

HPC in the Geosciences

Atmospheric Sciences

Breakout Group

Morning session

26 September 2006

Goal of the Workshop

The goal of the workshop is to provide the Geosciences directorate with a compelling, community-supported plan for implementing the “Petascale Collaboratory for the Geosciences” vision. Key questions to be addressed during the workshop include

Workshop Key Questions

- What is the best balance between capability and capacity for the Geosciences?
- What is the “sweet spot” for capability computing in the Geosciences?
- What research opportunities are being missed in the current computing climate?
- How should a high capability node be designed and optimized to support the Geosciences community?
- What governing and allocation model(s) make sense for the full Geosciences community?

Workshop Key Questions

- What is the best balance between capability and capacity for the ***Atmospheric Sciences***?
- What is the “sweet spot” for capability computing in the ***Atmospheric Sciences***?
- What ***Atmospheric Science*** research opportunities are being missed in the current computing climate?
- How should a high capability node be designed and optimized to support the ***Atmospheric Sciences*** community?
- What governing and allocation model(s) make sense for the full ***Atmospheric Sciences*** community?

Breakout Questions

1. How can a Geosciences HPC enterprise best be structured to meet the needs of your discipline?
2. What will be required to make such an enterprise function effectively?
3. What are the characteristics of a computing resource allocation model that will best meet the needs of your discipline?
4. What types of service and support will your discipline require and what are the types of applications that will need to be supported?

Breakout Questions

1. How can a Geosciences HPC enterprise best be structured to meet the needs of your discipline?

The 2005 Petascale Geoscience Collaboratory Science Frontiers report clearly shows that geoscience researchers are confronted by many grand challenge problems. Due to their interdisciplinary nature, the scale of the resources required, and/or their global importance, these problems need the focus, persistence, intellectual critical mass, and clout a national center can provide in order for measurable progress to be made. At the same time, smaller, distributed research enterprises offer distinct advantages in many situations in that they can be more specialized, nimble, and better able to meet focused research needs. The rapidly maturing state of scientific cyberinfrastructure now makes it possible to blend both the larger and smaller scale approaches via a distributed network of science and technology centers linked to a capability facility with significant scalable resources.

- **How could such a coordinated geoscience enterprise – composed of a spectrum of facilities that are linked by grid technology – be structured to be the most flexible and effective solution for addressing the needs of atmospheric sciences?**

Breakout Questions

2. What will be required to make such an enterprise function effectively?

In the “Establishing a Petascale Collaboratory for the Geosciences” reports (see http://www.joss.ucar.edu/joss_psg/meetings/petascale/), members of the community have outlined that a Geosciences Collaboratory is defined as a research environment that provides not only access to high-performance computers but also essential support services such as data and information management, visualization, and user interfaces and applications appropriate for addressing petascale computational problems. Within this framework, the efforts of scientists, facility providers, developers and providers of data, services, and applications, educators, computer and computational scientists, and scientific users from across all geosciences disciplines must be effectively managed and made available throughout the collaboratory.

- **From the perspective of atmospheric sciences, what are the key attributes of the collaboratory infrastructure systems (e.g., grid, or network, requirements, center capabilities, etc.) that will be required?**
- **What type of enterprise governance and oversight structures are appropriate and necessary for ensuring timely access to and effective usage of all collaboratory resources?**

Breakout Questions

3. What are the characteristics of a computing resource allocation model that will best meet the needs of your discipline?

Computing resources may include: access to computers, allocations of computer time, access to dedicated time and/or computer platforms, access to previously computed models and datasets, post-processing, storage (on-line, near-line, off-line), networks, and so on. A variety of allocation schemes may be appropriate for your discipline and may even vary depending on the resource.

- **At what frequency would allocations for major resources be made?**
- **Does a peer-review, panel-review, sponsor-review process make sense for atmospheric sciences?**
- **Can this type of review facilitate more experimental research?**
- **If community models are desirable, what would the most effective process for deciding models, parameter settings, physics?**
- **What would you like to happen if you exceed your allocation/ resources in the midst of project, or at a critical time?**
- **Are you willing to share an allocation with other atmospheric scientists?**
- **Does a per project, or per overall group allocation make better sense, or is another form of allocation more appropriate?**
- **What is the relative balance between needs for computer time, memory, storage, network bandwidth, etc.?**

Breakout Questions

4. What types of service and support will your discipline require and what are the types of applications that will need to be supported?

The types of service and support required can take many forms. Participants should take a holistic approach to defining the service and support needs of the entire scientific tool chain of their domain in their deliberations. In particular, both the data and modeling side of the problem should be considered. For models, both the forward and inverse modeling aspects should be discussed.

- **What additional support for existing or emerging community models is needed?**
- **To what extent is support for improved data portals, digital libraries, and other data- related services required?**
- **To what extent would access to application support be advantageous? General application support can range from simple porting and tuning assistance to very intrusive activities such as application parallelization or scalability improvements that might involve complete redesign. What types of domain-specific knowledge are required to support your application space?**
- **Would improved access to test platforms (e.g. different software stacks) be helpful?**
- **Would the availability of platforms enabling scalability tests be helpful?**