Cyberinfrastructure

11th Team Meeting
Fort Collins, CO
August, 2011

CMMAP
Reach for the sky

Cyberinfrastructure Working Group Agenda

- Resources
- Parallel I/O (Schuchardt)
- Model Code Portability
- Community Portal for MMF
- Data Interoperability
- Giga-???? (LES, VVM)
- Viz of Very-Large Datasets

CMMAP
Reach for the sky

Resources

The Teragrid is Dead. Long Live the Teragrid

News & Highlights

After 10 years of service to the national science and engineering community, the TeraGrid project has come to an end. It is succeeded by a new National Science Foundation program called XSEDE -- the Extreme Science and Engineering Digital Environment. See www.xsede.org for information.

✏ **Manage** ▾ ☑ **Toggle Edit Controls**   🏠 **Go to** ▾ | 👤 **John J Helly** (Sign Out)

XSEDE | USER PORTAL

Extreme Science and Engineering
Discovery Environment

Search XSEDE... 🔍

HOME   MY XSEDE   RESOURCES   DOCUMENTATION   ALLOCATIONS   TRAINING   USER FORUMS   HELP

Resource Monitor   Scheduled Downtimes   File Manager   Queue Prediction   Remote Visualization   Science Gateways

HPC SYSTEMS	ADVANCED VIS SYSTEMS	STORAGE SYSTEMS	SPECIAL PURPOSE SYSTEMS

NAME	INSTITUTION	SYSTEM	PEAK TFLOPS	MEMORY TBYTES	STATUS	LOAD	RUNNING JOBS	QUEUED JOBS	OTHER JOBS
Kraken	NICS	Cray XT5	1174.00	147.00	Up		503	312	142
Ranger	TACC	Sun Constellation Cluster	579.40	123.00	Up		411	509	170
Lonestar	TACC	Dell Linux Cluster	302.00	45.00	Up		0	309	78
Trestles	SDSC	Appro AMD Magny-Cours Cluster	100.00	20.25	Up		141	0	180
Steele	Purdue	Dell Intel 64 Linux Cluster	60.00	12.40	Up*				
Lincoln	NCSA	Dell/Intel PowerEdge 1950	47.50	3.00	Up		30	27	1
Blacklight	PSC	SGI UV	37.20	32.00	Up		94	164	1
Pople	PSC	SGI Altix 4700	5.00	1.54	Up*		40	0	21
Dash	SDSC	Appro Intel Nehalem Cluster	4.90	3.00	Up		2	3	0
		Total:	2310.0	387.19			1221	1324	593

*Indicates failure of one or more status test.

Hover mouse pointer over Resource Name, Resource Status, and headings to see additional information.

Thursday, August 11, 2011

http://www.sdsc.edu/us/resources/gordon/

gordon sdsc

Apple   Yahoo!   Google Maps   YouTube   Wikipedia   News ▾   Popular ▾

✉ Contact Us   ⊞ Site Map   ◈ Staff Directory

SDSC | User Support
ADVANCING SCIENCE THROUGH SERVICES

Search SDSC User Support   **GO**

| COMPUTE & DATA RESOURCES | CONSULTING | TRAINING | ALLOCATIONS | USER TOOLS |

User News   TAPP   SAC   Visualization Services   About User Services

SDSC > User Support > Compute & Data Resources > Gordon

Gordon User Guide: Technical Summary

Gordon User Guide

- ▸ Gordon Home & Technical Summary
- ▸ Logging In
- ▸ Jobs
- ▸ Storage
- ▸ Software

Gordon will be a powerful 1024-node supercomputer, unique in its extensive use of flash memory and virtual shared-memory "supernodes". Gordon's architecture will allow it to reduce solution times of data-intensive problems. It is scheduled to be installed in mid-2011.
Gordon is the follow-up to Dash, the first supercomputer to use flash memory.

Contact Us
consult@sdsc.edu
1-866-336-2357
(Toll-Free)

🔒 **Report a security incident**

Important Links
Quick Comparison: Storage Resources

Technical Summary

- 1024 compute nodes
- 32 supernodes, each with:
  - 32 compute nodes at 240 gigaflops per node
  - 2 I/O nodes with 4 TB of flash memory per node
  - 10 TB of memory (2 TB of DRAM and 8 TB of flash memory)
  - Peak performance of 7.7 teraflops
- Peak system performance of 245 teraflops
- 64 TB of DRAM
- 256 TB of flash memory and 4 PB of disk storage

Did You Get What You

Community Portal for MMF Experiments

Model Run Management App

Output returned to SDSC for post-processing analysis and publication

San Diego Supercomputer Center

MMF CMS module

CMMAP Digital Library

Subversion Code Repository

Job Submission and Monitoring

Subversion Working Copy

Remote HPCC Computer (e.g., Steele, Kraken)

Web Browser

Local terminal for microphysics code development

Extension to Data Analysis and Visualization on other HPCC resources such as Trestles and Gordon (SDSC) using new high capacity storage systems (Data Oasis) and virtual point-to-point network connections (DYNES)

Parallel I/O

CMMAP Data Services Update

Karen Schuchardt

Ft Collins

August 2011

Pacific Northwest
NATIONAL LABORATORY

IO Reminder Slide

- **PNetCDF**
  - Stable 1.2 release for quite some time
  - Only PNetCDF can process large files generated by PNetCDF
  - Relies strictly on collective IO

- **NetCDF4**
  - Current Version 4.1.3 not as stable or stress tested
  - Designed to sit on top of HDF5(more complicated to build)
  - still has large variable problem (fortran)
  - Can use collective or independent IO

- **BUT NetCDF4 build can specify to use PNetCDF underneath**

Pacific Northwest
NATIONAL LABORATORY

ZGrd IO Performance Profile - PNetCDF

▶ IO Configuration

- Hopper 80 OSTs, collective IO (multiple options)
- 100 model timesteps/write (every 2 minutes with 12 sec timestep)
- 8% of time in IO including averaging
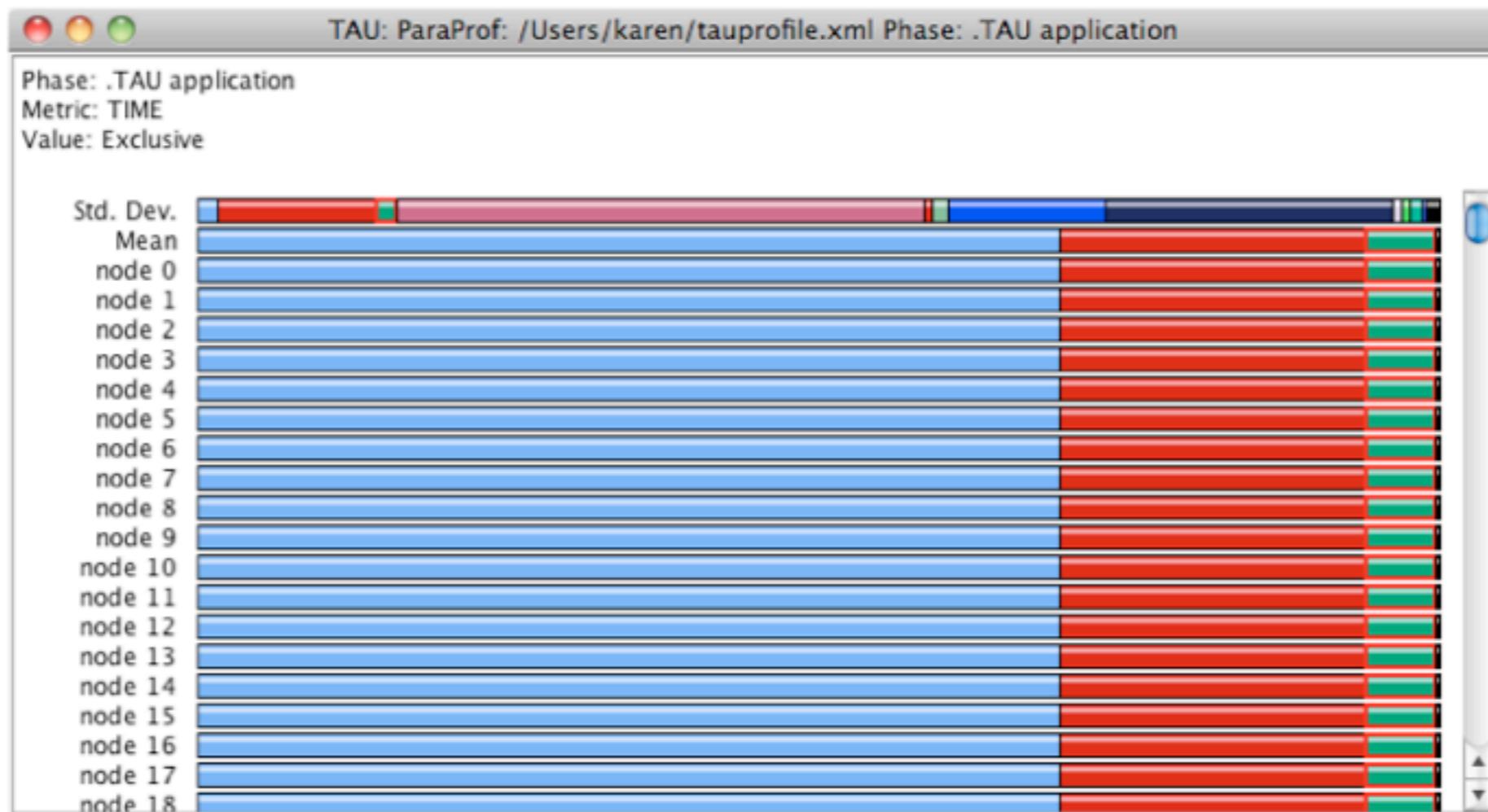
IO in ZGrd

- Supports PNetCDF, NetCDF4
- Restart is working with physics
- Supports instantaneous and averaged data
- Supports collective IO, Application Aggregation, and Psuedo non-blocking
- Supports per variable time series files or all variables in one file
- Testing on hopper, franklin; Intrepid (BG/P planned)
  - Up to 20,000 processors, primarily 10,000
  - Primarily very short runs for profiling
- History Output
  - 4km, 100 vertical (4 Billion cells, 8 B corners, 12 B edges)
  - 3 surface vars, ~ 12 center vars, ~ 3 corner vars (and growing)
  - ~400 GB per timestep

Pacific Northwest
NATIONAL LABORATORY

Community Support
Model Code Portability
and
Future Architectures
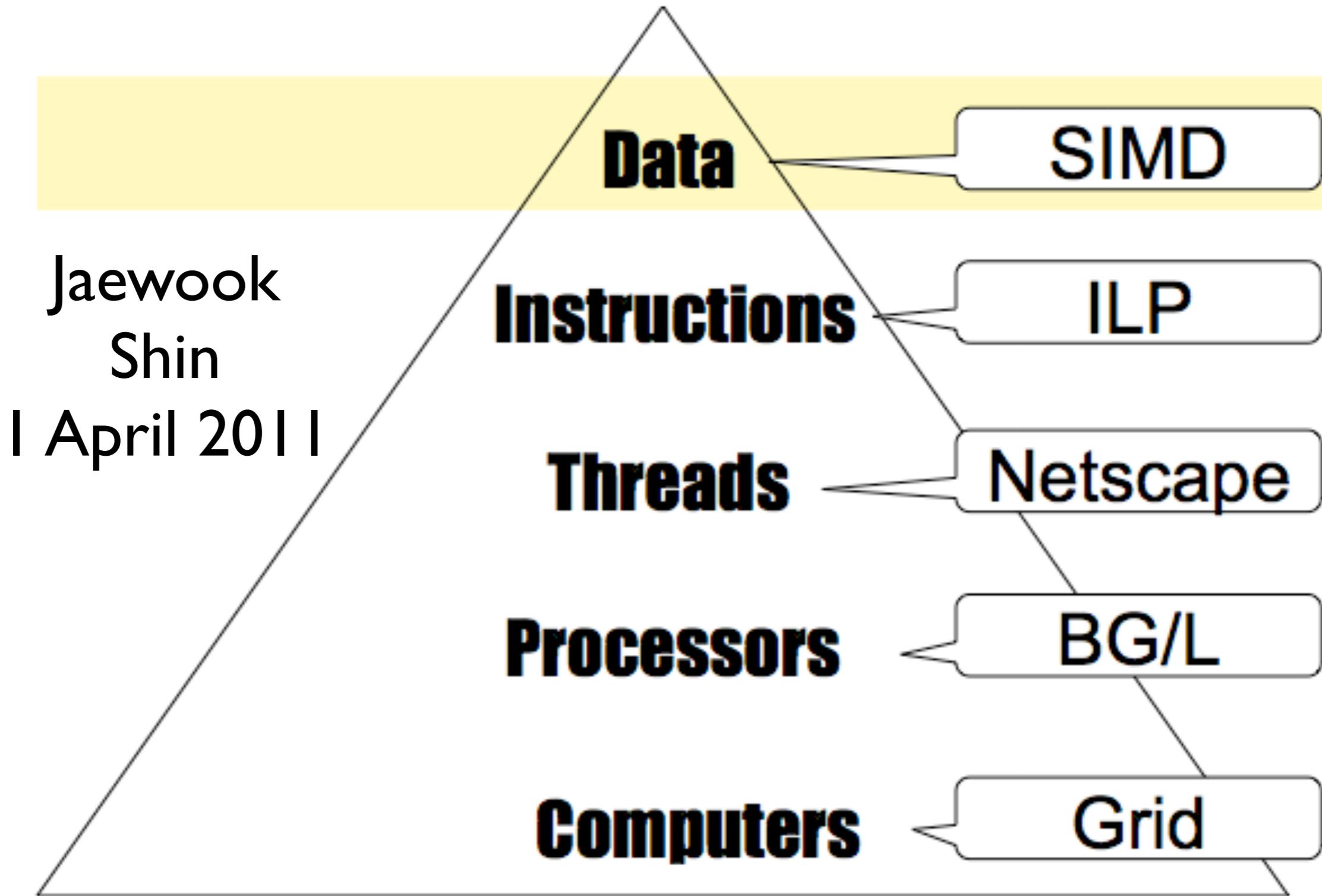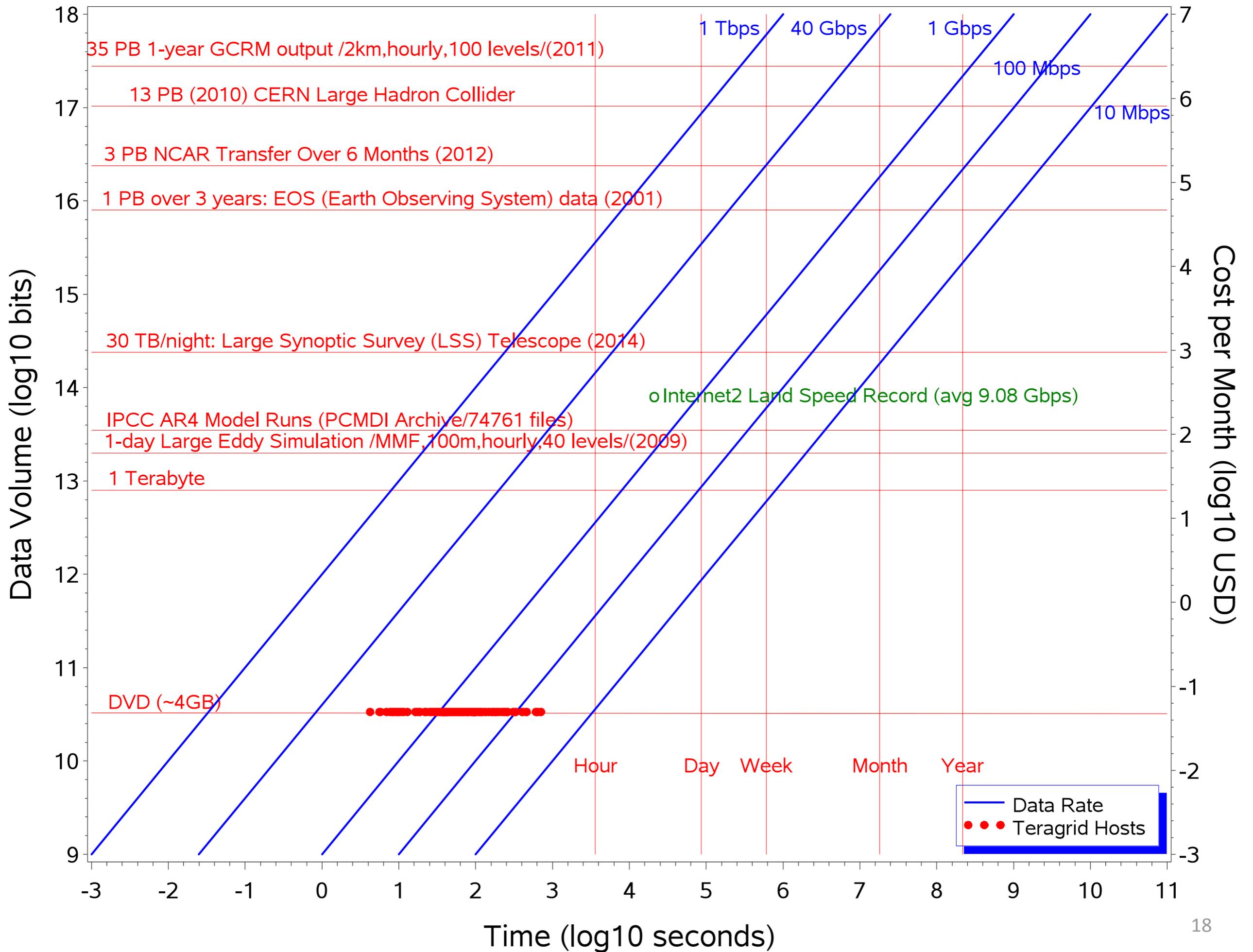
Community Support
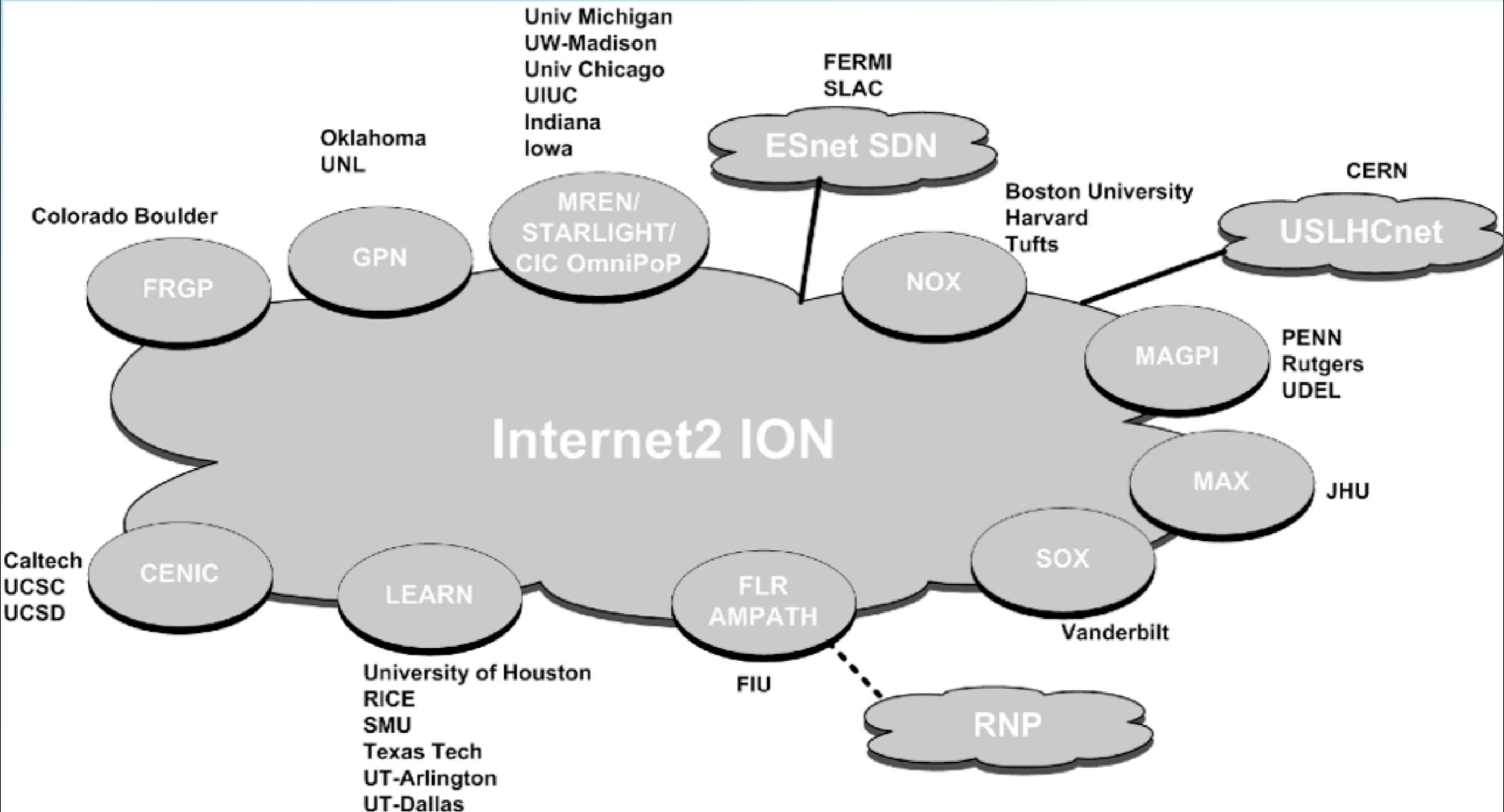# Data Interoperability

July 11th 2011 – Summer Joint Techs

Eric Boyd, Internet2 Deputy CTO

Jason Zurawski, Internet2 Research Liaison

NSF MRI-R2: DYnamic NEtwork System (DYNES, NSF #0958998)

DYNES Projected Topology (July 2011)

- Based on applications accepted
- Showing peerings to other Dynamic Circuit Networks (DCN)

What about Giga-LES/VVM?

- At UCB we talked about doing another Giga-LES run

- Given the lead time in getting computing allocation

  - Do we still want to do this (External Review)?

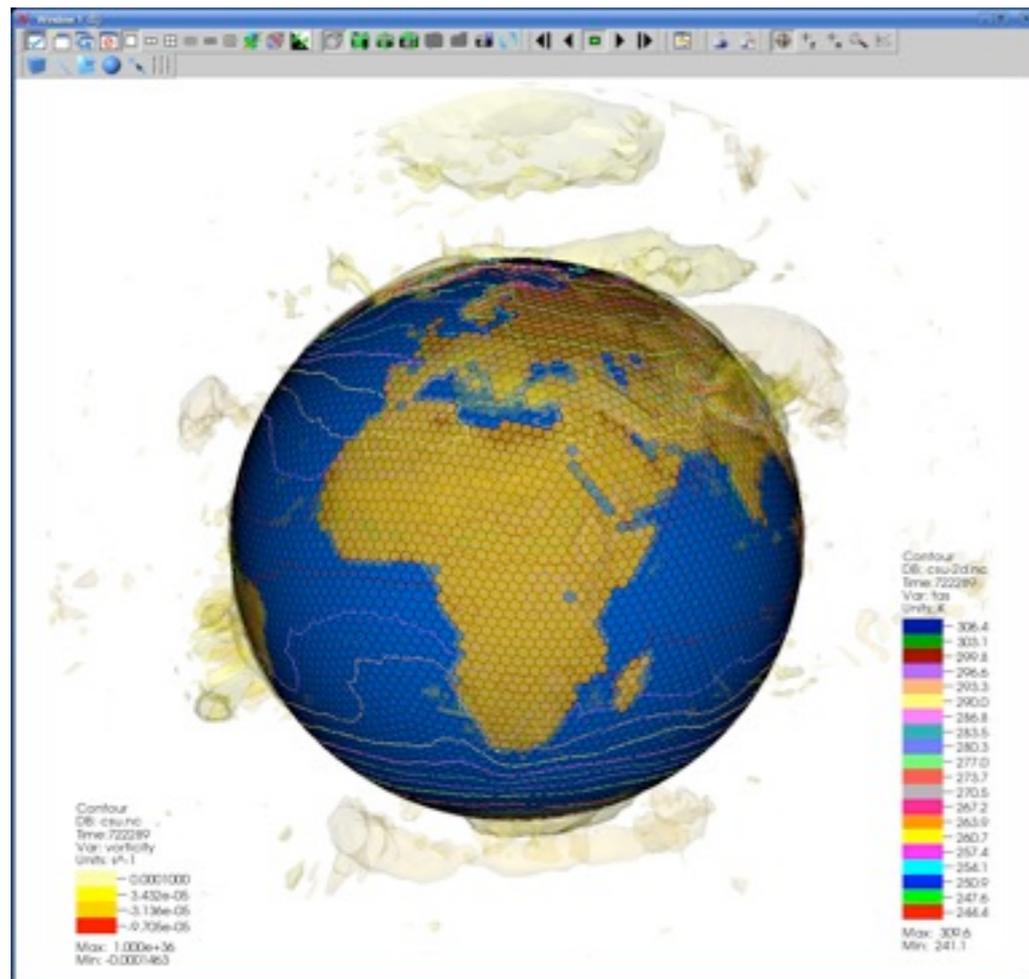 - Who's going to lead it?

  - What should be done?

Resources
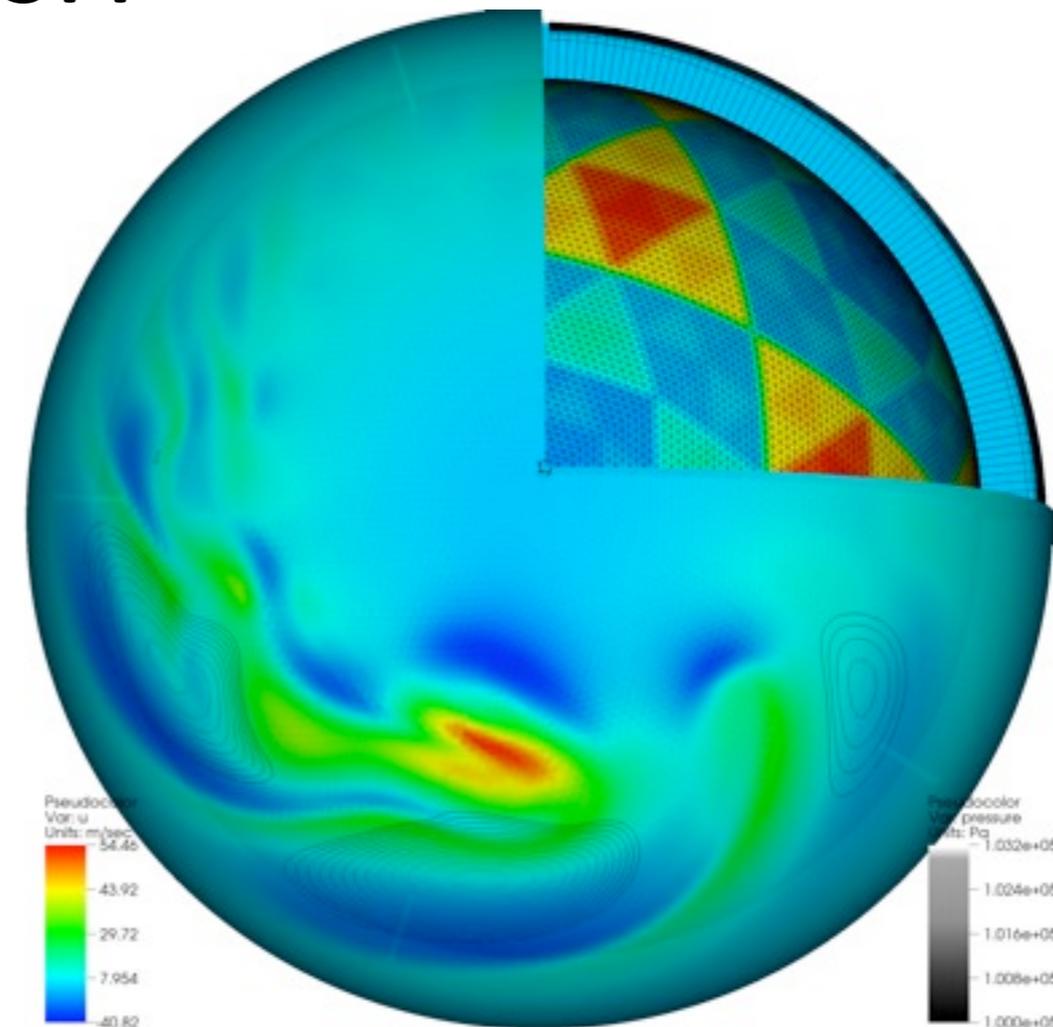# Visualization of Very Large Datasets

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)