Framework for Statistical Analysis of Homogeneous Multicore Power Grid Networks

Guanglei Liu, Jeffrey Fan*

Abstract— *In this paper, we propose a framework to analyze the large-scaled multicore power grid network statistically by first building a simplified multicore power supply distribution model. We then apply the Modified Nodal Analysis (MNA) method on a simplified power gird circuit. Under such a framework, most statistical approaches, including Monte Carlo (MC), Importance Sampling, and Stochastic Spectrum Analysis, can be applied to analyze the process-induced variation of homogeneous multicore power grid networks. In the experiment, we focus on the subthreshold leakage current variations, which are modeled as lognormal distribution random variables, by using MC approach as an example to demonstrate the feasibility of such a framework.* [1]

Index Terms – Multicore, Power Grid Network, Statistical Analysis, Process Variation

I. Introduction

In the past a few years, multicore processors have been widely used in the industry. By integrating several cores onto a single chip, significant performance improvement can be achieved. In addition, the multicore processor has a big advantage over the software parallelism, *i.e.* software running in parallel at the same time. Consequently, it has been attracting much attention lately. Until now, many multicore products have been released to the market. However, it is a troublesome task to achieve full performance from a multicore processor. Performance improvement needs to be achieved by increasing the clock frequency and numbers of transistors per die. It means designers will face the challenge of processor power consumption and heat dissipation problem. Because higher temperature can increase leakage current, thus may lead to degradation in performance [1]. As a result, the process-induced leakage current variations become a big concern.

The technology scaling in industry has continuously driven towards higher levels of integration, higher frequencies and lower operating voltages. For technologies down to 90nm, it is possible to increase performance, while reducing power from one processor generation to the next. However, for 45nm process induced variation, it has huge impacts on the circuit performance. One of the major variations comes from the leakage current, and it has a rapid increasing rate about five to ten times every technology generation [2]. It is because the leakage current has exponential relationship with threshold voltage. Hence, leakage current is highly sensitive to threshold voltage. It has become the number one concern in the design of multicore power grid network. Due to the important impacts of leakage current on the multicore power delivery networks, a number of works have been proposed to perform the stochastic analysis of power gird network under process-induced leakage current variations [3], [4], [5].

[1]Guanglei Liu and *Jeffrey Fan are with the Department of Electrical and Computer Engineering, Florida International University, Miami, FL 33174, USA, e-mail: {gliu002, jeffrey.fan}@fiu.edu

The main contribution of this paper is that we propose a framework to analyze the multicore power grid network performance considering the process-induced variation. With this framework, many statistical approaches, such as Monte Carlo, Importance Sampling and Stochastic Spectrum Analysis, can be applied. The experimental result illustrates the feasibility of our proposed framework. The remainder of the paper is organized as follows: We present our simplified multicore power grid model in Section II. The Monte Carlo and Importance Sampling method are briefly discussed on the power grid model to analyze the process-induced leakage current variations in Section III. We introduce the Hermit PC method to address the issue of leakage current modeled as a lognormal random variable in section IV. Experimental result is given in Section V, and we conclude this paper in Section VI.

II. Power Grid Model

In this section, we build a complete multicore power supply distribution model, which consists of a power mesh with C4 packages, resistance, capacitance, decoupling capacitors (decaps), and current sources on the grid. Thus a framework model can be shown in Figure 1. A homogeneous multicore system structure can be represented by several replicated core models as in Figure 2, where the deterministic component is a piecewise linear current source. The total leakage current is the sum of deterministic leakage current plus stochastic component, which is modeled as normalized lognormal distribution in consideration of process variations [7].

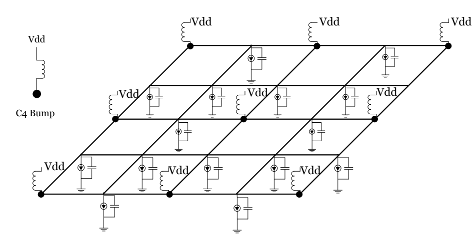

Fig. 1. Core model

Each core model is a collection of four connected base grid blocks. Each base grid is bounded by C4s as show in figure 3(a). To illustrate the framework structure, we assume each base grid is divided into a 2 $\times$ 2 circuit gird consisting of current sources and decoupling capacitances. Each based grid in a core model presents one component in the CPU module, such as CPU core, cache, decoders, and ALU. We decompose the core grid into smaller grids, while maintaining the systems behavior. Such decomposition can help us to analyze each base grid separately and reuse the result to analyze the core model.

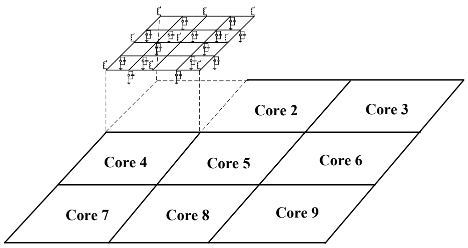

Fig. 2. Multi-core system model

In this framework of power grid networks, the current source and decoupling capacitances in a core grid are uniformly distributed over the core nodes, and decaps are separated proportionally among the neighboring bases. Decoupling capacitances of the base gird are extracted from the corresponding circuits. This analysis is also valid for non-uniform distribution of decaps and current sources. We assume the uniformity to simplify the explanation. Also, we use a core grid with four connected base grids for the same reason. It is valid for those cases with more than four base grids per core.

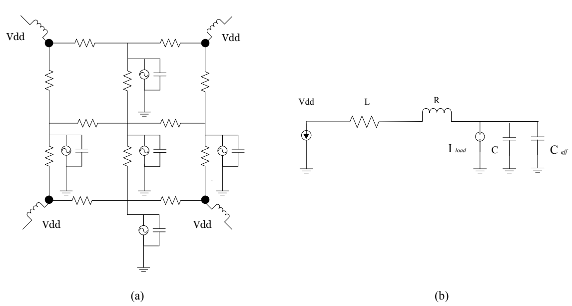

Fig. 3. (a) Base grid model (b) Simplified core model

Because each core grid consists of 4 connected base grids, we can analyze of the base grid to derive the node voltage of the core gird. For the core grid, each connected base girds can be represented by its simplified form [6], shown in figure 3(b). Replacing the base grids with the simplified circuit, the core model can be shown in figure 4(a), where the impedances $R_{12}, R_{13}, R_{24}, R_{34}$ in the circuit represent the local power gird branches between bases 1, 2, 3 and 4. Such representation can simplify the core grid structure and it is easier for us to analyze. The node voltages on the core grid can be expressed in terms of the node voltages derived from the base grid analysis. Equation (2) illustrates the relation between node voltage on core grid and node voltage on base grid in a matrix form.

$$A \cdot V_{core} = B \cdot V_{base} \tag{1}$$

Where the matrices A and B represent the conductance parameters for the core and base grids. The equation above is only valid when all bases on the same core are operational with the same switching frequency.

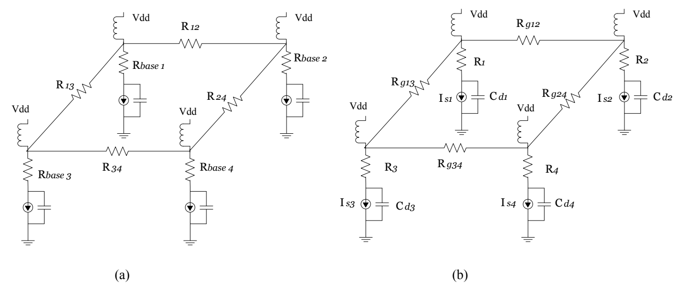

Fig. 4. (a) simplified core model (b) simplified multi core model

The simplified core model is used to build a global grid multicore system. We significantly reduce the complexity of the global grid without losing its generalization and accuracy by assuming the multicore grid is a 2×2 system. In this regard, Figure 4(b) shows the global grid structure with simplified core models for a 2 × 2 multicore system. $Rg_{12}, Rg_{13}, Rg_{24}$ and Rg_{34} are the impedances. For the global grid multicore system, each core will operate at different frequency with different workload assigned to the associated core. Under such a scenario, we can then implement the statistical analysis on the simplified global grid by using the statistical analysis, such as MC method, to consider the stochastic behavior of the power grid network, *i.e.* leakage current variations.

III. Problem Formulation

A. Power Grid Network Model

We implement the Modified Nodal Analysis (MNA) to the multicore power grid model, so the simplified circuit can be represented by matrix. Power grid networks are modeled as RC networks and the current sources are time-variant. Some nodes may have known voltage values as constant voltage sources. Also, for the C4 power grids mentioned in the last section, the known voltage nodes are internal nodes inside the power grid. Thus, we may obtain the node voltage by solving the following MNA equations.

$$Gv(t) + C\frac{dv(t)}{dt} = I(t) \tag{2}$$

where $I(t)$ is given deterministic vector of current sources, G is the conductance matrix, and C is the capacitance matrix. $v(t)$ is the vector of time varying node voltage and branch currents of voltage source. Since we can calculate the node voltage by using the MNA equation, we will introduce how to use MC method to generate normal distributed variations. Thus, we can then analyze the performance of such a multicore power grid system.

B. Monte Carlo Method Modeling Leakage Current

We apply a statistical approach for modeling manufacturing variations, such as metal width, gate length and thickness of the gate oxide, etc. Those factors will cause variations in conductance matrix G and capacitance matrix C. Meanwhile, variations in threshold voltage will cause leakage current variation. Consequently, taking into account of the aforementioned variations, the MNA equation (3) can be replaced by

$$G(\xi_g)v(t) + C(\xi_c)\frac{dv(t)}{dt} = I(t, \xi(\theta)) \quad (3)$$

The conductance and capacitance matrix can be expressed as

$$G(\xi_g) = G_1 + G_2\xi_g, C(\xi_c) = C_1 + C_2\xi_c \quad (4)$$

where G_1 and C_1 represent the deterministic components of conductance and capacitance of circuit, respectively. G_2 and C_2 represent sensitivity matrix of conductance and capacitance. ξ_g and ξ_c are normalized random variables with normal distribution. They represent process variations in conductance and capacitor. For illustration purpose, we perform MC analysis by randomly generating a set of lognormal distributed $G(\xi)$, $C(\xi)$ and $I(t\xi(\theta))$. Each of the matrices has the same size as the deterministic matrices and all of them will take part in the computation.

In the mathematical term, the expected value of a function $g(x)$ regarding the MC method can be express as

$$g(\xi) = \int_{x\in\alpha} g(x)f\xi(x) \quad (5)$$

where ξ is a random variable, $f\xi(x)$ is the probability function of ξ, and it is bigger than zero on any set of values α. Taking n samples of ξ, $(x_1, ..., x_n)$, we can calculate the mean and variance of $g(x)$ at any time instance.

$$\widetilde{g}_n(x) = \frac{1}{n}\sum_{i=1}^{n} g(x_i) \quad (6)$$

However, MC method requires a massive computation if the sample space is very large. To overcome this defect, we introduce a possible speed-up technique in the next subsection.

C. Importance Sampling

Importance sampling has long been recognized as a useful technique for increasing the efficiency of Monte Carlo algorithm. The concept is based on reducing the variance by choosing a proper density function. That is, certain values of the input random variables in a simulation have more impact on the parameter being estimated than others. If those important variables are sampled more frequently, we can get a new distribution with less variables. In this case, the number of computation can be reduced. Let $f_*(x)$ be an alternate density function, the importance sampling can be expressed as

$$g(\xi) = \int_{x\in\alpha} \frac{f(x)}{f_*(x)} f_*(x)dx \quad (7)$$

where $W(\xi) = \frac{f(x)}{f_*(x)}$ is the weighting function. As long as $f_*(x) \neq 0$ for any $x \in \alpha$ and $f(x) \neq 0$, we have Monte Carlo estimator

$$\widetilde{g}_n(\xi) = \frac{1}{n}\sum_{i=1}^{n}\frac{f(\xi_i)}{f_*(\xi_i)}, \xi_i \approx f_*(x) \quad (8)$$

The importance sampling technique focuses on finding a proper density $f_*(x)$, thus the variances of the importance sampling estimator is less than the conventional Monte Carlo Method.

IV. HERMITE POLYNOMIAL CHAOS FOR LEAKAGE CURRENT VARIATIONS

Using the Hermite Polynomial Chaos (Hermite PC) to model the lognormal distributed leakage current, we let $g(\xi)$ be the Gaussian random variable. $l(\xi)$ is the random variable obtained by taking the exponential of $g(\xi)$

$$l(\xi) = e^g(\xi), g(\xi) = ln(l(\xi)) \quad (9)$$

For equation (1), the MOS device leakage current I_{off} can be express by

$$I_{off} = cI_l(V_{th}) = ce^{-V_{th}} \quad (10)$$

where $I_l(V_{th})$ is a lognormal random variable. Let μ_g and σ_g^2 denote the mean and the variance of $g(\xi)$, then the mean and the variance of $l(\xi)$ can be expressed by

$$\mu_l = e^{(\mu_g + \frac{\sigma_g^2}{2})} \quad (11)$$

$$\sigma_l^2 = e^{(2\mu_g + \sigma_g^2)}[e^{\sigma_g^2} - 1] \quad (12)$$

The general Gaussian variable $g(\xi)$ can be expressed by the following form

$$g(\xi) = \sum_{i=0}^{n} \xi_i g_i \quad (13)$$

Note that ξ_i are Gaussian variables. This means $<\xi_i, \xi_j>= \delta_{ij}$, $< \xi_i >$ = 0 and ξ_0=1. We can use the Karhunen-Loeve orthogonal expansion method to get this form [9]. For illustration purpose, we can represent the lognormal random variable, $l(\xi)$, by using the Hermite PCs expansion

$$l(\xi) = \sum_{k=0}^{p} l_k H_k^n(\xi) \quad (14)$$

and l_0 = $\exp[\mu_g + \frac{\sigma^2}{2}]$. We apply Galerkin method [10] to find the coefficients on $l(\xi)$.

$$l_k(t) = \frac{< l(t,\xi) >, H_k(\xi)}{< H_k^2(\xi) >}, \forall k \in (0, ..., P) \quad (15)$$

Therefore, the lognormal process can be expressed by

$$l(\xi) = l_0(1+\sum_{i=1}^{n}\xi_i g_i + \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{(\xi_i\xi_j - \delta_{ij})}{\langle(\xi_i\xi_j - \delta_{ij})^2\rangle} g_i g_j + ...) \quad (16)$$

To simplify the explanation, we only consider one Gaussian Variable with Hermite PCs. Thus, we have $\xi = [\xi_1]$. For the second-order Hermite PC (P = 2), we implement equation (12) and obtain the following equation

$$l(\xi) = l_0(1 + \delta_g \xi_1 + \frac{1}{2}\sigma_g^2(\xi_1^2 - 1)) \quad (17)$$

From the above equation (13), we can find the desired Hermite PC coefficients $I_{0,1,2}$ that can be written as l_0, $I_0\sigma_g$ and $\frac{1}{2}l_0\sigma_g^2$. After combining the MNA equation (3) and equation (13), we can then calculate the mean and variance of node voltage easily.

V. Experimental Results

To illustrate the feasibility of the proposed framework, we apply Monte Carlo Method to our proposed system to model the leakage current as lognormal distribution random variables. Our simulation environment is based on Windows system with dual Intel Core 2 CPU with 1.86 GHz and 1GB memory. The proposed method has been implemented in Matlab. Because the MC sampling method has a error percentage of $1/\sqrt{N}$, where N is the number of samples. In this paper, we sample up to 10,000 times to get a 99% of accuracy.

For Monte Carlo sampling method, we first randomly select a set of normalized lognormal distribution value as the Gaussian variables, and add those variables to the G, C matrices as shown in equation (5), then the node voltage can be calculated. The whole process is set as one trail of stochastic node voltage value. Figure 5 shows the simulation result of the 10,000-sampling trail. Each blue line represents the node voltage considering the normal distributed random variables from 0 ns to 100 ns. The red lines indicate the upper bound ($\mu + 3\sigma$), mean (μ) and lower bound ($\mu - 3\sigma$) from top to the bottom, respectively. From the observation, we can obtain the normal distributed values, which are computed by using MC method and are perfectly matched in the middle of upper bound and lower bound.

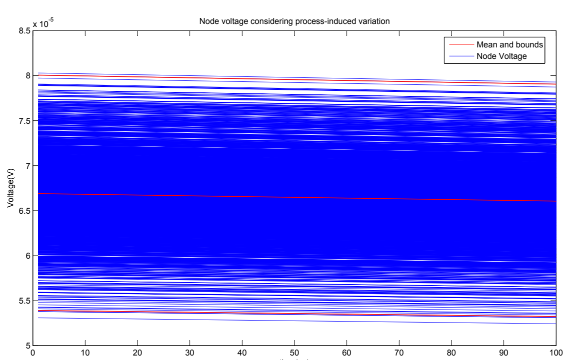

Fig. 5. Voltage distribution caused by the leakage current in a given node with one variable, from 0 to 100 ns.

Figure 6 shows the other experimental result of voltage distribution in a given node with one Gaussian variable. We set the time at 70 ns and collect the data after 10,000 times. The X axis indicates the given node voltage and Y axis denotes the number of occurrences with total trail size of 10,000. In figure 6, the middle vertical line denotes the mean value of node voltage without variations. The left vertical line is the lower bound and the right vertical line is the upper bound.

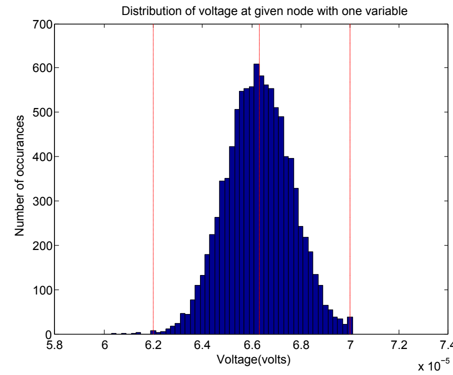

Fig. 6. Voltage distribution of a given node at time 70 ns with one Gaussian variable.

VI. Conclusion

In this work, we propose a framework for homogeneous multicore network. With this framework, many statistical approaches can be applied to analyze the multicore power grid network with process-induced variations. For illustration purpose, we model a framework for homogeneous multicore power gird network. The heterogeneous multicore modeling with more efficient statistical analysis will be analyzed in the future works.

References

[1] R. Bergamaschi, H. Guoling, A. Buyuktosunoglu, H. Patel, I. Nair, G. Dittmann, G. Janssen, N. Dhanwada, H. Zhigang, P. Bose, J. Darringer, *Exploring power management in multi-core systems*, Design Automation Conference, 2008. Volume, Issue , 21-24 March 2008 Page(s):708-713.

[2] V. De, S. Borkar, *Technology and design challenges for low power and high performance*, in Proc. Int. Symp. Low Power Electron. Design (ISPLED), Aug. 1999, pp. 163-68.

[3] P. Ghanta, S. Vrudhula, R. Panda, and J. Wang, *Stochastic power grid analysis considering process variations*, in Proc. Eur. Design Test Conf. (DATE), 2005, vol. 2, pp. 964-69.

[4] J.Wang, P. Ghanta, and S. Vrudhula, *Stochastic analysis of interconnect performance in the presence of process variations*, in Proc. Int. Conf. Comput.-Aided Design (ICCAD), Nov. 2004, pp. 880-86.

[5] S. Vrudhula, J. M. Wang, and P. Ghanta *Hermite polynomial based interconnect analysis in the presence of process variations*, IEEE Trans. Comput.-Aided Design Integr. Circuits Syst., vol. 25, no. 10, 2006.

[6] A Todri, J Kozhaya, M Marek-Sadowska *Power Supply Noise Aware Workload Assignment for Multi-Core Systems*, IEEE/ACM International Conference on Computer-Aided Design, 2008.

[7] J. Fan, N. Mi, S. Tan *Voltage drop reduction for on-chip power delivery considering leakage current variations*, IEEE 25th International Conference on Computer Design (ICCD'07), Lake Tahoe, CA, pp. 78-83, Oct. 7-10, 2007.

[8] Peter W. Glynn and Donald L. Iglehart *Importance Sampling for Stochastic Simulations*, Management Science, Vol. 35, No. 11, Focussed Issue on Variance Reduction Methods in Simulation (Nov., 1989), pp. 1367-1392.

[9] R. G. Ghanem and P. D. Spanos *Stochastic Finite Elements: A Spectral Approach*, Dover publications, 2003.

[10] R. G. Ghanem *The nonlinear Gaussian spectrum of log-normal stochastic processes and variables*, Journal of Applied Mechanics, 66:964-973, December 1999.