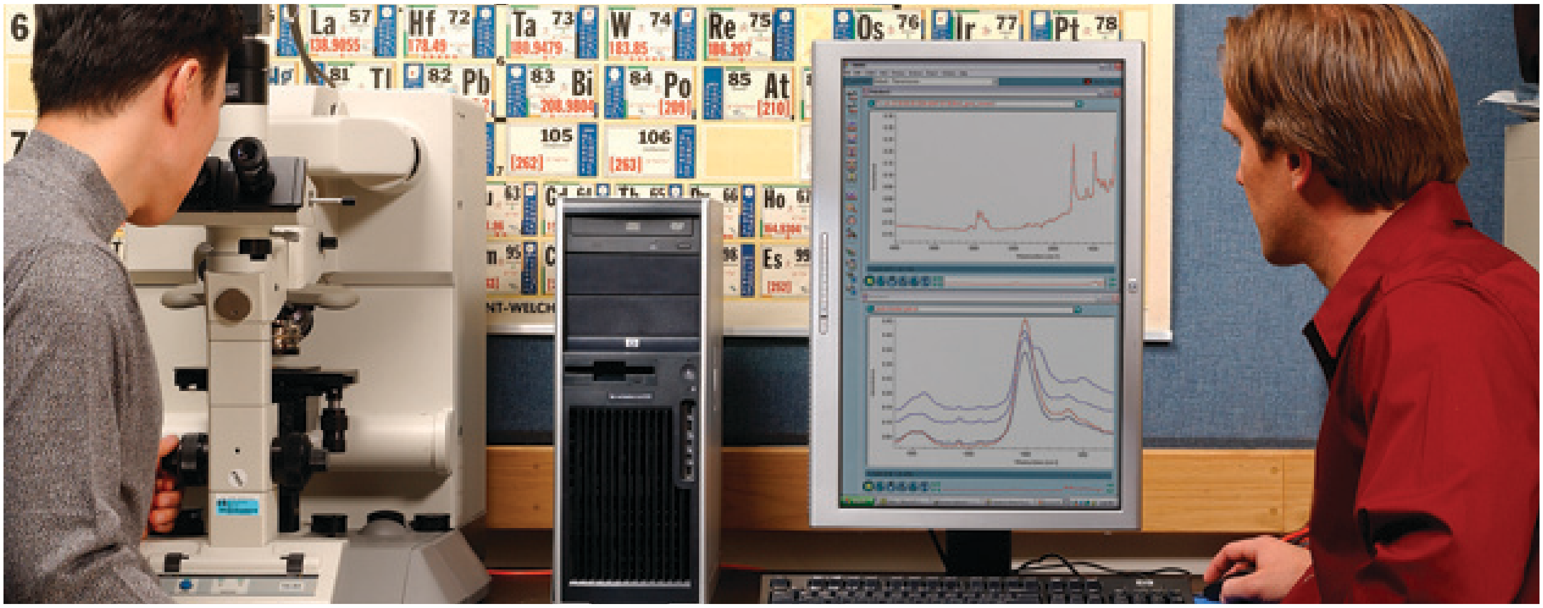

Compute clusters and parallel programming suites deliver supercomputing capabilities once reserved for the privileged few.

Do-it-yourself supercomputing

Financial modeling, computational fluid dynamics, molecular modeling, computational chemistry, structural stress analysis and other tasks requiring supercomputing assistance have in the past necessitated two types of individuals: the researchers, engineers and scientists who understand the science and comprehend the problems and dynamics of these complex application areas; and, the technology experts and programmers who must implement the requirement digitally, and analyze and test them in the computer realm.

According to Dan Cox, manager of High Performance Computing Programs for the Industry Standard Servers group at HP, this dynamic is changing. "Several developments are enabling the research world at-large in some very exciting ways," he notes. "Thanks to affordable computer clusters and new parallel programming suites, small companies, universities and other emergent research organizations now have access to supercomputing capabilities that once were reserved for the most prominent enterprises, governments and education institutions."

The possibility of what could be termed "do-it-yourself supercomputing" was born from the creation of dual core CPU chips, which enabled the aggregation, or clustering, of several computing systems based on multiprocessor architectures. These clusters deliver supercomputing capabilities at a much lower cost and complexity than traditional, standalone systems.

"Industry-standard compute clusters now can be purchased affordably and easily placed in a cubicle, delivering the same performance that previously necessitated a multimillion dollar supercomputer that filled a large room," exclaims Cox.

Affordable hardware clusters meet only half the challenge, however. The multiprocessor architectures on which compute clusters operate require parallel programming approaches, such as multithreading, openMP and message passing interface (MPI). According to Cox, parallel programming using these current industry standard approaches is difficult at best for the engineer, researcher and scientist who have not had to deal with this technology before. "There is a shortage of trained personnel who can effectively develop, maintain and enhance parallel applications.

invent

Even the experts are not very productive in these tasks," says Cox. Furthermore, software developers seek to create parallel multiprocessor applications that can run efficiently on all of the multiprocessor platforms.

"Creating an effective parallel version for even one platform is extremely difficult," says Martin Schultz, founder of Scientific Computing Associates Inc. (SCAI), an organization that pioneers the commercial use of parallel and distributed computing. "Creating parallel versions for multiple platforms is a daunting task that few developers are prepared to face." Which leads us back to the reliance of researchers, engineers and scientists on the select few technology experts and specialist programmers. In what has become a significant bottleneck, the productivity, efficiency and sheer availability of parallel application developers has threatened to stall the do-it yourself supercomputing movement.

Auspiciously, a new parallel programming approach resolves this dilemma and boosts the viability of commodity multiprocessor clusters. It's based on the generalization of the current fashion of processor virtualization, Schultz indicates, but goes one step further. Instead of the virtualization of a single CPU, this new approach considers the virtualization of entire multiprocessor systems. "As an alternative to using different parallel programming approaches to different architectures and being forced to maintain and enhance four 'parallelized' versions of each application, the developer only has to create a parallel code for one architecture," explains Schultz. "In other words, we create a single parallel source code, and when we wish to retarget the application to different parallel architectures, we recompile the code with a compiler for that architecture." Representing this new approach to parallel programming, SCAI's Linda Virtual Machine software suite (LindaVM) is widely used by scientists, researchers and non-programming professionals for parallel application development and processing. HP and SCAI are collaborating to deliver LindaVM tools on HP cluster platform solutions. The result is cost-effective supercomputing with greatly simplified parallel programming, making it possible for most developers, scientists and researchers to succeed, regardless of budget or programming expertise.

"Industry standard clusters with SCAI software tools provide an excellent alternative to difficult and time consuming message-passing programming for parallel and distributed computing projects," Schultz continues. "They save both programming time and hardware costs when working with applications that previously were designated only for supercomputer class machines. Developers and scientists are able to focus on their code and computational solutions — not expensive and complex systems integration efforts."

"Smaller entities, enterprises and research organizations now have access to supercomputing capabilities for the first time," says Cox. "Bringing supercomputing to the masses could very well lead to an explosion of research, learning and knowledge in much the same way that personal computing empowered the everyday productivity worker and revolutionized the ways in which we utilize and benefit from IT."

This article first appeared in the March 2006 High-Performance Computing Special Edition of HP's Transforming Your Enterprise publication. The current edition of this publication is available at: www.hp.com/go/transformHPC

To download a white paper on the HP Unified Cluster Portfolio,visit: www.hp.com/go/transformHPC

For more information on Scientific Computing Associates, visit: www.lindaspaces.com

4AA0-6633ENA, March 2006

invent