IBM Enterprise System/9000 Type 9121 Model 320 air-cooled processor technology

by V. L. Gani
M. C. Graf
R. F. Rizzolo
W. F. Washburn

The basic component of the new IBM Enterprise System/9000™ Type 9121 Model 320 processor is an air-cooled thermal conduction module (TCM). The fabrication of this module required the integration of new bipolar chips, CMOS SRAM chips, and ECL and DCS logic circuitry in a TCM that could dissipate heat by means of air cooling. The method and details of this process of integration are described and discussed.

Introduction

This paper discusses how the selection of bipolar logic chips, CMOS SRAM chips, and an enhanced TCM made possible the IBM Enterprise System/9000™ Type 9121 Model 320 (9121-320) air-cooled single-frame processor. The use of DCS circuits [1, 2] in a 2600-gate array and the integration of 128Kb CMOS SRAM [3, 4] chips on the TCM are described. The enhanced TCM, with 63 layers with top-surface thin-film metallurgy for redistribution [5], contains 121 chip sites and is capable of dissipating up to 600 W using air rather than water for cooling. The processor can provide a uniprocessor with vector facility or a dyadic processor in a single frame. Its performance is equivalent to that of an Enterprise System/3090™ Model 180E processor, or four times that of an air-cooled Enterprise System/4381™ Model 91E processor.

Bipolar logic chip

The bipolar logic chip contains over 350 different circuits from two independent and interchangeable circuit families: differential current switch (DCS) and emitter-coupled logic (ECL). The chip is designed to provide flexibility in selecting circuits with various speed–power options. A large number of logic functions are provided by each circuit family, most with two to four programmable speed–power options ranging from 100-ps, 7-mW circuits for performance-critical applications to 500-ps, 1-mW circuits for power-critical applications. A family of embedded arrays provide dense, fast on-chip-addressable storage. Conversion circuits are also provided for

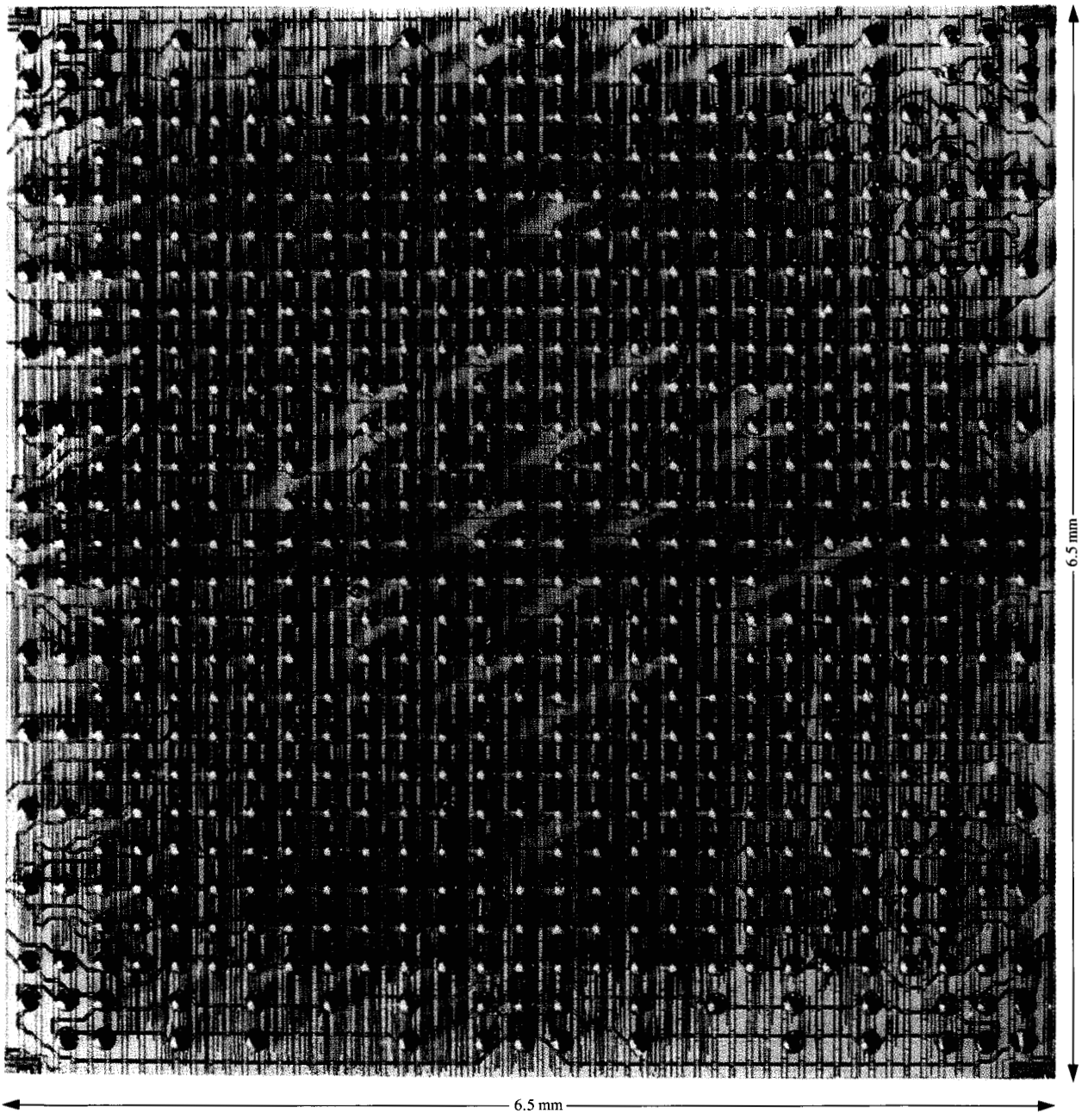

Figure 1

Bipolar logic chip front view.

communication between ECL and DCS circuits, to compensate for differing signal swings. Emitter-follower, push-pull, and bidirectional drivers are provided for off-chip communication.

The bipolar chip is designed with the IBM advanced semiconductor transistor process using 1-μm lithography technology with polysilicon and four levels of metal interconnect wiring on the chip. It measures 6.5 × 6.5 mm and contains 10 480 internal cells which can provide either 5200 ECL circuits or 2600 DCS circuits. Three power supplies are required: V_{CC} = 1.4 V, V_t = −0.7 V, and V_{EE} = −2.2 V. The reference voltage V_r is ground. There are a total of 549 chip I/O pads, of which 228 are signal I/O and 321 are power pads. **Figure 1** shows the front view of the chip with I/O pads. The bipolar logic chip design parameters are summarized in **Table 1**.

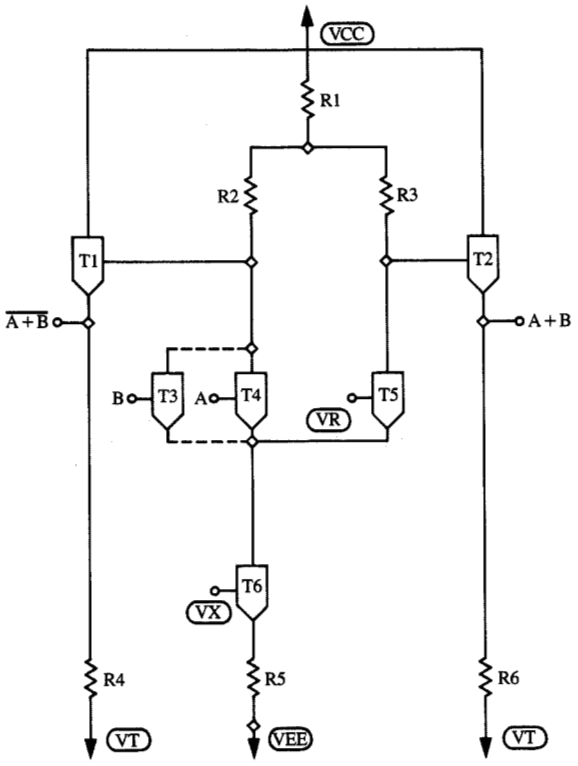

Figure 2

One-way/two-way medium-power ECL circuit.

Table 1 Design parameters of the bipolar logic chip.

Chip size:	6.5 × 6.5 mm
Process ground rule:	1 μm
Isolation:	polysilicon trench
Metallization:	four levels
NPN transistor:	Polysilicon base and emitter
Cell capacity:	5240 internal cells (ECL) Two internal cells = one DCS cell 360 I/O cells
Number of I/Os:	228 signal, 321 power
Maximum power:	10 W (on air-cooled TCM)
Power supplies:	V_{CC} (1.4 V), V_r (0.0 V), V_t (−0.7 V), V_{EE} (−2.2 V)
Circuit offerings:	Internal: emitter-coupled logic (ECL) differential current switch (DCS) embedded arrays (ECL I/O) External: Emitter-follower, Push-pull, and bidirectional drivers
Performance:	DCS 150 ps at 7.0 mW 200 ps at 3.8 mW 280 ps at 2.0 mW 440 ps at 1.3 mW

Emitter-coupled logic (ECL) has been used in the IBM mainframes in the past, starting with System/370™, System/3090™, and System/390™ water-cooled models. ECL technology provides speed but dissipates high power, requiring water cooling. The basic ECL circuit is shown in **Figure 2**.

A major consideration in an air-cooled system is the increased power dissipated by the logic circuit as the circuit density increases. The air-cooled models predominantly use DCS circuits because their speed/power ratio is superior to that of ECL. The basic DCS low-power internal two-way selector circuit is shown in **Figure 3**. The two circuit families are compared and the advantages of DCS over ECL for certain applications are described in [2]. DCS is a form of cascode logic using vertically stacked or cascoded current switches to provide more logic function per current source. DCS, however, can use the same power supply voltages as ECL. DCS circuits satisfy the low power dissipation requirements for the air-cooled mainframe and can be powered in a manner that is compatible with other ES/9000™ and ES/3090™ processors.

Multiple internal cells are used to implement the basic DCS circuit, which is a two-way selector (Figure 3). Slight modifications of this circuit produce a two-way XOR, a four-way select, a two-way AND, a two-way OR, or a two-port latch [2]. Other logic functions are obtained by the use of a powerful generalized dotting capability. This enables an AND, OR, XOR, or SELECT between any of the basic DCS functions without adding additional logic stages. The speed–power improvement results mainly from the reduced signal voltage levels made possible by the differential circuit design. However, the DCS circuit family requires two wires for every internal signal and needs more internal resistors and transistors than ECL to implement the logic. The DCS circuit provides 1.8 times more logic function than the ECL circuit.

128Kb CMOS SRAM chip

The integration of CMOS chips with bipolar chips in the TCM increases its circuit density while reducing its power dissipation. This is the first time that CMOS technology has been used in an IBM central processor. It was designed to interface with ECL signal levels, which required special CMOS drivers and receivers, and its footprint matches the common footprint of the TCM. The CMOS chip design details are described in [4].

To achieve the objectives of low power and high density at the system level, a single 128Kb SRAM with custom features was used in each of seven different applications: 1) writable control store, 2) L1 memory cache, 3) I/O processor cache, 4) store protect, 5) local work store, 6) vector array, and 7) memory configuration array. The custom features required to accommodate these applications are shown in **Figure 4**. The inputs are

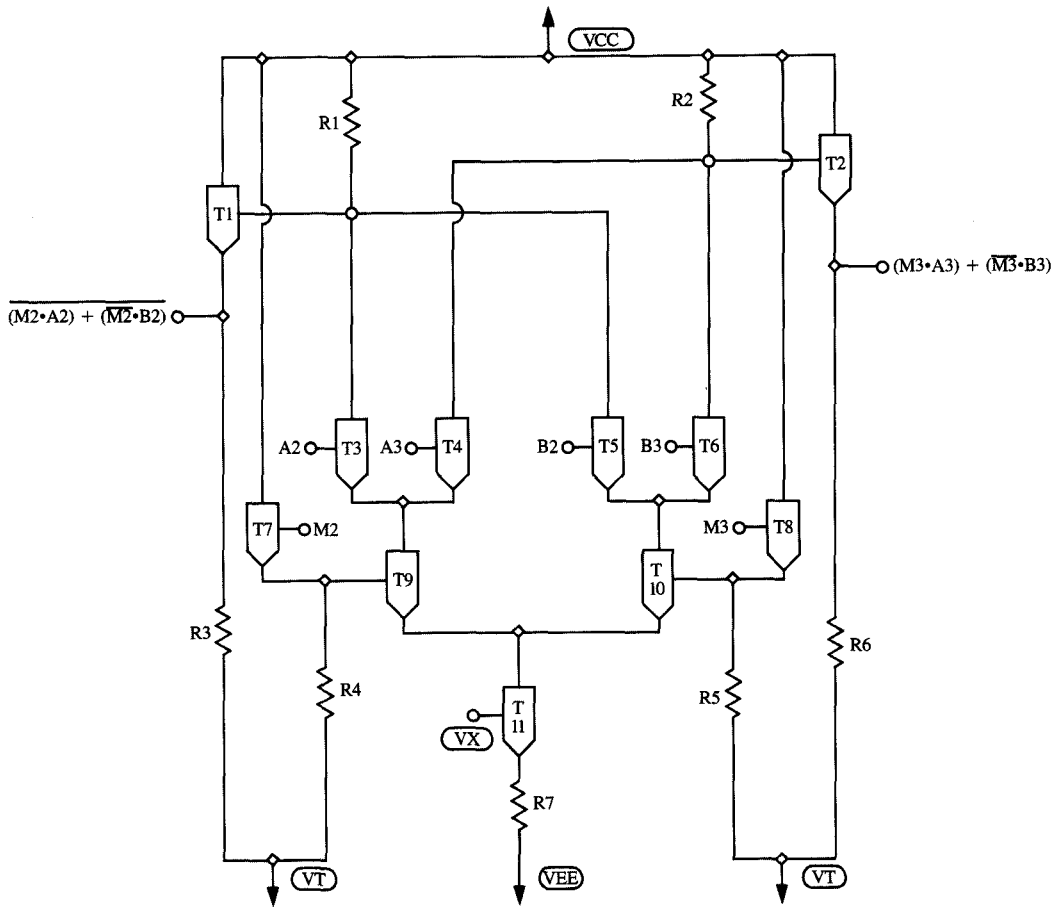

Figure 3

DCS low-power two-way selector internal circuit.

configured to permit operation in either an 8-bit or 32-bit mode. In addition, there are write controls which provide masking of any combination of single bits in 8-bit mode and any combination of nibbles (4 bits) in 32-bit mode. The chip outputs support either an 8-bit or 32-bit output mode and off-chip driver gating. Input and output modes are independently selectable.

The SRAM design features pipelining in both the address decoders and the output circuits, allowing the chip to be operated at a faster-than-access cycle time. The SRAM access time of less than 9 ns is essential to support a 15-ns system cycle time. The access time is important because the SRAM in its cache and control store applications must provide a single-cycle access. To achieve the required access time for all of these applications, innovation was required both in the physical design of the chip and in the circuit design and implementation. The chip is organized into four quadrants. One of eight subarrays in each quadrant is selected at a time, regardless of the read or write mode being used. Each subarray is designed with all of the flexibility required to read and write in either a 2-bit or 8-bit format and to provide a write mask of either one of two bits or one of two nibbles.

Integrating this chip within the TCM required a unique chip burn-in process called R3 (reduced radius removal) [4] to meet system reliability objectives. The R3 process screens the chips for defects before they are mounted on the TCM.

The CMOS SRAM chip measures 8.8 × 6.5 mm and has the same footprint as the bipolar chip. It has a total of 321 chip pads, of which 136 are used for signal I/O and 176 for power pads. The chip is designed with 1.2-μm N-well CMOS with 14-nm gate oxide, 0.5-μm effective channel length, and three levels of metal interconnection on the

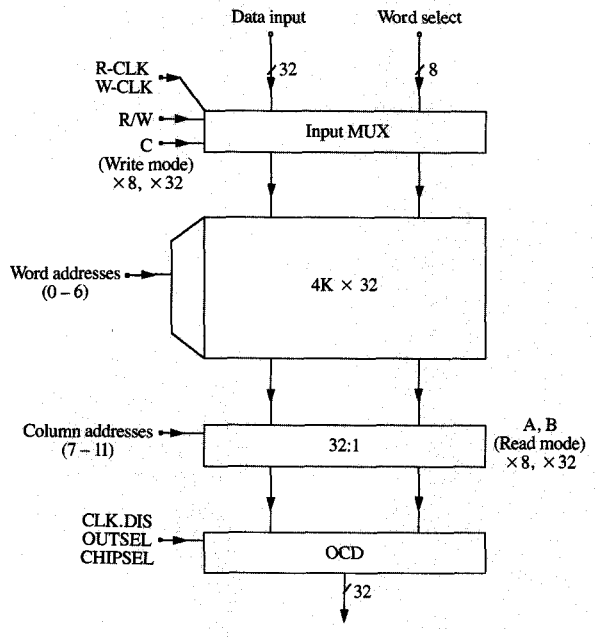

Figure 4

128Kb CMOS SRAM organization.

chip. The chip dissipates only 2 W maximum power, providing a worst-case access time of less than 9 ns. **Figure 5** is the front view of the chip, showing four quadrants with chip I/O pads and power distribution metal.

TCM package

The TCM package is an enhanced multilayer ceramic substrate [5] in which both bipolar chips and CMOS SRAM chips are placed on a common chip site (**Figure 6**). It is an air-cooled package capable of cooling a total of 600 W power dissipation at a maximum power of 10 W per chip. The air-cooled package does not require a separate cooling distribution unit as do water-cooled systems such as the 3090™ processors.

The TCM package required several technical innovations to meet its design objectives:

1. A "thin-film" copper wire layer on the top of the substrate distributes chip C4 (controlled collapse chip connection) pads. This helps optimize electrical signal characteristics, improve reliability, and reduce required wiring layers.

Table 2 Design parameters of the thermal conduction module.

Substrate size (mm)	127.5 × 127.5
Substrate material	Alumina/Mo
Number of chips	121
Chip size or edge (mm)	6.5
Chip pitch (mm)	9.9
Chip signal I/O	228
Cooling	Air
Chip power (W)	10
Module power (W)	600
Module I/O	2772
Module thickness (mm)	11.1
No. of layers	63
EC pads per chip site	152
EC wires	BECW*/Discrete
Top surface metallurgy	Thin-film Cu
Decoupling caps	144
Chip technology	Bipolar/CMOS
Piston material	Cu
Cooling media	Oil
Density (logic circuit count)	150–200K

* Buried EC wiring

2. Standard EC (engineering change) wires are "built in" to the ceramic. This allows faster response to design changes and better electrical signal attributes.
3. Decoupling capacitors are placed at each corner of the chip site (Figure 6) to reduce package switching noise.
4. Copper pistons, instead of the aluminum used in previous systems, provide improved cooling.
5. The thermal conduction region between the chip and piston is filled with oil rather than helium.

The TCM package description is summarized in **Table 2**. The multilayer ceramic substrate is made up of 63 layers with the thin-film redistribution circuitry on the top surface (**Figure 7**). Each TCM can contain 121 chips. Approximately 360 meters of interconnecting wire is used in this module. The TCM assembly cross section views (**Figures 8** and **9**) indicate the air flow on the attached heat sink. The copper pistons are immersed in oil. The heat generated by the chip is transferred through the piston to the heat sink, and from there it is forced outside the frame by the blowers located on top of the frame.

System elements

The control elements of the 9121-320 processor [6] comprise the following:

- CPE (central processor element).
- BCE (buffer control element).
- SCE (system control element).
- CCE (channel control element).
- VCE (vector control element).

In the ES/3090 processor, the CPE and BCE functions together required nine TCMs, the SCE and CCE functions

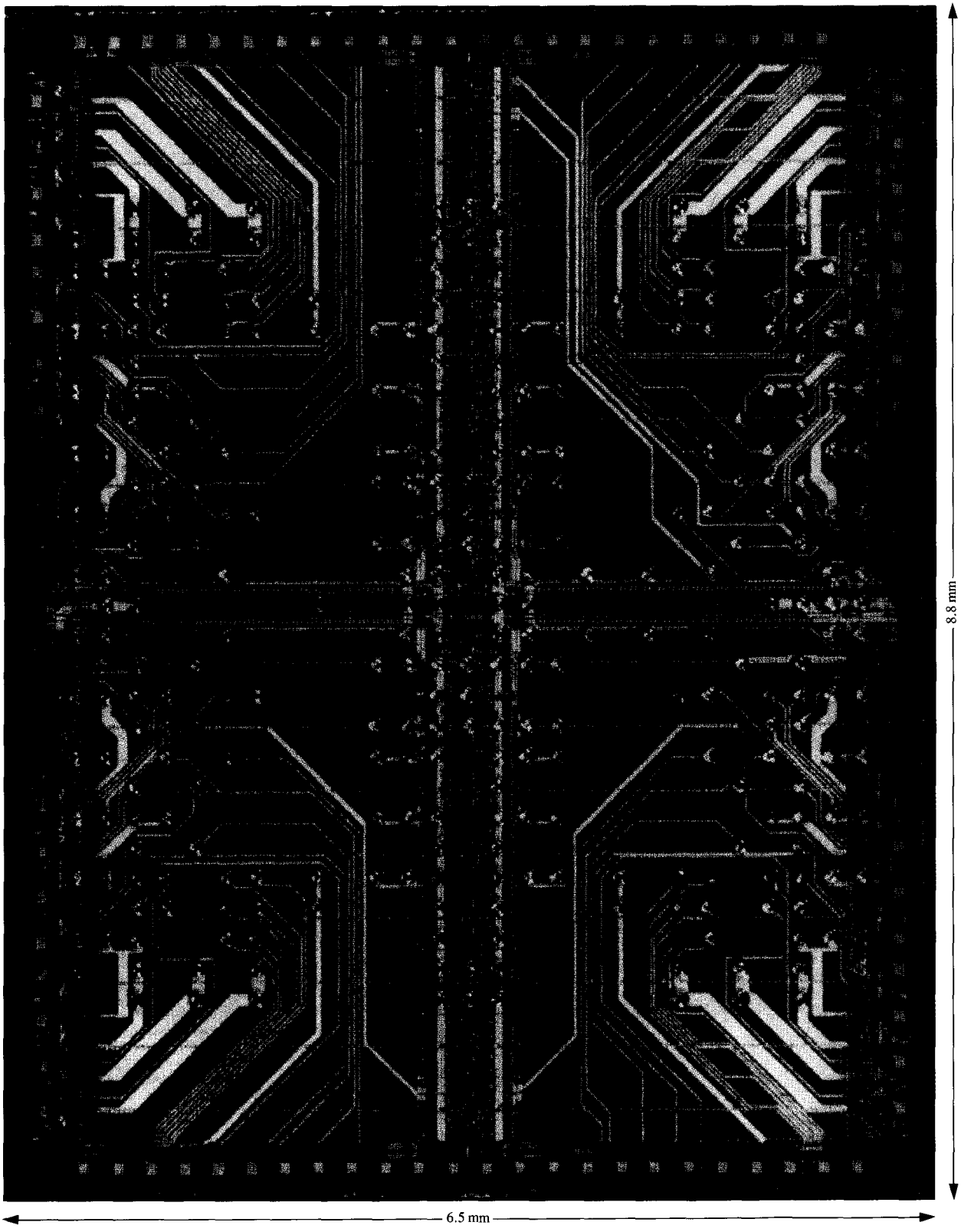

Figure 5

128Kb CMOS SRAM chip front view.

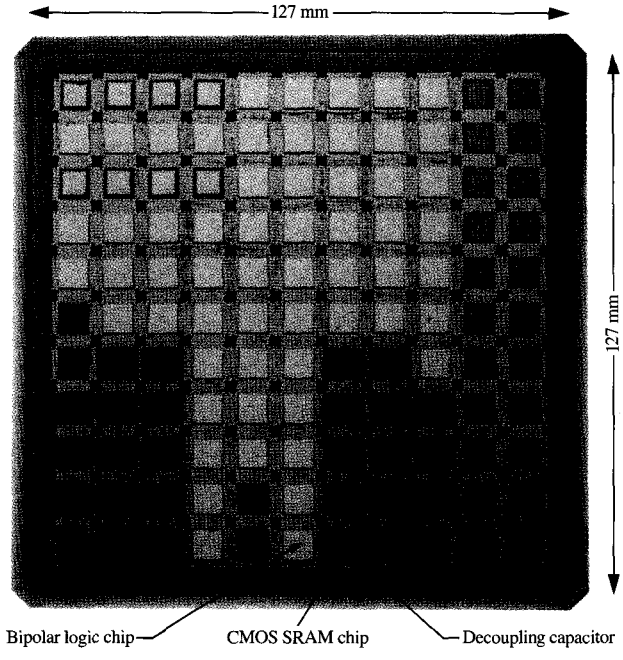

Figure 6

TCM package with CMOS SRAM chips and bipolar logic chips.

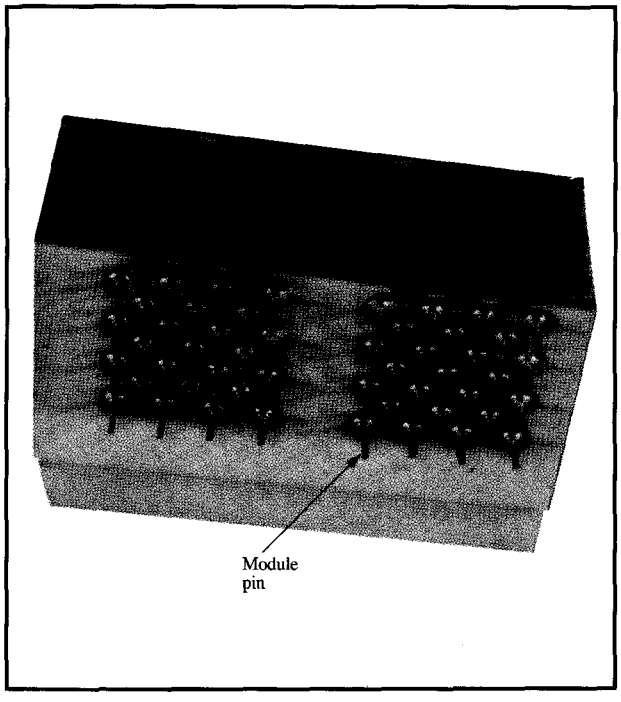

Figure 7

TCM substrate cross section.

each required six, and the VCE function three. Each of these elements is mapped into a single, separate TCM in the 9121 processor [7], reducing the number of TCMs required by a factor of almost 5.

9121-320 Mainframe

A mainframe holds three multilayer boards, each containing two TCMs (**Figure 10**). The physical size of these boards is one third that of the boards used in ES/3090 systems [8]. The tri-lead signal cables are used for off-board communication. The mainframe (**Figure 11**) shows all the components packaged. This frame contains memory cards, both parallel and serial channels to support ESCON connectivity using fiber-optic cables, and ac power distribution. The power supplies, which are located behind the board, are also air-cooled by the blowers located on top of the frame. Figure 11 shows a uniprocessor with one vector facility. The mainframe will support a dyadic processor without a vector facility. The dyadic processor with vector facility requires an additional frame to hold the vector TCM.

Summary

The high density and low power of the CMOS SRAM and the high logic function and low power of the bipolar DCS internal logic circuits made it possible to design an air-cooled 9121-320 processor. Technology innovations and package integration provided reductions in space, power, and weight with performance comparable to that of the much larger ES/3090 Model 180E, and four times the performance of the air-cooled IBM ES/4381™ Model 91E processor.

Acknowledgments

A development effort of this magnitude involves the cooperation of several sites within IBM: The bipolar chips and TCM modules were developed and built in East Fishkill, NY, the CMOS chips in Burlington, VT, and the TCM boards in Endicott, NY. The system product development and testing was done in the Mid-Hudson Valley (Poughkeepsie and Kingston, NY). The efforts of the people at all of these sites are acknowledged.

References

1. E. B. Eichelberger, S. E. Bello, R. O. Bergenn, W. M. Chu, J. A. Ludwig, and R. F. Rizzolo, "Two-Level Differential Cascode Current Switch Masterslice," U.S. Patent 4,760,289, July 26, 1988.
2. E. B. Eichelberger and S. E. Bello, "Differential Current Switch—High Performance at Low Power," *IBM J. Res. Develop.* **35**, 313–320 (1991, this issue).

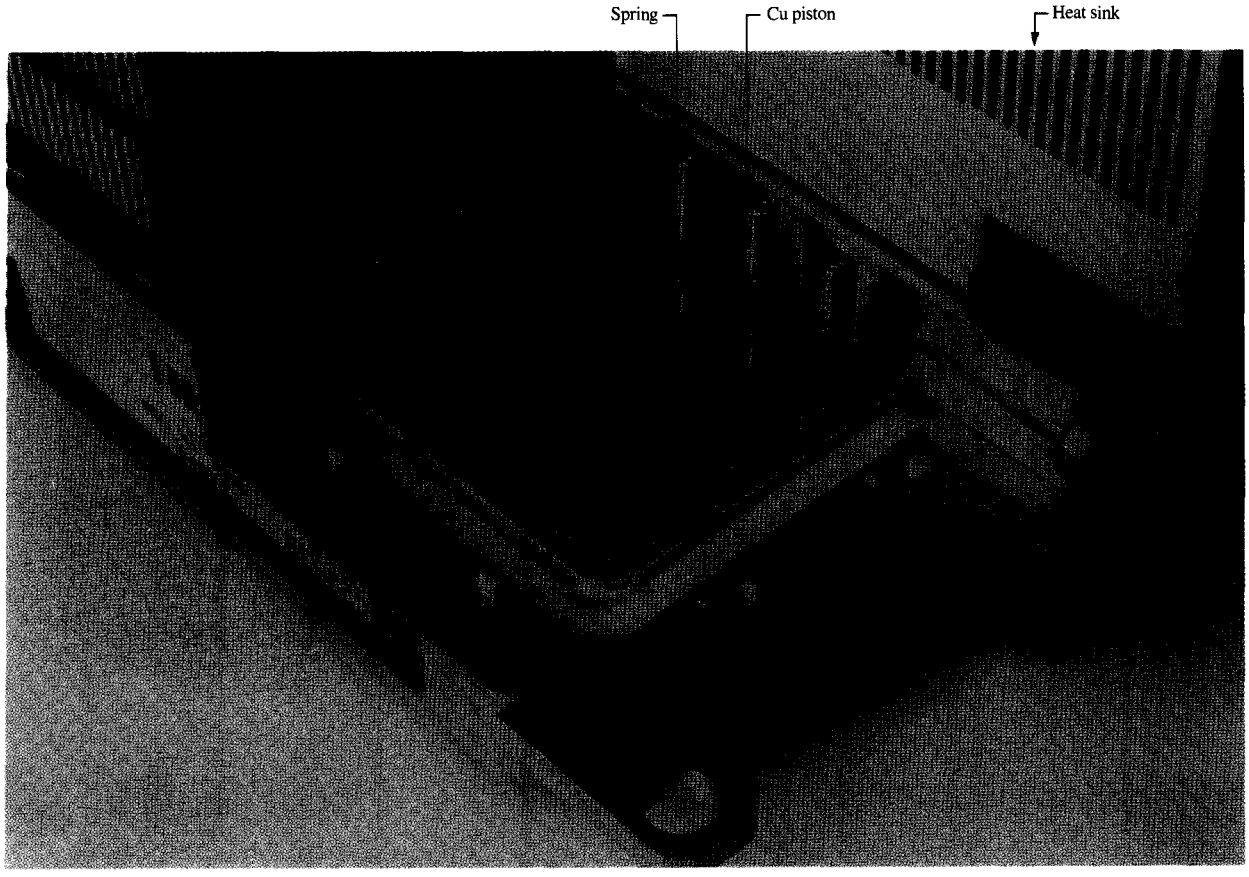

Figure 8

TCM assembly cross section.

3. F. Towler, J. Chu, R. Houghton, P. Lane, B. A. Chappell, T. I. Chappell, and S. E. Schuster, "A 128k 6.5ns Access/5ns Cycle CMOS ECL Static RAM," *Digest of Technical Papers*, International Solid State Circuits Conference, February 1989, pp. 30–31.
4. J. L. Chu, H. R. Torabi, and F. J. Towler, "A 128Kb CMOS Static Random-Access Memory," *IBM J. Res. Develop.* **35**, 321–329 (1991, this issue).
5. J. U. Knickerbocker, G. B. Leung, W. R. Miller, S. P. Young, S. A. Sands, and R. F. Indyk, "IBM System/390 Air-Cooled Alumina Thermal Conduction Module," *IBM J. Res. Develop.* **35**, 330–341 (1991, this issue).
6. S. F. Hajek, " The IBM Enterprise System/9000 Type 9121 Air-Cooled Processor," *IBM J. Res. Develop.* **35**, 307–312 (1991, this issue).
7. T. J. Slegel and R. J. Veracca, "Design and Performance of the IBM Enterprise System/9000 Type 9121 Vector Facility," *IBM J. Res. Develop.* **35**, 367–381 (1991, this issue).
8. Donald P. Seraphim, "A New Set of Printed-Circuit Technologies for the IBM 3081 Processor Unit," *IBM J. Res. Develop.* **26**, 37–44 (1982).

Received October 5, 1990

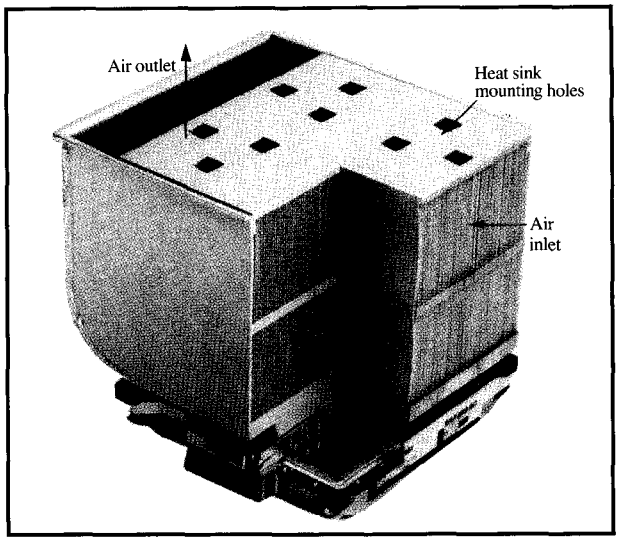

Figure 9

TCM assembly cross section with heat sink.

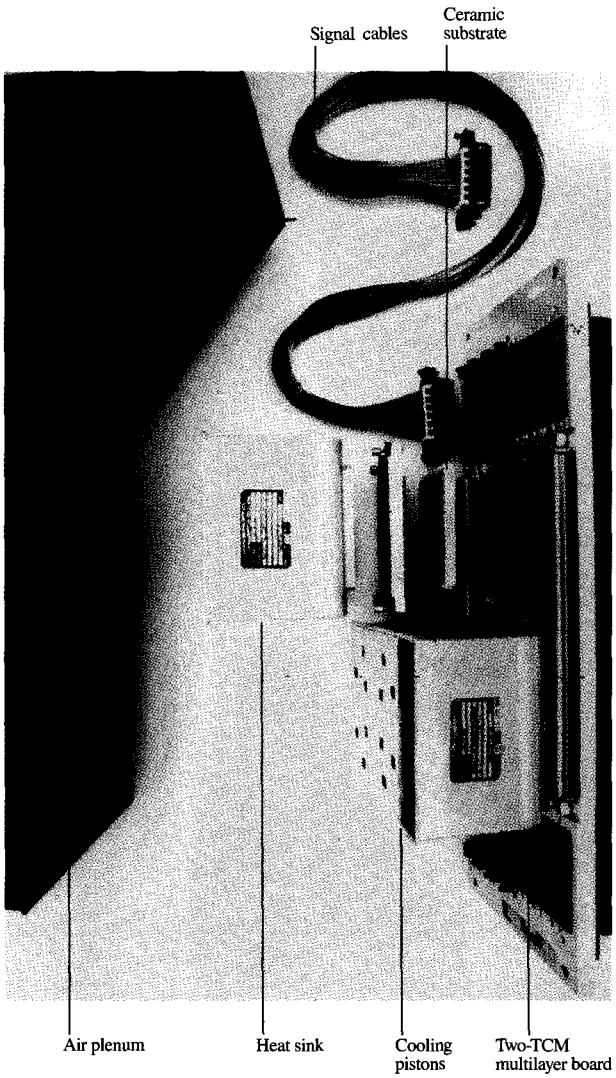

Figure 10

TCM substrate, heat sink, TCM board, and plenum assembly.

Matthew C. Graf *IBM Data Systems Division, P.O. Box 950, Poughkeepsie, New York 12602.* Mr. Graf received his B.S.E.E. degree from the University of Maine at Orono before joining IBM in 1978. He received the M.S.C.E. degree from Syracuse University in 1982. Mr. Graf managed the DCS circuit design group under IBM Fellow E. B. Eichelberger from 1985 to 1987. He assumed project management responsibility for the technology used in the ES/9000 high-end air-cooled processor and the package qualification for the ESCON™ channel subsystem. Mr. Graf is currently managing a technology project for future high-end processors in Poughkeepsie. He is a member of the Institute of Electrical and Electronics Engineers.

Richard F. Rizzolo *IBM Data Systems Division, P.O. Box 950, Poughkeepsie, New York 12602.* Mr. Rizzolo received his B.S. in physics in 1977 and his M.S.E.E. from Rensselaer Polytechnic Institute, joining IBM in 1978 in the Functional Products Advanced Technology group. In 1982 he moved to Advanced VLSI Technology and Testing, where his responsibilities included DCS circuit design and the evaluation and development of DCS testing techniques. Mr. Rizzolo is a co-inventor of a United States patent applicable to DCS circuitry (Reference 1). In 1987 he transferred to a system technology group involved in the development of the ES/9000 air-cooled processors, where he was technical co-leader with V. Gani. His responsibilities included training in DCS circuit usage, technology interface to East Fishkill, and critical path delay reduction. Mr. Rizzolo is currently working in Poughkeepsie on a follow-on to the water-cooled versions of the ES/9000 technology.

William F. Washburn *IBM Data Systems Division, Neighborhood Road, Kingston, New York 12401.* Mr. Washburn received his B.S.E.E. from Newark College of Engineering in 1965 and his M.S.E.E. from Syracuse University in 1977. He joined IBM in 1965 at the development laboratory in Poughkeepsie. Mr. Washburn has worked on integrated circuit development and system technology support for display systems and large systems in the Poughkeepsie and Kingston laboratories. He is currently managing the Advanced Technology group in Kingston. Mr. Washburn is a member of the Institute of Electrical and Electronics Engineers.

Venkappa L. Gani *IBM Data Systems Division, P.O. Box 950, Poughkeepsie, New York 12602.* Mr. Gani received his B.S.E.E. from Karnatak University, Dharwar, India, and his M.S.E.E. from the University of Colorado. He joined the IBM Components Division of East Fishkill in 1969, working in circuit and chip technology for the 3081 system. Mr. Gani managed the 3080 and 3090 mainframe system technology group in Poughkeepsie from 1981 to 1985. He was a technical leader in the system technology group that assisted in the development of the ES/9000 air-cooled processors in Kingston from 1987 to 1990. He is currently involved in the development of future high-end processors in Poughkeepsie. Mr. Gani is a senior member of the Institute of Electrical and Electronics Engineers and a member of the Computer Society.

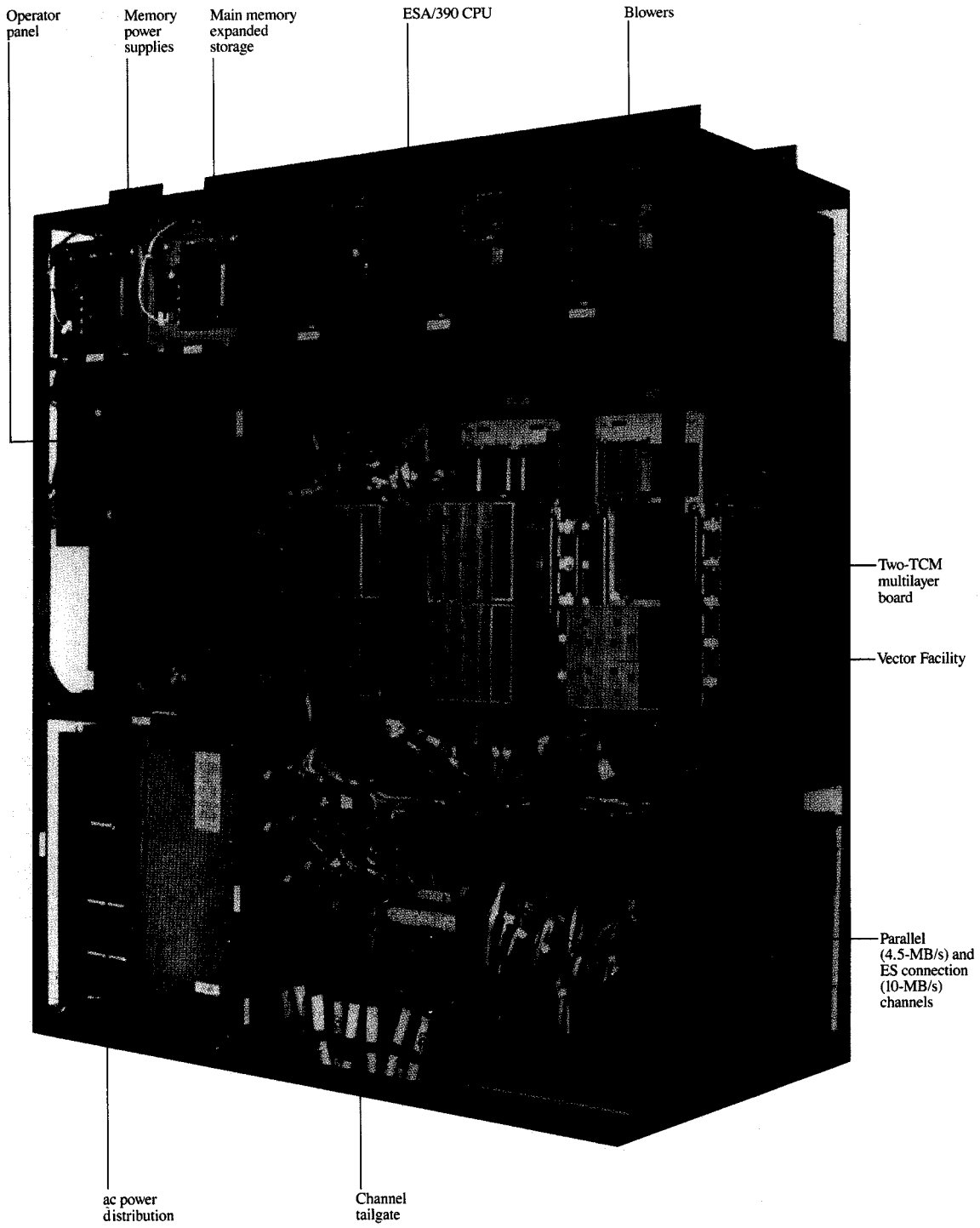

Figure 11

ES/9000 Type 9121 mainframe open view.