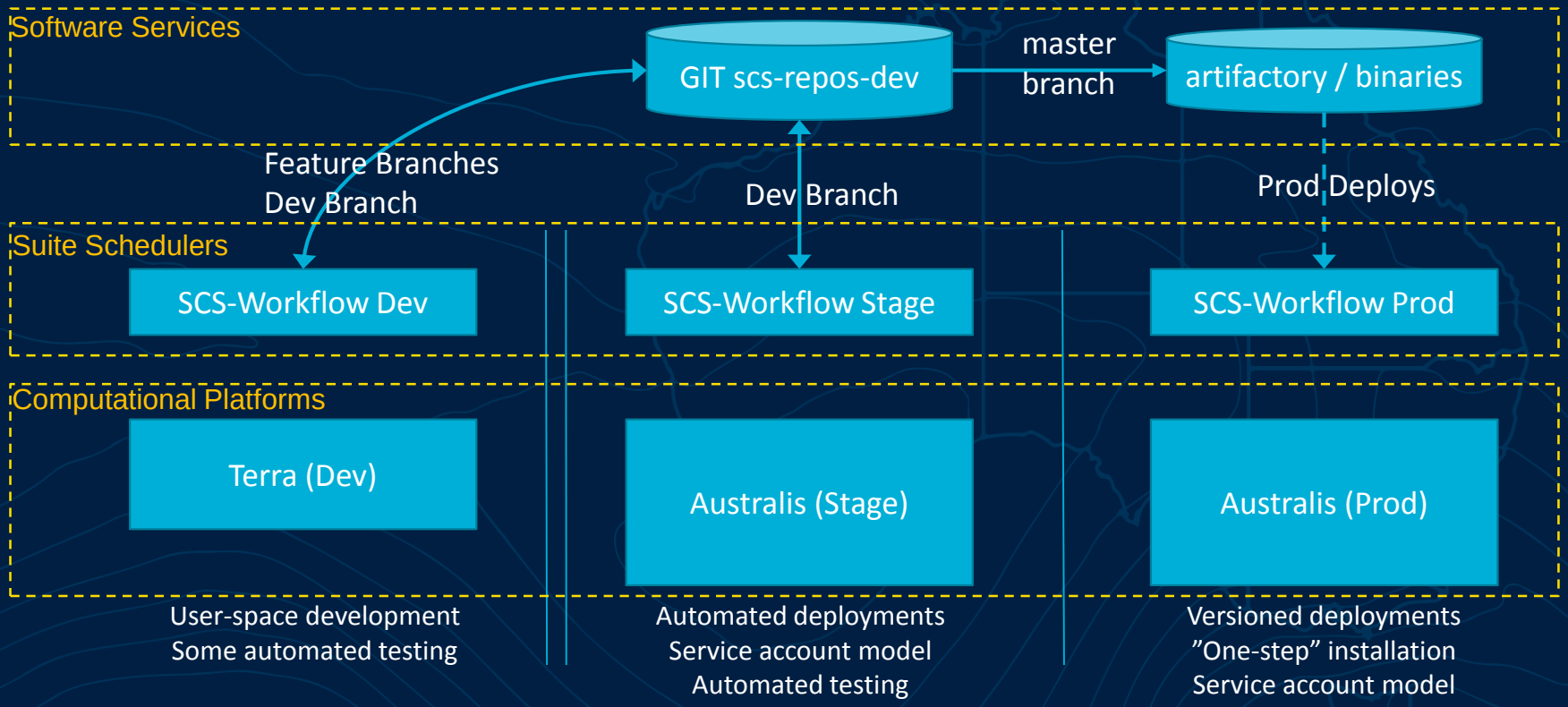
## Dev / Stage / Prod  Packaging

- Systems are “installable” through simple scripts
- Versioned binary components are deployed to the supercomputer
- Versioned scheduler scripts are deployed to the workflow machines
- Deployments are 100% hands-off automated in the staging environment

## Benefits

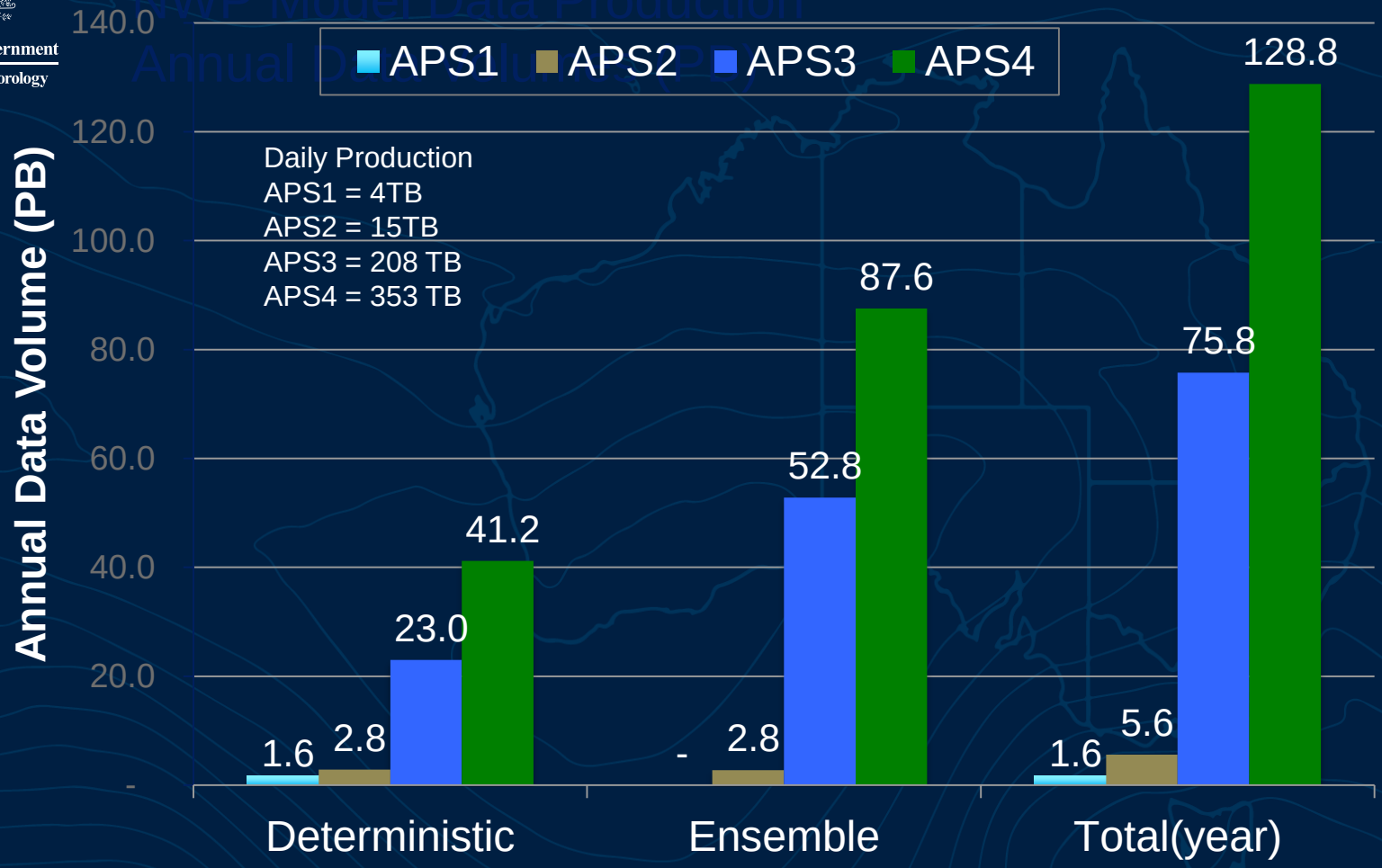
- Cuts down on issues around configuration of modelling systems
- Promotes standard approaches to system packaging and deployment
- Provides strong integration testing
- Reduction in manual effort for testing and deployment
- Reduces errors by forcing repeatable processes
- Reduces manually-introduced errors
- Increases the cost of preparation slightly, but significantly reduces cost of production deployments
- Provides the foundation for end-to-end automated integration testing
- Provides the possibility of hands-off production deployment

# Achievements in Modelling Automation



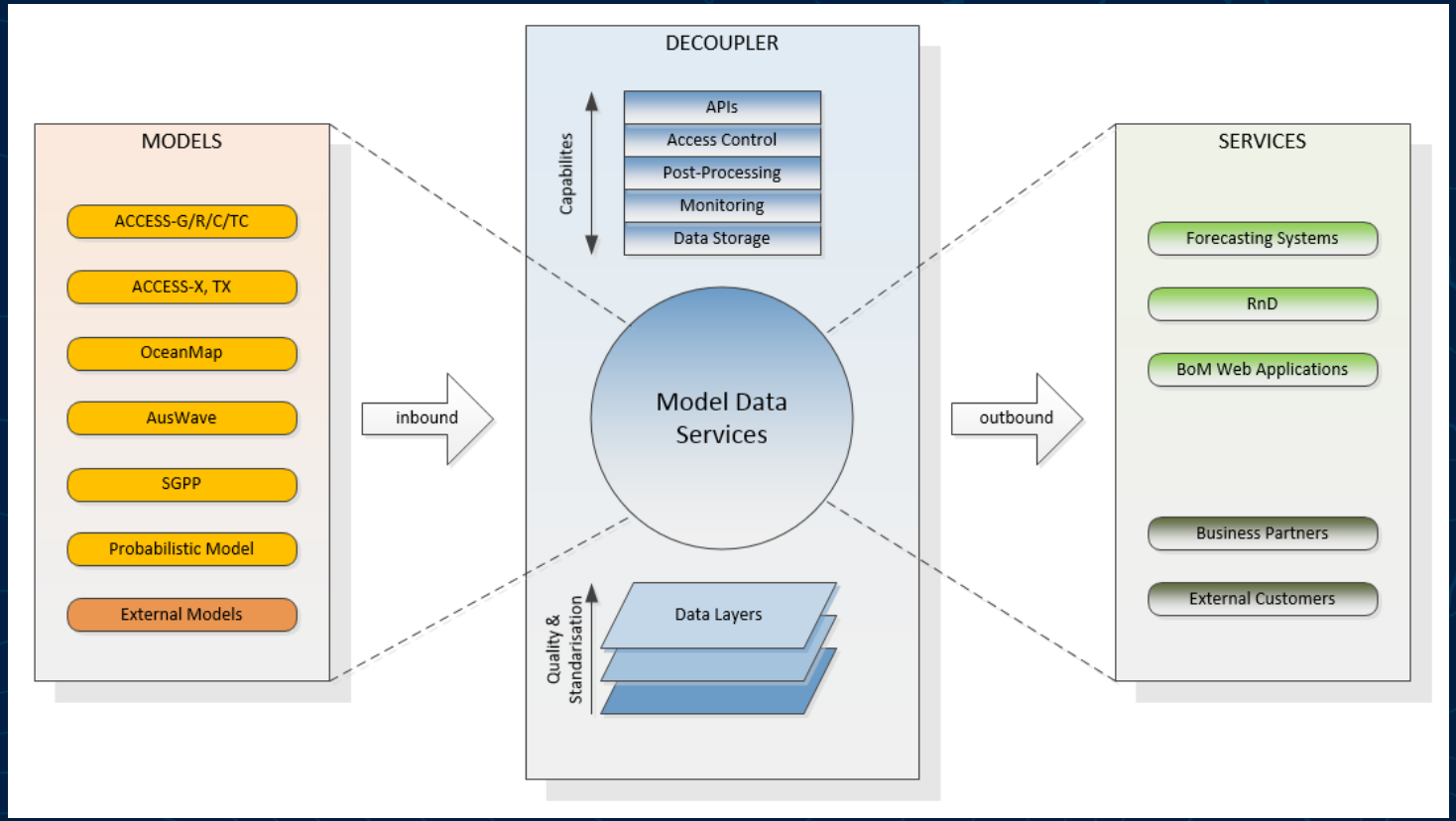
Australian Government
Bureau of Meteorology

# Australis Data Production (not storage)



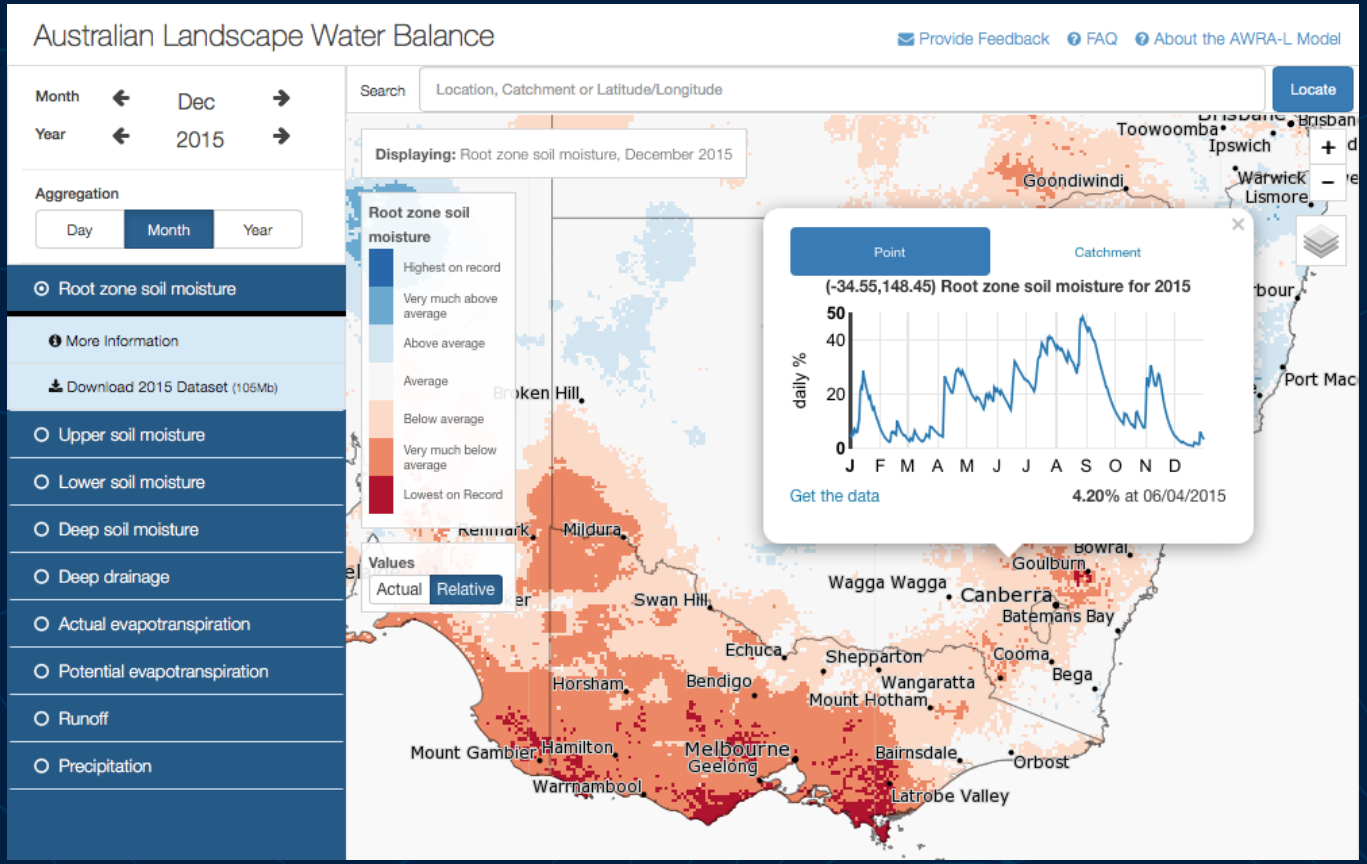
Australian Government
Bureau of Meteorology

# HPC Decoupler



# AWRA-L Web Application

- A unique service!
- Updated daily
- See all variables at daily, monthly or annual time slices
- Download the grids at a resolution of 5km x 5km
- Past 10 years data available to all
- Registered users access >100 years and tailored products




http://www.bom.gov.au/water/landscape

# Scalability Challenges

## Future architectures

- Compute processors (multi-core, many-core, memory pools, FPGA?)
- Data Intensive infrastructure (accelerators, networks, data storage cache)
- Programming paradigms including Hybrid, PGAS, Machine Learning, …
- HPC Clouds, containers and VMs – Dynamic allocation of resources and workflows

## Growth in Compute and Data

Only limited by what we can afford and manage…

This is where we focus on:
- Managing the data
- Improving workflows
- Accelerating outcomes

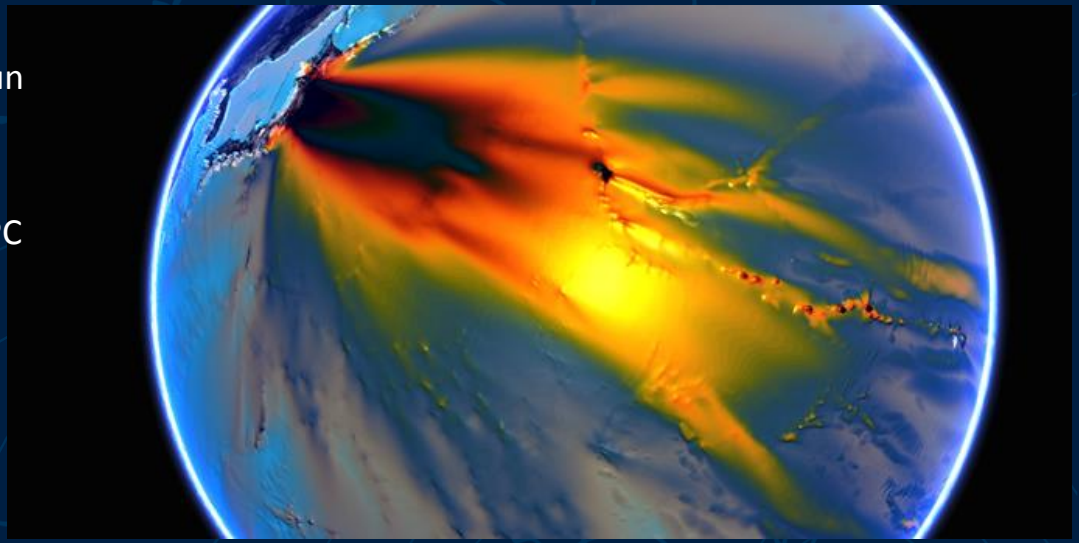
## Modernising Software - Design and Engineering

- Multi-threaded and hybrid parallelism (OpenMP, MPI) in applications
- Task Parallelism in workflows
- Data I/O programming (parallel, API access)

# Tsunami Events – Modelling Realtime Events

- Currently based on pre-computed scenarios
- Runtime of >60 minutes made it impossible to run a real time simulation during tsunami event
- ---
- Performance improvement in 6 weeks by two HPC programmers.
- Initial results of 24-hour simulation of tsunami wave propagation

| | |
|---|---|
| Serial code<br>Intel Xeon Haswell | > 3600 sec (53 min) |
| OpenMP, 24 cores<br>Intel Xeon Haswell | 262 sec (~4.4m) |
| CUDA, 1 GPU<br>NVIDIA K80 Telsa | 134 sec (~2.3 min) |
| CUDA, 8 GPUs<br>NVIDIA K80 Telsa | 22 sec (~1/3 min) |



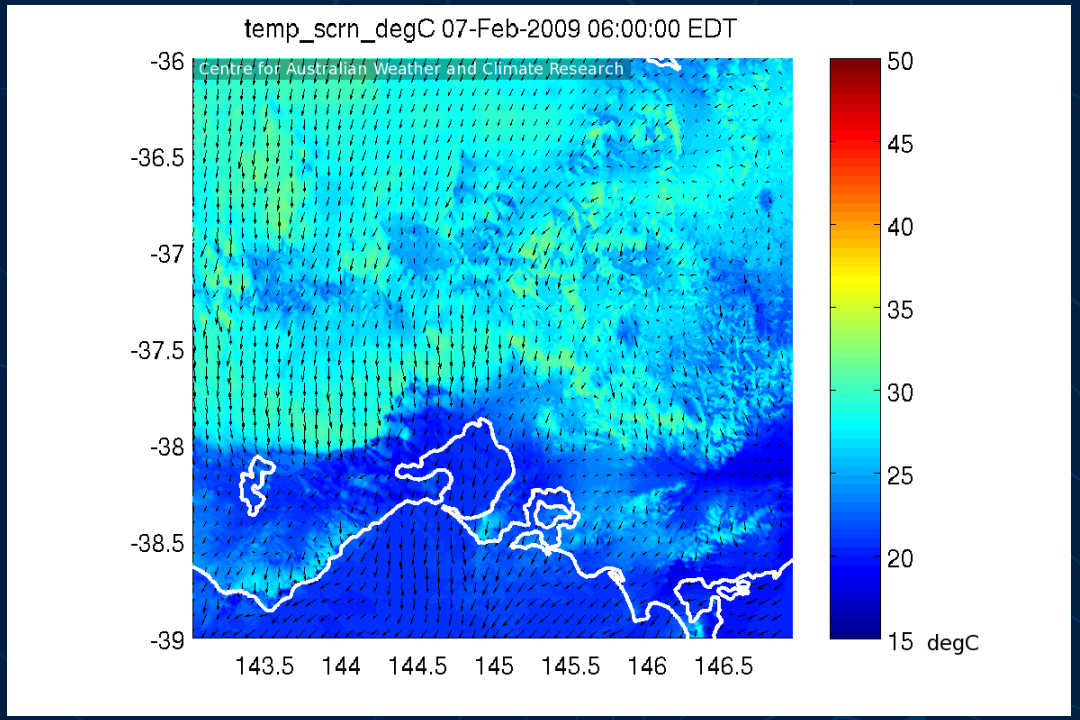
Parallel and GPU version allow on-demand simulation of Tsunami event

- More accurate forecasts of effects
- Ensemble modelling
- Better uncertainty estimation, improved risk map

# Will all this development be worth it?

**Black Saturday**
**7 Feb 2009**

Fire-weather/heat wave




**Model run at 400m**

**J. Kepert, R Fawsett**


# Thank you

## Any questions?

# ENVIRONMENTAL INTELLIGENCE

Conclusions drawn from environmental observations and models to guide decisions and actions by governments, businesses and individuals.

Public Safety

Societal Wellbeing

National Security

Economic Prosperity

Environmental Sustainability

# What increased high performance computing power gives you

## Model view improvements

- Finer grid resolution
- More frequent updating
- Greater resilience to failure

## Service view improvements

| Cyclones | Severe storms | Fires | Flooding rains |
|---|---|---|---|
| - track<br>- intensity<br>- structure | - location<br>- lead time<br>- intensity | - wind speed<br>- wind direction<br>- local detail | - location<br>- volumes<br>- lead time |

## User view improvements

| Industry | Community | Infrastructure | Emergency service |
|---|---|---|---|
| - Fewer false alarms<br>- Shorter closedown<br>- Better decisions | - More timely evacuation<br>- Fewer losses | - Increased preparedness<br>- Less damage | - More effective deployment<br>- Faster recovery |

Australian Government
Bureau of Meteorology

# Modelling Packaging and Automation

## Dev / Stage / Prod Packaging

- Systems are “installable” through simple scripts
- Versioned binary components are deployed to the supercomputer
- Versioned scheduler scripts are deployed to the workflow machines
- Deployments are 100% hands-off automated in the staging environment

## Automation is Based off Git

- Jenkins is the tool which monitors the repository
- New changes are deployed automatically, with any issues reported to developers
- SVN can also be used, but we have used git

## Modelling System Design Guidelines

- Systems must:
  - Compile automatically from a single script or Makefile
  - Deploy compute components (binaries, libraries and supporting scripts) automatically from a single script or Makefile
  - Deploy scheduling components (SMS or Cylc suites) automatically from a single script or Makefile
  - Not rely on having only a single instance (of the model or system) running at once per user
  - Ship with at least 20 automated tests (future goal for all systems post porting)

Australian Government
Bureau of Meteorology

# Modelling Packaging and Automation

## Benefits

- Cuts down on issues around configuration of modelling systems
- Promotes standard approaches to system packaging and deployment
- Provides strong integration testing
- Reduction in manual effort for testing and deployment
- Reduces errors by forcing repeatable processes
- Reduces manually-introduced errors
- Increases the cost of preparation slightly, but significantly reduces cost of production deployments
- Provides the foundation for end-to-end automated integration testing
- Provides the possibility of hands-off production deployment

# What is Jenkins?


- A glorified shell script scheduler OR
- A robot agent that watches you code and gives you feedback

- Builds, tests and compiles reports on your software
- Provides an at-a-glance readiness dashboard of your systems
- A basically ubiquitous industry standard build tool
- A nagging nanny that mainly just picks up typos and silly mistakes

# What is Artifactory?

JFrog Artifactory

- Like version control, but for binaries
- Suitable for large binary files, like topography
- Stores versioned binaries like compiled objects
- Allows you to know what version of code, and what versions of binaries are/were in use
- By testing specific versions of binaries, you can get good confidence that any non-code file changes are also safe to deploy