ISSN 0016-2523

FUJITSU

SCIENTIFIC & TECHNICAL JOURNAL

Summer 1991 VOL. 27, NO. 2

Special Issue on Supercomputer VP2000 Series

FUJITSU Sci. Tech. J., **27**, 2, pp. 147-232, Kawasaki, Summer, 1991

The Issue's Cover :
Light by Hirofumi OHKUMA
From nothing, an explosion of immense energy brings forth the universe.
This idea is our continual inspiration.

FUJITSU Scientific & Technical Journal is published quarterly by FUJITSU LIMITED of Japan to report the results of research conducted by FUJITSU LIMITED, FUJITSU LABORATORIES LTD., and their associated companies in communications, electronics, and related fields. It is the publisher's intent that FSTJ will promote the international exchange of such information, and we encourage the distribution of FSTJ on an exchange basis. All correspondence concerning the exchange of periodicals should be addressed to the editor.

FSTJ can be purchased through KINOKUNIYA COMPANY LTD., 3-17-7 Shinjuku, Shinjuku-ku, Tokyo 160-91, (Telephone : +81-3-3439-0162, Facsimile : +81-3-3706-7479).

The price is US$7.00 per copy, excluding postage.

FUJITSU LIMITED reserves all rights concerning the republication and publication after translation into other languages of articles appearing herein.

Permission to publish these articles may be obtained by contacting the editor.

FUJITSU LIMITED Tadashi Sekizawa, President
FUJITSU LABORATORIES LTD. Masaka Ogi, President

Editorial Board
Editor Takahiko Misugi
Associated Editors Shigeru Sato Hideo Takahashi
Editorial Representatives

Tadashi Hasegawa	Yoshimasa Miura	Makoto Mukai
Yasushi Nakajima	Hajime Nonogaki	Seiya Ogawa
Shinji Ohkawa	Shinya Okuda	Hirofumi Okuyama
Tohru Sato	Yoshio Tago	Shozo Taguchi
Jyun'ichi Tanahashi	Hirobumi Takanashi	Mitsuhiko Toda
Toru Tsuda	Itsuo Umebu	Yutaka Yamaoka
Akira Yoshida		

Editorial Coordinator
Yukichi Iwasaki

FUJITSU LIMITED 1015 Kamikodanaka, Nakahara-ku,
Kawasaki 211, Japan
Cable Address : FUJITSULIMITED KAWASAKI
Telephone : +81-44-777-1111
Facsimile : +81-44-754-3562

Printed by MIZUNO PRITECH Co., Ltd. in Japan
© 1991 FUJITSU LIMITED (June 20, 1991)

FUJITSU

SCIENTIFIC & TECHNICAL JOURNAL

Summer 1991 VOL.27, NO.2

Special Issue on Supercomputer VP2000 Series

CONTENTS

Featuring Papers

147 **Preface** ● Toshio Hiraguri

149 **System Overview of FUJITSU VP2000 Series**
● Nobuo Uchida ● Yuji Oinaga ● Hiroshi Tamura
● Kazuyuki Shimizu

158 **Hardware Technology for FUJITSU VP2000 Series**
● Akira Kaneko ● Kiyoshi Kuwabara ● Shun-ichi Kikuchi
● Takanobu Kano

169 **Semiconductor Devices for FUJITSU VP2000 Series**
● Ken-ichi Ohno ● Kazuo Ooami ● Hideo Itoh
● Katsuhiko Suyama

179 **Multilayer Glass-Ceramic-Composite Circuit Board for FUJITSU VP2000 Series**
● Koichi Niwa ● Yukichi Takeda

187 **Design Automation System for FUJITSU VP2000 Series**
● Hirofumi Hamamura ● Akihiko Hanafusa ● Minoru Saitoh
● Toshihiko Tada

197 **Basic Software for FUJITSU VP2000 Series**
● Koh-Ichiro Hotta ● Takashi Kunai ● Yoshio Honma

211 **Atomic-Scale Simulations for Semiconductors by Supercomputer**
● Minoru Ikeda ● Kumiko Furuya ● Takahiro Yamasaki
● Masuhiro Mikami

222 **Computational Fluid Dynamics and Computers**
● Satoru Ogawa ● Yoko Takakura

UDC 681.32
FUJITSU Sci. Tech. J., **27**, 2, pp. 149-157(1991)

System Overview of FUJITSU VP2000 Series

• Nobuo Uchida • Yuji Oinaga • Hiroshi Tamura
• Kazuyuki Shimizu

The demand for high performance supercomputing has been growing especially in the last several years. To meet this demand, Fujitsu has developed the FUJITSU VP2000 series pipelined supercomputers. This series has a maximum performance of 5 GFLOPS for a single processor. This paper describes the features, system configuration, and functional outline of the VP2000 series.

UDC 658.512.2:681.32
FUJITSU Sci. Tech. J., **27**, 2, pp. 187-196(1991)

Design Automation System for FUJITSU VP2000 Series

• Hirofumi Hamamura • Akihiko Hanafusa • Minoru Saitoh
• Toshihiko Tada

The performance of the FUJITSU VP2000 series computers is high because they incorporate advanced technologies; for example, ECL 15k-gate very-large-scale integrated circuits and multilayer glass-ceramic-composite circuit boards. The VP2000 series was developed using a design automation system (DA) that enabled the ultrahigh-speed technology used in the series to be fully exploited. This paper mainly describes the features of this DA system.

UDC 681.32
FUJITSU Sci. Tech. J., **27**, 2, pp. 158-168(1991)

Hardware Technology for FUJITSU VP2000 Series

• Akira Kaneko • Kiyoshi Kuwabara • Shun-ichi Kikuchi
• Takanobu Kano

Fujitsu has developed new packaging, cooling, and power supply technology for new high-speed and high-density LSIs. This paper introduces an ultra-miniature LSI package, a ceramic board on which two million gates can be mounted, and high-performance cooling and water-cooled power supply technology.

UDC 681.32.06
FUJITSU Sci. Tech. J., **27**, 2, pp. 197-210(1991)

Basic Software for FUJITSU VP2000 Series

• Koh-Ichiro Hotta • Takashi Kunai • Yoshio Honma

Since the first shipment of the FUJITSU VP-series, Fujitsu's policy regarding system software for supercomputers has been to supply systems that are as easy to use as a general purpose computer. Since then, the size of application programs has increased more and more, and progress made in computer networks has significantly changed the supercomputer environment. When the FUJITSU VP2000 series was released, new system software for the current environment was also developed. In this paper, the new system software is outlined.

UDC 621.3.049.774:681.32
FUJITSU Sci. Tech. J., **27**, 2, pp. 169-178(1991)

Semiconductor Devices for FUJITSU VP2000 Series

• Ken-ichi Ohno • Kazuo Ooami • Hideo Itoh
• Katsuhiko Suyama

Advanced silicon and GaAs technologies have been developed and used in the high-speed, high-density LSIs of the FUJITSU VP2000 series. The main LSIs are 15 000-gate ECL array with a 70 ps propagation delay, 3 500-gate ECL array with 64-Kbit STRAM and a 1.6 ns maximum access time, 1-Mbit static RAM with a maximum access time of 35 ns, and 1 200-gate GaAs line driver/receiver with a 60 ps propagation delay time.

UDC 548.5:681.32
FUJITSU Sci. Tech. J., **27**, 2, pp. 211-221(1991)

Atomic-Scale Simulations for Semiconductors by Supercomputer

• Minoru Ikeda • Kumiko Furuya • Takahiro Yamasaki
• Masuhiko Mikami

The electronic structure and crystal growth of semiconductor materials were studied using computer simulations. First, the energy bands and transition probability of Si-Ge superlattices were investigated. It was found that direct transition can be realized in $(Si)_5/(Ge)_5$ on a Ge substrate by zone folding and by breaking the inversion symmetry. Then, crystal growth by Molecular Beam Epitaxy was simulated for Lennard-Jones systems by applying molecular dynamics. The lattice mismatch produced various growth patterns, which agree well with experiments on metallic hetero-structures. It was found that Si adatoms on a reconstructed Si(100) surface diffuse anisotropically and make new stable dimers when they meet.

UDC 621.3.049:681.32
FUJITSU Sci. Tech. J., **27**, 2, pp. 179-186(1991)

Multilayer Glass-Ceramic-Composite Circuit Board for FUJITSU VP2000 Series

• Koichi Niwa • Yukichi Takeda

A new glass-ceramic-composite material has been developed and applied to the circuit board of the FUJITSU VP2000 series supercomputer. The composite material exhibits low dielectric constant and low thermal expansion coefficient. These two characteristics are capable of satisfying both present and future requirements of high-speed signal propagation and high-density packaging of LSIs. Copper conductors are used for the circuit wiring, which yields low resistivity conductor wiring for as many as 61 layers. The gate density of the VP2000 series is ten times higher than that of conventional organic printed circuit boards.

UDC 532.5.01:681.32
FUJITSU Sci. Tech. J., **27**, 2, pp. 222-232(1991)

Computational Fluid Dynamics and Computers

• Satoru Ogawa • Yoko Takakura

This paper describes the participation of computational fluid dynamics (CFD) and computers. First, the history and outline of CFD are briefly explained. Next, advanced researches (computations of transonic flow with large separation, supersonic flow around complex configulations, supersonic flow with combustion, and hypersonic flow of real gas) in Computation laboratory of National Aerospace Laboratory are shown. Finally, turbulence and the future problems of CFD are described as impacted by computer performance.

Preface

Special Issue on Supercomputer VP2000 Series

● Toshio Hiraguri
Managing Director, Computer Systems Group

Initially, supercomputers had only limited use in universities and national research centers, since then however, they have rapidly gained acceptance in a wide variety of industry applications. In universities, research centers and industries, research and development is being conducted making full use of leading edge technology. For the analysis of phenomena which is microscopic, momentary or heretofore unknown, the supercomputer plays an indispensable role. The challenge is to produce the highest speed computational machines, as well as a system environment that integrates supercomputers with workstations and personal computers to obtain the greatest benefit.

Fujitsu, a pioneer in developing supercomputers in Japan, shipped the FACOM 230-75 array processor in 1977. Based on the development experience gained with the predecessor, the FUJITSU VP-100 and VP-200 series were developed and shipped in 1983. Since then, the VP-30, VP-50, VP-400 and later the VP-E series have followed. Now Fujitsu offers the VP2000 series of supercomputers. Announced at the end of December 1988, the first shipments started in March 1990. About 30 of the VP2000 series have been installed around the world and over 120 systems are currently operating within the entire VP range. This success results from the high effective performance of application programs, compatibility with the FUJITSU M-series general-purpose computers and user-friendly systems. The new series incorporates these features and provides added power and functionality for today's research and development.

In this special issue, the development philosophy and features of the FUJITSU VP2000 series are described.

Super high-speed processing

The design target was to greatly improve not only the maximum performance but also the effective performance, as well as system expansibility. A performance of 5 GFLOPS, the highest order in the world for a single processor, has been achieved. Dual scalar processor (DSP) and quadruple scalar processor (QSP) multiprocessor architecture has also been implemented. The additional scalar unit significantly increases system performance, thereby dramatically improving the cost performance ratio and added flexibility to the system configuration. By using a new storage unit (system storage unit), the processing time of applications that have I/O intensive functions can be reduced, and the number of active TSS terminals using vector processing functions can be increased.

Adaption to the UNIX[Note)] environment

Workstations and open systems based on UNIX are spreading in research and development centers. To meet this requirement, Fujitsu's mainframe UNIX operating system, UXP/M, has a vector processor support option. All the operations for vector processing, from program development to high-speed vector execution, are

supported in a UNIX environment. MSP/EX, Fujitsu's proprietary operating system has also been enhanced for UNIX compatibility by incorporating UNIX functions into MSP/EX and can be run with UXP/M under advanced virtual machine (AVM).

Enhancement to language processing system

A high level optimizing function for the compiler is required to achieve a high effective performance for program execution. The new FORTRAN compiler has improved vectorizing and optimizing capability to achieve this. The compiler also provides a sophisticated parallel processing facility to reduce elapsed time of execution of programs when run in a DSP or QSP environment. Fujitsu also provides an enhanced interactive tuning tool to improve performance of applications.

This special issue also discusses new technologies introduced with the VP2000 series. The last part of the issue contains two papers on supercomputer applications. The first paper, *Atomic-Scale Simulations for Semiconductors by Supercomputer*, discusses the use of supercomputers in the study of semiconductor materials. The second paper, *Computational Fluid Dynamics and Computers*, describes the application of supercomputers to computational fluid dynamics (CFD), which is a method widely used in engineering fields. These are two areas which have greatly benefited through the introduction of supercomputers.

Researchers and developers place a continuous demand for higher and higher levels of performance and Fujitsu is now targeting performance levels in the tera-FLOPS range. Fujitsu will continue to develop leading edge hardware and software technology for massively parallel processing. Fujitsu also supports open systems and is enhancing the user's interface. Furthermore, Fujitsu will substantially increase the body of application software available.

Fujitsu is continually developing supercomputers that can meet the current requirements and future needs of our researchers and engineers.

Note: The UNIX operating system was developed and is licensed by UNIX System Laboratories, Inc.

UDC 681.32

System Overview of FUJITSU VP2000 Series

• Nobuo Uchida • Yuji Oinaga • Hiroshi Tamura • Kazuyuki Shimizu

(Manuscript received December 19, 1990)

The demand for high performance supercomputing has been growing especially in the last several years. To meet this demand, Fujitsu has developed the FUJITSU VP2000 series pipelined supercomputers. This series has a maximum performance of 5 GFLOPS for a single processor. This paper describes the features, system configuration, and functional outline of the VP2000 series.

1. Introduction

The rapidly growing field of technical calculation is boosting the demand for high-performance supercomputers. Many users in various fields require supercomputers that are more efficient and easier to use than conventional systems. This is especially true in the fields of fluid dynamics, image processing, resource exploration, meteorological forecasting, molecular dynamics, and energy analysis.

For more versatile applications, it has become necessary to link UNIX[Note] systems (prevalent in research and development areas), and to create easy-to-use open systems of linked workstations. The new FUJITSU VP2000 series are efficient, easy-to-use supercomputers. They have been developed using the latest technology and the experience gained from developing the preceding VP-series[1].

The VP2000 series consists of the following basic models (in descending order of vector performance): VP2600 (high-end model), VP2400, VP2200, and VP2100 (low-end model). There are three types of scalar processor configurations: the uni-processor (model 10), dual scalar processor (model 20), and quadruple scalar processor (model 40). In total, there are ten models, all of which can be upgraded in the field.

Note: The UNIX operating system was developed and is licensed by UNIX System Laboratories, Inc.

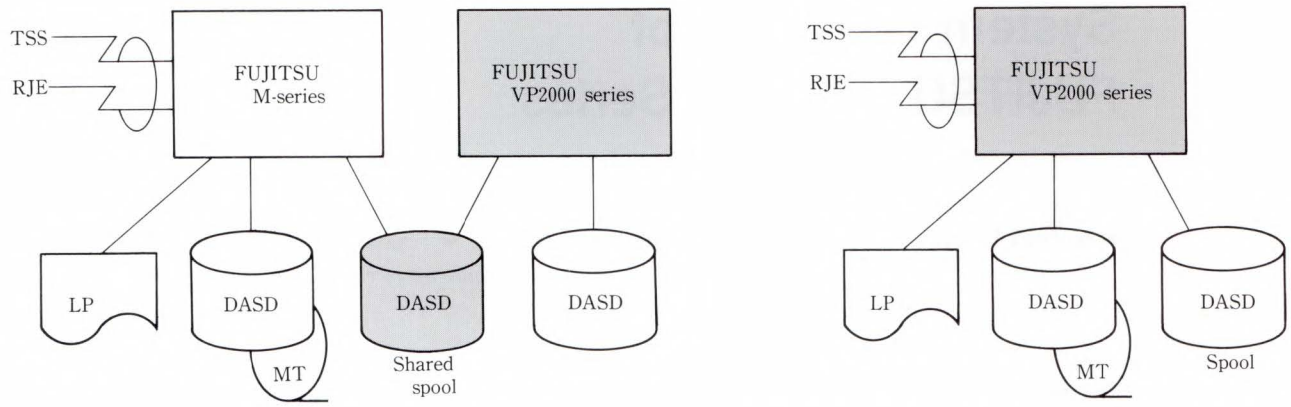
2. Features of VP2000 series

2.1 High-speed operation

The VP2000 series features superb performance based on super-high-speed, high-density technologies and an improved pipeline structure. The instruction set is efficient enough to extend the vector processing range[2]. Therefore, the parallel processing level of vector processing is raised and software can fully exploit the hardware functions. At 5 GFLOPS, the VP2600/10 has the highest performance for a single processor. This figure is about triple the performance of the previous VP-series.

2.2 Flexibility

There are two types of VP2000 series systems that operate under the MSP operating system. These are the stand-alone system and the loosely coupled back-end system (see Fig. 1). In the back-end system, the functions and workloads can be distributed optimally by connecting a front-end processor. This type enables the VP2000 series models to run

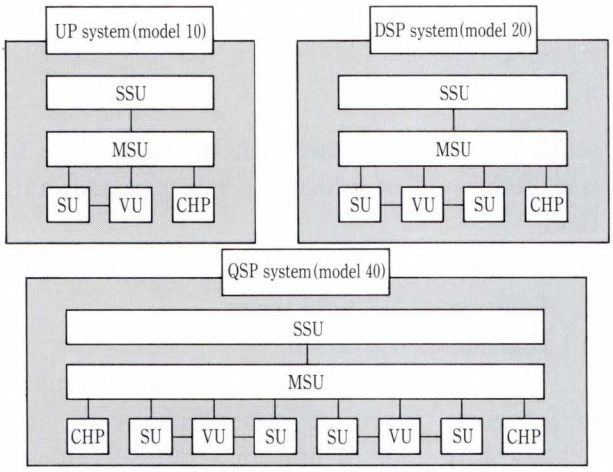


TSS : Time Sharing System
RJE : Remote Job Entry
LP : Line Printer
DASD : Direct Access Storage Device
MT : Magnetic Tape

a) Loosely coupled back-end system

b) Stand-alone system

Fig. 1—Configuration of VP2000 series systems.



SSU : System Storage Unit
MSU : Main Storage Unit
CHP : Channel Processor
SU : Scalar Unit
VU : Vector Unit

Fig. 2—Processor configuration of each system.

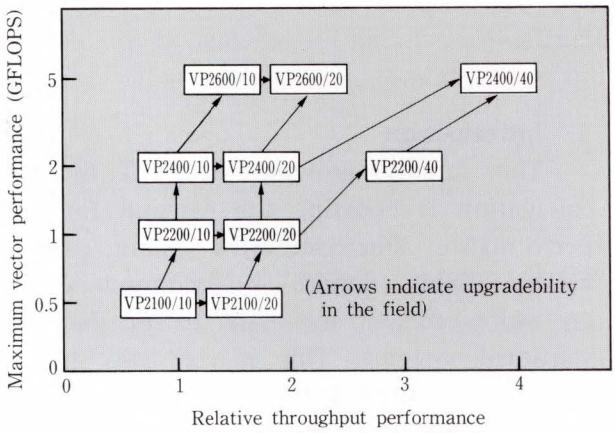


Fig. 3—Overall performance of VP2000 series models.

at their full super-high-speed as processors for high-speed operations. In the stand-alone system, the VP2000 series executes processing, from program development to high-speed execution, in the same way as a general purpose system.

The VP2000 series computers have one, two, or four scalar processors. These configurations are called the uni-processor (UP), dual scalar processor (DSP), and quadruple scalar processor (QSP) configurations respectively. The UP is the basic configuration, and has a scalar unit and a vector unit. The DSP is a new type of multiprocessor system in which two scalar units share a vector unit. The QSP is multiprocessor system having four scalar units and two vector units. These systems can be controlled by software in the same way as an ordinary multiprocessor system without the need to consider the vector unit shared between two scalar units.

The VP2000 series maintains complete upward compatibility with the previous VP-series and consists of ten models. Figure 2 shows the processor configuration of each system. Figure 3 compares the overall performance of the each models.

2.3 High cost-performance

The VP2000 series models are high performance supercomputers that require a small installation space and low operating power.

This series consumes about three-quarters of the power consumed by the previous VP-series uni-processor system. This reduction has been achieved by developing advanced LSI, packaging, and high-performance cooling technologies. These new technologies have nearly tripled the ratio of maximum performance to power consumption. They have also reduced the installation space to about 2/3 that of the previous VP-series. For example, the new scalar unit is contained on a single 24.5 cm square board[3]. The ratio of maximum performance to installation area has been increased 3.5 times.

3. Outline of VP2000 series models

3.1 System configuration

Figure 2 shows the processor configuration of each VP2000 series computer. Figure 3 compares the overall performance of the ten models. As the figure shows, overall performance increases with vector performance and the number of scalar units.

Since the DSP and QSP systems have two or four scalar units, the job load distribution can be optimized. For example, in the DSP system, scalar jobs not requiring the vector unit can be entered on a scalar unit while the other scalar unit uses the vector unit to execute highly vectorized jobs. Or, to ensure a high vector unit operating efficiency, medium-vectorized jobs can be entered on both scalar units in order to share the vector unit between jobs.

3.2 Components

The VP2000 series models consists of the following units:

1) Vector processing unit (VPU)

The VPU can be compared to the central processing unit (CPU) of a general-purpose computer. It consists of a scalar unit (SU) and a vector unit (VU) that contain 15 000-gate ECL LSIs, and 64-Kbit RAM & logic LSIs[4]. The SU fetches scalar and vector instructions, and executes scalar instructions, interrupt processing, and machine check processing. The SU has a 128-Kbyte high-speed buffer

Table 1 VP2000 series specifications

Item \ Model		VP2100/ 10, 20	VP2200/ 10, 20	VP2400/ 10, 20	VP2600/ 10, 20	VP2200/40	VP2400/40
Vector peak performance (GFLOPS)		0.5	1	2	5	2	5
CPU	Number of VUs	1				2	
	Number of SUs	1-2				4	
Main storage capacity (Mbyte)		64, 128 192, 256 384, 512 768, 1 024	128, 256 384, 512 768, 1 024	256, 512 768, 1 024	512, 1 024 1 536, 2 048	256, 512 768, 1 024 1 536, 2 048	512, 1 024 1 536, 2 048
System storage capacity (Gbyte)		1, 2, 4, 6, 8, 12, 16, 24, 32					
Channel	Number of CHPs	1				1-2	
	Number of channels	Max 128				Max 256	
	Types	MXC, BMC (4.5 Mbyte/s), High speed optical (36 Mbyte/s) Optical (9 Mbyte/s), HIPPI (100 Mbyte/s)					
	Throughput (Gbyte/s)	Max 1				Max 2	
Number of arithmetic pipelines	Multiply & add/logical	1	2			4	
	Divide	1	1			2	
Number of load/store pipelines		1	2			4	
Vector pipeline throughput/cycle		1		2	4	1	2
Capacity of vector register (Kbyte/SU)		32		64	128	32	64
Capacity of buffer storage (Kbyte/SU)		128					

storage for quick processing. Vector instructions are issued from the SU and executed in the VU. The VU consists of large-capacity vector registers and plural vector pipelines.

2) Main storage unit (MSU)

This unit uses high-speed 1-Mbit static RAMs to process multiple memory accesses from the VPU at high speed. The maximum capacity in the system is 2 Gbytes.

3) System storage unit (SSU)

The SSU is a large-capacity storage unit positioned above the MSU. This is the first time the SSU has been provided in a system. The SSU uses 4-Mbit DRAMs to achieve a maximum capacity of 32 Gbytes. It can be used as the swapping area for jobs, and as a virtual filling area for I/O operations to achieve a higher system throughput. GaAs LSIs with 1 200 gates are used in the data bus logic to attain a higher throughput between the SSU and the MSU.

4) Channel processor (CHP)

There are four types of channels: the 4.5 Mbyte/s electrical channel, 9 Mbyte/s optical channel, 36 Mbyte/s high-speed-optical channel, and 100 Mbyte/s HIgh Performance Parallel Interface (HIPPI)[Note] channel. A total throughput of 2 Gbyte/s can be attained using the maximum configuration of 256 channels. Peripheral devices can be placed as far as 2 km away using optical channels.

3.3 Performance and specifications

Table 1 gives the specifications of the VP2000 series. Figure 3 compares the overall performance of each model in the series and the four levels of vector performance that are available (i.e. from 0.5 GFLOPS to 5 GFLOPS).

4. Vector processing unit (VPU)

The VPU executes vector and scalar opera-

Note: An interface specification of the ANSI standards.

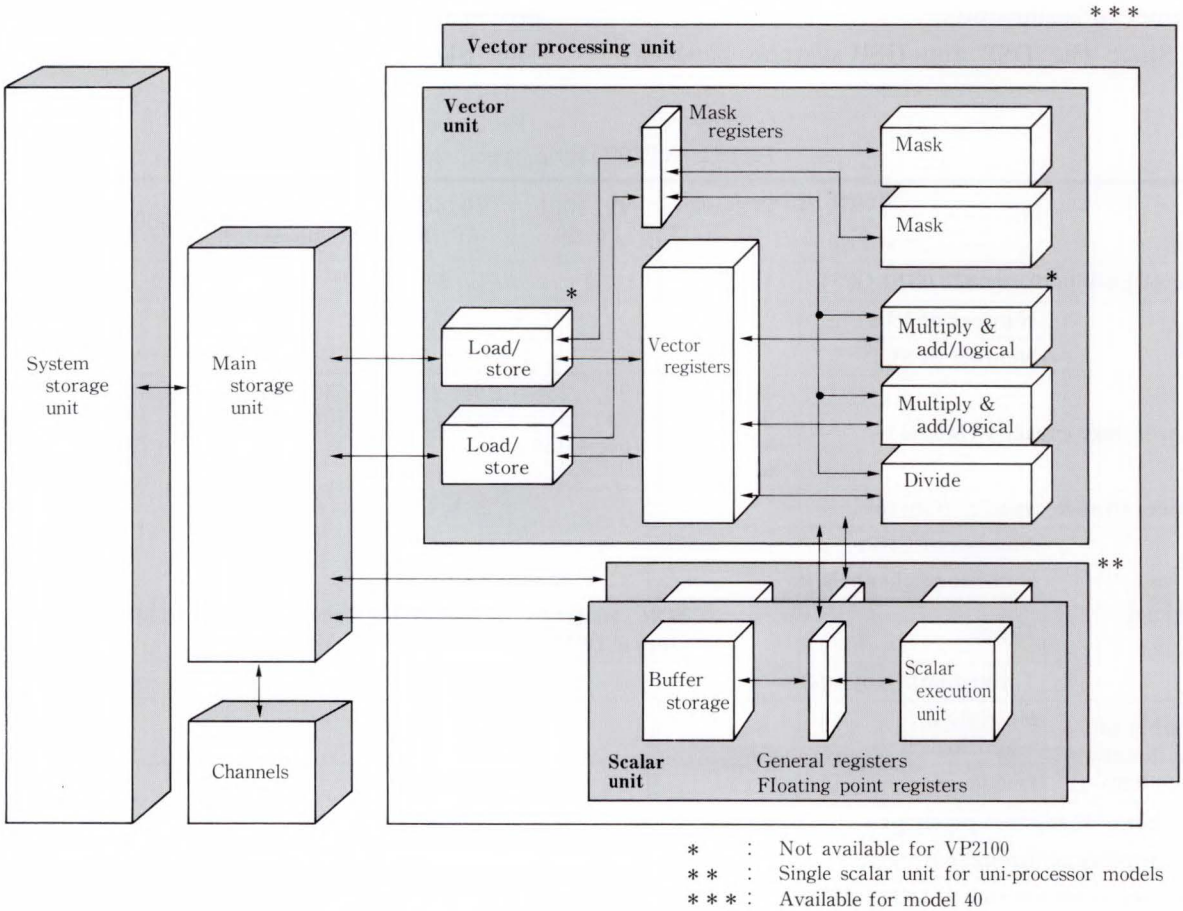


Fig. 4—VP2000 series hardware block diagram.

tions and has the same FUJITSU M-series architecture that is used in Fujitsu's general purpose computers. This chapter describes the features of the advanced hardware for high speed processing in the VPU.

4.1 Hardware configuration

Figure 4 shows the hardware block diagram of the VP2000 series. The VPU is connected to the MSU so that it can fetch instructions and data from the MSU and transfer data to the MSU. The SU has a buffer storage unit, registers, and a scalar arithmetic unit. Instructions and data from the MSU are stored in the buffer storage, and then transferred to the scalar arithmetic unit for high-speed processing of scalar instructions.

The VU has vector registers, mask registers, and plural vector pipelines. The basic configuration of vector registers can have up to 256 8-byte registers, numbered from 0 to 255. Each model has the specified number of elements per register (64 for VP2600, 32 for VP2400, and 16 for all other models). The vector register is between the MSU and the vector operation pipelines. This register contains a large amount of vector data so that access to the MSU can be minimized. As with previous supercomputers, to fully utilize the total capacity, the vector registers may be concatenated to form the following configurations: 64 (number of elements) x 256 (number of registers), 128 x 128, 256 x 64, ..., 2 048 x 8.

Up to 256 mask registers, numbered from 0 to 255, are available. These mask registers are 1-bit registers. As with the vector registers, each model consists of the specified number of elements, and the mask registers can also be concatenated.

As well as simple DO loops with only four arithmetic rules, a FORTRAN program can also have many complicated DO loops with conditional statements (IF statements). Therefore, masked operation, compress/expand, and list vector functions can be used as required to vectorize IF operations and to improve the execution efficiency[1]. The mask register keeps the mask data for the arithmetic mask function that specifies each masking or vectorization. (The vector pipelines execute operations on elements corresponding to vector data according to 1/0 patterns in the mask data.) This approach extends the vector processing range.

The load/store pipeline processes access to the main storage data specified by vector instructions. Load or store data is transferred between the MSU and the vector/mask registers. The logical addresses used for access to the MSU are rapidly translated into real addresses by the predetermined hardware address translation table under the specified page size[1].

The mask pipeline operates on mask data. All models have two mask pipelines: one for total summation/retrieval processing and the other for logical operations.

The multiply & add/logical operation pipeline is the so-called universal pipeline. This pipeline executes the multiply, add, multiply & add, first order recurrence, and logical operation functions. This pipeline enables hardware to execute vector instructions flexibly according to various program sequences.

The divide pipeline executes divide instructions exclusively.

4.2 Instruction control

Figure 5 shows the control instruction pipeline of the VPU. The VP2000 series VPU has arithmetic control pipelines to concurrently execute vector and scalar instructions, or vector and vector instructions.

1) Parallel instruction execution

Scalar instructions are executed in the SU. Vector instructions are issued to the VU. The VU controls instruction execution by transmitting the instruction to the vector control pipeline appropriate to the execution type. In the VU, two load/store pipelines (one for VP2100), two mask pipelines, and two of the three arithmetic operation pipelines (one of two for VP2100) can operate concurrently. The SU and at up to six of the seven vector pipelines (up to four for VP2100) can, therefore, operate concurrently (see Fig. 6). Also available is a linkage facility which allows two pipelines to be connected logically for continuous operations as

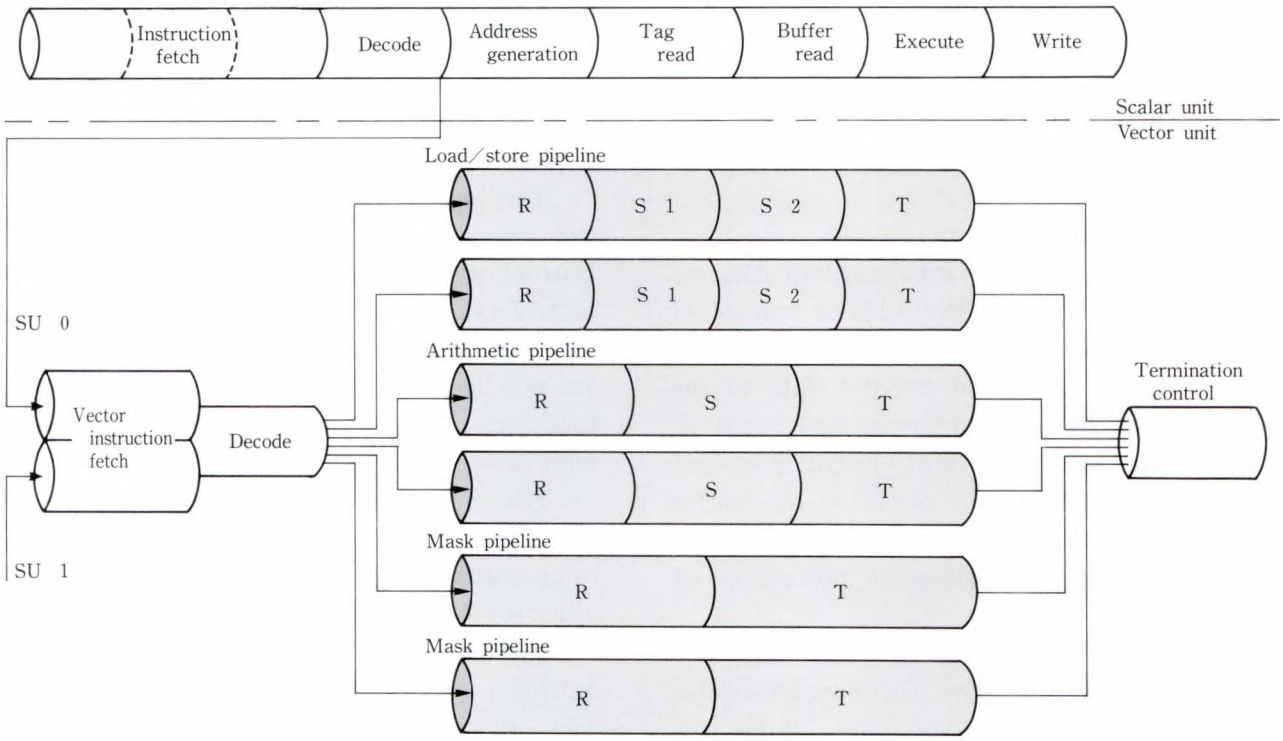
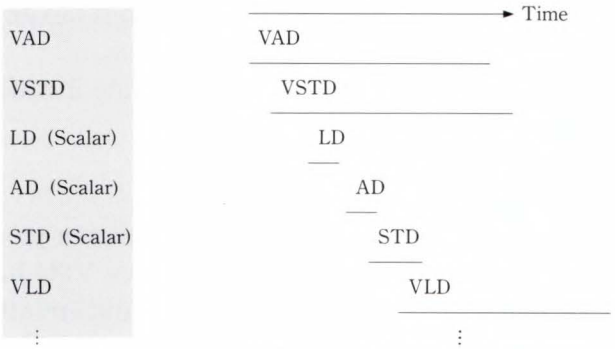
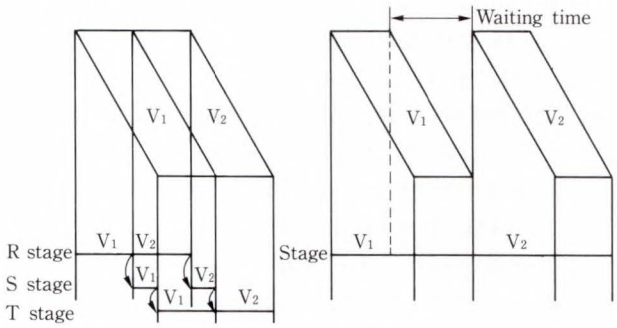


Fig. 5—Instruction control pipeline.



a) Program b) Instruction sequence

Fig. 6—Parallel processing.



a) Continuous processing b) Execution with only one stage

Fig. 7—Continuous execution.

if they were a single pipeline. By using this facility, a vector instruction can start to read the vector register or the mask register written by the preceding instruction without waiting for completion of the preceding instruction.

2) Continuous instruction execution

Each instruction control pipeline in the VU consists of an R, S, and T stage. A vector instruction is controlled and executed through the register at an appropriate stage in the pipeline. The R stage controls the fetching of operand data and the start of instruction execution. The T stage controls the storage of operand data and the termination of instruction processing. The S stage is an intermediate stage between R and T. When a vector instruction completes the R stage, the information required for instruction control is retained in the S stage until the instruction moves into the T stage. This allows the next instruction to be executed immediately in the R stage after the preceding instruction is issued (see Fig. 7).

4.3 High speed processing functions

Some vector instructions write data into a scalar register. When the SU detects this type of instruction, it checks whether the next scalar

instruction will need the register operand. If the next instruction needs the data, the instruction must wait until the vector instruction finishes writing data into the register. If the next instruction does not need the register operand, it can be executed immediately. If a subsequent instruction needs to read to scalar register written by the previous vector instruction, it can use the data by bypassing the next data to be written into the VU without having to wait until the register has been set.

A vector instruction that randomly accesses the main storage address may access the same address repeatedly. To avoid these wasteful accesses to the main storage unit in high-speed processing, only one access is required. In the VP2000 series, high performance is realized by using these unique functions to control the hardware.

5. Main storage unit (MSU)

A supercomputer must have a main storage unit with a large capacity, high-speed access, and high throughput. The RAS function is also important for stabilizing the system operation.

5.1 Capacity and performance of main storage

A capacity of 64 Mbytes per array card and 512 Mbytes per unit has been achieved using new technologies. Models VP2600/10, VP2600/20, and VP2400/40 have a maximum main storage capacity of 2 Gbytes. The large-capacity MSU is accessed from the SU, VU, and CHP. For the VU, the throughput between the MSU and the vector registers is important. That is, the MSU must supply the necessary data to the vector register according to the operational capacity of the vector operation pipelines. Therefore, each VP2000 series supercomputer has a data bus suitable for the operation capacity between the MSU and the VU. The VP2600 and VP2400/40 models also have sufficient interleave for 512 ways (units of independent access to memory).

5.2 RAS function

The MSU has an extended error checking and correction (ECC) function that automatically corrects single-bit errors and detects all double-bit errors and multi-bit errors in a single block (four bits). The MSU also has a function that detects single-bit fixed errors in a RAM and automatically replaces the RAM with a spare chip. This function, called the automatic alternate memory allocation function, assures a high immunity against fixed errors.

6. System storage unit (SSU)

The new SSU used in the VP2000 series has a large storage capacity for the vector processing swap area and the I/O virtual file area[5]. The SSU further enhances the system throughput.

6.1 Capacity and performance

The SSU has a capacity of 1 Gbyte to 32 Gbytes and can exchange data with the MSU. The speed of data transfer between the SSU and the MSU must be sufficient for large amounts of data. The VP2600 can transfer data between the SSU and the MSU at up to 2 Gbyte/s.

6.2 Transfer instruction

Two types of transfer instructions are used: synchronous and asynchronous.

1) Synchronous transfer instructions

In this type of instruction, the instruction and the transfer operate synchronously. Data transfer begins at the start of the transfer instruction, and the instruction terminates after the data transfer is completed. The VPU waits until the instruction finishes.

2) Asynchronous transfer instructions

This type of instruction transfers data asynchronously. Data transfer begins at the end of the transfer instruction. The data transfer is executed independently and its termination is reported by an interruption. When an asynchronous transfer instruction is issued, the VPU executes the next instruction independently without waiting for the termination of the data transfer. This releases the VPU from transfer instructions so that it can process high-speed arithmetic instructions concurrently with the data transfer.

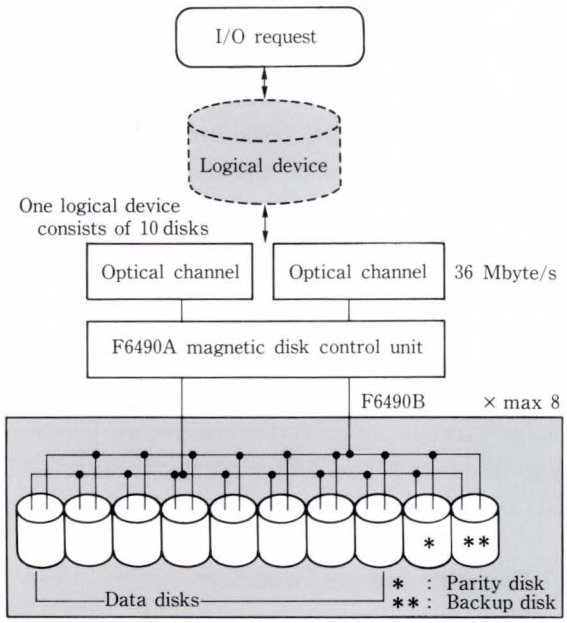


Fig. 8—Disk array subsystem.

6.3 RAS function

The SSU has an ECC function for automatic detection and correction of all single-bit errors and for detection of all double-bit errors.

To prevent a single-bit error from becoming a double-bit error, the memory patrol function periodically reads data from storage and corrects any detected single-bit errors. Also, if a single-bit fixed error is detected in the RAM, the automatic alternate memory allocation function replaces the RAM with a spare one. This function improves immunity against fixed errors.

7. New channels

The VP2000 series computers have two new types of high-throughput channels that improve the I/O transfer rate and simplify connection to an open network.

7.1 High-speed optical channel

The 36 Mbyte/s optical channel can connect the F6490 large-capacity high-speed disk unit[6]. F6490 contains an array of ten disks (see Fig. 8). The 8-byte data is distributed among eight of these disks (1 byte per disk). The remaining two disks are used as a parity disk and a backup disk. If a fixed error occurs in one of the eight disks, an alternate allocation function replaces the disk with the backup disk. This method enables a higher throughput and higher reliability than conventional type disk units.

7.2 HIPPI channel

A HIPPI channel with a 100 Mbyte/s throughput can be connected to the Ultra-Net[Note], the high-speed, multi-vender LAN. This connection enables high-speed data transfer between a VP2000 series supercomputer and workstations under a multi-vender environment.

8. Multiprocessor system

The multiprocessor system improves the overall system performance. There are two types of multiprocessor system: the dual scalar processor (DSP) and the quadruple scalar processor (QSP).

8.1 Dual scalar processor (DSP)

The DSP is a new architecture for multiprocessor systems. It increases system performance by attaching another scalar unit.

In ordinary vectorized application programs, the use-rate of the VU rarely reaches 100 percent. For example, a program with a vectorization ratio as high as 90 percent uses the VU for less than 50 percent of the total CPU time. Therefore, an extra SU that shares one VU can be attached in order to increase system performance.

The VU has a programmable register for each SU (see Table 1). This hardware allows software to control the system in the same way as in an ordinary multiprocessor system. Vector instructions issued by the two SUs are alternately selected by hardware (SU 0 and SU 1 each correspond to the SU in a DSP system), and are then transmitted to the pipelines for execution (see Fig. 5).

If all jobs are scalar jobs, a DSP system has double the throughput of a UP system. However, this advantage decreases in proportion to the vectorization ratio of programs. If all jobs are vector jobs, the throughputs of these two systems are the same. For example, if the vector vs scalar speedup factor is ten and there is no

Note: Ultra-Net is a registered trademark of Ultra Network Technologies, Inc. USA.

contention of vector unit usage, a DSP system will have double the throughput even if both programs running on the DSP system have a 91 percent vectorization ratio.

8.2 Quadruple scalar processor (QSP)

In the VP2200 and VP2400 models, there is another type of multiprocessor system, the QSP. The QSP system uses four SUs and two VUs, and is configured as a multiprocessor with two DSP systems tightly coupled to the MSU. In this configuration, the software can regard the system as an ordinary multiprocessor with four processors. In the QSP system, multi-tasking can be performed by distributing the tasks of a single job to the processors[5].

9. Conclusion

This paper introduced the features, system configurations, and functions of the VP2000 series. There is an enormous demand for high-speed processing of large-scale scientific and technical calculations. To satisfy this demand, faster and larger systems must be realized. This can be achieved by increasing the processing speed, developing new parallel processing architectures, creating parallel large-scale systems, generating advanced software, and developing new devices and technologies.

By incorporating high-speed channels, high-speed disks, and image processing for high-speed I/O operations, Fujitsu will continue to develop open systems that are adaptable to the standard UNIX culture. Also, Fujitsu intends to perfect a multi-vendor environment that will support equipment ranging from EWSs to supercomputers.

References

1) Tamura, H., Shinkai, Y., and Isobe, F.: The Supercomputer FACOM VP System. *FUJITSU Sci. Tech. J.* **21**, 1, pp. 90-108 (1985).
2) Uchida, N.: Architecture of VP2000 Series. InfoJapan'90 Information Technology Harmonizing with Society, Inf. Proc. Soc., Jpn., 1990, pp. 247-253.
3) Kaneko, A. et al.: Hardware Technology for FUJITSU VP2000 Series, *FUJITSU Sci. Tech. J.*, **27**, 2 (Special Issue on Supercomputer VP2000 Series), pp. 158-168 (1991).
4) Ohno, K., et al.: Semiconductor Devices for FUJITSU VP2000 Series, *FUJITSU Sci. Tech. J.*, **27**, 2 (Special Issue on Supercomputer VP2000 Series), pp. 169-178 (1991).
5) Hotta, K. et al.: Basic Software for FUJITSU VP2000 Series, *FUJITSU Sci. Tech. J.*, **27**, 2 (Special Issue on Supercomputer VP2000 Series), pp. 197-210 (1991).
6) Oyama, T., Ogawa, Y., and Sugiyama, K.: F6490 Magnetic Disk Subsystem: DIA. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 291-295 (1991).

Nobuo Uchida
Computer Engineering Dept.
Main Frame Div.
FUJITSU LIMITED
Bachelor of Electrical Eng.
Waseda University 1982
Specializing in Computer Engineering

Hiroshi Tamura
Computer Engineering Dept.
Main Frame Div.
FUJITSU LIMITED
Bachelor of Electrical Eng.
Yokohama National University 1969
Specializing in Computer Engineering

Yuji Oinaga
Computer Engineering Dept.
Main Frame Div.
FUJITSU LIMITED
Bachelor of Science Eng.
Hokkaido University 1974
Specializing in Computer Engineering

Kazuyuki Shimizu
Processor Development Div.
FUJITSU LIMITED
Bachelor of Electronic Eng.
Nagoya Institute of Technology 1966
Specializing in Computer Engineering

UDC 681.32

Hardware Technology for FUJITSU VP2000 Series

• Akira Kaneko • Kiyoshi Kuwabara • Shun-ichi Kikuchi • Takanobu Kano

(Manuscript received December 27, 1990)

Fujitsu has developed new packaging, cooling, and power supply technology for new high-speed and high-density LSIs. This paper introduces an ultra-miniature LSI package, a ceramic board on which two million gates can be mounted, and high-performance cooling and water-cooled power supply technology.

1. Introduction

Supercomputer performance depends primarily on hardware technology. Especially, semiconductor technology improvements and packaging technology advancements have contributed to the rapid progress of supercomputer performance by increasing the clock speed and making high-density gate packaging possible.

The clock cycle time is closely related to the signal propagation delay in LSIs and the board on which LSIs are mounted. In order to develop the FUJITSU VP2000 series high-speed supercomputer, high-speed LSIs, high-density ceramic boards, a high-performance cooling system and water-cooled power supply have been developed.

The major part of the VP2000 was built with ECL technology to achieve the fastest commercially available gate delay. In addition to the high-speed ECL LSI, a GaAs LSI was developed for signal transmission from large capacity strage unit to the vector processing unit. Details of the LSI technology are explained in another paper in this journal[1].

As the LSI gate count increases, it is necessary to increase the LSI pin count, and it is important to keep the board signal pattern length short for the high-speed clock system. A material that offers not only high density but also low dielectric constant is required for high-speed signal transmission on the board. In order to satisfy these two requirements, a glass-ceramic-composite board was developed.

As the gate count of the LSI increase, the LSI dissipates more power. On the other hand, LSIs are mounted on increasingly smaller board areas so that power density on the board has increased drastically. It was required to develop a high efficiency LSI cooling technology to effect improved cooling performance. Cooling technology capable of handling thirty watts per LSI and 4.6 kW per board was developed for the VP2000.

To keep the system cabinet small, it is also important to decrease the volume occupied by the power supply. A water-cooled power supply was, therefore, developed.

2. Packaging technology

The high-performance LSIs listed in Table 1 were developed for the VP2000. To accommodate the LSI's high speed, it was necessary to improve various factors in the new packaging design. Especially, the following two problems had to be solved.

First, the pattern density of the board had to be increased to compensate for the increase in pin density of the LSI. Second, it was desirable to reduce the dielectric constant of the board as well as to reduce the wiring length between LSIs. In order to solve these two problems, Fujitsu developed a new packaging technology for the VP2000.

2.1 Computer packaging history

Fujitsu has been developing unique high-density packaging technologies for four generations of computers, the FUJITSU M-190, M-380, M-780 and VP2000, to meet the demand for rapid increases in computer performance. Table 2 compares the packaging technologies for these four generations.

A planer packaging technology was developed for the FUJITSU M-190, which was introduced in 1974[2]. Forty-two printed circuit boards, called multi-chip carriers (MCC), were mounted on laminated bus plates, and were arranged in seven rows by three columns on both sides. Fourty-two LSI packages were mounted on each MCC. Each MCC had 664 signal pins, and the signal connections between the MCCs were provided by high-speed coaxial cables to minimize the signal propagation delay between the MCCs.

Table 1. LSI for VP2000

LSI	Specification	
Logic LSI	Circuit type	ECL
	Gate count	15 000 gates
	Propagation delay	70 ps/gate
	Power dissipation	30 W
	Package	462 pin PGA
RAM & logic LSI	Circuit type	ECL
	Capacity	64 Kbits+3 500 gates
	Address access time	1.6 ns
	Power dissipation	30 W
	Package	462 pin PGA
GaAs LSI	Circuit type	BFL
	Gate count	1 200 gates
	Propagation delay	60 ps/gate
	Power dissipation	5.5 W
	Package	180 pin FPT
Static RAM	Capacity	1 Mbits
	Address access time	35 ns

PGA: Pin Grid Array BFL: Buffered FET Logic
FPT: Flat lead Package Type

The FUJITSU M-380 was introduced in 1981. A unique three-dimensional stack structure was developed to shorten the signal wiring length between the MCCs[3]. Each MCC accommodated 121 LSI packages, and 12 MCCs (with 768 signal pins each) were stacked in a 50 cm cube. The signal connections and power supply were facilitated through two side panels.

The single-board CPU packaging technology was developed for the FUJITSU M-780 to eliminate the critical delay path between two boards. Up to 336 LSI packages were mounted on both sides of a 488 mm x 540 mm printed circuit board called the sub-system carrier (SSC)[4),5)].

Subsequently, the number of printed wiring boards (PWB) needed to form the CPU was significantly reduced over a ten-year period from 42 to only one, so that the size of the M-780 CPU unit could be reduced to 1/12 of the

Table 2. Trends in packaging technologies

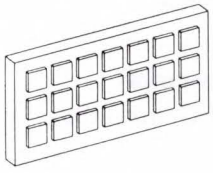
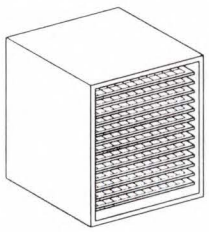
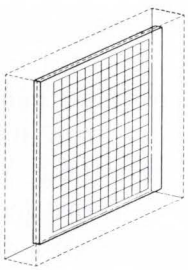
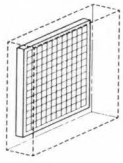
General-purpose computer	M-190	M-380	M-780	M-1800
Supercomputer	–	VP-100	–	VP2000
	Planer	Three-dimensional	Single-board	Single-board
CPU volume				
Size (cm x cm)	138 x 70	50 x 50	55 x 54	30 x 30
Number of PWBs	42	12	1	1
Volume ratio	100	13	8	1
Cooling	Forced air	Forced air	Conductive liquid	Conductive liquid

Table 3. Comparison of packaging complexity

Item	VP-100	M-780	VP2000
Pin density (pins/mm²)	0.16 (1.0)	0.28 (1.8)	1.29 (8.1)
Package density (%)	30.9 (1.0)	80.2 (2.6)	80.9 (2.6)
Silicon density (%)	4.0 (1.0)	15.0 (3.8)	51.0 (12.8)

M-190 CPU volume.

For the VP2000, Fujitsu developed high-density packaging with a glass-ceramic-composite board called the multilayer glass-ceramic-composite board (MLG)[6)-8)], to decrease board size and increase LSI pin density. It was possible to reduce the board size to 1/4 that of the M-780 SSC. The various units of the VP2000, including the vector execution unit (VXU), vector control unit (VCU), scalar unit (SU), input/output processor (IOP), and memory access control unit (MAC) are all accommodated on the MLG.

2.2 Packaging density

The packaging density trend of the three generations is listed in Table 3. The two factors to be considered for estimation of packaging density are pin density and silicon percentage, which is a unique new index, and it is calculated with Equation (1).

$$\text{Silicon percentage} = \text{SCHIP}/\text{SBOARD}, \quad (1)$$

where,

SCHIP: sum of LSI silicon die area on board,
SBOARD: board area − board peripheral area.

The VP2000's pin density per unit area on a board is five times greater than that of the M-780. Silicon percentage of the VP2000 is fifty one percent. This means that more than a half of the board surface is covered with silicon chips, and this value is difficult to achieve even if bare chips are embedded on the substrate. This high density is achieved by new packaging technology with ultra-miniature LSI packages and the MLG.

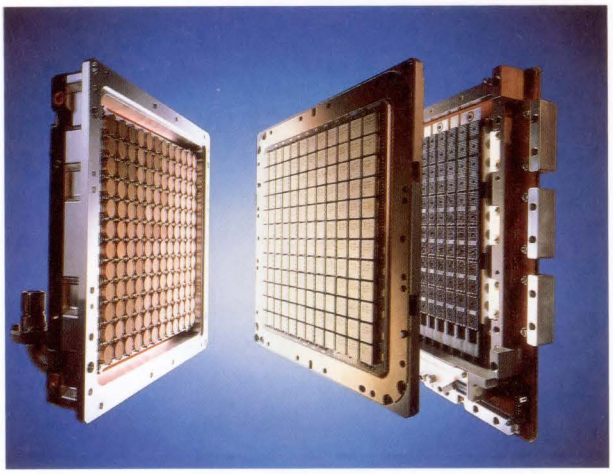


Fig. 1—MLG assembly.

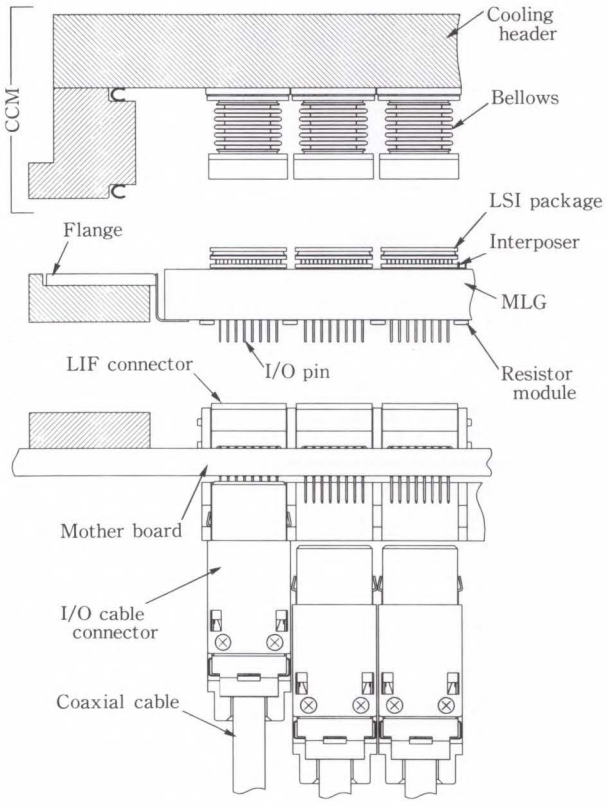


Fig. 2—Packaging overview.

2.3 Board assembly

The MLG assembly (MLA), mother board assembly, and conductive cooling module are shown in Fig. 1. The center portion shows the MLA. On the right-hand side, there is a mother board assembly, and on the left is a conductive cooling module. The MLG is mounted on a mother board, which supplies power to the MLG, with the light insertion

Table 4. LSI chip and package

Item	M-780	VP2000
Chip size	9.5 mm x 9.5 mm	13.5 mm x 13.5 mm
Package type	Flat lead package	Pin grid array package
Package size	22 mm x 22 mm	17 mm x 17 mm

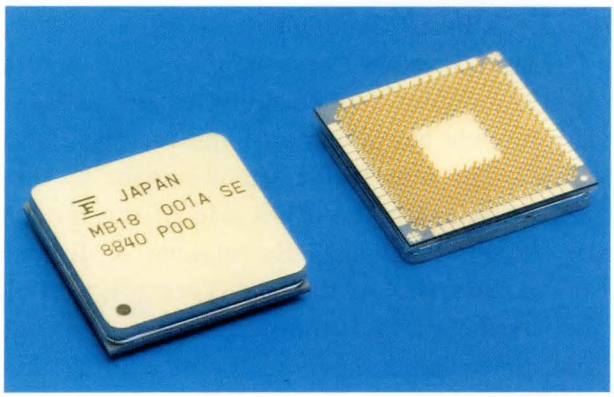


Fig. 3—LSI package.

force connector (LIF).

Figure 2 illustrates a cross-sectional view of the packaging hierarchy. On the top side of the MLG are mounted the LSI package and interposer, a very thin plate placed between the LSI package and the MLG. Interconnections between MLA units are made by using specially designed I/O connectors with high-speed coaxial cables.

2.4 Packaging components

2.4.1 LSI package

Table 4 compares LSI chips and packages between the M-780 and the VP2000. Although the silicon chip size of the VP2000's LSI increased to about twice that of the M-780's, it was possible to reduce the overall package size to about half that of the M-780 by developing the high-density pin grid array (PGA) type package.

Figure 3 shows a PGA type LSI package. It has 462 I/O pins; 320 are for signals, 140 for power and ground, and two dummy pins are provided for alignment. The package size is 17 mm square, with a height of 3.6 mm, and pins 1 mm long. I/O pins are arranged in a 0.45 mm staggered pitch pattern. The chip size is 13.5 mm square, with 100 μm pitch TAB leads. In order to minimize the thermal expansion mismatch between chip and MLG, aluminum nitride (AIN) is used for both heat sink and substrate. This LSI package is reflow-soldered by the butt-soldering method.

Table 5. MLG features

Item	Specification
Board size	245 mm x 245 mm
Board thickness	12.9 mm
Number of layers	61 (signal: 36)
DC resistance	100 mΩ/cm (Cu)
Dielectric constant	5.7
*T*pd	80 ps/cm
Number of LSIs	144

2.4.2 MLG and MLA

The primary advantage of using ceramic boards is the ability to make thick boards. When working with organic boards, it is difficult to make a thick board that can be drilled with fine drill sizes, which means that the organic board technology is limited from the standpoint of high density.

For the VP2000, glass-ceramic-composite was newly developed for the board material instead of conventional alumina material. Advantages of glass-ceramic-composite are as follows. First, alumina, the common material for printed circuit boards, has a relatively high dielectric constant, of about ten. The signal propagation speed on the board pattern was improved by about 25 percent by making use of the new glass-ceramic-composite material, the dielectric constant of which is 5.7.

The second advantage of glass-ceramic-composite is its low firing temperature. Because the firing temperature is around 1 000 °C, copper can be used for internal conductors. With the conventional alumina substrate, only tungusten (W) or molybden (Mo) could be used for the conductor material. The use of low resistivity copper assures low DC voltage drops, even when the board size becomes larger, which is required for enabling one-board CPU packaging.

The major features of MLG are listed in Table 5. The outer dimension is 245 mm square, and the board thickness is 12.9 mm. The total

layer count is 61, with 36 layers used for signals. The internal conductor material is copper, and the DC resistance of signal trace is about 100 mΩ/cm. The dielectric constant of this board is 5.7, and the propagation delay time (Tpd) is 80 ps/cm. Figure 4 shows the MLG. The basic grid pitch of this board is 0.45 mm, and there is one routing channel between 0.45 mm grids The via diameter is 80 μm, and the signal patterns is 95 μm wide and 45 μm thick.

LSI packages are arranged in a 12 x 12 matrix, the center spacing of the LSI package is 18.9 mm. On the bottom of the MLG, 312 resistor modules and 144 I/O pin blocks are mounted. The I/O pin blocks are arranged in a 12 x 12 matrix. There are 60 pins in each block, so the total I/O pin count is 8 640. The center spacing of I/O pin blocks is 18.9 mm.

2.4.3 Resistor module

Figure 5 shows the resistor module. A total of 312 modules can be mounted on an MLG. This module has 66 circuits, and each resistance is 65Ω. The resistance element is made by thin-film process. On the other side of this module are 83 solder bumps. By making a distance shorter than that of flat lead types, it is possible to decrease the inductance and noise. The outside dimension is 15 mm x 3 mm with a 1 mm height.

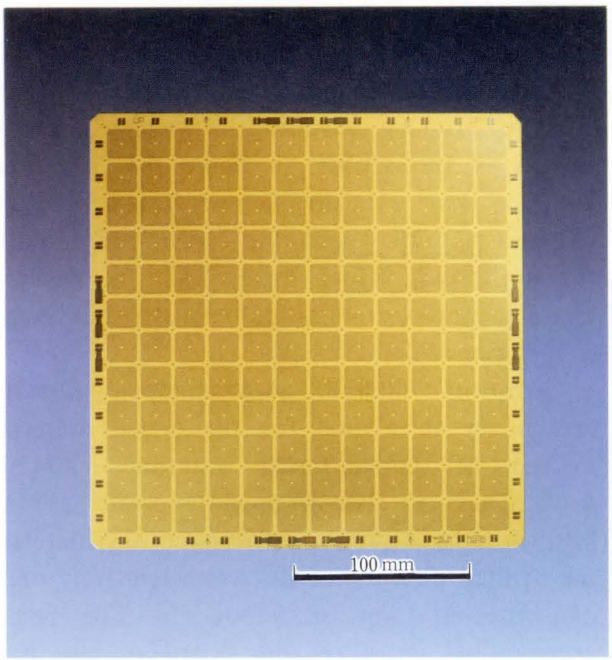


Fig. 4—MLG row board.

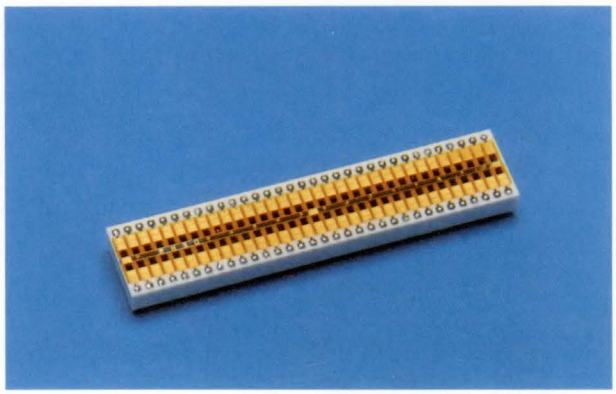


Fig. 5—Resistor module.

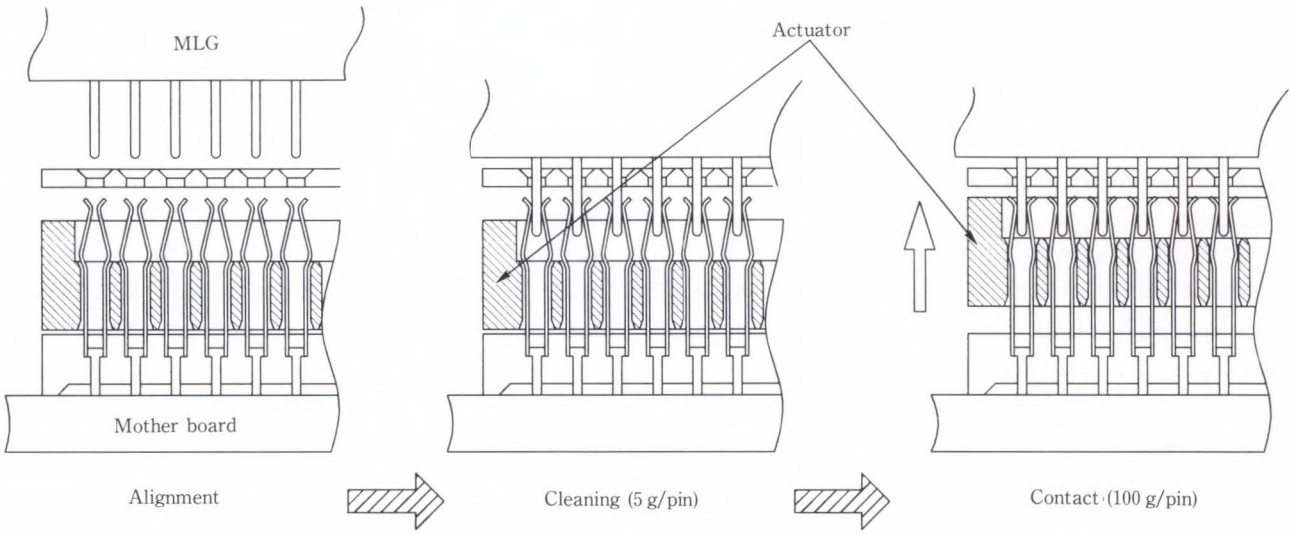


Fig. 6—New concept of connector.

2.4.4 Mother board

The fully assembled MLG is mounted on the mother board. On the mother board, LIF connectors are interconnected to the I/O pins of the MLG, and the coaxial connectors used for clock signals and terminals for supplying power are mounted.

The internal conductor material is copper, and 0.5 mm thick copper plates are used for the power planes. It was possible to hold the voltage drop in the MLG to less than 15 mV.

All electric components are reflow-soldered at one time. The dimensions of the board outline are 359 mm x 340 mm and 7 mm thick. The total layer count is 13.

2.4.5 LIF connector

A reliable connector must fulfill two basic requirements. One is to provide enough wiping action to remove surface contamination such as dust, and the other is to provide enough force for stable contact. It is not easy to satisfy these requirements, especially when the number of I/O pins to be interconnected becomes very large.

Figure 6 illustrates how these two issues were solved. The left-hand side shows the completion of alignment between the male pin and female spring. In this stage, a long span of the female spring gives a light but sufficient force to clean the surfaces. After insertion, a special mechanism moves an actuator upward, which applies force to the female contact, and it produces an adquate force. Eventually Fujitsu was able to clean the mating surface with a force as small as 5 g, and was able to get stable contact by applying a large force of about 100 g.

Figure 7 shows how 8 640 pins are mated. One connector module has 120 pins, corresponding to two LSI cells, and the connector modules are arranged in a 12 x 6 matrix on the mother board. A pair of protrusions on the slide cams enter grooves of the cam actuator, which moves up and down with a back and forth movement of two slide cams. Each lever operates for two rows at one time, and all pins can be connected within 20 seconds with a reasonable actuation force.

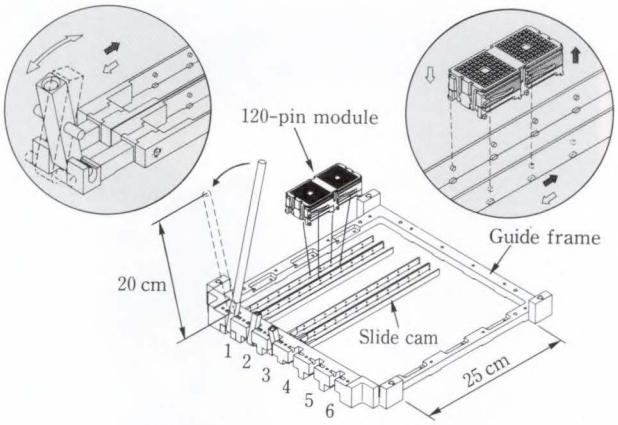


Fig. 7—Actuation mechanism.

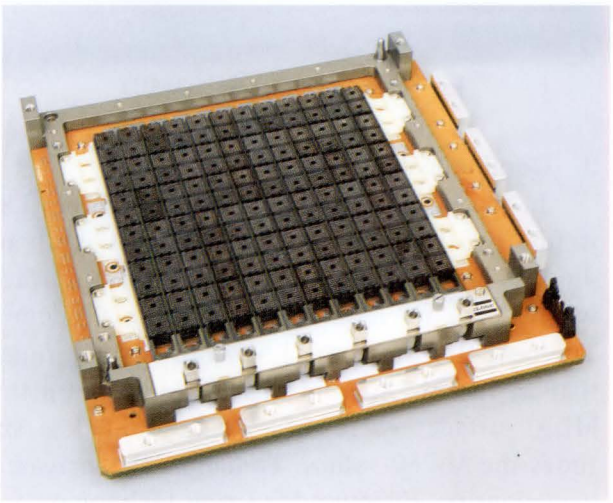


Fig. 8—Connector modules on mother board.

Figure 8 shows the fully assembled LIF connector modules on the mother board. A guide frame supports the assembled MLG.

2.5 MSU assembly

Figure 9 shows an MLG assembly of main storage unit (MSU). As with the other units, the MAC MLA is mounted on a mother board. And on the other side of mother board, a total of eight memory cards can be connected. These cards are cooled by forced air. The capacity of this memory card is 64 Mbytes, so the total memory capacity of the MSU is 512 Mbytes. For the maximum system configuration, the total memory capacity is 2 Gbytes.

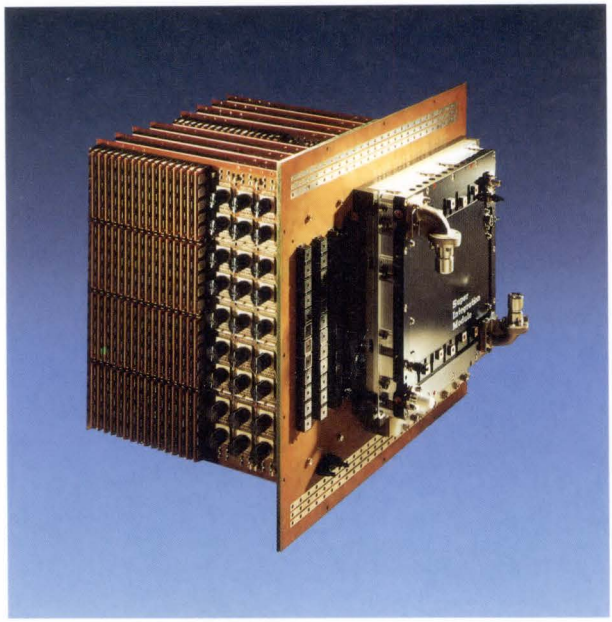


Fig. 9—Main storage unit (MSU).

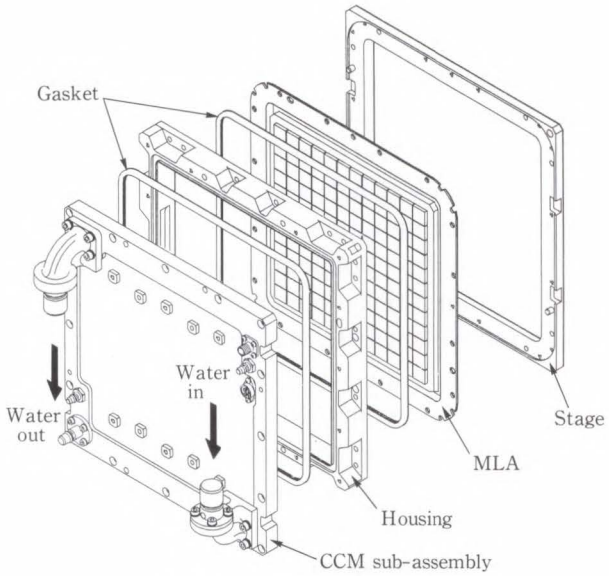


Fig. 10—SIM components.

3. Cooling technology

In the VP2000, the heat dissipation per LSI is up to 30 W and that of the MLA is up to 4.6 kW. The heat density on the MLA surface where LSIs are mounted goes up to 9 W/cm^2. The heat flux from the LSI chips is three times that in the M-780, and the heat density on the MLA surface where LSIs are mounted is six times the M-780 values. To handle the increased heat density, Fujitsu added new technology for heat transfer and greatly improved cooling performance.

3.1 Cooling mechanism

The new conductive cooling module (CCM) which was developed for the VP2000 consists of stage, gaskets, housing and CCM subassembly. The CCM sub-assembly also consists of a cooling-header and flexible thermal conductor (FTC) constructed as a micro-bellow. One FTC is used for each LSI mounted on the MLG.

The cooling structure, combining CCM with MLA, is shown in Fig. 10. This mechanism is basically the same as the one used in the M-780. It is an impinging jet flow which have been studied as practical application for very large-scale computers. The FTC contacts the top surface of the LSI through the medium of a thermal compound made to satisfy the requirements of special specification. An impinging jet of coolant from the nozzle in the FTC flows into the extended heat transfer plate of the FTC under lower compressive force and effectively removes the heat generated by the LSI.

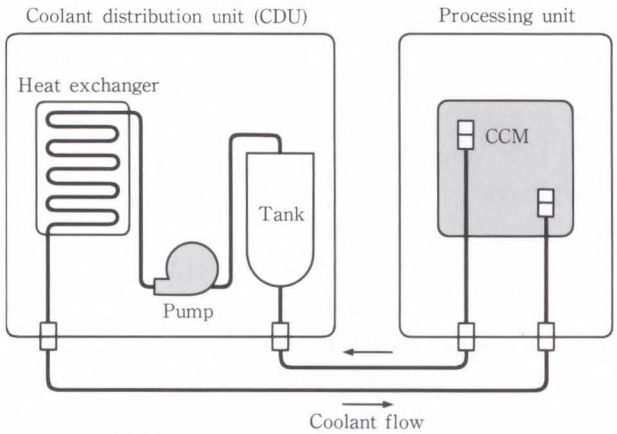


Fig. 11—Forced liquid circulation system.

Fujitsu uses water as coolant supplied from the coolant distribution unit (CDU) and piping loop which shapes the closed loop. The forced circulating system is sketched in Fig. 11.

3.2 Promotion of cooling performance

PN junction temperature (T_j) of the LSI is calculated from the heat of chip (P) and the thermal resistance from PN junction to coolant

($R_{j\text{-}w}$), the coolant temperature rise in CCM (ΔT_w) and coolant temperature supplied by CDU (T_w).

$$T_j = P \times R_{j\text{-}w} + \Delta T_w + T_w \quad \cdots\cdots\cdots (2)$$

ΔT_w changes with the location and heat flux of LSIs mounted on the MLG. Maximum value $\Delta T_{w\,max}$ depends primarily on the flow rate of coolant and the total heat of MLA. For example, when the flow rate is 10 ℓ/min and the total heat is 4.6 kW, ΔT_w is approximately equal to 6 degrees. Typically, $\Delta T_{w\,max}$ ranges from 0-6 degrees. T_w is obtained from the flow rate for CCM, the total heat on MLA and the cooling performance of CDU. Under ordinary operating conditions, T_w is equal to 25 °C.

$R_{j\text{-}w}$ in Equation (2) is a proper parameter used as a characteristic of cooling performance. In response to changes in the values of LSI and CCM, $R_{j\text{-}w}$ deviates in Fig. 12. These data

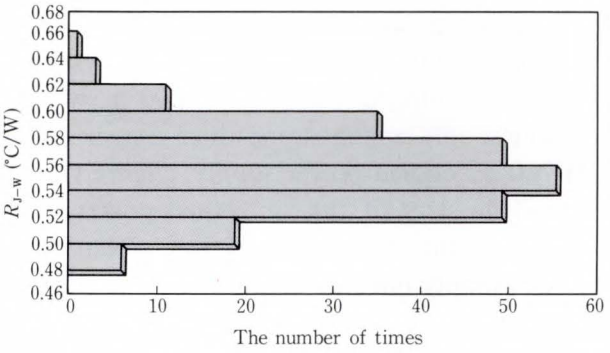


Fig. 12—$R_{j\text{-}w}$ evaluation data.

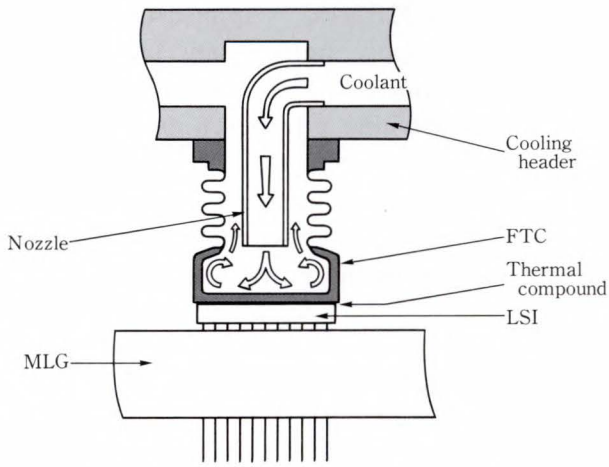


Fig. 13—Conductive cooling mechanism.

items are derived from experiments using special chips contained temperature-sensing diodes in the PN junction and can be obtained from the relation between the temperature in the PN junction and the voltage drop. Mean value of $R_{j\text{-}w}$ is equal to 0.56 °C/W. Even if the standard deviation (3σ) of $R_{j\text{-}w}$ is considered, the maximum value is less than 0.65 °C/W. It was possible to reduce the thermal resistance from PN junctions to the main flow of coolant to one-fourth that of the M-780 (2.4°C/W).

The basic mechanism of conductive cooling is shown in Fig. 13. R_{cond} indicates thermal resistance by the heat conduction between the PN junction of the chip and the heat transfer plate of FTC, R_{conv} indicates the thermal resistance by the heat convection between the heat transfer plate of FTC and the main flow of coolant. $R_{j\text{-}w}$ is the sum of R_{cond} and R_{conv}.

$$R_{j\text{-}w} = R_{cond} + R_{conv}. \quad \cdots\cdots\cdots (3)$$

R_{cond} is separated into three parts. The first is the internal thermal resistance of LSI, the second is the resistance by thermal connection between FTC and LSI, and the third is the thermal resistance of FTC heat transfer plate. Paying particular attention to the ratio of thermal resistances in the M-780, Fujitsu aimed for reduction of the thermal connection between FTC and LSI. To absorb the subtle gap which causes the thermal resistance between FTC and LSI, a thermal compound with a thermal conductivity exceeding 1 W/m·K was developed. Thus, 0.1 °C/W or less resistance was achieved in that thermal connection. And therefore, R_{cond} has been reduced to 0.22 °C/W, a level one seventh that of the M-780. As for R_{conv}, the nozzle shape and the shape inside the FTC for coolant flow were optimized. The average heat-transfer coefficient of the FTC heat transfer plate was raised to 14 000 W/m²·K and is 1.6 times the 9 000W/m²·K value of the M-780. Therefore, R_{conv} has been reduced to 0.34 °C/W.

3.3 Numerical analysis

In thermal design, the cooling performance was estimated with various analyses using

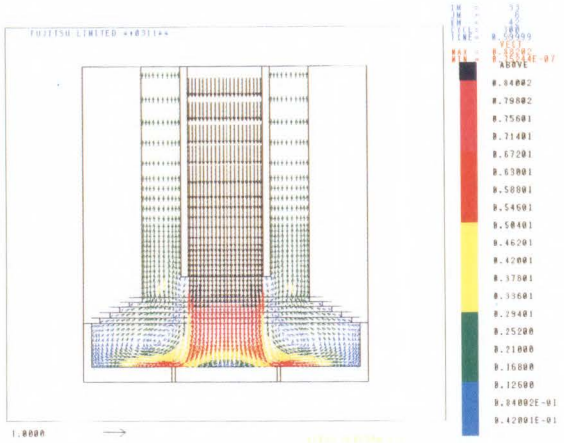


Fig. 14—Example of velocity vector inside FTC.

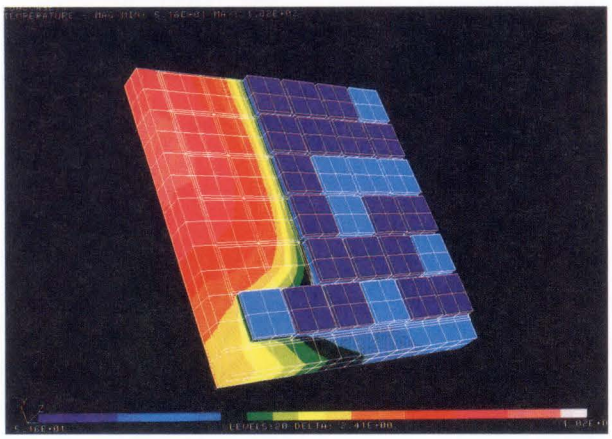


Fig. 15—Contour map example of temperature on MLA.

the finite element method (FEM). For example, Fig. 14 shows a coolant flow inside the FTC. This figure promoted an understanding of the vortex-like flow which disturbs the wall jet flow along the bottom of the FTC heat transfer plate. This flow improved the heat transfer coefficient at the bottom circumference.

Figure 15 shows the contour map of temperature on an MLA surface, depicted by NASTRAN. On this type of MLA the maximum number of LSIs is not mounted. Through changes in materials or dimensions of parts in simulations, the cooling mechanism was optimized and finally verified by the experiments.

4. Power supply technology

The power supplies for large-scale computers must supply the low voltages and heavy currents

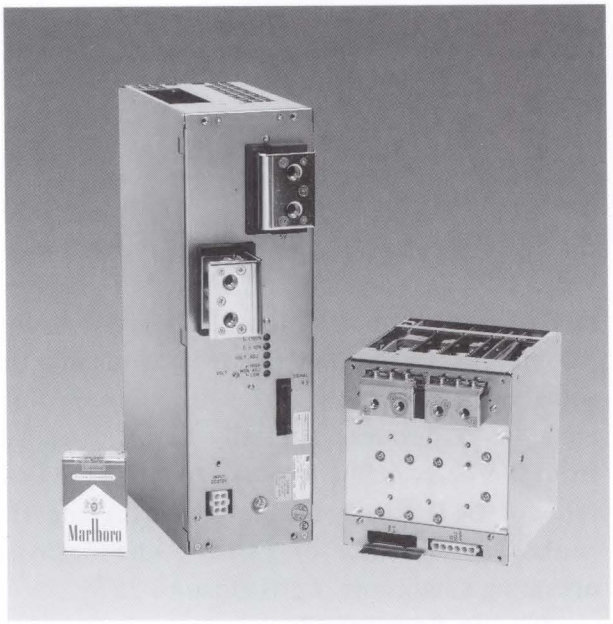


Fig. 16—Air-cooled unit (left) and water-cooled unit (5 V, 300 A).

required by the logic circuits. They must be located near the load as much as possible to avoid voltage drops and power losses in the bus bars. So, by circuit improvement and water cooling, Fujitsu developed for the VP2000 a switching power supply which is one third the size of the M-780 air cooled power supply. Figure 16 compares the sizes of the air cooled power supply unit for the M-780[9] and the water cooled power supply unit for the VP2000[10]

4.1 Power supply and cooling system

The three-phase 200 VAC to 240 VAC line power is rectified and filtered, and the resulting 270 VDC to 320 VDC is applied to the switching power supply units. They provide up to 530 A per unit of power to the logic circuits.

Each switching power supply is controlled by the unit power controller (UPC). The CDU supplies the water necessary for cooling both the MLA and each switching power supply. Figure 17 shows the power supply and cooling system.

4.2 Switching power supply

4.2.1 Circuit

The MOS FETs, which can be operated in parallel, are used for switching elements

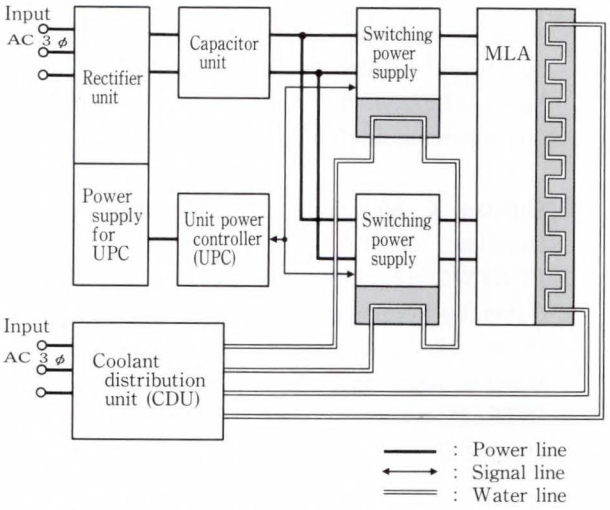


Fig. 17—Power supply and cooling system.

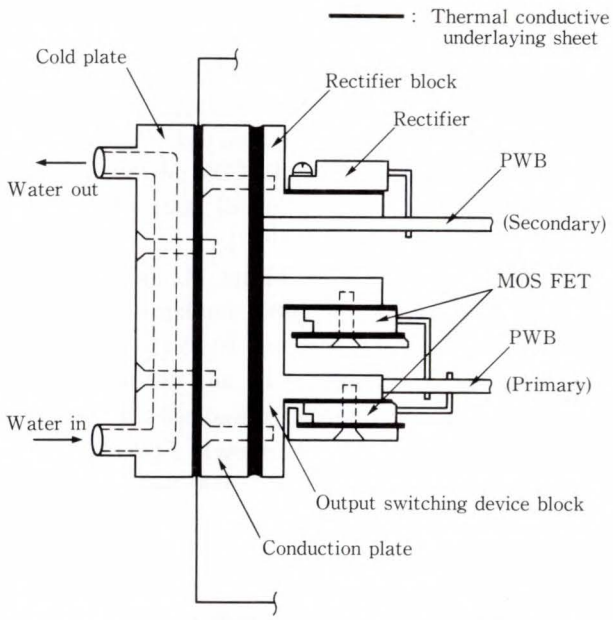


Fig. 18—Structure of the thermal conduction section.

to reduce switching losses. The switching transistors are driven by two-phase 200 kHz signals which are shifted 180 ° apart. The effective frequency seen at the output smoothing capacitors is 400 kHz. So it was possible to reduce the size of the capacitors to half that of the capacitors in the M-780.

4.2.2 Cooling method

The cooling method used in this switching power supply consists of both an indirect liquid cooling section using water as a coolant and a forced air cooling section using an external fan.

Figure 18 shows a cross-sectional view of the thermal conduction section. The heat sink part of the power supply, named the conduction plate, is clamped to the cold plate. This design facilitates installation and removal of the power supply unit. The output switching device block and the rectifier block are clamped to the conduction plate. The accumulated heat at the conduction plate is then transferred to the coolant via the cold plate.

The water cooling section handles approximately 70 percent of the power supply's overall heat, which is generated by the switching elements and rectifiers.

By careful choice of underlaying sheets and control of the clamping torque, the junction temperature of the semiconductors was able to be kept less than 60 percent of their approved maximum values.

5. Conclusion

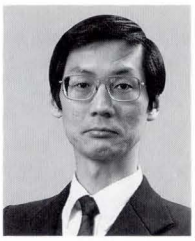
New LSIs and glass-ceramic-composite board are developed for the VP2000 system. The conduction cooling technology is developed for a 30 W LSI and the water-cooled power supply is also developed.

An ongoing demand for improvement in the supercomputer performance is increasing. Together with supercomputer architecture improvements like parallel processing, the development of hardware technology is becoming important. It may become necessary to develop a new generation of technology, such as GaAs LSI, bare chip packaging, and wafer scale integration. Fujitsu will continue developing new technology for the next generation supercomputer.

References

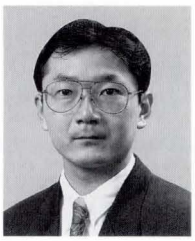
1) Ohno, K., Ooami, K., Itoh, H., and Suyama, K.: Semiconductor Devices for FUJITSU VP2000 Series. *FUJITSU Sci. Tech. J.*, **27**, 2, (Special Issue on Supercomputer VP2000 Series), pp. 169-178 (1991).
2) Yada, Y., Hiraguri, T., and Koike, Y.: Packaging technology of the M-190. (in Japanese), *NIKKEI ELECTRONICS*, 123, pp. 107-118 (1975).
3) Murase, T., Hirata, H., and Ueno, S. : High-density three-dimensional stack packaging for high-speed

computers. *ECC*, pp. 448-455 (1982).
4) Tsuchimoto, T., Shimizu, K., and Takamura, M.: The M-780 Large-scale Computer. (in Japanese), *NIKKEI ELECTRONICS*, **396**, pp. 179-209 (1986).
5) Nishihara, M., Murase, T., and Ogi, N.: Single-Board CPU Pakaging for the FACOM M-780. *FUJITSU Sci. Tech. J.*, **23**, 4 (Special Issue on Computer Systems), pp. 226-235 (1987).
6) Ogi, N., Murase, T., Nishihara, M., and Yamamoto, H.: Packaging for VP2000. (in Japanese), *NIKKEI MICRODEVICES*, June, pp. 50-55 (1989).
7) Kaneko, A., Seyama, K., and Suzuki, M. : LSI Packaging and Cooling Technologies for FUJITSU VP2000 Series. (in Japanese), *FUJITSU*, **41**, 1 (Special Issue FUJITSU VP2000 Series), pp. 12-19 (1990).
8) Niwa, K., and Takeda, Y.: Multilayer Glass-Ceramic-Composite Circuit Board for FUJITSU VP2000 Series. *FUJITSU Sci. Tech. J.*, **27**, 2 (Special Issue on Supercomputer VP2000 Series), pp. 179-186 (1991).
9) Shimizu, K., Miyazawa, T., and Tukuni, T.: System Overview of FACOM M-780. (in Japanese), *FUJITSU*, **37**, 2 (Special Issue :FACOM M-780 Computer Systems), pp. 93-103 (1986).
10) Hiraki, M., Kano, T., Kizu, H., and Satoh, H.: Water-cooled switching power supply. Proc. INTELEC 89 Inter. Telecom. Energy Conf. Session 20-8, 1989.



Akira Kaneko
Circuit Technology Dept.
Technology Development Div.
FUJITSU LIMITED
Bachelor of Electronics Eng.
Tokyo Institute of Technology 1972
Specializing in High-Speed Circuit Technology

Shun-ichi Kikuchi
Packaging Technology Dept.
Technology Development Div.
FUJITSU LIMITED
Bachelor of Mechanical Eng.
Yokohama National University 1981
Specializing in Packaging and Thermal Engineering of Computer Systems



Kiyoshi Kuwabara
Packaging Technology Dept.
Technology Development Div.
FUJITSU LIMITED
Bachelor of Mechanical Eng.
Tokyo University of Agriculture and Technology 1981
Specializing in Packaging Technology of Large-Scale Computers

Takanobu Kano
Computer Engineering Dept.
Mainframe Div.
FUJITSU LIMITED
Bachelor of Electronic Eng.
University of Electro-Communications 1974
Specializing in Power Supply Engineering

UDC 621.3.049.774:681.32

Semiconductor Devices for FUJITSU VP2000 Series

• Ken-ichi Ohno • Kazuo Ooami • Hideo Itoh • Katsuhiko Suyama

(Manuscript received November 29, 1990)

Advanced silicon and GaAs technologies have been developed and used in the high-speed, high-density LSIs of the FUJITSU VP2000 series. The main LSIs are 15 000-gate ECL array with a 70 ps propagation delay, 3 500-gate ECL array with 64-Kbit STRAM and a 1.6 ns maximum access time, 1-Mbit static RAM with a maximum access time of 35 ns, and 1 200-gate GaAs line driver/receiver with a 60 ps propagation delay time.

1. Introduction

Among the wide range of computers, supercomputers and high performance mainframes require the most advanced semiconductor devices to achieve the most powerful processing capabilities. Fujitsu's latest supercomputer FUJITSU VP2000 series uses many high performance devices. The key points of such devices are high speed and high density both in a chip and on a board. This paper describes the following main devices developed and used in the VP2000 series.

1) 70 ps ECL gate array and an array with 1.6 ns 64-Kbit RAM
2) 1-Mbit static RAM
3) 60 ps GaAs line driver/receiver

The above ECL gate arrays were developed based on advanced silicon bipolar technology with a sophisticated transistor structure called Emitter-base Self-aligned structure with Polysilicon Electrodes and Resistors (ESPER)[1], four-layer metallization and 0.8 μm scaled process. Further, the array chip is placed in a small package with 462 pins by 100 μm-pitch tape automated bonding (TAB) techniques.

The 1-Mbit static RAM in the main storage unit (MSU) is an advanced product of Si MOS technology. It uses three-layer polysilicon and two-layer metal process with 0.8 μm scaled CMOS technology. GaAs has the potential to drastically improve the performance of semiconductor devices. The driver/receiver used in the VP2000 series is the first step of GaAs LSIs toward possibly much under future usage in Fujitsu's computers. In the VP2000 series, the GaAs LSI is used for data bus logic in the system storage unit (SSU) in order to attain high transfer rate between SSU and MSU.

The contents of this paper on the above Si devices partially overlap the subject of a previous paper[2].

2. 70 ps ECL gate arrays

Two types of gate arrays have been developed. One is a 15 000-gate ECL logic gate array having a gate delay of 70 ps. The other is a 3 500-gate ECL gate array with 64-Kbit RAM having a maximum access time of 1.6 ns.

To meet the speed and integration objectives, all design and fabrication processes use the latest techniques in circuit and device design, wafer processing, and packaging. For circuit design, ECL was chosen for its speed and strong logic functions. For device design, gate array method was selected because it is best for production involving small lots of many different product types and short turnaround times. ESPER with 0.3 μm emitter and four-layer metallization are used for wafer processes. For assembly and packaging, 100 μm-pitch TAB technology and

Table 1. Evolution of Fujitsu ECL gate arrays for computers

System used	VP-400/M-380	M-780		VP2000 series	
Classification of array	Logic array	Logic array	Array with RAM	Logic array	Array with RAM
Number of gates	400	2 912	1 120	14 976	3 472
Number of output buffers		128	128	256	208
RAM size	–	–	16 384 bit	–	65 536 bit
Gate T_{pd}/RAM T_{AA}(typical)	350 ps/–	180 ps/–	280 ps/2.8 ns	70 ps/–	70 ps/1.4 ns
Power consumption per chip	2.7 W	8.5 W	9.5 W	30 W	
Supply voltage	−3.6 V	−3.6 V/−2.0 V		−3.6 V/−2.0 V	−3.6 V/−2.0 V −4.8 V
Chip size	4.6 mm x 4.4 mm	8.9 mm x 8.9 mm	9.4 mm x 9.5 mm	13.0 mm x 13.0 mm	13.5 mm x 13.5 mm
Package size	12.7 mm x 12.7 mm	22.5 mm x 22.5 mm		17.0 mm x 17.0 mm	
Package type	84-pin QFP	180-pin QFP		462-pin PGA	

QFP: Quad-line Flat Package PGA: Pin Grid Array package

462-pin high-density pin grid array (PGA) package with low thermal resistance are used[3].

Computer aided design (CAD) has been used in all LSI development and fabrication, including optimization based on estimated performance, mask data generation, and test data generation. The use of CAD helped reduce the time required for LSI development.

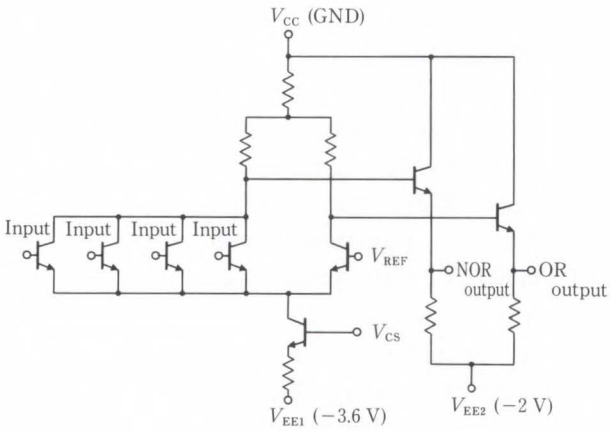
2.1 15 000-gate ECL gate array

2.1.1 Overview

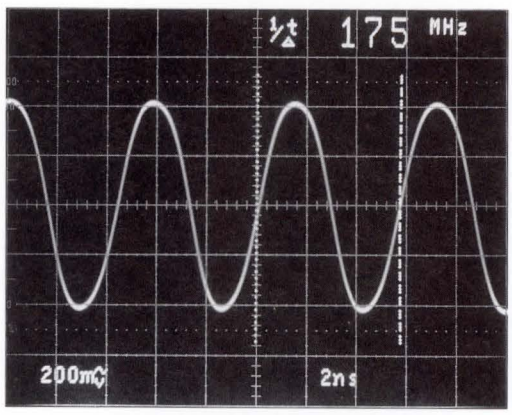
Previous Fujitsu computers have used gate arrays with integration levels from 100 gates to 3 000 gates and speeds from 700 ps to 180 ps[4),5]. The VP2000 series features a speed of 70 ps and an integration level of 15 000 gates. Table 1 lists the evolution of ECL gate arrays used in Fujitsu's computers. This table shows improvements in the array's speed, expressed by gate delay, and integration level, expressed by the number of gates. Compared with the FUJITSU VP-400 and the FUJITSU M-780, the array in the VP2000 is 5 times faster and 38 times more integrated, and 2.5 times faster and 5 times more integrated, respectively. These improvements mainly depend on the wafer process technology outlined in section 2.3

2.1.2 Basic ECL circuits

For the logic circuits, ECL was adopted for its superior speed characteristics and strong logic functions. The basic logic function block is a four-input OR/NOR gate, and the corresponding logic circuit is shown in Fig. 1a). The power



a) Basic ECL circuit



b) Output waveform of ring oscillator

Fig. 1—Basic circuit and ring oscillator waveform.

supply voltage to the circuit is −3.6 V and −2.0 V. One internal circuit typically consumes 1.8 mW and the signal amplitude is 550 mV. The delay times from the input to the complementary outputs, i.e. OR and NOR outputs, are nearly equal. The average *T*pd (propagation delay time) of the basic circuit is 70 ps.

Figure 1b) shows an example of the observed output waveform from a 41-stage ring oscillator on a fabricated ECL LSI. The oscillation frequency is 175 MHz, corresponding to 70 ps per circuit.

2.1.3 Structure

As explained in the preceding section, the basic circuit of the array is a four-input OR/NOR ECL gate designed to enable the integration of 14 976 circuits. Additionally, it has 256 output drivers for driving low-impedance inter-LSI device transmission lines. The threshold voltages of the internal circuits and output drivers are equalized to eliminate the need for input converters. The array has 320 signal pins, each of which can be assigned to an input, and/or an output, as needed.

The array has four-metallization layers. Layers 1 and 2 are mainly used for forming logic function blocks and/or interconnecting between the logic function blocks. Layers 3 and 4 are mainly for power distribution. The metallization pattern in layers 1 and 2 and through-hole placement between the two layers vary with the logic functions required by the system logic design. The other layers are common to all customized logic functions. About 2 200 routing channels in layer 1 and about 2 700 channels in layer 2 are provided for interconnecting the logic function blocks. This enables a nearly 100 percent circuit utilization rate by CAD routing system. Thus, ECL arrays with both high density and high power dissipation were made in a relatively small chip size. Because the interconnection length between the function blocks is short in the high density chip, the speed of loaded gate delay is fast.

Figure 2 shows the 13 mm square array chip. The pads on the chip periphery are gold bumps to enable TAB. Power comsumption is 30 W per LSI array.

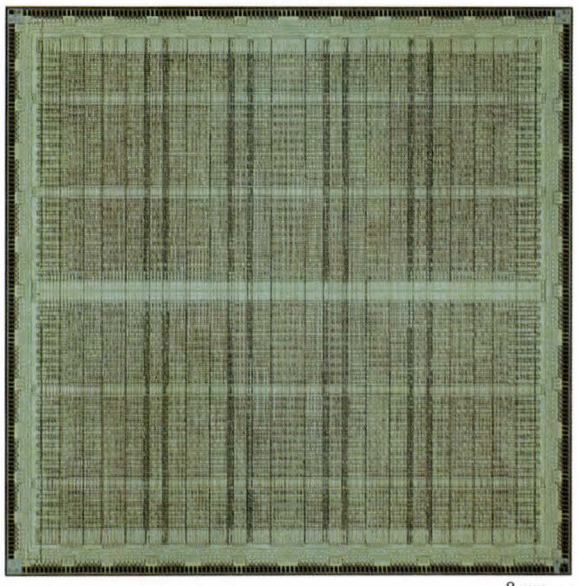


Fig. 2—15 000-gate logic array chip.

2.2 3 500-gate ECL array with 1.6 ns 64-Kbit STRAM[6]

2.2.1 Overview

A gate array with RAM combining high-speed bipolar RAM and logic array onto one chip was first developed for M-780[5),7)]. The advantages of this composite gate array over separate RAM and logic gate array include:

1) Reduced inter-chip wiring dealy
2) No output buffer of RAM required
3) Freely configurable RAM
4) Improved packaging density

These advantages help improve system performance and reduce power consumption.

For the VP2000, an improved gate array with RAM was developed. Compared with the array for the M-780, it features twice the speed, four times the memory size, and a simplified timing-controlled RAM called self-timed RAM (STRAM).

2.2.2 Embedded RAM

The RAM in a composite gate array is sometimes called an embedded RAM. The RAM for the VP2000 consists of sixteen 4-Kbit RAM macros (256 words by 16 bits each) as shown in Fig. 3. In contrast to conventional RAM, this RAM has latch circuits for all inputs and contains write pulse and internal clock pulse generators. The RAM is, thus, clock-controlled STRAM.

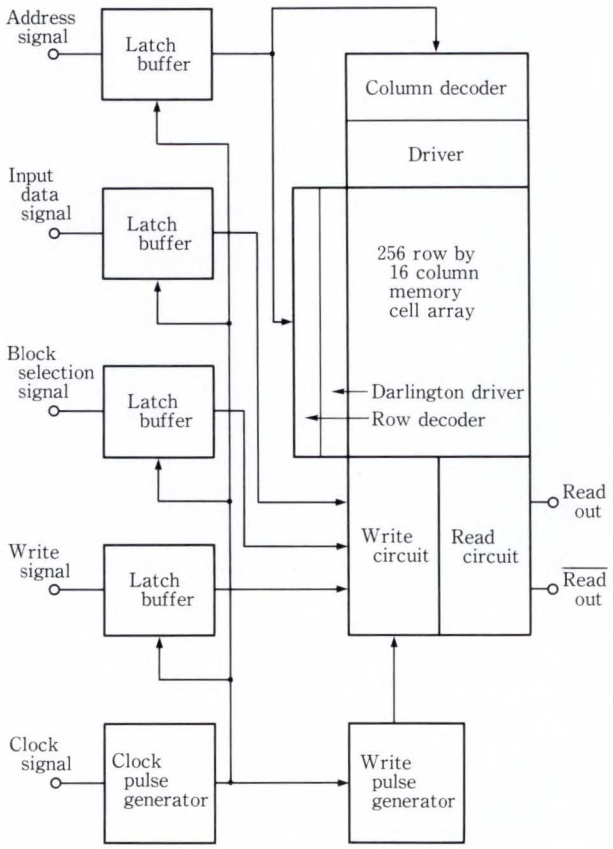


Fig. 3—Block diagram of STRAM macro in gate array with RAM.

STRAM features:

1) Simplified timing control because RAM is controlled only by the clock.
2) An increased number of gates available to the user because the peripheral circuit which would be implemented in the gate array section in conventional RAM is built within the STRAM.

The STRAM uses three new techniques: the combination of an input latch and an input buffer circuit into one circuit, Darlington word drivers, and complementary RAM outputs. These new features have shortened the address time to 1.6 ns at its maximum. This, together with the 1.95 ns maximum clock access time, shortens machine cycles.

2.2.3 Structure

The composite array chip (see Fig. 4) provides 64 Kbits of high-speed STRAM and contains 3 472 logic circuits and 208 output buffers. About 700 000 device elements, including NPN

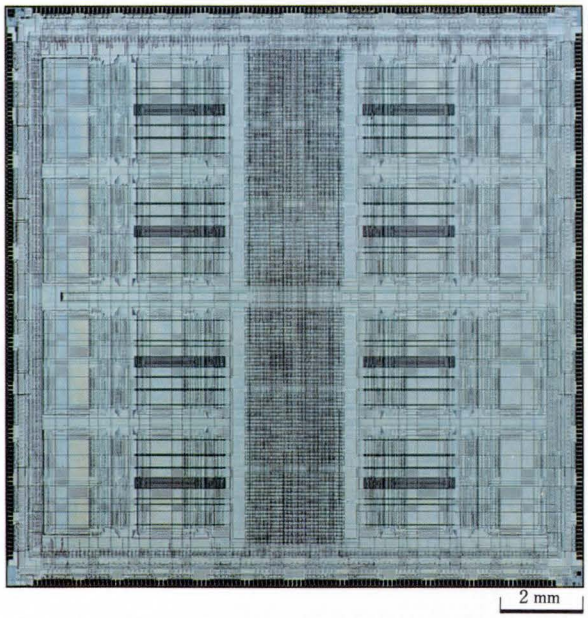


Fig. 4—3 500 gate array with 64-Kbit STRAM chip.

bipolar transistors, Schottky barrier diodes, capacitors, and resistors, are integrated on a 13.5 mm x 13.5 mm chip.

Three power supply voltages are used to reduce power consumption: −4.8 V, −3.6 V, and −2.0 V for STRAM, the −3.6 V and the −2.0 V for the logic circuits. The chip dissipates 30 W. In the composite array chip, the embedded RAM is composed of two RAM sections which are placed at either side, and the logic array section at the center.

Fixed segments of signal wiring are placed on the periphery of the chip and over the RAM section in the three-metallization layer to connect the logic section with the signal pads near the RAM section.

2.3 Wafer process technology

Process technology has played an important role in improving device integration density and speed. The previous ECL arrays for the M-780 used U-grooved isolation with thick field oxide (U-FOX) transistors with a minimum emitter width of 0.8 μm and three-layer metallization with 4 μm pitch signal wiring channels. For the new arrays used in the VP2000, more finely scaled, sophisticated technologies have been developed. The main technologies are ESPER transistors with a minimum emitter

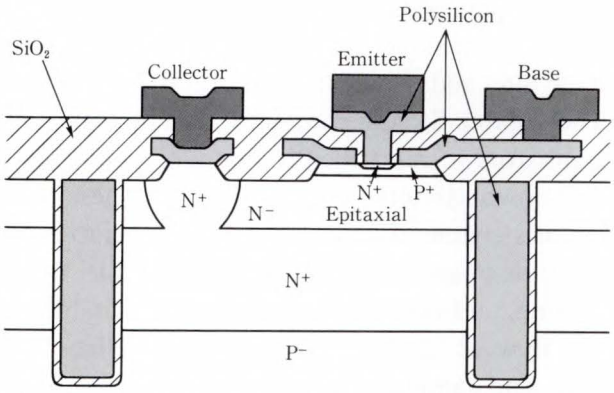
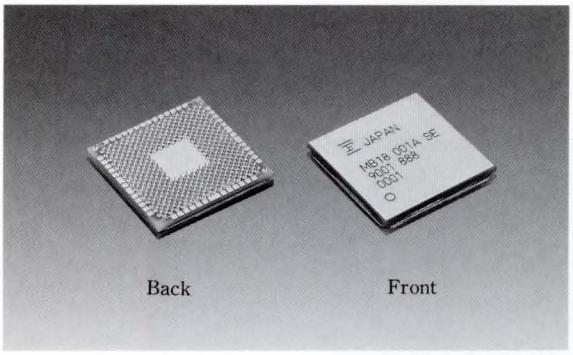
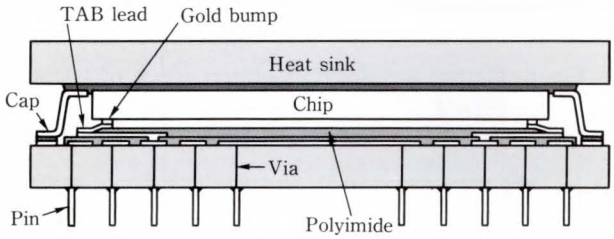


Fig. 5—Schematic cross-section of an ESPER transistor.



a) ECL gate array



b) Schematic cross-section

Fig. 6—ECL gate array and its schematic cross-section.

width of 0.3 μm[1], four-layer metallization with 2.6 μm pitch channels, and gold bumps for 100 μm-pitch TAB techniques. Typical of these new technologies is the ESPER structure which is outlined in the following.

By simulating ECL circuits, important device parameters sensitive to T_{pd} were found to be base-collector junction capacitance (C_{cb}), base resistance (r_B), and cutoff frequency (f_T).

Figure 5 shows the schematic cross-section of the ESPER transistor. This stacking double polysilicon structure for the base and emitter electrodes enables a dramatic reduction in the base-collector junction area and the distance between the base and emitter electrodes. Thus, smaller C_{cb}, and higher f_T can be obtained without increasing r_B or further with decreasing r_B. The f_T is about 15 GHz, a figure twice that of the U-FOX transistor.

2.4 Packing technology

The LSI package is a 462-pin, high-density PGA with a heat sink[3] (see Fig. 6). The size is 17 mm square, and pins are arranged in a zigzag pattern on 0.64-mm centers. The gold bumps on the chip are connected to the package substrate through the 100 μm pitch leads for TAB. The package substrate and heat sink are made of aluminum nitride, a material that offers superior thermal conduction qualities and a thermal expansion coefficient close to that of silicon. The package is hermetically sealed by the cap and heat sink.

3. 1-Mbit high-speed static RAM

3.1 Overview

Computers have long used DRAM for main storage, but to improve system performance, high performance computers and supercomputers increasingly are using SRAM for main storage. The VP2000 series uses high-speed 1-Mbit SRAM configured as 256K-words by 4 bits, and with a maximum access time of 35 ns. Compared with the 256-Kbit SRAM (a maximum access time of 55 ns), used for the M-780, this SRAM is superior both in density and access speed. These SRAM are packaged in 32-pin leadless chip carriers for greater package density.

3.2 Circuit technology

In general, the larger the memory, the weaker the signal read back from the memory cell, so it is important to detect and amplify such small signals quickly to reduce speed. The SRAM chip uses fine patterning and the following techniques for efficient transistor action: The memory cells are divided into eight blocks to reduce bit-line capacitance and word-line delay factors which cause slow operation.

This also greatly reduces power consumption because only one of the eight blocks is operating

during a given memory cycle. Fujitsu also added a substrate bias generator which biases the P-type substrate to about −2.5 V. This improves the N-channel transistor characteristics and reduces junction capacitance, ensuring high-speed operation.

Figure 7 shows the three-level differential sense amplifier which detects the small cell signals. One amplifier is assigned to each block and placed near the data bus, from which it receives inputs to reduce bus line capacitance. This configuration is thus suitable for quick sensing. The output of the third stage is simplified to nearly the CMOS level so that subsequent blocks can be selected easily. The sensing speed through the three stages is 4 ns. The sense amplifiers operate from a wide V_{CC} range (3 V to 7 V). This ensures both high speed and stable operations. A well designed address signal transition detector contributes to high-speed operation.

For asynchronous SRAM, the key to high-speed operation is quick erasure of records of previous memory cycles. That is, the bit line, data bus, and sense amplifier must all be reset. The 1-Mbit SRAM combines N- and P-channel transistors to reset the bit lines with appropriate pull-down circuits to reset the bus. These reset circuits enable efficient high-speed operation. A multiple-phase reset clock enables stable, quick sensing, and contains different delay suitable for the flow of signals read from the bit line, data bus, and amplifier's first, second, and third stages.

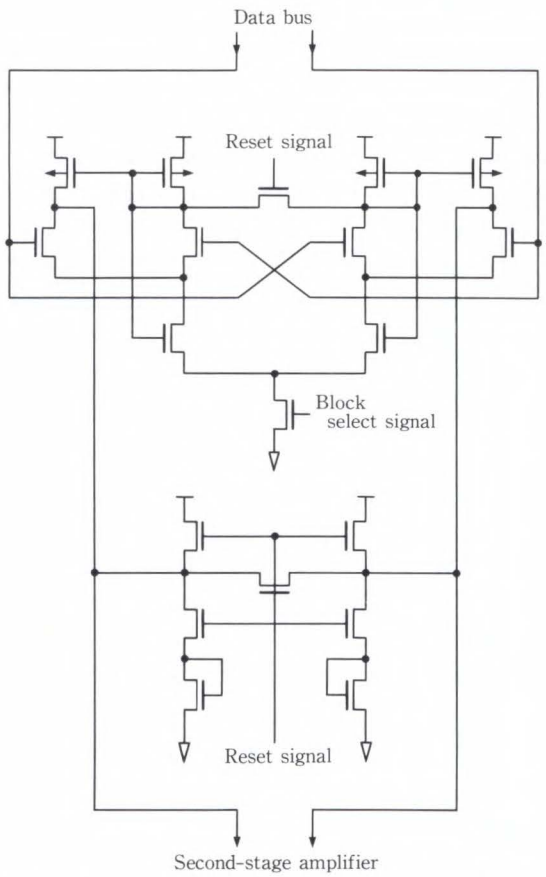


Fig. 7—Sense amplifier circuit of 1-Mbit SRAM.

3.3 Electrical characteristics

Figure 8 shows the data output waveform

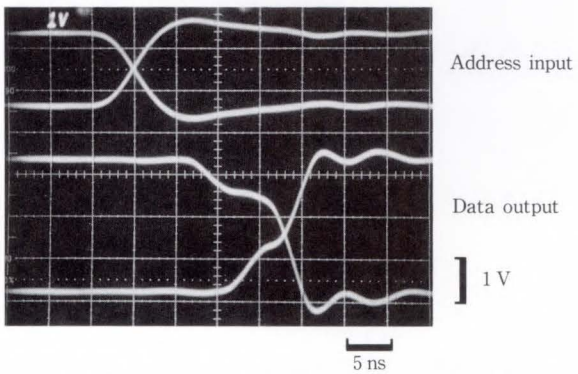


Fig. 8—Output waveform of address access of 1-Mbit SRAM.

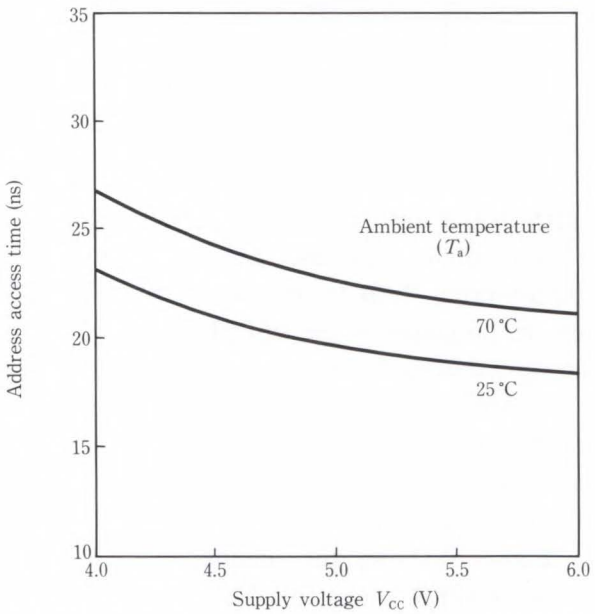


Fig. 9—Supply voltage dependence of access time of 1-Mbit SRAM.

obtained during access with 5 V (V_{cc}) and an ambient temperature of 25 °C. Under these conditions, the access time is about 20 ns.

Figure 9 shows the supply voltage dependence of the access time. Values in the graph are adequate to achieve the target access of 35 ns. Power consumption during operation is about 400 mW at a cycle time of 35 ns, V_{cc} = 5 V, and T_a = 25 °C. The stand-by power consumption is about 10 mW.

3.4 Process technology

To reduce memory cell size and raise performance, we used CMOS technology with three-layer polysilicon and two-layer metallization process. The minimum design rule is 0.8 μm. P-channel transistors are built up on N-type regions (N wells) formed on the P-type substrate. N-channel transistors for peripheral circuits are built up on the substrate. Memory cells are built up within the P wells to reduce the likelihood of soft errors. The gate length is 0.8 μm for N-channels and 1.1 μm for P-channels. The gate film is 20 nm thick.

A memory cell consists of four N-channel transistors and two load resistors. Power supply distribution to load resistors and cells is provided by polysilicon in layer 3 and distribution lines of V_{ss} are provided by polysilicon in layer 2. The layer 1 metal is used for the bit line, a cell selection line, and the layer 2 metal is used to reduce the line resistance of the polysilicon word line which is the other selection line. This reduces word line delay to a negligible level.

For the peripheral circuits, two-layer metallization helps increase device speed by reducing signal delay. We have, therefore, reduced the memory cell size to 4.8 μm by 8.5 μm, so that the chip (see Fig. 10) measures 7.5 mm x 12.0 mm.

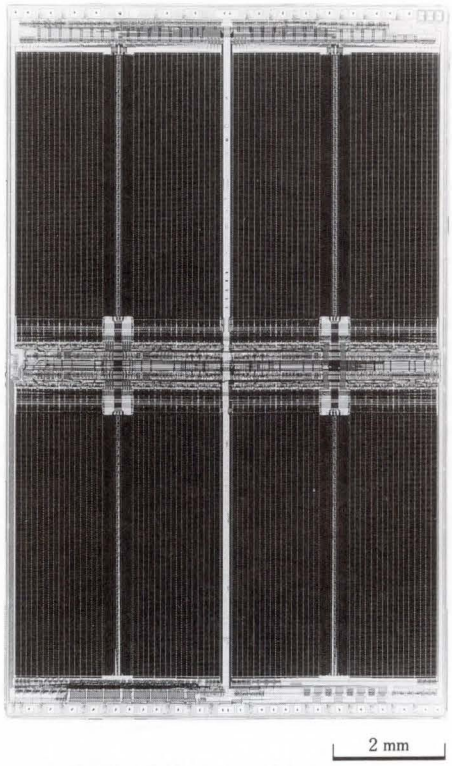


Fig. 10—1-Mbit SRAM chip.

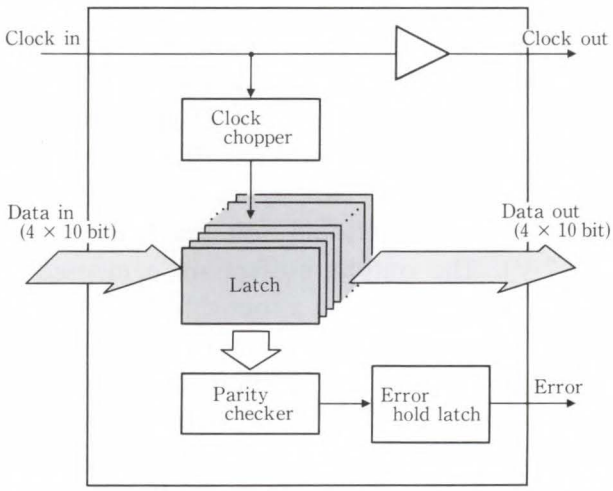


Fig. 11—Block diagram of GaAs line driver/receiver LSI.

4. GaAs line driver/receiver LSI

4.1 Overview

To improve system performance of supercomputers, it becomes important to increase the data transfer rate between system storage units and main storage units. Fujitsu developed an ultrahigh-speed GaAs line driver/receiver LSI which performs data transmission between the operational units in conjunction with ECL LSIs.

4.2 Structure and process

The GaAs line driver/receiver LSI with ECL compatible input output interface functions as either an off-board driver or a front-end receiver from another board. Figure 11 is the block diagram of the circuit. It consists of input buffers for ECL to GaAs level translation,

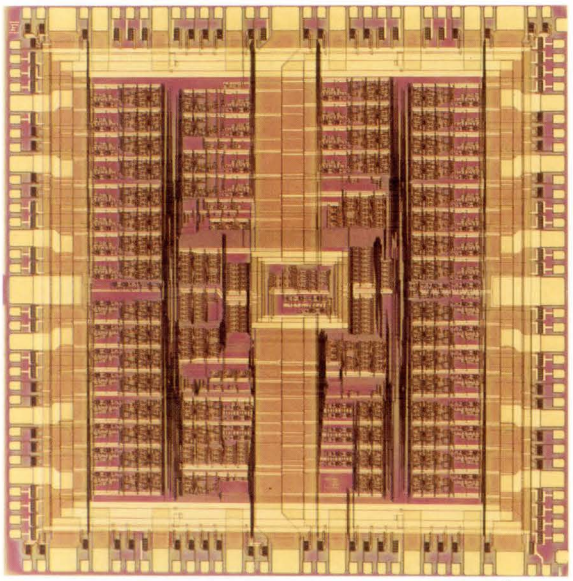


Fig. 12—GaAs line driver/receiver chip.

40-bit latches for data retiming and retention, output buffers for driving 50-ohm transmission lines, and peripheral circuits such as a clock distributor and a parity checker. The total gate count is about 1 200.

The input buffers consist of level shift circuits and a differential amplifier with on-chip ECL reference voltage (−1.3 V) generator. The internal circuits are buffered FET logic (BFL) which consist of two types of D-MESFETs with different threshold voltages (−0.3 V and −0.7 V). The output buffers are composed of a BFL driver stage and a source follower output stage. The chip is designed to minimize the difference in the propagation delay from the clock input to each data output. The chip contains 6 032 FETs, 1 417 diodes and 24 resistors, measure 6.16 mm × 6.26 mm and is surrounded by 100 signal and 32 power supply bonding pads. Twenty-four of the power supply pads are assigned to ground to suppress the current switching noise of the output stages.

Figure 12 is a microphotograph of the line driver/receiver LSI. Tungsten-silicide gate self-alignment MESFET and two-level gold wiring technologies are used to fabricate the LSI. The gate length is 1.2 μm. The line width/pitch is 2.5 μm/5 μm for layer 1 and 3.5 μm/7 μm for layer 2, respectively. The via size is 2 μm × 2 μm.

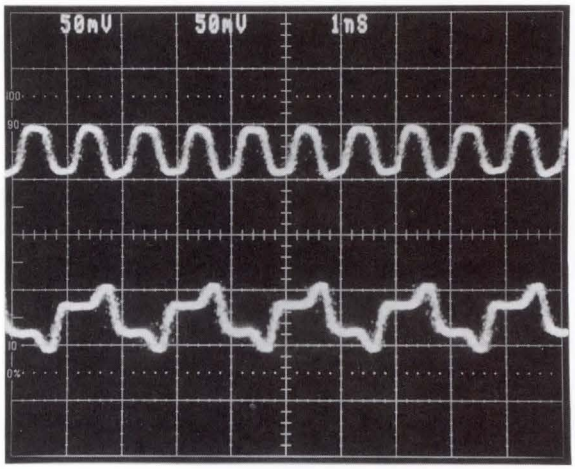


Fig. 13—Output waveform of line driver/receiver at 1 Gbit/s clock operation.

4.3 Characteristics

A complete function test was performed by an LSI tester at a 10 MHz clock frequency. The test pattern was developed by CAD system. The number of patterns was 10 412 and the testability of the pattern was 100 percent. The chip consumed 5.5 W power dissipation with the supply voltage of −2 V and −3.6 V. For the internal circuits, the gate delay time of 60 ps with 5 mW/gate power dissipation was obtained from the 51-stage ring oscillator experiment. Typical propagation delay time from the clock input to data output path was 1.3 ns.

Within the ten percent excursion of the supply voltages and the temperature range between 0 °C and 85 °C, the chip operated successfully. AC characteristics were also measured by using a pulse generator and a sampling oscilloscope. One output from a pulse generator was connected to the clock input and another output at half the frequency was connected to one of the data inputs of the chip. The clock output and one of the data outputs were monitored by sampling oscilloscope as shown in Fig. 13. Because a clock chopper circuit was adopted, stable operation of the latches was observed over a wide frequency range. The maximum clock rate was 1 Gbit/s.

5. Conclusion

The main semiconductor devices used for FUJITSU VP2000 series supercomputers are

outlined in this paper.

Two type of ECL gate arrays were developed based on an advanced silicon bipolar technology and a high-density packaging technique.

One is a 15k-gate ECL logic array with a gate propagation delay time of 70 ps. The other is a 3 500 gate 70 ps ECL composite array with 64-Kbit STRAM with a maximum access time of 1.6 ns. These arrays are more than two times faster and four times more integrated than the previous arrays used for FUJITSU M-780. These improvements resulted from the introduction of new techniques including ESPER structure, four-layer metallization, STRAM circuits, TAB, and a high-density 462-pin package.

Highly integrated 1-Mbit SRAM with a maximum access time of 35 ns was developed using silicon CMOS technology. The previous SRAM is 256-Kbit SRAM with a maximum access time of 55 ns. To improve the density and speed, the 1-Mbit SRAM uses new techniques including three-layer polysilicon and two-layer metallization process, 0.8 μm scaled process, and an amplifier circuit with higher sensitivity.

An ultrahigh-speed line driver/receiver with a 60 ps gate delay and 1 200 gate counts was developed using GaAs technology, which features tungsten-silicide gate self-alignment MESFET with 1.2 μm gate length and two-layer gold metallization. It is the first time that the Fujitsu's supercomputer has used GaAs LSIs.

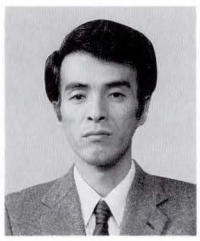
Semiconductor technologies for both silicon and GaAs devices will continue to progress rapidly in the future. And supercomputers and mainframe systems require ever higher performance.

Fujitsu will continue its efforts to develop and provide higher performance LSIs for future systems.

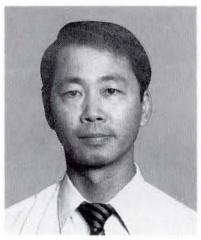
References

1) Deguchi, T., and Goto, H.: Ultra High-Speed Bipolar Process Technology: ESPER. *FUJITSU Sci. Tech. J.*, **24**, 4 (Special Issue on Semiconductors), pp. 379-383 (1988).
2) Ohno, K., Ooami, K., and Itoh, H.: Semiconductor Technologies for FUJITSU VP2000 Series. (in Japanese), *FUJITSU*, **41**, 1 (Special Issue: FUJITSU VP2000 Series), pp. 20-26 (1990).
3) Harada, S., Sugimoto, M., Matsuki, H., Sumi, Y., Ueno, S., and Murakami, S.: Fine-Pitch High-Density PGA Package. IEEE ISSCC Dig. Tech. Papers, 1990, pp. 154-155.
4) Ohno, K., and Takeda, H.: High-Speed Bipolar Logic IC. *FUJITSU Sci. Tech. J.*, **24**, 4 (Special Issue on Semiconductors), pp. 265-270 (1988).
5) Ohno, K., Tanaka, M., Noguchi, E., Ono, T., Yabu, T., and Sugimoto, M.: Semiconductor Technologies for FACOM M-780. *FUJITSU Sci. Tech. J.*, **23**, 4 (Special Issue on Computer Systems), pp. 216-225 (1987).
6) Kimoto, M., Shimizu, H., Itoh, Y., Kohno, K., Ikeda, M., Deguchi, T., Fukuda, N., Ueda, K., Harada, S., and Kubota, K.: A 1.4 ns/64 Kb RAM with 85 ps/3680 Logic Gate Array. IEEE CICC Proc., 1989, pp. 15.8.1-15.8.4.
7) Sugo, Y., Tanaka, M., Mafune, Y., Takeshima, T., Aihara, S., and Tanaka, K.: An ECL 2.8 ns 16 K-RAM with 1.2 K Gate Array. IEEE ISSCC Dig. Tech. Papers, 1986, pp. 256-257.

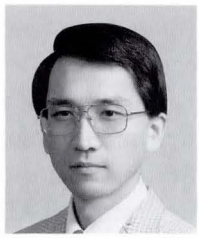
Ken-ichi Ohno
Advanced Gate Array Design Dept.
Gate Array Div.
FUJITSU LIMITED
Bachelor of Electric Eng.
Kyoto University 1967
Master of Electronics Eng.
Kyoto University 1969
Specializing in Gate Array Design



Hideo Itoh
ROM Design Dept.
Memory Div.
FUJITSU LIMITED
Bachelor of Electric Eng.
Iwate University 1969
Dr. of Electronics Eng.
Tohoku University 1975
Specializing in MOS Memory Design



Kazuo Ooami
Advanced Gate Array Design Dept.
Gate Array Div.
FUJITSU LIMITED
Bachelor of Applied Physics Eng.
Science University of Tokyo 1970
Specializing in Bipolar RAM Design



Katsuhiko Suyama
GaAs LSI Design Dept.
FUJITSU LIMITED
Bachelor of Electric Eng.
Keio University 1970
Master of Electric Eng.
Keio University 1972
Specializing in High Speed
LSI Design

UDC 621.3.049:681.32

Multilayer Glass-Ceramic-Composite Circuit Board for FUJITSU VP2000 Series

• Koichi Niwa • Yukichi Takeda

(Manuscript received December 4, 1990)

A new glass-ceramic-composite material has been developed and applied to the circuit board of the FUJITSU VP2000 series supercomputer. The composite material exhibits low dielectric constant and low thermal expansion coefficient. These two characteristics are capable of satisfying both present and future requirements of high-speed signal propagation and high density packaging of LSIs. Copper conductors are used for the circuit wiring, which yields low resistivity conductor wiring for as many as 61 layers. The gate density of the VP2000 series is ten times higher than that of conventional organic printed circuit boards.

1. Introduction

The demand for high-speed signal propagation in computer and communications systems has prompted dramatic progress in the technology of LSIs and circuit boards. To support high-speed switching of LSIs in a system, the circuit boards should have a low dielectric constant so that high-speed signals can propagate with a shorter delay. The low sintering temperatures of the circuit boards are also desirable so that the boards can be co-fired with high electrical conductivity materials such as copper or gold. Since high density bare chip packaging is also taken into consideration, the thermal expansion of the boards must be close to that of silicon chips to avoid chip breakage[1].

This paper describes the importance of ceramic materials having low dielectric constant and low firing temperature for the future of high-speed LSI circuit boards, with specific reference to conventional organic printed circuit boards and ceramic multilayer circuit boards. A new ceramic material composed of borosilicate glass and alumina[2] has been developed. The use of this composite material has successfully lowered the dielectric constant and optimum firing temperatures.

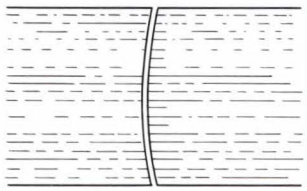
Multilayer circuit boards employing this composite insulator and using copper conductors has been applied to the FUJITSU VP2000 series supercomputer.

2. Conventional circuit boards

2.1 Printed circuit board

The circuit boards used in high-speed computer systems can be divided roughly into two types: organic printed circuit boards and alumina multilayer circuit boards.Organic printed circuit boards have a low dielectric constant of about 5, and the electrical conductivity of the laminated copper is excellent. This combination of insulation with low dielectric constant and copper conductors are adequate to satisfy the future electrical requirements. In the near future, however, highly integrated LSIs will require much more complex wiring in mutilayer printed circuit boards. To satisfy this requirement, organic printed circuit boards would be required to have increasingly large numbers of layers.

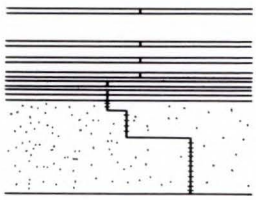
The through holes, which provide the electrical contact among the layers, are made by drilling holes through the laminated organic



High aspect ratio drilling

Through hole: After lamination

a) Through hole



Lamination of green sheets

Via: Before lamination

b) Via

Fig. 1—Formation of the through hole and via.

board. For fine, dense patterns, a small diameter drill bit is used. As the boards become thicker, however, the aspect ratio becomes larger and the drill bit tends to curve; this results in aberration from the nominal through hole point. Organic circuit boards include glass fiber to strengthen the board and reduce its thermal expansion, and the glass fiber, which is much harder than the surrounding polyimide or epoxy materials in the circuit board, often causes the bits to break during through hole drilling.

Organic circuit boards also present another difficulty: that of thermal expansion. In high density packaging, the direct mounting on the circuit boards of bare chips will soon become necessary. Addition of glass fiber can reduce the thermal expansion of organic materials, but it causes aberration of the through hole points and drill bit breakage {see Fig. 1 a)}.

2.2 Alumina multilayer circuit board

A multilayer ceramic circuit board has been used with high-speed circuits. The ceramic circuit board offers superior characteristics in that the process of multilayering is easier than with organic printed circuit boards, and the thermal expansion of the board is much closer to the characteristics of bare silicon chips. Heat resistance and thermal stability of the dimensions are also superior to those found in organic boards. Ceramic multilayer is manufactured in using green sheets. Via, by which electrical connections are made among layers of the patterns, are formed on the green sheet. First, holes for vias are punched in the green sheets, then the holes are filled with conductor paste. A number of green sheets with via and patterns are laminated and fired. The ease of forming via does not vary with the number of layers {see Fig. 1 b)}. Very fine diameter via are easy to prepunch in the thin green sheets, as opposed to the through holes for conventional organic boards which are formed by drilling the thick laminated boards. Because of this advantage, alumina multilayer circuit boards were used for high density, fine pattern circuits.

Table 1. Comparison between organic printed circuit board and alumina multilayer circuit board

Circuit board	Advantages	Disadvantages
Organic	Low K High conductivity	Limit for the number of layers High thermal expansion
Ceramic	No limit for the number of layers Low thermal expansion	High K Low conductivity

The dielectric constant K of alumina multilayer circuit boards, however, is high compared with organic materials. The resistivity of the patterns is also higher than the copper that is used in organic printed circuit boards (see Table 1). To satisfy the demand for hgih-speed LSI packaging, it was necessary to develop a new ceramic circuit board offering lower K and high electrical conductivity.

3. New ceramic circuit board

The requirements imposed on ceramic cir-

Table 2. Dielectric constant K, and softening point of glass

Glass	K (1 MHz)	Softning point (°C)	TEC ($\times 10^{-6}$/°C)
Borosilicate	4.1-4.9	700-850	3.0-4.8
Aluminum silicate	6.3	910	6.3
Barium borosilicate	5.8	840	4.6

TEC: Thermal expansion coefficient

Table 3. Properties of several ceramics

Ceramic	K (1 MHz)	TEC ($\times 10^{-6}$/°C)	Flexural strength (MPa)
Cordierite	5.3	2.2	100
Alumina	9.9	6.8	300
Aluminum nitride	8.9	4.4	400
Mullite	6.5	4.0	180
Forsterite	6.0	10.0	140
Zirconia	13.0	10.0	400

cuit boards is lower K and higher electrical conductivity. Glass material exhibits low K, and it becomes soft at low temperature. In the ceramic multilayer circuit board, the conductor must have a higher melting point than the optimum firing point T of the board, because the green sheets and the conductor patterns printed on the sheets are fired together. Alumina is usually fired at about 1 600 °C, so proper conductors that may be used with alumina are limited to materials with high melting point such as Mo or W. Regrettably, these conductors offer higher resistivity than either Cu or Au.

Glass itself has a low K and low softening point (see Table 2); borosilicate glass, in particular, combines low K and low softening point with high chemical stability. Furthermore, the properties are easy to control by altering the B_2O_3/SiO_2 ratio.

If glass is formed in a green sheet and then fired, the glass will shrink into a ball because of the surface tension when the glass becomes soft. If ceramic material is placed in a glass body, however, it will act as filler that inhibits curling of the formed sheet and strengthens the body.

Composition of glass-ceramic system, thus, is important in creating low K and low T substrate materials without shape distortion. But in the glass-ceramic composite, it is crucial to prevent crystallization during firing because the crystallized phase would clearly change the properties of the glass-ceramic-composite.

4. Glass-ceramic-composition

Tests were conducted with borosilicate glass, focusing on the crystallization of glass during firing, to determine a suitable combination for low K, low T sintering materials for future circuit boards.

4.1 Glass-cordierite system

Cordierite was chosen as the ceramic material for the initial examination. Cordierite has the lowest dielectric constant K listed in Table 3. In addition, the low thermal expansion of this ceramic is close to that of silicon.

Glass powder and ceramic powder were milled with binder and solvent for 20 hours in a plastic mill pot. The slurry was cast 500 μm thick using the doctor blade method. These green sheets were then laminated, giving a thickness of 1 mm, then fired in a N_2 atmosphere at 1 000 °C for five hours. The heating rate was 200 °C/h.

The formation of new crystal in the glass-ceramic composition was examined using X-ray analysis; Cu target, 30 kV, 10 mA. The thermal expansion coefficient was measured using a push rod dilatometer. Silica glass was used as standard specimen. For measuring the thermal expansion coefficient (TEC), the heating rate was 5 °C/h. The temperature ranged from room temperature to 300 °C.

The glass-cordierite system specimen exhibits high TEC. The calculated TEC of the system is $(2\text{-}3) \times 10^{-6}$/°C, which is slightly lower than that of silicon. The TEC actually measured for that system, however, is 17×10^{-6}/°C, or about seven times higher than the calculated value. The TEC curve is generally almost straight as temperature increases, but on the thermal expansion curve, the straight lines separate in two regions: below 100 °C and above 200 °C. On the 100-200 °C portion, the TEC shows a very steep curve similar to that of cristobalite[3].

Cristobalite is a polymorphism of silica

having an α-phase below 100 °C and a β-phase above 200 °C. Rearrangement of α-phase to β-phase is reversible, and the volume ratio changes from 0.431 to 0.451. This results in a rapid change in the curve of thermal expansion. Figure 2 shows the cristobalite formed in a glass matrix.

4.2 Glass-alumina system

The same tests as described above were performed on the glass-alumina system. No large difference was observed between the TEC values calculated and those measured for the glass-alumina system. The X-ray diffraction pattern of glass-alumina revealed no formation of cristobalite.

To clarify the inhibitory effect of alumina on crystallization of borosilicate glass, the alumina powder was added to a glass-cordierite system. The addition of alumina powder lowered the incidence of cristobalite. Figure 3 shows the TEC curve of the glass-cordierite system with alumina. The steep change in thermal expansion between 100-200 °C became more gentle, and the TEC decreased to 47 percent of that of the basic system. From these results, it is evident that the alumina can prevent the formation of cristobalite[4].

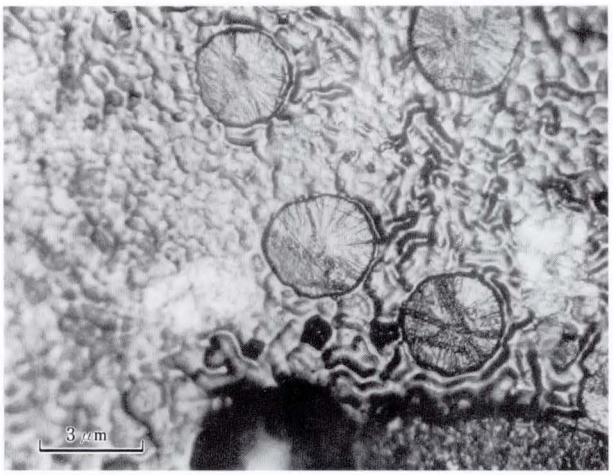


Fig. 2—Cristobalite crystal formed in glass matrix.

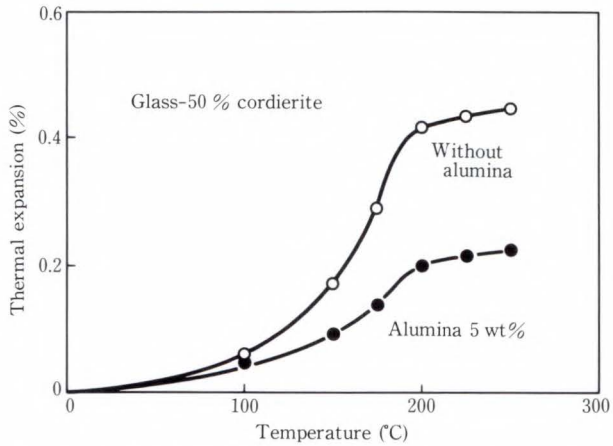


Fig. 3—Thermal expansion curve of glass-cordierite system and alumina added glass-cordierite system.

5. Properties of glass-alumina system

The densification mechanism of the alumina-glass system can be explained as liquid phase sintering. The alumina is uniformly dispersed in the glass matrix. Wettability between the glass and the ceramic powder is essential to densification. A proper amount of glass yields high density with high strength. An optimum firing point for the maximum densification also exists for any given glass-ceramic ratio. A combination suitable for co-firing with Cu or Au has an

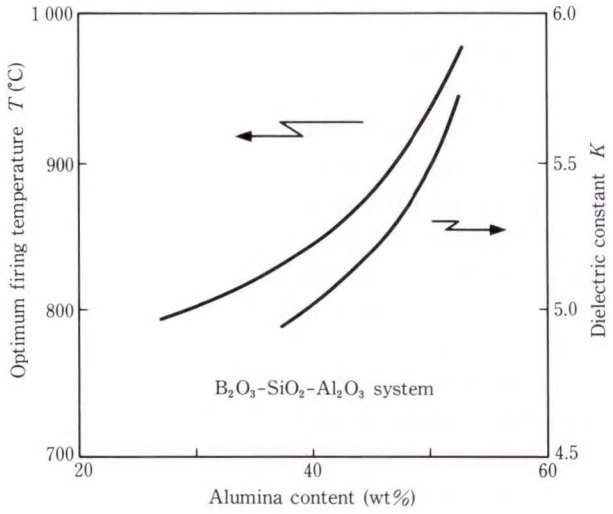


Fig. 4—Changes in K and T of glass-alumina system.

Table 4. Powder used for glass-alumina-composite material.

Powder	Specific gravity	Particle size (μm)	Specific surface area (m^2/g)
Cordierite	2.5	4.4	2.8
Borosilicate glass	2.2	4.5	4.2

alumina content between 40 wt% and 50 wt% (see Fig. 4).

The powder used for glass-alumina-composite materials is listed in Table 4.

6. Co-firing with copper

Co-firing with copper presents a major difficulty: burnout in an inert gas atmosphere. Copper wiring patterns are screen printed on the green sheets of glass-ceramic-composite, more than 60 of the green sheets are laminated, then they are fired. Tape-cast green sheets are generally fired in an air atmosphere to burn out the binder, but when copper wiring is used, an inert gas atmosphere must be used to avoid copper oxidation. The inert gas firing, however, causes a large amount of carbon residue in the fired multilayer body[5),6)].

To avoid carbonization of the binder, therefore, Fujitsu used a newly developed binder system that easily decomposes without oxygen when heated. The binder structure is shown in Fig. 5. Most binders generally used in ceramic boards show side chain reaction or random rupture, but the new binder shows depolymerization as illustrated in Fig. 5b). A 61-layer lamination of green sheets, however, makes the burning out of all binder difficult, so Fujitsu also developed a new firing process which can decrease the carbon residue in a fired body. If the binder and firing process are unsuitable for the co-firing of copper and glass-ceramic-composite, the final copper pattern will be oxidized as shown in Figs. 6a) and b), or a large amount of carbon will remain as shown in Fig. 6d). Figure 6c) shows Fujitsu's newly developed multilayer co-fired carbon- and oxide-free glass-ceramic circuit board.

Carbon residue exceeding 100 ppm clearly affects the densification of the fired body as well as the flexural strength. A fired body containing less than 30 ppm of carbon residue approaches 100 percent of theoretical density, but a specimen containing 1 000 ppm shows 93 percent of theoretical density. Carbon also affects the breakdown voltage; a specimen including more than 100 ppm of carbon shows zero kV/cm^2. Trapped carbon that is intermixed

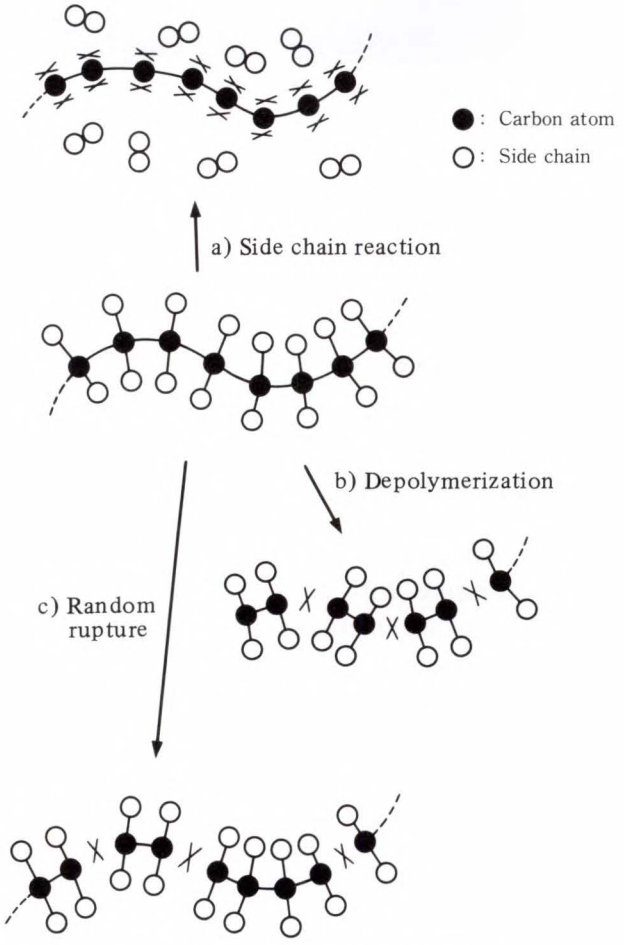


Fig. 5—The image of the binder.

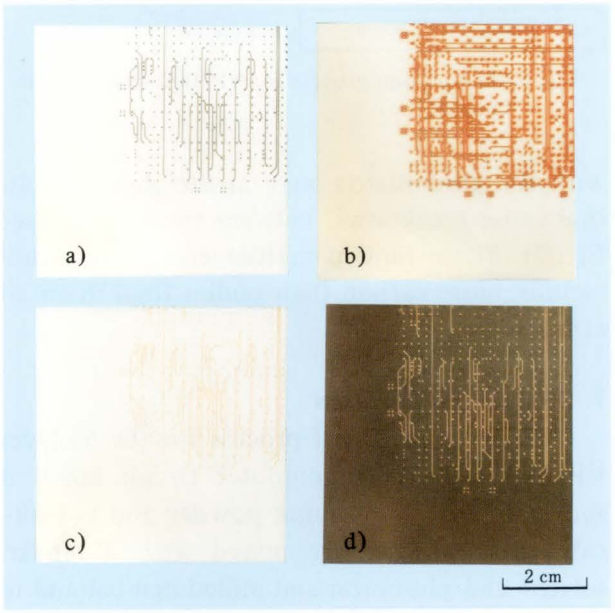


Fig. 6—Carbon residue and copper oxidation of co-fired multilayer board.

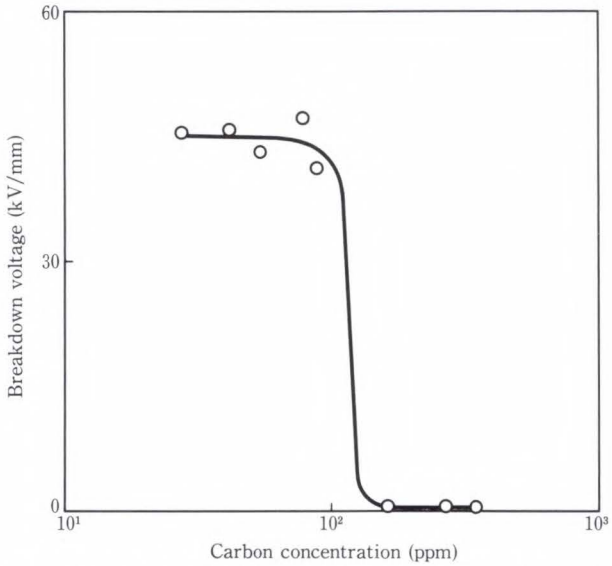


Fig. 7—The relation between carbon concentration and breakdown voltage.

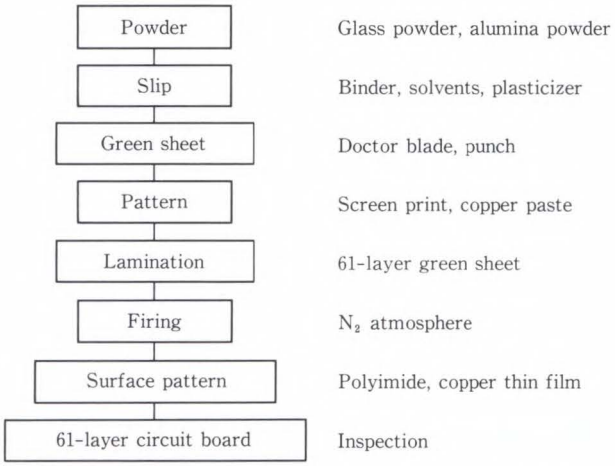


Fig. 8—Fabrication process for VP2000 series circuit board.

with the glass matrix may create short circuits that cause breakdown between the patterns (see Fig. 7). These Fujitsu multilayer bodies do not include more carbon than bodies fired in an air atmosphere.

7. Fabrication process

The manufacturing process for the 61-layer VP2000 series supercomputer circuit board is shown in Fig. 8. Alumina powder and borosilicate glass powder are mixed with a binder, solvent and plasticizer and milled in a ballmill to make a slurry. Slurry viscosity is adjusted before doctorblading by vaporizing solvent from the

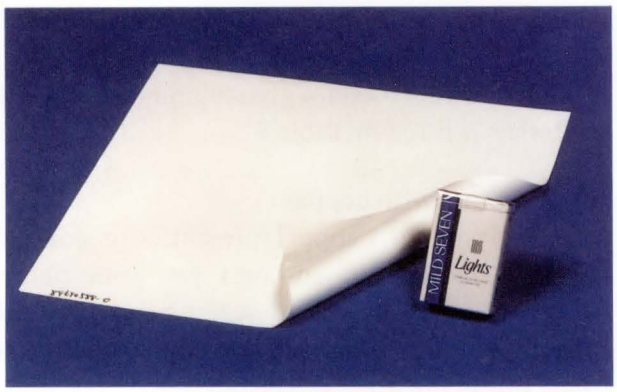


Fig. 9—Green sheet after dried.

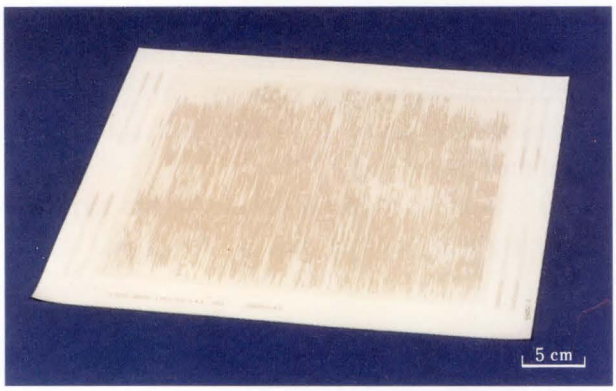


Fig. 10—Green sheet with copper pattern.

slurry. Tape-cast green sheets are dried in a continuous strip after doctorblading. The dried green sheet is cut to a proper size as shown in Fig. 9. Holes are punched for via, and copper paste is screen printed on the green sheets. The pattern width used for the VP2000 circuit board is 95 μm after firing. The paste viscosity, screen pressure on the green sheet, and velocity of squeegee are controlled for accurate pattern dimensions (see Fig. 10). More than 60 green sheets with copper paste wiring are laminated under pressure with heat. Before pressing, the green sheets are stacked, accurately adjusting via to via. The laminated green body is then fired in a nitrogen atmosphere. The surfaces of the fired multilayer body are polished to form a thin film layer of polyimide. The final process forms a thin polyimide layer with copper lands and via. Figure 11 shows the copper via jungle which is obtained by etching the glass-ceramic-composite body.

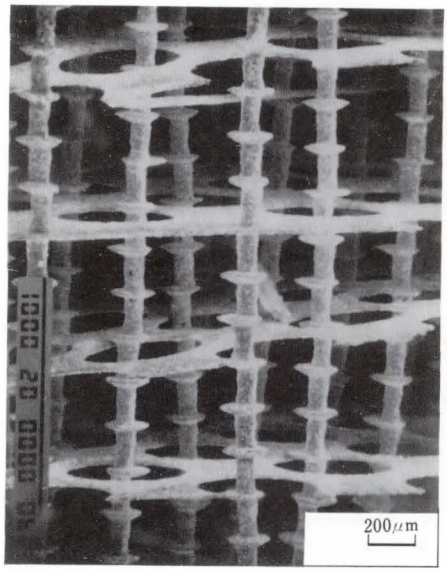
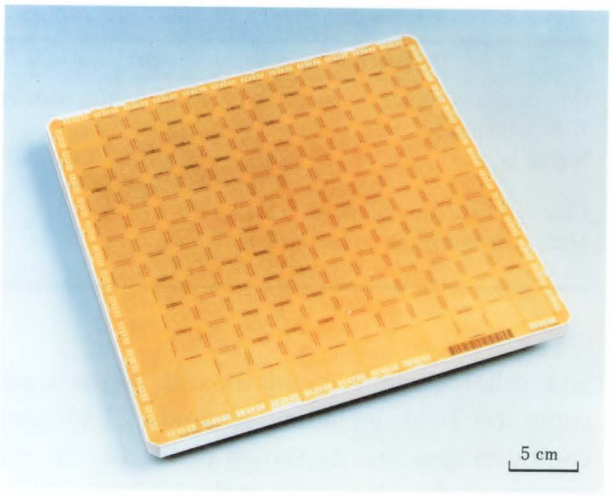
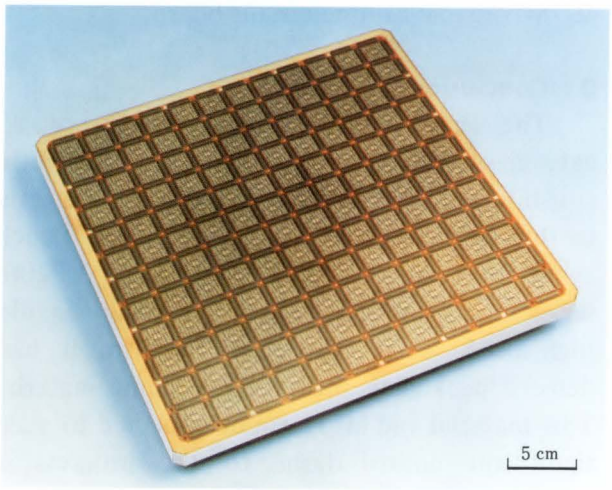


Fig. 11—Copper via jungle obtained by etching the multilayer body.

Fig. 12—Cross-sectional view of MLG.



a) Front side view



b) Back side view

Fig. 13—61-layer circuit board (24.5 cm x 24.5 cm).

8. VP2000 series circuit board

The multilayer circuit board used for the VP2000 series supercomputer is called Multi-Layer Glass ceramic circuit board (MLG). The maximum number of layers is 61, including 36 signal layers. About 40 000 signal patterns are formed on one layer, and the total wiring length of a signal pattern is about one kilometer. The size of the MLG is 24.5 x 24.5 cm and 13 mm thick. The characteristic impedance of the signal pattern is controlled to 65 Ω, and the resistance of the copper pattern is only 100 mΩ/cm. Two signal layers are sandwiched between ground and voltage layers to decrease the cross-talk noise and deviations in characteristic impedance.

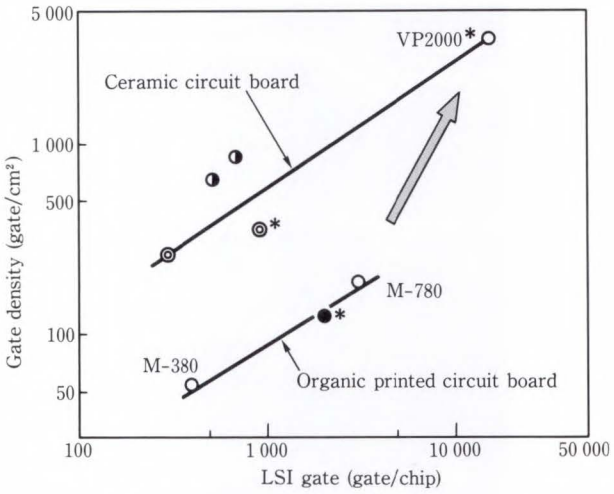
The thin surface layer of polyimide on the MLG is applied to adjust the shrinkage mismatch between the board and the nominal LSI surface terminations. A cross-sectional view of the MLG is shown in Fig. 12. On the front surface of the MLG, 144 LSI (12 x 12) chip terminations are formed on the thin polyimide layer {see Fig. 13a)}, and on the back surface, about 40 000 terminations for connector pins are formed {see Fig. 13b)}. A flange is brazed onto the MLG for attaching the cooling header and connector.

The properties of the MLG are listed in Table 5. The MLG made it possible to build a high density package that compares favorably with conventional organic printed circuit boards

Table 5. Properties of 61-layer VP 2000 circuit board

Properties		Specifications
Conductor: Copper	(mΩ/cm)	100 (95 μm width)
Wiring length	(m)	1 000
Signal layer		36 (40 000 wiring patterns)
Characteristic impedance	(Ω)	65
Dielectric constant	(1 MHz)	5.7
Grid	(mm)	0.45

and other ceramic circuit boards in the supercomputer field including large-scale, high-speed computer systems. As shown in Fig.14, the gate density of the MLG used in the VP2000 series is ten times greater than that of the FUJITSU M-780 organic printed circuit board.



*: Supercomputer
◑, ◎, ●: Systems of other computer companies

Fig. 14—Change in gate density of circuit board due to the LSI gate.

9. Conclusion

The ceramic material composed of borosilicate glass and alumina, with its low dielectric constant and low firing temperature, is expected to find application in a wide range of LSI packaging for many fields. The FUJITSU VP2000 series supercomputer system successfully enables high-speed signal propagation through high density packaging on a new ceramic material. This material and MLG can be expected to make significant contributions to the progress of packaging field for computer systems.

References

1) Niwa, K., Hashimoto, K., Kamehara, N., and Murakawa, K.: The Substrate for Large Si Chip and Full Wafer Packaging. Proc. NEPCON 1980, pp. 323-329.
2) Kamehara, N., Imanaka, Y., and Niwa. K.: Multilayer Ceramic Circuit Board with Copper Conductors. *Denshi Tokyo*, **26**, pp. 143-148 (1987).
3) Imanaka, Y., Yamazaki, K., Aoki, K., Kamehara, N., and Niwa, K.: Effect of Alumina Addition on Crystallization of Borosilicate Glass. *J. Ceram. Soc., Jpn.*, **93**, 3, pp. 309-313 (1989).
4) Imanaka, Y., Aoki, S., and Kamehara, N.: Thermal Expansion of Glass/Ceramic Composites for Multilayer Ceramic Circuit Board. *FUJITSU Sci. Tech. J.* **25**, 1, pp. 73-79 (1989).
5) Niwa, K., Kamehara, N., Yokoyama, H., Yokouchi, K., and Kurihara, K.: Multilayer Ceramic Circuit Board with Copper Conductor. *Advances in Ceramics*, **19** (Special Issue: Multilayer Ceramic Devies), pp. 41-47 (1987).
6) Niwa, K., Kamehara, N., Yokouchi, K., and Imanaka, Y.: Multilayer Ceramic Circuit Board with Copper Conductor. *Advanced Ceramic Materials*, **2**, 4, pp. 832-835 (1987).

Koichi Niwa
Materials Div.
FUJITSU LABORATORIES, ATSUGI
Bachelor of Physics Eng.
Chiba University 1964
Dr. of Eng. in Electronic Ceramics
Tokyo Institute of Technology 1988
Specializing in Electronic Ceramics

Yukichi Takeda
Printed Wiring Board Div.
FUJITSU LIMITED
Bachelor of Electrochemical Eng.
Tokyo Institute of Technology 1966
Specializing in Printed Wiring Board

UDC 658.512.2:681.32

Design Automation System for FUJITSU VP2000 Series

• Hirofumi Hamamura • Akihiko Hanafusa • Minoru Saitoh
• Toshihiko Tada

(Manuscript received November 28, 1990)

The performance of the FUJITSU VP2000 series computers is high because they incorporate advanced technologies; for example, ECL 15k-gate very-large-scale integrated circuits and multilayer glass-ceramic-composite circuit boards. The VP2000 series was developed using a design automation system (DA) that enabled the ultrahigh-speed technology used in the series to be fully exploited. This paper mainly describes the features of this DA system.

1. Introduction

In recent years, there has been an increasing need for ultrahigh-speed processing of computer graphics and scientific and technical computations. The FUJITSU VP2000 series, Fujitsu's newest supercomputers, fulfill these needs.

The VP2000 computers achieve a high performance by making the most of advanced technology; for example, high-density packaging of ECL 15k-gate very-large-scale integrated circuits (VLSI) and multilayer glass-ceramic-composite circuit boards. The performance of a system using ultrahigh-speed processing technology depends largely on the signal propagation delay caused by the wiring between elements. To reduce this delay, the whole system must be carefully designed. The design automation (DA) system for the VP2000 series was developed to enable the ultrahigh-speed technology used in the series to be fully exploited.

This paper mainly introduces the DA system for the VP2000 series. First, the DA system is outlined. Then, a logic simulation processor (SP), an ECL 15k-gate VLSI layout system, and a router system for multilayer ceramic boards are described.

2. Outline of the DA system

A DA system[1] automates certain processes between the logic and package design stage and the manufacture and inspection stages. It can be safely said that it is now impossible to design a computer without the help of a DA system. As the scale of computers has become larger, and their complexity and performance has increased, the functions required of a DA system have become increasingly advanced. Some examples of these functions are as follows:

1) Manipulation of large amounts of data
2) High-speed automatic processing
3) Highly functional automatic processing with consideration for constraints such as the propagation delay time
4) Highly precise checking, including the checking of electrical and thermal conditions
5) Short turnaround time for engineering changes
6) Excellent human-machine interface (HMI), especially for the execution of interactive programs
7) Centralized management of various kinds of design data at any level between the large-scale integrated circuit (LSI) level and the system level.

To satisfy these requirements, Fujitsu has developed an integrated DA system that supports all processing from logic data input to production data output at the LSI, multilayer glass ceramic assembly (MLA), and system levels. As shown in Fig. 1, the DA system consists

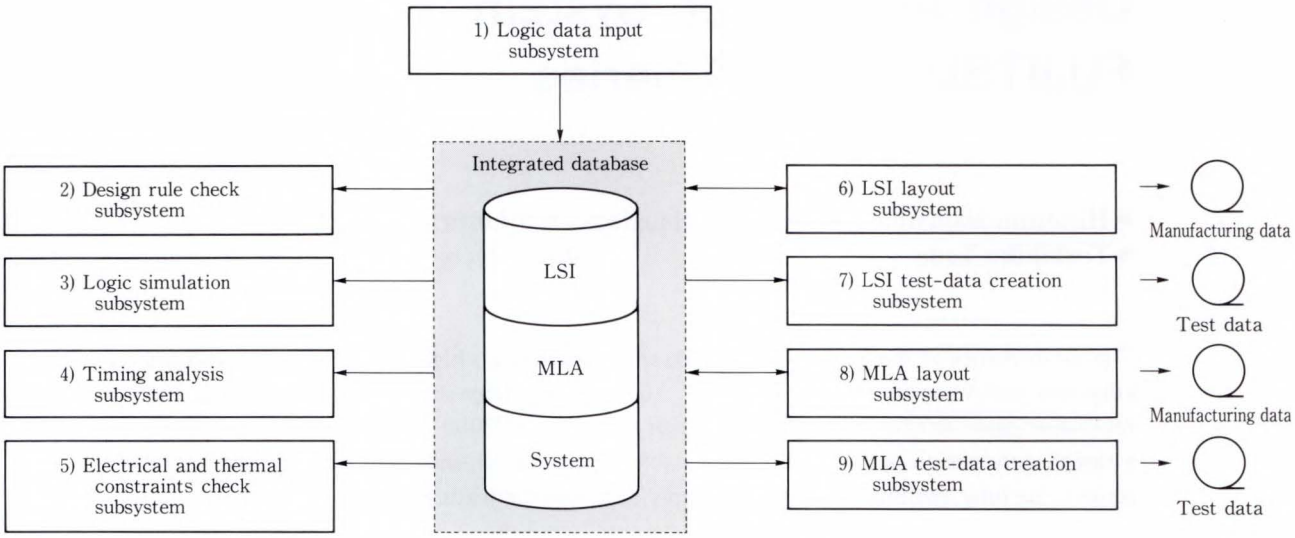


Fig. 1—DA system configuration.

of an integrated database and nine subsystems.

1) Logic data input subsystem

For logic data input, Fujitsu has developed a circuit editor for logic data that is input in the form of text and schema. This subsystem automatically generates scan circuits and clock circuits.

2) Design rule check subsystem

The design rule check subsystem checks whether a circuit conforms to the logic design rules; for example, constraints on the number of inputs and outputs connected to a gate.

3) Logic simulation subsystem

In addition to conventional simulation software, Fujitsu has developed a processor dedicated to logic simulation. This processor enables ultra high-speed simulation of several million gates at the system level.

4) Timing analysis subsystem

The timing analysis subsystem calculates the delay between flipflops, and checks whether the calculated delay is within the specified range. This subsystem can make delay checks anywhere between the LSI level and the system level.

5) Electrical and thermal constraint check subsystem

This subsystem makes electrical checks; for example, crosstalk and reflection noise checks, and thermal-condition checks for LSIs.

6) LSI layout subsystem

Fujitsu has constructed an LSI layout system, the main component of which is a floor plan editor that enables interactive placement using a graphic display. Using this floor plan editor, operations ranging from the rough placement of functional blocks to the detailed placement of cells can be performed. At the same time, delays and routing density can be evaluated. This subsystem also has a line length control function and a removal and rerouting function for automatic routing between cells. By using these functions, high routing densities can be achieved.

7) LSI test-data creation subsystem

The LSI test-data creation subsystem automatically generates functional test patterns and patterns for measuring path delays.

8) MLA layout subsystem

The MLA layout subsystem places LSIs on an MLA and performs routing between the LSIs. In particular, a technique for optimizing the design of new large-scale multilayer ceramic boards has been developed and applied to the router system.

9) MLA test-data creation subsystem

The MLA test-data creation subsystem checks the routing between LSIs on an MLA.

The next part of this paper explains the logic simulation processor of the logic simulation sub-

system, the LSI layout subsystem, and the multilayer ceramic board router system of the MLA layout subsystem.

3. Logic simulation processor

3.1 Logic simulation processor

For large-scale computers containing large numbers of VLSIs, engineering changes made after manufacturing has started significantly increase the development period and cost. Hence, it is important to determine whether logic devices will operate correctly by using full logic simulation before the manufacturing stage begins. However, full simulation of a large-scale circuit on a general-purpose computer is very difficult because it requires an enormous amount of computer time. The logic simulation processor[2)-4)] (referred to as the SP) is dedicated hardware that meets the need for high-speed simulation of large-scale logic devices.

3.2 Outline of the SP hardware

The dedicated hardware of the SP is based on the widely used event driven method. The event driven method performs computation only on the primitives in which a circuit operation has occurred. If the ratio of the number of operating circuits to the total number of primitives is low, the amount of computation performed by the event driven method is considerably less than the amount performed by methods that evaluate all the primitives in each clock cycle.

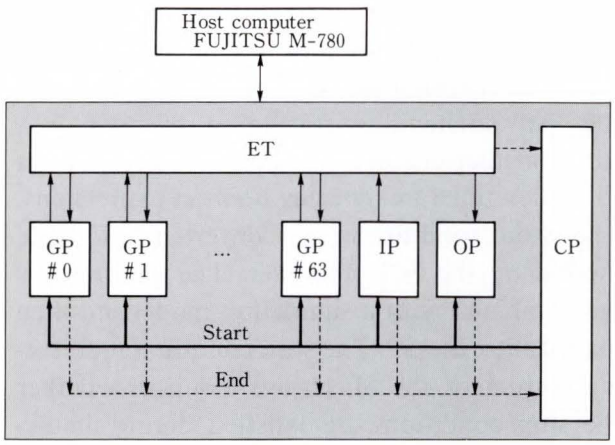


Fig. 2—SP configuration.

Figure 2 shows the SP configuration. The gate processor (GP) evaluates logic primitives having four inputs, one output, and memory primitives. Using pipeline processing of the algorithm for logic simulation, a single GP has performed processing 30 times faster than an equivalent software simulator run on the FUJITSU M-780. Also, under the same conditions, 64 GPs operating in parallel have performed processing up to 1 500 times faster.

The input processor (IP) holds the input patterns for simulation, and passes them to the GP together with the circuit model during simulation. The output processor (OP) stores the dynamic output values of specified primitives. The host computer reads the values contained in the OP after simulation. The OP also monitors the output values of particular primitives. When the output value reaches the specified level, the OP stops simulation. These functions of the dedicated input and output processors reduce the amount of communication between the host computer and SP.

The control processor (CP) controls all other processors. The event transmission network (ET) performs high-speed event communication between processors. The host computer divides a circuit model into sections and loads the sections into the individual GPs. The host computer also loads the input pattern into the IP and controls simulation by issuing commands to the SP.

3.3 Outline of the SP software

The most important purpose of the SP software system is to improve the overall system performance by making the best use of the high-speed processing capability of the SP hardware. For this purpose, Fujitsu has endeavored to speed up the preprocessing and postprocessing performed by the host computer and to minimize the overhead for communication between the host computer and SP during simulation.

Figure 3 shows the software system configuration. The entire software system resides in the host computer. The section that performs preprocessing for the simulation is called the model generation section. The model generation

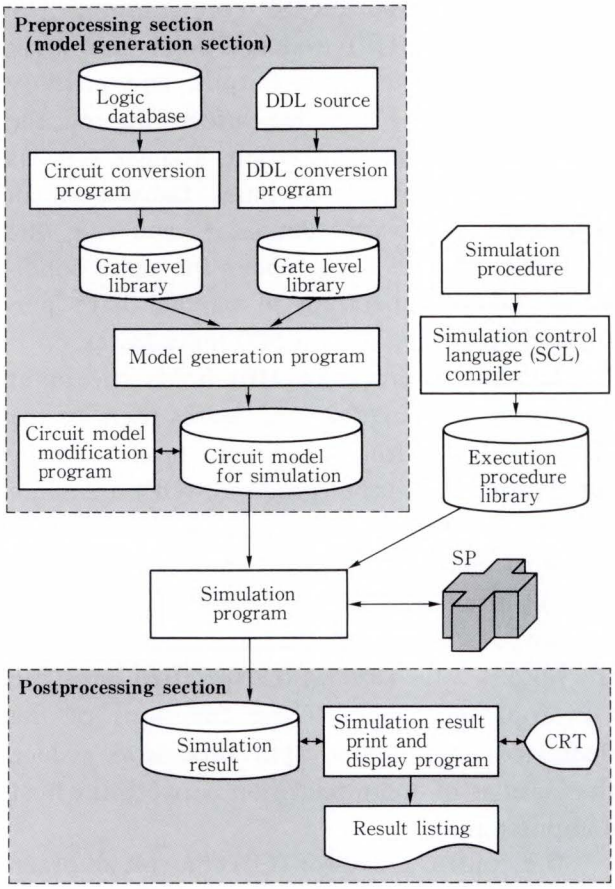


Fig. 3—SP software system configuration.

section consists of a circuit conversion program, a digital system design language (DDL) conversion program, a model generation program, and a circuit model modification program. This section generates circuit models for simulation. When the SP simulates a large-scale circuit having several million gates (the target circuit for the SP), the model generation section consumes most of the processing time. Hence, there is a great need to speed up the operation of the model generation section.

Each gate level library contains, in the form of a library, the circuit models at the gate level for individual LSIs or circuit units created by the circuit conversion program or DDL conversion program. If part of a circuit is changed, only that part needs to be re-created; therefore, the processing time is reduced. The model generation program links the gate level libraries and generates a model that the SP can simulate. The circuit model modification program partly changes a simulation circuit model. This program can modify a circuit in a maximum of one-tenth the processing time required by a model generation program for the same modification. In actual operation, minor modifications are made using the circuit model modification program. If the modification is not a small one, the logic device database is modified, and the circuit model is recreated by reexecuting the circuit conversion and model generation programs.

If the design is undetermined or detailed verification is unnecessary for part of a circuit, that part is described in the hardware function description langauge DDL. The DDL conversion program synthesizes a circuit model at the gate level from the function descriptions in the DDL. Then, this program outputs a gate level library in the same form in which the circuit conversion program is output. Finally, this part of the circuit is connected to the circuit that has been designed in detail, and the resulting circuit is simulated on the SP. This method involves a lower communication overhead between the host computer and the SP than methods in which the part described in the function descriptions and the part designed in detail are separately simulated on the host computer and the SP. Consequently, processing can be performed quickly even if the function description in the DDL and the detailed design are provided together.

The simulation control language (SCL) compiler converts an execution procedure written in the SCL into an execution procedure file containing SP control instructions. The simulation program performs simulation according to the instructions in the execution procedure file. An outstanding feature of the SCL is a function that enables simulation stop conditions to be described freely using boolean expressions. The stop conditions are converted into gate models by the SCL compiler. The gate models are combined with a simulation model and then loaded into the SP. The stop condition monitoring feature of the SP always monitors whether the stop conditions are satisfied during simulation. If the stop conditions are satisfied, the SP

stops. This function speeds up communication between the host computer and SP because the host computer does not have to interrupt the SP during simulation to check whether the stop conditions are satisfied.

3.4 Performance evaluation results

In a performance evaluation of a circuit having 410 000 logic primitives and 155 kbytes of memory, the SP (with 37 GPs) completed a simulation in 45 minutes that would have taken an estimated CPU time of 14 days on the M-780. In other words, the SP performed the simulation about 458 times faster than the estimated M-780 time. The important point here is that the SP can quickly execute a test program that would take much longer and be impossible to execute on a general-purpose computer.

The SP greatly reduced the development period of the VP2000 series and greatly improved the product reliability. The SP proved to be so valuable because it can quickly simulate an entire system.

4. LSI layout subsystem

4.1 LSI floor plan system

In recent years, as the scale of LSIs has become larger and their complexity has increased, LSI design has become more and more difficult and time consuming.

In conventional LSI design, logic and layout design are treated as independent fields. In logic design, priority is given to finding the combination of logic primitives that provides the correct logic function. In layout design, priority is given to determining how cells should be placed and connected together to improve the performance of the LSI. It takes a long time to evaluate the performance of an LSI because this is done after the placement and routing of the cells has been completed. Conventional LSI design is very inefficient because logic changes and even slight layout changes take up so much time.

To efficiently design an LSI in the environment described above, a design method and tool that enable the LSI to be evaluated after its logic has been designed is required. A method that has recently come into use is a hierarchical design method that uses a floor plan. A floor plan is a design tool used to determine general layout policies after the logic has been designed.

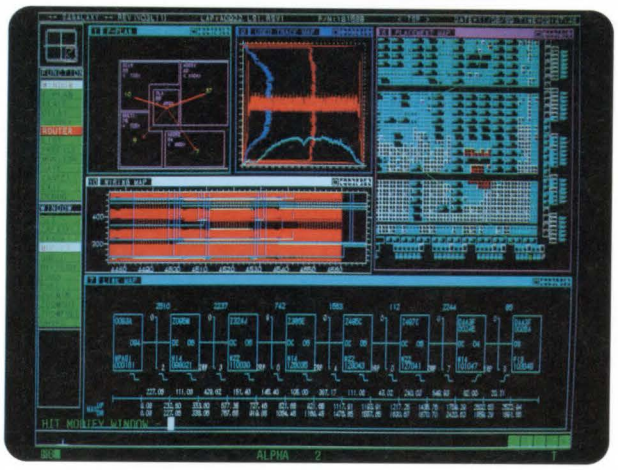


Fig. 4—Interactive floor plan system.

First, whilst taking the layout design into consideration, the function of the blocks (which contain dozens to thousands of cells) are defined in the logic design stage. Then, before the individual cells are placed and connected together, the placement of these blocks is determined on the basis of the block functions and on the connections between blocks. In this way, the layout of the entire LSI chip is optimized. Using the floor plan, the circuit size of each block and the routing density on the LSI chip can be easily estimated after the logic design stage. A logic change or placement modification can therefore be made in a short time. Furthermore, cells can be automatically placed in each block. Using this function, a good layout can be obtained in less time than when performing flat design of the entire chip.

The new floor plan is an interactive system that uses a graphic display and has an advanced HMI (see Fig. 4). The functions of this interactive floor plan are described below.

1) Specification of block shape and size

A block can be any shape defined by straight lines; for example, a rectangle or an L or T shape. A block must be large enough to

accommodate all the cells to be placed in it. When a block is created at the desired location on the LSI chip, the total number of cells in the block and the block size are displayed.

2) Evaluation of connection relationship between blocks

To enable the layout of blocks to be evaluated, the overall degree of congestion in the LSI chip and the connections between blocks are displayed after the blocks have been created.

3) Block modification

Modifications such as removing a cell from a block, moving a cell from one block to another, or replacing cells can be done easily on the graphic display. After the layout of the LSI chip has been evaluated, blocks can be replaced or the size and shape of a block can be changed.

4) Hierarchical block organization

By organizing blocks in a hierarcy, the floor plan can be used according to a hierarchical design method. For example, if an LSI chip is divided into four blocks, and the individual functions of each block are defined at a low hierarchical level, block placement can be evaluated at a high or low hierarchical level or at both hierarchical levels.

5) Automatic cell placement

Once the locations of individual blocks has been determined using the floor plan, cells can be placed automatically in each block. When placing cells, this system minimizes the total line lengths and conforms to the inter-cell line length limit. The results of the automatic cell placement can be checked on the graphic display.

The system displays manually placed cells and automatically placed cells in different colors. This function facilitates layout evaluation and modification after automatic cell placement.

The above concludes the outline of the floor plan and the functions of the system. Because the scale and complexity of LSIs is expected to continue to increase, design methods that use a floor plan are likely to become more important. Furthermore, LSI design work is expected to become more difficult. Fujitsu will therefore enhance the floor plan functions and improve the human machine interface (HMI) to further improve the efficiency of LSI design.

4.2 LSI router system

The ECL 15 k-gate LSIs used in the VP2000 series contain tens of thousands of sections that require routing, and more than a hundred different types of LSI have been developed for these machines. Moreover, the electrical restrictions on LSI wiring, for example, the line length limit and line spacing are very severe. Therefore, Fujitsu has developed an LSI router system that speeds up the development of LSIs. This router system has the following features:

1) Routing function with line length control

The length of clock signal routes and other routes that timing analysis has indicated to be critical can be kept within specified limits.

2) Routing with different line widths

Routing with different line widths can be performed in the same area of an LSI (see Fig. 5).

3) Remove and rerouting function

If previous routing prevents a required connection from being made, the section that is in the way can be removed and the required connection can be made (see Fig. 6). Then, the removed section can be rerouted.

4) Routing copy function

If the routing has to be altered because of a change in the logic or cell placement,

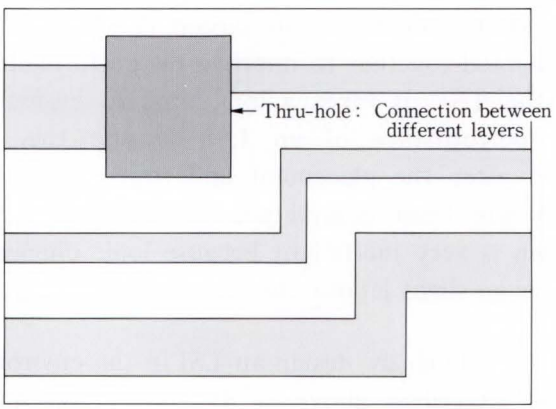


Fig. 5—Routing with different line widths.

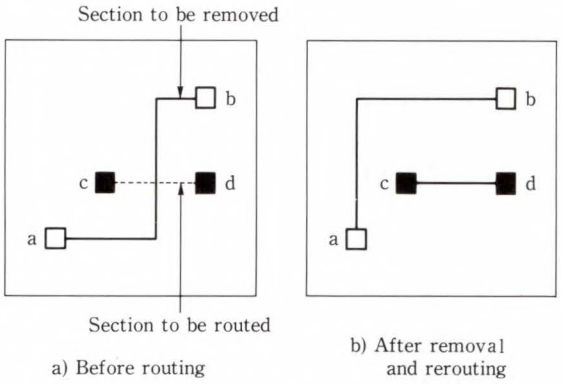


Fig. 6—Removal and rerouting.

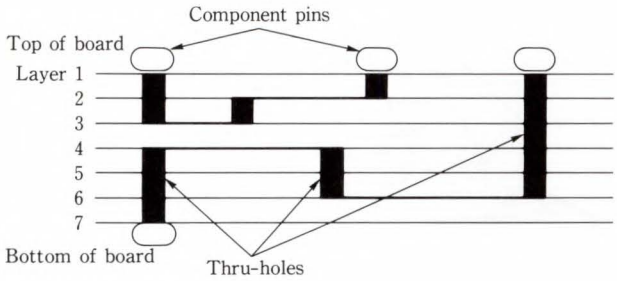


Fig. 7—Cross section of a ceramic board.

the LSI routing data that is not affected by the change in the logic or cell placement can be copied from the routing data before the routing is altered. This function enables the selective rerouting of a changed section, and thus speeds up the processing of design changes.

5) Interactive routing

If automatic routing cannot make the required connections, routing for the section containing the connection failure can be performed using the interactive router system on a graphic display. Various functions, such as routing with line length control, routing with different line widths, and removal and rerouting can be used with the interactive router system. The LSI router system therefore combines interactive and automatic processing.

5. Router system for multilayer ceramic boards

5.1 Features of ceramic boards

One of the features of ceramic boards is that thru-holes can be made to pass connections through any layer (see in Fig. 7). Therefore, in the routing of a ceramic board, checks

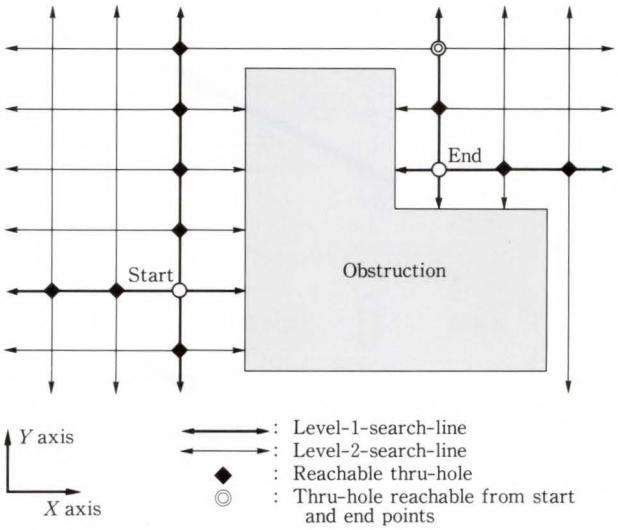


Fig. 8—Two-dimensional line-search.

must be made to determine which layer a thru-hole can pass through. Furthermore, because ultrahigh-speed LSIs are mounted on ceramic boards, the ratio of board thickness to board area is high. Hence, not only delay caused by the routing length, but also the delay caused by thru-holes must be considered.

5.2 Three-dimensional line search

The line search method is widely used to route between two points. Figure 8 shows the conventional two-dimensional line search method. First, level-1-search-lines are generated from the start point along the X and Y axes until an obstruction such as existing route is reached, then a search is made for reachable thru-holes. Next, level-2-search-lines are generated from the reachable thru-holes and a search is made for other reachable thru-holes. These operations are repeated from the end point. If a thru-hole that can be reached from both the start and end points is found, routing is performed between the start and end points. The routing path is determined by retracing the search lines from the common reachable thru-hole to the start and end points.

In this search method, searches are made only in the X-Y plane because the thru-holes pass through fixed layers. For the routing of a ceramic board, however, searches must

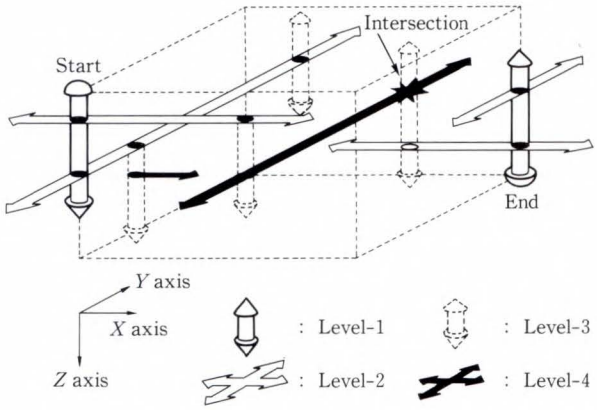


Fig. 9—Three-dimensional line-search.

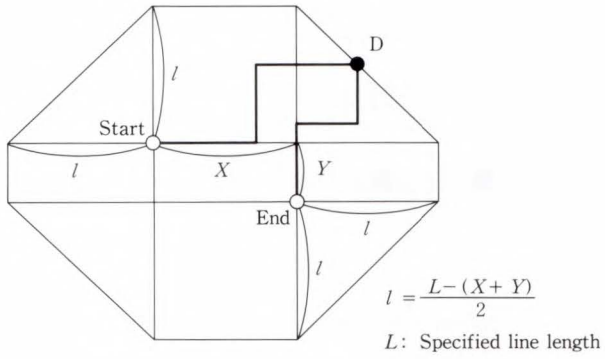


Fig. 10—Two-dimensional diamond.

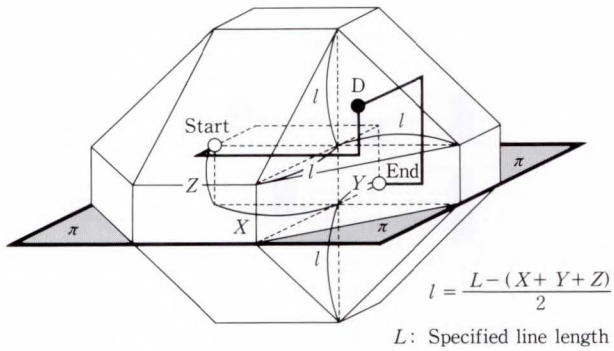


Fig. 11—Three-dimensional diamond.

also be made in the board's vertical direction (i.e. along the Z axis) to check which layer a thru-hole can pass through.

Figure 9 shows a three-dimensional line search method in which searches are also made along the Z axis. First, to find the layers that are reachable from the start point, a level-1-search-line is generated from the start point along the Z axis. Next, to find the points at which a thru-hole can be made, level-2-search-lines are generated along the X and Y axes in the layers that have been reached. To control the line length, this system records the length of each route form the start point to the points where a thru-hole can be made. (This length includes the thru-hole length.) Then, to find the layers that can be reached from these thru-hole points, level-3-search-lines are generated along the Z axis. These operations are repeated from the end point. If the lines generated from the start and end points intersect, routing can be performed between them.

5.3 Three-dimensional specified-length routing

If the delay for a route is specified (i.e. in the case of clock routes), the wiring between the start and end points must be controlled.

Two-dimensional specified-length routing has already been achieved by the router systems of the subsystem carriers (SSCs) for the printed wiring boards of the M-780[5),6)]. In these systems, an octagon (diamond) is drawn between the start and end points as shown in Fig. 10. Then, a detour point D is placed on one side of the diamond. If routing is performed using the shortest route between the start point and end point via D without detour, the length of the route is given by L.

If the diamond shown in Fig. 10 is extended to three dimensions, the solid shown in Fig. 11 is obtained. In this case, the detour point D is placed on the surface of the three-dimensional diamond. Again, if routing is performed using the shortest route between the start point and end point via D without detour, the routing between the start and end points is given by L, which includes the length in the board's vertical direction.

For the routing between LSI pins, only the partial solid of the diamond under the board surface π can be used because both the start and end points are on the π surface. If only the maximum length is specified, the maximum allowable area for a detour is the surface of the diamond.

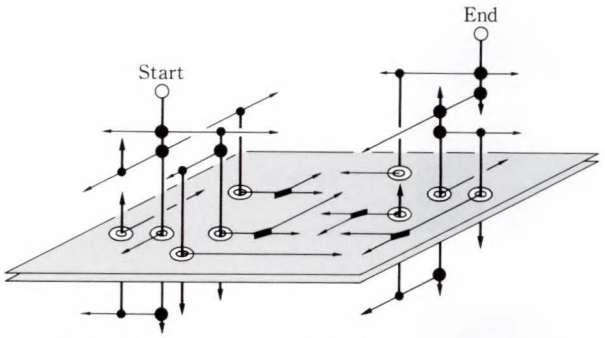


Fig. 12—Combining two- and three-dimensional routing.

5.4 Combing two- and three-dimensional routing

If complete three-dimensional routing is performed for all start and end points, the routing time is enormous because all routing layers will be searched. This method of routing is, therefore, impracticable. It was therefore decided to use the following two-stage routing method:

First, level-1-search-lines are generated from the start and end points along the Z axis to find the layers that are reachable from both points. Then, two-dimensional routing is performed for each pair of adjacent layers among the common reachable layers. If the line length is specified, priority is given to the layer pair that satisfies the specified line length (including the thru-hole length).

The three-dimensional routing shown in Fig. 12 is performed for those sections for which the above routing method has failed. First, three-dimentional searches are made from the start and end points. If the line length is specified, search lines are generated only within the extent of the diamond shown in Fig. 11. If search lines intersect and the line length condition is satisfied, that route is used. Next, a search is made for the layers that are reachable from both groups of thru-holes that can be reached from the start and end points. Lastly, two-dimensional routing is performed for each pair of adjacent layers among the common reachable layers. If the line length is specified, priority is given as in the previous stage.

This two-stage routing method, which combines two- and three-dimensional routing, has enabled the high-speed, high-density routing of tens of thousands of routing sections whilst satisfying the delay conditions.

6. Conclusion

This paper outlined the DA system for the development of the FUJITSU VP2000 series.

This DA system has fully met its original purpose of assisting in the development of Fujitsu's newest supercomputers, the VP2000 series. It is likely that the pursuit of performance improvements in the supercomputer field will continue endlessly, causing the design process to become increasingly complex.

Fujitsu will continue to improve its DA systems to support the design process.

References

1) Hamaguchi, T., Hamamura, H., Hori-i, Y.: Design Automation System for FACOM M-780. (in Japanese), *FUJITSU*, **37**, 2 (Special Issue: FACOM M-780 Computer Systems), pp. 135-139 (1986).
2) Hirose, F. et al.: Simulation Processor SP. Dig. Tech. Paper, IEEE Int. Conf. Comput. Aided Design (ICCAD-87), 1987, pp. 484-487.
3) Saitoh, M. et al.: LOGIC SIMULATION SYSTEM USING SIMULATION PROCESSOR (SP). Proc. 25th ACM/IEEE Design Autom. Conf., 1988, pp. 225-230.
4) Hirose, F. et al.: Simulation Processor: SP. (in Japanese), *FUJITSU*, **39**, 3 (Special Issue: FUJITSU TECHNOLOGY '88), pp. 198-205 (1988).
5) Tada, T., and Hanafusa, A.: ROUTER SYSTEM FOR PRINTED WIRING BOARDS OF VERY HIGH-SPEED, VERY LARGE-SCALE COMPUTERS. Proc. 23rd ACM/IEEE Design Autom. Conf., 1986, pp. 791-797.
6) Hanafusa, A., and Tada, T.: ROUTER SYSTEM FOR HIGH-DENSITY MULTI-LAYER PRINTED WIRING BOARDS. Tech. Paper, Printed Circuit World Conv. IV, 1987, pp.③2-16.

Hirofumi Hamamura
DA Development Dept.
FUJITSU LIMITED
Bachelor of Electrical Eng.
Keio University 1970
Specializing in Design Automation Systems

Minoru Saitoh
DA Development Dept.
FUJITSU LIMITED
Bachelor of Electronics Eng.
Tohoku University 1981
Specializing in Design Automation Systems

Akihiko Hanafusa
DA Development Dept.
FUJITSU LIMITED
Bachelor of Precision Eng.
The University of Tokyo 1981
Master of Precision Eng.
The University of Tokyo 1983
Specializing in Design Automation Systems

Toshihiko Tada
DA Development Dept.
FUJITSU LIMITED
Bachelor of Electrical Eng.
Kyoto University 1974
Master of Electrical Eng.
Kyoto University 1976
Specializing in Design Automation Systems

UDC: 681.32.06

Basic Software for FUJITSU VP2000 Series

• Koh-Ichiro Hotta • Takashi Kunai • Yoshio Honma

(Manuscript received January 28, 1991)

Since the first shipment of the FUJITSU VP-series, Fujitsu's policy regarding system software for supercomputers has been to supply systems that are as easy to use as a general purpose computer. Since then, the size of application programs has increased more and more, and progress made in computer networks has significantly changed the supercomputer environment. When the FUJITSU VP2000 series was released, new system software for the current environment was also developed. In this paper, the new system software is outlined.

1. Introduction

Fujitsu first shipped its FUJITSU VP-series supercomputer system in December 1983. The VP-series are easy to use high-performance supercomputers having architecture and system software compatible with the FUJITSU M-series of general purpose computers.

Fujitsu has developed the FUJITSU VP2000 series[1] as the latest processor of the VP-series. To support this new system, Fujitsu has also developed system software products, such as an operating system and a language processor system. These software products can cope with the increasing size of application programs and the changing computer environment.

This paper introduces the system software products for the VP2000 series.

2. Environments of system software for VP-series

2.1 History of system software for VP-series

When Fujitsu first shipped the VP-series in December 1983, the VP system was used as a back-end processor of a loosely coupled multiprocessor system and was regarded as a high speed calculator. The VSP special purpose operating system was developed with this function in mind. At the same time, a compiler having an automatic vectorization facility, the FORTRAN77/VP, was also developed.

Since the first shipment, the operating system for the VP-series has been enhanced to improve the use of resources and to simplify and improve operation. In 1985, the VPCF, a type of operating system attached to OS IV/F4 MSP, was developed. Using the VPCF system, a standalone VP system can be constructed that is as easy to use as a general purpose computer[2].

The language processor system has also been enhanced in terms of automatic vectorization, optimization, and usability (see Fig. 1).

2.2 Strategy for development of VP2000 system software

The use of supercomputers is expanding, and the increasing size of application programs is creating a shortage of system resources. The progress made in computer networks has enabled a wide variety of computer environments and applications.

The VP2000 system inherits the assets of the previous VP-series. Because the architecture has upward compatibility with the previous architecture, all application programs developed for the previous VP-series can be used on the VP2000 series without modification. A system storage is supported for large-scale programs, and a multiprocessor configuration is supported for

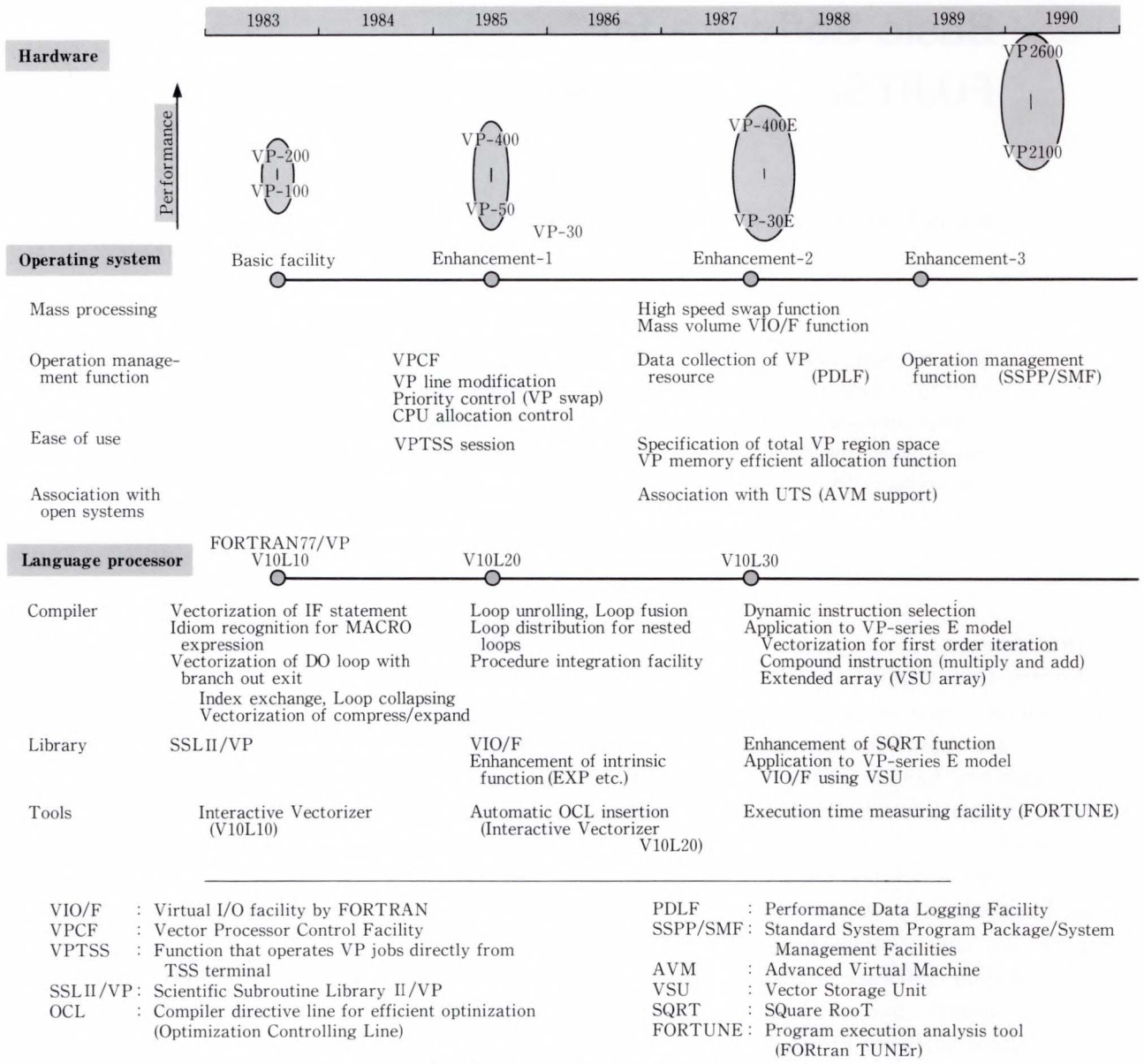


Fig. 1—History of VP software products.

the dual scalar processor (DSP) and quadruple scalar processor (QSP) systems. Vector programs can be run not only under MSP/EX system, but also, under UXP/M (based on UNIX[Note]) system. Regarding the language processor, the optimization and vectorization function has been enhanced and a parallelization function for the DSP and the QSP system has been added to achieve high performance.

The following chapter introduces the new facilities of the MSP/EX, UXP/M, and FORTRAN systems.

Note: The UNIX operating system was developed and is licensed by UNIX System Laboratories, Inc.

3. MSP/EX system

The MSP/EX system improves the throughput, expansibility, and flexibility of the VP2000 series by making full use of the hardware capability. The system also improves connectability to non-Fujitsu systems for easier construction of networks.

3.1 High throughput

3.1.1 System storage

A high-speed, large-capacity system storage was developed for the VP2000 series. The MSP/EX systems enables large-volume, high-speed input-output; high-speed swapping; and storage

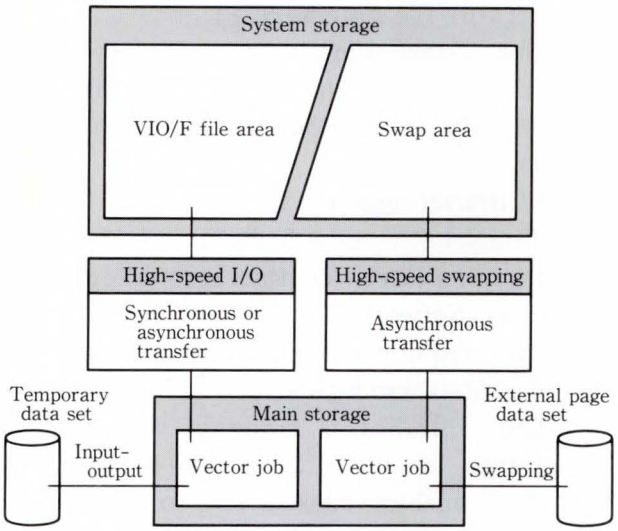


Fig. 2—System storage operating modes.

area of large-scale arrays. Figure 2 shows the operating modes of the system storage.

1) High-capacity

The high-speed performance of the VP series has been achieved by using a VIO/F input-output function that expands I/O files in main storage.

Because the amount of data that could be handled by the VIO/F input-output function was restricted by main storage limitations, the number of concurrent vector jobs could not be increased. However, with the release of the VP2000 series, it is now possible to expand the VIO/F file in the system storage. This has increased input-output speed and has solved the above problems. The system storage can hold up to 32 Gbytes of data. This system enables high-speed execution of application programs that process large amounts of I/O data. In addition, the VP2000 series can pass the system storage VIO/F file to subsequent job steps, and the VIO/F file is easier to use than a conventional VIO/F file that uses main storage.

2) High-speed swapping

The vector job and VPTSS sessions can be multiprocessed within the limitations of the main storage capacity. Also, to operate more vector jobs and VPTSS sessions, a swap function must be used. Swap processing, however, involves a large amount of data transfer. Consequently, if a direct access storage device (DASD) is used for external storage, swap processing is slow and the system overhead is increased.

The VP2000 series swapping function uses high-speed system storage. This enables very high-speed swapping, compared with conventional swapping in which swapping is done using a DASD. As a result, large-scale vector jobs that could only be executed at night (because they occupy main storage) can now be run at any time. Also, because the VPTSS session waiting the terminal input is swapped, the number of VPTSS sessions that can be used simultaneously can be increased beyond the restrictions imposed by the main storage capacity. Furthermore, because the system storage has an asynchronous transfer function that transfers data from and to main storage asynchronously with CPU processing, this function can be used to transfer swap data. The CPU can therefore be used exclusively for high-speed processing.

3) Storage area of large-scale arrays

As the amount of data processed by programs increases, large-scale arrays that are assigned to the VP job, but cannot be stored in main storage, are used more frequently. Conventionally, arrays are stored in external storage and only some of the arrays required for processing are transferred to main storage using an I/O instruction. However, the additional I/O statements complicate the program and performance is limited because of the increased number of I/O operations. The VP2000 series can store large-scale arrays in system storage, simplifies programs, and improves program execution time.

3.1.2 Array disk support

The VP2000 series can be connected to a disk array (F6490 magnetic disk subsystem) for high-speed, large-volume data transfer (36 Mbytes per second). The disk array enables high-speed parallel data transfer under hardware control, and radically reduces the I/O overhead as compared with data transfer using conventional external storage. In addition, the turnaround time of VP jobs is reduced and the throughput of the VP system is improved.

The MSP/EX system uses a disk array to support the following data sets:

1) FORTRAN I/O data set
2) VP swap page data set
3) SAVEHALT data set

3.2 Open system support

3.2.1 TCP/IP support

Since the MSP system uses the TCP/IP support package (TISP), the TCP/IP protocol (a de facto standard in the UNIX system) can be used directly. This enables construction of a network without the need to provide gateways, and enables the use of MSP system resources from a UNIX workstations.

The following functions can be used when the TISP is installed:

1) File transfer protocol (FTP)

This protocol is used to transfer files between the MSP system and UNIX workstations. File transfer can be specified from a workstation or the MSP system TSS terminal.

2) TELNET protocol (TELNET)

This protocol enables communication between UNIX workstations and the TSS and AIM of the MSP system. Remote login to another UNIX system is also possible from the MSP system TSS terminal.

3) Simple mail transfer protocol (SMTP)

This protocol is used to transfer mail between UNIX workstations and the MSP system. Mail transfer can be requested from workstations or the MSP system.

4) Network file system (NFS)

This system supports the NFS server function that can reference and transfer the MSP system file from UNIX workstations (NFS client).

3.2.2 UltraNet support

UltraNet[Note)] is a high-speed LAN (maximum transfer rate: 1 Gbit per second) developed by Ultra Network Technology Inc. By using UltraNet, requirements such as realtime transfer of image data and high-speed file transfer can be satisfied. UltraNet is currently supported by more than ten major companies in the world. UltraNet enables construction of distributed processing systems under a multivendor environment.

3.3 Expansibility and Flexibility

3.3.1 Multiprocessor support

The VP2000 series enables a multiprocessor configuration in addition to the conventional uni-processor configuration (UP). There are three types of multiprocessor configurations: the symmetric dual scalar processor (symmetric DSP), the asymmetric dual scalar processor (asymmetric DSP), and the quadruple scalar processor (QSP). These configurations improve cost effectiveness and enable high-speed processing by multitasking. Figure 3 shows these three

Note: A registered trademark of Ultra Network Technologies, Inc.

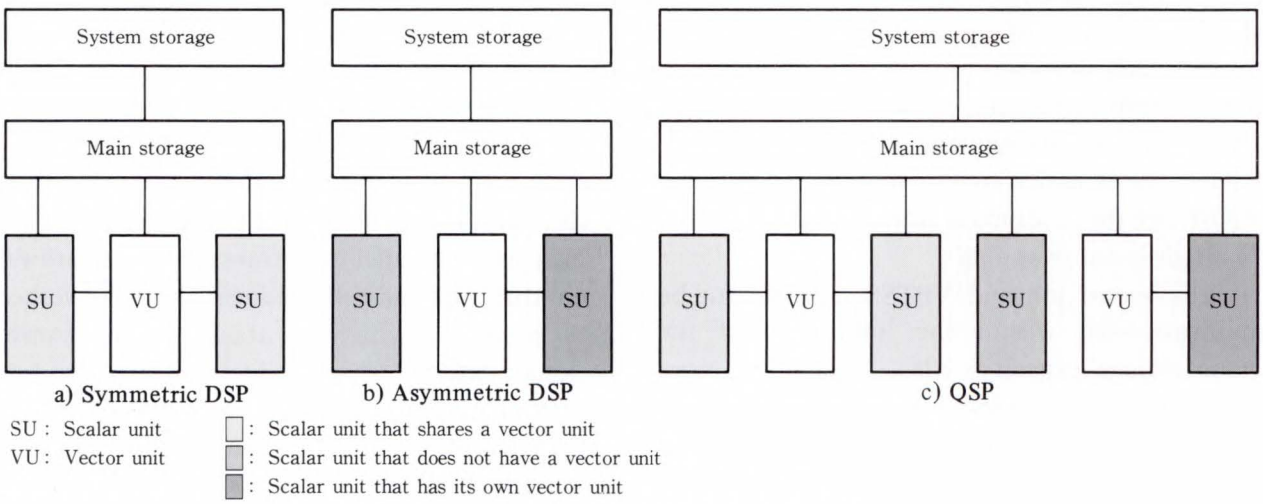


Fig. 3—Multiprocessor system configurations.

system configurations.

1) Symmetric DSP

In the symmetric DSP configuration, one vector unit is shared by two scalar units. The jobs processed in a general computer center do not always fully use the capability of the vector unit. Therefore, a way to fully use the hardware by improving the use-efficiency of the vector unit has been developed. In symmetric DSP, the two scalar units alternately use the vector unit. This increases the use-efficiency of the vector unit and also improves throughput and cost effectiveness.

2) Asymmetric DSP

In the asymmetric DSP configuration, one of the scalar units is separated from the vector unit. Asymmetric DSP provides the distributed functions of a conventional loosely-coupled, back-end system in a single machine. The independent scalar unit functions as a front-end processor and can be used for program development, compilation, and linkage editing. The scalar unit that has access to the vector unit can be used exclusively for vector jobs. Therefore, the MSP/EX system processes vector and scalar jobs separately. For vector jobs, the MSP/EX system uses the scalar unit that has access to the vector unit. For scalar jobs, the MSP/EX system uses the independent scalar unit.

Asymmetric DSP has the following advantages over a conventional loosely-coupled, back-end system:

i) System construction using a single machine
ii) Operation by a single system
iii) Cost reduction
iv) Exclusive control of shared DASDs, and the reduction of the overhead caused by exclusive control

Asymmetric DSP makes full use of the hardware and is suitable for a computer center that relies heavily on the vector unit. Both symmetric and asymmetric DSP can be dynamically selected during system operation.

3) QSP

The QSP configuration is a multiprocessor system in which two sets of DSP processors share the main storage. A QSP system achieves an exceedingly large throughput by allowing a total of four scalar units to execute jobs in parallel. Furthermore, by allowing the QSP system to use the multitasking function, a vector job can simultaneously use more than one vector unit; this greatly reduces the turnaround time of large-scale vector jobs.

The two DSPs can be independently configured as a symmetric or asymmetric DSP. For example, by configuring one DSP as a symmetric DSP and the other as an asymmetric DSP, the following configurations are possible: a scalar unit that uses the vector unit exclusively, a scalar unit that shares the vector unit, and a scalar unit that does not have a vector unit. Using a QSP, the optimum scalar unit can be assigned to the following jobs: vector jobs having a very high vectorization factor, vector jobs having a low vectorization factor, and scalar jobs.

The QSP configuration also improves the use-efficiency of scalar and vector units.

3.3.2 VP memory assignment function

Conventionally, when a VP job executed, the required size of VP memory is specified using JCL. However, because it is difficult to accurately estimate, more VP memory that is actually required is often specified and so the number of concurrent jobs that can be run is reduced.

To prevent this, a VP memory assignment function has been provided. When a program to be executed is compiled and linked as a VP job, the system calculates the VP memory size required to execute the program and stores the result in the load module. When the program is executed, the size requirement is read and the corresponding amount of VP memory is reserved.

This function conserves VP memory, and also improves system throughput because it allows the number of concurrent jobs in the VP system to be increased.

3.3.3 Parallel data contral facility (PDCF)

The PDCF enables parallel data access using system storage. Conventionally, when data is transferred between jobs using the temporary data set, the receiving job can read the data only after the data has been completely written. By

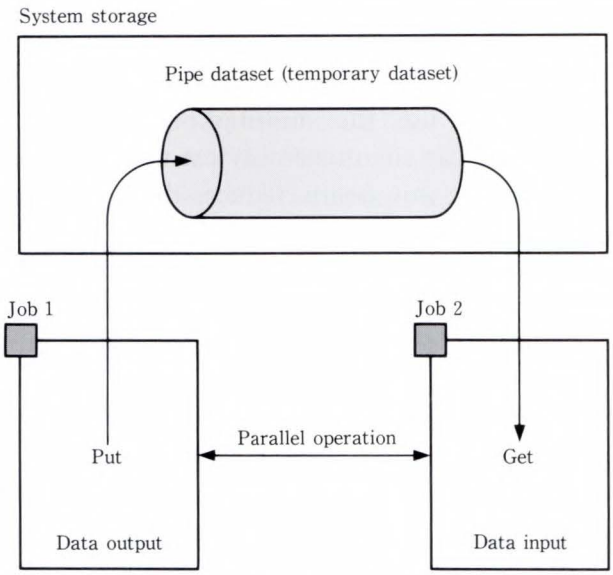


Fig. 4—Outline of PDCF function.

using the PDCF, the receiving job can read the data block immediately after the data block has been written to the temporary data set. This parrallel operation of output-input jobs improves the overall processing time. This function enables concurrent simulation and image processing on a realtime basis. Figure 4 outlines the PDCF function.

3.4 Advanced virtual machine/extended (AVM/EX)

The VP2000 series computers have the AVM/EX function which enables simultaneous, low-overhead operation of more than one system. In the AVM/EX basic section, hardware directly controls I/O instructions and reduces the overhead, thereby enabling high-speed operation. Also, the RAS function has been enhanced and its reliability improved.

The major characteristic of AVM/EX is that up to four guest OSs can simultaneously use the vector processing function. In the previous VP-E series, only one guest OS can use the vector processing function. This enables concurrent operation of MSP/EX and UNIX systems, and the workstation users can select the optimum system according to the job. Furthermore, system storage can be used by the guest OS, and the DSP system configuration and a wide range of other configurations can be constructed.

Faster file transfer and job transfer has been achieved by linking UNIX and MSP systems, and by directly transferring files and jobs using the main storage. As a result, workstations connected to the UNIX system can easily and efficiently use the MSP system resources and a large amount of the application software developed for the MSP system.

4. UXP/M system

In August 1990, Fujitsu announced that UXP/M, based on UNIX System V release 4 (SVR4), would be released in April 1991. UNIX is one of the key factors in supercomputers today. This chapter describes the roles of UNIX and Fujitsu's UXP/M operating system in the supercomputing environment. UXP/M has three major characteristics: a standard UNIX, a mainframe system, and a supercomputer operating system.

4.1 Supercomputers and UNIX

Engineering workstations are widely used in Research and Development (R & D) areas. Almost all computer processing, such as program development, document processing, electronic mailing services, graphics handling, and scientific calculations, can be carried out on these workstations. At the same time, improvements in the cost performance of supercomputers have enabled their use in new fields. Supercomputers are now common tools for R & D engineers.

Highly technical R & D environments, such as technical universities and research laboratories, use many different types of computer systems including workstations and supercomputers. Users are able to choose the best-fit computers for their applications, taking advantages of the features provided by each computer architecture. Therefore, the UNIX operating system has been adopted by many computers, because of its portability and architecture independency.

The UNIX environment offers the following advantages:

1) All computers that support UNIX have the same operation features and perform identical processing for the same application program. Therefore, experience on one

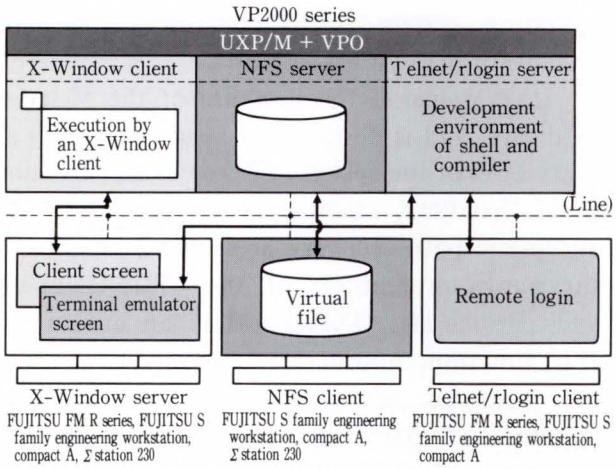


Fig. 5—UNIX network.

UNIX computer can be applied to all other UNIX computers.

2) A UNIX network can connect computers of different manufactures. Also, a UNIX network enables workstation users to login to a supercomputer and vice versa (see Fig. 5) For example, the result of a simulation generated by a supercomputer can be shown on a graphics processor connected to the UNIX network.
3) UNIX provides a standardized application interface, making application programs portable between computers that support UNIX.

4.2 Role of UXP/M[Note 1] as the standard UNIX

UXP/M is a standard version of the UNIX based on SVR4. To achieve network connectability and a standard application program interface, it is important to conform to the standard. UXP/M provides the same capabilities as the standard SVR4. Some examples are given below.

1) Unified UNIX

SVR4 is a unification of the standard System V UNIX and the Berkley Systems Distribution Version 4 (BSD4) UNIX. SVR4 is being promoted as the standard UNIX, and has the following features:

i) TCP/IP network functions common in workstations.
ii) Network applications such as the Network File System (NSF)[Note 2]
iii) File system functions such as the symbolic link, long path name, and quota system.
iv) SunOS[Note 3] functions including memory mapped files
v) BSD system commands such as job control SVR4 also includes the functions of the previous SVR3 version of UNIX, such as STREAM and dynamic link

2) Standardization

Conformity to the public standard has become a requirement for computer systems in the government agencies of western countries. SVR4 conforms to the standards shown below. UXP/M will also comform to these standards on its general availability.

i) SVIDS: UNIX standards
ii) Libraries and system call interfaces defined by POSIX P1003.1
iii) C language standard defined by ANSI X3J11
iv) XPG3: X/Open[Note 4] portability guide

3) Internationalization

An ultimate goal of internationalization is to provide a flexible environment in which a wide range of software can be used under any user-specified language conditions. SVR4 and UXP/M use a variable called 'locale' which enables the use of more than one language. It also enables message services in any language, including languages of multibyte characters.

4) Graphical user interface (GUI)

UXP/M supports X-Window[Note 5] V11R4, and also supports OpenLook[Note 6] as an advanced version of GUI.

4.3 Role of UXP/M as a mainframe system

UXP/M is based on the mainframe operating system, and when it is equipped with the vector processor support option (VPO) it also supports supercomputer vector processing. UXP/M offers the mainframe system capabilities listed below.

Note 1) UXP stands for UNIX Product.
Note 2) A registered trademark of Sun Microsystems.
Note 3) A registered trademark of Sun Microsystems.
Note 4) A registered trademark of X/Open Corporation.
Note 5) A registered trademark of the Massachusetts Institute of Technology.
Note 6) A trademark of AT & T.

These capabilities make UXP/M superior to versions of UNIX that have been implemented for ordinary workstations.

1) Supports of large-capacity, high-performance systems

The most notable feature of mainframe systerms are their high throughput. A mainframe system has much more real memory, better peripheral throughput, and higher CPU performance than minicomputers or workstations. Fujitsu's mainframe systems have superior performance in the following areas:

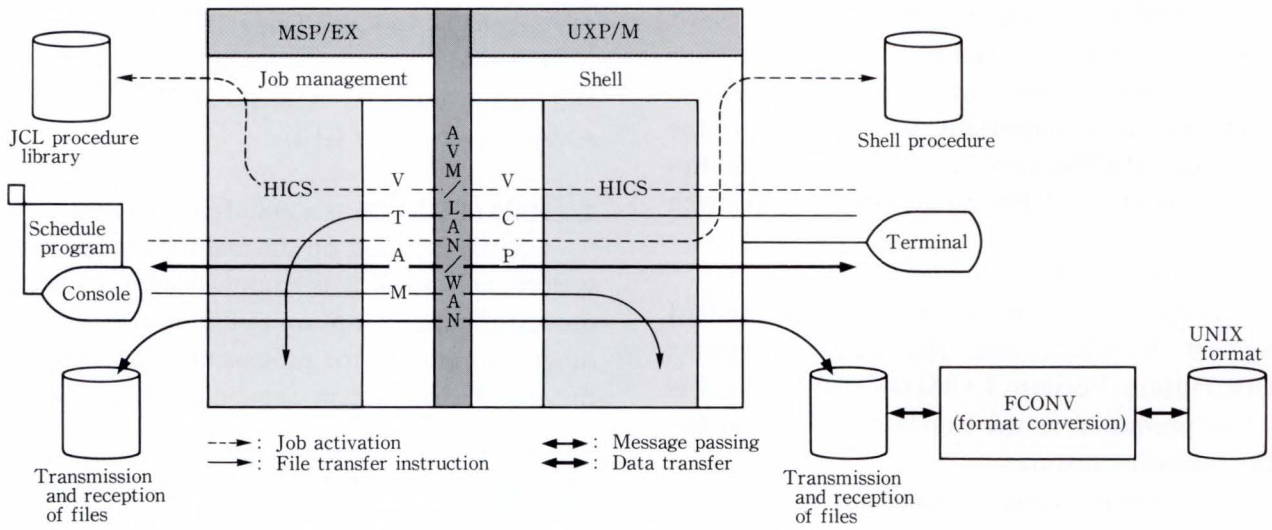
i) Multi processor configuration (including VP2000 series DSP/QSP)
ii) High I/O throughput (up to 128 high-throughput channels).
iii) Large capacity per magnetic disk spindle
iv) Large total capacity of magnetic disks connected to the system
v) Large network capacity (achieved by a communication control processor)
vi) High printer performance

2) Reliability and management functions

Because of their high throughput and large storage capacity, mainframe computers can accomodate many users and can store large amounts of data. One role of a mainframe operating system is to construct a secured system for many users.

UXP/M makes the best use of the M-series hardware, and is designed to assure reliability at every level of the system. CPU recovery, machine check handling, channel check handling, and automatic path recovery are some examples of the hardware features of the CPU, channels, and peripherals. UXP/M offers an automated IPL function, automated shutdown function, user resource control function, a file backup system, and many other functions to simplify the management of large systems.

UXP/M is often added to installations in which MSP/EX, Fujitsu's mainframe operating system, is already working. UXP/M is designed to be used with MSP/EX under a virtual machine (AVM/EX). UXP/M offers bi-directional file transfer, job activation, and message tranfer between guest operating systems under AVM/EX (see Fig. 6). The MSP/EX products can also be used under UXP/M. For example, the language processing system is highly portable, regardless of whether the operating system is MSP/EX or UXP/M. UXP/M supports the COBOL85, FORTRAN77, PROLOG, and LISP products developed under MSP/EX.



HICS : Hierarchical Information Control System
VTAM: Virtual Telecommunications Access Method
VCP : Virtual Communication Program
LAN : Local Area Network
WAN : Wide Area Network

Fig. 6—MSP/EX-UXP/M communication

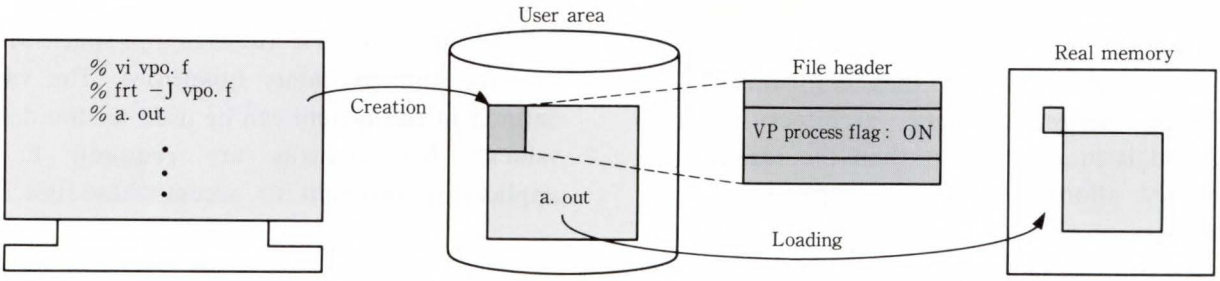


Fig. 7—VP load module.

4.4 Vector processor support option (VPO)

VPO is an optional software package that operates under UXP/M. VPO enables UXP/M to support the VP2000 series vector functions.

4.4.1 Features of VPO

VPO provides the extra function support necessary for the VP2000 series in the following ways:

1) Full vector performance

VPO's memory allocation method and process scheduling method are derived from the experience gained in developing vector support for MSP. The operating system is optimized so that vector programs can use the full power of vector functions. To attain full vector performance, a batch management function is provided (ordinary UNIX systems do not have this function). In addition, a large-capacity and high-speed file system is provided to cover shortcomings in the UNIX file system.

2) Complete support of UXP/M functions

The most significant feature of the VP2000 series is that it is based on the M-series general-purpose architecture. This provides the following benefits:

i) A standard UNIX, UXP/M reliability, and data security

ii) Scalar programs are guaranteed to run in exactly the same way as they run in the M-Series.

The VP2000 series can, therefore, be used as a special-purpose system for expert users or as a high-performance, general-purpose system for ordinary users.

4.4.2 Implementation of VPO

This section outlines the four major features of VPO.

1) Execution of VP programs

A vector version of the FORTRAN (FORTRAN77 EX/VP) compiler can be used to create vectorized application programs. The vector FORTRAN compiler indicates that the output is a VP module (see Fig. 7). When this module runs on a VP series, it is managed as a vector process which is permitted to use vector instructions. In order to optimize the machine performance, the scheduling algorithm for this process is set up for a longer slice time than that for non-VP (scalar) processes. The vector FORTRAN compiler can be used on the VP2000 series and on the M-series. This means that an M-series computer can be used as a front-end machine to develop software for a VP2000 series computer.

2) Allocation of memory resources

The VPLIMIT function controls the allocation of real memory to vector processes in the system. The VPLIMIT value restricts the maximum number and size of vector processes in a particular system. The VPLIMIT value is defined at system generation and should be the maximum value required for the site. The system administrator can change the VPLIMIT to any value within the value defined at system generation. For example, the system administrator can specify a large VPLIMIT value at system generation, and then decrease the value during the daytime when there are many terminal users and increase it at night when there are few terminal users.

Vector processes can run simultaneously on the system if the total memory used is less than the VPLIMIT value. The smaller the individual vector processes are, the greater the number of concurrent vector processes, even for the same

VPLIMIT value.

Memory is allocated using a 4-Mbyte page virtual address method that conforms to the VP2000 series memory architecture. This method is more efficient than the real address memory allocation method applied in other supercomputers.

A swap device can increase the number of concurrent vector processes from the VPLIMIT value up to the maximum capacity. Swap I/O devices are allowed to split their operation among more than one swap device. Swap speed can be increased by adding swap I/O resources of the system, such as devices and/or I/O paths. The use of a system storage unit (SSU) as a swap device will further increase swap speed.

3) System storage unit (SSU) support

Up to 32-Gbyte of SSU can be connected to a VP2000 series computer to function as a temporary file or a high-speed swap device.

The SSU can be used as part of the normal file system by defining a directory name (e.g. /tmp) and its capacity in the 'devicelist'. No changes are required in the application program. Specifying a file under the defined directory name will assure a transfer speed that is more than several hundred times higher than that achievable using an ordinary magnetic disk unit.

4) Very fast and large file system (VFL-FS)

To efficiently access small files, the ordinary UNIX uses a discrete file system and a buffer cache technique. This method, however, increases the overhead when large files are accessed. The reasons for this increase are:

i) The cache miss hit ratio is high for large files.
ii) The data is transferred to the system cache and then to the user data area.
iii) Single block accesses lower the disk transfer efficiency.

VFL-FS uses the virtual file system (VFS) functions introduced in SVR4 and provides the vfl file system in addition to s5, ufs, and nfs. The vfl file system with the mkfs option is differentiated from other file systems. The vfl file system supports a contiguous file system. Using VFL-FS, data is transfered directly from I/O devices to user application buffers. The vfl-alloc command or vfl-create system call allocates the vfl file area and specifies the primary and secondary quantities. The values defined in the system can be used as the default values. No changes are required in the application program to access these files.

4.5 Future of UXP/M and VPO

Fujitsu is developing UXP/M together with the activities of UNIX international, X/Open, IEEE POSIX, X consortium, and other international standardization bodies. UXP/M conforms to the standards of these bodies and also contributes to the development of such standards. SVR4 has improved the portability of BSD application programs, which are highly appreciated in R & D fields. SVR4 has plans to enhance security, networking, system administration and other system functions. SVR4 and its future releases will continue to be adopted by UXP/M as its base operating system. At the same time, UXP/M and VPO will also continue to be enhanced as a supercomputer operating system to maximize its usability and performance.

5. Language processing system

5.1 Concept of language processor development

The development target of the FORTRAN system for the VP-series is to improve the performance of application programs utilizing hardware functions so that the VP-series is as easy to

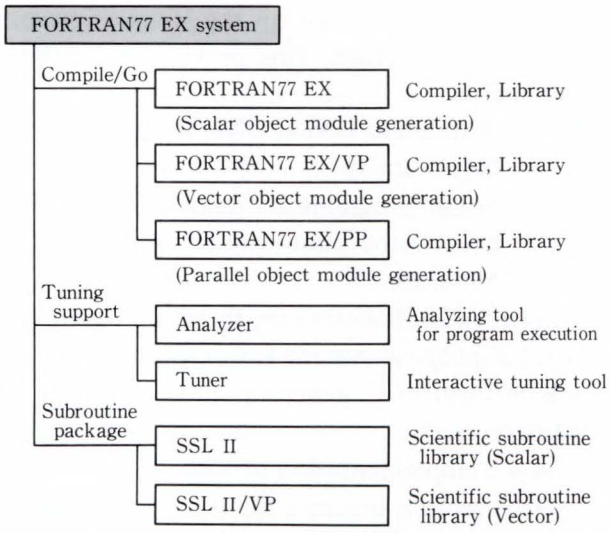


Fig. 8–FORTRAN77 EX system.

use as a general purpose computer. Before the first shipment of the FORTRAN77/VP compiler, Fujitsu had studied many scientific application programs to achieve this target and had implemented various kinds of optimization functions. In this sense, the FORTRAN77/VP compiler can be called an application program oriented system.

A new FORTRAN system has been developed for the new hardware facilities of the VP2000 series (see Fig. 8). FORTRAN77 EX/VP and FORTRAN77 EX/PP are new compiler systems for vector and parallel execution. FORTRAN77 EX/VP has new vectorization facilities which are based on the FORTRAN77/VP compiler. FORTRAN77 EX/PP generates object codes for parallel execution on the QSP or DSP system of the VP2000 series hardware. Also, the new tuning tools "Analyzer" and "Tuner" are developed for the FORTRAN77 EX/VP system.

5.2 Vectorization and optimization

To achieve high performance, it is important to first achieve a high vectorization ratio. The FORTRAN77/VP compiler carefully analyzes the statements in DO loops, and vectorizes not only simple patterns of statements but also complicated ones (e.g. IF statements, nested DO loops). In this way, the FORTRAN77/VP compiler achieves high vectorization ratio and good performance without modifications to source programs.

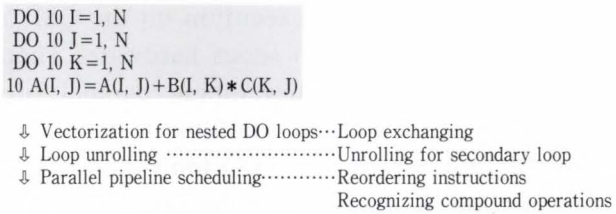
In the FORTRAN77 EX/VP compiler environment, the vectorization range has been extended and the performance enhanced. For example, the procedure integration facility has been enhanced so that a DO loop having CALL statements with a complicated argument interface can be vectroized after procedure integration. The FORTRAN77 EX/VP compiler not only vectorizes DO loops, but also optimizes vector object codes for VP2000 series hardware. The VP2000 series hardware has seven vector pipelines, six of which can work concurrently. Compound instructions (vector multiply & add operation) can be used on the multiply & add pipelines. To use VP2000 hardware efficiently, it is necessary to use as many pipelines as possible. Fujitsu has added new optimization facilities especially for the VP2000 hardware[3].

1) Parallel pipeline scheduling (PPS)

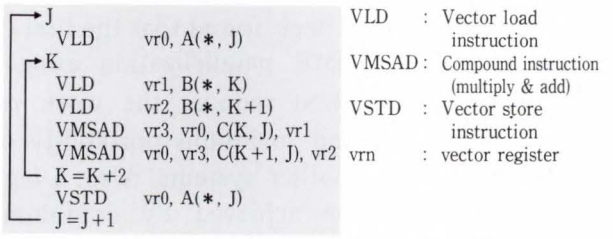
The PPS facility reorders the vector instructions and scalar instructions considering the hardware functions. A vector multiply operation and a vector add operation have been combined into a vector compound operation by using this facility.

2) Loop unrolling

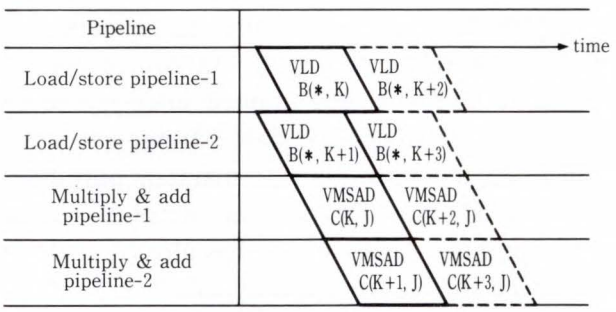
Loop unrolling is an optimization technique to reduce the iteration count of a loop and to duplicate the operations in the loop. The FORTRAN77/VP compiler unrolls secondary loops and increases the number of vector operations without changing the vector length. This loop unrolling optimization increases the number of vector instructions, therefore, promotes reordering of vector instructions by PPS and



a) Original program (matrix multiplication)



b) Object code (pseude instructions)



c) Execution (timing chart of vector pipelines)

Fig. 9—Example of matrix multiplication.

improves the efficiency of the vector pipeline use. The new unrolling facility completely unrolls the inner-most loops (the iteration counts of which are small) and vectorizes the outer loops which have long vector length and include unrolled parts. This function increases the vectorization ratio, improves the vector length, and improves the efficiency of pipeline use.
which include unrolled parts. This function increases the vectorization ratio, improves the vector length, and improves the efficiency of pipeline use.

Figure 9 shows an example of matrix multiplication after vectorization and optimization. In this example, the nesting of loops is restructured, the compound operations are fully used, and the pipelines are filled with the vector instructions.

5.3 Parallelization

5.3.1 Basic strategy

FORTRAN77 EX/PP is a language processor system for parallel execution on the QSP or DSP system of VP2000 series hardware. There are two methods of developing parallel programs. The first method is automatic parallelization for intra-procedural (DO loop) parallelism. The second is explicit parallelization using an optimization control line (OCL: compiler directive) for inter-procedural parallelism.

For users, it has been found that the best of these two is automatic parallelization using a compiler. This method reduces the work required in tuning, and maintains portability of source programs to other systems. Also, a high performance can be achieved by combining automatic vectorization and parallelization with other sophisticated techniques.

FORTRAN77 EX/PP also has a method to describe parallelism directly in source code in order to enable coarse grained inter-procedural parallel processing. The basic strategy to introduce such a description is to maintain portability with serial execution and to simplify parallel programming. OCL, which has already been introduced for vectorization by the FORTRAN77/VP compiler, is a suitable method of implementing this strategy. This is because an OCL is described as a comment line in a FORTRAN source program and is ignored by other FORTRAN compiler systems. The semantic model of parallel description is structured using the fork-join model. Because this model is well structured, parallel programming with OCL is very simple.

```
P     DO 10  J=1, M
PV        DO 10 I=1, N
PV            A(I, J) = B(I, J) + C(I, J)
PV 10 CONTINUE
```

a) Inner vectorization and outer parallelization

```
 V    DO 10  I=1, N
PV        DO 10 J=1, N
PV            A(I, J) = B(I, J) + C(I, J)
PV 10 CONTINUE
```

b) Optimal index selection

P : Parallelized statement/loop V : Vectorized statement/loop

Fig. 10—DO loop slicing.

```
!OCL PAR CALL
      CALL SUB1 (A)
      CALL SUB1 (B)
      CALL SUB2 (C)
!OCL END PAR CALL
```

Make new instances of SUBROUTINE, start parallel execution, and wait for all instances to end.

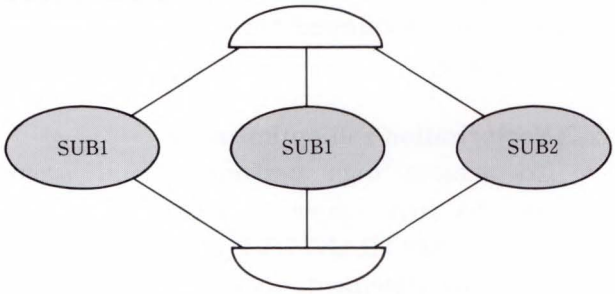


Fig. 11—Parallel CALL.

5.3.2 Automatic parallelization

Figure 10 shows examples of automatic DO loop slicing (parallelization). The standard operations in DO loop slicing are outer loop parallelization and inner loop vectorization {see Fig. 10 a)}. As an extended feature, the FORTRAN77 EX/PP compiler selects the optimal DO loop for slicing from the nested DO loops. This facility is a simple enhancement of the index exchanging function of vectorization for nested DO loops.

Table 1. LINPACK benchmark results

(MFLOPS)

	VP2600/10	VP2400/10	VP2200/10	VP2100/10
Order = 100	249	170	127	112
Order = 1000	4 009	1 668	842	445

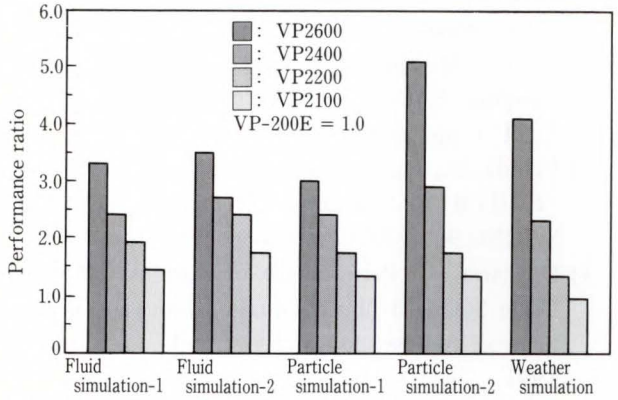


Fig. 12—Comparison of VP2000 series performance with VP-200E performance for various application programs.

5.3.3 Parallelism description

1) Fork-join

Figure 11 shows an example of an OCL that describes parallel execution of three subroutines. '!OCL PAR CALL' and '!OCL END PAR CALL' are the OCLs that execute concurrently the subroutines called between two OCLs. Because serial execution of these subroutines is a special case of parallel execution, OCL can be ignored and portability to other systems is maintained.

2) Synchronization

To synchronize parallel called subroutines, the FORTRAN77 EX/PP has three kinds of operation: mutual exclusion, barrier, and post/wait. Mutual exclusion is described by '!OCL MUTEX' and '!OCL END MUTEX'. Barrier and post/wait are described by CALL statements for intrinsic subroutines. In the case of serial execution, mutual exclusion can be ignored, because it keeps the meaning of a program. But barrier and post/wait change the meaning of a serial program.

5.4 Tuning

Program tuning on supercomputers such as the VP-series is more effective than on a general purpose computer. Programs for a FORTRAN77/VP system can be tuned using "VECTUNE". For a FORTRAN77 EX/VP system, 'Analyzer' and 'Tuner' are available. The basic method of program tuning is as follows:

1) Select a program section that consumes a lot of time.
2) Choose tuning methods to improve the selected part.

For the first procedure, 'Analyzer' calculates the cost of each statement, loop, and subprogram, and then outputs the results in a source program list which includes the optimization vectorization results. Programmers can determine the routines/loops required for tuning from these results.

For the second procedure, 'Tuner' shows users how to tune the loop.

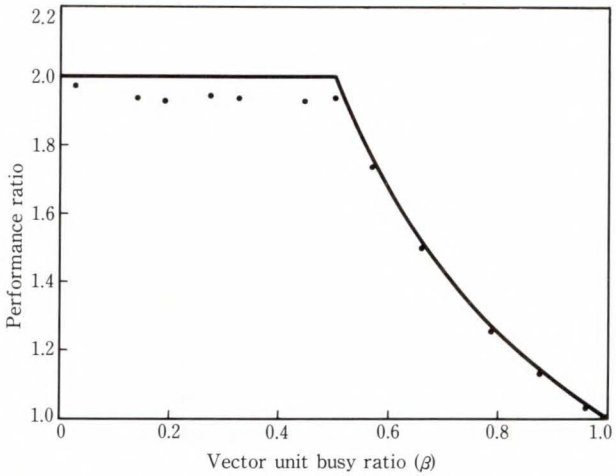


Fig. 13—CPU time/elapsed time on DSP.

5.5 Performance

On a VP2000 series uni-processor system, a high performance can be achieved using certain combinations of hardware and software. Table 1 shows the LINPACK[4] bench mark results for the VP2000 series. Figure 12 compares the performance of the VP2000 series with that of the VP-200E for various application programs.

The DSP system for the VP2000 series can be used as a multiprocessor system. Figure 13 shows the performance ratio of UP and DSP con-

figurations as measured by a model program[5]. The horizontal axis indicates the vector unit busy ratio β. The scalar unit is busy for a fixed time of $1-\beta$. If $\beta<0.5$, the vector unit can always be scheduled with minimum conflicts, and the performance ratio (total CPU time on DSP/elapsed time on DSP) is very close to 2.0. Hence, if $\beta>0.5$, the vector unit is always busy when the scalar units are in the execution or wait state, and the performance ratio decreases as β increases. Note that the vectorizing ratio is high (about 95 percent) when the vector unit busy ratio β is about 50 percent.

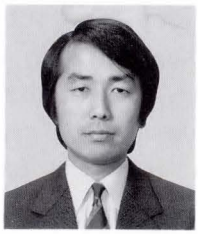
6. Conclusion

The system software of the VP2000 series enables high-performance execution of application programs having large amounts of data and many calculations. This software extends the range of high-performance applications and makes the supercomputer FUJITSU VP2000 series suitable for use in open systems environment.

There are now demands for even higher processing speeds. Fujitsu will continue to develop better and easier to use supercomputer systems.

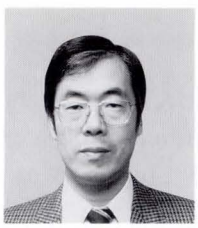
References

1) Uchida, N., Oinaga, Y., Tamura, H., and Shimizu, K.: System Overview of FUJITSU VP2000 Series. *FUJITSU Sci. Tech. J.*, **27**, 2 (Special Issue on Supercomputer VP2000 Series), pp. 149-157 (1991).
2) Tamura, H., Shinkai, Y., and Isobe, F.: The Supercomputer FACOM VP System. *FUJITSU Sci. Tech. J.*, **21**, 1, pp. 90-108 (1985).
3) Uchida, N., Hirai, M., Yoshida, M., and Hotta, K.: "FUJITSU VP2000 Series." Dig. papers, COMPCON SPRING 90, IEEE Comput. Soc. Press, pp. 4-11.
4) Dongarra, J.: Performance of Various Computers Using Standard Linear Equation Software. Electronic mail address: dongarra @ cs. ulk. edu., Feb. 16, 1991.
5) Miura, K., Nagakura, H., and Tamura, H.: "VP2000 SERIES DUAL SCALAR AND QUADRUPLE SCALAR MODELS SUPERCOMPUTER SYSTEMS —A NEW CONCEPT IN VECTOR PROCESSING—". Dig. papers, COMPCON SPRING 91, IEEE Comput. Soc. Press, pp. 294-302.



Koh-Ichiro Hotta
Development Dept.
Software Div.
FUJITSU LIMITED
Bachelor of Information Science
The University of Tokyo 1980
Specializing in Development of Language Processing Systems

Yoshio Honma
Development Dept.
Software Div.
FUJITSU LIMITED
Bachelor of Information Science
Tokyo Institute of Technology 1977
Master of Information Science
Tokyo Institute of Technology 1979
Specializing in Development of Operating Systems



Takashi Kunai
Development Dept.
Open System Development Group
FUJITSU LIMITED
Bachelor of Physics
Kyoto University 1972
Specializing in Development of Mainframe and Supercomputer UNIX Systems

UDC 548.5:681.32

Atomic-Scale Simulations for Semiconductors by Supercomputer

• Minoru Ikeda • Kumiko Furuya • Takahiro Yamasaki • Masuhiro Mikami

(Manuscript received December 21, 1990)

The electronic structure and crystal growth of semiconductor materials were studied using computer simulations. First, the energy bands and transition probability of Si-Ge superlattices were investigated. It was found that direct transition can be realized in $(Si)_5/(Ge)_5$ on a Ge substrate by zone folding and by breaking the inversion symmetry. Then, crystal growth by Molecular Beam Epitaxy was simulated for Lennard-Jones systems by applying molecular dynamics. The lattice mismatch produced various growth patterns, which agree well with experiments on metallic hetero-structures. It was found that Si adatoms on a reconstructed Si(100) surface diffuse anisotropically and make new stable dimers when they meet.

1. Introduction

The achievement of efficient light emission from Si at levels sufficient for optical devices would bring about a new generation of Si LSI technologies. However, the achievement of optical transition by Si with an indirect transition band structure is not expected.

Recently, Si-Ge superlattices, which consist of thin hetero-epitaxial layers, were shown to have a direct-transition band structure, and are therefore, able to emit light[1)-3)]. Various kinds of theoretical calculations have been made on this system for different layer structures and substrates[4)-6)]. These calculations indicate that $(Si)_6/(Ge)_4$ grown on a SiGe alloyed substrate shows direct transition[4),7)-9)]. In one experiment, a $(Si)_6/(Ge)_4$ direct transition superlattice was made, which emitted radiation of 0.85 eV in the photo-luminescence[3)].

Better controllability is required in the fabrication of these superlattices. Molecular Beam Epitaxy (MBE) and Metal Organic Chemical Vapor Deposition (MOCVD) techniques have enabled monolayer control of thickness[10)]. To further improve control, it is necessary to study the behavior of constituent and impurity atoms, and to study the generation and motion of defects during crystal growth at an atomic level. These requirements seem beyond the abilities of state-of-the-art experimental techniques.

The authors and other researchers have studied the crystal growth of the Lennard-Jones (L-J) system and the growth of Si by employing the Molecular Dynamics method(MD)[11)-17)]. MD is a suitable method because it shows how factors such as positions, velocities, and temperatures change with time. By using MD, the overlayer growth-patterns found in L-J lattice-mismatched systems and Si MBE growth mechanism on (100) and (111) surfaces were studied.

This paper discusses the possibility of radiation from Si-Ge superlattices, based on the theoretical calculations of energy band structures and optical transition probabilities. This paper continues with computer simulations of crystal growth for L-J systems and Si. Although the L-J system was designed for rare gases, it is quite useful because it exhibits the general features of crystal growth. Finally, this paper discusses surface reconstruction of Si and the behavior of Si adatoms as determined using the Stillinger-Weber (S-W) potential. All numeri-

cal calculations shown here were performed on FUJITSU VP-series supercomputers, and the graphic presentations of results were generated on general purpose FUJITSU M-series large-scale computers.

2. Calculation of electronic structure and optical transition in Si-Ge superlattices

2.1 Principles and models

For Si-Ge superlattices to be a practical material for optical radiation, the energy band structure must be of the direct transition type and the transition must be allowed quantum-mechanically.

Figure 1 shows the first Brillouin zone for bulk Si and Ge, where the Δ and L are points of the conduction-band minima for Si and Ge, and Γ is the maximum point of the valence band.

The first condition, making the Δ or L point coincide with the Γ point, is roughly satisfied by utilizing the characteristics of the artificial periodic structure, i.e. zone folding. In the superlattices grown on (100) substrates, five-times folding moves the Δ-point to $\Delta_f^{\perp}$ which is practically the same point as the Γ point.

Then, the energy levels at the other unfavorable points, i.e. $\Delta^{\parallel}$, L, and $X^{\parallel}$ are moved above $\Delta_f^{\perp}$ by the strain in the superlattice produced by the substrate. Therefore, a substrate of other than Si should be used, e.g. Ge or SiGe alloy. The symmetry of wave functions determines whether the transition is allowed. The direct transition superlattices can be found by varying the number of Si or Ge layers whilst

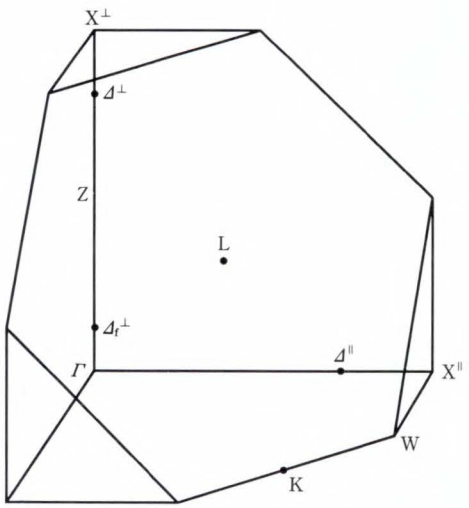
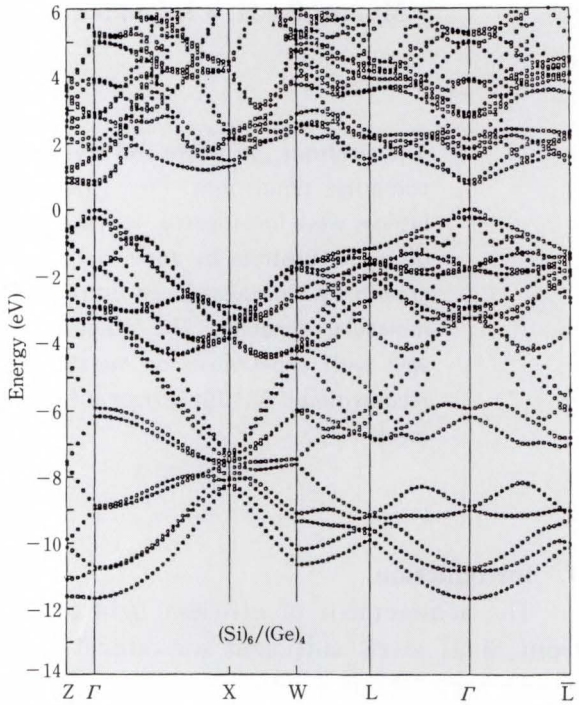
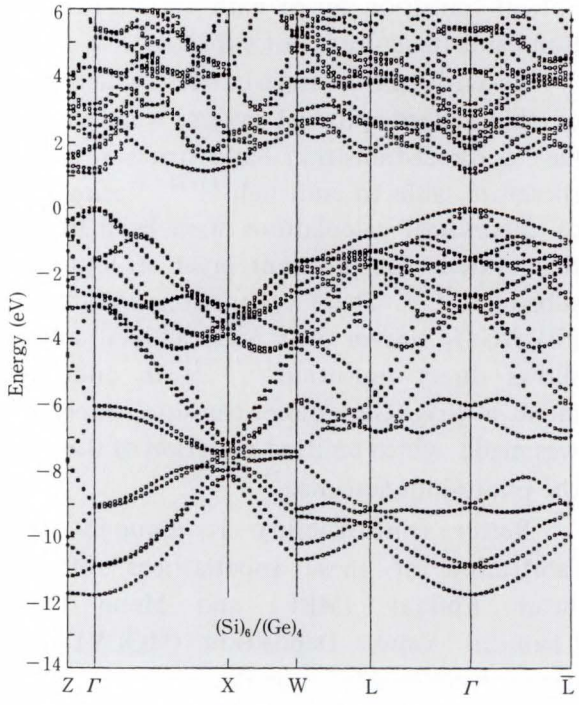


Fig. 1—First Brillouin zone for bulk Si and Ge. Longitudinal Δ-point is folded onto $\Delta_f^{\perp}$-point near Γ-point for superlattices grown on (100) substrate.



a) Ge substrate



b) SiGe alloyed substrate

Fig. 2—Energy band structures of $(Si)_6/(Ge)_4$ superlattice.

keeping the same period, or by alloying at the Si/Ge heterojunction interfaces.

Calculations were made for $(Si)_m/(Ge)_{10-m}$ superlattices with five folds and $m = 2$, 4, 5, 6, and 8. Also, $(Si)_2/(Ge)_2/(Si)_4/(Ge)_2$, and $(Si)_3/(Ge)_3/(Si)_2/(Ge)_2$ five-fold superlattices modulated from $(Si)_6/(Ge)_4$, aiming at closer mixing of the wave functions by the modulation were studied. For the superlattices with alloyed interface, we calculated for m = n = 4. Substrates used were Si, Ge, and $Si_m Ge_n$ alloy with (100) surfaces. Calculations were based on the LMTO method and are described elsewhere[9),18)].

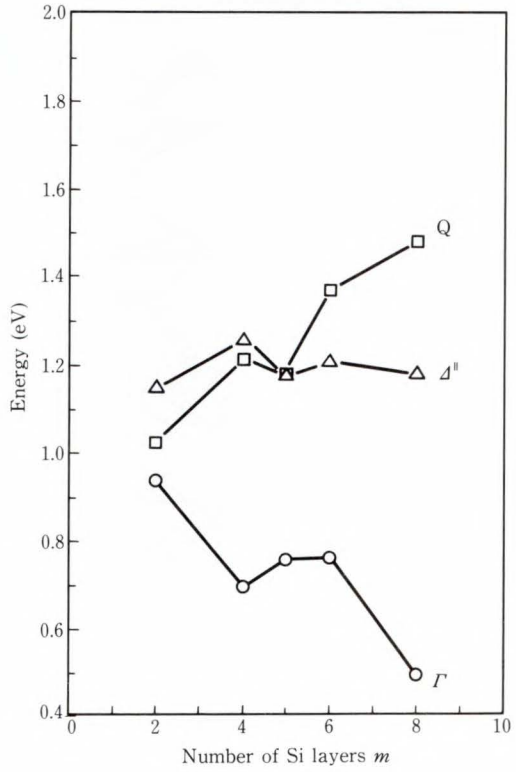
2.2 Results and discussion

1) $m + n = 10$

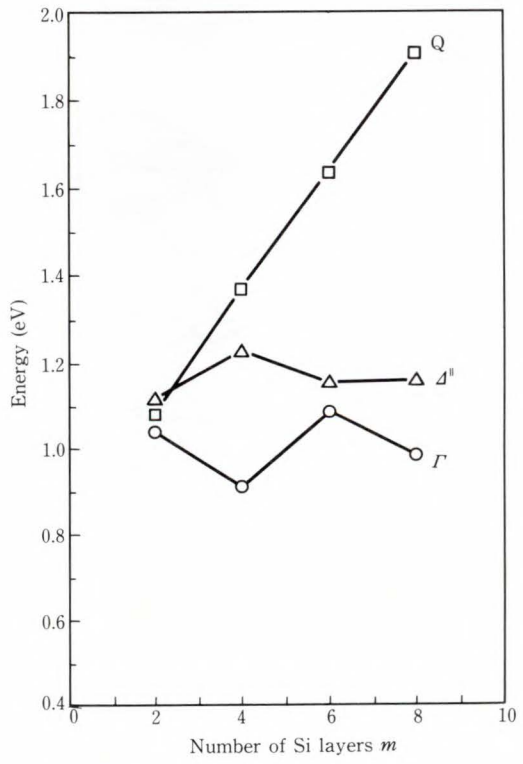
Calculations have shown that the folded Δ-point of superlattices having m = 2, 4, 5, 6, and 8 coincide with the Γ-point; therefore these lattices have direct transition band structures. Figure 2 shows the energy bands of $(Si)_6/(Ge)_4$ on Ge and SiGe alloyed substrates. The energy gap is 0.78 eV for the Ge substrate and 1.1 eV for the SiGe alloy substrate. Figure 3 plots the lowest energy of the conduction bands at the high symmetry points as functions of the number of Si layer, m. Q indicates the point on the W-L line where the energy is lowest. In the case of the Ge substrate, the energy gap of a superlattice changes from 0.9 eV to 0.5 eV as m increases. Of the superlattices for which calculations were made, only $(Si)_5/(Ge)_5$ has a finite transition probability. This is due to the nature of the wave functions. For an even value of m, the point group symmetry D_{2h}, inversion or mirror symmetry reduces the amplitude of the dipole matrix elements. For odd value of m, these symmetries are destroyed.

The modulated superlattice $(Si)_2/(Ge)_2/(Si)_4/(Ge)_2$ has a direct transition band structure for both Ge- and alloyed-substrates (see Fig. 4), however, these superlattices have no transition probability. The minimum point of the conduction band of $(Si)_3/(Ge)_3/(Si)_2/(Ge)_2$ is slightly away from the Γ-point, therefore, this superlattices is an indirect material.

The results of calculations for superlattices of $m + n = 10$ are summarized in Table 1. In this

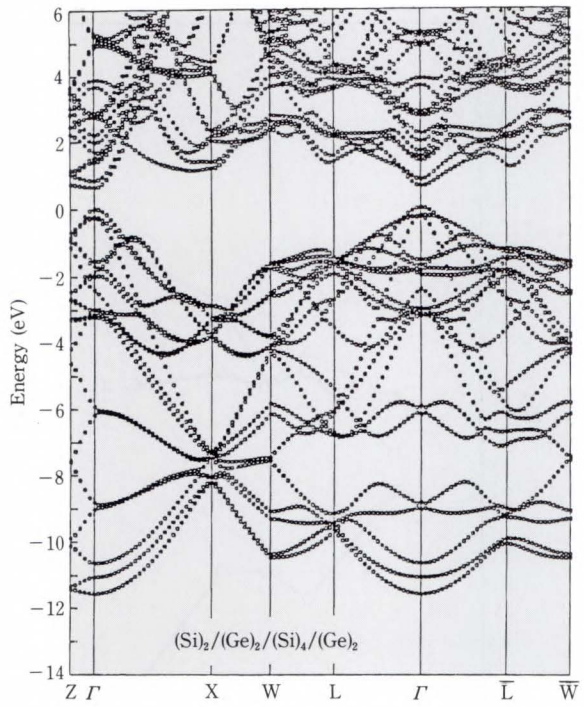


a) Ge substrate

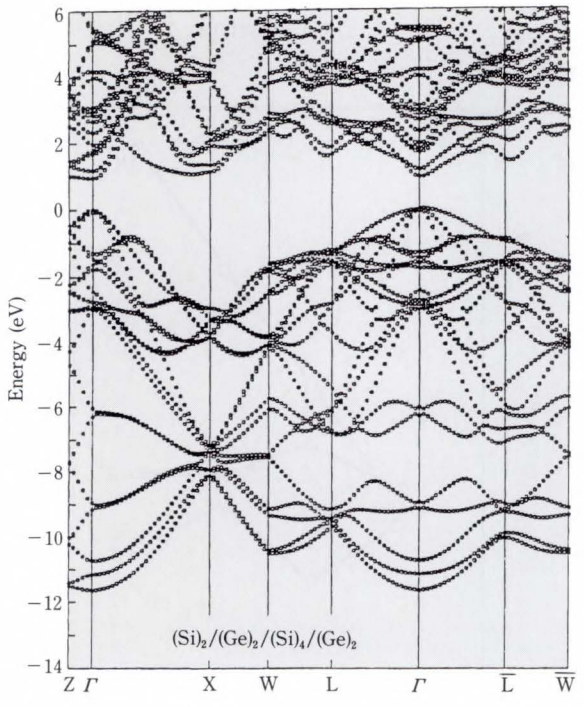


b) SiGe alloyed substrate

Fig. 3—Energy levels of conduction bands at Q-, Γ-, and transverse Δ-point vs. number of Si layers m.



a) Ge substrate



b) SiGe alloyed substrate

Fig. 4—Energy band structures of modulated superlattice, $(Si)_2/(Ge)_2/(Si)_4/(Ge)_2$.

table, D indicates a direct transition type, I an indirect transition type, T ≠ 0 an allowed transition, and T = 0 a forbidden transition.

Table 1. Summary types and probabilities of transition

Structure \ Substrate	Si substrate	Ge substrate	SiGe alloyed substrate
$(Si)_m/(Ge)_{10-m}$	(I)	D $T=0$	D $T=0$
$(Si)_5/(Ge)_5$	(I)	D $T\neq 0$	(D $T\neq 0$)
$(Si)_2/(Ge)_2/(Si)_4/(Ge)_2$	(I)	D $T=0$	D $T=0$
$(Si)_3/(Ge)_3/(Si)_2/(Ge)_2$	I	I	(I $T\neq 0$)

m: even number
D and I indicate direct and indirect band gap materials
$T = 0$ ($\neq 0$) indicates the forbidden(allowed) transition at the Γ-point. Items in parenthesis are conjectures.

2) Experimental viewpoint

The following experiments of Si-Ge superlattices have been reported: 0.85 eV photoluminescence from $(Si)_6/(Ge)_4$[3] and 0.76, 1.25, 1.70, and 2.20 eV allowed transition levels at the Γ-point found by electroreflectance for $(Si)_4/(Ge)_4$[2]. Theoretically, they are neither direct nor allowed transitions though the calculations give similar energy levels as the experiments: 0.78 eV for $(Si)_6/(Ge)_4$ on a Ge substrate and 1.1 eV on a Si substrate (both are forbidden transitions, and 1.0 (forbidden transition), 1.25, 1.65, and 2.20 eV (allowed transitions) for $(Si)_4/(Ge)_4$. There must, therefore, be additional mechanisms which destroy the inversion symmetry. To identify these mechanisms, an artificial two-layer Si-Ge compound was introduced to each Si/Ge interface of the $(Si)_4/(Ge)_4$ superlattices. This was done because this is believed to occur in actual superlattices[19]. As expected, an energy level close to the unidentified 0.76 eV appeared as an allowed transition level instead of the 1.0 eV forbidden transition. The disagreement between the theory and the experiment was, therefore, attributed to alloying at the hetero-interfaces.

3. Simulations of Molecular Beam Epitaxy

3.1 Model and methods

The model system of the MBE process consists of a substrate and a beam source. In the L-J system, the substrate consists of atoms arranged in the fcc (face-centered-cubic) structure. In the

Si system, the substrate consists of atoms arranged in the diamond structure with a (100) truncated surface. In both systems, the substrate has several atomic layers containing the square of n atoms. The atoms in the top four layers can move and interact, and those on the bottom four layers are bound at equilibrium positions. The temperatures of the substrate and beam atoms are kept at T_s and T_b, respectively. The source atoms are ejected at a rate of 1 atom/ps. The periodic boundary condition is applied in the horizontal direction. The parameters used in the simulation are shown in Table 2.

The calculation method is to solve the Newtonian motion equations for each atom at regular intervals. Determining the forces acting between atoms is a key step in MD simulation and this is discussed in the following sections.

3.2 Lennard-Jones system

The L-J potential is given as

$$V(r_{ij}) = 4\epsilon_{ij}\{(\sigma_{ij}/r_{ij})^{12} - (\sigma_{ij}/r_{ij})^{6}\}$$

where r_{ij}, ϵ_{ij} and σ_{ij} are the distance and parameters characterizing the binding energy and lattice constant between atoms i and j respectively. It is assumed that ϵ_{ij} is the same for all atoms, but that σ is σ_b or σ_s depending on whether interacting atoms are incident or substrate atoms. The σ_{sb} value, when an incident atom interacts with substrate atoms, is assumed arithmetic average $(\sigma_b + \sigma_s)/2$. The extent of lattice mismatch can be expressed by $\xi = \Delta\sigma/\sigma_s = (\sigma_b - \sigma_s)/\sigma_s$ and is from -0.4 to 0.4. The values of parameters σ and ϵ are those of Ar, i.e. $\sigma_s = 0.34$ nm, and $\epsilon = 1.67 \times 10^{14}$ erg.

3.2.1 Lattice-matched systems ($\xi = 0$)

Figure 5 is a side-view of the simultion at 400 ps. The blue atoms are the substrate atoms, and the red ones are deposited atoms. The white or yellow bars connect the neighboring atoms in each layer. Figure 6 shows the surface of Fig. 5.

The deposited atoms that reach the hollow sites in the surface remain there and become part of the square lattice of the substrate. Figures 5 and 6 show the results for $T_s = 10$ K. The results for $T_s = 30$ K and 50 K are almost the same.

3.2.2 Lattice-mismatched systems

1) $\xi = -0.4, -0.3$

In the first grown layer, incident atoms move into the hollow sites of the substrate and

Table 2 Conditions of MD simulations

Conditions / System	System size	Time step	Temperature
L-J	12 x 12 16 x 16 20 x 20	1×10^{-14} s	T_s = 10,30, 50 K T_b = 100 K
Si	6 x 6 12 x 12	2×10^{-15} s	T_s = 600 K T_b = 600 K

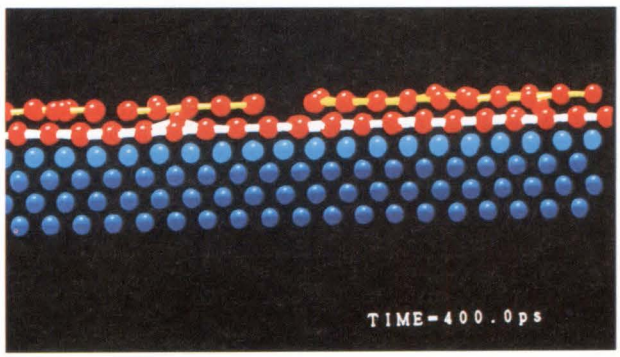


Fig. 5—Simulation for $\xi = 0$ at 400 ps seen from (110) direction. Two overlayers are partially formed. Blue atoms are substrate atoms, and red ones are deposited atoms. The latters are connected by white or yellow bars when they are in the first-neighbor distance within each layer.

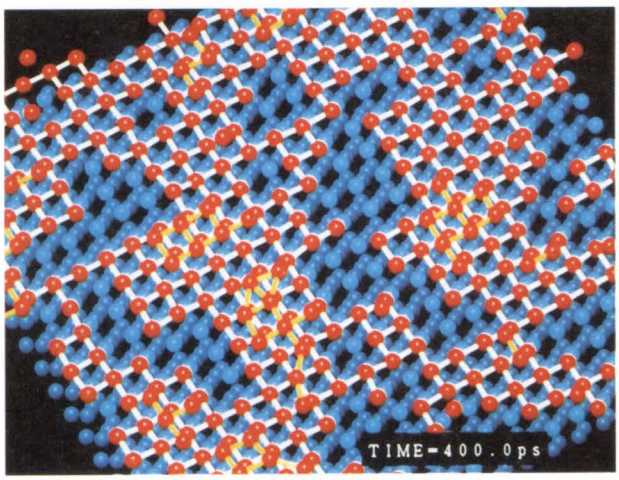


Fig. 6—Surface of the simulation shown in Fig. 5. Regular fcc lattice is formed by deposited atoms.

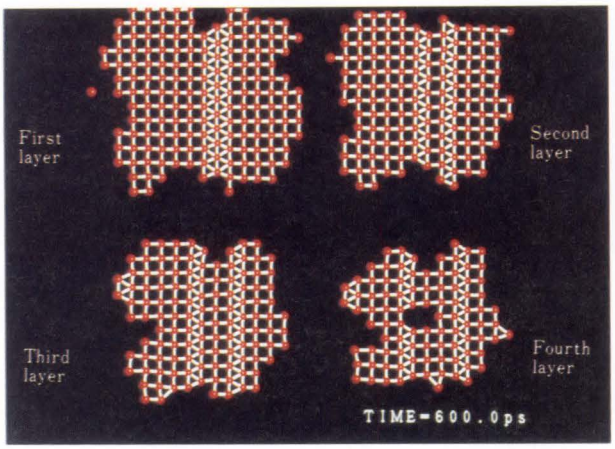


Fig. 7—Top views of simulation results for $\xi = 0.05$ at 600 ps.

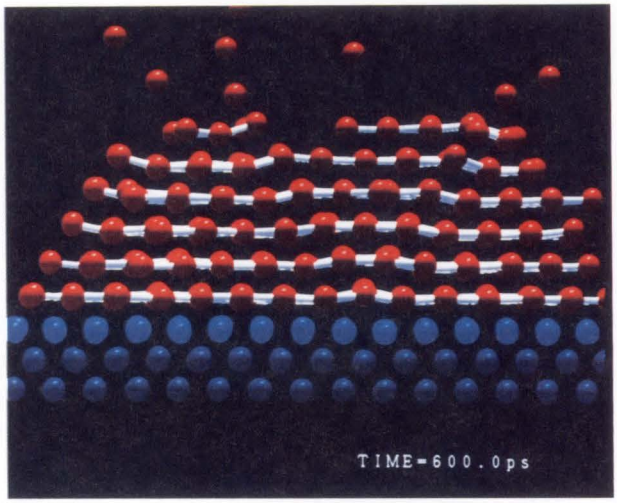


Fig. 8—Simulation for $\xi = 0.05$ at 600 ps. Side view from (110) direction. Central triangular part is raised against outsides at (111) planes.

become part of the lattice. In the second and higher layers, incident atoms take random positions as in an amorphous material. The distance between the grown layers is very short, and is about 40 percent of the distance between substrate layers. A similar result was obtained for $\xi = -0.3$.

2) $\xi = -0.2$

The deposited atoms take up positions that are slightly removed from the apexes of the square lattices. Therefore, the layers look like rows of chains with a (100) direction. We found it took place only when growth exceeds the fourth layer. Calculations for the energy of the system showed that four overlayers are sufficient to cause this structure when $\xi = -0.2$.

3) $\xi = 0.05$

Figure 7 shows the atomic arrangements of each layer after 600 ps of growth. Most of the structures are square lattices but discommensuration lines (DCLs: rows of triangular lattices) are also evident. A composite DCL on the first layer, and a pair of elemental DCLs over the second layer can be seen. The distance between a pair of DCLs increases toward the surface.

Figure 8 shows that the DCLs are in the (111) plane. The DCLs relax the stress in both the lateral and vertical directions. The DCLs moved as growth progressed, perhaps because of the accumulation of stress.

It was found that the number of DCLs depends on the simulation size; for example, two rectangular DCLs can reside stably for 20×20, whereas only one for 12×12.

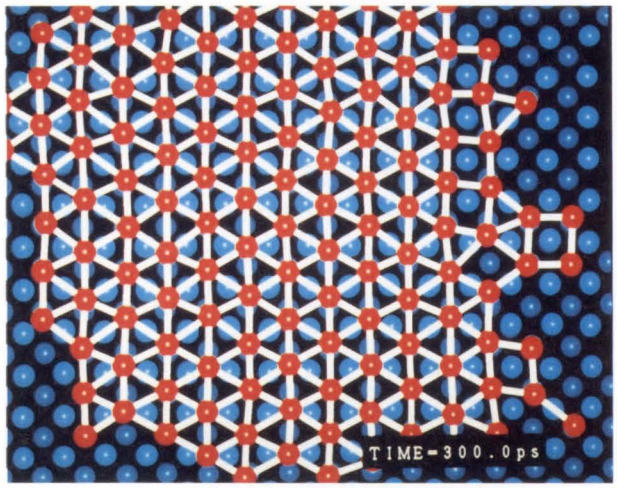


Fig. 9—First overlayer for $\xi = 0.1$ at 300 ps.

4) $\xi = 0.1$

In this lattice mismatch, a unit lattice on the overlayer becomes a triangle (see Fig. 9) that begins at the initial stage of the growth. One axis aligns exactly along the (110) direction, and the atomic spacing is about ten percent longer than that of the substrate.

5) $\xi = 0.4$

Figure 10 shows the first overlayer after 1 000 ps. This consists of two domains of square lattices, each of which fits into the substrate lattice by rotating 45° against the substrate. The

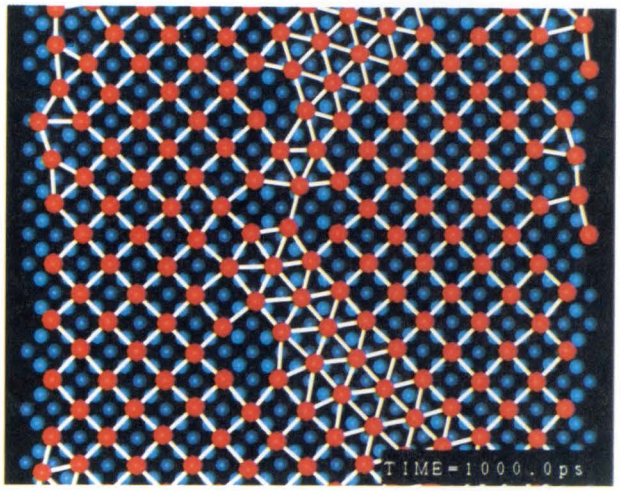


Fig. 10—First overlayer pattern for $\xi = 0.4$ at 1 000 ps.

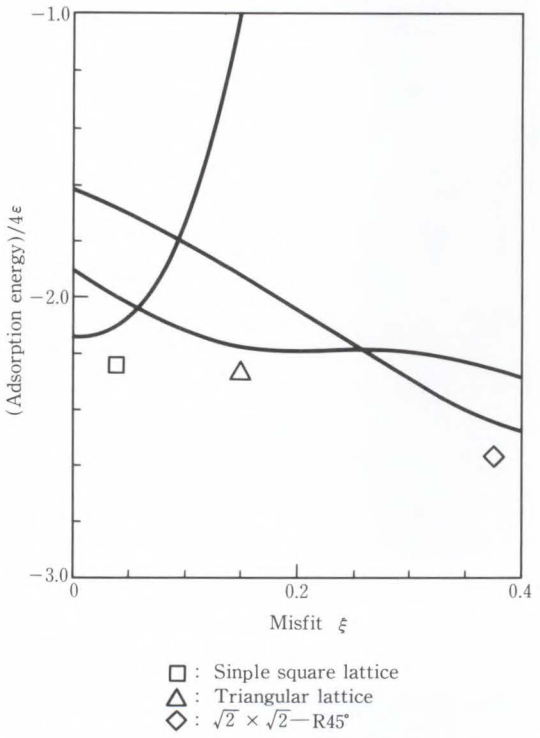
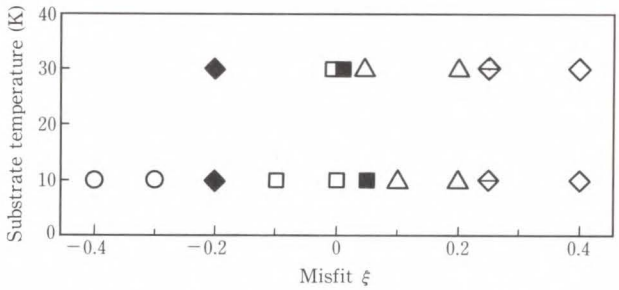
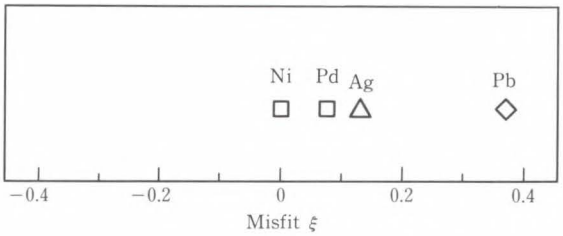


Fig. 11—Misfit ξ dependence of adsorption energy per overlayer atom for three typical growth patterns.

lattice constant of both these lattices equals the diagonal of the substrate's square lattice. This structure corresponds to $\sqrt{2} \times \sqrt{2} - \mathrm{R45}^\circ$ structure. Because there are two directions of rotation, two domains must have appeared. The layer spacing is 40 percent greater than that of the substrate. At $T_s = 30$ K, a single-domain structure is realized.



a) Growth patterns vs. misfit: calculation



b) Growth patterns vs. misfit: experiment

○: Amorphous-like ◆: Square lattice with chain-like modulation
□: Square lattice ■: Square lattice with discommensuration lines
△: Triangular lattice ◇: $\sqrt{2} \times \sqrt{2}$—R45°
◈: Mixed phases of triangular lattice and $\sqrt{2} \times \sqrt{2}$—R45°

Fig. 12—Summary of overlayer growth patterns obtained for L-J lattice mismatched system, and patterns for experimental results of metal heterostructures on Cu substrates.

3.2.3 Phase diagram

There were three atomic arrangements at $\xi \approx 0$, 0.1, and 0.4. The adsorption energy per overlayer atom calculated for these arrangements is plotted in Fig. 11. The lowest energy pattern changes from a square to a triangular lattice as ξ exceeds 0.05, and changes to $\sqrt{2} \times \sqrt{2} - \mathrm{R45}^\circ$ as it exceeds 0.25. Figure 12 summarizes all of the simulated results for $T_s = 10$ K and 30 K. The agreement between Figure 11 and 12 are sufficient to conclude that the complicated growth patterns obtained are energetically most stable under the growth conditions that were used. Figure 12 also illustrates the experimental results for fcc metal[20]. The substrate is Cu and the growth layers and their ξ values (extent of the lattice mismatch to the Cu substrate) are as follows: Ni ($\xi = -0.024$), Pd ($\xi = 0.075$), Ag ($\xi = 0.13$) and Pb ($\xi = 0.37$). The growth patterns of these materials agree well with the L-J system, which seems to justify the application of these results to metallic systems.

3.3 Silicon: surface reconstruction and adatom diffusion

3.3.1 Potentials and surface reconstruction for Si (100)

Many kinds of potential functions have been proposed for Si in the liquid, self-diffusion, and high pressure phases. However, since none of these are intended for use in surface-related studies, several potentials (Stillinger-Weber[21], Biswas-Hamann[22], Kaxiras-Pandey[23], and Tersoff[24] potentials) were evaluated to find a suitable potential for the study of surface phenomena of Si (100).

The 2×1 reconstruction energy per symmetric dimer was calculated for various combinations of distance between dimerized atoms and deviations in the Z direction from the ideal positions of surface atoms. The stable dimers have been observed using the Scanning Tunneling Miscroscopy (STM)[25]. The calculation results were compared with the quantum-mechanical results (norm-conserving pseudopotential method), and the Stillinger-Weber potential was found to be the most suitable.

For further testing of the S-W potential, an ideal Si (100) surface was annealed by applying an MD with the S-W potential incorporated.

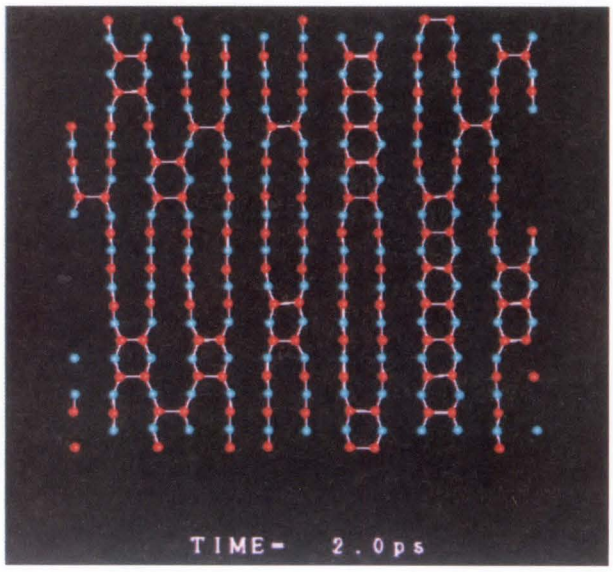


Fig. 13—Si(100) surface after 2 ps of annealing. Red and blue atoms indicate the first and second substrate layer.
2×1 surface reconstruction is evident.

Figure 13 shows the Si (100) surface after 2 ps of annealing. Figure 14 shows the top two surface layers after 10 ps. The 2×1 surface reconstruction proceeds with the annealing process. Most of the top layer atoms form dimers, but the dimer rows are not straight. Also, some atoms are still in the monomer states. The calculation using the S-W potential agrees rather well with the STM observation; therefore, this potential can, up to a point, be applied to the surface phenomena of Si.

3.3.2 Si adatom diffusion on Si (100) reconstructed surface

A Si (100) 2×1 reconstructed surface was

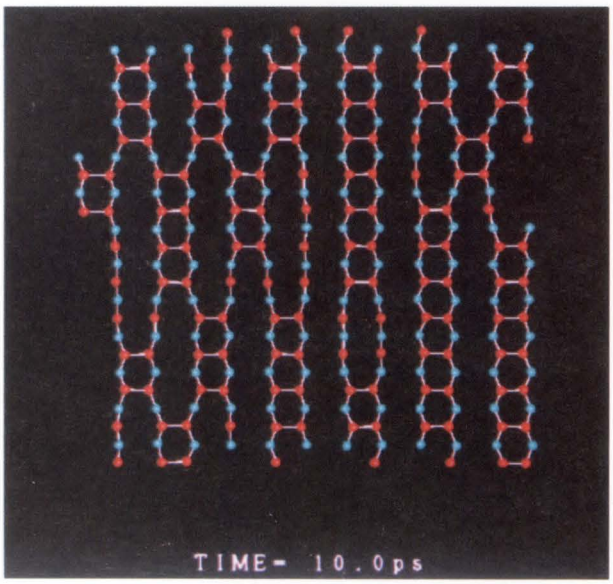


Fig. 14—Top view of Si(100) surface at 10 ps. 2×1 surface reconstruction proceeds with annealing.

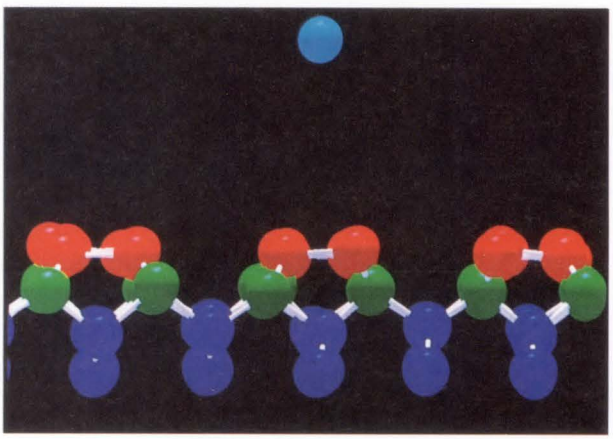


Fig. 15—View from (110) direction in which an incident Si atom is 1.5 nm above bridge site of a dimer.

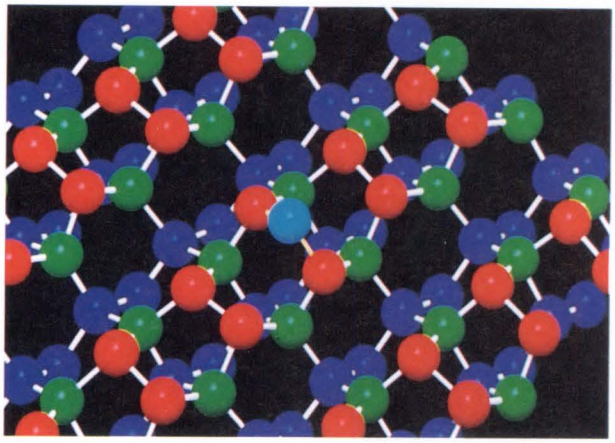


Fig. 16—Si adatom opening a dimer.

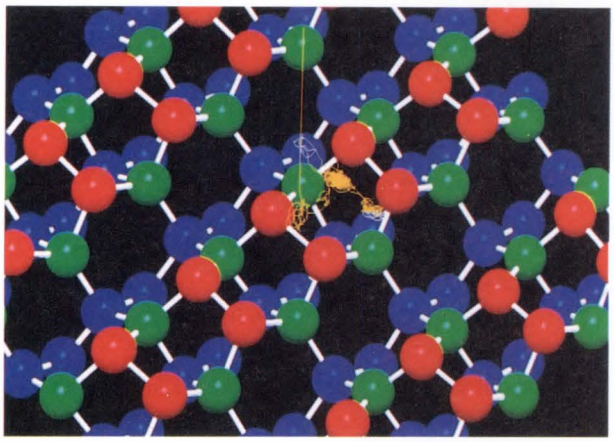


Fig. 17—Trajectory of Si adatom for 25 ps. Positions of substrate atoms are taken from initial state. An adatom moves along the dimer row after it reaches the surface, and then oscillates at the site between dimers.

prepared to investigate the dynamics of an adatom using the MD method with the S-W potential. As shown in Fig. 15, one Si atom appears at 1.5 nm above the bridge site of the dimer. When this atom reaches the dimer, the dimerized atoms cut their bonds, and return to the bulk crystalline positions to bond with the adatoms as shown in Fig. 16. After staying at this position for 2.7 ps, the adatom falls down to the hollow site, while the remaining two atoms (formerly dimerized and bonded to the adatom) bond to reform the dimer. The adatom at the hollow site oscillates for a while, and then propagates along a dimer row. The diffusion is anisotropic because the easy direction of diffusion is along the dimer rows. In Fig. 17, the yellow line shows the trajectory of the adatom for 25 ps. In this figure, the substrate atoms are at their initial positions. As can be seen, there is no stable position for the adatom.

When two adatoms meet, they form a new dimer on the reconstructed surface. In accordance with the experimental results observed by STM, the direction of this new dimer is perpendicular to the original dimer row. This newly-born dimer remains at its creation point throughout the 200 ps simulation. Thus, a created dimer serves as a seed for a one-dimensional island during crystal growth.

4. Conclusion

This paper covered recent studies on atomic-scale simulations of semiconductor materials.

In Si-Ge superlattices, direct transition is obtained by folding the Brillouin zone and applying a stress, and the finite transition probability is obtained by using materials that break the symmetry of the system e.g. $(Si)_5/(Ge)_5$. The calculated transition probability is very small, so further study is necessary to judge whether this superlattice is practicable. For superlattices with a period of ten and a modulated layer structure grown on (100) Ge and alloyed substrates, the band structure is direct but transition is forbidden. The disagreement between the theory and the experimental results (i.e. forbidden transition vs. allowed transition respectively) can be attributed to alloying at the hetero-interfaces.

The MBE process for two-component L-J systems was studied. Various growth patterns were observed, and the generation and behavior of defects in lattice-mismatched systems were observed. The calculation results agree well with experimental results for metallic hetero-structure. The lattice structures were explained in terms of energy, which suggests a wider application of the results to crystal growth in general.

The initial stage of Si crystal growth on Si (100) was studied using molecular dynamics with the Stillinger-Weber potential. (This potential well describes 2 x 1 surface reconstruction.) The diffusion of adatoms is anisotropic because they move more easily along the dimer rows. At 600 K, adatoms oscillate between two

neighboring dimers, or move from site to site with no particular stable point. When two adatoms meet, a new dimer forms and remains at the position of its creation throughout a 200 ps simulation. New dimers may seed surface steps.

The authors hope that this paper will encourage further studies in the material properties of condensed matter using simulation techniques by supercomputer.

This work was done during joint research with Prof. K. Terakura, Dr. H. Ishida of Institute for Solid State Physics, and Dr. T. Oguchi of National Research Institute for Metals.

References

1) Gnutzmann, U., and Clausecker, K.: Theory of Direct Optical Transitions in an Optical Indirect Semiconductor with a Superlattice Structure. *Appl. Phys.*, **3**, pp. 9-14 (1974).
2) Pearsall, T. P., Bevk, J., Feldman, L. C., Ourmazd, A., Bonar, J. M., and Mannaerts, J. P.: Structually Induced Optical Transitions in Ge-Si Superlattices. *Phys. Rev. Lett.*, **58**, 7, pp. 729-732 (1987).
3) Zachai, R., Friess, E., Abstreiter, G., Kasper, E., and Kibble, H.: Band structure and optical properties of strained symmetrized short period Si/Ge superlattices on Si (100) substrates. Zawadzki, W., ed., Proc. 19th Int. Conf. on Physics of Semiconductors, Vol. **1**, Warsaw, Polland, 1988, pp. 487-490.
4) Hybertsen, M. S., Friedel, P., and Schluter, M.: Theory of low energy optical transitions in Si/Ge strained layer superlattices. Zawadzki, W., ed. Proc. 19th Int. Conf. on Physics of Semiconductors, Vol. **1**, Warsaw, Polland, 1988, pp. 491-494.
5) Froyen, S., Wood, D. M., and Zunger, A.: Structural and electronic properties of epitaxial thin-layer Si_nGe_n superlattices. *Phys. Rev.*, **B37**, 12, pp. 6893-6907 (1988).
6) Hybertsen, M. S., and Schluter, M.: Theory of optical transitions in Si/Ge (001) strained-layer superlattices. *Phys. Rev.*, **B36**, 18, pp. 9683-9693 (1987).
7) Satpathy, S., Martin, R. M., and Van de Walle, C. G.: Electronic properties of the (100) (Si)/(Ge) strained-layer superlattices. *Phys. Rev.*, **B38**, 18, pp. 13237-13245 (1988).
8) Ikeda, M., Terakura, K., and Oguchi, T.: Theoretical study on band-gap character of $(Si)_m/(Ge)_n$ strained superlattices. Anastassakis, E. M., and Joannopoulos, J. D., ed., Proc. 20th Inte. Conf. on Physics of Semiconductors, Vol. 2, Thessalonike, Greece, 1990, pp. 889-892.
9) Ikeda, M., Terakura, K., and Oguchi, T.: THEORETICAL STUDY ON BAND-GAP CHARACTER OF $(Si)_m/(Ge)_n$ STRAINED SUPERLATTICES, Doyama, M., ed., Proc. Int. Conf. on Comput. Appl. to Materials Sci. and Eng. - CAMSE '90, Tokyo, Japan, 1990.
10) Tung, R. T., Dawson, L. R., and Gunshor, R. L.: Epitaxy of Semiconductor Layered Structures. Proc. Material Research Society symposia proceedings, 1987, **102**.
11) Schneider, M., Rahman, A., and Schuller, I. K.: Role of Relaxation in Epitaxial Growth: A Molecular-Dynamics Study. *Phys. Rev. Lett.*, **55**, 6, pp. 604-606 (1985).
12) Hara, K., Ikeda, M., Ohtsuki, O., Terakura, K., Mikami, M., Tago, Y., and Oguchi, T.: Molecular-dynamics simulations for molecular-beam epitaxy: Overlayer growth pattern in two-component Lennard-Jones systems. *Phys. Rev.*, **B39**, 13, pp. 9476-9485 (1989).
13) Schneider, M., Schuller, I. K., and Rahman, A.: Epitaxial growth of silicon: A molecular-dynamics simulation. *Phys. Rev.*, **B36**, 2, pp. 1340-1343 (1987).
14) Gawlinski, E. T., Gunton, J. D.: Molecular-dynamics simulation of molecular-beam epitaxial growth of the silicon (100) surface. *Phys. Rev.*, **B39**, 9, pp. 4774-4781 (1989).
15) Lampinen, J., Nieminen, R. M., and Kaski, K.: MOLECULAR DYNAMICS SIMULATION OF THE STRUCTURE AND MELTING TRANSITION OF THE Si (001) SURFACE. *Surface Sci.*, **200**, pp. 101-112 (1988).
16) Srivastava, D., Garrison, B. J., and Brenner, D. W.: Anisotropic Spread of Surface Dimer Openings in the Initial Stages of the Epitaxial Growth of Si on Si{100}. *Phys. Rev. Lett.*, **63**, 3, pp. 302-305 (1989).
17) Umebu, I., Ikeda, M., Yamasaki, T., Furuya, K., and Terakura, K.: MOLECULAR-DYNAMICS SIMULATION OF MOLECULAR-BEAM-EPITAXY USING LENNARD-JONES AND EMPIRICAL Si POTENTIALS. Doyama, M., ed., Proc. Int. Conf. on Comput. Appl. to Materials Sci. and Eng. - CAMSE '90, Tokyo, Japan, 1990.
18) Andersen, O. K.: Linear methods in band theory. *Phys. Rev.*, **B12**, 8, pp. 3060-3083 (1975).
19) Muller, E., Nissen, H. -U., Ospelt, M, and von Kanel, H.: Chemical ordering and boundary structure in strained-layer Si-Ge superlattices *Phys. Rev.*

Lett., **63**, 17, pp. 1819-1822 (1989).
20) Koma et al., ed., Handbook on Solid Surface Engineering (in Japanese), Maruzen, 1987, pp. 295-312.
21) Stillinger, F. H., and Weber, T. A.: Computer simulation of local order in condensed phase of silicon. *Phys. Rev.*, **B31**, 8, pp. 5262-5271 (1985).
22) Biswas, R., and Hamann, D. R.: New classical models for silicon structural energies. *Phys. Rev.*, **B36**, 12, pp. 6434-6445 (1987).
23) Kaxiras, E., and Pandey, K. C.: New classical potential for accurate simulation of the atomic processes in Si. *Phys. Rev.*, **B38**, 17, pp. 12736-12739 (1988).
24) Tersoff, J.: Empirical interatomic potential for Si with improved elastic properties. *Phys. Rev.*, **B38**, 14, pp. 9902-9905 (1988).
25) Tromp, R. M., Hamers, R. J., and Demuth, J. E.: Si (001) Dimer Structure Observed with Scanning Tunneling Microscopy. *Phys. Rev. Lett.*, **55**, 12, pp. 1303-1306 (1985).

Minoru Ikeda
Semiconductor Crystals Laboratory
FUJITSU LABORATORIES, ATSUGI
Bachelor of Science
Saitama University 1975
Dr. of Science
Osaka City University 1983
Specializing in Solid State Physics

Kumiko Furuya
Semiconductor Crystals Laboratory
FUJITSU LABORATORIES, ATSUGI
Bachelor of Science
University of Tsukuba 1984
Master of Science
University of Tsukuba 1986
Specializing in Solid State Physics

Takahiro Yamasaki
Semiconductor Crystals Laboratory
FUJITSU LABORATORIES, ATSUGI
Bachelor of Science
Kyoto University 1984
Dr. of Physics
Osaka University 1989
Specializing in Solid State Physics

Masuhiro Mikami
Sientific Systems Dept.
Systems Engineering Group
FUJITSU LIMITED
Bachelor of Physics
Rikkyo University 1975
Master of Chemical Eng.
Tokyo Institute of Technology 1978
Specializing in Computational Chemistry

UDC 532.5.01:681.32

Computational Fluid Dynamics and Computers

• Satoru Ogawa • Yoko Takakura

(Manuscript received December 3, 1990)

This paper describes the participation of computational fluid dynamics (CFD) and computers. First, the history and outline of CFD are briefly explained. Next, advanced researches (computations of transonic flow with large separation, supersonic flow around complex configurations, supersonic flow with combustion, and hypersonic flow of real gas) in Computation laboratory of National Aerospace Laboratory are shown. Finally, turbulence and the future problems of CFD are described as impacted by computer performance.

1. Introduction

Computational fluid dynamics (CFD) is the science of producing numerical solutions to a system of partial differential equations which describe fluid motions. Over the past several years, CFD has emerged as an extremely valuable scientific tool in various related fields due to the development of the supercomputer.

The authors are engaged in the application of CFD to the fields of aeronautics and astronautics, in which fields the usefulness of CFD is rapid. To develop a highly efficient aeroplane or engine today, the costs are high. The judicious application of CFD can greatly reduce developmental costs by partially replacing wind tunnel tests. In designing the aerodynamic configuration, parametric numerical computations can be performed very quickly so that configurations with poor performance may be discarded. Though wind tunnel tests measure only global characteristics and surface properties, computational solutions provide detailed information of flow properties throughout the entire flow-field. We can easily understand what happens in the flow fields by employing CFD.

Here the authors would like to present briefly the history, outline and examples of numerical computations. Finally, future directions of CFD are discussed, bearing in mind its relation to computer performance.

2. Outline of computational fluid dynamics

Because of drastic nonlinearity in the governing equations of fluid dynamics, analyzing the equations is difficult and many researchers have puzzled over it since the eighteenth century. Equations solvable by the classic analytic approach are limited to simplified ones containing such assumptions as symmetry and physical modelings. With the development of computers and numerical computing methods during the last thirty years, however, it has become possible to solve various flow fields without modelings. The appearance of supercomputers, in particular, has completely changed the quality and scale of CFD. The importance of CFD has, thus, expanded rapidly. The brief history of fluid dynamics and CFD in relation to Japanese developments is dealt with in chapter 6 Appendix.

The discipline of CFD, a large branch of scientific computing, has recently undergone rapid growth. It is composed of related disciplines: fluid mechanics, numerical analysis, geometry, and other specialities. The primary elements of CFD are briefly described in the next sections.

2.1 Governing equations

The motion of continuous material such as water or air can be described by the Navier-

Stokes equations. A characteristic of these equations is that the viscous stress is proportional to the velocity gradient, which is well confirmed experimentally. The equations for inviscid flow are called the Euler equations, and the assumption of non-vortical flow leads to the potential equation.

Though the Navier-Stokes equations are completely general, using them blindly does not always lead to proper solutions when they are applied to CFD techniques. For example, consider the high-speed flow around a body in which the flow fields may be characterized as primarily inviscid except for a very thin viscous region near the body. This viscous region is termed the boundary layer, the thickness is in the order of $(-1/2)$th power of the Reynolds number, and the change of the flow state is steep. Hence, an enormous number of grid points are required to resolve the thin viscous layer. If we solve the Navier-Stokes equations with coarse numerical grids, the solution may be miserably blunt. Therefore, various levels of approximations to the Navier-Stokes equations are used to obtain relatively efficient solutions. In some cases the use of inviscid equations produces a good solution in close agreement with the experimental data. Euler equations can save much computer time partly because there are fewer operations, and also because the number of grid points is drastically lower. Only after considering the physics of a given problem and the computer's limitations, should it be decided which equations to solve.

The Navier-Stokes equations are not universally applicable. It should be noted that these differential equations are meaningful only when the variations of physical values in time and space are so moderate that the differential is meaningful. For an extreme example, consider the atmospheric motion around the earth. When the entire earth is covered with the numerical grid, the fineness of the intervals would be, at best, only tens of kilometers, using the largest-performance computers available today. If numerical simulation using a grid interval of tens of kilometers is possible, numerical values that change almost linearly between the grid intervals would be obtained. But that is unacceptable, because this result indicates, for example, that almost the same wind blows all over Tokyo. In other words, it is fundamentally impossible to capture variations smaller than the grid interval. That introduces the necessity for the turbulence model in which turbulent viscosity is substituted for these small variations and by which physically reasonable solutions are obtained.

With existing limitations in computer performance, turbulence can be directly computed without using turbulence models in only a few cases where the Reynolds number is low. For problems in the field of aeronautics and astronautics, where flying planes generate high Reynolds numbers (10^7-10^8), we cannot avoid relying on turbulence models for the time being.

2.2 Grid generation

One of the most important steps in solving CFD problems accurately using finite difference procedures is the distribution of grid points in the flow region around the body. If a beautiful numerical grid is generated, it does not matter what kind of generation technique may be used. However, no definite criteria exist for judging the excellence of the numerical grid. In evaluating the quality of numerical grids, excellence is almost a subjective judgment.

To automatically generate numerical grids, many and many generation techniques have been proposed[1], i.e. elliptic, parabolic, hyperbolic, and algebraic methods. Each grid generation method has its own characteristics, and the suitable method is selected after considering such matters as the computational domain, topology of the numerical coordinates and the flow conditions. In the opinion of the authors, combining the algebraic method with the interpolation technique is simple and best, when contrasted with other methods for solving differential equations. Considering that the resolution of any solution is directly dependent on the grid point interval, the grid points must be concentrated where the difference of physical quantities is large, that is, the boundary layer, the shock wave and the contact discontinuity.

2.3 Numerical procedures

In the last 30 years, there has been remarkable progress in developing numerical algorithms to solve both inviscid and viscous flow equations. In the early stage of CFD, the Lax-Wendroff scheme[2] was the most successful, and many flow problems, primarily for Euler equations, were solved by using the scheme. This second-order scheme permitted greater resolution of the entire flow field, especially in the vicinity of the shock wave. Details of the development of numerical algorithms can be found in a textbook by Yazima and Nogi[3]. By the late 1970's computer performance had developed far enough to solve the Navier-Stokes equations, and primary approach at that time was the Beam-Warning scheme[4]. The characteristics of the Beam-Warming scheme are:

1) the fourth-order numerical dissipation is added to the central difference of numerical flux to suppress the numerical oscillation,
2) the implicit ADI method is used to increase the convergence rate to the steady state.

Though the Beam-Warming scheme itself was not revolutionary, it was an excellent and practical scheme overall. In actual computations, the most troublesome areas had to do with a number of technical procedures such as satisfying the boundary conditions or increasing the convergence rate to a steady solution. In NASA's Ames Research Center, the scheme has since been improved, and it has been used for many excellent computations.

Analytical methods for solving discontinuity in shock tubes have been studied in applied mathematics. It is Riemann's problem on the initial value in hyperbolic partial differential equations. Godunov's method[5], the numerical flux of which is calculated by using the exact solution of the Riemann problem, is a well-known and excellent numerical method that does not produce numerical oscillation near the shock wave. However, it was not used till quite recently because;

1) computation takes much time since the Riemann problem must be solved exactly at the grid point intervals,
2) its resolvability is poor because it is a first-order scheme.

In the 1980's the Godunov type scheme entered practical use in the aeronautical and astronautical fields due to several advances such as higher performance computers, the study of approximate Riemann solver by Osher[6] and Roe[7], and the higher-order scheme by Van Leer[8] and Harten[9] et al. Nowadays, the numerical approach called the Total Variation Diminishing (TVD) scheme, or upwind scheme, belongs to this type. These methods solve the flow without numerical oscillations near discontinuities such as the shock wave and the contact discontinuity. There is also no need to adjust the numerical dissipations in solving the flow for a wide range of Mach numbers. At present, the TVD scheme is the principal one used for solving compressible flow problems.

Current computer technology makes it possible to analyze fluid dynamics coupled with chemical reactions. However, incorporating a chemical reaction model with the fluid dynamics results in a number of difficulties. The first TVD method for treating the system of chemical reactions was the first-order scheme proposed by Eberhart and Brown[10]. Second-order TVD schemes have been proposed by Yee et al.[11] and Wada et al.[12]. An evaluation of TVD schemes that include chemical reactions is described in detail by Wada[13].

3. Examples of numerical computations

This chapter presents examples of numerical computations done recently by the group of the authors at National Aerospace Laboratory (NAL). These were computed by the numerical simulator system[14], NAL's computer system, designed around the FUJITSU VP-400 supercomputer with a memory of 1 Gbyte.

3.1 Flow around a three-dimensional wing

The purpose of this computation is to estimate finite-difference methods and to validate turbulence modeling methods. The ONERA-M6 wing[15] was taken as a case study, since this three-dimensional wing has a great store of experimental data, in the wind tunnel. The computations of this flow for estimation

and validation have been performed for more than three years. It would be difficult to find a more thoroughly analyzed example.

First, the computations of this flow were done to estimate the applicability of the TVD scheme (see chapter 6 Appendix). At that time it was said that the TVD schemes[16),17)] were of no use for the three-dimensional flow, partially because at that time the Beam-Warming scheme[4)] was thriving. But after the geometrical treatment of the TVD schemes was improved[18)], it was confirmed that the TVD schemes capture the shock wave without numerical oscillations and more clearly than does the Beam-Warming scheme. Their solutions also agree well with experimental data.

Next, turbulence models were validated under the same flow conditions. For turbulence models, the algebraic model[19)] has been in mainstream use, and nearly all researchers treating this model reported that solutions using the algebraic model agreed well with experiments. But it is known that when the separation region of flow is slightly larger, the discrepancy between the numerical solution and experiments becomes remarkable. Hence, the two-equation model[20),21)] and subgrid-scale model[22)] have been applied[23)]. The computation has been tried in cases where a triple shock wave, strong and weak shock waves and their united one, is formed and the interaction of shock wave with boundary layer is important because of the large separation, behind the united shock wave. The conclusion is that the two-equation model is promising[24)]. Figure 1 shows the pressure distribution in a solution using the two-equation model where the Mach number is 0.84 and the attack angle is 6.06°. In this figure the triple shock wave is observed on the wing-surface distribution and the interaction between a shock wave and boundary layer is also recognized in the spatial distribution. The computational results show good agreement with experiments in this case. However, creation of a universal turbulence model which could describe large separation in every case would be an important and indispensable study in the future.

3.2 Flow around a complexly configured vehicle[25)]

As the numerical computation has become more practicable, there have been efforts to solve numerically the flow around an entire plane. At the present level of computer performance, the maximum number of grid points usable in the flow calculation is about one million. One million points are too few to resolve the actual physical phenomena in the three-dimensional flow around an entire plane. Only rough solutions are calculated in the present step, therefore. Shown here is the computational example of supersonic flow around the combination of H-II rocket, booster and minishuttle HOPE planned by the National Space Development Agency of Japan. For a

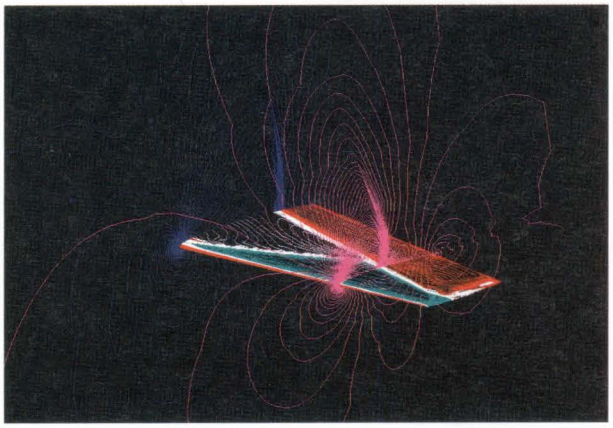


Fig. 1—Transonic flow around ONERA-M6 wing (solution by use of two-equation turbulence model).

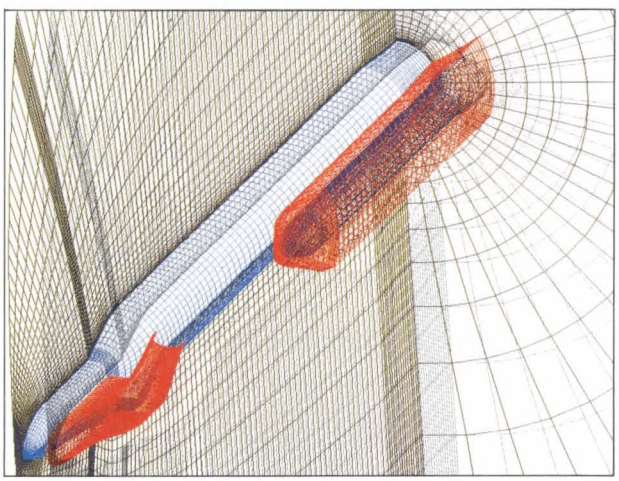


Fig. 2—View of embedded grids in multi-domain technique.

simple computational domain it is easy to cover the domain by a single grid, but for a complex domain it is difficult to generate a grid. As a compromise, the multi-domain technique is used. The whole is divided into three domains, i.e. those, surrounding the main body (H-II rocket and fuselage of HOPE), the booster and the wing of HOPE, For each domain, grids are generated (see Fig. 2).

On these grids, hypersonic flow has been numerically solved using the TVD scheme. Figure 3 shows the surface pressure distribution under inflow conditions of Mach number 1.8 and attack angle 3.0°. In spite of fairly rough computation, the fundamental flow phase is considered to be captured.

The multi-domain technique seems to be relatively simple and appears promising for future use. At the same time, the method[26] referred to recently, which uses Cartesian coordinates and evaluates the physical value correctly by making the grid fine near the wall, also seems to be gaining adherents. As computer performance continues to improve, there will be little difficulty in adapting flow calculations not only to simplified configurations but also to real configurations.

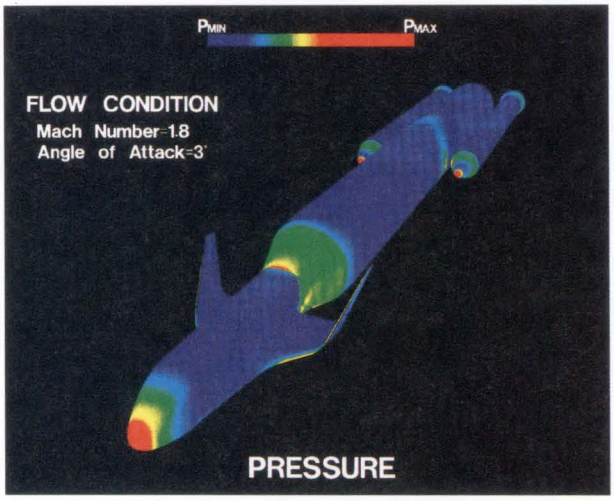


Fig. 3—Supersonic flow around combination of H-II rocket, booster and mini-shuttle HOPE.

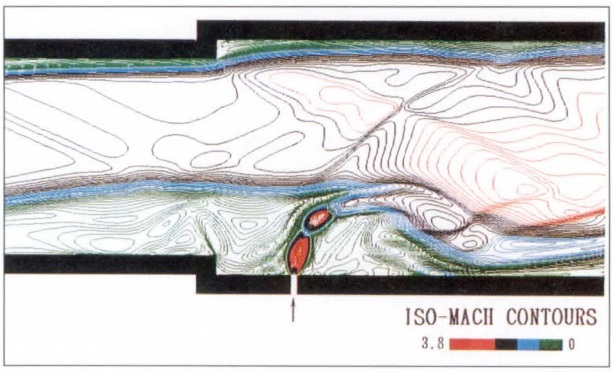


Fig. 4—Flow with supersonic combination in SCRAM jet engine.

3.3 Chemically reacting flow in a combustor[27]

NAL is pursuing basic research and development for a space plane which can go into space and return with ease. The most crucial element to this development is a supersonic combustion ramjet (SCRAM) engine. When flying more than ten times the speed of sound, the use of an ordinary jet engine with a compressor would drastically lower efficiency because of the shock wave on the compressor blades. The SCRAM jet engine is, therefore, planned because of its efficient operation which substitutes the ram pressure of high-speed gas for a compressor.

As the computation of chemically reacting flow has become possible with the progress of computers, the computational condition for this example almost coincide with the actual experimental condition for the SCRAM jet engine. In the fundamental physical phenomena of the SCRAM engine, the hydrogen blows up into a high temperature gas and burns. For the combustion, it is necessary to introduce the reacting model. In the westbrook reacting model, nine chemical species, i.e. N_2, H_2, O_2, OH, H_2O, H, O, H_2O_2 and HO_2, are considered and 17 elementary reaction steps are contained. The governing system is made up of 14 equations. Five equations correspond to the usual gas ones, and there are nine transport equations for each species, so there are more than three times as many operations as in the usual gas equations.

The chemically reacting flows in the combustor where the hydrogen blows up have been numerically solved using the TVD scheme for the governing equation system mentioned above. An example is shown in Fig. 4, where iso-Mach contours are shown. Thus, by using CFD, it is

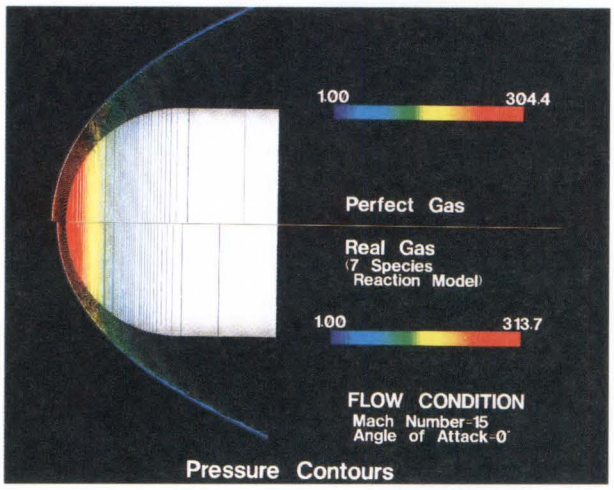


Fig. 5—Efect of real gas in hypersonic flow.

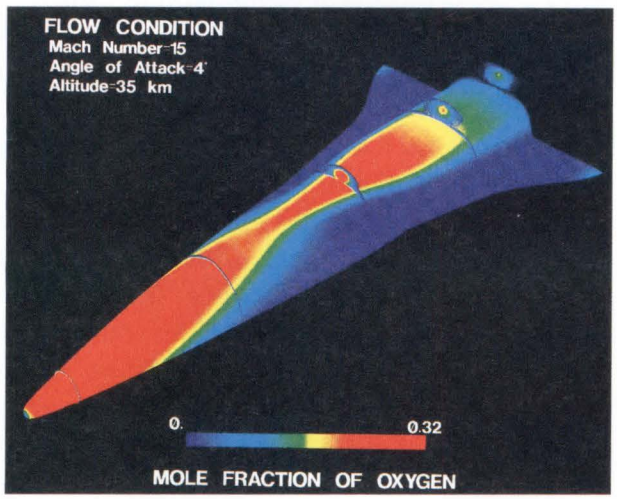


Fig. 6—Hypersonic flow of real gas around space plane.

possible to see the flow details such as the Mach disk formed by the blow. This is difficult to measure experimentally. Since the problem of turbulence in reacting flows remains largely unsolved, the reacting ratio does not show good agreement between the numerical solution and the experiments. The reason for this disagreement is that the reacting ratio greatly depends on the mixing of hydrogen and oxygen, which the turbulent diffusion governs. The problem of turbulence in reacting flows would be an important research theme in the future.

3.4 Hypersonic flow of real gas[28)]

Since the space plane flies more than ten times the sonic speed, an extremely strong shock wave appears ahead of it. The extreme compression following the strong shock wave causes a temperature of more than ten thousand degree near the plane. Thus, the nitrogen and oxygen in the atmosphere dissociate, consequently the usual assumption of a perfect gas no longer holds good and it becomes necessary to include the effect for the real gas. In this example, the elemental reactions for the seven components of N_2, O_2, N, O, NO, NO^+, and e^- are considered in the dissociation in order to include this real gas effect. Figure 5 shows the pressure distribution on a blunt body flying at a speed of Mach 15. The upper half shows the numerical solution for the usual perfect gas, while the lower half shows that for the real gas. The figure illustrates that in the real gas the shock wave is situated nearer to the blunt body than in the perfect gas, since the temperature is reduced by the endothermic reaction of dissociation.

When the shock wave strikes the wing of the space plane, the plane suffers severe aerodynamic heating. Therefore, predicting the position of the shock wave is very important, and correctly evaluating the real gas effect is necessary. Figure 6 shows the numerical solution for the hypersonic flow of Mach 15 around the space plane with the real gas effect included. The figure shows the mole fraction distribution of the atomic oxygen produced by the dissociation. The atomic oxygen near the nose of the plane is transported with the fluid and gathers near the center on both the upper and lower surfaces separately. On the upper surface, it is transported with the separated fluid and spreads over the plane again. It is extremely difficult to recreate this hypersonic flow in wind tunnel experiments, and except for actual in-flight experiments, using numerical computation alone can be a very powerful means to predict such flows.

As this example illustrates, numerical computation will become increasingly important for analyzing extreme situations which are impossible to predict experimentally.

4. Future prospects of computational fluid dynamics

CFD developments have made it possible to obtain numerical solutions that hold true to some extent for a large variety of problems. The problems for chemical reactions and both radiation and electro-magnetic fluid dynamics can be numerically solved without difficulty, though it demands much computing time. The problem still remaining, however, is the remarkable characteristic turbulence of nonlinear fluid motions. To capture the turbulent flow on a full scale by the direct simulation of the Navier-Stokes equations, supercomputers with higher performance and larger memory are highly desirable.

The turbulence consists of various scale eddies, with the size of the smallest turbulent eddy (Kormogorov scale) being proportional to the (−3/4)th power of the Reynolds number (Re). The computer performance required to simulate the turbulent eddies is estimated by Chapman[29]. In this estimate, he assumed 1) to simulate directly 90 percent of the kinetic energy of turbulence, 2) to locate at least five grid points to resolve an eddy, and 3) to use a nested grid, that is, locally refined grid, to resolve the viscous sublayer efficiently. To simulate three-dimensional flows at $Re = 10^7$, according to his estimation, about 4×10^8 grid points are needed for a wing with uniform section, and more than 10^{10} grid points are needed for a complex aircraft. Even using the maximum capacity of existing computers, however, numerical computations are limited to millions grid points at most. Since computing time is roughly proportional to the number of grid points, the simulation of turbulent eddies using a high Reynolds number is at present impossible under the conditions that Chapman posits.

According to the numerical simulation of turbulent flow between two parallel plates[30),31)], however, the turbulent structure is captured to some extent on a coarser grid than that prescribed by Chapman. Hence, it may be easier to capture the turbulent structure than Chapman estimates. For the time being, turbulence models should be vigorously researched, particularly with regard to the management of small eddies.

The use of supercomputers has made three-dimensional flow problems easily and accurately solvable within a reasonable time frame. The information on flow properties obtained by numerical computations is so enormous that it is difficult to understand flow fields without using devices such as graphic displays. It is even necessary to generate color movies of the solutions so that unsteady phenomena such as turbulence can be understood. Visualizing the solutions for two-dimensional flow problems is easy. However for three-dimensional information, we are usually troubled with how best to display the solution data for optimum understanding and analysis. To increase the usefulness of CFD, further development of graphic display devices and easy-to-use support software will be required.

5. Conclusion

The coming of the supercomputer has accelerated the development of computational fluid dynamics (CFD). With the use of the highly accurate numerical method, it has been possible to solve even the flow field around a complex configuration, which was previously impossible. However, the fluid is accompanied by minute variations in physical quantities such as separation and turbulent eddies, and by discontinuity of physical quantities such as shock waves. Trying to capture this complexity of fluid motion more minutely and more accurately by numerical simulations would mean a limitless demand for ever-greater computer power.

There was once a man who felt the mutability of life for bubbles floating on a river stream (Japanese classic literature). There was also a man who admired clouds for their genius (Japanese modern literature). Today, by using numerical simulation with computers under a single physical law, it has become possible to capture the diverse behavior of protean fluid, which has attracted people from ancient times. And with every new day, CFD will flourish. When people greet the new century, what high computer performance will have been achieved!

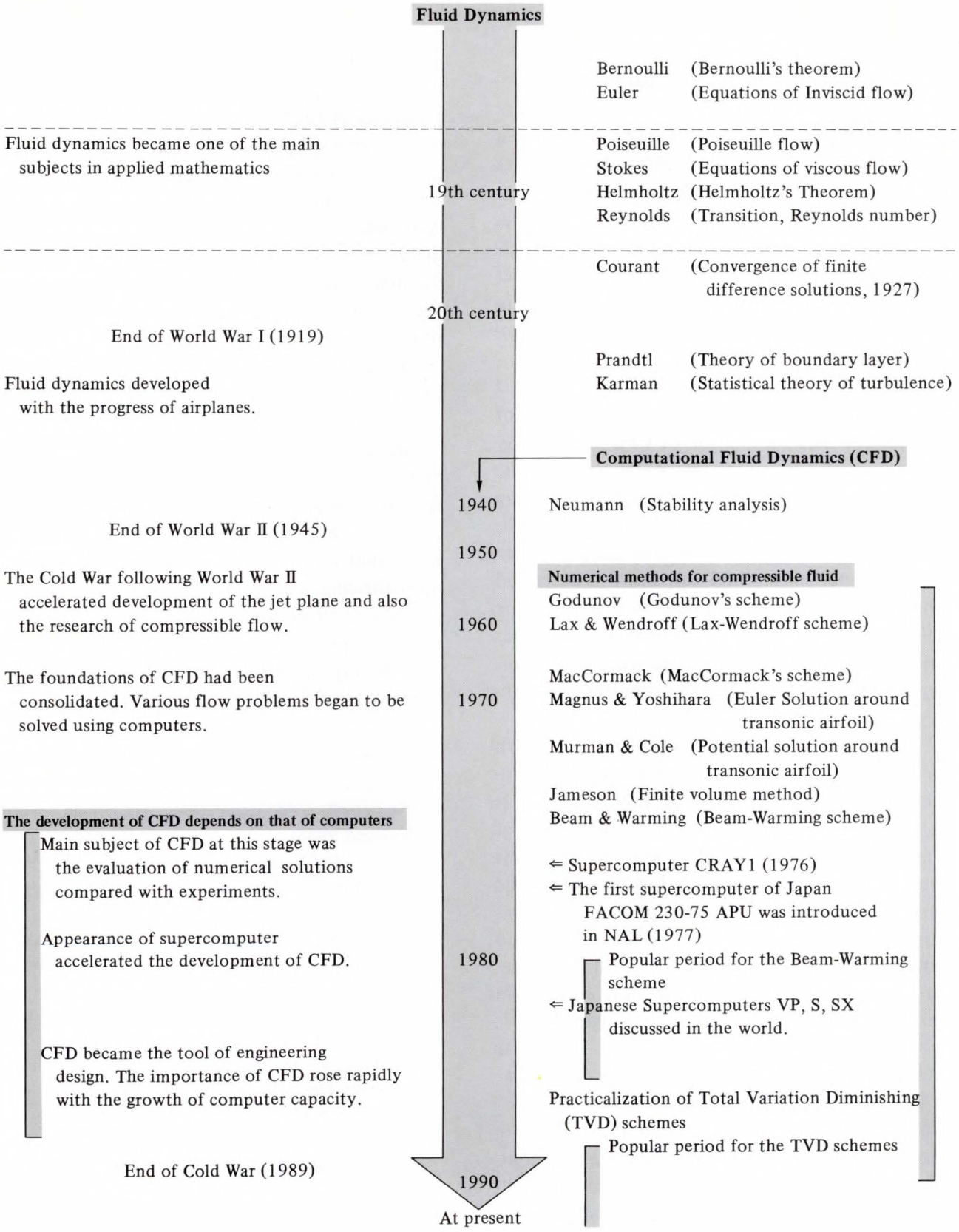


Fig. A1—Brief history of fluid dynamics and computational fluid dynamics.

What excellent numerical simulations there will be!

6. Appendix

6.1 Brief history of computational fluid dynamics

Computational fluid dynamics (CFD) is a branch of fluid dynamics (FD). It was after twentieth century that CFD appear on the stage of long history of FD since nineteeth century. Thereafter CFD have grown rapidly as a powerful support for FD. Figure A1 presents the brief history on the development of FD and CFD in a chronology, reflecting development of computers and social conditions.

As shown in Fig. A1, the developmental phase of CFD closely corresponds to that of computers. The foundation of CFD had already been established mathematically before the appearance of computers. The need to develop highly operational devices was strongly felt. Indeed, it is surprising to find that the prototype of the numerical methods used now is described almost exactly in a textbook[32] written in the 1960's. These numerical methods have, one after another, been put into practice with the appearance of full-scale computers supported by semiconductor technology.

As is well known, the pioneering CFD work in Japan was the flow simulation around a two-dimensional cylinder done by Kawaguchi[33]. It is said that he evolved the tiger computer over one and a half years. Similar numerical solutions would be obtained within a few seconds if the same computation were performed on a supercomputer, and further the numerical method which Dr. Kawaguchi used is still available, exclusive of accuracy. In the flow simulation in a two-dimensional nozzle by Ishiguro[34),35], which was the early work in Japan on regular numerical computation for compressible flow, it is said that Ishiguro calculated the converged solutions with a computational time of more than one hundred hours by dividing the computational domain into eight parts and repeatedly wrote to and read from files. Today, after twenty years, it is felt that computational methods have not advanced greatly, but that the performance of computers has made rapid progress.

The appearance of supercomputers has changed the quality and scale of CFD completely. Supercomputers with operational speeds that are scores of times faster than conventional computers have allowed three-dimensional calculations for problems on which only two-dimensional calculations have been possible before. And since it has enabled the application of highly accurate computational methods involving many operations and much computing time, the accuracy of solutions has remarkably improved.

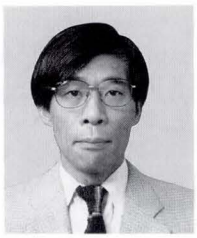
In NAL, Mr. Hajime Miyoshi, the director of Computational Sciences Division, has long maintained the importance of CFD, and the fastest computers in Japan have been installed at every opportunity. In this connection, it is worthy of special mention that the first Japanese supercomputer, FACOM 230-75 APU (22 MFLOPS), was installed in 1977. In this Laboratory, CFD is now thriving and being used increasingly as a partial substitute for wind tunnel experiments in the developmental design of airplanes.

The importance of CFD is rising rapidly with the growth of computer performance.

References

1) Tompson, J.F., Warsi, Z.U.A., and Mastin, C.W.: Numerical Grid Generation. Amsterdam, North Holland, 1985.
2) Lax, P., and Wendroff, B.: Systems of Conservation Laws. *Comuun. Pure Appl. Math.*, **XIII**, pp. 217-237 (1960).
3) Yazima, N., and Nogi, T.: Numerical Analysis of Evolutional Equations. (in Japanese), Iwanami, 1977.
4) Steger, J.L., and Warming, R.F.: Flux vector splitting of the Inviscid Gasdynamic Equations with Application to Finite Difference Methods. *J. Comput. Phys.*, **40**, pp. 263-293 (1981).
5) Godunov, S.K.: Finite-Difference Method for Numerical Computation of Discontinuous Solutions of the Equations of Fluid Dynamics. *Mat. Sbornik*, **47**, pp. 271-306 (1959).
6) Osher, S., and Solomon, F.: Upwind Difference Schemes for Hyperbolic Conservation Laws. *Math. Comput.*, **38**, 158, pp. 339-374 (1982).

7) Roe, P.L.: Approximate Riemann Solvers, Parameter Vectors and Difference Schemes. *J. Comput. Phys.*, **43**, pp. 357-372 (1981).
8) Van Leer, B.: Towards the Ultimate Conservative Difference Schemes. *J. Comput. Phys.*, **43**, pp. 101-136 (1981).
9) Harten, A.: High Resolution Schemes for Hyperbolic Conservation Laws. *J. Comput. Phys.*, **49**, pp. 357-393 (1983).
10) Eberhardt, S., and Brown, K.: A Shock Capturing Technique for Hypersonic, Chemically Reacting Flows. AIAA paper 86-0231, 1986.
11) Shinn, J.L., Yee, H.C., and Uenishi, K.: Extension of a Semi-Implicit Shock Capturing Algorithm for 3-D Fully Coupled, Chemically Reacting Flows in Generalized Coordinates. AIAA paper 87-1577, 1987.
12) Wada, Y., Kubota, H., Ogawa, S., and Ishiguro, T.: A Diagonalizing Formulation of General Real Gas-Dynamic Matrices with a New Class of Schemes. AIAA paper 88-3596-CP, 1988.
13) Wada, Y., Kubota, H., Ogawa, S., and Ishiguro, T.: A Generalized Roe's Approximate Riemann Solver for Chemically Reacting Flows. AIAA paper 89-0202, 1989.
14) Tsuchiya, M., and Morishige, H.: Supercomputer Operation Management System at National Aerospace Laboratory. (in Japanese) *FUJITSU*, **40**, 2, pp. 59-65 (1989).
15) Schmitt, V., and Charpin, F.: Pressure Distributions on the ONERA-M6 WING at Transonic Mach Numbers. AGARD AR-138-B1, 1979.
16) Yee, H.C., and Harten, A.: Implicit TVD Schemes for Hyperbolic Conservation Laws in Curvilinear Coordinates. AIAA paper 85-1513, 1985.
17) Chakravarthy, S.R., and Osher, S.: A New Class of High Accuracy TVD Schemes for Hyperbolic Conservation Laws. AIAA paper 85-0363, 1985.
18) Takakura, Y., Ishiguro, T., and Ogawa, S.: On the Recent Difference Schemes for the Three-Dimensional Difference Schemes. AIAA paper 87-1151, 1987.
19) Baldwin, B.S., and Lomax, H.: Thin layer Approximation and Algebraic Model for Separated Turbulent Flows. AIAA paper 78-257, 1978.
20) Jones, W.P., and Launder, B.E.: The Prediction of Laminarization with a Two-Equation Model of Turbulence. *Int. J. Heat and Mass Transfer*, **15**, pp. 301-304 (1972).
21) Coakley, T.J.: Turbulence Modeling Methods for the Compressible Navier-Stokes Equations. AIAA paper 83-1693, 1983.
22) Deardorff, J.W.: A Numerical Study of Three-Dimensional Turbulent Channel Flow at Large Reynolds Numbers. *J. Fluid Mech.*, **41**, pp. 452-480 (1972).
23) Takakura, Y., Ogawa, S., and Ishiguro, T.: Turbulence Models for 3-D Transonic Viscous Flows. AIAA paper 89-1952, 1989.
24) Takakura, Y., Ogawa, S., and Ishiguro, T.: Turbulence Models for 3-D Transonic Viscous Flows II. Proc. of ISCFD, Nagoya, 1989.
25) Ogawa, S., and Wada, Y.: Numerical Computations of Supersonic Flows around the Combined Configurations of Hope and H-II Rocket. Proc. of HOPE Workshop, NASDA, 1989.
26) Young, D.P., Melvin, R.G., Bieterman, M.B., Johnson, F.T., Samant, S.S., and Bussoletti, J.E.: A Locally Refined Rectangular Grid Finite Element Method. *J. Comput. Phys.*, **92**, pp. 1-66 (1991).
27) Ishiguro, T., Ogawa, S., Wada, Y., and Masuya, G.: Numerical Computations of Supersonic Chemically Reacting Flows using Hydrogen-Air Combustion Models. Proc. of ISCFD, Nagoya, pp. 611-616 (1989).
28) Wada, Y., Ogawa, S., and Ishiguro, T.: "Computation of Three-Dimensional Chemically Reacting Flows around Re-Entry Vehicles". *NAL SP-10*, 1989, pp. 59-64.
29) Chapman, D.R.: Computational Aerodynamics Development and Outlook. *AIAA Journal*, **17**, pp. 1293-1313 (1979).
30) Kawamura, T., and Kuwahara, K.: Direct Simulation of a Turbulent Inner Flow by Finite Difference Method. AIAA-85-1617, 1985.
31) Ohgishi, T., Tsuge, S., and Ogawa, S.: "The Accuracy of Direct Simulation of Wall Turbulent Flow and the Numerical Prediction on the Effect of Drag Reducing Devices". *NAL SP-9*, 1988, pp. 43-48.
32) Richtmyer, R.D., and Morton, K.W.: Difference Methods for Initial-value Problems. Interscience, 1967.
33) Kawaguchi, M.: Numerical Solution of the Navier-Stokes Equations for the Flow around a Circular Cylinder at Reynolds Number 40. *J. Phys. Soc. Jpn.*, **12**, pp. 747-757 (1953).
34) Ishiguro, T.: Numerical Computations for Two-Dimensional Unsteady Nozzle Flow using Finite Difference Schemes. NAL TR-205, 1970.
35) Ishiguro. T.: Several Finite Difference Schemes and their Evaluations for the Compressible Navier-Stokes Equations. NAL TR-310, 1973.



Satoru Ogawa

Computational Sciences Div.
NATIONAL AEROSPACE LABORATORY
Bachelor of Aeronautical Eng.
The University of Tokyo 1970
Dr. of Aeronautical Eng.
The University of Tokyo 1976
Specializing in Applied Mathematics

Yoko Takakura

Sientific Systems Engineering Dept.
FUJITSU LIMITED
Bachelor of Aeronautical Eng.
The University of Tokyo 1979
Specializing in Computational Fluid Dynamics

Overseas Offices

Abu Dhabi Office
P.O. Box 47047 Suite 802,
Al Masaood Tower, Seikh Hamdan
Street, Abu Dhabi, U.A.E.
Telephone : (971-2)-333440
FAX : (971-2)-333436

Algiers Office
9, Rue Louis Rougie Chateau Neuf,
EL Biar, Alger 16030, Algeria
Telephone : (213)-2-78-5542
Telex : 408-67522

Amman Office
P.O. Box 5420, Ammân, Jordan
Telephone : (962)-6-662417
FAX : (962)-6-6673275

Bangkok Office
3rd Floor, Dusit
Thani Bldg., 1-3, Rama IV,
Bangkok, Thailand
Telephone : (66-2)-236-7930
FAX : (66-2)-238-3666

Beijing Office
Room 2101, Fortune Building
5 Dong San Huan Bei-lu,
Chao Yang District, Beijing,
People's Republic of China
Telephone : (86-1)-501-3261
FAX : (86-1)-501-3260

Bogotá Office
Cra. 13 No. 27-50,
Edificio Centro Internacional
Tequendama, Oficina 326/328
Bogotá, D.E. Colombia
Telephone : (57-1)-286-7061
FAX : (57-1)-286-7148

Brussels Office
Avenue Louise 176, Bte 2
1050 Brussels, Belgium
Telephone : (32-2)-648-7622
FAX : (32-2)-648-6876

Hawaii Branch
6660 Hawaii Kai Drive, Honolulu,
Hawaii 96825, U.S.A.
Telephone : (1-808)-395-2314
FAX : (1-808)-396-0059

Indonesia Project Office
16th Floor, Skyline Bldg.,
Jalan M.H. Thamrin No. 9,
Jakarta, Indonesia
Telephone : (62-21)-3105710
FAX : (62-21)-3105983

Jakarta Office
16th Floor, Skyline Bldg.,
Jalan M.H. Thamrin No. 9,
Jakarta, Indonesia
Telephone : (62-21)-333245
FAX : (62-21)-327904

Kuala Lumpur Office
Letter Box No. 47, 22nd Floor,
UBN Tower No. 10, Jalan P.
Ramlee, 50250, Kuala Lumpur,
Malaysia
Telephone : (60-3)-238-4870
FAX : (60-3)-238-4869

Munich Office
c/o DV18 Siemens A.G.
Otto-Hahn-Ring 6, D-8000,
München 83, F.R. Germany
Telephone : (49-89)-636-3244
FAX : (49-89)-636-45345

New Delhi Office
1st Floor, 15 Katsurba
Gandhi Marg
New Delhi-110001, India
Telephone : (91-11)-331-1311
FAX : (91-11)-332-1321

New York Office
680 Fifth Avenue, New York,
N.Y. 10019, U.S.A.
Telephone : (1-212)-265-5360
FAX : (1-212)-541-9071

Paris Office
Bâtiment Aristote,
Rue Olof Palme 94006,
Creteil Cedex, France
Telephone : (33-1)-4-399-0897
FAX : (33-1)-4-399-0700

Shanghai Office
Room 1504 Ruijin Bldg.,
205 Maoming Road South,
Shanghai, People's Republic
of China
Telephone : (86-21)-336-462
FAX : (86-21)-336-480

Taipei Office
Sunglow Bldg., 66, Sung Chiang
Road, Taipei, Taiwan
Telephone : (886-2)-551-0233
FAX : (886-2)-536-7454

Washington, D.C. Office
1776 Eye Street, N.W.,
Suite 880, Washington, D.C.,

Overseas Subsidiaries

		Telephone	Facsimile
FKL Dong-Hwa Ltd.	338-13, Daehong-Ri, Sunghwan-Eub, Chunwon-Gun, Chungnam, Republic of Korea	(82-417)-581-0701	(82-417)-581-0700
Fujian Fujitsu Communications Software Ltd.	Wuliting Fuma Road, Fuzhou, Fujian Province, People's Republic of China	(86-591)-560070	(86-591)-560022
Fujitsu America, Inc.	3055 Orchard Drive, San Jose, CA 95134-2017, USA	(1-408)-432-1300	(1-408)-432-1318
Fujitsu Australia Ltd.	475 Victoria Ave., Chatswood, NSW 2067, Australia	(61-2)-410-4555	(61-2)-411-8603
Fujitsu Australia Software Technology Pty. Ltd.	1st Floor. Techway House 18 Rodborough Road, Frenchs Forest, N.S.W. 2086, Australia	(61-2)-936-1111	(61-2)-975-2899
Fujitsu Business Communication Systems, Inc.	3190 Miraloma Ave., Anaheim, CA 92806-1906, USA	(1-714)-630-7721	(1-714)-630-7660
Fujitsu Canada, Inc.	6280 Northwest Drive, Mississauga, Ontario, Canada L4V 1J7	(1-416)-673-8666	(1-416)-673-8677
Fujitsu Computer Packaging Technologies, Inc.	3025 Orchhard Parkway San Jose, CA 95134-2017, USA	(1-408)-432-1300	
Fujitsu Component (Malaysia) Sdn. Bhd.	No. 1, Lorong Satu, Kawasan Perindustrian Parit Raja, 86400 Batu Pahat, Johor, Malaysia	(60-7)-482-111	(60-7)-481-771
Fujitsu Component of America, Inc.	3545 North First Street, San Jose, CA 95134-1804, USA	(1-408)-922-9000	(1-408)-428-0640
Fujitsu Customer Service of America, Inc.	11085 N. Torrey Pines Rd. La Jolla, CA 92037, USA	(1-619)-457-9900	(1-619)-457-9968
Fujitsu Deutschland GmbH	Frankfurter Ring 211, 8000 München 40, F.R. Germany	(49-89)-323-780	(49-89)-323-78100
Fujitsu do Brasil Ltda.	Rua Manoel da Nôbrega, 1280-2Ã Andar, C.E.P. 04001, São Paulo, Brazil	(55-11)-885-2933	(55-11)-885-9132
Fujitsu Espana, S.A.	Edificio Torre Europa 5ª, Paseo de la Castellana, 95, 28046 Madrid, Spain	(34-1)-581-8000	(34-1)-581-8300
Fujitsu Europe Ltd.	2, Longwalk Road, Stockley Park, Uxbridge, Middlesex, UB11 1AB, England	(44-81)-573-4444	(44-81)-573-2643
Fujitsu Europe Telecom R&D Centre Limited	2, Longwalk Rd., Stockley Park, Uxbridge, Middlesex, UB11 1AB, England	(44-81)-756-0286	(44-81)-573-3602
Fujitsu Finance (U.K.) PLC	2, Longwalk Road, Stockley Park, Uxbridge, Middlesex, UB11 1AB, England	(44-81)-569-1628	(44-81)-573-2643
Fujitsu France S.A.	Bâtiment Aristote 17, Rue Olof Palme 94006, Creteil Cedex, Paris, France	(33-1)-4-399-4000	(33-1)-4-399-0700
Fujitsu Hong Kong Ltd.	Room 2521, Sun Hung Kai Centre, 30 Harbour Road, Hong Kong	(852)-8915780	(852)-5721724
Fujitsu Imaging Systems of America, Inc.	3 Corporate Drive, Commerce Park, Danbury, CT 06810, USA	(1-203)-796-5400	(1-203)-796-5665
Fujitsu International Finance (Netherlands) B.V.	Officia 1, De Boelelaan 7, 1083 HJ Amsterdam, The Netherlands	(31-20)-465996	(31-20)-427675
Fujitsu Italia S.p.A.	Via Melchiorre, Gioia No.8, 20124 Milano, Italy	(39-2)-657-2741	(39-2)-657-2257
Fujitsu Korea Ltd.	9th Floor, Korean Reinsurance Bldg., 80, Susong-Dong, Chongro-Gu, Seoul Special City, Republic of Korea	(82-2)-739-3281	(82-2)-739-3294
Fujitsu Microelectronics Asia Pte. Ltd.	No.2, Second Chin Bee Road, Jurong Town, Singapore 2261, Singapore	(65)-265-6511	(65)-265-6275
Fujitsu Microelectronics, Inc.	3545 North First Street, San Jose, CA 95134-1804, USA	(1-408)-922-9000	(1-408)-432-9044
Fujitsu Microelectronics Ireland Limited	Greenhills Centre, Greenhills Road, Tallaght, Dublin 24, Ireland	(353-1)-520744	(353-1)-520539
Fujitsu Microelectronics Italia S.r.l.	Centro Direzionale, Milanofiori, Strada No.4-Palazzo A2, 20090 Assago-Milano, Italy	(39-2)-824-6170	(39-2)-824-6189
Fujitsu Microelectronics Limited	Hargrave House, Belmont Road, Maidenhead, Berkshire SL6 6NE, U.K.	(44-628)-76100	(44-628)-781484
Fujitsu Microelectronics (Malaysia) Sdn. Bhd.	Persiaran Kuala Selangor, Seksyen 26,40000 Shah Alam, Selangor Darul Ehsan, Malaysia	(60-3)-511-1155	(60-3)-511-1227
Fujitsu Microelectronics Pacific Asia Limited	Rooms 616-617, Tower B, New Mandarin Plaza, 14 Science Museum Road, Tsimshatsui East, Kowloon, Hong Kong	(852-3)-723-0393	(852-3)-721-6555
Fujitsu Mikroelectronik Gm bH	Am Siebenstein 6-10, 6072 Dreieich Buchschlag, Germany	(49-6103)-690-0	(49-6103)-690-122
Fujitsu Network Switching of America, Inc.	4403 Bland Road, Somerset Park, Raleigh, NC 27609, USA	(1-919)-790-2211	(1-919)-790-8376
Fujitsu New Zealand Ltd.	6th Floor, National Insurance House, 119-123 Featherston Street, Wellington, New Zealand	(64-4)-473-3240	(64-4)-473-3429
Fujitsu Nordic AB	Torggatan 8, S-171 54 Solna, Sweden	(46-8)-764-7690	(46-8)-28-0345
Fujitsu Philippines, Inc.	2nd Floor, United Life Bldg., Pasay Road, Legaspi Village, Makati, Metro Manila, Philippines	(63-2)-85-49-51	(63-2)-817-7576
Fujitsu (Singapore) Pte.Ltd.	75, Science Park Drive, #02-06 CINTECH II Singapore 0511. Singapore	(65)-777-6577	(65)-778-2055
Fujitsu Systems Business of America, Inc.	2986 Oakmead Village Court, Santa Clara, CA 95051, USA	(1-408)-988-8012	(1-408)-492-1982
Fujitsu Systems Business (Malaysia) Sdn. Bhd.	No. 11-4 Right Angel Building, Jalan 14/22 Section 14, Petaling Jaya, Malaysia		
Fujitsu Systems Business (Thailand) Limited	492, 494 Mini Office, Rachada Complex, Rachadapisek Road, Bangkok, Thailand	(662)-513-9148	(662)-513-9147
Fujitsu Systems of America, Inc.	12670 High Bluff Drive, San Diego, CA 92130-2103, USA	(1-619)-481-4004	(1-619)-259-2603
Fujitsu (Thailand) Co., Ltd.	60/90 (Nava Nakorn Industrial Estate Zone 3) Moo 19, Phaholyothin Road, Tambon Klongnung, Amphur Klongluang, Pathumthani 12120, Thailand	(66-2)-529-2630	(66-2)-529-2581
Fujitsu Vitória Computodores e Serviços Ltda.	Avenida Nossa, Senhora da Penha, 570-8-S/801 Praia do Canto-Vitória-Espirito Santo, Brazil	(55-27)-225-0355	(55-27)-225-0954
Information Switching Technology Pty. Ltd.	Level 32, 200 Queen Street, Melbourne 3000, Australia	(61-3)-670-4755	(61-3)-670-6183
Intellistor, Inc.	2402 Clover Basin Drive, Longmont, Colorado 80503, USA	(1-303)-682-6400	(1-303)-682-6401
Tatung-Fujitsu Co., Ltd.	5 Floor Tatung Bldg., 225, Nanking East Road 3rd Section, Taipei, Taiwan	(886-2)-713-5396	(886-2)-717-4644

(The information here as of May 31, 1991)

FUJITSU LIMITED

6-1, Marunouchi 1-chome, Chiyoda-ku, Tokyo 100, Japan
Phone: National (03) 3216-3211 International (Int'l Prefix) 81-3-3216-3211 Telex: J22833 Cable: "FUJITSULIMITED TOKYO"