

IC Power Distribution Challenges

Sudhakar Bobba, Tyler Thorp, Kathirgamar Aingaran¹, and Dean Liu
Sun Microsystems, Inc. ¹Currently with Afara Websystems

Abstract

With each technology generation, delivering a time-varying current with reduced nominal supply voltage variation is becoming more difficult due to increasing current and power requirements. The power delivery network design becomes much more complex and requires accurate analysis and optimizations at all levels of abstraction in order to meet the specifications. In this paper, we describe techniques for estimation of the supply voltage variations that can be used in the design of the power delivery network. We also describe the decoupling capacitor hierarchy that provides a low impedance to the increasing high-frequency current demand and limits the supply voltage variations. Techniques for high-level power estimation that can be used for performance vs. power trade-offs to reduce the current and power requirements of the circuit are also presented.

1. Introduction

Computations in integrated circuits are driven by the energy from a DC power supply. In order to perform computations, this energy is transferred, stored and then dissipated as heat in the integrated circuits. Fig. 1 shows this process, from energy delivery to the heat transfer out of an integrated circuit. The rate of computations and the number of computing elements in an integrated circuit determine the rate of energy delivery or the power delivery requirements of the circuits. Since the rate of computations is time-varying, the power and current requirements of the integrated circuit will also vary with time. Delivering a time-varying current at a constant supply voltage with nominal variations is the goal of the power distribution network design.

Technology scaling over the past few decades has enabled integrated circuits to speed up the computation rate (with increasing clock frequencies) and increase the number of computing elements (through parallelism) at the cost of higher power dissipation [1][2][3][4][5]. The trend of increasing power and clock frequency while reducing power supply voltage causes the power supply network to experience larger di/dt noise. In modern deep-submicron technologies, the supply voltage variation greatly affects the delay of digital circuits and can push transistors out of saturation in analog circuits. Since device threshold voltage does not scale well with the reduced supply voltage in scaled technologies, the gate current drive becomes more sensitive to the supply voltage variations. As a result, circuit

designs require that the supply variation remain at least a fixed percentage of the supply voltage. With each technology generation, the larger di/dt noise makes it difficult to ensure that the supply variation does not exceed this fixed percentage.

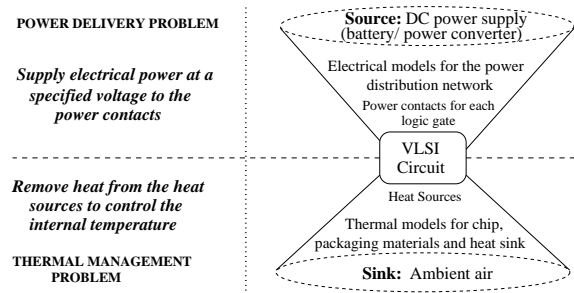


Fig. 1 Energy Cone

Process scaling allows faster switching devices that result in faster current transients which contain higher frequency components. The frequency spectrum of an integrated circuit's current demand usually contains large components at the clock frequency due to the synchronous switching events and at twice the clock frequency due to the clock distribution. For many integrated circuits, a significant fraction of the power is dissipated by the synchronous elements and the clock distribution network [5]. The impedance of the power distribution network at higher frequencies will have to be reduced in order to supply the higher frequency currents without causing significant supply voltage variations. One way to achieve this is to use more decoupling capacitors and place them closer to the switching transistors.

There are several techniques at different levels of abstraction that reduce the supply voltage variations and lead to the design of a robust power distribution network. High-level power estimation techniques that aid in the power vs. performance trade-off early in the design cycle and a good analysis of the power delivery network ensures a non-over-constraining design. Post-layout verification of the supply voltage variations and the use of decoupling capacitors to reduce these variations are required in the later design stage. In this paper, we describe several techniques and problems for the design of the power distribution network in high-performance integrated circuits.

This paper starts by describing power grid analysis

techniques in Section 2. This is followed by a discussion of the decoupling capacitor hierarchy in Section 3. In Section 4, we describe high-level power estimation techniques. In Section 5, we present power vs. performance trade-offs. Finally, we offer some concluding remarks.

2. Power Grid Analysis

The power distribution network is required to reliably deliver power over the lifetime of the integrated circuit with nominal voltage variations. The two main concerns in the design of the power grid are the transient voltage variations at nodes in the power grid and the long term reliability of the interconnects in the power grid. The transient voltage variations are caused by the flow of time-varying current through the power distribution network. These voltage variations degrade noise margins (possibly causing functional failure), increase the delay of logic gates, increase clock skew, and reduce gate oxide reliability. The interconnects of the power grid become susceptible to electromigration (EM) induced failure when they carry high currents over a sustained period of time. Several CAD tools exist to identify the voltage drop and EM problem locations. The power grid can then be redesigned to enhance its reliability. In this section, we describe the problems of supply voltage variations, EM, and design techniques for a reliable power grid.

2.1 Voltage variations in the power grid

Power grid voltage variations occur due to the flow of time-varying current through power grid interconnects that contain parasitic resistance, capacitance, and inductance. The dominant component of the current is drawn by logic gates in CMOS circuits while making logic transitions. Since the logic gates in a circuit share the same power distribution network, the switching of a set of gates can cause power supply voltage variations at another gate's power supply contacts. A drop in the power supply voltage at the logic gate power supply contact points can decrease the drive strength of the logic gate resulting in increased delay. The dependence of the delay of a gate on other switching gates is hard to model and increases the complexity of timing verification. In typical design flows, this undesirable electrical interaction is treated as power grid noise and a limit is put on the allowable voltage variations at nodes in the power grid. The power grid is then designed to meet this voltage variation budget, and the timing analysis is performed assuming a lower bound on the supply voltage.

High-performance integrated circuits require a robust power delivery network with nominal supply voltage fluctuations. The voltage variations are dependent on the power grid impedance at different frequencies. The design of a robust power distribution network requires the realization of a network with low impedance at all frequencies that

can be excited by the current waveforms. Fig. 2 shows the locations for the source of charge at different frequencies and the associated current loops. Accurate models are required to capture the worst-case transient voltage variations for all frequencies.

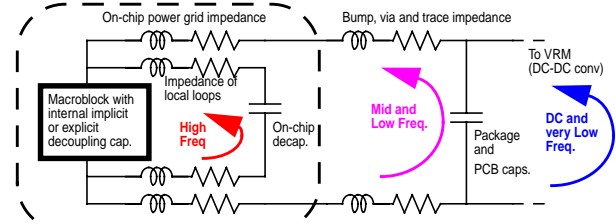


Fig. 2 Locations for source of charge at different frequencies

Estimation of the worst-case voltage variation is a difficult problem because of the size and complexity of the power supply network and the input pattern dependence of the current drawn by the circuit. The number of nodes in the power distribution network can be extremely large because the power distribution network connects to every transistor in an integrated circuit. However, a power grid is typically designed as a hierarchical structure in which the top-level power-grid connects to the macroblocks and the power distribution network inside the macroblock connects to the logic gates. The hierarchical analysis reduces the problem size and complexity.

The input pattern dependence of the current drawn by logic gates makes the problem of estimating the maximum current or maximum voltage drop difficult to solve. In order to reduce the complexity, the current drawn by the logic gates in a macroblock can be abstracted as a current waveform and the power distribution network can be analyzed with that waveform. This current waveform can be obtained using static approaches or dynamic approaches such as input vector simulation or by pattern independent techniques [6][7][8][9].

Existing techniques can be classified based on the models used for the power grid. The simplest electrical model for the power distribution network is the resistive power bus model. Although it simplifies the power bus analysis, the results are accurate only if the resistive effect dominates the capacitive and inductive effects. The power grid analysis techniques using the resistive model [10][11][12][13] compute the IR drop at nodes in the power distribution network by estimating macroblock currents in one of the following ways: a DC current obtained heuristically, a DC current obtained by logic simulations using input vectors, or a transient current waveform obtained by simulations using input vectors. A resistive model for the top-cell power distribution network does not take into account the presence of significant on-chip decoupling capacitance that helps to reduce the voltage drops. In

contrast, ignoring the inductance can result in an under-estimation of the supply voltage variations. The power grid analysis techniques using the RLC model [14][15] heuristically construct a triangular or trapezoidal macroblock current waveform using peak and average currents or the current waveform obtained by simulations for a few vectors. This macroblock current waveform is then used to estimate the voltage drop at nodes in the power distribution network through simulations or heuristic algorithm that performs a table lookup of a pre-characterized waveform library. Using a lower bound current waveform to excite the RLC power distribution network does not yield the maximum voltage variation in the power grid. In [16], pattern independent maximum envelope currents are used in a frequency domain technique for estimating the worst-case time-domain voltage variation using RLC models for the power distribution network. The results can be pessimistic if the envelope currents are not modeled accurately. In general, pattern dependent techniques generate a lower bound on the maximum voltage drop and the pattern independent techniques generate an upper bound on the voltage drop. The accuracy of each technique is dependent on the models and the current excitation used in the analysis.

2.2 Electromigration in power wires

Electromigration (EM) is the flow of metal ions under the influence of high electric current densities resulting in the depletion and accumulation of metal ions along the interconnect. Although metal migration causes voids and hillocks along the interconnect, electrical connectivity may still be maintained through the barrier metal layer which is resistive and more immune to electromigration. In power grid wires, the increased resistance due to EM can result in larger IR drops and degradation in gate delay.

Degradation and failure of a device are very complex and are commonly modeled as statistical phenomenon using empirical models based on experiments and/or simulations. The primary stress factors that accelerate EM induced degradation and failure of interconnects are the temperature and the current density through the interconnect. EM is also dependent on the length and width of wires.

The reduced dimensions of interconnects, contacts and via's with scaling can result in higher current densities through the interconnects and contacts. Current crowding in via's can cause local hot-spots and accelerate EM [17]. In order to meet the power requirements of future process generations, the power grid must be designed to withstand higher current densities. In addition, a higher operating temperature due to joule heating or thermal coupling accelerates EM induced degradation.

In order to identify interconnects that are susceptible to EM, estimates of the average current density (J_{avg}) for all the interconnects is required [18]. Existing CAD tools screen wires based on the average current density to identify

wires that are susceptible to EM induced failure. These wires can be resized to reduce the current density through them.

The failure mechanism for short wires is different from that of long wires. It is possible to design power grids with short wires that are immune from EM problems [19]. Tools and analysis techniques must comprehend the fundamental mechanism of failure and account for all parameters that affect the EM reliability of the interconnects. Tools should also take into account statistical EM budgeting rules to design EM reliable systems [20].

2.3 Design techniques for a reliable power grid

The reliability of the power grid can be enhanced by using more metal for the power grid. However, in high-performance integrated circuits, signal wires and power wires compete for routing resources. This implies that we need accurate estimation techniques to identify problem locations. Tools and techniques to predict these problems early in the design cycle are also required. Wire width adjustment and placement of decoupling capacitors are commonly used techniques to fix problems within the power grid.

3. Decoupling Capacitors (Decaps)

The general trend of increasing integrated circuit power and frequency while reducing power supply voltage is causing on-chip di/dt to increase with each technology generation [21]. In traditional on-chip power supply designs, where the package can supply current quickly to the chip, the voltage variations due to resistive IR drops primarily occur on-chip and the inductive noise primarily occurs in the package [14] [22]. As on-chip di/dt continues to increase, a more detailed analysis of the decoupling capacitor hierarchy extending well into the chip will be needed.

Fig. 3 shows a general decoupling capacitor hierarchy which can be used to give a specified target impedance across a broad range of frequencies. The components of the decoupling capacitor hierarchy can include the board, package, and on-chip decoupling capacitors. Several on-chip parasitic decoupling capacitors exist between metal wires and between nwell and substrate. A part of the output capacitance of non-switching logic gates appears between the supply rails and acts as implicit decoupling capacitance.

The decoupling capacitor hierarchy consists of several sections. Each section is composed of a parasitic inductor and the corresponding downstream (i.e. closer to the logic gate) capacitor. For instance, the section corresponding to the package would have the package inductor and the on-chip decoupling capacitor as part of the section. Each section of the off-chip power supply network acts as an under-damped, second order low pass filter. In contrast, the

on-chip sections have a dominant resistance and act as an over-damped, second order low pass filter.

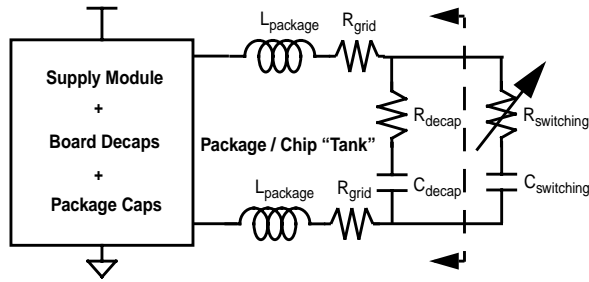


Fig. 3 Equivalent circuit for low freq.

The decoupling capacitor downstream must be large enough to supply high frequency current above the cutoff frequency of a section. If the downstream capacitor is not large enough, then the high frequency current that is not supplied by the capacitor will flow through the inductor of the section resulting in appreciable di/dt variations. The downstream portion of the network must also be sufficiently resistive at the resonant frequency of the section to damp any oscillations. In summary, decoupling capacitors of each section act as local charge reservoirs to reduce the peak current drawn through the inductor of each section, thereby reducing IR drop and di/dt noise.

Fig. 4 shows the equivalent circuit for the on-chip sections. When on-chip circuits switch, the initial current must be supplied through local on-chip decoupling capacitor. This can be viewed as a charge sharing event between the local decoupling capacitor and the logic gate load being charged. When logic gates switch, all the high frequency currents are supplied by the on-chip decoupling capacitors and the lower frequencies are supplied from the package. The inductive component of the on-chip interconnect impedance becomes larger with increasing on-chip di/dt and cannot be ignored in high frequency simulations [23].

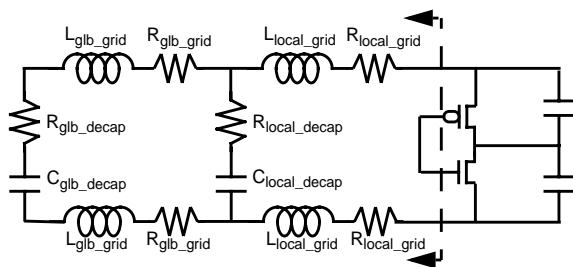


Fig. 4 Equivalent circuit at high-freq.

Recall that each off-chip section is an under-damped second order low pass filter. This implies that a resonance could be excited depending on the currents flowing through the section [24]. The section composed of the package

inductor and the on-chip decoupling capacitance, whose resonant frequency is closest to the clock frequency, is usually the most significant resonance. If the resonant frequency of this section is much higher than the clock frequency, then it is unlikely that oscillations would occur. However, if the resonant frequency is less than the clock frequency, then the resonance may be excited. Therefore, adding more on-chip decoupling capacitance is not always beneficial since it lowers the resonant frequency and may cause resonant oscillations in the power distribution network. On the other hand, increasing the on-chip decoupling capacitance also reduces the quality factor which in turn reduces the power supply impedance and therefore may be beneficial. When the power grid resonates dangerous voltage variations may occur that can cause gate-oxide failure. For designs with resonant frequency less than the clock frequency additional design constraints should be in place to guarantee that power supply network does not ring (e.g. any repeating sequence of instructions that cause current excitations at the resonant frequency should be avoided).

3.1 Decap implementation and verification

An effective way to implement on-chip decoupling capacitors is to use the gate capacitance of transistors. Each decoupling capacitor has a parasitic series resistance which impedes the flow of charge. A reasonable RC time constant for the decoupling capacitors can be achieved using a channel length that is roughly 10 times larger than the minimum channel length [25]. In scaled technologies, the dielectric leakage of thin gate oxide devices will be more significant and may become a limiting factor for its usage. Thicker gate oxide devices can be used to reduce the dielectric leakage at the cost of reduced capacitance per unit area and increased parasitic resistance per unit area. In order to dampen any resonance with the package inductance one can use a combination of high resistance and low resistance on-chip decoupling capacitors.

When logic gates switch, the required charge is delivered by the local decoupling capacitors close to the logic gate. As the local decoupling capacitor is moved farther away, the impedance of the current loops increase. This results in a smaller effective supply voltage to the logic gate which may increase its delay. In order to ensure that the delay of switching logic gates is not affected, sufficient local decoupling capacitors must be placed within a certain distance from all switching gates. This is the on-chip decoupling capacitor verification problem.

A first-order solution to this problem can be attained by partitioning the chip into regions. For each region, determine the amount of decoupling capacitance in the region and the amount of decoupling capacitance required within a certain distance from the region. This problem can then be formulated as a linear programming problem and

solved using traditional methods. Fig. 5 shows an example linear programming (LP) formulation for the decoupling capacitor verification problem. The number in the subscript of the variables denotes decoupling capacitors and the alphabet denotes drivers. For example, variable X_{1a} represents the percentage of the decoupling capacitance $Decap(1)$ attributed to the driver $Driver(a)$.

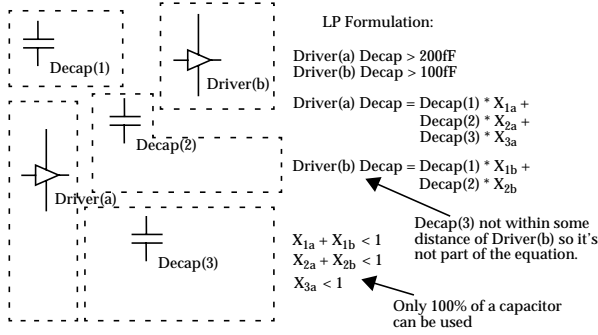


Fig. 5 LP formulation for decoupling capacitor verification

4. Power Estimation

Accurate estimates of power dissipation are necessary at various stages of the design cycle in order to make the correct architectural and implementation trade-offs. Circuit level power estimates can be obtained using SPICE-like simulators [26] on pre-layout or post-layout databases. Unfortunately, this type of analysis is done later in the design cycle when it may be too late to make architectural decisions. Therefore, methods for power estimation at higher levels of abstraction, namely, algorithmic, system, architectural, behavioral, and register transfer (RT) levels are required.

The dynamic power dissipated by CMOS logic gates can be expressed as, $P = \alpha CV^2f$, where α denotes the activity factor with respect to clock like signals, C denotes the switched capacitance, V denotes the supply voltage, and f denotes the clock frequency. Power estimation at the logic gate level requires the estimation of the activity factor for each logic gate. These values, along with the load capacitance for the logic gates, can be used to estimate the power dissipated by the circuit. Several gate-level power estimation techniques are described in [27]. Although gate-level power estimation is often very accurate, it may be too late or too expensive to go back and fix high power problems. In order to avoid the costly redesign steps, power estimation at higher levels of abstraction is needed.

At a higher level of abstraction, the power models consist of an estimate of the switched capacitance for different cases and estimates of the activity factors for each case [28][29][30][31][32][33][34]. For instance, instruction-level power estimation requires an estimate of the capacitance switched or energy consumed for each type of

instruction. These values can be weighted with the probability of each instruction to estimate the total power dissipation [34]. Activity factor estimation may be done using architectural information or RTL level toggle tracking tools while the capacitance values continue to evolve with each design integration. Taking into account only dynamic switching power neglects the contributions due to short circuit current, static leakage, and internal node glitching for which a correction factor may be added. Estimates of internal power can be provided by characterization and/or statistical estimation.

Increasingly, there is a need for fast high level power estimation, since power dissipation is now an important architectural consideration [1][31][35][36][37]. Power estimation based on architectural models uses more realistic estimates of block activity based on actual instruction traces [31][32][33]. However, it has no knowledge of the data behavior, which leads to errors in the measurement of data dependent power dissipation, where an average case has to be assumed all the time.

In summary, power estimation at the lower levels of abstraction involves circuit level and logic gate level simulations. Circuit level simulations estimate power dissipation given a set of input vectors. On the other hand, logic level simulations count the toggles, estimate the activity factors for the logic gate, and appropriately weigh the logic gate load capacitance to estimate the total power dissipation. At a higher level of abstraction, the block power (effective switched capacitance) can be estimated for different types of input patterns/accesses and the probability of each type of access can be used to estimate the total power dissipation. As the power estimates become more abstract, the accuracy decreases. However, high level power estimation facilitates high level design space exploration and can lead to significant power savings.

In addition to average power estimation, the power model may also be used to determine several other metrics relevant to power distribution. Tools and techniques are required to track cycle to cycle power variations, die temperature variations and to determine the amount of decoupling capacitance needed to stabilize the power supply.

5. Power vs. Performance Trade-offs

As emphasized throughout this paper, power dissipation is an important consideration in the design of an integrated circuit. The availability of architectural and implementation level power estimation tools can allow designers to make trade-offs between power and performance. However, in addition to power modeling and estimation tools, one needs objective criteria to determine the best *power/performance* point at which the design should be positioned. In this context, a "design" is a compound of architectural, microarchitectural, technological and physical features. As an

example, in the *power/performance* plot shown in Fig. 6, suppose we are evaluating two design alternatives A and B with relative *power/performance* metrics. Which is the better design choice, A or B?

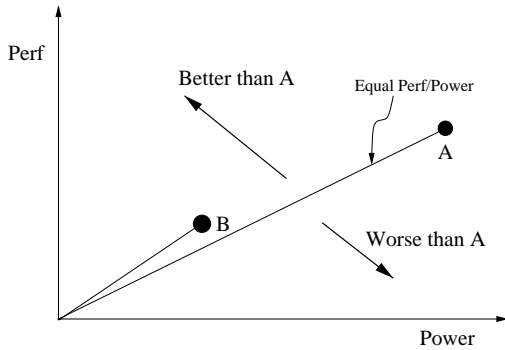


Fig. 6 Optimization with linear performance/power trade-off curve

Traditionally, *power/performance* (or power-delay) has been used to evaluate designs. Thus, all designs on the equal *power/performance* line in Fig. 6, correspond to a design as good as A, while designs on a steeper line such as B correspond to a better design. For evaluating an architecture, energy-delay (i.e. $power/[performance^2]$) may be a more suitable metric [38], possibly with a technology scaling factor also taken into account. Other metrics such as $power/[performance^3]$ have also been proposed as a better metric, which is independent of the supply voltage [37]. These metrics are for evaluating an architecture, but not for a "design" as defined above.

The optimization criteria for a compound "design" would in fact depend on the target application of the integrated circuit. For a given target application, a power related cost function, C_{power} , may be defined. For example, for a high performance server, the power related cost function is usually the capacity of the power input cone in Fig. 1. The server is limited by the amount of current (or power) that may be pumped into it. Thus, we define the cost function to be $C_{power} = power$. A similar cost function can be obtained for servers limited by the power output cone in Fig. 1, where only a given number of Joules of heat can be drawn out of the chip by the cooling system every second.

The power cost function should be optimized against a targeted value heuristic, V , which is important to the application. The value heuristic for a server, is pure performance and may be defined to be $V = performance$. Therefore, a server design should attempt to minimize $C_{power}/V = power/performance$, which may be achieved by the appropriate selection of voltage, frequency, technology and architecture. Design B in Fig. 6 is thus determined to be the better design choice for a server.

On the other hand, when the source of the energy cone

shown in Fig. 1 is the limiting factor (e.g. the energy utility bill incurred by the server is the primary cost), we would define $C_{power} = energy$ and $V = performance$. Therefore, the optimization function would be $C_{power}/V = energy/performance = power/[performance^2]$. This criteria is the same as in [38]. Energy-delay may also be used in energy limited mobile applications such as wireless modems where the battery life (i.e. energy) is the primary component of the cost function. Fig. 7 illustrates equal *power/performance* points for such an application. In this case, the design A is judged to be better than the design B.

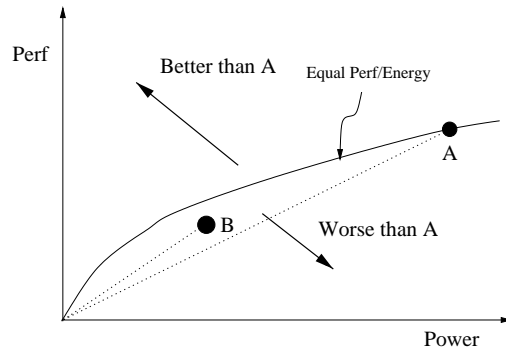


Fig. 7 Optimization with a quadratic performance/power trade-off curve

In other mobile devices, such as cell phones, the battery life or energy has to be traded off against the number of operations performed (i.e. minutes of talk time). In this case $C_{power} = energy$, $V = operations$, and $C_{power}/V = energy/operation$. Fig. 8 shows the trade off for laptop PCs where the cost function can be defined to be power, but the value heuristic is not a linear relationship with performance. It is assumed that the value heuristic is in fact proportional to the log of performance, since a doubling of performance in these devices typically leads only to an incremental increase in value. Therefore, $C_{power} = power$, $V = log(performance)$, and $C_{power}/V = power/log(performance)$.

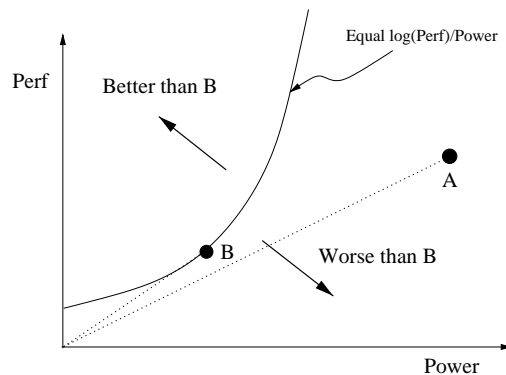


Fig. 8 Optimization with a logarithmic performance/power trade-off curve

Using this criteria, it is seen from Fig. 8, that once some minimum performance goal is met, it is better to optimize for power without much regard for performance. Design B in this case is the better design compared to A.

Interestingly, high performance design in the past has assumed $C_{power} = \text{constant}$, $V = \text{performance}$, $C_{power}/V = \text{constant}/\text{performance}$, which led to $1/\text{performance}$ being minimized (i.e. performance being maximized with no regard for power). In the above figures, this approach would determine A to be the better design choice.

While simple functions were used in the examples above to model C_{power} and V , it is acknowledged that they may be more complex for real applications. For example, C_{power} would typically have a ceiling value that cannot be crossed, while V would have a similar floor value. However, once these functions have been defined, a similar optimization could be done to determine the *power/performance* point most suited to the target application.

6. Conclusion

As the process technology continues to scale to smaller geometries, designing the power supply network becomes more challenging. In order to reliably deliver the power demanded by high-performance integrated circuits with nominal voltage variations, the power network must be properly analyzed with the appropriate amount of parasitic resistance, capacitance, and inductance included in the model. In addition, electromigration must be taken into account when sizing the power wires. With the faster switching devices and shorter clock cycles, the decoupling capacitor hierarchy extends closer to the switching logic gates. Avoiding resonant peaks near the clock frequency needs to be a constraint when designing the on-chip decoupling capacitors. We proposed formulating the decoupling capacitor verification problem as a linear programming problem to ensure that a sufficient amount of decoupling capacitors exists for all the on-chip switching devices. Part of the power network design process is to use architectural information or logic simulations to estimate the power requirement at the block or unit level during all stages of the design cycle. Finally, it is important to optimize the performance vs. power trade-offs using the appropriate cost/value metric. With each technology generation, the power supply network designer will face new challenges from all levels of the design hierarchy (i.e. devices to target applications).

7. Acknowledgements

The authors would like to thank R. Wheeler, R. Voelker, G. Yee, P. Trivedi, B. Amick, C. Gauthier, M. Blatt, J. Oh, J. Grinberg, H. Mau, and R. Heald for their contributions to this work.

8. References

- [1] R. Heald, et. al., "Implementation of a 3rd Generation SPARC V9 64b Microprocessor," *Proc. IEEE ISSCC*, pp. 412-413, 2000
- [2] P. Gronowski, W. Bowhill, R. Preston, M. Gowan, and R. Allmon, "High-Performance Microprocessor Design," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 5, pp. 676-686, Apr. 1998
- [3] J. Darnauer, D. Chengson, B. Schmidt, and E. Priest, "Electrical Evaluation of Flip-chip Package Alternatives for Next Generation Microprocessor," *Electronic Components and Technology Conference*, pp. 666-673, 1998
- [4] S. Borkar, "Low Power Design Challenges for the Decade," *Proc. of ISLPED*, 2000
- [5] V. Tiwari, D. Singh, S. Rajgopal, G. Mehta, R. Patel, and F. Baez, "Reducing Power in High-performance Microprocessors," *Proc. of Design Automations Conference*, 1997
- [6] G. Steele, et. al., "Full-Chip Verification Methods for DSM Power Distribution Systems," *Proc. of DAC*, pp. 744-749, 1998
- [7] R. Chaudhry, D. Blaauw, R. Panda, and T. Edwards, "Current Signature Compression For IR-Drop Analysis," *Proc. Design Automation Conference*, pp. 162-167, 2000
- [8] S. Bobba and I. N. Hajj, "Estimation of maximum current envelope for power bus analysis and design," *Proc. of ISPD*, pp. 141-146, Apr. 1998.
- [9] H. Kriplani, F. Najm, and I. Hajj, "Pattern Independent Maximum Current Estimation in Power and Ground Buses of CMOS VLSI Circuits: Algorithms, Signal Correlations, and Their Resolution," *IEEE Transaction on CAD*, vol. 14, no. 8, pp. 998-1012, Aug. 1995
- [10] D. Stark and M. Horowitz, "Techniques for calculating currents and voltages in VLSI power supply networks," *IEEE transactions on CAD*, vol. 9, no. 2, pp. 126-132, Feb. 1990
- [11] A. Dalal, L. Lev and S. Mitra, "Design of an efficient power distribution network for the UltraspARC-I microprocessor," in *Proc. of ICCD*, pp. 118-123, 1995.
- [12] G. Steele, D. Overhauser, S. Rochel and S. Z. Hussain, "Full-chip verification methods for DSM power distribution systems," *Proc. of DAC*, pp. 744-749, 1998.
- [13] A. Dharchoudhury, R. Panda, D. Blaauw and R. Vaidyanathan, "Design and analysis of power distribution networks in PowerPC microprocessor," *Proc. of DAC*, pp. 738-743, June 1998.
- [14] H. Chen and D. Ling, "Power supply noise analysis methodology for deep-submicron VLSI chip design," *Proc. of DAC*, pp. 638-643, 1997.
- [15] Y.-M. Jiang, K.-T. Cheng and A.-C. Deng, "Estimation of maximum power supply noise for deep sub-micron designs," *Proc. of ISLPED*, pp. 233-238, Aug. 1998.
- [16] S. Bobba and I. N. Hajj, "Maximum Voltage Variation in the Power Distribution Network of VLSI Circuits with RLC Models," *Proc. of ISLPED*, Aug. 2001.
- [17] J. Trattles, A. O'Neill, and B. Mecrow, "Three-Dimensional Finite-Element Investigation of Current Crowding and Peak Temperatures in VLSI Multilevel Interconnections," *IEEE Transactions on Electron Devices*, vol. 40, no. 7, pp 1344-1347, Jul. 1993
- [18] S. Rochel, G. Steele, J. Lloyd, S. Hussain, and D. Overhauser, "Full-Chip Reliability Analysis," *IEEE International Reliability Physics Symposium*, pp. 356-362, 1998
- [19] R. Wachnik, R. Filippi, T. Shaw, and P. Lin, "Practical Benefits of the electromigration Short-Length Effect, Including a New Design Rule Methodology and an Electromigration Resistant Power Grid with Enhanced Wireability," *Symposium on VLSI Technology Digest*, pp. 220-221, 2000
- [20] J. Kitchin, "Statistical Electromigration Budgeting for Reliable Design and Verification in a 300-MHz Microprocessor," *Symposium on VLSI Circuits Digest*, pp. 115-116, 1995
- [21] P. Larsson, "Power Supply Noise in Future IC's: A Crystal Ball Reading," *IEEE Custom Integrated Circuits conference*, pp. 467-474, 1999
- [22] H. Chen and S. Schuster, "On-chip Decoupling Capacitor Optimiza-

- tion for High-Performance VLSI Design," *International Symposium on VLSI Technology, Systems, and Applications*, pp. 99-103, 1995
- [23] H. Chen and J. Neely, "Interconnect and Circuit Modeling Techniques for Full-Chip Power Supply Noise Analysis," *IEEE Transactions on Components, Packaging, and Manufacturing Technology--Part B*, vol. 2, no. 3, Aug. 1998
- [24] P. Larsson, "Resonance and Damping in CMOS Circuits with On-Chip Decoupling Capacitance," *IEEE Transactions on Circuits and Systems--I*, vol. 45, no. 8, Aug. 1998
- [25] P. Larsson, "Parasitic Resistance in an MOS Transistor Used as On-Chip Decoupling Capacitance," *IEEE Journal of Solid-State Circuits*, vol. 32, no. 4, Apr. 1997
- [26] S.M.Kang, "Accurate Simulation of Power in VLSI circuits," *IEEE Journal of Solid State Circuits*, Oct 1986.
- [27] F. N. Najm, "A survey of Power Estimation Techniques in VLSI Circuits," *IEEE Transactions on VLSI*, Dec 1994.
- [28] H. Mehta, R. M. Owens, M. J. Irwin, "Energy Characterization Based on Clustering," *33rd Design Automation Conference*, June 1996.
- [29] E. Macii, M. Pedram and F. Somenzi, "High Level Power Modeling and Estimation," *IEEE Transactions on Computer Aided Design of Integrated Circuits and Systems*, vol 17, November 1998.
- [30] S. Gupta and F. N. Najm, "Power Macromodeling for High Level Power Estimation," *34th Design Automation Conference*, June 1997.
- [31] D. Brooks, V. Tiwari, and M. Martonosi, "Wattch: A Framework for Architectural-Level Power Analysis and Optimizations, " *Proc. of International Symposium on Computer Architecture*, pp. 83-94, June 2000
- [32] G. Cai and C.H. Lim, "Architectural Level Power/Performance Optimization and Dynamic Power Estimation," *Cool Chips Tutorial, MICRO32*, November 1999.
- [33] A. Dhodapkar et al, "TEM2P2EST: A Thermal Enabled Multi Model Power/Performance Estimator", *Workshop on Power-Aware Computer Systems, ASPLOS-IX*, November 2000.
- [34] V. Tiwari, S. Malik, and A. Wolfe, "Power Analysis of Embedded Software: A First Step Toward Software Power Minimization," *IEEE Trans. VLSI Syst.*, vol2, no. 4, pp.437-445, 1994
- [35] M. Gowan, L. Biro, and D. Jackson, "Power Considerations in the Design of the Alpha 21264 Microprocessor," *Proc. of Design Automation Conference*, 1998
- [36] S. Ghiasi and D. Grunwald, "A comparison of two architectural power models," *Workshop on Power-Aware Computer Systems, ASPLOS-IX*, November 2000.
- [37] D. Brooks et al, "Power-Performance Modeling and Tradeoff Analysis for a High End Microprocessor," *Workshop on Power-Aware Computer Systems, ASPLOS-IX*, November 2000.
- [38] R. Gonzalez and M. Horowitz, "Energy Dissipation In General Purpose Microprocessors," *IEEE Journal of Solid State Circuits*, vol 31, September 1996.