UCAR Member’s Meeting

Cyberinfrastructure Breakout

Co-chairs: Jim Hansen (MIT) & Don Middleton (NCAR/SCD)

Notes provided by Markus Stobbs (NCAR/SCD) and Mohan Ramamurthy (UCAR/Unidata)
We Provided an Overview of Cyberinfrastructure and Discussed the “Atkins Report”

http://www.cise.nsf.gov/evnt/reports/atkins_annc_020303.htm

Thanks to NSF’ers Jarvis Moyers and Cliff Jacobs for additional insight!
What is cyberinfrastructure?

- Think of parallels to societal infrastructure (roads, communications, power) that serve as the digital foundations that enable science to progress.
  - People
  - Data/Libraries
  - Systems
  - Software
  - Hardware
Questions

• What are the CyberRoadblocks stopping you from doing the science you’d like to be doing?

• What organizational structure is needed to provide long-term support for data storage, access, model development, and services for a global clientele of researchers, educators, policy makers, and citizens?

• How can communication and coordination among computer scientists and environmental researchers and educators be enhanced to develop this innovative, powerful, and accessible infrastructure?

• To what extent are atmospheric science needs unique?
CyberRoadblocks: Collaboration technology

- Access Grid wonderful tool, great for committees, group meetings, but does not yet work well for impromptu scientist-to-scientist collaboration
- Need for ubiquitous, easy-to-use collaboration technology
- Personal collaboration technology that interoperates with systems like the AG could be useful
CyberRoadblocks: Need for local IT/programming support

• Rapidly growing need for local technical staff to install, support, and build the “on-ramps” to cyberinfrastructure
  – Idea: Fund local centers of support on each campus that will support a number of individual scientists and groups. There was some lack of confidence expressed relative to university administration in this regard.
CyberRoadblocks:
Seamless access to resources

- Security is increasingly complex and a growing problem for researchers
- Individual researchers and teams need to be able to use multiple computational and storage systems without fighting with firewalls and security systems
- Grid technology is aimed squarely at this problem
CyberRoadblocks: The “last mile”, in a general sense

- There is a great deal of interesting and potentially useful cyberinfrastructure
- But there has been a fairly consistent disconnect between the good ideas and the deployment of useable end-user environments
- Note that this is what the “cyberinfrastructure philosophy” is aimed at addressing
- And yes, the “last mile” to the scientist’s office is still an issue
CyberRoadblocks: Access to Scientific Data

- Data volumes, complexity, and diversity present a growing challenge
- We demonstrated and discussed one of NCAR’s efforts, the Community Data Portal
  - [http://dataportal.ucar.edu](http://dataportal.ucar.edu)
- Such environments will be of greatest benefit if the IT can be shared with universities
Organizational Issues & Ideas

- There is much commonality among disciplines such as oceanography, climate, etc. - this suggests common CI
- Lack of trust that institutions will provide useful technology even if well funded
- CI initiative should not “squeeze the balloon” of existing successful efforts (e.g. Unidata). Note that the balloon is expected to be larger.
What makes atmospheric sciences unique?

• **Integrative multidisciplinary nature, particularly in the context of climate change and societal/environmental studies**
• Realtime weather observation and forecasting
• Need to understand extremely diverse data (e.g. sociologists using climate data)
• Computational complexity
• Size, scope, and diversity of data - and a track record of sharing widely
Personnel issues

- Atmospheric scientists, educators, computer scientists, and environmental people must team to define and build useful CI
- There needs to be a better reward model for tenure-track researchers to contribute to the development of CI
- There is a growing need for cross-over people, who understand both the science and the technology very well (note Purdue Coalesce)
- Challenge of communicating the needs/desires of individuals/teams to the builders of the infrastructure
Carrying this conversation forward...

• This breakout group was an opportunity to heighten awareness of the CI thrust and initiate discussion on its future relative to atmospheric sciences

• CyRDAS is a forum where these initial conversations may be continued and leveraged
CyRDAS focus group dates

- Mountain region, NCAR, Oct 10
- Midwest region, NCSA, Oct 13
- Northeast region, Access Center (DC), Oct 15
- Southwest region, SDSC, Oct 21
- Southeast region, Georgia Tech, Oct 28
- Northwest region, Univ. Washington, Oct 30

Details and registration at
www.cyrdas.org
Get Involved!

http://www.cyrdas.org
End
Cyberinfrastructure Drivers from the UCAR perspective

- **Atmospheric science** (observations and model data, observational systems, model complexity, DA, collaboration, etc.)

- **Atmospheric science education** (collaboration tools for education, real time data for the classroom, digital libraries, distance learning, etc.)

- **Computer/networking hardware** (big computers, big storage systems, big pipes, wireless, etc.)

- **Computer software** (professional code, distributed code, distributed data, visualization, etc.)
CyRDAS – CI Planning for the Atmos Sci Community

- Seeks broadest possible representation of the atmospheric sciences community
- The beginning of a process for integration of CI planning for the atmospheric sciences into the larger planning processes for geosciences and environmental sciences
- Seeks broadest possible dissemination of findings and recommendations
Many reports on supercomputing, networking etc. over last two years have suggested that scientific investigation has added a new methodology to observation, experimentation and theory: numerical investigation.

Most recently, the Atkins report recommended an additional $1B/year support of cyberinfrastructure (CI) to support numerical investigation.

Several initiatives are getting underway, including geosciences CI and environmental CI activities.
Cyberinfrastructure Enablers
CS, IT and Math

• Research into data storage, data distribution, data transport, and knowledge discovery, e.g., data mining, pattern recognition, etc.
• Dynamical systems theory, stochastic models, probabilistic treatment, discrete systems and more sophisticated (probabilistic and statistical) analysis strategies
• Development of model-data fusion methods
• Advanced visualization and virtual reality methodologies
• Faster, cheaper computers and networking
• The Grid
CI Goals for the Atmospheric Sciences

- Using cyberinfrastructure to lead to more rapid and more substantial progress in research and more efficient and effective education.
- Removing cyberinfrastructure barriers that are impeding scientific progress.
- Identifying the central issues that atmospheric scientists, educators and technologists consider most important, and helping them address those issues.