

Preface

The introduction of the new zSeries® flagship mainframe z990 in June 2003 marked a significant milestone in the IBM e-business on demand™ initiative. Although it was built on the e-business legacy of its zSeries predecessor, the IBM eServer™ z900, the z990 server was substantially enhanced to meet the unique computing challenges presented by e-business on demand.

Particular focus was placed on the overall system design to overcome system constraints within existing zSeries servers and in packaging concepts, allowing for a more modular base manufacturing cost (BMC) structure.

The microprocessor was completely redesigned in order to introduce a superscalar microarchitecture, with several enhancements to run new workloads such as zLinux and WebSphere® faster on the zSeries platform. In addition, the z990 introduced the IBM 130-nm CMOS SOI (silicon-on-insulator) technology into the microprocessor subsystem. As a result, the uniprocessor performance improved more than 50% compared with the predecessor system.

Packaging density increased significantly, allowing for up to 32 processors running in one system and tripling the overall system capacity. System configurations range from one active processor up to 32, giving a customer an unprecedented range of capacity growth within one system. All configurations can be concurrently upgraded to the maximum system size. To enable this flexibility, the z990 system design changed toward a nodal concept, with up to four processor books being connected by means of an interconnect ring. Each processor book contains up to eight customer-usable processors, two spare processors, two I/O engines, and 64-GB memory.

The I/O subsystem was greatly enhanced to cope with the requirements of a truly e-business on demand work environment as well as the increased processing power.

The z990 is the first zSeries server with more than 256 I/O channels. To achieve this density, up to four logical channel subsystems (LCSSs) with 256 channels each are designed into the system. Channels can be shared among the individual LCSSs to allow maximum flexibility. In addition, the number of attachable I/O device addresses has quadrupled.

The number of logical partitions (LPARs) has doubled, to a maximum of 30 partitions. The Ethernet family of adapters received a new member with the addition of the new 10-Gb Ethernet channel, and the packaging density of FICON® channel cards was doubled to increase the number of FICON ports. This substantial upgrading of I/O capacity required the redesign of the data link between the processor subsystem and the I/O subsystem. The capacity of each link was doubled to 2 GB/s, and the number of links was increased to 12 per processor book, totaling 48 for a fully configured system. In addition, the new link can be used as an enhanced integrated cluster

bus (ICB) for high-speed coupling between different z990 systems.

Heightened security demands required us to redesign the existing cryptographic adapters and introduce a new flexible cryptographic I/O card called xCrypto with more performance and flexibility for upcoming encryption requirements. The design of the card has been kept flexible so that it can easily be integrated into other IBM eServers.

However, e-business on demand is not just about performance and bandwidth. Flexible system capacity and continuous availability are base requirements. The z990 is setting new standards in this category. In past zSeries servers, IBM introduced capabilities to temporarily increase the processor capacity for disaster recovery or workload peaks. In the z990 this concept was enhanced with the introduction of OoCoD (on/off capacity on demand), which allows customers to concurrently increase and decrease the number of processors in a fully flexible manner according to their business needs. In the rare event of a processor book or memory card failure, the failing hardware can be exchanged while the remainder of the system continues its operation. Special effort was put into the design quality of the various components by enhancing existing simulation concepts, from the smallest transistor cell simulation to complete system emulation.

This double issue of the *IBM Journal of Research and Development* contains 23 papers that describe many aspects of the z990 server, ranging from its architecture and design to its leading-edge implementation of several autonomic computing concepts.

The first paper, by Slegel et al., describes the microprocessor design with its new microarchitecture. The conceptual design of the processor comprises several elements such as superscalar architecture, pipelining, and partial out-of-order command execution, making it an ideal vehicle for running modern applications as well as legacy solutions on the z990 system.

One key element of the new processor, the floating-point unit, is outlined in a separate paper by Gerwig et al. The introduction of the nodal design required a complete redesign of the existing zSeries cache and memory hierarchy. The paper by Mak et al. summarizes this sophisticated concept and its realization. Each processor book has its own clock chip, which not only provides an accurate clocking path to the various processor and cache chips, but also functions as the interface to an attached service processor. With multiple clock chips within one system, a unique run-control concept is required; this is described in the paper by Webel et al.

Special focus was placed on the simulation of the processor subsystem. The paper by Bair et al. describes the challenges and the concepts of the processor subsystem simulation. A key challenge for a successful

hardware system simulation is the exact representation of the hardware in the simulation model. The next paper, by Anderson et al., outlines the unique challenges and solutions to accurately reflect the hardware in simulation.

Two papers describe the new z990 packaging. Higher processor speeds and denser design packages require a thorough electrical analysis. Winkel et al. illustrate the concepts and results of their analysis as well as the implications for the logic card packaging. In the paper by Parrilla et al., the electromechanical challenges of packaging the CEC cage are presented. The scheme for connecting the land grid array (LGA) multichip module (MCM) to a printed wiring board and the mechanical isolation of the MCM evaporator/heat sink mass from the LGA contacts are discussed.

Increases in package density and eServer performance require additional cooling. The IBM zSeries invested early in closed-loop cooling concepts, and the implementation for the z990 system is an example of one of the most advanced cooling systems in commercial computing. The concept is described in the paper by Goth et al.

The firmware plays a pivotal role in ensuring flawless processor operation. The paper by Heller and Farrell outlines the unique design elements of a modern processor code. High-end servers such as the z990 are attached to external service processors to load code into the machine and allow service personnel to interact with the system. The interface between the service processor and the system itself must be fast, reliable, and fail-safe. The paper by Axnix et al. describes the z990 service processor interface.

Several papers outline the new I/O subsystem of the z990 server. The paper by Chencinski et al. documents a generic overview of the new I/O hardware, while the paper by Hoppe et al. describes the unique challenges and solutions of the z990 I/O hardware simulation. Cryptographic functionality is becoming increasingly important, requiring redesign of the existing encryption/decryption concepts. The paper by Arnold and Van Doorn describes this new xCrypto chip design. The new LCSS concept is a breakthrough in zSeries I/O bandwidth. The implementation was made very flexible to allow for future enhancements. The paper by Wyman et al. describes this new concept, including the changes for the operating system.

With the introduction of zLinux, the requirements for SCSI (small computer system interface) channels grew. The paper by Banzhaf et al. explains the SCSI implementation for the z990 server.

Reliability, availability, and serviceability (RAS) remain a cornerstone of each zSeries server. Like its predecessor, the z990 sets the standard in server availability; its high level of error detection and correction as well as its unique recovery capabilities make it the platform of

choice for critical IT environments. The paper by Fair et al. offers an overview of the z990 RAS capabilities.

Siegel et al. address the new logical partition design of the z990. More partitions and more hardware resources, such as 32 processors and four channel subsystems, require several enhancements to achieve maximum configuration flexibility.

The extension of the z990 firmware address space beyond 2 GB requires a 64-bit compiler for the PL8 firmware code language. The paper by Gellerich et al. illustrates the transformation of PL8 which enabled use of the 64-bit GNU compiler for the z990 firmware development.

For fast error analysis, it is crucial to collect and store sufficient information on failure occurrences for later analysis and, ultimately, solution of the problem. A concept called first error data capture (FEDC) guarantees that necessary error information will be collected and made available to the development team. The FEDC concept is described in the paper by Koerner et al.

The last two papers describe the tools and methods used for system and firmware simulation. The paper by Schubert et al. presents an overview of the hardware and firmware integration on a hardware emulator. A description of the unique zSeries firmware simulation environment, called CECSIM, which enables immediate testing of firmware designs, is available in the paper by Stetter et al.

Finally, I thank the authors in the IBM Systems and Technology Group and in IBM Research who have taken the time to document their outstanding accomplishments.

Rolf Schmidt
z990 Program Manager
IBM Systems and Technology Group

Guest Editor