Interactive Remote and Distributed Visualization of Fusion Simulation Results

Problem Statement

The National Energy Research Scientific Computing Center (NERSC) is DOE's premier facility for large-scale computational science. One of the many projects that rely on NERSC for computing capabilities is the M3D (Multilevel 3D) Project. It aims to accurately model plasmas, a magnetic fusion phenomenon viewed as a potential source of energy in the future. Modeling plasmas is a numerically challenging endeavor due to the complex physics required to accurately to produce a meaningful simulation. Like many contemporary computational science efforts, the M3D project generates a substantial amount of simulation output that must be analyzed. The M3D team is distributed across several organizations: the Princeton Plasma Physics Laboratory, New York University and the Massachusetts Institute of Technology. NERSC, on the other hand, is located in Oakland, California. Related to the M3D code is the Gyrokinetic Toroidal Particle-in-Cell code (GTC). Output from a series of M3D or GTC runs can be as large as a Terabyte. In order to analyze simulation results, one must either transfer the entire simulation output from NERSC to a local machine, or perform analysis on NERSC machines. Transfer of such large datasets across the network introduces delay into the time between thesis and confirmation. An alternative approach is to perform analysis on NERSC machines, thereby eliminating the cost of very large data transfer. The challenge in such an approach is maximizing the interactivity of visual data analysis tools for remote users.

Approach

The NERSC center has unique resources that are especially tailored for interactive, high performance remote visualization. The central interactive visualization resource is *Escher.nersc.gov*, which is a large symmetric multiprocessor equipped with multiple gigabit interconnects to NERSC storage systems, as well as a large amount of main memory and substantial directly attached storage. Using Escher, we employ a *pipelined* architecture to support the demands of interactive, high-performance remote visualization. The pipelined architecture refers to an assembly-line organization of software "workers" that each contribute to an overall work flow. The first worker in the assembly line – the server – runs interactively in parallel on Escher. Parallel execution of the server provides substantial I/O and processing capabilities where it is needed – close to the data. The second worker in the assembly line runs on the scientist's workstation in their office, and performs 3D rendering at interactive rates. Visualization results – geometry – are transmitted between the two stages in the pipeline. This combination has proven effective at providing interactive 3D scientific visualization capabilities to remotely located NERSC users. Amortizing data I/O and visualization processing over parallel processors located close to the data provides capabilities that are simply not available on any desktop platform. These capabilities are an example of NERSC's commitment to providing the best possible tools and infrastructure to the computational science community.

Benefits

The pipelining architecture provides new capabilities that accelerate the analysis and understanding of computational science efforts. It enables the ability to quickly perform interactive 3D visual analysis of complex data from a remote location without having to download large datasets from NERSC. The cost savings are difficult to quantify because they vary as a function of parameters that are specific to each dataset and visualization technique. Generally speaking, the savings can be characterized as the ability to quickly run a parallel visual data analysis tool on a machine "close to" the data, thereby leveraging expensive parallel interactive resources, while sending only a fraction of the data to the desktop where it may be interactively manipulated. In the case of M3D simulations, which produce complex time varying 3D output, having the ability to perform such interactive visual analysis is a crucial part of the scientific process. Reducing the

time required for such analysis from hours or days to seconds or minutes represents a substantial breakthrough for the NERSC computational science community.
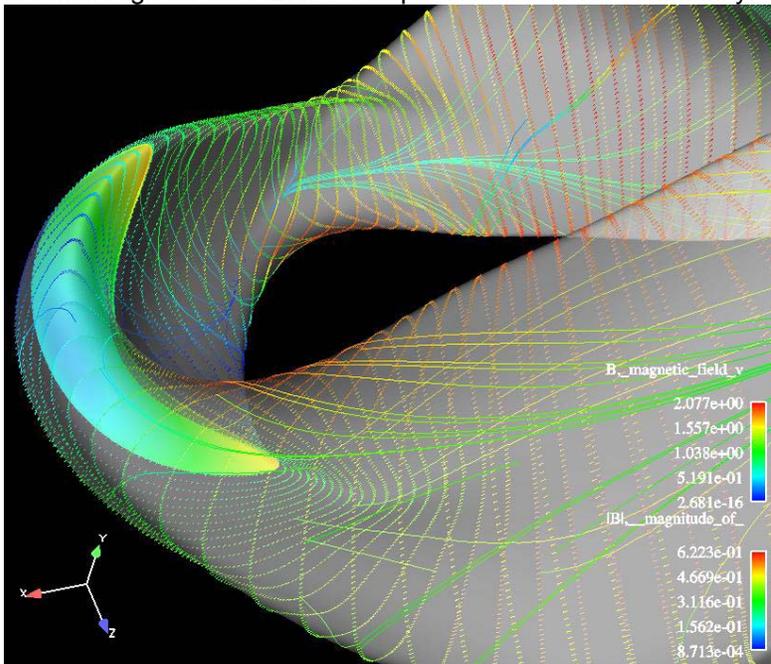

Figure 1.   M3D Simulation Visualization: traces along the magnetic field with particles traces emitted from seed points that lie on a plane. Colors indicate the magnitude of the magnetic field. The visualization software used these images is Ensight Gold, which is a commercially available product. *Cristina Siegerist, LBNL/NERSC and Scott Klasky, PPPL.*
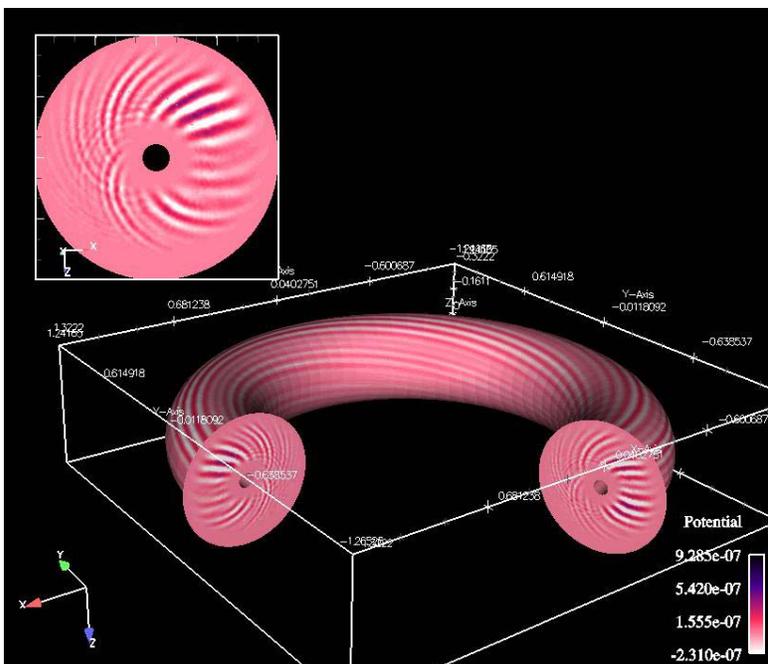

Figure 2.   GTC Simulation Visualization: poloidal planes and flux surface colored by potential. *Cristina Siegerist, LBNL/NERSC and Scott Klasky, PPPL.*
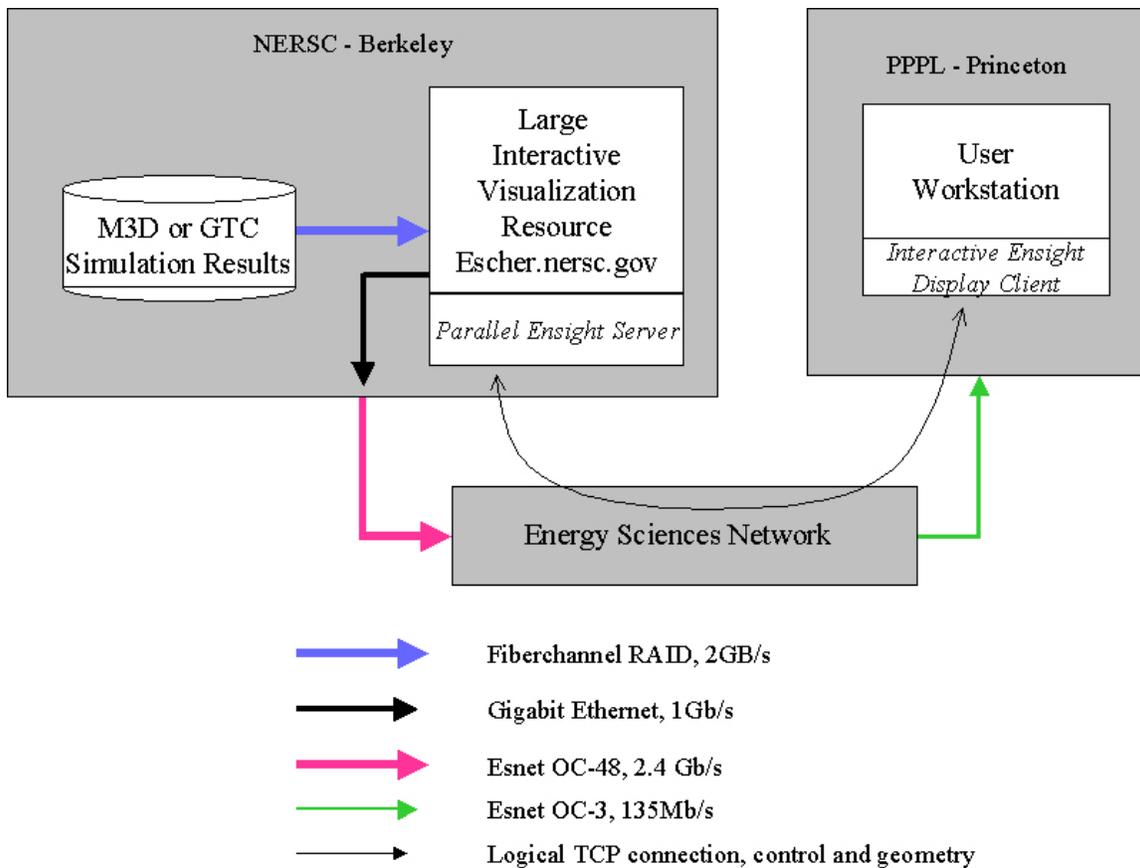
Figure 3. Simulation data is read from a Fiberchannel disk array attached directly to an interactive resource. The simulation data is processed interactively in parallel by the Ensight visualization server component. The geometry produced by the server component leaves Escher over a Gigabit Ethernet link, then leaves NERSC over an OC-48 connection, and then transits the ESnet backbone, finally entering PPPL over an OC-3 connection. A user at a workstation in Princeton then may perform interactive transformations over the geometry once it fully arrives.

Contact Information
Cristina Siegerist, LBNL/NERSC. cesiegerist@lbl.gov, 510-495-2520.