



Biosphere-relevant earth system model

CYBER-INFRASTRUCTURE WORKGROUP 1.C PROGRESS REPORT

Presented by Joe Vaughan
All-Hand Meeting
June 26-27, 2012
Pullman, WA



CYBER-INFRASTRUCTURE REPORT OUTLINE

- ◎ Individuals in CI working group
- ◎ Scope of BioEarth Cyber-Infrastructure
- ◎ Computing Resources
- ◎ Milestones from 2011 and Status of Progress
- ◎ Peer Review Plan
- ◎ Revised 2012-2013 Milestones

WORKING GROUP 1.C: CYBER-INFRASTRUCTURE

Cyberinfrastructure Team

Jennifer Adam, WSU

Serena Chung, WSU

Ananth Kalyanaraman, WSU

Mingliang Liu, WSU

George Gruber, WSU

*Joe Vaughan, WSU

CYBER-INFRASTRUCTURE SCOPE

Topics of Concern

Computing Resources

Forum for BioEarth Project Activities

Web Support for Teleconferencing

Code Acquisition, Storage and Version Tracking

Workflow Automation



COMPUTING RESOURCES: WSU CEE/EECS AEOLUS CLUSTER

24 8-core, 8GB Ram Intel nodes and

6 16-core, 32GB ram nodes, for total of 288 cores.

98 Terabytes RAID Storage

(Planned upgrade adds:

4 64-core, 128GB nodes for 256 more cores,

1 dual-Tesla 1060 GPU node.)

10GB Ethernet between all nodes and storage

Aeolus Storage of 70 TB, with work in progress sharing to Central Cluster.



COMPUTING RESOURCES: WSU CENTRAL CLUSTER

164 12-core nodes for 1,968 cores.

3 large memory nodes with 512 GB per node.

QDR Infiniband at 40GB between all nodes.

The 'in progress' connection from Aeolus's RAID Storage to Central Cluster will likely start at 1GBE, with increase possible toward 10GBE.

MILESTONES FROM 2011 AND PROGRESS / STATUS

SUPPORT FOR BioEarth COMMUNICATIONS:

1.A By Dec 2012 all users will be using the forum as a major means of project communication and generating memoranda documenting progress.

Not accomplished – perhaps ill conceived?

1.B Solve need for web-conferencing by February 2012.

Using Adobe Connect, courtesy of Chad Kruger.



MILESTONES FROM 2011 AND PROGRESS / STATUS

SUPPORT FOR BioEarth CODES:

2.A By February 2012, to solicit Individual project members and negotiate a plan and schedule for code acquisition. Goal to have code collected by December 2012.

Not accomplished, yet.

2.B By March 2012, conduct training on use of subversion for code storage and versioning.

Not accomplished, yet.



MILESTONES FROM 2011 AND PROGRESS / STATUS

SUPPORT FOR BioEarth CODES (continued):

2.C By March 2012, to implement Workflow test for WRF.

Late but in progress June 2012, contingent on WRF.

2.C By June 2012, implement sequential Workflow test for more than one model / program.

In progress June 2012, contingent on WRF.

2.C By December 2012, implement Provenance features for Workflows.

Still reachable for December 2012.

DISCUSSION

Forum

What's New?

New Posts Private Messages FAQ Calendar Community Forum Actions Quick Links

Forum

BioEarth Integrated Regional Modeling Initiative Forums

Welcome to the BioEarth Integrated Regional Modeling Initiative Forums.

Logistics and Scheduling



Logistics and Scheduling

Scheduling of all meetings, coordination within and between teams, etc.

Research Topics



Model Integration



Atmospheric Modeling

WRF: Meteorology, CMAQ: Atmospheric Chemistry



Terrestrial Modeling

Cropsyst, RHESSys: Terrestrial nutrient dynamics, VIC: hydrology, MEGAN: Biogenic Emissions



Aquatic modeling

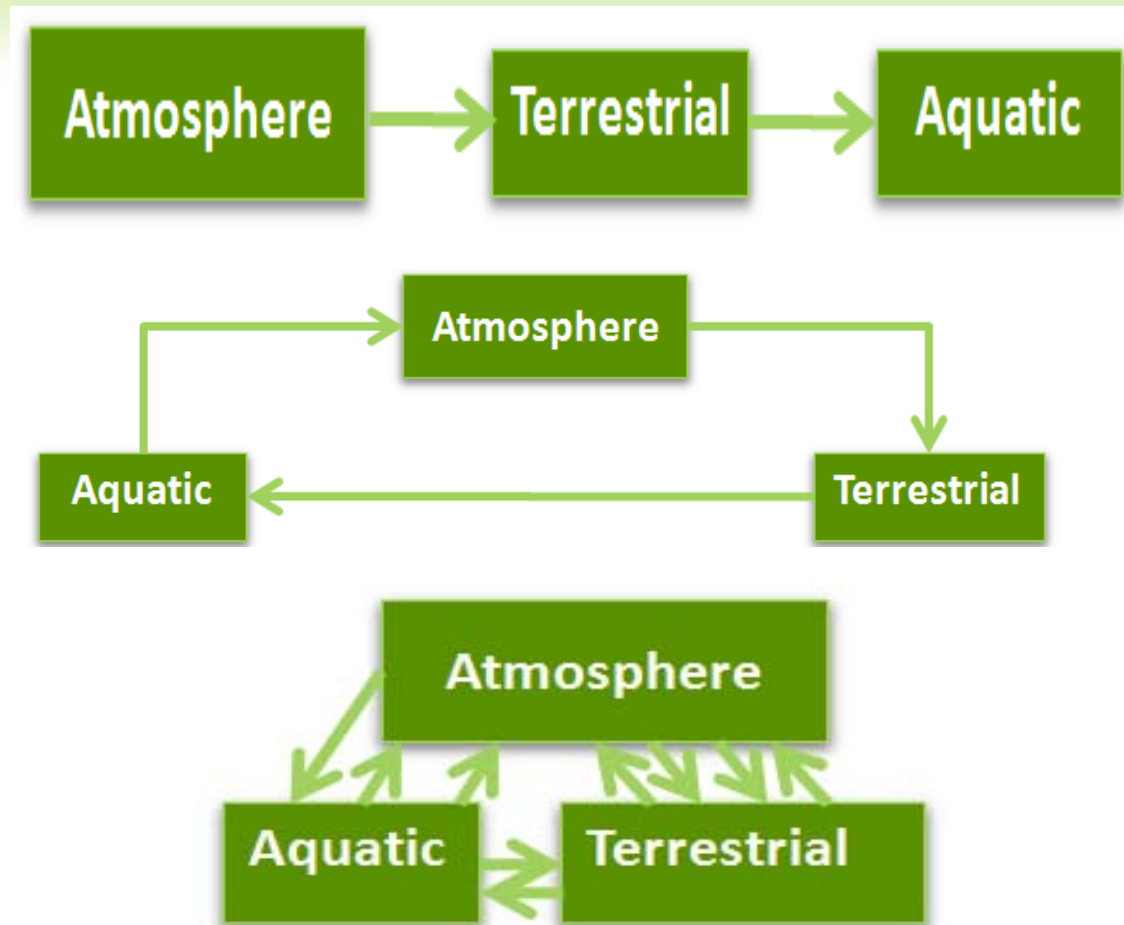
ColSim: Reservoir Operations and Withdrawals, Global NEWS: Nutrient Transport in Channels

“OK to start conversations by email, but then post to Forum as they develop and email the new thread to conversation participants.”

Forum communications may be more salient to Stakeholders, as they may not use email constantly as do many BioEarth academics.

REMINDER: BIOEARTH COUPLING TERMINOLOGY

- SEQUENTIAL
- ITERATIVE
- FULLY --
COUPLED



KEPLER WORKFLOW JOB INITIALIZATION:



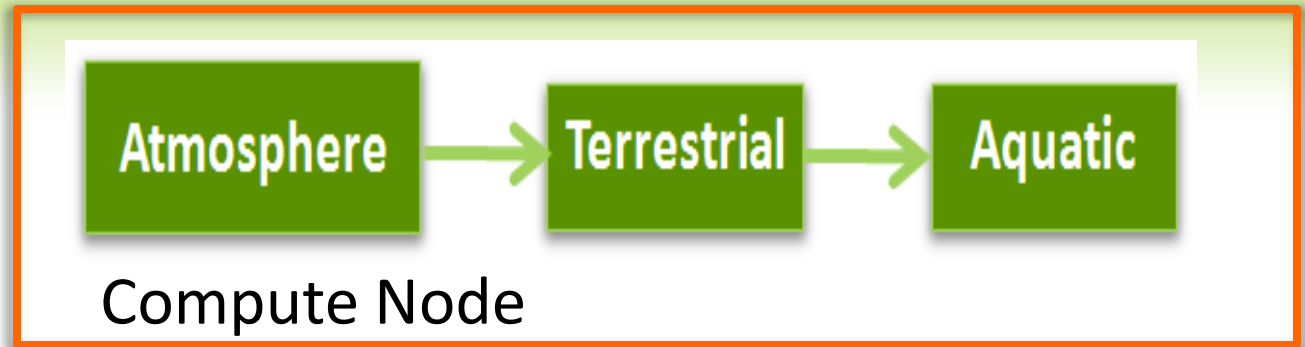
○ SEQUENTIAL



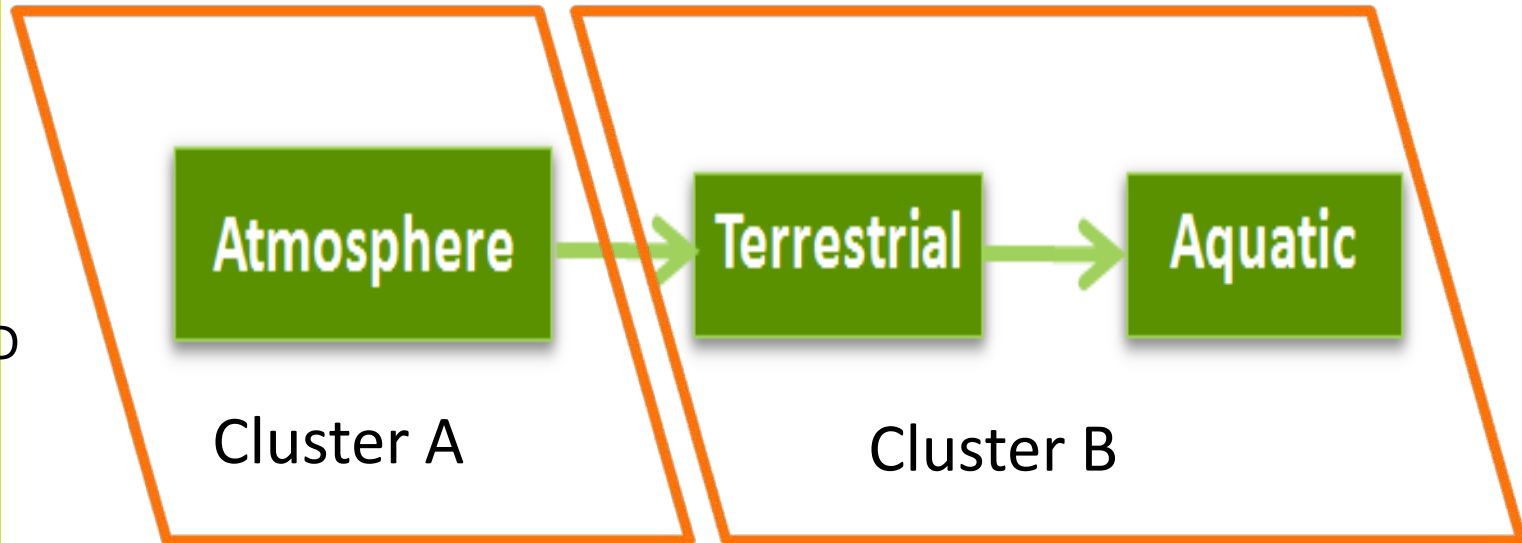
○ PARALLEL

KEPLER WORKFLOW JOBS: SINGLE MACHINE OR DISTRIBUTED

○ SINGLE
MACHINE

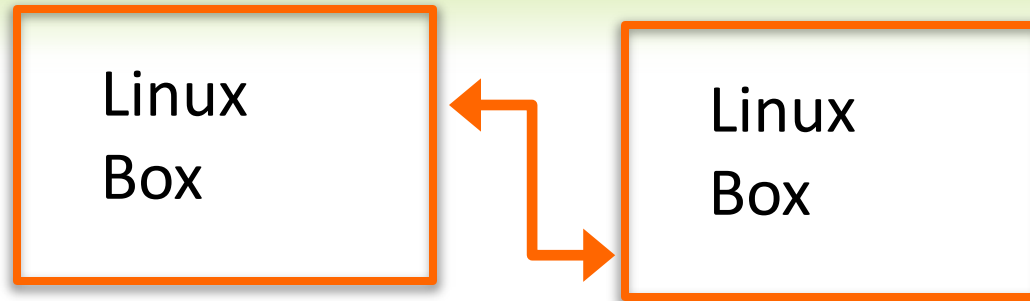


○ DISTRIBUTED

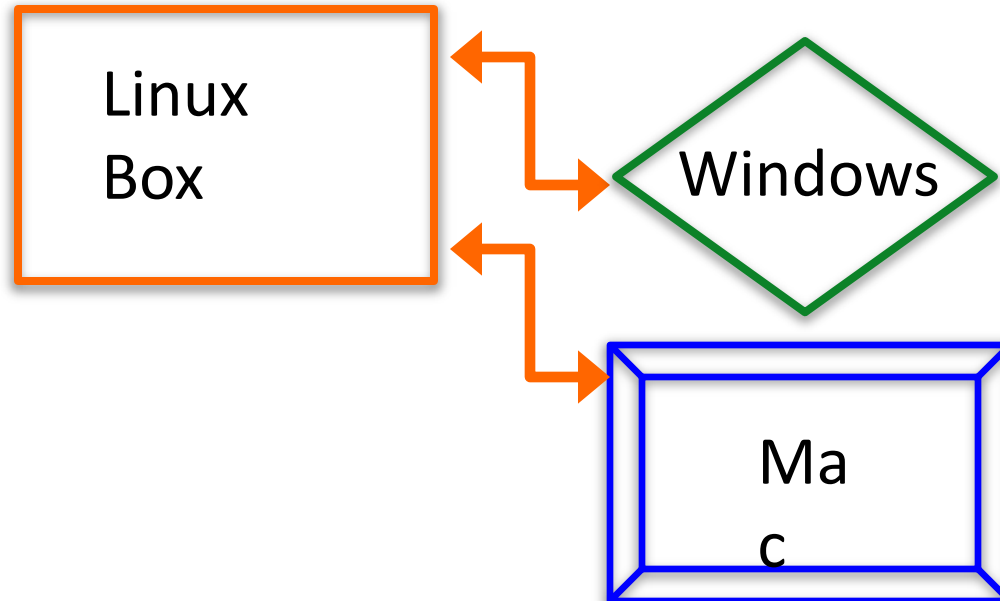


KEPLER DISTRIBUTED WORKFLOWS: HOMOGENEOUS OR HETEROGENEOUS

Homo-
geneous



Hetero-
geneous



CI PEER REVIEW PLAN

- ◎ Ilkay Altintas, Deputy Coordinator for Research,
San Diego Supercomputer Center (SDSC)
Lab Director, Scientific Workflow Automation Technologies
University of California, San Diego
- ◎ Kepler Project for Workflow Automation
- ◎ *Kepler is apparently still in active development*
- ◎ *Early Kepler users will likely help drive development through presenting needs and problems to the development team.*
- ◎ Peer Review visit by Ilkay Altintas planned for January/February 2013. Teleconferences may be arranged in meantime as consultation is needed.



REVISED MILESTONES FOR 2012-2013

SUPPORT FOR BioEarth COMMUNICATIONS:

- 1.A By February 2013 be providing support to Stakeholders in using the forum.

- 1.B Through 2013, Monitor adequacy of Using Adobe Connect and parallel telephone links to support project activities.



REVISED MILESTONES FOR 2012-2013

SUPPORT FOR BioEarth CODES:

- 2.A By August 2012, implement Workflow test for WRF through CMAQ, sequential run.
- 2.B By October 2012, understand Provenance and implement some features in Workflows.
- 2.C By November 2012, offer training in basic Workflow coding and testing to BioEarth Investigators.

REVISED MILESTONES FOR 2012-2013

SUPPORT FOR BioEarth CODES (continued):

- 2.D By December 2012, conduct training on use of subversion for code storage and versioning.
- 2.E By December 2012, Distributed Workflow testing for running across multiple homogenous hosts.
- 2.F By December 2012, capture an archive of codes being tested in the central cluster, using subversion.
- 2.G By February 2013, Distributed Workflow testing for running across multiple heterogeneous hosts.

CYBER-INFRASTRUCTURE REPORT CONCLUSION

- ⊙ Individuals in CI working group
- ⊙ Scope of BioEarth Cyber-Infrastructure
- ⊙ Computing Resources
- ⊙ Milestones from 2011 and Status of Progress
- ⊙ Peer Review Plan
- ⊙ Revised 2012-2013 Milestones

⊙ Questions & Comments?