CESM Update 2014

Mariana Vertenstein

NCAR Earth System Laboratory
CESM Software Engineering Group (CSEG)

NCAR is sponsored by the National Science Foundation

Outline

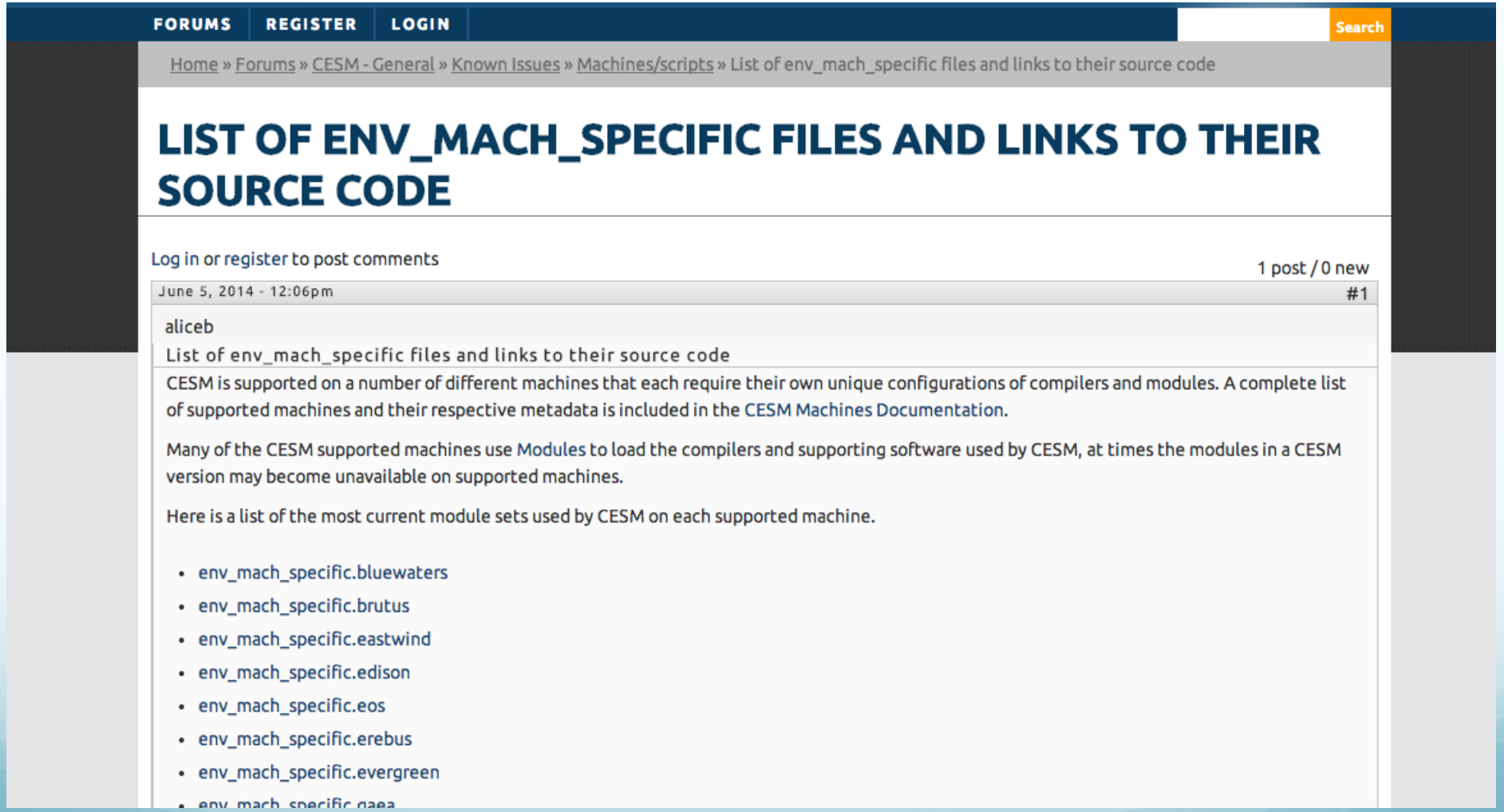
- **Community Releases and Simulations**
- **Infrastructure Updates**
  - CPL7
  - System Testing/Unit Testing/PIO
  - Workflow
  - User Defined Grids
- **New Software for New Science**
  - CVMix (POP/MPAS-O)
  - Dynamic Landunits
  - Ecosystem Demography-ED (CLM)
  - Subcolumns (CAM)
- **Brief Summary of other Component Updates**

Support for CESM Community

Community Releases and Support

- Community releases this year
  - CESM1.0.6 (May 2014)
  - CESM1.2.2 (June 2014)
- Community Support - *DiscussCESM Forums Bulletin Board* http://bb.cgd.ucar.edu
  - 1793 Total Registered users
  - 14 Top level forums with 77 sub-forums
  - 5229 Total posts - 800 posts since March 2014
  - Recommended as *ONLY* way to communicate release based issues - CSEG routinely monitors and replies to posts

New Forums Machines Page



http://bb.cgd.ucar.edu/list-envmachspecific-files-and-links-their-source-code

CESM Community Simulations

- Large Ensemble: 62+ simulations, ~7,500 years
- Last Millennium Ensemble: 26 simulations, ~32,000 years
- Data management is challenging!
- Both employing new CESM workflow changes
- Exploring assigning DOIs (digital object identifiers) to output datasets using these two projects as testbeds
  - Interesting implications for reproducibility, data provenance, data versioning

    - Acknowledgement: Andy Mai and Gary Strand

Infrastructure
CPL7

CPL7 Improvements

(Tony Craig, Jim Edwards, Mariana Vertenstein)

- New coupling interactions added – not all are being actively leveraged yet
  - *glc -> ice*
  - *glc -> ocn*
  - *riv -> ice*
  - *riv->lnd (being used for feeding flooding back to CLM)*
- New options for coupler sequencing
  - reduce lags
  - improves coupling field consistency
- Major refactor for modularity

CPL7 Refactorization

- Why?
  - driver code hard-coded to contain MCT types
  - one large routine (~6K lines) with functionality often repeated
  - difficult to understand and modify flow through system
  - difficult to implement and evaluate possible alternative coupling architecture (e.g. ESMF)
- Introduced a new *abstraction layer* between driver and components - *driver now has no references to either MCT or ESMF types*
  - Driver lines of code reduced from 6k -> 3k
  - Much easier to understand and modify flow
  - Much easier to incorporate ESMF/NUOPC driver and components

Infrastructure Testing/PIO

System Testing Speedup

(Jay Shollenberger and Jim Edwards)

- Goal: Speed up testing by leveraging sharing component build libraries (where possible) across executable builds
  - MCT, GPTL, PIO (CLM, Data Models are next)
- Extensive modifications made to both CESM build scripts and Makefile
- Results: Yellowstone development code test suite is currently 103 separate test cases:
  - 11 hour build time *without* new shared builds.
  - 7 hour build time *with* shared builds enabled.

System Testing Management

(Jay Shollenberger)

- New perl/xml-based test list management utility
  - Responsible for adding, deleting, and syncing changes to CESM test lists.
  - Fully unit and integration tested
  - Easier to use and fully documented
  - In development code only right now – will be part of next release

New Unit Testing Framework

(Sean Santos)

- Written leveraging pfUnit and Cmake (Sean Santos)
- Implemented in CAM, CLM and csm_share
- Being extended by CSEG

New PIO in C

(Jim Edwards)

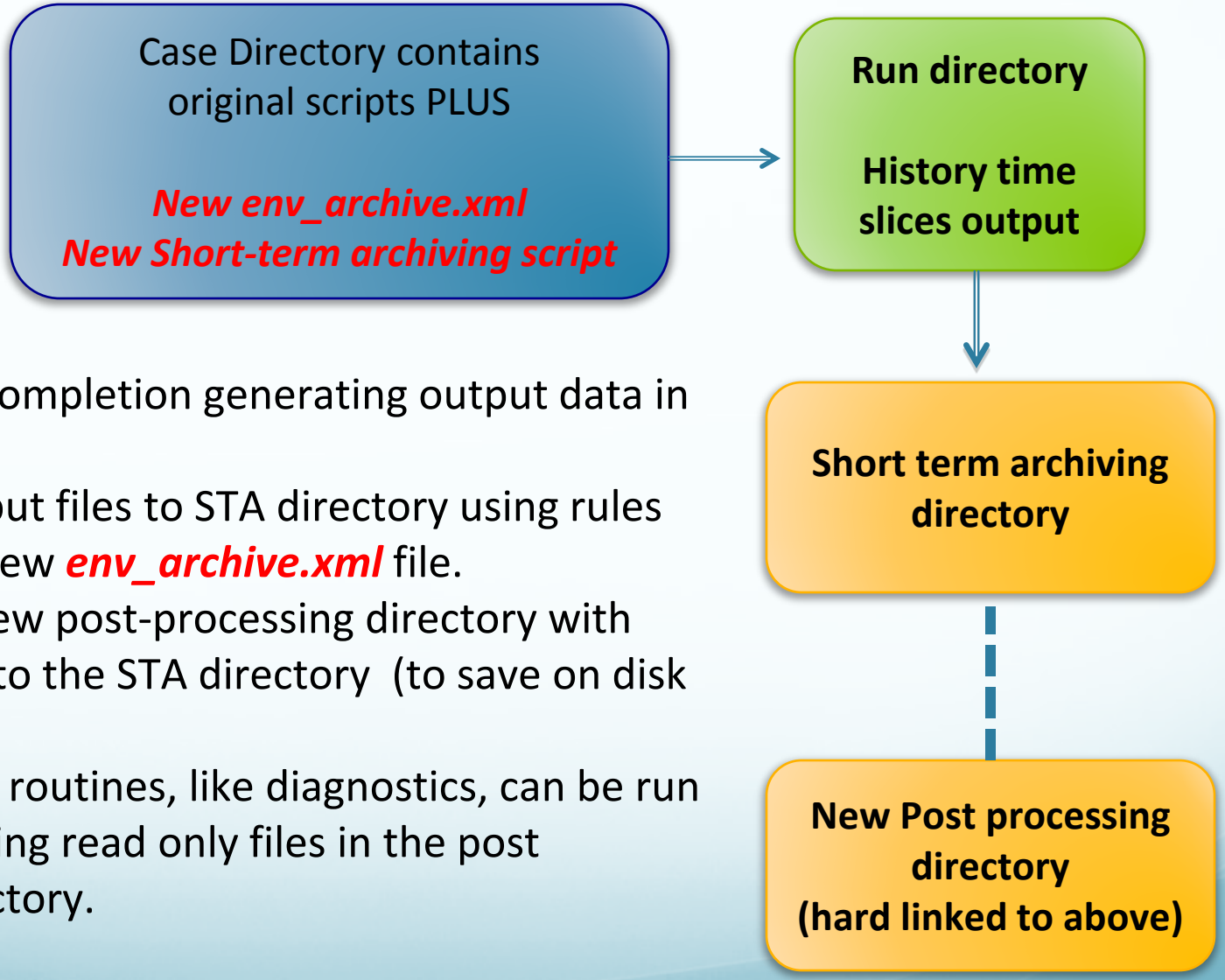
- Rewrite will enable PIO to easily be leveraged by new utilities (e.g. MOAB)
- More robust interface for netcdf

Infrastructure Workflow

Archivers Refactoring Project
(Alice Bertini)

- Improved short-term and long-term archiving
  - preserve a ***working copy*** of model run data on disk for post-processing ***during*** model run
- New interaction with new time-slice->time series generation (needed by diagnostics)
  - controls time-series generation from time-slice history field output as part of short-term archiving
- Improved model provenance
  - "push" metadata from case directory to the experiment database (https://csegweb.cgd.ucar.edu/expdb/cgi-bin/login.cgi) via XML and CGI
- Written in Perl and Python and totally XML based for easy configuration

Short Term Archiving (STA) Capabilities



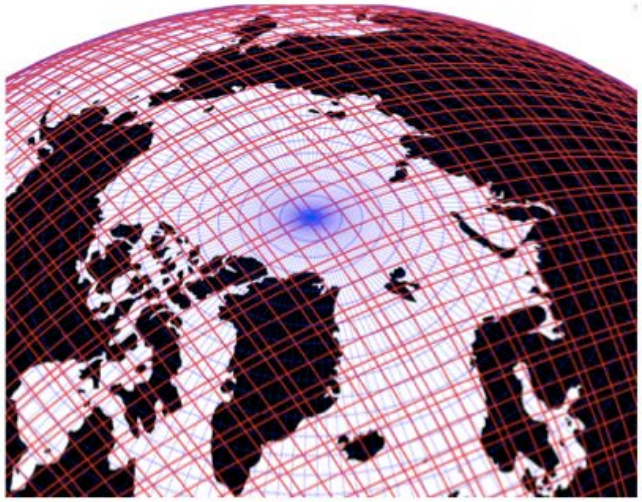
1. Model runs to completion generating output data in run directory.
2. STA moves output files to STA directory using rules defined in the new ***env_archive.xml*** file.
3. STA creates a new post-processing directory with hard links back to the STA directory (to save on disk space).
4. Post-processing routines, like diagnostics, can be run concurrently using read only files in the post processing directory.

Infrastructure
User Defined Grids

User Defined Grids - Example

- Continue to improve tool chain and capability to simplify introduction of new grids into CESM
- Recent example - Fully coupled model with CAM SE dycore ocean shows much slower spin up time from Levitus than with CAM FV dycore

Grid differences at high latitudes



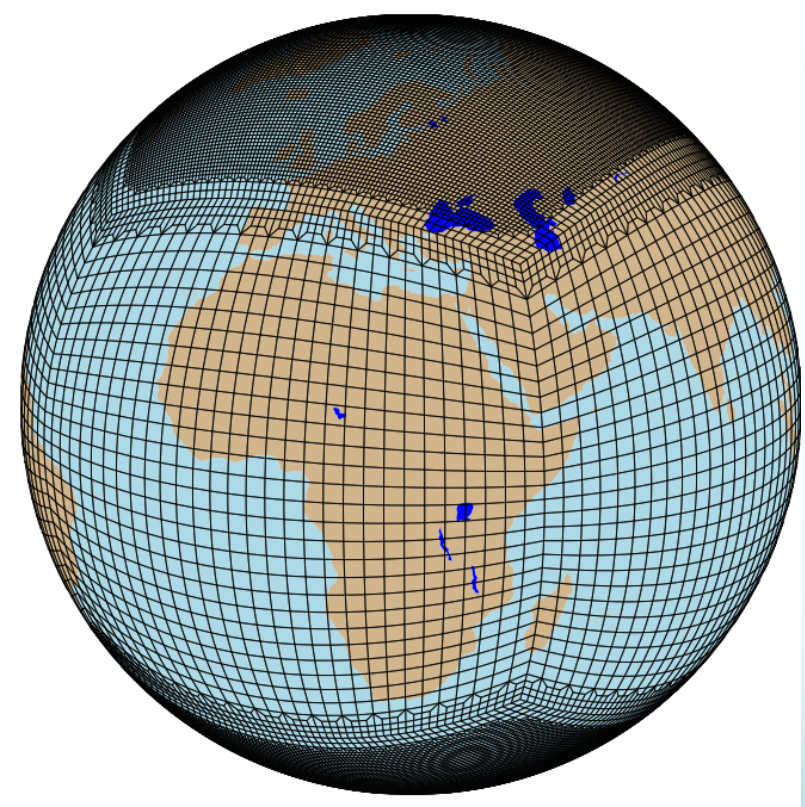
Red: CAM-SE grid
Blue: CAM-FV grid
(at about 2 degree)

Courtesy:
Peter Lauritzen

Remapping differences (ocn ⇔ atm)

State variables: FV uses "bilinear" and SE "native"

Want to examine this further by creating new polar refined grid



Created new atm/land grid and mapping files and carried out new coupled Model experiment on this grid in less than 2 weeks!

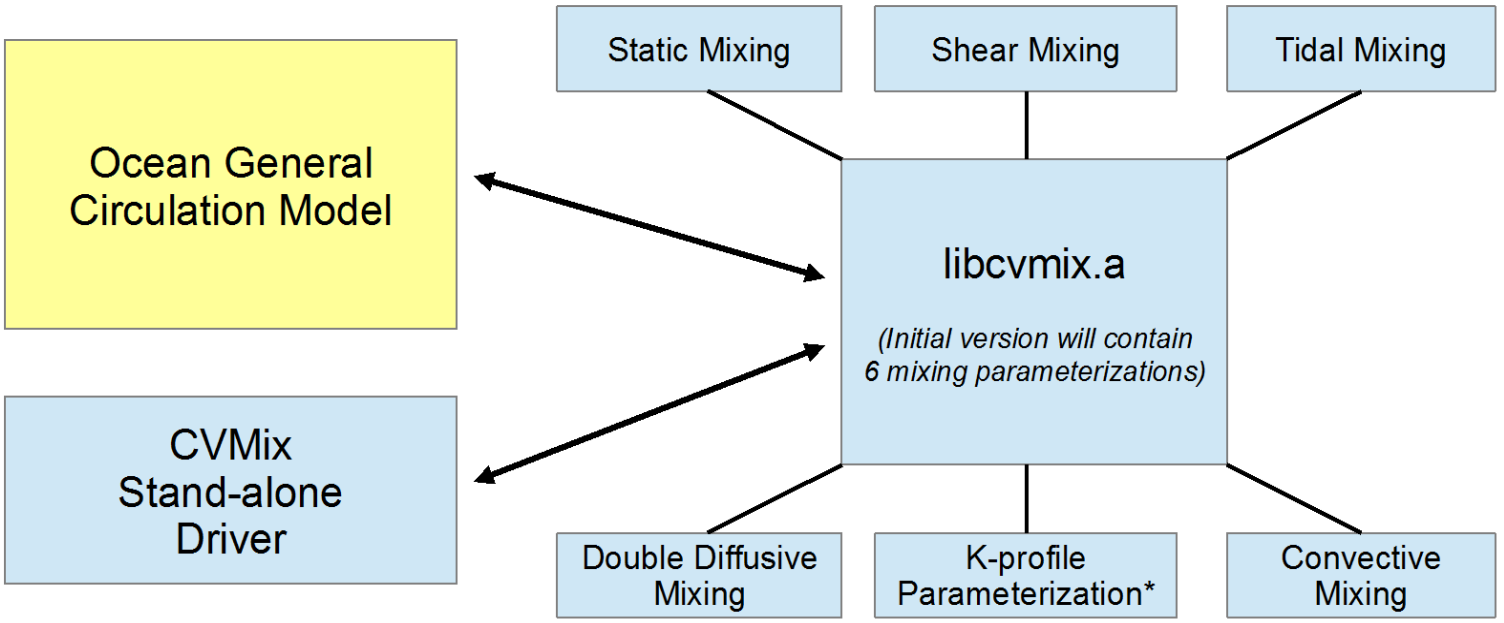
Result showed that grid was not not problem

New Software/
New Science

Vertical Mixing in Ocean Models
(Mike Levy)

- Current State
  - Numerous techniques for parameterizing the mixing process
  - Model developers choose their favorite parameterization(s) and code them up as part of their ocean model – could not easily be shared among different ocean models
- Community Vertical Mixing (CVMix) Project
  - Joint project between NCAR, LANL, GFDL
  - Will be included in POP2, MPAS-O and MOM6
  - Primary goal: produce an easy-to-use library containing a range of parameterizations that is not model dependent
  - Secondary goal: provide a stand-alone driver to test the library on its own

CVMix (Current State)

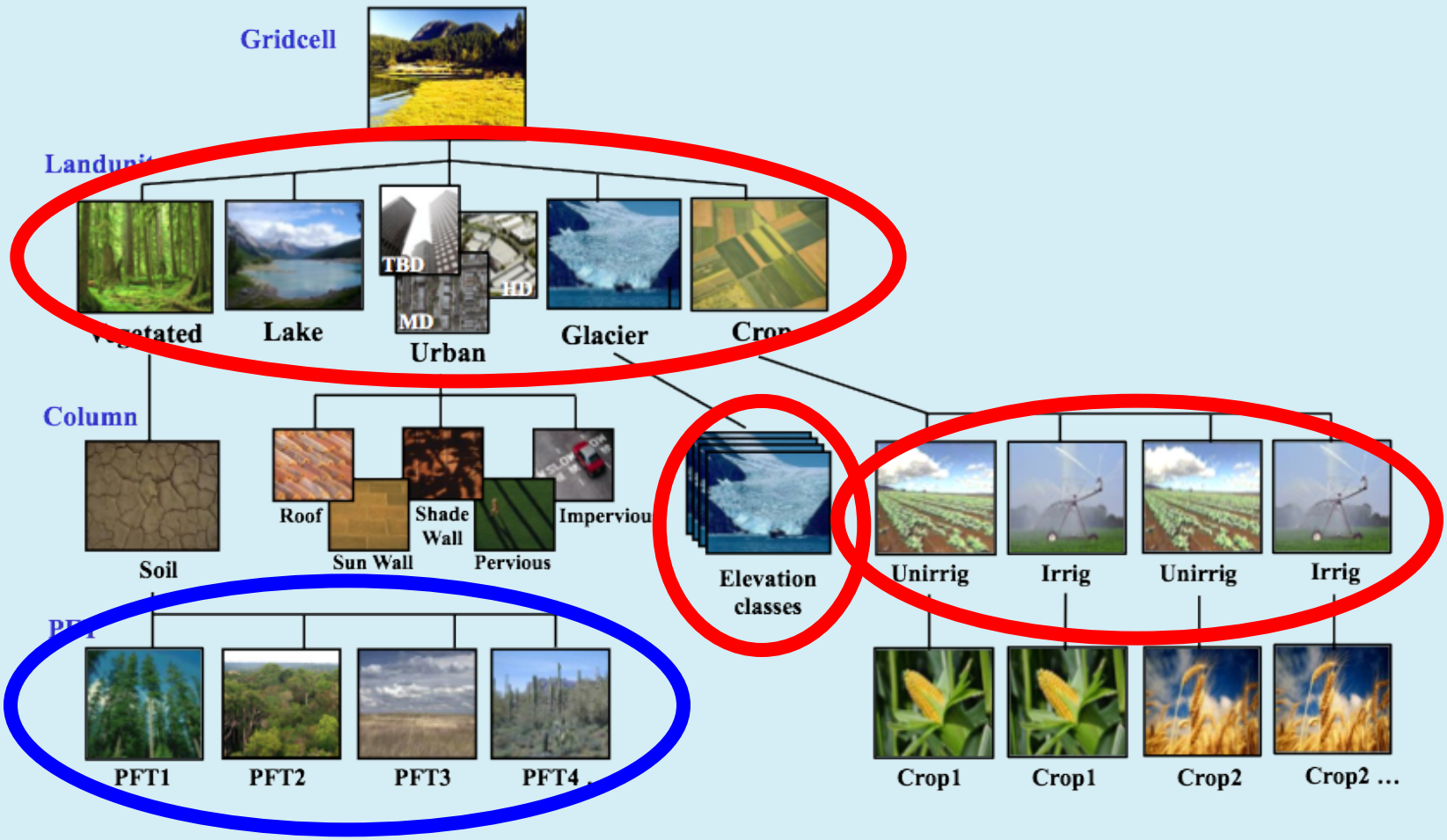


*Currently debugging KPP mixing, last hurdle before beta release (anticipated June 30th)

- Beta testing to begin by June 30, 2014
- Version 1.0 release on September 30, 2014

CLM Gridcell Organization and Dynamic Landunits

(Bill Sacks)



Dynamic Landunits - Near-term Science Targets

Glacier dynamics

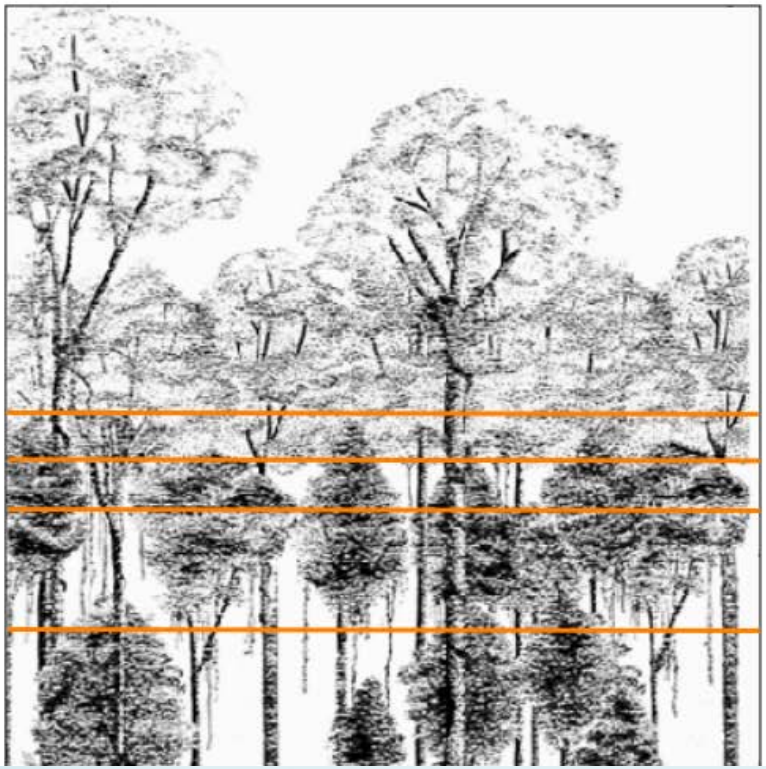
- Future stability and potential for recovery of the Greenland Ice Sheet, considering feedbacks
- Transient simulations of the Last Interglacial and deglaciation following the Last Glacial Maximum

Crop dynamics

- Transient simulations with managed crops
- Coupling CESM to the iESM integrated assessment model

New Ecosystem Demography (ED)

Plant competition as an emergent property of plant physiology…



'Cohorts' of plants,
Binned according to:

1. Successional stage of land
2. Functional type of plant
3. Height class of plant

Provides a framework to simulate

1. Competition for light
2. Coexistence between plant types…
3. Post-disturbance succession
4. Variable vegetation complexity

Courtesy of Rosie Fisher

CLM ED implementation

(Stefan Muszala)

- New CLM/ED data structures
  - Each gridcell contains patches (oldest->youngest)
  - Each patch contains cohorts (shortest->tallest)
  - Implemented as a linked list
- Since all other CLM data structures are vector data – introduced generalized method to translate
  - Linked list <-> current vector data structures
  - Needed for history and restart output
- Implementation has been modular and extensible
  - Touches very few lines of CLM code
- Status
  - Implemented in development trunk – support for single point, regional and some global resolutions

Subcolumns in CAM

(Cheryl Craig and Steve Goldhaber)

Approaches

- Allow for finer granularity than grid box by spatially sub-dividing the column (grid cell)
- Use a statistical approach to sample subgrid variability within a single column (grid cell)

Features

- Avoid having to increase resolution of entire model to study one parameterization
- Share subcolumn data between parameterizations (unlike current sub-grid-scale parameterizations such as radiation)

Status

- Subcolumn infrastructure complete on CAM trunk
- Subcolumn support in CAM microphysics on CAM trun

A Few Other New Component Capabilities

CAM

- Substantial refactoring of CAM's aerosol interface code. - will facilitate adding new aerosol models.
- Integrated the PORT offline radiation code into CAM - provides a framework to build offline modes for testing other parameterizations.
- CAM and WACCM gravity wave modules refactored and merged.

WACCM

- New WACCM chemistry for the whole atmosphere (troposphere through lower thermosphere), developed for the Chemistry-Climate Model Initiative (CCMI).
- Updated specified chemistry (SC-WACCM) compsets in recent releases
- Examining ways to make history output more convenient for WACCM and CAM-CHEM (more of fields you want, less of the ones you don't).

CICE

- CICE5 is currently being brought into CESM. Coding is mostly complete and validation is underway.

MPAS-A (atmosphere)

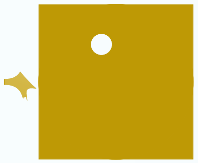
- MPAS non-hydrostatic dynamics implementation in CAM development branch with CAM4 and CAM5 physics (Michael Duda and Sang-Hun Park)
  - tested so far with aquaplaned and Held-Suarez idealized cases
- Work on AMIP simulations starting – will serve as a stepping stone to coupled NWP-type simulations

MPAS-O (ocean)

- Implemented as alternative prognostic ocean component in development code (Doug Jacobson)
- Core forcing runs have started

NCAR
Earth System Laboratory

U.S. DEPARTMENT OF ENERGY
Office of Science



Questions?

Appendix

CESM Verification Tool

- Goal:
  - Verify that port to new machine and/or changes to compilers produces same climate as a run on a trusted machine.
- Approach:
  - Creation of "trusted baseline": Run 101 one year runs that differ only in initial atmosphere temperature field (round-off level changes). Look at annual averages of CAM output and compute distributions of RMSZ scores and global means for each variable.
  - Verification test: Randomly select three ensemble members to run on the new machine, compare RMSZ and global means from each run to the original ensemble.
- Status:
  - Currently performing verification test (for B1850C5CN at ne30_g16) as part of the CESM test suite for each new beta development tag – create new baselines each time climate changes.
  - Still an open research question, statisticians and scientists are working with us to develop pass / fail criteria that we can trust.