國立交通大學

電子工程學系 電子研究所

碩 士 論 文

用於矽穿孔之三維積體電路完整電源供應之分析

Power Integrity in TSV 3D Integration

研 究 生：林天鴻

指導教授：黃 威 教授

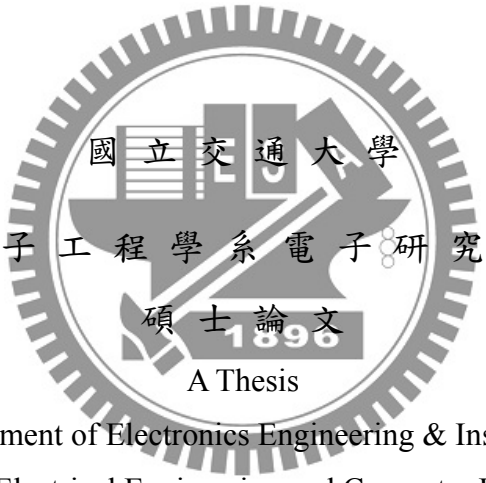
中 華 民 國 九 十 九 年 七 月

用於矽穿孔之三維積體電路完整電源供應之分析

Power Integrity in TSV 3D Integration

研 究 生：林天鴻　　Student：Tien-Hung Lin

指導教授：黃　威 教授　　Advisor：Prof. Wei Hwang



國 立 交 通 大 學

電 子 工 程 學 系 電 子 研 究 所

碩 士 論 文

A Thesis

Submitted to Department of Electronics Engineering & Institute of Electronics

College of Electrical Engineering and Computer Engineering

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Electronics Engineering

July 2010

Hsinchu, Taiwan, Republic of China

中華民國九十九年七月

用於矽穿孔之三維積體電路完整電源供應之分析

學生：林天鴻　　指導教授：黃　威　教授

國立交通大學電子工程學系電子研究所

摘　　要

三維積體電路不但非常適合異質整合，更能夠提升系統整體的速度以及降低功率消耗。然而當大量的電流同時流過封裝及矽穿孔(TSV)所造成的雜訊是非常可怕的，因此如何設計一個穩定的電源供應系統是非常重要的。在本篇論文中，我們提出了一個有效率的設計電源矽穿孔(power TSV bundle)的方法。並且根據矽穿孔之三維積體電路的特性提出了一個使用主動藕荷電容的電源雜訊抑制的電路機制。最後提出一個低靜態電流的低壓降線性穩壓器(LDO)提供可調的低雜訊電源供應電壓，並使用基底雜訊抑制的機制來減少藉由基底所傳出的雜訊。利用 UMC 65nm CMOS 以及 TSV 模型技術來實現電路設計與佈局。

Power Integrity in TSV 3D Integration

Student : Tien-Hung Lin Advisors : Prof. Wei Hwang

Department of Electronics Engineering & Institute of Electronics

National Chiao-Tung University

ABSTRACT

Three-dimensional (3D) nanosystems can provide enormous advantages in achieving multi-functional integration, improving system speed and reducing power consumption for future generations of ICs. The robust power delivery system is very important in 3D ICs. In 3D integration, the increasing supply current through both package and through-silicon-via (TSV) would lead to a large simultaneous switching noise potentially. We will introduce basic 3D technology and a proposed design method for placing the TSV bundle in 3D IC under the efficiency condition. Next, the active supply noise regulation scheme is introduced; we propose the power noise suppression technique using active decoupling capacitor (DECAPs) for TSV 3D integration characteristic. Finally, a low quiescent current linear drop regulator (LDO) is proposed to provide difference supply voltage with low noise, and substrate noise canceller is used to suppress the impact of substrate noise. All simulations are based on UMC 65nm CMOS technology and TSV model at 1V supply voltage.

致謝

感謝許多人的幫助，讓我完成了這一篇論文。

首先，我要感謝我的指導教授黃威，在他的指導下讓我對自己研究的領域有更深入的瞭解，建立了研究的興趣與自信心。黃威教授提供了一個常優良的研究環境與充足的研究資源，讓我能夠充分發揮自己的能力完成這一篇論文。

感謝實驗室的成員，在過去的兩年當中在生活上與研究上對我的幫助。感謝黃柏蒼、張銘宏、謝維致和楊浩義學長對於我在研究上的幫忙，讓我能夠有更加優良的研究成果。也要感謝陳璽文、邱議德、張雍和謝忠穎同學彼此互相的關懷。

最後我要感謝我的家人對我在生活上的關心與幫助，讓我能夠順利的完成碩士的論文研究。

Contents

List of Figures

List of Tables

Chapter 1
Introduction

1.1 Motivation

Moore's law describes a long-term trend in the history of computing hardware, in which the number of transistors that can be placed inexpensively on an integrated circuit has doubled approximately every two years. With process scaling down continuously, it will be a big problem about limitation of physical phenomenon. Therefore, 3D-ICs represent the trend of future.

Three-dimensional (3D) integration technology can provide enormous advantages in achieving multi-functional integration, improving system speed and reducing power consumption for future generations of ICs [1.1]. However, stacking multiple dies would face a severe problem of the power integrity [1.2]. Fig. 1.1 shows the power integrity for the TSV 3D integration. It is shown that heavy current density of through-silicon-via (TSVs) and packages exists in the power network and further increases the power supply noise. Moreover, the supply impedance response is dominated by both the packages and TSVs [1.3]. In view of these, noise suppression will become one of the critical design problems for TSV 3D integration.

First problem is how to decide the power TSV parameter. Generally, increasing the TSV cross section area will reduce the impedance of the power delivery network (PDN) structures and as a result mitigate the power noise. However, increase in the

dimension and density will reduce the routable area of the stacked dies.

Second, to suppress the power noise, decoupling capacitors (DECAPs) are widely used. DECAPs perform as a local reservoir of charge, which is released when the current load varies. Since the inductance of packages scales slowly, the DECAPs significantly affect the design of the power/ground (P/G) networks in high performance ICs and TSV 3D integration. At higher frequencies, DECAPs are distributed on chips to effectively manage the power supply noise. However, the usage of the on-chip passive DECAPs is limited by two major constraints, including a great amount of gate tunneling leakage and large area occupation [1.4].

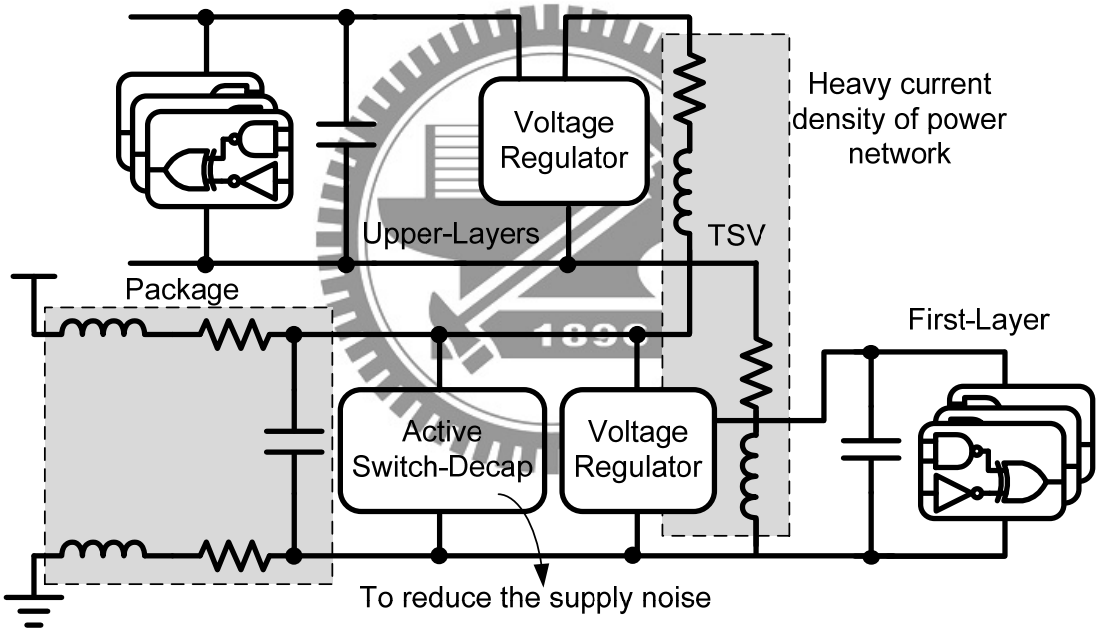


Fig. 1.1 Simplified supply impedance model in 3D IC.

The Hierarchical distributed power delivery architecture is shown as Fig. 1.2. First, a global regulator will suppress the power supply noise and deliver the supply voltage to local regulators in each layer. The number of layer is according to how many chip is stacking on. And, the local regulators will deliver power to each independent block. In other words, there is also a distributed power delivery system in each layer. Therefore, different circuit blocks have its power supply source, and it is

more efficiency to build a power management system.

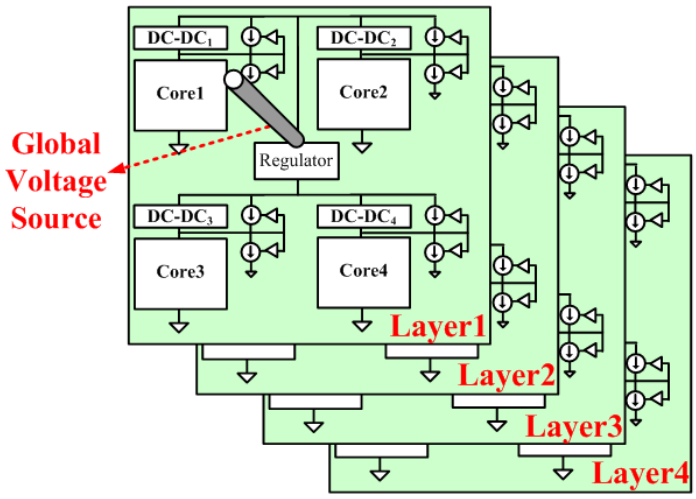


Fig. 1.2 Hierarchical distributed power delivery architecture

1.2 Research Goals and Major Contributions

The goal of this research is to design and implement a robust power delivery system for TSV 3D integration. This includes the design of power TSV placement and optimization, active DECAPs noise suppression technique, low drop regulator and power delivery system in TSV 3D integration.

The major contributions of this thesis are list as follow:

1. A robust power delivery system for TSV 3D integration is proposed. We propose design methodology for placing the power TSV bundle in TSV 3D integration under the efficiency condition.

2. A novel power noise suppression technique using active decoupling capacitors (DECAPs) is proposed based on TSV 3D integration characteristic.

3. The supply noise and substrate noise are both considered and solved in our power delivery system for TSV 3D integration. A low quiescent current low drop

regulator (LDO) is proposed to provide difference supply voltage with low supply noise, and substrate noise canceller is used to suppress the impact of substrate noise.

1.3 Organization

The rest of this thesis is organized as follows. A preview of 3D integration technology is introduced in Chapter 2. Section 2.1 will present the advantage and evaluation of 3D integration. Different kinds of fabrication of 3D integration will be introduced in section 2.2. Especially, the key technology of TSV 3D integration will also be discussed detailed in section 2.3. Although 3D ICs offer many advantages over 2D ICs, many challenges should be overcome before volume production of TSV 3D ICs, and these challenges will be presented in the final of this chapter.

The power TSV placement and optimization in TSV 3D integration is proposed in Chapter 3. The essential power grid analysis of 2D system has been invested at first. And the TSV electrical characteristic and TSV modeling will be introduced to consider the power supply noise behavior in section 3.2 and 3.3. Increasing the TSV cross section area will reduce the impedance of the PDN structures and as a result mitigate the power noise. However, increase in the dimension and density will reduce the routable area of the stacked dies. Therefore, a design method of power TSV in TSV 3D integration is proposed to derive the adequate TSV parameter in order to satisfy the efficiency condition between power noise reduction and area overhead.

Analysis supply noise regulation in TSV 3D integration is realized in Chapter 4.The optimization of passive decoupling capacitor for reducing power supply noise in 2D system will be discussed in the section 4.1. Section 4.2 describes another scheme for reducing power supply noise, which is using active circuits. In section 4.3,

a noise suppression technique using low power active decoupling capacitors (DECAPs) is proposed for TSV 3D integration. Through the latch-based noise detection circuitry, the power supply noise can be detected and regulated via active DECAPs.

Design methodology of power delivery system in TSV 3D integration is implemented in Chapter 5. Substrate noise canceller and proposed voltage regulator will be introduced in section 5.1 and 5.2. The power delivery system in TSV 3D integration is implemented UMC 65nm technology and TSV model, its simulation result and discussion is realized in section 5.3. Finally, the overall investigation results and conclusions will be discussed in Chapter 6.

Chapter 2
Preview of 3D Integration Technology

In this chapter, it introduces the preview of 3D integration technology. The motivation of IC design from 2D to 3D would be described in Section 2.1. Section 2.2 shows the general categories of 3D integration technology. In addition, the key fabrication technology of TSV 3D integration would be detailed in Section 2.3. Finally, Section 2.4 would give an introduction for the challenge of TSV 3D integration.

2.1 Why 3D?

As the semiconductor roadmap strides on, packaging and interconnection technologies are required to follow. In order to stay in pace with system demands on scaling, performance and functionality 3D integration is gaining a lot of interest as a solution to this demand [2.1]. The reasons and requirements for 3D integration are however very diverse and often application specific.

A basic reason for 3D-integration is system-size reduction. Traditional assembly technologies are based on 2D planar architectures. Die are individually packaged and interconnected on a planar interconnect substrate, mainly printed circuit boards. The area-packaging efficiency (ratio of die to package area) of individually packaged die is generally rather low (e.g. 5x5mm die in 7x7mm package: 50% area efficiency) and an additional spacing between components on the board is typically required, further

reducing the area efficiency (for example above e.g. 1mm clearance: 30% area efficiency). If we consider the volumetric packaging density, the packaging efficiency drops to very low levels. If in the previous example, we consider the active area of a die to be about 10 μm, and the combined package and board thickness to be 2 mm, the volumetric packaging density is only 0.15%. There is clearly room for improvement of the packaging density.

A different reason for looking at 3D integration is performance driven. Interconnects in a 3D assembly are potentially much shorter than in a 2D configuration, allowing for a higher operating speed and smaller power consumption. This is of particular interest for advanced computing applications. Due to the rising on-chip clock speeds, only a limited distance may be traveled by a signal in a synchronous operating mode. Using 3D-IC stacking techniques, more circuits may be packed in a single synchronous region. This requires a technology with 3D interconnects with low parasitics; in particular low capacitance and inductance are needed to avoid additional signal delay. The interconnection of circuit elements can be performed at several levels of the on-chip hierarchy. Of particular interest is the 3D stacking at the so-called “tile-level”. As shown in Fig. 2.1, typical system-on-chip, SOC, devices are constructed of a number of functional blocks. The longest on-chip lines are those that are used to interconnect these tiles. Functional ‘tiles’ on the die are rearranged in multiple die that are vertically interconnected, resulting in much shorter global interconnect These lines are typically in the top-on-chip interconnect layers and are referred to as “global” interconnects in the on-chip wiring hierarchy. Within the tiles, “local“ and “intermediate“ wiring hierarchy levels are mainly used. In a 3D approach, the large die is split in a number of smaller die, using the 3D interconnects as “global” interconnects between the tiles on both die. As this interconnect goes one

or more levels down the traditional IC-pad level, a very high 3D interconnect density is required for such an application.

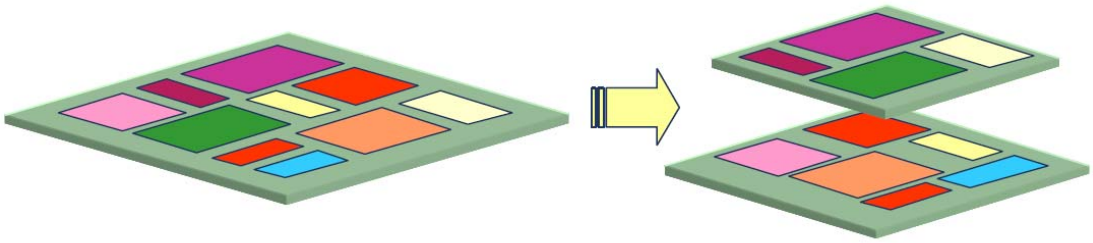


Fig. 2.1 Conceptual view a 3D stacked SOC.

A third, and maybe most important, reason to consider 3D integration is so-called hetero-integration. As silicon semiconductor technologies continue to scale (vertical scaling), the realization of true SOC devices with a large variety of functional blocks becomes very difficult to achieve. Technologies need specific optimization for logic, analog, memory etc. to reach the desired performance levels and circuit density.

Furthermore, the substrates used to build active devices may vary significantly between technologies, including non-silicon substrates, e.g. compound semiconductors. Also systems may contain other planar components, such as MEMS and integrated passive devices. Besides the 'vertical' scaling we are also experiencing a 'horizontal' scaling. Realizing the full system on a single SOC die is becoming increasingly difficult and often not economically justified. If however a high-density 3D technology is available, a "3D-SOC" device could be manufactured, consisting of a stack of heterogeneous devices. This device would be smaller, lower power and higher performance than a monolithical SOC approach. Such an approach is the obvious choice for many sensor-array applications. Many sensor applications use particular substrate materials, such as IR and X-ray sensing, that are incompatible

with Si-CMOS processing. These applications require however high-density circuits to read-out the signals from individual sensor pixels, a requirement best met with advanced CMOS technologies. The solution therefore consists in flip-chip (3D) mounting the sensor-array on a read-out electronics chip. Another possible application for this approach is the combination of logic and memory, which is shown as Fig. 2.2. The left one is 2D interconnect between logic and memory die, and the center is present (2D-SOC) combined logic and memory device, and the Right one is shown "heterogeneous 3D-SOC" stacking of a memory and logic device with 3D interconnects between individual logic tiles and memory banks.

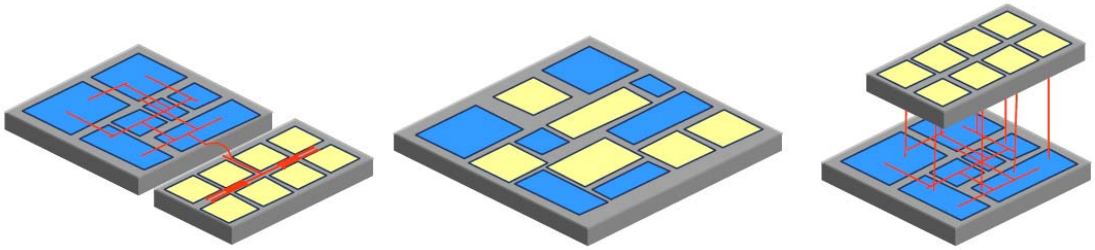


Fig. 2.2 Different approaches for combining logic and memory

Most applications require a combination of logic and memory. When large amounts of memory are needed, the memory is realized as a separate die, using a high density, optimized memory technology. Due to the use of large busses on the logic and memory die and the use of off-chip interconnects, only a relatively slow and power-hungry interconnect between memory and logic is possible. To overcome these limitations, e.g. for real-time data processing applications, a SOC approach is typically used. Although not optimal for the integration of high-density memory, the IC logic technology is used for integrating large amounts of memory. This allows for allocating smaller pieces of memory (memory-banks) to specific logic blocks. Distance between logic and memory is short, resulting in the required performance.

The integrated memory is however of the same performance as dedicated

memory technologies would offer. In particular, a much larger die area is consumed by the memory cells, resulting in a die are that is significantly larger than the case with 2 die solutions. 3D interconnect technology may solve this problem, by allowing for logic 'tiles' on a first die to directly access memory banks on a memory chip. In this case the number of 3D connections required from the memory die to the logic die will increase by an order of magnitude compared to the I/O count of standard memory devices. Similarly as for the example shown in figure 1, this approach uses 3D interconnects as "global-on chip" interconnect layers to realize a "heterogeneous 3D-SOC" structure.

A new electronics era has begun to emerge, the focus of which is on 3D ICs instead of monolithic integration of heterogeneous functions. While the impact of this approach is profound, it addresses a small part of the system. Therefore, another paradigm shift is illustrated in Fig. 2.3. 3D Systems are leading to unparalleled miniaturization, functionality and cost at system level.

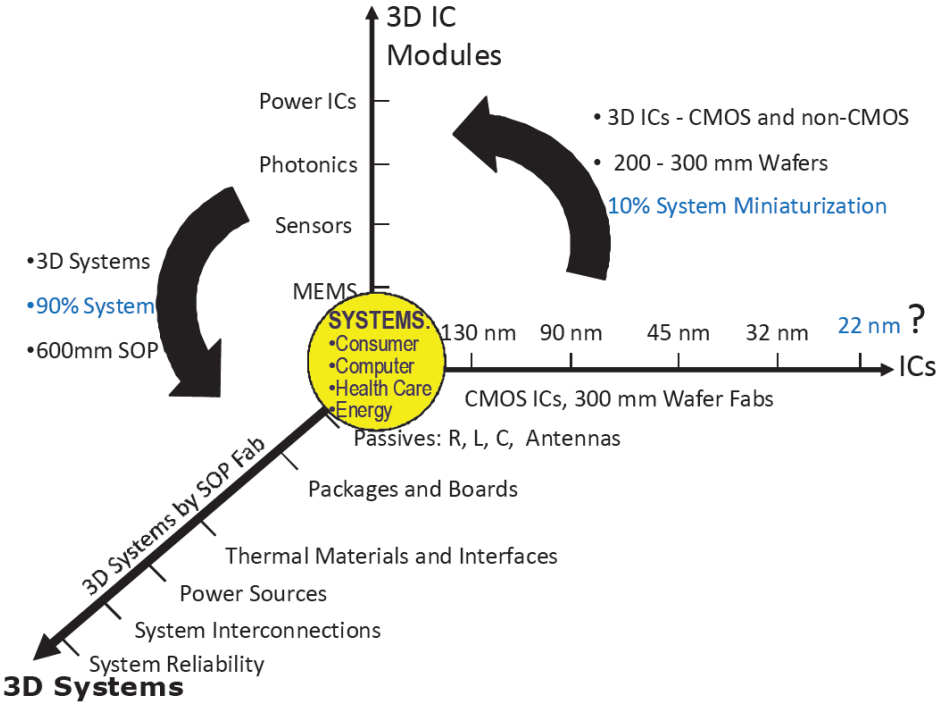


Fig. 2.3 3D Systems from ICs and 3D ICs.

To conclude, there are different motivations for the development of 3D IC solutions:

- *Form factor*: It can increase density, achieve the highest capacity and volume ratio.
- *Increased electrical performances*: Which includes shorter interconnects length and improves device speed, and it achieves better electrical insulation (to reduce electrical parasitances in RF applications).
- *Heterogeneous integration*: Integration of different functions in a 3D IC is available. (RF + memory + logic + sensor + imagers + different substrate materials + …)
- *Cost* : Cost of 3D integration may be cheaper than to keep shrinking 2D design rules following the ITRS / Moore law

2.2 Categories of 3D integration technology

3D integration is generally defined as fabrication of stacked and vertically interconnected device layers. The large spectrum of 3D integration technologies can be reasonably classified mainly in three categories [2.3][2.4][2.6][2.7][2.8][2.9]:

1. Stacking of packages and Die stacking (without TSVs)
2. TSV technology
3. Monolithic 3D

Fig. 2.4 is a representative schematic illustration of the 3D integration technologies that have been proposed to date and consists of three categories. The first category consists of 3D stacking technologies that do not utilize TSVs and are shown

in Fig. 2.4 (a)-(c). The second category consist of 3D integration technologies that require TSVs (Figure 2 (d)-(e)), and the third category consists of monolithic 3D systems that make use of semiconductor recrystallization to form active levels that are vertically stacked (with on-chip interconnects possibly between). Of course, a combination of all these technologies is possible.

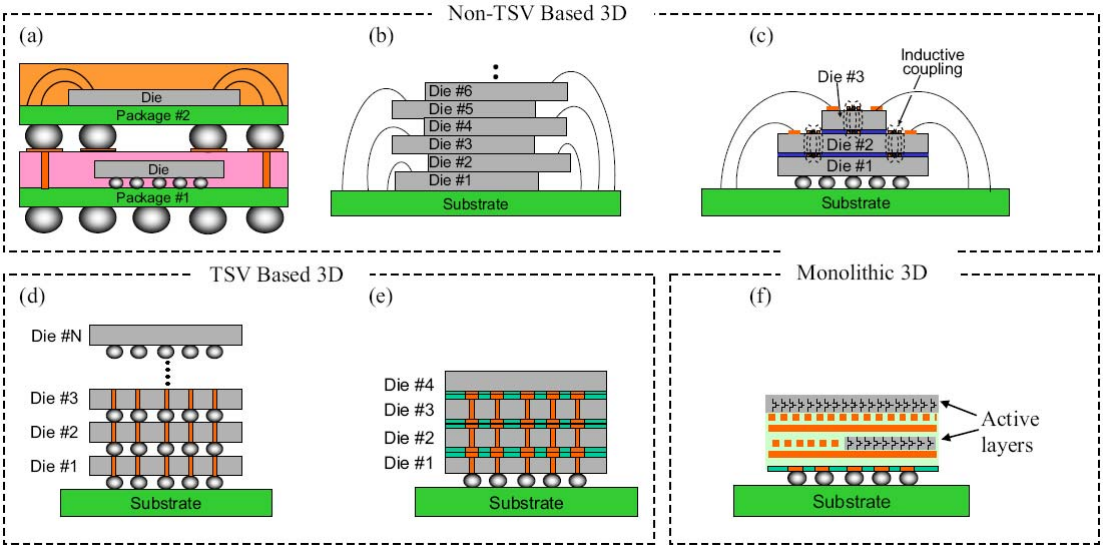


Fig. 2.4 Schematic illustration of various 3D integration technologies

A. Stacking of packages and Die stacking (without TSVs)

The non-TSV 3D systems span a wide range of different integration methodologies [2.4] and [2.5]. Fig. 2.4 (a) illustrates stacking of fully packaged dice. Although this may offer the advantages of being low cost, simplest to adopt, fastest to market, and modest form-factor reduction, the overhead in interconnect length and low-density interconnects between the two die do not enable one to fully exploit the advantages of 3D integration. Fig. 2.4 (b) illustrates the most common method to stacking memory die, which is based on the use of wire bonds. Naturally, this 3D technology is suitable for low-power and low-frequency chips due to the adverse effect of wire bond length, low density, and peripheral limited pad location for

signaling and power delivery. Fig. 2.4 (c) illustrates the use of wireless signal interconnection between different levels using inductive coupling (capacitive coupling is also possible, but more limiting). This approach is quite elegant for low-power chips that require high-data rate signaling (without the need for TSVs). Power delivery, however, requires use of wire bonds for top dice in the stack, which are not applicable for high-performance/power chips. There are several derivatives to the topologies described above, such as the die embedded in polymer approach. This approach, although different from others discussed, makes use of a redistribution layer and vias through the polymer film, and thus is a hybrid die/package level solution. It is important to note that all non-TSV approaches rely on stacking at the die/package level (die-on-wafer possible for inductive coupling and wire bond) and thus do not utilize wafer-scale bonding. This may serve to impose limits on economic gains from 3D integration due to cost of the serial assembly process.

B. TSV technology

Fig. 2.4 (d) and (e) illustrate 3D integration based on TSVs. The former figure illustrates bonding of dice with C4 bumps and TSVs. The short interconnect lengths and high density of interconnects that this approach offers are important several orders of magnitude larger number of interconnects. Although it is possible to bond at the wafer level, this approach is most suitable for die-level bonding (using a flip-chip bonder) and thus faces some of the same economic issues described above. Fig. 2.4 (e) illustrates 3D stacking based on thin-film bonding (metal-metal or dielectric-dielectric). Not only are solder bumps eliminated in this approach, but also increased interconnect density and tighter alignment accuracy can be achieved when compared to the previous approach due to the fact that these approaches are based on wafer-scale bonding. Thus, they utilize semiconductor based alignment and

manufacturing techniques.

C. Monolithic 3D

Finally, Fig. 2.4 (f) illustrates a purely semiconductor manufacturing (non-packaging) approach to 3D integration. The main enabler to this approach is the ability to deposit an amorphous semiconductor film (Si or Ge) on a wafer during the IC manufacturing process and re-crystallize to form a single-crystal film using a number of techniques. Ultimately, this approach may offer the most integrated system with least interconnects possible but may not provide chip-size areas for device fabrication in the stack.

Besides, Fig. 2.5 shows the functionality and density of those advanced packaging technology that we have mention above. We can see the TSVs can deliver the highest performance and functionality and be cost effective.

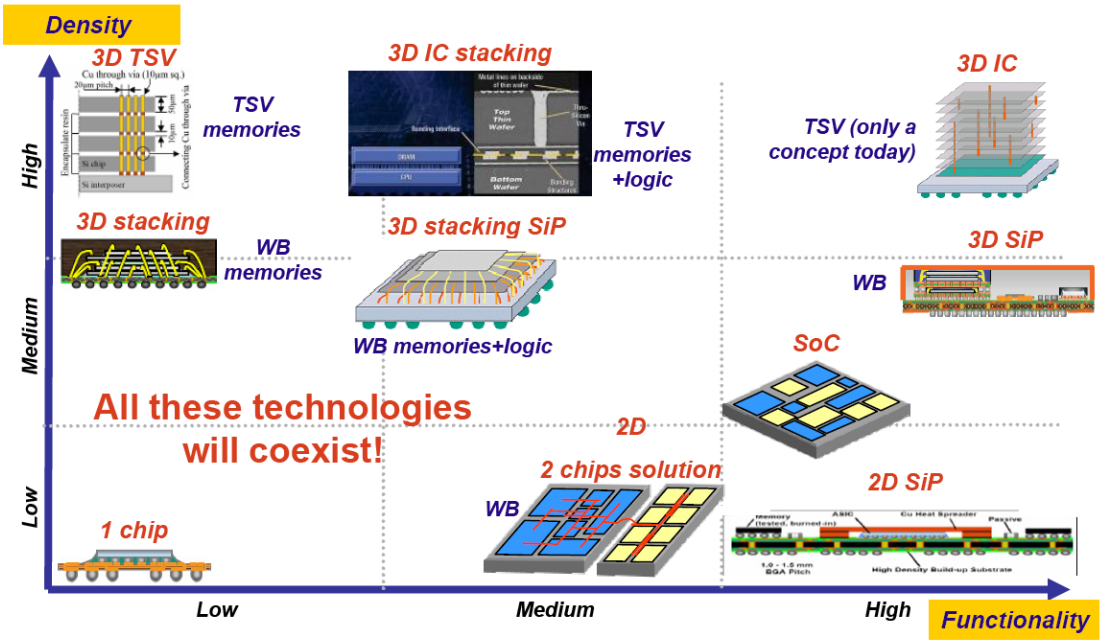


Fig. 2.5 Advanced packaging trends

It is important to note that none of the above described 3D integration

technologies address the need for cooling in a 3D stack of high performance chips. This is a significant omission and imposes a constraint on the ability to fully utilize the benefits of 3D technology. As such, new 3D integration technologies are needed for such applications.

2.3 Key Technologies and Design Choice of TSV 3D Integration

A number of 3D integration schemes or architectures are available to the system designer. The various 3D architectures for TSV 3D integration are based on the key process technologies of the following as shown as Fig. 2.6 [2.3][2.6][2.7][2.8][2.9]:

1. Stacking approach
2. Wafer (Substrate) selection
3. Bonding method
4. Direction of stacking
5. Fabrication of Thru-Si-Via (TSV)

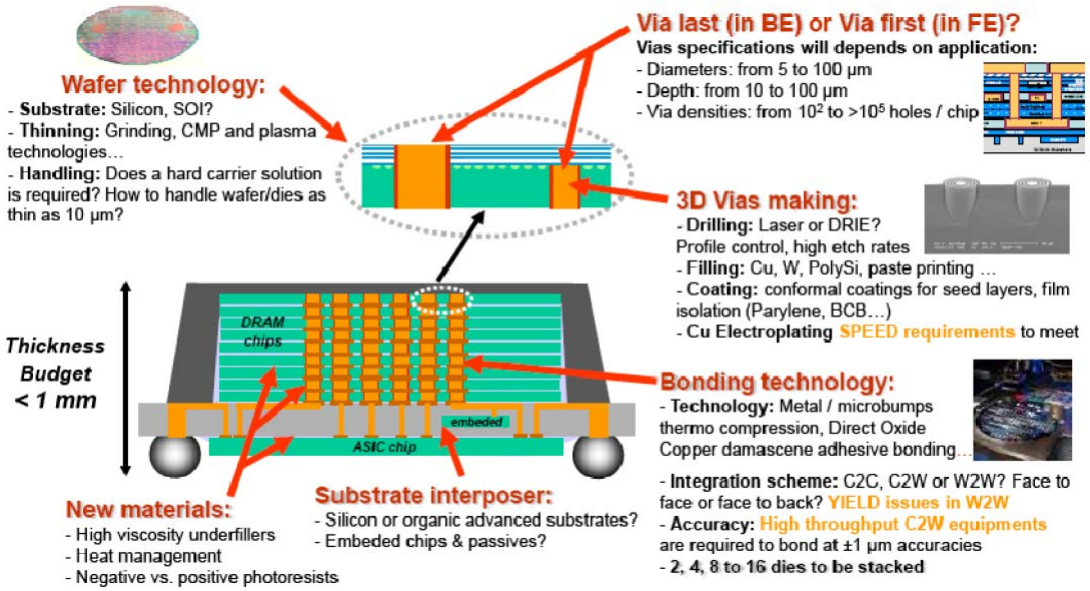


Fig. 2.6 Enable technology of TSV 3D integration

A. Stacking approach

There are three kinds of stacking approach in TSV 3D integration (see Fig. 2.7):

1. Die to Die or Chip to Chip (D2D)

2. Die to Wafer (D2W)

3. Wafer to Wafer (W2W)

If we using D2D or D2W technology, will the yields be higher than W2W. Because we can test the chip and find know good die (KGD) before bonding. Moreover, W2W technology is only suitable for common size. Although D2D or D2W is suitable for both common and dissimilar size, and it has high yields, it has two main shortcomings: handling problem and low throughput.

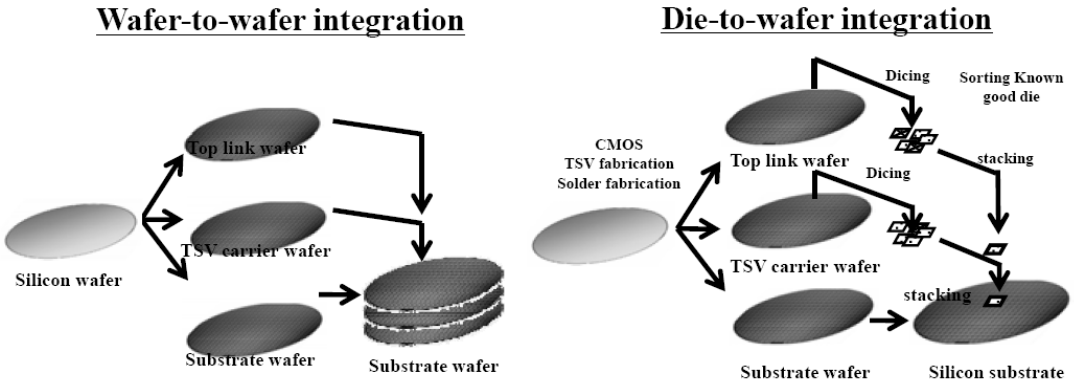


Fig. 2.7 Stacking approach

B. Wafer (Substrate) selection

There are two kinds of wafer selection has been used today. One is Bulk Si, which includes Si, Ge, or GaAs, and anther is SOI wafer which is shown in Fig. 2.8. High aspect ratio TSV are required in Buck Si wafer, the target length of TSV is equal to 50μm. And it's the most developed approach today. On the other hand, SOI simplify TSVs formation, avoid the need of a temporary carrier, and allow to stack extremely

thin layers. The BOX layer can be used as stopping layer, so the thickness of 2nd layer will be more uniform. However, it's very expensive. It seems that this approach is not cost-effective.

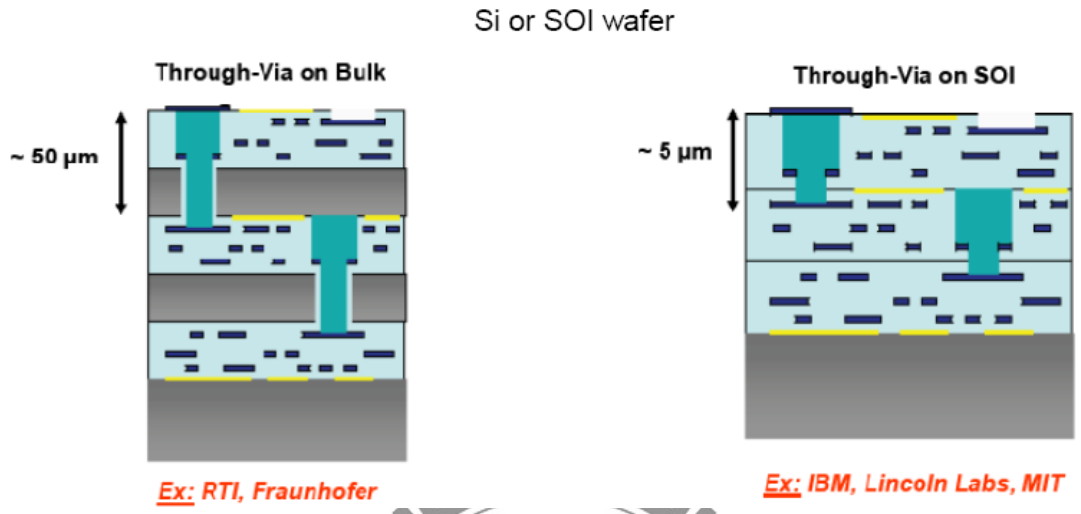


Fig. 2.8 Wafer selection

C. Bonding method

A major 3D bonding architectural choice is between dielectric bonding (Oxide-to-Oxide or Polymer-to-Polymer) and metallic bonding (Metal-to-Metal), which illustrate in Fig. 2.9. In addition to the differences in bonding materials, this choice also has a substantial impact on the details of the interstratum connections. In dielectric bonding, the interstratum connections are completed after bonding by using TSVs to pass through the top die and to connect to the conventional interconnect in the adjacent strata. In metallic bonding, the interstratum connections are completed by bonding pre-existing microconnects, and the interstratum connection may include TSVs. Another major option in the bonding of strata is the choice of wafer-to-wafer, die-to-wafer, and die-to-die bonding. Dielectric bonding typically uses wafer-to-wafer bonding, while metallic bonding is commonly associated with any of the three. Other detail characteristic is shown in Table 2.1.

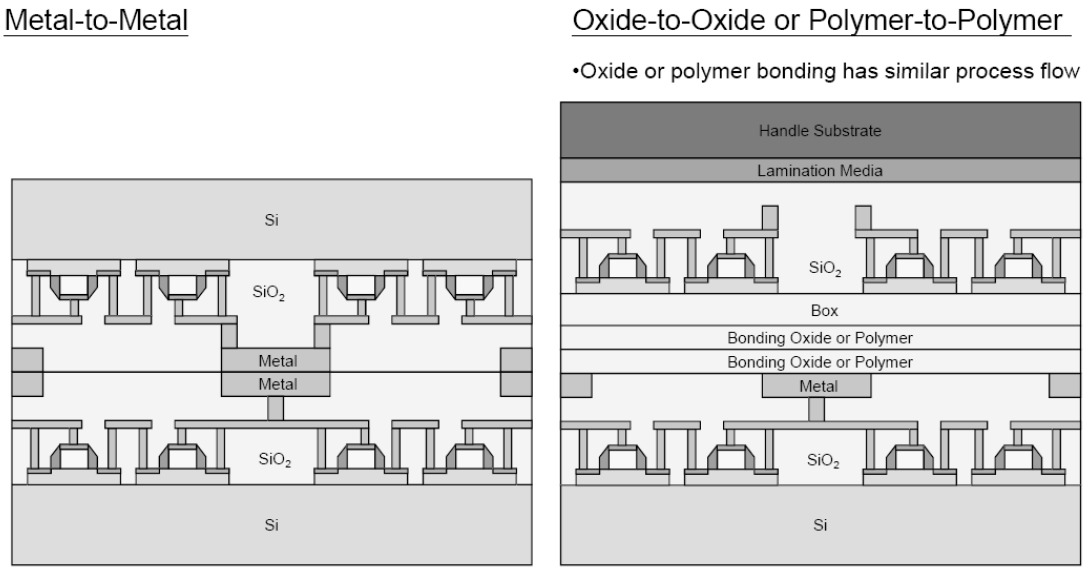


Fig. 2.9 Bonding Method

Table 2.1 Comparison of boding method

	Pro.	Con.
Metal-to-Metal	1.Metal bonding can be used as extra metal layer 2.Better heat dissipation 3.Less cleanness requirement	1. Large pitch(Misalignment) 2. How to deal the un-Cu are?
Oxide-to-Oxide	1.Possible tight pitch 2.Everywhere is oxide-bonded	1.High cleanness requirement 2.Heat dissipation
Polymer-to-Polymer	1.Possible tight pitch 2.Everywhere is polymer-boned 3.Stronger bond strength than oxide	1.Good cleanness requirement 2.Heat dissipation 3.Possible polymer contamination issue

D. Direction of stacking

Further 3D bonding architectural choices relate to the relative orientation of the dice in a 3D stack. The bonding scheme can be face-to-back, or face-to-face, where face refers to the surface on which transistors and the primary interconnect layers are formed and back refers to the Si substrate side of a die. Fig. 2.10 is a schematic illustration of a 3D chip stacking where the left one are bonded face-to- back and the right one is bonded face-to- face.

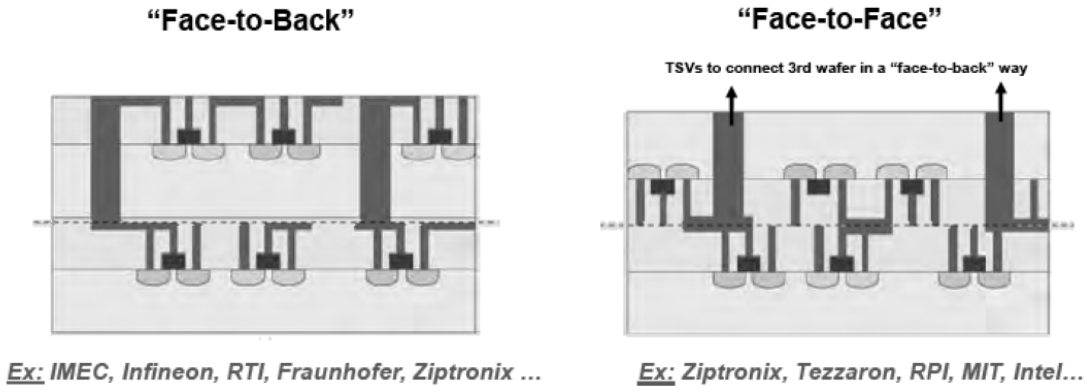


Fig. 2.10 Direction of stacking

E. Fabrication of Thru-Si-Via (TSV)

According to Fig. 2.11, Fabrication of TSV can be separate into via first and via last. It depends on the via fabrication step before or after the BEOL process. Via-first approach is challenging by CMOS process. There will be issues with the subsequent CMOS steps at different temperature ranges, so the materials must be CMOS compatible. But it has no yield issue, only good wafers are used. And it has lower cost than via-last. Via-last will not being thermal stress issues, but where the vias etching must be carefully done. However, the yield of the TSV process affects the full process, and it will lower the total yield.

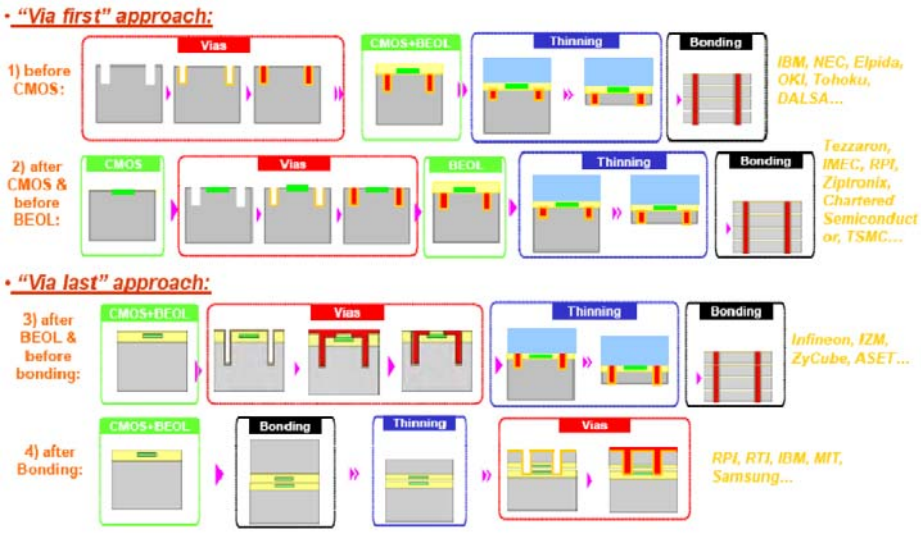


Fig. 2.11 Fabrication of Thru-Si-Via (TSV)

2.4 Challenge of TSV 3D integration

Although 3D ICs offer many advantages over 2D ICs, many challenges should be overcome before volume production of TSV-based 3D ICs becomes possible. These challenges include technological challenges, yield and test challenges, thermal challenges, infrastructure challenges [2.12], etc.

1. Thermal issue—Although the power consumption of a die within a 3D IC is expected to decrease due to the shorter interconnects, the heat removing of a 3D IC is much more difficult than that of a 2D IC. The cause is that the ambient environment of the die of a 2D IC is the cooling material, but the ambient environment of a die within a 3D IC may be another die which also generates heat. Therefore, the thermal issue of a 3D IC is much severer than that of a 2D IC.

2. Yield issue—3D integration technology may benefit the yield of 3D ICs but may deteriorate the yield of 3D ICs on the other hand. For W2W bonding technology, the yield of a 3D IC is the product of the yields of multiple die and the yield of stacking process. When combining n untested die from wafers with a die yield Y_i, then the compound yield of the 3D structure Y_m can be expressed as $Y_m = Y_s^{n-1} \times Y_i^n$, where Y_s is the yield of the stacking process . Apparently, the yield of a 3D IC is dramatically reduced. But, for D2W and D2D bonding technologies the yield of 3D ICs can be remained at high level if the known-good die (KGD) is done. On the other hand, 3D integration technology inherently increases the yield since heterogeneous structures can be fabricated on separate wafers using individually optimized fabrication process and materials. This is impossible for integrating these heterogeneous structures in a 2D IC.

3. Test issue—For achieving a high yield of 3D ICs, high-quality KGD must be done.

Wafer-level KGD for 3D ICs is more difficult than existing KGD approaches for system-in-package (SiP). The cause is that the die for SiP has I/O pads but the die for 3D IC may only has TSVs. The pitch between TSVs is much smaller than that of I/O pads. In wafer-level testing, the probe of TSVs becomes a challenge. On the other hand, the TSV of each die before bonding is a partial circuit and these results in that the pre-bond testing is also a challenge. Furthermore, the test optimization and integration for pre-bond and post-pond testing are also an important issue.

4. Technological issue—As aforementioned, different bonding technologies have different impact on the final yield of 3D ICs. In addition, each step of the overall 3D integration process has heavy impact on the final yield. For example, the alignment accuracy, wafer thinning, TSV formation, and so on. All of these should be investigated and developed further before the 3D integration technology is mature enough for high-volume production.

5. Infrastructure issue—Although many 3D integration technologies have been investigated and demonstrated, an effective design flow for 3D ICs has not be developed. Computer-aided design (CAD) algorithms and tools for 3D ICs thus are required. For example, floorplanning, placement, and routing tools for 3D ICs must be developed.

2.5 Thermal and Power Delivery Challenge in TSV 3D integration

One of the primary advantages of 3D chips stems from their ability to pack circuitry more densely than in 2D. However, this increased level of integration also results in side effects in the form of new limitations and challenges to the designer as we

mentioned in section 2.4. Thermal and power delivery problems, which are most serious, can both be traced to the fact that a *k*-tier 3D chip could use *k* times as much current as a single 2D chip of the same footprint while using substantially similar packaging technology. The implications of this are as follows:

- First, the 3D chip generates *k* times the power of the 2D chip, which implies that the corresponding heat generated must be sent out to the environment. If the design technique is thermally unaware and the package thermal characteristics for 2D and 3D circuits are similar, this implies that on-chip temperatures on 3D chips will be higher than for 2D chips. Elevated temperatures can hurt performance and reliability, in addition to introducing variabilities in the performance of the chip. Therefore, on-chip thermal management is a critical issue in 3D design.
- Second, the package must be capable of supplying *k* times the current through the power supply (*Vdd* and ground) pins as compared to the 2D chip. Moreover, the power delivery problem is worsened in 3D ICs as through-silicon vias (TSVs) contribute additional resistance to the supply network. Given that reliable power grid design is a major bottleneck even for 2D designs, this implies that significant resources have to be invested in building a bulletproof power grid for the 3D chip.

2.5.1 Thermal Issues in 3D ICs

2.5.1.1 The Thermal PDE

Full-chip thermal analysis involves the application of classical heat transfer theory. The differences lie in incorporating issues that are specific to the on-chip context: for

example, on-chip geometries are strongly rectilinear in nature and involve rectangular geometric symmetries; the major sources of heat, the devices, lie in either a single layer in each 3D tier, and the points at which a user is typically interested in analyzing temperature are within the device layer(s).

Conventional heat transfer in a chip is described by Fourier's law of conduction [2.13], which states that the heat flux, q (inW/m2), is proportional to the negative gradient of the temperature, T (in K), with the constant of proportionality corresponding to the thermal conductivity of the material, k_t (in W/(m K)), i.e.,

$$q = -k_t \nabla T \tag{2.1}$$

The divergence of q in a region is the difference between the power generated and the time rate of change of heat energy in the region. In other words,

$$\nabla q = -k_t \nabla \cdot \nabla T = -k_t \nabla^2 T = g(r,t) - \rho c_p \frac{\partial T(r,t)}{\partial t} \tag{2.2}$$

Here, r is the spatial coordinate of the point at which the temperature is being determined, t represents time (in s), g is the power density per unit volume (in W/m3), c_p is the heat capacity of the chip material (in J/(kg K)), ρ is the density of the material (in kg/m3). This may be rewritten as the following heat equation, which is a parabolic PDE:

$$\rho c_p \frac{\partial T(r,t)}{\partial t} = k_t \nabla^2 T(r,t) + g(r,t) \tag{2.3}$$

The thermal conductivity, k_t, in a uniform medium is isotropic, and thermal conductivities for silicon, silicon dioxide, and metals such as aluminum and copper are fundamental material properties whose values can be determined from standard tables. The solution to Equation (2.3) corresponds to the transient thermal response. In the steady state, all derivatives with respect to time go to 0, and therefore, steady-state

analysis corresponds to solving the PDE:

$$T(r,t) = -\frac{g(r,t)}{k_t \nabla^2} \tag{2.4}$$

This is the well-known Poisson's equation.

The time constants of heat transfer are of the order of milliseconds and are much longer than the subnanosecond clock periods in today's VLSI circuits. Therefore, if a circuit remains within the same power mode for an extended period of time and its power density distribution remains relatively constant, steady-state analysis can capture the thermal behavior of the circuit accurately. Even if this is not the case, steady-state analysis can be particularly useful for early and more approximate analysis, in the same spirit that steady-state analysis is used to analyze power grid networks early in the design cycle. On the other hand, when greater levels of detail about the inputs are available or when a circuit makes a number of changes between power modes at time intervals above the thermal time constant, transient analysis is possible and potentially useful.

To obtain a well-defined solution to Equation (2.3), a set of boundary conditions must be imposed. Typically, at the chip level, this involves building a package macromodel and assuming that this macromodel interacts with a constant ambient temperature.

2.5.1.2 Steady-State Thermal Analysis Algorithms

Next, we will describe steady-state analysis techniques based on the application of the finite difference method (FDM) and the finite element method (FEM). Both methods discretize the entire chip and form a system of linear equations relating the temperature distribution within the chip to the power density distribution. The major

difference between the FDM and FEM is that while the FDM discretizes the differential operator, the FEM discretizes the temperature field. The primary advantage of the FDM and FEM is their capability of handling complicated material structures, particularly nonuniform interconnect distributions in a VLSI chip.

The FEM and FDM methods both lead to problem formulations that require the solution of large systems of linear equations. The matrices that describe these equations are typically sparse (more so for the FDM than the FEM, as can be seen from the individual element stamps) and positive definite. There are many different ways of solving these equations. Direct methods typically use variants of Gaussian elimination, such as LU factorization, to first factor the matrices and then solve the system through forward and backward substitutions. The cost of LU factorization is $O(n^3)$ for a dense $n \times n$ matrix but is just slightly superlinear in practice for sparse systems. This step is followed by forward/backward substitution, which can be performed in O(n) time for a sparse system where the number of entries per row is bounded by a constant. If a system is to be evaluated for a large number of righthand side vectors, corresponding to different power vectors, LU factorization only needs to be performed once, and its cost may be amortized over the solution for multiple input vectors.

Iterative methods are seen to be very effective for large sparse positive definite matrices. This class of techniques includes more classical methods such as Gauss–Jacobi, Gauss–Seidel, and successive overrelaxation, as well as more contemporary approaches based on the conjugate gradient method or GMRES. The idea here is to begin with an initial guess solution and to successively refine it to achieve convergence. Under certain circumstances, it is possible to guarantee this convergence: in particular, FDM matrices have a structure that guarantees this property. For FDM

methods, the similarity with the power grid analysis problem invites the use of similar solution techniques, including random walk methods [2.14] and other methods such as multigrid approaches [2.15].

2.5.1.3 Thermal Optimization of 3D Circuit

The illustration in Fig. 2.12 shows a simple thermal model for a 3D circuit and outlines techniques for overcoming thermal challenges in these structures. Fig. 2.12 shows a schematic of a 3D chip sitting atop a heat sink: this is modeled using a distributed power source feeding a distributed resistive network connected to a thermal resistance that models the heat sink. Although this is a coarse model, it suffices for illustrative purposes. By the thermal–electrical analogy, the voltage in this network represents the temperature on the chip. The temperature can therefore be reduced using the following schemes:

- Through low-power design: By reducing the power dissipation of the chip, the thermal current injected into the network can be reduced, controlling the IR drop, and therefore, the voltage.

- By rearranging the heat sources: The locations of the heat sources can be altered through physical design (floorplanning and placement) to obtain improved temperatures. Coarsely speaking, this implies that high-power modules should be moved away from each other and closer to the heat sink.

- By improving thermal conduits: The temperature may also be reduced by improving the effective thermal conductivity of paths from the devices to the heat sink. An effective method for achieving this is through the insertion of thermal vias: thermal vias are structurally similar to electrical vias but serve no

electrical purpose. Their primary function is to conduct heat through the 3D structure and convey it to the heat sink.

- By improving the heat sink: An improved heat sink results in an improvement in the value of R_{sink}, which can help reduce the temperature.

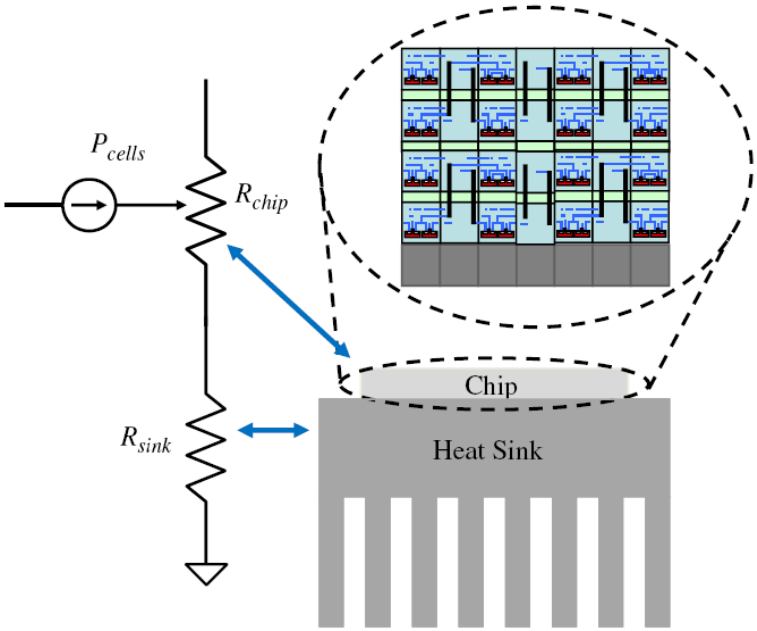


Fig. 2.12 A simple thermal model for a 3D chip

2.5.2 Power Delivery in 3D ICs

Despite the recent surge in 3D IC research, there has been very little work from the circuit design and automation community on power delivery issues for 3D ICs. On-chip power supply noise has worsened in modern systems because scaling of the power supply network (PSN) impedance has not kept up with the increase in device density and operating current due to the limited wire resources and constant RC per wire length, and as stated earlier, this situation is worsened in 3D ICs. The increased IR and Ldi/dt supply noise in 3D chips may cause a larger variation in operating speed leading to more timing violations. The supply noise overshoot due to inductive

parasitics may aggravate reliability issues such as oxide breakdown, hot carrier injection (HCI), and negative bias temperature instability (NBTI) (which are also negatively affected by elevated temperatures). Consequently, on-chip power delivery will be a critical challenge for 3D ICs.[2.10]

2.5.2.1 The Basics of Power Delivery

According to scaling roadmaps, future high-performance ICs will need multiple, sub-1V supply voltages, with total currents exceeding 100 A/cm2 even for 2D chips [2.16]. Conventional power delivery methods for high-performance ICs employ a DC– DC converter known as a voltage regulator module (VRM). The VRM is typically mounted on the motherboard, with external interconnects providing the power to the chip, as depicted in Fig. 2.13[2.17]. The intrachip power delivery network is shown in Fig. 2.13b, which shows a part of the modeled PSN of a microprocessor [2.18]. The package parasitics, contributed by the I/O pads and bonding wires, are modeled as an inductance and resistance in series. The decoupling capacitors (DECAPs) shown in the figure are intended to damp out transient noise and include the external decap as well as the capacitance due to the various circuit components such as the MOS gate capacitance.

The chip acts as a distributed noise source drawing current in different locations and at different frequencies, causing imperfections in the delivered supply. The supply that reaches the processor is affected by IR and Ldi/dt drop across the package constituting the supply noise: the package impedance has largely remained unaffected by technology scaling. Scaling does, however, result in some unwanted effects on-chip, namely, increased currents and faster transients from one technology node to the next. The former aggravate the IR drop, while the latter worsen the Ldi/dt

drop [2.10]. Over and above these effects is the issue of global resonant noise in which the supply impedance gets excited to produce large drops on supply at or near the resonant frequency. With these increased levels of noise and reduced noise margins, as Vdd levels scale down, reliable power delivery to power-hungry chips has become a major challenge.

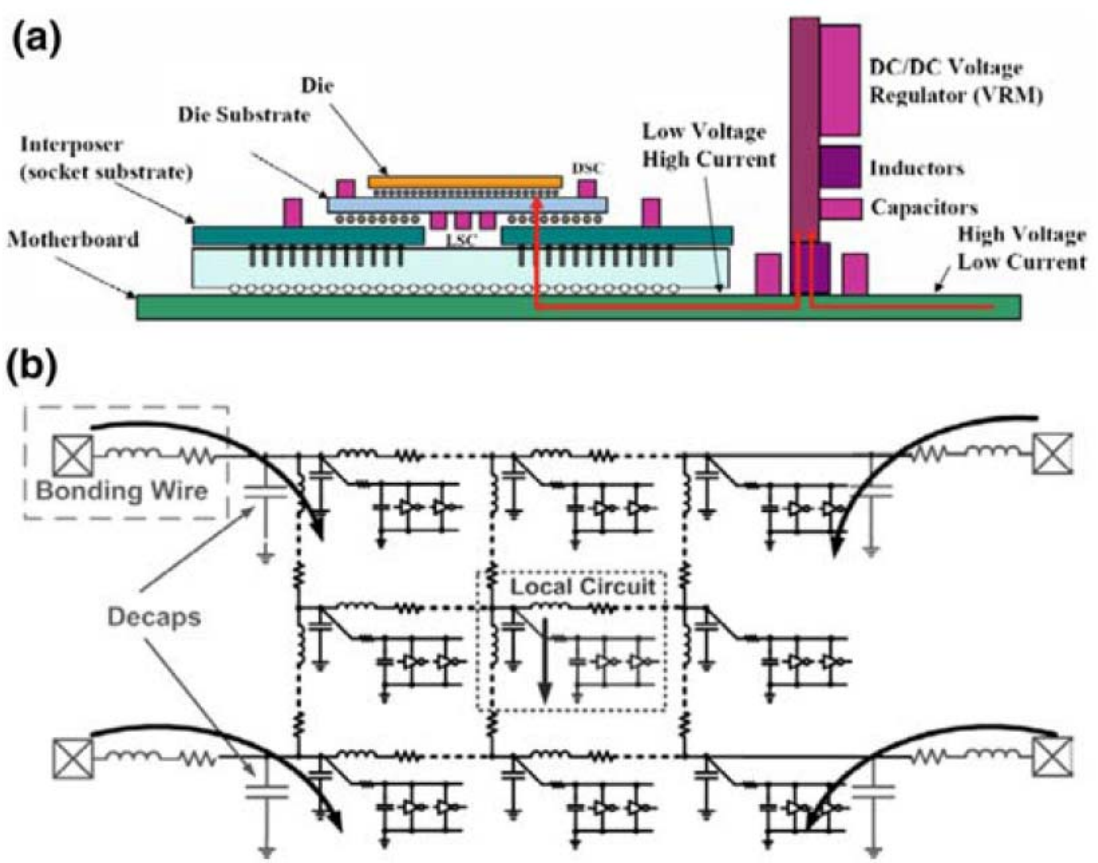


Fig. 2.13 (**a**) Conventional power delivery architecture (**b**) On-chip power grid

The noise spectrum for a typical power grid is shown in Fig. 2.14a. The DC mcomponent of the noise is given by IR drop across the package and power grid. The first peak in the figure corresponds to the resonant frequency, given by $f\text{res} = 1/(2\pi\sqrt{LC})$, which typically appears in the range of 100–300 MHz. An excitation at this frequency can be triggered during microprocessor loop operations or wakeup. Several other peaks are seen in the figure due to switching at clock frequency and its higher harmonics or due to local resonance: the corresponding noise is

typically an order less in magnitude than the resonant peak. Fig. 2.14b shows a measured supply impedance profile of a separate test structure, which validates the simulation model developed in Fig. 2.13. The noise at a particular frequency is estimated by multiplying the impedance with the current component at that frequency [2.19].

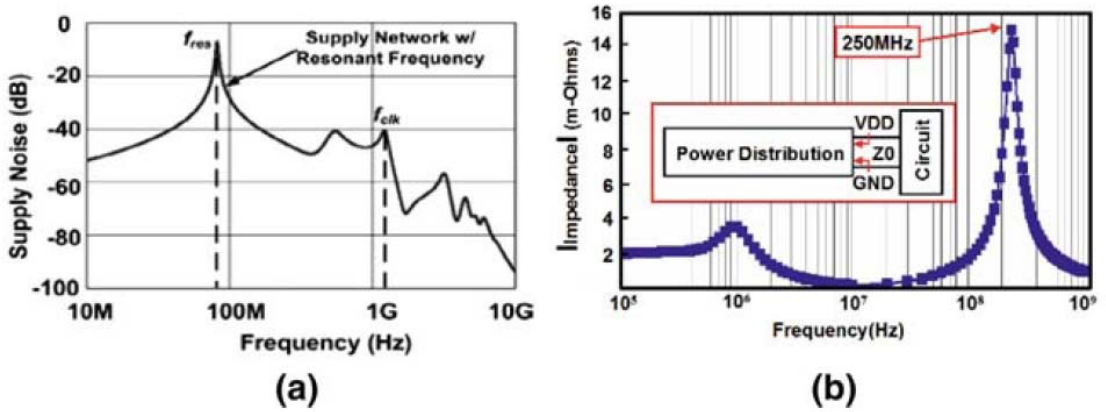


Fig. 2.14 (**a**) Simulation of the supply noise spectrum. (**b**) Measurement results for supply noise

2.5.2.2 Three-Dimensional IC Power Delivery: Modeling and Challenges

A model for 3D ICs, based on distributed models of the on-chip and package power supply structures, is shown in Fig. 2.15[2.20]. Power is fed from the package through power I/O bumps distributed over the bottom-most tier and travels to the upper tiers using TSVs. The footprint of the chip can be divided into cells, which are identical square regions between a pair of adjacent power and ground pads, as shown in Fig. 2.15a. The cells are connected in Fig. 2.15b in the form of a grid formed by several subcells between adjacent TSVs. Electrically, each TSV is modeled as a series combination of resistance and inductance. The planar square cells use a lumped model,

where R_{si}, J_i, and C_{di} represent, respectively, the grid resistance, effective current density, and chip decap on a per-unit basis. Since each pad is shared by four independent cells, the package parameters are normalized by a factor of four. The subcell can be then repeated multiple times to realize the complete 3D IC functional block.

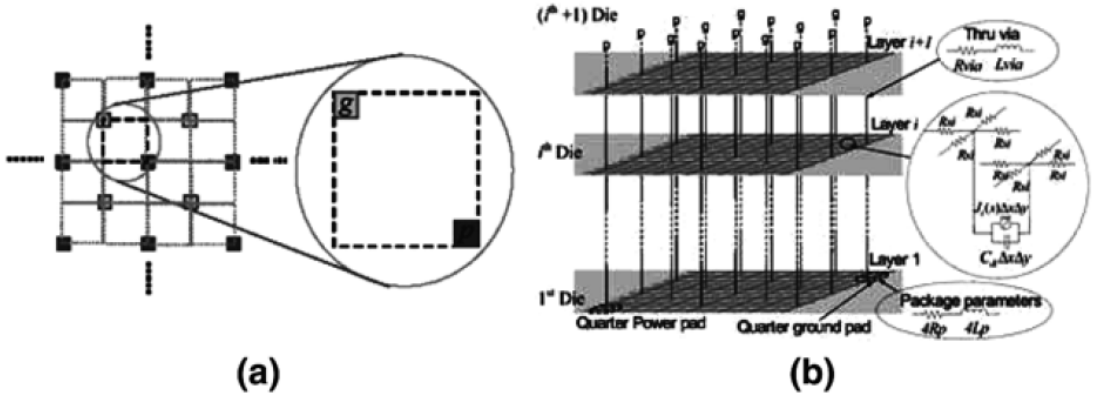


Fig. 2.15 Distributed model for 3D IC. (**a**) Division of the power grid into independent cells. (**b**) A model for one such cell.

The power grid model must necessarily be tied to a real 3D process. Fig. 2.16a depicts a 3D IC cross-sectional model of a production level 0.18μm 3D process from MIT Lincoln Laboratory [2.20]. This process has three tiers. The bonding pads are on the top tier, while the heat sink is typically below the bottom tier. Processors or other power intensive circuits would ideally be placed on the bottom tier in close proximity with the heat sink. The tiers are interconnected through TSVs for electrical and thermal conduction. Fig. 2.16b shows the cross-sectional scanning electron microscope (SEM) photograph of a stacked TSV connecting the back metal of the top tier with the top level metal of the bottom tier. A simplified resistance model is superimposed. Based on actual parameter extraction [2.20], each stacked cone-shaped TSV has a resistance of 1Ω in this process. The top and middle tiers are aligned face-to-back, while the middle and bottom tiers one face-to-face, making the path

from the top to middle tier longer and more resistive. This configuration can be modeled by breaking up the total 1Ω-stacked via resistance into chunks of 0.25Ω, 0.5Ω, and 0.2Ω, as shown in Fig. 2.16b. The values of TSV inductance and capacitance can be ignored as their values, found experimentally, are fairly small.

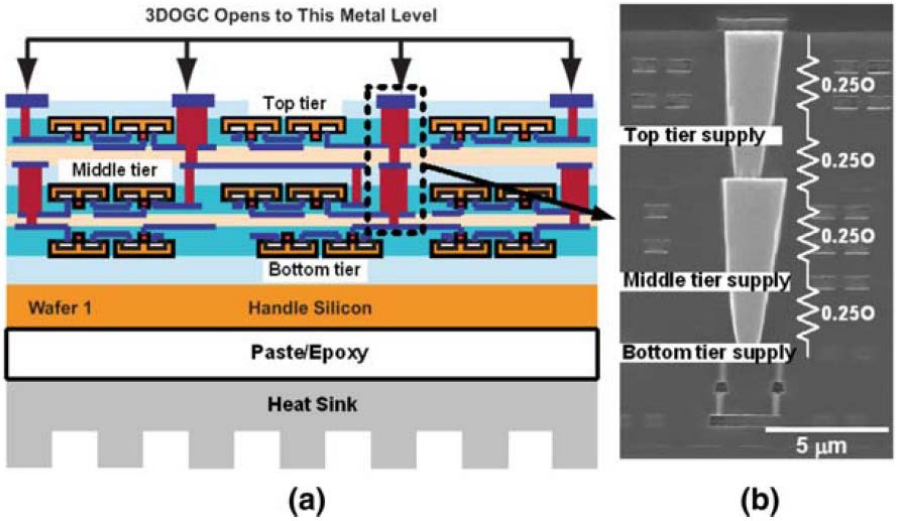


Fig. 2.16 (**a**) Cross section of 3D FD-SOI process. (**b**) Simplified via resistance model aligned with a cross-sectional SEM photograph.

The TSV resistance in the supply path potentially imposes new challenges in 3D power delivery vis-à-vis the conventional 2D case [2.22]. First, the lower tiers experience worsened PSN noise due to the increased resistance in the PSN. Moreover, power intensive circuits have to be placed at the bottom tier, which makes reliable power delivery further difficult.

In 3D, there are two significant points of departure, in comparison with models for conventional 2D chips. First, for the same circuitry, the reduced footprint of the 3D die effectively increases the package parasitics: since ratios of the number of supply pins and bonding wires to the supply current are reduced, the role of the package resistance and inductance is increased. Second, the noise characteristics in each tier are affected by the additional TSV resistance in the supply path.

Fig. 2.17 shows the circuit models developed to compare the 3D and 2D cases. The models are based on curve fits with the impedance profile of a distributed supply network model, along with typical decap and package parasitic values. In 3D, we see that the supply path would be dominated by the TSVs. The overall chip capacitance (3nF in the 2D case) within an equal footprint is assumed to be split equally in the 3D IC between its three tiers. Moreover, due to the reduced footprint of the 3D die, the number of power pins is assumed to be a third of the 2D case, leading to 3× increase in package parasitic inductance and resistance values.

Since the noise at the bottom tier is predictably worst, we compare the impedance response of this tier with the 2D case. The normalized impedance comparison is shown in Fig. 2.17b, which illustrates the following:

- *Low-frequency impedance*: At low frequencies, the capacitors and inductors are open and short circuited, respectively. Therefore, the 2D model has an impedance of 2(0.01+0.03)=0.08Ω, while the 3D model has an impedance of 2(0.03+0.05+0.1+0.05)=0.46Ω. This indicates that for the same amount of current, the 3D chip will have 5.75× more IR drop compared to 2D.

- *Resonant peak impedance*: The resonant peak is determined by the amount of damping and the value of inductance. Here, the increased role of inductance in 3D is counteracted by the increased damping provided by the larger resistance drop to the bottom tier, and the peaks show comparable values.

- *Resonant frequencies*: Two-dimensional circuits typically have a resonant frequency of around 50–300 MHz, given by $f_{res} = 1/(2\pi\sqrt{LC})$. If the equivalent capacitance in 3D is same as in our model, due to the increased L, the peak is shifted to a lower frequency as seen in Fig. 2.17c.

- *High-frequency impedance*: At high frequencies, 2D and 3D impedances become comparable, and this is attributed to the shielding effect of the bottom tier capacitance, due to the fact that the capacitance becomes virtually a short circuit at high frequencies.

Clearly, it can be seen that DC supply noise becomes a greater concern in 3D designs as compared to its 2D counterpart.

To understand the supply noise behavior in different tiers, we analyze the impedance spectrum (see Fig. 2.17c) across different tiers obtained by simulating the 3D IC model. The key results are as follows:

- *Low-frequency impedance*: As expected, the DC- and low-frequency impedances, which are governed by the TSV resistances, show a worsening trend for the lower level tiers.
- *High-frequency impedance*: At high frequencies, the top tier has the largest impedance, while the middle tier has the minimum AC impedance. Although this seems to be counter intuitive, it can be explained by the shielding/decap effect of the adjacent tier capacitances, which causes the effective damping resistances to be the largest for the middle tier and smallest for the top tier. The above trend is more noticeable at high frequencies beyond the resonance peak.
- *Resonant behavior*: Since the shielding effect mentioned above is not significant at mid-frequencies, the resonance peak follows the low-frequency trend, with the bottom tier being the worst case. However, there is a reduced noise offset as noted

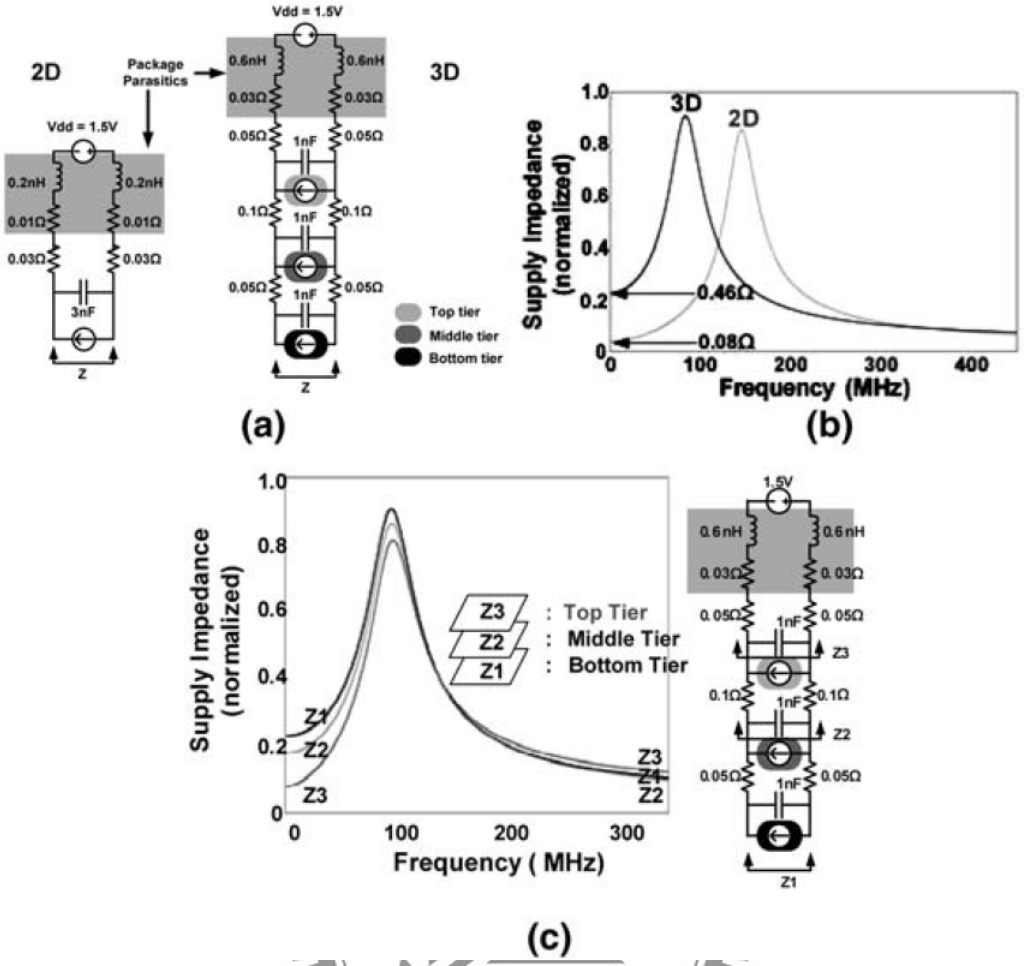


Fig. 2.17 (**a**) Simplified PSN models for comparing impedance response in 2D and 3D. (**b**) Impedance response comparison between 2D and 3D. (**c**) Impedance response of the three tiers in a 3D IC

In summary, the AC impedance is worst for the bottom tier until the resonant frequency, while beyond this point, the top tier has a slightly larger impedance value. Since thermal constraints dictate that the bottom tier is likely to contain circuit blocks with large current consumption, the supply noise in the bottom tier (i.e., the product of current and impedance) will become a significant concern for 3D implementations.

The aim of the above discussion was to provide some quantitative understanding of power delivery in 3D ICs. It should be pointed out that these numbers are tied to a specific process and will change depending on the process. For example, if the technology allows TSVs with much lower resistance or area, then the impedance

bottleneck in a path may be due to the supply pads, and the PSN models should account for that. However, regardless of this, it remains likely that PSN will be a key problem in 3D designs.

2.5.2.3 Design Techniques for Controlling Power Supply Network Noise

The presence of severe power delivery bottlenecks necessitates a look at entirely novel power delivery schemes for 3D chips. We introduce several possible approaches for this purpose.

A. On-Chip Voltage Regulation

One way of dealing with the power delivery problem in 3D ICs (and also in conventional 2D ICs) is to bring the DC–DC converter module closer to the processor, conceptually shown in Fig. 2.18 [2.23]. Boosting the external voltage and locally downconverting it ensures that the current through external package, I_{ext}, is small, and relaxes the scaling requirement on external package impedance. Moreover, this point of load (PoL) regulation isolates the load from global resonant noise from external package and decap. Traditionally, the efficiency of monolithic DC–DC converters has been limited by the small physical inductors allowed on-chip. Typical off-chip DC–DC conversion requires high-Q inductors of the order of 1-100 μH, which are difficult to implement on-chip due to their area requirements. With growing power delivery problems, the focus has been on building compact inductors through technologies like thin film inductors or on more efficient, but costly, DC–DC converters through multiphase/interleaving topologies. Clearly, there is an onus to incorporate these on-chip, which calls for a different process altogether. The

possibility to stack different wafers with heterogeneous technologies, as offered by three-dimensional wafer-level stacking in 3D ICs, is thus the natural solution for realizing on-chip switching converters.

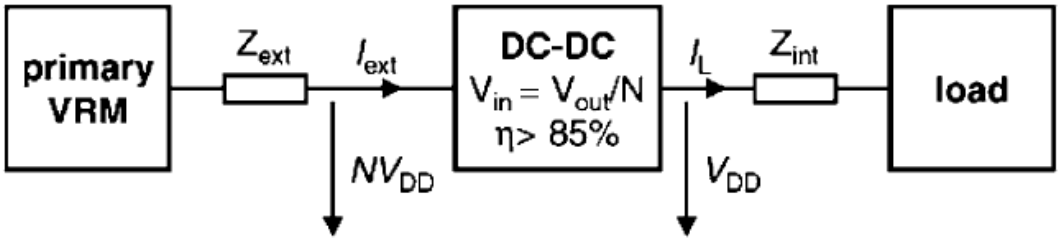


Fig. 2.18 Insertion of a DC–DC converter near the load.

B. Z-axis Power Delivery

Z-axis or 3D power delivery, in which the PSN is vertically integrated with the processor in a 3D stack, promises an attractive solution for on-chip DC–DC conversion. Fig. 2.19 shows the schematic visualization [2.17] of such a *Z*-axis power delivery technique using wafer–wafer integration. This still requires that all passives, including the inductors and output capacitors, must be monolithically integrated with the power switches and control circuitry. The idea is gaining traction in research, and implementation of such a structure, using two interleaved buck converter cells each operating at 200MHz switching frequency and delivering 500mA output current has been reported [2.17]. In the future, we may see a 3D IC with several tiers, with one whole tier dedicated to voltage regulation, incorporating various passives and other circuitry.

One main issue with *Z*-axis power delivery is the area overhead in dedicating a tier to an on-chip DC–DC converter, whose footprint should be at par with the processor in a wafer–wafer 3D process. Moreover, high-efficiency switching regulators for DC–DC conversion require monolithic realization of bulky passive

components. On the other hand, typical linear regulators, though less bulky, suffer from efficiency loss.

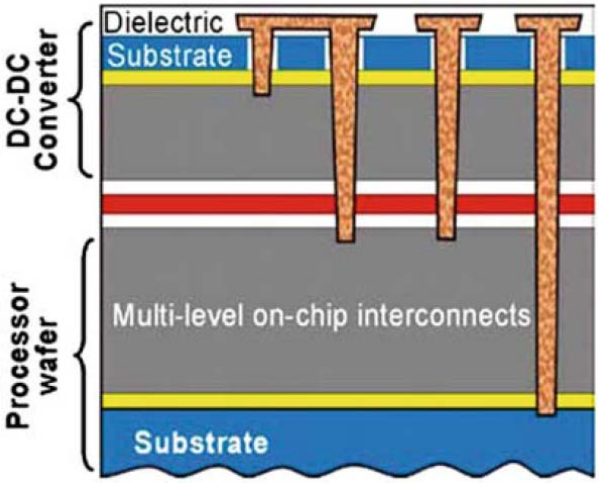


Fig. 2.19 *Z*-axis power delivery based on monolithic power conversion and wafer–wafer bonding .

C. Multistory Power Delivery

A promising technique for achieving high-efficiency on-chip DC–DC conversion and supply noise reduction is the multistory power delivery (MSPD) scheme [2.22]. It has been demonstrated in [1.3] that the idea becomes particularly attractive for 3D IC structures involving stacked processors and memories. Fig. 2.20 demonstrates the basic concept of MSPD. A schematic of a conventional supply network is shown in Fig. 2.20a, where all circuits draw current from a single power source. Fig. 2.20b shows the multistory supply network, with subcircuits operating between two supply stories. The concept of a "story" is merely an abstraction to illustrate the nature of the power delivery scheme, as opposed to the 3D IC architecture, where circuits are physically stacked in tiers. In this scheme, current consumed in the "2V*dd*–V*dd* story" is subsequently recycled in the "V*dd*–Gnd story." Due to this internal recycling, half as much current is drawn compared to the conventional scheme, with almost the same total power consumption. A reduced current is beneficial since it cuts down the supply

noise. Thus, in the best case, if the currents in the two subcircuits are completely balanced, the middle supply path will sink zero current. This results in minimal noise on that rail, as also illustrated in Fig. 2.20

However, the main issue with this technique is the requirement of separate body islands. This may be difficult in typical bulk processes. If and only if we consider 3D ICs, the tiers are inherently separated electrically, which makes MSPD particularly attractive.

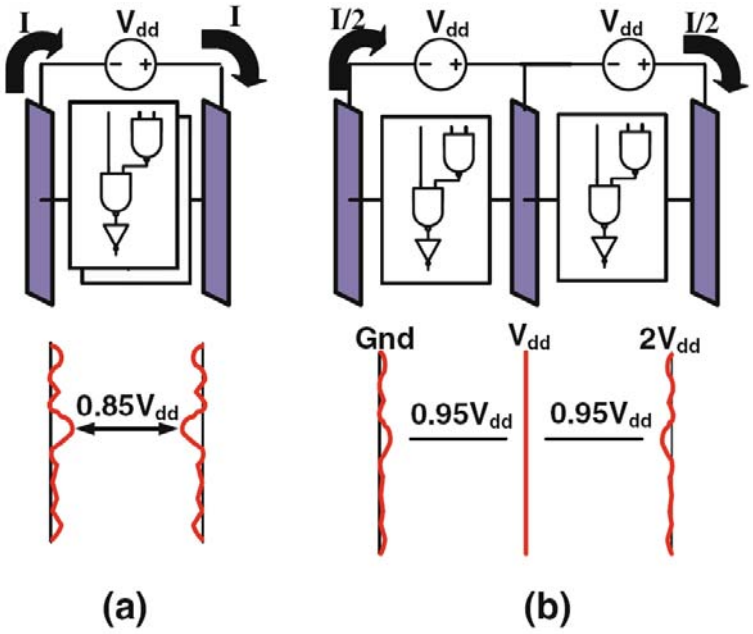


Fig. 2.20 The (**a**) conventional and (**b**) multistory power delivery schemes.

Chapter 3
Power TSV Placement and Optimization for TSV 3D Integration

As technology advances towards Gigascale Integration (GSI), chips require higher current densities and lower supply voltages in their power distribution networks. The tolerable power supply noise of circuits, as a result, decreases and makes design of power distribution networks more challenging.

IR-drop and Simultaneous Switching Noise (SSN) are the two main components of power supply noise. IR-drop results from the supply current passing through the parasitic resistance of the power distribution network. SSN is caused by inductance of the power delivery system, and occurs when a group of circuits switch simultaneously. Among three distinct droops of SSN, the first droop has the shortest duration and largest magnitude, thus it influences chip performance most severely [3.1]

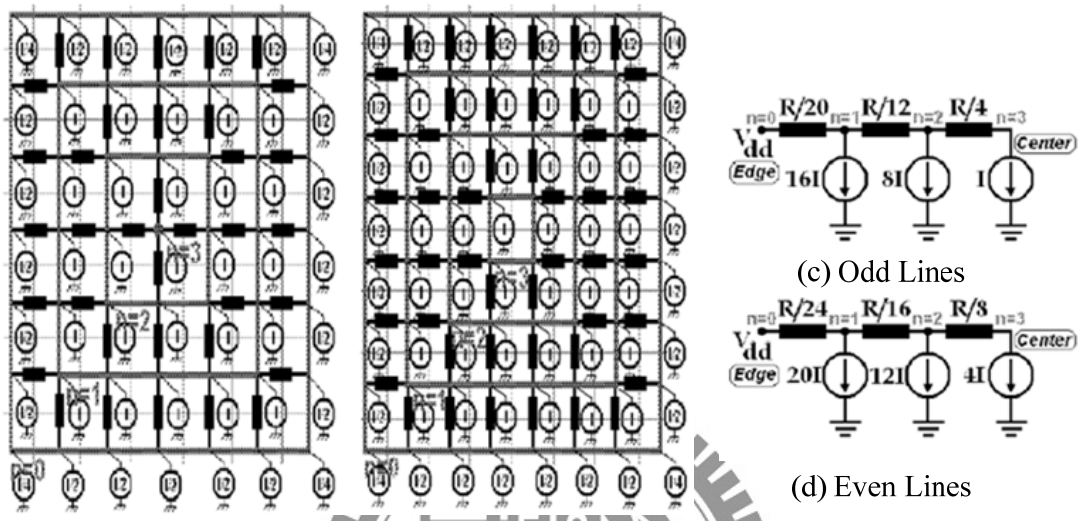
3.1 Analysis Power Supply Grids on 2D System

In section 3.1, we will introduce the optimization method for power supply grids on 2D system.

3.1.1 Optimum Sizing of Power Grids for IR Drop

As IR drop is only concerned in [3.2], only the resistive nature of the power grid lines will be considered, and a constant DC current load will represent the average load

current. The grid structure exhibits large symmetry around the origin. The power grid now has a group of square-shaped concentric nodes. Grouping parallel resistors and current sources in Fig. 3.1(a) and (b), the corresponding reduced grid models in Fig. 3.1(c) and (d) are obtained.



(a) Odd Lines (b) Even Lines

(c) Odd Lines

(d) Even Lines

Fig. 3.1 Power grid structure and Reduced models

This paper targeted the initial design of power grids. It presented simple reduced models based on equipotential-nodes approximation, and simple yet accurate analytical optimum line width to uniform one formula to calculate the IR drop with error less than 0.1%. The models and the IR drop formula accurately captured the effects of the design parameters on IR drop. Thus, they have high fidelity in comparing different grid designs regarding the IR drop. This paper also derived the optimum sizing scheme of power grid lines for minimum IR drop at the grid center. Using this technique, a uniform grid was optimum for uniform load current profiles. For real chips examples, the paper showed a reduction of 14% in IR drop by using the optimum sizing rather than the uniform one.

3.1.2 LdI/dt Drop on Power Grids

However, with increased clock frequencies and power supply demands, on-chip inductance has become a significant factor in the total LdI/dt drop. In [3.3], the author presents two new power grid topologies that reduce the voltage drop induced by on chip inductance. The original power grid design, which we refer to as the interdigitated topology, is shown in Fig. 3.2 (a) and consists of alternating power and ground lines that are equally spaced. However, due to the large spacing of the power grid wires, this topology results in large inductance and hence, a high supply voltage drop. The topology is shown in Fig. 3.2 (b), which is refer to as the single layer paired topology, reduces the spacing between power and ground lines by adding orthogonal routs between the Vdd supply pad the Vdd wires at the top layer. According to [3.3], this configuration significantly reduces the total voltage drop. However, this topology significantly reduces the routability of the top interconnect layer since it requires orthogonal supply wires. Hence we proposed a so-called multi-layer paired topology in Fig. 3.2 (c). In this topology the spacing between power and ground lines is reduced, similar to that in the single paired topology, however, the routability of the design is not reduced since we use paring of the supply lines at different layers.

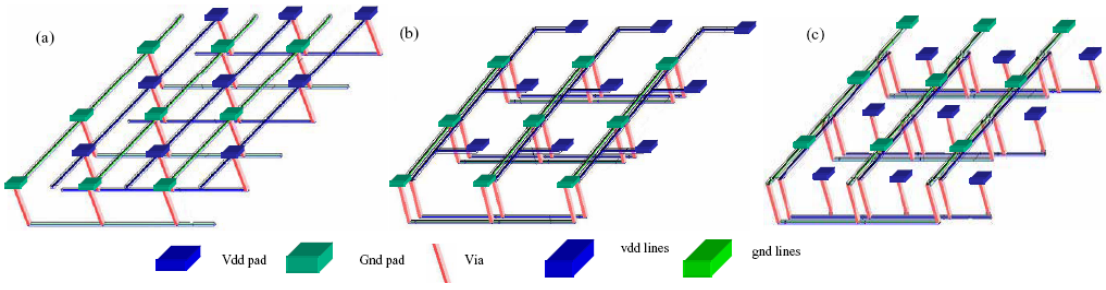


Fig. 3.2 Three designs of power distribution grids. (a) interdigitated power grid, (b) single paired power grid, and (c) multi paired power grid.

The voltage drop due to on-chip inductance was reduced by 70% using the

proposed topology. The author examines the layer dependency of the proposed paired topology. As expected, applying the paired power grid topology at the top layers yields the most significant reduction in Ldi/dt drop.

3.2 Though-Silicon Via (TSV) Electrical Characteristic

Most previous work characterizing 3-D vias has focused on bulk silicon and emphasized the experimental extraction of the via resistance and capacitance. Due to the large variation in the 3-D via diameter, length, dielectric thickness, and fill material, a wide range of measured resistances, capacitances, and inductances have been reported in the literature. Single 3-D via resistance values vary from 20 mΩ to as high as 350 Ω, while reported capacitances vary from 40 fF to over 1 pF. A few researchers have reported measured via inductances that range from as low as 42 pH to as high as 255 pH. Alternatively, 3-D via modeling has primarily been applied to simple structures to verify the measured RLC values while providing some insight into 3-D via to 3-D via capacitive coupling.

3.2.1 TSV Electrical Characteristic

Electrical characterization of a single via for several 3-D via configurations is described in [3.4]. Multiple 3-D via configurations has been simulated with and without the ground plane as illustrated in Fig. 3.3. These configurations include signal propagation from T1 to T2, T2 to T3, and T1 to T3. The 2 μm thick aluminum ground plane is located at the top of the silicon substrate.

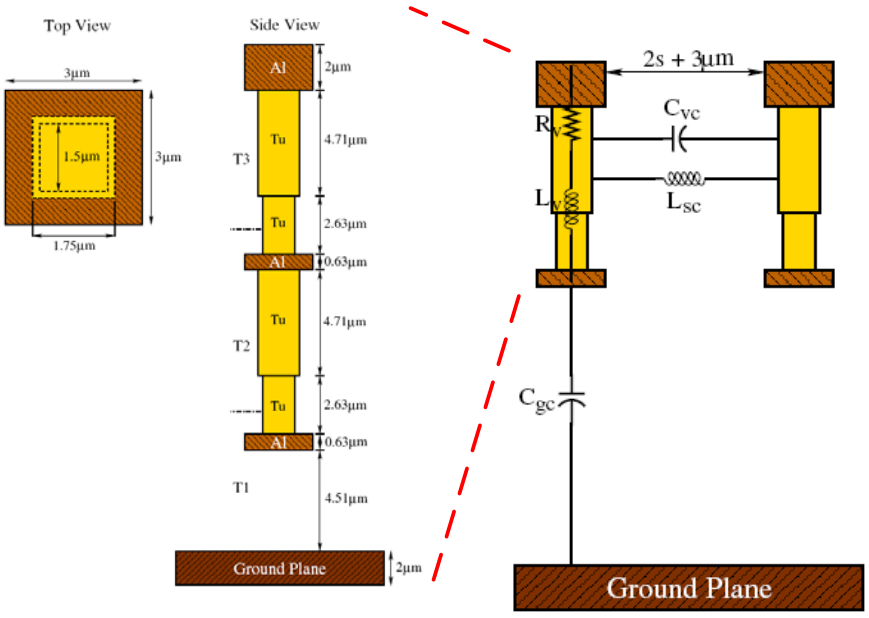


Fig. 3.3 Circuit model for RLC extraction of two 3-D vias

The L/R time constant is approximately four orders of magnitude greater than the RC time constant. In addition, the ground plane does not affect both the resistance and inductance, but the total via capacitance is significantly affected, with the capacitance increasing by as much as 37%.The L/R time constant is shown to dominate the RC time constant, implying that a 3-D via is inductively limited.

Once a single 3-D via is electrically characterized, the capacitive and inductive coupling between two 3-D vias, as shown as Fig. 3.3, can be investigated. This analysis has been performed on various configurations with increasing separation between the two 3- D vias. A ground plane is present in all cases. The tabulated results reveal several interesting characteristics in addition to the expected trend of decreasing coupling capacitance (C_{vc}) and inductance (L_{vc}) with increasing via separation.

Analysis of two 3-D via system reveals that with increasing spacing between the two vias, the RLC impedances approach those of a single 3-D via. In addition, the

analysis of two 3-D via systems verifies that inter-plane coupling must be considered when properly modeling 3-D integrated structures.

3.2.2 Closed-Form Expressions of TSV

In [3.10], Closed-form expressions of the resistance, inductance, and capacitance of a 3-D via have revealed good agreement with full-wave electromagnetic simulation. Errors of less than 6% between the closed-form models and simulation have been demonstrated for both the resistance and inductance of a 3-D via. Errors of less than 8% for the capacitance have also been reported. The use of these closed-form expressions rather than full-wave electromagnetic simulations to estimate the 3-D via impedances enhances the system-level design process. These models of the via impedance are accurate over a wide range of diameters, lengths, dielectric thickness, and spacing.

However, the dimensions of a 3-D via are highly technology dependent. For SOI processes, where the buried oxide behaves as a natural stop for wafer thinning, the length L of the TSVs is much shorter than the bulk counterparts. The diameter D of the vias follows a similar pattern, where SOI processes utilize smaller diameters than bulk technologies. The 3-D via models developed here for an SOI process therefore consider diameters of 1, 5, and 10 μm, while the 3-D vias implemented in bulk silicon have diameters ranging between 20 and 60 μm. A 20 μm diameter is chosen for the comparison between the 3-D via model and the equivalent cylindrical structure. In all cases, the via lengths are chosen to maintain an aspect ratio L/D of 0.5, 1, 3, 5, 7, and 9. In addition, the dielectric thickness for the bulk processes is 1 μm, and the material filling the 3-D vias in both SOI and bulk technologies is tungsten.

At DC situation, a closed-form expression of the 3-D via resistance is presented

next as (1). The DC resistance is only dependent on the length L, radius R, and conductivity σ. Expressions for the DC partial self-(L_{11}) nd mutual (L_{21}) inductances are provided in (2)–(3). The inductance models with a fitting parameter to adjust for inaccuracies in the Rosa expressions.

$$R_{DC} = \frac{1}{\sigma}\frac{L}{\pi R^2} \tag{3.1}$$

$$L_{11} = \alpha\frac{\mu_0}{2\pi}\left[\ln\left(\frac{H+\sqrt{H^2+\left(\frac{D}{2}\right)^2}}{\frac{D}{2}}\right)H + \frac{D}{2} - \sqrt{H^2+\left(\frac{D}{2}\right)^2} + \frac{H}{4}\right] \tag{3.2}$$

where $\alpha = 1 - e^{\frac{-4.3H}{D}}$

$$L_{12} = \frac{\mu_0}{2\pi}\left[\ln\left(\frac{H+\sqrt{H^2+P^2}}{P}\right)H + P - \sqrt{H^2+P^2}\right] \tag{3.3}$$

Equation (3.4) accounts for both the formation of a depletion region surrounding a p-type bulk substrate and the termination of the electrical field lines on a ground plane below the 3-D via. The termination of the field lines from the 3-D via to the ground plane forms a capacitance to the on-chip metal interconnect,

$$C = \alpha\beta\frac{\varepsilon_{SiO_2}}{t_{diel}+\frac{\varepsilon_{SiO_2}}{\varepsilon_{Si}}X_dT_p}2\pi RH \tag{3.4}$$

Note that (44) is dependent on the depletion region depth X_dT_p in doped p-type silicon (the doped acceptor concentration N_A is 10^{21} m^{-3} in this case). The depletion region is, in turn, dependent on the p-type silicon work functionφ_{fp}. The intrinsic semiconductor concentration ni is 1.5×10^{16} m^{-3}, and the silicon permittivity is $11.7 \times (8.85 \times 10^{-12})$ F/m. The thermal voltage kT/q at T = 300 K is 25.9 mV, where q is the electron charge (1.6×10^{-19} C) and k is the Boltzmann constant, 1.38×10^{-23} J/K,

$$X_dT_p = \sqrt{\frac{4\varepsilon_{Si}\phi_{f_p}}{qN_A}} \tag{3.5}$$

$$\phi_{f_p} = V_{th} \ln\left(\frac{N_A}{n_i}\right) \tag{3.6}$$

The fitting parameters α and β are used to adjust the capacitance for the two physical factors. The β parameter adjusts the capacitance of a 3-D via since a smaller component of the capacitance is contributed by the portion of the 3-D via farthest from the ground plane. A decrease in the growth of the capacitance therefore occurs as the aspect ratio increases. The α term is used to adjust the capacitance based on the distance to the ground plane Sgnd. As Sgnd increases, the capacitance of the 3-D via decreases. The α and β terms are

$$\alpha = \left(-0.0351\frac{H}{D} + 1.5701\right) S_{gnd_{\mu m}}^{0.0111\frac{H}{D}-0.1997} \tag{3.7}$$

$$\beta = 5.8934 D_{\mu m}^{-0.553} \left(\frac{H}{D}\right)^{-\left(0.0031 D_{\mu m}+0.43\right)} \tag{3.8}$$

An expression for the coupling capacitance between two 3-D vias over a ground plane is presented. The expression for the coupling capacitance between two 3-D vias is

$$C_c = 0.4\alpha\beta\gamma\frac{\varepsilon_{Si}}{S}\pi DH \tag{3.9}$$

The 0.4 multiplier in (9) adjusts the sheet capacitance between two TSVs when assuming that all electric field lines originating from half of the surface of one TSV terminate on the other TSV. Each fitting parameter (α, β, and γ) is used to adjust the coupling capacitance for a specific physical factor. The pitch P, which is the sum of the distance between the two vias and a single TSV diameter (P = S + D).

$$\alpha = 0.225 \ln\left(0.97\frac{H}{D}\right) + 0.53 \tag{3.10}$$

$$\beta = 0.5711\left(\frac{H}{D}\right)^{-0.988} \ln\left(S_{gnd_{\mu m}}\right) + \left(0.85 - e^{-\frac{H}{D}+1.3}\right) \tag{3.11}$$

$$\gamma = 1, \text{ if } \frac{S}{D} \leq 1 \tag{3.12}$$

3.3 Modeling and Suggestion for TSV 3D Integration

3.3.1 Physical and Electrical Modeling of TSV

According to [3.5], three chips in 3D stack interconnected through TSVs along periphery of the chips which is shown as Fig. 3.4(a). Fig. 3.4 (b) shows sample 3X3 supply grid of orthogonal layers connecting two chips through peripheral TSVs at crossing points, and the model based on orthogonal paired power and ground grids. The system of matrices shows each node on a chip connected to n neighboring nodes on same chip through supply line segment and corresponding node on other chip through TSV. And the Fig. 3.4 (c) shows TSV interconnecting three chips where each TSV includes bump and pad pair at each side which is not shown here for the sake of simplicity.

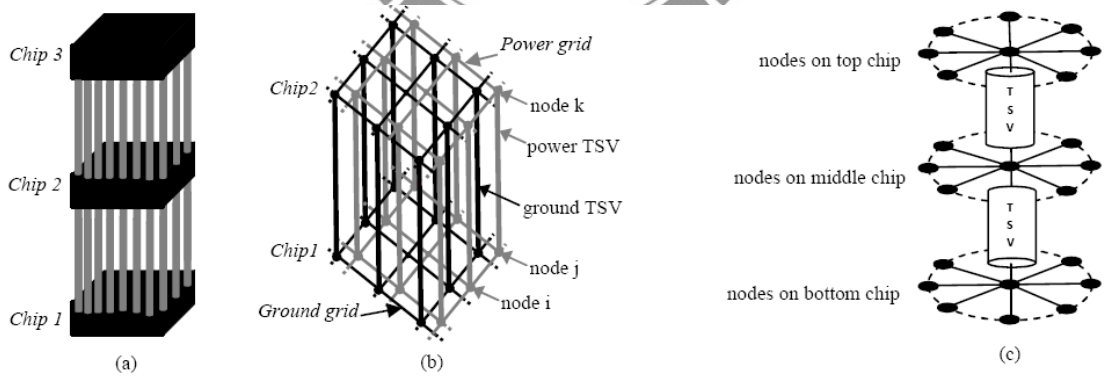


Fig. 3.4 TSV physical model

In a 3-D chip stack, depending on the TSV density, a TSV structure may be relatively isolated with no significant capacitive coupling to other TSVs, or there may be significant coupling to adjacent TSVs (refer Fig. 3.5). Also they may be laid out in a row, or within a regular matrix depending on the application and technology

constraints. The representative unit for TSVs in a regular matrix is a 3×3 bundle (see Fig. 3.5 (c)). Roshan et al. [3.6] models all of these likely structures, and present close-form equations to extract resistive, capacitive and inductive components including coupling terms as a function of the TSV geometry. That has been achieved by simulating all structures in a 3-D/2-D quasi-static electromagnetic-field solver specifically used for parasitic extraction of electronic components.

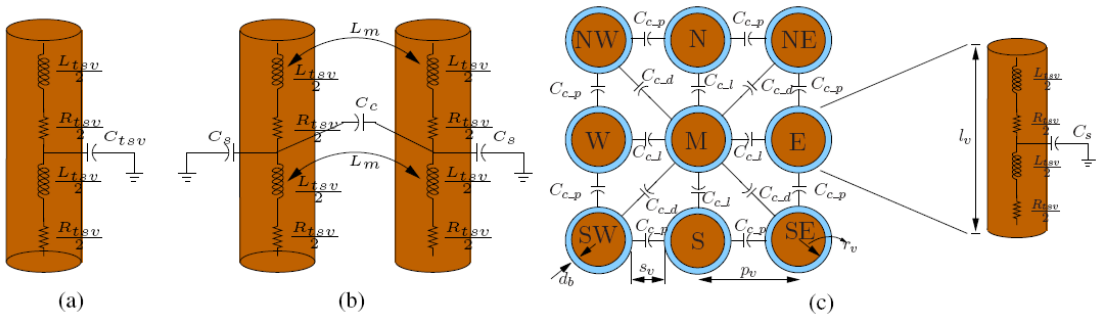


Fig. 3.5 TSV structures and electrical model.

3.3.2 TSV's parameter and optimization

According to [3.7], increase the density of the TSV placement in the stacked ICs such that the power noise is reduced below an acceptable voltage drop tolerance. Increasing the density of TSV reduces the voltage drop caused by the TSV failure. Fig. 3.6, demonstrates iterations among the dimensions of the TSV and the changes of the aspect ratio (height to cross section) of the TSV. Increasing the TSV cross section area will reduce the impedance of the PDN structures and as a result mitigate the power noise. However, increase in the dimension and density will reduce the routable area of the stacked dies. According to the Design Rule Check, there is a keep out area surrounding the TSV in the contact of the TSV and tier power grids where no signal or power could be routed and no hard macro could be placed.

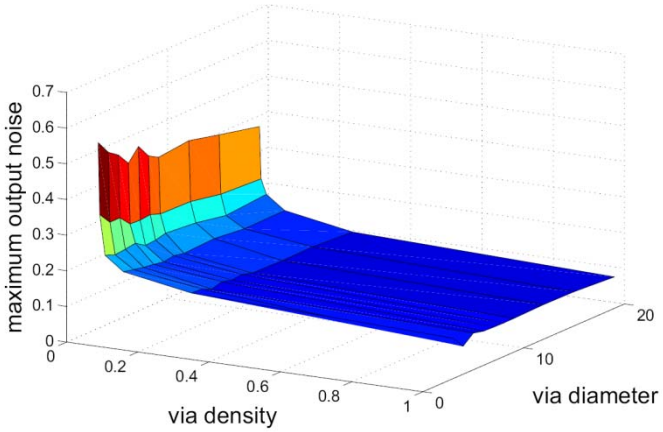


Fig. 3.6 TSV design with TSV pitch and width scaling.

Selvanayagam et al. [3.8] performed comprehensive study on thermo-mechanical reliability of the TSV for different TSV dimensions. Their study shows the relation between the thermal strain and diameter of the TSV is derived. The increase in the TSV diameter will increase the thermo-mechanical strains and as a result the reliability is reduced. The thermal resistance of the stacked 3D PDN reduces with the increase in the TSV densities.

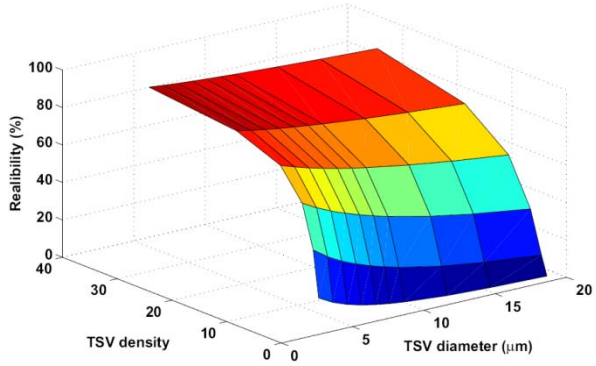


Fig. 3.7 Normalized TSV reliability factor.

The relative normalized TSV reliability factor for fixed number of operation cycles is derived in Fig. 3.7. Fig. 3.7 shows that as the TSV diameter increases, the relative reliability factor is reduced. The 3D PDN reliability is also increased with the increase in the density of the TSV in the stacking. In addition, from Fig. 3.6 it is

observed that we have the minimum power noise for the largest TSV diameters and highest density.

These conclude all observation in Fig. 3.8(a) which represents the cost function: Failure rate × max voltage drop. There is a tradeoff between the reliability and the power noise reduction as the TSV diameters increases. Increases in the diameter will reduce the voltage noise but decrease the reliability factor which is demonstrated in Fig. 3.8 (a). Besides, Fig. 3.8 (b) concludes the TSV analysis framework. To obtain the minimum cost function (Failure rate × max voltage drop) with the minimum feasible TSV block out area, the author minimize the failure rate × maximum voltage drop × block out area . The minimum, in Fig. 3.8 (b), is the configuration where the 3D PDN has the least cost function (failure × noise) under local TSV failures and the block out region is minimized.

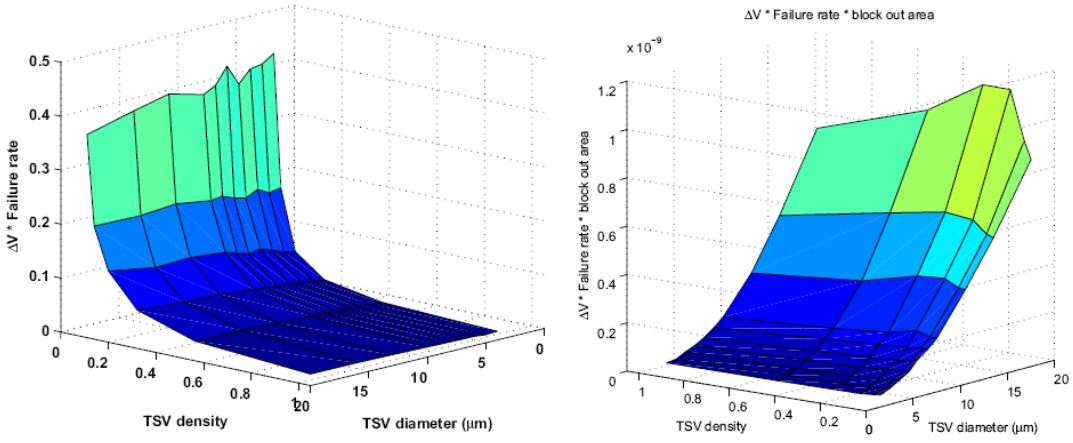


Fig. 3.8 (a)Maximum ΔV × Failure rate vs. TSV pitch and size (b) Max ΔV × Failure rate × Block out area.

However, various techniques can impact the quality of power delivery in 3D ICs. These include through-silicon via (TSV) size and spacing, controlled collapse chip connection (C4) spacing, and a combination of dedicated and shared power delivery. In [3.9], their evaluation system is composed of quad-core chip multiprocessor,

memory, and accelerator engine. Each of these modules is running representative SPEC benchmark traces. And they present a set of guidelines for designing and optimizing power delivery networks in future 3D designs:

- Locality in the vertical dimension impacts both IR drop and Ldi/dt voltage droop trends in a 3D PDN. A voltage droop at a node in 3D can get current from decoupling caps in the vertical neighbors as well as from the ones in the same plane. The resulting behavior is dependent on the locality of the droop as well as the state of the neighboring nodes. Therefore, a detail 3D PDN analysis with architecture or module level placement using representative workloads is necessary during 3D chip design.

- A critical observation in [3.9] is the saturation trend of IR drop in 3D PDNs with increased TSV size. This suggests the need for first finding the optimal TSV size given the on-chip grids in 3D stacked layers such that the least amount of silicon area penalty is incurred.

- While it is generally expected that the power delivery would be affected most in the die stacked furthest away from the C4 connections, the author report that percentage degradation in power delivery is in fact worse in lower level dies closer to C4s. This is particularly true when a highly active module, such as PROC, is placed next to heat sink for thermal concerns and furthest away from C4 connections. Therefore, 3D PDN analysis needs to carefully consider the impact in all the dies while optimizing the grid.

- Increasing the TSV granularity or equivalently decreasing the TSV spacing in 3D PDN improves the standard deviation in IR drop and Ldi/dt voltage droop most, with marginal improvements in maximum and average values. Therefore,

physical design for 3D PDN must consider this impact and choose TSV granularity accordingly.

- Despite selecting the optimal TSV size and TSV spacing, 3D PDN performs worse in both IR drop and Ldi/dt voltage droop compared to 2D PDN if the package connection, such as C4, pitch or granularity is maintained the same as in the 2D case. This study shows that improving off-chip component of the 3D PDN, for example through reducing C4 pitch for a higher number of C4s, has the highest relative impact on power grid metrics that enables 2D like or even better quality 3D PDN.

- A combination of shared and dedicated TSV power delivery can be used, as illustrated in 3D TAP configuration in Fig. 3.9 , to achieve improvements in both IR drop and Ldi/dt voltage droop.

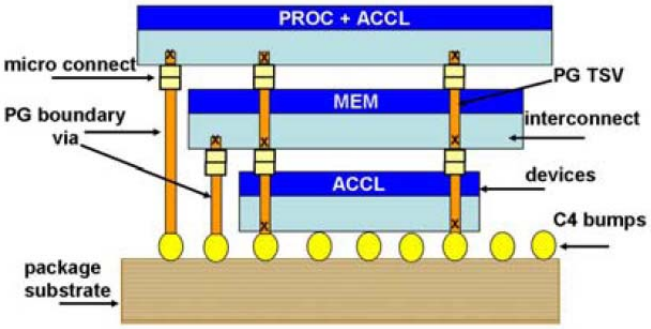


Fig. 3.9 Tapered Stacked (TAP) 3D Configuration

3.4 Proposed Methodology of Power TSV Placement and Optimization in TSV 3D Integration

Increasing the TSV cross section area will reduce the impedance of the PDN structures and as a result mitigate the power noise. However, increase in the dimension and density will reduce the routable area of the stacked dies. Therefore, a design method of power TSV in TSV 3D integration is proposed to derive the adequate TSV parameter in order to satisfy the efficiency condition between power noise reduction and area overhead.

3.4.1 Power Grids Noise Estimation

In [3.13], an equivalent circuit model to investigate the noise behavior in terms of the transition time is shown in Fig. 3.10, where R, L represent the power and ground impedances, C is the decoupling capacitor. The parasitic resistance and inductance of the power and ground networks are assumed to be equal due to the symmetry of these two TSV bundles. Note that this model does not consider the feedback effect of the power noise since in this model the current is independent of this noise.

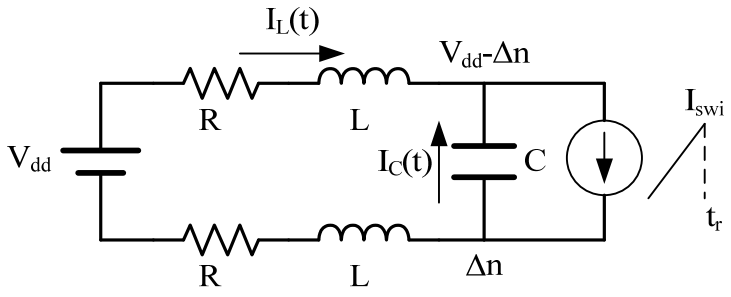


Fig. 3.10 Equivalent circuit model of power grid

The current provided by the decoupling capacitance $I_C(t)$ and the current flowing through the parasitic inductance $I_L(t)$ from the power supply are, respectively

$$I_C(t) = -C\frac{\partial V_c}{\partial t} \tag{3.13}$$

$$I_L(t) = \frac{1}{L}\int_0^t V_L(t)\,\partial t \tag{3.14}$$

Where VC(t) and VL(t) are, respectively

$$V_C(t) = V_{dd} - 2\Delta n(t) \tag{3.15}$$

$$V_L(t) = \Delta n(t) - I_L(t)R \tag{3.16}$$

A ramp function is assumed for the noise $\Delta n(t)$ as

$$\Delta n(t) = \frac{V_{noise}}{t_{r,v}}t \tag{3.17}$$

where V_{noise} is the peak noise voltage and $t_{r,v}$ is the transition time of the noise spike. Then we can get the result in the following differential equations:

$$I_C(t) = \frac{2CV_{noise}}{t_{r,v}} \tag{3.18}$$

$$\frac{\partial I_L(t)}{\partial t} = \frac{V_{noise}t}{t_{r,v}L} - \frac{R}{L}I_L(t) \tag{3.19}$$

Solving these differential equations with the initial conditions $I_C(0) = 0$ and produces the $I_L(0) = 0$ inductive and capacitive current, respectively, as

$$I_C(t) = V_{noise}\frac{2C}{t_{r,v}} \tag{3.20}$$

$$I_L(t) = V_{noise}\left[\frac{t}{t_{r,v}R} - \frac{L}{t_{r,v}R^2}\left(1 - e^{\frac{-tR}{L}}\right)\right] \tag{3.21}$$

Assume peak noise occurs when current reaches the maximum current, we can derivate the V_{noise} equation as

$$I_{swi} = I_C(t_r) + I_L(t_r) = V_{noise}\left\{\frac{2C}{t_r} + \left[\frac{t_r}{t_rR} - \frac{L}{t_rR^2}\left(1 - e^{\frac{-t_rR}{L}}\right)\right]\right\} \tag{3.22}$$

$$V_{noise} = \frac{I_{swi} t_r R^2}{2CR^2 + t_r R - L\left(1 - e^{\frac{-t_r R}{L}}\right)} \tag{3.23}$$

3.4.1 Topology of power TSV Bundle

A simplified circuit model can be used to analyze this power distribution network, as shown in Fig. 3.11. However, consider a power transmission circuit as shown in Fig. 3.12.

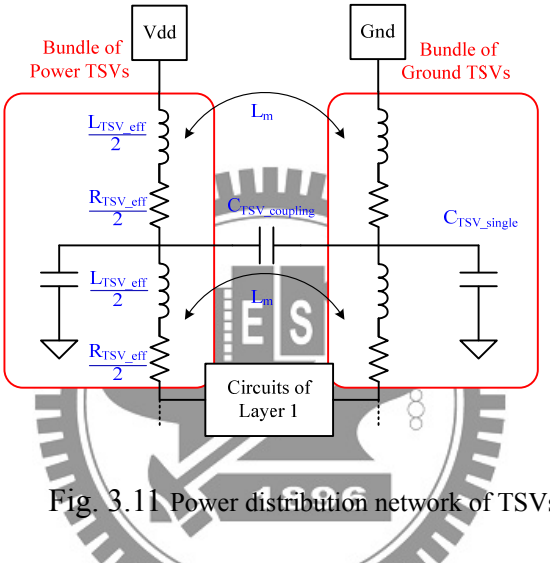


Fig. 3.11 Power distribution network of TSVs

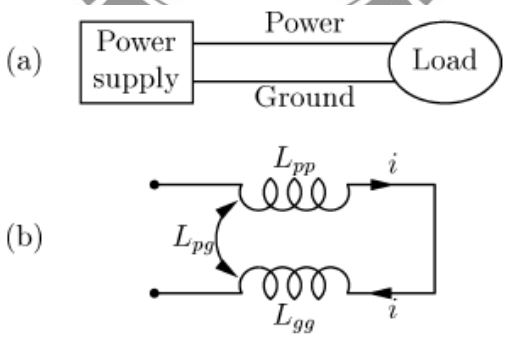


Fig. 3.12 A simple power transmission circuit. (a) block diagram, (b) the equivalent inductive circuit

The circuit consists of the forward current (power) path and the return current (ground) path forming a transmission current loop between the power supply at one end of the loop and a power consuming circuit at the other end [3.11][3.12]. In a simple case, the forward and return paths each consist of a single conductor. In

general, a “path” refers to a multi-conductor structure carrying current in a specific direction (to or from the load), which is the case in power distribution grids. The power transmission loop consists of the forward and return current paths. The equivalent inductive circuit is depicted in Fig. 3.12(b). The partial inductance matrix for this circuit is

$$L_{ij} = \begin{bmatrix} L_{pp} & -M_{pg} \\ -M_{pg} & L_{gg} \end{bmatrix} \tag{3.24}$$

where L_{pp} and L_{gg} are the partial self inductances of the forward and return current paths, respectively, and Mpg is the absolute value of the partial mutual inductance between the paths. The loop inductance of the power transmission loop is

$$L_{loop} = L_{pp} + L_{gg} - 2M_{pg} \tag{3.25}$$

The mutual coupling Mpg between the power and ground paths reduces the loop inductance. This behavior can be formulated more generally: the greater the mutual coupling between antiparallel (flowing in opposite directions) currents, the smaller the loop inductance of a circuit. The effect is particularly significant when the mutual inductance is comparable to the self inductance of the current paths.

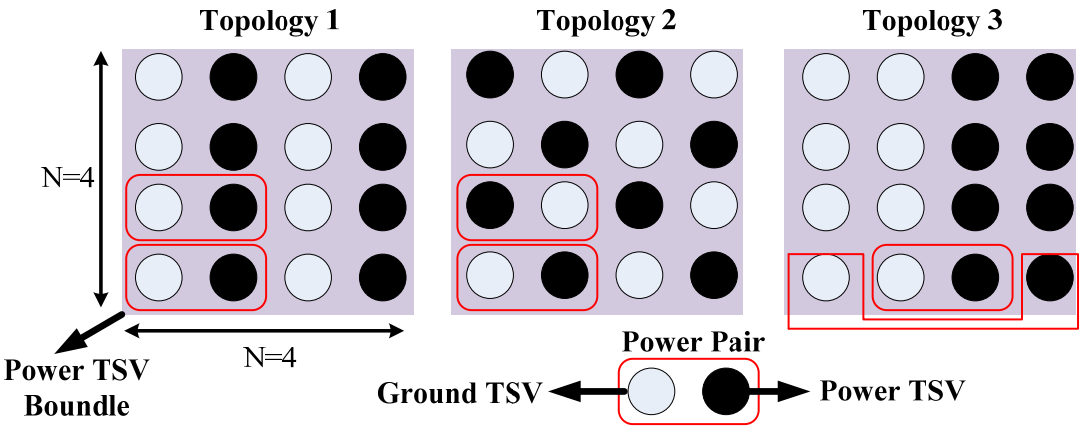


Fig. 3.13 Topology of TSV bundle

Therefore, we propose three topologies for placing TSVs into a bundle. As Fig.

3.13 is shown, it's an example of three TSV topologies in a 4x4 TSVs bundle. The first topology places the power TSVs and ground TSVs column by column. The second topology interleaves the power TSVs and ground TSVs both in every column and row, such that every power TSV is neighbor with a ground TSV. The last topology put power TSVs at one side and ground TSVs at another side.

In the bundle of TSVs, the white circle is presented as power TSVs and black circle is presented as ground TSVs. The red circle which including a pair TSV (power & ground) is presented as one pair of two parallel TSVs with opposite current flow. In this example of 4x4 TSVs bundle, there are eight pairs. Then, we can calculate the inductance of each pair (two TSVs). The inductance of the first pair is

$$L_1 = L_{1_p} + L_{1_g} + \sum_{i=2}^{8}\left(M_{1_p i_p} + M_{1_g i_g}\right) - \sum_{i=2}^{8}\left(M_{1_p i_g} + M_{1_g i_p}\right) \quad (3.26)$$

where subscripts p and g represent power and ground, respectively.

We measure the power supply noise from 100MHz to 200MHz of the bundle of TSVs with different topologies. In each case, we demonstrate not only with different Diameter of TSVs but also including different length of TSVs (Height). Evan more, we also consider the electrical characteristic with two kinds of filled material, copper and tungsten. The simulation result of power supply noise is shown as Table 3.1 and Table 3.2.

Filled by Tungsten

Table 3.1 Normalized supply noise with TSV filled by Tungsten

(a) Normalize Power Supply Noise with H=60 filled by Tungsten

H=60	**D(μm)**					
Topology	10	20	30	40	50	60
1	105.85%	111.74%	115.97%	114.35%	113.62%	112.33%
2	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
3	120.13%	164.09%	180.09%	166.55%	161.43%	156.48%

(b) Normalize Power Supply Noise with H=80 filled by Tungsten

H=80	D(μm)					
Topology	10	20	30	40	50	60
1	103.91%	114.45%	118.03%	119.14%	115.20%	114.30%
2	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
3	120.85%	177.99%	191.26%	186.96%	180.23%	173.52%

(c) Normalize Power Supply Noise with H=100 filled by Tungsten

H=100	**D(μm)**					
Topology	10	20	30	40	50	60
1	104.45%	115.05%	118.69%	117.58%	111.82%	125.26%
2	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
3	122.87%	185.70%	198.15%	189.62%	173.68%	186.48%

Filled by Copper

Table 3.2 Normalized supply noise with TSV filled by Copper

(a) Normalize Power Supply Noise with H=60 filled by Copper

H=60μm	D(μm)					
Topology	10	20	30	40	50	60
1	114.77%	119.57%	117.29%	114.43%	113.29%	112.34%
2	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
3	181.75%	202.98%	186.53%	162.72%	160.91%	155.65%

(b) Normalize Power Supply Noise with H=80 filled by Copper

H=80μm	D(μm)					
Topology	10	20	30	40	50	60
1	115.15%	118.29%	123.76%	115.28%	115.59%	113.29%
2	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
3	189.23%	202.76%	203.87%	188.59%	178.85%	168.44%

(c) Normalize Power Supply Noise with H=100 filled by Copper

H=100μm	D(μm)					
Topology	10	20	30	40	50	60
1	115.57%	124.80%	122.68%	121.41%	121.05%	116.62%
2	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
3	195.15%	221.86%	206.58%	195.36%	188.66%	172.26%

We simulate these three topologies with different Height, Diameter, filling material of TSV, seems the topology 2 is the most suitable topology for placing TSVs into a bundle in 3D power system. And the connection between power TSV bundle to the 2D power system is also be considered as shown in Fig. 3.14. We assume the number of a TSV bundle is 4x4 in this section. However, the number of a TSV bundle

is also a parameter we need to choice. In next section, we propose a method to decide the suitable parameter of the number of TSV bundle (N) and the Diameter of TSV (D).

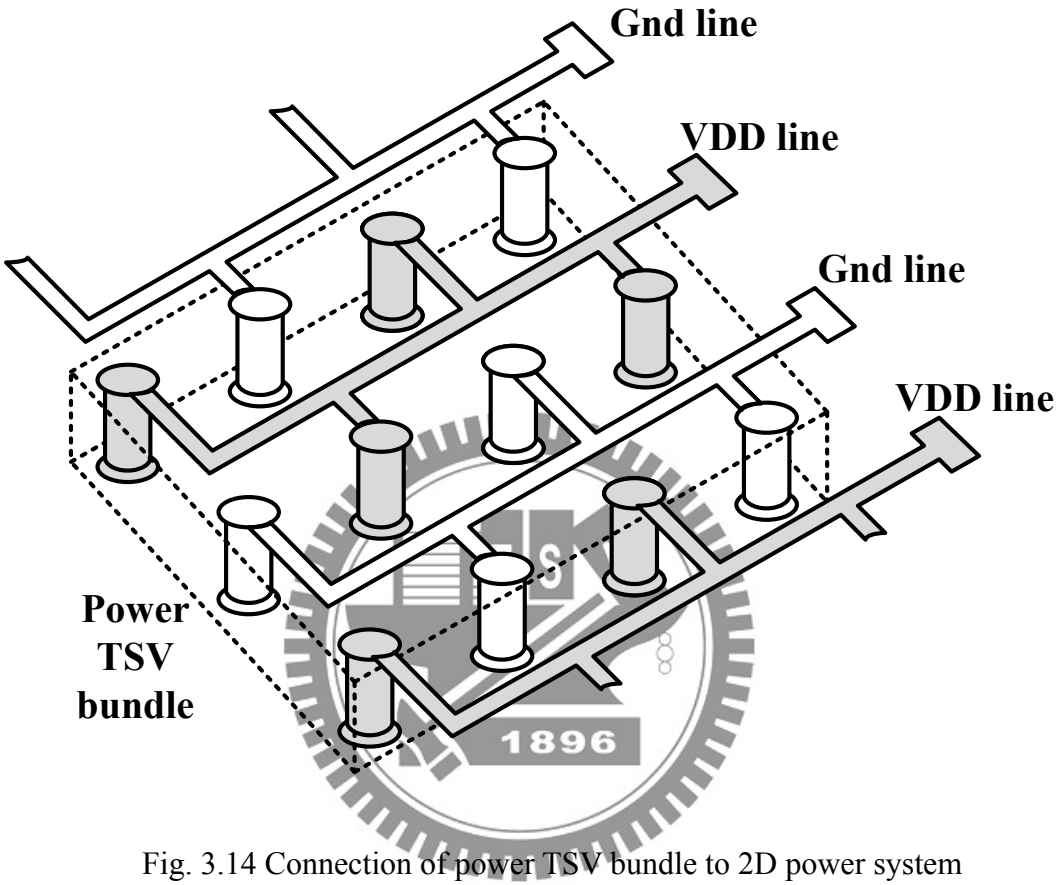


Fig. 3.14 Connection of power TSV bundle to 2D power system

3.4.3 Simplification of the Effective Electrical Model of TSV Bundle

We assume the effective inductance of the every one TSV is only dependent with different topology and independent with number of TSV bundle (N). In the section 3.3.2, the topology 2 has the lowest noise of power supply. The following

Table 3.3 is shown the normalize value of effective inductance difference of the every one TSV with different number of TSV bundle (N).

Table 3.3 Normalized inductance difference on TSV bundle.

(a) Error of each number of TSV bundle (N) compare to 2x2 with H=60μm

H=60μm	D(μm)					
N	10	20	30	40	50	60
2X2	-	-	-	-	-	-
4X4	4.03%	3.56%	3.16%	2.84%	2.61%	2.44%
6X6	5.25%	4.62%	4.09%	3.68%	3.38%	3.16%
8X8	5.81%	5.11%	4.52%	4.06%	3.73%	3.48%
10X10	6.13%	5.38%	4.75%	4.28%	3.92%	3.67%

(b) Error of each number of TSV bundle (N) compare to 2x2 with H=80μm

H=80μm	D(μm)					
N	10	20	30	40	50	60
2X2	-	-	-	-	-	-
4X4	4.15%	3.79%	3.45%	3.16%	2.91%	2.72%
6X6	5.41%	4.93%	4.48%	4.09%	3.77%	3.52%
8X8	5.99%	5.45%	4.95%	4.52%	4.16%	3.88%
10X10	6.32%	5.74%	5.21%	4.75%	4.38%	4.09%

(c) Error of each number of TSV bundle (N) compare to 2x2 with H=100μm

H=100μm	D(μm)					
N	10	20	30	40	50	60
2X2	-	-	-	-	-	-
4X4	4.22%	3.94%	3.65%	3.39%	3.16%	2.96%
6X6	5.50%	5.12%	4.74%	4.40%	4.09%	3.83%
8X8	6.10%	5.67%	5.24%	4.85%	4.52%	4.23%
10X10	6.43%	5.97%	5.52%	5.11%	4.75%	4.45%

The maximum difference in the different number of TSV bundle (N) compare to 2x2 TSV bundle is less than 6.5%. The effective inductance of each TSV in a TSV bundle under the topology 2 can be presented as

$$L = L_0 + M_{S=\sqrt{2}D} - 2M_{S=D} \tag{3.27}$$

Replace the equation (3.2),(3.3) into equation (3.27)

$$L = \alpha\frac{\mu_0}{2\pi}\left[\ln\left(\frac{H+\sqrt{H^2+\left(\frac{D}{2}\right)^2}}{\frac{D}{2}}\right)H + \frac{D}{2} - \sqrt{H^2+\left(\frac{D}{2}\right)^2} + \frac{H}{4}\right] \tag{3.28}$$

$$+\frac{\mu_0}{2\pi}\left[\ln\left(\frac{H+\sqrt{H^2+5.8D^2}}{2.4D}\right)H + 2.4D - \sqrt{H^2+5.8D^2}\right]$$

$$-2\frac{\mu_0}{2\pi}\left[\ln\left(\frac{H+\sqrt{H^2+4D^2}}{2D}\right)H + 2D - \sqrt{H^2+4D^2}\right]$$

The equation (3.28) is too complex to simplify. However, if we only consider a small region of the Diameter of TSV (10μm~60μm) and small region of the Height of TSV (60μm, 80μm and 100μm), the effective inductance could be approached to a linear function of D. (Curve fitting is used to get the linear function in this part.)

$$L = (aD_{\mu m} + b) \times 10^{-12} \tag{3.29}$$

With different Height of TSV, different parameter of ‘a’ and ‘b’ are fitted.

Table 3.4 Parameter of inductance

Parameter	H=60μm	H=80μm	H=100μm
a	-0.113914	-0.132682	-0.146685
b	15.93	21.93	27.9

After simplification of inductance of TSV, we need to simplify the resistance and the capacitor of TSV. The resistance is exactly quadratic function of D which is derivate by equation (3.1), but with different parameter ‘e’ when using different filling

material or different height of TSV.

$$R = \frac{e}{D_{\mu m}^{\ 2}} \times 10^{-3} \tag{3.30}$$

Table 3.5 Parameter of resistor

Parameter "e"	H=60μm	H=80μm	H=100μm
Filling by Tungsten	4039.7	5386.3	6732.9
Filling by Copper	1222.3	1629.8	2037.2

3.4.4 Power Grids Noise Estimation on the Bundle of TSVs

Review following equation,

$$V_{noise} = \frac{I_{swi} t_r R^2}{2CR^2 + t_r R - L\left(1 - e^{\frac{-t_r R}{L}}\right)} \tag{3.31}$$

We assume $t_r = 2.5 \times 10^{-9}$ (equal to the rising time of 200MHz). Besides, we found that following situation $\left(\frac{-t_r R}{L}\right) < 1$ and $2CR^2 \ll t_r R$. Hence, the equation (3.30) can be expanded using a Taylor series expansion to

$$V_{noise} = \frac{I_{swi} t_r R^2}{2CR^2 + t_r R - L\left[1 - \left(1 + \frac{\left(\frac{-t_r R}{L}\right)}{1!} + \frac{\left(\frac{-t_r R}{L}\right)^2}{2!} + \frac{\left(\frac{-t_r R}{L}\right)^3}{3!}\right)\right]} \tag{3.32}$$

Simplifying the equation (3.31) in

$$V_{noise} = \frac{6 I_{swi} L^2}{3 t_r L - t_r^{\ 2} R} \tag{3.33}$$

Replacing equation (3.29) into equation (3.30)

$$V_{noise}(mV) = 2 \times I_{swi} \times \frac{(aD_{\mu m} + b)^2}{t_{r(ns)}(aD_{\mu m} + b) - \frac{t_{r(ns)}^{\ 2} e}{3 D_{\mu m}^{\ 2}}} \tag{3.34}$$

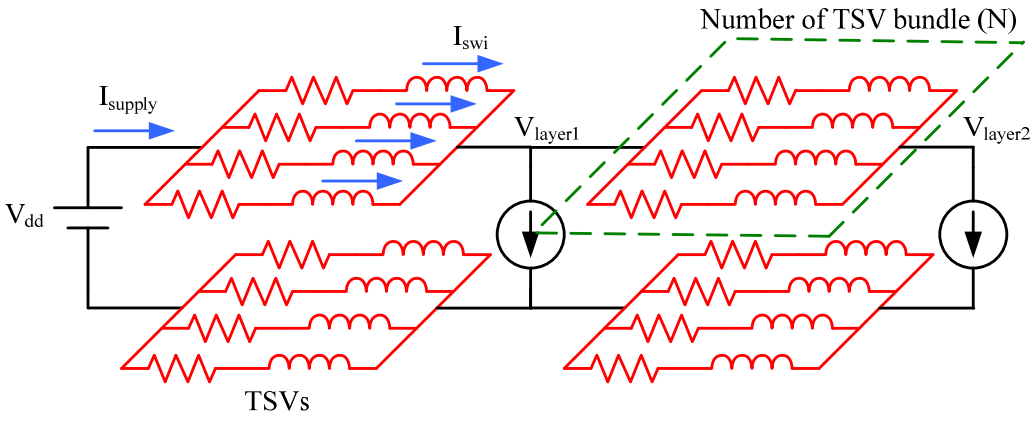


Fig. 3.15 The parameter of the TSV bundle

The parameter of the TSV bundle is shown as Fig. 3.15. The I_{swi} is the current of each TSV in a TSV bundle, the total current of the TSV bundle can be presented as I_{supply}. The parameter 'N' is the number of TSV bundle which will impact the I_{swi} on each single TSV. Such that,

$$I_{swi} = \frac{I_{supply}}{N} \tag{3.35}$$

$$V_{noise}(mV) = 2 \times \frac{I_{supply}}{N} \times \frac{(aD_{\mu m}+b)^2}{t_{r(ns)}(aD_{\mu m}+b) - t_{r(ns)}^2 \frac{e}{3D_{\mu m}^2}} \tag{3.36}$$

3.4.5 Design Method for Choosing Height and Number of TSV Bundle

According to equation (3.36), the noise of power supply is the function of D (Diameter of TSV) and N (number of TSV in a bundle). As the simulation result, we found that the V_{noise} decrease when increasing the Diameter of TSV, but the V_{noise} will be saturation when D is higher than a threshold value (D_{limit}). Therefore, we find out the D_{limit} by solving the following equation

$$\frac{\partial V_{noise}(mV)}{\partial D_{\mu m}} = -1 \tag{3.37}$$

The result is shown as the Table 5. We find that the threshold value (D_{limit}) is more related to the filling material instead the Height of TSV.

Table 3.6 Power TSV design method

D_{limit}	Filling by Copper	Filling by Tungsten
H=60μm	$D_{limit} = 5.5 \times I_{swi} + 10.3$	$D_{limit} = 8 \times I_{swi} + 17.9$
H=80μm	$D_{limit} = 6.3 \times I_{swi} + 10.5$	$D_{limit} = 9.2 \times I_{swi} + 17.9$
H=100μm	$D_{limit} = 7.1 \times I_{swi} + 10.7$	$D_{limit} = 10.3 \times I_{swi} + 18.1$

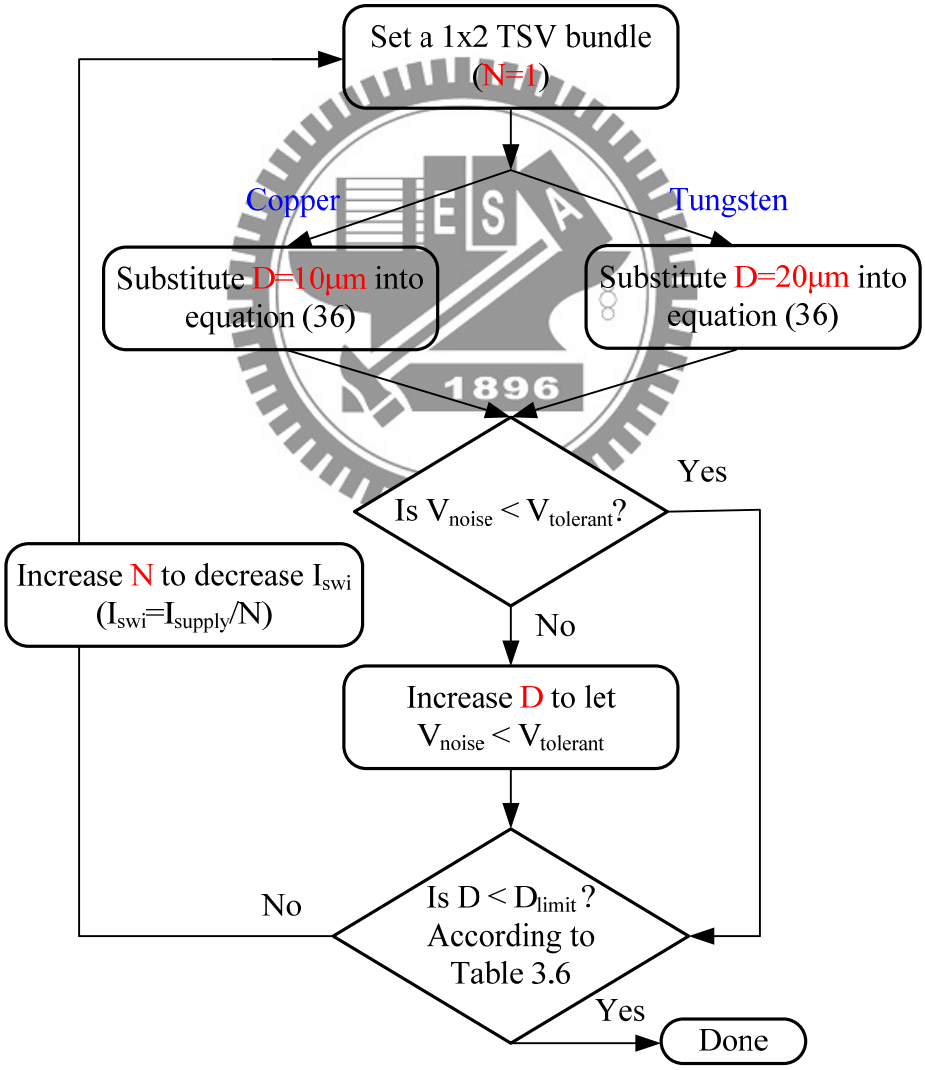


Fig. 3.16 Design Flow to select the Number (N) and Diameter (D)

Hence, we can follow the flowchart, as shown in Fig. 3.16, to choose the suitable Height of one TSV and Number of the TSV bundle. First, setting a TSV bundle with N=2 to estimate I_{swi} of a TSV. According to different filling material to select different minimum Diameter to calculate the V_{noise}. If V_{noise} is less $V_{tolerant}$, you can increase D or increase N to lower the V_{noise}. Both in each case, you always need to check that is the D large than D_{limit} or not, in order to insure that is efficiency when you are increasing the diameter of the TSV.

Finally, the example of power TSV design method is shown in Fig. 3.17. This example shows the following environment: the supply current is 1 A, and the tolerant noise supply voltage is 50mV, and the TSV is filled by copper and its height is 100μm. Fig. 3.17 shows procedures of choosing the power TSV parameter step by step.

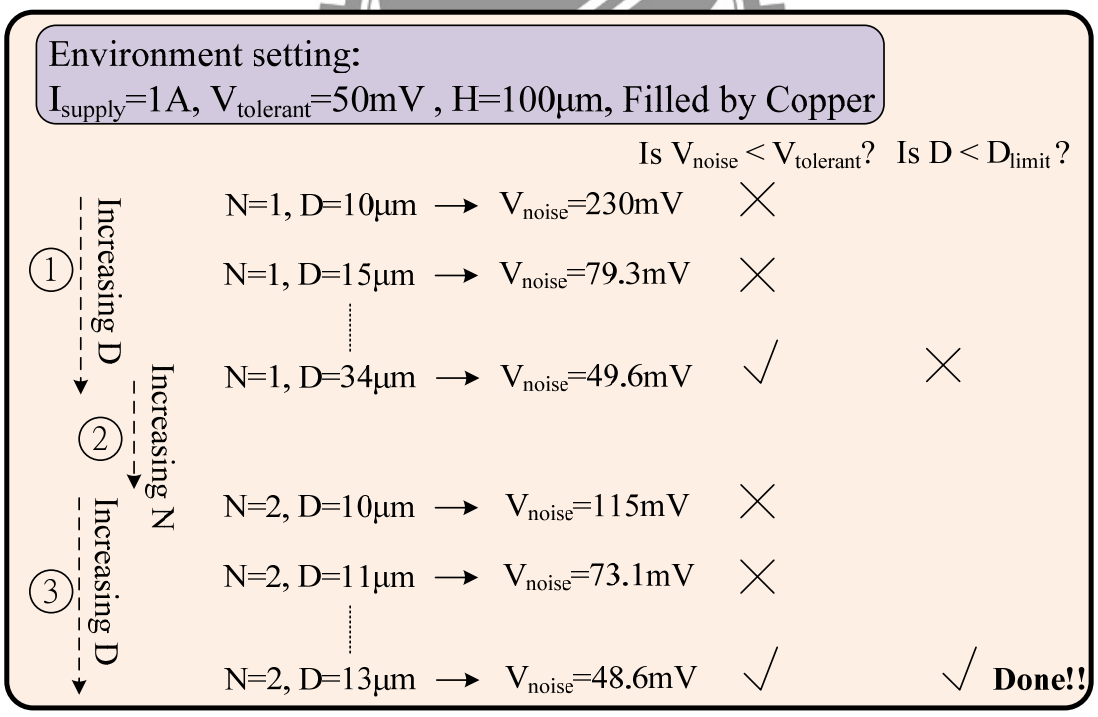


Fig. 3.17 Example of the power TSV design method

Chapter 4
Analysis Supply Noise Regulation in TSV 3D Integration

In chapter 3, we already have discussed the power grids system in 2D system, and the characteristic of a TSV. We also estimate the noise of the power grids on a TSV, and then we propose a method to choose the parameters of a TSV bundle. Next, we will discuss the technical of supply noise regulation for providing a robust supply voltage. Robust power delivery is also considered as one of the grand challenges. As predicted in the roadmap for integrated circuit (IC) development from the ITRS-2004 update [4.1], the chip working frequency and supply voltage will continue to scale aggressively, both of which will lead to significant Ldi/dt noises and IR drops on power/ground (P/G) networks, which will further affect the performances and reliabilities of VLSI chips.

Decoupling capacitors are widely used to manage power supply noise. Decoupling capacitors are an effective way to reduce the impedance of power delivery systems operating at high frequencies. A decoupling capacitor acts as a local reservoir of charge, which is released when the power supply voltage at a particular current load drops below some tolerable level. Since the inductance scales slowly, the location of the decoupling capacitors significantly affects the design of the power/ground (P/G) networks in high performance integrated circuits (ICs) such as

microprocessors. At higher frequencies, a distributed system of decoupling capacitors is placed on-chip to effectively manage the power supply noise.

However, two major constraints limit the usage of passive DECAPs in scaled technologies. First, adding on-chip DECAPs consumes a large amount of die area. Usually, a total on-chip DECAP of ~100 s of nFs or more have to be deployed on a microprocessor die to keep the supply noise within the target range [4.2]. As a result, in some high-end microprocessor chips more than 20% of the total area has been occupied by DECAPs leading to a significant waste of active die area. Second, adding on-chip DECAPs introduce large amount of gate tunneling leakage that eats into the power budget.

Analysis supply noise regulation in TSV 3D integration is realized in this Chapter. The optimization of passive decoupling capacitor for reducing power supply noise in 2D system will be discussed in the section 3.1. Section 3.2 describes another kind of reducing power supply noise, which all are using active circuits. In section 3.3, a noise suppression technique using low power active decoupling capacitors is proposed for TSV 3D integration. Through the latch-based noise detection circuitry, the power supply noise can be detected and regulated via active DECAPs.

4.1 Optimizations for Passive Decoupling Capacitor

As technology advances the operating voltage levels are reduced aggravating these problems as noise does not scale with technology. The most commonly used solution is to incorporate on-chip decoupling capacitors to the power delivery network to keep the noise within a tolerance margin ensuring circuits functionality. First, we will discuss how to place DECAPs to get a better performance. And then we will introduce the method for MOS decoupling capacitor optimization.

4.1.1 Placing the Passive Decoupling Capacitor

A methodology for efficiently placing distributed on-chip decoupling capacitors to replace one large capacitor has been proposed in [4.3]. A distributed on-chip decoupling capacitor network is an efficient solution for providing the required on-chip decoupling capacitance under existing technology constraints in nanoscale ICs. In a system of distributed on-chip decoupling capacitors, each capacitor is sized based on the parasitic impedance of the power delivery system. Hierarchically allocating the on-chip decoupling capacitors greatly relaxes the technology constraints for physically distant capacitors as the Fig. 4.1. The magnitude of the decoupling capacitors is based on the impedance of the interconnect segment connecting a specific capacitor to a current load, and this scheme makes the distant on-chip decoupling capacitors more effective [4.4].

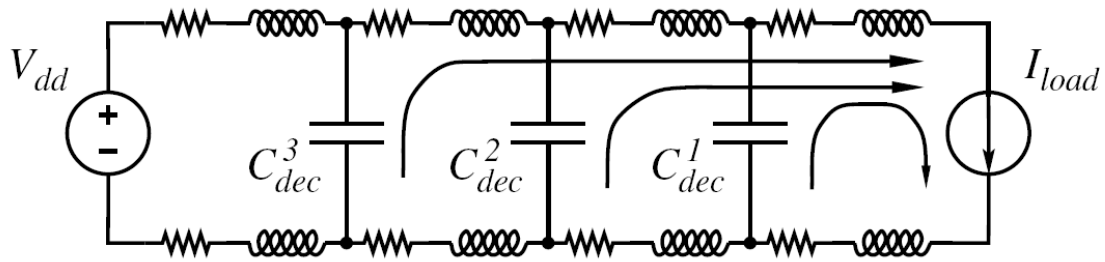


Fig. 4.1 A network of distributed on-chip decoupling capacitors.

In [4.5], the paper show that inserted DECAPs can be detrimental if its distance to the power supply pins is not considered properly for both single line and mesh type circuits. The author propose two metrics: “distance DECAP-switching source” and “distance DECAP-supply pin”, in Fig. 4.2, to determine the correct insertion position for a DECAP to be effective in reducing power supply noise. Fig. 4.2(a) shows DECAP inserted closer to switching circuit but farther away from supply (configuration A), and Fig. 4.2(b) shows DECAP inserted near supply and switching

circuit (configuration B).

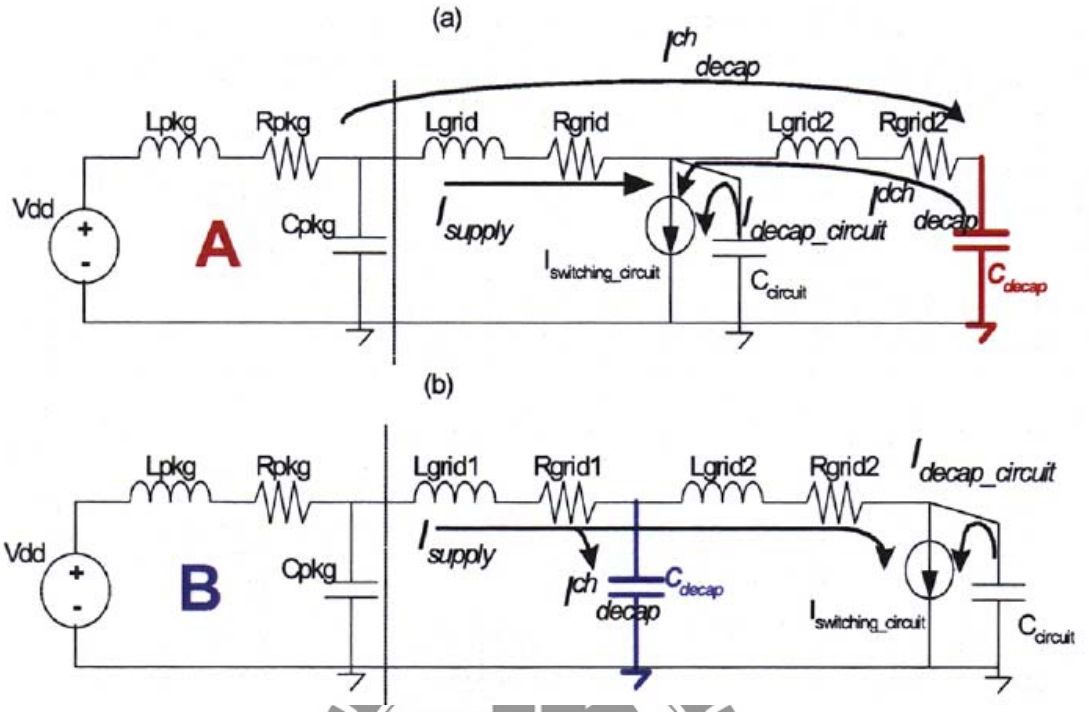


Fig. 4.2 Different location of DECAP

The author calculates the effective distance coefficients: coefficient a is the distance between Vdd pin and DECAP, coefficient a is the distance between switching source and DECAP. The author also observes that the distance DECAP-switching source (coefficient b) increases with the increase in the DECAP size and the distance DECAP-Vdd pin (coefficient a) decreases with an increase in DECAP size. These observations are intuitive: for coefficient *a,* a greater DECAP will have a greater effective distance from the source; for coefficient *b,* a greater DECAP has to be closer to the power supply pin in order to be recharged.

4.1.2 MOS Decoupling Capacitor Optimization

With technology scaling down to 90 nm and below, gate leakage current of DECAPs becomes so significant that we have to explicitly consider DECAPs leakage currents for robust P/G grid design. This is especially the case t_{ox} when shrinks below 2nm.

In [4.6], the author take a first look at the impacts of gate leakage of practical MOS-based DECAPs on P/G grid designs. To clearly show the influence of gate leakage currents of DECAPs on the P/G grids, the author propose a leakage current model for practical DECAPs. Then, analyzing the effect of leakage currents and show that practical (leaky) DECAPs may increase power consumption significantly and thus demand more routing resource or die area for achieving robust power delivery. The author proposes a more effective two-stage optimization approach to efficiently optimize P/G grids in the presence of DECAP leakage currents.

Based on minimizing the cost function of implementing DECAPs, a novel optimization technique for determining the optimum channel length of the MOS DECAPs was presented in [4.7]. This approach was applied to the 45nm and 32nm technologies, and the results showed that, on contrary to the conventional belief of setting the DECAP channel length to about 10 times the minimum channel length, the optimum channel length is different from this value and should be calculated individually for each technology node and operating frequency. Next, two optimum DECAP configurations were discussed. One only used NMOS transistors and was suitable for the applications where the area was the main concern. The other was built only by PMOS transistors and was a good candidate for low leakage applications. The comparison table is shown as Table 4.1.

Table 4.1 Specifications of DECAP configurations

Config.	C_{eff} (fF)	Area (μm^2)	C_{eff} / Area (fF / μm^2)	I_{gate} (nA)	I_{gate} / C_{eff} (nA / fF)
1st	15.09	1.637	9.22	140.01	9.28
2nd	21.04	1.637	**12.85**	284.22	13.51
3rd	9.023	0.902	10.00	39.55	**4.38**

4.1.3 Decoupling Capacitor Allocation in 3D IC

Conventional technologies for implementing DECAPs are based on SiO2-based structures that are widely used in robust power delivery network design. Three-dimensional power grid optimization has been studied in [4.8]. Unlike the 2D case, new considerations come into play while optimizing a 3D power grid using CMOS DECAPs:

Since CMOS DECAPs are usually fabricated using white space on the device layer, they must compete for area with TSVs, or with the landing pads of 3D vias, for the limited white space. This leads to a new resource contention problem. One way to resolve this contention problem is to increase the chip size in order to make room for CMOS DECAPs. However, one of the advantages of 3D circuits over 2D implementations is that they result in a reduced chip footprint: increasing the chip size may counteract this benefit.

Leakage power is an important issue in 3D circuit design. The CMOS DECAPs added to the 3D circuit will consume extra leakage power and make things worse. While new high-k dielectrics have been proposed, they are not in widespread use yet and even when they are deployed, they will provide temporary relief to the gate leakage problem.

The approach in [4.9] presents an approach for decap allocation in 3D power grids, using both conventional CMOS DECAPs and metal–insulator–metal (MIM) DECAPs. Unlike CMOS capacitors that are built in the device layer, MIM capacitors are fabricated between metal layers. These structures have high capacitance density and low leakage current density. Fig. 4.3 shows the positions of CMOS and MIM DECAPs in a 3D circuit. MIM DECAPs are usually fabricated between the top two

metal layers in each 2D tier. A significant advantage of MIM DECAPs lies in their extremely low leakage: in [4.9], the leakage current for the 250 nF MIM DECAPs is reported to be about 1.0 × 10−8 A (with leakage density of 3.2 × 10−8 A/cm2), while the leakage current for a 25 nF CMOS decap in parallel with MIM is approximately 3.2 × 10−6 A (with leakage density of 1.45 × 10−4 A/cm2).

However, MIM DECAPs cannot be used unconditionally to replace CMOS DECAPs, since their use incurs a cost: they present routing blockages to nets that attempt to cross them. In [4.9], the decap budgeting problem, using both CMOS and MIM DECAPs, is formulated as a linear programming (LP) problem, and an efficient congestion-aware algorithm is proposed to optimize the power supply noise while trying to find a balance between the routing congestion deterioration and leakage power increase.

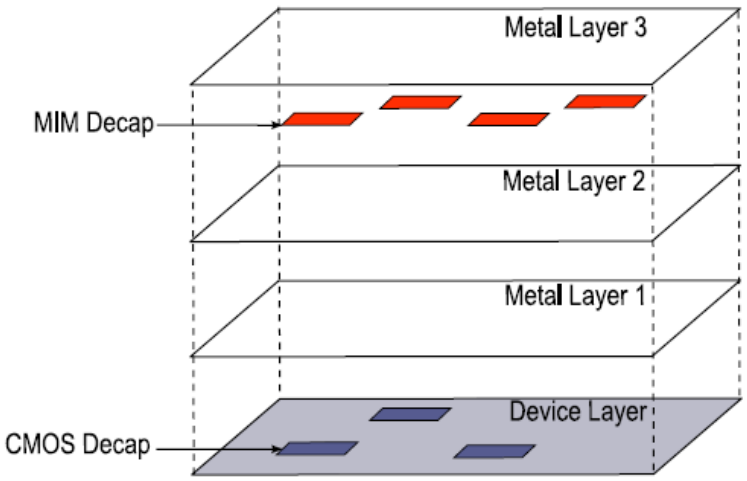


Fig. 4.3 MIM and CMOS DECAPs in one 2D tier with three metal layers.

An iterative flow is used to solve the decap allocation problem. In each iteration a relatively small amount of decap is allocated to the current circuit for two reasons. First, the decap allocation problem is highly nonlinear, and this iterative approach permits the optimization process to be controlled by solving a sequence of linear

programs, one in each iteration. In order to enable the formulation of these linear programs, it is necessary to model the noise violation and the congestion using models that are linear in the decap value. The second reason is related to this approximation: it avoids the excessive allocation of DECAPs that could invalidate the approximate linear model of congestion and noise violation used in the algorithm; these models are predicated on the assumption of small perturbations.

Experimental results demonstrate that the use of CMOS DECAPs alone is insufficient to overcome the violations; the use of MIM DECAPs results in high levels of congestion; and the optimal mix of the two meets both congestion and noise constraints with low leakage.

4.2 On-Die Noise Suppression with Active Circuits

The traditional solution for reducing power supply noise is to use on-chip decoupling capacitors, along with on-package and on-board capacitors to supply instantaneous current demand. However, the large consumption of die area and gate leakage has limited the total amount of DECAPs that can be deployed on a chip for noise reduction. In this section, we will introduce the noise suppression mechanical with active circuits which is more efficiency to reduce to noise on power delivery network.

4.2.1 On-Die Noise Suppression by Improving Impedance of Power Grid

The lower the resistance, the lower the overshoots; however that comes with an increased power consumption trade-off. Adding a conventional resistor in parallel to the power grid network would increase the power consumption significantly. In reality, damping is only required in the frequency domain around the resonance frequency,

rather than at all frequencies as provided by a conventional resistor. An active resistor in parallel with the on-chip decoupling capacitors [4.10] for damping has been presented. And achieving 40% peak noise reduction for overshoots and 15% reduction for undershoot. However, the increased IR droop limits the usefulness of this technique.

Active DECAPs were previously used for the suppression of substrate crosstalk in mixed-signal ICs [4.11]. Fig. 4.3 shows the principle of our proposed active DECAP circuit that is targeted towards the suppression of supply noise due to finite on-chip and off-chip parasitic impedances in digital ICs. The proposed circuit consists of an operational amplifier (Opamp) and a passive load capacitor C_{load}. The feedback loop detects changes in the voltage VDD-Gnd and drives the load capacitance, effectively amplifying the capacitance seen at the VDD input to $(1+A(\omega))\cdot C_{load}$ via the Miller effect, where $A(\omega)$ represents the gain of the opamp. In [4.12] and [4.13], the same idea but the new active DECAP circuit is proposed. The proposed circuit is designed suitable for digital implementation by using self-biasing schemes and has maximized its output swing to obtain sufficient noise regulation range. Simulations are performed to exam the DECAP boosting performance under various capacitive load conditions. Results show that the proposed circuit not only can boost the DECAP value by more than 10 but also exhibits significant advantage of suppressing the dominant resonant noise in an IC chip.

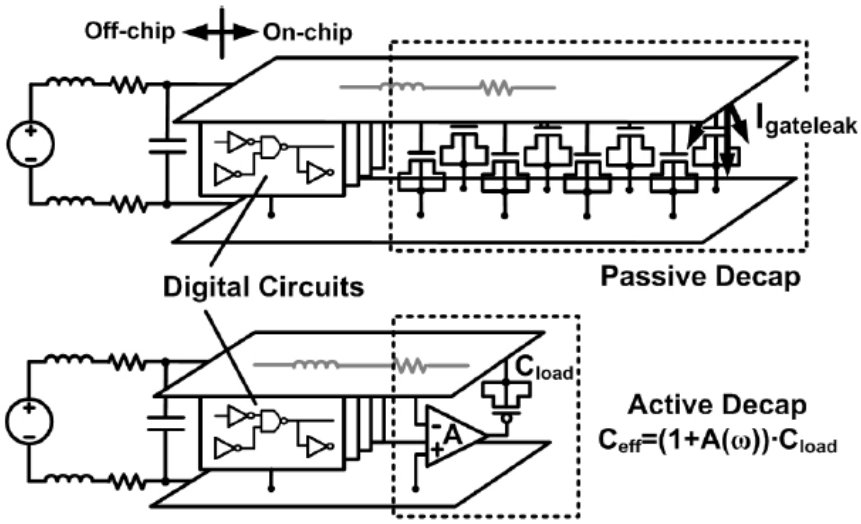


Fig. 4.4 Principle of proposed active DECAP circuit for supply noise suppression.

4.2.2 On-Die Noise Suppression by Providing Additional Current

Unlike the section 4.2.2, we introduce the supply noise suppression technology by proving additional current. In [4.14] and [4.15], an active circuit to detect and suppress excessive supply-voltage undershoots and overshoots caused by large current transients or by excitation of supply resonance. A nominal-voltage active supply, V_{DDA}, is used to inject extra charge into the power grid during excessive undershoots. The use of a nominal-voltage V_{DDA} eliminates the need for any high-voltage supplies and enables use of DECAPs and transistors with nominal oxide thickness, which is shown as Fig. 4.5. A similar circuit is proposed in [4.16], Fig. 4.6 shows the schematic of the noise suppression. When the logic circuit wakes up, the switch between V_{DDH} and V_{DD} is turned on by the level-shifter and the current from V_{DDH} substitutes the current flowing through the bonding wire and the onboard supply lines of V_{DD}. But this technical has a disadvantage that is trigger by the switching circuit rather than detecting the power supply noise.

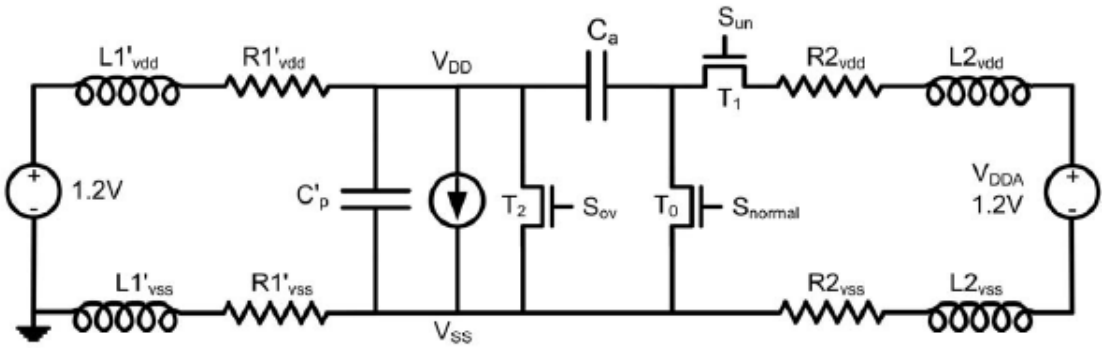


Fig. 4.5 Power delivery model with active noise suppression.

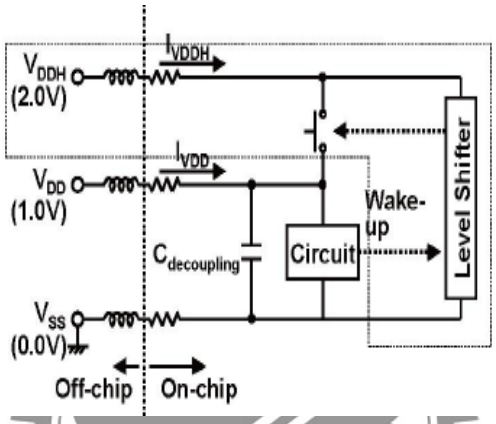


Fig. 4.6 (a) Configuration of the noise suppression

Both the technical we mention above need an additional power supply, Jianping presents an on-die resonance suppression circuit technique that uses band-limited active damping to reduce resonance-induced voltage fluctuations [4.17] and [4.18]. Fig. 4.7 shows an on-die resonance-suppression circuit (RSC) technique that uses band-limited active damping to reduce resonance-induced voltage fluctuations. The RSC consists of a supply noise amplifier with an integrated band-pass filter that lowers RSC sensitivity to out-of-band supply noise. The amplified supply noise in the pass band feeds a comparator that drives a current generator, which produces the damping current. When the load current has a large component at the resonant frequency, the supply voltage variation is 180° phase-shifted compared to the load current as a rising load current induces a falling

supply voltage. The RSC monitors the supply voltage and generates an AC current in phase with the voltage fluctuations, or 180° out of phase with the in-band load current, effectively clamping the voltage overshoots.

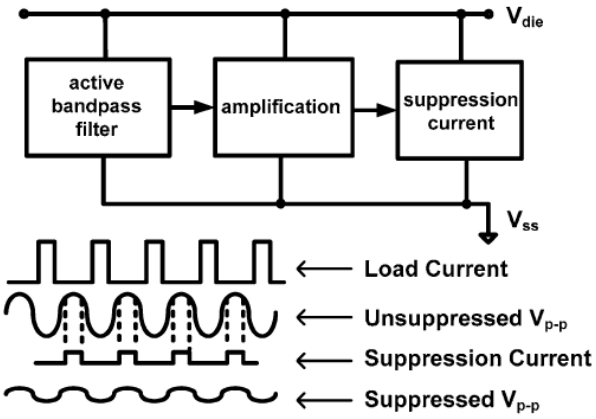


Fig. 4.7 block diagram of RSC

4.2.3 On-Die Noise Suppression by switching DECAP

We mention the noise suppression technical by proving additional current in section 4.2.2. However, it is not a good idea to provide addition current to reduce the power supply noise in power efficiency view.

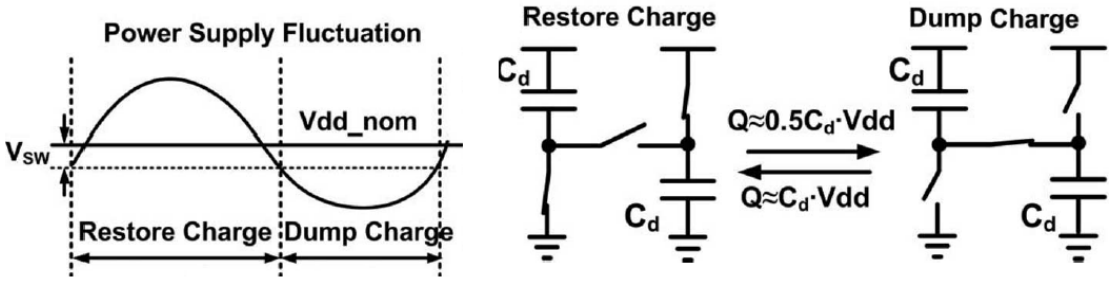


Fig. 4.8 Principle of resonant noise suppression using switched DECAPs

The Fig. 4.8 shows the principle of using switched DECAPs to boost the amount of charge that is delivered by the conventional DECAPs. Two passive DECAPs are connected in parallel during normal condition and serve as conventional DECAPs. When supply noises undershoot reaches a switching trigger threshold $V_{SW,}$ the

DECAPs are switched into a series connection where charge is dumped into the supply network. In the supply overshoot cycle, the capacitors are switched back into parallel and charge is restored to the capacitors from the supply network.

Fig. 4.9 illustrates the complete active DECAP design containing four blocks: a reference voltage generator, a pair of high-pass filters, two comparators, and the switched DECAPs in [4.19] and [4.20]. The user logic circuit block shown in the figure is considered to be the main cause of power supply noise violation. This active DECAP using latch-based comparators in 90 nm CMOS is able to switch in 0.5 ns and consumes a relatively low power of 2.8mW, which is about 5X lower than a previous design running at approximately the same speed. Even more, the author investigate these higher stack height configurations to assess the degree of improvement in supply noise reduction [4.21]. Extensive simulation results correlated with a test chip indicate that an active DECAP with a stack height of three provides the best noise reduction if the supply noise level is between 7%-14%, but a stack height of two is best if the noise level is between 14%-16%.

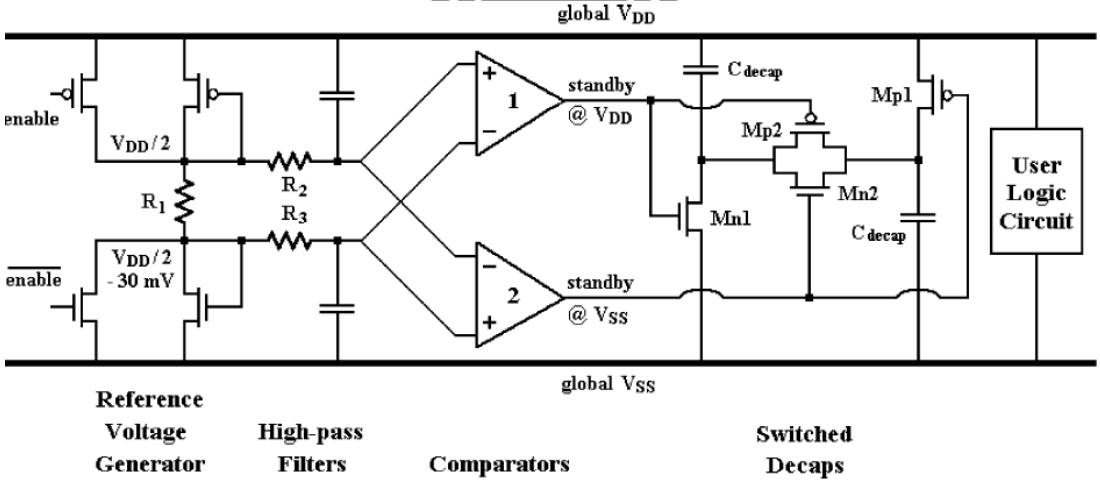


Fig. 4.9 Active DECAP architecture

On the contrary, Fig. 4.10 shows the schematic of the switched DECAP regulator with a digital resonant detection scheme which is proposed in [4.22]. The noise

detection is realized by comparing the delay of a constant delay line (CDL) and a variable delay line (VDL). The CDL forms a ring oscillator with a frequency of 2 GHz to continuously trigger the comparison. The supply of the CDL (Vdd') is low-pass filtered so that the delay is insensitive to supply noise above 10 MHz which is the low cut-off frequency of our resonant regulation. The supply of VDL is directly connected to the noisy Vdd so its delay varies with supply fluctuations. Measurements from a 0.13 m test chip confirm an effective DECAP boost up to 11X which translates into an 89% saving of passive DECAP area. Owing to the digital implementation, the proposed circuit consumes 91% less quiescent power with a better tolerance to PVT variation compared with the previous analog scheme.

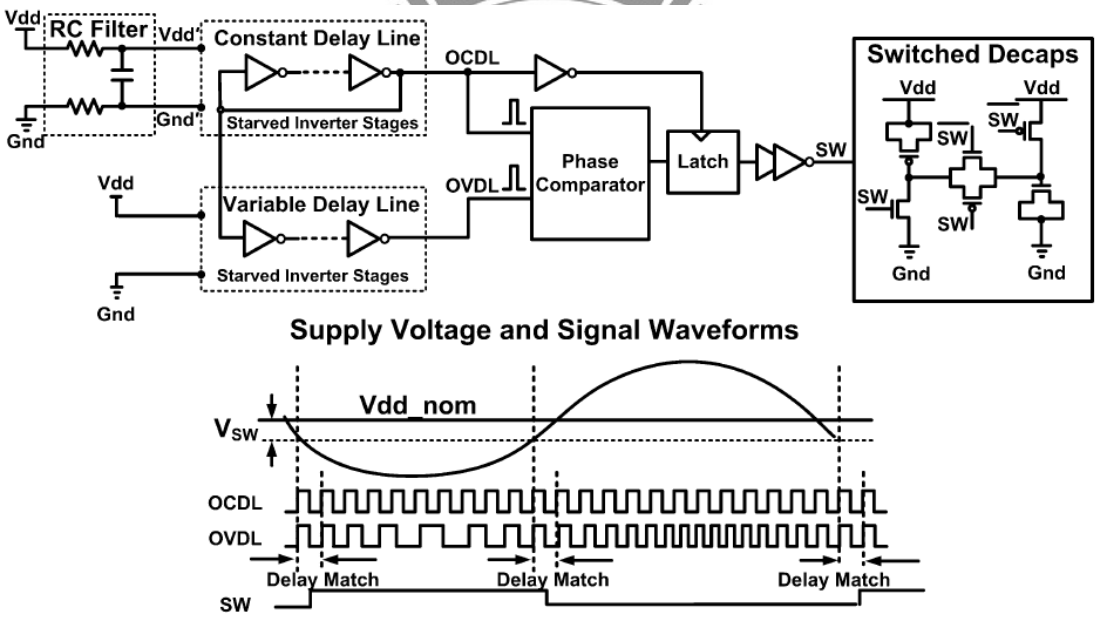


Fig. 4.10 Schematic of the proposed switched DECAP circuit with digital resonant detection

4.3 Proposed Active Decoupling Capacitor for Supply Noise Regulation for TSV 3D Integration

Fig. 4.11 shows the power integrity for the TSV 3D integration. It is shown that heavy

current density of through-silicon-via (TSVs) and packages exists in the power network and further increases the power supply noise. The supply current per package pin in 3D designs is significantly higher than in 2D designs. Therefore, the power supply noise problem, already a major issue in 2D, is even more severe in 3D chips. Moreover, the supply impedance response is dominated by both the packages and TSVs. In view of these, noise suppression will become one of the critical design problems for TSV 3D integration.

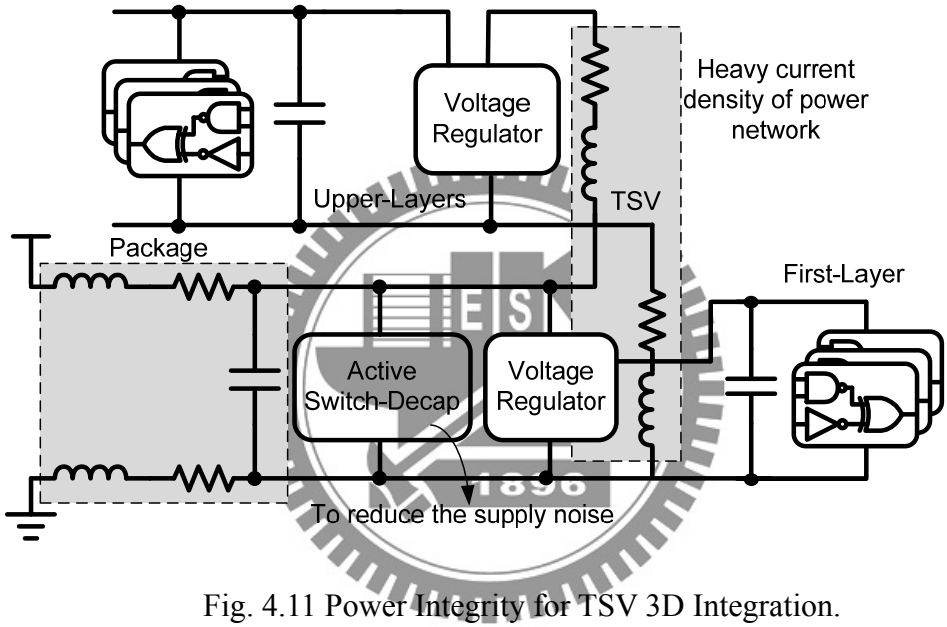


Fig. 4.11 Power Integrity for TSV 3D Integration.

To suppress the power noise, decoupling capacitors (DECAPs) are widely used. DECAPs perform as a local reservoir of charge, which is released when the current load varies. Since the inductance of packages scales slowly, the DECAPs significantly affect the design of the power/ground (P/G) networks in high performance ICs and TSV 3D integration. At higher frequencies, DECAPs are distributed on chips to effectively manage the power supply noise. However, the usage of the on-chip passive DECAPs is limited by two major constraints, including a great amount of gate tunneling leakage and large area occupation. Therefore, current suppression techniques have been proposed to reduce power supply noise, and the resonant supply

noise is suppressed via the delay-line-based and OP-based detection circuits with switched DECAPs, respectively. However, the efficiency of these noise suppression techniques would be reduced significantly by the leakage current in nano-scale technologies. In this paper, a noise suppression technique is proposed for TSV 3D integration based on UMC 65nm CMOS technology. This noise suppression technique reduces the supply noise using a latch-based comparator and switched DECAPs.

4.3.1 Power Noise Suppression for 3D Integration

Depending on the heavy current loading of P/G networks in 3D integration, the supply noise is a serious problem for power integrity. In view of this, Fig. 4.12 shows the proposed architecture of the noise suppression technique to reduce the supply noise. This architecture contains four blocks: a low pass filter, a latch-based comparator, a charge pump, and switched DECAPs. The prior three blocks are designed to detect the resonant supply noise and to control the switches of the switched DECAPs. The details of each block are described as follows.

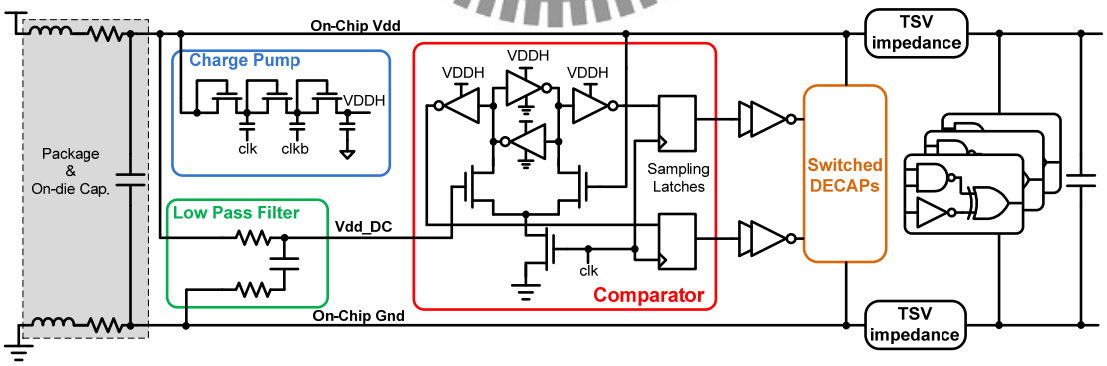


Fig. 4.12 Architecture of the Noise Suppression Technique.

A. Switched DECAPs

The switched DECAPs are designed to suppress the resonant supply noise. Fig. 4.13 illustrates the resonant noise suppression using switched DECAPs. If the power

supply is overshooting than the Vdd_DC, excess charge would be transferred to the capacitors, C_{d1} and C_{d2}, respectively. On the contrary, if the power supply is undershooting than the Vdd_DC, the DECAPs would be connected in series. And thus, the boosted voltage would be twice Vdd_DC. The additional charge can be provided for the power supply from DECAPs to reduce the supply noise. Additionally, the hysteresis voltage levels are the high/low boundary conditions for switching the DECAPS from series to parallel and parallel to series, respectively. The switched DECAPs would be switched only when the supply voltage is higher or lower than the hysteresis voltage levels. In other words, the hysteresis voltage provides a tolerant interval and avoids frequent switches with a small supply noise.

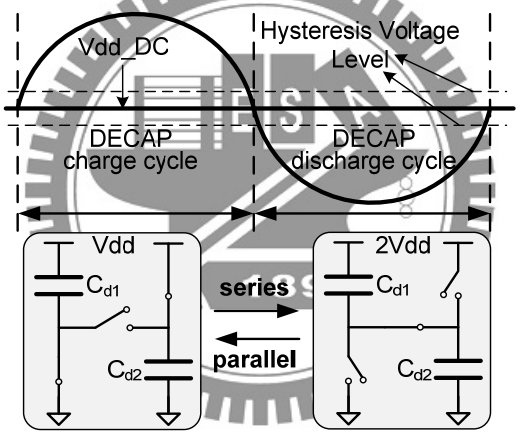


Fig. 4.13 Resonant noise suppression using switched DECAPs

B. Low pass filter

The RC low pass filter provides the reference voltage for the latched-based comparator. However, the current through the resistor of the RC filter induces the IR drop and further decreases the reference voltage. The current contains the leakage current of the capacitance and the load current of the comparator. Therefore, the demand current of the comparator should be small to generate a low-drop reference voltage.

C. Latch-based Comparator

In order to switch the DECAPs, a latch-base comparator is designed to detect the resonant noise. The supply voltage (Vdd) and the reference voltage (Vdd_DC) are compared via the latch-based comparator. This comparator achieves not only good noise tolerant interval but low power consumption. Fig. 4.14 shows the schematic of the latch-based comparator via High-Vt transistor to reduce the leakage power in nano-scale technologies. The operation voltage of this latch-based is higher than the on-chip supply voltage to detect the resonant supply noise. Additionally, the demand current of the comparator is very small because the reference voltage (Vdd_DC) is connected to the gate of M1.

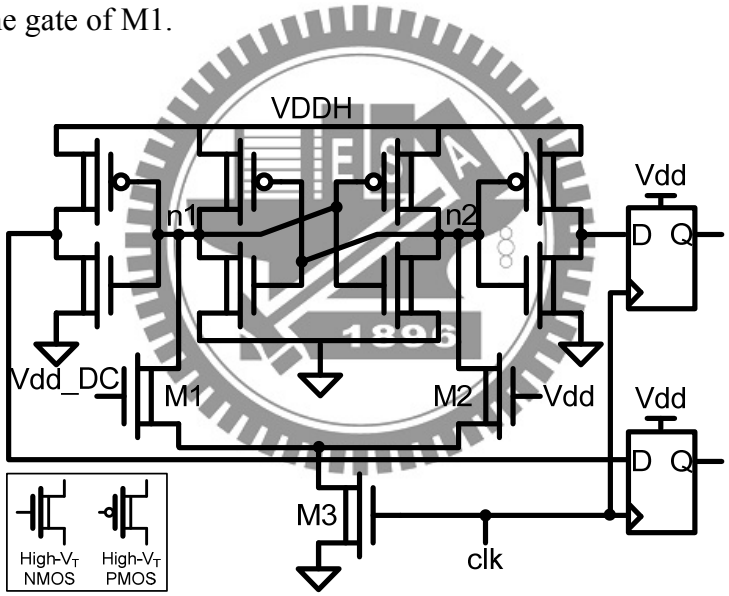


Fig. 4.14 Latch-based comparator with High-Vt MOS

The frequency of the clock determines the sampling rate of the comparison results. When the clock is high, two discharging paths exist to pull down n1 and n2. After a half clock cycle, M3 would be turned off by the low level of the clock. And thus, the discharging paths would disappear. Therefore, the data in the weak back-to-back invertors would be determined according to the charge in n1 and n2. Additionally, two sampling latches capture the comparison results at the positive edge of the clock.

The latched-based comparator switches the switched DECAPs with a hysteresis voltage as shown as Fig. 4.15. Assume Vdd is increasing from the undershooting state to the overshooting state, and the initial voltage of n2 is high. While the vdd is a little larger than the vdd_DC, the discharging time is not enough to flip the data in the back-to-back inverter although the drain current of M2 is larger than that of M1. Until the Vdd is larger than the reference voltage plus a hysteresis voltage, M2 has enough driving ability to flip the data. Therefore, the data in the back-to-back inverter would be changed and the DECAPs would be switched into the series stack. If Vdd is decreasing from the overshooting state to the undershooting state, the switched DECAPs would be changed to the parallel mode as well as the Vdd is smaller than the reference voltage minus a hysteresis voltage.

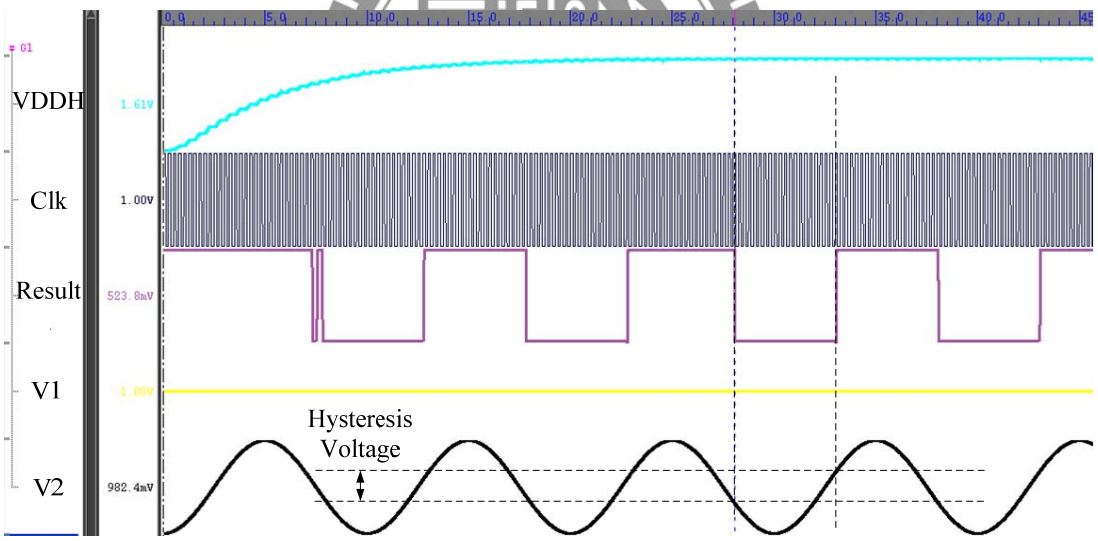


Fig. 4.15 Simulation result of latch-based comparator

D. Charge pump with improving body effect

For the latched-based comparactor, an additional higher voltage (VDDH) is applied to ensure that the transistors, M1 and M2, are operated in the saturation region. Additionally, the hysteresis voltage level is influenced by the level of VDDH and the

widths of M1 and M2. With the increasing of VDDH, both the hysteresis voltage and power consumption increase. Therefore, a modified Dickson charge pump is designed to pump the VDDH to 1.6V. Fig. 4.16 shows the modifed circuits from the Dickson charge pump [4.23]. The bodies of MN1, MN2 and MN3 are connected to their drains to adjust the threshold voltage and further increase the efficiency of the charge pump.

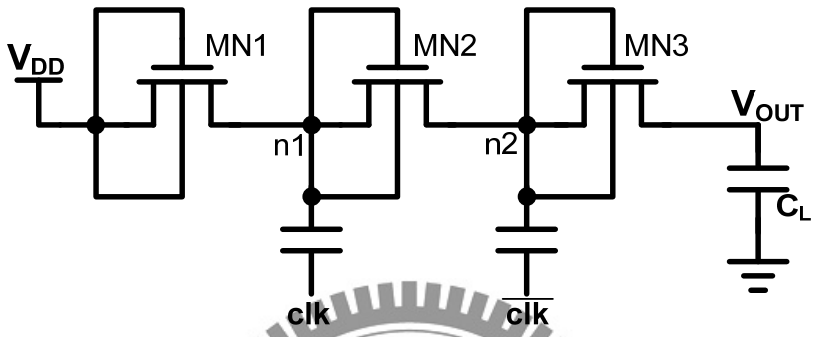


Fig. 4.16 Modified Dickson charge pump.

When the clock is low, the node n1 is charged to Vdd-$V_{tn1\text{-}low}$, where $V_{tn1\text{-}low}$ is shown as Eq. (4.1). Therefore, $V_{tn1\text{-}low}$ is smaller than the threshold voltage when the body is connected to ground.

$$V_{tn1-low} = V_{t0} + \gamma\left[\sqrt{(-V_{DD}) + 2\varphi_f} - \sqrt{2\varphi_f}\right] \quad (4.1)$$

When CLK is high, the node n1 is pumped to 2Vdd-$V_{tn1\text{-}low}$. And thus, the threshold voltage of MN1 would be adjusted as Eq. (4.2) which is smaller than the threshold voltage when the body is connected to ground. In view of this, the leakage current from n1 to Vdd is reduced. In the following stages of the charge pump, the threshold voltage of NMOS in each stage is the same as V_{tn1}. The body bias of NMOSs can both improve the pumping efficieny and reduce the leakage current by adjusting the threshold voltage.

$$V_{tn1-high} = V_{t0} + \gamma\left[\sqrt{(V_{DD} - V_{tn-low} + 2\varphi_f} - \sqrt{2\varphi_f}\right] \quad (4.2)$$

The comparison of modified charge pumpwith previous work is shown in Fig. 4.17. The simulation results show that the modified chrage pump has better loading capability and power effiency than others.

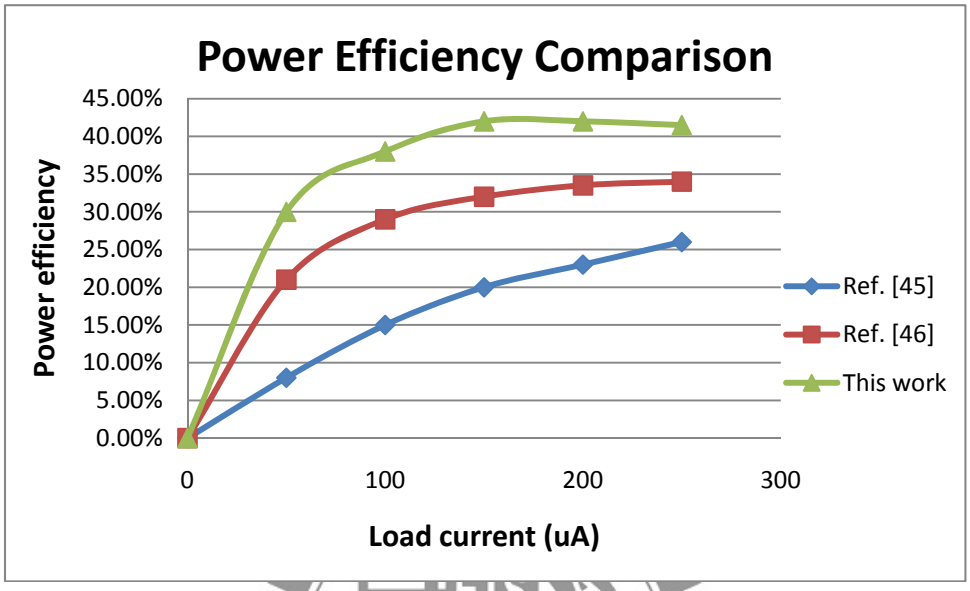


Fig. 4.17 Simulation Result of modified charge pump.

4.3.2 Simulation Result

The noise suppression technique with low power active DECAPs is implemented via UMC 65nm CMOS technology and the TSV model [4.24]. According to the speed of the resonant noise, the frequency of the comparator is 2GHz. This clock source is also provided to the charge pump at 1GHz by a frequency divider. Fig. 6 shows the layout view and the floorplan of the noise suppression circuit. The total value of DECAPs is 200pF. Moreover, 84% area is occupied by the switched DECAPs which are implemented by MOS capacitors. The size of the proposed noise suppression circuit is 170x230 μm^2.

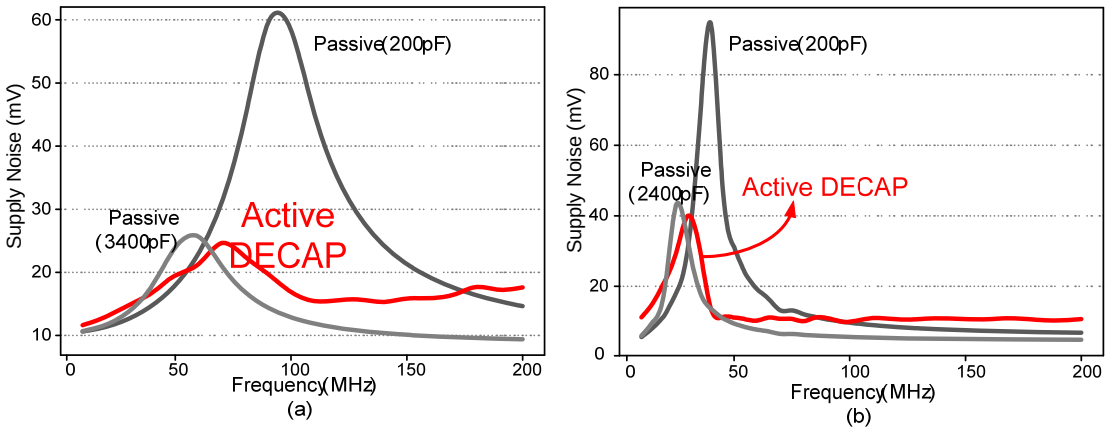


Fig. 4.18 Noise suppressions of the active and passive DECAPs for (a) high performance IC (b) TSV 3D integration.

Fig. 4.18(a) and (b) show the suppressed noises resonated at 100MHz and 40MHz, respectively. The two configurations of the pads are set as L=0.75nH, C=1.69nF, R=0.14Ω and L=4.5nH, C=1.69nF, R=0.28Ω, which represent a typical supply impedance for high performance chips [4.22] and the footprint and TSV impedance in 3D integration. Since the number of the power pins would be limited in 3D integration, and the inductance of the package would be larger than the 2D-ICs. Compared to the same value of the passive capacitance, the active DECAP can realize the improvements of 55% (6.9dB) and 57.6% (7.4dB) noise reductions for the high performance IC and TSV 3D integration, respectively. Additionally, in order to evaluate the boost factor of the proposed active DECAP, a great amount of passive DECAPs are traced to achieve the similar noise suppression. Therefore, the passive DECAPs with 3400 pF and 2400 pF are deployed for the similar noise regulation. And thus, 17X and 12X boost factor can be achieved based on the proposed noise suppression circuit. Moreover, the leakage power can also be reduced by 71% and 59% due to the reduced DECAP area.

Table 4.2 lists the comparisons between the proposed active DECAP and other approaches [4.20] and [4.22]. The hysteresis voltage for the latch-based comparator is

17mV to avoid the continuous switches if the supply noise is small. The range of hysteresis is between 9mV to 31mV across different process corners from -50~125°C. The delay-line- based active DECAP uses two delay line with biasing starved inverters to compare the supply noise. However, the delay line using the bias scheme is too sensitive when the noise is large. The comparisons would be wrong since the difference between the two delay lines is more than one clock cycle. Moreover, this approach would face leakage problems when shrinking to nano-scale technologies. The static power of the proposed scheme is 0.55 mW. In the latch-based comparator, the demand current from the RC filter is small because the Vdd_DC is connected to the gate of M1 as shown in Fig. 4.14. For the delay-line-based comparator, the current of the constant delay line flows through the resistor and induces 40mV drop. It not only affects the switching timing but decreases the boost factor. Therefore, the proposed scheme can realize not only good noise tolerant interval but low power consumption.

Table 4.2 Comparisons of active DECAPs

	Delay Line *[JSSC'09][4.22]*	*Analog OP* *[JSSC'09][4.20]*	*Latch-based Comparator*
Technology	0.13 μm	90 nm	65 nm
Static Power	0.65 mW	2.9 mW	0.55 mW
Hysteresis Voltage	0mV	5mV	17mV
Triggered Voltage	40mV	50mV	17mV

The noise suppression technique with low power active DECAPs is implemented via UMC 65nm CMOS technology and the TSV model. According to the speed of the resonant noise, the frequency of the comparator is 2GHz. This clock source is also provided to the charge pump at 1GHz by a frequency divider. Fig. 4.19 shows the

layout view and the floorplan of the noise suppression circuit. The total value of DECAPs is 200pF. Moreover, 84% area is occupied by the switched DECAPs which are implemented by MOS capacitors. The size of the proposed noise suppression circuit is 170x230 μm^2.

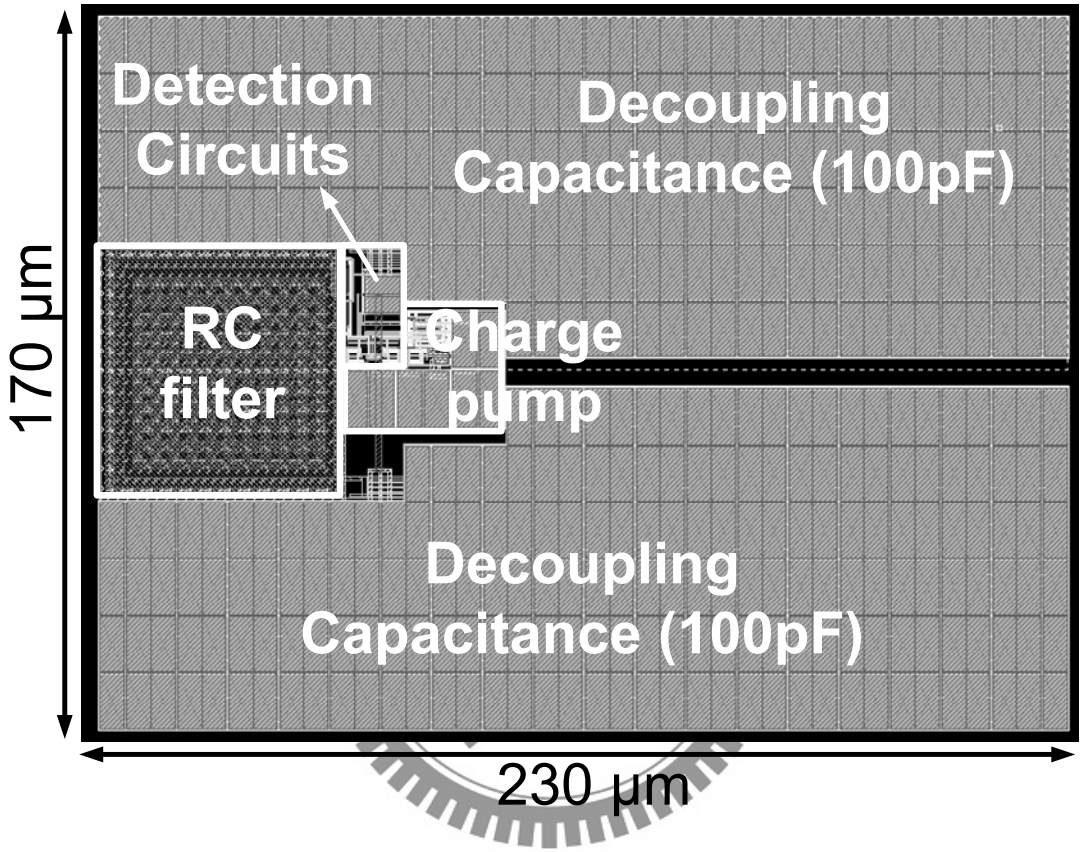


Fig. 4.19 Layout view of the noise suppression circuit

Chapter 5
Design Methodology of TSV 3D Integration Power Delivery System

In this chapter, we start from substrate noise canceller scheme. The low drop voltage regulator (LDO) is proposed in Section 5.2. In order to design a robust power system, we proposed power delivery system in TSV 3D integration. Furthermore, the layout is implemented in 65nm CMOS technology and TSV prediction model and post-simulation is done in Section 5.3.

5.1 Substrate Noise canceller

The main cause of substrate noise is a coupling from digital ground lines. At higher frequency, ground bounce is affected by inductance of power/ground line strongly. Substrate noise is proportional to di/dt. Generally guard rings are used to suppress substrate noise. However, the parasitic impedance of guard line prevents the effect of noise absorption in the case of high speed LSIs. Another method is an active noise canceling technique, and this technique uses operational amplifiers to create a compensation signal shifted in phase by 180° .

The detail structure of substrate noise canceller [5.1] is shown as Fig. 5.1, A power supply current of the internal circuit goes through parasitic inductance L_1 of the power supply line. A pickup inductance L_2 of the di/dt detector coupled to L_1 with a coupling coefficient K induces a di/dt proportional voltage. A noise tolerant

amplifier amplifies the induced voltage and generates anti-phase di/dt proportional current. This di/dt detector can be applicable to cancelling the substrate noise by injecting anti-phase noise current into substrate.

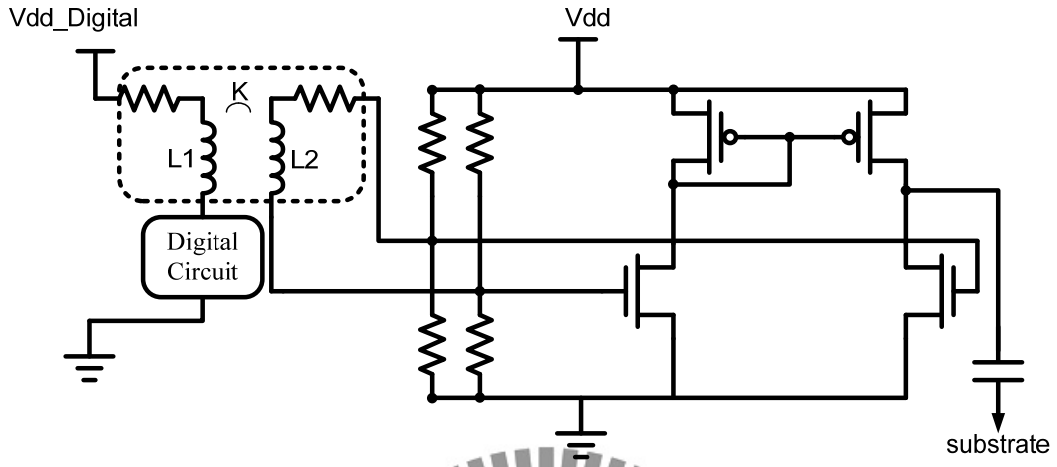


Fig. 5.1 Diagram of the di/dt substrate canceller

Since the substrate is tied to the ground voltage, the coupling capacitor Cc is inserted at the output of the amplifier to prevent the bias voltage change of the node n2. Here, the input impedance of the substrate from the injection point is assumed to be pure resistive. In order to inject the current with the appropriate phase, the impedance of the coupling capacitor Cc should be small enough compared with the substrate resistor impedance. A standard block diagram of feedforward active substrate noise cancelling system is shown in Fig. 5.2. If the complete anti-phase noise signal is injected into substrate, the original substrate noise can be canceled out.

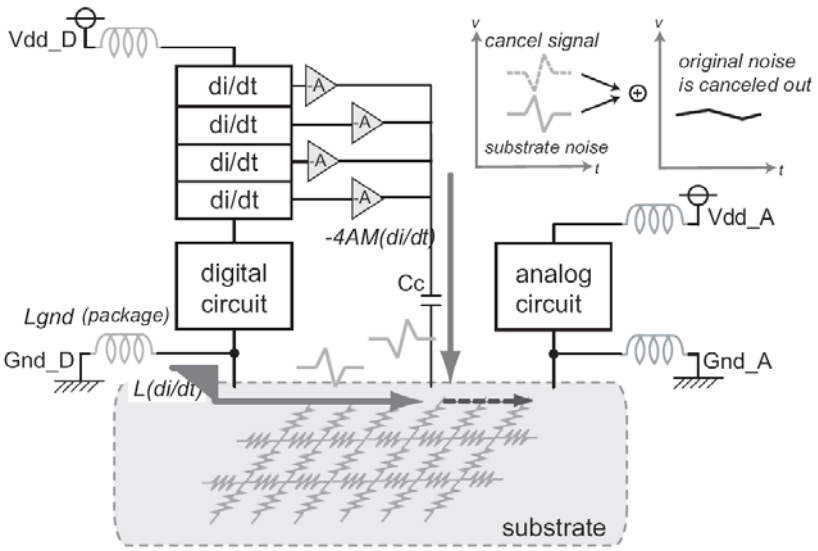


Fig. 5.2 Design diagram of active substrate noise canceller

5.2 Voltage Regulator

5.2.1 Stability Scheme of Voltage Regulator

A typical structure of a low-dropout regulator shows in Fig. 5.3 which consists of an error amplifier comparing the output voltage to the reference voltage V_{ref}, a PMOS pass transistor M_p, and the output buffer stage driving M_p. There are three different poles in the voltage regulator structure located at the output node of the error amplifier (N1), the output node of the buffer (N2), and the output node of the voltage regulator (V_{out}). In particular, these poles are given by

$$P_1|N_1 = \frac{1}{r_{o1}C_1} \tag{5.1}$$

$$P_2|N_2 = \frac{1}{r_{ob}C_p} \tag{5.2}$$

$$P_o|N_{out} = \frac{1}{r_{oeq}C_L} \tag{5.3}$$

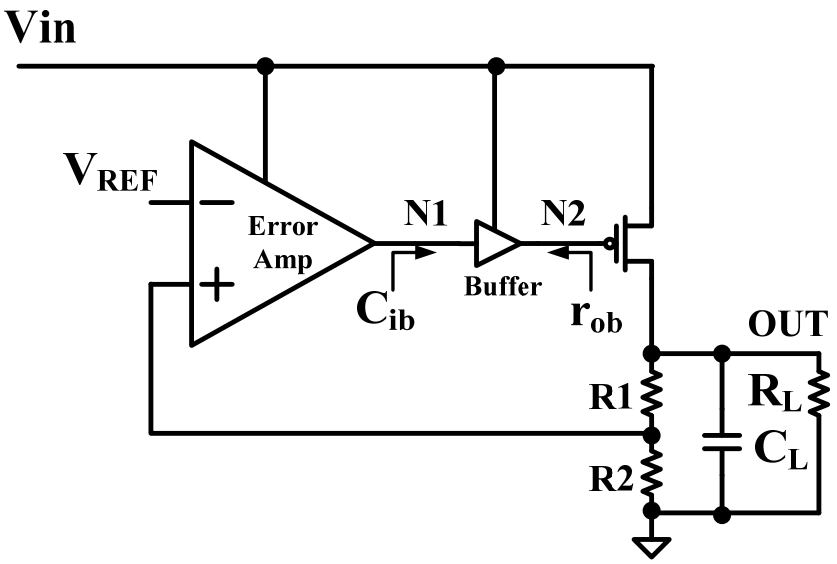


Fig. 5.3 Typical structure of a low-dropout regulator with an intermediate buffer stage.

The ro1 is the output resistance of the error amplifier, C1 is the equivalent capacitance at N1 which is dominated by the input capacitance of the buffer Cib, rob, is the output resistance of the buffer, Cp is the input capacitance of Mp, and roeq is the equivalent resistance seen at the output of the voltage regulator. Ideally, both Cib and rob should be very small in order to achieve single-pole loop response by locating both p1 and p2 at frequencies much higher than the unity-gain frequency of the regulation loop.

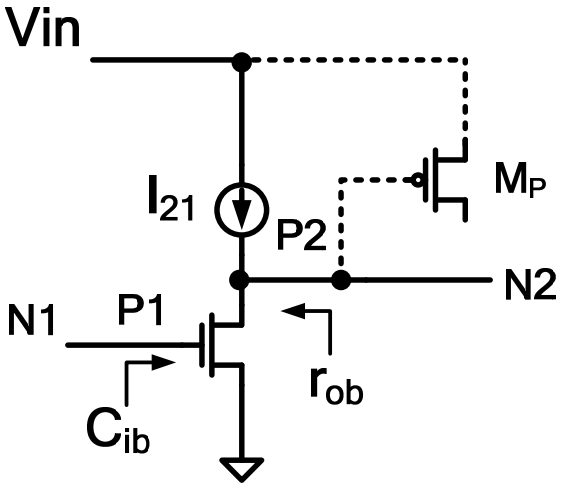


Fig. 5.4 Source-follower implementation of the intermediate buffer stage.

In order to construct the required output buffer stage, a simple PMOS source-follower is first considered for implementing the output buffer and its structure is shown in Fig. 5.4 [5.2]. The PMOS source-follower provides near complete shutdown of the pass device when under the light-load conditions. Because of the output resistance rob of the source-follower is given by 1/gm21, it is necessary to increase gm21 in order to decrease the value of rob and allow p2 to be located at frequencies much higher than the unity-gain frequency of the regulation loop. Transconductance gm21 can only be increased either through using a larger W/L ratio of transistor M21, or through increasing the DC biasing current I21 through M21, or both. However, increasing I21 would increase the total quiescent current of the regulator, and the current efficiency of the voltage regulator is degraded. Using a larger W/L ratio of M21 would increase the input capacitance Cib of the buffer, which is in turn pushes p1 to a lower frequency and the stability would be poorly affected. A simple PMOS source-follower need, therefore, to design carefully.

5.2.2 Voltage Regulator and Variable Voltage Reference

In order to decrease the bias current of op amp and stabilize the transient response of regulator simultaneous, we proposed a low quiescent current low drop regulator (LDO). The schematic of proposed voltage regulator is shown in Fig. 5.5.

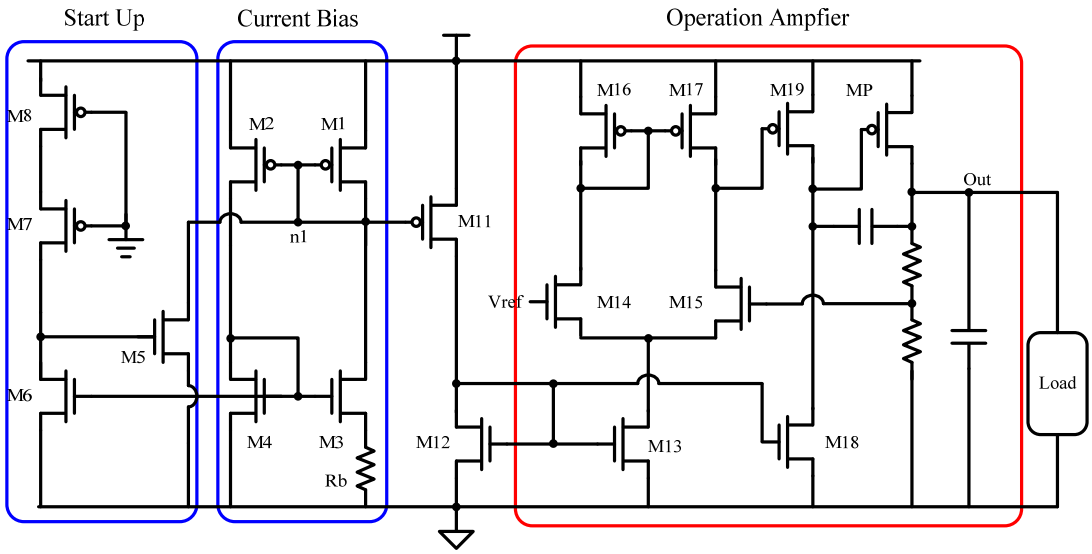


Fig. 5.5 Low quiescent current voltage regulator

The voltage regulator is composed of op amp, a voltage buffer, a power PMOS transistor (M_{POUT}), current bias circuit and a start up circuit. The op amp is composed of M_{12} to M_{17}. The voltage buffer is composed of M_{18} and M_{19}. The self current biased circuit is composed of M_1 to M_4, and it can provide a bias current to M_{11}. Finally, the start-up stage will pull node n1 to the ground level ensure the transistors, from M_1 to M_4, will be in the saturation region.

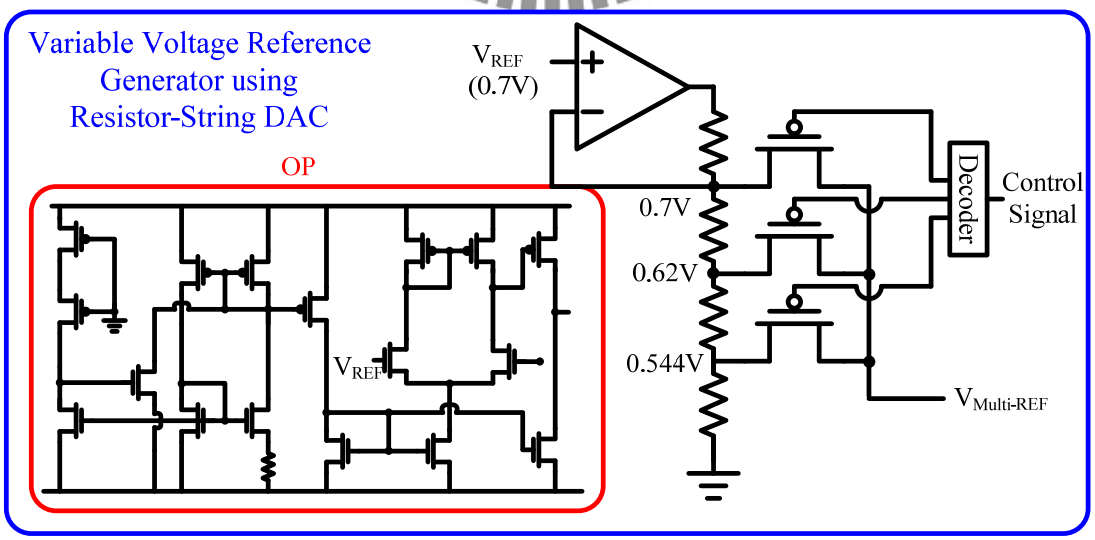


Fig. 5.6 Resistor-string DAC for multi-voltage reference generation.

In order to produce different out voltages, we need different reference voltages. However, use too many Bandgap reference voltage circuit is not an efficiency solution. Hence, the multi-voltage reference (MVR) is generated by resistor-string DAC in Fig. 5.6. The decoder selects the output voltage for variable voltage reference. Fig. 5.7 and Fig. 5.8 show the variable voltage reference with different Vin and temperature. For providing three different output voltages (0.9V, 0.8V and 0.7V), three different voltage references need to be generated, which are 0.7V, 0.62V and 0.54V, respectively. The maximum variation of output voltage is less than 4mV, over a 10% Vin variation range. And the maximum variation of output voltage is less than 3mV in the temperature range from -50℃ to 125 ℃.

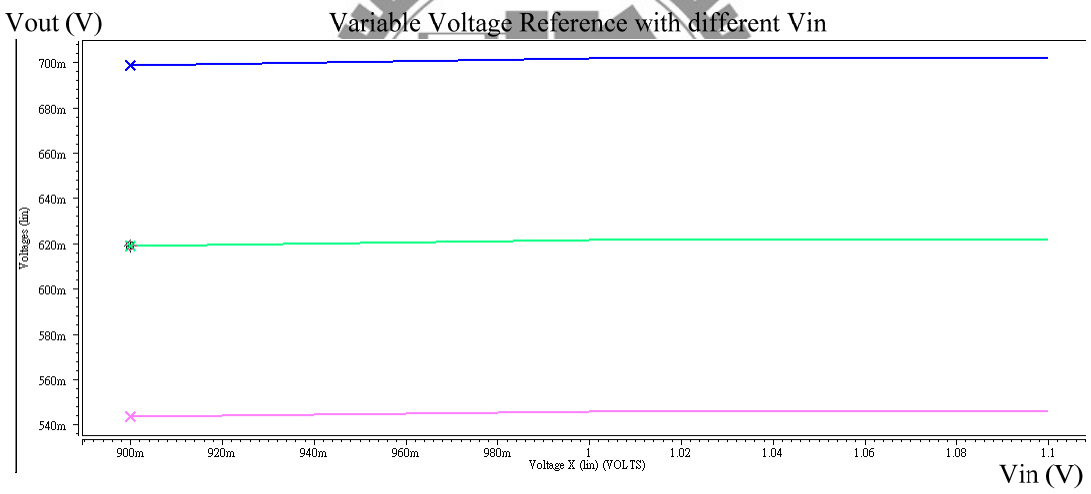


Fig. 5.7 Variable voltage reference with different Vin

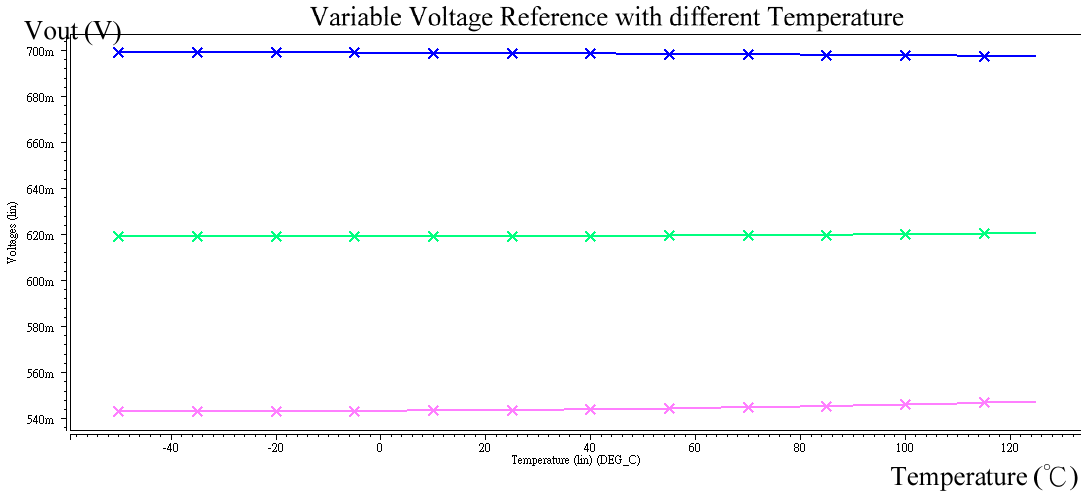


Fig. 5.8 Variable voltage reference with different temperature

The simulation results of low quiescent current low drop regulator (LDO) are shown in Fig. 5.9. When load current is 10mA, the voltage regulator outputs 901.3mV and the quiescent current is 88μA. When load current is 200mA, the output voltage of voltage regulator will drop down and is stable in 897.8mV. The current efficiency is 99.95%.

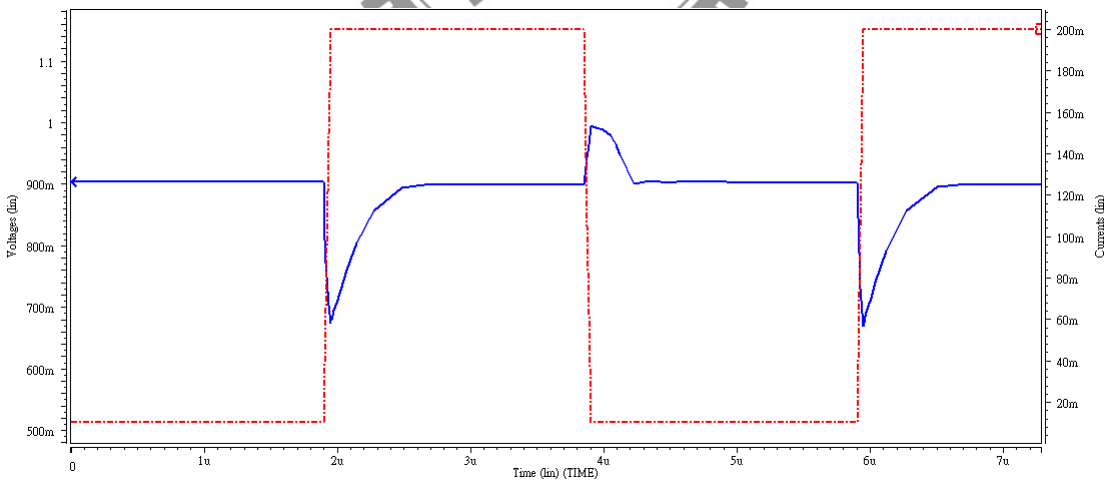


Fig. 5.9 Output voltage of LDO under different load current

To ensure the stability of LDO, the AC simulation is also required. Fig. 5.10 and Fig. 5.11 show the AC simulation of the LDO in light load (10mA) and heavy load (200mA) condition, respectively. The DC gain of the LDO is 51.3dB and 35.3dB,

and the phase margin (PM) is 45∘and 62∘in the light load and heavy load condition, respectively.

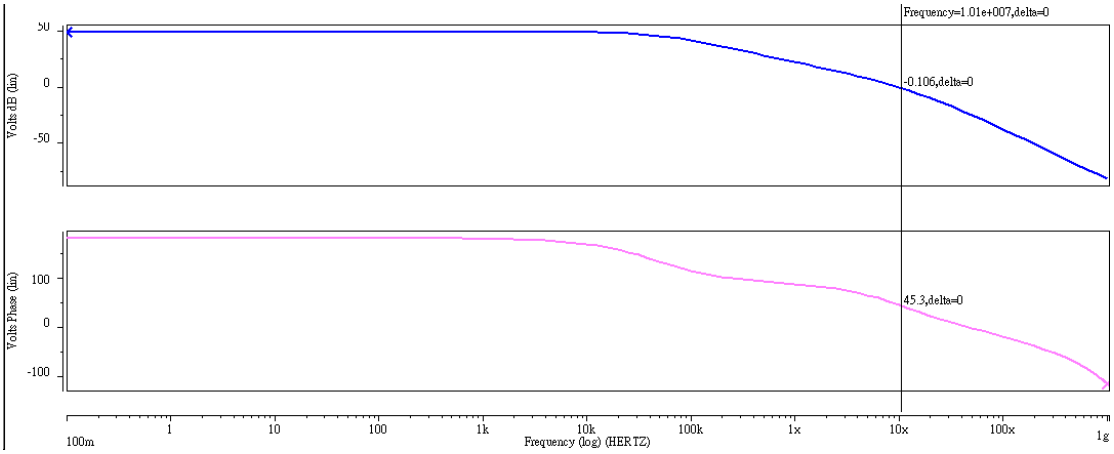


Fig. 5.10 AC simulation of LDO with light load (10mA)

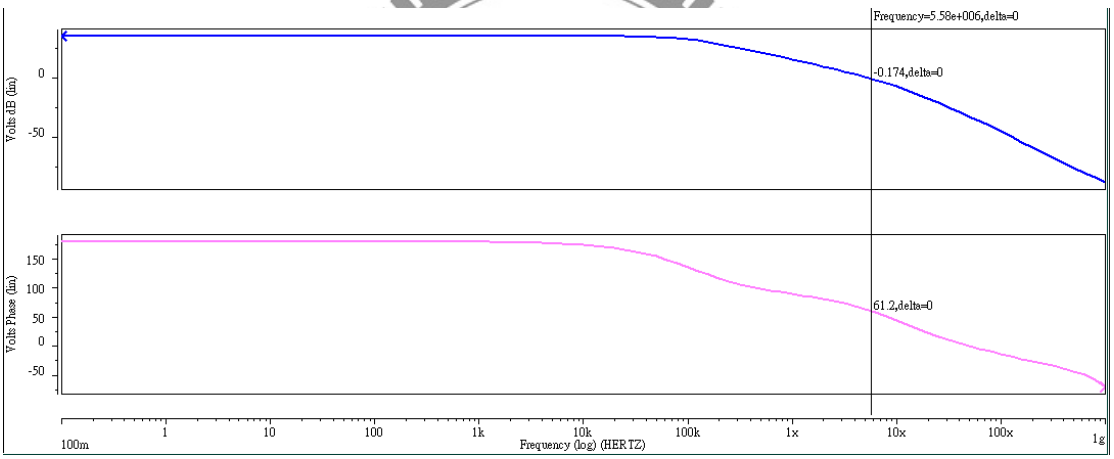


Fig. 5.11 AC simulation of LDO with heavy load (200mA)

The comparison of the d low quiescent current voltage regulator with previous work ([5.3], [5.4] and [5.5]) is shown in Table 5.1. Because our regulator only needs 88μA quiescent current, these LDO has the best figure of merit.

Table 5.1 Comparison of voltage regulator

	Ref.[5.3]	Ref. [5.4]	Ref. [5.5]	This work
Technology (nm)	90	90	90	65
Input Voltage	1.2V	2.4V	1.2V	1V
Output Voltage	0.9V	1.2V	1V	0.9V
Output droop ΔVOUT	90mV	120mV	96mV	203mV
MAX Load Current	100mA	1A	150mA	200mA
IQ (quiescent current)	6mA	25.7mA	1mA	88uA
Current Efficiency	94.30%	97.50%	99%	99.95%
Decoupling Cap.	0.6nF	2.4nF	0.4nF	0.4nF
Response time (T_R)	540ps	288ps	256ps	406ps
FOM (figure of merit)	32ps	7.4ps	1.7ps	179fs

$$Current\ efficiency = \frac{I_{MAX}}{I_{MAX} + I_Q} \qquad T_R = \frac{C * \Delta V_{OUT}}{I_{MAX}} \qquad FOM = T_R \frac{I_Q}{I_{MAX}} = \frac{C * \Delta V_{OUT}}{I_{MAX}} * \frac{I_Q}{I_{MAX}}$$

The linear low drop regulator with low quiescent current is implemented via UMC 65nm CMOS technology. For providing the different supply voltages in each layer, the MVR is also designed in this voltage regulator and it occupy very small area of the LDO. Fig. 5.12 shows the layout view and the floorplan of the voltage regulator. The total value of decoupling capacitor is 400pF. The size of the proposed low quiescent current voltage regulator is 312x326 μm^2.

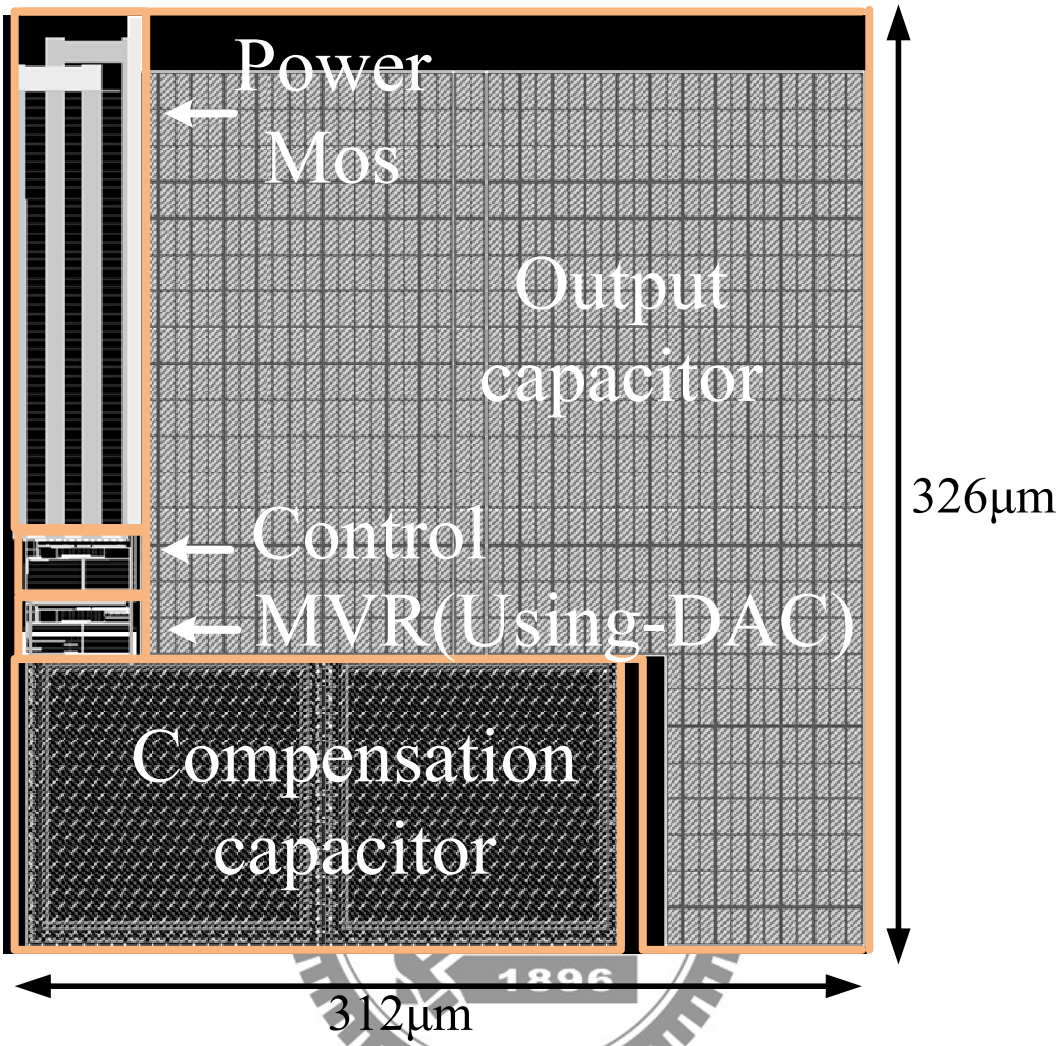


Fig. 5.12 Layout view of low quiescent current voltage regulator

5.3 The Scheme of Power Delivery System in TSV 3D Integration

Three-dimensional (3D) integration technology can provide enormous advantages in achieving multi-functional integration, improving system speed and reducing power consumption for future generations of ICs. However, stacking multiple dies would face a severe problem of the power integrity. In this chapter, we develop a power delivery system in TSV 3D integration, which focus on the noise of the power system in TSV 3D integration, as shown as Fig. 5.13.

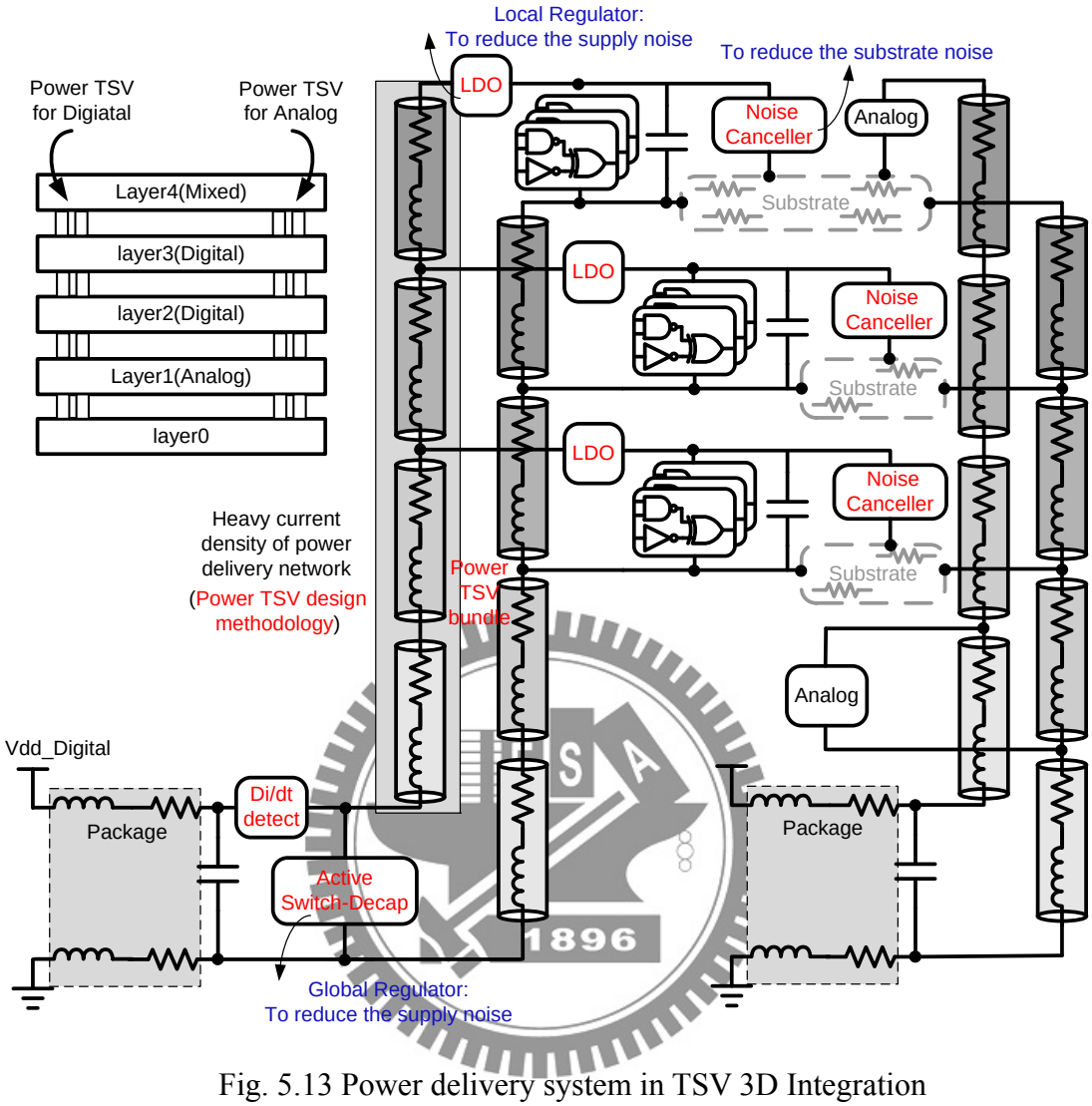


Fig. 5.13 Power delivery system in TSV 3D Integration

First, the proposed active switched-DECAP, which is discussed in chapter 4, is used to reduce the resonant noise caused by pin of package. Second, according to the power TSV design method, designer can choose the suitable parameter of the power TSV. Third, voltage regulator in each layer can not only provide the different supply voltage but also reduce the supply voltage noise cause by the power TSV and package. Finally, substrate is also considered. A substrate noise canceller is used to reduce substrate noise coupling from digital domain to analog domain.

As shown as Fig. 5.14, the impedance of TSV 3D IC is dominant by two sources, packaging and power TSV bundle. We use the power TSV design method which is proposed in chapter 3 to design the power TSV. The impedance to each layer is similar at the resonant frequency caused by packaging. In other words, we can ignore the impedance of power TSV bundle at this frequency. On the other hand, the impedance to each layer is totally different at the resonant frequency caused by power TSV bundle.

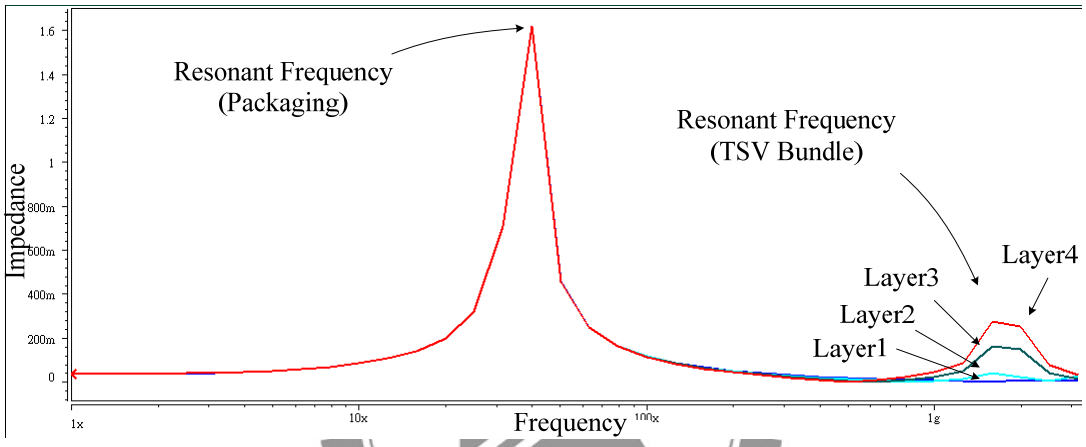


Fig. 5.14 Impedance of TSV 3D integration

In TSV 3D integration, the supply noise would be increased with the increasing current density of the power network. A noise suppression technique for TSV 3D integration is proposed to reduce the supply noise by the latch-based active DECAPs. Based on UMC 65nm CMOS technology and TSV model, the proposed scheme can realize maximum 7.4dB supply noise reduction and 12X boost fact at the resonant frequency. The simulation result is shown as Fig. 5.15, the Layer4_vdd and Layer4_Gnd means the power supply voltage and ground line voltage on the fourth layer, respectively.

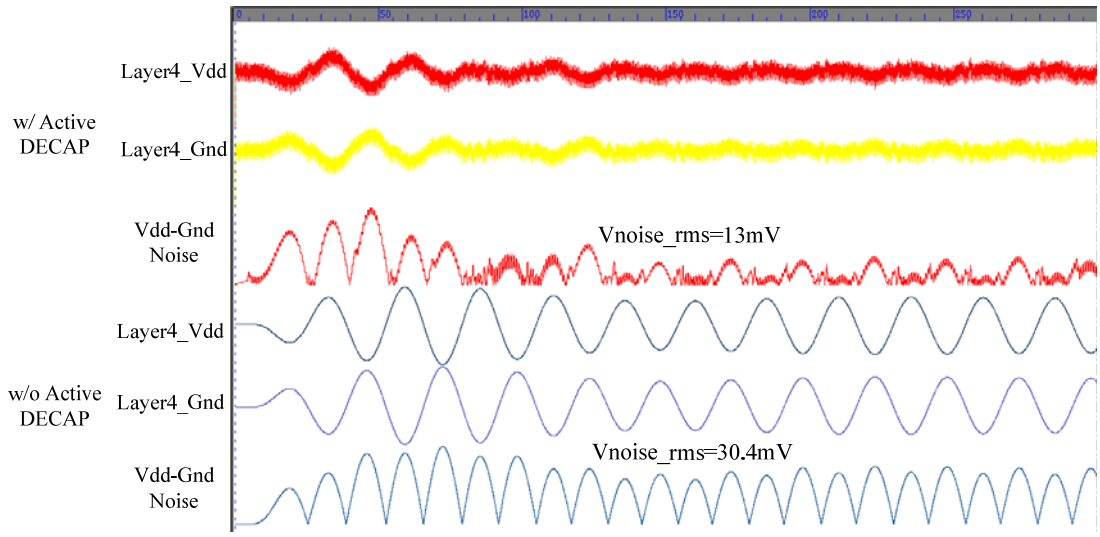


Fig. 5.15 Simulation of power noise on each layer.

The low quiescent current LDO in each layer can not only provide the different supply voltage but also reduce the supply voltage noise cause by the power TSV and package. We assume the power source is 1V, and the proposed LDO can provide 0.9V, 0.8V or 0.7V on each layer. Fig. 5.16 shows the simulation result of our LDO on the layer 4.

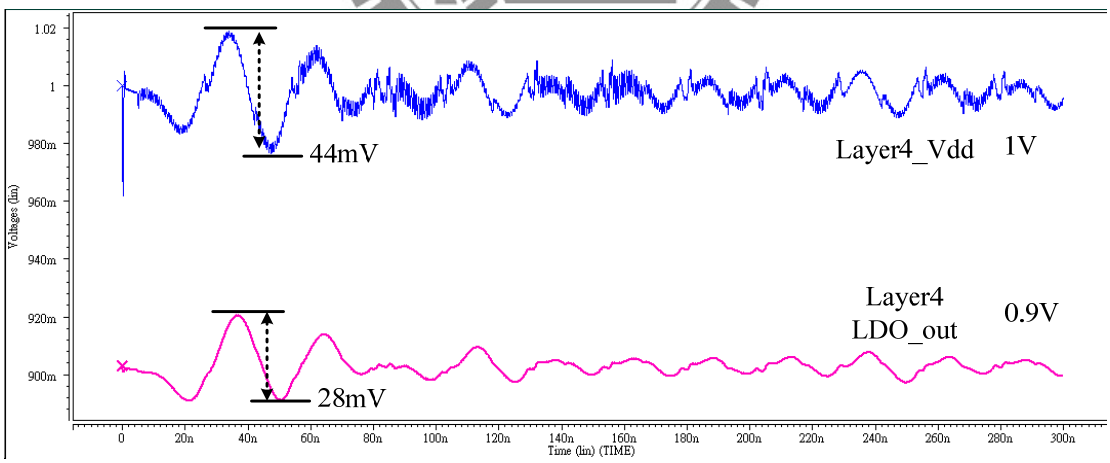


Fig. 5.16 Simulation of LDO on fourth layer.

The variable voltage reference is proposed in chapter 5.2 is used to generate variable supply voltage in each layer according to the requirement of circuit domain. The proposed LDO can provide three supply voltages, 0.9V or 0.8V or 0.7V,

according to different control codes.

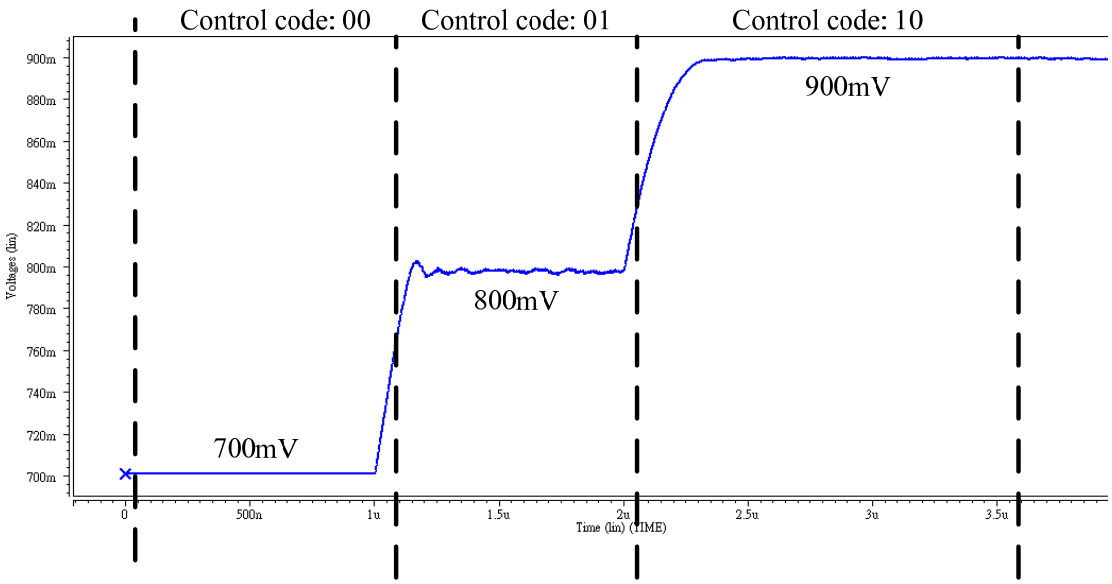


Fig. 5.17 Dynamic voltage scaling of LDO

Substrate coupling noise can cause important performance degradation of the analog circuits. There are two main kinds of substrate noise in TSV 3D Integration. First, if the digital circuit and analog circuits are in the same layer, similar to the mixed-signal IC, will the digital switching noise couple to sensitive analog circuits. The shared silicon substrate is one important medium through which it ("substrate coupling" or "substrate noise coupling") occurs. A digital state transition causes fluctuations in the underlying voltage that spread through the substrate and cause substrate potential variations at sensitive analog [5.3]. Moreover, the digital switching noise will become serious due to the impedance of power TSV, as shown as Fig. 5.18.

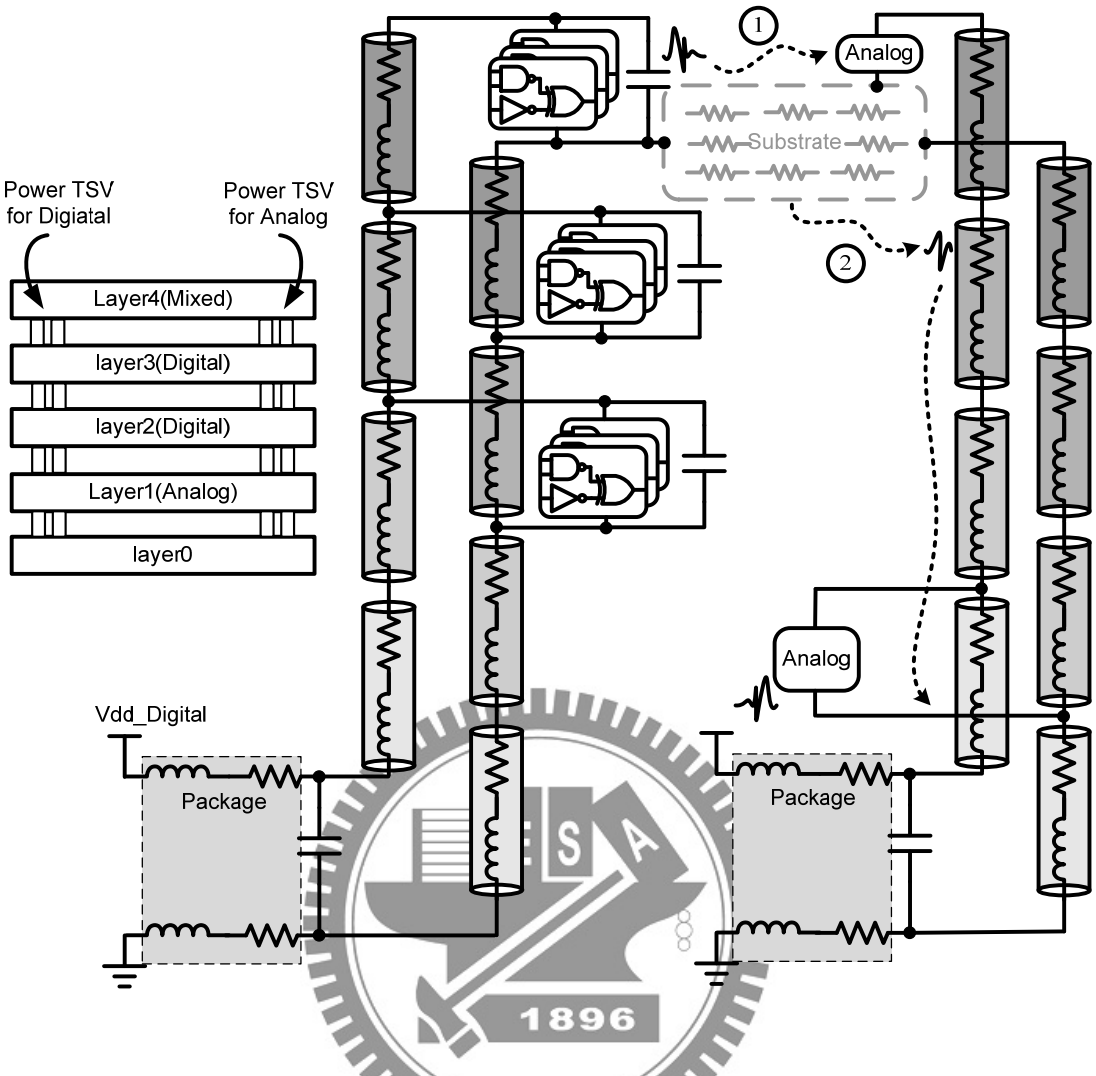


Fig. 5.18 Substrate noise in TSV 3D integration

Second, for DC isolation between the power TSV and the highly conductive silicon substrate, a SiO_2 insulation layer is formed surround the power TSV structure. The insulation layer has a thickness of a few hundredths of a nanometer, which results characteristic high capacitance between the TSV and the silicon substrate. Therefore, high frequency noise in silicon substrate can be coupled to the TSV through the large capacitance [5.7]. The coupled noise propagates through the TSV and corrupts the circuits connected to the TSV which degrades circuit performance as illustrated in Fig. 5.18. In a traditional 2D system, several noise isolation techniques have been introduced: split power planes, the deep-nwell process, and the guard ring structure.

The techniques have been proven for 2D noise isolation, but they are problematic if the TSV structure is connected for 3D integration.

Fig. 5.19 shows all substrate noise in TSV 3D integration. The dot lines present the substrate noise on each layers, this noise is because digital circuits and analog circuits are in the same layer, similar to the mixed-signal IC. The substrate noise is large and large when the chip is stacking high, approximant 0.8 time large per every layer stacking, and almost three times large on the layer 4 rather than layer. It shows that the power TSV impedance should be considered at this resonant frequency.

Another, the solidly line presents the analog power TSV noise, which is coupling from the substrate noise caused by digital switching. For example, there is only analog circuit in the layer1 in Fig. 5.18, so no digital switching noise will directly couple into this substrate. However, the substrate noise, in layer2 and layer3 and layer4, will propagate through the TSV and corrupts the analog power TSV which degrades the analog circuit performance in layer 1. Table 5.2 shows the normalization of these two kinds substrate noise. According this result, although the digital circuit and analog circuit are in the different layer, still need to consider the substrate noise coupled from other layers.

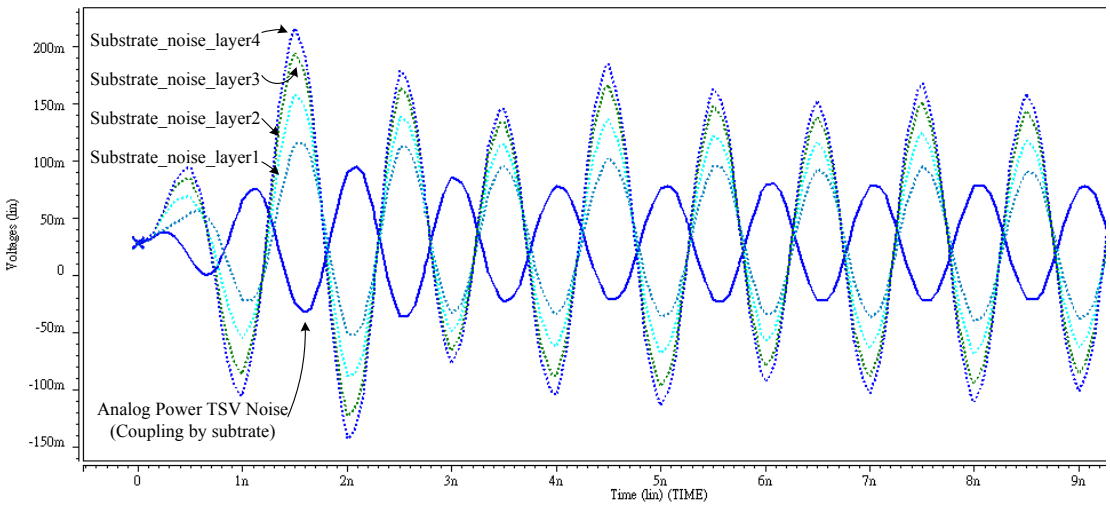


Fig. 5.19 Substrate noise in TSV 3D integration

Table 5.2 Normalization of substrate noise in TSV 3D integration

Location	**Substrate**				**Analog Power TSV**
Normalization	Layer1	Layer2	Layer3	Layer4	Analog Ground
	1x	1.9x	2.7x	3.2x	0.53x

The active substrate noise cancelling system is used to reduce the both kind of substrate noise which is mentioned above. The anti-phase noise signal is injected into substrates in each layer. Fig. 5.20 presents the improvement of worst case of first kind substrate noise and the analog power TSV noise. Table 5.3 list the improvement in each layer substrate, and it has 30% reduction at least.

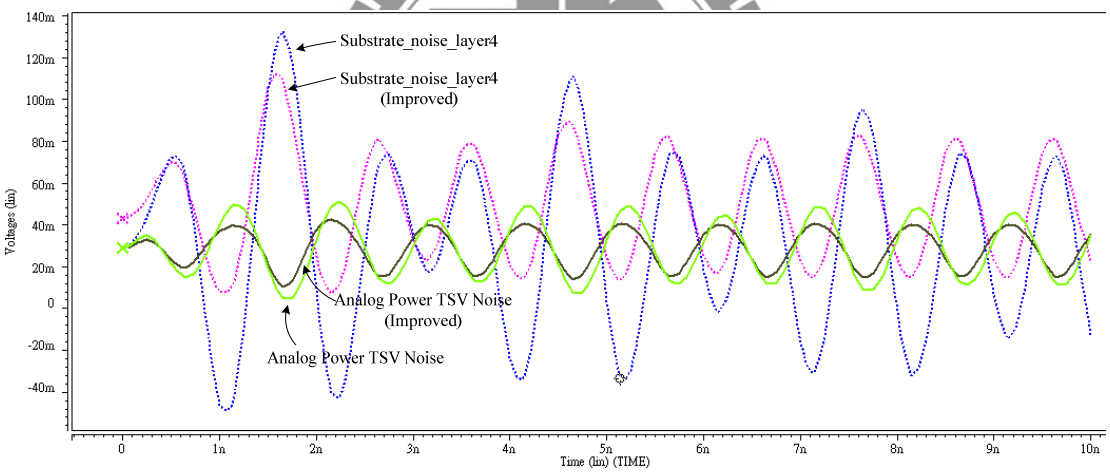


Fig. 5.20 Improvement of substrate noise in TSV 3D integration

Table 5.3 Improvement of substrate noise in TSV 3D integration

Location	**Substrate**				**Analog Power TSV**
Improvement	Layer1	Layer2	Layer3	Layer4	Analog ground
	32.7%	42.1%	42.7%	42.1%	30.0%

Chapter 6
Conclusions and Future Work

6.1 Conclusions

Three-dimensional (3D) integration technology can provide enormous advantages in achieving multi-functional integration, improving system speed and reducing power consumption for future generations of ICs. However, stacking multiple dies would face a severe problem of the power integrity. The power to be delivered to TSV 3D integration, per package pin, per power TSB, is tremendously increased, leading to significant complications in the task of reliable power delivery. This thesis presents an overview of both of these problems and outlines solution schemes to overcome the corresponding bottlenecks.

In this thesis, we proposed a design method for power TSV bundle in Chapter 3. It derives the adequate TSV parameter in order to satisfy the efficiency condition between power noise reduction and area overhead. A new scheme of supply noise regulation using active decoupling capacitor for TSV 3D integration is proposed in Chapter 4. Through the latch-based noise detection circuitry, the power supply noise can be detected and regulated via active DECAPs. In Chapter 5, we proposed a power delivery system for TSV 3D integration. Both substrate noise canceller and proposed low quiescent current voltage regulator are used to reduce the supply noise and substrate noise respectively. The whole system is able to solve the noise problem in the power system in TSV 3D integration.

6.2 Future Work

Power integrity is expected to be a major physical design concern in the TSV 3D integration due to higher power density and package asymmetries. In this thesis, we build a power delivery system to solve the supply noise problem in TSV 3D integration. However, 3D IC is able to integrate many applications together as shown in Fig. 6.2. Hence, we need to design a power management system to provide different supply voltages to each application. For example, 1.8V~2.0V for analog circuitry, 0.5V~1V for low power circuitry, -1.2V for memory circuitry, and 5V for I/O components. Moreover, the thermal problem may be aggravated in 3-D ICs because more transistors are packed, resulting in higher power density, and materials (such as on-chip interconnects and bonding materials) between chips have poor thermal conductivity. Design partitioning is a way to address the issues by placing the circuits (e.g., processors), which generate more heat, on the bottom layer directly attached to the heat sink and the circuits (e.g., memories), which generates less heat, on top of the stack. Insertion of thermal TSVs and thermal spreader between layers can also alleviate the thermal issues. Therefore, design a dynamical 3-D power delivery system with thermal feedback and control with the processor may greatly reduce the overall thermal constraints in 3-D integration as shown as Fig. 6. 1.

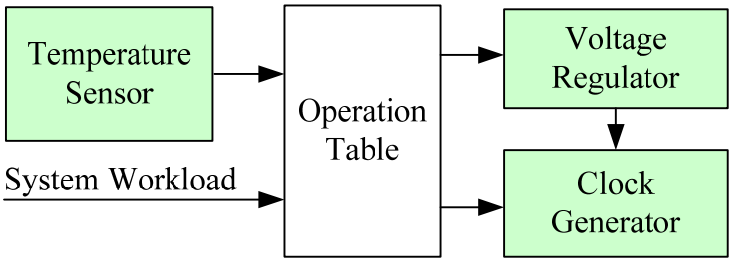


Fig. 6. 1 Thermal-Power management diagram

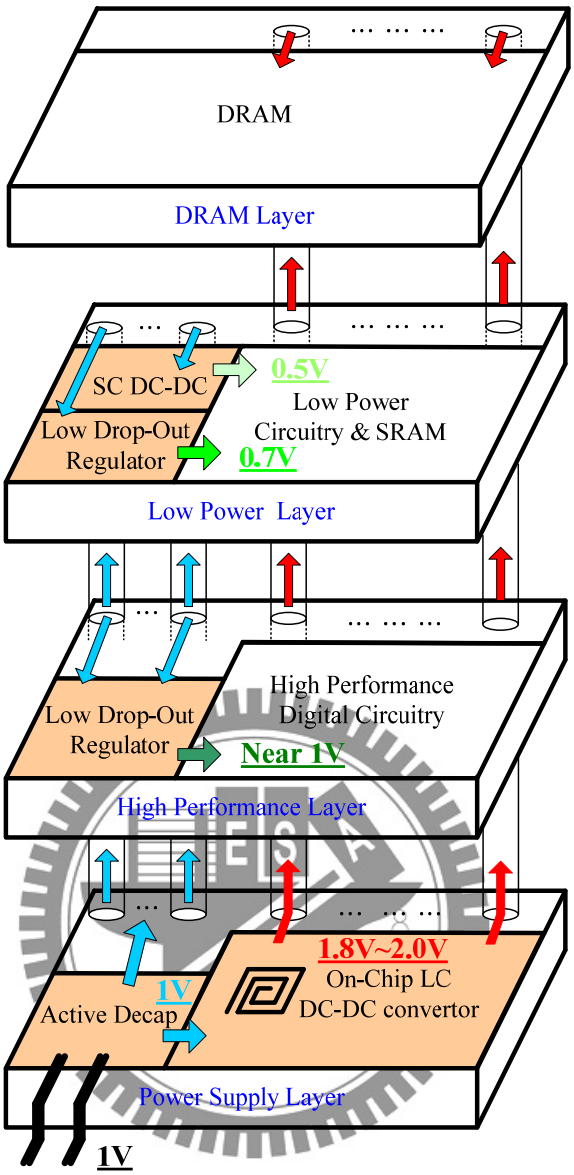


Fig. 6.2 3D IC application

Bibliography

[1.1] G. Huang, M. Bakir, A. Naeemi, H. Chen, and J.D. Meindl, "Power Delivery for 3D Chip Stacks: Physical Modeling and Design Implication," Proc. IEEE Conf. Electrical perform. Electron. Packaging Syst., pp. 205-208, Oct. 2007.

[1.2] W. Ahmad, L.R. Zheng, R. Weerasekera, Qiang Chen, A.Y. Weldezion, and H. Tenhunen, "Power Integrity Optimization of 3D Chips Stacked through TSVs," Proc. IEEE Conf. Electrical perform. Electron. Packaging Syst., pp. 105-108, Oct. 2009.

[1.3] P. Jain, T.H. Kim, J. Keane, and C.H. Kim, "A Multi-Story Power Delivery Technique for 3D Integrated Circuits," Proc. IEEE Int. Symp. Low Power Electron. and Des., pp. 57-62, Aug. 2008.

[1.4] X. Meng and R. Saleh, "An Improved Active Decoupling Capacitor for Hot-Spot Supply Noise Reduction in ASIC Designs," IEEE J. Solid-State Circuits, Vol. 44, No. 2, pp. 584-593, Feb. 2009.

[2.1] E. Beyne, "The Rise of the 3rd Dimension for System Integration," IEEE IITC, pp.1-5, July 2006.

[2.2] R. R. Tummala, V. Sundaram, R. Chatterjee, P.M. Raj, N. Kumbhat, V. Sukumaran, V. Sridharan, A. Choudury, Qiao Chen, and T. Bandyopadhyay, "Trend from ICs to 3D ICs to 3D Systems," Proc. IEEE Conf. Custom Integrated Circuits Conference, pp. 439-444, Sept. 2009

[2.3] Yole Development. (2007). 3DIC & TSV Report [Online]. "http://www.yole.fr/pagesan/products/report_sample/3dic.pdf"

[2.4] M. S. Bakir, C. King, D. Sekar, H. Thacker, Bing Dan, Gang Huang, A. Naeemi, and J.D. Meindl, "3D Heterogeneous Integrated Systems: Liquid Cooling, Power Delivery, and Implementation," Proc. IEEE Conf. Custom Integrated Circuits Conference, pp. 663-670, Nov. 2008

[2.5] T. Whipple, T. Kukal, K. Felton, and V. Gerousis, "IC-Package Co-design and Analysis for 3D-IC Designs," Proc. IEEE Conf. 3DIC, pp. 1-6, Oct. 2009.

[2.6] V.F. Pavlidis and E.G. Friedman, "Interconnect-Based Design Methodologies for Three-Dimensional Integrated Circuits," Proceedings of the IEEE, Vol. 97 No. 1, pp. 123-140, Jan. 2009

[2.7] M. Motoyoshi, "Through-Silicon Via (TSV)," Proceedings of the IEEE, Vol. 97, No. 1, pp.43-48, Jan. 2009

[2.8] M. Koyanagi, T. Fukushima, and T. Tanaka, "High-Density Through Silicon Vias for 3-D LSIs," Proceedings of the IEEE, Vol. 97 No. 1, pp. 49-59, Jan. 2009

[2.9] P. Marchal, B. Bougard, G. Katti, M. Stucchi, W. Dehaene, A. Papanikolaou, D. Verkest, B. Swinnen, and E. Beyne, "3-D Technology Assessment: Path-Finding the Technology/Design Sweet-Spot," Proceedings of the IEEE, Vol. 97, No. 1, pp. 96-107, Jan. 2009

[2.10] Tong Zhang, R. Micheloni, Guoyan Zhang, Z.R. Huang, and J.J. Lu, "3-D Data Storage, Power Delivery, and RF/Optical Transceiver-Case Studies of 3-D Integration From System Design Perspectives" Proceedings of the IEEE, Vol. 97, No. 1, pp. 161-174, Jan. 2009

[2.11] S. M. Alam, R.E. Jones, S. Pozder, R. Chatterjee, S. Rauf, and A. Jain, "Interstratum Connection Design Considerations for Cost-Effective 3-D System Integration," IEEE TVLSI, Vol. 18, No. 3, pp. 450-460 Feb. 2010.

[2.12] Jin-Fu Li and Cheng-Wen Wu, “Is 3D Integration an Opportunity or Just a Hype,” Proc. IEEE Conf. ASPDAC, pp. 541-543, Feb. 2010.

[2.13] M. N. Özisik. “Heat Transfer: A Basic Approach.” McGraw-Hill, New York, NY, 1985.

[2.14] H. Qian and S.S. Sapatnekar, “Hierarchical random-walk algorithms for power grid analysis.” In Proceedings of the Asia-South Pacific Design Automation Conference, pp. 499–504, 2004.

[2.15] P. Li, L.T. Pileggi, M. Asheghi, and R. Chandra, “IC thermal simulation and modeling via efficient multigrid-based approaches.” IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, Vol. 25, No. 9 pp. 1763–1776, September 2006.

[2.16] Semiconductor Industry Association. International technology roadmap for semiconductors (ITRS), 2004. http://public.itrs.net/.

[2.17] J. Sun, J. Lu, D. Giuliano, T. P. Chow, and R. J. Gutmann, “3D power delivery for microprocessors and high-performance ASICs.” In Proceedings of IEEE Applied Power Electronics Conference, pp. 127–133, Feb. 2007.

[2.18] N. Na, T. Budell, C. Chiu, E. Tremble, and I. Wernple, “The effects of on-chip and package decoupling capacitors and an efficient ASIC decoupling methodology.” In Proceedings of the IEEE Electronic Components and Technology Conference, pp. 556–567, June 2004.

[2.19] E. Hailu, D. Boerstler, K. Miki, J. Qi, M. Wang, and M. Riley, “A circuit for reducing large transient current effects on processor power grids.” In Proceedings of the IEEE International Solid-State Circuits Conference, pp. 2238–2245, Feb. 2006.

[2.20] G. Huang, M. Bakir, A. Naeemi, H. Chen, and J.D. Meindl, “Power Delivery for 3D Chip Stacks: Physical Modeling and Design Implication,” Proc. IEEE Conf. Electrical perform. Electron. Packaging Syst., pp. 205-208, Oct. 2007.

[2.21] J. A. Burns, B. F. Aull, C. K. Chen, C. L. Keast, J. M. Knecht, V. Suntharalingam, K. Warner, P.W.Wyatt, and D. Yost, “A wafer-scale 3-D circuit integration technology” IEEE Transactions on Electron Devices, Vol. 53, No. 10, pp. 2507–2516, October 2006.

[2.22] J. Gu and C. H. Kim, “Multi-story power delivery for supply noise reduction and low voltage operation.” In Proceedings of the ACM International Symposium on Low Power Electronics and Design, pp. 192–197, Aug. 2005.

[2.23] G. Schrom, P. Hazucha, J. Hahn, V. Kursun, D. Gardner, S. Narendra, T. Karnik, and V. De, “Feasibility of monolithic and 3D-stacked DC-DC converters for microprocessors in 90nm technology generation.” In

Proceedings of the ACM International Symposium on Low Power Electronics and Design, pp. 263–268, Aug. 2004.

[3.1] G. Huang, D.C. Sekar, A. Naeemi, K. Shakeri, and J.D. Meindl, "Compact Physical Models for Power Supply Noise and Chip/Package Co-Design of Gigascale Integration," Proc. IEEE Conf. ECTC, pp 1659-1666, June 2007.

[3.2] D. E. Khalil and Y. Ismail, "Optimum Sizing of Power Grids for IR Drop," Proc. IEEE Conf. ISCAS, pp 480-484, Step. 2006.

[3.3] W.H. Lee, S. Pant, and D. Blaauw, "Analysis and Reduction of On-Chip Inductance Effects in Power Supply Grids," Proc. IEEE Conf. ISQED, pp. 131-136, Aug. 2004.

[3.4] I. Savidis and E. G. Friedman, "Electrical Modeling and Characterization of 3-D Vias," Proc. IEEE Conf. ISCAS, pp. 784-787, May 2008.

[3.5] W. Ahmad, Li-Rong Zheng, R. Weerasekera, Qiang Chen, A.Y. Weldezion, and H. Tenhunen, "Power Integrity Optimization of 3D Chips Stacked Through TSVs," Proc. IEEE Conf. EPEPS, pp. 105-108, Nov. 2009.

[3.6] R. Weerasekera, M. Grange, D. Pamunuwa, H. Tenhunen, and Li-Rong Zheng; "Compact Modelling of Through-Silicon Vias (TSVs) in Three-Dimensional (3-D) Integrated Circuits," Proc. IEEE. Conf. 3DIC, pp. 1-8, Sept. 2009

[3.7] A. Shayan, Xiang Hu, He Peng, Chung-Kuan Cheng, Wenjian Yu, M. Popovich, T. Toms, Xiaoming Chen, "Reliability Aware Through Silicon Via Planning for 3D Stacked ICs," Proc. IEEE Conf. DATE, pp. 288-291, April 2009.

[3.8] C.S. Selvanayagam, J.H. Lau, Xiaowu Zhang, S. Seah, K. Vaidyanathan, T.C. Chai, "Nonlinear Thermal Stress/Strain Analysis of Copper Filled TSV (Through Silicon Via) and Their Flip-Chip Microbumps," Proc. IEEE Conf. ECTC, pp. 1073-1081, May 2008.

[3.9] N.H. Khan, S.M. Alam, S. Hassoun, "System-Level Comparison of Power Delivery Design for 2D and 3D ICs," Proc. IEEE Conf. 3DIC, pp. 1-7, Sept. 2009.

[3.10] I. Savidis and E.G. Friedman, "Closed-Form Expressions of 3-D Via Resistance, Inductance, and Capacitance," IEEE Transaction on Electron Device, pp. 1873 – 1881, Sept. 2009

[3.11] A.V. Mezhiba and E.G. Friedman," Inductive properties of high-performance power distribution grids," IEEE TVLSI, Vol. 10, No. 6, pp. 762-776, Dec. 2002

[3.12] R. Jakushoksa and E.G. Friedman," Inductance Model of Interdigitated Power and Ground Distribution Networks," IEEE TCSII, Vol. 56, No. 7, pp. 585-589, July 2009

[3.13] E. Salman, E.G. Friedman, R.M. Secareanu, and O.L. Hartin, "Worst Case Power/Ground Noise Estimation Using an Equivalent Transition Time for Resonance," IEEE TCAS I, Vol. 56, No. 5, pp. 997 - 1004, May 2009

[4.1] Semiconductor Industry Association. International technology roadmap for semiconductors (ITRS), 2007. http://public.itrs.net/.

[4.2] H. H. Chen, J. S. Neely, M. F. Wang, and G. Co, "On-Chip Decoupling Capacitor Optimization for Noise and Leakage Reduction," Proc. IEEE Conf. SBCCI, pp. 251-255, Sept. 2003

[4.3] M. Popovich and E.G. Friedman, "Nanoscale On-Chip Decoupling Capacitors," Proc. IEEE Conf. SOCC, pp. 51-54, Sept. 2008

[4.4] M. Popovich, E.G. Friedman, R.M. Secareanu, and O.L. Hartin, "Efficient Distributed On-Chip Decoupling Capacitors for Nanoscale ICs," IEEE TVLSI, Vol. 16, No. 12, pp. 1717-1721, Dec. 2008

[4.5] A. Todri, M. Marek-Sadowska, F. Maire, and C. Matheron, "A Study of Decoupling Capacitor Effectiveness in Power and Ground Grid Networks," Proc. IEEE Conf. ISQED, pp. 653-658, March 2009

[4.6] Y. Cai, J. Fu, X. Hong, S. X.-D Tan, and Z. Luo, "Power/Ground Network Optimization Considering Decap Leakage Currents," IEEE TCAS II, Vol. 53, No.10, pp. 1012-1016, Oct. 2006

[4.7] B. Bozorgzadeh and A. Afzali-Kusha, "Novel MOS Decoupling Capacitor Optimization Technique for Nanotechnologies," Proc. IEEE Conf. VLSI, pp 175-180, Jan. 2009

[4.8] E. Wong, J. Minz, and S. K. Lim. "Decoupling capacitor planning and sizing for noise and leakage reduction," In Proceedings of the IEEE/ACM International Conference on Computer- Aided Design, pp. 395–400, Nov. 2006.

[4.9] P. Zhou, K. Sridharan, and S.S. Sapatnekar. "Congestion-aware power grid optimization for 3D circuits using MIM and CMOS decoupling capacitors," In Proceedings of the Asia-South Pacific Design Automation Conference, pp. 179–184, Jan. 2009.

[4.10] G. Keskin, X. Li, L. Pileggi, "Active On-Die Suppression of Power Supply Noise," Proc. IEEE Conf. CICC, pp. 813-816, Sept. 2006

[4.11] T. Tsukada, Y. Hashimoto, K. Sakata, H. Okada, and K. Ishibashi, "An On-Chip Active Decoupling Circuit to Suppress Crosstalk in Deep-Submicron

CMOS Mixed-Signal SoCs," IEEE J. Solid-State Circuits, Vol. No. 1, pp. 67-79 ,Jan. 2005

[4.12] J. Gu, R. Harjani, and C. Kim, "Distributed Active Decoupling Capacitors for On-Chip Supply Noise Cancellation in Digital VLSI Circuits," Proc. IEEE Conf. VLSIC, pp. 216-217, Feb. 2007

[4.13] J. Gu, R. Harjani, and C.H. Kim, "Design and Implementation of Active Decoupling Capacitor Circuits for Power Supply Regulation in Digital ICs," IEEE TVLSI, Vol. 17, No. 2, pp. 292-301, Feb. 2009

[4.14] S. Pant and D. Blaauw, "Circuit Techniques for Suppression and Measurement of On-chip Inductive Supply Noise," Proc. IEEE Conf. ESSCIRC, pp. 134-137, Sept. 2008

[4.15] S. Pant and D. Blaauw, "A Charge-Injection-Based Active-Decoupling Technique for Inductive-Supply-Noise Suppression," Proc. IEEE Conf. ISSCC, pp. 416-624, Feb. 2008

[4.16] Y. Nakamura, M. Takamiya, and T. Sakurai, "An On-Chip Noise Canceller with High Voltage Supply Lines for Nanosecond-Range Power Supply Noise," Proc. IEEE Conf. VLSIC, pp. 124-15, June 2007

[4.17] Ji. Xu, P. Hazucha, M. Huang, P. Aseron, F. Paillet, G. Schrom, J. Tschanz, C. Zhao, V. De, T. Karnik, and G. Taylor, "On-Die Supply-Resonance Suppression Using Band-Limited Active Damping," Proc. IEEE Conf. ISSCC, pp. 286-603, Feb. 2007

[4.18] J. Xu, P. Hazuch, Z. Wu, P. Aseron, M. Huang, F. Paillet, G. Schrom, J. Tschanz, V. De, T. Karnik, and G. Taylor, " A Band-Limited Active Damping Circuit With 13 dB Power Supply Resonance Reduction." IEEE J. Solid-State Circuits, Vol. 43, No. 1, pp. 61-68, Jan. 2008

[4.19] X. Meng, K. Arabif, and R. Saleh, "A Novel Active Decoupling Capacitor Design in 90nm CMOS," Proc. IEEE Conf. ISCAS, pp. 657-660, May 2007

[4.20] X. Meng and R. Saleh," An Improved Active Decoupling Capacitor for "Hot-Spot" Supply Noise Reduction in ASIC Designs," IEEE J. Solid-State Circuits, Vol. 44, No. 2, pp.584-593, Feb. 2009

[4.21] X. Meng and R. Saleh, "Active Decap Design Considerations for Optimal Supply Noise Reduction," Proc. IEEE Conf. ISQED, pp. 765-769, March 2009

[4.22] J. Gu, H. Eom, and C.H. Kim," On-Chip Supply Noise Regulation Using a Low-Power Digital Switched Decoupling Capacitor Circuit," IEEE J. Solid-State Circuits, Vol. 44, No. 6, pp. 1765-1775, June 2009

[4.23] J.-T. Wu and K.-L. Chang, "MOS Charge Pump for Low-Voltage Operation," IEEE J. Solid-State Circuits, Vol. 33, No. 4, pp. 592-597, Apr.1998.

[4.24] I. Savidis and E.G. Friedman, “Closed-Form Expressions of 3-D Via Resistance, Inductance, and Capacitance,” IEEE Transaction on Electron Device, pp. 1873 – 1881, Sept. 2009

[4.25] J. Shin, I.-Y. Chung, Y.J. Park, and H.-S. Min, “A New Charge Pump without Degradation in Threshold Voltage due to Body Effect,” IEEE J. Solid-State Circuits, Vol. 35, No. 8, pp. 1227-1230, Aug. 2000.

[4.26] J. F. Dickson, “On-chip high-voltage generation in MNOS integrated circuits using an improved voltage multiplier technique,” IEEE J. Solid- State Circuits, Vol. 1, No. 3, pp. 374–378, Jun. 1976.

[5.1] T. Kazama, T. Nakura, M. Ikeda, and K. Asada, “Design of Active Substrate Noise Canceller using Power Supply di/dt Dectect,” Proc. IEEE Conf. ASPDAC, pp. 100-101, Jan. 2007

[5.2] M. Al-Shyoukh, H. Lee, and R. Perez, “A Transient-Enhanced Low-Quiescent Current Low-Dropout Regulator With Buffer Impedance Attenuation,” IEEE J. Solid-State Circuits, Vol. 42, No. 8, pp.1732-1742, Aug. 2007.

[5.3] P. Hazucha, T. Karnik, D. Finan, and S. Borkar, “Area-Efficient Linear Regulator With Ultra-Fast Load Regulation,” IEEE J. Solid-State Circuits, Vol. 40, No. 4, pp.933-940, Apr. 2005.

[5.4] P. Hazucha, S. T. Moon, G. Schrom, F. Paille, D. Gardner, S. Rajapandian, and T. Karnik, “High Voltage Tolerant Linear Regulator With Fast Digital Control for Biasing of Integrated DC-DC Converters,” IEEE J. Solid-State Circuits, Vol. 42, No. 1, pp.66-73, Jan. 2007.

[5.5] C. –Y. Wu, “High Efficiency Power Management System for Solar Energy Harvesting Applications,” Thesis of NCTU, Sept. 2009

[5.6] A. Afzali-Kusha, M. Nagata, N.K. Verghese, and D.J. Allstot, “Substrate Noise Coupling in SoC Design: Modeling, Avoidance, and Validation,” IEEE JPROC, Vol. 94, No. 12, pp. 2109-2138, Dec. 2006.

[5.7] Jonghyun Cho, Jongjoo Shim, Eakhwan Song, Jun So Pak, Junho Lee, Hyungdong Lee, Kunwoo Park, and Joungho Kim, “Active Circuit to Through Silicon Via (TSV) Noise Coupling,” Proc. IEEE Conf. EPEPS, pp. 97-100, Oct. 2009.

Vita

林天鴻 Tien-Hung Lin

PERSONAL INFORMATION

Birth Date: October. 17, 1986

Birth Place: Taipei, TAIWAN.

E-Mail Address: skyredd@gmail.com

EDUCATION

09/2008 – 07/2010 M.S. in Electronics Engineering, National Chiao Tung University
Thesis: Power Integrity in TSV 3D Integration

09/2004 – 06/2008 B.S. in Electrical Control Engineering, National Chiao Tung University

PUBLICATIONS

Tien-Hung Lin, Po-Tsang Huang, and Wei Hwang, "Power Noise Suppression Technique using Active Decoupling Capacitor for TSV 3D Integration" in Proc. IEEE System On Chip Conference, SOCC, Sept. 2010. (Accepted)

PATENTS

Tien-Hung Lin, Po-Tsang Huang, and Wei Hwang, "Power Noise Suppression Technique using Active Decoupling Capacitor for TSV 3D Integration" US/TW Patent Pending (Submitted)