OCCASION

This publication has been made available to the public on the occasion of the 50th anniversary of the United Nations Industrial Development Organisation.

DISCLAIMER

FAIR USE POLICY

CONTACT

Please contact publications@unido.org for further information concerning UNIDO publications.

For more information about UNIDO, please visit us at www.unido.org

UNITED NATIONS INDUSTRIAL DEVELOPMENT ORGANIZATION
Vienna International Centre, P.O. Box 300, 1400 Vienna, Austria
Tel: (+43-1) 26026-0 · www.unido.org · unido@unido.org

Dear Prof. Szymanski :

Following is the ASCII version of my chapter which does not have the figures. The figures are to be taken from the hardcopies. Please acknowledge receipt of the same.

Regards

21339

L M Patnaik

HIGH PERFORMANCE COMPUTING IN INDIA AND FAR-EAST

L.M. Patnaik
Microprocessor Applications Laboratory
Indian Institute of Science
Bangalore - 560012
India
Email: lalit@micro.iisc.ernet.in

Summary

With a vision for developing its own advanced computing technology based on strong intellectual resources, India launched a major initiative in parallel computing in 1988. The launch of the Centre for Development of Advanced Computing (C-DAC) and concurrently other efforts at National Aerospace Laboratory (NAL), Bangalore, Advanced Numerical Research & Analysis Group (ANURAG), Hyderabad, Bhabha Atomic Research Centre (BARC), Bombay, Centre for Development of Telematics (C-DOT), Bangalore, marked the beginning of high performance computing in India. Today, India has designed its own high performance computers in the form of PARAM by C-DAC, FLOSOLVER by NAL, PACE by ANURAG, ANUPAM by BARC and CHIPPS by C-DOT. These machines are contemporary in architecture, brought down the cost of equivalent international machines in the Indian market place, created high performance computing culture in India, and have contributed to several mission critical advanced research programmes. C-DAC is now on the way to develop its own massively parallel teraflops architecture by 1998, placing India in league with several advanced countries as far as supercomputing technology and applications are concerned. This chapter reviews these developments in India, other than highlighting the future directions in this area. A brief discussion on the high performance computing activities in some other countries in the far-east is also included.

INTRODUCTION

Although the performance of single processors has been steadily increasing over the years, the only way to build the next generation teraflop architecture supercomputers seems to be through the root of parallel processing.

Even with today's workstation class high performance processors crossing 100 mega-flops, thousands of processors are required to build the teraflop architecture machine. One of the grand challenges of computing technology of the 1990s is to build high-performance computers in the teraflops range. The high-performance computers are required to solve the so-called grand challenge problems of science and engineering. Several such grand challenge problems have now been posed and the lot is continously growing.

From the applications point of view, there are several motivations to build teraflop machines to solve grand challenge problems in science and engineering. These grand challenge problems include designing aircrafts,

simulating semiconductor materials, analyzing fuel combustion, rational design of drugs, understanding catalysis, designing protein structures, researching the human anatomy, imaging, predicting the weather, studying air pollution, ocean modeling, evaluating ozone depletion and oil exploration. The benefits that result from application of these high speed machines will touch every facet of life such as the air we breathe, the weather, the ozone and global-warming problems which are threatening humankind, design of new generation aircraft, drug research, improved medical therapies, more efficient engines and cars, new chemicals, stronger structural materials, faster VLSI chips. The goal is to model the reality more and more accurately on the supercomputers and to simulate the real phenomena in unprecedented details of chemistry, physics, biology and other sciences. Supercomputing has now become the third mode of scientific investigation in the 90s, complementing theory and practice. It is now widely believed that the supercomputing technology at its cutting edge will hold the key to future competitivity of nations in advanced science and technology, business and security.

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 delays in getting a CRAY XMP for weather forecasting on one hand and a firm belief that India can develop its own high performance computing technology leveraging upon its high-class intellectual resources on the other. Right from the beginning, it was clear that the destiny of supercomputing in India would be carved through the parallel processing route. The Centre for Development of Advanced Computing (C-DAC) was set up for this purpose in August 1988 with the First Mission 3-year budget of Rs. 375 million (approximately US $ 12 million). The launch of C-DAC marks the beginning of high performance computers in India.

The only high performance computers in India at that time were IBM 3090, DEC 10/20, CDC Cyber 730/830, ICL 2100, NEC S1000 series mainframe computers, NORSK DATA ND1000/550 superminis and HP 1000, DEC PDP 11 and MicroVAX minicomputers. The peak computing power of these mainframe computers, superminis, and minis was less than 10 megaflops. Supercomputers were viewed as a strategic resource for India's advanced education and research programme.

C-DAC's First Mission was directed to deliver 1000 MFlops parallel supercomputer by 1991. Simultaneously, several other complementary projects were initiated to develop high-performance parallel computers at National Aerospace Laboratory of the Council of Scientific and Industrial Research (CSIR), the Centre for Development of Telematics (C-DOT), Advanced Numerical Research & Analysis Group (ANURAG) of Defence Research and Development Organisation (DRDO) and Bhabha Atomic Research Centre (BARC). India's first generation parallel computers were delivered starting from 1991. C-DAC has already achieved significant commercialization of its advanced computing technology. India is now in the process of announcing its second generation parallel computer and C-DAC has already began its Second Mission to develop teraflop architecture massively parallel supercomputers. This article presents the current state-of-the-art of Indian high performance computers which have come out of the efforts of C-DAC, BARC, NAL, ANURAG and C-DOT.

C-DAC's PARAM 9000

Background and Evolution

C-DAC formally launched its first mission in August 1988 to deliver the Gigaflops range parallel machines. C-DAC had started almost from scratch with very little knowledge in this fast advancing field. But under the dynamic leadership of its executive director Dr.V.P.Bhatkar, it came out with the first 64 node prototype in barely two years. The machine was taken to Zurich in CONPAR 90, a major parallel processing conference in

Europe, and the international scientific community saw that India had developed a parallel computer comparable to the machines developed in Europe. Indeed, the target machine of 256-node complete with parallel disk array storage, comprehensive parallel programming environment and multi-user and multi-host capability was delivered in August 1991 without time or cost overruns.

C-DAC's parallel supercomputers have been named PARAM, meaning in Sanskrit "Supreme". It also made a nice acronym for a PARAllel Machine. The programming environment is called PARAS (the mythical stone which can turn iron into gold by mere touch) which gave a golden touch to the underlying machine and made the job of programmer or user considerably easy. The first PARAM series scalable supercomputers were based on INMOS Transputers 800/805 as computing nodes, and the first PARAM models were called PARAM 8000 series systems.

During the first mission, C-DAC had launched fairly large applications development programme spanning over 20 academic institutions, research laboratories and industries. Through a series of workshops, over 200 scientists and engineers were trained to use PARAM. A lot of hand-holding was done to parallelize the sequential codes on PARAM. Initial results were stunning, the performance of PARAM 8000's single node sometime exceeded the then prrformance of the available popular workstations like microVAX, and even mainframes. The scalability of PARAM 8000 was demonstrated in a variety of applications like Finite Element Methods, Computational Fluid Dynamics, Computational Physics, Computational Chemistry and Monte Carlo Simulation.

Although the theoretical peak-performance of 256 node PARAM machine was 1 GigaFlops (single node T805 claiming 4.25 MFlops), its sustained performance in actual application turned out to be between 100 to 200 MFlops. The scalability in many applications was, however, excellent.

At the beginning of 1992, the basic compute node of PARAM 8000 was found to be underpowered. INTEL had announced their i860 RISC processor which claimed a theoretical peak-performance of 60 MFlops. A decision was taken at this time to integrate i860 into the PARAM architecture. C-DAC's objective was to preserve the same application programming environment and provide straightforward hardware upgradability by just replacing the compute node boards of PARAM 8000. This resulted in the next architecture with i860 as a main processor with 4 transputers acting as communication processors each with 4 built-in links. PARAS programming environment was extended to PARAM 8600 to give an identical user view as PARAM 8000. C-DAC succeeded during 1992 and 1993 in building a scalable parallel machine, which was called PARAM 8600.

The computing power of 4 compute clusters of PARAM 8000 could now be realized in a single compute cluster of PARAM 8600. The sustained performance of 16 node PARAM 8600 was in the range of 100-200 MFlops, depending on the application. Thanks to PARAS 8600, PARAM 8000 applications could be easily ported on PARAM 8600 machines.

The effort required to deliver the PARAM 8000/8600 machines exceeded 300 man years. Everything was developed from the root level, only the chips were imported. The software effort extended to over a million lines of source code! In this process, C-DAC built strong foundations for undertaking next challenges.

In August 1988 C-DAC began from scratch with a small core team; at the end of 1992, a strong institution of 200 outstanding scientists and engineers was created with Headquarters at Pune and centres at Bangalore and Delhi,in India. C-DAC had also installed a world class Electronic Design Automation (EDA) and software development environments. The institution became well-known not only in India but also in the world. Several interational alliances were formed with leading interational universities from Russia, Europe and North America.

Second Mission

It is in this backdrop of the accomplishment of the First Mission. C-DAC started its Second Mission. While in the First Mission C-DAC delivered the Gigaflops range parallel supercomputers, the goal of the Second Mission was to deliver the teraflops range massively parallel grand challenge supercomputer. This meant that C-DAC was aiming for a supercomputer with 1000 times more performance than what was delivered in the First Mission. The motivations for this goal were many. C-DAC wanted to consolidate the gains of the First Mission and run the race for the teraflops range supercomputers. India became the only country running this race outside USA, Europe and Japan.

Within 18 months of its Second Mission, C-DAC announced its second generation machine called PARAM 9000. The machine was exhibited in Supercomputing '94 at Washington. Progressively, through the successive years, PARAM 9000 will be scaled up to teraflops level.

System Architecture

C-DAC has advented the OpenFrame Architecture for PARAM 9000 series systems which heralds the era of flexible supercomputing. The OpenFrame Architecture brings a new and innovative concept in scalable parallel supercomputing. The distinctive features of the architecture are processor independence and unification of cluster and massively parallel computing. The new architecture allows the integration of different processors as compute nodes and unifies both cluster and massively parallel computing within a single framework, thus providing a truly open framework. The inherent property of OpenFrame Architecture permits continous technology upgradation as new processors and advances in interconnect technology become available.

The OpenFrame systems support the popular and powerful RISC processor based compute nodes, industry standard networking, I/O devices, standards based parallel programming environment. The OpenFrame Architecture provides for a variety of scalable I/O and networking interfaces, with Ethernet and fast and wide SCSI as standard. Available options are multiple HiPPI, multiple FDDI, multiple SCSI, Ethernet and ATM interfaces.

The nucleus of the OpenFrame Architecture is a modularly scalable multistage interconnect network accommodating more than thousand heterogeneous processing nodes. The internode communication is via a low latency, high bandwidth point-to-point links. The multistage interconnect network of PARAM 9000 uses a packet switching wormwhole router as the basic switching element. Each switch is capable of establishing 32 simultaneous non-blocking connections to provide a sustainable bandwidth of 320 MBytes/sec. The communication links of PARAM 9000/SS conform to the IEEE P1355 standards for point-to-point links.

The interconnect network is non-blocking and provides full connectivity to all the nodes, thus freeing the programmers from having to concern themselves with topologies. All the nodes are equidistant, resulting in predictable and repeatable performance. The dynamic adaptive routing ability of the interconnect network avoids hot spot build-up, which is typical in irregular problems. If the targeted output link is busy, other links in the group are dynamically and automatically selected for message routing, thereby removing the congestion. Future additions to the interconnect network include hardware support for HPF and MPI. Planned features are for the global reduction, broadcast/multi communications, barrier synchronization and data distribution. The communication fabric allows the partitioning of the system into production and development environments in any arbitrary manner and can be changed by the system administrator to suit the dynamic needs. Even within these two partitions,

multi-user access is provided and these partitions are non-blocking.

The PARAM 9000 carries the philosophy of flexibility into the node architecture. As new technologies in processors, memory and communication links advance and become available, these can be upgraded in the field. First offering of PARAM 9000 architecture is PARAM 9000SS system based on SuperSparc series processors. The complete node is realized using the SuperSparc II processor with 1 MB of external cache, 16 to 128 MB of memory, one to four communication links and related I/O devices. The current operating speed of the processor is 75 MHz and as and when new MBus modules with higher frequencies become available, they can be field-upgraded. MBus speed is 50 MHz. Each node can be configured with 16, 32, 64, 80 or 128 MB high speed, error correcting memory and can be field-upgraded.

Communication links can be scaled from 1 to 4 links, providing a bandwidth of 10, 20 or 40 MBytes/sec. The processor is used only for the initiation of the message and is interrupted whenever a programmable number of messages have been received. The DMA engine on the communiction interface performs the message packetization, routing of packets on all the available links and reassembly of packets. The communication protocol defines three message types based on the message length; very short message, short message and long message. While the very short message is typically used by the kernel and can be assigned priority over others, the other two types of messages are packetized and time-multiplexed.

One out of every four nodes can be configured as an I/O or server node, satisfying a variety of I/O needs. While the service nodes run Solaris operating system, the compute nodes run the PARAS microkernel.

Users can now integrate the SPARC workstations into PARAM 9000/SS by just adding the SBus based network interface card. Each network interface card supports 1, 2 or 4 communication links. C-DAC also provides the necessary software drivers.

The OpenFrame Architecture supports a variety of scalable networking and mass storage options conforming to industry standards. These interfaces are connected to the I/O nodes running Solaris operating system. PARAM 9000/SS supports multiple fast and wide SCSI channels for connection to external RAID boxes. Any off-the-shelf SCSI-in-SCSI-out RAID boxes can also be used to realize a range of storage needs, including C-DAC's own scalable I/O system. Through the Solaris operating system, the user can perform all the regular file operations including NFS, remote file system, etc. The system's I/O nodes support standards conforming interfaces to ethernet, FDDI and HiPPI. Via the Solaris operating system, the standard software utilities and protocols like FTP, TCP/IP and sockets, telnet, NFS, etc. are supported.

Programming Environment

A high performance computing system is only as open and flexible as its software. The open and flexible software environment is an integral part of the OpenFrame Architecture advented by C-DAC. The architecture permits the parallel processing system to be viewed as an ensemble of independent workstations, a cluster of workstations or as massively parallel systems connected through a scalable high bandwidth network or any combination of these.

The users of parallel programming environment are looking for openness, industry standards, a spectrum of sophisticated development tools and flexibility of composition depending on requirements for robustness, performance and ease-of-use. The software for OpenFrame Architecture meets this requirement with a comprehensive range of software tools and utilities which can cater to the different needs of the user.

PARAS 9000/SS is C-DAC's parallel program development environment, for C-DAC's Sparc processing nodes based scalable massively parellel system - PARAM 9000/SS. The software environment, PARAS 9000/SS seamlessly blends industry standards with parallel programming extensions to provide both high performance and ease of use. High performance is achieved through optimized microkernel, parallel high performance file system, standard and enhanced compiler optimizations and parallel libraries.

The PARAS 9000/SS supports the two main massively parallel programming models: data parallelism and multiprocess parallelism. Data parallelism is automatically realized through the support for HPF, while CORE, PVM and MPI message passing interfaces provide the required support for multiprocess parallelism.

Solaris is the preferred program development OS environment of PARAM 9000/SS and runs on all the service nodes of the system. For performance and efficiency during production runs, the PARAS microkernel is replicated on the compute and I/O nodes. The microkernel provides all the necessary services to the application program, while dispensing with the overheads of a standard OS. The result is better performance on the production codes and ease of development through a standard OS. The salient features of PARAS microkernel are Mach like process management, enhanced exception handling support, simple virtual memory model, port based interprocess communication. The abstractions that are supported are tasks and threads, ports and port groups and virtual memory regions. To enhance the portability of applications, popular message passing interfaces PVM and MPI are provided. Applications written using the PVM and MPI libraries can be ported on to PARAM 9000/SS effortlessly.

The operating system configures the system into service, compute and I/O partitions. The user logs onto a service node and uses the PARAS program development tools to develop a parallel application. Upon request from the user, the resource manager allocates a pool of processors and maps the application onto the allocated nodes. The user's parallel application can be distributed betweeen the compute and service partitions This flexibility allows the user to configure Solaris servers to provide specialized services to the compute node tasks of the parallel application. The operating system space shares the compute partition across multiple users. Tha nodes of the service partition run Solaris 2.x while the nodes of the compute and I/O nodes run PARAS microkernel along with appropriate servers.

The languages supported by PARAM 9000/SS include ANSI C, C++, Fortran-77, Fortran-90 and the emerging HPF. Either C-DAC's own compilers or third party compilers can be used for program development with standard and enhanced optimizations. The interconnect network of PARAM 9000/SS will provide hardware support for MPI and HPF for speedier execution of the applications. The features that will be supported are global reduction, broadcast/multicast communications, barrier synchronization and data distribution.

PARAS 9000/SS provides a variety of tools for program development and debugging. FORGE 90, the program restructuring tool from Applied Parallel Research, Inc., AIDE and PET from C-DAC are the development tools available under PARAS. C-DAC will also provide Total View debugger for cluster computing environment soon. PARUL - the parallel libraries from C-DAC contain more than 400 routines for dense linear equation solution and eigen value determination, sparse linear equation solution, BLAS level 3 - basic linear algebra sub-routines, Poisson's equation solution, 1-D and 2-D FFTs and general purpose sorting.

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

In addition to accessing the file system of the service nodes, a separate dedicated and high performance scalable mass storage server based on multiple I/O nodes is also supported to provide high bandwidth and large storage capacity. Parallel applications that run on compute nodes under PARAS can use both UNIX file I/O calls, or parallel I/O primitives which efficiently and transparently map structured data such as matrices over multiple I/O nodes. Reliability and speed are enhanced by connecting RAID systems to the I/O nodes. The mass storage server follows the IEEE Mass Storage System Reference model.

A distributed visualization environment will be supported on PARAM 9000/SS. KHOROS, one of the most widely used visualization system by computation scientists will be provided as the basic environment. Other standard visualization packages like IRIS Explorer will also be made available.

The PARAM 9000/SS system can be partitioned mainly into batch, interactive and production partitions.These partitions can be created by systems administrator depending on the need. C-DAC supports CODINE, the distributed computing management utility for the management of heterogeneous workstation clusters which are integrated with vector and parallel compute servers.

C-DAC is committed to maintaining backward compatibility with its earlier generation of PARAM series of parallel machines. Applications developed using PARAS 8000/8600, can be recompiled and executed on the PARAM 9000/SS machines. Applications developed on workstation clusters and other parallel machines using the PVM and MPI message passing interfaces just require to be recompiled for execution on PARAM 9000/SS.

C-DAC will enhance the PARAS 9000/SS programming environment to provide a single system image of UNIX. The directions for parallel programming environment are evolving continously and new standards are being defined. The parallel programming environment as envisioned in the OpenFrame Architecture is planned to conform to these emerging standards.

Applications on PARAM

Development and porting of parallel processing applications is a major driving force of C-DAC Mission. In the First Mission, over 40 application kernels were developed in collaboration with user agencies and demonstrated on PARAM 8000 and 8600 series machines. Emerging application areas such as parallel database management, complex query decision support system and video-on-demand will also be supported. C-DAC's application development program includes development of parallel libraries, application kernels and benchmarking, parallelisation and porting of production quality industry standard packages, abinitio development of parallel application packages in select areas, wide scale catalysis and education and research in parallel processing and installation and networking of parallel supercomputing facilities.

Several application kernels have been developed on PARAM 8000/8600 series machines in the areas of computation fluid dynamics, finite element analysis, oil reservoir modelling, seismic data processing, image processing, remote sensing, medical imaging, signal processing, radio astronomy, molecular modelling, biotechnology, quantum molecular dynamics, quantum chemical calculations, semiconductor physics, composites and special materials, power systems analysis and energy management, and discrete optimization. In the Second Mission, attention has been focussed on the parallelisation and porting production quality industry standard codes in collaboration with various organizations.

Background and Evolution

Bhabha Atomic Research Centre is a premier national centre for nuclear science and allied disciplines founded by Dr. Homi Bhabha and has since then been at the forefront of India's Atomic Energy Programme. Over 2000 scientists and engineers here make use of the central computing facility for solving their computational problems. Until 1992, the computers available to BARC scientists were the last generation Norsk Data Computers, which took hours of computational time over moderate problems of BARC scientists and engineers. BARC began some initial work on parallel processing by configuring transputer network for image processing based on hardware procured from INMOS and C-DAC. Through 1991 and 1992, BARC computer facility members started interacting with C-DAC to have a high-performance computing facility. It was estimated that a machine of 200 MFlops of sustained computing power would be needed to solve the current problems. In view of the vulnerability of this programme, BARC decided to build their own parallel computer.

In 1992 BARC system integrated their own parallel computer based on the standard Multibus II i860 hardware. Initially an 8-node machine was announced which was expanded to 16 node and then to 24 and 32 node systems. BARC's parallel machine was called ANUPAM, meaning "unparalleled" in Sanskrit. BARC transferred the technology of ANUPAM to the Electronics Corporation of India Ltd. (ECIL) located in Hyderabad, a public sector unit manufacturing electronic systems under the umbrella of Department of Atomic Energy of the Govt. of India.

System Architecture

BARC's ANUPAM is an MIMD architecture (Fig.1) machine realized through standard off-the-shelf MULTIBUS II i860 cards and crates supplied by WIPRO. Each node is a 64 bit i860 processor with 64 KB cache and a local memory of 16-64 MB. The peak computing power of a single node is 100 MFlops, although the sustained power is much less. The first version of the machine was announced with 8 nodes in a single cluster (or Multibus II crate). There is no need for seperate host (Node 0 acts as the host processor).

ANUPAM is scalable to 64 nodes. The inter-cluster message passing bus is 32 bit MULTIBUS II backplane bus operating at 40 MB/sec. peak. This bus is shared by 8 nodes within a cluster. The communication between clusters is achieved through two 16 bit wide SCSI busses -- one in X and other in Y direction, forming a 2D mesh. The inter-cluster bus realized through standard SCSI controller chips has a peak speed of 20 MB/sec. Standard topologies like mesh, ring, hypercube, etc. can be easily mapped on the 2D mesh topology. The interconnection network is scalable to 64 processors with 8 MULTIBUS II crates. Each cluster also handles the I/O requirements of the system.

Programming Environment

The ANUPAM parallel programming environment is simple and straightforward. The master processor runs UNIX SVR4 operating system. The user writes his parallel program in FORTRAN 77 or ANSI C language. A FORTRAN vectorizer is provided. To aid program development and debugging, a profiler and a debugger are provided. The parallel simulator (PSIM) which runs on any UNIX machine provides a simulated parallel environment of ANUPAM and helps users in parallelizing and debugging algorithms without the need for a target machine.

The parallel processing paradigm makes use of Hoare's Communicating Sequential Processes CSP model. In this model, the underlying computing system is a collection of concurrently executing sequential processes on

multiple processors, communicating with each other via explicit messages. Each processor has its own code, data and stack.

The message passing library consists of send/receive calls to facilitate the writing of parallel codes by calling the library routines either in Fortran or C language. Apart from this, Program Scheduler, Batch que manager and parallel library for scientific and image processing routines are a part of ANUPAM software environment.

The graphics and visualization support is through standard graphics workstation like SGI connected on ethernet. The image processing software libraries developed at Computer Division of BARC make use of client/server model assembly ANUPAM as a server for image analysis and enhancement and SGI workstation as a client for image display.

Applications

The ANUPAM parallel computer was installed in the BARC Computer Centre, the configuration was scaled from 8 to 32 nodes along with hardware enhancements in the later models. The parallel computing facility has been used extensively by BARC scientists and engineers as well as by other users from outside Institutions. More than 50 applications have been parallelized and are being used by users. Studies have also been carried out on algorithms employing data domain decomposition, algorithmic parallelisation, geometric parallelisation, event parallelisation, etc.

For standard Linpack benchmark with 1000 x 1000 matrix size in double precision calculations, ANUPAM 8 gave 52 Mega Floating Point operations per second (MFlops) employing 8 nodes with 80% efficiency. Matrix multiplication with 1000 x 1000 matrix size obtained 106 MFlops performance on 8 node ANUPAM system with 88% efficiency. A molecular dynamic program to calculate electronic structures using linear approximations in BAND theory demonstrated efficiencies of the order of 81% on 16 node configuration. A program called PROLSQ to calculate Protein structure refinement by the method of least square fit, runs on ANUPAM 8 giving 7.45 times speed as compared to a single node. The computational problems in multidisciplinary aerospace science are also running successfully on ANUPAM 8, ANUPAM 16 and ANUPAM 24 systems displaying better efficiencies. The VASBI (Viscous Analysis of Symmetric Bifurcated Intake) code used for calculating airflow around an aircraft exhibited 5.4 times higher speed than IBM RS 6000/560 using 24 nodes of ANUPAM (measuring over 160 MFlops).

Future

It is understood that the ANUPAM system is evolving with i860 XP processor with per node memory extended to 128 and 256 MB. Intercluster bus is being expanded to fast and wide SCSIs. An X-Y router chip is planned. The present slow speed ethernet link to graphics workstation will be replaced by a highspeed graphics adapter sitting directly on a Multibus II. The graphics adapter will make use of i860/XP graphics processor and will provide TCP/IP support over Multibus II, thus maintaining compatibility with the existing graphics interface.

ANURAG's PACE

Background and Evolution

The Advanced Numerical Research and Analysis group (ANURAG) located in Hyderabad is a recently created Laboratory of the Defence Research Development Organization (DRDO) focussed on the R & D in parallel computing and VLSIs and applications of High Performance Computing in CFD, medical imaging, and other areas. ANURAG has developed PACE, a

loosely-coupled, message-passing parallel processing system. PACE is an acronym for Processor for Aerodynamic Computations and Evaluation. ANURAG's PACE programme began in August 1988. As its name suggests, PACE was originally designed to cater to the needs of Computational Fluid Dynamics (CFD) requirements of aircraft design. However, since CFD essentially involves the solution of partial differential equations, PACE is also targeted at scientific computations.

The initial prototypes of PACE were based on the Motorola MC 68020 processor. The hardware was supplied by ECIL. A 4 node prototype based on the MC 68020 processor (working at 16.67 MHz) was first established. This used the VME bus for communication. The VME backplane is `natural' to the Motorola family of processor and was found to provide the necessary bandwidth and operational flexibility. Later, an 8-node prototype based on MC 68030 processor (working at 25 MHz) was developed. This 8-node Cluster forms the backbone of the PACE architecture. The 128 node prototype is based on the MC 68030 processor working at 33 MHz. In order to enhance the floating point speed, ANURAG has developed its own custom floating point processor, ANUCO. The processor board has been specially designed to accomodate the MC 68881, MC 68882 or the ANUCO floating point accelerators.

Subsequently, ANURAG configured a 2-node version of PACE based on the Intel i860. Since the specific CPU boards used were not configured on a standard bus, the communications were established through a common multiported memory. This was basically done as a demonstration to prove that the concept was portable. Recently, the PACE architecture was sported on the SPARC II processor working at 40 MHz. The hardware was procured from Themes/France through a Bangalore-based company called UBM.

System Architecture

PACE consists of a Front-End-Processor (FEP) which is connected to 4 Super-Clusters by means of VME-to-VME communication links. These VME-to-VME links provide high speed parallel (32 bit wide) communications between the two VME backplanes. The Super-Clusters are completely connected to VME-to-VME links. Each Super-Cluster has two CPUs exclusively devoted to communications. One CPU handles intra-Super-Cluster messages while the other handles inter-Super-Cluster messages.

Each Super-Cluster has four Clusters connected to it. Each Cluster has 8 CPUs connected on a VME backplane. The Clusters are linked to the Super-Clusters by VME-to-VME links. Thus each Super-Cluster has 32 CPUs and 4 Super-Clusters can accommodate 128 CPUs. The CPUs within the Cluster are completely connected over the VME bus. They communicate with each other by directly writing the messages into the appropriate buffer space over the VME bus. For communication across clusters within the same supercluster, the CPUs within a cluster pass the message on to the Super-Cluster which then passes on the message to the node in the destination cluster. Communication between nodes in different superclusters, takes three hops which involves two superclusters, the source cluster and the destination cluster.

The latest offering of PACE is called PACE+ and is based on the HyperSPARC node running at 66 MHz. The memory per node is expandable upto 256 MB. The basic hardware again comes from Themes, France.

Programming Environment

The programming environment of PACE is simple and straightforward. The user interacts with the frontend processor (FEP) which is a standard UNIX engine with HyperSPARC processor running Solaris. The parallel processor is treated as a resource of the Front-End-Processor. The user writes his program in a sequential fashion (this is called the `host' program). All

computationally intensive portions of the programs are written as subroutines which are executed in parallel on the parallel processor. The user therefore needs to parallelize only the computationally intensive parts of the program which are treated as subroutines (called the `node' program) to be called by the host program.

In order to enable the user to create, debug and execute his programs, ANURAG has written a parallel programming environment called ANUPAM (ANURAG's Parallel Applications Manager). ANUPAM runs under UNIX and consists of several modules and utilities. These include Preprocessor, Simulator, Queue Manager, Run-Time Libraries, Communications Debugger, Source-Level Debugger, Parallel Library and other utilities. The ANUPAM software only depends on the availability of UNIX at the front-end. The software is portable across machines with very few modifications (the modifications relate to the physical addresses of the CPU boards).

Applications

Several application programs have been run on the various models of PACE. These include Linpack, FFT, neural networks simulation, FEM codes and several CFD codes. The PACE 128 system based on the Motorola 68030 processor and MC 68882 co-processor delivered over 30 MFlops speed for large problems. The speed per processor node was 0.33 MFlops. Later, this was enhanced to 0.75 MFlops per node incorporating ANURAG's custom FPU ANUCO. With the SPARC II processors, the speed is 4.5 MFlops per node. The latest SPARC processors will offer higher performance.

NAL's FLOSOLVER

Background and Evolution

The National Aerospace Laboratories (NAL) located at Bangalore is a major national laboratory of the Council for Scientific & Industrial Research of the Govt. of India. In 1986, NAL started the project to design, develop and fabricate suitable parallel processing systems to solve fluid dynamical and aerodynamical problems. The project was motivated by the need for a powerful computer in the laboratory and was influenced by similar developments internationally. The existing UNIVAC system then with NAL could not meet the computational requirements of NAL.

The parallel computer of NAL is called Flosolver which was the first Indian parallel computer which became operational in 1986. Since then, a series of parallel computers have been built in NAL, which include Flosolver Mk1 and Mk1A, which were four processor systems based on 16-bit Intel 8086/8087 processors, Flosolver Mk1B, an eight processor system in this series, Flosolver Mk2, based on 32-bit Intel 80386/80387 processors and the latest version, Flosolver Mk3, based on the RISC processor i860 from Intel.

System Architecture and Programming Environment

The present version of Flosolver Mk3 is based on eight i860 RISC processors with an on board memory of 64 MB per processor. The system bus is Multibus II and has a bandwidth of 40 MB/sec. The communication between the processors is assisted through the message passing co-processor (MPC) and high speed direct memory access (DMA) controllers available on each of the boards.

Thus, it can be seen that the system architecture of NAL's Flosolver is very similar to the architecture of BARC's ANUPAM single cluster. The major software development carried out for Flosolver includes a simulator which can run on any UNIX machine which can be used as a front-end. The parallel FORTRAN pre-processor has also been developed. The concurrent

executive, a varient of Intel's iMRX, runs on each node. The standard C and FORTRAN compilers are supported. A vectorizer is also available.

Applications

The application of NAL's Flosolver is dominantly focussed on the weather forecasting code T80 under a project from the Department of Science and Technology of the Govt. of India. The Flosolver has also been extensively used by the scientists of NAL for solving their computational fluid dynamics problems.

With the sustained speed of over 15 MFlops, NAL scientists have actively used Flosolver for CFD problems. A panel code to compute aerodynamic coefficients on an entire aircraft, with over 6000 panels, has been completed in less than an hour. Direct numerical simulation of the initial evolution of a turbulent axisymmetric wake and a 3D Euler computation of flow past a wing have been done in this machine. In one of the latest applications, Global Circulation Model code for weather forecasting running on Cray has been sucessfully parallelised on Flosolver.

C-DOT's CHIPPS

Background and Evolution

The Centre for Development of Telemaatics (C-DOT) was launched by the Government of India as a mission project to develop indigenous Digital Switching technology. C-DOT completed its first Mission in 1989 by delivering technologies of RAX for Rural Exchanges, and MACS for secondary switching areas. In Febraury 1988, a development contract was signed by the Department of Science and Technology and C-DOT under which C-DOT was to design and build a 640 MFlops 1000 MIPS peak parallel computer. C-DOT set a target of 200 MFlops for sustained performance.

System Architecture

The CHIPPS, C-DOT's High Performance Parallel Processing System is based on the single algorithm multiple data architecture. Drawing advantages of both SIMD and MIMD, the architecture provides coarse grain parallelism with barrier synchronization. It also provides uniform start-up and simultaneous data distribution across all configurations and it is realized using off-the-shelf hardware and software technology tools. The CHIPPS is designed to support large, medium and small applications. The range includes a large 192-node machine, a 64-node machine and a compact 16-node machine.

A flexible Interconnection Network (ICN) configured by the Main Controller (MC) interconnects Processing Elements (PEs) and global memory (MDM) banks. While raw data for computation is written by the MC into the memory mapped banks, the program is broadcast to the PEs through high speed serial links. The computed results are transferred simultaneously to the banks which are then stored onto the disks. The MC broadcasts the commands to the PEs and MDMs and synchronizes all operations at task levels.

Programming Environment

The CHIPPS provides Data Parallelism as its program paradigm. Application programs written for sequential machines can be easily parallelised by identifying the tasks which are data-independent. The application program is functionally divided as tasks viz, I/O, data management, computation. The computation is performed at PE and the control and I/O at MC.

The user interface is provided through system library to perform tasks like data transfer, switch setup, execution control, etc. The system library, a collection of high level language callable routines coordinates with the kernels at the PE and the MDMs and thus helps the user to parallelize the application program in the UNIX environment at the MC. A debugger at the MC which allows setting of break and watch points, viewing the values of selected variables etc. is an additional feature for the application program development.

Applications

The CHIPPS was originally designed primarily for weather forecasting and radio astronomy applications. Recently, several scientific kernel codes were ported on the machine to demonstrate its general applicability in scientific and engineering applications.

Relative Performance of Indian Parallel Computers

In terms of speed, the Indian high performance computers may not be comparable to the best machines such as IBM SP-2 and SG Challenger. But the Indian machines may turn out cheaper compared to machines with lower power, for a similar performance range. For example, CDAC's PARAM 9000/SS model with Super Sparc has a peak performance of 0.96 GFlops for a 16 node system, whereas SGI Power Challenge has a peak performance of 5.76 GFlops for 16 processors and IBM Power 2 model 590 has a peak performance of 4.22 GFlops for 16 processors. The future generation teraflop machines of CDAC based on DEC Alpha processors are supposed to be a match to the performance of the best parallel machines available. The projected performance of the future CDAC high performance computers is shown in fig 2.

More details on the applications and architectures of the machines discussed above may be found in [8].

A Glimpse of Certain Indian Research Efforts

The five Indian Institutes of Technology (IITs) located at Bombay, Madras, Kharagpur. Delhi, and Kanpur, the Indian Institute of Science (IISc) located at Bangalore, the Centre for Mathematical Modeling and Computer Simulation (CMMACS) located at Bangalore and a number of other organizations have made several significant research contributions to high performance computing. Notable among them are, the novel architecture based on hierarchical network of hypercubes [1], parallelizing diverse applications in the areas of medical imaging [2], VLSI layout [3], logic programming [4], scientific computation [5], multiprocessor operating systems [6], and interconnection networks [7]. There have been several success stories of building prototype/experimental parallel machines in some of the above academic institutions.

Support for Indian High Performance Computing Activities

Typical high performance computing facilities in India include Cyber 992, CD4360, R6000/580 cluster, DEC 10000/620, CDAC's PARAM, and Power Challenge at the Indian Institute of Science, Bangalore; Convex C220 at the Indian Institute of Technology, Kanpur; CDAC's PARAM at the Indian Institute of Technology, New Delhi; Cyber 180 at the Indian Institutes of Technology, Bombay and Kharagpur; and Convex 3820 at the Centre for Mathematical Modelling and Computer Simulation, Bangalore, and CRAY XMP/216 at the National Centre for Medium Range Weather Forecasting (NCMRWF) at New Delhi. Most of these computing systems have support of

PVM, and MPI for carrying out extensive studies on parallel algorithms/architectures. Most of the parallel software development is carried out using Parallel C, Parallel Fortran and Occam on transputer-based systems.

At present, the major funding for high performance computing activities in India comes from the government sources such as the Department of Electronics, Department of Science and Technology, Department of Scientific and Industrial Research, in the form of sponsored research projects. Realising the significance of this activity, the Department of Electronics of the government of India took all initiatives to set up the Centre for Development of Advanced Computing (CDAC), in a mission mode. Several Indian and multinational firms such as Tata Elxsi (India) Ltd., Tata Information Systems Limited, Motorola India Electronics Pvt. Ltd., Digital Equipment (India) Ltd., Silicon Graphics, have initiated strong marketing and software development activities in India in recent months. Diverse applications drawn from scientific computing particularly in the areas of signal processing, transaction processing, multimedia, virtual reality, and simulation are of interest to several of these private industries. The Indian software houses at present don't significantly contribute to parallel software development efforts, but the scenario may change in future. The government-funded organizations primarily look at applications in the areas of weather modelling, aerodynamic simulation, finite element analysis. Some other applications are in the areas of analysis of aerospace, automotive, and offshore structures; molecular modelling for drug, pesticide, and biotechnology, weather prediction, pollution studies, and turbulence. Academic Institutions such as the Indian Institutes of Technology (IITs), Indian Institute of Science (IISc), Birla Institute of Technology and Science (BITS), and several Regional and University Engineering colleges offer courses at the introductory and advanced level in the area of Parallel Processing. Most of these Institutions have excellent research programs in the areas of parallel architecture, parallel algorithms, compilers and operating systems for parallel computers and mapping diverse aplications to parallel machines. Future applications will broadly address the above areas, with a possible emphasis by private industries on commercial exploitation of high performance computing. The future government funding in this area may not increase significantly because the present emphasis is more on industrial collaborative efforts. Industrial participation and support particularly from the multinational firms is anticipated for an active promotion of this area. The general feeling is that most of the present applications are of interest to academic institutions and government industries. Unless sufficient applications of high performance computing are demonstrated in terms of their commercial viability, the future funding to support this important activity may be adversely affected in India.

A Brief Overview of High Performance Computing Activities in Other Countries from the Far-East Region

The National Supercomputing Research Centre (NSRC) is the focus of High Performance Computing activities in Singapore. The emphasis in Singapore is to use high performance computing to promote innovation and industrial development in Singapore, and help its industry adopt advanced computing technologies, rather than for more basic research. NEC SX-3 and IBM SP2 are the major facilities at the NSRC. The NSRC plans to procure a CRAY T90 by the end of 1995. Research focus ranges from traditional scientific areas to defence science, weather prediction and commercial applications. The NSRC in Singapore is funded by the National Science and Technology Board. Silicon Graphics and NSRC signed a memorandum in August 94 to promote visualization technology, by establishing a Supercomputing Visualization Laboratory at the NSRC. The result of this memorandum is the supply of a pair of R8000 CPUs, two Indy workstations, an Onyx workstation, and a variety of software by SGI. Other than this, there is a new program in Computational Science at the National University of Singapore. Work is also carried carried out at the Institute of Systems

Science of the National University of Singapore, and Nanyang Technological University, Singapore. Some industrial collaborative applications being investigated in Singapore are molecular simulation, scientific visualization, integrated land use and transportation modelling, financial modelling, and forecasting. The National University of Singapore currently operates a Cray J916 Supercomputer in addition to a few SGI and Convex machines. The Nanyang Technological University has an IBM SP2. The National Centre for High Performance Computing (NCHC) in Taiwan is equipped with IBM ES-9000/860 (5 CPU), Convex C3840 (4 CPU), IBM SP1 (16 nodes), Convex SPP (8 nodes), Convex Meta (8 nodes), and handles projects in the areas of scientific databases, and network applications. The activities are primarily supported by academic institutions, and to some extent by industry and NCHC. The Tsinghua University in China emphasizes on climate modelling, ocean circulation, turbulent flow, vision, and cognition as the main applications of high performance computing. In China, Galaxy 1 and IBM 3084 (2 processors) are being used for seismic processing at CNPC; Fujitsu M150, Galaxy 2 (4 processors), CRAY YMP (2 processors) for weather prediction at National Weather Bureau (Beijing); and IBM ES 9000 for financial applications at Commercial Bank (Sanghai). Other than universities, the Chineese government plans to promote this activity through projects such as National Natural Science Foundation.

Some of the high performance computers used in Australia are, CRAY YMP EL (Moldflow and BHP in Sydney), Convex 300 (Biochemical Research, Melbourne), CRAY YMP 464 (CSIRO, Melbourne), Intel Paragon and iPSC/860 (University of Melbourne), Fujitsu VP2000 and CM-5 (Australian National University). These systems are being used in the areas of image recognition and analysis, molecular modelling, visualization, climate research, polymers, tomography, geophysical and chemical applications. The support for such activities in Australia is through Universities and research schemes such as ARC and CRC (Collaboration Research Centres). The Korean high performance computing systems under use are primarily the CRAY Y-MP systems, and application software in the areas of weather forecasting, computational fluid dynamics, computational chemistry, genetic engineering, and nuclear power plant safety analysis, have been developed. Some interesting research and development projects in Korea are the KAICUBE project developed by KAIST (Korean Advanced Institute of Science and Technology) in 1993 and the TICOM IV Project sponsored by Ministry of Communication and Science and Technology. An 8 CPU (i860) machine KAICUBE - 860/8 with a 320 MFLOPS rating is the target of the KAICUBE Project whereas the TICOM IV Project (1994 - 1998) aims at 256 processors (P6) with 20 GIPS performance, 1 GB/node memory for large online transaction processing applications. In New Zealand, DEC Alpha 2000/300 AXP (University of Waikato), Silicon Graphics Iris Indigo Workstations and Sparc 10 dual processor (Victoria University), DEC Station 5000/200 (Lincoln University) and other high performance computers are being used for application software development in the areas of finite element analysis, image processing, seismology, combustion modelling, weather forecasting, and computational chemistry applications.

In Hong Kong, its University of Science and Technology is equipped with an Intel Paragon with 140 processors (75 Mflops peak performance per processor) and 5 GB memory and 35 GB parallel disk array; and an SGI Onyx parallel processor with 8 processor system (MIPS R4400 chip), 512 MB memory (shared), 10 GB disk memory. The Chineese University of Hong Kong has a DEC mpp 12000 massively parallel computer with 8192 processors, with Maspar parallel C, parallel FORTRAN, and parallel math library. The computing facilities at the Hong Kong University and City Polytechnic of Hong Kong are configured around an IBM SP1 with 8 processors and a CRAY YMP8/86A with 8 processors respectively. Some of the interesting applications studied in the various Universities of Hong Kong include, convection induced turbulence simulation, parallelizing large linear programming problems, application of two level finite element method, Monte Carlo and molecular dynamics studies, and parallel implementation of neural networks. Most of the high performance activity in Hong Kong is supported in Universities.

Some Future Predictions

One of the toughest things is to predict the future, more so for high performance computing. In future, the Network Of Workstations (NOW) will be an attractive alternative to parallel machines as far as high performance computing solutions are concerned. Such an alternative will be more realistic in future with fast advances in communications technology, particularly in the areas of high speed switching and ATM networks. But the interest in massively parallel machines and heterogeneous computers wil continue for teraflop range applicatioms. The support for such highly specialised applications may have to come through government sources since such applications will be from the research community, rather than being from the industry. The trend may be to network several parallel machines (even of different architecture) through high speed networks. Future parallel systems will be networks of heterogeneous computers comprising some of the following: workstations, PCs, shared-memory multiprocessors, and special-purpose machines. There will be greater integration of parallel computation, high-performance networking, and multimedia systems. The popular parallel programming languages will be based on C, C++ or some object-oriented paradigm.

Standard parallel languages must be developed if parallel computers are to achieve the same level of popularity as that of sequential machines. Future compilers or operating systems will take charge of distributing parallelism onto different processors and also of exploiting levels of parallelism in an application. There will be more attempts to run databases like Oracle, Informix, on parallel machines. Other than business applications, recreation too will drive parallel computing.

Some "functional" version of C and Fortran will facilitate parallelizing compilers. Tools will also play a very important role in developing parallel applications and making them portable across different architectures. Coarser grained objects than we have today may emerge. Multiple objects will work in parallel to solve a problem. These objects will use the parallel processing constructs.

In the next decade or so, machines in the commercial arena will be shared-memory multiprocessors. Hopefully, we will have multiple nodes with independent memory virtualized as a shared-memory system by the operating system. A multiparadigm, portable, standard substrate is essential if parallel computers are ever to flourish.

The massively parallel machines will continue to be used for the same type of applications broadly classified under the HPCC (High Performance Computation and Communication) project in the USA. However, there will be more interest in the use of High Performance Computing techniques for commercial applications, e.g. on-line transaction processing, multimedia, etc. High speed networking will enable users to access teraflop machines across continents. Applications involving image processing, virtual reality, and simulation with possible emphasis on defense strategies will assume more significance. Data-intensive business applications such as videoconferencing, advanced graphics, multimedia, will take advantage of parallelism. To sum up, high performance computing may not be projected to be a hype any more!

Acknowledgements

Dr. Vijay P. Bhatkar, Executive Director, Centre for Development of Advanced Computing, Pune, India is profusely thanked for providing details of C-DAC's parallel machines through his technical report. Dr. K. Neelakantan, ANURAG, Mr. M. Periasamy, C-DOT, and Dr. U. N. Sinha, NAL, are also acknowledged for providing details about the parallel machines in their organisations.

References

[1] J. Mohan Kumar and L.M. Patnaik, Extended Hypercube: A Hierarchical Interconnection Network of Hypercubes, IEEE Transactions on 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 Transactions on Nuclear Science, Vol.41, No.5, October 1994 pp.1721-1728.

[3] B.B. Prahlad Rao, L.M. Patnaik, and R.C. Hansdah, Parallel Genetic Algorithm for Channel Routing, Proceedings of the Third Great Lakes Symposium on VLSI, Kalamazoo, Michigan, USA, March 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 Transactions on Computers, Vol.39, No.5, May 1990, pp.653-665.

[6] Pradeep Hatkanagalekar, Concurrency Issues in Shared-Memory Multiprocessor Operating System Kernels, Ph.D Thesis, Department of Computer Science and Engineering, Indian Institute of Technology, Bombay, 1995.

[7] Ravi Mittal, B. N. Jain, R. K. Patney, "Link Augmented Binary (LAB) -Tree Architecture, IEE Proceeedings E: Computer and Digital Techniques, Vol. 140, No.2, March 1993, pp. 127-133.

[8] Vijay P.Bhatkar, Parallel Computers Galore in India, Technical Report, Centre for Development of Advanced Computing, Pune, India.