System performance management for the S/390 Parallel Enterprise Server G5

by R. F. Rizzolo
G. Hinkel
S. Michnowski
T. J. McPherson
A. J. Sutcliffe

System performance management is a broad category of techniques that cover all aspects of obtaining maximum performance or speed from a given design. Items such as sorting methodology, critical path improvements, semiconductor line optimization, power-supply optimization, clock tuning, and cooling are part of performance management. Logic or architecture improvements that have a major effect on design, such as increasing cache size, are usually not included.

Introduction

The strong performance of today's S/390* systems is to a large extent based on the development and implementation of enhancements resulting from the system performance management techniques described in this paper. We estimate that the implementation of these techniques has increased performance by more than 35% over that of an unsorted and unoptimized design in the same semiconductor technology. The techniques employed are wide-ranging, covering design, test, and sorting. Continuous planning and communication among the design, semiconductor process, burn-in, and test teams was vital to ensure timely implementation and adequate manufacturability.

Mainframe design in the 1980s was completely based on bipolar transistor technology. Bipolar processing was considered to have less than 20% variation from the fastest to the slowest chips. In addition, integration levels were low. Most critical (system-performance-limiting) paths were "wire-dominated" paths extending long distances. Timing predictions were based on worst-case process models. There was no cycle-time differential between models of a given machine—chip sorting was not done and probably would not have been very effective. The extent of performance management was reaction to unexpected empirical test results. The final cycle time was based on testing a sample of 20 to 30 machines. As long as that cycle time was faster than the predictions, it was considered to be no problem. There was a dichotomy—a three-sigma worst-case design was done and predicted a "slow" cycle time, but shipment was based on 20 to 30 machines and performance was significantly faster.

Performance management techniques originated in the early 1990s on the last generations of bipolar mainframe systems. Given the limited time between design functional

0018-8646/99/$5.00 © 1999 IBM

fixes, a "cycle-time reduction" team was formed to identify and fix as many critical paths as possible before the next chip/module/board release. Tasks included defining test requirements, investigating and fixing empirical performance-limiting paths, correlation with timing predictions, and root-cause determination of any differences. It was observed, as on previous machines, that the performance-limiting paths were not predicted by the timing tool, and delay was predominantly due to off-chip wiring. Because these paths were empirically discovered, it was too late to "fix," or reduce the delay of, many of them. Almost all could have been fixed if they had been anticipated. The root cause of the problem was three-sigma process modeling. Machines were faster than the timing tools predicted; however, unanticipated paths dictated the cycle time. These paths were primarily interconnections between chips which could easily have been fixed in the design phase if they had been anticipated. The result of this conservative philosophy was, paradoxically, a machine significantly slower than it could have been. In more recent development programs, the following design and methodology changes have been implemented:

- Less conservative process models are used in the design phase (one sigma slow was used then, nominal now).
- Programmable clocks have been added to the chip design to identify and correct timing problems.
- Differences between timing simulation and actual hardware are expected and planned for.
- All significant differences between predictions and actual hardware must be explained and accounted for in subsequent designs.
- Early parts to be tested must represent a wide process distribution, sufficient to extrapolate to a full process sample.
- Chips and MCMs are all tested to failure to determine margins.
- A cycle-time reduction team is formed for each project to manage these activities.

The advent of CMOS brought higher integration levels and produced a process variation of more than double that of bipolar technology. A "nominal" design actually could produce individual chips that are 25% slower than the nominal, and since most of the function was now on a single chip, the entire system would be proportionately slower. A chip-sort methodology was clearly needed. Timing correlation experiments determined that the existing on-chip performance indicator, a small ring oscillator, was inadequate. Frequency did not correlate well with chip performance, and the characteristics of circuits in the ring did not match those in a typical critical path.

In 1995, the IBM S/390 design and technology teams jointly developed an elegant CMOS performance-sorting methodology which allowed chip sorting at the wafer level. The delay through a portion of the latch scan chain in so-called "flush mode" was shown to correlate extremely well with the chip's performance in the system. All S/390 chips use scan-based design and test, and no special sort structures have to be added to the design. Not only did this sorting methodology help sidestep a major perceived disadvantage of CMOS, but it also allowed for faster-than-nominal chip sorts to be practical. It is important to note that this sort measurement does not use a "functional" test nor a "structural" test. A series of meetings was also held between the S/390 system developers and the technology development group to evaluate and negotiate semiconductor process improvements. This led to the first commitment by IBM Microelectronics to introduce CMOS performance sorting. Agreements were struck to improve the process while committing to "buy all" of the manufactured chips. Additional system models with different cycle times (including a high-end "turbo") were offered that fully utilized all chip speeds. The resulting improvement in the performance of the S/390 G3 was more than 20% over that of an unsorted machine.

Over the past two generations of CMOS mainframes, G4 and G5, these testing, methodology, design, and sorting techniques have been combined and further refined. These techniques, combined with robust circuit design and thorough design for testability, have contributed to the dramatic performance improvements seen in the last two years [1, 2].

Design criteria

During the design phase of each chip in the S/390 G5 system, certain value distributions are assumed for process parameters, chip temperature, and circuit voltage. (There are a number of other parameters such as capacitance and noise that are difficult to calculate and also affect circuit performance.) Finally, the tools that are used for performance prediction have accuracy limitations, with associated guard bands. On the basis of these assumptions and calculations, a cycle time is chosen as a design point. This is the fundamental clock period for the chip being developed and, generally speaking, represents the time limit for data to propagate from one state latch to another state latch (see **Figure 1**).

Extensive test characterization and diagnostic work have shown that actual physical chips can have speeds significantly different from predictions, and what limits the cycle time is often different from what was expected. This is due both to timing tool inaccuracy and to the process spread around the timing tool design point. Timing simulation is generally accurate to within 5%. A 5% cycle-time improvement, however, is significant, and once chips

arrive there is an intense effort not only to verify functionality but to maximize performance by adjusting voltage, temperature, process—in fact, whatever variable can be adjusted in the short, several-month functional evaluation period. The extent to which these variables are adjusted depends on existing design margins, how quickly changes can be made, and the ability to change each parameter. It is fundamentally an empirical, iterative process because of the limitations of simulation and modeling. A major part of performance optimization plans for these iterations.

Performance improvement is not automatic. If iterations are not planned for, cycle time may be "locked in" by long-lead-time items that would have been simple to change if only they had been anticipated. A simple example would be a system power supply with a 2.1-V maximum specification based on process models, power estimates, and cooling estimates given early in the design cycle. If the chips and the cooling system empirically support 2.3 V, the cycle time could be locked in by the power supply and not the chips—a very undesirable situation. Since these methods are empirically based, it is crucial that a representative sampling of chips be used in the evaluation. A major activity of performance management is the sorting methodology. The S/390 G5 chips use CMOS technology, and as such have significant performance variation from chip to chip. An effective sort methodology separates the population into speed "bins," each of which supports manufacturable volumes. It also addresses controlling certain process parameters to maximize yield and deliver samples of these bins early in the empirical evaluation process.

As shown in **Table 1**, there are six major chip types used in the S/390 G5. The design target for this system was 5.0 ns, with the processor running twice as fast at 2.5 ns. Chip timing was calculated using static timing tools. A static timer does not model logic function, so some predicted critical paths may be "false" (logically impossible). Timing simulation relies on process and parasitic models for accuracy. However, simplified models are used in order to keep simulation times practical. It is important to understand the philosophy of design—models

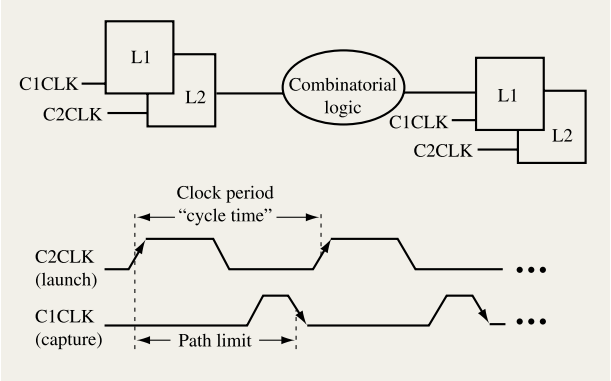

Figure 1

Typical latch-to-latch path and cycle-time definition.

are set up such that it is reasonably certain that the actual product will not be slower than the design-point cycle time. The cycle time predicted by the timing tool is used as a commitment to the business.

A number of factors enter into the accuracy of timing tool predictions. Some of these are observable and controllable, and some are not. The most important are

- Process models.
- Capacitive and resistive loading on nets.
- Voltage at the circuit.
- Clock skew.
- Temperature at the circuit.
- Electrical noise.
- Relative timing of switching inputs.

Expected ranges are given for these parameters in the design phase, but any uncertainties are usually biased on the conservative side—that is, they will overpredict delays. As an example, the calculated circuit temperature is based on chip power (with certain assumed switching factors), cooling capability, heat sink material, and thermal paste material. Because the design of the thermal

Table 1 Characteristics of S/390 G5 CMOS chip types.

Chip	*Quantity (fully configured)*	*Function*	*Process*	*Transistors* (million)	*Target cycle* (ns)
CLK	1	System clock	CMOS 6S2	1.4	5.0
MBA	4	Bus adapter	CMOS 6S2	5.9	5.0
SC	2	L2 control	CMOS 6S2	8.6	5.0
SD	8	L2 cache	CMOS 6S2	59	5.0
CP	12	Processor	CMOS 6X	25	2.5
CRY	2	Encryption	CMOS 5X	1.8	>5.0

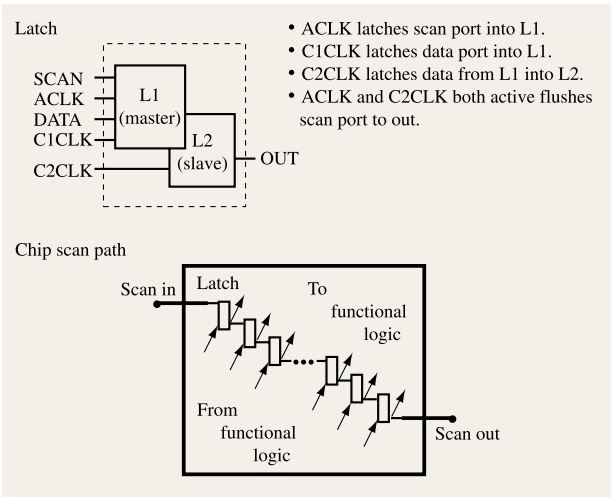

Figure 2

Basic latch and scan path.

system is done in parallel, not all information may be available. An aluminum heat sink may have been the plan of record, for instance, but a copper one might be chosen at the end. Cost and vendor supply agreements are not closed at the time the design is being done. The cooler temperature (which improves chip performance) is thus not considered in the design phase.

Additional design practices enable us to empirically optimize cycle time more effectively. The latches on the G5 processor have many independently programmable clocks; that is, the latch-capture and launch clocks can be delayed in multiple increments to allow both hardware diagnosis and performance improvement. The critical paths can be determined in hardware by a combination of the failure state of the processor, timing simulation, and hardware path stressing. For example, assume that the processor clock rate is increased until a failure occurs, and the error data indicates that an instruction was being decoded at the time of failure. The timing reports can be analyzed to determine the timing margin of the instruction decode path, and a final hardware proof of the failing path can be done by programming clocks to delay the launch clock of the instruction decode path. Empirical determination of specific critical paths provides designers with the option of further optimizing the design critical paths as schedule allows. Also, the timing tools can be correlated with the actual hardware. Another benefit of the G5 programmable clocks is hardware performance improvement. After the critical paths have been determined in hardware, the critical path-capture clocks are delayed to allow more time for the data to arrive. Only the capture clock is delayed; the launch clock remains fixed, resulting in capture/launch-clock overlap. This effectively puts the latch in flush mode and allows cycle stealing on the critical paths. The clock skew penalty is eliminated with overlapped clocks. The G5 cache and memory control chips also have programmable clocks for performance optimization and empirical margin determination.

The most critical chip is the central processor (CP), because it must operate at twice the speed of the other chips. Since sorting was planned for, a design-for-sorting strategy is applied to the processor chip. A chip with a silicon-dominated critical path will achieve greater improvement from sorting. During the timing-reduction design phase, wire-dominated paths are "over-fixed" to have more timing margin than silicon-dominated paths. Wire-dominated paths with the same timing margin as silicon-dominated paths may become the hardware-critical path on sorted chips, since sorting does not improve wire delay.

Chip performance estimation

Chip timing predictions are used as commitments to the business, and as such become minimum requirements. Much attention is focused on how the hardware compares to the timing tool prediction. The chip timing analysis is performed at the transistor level, and the G5 hardware was within 5% of the timing tool prediction when temperature, voltage, and other parameters were at the assumed values. The memory and cache control chips use a gate-level timing methodology which is typically more conservative. The timing tool cycle time target contains a guard band to account for manufacturing test margin and design uncertainties such as phase-locked-loop (PLL) jitter and clock skew. For example, if the product is required to run at 2.5 ns, the design timing tool goal would be near 2.3 ns. As discussed before, the timing tool calculations usually have built-in conservatism for good reason. However, "what-if" calculations are needed in order to predict what the cycle time would be if parameters were more on the favorable side and what actions it would take to achieve that goal. We call this the potential performance analysis. This cycle time is usually much faster than that provided by the timing tool, and there is no guarantee that such a cycle time is achievable. The advantage of this approach is that specific items that *may* limit the cycle time are identified. These items are then confirmed as fundamental limits, or are improved so that they no longer represent a limit. Since the processor chip is the most complex design and must run at twice the speed of the other chips, it is expected to be the only critical path in the system. Most of the performance-tuning effort is concentrated on optimizing chip performance and, thus, system performance. Typically,

(a) (b)

Figure 3

Physical location of latches in different scan paths on the same chip: (a) Good chain for sorting (boundary-scan chain); (b) poor chain for sorting.

minimal sorting is required on memory and cache control chips.

Sorting methodology

All S/390 chips are based on a full-scan design in which every latch is controllable and observable through scan ports on the chip. Latches are connected serially by a scan path and are clocked serially by scan clocks (see **Figure 2**).

Chips are sorted for performance on the basis of a "flush" delay measurement through a series of latches in the scan chain of each chip. Scan clocks are held in their active state, and a data transition on the chip scan-in port "flushes" through the chain to the scan-out port. With this configuration, special sort structures such as recirculating loops are not needed.

The sort-chain selection process is one of the key elements of CMOS performance sorting. The scan chain is configurable and can be set as one long chain or as multiple smaller chains, depending on test needs [3]. More than ten chains are available to use for sorting. During the bring-up of the S/390 G2 and G3 systems, three scan chains each on several thousand chips were tested at wafer final test, and the chip's performance was then determined in the system. Nine different chip types on two different carriers were analyzed. The best chains yielded a sorting precision of ±3%, the worst chains ±8%. **Figure 3(a)** shows a good chain for sorting, **Figure 3(b)** a poor chain for sorting. Sorting precision is the accuracy with which a chip's performance in the system can be predicted on the basis of the flush delay tested at wafer final test. The following sort-chain criteria resulted from this large database:

1. The chain must be evenly distributed over the entire chip area.
2. The latches in the chain must be "standard latches"; i.e., no latches in optimized macros such as "growable" register arrays or SRAMs.
3. There must be several different latch types in the chain, with a statistically significant number of latches per type (>50).
4. The delay through the chain (in a flush mode) should be of the order of one microsecond or more, to avoid precision loss during test (typical tester precision is ±3 ns).
5. Loading and wire lengths must be comparable to those of typical critical paths.

The chain chosen for the chips in the G5 system is the boundary-scan chain. Since its primary use is to isolate chip inputs and outputs for testing, it tends to be widely dispersed.

Huisman [4] has shown that performance correlation is improved when the path used for sorting covers a large area, because local variations average out. Our experience supports this. Both flush 1 and flush 0 propagation delays are measured at wafer level under controlled conditions of temperature and voltage.

The intent of establishing a sorting methodology is to coordinate model offerings and speed "binning" in a way

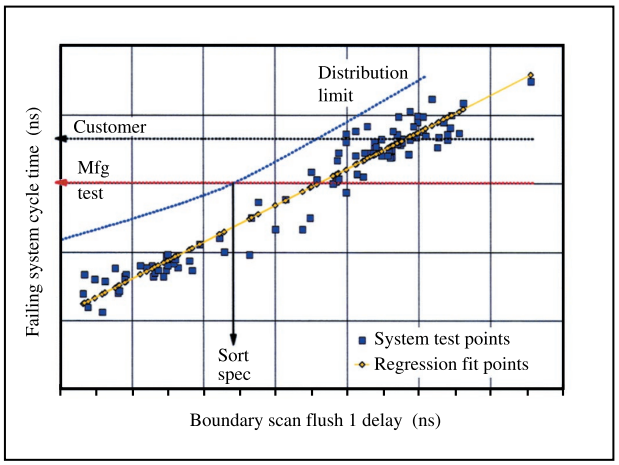

Figure 4

Typical correlation between flush delay and system cycle time.

that markets both aggressive and low-end machines with a manufacturable volume in order to use all of the chip-speed distribution. Chip yield at each sort specification must meet target objectives. To accomplish these goals, sort-point determination starts after specially constructed hardware has been system tested for failing cycle time for each chip.

Special sort hardware

Empirical evaluation of design margin and cycle-time potential is critically dependent on the samples that are evaluated. If only slow chips are tested, it is risky to project how fast chips will behave. Perhaps there will be different critical paths, perhaps there will be voltage sensitivity causing circuit malfunction, perhaps there is a process problem that sharply limits yield. A major effort of chip performance management is coordinating activities which manufacture and test hardware that is representative of the full process window and beyond, early in the evaluation process. The first process lots are built in a special way that varies the effective n-FET and p-FET gate lengths (L_{eff}) by the full process range across each wafer. This parameter is the major factor affecting circuit speed. Special lots are built with other process parameters that affect performance intentionally varied over the full specified range. Chips from these special "process-split lots" are taken through every level of the test process from wafer test through burn-in to system test. The intent is to

1. Ensure that there will be some yielding chips for functional debug.
2. Characterize product yields and margins over the full process window.
3. Understand the failures that occur and add design fixes as soon as possible.
4. Work around those failures until a "wall" is found.
5. Achieve process learning in controlling parameters that affect speed and yield.

In this way, if there are any chip engineering changes needed to correct functional problems, fixes to improve performance and yield will be combined in those changes.

The setting of the precise sort point to establish optimum system performance within a given chip's speed range requires the construction of special-sort multichip modules (MCMs). The building of the sort-point MCMs begins with selection of chips with a large flush-delay range, with concentration at a specific (usually fast) sort point. These chips are obtained from L_{eff}-split wafers. Sort-point MCMs contain chips representing multiple wafer lots to account for lot-to-lot variations. The intent is to develop a database of characterized MCMs. From this database, a trend line can be created that correlates system cycle time with chip delay performance across a wide distribution of chip delays.

The flush-path delays are from the boundary-scan chain on each chip. The flush delay closely correlates with the overall speed of the critical paths in the chip, which is imperative for predicting system cycle times. The flush delays also correlate with critical process parameters to predict nominal to ± 3 sigma of the process. After chip test, the flush-delay test results are logged into a database for subsequent use.

The balance of the non-sort chips on the sort MCM should include chips that have the fastest flush delays permitted by the wafer yields and process. The goal is for those chips selected for the sort point to be the cycle-time limiters and the first chips to fail during system cycle-time profiling.

Engineering test requirements

Every customer's application and environment are unique. Most large-system customers need twenty-four-hour-a-day availability, seven days a week; failures are intolerable. In manufacturing, systems are tested to conditions beyond those any customer would experience. These stresses include voltage, cycle time, and temperature. The goal of the manufacturing test is to protect the customer from end-of-life degradation, system clock variation, part-to-part normal distribution, and manufacturing test programs on the customer's application and power-supply regulation.

Engineering test requirements go beyond manufacturing. The extended goals are as follows:

1. Characterize all functional failure modes and diagnose root causes.

2. Test voltage, temperature, and cycle-time margins to failure and identify causes of failure.
3. Determine optimum clock placement and sensitivity.
4. Determine the cycle time and guard band.
5. Verify all manufacturing tests.

Engineering must test systems with a large test suite before release to manufacturing. This suite includes the manufacturing test programs, self-test, special architectural exercises, and customer operating-system programs such as MVS. Our focus is limited to failures that are a function of process, speed, voltage, or temperature. All G5 components (chips, MCMs, memory cards, etc.) are rigorously tested prior to assembly; however, test coverage is rarely perfect, and there are some failure modes that may be "system only." In addition, some parameters, such as cycle time, are very difficult to test at lower levels of assembly. The precise path may not be activated in a chip test, the voltage and temperature at the circuit may be different in the system, and switching noise may be different. If engineering test identifies significant design margin, there is an opportunity for cycle-time improvement.

The special sort-point MCMs go through rigorous engineering testing. Twenty-nine chips, including 12 processor chips, are mounted on an MCM, which is first tested in a full 12-processor configuration. Using a test program with essentially functional characteristics, cycle time is reduced until a chip fails (usually one of the processors). The failing path is examined to confirm that it is a known cycle-time-limiting path. After the test is complete, that chip is electronically removed from the configuration and the testing continues (with eleven processors if a processor chip failed). This is repeated until the system is reduced to the minimum configuration. The failing cycle times and the flush delay for each chip of the same chip type are entered into a database. The cycle times are then adjusted according to the MCM chip configuration at the time of failure, because lower configurations run cooler and have less voltage drop.

A regression analysis is applied to the system cycle-time and flush-delay database to determine the strength of the correlation. The regression analysis produces a "best-fit" trend line of cycle time versus flush delay. With knowledge of the target "ship" cycle time, the objective is to find the corresponding flush-delay or cut-point specification at which the chip will be speed-sorted. A guard band must be added to account for manufacturing test margin and the scatter of the cycle time vs. flush delay (see **Figure 4**). After the ship cycle time has been properly guard-banded, the corresponding flush delay can be extracted from the plot. This flush delay becomes the specification against which all chips will be sorted. The sort-point specification

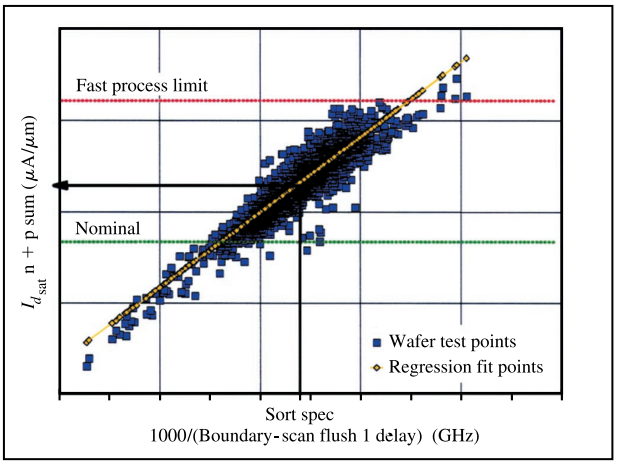

Figure 5

$I_{d_{sat}}$ vs. flush delay.

ensures a specific ship cycle time, and all flush delays less than the sort specification will also support the same ship cycle time.

Chip yield must support manufacturable volumes for each sort point. Yield is determined by correlating flush delay with a fundamental measurable process parameter called $I_{d_{sat}}$, which is in turn correlated with speed. $I_{d_{sat}}$ is the sum of the saturation currents of an n-FET transistor of typical size plus a p-FET transistor. These transistors, along with other process and yield monitors, are located in the kerf (the area between chips on the wafer). The $I_{d_{sat}}$ process parameter is an excellent indicator of chip performance and is measurable weeks before flush delay, while the wafer is still being processed. If a high $I_{d_{sat}}$ is measured, the device channel lengths are effectively short, and chip speed will be toward the fast end. If a low $I_{d_{sat}}$ is measured, the device channel lengths are long, and chip speed will be toward the slow end.

During wafer testing, $I_{d_{sat}}$ is measured on the kerf while its corresponding chip product site is tested for flush delays. Both $I_{d_{sat}}$ and the flush delays are entered in a database, where regression analysis is performed. **Figure 5** shows $I_{d_{sat}}$ vs. 1000/flush delay (frequency) to linearize the curve fitting. The value of $I_{d_{sat}}$ necessary to meet the flush-delay sort specification can be selected from the graph.

$I_{d_{sat}}$ is an indicator of the chip's ± 3 sigma process distribution window. Thus, the required value of $I_{d_{sat}}$ can then be compared against the nominal $I_{d_{sat}}$. The resultant difference can be divided by one standard deviation (sigma) to determine the number of sigmas by which the selected $I_{d_{sat}}$ is separated from the nominal $I_{d_{sat}}$. A projection of chip yield is based on the value of this number of sigmas from a nominal process. Sorting at a

nominal process, or 0 sigma, would produce chips at a 50% yield. Sorting at greater than nominal (assume +1.5 sigma) would yield chips in the 93% range taken from a statistical normal distribution chart. Sorting at less than nominal (assume −1.5 sigma) shows a chip yield of only 7%.

Chip performance optimization in manufacturing

Once the design has been demonstrated to be robust and the semiconductor process mature, two basic performance-related improvements are possible: a process movement to the short-channel side called *line center shift*, and a combination of line center shift with *line tailoring*. Line centering is a shift in the process to the short side of L_{eff} within the current process specifications that has minimal yield implications. Tighter tracking is implied, but there is no major requalification effort. Line tailoring is a more aggressive shift, together with a tightened distribution of L_{eff}, and requires enhanced process controls and usage of dedicated tools; i.e., there can be no change of critical tools in key processes. Additional qualification of deep-sorted hardware may be required. The effect of tailoring is a substantial compression of the channel-length spread and consequently a tighter chip performance distribution. Cost may be substantial, however, and cost–benefit tradeoffs must be optimized. Line tailoring yields the optimum performance positioning.

Ideally, the achievable performance improvement is gated by fundamental semiconductor process limitations and not by MCM wiring, circuit margins, or other nonprocess parameters. These limitations are all verified empirically.

Chip performance improvement implementation process

The goal of the S/390 Division is to have the best possible system performance at general availability (GA) time, when system shipments to external customers begin. The time from the beginning of chip design to the beginning of external shipments is the time period during which improvements can be identified and implemented. It is necessary to have close collaboration among experts in logic design, manufacturing engineering, test and qualification, process engineering, and marketing in order to identify improvements that can be implemented in a timely fashion. An improvement-implementation process which includes several workshops between "chip design release-to-manufacturing" and GA has been introduced and is strictly adhered to. It ensures that the key members of the improvement team from several locations generate clear agreements regarding potential improvements and define all necessary activities and clarifications, including a timetable for completion of work items and decisions. Technical and business aspects are addressed. The most important aspects dealt with are the following:

- Compatible speed for all chips which are tied together in the processor cycle domain.
- Chip design improvements via logic and physical design for the slowest chips in the set.
- Chip characterization to identify limitations (voltage, line center, burn-in, etc.).
- Analysis of yields (such as yield roll-off for chips with very short channel length).
- Chip design robustness for optimized line or application conditions.
- Business aspects (such as line tailoring for some chips).

Voltage and temperature optimization

The performance of the processor and its surrounding chips also depends on the voltage and temperature under which the chips operate. Optimization in these conditions is also an important performance management task.

The process development organization typically defines voltage ranges for its CMOS technologies of nominal voltage ± 8%, i.e., 2.5 V ± 0.20 V for CMOS 6S2 and 1.8 V ± 0.15 V for CMOS 7S.

Experience with power-supply voltage regulation, voltage drops from power supply to the operating circuit, and electrical noise show that the voltage variation at the circuit in a G5 system is significantly less than 8%. The operating voltage for the chips can therefore often be increased without exceeding the specified maximum level. Extensive characterization of G5 chips from wafer test through burn-in and system test confirm that the voltage could be raised higher than the specified technology maximum for the chip with the most critical performance. This change improves chip and system performance by more than 8%. Measurements have repeatedly shown that the chip's performance gain grows nearly one-to-one with the increased voltage; i.e., a 5% higher voltage will create a chip performance gain of almost 5%.

The use of a very effective chip/module cooling system (chiller) in the G5 prime system reduced the worst-case chip temperature by approximately 50°C. The corresponding performance gain for the lower chip temperature is approximately 8% using the verified performance dependency of chip performance on temperature.

The total performance gain via optimization of the operating voltage and the chip temperature is more than 15%.

Conclusions

An integration of techniques from a wide variety of disciplines incorporating test, design, timing, and sorting has significantly improved the performance of the S/390 G5.

These techniques have a strong empirical foundation, requiring early testing of a wide process range of hardware. The philosophy is to explore design and process margins, test to failure, and plan to iterate. It is built on a foundation of robust circuit and chip design. We estimate that the implementation of these techniques has increased performance by more than 35% over that of an unsorted and unoptimized design in the same semiconductor technology. Continuous planning and communication among the semiconductor process, design, and test teams are vital to ensure timely implementation and adequate manufacturability.

Acknowledgments

Many individuals contributed to the development of these techniques and the success of this effort: in IBM Microelectronics, Karre Green and Kort Longenbach for process characterization and monitoring, Greg O'Malley for testing, Deb Hamm for burn-in and high-voltage characterization, Rich Hamilton for sort MCM assembly, Dawn Tudryn and Glen Froese for logistical control, and Moe Hamel for general technical coordination; in S/390 Boeblingen, Germany, Peter Bunse for sort specifications and interfacing with IMD, Herb Kuestner for characterizing MCMs, Uwe Fassnacht for chip timing and layout extraction, and Axel Prucha for sort-point module build. MiCrus Corporation Test Development Engineering personnel Franco Motika, Ray Kurtulik, Donato Forlenza, Orazio Forlenza, Wendy Chong, and Adrian Anderson were responsible for chip characterization and hardware bring-up. In IBM S/390 Poughkeepsie, New York, Pat Meaney was responsible for SC and SD chip characterization, Sean Carey for chip timing consultation, Wilt McClain for interfacing with Micrus and MD, and Phil Wu for overall test management. We would also like to thank the members of the G5 test team for their assistance: Uli Baur, Tom Foote, Bill Huott, Mary Kusko, Larry Lange, Bryan Robbins, and Peilin Song.

*Trademark or registered trademark of International Business Machines Corporation.

References

1. T. Slegel, R. M. Averill, M. A. Check, B. C. Giamei, B. W. Krumm, C. A. Krygowski, W. H. Li, J. S. Liptay, J. D. McDougall, T. J. McPherson, J. A. Navarro, E. M. Schwarz, and C. F. Webb, "IBM S/390 G5 Microprocessor," presented at the 1998 Hot Chips Symposium, Stanford, CA, August 1998.
2. D. E. Hoffman, R. M. Averill, B. Curran, Y. H. Chan, A. Dansky, R. Hatch, T. McNamara, T. McPherson, G. Northrop, L. Sigal, A. Pelella, and P. M. Williams, "Deep Submicron Design Techniques for the 500MHz IBM S/390 G5 Custom Microprocessor," *Proceedings of the 1998 International Conference on Computer Design*, 1998, pp. 258–263.
3. W. V. Huott, T. J. Koprowski, B. J. Robbins, M. P. Kusco, S. V. Pateras, D. E. Hoffman, T. G. McNamara, and T. J. Snethen, "Advanced Microprocessor Test Strategy and Methodology," *IBM J. Res. Develop.* **41,** No. 4/5, 611–627 (July/September 1997).
4. L. M. Huisman, "Correlations Between Path Delays and the Accuracy of Performance Prediction," *Proceedings of the 1998 International Test Conference,* October 1998, pp. 801–808.

Received April 22, 1999; accepted for publication July 9, 1999

Richard F. Rizzolo *IBM System/390 Division, 522 South Road, Poughkeepsie, New York 12601 (rizzolo@us.ibm.com).* Mr. Rizzolo is a Senior Engineer and served as test team leader for the S/390 G4 and G5 projects. He also had primary responsibility for sorting for S/390 chips designed in Poughkeepsie. He received his B.S. degree in physics from Rensselaer Polytechnic Institute in 1977 and his M.E. degree in electrical engineering from Rensselaer in 1980. Since joining IBM in 1978, Mr. Rizzolo has worked on bipolar and CMOS projects in the areas of design for testability, high-frequency design and timing analysis, diagnostics, and circuit design. In 1994 Mr. Rizzolo was assigned to Boeblingen, Germany, to develop CMOS timing methodology. He is a member of the IEEE and holds three patents and several published inventions. Mr. Rizzolo is a co-author of several papers in the field of testability and diagnostics; he has received three Outstanding Technical Achievement Awards and two Outstanding Innovation Awards.

Guenter Hinkel *IBM System/390 Division, Boeblingen, Germany (hinkel@ibmde).* Mr. Hinkel joined IBM in 1964 after graduating in physics from the university in Hamburg, Germany. He began work in the Central Component Evaluation and Application Laboratory in Sindelfingen, Germany, in the areas of memory chip characterization and qualification. After two one-year assignments, in Kingston/Poughkeepsie and East Fishkill, he was appointed to a management position in 1969. In 1978 Mr. Hinkel moved to the Boeblingen Development Laboratory as manager in Product Assurance and then in VLSI Development. His assignments also included leading the Boeblingen CMOS pilot line yield enhancement task force in 1986 and 1987 and development and implementation of the performance sorting methodology for S/390 CMOS chips in cooperation with the S/390 team and MD Burlington from 1994 until 1999. He has received several major awards, including a Corporate Award for his outstanding contributions in the area of CMOS chip performance management. Mr. Hinkel is also a member of the German ITG/VDE in the expert group on quality and reliability of high-density semiconductor devices,

Steven Michnowski *IBM System/390 Division, 522 South Road, Poughkeepsie, New York 12601 (michnows@us.ibm.com).* Mr. Michnowski is an Advisory Engineer in S/390 Test Development. He received his Bachelor of Science degree in electrical engineering from the New Jersey Institute of Technology in 1984. After joining IBM in 1984, he was responsible for custom I/O circuit design and analysis for memory interface products. His present responsibilities include establishing system cycle-time performance criteria and test-pattern generation for the L2 cache chip. Mr. Michnowski has received an Outstanding Technical Achievement Award and several informal awards; he has several published invention disclosures.

Thomas J. McPherson *IBM System/390 Division, 522 South Road, Poughkeepsie, New York 12601 (mcpherso@us.ibm.com).* Mr. McPherson is an Engineering Manager in S/390 Microprocessor Development. He received his B.S. degree in electrical engineering from Rutgers University in 1990 and his M.S. degree in computer engineering from Syracuse University in 1994. Mr. McPherson joined IBM in 1990 as a logic designer for CMOS ASIC designs. Since 1994 he has worked on CMOS microprocessor designs. Mr. McPherson was the Frequency Leader for the G5 and G6 microprocessors; he is currently working on future S/390 microprocessors.

Allen J. Sutcliffe *IBM System/390 Division, 522 South Road, Poughkeepsie, New York 12601 (asutclif@us.ibm.com).* Mr. Sutcliffe is an Advisory Engineer in System Technology Engineering Test and Validation, where he is the team leader. His primary responsibility is qualifying S/390 CMOS servers with respect to cycle time and manufacturability. He received a B.S. degree in electrical engineering from the State University of New York at Binghamton in 1979, and a B.S. degree in computer science with emphasis on hardware from Union College in 1984. Since joining IBM in 1979, Mr. Sutcliffe has worked on voltage, frequency, and temperature characterization and resolution of engineering, manufacturing, and field problems for bipolar and CMOS systems. He has received four Outstanding Technical Achievement Awards and several informal and team awards.