A COMPACT MODEL FOR PROJECTIONS OF FUTURE POWER SUPPLY DISTRIBUTION NETWORK REQUIREMENTS

James W. Joyner and James D. Meindl

Georgia Institute of Technology
Atlanta, GA 30332-0269
Phone: (404) 894-9910 Fax: (404) 894-0462 E-mail: joyner@ece.gatech.edu

ABSTRACT

A closed-form worst-case IR drop model is developed to enable projections of the wiring resource requirements of the power supply distribution networks of future microprocessors and ASIC's. The model is then used to highlight the design trade-off between I/O requirements and global wiring area needed for power distribution across three generations. For the 2013 technology generation, the global wiring resources exceed 40% of the available if fewer than 10,000 I/O's are utilized for power and ground.

INTRODUCTION

Power supply distribution network design is an area of growing concern for future gigascale integrated circuits. With each new technology generation, the active power is expected to increase by 2.7x while the leakage power is projected to grow at an even more alarming rate [1]. Coupling this with an expected 15% reduction in supply voltage per generation [1] leads to a commensurate increase in current density with each new technology generation. Such an increase places a large burden on the design of the distribution network.

Understanding the design trade-offs of the power supply distribution network for future ASIC's and microprocessors is of critical importance to improving both the performance and cost, and thus the cost effectiveness, of the products. With supply voltages constantly scaling downward with each new technology generation [2], local variations in IR drop produce performance variations in the realm of magnitude as those seen as a result of process variations [3]. These variations may become fractionally larger as the supply voltage is scaled down. Local IR drop creates a source-to-substrate bias that effectively increases the threshold voltage and creates larger timing delays [3]. In addition, power supply variations are a root cause of clock skew, which can further reduce performance [4]. More wiring resources, such as metal area per metal level, number of metal levels, and number of vias, can be used to negate these effects, but such a conservative design sees decreased process yield and increased cost per die [3]. The goal of an optimum power supply distribution network design is to minimize the wiring resources required to maintain system performance.

The total IR drop of a network must be modeled in a generic fashion to enable projections of network design requirements for future gigascale integrated circuits. Previous work in this field falls primarily in two categories: 1) closed-form modeling and 2) simulation. As solving for the voltage drop in even the most generic of distribution networks is computationally expensive, most closed-form IR drop models such as [3],[5],[6] are based only on the percentage area used for the network. On the other hand, simulation-based methods such as [7],[8] are targeted for use in a design methodology and do not result in an analytical expression which may highlight design trade-offs more readily. The model presented here makes use of the more complex simulation methods to highlight more dependencies yet results in a closed-form expression for the worst-case voltage drop needed to enable projections for future technology generations.

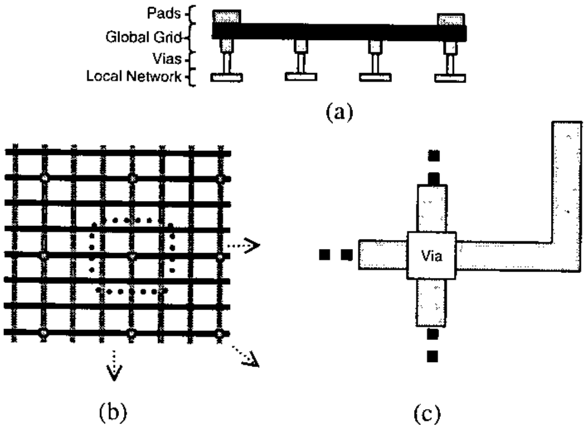
GENERIC POWER SUPPLY DISTRIBUTION NETWORK

The power supply distribution network is modeled after that presented in [3]. It consists of a single grid at the global level and feeder lines in a star network at the local level. The grid is located in the top two orthogonal metal levels, which have equal metal and dielectric thicknesses. Each intersection point of the global grid is connected to a star network on the lowest metal level (M1) through stacked vias that bore through the intermediate metal levels. A star network in M1 consists of a tree of metal lines emanating from each of the four sides of the via contact. An illustration of such a generic power supply distribution network is found in Figure 1.

The grid is characterized by two key design parameters: 1) the grid fineness f_{grid}, the number of parallel lines in a unit cell of the grid, and 2) the resistance of each grid segment R_{seg}. The number of parallel lines in a unit cell of the grid is assumed to be odd so that the power and ground networks can be optimally interlaced. The metal lines used for the star networks in M1 are assumed to have a width equal to that of the minimum feature size F. The area serviced by

0-7803-7494-0/02/$17.00 © 2002 IEEE

an individual star network is defined by the grid fineness of the global grid. Thus, the local power distribution network design is completely defined by the design at the global levels.



(a)

(b)

(c)

Figure 1. A generic one-grid power supply distribution network: (a) The cross-sectional view illustrates how the global grid matches the pad pitch to the via pitch. (b) The dotted rectangle in the top-down view of the global grid represents a grid cell centered about a single I/O pad. (c) The top-down view of a local star network shows the four escape paths, one of which is a worst-case longest path.

WORST-CASE IR DROP MODEL

In modeling the total IR drop of a power supply distribution network, the case in which a single global grid is utilized at the two uppermost metal levels as described in the previous section is considered. Assuming a grid fineness f_{grid} and a number n_{pg} of uniformly distributed power and ground I/O pads, a complete nodal analysis would include roughly ($0.5\ n_{pg}\ f_{grid}^2$) nodes, a number that can grow enormously for realistic future designs. Since only the worst-case IR drop is sought, the voltage at the center of the cell at the center of the chip is principally of interest. Based upon simulation of a single-grid network, the cells near the edge of the chip do not experience as severe an IR drop as those near the center of the chip as shown in Figure 2. Since the cells along the edge of the chip are bounded only on 2 or 3 sides rather than 4 as for the central cells, they experience less immediate loading and thus have less resistive voltage drop. Likewise, those cells adjacent to the outer cells experience less loading than the central cells but more than the outer cells.

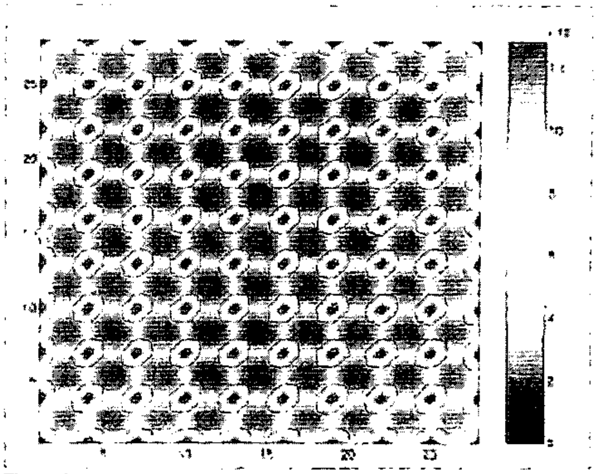


Figure 2. IR drop, given in millivolts, for a single-grid power supply distribution network with a grid fineness of three. The cells near the center of the chip see a larger load and experience the most severe worst-case IR drop.

Assuming that the number of cells is sufficiently large, the cells near the center of the chip see a load such that they are approximately equal to that if they were in an infinite plane. Assuming that the grid grows toward infinity in extent with a commensurate increase in the number of I/O pads utilized for power and ground, the IR drops at the nodes of each cell begin to mirror those of its neighboring cells. As the node voltages at the edge of a cell mirror those of its neighboring cells, the current conducted through the grid segments connecting the cells rapidly approaches zero, thus electrically isolating each individual cell. Assuming that a large number of I/O pads are dedicated to the power supply distribution network, therefore, leads to approximation of the entire system by a single cell. This reduces the number of nodes in a nodal analysis to simply (f_{grid}^2).

A single cell of the power supply distribution network had three-fold symmetry in the Cartesian plane referenced as in Figure 3: 1) about the x-axis, 2) about the y-axis, and 3) about the line $y=x$. Taking advantage of this symmetry, the possibility of only solving for the voltages in a single octant as shown in Figure 3 becomes available with no loss in accuracy. Such use of symmetry further reduces the number of nodes required for nodal analysis by roughly a factor of eight.

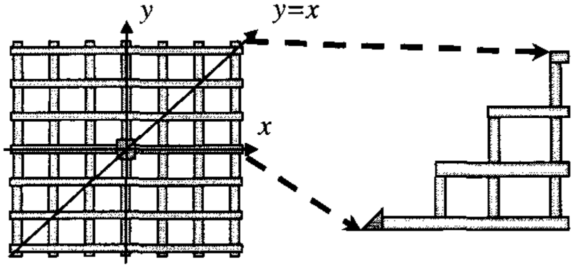


Figure 3. Symmetry about the x-axis, the y-axis, and the line $y=x$, shown on the left, allows a single octant, shown on the right, to completely describe the node voltages of a central grid cell.

Given the node naming scheme shown in Figure 4 and setting $V_{00}=V_{dd}$, the nodal system of equations can be expressed as

$$I_{i,j}=\begin{bmatrix}\dfrac{V_{i-1,j}-V_{i,j}}{R_{(i-1,j)(i,j)}}u_0\big(i-(j+1)\big)\\[2ex]+\dfrac{V_{i+1,j}-V_{i,j}}{R_{(i+1,j)(i,j)}}u_0\big(n-(i+1)\big)\\[2ex]+\dfrac{V_{i,j-1}-V_{i,j}}{R_{(i,j-1)(i,j)}}u_0(j)\\[2ex]+\dfrac{V_{i,j+1}-V_{i,j}}{R_{(i,j+1)(i,j)}}u_0\big(i-(j+1)\big)\end{bmatrix}\quad\text{for}\begin{Bmatrix}0\le j\le i\le n\\ i>0\end{Bmatrix}\quad(1),$$

where $I_{i,j}$ is the sink current from the node to ground, $R_{(i,j)(k,l)}$ is the resistance directly connecting the i,j node to the k,l node, and $u_0(x)$ is the unit step function of x. In Figure 4, I_s is the sink current seen at each node in a full cell. Similarly, R_{seg} is the resistance of a grid segment connecting two nodes.

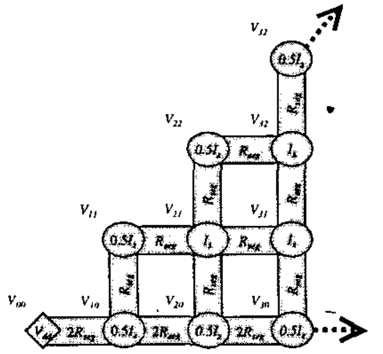


Figure 4. Representation of an octant of a power supply distribution network cell. The diamond node represents the supply pad. The rectangles represent grid segments with effective resistances equal to those indicated. Each oval node represents the intersection of two grid lines and contains the value of the effective sink current.

This system of equations in (1) can be readily solved using a matrix solver such as Matlab. Sweeping f_{grid} and normalizing the worst-case global IR drop $V_{IR,global}$ to the product of R_{seg} and the current supplied by each pad I_{cell} results in the plot seen in Figure 5. Using a curve-fitting algorithm, the IR drop is found as

$$V_{IR,global}=\frac{1}{2\pi}ln\left(1.917f_{grid}\right)I_{cell}R_{seg}\qquad(2)$$

with fitness parameter $r^2=1$. The maximum deviation of the expression to the numerical data is less than 0.2%.

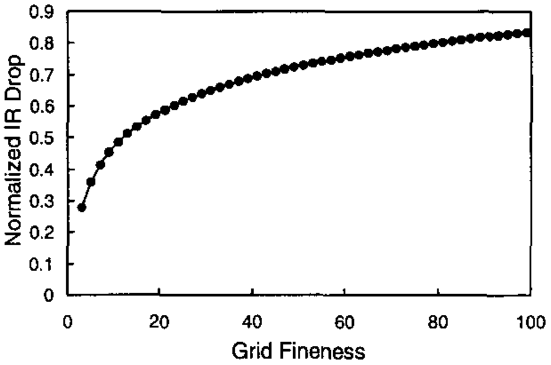


Figure 5. Numerical solution of global IR drop versus grid fineness normalized to the product of cell current and grid segment resistance. This curve can be accurately represented through a natural logarithmic function.

The cell current can be expressed in terms of the total chip current I_{total} as

$$I_{cell}=\frac{2I_{total}}{n_{pg}}\qquad(3).$$

The grid segment resistance can be expanded as

$$R_{seg}=\frac{\rho l_{seg}}{w_{pg}h}\qquad(4),$$

where ρ is the metal resistivity, w_{pg} is the width of the power and ground lines, h is the global metal thickness, and the grid segment length is expressed in terms of the chip area A_{chip} as

$$l_{seg}=\frac{1}{f_{grid}}\sqrt{\frac{A_{chip}}{n_{npg}}}\qquad(5).$$

Substituting and simplifying,

$$V_{IR,global}=\frac{1}{\pi f_{grid}}ln\left(1.917f_{grid}\right)\frac{I_{total}\rho}{w_{pg}h}\sqrt{\frac{A_{chip}}{n_{pg}^{\;3}}}\qquad(6).$$

The local power supply distribution network, defined here as that which lies in M1, consists of a star network emanating from each via boring down through the intermediate metal levels from the global grid. The worst-case voltage at the via contact is known from the worst-case global IR drop model described previously. From the via contact, the star network branches out in four directions. Assuming that each branch sinks equal current, only one branch must be considered. Although the devices sinking the current from the local feeder line may be distributed anywhere in the surrounding area, it is assumed that they all lie at the end of the line as a worst-case model. Thus, the current travels along a minimum feature size (F) feeder line for a distance of at most the via pitch, the distance between adjacent vias feeding the local star networks. Thus, the local worst-case IR drop is

$$V_{IR,local} = \frac{1}{4}\frac{I_{cell}}{f_{grid}^{\ 2}} R_{star} \qquad (7),$$

where the resistance of a line in the star network is

$$R_{star} = \frac{\rho l_{seg}}{F^2} \qquad (8).$$

Substituting and simplifying,

$$V_{IR,local} = \frac{1}{2}\frac{I_{total}\rho}{F^2}\frac{1}{f_{grid}^{\ 3}}\sqrt{\frac{A_{chip}}{n_{pg}^{\ 3}}} \qquad (9).$$

Summing, the total worst-case IR drop of a power supply distribution network is

$$V_{IR} = \frac{I_{total}\rho}{f_{grid}}\sqrt{\frac{A_{chip}}{n_{pg}^{\ 3}}}\left[\frac{1}{\pi w_{pg} h}\ln\left(1.917 f_{grid}\right) + \frac{0.5}{\left(F f_{grid}\right)^2}\right] \qquad (10).$$

A POWER SUPPLY DESIGN CONSTRAINT

As shown in (10), two power supply design parameters affect both the local and global IR drops significantly: 1) the grid fineness f_{grid} and 2) the number of power and ground I/O pads n_{pg}. As pointed out in [6], wire-bonding is not a scaleable packaging solution, and an area-array technology (e.g., flip-chip) is required to meet the needs of future high-power applications. Many new packaging technologies that supply high-density area-array I/O's at a low cost are being investigated [9].

To determine the number of I/O's needed to meet a target maximum IR drop $V_{IR,max}$, the model in (10) is constrained as

$$\frac{I_{total}\rho}{f_{grid}}\sqrt{\frac{A_{chip}}{n_{pg}^{\ 3}}}\left[\frac{1}{\pi w_{pg} h}\ln\left(1.917 f_{grid}\right) + \frac{0.5}{\left(F f_{grid}\right)^2}\right] \le V_{IR,max} \qquad (11).$$

Solving (11) for n_{pg},

$$n_{pg} \ge \left[\frac{I_{total}\rho}{f_{grid} V_{IR,max}}\sqrt{A_{chip}} \cdot \left(\frac{1}{\pi w_{pg} h}\ln\left(1.917 f_{grid}\right) + \frac{0.5}{\left(F f_{grid}\right)^2}\right)\right]^{\frac{2}{3}} \qquad (12).$$

Figure 6 illustrates this constraint on the number of I/O pads versus the grid fineness for the International Technology Roadmap for Semiconductors (ITRS) parameters shown in Figure 7. In addition, the percentage of the wiring resources of each global metal level required for power distribution is plotted in Figure 8. Figure 9 combines Figures 6 and 8 to illustrate the design trade-off between I/O requirements and on-chip metal resources. As demonstrated in the figure, a reduction in I/O requirements results in a steep increase of on-chip global wiring resources dedicated to power supply distribution. For the 2013 technology generation, the global wiring resources exceed 40% of the available if fewer than 10,000 I/O's are utilized for power and ground.

CONCLUSION

A closed-form worst-case IR drop model has been developed to enable projections of the wiring resources required for power supply distribution networks of future gigascale integrated circuits. Application of the model highlights the design trade-off between I/O requirements and global wiring area needed for power distribution across three generations. As the number of I/O's dedicated to power supply distribution decreases, the global wiring resources needed to maintain a target IR drop increases, highlighting the need for high-density I/O packaging solutions.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the support of the Defense Advance Research Project Agency (DARPA F33615-97-J-1132) and the Semiconductor Research Corporation Center for Advanced Interconnect Science and Technology (SRC-CAIST A70771). The authors also thank Payman Zarkesh-Ha of LSI Logic for many fruitful discussions.

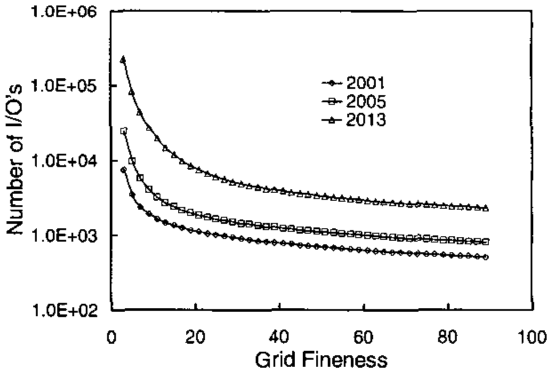


Figure 6. The number of power and ground I/O's required to meet a target overall IR drop versus grid fineness for three technology nodes.

Year	F (nm)	V_{dd} (V)	P_{total} (W)	A_{chip} (cm^2)	w_{pg} (μ)
2001	150	1.1	130	3.1	1
2005	80	0.9	170	3.1	1
2013	32	0.5	251	3.1	1

Figure 7. Technology parameters as outlined in [2].

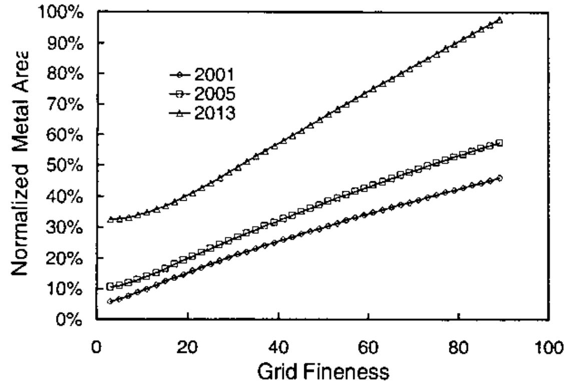


Figure 8. The percentage of global wiring resources needed to meet a target overall IR drop versus grid fineness for three technology nodes.

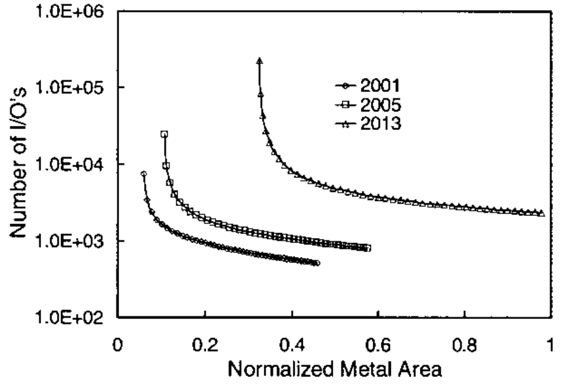


Figure 9. The number of power and ground I/O's versus the percentage of global wiring resources needed to meet a target overall IR drop for three technology nodes.

References

[1] S. Borkar, "Obeying Moore's Law beyond 0.18 micron," *IEEE Intl. ASIC/SOC Conf.*, 2000, pp. 26-31.

[2] Semiconductor Industry Association, "International Technology Roadmap for Semiconductors (ITRS)," 2001.

[3] L. A. Arledge and W. T. Lynch, "Scaling and Performance Implications for Power Supply and Other Signal Routing Constraints Imposed by I/O Pad Limitations," *IEEE Symp. IC/Package Design Integration*, 1998, pp. 45-50.

[4] S. Sauter, D. Schmitt-Landsiedel, R. Thewes, and W. Weber, "Effect of Parameter Variations at Chip and Wafer Level on Clock Skews," *IEEE Trans. Semiconductor Manufacturing*, vol. 13, no. 4, pp. 395-400, 2000.

[5] P. Zarkesh-Ha and J. D. Meindl, "Optimum On-chip Power Distribution Networks for Gigascale Integration (GSI)," *IEEE Intl. Interconnect Technology Conf.*, 2001, pp. 125 -127.

[6] D. Sylvester and K. Keutzer, "A Global Wiring Paradigm for Deep Submicron Design," *IEEE Trans. CAD-ICS*, vol.19, no.2, pp. 242-252, Feb. 2000.

[7] S. R. Nassif and J. N. Kozhaya, "Fast Power Grid Simulation," *Design Automation Conf. (DAC)*, 2000, pp. 1-6

[8] R. Panda, et al., "Model and Analysis for Combined Package and On-chip Power Grid Simulation," *Intl. Symp. Low Power Elec. Design (ISLPED)*, 2000, pp. 179-184.

[9] M. S. Bakir, H. A. Reed, P. A. Kohl, K. P. Martin, and J. D. Meindl, "Sea of Leads Ultra High-Density Compliant Wafer-Level Packaging Technology," to be published *IEEE Electronic Components and Technology Conf. (ECTC)*, San Diego, CA, May 28-31, 2002.