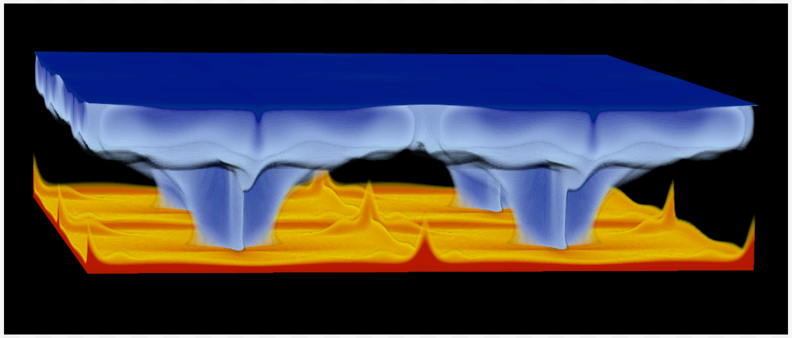
Real-Time Interactive Visualization of Three-Dimensional Mantle Convection

Dave Yuen, Megan Damon, E.O.D. Sevre
Minnesota Supercomputing Institute
University of Minnesota, Minneapolis, Minnesota

Mike Knox and David Henry Porter
Laboratory of Computational Science and Engineering
University of Minnesota, Minneapolis, Minnesota

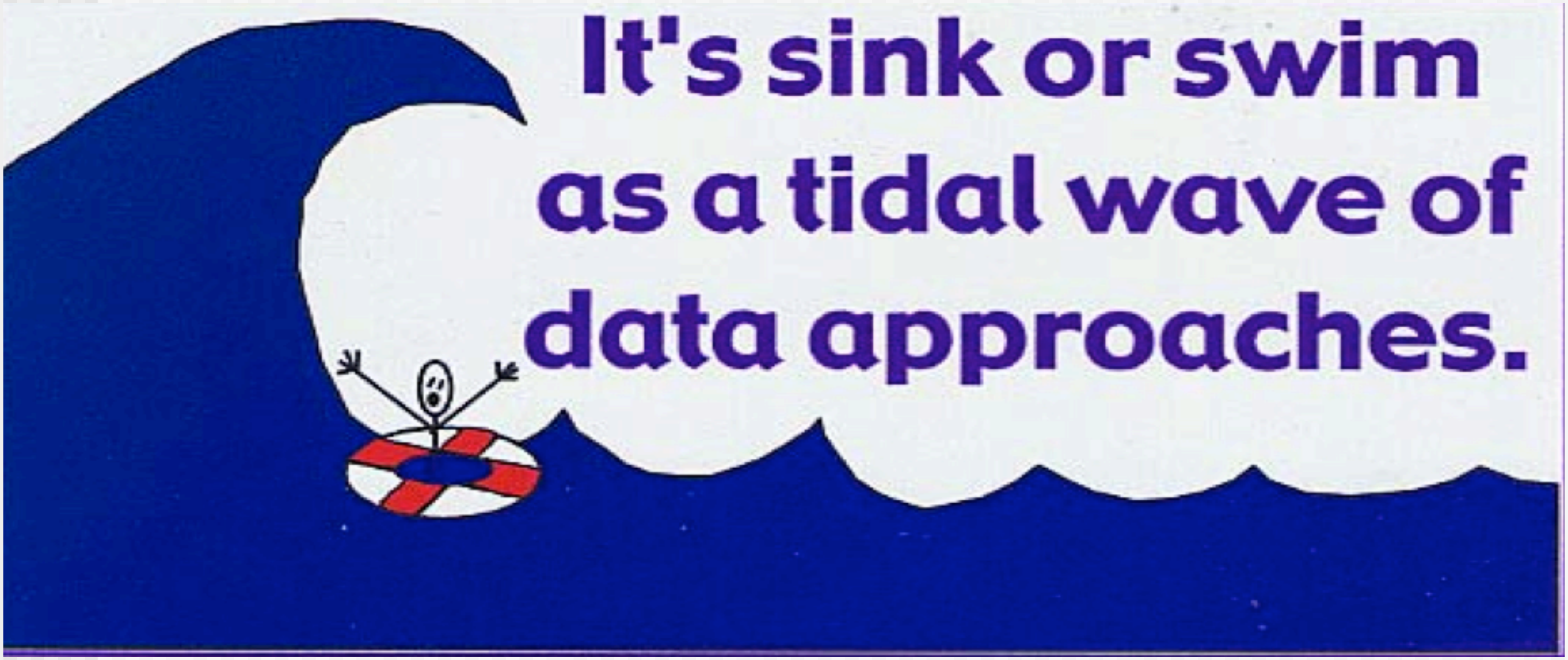
M. Charley Kameyama
Earth Simulator Center, JAMSTEC, Yokohama, Japan



Outline

1. Introduction: raison d’être, onslaught of petascale computing, the “data tsunami”
2. LCSE visualization system
3. 3-D numerical model ACuTEMan, by Charley Kameyama
4. Demands of interactive visualization
5. Examples
6. Emerging paradigm in numerical modeling and visualization

Petascale Computing

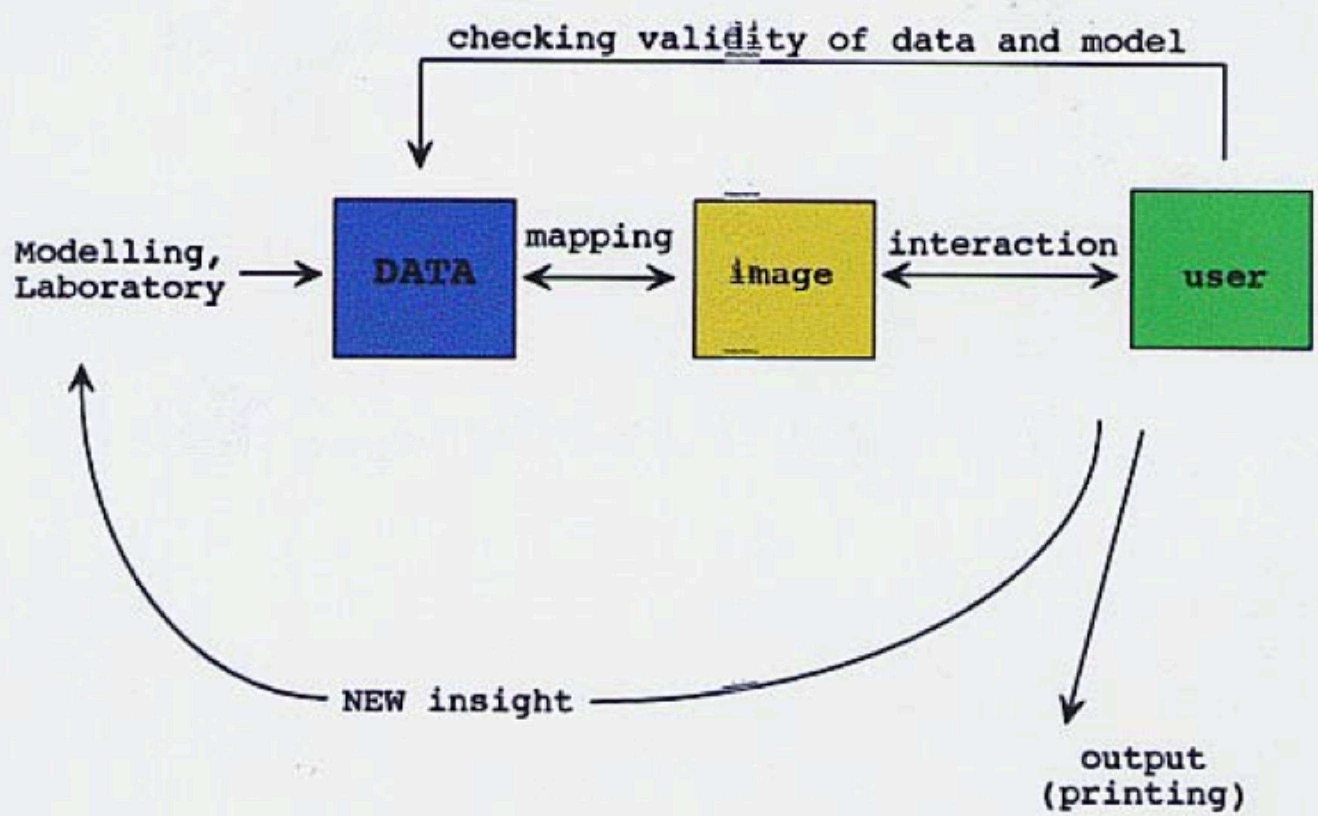


The Devil Is in the Details

Large 3-D Data Sets

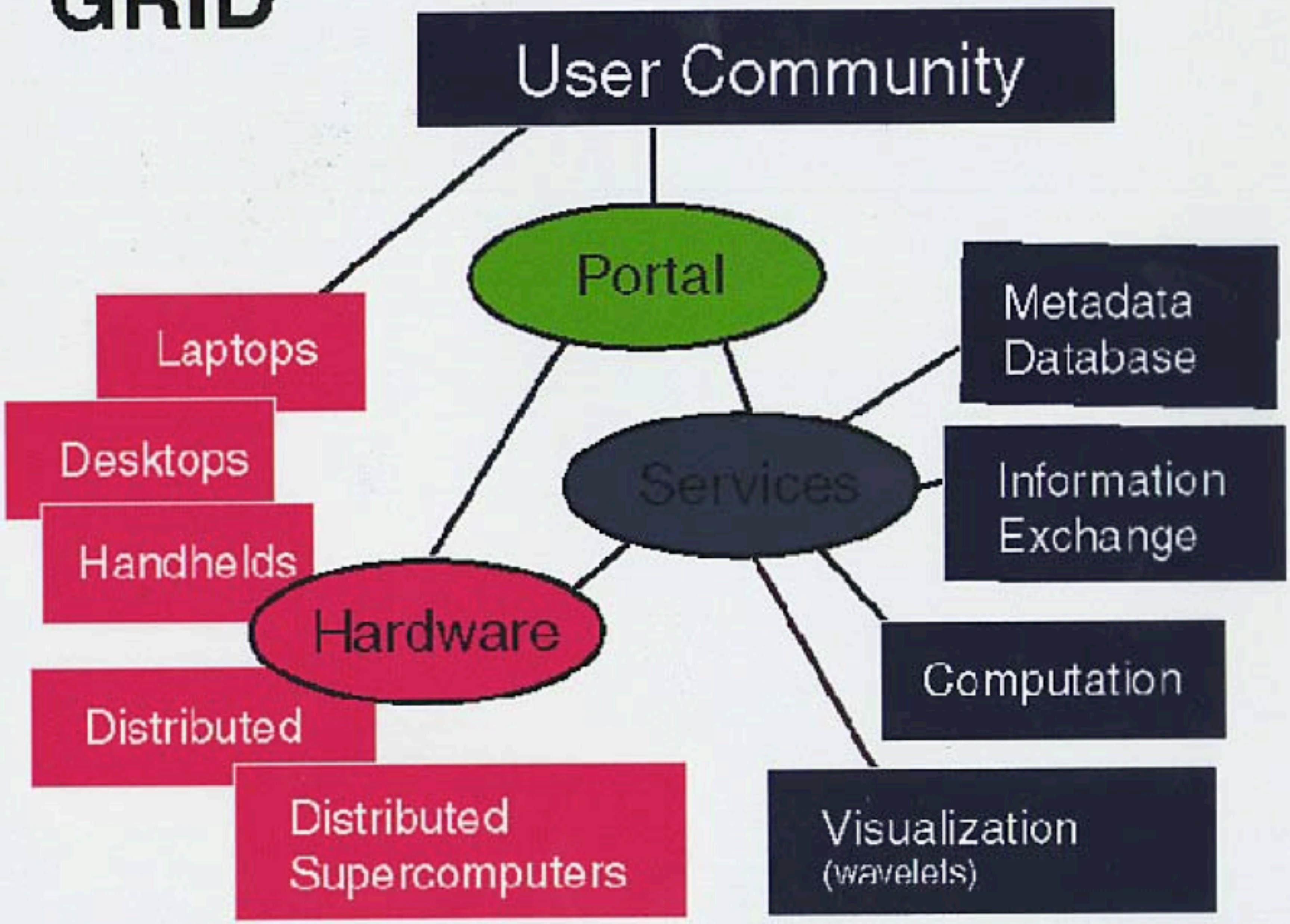
- Can 3-D time dependent data sets be visualized on the fly?
- Seismology advancements: **2000x2000x2000** points from 500x500x1000 points!
- Wave patterns must be extracted from seismic calculations.
- This is more manageable using curvelets (De Hoop) for decomposition.

Challenges of Large Data Sets



Can we do this interactively and collaboratively?

GRID



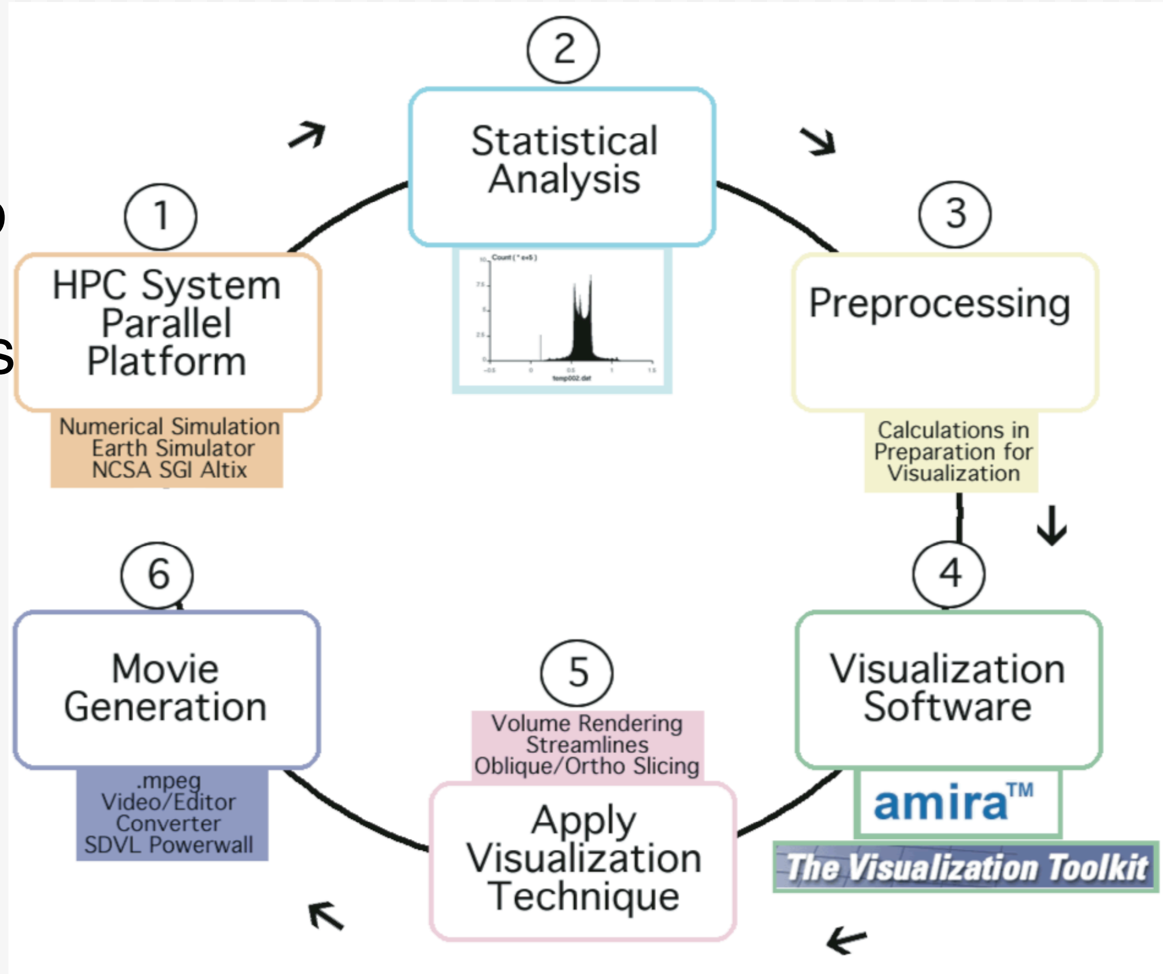
Fox et al. Grid Services for Earthquake Science, ACES 2001: Special issue of *Concurrency & Practice.*

Why Visualization?

- Numerical simulations and field experiments produce extremely large datasets
- Petascale computing: in 2002 it was a dream, but by 2012 it is a reality, like Global Warming
- The size of these datasets are increasing exponentially fast, with resolution
- Numerical output (e.g. tables) does not lend itself to easy comprehension
- We need new dynamical display of fields for unraveling new physics

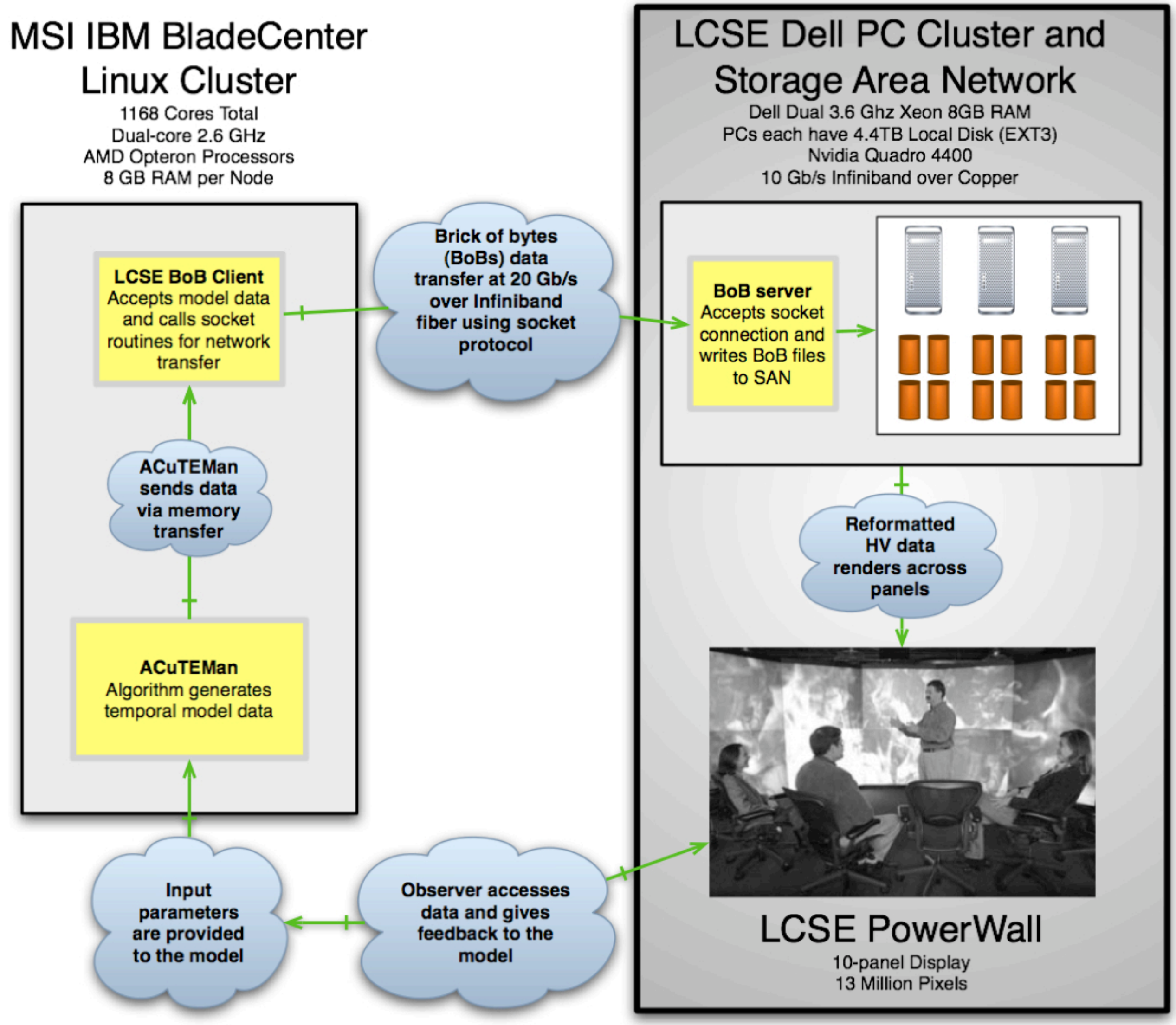
Post-Processing Visualization

- Start with raw data files from HPC system
- Results are analyzed to determine the best visualization techniques
- Human spends a lot of time learning and using software packages
- **There is a need to automate these steps for a much faster turn around**

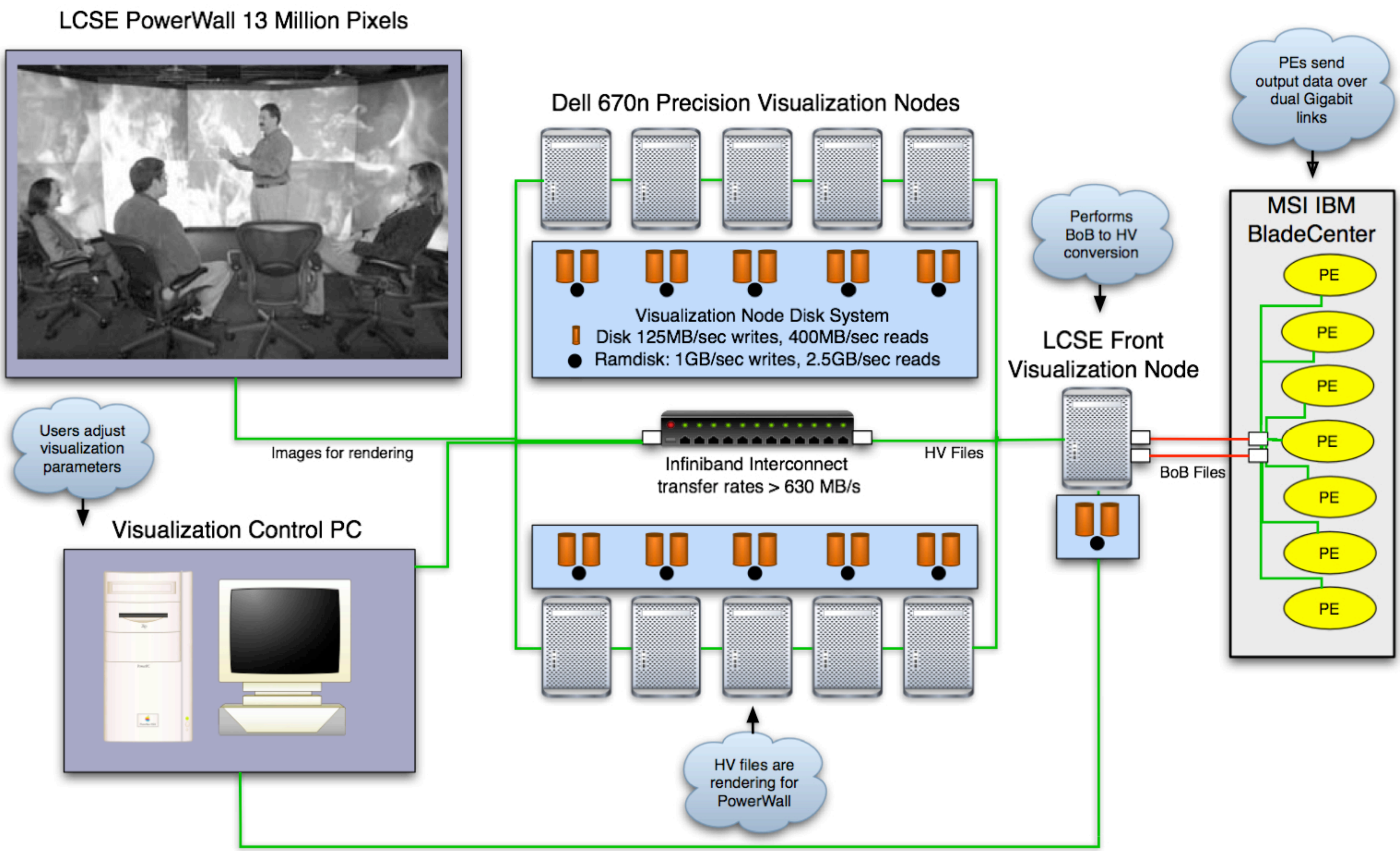


LCSE Visualization System

- Raw data from the BladeCenter is transferred over Infiniband using a client-server application
- 3-D results are rendered as a volume at LCSE
- Researchers use DSCVR software to interact with data
- Data can be displayed on high resolution Powerwall



Closer look at LCSE system

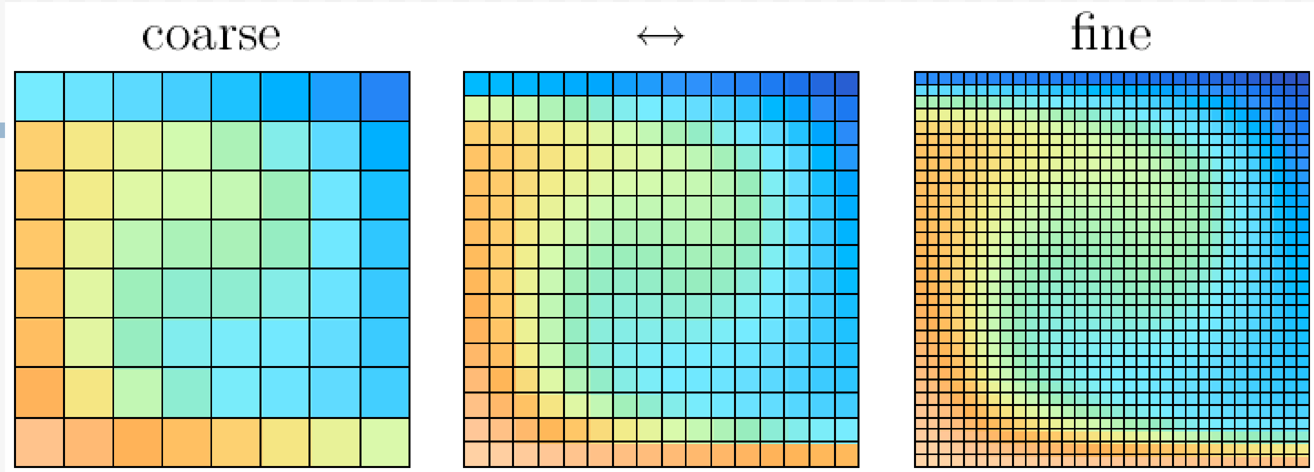


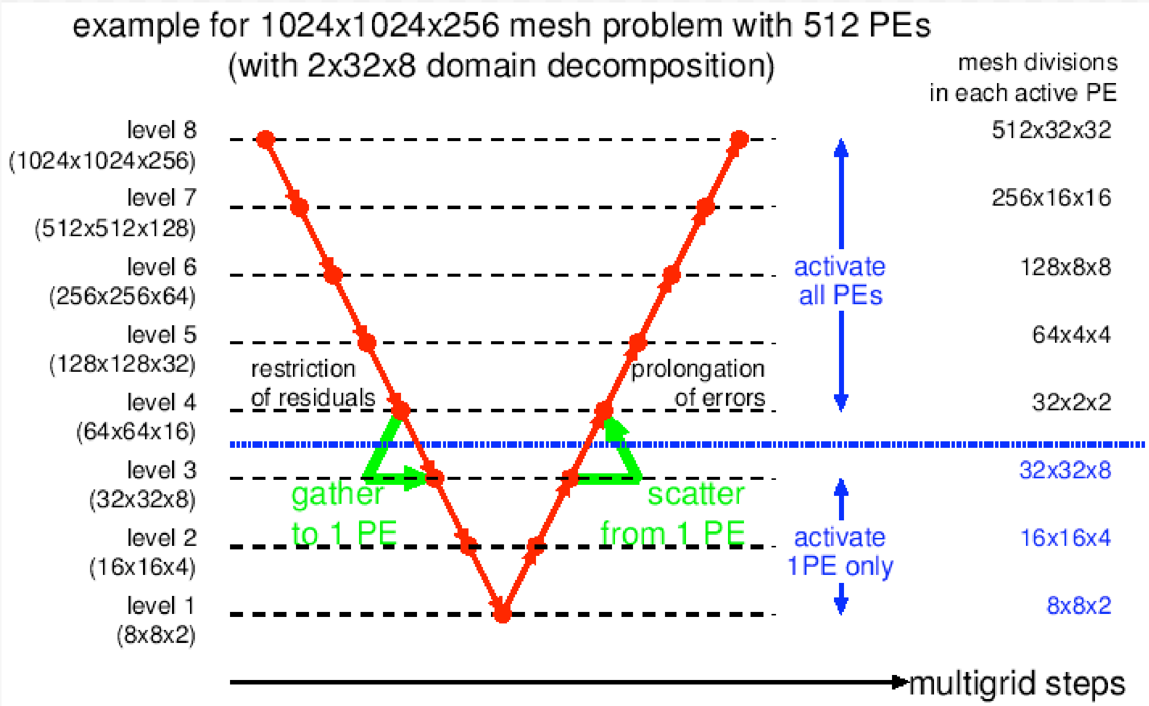
ACuTEMan

- Written by Charley Kameyama, who is moving to Ehime University in September 2007
- Parallel, multigrid cartesian, finite volume code, written for vector parallel computing using MPI-2
- Benchmarked at 3.4 Teraflops a Supercomputing 2005 in Pittsburgh
- Generalized to Yin-Yang grid for 3-D spherical geometry and recently executed on the Earth Simulator. Variable viscosity contrast up to ten thousand across adjacent grid points

*Kameyama, Kageyama, Sato, J. Computational Physics, 2005 Kameyama, J. Earth Simulator, 2006.

Multigrid Method





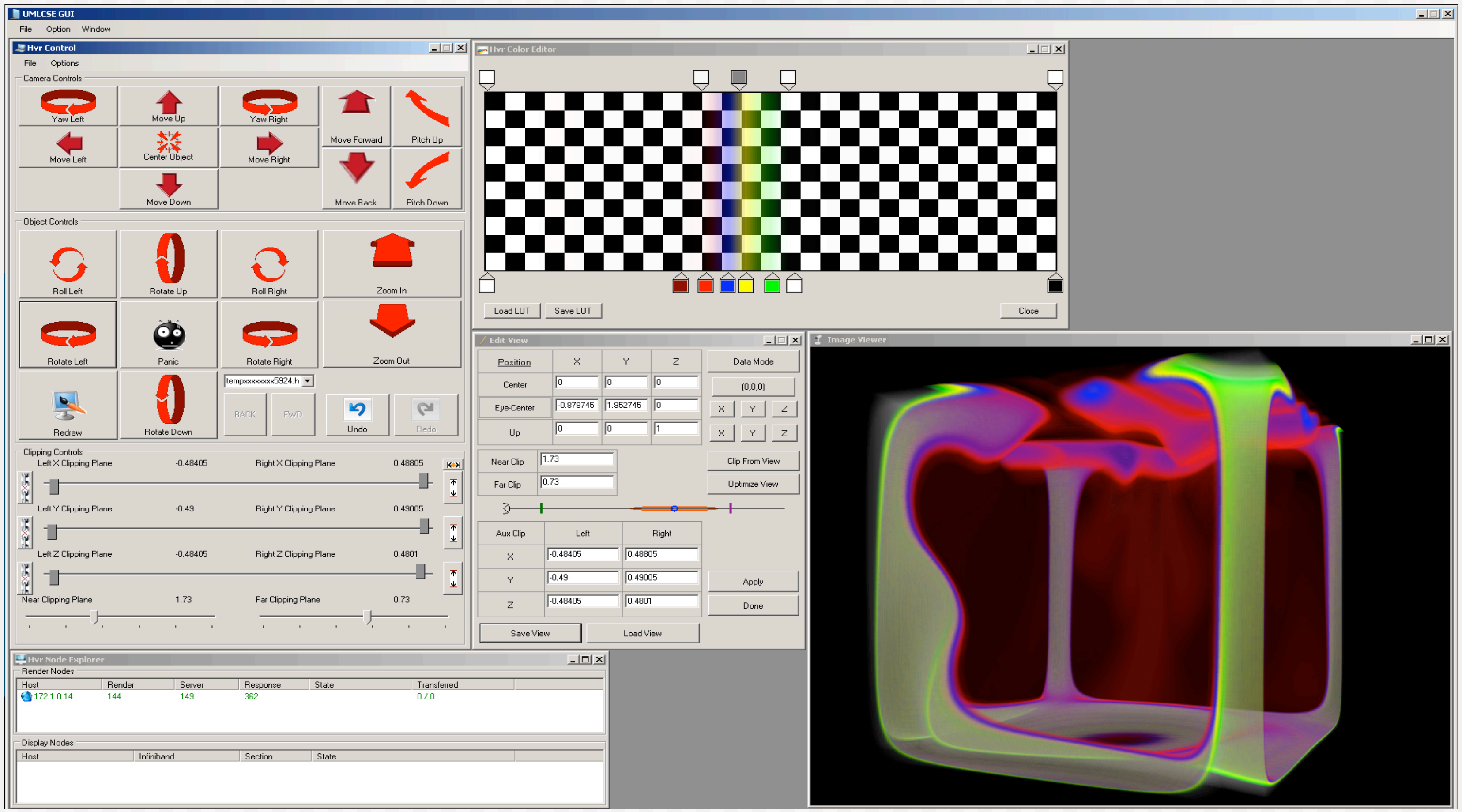
Demands of Petascale Computing

- 5 hours on 3-D grid with 2000x2000x2000 grid points
- 5 grid fields x 4 bytes storage every 5 minutes (or thousand time steps) on a linear scale with 5,000 cores
- 2000x2000x2000x5x4x60 = 9.6 TB
- 40 runs -> **384 TB of raw data** - AND we are only storing every 1,000 time steps (stingy)!
- Solutions:
  - Stream raw output data to visualization system in real time to generate movies - data compression
  - Interactive visualization with real-time simulation - avoid saving most of the data

Need for Software & Hardware Solutions

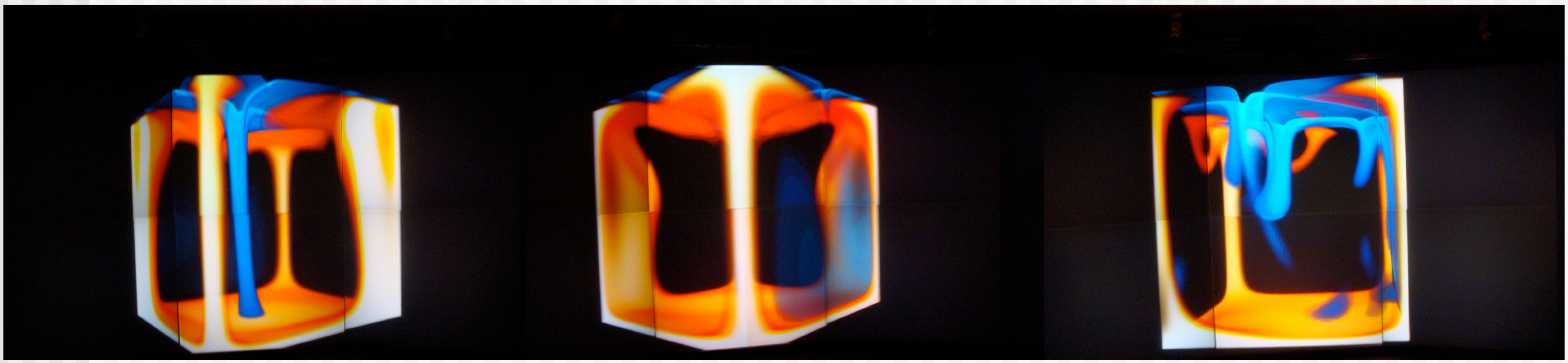
1. Collaborative mode of visualization whereby we can communicate readily upon viewing a common image
2. Storage capabilities at a central site on demand
3. Web-portal services for controlling our desires

Results - Real-time Interactive Visualization with LCSE DSCVR



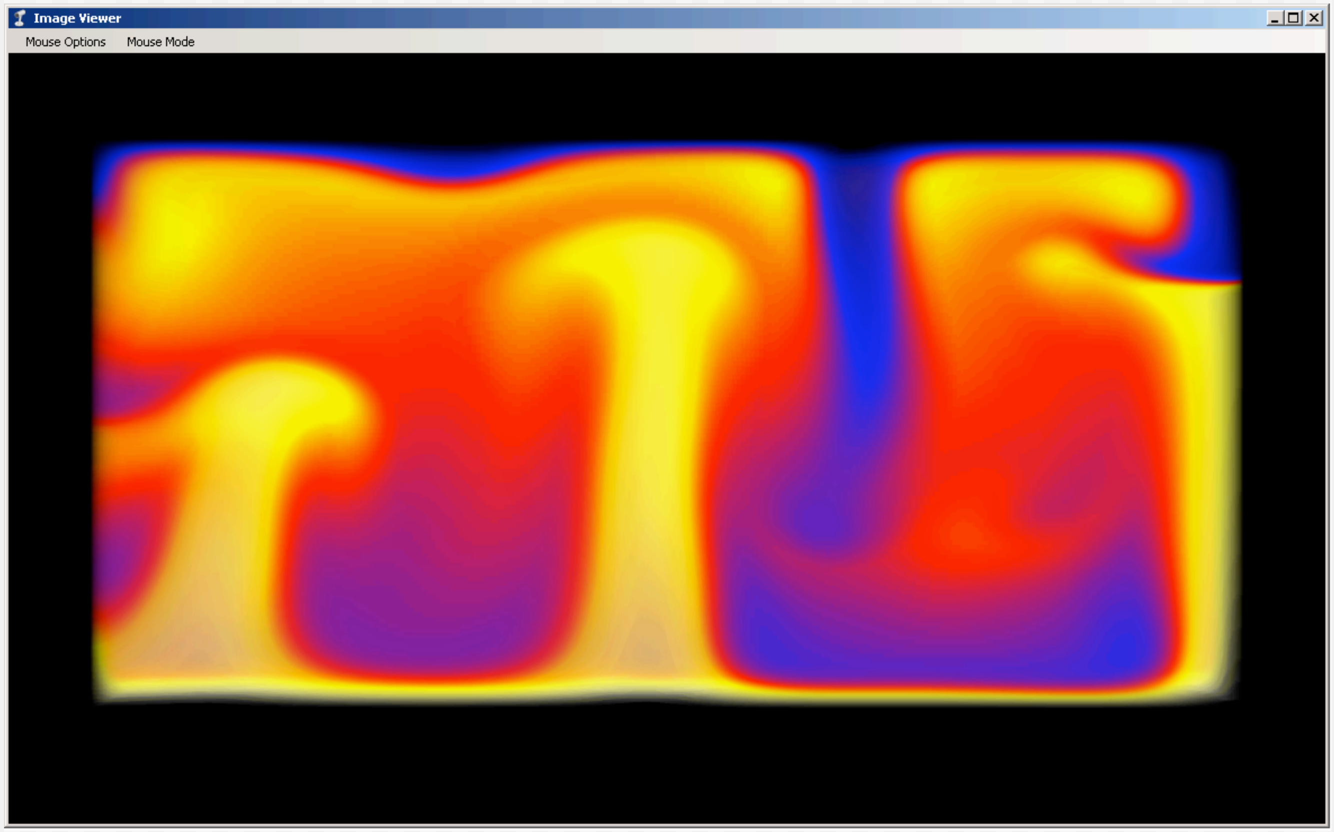
128x128x128 grid points rendering in real time across 128 PUs

Results - Real-time Interactive Visualization with LCSE Powerwall



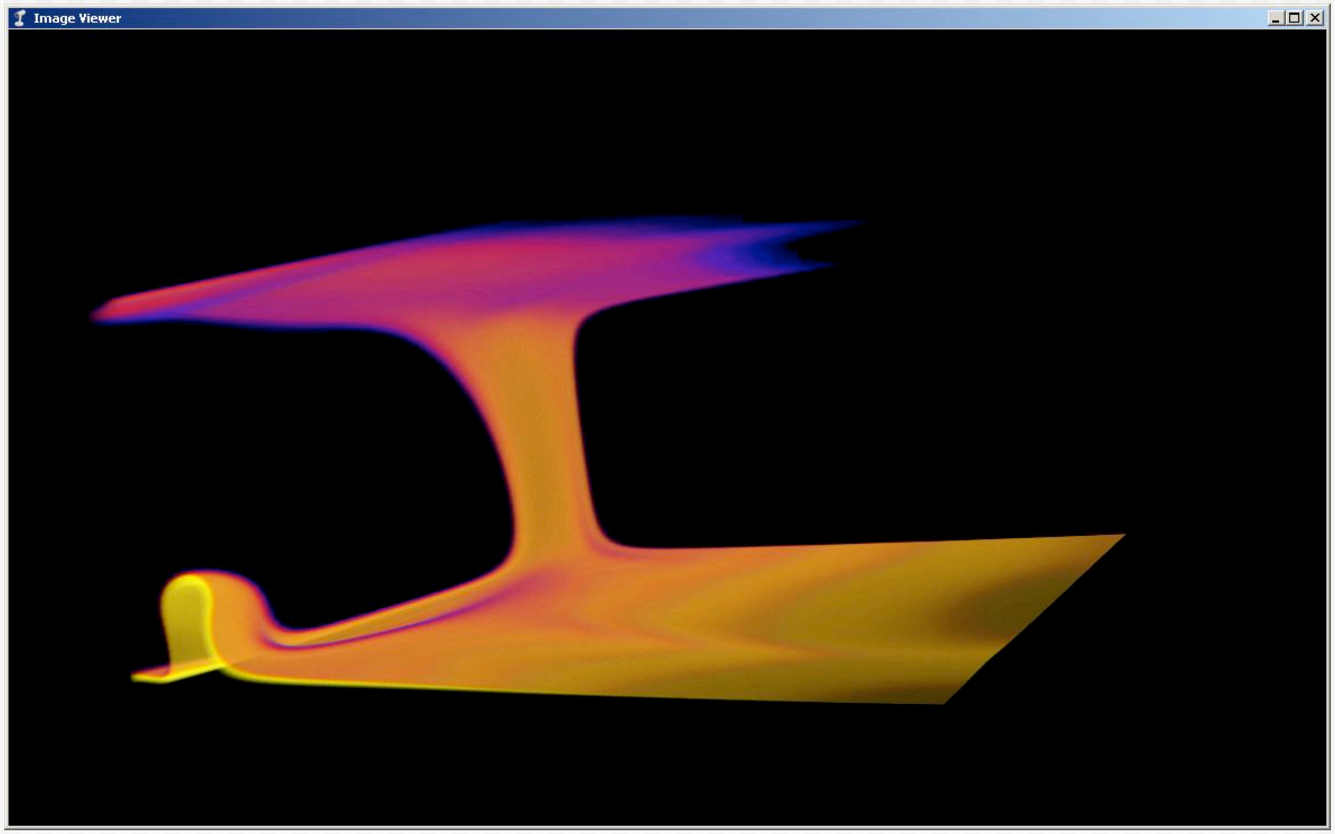
128x128x128 grid points rendered in real time across 128 PUs
Displayed on 13 megapixel Powerwall
Rayleigh number = 10^6

Results - Larger Grid, More Processors, Higher Rayleigh Number



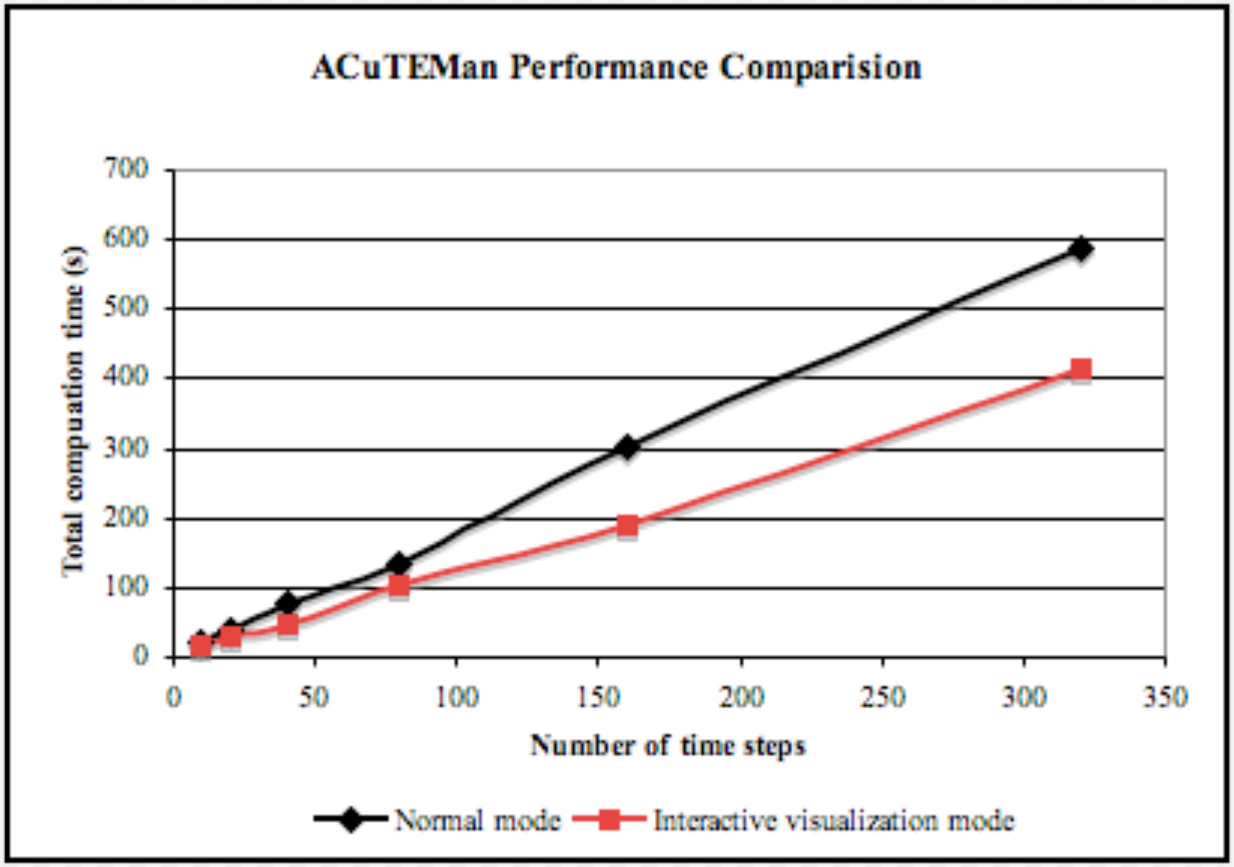
128x256x256 grid points rendered in real time across 256 PUs
Time step 2,700
Rayleigh number = $4x10^6$

16,000 time steps later…



128x256x256 grid points rendering in real time across 256 PUs
Time step 19,000
Rayleigh number = $4x10^6$

Results - Performance gains



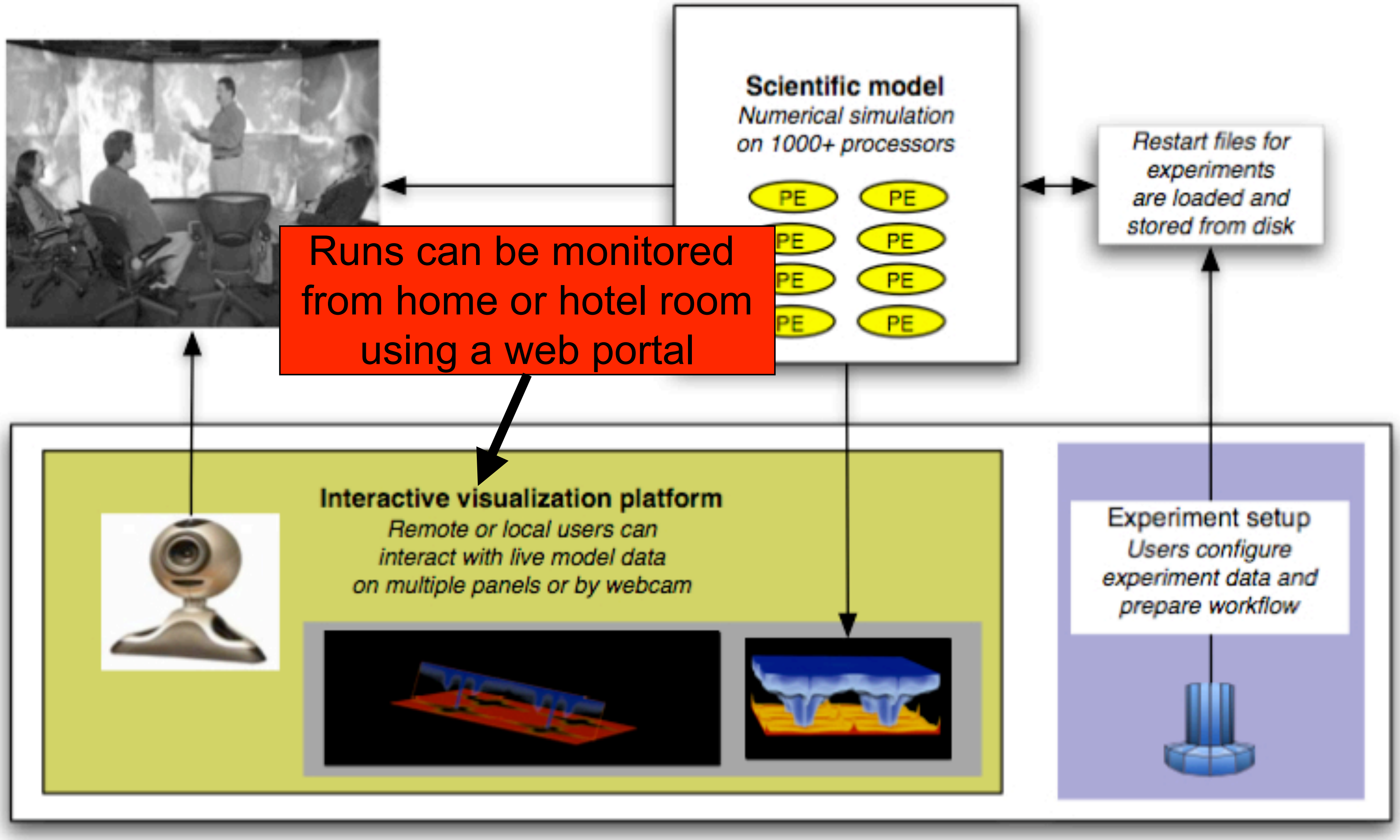
Total compute time reduced on average by 30 %
Local disk space on Supercomputer also reduced

Movie of Interactive Visualization

Higher quality: http://www.msi.umn.edu/~esevre/lcse/interactive/megan-powerwall-v1.0iPod.m4v

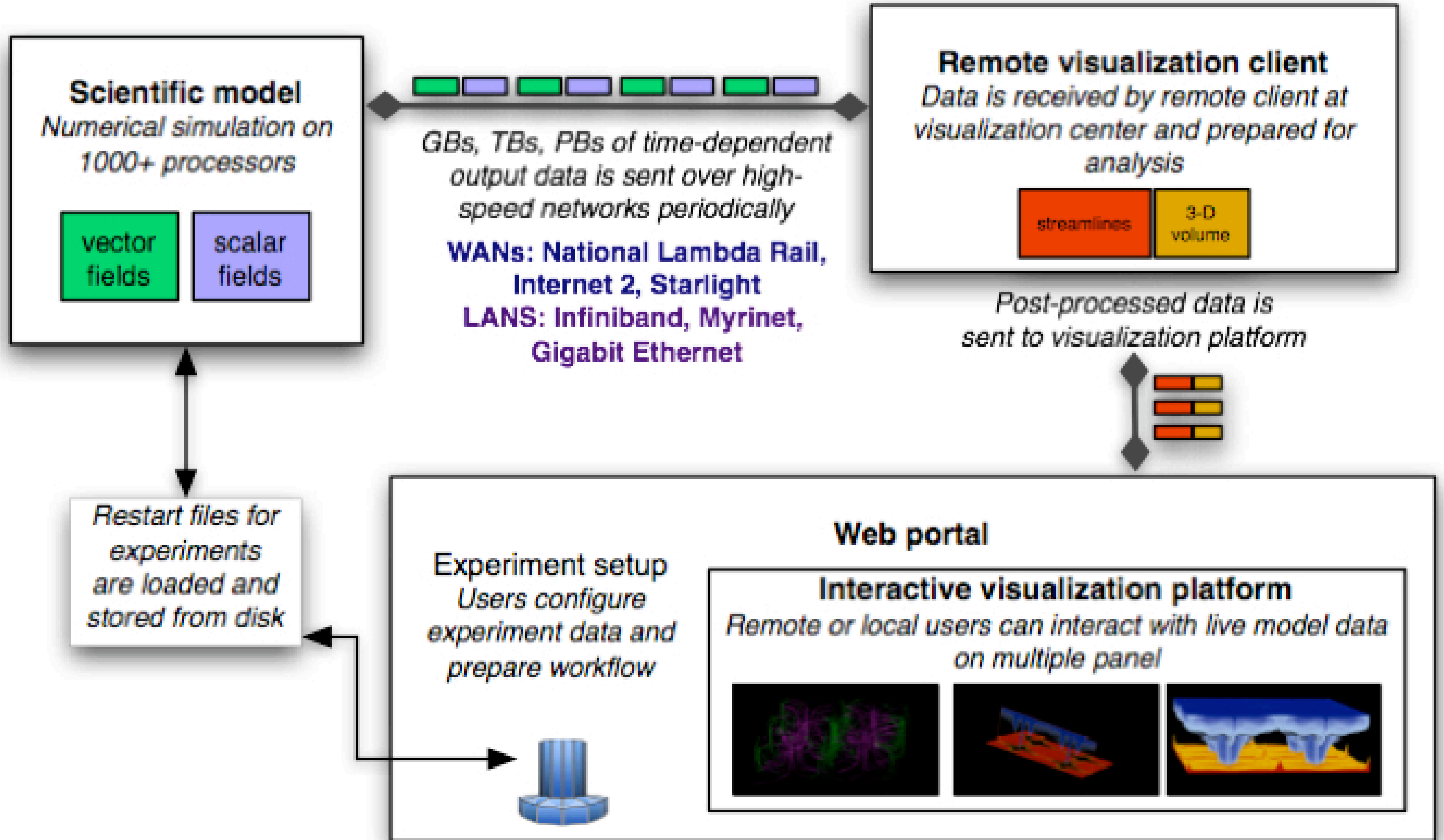
Interactive Computing

- Limited time duration, perhaps 1 or 2 hours
- Balance between grid resolution and number of processors
- Few frames to one frame per second
- Runs can be monitored from remote locations


Scientific model
Numerical simulation
on 1000+ processors
PE
PE
PE
PE
PE
PE
PE
PE
Restart files for
experiments
are loaded and
stored from disk
Runs can be monitored
from home or hotel room
using a web portal
Interactive visualization platform
Remote or local users can
interact with live model data
on multiple panels or by webcam
Experiment setup
Users configure
experiment data and
prepare workflow

New Paradigm in Numerical Modeling

- Disks are a well-known bottlenecks that slow research
- We need real-time computing combined with interactive visualization for fast feedback and results
- We can leverage high speed networks instead of disks to reduce storage requirements
- Web services for remote collaborations
- Streamlined visualization techniques
- This will drastically reduce storage requirements and save compute time
- We can run a model to analyze results instead of pulling results from an archive
- Short interactive periods


Scientific model
Numerical simulation on 1000+ processors
vector fields
scalar fields
GBs, TBs, PBs of time-dependent output data is sent over high-speed networks periodically
WANs: National Lambda Rail, Internet 2, Starlight
LANS: Infiniband, Myrinet, Gigabit Ethernet
Remote visualization client
Data is received by remote client at visualization center and prepared for analysis
streamlines
3-D volume
Post-processed data is sent to visualization platform
Restart files for experiments are loaded and stored from disk
Web portal
Experiment setup
Users configure experiment data and prepare workflow
Interactive visualization platform
Remote or local users can interact with live model data on multiple panel

Summary

1. The imminent arrival (2011) of petascale computing makes it imperative that we develop some strategy for visualizing the flood of time-dependent 3-D data and simulations.
2. Interactive visualization of 3-D convection is feasible now with the deployment of hundreds of processors together with a fast network, multi-Terabyte storage at each visualization node. Large-display devices (CAVE, PowerWall) and web-portal services will allow for collaborative research.
3. This mode of operation (via a client-server application) will eventually be used by many parties around the world. A new paradigm in large-scale numerical modeling is now at hand.