

SESSION 3: MICROPROCESSORS

WPM 3.5: System, Process, and Design Implications of a Reduced Supply Voltage Microprocessor

Randy Allmon, Brad Benschneider, Michael Callander, Linda Chao, Dan Dever, Jim Farrell, Nanette Fitzgerald, Joel Grodstein, Soha Hassoun, Larry Hudepohl, David Kravitz, Jim Lundberg, Rich Marcello, Suzanne Marino, Jeff Pickholtz, Ron Preston, Maurice Richesson, Sri Samudrala, Douglas Sanders
Digital Equipment Corporation
Hudson, MA

This paper describes the system, process and design implications of converting a microprocessor chipset originally implemented in a 5V 1.5 μ m (drawn) CMOS process to one implemented in a 3.3V 1.0 μ m (drawn) CMOS process. The chipset is 75% faster than the previous generation and is comprised of a processor chip, a floating point chip, a cache controller chip, and a clock chip¹. It operates at 62.5MHz under worst-case conditions. Figures 1-4 contain micrographs of each design. Table 1 describes power and packaging specifications for each chip. Table 2 describes the 3.3V 1.0 μ m (drawn) process specifications. Figure 5 shows a high-temperature schmoo plot for the CPU chip.

The 1 μ m/3.3V process is derived from the 1.5 μ m/5V process by scaling down all lateral dimensions and the gate oxide by 67% and reducing VT's proportionally with supply voltage. The supply voltage is reduced primarily to reduce overall power consumption but also to improve reliability. Other process enhancements include a third level of Al interconnect, low-resistance source drains, and precision resistors. The third level of metal is added for improved power distribution and to maintain acceptable electrical integrity. The TiN component of Metal-3 can also act as a fuse layer if redundancy is incorporated in the design. When optimized, this 3.3V process is as fast as a comparable 5V process. The change in VT necessitates extreme care in designing dynamic circuits subject to absolute noise, such as input buffers, even though TTL level conversion is much simpler. The clock-chip oscillator input provides an example. A differential ECL oscillator is ac coupled to the oscillator input. Pullup and pulldown resistors of equal value are used to bias the input to VDD/2. The input is then fed directly into the differential amplifier, as shown in Figure 6. The circuit is able to resolve a voltage difference of 300mV.

As a result of the Tox reduction, an on-chip decoupling capacitor ring (.012 μ F) is added for improved signal integrity. The ring must supply sufficient charge during each 4ns phase to decouple all switching events. The capacitor is implemented as parallel NMOS devices with VDD attached to the gate and VSS attached to both source and drain. The dimensions of the device (12.5 μ m channel length by 150 μ m channel width) are chosen to maximize gate area and limit worst-case RC delay to 0.16ns in both poly-silicon and the channel. Internal noise is reduced by 75% at the expense of a predicted 3.7% yield reduction.

Reduced geometries make it possible to increase the size of a number of arrays in the chipset. Row redundancy is added to the on-chip cache for improved yield. The TiN Metal-3 fuse circuit designed for laser programming during wafer sort is shown in Figure 7. In a processor chip, a fraction of the area is devoted to the cache array and therefore adding redundancy can only increase the overall yield to the level of the remaining non-cache logic. Yield analysis shows that adding row redundancy to the cache design could double the expected yield during prototype manufacturing.

The chipset interface to 5V CMOS or TTL I/O poses several problems. One problem is latchup, which can be resolved by biasing appropriate wells in the output driver to 5V. Another problem occurs when the output of a tristated driver sees a voltage greater than 3.3V plus the magnitude of the PMOS threshold, resulting in the output PMOS pullup of the driver turning on and sinking current into the power supply. This current can exceed an ampere when interfacing to a wide data bus.

The driver, shown in Figure 8, uses only two devices more than a conventional I/O driver - a PMOS shunt device (T4) and an NMOS device (T2). In tristate, when the output node is driven above 3.3V plus VTP, the shunt device turns on. NMOS T2, by body effect and channel IR drop, allows the gate of the output PMOS pullup to follow the output voltage to within 100mV and also T2 holds node PASS to within 100mV of VDD3. This keeps the output PMOS pullup off and limits current into the 3.3V supply to less than 1.8mA per driver. The driver suffers neither the area penalty or increased predrive loading of other solutions using cascaded output devices².

ESD protection of the circuits is complicated by the presence of silicided source drains, because the output drivers are not able to absorb sufficient charge to protect the chip. Consequently, ESD devices were added to each output. A grounded-gate MOSFET clamp was placed on the pad side of an impedance matching resistor to prevent a large ESD current from flowing into the output-driver MOSFETs.

The change in supply voltage also has system implications. Proper sequencing between the 3V supply and the 5V supply is required to avoid excess current injection. The reduced power supply voltage, and the consequent reduction in power dissipation, allow the chipset to be used in a wide variety of environments without customized packaging for thermal management. Finally, overall system signal integrity is improved as a result of two factors. First, voltage swings on the multi-drop system bus are reduced and consequently, the interchip communication bus settles more quickly. Second, the precision resistor used as a source termination device in each driver makes it possible to terminate the line more effectively than with previous processes. The end result is the ability to drive a seven-drop bus within 12ns, more than fast enough to meet overall cycle time requirements.

Acknowledgement

The authors acknowledge the technical contributions of S. Bhatt, A. Black, M. Butler, G. Cheney, C. Dobriansky, E. Gomes, C. Herbert, E. Kagan, K. Kuchler, S. Martin, R. Meeks, M. Minardi, T. Shedd, B. Supnik, B. U'pham.

¹Allmon, R., et al., "CMOS Implementation of a 32b Computer", ISSCC DIGEST OF TECHNICAL PAPERS; p80-81; Feb., 1989.

²Roberts, A., et al., "A 256K SRAM with On-chip Power Supply Conversion", ISSCC DIGEST OF TECHNICAL PAPERS; p252-253; Feb., 1987.

FIGURES 1, 2, 3, 4 - See page 263

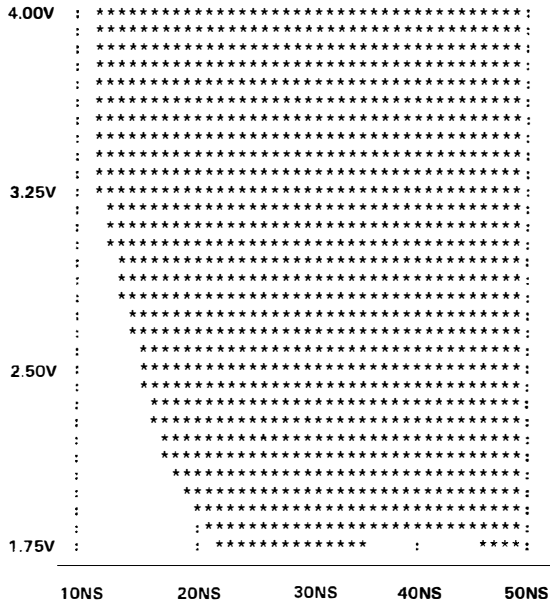


FIGURE 5 - Schmoo plot of a CPU chip at high temperature

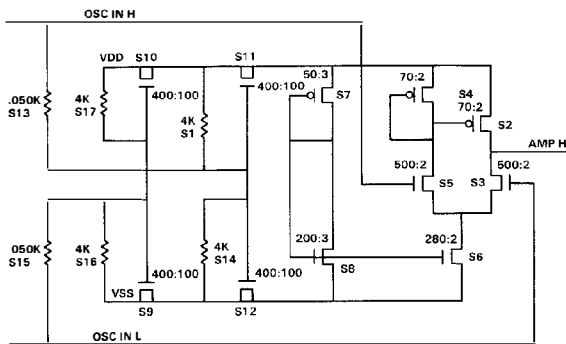


FIGURE 6 - AC-coupled differential oscillator input buffer

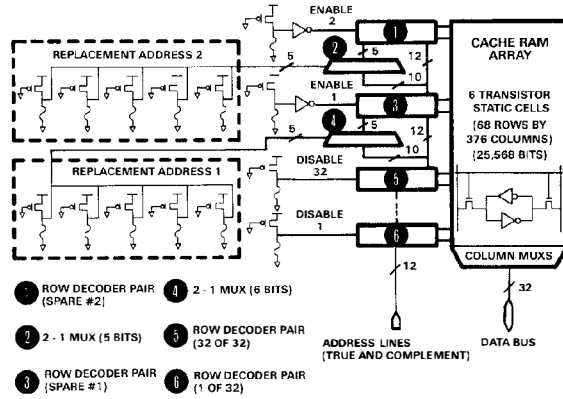


FIGURE 7 - P cache row redundancy scheme

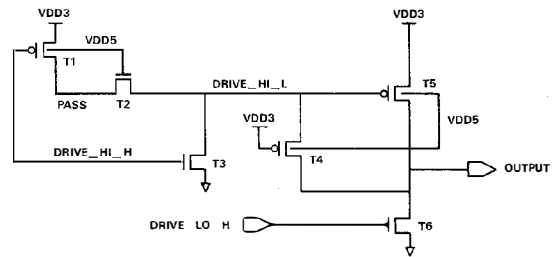


FIGURE 8 - 3V/5V output buffer

	Maximum Power Dissipation	Package Type
Processor Chip	5W	224 LDCCW
Floating Point Co-processor	3.5W	224 LDCCW
Cache Controller Chip	3.5W	224 LDCCW
Clock Chip	1.5W	68 Cerquad

TABLE 1 - Chipset power and package specifications

Effective N Channel length	0.7µm
Effective P Channel length	0.55µm
Metal 1	2.0µm Width, 1.0µm space
Metal 1 Contact	1.0 x 1.0µm
Metal 2	2.5µm Width, 1.0µm space
Metal 2 Contact	1.0 x 1.0µm
Metal 3	4.0µm Width, 6.0µm space
Metal 3 Contact	4.0 x 4.0µm
Gate Oxide Thickness	150Å
Metal 1 Field Oxide	0.75µm
Metal 2 Field Oxide	1.75µm
Polycide Resistivity	3.0 Ohms/square

TABLE 2 - Summary of process characteristics

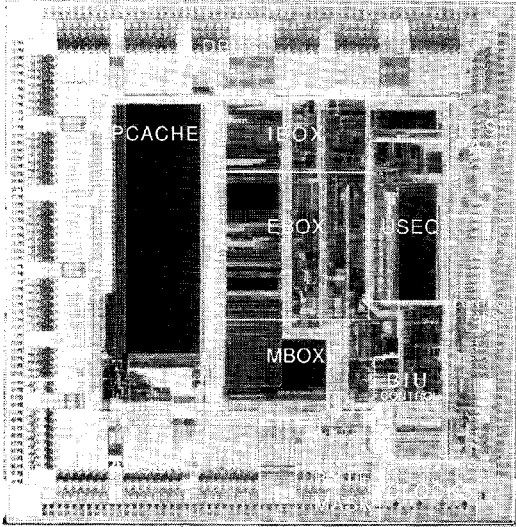


FIGURE 1 – Processor chip

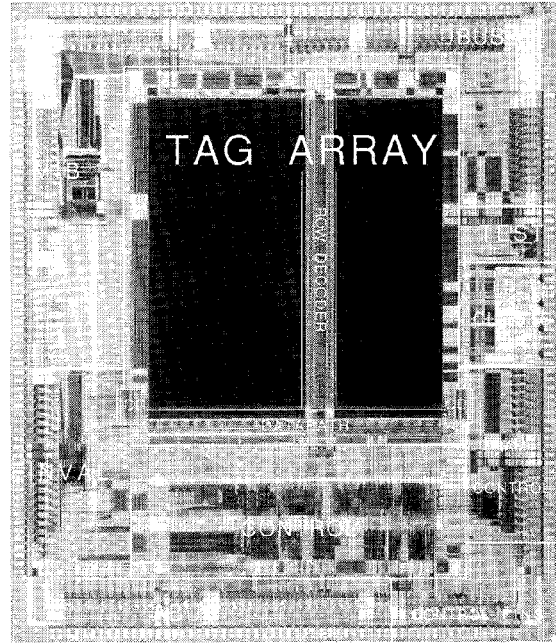


FIGURE 3 – Cache controller chip

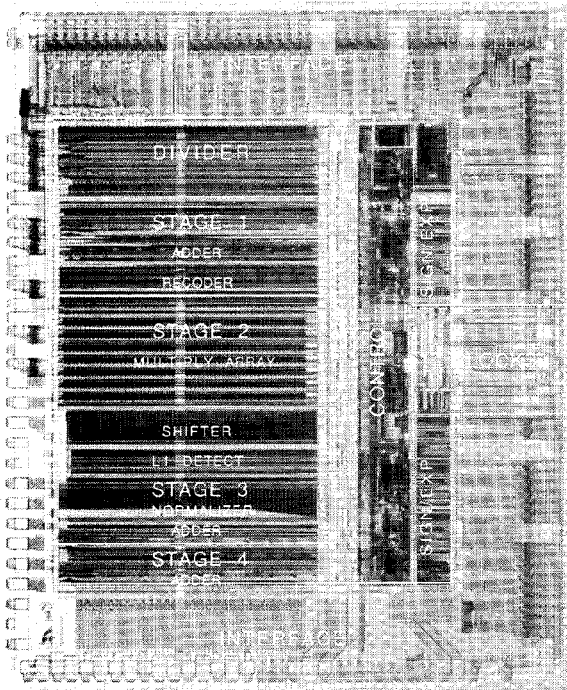


FIGURE 2 – Floating-point chip

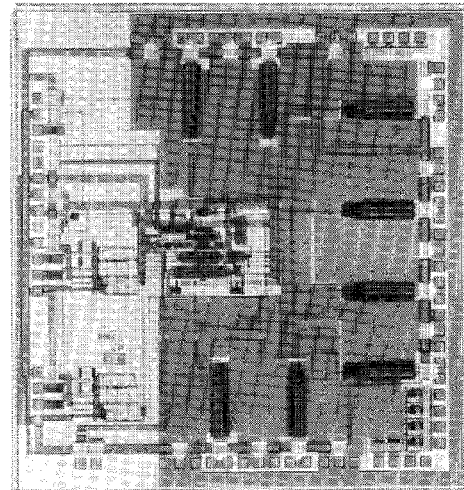


FIGURE 4 – Clock chip