The Climate and Weather Science Laboratory
an integrated facility for our research community

www.cawcr.gov.au

Tim F. Pugh & Ben Evans
NeCTAR Work Package Leaders

EmPOWER eResearch Conference, Sydney
31 October 2012
Objective:
The virtual laboratory is a new community project to establish an integrated national facility for research in climate and weather simulation and analysis.

Location:
Australian National University’s National Computational Infrastructure (NCI)

Development Organizations:
Australian Bureau of Meteorology (www.bom.gov.au)
Australian National University (nci.org.au)
CSIRO Marine and Atmosphere Research (www.csiro.au/cmar)
Centre for Australian Weather and Climate Research (www.cawcr.gov.au)
ARC Centre of Excellence for Climate System Science (www.climatescience.org.au)

Goals:
• To reduce the technical barriers to using state of the art tools,
• To facilitate the sharing of experiments, data and results,
• To reduce the time to conduct scientific research studies, and
• To elevate the collaboration and contributions to the development of the Australian Community Climate Earth-System Simulator (ACCESS)
Objectives

- To provide a **national facility for climate and weather prediction** in uncoupled and coupled ACCESS model development and simulation.

- To provide a **facility for scientific workflows in climate and weather model analysis**.

- To provide **collaborative content management and data access services**, including a geospatial web-based graphics capability in support of climate and weather research.

- To provide a **data library management and communication system** to integrate and interoperate with national and international climate and weather data catalogues and repositories.

Benefits

- The preparation and run of coupled and uncoupled model experiments within a framework designed with reproducibility, ease of use, support, and sharing of code and experiments.

- A scientific workflow to easily run analyses, share and reproduce analyses, help enforce a controlled vocabulary, and create metadata for traceability and reproducibility while reducing the need for specialist skills.

- Established mechanisms for users to align with managed approaches in simulation, data analysis, and data management to further improve research collaborations and applications.

- Improved data discovery and access to climate and weather repositories through community standard protocols and services.
Names behind the Virtual Laboratory

Work Packages and Leaders

WP1. Model and Simulation Service
WP2. Model Analysis Service
WP3. Data Library & Facility Services
WP4. Website and Integration

<table>
<thead>
<tr>
<th>Name</th>
<th>Location</th>
<th>Position</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Martin Dix</td>
<td>CSIRO, Aspendale VIC</td>
<td>CAWCR ESM IT Leader</td>
<td>Model and Simulation Services Leader of Work Package #1</td>
</tr>
<tr>
<td>Tom Beers</td>
<td>CSIRO, Aspendale VIC</td>
<td>CAWCR CVC Program Leader</td>
<td>Model Analysis Services Leader of Work Package #2</td>
</tr>
<tr>
<td>Ben Evans</td>
<td>ANU/NCI, Canberra, ACT</td>
<td>National Facility Manager</td>
<td>Data Library and Facility Interfaces and Services Leader of Work Package #3</td>
</tr>
<tr>
<td>Tim Pugh</td>
<td>BoM, Melbourne VIC</td>
<td>CAWCR IT Manager</td>
<td>Website and Integration Leader of Work Package #4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Staff Name</th>
<th>Location</th>
<th>Position</th>
<th>Duties and Roles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Say Teong Ng</td>
<td>CSIRO, Aspendale VIC</td>
<td>ACCESS Climate infrastructure</td>
<td>ACCESS services and user interfaces</td>
</tr>
<tr>
<td>Wenming Lu</td>
<td>BoM, Melbourne VIC</td>
<td>ACCESS NWP infrastructure</td>
<td>Model preparation, scheduling and interfaces</td>
</tr>
<tr>
<td>Asri Sulaiman</td>
<td>BoM, Melbourne VIC</td>
<td>ACCESS NWP infrastructure</td>
<td>Model source code and build management</td>
</tr>
<tr>
<td>Zhihong Li</td>
<td>BoM, Melbourne VIC</td>
<td>ACCESS NWP infrastructure</td>
<td>Data assimilation source code and build management</td>
</tr>
<tr>
<td>Li Xiao</td>
<td>BoM, Melbourne VIC</td>
<td>ACCESS NWP infrastructure</td>
<td>Data assimilation build and run management</td>
</tr>
<tr>
<td>Ilia Bermous</td>
<td>BoM, Melbourne VIC</td>
<td>ACCESS NWP infrastructure</td>
<td>Model infrastructure for verification and validation</td>
</tr>
<tr>
<td>Tim Erwin</td>
<td>CSIRO, Aspendale VIC</td>
<td>Climate analysis development</td>
<td>Technical expert in workflows, analysis, and data interoperability</td>
</tr>
<tr>
<td>David Kent</td>
<td>CSIRO, Aspendale VIC</td>
<td>Climate analysis development</td>
<td>Technical expert in workflows, analysis, and data interoperability</td>
</tr>
<tr>
<td>Aurel Moise</td>
<td>BoM, Melbourne VIC</td>
<td>CAWCR CVC Scientist</td>
<td>Climate scientist involved in IPCC model analysis and workflow tools.</td>
</tr>
<tr>
<td>Lawson Hanson</td>
<td>BoM, Melbourne VIC</td>
<td>Climate analysis development</td>
<td>Climate model analysis tools and data interoperability</td>
</tr>
<tr>
<td>Kevin Pulo</td>
<td>ANU/NCI</td>
<td>Data analytics manager</td>
<td>Data analytics and visualisation methods</td>
</tr>
<tr>
<td>Joseph Antony</td>
<td>ANU/NCI</td>
<td>NCI cloud manager</td>
<td>Cloud system specification and laboratory services</td>
</tr>
<tr>
<td>Muhammad Atif</td>
<td>ANU/NCI</td>
<td>Data specialist</td>
<td>ESG development and deployment</td>
</tr>
<tr>
<td>Stuart Ramsden</td>
<td>ANU/NCI</td>
<td>Data analytics</td>
<td>Core data, analysis, and graphics tools for laboratory integration</td>
</tr>
<tr>
<td>Ahmed El Zein</td>
<td>ANU/NCI</td>
<td>Storage cloud manager</td>
<td>Storage system specifications and laboratory services</td>
</tr>
<tr>
<td>David Singleton</td>
<td>ANU/NCI</td>
<td>HPC system manager</td>
<td>System resources and scheduling for computing</td>
</tr>
<tr>
<td>Gareth Williams</td>
<td>CSIRO, Melbourne VIC</td>
<td>eResearch advisor</td>
<td>eResearch consultant and advisor to the project from CSIRO eResearch</td>
</tr>
</tbody>
</table>

...plus many more

The Centre for Australian Weather and Climate Research
A partnership between CSIRO and the Bureau of Meteorology
THE SCIENCE AND THE NEED
Recognizing the direction and need of climate and weather simulation science

Moving from Open Source Software to Open Research Services
Earth System Simulator

The preparation and run of coupled and uncoupled model experiments within a framework designed with reproducibility, ease of use, support, and sharing of code and experiments.

ACCESS
The Australian Community Climate and Earth-System Simulator

(http://www.accessimulator.org.au)
To achieve insight and foresight into Climate and Environmental processes and prospects... future developments will be underpinned by increasing model complexity, compute capability, and data collections.

Provide a national approach to climate and weather prediction model development
Consistent with the strategy of **seamless prediction**

- **OBS**
- **Assimilation (4DVAR)**
- **Atmosphere (UM)**
- **Chemistry (UKCA)**
- **Wave (WaveWatch3)**
- **Sea-ice (CICE4)**
- **OCEAN** (MOM4)

**Drivers of variability and Change (SCIENCE):**
- Fronts, Convection, Local storms
- **Fronts**
- **Large scale weather systems**
- **Cut off lows**
- **Tropical cyclones**
- **Tropical air/sea interactions:**
  - **El Niño**
  - **Indian Ocean Dipole**
- **Southern Annular Mode**
- **Blocking**
- **Anthropogenic forcing on mean climate**
- **Other predictable sources**
- **Anthropogenic forcing**

**Social, economic and environmental targets (BENEFITS):**
- **Emergency response**
- Fire weather
- **Flash floods**
- **Safety of life**
- **Emergency management**
- **Flood and storm warnings**
- **Irrigation scheduling**
- **Short-term planning**
- **Emergency preparedness**
- **Water resource and crop management**
- **Drought preparedness and management**
- **Adaptation measures, Biodiversity and ecosystem conservation**
- **Resource development**
- **Emissions reduction**
- **Strategic resource planning**

**Time scale:**
- Days
- Weeks
- Months
- Years
- Decades
- Centuries

**Seamless Prediction**
- **Coupler (OASIS)**
- **Atmosphere (UM)**
- **Ocean (MOM4)**
- **Sea-ice (CICE4)**
- **Assimilation (BODAS)**
- **OBS**
- **Dynamic Veg. (LPJ)**
- **Land surface (CABLE)**
- **Ocean Carbon (CSIRO)**
- **Assimilation**
- **Emergency management and crop management**
- **Emergency preparedness**
- **Water resource and crop management**
- **Drought preparedness and management**
- **Adaptation measures, Biodiversity and ecosystem conservation**
- **Resource development**
- **Emissions reduction**
- **Strategic resource planning**

**Prediction Problem**
- **Nowcasting**
- **Numerical Weather Prediction**
- **Intra-seasonal Prediction**
- **Seasonal Prediction**
- **Decadal Prediction**
- **Climate Change Projections**

**Prediction Problem**
- Social, economic and environmental targets (BENEFITS)
- Drivers of variability and Change (SCIENCE)

**Prediction Problem**
- **Prediction Problem**
- **Social, economic and environmental targets (BENEFITS)**
- **Drivers of variability and Change (SCIENCE)**
Hurricane Sandy and Nor’ester cold front (Frankenstorm & Super Storm)

The Centre for Australian Weather and Climate Research
A partnership between CSIRO and the Bureau of Meteorology
THE INFRASTRUCTURE TO SUPPORT THE SCIENCE
Laboratory Work Packages

- **Web site**
  - Climate & Weather Science and Simulation Laboratory

- **Portal Service**
  - ACCESS Simulation Service
  - Model Analysis Service
  - Community Data Portal Service
  - Future Research Services

- **Framework & Toolkits**
  - ACCESS Framework
  - Model Analysis Framework
  - Data Services Framework
  - Data Toolkit
  - Service Framework
  - Service Tools & Methods
  - Modelling Toolkit
  - Model Analysis Toolkit

- **Core Data, Analysis, and Graphics Layer**

- **Collab Service APIs**
  - Core Data, Analysis, and Graphics Layer

- **Facility Interfaces and Data Library Services**

- **National Facilities**
  - NCI Storage and Compute Facility
  - RDSI Storage Facility
  - IMOS & TERN Facilities
  - RDSI & ESG nodes
Current Unified Modelling System

Bureau of Meteorology

Slave Repository

umui client

UM Target

CAWCR Users

umui server

Master Subversion Repository

umui server

Shared System @ NCI

University Users

Courtesy of Mike Rezny, CoE CSS
Setting up your UMUI suite of jobs

1. Copy an experiment or model release
2. Delete the jobs you don’t need from your experiment
3. Make sure that the start data is correctly specified
   • E.g. Global ECMWF reanalyses
4. Process all the jobs in your suite.
   • Assumes you’re happy with output data product listing
5. Submit the first job in the suite
   • And that’s it!
Met Office: Rose is a new Modelling Infrastructure
Planned Model Analysis Service Features

1. Easy of use: graphical user interface to create and visualize scientific workflows
   - Ability to use command line when necessary
   - Removes learning requirements of high performance computing (HPC) environments

2. Extensibility: ability of user to make their own workflow components and analysis.
   - Allows a simple analysis to be conducted over large datasets
   - Supports programming language features
   - Distributed and collaborative workflows

3. Traceability: easily reproduce analysis
   - Integration with relational database management system
Laboratory Goals

• Improve processing practices
  – Focus on data analysis rather than issues such as data access, tape recalls and job management
  – Easily incorporate new or change processing elements into the workflow
  – Sharing of analysis methods
  – Traceability and reproducibility

• Reduce data duplication
  – Common core processing can reduce data duplication
  – Sharing of processed data
Scientific Workflow

• Why not use existing software?
  – Current scientific workflow software treat data as:
    • Collection of files
    • Particular data set such as 20\textsuperscript{th} century temperature
    • Don't understand dimensions, attributes, meta-data etc
  – Can re-use a workflow but need to modify the data source or have multiple data sources.

• Prototype:
  – Workflow needs to be aware of attributes of data such as model, experiment and variable
  – Create a single data set for CMIP5
  – Allow 'sub-setting' the data on the fly
  – Creating standard names and help enforcing better standards
VisTrails Workflow

• Intuitive, drag-and-drop graphical interface

• Tightly integrated with Python language.

• Module structure means that plug-ins can be written easily to add the extra functions that we require, like integration with a database and climate-specific data naming.
Data Library and Facility Services

Data library

Hosting

Interfaces

Services

Data analytics

Cloud services

Visualisation

Management Tools
Data services and repository

- Hosting of data library
  - Based on the Earth Systems Grid Federation (ESGF) Peer-to-Peer system (supporting CMIP5 work)
  - ANL, BADC, CMCC, DKRZ, IPSL, NASA, NCI, ORNL, PCMDI

- Data library interfaces
  - Supporting searching, browsing, downloading, online use, and submitting/depositing datasets
  - Web based (browser and web services API)
  - Command line based (on NCI compute systems)

- Data analytics
  - Advanced web-based visualisation of datasets from data library

- Interoperability with local research community and facilities
  - IMOS, TERN, NEII, RDSI, …
  - Building a data services roadmap
  - Moving modelling towards community data and metadata standards

- Note, Service development is limited by available funding and staff
Cloud services and infrastructure

• Supporting users and developers of climate models on peak HPC facilities
  • Users: Cloud-based production environment for interactive model setup, submission, monitoring, post-processing, metadata validation and data library submission
  • Developers: Private, per-developer cloud-based VMs that mimic production environment
Laboratory Content

- Project Database
- Software Repository
- Experiment Database
- Data Repository

Australian Government
Bureau of Meteorology
CSIRO
Community Feedback

• Community Engagement
  • Please invite us to your conferences and workshops
    • AMOS 2013, Special Session 10 for CMIP5

• Community Reference Group
  • Express interest to join our community reference group

• Laboratory website to be establish by end of the 2012
  • Helpdesk and feedback

• Laboratory infrastructure releases every 6 months until end of 2013
  • Minor updates about every 3 months

• Supported until end of 2015
  • Bureau of Meteorology, CSIRO, CoE CSS, NCI

The Centre for Australian Weather and Climate Research
A partnership between CSIRO and the Bureau of Meteorology
Thank you
Governance Structure

The Centre for Australian Weather and Climate Research
A partnership between CSIRO and the Bureau of Meteorology

Stakeholder & Community Reference Group

Project Board

Project Manager

IT Lead

Science Lead

Project Implementation Committee

WP1

WP2

WP3

WP4

Project Support Office
# Project Steering Board

<table>
<thead>
<tr>
<th>Role</th>
<th>Name</th>
<th>Organization</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chair</td>
<td>Tom Keenan</td>
<td>CAWCR Director</td>
<td>Melbourne, VIC</td>
</tr>
<tr>
<td>Member</td>
<td>Helen Cleugh</td>
<td>CAWCR, CSIRO Deputy Directory</td>
<td>Canberra, ACT</td>
</tr>
<tr>
<td>Member</td>
<td>Peter May</td>
<td>CAWCR, BoM Deputy Directory</td>
<td>Melbourne, VIC</td>
</tr>
<tr>
<td>Member</td>
<td>Andy Pitman</td>
<td>CoE Climate System Science Director</td>
<td>Sydney, NSW</td>
</tr>
<tr>
<td>Member</td>
<td>Lindsay Botten</td>
<td>ANU / NCI Director</td>
<td>Canberra, ACT</td>
</tr>
<tr>
<td>Observer</td>
<td>Glenn Moloney (or delegate)</td>
<td>NeCTAR Director, Uni of Melbourne</td>
<td>Melbourne, VIC</td>
</tr>
<tr>
<td>Observer</td>
<td>Clare Richards</td>
<td>CAWCR Project Manager</td>
<td>Melbourne, VIC</td>
</tr>
<tr>
<td>Observer</td>
<td>Neville Smith (Sponsor)</td>
<td>Bureau of Meteorology, Deputy Director of Research and Computing</td>
<td>Melbourne, VIC</td>
</tr>
</tbody>
</table>