"Standing on the Shoulders of Giants – Becoming The Giants"

Patricia Crossno*, Ph.D., Sandia National Laboratories
Lee Ann Fisk, Sandia National Laboratories
Kenneth Moreland, Sandia National Laboratories

Summary

Visualization researchers at Sandia National Laboratories, funded by the Office of Science, are working to enable scientists to interactively view the results of extremely large and detailed simulations of physical phenomena. The software that has been developed is being shared with the larger research community to act as a foundation upon which others can build.

As physicists explore the nature of dynamic events, such as how an object breaks apart when hit by a projectile or how a wire vaporizes when large amounts of current flow through it, they develop theories about how the universe works. Previously, the only way for scientists to test their theories was through experimentation. Now, due to recent advances in technology, scientists can also use computer simulations to model their ideas and perform experiments that would be difficult to do in real life.

In order to accurately perform one of these simulations, the computer model needs to be very detailed and therefore very large. The simulation must run on groups of hundreds, or even thousands, of computers that work in concert to perform the experiment. This is known as parallel processing. When the simulation is finished, the scientist needs some way to transform the numbers or data generated by the computer into something that is easier for a human being to understand in order to evaluate the results. A branch of computer graphics known as scientific visualization provides various techniques to represent the simulation visually through one or more pictures.

Since commercial visualization software packages generally cannot handle data and models of this magnitude, researchers at the national laboratories and universities have had to write their own visualization software to view these models, typically also using parallel processing. Although the software is frequently based on previously published techniques, the need to rewrite basic visualization software (with only some minor differences) at each institution has kept visualization researchers from investing that time in developing new techniques.

With the open source model, software is shared and improved upon by the entire research community. Instead of rewriting functions that already exist in proprietary software packages, programmers can download the software from an Internet site and either change it or add new functionality to it without having to rewrite everything from scratch. The improved software can then be returned to the repository to be shared with others. In this manner, software contributed to an open source library builds a foundation upon which everyone can build – in effect becoming the giant on whose shoulders others can stand.

* (505) 845-7506, pjcross@sandia.gov

The Visualization Toolkit (VTK) is an open source library of scientific visualization software. It has been adopted by the Department of Energy (DOE) Tri-Labs (Sandia*, Los Alamos, and Lawrence Livermore National Laboratories) as the vehicle for sharing visualization software between the laboratories and with other research institutions. Since VTK did not originally perform parallel processing, this functionality has had to be added in order to deal with the types of data generated by the Tri-Labs.

A fundamental component of parallel processing is the software that decides how to divide up the model, or data, amongst the available computers. Researchers at Sandia National Laboratories have developed this functionality in D^3 (Distributed Data Decomposition) and contributed it to the VTK library. D^3 can be used to divide and distribute data prior to parallel execution of any method in the VTK library.

In designing D^3, there were two important considerations. First, each computer needs to have an equal amount of data so work on all the computers is completed at about the same time. Second, the software also needs to be able to divide the data up using various methods, such as slices versus cubes.

D^3 uses the spatial location of parts to group them. It is based on a subdivision method known as a k-d tree. A two-dimensional data set consisting of points and the associated k-d tree are shown in Figure 1. The k-d tree is built by splitting the data into two groups of equal size in the horizontal direction. Then each half is split into two equal groups in the vertical direction. The subdivision continues, alternating between horizontal and vertical divisions, until the size of each group reaches a predetermined level. In Figure 2, the results of performing a similar subdivision in three-dimensions is shown, with the data that is allocated to one particular computer shown in red.

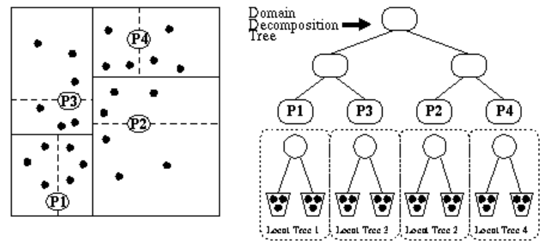

Figure 1. Two-dimensional point set and associated subdivision into a k-d tree.

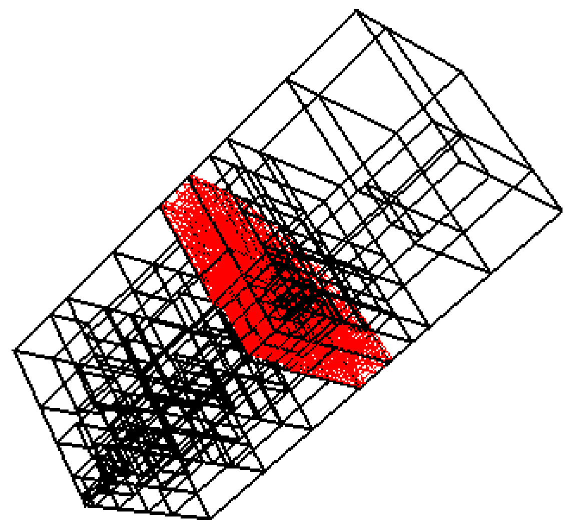

Figure 2. Three-dimensional data set with subdivisions shown as black outlines and one computer's portion of the data in red.

Combining D3, VTK, and earlier work using graphics cards from videogames to speedup visualization methods, Sandia researchers are currently building an interactive viewer for very large simulation data. Scientists will greatly benefit from being able to dynamically experience and explore their simulation results.

For further information on this subject contact:
Dr. Fred Johnson, Program Manager
Mathematical, Information, and Computational Sciences Division
Office of Advanced Scientific Computing Research
Phone: 301-903-3601
fjohnson@er.doe.gov

*Sandia is a multi-program laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.