A 480-MHz RISC Microprocessor in a 0.12-μm $L_{\rm eff}$ CMOS Technology with Copper Interconnects

Chekib Akrout, John Bialas, Miles Canada, Duane Cawthron, James Corr, Bijan Davari, *Senior Member, IEEE*, Robert Floyd, Stephen Geissler, Ronald Goldblatt, Robert Houle, Paul Kartschoke, Diane Kramer, Peter McCormick, *Member, IEEE*, Norman Rohrer, Gerard Salem, Ronald Schulz, Lisa Su, *Member, IEEE*, and Linda Whitney

***Abstract*—This paper describes the performance improvements of a reduced instruction set computer (RISC) microprocessor that has migrated from a 2.5-V technology to a 1.8-V technology. The 1.8-V technology implements copper interconnects and low V_t field-effect transistors in speed-critical paths and has an $L_{\rm eff}$ of 0.12 μm. Global clock latency and skew are improved by using copper wires, and early mode timings are improved by reducing clock skew and adding buffers. These enhancements, along with an environment of 2.0 V, 85°C, and with a fast process, produced a 480-MHz RISC microprocessor.**

***Index Terms*—CMOS, copper, low threshold, microprocessor, PowerPC, reduced instruction set computer (RISC).**

I. Introduction

A 32-BIT 480-MHz PowerPC[1] reduced instruction set computer (RISC) microprocessor has been migrated into an advanced 0.2-μm CMOS technology with copper for its interconnects and multithreshold transistors. The technology features help to increase the microprocessor's internal clock frequency up to 480 MHz at 2.0 V and 85°C, and at the fast end of the process distribution. When operating at room temperature, the clock frequency increases to over 500 MHz.

II. Processor Architecture

The microprocessor's architecture can dispatch two instructions per cycle by using two 32-KB L1 caches, one for data and one for instructions. An integrated L2 cache controller can work with L2 cache sizes of 256 KB, 512 KB, or 1 MB, and the I/O's interface with the external bus using an industry standard of 3.3 V. The 60× bus has ratios of 1× as well as 2× to 8× with 0.5× increments. The L2 cache interface contains ratios of L2 bus frequency to internal clock frequency of 1 : 1, 2 : 3, 1 : 2, 2 : 5, and 1 : 3. The microprocessor has been successfully implemented in a 2.5-V CMOS technology [1] and has migrated to a 1.8-V CMOS technology (Fig. 1). The 1.8-V microprocessor has estimated SPECint95 and SPECfloat95 ratings of 20 and 12, respectively, at 480 MHz with a 1-MB L2 cache running at 240 MHz and a 60× bus frequency of 96 MHz. The 60× bus is the high-speed 32- and 64-bit processor bus interface used throughout the PowerPC family of processors.

Manuscript received July 10, 1998; revised September 22, 1998.

C. Akrout, D. Cawthron, D. Kramer, and L. Whitney are with IBM Microelectronics Division, Austin, TX 78758 USA.

J. Bialas, M. Canada, J. Corr, R. Floyd, S. Geissler, R. Houle, P. Kartschoke, P. McCormick, N. Rohrer, and G. Salem are with IBM Microelectronics Division, Essex Junction, VT 05452 USA (e-mail: pkartsch@btv.ibm.com).

B. Davari, R. Goldblatt, R. Schulz, and L. Su are with IBM Microelectronics, Hopewell Junction, NY 12533 USA.

Publisher Item Identifier S 0018-9200(98)07042-5.

III. Technology

The 2.5-V CMOS technology, with a field-effect transistor (FET) $L_{\rm eff}$ of 0.18 μm, has five interconnect levels made of aluminum, a tungsten local interconnect, and tungsten vies. The local interconnect contacts both diffusion and polysilicon by overlap. The contacted wiring pitch for the third level of metal is 1.26 μm. The metal thicknesses are 0.54 μm for M1, 0.73 μm for M2, M3, and M4, and 2.07 μm for M5. The nominal internal clock frequency is 221 MHz; but the frequency increases to greater than 275 MHz when the $L_{\rm eff}$ line is tailored shorter, the voltage is increased, and the maximum operating temperature is reduced. The I/O's were designed to interface from the internal voltage of 2.5 V to the 3.3-V external bus. The 6.4 million transistors fit into a 67 mm^2 chip area.

The 1.8-V CMOS technology has a nominal NFET $L_{\rm eff}$ of 0.12 μm. This technology is a linear shrink from the 2.5-V technology with some changes to the ground rules for the local interconnect and the Si/local interconnect spacing. In this technology, the microprocessor has six levels of metal (Fig. 2). Metal-level thicknesses are 0.40 μm for M1 and M2, 0.55 μm for M3 and M4, and 1.20 μm for M5 and M6. Beginning with M1, all levels of metal and vias are fabricated with copper. Again, a tungsten local interconnect with a tungsten via is used to increase the overall density. The copper wires, manufactured with a damascene process, are deposited via electroplating, which provides lower costs and better connections to the vias than the tungsten via and metal etch-back process used for aluminum. A single damascene process is used for M1 and a dual damascene process is used for the remaining metal levels. The dual damascene process produces the via connections and the metal level, in copper, at the same time. The third-metal level has a reduced contacted pitch of 0.81 μm. This microprocessor's SRAM cell, used in several custom circuits, is a migrated design with new ground-rule waivers measuring 7.6 μm^2. The smallest SRAM cell available in the 1.8-V technology, which measures 6.84 μm^2, was not selected in order to maintain the M1 word line and not adversely affect the

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

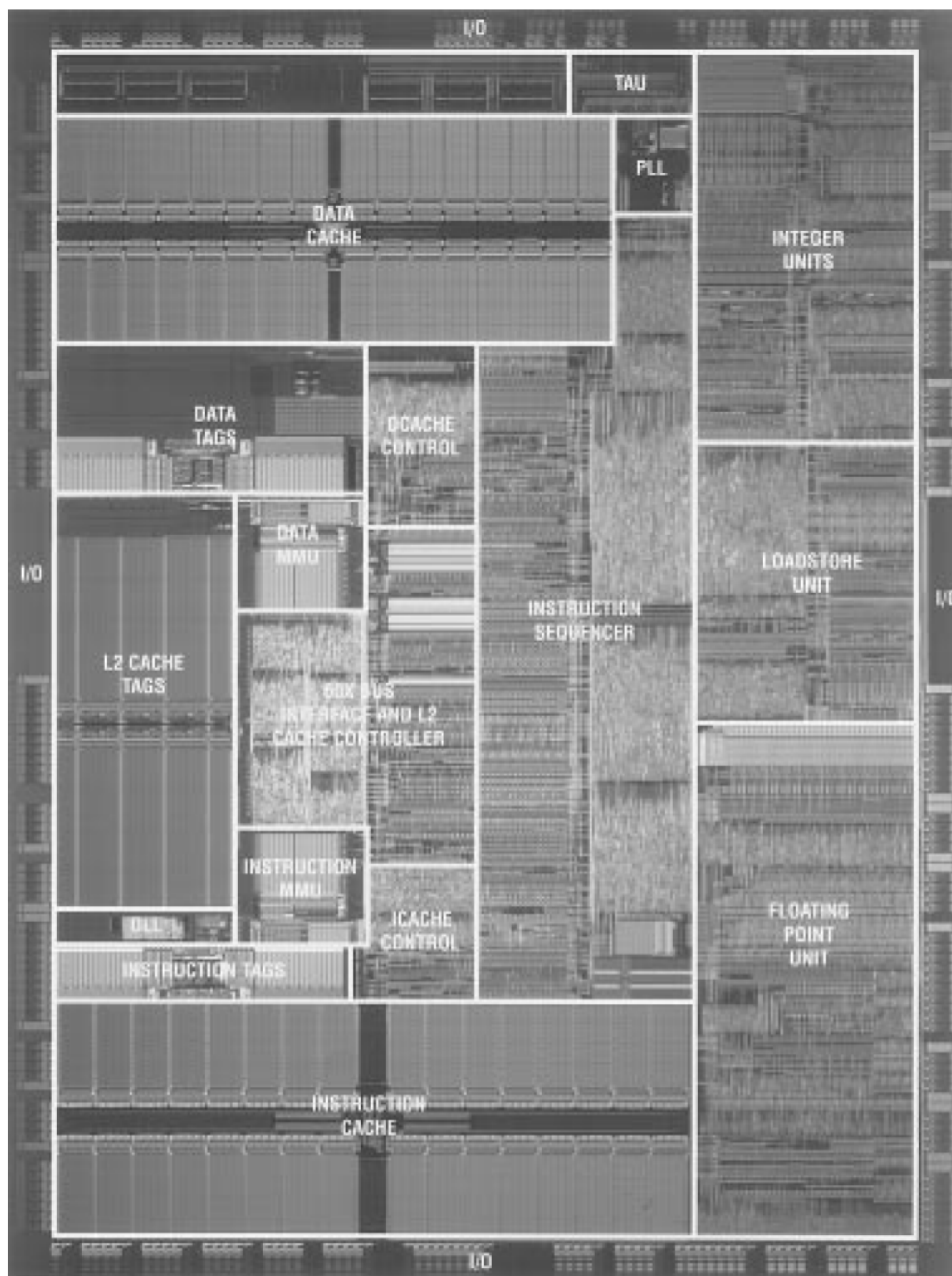

Fig. 1. A 480-MHz RISC microprocessor with copper interconnect.

distinct threshold voltages (V_t) for both NFET's and PFET's. At long channel lengths, the low threshold voltage NFET has a V_t that is 100 mV lower than the nonlow V_t device of 310 mV. The low V_t devices are provided to increase performance. The microprocessor's present chip size is 40 mm^2. Table I lists key microprocessor and technology parameters for both the 2.5- and 1.8-V technologies.

IV. Copper Interconnect

Copper interconnects provide a 40% decrease in RC delay over the traditional aluminum wires due to the reduced resis-

tradeoff of the resistance or capacitance by thinning the wire. The sheet resistance of the previous interconnect technology is maintained because a thinner wire improves manufacturability and yield and because the aspect ratio of the wire is not as extreme. Reliability of these thinner wires is not compromised because copper has less electromigration than aluminum. For the 1.8-V technology, the RC reduction of copper is mainly caused by the reduced capacitance; the sheet resistance for a given interconnect remains nearly constant. The total RC reduction is 23%. Because of lower capacitance on the wires, intralayer coupling capacitance between the third-metal-level wires is reduced by 27%. Fig. 3 shows the late-mode timing

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

TABLE I
TECHNOLOGY FEATURES

Property of Microprocessor	2.5V Technology (Aluminum)	1.8V Technology (Copper)
Voltage	2.5 ± 0.2V	1.8 ± 0.15V
I/O Voltage	3.3 ± 0.3V	3.3 ± 0.3V
L2 Cache I/O Voltage	3.3V	3.3V or 1.8V
L_{eff} for NFET/PFET	0.18µm/0.18µm	0.12µm/0.15µm
T_{ox}	5.0nm	3.5nm
Contacted Local Interconnect Pitch	0.84µm	0.55µm
Contacted M1 Pitch	0.98µm	0.63µm
Contacted M2, M3, M4 Pitch	1.26µm	0.81µm
Contacted M5, M6 Pitch	2.52µm	1.62µm
M1 Thickness	0.54µm	0.40µm
M2 Thickness	0.73µm	0.40µm
M3 Thickness	0.73µm	0.55µm
M4 Thickness	0.73µm	0.55µm
M5 Thickness	2.07µm	1.2µm
M6 Thickness		1.2µm
SRAM Cell Area	18.4µm^2	7.6µm^2
Chip Size	67mm^2	40mm^2
Power	5.4W @ 300MHz	3.0W @ 400MHz

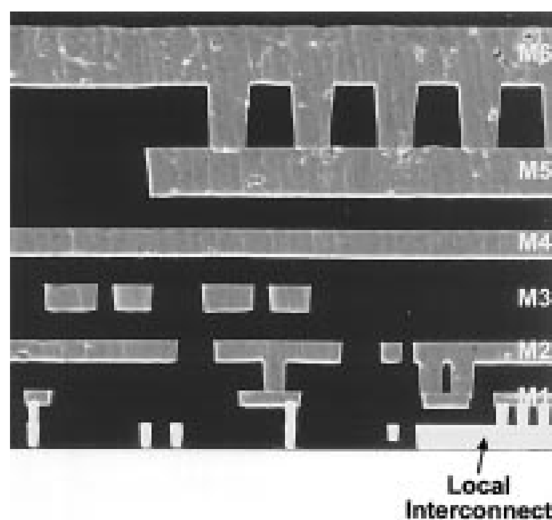

Fig. 2. Scanning electron microscope of six copper metal levels and local interconnect.

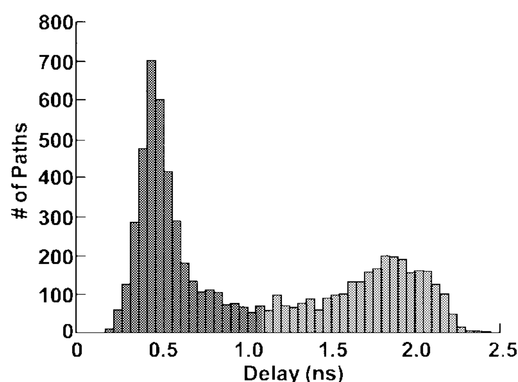

Fig. 3. Histogram of path delays with Al interconnects.

wired with aluminum. Fig. 4 presents the same data with the metal levels fabricated in copper; the path delay near 2.0 ns is reduced because the RC delay percentage on the critical paths is decreased. Path delays less than 1.0 ns did not change significantly because the paths contained little RC delay. The top 5% of the paths have RC delay ranging from

from copper is noticeable. The top 5% of the paths typically contain 18–23 gates/cycle. With a higher percentage of the critical path delay dominated by intrinsic gate delay, the frequency of the microprocessor more easily increases at the fast end of the process distribution, with either higher voltage or lower temperature. Some of the late-mode critical timing paths, which still contain a large amount of RC, are tripled in width to decrease the RC delay further. Decreased cycle time is accomplished without adding a pipe stage, so a high frequency can be reached while still preserving the number of instructions executed per cycle. A similar timing for early

DOCKET ALARM
Find authenticated court documents without watermarks at docketalarm.com.

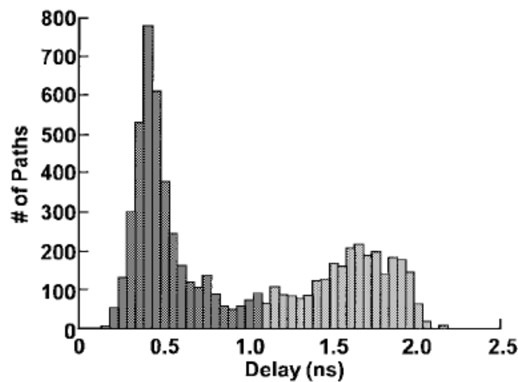

Fig. 4. Histogram of path delays with Cu interconnects.

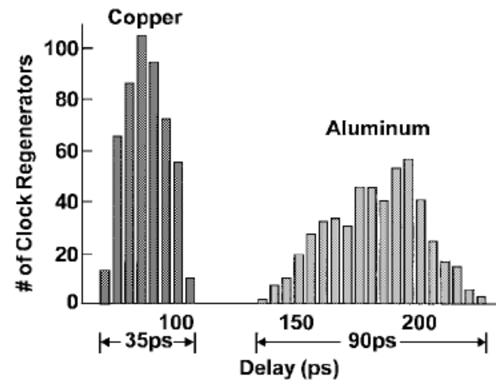

Fig. 5. Histogram of the clock tree latency with Al and Cu interconnects.

data paths. If a latching element has data inputs that transition too fast or a clock that transitions too slow, the latch can be corrupted. Because copper interconnects reduce RC delay and can produce early mode paths, buffers are added to slow down and eliminate these paths.

V. Low V_t FET Implementation

The microprocessor's internal clock frequency is further improved with low V_t FET's. With our timing and floor-planning methodology, low V_t transistors are inserted manually into custom macros to improve delay through a custom circuit. The remaining logic on the chip is implemented with 4-bit nibbles of random logic gates (extended custom standard-cell library) and individual 1-bit standard-cell books. A duplicate low V_t library has been developed for these standard cells. To allow nonlow V_t and low V_t books to be interchangeable, the layout of the low V_t books is identical to that of the nonlow V_t books in size and wiring. After timing the chip at 500 MHz, any part of a late-mode timing path that contains standard-cell books and does not meet the timing requirements is converted to a low V_t book and the chip is retimed. Because low V_t books are used only for the critical paths, the number of transistors with the reduced threshold voltage is limited. For the whole chip, 12.3% of the 4-bit standard-cell books and 5.0% of the 1-bit standard-cell books were converted to low V_t books. The I_{dsat} for the low V_t FET's is 10% higher than for the standard FET's. The low V_t FET leakage is 8–10× higher than standard FET's. Therefore, to balance power versus speed, only 4.2% of the total transistors on the microprocessor are converted to low V_t, which accounts for a performance increase of 6.5%. To minimize risk and maintain noise margins, this design does not implement low V_t FET's within dynamic circuits.

VI. Clock Distribution and Phase-Locked Loop Design

Increasing the processor speed with a technology remap required a proportional improvement in the clock design and distribution because clock skew subtracts from the processor cycle time. A hierarchical clock-distribution methodology included variable wire widths to minimize delay, automated clock load balancing, and local clock buffer clustering. The clock distribution network contains four stages: a large global clock buffers. The local clock splitter converts a global single-phase clock into complementary unbuffered C1 and C2 clocks. The local clock buffers use the split outputs to produce a buffered master latch clock C1 and a buffered slave latch clock C2.

The clock tree is designed in a pseudogrid-spine network. Thick metal levels M6 and M5 are used for the primary global clock wiring. Four M5 spines run parallel for 3 mm, and M6 is used to connect the spines to form a grid network. Wiring levels M4 through M2 are used to connect clusters of clock splitters to the spines, and total wire lengths are restricted to 500 μm. If the 2.5-V-technology microprocessor implements six levels of metal, the improved RC delay from copper interconnect reduces clock latency from 170 to 85 ps and improves the clock tree skew from 55 to 35 ps, as shown in Fig. 5. This improvement in clock skew helps reduce the number of early mode fast paths. The copper interconnect allows the use of narrow metal lines, which reduces inductance effects; an aluminum interconnect could produce similar RC skew by using wider lines, but the increased inductance would affect skew significantly.

Because additional clock skew directly impacts the process cycle time, the local clock splitter and local clock buffer must deliver accurate clock signals to the latches. Therefore, the number of local clock buffers was increased from eight to 24, with each buffer driving a smaller restricted load range. The restricted buffer load capacitance range reduces clock buffer skew from 100 to 20 ps. Software tools allow automated repowering based on wire and latch capacitances. The low RC interconnect and smaller clock buffer load capacitance range improves the distribution of a falling C1 edge at the latch boundaries, as shown in Fig. 6, which compares the 2.5-V technology to the 1.8-V technology.

Critical components must be adjusted when a phase-locked loop (PLL) is remapped from a 2.5-V to a 1.8-V technology. The PLL [3] was originally designed for a 2.5-V CMOS technology. The critical components examined during a technology remap are the phase detector gain, voltage reference center current, charge-pump current sources, feed-forward current sources, and voltage-controlled oscillator (VCO) gain. Statistical analysis is performed for various parameter changes such as channel length, threshold voltage, temperature, power

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

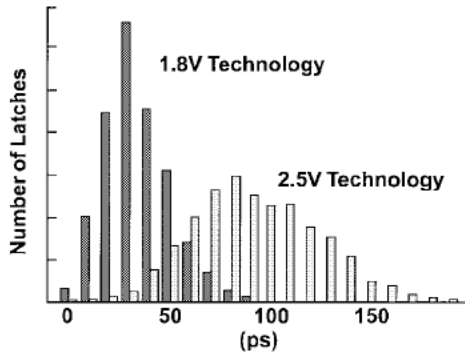

Fig. 6. Clock skew of all falling C1.

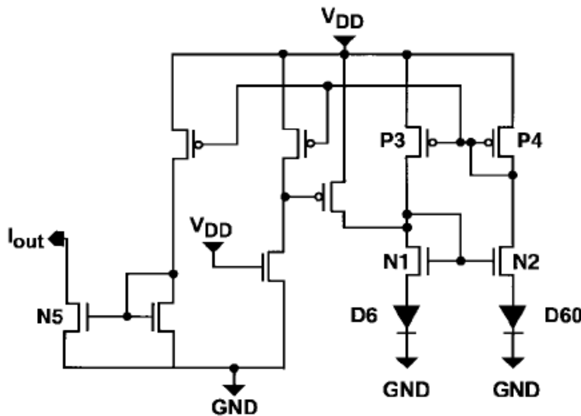

Fig. 7. Voltage reference circuit.

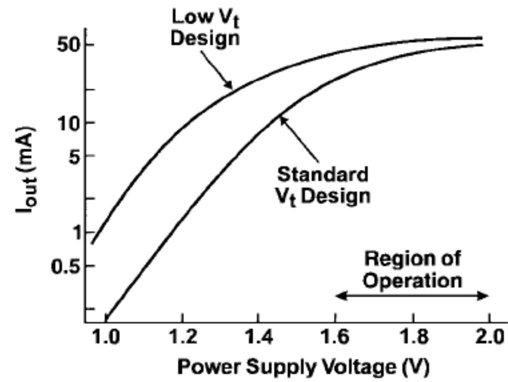

Fig. 8. Voltage reference current source N5.

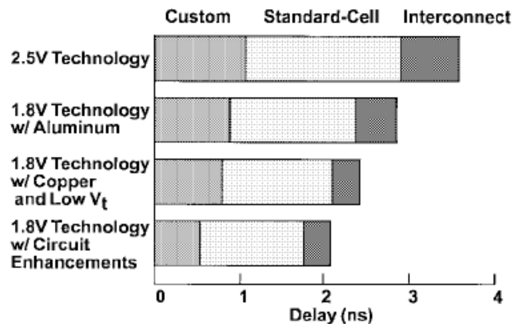

Fig. 9. Typical path delays.

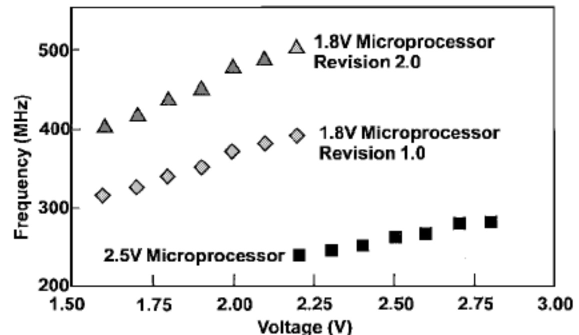

Fig. 10. $F_{\max}$ versus voltage for critical path delays.

closed-loop circuit simulations as well as closed-loop behavioral time-domain models are used to optimize the PLL for damping coefficient and to improve stability of the VCO pole, phase error, and jitter.

Analog circuit operation becomes difficult when the power supply is scaled because MOS device threshold voltage and dc operating points are not scaling proportionally. The voltage reference circuit consists of a pair of ratioed diodes (D6 and D60), a pair of ratioed NMOS transistors (N1 and N2), and a pair of load PMOS transistors (P3 and P4), as shown in Fig. 7. As the technology is scaled, the diode forward bias remains constant, the MOS device dc bias condition (V_{gs}–V_t) scales by 0.85×, and the power supply scales by 0.72×. Thus, as the power supply is lowered, the voltage-reference-transistor operating points move from saturation into the linear region. When the transistors are in the linear region of their operation, the charge pump stops functioning as an infinite current source and operates with poor stability over process, temperature, and voltage. The voltage-reference-transistor operation in the saturation region is guaranteed by converting NMOS transistors (N1 and N2) from standard threshold devices to zero V_t devices and PMOS transistors (P3 and P4) from standard V_t to low V_t devices. The improvement in the voltage reference circuit is measured through transistor N5 (Fig. 8).

VII. Timing and Performance

The 1.8-V technology improves the frequency of the pro-

enhancements contribute a 17% improvement. In addition, the copper wires for this design result in a 12% improvement in path delay, while low V_t devices add another 6.5% to the maximum frequency. Total improvement is 77%. One of the critical path delays is shown in Fig. 9. The timing delay is segmented into custom gate delay, standard-cell gate delay, and interconnect delay for the 2.5-V technology, 1.8-V technology without copper interconnects or low V_t FET's, 1.8-V technology with copper interconnects and low V_t devices, and 1.8-V technology with circuit enhancements.

A timing analysis was performed at 480 MHz, 2.0 V, 85°C, and all process parameters were centered at 1-sigma fast—except for $L_{\rm eff}$, which was centered at 3-sigma fast—to represent a chip design line-tailored to the fast corner. The fast corner of the process distribution also shows more RC effects. Fig. 10 shows the late-mode timing for revisions 1.0

DOCKET ALARM
Find authenticated court documents without watermarks at docketalarm.com.

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.