

MINIS AND MAINFRAMES

Microcomputers as we know them may disappear as sophisticated microprocessors reach toward mainframe capability; more powerful supercomputers are introduced

Fifty years from now, historians of computer science may regard the VAX 11/780 as the last true minicomputer. Spiraling capabilities of single-chip VLSI processors and declining memory costs are combining to blur the distinctions between the high end of microprocessor-based computers and the multiboard minis of the past.

"The definition of a mini used to be a computer that fit in a single rack and had 16-bit data paths," said Nick Matelan, president of Flexible Computer Corp. in Dallas, Texas. "DEC invented it in the late 1960s," he noted, referring to Digital Equipment Corp.'s PDP-11. "Today minis are just getting squeezed."

Pushed from the bottom, several commercial middle-range computers moved upward during 1985 in the fastest practical way: by combining several or even tens of very large-scale integrated (VLSI) processors to get raw processing power like that of the traditional mainframe. At the same time, the upper limit on processing power available from the most powerful commercial machines also moved upward with the introduction of the Cray-2 from Cray Research Inc. of Minneapolis, Minn. Capable of executing 64-bit floating-point operations at the rate of more than 1 billion per second, the machine has reaffirmed to many people Cray's strength in state-of-the-art computing.

Multiprocessor designs compete

Often the early commercial stages of a new technology are characterized by a confusing multiplicity of approaches and claims, and the introduction of multiple-processor systems is no exception. Although supercomputer-like performance is theoretically possible with some of the commercial architectures, most currently fall into the range between high-speed, special-purpose machines like array processors and superminicomputers like the VAX 11/780.

All of the machines use either bus-based or nonbus-based architectures. Within the bus-based group are two subcategories: tightly coupled and loosely coupled systems. The tightly coupled systems, sometimes called multiprocessors, have multiple processors and a common, or global, memory. The processors and memory are connected by one or more high-speed buses. Loosely coupled systems, sometimes called multicomputers, have local memories for each processor, although like multiprocessors they sometimes have global memory for shared data.

The main commercial architecture in the nonbus-based category is the hypercube, based on theoretical work done at the University of Michigan in the early 1960s. That research eventually led to the construction of the "cosmic cube" at the California Institute of Technology in 1983. Instead of buses, hypercubes rely on direct-memory-access channels between neighboring processors and their memories. Last year saw the introduction of three commercial hypercubes.

Paul Wallich, Glenn Zorpette Associate Editors



In the bus-based world, one of last year's notable events was the introduction of the long-awaited Multimax computer system from Encore Computer Corp. of Marlborough, Mass. The company made news when it was founded two years ago by three of the top names in the minicomputer world: Kenneth G. Fisher, formerly president of Prime Computer; C. Gordon Bell, a former vice president at the Digital Equipment Corp.; and Henry Burkhardt III, a cofounder of Data General Corp. Encore was back

in the news in October when Burkhardt resigned from the company and attention was focused on the company's difficulties in bringing its technologically ambitious machine to the market on time. The first Multimax, a full 20-processor system, was finally shipped in mid-October to a U.S. government installation.

Encore's machines joins the Balance computer from Sequent Computer Systems Inc. of Beaverton, Ore., in the bus-based multiprocessor category. The Multimax has a 100-megabyte-per-second bus and can be expanded from two to 20 microprocessors with 4 to 32 megabytes of common memory. (Sequent's system, introduced in September 1984, expands from two to 12 processors and up to 28 megabytes of main memory. It has a 26.7 mega-

C. GORDON BELL: EXPERT OPINION



"Processors with pared-down instruction sets are an alternative for high-performance single-processor designs."

The taxonomic tree of today's computers shows exciting new branches based on parallel processing as well as more leaves on all current branches. Three areas in particular showed gains last year: high-speed conventional processors based on gate arrays and other custom chips; multiprocessor superminis—"multis"—based on off-the-shelf microprocessors; and massively parallel computers with all elements operating in synchrony.

Conventional processors based on complex-instruction-set architectures (including floating-point, decimal arithmetic, and character-string data types to support Cobol and Fortran) continue to evolve, becoming more powerful and shrinking in size. A fairly extreme example is Digital Equipment Corp.'s MicroVAX II, a single-chip processor introduced last spring, which provides roughly the same performance as the VAX-11/780 superminicomputers.

Processors with pared-down instruction sets—RISCs—are an alternative for high-performance single-processor designs. They do not require a microprogrammed architecture and are thus simpler to design and build [see "Toward simpler, faster computers," *Spectrum*, August 1985, p. 38]. The most impressive, Fair-

byte-per-second bus.)

Like the processing units in the Sequent machine, Encore's processors are based on National Semiconductor's NS32032 microprocessors, for which Encore's engineers have devised a system that expands the addressing from 24 bits to 32. On each processor card in the system are two processors, which share a 32-kilobyte cache memory based on 45-nanosecond static RAMs. Encore and Sequent are both working with National Semiconductor's second-generation 32-bit microprocessor, the NS32332, and other 32-bit chips, so machines with significantly higher performance could appear as soon as the end of this year.

The Multimax was introduced with an operating system called Umax 4.2, Encore's multiprocessing version of Unix that is compatible with 4.2 BSD (for Berkeley Systems Distribution—one of several versions of the Unix operating system). Umax 4.2, like most parallel operating systems, makes shared memory and interprocessor synchronization available in high-level languages. The company has compilers for C, Fortran 77, and Pascal.

Sequent, meanwhile, strengthened its software line for the Balance systems. Like Encore, the company has a 4.2-BSD-derived operating system and compilers for C, Fortran, and Pascal. In November, Sequent added a parallel debugger. The tool helps programmers troubleshoot parallel programs by running all of the processes in an application on multiple processing units, as in a regular task, but with the clock under their control. This enables programmers to single-step through the application—to see the results of each clock cycle for all of the processors.

Automatic parallel-processing software, which would allow programmers to work with multiple-processor systems as they would with a single processor, is not yet a reality, software engineers say. Programmers have to begin by studying the application at hand for points where the task can be divided among several processors, and they must then decide which of several synchronization mechanisms is best suited to the problem.

Flexible Computer shipped its second and third Flex/32 multi-computers in February and March last year. The system has dual 20-megabyte-per-second common buses and separate 4-megabyte-per-second local buses for each of the processor and memory cards; the aggregate bus bandwidth within a system can be as

high as 120 megabytes per second. Up to 20 processor and memory cards can be inserted on the local buses with access to the common buses. The processor cards have an NS32032 microprocessor, a floating-point arithmetic processor, 1 or 4 megabytes of RAM, up to 128 kilobytes of electrically programmable ROM, and support for the VME bus, an industry standard. The memory card adds up to 8 megabytes of 150-nanosecond dynamic RAMs. Two of the cards share dual 4-megabyte-per-second local buses for communication with the common (20 megabytes-per-second) system buses.

Flexible has a Unix System V-derived operating system and a Unix-compatible real-time operating system called MMOS (for multicomputing, multitasking operating system). High-level languages include C, Fortran, and Ada.

Hypercube hype heats up

Although bus-based architectures can be made extremely powerful by using very high-speed buses, they are always limited to a certain number of processors by the bandwidth of the buses. Hypercubes, on the other hand, avoid this limitation by eliminating buses. With this architecture, each processing unit, called a node, can communicate directly with its nearest neighbors in the n-dimensional space in which it has been designed and built.

For example, a two-dimensional hypercube would have four nodes, each at a corner of a simple square, and each node would be able to communicate directly with two other nodes. In three dimensions, there would be eight nodes, each at one corner of a cube. Each would communicate directly with three other nodes. In four or more dimensions the nodes are not at the corners of any easily perceivable shape, but topological considerations aside, the extrapolations are simple: in a four-dimensional cube (16 nodes), there are four nearest neighbors; in five dimensions (32 nodes), there are five. [See illustration, p. 39.] One drawback with this communications scheme is that if a processor needs to communicate with a node that is not one of its nearest neighbors, the data must be routed via intervening processors; this can slow overall processing rates if it occurs often.

Two of the three companies that introduced commercial

child's Clipper microprocessor chip set, consists of five chips and executes instructions at roughly the same rate as the VAX 8600. The processor being built by Mips Computers Inc. exploits work on simple, fast architectures done at Stanford University. The U.S. Defense Advanced Research Projects Agency has let a contract for a gallium arsenide version intended to reach 100 million instructions per second (MIPS).

A new class of high-performance small computers is the multi-computer—a shared-memory multiprocessor with a single bus for interconnection and cache memory for each processor to reduce delays caused by bus traffic. Computers of this type were introduced by Elite, Encore, and Sequent, with up to 6, 20, and 12 processors respectively. They offer performance up to 20 million instructions per second. These machines, which are based on off-the-shelf microprocessors, generally assign one or more tasks to each processor, although processors can work together on a single task under specialized conditions.

At the top end, uniprocessors have given way to small-scale parallelism: the Cray XMP and Cray-2 supercomputers each contain four processors. They use multitasking to improve performance by applying both coarse- and fine-grain parallelism. In coarse-grain parallelism, different segments of a job are divided between processors. In fine-grain parallelism, elements of a single operation, like matrix manipulation, are performed concurrently.

An important development last year was the appearance of very powerful superminicomputers, dubbed minisupercomputers. They implemented supercomputer architectures with cheaper, slower components, offering one-third to one-half the performance of a top-end machine for one-tenth the price. Convex, for example, has a 4-MIPS, 50-megaflop computer priced in the same range as a

supermini—about half a million dollars.

Larger-scale parallelism is exemplified by the Connection Machine, from Thinking Machines Corp. of Cambridge, Mass. It operates at a peak of 1 billion operations per second on 32-bit data. It uses a single instruction to control all the processing elements in lockstep fashion, although some processors can disable themselves and rejoin the instruction stream at a later time. Each of the 64 000 1-bit processing elements has 4096 bits of memory, and information can be transferred serially to any other element in the array either through a global routing mechanism or by local connections. IBM's GF11 is an 11-gigaflop special-purpose supercomputer designed for solving problems in quantum chromodynamics. It contains 576 20-megaflop processing elements interconnected by a three-stage switching network.

Smaller-scale special-purpose multicomputers include the Butterfly, built by Bolt, Beranek, & Newman Inc., of Cambridge, Mass., and the iPSC, built by Intel Corp. in Beaverton, Ore. The Butterfly is a collection of up to 128 microprocessors—each with its own memory—that communicate through a high-speed packet-switching network. Intel's iPSC contains 32, 64, or 128 processors connected in a hypercube by 10-megabit-per-second links. The Intel machine has been delivered to research groups that hope to use it as a relatively inexpensive vehicle for studying parallel algorithms and achieving near-supercomputer speeds.

C. Gordon Bell (F) is vice chairman for technology at Encore Computer Corp. Prior to joining Encore he was vice president for engineering at Digital Equipment Corp., where he was responsible for the architecture of the VAX superminicomputer.

hypercube systems last year are using processing nodes based on the Intel 80286 microprocessor and its floating-point chip, the 80287. One of the companies is Intel's own Scientific Computers Division in Beaverton, Ore., which unveiled its iPSC (for Intel parallel supercomputer) in February 1985. The computer is expandable from a 32-node, five-dimension machine to 128 nodes—seven dimensions. Besides the 80286 and 80287 processors, the nodes of the Intel system have 512 K of RAM and 64 K of ROM, as well as controller chips for the Ethernet that is used for internode communications.

Another hypercube using nodes based on the 80286-80287 combination is the System 14 from the Computer Research Division of Ametek Inc. of Arcadia, Calif. This machine is expandable from four dimensions (16 nodes) to eight (256 nodes). The nodes have an 80186 microprocessor for communications with other nodes, and 1 megabyte of RAM, which is accessible to all three processors in the node.

The most ambitious of the three commercial machines is the Ncube/ten, announced last November by Ncube of Beaverton, Ore. The company was founded in the summer of 1983 by several former Intel employees, including John Palmer, the architect of the 8087, and William Richardson and Stephen Colley, who worked on the iAPX 432.

The Ncube/ten derives its name from the fact that it can be expanded to 10 dimensions—1024 nodes. In its most powerful configuration it should be able to achieve a processing rate of about 500 million floating-point operations per second for an efficient program, according to Palmer.

To enhance the reliability of a machine with so many nodes, the company decided to minimize the number of chips required

for each node by integrating as many functions as possible on a single, custom-designed processor chip. The result was a 2.5-micron, 160 000-transistor, 10-Megahertz NMOS chip that integrates a 32-bit general-purpose processor, a 64-bit floating-point processor, a memory interface unit with error correction, and a communications interface based on 22 independent direct-memory access channels—in short, all of the hardware needed to build a node except memory. With six 256-K dynamic RAM chips and the single processor-communications chip, the Ncube node has far fewer than the approximately 50 chips used in the nodes of the Intel and Ametek machines.

Relatively little software has been written for the hypercubes so far. All three have Unix-compatible or Unix-like operating systems and high-level languages like C and Fortran. Some have simulation tools that permit programmers to trace the execution of programs on a single-processor, front-end system to make sure that it runs correctly before dividing it up among the nodes of the hypercube. Still, given the complexity of designing parallel algorithms and matching them to an optimum topology within the machine, programming hypercube-based computers is no picnic.

"Precious few people know how to use machines of this sort," admitted an executive of one of the hypercube makers. However, this situation could change as many, if not most, of the first hypercubes go into universities. The pool of trained users could expand, and more sophisticated software tools may flow from university work.

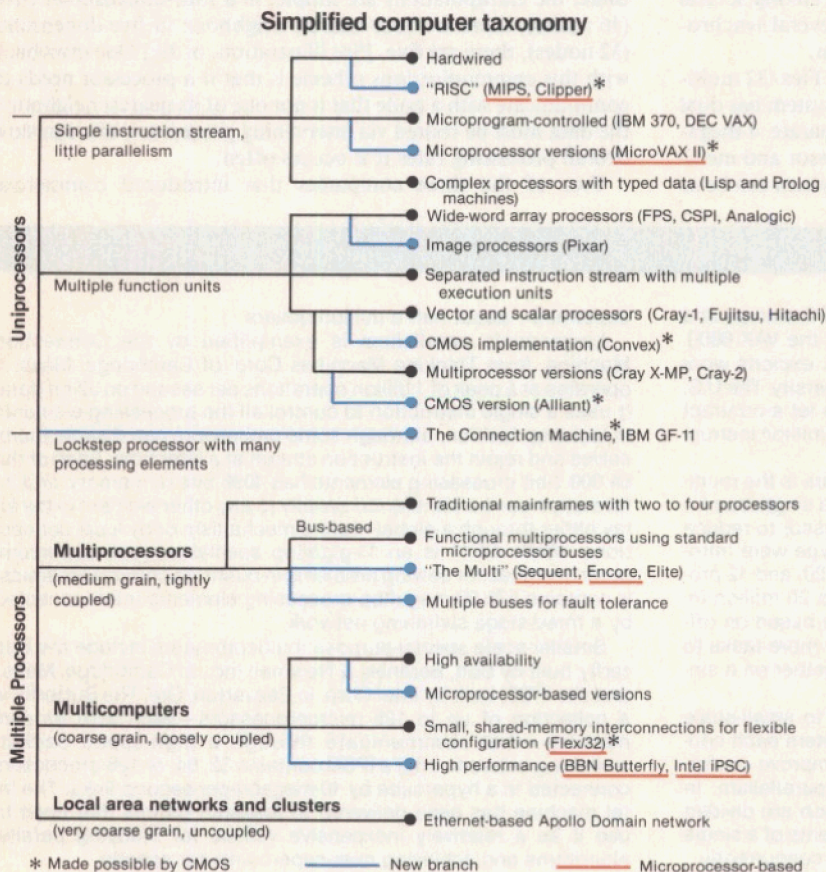
The other important continuing trend in the design of middle-range computers is the streamlining of instruction sets. This began several years ago with the introduction of VAX-class machines by Ridge Computers Inc., Pyramid Technology Corp.,

and others. This year, both Ridge and Pyramid expanded their lines with more powerful versions of their original systems. Pyramid, for example, introduced single- and dual-processor systems with 100-nanosecond cycle times, down from 125 on its previous machines. Harris Corp. of Melbourne, Fla., introduced its HX-7, a RISC-like computer optimized for the Unix operating system. Implementations of the single-chip RISC (reduced-instruction-set-computer) also appeared in 1985 [see "Microprocessors," p. 46]. Among the major manufacturers, Digital Equipment and IBM Corp. are rumored to have RISC projects underway, and Hewlett-Packard has publicly announced its commitment to RISC with the Spectrum minicomputer project, but has refused to release any details. [For more information on processors with reduced instruction sets, see "Toward simpler, faster computers," *Spectrum*, August 1985, p. 38.]

Large machines also advance

In addition to powerful processors based on collections of many relatively conventional single-chip processors, there were advances last year in supercomputers and near-supercomputers based on relatively small numbers of very powerful processors. The prime example is the Cray-2, built by Cray Research Inc., with four processors, a 4.1-nanosecond cycle time, and a peak calculation rate of well over 1 billion floating-point operations per second. The Minneapolis, Minn., company delivered its first fully configured Cray-2 to the National Aeronautics and Space Administration's Ames Research Center, Moffett Field, Calif., last summer, for use in the agency's Numerical Aerodynamic Simulation (NAS) program.

The Cray-2 is remarkable not only for its speed but also for its memory capacity: 256



All the computer architectures can be viewed as variations on a few themes: single processor versus multiple processors or single-chip versus multiple-chip implementations, for instance. These variations fan out in a family tree of architectural types, showing where a particular kind of computer may have come from, and where it may be going.

million 64-bit words, or 2 gigabytes. In the past, the high cost and low density of high-speed memory limited supercomputers to a few million words of main memory. The total number of bits per chip and the number of chips that could be squeezed into a small volume without overheating limited the total memory available. To achieve high clock speeds, the wires in a supercomputer must be short so signals are not unduly delayed in transit. A large memory array would previously have required wires too long for high-speed operation.

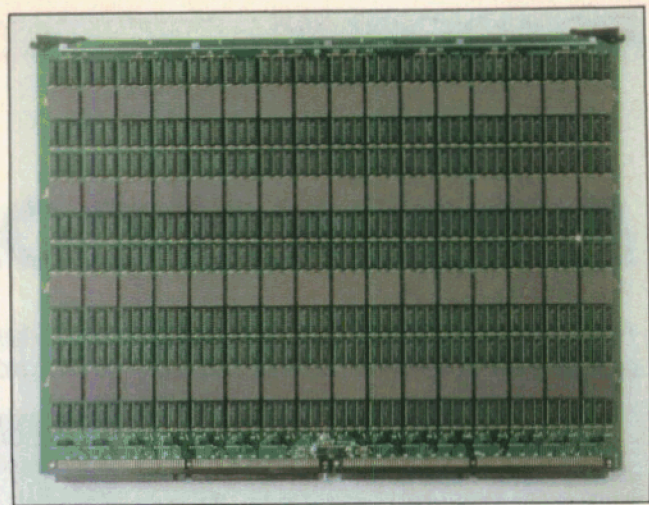
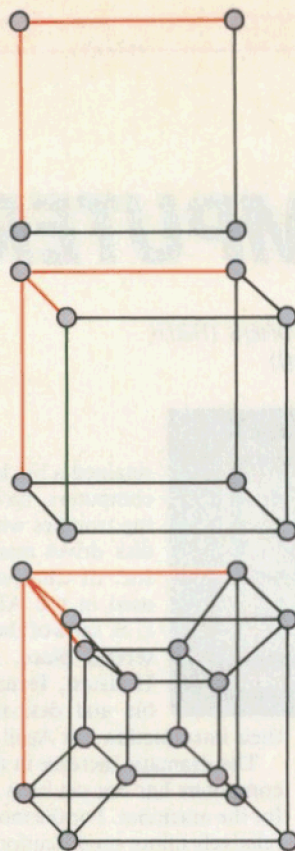
Electronics packaging is perhaps equally as important to the performance of the Cray as electronic design; each of its more than 300 4-by-8-by-1-inch circuit modules consists of eight circuit boards with vertical as well as horizontal connections. The modules dissipate between 300 and 500 watts each, for a total power dissipation of more than 150 kilowatts in a volume of approximately 1 cubic meter—roughly equivalent to an array of 2000 tightly packed 75-watt light bulbs. A liquid fluorocarbon coolant that is pumped through the circuit modules keeps the operating temperatures lower than 25°C.

The architecture of the Cray-2 is much like that of earlier Cray computers, with the exception of a local memory in place of auxiliary registers attached to each processor. Previously the eight scalar registers and eight address registers of a Cray central processing unit had a backup set of 64 registers each to hold data for temporary storage rather than transfer it all the way back out to main memory. Each Cray-2 processor has 16 kilowords of local memory, accessible to the vector registers as well as to the scalar and address registers, for temporary storage. Main memory is organized in four quadrants of 32 banks each, and each processor has access to one quadrant during each clock cycle.

A number of "minisupercomputer" manufacturers have chosen to mimic the Cray's architecture, but they use slower, lower-power components to avoid the packaging problems. Convex Computer Corp. of Richardson, Texas, for example, in its C-1 computer, uses a vector execution unit similar to that in the Cray-1. But the C-1 uses off-the-shelf TTL parts and CMOS gate arrays, instead of emitter-coupled logic, thus allowing it to fit in a standard 19-inch rack with fans for cooling. Of course, the Convex performs at between 4 and 5 percent of the speed of the Cray-2, but its \$500 000 price is only about 3 percent of the Cray's.

Other minisupercomputer makers are taking new architectural approaches—Culler Scientific Corp., of Santa Barbara, Calif., for example, uses multiple execution units within a single central processor and a 96-bit-wide instruction word to schedule several operations in parallel, achieving a maximum of a dozen operations in a single clock cycle and thus 7 million instructions per second (MIPS) and 2 megaflops with a single processor. For example, integer arithmetic and floating-point arithmetic are performed by two different units, and additional hardware fetches data from memory. A four-processor version of the machine is said to achieve 36 MIPS and 16 megaflops.

Alliant Computer Systems Corp. of Acton, Mass., has developed a multiprocessor unit, the FX/8, capable of delivering ap-



Ncube

Hypercubes are multiple-processor computers that have no bus. The processor nodes (the spheres in the diagram at left) have dedicated communications channels to their "nearest neighbors" in the n-dimensional space in which they were designed. There are two nearest neighbors for each node in a two-dimensional hypercube (left, top); three for a three-dimensional system (middle); and four for a 4-D machine (bottom). One of the three companies selling hypercubes, Ncube of Beaverton, Ore., uses state-of-the-art VLSI technology to cram a 6-D hypercube—64 nodes—onto a single 16-by-22-inch eight-layer printed-circuit board (above). Each of the 64 nodes has a custom-designed processor chip with built-in floating-point and communications logic, and six 256-K dynamic RAMs.

proximately 1 million floating-point operations per second from each processor up to a maximum of eight. It uses advanced compiler technology to detect concurrency within loops and also to detect data dependencies—information that must be present for a computation to continue—so that it can schedule multiprocessor execution to minimize bottlenecks.

While there are a number of new entrants in the high-performance computing arena, some old ones have dropped out as well. Denelcor, for example, whose HEP-1 (heterogeneous-element processor) computer embodied an innovative parallel architecture, found itself unable to sell enough machines to cover operating expenses, or to attract funds for development of a second-generation version.

Mainframe manufacturer IBM has also gotten into the numeric-processing business, introducing a vector-processing unit for its recently announced 3090 mainframe series. The vector unit plugs directly into the 3090, effectively adding instructions to its instruction set rather than acting as a separate processor. It executes one 64-bit floating-point multiplication and one addition in 18.5 nanoseconds for a theoretical maximum rate of 108 million floating-point operations per second.

The actual performance of the vector unit—about 12 megaflops on the commonly used Linpack benchmarks—is substantially lower because of the overhead incurred in setting up vector calculations and fetching instructions and operands from memory, according to Jack Dongarra, a computer scientist at Argonne National Laboratory in Illinois, who consulted for IBM and had access to early models of the processor. One reason for this performance drop, Dongarra said, is that all memory references pass through the 3090's relatively small cache memory, causing delays whenever data is not in the cache and must be fetched from main memory. Since vector calculations generally operate on very large amounts of data, cache misses are frequent. Machine-language routines that reformulate their algorithms to take the cache into account reach speeds of 65 to 80 megaflops, Dongarra noted. ◆