

3D Power Distribution Network Co-design for Nanoscale Stacked Silicon ICs

Amirali Shayan¹, Xiang Hu², He Peng¹, Mikhail Popovich³, Wanping Zhang¹
Chung-Kuan Cheng¹, Lew Chua-Eoan³, Xiaoming Chen³

¹CSE Dept., ²ECE Dept., University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0404

Tel: 1-858-534-8174, Fax: 1-858-534-7029, Email: {amirali, x2hu, hepeng, w7zhang, ckcheng}@ucsd.edu

³Qualcomm Inc., San Diego, CA 92121, Email: {mikhailp, xiaoming, lewc}@qualcomm.com

Abstract: In this paper, we propose an efficient flow for the analysis and co-design of large 3D power distribution networks (3D PDN). In this flow, the network is modeled in frequency domain and thus can take advantage of parallel computing. The proposed flow significantly reduces the CPU time while obtaining accurate results as compared to commercial simulation tools. In the established 3D PDN model, we incorporate the on-chip voltage regulator module (VRM) and effect of on-chip inductance. The impact of each design parameter of the 3D PDN on simultaneous switching noise (SSN) is investigated based on the model¹.

1. Introduction

Early design planning is a crucial step as the noise margins become tighter in below 45nm technology and the clock frequency increases. It is important to perform this task in the early design stages along with the through silicon stacking design in order to prevent logic and chip failures [1].

In 3D stacked ICs, about 30-40% of the available die to die vias are used for power delivery and the rest are used for signals [5]. The supply current flows through the inductive solder bumps and narrows through silicon vias with considerable parasitic inductances, results in a significant SSN. The 3D integration requires less than the original 2D design footprint and therefore the current density per pin is increased. As a result the power distribution network in 3D systems needs to be accurately modeled and designed in order to satisfy the noise margins and power demand. Authors in [2] proposed an analytical PDN model for the 3D stacked ICs considering the via inductance and stacked decoupling. However the model does not include the effect of on-chip inductance. As the clock frequency increases to GHz range, SSN from the switching dice will be as significant as the IR drop. Therefore, in the GHz range on-chip inductance in the PDN model should be considered.

In this work, a 3D PDN model is presented considering the effect of on-chip inductance. And on-chip dc-dc converter is adopted as a possible alternative to off-chip VRM. Higher circuit density in 3D PDN results in even larger circuit netlist, which makes the simulation of huge 3D PDN a challenging task. The transient simulation result at one time point depends on the results from previous time points. Thus, it will take a long time to simulate large networks. A parallel flow using Message Passing Interface (MPI) on clustered Linux machines is exploited in this paper which significantly reduces the simulation time.

The rest of the paper is organized as follows: The parallel flow is introduced in section 2. The 3D PDN model is discussed in section 3. In section 4, we will discuss our experiments result based on the grid design parameters are discussed. Conclusions are drawn in section 5.

2. Parallel 3D PDN Analysis Flow

For this work, we have implemented an efficient parallel processing analysis flow for the full power distribution network in both frequency and time domain. Many conventional simulators fail to simulate the system in a reasonable time because of the large scale number of grid in the 3D PDN.

Zhang *et al.* in [5] proposed an analysis method for the on-chip *RLC* power ground models based on the frequency domain for single processor. We adopt their technique, and extend and parallelize it to deal with the large 3D grid size. The frequency domain analysis is performed in a broadband range from DC to GHz to identify the hot spot locations and meet the target impedance for the entire range. The impedance of the PDN is one of the most important design variables. The impact of grid parameters on the 3D PDN impedance is explored in this study. The time domain waveform is recovered from the frequency domain to examine the voltage droop of the PDN during chip performance.

We implement the package using MPI and run the flow on FWgrid infrastructure (<http://fwgrid.ucsd.edu>) using multiprocessor to speedup the simulation time. We run the parallel flow on clustered Linux machine and our results show speedup of up to 22 times with single processor and more than 430 times by using up to 200 processors over HSPICE transient simulation. The PDN simulation time is reduced from hours to less than hundreds of seconds.

3. Design Planning for the 3D Power Grid

In this study, we explore efficient via placement to reduce the power noise. In 3D stacked circuits the power/ground wires are routed by using orthogonal interconnection. The current increases with the new high performance demand. This current goes through the inductive solder bumps and through silicon vias and results in severe power integrity problem. The die to die through silicon vias serve both for signal routing and power delivery. Also, dummy vias are used as heat conductance between the stacked dice [3].

¹The authors would like to acknowledge Qualcomm Inc. for the support of this work.

The on-chip stacked VRM solution will also be considered in 3D model. We vary the location of the on-chip regulator in the stacked die from bottom to center. In the next section, we describe the established 3D PDN model.

3.1 3D Power Distribution Model

The 3D power delivery model that we use in the experiments is a stacked 3D RLC grid, as shown in Fig. 1. In each stack layer a $10 \times 10 \text{ mm}^2$ active die is modeled with RLC network with the current sink stimuli in the center, representing the active switching transistors. The RL through silicon vias with the length of $200 \mu\text{m}$ connect each layer of the stacked dies. The RL values are based on the via processing technology. The SSN is generated by inductive grid. Next, we propose our scheme for on-chip VRM.

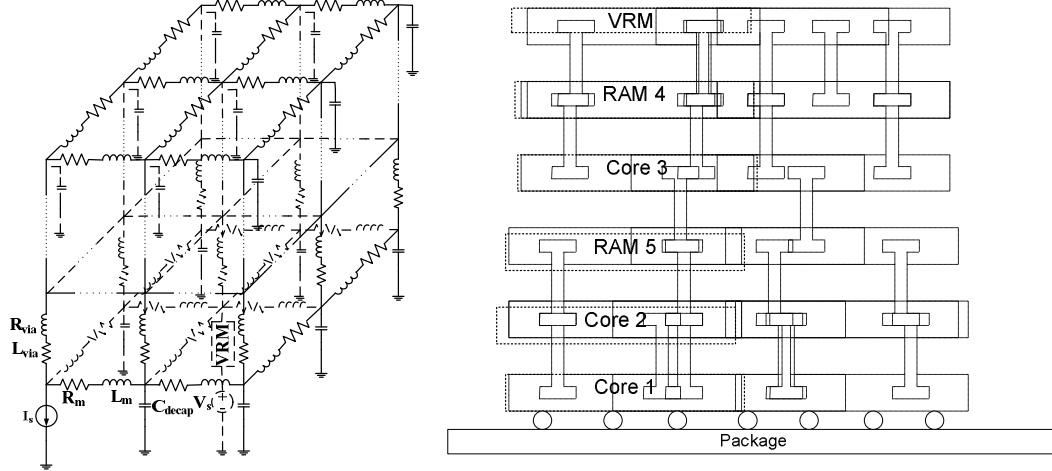


Figure 1. 3D PDN model with on-chip inductance.

3.2 On-chip Voltage Regulator

To reduce the di/dt event from supply voltage regulator variation, an integrated on-chip voltage regulation network in 3D stacked IC is adopted. In conventional power delivery systems, mounted VRM on the board is used which has critical problems: long interconnect path, parasitic inductance of the package-board which generate SSN, large decap demand and large number of the power and ground pins required by chip which takes expensive area and make the packaging complex. The stacked dc-dc converter in the 3D integration benefits from minimum interconnect parasitics, wider bandwidth and easy distribution to multiple domains [4].

It is practical to integrate the regulation circuitry and use small, discrete external inductors mounted close to the die as switching mode power supply. A large number of these external inductors would be required, at least one per PDN node. The inductance can be embedded in the package as well. In addition, these physically small components have limited inductance [1]. We use the close loop impedance of the VRM model in the analytical flow.

4. Experiment Results and Analysis

In this section, we will discuss our experimental results using the proposed 3D PDN model. The changes of the through silicon vias and PDN grid parameters are explored.

4.1 Through Silicon Via Distribution Density

Each die grid is divided into equal quadrant based on the voltage droop and pitch. We allocate two via to the corner of each quadrant in the PDN simulation. The via placement is modeled based on uniform distribution in each tier. We change the via distribution density from 10% to 80% of the total grid geometry. In Fig. 2, we see for low frequency higher via density results in less output impedance but in high frequency the PDN with higher via density has larger impedance peaks. That is because reduced resistance increases the Q factor, which leads to higher anti-resonance.

4.2 3D PDN With and Without On-chip Inductance

Fig. 3 illustrates our 3D PDN impedance profile with and without on-chip inductance. For higher frequencies in the range of GHz the impedance increases up to 8 times with on-chip inductance compared to RC on-chip power grid. This tells us that on-chip inductance should no longer be ignored when operation frequency increases.

4.3 Scaling On-chip Inductance in Tiers

In this section, we scale 3 orders of magnitude the 3D grid on-chip inductance. The low frequency impedance is same but as the frequency increases the impedance is increased up to 10 times (Fig. 4(a)).

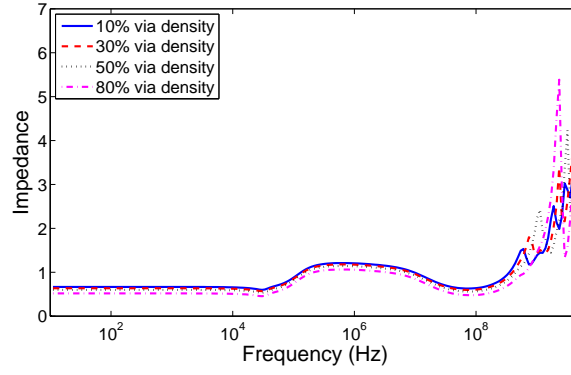


Figure 2. 3D PDN impedance profile v.s. via distribution density.

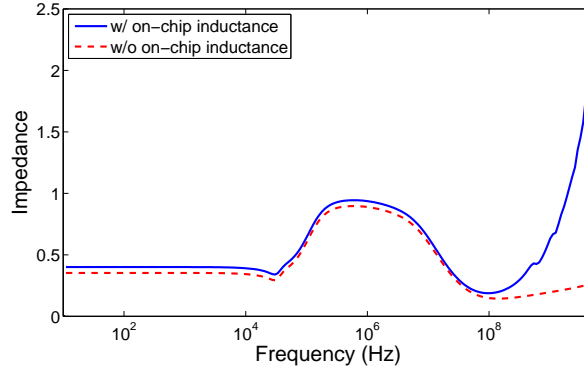


Figure 3. 3D PDN with and without on-chip inductance.

4.4 Scaling 3D PDN On-chip Resistance

It is observed that as the resistance of the tier grid increases in low frequencies, the 3D PDN impedance is high. This can be explained by the fact that as the frequency increases the damping effect of the R is dominant and compensate the increase in the low frequency impedance (Fig. 4(b)).

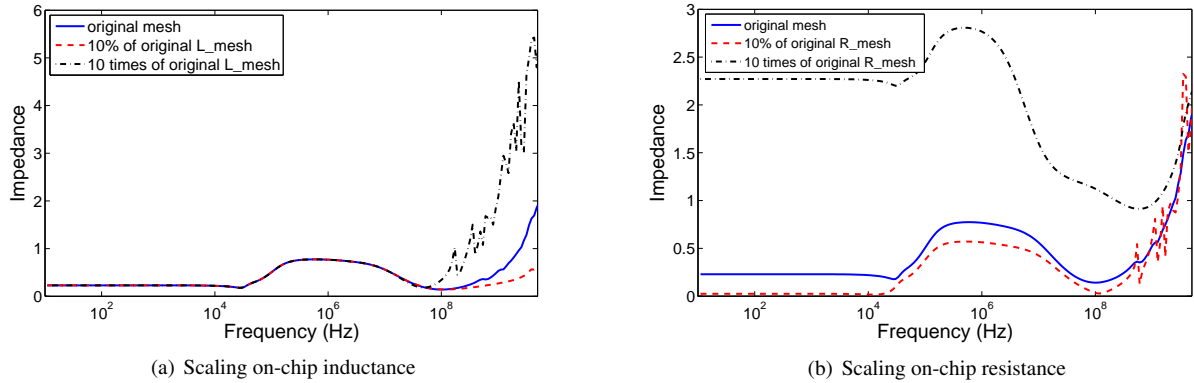
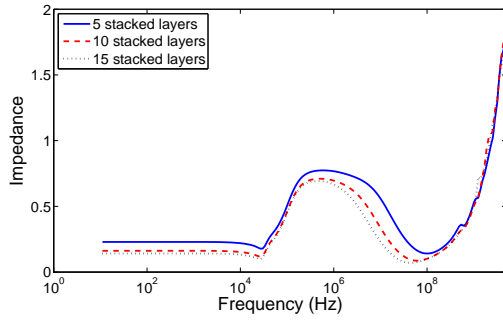


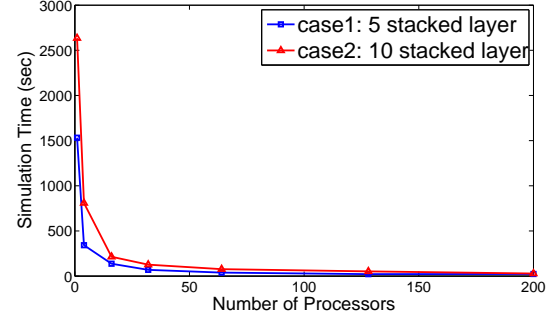
Figure 4. Scaling 3D PDN on-chip parameters.

4.5 3D Stacked Layers

In this section, we show how increasing the number of stacked layers from 5 to 10 dice affects the impedance profile of the stacked silicon. We run the flow on the 3D power grid structures with the dimension of 32×119 nodes in the grid and 5 to 15 stacked die layers. From Fig. 5(a) we observe the impedance is reduced with more stacked layer. Fig. 5(b) shows simulation time versus number of processors for both cases on FWgrid machine with up to 200 processors where simulation times is reduced significantly to less than a minute. Simulations of the case 1 and 2 in HSPICE took 9911 sec and 110479 sec respectively for 300 nsec duration (Table 1).



(a) The impedance profile for multi stacked layers



(b) 3D PDN simulation CPU time v.s. number of processors

Figure 5. The interconnect structure used in this work.

Table 1. 3D grid simulation time v.s. number of processors on FWgrid for the flow and HSPICE

Simulation time	HSPICE(sec)	Parallel flow running on clustered linux(No. of processors)						
		1 prcs.(sec)	4 prcs.(sec)	16 prcs.(sec)	32 prcs.(sec)	64 prcs.(sec)	128 prcs.(sec)	200 prcs.(sec)
case1	9911.32	1531.80	342.20	136.70	68.70	38.57	21.35	23.01
case2	110479.36	2635.06	808.90	214.40	126.50	77.00	52.30	28.75

4.6 Analysis of 3D PDN in Time Domain

As concluded in [2], the higher via density will result in lower voltage noise. However, this is not always the case. In figure 2, we can see that for a certain via placement strategy, higher density may lead to lower impedance at low frequencies, it may also result in higher resonance peak at higher frequencies. This is because more parallel vias reduce effective resistance, and thus increase the quality factor Q of the RLC tank which makes anti-resonance more pronounced. As a result, the output voltage noise depends on the frequency-domain profile of the input noise current. Fig. 6 shows the output voltage noise for different via densities. From Fig. 6, it is observed that higher via density may result in larger voltage noise.

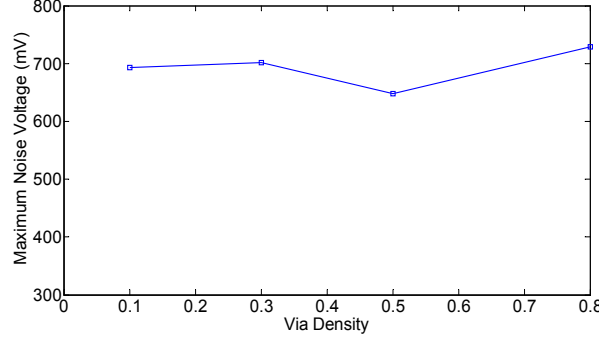


Figure 6. 3D PDN time domain noise magnitude v.s. via density.

5. Conclusions

In this work, we developed an efficient parallel flow for the analysis of the 3D PDN. We explored different design parameters including the impact of the on-chip inductance, on-chip regulator and inductive through silicon via in the 3D PDN .

References

- [1] M. M. Budnik and K. Roy. A Power Delivery and Decoupling Network Minimizing Ohmic Loss and Supply Voltage Variation in Silicon Nanoscale Technologies. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 14(12):1336–1346, Dec. 2006.
- [2] G. Huang, M. Bakir, A. Naeemi, H. Chen, and J. D. Meindl. Power delivery for 3d chip stacks: Physical modeling and design implication. In *IEEE Electrical Performance of Electronic Packaging, 2007*, pages 205–208, Oct. 2007.
- [3] C. Ryu, D. Chung, J. Lee, K. Lee, T. Oh, and J. Kim. High Frequency Electrical Circuit Model of Chip-to-chip Vertical Via Interconnection for 3-D Chip Stacking Package. In *IEEE 14th Topical Meeting on Electrical Performance of Electronic Packaging, 2005*, pages 151–154, Oct. 2005.
- [4] J. Sun, J.-Q. Lu, D. Giuliano, T. P. Chow, and R. J. Gutmann. 3D Power Delivery for Microprocessors and High-Performance ASICs. In *IEEE 22nd Applied Power Electronics Conference, 2007*, pages 127–133, Mar. 2007.
- [5] W. Zhang, L. Zhang, R. Shi, H. Peng, Z. Zhu, L. Chua-Eoan, R. Murgai, T. Shibuya, N. Ito, and C. K. Cheng. Fast Power Network Analysis with Multiple Clock Domains. In *25th International Conference on Computer Design, 2007*, pages 456–463, Oct. 2007.