SCIENCE/TECHNOLOGY
January 10, 2000
Volume 78, Number 2
CENEAR 78 2 pp. 27-31
ISSN 0009-2347

[Previous Story][Next Story]

CRUNCHING DATA WITH CLUSTERS

Low-cost Beowulf-class computer systems can perform as well as supercomputers

Elizabeth K. Wilson

C&EN West Coast News Bureau

You're a chemist with a meaty computational problem to solve. But you can't get any time on the already overloaded supercomputing facilities at your university, and you certainly can't afford to buy a supercomputer on your limited budget. What do you do?

A growing number of scientists say they have the answer: Gather together a few inexpensive personal computers, hook them together, and voila!-you have your very own supercomputer.

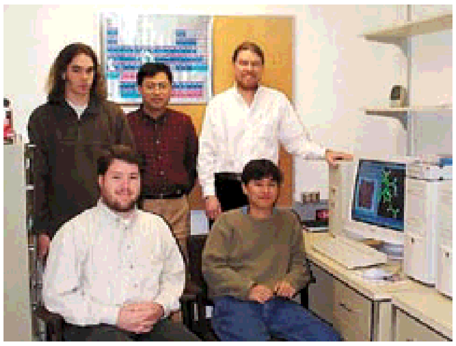
Known popularly as Beowulf-class clusters, these groups of linked, off-the-shelf computers are taking the scientific world by storm. Costing mere tens of thousands of dollars, computer clusters are capable of performing high-speed, complex calculations and simulations, frequently on par or even better than their more expensive supercomputing cousins.

Above, Rochester's Dellago (left) and coworker Thjis Vlugt in front of their 48-node cluster. Seated at left, Cincinnati graduate students Matthew Hafley (left) and Nobu Matsuno and, standing from left, graduate student Jason Clohecy, Director of Molecular Modeling Anping Liu, and Beck show off their Linux-based cluster.

"This is something a poor assistant professor like me can afford," says Cristoph Dellago, a theoretical chemist at the University of Rochester, whose group's recently constructed cluster cost $55,000.

What's more, when you build a cluster, it's all yours. "One of the advantages is that you don't have to share it with a million other groups, and you can manage it as you want to," Dellago says. His group uses its cluster for complex systems dynamics, such as water dissociation simulations.

Tom Ziegler, a chemistry professor at the University of Calgary, Alberta, whose group developed a cluster dubbed COBALT (Computers On Benches All Linked Together), echoes the sentiment: "If you want to rotate the Milky Way or solve something with proteins, you can build a cluster around the problem and optimize it for that-and cut away all the administration."



Seated at left, Cincinnati graduate students Matthew Hafley (left) and Nobu Matsuno and, standing from left, graduate student Jason Clohecy, Director of Molecular Modeling Anping Liu, and Beck show off their Linux-based cluster.

Beowulf clusters aren't even 10 years old, so it's not surprising that, initially, only groups staffed with computer whizzes attempted to build them. With little support or documentation to assist them, scientists frequently spent months of intensive effort constructing a cluster. But as the technology's popularity is growing, there's now a wealth of resources available for the nonexpert. Books, web pages, conferences, and software devoted to facilitating cluster-building are helping to draw more and more researchers into the fold.

"Even four-year colleges can build these clusters-even high schools can build them," says Thomas L. Beck, a chemistry professor at the University of Cincinnati, whose group put together a $25,000 cluster. "I think that's an exciting development."

Most bigger research computing facilities, such as the Ohio Supercomputer Center in Columbus; Pacific Northwest National Laboratory in Richland, Wash.; the National Energy Research Scientific Computing Center at Lawrence Berkeley National Laboratory, Berkeley, Calif.; as well as many university computing centers also have jumped on the bandwagon, building cluster systems to include in their machine lineups.

Several ultra-high-powered "self-made" clusters have made a Top 500 supercomputer site list, including the CPlant cluster at Sandia National Laboratories, Albuquerque (number 44 on the list), and the Avalon cluster at Los Alamos National Laboratory in New Mexico (number 265).

Even the companies that manufacture high-end supercomputers, such as IBM and Silicon

Graphics, are now constructing Beowulf-style systems.

At Supercomputing 99, a trade show for supercomputer vendors and technologists held last November in Portland, Ore., clusters were the center of attention, according to numerous scientists.

Not bad for a concept that originated in 1993. Back then, computer scientist Thomas Sterling at the National Aeronautics & Space Administration's Goddard Space Center in Greenbelt, Md., was challenged to come up with a low-cost yet high-powered computer system that could be made available to individual researchers.

Sterling, now a faculty associate at California Institute of Technology's Center for Advanced Computing Research, says a confluence of ideas and events led to Beowulf. One factor was the then-recent emergence of the Linux operating system. Noted for its extraordinary stability, Linux is a close relative of the UNIX operating system. Whereas UNIX was developed for workstations and mainframe computers and costs thousands of dollars, Linux was developed for PCs and is free.

Sidebar: Linux Software for the people?

Sterling saw that it might be possible to string groups of PCs together. He had colleagues who were able to develop software to enable the computers to communicate with each other. The only thing left was to come up with a name for the system, which they had come to view as a triumphant underdog.

"I was looking for a David and Goliath metaphor," Sterling says. "But I didn't think we'd ever get famous with a computer system named David."

Sterling: developing a triumphant underdog

Then he remembered Beowulf, hero of the medieval epic poem, who slew the monster Grendel. The name and the concept caught on quickly. In just a few years, clusters of all sizes began appearing, ranging from just a few PCs hooked together to hundreds of high-powered workstations.

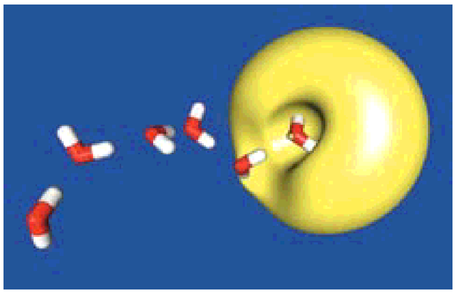
In addition to their low cost, clusters are flexible: Processors can be added or subtracted without much trouble. Many of the clusters are based on Linux, although some clusters are now being designed to work on Windows NT.

However, Beowulf clusters, scientists who use them say, are not a simple substitute for supercomputers and, indeed, are not likely to supplant them. There are a number of technological issues that cluster designers must contend with.

For example, the goal computer scientists strive to achieve in systems that involve any sort of parallel processing-be it a Beowulf cluster or a prefabricated supercomputer-is scalability. This refers to the ability to get out in processing speed what you put in processors. A computer with 10 processors that scales ideally would compute a result 10 times as fast as a computer with one processor.

But scaling becomes more and more difficult as the number of processors increases. "Only in special applications do you actually get true scalability," says Mark S. Gordon, a theoretical chemistry professor at Iowa State University, Ames, and director of the Scalable Computing Laboratory (SCL) at the Department of Energy’s Ames Laboratory.

There are several reasons for scalability problems. A significant bottleneck in limiting performance is the time it takes to communicate between the groups of one or more processors in a single box, known as nodes. Passing information back and forth between nodes contributes to what’s known as overhead. "That’s taking away from scalability," Gordon explains.



Calculations performed on Jordan’s computer cluster show the linear structure of an anionic water cluster, with excess electron charge distribution (yellow) bound at the end.

Supercomputers have hundreds of processors right next to each other that can communicate extremely quickly, whereas cluster nodes aren’t so easily connected. Frequently, cluster builders use what’s known as fast Ethernet with the connections being handled by a fast Ethernet switch, which transmits data between nodes at a rate of 100 megabits per second. Even faster communication is possible with gigabit Ethernet, but this is much more expensive.

Other factors that affect cluster performance include latency, which is the time lag before sent information actually transmits, and bandwidth, which determines the amount of data that can be sent at one time.

Clusters also need software to optimize the data exchange, such as the popular message passing protocol MPI or systems known as Linda, PVM, or TCGMSG.

Because switching between nodes eats up valuable time, certain types of problems lend themselves better than others to running on parallel computer systems. Those that work best require less communication between nodes during calculations. Fluid dynamics calculations, for example, scale quite well, whereas weather forecasting can be very difficult.

"There are applications where you have to work really hard [to get them to scale], and one is quantum chemistry," notes Gordon, who with his colleagues has developed the academic ab initio quantum chemistry program GAMESS. Over the past 10 years, however, a number of computational chemistry groups have made great inroads on parallelizing their codes.

In general, Hartree-Fock and density functional theory-the basic electronic structure theories-are fairly well parallelized, as is second-order perturbation theory. More sophisticated calculations-using coupled cluster theory, for example-are more difficult. Companies such as Gaussian and Q-Chem offer parallelized versions of their software that can run on clusters. Other codes that have been parallelized include the academic macromolecular simulations program CHARMM.

Halstead (left) and Gordon develop clusters at Ames Laboratory, such as the 64-node ALICE cluster at right. The group is now working on even bigger clusters.

One ab initio quantum chemistry program, known as NWCHEM, is at its core a parallelized program. When theoretical chemists at Pacific Northwest National Laboratory began writing the program years ago, they designed it specifically to run on massively parallel platforms.

Other chemistry groups are joining computer scientists in pushing the boundaries of cluster technology, frequently with grants from government agencies and companies. At the University of Pittsburgh, chemistry professor Kenneth D. Jordan and his colleagues in chemistry and chemical engineering have developed a high-powered cluster consisting of 25 dual-processor IBM 43P 260 workstations, 16 of which are connected by a gigabit switch.

With a cost of about $900,000, the cluster is more expensive than many. But with 20 gigabytes of memory and 250 gigabytes of high-speed disk space, it’s "ideally suited for applications with large memory and disk demands," Jordan says. The National Science Foundation and IBM funded the cluster, which serves the university’s new Laboratory for Molecular & Materials Simulations.

Pulay [Photo by Elizabeth Wilson]

Ziegler’s COBALT system, which cost about $235,000, is a cluster of 94 Compaq Alpha Personal Workstations, funded by the Canadian Foundation for Innovation. The cluster’s performance, Ziegler says, is 1 gigaflops (1 billion calculations per second) per node. About 20 scientists use COBALT, which runs quantum chemistry programs such as ADF (Amsterdam Density Functional) and the ab initio molecular dynamics code PAW (Projector Augmented Plane Waves).

Sidebar: Computer cluster and Linux-related web sites

Gordon, SCL theoretical chemist David M. Halstead, and their colleagues at Ames developed a cluster they call ALICE (Ames Lab/ISU Cluster Environment), a 64-node system of PCs, which cost about $300,000. It is four times as fast as the largest commercial supercomputer at SCL, for a third of the price.

Peter Pulay, a chemistry professor at the University of Arkansas, Fayetteville, and colleagues have even gone into the cluster business. After developing their own small cluster several years ago, they were so pleased with the results that they started a company called Parallel Quantum Solutions that markets systems with four to eight Pentium III processors that are geared specifically toward chemists.

Rochester’s Dellago says his 48-node cluster of AMD processors performs better than a supercomputer in terms of speed per dollar by a factor of five to 10. With his collaborators, he’s studying such problems as proton transfer in small water clusters and the diffusion of alkanes in zeolites.

And Beck uses his small cluster of four dual-processor Pentium machines to help develop a method to improve the scaling of density functional theory. He’s also collaborating with other groups, running simulations of proteins and carrying out quantum mechanical studies of transition-metal atoms in the active sites of enzymes.

Clusters, Beck says, "allow us to study much larger systems than we could before."

[Previous Story][Next Story]

Chemical & Engineering News