

The Ever-Present Packaging Challenge

Lloyd M. Thorndyke and John P. Riganati

One supercomputer characteristic that has remained constant with time is the packaging challenge. In the 1960s, the Control Data 6600 used three-dimensional (3D) discrete transistor and resistor logic modules with wiring tuned by lengths and Freon cooling, which permitted a reduction in the physical size of the computer and a significant performance improvement. Its successor, the CDC7600, continued that evolution with even greater packing density by using extremely small transistors and resistors packaged in larger 3D logic modules. Cray Research used early integrated circuits in much smaller packages and produced the Cray 1a with impressive performance and small size.

Those early, complex high-performance computers shared the characteristic of innovative packaging and cooling, allowing them to operate at millions of floating-point operations (flops) per second. Although these systems realized some performance gain from reduced interconnect distance, most of the speed improvements came from ever faster logic and memory. However, as the rate of increase in logic speed gradually slowed, computer architects were forced to reduce the interconnect distances to gain further performance.

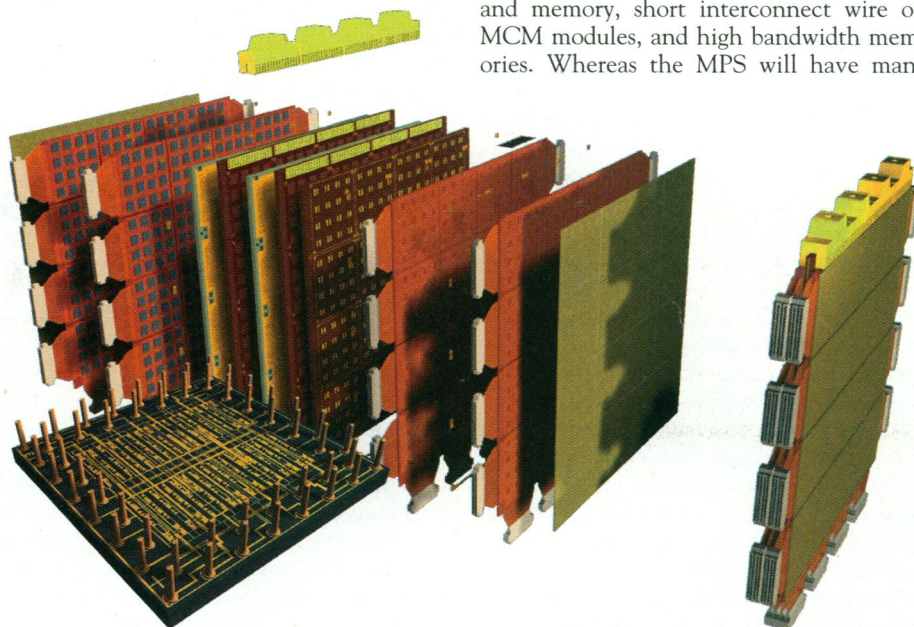
In seeking more performance, the computer architect looks at several factors: (i) faster logic gates, (ii) greater logic density as measured by gates per chip, and (iii) denser packaging to reduce the interconnect delays. This pressing need started the trend that will lead to a few large-scale integration (LSI) custom "bare dies" (that is, unpackaged circuits) per processor accompanied by bare-die memory. Logic and memory will be included in a multichip module (MCM) to achieve a compact high-performance superprocessor. Multiple MCMs can be packaged into a physically small parallel system that has impressive performance and is scalable to the needs of the marketplace. Clearly, this level of integration will be a serious contender in the supercomputer design race. However, this compact 3D package has significant challenges remaining to be solved. Problems—such as

power distribution, ground bus noise, removal of heat, interconnect reliability and impedance control, and ready parts replacement without service interruption—place new demands on the packaging engineer.

Today we are seeing the emergence of the Massively Parallel Systems (MPSs) that use standard packaged logic and memory chips. The typical MPS, with a few thousand standard microprocessor chips and tens of thou-

ing will result in bare microprocessors and memory die, drastically reducing the processor-to-memory distance and increasing system performance. This packaging will gradually evolve into a 3D structure that closely resembles some current supercomputers. When this occurs, the costs of a massively parallel engineering and manufacturing development will approach that of supercomputers—they both will be expensive. However, the production costs for the MCM-based systems will be lower because of the automated MCM assembly tooling and testing and the automated checkout of the system.

This projected trend will move the MPS and the traditional supercomputer into the same technology path in which both will have custom packaging of bare die for logic and memory, short interconnect wire on MCM modules, and high bandwidth memories. Whereas the MPS will have many



A tidy package. Exploded view of the gallium arsenide module designed for the Cray-3 supercomputer. Each module is a 7-mm-thick sandwich of multilayer circuit boards that are 121 mm by 107 mm. The module contains four layers of 16, 25-mm square, multilayer printed circuit boards, these in turn holding 16 gallium arsenide integrated circuits or 12 silicon memory circuits. Five larger multilayer circuit boards make up a plate assembly at the center of the module for distribution of power and logic signals to the smaller circuit boards. The assembled package is shown at right. Also shown in lower left is the Cray-3 gallium arsenide unpackaged chip, which measures 3.8 mm square by 0.2 mm thick. [Courtesy Tom Sibert, Cray Computer Corporation]

sands of memory chips, has grown to large physical proportions and looks like the supercomputers of the 1970s. In the 1980s, the supercomputer packaging technologists recognized that this large size required long interconnect wires, restricting the performance gains in supercomputers. The MPS designers must recognize that they need to follow the supercomputer packaging trends of the 1990s and embrace bare-die packaging to reduce the interconnect distances necessary to achieve a smaller and lower cost system.

The MPS already leads the trend to LSI logic through the use of standard microprocessors. The application of MCM packag-

microprocessor dies with a multitude of processors, the traditional supercomputer will have a few custom dies per processor and many processors. The market price of both systems in the same performance range will be about equal, with variations attributable to software costs and profit margins needed to sustain the business.

The continuing trend in semiconductor technology could well lead to a 1-gigabit memory chip by the year 2000. This same processing technology applied to a custom supercomputer processor has the potential to place multiple processors and supporting memory on a single chip (however, the ar-

L. M. Thorndyke is at DataMax, Edina, MN 55435. J. P. Riganati was at the Supercomputing Research Center, Institute for Defense Analyses, Bowie, MD 20715. His present address is the David Sarnoff Research Center, Princeton, NJ 08543. They are both members of the Institute of Electrical and Electronics Engineers (IEEE) Computer Society Scientific Supercomputing Subcommittee.

chitecture scalability and serviceability requirements may preclude this approach, except in small systems). The MPS designers will have to package their standard microprocessors in a custom die, just as in the supercomputer, making the two systems very similar. When this eventually occurs, the supercomputer company will either be vertically integrated—that is, owned by a semiconductor manufacturer—or have vertical cooperation agreements to allow manufacturing of the custom dies on a memory processing line. This shared use allows the semiconductor company to recover some of its very large capital investment in a memory processing line for those situations in which some excess capacity is available.

Memory bandwidth is one of the key parameters determining the performance of parallel systems. From the Cray Research Y-MP to the Cray C90 to the Cray C95, there are rather dramatic advances in the memory bandwidth being provided: a factor of 6 from the Y-MP to the C90 and a factor of 4 expected from follow-on systems. Looking at specific cases of MPS suppliers, we find, in the case of Thinking Machines, that as they moved from the CM1 to the CM2 to the CM5, there has been a decrease in the number of processors and an increase in the system's power. In the case of MasPar, there has been an eightfold increase in the power of the processor with the same number of processors.

If the term "massively" refers only to the number of processors (independent of the other complex and far more important system considerations), it is misleading because current trends indicate that the industry is seeking a balance between numbers and power, not an unusual situation in the history of science. It is the simple pendulum effect. The efforts started with extremely powerful single processors that were internally highly parallel, followed by a multiplicity of small, lower power processors. We are now seeking a balance between the two extremes. Nothing could be more natural. This trend influences packaging dramatically and is consonant with the advantages of compactness and low cost in MCMs.

As clock periods have been reduced, there has been a trend toward paying ever more attention to impedance matching at all levels of the interconnect, within the module and between modules, including through the connectors. Although there has not yet been major concern with impedance matching on the chip, there is little question about the need to deal with this in the very near future for the majority of packaging techniques used today.

One current challenge is to build a balanced supercomputer processor consisting of about 10 million gates with a peak performance well over 1 gigaflop. The proces-

sor must also possess sufficient bandwidth to supply the functional units with multiple data words to and from memory at every clock period. Key systems decisions involving tradeoffs that must be made include such considerations as the use of custom chips housed in small MCMs, as opposed to gate arrays and sophisticated MCMs and combinations thereof. Custom logic is useful if there are few options per system and a mix of storage and logic is required. One must also assume that there is enough production volume to justify a return on investment from the nonrecurring costs of a custom approach. The success of a custom design requires compatibility between a suite of excellent computer-aided design (CAD) tools and a cooperative semiconductor supplier with advanced processes.

Requirements for CAD systems include the ability to handle thermal analysis, to interconnect designs, and to address mechanical considerations. Today one must consid-

er memory versus smart memory or combined memory and logic on the same die. This is a function of the speed and on-off cycles of the chip, the kind of special functions being supported, the volume, and the level of cooperation from an integrated circuit supplier. In the near future, the industry expects 500 to 1000 input-output pins, diamond conduction cooling, liquid cooling, impedance-controlled MCMs, and impedance-controlled high-density interconnects.

We believe that the current supercomputer companies have the requisite systems-integration technology and the packaging experience as demonstrated by compact physical size, power, cooling, and interconnectability. It will be easier for such companies to move into the MPS market quickly than for the current supplier of MPSs to move to the sophisticated packaging required of the highest performance computers. The MPS supercomputer game, therefore, is for the incumbent companies to win or lose.

Workstation Clusters Rise and Shine

Bill Buzbee

Not very long ago, there was only one option for researchers interested in high-performance computing: the supercomputer. But these powerful machines are extremely expensive—so much so that only large research facilities can afford to buy and maintain them. Researchers at other locations can use these supercomputers by working over high-speed networks, but the number of users usually exceeds the available resources. Recently, a lower cost alternative to single-site supercomputing has become practical, with comparable performance: the workstation cluster.

Workstation clusters consist of an ensemble of workstations or high-performance microprocessor systems that are networked together in some fashion and that often appear to the user as a single resource. The equipment can be all of one type, or a mixture of different workstations and several different networks can be used. Potential benefits of workstation clusters include (i) a cost-effective alternative to mainframe systems, (ii) a cost-effective alternative to providing a workstation to each scientist and engineer in an organization, (iii) an approach to utilizing otherwise unused cycles on personal workstations, and (iv)

loosely coupled parallel capability.

All of these benefits are a consequence of the steady and remarkable progress in very large scale integrated-circuit (VLSI) technology. The cost performance, measured in millions of floating-point operations (flops) per dollar, for top-of-the-line workstations has been growing at a compounded rate of 38% per year, in contrast with 10 to 15% for other systems (see figure) (1). It is no surprise that top-of-the-line microprocessors are sometimes referred to as "killer micros," owing to their tendency to devour other systems in the marketplace. Today, a top-of-the-line microprocessor often matches the scalar performance of a single central processing unit (CPU) in a supercomputer, and even in vector mode, a supercomputer CPU seldom outperforms a top-of-the-line micro by more than an order of magnitude. Also, thanks to progress in VLSI technology, microprocessor systems can be cost-effectively equipped with megawords of memory.

These technology trends combined with semiconductor standardization and high-volume production make possible microprocessor systems that cost much less than mainframe and supercomputers. The resultant cost performance advantages are the basis of growing interest in and use of workstation clusters.

A recent acquisition at Lawrence Liver-

The author is director of the Scientific Computing Division, National Center for Atmospheric Research, Boulder, CO 80307.