A Review on Low Leakage Techniques for VLSI Circuits

Vivek Bansal
M.Tech Student, Shri Ram College of Engineering and Management, Banmore, INDIA

Abstract: **In today's battery powered era, to reduce power is the basic aim in the mind of designers. Power dissipation is an important issue in VLSI circuits. Various techniques are adopted at various levels of abstraction to reduce power consumption and hence to improve the overall circuit performance. In this paper we discuss various low power reduction techniques and find out the best suitable technique.**

1. INTRODUCTION

It's no secret that power is emerging as the most critical issue in system- on-chip (SoC) design today. Power management is becoming an increasingly urgent problem for almost every category of design, as power density-measured in watts per square millimeter rises at an alarming rate. From a chip-engineering perspective, effective energy management for a SoC must be built into the design starting at the architecture stage; and low power techniques need to be employed at every stage of the design, from RTL to GDSII. Fred Pollack of Intel first noted a rather alarming trend in his keynote at MICRO-32 in 1999. He made the now well-known observation that power density is increasing at an alarming rate, approaching that of the hottest man-made objects on the planet, and graphed power density as shown in **Fig. 1.1**

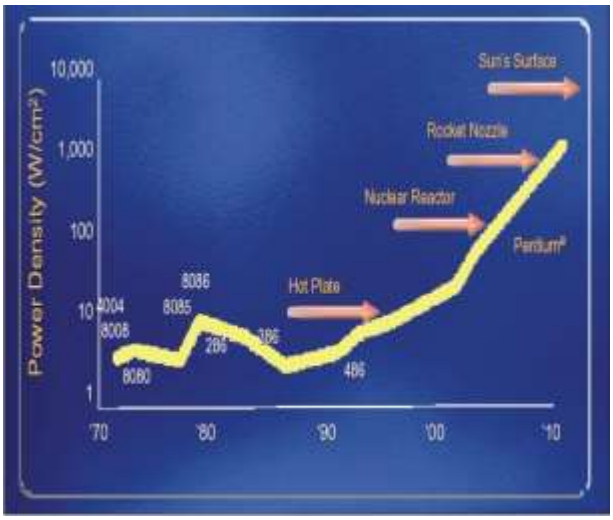

Figure 1.1: Power density with shrinking technology. Courtesy Intel Corporation [1]

Power management is a must for all designs of 90nm and below. Both market forces and process technology have driven power to the forefront of all factors constraining electronic design. The increasing demand for high-performance, battery-operated, system-on-chips (SoC) in communication and computing has shifted the focus from traditional constraints (such as area, performance, cost, and reliability) to power consumption. Just as important, though not so obvious, is the need to reduce power consumption for non-portable systems, such as base stations, where heat dissipation and energy consumption are critical concerns. At process nodes below 100 nm, power consumption due to *leakage* (static energy loss) has joined switching activity (dynamic power consumption) as a primary power management concern. It has been reported that leakage constitutes over 40 percent of total power expenditures at the 65 nm technology node. The quadratic dependency of power leakage on total transistor count easily qualifies leakage optimization as a key design objective. Consequently, designers have redirected their efforts toward exploring various techniques that reduce leakage and, thereby, increase battery life in the final product.

In this paper we discuss and reviewed the prevalent leakage reduction techniques. Section II discuss the brief overview of leakage power, section III did the comparative analysis of leakage reduction techniques and finally the paper is concluded in Section IV.

II. Leakage Power

Leakage has become one of the most dominant factors of power management of nanometer scale integrated circuits. The exponential increase in the leakage component of the total chip power can be attributed to threshold voltage scaling, which is essential to maintain high performance in active mode, since supply voltages are scaled. The main contributor to leakage current is the sub threshold leakage current.

Leakage currents are especially important in burst mode type IC where the majority of the time the system is in idle or sleep mode **[2],** where no computation is taking places. For example, cell phones, pagers will spend upwards of 90% of time in standby mode where the processor is waiting for user

input. At process nodes below 100 nm, power consumption due to *leakage* (static energy loss) has joined switching activity (dynamic power consumption) as a primary power management concern. It has been reported that leakage constitutes over 40 percent of total power expenditures at the 65 nm technology node [J. Kao, et al. 2002]. The leakage power is the product of the supply voltage and leakage current $I_{leakage.}$

$$P_{leakage} = V_{DD}\, I_{leakage} \quad (1)$$

Where, $I_{leakage}$ corresponds to the sub threshold current. Generally there are four main sources of leakage currents in a CMOS gate **(Fig. 2.)**

- **GATE LEAKAGE (I_{GATE}):** The current which flows directly from the gate through the oxide to the substrate due to the gate oxide tunneling and hot carrier injection.
- **GATE INDUCED DRAIN LEAKAGE (I_{GIDL}):** The current which flows from the drain to the substrate induced by the high field effect in the MOSFET drain caused by a high V_{DG}.
- **REVERSE BIAS JUNCTION LEAKAGE (I_{REV}):** This is caused by minority carrier drift and generation of electron/holes pairs in the depletion regions

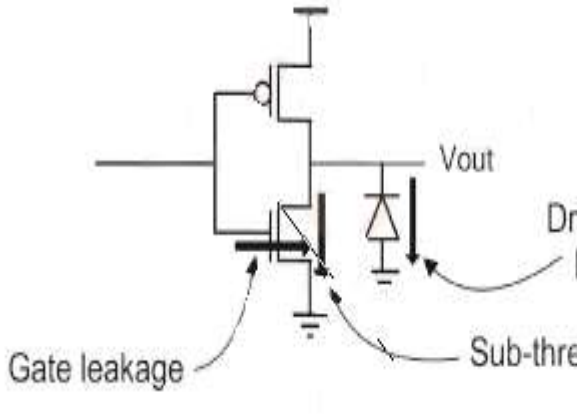

Figure 2. Leakage Currents [3]

Sub threshold leakage occurs when a CMOS gate is not turned completely OFF. Its value is given by-

$$I_{SUB} = \mu C_{OX}\, V^2_{TH}\, W/L\, e^{V_{GS} - V_T/\eta V_{TH}} \quad \textbf{(2)}$$

Where W and L are the dimensions of the transistor and V_{TH} is the thermal voltage KT/q. The parameter n is a function of the device fabrication process and ranges from 1.0 to 2.5. The sub threshold leakage current increases exponentially with temperature. This greatly complicates the problem of designing low power (LP) systems.

To manage the growing problem of leakage power dissipation, a proven approach is to use a complementary set of two or more libraries containing matching sets of logic cells. The sets of cells are the same except for the threshold voltages (V_T). Cell libraries with higher Vt operate at lower frequency and dissipate significantly less leakage power than libraries with lower Vt. The aim of this design methodology is to restrict use of lower-Vt cells to critical paths and use high-Vt cells on noncritical paths during both the logic and physical synthesis stages. The typical approach had been to apply multiple-pass synthesis, using one library at a time to either fix timing or reduce power, but not at the same time. There were also published utility scripts for manually swapping cells. The disadvantage of these approaches is the difficulty in finding the right balance for optimized timing and power. An overly aggressive approach with high-Vt cells could cause a timing closure problem, while very passive use of high-Vt cells would leave the design under optimized regarding leakage power consumption. Leakage power optimization can reduce leakage power dissipation 30% to 70%, while enabling replacement of 50% to 95% of the cells with high-Vt cells.

2.1 Diode Leakage

Diode leakage current occurs when a transistor is turned OFF, and another active transistor charges up/down the drain with respect to the former's bulk potential. In the case of an inverter with a high input voltage, the output voltage becomes "0" because the NMOS transistor is "ON". The PMOS transistor is turned OFF, but the drain to bulk voltage is equal to the supply voltage ($-V_{dd}$). The resulting diode leakage current is approximately

$$I_L = A_D \,.\, J_S$$

Where A_D is the area of the drain diffusion and J_S is the leakage current density set by the technology. Since the diode reaches the maximum reverse bias current for a relatively small reverse bias potential, the leakage current is roughly independent of the supply voltage. The leakage current is proportional to the diffusion area and the perimeter of the drain. Therefore, it is preferred to minimize the diffusion area and the perimeter in the layout. The leakage current density is exponentially proportional to temperature as well, so that J_s increases dramatically at higher temperature.

2.2 Sub threshold Leakage

This occurs when the gate-source voltage, Vgs, has exceeded the weak inversion point but is still below the threshold voltage Vth. In this region, the MOSFET behaves similar to a bipolar transistor, with

it's exponential characteristics. The current in the sub-threshold region is given by:

$$I_{SUB} = K\left(\frac{W}{L}\right) e^{\frac{Vgs-Vth}{nVT}} \left(1 - e^{Vds\text{-}VT}\right) \quad \textbf{(3)}$$

Where n and K are technology parameters, and Vds is the drain-source voltage.

Scaling down the supply voltage in CMOS requires also to scale down the threshold voltage, Vth, in order to maintain the performance of the scaled down logic. From the equation above, it becomes clear that the reduction of the threshold voltage increases the sub-threshold leakage current significantly. Sub-threshold leakage current along with reverse biased pn junction current are currently the most important components of leakage current.

2.3 Gate Induced drain Leakage (GIDL)

Gate induced drain leakage (GIDL) current (IGIDL) arises in the high electric field under the gate/drain overlap region causing deep depletion. GIDL occurs at low VG and high VD and generates carriers into the substrate and drain from surface traps or band-to-band tunneling.

2.4 Punch through

Punch through occurs when the drain and source depletion region approach each other and electrically "touch" deep in the channel. Punch through current (IPT) varies quadratically with drain voltage.

2.5 Gate Tunneling

Gate oxide tunneling current (IG) is present when the electric field at the gate is high enough to tunnel through the gate oxide layer. This phenomenon is common in scaled down devices with reduced oxide thickness.

III. Leakage Reduction Techniques

In nanoscale CMOS circuits, besides dynamic power consumption, energy is dissipated in terms of various forms of leakage, such as gate-oxide leakage, sub threshold leakage, junction tunneling leakage. The leakage power dissipation occurs both in active (ON) and passive (OFF) states of a device. Hence, leakage reduction has become more critical for nano-CMOS transistors and circuits. Leakage power dissipation will be additive to instantaneous as well as average power and hence will have impact on most of the power profile metrics. However, it is of the highest significance in devices designed for applications where the system goes to standby mode very often such as mobile phones and PDAs.

3.1 Multiple Threshold CMOS (MTCMOS) Technique

One of the natural ways to reduce the leakage of a circuit is to gate the power supply using power gating transistors (also called sleep transistors). Typically high VT power-gating transistors are placed between the power supplies and the logic gates. This is called MTCMOS (Multi-threshold CMOS) technique. In standby, these power-gating transistors are turned off, thus shutting off power to the gates of the circuit. In multiple threshold CMOS (MTCMOS) technique, two high-V_{Th} transistors are put in series with a low-V_{Th} logic block to cut off the path from power supply rail to ground rail during the idle or standby period to reduce the leakage as shown in **Fig. 3**.

This technique was proposed by inserting high threshold devices in series to low Vth circuitry. **Fig. 3(a)** shows the schematic of a MTCMOS circuit.. **Fig. 3 (b)** and **(c)** show the PMOS insertion and NMOS insertion schemes, respectively. The NMOS insertion scheme is preferable, since the NMOS on-resistance is smaller at the same width and it can be sized smaller than corresponding PMOS. However MTCMOS can only reduce the standby leakage power and the large inserted MOSFETs will increase the area and delay.

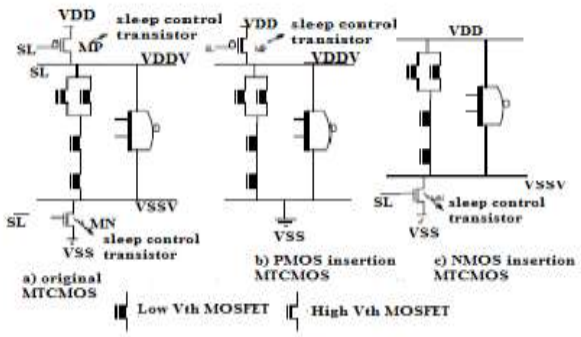

Figure 3 MTCMOS Technique

A conservative method to sizing the sleep transistors would be to first estimate the width of the sleep transistor required for each gate in a design such that the delay of the individual gate is within a specified bound and then add up the sleep transistor widths for all gates to come up with the total sleep transistor width required.

3.2 Dual Threshold CMOS Technique

For a logic circuit, a higher threshold voltage can be assigned to some transistors in non-critical paths so as to reduce leakage current, while the performance is maintained due to the low threshold transistors in the critical path(s). Therefore, no additional transistors are required, and both high performance and low

power can be achieved simultaneously. **Fig. 4** shows the schematic of a dual-Vth circuit. This dual threshold technique is good for both standby and active modes.

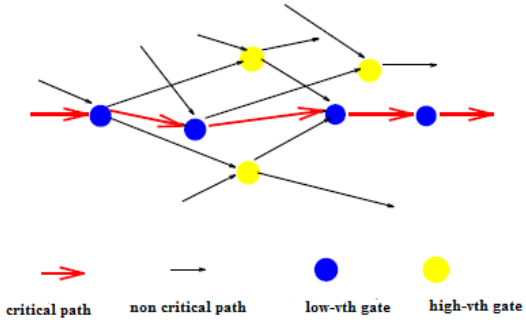

Figure 4 Dual Vth CMOS circuit

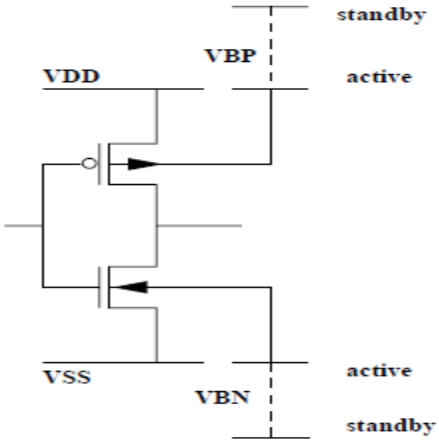

Figure 5 Variable threshold CMOS circuit

3.3 Variable Threshold CMOS (VTCMOS)/ Body Biasing Technique

Variable threshold technique is a body biasing design technique **[5]. Fig. 5** shows the VT-CMOS scheme. In order to achieve different threshold voltages, a self-substrate bias circuit is used to control the body bias. In active mode, a nearly zero body bias is applied. While in standby mode, a deeper reverse body bias is applied to increase threshold voltage and cut off leakage current. This scheme has been used in two dimensional discrete cosine transform core processor **[14].** Furthermore, in active mode, a slightly forward substrate bias can be used to increase the circuit speed while reducing short channel effect **[13], [15].**

In this technique the substrate bias voltage is dynamically varied to control the threshold voltage. All the transistors initially have low Vth and the substrate bias is altered for two main reasons: to reduce the leakage current in the standby mode and to compensate for the Vth fluctuations in the active mode and consequently minimizing the delay variations. This technique is very effective way of mitigating standard by leakage power. The body effect equation can be written as:

$$V_{TN} = V_{TO} + \gamma\,(\sqrt{V_{SB}} + 2\phi - \sqrt{2\phi}$$

Where V_{TN} is the threshold voltage with substrate bias present, and V_{TO} is the zero- V_{SB} value of threshold voltage, γ is the body effect parameter, and $2\,\phi$ is the surface potential parameter. Hence, this method of controlling the threshold voltage of transistors through body biasing is often referred to as the VTCMOS technique.

3.4 Dynamic Threshold CMOS (DTMOS) Technique

For dynamic threshold CMOS, the threshold voltage is altered dynamically to suit the operating state of the circuit. A high threshold voltage in the standby mode gives low leakage current, while a low threshold voltage allows for higher current drives in the active mode of operation. Dynamic threshold voltage can be achieved by tying the gate and body together- DTMOS. **Fig. 6** shows the schematic of DTMOS circuit.

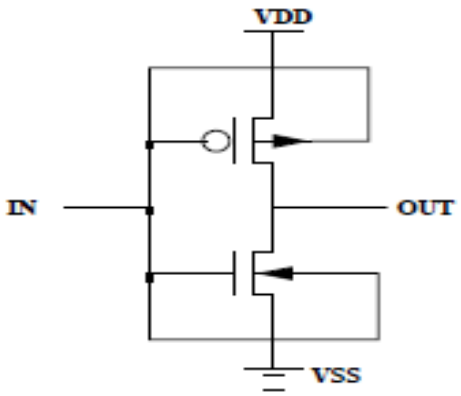

Figure 7 Schematic of DTMOS circuit

Fig. 8 shows the SOI DTMOS structure and layout. From **[15],** the excellent DC inverter characteristics down to 0.2V and good ring oscillator performance down to 0.3V can be achieved using this method.

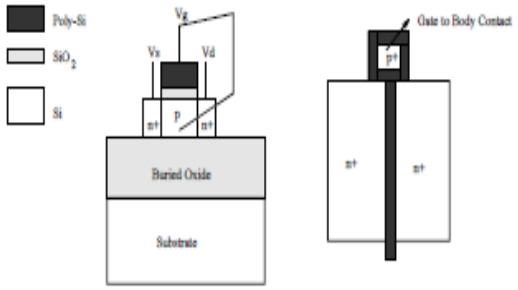

Figure 8 SOI DTMOS structure and layout

3.5 Multiple VDD CMOS Design Technique

Multiple supply voltage technique has been proposed for low voltage circuit design [**19], [20].** The high supply voltage can be assigned to the gates in critical path(s), while some gates in non-critical path will have lower supply voltage. **Fig. 9** shows the structure of the cluster voltage dual-VDD (CVS) scheme. There are two circuit clusters: one is for high VDD, the other is for low VDD. The high VDD cluster is placed in front of the low VDD cluster to avoid direct path leakage current and level converters. Dual-Vdd and dual-Vth techniques can be combined to further reduce the leakage power **[21].**

3.6 Stack Effect

Subthreshold leakage current flowing through a stack of series- connected transistors reduces when more than one transistor in the stack is turned off. This effect is known as the Stacking effect. This effect is best understood by considering a two- input NAND gate as shown in **Fig. 10**

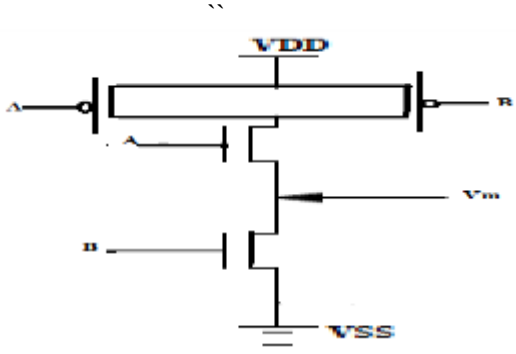

Figure 10 Stacking effect in two-input NAND gate

When both M1 and M2 are turned off, the voltage at the intermediate node (Vm) is positive due to small drain current. Positive potential at the intermediate node has 3 effects:

1. Due to the positive source potential Vm, gate – to – source voltage of M1 becomes negative. Hence, the sub threshold current reduces substantially.
2. Due to Vm > 0, body-to-source potential of M1 becomes negative, resulting in an increasing threshold voltage of M1, and thus reducing the sub threshold leakage.
3. Due to Vm < 0, the drain-to-source potential of M1 decreases, resulting in the increase in the threshold voltage of M1 and thus reducing the sub threshold leakage.

3.7 Power Gating Technique

Power gating is effective for reducing leakage power. Power gating is the technique wherein circuit blocks that are not in use are temporarily turned off to reduce the overall leakage power of the chip. This temporary shutdown time can also called as "low power mode" or "inactive mode". When circuit blocks are required for operation once again they are activated to "active mode". These two modes are switched at the appropriate time and in the suitable manner to maximize leakage power by temporarily cutting power off to selective blocks that are not required in that mode.

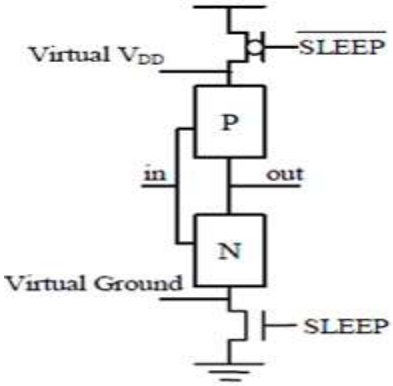

Figure 11 Power Gating Circuit

The most natural way of lowering the leakage power dissipation of a VLSI circuit in the STANDBY state is to turn off its supply voltage. This can be done by using one PMOS transistor and one NMOS transistor in series with the transistors of each logic block to create a virtual ground and a virtual power supply as shown in **Fig. 11** but in practice only one transistor is necessary. Because of their lower on-resistance, NMOS transistors are usually used.

In the active state, the sleep transistor is on. Therefore, the circuit functions as usual. In the standby state, the transistor is turned off, which disconnects the gate from the ground. It is to be keep in mind that to lower the leakage, the threshold voltage of the sleep transistor must be large. Otherwise, the sleep transistor will have a high leakage current, which will make the power gating less effective. Additional savings may be achieved if the width of the sleep transistor is smaller than the combined width of the transistors in the pull- down network. In practice, Dual VTH CMOS or multi-threshold CMOS is used for power gating. In these technologies there are several types of transistors with different VTH values. Transistors with a low VTH are used to implement the logic, while high-VTH devices are used as sleep transistors.

IV CONCLUSIONS

In this paper various leakage reduction techniques have been reviewed. The power gating technique is considered as best technique.

REFERENCES

[1] R. Bhanuprakash, Manisha Pattanaik, S.S Rajput and Kaushik Mazumdar, " Analysis & reduction of ground bounce noise and leakage current during mode transition of stacking power gating logic circuits" , proceedings of IEEE TENCON Singapore, pp. 1-6, 2009.

[2] "International Technology Roadmap for Semiconductors," Semiconductor Industry Association, 2005. [Online].Available: http://public.itrs.net

[3] Y. Chang, S. K. Gupta, and M. A. Breuer, "Analysis of ground bounce in deep sub-micron circuits," in Proc. 15th IEEE VLSI Test Symposium 1997,pp. 110–116

[4] S. Kim et al., "Understanding and minimizing ground bounce during mode transition of power gating structure," in Proc. Int. Symposium. Low-Power Electron. Design, August 2003, pp. 22–25

[5] S. Kim, S. V. Kosonocky, D. R.. Knebel, K. Stawiasz, D. Heidel, and M. Immediato, "Minimizing inductive noise in system-on-a-chip with multiple power gating structures," in Proceedings of European Solid-State Circuits: pp. 16-18, 2003.

[6] A. Kabbani and A. J. AI-Khalili, "Estimation of ground bounce effects on CMOS circuits," IREE Transactions on Components and Packaging Technology, vol. 22, pp. 316-325, June 1999.

[7] K. Agrawal et al. "Power gating with multiple sleep modes proceedings of 7th IEEE ISQED January 2006.

[8] M.D Pang and E.G Friendman, "An architectural solution for the inductive noise problem due to clock-gating", proceedings of International symposium on low power electronics and design, 263-266, 2002.

[9] Suhwan Kim, Chang Jun Choi, Deog-Kyoon Jeong ,Stephen V. Kosonocky, and Sung Bae Park" Reducing Ground-Bounce Noise and Stabilizing the Data-Retention Voltage of Power-Gating Structures" IEEE transactions on electron devices, vol.. 55, no.. 1, January 2008.

[10] Suhwan Kim; Kosonocky, S.V.; Knebel, D.R.; Stawiasz, K.," Experimental measurement of a novel power gating structure with intermediate power saving mode," Proc. of the Int. Symposium on Low Power Electronics and Design, 2004, pp. 20- 25.

[11] K. Agrawal, H. Deogun, D. Sylvester and K. Nowka, "Power Gating with Multiple Sleep Modes", Proc. of 7th IEEE ISQED, January, 2006

[12] S. V. Kosonocky, M. Immediato, P. Cottrell, T. Hook, R. Mann, and J. Brown, "Enhanced multi-threshold (MTCMOS) circuits using variable well bias," in Proceedings of International Symposium on Low-Power Electronics and Design, pp. 165-169, Aug. 2001.

[13] K. Kumagai, J. Iwaki, H. Suzuki, T. Yamada, and S. Kurosawa, "A novel powering-down scheme for low Vt CMOS circuits," in Proc. IEEE Symposium VLSI Circuits, 1998, pp. 44–45.

[14] C. Changbo and L. He, "Distributed sleep transistor network for power reduction", IEEE transactions on very large scale integration (VLSI) systems, vol.12, no. 9, pp. 937- 946, September 2004.

[15] Jun Cheol Park and Vincent J. Mooney" Sleepy Stack Leakage Reduction" IEEE transactions on very large scale integration (VLSI) SYSTEMS, VOL.14, no.11 November 2006.

[16] Piguet, C. Low Power Electronics Design, CRC press, 2005.

[17] Charbel J. Akl, Rafic A. Ayoubi, Magdy A. Bayoumi, " An effective staggered- phase damping technique for suppressing power- gating resonance noise during mode transition", 10th International Symposium on Quality of Electronic Design, pp. 116-119, 2009.

[18] Ku He, Rong Luo, Yu Wang, " A power gating scheme for ground bounce reduction during mode transition", in ICCD07, pp. 388-394,2007.

