

**Parallel Computing in Industry -  
An Initial Survey**

*Geoffrey Fox*

**CRPC-TR92219  
May 1992**

Center for Research on Parallel Computation  
Rice University  
P.O. Box 1892  
Houston, TX 77251-1892



# Parallel Computing in Industry — An Initial Survey

Geoffrey C. Fox

Syracuse University  
Northeast Parallel Architectures Center  
Syracuse, New York 13244-4100  
gcf@nova.npac.syr.edu

## Abstract

We describe some issues in and examples of the use of high performance computing in industry and government applications.

## I: Background

We believe that the progress in parallel computing will be determined by the way and time scale in which it is used by industry. We also believe that high performance computing offers an important opportunity for industry to obtain economic advantage in new and old products. High performance computing is a technology where currently the U.S. has a major lead over both Europe and Japan.

There is much experience with parallel computing, but so far this is mainly in research and academic applications. How do these differ from the "real world" of industrial and government problems? The latter tend to much larger—perhaps one hundred thousand to a million lines of code, not the one to ten thousand line size typical of research academic problems. Correspondingly, industrial codes last longer and evolve more slowly. It is not clear what implication these size and time line differences make for the software support needs. We note that "software engineering" is an underlying principle behind many new government and large industrial software projects. I see little communication between the national high performance computing (HPC) and software engineering communities—this seems a serious issue which needs to be addressed. DoD estimates that the costs of a typical software project divide into 10% initial development, 20% testing and 70% maintenance during the life time of the code. Most HPC efforts (including mine) address the initial 10%—is this

wise? One could argue that the low ratio for initial development costs could suggest that one should not worry about redeveloping codes from scratch and this should be preferred to porting existing codes. This conclusion is countered by the observation that the existing code embodies "expertise" of previous developers and this expertise is not easily available except from the code as the original designers quite likely have moved on. Such nontechnical issues are critical in evaluating the introduction of HPC into industry.

In the area of simulation, e.g., industrial design of new materials, air frames, etc., the needed parallel algorithms have probably already been studied in the research community and so one can be confident that parallel computing will be effective as long as we can address the software engineering issues coming from the special demands of the larger and longer lived industrial codes. On the other hand, one initial conclusion of my evaluation of industrial opportunities for HPC is that simulation is very promising, but not the largest market in the long run. Rather, we live in the "information area" and it is in the processing of information where HPC will have its largest opportunity. This is not (just) transaction processing for the galaxy-wide network of automatic teller machines; rather, it is the storage and access of information followed by major processing ("number-crunching"). Examples include the interpretation of data from NASA's "mission to planet earth" where the processing is large scale image analysis; the scanning and correlation of technical and electronic information from the world's media to give early warning for economic and social crises; the integration of medicaid databases to lower the burden on doctors and patients and identify inefficiencies. Interestingly, such information processing is currently not stressed in the national HPC initiative.

## II. Examples

New York State has funded a project ACTION-NYS (Advanced Computing Technology is an Innovative Opportunity Now for New York State) which is designed

to accelerate the integration of HPC into New York industry. Table 1 presents some initial results [1] of a survey dating from the Summer of 1991 which is currently being extended. This quantifies some of the issues described in Section I. The fourth column "problem class" is an initial classification of the problem architecture which can be used to identify appropriate hardware and software paradigms. These issues are summarized in Table 2 and have been described elsewhere, and will not be discussed in this paper [2-5]. In the following, we refer to the numerical label (item number) in the first column of Table 1.

Items 1, 4, 14, 15, and 16 are typical of major signal processing and feature identification problems in defense systems. Currently, special purpose hardware—typically with built in parallelism—is used for such problems. We can expect that use of commercial parallel architectures will aid the software development process and enhance re-use. HPC in acoustic beam forming (item 1) should allow adaptive on-line signal processing to maximize signal to noise ratio dynamically as a function of angle and time. Currently, the INTEL iWarp is being used although SIMD architectures would be effective in this and most low level signal processing problems. A SIMD initial processor would be augmented with a MIMD machine to do the higher level vision functions. Currently, JointStars (item 4) uses a VAX for the final tracking stage of their airborne synthetic aperture radar system. This was used very successfully in the gulf war. However, HPC could enhance the performance of JointStars and allow it to track many moving targets—one may remember the difficulties in following the movement of skud launchers in the gulf war. We already know good MIMD algorithms for multi-target tracking [6].

We can expect DoD to reduce the purchases of new planes, tanks and ships. However, we see a significant opportunity to integrate new HPC systems into existing systems at all levels of defense. This includes both avionics and mission control in existing aircraft, and the hierarchy of control centers within the armed services. HPC can be used both in the delivered systems and perhaps even more importantly in the simulation of their performance.

Modelling of the ocean environment (item 2) is a large scale partial differential equation problem which can determine dynamically the acoustic environment in which sonar signals are propagating. Large scale (teraflop) machines would allow real time simulation in a submarine and lead to dramatic improvement in detection efficiency.

Computational fluid dynamics, structural analysis and electromagnetic simulation (item 3) are a major emphasis in the national high performance computing initiative—especially within NASA and DOE. However,

the industries that can use this application are typically facing major cutbacks and the integration of new technology faces major hurdles. How do you use parallelism when the corporation would like to shut down its current supercomputer center and, further, has a hiring freeze preventing personnel trained in this area enter the company? We are collaborating with NASA in helping industry with a new consortium GCEFSIC (Grand Challenges in Electromagnetics, Fluids, and Structures Industrial Consortium) where several companies are banding together to accelerate the integration of parallelism into their working environment. An interesting initial concept was a consortium project to develop a non-proprietary software suite of generic applications which would be modified by each company for their particular needs. One company would optimize the CFD code for a new commercial transport, another for aircraft engine design, another for automobile air drag simulation, another for automobile fan design, another for minimizing noise in air conditioners (item 7) or more efficient exhaust pumps (item 6). The electromagnetic simulation could be optimized either for stealth aircraft or the simulation of electromagnetics properties for a new high frequency printed circuit board. In the latter case, we use simulation to identify problems which otherwise would require time consuming fabrication cycles. Thus, HPC can accelerate the introduction of products to market and so give competitive edge to corporations using HPC.

Power Utilities (item 9) have several interesting applications of HPC including nuclear power safety simulation, as well as gas and electric transmission problems. Here the huge dollar value of power implies that small percentage savings can warrant large HPC systems. There are many electrical transmission problems suitable for HPC which are built around sparse matrix operations. For Niagara Mohawk, a New York utility, the matrix has about 4,000 rows (and columns) with approximately 12 nonzero elements in each row (column). We are designing a parallel transient stability analysis system now. Niagara Mohawk's problem (matrix size) can only use a modest (perhaps 16 node) parallel system. However, one could use large teraflop machines (10,000 nodes?) to simulate larger areas—such as the sharing of power over a national grid.

In a completely different area, the MONY Insurance Company (item 10) spends \$70M a year on data processing—largely on COBOL applications where they have some 15 million lines of code and a multi-year backlog. They see no immediate need for HPC, but surely a more productive software environment would be a great value! Similarly, Empire Blue Cross/Shield (item 11) processes 6.5 million medical insurance transactions every day. Their IBM 3090-400 handles this even with automatic optical scanning of all documents.

HPC could only be relevant if one could develop a new approach with perhaps HPC examining the database with an expert system or neural network to identify anomalous situations. The states and federal government are burdened by the major cost of medicaid and small improvements would have great value.

The major computing problem for Wall Street (items 12, 13) is currently centered on the large databases. SIAC runs the day-to-day operation of the New York and American Stock exchanges. Two acres (about 300) of Tandem computers handle the calls from brokers to traders on the floor. The traders already use an "embarrassingly parallel" decomposition with some 2000 stocks of the New York Stock Exchange decomposed over about five hundred personal computers with about one PC per trader. For SIAC, the major problem is reliability and network management with essentially no down time "allowed". HPC could perhaps be used as part of a heterogeneous network management system to simulate potential bottlenecks and strategies to deal with faults. The brokerages already use parallel computers for economic modelling [7,8]. This is obviously glamorous with integration of sophisticated optimization methods very promising.

As our final example (item 17), we have the entertainment and education industries. Here HPC is linked to areas such as multimedia and virtual reality with high bandwidth and sophisticated visualization and delivery systems. Some applications can be viewed as the civilian versions of military flight simulators with HPC systems replacing the special purpose hardware now used. Parallelism will appear in the low end with future extensions of Nintendo like systems; at a medium scale for computer generated stages in a future theater; at the high end with parallel supercomputers controlling simulations in tomorrow's theme parks. HPC should also appear in all major sports stadiums to perform image analysis as a training aid for coaches or providing new views for cable TV audiences. We can imagine sensors and tracking systems developed for Strategic Defense Initiative being adapted to track players' on a football field rather than a missile launch from an unfriendly country. Many might consider this appropriate with American football as aggressive as many military battles!

Otis (item 5) is another example of information processing discussed generally in Section I. They are interested in setting up a database of elevator monitoring data which can be analyzed for indicators of future equipment problems. This would lead to improved reliability—an area where Otis and Japanese companies compete. In this way, HPC can lead to competitive advantage in the "global economic war".

Table 2: Architectures for Five Problem Classifications

- **Synchronous: Data Parallel**

Tightly coupled. Software needs to exploit features of problem structure to get good performance. Comparatively easy, as different data elements are essentially identical. *Candidate software paradigms: High Performance Fortran, Parallel Fortran 77D, Fortran 90D, CMFortran, Crystal, APL, C++.*

- **Loosely Synchronous**

As above but data elements are not identical. Still parallelizes due to macroscopic time synchronization. *Candidate software paradigms: may be extensions of the above. C (Fortran) message passing is currently only guaranteed method!*

- **Asynchronous**

Functional (or data) parallelism that is irregular in space and time. Often loosely coupled and so need not worry about optimal decompositions to minimize communication. Hard to parallelize (massively). *Candidate software paradigms: PCN, Linda, object-oriented approaches.*

- **Embarrassingly Parallel**

Independent execution of disconnected components. *Candidate software paradigms: Several approaches work? PCN, Linda, Network Express, ISIS.*

- **A=LS (Loosely Synchronous Complex)**

Asynchronous collection of loosely synchronous components where these program modules can be parallelized. *Candidate software paradigms: PCN, Linda, ADA, controlling modules written in synchronous or loosely synchronous fashion.*

## Acknowledgments

This work was supported by the State of New York and the Center for Research on Parallel Computation with National Science Foundation Cooperation Agreement No. CCR-8809165—the Government has certain rights to this material.

## References

- [1] Pat Carroll, "Industrial Applications for Parallel Computing", *Numerical Aerodynamic Simulation Program Newsletter*, Vol. 7, No. 3, April 1992.
- [2] Geoffrey C. Fox, "FortranD as a Portable Software System for Parallel Computings", in *Proceedings of Supercomputing USA/Pacific 91*, held in Santa Clara, California. Syracuse University Technical Report SCCS-91, June 1991.
- [3] Geoffrey C. Fox, "The Architecture of Problems and Portable Parallel Software Systems". Syracuse University Technical Report SCCS-134, July 1991.
- [4] Geoffrey C. Fox, "Lessons from Massively Parallel Architectures on Message Passing Computers" in *The 37th Annual IEEE International Computer Conference, COMPCON '92*, IEEE Computer Society Press. Syracuse University Technical Report SCCS-214, December 1991.
- [5] Alok Choudhary, Geoffrey Fox, Seema Hiranandani, Ken Kennedy, Chuck Koelbel, Sanjay Ranka and Joel Saltz, "A Classification of Irregular Loosely Synchronous Problems and Their Support in Scalable Parallel Software Systems", in *Proceedings from Darpa Software Technology Conference*, 1992, pp. 138-149.
- [6] T. D. Gottschalk, "Concurrent Multiple Target Tracking", in *The Third Conference on Hypercube Concurrent Computers and Applications, Volume 2*, ACM Press, New York, 1988, pp. 1247-1268. California Institute of Technology Technical Report C3P-567.  
  
T. D. Gottschalk, "Concurrent Multi-Target Tracking", in *The Fifth Distributed Memory Computing Conference, Volume I*, D. W. Walker and Q. F. Stout, editors, IEEE Computer Society Press, 1990, pp 85-88. California Institute of Technology Technical Report C3P-908.
- [7] Kim Mills, Michael Vinson, Gang Cheng, Thomas Finucane, "A Large Scale Comparison of Option Pricing Models with Historical Market Data" submitted to the *Fourth Symposium on the Frontiers of Massively Parallel Computing*, March 1992. Syracuse University Technical Report SCCS-260.  
  
Kim Mills, Gan Cheng, Michael Vinson, Sanjay Ranka, Geoffrey Fox, "Load Balancing, Communication, and Performance of a Stock Option Pricing Model on the Connection Machine-2 and DEC MPP-1200", March 31, 1992. Syracuse University Technical Report SCCS-273.
- [8] Stavros Zenios, "Massively Parallel Computations for Financial Planning under Uncertainty", Chapter 18 in *Very Large Scale Computation in the 21st Century*, edited by Jill P. Mesirov, SIAM, Philadelphia, 1991.

Item	Company/Agency	Applications	Current Big Computer	Problem Class	Can One Use HPC?	Relevant Dusty Decks Exist?
1	General Electric (Syracuse)	Acoustic Beam Forming	Special & iWarp	Synchronous	yes	no
2	General Electric (Syracuse)	Ocean Environmental Modeling	many	Loosely Synchronous	yes	maybe
3	Gruddman United Technologies	CFD & Structures Electromagnetic Simulations	Cray & CM-2	Synchronous/Loosely Synchronous	yes	yes
4	Gruddman Par Systems	Avionics & Command and Control (JointStars)	Special & VAX	Loosely Synchronous Complex	yes	probably not
5	Otis	Manage Repair Database	Not Implemented Yet	Loosely Synchronous Complex	yes	no
6	Lamson (small Syracuse Company)	Design Exhaust Pump	None	Loosely Synchronous	?	yes
7	Carrier	Design Air Conditioners	VAX	Loosely Synchronous	? (is physics or noise understood)	yes
8	Atlantic Research (ARC)	Antenna Fields	VAX	Embarrassingly Parallel	yes	yes (little change)
9	Niagara Mohawk	Control and Plan Electrical Power Grid	IBM 3090	Loosely Synchronous	yes	no
10	MONY Insurance Co.	Process Insurance Policies	IBM 3090	Embarrassingly Parallel	no	yes

Item	Company/Agency	Applications	Current Big Computer	Problem Class	Can One Use HPC?	Relevant Dusty Decks Exist?
11	Blue Cross/Blue Shield	Process Medicare Claims	IBM 3090	Embarrassingly Parallel	no	yes
12	Wall Street Brokers	Predict Values of Securities	Maspar, iPSC/860 & Hypercube	Some Embarrassingly Parallel	?	yes (small)
13	SIAC	Control Trading on Floor of NY and American Stock Exchange	Tandem & PC	Embarrassingly Parallel	?	yes
14	Electronic Systems Division ESD (USAF)	Command and Control & Battlement Management for SDI	Many	Loosely Synchronous Complex	yes	probably not
15	Army ASAS Project	Data Fusion	VAX	Loosely Synchronous Complex	yes	somewhat
16	Par Systems	Infrared Sensors in Aircraft	Special Purpose	Synchronous	yes	no
17	AGE	Entertainment, Toys, Theme Parks, Virtual Reality	Must be low cost (toys) Supercomputer (theme park)	Loosely Synchronous (complex)	yes	no

Table 1: An Initial Survey of Industry and Government Opportunities for High Performance (Parallel) Computing