LA-UR- 08-7338

*Approved for public release; distribution is unlimited.*

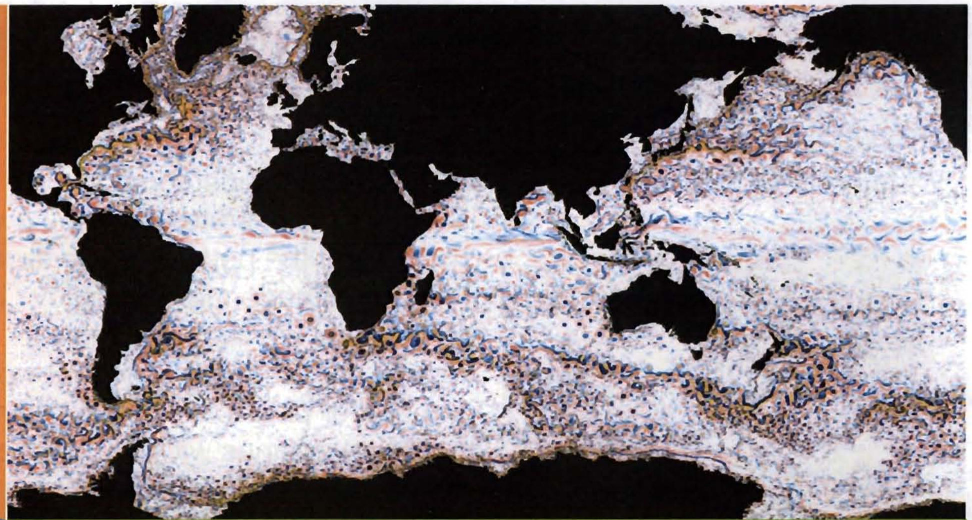
*Title:* Scientific Visualization at Los Alamos National Laboratory

*Author(s):* James Ahrens

*Intended for:* IEEE Supercomputing 2008
Austin, TX
November 16-21, 2008

Los Alamos NATIONAL LABORATORY
EST. 1943

# Scientific Visualization at Los Alamos National Laboratory



James Ahrens (team lead)
Patricia Fasel
Chung-Hsing Hsu
Li-Ta Lo
Patrick McCormick
John Patchett
Jonathan Woodring

Supercomputing 2008
November 16-24, 2008

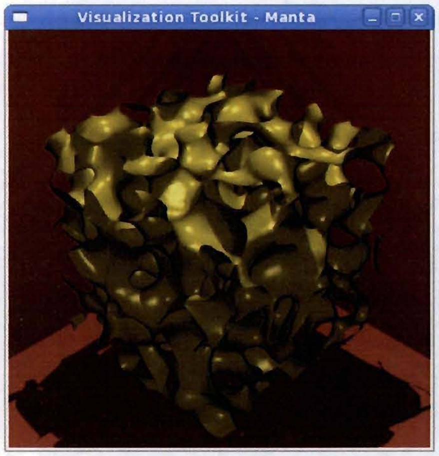
## Visualization research with open-source impact

**Scientific visualization is an essential tool for understanding the vast quantities of large-scale, time-dependent data produced from high performance computer simulations. Required research spans the areas of traditional computer graphics, scientific visualization and computer systems software.**

In the late nineties, there were no open-source or commercial visualization packages that could effectively visualize large datasets. This was a significant concern to the scientific simulation community because large-scale results were being generated and needed analysis. James Ahrens at Los Alamos led an effort to change this by researching and developing an open-source, multi-platform parallel visualization application and toolkit in order to process large data sets. The vision was to have an open-source package so members of the visualization community at the National Laboratories, academia and industry could participate and contribute to the effort. An open-source object-oriented visualization library, developed at Kitware Inc., was available, the Visualization Toolkit (VTK) (www.vtk.org). Using VTK as a starting point we extended the toolkit to run in parallel. By modifying the systems infrastructure of VTK to support data streaming (the ability to incrementally process a dataset), data parallelism and distributed computing, the new parallel toolkit was able to extend the existing full range of visualization, imaging and rendering algorithms available in VTK. In addition, parallel rendering algorithms were added. Ahrens also led a second R&D effort, created an end-user visualization tool for large-scale data called ParaView (www.paraview.org) for use by scientists. Scientists are not interested in writing C++-based visualization code to analyze their data. ParaView encapsulates the details of configuring parallel visualization and rendering algorithms using a graphical user interface or scripting language. ParaView is designed to support the day-to-day workflow and facilities issues that laboratory and academic scientists encounter.

# Petascale Visualization

With the advent of the first multi-core petascale supercomputer, Los Alamos's Roadrunner, there is a pressing need to address how to visualize petascale data. The crux of the petascale visualization performance problem is interactive rendering. To achieve high-performance on multi-core processors, we tested with multi-core optimized ray-tracing engines for rendering. For real-world performance testing and to prepare for petascale visualization tasks we interfaced these rendering engines with vtk and ParaView. Initial results show that rendering software optimized for multi-core CPU and Cell processors provides competitive performance to GPU clusters, for the parallel rendering of massive data. As part of this project, we interfaced the multi-core optimized Manta parallel raytracer as a rendering back-end for VTK and ParaView and it will be available in the VTK and ParaView repository soon.




Presentation by J. Ahrens at Ultrascale Visualization Workshop – Sunday, Nov. 16th, 9:30 AM

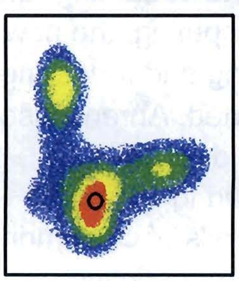
# Distance Visualization

The simulations that run on the next-generation of supercomputers will advance science. The scientists that work on these supercomputers typically access them via a network. The sizes of their petascale results will most likely overwhelm planned network upgrades. Visualizing, analyzing and understanding results are key to effective science. Thus, the distance visualization challenge, addressed by this proposal, is to alleviate the barriers to visualizing results over wide area networks. We have worked to improve our open-source software infrastructure to better facilitate distance visualization. ParaView supports parallel remote visualization using remote servers and a local client providing a user interface. ParaView is a natural base to add increased functionality to improve the user's capability for distance visualization. Specfically, we recently released a new version of ParaView that supports prioritized streaming. A massive dataset can be incrementally read and visualized directly from disk at a remote supercomputing site, streamed over the network, and progressively rendered on the scientist's local desktop display.

Presentation by J. Ahrens at Oak Ridge National Lab booth – Tuesday, Nov. 18th, 1:00 PM

# Comparative and Quantitative Visualization

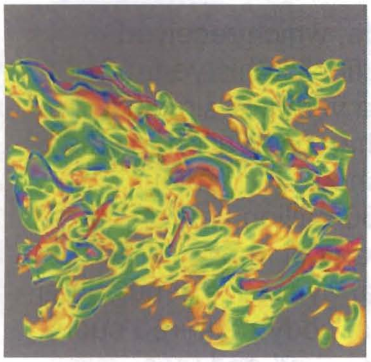
Our universe contains 22% invisible dark matter and 74% mysterious dark energy. To understand the underlying physics, scientists have proposed different simulation-based predictive models that balance between precision and execution efficiency at different levels. Our work provides scientists an interactive framework within ParaView to insightfully compare these different simulations models. The framework supports qualitative and quantitative comparisons in a coordinated multi-view fashion that assist scientists to rapidly identify promising hypotheses and to firmly verify them. We believe that this approach, the iteration of the hypothesis-verification process will help to scientists easily and accurately characterize difference between these simulation models.



Presentation by J. Ahrens, D. Daniel and C. Hsu at ASC booth – Tuesday, Nov. 18th, 4:00 PM

## Time-Varying Data

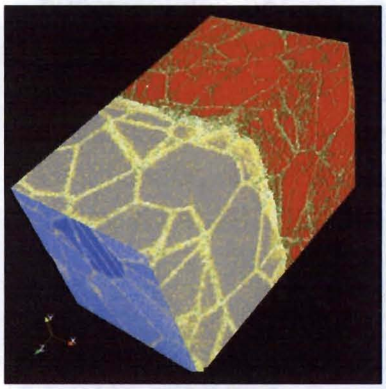
Visualizing time-varying data has traditionally had difficulty in providing a color map for an entire time sequence. This is due to the dynamic value range of the data set over the time sequence makes it hard to map values to colors that are meaningful for the entire sequence. To address this issue, we analyze the time sequence, through parallel clustering and time sequencing, to create a color map that reflects the change in value over time. The software to create color maps (transfer functions) for time-varying visualization will be freely released to the public through SciDAC (scidac.gov), in cooperation with The Ohio State University and the UltraScale Visualization Institute.



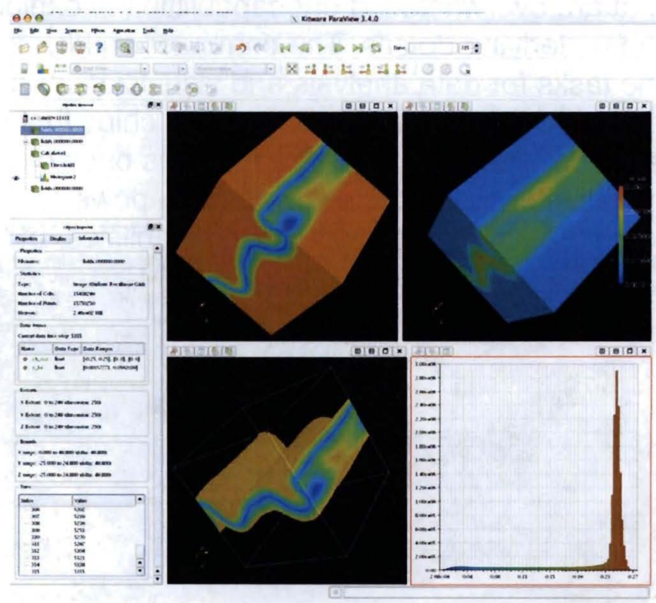
## Scout: Visualization and Data Analysis on Emerging Architectures

Many simulation and modeling efforts are continuing to produce data at ever increasing rates that are overwhelming our capabilities to explore, hypothesize, document, and thus fully interpret the underlying details. The computer architectures that support these computations, including the tasks for data analysis and visualization, are undergoing a revolutionary change as manufacturers transition to building chips that use an increasing number of processor cores. In addition, graphics hardware that was once designed entirely for the rendering of polygonal primitives has rapidly evolved into a powerful general-purpose processor. While these trends have the ability to provide new capabilities and increase performance, they will do so in a disruptive manner, potentially placing a significant strain on software development activities. The goal of our effort is to exploit these architectures to provide an integrated environment for high-performance data analysis and visualization. In addition, we provides scientists with a simplified programming model for these processor architectures, enabling them to perform both the numerical and visualization tasks necessary for the study of large and complex data sets.

# Visualization Demos at Supercomputing



LANL's SPaSM (Scalable Parallel Short-range Molecular dynamics) code, which received Gordon Bell Prizes in 1993 and 1998, is a finalist this year after being ported to Roadrunner. SPaSM's is a versatile simulation, and has a great breadth of phenomena it has been used to study. The visualization of SPaSM in ParaView shows the various crystalline forms that iron can adopt, in a 30-million particle time-series simulation, and underscores one many of iron's unique properties. SPaSM's large-scale molecular dynamics simulations reveal a rapid transformation from the initial body-centered cubic structure to a mixture of hexagonal close-packed and face-centered cubic products in polycrystalline iron under shock loading, which can be seen in the visualization.



LANL's VPIC, another Gordon Bell finalist, is a highly optimized, state-of-the-art, particle-in-cell (PIC) implementation for full relativistic, 3D, explicit, kinetic plasma simulation, which has also been ported to Roadrunner. VPIC has been applied with success in both basic and applied science. The visualization of VPIC in ParaView shows the electron density resulting from magnetic reconnection of a Harris geometry, which is a plasma equilibrium with a reversal of magnetic field across a neutral layer, with pressure equilibrium established by an initial high density at the center. The visualization shows the time-series dynamics of a 16 billion particle simulation on a 1000x1000x1000 mesh.

Visualization is providing the capability to view the results of a simulation, as it occurs, to determine the integrity and correctness of a particular simulation. One of these pieces is a VPIC reader, which can decide how simulation partitions relate to each other, and repartition the data over the visualization processes. The reader knows how to access the data for a particular time step and the offsets for any piece of data, along with shared memory buffers between the simulation and visualization to eliminate the need to write to disk. Both ParaView and EnSight are supported, so that the user can choose the visualization platform.

Additionally, VTK and ParaView have been updated to support the DLP 3D stereo technology found in newer HDTVs, such as the demonstration display in the LANL booth. The checkerboard stereo image format, which DLP 3D uses, interleaves the left and right eye images in a checkerboard fashion, providing both stereo eye images in a single frame. The DLP 3D stereo image format will be supported by VTK and ParaView in an upcoming release, and freely available to the public through Kitware.

Graphic for Distance Visualization Section
-this is Salinity @ ocean surface in tropical pacific from Mat Maltrud