Cyberinfrastructure

John Helly
San Diego Supercomputer Center
University of California, San Diego

Outline

- CMMAP Data and Computing Architecture Goals
- Management Approach: Cyberinfrastructure Working Group
- Cyberinfrastructure-based Research
- Computing futures
- GCRM Data Challenges (K. Schuchardt)
- Ideas & Suggestions

Preface

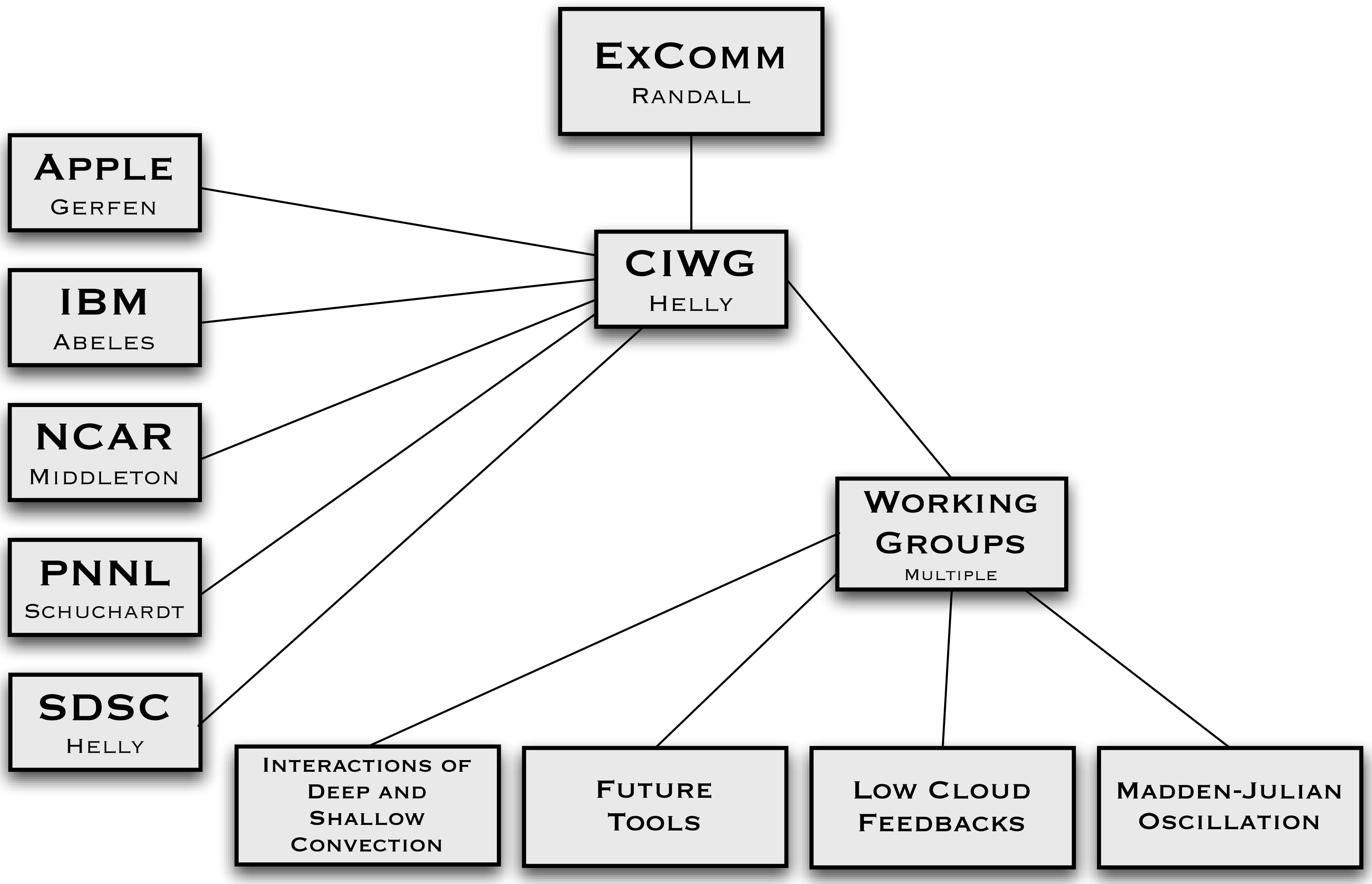
- CMMAP cyberinfrastructure is shared, not owned
  - Collaboration and leveraging are key to acquisition and efficient use of resources
  - CMMAP-owned computing and data resources will come from separate proposals
- Training the next generation of scientists in a state-of-the-art computing environment is essential
  - Want graduate student involvement in every aspect of computing activities

Management Approach Cyberinfrastructure Working Group

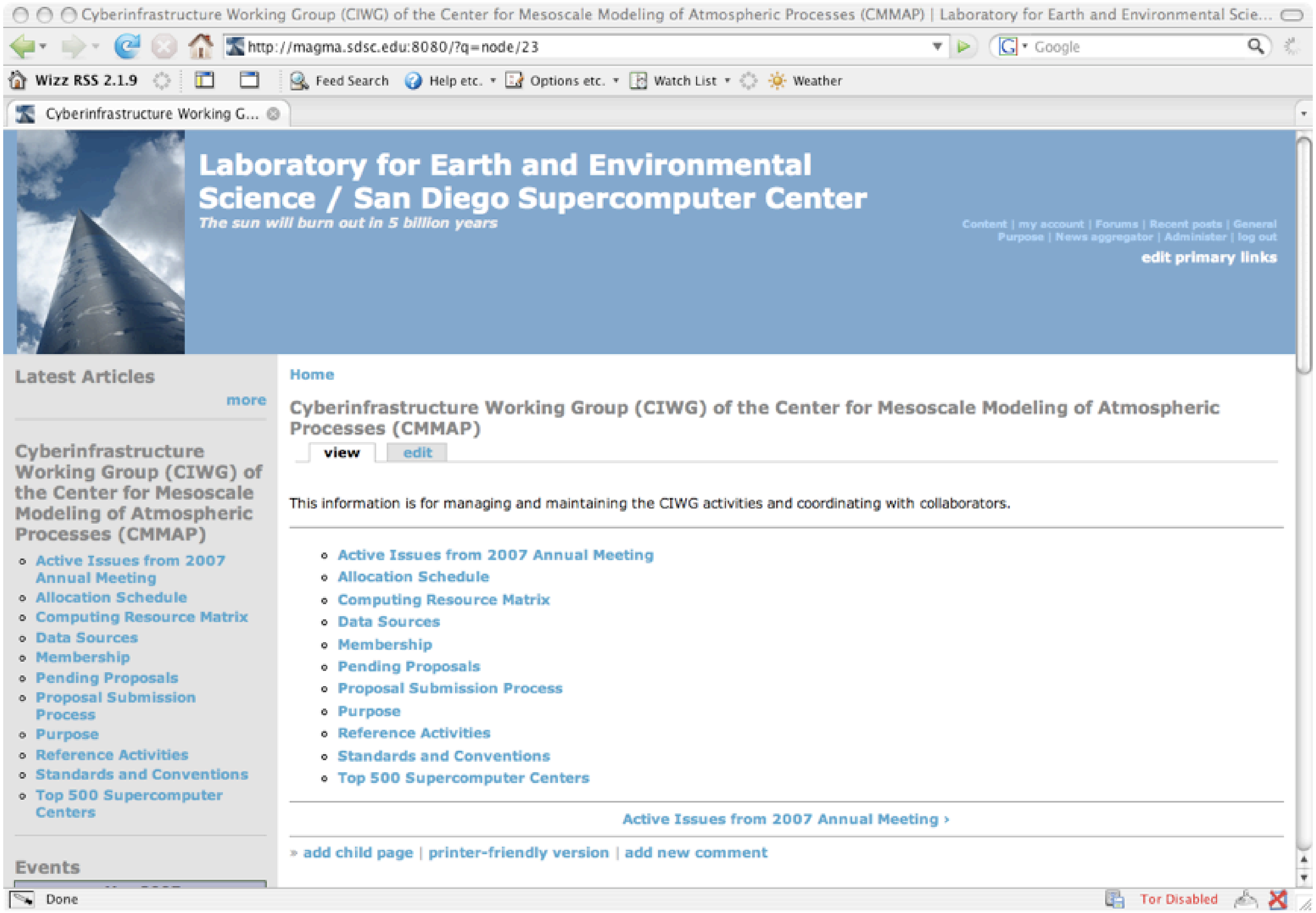
CIWG Objectives

- Make efficient use of computing and data resources
  - acquire resources
  - coordinate resource utilization
  - collaborate to leverage joint efforts
- Validate goals and provide advice and consent to Executive Committee

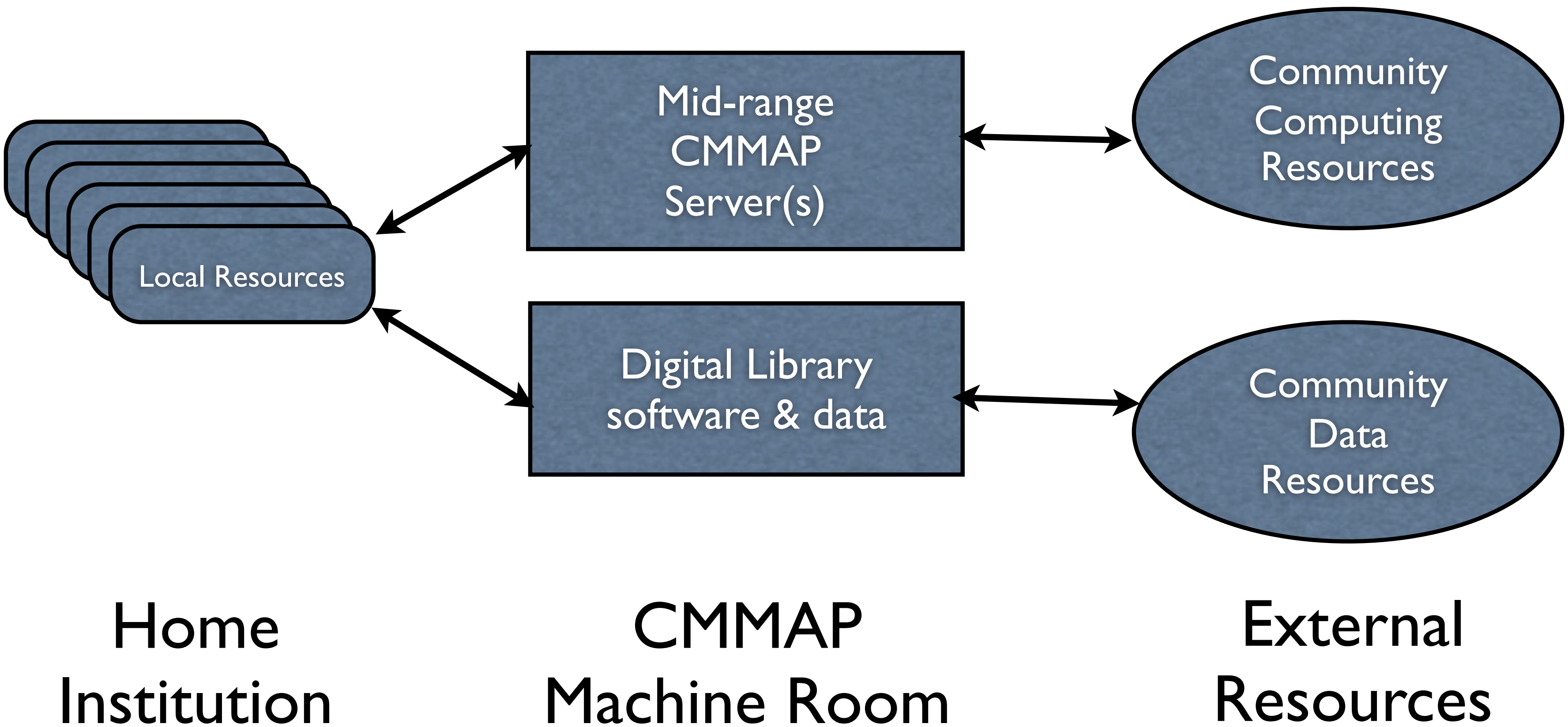
Cyberinfrastructure Organization

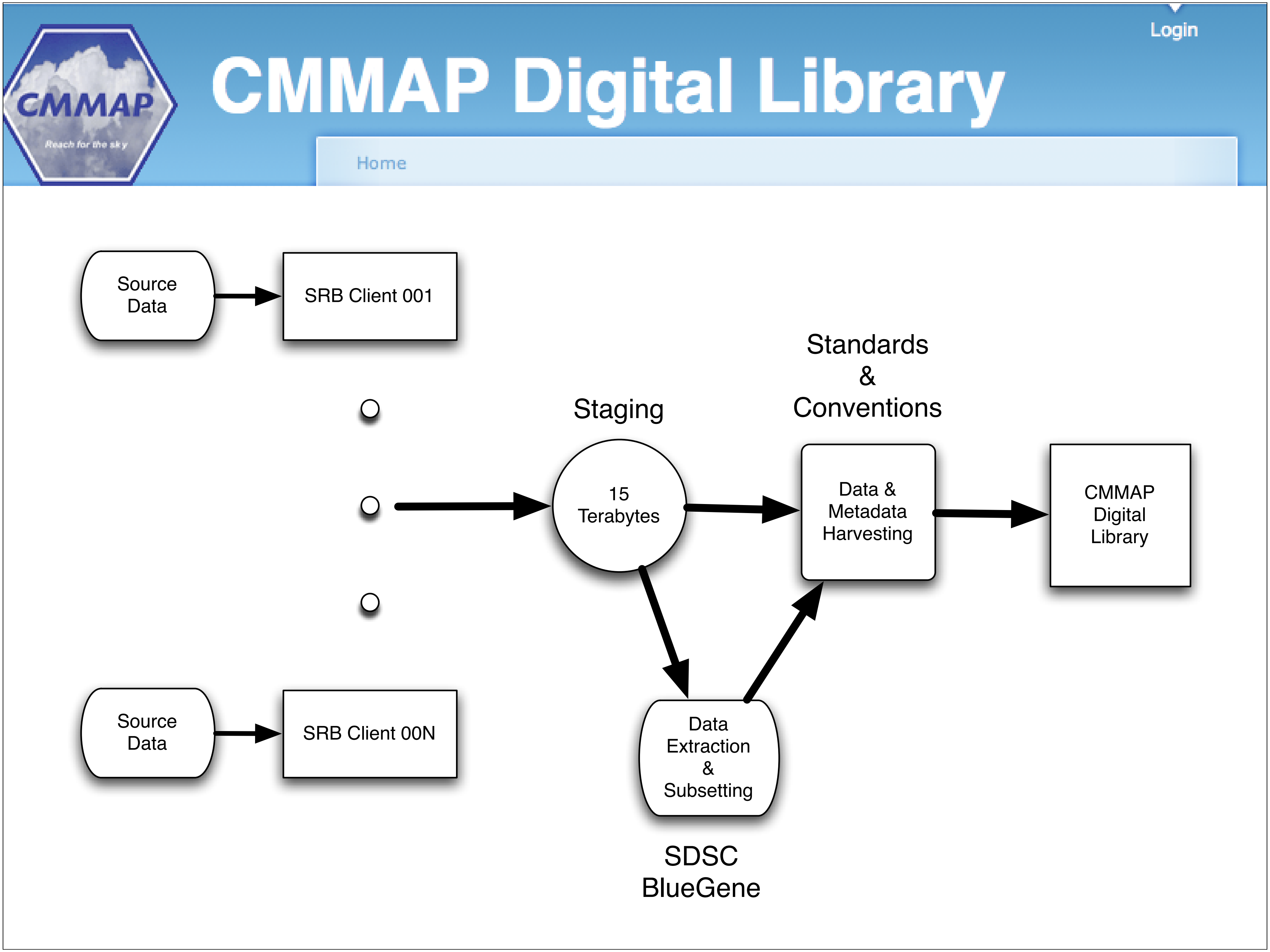


CIWG Management Resources



Architectural Goals




CMMAP
Reach for the sky
CMMAP Digital Library
Login
Home
Source Data
SRB Client 001
Source Data
SRB Client 00N
Staging
15 Terabytes
Standards & Conventions
Data & Metadata Harvesting
CMMAP Digital Library
Data Extraction & Subsetting
SDSC BlueGene

Adopting Community-based Standards and Conventions

Table 2. Tabulation of resources of interest identified by workshop participants.

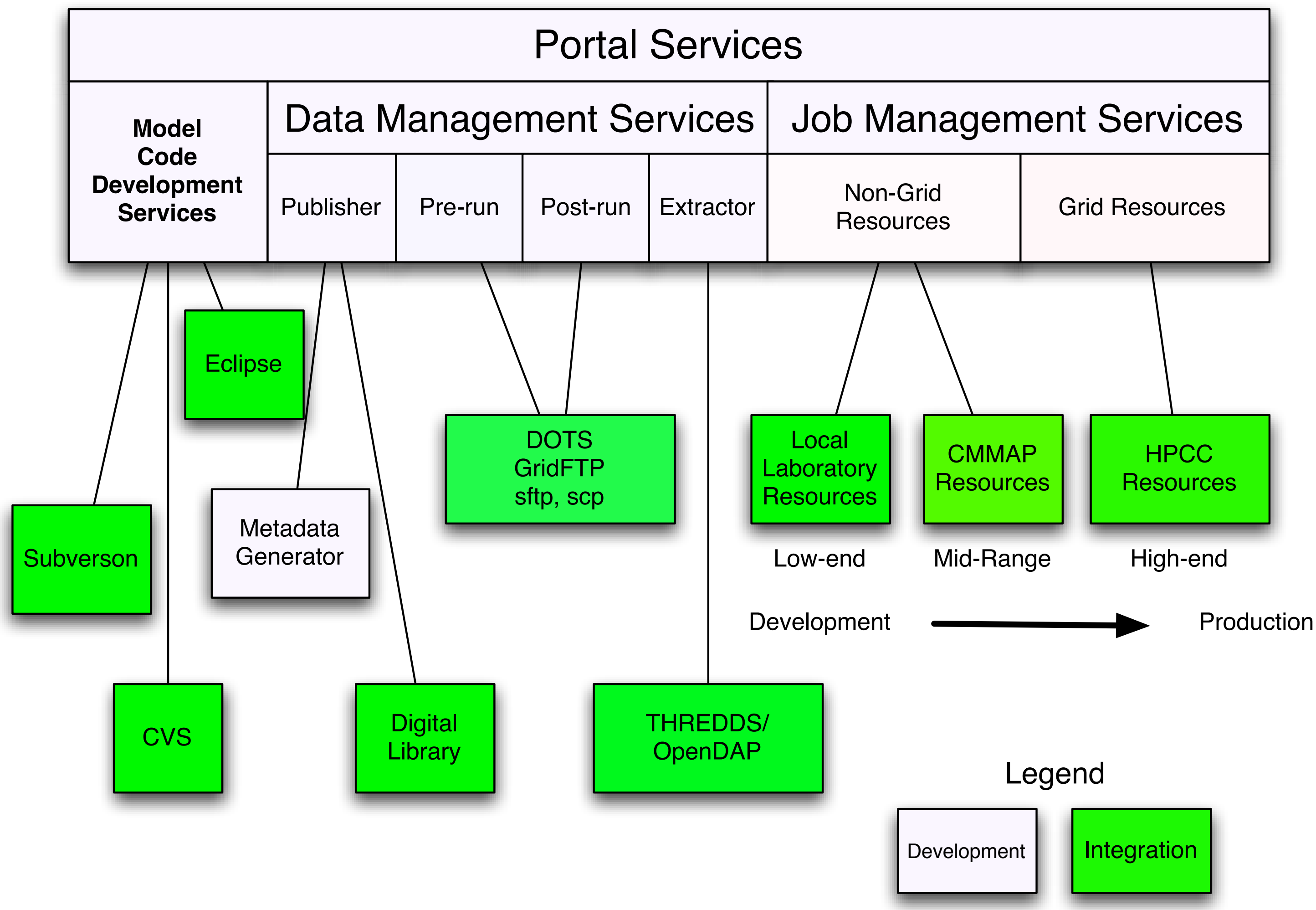
Category	Common Name	Resource Description or Reference
Standards and Conventions	Climate and Forecast (CF) Standard Names	http://www.cgd.ucar.edu/cms/eaton/cf-metadata/CF-current.html#sname
	COARDs	ftp://ftp.unidata.ucar.edu/pub/netcdf/Conventions/COARDS
	Digital Object Identifier System	http://www.doi.org/
	Dublincore	http://dublincore.org/index.shtml
	EML	http://knb.ecoinformatics.org/software/eml/eml-2.0.1/index.html
	Fedora	http://www.fedora.info
	FGDC	http://www.fgdc.gov/
	Grib	http://www.wmo.ch/web/www/WDM/Guides/Guide-binary-2.html
	Geographic Markup Language	http://www.opengis.net/gml/
	Geoscience Markup Language	http://www.opengis.net/GeoSciML
	HDF	http://hdf.ncsa.uiuc.edu/
	ISO 19115	Defines the schema required for describing geographic information and services.
	ISO 19139	Geographic information -- Metadata -- XML schema implementation
	MMI	http://marinemetadata.org
	NetCDF	http://www.unidata.ucar.edu/software/netcdf/
	OAI	http://www.openarchives.org/
	OBIS	http://www.iobis.org/
	OpenGIS	http://www.opengeospatial.org/
	SensorML	http://www.opengeospatial.org/projects/groups/sensorweb
	STD-DOI	http://www.std-doi.de
Transport Protocol	GridFTP	http://www.globus.org/grid_software/data/gridftp.php
	HTTP	http://www.w3.org/Protocols/
	OPeNDAP	http://www.opendap.org/
	REST	http://en.wikipedia.org/wiki/Representational_State_Transfer#References
	SOAP	http://www.w3.org/TR/soap/

CMMAP Best Practices Evolution

- Learn from CCSM
- Languages
  - F90, F95,F77, C,C++,....
- Programming models
  - MPI,OpenMP, HPF,...
  - Global Arrays
- Data formats
  - netCDF (in various flavors)
- Fault tolerance strategies for model codes

CMMAP Modelers' Workbench

(separately pending proposal)



High-Performance Computing Futures: Big Issues

- Movement of vendors to multi-core chips is problematic for legacy codes and probably future codes
  - creating problems with memory limitations
  - big memory machines are becoming increasingly scarce How to deal with multi-core chips?
  - MPI is considered by some to be a failure of the computer science community as it is too hard for general use.
  - probably lead to a hybrid computing model related to earlier approaches
    - OpenMP
    - High Performance Fortran (HPF)
- How to connect these multi-core chips in a network (within a machine)?
  - Infiniband BW (~1GB/sec nominal called 4x SDR single-data-rate) is not keeping pace with multi-core and leads to cabling problems.
  - switches are also a problem
- Compilers
  - don't deal with multiple cores well
  - so, the burden is on the programmer

Cyberinfrastructure-based Research

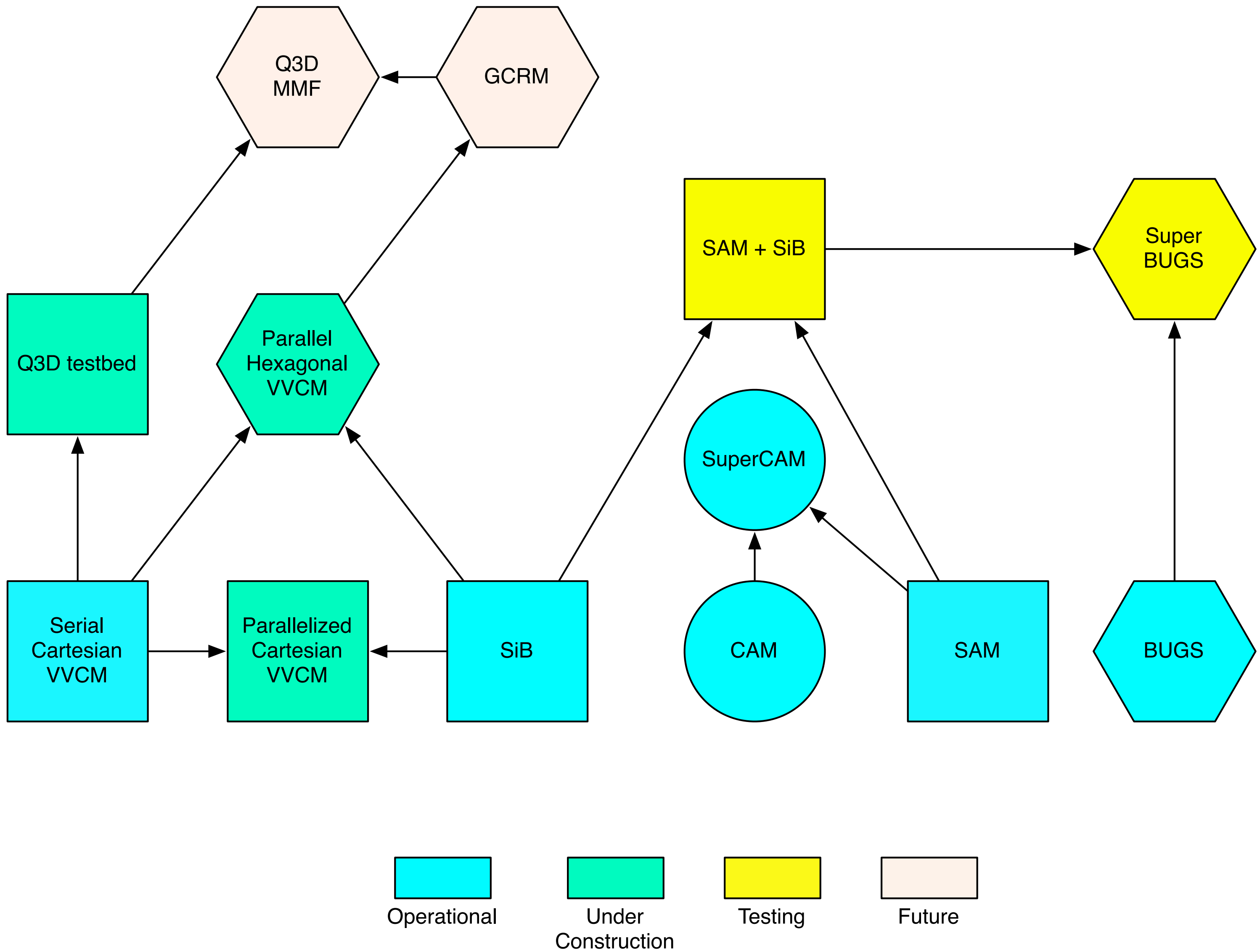
Leveraging National & Partner Resources

	Organization	Resource	Amount
Data Allocations	San Diego Supercomputer Center (SDSC)	Disk	15 Terabytes
		BlueGene	30,000 SUs*
Computing Allocations	Teragrid (multi-institution)	SDSC DataStar (IBM SP4)	600,000 SUs
		Grid Roaming	600,000 SUs
	Lawrence Berkeley National Laboratory (LBNL)	National Energy Research Scientific Computing Center (NERSC)	700,000 SUs
	National Center for Atmospheric research (NCAR)	BlueIce IBM Power5	500,000 SUs
	IBM Watson Research Center	BGW - eServer Blue Gene Solution	TBD
	Stonybrook		TBD

Working Group-defined Computational Experiments

ID	Title	Description	POC
MJO001	MJO Forecasting	Case studies of MJO events on the scale of the Indian Ocean/Western Pacific using NWP models (probably nested models forced at the lateral boundaries by global model analysis and compare with satellite measurements (e.g., A-train, TRMM, AIRS etc). Full orographic complexity, inner domains have explicit convection, outer domains convective parameterization.	Mitch Montcrieff (NCAR) Marat Khairoutdinov (STONYBROOK)
MJO002	MJO Hindcasting	Hindcasts using MMF of observed MJO events	Mitch Montcrieff (NCAR) Marat Khairoutdinov (STONYBROOK)
MJO003	MJO aquaplanet	Simulations of MJOs in an aquaplanet version of MMF (e.g., sensitivity to microphysics, SST distribution, convective transport and mesoscale momentum transport.	Mitch Montcrieff (NCAR) Marat Khairoutdinov (STONYBROOK)
MJO004	MJO Multi-scale Evaluation	Use the aquaplanet simulations to evaluate multi-scale analytic models.	Mitch Montcrieff (NCAR) Marat Khairoutdinov (STONYBROOK)

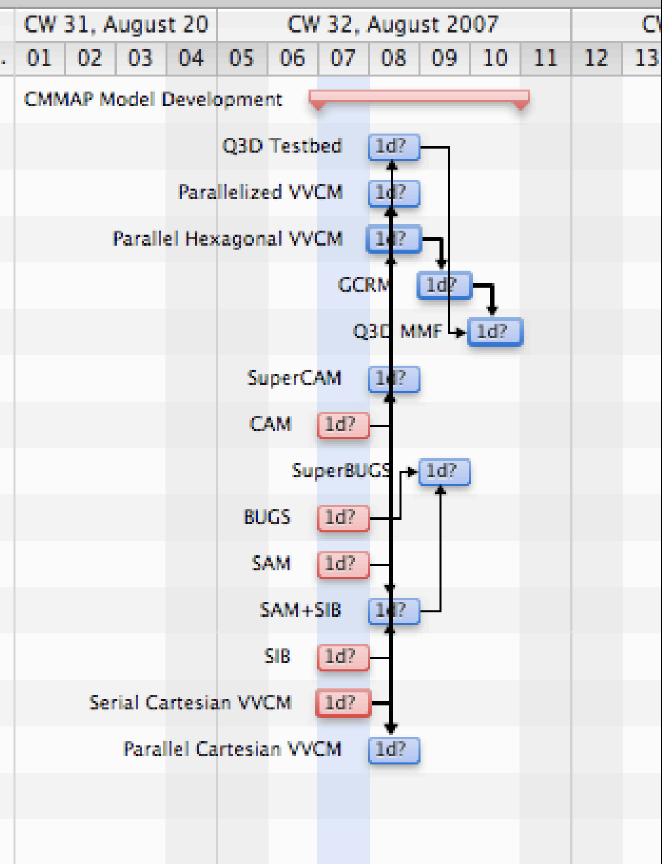
Model Taxonomy



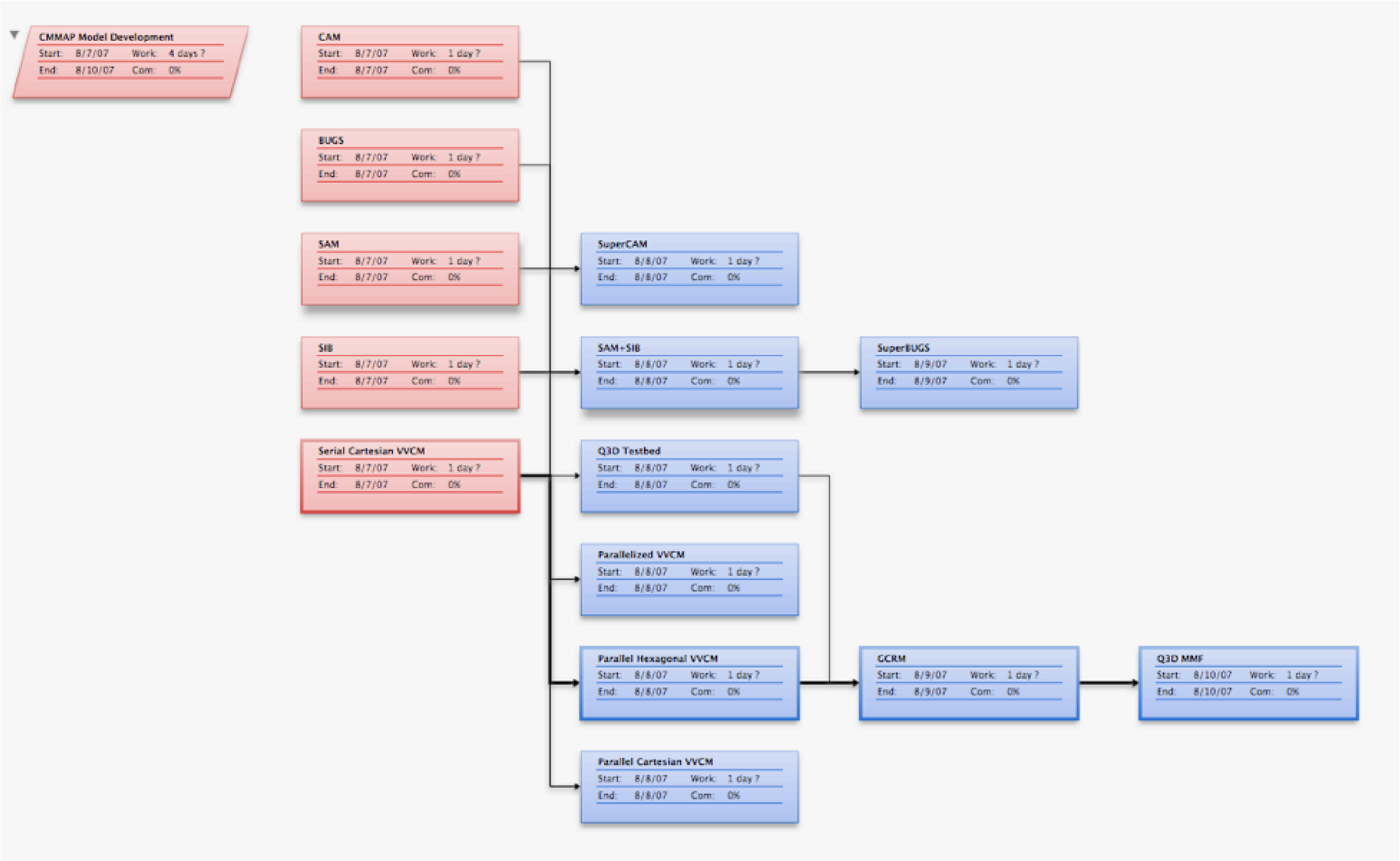
Coordination of Resources and Development

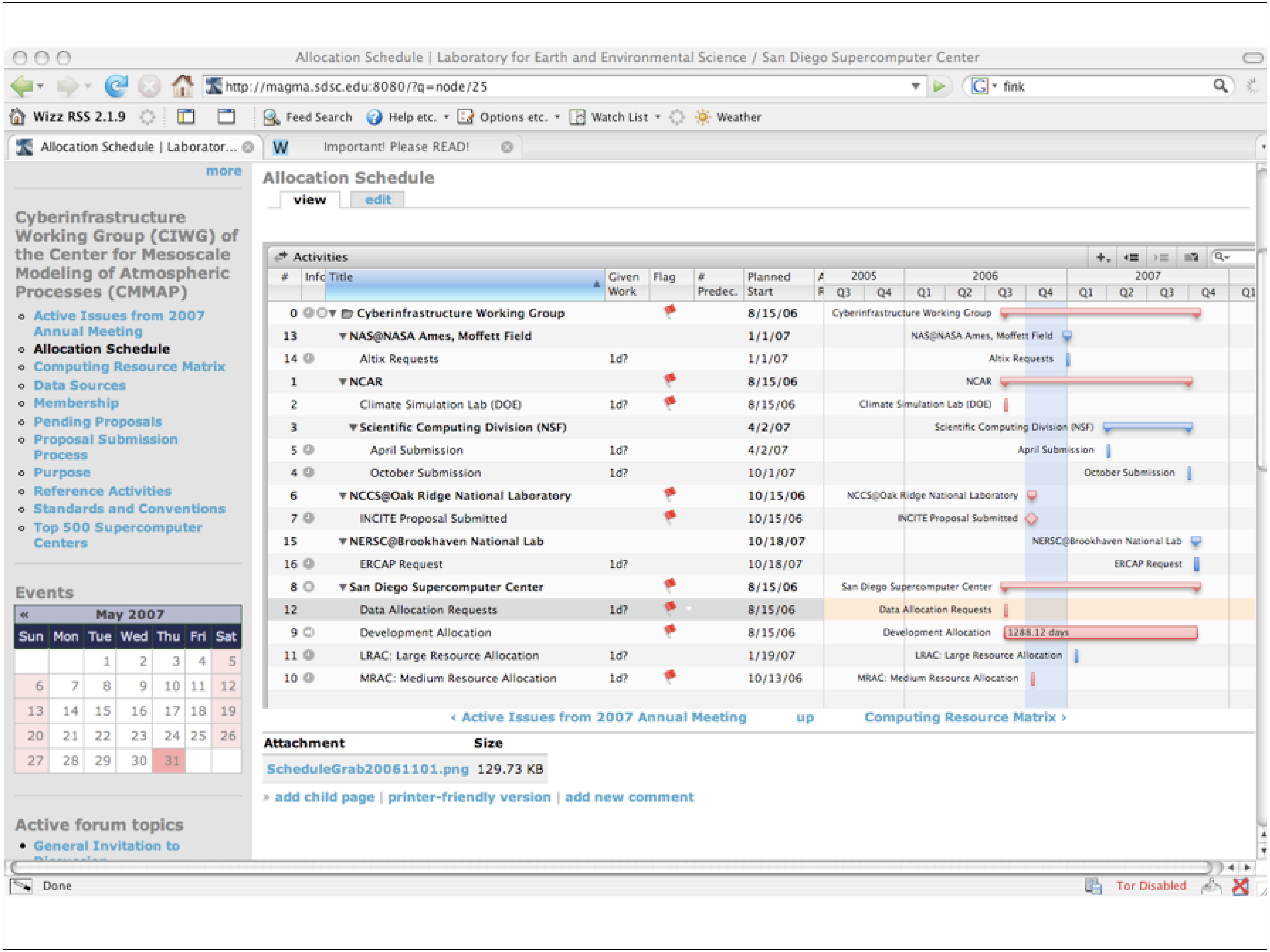
Activities

Title	Given Work	Flag	# Predec.	Expected Start	Assigned Resources	% Compl.
▼ CMMAP Model Development				8/7/07		0%
Q3D Testbed	1 day ?		13	8/8/07		0%
Parallelized VVCM	1 day ?		12; 13	8/8/07		0%
Parallel Hexagonal VVCM	1 day ?		13	8/8/07		0%
GCRM	1 day ?		3	8/9/07		0%
Q3D MMF	1 day ?		1; 4	8/10/07		0%
SuperCAM	1 day ?		7; 10	8/8/07		0%
CAM	1 day ?			8/7/07		0%
SuperBUGS	1 day ?		9; 11	8/9/07		0%
BUGS	1 day ?			8/7/07		0%
SAM	1 day ?			8/7/07		0%
SAM+SIB	1 day ?		10; 12	8/8/07		0%
SIB	1 day ?			8/7/07		0%
Serial Cartesian VVCM	1 day ?			8/7/07		0%
Parallel Cartesian VVCM	1 day ?		12; 13	8/8/07		0%



Model Development Workflow





Allocation Schedule | Laboratory for Earth and Environmental Science / San Diego Supercomputer Center
http://magma.sdsc.edu:8080/?q=node/25
fink
Wizz RSS 2.1.9
Feed Search
Help etc.
Options etc.
Watch List
Weather
Allocation Schedule | Laborator...
Important! Please READ!
more
Cyberinfrastructure Working Group (CIWG) of the Center for Mesoscale Modeling of Atmospheric Processes (CMMAP)
Active Issues from 2007 Annual Meeting
Allocation Schedule
Computing Resource Matrix
Data Sources
Membership
Pending Proposals
Proposal Submission Process
Purpose
Reference Activities
Standards and Conventions
Top 500 Supercomputer Centers
Events
May 2007
Sun Mon Tue Wed Thu Fri Sat
1 2 3 4 5
6 7 8 9 10 11 12
13 14 15 16 17 18 19
20 21 22 23 24 25 26
27 28 29 30 31
Active forum topics
General Invitation to
Allocation Schedule
view
edit
Activities
Info Title Given Work Flag # Predec. Planned Start
2005 2006 2007
Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1
0 Cyberinfrastructure Working Group 8/15/06
13 NAS@NASA Ames, Moffett Field 1/1/07
14 Altix Requests 1d? 1/1/07
1 NCAR 8/15/06
2 Climate Simulation Lab (DOE) 1d? 8/15/06
3 Scientific Computing Division (NSF) 4/2/07
5 April Submission 1d? 4/2/07
4 October Submission 1d? 10/1/07
6 NCCS@Oak Ridge National Laboratory 10/15/06
7 INCITE Proposal Submitted 10/15/06
15 NERSC@Brookhaven National Lab 10/18/07
16 ERCAP Request 1d? 10/18/07
8 San Diego Supercomputer Center 8/15/06
12 Data Allocation Requests 1d? 8/15/06
9 Development Allocation 8/15/06
11 LRAC: Large Resource Allocation 1d? 1/19/07
10 MRAC: Medium Resource Allocation 1d? 10/13/06
1288.12 days
‹ Active Issues from 2007 Annual Meeting
up
Computing Resource Matrix ›
Attachment Size
ScheduleGrab20061101.png 129.73 KB
» add child page | printer-friendly version | add new comment
Done
Tor Disabled

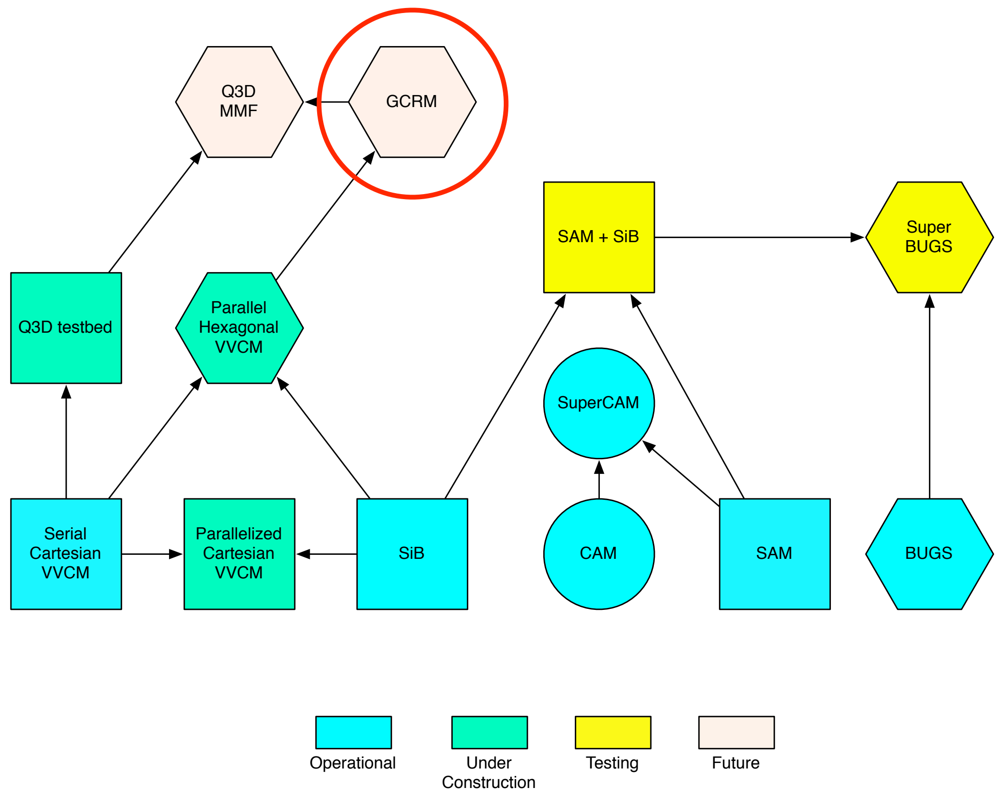
Collaborations

Participants	MMF Bluegene	Data Extraction & Subsetting	Data & Metadata Interoperability
IBM			
NCAR			
PNNL			
SDSC			
Stonybrook			

Data Extraction & Subsetting

Transpose Method

- Take 1000s of files containing timesteps of all variables into 10-100s of file containing full time-series of a single variable
- Discussions at IBM Watson with Bluegene group led to idea to use the BG memory as a file system and do the extraction using one processor per file then combining results
    - BG has limited memory so this is a good problem for this type of architecture
- Parallelized netCDF transposer code is being tested at SDSC using AMIP data from Marat



Data Challenges of GCRM

Karen Schuchardt / PNNL

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)

Data Challenges of the GCRM (cont)

- Current scalar analysis tools break down
  - Insufficient memory
  - Assumptions made for smaller data sets no longer valid
  - Screen resolution exceeded

Current thrust Areas

- IO Strategies for the GCRM
  - Benchmarking, APIs, Data Layout
- Evaluation and Adaptation of data "manipulation and analysis" tools
- Web Portal
- Enhanced metadata including data signatures

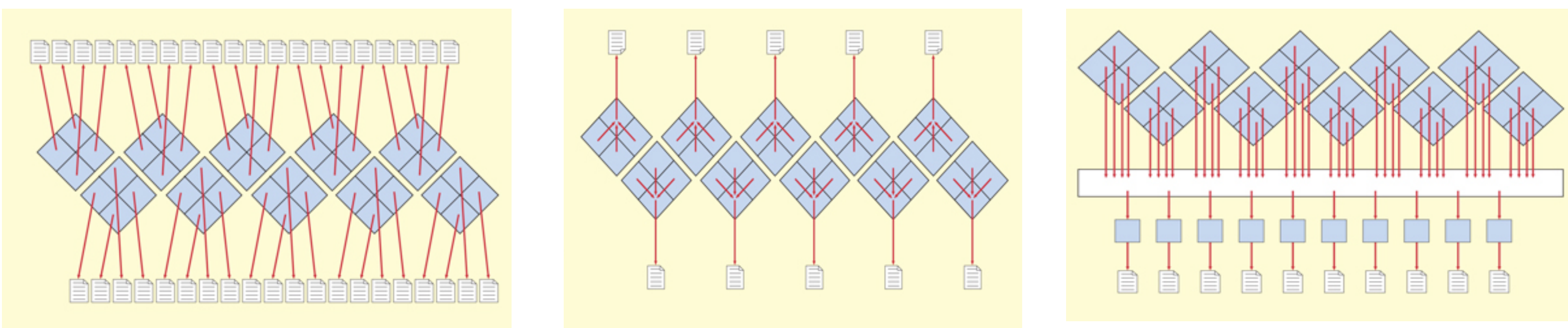
IO Strategies

Tradeoffs:

- Minimize blocking of computation
- Maximize bandwidth to file system
- Mimimize memory requirements
- use format that supports data access and analysis

Progress

- Developing and Benchmarking IO codes
  - XT4, BlueGene, HP cluster
- Evaluating writing of one variable per file
- Evaluating parallel IO libraries

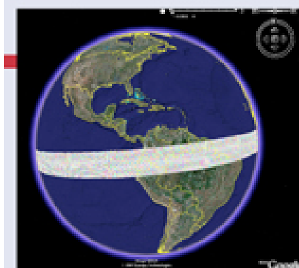


Data Manipulation & Analysis

Objective Evaluate existing tools with respect to their capability to handle large, high resolution data; identify where parallelization will be necessary and what type of parallelization to use

Progress

- Generated evaluation data sets up to 3 km resolution (randomized)
- Scalar tools break down at 30 km resolution; targetting 2-4 km
- Creating hyperslabs based on coordinate values not supported
- Completeness of aggregation strategies not clear (CDAT, ESG)

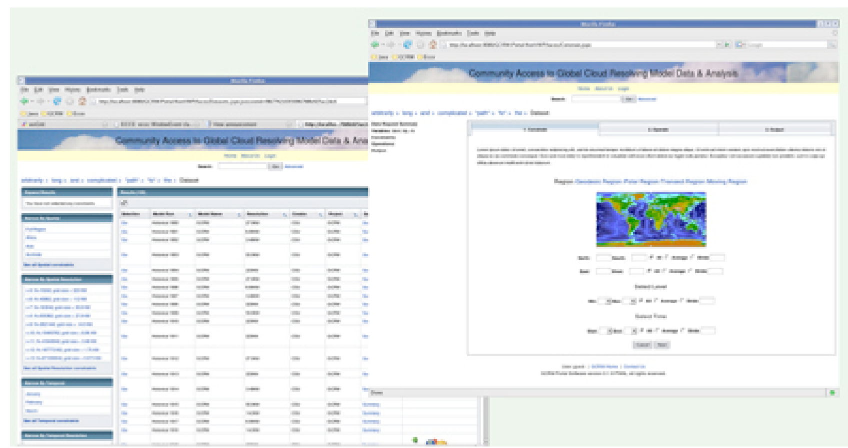


Results of subsetting geodesic data set to extract one variable at a single level over +/- 5 degrees latitude. An equatorial slice is a worst-case performance scenario.

Resolution *(km)*	# slabs	Minutes *(1 processor)*
3.49	13926	~75
6.98	7391	~10
14.0	3636	~1
27.9	1751	~ 0.2

Web Portal

Objective Develop web interface that facilitates sharing and disovery of data and provides access to reduced data sets, visualizations, and ultimately in-situ analysis

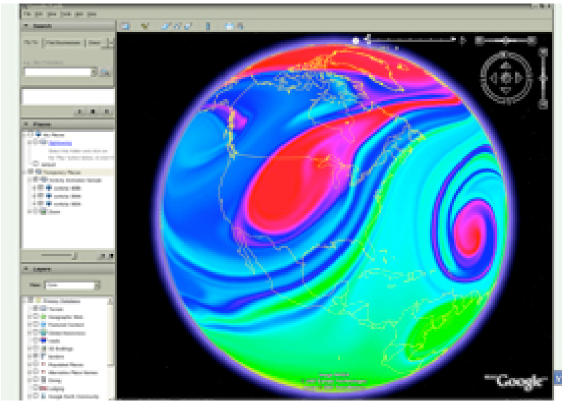


Progress

- Preliminary architecture defined
- Development of initial prototype started
- Conversing with folks such as ESG, Curator, etc to figure out best leverage each other work

Data Signatures

- Need to generate more metadata that characterizes data sets to minimize need to get all the data from storage
  – Graphs
  – GoogleEarth images



GCRM Data Summary

- Efforts in multiple areas
  - IO Strategies
  - Data manipulation tool enhancements
  - Web portal
  - Data signatures
- Leveraging
  - Other efforts in earth systems portals (ESG, etc)
  - Numerous existing analysis packages
  - Other SciDAC efforts (Parallel NetCDF)
- Interested to apply this work to other CMMAP models
  - Looking for existing high resolution data sets to work with

Ideas, Suggestions, Issues

- Model usage workshop (Steve Kruger)
- Participation in verification objective group (Bill Rossow)
- CMMAP policy for code management and documentation (issue raised by site visitors)
- Effective programming techniques for programmers and graduate students
- Integration/relationship to international archiving activities (e.g., TIGGE)