

ORNL's Path to Teraops Computing

In January 1995, ORNL had the fastest parallel supercomputer in the world—the Intel Paragon XP/S 150. It would run for less than a day at a time, but it was the only tool available for performing the most complex calculations of the day. By partnering with Intel, the staff of ORNL's Center for Computational Sciences (CCS) turned this and another Paragon into reliable supercomputers. These machines rapidly perform calculations to help scientists better understand complex phenomena, such as the movements of pollutants in groundwater and of particles in a gas, effects of growing atmospheric concentrations of carbon dioxide on future climate, and the behavior of materials ranging from melting to magnetism.

In the past four years, the reliability of the Intel Paragons has been strongly improved, stamping the CCS as a major Department of Energy (DOE) computational development center. However, largely because the speed of processor chips continues to double every 18 months (as predicted by Moore's law), the position of the Intel Paragon XP/S 150 in the hierarchy of powerful machines has slipped to #24. Our Intel Paragon XP/S



The SRC-6 is a machine of an innovative architecture being built by a company started by Seymour R. Cray. It will be installed at ORNL in 1998 for evaluation. *Model rendering by Ross Toedte.*

150, which once led the world in computing speed, can perform 150 billion calculations per second. But this multigigaops computer has been eclipsed by the superb Intel Teraflops machine at Sandia National Laboratories (SNL) in Albuquerque, which reliably provides a peak computing level near two teraops—trillion operations (arithmetic calculations) per second. The Intel Teraflops computer is used for classified defense work. SNL also has an Intel Paragon, which has been linked by a high-speed network data line to the ORNL Intel Paragons to leverage the ability of the three computers to solve complex problems in energy, environmental, and materials research.



New Strategy

Recent technological advances, the decision by Intel to halt production of new supercomputers, and the realization that our Intel Paragons have a limited lifetime and limited appeal to computational scientists looking for the fastest machines has compelled a new perspective. ORNL has decided to pursue a different path to achieving a teraops level of computing. This path will continue to offer challenging opportunities for CCS as a computational development center.

We have received DOE funding to purchase and install in 1998 a new parallel supercomputer built by SRC Computers, Inc., the company Seymour R. Cray (the inventor of the Cray supercomputer) was heading at the time of his death. Our plan is to evaluate comprehensively the performance of this machine, called the SRC-6. We hope to incorporate it into ORNL's collaboration with SNL on the networking of high-performance computers that are widely separated geographically. Anticipating top-quality performance for this supercomputer because of its innovative architecture, in 1999 we will write specifications for its multiteraops successor, called the SRC-7, if funded. In 2000, we will install an SRC-7 system in stages.

Such a powerful computing system will be needed at ORNL to help our scientists and engineers solve difficult Grand Challenge problems. Computer modeling on a teraops computer will lead to a more detailed understanding of the structure and mechanical behavior of metals and alloys. As a result of this information, these materials can be made stronger and more resistant to the effects of aging, radiation, and stress-corrosion cracking. Modeling may also be used to study an alloy's magnetic properties that affect its behavior.

Faster computer models are needed to tease out the details of combustion in spark-ignited and diesel engines. The information could speed the development of lean-burn natural gas vehicles that will use less fuel, reduce U.S. dependence on foreign oil, and produce lower carbon dioxide emissions, slowing the buildup of greenhouse gases in the atmosphere. The effects on future climate of increasing atmospheric concentrations of greenhouse gases requires complex simulations that teraops computing could make possible.

Advances have been made in modeling the oceans and the atmosphere to pin down sources and sinks for greenhouse gases, but the influences of polar ice and land masses on future climate have not yet been factored in. Accurate models of the effects of pollution and greenhouse gases on future climate are needed to guide wise decisions by technical and political leaders.

How the chemical bases of long DNA strands containing genes are arranged, how genes produce proteins, how strings of amino acids fold into proteins, and how enzymes interact with protein receptors and with nucleic acids (DNA and RNA) are not well understood. Because of the huge amount of data from experimental approaches to these questions, sophisticated computational analysis is needed to make sense of this information. The more data there are, the more teraops computing will be needed to solve problems such as locating a particular gene and determining its structure and function, finding the set of genes that cause a fatal disease or unhealthy condition such as obesity, and designing a new drug.

Predicting the properties of plastics directly from their atomic structure has long been an elusive goal of polymer science, and it will continue to be. But calculations at the teraops rate will enable ORNL and University of Tennessee researchers to bridge from the results of atomistic calculations to those of calculations using approximate models that simulate polymer properties. The U.S. chemical industry has identified advanced computational modeling as a key technology that will enable molecular design of new plastics and other lighter, stronger materials of the future.

What features will enable the SRC supercomputers to perform more calculations at once and produce solutions faster than our Intel Paragons? For one thing, the central processing unit (CPU) chips that do computations will be much faster simply because they will be newer (remember Moore's law?). The Intel Paragons have Intel's older i860 chips, which aren't nearly as fast as the Pentium Pro chips in SNL's Intel Teraops computer. The SRC-7 machine will use the fastest CPU chips ever designed by Intel—the not-yet-released Intel Merced chips.

A second key improvement will be in the handling of computer memory. The multigigaops Intel Paragons in CCS have distributed memory—that is, each CPU has its own share of the total system memory. In this arrangement, any CPU that requires information stored in the memory of another must send a message requesting this information, which is then supplied by a response message through a sophisticated "message passing" system. The SRC supercomputers will have shared memory instead of distributed memory because all CPUs will be in direct contact with the entire machine memory. It will be easier for computer scientists to program a parallel machine with a shared memory than one with a message-passing distributed memory system. A shared memory will also allow faster retrieval of information.

A third feature will be the field programmable gate arrays (FPGAs). These FPGAs act as configurable special-purpose processors. They can execute multiple CPU instructions in one clock cycle. The reconfiguring of these FPGAs is easy, taking only 10 to 20 microseconds. They can be reprogrammed as often as desired.

SRC machines will have a fourth feature that will increase the speed of entering and retrieving information in a usable form (input-output, or I/O). SRC Computers, Inc., has placed an emphasis on the design of the I/O. By contrast, the Paragon XP/S 150 at ORNL has only one service processor for every eight computational processors. The improved processor ratio in the SRC machines should speed up I/O operations.



Evaluating Performance of the SRC-6

SRC machines are built from units called segments. Each segment is housed in a standard rack enclosure. Segments may be combined by linking the memory crossbar switches together to form larger shared-memory systems.

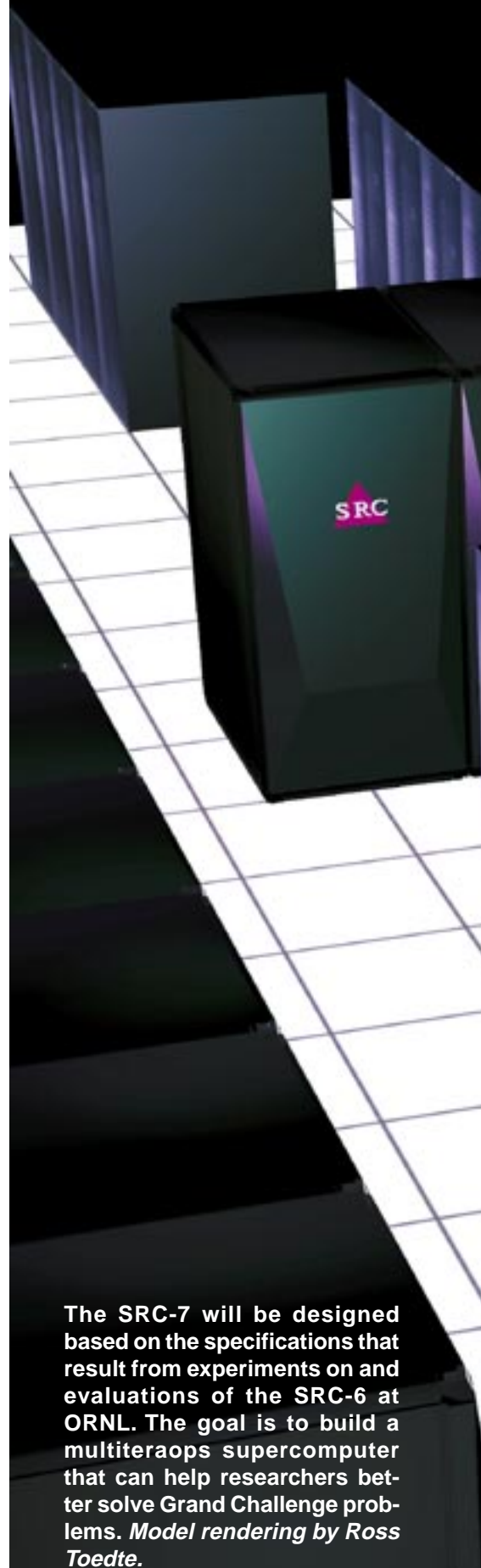
We will acquire two SRC-6 segments (see sketch on p. 20). We will evaluate their components and systems to determine strategies and mechanisms for interconnecting multiprocessor units. We will assess the use of FPGAs in various communication systems as well as options for programming frequently used application algorithms into FPGAs. This development effort by the CCS will be coordinated with other ORNL divisions and SRC.

CCS will initially evaluate the segment to determine if CPU performance, memory bandwidth, and I/O performance are properly balanced. We will then connect two segments using SRC's memory crossbar to form a larger, shared-memory machine. We will evaluate the new configuration, looking for weaknesses in the architecture that would prevent the machine from being scaled up to SRC's stated objective. In addition to testing the hardware, we will also test the scalability and performance of the system software and the user development environment provided by SRC. We will port a variety of our computer codes from the Paragon environment to the SRC system to see how well it performs on real codes.

Once the performance of a single node is well understood, we will split the machine back into two segments and explore a variety of strategies for providing intermachine communications. We intend to work with SRC to develop several memory-connected communications channels to the SRC memory subsystem in a quest for highest performance. If our evaluation of SRC-6 verifies the excellence of its performance, as expected, we will write specifications and acceptance plans for an SRC-7 in FY 1999.

It should be emphasized that a shared-memory architecture provides a programming environment that hosts both shared-memory and message-passing applications. Shared-memory applications will port to the SRC system with a minimum of code changes. However, a major difference between shared- and distributed-memory architectures is that implementing a message-passing programming model on a shared-memory system is straightforward and efficient if sufficient memory bandwidth and capacity are present. However, implementing a shared-memory programming model on a distributed-memory system is not very efficient. Because the DOE energy research community has many codes developed for both types of parallel programming models, it is important that any large system support both types of programming model. Thus, it is essential that a shared-memory architecture with hundreds of processors (e.g., the planned SRC-7) has a memory subsystem with enough extra bandwidth to handle the message-passing traffic, as well as the memory requirements of the processors. In our evaluation of the SRC-6, we will determine the requirements for the SRC-7 in this area.

A multiteraops computer possessing a shared-memory architecture has long been a high-performance computing goal. A properly designed SRC-7 should meet that goal. We are focused now on the path to the SRC-7, which involves a detailed technical appraisal of the intriguing architecture incorporated in the SRC-6. This work is sponsored by DOE's Office of Energy Research, Office of Computational and Technology, Information, and Computational Sciences. The CCS-ORNL team that made the Intel Paragons hum is most pleased that the initial step on this path has been taken and looks forward to the technical challenges of the comprehensive evaluation of the SRC-6. The journey should help ORNL reclaim its position near the top of the high-performance computing hierarchy by the start of the next millenium.—Based on information provided by Ken Kliewer, director of the CCS, and Arthur S. (Buddy) Bland, also of the CCS



The SRC-7 will be designed based on the specifications that result from experiments and evaluations of the SRC-6 at ORNL. The goal is to build a multiteraops supercomputer that can help researchers better solve Grand Challenge problems. Model rendering by Ross Toedte.

