Design Space Exploration for Robust Power Delivery in TSV Based 3-D Systems-on-Chip

Suhas M. Satheesh
High-Speed Fabrics Team
NVIDIA
Santa Clara, California 95050
ssatheesh@nvidia.com

Emre Salman
Department of Electrical and Computer Engineering
Stony Brook University
Stony Brook, New York 11794
emre@ece.sunysb.edu

Abstract—3-D integration technologies offer significant advantages to develop multiprocessor systems-on-chip with embedded memory. Reliable power distribution is a challenging issue in these systems due to multiple planes and through silicon vias (TSVs). The two primary TSV technologies, via-first and via-last, have been evaluated for power delivery in a 32 nm 3-D system with eight memory planes and one processor plane. Since the impedance characteristics of via-first and via-last based TSVs are significantly different due to distinct filling materials and dimensions, the power distribution network in each case exhibits different design requirements. A valid design space is identified for both cases. Despite the low parasitic resistance of a via-last TSV, a power network based on via-last TSVs produces signal routing blockages. Furthermore, via-last TSVs exhibit high inductive behavior, producing a non-monotonic design space. It is demonstrated that via-first TSVs can satisfy the power supply noise at the expense of 7.5% additional area as compared to via-last TSVs.

I. INTRODUCTION

In the past decade, three-dimensional (3-D) integration has emerged as a promising technology that maintains the benefits of miniaturization by enabling higher integration density and enhancing system performance [1]. An important application of heterogeneous 3-D integration is stacked processor-memory systems to alleviate the existing gap between logic blocks and memory units in high performance microprocessors. Additional dynamic random access memory (DRAM) is stacked with the processor cores, as depicted in Fig. 1. Such a system significantly increases the memory bandwidth and reduces memory access time. A significant circuit- and physical-level challenge in this system is to design a robust power distribution network that achieves reliable power delivery to each die.

Maintaining the power network impedance smaller than a target impedance is a difficult task due to reduced operating voltages, increased current magnitudes, and the existence of multiple dies and TSVs. Previous work on 3-D power delivery has focused on different power distribution topologies, effect of TSV geometry, and effect of core versus coaxial TSVs [2]–[5]. Circuit-level analysis and comparison of the two primary TSV technologies, i.e., via-first and via-last, however, has not received much attention, primarily from the design perspective.

The primary contribution of this paper is two-fold: (1) the use of via-first and via-last TSV technologies for power distribution in 3-D processor-memory systems is explored in a comparative manner, (2) design space that satisfies power supply noise while minimizing the physical area is determined for both via-first and via-last TSVs. Different design requirements are identified for both cases. Since via-first and via-last methods have different advantages and limitations both at the fabrication and circuit levels, exploring design tradeoffs provides valuable guidelines in developing 3-D power networks.

The rest of the paper is organized as follows. The characteristics of the via-first and via-last technologies are summarized in Section II. The proposed electrical models for 3-D power distribution for both technologies are described in Section III. The results of the proposed analysis are provided in Section IV. Finally, the paper is concluded in Section V.

II. VIA-FIRST VERSUS VIA-LAST TSV TECHNOLOGIES

In the via-first method, TSVs are fabricated before the transistors are patterned in silicon, i.e., prior to front-end-of-line (FEOL) [6]. Thus, TSVs fabricated with the via-first technique do not pass through the metalization layers, as depicted in Fig. 2. The TSV of a plane is connected between the first metal layer of the same plane and the top most metal layer of the previous plane. Polysilicon is typically used as the filling material due to its ability to withstand high temperatures [6]. Alternatively, in the via-last approach, TSV formation occurs after the metalization layers are fabricated, i.e., after back-end-of-line (BEOL) [6]. Thus, via-last based TSVs pass through the metal layers, as depicted in Fig. 2. A lower resistivity filling material such as copper is used [6].
Silicon
Via-last TSV
(Copper)

Fig. 2. Illustration of via-first and via-last techniques to fabricate a TSV in 3-D technologies.

The connection between the TSV and metal layers is typically achieved at the top most metal layer. The dimensions of the via-last based TSVs are also typically greater than the via-first based TSVs [7]. Despite the significant advantage of having a low resistivity filling material, via-last based TSVs have several circuit and process related limitations: signal routing blockages, higher inductive characteristics, requirement for backside lithography and low temperature process, and high sensitivity to contamination [6], [8]. Alternatively, via-first TSVs require a high resistivity filling material. Due to these different aspects, a circuit-level analysis is performed to evaluate and compare the performance of via-first and via-last based TSVs in power delivery. A design space is determined for both cases to satisfy power supply noise while minimizing the physical area.

III. ELECTRICAL MODELS FOR 3-D POWER DISTRIBUTION

A 3-D system designed with a 32 nm CMOS technology consisting of eight memory planes and one plane for the processor cores is assumed. The power supply voltage is equal to one volt. Note that the processor plane is placed closest to the heat sink (due to high switching activity), and therefore farthest from the power supply pads.

Each memory plane has one gigabit DRAM, producing a total of one gigabyte memory. The one gigabit memory in each die is divided into 32 modules, each consisting of 32 megabit memory and consumes an area of $1500 \times 2500 \mu m^2$. For a via-last based power distribution network, the power and ground TSVs are distributed on both sides of each memory module, as depicted in Fig. 3. Alternatively, in a via-first based power network, power and ground TSVs are located within the memory module, i.e., beneath the active devices rather than on the sides of the memory module.

In the rest of this paper, the analysis is performed for a total of 32 megabyte DRAM (eight planes times 32 megabit). Part of the processor die that corresponds to the area of a 32 megabit memory is also considered in the analysis, producing a total of nine stacks. The procedure is similar for the remaining 31 modules.

An equivalent electrical model that represents power distribution within a single module is shown in Fig. 4. The model consists of power and ground TSVs, the substrate between these two TSVs, power distribution network within a plane, and switching circuit, as described in the following sections. Note that in addition to these impedances, the parasitic package resistance and inductance are, respectively, 3 mΩ and 100 pH at both the power and ground supplies. Also note that the physical interconnect parameters (such as number of metal layers within a plane, aspect ratio, thickness, and pitch) are determined based a 32 nm technology [9].

A. TSV Model

A TSV is typically represented as a cylinder with a diameter $W$ and depth $H$. Aspect ratio of a TSV is given by $H/W$. The minimum distance between the two TSVs is determined by the pitch $P$. The TSV model consists of a resistance $R_{tsv}$ and inductance $L_{tsv}$ due to the filling material, and a capacitance to the substrate $C_{tsv}$ due to the thin dielectric layer [10], [11]. $R_{tsv}$ is determined by [12],

$$R_{tsv} = \sqrt{(R_{DC}^{tsv})^2 + (R_{AC}^{tsv})^2}, \quad (1)$$

where the DC resistance $R_{DC}^{tsv}$ and AC resistance $R_{AC}^{tsv}$ are, respectively,

$$R_{DC}^{tsv} = \frac{\rho H}{\pi (W/2)^2}, \quad (2)$$
TABLE I

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Via-first</th>
<th>Via-last</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diameter W</td>
<td>4 (\mu m)</td>
<td>10 (\mu m)</td>
</tr>
<tr>
<td>Height H</td>
<td>10 (\mu m)</td>
<td>60 (\mu m)</td>
</tr>
<tr>
<td>Pitch P</td>
<td>8 (\mu m)</td>
<td>20 (\mu m)</td>
</tr>
<tr>
<td>TSV resistance (R_{tsv})</td>
<td>5.7 (\Omega)</td>
<td>20 (\mu m)</td>
</tr>
<tr>
<td>TSV inductance (L_{tsv})</td>
<td>4.2 (nH)</td>
<td>35 (nH)</td>
</tr>
<tr>
<td>TSV capacitance (C_{tsv})</td>
<td>23 (fF)</td>
<td>283 (fF)</td>
</tr>
<tr>
<td>Material resistivity (doped polysilicon/copper) (\rho_f)</td>
<td>7.2 m(\Omega)</td>
<td>16.8 m(\Omega)</td>
</tr>
</tbody>
</table>

For a via-first based power network, since the TSVs connect the first metal layer of the same plane and the top most metal layer of the previous plane, the connection between the first and the top most metal layers within the same plane is achieved by a stack of vias. This resistance is modeled with \(R_{vertical} = 1.97 \Omega\). Furthermore, since the TSV lands on the first metal layer, power supply voltage can be directly distributed through the first metal layer, also referred to as an alternative current path. This resistance is modeled with \(R_{M1} = 73 \Omega\). For via-last TSVs, \(R_{vertical}\) is not required since the TSVs pass through the metal layers. Also, since via-last TSVs land at the top most metal layer, alternative current paths do not exist, i.e., \(R_{M1}\) approaches infinity. Finally, the equivalent resistance of the power distribution network from the top metal layer to the devices is modeled with \(R_{Network}\), which is equal to 1.97 m\(\Omega\) for a via-first power network and to 8.51 m\(\Omega\) for a via-last power network. These values are determined by modeling the entire power network of a plane with a two-layer mesh, and assuming 32 nm technology parameters.

D. Switching Circuit

The switching circuit for both memory and processor is modeled as a piecewise linear current source, as shown in Fig. 5. Note that the current has a DC magnitude of \(I_{DC}\), peak value, and rise (100 ps) and fall (150 ps) times. For the processor plane, \(I_{DC}\) is determined from the leakage power consumption (assumed to be 30% of the overall power consumption) as 844 mA whereas the peak current is determined from the overall power consumption (90 W) as 7.1 A. The current profile for each memory plane is obtained similarly, assuming that the power consumption of the one gigabyte DRAM is 3 W. Note that the values of \(I_{DC}\) and peak current listed here are only for a single module that corresponds to an area of 1500 \(\times\) 2500 \(\mu m^2\), as previously mentioned.

IV. ANALYSIS RESULTS

A. Valid Design Space to Satisfy Power Supply Noise

Power supply noise is observed across the current source located at the processor plane which is farthest from the supply pads. To evaluate the valid design space, the models described in the previous section are utilized and the power supply noise is analyzed as a function of decoupling capacitance and number of TSVs. These results are depicted in Figs. 6(a) and 7(a) for, respectively, via-first and via-last TSVs. Note that the number of TSVs in these figures refers to only power and ground nodes is determined, and a contour graph is obtained from the noise surface. These contours are depicted in Figs. 6(b) and 7(b) for, respectively, via-first and via-last TSVs.
Each point on the curve in Fig. 6(b) represents a valid pair of decoupling capacitance and number of TSVs for a via-first based power network. Alternatively, any point under the curve does not satisfy the target power supply noise. For example, a 5 nF decoupling capacitance and 2500 TSVs produce a maximum of 100 mV power supply noise. The target noise can also be achieved by decreasing the decoupling capacitance to 0.7 nF, and increasing the number of TSVs to 3500.

The valid design space in via-first TSVs is continuous due to the monotonic relationship between power supply noise, decoupling capacitance, and number of TSVs. This characteristic is due to the highly resistive behavior of via-first TSVs where the inductive effects are suppressed and the power distribution network is overdamped. Alternatively, in a power distribution network utilizing via-last TSVs, the design space exhibits discrete characteristics due to the underdamped behavior. Decoupling capacitance, TSV inductance, and package inductance produce multiple resonant frequencies where the power supply noise exceeds the design objective. Thus, the amount of decoupling capacitance should be carefully determined in power distribution networks with via-last TSVs to avoid resonance [15]. For example, at 1 nF of decoupling capacitance and 250 TSVs, the power supply noise is below 100 mV. However, if the decoupling capacitance is increased to 2 nF, the noise increases since the resonant frequency is reduced. Alternatively, if the number of TSVs is increased to 350, the power supply noise increases due to a lower damping factor. Thus, in via-last TSVs, if the resonant behavior is not considered, the power supply noise can dramatically increase, as depicted in Fig. 7(a). The amount of decoupling capacitance and number of TSVs should therefore fall within the valid region, as illustrated in Fig. 7(b), to ensure robust power delivery.

B. Optimum Number of TSVs and Decoupling Capacitance

Since multiple valid points exist, the amount of decoupling capacitance and number of TSVs can be chosen to minimize physical area while satisfying the power supply noise. Assum-
Ing a MOS-C with a capacitance density of 39.4 fF/µm² [9], this optimum point corresponds to 2750 TSVs and 2.7 nF of decoupling capacitance for a power distribution network with via-first TSVs. For via-last TSVs, the optimum number of TSVs and decoupling capacitance are reduced, respectively, to 76 and 0.39 nF. Note that 6.25 times more via-first TSVs than via-last TSVs can be placed within a constant area due to smaller via-first TSV dimensions. The area penalty of the via-first TSVs is approximately 7.5% higher than the via-last TSVs. These results are listed in Table II.

V. CONCLUSIONS

Two different TSV technologies, via-first and via-last, have been evaluated to distribute power in a 32 nm stacked 3-D system with eight memory planes and one processor plane. A design space is developed to satisfy the target power supply noise for both via-first and via-last based power distribution networks. The optimum number of TSVs and amount of decoupling capacitance that minimize physical area are also determined. Despite the low parasitic resistance of a via-last TSV, a power network based on via-last TSVs produces signal routing blockages and exhibits high inductive behavior, producing a non-monotonic design space. A power network based on via-first TSVs is overdamped and satisfies the power supply noise at the expense of 7.5% additional area as compared to via-last TSVs.

REFERENCES


<table>
<thead>
<tr>
<th>TSV type</th>
<th>Number of TSVs</th>
<th>Decoupling capacitance (nF)</th>
<th>Area overhead</th>
</tr>
</thead>
<tbody>
<tr>
<td>Via-first</td>
<td>2750</td>
<td>2.7 nF</td>
<td>9%</td>
</tr>
<tr>
<td>Via-last</td>
<td>76</td>
<td>0.39 nF</td>
<td>1.5%</td>
</tr>
</tbody>
</table>