MINTEK

Practical desktop supercomputing

by R.T. Jones

Third National Symposium on Computational Fluid Dynamics
30 June - 2 July 1993, University of Stellenbosch

Mintek Paper No. 8235

Practical desktop supercomputing

by **R.T. Jones**
Principal Engineer, Pyrometallurgy Division
Mintek, Private Bag X3015, Randburg, 2125, South Africa

<u>Synopsis:</u> Scientists and engineers studying computationally-intensive problems are confronted with a wide choice of computer hardware. The range of options encompasses personal computers (possibly equipped with specialised numeric processing boards, such as those based on the Inmos T800 transputer or those using the Intel i860 RISC chip), workstations, and supercomputers.

Unfortunately, direct comparisons of 'number-crunching' power are seldom straightforward. Manufacturers commonly present benchmark tests using code that has been highly optimized for a particular system. Often these theoretical figures have no bearing on the performance that can be expected when running real-world applications.

For the purpose of making direct comparisons between the performance of different machines, a finite-volume heat-transfer problem in two-dimensional cylindrical geometry was coded using a readily-portable subset of the popular 'C' language. This code was compiled and run on various computer systems. The performance of a wide range of machines has been compared on the basis of the running times of this benchmark test.

1 INTRODUCTION

In less than 50 years, computer technology has progressed from the electromechanical simplicity of the ENIAC Mark-2 (1946), which could carry out a single floating-point operation in one second, to that of supercomputers such as the Cray 2 operating at over 1600 million floating-point operations per second (1600 MFLOPS).[1] Machines with even higher speeds, based on many different computer architectures, continue to be announced. For example, Cray Research has recently announced[2] a new parallel machine based on the DEC Alpha microprocessor, to complement their base of vector-processor machines. Rapid progress also continues to be made in the area of reduced instruction set computing (RISC), as well as parallel processing.

Scientists and engineers studying computationally-intensive problems are confronted with an increasingly wide choice of computer hardware on which to run their programs. Most people have access to one of the newer IBM-compatible personal computers, but are unsure as to how these perform relative to specialised numeric processing boards (such as those based on the Inmos T800 transputer or those using the Intel i860 RISC chip), workstations, or supercomputers. Unfortunately, direct comparisons of 'number-crunching' power are seldom straightforward. Manufacturers often present benchmark test results from code that has been highly optimized for a particular system. On the other hand, compilers are often specifically optimized to show their best performance on a few well-known benchmarks. The best-known floating-point benchmark test is the LINPACK benchmark, which solves a number of linear simultaneous equations. This exemplifies double-precision vector operations. Upon investigation, it is apparent that these theoretical figures have little bearing on the performance that can be expected when running real-world applications.

In order to choose between various computing platforms, it would be helpful to have a good idea of how the various systems perform on a task that approximates the kind of calculations performed in real life.

2 CFDTEST BENCHMARK

For the purpose of a direct comparison between the floating-point calculational performance of different computers, a finite-volume heat-transfer problem in two-dimensional cylindrical geometry was coded using a readily-portable subset of the popular 'C' language. This code has been compiled (with optimization for speed) on various platforms. As a reference, the popular 'industrial-strength' MS-DOS compiler, Borland C++ version 3.1, was used for compilation and running on a 50 MHz i486 personal computer. A comparison has been made of the running times of this benchmark test on a number of computers.

The CFDTEST program simulates two-dimensional heat transfer in a composite system of cylindrical geometry, in which a bath of molten slag is contained in a vessel lined with refractory bricks. The thermal conductivity is given as a quadratic function of temperature. The test can be run a number of times, with each test automatically using a slightly different boundary temperature in the range 1700 to 1800°C (to prevent smart compilers from avoiding duplicated code). A step size of 20° was specified for the final tests, thereby running the problem six times. Grid sizes were also varied in the initial round of tests, with a 21 x 21 grid being used for the results presented here. This grid is sufficiently coarse to allow all data for all of the elements to reside in memory simultaneously, without having to be spooled to auxiliary memory or to a disk. The steady-state temperature at a particular point is the only substantial output from the program. Provision is made for switching off the display. Alternatively, this can be done via the operating system (by redirecting the output to a null device). No particular significance should be placed on the actual running time for the benchmark - the test was devised merely to distinguish between the relative performance of different machines.

Most of the computational work in the test program involves the iterative solution of a system of equations that can be broken down into sets of simultaneous linear equations expressed in tridiagonal matrix form. These equations are solved using the Thomas algorithm, also called the TDMA (TriDiagonal Matrix Algorithm).[3]

The tests were carried out by compiling the code with the best available speed-optimization switches set on the various compilers. No alterations were made to the actual source code (beyond the very occasional trivial changes required in the variable or function declarations, to satisfy an old compiler). The actual elapsed running time, from the answered prompt of the running program to the final appearance of the results on the screen, was measured on a stopwatch, and the CPU time was measured by using a timing command, if one was supplied by the operating system. The tests were usually run about four times to ensure consistency of results. The results of the calculations and the number of iterations were checked to ensure that they were the same in all cases. Where a graphically windowed display was used, a test was also run with the results redirected to a null device, in order to ensure that even the minimal screen display did not slow down the test. For computers running a single task, the elapsed time and the CPU time are by definition the same. In the case of multiprocessing

The i486 is the first complex instruction-set computer (CISC) chip to challenge RISC chips in performance. The 80486 is a combination of an enhanced 80386 processor, an enhanced 80387 floating-point unit (FPU), a memory-management unit, a cache controller, and an 8 kilobyte cache. Most importantly, it is compatible with the earlier 80386 and 80286, and even the 8088 chips. The 50 MHz version of the 486 is rated at 30 MIPS.

Intel's i586 (also called the P5 or Pentium) chip will form the basis of the next generation of PCs. This chip, rated at 100 MIPS (million instructions per second), has just begun to be produced, and should be more widely available during the course of 1993.

Motorola's 68040 microprocessor is another reasonably powerful performer that is used in some personal computers, notably the Apple Macintosh.

IBM-compatible personal computers (usually equipped with a fast i486 processor) can also act as a host to specialised numeric processing boards (such as those based on the Inmos T800 transputer or those using the Intel i860 RISC chip). These boards are supplied with system-specific compilers.

4 TRANSPUTERS

Inmos launched the first transputer in 1985, as the world's first microprocessor built for parallel processing. Each extra transputer chip added to a parallel network adds extra computational power and communications bandwidth to keep these roughly in balance as the network grows, thus preventing communications bottlenecks from developing between processors. The Inmos T800 floating-point transputer chip was launched in 1987. The 20 MHz version of this chip is rated at 1.5 MFLOPS. Since the introduction of this fast chip, newer RISC and CISC (Complex Instruction Set Computing) processors from Sun, Mips Computer Systems, Intel, and DEC have obliterated the initial performance edge once held by the transputer. Some vendors of parallel computers have switched to the Intel i860 chip, despite its conventional bus-based communications between chips. Inmos announced the T9000 transputer (200 MIPS, 25 MFLOPS, 50 MHz, two million transistors, 10 times as fast as a T800-20) in mid-1991, but delays in its introduction have led to much disappointment and scepticism among transputer users.

Parallel processing, at first sight, seems to be the obvious way to tackle CFD problems, by solving equations for many gridpoints in parallel. However, use of this capability requires changes to existing computer programs, as well as to the way we think about and formulate solutions to problems. Most scientists and engineers are trained to think serially, and to define problems serially in conventional programming languages. Most researchers are not interested in the minute details of how to optimize the computer's performance, and prefer rather to concentrate on better definition of their problems. This has been one of the barriers to wide acceptance of transputers.

The test results show that a single T800-20 transputer is just over one third of the speed of an i486-50. Obviously, more than three transputers would be required to compete with an i486-50. A transputer system could be built with a base cost of about R5 000, and an additional cost of R4 000 per transputer.

number and nature of the other jobs that are running at the same time. The detailed results are presented in Table I, and are discussed in the following sections.

Table I: Comparison of performance of various computer systems, showing the time taken to run the CFDTEST benchmark, and the speed index relative to a 486-50 PC

Computer system	Time	Speed index
HP 9000 Model 350	228.0 s	0.08
387-16	169.4 s	0.10
INMOS Transputer T800-20	50.8 s	0.35
486-33	25.5 s	0.69
Sun SPARCstation 2	17.8 s	0.99
486-50	17.6 s	1.00
Convex 210	9.8 s	1.79
IBM RS/6000 320H	9.7 s	1.82
Intel i860-40	7.0 s	2.51
ICL DRS 6000 Model 620	6.3 s	2.79
Silicon Graphics Personal IRIS 4D/35	5.5 s	3.20
Sun SPARCstation 10 Model 30	5.1 s	3.45
HP Apollo 710	3.5 s	5.07
Cray 2	3.4 s	5.25
Cray XMP	2.6 s	6.89
Cray M92	2.4 s	7.26

The test results reflect, as purely as possible, the relative ability of the various computers to perform floating-point calculations of the type usually encountered in computational fluid dynamics (CFD) and related areas. The computation-bound execution rate depends on three factors:

i)   the number of compiler-generated machine instructions produced to perform the calculations;

ii)   the sustained rate of instructions completed per clock cycle; and

iii)   the clock rate (made possible by the underlying integrated-circuit technology).

High performance requires advances on all three fronts. All designs will benefit from better compilers and a high clock rate.

3 PERSONAL COMPUTERS

The original IBM PC (based on Intel's 4.77 MHz 8088 microprocessor) was launched in 1981, and would certainly have been woefully inadequate for serious CFD work, even when equipped with an 8087 math coprocessor. However, personal computers have evolved extremely rapidly. The 50 MHz i486 chip (also called an 80486, but designated here as 486-50) runs numeric problems more than eighty times as fast as the original PC (8088-4.77 with 8087 math coprocessor). Computers based on the 486-50 are widely available, and are relatively inexpensive, costing in the region of R7 000. This was chosen as a suitable

5 i860 BOARDS

The i860 CPU (launched in 1989) was the world's first one-million-transistor microprocessor. Unlike other RISC processors (which are usually chip sets), the i860 is a single chip. It integrates a RISC integer core, a floating-point unit, a 3D graphics processor, data and instruction caches, memory management, and a bus interface and cache control. The floating point unit can deliver a peak performance rate of 80 MFLOPS (in the 40 MHz version of the chip). Since all the sub-systems are on a single chip, the transfer of information between systems is allowed to keep up with the CPU's internal speed. Another very sizeable part of the i860's speed comes from its parallel design. The chip is designed to do three things at once: integer calculations, floating-point addition, and floating-point multiplication. This is possible because the three mathematical units are separate, and can work on different problems at the same time. In addition, each of the three mathematical units is designed for pipelining, where a set of instructions (residing in an instruction cache) can be performed on each element in the data cache. To take advantage of this feature, tricky hand- or machine-generated code is needed to optimize the instruction order for the pipeline. Smart compilers are certainly needed here. Intel's claim of 80 MFLOPS is based on hand-coded matrix-multiply pipelined routines.

The Microway Number Smasher-860 is a coprocessor board hosted inside a PC/AT or compatible computer. This board was tested and found to be over 2.5 times the speed of an i486-50. An i860 processor-based system (using a plug-in coprocessor card for a personal computer) with something approaching supercomputing power can be built with a system price of about R30 000. This system is very powerful and cost-effective, although it delivers only about one fifteenth of the performance that could be expected if it was to run at its theoretical maximum of 80 MFLOPS. Very high performance levels can certainly be achieved, but only with special attention and hand-coding. However, some simple optimization can be carried out if the source code is to be written specifically for this system. For example, there is no intrinsic divide instruction built into the i860 hardware, so division operations should be replaced by suitable multiplication operations wherever possible.

Microway has recently developed the QuadPuter-860, and rates this at 200 MFLOPS. This board contains four i860 chips. Five of these boards can be built into a computational server costing about R200 000. This machine would be rated at 1000 MFLOPS.

6 WORKSTATIONS

Workstations running variants of the Unix operating system provide another popular computing platform. These workstations are often supplied with expensive high-end graphics capabilities and large storage capacities, and are usually centred around a RISC processor.

RISC has emerged over the last 15 years as the model of choice for the design of general-purpose processors[4]. The basic concept involves reducing the instruction set to a bare minimum of only the most frequently used, simple, single-stage instructions that can each be performed in a single clock cycle, and giving the language compilers greater responsibility for handling complex operations. Rapid progress is being made in this area, with semiconductor companies continually announcing new CPUs such as the Intel 860, Sun

SPARC, DEC Alpha, Inmos T9000 transputer, IBM RISC System/6000, and Mips R3000 and R4000.

One of the most widely used RISC processors is the SPARC (Scalable Processor ARChitecture) chipset, as used in the Sun SPARCstation 10 and the ICL range. Running second in popularity is the Mips processor (as distinct from the abbreviation for 'Million Instructions Per Second'), which is the CPU for the Silicon Graphics Personal IRIS workstation. Other contenders that have been available for a few years include IBM's own RISC CPU and Motorola's 88000.

Interestingly, like the i860, the SPARC system performs division operations far slower than multiplication. Only 1 cycle is required per multiply instruction, but 4 are needed for divide (single precision) and 7 for divide (double precision).

Digital Equipment Corporation announced its Alpha RISC processor during 1992. The Alpha chip is now widely regarded as being the world's top performer. It operates at 150 MHz, and achieves an integer processing level of 300 MIPS, and a floating-point performance of up to 150 MFLOPS.

A number of prominent workstation systems were tested, so that users of computers in a particular range should better be able to gauge the relative performance of other machines. However, these tests should not be regarded as sufficiently discriminating to determine subtle differences in performance by systems belonging to various manufacturers. This would require a lot more attention to the individual configurations of the various models.

The machines tested fall into a price range between about R60 000 and R120 000. None of the workstations were at the very top of the range, and it should be quite possible to obtain machines that perform even better than those tested. The HP Apollo 710 delivered almost identical performance to that of the Cray 2; this was the fastest workstation amongst those tested, and can certainly be regarded as a personal supercomputer. It is also interesting to note that the model 735 is rated at more than three times as fast as the 710. If this is the case, better performance could be expected from this machine than from the Cray supercomputers that were tested. The Silicon Graphics Personal IRIS 4D/35 and Sun SPARCstation 10 also turned in very creditable performances. The ICL DRS 6000 Model 620 and IBM RS/6000 320H were both running many other jobs when tested. This showed up in the ratio of elapsed time to CPU time, which varied between two and three. Again, these machines are not at the top of their ranges, and significantly better performance is expected from the top-of-range machines. For example, the IBM RISC System/6000 models 360/5 and 370/5 are about twice the speed of the 320H. The dramatic improvements over the space of a few years can be seen by a comparison of the performances of the early models with the more recent ones. For example, the HP Apollo 710 is about 65 times as fast as the HP 9000 Model 350, and the Sun SPARCstation 10 Model 30 is about 3.5 times the speed of the Sun SPARCstation 2.

However, for a detailed comparison of the strengths of individual manufacturer's machines, the full range of products would have to be tested. This is beyond the scope of the present investigation, as the primary intention here is to provide some idea of how the various categories of computer perform relative to each other. Machines were tested on the basis of

However, it should be easy (using manufacturers' own comparative benchmarks) to scale the performance of computers within a particular manufacturer's product range.

7 SUPERCOMPUTERS

The first Cray supercomputer (the Cray 1) was launched in 1976. It had one processor, and required an elaborate refrigeration system to keep it cool. The Cray 1 is capable of carrying out 190 million floating point operations per second (MFLOPS). In 1982, Cray released its first multiprocessor machine, the Cray X-MP with four processors, capable of 713 MFLOPS.[1] The Cray 2, which followed in 1985, is capable of 1600 MFLOPS.[5] Other, even faster, machines have been developed since then.

Tests were conducted on various Cray supercomputers at Minnesota Supercomputer Center Inc. in Minneapolis, an affiliate of the University of Minnesota. In 1981, the University of Minnesota became the first university in the United States to acquire a supercomputer (a Cray 1). Other computers available are the four-processor Cray 2/4-512 (1985), with a top speed of over 1600 MFLOPS, and the four-processor Cray X-MP/416 (1989). An eight-processor Cray C90 is expected to be installed in April / May 1993; this machine should be four to five times as fast as a Cray 2.

These machines obviously have enormous potential processing power. They were also the fastest and most impressive of the machines tested. However, the factor by which they were faster was much less than might be expected. To take full advantage of the innate processing power, the source code would need to be developed with the Cray in mind, in order to use the vector- and parallel-processing features of these machines. This would be appropriate for large computationally-intensive problems that are well defined. However, for research code, it is not altogether practical to spend a significant amount of time on code optimization when the algorithms are still under development.

The performance benefits of supercomputers can be seen when moving from scalar processing (calculations on individual numbers) to vector processing (simultaneous calculations on lists of numbers). The key to conventional supercomputer performance is vector processing, performed by a set of pipelined functional units, and special vector instructions designed to operate in rapid succession on many related data items. For efficient vector operations, the memory system must be able to produce and accept strings of data items very quickly. The most common way to achieve this is to use interleaving, a technique by which many independent memory banks are organised so that successive addresses are routed to a sequence of different banks. This allows the multiple banks to operate essentially concurrently, thus seeming to be producing or accepting data items much faster than the basic cycle time of the memory chips.[6]

Apart from the obvious drawback of the expense of these machines, or the difficulties in obtaining access, the actual running time of a program will almost always be much greater than the amount of CPU time utilised, because the computer's resources are shared between a number of simultaneous users. It should be stressed that although supercomputers would probably not be the best choice for the simple task presented here, they clearly still provide

possible for users to experience long 'time-to-solution' intervals. Realistically, development work should be carried out on other machines, and the supercomputer should be used for final detailed production runs.

It is interesting to note that, according to the official benchmarks, the Cray 2 is rated as being more than twice as fast as the Cray X-MP. However, the Cray X-MP actually ran the test program 30 per cent faster than the Cray 2.

A Convex 210 supercomputer was tested at the University of the Witwatersrand. This heavily-used machine yielded only mediocre performance, with real elapsed time being significantly longer than on a PC. Even the CPU time used was only half as much as that taken by an i486-50.

The classic application of supercomputers is that of weather prediction, where the system's equations need to be solved as fast as possible. It has been said that 'There is a very limited market for predictions of yesterday's weather'.[6]

8 PARALLEL PROCESSING

If you have an application that requires 1000 MFLOPS of computing power, you could wait for the development of a more powerful processor, or you could use twenty i860 chips running in parallel today. 'It's a lot easier to harness 100 horses than to grow one that's 100 times bigger', according to Michael Dertouzos, Director of MIT's Laboratory for Computer Science.[2]

Parallel processing architectures are traditionally divided into two broad camps: SIMD (pronounce *sim-dee*), or single instruction / multiple data, and MIMD (pronounced *mim-dee*), or multiple instruction / multiple data. The most obvious topology for a SIMD machine is an array processor, which calculates a large two-dimensional matrix of values in parallel by applying the same operation to every element. The best-known and most commercially successful SIMD machine is the Connection Machine from Thinking Machines, which can be supplied with as many as 64 000 processing elements[2]. The latest version of the Connection Machine is the CM-5, which has a peak processing speed of 130 GFLOPS[7]. Each node in the machine contains a SPARC processor, built by Sun Microsystems for its high-performance workstations. SIMD machines are well suited for speeding up the matrix calculations used in computational fluid dynamics, although they are less suitable for general computing tasks in which the data is not naturally expressible as a large uniform array.

MIMD involves connecting a number of processors that run different programs or parts of a program on different sets of data. Communications between the processors is crucial. If two processors must cooperate on different parts of the same problem, they must communicate results to each other. (The class of MIMD computers can be further subdivided into shared-memory and distributed-memory machines.) This approach holds a pitfall for the programmer, who must keep track of what each processor is doing at all times. Some users prefer to program for a machine such as the Cray Y-MP C90, which has 16 very powerful processors, rather than for a massively parallel machine. Robert Hyatt, a computer scientist at the University of Alabama, poses the question 'Do you want to use 16 elephants to

A group at Los Alamos reworked a standard ocean model for a CM-5, which then produced results five to 20 times as fast as a Cray Y-MP, a machine that was the world's fastest supercomputer when it was introduced in 1988.

Like other specialised machines, when assigned to solve problems from the real world, with software that was originally written for traditional computers, massively parallel machines often achieve only a fraction of their advertised speeds. In order for parallel computers to reach their potential, programmers have to rethink fundamentally how they want to attack a problem, and rewrite programs to take advantage of any parallelism.

Unfortunately, the CM-5 machine at the Minnesota Supercomputer Center Inc. was not available for testing, nor was any other massively parallel machine.

9 CONCLUSIONS

A comparison of performance of the various classes of machine is shown in Figure 1. This simplified comparison uses representative values for the time taken by each class of machine. It should be stressed that this comparison is valid only for computationally-intensive tasks. Clearly, the results might be quite different for a comparison of database manipulation, graphics rendering, or tasks requiring vast amounts of memory or disk space.
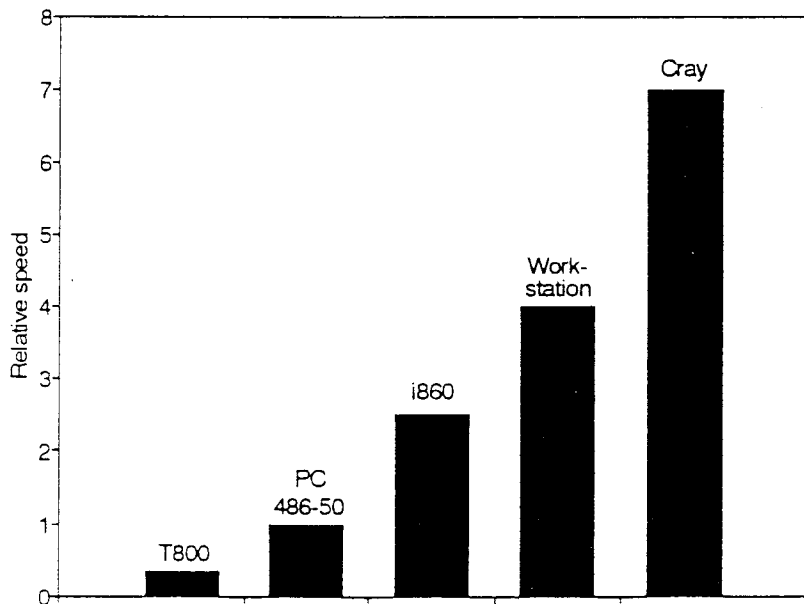

Figure 1: Representative floating-point performance comparison between classes of machines

Figure 1 shows just how much single-transputer systems have been overtaken by other systems. The fact that transputer development has slowed down makes it difficult to recommend the use of this approach today. Most noticeable from Figure 1 is the fact that the ratio of Cray supercomputer performance to that of the i486-50 PC is only about seven to one.

to-performance ratio is about 70 per cent higher than that of the PC. Workstations provide performance approaching that of supercomputers, although this comes with a cost-to-performance ratio of about three times that of the PC (based on the assumption of a single function to be performed by the workstation). The enormous cost of supercomputers, and the fact that they are not usually bought for single tasks, makes it difficult to compare the cost-to-performance ratio, which appears to be about 200 times that of the PC. Obviously though, there are many tasks where supercomputers are the most suitable tool for the job.

For many computing tasks similar to that modelled by the CFDTEST benchmark, an i486-50 (or the next generation) PC (possibly fitted with an i860 board) is the most cost-effective solution. It can even provide the quickest turnaround time, because the machine is not shared between users. Supercomputers have enormous computing power — but to harness this requires special and dedicated effort, which is not appropriate for experimental or research-based code that is undergoing continual modification.

I believe it is true to say that 'Desktop supercomputing has become a practical reality'.

10 ACKNOWLEDGEMENTS

11 REFERENCES

1. GALEA E. Supercomputers and the need for speed, *New Scientist*, 12 November 1988, pp. 50-55.

2. POUNTAIN D. and BRYAN J. All Systems Go, *Byte*, August 1992, pp. 112-136.

3. GERALD C.F. *Applied Numerical Analysis*, 2nd ed., Addison-Wesley, Reading, Massachusetts, 1978, p. 133.

4. SITES R.L. RISC Enters a New Generation, *Byte*, August 1992, pp. 141-148.

5. Minnesota Supercomputer Center, Inc. brochure.

6. SERLIN O. Shakeout in the supercomputer market, *Unix World*, November 1988, pp. 68-78.

7. CHARLES D. The thinking machine's guide to computing, *New Scientist*, 5 September 1992, pp. 26-30.