Algorithms, Frameworks and Toolsets for High Performance, Remote and Distributed Visualization

E. Wes Bethel and John M. Shalf, Lawrence Berkeley National Laboratory

Summary

Modern computational and experimental sciences have entered an age in which information management and understanding are critical bottlenecks in the scientific process: we are literally generating and collecting data much faster than we can understand it. Since the scope of the problem is much greater than can addressed by an individual research effort, the Office of Science sponsors several research projects that take aim at different aspects of the larger problem. The Berkeley Lab Visualization effort focuses on technologies for petascale data visualization and understanding with a particular emphasis upon architectural and implementation issues germane in remote and distributed computing environments. Recent accomplishments include novel approaches for petascale visualization, including a revolutionary new approach called "Query-Driven Visualization," along with a new technical capability that delivers interactive, multiresolution 3D visualization over network links to remote researchers.

Query-Driven Visualization

Scientific visualization – the transformation of abstract data into readily comprehensible images – plays a key role in the scientific process by offering the means to quickly gain understanding of phenomena hidden in data produced by scientific experiments and simulation. A central challenge faced by modern science is detecting and understanding features hidden in large and complex datasets.

Query-Driven Visualization (QDV) is a revolutionary new approach to large and complex data analysis and visualization. In a nutshell, it focuses visualization and analysis processing only on data deemed to be "interesting" by the scientific researcher. "Interesting" is defined quantitatively through compound Boolean expressions like (Temperature > 1500) AND (CH_4 > 30%).

QDV is the intersection of scientific visualization and scientific data management. The LBNL prototype implementation shown first at SC04[1], named "DEX" (short for dextrous data explorer), uses the patented FastBit bitmap indexing technology[2] for performing efficient data queries. Our SC04 demo illustrated the power and flexibility of the approach to quickly visualize interesting scientific features hidden in massive and complex datasets. Our SC04 demonstration used core-collapse supernovae simulation data produced by the TeraScale Supernova Initiative[3] SciDAC project. QDV offers a promising approach to large and complex data visualization and analysis. It helps accelerate scientific insight by quickly locating and visualization only those features of interest in scientific data. QDV.

Current and future research activities aim to refine and extend the core QDV computer science algorithms and related visualization algorithms to a diverse set of scientific data and applications. Applying the principles to

[1] http://vis.lbl.gov/Events/SC04/Dex/index.html
[2] LBNL Scientific Data Management Research. See http://sdm.lbl.gov/fastbit/.
[3] http://www.phy.ornl.gov/tsi/

[*] Telephone Number, E-mail address

multivariate datasets that are on the scale of tens of terabytes will provide insight into their effectiveness on problems of scale as well as provide new visualization capabilities to the scientific research community. The techniques will be extended to accommodate multiresolution queries, which will result in performance increases drastically needed to tackle petscale data sizes. The concepts are logically extensible to temporal visualization and analysis, as well as comparative visualization and analysis, both of which are capabilities acutely needed both computational and experimental scientific research projects.

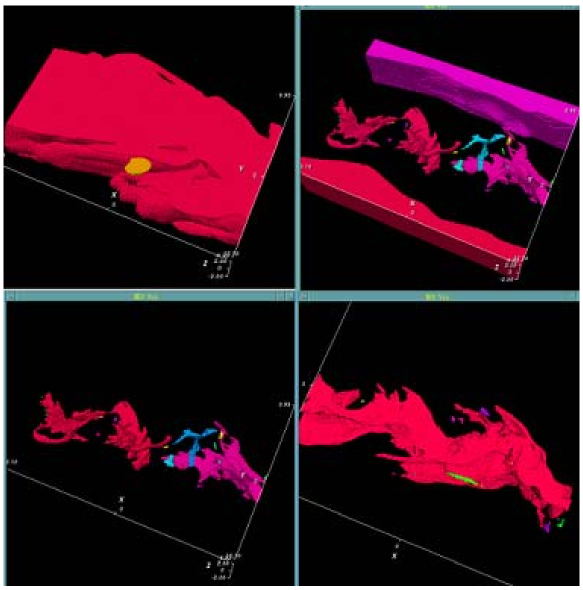

Figure 1. Query-Driven Visualization concepts applied to combustion simulation data. In the upper-left image, we first extract and visualize an interesting datasubset where the concentration of CH_4 is greater than 30%. The upper right image shows grid cells where temperature is less than T_1. The lower left image shows the grid cells where both conditions $CH_4 > 30\%$ and Temperature $> T_1$ are true. In the lower right image, we have relaxed the constraints to include more grid cells.

Remote Delivery of Interactive, Multiresolution 3D Visualization

The concept of "latency tolerant remote and distributed visualization" refers to an implementation whereby desktop interactivity is decoupled from the speed of the underlying network that connects distributed visualization components. Our recent research has focused on leveraging existing media standards as the basis for delivering latency tolerant content to a remote viewer. Specifically, we explored using QuickTime VR (QTVR) Object movies as the delivery vehicle for interactive, 3D visualization content.

The Mbender project demonstrates use of the QTVR Object Movie media format to deliver interactive, time-varying 3D precomputed visualization results to a remote user. With a standard web browser or QTVR player, a user can interact with precomputed 3D and 4D visualization results. The approach supports arbitrary view orientation changes, zoom-in and zoom-out, and temporal browsing. Such changes have been shown by studies in cognitive psychology to improve a viewer's understanding of 3D depth relationships.

Current research focuses on a combination of caching and client-pull data transfer algorithms that support on-demand increased resolution when needed.

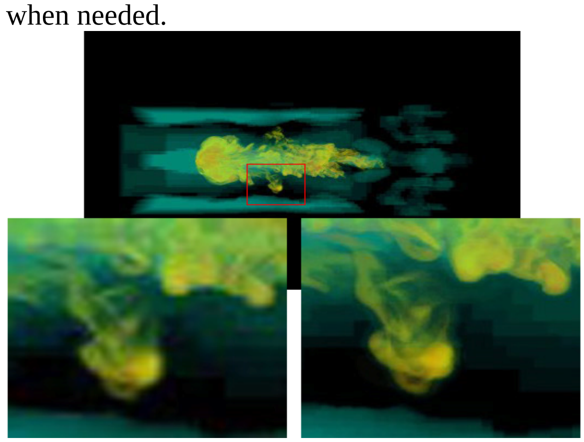

Figure 2. The fixed resolution of contemporary media formats results in a loss of crucial detail during remote, retained-mode viewing. A zoomed in view of a high resolution visualization (lower right) contains more visual detail than a fixed-resolution representation (lower left) of a contemporary scientific visualization (top image).

For further information on this subject contact:
Name: E. Wes Bethel
Org: Lawrence Berkeley National Laboratory
Email: ewbethel at lbl dot gov
Phone: 510-486-7353