Simultaneous Switching Noise Consideration for Power/Ground Network Optimization

Le Kang[1], Yici Cai[1], Jin Shi[1], Xianlong Hong[1], Sheldon X.-D. Tan[2], and Xiaoyi Wang[1]
[1]*Department of Computer Science and Technology, Tsinghua University, Beijing, P.R. China*
[2]*Department of Electrical Engineering, University of California at Riverside, USA*
kl@mails.tsinghua.edu.cn

Abstract

With the rapid development of semiconductor technology, the working frequency of chips increases dramatically. Thus simultaneous switching noise (SSN) must be considered for robust power/ground (P/G) network design. In this paper, we mainly focus on the SSN effects for P/G network optimization. We first point out the drawbacks of the P/G optimization process without considering the SSN, by analyzing the optimized P/G grids. Then we propose a random walk based technique to consider SSN by adding decoupling capacitor (decap) prior to the nonlinear optimization process. This additional decap allocation phase constructs good current return path for the switching current caused by clock buffers and then reduces the dynamic voltage drop. Experiment results show that the proposed method achieves 2X speed up over the original approach without adding decaps in advance while the decap budget overhead is acceptable.

1. Introduction

Signal integrity in the power/ground (P/G) network is becoming a serious issue in nano-era chip design. The robust P/G delivery network design is considered as a major challenge as the manufacture technique advances [1]. Improper power distribution system design can reduce the chip reliability, or even worse, cause functional failures. As the chip complexity increases, clock switching frequency becomes higher and causes larger on-chip dynamic current within a shorter period of time. If this current propagates along with the power rail, the supply voltage will decrease significantly, especially when many macro-blocks or output buffers switch at the same time [2] [3]. The simultaneous switching noise (SSN) can cause logic errors and adversely affect the performances and reliabilities of VLSI circuits.

Power delivery system may suffer power drops due to resistive and inductive impedance. In general, due to the higher power consumption of the VLSI chips, wire sizing technique is typically used to reduce the static *IR* drops [5] [6]. However dynamic voltage fluctuations may still occur on the P/G network even if the wire sizing strategy is performed. In this case, adding decoupling capacitors (decaps) is regarded as the most efficient way to reduce dynamic noises [7] [8] [9]. Traditionally, spare area generated in the placement phase can be reserved for adding decaps for noise reduction. However, with the increase of clock frequency and the number of clock buffers, if lots of gates switch at the same time, the overall current surge can generate large dynamic drops [2], and the SSN in VLSI circuits will become more and more significant [3]. Therefore, the SSN can no longer be ignored during the P/G network optimization.

In this work, we first point out the drawbacks of the optimization process without considering the SSN. Second, we focus on how to optimize the SSN in the P/G network efficiently and propose a predecoupling approach based on random walks to reduce the noises caused by drastic current generated by clock buffers. In our algorithm, we use preconditioned conjugate gradient (PCG) method to get the transient response of power grid, and then we perform a random walk process starting from the nodes in power grid, which connect to the clock buffers and would have large impact on the performance of the power network when all of them switch at the same time. During this process, decaps are pre-added to the P/G network according to the current value of the clock buffers which is usually larger than that of the normal functional macro-blocks. After that, we use the nonlinear programming method to optimize the pre-optimized power grids. The experimental results demonstrate that the predecoupling method achieves 2X speed up over the approach which tries to directly

978 -1-4244-1579-3/07/$25.00 © 2007 IEEE

optimize the networks, while just use a little more budget area as penalty.

This paper is organized as follows. Section 2 formulates the simultaneous switching current and analyzes the effects of SSN; section 3 describes predecoupling optimization approach; section 4 gives out experimental results and finally, section 5 concludes the paper.

2. Analysis of the optimization results considering SSN

In this section, we first introduce the modeling of P/G delivery network and nonlinear programming process for P/G network optimization. Then we mainly analyze the non-SSN-aware optimized power delivery network with the simultaneous switching activities, and give out the drawbacks of the optimization process which does not consider the SSN separately.

2.1. Power delivery network modeling

The power network is typically designed as a hierarchical structure in which global all the top-level macro-blocks and clock buffers connect to the global power grid while logic gates inside the top-level macro-blocks connect to the inner power grid which is located inside the global power grid. On the other side, power grid analysis techniques needs current waveform of each macro cell or buffer as input, thus, the normal analysis process is always divided into two parts. In the first part, the current waveform should be analyzed. Then all the current information is used to estimate the voltage variation waveforms in the power grid. And the complete power grid model can be considered as a mesh-like RLC network illustrated in Figure 1.

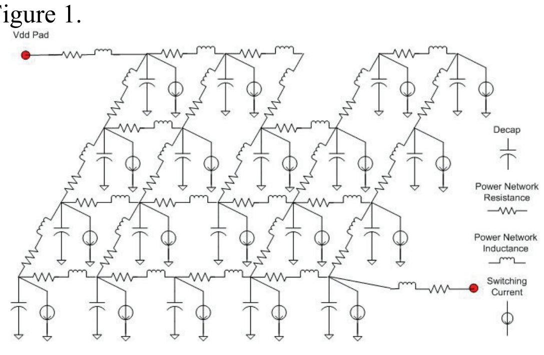

Figure. 1. A general modeling of P/G network.

2.2. Nonlinear programming optimization

The optimization techniques presented in previous work [7] [8] [9] add on-chip decaps to reduce transient voltage drops using the dynamic power consumption excited by the time-varying input vectors based on empirical values of the switching current.

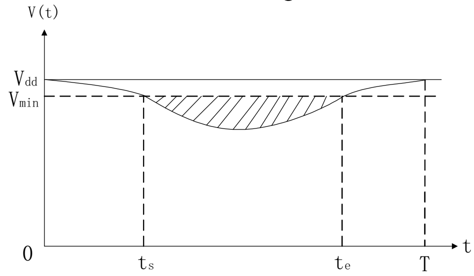

Figure. 2. Illustration of IR drop violation.

The normal category is to obtain the optimal budget of decaps which provide reservoir capacitors between power and ground as shown in Figure 1. In previous work [7], the violation area at node j is expressed as equation (1) according to the shaded area within one clock cycle illustrated in Figure 2:

$$g_j(C_1, C_2, ..., C_n) = \int_0^T \max(V_{\min} - V_j(t), 0)dt \qquad (1)$$

Recent work [8] formulates the total decap area using the bellowing objective function:

$$\min A = \left(\sum_{i \in alldecapnodes} C_i\right) + \lambda\left(\sum_{j \in allviolaitonnodes} g_j^2\right) \qquad (2)$$

Then objective function is minimized by traditional conjugate gradient (CG) method in each optimization iteration. The weighting factor λ in equation (2) is used to balance the two terms, and will keep changing in each CG iteration. In order to get the gradient direction of the objective function to each decap value, sensitivity of the objective function respect to each decap value is computed by the merged adjoint network method and it can be calculated in only two transient simulations within each iteration [8]. In each step, a binary search is still needed for CG method to update decaps along the current gradient. The CG process stops when the constraints are satisfied.

2.3. Drawbacks of the non-SSN-aware optimization results after considering SSN

As described above, the optimization techniques in previous work [7] [8] [9] add on-chip decaps to reduce transient voltage drops given the dynamic power consumption excited by the time-varying input vectors based on empirical values of the switching current. In almost previous works for P/G optimization, the SSN is not noticeable. When many circuit switch at the

Table 1. Violation node (VN) statistics about non-SSN-ware optimized P/G network when a number of circuits switch at the same time.

Power Gird	Cell/Circuit Num	Power Node Num	Eliminated VN Num	Added Decaps(nf)	Circuit Switch Simultaneously	Newly VN Num
1	15	26	0	0	3	4
2	744	806	139	0.101960	148	157
3	3492	4032	952	0.443672	982	681

same time, the current supplied by the power networks can change rapidly, and the voltage drop along the power line can cause the power supply level to go down.

Here we select three P/G grids, and the transient current signature of each macro-block or buffer, which mimics the waveforms of the actual circuits, is acquired by the equivalent circuit inside. So the circuit can be replaced by the piecewise linear current, such as the triangular or trapezoidal current waveform [4].

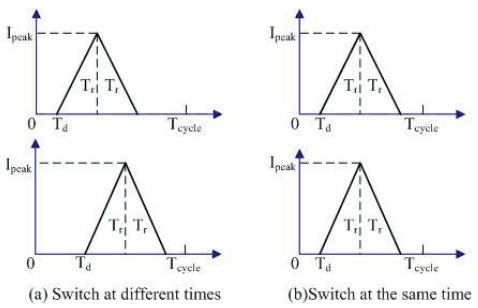

Figure. 3. Illustration of two circuits switching.

Firstly, the power girds are analyzed and optimized based on empirical values of the switching current, i.e. in the optimization process, the input vectors of switching current for different circuits usually have different phases, for instance, Figure 3(a) gives an illustration of two circuits which switch at different times. In this case, the potential voltage drop along the power line is not particularly evident and the decap allocation acquired from optimization process [8] may not be perfect.

Then we consider some circuits including the clock buffers and some macro-blocks turn on simultaneously. As shown in Figure 3(b), the equivalent transient currents of two circuits, which connected to the P/G network switch at the same time, have the same delay time T_d. So we use the transient analysis tool to verify the actual voltage drop in the P/G network that is optimized by the non-SSN-aware optimization process.

Table 1 summarizes the analysis results. From the results, we can find that the violation node (VN) number is reduced to zero after adding decaps in the non-SSN-aware power network. However, some new violation nodes appear when SSN is considered, which may cause logic errors and affect the performances of the chips.

So one observation we can have from the data is that it is necessary to consider SSN during the P/G optimization process to maintain a robust optimization.

3. Predecoupling technique for P/G network optimization

One simple way to optimize the P/G network is to introduce the SSN directly to the optimization method. In this case, optimization algorithms only try to insert decaps when power node voltage falls below the required boundary, but the features of the SSN is ignored. In this section, we propose the predecoupling technique by adding decaps prior to the nonlinear optimization process. The technique for P/G network constructs good current return path for the drastic switching current caused by clock buffers and reduces the dynamic voltage drop, which will improve the run time the optimization process.

3.1. Predecoupling for the SSN of buffers

Our main idea is based on the fact that the switching activity of high speed CMOS circuit may produce large currents or current derivatives [2]. This is especially true for the clock buffer which is used for reducing interconnect delay in the clock network. So the SSN due to output buffers switching current is larger than the SSN due to the internal switching in the macro-blocks. In other words, the voltage drops due to SSN in the P/G grid are mainly caused by the simultaneous switching activities of the buffers that connect to the P/G network.

Usually, due to the frequency of noises in P/G network is high, enough decaps will provide good fast return paths for noise current. On the other hand, under a special noise frequency, we can calculate how many decaps are enough for the transient current of the buffers to provide isolation in P/G network even in time domain. So in order to compensate the noise caused by the buffers, we can try to plant a certain decap direct to or around each buffer to reduce the path length. And we can calculate the capacitor size C_i as follows:

$$C_i \times U_i = Q_i = \int I_i(t)dt \quad (3)$$

In the equation above, $I_i(t)$ is the switching current of buffer *i*, and U_i is the voltage.

The planted decap will provide the return path for each buffer in isolation, and each of them can be taken into consideration independently. And we call this predecoupling technique for the buffers.

After adding decaps in advance, we still introduce the nonlinear optimization process to the P/G network, which continues to reduce other noises until the power grid is robust. And an efficient and novel approach is given out in next part to describe how to plant the definite decap for a buffer.

3.2. Random walk based decap planting strategy

In this part, decaps are planted during a modified random walk process. Random walk [10], a classical stochastic technique, is widely used to solve many engineering problems. For example, random walk is adopted for circuit simulation such as the power grid verification [11], the analysis of the circuit with the resistor and power supply can be transformed into a random walk on the circuit network via using walking probability determined by the electrical parameters on one path. No solving of matrix equation is needed, but the performance of random walk is poor when the grid has few Vdd pads or the grid is too large, and it is also difficult to use random walks to simulate a whole P/G gird or solve the transient analysis problem [12].

Our decap planting approach is mainly refer to the previous work [11], and take the advantage of random walks.

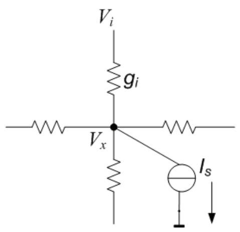

Figure. 4. A representative part of P/G networks.

First, we treat the power nodes in the network that directly connects the buffers as the starting nodes in the random walk process.

Then the modified walking process is performed according to the probability mainly decided by the resistors and power supplies. For a part of P/G network shown in Figure 4 following formula can be written according to Kirchoff current law and nodal equation.

$$V_x \sum g_i - \sum g_i V_i = -I_s \quad (4)$$

Equation (4) can be reformulated as follows:

$$V_x = \sum \frac{g_i}{\sum g_i} V_i - \frac{I_s}{\sum g_i} \quad (5)$$

The coefficient $g_i / \sum g_i$ can be treated as the probability of walking from node *x* to node *i*. The constant $-I_s / \sum g_i$, which is treated as the fee at node *x* and accumulated at each walking step to calculate the voltage of the starting node in the random walk simulation, can be ignored in our process. But differently, the walking process doesn't need to terminate at pad nodes [11], instead, the termination condition of the walks is controlled by the parameter which affects the scope in the network for adding decaps to provide the fast return path for the buffer.

Finally, we can carry out the modified walking process several times from the node that connects the buffer, and the accessed times for the visited nodes are gained during the walks. For a specific buffer *i*, the decap C_i is divided, and planted at several nodes according to the visited times of them in the walks. The capacitance of a visited node *j* of buffer *i* can be formulated as following:

$$C_j^{(i)} = \frac{Count_j^{(i)}}{\sum_j Count_j^{(i)}} C_i \quad j \in \text{all visited nodes of buffer } i \quad (6)$$

where C_i, $C_j^{(i)}$ are total decap size for the buffer and decap size for the visited node *j*. $Count_j^{(i)}$ is the visited times of node *j* during the walking processes starting from the node connecting to the buffer *i*. Thus, $Count_j^{(i)} / \sum_j Count_j^{(i)}$ can be treated as the probability that node *j* is hinted in the modified walking process.

In general, for each buffer, decaps are planted according to the probability in random walks independently, where the decap size for a buffer is

Table 2. Analysis results of P/G network after adding decaps based on random walks.

Power Gird	#Circuit	Power Node	VN Num	Pre-added Decaps(nf)	VN Num after Adding Decaps
1	15	26	4	0.576113	0
2	744	806	277	0.093449	175
3	3492	4032	1411	0.181388	1236
4	32112	32952	9988	1.802553	3791

estimated according to that given in equation (3). Table 2 analyzes the P/G network before and after the predecoupling considering the SSN. The column 4 5 6 represent the violation node number in the original network, the total predecoupling capacitor size and the violation node number after the predecoupling. From the results we can observe that the violation node number decreases significantly after random walk based predecoupling strategy on the buffer when considering SSN.

According to the characteristics of the SSN in P/G network, the definite decap for each buffer to compensate the drastic changes in the currents. So after the decap planting for each buffer, nonlinear programming optimization process is still performed to optimize the pretreated power gird if necessary. Compared to the direct nonlinear programming approach for P/G network, our proposed method has an advantage in running time occurring only a little more decaps, and we will also discuss why random walk is used for decap planting during the predecoupling strategy via the experimental results in next section.

4. Experimental results

Our predecoupling optimization algorithm is implemented using C++ language. All test cases are generated randomly according to real industry standard-cell circuits, and buffers are inserted during the macro-blocks.

4.1. Two factors affecting the predecoupling

The efficiency of our proposed algorithm depends on two factors in the predecoupling stage. First one is the amount of the decap we add for each drastic current, the other one is the terminal parameter k of random walks introduced above. Figure 5 shows the optimization results of different amounts of pre-added decaps, where the decap size from the equation (3) is considered as one on the x-axis and the markers are relative errors for different pre-added decaps. The results shown in Figure 5 demonstrate that the pre-decap calculated in the equation is sufficient, and if the amount increases, the over-head of the optimization results may be too large, or if the pre-added decap size is very small, it will take more time to carry out the nonlinear programming but the decrease of the errors isn't obvious.

To demonstrate the efficiency of the terminal parameter in random walk based predecoupling stage, none predecoupling strategy, directly adding predecoupling capacitors on the equivalent drastic currents of buffers, and random walk based predecoupling technique are all performed to optimize a same P/G delivery network considering the SSN.

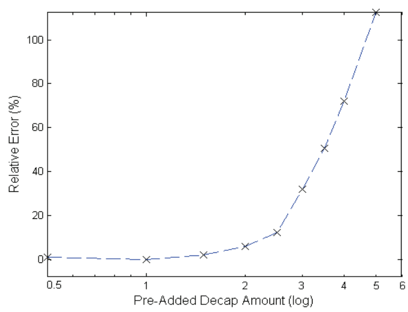

Figure 5. Relative errors of different pre-added decap amounts.

Table 3 summarizes the comparison of the decap amount of the optimization results, where the decap amount acquired without adding predecoupling capacitors can be considered as the exact result. From the results, we can observe that the relative error of random walk based predecoupling optimization increases when the terminal parameter increases, but the optimization results getting from random walk based method may be better than that acquired from directly adding decaps at the drastic currents, when the parameter k isn't too large. Because the random walk based approach takes the advantage of the Kirchoff law and it can find the more sensitive nodes to plant decaps, but the sensitivity may be poor if the scope of the random walks is too large.

4.2. Acceleration using predecoupling strategy

We apply our random walk based predecoupling optimization approach for the test cases with different sizes. As shown in Table 4, comparisons are made between the non pre-added optimization method and our proposed approach considering the SSN. Columns

Table 3. The optimization results comparison with different parameters on the same testcase.

	None Predecoupling	Direct Predecoupling	Random Walk based Predecoupling with Different k				
			2	3	5	8	10
Added Decap (nf)	0.229297	0.231994	0.229394	0.233949	0.235260	0.237079	0.240833
Relative Error	--	1.18%	0.04%	2.03%	2.60%	3.39%	5.03%

1, 2, 3, 4 represent the power grid, the node of the test case, total number of macro-blocks and the number of clock buffers connecting to the power network respectively. The budget of the non pre-added method can be considered as the optimal solution for P/G network optimization. The results show that our proposed approach only has approximate 2% decap area overhead in average, and it may get better budget with the same tolerance in the CG iteration.

On the other hand, the time complexity of the CG optimization process [8] [9] for P/G network which has been introduced in section 2 can be expressed as $O(n^{1.5}ml(2+r^i))$, where n is the power node count, $n^{1.5}$ is the typical time complexity for solving sparse matrices [9], m is the number of time points in time-domain, and l is optimization iteration number. Here *2* denotes two times transient simulations for calculating the sensitivity and r^i is the complexity of binary search at iteration i. The random walk based predecoupling approach achieves a 2x speed up over none pre-added approach because our proposed method reduces the number of the CG iterations after pre-adding decoupling capacitors. Generally, the steps of binary search in each iteration may also be reduced after the predecoupling strategy. Thus the faster convergence in conjugate gradient leads to less computation cost and the run time reduction.

5. Conclusion

In this paper, in order to clearly show the influence of the SSN, we first analyze the P/G grids optimized without considering SSN to show the impacts when many circuits turn on simultaneously. Then, we propose a random walk based predecoupling approach for P/G optimization. The approach relies on the fact that the equivalent switching current of clock buffer is more drastic than that of normal macro-block. Thus we add decaps prior to the optimization process for reducing the dynamic voltage drops caused by clock buffers. According to the experimental results, our proposed approach has a faster convergence in conjugate gradient compared with the method that carries out the optimization process without adding decaps in advance. And the decap area overhead in our proposed approach is acceptable.

Acknowledgement

This work is supported by the National Natural Science Foundation of China (NSFC) 90307017.

References

[1] The International Technology Roadmap for Semiconductors (ITRS), http://public.itrs.net/, 2005 update.

[2] G. Bai and I. N. Hajj, "Simultaneous Switching Noise and Resonance Analysis of On-Chip Power Distribution Network," *Proc. of ISQED*, pp. 163 - 168, 2002.

[3] K. Tang and E. Friedman, "Simultaneous Switching Noise in On-Chip CMOS Power Distribution Networks," IEEE Transactions On Very Large Scale Integration (VLSI) Systems, vol. 10, no. 4, August 2002, 487-493.

[4] H. Chen and D. Ling. "Power Supply Noise Analysis Methodology for Deep-submicron VLSI Chip Design," *Proc. of the 34th DAC*, pp. 638-643, 1997.

[5] T. Mitsuhashi and E.S. Kuh, "Power and Ground Network Topology Optimization for Cell Based VLSIs," *Proc. of the 29th DAC*, pp. 524~529, 1992.

[6] X. Wu, X. Hong, Y. Cai, and et al. "Area minimization of power distribution network using efficient nonlinear programming techniques," *Proc. of ICCAD,* pp.153-157, 2001.

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

[8] J. Fu, Z. Luo, X. Hong, Y. Cai, S. Tan, Z. Pan, "A fast decoupling capacitor budgeting algorithm for robust on-chip power delivery," *Proc. of ASPDAC*, pp. 505-510, 2004.

[9] H. Li, Z. Qi, S. X-D. Tan, L. Wu, "Partitioning-based approach to fast on-chip decap budgeting and minimization," *Proc. of the 42th DAC,* pp. 170-175, 2005.

[10] P.G. Doyle and J.L. Snell, "Random Walks and Electric Networks." arXiv: math.PR/0001057

[11] H. Qian, S. Nassif, and S. Sapatnekar, "Random walks in a supply network." *Proc. of the 40th DAC*, pp. 93-98, 2003.

[12] L. Kang, Y. Cai, J. Shi and X. Hong. "In-depth Exprimental Study of Power Grid Network Analysis using Random Walks Algorithm," *ICCCAS Conf.* pp.2401-2405, 2006.

Table 4. Comparison with none pre-added approach for P/G optimization considering the SSN.

Power Gird	Power Node	Macro Block	Clock Buffer	None Pre-Added		Random Walk based Predecoupling			Speed Up	Over-head
				time(s)	decap(nf)	time(s)	Preadded(nf)	decap(nf)		
1	26	13	2	2.49	0.685745	0	0.576113	0.576113	--	--
2	806	642	102	27.97	0.229297	10.29	0.0934491	0.229394	2.72	0.04%
3	4032	2744	748	63.72	0.797331	29.33	0.181388	0.806774	2.17	1.18%
4	32952	25690	6422	2031.96	5.30618	1031.81	1.80255	5.07767	1.97	--
5	108392	89375	23017	8824.35	6.51945	2528.57	2.33964	6.86403	3.49	5.29%
								Average	2.59	2.17%