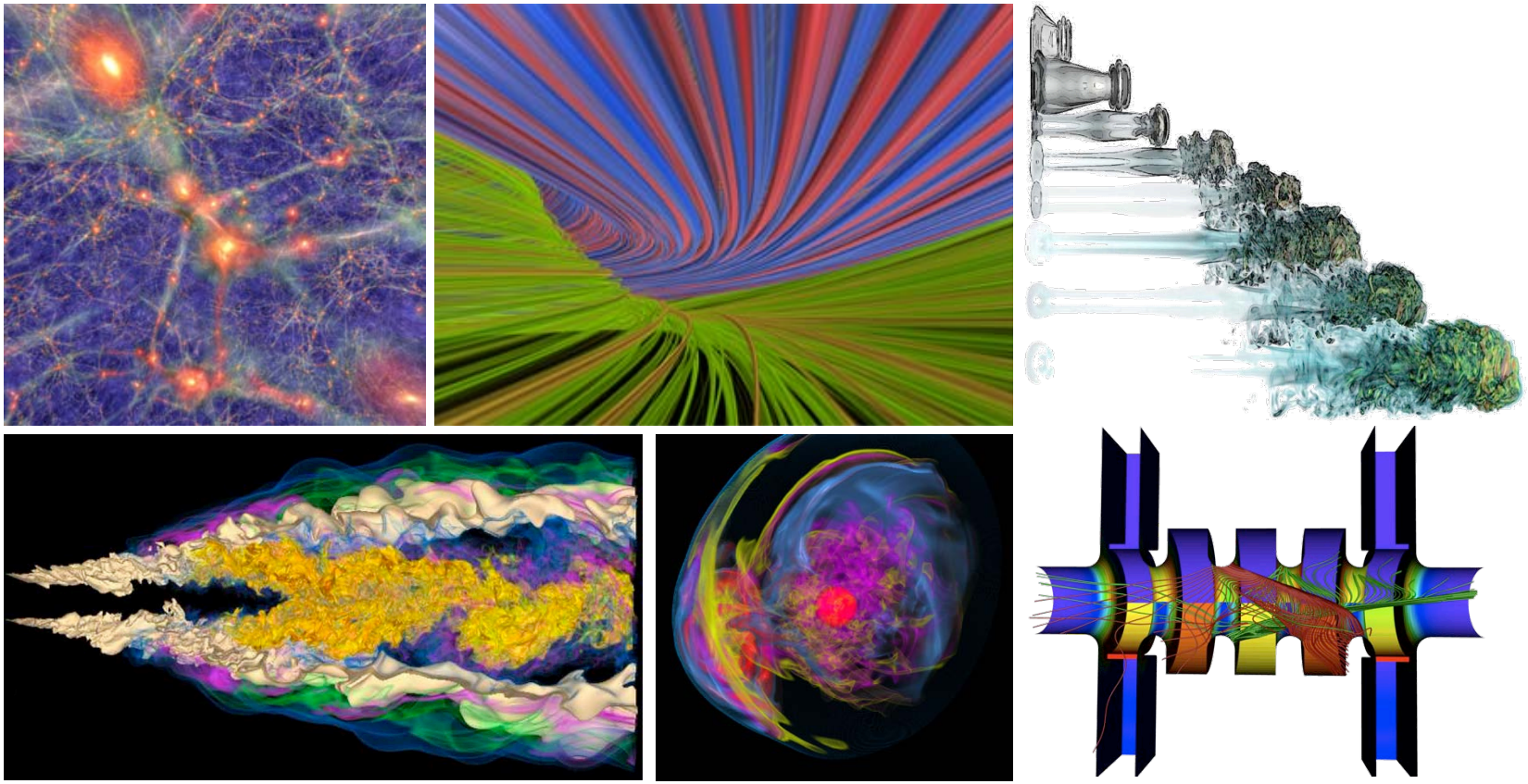
New Perspectives for Large Data Visualization Research

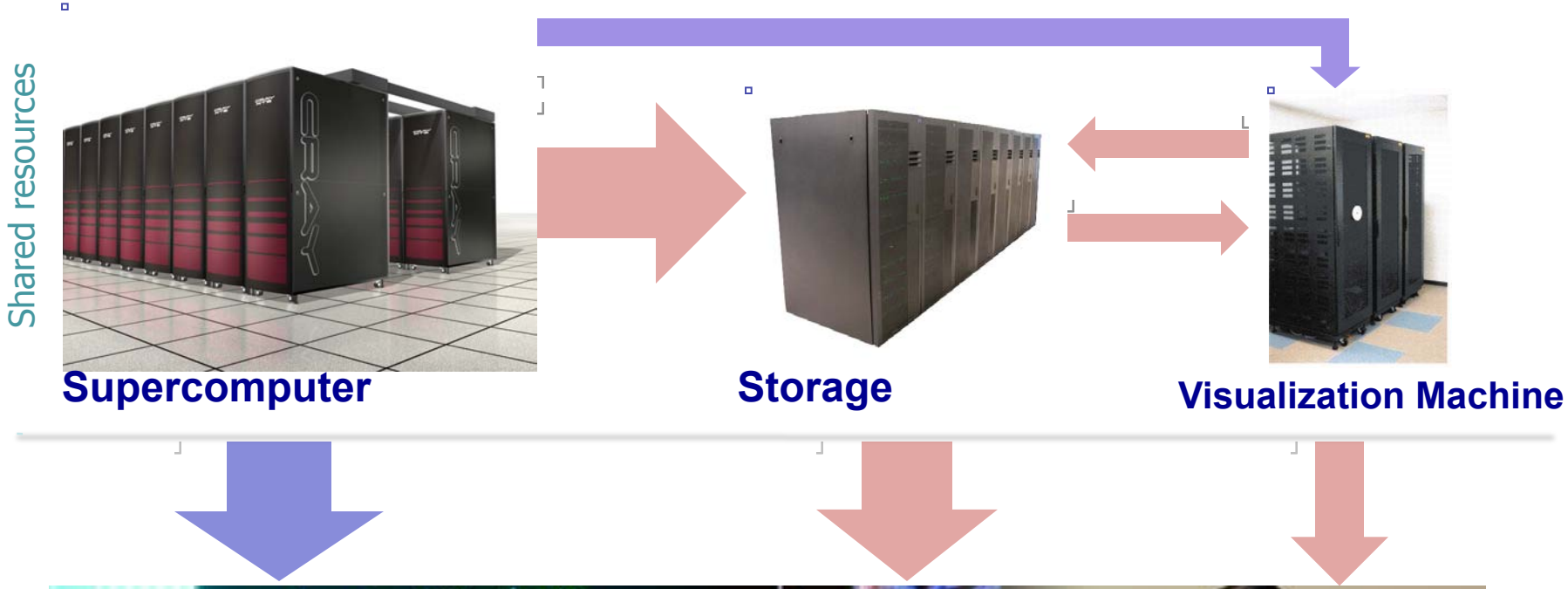
Kwan-Liu Ma
Department of Computer Science
University of California at Davis

Topics

- In Situ Data Processing and Visualization
- Scientific Storytelling



Large Data Visualization



See your data anytime?

Large Data Visualization

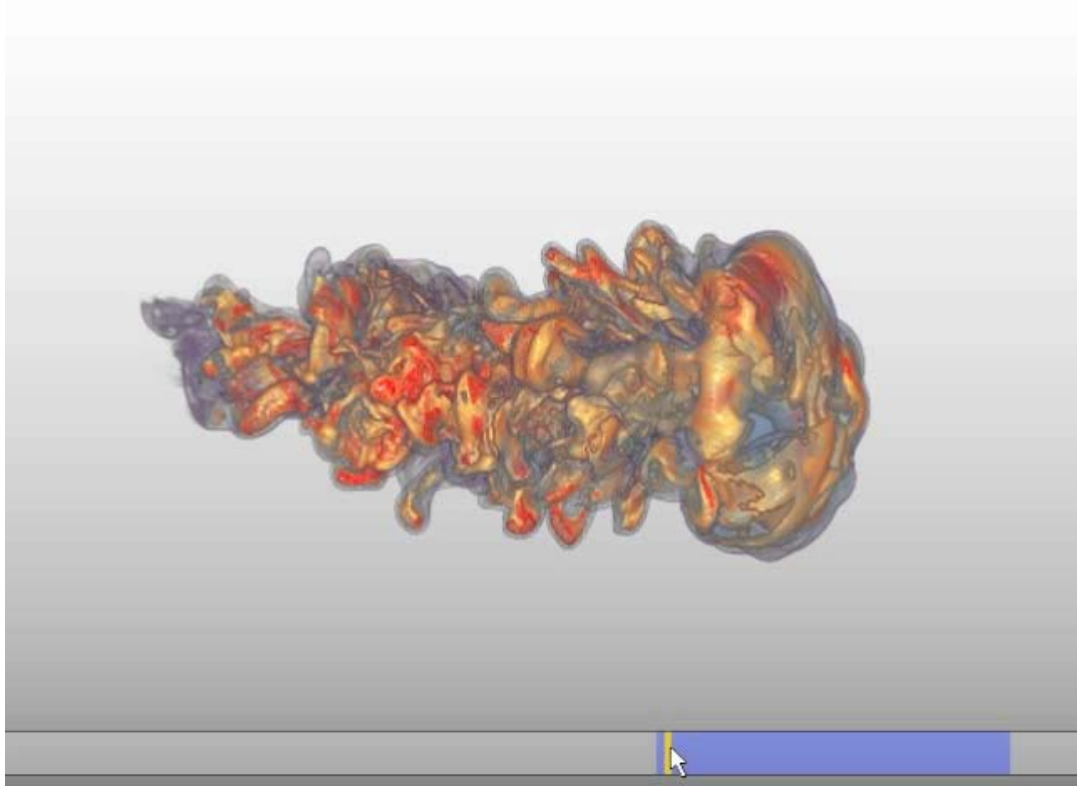
Supercomputer

Storage

Visualization Machine

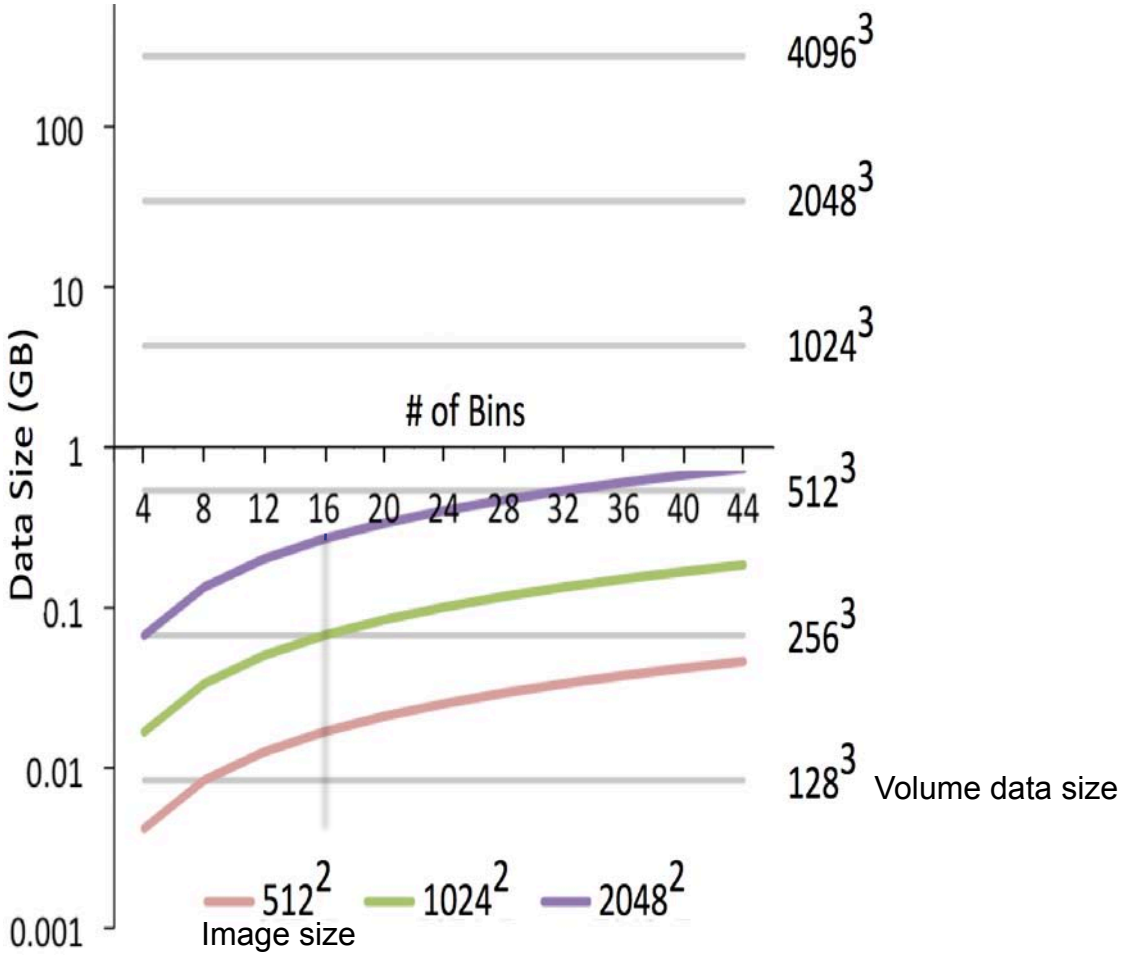
Visualization by Proxy aka Explorable Images

- A compact, intermediate representation of the data of interest
- Visualization by proxy for deferred interaction
- Explorable in
  - Spatial domain
  - TF space
  - Temporal domain
  - Rendering space

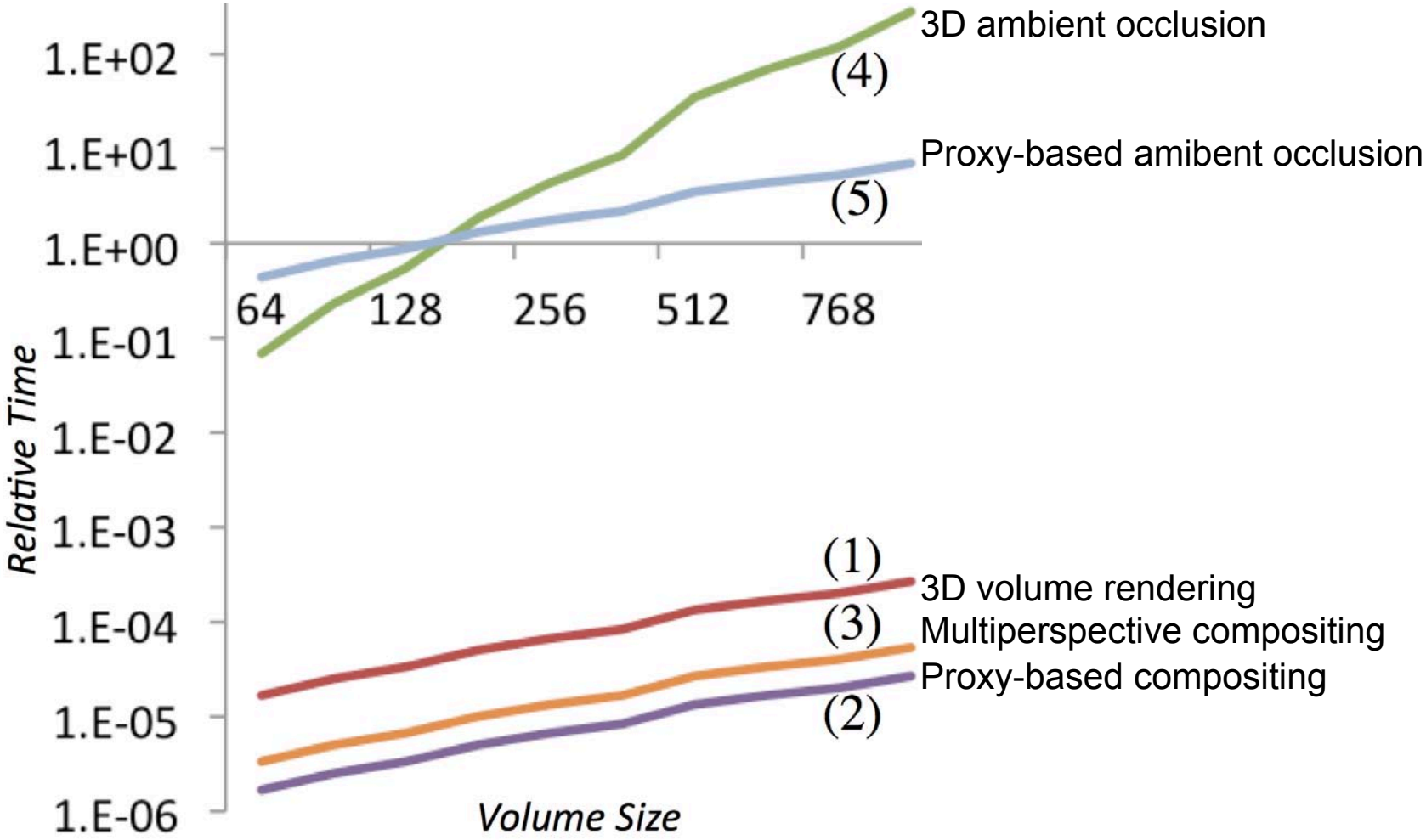


Storage Cost

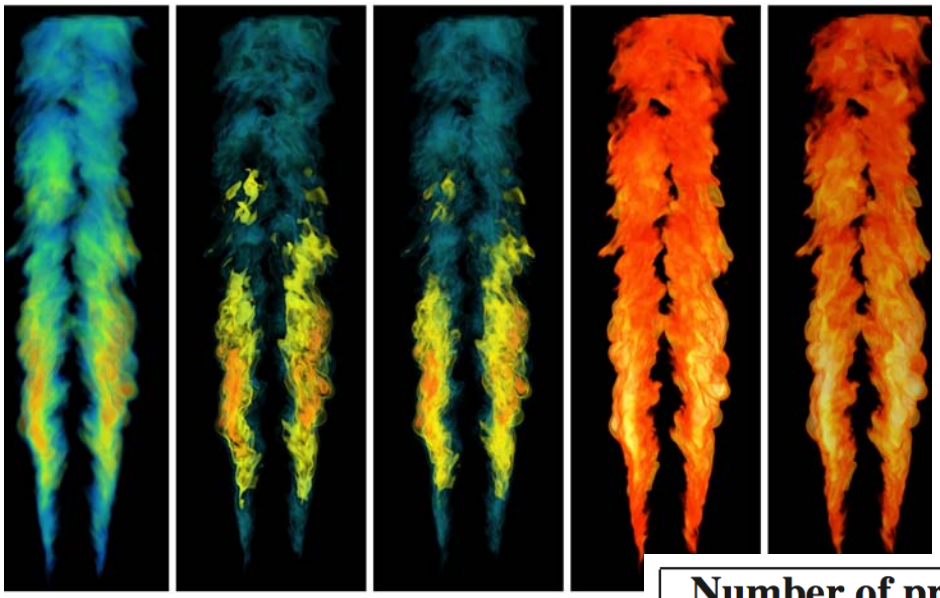
An explorable image is larger than a regular image,
but much smaller than the raw data and a typical video



Computational Cost

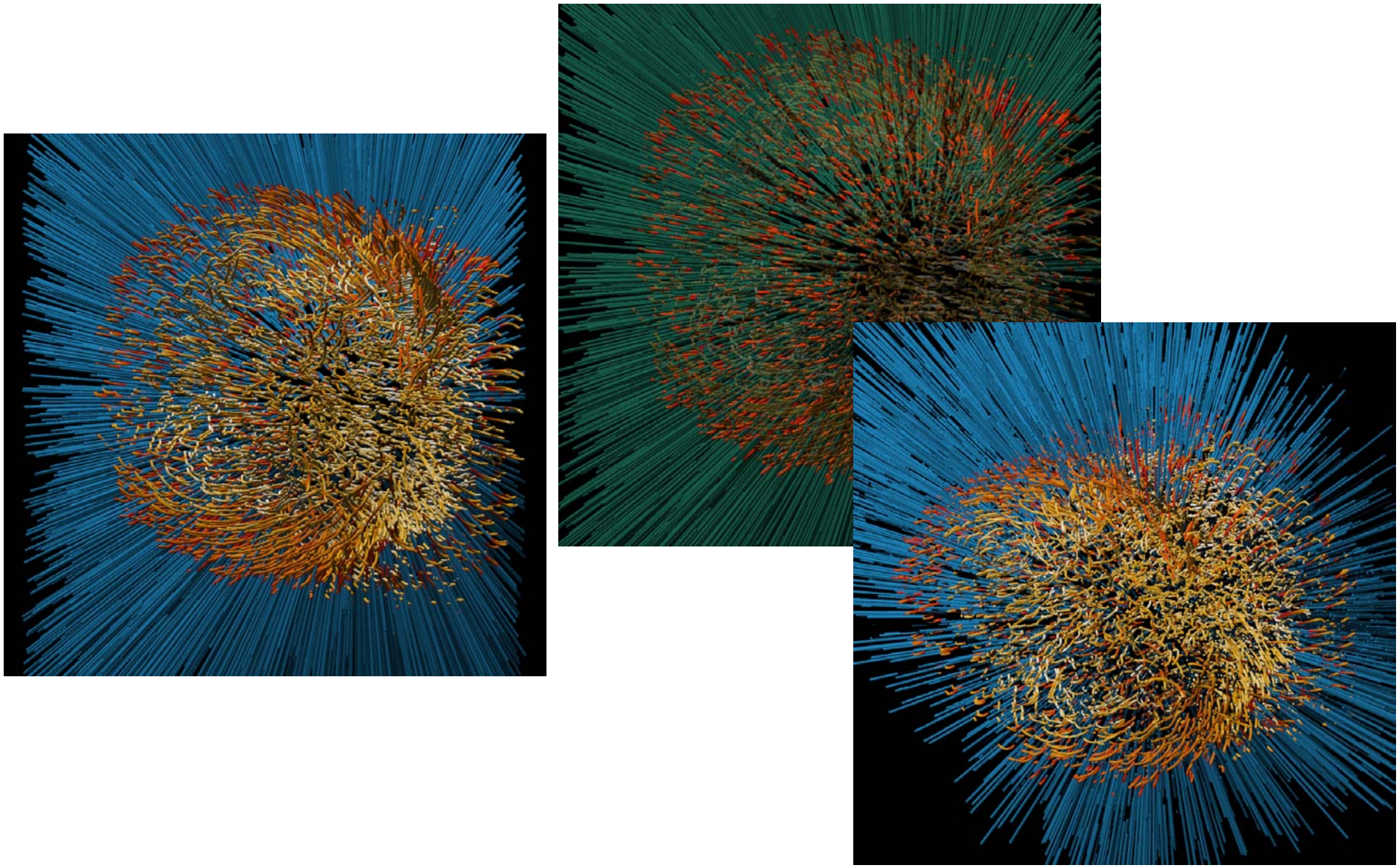


Visualization by Proxy



Number of processors	240	1920	6480
Simulation time (sec)	8.7204	9.3393	9.5573
I/O time (sec)	9.4563	26.051	52.565
Total volume rendering time (sec)	0.3817	0.6155	0.7359
Boundary voxel exchange	0.0042	0.0059	0.0064
Ray casting	0.0226	0.0148	0.0095
Image compositing	0.3549	0.5948	0.7200
Total IAF computation time (sec)	1.2775	1.3938	1.3973
Boundary voxel exchange	0.0026	0.0066	0.0068
IAF construction	0.0806	0.0729	0.0450
IAF compositing	1.1943	1.3143	1.3455

Particle Path Visualization by Proxy

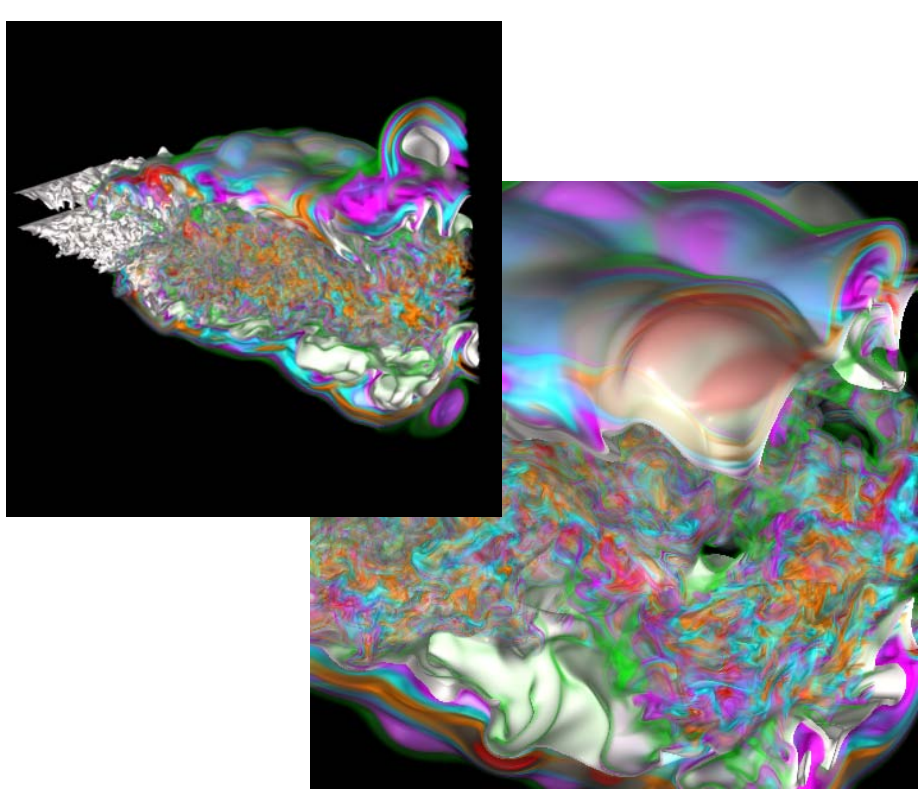


Visualization by Proxy

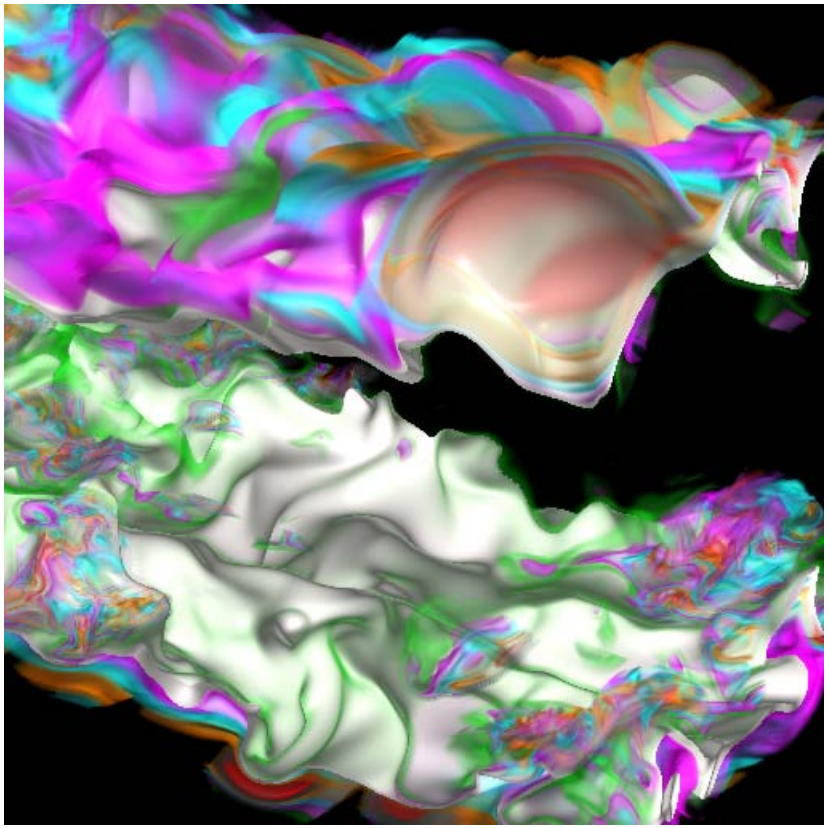
- A viable solution for interactive visualization of large, complex data
- In-situ generation of explorable images as a remedy to data reduction, a previewing solution, or a solution to make visualization that is impossible/ impractical to make after the simulation
- Data visualization on mobile devices or via a web browser
- An idea generalizable for different types of visualization

In Situ Data Reduction

- Lossless compression for floating-point data
- Importance based visualization and data reduction
- Other feature extraction and tracking methods

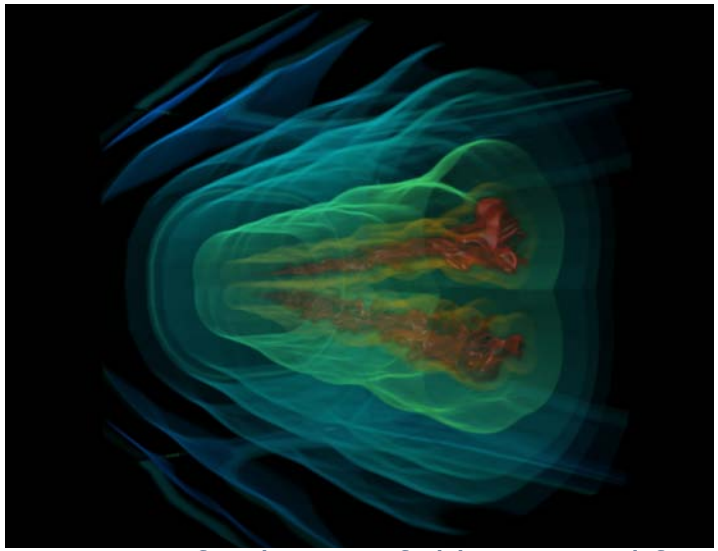


Small eddies are hidden in the multi-layer flow



Distance Field Computing

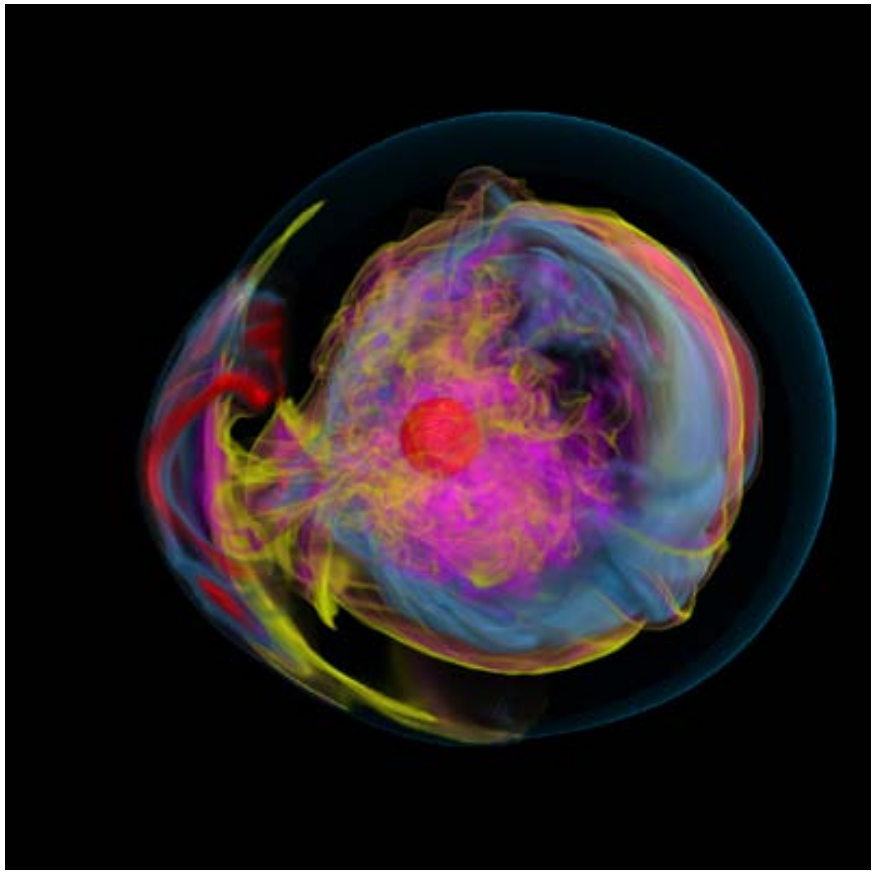
- Distance fields can be used as importance fields to guide rendering, data compression, sampling, and feature-based optimizations.
- Scalable distance field calculations have been achieved
- Tests on the parallel implementation show the data that must be exchanged is under 0.01% of the total data, and the cost to exchange the data is under 0.2% of the overall time.

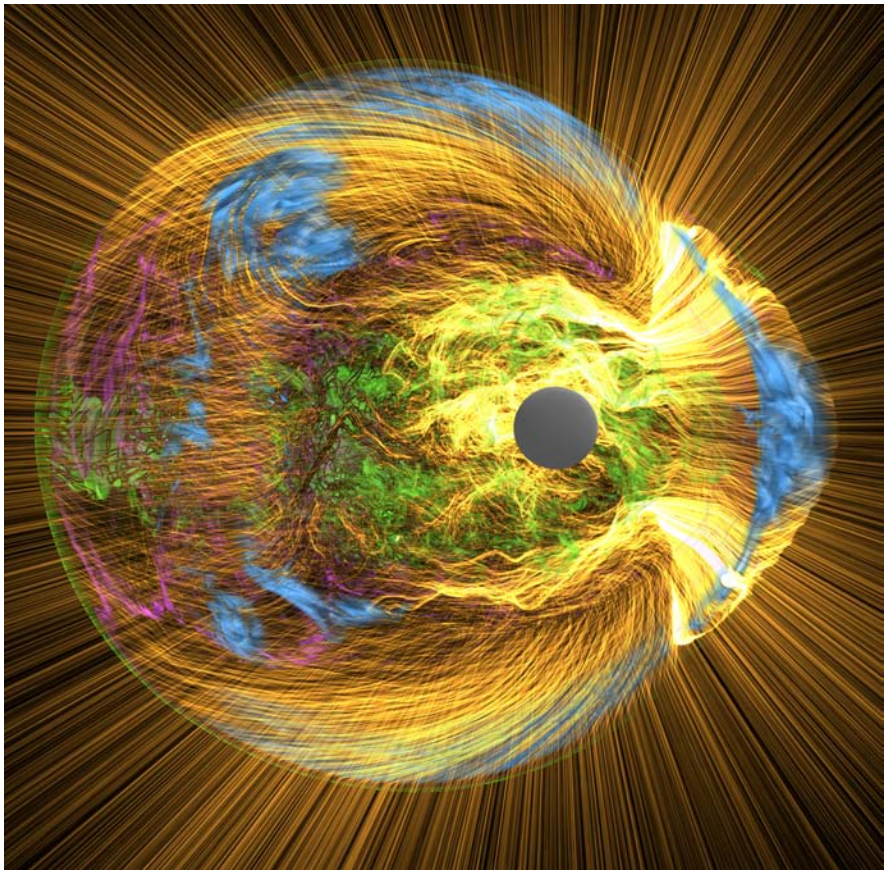


Depiction of a distance field computed from a feature surface in data generated by a combustion simulation.

Support for Scientific Storytelling

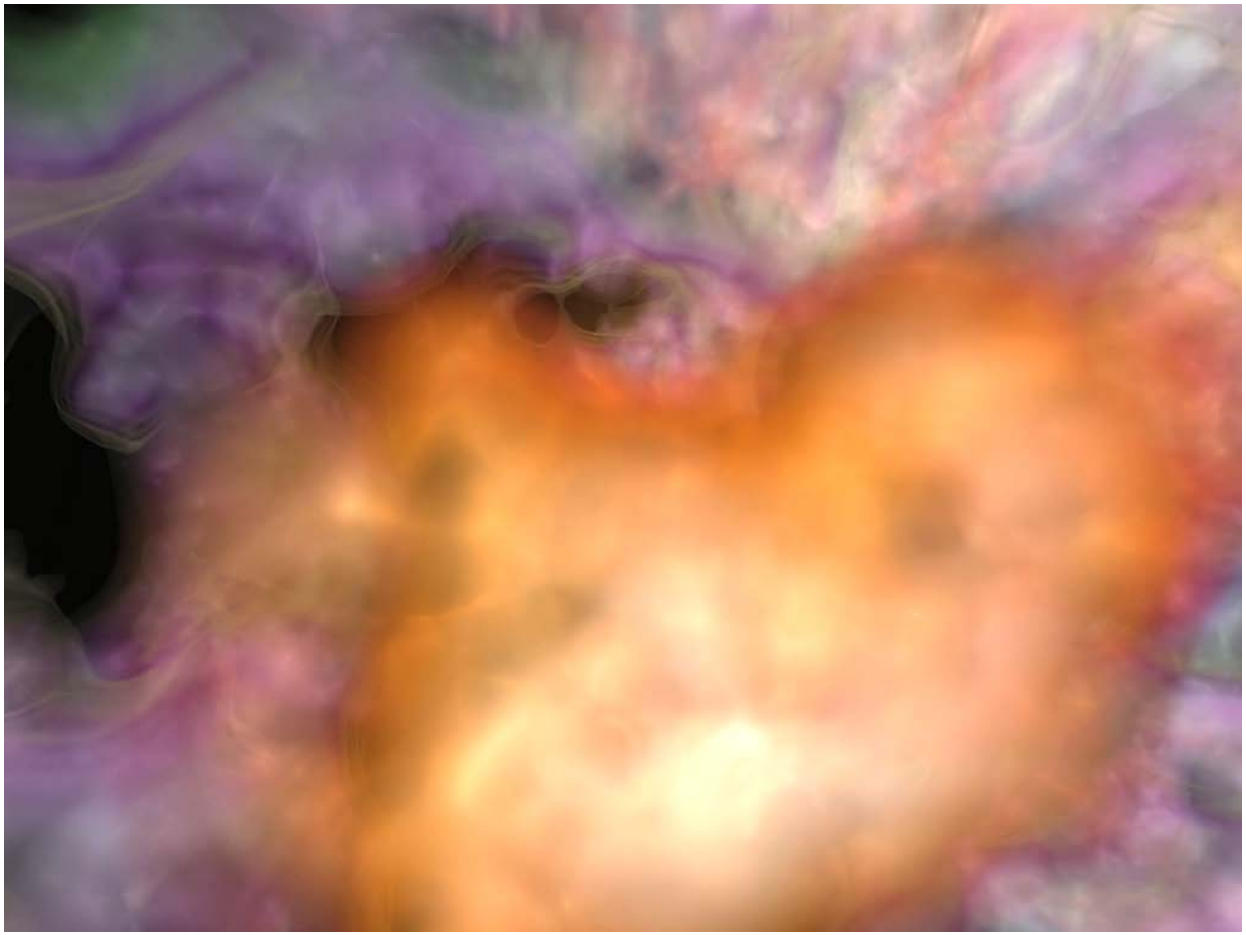
- High performance and high quality rendering





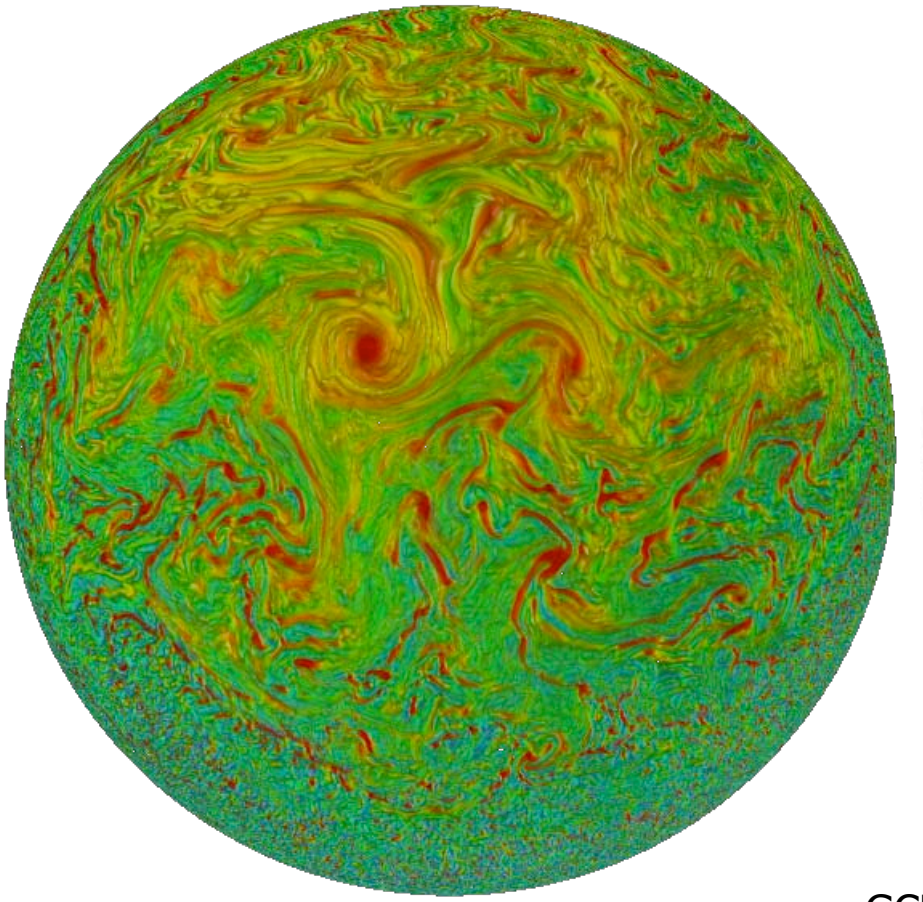
Support for Scientific Storytelling

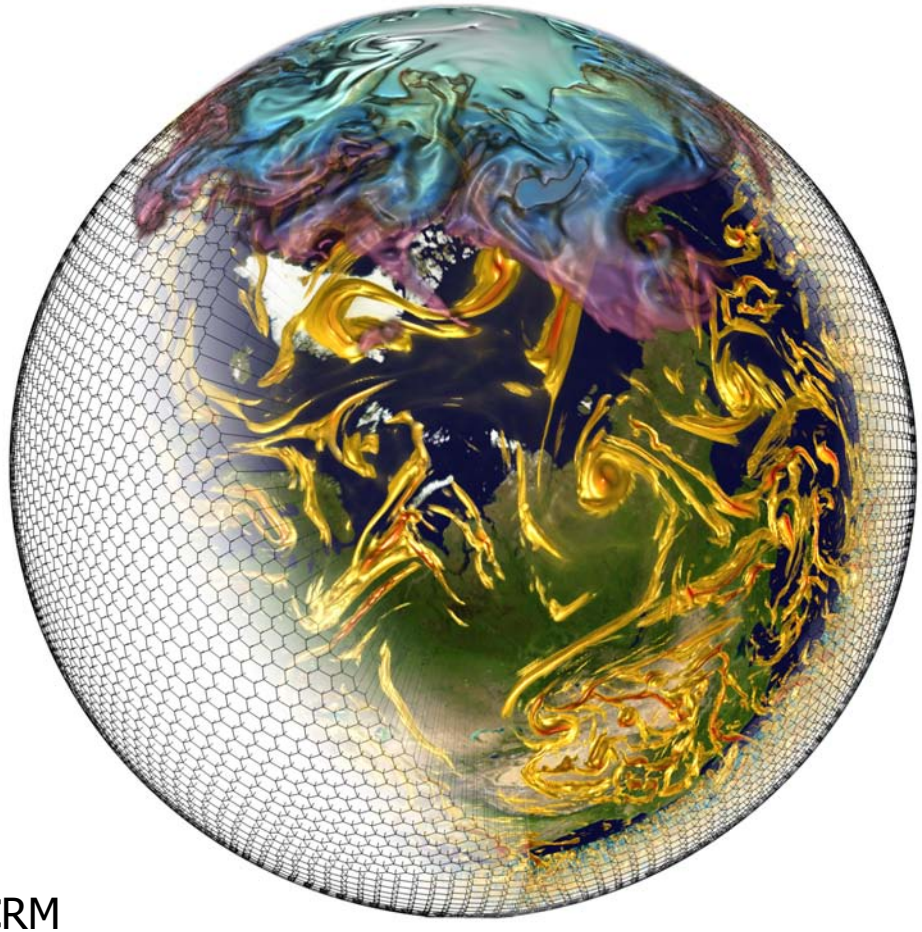
- High performance and high quality rendering



Support for Scientific Storytelling

- High performance and high quality rendering

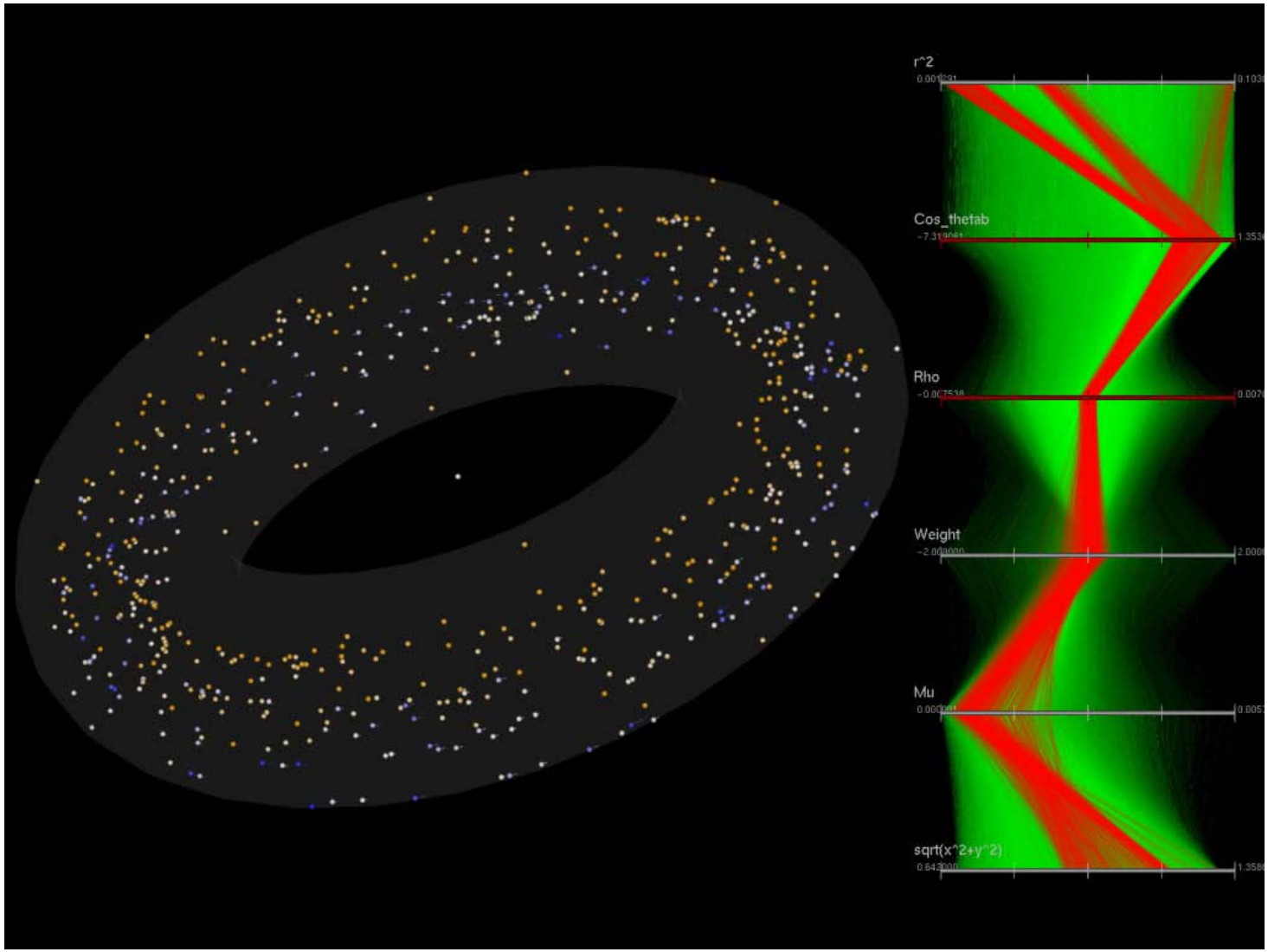




GCRM

Support for Scientific Storytelling

- Visualization interfaces

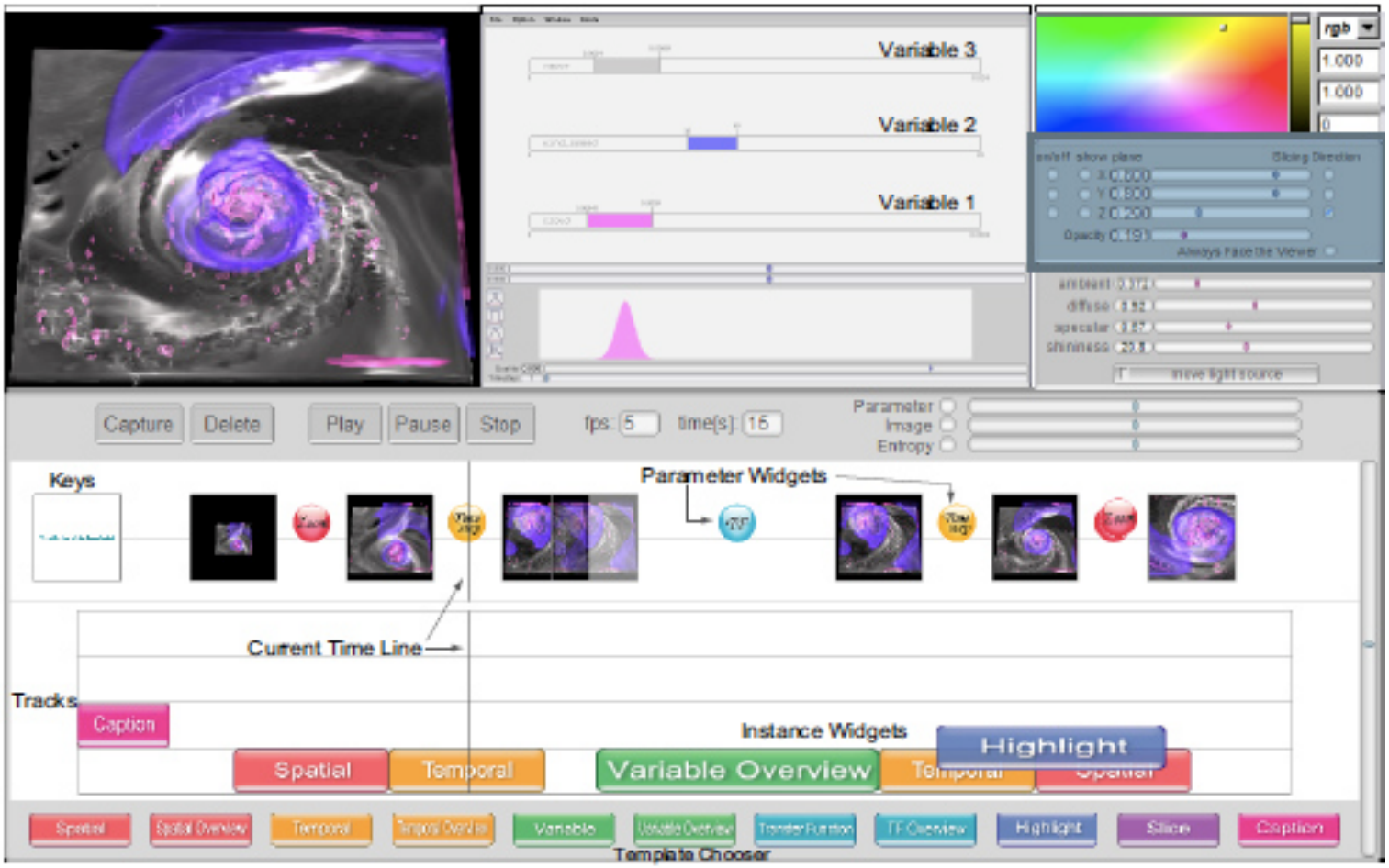


Support for Scientific Storytelling

- Visualization interfaces

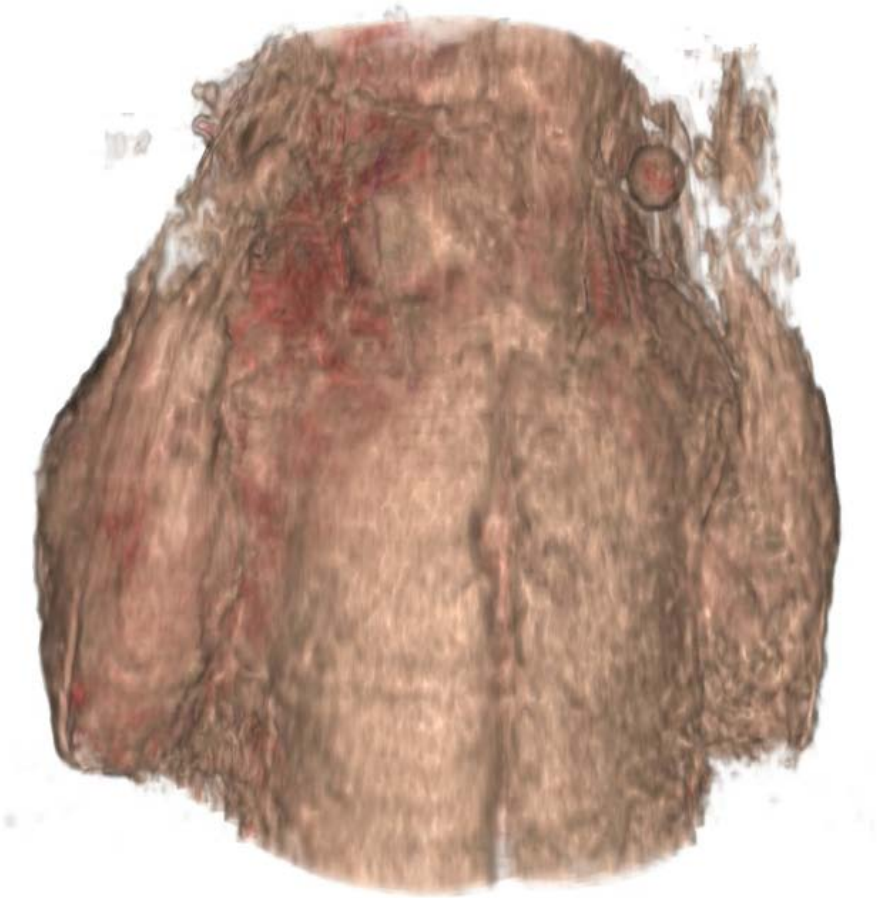
Support for Scientific Storytelling

- Visualization interfaces



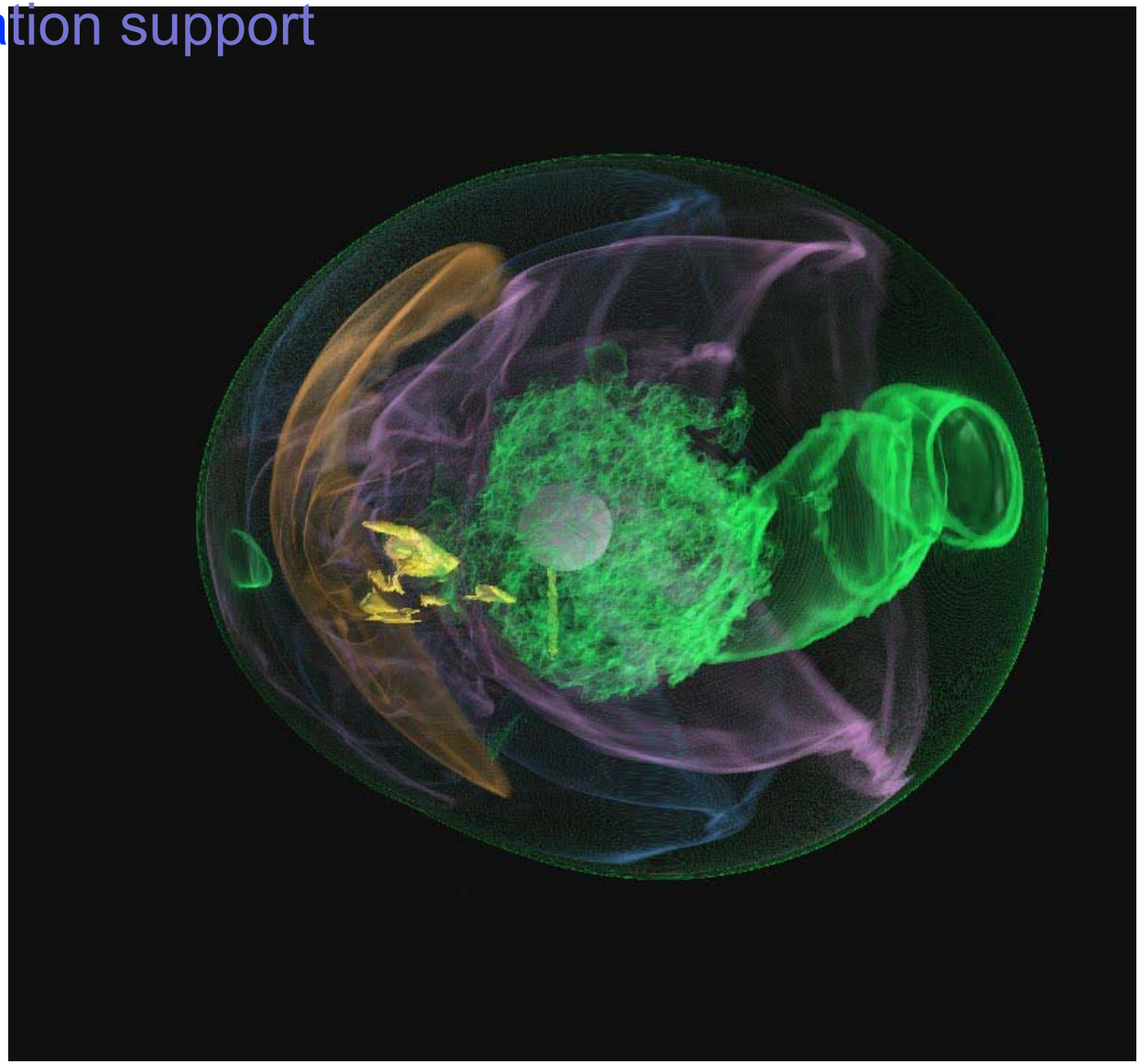
Support for Scientific Storytelling

- Animation support



Support for Scientific Storytelling

- Animation support



Support for Scientific Storytelling

- Video narratives