On-chip Power Supply Network Optimization using Multigrid-based Technique

Kai Wang and Malgorzata Marek-Sadowska
Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106
{wk,mms}@ece.ucsb.edu

Abstract

In this paper, we present a novel multigrid-based technique for on-chip power supply network optimization. We reduce a large-scale network to a much coarser one which can be efficiently optimized. The solution for the original network is then quickly computed using a back-mapping process. We model the power grid by an RLC network and use time-varying current sources to capture the on-chip switching. Our technique is capable of optimizing power grid and decoupling capacitance simultaneously. Experimental results show that the proposed technique provides more robust and area-efficient solutions than those obtained by the earlier approaches. It also provides a significant speed-up and brings up a possibility of incorporating power supply network optimization into other physical design stages such as signal routing.

Categories and Subject Descriptors

B.8.2[**Performance and Reliability**]: Performance Analysis and Design Aids

General Terms

Algorithms

Keywords

power supply noise, multigrid, congestion-aware

1. Introduction

The goal of the power supply network design is to deliver a time-varying current at a constant supply voltage with nominal variations. Technology scaling over the past few decades has enabled integrated circuits to speed up the computation rate and increase the number of computing elements at the cost of higher power dissipation. The trend of increasing power and clock frequency while reducing power supply voltage causes the power supply network to experience larger noise due to static IR-drop and dynamic Ldi/dt noise. In modern deep-submicron technologies, the power supply voltage variation greatly affects the delay of digital circuits. Additionally, device threshold voltage does not scale well with the reduced supply voltage. As a result, the power supply noise margin has become very tight, which makes the circuit performance more sensitive to the power supply noise.

On-chip power supply network consists of two main parts: power grid and decoupling capacitance (decap), both of which usually occupy a large portion of the chip area. To balance the system performance and reliability requirements, high-quality on-chip power supply network must be designed with exceptional care. The goal is to utilize the chip area efficiently, eliminate potential electromigration failures, and avoid excessive voltage drops. Power grid wire sizing ([2], [5], [7], [17], [18]) is an important method in power supply network optimization. Static IR-drop noise can be effectively reduced by careful adjustments of power grid wires widths. However, if switching current fluctuations are taken into account, it is difficult to design a robust power supply network purely by power grid wire sizing. Dynamic Ldi/dt noise becomes more pronounced as CMOS technology scales, and because inductance scales poorly with sizing [1], decap deployment ([15], [19]) becomes an indispensable technique for robust power supply network design.

In the previous literature, simultaneous power grid and decoupling capacitance optimization have only been applied to analog/mixed-signal ICs [13]. For large VLSI designs, power grid and decoupling capacitance optimization are treated separately. Most of the power grid wire sizing algorithms adopt a static resistive-only model for power supply network and do not consider the effect of decaps. They assume that the switching current drawn by each current source is constant, modeled by the average value of the enveloping current waveform. In other words, these power grid wire sizing algorithms do not take switching current fluctuations and Ldi/dt noise into account, thus in practice, their solutions are not area-efficient and robust. On the other hand, the previous decap optimization methods usually assume the power grid has been pre-designed and fixed. However, the performance and configuration of power grid and decoupling capacitance are closely related and greatly depend on each other. By simultaneously optimizing power grid and decoupling capacitance, a more area-efficient and robust solution can be found. However, in such a case, we have to deal with a huge number of variables because IR-drop and Ldi/dt noise have to be considered at the same time. The existing power supply network optimization methods are computationally intensive. Consequently they are unable to handle truly large-scale problems in a timely manner.

In this paper, we investigate the problem of on-chip power supply network optimization (we discuss only the V_{DD} network; the GND network can be treated similarly). To overcome the difficulties discussed above, we propose a novel multigrid-based technique to reduce the original large-scale network to a much coarser one. The reduced network can be optimized very efficiently due to its smaller dimensions. The solution of the original network is then quickly computed using a back-mapping process. The advantages of our technique are:

1) We use an accurate RLC power grid model and time-varying switching current profiles.

2) Power grid and decoupling capacitance are optimized simultaneously.

3) We achieve significant speed up of the optimization process without compromising the solution quality and we are capable of handling large-scale problems.

4) Our method of power supply network optimization can be combined with other design objectives such as signal routing.

The rest of the paper is organized as follows. In Section 2 we discuss the problem formulation. In Section 3 we review briefly the

DAC 2003, June 2-6, 2003, Anaheim, California, USA.

multigrid method for power supply network analysis. In Section 4 we give an overview of our technique. In Sections 5, 6, and 7 we describe the main three steps of our technique. In Section 8 we present the experimental results. Section 9 concludes the paper.

2. Problem Formulation

2.1 Power Supply Network Modeling

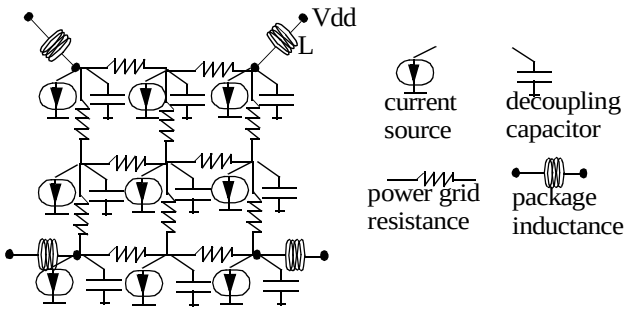

Fig. 1: Power supply network model

Power supply network consists of two main parts: power grid and decoupling capacitance, which are illustrated in Fig.1. Power supply network optimization is usually carried out after placement. This research targets optimization problems for global power supply networks. We make the following assumptions:

1) We consider only mesh topology of a power grid. Mesh is the most widely used structure in practical designs. Also, at the beginning of a systematic power supply network design process, it is advisable to have a regular structure such as mesh which can possibly be modified later into some irregular network

2) The package is predominantly inductive, and is modeled by serially connected inductances through power supply sources to their connection points in the power grid. The power grid itself is modeled as a resistive-only network because the inductance of the power grid is still much smaller than the package inductance, even for frequencies in a GHz range [8]. The capacitance of the grid is negligible compared to the capacitance of the active devices [12]. However, there is a decoupling capacitor between each power grid node and ground.

3) A time-varying current source is connected between each power grid node and ground. Those current sources represent switching of the circuit. The switching envelop current profile is modeled by a piece-wise-linear (PWL) waveform, which can be derived by an off-line logic block simulation and current-signature compression technique [3].

2.2 Cost Function

Our objective is to minimize the total area of power grid and decoupling capacitance subject to constraints. We define the cost function as a weighted sum of the power grid area and total cost of decoupling capacitance:

$$F = \alpha \sum_{i=1}^{p} l_i w_i + \beta \sum_{j=1}^{n} Cd_j \qquad \text{(EQ1)}$$

where α and β are two user-defined weights, p is the number of power grid wire segments, n is the number of power grid nodes, l_i and w_i are the length and width of wire segment i and Cd_j is the cost of decoupling capacitor at node j. Because the power grid topology is fixed, l's are fixed and variables are w's and Cd's.

2.3 Constraints

1) **Voltage drop constraints**. Instead of directly using $V_i(t) \geq V_{min}$ for each power grid node i, we use another noise metric proposed in [6] for voltage drop constraints:

$$\sum_{i=1}^{n} \int_0^T \left| min(0, V_i(t) - V_{min}) \right| dt = 0 \qquad \text{(EQ2)}$$

2) **Current density constraints**. Electro-migration in a wire segment sets an upper bound on the average current density. For a fixed thickness σ of a layer, this constraint for each wire segment i can be expressed as $\left| \frac{1}{T} \int_0^T I_i(t)\, dt \right| \leq w_i \sigma$, and can be expressed as nodal voltage constraints:

$$\left| \frac{1}{T} \int_0^T (V_i(t) - V_j(t))\, dt \right| \leq \rho\, l_i \sigma \qquad \text{(EQ3)}$$

where ρ is the sheet resistance and l_i is the length of the wire segment.

3) **Minimum width constraints**. The widths of the p/g wire segments are limited by the process technology to the minimum width allowed in the layer. We have:

$$w_i \geq W_{MIN} \qquad \text{(EQ4)}$$

4) **Decoupling capacitance constraints**. The size of every decoupling capacitor has its own lower and upper bounds:

$$C_{min,\,j} \leq Cd_j \leq C_{max,\,j} \qquad \text{(EQ5)}$$

These constraints can be derived from the layout information after placement. The lower bound depends on the estimated intrinsic decap value in the local region of the power node. The upper bound is determined by the cost of white space for extra decoupling capacitance allocation.

5) **Geometry regularity constraints**. All the power grid wire segments on the same horizontal or vertical line have the same width. The usefulness of this constraint will be discussed in section 5.2.

In summary, the problem of power supply network optimization is stated as follows: given a power supply network structure modeled as in section 2.1, decide the width of each power grid wire and the value of each decoupling capacitor such that the cost function F described by EQ1 is minimized subject to all the constraints represented by EQ2-EQ5.

3. Multigrid Method for Power Supply Network Analysis

General multigrid methods are important techniques for solving many large-scale problems. Due to the space limitations, we do not review them here. Please refer to [4], [9] and [14] for further details.

A multigrid-like technique was first applied for power supply network analysis in [11], then refined in [10]. The use of multigrid methods in power supply network analysis has been motivated by two considerations:

1) Well-designed power supply networks are characterized by voltage distributions which are spatially smooth [11].

2) The system of linear equations resulting from the analysis of power networks is structurally identical to that of a finite element discretization of a two-dimensional PDE.

The analysis technique proposed in [10] and [11] follows the steps of a general multigrid method. The basic idea is to coarsen the network until the problem becomes small enough to be solved exactly using a direct approach, and then map the solution back to the original fine network. Correspondingly there are three steps in that technique:

1) Grid reduction:

The objective is to remove as many nodes as possible while maintaining the ability to estimate voltage at the removed nodes.

2) Solving reduced grid:

The reduced grid can be solved exactly using a direct solver.

3) Interpolation:

The solution of the reduced grid is mapped back to the fine grid. The voltages at the removed nodes are estimated by interpolation based on the resistances to their neighboring nodes.

4. Overview of the Proposed Technique

Experimental results in [10] show that the multigrid-like analysis technique provides very accurate simulation results for DC as well as transient analysis of the power supply network, achieving significant speed-up over the traditional analysis techniques. From the geometrical interpolation of the algebraic multigrid reduction operation, an important property of the multigrid-like technique for the case of mesh structure can be derived and proved: during the multigrid reduction process, the area (thus the resistance) of the entire power mesh and the total value of decoupling capacitors remain constant. This constant-area-decap property constitutes the basis of our work. Let us consider a particular instance of a power supply network optimization problem. A feasible solution for the considered instance is a network with a given mesh topology and decided wire segment widths and decaps values, such that all the constraints are fulfilled. The set of all such solutions constitutes a feasible solution space for the initial network, which is denoted as FSP. If we apply the multigrid reduction technique to every solution in FSP, we obtain the set of all reduced solutions, denoted as RFSP and called the "reduced solution space". Every feasible solution in FSP can be mapped to a solution in RFSP. During the reduction process, the electrical characteristics are maintained by the multigrid technique, so all solutions in RFSP are also feasible. Due to the constant-area-decap property of the reduction process, the reduced solution in RFSP mapped from the optimal solution in FSP must also be the optimal solution in RFSP. In other words, if we can find the optimal solution in RFSP, then by using a back-mapping process, we can determine the optimal solution in the original solution space FSP. Because the network after reduction is much coarser than the original one, the process of searching for the optimal solution in the reduced space becomes much easier than searching in the original space.

Our optimization technique follows the basic procedures of the general multigrid method. We also have three steps, which will be discussed in detail in the following sections:

1) Power supply network reduction.

2) Solving the reduced network.

3) Back-mapping process.

5. Power Supply Network Reduction

In this section, we will first discuss the reduction process of a power grid with a mesh structure; then we will show how to handle decoupling capacitors, current sources, and voltage sources.

5.1 Power Mesh Reduction

As suggested in [11], a natural method for efficient single-level reduction of a uniform mesh, inspired by the standard multigrid method, is to skip every other row (or column) and to double the width of each remaining row (or column). This yields a significant reduction of the mesh size, almost a factor of four. Just as the general multigrid method can handle nonuniform grids [4], our work extends the "doubling" approach to handle general mesh structure in which the distances between columns (or rows) and the widths of columns (or rows) may differ.

For each level of reduction, given the topology of the original mesh, we can immediately decide the topology of the reduced mesh by skipping every other row (or column), and generating the relations between the widths of wire segments in the original and the reduced mesh. These relations are expressed using two reduction matrices, one for the row reduction and the other for the column reduction. The processes for constructing these two matrices are the same, so we will discuss only the construction of the column reduction matrix. Let us first consider a simple case of one level reduction. Suppose the number of columns in the original mesh is m, and the reduced mesh has $m/2$ columns. A is the column reduction $m/2 \times m$ matrix. The widths of the reduced mesh columns can be expressed as a linear combination of their original widths and the widths of their neighboring columns, now deleted, in the original mesh. Therefore, there is a total of $m/2$ linear equations. The reduction matrix A is the coefficient matrix of this set of linear equations. Each element in A is determined by the locations of the columns in the original mesh. Fig. 1 shows a simple case of removing one column:

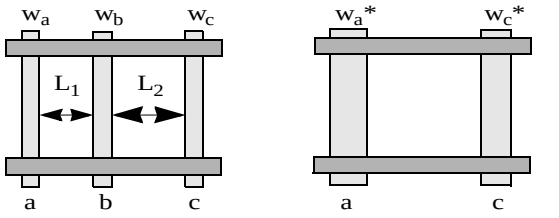

Fig. 2: A simple case

In the original mesh, the widths of columns a, b, and c are w_a, w_b, and w_c respectively. L_1 and L_2 are the distances from a to b and from b to c, respectively. After removing column b in the reduced mesh, the widths of column a and b become w_a* and w_c*. To keep the total area unchanged, the increments of the widths can be expressed as follows:

$$\Delta w_a = w^*_a - w_a = w_b \cdot \frac{L_2}{L_1 + L_2}$$

$$\Delta w_c = w^*_c - w_c = w_b \cdot \frac{L_1}{L_1 + L_2}$$

The example in Fig.2 shows the complete construction process of the column reduction matrix. We have the following set of linear equations:

$$\begin{bmatrix} 1 & \frac{l_2}{l_1 + l_2} & 0 & 0 & 0 \\ 0 & \frac{l_1}{l_1 + l_2} & 1 & \frac{l_4}{l_3 + l_4} & 0 \\ 0 & 0 & 0 & \frac{l_3}{l_3 + l_4} & 1 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \end{bmatrix} = \begin{bmatrix} w_1^* \\ w_2^* \\ w_3^* \end{bmatrix}$$

Please note that the above discussion is applicable only to single-level reduction. Now we consider the case of multi-level reduction. Supposing that k is the number of reduction levels, the original mesh is denoted L_0, and after the ith ($1 \le i \le k$) level reduction, the resulting mesh is L_i. The reduction matrix associated with the L_i mesh, ARM(i), is calculated as a product of all the previous single level reduction matrices:

$$ARM(i) = \prod_{j=0}^{i-1} A_{j \to j+1} \qquad (1 \le i \le k) \qquad \text{(EQ6)}$$

where $A_{j \to j+1}$ is a single level reduction matrix from L_j to L_{j+1}. Specifically, when $i = k$, the final reduction matrix A_F is given by:

$$A_F = ARM(k) = \prod_{j=0}^{k-1} A_{j \to j+1} \quad \text{(EQ7)}$$

We point out that the constant area property holds only for the mesh structure. For a general topology, even though the constant area might hold in some cases, it is not a necessary condition. Besides, the multigrid reduction operation on a system matrix cannot be interpolated geometrically for an irregular grid [8].

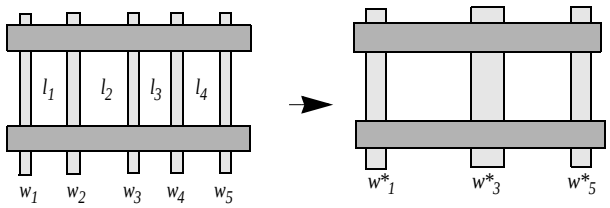

Fig. 3: A complete example

5.2 Splitting Decaps and Current Sources

We handle decaps and current sources in the same way as wires. A decap (or a current source) connecting to a removed node *i*, is split into decaps (current sources) connecting the neighboring nodes from which *i* will be interpolated. This splitting will be proportional to the resistances (thus the distances if we take into account the geometry regularity constraints in section 2.3) between the node *i* and its neighboring nodes. The total values of decaps and current sources are kept unchanged. Therefore, reducing from L_j mesh to L_{j+1} mesh, we have:

$$SP_{j \to j+1} Cd_j^T = Cd_{j+1}^T SP_{j \to j+1} I(t)_j^T = I(t)_{j+1}^T \quad \text{(EQ8)}$$

where $SP_{j->j+1}$ denotes the splitting matrix that can be derived from the topological structure of L_j and L_{j+1} meshes. Cd_j^T and $I(t)_{j+1}^T$ are the power node decaps and current source vectors, respectively. Moreover, it is easy to see that if we replace Cd_j^T with $Cd_{min,j}^T$ or $Cd_{max,j}^T$ (minimum or maximum decaps constraints vectors), the above equations still hold.

5.3 Handling Voltage Sources

The remaining problem we have to address is how to handle voltage sources during the reduction process. As in [11], we keep the wire segments that have voltage sources attached to them, so that the voltage sources always remain in the mesh during the reduction. Even in the flip-chip technology, the number of voltage sources in a typical power supply network is much smaller than the total number of nodes. Therefore this restriction does not seriously affect the efficiency of the reduction.

6. Optimizing Reduced Power Supply Network

6.1 Additional Linear Constraints

The linear relations between *W*, the widths of the wire segments in the original mesh, and *W**, the widths of the wire segments in the reduced mesh, can be expressed as:

$$\sum_{i=1}^{n} a_{ij} \cdot w_i = w^*_j \qquad j = 1 \ldots m \quad \text{(EQ9)}$$

where *n* is the number of columns (rows) in the original mesh and *m* is the number of columns (rows) in the reduced mesh. a_{ij} is an element in A_F. *W** is decided by the optimization engine, and *W* is computed by the following back-mapping process. We notice that according to the minimum width constraints in the problem formulation, the above linear equations must have a solution whose each element is larger than W_{MIN}. However, if the optimization engine is unaware of this requirement, it may produce a solution for *W** which does not fulfill the minimum width constraints. Therefore another set of constraints on *W** is added to guarantee a feasible solution of *W*.

We perform the following substitutions to (EQ9) and we have:

$$\sum_{i=1}^{n} a_{ij} \cdot x_i = b_j \qquad j = 1 \ldots m \quad \text{(EQ10)}$$

where $x_i = w_i - W_{MIN}$ and $b_j = w^*_j - \left(\sum_{i=1}^{n} a_{ij} \right) W_{MIN}$

We want to find a sufficient condition for the vector $B = [b_1,..b_m]^T$ such that *X* has at least one solution in which every x_i is larger than or equal to zero. For a general structure of *A*, finding such a condition for vector *B* is not trivial. Fortunately in our case, the reduction matrix A_F has two properties:1) A_F is very sparse; and 2) The number of variables present in the neighboring equations is very small. Based on these two properties, it can be proved that the necessary and sufficient condition for *B* such that *X* has at least one solution in which every x_i is larger than or equal to zero is:

$$w^*_j \geq \left(\sum_{i=1}^{n} a_{ij} \right) W_{MIN} \qquad j = 1 \ldots m \quad \text{(EQ11)}$$

Therefore, besides the constraints listed in section 2.3, the additional set of linear constraints (EQ11) must be taken into account during the optimization process to find the optimal solution for the reduced mesh. Note that for any intermediate L_i mesh, there is also a set of additional linear constraints on its wire segment widths, which can be similarly derived from the associated reduction matrix *ARM*(*i*) as we have done for the final reduced mesh.

6.2 Sequence of Quadratic Programing

The optimization problem for power supply network is a constrained non-linear optimization problem in time domain, which is computationally difficult, especially for large-scale problems. However, after multigrid reduction, the reduced problem can be solved efficiently. We use a sequence of quadratic programing package to solve it. Sensitivities of the noise constraints with respect to network parameters are calculated using transient adjoint network analysis technique. A similar method was used in [15] to optimize the decoupling capacitors size and placement in a standard-cell style layout. Due to the space limitations, we do not discuss the algorithm here. Please refer to [15] for more details.

7. Back-mapping Process

Once the optimal solution for the reduced network is obtained by the optimization engine, the total area of the power mesh and the total value of decaps are determined. The final step is to distribute the mesh area and decaps into the original power supply network by computing the solution for the original network using a back-mapping process. It is imperative for the resulting solution of the original network to fulfill all the reliability constraints. To achieve this objective and decrease the approximation error brought in by the multigrid reduction process, we adopt a corresponding multi-level back-mapping process. For each level, a set of linear equations is solved. From L_i network to L_{i-1} network, the set of linear equations can be expressed using the one-level mesh reduction matrix and decaps-splitting matrix. We note that these equations have more than one solution. This is so because multiple solutions

for the L_{i-1} network can be reduced to the same solution for the L_i network. This property gives us a flexibility in choosing the configuration of the L_i network. Therefore the choice of the solution for the original network can be affected by some other design issues. As mentioned in the introduction, one approach is to combine power supply network optimization with signal routing. Because, module placement has been finished by this time, we can utilize the resulting placement information to do congestion estimation. For the chip regions with higher congestion, the power supply network should occupy a smaller area so that the routing overhead can be effectively reduced. On the other hand, for the regions with lower congestion, the power supply network may have a relatively larger area. A similar idea is used in [16] We call this congestion-aware back-mapping. In our work, one level of back mapping problem is formulated as follows:

Congestion-aware back-mapping from L_i network to L_{i-1} network ($i = m,..., 1$)

Given the nodal voltage waveforms (and thus their time integrals), node decap values, and wire widths in L_i network,

$$\text{minimize} \quad \alpha \sum_{j}^{N} \gamma_j \cdot w_{i-1} + \beta \sum_{k}^{M} \gamma_k \cdot Cd^k{}_{i-1}$$

subject to the following constraints:

1) $A_{i-1 \to i} \cdot W_{i-1} = W_i$;

2) $SP_{i-1 \to i} Cd^T{}_{i-1} = Cd_i{}^T$

3) All the newly added vertical (horizontal) wire segments at the same location in the L_{i-1} mesh have the same width.

4) Voltage drop constraints (EQ2) are fulfilled for all nodes in the L_{i-1} mesh.

5) Current density constraints (EQ3) are fulfilled for the power wires in the L_{i-1} mesh.

6) $w^j{}_{i-1} \geq \alpha^j{}_{i-1} \cdot W_{MIN} \qquad (j = 1 \ldots N)$

7) $Cd^T{}_{min,i-1} \leq Cd^T{}_{i-1} \leq Cd^T{}_{max,i-1}$

In the above formulation, α and β are the same weights as in (EQ1); γ_j and γ_k are the weight coefficients assigned to the power wire j and decoupling capacitor k. They reflect the congestion at the location where the wire and decap are placed; larger γ implies higher congestion. With the progressing of the back-mapping process, the congestion resolution can be gradually increased.

From EQ(2) and EQ(3), we observe that all the reliability constraints can be expressed as linear functions of the time integrals of the nodal voltages. By enforcing the constraint 3) in the above formulation and using interpolation, we can express the time integrals of nodal voltages in L_{i-1} network using only a linear combination of the time integrals of nodal voltages in the L_i network and the topological information (wire lengths), which are known. Therefore all the constraints in 4) and 5) are linear. Although constraint 3) may limit the solution space, this effect is inconsequential based on our experiments. The constraints 6) come from the consideration that L_{i-1} mesh must be back-mapped further to obtain the original mesh. All the α values in 6) can be computed from *ARM*(i-1), the reduction matrix of the L_{i-1} mesh discussed in section 6. Now we can conclude that the above one-level back-mapping problem is a linear programming problem that can be solved very efficiently. After solving this problem, the widths of the wire segments and the decaps values in L_{i-1} network are decided. The information about L_{i-1} network can be further used to solve the next-level back-mapping problem.

In summary, the back-mapping process starts from the optimized final reduced power supply network, and by solving a sequence of one-level back-mapping problems from level to level, the optimal solution for the original power supply network can be computed without violating the reliability constraints.

8. Experimental Results

Based on the proposed multigrid-based technique, we have developed our prototype tool in C++ programming language. The benchmark circuits are implemented in a 0.18μm technology. All experiments were carried out on a PC running Linux operating system with P4 2.4GHz processor and 1GB memory.

To demonstrate the efficiency of our technique, we also implemented a direct two-step power supply network optimization approach to compare against our technique. In this two-step optimization approach, we perform optimization directly on the original network without multigrid reduction. We first adopt the sequence of linear programs (SLP) [17] method to decide the power-grid wire widths using the previous power grid sizing formulation [5], then decide the values of decoupling capacitors using the optimization method discussed in section 6.2.

The results are summarized in Table 1. Columns 1 to 3 list the circuit name, number of nodes and number of wire segments. Column 4 shows the percentage of the nodes and wire segments with reliability violations with respect to the total number of nodes and wire segments in the network resulting from the direct method. Column 5 shows the CPU time in seconds for the direct method. The power grid area, decoupling capacitors, CPU time, and speed up for our multigrid reduction technique are reported in column 6, 7, 8 and 9 respectively. For every tested mesh, the power grid area and decap values obtained from direct method are normalized to 1. There are no reliability violations in the resulting network determined by our technique, which is verified by SPICE simulation.

We make several observations based on the results shown in Table 1:

1) Our technique is several orders of magnitude faster than the direct method. The speed-up is even more dramatic for larger circuits.

2) Our technique can produce robust and more area-efficient solutions in terms of power grid and decoupling capacitance areas. By using the multigrid-based technique, the reduced mesh can be optimized considering temporal correlations.

3) Our technique can handle truly large circuits. A network with 1000x1000 nodes can be solved in about 5 minutes. In contrast, the direct method cannot produce any results for the tested network even after running 16 hours.

9. Conclusions

In this paper, we have presented a novel multigrid-based technique for the power supply network optimization, subject to reliability constraints. Using an accurate RLC power supply network and time-varying switching-current models, our technique is capable of simultaneously optimizing power grid and decoupling capacitance. Experimental results show that the proposed technique not only provides a more robust and area-efficient solution with significant speedup, but also opens up the possibility of incorporating the power supply network optimization into other physical design stages such as signal routing.

Acknowledgement

This work was supported in part by the NSF grant CCR 0098069 and in part by the California MICRO program and Cadence Design Systems. The authors also would like to thank Sani R. Nassif and Haihua Su for providing helpful suggestions.

References

[1] H. B. Bakoglu, *Circuits, Interconnects and Packaging for VLSI*. MA: Addison-Wesley, 1990.

[2] S. Boyd, L. Vandenberghe, A. El Gamal, and S. Yun, "Design of Robust Global Power and Ground Networks", *Proc. of ISPD'01*, pp. 60-65, 2001.

TABLE 1. Comparison of our multigrid-based technique against the direct method

Circuit	# of nodes	# of edges	direct method		multigrid-based technique			Speedup factor
			reliability violations	CPU time(s)	power grid area	decoupling capacitance	CPU time(s)	
circuit1	100	180	10.17%	19.60	0.91	0.82	5.75	3.4
circuit2	400	760	5.66%	76.05	0.89	0.93	8.85	8.6
circuit3	1600	3120	7.82%	549.6	0.73	0.65	27.8	19.7
circuit4	10000	19800	17.93%	2335.65	1.01	0.75	56.75	41.3
circuit5	1000000	1998000	N/A	N/A	N/A	N/A	308.35	>194.5

[3] R. Chaudhry, D. Blaauw, R. Panda, T. Edwards, "Current signature compression for IR-drop analysis", *Proc. of Design Automation Conference*, June 2000.

[4] W. L. Briggs. "A Multigrid Tutorial", *SIAM*, 1987.

[5] S. Chowdhury, "Optimum Design of Reliable IC Power Networks Having General Graph Topologies,"*Proc. of Design Automation Conf.*, pp. 787-790, 1989.

[6] A. R. Conn, R. A. Haring, and C. Visweswariah. "Noise Consideration in Circuit Optimization", *Proc. ICCAD*, pp. 220-227, Nov. 1998.

[7] R. Dutta and M. Marek-Sadowska, "Automatic sizing of power/ground (P/G) networks in VLSI," *Proc. of Design Automation Conf.*, pp, 783-786, 1989.

[8] K. Gala, V. Zolotov, R. Panda, B. Young, J. Wang, and D. Blaauw, "On chip inductance modeling and analysis", *Proc. Design Automation Conference*, June 2000.

[9] W. Hackbusch, "Multi-Grid Methods and Applications", *Springer-Verlag* Berlin, 1985.

[10] J. N. Kozhaya, S. R. Nassif, and F. N. Najm, "Multigrid-like Technique for Power Grid Analysis", *Proc. of ICCAD*, pp. 480-487, Nov. 2001.

[11] S. R. Nassif and J. N. Kozhaya, "Fast power grid simulation", *Proc of Design Automation Conf.*, pp.156-161, 2000.

[12] S. R. Nassif and O. Fakhouri, "Technology Trends in Power-Grid-Induced Noise", *Proc. of SLIP'02*, pp. 55-59, Apr. 2002.

[13] B. Stanisic, R. Rutenbar, and L. Carley, "Addressing Noise Decoupling in Mixed-Signal IC's: Power distribution Design and Cell Customization", *IEEE J. Solid State Circuits*, vol. 30, no. 3, pp. 321-326, Mar. 1995.

[14] K. Stiiben, and U. Trottenberg. "Multigrid Methods: Fundamental Algorithms, Model Problem Analysis and Applications", *Multigrid Methods Proc.*, 1982.

[15] H. Su, S. S. Sapatnekar, and S. R. Nassif, "An Algorithm for Optimal Decoupling Capacitor Sizing and Placement for Standard Cell Layouts", *Proc. of ISPD'02*, pp. 68-73, Apr. 2002.

[16] H. Su, J. Hu, S. S. Sapatnekar, and S. R. Nassif. "Congestion-driven Codesign of Power and Signal Networks", *Proc, Design Automation Conference*, June 2002.

[17] X.-D. S. Tan, C-J. R. Shi, D. Lungeanu, J.-C. Lee, and L.-P. Yuan, "Reliability-Constrained Area Optimization of VLSI Power/Ground Networks via Sequence of Linear Programming", *Proc. of Design Automation Conf.*, pp. 78-83, 1999.

[18] X. Wu, X. Hong, Y. Ca, C. K. Chen, J. Gu, and W. Dai, "Area minimization of power distribution network using efficient nonlinear programming techniques", *Proc. of ICCAD*., pp. 153-157, 2001.

[19] S. Zhao, K. Roy, and C.-K. Koh, "Decoupling Capacitance Allocation and Its Application to Power-Supply Noise-Aware Floorplanning", *Computer-Aided Design of Integrated Circuits, IEEE Trans*. on, pp. 81-92, Jan. 2002.