High Performance Computing in the U.S. – An Analysis on the Basis of the TOP500 List

Horst D. Simon[1]
Applied Research Branch
NASA Ames Research Center
Mail Stop T27A-1
Moffett Field, CA 94035-1000

May 26, 1994

Abstract

In 1993 for the first time a list of the top 500 supercomputer sites worldwide has been made available. The TOP500 list allows a much more detailed and well founded analysis of the state of high performance computing. Previously data such as the number and geographical distribution of supercomputer installations were difficult to obtain, and only a few analysts undertook the effort to track the press releases by dozens of vendors. With the TOP500 report now generally and easily available it is possible to present an analysis of the state of High Performance Computing (HPC) in the U.S. This note summarizes some of the most important observations about HPC in the U.S. as of late 1993 or early 1994, in particular the continued dominance of the world market in HPC by the U.S, the market penetration of massively parallel systems (MPP), and the growing industrial use of supercomputers.

[1]The author is an employee of Computer Sciences Corporation. This work is supported through NASA Contract NAS 2-12961. e-mail: simon@nas.nasa.gov.

1 Introduction

In 1994 high performance computing in the U.S. appears to be undergoing a major transition. With the introduction of powerful new "massively parallel processors[2]" (MPP) by established vendors in the high performance computing field such as the T3D by Cray Research, the SP-2 by IBM, and the Exemplar SPP by Convex, there are strong indications that MPP is now finally in the mainstream of scientific supercomputing. These new entries compete with existing MPP vendors such as Intel, Kendall Square, Ncube, and Thinking Machines for market share. This happens at a time, when demand for MPP systems is not as rapidly increasing as expected. Traditional parallel vector supercomputes such as the Cray C90 continue to dominate both performance benchmarks [1], and show strong results when the breadth and complexity of applications is concerned (see [8]). The strength of the market (more in terms of expectations, than in real dollars) can be seen from the list of vendors in Table 1, which is updated from Smaby [10]. The consequences of such a large number of vendors competing for such a small (but highly visible and important) market are widely discussed [9].

At the same time the federal High Performance Computing and Communications Program (HPCCP) is in high gear, and considerable progress has been made as documented in the famous "Blue Book" [3]. On the other hand there has been criticism of this very program, e.g., in the form of a report by the Congressional Budget Office [4]. The discussion about HPC in the commercial and in the government market place is mainly based on beliefs and impressions, and often lacks hard data. It is surprising that a field such as HPC that is deemed so critically important to the national agenda lacks almost completely any quantitative assessment of its progress.

This report attempts to shed some light on recent developments in HPC in the U.S. and present some quantitative data on the type and distribution of HPC technology. All the information here is based on the TOP500 list of November 1993. The report [6] ranks the 500 top performing supercomputers worldwide. The measure of performance is the maximal achieved $Rmax$ value for the computer on the LINPACK benchmark as reported in [5]. Using this

[2] I have argued elswhere that "massive" is a misnomer since none of these systems are truly massively parallel, i.e., have in excess of 1024 processors. However, currently MPP has become a commonly accepted term, denoting scalable, distributed memory, parallel computers, and I will use it here in this sense

Table 1: **Commerical HPC Vendors in the U.S (early 1994)**

Status	Vendors
Out of business	Alliant, American, Ametek, Culler, Cydrome, Denelcor, Elexsi, Multiflow, Myrias, Prisma, Saxpy, SCS, SSI(2), Trilogy, Wavetracer
Division closed	Astronautics, BBN, CDC/ETA Systems, E&S, Gould, Vitesse
Merged	Celerity, FPS, Key, Supertek
Down, not out	AMT(Cambridge), CHoPP, Encore, Stardent/Kubota
Currently active	Convex/HP, Cray Computer, Cray Research, Fujitsu, IBM, Intel, KSR, nCUBE, Meiko, NEC, Parsytec, SGI, Tera, TMC

measure the cutoff to make the list of the TOP500 systems worldwide is a performance of 472 Mflop/s on the LINPACK benchmark. Six Fujitsu VP-200E computers take on the ranks from 495 to 500. Interestingly the top ranked machine is also made by Fujitsu: the 140 processor specially built computer for the Numerical Windtunnel project at NAL in Japan is rated at 124.5 Gflop/s.

Before investigating some of the data in [6] in more detail, it is important to understand the limitations of the TOP500 study. These limitations can be summarized as follows:

- The list is a snapshot in time. The TOP500 list reflects the state of supercomputer installations as of November 1993. Since the market is extremely dynamic, such a study is almost out of date by the moment it is published. For example the November 1993 report does not list any of the Cray T3D installations; by March 94 there were about 15. It lists only six IBM SP-1 machines, and by March 94 there were more than 75 installed. However, since the data were collected consistently, we have at least an accurate reflection of the distribution at this time.

- The LINPACK *Rmax* value overestimates the performance of parallel machines. It is known that solving dense linear systems is a task that is well suited to highly parallel systems. MPPs therefore tend to exhibit a proportionally much higher performance than parallel vector processors (PVP). Even though the NAS Parallel Benchmarks [2], for example, would be a more accurate measure of parallel system performance, LINPACK figures are more widely available. They also provide an upper limit of the best possible performance of a computer on *any* application. If understood this way, the results in TOP500 take on a different, but at least consistent meaning.

- A large number of systems are in classified or highly proprietary areas. Both government and industrial users of HPC are, for different reasons, reluctant to provide details on the type and number of systems used. This may lead to double counts in some cases, and to undercounts in others, for example a large nCUBE2 system to be rumored to be installed at Shell is not on the list. On the other hand the TOP500 list contains many "government" listings where location and use is competely unknown, and where is it possible that the same machine is counted twice.

- The TOP500 does not reflect actual usage of machines. For example several of the Intel Paragon machines are listed with high $Rmax$ value, however, it is known that many of the larger Paragon systems in 1993 were not yet able to deliver their full compute power for application users (see for example [11] for machine #48). Other systems, such as the Cray 2 at Phillips Laboratory (#235) have been decommissioned and are no longer in active use. However, there is no reasonable way to consistently estimate the actual compute cycles delivered by any of the computers on the list.

In spite of these inherent limitations the TOP500 can provide extremely useful information, and valuable insights. It is more accurate than many marketing studies, and the possible sources of error discussed above are probably statistically insignificant, if we consider only summary statistics, and not individual data. All Mflop/s or Gflop/s performance figures here refer to performance in terms of Linpack $Rmax$.

In the analysis of geographical distribution, machines in Canada have been included in the figures for the U.S., and the figures for Europe include all European countries, not just EC members. The other country category includes mostly countries of the Pacific Rim with the exclusion of Japan, and a few Latin American Countries.

2 U.S. Dominance of the World Wide HPC Market

The TOP500 list clearly demonstrates the dominant position the U.S. assumes in the world both as producer and as consumer of high performance computers. In Table 2 the total number of installed systems in the major world regions is given with respect to the origin of the computers.

If one considers in Table 2 the country of origin then it is striking that 389 out of the TOP500 systems are produced in the U.S., which amounts to 78% of all installed systems. Japan accounts for 18% of the systems, and Europe produces only 4%. The extent of the American dominance of the market is quite surprising. For years, in particular in the mid 80's, there were ominous and ubiquitous warnings that the American supercomputer industry (which was essentially Cray Research at that time) is highly vulnerable to an "attack" by the Japanese vertically integrated computer giants Fujitsu, NEC, and Hitachi. Obviously this has not happened. How much various efforts such as the NSF Supercomputing Initiative in the mid 80's, or more recently the HPCC Program have contributed to the current vast superiority of the U.S. high performance computing industry, remains to be investigated. It is interesting to note that one view expressed outside the U.S. [12] is that strengthening the U.S. HPC industry and easing the transition to MPP was the *only* rationale for the HPCC Program.

The numbers for Europe are actually worse than Table 2 indicates, since 14 of the 21 "European" machines are actually Fujitsu VP products, which are resold in Europe by Siemens/Nixdorf. Hence the true European production is a total of only 7 machines (4 Parsytec and 3 Meiko), all of which are installed in Europe. In spite of a recent installation of a European system at a U.S. government laboratory (a Meiko CS-2 at Lawrence Livermore Natl. Lab.) the situation in Europe remains bleak. With lack of immediate access to the newest hardware, and the absence of the close interaction of users with vendors as is prevalent in the U.S., the best the European High Perfor-

Table 2: **US Share of Total Number of Installed TOP500 Systems**

Systems Manufactured In	Systems Installed In U.S.	Japan	Europe	Other	Total
U.S.	236	31	106	16	389
Japan	5	75	6	4	90
Europe	0	0	21	0	21
Total	241	106	133	20	500

Table 3: **US Share of Total Rmax (in Gflop/s) of Installed TOP500 Systems.**

Systems Manufactured In	Systems Installed In U.S.	Japan	Europe	Other	Total
U.S.	853	46	143	33	1075
Japan	30	307	21	5	363
Europe	0	0	30	0	30
Total	883	353	194	38	1468

mance Computing and Networking Initiative can accomplish is maintaining the status quo of Europe as a distant third in high performance computing technologies.

Table 3 is analogous to Table 2, but instead of the number of systems, the aggregate performance in *Rmax*-Gflop/s is listed. Table 4 lists the ratio of the corresponding entries in Tables 2 and 3, i.e., the average *Rmax*-Gflop/s per machine. From Table 4 we can see that on average the machines manufactured in Japan have higher performance ratings than machine manufactured in the U.S. Also machines installed both in Japan and the U.S. appear to be more powerful than the machines installed in Europe or in other countries.

A more interesting analysis of Tables 2 and 3 addresses the ongoing question of who has the higher trade barriers with respect to high performance computing, the U.S. or Japan? Table 2 shows that obviously both countries favor their own machines over their competitor's. But let us assume for the

Table 4: **Average Gflop/s per machine.**

Systems Manufactured In	Systems Installed In U.S.	Japan	Europe	Other	Total
U.S.	3.6	1.5	1.3	2.1	2.8
Japan	6.0	4.1	5.3	1.3	4.0
Europe	-	-	1.4	-	1.4
Total	3.7	3.3	1.5	1.9	2.9

sake or argument that the European market is equally open and accessible to both American and Japanese machines, i.e., the distribution of supercomputers in Europe reflects an open market. According to the distribution in Europe the split in an open market should be 83% American and 17% Japanese. Assuming these market shares worldwide, the number of Japanese machines in the U.S. would be 41 (as opposed to 5), and the number of U.S. made supercomputers installed in Japan would be 88 (as opposed to 31). Clearly both countries seem to have mechanisms in place which restrict competition in favor of the local products. The reader can decide whether only 5 out of 41 or only 31 out of 88 constitutes a more biased situation.

3 Market Penetration of MPP

The penetration of the supercomputer market by MPPs is another often debated issue. The TOP500 list again delivers hard data, which come somewhat as a surprise. In Table 5 the share in % in number of installations for MPP machines among the TOP500 is given. All other machines will be counted as parallel vector processors (PVP). The PVP category includes all machines manufactured by Cray Research (with the exception of the Cray Superserver and T3D machines), all machines manufactured by IBM (with the exception of SP-1 machines), as well as all machines from Convex, Fujitsu, Hitachi, NEC, and Siemens-Nixdorf (Fujitsu). All others are counted as MPPs. The MPP category thus includes a number of machines which should rather be labeled symmetric multiprocessors (SMPs), but their inclusion does not change the overall picture significantly.

Overall MPPs account for 37% of all installed supercomputers. The installed base of MPPs in the U.S. is significantly higher. In the U.S. MPPs are at 51% already in the majority, whereas in Japan MPPs are at 13% considerably underrepresented. Table 6 shows that irrespective of the country of installation, MPPs are generally more powerful than PVPs in terms of $Rmax$-Gflop/s.

The considerably higher number of MPPs in the U.S. compared to the world average is clearly a direct consequence of the U.S. HPCC program. How many *additional* machines in the U.S. have been installed because of the HPCC support, and other subsidies for parallel machines, e.g. (D)ARPA funding? Again this is difficult to estimate, but as before we can assume that Europe is an objective test market. This is probably a fair assumption, since at the time of this study the European HPCN Initiative probably had no effect on the selection of MPP machines versus PVPs. Under this assumption then the share of MPP machines in the U.S. would be only 32%, which is about 77 machines. Following this analysis one therefore could credit the installation of about 45 machines in the U.S. to the strong government support for massive parallelism. With the data in Table 3 one can conclude that about 139 $Rmax$-Gflop/s of the installed base of machines in the U.S. are due to HPCC funding. At about 0.5$MperGflop/sthiswouldamountto$70M of direct government support for advanced architecture hardware in the U.S. from 1990 to 1993. This figure appears to be in the right order of magnitude. For example the General Accounting Office report on ARPA [7] lists a total of $55M spent in connection with procurements of Intel and TMC supercomputers alone. The list of ARPA supported machines ends with installations in late 1992, and does not include any of the more recent large scale Paragon or CM-5 machines, which have made the TOP500 list, and which are also part of the HPCC program. Thus as a result of direct government support, the U.S. is clear world leader in the adaptation and use of MPP technology.

4 HPC usage in Government, Universities, and Industry

The discussion of the high level of government support for MPPs in the U.S. in section 3 should imply that there is a large installed base of supercomputers

Table 5: **Market Share of MPPs (in number of installations).**

Region	PVP	MPP	% MPP
U.S.	119	122	51
Japan	92	14	13
Europe	90	43	32
other	12	8	40
Total	313	187	37

Table 6: **Market Share of MPPs (in computational power in Rmax-Gflop/s).**

Region	PVP	MPP	% MPP
U.S.	355	528	60
Japan	198	155	44
Europe	128	66	34
other	33	5	13
Total	714	754	51

Table 7: **Distribution of TOP500 by Sector (Number of Installations).**

	Gov	Uni	Ind	Ven	Total
U.S.	86	71	62	22	241
Japan	23	22	43	18	106
Europe	52	53	27	1	133
other	6	11	3	0	20
Total	167	157	135	41	500

Table 8: **Distribution of TOP100 by Sector (Number of Installations).**

	Gov	Uni	Ind	Ven	Total
U.S.	28	16	10	5	59
Japan	11	5	6	3	25
Europe	7	5	1	0	13
other	2	1	0	0	3
Total	48	27	17	8	100

in the government sector in the U.S. This is indeed the case as Tables 7 - 9 show.

In Tables 7 - 9 first the distribution of the TOP500 and then the distribution of the TOP100 machines across different sectors is shown with respect to number of installations and performance in $Rmax$-Gflop/s. Here government refers to all government agencies and laboratories, which ranges from basic research environments to more production oriented machines as in the national weather bureaus. The categories university and industry are self-explanatory. The column with vendor machines includes all the machines used by supercomputer vendors in house for benchmarking and software development. The distribution of both the TOP500 and of the TOP100 is given, because there are considerable differences.

¿From the data in Table 8 it is clear that the distribution of the largest

Table 9: **Distribution of TOP100 by Sector (Performance in Rmax-Gflop/s).**

	Gov	Uni	Ind	Ven	Total
U.S.	385	160	55	42	642
Japan	176	25	26	36	263
Europe	43	21	3	0	67
other	17	4	0	0	84
Total	621	210	84	78	993

supercomputers in the U.S. is fairly consistent with the distribution in Japan and Europe. In all three regions/countries government accounts for about half of the largest supercomputers and industry for around 20 %.

Comparing the distribution of the TOP100 with the TOP500 is quite intriguing. First it is notable that industrial installations with about have world wide a much larger share among the TOP500 systems (27% compared to 17%). This is to be expected since the really large machines will be predominant in research environments. However, what is surprising is that the distribution between industrial and government installations is quite different in the corresponding countries/regions. In Japan only about 21% of the supercomputers are installed in government laboratories, whereas 40% Europe is exactly reverse with 39% government and 20% industrial use. The U.S. takes a middle position, somewhat closer to the situation in Europe, with 33% government use and 25% industrial use.

These results seem to imply that industrial use of supercomputers in Japan is considerably more widespread than in the U.S. While this is probably true in general, the situation in the U.S. is biased because of the effects of HPCC. It may be that U.S. industry is not using significantly less supercomputers than Japan, but rather that U.S. government laboratories due to government spending are using considerably more supercomputer power than their Japanese counterparts. It is the goal of the HPCC program that this heavy use in government laboratories will eventually have a positive effect on the industrial use of high performance computing technologies. If this type of technology transfer will occur remains to be seen. Future version of the TOP500 report will provide the necessary data.

Table 10: **Distribution of Industrial Supercomputers by Sector.**

Industry	PVP		MPP		Total
	U.S.	Other	U.S.	Other	
Automotive	4	25	0	0	29
Geophysics	15	5	6	3	29
Aerospace	6	3	12	1	22
Electronics	2	10	3	4	19
Metal/Constr.	0	10	0	0	10
Chem./Bio.	5	2	1	0	8
Econ./Fin.	0	2	3	0	5
Energy	1	2	1	0	4
Other	0	2	3	4	9
Total	33	61	29	12	135

5 Commercial Use of Supercomputers.

One of the strategic goals of the HPCC program was to "spur gains in U.S. productivity and industrial competitiveness by making high performance computing and networking technologies an integral part of the design and production processes". The TOP500 data can give some indication what progress has been made towards this goal. In Table 10 the distribution of the 135 industrial supercomputers worldwide is given by industry segment, as well as by MPP versus PVP, and by use in the U.S. versus all other countries.

The most surprising fact is that in the U.S. the number of MPPs in industry is almost the same as the number of PVPs. Thus even in the most competitive environment MPPs have already made major inroads. These data contradict some of the conclusions made in a controversial study [4], where it is maintained that "conventional supercomputers ... (are) hard to displace in the market". The TOP500 data indicate that displacement of PVP has already occurred to a much larger extent than generally believed in areas that have been traditionally claimed by PVPs. The two industry segments where MPPs have made most progress in displacing PVPs are aerospace and geophysics. The success of MPPs in these fields can be easily explained by the nature of some of the typical applications. Seismic processing in the oil industry is an application where MPPs excel, since most of the computation is of explicit nature and done on structured grids or often is I/O bound. Similarly, electromagnetics applications using methods of moments algorithms, and image and signal processing applications account for the use of MPPs in the aerospace community. The parallel applications taxonomy in [9] explains this success of MPPs.

Similarly there are areas where MPPs had no success whatsoever. Not surprisingly these are for example in the automotive industry. This industry is characterized by using unstructured finite element codes with implicit solution algorithms. These types of applications are very difficult to solve efficiently on MPPs. Another characteristic of the automotive industry is reliance on third party applications such as crash codes and structural analysis packages. So far third party developers, mostly small software companies, have been reluctant to port their software to unproven new technology. In contrast, most oil companies employ proprietary in-house codes, and have been willing to make the investment in porting applications. Thus the data in Table 10 reflect exactly the difficulty in porting the applications in different industry segments.

Another observation is that MPPs so far have been replacing PVPs in established field of high performance computing use. There have been no significant inroads into new applications markets yet. Even though there are several MPPs now in financial institutions, or in chemical and biological applications, their numbers are still small. Hence MPPs have not (yet) opened up new markets for HPC technology.

There are a few other observations one can make from Table 10 concerning HPC use in industry comparison. For example all 10 supercomputers in the metal working and construction industry are installed in Japan, whereas in aerospace with 18 out of 22 installations the U.S. is the clear leader. The supercomputers in the automotive industry are almost evenly distributed worldwide. Every major car manufacturer owns a supercomputer (with the exception of Volvo). One can therefore assume that the use of supercomputers in industrial applications is a *consequence* of the leadership position each country enjoys in the respective applications field (and not vice versa). However, the topic of supercomputer use and industrial leadership deserves some further detailed investigation.

6 Conclusions

The analysis of the data provided by the TOP500 report has led us to a number of conclusions concerning the state of HPC in the U.S. Some of these conclusions are:

- The U.S. is the clear world leader both as producer and as consumer of high performance computers.
- Both the U.S. and Japan seem to have mechanisms in place which restrict competition in favor of the local products.
- MPPs account for more than a third of all installed supercomputers worldwide. Market penetration by MPPs in the U.S. is significantly higher, where MPPs constitute already now a majority of the installed supercomputers. A sizable portion of this market share can be attributed to government spending and the HPCC program.
- Government installations dominate the market for large systems both in the U.S. and worldwide. Overall, the U.S. lags somewhat in industrial use of supercomputers.
- MPPs have made major progress in replacing PVPs in several industries. In the U.S. MPP usage in aerospace and in the oil industry is very high. MPPs appear to be considerably more widespread in industrial environments than commonly thought.

Generally the TOP500 list has proven itself to be an extremely valuable tool for evaluating trends in the HPC market. Future releases of this report should enable the HPC community to track important developments much more accurately than in the past.

References

[1] D. H. Bailey, E. Barszcz, L. Dagum, and H. D. Simon. NAS parallel benchmark results 03-94. Technical Report RNR-94-006, NASA Ames Research Center, Moffett Field, CA 94035, March 1994.

[2] D. H. Bailey, J. Barton, T. Lasinski, and H. Simon (editors). The NAS parallel benchmarks. Technical Report 103863, NASA Ames Research Center, Moffett Field, CA 94035, July 1993.

[3] Committee on Physical, Mathematical, and Engineering Sciences. *High Performance Computing and Communications: Toward a National Information Infrastructure.* Office of Science And Technology Policy, Washington, D.C., 1994. (available via Mosaic from the National Coordination Office for HPCC, URL `http://www.hpcc.gov/`).

[4] Congress of the United States, CBO. Promoting high performance computing and communications. Technical Report ISBN 0-16-041828-3, U.S. Government Printing Office, Mail Stop SSOP, Washington D.C. 20402-9328, June 1993.

[5] J. J. Dongarra. Performance of various computers using standard linear equations software. Technical Report CS-89-85, Department of Computer Science, Univ. of Tennessee, November 1993.

[6] J. J. Dongarra, H. W. Meuer, and E. Strohmeier. TOP 500 supercomputer sites. Technical Report RUM 33/93, Computing Centre, Univ. of Mannheim, 68131 Mannheim, Germany, November 1993. (The TOP500 report is available via anonymous ftp by `ftp ftp.uni-mannheim.de`, then `cd info/rumdoc`, and `get rum3593.ps`).

[7] General Accounting Office. High performance computing: Advanced Research Projects Agency should do more to foster program goals. Technical Report GAO-IMTEC-93-24, United States GAO, Washington, D.C. 20548, May 1993.

[8] Alan H. Karp, Don Heller, and Horst Simon. 1993 Gordon Bell Prize winners. *IEEE Computer*, 27(1):69–75, January 1994.

[9] H. D. Simon. High performance computing: Architecture, software, algorithms. Technical Report RNR-93-018, NASA Ames Research Center, Moffett Field, CA 94035, December 1993.

[10] G. P. Smaby. Crystal balls and rearview mirrors: A wall street view on supercomputing. Parallel Computing in the 90s Seminar presented at NASA Ames Research Center, January 1992.

[11] B. Traverstat and D. McNab. Acceptance test for the Intel Paragon XP/S15. Technical Report RND-94-004, NASA Ames Research Center, Moffett Field, CA 94035, February 1994.

[12] Hans-Martin Wacker. The European HPC initiative. In H.-W. Meuer, editor, *Supercomputer '92*, pages 239 – 239, Berlin, Heidelberg, New York, 1992. Springer Verlag.