Outline

• Management
  • CIWG (Cyberinfrastructure Working Group)
• Data Policy
• Software Policy
• Resources
  • Computing
  • Data
• Community Support
  • Data interoperability
  • Model code portability
Objective 6

- Management, analysis and visualization of very large model output datasets - creation of infrastructure
- Resource procurement & distribution
- Data management
- Analysis and visualization of model results
Management
CIWG

Wednesday, January 13, 2010
CIWG Objectives

• Make efficient use of computing and data resources
• acquire resources
• coordinate resource utilization
• collaborate to leverage joint efforts
• Provide technology look-ahead
• Validate goals and provide advice and consent to Executive Committee
URL data retrieval

OAI metadata
200 GB/snapshot = 1/2 hr

MMF-LES
Model Data
Output
Budget

20 TB/day
Simple Methods of Data Download (2) and Upload (1)

Storage Resource Broker (SRB)

SAM–QFS has 16 tape drives, an 304–TB disk cache
- Meta Data Servers: 2 x Sun x4600
- Data Login Servers: 2 x Sun v40z /
- 6 x 1GbE connections to the TeraGrid/HPC network
- Tape Drives
- 16 STK 9940–B tape drives
- 24 IBM J2 3592 tape drives
The GCRM Tsunami

2 km, 100 levels, hourly data
• ~4 TB / simulated hour
• ~100 TB / simulated day
• ~35 PB / simulated year

4 km, 100 levels, hourly data
• ~1 TB / simulated hour
• ~24 TB / simulated day
• ~9 PB / simulated year

Other Data Tsunamis
• 30 TB/night: Large Synoptic Survey (LSS) Telescope (2014)
• 15 PB/year: CERN’s Large Haydron Collider (May 2008)
• 1 PB over 3 years: EOS (Earth Observing System) data (2001)

(from K. Schuchart presentation January, 2008)
Parallel Input/Output Technology Progress

Internet Data Transfer Capacity

Timeline

NetCDF 3.0

NetCDF 3.6

NetCDF 4.0/4.0.1beta

Pnetcdf 1.0

Pnetcdf 1.0.3

HDF5

Netcdf4.0 beta

ADIOS alpha

64 bit PNetCDF/NetCDF?
Architectural Goals

Local Resources → Mid-range CMMAP Server(s) → Digital Library software & data → Community Computing Resources

Local Resources → CMMAP Machine Room

Home Institution

CMMAP Machine Room

External Resources
Management

Data Policy
Data Policy

- Data published on the CMMAP Digital Library is in the public domain but registration and authorization required to access it
  - this is to prevent hacking and bot-crawling and
  - provide tracking of who is accessing the data
- All metadata is public
  - will be published via a new OAI (Open Archive Initiative) service to be instituted this year
- CF metadata conventions are followed
- We are investigating the use of DOIs (digital object identifiers) for data consistent with the scholarly publication process
Management

Software Policy
Software Policy

• Software developed with CMMAP funding is covered by an intellectual property agreement between CSU and CMMAP partners

• MMF code is controlled by the author of the code

• GCRM code is very new and developed under SCIDAC funding

• Subversion source control system has been provisioned but has not yet been required

• CCSM coding conventions are encouraged
Community Sea Ice Model (CSIM) Developer’s Guide
Code Reference for Version 5.0

Released with CCSM3.0

Julie Schramm
Cecilia Bitz
Bruce Briegleb
Marika Holland
Elizabeth Hunke
Bill Lipscomb
Dick Moritz

5.1 Style Guidelines

Subsections
- General Guidelines
- Modules
- Subroutines
- Loops
- Array Syntax
- Allocatable Arrays
- Variable Names
- Variable Declarations
- Code Indentation
- Commenting of code
- Portability
- Incomplete and dead code
- Miscellaneous

5.1 Style Guidelines

General Guidelines

- Fortran 90 Standard CSIM uses the F90 standard.
- Preprocessor The C preprocessor is required as the macro processor (cpp). All tokens should be upper case to distinguish them from fortran code. The use of ifdef constructs should be avoided wherever possible. The option should be set in the namelist instead.
- Fixed-Form Source Fixed-form source will be used for readability. Columns 7 to 72 of a line may contain a Fortran statement, while columns 1 to 6 are reserved for special purposes. Blanks may be used freely anywhere. Exclamation marks should be used to denote comments. A statement may include a maximum of 19 continuation lines.
- Bounds Checking All code should be able to run when the following compiler options are set:
  o Array bounds checking
  o Automatic variables are initialized to NaN
  o Floating point exception conditions: overflow, division by zero, and invalid operations

Modules

The use of modules is strongly encouraged since they provide global access to variables.
- Modules must have the same name as the file in which they reside, due to dependency rules used by "make" programs. This implies that multiple modules are not allowed in a single file.
- Module names must begin with "ice_". This will provide unique module names when CCSSM is released as a single executable version.

Related subroutines and variables should be placed in a single module file, rather than to maintain a single routine per file.
http://cmmmap.sdsc.edu/wiki/

Documentation
Resources
Computing
### Allocation Types, Limits, and Deadlines

<table>
<thead>
<tr>
<th>Units Requested</th>
<th>Startup/Educational Allocation</th>
<th>Research Allocation (TRAC)</th>
</tr>
</thead>
</table>
| Service Units (SU) on Computer Resources | - Most systems have a limit of 30,000 SUs, though some may be as high as 200,000 SUs. Please see the Resource Catalog for specific resource limits.  
- Maximum Startup request for TeraGrid Roaming: 50,000 SUs  
- Aggregate request for multiple compute resources cannot exceed 200,000 SUs. | - 30,000 – Unlimited |
| Terabytes (TBx) on Data Resources | - Storage on disk: 5 TB  
- Storage on tape: 25 TB | |

### Deadlines

<table>
<thead>
<tr>
<th>Open Submissions</th>
<th>Close Submissions</th>
</tr>
</thead>
</table>
| N/A              | Jan. 15  
| Dec. 15          | Apr. 15  
| Mar. 15          | Jul. 15  
| Jun. 15          | Oct. 15  
| Sept. 15         |                 |

### Allocations Begin

| Two weeks after submission | April 1  
|                           | July 1  
|                           | October 1  
|                           | January 1  

### Review Cycle

Within one week  
Quarterly  

### Typical Use

Classroom or training accounts and startup accounts requiring small amounts of time  
Experienced users with research projects

---

1 The TRAC Meeting schedule is available for reviewers on the POPS Review page.  
2 See the Past Allocations for a list of previous awards.
## Leveraging National & Partner Resources

<table>
<thead>
<tr>
<th>Organization</th>
<th>Resource</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Allocations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>San Diego Supercomputer Center (SDSC)</td>
<td>Disk</td>
<td>15 Terabytes</td>
<td>15 Terabytes</td>
<td>30 Terabytes</td>
<td>45 Terabytes</td>
</tr>
<tr>
<td></td>
<td>BlueGene</td>
<td></td>
<td></td>
<td></td>
<td>30,000 SUs*</td>
</tr>
<tr>
<td>Teragrid (multi-institution)</td>
<td>SDSC DataStar (IBM SP4)</td>
<td>600,000 SUs</td>
<td>1,200,000 SUs</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Grid Roaming</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LSU Steele</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lawrence Berkeley National Laboratory (LBNL)</td>
<td>National Energy Research Scientific Computing Center (NERSC)</td>
<td></td>
<td></td>
<td>700,000 SUs</td>
<td></td>
</tr>
<tr>
<td>Oak Ridge National Laboratory (ORNL)</td>
<td>Cray XT</td>
<td></td>
<td></td>
<td>2,000,000 hrs</td>
<td>3,000,000 hrs</td>
</tr>
<tr>
<td>National Center for Atmospheric research (NCAR)</td>
<td>Bluelce IBM Power5</td>
<td></td>
<td></td>
<td></td>
<td>500,000 SUs</td>
</tr>
<tr>
<td>IBM Watson Research Center</td>
<td>BGW - eServer Blue Gene Solution</td>
<td></td>
<td></td>
<td></td>
<td>TBD</td>
</tr>
<tr>
<td>Stonybrook</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TBD</td>
</tr>
<tr>
<td>Computing Allocations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Resources

Data
Published in Digital Library

Stored in Archival File System

Conversion of SAM6.7SR com3D to netCDF

*AI: Rename SUPERLES to SPCAM
http://cmmmap.sdsc.edu/wiki/

Same as Wikipedia
Community Support

Data Interoperability
INCITE Resources & Allocations
Mon, 05/18/2009 - 23:25 — hellyj

2010 INCITE Call for Proposals

Add new comment  Read more

Teragrid Resources & Allocations
Mon, 05/18/2009 - 23:14 — hellyj

Add new comment  Read more

Sample Fortran program to read GIGALES netcdf 3D snapshot data

This program will read one variables for a subset of the horizontal domain for all vertical levels of a given 3D snapshot time.

1 attachment
Data Conversion (com3D->netCDF)

by Mark Branson/CSU
Circulating External Disk Drives

- FAT32 volume format
- USB 2.0 connector
- SRB-client can also be used for non-web data transfer ala rsync
- List-base scripts are also possible if they are desired
Resources

Visualization of Very Large Datasets
Testing (preliminary data point)

- Remote viz on Teragrid
  - TACC (UT Austin) / SPUR
  - Super-LES data
  - ParaView (client-server, remote X-session over ssh)
- TACC<-->Cox Cable (residential ~10Mbps)
  - ‘works’ but not usable
- Need better bandwidth for both data transfer and interactivity
3D visualization of geodesic data

VisIT

3D isocontours of vorticity.

Composite plot of multiple mesh types and variables in the geodesic grid. Cell area (2D cell-centered data) and wind velocity (3D corner-centered on layers) data is shown by pseudocolor plots. Pressure (3D cell-centered on layers) is shown by contour lines.

Plots and movies courtesy of Prabhat (lbnl)
Future Testing

• Remote viz on Teragrid
  – TACC (UT Austin) / SPUR
  – Super-LES data
  – ParaView (client-server, remote X-session over ssh)

• UCSD Cave

• SDSC high-capacity network connections

• CSU network connections

• Other interested parties?
Community Support

Model Code Portability
SAM

CAM

SP-CAM (an MMF)

POP & CICE

CAM

SAM

CCSM

CCSM
Model Taxonomy & Software Configuration Management

• Model codes
• Utility codes
Connecting to TeraGrid

Community Accounts

To address scalability issues, many gateways provide access to TeraGrid resources through a community account rather than setting up unique TeraGrid accounts for each gateway user.

A community account has the following characteristics:

- Only a single community user account (i.e., a TeraGrid username/password) is created.
- The Science Gateway uses the single TeraGrid community user account to launch jobs on the TeraGrid.
- The gateway user running under the community account typically has privileges to run only a limited set of applications.

The chief difference between an individual and a community account is that a community account is essentially a single username on the TeraGrid shared by many (human) users. While this eliminates the need for individual gateway users to request their own TeraGrid accounts, it places additional accounting and security burdens on the gateway developers. To distinguish one gateway user from another, the gateway developer has to institute a user registry and gateway authentication mechanism.

Advanced Support for TeraGrid Applications (ASTA)

Advanced Support for TeraGrid Applications (ASTA) provides collaboration between Advanced User Support (AUS) staff and users of TeraGrid resources. The objective of the program is to enhance the effectiveness and productivity of scientists and engineers. As a part of the ASTA program, guided by the allocation process, one or multiple AUS staff will join the principal investigator’s (PI’s) team to collaborate for up to a year, working with users’ applications.

AUS staff from TeraGrid resource provider sites have expertise and experience in many areas of high performance computing, domain sciences, data management, and grid computing. Collaborative work can include any of the following:

- Porting applications to new resources
- Implementing algorithmic enhancements
- Implementing parallel math libraries
- Improving the scalability of codes to higher processor counts
- Optimizing codes to efficiently utilize specific resources
- Providing help for portal and gateway development
- Assisting with visualization, workflow, and data analysis/transfer

In addition to requesting ASTA at the four quarterly allocations meetings, ASTA can also be requested throughout the year, as a Startup or Supplemental request, via POPS. Please visit the ASTA Project List page to view a list of current ASTA projects.
Backup
### Remote Visualization of Clouds

<table>
<thead>
<tr>
<th>Allocations</th>
<th>Start Date</th>
<th>End Date</th>
<th>Resource</th>
<th>SUs Remaining</th>
<th>SUs Awarded</th>
<th>My Usage (SU)</th>
<th>% Remaining</th>
<th>Alloc. Type</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2009-11-07</td>
<td>2010-11-07</td>
<td>Spur</td>
<td>30,000</td>
<td>30,000</td>
<td>0.0</td>
<td>100%</td>
<td>new</td>
<td>active</td>
</tr>
</tbody>
</table>

### Regionalization of Anthropogenic Climate Change Simulations

<table>
<thead>
<tr>
<th>Allocations</th>
<th>Start Date</th>
<th>End Date</th>
<th>Resource</th>
<th>SUs Remaining</th>
<th>SUs Awarded</th>
<th>My Usage (SU)</th>
<th>% Remaining</th>
<th>Alloc. Type</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2009-04-01</td>
<td>2010-03-31</td>
<td>Ranger</td>
<td>1,989,328</td>
<td>3,950,000</td>
<td>0.0</td>
<td>50%</td>
<td>new</td>
<td>active</td>
</tr>
<tr>
<td></td>
<td>2009-04-01</td>
<td>2010-03-31</td>
<td>Spur</td>
<td>500</td>
<td>500</td>
<td>0.0</td>
<td>100%</td>
<td>new</td>
<td>active</td>
</tr>
</tbody>
</table>

### Modeling Global Climate Variability with the Multi-scale Modeling Framework: The Boundary-layer Cloud Problem

<table>
<thead>
<tr>
<th>Allocations</th>
<th>Start Date</th>
<th>End Date</th>
<th>Resource</th>
<th>SUs Remaining</th>
<th>SUs Awarded</th>
<th>My Usage (SU)</th>
<th>% Remaining</th>
<th>Alloc. Type</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2009-04-01</td>
<td>2010-03-31</td>
<td>Steele</td>
<td>812,564</td>
<td>950,000</td>
<td>0.0</td>
<td>86%</td>
<td>new</td>
<td>active</td>
</tr>
</tbody>
</table>

---

**Wednesday, January 13, 2010**
Data Challenges of GCRM

• Extremely high volumes of data
  – 10 GB/ variable / step
  – 1-10 petabytes / simulated year
  – Can’t just move data to local systems
  – Data will have to be on-offline
  – 4 byte offsets exceeded
  – Huge number of files per simulation
  – Model for running analysis on the entire data set needed
Data Challenges of the GCRM (cont)

• Geodesic Grid
  – Preliminary (but not sufficient) support in some analysis tools
  – Standards for complete description not defined
  – Hyperslab-ing on coordinate values not supported and very costly
  – Grid itself is large (~ 2GB)
CMMAP’s Stable of Models

- Current
  - Super-CAM
  - SAM
  - VVM
- Under development
  - GCRM
  - Q3D MMF

What do we mean when we say we ‘...are running the MMF’

Who’s using what and what modifications are being made?
Internet Data Transfer Capacity

The graph illustrates the data transfer capacity over time for various transmission speeds. The x-axis represents time in days, while the y-axis represents data volume in bytes. The graph includes data for IPv6 2008 Internet2 Land Speed Record, MMF LES Super-run, and other transmission speeds such as 1 Mbps, 100 Mbps, 1 Gbps, 10 Gbps, 100 Gbps, and linear speeds at 100 Gbps, 10 Gbps, 1 Gbps, and 20.42 TB@300 min. The graph shows the capacity to transfer large volumes of data efficiently.
Conventions for Naming Datasets

spcam_climo_-3C-12C_1e-5_re7-10_vt01----> CMMAP_SUPERLES

- Standardized names aid communication and data management
- Current convention is adapted from oceanographic collections done for SIO and IODP
- None apparent for climate modeling (e.g., AMIP, CF)
- DOI (digital object identifiers) are used in the publishing world
- This is similar to a DOI and we may add a DOI if a DOI authority emerges for atmospheric data
- In the meantime we should have something equivalent

CMMAP_SUPERLES_20080725014445402_20080725014445402_spcam3.clm2.r.0003-01-01-00000.mif

- Standardized prefix is added to facilitate search and identification
- Original filename is untouched
Timeline

- NetCDF 3.0
  - 1996
- NetCDF 3.6
  - 2000
- NetCDF4.0/4.0.1beta
  - 2004
- NetCDF4.1?
  - 2006
- Pnetcdf 1.0
  - 2008
- ADIOS alpha
  - 2010
- HDF5
  - 1998
- Netcdf4.0 beta
  - 2002
- 64 bit PNetCDF/NetCDF?
  - 2004
- Pnetcdf 1.0.3
  - 2006
- Wednesday, January 13, 2010
Action Items

- Data conversion code: com3d-netCDF (Helly, Branson)
- add sample dataset
- Data Distribution of LES dataset (Helly)
- NSF LRAC Allocation Proposal w/technical support (Helly, Randall, Ackerman/Bretherton, Krueger)
- Digital Library search interface extension (Helly)
- OAI metadata server for external interoperability (Helly)
- Intercalibration observations for digital library (Helly, Ackerman)
- Prepare version control for MMF 2.0 and GCRM
NetCDF Status

Real large data support coming soon!

Netcdf4.0
Limited Large data

today

Netcdf4.x
[largedata]

Netcdf3.x
[largedata]

Netcdf3.6
[largefiles]

Netcdf2

Pnetcdf1.0
[largefiles]

Pnetcdf1.x
prerelease
[largedata]

Caution:
Pnetcdf output not currently compatible with other formats