# How Useful Are Today's Parallel Computers?

**David H. Bailey**
**NAS Applied Research Branch**
**NASA Ames Research Center**

**Ref: Computers in Physics, vol. 6, no. 2 (Mar./Apr. 1992), pg. 216.**

In the last few years, NASA Ames, along with many other large scientific laboratories, has recognized that conventional supercomputer technology is maturing, and that the multi-teraflops sustained performance that our scientists have requested by the year 2000 can only be provided by employing highly parallel computer systems. In response to these challenges, NASA Ames instituted a program in 1988 to acquire and test these systems.

NASA Ames presently operates a Connection Machine 2, with 1,024 64-bit floating point processors and four gigabytes of memory, and an Intel iPSC/860, with 128 i860 processors and one gigabyte of memory. Additional systems are to be acquired in 1992 and 1993. During the past few years, scientists at both NASA Ames and other NASA centers have been exploring the usability of these systems for important aerophysics applications. Some full scale state-of-the-art applications have been successfully ported to these systems, and their performance characteristics have been studied.

What have we learned about parallel scientific computers in this process, and what are the prospects for the future?

First of all, it is clear to us that these systems have great promise. Wray and Rogallo of Ames ported a direct simulation of turbulence application to the Intel and achieved over 1.6 gigaflops. This performance level is competitive with an eight processor Cray system and represents a significant cost-performance advantage over Cray systems.

Quite a number of other scientists have also been successful in porting applications. However, in most cases performance levels are only in the range of 200 to 400 MFLOPS, and cost-performance levels are only on a par with current Cray systems. Furthermore, the parallel systems are obviously much harder to program than conventional systems, and this difficulty is compounded by system environments that are often primitive and unstable.

Thus while we are still basically optimistic that highly parallel computers will predominate in the future, it is clear that some improvements must be made in the current designs. First, actual sustained performance on real aerophysics problems must be substantially increased. The current systems deliver only two to five percent of peak performance on our problems -- clearly there is much room for improvement here. The bottlenecks to improved computational performance are not hard to identify: weak compilers, limited bandwidth between processors and main memory, limited internode bandwidth, and high latency for non-local data accesses.

With respect to internode communication weaknesses, many have suggested that the solution here is to employ alternative algorithms that do not require intensive communication. We have found instead that the challenging aerophysics problems we deal with require implicit algorithms, which are inherently non-local and require intensive communication. It is pointless to employ a numerically inefficient algorithm just to exhibit artificially high megaflops rates on a highly parallel system.

With respect to programming languages, we are now convinced that for a majority of all scientific applications the Fortran-90 language, in particular the array constructs, will be the language of choice, and vendors must support this language for multiprocessor computation, not just on a single node. A principal reason for this is the expected portability of Fortran-90 codes between systems of different architectures. However, we feel that the current Fortran-90 standard is not the entire answer -- standards for specifying data layout need to be imposed, and facilities for non-rectangular data structures need to be added.

Although a high level of sustained performance on real scientific applications, coded with reasonable effort, is the most important consideration in a highly parallel scientific computer, it is clear from our experience that other aspects of these systems also need to be improved. If parallel computers are ever to be widely used in scientific computation centers, then they must be capable of handling a fairly large number of interactive users (say 50), especially during daytime hours when users are debugging and tuning their codes and using interactive visualization tools. Mass storage, interfaces to external data networks, and archival storage are also weak points in current systems that must be improved in the future.

We recognize that the development of truly high performance, highly usable parallel computer systems will be a challenging task. But we feel that without major improvements in both hardware and software, there is a risk that scientists will come to view highly parallel computers as an oversold technology, and they will be reluctant to expend the effort to use them. We look forward to the new generation of highly parallel systems, which are now being announced, and hope that they will exhibit many of these improvements.