

# Science Drivers

# Climate

- Single model runs will be O(100TB)
- Data management requirements are critical
- IPCC and other research will involve enormous distributed data intercomparison
- Climate research is a global, multidisciplinary effort and requires substantial collaboration capability
- Pressing need for coupled analysis & visualization tools that support complex data fusion across simulated, observed, □ □ □ and GIS with a rapidly growing set of complex grid topologies
- Growing desire to explore coupled models in 3D as resolution grows and regional climate studies progress.

# Fusion Simulation

- Single model runs will be  $O(10-100\text{TB})$
- 3D visualization of time-varying data required
- Data management requirements are critical
- Effective handling of complex grid topologies is critical
- Widely-varying spatiotemporal scales
- Visualization of higher-dimensional phase space particle distribution functions

# Fusion Experimentation

- Single experiments □□□□ will produce  $O(10K)$  signals) with a total dataset size  $O(10GB)$
- Low requirements for 3D visualization
- Data management requirements are critical
- World wide collaborative remote visualization
- Severe real-time requirement
- Disparate display types ranging from walls to desktops
- Need for a Python-like scripting language to create end-user workflow environments

# Nano, materials

- Single simulation dataset size O(100s of GB)
- Collaborative, multisite science projects
- Data management requirements are critical
- Close coupling between modeling, simulation and experiment

# Combustion

- Single simulation dataset size O(10s-100s of TB)
- Ability to visualize and understand flame processes at many scales
- Capabilities to compare 3D simulations and experimental measurements
- Explore and gain insight from multiple terabyte simulation runs in a coupled analysis and visualization environment
- Industry CRADAs require secure environments
- Effective handling of pedigree and metadata for simulations and their underlying reaction data from many locations

# Astrophysics

- Single simulation dataset size  $O(10\text{s}-100\text{s of TB})$ , vast spatial domains
- Strong need for coupled visualization and statistical analysis capabilities
- Need collaborative capabilities for multiple sites with very different visualization capabilities (walls vs. desktop)
- Huge data management issues
- Severe multiscale problem with length scales ranging from centimeters to parsecs
- Need to visualization radiation datasets in 7D phase space

# Genomics/biology

- Genomics Data O(small), Computational biology O(10-100TB)
- Pushing the frontiers of information visualization
- Very strong requirements for the elements of collaboration: data and metadata interchange formats, collaboration, and remote capabilities
- Very strong needs for data management along with complex metadata and data provenance issues

# Subsurface, seismic

- Simulation datasets in the 10s of TB range
- Complex geometries and grids
- Large-scale deployment of sensor networks
- Multiscale from molecular to kilometers
- Visualization may help with uncertainty quantification and parameter studies
- Increasing geographic collaboration and collaborative visualization
- Data management requirements are strong, and similar to many field-oriented projects.
- Metadata?