

<http://www.latimes.com/news/printedition/front/la-071201distrib.story>

COLUMN ONE

Tackling Puzzles PC by PC

Millions volunteer their home and office computers to crunch data over the Net to aid scientific and medical studies. For some, it's a heated competition.

By KAREN KAPLAN
Times Staff Writer

July 12 2001

Inside a country inn at the base of the Pocono Mountains in Pennsylvania, a handful of home computers belonging to innkeeper Michael Kelly are whirring away on one of the gargantuan problems of modern medicine.

Kelly's five PCs are using their excess computing power to analyze hundreds of millions of genes to understand their role in diseases such as AIDS and cancer. On a good day, he can plod through eight or nine genes--a rate that would allow him to finish in a few eons.

But joining Kelly's humble computers are 10,000 other PCs scattered around the globe that have been woven together over the Internet into a makeshift supercomputer. Like an army of ants gnawing through an elephant's carcass, they are methodically testing every potential gene. They expect to be done in about two years.

Kelly is part of the most ambitious grass-roots movement in technology: a mass-mobilization of PCs known as distributed computing. Its goal is to harness the power of tens of millions of desktop PCs to create virtual supercomputers capable of attacking some of the most brutal problems in modern computation.

"We have the equivalent of a million Cray supercomputers sitting idle when they could be put to work alleviating suffering," Kelly said.

The process involves dividing a problem into petite pieces and distributing them to as many individual computers as possible. After each PC works on its chunk of the problem, the results are combined into one giant solution.

There have been scattered distributed computing projects in the past, but the spread of the Internet and the increasing power of personal computers have sparked a renaissance in the field. Researchers are using the strategy to battle disease, forecast global climate changes and predict volatility in the stock market.

Swedish radiation scientist Peter Jansson, for example, has been masterminding billions of calculations to measure gamma radiation leakage from a skinny 14½-foot-long, stainless steel canister designed to hold radioactive waste. When he knows the answer, he'll know how much waste each canister can safely hold.

Jansson's computer simulation measures radiation at tens of millions of points around the canister, each of which is bombarded with several hundred thousand gamma rays. The calculations would take one PC about 50,000 years to complete.

After running his experiment for two years with 20,000 volunteers, the number crunching for the canister is nearly finished. "The calculations are easy, but there are a lot of them," Jansson said.

Distributed Computing Idea Emerged in '70s

Distributed computing is still a somewhat esoteric movement, with its participants numbering just a few million out of the hundreds of millions of Internet users who could join.

But it has swept across the usually anarchic world of the Internet in a remarkably short time, spawning an odd techno culture that is part philanthropy, part high-tech environmentalism and part video-game-style shoot 'em up over who is the biggest, baddest number cruncher in cyberspace.

Groups of users have formed into fiercely competitive computing brigades that have become the geekiest manifestation of team sports on the Internet.

Henry Staples, an Arlington, Va., Web developer, became so obsessed with distributed computing that he gathered a dozen castoff PCs to increase his productivity ranking on an array of distributed projects.

"You see your name, and you say, 'I want my name on the top,' " said Staples, a volunteer for a project on stock market volatility and another on how proteins transform themselves from strings of amino acids into three-dimensional shapes.

The idea behind distributed computing has been around since the earliest days of network computing. The idea was hatched in the 1970s at Xerox Corp.'s Palo Alto Research Center, where networking scientists created a program called a "worm" that roamed among the center's 200 machines in search of idle computers that could be put to use.

The project turned out to be a disaster. A test program run by the worm ended up crashing computers. As each machine crashed, the worm spread to a new host, crashing it as well.

Today, most people have forgotten the experiment as the beginning of distributed computing and instead remember it as the first instance of a destructive, self-propagating computer virus--the forerunner of last year's crippling "Love Bug."

The idea went into hibernation but was revived in the mid-1990s as the Internet became a public phenomenon, making millions of PCs accessible through a single network.

Researchers and math gurus launched a number of prime-number searches and code-cracking efforts that were tailor-made for the method because they involved an unthinkable number of relatively simple calculations.

What eventually brought distributed computing to the masses was a quixotic project named SETI@home. The UC Berkeley-led effort assists in the Search for Extraterrestrial Intelligence program by letting individual computer users crunch radio telescope data from the Arecibo Observatory in Puerto Rico in search of any unusual signals that might be characteristic of an advanced civilization.

Though most serious scientists doubt that the searchers will find anything, the prospect of intercepting alien communications is so tantalizing that more than 3 million volunteers have downloaded SETI@home's data-crunching screen saver in the two years it has been available.

Now, with desktop computers approaching the power of old mainframes, researchers from a variety of fields are launching SETI-like projects with more practical goals.

At Stanford University, chemistry professor Vijay Pande is spearheading the Genome@home project that occupies Kelly's computers in the Pocono Mountains. With the help of 18,000 volunteers, he also is running a related project that simulates, nanosecond by nanosecond, the behavior of proteins as they fold into three-dimensional shapes.

Each day, a volunteer's computer can calculate about one-billionth of a second in the progression of a protein. It takes about three weeks with all 18,000 computers blazing away to get some usable results. Then they start over with another set of proteins.

Myles Allen, a climate dynamics researcher at the Rutherford Appleton Laboratory in Oxfordshire, England, is attempting to predict global weather in 2050.

This summer, Allen will dole out climate prediction models to more than 17,000 volunteers. It will take each of them at least 200 days to churn through one model on a desktop PC.

But distributed computing will allow him to simulate thousands of scenarios and evaluate the effects of more than 100 parameters, such as temperature, on global climate. On a supercomputer, he could afford to run through only about a dozen scenarios.

Such research efforts already have begun to bear fruit. A group of 95,000 volunteers have won several encryption contests and discovered new prime numbers. Stanford University's Pande has published two articles in scientific journals based on results from his ongoing experiment studying proteins. At Scripps Research Institute in La Jolla, a project called FightAIDS@home has sped up the process of pre-screening potential HIV protease-inhibitor drugs by a factor of 1,000.

But for all the practical benefits, what has really propelled the movement is that--in an era of increased online pornography, EBay fraud and junk e-mail--it has tapped into a spirit of altruism and scientific exploration that hark back to the pure, early days of the Internet.

"I lost both parents to cancer, and I saw my participation as a way of 'doing my bit' to combat this and many other terrible diseases," said Steve Taylor, a volunteer for Genome@home, the gene-synthesizing experiment at Stanford.

Taylor built six PCs for the project. "I find this far more satisfying than simply putting money in a charity envelope and never knowing where it goes."

The beauty of his technological altruism is that it costs almost nothing.

Commercial Interests Enter the Picture

To participate in a project, ordinary computer users volunteer to download a small program that performs the calculations and coordinates communication with the servers that send out blocks of data to be processed.

The programs are designed to use only surplus computer cycles. That means they go faster when a computer is sitting idle and slower when the PC is downloading a movie from the Internet.

"A computer cycle isn't something you can store or reclaim," said David McNett, president and co-founder of Distributed.net, a 4-year-old nonprofit research foundation dedicated to pushing the limits of distributed computing. "You have to make use of it while it happens. If you can put it to good use, that's a very compelling concept. I wish we could do that with all of our resources."

But just like the Internet's gradual migration from ivory towers to the marketplace, commercial interests have stumbled on the promise of distributed computing--and have stirred controversy in the process.

Free Internet service provider Juno Online Services recently roiled some of its customers by reserving the right to require them to leave their computers on around the clock so the cash-strapped company could sell huge amounts of processing power to research institutions and others.

A handful of start-up companies are building their own PC networks and are trying to lure volunteers by entering them in prize lotteries and offering them cash.

Parabon Computation Inc. in Fairfax, Va., plans to pay people based on how much computing power they contribute to the network. A high-end workstation might bring in \$500 a year, said Chief Executive Steven Armentrout.

But for most volunteers, it's not about money. It's about results.

Database administrator Michael Hotek was frustrated by his inability to move up in the SETI@home rankings, a listing of participants according to how much work they contribute to the project. The rankings have no more scientific value than the enumeration of high scorers displayed on arcade video games, but for participants they are a major source of pride. Hotek decided to quit the SETI@home project and instead joined two newer projects that offered a better chance of upward mobility.

He joined a team working on FightAIDS@home and broke into the top 100. Then he joined the Cancer Research Project sponsored by Intel Corp. and United Devices Inc., a distributed computing firm in Austin, Texas, and rose as high as eighth place.

"If I find something, I find something, but I was much more interested in getting points," said Hotek, who has four black belts in martial arts and a considerable competitive streak.

Hotek eventually burned out on the competitive aspect, but he has kept his machine running. "My machines stay on 24 hours a day," said Hotek, who lives in Columbus, Ohio. "If I'm not using them, I might as well have them doing something useful."

The competition has sharpened to the point that some of the projects have begun to look like the

Net's version of the USC-UCLA rivalry. Competing groups, such as Wicked Old Atheists, Czech Republic and Beer, have been fighting tooth and nail for their piece of computing glory.

Competition Builds Among Altruistic Teams

Kelly, the innkeeper, is a member of the Overclockers Network Hellspawns team. This ragtag bunch, whose approximately 65 members range from technology professionals to high school students, recently passed a rival team named for Star Trek character Capt. Jean-Luc Picard to take third place in the Stanford gene-synthesizing project.

Now they are closing in on an Italian team, which is more than 2,000 genes ahead. It would take the Overclockers about five days to process that many.

Kelly says his team is "pulling machines out of the woodwork" in an effort to catch up. Members are wrangling valuable computing cycles out of machines in their offices, and a handful of Internet cafes have been drafted for the effort. One member of the team is writing a SETI@home-style screen saver program to attract support from a more low-tech crowd.

Though the Italians have been able to maintain their lead, Kelly is confident of catching them.

That would leave only Ars Technica Team Primordial Soup, the first-place team that is about 12,400 genes ahead of the Overclockers. But he acknowledges that chances of catching them are slim.

"They are like monsters," Kelly said. "They have hundreds and hundreds of members, far more than we could ever touch."

Copyright 2001, Los Angeles Times