used for large programs, a balanced system could require at least 4 megabytes of memory per million instructions per second.
16. Digital Equipment Corporation's VAX-11/780 Project began in April 1975 and the first delivery was in March 1978.
17. T. Kidder, *Soul of a New Machine* (Little, Brown, Boston, 1981).
18. The simplest way to ensure cache memories do not contain stale data is for processors to write through the cache memory each time a processor issues a write command. All cache memories monitor the bus for data being written back to memory. If the cache memory contains a copy of data that has been modified, it deletes the stale data. This simple write-through scheme requires much higher bus bandwidth, however; and the system's performance is limited by the ability of a cache memory to monitor the write-through commands from all the other cache memories.
19. In a modern operating system a user may initiate and control several jobs. A single job is often composed of a set of independent processes that, in turn, may each be composed of several tasks. Whereas jobs, processes, and tasks are all candidates for parallel execution, I will use the term job to represent generic computational activity.
20. G. C. Fox, in *IEEE Computer Society Proceedings, Compcon 84* (IEEE Computer Society, New York, 1984), pp. 70–73.
21. A. K. Jones and E. Gehringer, *The Cm* Multiprocessor Project: A Research Review* (Carnegie-Mellon University, Pittsburgh, 1980).
22. D. Vrsalovic, D. P. Siewiorek, Z. Segall, E. Gehringer, *Performance Prediction and Calibration for a Class of Multiprocessor Systems* (Carnegie-Mellon University, Pittsburgh, 1984).
23. E. A. Torrero, Ed., *Spectrum* **20** (November 1983).
24. I thank H. Burkhardt III, D. Schanin, R. Moore and S. Frank for their contributions to this article and D. P. Siewiorek and G. Bell for their assistance in preparing the article.

Workstations in Science

William Joy and John Gage

Summary. Recent advances in microprocessor technology are making a new generation of personal computers feasible and affordable. These computers, called "workstations," can be connected with other workstations, with mainframe computers, with supercomputers, and with remote networks. Workstations provide the graphical interface to supercomputers and can run applications that formerly required the use of mainframe computers. Emerging standards for communications, graphics, databases, numerical algorithms, and languages allow workstations to share programs with various types of computers, including mainframe and personal computers.

In the past several years, a class of computers has emerged that provides the scientific and engineering user with inexpensive personal computing power comparable to that of commercially available superminicomputers such as the VAX-11/780. These machines have a sophisticated graphical user interface, combined with the capacity to utilize specialized computing services provided over a high-speed network.

Workstations are microprocessor-based desktop computers with the hardware architecture of superminicomputers, combining virtual memory with sophisticated operating systems, communications protocols, advanced languages, and high-resolution graphics displays. They are able to run applications from larger machines directly and interactively.

Furthermore, workstations are designed to be connected with other workstations, with specialized processors, with mainframes, and with supercomputers, as well as with large mass storage devices, specialized input and output devices, graphic input and output devices, and remote networks. Workstation users belong to a community, linked by electronic mail, conferencing and communication tools.

Workstations range in cost from $9,000 to $50,000 dollars but are expected to decrease in cost to half that in the next 2 years. The IBM PC/AT, the Apple Macintosh, and other smaller machines perform many of the functions of workstations, although they lack virtual memory and other characteristics of mainframe machines that workstations provide. Both the workstations and the smaller machines will increase in capability in the next few years. According to figures obtained from the larger vendors, including Apollo Computers, Sun Microsystems, Symbolics, Masscomp, Xerox, and Tektronix, more than 30,000 workstations have been purchased in the last 4 years, most of them for scientific and engineering uses.

Workstations in a Scientific Environment

As part of their day-to-day research, Eric Arens and Michael Peck at the Space Sciences Laboratory of the University of California, Berkeley, use an inexpensive desktop workstation to perform 48-megabit-per-second data acquisition, instrument control, signal processing, statistical analysis, digital image enhancement, database storage and retrieval, and color image animation. On the same system, Peck designs and lays out printed circuit boards to control hardware used in his research, writes signal-processing algorithms, typesets research papers, and corresponds with dozens of colleagues over international electronic mail networks. One window on the screen is logged in to the Cray supercomputer at Lawrence Livermore National Laboratory for large data analyses. When their scheduled observation times at Kitt Peak National Observatory occur, Arens and Peck disconnect the workstation from the university's network, move it to the floor of the telescope, and use it to control the telescope, the image acquisition system, and the image display system. Before workstations, these applications would have been performed on several separate superminicomputers, none of them easily transportable.

Development of Workstations

In the 1970's, integrated circuit technology began to advance rapidly. The results of these advances—inexpensive microchip computer processors and inexpensive semiconductor memory—made personal computers and personal workstations possible. These technologies were used at a number of different sites to create prototypes of the workstations, personal computers, and computer networks that are available today.

The first workstation, the Alto, was

William Joy is Vice President for Research and Development, and John Gage is Science Officer of Sun Microsystems, Inc., Mountain View, California 94043.

built at the Xerox Palo Alto Research Center (PARC) in 1973 (*1*). The Alto supported high-performance graphics on a full page bit-map display. It had a mouse, a removable disk, and an Ethernet local-area communications interface. By means of the Ethernet, the Altos shared resources, which included the first laser printers and remote file servers. Although the Alto's 16-bit address space was not adequate for running the large applications common on the mainframe machines of the time, it had adequate performance to develop the bit-mapped and windowed user interfaces now available on most workstations and personal computers.

For many years the work done at Xerox was cloaked in secrecy. Some of the world's best computer scientists worked on projects such as Smalltalk, Mesa, Cedar, Courier, Grapevine, Cypress, Interlisp-D, Pilot, Ethernet, Xerox Network Services, and others (*2–6*) but, except for a small community of computer scientists, the scientific world in general was largely uninformed about the programming languages, graphic user interfaces, and systems developed at Xerox PARC.

The Alto was never a commercial product, but a follow-on product was introduced by Xerox in 1981, the Star 8010, a multiwindowed workstation with multiple fonts and an editor that typeset to the screen.

University Contributions

In the late 1970's, a group at Carnegie-Mellon University proposed the development of a scientific computing system based on workstation technology. The Spice proposal (*7*) set goals largely implemented by today's available machines. Much of their development work emerged on the PERQ workstation, manufactured by Three Rivers Computers. At the same time, a group of researchers at Massachusetts Institute of Technology were developing the Nu machine, which had the same general features as the Alto. The design for this machine was licensed to several companies, including Zenith, Western Digital, and most recently Texas Instruments. In 1979, at Stanford University, the computer science department installed a number of the Xerox Alto personal computers, a Xerox laser printer, and an Ethernet linking them to a file server, to existing VAX computers, and to other large computers. To augment this network, in 1981 a group of researchers built a workstation that made use of commercially available microprocessor technology (*8*). This workstation was subsequently licensed to Sun Microsystems.

While the Carnegie, MIT, and Stanford groups were creating computing environments based on specialized hardware, a group at the University of California, Berkeley, was enhancing an existing and "portable" software operating system, UNIX, to meet research computing needs; UNIX is portable in that it can be used with various commercially available machines. UNIX System III (today, System V) had no provision for networking or virtual memory, fundamental requirements for demanding scientific problems. The Berkeley group, with the support of the Defense Advanced Research Project Agency (DARPA), added capability for virtual memory, communications, and networking, but provided no graphic user interface. Using these enhancements, the Berkeley group built a distributed network of small superminicomputers (VAX 11/750's) designed for single users. The resultant UNIX system, called 4.2BSD, became a standard for scientific computing on several thousand machines (*9*).

Under DARPA sponsorship, long-distance networks that linked computing resources at major university research centers and the national laboratories were developed. The construction of large networks at the national laboratories led to new system designs to provide all users with uniform access to all system resources.

Commercialization

In 1980, a group within Prime Computer attempted to persuade management that workstation technology was worth investing in. When their project failed to be funded, they began their own company. This company developed the Apollo Domain workstation, the first commercial workstation to offer diskless virtual memory (*10*). Within a few years, a number of other vendors, including Digital Equipment Corporation, Silicon Graphics, Tektronix, Apollo, and Sun Microsystems, offered workstations based on the 4.2BSD UNIX system.

Technological Advances

Semiconductor technology, which made workstations possible, continues to advance rapidly. The fastest available microprocessors are increasing in speed by an average of 50 percent per year. Main memory components, which store programs and data while the processor is executing programs, have been increasing in density by an average of 60 percent per year (*11*). Today, an inexpensive personal computer like the IBM PC or Apple Macintosh, which can be purchased for a few thousand dollars, has the memory capacity and processor power of the fastest minicomputers available in 1975, which then cost over $50,000. The VAX 11/780 minicomputer, which has been a staple in the scientific computing environment for many years, is only about 50 percent faster than today's single-chip microprocessors. Not all technologies are advancing at the same rate, however. The technology used to build large-scale machines, with bit-slice components and high-speed emitter-coupled logic, are getting faster each year, but only by roughly 20 percent. Thus, these more expensive technologies will be outpaced by microprocessors for many applications.

Microprocessors have caught up with the bit-slice technology machines in performance, at a much lower cost, and at about the same speed as a lightly loaded superminicomputer. But the next generation of microprocessors should far outstrip the current ones. With the introduction of the Motorola 68020 processor, as well as chips from Intel and National, workstations are expected to run three times as fast as last generation's superminicomputers. Within a few years, this will probably increase by another factor of 3 with the use of the reduced instruction-set computer (RISC) (*12*).

Current Workstations

Workstations today provide three major functions. They run a wide range of applications software; they provide a graphical and multicontext user interface for this software; and they communicate at high bandwidth with the other machines that are components of the system. The synergy of these three capabilities characterizes workstations and makes them important in the emerging computing environment.

Applications Software

Workstations run the operating systems and software of the superminicomputer world, which, in turn, includes software from the mainframe world. This software is usually written in a high-level programming language. Such operating systems allow a multitasking; that is, the user can run several programs concurrently. UNIX, the first operating system to run on computers ranging in capability from the IBM PC to the Cray, is well suited to the workstation environment. It has been running for nearly a decade on minicomputers and superminicomputers, and allows the machine to be shared efficiently among different concurrent activities.

Workstations provide number-crunching capability necessary for much scientific computation. As the IEEE floating point arithmetic standard is implemented on chips, workstations will incorporate them, thus gaining in accuracy and correctness. (IEEE is the acronym for the Institute of Electrical and Electronics Engineers.) Workstations are rapidly increasing in their ability to run large scientific codes and can be effectively improved with small-array processors. Hundreds of demanding scientific and engineering applications recently available only on superminicomputers or mainframes now run on workstations. Seismic analysis, chromosomal sequence analysis, spectroscopy, nuclear magnetic resonance analysis, crystallographic analysis, three-dimensional molecular analysis, finite element analysis, image processing, exploration of properties of cellular automata, computer algebra, expert system development—all are being performed on desktop workstations. Workstations, however, are not enough to run the most demanding scientific applications. The computation-intensive portions of a program are run over the network on specialized large processors and database machines. The display and analysis portions of the program run on the workstation.

User Interface

The ability of workstations to run multitasking operating systems has encouraged engineers to develop user interfaces that exploit multiple processing. A workstation display allows the user to interact with several tasks simultaneously. For example, someone may read a piece of mail from a colleague in one window while examining a dynamic graphical display of related data in another window.

As a more interesting example, in a current application at the DARPA Center for Seismic Studies, one window displays a portion of a wave form so that parameters may be measured, while another window lists the parameters that remain to be extracted. Several other windows allow the analyst to see a graphic portrayal of the position in time of the data with relation to eight other streams of data, to control the transmittal of the extracted parameters to other investigators over the World Meteorological Network, to examine a large database for events with similar parameters, to monitor other wave forms arriving over the satellite links from the seismometers, and perform other tasks. Each window presents an active portrayal of a different aspect of the task, allowing the analyst to choose which is most meaningful or useful at the time.

To make best use of the normal million-pixel screens, workstations are beginning to develop the speed for kinematics. By storing multiple images in the display memory of a workstation, the output of a simulation run that produced these images can be viewed on the workstation screen. Making movies in this way can show dynamic effect without the expense and delay of developing film. For high-resolution images, the data involved is enormous. A high-resolution frame occupies 1 million bytes of memory if it stores 8 bits per pixel. At ten frames per second, a 10-second sequence would involve 100 million bytes of information. Since mass-produced disk drives transfer a maximum of 1 million bytes per second, the ability to do this kind of animation on every desktop is beyond the capabilities of current mass-produced hardware. However, the memory and bandwidth required is quadratic in the resolution, so displaying four frames a second at a resolution of 512 lines and columns can be done today.

Workstation Communications

Workstations have high-bandwidth connections to other machines via local-area networks. The most commonly used network is Ethernet, which runs at 10 million bits per second, and can send data over several kilometers. Because low-cost Ethernet interfaces are available for personal computers, workstations, minicomputers, and mainframes, it is a useful base-level network for connecting the components of a computer environment. Local area networks make it possible to share resources as well as information. Workstations linked by high-bandwidth networks have access to shared disk storage on file servers, machines used exclusively for the storage of files. Thus, workstations may use large files without local disk storage, and this reduces system cost. The connection to local area networks and then to long-haul networks also allows mail to be sent electronically to colleagues and thus makes long-distance collaboration and rapid exchange of information possible.

In addition, high-bandwidth links to supercomputer centers allow data and programs developed on local workstations or other local resources to be shared with programs and data on the supercomputers. The results of computations made on the central machines can be brought back to the workstations for local analysis and interactive display. The interconnection of these machines can lead to a worldwide network similar to the one that now links thousands of UNIX installations in Europe, North America, the Far East, and Australia. Effective use of satellite-redistribution methods for such material can keep the cost of connecting to a network down, and the entry cost low.

Mainframes and Microcomputers

Until recently, scientific computing took place almost exclusively on large mainframe computers in centralized facilities; the resources of a computer center were available only to those associated with the institution that paid for the high costs of maintenance and administration (*13*). With the advent of inexpensive microcomputers, some work previously accomplished on the mainframes moved to the desktop, the laboratory, or the home. For many researchers, the availability of word processing, spread sheets, elementary graphics, and terminal emulation on a local, stand-alone personal computer costing less than $3000 represented freedom from the sometimes restrictive environment of the computer center.

For the user deciding how to allocate his work load, a major advantage of local computing power is predictable response time. Typically, mainframe users developed work patterns designed to avoid the midmorning to midafternoon "brown-out," when machine resources were so slowed by being shared among many users that any attempt at interactive development or analysis resembled batch-mode computing. Batch-mode work made sense for large-scale simulations, but not, for example, for word processing. The availability of inexpensive single-purpose processors greatly improved interactive response and therefore productivity. Small machines proliferated. Many scientists attempted to move serious computation to their convenient personal computers; some succeeded, most did not.

The personal computer lacked large, powerful programs, communications capability, well-tested numerical algorithms, standard languages or graphics capability, a known community of users, public protocols, sophisticated operating systems, and support services—the elements taken for granted in the more sophisticated world of the mainframe computer center. In short, two worlds developed, each unable to utilize the strengths of the other. The scientific user needs both.

Supercomputers and Workstations

The increased availability of supercomputer power can extend the scientist's ability to simulate, model, and produce images of complex phenomena in new domains. However, the architectural specialization of supercomputers makes them difficult to use for the design of application programs. Machine cycles are used most efficiently in providing high computational rates, not in providing quick interactive response. For the most effective use of these resources, most of the user interface can be accomplished with workstations.

With the rapid increases in performance that are to come in the next few years, the range of applications that can be run on the desktop will increase. Workstations are powerful enough today to run most applications. They are an important component of the next-generation computing environment, which will consist of a network of personal workstations linked with servers for file storage, printing, and other services, and, in addition, large supercomputer processing nodes (*14, 15*).

In a vision of the new scientific environment, each scientist has a powerful personal machine, can run interactive graphic tasks locally, can easily move applications through a hierarchy of more powerful computers, can exchange information with other researchers, can replace his machine with better ones as they become available, and can run other scientists' software. This environment is made possible by the research work on local-area and long-haul networks, on portable software, and by the emergence of important communications standards such as TCP/IP and Ethernet and portable operating systems such as UNIX.

References and Notes

1. C. P. Thacker, E. M. McCreight, B. W. Lampson, R. F. Sproull, D. R. Boggs, "Alto: A personal computer," in *Computer Structures: Readings and Examples*, C. G. Bell and A. Newell, Eds. (McGraw-Hill, New York, 1979).
2. A. Goldberg and D. Robson, *Smalltalk-80: The Language and Its Implementation* (Addison-Wesley, Reading, Mass., 1983).
3. T. Linden and E. Harslem, "Office Systems Technology: A Look into the World of the Xerox 8000 Series Products: Workstations, Services, Ethernet, and Software Development" (Xerox Office Systems Division, November 1982).
4. D. R. Boggs, J. F. Shoch, E. A. Taft, R. M. Metcalfe, R. Pup, "An Internet Architecture" (Xerox Palo Alto Research Center, July 1979).
5. R. G. G. Cattell, "Design and Implementation of a Relationship-Entity-Datum Data Model" (Xerox Palo Alto Research Center, May 1983).
6. W. Teitelman, "A Display Oriented Programmer's Assistant" (Xerox Palo Alto Research Center, March 1977).
7. A. Newell, S. E. Fahlman, R. F. Sproull, H. D. Wactlar, "Proposal for a Joint Effort in Personal Scientific Computing" (Carnegie-Mellon University, Pittsburgh, 23 August 1979).
8. A. Bechtolsheim, F. Baskett, V. Pratt, "The SUN Workstation Architecture" (Technical Report Number 229, Stanford Computer Science Department, Stanford, Calif., March 1982).
9. B. Joy, "Berkeley 4.2 gives UNIX operating systems network support," *Electronics* **56**, 113 (1983).
10. D. L. Nelson and P. J. Leach, "The Evolution of the Apollo Domain" (Hawaii International Conference on Systems Science, Honolulu, 1984).
11. C. G. Bell, *IEEE Computer* **17**, 14 (October 1984).
12. D. A. Patterson, "Reduced instruction set computers," *Commun. ACM* **28**, 8 (1985).
13. B. L. Buzbee and D. H. Sharp, *Science* **227**, 591 (1985).
14. W. Joy, "The UNIX system in the laboratory," *UNIX/World* **1** (No. 4), 32 (1984).
15. J. P. Riganati and P. B. Schneck, *IEEE Computer* **17**, 97 (October 1984).