EXECUTIVE SECRETARIAT
ROUTING SLIP

TO:			ACTION	INFO	DATE	INITIAL
	1	DCI				
	2	DDCI		X		
	3	EXDIR				
	4	D/ICS				
	5	DDI		X		
	6	DDA				
	7	DDO				
	8	DDS&T				
	9	Chm/NIC				
	10	GC				
	11	IG				
	12	Compt				
	13	D/OCA				
	14	D/PAO				
	15	D/PERS				
	16	D/Ex Staff				
	17	C/TTAC/OSWR		X		
	18	D/OGI/DI		X		
	19	NIO/ECON		X		
	20					
	21					
	22					

ER

SUSPENSE _____
Date

Remarks

STAT

ER 88-3570X

Executive Secretary

14 SEP 88
Date

3637 (10-81)

Page Denied

Next 1 Page(s) In Document Denied

IEEE/USAB COMMITTEE ON COMMUNICATIONS AND INFORMATION POLICY
UNITED STATES ACTIVITIES BOARD

John M. Richardson
Chairman
(202) 334-2844

Claud M. Davis
Vice Chairman
(914) 742-5929

Richard VanSlyke
Vice Chairman
(718) 260-3050

Heidi F. James
Executive Secretary
(202) 785-0017

PLEASE REPLY TO:
1111 19th Street, NW
Suite 608
Washington, DC 20036-3690 USA

U.S. SUPERCOMPUTER VULNERABILITY

A REPORT PREPARED BY:

SCIENTIFIC SUPERCOMPUTER SUBCOMMITTEE
COMMITTEE ON COMMUNICATIONS AND INFORMATION POLICY
UNITED STATES ACTIVITIES BOARD
INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

AUGUST 8, 1988

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

U.S. SUPERCOMPUTER VULNERABILITY
BY THE
IEEE SCIENTIFIC SUPERCOMPUTER SUBCOMMITTEE [1]
COMMITTEE ON COMMUNICATIONS AND INFORMATION POLICY

INTRODUCTION

The U.S. created the supercomputer industry. Today U.S. firms jointly are leaders in world supercomputer markets. Nonetheless, U.S. supercomputer firms are vulnerable to a focused strategy by the Japanese targeted on their industry. As documented below, the reasons for the vulnerability of the U.S. firms are complex. To overcome their vulnerability will require a systems solution: an integrated, cooperative effort among industry, universities and government. It appears that such a solution will require coordination by government to a degree seldom if ever achieved in this country in peace time. But the threat is real, and it is not limited to supercomputers. Supercomputers appear to represent only a next step in an on-going process that is gaining momentum. As one can see from the "Visions" for the future promulgated by the Ministry of International Trade and Industry (MITI), Japan seeks a dominant position in the information industries--the computer and communications industries of the world. And it is well on the way to achieving that position.

Many doubt that the U.S. government could play an effective, leadership, policy role in relation to civilian technology/industry problems of this kind. Yet the Japanese government clearly does play such a role in its economy. And Japan is a market-oriented, capitalistic, democracy. By using its approach Japan is overtaking, and often surpassing, the U.S. in field after field. How long can we afford to wait before we respond?

Perhaps the most important question for the U.S. to ponder is: What is the alternative to effective government leadership and support in vital technology/industry matters such as the ones discussed here? Current approaches are not working. The discussion that follows is focused on the supercomputer issue as compounded by the semiconductor issue. Yet, these two issues appear to be only special cases of the broader, generic problem.

---

[1] The IEEE Scientific Supercomputer Subcommittee is a technology-policy subcommittee of the Committee on Communications and Information Policy under the United States Activities Board of the Institute of Electrical and Electronics Engineers, 1111 19th Street, NW, Washington, DC 20036. The members of the Subcommittee are:

Sidney Fernbach, Chairman	Duncan Lawrie	Stewart Saphier
Lara Baker	Alan McAdams	Paul B. Schneck
Vito Bongiorno	Kenneth W. Neves	Lloyd Thorndyke
Alfred E. Brenner	John Ranelletti	Kenneth F. Tiede
James F. Decker	John P. Riganati	Hugh Walsh

John M. Richardson, Chairman, Committee
  on Communications and Information Policy

ANALYSIS

Cray and Control Data Corporation (CDC) with its ETA subsidiary are the two long-term U.S. manufacturers of supercomputers. Cray has the largest world market share. It is a stand-alone company, solely dependent on the supercomputer business. CDC and ETA are emerging from financial difficulties and, while more diversified, are not yet nearly as strong as Cray in the supercomputer field. Both Cray and ETA are in a vulnerable position.

Several important factors contribute to the vulnerability of the U.S. supercomputer firms: (1) the current situation in the semiconductor industry and the recent history of US-Japan trade problems in the industry; (2) the rapid strides the Japanese supercomputer firms have made; (3) the fact that Cray and ETA rely on high performance memory devices that now are available only from Japan; (4) the implications of the recent supercomputer trade agreement with Japan; and (5) the fact that U.S. firms are losing a primary market advantage, their differential access to the huge reservoir of applications software that currently is optimized primarily for their installed base of supercomputers.

These factors are discussed in relation to their specific impact on U.S. supercomputer manufacturers; but the process is generic Japanese. As stated in SCIENCE in November, 1985:[2]

> "...In this process, the Japanese begin by picking a few key, high-volume components over which to compete. Then, as the U.S. producers retreat under extreme price-cutting assaults, the Japanese companies extend the fight step-by-step to other items until they have swept competitors out of the most profitable areas. (Emphasis added.)
>
> This has happened in other industries, and in semiconductors the Japanese have all but taken over the field of memory chips..."

Conditions are ripe for supercomputers to represent the next step in the process.

RECENT SEMICONDUCTOR HISTORY

The U.S. created the semiconductor industry. The U.S. invented the VLSI (very large scale integration) Chip. For decades U.S. firms jointly were the leaders in world semiconductor markets. Nonetheless, U.S. semiconductor firms were vulnerable to a focused strategy by the Japanese on DRAMs, dynamic random access memories.

The Japanese developed and manufactured excellent, highly reliable, high quality DRAMs. Then, (as adjudicated by the U.S. Foreign Trade Commission) for more than two years the Japanese dumped these products (sold them below cost) in the U.S. and third markets. During the time that the Japanese were flooding the market with DRAMs, U.S. firms lost in excess of two billion dollars. U.S. merchant manufacturers of semiconductors are stand-alone companies. Effectively they were forced to withdraw from domestic production of high-performance DRAMs.

---

[2] Eliot Marshall, "Fallout from the Trade War in Chips," SCIENCE, November 22, 1985, p. 918.

In other product lines, SRAMs, static random access memories; and EPROMs, erasable programmable read only memory devices; U.S. firms also experienced duress. The U.S. firms responded with legal (as opposed to technological) actions that led to the U.S. Government's imposing tariffs on the Japanese as sanctions. This government effort, while well intended, raised the prices of chips and thus resulted in increased costs to U.S. supercomputer manufacturers and thus has tended to reduce their competitiveness.

More recently, the problem of dumping by the Japanese has been replaced by the current problem of severe shortages of DRAMs in the U.S. These shortages have been shown by experts from the Brookings Institution to result from reduced actual levels of Japanese output that "happen" to accord almost perfectly with the "forecasts" of production levels that had been made by MITI several months before. DRAM prices in the U.S. have sky rocketed--sometimes tenfold, and more. The DRAM business is now enormously profitable to the Japanese producers of semiconductors. This is a classic pattern for those who achieve effective dominance of a market.

During the period of DRAM shortages, exchange rates have been highly favorable to U.S. domestic production. Furthermore, sales of DRAMs in the U.S. market have been highly profitable to suppliers. Nonetheless, (at the time of this writing) U.S. firms were not re-entering the market, either by investing in new plants or even by reopening "mothballed" DRAM plants. Perhaps they recall their recent huge losses too vividly, or perhaps the rapid "decay" of know-how and expertise in so demanding a business explains their lack of response. In any case, a market that is difficult to re-enter is the type of market in which the classic pattern is most successful.

The problems of U.S. firms are rooted in the new-found power of Japanese firms in important aspects of the semiconductor business. In presentations to the authoring IEEE groups, expert observers have stated variously that Japan has "won the silicon war," (August, 1985); and that, "We (the U.S.) failed. We had it (semiconductor leadership) and we lost it!" (May, 1988). Despite these pessimistic assessments and the current dire situation, we believe that this need not (yet) be the case. But the hour is late, and strong action will be required to prevent it.

JAPANESE SUPERCOMPUTER COMPETITION

Cray and CDC/ETA face the three Japanese manufacturers of supercomputers shown on Table I. Not coincidentally, these three firms are also among the leading Japanese manufacturers of semiconductors. Cray and ETA are tiny in comparison to their Japanese rivals--they are small even in relation to the leading U.S. merchant semiconductor manufacturers. (These firms are also shown on Table I.) Given the experience of the semiconductor industry, plus the facts cited in the next paragraphs, it is not hard to understand why the U.S. supercomputer firms are vulnerable to the "next step" by the Japanese.

The sum of the current annual supercomputer outputs of the three Japanese giants approximates the number, though not yet the dollar value, of Cray's output, and greatly exceeds that of CDC/ETA. As Table I demonstrates, the Japanese firms are not stand-alone manufacturers of supercomputers. Their operations are integrated over multiple production stages and across several product lines.

Table II shows the number of supercomputers of each manufacturer installed and on order as of late 1987. It is important to note that both the CDC 205 and the Cray 1 are now obsolete, while the ETA 10's and Cray Y-MPs (not listed) are just becoming operational. Table III shows the peak megaflops per processor for each manufacturer as well as the megaflops per total (multiprocessor) system.

-4-

Table I

Market Segments Participated in By Selected U.S. And Japanese Manufacturers of Supercomputers and/or Semiconductors

	CRAY	CDC/ETA	TI INTEL Motorola AMD ETC.	AT&T	IBM	NEC	FUJITSU	HITACHI
Super-computers	x	x			RE*	x	x	x
Main Frame Computers		x			x	x	x	x
Intermediate Computers					x	x	x	x
Mini Computers		x		x	x	x	x	x
Micro Computers					x	x	x	
Consumer Electronics						x	x	x
Semiconductors								
– Merchant			x			x	x	x
– Captive				x	x	x	x	x

* Re-entering.

-5-

Table II
Supercomputer Installations

System	No. Installed	No. on Order	FCS
CDC/ETA			
CDC Cyber 205	37	0	1980
ETA 10 series	2	8	1987
Cray			
Cray-2	8	4	1985
X-MP series	117	43	1984
Cray-1	58	0	1976
Fujitsu			
VP series	56	5	1984
Hitachi			
S-810 & 820	19	4	1983
NEC			
SX2 series	12	7	1985

Source: Adapted from The Gartner Group, November, 1987.

Table III
Peak Performance Rates

COMPUTER	SINGLE CPU PEAK 64-BIT MFLOP RATE	ALLOWING MULTIPLE CPUs	
CRAY-1	160	160	
CRAY X-MP	233	932	(4-CPUs)
CRAY-2	488	1,952	(4-CPUs)
CRAY-3 (1989)	1,000	16,000	(16-CPUs) est.
CYBER 205	200 (2-PIPE)	400	(4-PIPE)
ETA-10 (1986)	350	1,400	(4-CPUs)
ETA-10/E	415	1,660	(4-CPUs)
ETA-10/G (1988)	643	5,142	(8-CPUs)
FUJ. VP 100	271	271	
FUJ. VP 200	533	533	
FUJ. VP 400	1,067	1,067	
HIT. S-810/20	630	630	
HIT. S-820/80	2,000	N/A	
IBM 3090/VF	116	696	(6-CPUs)
NEC SX-1	570	570	
NEC SX-2	1,300	1,300	
NEC SX-3 (1989)	5,000	20,000	(4-CPUs)

Source: Adapted from "Computational Fluid Dynamics: Algorithms and Supercom-puters," by W. Gentzsch and K. Neves, NATO AGARDograph No. 311, March, 1988.

The facts in these tables are consistent with the following summary. The Japanese manufacturers, especially Fujitsu, have made rapid progress in installing systems during the four short years they have been producing super-computers (Table II). Their systems are to date single processor systems (Table III). This latter point relates to two elements: (A) the individual processors in the Japanese supercomputers are more powerful today than any U.S. manufacturer currently expects to deliver in less than half a decade; but, (B) the Japanese manufacturers today lag somewhat in adapting to parallel processing techniques (though they appear to lead in the complementary field of vectorization and vector processing; see below).

These facts imply that multiprocessor versions of the Japanese supercomputers, especially the newest Hitachi and NEC (Nippon Electric Corporation) machines, could far outstrip the best of the U.S. machines once the Japanese catch up in parallel processing, an area where they are hard at work. NEC has announced parallelism for its SX-2 series and is said to have completed the circuit design for its 20 gigaflop, four processor SX-3 series currently in development.

In a recent special issue (March 3, 1988) of the trade journal Electronics entitled "Inside Technology,"[3] these facts are reiterated in very strong language:

> "...The Japanese stick to simpler single-processor architectures. And their supercomputers are the speediest one-processor machines in the world--by far--thanks primarily to advanced semiconductor technology. The only way U.S. supercomputers can come near their processing rates is in multiprocessor configurations."

The following quotes also are in the article in a box entitled, "Where Japan Shines: Vectorizing Compilers."[4]

> "...Japanese...compilers make it relatively easy for a user with his own Fortran program to wring near maximum performance from a supercomputer's vector processor without resorting to special vector-processor programming methods.
>
> Automatic vectorization has been a Japanese strong suit for some time, and it is made easier by the Japanese's use of single-processor architectures in supercomputers. Compiler writers need not face the difficulties of automatic parallelization."

The latter quote points up an additional area of Japanese strength; the systems software of the Japanese supercomputers is already world class. NEC's Fortran compiler is reputed to be the best in the world. If that is in fact the case, it would have surpassed the previous best--that from Hitachi.

---

[3] Charles L. Cohen, "Japan Focuses on Simple but Fast Single-Processor Supercomputers," Electronics, March 3, 1988, p. 57.

[4] Ibid. p. 58.

U.S. SUPERCOMPUTER FIRMS RELIANCE ON JAPANESE SEMICONDUCTORS

Already the aggressive pricing strategies of the Japanese in supercomputers (discounts of up to 80% to universities and others) combined with inversely aggressive semiconductor strategies (refusal to export their latest high performance component devices), threaten the existence of Cray and ETA. The highest performance memory and bipolar logic components useful for supercomputers are no longer manufactured in the U.S.; they are available only from Japan. The managements of Cray and ETA have been quoted in the press at various times as stating that these Japanese components are, "not yet available for export" from Japan to Cray or ETA as devices--but they are available to end users in the Japanese supercomputer systems. Those systems are definitely available for export.

This continues the familiar, oft-repeated pattern. Japanese firms plan and act in accord with long-range strategic goals. When they achieve a technological advantage in one area, they use that advantage to ensure their advance into new areas. They have targeted supercomputers as the next high tech area in which to establish a dominant position, and they have made enormous strides toward that goal as Tables II and III demonstrate.

THE IMPLICATIONS OF THE SUPERCOMPUTER TRADE AGREEMENT WITH JAPAN

The supercomputer trade agreement with Japan contains the seeds which can further undermine the position of U.S. manufacturers. It requires of Japan a number of actions which, if made reciprocal to the U.S., could facilitate the entry of Japanese supercomputers into the U.S. market.

The reasoning that supports this conclusion proceeds through the following steps. Almost by definition, no one knows how to use a new generation of supercomputers efficiently. New systems generally embody the latest technology and architectural design in order to achieve their state-of-the-art computing speeds. In this country, the national laboratories historically have been "partners" in supercomputer development. The laboratories have developed the applications and forced the evolution of the operating systems that together permit the machines to work effectively. In other words, the laboratories' expert users have assisted greatly in facilitating improved productivity of the systems as these users learned more about the characteristics of the new systems and how to make optimal use of them. Now with the new NSF programs that have made supercomputers available to universities, universities, too, are becoming partners in these efforts. Often the U.S. manufacturer has recognized the "partnership contribution" of these U.S. institutions through price concessions, "buybacks" of time or other means. The same has been true in Japan.

The trade agreement requires the Japanese in procurements of supercomputers by their government or their universities to give equal preference to U.S. manufactured supercomputers. In essence, the Japanese have agreed to avoid "unfair" pricing on the part of the Japanese manufacturers. But what may seem "unfair" to a trade administrator looks to a manufacturer like "recognition of partnership" with a government agency or university.

While negotiating this agreement with one hand, the U.S. government has blocked the sale of a NEC supercomputer to M.I.T. with the other. To date Japanese supercomputers have been virtually shut out of U.S. government and university markets. It appears that government agencies, especially the Department of Defense, intend to keep it that way.

How long can such a pattern last? Is it realistic to believe that the Japanese can long be required to give unbiased consideration to U.S. supercomputers while Japanese machines are foreclosed from U.S. institutions? If the requirements we now impose on the Japanese were made reciprocal for U.S. universities and government laboratories, Japanese supercomputers would have to be acceptable to those agencies on non-discriminatory terms. The Japanese systems are already a match for the U.S. systems and soon could surpass them in performance. Price thus is likely to be a major determinant of choice; the Japanese manufacturers have been more than willing to sacrifice short-term profit for long-term market penetration.

The new requirements could greatly facilitate the entry of the Japanese into the U.S. At the same time such requirements could disrupt the implicit partnership between U.S. manufacturers and U.S. government laboratories and/or U.S. universities, and transfer the benefits of partnership to Japanese firms. In response to a low bid, U.S. National Laboratories could be required to become partners to Japanese firms in perfecting their systems for penetration of U.S. markets.

INCREASING "PORTABILITY" OF SOFTWARE

Paradoxically, desirable developments long sought by users on another front are hastening the progress of the Japanese. It is becoming increasingly easy to move applications from the machine of one supercomputer vendor to that of another. There are four elements which are facilitating software "portability:"

(1) Standardized versions of FORTRAN;
(2) Improved FORTRAN optimizing compilers;
(3) Broad-spectrum libraries of algorithms optimized for each vendor's systems;
(4) A common operating system--UNIX--now being implemented on supercomputers.

While in the past each supercomputer vendor had a "captive audience" of users, now it is increasingly possible to move applications between and among vendors. As the industry leader with the largest installed base, Cray had benefited from the prior situation in the industry; CDC did also, but to a lesser extent. But portable software does provide substantial benefit to users of supercomputers, it greatly diminishes this advantage previously enjoyed by the U.S. firms.

New entry into the field of supercomputing also is facilitated by increased applications portability. As the most recent entrants, the Japanese vendors are the greatest beneficiaries of these changes.

Cray itself also benefits from portability since portability helps Cray provide compatibility to its own multiple product lines. Use of UNIX permits Cray to focus on a single operating system thus reducing its development costs. The most significant effects, however, are those on the new entrants.

A further benefit to users is also achieved. Through UNIX, an operating environment uniform from work station to minicomputer, to minisupercomputer, to supercomputer would become possible. The Los Alamos National Laboratory is one of the organizations working hard to bring this about. One of its major objectives is to bring the myriad, powerful software innovations pioneered in the highly competitive fields of commercial computing over into the more rarified arena of supercomputing--with benefit to U.S. manufacturers as well as to users. The other side of the coin, of course, is that these benefits accrue to all vendors of supercomputers--especially to the Japanese.

THE IMPLICATIONS OF THE PHENOMENA OF SEMICONDUCTOR DISADVANTAGE AND THE PRESENCE OF PORTABILITY

At this point two threads come together. Once general portability of applications is achieved, the avenue to continued market leadership by U.S. firms would be solely through technological leadership. But how can U.S. firms simultaneously rely on componentry manufactured by their competitors, the Japanese, and assure their customers that they, the U.S. firms, can maintain technological leadership? The Japanese dominate completely in semiconductor memory (and in bi-polar logic). The American firms have each sought logic components of usual technology to advance their competitive positions: Cray with Gallium Arsenide chips for the Cray III, ETA with ultra-dense C-MOS chips cooled with liquid nitrogen for the ETA-10. (The latter chips are available from Honeywell as a result of that firm's participation in the VHSIC "Very High Speed Integrated Circuit" program of the Department of Defense. To date, these have not been manufactured to the specifications required for ETA to meet its advertised cycle-time of 7 nanoseconds.) But we see also that despite these efforts, at their rated performance levels the U.S. systems lose the megaflops/processor race (Table III). And each firm's effort is based on a favorable "spike" in an otherwise dismal U.S. semiconductor picture--which as yet shows no signs of improvement sufficient to reverse its overall, rapid decline.

From experience we can see that once the Japanese fully consolidate their semiconductor technology leadership, in time U.S. firms would be most fortunate if they found themselves to be only a generation or so behind them. The experience of Asia's "Four Tigers," Korea, Taiwan, Hong Kong, and Singapore confirms this point. The Japanese have consistently resisted the requests of the tigers for newer Japanese technologies. The tigers are explicitly told by the Japanese that this is because Japan looks upon them as "potential future competitors." As a rule, the Japanese refuse to license them for any technology less than five years old.

LONGER RUN IMPLICATIONS

The "vector facility" which provides vector capability to IBM's 3090 Mainframes, today blurs the boundary between the mainframe and the supercomputer. In any case, IBM has demonstrated its renewed interest in the field of supercomputing. This confirms the field's growing commercial importance--but IBM is not yet well established as a vendor of supercomputers.

The Japanese are established in supercomputing. They are challenging IBM in mainframes: Hitachi, with its plug compatible CPU's marketed by National Semiconductor's subsidiary, NAS; Fujitsu/Amdahl with their CPU's marketed through Amdahl Inc. (which recently experienced its best year). NEC's ACOS series is marketed in the U.S. through Honeywell. The mainframes from all three of these vendors also have some vector capability. Most Japanese supercomputers are "IBM-compatible." This is especially true of those from Fujitsu and Hitachi. (And the Fortran for NEC's SX-2 series is known as its "IBM version." For the SX-3 series NEC plans both an "IBM version" and a "Cray version.") Mainframes are growing larger and faster. But the Japanese appear to have intercepted that market from above through their supercomputers.

CONCLUSION

Once they were to achieve preeminence in supercomputers, the Japanese would have established themselves as the manufacturers of the most powerful computer systems, with supercomputers installed with the "Gold Chip" accounts of the world. They could then use that base and prestige to cement their move into the more lucrative market for commercial data processing systems. That is a long-term goal.

The threat is not just to supercomputers. The threat is to the entire U.S. base of computer systems. It is ironic that such cascading impacts could originate from dominance of a product area--DRAMs--that has reached almost commodity status. Yet such impacts appear not only possible, but likely. There is no denying the enormous impact that current DRAM shortages are already having on U.S. systems manufacturers. Their component costs are rising sharply. Many are unable to meet their orders due solely to lack of DRAMs.

The most recent developments with Sematech further confirm the depth of the crisis in U.S. semiconductor manufacturing at two levels. First, IBM and AT&T each has agreed to donate an advanced technology to Sematech. For IBM it is the four megabit DRAM. For AT&T it is the technologically equivalent one megabit SRAM. That these two arch rivals would see their way clear to share technology with each other and with others in the U.S. semiconductor industry demonstrates the depth of their concern. In part this is because these integrated firms buy their equipment for manufacturing semiconductors from others. And the Japanese are dangerously close to becoming the sole source of that equipment. Second, it is not encouraging that (as of this writing) this fledgling effort at cooperation by the semiconductor industry, Sematech, remains understaffed, organizationally in disarray, and well behind its self-imposed schedule.

At another level, it must be recognized that there is no assurance that even a Sematech highly successful in relation to its stated goals was set up to be coordinated with needed actions in relation to supercomputers--or with any other elements required for an effective national technology policy. This brings us back to the point made in the opening paragraph.

REQUIRED: A SYSTEM SOLUTION

The solution to the problems of the U.S. supercomputer manufacturers lies not in imploring the Japanese not to pursue their advantage, but in the U.S. taking positive actions at home to insure that the Japanese don't succeed at our expense. The U.S. must get its technology base in order: in semiconductors, in supercomputers and in software. No current initiative--including Sematech or IBM's own recent focus on the field of supercomputers--assures that these objectives will be achieved. A system solution cannot be achieved through discrete, uncoordinated initiatives, no matter how worthy.

Several studies over recent years make it clear that a solution to the problems of U.S. civilian technology/industry/trade policy, such as the matters discussed in this paper, will require expert coordination and leadership.[5] It seems clear

---

[5] For example, see "Global Competition: The New Reality," the Report of the President's Commission on Industrial Competitiveness, January, 1985; Superintendent of Documents, Washington, DC, 20402; and "The Technological Dimensions of International Competitiveness," May, 1988, Office of Administration, Finance, and Public Awareness, National Academy of Engineering, 2101 Constitution Ave, NW, Washington, DC 20418.

that to be effective, these can only by provided at the national level. Only at this level would it be possible for a true systems approach to be taken to the problems facing this country in the international competitiveness of its industry and of its high technology base. Action is required on long term R&D policies: it is necessary to integrate the activities of government, industry and universities; it is necessary to achieve timely, effective technology transfer from stage to stage in the process; it's necessary to ensure careful attention to the provision of suitable people-power to each stage in the process, an issue involving not just numbers, but also appropriate intellectual capital.

The challenge is to find an acceptable institutional framework in which government, industry and academia can pursue these objectives to the long-run benefit of the nation as a whole. A necessary requirement of such a framework is: it must ensure that economic and technological decisions be taken on the basis of economic and technological--as opposed to mainly political or military-- criteria. The answer may well require that the coordination and leadership functions be vested in a new, lean, expert, civilian agency of government that is capable of focusing on the longer term national interest. Only through a coordinated approach to all of the above issues will we be able to ensure a strong U.S. base for innovation, productivity, and international competitiveness--a base in which supercomputers constitute a vital factor.

For more information on the IEEE Scientific Supercomputer Subcommittee, contact:

Heidi F. James
IEEE/USAB
1111 19th Street, NW
Suite 608
Washington, DC 20036
(202) 785-0017

NEWS RELEASE

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

IEEE

FOR IMMEDIATE RELEASE

Contact: Pender M. McCarter
202/785-0017
or
Jayne F. Cerone
212/705-7847

VARIOUS JAPANESE TECHNICAL, MARKETING STRATEGIES

MAKE U.S. SUPERCOMPUTER FIRMS VULNERABLE:

IEEE SUPERCOMPUTER SUBCOMMITTEE REPORT

U.S. Lead in Supercomputing Could Be Lost;

IEEE Group Urges Focus on 'Longer-Term National Interest'

to Ensure Strong Technology Base

WASHINGTON, DC, August 8: Various Japanese technical and marketing strategies make U.S. supercomputer firms "vulnerable to loss of their world leadership," according to a report issued today by the Scientific Supercomputer Subcommittee of the IEEE Committee on Communications and Information Policy (CCIP). The report, titled "U.S. Supercomputer Vulnerability," was prepared by the CCIP of The Institute of Electrical and Electronics Engineers, Inc. (IEEE). It cites Japan's introduction of advanced machines and adoption of aggressive marketing techniques, including use of strategic delays in marketing high-speed computer chips in the United States. And the report recommended that the U.S. focus on the longer-term national interest to ensure a strong technology base.

If allowed to become preeminent in supercomputers, Japanese computer companies could then use that base and prestige to increase their role in the more lucrative market for processing systems, says the IEEE

-more-

NEW YORK, N.Y. • 345 EAST 47TH STREET, 10017
WASHINGTON, D.C. • 1111 NINETEENTH STREET, N.W. 20036

-2-

group of academic, government, and industry experts. It also observes that these companies include operations integrated over multiple production stages and across several product lines.

The report adds that, although the most advanced individual processors in Japanese supercomputers are superior to those used in U.S. machines, the Japanese still lag somewhat in their ability to perform parallel processing; that is, solving problems faster by dividing them into parts that can be handled by a number of processors simultaneously. However, new multiprocessor versions of Japanese machines "could far outstrip the best of the U.S. machines once the Japanese catch up in parallel processing, an area where they are hard at work," the report states.

Paradoxically, the increased "portability" of programs among supercomputers of different manufacturers, aided by a standardized FORTRAN computer language and a common UNIX operating system -- otherwise desirable, is also "hastening the progress of the Japanese," according to the IEEE group. It summarizes: "While in the past each [U.S.] supercomputer vendor had a 'captive audience' of users, now it is increasingly possible to move applications between and among vendors. Portable software does provide substantial benefit to users of super-computers, but it greatly diminishes the advantage previously enjoyed by U.S. firms." In addition, the report says, Japanese dominance in producing high-performance chips could mean that, "in time, U.S. firms would be most fortunate if they found themselves . . . [only to be] a generation or so behind [the Japanese]."

-more-

-3-

The document also addresses the recent U.S. supercomputer trade agreement with Japan. It states that "Japanese manufacturers have been more than willing to sacrifice short-term profit for long-term market penetration." In this way, the report continues, the trade agreement could work to the disadvantage of the U.S. as Japanese supercomputers become more competitive with U.S. machines. The document also notes the seriousness of the Japanese challenge, pointing to a slow start for Sematech, the industry consortium established to encourage technological advances in the semiconductor industry.

The IEEE-CCIP subcommittee stresses that "the solution to the problems of . . . U.S. supercomputer manufacturers lies not in imploring the Japanese not to pursue their advantage, but in the U.S. taking positive actions at home to ensure that the Japanese don't succeed at our expense." The group calls for the creation of Federal research and development policies that will integrate efforts of government, industry, and universities -- all focusing on the longer-term national interest to ensure a strong U.S. base in supercomputer and other critical technologies.

The IEEE-CCIP is chaired by Dr. John M. Richardson of the National Research Council. CCIP's Scientific Supercomputer Subcommittee is headed by Dr. Sidney Fernbach, a pioneer supercomputer user, former director of the computer center at the Lawrence Livermore National Laboratory, and now a consultant to Control Data Corporation.

# # #

[Note to Editors: Copies of the report are available from Pender M. McCarter, Manager, Public Relations, IEEE Washington Office, telephone (202) 785-0017; and Jayne F. Cerone, Coordinator, Media Relations, IEEE Headquarters, telephone (212) 705-8747.]

8/8/88

News in Perspective

SUPERCOMPUTING

IEEE Warns of the Japanese Supercomputer Threat

A new study finds that the U.S. must act now to stem the perceived Japanese invasion and suggests a new civilian agency focused on long-term national interests.

BY WILLIE SCHATZ

The IEEE's Committee on Communications and Information Policy has made available to DATAMATION a new report on how the Japanese are wiping out the U.S. supercomputer industry. Entitled "U.S. Supercomputer Vulnerability," it leads to the inexorable conclusion that "we better do something," according to its principal author, Alan McAdams, an assistant professor of managerial economics at Cornell University.

The Scientific Supercomputer Subcommittee of IEEE contends that the U.S. supercomputer industry is in deep trouble thanks to a focused market strategy by the Japanese. To overcome U.S. firms' vulnerability will require coordination by government to a degree seldom achieved in peacetime—and time is of the essence.

What's new about this? "The IEEE has never taken a position like this before," McAdams says. "This is the nonpartisan IEEE taking a policy position. Things must be pretty bad for them to start screaming."

Finding the Framework

The something McAdams refers to is finding "an acceptable institutional framework in which government, industry, and academia can pursue these objectives to the long-run benefit of the nation as a whole." Were that to occur, the institutional framework would still be useless without a guarantee that economic and technological decisions be made according to economic and technological—not military and political—criteria.

"The answer may well require that the coordination and leadership functions be nested in a new, lean, expert civilian agency of government that is capable of focusing on the longer-term national interest," find McAdams and friends. "Only through a coordinated approach to all these

CORNELL'S McADAMS: The threat isn't limited to supercomputers.

issues will we able to ensure a strong U.S. base for innovation, productivity, and international competitiveness."

Grass-Roots Commitment

"A new agency is fine if you've got a professional government with a long-term interest," says a government official intimately involved in the supercomputer industry. "But if Congress thinks it can just create it with a few pieces of paper, then it will be another bureaucracy that won't work. What you really need is a grass-roots commitment.

Hopefully, the new administration will have more of a commitment to understanding high tech. But it's still not worth creating another bureaucracy."

That opinion isn't confined to the government. "Any proposal to establish a new government agency isn't something I favor," says Sid Karin, director of the National Science Foundation's (NSF's) San Diego Supercomputer Center (SDSC). "I've heard all this before. There's nothing earthshaking in here."

The IEEE begs to differ. "People aren't realizing the real crisis that exists," McAdams contends. "The whole U.S. technological base is at risk, and it's getting worse.

To emphasize that this report is different from all the others that have reached the same conclusion, McAdams rests his case on economics and technology.

"To overcome [U.S. supercomputers'] vulnerability will require a systems solu-

tion: an integrated cooperative effort among industry, universities, and government," the report says. "It appears that such a solution will require coordination by government to a degree seldom if ever achieved in this country in peacetime. But the threat is real, and it is not limited to supercomputers. Supercomputers appear to represent only a next step in an ongoing process."

Banging the Drum Slowly

Thus, while Cray and ETA may say that Japanese components are not yet available for export to those two companies, the devices are readily available to end users in the Japanese supercomputer systems.

"This continues the familiar, oft-repeated pattern," the report contends. "Japanese firms plan and act in accord with long-range goals. When they achieve a technological advantage in one area, they use that advantage to insure their advance into new areas. They have targeted supercomputers as the next high-tech area in which to establish a dominant position."

You couldn't tell it from looking at the Japanese supercomputers in the U.S., though. The only one is an NEC SX 2 leased by the Houston Area Research Consortium (HARC). There have been many other efforts to land a Japanese supercomputer, but none has succeeded (see "Supercomputer Dumping Alleged at U.S. Universities," Sept. 15, 1987, p. 17).

The SDSC would just as soon keep it that way, but it sees the U.S. government tripping all over itself.

"The supercomputer agreement with Japan contains seeds which can further undermine the position of U.S. manufacturers," the report contends. "It requires of Japan a number of actions which, if made reciprocal to

Photograph by Bob Mahoney/Picture Group

the U.S., could facilitate the entry of Japanese supercomputers into the U.S. market."

Supercomputer Partnerships

Now, the supercomputer trade agreement requires the Japanese in government or university supercomputer procurements to give equal preference to U.S. manufactured supercomputers. The Japanese essentially have agreed to avoid unfair pricing by their manufacturers. According to McAdams, however, what looks unfair to a trade administrator looks to a manufacturer like recognition of partnership with a government agency or university.

"This sounds like a complete misunderstanding of the trade agreement," says Lauren Kelley, a supercomputer analyst in the Department of Commerce (DOC) Office of Computer and Business Equipment. "The agreement is based on the international GATT [General Agreement on Tariffs and Trade] government procurement code. In fact, nothing in the agreement is different from the standard General Services Administration procedures. To say this is a one-sided arrangement is completely untrue."

Nonetheless, the IEEE thinks a hard rain's gonna fall. Here's the U.S. telling the Japanese to open their markets or else, while simultaneously blocking the Massachusetts Institute of Technology (MIT) from purchasing an SX 2 from NEC. Of course, even a Freedom of Information Act search wouldn't uncover a written policy on the subject, but you can bet the national debt that government agencies aren't about to open their doors to the Japanese.

(The Department of Defense, which sees national security in every byte, is legislatively prohibited from buying

Peak Supercomputer Performance Rates

	SINGLE CPU PEAK 64-BIT MFLOP RATE	ALLOWING MULTIPLE CPUS
Cray-1	160	160
Cray X-MP	233	932 (4 cpus)
Cray-2	488	1,952 (4 cpus)
Cray-3 (1989)	1,000	16,000 (16 cpus) est.
Cyber 205	200 (2 pipe)	400 (4 pipe)
ETA 10 (1986)	350	1,400 (4 cpus)
ETA 10/E	415	1,660 (4 cpus)
ETA 10/G (1988)	643	5,142 (8 cpus)
Fujitsu VP 100	271	271
Fujitsu VP 200	533	533
Fujitsu VP 400	1,067	1,067
Hitachi S-810/20	630	630
Hitachi S-820/80	2,000	N/A
IBM 3090/VF	116	696 (6 cpus)
NEC SX 1	570	570
NEC SX 2	1,300	1,300
NEC SX 3 (1989)	5,000	20,000 (4 cpus)

Source: "U.S. Supercomputer Vulnerability," IEEE's Scientific Supercomputer Subcommittee.

any foreign—Congress meant Japanese—supercomputers in 1988.)

"How long can such a pattern last?" the paper asks. "Is it realistic to believe that the Japanese machines are foreclosed from U.S. institutions? If the requirements we now impose on the Japanese were made reciprocal for U.S. universities and government laboratories, Japanese supercomputers would have to be acceptable to those agencies on nondiscriminatory terms."

For some supercomputer users, that day can't dawn soon enough.

"We Want the Best Product"

"The economic leverage of supercomputers is irrelevant and always will be," NSF's Karin says. "What matters is the use of supercomputers. We need the best supercomputers, and we need to make the best use of them. Who makes a supercomputer is far less important than how

it's used."

"There's a genuine concern here about foreign competition," says a user at a major federal lab. "But when it comes to computers, we just want the best product. And by keeping out the Japanese, the government and the U.S. supercomputer industry are pretending the situation is better than it really is."

However, by letting in the Japanese, the IEEE sees the industry living on desolation row.

"The new requirements could greatly facilitate the entry of the Japanese into the U.S. At the same time such requirements could disrupt the implicit partnership between U.S. manufacturers and U.S. government laboratories and/or U.S. universities, and transfer the benefits of partnership to Japanese firms. In response to a low bid, U.S. national labs could be required to become partners to Japanese firms in perfecting their

systems for penetration of U.S. markets."

No U.S. lab would want to do that, at least on the record. But the Japanese clearly have the fast single-processor system and are expected to increase that lead with their next generation product expected next year (see "Peak Supercomputer Performance Rates"). So how much longer can users be shut down at their expense? Not very. So it's only a matter of time before HARC has company.

Japan's Software Is Lacking

That could be very soon if the Japanese get their software act together. Their software isn't quite up to their hardware, but therein lies the danger.

"Once general portability of applications is achieved, the avenue to continued market leadership by U.S. firms would be solely through technical leadership," the report says. "But how can U.S. firms simultaneously rely on componentry manufactured by their competitors, the Japanese, and assure their customers that they, the U.S. firms, can maintain technological leadership?"

They can't.

"As soon as Japanese firms have software that U.S. companies need, they'll be selling heavily here," the government official says. "NEC is working its butt off to develop software. When those developments take place, we have no laws to restrict them." This is generally expected to be sooner rather than later.

So why not let them come and fight it out nanosecond-to-nanosecond in the tried-and-true capitalist tradition?

"Because our entire economy is at risk," McAdams contends. "Supercomputers are the key to industrial design. If you lose supercomputers, you're in real trouble." ∎

SOFTWARE FOR SUPERCOMPUTERS
— A REPORT —

Prepared by the Scientific Supercomputer Subcommittee
of the
Committee on Communications and Information Policy
United States Activities Board

INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS

1988

SOFTWARE FOR SUPERCOMPUTERS
A REPORT

EXECUTIVE SUMMARY

This is a summary of a report prepared by the Scientific Supercomputer Subcommittee of the Committee on Communications and Information Policy, United States Activities Board, Institute of Electrical and Electronics Engineers.

Early supercomputers had poor software.

The first truly unique supercomputer architectures started to appear in the early 1970s. The Burroughs Illiac IV, the Texas Instruments ASC, and CDC's Star 100 were built in small quantities and the software was not highly developed. In many cases the purchasers provided most of the software themselves. It is important to understand that the very uniqueness that allows these computers to yield very high performance also forces users to expend significantly more effort in optimizing their codes to achieve even a fraction of this potential power. These early machines were very difficult to use and the software was not easily optimized. When the current class of supercomputers (called Class VI machines) started to make their appearance in the late 1970s and early 1980s, new software had to be provided. Because of difficulties in formulating and optimizing the programs appropriately for these newer vector computer architectures, the efficiency of use of Class VI architectures suffered (and still does).

Portability, optimization, algorithms.

We have three main needs associated with supercomputer software: portability; language and compiler related software, especially automatic optimization; and architecture-appropriate algorithms.

Portability of code was and is an important problem.

The ability to take programs from one manufacturer's machines to another or even to move code to later generations of the same equipment is known as portability. This has been a continuing problem in the computing industry, especially in supercomputing. Lack of portability not only causes premature obsolescence of users' codes, but also shortens the lifetime of system code supplied by the manufacturer, thus making it even more difficult to justify the heavy cost of system code development. Some early government users of supercomputers have tried over the years to provide continuity from one generation to the next. For example, a group at the Lawrence Livermore National Laboratory designed the Livermore Time Sharing System (LTSS) operating system in 1963 for the

Control Data Corporation 1604 computer and continued its development through later CDC computers, including the CDC 3600, 6600 and 7600, thus providing compatibility from generation to generation. The version that runs on Cray computers is now called CTSS and is being used at Livermore and by other DOE laboratories. This is one example of how a national software center could help to provide some leadership in the area of portability.

Supercomputer software is expensive, difficult.

Writing good supercomputer software is especially difficult because of the need to take advantage of the complex architectures needed for high performance. For example, program optimization is an area of crucial importance to supercomputers. Good automatic optimization is required to achieve a higher percentage of the potential speed of supercomputers, better utilization of scarce manpower, and better portability. A good deal of work is now being done on automatic program optimization. But even after using the best of these optimizers, the performance typically obtained from the machine is far less than the peak performance possible. Not only are vector optimizers not sophisticated enough, but simply applying vectorization techniques to make use of multiprocessors is not enough. Whole new techniques are needed to get acceptable multiprocessor performance.

Newer (multiprocessor) architectures not only make the problem worse, they create whole new problems.

Machines with new architectures possessing highly parallel structures are now being designed and built. At the moment, we are exploring the capabilities of high performance systems containing only a few parallel processors. A number of supercomputer systems being planned are somewhat larger, having up to 16 processors. Yet good optimization software does not yet exist even for these low levels of parallelism. Machines with new architectures possessing highly parallel structures including hundreds, even thousands of processors, are now being designed and built. Optimization for these machines promises to be even more difficult and labor intensive than the last generation of machines. This optimization is not just harder, it poses new problems not encountered before. Efforts to design automatic optimization software to alleviate this problem are at a very early stage and the costs involved in developing this software are so high and the efforts to develop it are so fragmented, that very little may ever see the light of day.

Better parallel algorithms are needed.

Better algorithms can make a major difference in the feasibility of some applications. One only has to think of Fast Fourier Transforms (FFT) and the Simplex method to recognize the impact better algorithms can have. Algorithms are especially important on supercomputers because they need to be specially designed to take advantage of the vector and multiprocessor parallelism. Fortunately, some more economical computers now available make it possible to experiment with new parallel

algorithms, and many of these systems are being used in this fashion. However, we are a long way from saying that we know how to use parallel processors efficiently for most problems. We now have the tools to study parallel algorithms and we must make these tools available to the algorithms community.

Need for coordination. Many of our difficulties stem from a lack of coordination of effort. Manufacturers were reluctant to cooperate out of fear of antitrust laws, and they were reluctant to finance significant software development for what they viewed as a short term product. When the Japanese entered the supercomputer competition, they took a more global approach to supercomputers and treated them as important, highly marketable entities. Having established supercomputers as a national priority, they were able to take a longer term view of software development for these machines, a policy which is proving to be successful. The performance of Japanese machines shows the results of their foresight by outperforming the U.S. supercomputers in many instances (see Table 2), despite their relatively recent entry into the supercomputer arena. One should wonder how the performance of U.S. machines will compare with our competition in a few years.

Recommendations.

Risk sharing, coordination, and national laboratories. This subcommittee makes five recommendations aimed at improving the health of the U.S. supercomputer industry. These recommendations include more participation from the federal government in the financial risks involved with supercomputer hardware and software development, better coordination of U.S. activities in this area, and the establishment of national software laboratories whose purpose would be to provide software specific to supercomputing.

-3-

SOFTWARE FOR SUPERCOMPUTERS

This is a report prepared by the Scientific Supercomputer Subcommittee of the Committee on Communications and Information Policy, United States Activities Board, Institute of Electrical and Electronics Engineers.

1. Introduction.

IEEE.

The Institute of Electrical and Electronics Engineers (IEEE) is the world's largest engineering society, with over 295,000 members world-wide, 223,000 of whom live and work in the United States. The United States Activities Board of the IEEE takes the position that advanced scientific computing capability is a technology that must be accelerated in the United States. This technology is crucial for national defense, economic growth, and advances in engineering and science.

Background.

In 1983, the United States Activities Board of the IEEE set up a special committee to examine the U.S. position in supercomputer development and to recommend actions the government should take to ensure continued preeminence of the U.S. in this challenging and complex field. Among its five recommendations, the committee proposed that the Federal government make a long-range commitment to maintaining leadership in supercomputer development and take an active role in fostering development of new systems.[1] The establishment of this committee on supercomputers and the continuing support of the its recommendations demonstrate the continuing deep concern of the IEEE on this matter.

Need for strong government role.

This subcommittee reiterates the position taken in the prior report and restates, in the strongest terms, its own belief that the government should pursue an active role. Supercomputer software is not just important to the position of the U.S. supercomputer industry in the world market, but it is also crucial to a much broader spectrum of industries that depend on supercomputers for the design of competitive products, not to mention its strategic value to U.S. defense. Toward this end, the government must allocate sufficient research funds to improve components, architecture, systems software, applications software, and very high performance peripheral equipment.

---

[1] "Scientific Supercomputer Committee Report," produced by the Scientific Supercomputer Subcommittee of the IEEE U. S. Activities Board, Sidney Fernbach, Chairman, October, 1983.

Slow progress since 1983.

In the intervening years since October 1983, much progress had been made on many fronts. New programs budgeted in the tens of millions of dollars have been started in several agencies including DOE, NSF, NSA, and NRL. Prior efforts have been redoubled in all three arenas: the universities, the private sector, and the Federal government. These are hopeful auguries, but the advances are only incremental. Neither these steps nor those currently being considered will be enough. Thus, this subcommittee further proposes a particular action agenda to enable the universities, the private sector, and the government to work together to insure our preeminence in supercomputers.

2. Historical Background.

Early machines evolutionary, good software evolved from existing products.

The definition of a supercomputer is that it is the most powerful scientific system available at any given time.[2] Early machines in this category, e.g., the IBM 7094, CDC 6600, and IBM 360/91, topped off the standard line of the manufacturer. Hence, based upon the fairly large sales volume for the low end of the equipment, a reasonable amount and quality of software was supplied with them.

Limited-production machines had poor software.

With the advent of "one-of-a-kind" machines, the situation started to change somewhat. The IBM NORC (only one existed), the Sperry-Univac LARC (only two were built) and the IBM STRETCH (nine were built) each had unique software that was incompatible with that of other systems. These machines tended to be revolutionary rather than evolutionary, and they lacked the market volume to permit the extensive software development required for a revolutionary machine.

Exception: 6600 successful, users helped with the software.

An exception to the rule, and perhaps the first highly successful supercomputer was the CDC 6600. Control Data Corporation (CDC), in an important symbiotic relationship with some of its customers (the national laboratories), was able to provide software for the system and went on to sell a large number of machines. The successor to the 6600 was the CDC 7600. The compiler for the 6600 formed the basis for the 7600 compiler, and with this head start and more help from a few customers, adequate software was again available. Thus many copies of the CDC 7600 were sold, and it too became quite successful.

Newer architectures require more effort to optimize.

Then even newer architectures began to emerge, architectures which required far greater effort to restructure or optimize programs to utilize the higher power available in these machines. When the current class of supercomputers (called Class VI machines) started to make their appearance in the late 1970s and early 1980s, the software had to be redone yet

---

[2] *Supercomputing — An Informal Glossary of Terms,* IEEE, 1111 Nineteenth Street, NW, Washington, DC, 1987.

again. These machines were very difficult to use and the software was not easily optimized. In many cases the purchasers provided most of the software themselves. Because of the difficulties of restructuring problems to take advantage of the vector oriented machine architecture, much of the potential of these machines failed to be realized. It is not unusual for sustained performance on these machines to be less that 20% of the peak performance. Once again we see that industry has not had the immediate market base to justify this software development, nor the lead time to accomplish this development.

3. What do we need?

Why manufacturers have not provided good software for supercomputers.

Because of their cost and complexity, the market for supercomputers has been small and, additionally, writing systems software for these machines has been particularly difficult. The result has been a significantly lower standard for scientific supercomputer software, a standard we have been able accept only because of a small and sophisticated cadre of users. For years manufacturers were able to deliver their products to the few customers who could afford them--customers who could also afford to do most of the software development themselves. Because of the limited volume of the market, short viewpoints of the financial community, and limited lifetime of the software (due to lack of portability), expensive software development projects could not be justified.

Broadening market and foreign competition require better software.

Industry is now beginning to realize that there are important economic advantages in exploiting modern supercomputers. Many new supercomputer installations are being set up to satisfy the needs. See Table 1 for current or currently planned installations. Now that the market for supercomputers is broadening, new classes of users require a much higher standard, especially if our industries are to benefit from the superior design capabilities afforded by supercomputers.

Why the Japanese are successful.

When the Japanese started to build supercomputers, they recognized this broadening market and they took a more global approach to supercomputers, treating them as another important, marketable commodity. Having established supercomputers as a national priority, they were able to take a longer term view of software development for these machines, a policy which is proving to be successful. (See Table 2 for a comparison of performance of certain computers as on the LINPACK programs as evaluated by Jack Dongarra of Argonne National Laboratory.) Obviously, the Japanese manufacturers have recognized the areas in which they should apply strong efforts. *One should wonder how the performance of U.S. machines will compare with that of our competition in a few years.*

TABLE 1

FREE WORLD DISTRIBUTION OF SUPERCOMPUTERS

(INSTALLED OR ON ORDER AS OF 12/31/87)

COUNTRY	NUMBER OF SUPERCOMPUTERS
UNITED STATES	134
JAPAN	70
UNITED KINGDOM	19
GERMANY	17
FRANCE	16
CANADA	7
HOLLAND	3
NORWAY	3
SWEDEN	2
SWITZERLAND	2
ABU DHABI	1
AUSTRALIA	1
ITALY	1
SAUDI ARABIA	1
TAIWAN	1

FREE WORLD DISTRIBUTION BY APPLICATION

(APPROXIMATE)

APPLICATION:	NUMBER OF SUPERCOMPUTERS
Research:	81
Defense:	45
Universities:	45
Aerospace:	32
Petroleum:	30
Weather:	16
Nuclear Energy (weapons, reactors):	12
Automotive:	11
Service Bureaus:	9

-7-

TABLE 2

COMPUTER PERFORMANCE

SOLVING A SYSTEM OF LINEAR
EQUATIONS WITH LINPACK†
(FULL PRECISION--ALL FORTRAN)

SYSTEM	MFLOPS
*ETA 10-E (1 proc. 10.5ns)	52
*NEC SX-2	43
*CRAY X-MP-4 (1 proc. 8.5 ns)	39
*NEC SX-1	36
*NEC SX-1E	32
*CRAY X-MP-2 (1 proc.)	24
*CRAY-2 (1 proc.)	21
*AMDAHL 1200	19
*CDC CYBER 205 (2-pipe)	17
*FUJITSU VP-200	17
*HITACHI S-810/20	17
*CRAY 1S	12
*IBM 3090/180 VF (1 proc.)	12
FUJITSU M-380	6.3
CDC CYBER 875	4.8
AMDAHL 5860 HSFPF	3.9
CDC 7600	3.3
IBM 3090/120E	3.1
CONVEX C-1/XP	3.0
FPS-264/20 (M64/50)	3.0
IBM 3081 K (1 proc.)	2.1
HONEYWELL DPS 8/88	1.7
AMDAHL 470 V/8	1.6
IBM 370/168 (fast mult.)	1.2
AMDAHL 470 V/6	1.1
ELXSI 6420	1.5
DEC VAX 8600	.48
IBM PC (W/8087)	.012

*These machines are generally acknowledged to be supercomputers.

†Data from J. Dongarra in "Performance of Various Computers Using Standard Linear Equations Software in a Fortran Environment," Table 1 (February 2, 1988, Linpack, full precision, no Blas), *Computer Architecture News*, Vol. 16, No. 1, March, 1988.

Supercomputer software is difficult to write.

Writing good supercomputer software is especially difficult because of the need to take advantage of the complex architectures needed for high performance. For example, program optimization is an area of crucial importance to supercomputers. Good automatic optimization is required to achieve a higher percentage of the potential speed of supercomputers, better utilization of scarce manpower, and better portability. A good deal of work is now being done on automatic program optimization. But even after using the best of these optimizers, the performance typically obtained from the machine is far less than the peak performance possible. Not only are vector optimizers not sophisticated enough, but simply applying vectorization techniques to make use of multiprocessors is not enough. Whole new techniques are needed to get acceptable multiprocessor performance.

Newer (multiprocessor) architectures make the problem worse.

Machines with new architectures possessing highly parallel structures are now being designed and built. At the moment, we are exploring the capabilities of high performance systems containing only a few parallel processors. A number of supercomputer systems being planned are somewhat larger, having up to 16 processors. Yet good optimization software does not yet exist even for these low levels of parallelism. Machines with new architectures possessing highly parallel structures including hundreds, even thousands of processors, are now being designed and built. Optimization for these machines promises to be even more difficult and labor intensive than the last generation of machines. This optimization is not just harder, it poses new problems not encountered before. Efforts to design automatic optimization software to alleviate this problem are at a very early stage and the costs involved in developing this software are so high and the efforts to develop it are so fragmented, that very little may ever see the light of day.

What do we need?

What do we need? 1) Better portability, so that software has a longer lifetime and can therefore sustain more development; 2) Better program optimizers, so that users can spend their time more productively, and get a higher fraction of the potential power of the supercomputer; 3) Better algorithms, again to achieve a higher fraction of the potential power; 4) Better languages and operating systems, to improve ease of use, expressive efficiency, and execution efficiency;

Parallel algorithms.

Better algorithms can make a major difference in the feasibility of some applications. One only has to think of Fast Fourier Transforms (FFT) and the Simplex method to recognize the impact better algorithms can have. Algorithms are especially important on supercomputers because they need to be specially designed to take advantage of the vector and multiprocessor parallelism. Fortunately, some more economical computers now available make it possible to experiment with new parallel

algorithms, and many of these systems are being used in this fashion. However, we are a long way from saying that we know how to use parallel processors efficiently for most problems. We now have the tools to study parallel algorithms and we must make these tools available to the algorithms community.

Better languages are needed.

Once algorithms have been designed, better languages must be provided to allow more efficient expression of these algorithms and to allow their efficient execution. Fortran is the traditional language of scientific computation and a new standard incorporating vector extensions is expected to be out soon. Two reasons for the popularity of Fortran, even in the face of its age, are its execution efficiency and its portability. Unfortunately its portability is limited: It is easy to move a Fortran program from one machine to another, but the level of the machine dependent detail which users put into their program to gain execution efficiency often does not prove effective on another machine. Low level details included in a program to improve execution efficiency on one machine may (and often do) prove detrimental on another machine. Thus the program may execute correctly on another machine, but with considerable loss of efficiency. In any event, other languages and programming paradigms are proving to be as portable as Fortran and considerably more expressive.

Portability.

One does not wish to program all problems for all machines, especially when it means reprogramming each program in order to get optimum performance. But there has been little or no compatibility between supercomputer systems. Originally, the manufacturers desired it this way. It was easier to lock a customer into a given series of machines by making conversion to another much too difficult. This may be good for individual manufacturers, but certainly makes for great difficulty at the user level, whether it be in government or industry. The resulting loss of productivity is not good for either the customer or the U.S. economy. Fortunately, driven by forces like the popularity of the Unix operating system and the demands of users, manufacturers are changing their ways. Portability has a chance of becoming a reality. But portability requires much more to be done both in the area of standards and optimization.

Advantages of portability.

Portability will greatly extend the lifetime of software, both applications software and systems software. One only has to look at the growth of Unix to see the potential for portable operating systems. Given these longer lifetimes, we can afford considerably more effort to achieve better software, and users will be able to spend more of their time on truly creative work like designing new algorithms.

Optimization.   True portability requires a high level representation of algorithms, with no machine dependent semantics. However, a sophisticated program optimizer is needed to get the required machine efficiency. Thus, for example, details of memory hierarchy management—vector registers, cache, virtual memory— should not be part of the user's program but must be optimized by the software itself. This allows true portability, and also permits more productive use of the programmers' time.

4. How do we get what we need?

Earlier reports and actions.   During the early 1980s it was realized that the U.S. scientific effort was suffering from a lack of adequate facilities to carry out large scale research projects. The Lax Report[3] to the National Science Board pointed out the lack of availability of supercomputer resources for researchers in our universities. Through supercomputer center grants from NSF, this situation is beginning to be corrected. We are placing hardware at many sites, so that availability is being increased significantly. The Lax Report also urged the need for training in the use of supercomputers, and for research and development to plan future generations of supercomputers. Little has been done in these areas, as yet. A more recent report[4] reiterates these concerns.

Standards.   The IEEE and similar organizations should play a larger role in establishing standards for languages and operating system software that would improve portability, similar to the role played in the establishment of the Posix standard for Unix. In the supercomputer marketplace, the government's share of the market is close to 41%. It is expected that the U.S. Government would do its part by adopting these standards and enforcing them on government purchases. Such standards will enhance portability, longer software lifetimes, and broader markets.

Incentives and disincentives.   In the past, most software was written by the hardware vendor or by the customers, but this is beginning to change. Software is increasingly provided by third parties, independent of the vendor's particular hardware. Given a larger, multivendor market for software, and the resulting increase in motivation for portability, better software results. But the development costs for supercomputer software are still enormous, and most software houses prefer to concentrate on markets like the personal

---

[3] *Report of the Panel on Large Scale Computing in Science and Engineering,* Peter D. Lax, Chairman, National Science Foundation, December, 1982.

[4] *A National Computing Initiative: The Agenda for Leadership,* produced under the auspices of the Federal Coordinating Council for Science, Engineering, and Technology (FCCSET), sponsored by NSF and DOE, and published by SIAM, Philadelphia, 1987.

computer market where the volume is high and the software easy. Incentives from the government such as guaranteed purchases (software and hardware) and even direct development contracts are needed to focus some of the attention of the third party software vendors on supercomputer software. Additionally, vendors' investments in software need adequate protection under the law. And finally, not only should any impediments to cooperation between vendors (hardware and software) and customers (especially national laboratories) be removed but such cooperation should be encouraged. Stimulation of better software for supercomputers and the protection of the investment in this software development involves many complex issues beyond the scope of this report. These issues need to be discussed and resolved.

National software centers.

Additionally, a mechanism needs to be set up to both better focus our efforts in the area of supercomputer software and to take on the most expensive and risky elements of software development, much as the national laboratories take on these risks in the areas of energy, weather research, and health.

5. Recommendations.

Importance of super-computers and need for government directed focus.

Supercomputer software is important not just to the position of the U.S. supercomputer industry in the world market, but is also crucial to a broad spectrum of industries that depend or will come to depend on super-computers for the design of competitive products. This subcommittee believes that the importance of supercomputers in government and industry is just being recognized. Nevertheless, software for supercomputers remains underdeveloped due to the relatively small size of the supercomputer software marketplace (compared, for example, to the market for workstation and personal computer software) and the fragmented and uncoordinated efforts in this area. Although some attempts have been made to remedy the situation, we believe that it would be in the best interests of the United States if the government were to provide more focus on this problem through the following actions:

(1) Stimulate the supercomputer industry by underwriting some of the costs and risks of hardware and (especially) software development. This might be done through a program where the government, during the early stages of the development cycle, commits to purchase supercomputer systems (providing they meet certain performance requirements). Not only would this help to underwrite risks associated with these machines, but would provide a stronger voice in their design.

(2) Improve the state of supercomputer software by direct research and development contracts and grants to industry and government laboratories.

(3) Increase basic research funding in supercomputer software.

(4) Establish a formal coordinating body to better focus existing development efforts through standards for software portability, and to provide interagency coordination of Federally funded research efforts.

(5) Establish several laboratories, such as the National Supercomputer Software Research and Development Institutes recommended in the SIAM report[5] and in an earlier report from this subcommittee,[6] to be associated with existing supercomputer centers, both academic and at national laboratories. The early successes of some of the present national laboratories is evidence of the potential for success of such institutes. These institutes should have the following goals.

- Advise the Federal government on matters relating to supercomputing;

- Set common software specifications for supercomputers;

- Carry out practical research in structuring algorithms and applications for supercomputers, including parallel (multiprocessor) algorithms;

- Develop software packages, including operating systems and compilers, that would be suited for a wide variety of supercomputers;

- Devise performance measures for supercomputers; and

- Package these products for U.S. government, educational, and industrial use.

---

[5] *Ibid.*

[6] *Software for High Performance Computers,* Prepared by the Subcommittee on Supercomputers of the Committee on Communications and Information Policy of the Institute of Electrical and Electronics Engineers, December, 1985, Washington D.C.

Participating Members of the Subcommittee on Scientific Supercomputing

Sidney Fernbach, Chairman
Heidi F. James, Committee Secretary
Alfred E. Brenner
James F. Decker
Duncan H. Lawrie
Alan McAdams
Kenneth W. Neves
John P. Riganati
Stewart Saphier
Paul B. Schneck
Lloyd Thorndyke
Hugh Walsh

The IEEE Subcommittee on Supercomputers of the Committee on Communications and Information Policy has produced a number of position papers on supercomputers. For further information, or to be placed on a continuing mailing list, contact:

Heidi F. James
IEEE Washington Office
1111 19th Street, N.W.
Suite 608
Washington, D.C. 20036
(202) 785-0017

Production: — Fri May 20 17:29:06 CDT 1988
File: uicsrdsw:/xd00f/homes/lawrie/IEEE-CS/Fernbach/software.position

Science Revision: 5/16/88

THE COMPUTER SPECTRUM:

A PERSPECTIVE ON THE EVOLUTION OF COMPUTING

IEEE Scientific Supercomputer Subcommittee

Since the mid 50's, the modern computer has rapidly evolved to satisfy the computational needs of an increasingly large fraction of society. This evolution is discussed and the development of the various classes of computers that have come into being to satisfy the increasingly diverse computational needs which have developed is explored. Based upon this historical analysis, projections for the near future directions of growth are offered.

Computers in widespread use today span a range of sizes and a range of applications much greater than that of any other manufactured product in our highly technological society. Quantitatively, the costs of manufactured items which might properly be called computers range from $20 for hand held programmable devices to more than $20 million for the largest supercomputer systems in the marketplace today - a factor exceeding 10^6 in cost and a similar factor in terms of the computational power and storage capacity. Although their range is not quantifiable, the breadth of applications for which these computers are put to work is equally impressive.

The IEEE Scientific Supercomputer Subcommittee is a technology-policy subcommittee of the IEEE United States Activities Board, 1111 19th Street, N.W., Washington, D.C. 20036. The subcommittee members participating in the development of this paper are:

Alfred E. Brenner, Supercomputing Research Center, Lanham, MD
Sidney Fernbach (Chairman), Consultant, San Jose, CA
Duncan Lawrie, Center for Supercomputing R&D, University of Illinois, Urbana, IL
Alan McAdams, Cornell University, Ithaca, NY
Kenneth W. Neves, Boeing Computer Services, Seattle, WA
John M. Richardson, National Research Council, Washington, D.C.
John P. Riganati, Supercomputing Research Center, Lanham, MD
Stewart Saphier, Department of Defense, Washington, D.C.
Paul B. Schneck, Supercomputing Research Center, Lanham, MD
Lloyd Thorndike, ETA, St. Paul, MN
Hugh Walsh, IBM Data Systems Division, Kingston, NY

To satisfy this broad spectrum of use and to accommodate to the pocketbooks of the broad range of potential customers, manufacturers of computers continue to develop at a dizzying rate new types of computers and specialized software to satisfy particular market niches. In fact, the options available are an elegant example of the diversity that mass production has produced in the modern world as so pointedly observed by Alan Toeffler in his book Future Shock (1). Although there is a multitude of terms in use today identifying classes of computers, there seems to be little precision or even agreement on these classifications.

This paper offers a perspective on appropriate classifications for computer systems and gives a historical review of their evolution to the present time. Then, with the hindsight of this historical analysis, projections into the near future are made giving what further changes and evolutions may be expected.

The First Two Computer Classes: Business and Scientific

In the early days of the modern computer era, i.e., during the decade of the 1950's, all machines were thought to be equally useful on all problems. It should be noted, however, that while the range of applications was quite diverse, the extent of use at that time was quite limited. The computers used either binary or decimal numerical representations. Soon, two marketplaces developed. One for business applications and the other for scientific applications. In the former case, fixed-point decimal arithmetic seemed in order; for the latter, floating-point binary arithmetic seemed preferable. These differences persist, more or less, until the present day.

During the latter part of the 50's and early into the 60's, the two marketplaces were developed more or less as separate entities with separate cultures. Typically computer architectures focused on one or the other of the two marketplaces. However, starting with the announcement (2) by IBM of the System 360 in 1964, there was an attempt by manufacturers to coalesce at least the hardware (3) satisfying these two marketplaces (4). Thus the architecture of the medium to large-scale computers developed in the mid-60's

included both decimal and floating-point arithmetic. Although unified operating systems were introduced to satisfy both communities, separate high level languages and applications packages were developed to satisfy the specific needs of each of these two major marketplaces. These machines continued to evolve through the 70's and 80's into quite general purpose computing engines, approximately equally capable of serving both classes of use. Today these have come to be called mainframe computers.

The Emergence of New Computer Classes: the Supercomputer and the Minicomputer

Starting in the 60's, there was a persistent and major demand for scientific computers yet more powerful than the mainframes then extant. This new market niche, driven by the needs of the U.S. Government for national security purposes and by the need of the meteorological and seismic communities, was developed at the very high end of the scientific computing spectrum. The computers serving this market niche, the most powerful scientific computational systems available, have come to be called "supercomputers" (5).

The demand for ever faster computer performance for scientific problems has persisted for over four decades. Approximately in the mid 1970's, the capabilities of supercomputers crossed a critical threshold allowing "computational science" to develop explosively. Computational science now stands, *de facto,* alongside the theoretical and experimental sciences as a fully legitimate field. In almost every scientific and engineering discipline, technology development is coming to rely more and more heavily on computer simulation, since only a small fraction of desirable experiments can be carried out physically in a cost-effective or completely realistic way. This has fueled the ever-increasing demand for more scientific computing power.

Early in the modern computer era, the technical and business communities quickly rose to the challenge of putting computers to active use. However, the early systems were all physically quite large and very costly. High cost inhibited the proliferation of these

(mainframe) systems. Yet, a very large potential marketplace awaited computers with price tags substantially below those of the then-available mainframes. A new class of computer system, the forerunners of today's minicomputers, with lower computational speed but more than proportionally reduced cost were introduced, primarily by new companies. Early examples of these were Digital Equipment Corporation's PDP 1 (6) and IBM's 650 (7). Other entries from additional manufacturers entering the field soon followed. Thus, by the mid-60's, although the general scientific and business marketplaces were still distinct, there were now three products available in the marketplace: minicomputers, mainframes, and supercomputers.

As with all new products, variations soon began to appear to meet the requirements of different niches in the marketplace. The major categories remained, but new systems were introduced with variations in performance, size and kind of memory, word length and types of instructions. Software was further diversified in an attempt to effectively address the particular needs of the wide variety of users in the marketplace. Price, size, ease of use, and relative performance remained the major factors determining the success of the various offerings. The price was the most important single factor, a feature that persists to this day. The minicomputers, which established a price niche, have persisted to this day with price tags which span a very wide range, from quite small to the high end superminicomputers which are larger and more capable than many of the current low-end mainframe systems.

The Evolution of the Supercomputer

In the 1970's, two new architectural features - vector processing and parallel processing - that would have widespread repercussions made their first appearances. These were typified by the STAR 100 (9) - a vector machine (10) built by Control Data Corporation, and the Illiac IV (11) - a parallel processor (12) designed at the University of Illinois and built by Burroughs Corp. Vector processing was the first of the two to be introduced into commercially available supercomputers. It gave rise to major performance

5

improvements. However, additional gains achievable through architectural, technological, and software improvements to this approach are reaching saturation. It is now quite apparent that major additional increases in computational power using current technology must be derived primarily from parallel processing approaches. All high performance computer vendors have begun to introduce this approach into their product lines.

Over the years, continuing advances in technology have allowed for increasing supercomputer performance while total system prices have remained essentially constant. Typically, supercomputer acquisition costs are in the range of from 10 to 20 million dollars. Despite the rapidly falling price/performance ratio, the high acquisition cost continues to be a factor inhibiting more widespread use of such systems, even in those cases where the life cycle payoff may far exceed the investment.

In an attempt to make supercomputers more readily available to a broader group of users, high performance, wide-area networking is now being introduced to allow for effective remote access to these costly facilities. Communication lines operating at the highest speed commercially available interconnect user equipment at local sites to remotely located supercomputers. Most often, the communications are handled by front-end computers or other smaller computer systems at the supercomputer site so as not to burden the supercomputer with communications overhead. The computers used as front-ends span a wide spectrum of capabilities, depending on what is readily available or affordable to the given user.

Because of the high cost of supercomputers and/or the additional complications involved in accessing such facilities remotely, there are continuing strong pressures to find alternatives to solving problems by means other than supercomputers. This gives rise to a continuous search for other more readily available, typically less expensive, computing resources. When smaller mainframes or even minicomputers are adequate to the

problem solution, such alternatives do work well, but only because these problems do not require supercomputers. Nevertheless, there remain large (growing) numbers of problems which are sufficiently demanding in their computing requirements that only supercomputer performance level facilities are appropriate.

The Emergence of the Minisupercomputer

By the 1980's, these pressures gave rise to yet another category of computer, the "minisupercomputer." These are relatively inexpensive, readily available, high performance computing engines which are architecturally similar to supercomputers, but typically use the manufacturing methods of minicomputer vendors. Their cost range makes them available to small groups of users (departments) within a company or university. Great incentives existed for vendors to satisfy this newly developing market niche. As has happened in the past, new players perceived the developing market niche and quite a large number of new firms are trying to fill the gap.

The minisupercomputer is a high performance computing engine with a much lower cost than that of a leading edge supercomputer. In many senses, this gives rise to the "personal" supercomputer. The engineering trade-off here is in favor of economy - not performance - in contrast to the performance over cost trade-off which drives supercomputer design. There is a very high demand from the research community and industry for such economical machines. It is too early to tell how broad is the funding base to support this class of need. This approach appears to be the "poor" (not wealthy) scientist's and engineer's preferred alternative to the remote access to a large supercomputer. A major factor in this reference is the advantage of local rather than remote control of the resource.

Two technological factors are having a major impact on the continuing development of the minisupercomputer. These are the microprocessor and parallel processing. The microprocessor - a full fledged processor on a chip - made possible the development of

the low-cost personal computer whose sales now number in the millions of units. Revenue from these, in turn, have fueled the development of even more powerful microprocessors. Secondly, the advancing technology utilized by successive generations of supercomputers has slowly been reaching the natural limits imposed by the speed of light and material limitations for the processees which have become the basis for increasing technological improvements. Consequently, major future improvements must come from an approach which attains high performance by architectural designs capable of parallel processing on a single computation. This is the direction taken by all manufacturers involved in high performance computing, and includes large mainframes as well as supercomputers. Also, the juxtaposition of microprocessors into parallel processing architectures have made it possible for many universities and many small firms to engage in low-cost experimentation on parallel processing. This has given rise to a large number of quite modest, mostly start-up, commercial entities competing for a broadening minisupercomputer niche.

The minisupercomputer will not supplant the supercomputer. There will always be a need for larger and faster machines. However, there will also be a need for less costly systems (minisupercomputers) to serve as powerful "front-ends," stand alone systems, and distributed departmental level systems.

Other Important Classes; The Personal Computer and the Applications Workstation

Although not the first entry in the arena, Apple Computer introduced its inexpensive user-friendly personal computer (13) in 1977. It rapidly established a new niche in the marketplace. This resulted from the low cost and quite powerful processing capability of the product. Most importantly, the new systems required very little prior education on the part of the user for their successful application to problems. IBM, in 1981, recognizing the major new market potential, also entered this mass market low-cost

product arena. These two firms are now the major players in this niche, a niche quite removed from all previously identifiable classes of computers.

The personal computer started as an expensive toy. Electronic games were a major use in their early days. With time, more serious applications activities, especially word processing, spreadsheets, and data base procedures, quickly became the primary function for these machines. Now personal computers can be found in many homes. They have been the basis for the transformation of the office into today's "electronic office." The breadth of applications to which these machines have been put has increased enormously. Their range of performance and price now also spans a wide range.

Even more recently, specialized, very powerful classes of personal computers have entered the marketplace. Initially these were "scientific workstations," but more generally "applications workstations." These machines span computing performance levels from the high end of the personal computer well into the arena heretofore covered by the minicomputer and low end mainframes. The driving force for the development of this new class of system is primarily the rapid growth in the performance capability of mass produced microprocessors. Powerful processors on a single silicon chip that are mass produced at remarkably low cost have made it possible to place significant computing capability into a desk-top package. Only a few years ago, systems of similar power required investments in excess of $100,000 and took up a good fraction of a room. Scientific and engineering workstations have blossomed into a major new industry over the last few years. This niche, too, is now well established. Low cost, high performance, personalization of such workstations to an individual's needs, and their dedicated availability to an individual make them a very attractive product.

Historical Analysis

In this section, the evolution of the computer industry is traced through the stages of its development. The bases for this representation of the industry are: the applications

served, the cost of an individual system and the total annual revenues achieved in each market segment. Earlier, particular niches were identified as they developed, and the factors that influenced their development, growth and evolution were discussed. Figure 1 shows snapshots of computer industry development representative of each period.

In these graphs, the "applications" axis is a qualitative one, divided into "scientific" and "business" applications. The "unit cost" axis (14) is the price of the various systems stated in thousands of dollars of the year of the given snapshot. The vertical axis is the "total annual revenues" for each of the market segments depicted, again stated in dollars of the year of the given snapshot. The latter two axes use logarithmic scales to accommodate the range.

Figure 1(a) shows the situation as of 1960. The dollar volume of computer revenue that year was almost $700,000,000. At that time, there was a well-defined bifurcation of scientific and data processing (or commercial) computing, each served by different computer models.

In the ten year period between 1965 and 1975, the minicomputer market segment emerged and began to take on a substantial fraction of the workload. Also during that period, as shown in Figure 1(b) for 1970, the emergence of the general purpose mainframe allowed for the coalescence of the scientific and commercial marketplaces.

During the period 1975 to 1980, with 1977 shown in Figure 1(c) as a representative year, the supercomputer market segment emerged. It responded to a need at the high end of the scientific computing applications space. Also shown is the beginning of the personal computer market segment, quite separated in its early days from the rest of the "serious" computing systems.

During the first half of the eighties, there was continued growth in all market segments; Figure 1(d) depicts the situation for 1982. The scientific workstation emerged spanning

the space between the personal computer and the minicomputer. Finally, the current state of affairs is shown in Figure 1(e); 1987 is shown, although the data here is preliminary. There is continued growth in all computing market segments and the development of the latest niche -- the minisupercomputer -- has started.

It should be noted that at every step along the way, there have been other classes of computers that were developed, but failed to achieve significance as market identifiable segments. Some of these, e.g., attached array processors, are still being marketed. Because their dollar volume has remained limited, however, they do not appear in the graphs. In many cases, uses of array processors have gradually been satisfied by other products.

The current annual revenues of the U.S. computing industry are in excess of 100 billion dollars. The industry continues as one of the fastest growing components of the U.S. economy. Given the entrepreneurial spirit, especially strong in the electronics and computer components of our industrial society, there continues to be new attempts to develop heretofore unrecognized niches. Many of these entrepreneurial efforts fail either because the ideas are unsound or badly implemented, or as frequently happens the timing is not right. Nevertheless, the activity can be expected to continue, evolving in the marketplace, and establishing new niches.

The Missing Element in the Analysis: Software

Making effective use of any hardware system requires appropriate software. In this arena, the level of maturity is well behind that of the hardware. To some extent, this is related to the lower level of maturity of our software experience relative to our hardware experience. The construction of electronic hardware items in our industrial base goes back over half a century. Our schools address hardware in a mature engineering and production sense in the light of this experience. This is not the case with software.

One of the problems is that not only is software much newer in concept, but also the options available in its implementation are much greater than those in the hardware arena. As contrasted to the situation twenty or more years ago when the early computer systems were delivered, the software efforts required to build a system today are much more demanding than are the efforts for the hardware development. Most firms have not adequately integrated this new relationship into their planning and development cycles.

It is the software more than the hardware that the user sees in interactions with computer systems. Sensible developments of software are further hampered by the fact that there are few well agreed upon standards to which the designers must conform. In many cases, to differentiate their products, manufacturers will resist agreeing to proposed industry-wide standards. Some will explicitly avoid using a standard even when it does exist. These are serious problems which are hampering the growth of the computing industry.

It is important to facilitate the maturation of the software discipline to bring it up to the level of the hardware disciplines in use today. Until this is accomplished, investments in computing will not give rise to optimum return. Since supercomputers play an important role in advancing U.S. technology, it is imperative that these problems be solved. Unfortunately, the short-term need of profitability on the part of industrial providers is not always consistent with fostering these developments. Thus, because of the importance of software, there is a need for government to take an active part in fostering research, development, and industrial coordination to bring this about (16).

The Future Spectrum

What projections can be made for the evolution of the computer industry during the next decade? One of the important developing arenas is that of universally available networking. Computer networks are becoming important in every facet of research and industry. The merging of the communications and the computing industries continues slowly and erratically, but inexorably. The growth of very high bandwidth

communications links across the country is driven by the need for the transmission of large quantities of data by increasingly large numbers of customers. These include customers with science, research, commercial and consumer based needs. Especially among the latter are those needs arising from the entertainment industry. We project that the collective demands of these users will ensure the development of a basic affordable capability to accomplish universally available networking.

At the present time, however, these concepts are still quite immature. Although, conceptual standards have been adopted, the situation is still quite chaotic. Networking is passing through a Tower of Babel building period where communications between different entities invariably have major complications. Each manufacturer generally solves his own product line problems reasonably quickly and reasonably well. However, communication between units of disparate manufacture is frequently an unsolved problem. It may very well take to the end of the century to bring about a level of universality of networked computer communications approaching that of our telephone system today.

As communication problems are solved, one can expect a coalescence of some of the niches which have developed heretofore. Certainly, there will continue to be very powerful computing engines, the supercomputers, for solving the most demanding computational problems. Also, there will continue to exist mainframes appropriate to massive data processing and the management of large files whether centralized or distributed. The largest of each of these systems will certainly rely upon parallel processing approaches to attain the required levels of performance. They may or may not utilize vector processing approaches as well. What the architecture might look like is a question that only the successful conclusion of currently active research and product development programs will answer. Whether the level of parallelism is only modest (a few to a few hundred processors) or where the processor count is in the thousands is an open question. Much progress must be made, especially in the software area, before

these questions can be settled. For a long time, however, it is likely that a wide range in the level of parallelism will be explored.

Because of their cost niches, minisupercomputers and also minicomputers are likely to continue to be identifiable market niches. However, the demarcation between the minicomputer and the applications workstation and possibly between the minicomputer and the larger mainframes may very well disappear.

Good networking and good inter-unit communications will make the breadth of options quite wide. Which options are utilized will be determined by: cost, the nature of the problem, the size and type of databases accessed, organizational division lines, and, as always, history.

Especially for the scientific and engineering communities, very powerful desktop graphics will certainly play a major role. For almost all users, graphics (including color) will play an important role. Also, the user friendly tools that have become the norm for the personal computer will continue to be extended and span almost all human/computer interfaces.

Applications software will continue to evolve explosively. The user will become further and further relieved of understanding the working innards of the computer. Which computer in a network is being used will become of less interest to the user; possibly knowable only on explicit demand. Higher level languages will continue to evolve to accomplish these objectives.

Ideally, what a user wants in any discipline is a nonobtrusive desktop device through which requests or inquiries may be made and responses received in as short a time as possible. Ideally, the workstation or the device on the users desk would be inexpensive, generate little noise and heat, take up little space on the desk, and would yield output as required, and make hard copy inobtrusively and quickly. Whether the actual processing

components are in the box, down the hall, or in a central location some distance away (possibly across the country), should be of no consequence to the user.

What will the computer spectrum of 1993 or 2000 look like? It is safe to predict that the trends of the last 30 years will continue. Niches will continue to evolve and revolutions in thought will continue to be ameliorated by the pragmatic realities of the marketplace. As Toeffler warned, dramatic change has itself become a stable process. Nowhere is this more evident - and more satisfying - than in the computer spectrum.

NOTES AND REFERENCES

1. Alan Toeffler, *Future Shock*, (Random House, New York, NY, 1970).

2. The internal struggles within IBM that preceded the decision to commit to the System 360 as a single product line were enormous. The story is well documented, see e.g., Pugh, E.W., "Memories That Shaped an Industry" (MIT Press, Cambridge, MA, 1984); also the articles in Fortune magazine, September 1966, p. 118 and October 1966, p. 140.

3. Amdahl, G.M., Blaauw, G.A., Brooks, F.P. Jr., IBM Journal of Research and Development, 8, 87-101 (1964); Amdahl, G.M., Blaauw, G.A. Brooks, F.P. Jr., Padegs, A., Stevens, W.Y., IBM Systems Journal, 3, 119-261 (1964).

4. Evans, B.O., "System/360: A Retrospective View," *Annals of the History of Computing* (AFIPS Press, Chicago, IL, 1986).

5. A definition of supercomputer(s) as given in *Supercomputing, An Informal Glossary of Terms*, prepared by the Scientific Supercomputer Subcommittee of the Committee on Communications and Information Policy (Institute of Electrical & Electronics Engineers, Inc., New York, NY, 1987): "At any given time, that class of general-purpose computers that are both faster than their commercial competitors and have sufficient central memory to store the problem sets for which they are designed. Computer memory, throughput, computational rates, and other related computer capabilities contribute to performance. Consequently, a quantitative measure of computer power in large-scale scientific processing does not exist and a precise definition of supercomputers is difficult to formulate."

6. Programmed Data Processor -1 Handbook (Digital Equipment Corp., Maynard, MA, 1960).

7. IBM 650 DP Systems Bulletin, "General Information, Console Operations, and Special Devices," G24-5000-0 (IBM Corp., 1958).

8. Riganati, J.P., Schneck, P.B., *Computer,* **17**, 97-113 (1984).

9. Control Data Corporation, *Control Data STAR-100 Computer*, St. Paul, Minnesota, 1970.

10. Cray, S.R. Jr., "Computer Vector Register Processing," *United States Patent,* No. 4,128,880 (1978).

11. Barnes, G.H., Brown, R.M., Kato, M., Kuck, D.J., Slotnick, D.L., Stokes, R.A., *IEEE Transactions on Computers,* **C-17**, 746, (1968); Bouknight, W.J., Denenberg, S.A., McIntyre, D.E., Randal, J.M., Sameh, A.H., Slotnick, D.L., *Proceedings of the IEEE,* **60**, 369, (1972); *ILLIAC IV Systems Characteristics and Programming Manual*, Burroughs Corporation (1972).

12. Hockney, R.W., Jesshope, C.R., *Parallel Computers* (Adam Hilger Ltd, Bristol, 1981).

13. For a recounting of the early personal computer story, see e.g., Rogers, E.M. and Larsen, J.K., "Silicon Valley Fever" (Basic Books, Inc., New York, NY, 1984); Freiberger, P. and Swaine, M., "Fire in the Valley" (Osborne/McGraw-Hill, Berkeley, California, 1984).

14. Although "unit cost" and "annual dollar revenue" may be described in quantitative terms, these numbers are difficult to obtain with precision.
There are questions such as whether or not the cost of peripherals and software are included. Indeed, is a disk - necessary for a system to function - included as part of the system or as a separate peripheral? Also, the demarcations between

17

some of the categories, e.g., minicomputers and mainframes, are likely to be made differently by different manufactures. Thus, the more quantitative axes also must be interpreted as having quite large uncertainties.

15. The primary data for generating this figure is from Dataquest. Additional and corroborating data was obtained from Hambrecht and Quist, and the Gartner Group.

16. A report on this subject is in preparation by the IEEE Scientific Supercomputer Subcommittee.

Acknowledgements: We would like to acknowledge the assistance of Neil Coletti for the graphical presentation, Diana Evans for the manuscript preparation, Linda Bringan for library assistance in the research and Heidi James, the IEEE staff member responsible for innumerable activities that made this paper possible.

FIGURE CAPTIONS

FIGURE 1. The changing profile of computing space, over the last three decades: (a) 1960; (b) 1970; (c) 1977; (d) 1982; (e) 1987. The Annual $ Revenue peak represents total sales worldwide in dollars current for the year listed. See Reference (15) for sources of data. Note that log scales are used for the Unit Cost and Annual $ Revenue axes.