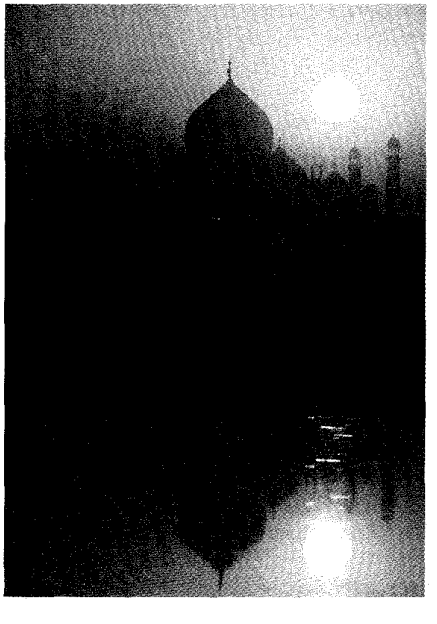
NEW Global Broadcast

Parallel computing in India

David K. Kahaner
Asian Technology
Information Program

India has made significant strides in developing high-performance parallel computers. Many Indians feel that the presence of these systems has helped create a high-performance computing culture in India, and has brought down the cost of equivalent international machines in the Indian marketplace. However, questions remain about the cost-effectiveness of the government funding for these systems, and about their commercial viability.

In 1987, India decided to launch a national initiative in supercomputing in the form of a time-bound mission to design, develop, and deliver a supercomputer in the gigaflops range. The major motivation came from the political delays in getting a Cray XMP for weather forecasting. India's government decided to support the development of indigenous parallel processing technology. In August 1988 it set up the Center for Development of Advanced Computing (C-DAC), located in Pune, with a three-year budget of 375 million rupees (approximately $12 million US).



The C-DAC's First Mission was to deliver a 1-Gflop parallel supercomputer by 1991. Simultaneously, the Bhabha Atomic Research Center (BARC), the Advanced Numerical Research & Analysis Group (Anurag) of the Defense Research and Development Organization, the National Aerospace Laboratory (NAL) of the Council of Scientific and Industrial Research, and the Center for Development of Telematics (C-DOT) initiated complementary projects to develop high-performance parallel computers. Delivery of India's first-generation parallel computers started in 1991.

In this article, I'll summarize the status of these projects, and assess the viability of these efforts both in India and as part of the worldwide picture of advanced computing products. (For more details on the applications and architectures for some of the machines discussed in this article, see Vijay Bhaktar's technical report.[1]) The sidebar provides information on other Indian research in parallel computing.

Param

The C-DAC's mission to build and deliver a 1-Gflop parallel machine started almost from scratch, but came out with its first 64-node prototype in two years. V.P. Bhatkar, the C-DAC's executive director, is generally credited with sparkplugging this effort.

The C-DAC's computers are named *Param* (*para*llel *m*achine), which means "supreme" in Sanskrit. The first Param systems, called the 8000 series, used Inmos 800 and 805 Transputers as computing nodes. Although the theoretical peak-performance of a 256-node Param was 1 Gflop (a single-node T805 performs at 4.25 Mflops), its sustained performance in an actual application turned out to be between 100 and 200 Mflops. The C-DAC named the programming environment *Paras*, after the mythical stone that can turn iron into gold by mere touch.

Early in 1992, the C-DAC realized that the Param 8000's basic compute node was underpowered, so it integrated Intel's i860 chip into

Other research

The five Indian Institutes of Technology, at Bombay, Madras, Kharagpur, Delhi, and Kanpur; the Indian Institute of Science, at Bangalore; the Center for Mathematical Modeling and Computer Simulation, at Bangalore; and several other organizations have made research contributions to high-performance computing. Worth noting among them are an architecture based on a hierarchical network of hypercubes;[1] parallelized applications for medical imaging;[2] and research on VLSI layout,[3] logic programming,[4] scientific computation,[5] multiprocessor operating systems,[6] and interconnection networks.[7] Some of these institutions have built successful prototype or experimental parallel machines.

References

1. J. Mohan Kumar and L.M. Patnaik, "Extended Hypercube: A Hierarchical Interconnection Network of Hypercubes," *IEEE Trans. Parallel and Distributed Systems*, Vol. 3, No. 1, Jan. 1992, pp. 45–57.
2. K. Rajan, L.M. Patnaik, and J. Ramakrishna, "High-Speed Computation of the EM Algorithm for PET Image Reconstruction," *IEEE Trans. Nuclear Science*, Vol. 41, No. 5, Oct. 1994, pp. 1721–1728.
3. B.B. Prahlada Rao, L.M. Patnaik, and R.C. Hansdah, "Parallel Genetic Algorithm for Channel Routing," *Proc. Third Great Lakes Symp. VLSI*, IEEE Computer Society Press, Los Alamitos, Calif., 1993, pp. 69–70.
4. A.V.S. Sastry and L.M. Patnaik, "OR-Parallel Evaluation of Logic Programs on a Multi-Ring Dataflow Machine," *New Generation Computing*, Vol. 10, No. 1, 1991, pp. 23–54.
5. P.C. Mathias and L.M. Patnaik, "Systolic Evaluation of Polynomial Expressions," *IEEE Trans. Computers*, Vol. 39, No. 5, May 1990, pp. 653–665.
6. P. Hatkanagalekar, *Concurrency Issues in Shared-Memory Multiprocessor Operating System Kernels*, PhD thesis, Dept. of Computer Science and Engineering, Indian Inst. of Technology, Bombay, India, 1995.
7. R. Mittal, B.N. Jain, and R.K. Patney, "Link Augmented Binary (LAB)-Tree Architecture, *IEE Proc. E: Computer and Digital Techniques*, Vol. 140, No. 2, Mar. 1993, pp. 127–133.

the Param architecture. The objective was to preserve the same application programming environment and provide straightforward hardware upgrades by just replacing the Param 8000's compute-node boards. This resulted in the Param 8600, an architecture with the i860 as a main processor and four Transputers as communication processors, each with four built-in links. The C-DAC extended Paras to the Param 8600 to give a user view identical to that of the Param 8000. Param 8000 applications could easily port to the new machine. The C-DAC claimed that the sustained performance of the 16-node Param 8600 ranged from 100 to 200 Mflops, depending on the application.

Both the C-DAC and the Indian government considered that the First Mission was accomplished and embarked on the Second Mission, to deliver a teraflops-range parallel system capable of addressing grand challenge problems. This machine, the Param 9000, was announced in 1994 and exhibited at Supercomputing '94. The C-DAC plans to scale it to teraflops-level performance.

System architecture

The Param 9000's multistage interconnect network uses a packet-switching wormhole router as the basic switching element. Each switch can establish 32 simultaneous nonblocking connections to provide a sustainable bandwidth of 320 Mbytes per second. The communication links conform to the IEEE P1355 standards for point-to-point links.

The Param 9000 architecture emphasizes flexibility. The C-DAC hopes that, as new technologies in processors, memory, and communication links become available, those elements can be upgraded in the field. The first system is the Param 9000/SS, which is based on SuperSparc processors. A complete node is a 75-MHz SuperSparc II processor with 1 Mbyte of external cache, 16 to 128 Mbytes of memory, one to four communication links, and related I/O devices. When new MBus modules with higher frequencies become available, the computers can be field-upgraded. Users can integrate Sparc workstations into the Param 9000/SS by adding an Sbus-based network interface card. Each card supports one, two, or four communication links. The C-DAC also provides the necessary software drivers.

Programming environment

According to the C-DAC, Paras 9000/SS blends industry standards with parallel programming extensions to provide high performance and ease of use (but see my comments in "An assessment"). The environment achieves high performance through an optimized microkernel, parallel file system, standard and enhanced compiler optimizations, and parallel libraries.

The Param 9000/SS supports ANSI C, C++, Fortran 77, Fortran 90, and High Performance Fortran. Developers can use either the C-DAC's own compilers or third-party compilers for program development with standard and enhanced optimizations.

For flexibility in client/server environments, the system's service nodes support standards-conforming interfaces for Ethernet and FDDI connections and the C-DAC's interconnect network. The C-DAC expects to add support for ATM networks soon. Networking software includes standard utilities and protocols such as ftp, TCP/IP, sockets, telnet, and NFS.

Applications

The C-DAC says that application kernels have been developed for computational fluid dynamics, finite-element analysis, oil reservoir modeling, seismic data processing, image processing, remote sensing, medical imaging, signal processing, radio astronomy, molecular modeling, biotechnology, quantum mol-

ecular dynamics, quantum chemical calculations, semiconductor physics, composites and special materials, power systems analysis and energy management, and discrete optimization.

Anupam

The BARC, founded by Homi Bhabha and located in Bombay, is India's major center for nuclear science and is at the forefront of India's Atomic Energy Program. Through 1991 and 1992, BARC computer facility members started interacting with the C-DAC to develop a high-performance computing facility. The BARC estimated that it needed a machine of 200 Mflops sustained computing power to solve its problems. Because of the importance of the BARC's program, it decided to build its own parallel computer.

In 1992, the BARC developed the *Anupam* (Sanskrit for "unparalleled") computer, based on the standard Multibus II i860 hardware. Initially, it announced an eight-node machine, which it expanded to 16, 24, and 32 nodes. Subsequently, the BARC transferred Anupam to the Electronics Corporation of India, which manufactures electronic systems under the umbrella of India's Department of Atomic Energy.

System architecture

Anupam has a multiple-instruction, multiple-data (MIMD) architecture realized through off-the-shelf Multibus II i860 cards and crates. Each node is a 64-bit i860 processor with a 64-Kbyte cache and a local memory of 16 to 64 Mbytes. A node's peak computing power is 100 Mflops, although the sustained power is much less. The first version of the machine had eight nodes in a single cluster (or Multibus II crate). There is no need for a separate host.

Anupam scales to 64 nodes. The intercluster message-passing bus is a 32-bit Multibus II backplane bus operating at 40 Mbytes/s peak. Eight nodes in a cluster share this bus. Communication between clusters travels through two 16-bit-wide SCSI buses that form a 2D mesh. Standard topologies such as a mesh, ring, or hypercube can easily map to the mesh.

Programming environment

Anupam's parallel programming environment is simple and straightforward. The master processor runs the Unix SVR4 operating system. Users can write parallel programs in Fortran 77 or ANSI C. The environment provides a Fortran vectorizer. To aid program development and debugging, it also includes a profiler and a debugger. The environment uses Hoare's Communicating Sequential Processes model.

The message-passing library consists of send and receive calls to facilitate the writing of parallel codes by calling the library routines in either Fortran or C. The environment also includes a program scheduler, batch queue manager, and parallel library for scientific and image-processing routines.

Applications and performance

BARC scientists and engineers, as well as users from outside institutions, have extensively used Anupam. BARC claims that more than 50 applications have been parallelized and are being used.

For the standard Linpack benchmark with a 1,000×1,000 matrix in double-precision calculation, the eight-node Anupam achieved 52 Mflops, with 80% efficiency. For matrix multiplication, the same configuration achieved 106 Mflops, with 88% efficiency. A molecular dynamics program to calculate electronic structures using linear approximations in band theory demonstrated 81% efficiency on a 16-node configuration.

PACE

Anurag, located in Hyderabad, focuses on R&D in parallel computing; VLSIs; and applications of high-performance computing in computational fluid dynamics, medical imaging, and other areas. Anurag has developed the *Processor for Aerodynamic Computations and Evaluation*, a loosely-coupled, message-passing parallel processing system.

The PACE program began in August 1988. The initial prototypes used the 16.67-MHz Motorola MC 68020 processor. The first prototype had four nodes and used a VME bus for communication. The VME backplane works well with Motorola processors and provided the necessary bandwidth and operational flexibility. Later, Anurag developed an eight-node prototype based on the 25-MHz MC 68030. This cluster forms the backbone of the PACE architecture. The 128-node prototype is based on the 33-MHz MC 68030. To enhance the floating-point speed, Anurag has developed a floating-point processor, *Anuco*. The processor board has been specially designed to accommodate the MC 68881, MC 68882, or the Anuco floating-point accelerators.

PACE+, the latest version, uses a 66-MHz HyperSparc node. The memory per node can expand to 256 MBytes.

Programming environment

PACE's programming environment is simple and straightforward. The user interacts with the front-end processor, which is a standard Unix engine with a HyperSparc processor running Solaris. PACE treats parallel processors as a resource of the front-end processor. The user writes a sequential host program. The user parallelizes the computationally intensive parts, which are treated as subroutines (node programs) to be called by the host program.

Anurag's parallel programming environment is called *Anupam* (Anurag's Parallel Applications Manager). Anupam runs under Unix and consists of several modules and utilities, including a preprocessor, a simulator, a queue manager, runtime libraries, a communications debugger, a source-level debugger, and a parallel library. The Anupam software depends only on the availability of Unix at the front end. The software is portable across machines, with very few modifications (relating to the physical addresses of the CPU boards).

Applications and performance

Several application programs have been run on the various models of PACE. These include Linpack, fast Fourier

transform, neural network simulations, finite-element method codes, and several computational fluid dynamics codes. The 128-node PACE with an MC 68030 and an MC 68882 coprocessor delivered over 30 Mflops for large problems, with 0.33 Mflops per processor node. Later, Anurag enhanced the performance to 0.75 Mflops per node by incorporating Anuco. With the Sparc II processors, the speed is 4.5 Mflops per node. The latest Sparc processors should offer higher performance. The 128-node configuration is supposed to provide a Linpack (1000×1000) rating of 375 Mflops (single precision).

Flosolver

In 1986, the NAL, located in Bangalore, started a project to design, develop, and fabricate suitable parallel processing systems to solve fluid dynamics and aerodynamics problems. The project was motivated by the need for a powerful computer in the laboratory and was influenced by similar international developments.

Flosolver, the NAL's parallel computer, was the first operational Indian parallel computer. Since then, the NAL has built a series of updated versions, including Flosolver Mk1 and Mk1A, four-processor systems based on 16-bit Intel 8086 and 8087 processors; Flosolver Mk1B, an eight-processor system; Flosolver Mk2, based on Intel's 32-bit 80386 and 80387 processors; and the latest version, Flosolver Mk3, based on Intel's i860 RISC processor.

System architecture and programming environment

The current version of the Mk3 uses eight processors with an onboard memory of 64 Mbytes per processor. The system bus is a Multibus II with a bandwidth of 40 Mbytes/s. A message-passing coprocessor and high-speed direct-memory access controllers on each board assist communication between processors.

Applications

NAL scientists have used Flosolver primarily for the weather forecasting code T80 under a project from India's Department of Science and Technology. They've also used it to solve computational fluid dynamics problems.

Chipps

The Indian government launched the C-DOT to develop indigenous digital switching technology. The C-DOT, located in Bangalore, completed its First Mission in 1989 by delivering technologies for rural exchanges and secondary switching areas. In February 1988, the C-DOT signed a contract with the Department of Science and Technology to design and build a 640-Mflop, 1,000-MIPS-peak parallel computer. The C-DOT set a target of 200 Mflops for sustained performance.

There is considerable sentiment in the Indian scientific community that the government support of these projects has been too costly, and the results not cost-effective.

System architecture

C-DOT's High Performance Parallel Processing System (Chipps) is based on the single-algorithm, multiple-data architecture. Such an architecture provides coarse-grain parallelism with barrier synchronization, and uniform startup and simultaneous data distribution across all configurations. It also employs off-the-shelf hardware and software. Chipps supports large, medium, and small applications. The system has three versions: a 192-node, a 64-node, and a 16-node machine.

Programming environment

Chipps employs data-parallelism. Programmers can easily parallelize applications written for sequential machines by identifying the tasks that are data-independent. The environment divides application program functionally as tasks—that is, I/O, data management, and computation. The processing element performs the computation, and the microcontroller handles control and I/O.

Applications

The C-DOT designed Chipps primarily for weather forecasting and radio astronomy applications. Recently, they ported several scientific kernel codes on the machine to demonstrate its general applicability in scientific and engineering applications.

An assessment

The technological developments in parallel computing in India have been considerable. They show that the general evolution of society in India is such that if appropriate financial resources are available, parallel computer systems can be designed and built around available microprocessor chips along with system software, and many application packages can be ported onto these machines. That five such efforts in different organizations and cities have borne fruit suggests that there is no shortage of leadership, technical, and organizational skills. Furthermore, these successes have not only enhanced the self-confidence of the participating organizations, but also helped relieve bottlenecks in scientific research and technological development. However, all is not positive.

There is considerable sentiment in the Indian scientific community that the government support of these projects has been too costly, and the results not cost-effective. One reason for the different organizations' simultaneous efforts in parallel computing is the emphasis on parallelizing different applications. These organizations, which report to independent government departments, possibly did not have strong ties in the initial stages of their design and development activity. Some scientists feel that a single organization concentrating on parallel computing would have had a much stronger impact. More applica-

tions could have been ported onto a single machine, with significant performance on standard benchmarks. Such an approach could have given Indian high-tech products the necessary international visibility and commercial viability.

In terms of performance and software support, the Indian high-performance computers hardly compare to the best commercial machines. For example, the C-DAC's 16-node Param 9000/SS has a peak performance of 0.96 Gflops, whereas Silicon Graphics' 16-processor Power Challenge has a peak performance of 5.96 Gflops, and IBM's 16-processor Para2 model 590 has a peak performance of 4.22 Gflops. However, the C-DAC hopes that a future Param based on DEC Alpha processors will match such performance.

In addition, we have not seen uptime and utilization statistics for these systems, and it is highly unlikely that they will compare favorably with statistics for commercial products. One such product is the Convex C 3820 that was installed at the Center for Mathematical Modeling and Computer Simulation. Its uptime and utilization were between 95% and 100%, based on 23,000 CPU hours utilized during the most recent quarter. Also, the prices of the supercomputers from companies such as Cray, IBM, and DEC have fallen significantly in India, making these commercial systems relatively more affordable.

The Indian parallel computers have had trouble meeting the users' requirements. A number of experts who have access to these computers believe that the systems are not easy to use for many scientific users, primarily because systems software and software-development tools are not fully developed or mature. In 1993, the Indian Institute of Science at Bangalore installed an earlier version of Param. Its use has not been as high as expected, because of unfamiliar software and development tools, and difficulties with functioning in a network environment. At least one large Indian computing establishment has considered Indian systems as serious candidates in an expansion plan, and included them as part of an objective procurement exercise, where bids were accepted from all parties. In that case, in the first screening the decision-making committee rejected the one Indian system that was bid, giving nearly a dozen reasons why it did not meet the technical requirements.

These systems have suffered from poor sales. About 25 Param installations are in India, at various academic and R&D organizations; however, this includes some very small systems. The few Indian parallel computers sold outside India have been primarily to research institutes already collaborating in development activities. Notable C-DAC installations outside India are in Russia, Germany, and Canada. No significant number of parallel machines from other Indian manufacturers have been installed, in or outside India.

To be marketable, a system designed and manufactured around a general-purpose microprocessor chip must appear in a year or so after the chip hits the market.

The Indian government has possibly spent tens of millions of dollars on parallel computing development, with only a very small fraction of that returning as sales. Most of the sales were for Params, and most of these were sold to India's defense community. (Wipro, one of India's largest information technology companies, has been trying to market Flosolver, but no sales have been reported.) Unbiased evaluation is difficult because, for many defense technology applications (composites, avionics, cryogenics, superconductivity, and so on), acquisition cost is not the dominant factor. The important thing is to show that the job can be done; Param and others have proved that. However, defense budgets are coming under pressure, and there is a considerable push to commercialize the parallel computing projects. That is where most will meet their end.

As I mentioned earlier, the C-DAC designed the Param 9000 to be processor-independent, so that new and improved chips can be easily incorporated. However, the pace of development is very rapid. To be marketable, a system designed and manufactured around a general-purpose microprocessor chip must appear in a year or so after the chip hits the market. This usually requires access to beta-version chips and advance technical information from the chip manufacturer. At present, the Indian organizations involved in parallel computing do not seem to have this type of relationship with the chip manufacturers, and their machines appear on the market fairly late. They could get around this disadvantage by building functionality to provide individuality and market appeal. But without such a strategy, the chances are small that Indian organizations can produce a commercially successful product based on a message-passing architecture.

With enough funding and effort, India can develop its own teraflops supercomputers, perhaps in the next few years. Whether or not India can sustain such an effort will depend on how well it can surmount these problems.

Acknowledgment

L.M. Patnaik developed a significant amount of the factual material for this report. Contact him at the Microprocessor Applications Lab, Indian Inst. of Science, Bangalore 560 012, India; phone 91 (80) 334-2451 or 334-4411, extension 2520 or 2368; fax 91 (80) 334-1683; lalit@micro.iisc.ernet.in.

Reference

1. V.P. Bhatkar, "Parallel Computers Galore in India," tech. report, Centre for Development of Advanced Computing, Pune, India.

David K. Kahaner is the director of the nonprofit Asian Technology Information Program. He can be contacted at the ATIP, Harks Roppongi Bldg. 1F, 6-15-21 Roppongi, Minato-ku, Tokyo 106, Japan; kahaner@atip.or.jp.