

The Marketplace of High Performance Computing

Erich Strohmaier^{a 1}, Jack J. Dongarra^{a,b 2}, Hans W. Meuer^{c 3}
and Horst D. Simon^{d 4}

^a *Computer Science Department, University of Tennessee, Knoxville, TN 37996*

^b *Mathematical Science Section, Oak Ridge National Lab., Oak Ridge, TN 37831*

^c *Computing Center, University of Mannheim, D-68131 Mannheim, Germany*

^d *NERSC, Lawrence Berkeley Laboratory, 50A, Berkeley, CA 94720*

Abstract

In this paper we analyze the major trends and changes in the High Performance Computing (HPC) market place since the beginning of the journal ‘Parallel Computing’. The initial success of vector computers in the seventies was driven by raw performance. The introduction of this type of computer systems started the area of ‘Supercomputing’. In the eighties the availability of standard development environments and of application software packages became more important. These criteria determined next to performance the success of MP vector systems especially at industrial customers. MPPs became successful in the early nineties due to their better price/performance ratios which was enabled by the attack of the ‘killer-micros’. In the lower and medium market segments the MPPs were replaced by microprocessor based SMP systems in the middle of the nineties. This success was the basis for the emerging cluster concepts for the very high end systems. In the last few years only the companies which have entered the emerging markets for massive parallel database servers and financial applications attract enough business volume to be able to support the hardware development for the numerical high end computing market as well. Success in the traditional floating point intensive engineering applications seems to be no longer sufficient for survival in the market.

Key words: High Performance Computing, HPC Market, Supercomputer Market, HPC technology, Supercomputer market, Supercomputer technology

¹ e-mail: erich@cs.utk.edu

² e-mail: dongarra@cs.utk.edu

³ e-mail: meuer@rz.uni-mannheim.de

⁴ e-mail: simon@nersc.gov

1 Introduction

“The Only Thing Constant Is Change” — Looking back on the last decades this seem certainly to be true for the market of High-Performance Computing systems (HPC). This market was always characterized by a rapid change of vendors, architectures, technologies and the usage of systems. Despite all these changes the evolution of performance on a large scale however seems to be a very steady and continuous process. Moore’s Law is often cited in this context. If we plot the peak performance of various computers of the last 5 decades in figure 1 which could have been called the ‘supercomputers’ of there time [4,2] we indeed see how well this law holds for almost the complete lifespan of modern computing. On average we see an increase in performance of two magnitudes of order every decade.

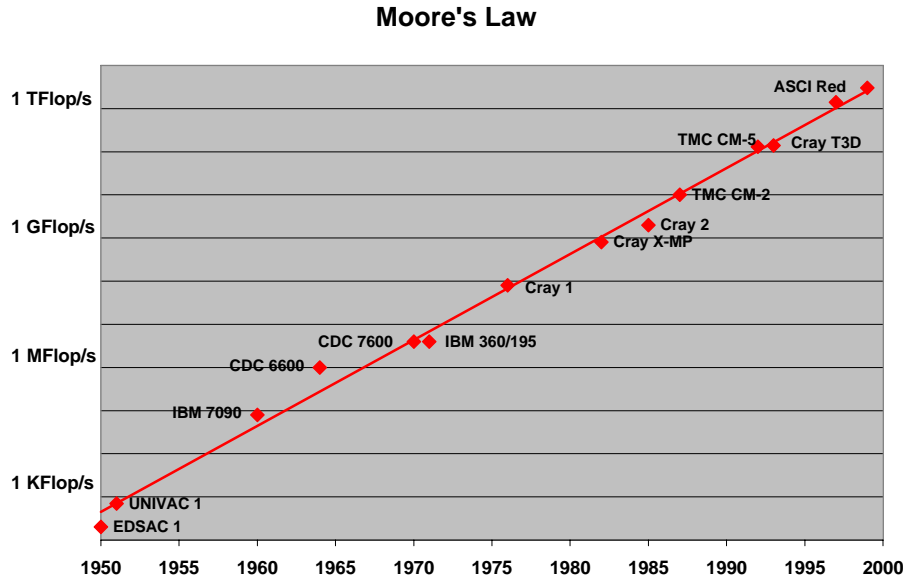


Fig. 1. Performance of the fastest computer systems for the last 5 decades compared to Moore’s Law.

In this paper we analyze the major trends and changes in the HPC market for the last three decades. For this we focus on systems which had at least some commercial relevance. Historical overviews with different focus can be found in [8,9].

In the second half of the seventies the introduction of vector computer systems marked the beginning of modern Supercomputing. These systems offered a performance advantage of at least one order of magnitude over conventional systems of that time. Raw performance was the main if not the only selling argument. In the first half of the eighties the integration of vector system

in conventional computing environments became more important. Only the manufacturers which provided standard programming environments, operating systems and key applications were successful in getting industrial customers and survived. Performance was mainly increased by improved chip technologies and by producing shared memory multi processor systems.

Fostered by several Government programs massive parallel computing with scalable systems using distributed memory got in the focus of interest end of the eighties. Overcoming the hardware scalability limitations of shared memory systems was the main goal. The increase of performance of standard micro processors after the RISC revolution together with the cost advantage of large scale productions formed the basis for the “Attack of the Killer Micro”. The transition from ECL to CMOS chip technology and the usage of “off the shelf” micro processor instead of custom designed processors for MPPs was the consequence.

Traditional design focus for MPP system was the very high end of performance. In the early nineties the SMP systems of various workstation manufacturers as well as the IBM SP series which targeted the lower and medium market segments gained great popularity. Their price/performance ratios were better due to the missing overhead in the design for support of the very large configurations and due to cost advantages of the larger production numbers. Due to the vertical integration of performance it was no longer economically feasible to produce and focus on the highest end of computing power alone. The design focus for new systems shifted to the market of medium performance systems.

The acceptance of MPP system not only for engineering applications but also for new commercial applications especially for database applications emphasized different criteria for market success such as stability of system, continuity of the manufacturer and price/performance. Success in commercial environments is now a new important requirement for a successful Supercomputer business. Due to these factors and the consolidation in the number of vendors in the market hierarchical systems build with components designed for the broader commercial market are currently replacing homogeneous systems at the very high end of performance. Clusters build with components of the shelf also gain more and more attention.

2 1976–1985: The first Vector Computers

If one had to pick one person associated with Supercomputing it would be without doubt Seymour Cray. Coming from Control Data Corporation (CDC) where he had designed the CDC 6600 series in the sixties he had started his own company ‘Cray Research Inc.’ in 1972. The delivery of the first Cray 1

vector computer in 1976 to the Los Alamos Scientific Laboratory marked the beginning of the modern area of ‘Supercomputing’. The Cray 1 was characterized by a new architecture which gave it a performance advantage of more than an order of magnitude over scalar systems at that time. Beginning with this system high-performance computers had a substantially different architecture from main stream computers. Before the Cray 1 systems which sometimes were called ‘Supercomputer’ like the CDC 7600 still had been scalar systems and did not differ in their architecture to this extent from competing main stream systems. For more than a decade supercomputer was a synonym for vector computer. Only at the beginning of the nineties would the MPPs be able to challenge or outperform their MP vector competitors.

2.1 Cray 1

The architecture of the vector units of the Cray 1 was the basis for the complete family of Cray vector systems into the nineties including the Cray 2, Cray X-MP, Y-MP, C-90, J-90 and T-90. Common feature was not only the usage of vector instructions and vector register but especially the close coupling of the fast main memory with the CPU. The system did not have a separate scalar unit but integrated the scalar functions efficiently in the vector cpu with the advantage of high scalar computing speed as well. One common remark about the Cray 1 was often that it was not only the fastest vector system but especially also the fastest scalar system of it’s time. The Cray 1 was also a true Load/Store architecture. A concept which later entered mainstream computing with the RISC processors. In the X-MP and follow on architecture, three simultaneous Load/Store operations per CPU were supported in parallel from main memory without using caches. This gave the systems exceptionally high memory to register bandwidth and facilitated the ease of use greatly.

The Cray 1 was well accepted in the scientific community and 65 systems were sold till the end of it’s production in 1984. In the US the initial acceptance was largely driven by government laboratories and classified sites for which raw performance was essential. Due to it’s potential the Cray 1 soon gained great popularity in general research laboratories and at universities.

2.2 Cyber 205

Main competitor for the Cray 1 was a vector computer from CDC the Cyber 205. This system was based on the design of the Star 100 of which only 4 systems had been build after it’s first delivery in 1974. Neil Lincoln designed the Cyber 203 and Cyber 205 systems as memory to memory machines not using any registers for the vector units. The system also had separate scalar

units. The system used multiple pipelines to achieve high peak performance and a virtual memory in contrast to Cray's direct memory. Due to the memory to memory operation the vector units had rather long startup phases which allowed to achieve high performance only on long vectors.

CDC had been the market leader for high performance systems with its CDC 6600 and CDC 7600 models for many years which gave the company the advantage of a broad existing customer base. The Cyber 205 was first delivered in 1982 and about 30 systems were sold altogether.

2.3 Japanese Vector Systems

End of the seventies the main Japanese computer manufacturers (Fujitsu, Hitachi and NEC) started to develop their own vector computer systems. First models were delivered in late 1983 and mainly sold in Japan. Fujitsu had early decided to sell their vector systems in the USA and Europe through their mainframe distribution partners Amdahl and Siemens. This was the main reason that Fujitsu VP100, and VP200 systems could be found in decent numbers early on in Europe. NEC tried to market their SX1 and SX2 systems by themselves and had a much harder time to find customers outside of Japan. From the beginning Hitachi had decided not to market the S810 system outside of Japan. Common feature of the Japanese systems were separate scalar and vector units and the usage of large vector registers and multiple pipelines in the vector units. The scalar units were IBM 370 instruction compatible which made the integration of these systems in existing computer centers easy. In Japan all these systems were well accepted and especially the smaller models were sold in reasonable numbers.

2.4 Vector Multi-Processor

At Cray Research the next steps to increase performance were not only to increase the performance and efficiency of the single processors but also to build systems with multiple processors. Due to diverging design ideas and emphasis two design teams worked parallel in Chippewa Falls.

Steve Chen designed the the Cray X-MP system first introduced with 2 processors in 1982. The enlarged model with 4 processor was available in 1984. The systems were designed as symmetrical shared memory multi processor systems. The main emphasis of the development was the support of multiple processors. Great effort went into the design of an effective memory access subsystem which was able to support multiple processors with high bandwidth. While the multiple processors were mainly used to increase the throughput of

computing centers, Cray was one of the first to offer a means for parallelization within a user's program using features such as Macrotasking and later on Microtasking.

At the same time Seymour Cray focused at Cray Research on the development of the Cray 2. His main focus was on advanced chip technology and new concepts in cooling. The first model was delivered in 1985. With its four processors it promised a peak performance of almost 2 GFlop/s more than twice as much as a 4 processor X-MP. As its memory was built with DRAM technology the available real main memory reached the unprecedented amount of 2 GByte. This memory size allowed for a long running programs not feasible on any other systems.

The Japanese Vector computer manufacturers decided to follow a different technology path. They increased the performance of their single processors by using advanced chip technology and multiple pipelines. Later the Japanese manufactures announced multiprocessor models typically had 2 or at most 4 processors.

3 1985–1990: The Golden Age of Vector Computers

The class of symmetric multi-vector processor systems dominated the super-computing arena due to its commercial success in the eighties.

3.1 Cray Y-MP - Success in Industry

The follow up of the Cray X-MP, the Cray Y-MP, was a typical example for the sophistication of the memory access subsystems, which was one of the major reason for the overall very high efficiency achievable on these systems. With this product line later including the C-90 and T-90, Cray Research followed the very successful path to higher processor numbers always trying to keep the usability and efficiency of their systems as high as possible. The Cray Y-MP first delivered in 1988 had up to 8 processors, the C-90 first delivered in 1991 up to 16 processors. and the T-90 first delivered in 1995 up to 32 processors. All these systems were produced in ECL chip technology.

Beginning of the eighties the acceptance of the Cray 1 systems was strongly helped by the easy integration in computing center environments of other vendors and by standard programming language support (Fortran77). After 1984 a standard UNIX operating system, UNICOS, was available for all Cray systems which was quite an innovation for a mainframe at that time. With

Year	Cray	CDC	Fujitsu	Hitachi	NEC	Total
1986	118	30	31	8	2	187
1987	148	34	36	9	8	235
1988	178	45	56	11	10	300
1989	235	62	72	17	18	404
1990	248	24	87	27	27	413
1991	268		108	36	35	447
1992	305		141	44	40	530

Table 1
Vector computer installations worldwide [3].

the availability of vectorizing compilers in the second half of the eighties more independent software vendors started to port their key applications on Cray systems which was an immense advantage to sell Cray systems to industrial customers. Due to these reasons Cray vector systems started to have success in industries such as automotive industry and oil industry. Success in these market ensured the dominance of Cray Research in the overall supercomputer market for more than a decade.

Looking at the Mannheim supercomputer statistics [3] in table 1 which counted the worldwide installed vector systems we see the dominance of Cray with a constant market share of 60%. This was confirmed by the first TOP500 list from June 1993 [2] which included MPP systems as well. Cray had an overall share of 40% of all the installed systems which was equivalent to 60% of the included vector systems.

3.2 Cray 3

Seymour Cray left Cray Research Inc. in 1989 to start Cray Computer and to build the follow up to the Cray 2 the Cray 3. Again the idea was to use the most advanced chip technology to push single processor performance to it's limits. The choice of GaAs technology was however ahead of its time and lead to many development problems. In 1992 a single system was delivered. The announced Cray 4 system was never completed.

3.3 *ETA*

In 1983 CDC decided to spin off its supercomputer business in the subsidiary 'ETA Systems Inc'. The ETA10 system was the follow up to the Cyber 205 on which it was based. The CPU's had the same design and the systems had up to 8 processors with a hierarchical memory. This memory consisted of a global shared memory and local memories per processor, all of which were organized as virtual memory. CMOS was chosen as basic chip technology. To achieve low cycle times the high end models had sophisticated cooling system using liquid nitrogen. First systems were delivered in 1987. The largest model had a peak performance of 10 GFlop/s well beyond the competing model of Cray Research.

ETA however seem to have overlooked the fact that raw performance was no longer the only or even most important selling argument. In April 1989 CDC terminated ETA and closed its supercomputer business. One of the main failures of the company being that they overlooked the importance of standard operating system and standard development environments. A mistake which did not only brought ETA but CDC itself down.

3.4 *Mini-Supercomputer*

Due to the limited scalability of existing vector systems there was a gap in performance between traditional scalar mainframes and the vector systems of the Cray class. This market was targeted by some new companies who started in the early eighties to develop the so called mini-supercomputer. Design goal was one third of the performance of the Cray class supers but only one tenth of the price. The most successful of these companies was Convex founded by Steve Wallach in 1982. They delivered the first single processor system Convex C1 in 1985. In 1988 the multiprocessor system C2 followed. Due to the wide software support these systems were well accepted in industrial environments and Convex sold more than 500 of these systems worldwide.

3.5 *MPP - Scalable Systems and the Killer Micros*

In the second half of the eighties a new class of system started to appear - parallel computers with distributed memory. Supported by the Strategic Computing Initiative of the US Defense Advanced Research Agency (DARPA - 1983) a couple of companies started developing such system early in the eighties. Basic idea was to create parallel systems without the obvious limitations in processor number shown by shared memory designs of the vector

multiprocessor.

First models of such massive parallel systems (MPP) were introduced in the market in 1985. At the beginning the architectures of the different MPPs were still quite diverse. Major exception was the connection network as most vendors choose a hypercube topology. Thinking Machine Corporation (TMC) demonstrated their first SIMD system the Connection Machine 1 (CM1). Intel showed their iPSC/1 hypercube system using Intel 80286 cpu and and Ethernet based connection network. nCube produced the first nCube/10 hypercube system with scalar Vax-like custom processors. While these system still were clearly in the stage of experimental machines they formed the basis for broad research on all issues of massive parallel computing. Later generations of these system were then able to compete with vector MP systems.

Due to the conceptual simplicity of the global architecture the number of companies building such machine grew very fast. This included the otherwise rare European efforts to produce supercomputer hardware. Companies who started to develop or produce MPP system in the second half of the eighties include: TMC, Intel, nCube, FPS (Floating Point Systems), KSR (Kendall Square Research), Meiko, Parsytec, Telmat, Suprenum, MasPar, BBN, and others.

3.6 Thinking Machines

After demonstrating the CM-1 (Connection Machine) in 1985 TMC soon introduced the follow on CM-2 which became the first major MPP, designed by Danny Hillis. In 1987 TMC started to install the CM-2 system. The Connection Machine model were single instruction on multiple data (SIMD) systems. Up to 64k single-bit processors connected in a hypercube network together with 2048 Weitek floating point units could work together under the control of a single front-end system on a single problem. The CM-2 was the first MPP system which was not only successful in the market but which also could challenge the vector MP systems of it's time (Cray Y-MP), at least for certain applications.

The success of the CM-2 was great enough that another company, MasPar, which started producing SIMD systems as well. It's first system the MasPar MP-1 using 4-bit processors was first delivered in 1990. The follow on model MP-2 with 8-bit processors was first installed in 1992.

Main disadvantage of all SIMD system however proved to be the limited flexibility of the hardware which limited the number of applications and programming models which could be supported. Consequently TMC decided to design their next major system the CM-5 as MIMD system. To satisfy the existing

customer base this system could run data-parallel program as well.

3.7 Early MPPs

Competing MPP manufacturer had from the start decided to produce MIMD systems. The more complex programming of these systems was more than compensated by the much greater flexibility to support different programming paradigms efficiently.

Intel build systems based on the different generations of Intel micro processors. The first system iPSC/1 was introduced in 1985 and used Intel 80286 processors with an Ethernet based connection network. The second model iPSC/2 used the 80386 and already had a circuit switched routing network. The iPSC/860 introduced in 1990 finally featured the i860 chip. For Intel massive parallel meant up to 128 processors which was the limit due to the maximum dimension of the connection network.

In contrast to using standard off-the-shelf microprocessor nCube had designed their own custom processor as basis for their nCube/10 system introduced in 1985 as well. The design of the processor was similar to the good old Vax and therefore a typically CISC design. To compensate for the relatively small performance of this processor the maximal number of processors possible was however quite high. Limitation was again the dimension 13 of the hypercube network which would have allowed up to 8096 processors. The follow-up nCube/2 again using this custom processor was introduced in 1990.

3.8 Attack of the Killer Micros

Beginning of the nineties one phrase showed up on the front pages of several magazines: “The Attack of the Killer Micro”. Coined by Eugene Brooks from Livermore National Laboratory this was the synonym for the the greatly increased performance levels achievable with microprocessors after the RISC/superscalar design revolution. The performance of micro processors seemed to have reached comparable level to the much more expensive ECL custom mainframe computer. But not only the traditional mainframes started to feel the heat, even the traditional supercomputer the vector multi-processors got under attack.

Another slogan for this process was widely used when Intel introduced it’s answer to the RISC processors, the i860 chip: “Cray on a chip” Even as sustained performance values did not always come up to PAP (peak advertised performance) values the direction of the attack was clear. The new genera-

tion of microprocessors manufactured relatively cheap in great numbers for the workstation market offered the much better price/performance ratios. Together with the scalable architecture of MPPs the same high performance levels as with vector multiprocessor could be achieved for a better price.

Aggravated was this contrast greatly by the fact that most mainframe manufacturers had not seen early enough the advantage of CMOS chip technology and were still using the (little) faster but much more expensive ECL technology. Cray Research was no exception in this respect. Under great pressure all mainframe manufacturers started to switch from ECL to CMOS. At the same time they also started to produce their own MPP systems to have competing products with the up and coming new MPP vendors.

Hidden behind these slogans were actually two different trend working together, both of which effects can clearly be seen in the TOP500 data. The replacement of ECL chip technology by CMOS is shown in figure 2. The rapid decline of vector based systems in the nineties can be seen in the TOP500 in figure 3. This change in the processor architecture is however not as generally accepted as the change from ECL to CMOS. The Japanese Manufacturers together with SGI still continue to produce vector based systems.

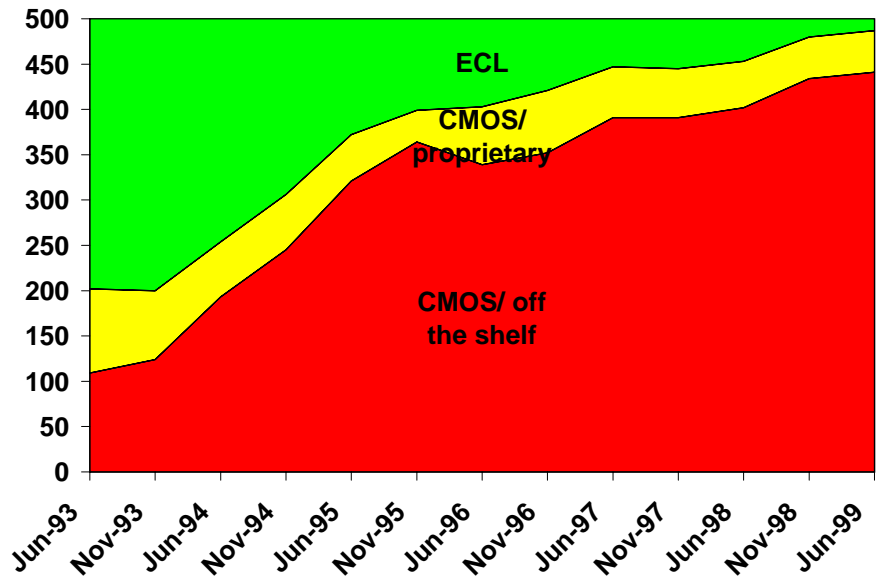


Fig. 2. Chips technologies used as seen in the TOP500.

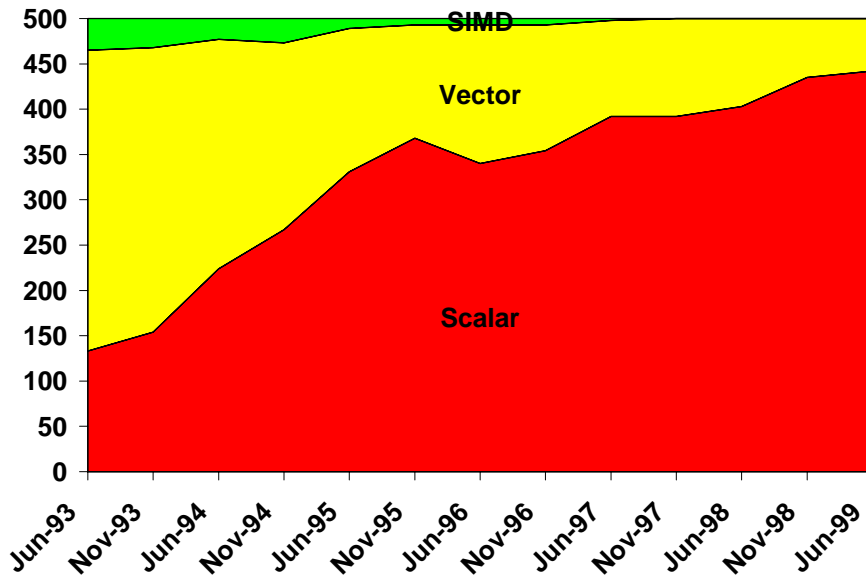


Fig. 3. CPU design used as seen in the TOP500.

4 1990–1995: MPP come to age

Beginning of the nineties while the MP vector systems reached their widest distribution, a new generation of MPP system came on the market with the claim to be able to substitute or even surpass the vector MPs. The increased competitiveness of MPPs made it less and less meaningful to compile ‘Supercomputer’ statistics by just counting vector computers. This was the major reason for starting the TOP500 project[2]. Here we list since 1993 twice a year the 500 most powerful installed computer systems ranked by the best LINPACK performance [5]. In the first TOP500 list in June 1993 there were already 156 MPP and SIMD systems present (31% of the total 500 systems).

The hopes of all the MPP manufacturers to grow and gain market share however did not come true. The overall market for HPC systems did grow only slowly and mostly in directions not anticipated by the traditional MP vector or MPP manufacturers. The attack of killer micros went into its next stage. This time the large scale architecture of the MPPs seen before would be under attack.

One major side-effect of the introduction of MPP system in the market for supercomputer was that the sharp performance gap between supercomputers and mainstream computers no longer existed. MPP could - by definition - be scaled by one or two magnitudes of order bridging the gap between high-end multi-processor workstations and supercomputers. Homogeneous and mono-

lithic architectures designed for the very high end of performance would have to compete with clustered concepts based on shared memory multi-processor models from the traditional UNIX server manufacturer. These cluster offered another level of price/performance advantage due to the large supporting industrial and commercial business market in the background and due to the reduced design cost not focusing on the very high end of performance any more. This change towards the usage of standard components widely used in the commercial market place actually widened the market scope for such MPPs in general. As a result the notion of a separated market for floating point intensive high-performance computing no longer holds. The HPC market is nowadays the upper end of a continuum of systems and usage in all kind of applications.

4.1 Playground for Manufacturers

With the largely increased number of MPP manufacturers it was evident that a “shake-out” of manufactures was unavoidable. In table 2 we list vendors which have been active at some point in the HPC market [6,1]. In Figure 4 we try to visualize the historic presence of companies in the HPC market. After 1993 we included only companies which had entries in the TOP500 and were actively manufacturing.

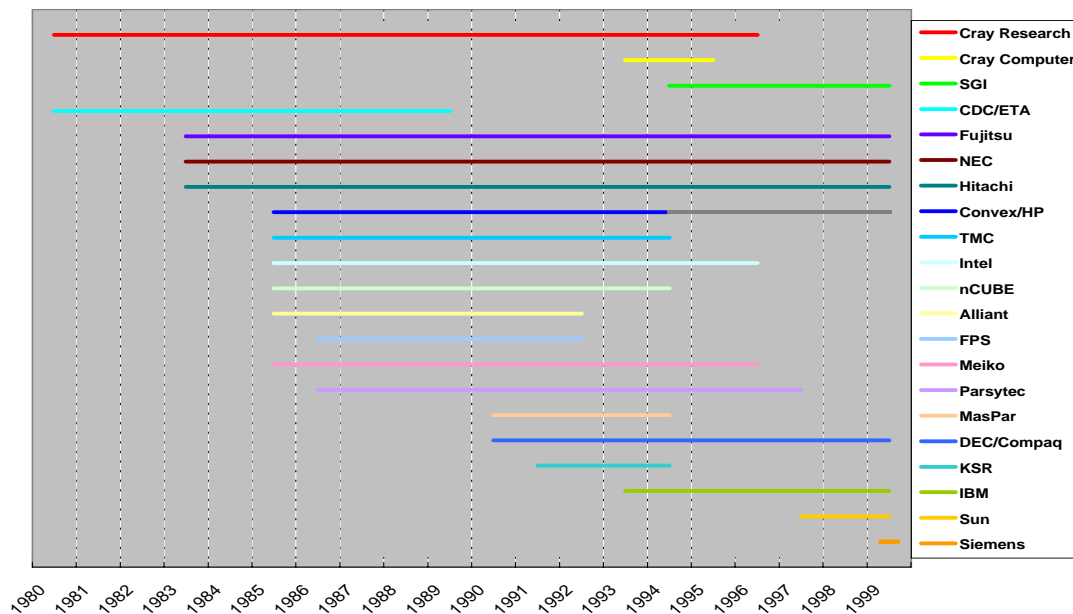


Fig. 4. Main manufacturer active in the HPC market.

From the 14 major companies in the early nineties only 4 survived this decade

Status	Vendors
Out of business	Alliant, American Supercomputer, Ametek, AMT(Cambridge), Astronautics, BBN Supercomputer, Biin, CDC/ETA Systems, Chen Systems, CHoPPCHoPP, Cogent, Cray Computer, Culler, Cydrome, Denelcor, Elexsi, Encore, E&S Supercomputers, Flexible, Goodyear, Gould/SEL, Intel Supercomputer Division, IPM, iP-Systems, Key, Kendall Square Research, Multiflow, Myrias, Pixar, Prevec, Prisma, Saxpy, SCS, SDSA, Suprenum, Stardent (Stellar and Ardent), Supercomputer Systems Inc., Synapse, Thinking Machines, Trilogy, VIttec, Vitesse, Wavetracer
Merged	Celerity, Convex (with Hewlett-Packard), Cray Research (with SGI), DEC (with Compaq), Floating Point Systems (with Cray Research), Key, MasPar (with DEC), Meiko, Supertek (with Cray Research)
Changed market	nCUBE, Parsytec
Currently active	Compaq, Fujitsu, Hewlett-Packard, Hitachi, IBM, NEC, SGI, Siemens, Sun, Tera

Table 2
Commercial HPC Vendors

on their own. These are the 3 Japanese vector manufacturer (Fujitsu, Hitachi, and NEC) and IBM which due to it's marginal HPC presence at the beginning of the nineties could even be considered a newcomer. Four other companies entered the HPC market either by buying some companies or by developing their own products (Silicon Graphics, Hewlett-Packard, Sun, Compaq). None of these is a former HPC manufacturer. All are typical workstation manufacturer which entered the HPC market (at least initially) from the lower end with high-end UNIX-server models. Their presence and success already indicates the change in focus from the very high end to markets for medium size HPC systems.

4.2 *Kendall Square Research*

In 1991 first models of a quite new and innovative systems were installed. The KSR1 from Kendall Square Research. The hardware was build like other MPP with distributed memory but gave the user the view of a shared memory system. The custom design hardware and the operating system software were responsible for this virtual shared memory appearance. This concept of virtual shared memory could later on be found in other systems such as the Convex SPP series and lately in the SGI Origin series. The KSR systems organized the complete memory on top of the VSM as cache only memory. By this the data had no fixed home in the machine and could freely room to the location where they were needed. ‘Management mistakes’ brought the operations of KSR to an abrupt end in late 1994.

4.3 *Intel*

In 1992 Intel started to produce the Paragon/XP series after delivering the Touchstone Delta system to Caltech in 1991. Still based on the i860 chips the interconnection network was changed to a two dimensional grid which now allowed up to 2048 processors. Several quite large system were installed and in June 1994 a system at Sandia National Laboratory which achieved 143 GFlop/s on the LINPACK benchmark was the number one in the TOP500. Intel decided to stop it’s general HPC business in 1996 but still build the ASCI Red system afterwards.

4.4 *Thinking Machines*

Also in 1992 TMC started to deliver the CM-5 a MIMD system designed for the very high end. Theoretically systems up to 16k processors could be build. In practice the largest configurations reached 1056 processors at the Los Alamos National Laboratory. This system was at the number one spot of the very first TOP500 list in June 1993 achieving almost 60 GFlop/s. A Sun SPARC processor was chosen as basic node cpu. Each of these processors had 4 vector coprocessors to increase the floating-point performance of the system. Initial programming paradigm was the data-parallel model familiar from the CM-2 predecessor. The complexity of this node design however was more than the company or customers could handle. Due to the design point being the very high end the smaller models also had problems competing with models from other companies which did not have to pay for the overhead of supporting such large systems in their design. The CM-5 would be the last

TMC model before the company had to stop the production of hardware in 1994.

The raw potential of the CM-5 was demonstrated by the fact that in June 1993 the position 1–4 were all held by TMC CM-5 systems ahead of the first MP vector system an NEC SX-3/44R. The only other MPP system able to beat a Cray C-90 at that time was the Intel Delta system. The performance leadership however had started to change.

In June 1993, still 5 of the first 10 system were MP vector systems. This number decreased fast and the last MP vector system which managed to make the top 10 was a NEC SX-4 with 32 processors in June 1996 with 66.5 GFlop/s. Later only systems with distributed memory made the top 10 list. Japanese MPPs with vector cpu managed to keep their spot in the top 10 for some time. In the November 1998 list however the top 10 positions were for the first time all taken by microprocessor based ASCI or Cray T3E systems. The first system with vector CPU's was a NEC SX-4 with 128 processor at number 18.

4.5 IBM

In 1993 IBM finally joined the field of MPP producers by building the IBM SP1 based on their successful workstation series RS6000. While this system was often mocked being a workstation cluster and not a MPP it set the ground for the reentry of IBM in the supercomputer market. The follow on SP2 with increased node and network performance was first delivered in 1994. Contrary to other MPP manufacturers IBM was focusing on the market for small to medium size machines especially for the commercial UNIX server market. Over time this proved to be a very profitable strategy for IBM who managed to sell models of the SP quite successful as a database server. Due to the design of the SP2, IBM is able to constantly offer new nodes based on the latest RISC system available.

4.6 Cray Research

In 1993 Cray Research finally started to install their first MPP system, the Cray T3D. As indicated by the name the network was a 3 dimensional torus. As CPU Cray had chosen the DEC alpha processor. The design of the node was completely done by Cray itself and was substantially different from a typical workstation using the same processor. This had advantages and disadvantages. Due to their closely integrated custom network interface the network latencies and the bandwidth reached values not seen before and allowed very efficient parallel processing. The computational node performance itself was however

greatly affected by the missing 2nd level cache. The system was immediately well accepted at research laboratories and was even installed at some industrial customer sites. The largest configuration known is installed at a classified government site in the USA with 1024 processors and just breaking the 100 GFlop/s barrier on the LINPACK.

4.7 Convex

In 1994 Convex introduced its first true MPP, the SPP1000 series. This series was also awaited with some curiosity as it was after KSR the second commercial system featuring a virtual shared memory. The architecture of the system was hierarchical. Up to 8 HP microprocessors were connected to a shared memory with crossbar technology similar to the one used in the Convex vector series. Multiple of these SMP units would then be connected in a distributed memory fashion. The operating system and the connection hardware would provide the view of a non-uniform shared memory over the whole machine. In the following years a series of follow on models was introduced the SPP1200 in 1995, the SPP1600 in 1996, and the SPP2000 renamed by HP as Exemplar X-Class in 1997.

4.8 The role of MP vector systems during 1990–1995

Cray Research continued building their main MP vector system in traditional style. The Triton, known as T-90, was introduced in 1995 and built in ECL very much along the line of the Y-MP and C-90 series. The maximum number of processors was increased to 32. This gave a full system a peak performance of 58 GFlop/s. Realizing that it needed a product for lower market segment Cray had bought the company Supercomputer which had developed Cray Y-MP compatible vector systems in CMOS technology. It was marketed by Cray starting in 1993. The next system in this series developed by Cray Research itself was the J-90 introduced in 1995 as well. With up to 32 processors it reached a peak performance of 6.4 GFlop/s which was well below the ECL systems from Cray and unfortunately not much above the performance of best microprocessor available.

Convex introduced the C3 series in 1991 and the C4 in 1994 before the company was bought by Hewlett-Packard the same year. After this merger the unit focused on its MPP products.

In Japan, Fujitsu had introduced the single processor VP2600 in 1990 and would market this series till the introduction of the CMOS based VPP500 in 1994. NEC introduced the SX-3 series in 1990 as well. This system reached

with it's up to four processor a peak performance of 26 GFlop/s. NEC subsequently implemented their vector series in CMOS and introduced the NEC SX-4 in 1994. Up to 32 processors can be installed as conventional shared memory MP vector system. Beyond this up to 16 of these units can be clustered in a distributed memory fashion. The largest configurations known to be installed have 128 processors with which they gained positions 29 and 30 in the June 1999 TOP500 list. These are at present the largest vector based systems with traditional design.

4.9 Fujitsu's MPP-Vector Approach

Fujitsu decided to go it's own way into the world of MPPs. They build their commercial MPP system with distributed memory around the node and cpu of their successful VP2600 vector computer series. However Fujitsu was the first Japanese company who implementing their vector design in CMOS, the VPP500. A first 'pre-prototype' was developed together with the National Aerospace Laboratories (NAL). The installation of this system named the Numerical Wind Tunnel (NWT) started in 1993. Due to it's size this system managed to gain the number 1 position in the TOP500 an unchallenged four and number 2 position three times from November 1993 to November 1996. Delivery of VPP500 systems started in 1993.

5 After 1995: Old and New Customers

The year 1995 saw some remarkable changes in the distribution of the systems in the TOP500 for the different types of customer (academic sites, research labs, industrial/commercial users, vendor installations, and confidential sites) (see Fig. 5).

Until June 1995, the major trend seen in the TOP500 data was a steady decrease of industrial customers, matched by an increase in the number of government-funded research sites. This trend reflects the influence of the different governmental HPC programs that enabled research sites to buy parallel systems, especially systems with distributed memory. Industry was understandably reluctant to follow this step, since systems with distributed memory have often been far from mature or stable. Hence, industrial customers stayed with their older vector systems, which gradually dropped off the TOP500 list because of low performance.

Beginning in 1994, however, companies such as SGI, Digital, and Sun started to sell symmetrical multiprocessor (SMP) models of their major workstation

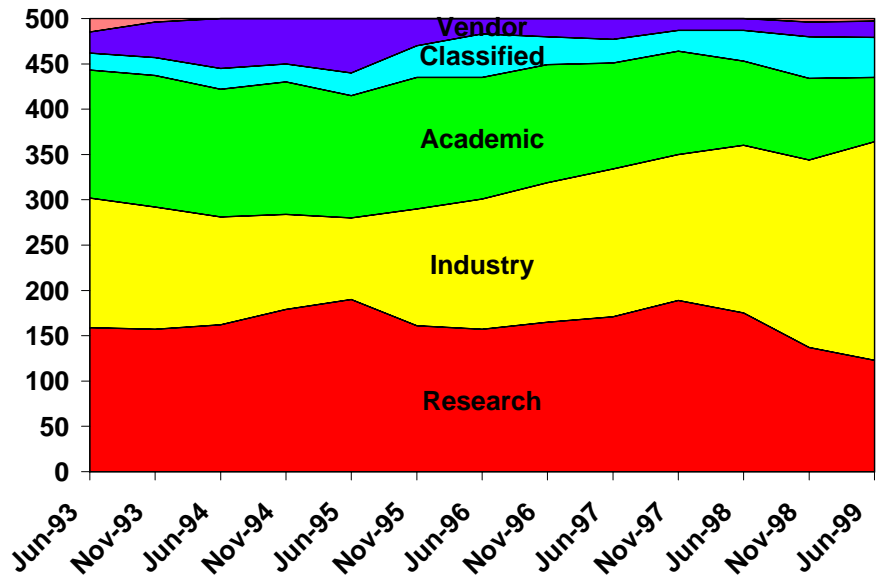


Fig. 5. The number of systems on the different types of customers over time.

families. From the very beginning, these systems were popular with industrial customers because of the maturity of these architectures and their superior price/performance ratio. At the same time, IBM SP2 systems started to appear at a reasonable number of industrial sites. While the SP initially was sold for numerically intensive applications, the system began selling successfully to a larger market, including database applications, in the second half of 1995.

Subsequently, the number of industrial customers listed in the TOP500 increased from 85, or 17%, in June 1995 to about 241, or 48.2%, in June 1999. We believe that this is a strong new trend because of the following reasons.

- The architectures installed at industrial sites changed from vector systems to a substantial number of MPP systems. This change reflects the fact that parallel systems are ready for commercial use and environments.
- The most successful companies (Sun, IBM and SGI) are selling well to industrial customers. Their success is built on the fact that they are using standard workstation technologies for their MPP nodes. This approach provides a smooth migration path for applications from workstations up to parallel machines.
- The maturity of these advanced systems and the availability of key applications for them make the systems appealing to commercial customers. Especially important are database applications, since these can use highly parallel systems with more than 128 processors.

Figure 6 shows that the increase in the number of systems installed at indus-

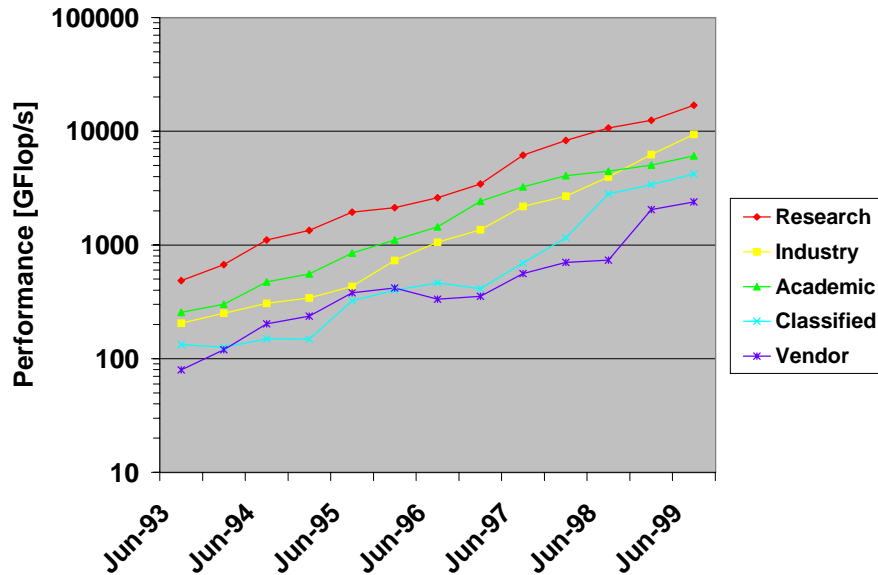


Fig. 6. The accumulated performance of the different types of customers over time.

trial sites is matched by a similar increase in the installed accumulated performance. The relative share of industrial sites rose from 8.7% in June 1995 to 24.0% in June 1999. Thus, even though industrial systems are typically smaller than systems at research laboratories and universities, their average performance and size are growing at the same rate as at research installations. The strong increase in the number of processors in systems at industrial sites is another major reason for the rise of industrial sites in the TOP500. The industry is ready to use bigger parallel systems than in the past.

5.1 Architectures

The changing share of the different system architectures in the HPC market as reflected in the TOP500 is shown in figure 7. Besides the fact that no single processor systems are any longer powerful enough to enter the TOP500, the major trend is the growing number of MPP systems. The number of clustered systems is also growing and for the last two years we see a number of PC or workstation based ‘Network of Workstations’ in the TOP500. It is an interesting and open question which share of the TOP500 such NOWs will capture in the future.

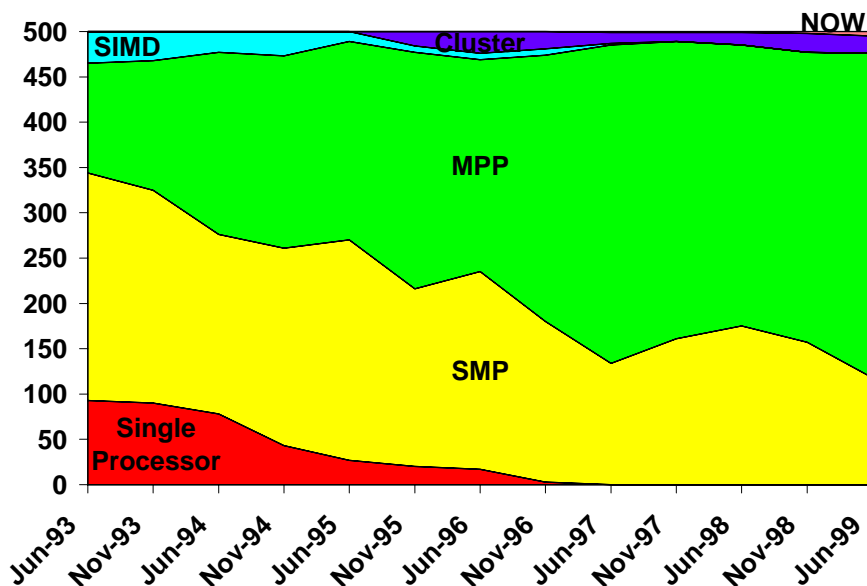


Fig. 7. Architectures of systems in the TOP500.

5.2 Vector based Systems

Cray Research introduced their last ECL-based vector system the T-90 series in 1995. Due to the unfavorable price/performance of this technology the T-90 was not an economical success for Cray Research. One year later in 1996 SGI bought Cray Research. After this acquisition the future of the Cray vector series was in doubt. The joint company announced plans to produce a joint macro architecture for its microprocessor and vector processor based MPPs. In mid 1998 the SGI SV1 was announced as future vector system of SGI. The SV1 is the successor both to the CMOS based Cray J-90 and the ECL based Cray T-90. The SV1 is CMOS based which means that SGI is finally following the trend set in by Fujitsu (VPP700) and NEC (SX-4) a few years earlier. First user shipments are expected by the end of 1999. One has to wait and see whether the SV1 can compete with the advanced new generation of Japanese vector systems, especially the NEC SX-5 and the Fujitsu VPP5000.

Fujitsu continued along the line of the VPP system and introduced in 1996 the VPP700 series featuring increased single node performance. For the lower market segment the VPP300 using the same nodes but a less expandable interconnect network was introduced. The recently announced VPP5000 is again a distributed-memory vector system, where 4 up to 128 (512 by special order) processors may be connected via a fully distributed crossbar. The theoretical peak performance is ranges from 38.4 GFlop/s up to 1.229 TFlop/s, in special configurations even 4.915 TFlop/s.

NEC had announced the SX-4 series in 1994 and continues along this system line. The SX-4 features shared memory up to a maximum of 32 processors. Larger configurations are build as cluster using a proprietary crossbar switch. In 1998 the follow up model SX-5 was announced for first delivery in early 1999. In 1998 the follow up model SX-5 was announced for first delivery in late 1998 and early 1999. In the June 99 TOP500 list the SX-5 was not yet represented. In contrast to its predecessor, the SX-4, the SX-5 is not offered anymore with faster, but more expensive SRAM memory. The SX-5 systems are exclusively manufactured with synchronous DRAM memory. The multi-frame version of the SX-5 can host up to 512 processors with 8 GFlop/s peak performance each, resulting in a theoretical peak performance of 2 TFlop/s. More information about all these current architectures can be found in [7].

5.3 Traditional MPPs

Large scale MPPs with homogeneous system architectures had matured during the nineties with respect to performance and usage. Cray finally took the leadership here as well with the T3E system series introduced in 1996 just before the merger with SGI. The performance potential of the T3E can be seen by the fact that in June 1997, 6 of the top 10 positions in the TOP500 were occupied by T3E's. End of 1998 the top 10 consisted only of ASCI systems and T3E's.

Hitachi was one of the few companies introducing large scale homogeneous system in the late nineties. It announced the SR2201 series in 1996 and tries to sell this system outside of Japan as well.

The first of the ASCI system, the ASCI Red at Sandia National Laboratory, was delivered in 1997. It took immediately the first position in the TOP500 in June 1997 being the first system to exceed 1 TFlop/s LINPACK performance. ASCI Red also ended several years during which several Japanese systems ranked as number one.

IBM followed the Path of their SP series introducing new nodes and faster interconnects as well. One major innovation here was the usage of SMP nodes as building blocks which further demonstrates the proximity of the SP architecture to clusters. This design with SMP nodes was also chosen for the ASCI Blue Pacific systems.

5.4 *SMPs and their successors*

Beginning in 1994, however, companies such as SGI, Digital, and Sun started to sell symmetrical multiprocessor (SMP) models of their major workstation families. From the very beginning, these systems were popular with industrial customers because of the maturity of these architectures and their superior price/performance ratio. At the same time, IBM SP2 systems started to appear at a reasonable number of industrial sites. While the SP initially was sold for numerically intensive applications, the system began selling successfully to a larger market, including database applications, in the second half of 1995.

SGI made a strong appearance in the TOP500 in 1994 and 1995. Their PowerChallenge systems introduced in 1994 sold very well in the industrial market for floating point intensive applications. Cluster build with these SMPs appeared in a reasonable number at customer sites.

In 1996 the Origin2000 series was announced. With this system SGI took the step away from the bus based SMP design of the Challenge series. The Origin series features a virtual memory system build with distributed memory nodes up to 128 processors. To achieve higher performance these systems can be clustered again. This is the basic design of the ASCI Blue Mountain system.

Digital was for a long time active as producer of clustered systems for commercial customers. In 1997 the Alpha Server Cluster was introduced which was targeted towards floating point intensive applications as well. One year later Compaq acquired Digital giving Compaq as first PC manufacturer an entry in the HPC market.

Hewlett-Packard continued producing systems along the line of the former Convex SPP systems targeting mainly the midsize business market where the company has good success. The very high end - which had never been a target for Convex or HP - still seems to be of minor interest to HP.

Sun was the latest company who entered the TOP500. After the merger of SGI and Cray Research in 1996, Sun bought the former business server division of Cray which produced SPARC based SMP systems for several years. In 1997 Sun introduced the HPC 10000 series. This SMP system is build around a new type of switched bus which allows to integrate up to 64 processors in an efficient way. To due it's wide customer base and good reputation in the commercial market Sun was able to sell these SMPs very well especially to commercial and industrial customers. For the very high end market clusters build with these SMP were introduced in 1998.

5.5 New Application Areas

For research sites or academic installations, it is often difficult — if not impossible — to specify a single dominant application. The situation is different for industrial installations, however, where systems are often dedicated to specialized tasks or even to single major application programs. Since the very beginning of the TOP500 project, we have tried to record the major application area for the industrial systems in the list. We have managed to track the application area for almost 90% of the industrial systems over time.

Since June 1995 we see many systems involved in new application areas entering the list. Figure 8 shows the total numbers of all industrial systems which is made up of three components, traditional floating point intensive engineering applications, new non floating point applications, and unknown application areas.

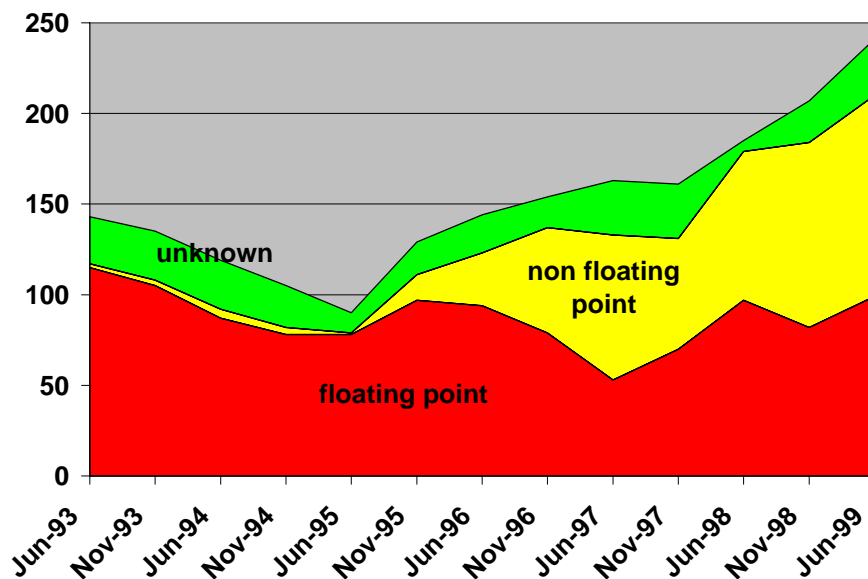


Fig. 8. The total number of systems at industrial sites together with the numbers of sites with traditional engineering applications, new emerging application areas and unknown application areas.

In 1993, the applications in industry typically were numerically-intensive applications, for example,

- geophysics and oil applications,
- automotive applications,
- chemical and pharmaceutical studies,
- aerospace studies,

- electronics, and
- other engineering including energy research, mechanical engineering etc.

The share of these areas from 1993 to 1996 remained fairly constant over time. Recently, however, industrial systems in the TOP500 have been used for new application areas. These include

- database applications,
- finance applications, and
- image processing.

The most dominant trend is the strong rise of database applications since November 1995. These applications include on-line transaction processing as well as data mining. The HPC systems being sold and installed for such applications are large enough to enter the first hundred positions—a clear sign of the maturity of these systems and their practicality for industrial usage.

It is also important to notice that industrial customers are buying not only systems with traditional architectures, such as the SGI PowerChallenge or Cray T-90, but also MPP systems with distributed memory, such as the IBM SP2. Distributed memory is no longer a hindrance to success in the commercial marketplace.

6 2000 and beyond

Two decades after the introduction of the Cray 1 the HPC market has changed its face quite a bit. It used to be a market for systems clearly different from any other computer systems. Nowadays the HPC market is no longer an isolated niche market for specialized systems. Vertically integrated companies produce systems of any size. Similar software environments are available on all of these systems. This was the basis for a broad acceptance at industrial and commercial customers.

The increasing market share of industrial and commercial installations had several very critical implications for the HPC market. The manufacturers of supercomputers for numerical applications face in the market for small to medium size HPC systems the strong competition of manufacturers selling their systems in the very lucrative commercial market. These systems tend to have better price/performance ratios due to the larger production numbers of systems accepted at commercial customers and the reduced design costs of medium size systems. The market for the very high end systems itself is relatively small and does not grow. It can not easily support specialized niche market manufacturers. This forces the remaining manufacturers to change

the design for the very high end away from homogeneous large scale systems towards cluster concepts based on medium size 'off-the-shelf' systems.

Currently the debate still goes on if we need a new architecture for the very high end supercomputer such as the multithreaded design of Tera. At the same time we see a new architecture appearing in growing number in the TOP500 - the 'Network of Workstations' including PC based systems. Depending where one draws the line between a cluster of SMPs and a network of workstation we have about 6 such NOWs in the June 1999 edition of the TOP500. As there is not a single manufacturer who provides LINPACK measurements for such systems we certainly underestimate the actual number of NOW large enough to fit in the TOP500. The potential of this class of HPC architecture with excellent price/performance ratios should not be underestimated and we expect to see more in the future. If they however will seriously challenge the clusters of large SMPs for a wide range of application is an open question.

6.1 Dynamic of the Market

The HPC market is by its very nature very dynamic. This is not only reflected by the coming and going of new manufacturers but especially by the need to update and replace systems quite often to keep pace with the general performance increase. This general dynamic of the HPC market is well reflected in the TOP500. In table 3 we show the number of systems which fall off the end of the list within 6 months due to the increase in the entry level performance. We see an average replacement rate of about 160 systems every half year or more than half the list every year. This means that a system which is at position 100 at a given time will fall off the TOP500 within 2-3 years.

6.2 Consumer and Producer

The dynamic of the HPC market is well reflected in the rapidly changing market shares of the used chip or system technologies, of manufacturers, customer types or application areas. If we however are interested in where these HPC systems are installed or produced we see a different picture.

Plotting the number of systems installed in different geographical areas in figure 9 we see a rather steady distribution. The number of systems installed in the US is slightly increasing over time while the number of systems in Japan is slowly decreasing.

Looking at the producers of HPC systems in figure 10 we see an even greater dominance of the US, which actually slowly increases over time. European

List	Last System	Entry level R_{max} [GFlop/s]	Replaced Systems
6/93	Fujitsu VP200	0.422	
11/93	Fujitsu VP200EX	0.472	83
6/94	Cray X-MP/4	0.822	121
11/94	Cray Y-MP/M4	1.114	115
6/95	SGI PC/8	1.955	216
11/95	Cray C94/3	2.489	144
6/96	Convex SPP1000/32	3.306	137
11/96	SGI PC/18	4.620	183
6/97	IBM SP2/32	7.680	245
11/97	Sun Ultra HPC 10000	9.513	129
6/98	Sun HPC 10000	13.390	180
11/98	Sun HPC 10000	17.120	162
6/99	SGI T3E900	24.730	194
Average			159

Table 3

The replacement rate in the TOP500 defined as number of systems omitted because of their performance being too small.

manufacturers do not play any substantial role in the HPC market at all.

6.3 Government programs

The high end of the HPC market was always the target for government program all over the world to influence the further development of new systems. In the USA there are currently several government projects on the way to consolidate and advance the numerical HPC capabilities of US government laboratories and the US research community in general. The most prominent of these is the ‘Accelerated Strategic Computing Initiative (ASCI)’. Goal of this program is “to create the leading-edge computational modeling and simulation capabilities that are essential for maintaining the safety, reliability, and performance of the U.S. nuclear stockpile and reducing the nuclear danger”⁵.

Three main laboratories were selected as sites for deploring parallel computing

⁵ www.llnl.gov/asci/

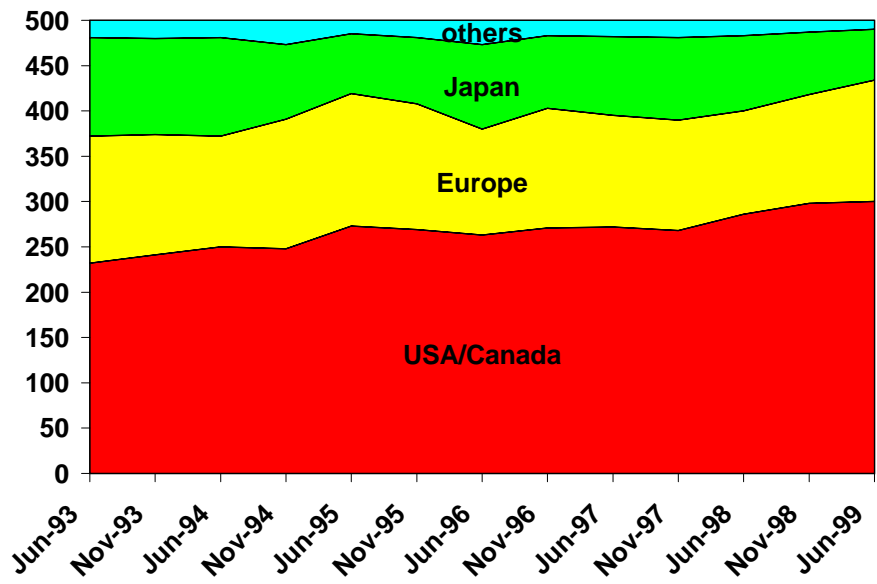


Fig. 9. The consumers of HPC systems in different geographical regions as seen in the TOP500.

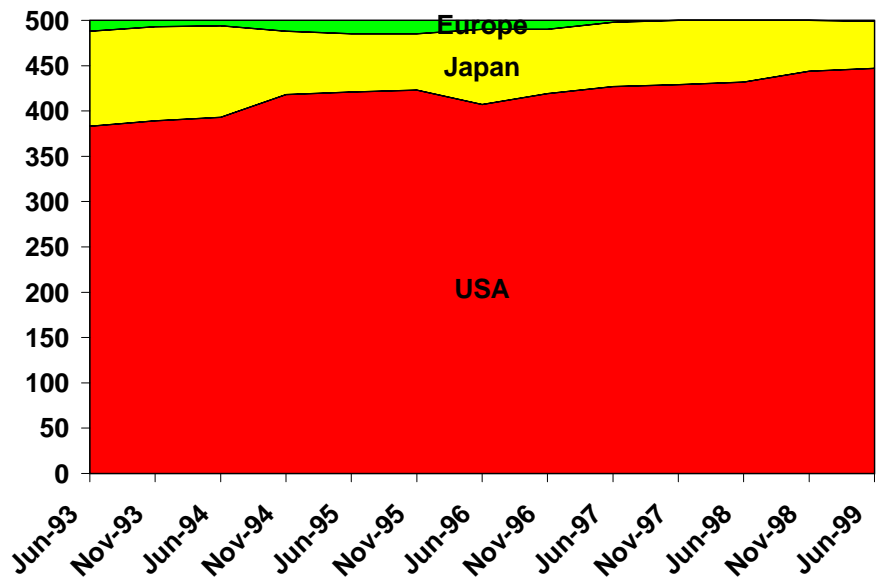


Fig. 10. The producers of HPC systems as seen in the TOP500.

systems of the largest scale technically possible. Currently the system ‘ASCI Red’ is installed at Sandia National Laboratory. This system is produced by Intel and has currently 9472 Intel Pentium Xeon processors. It was the first system to exceed the 1 TFlop/s mark on the LINPACK benchmark in 1997. Since then ASCI Red is the number 1 on the TOP500 with currently 2.1 TFlop/s. ‘ASCI Blue Mountain’ installed at the Los Alamos National Laboratory is a cluster of Origin2000 systems produced by SGI with a total of 6144 processors. It achieves 1.6 TFlop/s LINPACK performance. ‘ASCI Blue Pacific’ installed at the Lawrence Livermore National Laboratory is an IBM SP system with a total of 5856 processors. The future plans of the ASCI program call for a 100 TFlop/s system by April 2003.

The Japanese government decided to fund the development of an ‘Earth Simulator’ to simulate and forecast the global environment. In 1998 NEC was awarded the contract to develop a 30 TFlop/s system to be installed by 2002.

6.4 Performance Growth

While many aspects of the HPC market change quite dynamically over time, the evolution of performance seems to follow quite well some empirical law’s such as Moore’s law mentioned at the beginning of this article. The TOP500 provides an ideal data basis to verify an observation like this. Looking at the computing power of the individual machines present in the TOP500 and the evolution of the total installed performance, we plot the performance of the systems at positions 1, 10, 100 and 500 in the list as well as the total accumulated performance. In Figure 11 the curve of position 500 shows on the average an increase of a factor of two within one year. All other curves show a growth rate of 1.8 ± 0.07 per year.

To compare these growth rates with Moore’s Law we now separate the influence from the increasing processor performance and from the increasing number of processor per system on the total accumulated performance. To get meaningful numbers we exclude the SIMD systems for this analysis as they tend to have extreme high processor numbers and extreme low processor performance. In Figure 12 we plot the relative growth of the total processor number and of the average processor performance defined as the quotient of total accumulated performance by the total processor number. We find that these two factors contribute almost equally to the annual total performance growth factor of 1.82. The processor number grows per year on the average by a factor of 1.30 and the processor performance by 1.40 compared 1.58 of Moore’s Law.

The average growth in processor performance is lower than we expected. A

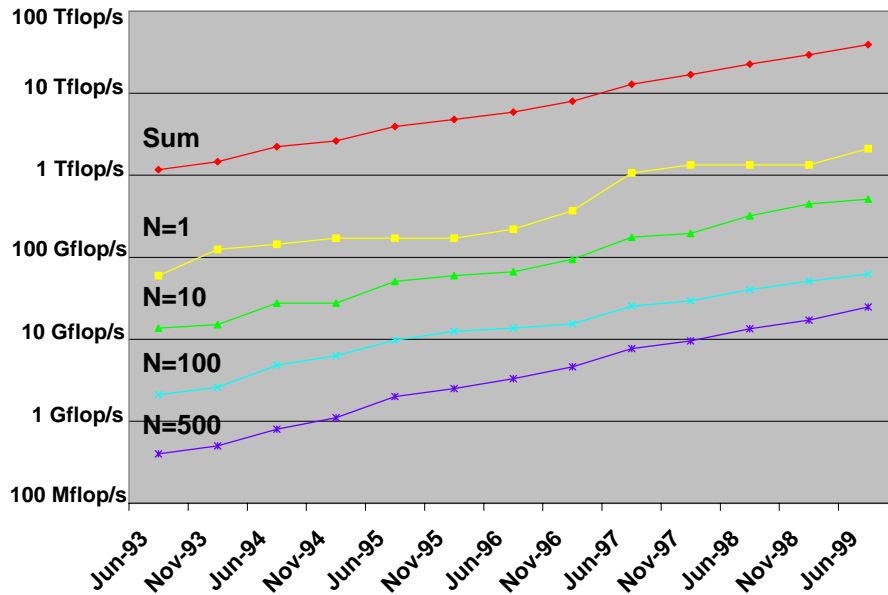


Fig. 11. Overall growth of accumulated and individual performance as seen in the TOP500.

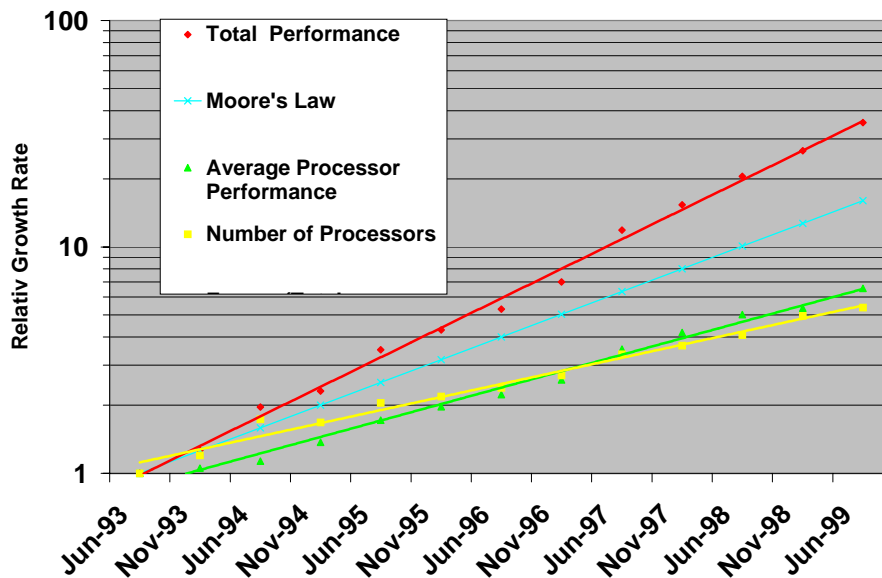


Fig. 12. The growth rate of total accumulated performance, total number of processors and single processor performance for non-SIMD systems as seen in the TOP500.

possible explanation is that during the last years powerful vector processors got replaced in the TOP500 by less powerful super-scalar RISC processor. This effect might be the reason why the TOP500 currently does not reflect the full increase in RISC performance. Once the ratio of vector to super-scalar processors in the TOP500 stabilizes we should see the full increase in processor performance reflected in the TOP500. The overall growth of system performance is however larger than expected from Moore's Law. This comes from growth in the two dimensions processor performance and number of processors used.

6.5 Projections

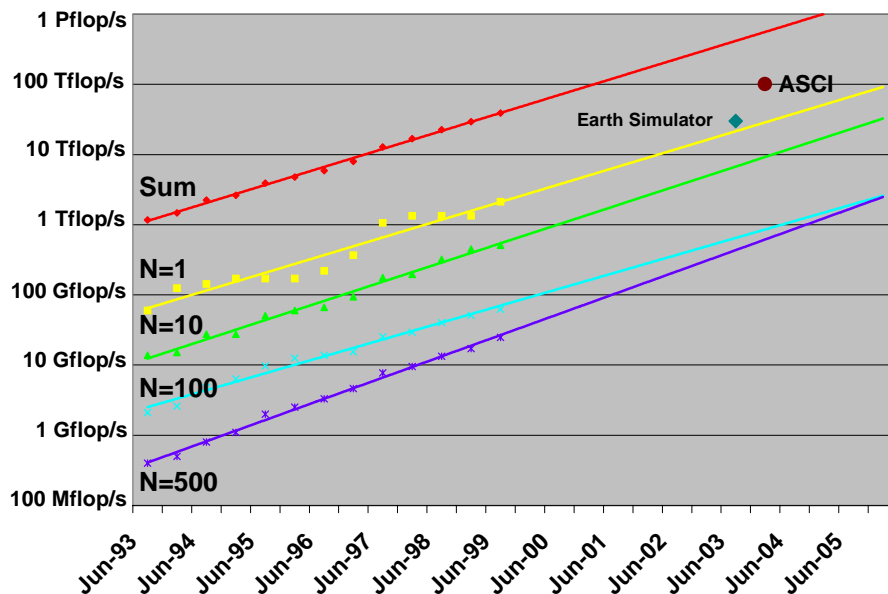


Fig. 13. Extrapolation of recent growth rates of performance seen in the TOP500.

Based on the current TOP500 data which cover the last 6 years and the assumption that the current performance development continue for some time to come we can now extrapolate the observed performance and compare these values with the goals of the mentioned government programs. In figure 13 we extrapolate the observed performance values using linear regression on the logarithmic scale. This means that we fit exponential growth to all levels of performance in the TOP500. These simple fitting of the data shows surprisingly consistent results. Based on the the extrapolation from these fits we can expect to have the first 100 TFlop/s system by 2005 which is about 1–2 years later than the ASCI path forward plans. By 2005 also no system smaller then 1 TFlop/s should be able to make the TOP500 any more.

The only performance level which deviates with its growth rate from all others is the performance of position 100. In our opinion this is actually a sign of a lower number of centers being able to afford a very high end system. This concentration of computing power in fewer centers can be seen in the US for a couple of years already in basically all government programs.

Looking even further in the future we could speculate that based on the current doubling of performance every year the first Petaflop system should be available around 2009. Due to the rapid changes in the technologies used in HPC systems there is however at this point in time no reasonable projection possible for the architecture of such a system at the end of the next decade. Even as the HPC market has changed its face quite substantially since the introduction of the Cray 1 three decades ago, there is no end in sight for these rapid cycles of re-definition. And we still can say that in the High-Performance Computing Market “The Only Thing Constant Is Change”.

References

- [1] G. Bell, The Next Ten Years of Supercomputing, *Proceedings 14th Supercomputer Conference*, Mannheim, June 10 -12, 1999, Editor: Hans Werner Meuer, CD-ROM (MaxionMedia), ISBN Nr. 3-932178-08-4
- [2] J. J. Dongarra, H. W. Meuer, and E. Strohmaier, TOP500, www.top500.org.
- [3] H. W. Meuer, The Mannheim Supercomputer Statistics 1986–1992, *TOP500 Report 1993*, (University of Mannheim, 1994), 1–15
- [4] R. W. Hockney, C. Jesshope, *Parallel Computers II: Architecture, Programming and Algorithms*, Adam Hilger, Ltd., Bristol, United Kingdom, 1988
- [5] see: www.netlib.org/benchmark/performance.ps
- [6] H. D. Simon, High Performance Computing in the U.S., *TOP500 Report 1993*, (University of Mannheim, 1994), 116–147
- [7] A. J. van der Steen and J. J. Dongarra, Overview of Recent Supercomputers, www.phys.uu.nl/~steen/overv99-web/overview99.html.
- [8] G. V. Wilson, Chronology of major developments in parallel computing and supercomputing, www.unipaderborn.de/fachbereich/AG/agmadh/WWW/GI/History/history.txt.gz
- [9] P. R. Woodward, Perspectives on Supercomputing, *Computer* **29** (October 1996), no. 10, 99–111