

It was delightful to learn (*HPCC Week*, 31 March 1994) that one of our MPP State Computer vendors, Thinking Machines, has finally discovered commercial database software produced by an Independent Software Vendor (ISV) and sees this as an important application that could lead to profitability and new applications. Teradata aka NCR aka AT&T have delivered MPP systems for databases for about a decade that generate revenue that's equal to the MPP technical market. Your readers, who make policy and fund computer systems development at ARPA, DOC and DOE should note the following ironies about this discovery:

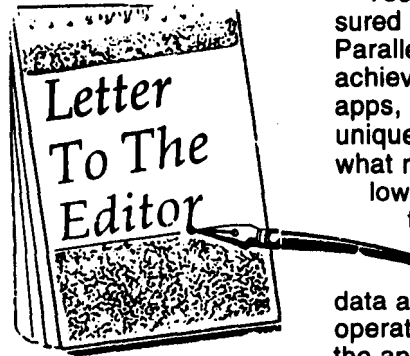
1. After 11 years of investment of \$1+-0.5 billion (private and federal R&D, federally stimulated purchases and customer funded systems and apps) to just one of the State companies, it might become profitable to pay off the investment. Perhaps after the Paragon works better, (*HPCC Week*, April 7, 1994), Intel may also "discover" the commercial MPP market.

2. The database port was simple and straightforward, if one believes TMC's marketing person (so what took them so long?), because Oracle has been working on a parallel, distributed database and transaction processing monitor for many years for workstations, multiprocessors and multicomputers (Meiko and especially nCUBE). Larry Ellison, Oracle's founder and CEO, is a substantial owner of nCube. nCube has not benefitted from HPCC and has probably been discriminated against as a non-State Computer vendor. Thus, an outcast just might be the biggest HPCC contributor.

3. Some conclusions. It's difficult to pick winning computers and applications. Federally funded, "supply-side" R&D to companies is most likely to induce brain damage in a company. Funding, like drugs, becomes an addiction that inhibits self-sustenance. The very technically bright people at Thinking Machines may have figured out how to get off the drug. As an investor, I'd like to believe it.

It should be no surprise that MPP may be more amenable to solving commercial, rather than the technical apps that the federal government has focused on. Commercial applications have implicit parallelism by being able to access a database in parallel and with multiple, independent transactions. Only a few parallel application environments needed to be ported and users write apps in a higher 4GL fashion. Technical apps require users to deal explicitly with parallelism and every application is unique.

The commercial market is much larger although it has what may seem like unreasonable demands for the technical market...products have to work, be cost-effective and retain data forever.



Why commercial MPP may be a market and technical MPP is not.*

Technical computers are measured in Mflops, Mips, and Mpixels/s. Parallelism is explicit, difficult to achieve across a wide range of apps, and every application is unique. Users expect little from what may be marginal systems with low MTBF, because they expect to pay nothing even though they are subsidized.

Simple file systems store data and scratch files, and the operating system is not involved in the application. While the current multicomputers don't share

memory, software (overhead) is required to make a shared memory for programmability.

A user typically works on a single, large problem doing design, simulation and analysis for a long time. The system attempts to minimize the time for each user to get an answer using a batch environment. Users require a wide range of resource and application programs. Applications are unique using basic packages such as Nastran, lab grown models such as NCAR's, or researcher-specific (aka graduate student) programs.

The many ISVs in a broad scientific and engineering marketplace have ignored MPP in favor of apps that run on widely available platforms e.g. workstations, multiprocessor servers, and supers because of the small market size, porting difficulty, and small mPP (minimally Parallel Processing) gain. The funding for large scale MPPs from the federal government seems unlimited. Industrial users are much more cautious about MPP compared with the many, cost-effective alternatives.

Commercial computers are measured in disk accesses/s and transactions/s against a fixed database system by multiple independent (parallel) users with many relatively simple parallel requests. Parallelism is implicit in the ported database or occurs through multiple, independent users. A distributed memory works just fine. Users are concerned with persistence of their data, measure MTBF in weeks and the cost/transaction or data-access.

Many users access a common database so the system's goal is to maximize the number of transactions or accesses for all users. The operating system and database systems are tightly integrated with most of the computation occurring in the ported database. Only a few ISVs supply a basic, albeit large, package that is ported to a computer. Users buy portable apps from a single vendor that accesses the vendor's database. Thus only a few packages need to be ported and users write in very high level 4 GL languages. Surveys show that few buyers are investigating MPP.

* Based on a presentation by Frederica Damema and Francis Parr, IBM Research.

Sincerely,
GORDON BELL