Power Considerations in the Design of the Alpha 21264 Microprocessor

Michael K. Gowan, Larry L. Biro, Daniel B. Jackson
Digital Equipment Corporation
Hudson, Massachusetts

1.   ABSTRACT

Power dissipation is rapidly becoming a limiting factor in high performance microprocessor design due to ever increasing device counts and clock rates. The 21264 is a third generation Alpha microprocessor implementation, containing 15.2 million transistors and operating at 600 MHz. This paper describes some of the techniques the Alpha design team utilized to help manage power dissipation. In addition, the electrical design of the power, ground, and clock networks is presented.

2.   INTRODUCTION

Digital introduced the Alpha 21064 [1] microprocessor in 1992, thus delivering the industry's highest performance at that time. Manufacturing process technology advancements, architectural innovations, and full-custom circuit design techniques have been significant contributors to Digital's delivery of two additional generations of performance leadership Alpha microprocessors [2,3]. Between the first generation Alpha 21064 and the third generation Alpha 21264 designs, device counts have grown by a factor of 9, from 1.68 million to 15.2 million, while clock frequencies have increased from 200 MHz to 600 MHz, or a factor of 3. These rapid increases in available chip real estate and clock rates, combined with Digital's goal of delivering the highest performance possible, have caused significant increases in power dissipation and supply currents, which have had far-reaching effects on the clock, power, and ground networks, and on reliability. Consequently, power dissipation has joined clock frequency and die size as a first order design constraint. As die sizes and clock rates continue to increase, power dissipation may eventually limit the amount of hardware included on a microprocessor.

3.   SOURCES OF POWER DISSIPATION

Power is dissipated in CMOS circuits during both static and dynamic operating conditions. Ideally, a non-transitioning complementary CMOS gate dissipates no power because the PMOS and NMOS devices are never simultaneously conducting. However, in reality, a small leakage current flows through the reverse biased diodes formed by the source and drain diffusions and the substrate. More significant static leakage currents result from subthreshold conduction of the PMOS and NMOS devices. High performance CMOS processes specifically tuned for speed critical applications often feature very low threshold voltage devices. These low threshold devices have significantly increased subthreshold currents. Circuits designed with static loads may also dissipate power under DC operating frequencies.

Dynamic power is dissipated when the device output capacitance is charged to VDD through the PMOS device and discharged to VSS through the NMOS device. This power is given by the well-known equation $p = cv^2 f$. Note that while dynamic power dissipation is directly proportional to switching capacitance and frequency, there is a quadratic dependence on supply voltage. An additional source of dynamic power dissipation results from finite input signal edge rates. As an input signal transitions, there is a region of input voltages where the PMOS and NMOS devices are conducting simultaneously, thereby creating a low-resistance path from VDD to VSS. Careful control of the input and output edge rates of CMOS circuits can keep this crossover current component to less than 20% of the total dynamic power dissipation. Dynamic power also results from the glitching hazards inherent in cascaded complementary logic chains. Finite propagation delays in the logic stages can cause spurious glitches at the outputs of the complementary gates. In some cases, this glitching activity can be a significant source of power dissipation. Dynamic power is also consumed when subthreshold leakage currents experience exponential increases due to power supply noise.

4. ALPHA STATISTICS AND POWER AND CURRENT TRENDS

Table 1 lists some of the key features of the three generations of Alpha microprocessors.

	21064	21164	21264
Introduction	1992	1995	1998
SpecInt / FP95	3.6 / 4.0	7.5 / 12.0	40 / 60 (est.)
Transistor Count (millions)	1.68	9.3	15.2
Die Size (cm^2)	2.33	2.99	3.14
Process Technology (μm)	0.75	0.50	0.35
Power Supply (Volts)	3.3	3.3	2.2
Avg. Power Dissipation (Watts)	30	50	72
Avg. Supply Current (Amps)	9.1	15.2	32.7
Target Design Frequency (MHz)	200	300	600
Typical Gate Delays / Cycle	16	14	12
On-chip Cache	8 KB L1-I	8 KB L1-I	64 KB L1-I
	8 KB L1-D	8 KB L1-D	64 KB L1-D
		96 KB L2-U	
Instruction Issue / Cycle	2	4	6
Execution Flow	In-Order	In-Order	Out-of-Order

Table 1: Alpha Statistics and Trends

Process scaling has enabled a significant growth in the number of devices on a die. For example, a reduction in feature size from .50μm to .35μm results in an increase of approximately 60% more devices on a similarly sized die. Also note the nearly linear increase in average power dissipation of the microprocessors over time, despite the use of advanced processing technologies and scaled power supply voltages. Power supply currents have grown at an even faster rate. The table shows that supply current is nearly doubling with each new Alpha microprocessor generation. This trend of increased supply currents can be explained by a number of factors.

Each succeeding generation of microprocessors has been designed with an increasing number of execution units. The 21264, for example, integrates 6 execution units that can operate simultaneously. In addition, architectural features such as out-of-order execution, speculative execution algorithms, improved branch prediction, and compiler enhancements insure that these units are more effectively utilized than in previous generations. Therefore, relative to previous generation Alpha microprocessors, the 21264 has a higher switching activity rate in addition to the increased number of devices.

Another cause for the increase in currents is that each succeeding Alpha microprocessor generation has relied on advancements in circuit design techniques to increase the operating frequency by more than would be predicted by scaling alone. This effect can be seen by the reduction in typical gate delays per cycle from 16 for the 21064 to 12 for the 21264. This reduction in the number of gate delays per cycle was accomplished without reducing the effective work per cycle. The extensive use of large dynamic arrays, dual-rail dynamic logic, static ratioed logic, and other advanced circuit techniques has significantly contributed to this increased performance at the expense of increased power consumption. In addition, the 21264 has a larger software dependent variation of supply currents than in previous generations. These large increases in power and currents, combined with the increased variation of the magnitude of the currents, require careful design of the clock, power, and ground networks.

5. ALPHA 21264 POWER RESULTS AND ANALYSIS

A plot of average power versus operating frequency for the 21264 at a junction temperature of 100°C when running typical programs is shown in Figure 1.
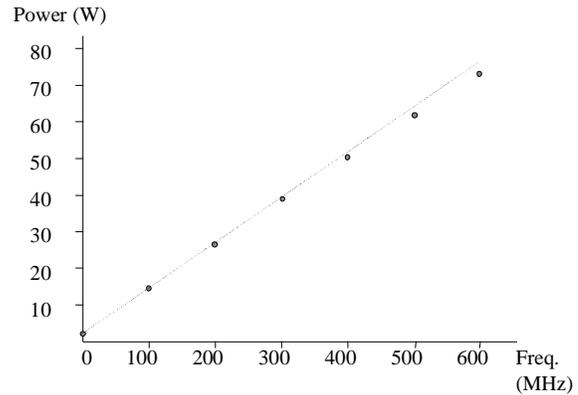

Figure 1: Alpha 21264 Power vs. Frequency

Note that power dissipated by the 21264 is approximately 2 Watts at static operating conditions. DC power dissipation of the 21264 is nearly fully attributed to subthreshold conduction. A limited number of circuits on the 21264, such as portions of the analog Phase Locked Loop (PLL) and pad input receivers, dissipate power continuously, independent of clock frequency. With the aid of CAD tools, the 21264 designers insured that no unwanted conduction paths exist between power and ground at DC operating conditions. In order to maintain functionality and limit power at DC frequencies, keeper devices were added to all dynamic nodes. These keeper devices supply the current necessary to keep the voltage of the dynamic nodes from varying significantly from the power supply rails. Power supply current at low frequencies would be expected to be much higher without these keeper devices because the degraded dynamic nodes driving other structures would otherwise create large crossover currents.

As expected, the graph shows power increasing nearly linearly with increasing operating frequency, with the power reaching approximately 50 Watts at 400 MHz. Power then starts to rise slightly less rapidly with increasing frequency until it reaches 72 Watts at 600 MHz. This slight fall-off in the rate of increasing power may be explained by the extensive use of low-swing differential

busses in the design. As the chip frequency approaches its functionality limit, these busses experience reduced voltage swings.

Table 2 estimates the components of total power in the 21264 design at the maximum operating frequency.

Global Clock Network	32%
Instruction Issue Units	18%
Caches	15%
Floating Execution Units	10%
Integer Execution Units	10%
Memory Management Unit	8%
I/O	5%
Miscellaneous Logic	2%

Table 2: Alpha 21264 Power Components

Note that the clocking network accounts for about one third of the chip's power dissipation. For previous generation Alpha microprocessors, nearly 40% of the chips power was attributed to the clocks. As will be described later, the reduction of clock power was the primary power savings focus for the 21264 designers. The table also shows that the out-of-order instruction issue logic dissipates nearly 20% of the total power, or about as much as the floating and integer execution units combined.

It is difficult to precisely determine the maximum power dissipation of a large microprocessor because the effective operating frequency of each circuit is dependent on its switching activity. A running program that maximally utilizes all of the circuits in a microprocessor will result in substantially higher power dissipation than a program that leaves much of the processor idle. In order to maximize switching activity within the 21264 to create a worst case for power testing, a special executable was created. This executable attempts to maximize switching activity of all internal components of the microprocessor by fully and continuously utilizing the instruction issue logic, all of the execution units, and major busses. It should be noted that this program performs no useful task and it generates no output. Because the development of this program required an intimate understanding of details of the internals of the microprocessor, the level of switching activity accomplished with this program is far higher than would be generated while running normal applications. The maximum power dissipated in the current 0.35µ process under these absolute worst case conditions was found to be about 95 Watts, which approaches the limit at which conventional air-cooling techniques can be used. The 21264 when running normal programs dissipates a far lower average power, as previously described.

In the next generation process, a VDD reduction significantly lowers chip power.

6. DESIGN TECHNIQUES TO REDUCE POWER ON THE ALPHA 21264

The 21264 designers used a number of techniques to help limit power dissipation. Lowering VDD has the largest impact on reducing power, but clock frequency is negatively impacted. As previously noted, the reduction of clock power is important to the management of overall chip power. The use of a hierarchical clocking scheme with conditional clocks was the primary design choice used to lower clock and logic power. Full-custom design methodologies also help enable power reduction. When circuits are designed at the transistor level, it is possible to use more innovative power reduction techniques

Previous Alpha designs utilized a single large clock grid that generated a timing reference for nearly all latches in the design. The Alpha 21164 did include some conditional clocking, however its use was limited to the large on-chip secondary cache. While straightforward and robust, a single wire clocking scheme causes many nodes to switch every cycle even though no useful work is being performed. The main benefit of conditional clocking is the lower average power that results from turning off sections of the chip that are not needed on a cycle-by-cycle basis. Moreover, peak power is also lowered because of logical exclusivities inherent in the design. Conditional clocking significantly complicates timing analysis of the design [4], however its use was justified because of the average and peak power savings.

The 21264 Floating Point Unit (Fbox) provides an example of the use of conditional clocking in datapaths. The basic clocking scheme for each operational unit (multiply, add, etc.) is indicated conceptually in Figure 2.
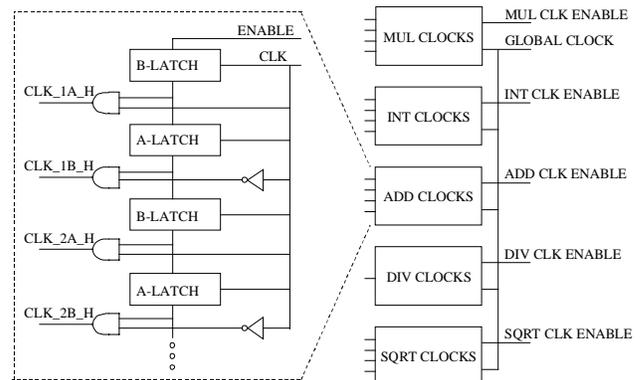

Figure 2: Floating Point Datapath Clocking

If, for example, the chip were to execute a single floating point ADD, the control logic would assert the adder clock-enable line for one cycle. The pulse would then propagate

through a chain of alternating active-low (B-latch) and active-high (A-latch) level-sensitive latches. Each latch output is used to enable the latch bank clock driver, indicated conceptually by the AND buffer, for the corresponding datapath pipeline stage. The data flow down the datapath with the clock enable signal preceding them so the latch banks at successive stages are clocked as needed. As pipeline stages of the datapath are no longer required, the de-assertion edge of the enable input signal propagates down the latch chain, successively turning off the latch bank clocks and saving power that would otherwise be wasted in performing unneeded work. If the Fbox were performing only a floating point ADD, clocks for the other operational units would remain disabled. The Fbox control logic is unconditionally clocked so it can respond to incoming floating point instructions with minimal delay.

Due to difficulties in managing circuit race issues, the entire divider datapath is operated on one conditional clock, rather than by pipeline stage as in the other operational units. The net power saved by these techniques depends on the floating point instructions being executed. For the case where no floating point instructions are being executed, Fbox power dissipation is estimated to be about 25% of what it would have been without conditional clocking. As floating point instructions are added to the instruction stream, Fbox power dissipation increases.

Low-swing busses were used extensively on the 21264 in order to help meet performance goals. An additional benefit of these busses is their reduced power consumption. Low-swing busses save power by reducing the voltage excursions on large capacitances to hundreds of millivolts instead of the full power supply range. The Execution Unit (Ebox), as shown in Figure 3, provides an example of the usage of low-swing busses in the 21264 design.
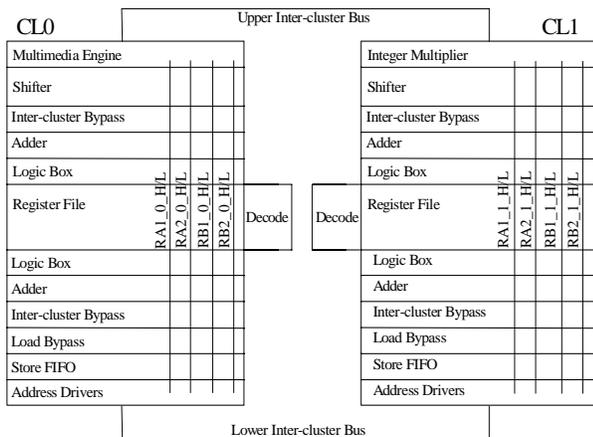

Figure 3: Execution Unit Low-Swing Busses

The Ebox is divided into two clusters, CL0 and CL1. Each cluster contains an 80-entry register file plus two independent execution pipelines containing arithmetic, shift, and multimedia functional units. Functional unit results, register file contents, load data, and inter-cluster results drive 64-bit limited-swing, differential operand busses (depicted as RA1_0, RA2_0, RB1_0, RB2_0, RA1_1, RA2_1, RB1_1, RB2_1 in the diagram).

Had these busses been implemented with standard full-swing techniques, it is estimated that their worst case dynamic power dissipation at the 600 MHz target operating frequency would have been about 1.5 Watts. The resulting power estimate with the use of limited swing busses under these same conditions is less than 15mW, or a reduction of two orders of magnitude. The bus receivers were designed to operate at 200mV of differential swing, thus requiring careful circuit and layout design to minimize all differential noise effects.

Performance requirements dictated the extensive use of high performance dynamic structures. However, dynamic structures have several properties that can help reduce power. Dynamic circuits have reduced fanin and fanout capacitances because there is no need for the complementary PMOS network. In addition, dynamic circuits have smaller layout areas relative to equivalent static complementary structures, resulting in lower wiring capacitances. Dynamic circuits also generate monotonic signal transitions, so power is not dissipated by glitching hazards. On the 21264 design, many dynamic nodes are conditionally precharged only a cycle or phase before the circuit is to evaluate. This optimization lowers the effective switching capacitance of the clock network by keeping the clock node from needlessly driving the gates of the precharge devices. In addition, extensive device-level simulations were performed to insure that the overlap between dynamic node precharge and evaluation is minimized, thereby reducing crossover currents.

Static complementary logic was used in many large control sections of the chip. Spurious input transitions to large logic sections were limited in order to keep unused signal transitions from propagating through the networks. Additionally, significant power savings was also achieved by reducing the I/O voltage from 3.3 Volts to 2.2 Volts.

7.   IMPLEMENTATION TECHNIQUES FOR HIGH PERFORMANCE DESIGN
7.1   Power and Ground Distribution
As previously described, supply current is roughly doubling with each new generation. This dramatic increase in current has required the implementation of increasingly complex on-chip power distribution schemes. In addition, the scaling of power supply voltages to limit

power dissipation has also reduced the amount of absolute supply noise that can be tolerated. A high performance power distribution scheme must allow for all circuits on the die to receive the same voltage reference. A variation of supply reference across the die can lead to problems such as reduced noise margins, reduced device switching speed, increased subthreshold conduction, or possibly even latchup. An additional factor which stresses the 21264 power distribution scheme is the extensive use of conditional clocks, which exaggerates cycle-to-cycle current variations. The 21264 power supply network was designed to handle a maximum variation in power supply current of up to 25 amps between adjacent cycles.

With these severe constraints, it was demonstrated that a standard distribution scheme of a two dimensional grid of course pitch interconnect for routing power and ground would be insufficient. The grid would not adequately prevent a power supply drop in the chip's center due to the chip's large average current. In addition, the program dependent variable peak currents further exacerbate the problem. For this reason two thick, low resistance aluminum reference planes were added to the .35μm 21264 process. Reference plane 1 was added between metal 2 and metal 3 and was connected to VSS. Reference plane 2, which was connected to VDD, was added above metal 4. The use of power planes had several beneficial effects. First, a stable power reference is available to circuits situated anywhere on the die. Secondly, the lower reference plane inductively and capacitively isolates the metal 2 and metal 3 signal lines, effectively reducing crosstalk.

The additional power plane layers significantly reduce the DC IR voltage drop to the center of the die, but the high clock frequencies and conditional clock buffers cause large cycle-to-cycle fluctuations in supply current. These variations of current flowing though the package and bond wires can cause significant power supply ringing on the chip. Therefore, substantial amounts of on-chip decoupling capacitance are used to smooth out the supply current fluctuations though the bond wires. These decoupling capacitors, which are constructed using the gate oxide of an NMOS transistor, are used extensively around large clock drivers, under large signal busses, and in other unoccupied areas of the die. A total of 0.32μF of decoupling capacitance was added the 21264 die, and it occupied roughly 15-20% of the die area. Extensive simulations showed that the on-chip decoupling capacitance alone was not sufficient to control the power and ground fluctuations. Therefore an additional 1μF, 2cm^2 Wire-bond Attached Chip Capacitor, or WACC, was added to the chip and package network. It should be noted that advanced packaging techniques such as flip chip with solder bumps can be used to provide a more robust power distribution scheme than can be achieved with stitch wire bonding techniques.

7.2 Clocking Network

The high frequencies of the Alpha microprocessors have required the generation and distribution of very high quality clock signals. As more aggressive circuit techniques and complex micro-architectural features were implemented on the 21264, power consumption became a major concern in designing the clocking system. The 21264 uses a single wire global clock, called GCLK, as the chip's main timing reference [3,5]. Performance goals required that the overall GCLK skew be limited to less than 75ps, with steep, uniform edge rates, and a duty cycle approaching 50%. These goals were accomplished with a modified H-tree GCLK distribution scheme shown in Figure 4.
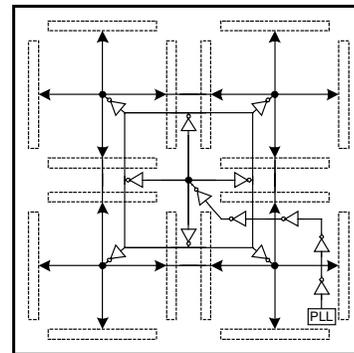

Figure 4: Alpha 21264 GCLK Network

The PLL output is repeated, driven to the approximate die center, and then routed vertically and horizontally to buffers midway between the die center and die edges. From these locations, the clock is driven to the centers of each quadrant of the chip where a second set of buffers is located. Here, the clock is inverted and driven to the peripheries of the quadrants where sixteen 2-stage distributed buffers drive the GCLK grid. The GCLK grid consumes approximately 3% of the total available metals 3 and 4.

GCLK was used as the basis of a hierarchy of more than ten thousand buffered and conditioned clocks used across the chip. Most clock load is isolated from GCLK by double buffering the load onto large buffered or conditioned section clocks. These buffered section clocks are also the basis for additional, smaller buffered or conditioned clocks. This clock distribution scheme allowed designers freedom to use conditioned clocks wherever significant power savings could be accomplished. In addition, the improved locality of the clock drivers with respect to their loads allowed for significant reductions in skews and metal usage. This clock generation and

distribution scheme results in significantly lower metal 3 and 4 clock usage, saving an estimated 10 Watts over previously used gridded clocking schemes. Note that the choice of skew targets greatly affects overall clock network loading. A less aggressive skew target can be achieved with a sparser grid and smaller drivers.

An additional benefit of this distributed clock generation scheme is the reduced thermal gradient over the centralized clocking approach. The 21264 achieves a thermal resistance from the junction to the heatsink of 0.3°C/W, whereas the previous Alpha design using the same packaging technology achieved a θ of .4°C/W. The distributed clock drivers also relieve stress on the power and ground distribution schemes by spreading out the large drivers across the die. As previously described, the major disadvantages of this hierarchical clocking scheme is the potential for large current gradients on a cycle by cycle basis as the conditional clocks are successively enabled and disabled, and the increased complexity of the timing analysis.

7.3   Electromigration Reliability

The increase in supply currents in each succeeding generation of microprocessor design has necessitated more sophisticated electromigration analysis, especially in the power, ground, and clock networks. Traditionally, current densities through all interconnect segments of a signal were required to be below defined threshold limits that depended on the interconnect type. These thresholds were based on current densities that had been historically determined to be safe. If all interconnect segments of a design met this electromigration specification, the entire design was assumed to be safe against electromigration risk. Increased currents have caused this approach to be increasingly difficult to implement. Additionally, this approach does not take into account the increased risk of premature electromigration failure if an excessive number of interconnect segments of a design are at or near the maximum current limit.

The 21264 microprocessor design team used a technique called Statistical Electromigration Budgeting (SEB) [6]. SEB takes into account the statistical nature of electromigration failures. With SEB, an electromigration failure probability, which depends on the current density through the interconnect segment, is calculated for each interconnect segment. The probability of an electromigration failure of the entire design is then a function of the failure probabilities of all the individual interconnect segments. SEB allows a limited number of interconnect segments to have larger currents. Although

any individual interconnect segment with a larger current density has a higher probability of failure, a small number of these segments will have little affect on the failure probability of the whole design. SEB does impose an absolute maximum current density limitation on each interconnect segment, which if exceeded, may cause a catastrophic failure due to Joule heating.

8.   CONCLUSIONS

As demonstrated in the three generations of high performance Alpha microprocessors, increased clock frequencies and dramatically higher device counts have caused rapid increases in power and current consumption. In order for microprocessor performance to keep increasing at historical rates, careful consideration must be given to power consumption, power and ground routing, and the clocking networks. One of the main challenges in future microprocessor design will center around devising new circuit techniques that help lower power without impacting overall circuit performance.

9.   ACKNOWLEDGMENTS

10.   REFERENCES

[1] Dobberpuhl, D., et al., "A 200 MHz 64b Dual-Issue CMOS Microprocessor," IEEE Journal of Solid State Circuits, vol. 27, no. 11, Nov., 1992.

[2] Benschneider, B., et al., "A 300-MHz 64b Quad-Issue CMOS RISC Microprocessor," IEEE Journal of Solid State Circuits, vol. 30, no. 11, Nov., 1995.

[3] Gieseke, B., et al., "A 600 MHz Superscaler RISC Microprocessor With Out-of-Order Execution," ISSCC Digest of Technical Papers, pp. 222-223, Feb., 1996.

[4] Grundmann, W., et al., "Designing High Performance CMOS Microprocessors Using Full Custom Techniques," Proceedings of the 34th Design Automation Conference, pp. 722-727, June, 1997.

[5] Fair, H. and Bailey, D., "Clocking Design and Analysis for a 600 MHz Alpha Microprocessor," ISSCC Digest of Technical Papers, pp. 398-399, Feb., 1998.

[6] Kitchin, J., "Statistical Electromigration Risk Budgeting for Reliable Design and Verification in a 300MHz Microprocessor," Digest of Technical Papers, VLSI Circuits Symposium, 1995.