A. J. Blodgett
D. R. Barbour

Thermal Conduction Module: A High-Performance Multilayer Ceramic Package

Innovations in package design coupled with major advances in multilayer ceramic (MLC) technology provide a high-performance LSI package for the IBM 3081 Processor Unit. The thermal conduction module (TCM) utilizes a 90 × 90-mm MLC substrate to interconnect up to 118 LSI devices. The substrate, which typically contains 130 m of impedance-controlled wiring, provides an array of 121 pads for solder connections to each device and an array of 1800 pins for interconnection with the next-level package. A unique thermal design provides a cooling capacity of up to 300 W. This paper describes the TCM design and outlines the processes for fabrication of these modules. The TCM is compared to prior technologies to illustrate the improvements in packaging density, reliability, and performance.

Introduction

The key challenge to the component packaging engineer is to provide a package which allows the system designer to exploit the on-going advances in high-performance semiconductor technology. The main driving forces are increased circuits per chip, increase in chip-to-module interconnections, and increase in chip power. These factors place new demands on the package for improved electrical performance and reliability at a reduced cost.

In 1979, IBM introduced a *multilayer ceramic* (MLC), *multichip module* (MCM) technology with the announcement of the 4300 processors [1, 2]. The 35-mm and 50-mm modules used in the 4300 series processors represent a revolutionary advance in packaging state of the art and provide the high-density, high-performance wiring needed to exploit the gains achieved with today's LSI logic devices. These modules contain up to nine 704-circuit, high-performance chips interconnected on substrates containing up to 23 layers of metallized ceramic. A typical module utilizes 4000 circuits and has a maximum cooling capability of 9 W.

This paper describes the *thermal conduction module* (TCM), which significantly extends the MLC technology and is utilized in the recently announced IBM 3081 Processor Unit. The TCM, shown in Fig. 1, is approximately 15 × 15 × 6 cm and contains up to 118 logic and array chips. The TCM utilizes a multilayer ceramic substrate and a new cooling technology which allows the package to dissipate up to 300 W. A typical TCM contains over 25 000 logic circuits and 65 000 array bits of storage. It has 1800 pins to interface with the second-level package.

In the following sections of this paper, the major TCM design features are first described. Then the process for fabricating the TCM is outlined, including substrate fabrication, module assembly, and engineering change methods. The last main section presents a comparison of the TCM to prior technologies in terms of packaging density, reliability, and performance.

Thermal conduction module design

- *General features*

The components of a TCM assembly are shown schematically in Fig. 2, and in actual cross section in Fig. 3. A key element of the assembly is a 90-mm multilayer ceramic substrate capable of providing power distribution and wiring for up to 100 logic chips (with up to 704 circuits per chip) or a combination of 118 logic or array chips. 1800 pins are brazed to the back side of the substrate and connect to the next-level package (which is described in depth elsewhere in this issue [3]) through a connector assembly. Spring-loaded

Copyright 1982 by International Business Machines Corporation. Copying is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract may be used without further permission in computer-based and other information-service systems. Permission to *republish* other excerpts should be obtained from the Editor.

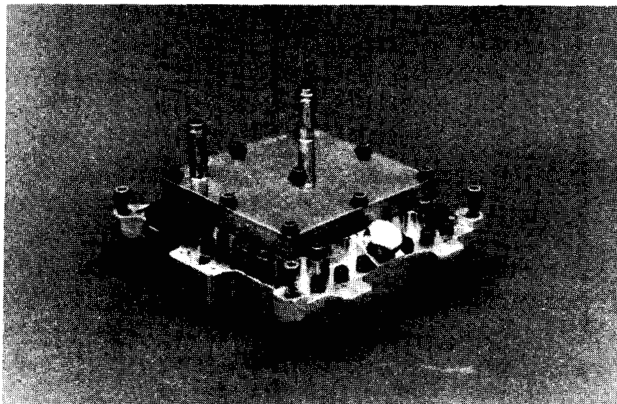

Figure 1 The thermal conduction module (TCM).

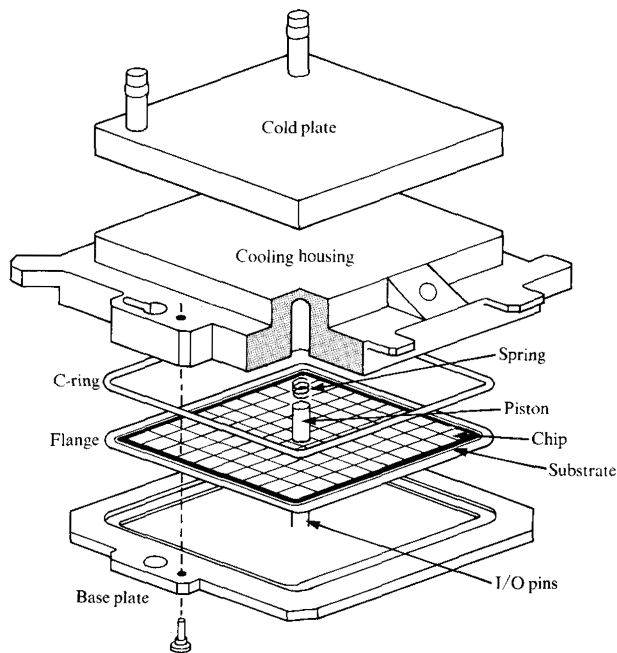

Figure 2 Exploded view of the TCM assembly.

pistons contact the back of each chip and provide the main thermal path to the water-cooled housing. A C-ring, clamped between the cover and a flange which is brazed to the substrate, seals the assembly.

- *Multilayer ceramic substrate*

The 90-mm substrate (see the cross section shown in Fig. 4) consists of up to 33 molybdenum metallized alumina layers which are required for power distribution, for impedance-controlled interconnection of up to 12 000 chip pads, and for wiring to the 1800 module pins. A typical substrate contains 350 000 vias for layer-to-layer connections and 130 m of x-y wiring. The total substrate thickness is approximately 5.5 mm.

The ceramic formulation, consisting of approximately 90% alumina and 10% glass, was selected considering strength requirements, sintering characteristics, and shrinkage compatibility with the molybdenum metallurgy.

A key design feature is the routing of all signal connections from the chip (0.25-mm centers), through the uppermost layers in the substrate, to an array of surface pads, which in turn are connected to internal wiring layers. This design (Fig. 5) provides the ability to delete connections to internal layers by deleting a connection on the top surface and to substitute a surface discrete wire when a wiring change is needed. The discrete wires are ultrasonically bonded to the gold-plated pads.

The center layers of the substrate include 16 wiring planes arranged in x-y pairs to maximize wiring efficiency. Metallized, 0.12-mm-diameter vias on 0.5-mm centers are used for x-plane-to-y-plane connections. Voltage reference planes are appropriately interspersed for signal wiring impedance control.

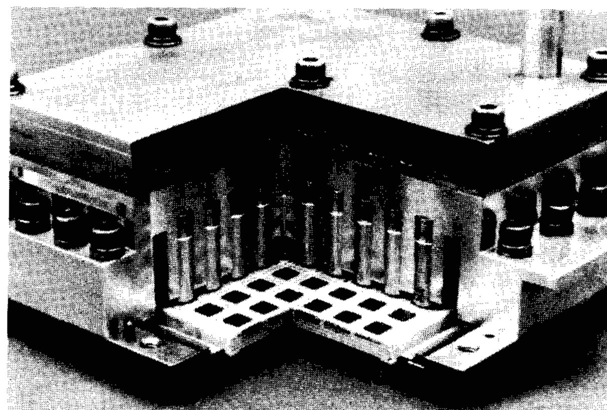

Figure 3 Cross section of the TCM.

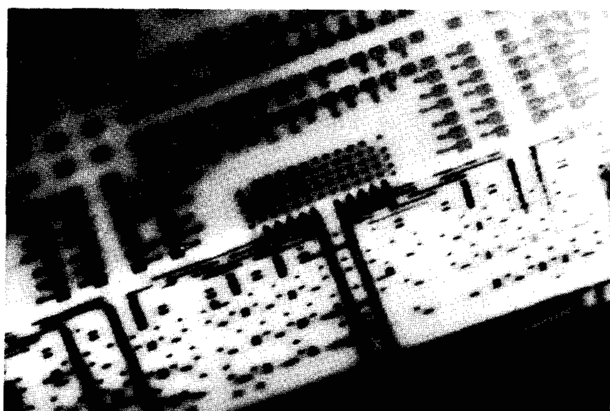

Figure 4 Cross section of a multilayer ceramic substrate.

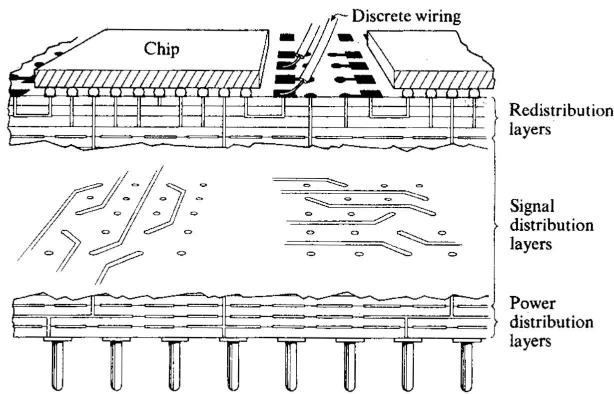

Figure 5 A multilayer multichip module.

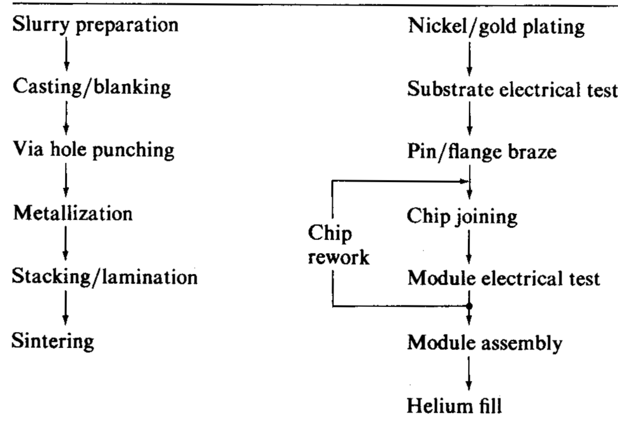

Figure 6 Outline of the process flow for fabricating TCM modules.

The bottom layers are used for power distribution and rerouting of signal connections from the 0.5-mm internal grid to I/O pin positions on a 1.25 × 2.5-mm grid. Power is distributed from the power planes directly to the chips through via columns as shown in Fig. 4.

Electrical parameters

The signal-line reference planes have been designed to provide a characteristic impedance of 55 Ω for the selected line geometry by the use of a combination of nominally 0.15-mm and 0.2-mm-thick ceramic layers with a nominal dielectric constant of 9.4. The bulk resistivity of the sintered molybdenum metallurgy is approximately 10 $\mu\Omega/\mathrm{cm}^2$.

Power is distributed from the substrate pins to the power planes and in turn through parallel paths to chip power pads on the surface of the substrate. The maximum voltage drop in the substrate is 16.5 mV, and the power distribution design allows a minimum of 450 off-chip drivers to be switched simultaneously.

Cooling

The thermal conduction module has been designed to maintain the chip junction temperature below 85°C with a maximum thermal load of 4 W/chip and a total of 300 W/module. The heat sink is a water jacket which mounts to the module cover assembly. The cover assembly contains 5.5-mm-diameter spring-loaded pistons which contact the back side of each chip and form the main thermal path to the water jacket (Fig. 3). The TCM assembly is charged with helium to a pressure of 0.16 MPa (1.6 atm) to minimize the thermal resistances at the chip-to-piston and piston-to-housing interfaces. The total thermal resistance from the chip surface to the module housing/water jacket interface is a maximum of 10°C/W. Papers detailing the thermal design are contained in this issue [4, 5].

Module fabrication

The process for fabricating TCM modules is outlined in Fig. 6.

Substrate

Ceramic green sheet

The basic building block used in the MLC process is the ceramic "green sheet," nominally 0.2 mm or 0.28 mm thick (unfired), which is a mixture of ceramic and glass powder suspended in an organic binder. A key factor in achieving acceptable yields is the formulation of a green sheet that exhibits the necessary strength for handling and processing. In addition, the green sheet must be dimensionally stable to ensure accurate plane-to-plane registration when the layers are stacked and laminated. Strength and stability have been achieved by the proper selection of binder constituents, a controlled casting process, and the use of a molybdenum paste vehicle that does not interact with the green sheet binder.

The organic raw materials are combined with the alumina and glass in a ball mill to achieve a uniform dispersion. This slurry is fed to a continuous caster and deposited on a constantly moving plastic web to form 200-mm-wide ceramic tapes. The caster is fitted with a constant-level slurry reservoir and a doctor blade assembly to control green sheet thickness. The ceramic webs are passed through a series of drying ovens and are subsequently separated from the plastic carrier and spooled.

Each spool of ceramic tape is fully inspected and cut into square blanks in preparation for personalization. A representative sample of the blanks is evaluated for density, compressibility, bond and yield strength, and shrinkage.

Green sheet personalization
IBM-designed, computer-controlled step-and-repeat equipment is used to punch via holes in each green sheet layer. A pallet, used for mounting the green sheets, is an integral part of a precision x-y table that moves the green sheet relative to a stationary die set containing 100 punches at a rate of approximately nine steps per second. A location hole in each corner of the green sheet is used to position the green sheet on the punch pallet. Up to 36 000 via holes are punched in a single green sheet layer.

The screening pastes are tailored for the various patterns shown in Fig. 7 and consist of molybdenum powder uniformly dispersed in a resin and solvent mixture. Metallization of the green sheet is accomplished by extruding the molybdenum paste through a nozzle as the nozzle traverses a metal mask in contact with the green sheet. The vias are filled and the pattern is defined on the surface of the green sheet simultaneously (Fig. 8).

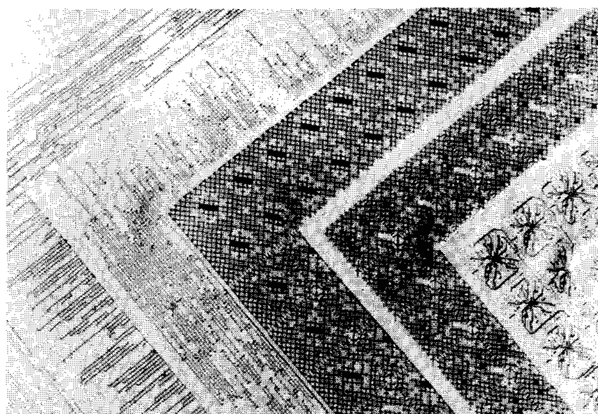

Figure 7 Various screened green sheet patterns.

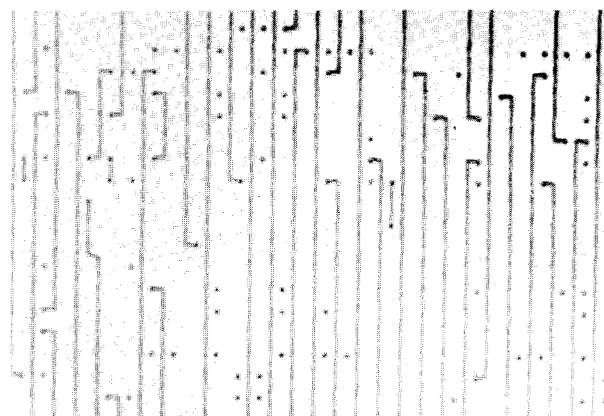

Figure 8 Magnified portion of a personalized green sheet.

Following screening, the metallized layers are dried in a forced-air circulating oven. The drying cycle has been optimized, and is carefully controlled, to avoid dimensional change and damage to the green sheet, particularly on sheets with dense via arrays.

One of the key advantages of the multilayer ceramic process is the ability to inspect and repair individual metallized layers prior to stacking and lamination. Inspection is accomplished automatically with a system which detects deviations in the screened pattern relative to an optimum configuration.

Substrate lamination and sintering
The inspected green sheets are stacked in the desired sequence using the four corner-location holes to ensure accurate layer-to-layer alignment. The stacked sheets are then mounted on a lamination die which is precision machined and assembled to ensure parallelism and flatness, an operation which is critical to achieving uniform green laminate density and uniform shrinkage during sintering.

During lamination, the substrate is sheared from the green sheet stack and compressed at 75°C at a pressure of up to 25 MPa. The substrates are sintered on support plates in a batch kiln over a period of 33 hours. The lamination and sintering processes have been optimized to control shrinkage and ensure precise location of surface features. The linear shrinkage during sintering is approximately 17% and corresponds to a 31% reduction in area.

Residual solvents evaporate and the organic materials start to decompose in a dry hydrogen environment during the initial phase of the sintering cycle. Wet hydrogen is then introduced to accelerate oxidation of residual carbon and, in the range of 1250°C to 1560°C, sintering of the substrate occurs. The substrate is maintained for a period at the peak temperature to ensure maximum densification and the formation of a firm bond between the ceramic and metallurgy. The wet-hydrogen atmosphere is maintained throughout the sintering phase and the initial part of the cooling cycle; dry hydrogen is introduced during the final stage of cooling to room temperature.

Substrate finishing
Nickel is plated on the surface features and diffusion-bonded to the molybdenum base metal to enhance adhesion. Following nickel diffusion, a layer of gold is applied to prevent formation of nickel oxide and to enhance wettability during subsequent soldering and brazing processes. The final plating step is the application of gold on the wiring pads

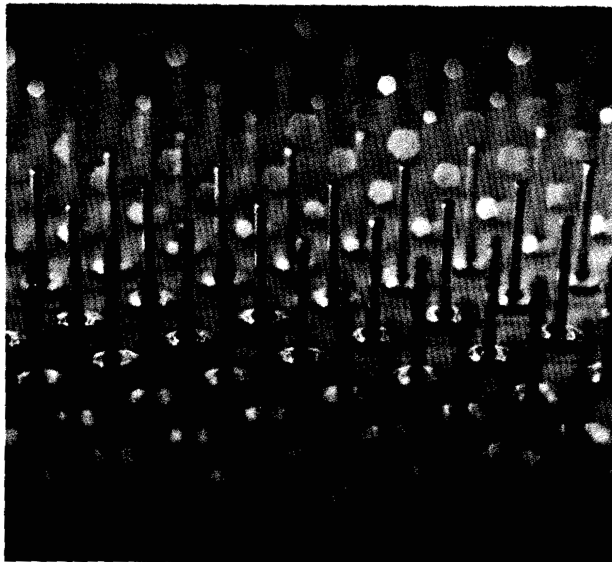

Figure 9 Detail of pins brazed to MLC substrate.

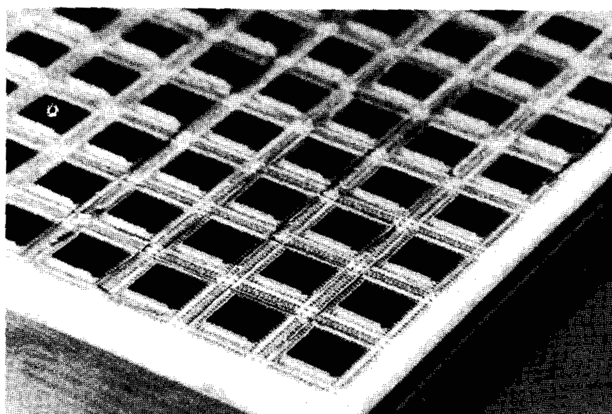

Figure 10 Detail of substrate with reflowed chips.

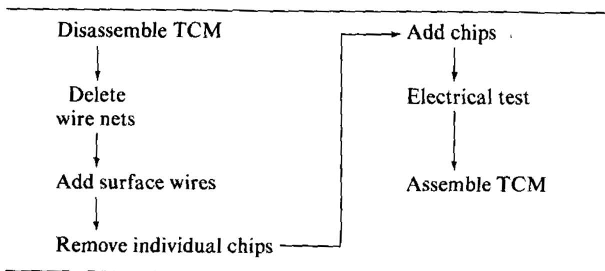

Figure 11 Engineering change process flow.

surrounding each chip site to accommodate ultrasonic joining of discrete engineering-change wires on the module surface. After substrate electric test, the plated pins and flange are simultaneously brazed to the substrate in a hydrogen atmosphere using a gold-tin alloy (Fig. 9).

• *Module assembly*

Chips are joined to the substrate using the flip-chip solder-reflow process first introduced in IBM System/370 [6, 7], as illustrated in Fig. 10.

The chip sites are fluxed with a water-white rosin, and the LSI devices, which are terminated in 120 lead-tin solder pads, are positioned semiautomatically on the corresponding substrate pads. Up to 118 chips are positioned on the substrate in this manner and reflowed in a nitrogen environment. Thousands of chip-to-module connections are made simultaneously in this manner. Following chip joining, the assembly is submerged in a nonconducting coolant and is tested by contacting the I/O pins and probing the surface pads with power applied to the substrate.

Processes have been developed for the replacement of individual chips. Replacement is accomplished by mechanical removal of the chip, thermal rework of the substrate pads, and reflow of the replacement chip either in a furnace or by a localized heat source.

After final testing of the substrate with chips in place, the cooling housing is mounted in place and the entire assembly is charged with helium to a pressure of 0.16 MPa (1.6 atm). A C-ring, clamped between the cooling housing and the substrate frame, seals the assembly.

• *Engineering changes*

A key feature of the TCM design and process is the ability to rework the assembly to make *engineering changes* (ECs). This feature is essential during the design and debug phase of the system development.

The process flow, Fig. 11, utilizes the individual chip replacement techniques just described. Wire nets internal to the MLC substrates are disconnected at the surface. Surface wires are added in the desired pattern using ultrasonic bonding. The photograph of a chip site in Fig. 12 shows the EC pads and surface wiring. The module is tested, then reassembled using a new C-ring and filled with helium.

Technology comparison

• *Hardware*

A typical TCM contains 52 logic chips, 34 array chips, and 5 terminator chips, and utilizes over 25 000 logic circuits and

65 000 array bits. The same function designed in the technology used in the IBM 3033 processor would require 1880 single-chip logic modules, 80 array modules with associated terminating devices, 52 multilayer printed-circuit cards, four large multilayer printed-circuit boards, and associated interconnecting cables. This large reduction in hardware allows a significant reduction in overall processor costs.

- *Reliability*

Interconnections between device and module and between different levels of the package are the main source of reliability failures in electronic assemblies. By utilizing the dense wiring capability offered by the LSI chips and providing the chip-to-chip interconnections on the first-level package, the average number of logic interconnections is reduced by an order of magnitude, as shown in Fig. 13. These statistics refer to the number of logic interconnections between package levels and exclude the interconnections required for power distribution and array devices.

Further, the majority of the interconnections are now chip-to-substrate solder reflow connections—a proven reliable interconnecting system that was first introduced in System/370 in 1969 [6, 7].

- *Performance*

The time-of-flight circuit delay introduced by the package elements becomes a significant factor as the performance of LSI chips increases. In comparison to the 3033 technology, the TCM provides a factor of seven reduction in logic wiring length at the module/card level, with no increase at the chip level, as shown in Fig. 14. Further, the wiring associated with each TCM at the cable and board level is only 150 m compared to 1440 m in the equivalent 3033 technology. This not only significantly reduces the package delay but also reduces the power requirements on the drive circuits.

The impact of the TCM design on the system cycle time is shown in Fig. 15. The elimination of the cards and the order of magnitude reduction in wiring length at the cable and board levels reduce the package delay by a factor of four from that of the 3033 technology.

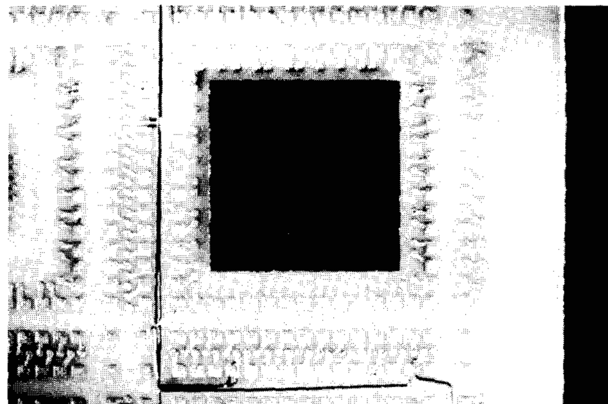

Figure 12 Detail of chip site on MLC substrate showing chip, engineering change pads, and surface wiring.

	3033 Technology equivalent	*TCM*
Chip-to-module	22 560	4368
Module-to-card	22 560	670 (no card)
Card-to-board	4000	
Total	49 120	5038

Figure 13 Average number of logic interconnections between packaging levels in two technologies.

	3033 Technology equivalent	*TCM*
Chip level	30 m	25 m
Module/card level	935 m	130 m
Associated cable/board wiring	1440 m	150 m
Total	2405 m	305 m

Figure 14 Comparison of typical logic wiring lengths of 3033 equivalent and TCM technologies.

Summary

The thermal conduction module represents a revolution in high-performance package design and concept, in process and material technologies, and in manufacturing capability. By providing true LSI capability in the first-level package, it allows the same advances in cost, performance, and reliability achieved at the chip level to be achieved in the package.

Some of the more significant elements of this work have included the following:

- A complex process and material system has been developed and introduced to manufacturing which allows the production of large, complex, tight-tolerance, multilayer ceramic structures. This provides the required LSI interconnection capability, high-performance wiring, and power distribution.
- A novel cooling concept has been devised and reduced to practice which removes the thermal limitation to high-performance LSI at the package level.

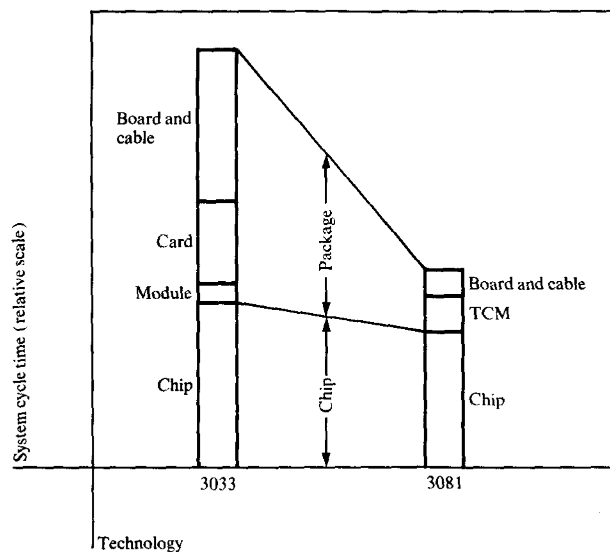

Figure 15 Comparison of system delay distribution of 3033 and 3081 technologies.

- A design has been developed which capitalizes on these advances and extensions of solder-reflow chip-joining technologies. The design features large area array interconnection at both the chip and board level. The area array interconnections provide short wiring paths and engineering change capability, and allow the system designer to realize the advantages offered by LSI.
- Processes and tooling have been developed for engineering change and rework capability: individual chip replacement, wire deletion, and reworkable helium seal.

Although this paper has addressed a specific design, the technology base established is flexible and extendible to other designs and applications.

Acknowledgments

The thermal conduction module represents the culmination of the efforts of many contributors in many areas over a significant time span. Some roots extend back 18 years in time [8, 9]. Others extend geographically to efforts in IBM Germany and IBM France. It is not possible to list all these here. However, the authors hereby acknowledge these contributors and appreciate the opportunity to present a summary of their work.

References

1. A. J. Blodgett, "A Multilayer Ceramic Multi-Chip Module," *Proceedings of the Electronic Components Conference,* IEEE, New York, 1980, pp. 283–285.
2. B. T. Clark and Y. M. Hill, "IBM Multichip Multilayer Ceramic Modules for LSI Chips—Designs for Performance and Density," *IEEE Trans. Components, Hybrids, Manuf. Technol.* **CHMT-3,** 89–93 (1980).
3. Donald P. Seraphim, "A New Set of Printed-Circuit Technologies for the IBM 3081 Processor Unit," *IBM J. Res. Develop.* **26,** 37–44 (1982, this issue).
4. S. Oktay and H. C. Kammerer, "A Conduction-Cooled Module for High-Performance LSI Devices," *IBM J. Res. Develop.* **26,** 55–66 (1982, this issue).
5. R. C. Chu, U. P. Hwang, and R. E. Simons, "Conduction Cooling for an LSI Package: A One-Dimensional Approach, *IBM J. Res. Develop.* **26,** 45–54 (1982, this issue).
6. L. F. Miller, "Controlled Collapse Reflow Chip Joining," *IBM J. Res. Develop.* **13,** 239–250 (1969).
7. P. A. Totta and R. P. Sopher, "SLT Device Metallurgy and its Monolithic Extension," *IBM J. Res. Develop.* **13,** 226–238 (1969).
8. B. Schwartz and D. L. Wilcox, "Laminated Ceramics," *Proceedings of the Electronic Components Conference,* IEEE, New York, 1967, pp. 17–26.
9. H. D. Kaiser, F. J. Pakulski, and A. F. Schmeckenbecher, "A Fabrication Technique for Multilayer Ceramic Modules," *Solid State Technol.* **15,** No. 5, 35–40 (May 1972).

Received August 26, 1980; revised July 7, 1981

A. J. Blodgett is located at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598. D. R. Barbour is located at the IBM General Technology Division Laboratory, East Fishkill Facility, Hopewell Junction, New York 12533.