

2011

On-chip current measurement for multi-site electromigration monitoring

Tianhan Wang
Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/etd>

 Part of the [Electrical and Computer Engineering Commons](#)

Recommended Citation

Wang, Tianhan, "On-chip current measurement for multi-site electromigration monitoring" (2011). *Graduate Theses and Dissertations*. 10451.
<https://lib.dr.iastate.edu/etd/10451>

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

**On-chip current measurement for multi-site
electromigration monitoring**

by

Tianhan Wang

A thesis submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Electrical Engineering

Program of Study Committee
Degang Chen, Major Professor
Randall L. Geiger
Jiming Song

Iowa State University

Ames, Iowa

2011

Copyright © Tianhan Wang, 2011. All rights reserved.

TABLE OF CONTENTS

LIST OF FIGURES.....	iv
LIST OF TABLES	v
ABSTRACT.....	vi
CHAPTER 1 INTRODUCTION	1
1.1 Thesis Outline.....	4
CHAPTER 2 BACKGROUND	5
2.1 Introduction	5
2.2 IC Reliability and Testing	5
2.3 Electromigration Lifetime Problem	8
2.3.1 Electromigration Lifetime Models for Current Density.....	10
2.4 Build-in Current Sensor	12
CHAPTER 3 CURRENT SENSOR ARCHITECTURE	16
3.1 Introduction of Proposed Current Sensor	16
3.2 Basic Concepts and Working Scenarios	17
3.2.1 Instant Current Measurement	18
3.2.2 Wear Out Lifetime Electromigration Monitoring.....	19
3.2.3 Offset Calibration	23
3.3 Current Sensor Architecture.....	24

3.4 Limiting Problem.....	28
3.4.1 Comparator Offset	29
CHAPTER 4 CIRCUIT IMPLEMENTATION AND SIMULATION RESULTS.....	31
4.1 Top Level of Current Sensor	31
4.2 Sub-Circuits.....	36
4.2.1 Comparator.....	36
4.2.2 Offset Cancellation Technique	39
4.2.3 Noise Reduction On the Current Sensor	42
4.2.4 M/2M Ladder DAC.....	43
4.3 Simulation Results	45
4.4 Summary	50
CHAPTER 5 CONCLUSIONS AND FUTURE WORK	51
ACKNOWLEDGEMENT.....	52
BIBLIOGRAPHY	53

LIST OF FIGURES

Figure 1-1: Processor power density over time	2
Figure 2-1: IC reliability bulb curve	6
Figure 3-1: Typical IR drop and working range of on-chip current sensor.....	18
Figure 4-1: Circuit of on-chip current sensor	31
Figure 4-2: Complete schematic of comparator.....	37
Figure 4-3: Comparaotr premap and small signal circuit diagram	38
Figure 4-4: Monte Carlo analysis of preamp offset	39
Figure 4-5: Offset cancellation technique	40
Figure 4-6: Comparator transient simulation waveform	41
Figure 4-7: Chopping technique to reduce 1/f noise.....	42
Figure 4-8: M/2M ladder DAC	44
Figure 4-9: Input-output transfer characteristic of current sensor	46
Figure 4-10: Absolute accuracy	46
Figure 4-11: Absolute error	47
Figure 4-12: DNL&INL	47
Figure 4-13: SAR Convergence process at different corners at IR = 0.....	48
Figure 4-14: SAR Convergence process at different corners at IR = 10mV	49
Figure 4-15: SAR Convergence process at different corners at IR = 20mV	49

LIST OF TABLES

Table 4-1: Transistors size for current sensor.....	36
Table 4-2: Transistors size for comparator.....	38
Table 4-3: Summary of performance	50

ABSTRACT

Power consumption and die heating are of major concern in high-density high-speed integrated circuits. The performance of modern IC designs is limited by power consumption and thermal issues. Direct and continuous measurement of on-chip currents is becoming increasingly common in interconnects at multiple locations on a die with real time feedback to the controller for efficient load management and/or load balancing and system performance optimization. The demand for good on-chip current sensors to support such applications is growing. These on-chip current sensors should have sufficient accuracy as well as compact area and low power consumption.

This thesis introduces on-chip current measurement method providing performance improvement as well as lifetime electromigration management. The inherent voltage drop across existing interconnects is used to determine the current flow rather than inserting shunts in the current-flow paths for creating voltage drops,. Current is measured with a MOSFET-only sensing circuit providing 9 bits of resolution with midrange current levels at the threshold where electromigration concerns become relevant. This current sensor can be used for sensing currents in either VDD or VSS busses and is targeted for use in power/thermal management units of integrated circuits. Simulation results show the DNL/INL of this sensor is within $+0.15/-0.3$ LSB. The current sensor is proved still useful with respect to local mismatches. The small area and low power dissipation make the structure suitable for multiple-site on-chip current measurements.

CHAPTER 1 INTRODUCTION

The computer industry has faced relentless pressure to continually improve performance and reduce cost. Through the 80's and 90's, much of this improvement was attributable primarily to advances in technology with smaller devices, faster clocks, and lower supply voltages. In recent years, however, power density in microprocessors and related SoC scale circuits, which is now dramatically increasing with decreasing feature sizes and increasing clock speeds, has emerged as one of the most daunting design challenges and has forced designers to abandon the decades-old approach of increasing clock frequencies to enhance performance. Nowadays, the greater emphasis was on both performance and life time improvements. In all electronics applications, it is important to prolong the chip's life as much as possible. And now, with the growing trend towards computing tasks, power dissipation has become one of the most critical factors in the continued development of the microelectronics technology. To continue to improve the performance of the circuits and to integrate more functions into each chip, feature size has to continue to shrink. As a result, the magnitude of power per unit area is growing and the accompanying problem of heat removal and cooling is worsening. Even with the scaling down of the supply voltage, power dissipation has not come down. Figure 1-1 shows the power density for several commercial processors. As it is shown in the figure, the trend is to increase the power density to levels where the cooling mechanisms are unlikely to be effective enough. As a result, today, it is widely accepted that power efficiency is a design goal as important as miniaturization and performance. The practice of low-power and multi-core design methodologies is being

adopted. Minimizing power dissipation and energy consumption calls for delicate effort at each abstraction level and at each phase of the design process [3].

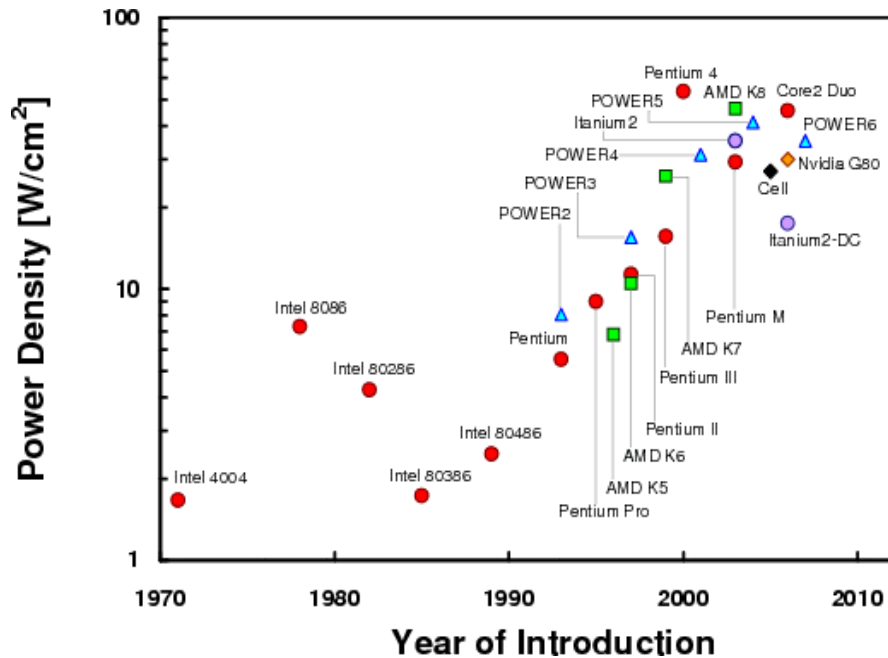


Figure 1-1: Processor power density over time[1]

Currently, Performance improvements in emerging processes are coming at the architectural level by using multi-core structures along with power management techniques that include combinations of measurement-driven dynamic supply voltage scaling, dynamic clock frequency scaling, and pre-calculated or dynamic task assignments. The first microprocessor, the Intel 4004[1], ran at a clock speed of 784 KHz while microprocessors of today run comfortably in the GHz range due to use of significantly smaller and faster transistors. Traditional single-microprocessor designs have reached a bottleneck in which doubling the number of transistors in a serial CPU results in only a modest increase in performance at a significant increase in energy. This bottleneck has motivated engineers to move into multi-core technology. With multi-core, we can achieve higher speed and

performance with acceptable power consumption [2], even the core-level frequency of a multi-core processor may be lower than that of a serial CPU. To continue this type of scaling performance, more and more cores are being placed on a die/chip; examples include an Intel Single-chip Cloud Computer with 48 cores [4] and a Tiler processor with 64 cores [5]. With this trend toward an increasing number of cores on a chip, the previous performance trend—doubling performance per processing element every two years or so—has ended. Instead, we can observe that doubling the number of processors can result in many programs doubling their speed [6]. The increase in performance has historically been consistent with Moore's law that states that the number of transistors on the processor die keeps doubling every eighteen months due the reduced transistor size with every successive step in process technology.

However, the price that we pay for getting higher performance has been rapidly increasing as well. For example, as Horowitz et al [7] show, the power cost for extracting a given additional amount of performance has been going up linearly with processor performance. This is a super-exponential increase over time. The measurement-driven power/thermal management approach is based upon thermal measurements at both multiple strategic sites in the cores as well as at other critical locations scattered throughout the remainder of the die. Indeed, Power Management has been identified as one of the Grand Challenges in the most recent International Technology Roadmap for Semiconductors (ITRS) [8]. In the context of multi-core technologies, the primary purpose of power and thermal management is to maximize system performance under power and thermal constraints mandating that the chip working condition must be maintained below an acceptable level

throughout the integrated circuit. The measurement-driven power management approach is based upon power measurements at multiple strategic sites in the cores as well as at critical lines throughout the remainder of the die.

1.1 THESIS OUTLINE

The research presented in this thesis primarily focus on design of current sensor that also implement electromigration monitoring. This thesis addresses IR drop to digital converter with two issues comprising current measurement and long time electromigration monitoring for multi-core system in 0.18 μm CMOS. This new current-sensor circuit structure places emphasis on reduction of area and power while maintaining accuracy needed for practically optimizing a measurement-driven power/thermal management strategy. A brief outline of each chapter follows.

Chapter 2 is comprised of an introduction the build-in current measurement and challenges with respect to the electromigration issue.

Chapter 3 includes a detailed examination of current-sensor operation that is comprehensively analyzed in chapter 4, with sub-circuits such as a comparator and current mode DAC which is also included in this analysis.

Chapter 5 presents conclusions from the work done described in this thesis as well as discussion of directions for future research.

CHAPTER 2 BACKGROUND

2.1 INTRODUCTION

Reliability of a product describes the probability that it functions as intended over a specified period of time. For an integrated circuit (IC), reliability represents a critical product specification under today's aggressive technology-scaling, even though it has always been very difficult and costly both to measure and to achieve in leading-edge technology. The work here was motivated by the potentially considerable benefits associated with efficient reliability evaluation and reliable circuit design. It is important to have a efficient method to provide the operating status of a circuit, and then provide insight into both performance and reliability information of the circuit we interested in.

2.2 IC RELIABILITY AND TESTING

Reliability of an item is defined by the International Electrotechnical Commission [I.E.C., 1974] as its ability to perform a required function, under stated conditions, for a specified period of time. The term reliability is also used to describe a reliability characteristic in terms of a probability of success or a success ratio [9]. In IC-manufacturing practice the reliability can be generally specified either by the lifetime over which an IC is expected to perform its designed functions, or by the failure rate, namely, the instantaneous probability that IC fails to perform its functions at a given time. IC reliability failures can occur due to either material wear-out or defects, and they occur after the ICs are delivered to customers. An overview is given by Figure 2-1, which depicts a typical IC reliability bathtub curve expressing failure rate as a function of product lifetime.

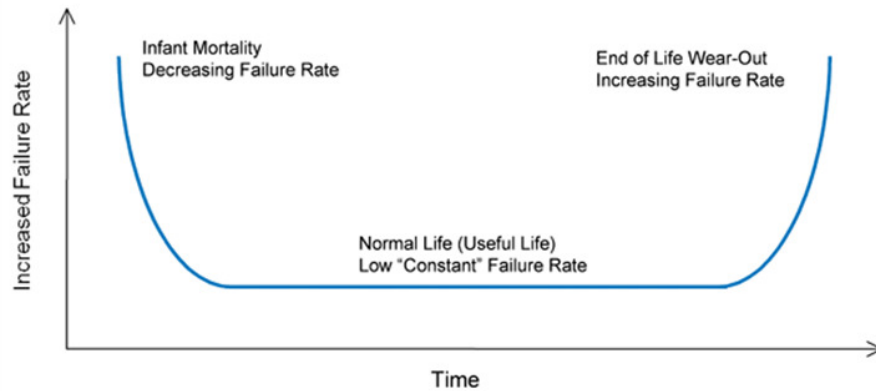


Figure 2-1: IC reliability bulb curve[9]

In Figure 2-1, the early infant mortality period is generally attributed to defective material [10]. In this stage the failure rate can be quite high, and usually a very expensive burn-in test is performed before product delivery to screen out severe-defective components. The next region of the curve is one corresponding to chance failure, and in this region the failure rate is low and nearly constant, representing the useful IC lifetime. Failures in this range are principally due to low-level residual defects or electrical overstress/electrostatic discharge events. A qualification test is customarily performed by IC reliability engineers to predict both the failure rate and the associated IC lifetime. The final increase in failure rate at the right-most part of the curve occurs as the result of intrinsic material wear-out. For a mature process this region may not actually be encountered because the IC product has already been replaced by a new one before it is entered.

For IC design, there is always a trade-off between reliability margins and performance. In order for ICs to be faster and smaller, feature size has been dramatically shrinking. For example, the feature size of an Intel processor has decreased from 3.0 μm to 0.09 μm in the past tens of years, and a significant increase in power consumption for

microprocessors, due to the increased operating frequency and transistor count, was reported from the same source.

Other key device dimensions such as oxide thickness and interconnect width have also decreased proportionally. The overall result is that, by accepting shrinking device dimensions and resulting high operating temperature, the IC has become much more vulnerable to failure mechanisms. Serious reliability challenges have been generated by aggressive technology scaling.

ICs are degraded by various failure mechanisms, with electromigration (EM) as one of the ruling mechanisms in terms of interconnect failure. Under elevated current density and temperature, EM can generate voids that can ultimately result in interconnect breakage in interconnect traces. For device degradation, hot-carrier and oxide wear-out are two major mechanisms. The former is initiated by a channel's electric field and can cause permanent oxide damage, resulting in degradation of parameters like threshold voltage shift. The latter is due to the oxide layer's electric field and can generate defects inside the oxide that could induce catastrophic oxide breakdown. These three failure mechanisms are major causes of IC failures and become more serious with scaled-down technology. With such rapid dimension-shrinking, if the power supply does not scale proportionally, virtually every aspect of the circuit becomes more fragile. Unfortunately, this behavior is exactly what has been happening. Interconnects elements also suffer from EM damage due to increased driving-current densities. The current density inside interconnects for Intel processor chips increases by a factor of 1.5 per generation for current and future technology. This serious fact has strongly suggested that further EM monitoring and improvement is definitely needed for our

technologies and even for material interconnections in other applications, as described in [11]. In this thesis, the emphasis is primarily on electromigration.

Most reliability tests consist of extremely expensive and time-consuming accelerated stress tests. Two major processes usually conducted by reliability engineers are burn-in tests to discover infant mortality, and qualification tests addressing the chance failure region, as shown in Figure 2.1. Both use elevated temperature and voltage to make ICs fail earlier, and they can take days or even months to finish. The accuracy of accelerated stress tests may, however, be doubtful, since parameters such as activation energy are usually determined under stress conditions completely different from actual use conditions. It has long been known that various physical defects may be generated inside the IC [12] due to uncontrollable processes present during fabrication.

2.3 ELECTROMIGRATION LIFETIME PROBLEM

Electromigration (EM) is an important failure mechanism inherent to silicon chips which is always used for metal reliability. Keeping pace with the shrinking of MOSFET physical dimensions, interconnecting layers and geometries of on-chip metallization both scale very quickly. This leads to higher current density flowing through interconnects and exacerbation of EM wearout effects on circuit performance and reliability. As a result, even though some new materials with better immunity to EM failures have been used to replace Aluminum (Al), as on-chip interconnects, EM is still a major reliability concern, and designers need accurate EM lifetime models as well as information to correctly predict device failure rates for long-life applications.

Electromigration (EM) has been considered a major failure mechanism in discrete solid-state devices and integrated circuits. Its classical definition refers to the structural damage caused by ion transport in metal thin films experiencing high current densities. EM damage occurs in the forms of voids and hillocks on interconnect traces, with voids being the major concern due to the accompanying increased current density. In addition to current density, temperature and material properties also play a critical role. As a major failure mechanism long known within the IC industry, EM is still with us today, and has been becoming an increasingly serious concern in terms of interconnect reliability with continuation of technology downscaling [13].

Besides runtime efforts to reduce the power consumption that increases internal temperature of a system, special attention must also be paid to runtime management needed to guarantee the expected system lifetime. We should investigate how runtime changes in parameters such as current density affect system reliability, and how to guide runtime management to maximize system performance without reliability violations. EM is a process of self-diffusion due to momentum exchange between electrons and atoms in the metal interconnects. As a result of electromigration, short or open circuit failures may occur due to the formation of hillocks or voids in interconnects. Several EM models are also available [14, 15]. In this thesis, the empirical current density and temperature dependent equation model by Black [16] that has been widely used in predicting interconnect lifetime related to EM failures is adopted.

2.3.1 Electromigration Lifetime Models for Current Density

A practical EM lifetime model must realize two important functions. First, it must identify critical stress parameters, provide guidelines to perform tests, and account for the relationships between test results and actual use conditions. Second, EM failure behaviors and physics of different metal structures must be taken into account, so that the test results can be extended to real and complex circuits to enable proper estimation of product reliability. Traditionally, these two functions have been treated separately, but new experimental and research work has led to a general model unifying these two aspects into one framework.

In a paper published by Black in 1969[16], the relationship between current density, electrical stress, and median-time-to-failure (MTF) associated with electromigration in aluminum interconnects was discussed. The original Black model is the first accepted EM lifetime model. It is an empirical model for grain-boundary-controlled EM failures and fits field data well. A generalized Black model has been proposed to characterize EM failure behaviors [17]:

$$MTF(J,T) = A_0 (J - J_{CRIT})^{-n} \cdot \exp\left(\frac{E_a}{kT}\right) \quad (2.1)$$

where A_0 is a constant based on the interconnect geometry and material, J is the average current density, k is Boltzmann's constant, T is temperature in Kelvin, and E_a is an experimentally determined activation energy (e.g., 0.7 eV for grain boundary diffusion in aluminum, from Wikipedia). The current exponent, n , has different values depending the actual failure mechanism. Black's equation is widely used in lifetime reliability analysis and design. For example, Hunter [18] derived a self-consistent allowable current density upper

bound for achieving a reliability goal by taking into account interconnect self-heating effects using Black's equation.

Evolution of J_{\max} (from device requirement) and J_{EM} (from targeted lifetime).
 J_{\max} will increase with frequency and reducing cross-section, while J_{EM} will scale with the product $w \cdot h$ according to EM lifetime dependence on wiring width.

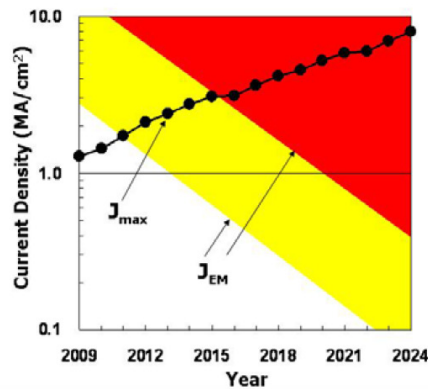


Figure 2-2: Evolution of J_{\max} (from device requirement) and J_{EM} (from targeted lifetime) [19]

Electromigration failures described with Black's equation determine the maximum current density (J_{EM}) that can safely flow in a wire in Figure 2-2. The most common metals used in today's ICs are aluminum and copper. Cu interconnects were introduced in 1997 in a damascene scheme to reduce wiring delay, but Al interconnects are still used for specific applications and at several levels in multilevel interconnects. Electromigration-limited current in Al interconnects can be determined using the usual EM physics rules [19].

J_{EM} is the maximum current density providing the targeted lifetime, and it scales with the product width and height. J_{\max} is defined as the maximum equivalent dc current expected to appear in a high-performance digital circuit divided by the cross-sectional area of an intermediate wire. The comparison of the evolution of J_{\max} and J_{EM} limited by the interconnect geometry scaling is shown in Figure2-3. J_{\max} increases with scaling due to

reduction in the interconnect cross-section and the increase in the maximum operating frequency. J_{\max} will exceed the J_{EM} limit of conventional copper interconnects.

2.4 BUILD-IN CURRENT SENSOR

To further extend IC operating life and optimize chip performance, we must better manage power consumption. Dynamic voltage and frequency scaling and low-power idle states have become standard techniques in modern computing systems. Several techniques have been proposed to reduced power consumption. however, they all have one need thing in common: the capability to continuously monitor the current/power management system to dynamically adjust the operating frequency or supply voltage, or block some processes which cause energy quotas to be exceeded.

Built-in current sensors (BICS) have been designed to provide useful on-chip current measurements for current testing. Since each BICS has limited resolution, the circuit under test (CUT) is required to separate into several testing blocks such that each block lies within the testing capability of BICS. Once a built-in current sensor (BICS) is fabricated on an IC chip, it cannot be removed after current tests are performed. This implies a relatively high silicon cost for built-in current testing without considering the increased design complexity. Therefore, area and design complexity are the most important aspects when considering the tradeoffs involved. Maly and Nigh established the feasibility of testing integrated circuits using on-chip monitoring of supply current (IDD) and introduced one of the first built-in current sensors (BICS) [20,21] there has been considerable research on developing on-chip current-measurement strategies. Although some of these research results have extended the

concept of on-chip current sensing to the testing of analog, mixed-signal, and RF circuits, most have been focused on hard-defect testing of digital circuits.

Current measurements have been used to screen potential IC defects such as shorts or opens between signals and power supply lines. Quiescent current test techniques (I_{DDQ}) have proved to be very effective as current testing techniques are increasingly challenged by the growing complexity of circuit designs and as CMOS has evolved to high-level background leakage currents and large variability. I_{DDQ} testing has been a very useful test screen for CMOS circuits. In ideal CMOS circuits the NMOS and PMOS circuits do not conduct at the same time, so at any given time there is no conducting path from VDD to VSS (or GND) and thus no current flow. In reality transistors are not ideal, and MOSFETs do simultaneously conduct for a short period of time when input values switch and capacitors are charged and discharged. Once input values have changed, the circuit settles down to a stable state called the quiescent state of the circuit, and I_{DDQ} is the total current that flows in a CMOS circuit when all nodes are in such a quiescent state. In ideal CMOS circuits I_{DDQ} should be zero. However, due to various transistor leakage mechanisms there is always a current flowing from VDD to VSS. In a fault-free circuit this current is typically very small, but in a circuit with a fault this current may appreciably increase. Thus, by measuring this current we can differentiate faulty from fault free circuits [22].

However, with each technology advance the background leakage of chips is rapidly increasing. As a result it is becoming more difficult to distinguish between faulty and fault-free chips using I_{DDQ} testing. For deep submicron technologies it is becoming even more difficult to draw a line between faulty and fault-free chips based on full chip I_{DDQ} . The

current-limit setting problem respect to I_{DDQ} testing involves deciding on a pass/fail current value that can discriminate between good and bad chips. Setting the threshold limit (I_{th}) to too high a high value may allow faulty chips to be shipped, while on the contrary, setting the limit too low may cause good chips to be discarded (a situation referred to as overkill). Figure 2.3 shows how the I_{DDQ} limit-setting problem is becoming more difficult with Deep Submicron (DSM) technologies. Region A represents the overkill situation and region B the test escapes situation.

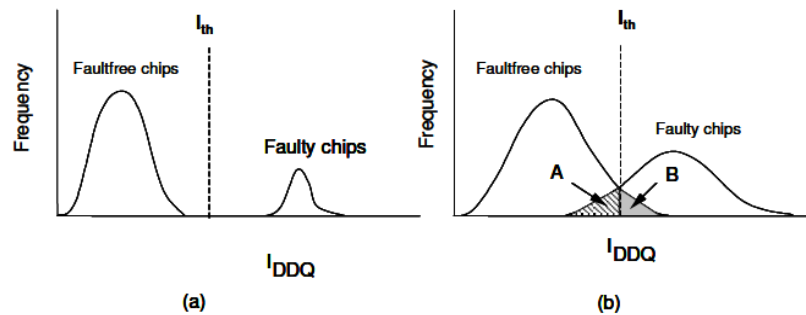


Figure 2-3: Limit Setting Problem for I_{DDQ} Testing[22]

On the other hand, the most damaging aspect of existing built-in current sensors is permanent performance degradation imposed on the CUT beyond the testing process. In many approaches have appeared in the literature, the current to be measured is first converted into a voltage signal by feeding it into a transistor. A typical BICS has a sensing element inserted in series with either the V_{dd} or G_{nd} power busses of the CUT. This sensing element may be a MOS transistor, a bipolar transistor, a resistor, or a diode, etc as shown in Figure 2-4[23]. Regardless of the type of sensing element used, a permanent load is imposed on the CUT's power supply line. Because of the voltage drop across such sensing elements and the

consequently increased parasitic load, significant performance degradation of the CUT may occur.

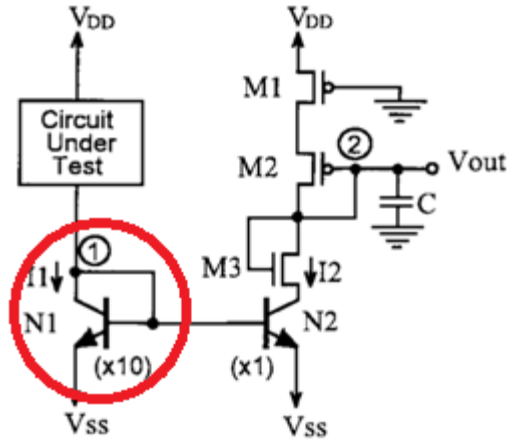


Figure 2-4 Current Shunt Insert Into Current Path

New on-chip current and power measurement techniques are motivated to provide solutions for the problems described above. The first BICS designs that claim to induce no performance degradation are described in [24]. The techniques described measuring very small voltage drops along a segment of the power supply line of the CUT. It demonstrated that the I-R voltage drop on IDD lines can be used to measure I_{DD} itself. This represented a major development in on-chip current measurement and its applications, and spawned considerable research activity. This concept offers a new direction for parametric measurement of currents in contrast to earlier work that focused on pass/fail current testing. With this new approach, an accurate matching of several transistor ratios and resistance ratios is critical and also particularly challenging in emerging processes where reduced feature sizes lead to increased variability.

CHAPTER 3 CURRENT SENSOR ARCHITECTURE

3.1 INTRODUCTION OF PROPOSED CURRENT SENSOR

Power consumption and die heating are of major concern in high-density high-speed integrated circuits since performance of modern IC designs is limited by power consumption and thermal issues. Performance improvements in emerging processes are being accomplished at the architectural level by using multi-core structures along with power management techniques that include combinations of measurement-driven dynamic supply-voltage scaling, dynamic clock-frequency scaling, and pre-calculated or dynamic task assignments. Direct and continuous measurement of on-chip interconnect currents is becoming increasingly common at multiple locations on a die to support power management and system performance optimization. Traditional off-chip current measurement methods are not practical for continuous-time monitoring of current in critical interconnects distributed throughout a chip, and demand for good on-chip current sensors needed to support these applications is growing. Requirements for these on-chip current sensors include not only good accuracy but compact area and low power consumption as well.

For the purpose of electrothermal stress management, a pass/fail or over-the-limit indicator is not sufficient. On-chip real-time current measurement at a modest accuracy level in those interconnectors where the current density is near the electromigration limit is important for developing a good power/thermal management strategy. A current sensor with more than 6-bit efficient resolution and a current range of 200% of the maximum acceptable operating current should provide sufficient resolution to support the power/thermal management system. The proposed current sensor will be non intrusive, that is, no switches,

transistors, or resistors will be inserted into the path of current flow. The natural voltage drop due to current flowing through parasitic metal resistance will be detected, amplified, and converted by the current sensor.

3.2 BASIC CONCEPTS AND WORKING SCENARIOS

The basic concept for the proposed current measurement strategy is illustrated in Figure 3-1. The strip represent the metal interconnect carrying a critical current to be measured and the boundary may have some irregular edge due to process variation. As the current flows through this metal wire, an inherent IR drop will be created between two locations that are separated by a known distance. It has the advantage of converting current into voltage in a linear manner that inherently follows Ohm's law:

$$V = I \cdot R \quad (3.1)$$

With typical metal wire resistivity and typically metal line width, a voltage drop of a few mV can be expected over a distance of 100 μm in interconnects where concerns about electromigration exist.

For electromigration-safe operation of aluminum at 125°C, the average DC voltage drop must be less than approximately 0.1 mV/ μm :

$$V(L) < V_{safe}(L) = \frac{0.1\text{mV}}{\mu\text{m}} \cdot L \quad (3.2)$$

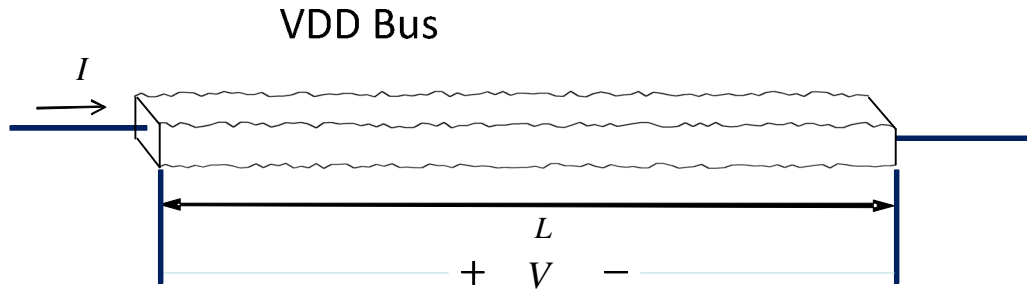


Figure 3-1: Typical IR drop and working range of on-chip current sensor

In an aluminum interconnect of length $100\mu\text{m}$, $V(L)$ is around 10mV , and this voltage is large enough to measure practically without adding a current shunt. This result is independent of both the thickness and the width of interconnect. Sensing IR in interconnect of such length requires no “shunt” for measuring current, with a voltage drop of 10mV occurring near the electromigration threshold.

3.2.1 Instant Current Measurement

Within a short periods, resistance of interconnect R could be considered as constant value. Therefore, from (3.1), we can get

$$V \propto I \quad (3.3)$$

In multi-core circuit systems, computation tasks/loads are typically distributed by a controller. During regular operation, some cores are operating at maximum capacity while others are either handling moderate loads or idling. High-load cores are typically pushed to the limit with respect to core temperature and current density, and the die may be stressed. Hence, it is important to sense the health of the cores and balance the load using an effective power management system. This current sensor has been designed to measure currents that

the vicinity of the working threshold and give instant feedback to Power Management Unit. Detailed instant current measurement principal will be introduced in Chapter 4.

3.2.2 Wear Out Lifetime Electromigration Monitoring

Measured-structure failure is always defined as a percentage increase in resistance, commonly 20%, plotting with a resistance versus time plot during an electromigration test. The 20% resistance increase is a commonly-used figure of merit selected as a failure criterion by the industry. The size of the jump in resistance depends on the length of the grain electromigrated away and the redundancy features of the test structure. Although the shape of the curve can vary substantially, and depends strongly on the metallization and structure being tested, The failure- time distributions vary to some degree depending on the structure, but the general trends and final conclusions are not changed [25]. If IR drop is sampled at same situation and environment such as applying particular tasks on chips at 3AM in the morning, current I could be considered as constant value. Then

$$V \propto R \quad (3.4)$$

and the resistance information could be used in electromigration monitoring.

On the other hand, the Black model introduced above is widely used throughout the industry for the accelerated-lifetime testing needed to predict product reliability. But when J and T become time-dependent, which naturally happens when a multi-core processor is used, this equation no longer predicts the MTF. If J and T are time-dependent, the MTF can be expressed as

$$MTF = \int_{t=0}^{\infty} t \cdot f(J(t), T(t)) dt \quad (3.5)$$

where f is a time-dependent failure density function. Unfortunately, even the original work of Black [19] did not discuss the form of this failure density function nor consider the time-dependence of J and T . If the device has been deployed, the failure density function actually changes with time if a device has not yet failed. In this situation, the MTF at any time t_1 can be expressed as

$$MTF = \int_{t=t_1}^{\infty} t \cdot f_c(t_1, J(t), T(t)) dt \quad (3.6)$$

where the function f_c becomes a conditional failure-density function. This functional form for the expression for the MTF applies regardless of whether the interest is electromigration, dielectric degradation, or other failure mechanisms.

Black described the process of electromigration and presented a model of this effect in which current density is exponentially related to an decrease in MTF. It follows from this exponential relationship that small increases in current density can result in dramatic decreases in the MTF and that, if the unit is operated at a current density higher than some critical threshold, the MTF will be unacceptably short. Thus, if sensors can accurately tell when a critical current density or critical resistance is met, electromigration failure can be probabilistically reduced to an acceptable level or, equivalently, the MTF of a processor can meet a predetermined target value.

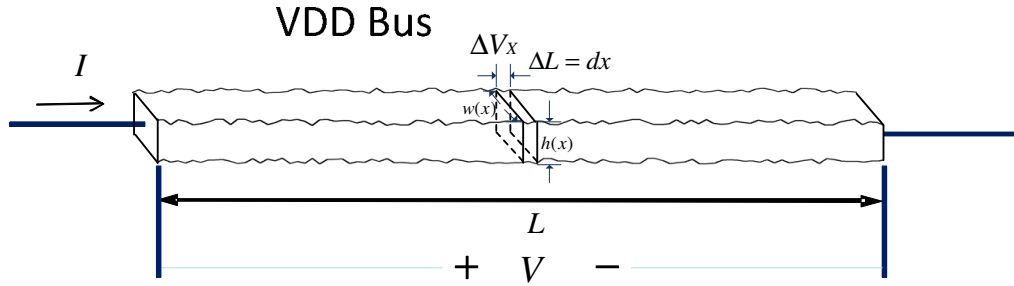


Figure 3-2: Voltage Drop Partition Along Interconnect

Unfortunately, there is no practical way that we can know the current density information at specific point. However, we are able to get such kind of information through the voltage drop $V(L)$ if we separate it into small partitions as shown in Figure 3-2. Then the total voltage drop $V(L)$ could be the summation of $\Delta V(x)$:

$$V(L) = \Sigma \Delta V(x) = \int_0^L dV(x) \quad (3.7)$$

Since the current flow through the critical interconnect is same at different location, then the formula above can be transform into

$$\begin{aligned} V(L) &= \int_0^L I(x) dR(x) \\ &= \int_0^L J(x) A(x) dR(x) \end{aligned} \quad (3.8)$$

where A and R denotes cross-section area ($w(x) \cdot h(x)$) and resistance corresponding.

As we know

$$R = \rho \frac{L}{A}, \quad (3.9)$$

Where ρ is the electrical resistivity of the material which can be considered as constant if temperature doesn't change. Then equation (3.8) can be expressed as:

$$V(L) = \rho \int_0^L J(x) dx \quad (3.10)$$

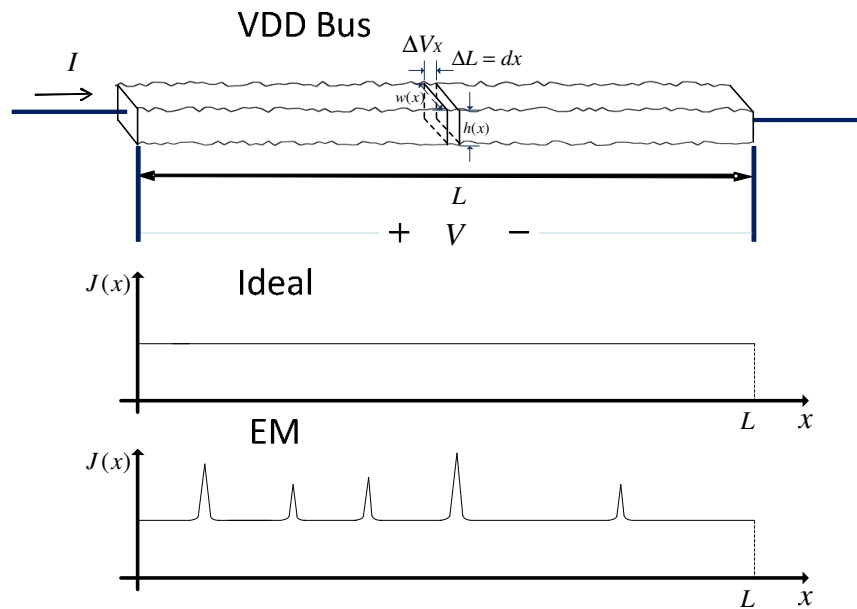


Figure 3-3: Typical IR drop and Working Range of On-Chip Current Sensor

Therefore, the integral of current density along the certain distance of interconnect can be got from measuring the total IR drop. can be implemented as a difference function in the current sensor with a sampling period T_x which could be as long as days or even months. The $J(kT_x)$ values will be obtained from the interconnect voltage drop measurements then compared with ideal case as shown in Figure 3-3. The time-dependent parameters that characterize the functional form of the density function will be periodically updated and stored in external memory so that history information is not lost if the multi-core processor

is powered down. The sequence of current density samples is also able to give us thermal information since power density and heat are highly related.

With a measurement-driven real-time wear-out model for the time-cumulative effects of stress the remaining life of a system can be estimated through the conditional time-history dependent failure density function and, further, related to the time-dependent system reliability. Measurements obtained from the current sensors can be used as inputs to the real-time wear-out model. During normal operation, some “wear” will occur whenever the current approaches or exceeds the electromigration threshold. Wear-out mechanisms, which invariably establish constraints on the relationship between performance and temperature at the design stage, can be practically reflected in a measurement-driven power/thermal management unit that can dramatically improve the reliability of a system throughout its target life.

In one word, accumulative “wear” can be assessed by creating a time-history of the IR drop data. Electromigration monitoring for the purpose of assessing and managing “wear” will require many IR drop measurements over long time intervals measured in months or perhaps even years.

3.2.3 Offset Calibration

Of course the process variations and device mismatch will be interpreted as a digital offset error it will be assumed that taking calibration will be used to eliminate or minimize this error. C_1 and C_2 are considered as reference value in the equation below.

$$\begin{cases} I = D \times C_1 + Offset_1 \\ R = D \times C_2 + Offset_2 \end{cases} \quad (3.11)$$

3.3 CURRENT SENSOR ARCHITECTURE

A new on-chip current sensing technique requiring no current shunt but instead measuring the voltage drop across an existing interconnect is introduced in this thesis. The proposed technique is directed toward a MOSFET-only non-intrusive method that supports power/thermal management circuits and that can provide real-time current levels suitable for electromigration management. This low-power current sensor has good accuracy, small area, and is suitable for multi-site current monitoring in multi-core systems. This sensor design has been derived through simulation results for a circuit designed in a TSMC 0.18 μm CMOS process.

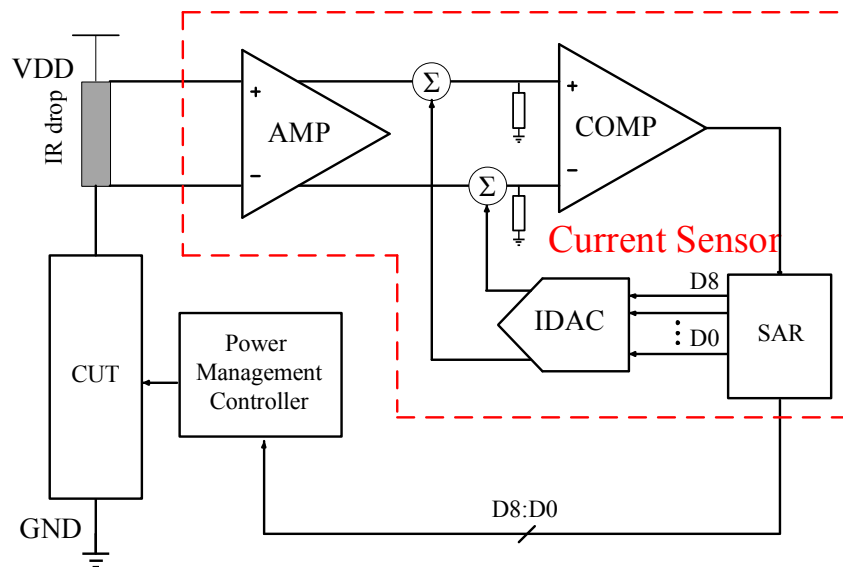


Figure 3-4: Top Level of On-chip Current Sensor

The distributed I-R voltage drop due to current flowing through parasitic metal resistance will be detected, amplified, and converted into digital codes.

The resistors in the figure represent the metal interconnect resistance carrying a critical current to be measured. As the current flows through the metal wire, an IR drop will be created between two locations separated by a known distance.

A Successive Approximation Register (SAR)-based analog-to-digital converter has demonstrated great amenability to deeply-scaled CMOS technologies due to its inherent architectural advantages. Compared to flash converters, increasing the speed/accuracy of a SAR structure requires only the design of faster or more accurate components rather than an increase in overall complexity. Hence, the SAR converter is much less complex and more efficient in terms of both area and power consumption. Compared to pipeline architecture, SAR has the advantage of being free from necessity of designing high-performance op amps in sub-100nm CMOS. Also, the traditional 8 - 10b accuracy conversion rate specifications territory long dominated by pipeline ADCs [41] has come well into the range of SAR ADCs as development of fast thin-oxide sub-100nm transistors supports the design of relatively high-speed comparator and digital circuits with lower power supply voltages. Finally, the fact that a large fraction of the SAR ADC is digital circuitry renders this architecture even more attractive for use in deeply-scaled CMOS technologies.

The successive-approximation algorithm can be easily implemented through charge redistribution in a binary-weighted capacitor array. Assuming all capacitors are ideally matched, the resolution of a SAR ADC is determined by the total number of capacitors included in the DAC. To increase the resolution, the simplest approach is to increase the number of capacitors. However, due to the binary-weighted nature; this approach will cause an exponential increase in the DAC's area consumption. Additionally, the ratio of the MSB

component to the least significant bit (LSB) component becomes impractically large as a result of increases in capacitor array size, so the required matching accuracy between the two may be significantly degraded. A mismatch ultimately limits the maximum achievable resolution of this SAR ADC.

A MOSFET-only SAR structure is proposed in our design to achieve acceptable resolution in a SAR converter without sacrificing too much area,. The structure consists of four principal subcircuits, including:

1. An analog voltage comparator to compare the voltage before and after the measured IR drop and present the the result of the comparison to the SAR;
2. A successive approximation register array designed to supply an approximate digital value of V_{in} to the internal current mode DAC(IDAC) and to store the output of the digital output;
3. A pair of common gate amplifiers to amplify the very small voltage drop;
4. An internally-referenced complementary current mode DAC (IDAC) to supply the comparator with an analog current equivalent of the digital code output of the SAR for comparison with the IR drop.

The processing procedure is described as following. Before real-time measurement, resistance calibration would be executed for the determined distance between two selected locations. A certain bandgap current reference would be applied so that an accurate parasitic metal resistance could be determined. Since the calibration is performed only in background during power-on, no extra latency or clock cycles are introduced due to the sensor's normal

operation. The distributed I-R voltage drop due to current flowing through parasitic metal resistance will first be detected and amplified by a paired common-gate amplifier because the IR drop is a truly very small signal. Next, for each conversion, the SAR converts a continuous IR drop waveform into a discrete digital representation via a binary search through all possible quantization levels before finally converging upon a digital output. The successive approximation process is initialized in mid-range so that the most significant bit (MSB) is initially equal to a digital 1. This code is fed into the IDAC, which then supplies two equivalent of this digital code into the comparator circuit for comparison. If this node voltage is greater than another, the comparator causes the SAR to reset this bit; otherwise, the bit is left at the 1 level. Then the next bit is set to 1 and the same test performed, continuing this binary search until every bit in the SAR has been developed. The resulting code is the digital approximation of the input voltage and is finally output by the IDAC at the end of the conversion process.

In order to make on-chip real-time current measurement at a modest accuracy level, this current sensor is designed with 9-bit resolution over a range of from 25% to 200% of the electromigration threshold, which should provide sufficient information for adequately supporting the power/thermal management system.

In addition, this structure could easily be worked as threshold current indicator with rarely changing as replace the SAR logic with latch. This indicator could tell power manager unit if current is exceed the threshold value and is shown in Figure 3-5.

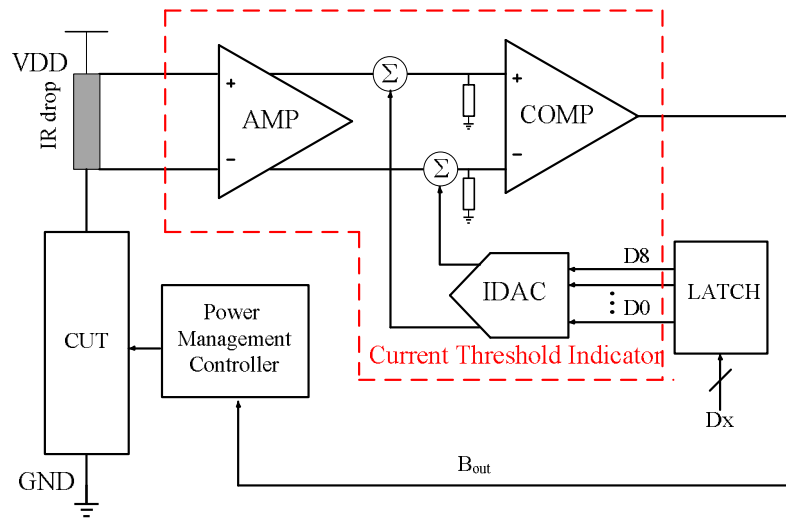


Figure 3-5: Top Level of Threshold Current Indicator

3.4 LIMITING PROBLEM

With inaccurate current sensors, on-chip power management systems cannot optimally manage performance to avoid stress-induced damage such as electromigration. If current sensors could be made small enough with sufficiently low power requirements while still retaining acceptable accuracy, multiple sensors could be placed at critical locations in each core and in critical paths where heating is likely to cause excessive stress and wear on the integrated circuit. With such real-time current and resistance information, combined with temperature data, power management algorithms that will maximum integrated-circuit performance while still maintaining target reliability and lifetime metrics can be developed. With better power and temperature sensors integrated into the power management unit (PMU), to obtain increased power and temperature information, cores can be better utilized, and the load spread to multiple cores without overloading either any particular core or any critical sub-units within a core or between cores. In this thesis, focus will be restricted to the

current sensors themselves. These sensors are intended to provide enabling technology for \ designers having responsibility for either PMU or system architecture design.

3.4.1 Comparator Offset

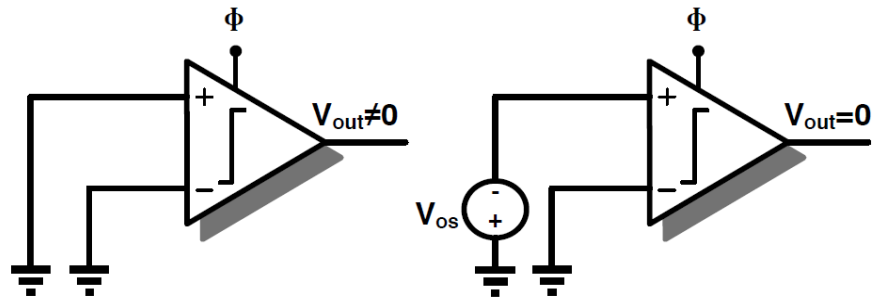


Figure 3-6: Comparator input referred offset

The input-referred offset voltage of the comparator is by definition the differential input voltage required to make the output voltage zero, as shown in Figure 3-6, and is mainly a result of the mismatch and process variation of the two comparator input transistors. Pelgrom[29] had shown that distribution of V_t is inversely proportional to the square root of the transistor area (WL), as shown in eq. 4.15, where the proportionality constant A_{VT} is technology dependent. Therefore, as CMOS scales into the nanoscale region, offset has become a more critical issue in analog circuit design.

$$\sigma_{\Delta V_T} = \frac{A_{VT}}{\sqrt{WL}} \quad (3.12)$$

The offset in the core comparator of the current sensor may limit the system's performance because any input signal that is smaller than the input offset will generate an erroneous code, so comparator offset needs to be limited to less than 0.5LSB before SAR conversion begins. There are many well-developed techniques for cancelling the comparator

offset in switched-capacitor circuit techniques, including input offset storage and output offset storage [26].

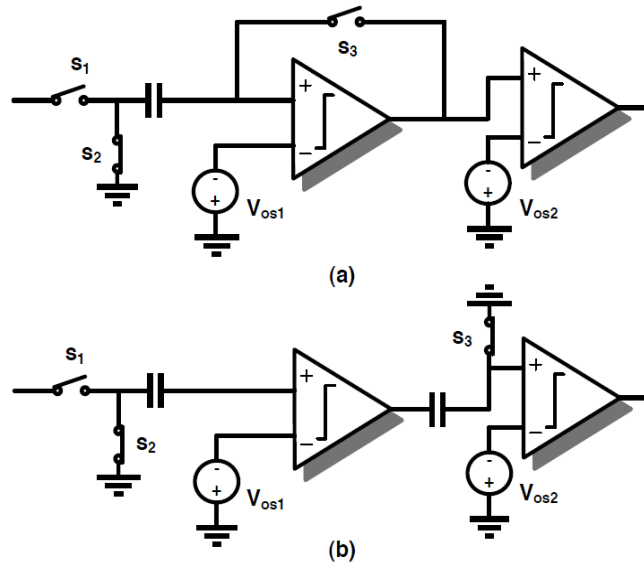


Figure 3-7: (a) Input offset storage (b) output offset storage

In the input offset storage technique, depicted in Figure 3-7(a), cancellation is performed by closing a unity-gain loop around the preamplifier and storing the offset on the input-coupling capacitors. In the output-offset storage technique of Figure 3-7 (b), the offset is cancelled by short-circuiting preamplifier inputs and storing the amplified offset on the output coupling capacitor. However, these traditional switched-capacitor techniques face challenges posed by a deeply-scaled CMOS process. For example, a capacitor will always occupy a large layout area, which is unacceptable in our sensor design.

CHAPTER 4 CIRCUIT IMPLEMENTATION AND SIMULATION

RESULTS

The previous chapter reviews the architecture and design challenges of an area-efficient build-in current sensor. This chapter focuses on a transistor-level implementation of this current sensor. It begins with top level of the circuit technique describing how an analog IR drop is linearly converted into digital code. It will then focus on the designs of the important sensor building blocks such as a comparator and supporting circuits that overcome major resolution-limiting factors: comparator offset, and $1/f$ noise. Since a large portion of the ADC is digital, mixed-signal design methodology and digital synthesis will also be discussed. Finally, the chapter will conclude with detailed simulation results.

4.1 TOP LEVEL OF CURRENT SENSOR

The top transistor-level implementation for the proposed current-measurement strategy is illustrated in Figure.1. The distributed I-R voltage drop due to current flowing through a parasitic metal resistance will be detected, amplified, and converted to a digital code.

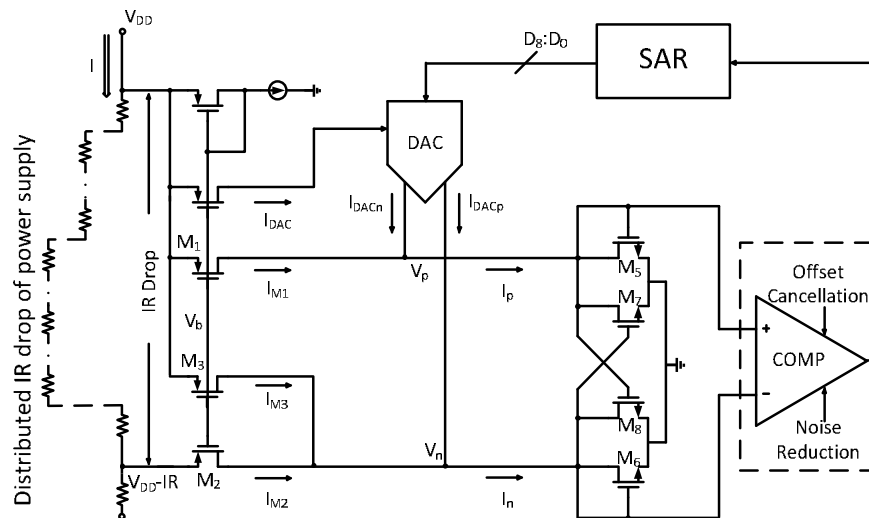


Figure 4-1: Circuit of on-chip current sensor

To amplify this small signal and bring it to a manageable level, an open-loop common-gate PMOS amplifier comprised of M1-M2 is used. The gate voltage of the amplifier will be biased at a constant voltage V_b generated by a simple current source. Hence, the voltage drop across the metal wire will serve as a small-signal input to the gain stage. The amplifier has diode-connected NMOS devices M5-M6 with a positive-feedback pair M7-M8 providing a load to enhance the small-signal voltage gain. The V_p and V_n nodes in the schematic are self-biased through the diode-connected devices. The small-signal gain can be expressed as $g_{m1}/(g_{m5} - g_{m7})$. Under normal operation, the currents flowing through the two sides are each less than $100\mu\text{A}$, relatively small values compared with power-supply currents of typically several amps. With the gain stage, the full-scale output signal going into the comparator will be in the hundreds-of-millivolts range and therefore near the electromigration limit, which greatly relaxes the comparator's resolution requirement.

To reduce both circuit complexity and silicon area, a successive-approximation register (SAR)-based method is used to implement a DAC in our structure. To further reduce area, the main DAC in the design is a MOSFET-only R-2R ladder [27]. Sizing of the DAC is based on matching requirements. The output currents of the DAC will balance the current difference between the two drain currents in the common gate input pair. Offset cancellation and noise-reduction techniques are applied on the comparator and the SAR logic driven by the averaged output of the comparator. A binary search is used to equate V_p and V_n or, equivalently, to equate I_p and I_n . When this condition is reached, we obtain the relationship

$$I_{DAC_p} + I_{M_2} + I_{M_3} = I_{M_1} + I_{DAC_n} \quad (4.1)$$

I_{DAC_p} and I_{DAC_n} can be expressed as

$$I_{DAC_p} = \frac{D}{2^n} I_{DAC} \quad (4.2)$$

$$I_{DAC_n} = I_{DAC} - \frac{D}{2^n} I_{DAC}. \quad (4.3)$$

where I_{DAC} is the DAC's reference current and D is the decimal equivalent of the 9-bit binary word $D_8:D_0$. Assume that all PMOS transistors have the same threshold voltage and neglecting channel-length modulation, V_b is a fixed dc bias voltage, R is the metal resistance, and I is the current to be measured.

The square-law model is often used to model the strong-inversion operation of a MOS transistor, A straightforward analysis of the circuit structure of Figure 4-1 based upon the standard square-law model of the MOS transistors follows. In this analysis, it will be assumed that all transistors are operating in the strong-inversion saturation region. The current on each flow can be given by

$$I_{M_1} = \frac{\mu_p C_{ox} W_1}{2 L_1} (V_{DD} - V_b - V_{TH})^2 \quad (4.4)$$

$$I_{M_2} = \frac{\mu_p C_{ox} W_2}{2 L_2} (V_{DD} - I \cdot R - V_b - V_{TH})^2 \quad (4.5)$$

$$I_{M_3} = \frac{\mu_p C_{ox}}{2} \frac{W_3}{L_3} (V_{DD} - V_b - V_{TH})^2 = I_{DAC} \quad (4.6)$$

Substitute into the above equation with (3.3), we obtain

$$\begin{aligned} \frac{D}{2^n} I_{DAC} - \left(I_{DAC} - \frac{D}{2^n} I_{DAC} \right) + I_{M_3} = \\ \frac{\mu_p C_{ox}}{2} \frac{W_1}{L_1} (V_{DD} - V_b - V_{TH})^2 - \frac{\mu_p C_{ox}}{2} \frac{W_2}{L_2} (V_{DD} - I \cdot R - V_b - V_{TH})^2 \end{aligned} \quad (4.7)$$

Assuming that $(W/L)_1 = (W/L)_2$,

$$\begin{aligned} \left(\frac{2D}{2^n} - 1 \right) I_{DAC} + I_{M_3} = \frac{\mu_p C_{ox}}{2} \frac{W_{1,2}}{L_{1,2}} [(V_{DD} - V_b - V_{TH}) - (V_{DD} - I \cdot R - V_b - V_{TH})] \cdot \\ [(V_{DD} - V_b - V_{TH}) + (V_{DD} - I \cdot R - V_b - V_{TH})] \end{aligned} \quad (4.8)$$

This can be further simplified to

$$\left(\frac{2D}{2^n} - 1 \right) I_{DAC} + I_{M_3} = \frac{\mu_p C_{ox}}{2} \frac{W_{1,2}}{L_{1,2}} I \cdot R \cdot [2(V_{DD} - V_b - V_{TH}) - I \cdot R] \quad (4.9)$$

For a nominal IR drop of 10mV at the electromigration limit, we could take

$V_{DD} - V_b - V_{TH} \gg IR/2$, so

$$[2(V_{DD} - V_b - V_{TH}) - I \cdot R] \approx 2(V_{DD} - V_b - V_{TH}) \quad (4.10)$$

(4.1) can be simplified to

$$\left(\frac{2D}{2^n} - 1 \right) I_{DAC} + I_{M_3} = \frac{\mu_p C_{ox}}{2} \frac{W_{1,2}}{L_{1,2}} I \cdot R \cdot 2(V_{DD} - V_b - V_{TH}) \quad (4.11)$$

Since $I_{DAC} = I_{M_3}$, we can obtain

$$\frac{2D}{2^n} I_{DAC} = \frac{\mu_p C_{ox} W_{1,2}}{2 L_{1,2}} \cdot I \cdot R \cdot 2(V_{DD} - V_b - V_{TH}) \quad (4.12)$$

Substitute from (4.6), it thus follows that

$$\frac{2D}{2^n} \frac{\mu_p C_{ox} W_3}{2 L_3} (V_{DD} - V_b - V_{TH})^2 = \frac{\mu_p C_{ox} W_{1,2}}{2 L_{1,2}} \cdot I \cdot R \cdot 2(V_{DD} - V_b - V_{TH}) \quad (4.13)$$

Simplifying to get

$$IR = \frac{D}{2^n} \frac{(W/L)_3}{(W/L)_{1,2}} (V_{DD} - V_b - V_{TH}) = \frac{D}{2^n} V_{ref} \quad (4.14)$$

$$\text{Where } V_{ref} = \frac{(W/L)_3}{(W/L)_{1,2}} (V_{DD} - V_b - V_{TH}) \quad (4.15)$$

As (4.14) demonstrates, the IR drop is proportional to the digital word. The full-scale reference voltage V_{ref} can be set by adjusting the size ratio of M_3 and $M_{1,2}$. For current measurement, certain bandgap current sources will first be applied to the power-supply line for resistance calibration. Since the calibration is only performed in background during power-on, no extra latency or clock cycles are introduced to the sensor's normal operation. After calibration, the relationship between the IR voltage drop and the actual power supply current will be known, so real-time current measurement data can be obtained and used to improve the effectiveness of overall power and thermal management algorithms.

Table 4-1: Transistors size for current sensor

Transistor	Size (μm)
M1	$W_1/L_1 = 20 \times 1/6$
M2	$W_2/L_2 = 20 \times 1/6$
M3	$W_3/L_3 = 1 \times 1/6$
M4	$W_4/L_4 = 1 \times 1/6$
M5	$W_5/L_5 = 2 \times 0.22/0.18$
M6	$W_5/L_5 = 2 \times 0.22/0.18$
M7	$W_5/L_5 = 2 \times 0.22/0.2$
M8	$W_5/L_5 = 2 \times 0.22/0.2$

4.2 SUB-CIRCUITS

4.2.1 Comparator

The dynamic comparator is a critical component in this system. First of all, it must have an extremely large open-loop gain ($> 73\text{dB}$) in order to amplify a LSB voltage of less than $400\mu\text{V}$ all the way up to digital level (1.8V) in a relatively short time period ($\sim 50\text{nS}$). However, the low inherent gain associated with $0.18\mu\text{m}$ transistors limits the maximum achievable gain for one amplifier stage, and the reduced power supply voltage prevents use of the cascade technique for gain enhancement. To reduce the offset voltage and settling time of the comparator, a preamp is used at its input. This preamp uses diode-connected PMOS devices and a positive-feedback PMOS pair as its load to enhance the gain. This preamp stage has advantages of high differential-mode gain and common-mode rejection. It also doesn't require common-mode feedback because the output quiescent point is self-biased through the diode-connected devices.

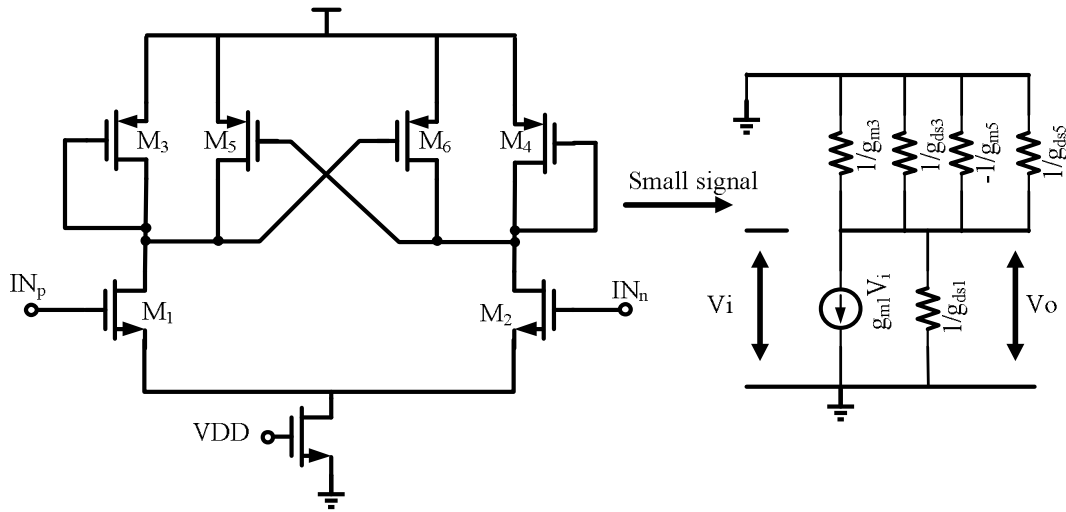


Figure 4-3: Comparator preamp and small signal circuit diagram

In this design, a preamp is applied to achieve an overall gain of about 50dB. In addition to this moderate gain, the preamp stages also have reasonable bandwidth to allow fast settling.

In latch mode, Clk goes low, disabling the input pair and turning off M13- M14. The positive feedback regenerates the analog signal into a full-scale digital signal. Complete sizing information is shown in Table 2 below.

Table 4-2: Transistors size for comparator

Transistor	Size (μm)
M1 & M2	$W_{1,2}/L_{1,2} = 3/0.18$
M3 & M4	$W_{3,4}/L_{3,4} = 0.22/0.18$
M5 & M6	$W_{5,6}/L_{5,6} = 0.22/0.2$
M7 & M8	$W_{7,8}/L_{7,8} = 0.22/0.18$
M9 & M10	$W_{9,10}/L_{9,10} = 2/0.18$
M11 & M12	$W_{11,12}/L_{11,12} = 0.22/0.18$
M13 & M14	$W_{13,14}/L_{13,14} = 0.22/0.18$
M15	$W_{17}/L_{17} = 0.22/0.44$

4.2.2 Offset Cancellation Technique

The accuracy of a current sensor relies heavily on the performance of the comparator, because any input signal smaller than the input offset will not generate a correct decision to drive the SAR logic. The dynamic comparator is vulnerable to device mismatch, so the comparator offset must be limited to less than 0.5LSB before the SAR conversion begins. The input referred offset of the comparator is dominated by the offset voltage of the preamp.

In submicron CMOS, the comparator offset can vary from a few mV to more than 100 mV, depending on transistor size. The design strategy used this research is to first optimize comparator sizing and then use Monte Carlo analysis to determine the distribution of the input offset. Because of a lack of area-efficient specification, capacitor cancellation is not capable of satisfactory performance in our design.

The input-referred comparator offset results from a combination of preamp input offset and latch offset. During comparator design, the offset was analyzed by running separate Monte Carlo analyses on the preamp and the latches. Figure 4.4 shows the Monte Carlo simulation results for the comparator preamp.

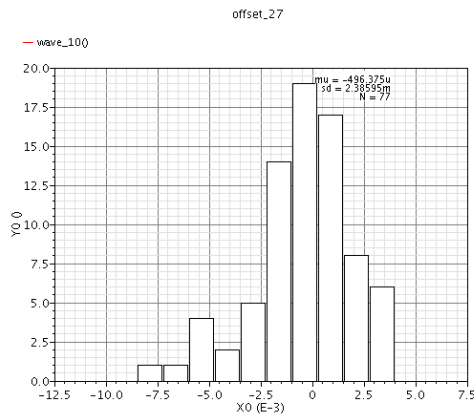


Figure 4-4: Monte Carlo analysis of preamp offset

Comparator input offset is determined by dividing the accumulated gain from the previous stage. The total input-referred offset of the final design is expressed in eq. 4.17 as

$$V_{os}^2 = V_{os,preamp}^2 + \frac{V_{os,latch}^2}{A_{v,preamp}^2} \quad (4.17)$$

Since the preamp's gain is high, the above equation indicates that the total comparator offset is approximately equal to the preamp offset. The estimated offset value, based on the Monte Carlo analysis, is about 2.5mV.

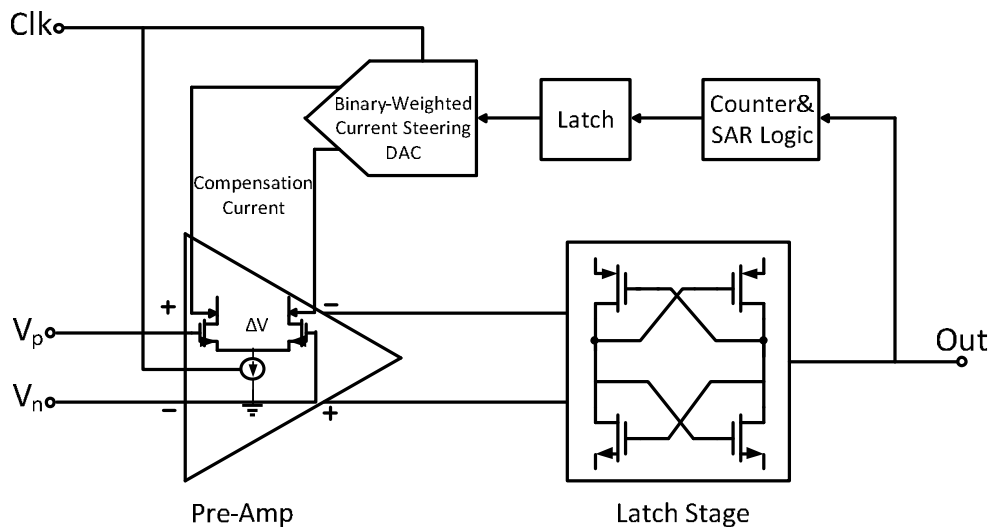


Figure 4-5: Offset cancellation technique

To correct this offset, a successive-approximation-based offset-cancellation method was implemented to control the input differential pair's current, as shown in Figure. 4.5. When some mismatch occurs, it creates a difference voltage between the input pair terminals, and the circuit will lose balance. Driven by SAR logic, the offset DAC successively generates compensation current. In the offset-cancelling phase, the two inputs are connected together. Without offset, there is randomly either 'high' or 'low' voltage at the output node if the two

inputs are equal, and the output won't change until the input difference reaches the offset voltage. Therefore, the count and binary search process injects appropriate compensation current into the differential pair to achieve approximately 50% 'high' and 50% 'low' values at the output while taking a certain amount of time for the comparison. With this approach, the offset will be cancelled out if the output of the comparator provides a random value. This offset-cancellation technique is only enabled when the preamp is on. It doesn't insert any compensation current into the latch stage, so it doesn't affect the latch operation in latch mode. If multiple current sensors are implemented, the SAR logic used for offset cancellation can be shared between the sensors. In the offset-cancellation scheme, a maximum of $\pm 20\text{mV}$ offset can be accommodated.

Figure 4.6 shows the transient simulation waveform of the comparator. At a reset frequency of 20MHz, the comparator is able to resolve a minimum input voltage of $\sim 200\mu\text{V}$ after offset cancellation.

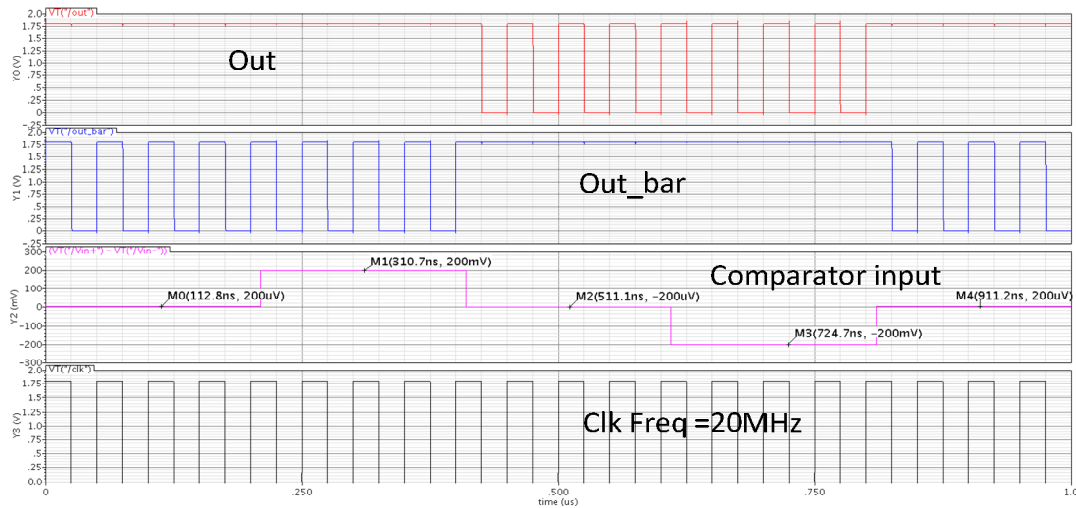


Figure 4-6: Comparator transient simulation waveform

4.2.3 Noise Reduction On the Current Sensor

As with any analog and mixed-signal circuit, accounting for the effects of noise on the performance of the current sensor must be considered as part of the design process. In the case of a current sensor where voltage drop information is converted into digital output voltage, the noise voltage at the output would ultimately be converted into noise-containing data. Therefore, the comparator in this design incorporates a chopping technique for noise reduction. Since the current sensor will be operating relatively slowly, $1/f$ noise is expected to dominate noise performance. The differential input pair should first have relatively large size to minimize $1/f$ noise effects. Figure 4.7 provides a conceptual diagram of the $1/f$ cancellation scheme integrated into the comparator.

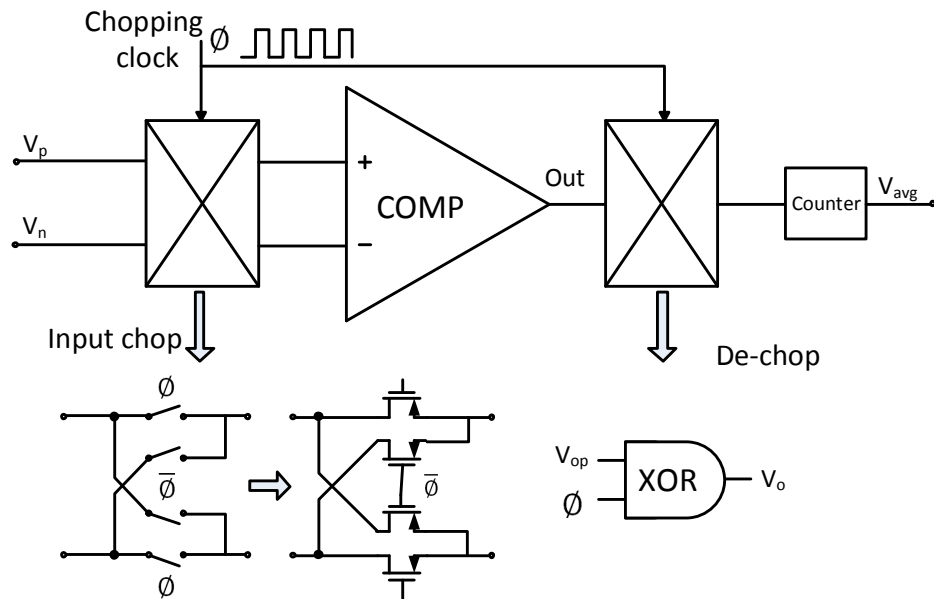


Figure 4-7: Chopping technique to reduce $1/f$ noise

As shown in Figure. 4.7, two inputs of the comparator are swapped four times during each comparison cycle. The swapping ideally causes the comparator to make complementary decisions. In the time domain, the chopping behavior can be explained as follows: Since input

is slow-varying, two consecutive samples in opposite directions should lead to two digital words with the same magnitude but opposite signs after the conversion. However, since it is a low-frequency signal, the $1/f$ noise should have approximately the same contribution during both cycles. If the two consecutive outputs are digitally subtracted from one other, the $1/f$ noise contribution should be cancelled out. The input can be recovered by de-chopping using a simple XOR gate. The swapping causes the comparators to have two opposite offsets that are chosen independently of the input signal. If the accumulated output is averaged, the noise contribution should be reduced.

4.2.4 M/2M Ladder DAC

In 1992 Bult and Geelen [28] introduced an interesting and simple two-transistor current divider that was claimed to be “inherently linear” with a current division factor dependent only on device geometries. The authors observe that this divider is attractive because, in addition to its small size, the attenuation factor can be accurately controlled in most semiconductor processes and because the full-scale input is large, essentially extending to a level that will cause one of the two transistors to enter weak inversion of operation. Using these properties, one can design an R/2R ladder using MOS pseudo-resistors instead of simple resistors. This network is called a M/2M ladder.

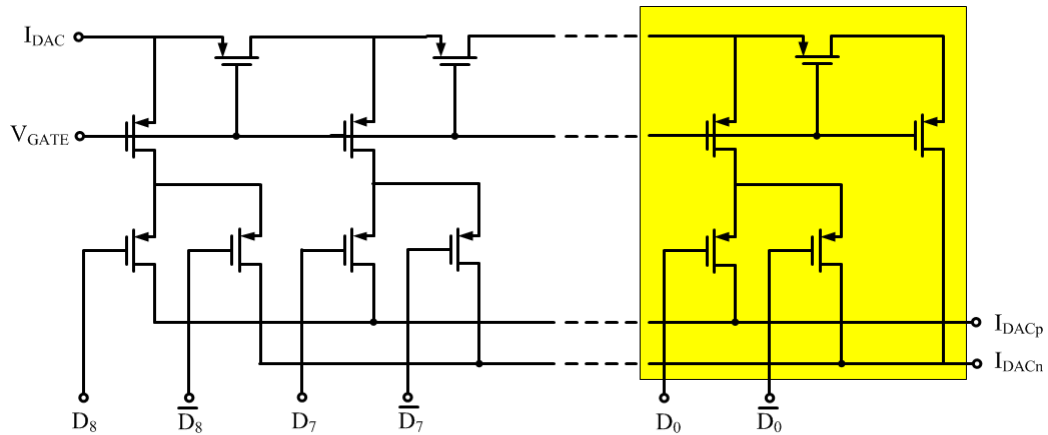


Figure 4-8: M/2M ladder DAC

The architecture proposed by [27] could satisfy the pressing area-efficient requirement in our design. It is based on the M-2M ladder, while the current-mode, MOSFET-only version is depicted in Fig.4.8. In this architecture all MOSFETs are related to the R-2R topology and are of the same size, and they operate in the linear region with R_{ON} resistance. Each transistor acts as a single pseudo-resistor having a value equal to R . Each of the transistors in the ladder has the same dimensions W and L , and each gate voltage is equal to V_{GATE} . For the bottom-most transistors, the gate voltage has two possible values: If the switch should be open, a gate voltage ensuring the blocking of the transistor is applied. When the switch is closed, it acts as a pseudo-resistor with the same value as the other transistors in the ladder, because the voltage V_{GATE} is applied to each gate. If one chooses $V_{GATE} = VSS/GND$, it is possible to directly use logic levels to drive the switch transistor gates. A low logic level (gate voltage equal to 0) makes the corresponding transistor act as a pseudo-resistor, whereas a high logic level (gate voltage equal to 1.8V) forces it to act as an open switch.

The M/2M ladder can be used to realize digital-to-analog converters. One drawback of this structure is that the ladder may be sensitive to a voltage difference between the two

current-collector nodes, so a regulator circuit could be implemented as in previously-published research. Fortunately, the ultimate goal in our current sensor design is to make the current collector node's voltage V_p and V_n equal in Figure 4.1. Thus we won't suffer this problem in our design.

Another problem is due to a second-order effect, i.e., linearity that limits the achievable resolution. We chose $W/L = 1\mu\text{m}/5\mu\text{m}$ from other analyses regarding the impact of the mismatch of the transistors in a M/2M ladder on the best-achievable resolution of digital-to-analog converters [27].

4.3 SIMULATION RESULTS

This current sensor was designed in a $0.18\mu\text{m}$ CMOS process with a 1.8V power supply voltage. An IR drop ramp from 0 to 20mV was applied to simulate static performance as well as linearity. The simulation was performed with a system clock rate of 8MHz, equivalent to a sampling rate of 100KS/s. At this sampling rate, 6000 samples were collected. The measured transfer curve and absolute accuracy is shown in Figure 4-9 and Figure 4-10 where blue and red line represent measured IR drop and actual IR drop respectively. Then absolute error is calculated through Matlab as shown in Figure 4-11. Maximum absolute error is 0.04mV which is approximately equal to 1 LSB.

Linearity performance DNL and INL are shown in Figure 4-12. The maximum DNL is 0.15 LSB and the minimum DNL is -0.12 LSB. The maximum INL is 0.05 LSB and minimum INL is -0.3 LSB.

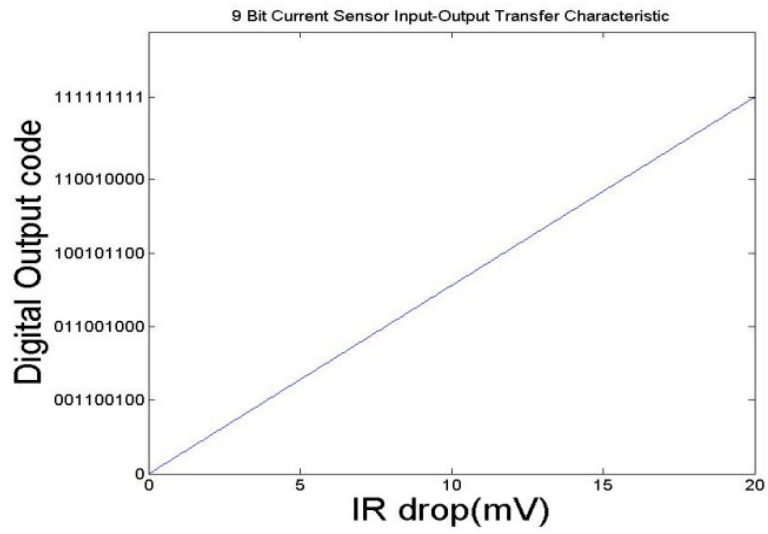


Figure 4-9: Input-output transfer characteristic of current sensor

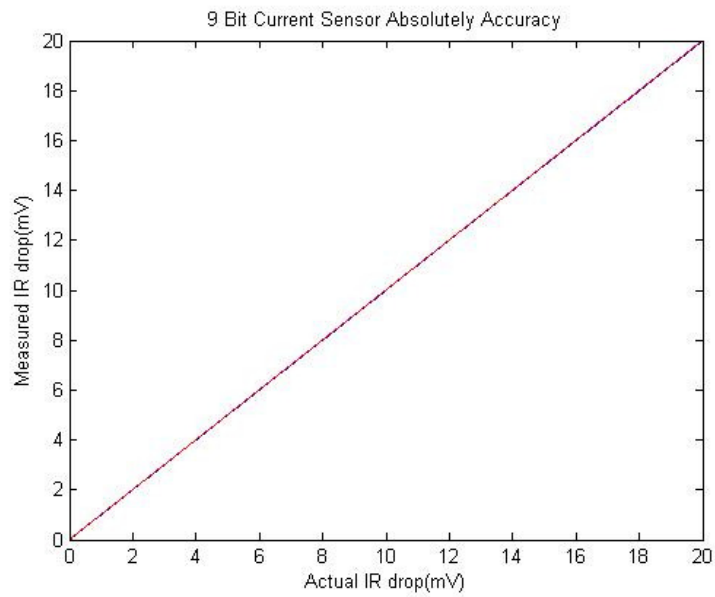


Figure 4-10: Absolute accuracy

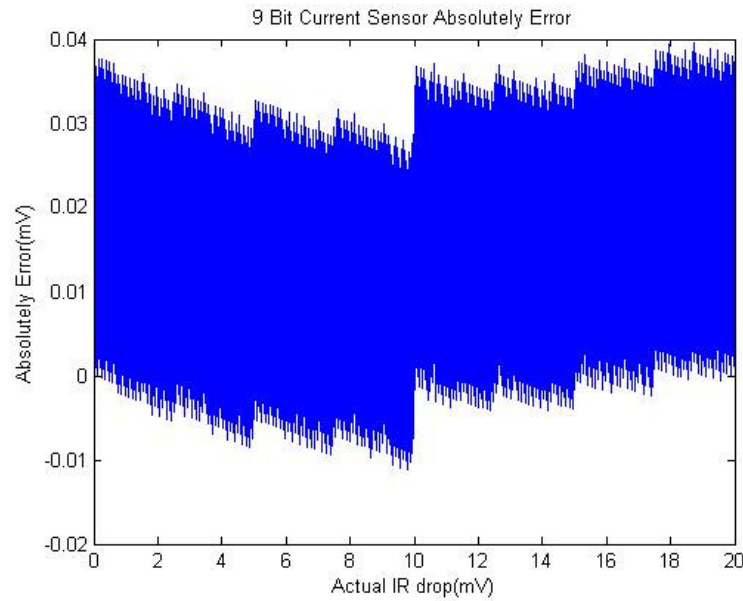


Figure 4-11: Absolute error

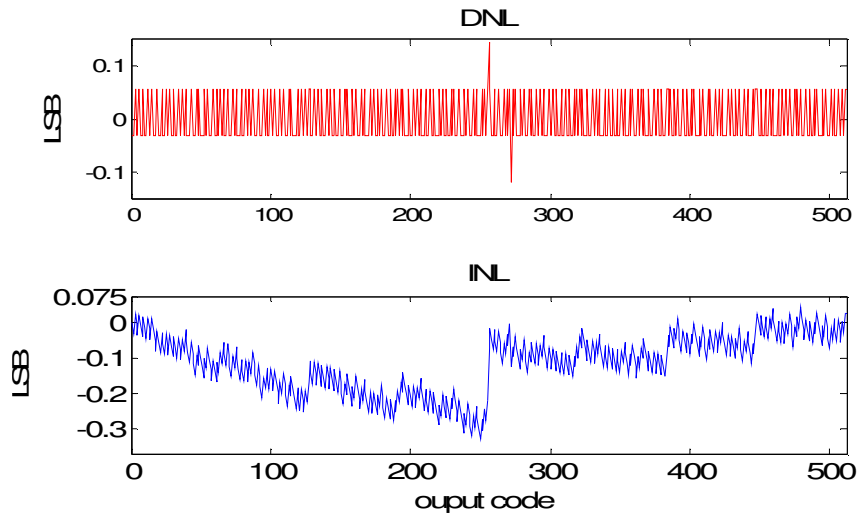


Figure 4-12: DNL&INL

To gain insight into the effects of process variations on the circuit performance, the circuit was simulated at typical (denoted as TT) and over different process corners (FF: Fast NMOS Fast PMOS, SS: Slow NMOS Slow PMOS, FS: Fast NMOS Slow PMOS, SF: Slow

NMOS Fast PMOS) with nominal V_{DD} of 1.8V. Simulation results showing transient SAR convergence process with corner variation at the critical points which IR drop is equal to 0, nominal case and full range are shown in Figure 4-13 to Figure 4-15. In graph, the absolute value of Y-axis is digital output which already convert to decimal format and X-axis is the time. With zero input, the digital output gave same correct result at four process corners. At nominal case which mean the IR drop is equal to 10mV, the digital output range vary from 011111111 to 100000010 due to process variations at a fixed supply voltage. The absolutely error is 0.117mV with the full range of 20mV. At full range, the digital output range vary from 111111000 to 111111111 then the absolutely error is 0.273mV with the full range of 20mV. From the simulation results, the error due to process variations can still achieve at least 6 efficient bits accuracy even no calibration technique is applied so far.

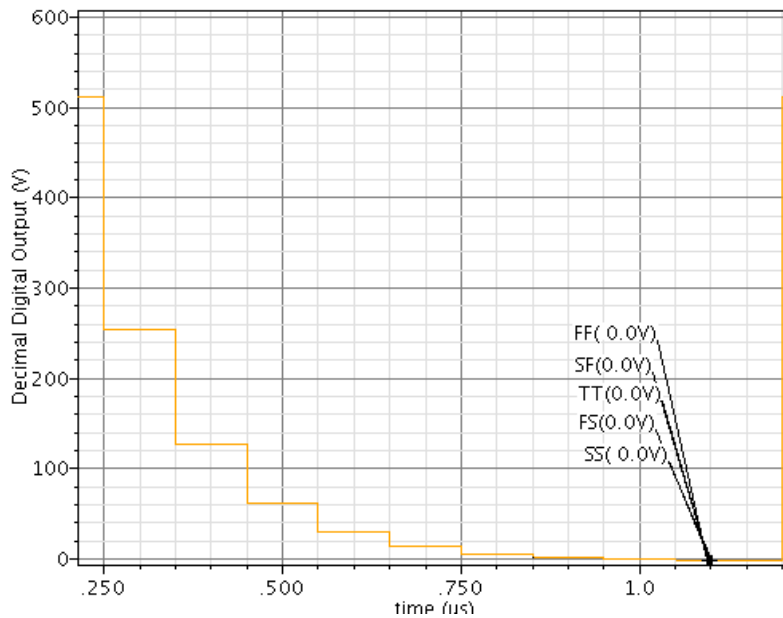


Figure 4-13: SAR Convergence process at different corners at IR = 0

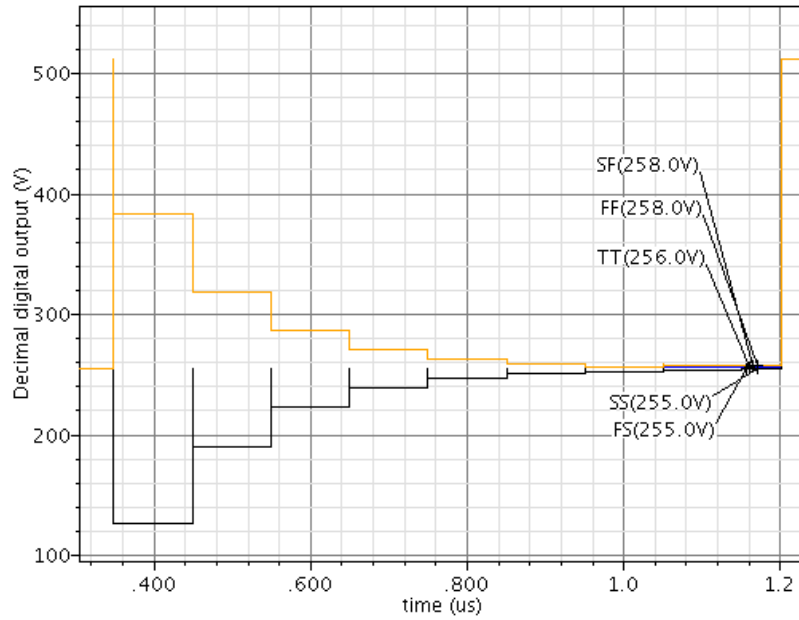


Figure 4-14: SAR Convergence process at different corners at IR = 10mV

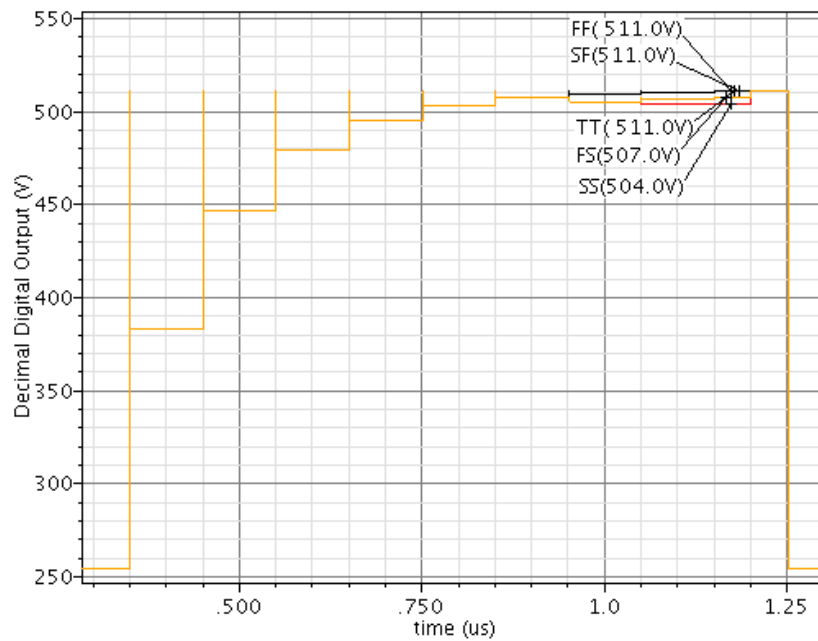


Figure 4-15: SAR Convergence process at different corners at IR = 20mV

4.4 SUMMARY

This chapter presents a new on-chip current sensing technique providing current measurements near the electromigration threshold. The proposed technique is a MOSFET-only non-intrusive measurement approach requiring only a small silicon area and low power. Simulation results show this circuit has good accuracy and linearity with respect to DNL/INL within +0.15/-0.3 LSB. The summary of performance is shown in Table 3. This structure is suitable for multi-site on-chip current measurement needed to support power/thermal management of large integrated circuits.

Table 4-3: Summary of performance

Parameters	Values
Process	1P6M 0.18 μm
Power Supply	1.8V
Power Dissipation	131.4 μW
Active area without digital parts	57.2 μm^2
Sampling Frequency	100KHz
Absolute Error	0.04mV(1 LSB)
INL	+0.05/-0.3 LSB
DNL	+0.15/-0.12 LSB

CHAPTER 5 CONCLUSIONS AND FUTURE WORK

In this work, circuit and system design for a SAR-based current sensor to be used in multi-core systems is investigated. An on-chip current measurement method suitable for electromigration management is introduced. Rather than inserting a shunt in the current flow path for creating a voltage drop, the voltage drop across existing interconnects is used to determine the value of current flow. Current is measured with a MOSFET-only sensing circuit providing 9 bits of resolution with midrange current levels at the threshold where electromigration concerns become relevant. This current sensor can be used for sensing currents in either VDD or VSS busses and is targeted for use in power power/thermal management units in integrated circuits. Simulation results show the static performance are all within acceptable range. The current sensor is robust with respect to local mismatching. The small area and low power dissipation makes the structure suitable for multiple-site on-chip current measurements.

As a continuation of the work presented in this thesis, the priority of future study is to investigate the mismatch in both the main sensor and each of the sub-circuits. To increase the accuracy of the current sensor, more calibration techniques may be considered. In addition, with respect to the proposed current sensor design, more effort can be exerted toward the goal of optimizing key performance parameters such as speed based upon the analysis done in this work.

ACKNOWLEDGEMENT

First of all, I would like to express my eternal gratitude to my parents, Mr. Zhenhua Wang and Ms. Lu Yu, for raising and supporting me throughout my life. Their unconditional love and motivation have pushed me toward my goals and to succeed further in life to come.

Most importantly, I would like to express my sincere and deep appreciation to my academic advisor, Dr. Degang Chen, for allowing me to enjoy rewarding years in my life at Iowa State. Without his continuous insightful guidance and generous support, it would be impossible for me to complete the study. I would like to appreciate Prof. Randall Geiger for providing constructive suggestions to this research. The knowledge he imparted to me will serve as the foundation to my future career path. I would also like to thank Dr. Jiming Song for his trust as well as for support throughout my graduate years, and especially for serving on my guidance committee.

I am also thankful for my good colleagues and friends in Iowa State University, including Chen Zhao, Jingbo duan, Geyuan Liu, Yulong Shi, Yen-Ting Wang, Alex Lee, Tao Zeng, Ziyue Liu, and Wei Fu. They are the best peers and friends in my life.

Last but not least, I would like to express my thanks to all my friends and family members whom names are not mentioned here. They have always given me the best memories to date along with wonderful moral support.

Sincerely,

Tianhan Wang

BIBLIOGRAPHY

- [1] http://www.intel.com/pressroom/kits/quickre_am.htm.
- [2] Jan M. Rabaey. The quest for ultra-low energy computation opportunities for architectures exploiting low-current devices. April 2000.
- [3] Ramon Canal, "Power- and Performance-Aware Architectures," Phd. Thesis. Universitat Politècnica de Catalunya.pp1-3.
- [4] Intel Corporation. The SCC Platform Overview. [Online]. http://techresearch.intel.com/spaw2/uploads/files/SCC_Platform_Overview.pdf
- [5] Tilera Corporation. Technical document.[Online]. http://www.tilera.com/sites/default/files/productbriefs/PB010_TILE64_Processor_A_v4.pdf
- [6] Jim Laudon. Performance/watt the new server focus. In Proceedings of First Workshop on Design, Architecture, and Simulation of Chip-Multiprocessors (dasCMP), November 2005.
- [7] M. Horowitz, E. Alon, D. Patil, S. Na_ziger, R. Kumar, and K. Bernstein. Scaling, power, and the future of cmos. In IEEE International Electron Devices Meeting, December 2005.
- [8] ITRS. (2010) ITRS 2010 Update. [Online]. <http://www.itrs.net/Links/2010ITRS/Home2010.htm>
- [9] A. Pages and M. Gondran, translated by E. Griffin, System Reliability Evaluation and Prediction in Engineering, New York: Springer-Verlag, 1986.

- [10] C. F. Hawkins, J. Segura, J. M. Soden and T. A. Dellin, "Test and Reliability: Partners in IC Manufacturing," IEEE DESIGN & TEST OF COMPUTERS, Oct.,1999, pp.66-73.
- [11] M. Bohr, "High Performance Logic Technology and Reliability Challenges," IEEE Int. Reliab. Phys. Symp., Key Note Presentation, 2003.
- [12] J. P. Shen, W. Maly and F. J. Ferguson, "Inductive Fault Analysis of MOS Integrated Circuits," IEEE Design and Test, December 1985, pp. 13-26.
- [13] J. R. Lloyd, "Reliability Modeling For Electromigration Failure," Quality and Reliability Engineering International, vol.10, pp.303-308, 1999.
- [14] Y.-J. Park, V. K. Andleigh, and C. V. Thompson, "Simulations of stress evolution and the current density scaling of electromigration-induced failure times in pure and alloyed interconnects", J. Appl. Phys., 85(7):3546–3555, April 1999.
- [15] M. E. Sarychev, Y. V. Zhitnikov, L. Borucki, C.-L. Liu, and T. M. Markhviladze, "General model for mechanical stress evolution during electromigration," J. Appl. Phys., 86(6):3068–75, September 1999.
- [16] Black, J.R., "Electromigration—A brief survey and some recent results," , IEEE Transactions on Electron Devices, 16(4):338 - 347 , April 1969.
- [17] Black, J., "Mass Transport of Aluminum by Momentum Exchange with Conducting Electrons", Proc. 6th Annual Reliability Physics Symposium, pp. 148-159, Nov. 1967

- [18] W. R. Hunter, "Self-consistent solutions for allowed interconnect current density part I: Implications for technology evolution. IEEE Trans. on Electron Devices", 44(2):304–309, February 1997.
- [19] ITRS. (2009) ITRS 2009 Update. [Online].
http://www.itrs.net/links/2009itrs/2009chapters_2009tables/2009_Interconnect.pdf
- [20] Maly, W. and Nigh, P., "Built-In current Testing – Feasibility Study", IEEE International Conference on Computer Aided Design (ICCAD), pp. 340-343, Nov. 1988.
- [21] Maly, W. and Nigh, P., "Built-in Current Testing of Integrated Circuits", US Patent 5025344, June 18, 1991.
- [22] Abhijit Prasad, "Power Supply Partitioning For Placement of Built-In Current Sensors for IDDQ Testing," Master of Science Thesis, Texas A&M University, 2003.
- [23] S. Siskos and A. A. Hatzopoulos, "A Simple Built-In Current Sensor for Current Monitoring in Mixed-Signal Circuits," IEEE Transactions On Instrumentation And Measurement, vol. 46, no. 6, pp. 1301-1304, December 1997.
- [24] Maidon Y, Deval Y, Fremont H, Dom JP, "Using IDD to analyze Analogue Faults and Development of a sensor," Proceeding of MM'96 Symposium, Austin, USA, pp 100–103, 1996.
- [25] Baozhen Li, Cathryn Christiansen, Jason Gill, Timothy Sullivan, Emmanuel Yashchin et al. "Threshold electromigration failure time and its statistics for Cu interconnects," Journal of Applied Physics, 2006

- [26] B. Razavi and B. A. Wooley, "Design Techniques for High-Speed, High-Resolution Comparator," IEEE Journal of Solid-State Circuits, vol. 27, no. 12, pp. 1916-1926, Dec. 1992.
- [27] Hammerschmied, C.M. and Qiuting Huang, "Design and implementation of an untrimmed MOSFET-only 10-bit A/D converter with -79-dB THD," IEEE Journal of solid-State Circuits, vol.33 Issue:8, pp. 1148 – 1157, Aug 1998.
- [28] K. Bult and G. Geelen, "An Inherently Linear and Compact MOST Only Current Division Technique", IEEE J. Solid-State Circuits, Vol. 26, pp. 1730-1735, Dec. 1992.
- [29] M. J. M. Pelgrom, A. C. J. Duinmaijer, and A. P. G. Welbers, "Matching Properties of MOS Transistors," IEEE J. Solid-State Circuits, vol. 24, pp. 1433-1439, Oct. 1989.