Scientists, Engineers See Beowulf in Their Future

ASCI teams with CEI to map future of visualization computing

by Anna Turnage

A computer room in Apex, N.C., houses a study of contrasts: Against one wall is a sleek purple SGI Onyx 2 supercomputer larger than an average-sized refrigerator. Next to it is a rack of 16 plain-vanilla PCs connected by Ethernet cables.

Despite their differences, both systems are important to scientists working on the Department of Energy's Accelerated Strategic Computing Initiative (ASCI). ASCI officials have contracted with Apex-based CEI to scale its EnSight Gold software to visualize datasets containing billions of cells. The Onyx 2 is for a $1.8-million contract with Los Alamos National Labs. The PCs, linked in what is known as a Beowulf cluster, are being used for a $1.4-million contract with Sandia, Los Alamos and Lawrence Livermore National Labs.

Visualization Beyond Imagination

In 1996, the Comprehensive Test Ban Treaty was signed, placing a moratorium on underground nuclear testing. But the Department of Energy and its national laboratories are still responsible for guaranteeing the safety and reliability of the nation's nuclear stockpile. ASCI was formed to develop the high-resolution, three-dimensional physics modeling needed to evaluate the aging stockpile and accurately predict how time will affect different weapon components.

CEI's Beowulf Cluster

ASCI scientists are faced with the challenge of visualizing computational models of a size that could barely be imagined a few years ago. The animated simulations reveal exactly how individual atoms and molecules interact with each other in the bombs. The fusion reactions studied by ASCI can occur in just three-billionths of a second.

"Since we can no longer conduct live testing, these computer simulations must be a much higher resolution than what we have dealt with in the past," says Jeff Jortner, principal member of the technical staff at Sandia National Labs. "There are a lot of variables that go into viewing these large datasets. We must have a finer fidelity of the meshes and there are more time steps involved. A large part of our research is to develop supercomputers and software that can handle these datasets."

Higher Stakes

Jortner says CEI and its EnSight Gold software were chosen for these studies based on a track record for being able to visualize large datasets. Last August, EnSight Gold was used at Los Alamos National Labs to generate an image for a computational test model containing 11.5-billion cells. ASCI researchers use EnSight Gold to visualize a number of components within a data set, including scalar fields, vector fields, cell-centered variables, vertex-centered variables, and polygon information.

But as the research continues, the stakes get higher. Before long, scientists will need software that can visualize a terabyte or more of data.

The SGI Onyx Supercomputer

"A model on the terabyte scale is incomprehensible unless it can be visualized, and the only way to fully understand complicated 3D calculations is through sophisticated graphics software," says Mike Krogh, senior developer at CEI. Both ASCI contracts call on CEI to add features that will improve the speed and efficiency of handling massive amounts of data for rendering on normal displays and in VR environments such as CAVEs and PowerWalls.

One Goal, Two Paths

Although the goals of the two contracts are the same, they require CEI to scale the software on two different computing tracks - traditional shared-memory processing (SMP) such as that used on the SGI Onyx 2, and the Beowulf cluster. The research will help determine the most cost-effective and efficient way to meet scientists' visualization needs.

The Los Alamos research requires an Onyx 2 with eight 195-MHz MIPS R10000 processors, 10-GB RAM, 500-GB RAIDed disk and three InfiniteReality 3 graphics engines.

The Tri-Labs (Sandia, Los Alamos and Lawrence Livermore) contract focuses on parallel computing using commodity-based PCs and associated components. CEI's Beowulf cluster is connected together with 100-Mbit Ethernet and an HP ProCurve 4000 switch. It runs on RedHat 6.2 Linux software. The rack of PCs contains:

-- One 733-MHz Pentium III processor in each computer

-- 1-GB RAM in each of the 15 nodes and a root node with 2-GB of RAM

-- 40-GB RAIDed disk in each of the 15 nodes and a 55-GB disk in the root node

-- One NVIDIA GeForce2 graphics card in each computer, each of which can render about 20-million triangles per second

David Vs. Goliath?

The name Beowulf cluster was coined by two NASA Goddard Space Flight Center scientists after the sixth-century hero who freed the Danes of Heorot by destroying the oppressive monster Grendel. When linked together correctly, Beowulf clusters can reach high rates of computation speed comparable to a traditional supercomputer at a much lower cost. CEI's cluster cost $35,000 compared to the $400,000 SGI system.

Not only are Beowulf clusters less expensive to buy and maintain, they offer greater flexibility for increasing functionality, according to Jortner. "With a cluster, you can add a new machine, more memory or more processors and it's much less expensive than the traditional supercomputers of the past." That's why the majority of computers at the national labs are now clusters of some sort. Some fill entire rooms, such as IBM's ASCI White, which is capable of running at speeds of 12.3 teraflops per second. Other cluster-based systems include Los Alamos' Avalon, a 140-processor Beowulf cluster, and Loki, a 16-processor Beowulf.

"Computational codes are all being moved to clusters," says Jortner. "Right now they're no longer building single machines that can do computations of this size, so we are constantly looking for more power. But that doesn't always mean it has to be more expensive."

A Visualization Challenge

While the idea of linking commodity PCs together to emulate a visualization supercomputer sounds obvious, the programming involved is not. A cluster is a distributed memory system, which is more difficult to program for parallel processing than a shared-memory system such as the Onyx 2. If not programmed properly, a cluster can be tediously slow, eliminating any of its benefits.

Traditionally, scaling visualization software with a Beowulf cluster required fragmenting the simulation data across the individual PCs. The separate pieces of the solution are then "glued" back together into a datafile, then postprocessed the traditional way, as it would be with a supercomputer.

The problem is that as the aggregate model size increases, the combined results file becomes too large and unwieldy to process on a single system. CEI's goal is to program its EnSight Gold software to have the intelligence to find these individual pieces and spawn multiple functions. If the user wants to take a slice through the model, for example, the code would have to issue a command to each individual PC to do its part of this slice, then collect all the data and display it on the user's monitor. All this must be transparent to the user, as if he or she is dealing with a single fileset on a single computer. And it must be done as fast, or faster, than it would be on a supercomputer.

In addition to programming challenges, a Beowulf cluster can entail increased labor costs, according to CEI's Krogh. "The significant cost in a cluster is the human cost to assemble it and get it running. It's the same as managing any network of computers, but if you don't have the staff to handle it, it could be costly."

If programmed and maintained properly, the price/performance ratio of a cluster can be at least a magnitude better than that for a supercomputer, according to Krogh. PC-based systems are built using cheap commodity components costing tens to hundreds of dollars, significantly lowering maintenance costs. Most PCs come with a one-to-three-year warranty on parts and they can be fixed by someone with a little experience. A traditional supercomputer, on the other hand, requires an expensive maintenance contract with a vendor field engineer to perform the repairs. Individual parts can costs thousands or tens of thousands of dollars.

Supercomputing for the Future

The cost-saving factor of Beowulf clusters is beginning to draw the attention of the mainstream world. At first their use was limited to universities and government labs for large-scale biological research, engineering, astrophysics, meteorology and projects such as ASCI. But now many major corporations, such as Amerada Hess Corp. and Conoco, are beginning to use them as well. And large computer companies, such as IBM, SGI and Hewlett-Packard, are working to offer clusters in the mainstream computing community. There are, in fact, individuals who are using clusters at home since they are relatively easy to build. There are now numerous websites, books and articles on how to put one together.

"In my opinion, the only thing slowing cluster ubiquity is the shortage of application software," Krogh says. "But this is starting to improve." Krogh points out that software developers get an added bonus when developing code for distributed computing environments: the software will most likely run better on traditional systems, which at heart are distributed memory systems whose inner workings are largely hidden from users.

"It's more work to write software for a PC cluster," says Krogh. "But if you write it well, it will also run well on Onyx and other graphics workstation clusters."

If CEI's cluster is successful, it will be a big step not just for the ASCI project, but for visualization technology in all areas of research and development, says CEI president Kent Misegades.

"This project will help advance state-of-the-art, large-scale visualization," he says. "In particular, it enables us to adapt our technology to advanced parallel computing architectures, which we believe will be the systems of choice in the future for scientists and engineers. It will mean major innovations for all types of research and disciplines that have a need for supercomputing power."

###

Anna Turnage is a writer specializing in computer graphics and other technology issues. She works for Cramblitt & Company in Cary, N.C.