

First- and second-level packaging of the z990 processor cage

T.-M. Winkel
W. D. Becker
H. Harrer
H. Pross
D. Kaller
B. Garben
B. J. Chamberlin
S. A. Kuppingner

In this paper, we describe the challenging first- and second-level packaging technology of a new system packaging architecture for the IBM eServer™ z990. The z990 dramatically increases the volumetric processor density over that of the predecessor z900 by implementing a super-blade design comprising four node cards. Each blade is plugged into a common center board, and a blade contains the node with up to sixteen processor cores on the multichip module (MCM), up to 64 GB of memory on two memory cards, and up to twelve self-timed interface (STI) cables plugged into the front of the node. Each glass-ceramic MCM carries 16 chips dissipating a maximum power of 800 W. In this super-blade design, the packaging complexity is increased dramatically over that of the previous zSeries® eServer z900 to achieve increased volumetric density, processor performance, and system scalability. This approach permits the system to be scaled from one to four nodes, with full interaction between all nodes using a ring structure for the wiring between the four nodes. The processor frequencies are increased to 1.2 GHz, with a 0.6-GHz nest with synchronous double-data-rate interchip and interblade communication. This data rate over these package connections demands an electrical verification methodology that includes all of the different relevant system components to ensure that the proper signal and power distribution operation is achieved. The signal integrity analysis verifies that crosstalk limits are not exceeded and proper timing relationships are maintained. The power integrity simulations are performed to optimize the hierarchical decoupling in order to maintain the voltage on the power distribution networks within prescribed limits.

Introduction

A super-blade packaging architecture has been developed for the IBM z990 server. This is a system structure based on packaging the system processor cores, the memory hierarchy, and the ports that connect to the system I/O in such a manner that the system capacity can be increased by plugging additional daughter printed-circuit boards into a common center board. The daughter printed-circuit board with its associated electronics and mechanical packaging is referred to here as a *node*. In the system,

the nodes plug into a center board in the center of the processor cage. There are slots for four processor nodes on the front side of the center board. We consider this to be a super-blade because this architecture provides scalability superior to that of other blade processors.

The back of the center board has slots for eight DCAs (the dc-to-dc voltage converter assemblies) which power the four nodes, as well as redundant oscillator (OSC) cards that provide the clocking function for the system and redundant external time reference (ETR) cards

©Copyright 2004 by International Business Machines Corporation. Copying in printed form for private use 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, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to *republish* any other portion of this paper must be obtained from the Editor.

0018-8646/04/\$5.00 © 2004 IBM

enabling the coupling to other systems. The processor cage with the four nodes, center board, and eight DCAs, including two OSC and two ETR cards, is then packaged in a frame along with the bulk power and I/O cage. A second cage is added to provide the remainder of the I/O components required to make up the z990 system.

This three-dimensional packaging structure is one of the factors providing the opportunity to increase the volumetric density of computing power. In a system cage and frame of similar size, the z900 [1], introduced in 2000, had 20 processor cores; the p690, introduced in 2001, had 32 processor cores; and the z990, introduced in 2003, is capable of housing 64 processor cores.

Two other factors are significant with respect to the increase in processor density. The z990 has taken advantage of advances in silicon technology to place two processor cores on a single die, where the z900 has only one. Also, the MCM housing the processor has reduced the via-to-via and line-to-line pitch from 396 μm in the z900 to 375 μm in the z990. The connection to the processor printed-circuit board (PCB) used the 1-mm land grid array (LGA) technology from the p690 [2], where the z900 used the pin grid array (PGA) technology on a 1.7-mm pitch.

However, this blade approach creates challenges to meet the electrical, mechanical, and thermal demands of the system for all elements of the system package design. The thermal challenge emanates from the 5.5-in. (140-mm) node pitch needed to fit the node within the 24-in. (610-mm) cage. The higher density of circuits running at higher frequencies provides a heat load that is challenging to manage. The choice was made to modify the z900 modular cooling unit technology (MCU) to meet the demands of the z990 system. The modification was to use an air-cooled backup in case of MCU failure instead of redundant MCUs [3]. This modified cooling technology maintains the average junction temperature of the chips on the MCM below 50°C.

The electrical challenges were twofold. One is the interface signaling, which was stressed at every level of the system packaging. Communication among the four nodes must have minimum latency and sufficient bandwidth to maintain the performance expected with the z990 system; because of the resulting scalability, we call it a super-node. The 600-Mb/s signal interfaces between nodes traverse a total of 35 in. (89 cm) of printed-circuit-board wiring. A low-loss printed-circuit-board material is selected to minimize voltage loss on these signal interfaces. The second electrical challenge is the distribution of the current demand to chips. Compared with previous systems, the current demands have increased while the voltage rails have decreased [4]. Managing the dc distribution to the four nodes is critical for maximizing the performance of the chips and the interfaces between the chips. The speed

performance of the chips is dependent on maintaining the voltage between the voltage and ground rails above the minimum specified value. The speed performance of the signal interfaces between nodes is dependent on keeping the voltage rails within a specified tolerance in order to maintain the proper voltage reference between the driving and receiving circuits. A major component of the voltage tolerance is the ΔI noise, that is, the noise due to the switching circuits in the system. The power rails and the choice and placement of decoupling capacitors are designed to maintain the voltage variation at the required level over a frequency range extending from dc to several GHz. An 18-in. node-board-to-center-board connector is selected to provide 1080 signal pins and the power pins for the current and signal requirements of the node.

The mechanical designers then have the challenging task of fitting all of the components into the volume allowed them. The center-board connector and the MCM LGA connector both contain a large number of pins on a dense grid. The weight of the MCM heat sink and the node adds to the challenge. The connector design must provide sufficient force to insert the components, enough strength to hold the components in place, and guidance with tight tolerances. When completed, the mechanical design must allow the system to be serviced or upgraded at the customer installation.

This blade architecture, in combination with advanced packaging technologies, yields a processor cage with $2.75\times$ cost-performance improvement over the previous z900 system, $2.6\times$ improvement in operational processor volumetric density, and 56% improvement in component bandwidth, or $2.9\times$ improvement in signal bandwidth per unit area. In addition, it makes available granular growth in enterprise servers similar to that provided by other blade servers. These improvements are enabled through the use of advanced silicon chips, a glass-ceramic MCM, PCB technology, and a high-performance center board and LGA connectors.

In the following sections, we describe the details of the packaging that enabled this density and performance to be achieved.

Logical system structure

The new processor packaging design of the z990 series system accommodates up to 64 processors. The memory size can be increased to a maximum of 256 GB, and the connectivity has been increased to 48 self-timed interface (STI) I/O cables, each with a capability of 2 GB/s.

The increase in the number of processors, memory size, and connectivity allows us to achieve the desired symmetrical multiprocessor (SMP) performance, but it generates a significant increase in the total number of interconnections between the chips. In combination with all packaging restrictions, this results in an increased

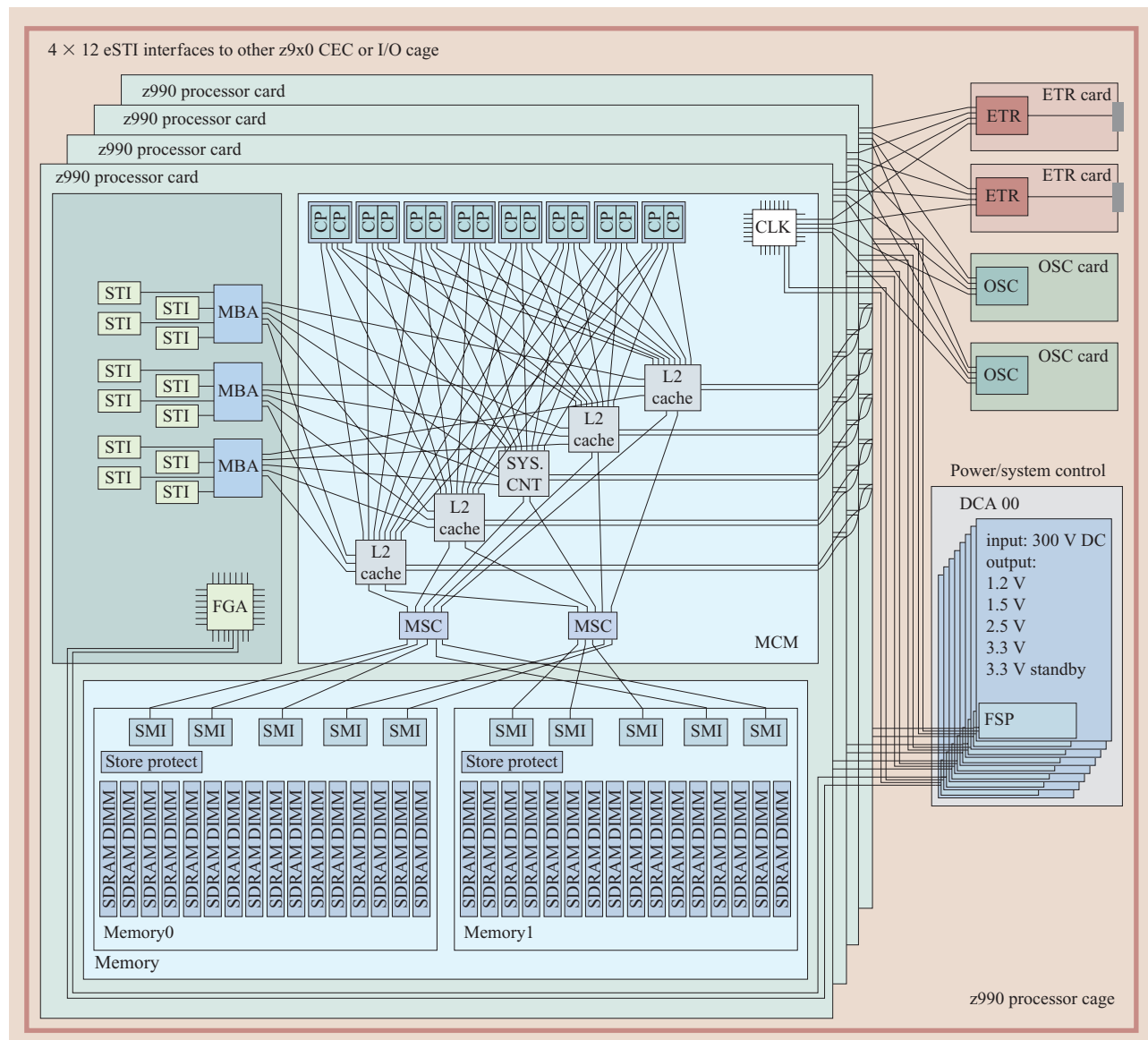


Figure 1

Logical structure of the z990 processor cage.

complexity of the first- and second-level package—the MCM and the printed wiring cards and boards.

The high-level logical structure of the z990 system is shown in **Figure 1**. The 64-processor system is divided into four processor cards, each containing 16 processors, up to 64 GB of main memory, and 12 STI interfaces. Two high-availability redundant power supplies, called distributed converter assemblies (DCAs), provide redundant power to each processor card. One flexible service processor (FSP) card, hosted in the DCA, controls one processor card in conjunction with the field-replaceable unit gate array (FGA); the second FSP, hosted in the second DCA, is for

redundancy. A maximum of eight DCA cards with eight FSP daughter cards are used in each central electronic complex (CEC) cage. Two redundant oscillator (OSC) cards generate all clocks required to run all 64 processors synchronously. In addition, two redundant external time reference (ETR) cards are implemented to allow the z990 system to be coupled to another system. OSC and ETR cards are connected to all four clock (CLK) chips, each one located on one processor card.

Specifically, the center of each processor card is the MCM, containing eight processor chips with two z990 processor cores each, resulting in a total of 16 processors

per card. The Level 2 (L2) cache structure demands four L2 cache data chips to provide a total of 32 MB, and one system control chip, which is a separate chip because of the high pin count connectivity requirement. Each processor chip is connected with a 16-byte-wide bidirectional bus to each L2 cache and system control chip on each MCM.

The system control chip is an I/O-limited chip. It requires 1,666 signal connections, leading to a chip size of 16.8×16.8 mm. Because this chip arbitrates the system communication, it is connected to all processor and L2 cache chips on the same node as well as to the corresponding system control chip on each of the other nodes, resulting in the extremely high signal I/O number.

The memory connection is made from the L2 cache chips by two main store control (MSC) chips, also soldered onto the MCM. Each MSC chip communicates with five synchronous memory interface (SMI) chips, which are located on two memory cards. Each memory card contains 16 dual inline memory modules (DIMMs) and the storage protection required for IBM z/Architecture*. This provides a maximum memory size of 32 GB per memory card, 64 GB per processor node, and a total of 256 GB per system.

The interfaces between the two MSC and ten SMI chips are extremely critical to the system performance; it must run on a 2:1 frequency ratio with respect to the operating frequency of the processor.

The STI I/O connection, used to hook up zSeries I/O cages or connect to other zSeries systems, is generated by three memory bus adapter (MBA) chips, located on an I/O card which is mounted on the processor card. Each MBA is connected to all L2 cache data and system control chips. These interfaces operate at one fourth of the processor chip operating frequency. Four STI cable channels, each supporting 2 GB/s of unidirectional links, are driven from each MBA. This results in 12 STIs per processor card and up to 48 STI I/O paths per system.

The clock chip located on the MCM receives a reference clock from both oscillator cards. The reference clock is buffered and distributed to all processor, cache, system control, memory bus adapter, and memory storage controller chips. The additional function of the clock chip is to read and distribute the internal machine status by shifting bit chains out of the system. The external time reference (ETR) is needed to synchronize up to 32 zSeries systems in a Parallel Sysplex* configuration. The ETR function is implemented on the clock chip in order to achieve chip size optimization for the CEC chip set. The clock chip also connects to the two flexible service processor (FSP) cards. The FSP cards, hosted in the DCA, control the processor cage and communicate with the system console.

The connections among the four processor cards are provided through the cache and system control chips. The four processor cards are plugged into the center board. Each L2 cache chip and all system control chips are connected to all other corresponding L2 cache and system control chips in the different nodes using a closed-loop approach. They are wired from processor card slot 1 to slot 3, from slot 3 to slot 4, from slot 4 to slot 2 and from slot 2 back to slot 1. This provides the advantage of having three processor cards connected at all times if one processor card requires maintenance. The cache chips are connected by means of an 8-byte unidirectional store bus and an 8-byte unidirectional fetch bus. To achieve the operation frequency of all off-chip buses, a special so-called elastic interface must be used. This elastic interface is described later, in the section on elastic interface bus signaling.

The redundantly designed system components, such as power distribution, OSC, ETR, and FSP control cards, as well as the divided power domains for each processor card and the single processor card maintenance, result in the excellent reliability, availability, and serviceability (RAS) which are the hallmark of all S/390* mainframes and enterprise zSeries servers.

MCM

The MCM is a 93-mm $\times$ 93-mm high-performance glass-ceramic substrate. The floorplan shown in **Figure 2** illustrates the 16-chip MCM, which consists of eight dual-core processor chips labeled PU0–PU7 in the figure, four L2 cache chips denoted as SD0–SD3¹ creating 32 MB of shared second-level cache, a system controller chip (SCC), two memory controller chips (MC0, MC1), and a clock chip (CLK), which provides the clock distribution as well as pervasive function for the MCM. The MCM was designed using the methodology described in [1, 5] with upgraded noise-checking tools and routing capability.

The MCM technology has undergone significant evolution from the z900 system to enable the dense processor packaging and provide the cost reduction required to remain competitive. The attributes of the z990 MCM have been added to a table from [6] shown here as **Table 1**. The areas of significant change are indicated with bullets in the 2003 z990 column. Leveraging the MCM technology developed for the IBM pSeries* p690 server [7], a glass-ceramic MCM without thin film is chosen with a 5,184-contact land grid array (LGA) attachment to the second-level processor board. The 93-mm $\times$ 93-mm form factor requires almost 50% less area than the z900 MCM, enabling the large increase in packaging density and the higher level of silicon circuit integration and allowing two cores per processor chip. The reduction of MCM line

¹ Also designated as SCD chips.

pitch from 396 μm to 375 μm enables the design of an MCM without the additional thin-film layers of the z900, thereby providing a processor module with an improved performance per unit cost.

Second-level packaging for the central electronic complex

The physical partitioning of the logical structure is shown in **Figure 3**. This section describes the main second-level package component, i.e., the processor board. The resulting processor subsystem is integrated in a mechanical structure known as the central electronic complex (CEC) cage, whose backbone is a printed wiring center board. Four processor cards, eight DCA/FSP cards, two OSC, and two ETR cards can be plugged into it. The board also contains the vital product data (VPD) for the processor cage. The following sections describe the main cards.

Processor card

The processor card holds the 16-way processor MCM, two memory card slots, and the I/O interface card. The card can be plugged into the center board using a 1,080-signal-pin connector. Compliant pin card connectors are used for the connection from the processor card to the center board as well as to the memory card and the I/O card. Surface-mount technology (SMT) ceramic capacitors are soldered on both sides of the processor card. The on-card decoupling hierarchy from high frequency to low frequency is completed by pin-in-hole electrolytic capacitors.

Center board

This board contains compliant pin connectors instead of soldered components. The four processor cards are plugged from one side, while the eight DCA/FSP cards and the two OSC and two ETR cards are plugged from the opposite side. The center board also holds the VPD card (not depicted in Figure 3), which contains the board serial number and other system-related information.

Memory card

The maximum amount per memory card is 32 GB using 16 double-data-rate type 1 (DDR1) dual inline memory modules (DIMMs), each 2 GB in size. Each memory card contains 16 soldered DIMMs and five synchronous memory interface (SMI) chips as well as store-protect memory chips. The memory card is treated as a field-replaceable unit (FRU).

I/O card

Each of the three MBA chips located on the I/O card drives four STI interfaces. This gives us a connectivity of 12 STIs per processor card and up to 48 STI interfaces per system. All STI cables can be plugged and unplugged

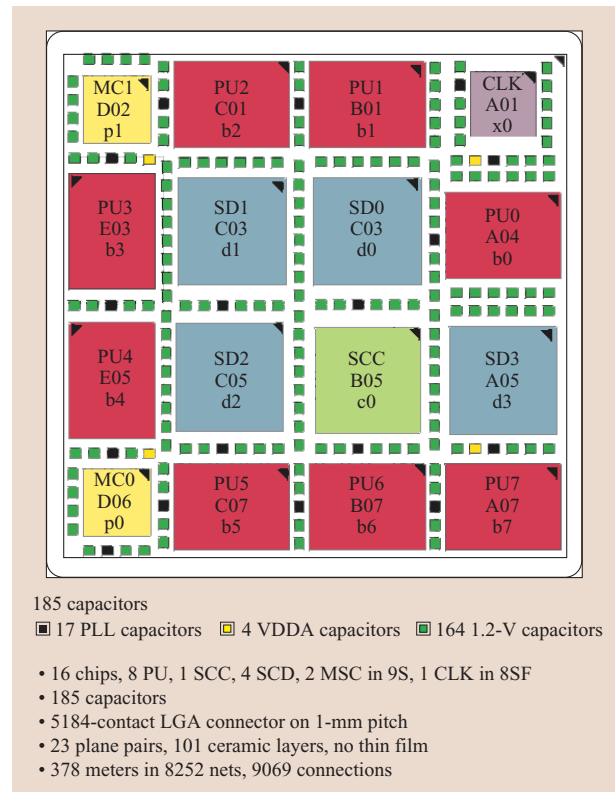


Figure 2

Floorplan of the 16-chip, 16-core MCM for IBM eServer z990.

without shutting off power, allowing concurrent maintenance, upgrade, and reconfiguration capability.

OSC card

Two OSC cards are plugged into the z990 center board. One card is always active, while the other one is a redundant backup. Each card provides the clock signal generator for various master clocks for the clock chip on the CEC module. For enhanced system reliability, the clock chip has two independent inputs; during the system power-on process, the clock chip selects which OSC card becomes the master and which one the backup after the power-on sequence is completed. All four processor cards are controlled by the same OSC card.

ETR card

Two ETR cards are plugged into the z990 center board. One card is always active, while the other one is a redundant backup. Each card provides the external time reference (ETR) optical receiver function for the clock chip on each processor MCM. Each card contains receivers/drivers for optical fiber cables. For enhanced system reliability, the clock chip has two independent

Table 1 Attributes of zSeries processor MCMs.

<i>Year--server</i>	<i>1997--G4</i>	<i>1998--G5</i>	<i>1999--G6</i>	<i>2000--z900</i>	<i>2003--z990</i>
<i>Number of chips</i>	30	29	31	35	•16
<i>Number of processor unit chips</i>	12	12	14	20	•8 dual core •16 processors
<i>Processor frequency (Processor voltage)</i>	370 MHz (2.7 V)	500 MHz (2.0 V)	637 MHz (2.0 V)	770 MHz (1.7 V)	1.2 GHz (1.35 V)
<i>L2 cache (MCM voltage)</i>	3 MB (2.7 V)	8 MB (2.6 V)	16 MB (2.0 V)	32 MB (1.7 V)	32 MB (1.35 V)
<i>MCM power capacity (W)</i>	1,050	800	900	1,400	800
<i>Processor nominal* junction temperature (°C)</i>	45	20	15	0	•40
<i>MCM technology</i>	127.5 mm × 127.5 mm alumina/thin-film redistribution	127.5 mm × 127.5 mm glass-ceramic/thin-film wire	127.5 mm × 127.5 mm glass-ceramic/thin-film wire	127.5 mm × 127.5 mm glass-ceramic/thin-film wire	•93 mm × 93 mm glass-ceramic
<i>Total number of C4s/wire length (m)</i>	83,000/460	80,000/600	85,000/640	101,000/920	•55,000/378
<i>Max. number of signal C4s per chip</i>	1,244	1,244	1,244	1,186	1,816
<i>Ground rules for the thin-film layers</i>	Four layers/56 μm	Six layers/45 μm	Six layers/45 μm	Six layers/33 μm	N/A
<i>Glass-ceramic ground rules. (Note: G4 used alumina.)</i>	87 layers/450 μm	75 layers/450 μm	87 layers/450 μm	101 layers/396 μm	•101 layers/375 μm
<i>MCM I/O technology</i>	3,528 total pins/1,764 signal	4,224 total pins/2,450 signal	4,224 total pins/2,450 signal	4,224 total pins/2,489 signal	•5,184 total pins/2,930 signal
<i>Nominal characteristic impedance, Z₀ (Ω)</i>	9,211–50	TF - 39 GC - 60	TF - 39 GC - 60	TF - 43 GC - 55	GC - 55
<i>Nominal signal delay (ps/mm)</i>	11.5	TF - 6.4 GC - 7.8	TF - 6.4 GC - 7.8	TF - 6.4 GC - 7.8	GC - 7.8
<i>Nominal signal-line resistance (Ω/mm)</i>	0.035	TF - 0.2 GC - 0.022	TF - 0.2 GC - 0.022	TF - 0.24 GC - 0.022	GC - 0.022
<i>Signal-line cross section</i>	90 μm × 30 μm	TF - 18 μm × 6 μm GC - 70 μm × 25 μm	TF - 18 μm × 6 μm GC - 70 μm × 25 μm	TF - 16 μm × 4.5 μm GC - 70 μm × 25 μm	GC - 70 μm × 25 μm
<i>Dielectric thickness (fired) (mm)</i>	0.2/0.15	TF - 0.01 GC - 0.111	TF - 0.01 GC - 0.111	TF - 0.01 GC - 0.092	GC - 0.092
<i>On-MCM capacitors</i>	100 nF/cap 220 capacitors	200 nF/cap 177 capacitors	200 nF/cap 202 capacitors	300 nF/cap 366 capacitors	300 nF/cap 185 capacitors

*The term *nominal* refers to the expected value of the respective parameters.

inputs; during the system power-on process, the clock chip selects which ETR port becomes the master and which one the backup after the power-on sequence is completed. The timing synchronization of coupled multiple central electronic complexes is done by the ETR electronics on

the ETR card. It is connected to the external control function, called a sysplex. This allows the use of up to 32 systems, each one having a 64-way zSeries node, which may result in a maximum sysplex of 2048 processors within a single system image.

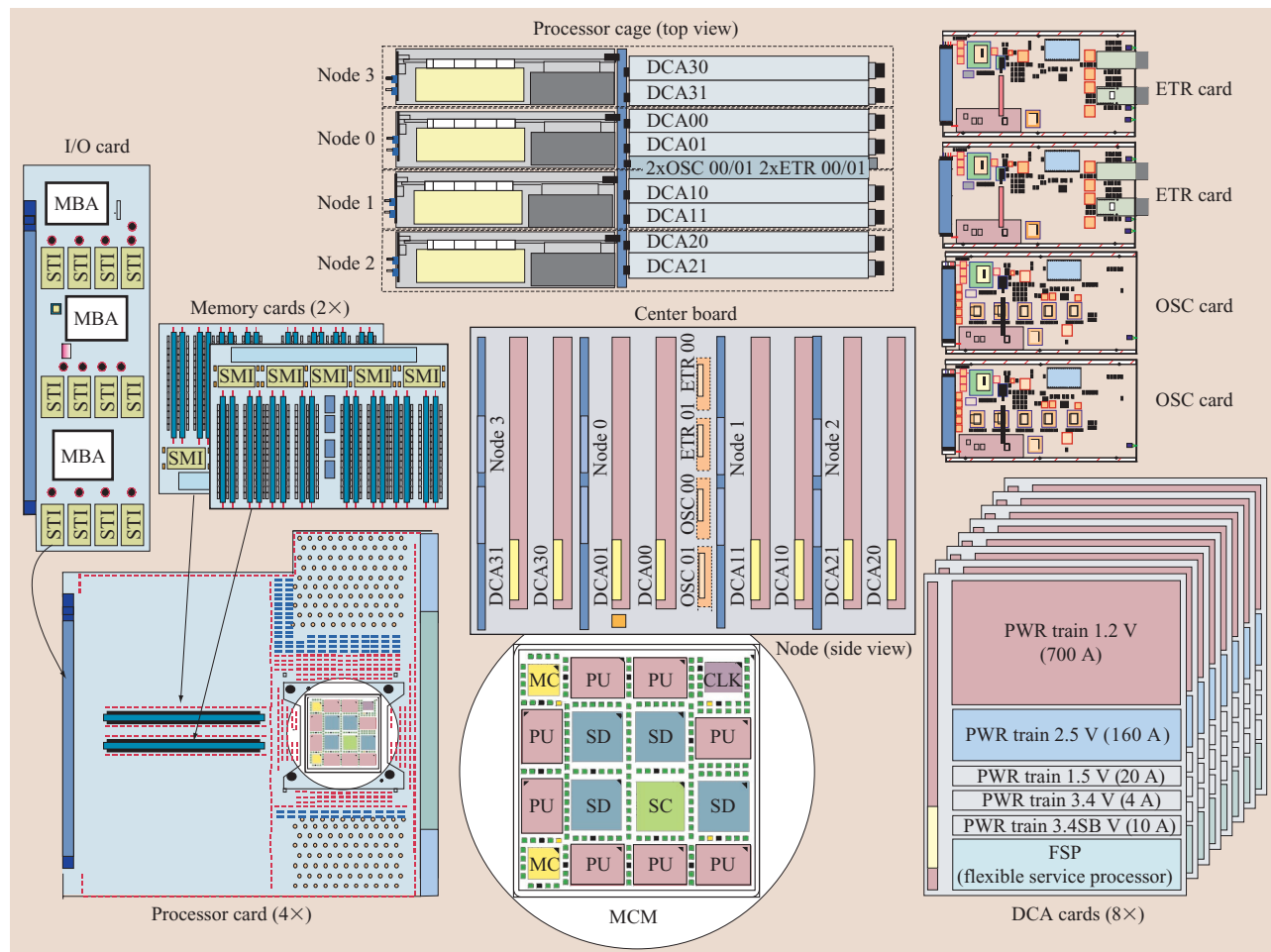


Figure 3

Second-level package partitioning for the central electronic complex. (FSP: flexible service processor.)

DCA/FSP card

Four groups of two dc–dc adapter (DCA) cards are plugged into the z990 center board to provide $n + 1$ redundant power supplies for all logic voltages (1.2 V, 1.5 V, 2.5 V, 3.3 V, and 3.4 V standby). One card is required by the electrical load, while the second one provides the required redundancy for each processor node. Each node can be powered on/off individually to perform maintenance on a single processor card base. All eight DCA cards are hot-pluggable (i.e., they can be plugged in while the system is operating). Each of the eight DCA cards hosts one flexible service processor (FSP) card. The FSP card controls the processor card infrastructure; i.e., it reads VPD and configuration data, generates reset signals, initializes the processor through a high-speed interface into the CLK chip, and runs the DCAs. Each FRU contains a control chip to control the logic on each card and a

VPD data module, which contains information such as the card type, the part number, and the serial number.

Processor card

The processor card is 508 mm wide, 470 mm high, and 3.5 mm thick. The card connectors are mainly VHDM** (Very High Density Metric**) from Teradyne, Inc. They are six- and eight-pin-row high-speed connectors with matched impedance (50 Ω) for the signal paths. Each signal row is separated by internal shield plate contacts, providing an excellent high-frequency return path. All connectors are pressed into plated-through holes (PTHs) in the board.

The MCM land grid array (LGA) connector is a press-on connector which uses an insulated interposer containing 5,184 conductive buttons. The contact springs are divided into four quadrants, with 1,296 (36 $\times$ 36) contacts

Table 2 Pin count summary.

<i>MCM/card/board type</i>	<i>How often/cage</i>	<i>Signal pins</i>	<i>Power/Ground pins</i>	<i>Pins/Component</i>	<i>Total contacts/system</i>
Processor MCM	4	2,970	2,214	5,184	20,736
Memory cards	8	504	706	1,210	9,680
I/O cards	4	720	710	1,430	5,720
OSC cards	2	120	124	244	488
ETR cards	2	120	124	244	488
STI cable connectors	48	42	24	66	3,168
SEEPROM card	1	8	8	8	8
DCA card	8	120	348	468	3,744
Total component PTHs in processor cage					44,032

each. The LGA pads are placed on a 1-mm rectangular grid.

The surface-mount technology (SMT) capacitors are soldered on both sides of the z990 processor board. This board assembly comprises 35 large components and 5,066 decoupling capacitors. The board contains a total of 3,100 signal wiring nets with 15,000 signal and power pins. The total wiring length for all of the interconnections in the z990 processor card is 767 m (30,231 in.). All of the onboard connections are as short as possible, but they must be length-adjusted within one bus and clock group to meet the elastic interface specification.

Ten signal layers are used to fan out the MCM. Eighteen power layers are used to distribute the ground and the logic voltages (1.2 V, 1.5 V, 2.5 V, 2.5 V memory, 3.3 V and 3.3 V standby). Because of performance requirements, all of the buses and most of the control signals are wired point-to-point. The average pin density is about 39.7 pins/in.², while the pin density under the MCM is 625 pins/in.² and 250 pins/in.² under the VHDM connector. An overview of the components with their signal/power pins is given in **Table 2**.

Printed-circuit-board (PCB) technology

Choosing an appropriate board technology is one of the challenges in designing a complex system. Different aspects must be considered when defining a board cross section to meet performance and manufacturing requirements. The signal routing requirements define the number of signal layers that are needed, while the electrical requirements define the dielectric thickness between the signal and voltage layers for impedance and crosstalk control and to provide proper signal reference. In addition, the power distribution must contain enough copper to maintain acceptable voltage levels.

The manufacturing design is limited by board size, because of its impact on the board yield, as well as by board thickness, because of drilling and copper plating limitations. The maximum board thickness is a function

of the minimum plated-through-hole diameter. The plated-hole diameter is defined by the system wirability requirements and the connector technologies. Two distinct factors limit the amount of copper that can be connected to a given plated-through hole (PTH). The PTH drilling limitations are strongly linked to the amount of copper being drilled through. Increased drill breakage and poor PTH quality can increase board costs and reduce reliability if too much copper is connected to a PTH. Soldering and reworking of pin-in-hole components are much more difficult because of the faster heat dissipation in such massive power planes. This can increase costs by reducing the assembly yield.

Power and ground planes provide the signal reference in the processor board. The thickness of these conductive planes is generally 0.03 mm (1.2 mil) using the so-called one-ounce-per-square-foot copper foils. The minimum line width is 0.076 mm (3.0 mil), and the line spacing is kept at or above 0.10 mm (4 mil). This represents the optimization of electrical requirements and the number of signal layers.

In general, the technology selection was completed with a continuous interaction between development and manufacturing to ensure that no parameters were unduly increasing manufacturing costs.

The main design considerations with respect to power distribution are the number of voltage planes and the dc voltage drop. The number of voltage planes is a function of the number of different voltages used on the board, the amount of copper required per voltage domain, the maximum copper thickness, the number of reference planes required by the signal planes, and any restriction on split-voltage planes. Since the board-resistive heating was not a concern in this design, the voltage drop limit was set by the voltage variation tolerated by the active components. This defined the total amount of copper required for each voltage level.

Further considerations are based on the signal layer design. In the board, the number of signal layers is determined by the wiring channels needed for the

Table 3 Overview of elastic interface (EI) buses in z990 eServer.

<i>EI type</i>	<i>Max. PCB lengths</i>	<i>No. of connectors</i>	<i>Bit rate (Mb/s)</i>
MSC ↔ MBA	26.5	1	625
MSC ↔ SMI	21.1	1	500
SCD/SCC ↔ SCD/SCC	34.2	2 or 4	625

fabricated using a 1.0-oz copper foil with a thickness of 0.036 mm (1.4 mil).

Although 100% wiring efficiency, with no wiring restrictions, would allow fanning out the MCM in fewer signal layers, noise magnitude and length restrictions on the interconnections made the wiring of this board in ten signal planes challenging. Specifically, the coupled noise magnitude constraints allowed a maximum of two wires between two adjacent module pins. With an average signal-pin-to-power-pin ratio of 1.6 to 1 and a pin quadrant depth of 24 pins, at least 15 signals had to be brought out in one vertical set of wiring channels across all ten signal layers. To meet the system wiring demand in ten signal planes requires an optimized MCM pin assignment.

As shown in Figure 4, most of the power planes were placed at the top and at the bottom of the board. In the center of the board, signal-plane pairs were separated by 1.2 V or by ground planes. This approach was successfully used in many previous S/390* generations because it offered two main advantages. First, the voltage/ground planes in the board center adjacent to the signal planes ensure a good high-frequency signal return path for all signals. As a result, the discontinuities between the cards and the connectors are minimized, and the signal behavior is improved. Second, the 1.2-V/GND plane pairs on the top and on the bottom of the board minimize the via length between these planes and the decoupling capacitors mounted on the board surface. Thus, the via inductance is reduced. As in the previous system [1], a special pad via design with a minimized parasitic inductance was used, resulting in an improvement in the effectiveness of the decoupling capacitors.

Electrical system design

The new system architecture of the IBM eServer* z990 demands a new method for signal transmission, the so-called elastic interface bus [8], to meet system performance goals. This means that special requirements must be taken into account when choosing the cross sections of the different cards and the center board. The complex bus structures connecting logical components on

different cards require a new closed form of timing and noise verification of all signal wiring in the system in parallel after physical design. Because of increased currents and decreased voltage levels and thus higher sensitivity to voltage fluctuation of the logic, the voltage stabilization presents another challenge to system design.

Elastic interface bus signaling

The electrical design of the first- and second-level package structure for a complex system such as the z990 server demands a large number of simulations and modeling of packaging structures. In an early pre-physical-design phase, this is needed to assess timing and noise requirements for the different types of driver and receiver characteristics and net topologies. With respect to the increased length of the ring bus that is used to connect the four nodes, the need to accelerate the data rates for the off-MCM buses has already been addressed in the high-level design phase.

The elastic interface, as described in [8, 9], is used for the high-speed off-MCM buses. In this interface, a bundle of single-ended data nets is grouped with a differential clock. The bus consists of a certain number of such groups. The synchronization of this bus is performed in three steps: First, the data lines of a group are de-skewed; second, the clock is centered with respect to the data bits in that clock group; third, the target cycle is set on the basis of the latency of the longest bit of the bus. This initializing is executed once at the power-on sequence of the system in the interface alignment procedure (IAP). This allows compensation of the skew that results from length differences and the timing impact of process and environmental conditions. **Table 3** shows an overview of the elastic interfaces used in the z990 eServer, with their different lengths, topology, and cycle times. The most challenging interface here is the node-to-node interface between the system controller chip (SCC) and the system controller data chip (SCD). This interface is designed as a double-ring interface so that communication with adjacent nodes is possible in both directions. Especially in a two-node system, this ring must be closed with a passive pass-through card plugged into the non-populated slots. This increases the maximum PCB length to more than 34 in., and the net path consists of four card connectors instead of two.

A tremendous effort has been made to meet the requirements, specifically to minimize delay differences and to achieve the needed signal eye opening at the receiving chip. The signal eye opening defines the signal delay margin (vertical eye opening) and the signal-amplitude-to-threshold margin (horizontal eye opening). One demand for the critical buses was to keep the number of discontinuities as small as possible to avoid having to change the signal layer in one card to avoid vias in the

signal path. Therefore, the pin assignment of the MCM and VHDM board connector was optimized to prevent crossing lines causing a change in signal layer. In addition, this provides the options of wiring the longest nets on the processor card using 45° wiring instead of only orthogonal wiring, without blocking other wiring channels, and of reaching a length below the minimum wiring length (manhattan distance). This helps reduce the maximum latency and attenuation of the longest nets to meet system requirements.

To control the number of layers and the wiring utilization in the MCM, no length adjustment was done for the off-MCM nets on the MCM itself. With a specially defined signal-length-adjusted wiring (called “trombone wiring”) on the processor card and a semiautomatic adjustment of these wires using a newly introduced Cadence subprogram (skill code) for the router, the skew targets were met and a nominal delay difference of less than 150 ps was achieved on the different elastic interface groups. To meet the noise specification, a line-to-line spacing of 12 mils was reached in the open wiring area of the PCBs. A key issue for the signal eye opening on the long interconnections is the card/board signal line attenuation. This attenuation is primarily a function of the bulk resistance and the surface area of the wire, as well as the loss tangent of the used laminate material. Fulfilling the performance goals required the use of advanced low-loss material (with the characteristic material-related loss factor $\tan \delta < 0.13$ @ 1.5 GHz) with 1-oz copper lines. This reduced the attenuation from 0.104 dB/cm at 1 GHz for FR4 standard material with 0.5-oz copper lines to less than 0.082 dB/cm. In addition, this allowed the use of standard series-terminated signal nets driven by 35- Ω drivers and terminated with diode clamps, avoiding the design of complex I/O circuits with restrictive layout and routing constraints, such as very small on-chip resistance limits. Further analysis proved that a tight control of the impedances in the PCBs was essential to reach the target eye opening. In the worst case, the signal wiring passes five cards and boards with the associated connectors. It turned out that the reflections on the different interface boundaries between the card and boards degrade the signal levels significantly. Therefore, the impedance variation was kept in the range of nominal $50 \Omega \pm 10\%$.

Timing and noise verification

With the greater packaging complexity and higher frequencies presented by the design of the z990, it is imperative to verify the electrical quality of all signals in the system once the physical design has been completed. The ideal method of verification is to run a circuit simulation for each signal; however, the long process times of conventional circuit simulators such as Powerspice make it impossible to perform an analysis on a whole

system, which contains thousands of signals, in a reasonable amount of time. An alternate simulation algorithm, the so-called Fastline [10] analysis package, was invented to overcome this hurdle and greatly reduce the time and resource needed to perform system simulation.

Fastline is a linear simulator that makes use of linearized I/O models and triangular impulse responses for various lengths of coupled transmission lines with frequency-dependent inductance, resistance, capacitance, and conductance terms to perform its signal simulation. All nonlinear Powerspice I/O models for the system are first translated into linearized Fastline I/O models. For every transmission-line model used in a system, a database is created which contains various triangular impulse responses for a set of incremental lengths of transmission line. Rules files are created to describe which models, or triangular impulse response databases, should be used, given certain wiring scenarios for each component in the system. A description of the system and its interconnections and package hierarchy is also supplied as input to the Fastline process. Fastline is able to trace each circuit in the system and extract the appropriate information from these databases as it performs a simulation at much greater speeds than Powerspice.

Two types of analyses can be performed with Fastline. Fastline is able to report delays and timing slacks in order to ensure that the system signals are meeting their timing requirements, and it also calculates total noise from the combinations of pairwise-coupled segments along the length of each net in order to ensure that system noise criteria are met. The timing simulation for a system, ~12,000 nets, takes about four or five hours to complete, and the noise simulation for the same number of nets takes about 24 hours. This turnaround time is very short compared to Powerspice. The achieved accuracy is very satisfying, since a comparison of Fastline and Powerspice typically shows less than 10% difference between the two simulation methods.

Both timing and noise analyses were performed for the z990 after initial wiring was complete. Potential problems were identified as a result of the Fastline analysis, improvements were made, I/O books were changed, and wiring was rerouted. Fastline was rerun, and the process was iterated until the Fastline results showed that all design requirements for timing and total system noise had been met. The ability to improve the physical design in such a short amount of time instead of discovering problems in hardware saved the program significant time and money.

Signal referencing and transmission-line impedance control

Controlling signal delay and coupled signal noise is key to meeting the required system performance. In particular,

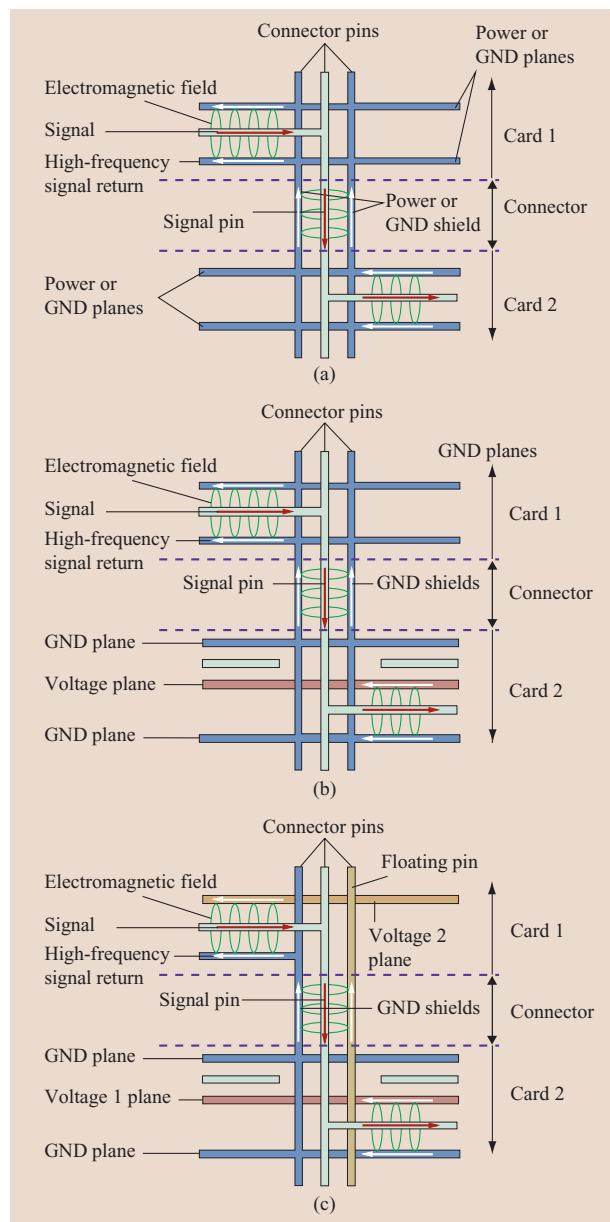


Figure 5

(a) Case 1: High-frequency signal return (ideal case). (b) Case 2: High-frequency signal return (quasi-ideal case). (c) Case 3: High-frequency signal return (worst case).

the ring bus wired through different packaging levels as chips, MCMs, cards, and connectors requires a close control of the signal-line impedance and the discontinuities in the signal path. Multiple changes of signal-line impedance along a signal path cause signal reflections, limiting the signal eye and thus system performance. The signal-line impedances in all packaging levels were kept in a tolerance range of $50 \Omega \pm 10\%$.

Specifically, the signal-line impedances of the cards, the center board, and the connectors were kept in a range of $50 \Omega \pm 10\%$, while the impedance of the MCM was $54 \Omega \pm 10\%$.

Special attention must be given to the signal reference in different packaging levels, particularly when signal wires are changing packaging levels, e.g., from card to connector or from MCM to card. The electromagnetic fields surrounding each signal wire carrying a signal cause coupled noise in adjacent signal wires and high-frequency signal return currents in all adjacent voltage or ground planes or pins. It is well known that the coupling between signal wires increases with decreasing spacing. The amplitude of the high-frequency signal return current in adjacent planes, connector pins, or card vias in a complex system design is also a function of the distance to the signal wire. In order to minimize discontinuities when the signal path changes packaging levels, e.g., from card to connector, the high-frequency signal return path must be closed. This demand can be satisfied by using the same signal reference, e.g., ground and/or voltage, in all packaging levels. As depicted in **Figure 5(a)** for the ideal case, the signal wire is embedded either between ground layers in both cards and ground shields in the connector or between voltage layers in both cards and voltage shields in the connector.

In the center board, the voltages for all nodes are separated in order to avoid compensation currents between different nodes, while the ground for all nodes is common. Thus, signal layers are simply kept between ground layers. Also, the shields in the node connector to the processor unit (PU) card are connected only to ground. Special power connectors are used to distribute the currents to the cards carrying the MCM PU cards.

Because of power distribution requirements in different packaging levels in this complex system, the signal wiring cannot be kept only between ground planes on all packaging levels. The pins in the MCM mounted on the PU card have a 1:1 or 2:1 signal to 1.2-V power/ground pin assignment. In order to close the high-frequency signal return between the MCM and the PU card, the signal layers in the PU card were alternated between 1.2-V voltage planes and ground planes. As depicted in **Figure 5(b)** for Case 2, the high-frequency signal return for the ring nets wired in card 2 is partially broken in the ring connector area because the shields in the connector are connected only to ground, while the signal wire in card 2 is wired between voltage and ground layers. Because of the good plane capacitance between the 1.2-V voltage layers and the adjacent ground layers, the high-frequency signal return is directly coupled to the ground planes and thus connected to the ground shields of the connector. To quantify the impact of the discontinuity in the high-frequency return path, measurements were performed on

Table 4 Total coupled noise as distributed from all adjacent signal lines.

	<i>Measured coupled noise (mV)</i>	
	<i>Near end</i>	<i>Far end</i>
Case 1 Figure 5(a)	74	14
Case 2 Figure 5(b)	67	16
Case 3 Figure 5(c)	272	165

a special connector test card. The main effect that was detected was a changed signal-coupling behavior. The results are summarized in **Table 4**.

In comparison, the near-end and far-end coupling are nearly equal for Cases 1 and 2, while the coupling for Case 3 is significantly increased. In the system design, a floating connector pin as depicted in **Figure 5(c)**, Case 3, is not allowed, but a connector voltage pin connected to a voltage plane on the top or bottom of a card far away from the internal signal seems to be nearly floating for the high-frequency return current. Therefore, if the high-frequency return current has to change reference planes, this must happen close to the signal. The internal plane capacitance of adjacent voltage/ground planes is most efficient for that purpose. All card cross sections and connector shields in this system were assigned to power and ground with respect to this demand.

Power integrity

Appropriate design of the complete power delivery system, including the voltage regulator, processor board, module, and chips with decoupling capacitors at all levels (hierarchical decoupling), is to keep the voltage at circuit level within the specified range of $1.2 \text{ V} \pm 100 \text{ mV}$ for this processor cage. The design has been supported and checked with early calculations and simulations which require different tools and granularity of the power delivery system models in the range between dc and clock frequency. The main emphasis in this section is the power integrity in the mid-frequency range (1 MHz–100 MHz). The goal was to keep the mid-frequency noise below 65 mV to maintain the $\pm 100\text{-mV}$ tolerance. Power noise in the mid-frequency range is caused by sustained changes in chip activity for more than one cycle which occur during system power-on and clock start, self-test, and normal system operation owing to simultaneous switching output drivers, logic switching activity variations, and varying clock gating to macros for reduction of power dissipation.

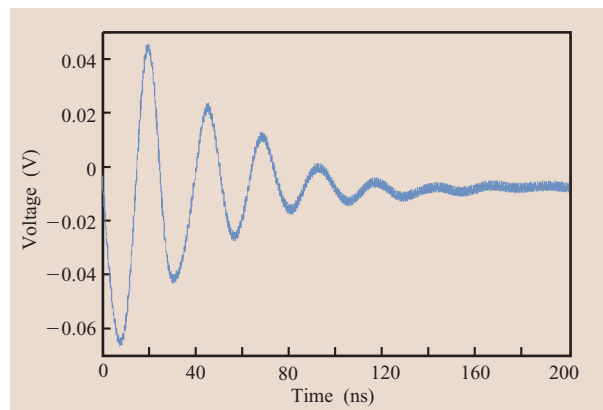


Figure 6

Change of supply voltage on the corner PU chip (PU7) after a 20% change of on-chip switching activity at time zero (SPEED2000 simulation).

The placement of the appropriate decoupling capacitors is essential for power noise containment in all three cases. Moreover, a staged power-on sequence is used, and dead cycles are introduced to reduce the power noise during system power-on and self-test.

The hierarchical decoupling for this processor cage consists of the 713 nF of on-chip capacitance on the processor chip, which includes the intrinsic capacitance plus added on-chip decoupling capacitance and 164 low-inductance 265-nF Low Inductance Capacitor Array (LICA**) capacitors on the MCM module; on the processor board there are 1,500 1- μF and 404 10- μF ceramic capacitors and 145 4.7-mF electrolytic capacitors.

The simulation program SPEED2000 from Sigrity, Inc. [11] was used to analyze the mid-frequency noise and optimize the hierarchical decoupling. The package model included the power-delivery system of the board and MCM module with capacitors, and with current sources, on-chip capacitors, and shunt resistors at the module surface to model the on-chip continuous logic switching and the logic switching activity variations on the chips. A 20% change in the switching activity (9.2 A for each dual-core PU chip, 91.8 A total for all chips on MCM) was assumed for the simulations owing to chip and system operation analysis. **Figure 6** shows the change of the supply voltage on the corner PU (chip PU7 in Figure 2). The voltage change exhibits a damped mid-frequency oscillation with 42 MHz and 65-mV peak, which is just within the specified range for the mid-frequency noise. A detailed sensitivity analysis [12] has shown that the noise peak is determined for each chip on this MCM by the chip ΔI , the on-chip capacitance and the inductance loop from

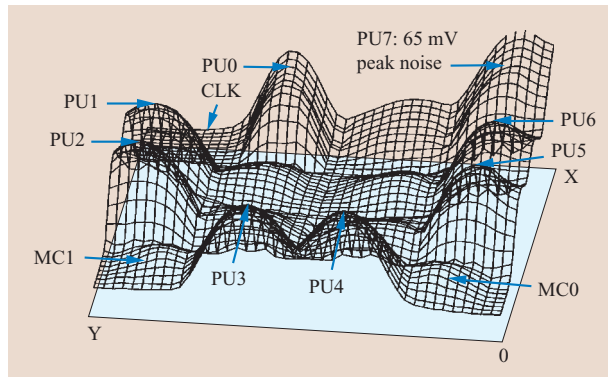


Figure 7

Two-dimensional peak noise distribution for the top power/ground planes of the MCM after a 20% change of on-chip switching activity at time zero (SPEED2000 simulation).

the on-chip to the nearest on-MCM capacitors. This loop inductance affects the mid-frequency oscillation period and determines the effectiveness of the on-MCM capacitors.

Figure 7 shows the two-dimensional peak noise distribution for the top power/ground planes of the MCM. PU7 (Figure 2) has the largest mid-frequency noise because this chip has module capacitors at two chip edges only, which increases the loop inductance [12]. The module capacitor locations are clearly indicated by the low peak noise in Figure 7. However, placing more module capacitors at PU7 was not possible for the present module technology because of area and timing restrictions. Therefore, increasing the on-chip capacitance from 220 nF for the PU chip in [1] to 713 nF in this processor cage was essential for noise reduction.

Summary

In this paper we have described the packaging of the z990 system. The achievement of a factor of $2.6\times$ increase in the processor volumetric density while at the same time providing a $2.75\times$ cost-performance improvement is the result of this effort. The cost part of the cost-performance improvement was significantly assisted by leveraging the improved glass-ceramic ground rules and dual-core processor chips to provide an MCM at half the size of the z900 MCM while removing the additional steps of adding the thin-film wiring layers on the top of the MCM as used in the five previous generations of CMOS zSeries servers.

The low-loss printed-circuit-board technology and advanced connector technology were developed to support the signal and power distribution demands of the 64-processor-unit z990 design.

Acknowledgment

The authors would like to thank Roland Frech for performing the VHDM connector measurements and for providing the coupling data in Table 4.

*Trademark or registered trademark of International Business Machines Corporation.

**Trademark or registered trademark of Teradyne, Inc. or AVX Corporation.

References

1. H. Harrer, H. Pross, T.-M. Winkel, W. D. Becker, H. I. Stoller, M. Yamamoto, S. Abe, B. J. Chamberlin, and G. A. Katopis, "First- and Second-Level Packaging for the IBM eServer z900," *IBM J. Res. & Dev.* **46**, No. 4/5, 397–420 (July/September 2002).
2. J. S. Corbin, C. N. Ramirez, and D. E. Massey, "Land Grid Array Sockets for Server Applications," *IBM J. Res. & Dev.* **46**, No. 6, 763–778 (November 2002).
3. G. F. Goth, D. J. Kearney, U. Meyer, and D. W. Porter, "Hybrid Cooling with Cycle Steering in the IBM eServer z990," *IBM J. Res. & Dev.* **48**, No. 3/4, 409–423 (May/July 2004, this issue).
4. J. C. Parrilla, F. E. Bosco, J. S. Corbin, J. Loparco, P. J. Singh, and J. G. Torok, "Packaging the IBM eServer z990 Central Electronic Complex," *IBM J. Res. & Dev.* **48**, No. 3/4, 395–407 (May/July 2004, this issue).
5. G. A. Katopis, W. D. Becker, T. R. Mazzawy, H. H. Smith, C. K. Vakirtzis, S. A. Kuppinger, B. Singh, P. C. Lin, J. Bartells, Jr., G. V. Kihlmire, P. N. Venkatachalam, H. I. Stoller, and J. L. Frankel, "MCM Technology and Design for the S/390 G5 System," *IBM J. Res. & Dev.* **43**, No. 5/6, 621–650 (September/November 1999).
6. P. Singh, S. J. Ahladas, W. D. Becker, F. E. Bosco, J. P. Corrado, G. F. Goth, S. Iruvanti, M. A. Nobile, B. D. Notohardjono, J. H. Quick, E. J. Seminaro, K. M. Soohoo, and C. Wu, "A Power, Packaging, and Cooling Overview of the IBM eServer z900," *IBM J. Res. & Dev.* **46**, No. 6, 711–738 (November 2002).
7. J. U. Knickerbocker, F. L. Pompeo, A. F. Tai, D. L. Thomas, R. D. Weekly, M. G. Nealon, H. C. Hamel, A. Haridass, J. N. Humenik, R. A. Shelleman, S. N. Reddy, K. I. Prettyman, B. V. Fasano, S. K. Ray, T. E. Lombardi, K. C. Marston, P. A. Coico, P. J. Brofman, L. S. Goldmann, D. L. Edwards, J. A. Zitz, S. Iruvanti, S. L. Shinde, and H. P. Longworth, "An Advanced Multichip Module (MCM) for High-Performance UNIX Servers," *IBM J. Res. & Dev.* **46**, No. 6, 779–804 (November 2002).
8. E. Cordero, F. Ferriolo, M. Floyd, K. Grower, and B. McCredie, "A Synchronous Wave-Pipeline Interface for POWER4," presented at the IEEE Computer Society HOT CHIPS Workshop, Stanford University, California, August 15–17, 1999.
9. P. Mak, M. A. Blake, K. W. Kark, V. K. Papazova, A. E. Seigler, G. E. Strait, G. A. Van Huben, L. Wang, and G. C. Wellwood, "Processor Subsystem Interconnect Architecture for a Large Symmetric Multiprocessing System," *IBM J. Res. & Dev.* **48**, No. 3/4, 323–337 (May/July 2004, this issue).
10. Z. Chen, W. D. Becker, and G. Katopis, "A Fast Simulation Method for Single and Coupled Lossy Lines with Frequency-Dependent Parameters Based on Triangle Impulse Responses," *Proceedings of the 8th IEEE Topical Meeting on Electrical Performance of Electronic Packaging*, San Diego, October 25–27, 1999, pp. 257–260.
11. Sigrity, Inc., SPEED2000, available online at <http://www.sigrity.com/>.

12. B. Garben, G. A. Katopis, and W. D. Becker, "Package and Chip Design Optimization for Mid-Frequency Power Noise Decoupling," *Proceedings of the 11th IEEE Topical Meeting on Electrical Performance on Electronic Packaging*, Monterey, CA, 2002, pp. 245-248.

Received October 27, 2003; accepted for publication April 22, 2004; Internet publication June 7, 2004

Thomas-Michael Winkel *IBM Systems and Technology Group, IBM Deutschland Entwicklung GmbH, Schoenaicherstrasse 220, 71032 Boeblingen, Germany (winkel@de.ibm.com).* Dr. Winkel received his Diploma in electrical engineering in 1989 and his Ph.D. degree in 1997 from the University of Hannover, Germany. His research activities covered the area of characterization and modeling of on-chip interconnects using high-frequency measurements. In 1996 he joined the IBM development laboratory in Boeblingen, Germany. He is currently an Advisory Engineer in the IBM Server Group in Boeblingen, leading the electrical design team for second-level packaging. Dr. Winkel's current focus is electrical packaging design with respect to high-frequency signal distribution, power noise, and system aspects for the pSeries and zSeries. He is also interested in high-frequency on- and off-chip measurements and modeling of signal as well as power and ground lines. He has published multiple papers and holds four patents.

Wiren D. (Dale) Becker *IBM Systems and Technology Group, 2455 South Road, Poughkeepsie, New York 12601 (wbecker@us.ibm.com).* Dr. Becker received his Ph.D. degree from the University of Illinois. He is currently a Senior Technical Staff Member in the IBM Systems and Technology Group, leading the package design team that integrates and implements the multiprocessor design for IBM's large server platforms. Dr. Becker's current interests focus on robust electrical design methodologies for a high-frequency CMOS processor system. He specializes in the application of electromagnetic numerical methods to the issues of signal integrity and simultaneous switching noise in electronic packaging, the measurement of these phenomena, and the verification of the models. He has published four papers in the *IBM Journal of Research and Development*, as well as more than 50 conference and journal papers. Dr. Becker has reached the second invention plateau at IBM.

Hubert Harrer *IBM Systems and Technology Group, IBM Deutschland Entwicklung GmbH, Schoenaicherstrasse 220, 71032 Boeblingen, Germany (hharrer@de.ibm.com).* Dr. Harrer is a Senior Technical Staff Member working in the IBM Server Group. He received his Dipl.-Ing. degree in 1989 and his Ph.D. degree in 1992 from the Technical University of Munich. In 1993 he received a DFG research grant to work at the University of California at Berkeley. Since 1994 he has worked for IBM in the Boeblingen Packaging Department. In 1999 he was on international assignment at IBM Poughkeepsie, New York. Dr. Harrer's interests currently focus on zSeries first- and second-level packaging design for the IBM Server Division. He has published multiple papers and holds six patents in the area of first-level and second-level packaging.

Harald Pross *IBM Systems and Technology Group, IBM Deutschland Entwicklung GmbH, Schoenaicherstrasse 220, 71032 Boeblingen, Germany (hpross@de.ibm.com).* Mr. Pross is an Advisory Engineer working in the IBM Server Group. After joining IBM in 1975, he studied applied physics at the Fachhochschule Heilbronn from 1979 to 1984, graduating with a B.S. degree in physics in 1984 and joining the IBM Boeblingen laboratory that same year. He held various technical positions in S/390 processor packaging design in Boeblingen. In 1993 he moved to Rochester, Minnesota, to lead the card/board packaging design of the new high-end

AS/400 multiprocessor system. Returning to Boeblingen in 1994, he worked on various printed wiring board designs for S/390 processors. In 1997 Mr. Pross became the project leader for the Boeblingen processor card and board designs. He led the development of the zSeries 900 processor cage in 2000 and has worked on packaging for the zSeries 990 processor. He is currently responsible for packaging implementation of the future high-performance, high-availability pSeries processor card and board development.

Dierk Kaller *IBM Systems and Technology Group, IBM Deutschland Entwicklung GmbH, Schoenaicherstrasse 220, 71032 Boeblingen, Germany (dkaller@de.ibm.com).* Mr. Kaller is an Advisory Engineer working in the IBM Server Group. He received his Diploma in electrical engineering in 1995 from the University of Hannover, Germany, joining the IBM development laboratory in Boeblingen that same year. Since then he has worked on various S/390 packaging designs in packaging development. He is currently responsible for the signal integrity of the first- and second-level packaging of the zSeries and pSeries high-end server.

Bernd Garben *IBM Systems and Technology Group, IBM Deutschland Entwicklung GmbH, Schoenaicherstrasse 220, 71032 Boeblingen, Germany (garbenb@de.ibm.com).* Dr. Garben is an Advisory Engineer in the IBM Server Group. He received his Ph.D. degree in physics from the University of Goettingen, Germany, in 1972. He joined the IBM Boeblingen Laboratory in 1974 and was on assignment at IBM East Fishkill from 1982 to 1984. Dr. Garben has worked in the areas of physical materials analysis, bipolar and CMOS semiconductor technology development, CMOS microprocessor failure analysis, and packaging development; he holds seven patents in those areas. His current field of interest is simultaneous switching noise analysis in electronic packaging.

Bruce J. Chamberlin *IBM Integrated Supply Chain Division, 1701 North Street, Endicott, New York 13760 (chamberb@us.ibm.com).* Mr. Chamberlin is a Senior Engineer in the IBM Integrated Supply Chain Division. He holds a B.S. degree in mechanical engineering from Clarkson University. Mr. Chamberlin holds six patents, three of which are utilized in production zSeries machines. He has been the product engineer for Clark raw boards and for the Clark board assembly, and the program manager for development and qualification of high-end PWBs supporting zSeries machines. He currently works in the Electronic Materials/Processes Department within Procurement, responsible for support of the high-end iSeries, pSeries, and zSeries circuit boards and assemblies.

Scott A. Kuppinger *IBM Systems and Technology Group, 2455 South Road, Poughkeepsie, New York 12601 (skuppinger@us.ibm.com).* Mr. Kuppinger is an Advisory Engineer in the Electronic Packaging and Integration Department at the IBM Poughkeepsie development laboratory. He joined IBM in 1989 after receiving a B.S. degree in electrical and computer engineering from Clarkson University.