The IBM Enterprise System/9000 Type 9121 air-cooled processor

by S. F. Hajek

The IBM Enterprise System/9000™ Type 9121 air-cooled processor achieves, with the same or reduced physical floor space and power levels, a performance level equal to or greater than those of previous IBM processors. This performance level was attained by a combination of design innovations: a new air-cooled thermal conduction module (TCM), integration of bipolar and CMOS technology in this TCM, the design and implementation of a differential current switch bipolar circuit family, integrated programmable memory subsystem design, and extensive use of VLSI technology. The result of these innovations was a 15-ns-cycle air-cooled machine. The salient features and an overview of the machine design are presented.

Introduction

Three different types of processors are used in the new IBM Enterprise System/9000™ (ES/9000™) family: water-cooled processors (Type 9021), air-cooled processors (Type 9121), and an air-cooled rack processor (Type 9221). The Type 9121 processors use a thermal conduction module (TCM) [1] which is air-cooled. This radical departure from previous TCMs required extensive redesign of the internal and external TCM cooling system, TCM substrate, substrate wiring, and top-surface distribution wiring; new chip logic circuitry; and the use of bipolar and CMOS chips in the same TCM. In addition, new ancillary components such as system logic and air cooling were designed to meet the enhanced performance, space, and power-consumption specifications. This issue of the *IBM Journal of Research and Development* describes the design innovations embodied in the ES/9000 Type 9121 air-cooled processor.

Design advances

The following design improvements have been realized in the ES/9000 Type 9121 processor:

- A new differential current switch (DCS) bipolar circuit family has been used to obtain the circuit density and power/speed performance required for an air-cooled TCM design. High speed and low power are achieved through reduced signal swing and an active rising-signal transition.

- An air-cooled TCM has been developed which uses a "thin-film" copper wire layer for redistribution, 63 wiring layers, spring-loaded copper pistons, and a finned heat sink to cool up to 121 logic chips which generate up to 10 W individually and 600 W collectively [2].
- A new, high-speed CMOS SRAM meets the power and speed requirements for the internal arrays of the processor. A 10-ns-access, 128Kb chip design has been achieved [3].
- CMOS, ECL bipolar, and DCS bipolar technologies have been integrated within a single air-cooled TCM, and a high-power ECL-compatible CMOS driver has been developed [3] to obtain a 15-ns processor cycle time. "Cycle-stealing" design techniques are used to compensate for the slower CMOS SRAM arrays [4]. This constitutes the first use of CMOS technology within a TCM in any machine design.
- Central and expanded storage have been integrated into a single physical storage. Advanced design techniques such as single-level buffers have been implemented to achieve the performance levels of independent storages [5]. A programmable design technique has also been introduced which allows the user to dynamically configure the storages at initial program load (IPL).
- A vector facility has been designed for the air-cooled processor [6].
- VLSI/CMOS technology is used extensively in the design of the I/O channels.
- A single CMOS multichip module (MCM) implements either the parallel original equipment manufacture interface (OEMI) or the fiber-optic-based Enterprise Systems Connection Architecture™ (ESCON™) serial channel. The common design of these new channels has allowed the 9121 processors to offer channel subsystem mixtures of both channel types within a single frame.

As a result of these design improvements, a six-TCM dyadic processor achieves a performance level more than four times that of previous air-cooled frame designs such as the IBM ES/4381™ Model 92E. The extensive use of VLSI technology provides as many as 28 channels and 1 GB of storage within a single frame. (A frame is the structure which contains the essential elements of the system.)

Figure 1 is a logical view of the system. The channel control element (CCE) TCM controls the channel subsystem. The channel subsystem comprises the CCE TCM, the input/output processor (IOP) contained within the CCE TCM, the channel adapter cards, the parallel OEMI channels, the ESCON serial channels, and the associated cabling. A circuit board can contain either 12 parallel OEMI and four ESCON serial channels, or 16 ESCON serial channels. The base frame can contain up to two channel circuit boards, resulting in 32 channels. An additional 16 channels on two channel circuit boards can be contained in an expansion frame, giving the user a total of 48 channels. The channel adapter card converts the common parallel OEMI and ESCON serial channel interface to that of the CCE TCM interface. Individual I/O buffers are provided for each attached channel. The ESCON serial channel utilizes high-speed fiber-optic technology to increase distances, provide higher data rates, and reduce cable bulk to the processor-attached I/O control units and devices. ESCON serial channels support speeds of 10 MB/s as compared to the 4.5 MB/s of parallel channels.

The system control element (SCE) TCM controls the flow of data among the central storage, the expanded storage, the processors, and the channel subsystem. This TCM also controls the central and expanded storage configurations of the single physical processor storage. Configuration parameters are selected at power-on reset. Storage protection keys are maintained within this TCM.

The buffer control element (BCE) TCM contains either a 64KB or a 128KB high-speed buffer. The high-speed buffer data array is composed of SRAM CMOS 128Kb chips. The BCE is organized with an eight-double-word line size and four-way set associativity. The BCE utilizes a byte error-correcting algorithm (ECC) complemented by a line-delete algorithm. No single failure point can result in more than one bit error in any given ECC word.

The central processor element (CPE) TCM executes all instructions defined by the ESA/390™ architecture. The Processor Resource/Systems Manager™ (PR/SM™) [7] function within the processor allows several independent logical processors to share the physical processor. System/370™ operating systems are supported under one or more partitions of the PR/SM facility. The internal cycle time is 15 ns for all models. The internal performance level of the processor is determined by the size of the high-speed buffer contained within the BCE TCM and other design parameters.

The vector control element (VCE) TCM is an optional feature which executes 171 vector instructions and contains the vector registers. Vector instructions significantly improve numerically intensive applications in that the same instruction operates on multiple sets of data. In contrast, scalar operations require a set of instructions for each set of data. The 16 vector registers which together form the vector array are each 256 elements deep and 32 bits wide [6].

A single storage configuration contains both central and expanded storage [5]. A total of 1024 MB of storage can be contained on one storage circuit board. Up to eight storage array cards of varying sizes are used to form various storage configuration sizes, up to the maximum of 1024 MB. The storage is logically configured by programmable means into central and expanded storage at power-on

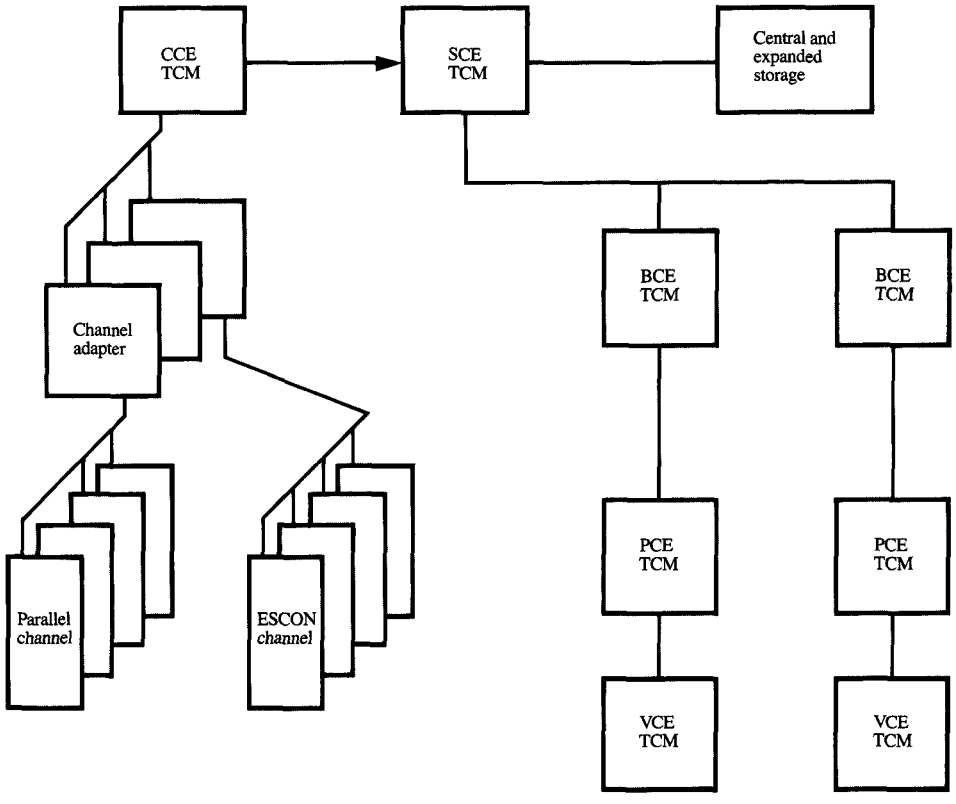

Figure 1

ES/9000 air-cooled frame processor: logical view.

reset. A processor can have up to 256 MB of central storage. Either 1Mb or 4Mb storage chips are used for the physical storage. Error-correcting code bits are stored with the data.

The ES/9000 Type 9121 air-cooled processor is packaged within two frames, a base frame and an optional expansion frame. The 9121 base frame accommodates the following:

- 1024MB central and expanded storage and SCE and CCE TCM subsystem.
- Two processor (CPE/BCE TCM) subsystems without VCE TCM, or one processor subsystem with VCE TCM.
- Up to 32 parallel and/or ESCON channels.
- An inboard service processor.
- A power/thermal subsystem.

The expansion frame, which measures approximately 2/3 the length of the base frame, accommodates VCE TCMs for the dyadic processor and additional channels to a maximum of 48 channels per system.

Figure 2 provides a floor-space comparison of the ES/3090™ [8] and ES/4381 processors and the ES/9000 Type 9121 air-cooled processor.

Figure 3 shows the location of the six TCMs of the dyadic processor within the base frame. Below the TCMs the channels are located on two channel circuit boards with the I/O panel adjacent. To the left of the TCMs is the operator panel, which is accessible when the covers are in place. Located below the operator panel is the processor storage, which is packaged on a single board. Located across the top of the frame are power supplies and blowers used for cooling the components.

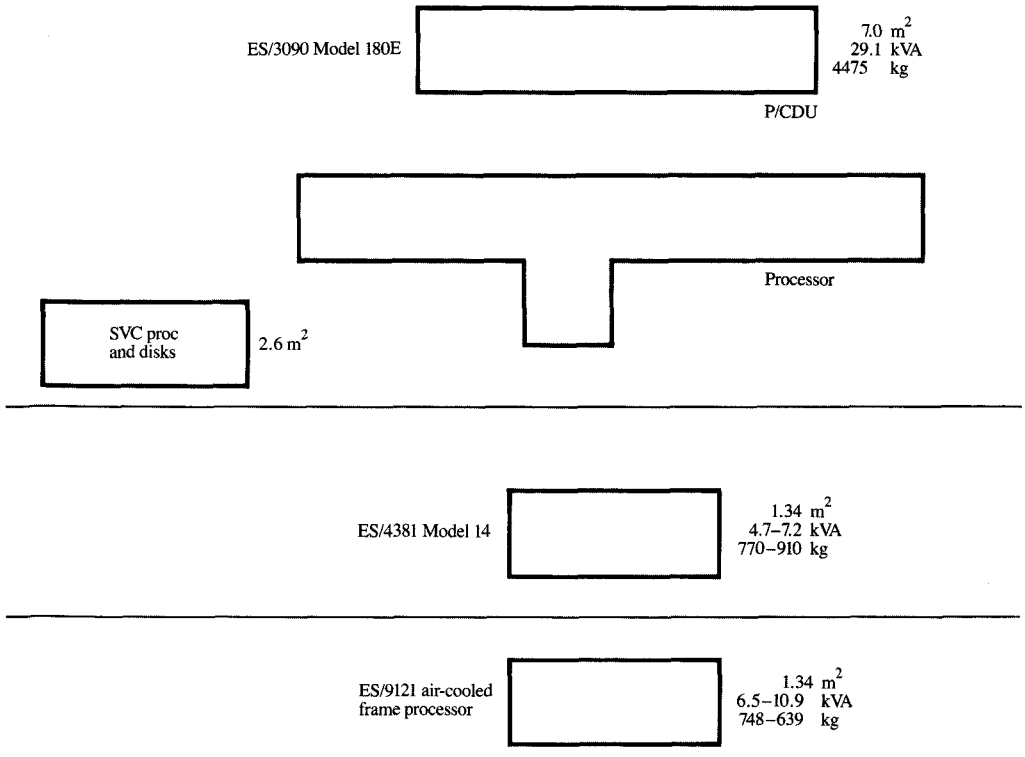

Figure 2

ES/9000 air-cooled frame processor: floor-space comparison.

Summary

Reliability, availability, and serviceability (RAS) targets were achieved in the design of the ES/9000 Type 9121 processor. Increased circuit densities have greatly reduced the number of external interconnections, a prime source of reliability failures. The machine design has also benefited from the reduced intrinsic failure rate of the base technologies. Machine design has complemented these technology improvements with extensive error-correcting design techniques throughout the design. The high-speed buffer array, for example, uses an error-correcting code on each byte of stored data. Other error-correcting techniques are used on the other data arrays contained within the processor complex. Error-recovery designs are also implemented throughout the processor.

An inboard processor controller contained within the base frame provides recovery and support services to the machine. An outboard PS/2®-based system is utilized as an external support facility. All necessary machine functions are executed by the inboard processor. In fact, the PS/2 can be disconnected without disrupting the operation of the machine.

Many advances were made in designing the power thermal subsystem for this air-cooled machine. In contrast to the IBM 3090™ processors [8], a motor-generator set is not required. Three-phase power of either 50 or 60 Hz is converted into various regulated dc voltages required by the logic technologies. A maximum of 11 kVA is required for a fully configured dyadic processor. The thermal system utilizes forced convection cooling technology to provide the appropriate cooling for all of the logic technologies.

Advances in semiconductor technology played a key role in achieving success. A differential current switch circuit family was developed which made it possible to meet the power/speed requirements of a 15-ns machine design point. This circuit technology is supported by advanced physical design, circuit design, and chip design

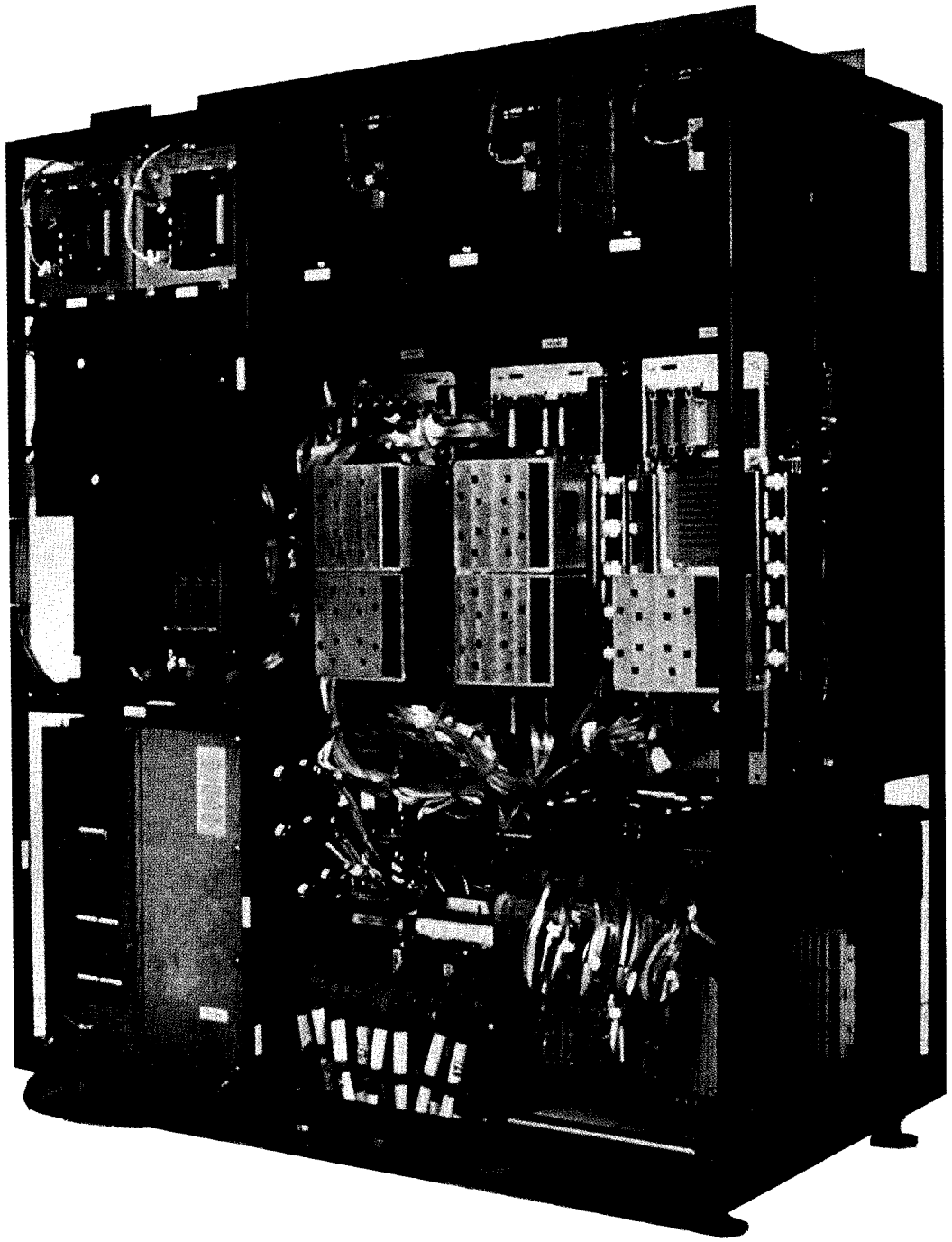

Figure 3

ES/9000 air-cooled frame processor: base frame front view.

practices. Corresponding advances in TCM technology were required to support increased chip densities and cooling requirements.

Advances in logic design verification and manufacturing test verification were also required to make this machine a reality. The Engineering Verification Engine (EVE) [9] was heavily utilized in verifying the logic design and contributed significantly to the reduction of the development cycle for this product. EVE allowed the design engineers to run the System Architectural Kernel (SAK) program before releasing the hardware design. Successful execution of the SAK program on the test floor was an announcement criterion. Self-test techniques were used in the TCM design [10] to satisfy manufacturing requirements.

The Enterprise System/9000 Type 9121 processor has advanced the state of the art in air-cooled mainframe processors. A performance level four times greater than those of previous processors was achieved within the same volumetric and power characteristics. For instance, the IBM 3090 System Model S, a water-cooled mainframe, required 36 TCMs to achieve the 9121 performance level. Bipolar and CMOS technology were integrated within a TCM for the first time. Three different circuit families, CMOS, ECL bipolar, and DCS bipolar, were utilized to meet the performance and power requirements to achieve a 15-ns machine cycle design point.

Appendix: Glossary

ES/9000 Type 9121 air-cooled processor One of three mainframe types in the ES/9000 family of processors announced by IBM on September 5, 1991. The Type 9121 processors incorporate air-cooled TCMs to implement the design point.

ESA/390 The system architecture announced by IBM on September 5, 1991. All hardware and software elements of the ES/9000 Type 9121 processors support this architecture. System/370 architecture is supported as a subset.

Differential current switch (DCS) A bipolar circuit design technology which uses differential circuit design techniques to achieve the desired density, power, and performance.

Emitter-coupled logic (ECL) A bipolar circuit design technology which uses emitter-coupled circuit design techniques to achieve high performance.

Acknowledgments

It is impossible to acknowledge individually all of the engineers responsible for the design of this machine. However, I would like to acknowledge William Spraul, engineering design manager, and his talented team of engineers and managers who made this machine a reality. I would also like to acknowledge Columbo Dalmas and Paul Simoneau for their tireless efforts in performing the verification and simulation efforts for this machine. Virgil McIntosh and his physical design team spent countless hours performing the physical design of the chips, TCMs, boards, and cards for this machine.

References

1. A. J. Blodgett and D. R. Barbour, "Thermal Conduction Module: A High-Performance Multilayer Ceramic Package," *IBM J. Res. Develop.* **26,** 30–36 (1982).
2. V. L. Gani, M. C. Graf, R. F. Rizzolo, and W. F. Washburn, "IBM Enterprise System/9000 Type 9121 Model 320 Air-Cooled Processor Technology," *IBM J. Res. Develop.* **35,** 342–351 (1991, this issue).
3. J. L. Chu, H. R. Torabi, and F. J. Towler, "A 128Kb Static Random-Access Memory," *IBM J. Res. Develop.* **35,** 321–329 (1991, this issue).
4. P. R. Turgeon, A. R. Steel, and M. R. Charlebois, "Two Approaches to Array Fault Tolerance in the IBM Enterprise System/9000 Type 9121 Processor," *IBM J. Res. Develop.* **35,** 382–389 (1991, this issue).
5. B. W. Curran and M. H. Walz, "IBM Enterprise System/9000 Type 9121 System Controller and Memory Subsystem Design," *IBM J. Res. Develop.* **35,** 357–366 (1991, this issue).
6. 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).
7. *ES/3090 Model S Processor Complex PR/SM Feature*, Order No. GA24-3416-00; available through IBM branch offices.
8. S. G. Tucker, "The IBM 3090 System: An Overview," *IBM Syst. J.* **25,** 4–19 (1986).
9. D. K. Beece, G. Deiberg, G. Papp, and F. Villante, "IBM Engineering Verification Vehicle," *Proceedings of the IEEE 25th ACM/IEEE Design Automation Conference*, 1988, pp. 218–224.
10. S. Sarma, "Enhanced Self-Test Techniques for VLSI Systems Applied to the IBM Enterprise System/9000 Type 9121 Processor," *IBM J. Res. Develop.* **35,** 390–399 (1991, this issue).

Received October 10, 1990

Steven F. Hajek *IBM Data Systems Division, Neighborhood Road, Kingston, New York 12401.* Mr. Hajek is the Engineering Program Manager for the ES/9000 Type 9121 air-cooled processor. He joined the Data Systems Division in Kingston in 1968 and has since worked in various technical and managerial assignments both within DSD and other IBM divisions. Mr. Hajek received a B.S. and an M.S. in electrical engineering from Purdue University in 1966 and 1968. He is a member of the Institute of Electrical and Electronics Engineers.